



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

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

(11)



1-0038942

(51)^{2020.01} C12N 15/70; C12P 13/04; C12N 9/10;
C12N 9/12; C12N 15/77; C12N 15/81

(13) B

(21) 1-2020-05284

(22) 22/02/2019

(86) PCT/KR2019/002242 22/02/2019

(87) WO2019/164351 29/08/2019

(30) 10-2018-0022185 23/02/2018 KR

(45) 26/02/2024 431

(43) 25/12/2020 393

(73) CJ CHEILJEDANG CORPORATION (KR)

330, Dongho-ro, Jung-gu, Seoul 04560, Republic of Korea

(72) KIM, Sol (KR); SEOK, Jong-cheol (KR); LEE, Kyusung (KR); JANG, Jae Woo (KR).

(74) Công ty TNHH Sáng chế ACTIP (ACTIP PATENT LIMITED)

(54) VI SINH VẬT SẢN SINH AXIT AMIN DẠNG MYCOSPORIN VÀ PHƯƠNG
PHÁP SẢN XUẤT AXIT AMIN DẠNG MYCOSPORIN BẰNG VI SINH VẬT

(57) Sáng chế đề cập đến vi sinh vật sản sinh axit amin dạng mycosporin và phương pháp sản xuất axit amin dạng mycosporin bằng vi sinh vật này. Do vi sinh vật theo sáng chế tăng khả năng sản sinh axit amin dạng mycosporin, nó có thể được sử dụng hiệu quả trong sản xuất axit amin dạng mycosporin.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến vi sinh vật sản sinh axit amin dạng mycosporin, và phương pháp sản xuất axit amin dạng mycosporin bằng vi sinh vật.

Tình trạng kỹ thuật của sáng chế

Bức xạ tia cực tím (UV) được phát ra từ mặt trời chứa UVA (có bước sóng trong khoảng từ 320 nm đến 400 nm), UVB (có bước sóng trong khoảng từ 290 nm đến 320 nm) và UVC (có bước sóng trong khoảng từ 100 nm đến 280 nm). Tia UVA được biết xâm nhập vào lớp da, chủ yếu gây nám và lão hóa da, và liên quan đến sự xuất hiện của bệnh da nhạy cảm với ánh sáng, trong khi tia UVB là tia năng lượng cao xuyên qua lớp biểu bì và lớp nền của hạ bì và có liên quan đến việc gây cháy nắng, nám da và sự xuất hiện của ung thư da.

Các nỗ lực được thực hiện để ngăn chặn các tia UV nhằm ngăn chặn các tác động phụ gây ra bởi các tia UV. Các loại chất chống nắng bao gồm chất chống nắng hóa học và chất chống nắng vật lý. Chất chống nắng hóa học ngăn chặn sự xâm nhập của các tia UV chủ yếu bằng cách hấp thụ các tia UV, trong khi chất chống nắng vật lý ngăn chặn sự xâm nhập của tia các UV thông qua phản xạ và tán xạ.

Các thành phần được biết là chứa trong các chất chống nắng hóa học có thể bao gồm các thành phần chủ yếu hấp thụ các tia UVB (ví dụ, PABA, các PABA este (amyl dimetyl PABA, octyl dimetyl PABA), các xinamat (xinoxat), salixylat (homomentyl Salixylat), long não, v.v.); và những chất chủ yếu hấp thụ các tia UVA (ví dụ, các benzophenon (oxybenzon, đioxybenzon và suliso benzen), dibenzoyl metan, anthranilat, v.v.). Mặc dù các chất chống nắng hóa học này có thể mang lại tác dụng hấp thụ và ngăn chặn các tia UV, nhưng được biết rằng một số chất chống nắng hóa học này có thể gây kích ứng da hoặc mắt, và cụ thể là, PABA, các PABA este, các benzophenon và các xinamat, v.v. có thể gây viêm da tiếp xúc. Ngoài ra, một số chất chống nắng hóa học khác được báo cáo là gây ra phản ứng nhạy cảm với ánh sáng ở da, v.v. Theo đó, ở một số quốc gia, việc sử dụng hoặc lượng sử dụng các chất chống nắng hóa học bị hạn chế.

Các thành phần được biết là chứa trong các chất chống nắng vật lý có thể bao gồm

titan đioxit, talc (magie silicat), magie oxit, kẽm oxit, cao lanh, v.v. Các chất chống nắng vật lý có ưu điểm là chúng không gây ra tác dụng phụ, như viêm da tiếp xúc và chúng không dễ dàng bị rửa trôi bởi nước. Tuy nhiên, chúng có nhược điểm là khó duy trì hàm lượng hiệu quả trong khi thực hiện các công thức mong muốn và có thể xảy ra hiện tượng để lại một lớp trắng khi thoa lên da, v.v.

Các axit amin dạng mycosporin (MAAs) là các vật liệu có trong các sinh vật tự nhiên và chúng được biết là hấp thụ hiệu quả tia UVA và UVB. Hơn 35 loại MAAs được biết là có trong tự nhiên (Sinh vật biển., 1991, 108: 157-166; Tạp chí về Cây thuốc 2015. 81: 813-820). Gần đây, MAAs được gắn các loại đường khác nhau, đã được báo cáo là tồn tại trong vi tảo và chúng có chức năng chống oxy hóa tốt (Tạp chí Quang hóa và Quang sinh học, 2015, 142: 154-168). Ngoài ra, MAAs được biết là không chỉ cung cấp khả năng ngăn chặn các tia UV mà còn cho thấy khả năng chống lại quá trình oxy hóa, thẩm thấu, căng thẳng nhiệt, v.v. (So sánh sinh hóa và sinh lý học, phần C: Độc tính và Dược lý 2007, 146: 60-78; Tạp chí Quang hóa và Quang sinh học, phần B 2007, 89: 29-35).

Tuy nhiên, lượng MAAs được sản xuất trong vi tảo là rất nhỏ ở mức vài μg , và các điều kiện để phân tách, tách chiết và tinh sạch MAAs bằng cách nuôi cấy vi tảo rất phức tạp, và do đó, sẽ khó sản xuất MAAs trong sản xuất quy mô lớn.

Tài liệu kỹ thuật tham khảo

Tài liệu non-patent

Tài liệu non-patent 1: So sánh sinh hóa và sinh lý học, phần B 1995, 112: 105-114.

Tài liệu non-patent 2: Báo cáo của Liên đoàn hiệp hội Vi sinh vật châu Âu 2007, 269: 1-10.

Tài liệu non-patent 3: Đánh giá thường niên về sinh lý học 2002, 64: 223-262.

Tài liệu non-patent 4: Sinh vật biển 1991, 108: 157-166.

Tài liệu non-patent 5: Tạp chí quang hóa và quang sinh học, phần B: Sinh học 2015, 142: 154-168

Tài liệu non-patent 6: Đánh giá sinh học 1999, 74: 311-345.

Tài liệu non-patent 7: Sinh học phân tử và tiến hóa 2006, 23: 1437-1443.

Tài liệu non-patent 8: Khoa học 2010, 329: 1653-1656.

Tài liệu non-patent 9: Hệ gen 2010, 95: 120-128.

Tài liệu non-patent 10: Địa sinh học. J. 1997. 14: 231-241.

Tài liệu non-patent 11: So sánh sinh hóa và sinh lý học, phần C: Độc tính và dược lý 2007. 146: 60-78.

Tài liệu non-patent 12: Tạp chí thực vật học Canada 2003. 81: 131-138.

Tài liệu non-patent 13: Tạp chí quang hóa và quang sinh học, phần B 2007, 89: 29-35.

Tài liệu non-patent 14: Tạp chí vi khuẩn học 2011. 193(21): 5923-5928.

Tài liệu non-patent 15: Tạp chí về cây thuốc 2015. 81: 813-820

Tài liệu non-patent 16: Vật liệu ứng dụng & giao diện của Hiệp hội hóa học Hoa Kỳ 2015. 7: 16558-16564

Tài liệu non-patent 17: Ứng dụng và vi sinh môi trường nuôi cấy 2016, 82(20): 6167-6173

Tài liệu non-patent 18: Tạp chí hóa học và sinh học 2015, 16: 320-327

Tài liệu non-patent 19: Các phương pháp trong sinh học phân tử 2013, 1073: 43-7

Tài liệu non-patent 20: Đánh giá thiên nhiên 2011, 9: 791-802

Các tác giả sáng chế đã nỗ lực rất nhiều để tăng sản sinh MAAs trong các vi sinh vật. Kết quả là, đã xác nhận rằng việc sản sinh MAAs có thể được tăng lên trong các vi sinh vật sản sinh MAAs thông qua các nghiên cứu khác nhau về việc tăng cường hoạt tính của 2-đehydro-3-đeoxyphosphoheptonat aldolaza, phosphoenolpyruvat synthetaza và các transketolaza protein trong các vi sinh vật, nhờ đó hoàn thành sáng chế.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cung cấp vi sinh vật sản sinh axit amin dạng mycosporin (MMA), trong đó hoạt tính của ít nhất một protein được chọn từ nhóm bao gồm 2-đehydro-3-đeoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza được tăng cường.

Mục đích khác của sáng chế là đề xuất phương pháp sản xuất axit amin dạng

mycosporin, bao gồm nuôi cấy vi sinh vật trong môi trường nuôi cấy; và thu hồi axit amin dạng mycosporin từ vi sinh vật được nuôi cấy hoặc môi trường nuôi cấy.

Mục đích khác nữa của sáng chế là đề xuất việc sử dụng vi sinh vật để sản xuất axit amin dạng mycosporin.

Hiệu quả đạt được của sáng chế

Vì vi sinh vật theo sáng chế tăng khả năng sản sinh axit amin dạng mycosporin (MMA), do đó có thể được sử dụng hiệu quả trong việc sản xuất các axit amin dạng mycosporin.

Mô tả chi tiết sáng chế

Sáng chế được mô tả chi tiết dưới đây. Trong khi, các phần mô tả và phương án tương ứng được bộc lộ trong sáng chế có thể được áp dụng cho các phần mô tả và phương án khác. Do đó, tất cả sự kết hợp của các yếu tố khác nhau được bộc lộ ở đây nằm trong phạm vi sáng chế. Ngoài ra, phạm vi của sáng chế không bị giới hạn bởi các phần mô tả cụ thể dưới đây. Hơn nữa, người có hiểu biết trung bình trong cùng lĩnh vực kỹ thuật có thể nhận biết hoặc xác định một số đương lượng đối với các khía cạnh nhất định của sáng chế chỉ bằng thử nghiệm thông thường. Theo đó, các đương lượng này cũng thuộc phạm vi của sáng chế.

Để đạt được các mục đích trên, một khía cạnh của sáng chế cung cấp vi sinh vật sản sinh axit amin dạng mycosporin (MMA), trong đó hoạt tính của ít nhất một protein được chọn từ nhóm bao gồm 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza được tăng cường.

Như được sử dụng trong sáng chế, thuật ngữ “2-dehydro-3-deoxyphosphoheptonat aldolaza” đề cập đến enzym xúc tác cho phản ứng thuận nghịch được biểu thị bởi sơ đồ phản ứng dưới đây, và cụ thể là, có thể đề cập đến enzym để tổng hợp 3-deoxy-D-arabino-heptulosonat-7-phosphat (DAHP), nhưng sáng chế không giới hạn ở đây.

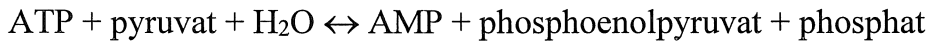
Sơ đồ phản ứng

Phosphoenolpyruvat + D-erythroza-4-phosphat + H₂O ↔ 3-deoxy-D-arabinoheptulosonat-7-phosphat + phosphat

Theo sáng chế, 2-dehydro-3-deoxyphosphoheptonat aldolaza có thể sử dụng thay thế cho 3-deoxy-D-arabino-heptulosonat-7-phosphat (DAHP) synthaza.

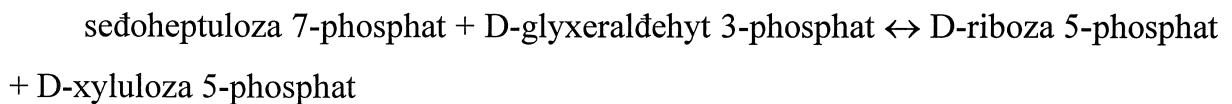
Như được sử dụng trong sáng chế, thuật ngữ “phosphoenolpyruvat syntetaza” đề cập đến enzym xúc tác cho phản ứng thuận nghịch được biểu thị bởi sơ đồ phản ứng dưới đây, và cụ thể là, có thể đề cập đến enzym để tổng hợp phosphoenolpyruvat, nhưng sáng chế không giới hạn ở đây.

Sơ đồ phản ứng

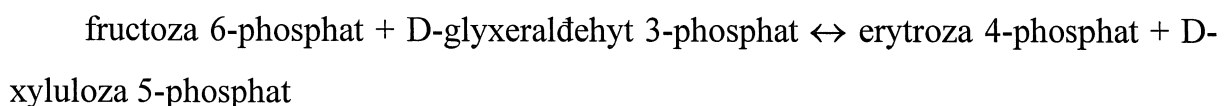


Như được sử dụng trong sáng chế, thuật ngữ “transketolaza” đề cập đến enzym xúc tác cho phản ứng thuận nghịch được biểu thị bởi sơ đồ phản ứng dưới đây.

Sơ đồ phản ứng



hoặc



Thông tin di truyền của 2-đehydro-3-đeoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza có thể thu được từ các cơ sở dữ liệu đã biết (ví dụ, Ngân hàng gen của Trung tâm thông tin công nghệ sinh học Quốc gia (NCBI), v.v.), nhưng sáng chế không giới hạn ở đây.

2-đehydro-3-đeoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza không bị giới hạn về nguồn gốc hoặc các trình tự của chúng vì có các trường hợp các protein biểu hiện các hoạt tính khác nhau trong các trình tự axit amin của chúng tùy thuộc vào các loài vi sinh vật hoặc chính vi sinh vật đó.

Cụ thể là, 2-đehydro-3-đeoxyphosphoheptonat aldolaza có thể là protein chứa trình tự axit amin được biểu thị bởi SEQ ID NO: 2, SEQ ID NO: 37, hoặc SEQ ID NO: 124; phosphoenolpyruvat syntetaza có thể là protein chứa trình tự axit amin được biểu thị bởi SEQ ID NO: 19 hoặc SEQ ID NO: 98; và transketolaza có thể là protein chứa trình tự axit amin được biểu thị bởi SEQ ID NO: 24, SEQ ID NO: 96, hoặc SEQ ID NO: 123, nhưng các trình tự axit amin của các protein này không bị giới hạn ở đây. Như được sử dụng trong sáng chế, thuật ngữ “protein chứa trình tự axit amin” có thể sử dụng thay thế

cho sự biểu hiện “protein có trình tự axit amin” hoặc “protein bao gồm trình tự axit amin”.

Ngoài ra, theo sáng chế, các enzym có thể bao gồm các protein chứa các trình tự axit amin được biểu thị bởi các SEQ ID NO được mô tả ở trên cũng như protein có sự tương đồng hoặc đồng nhất tối thiểu 80%, tối thiểu 85%, tối thiểu 90%, tối thiểu 91%, tối thiểu 92%, tối thiểu 93%, tối thiểu 94%, tối thiểu 95%, tối thiểu 96%, tối thiểu 97%, tối thiểu 98%, tối thiểu 99% với các trình tự axit amin ở trên, miễn là các protein có hoạt tính sinh học giống hệt hoặc tương ứng với từng enzym.

Ngoài ra, rõ ràng là bất kỳ protein nào có trình tự axit amin bị xóa, biến đổi, thay thế, hoặc thêm một phần của trình tự cũng có thể được bao gồm trong phạm vi của sáng chế, miễn là trình tự axit amin có sự tương đồng hoặc đồng nhất với các trình tự SEQ ID NO được mô tả ở trên và có hoạt tính sinh học về cơ bản giống hệt hoặc tương ứng với các protein enzym của các trình tự SEQ ID NO được mô tả ở trên.

Như được sử dụng trong sáng chế, thuật ngữ “tương đồng hoặc đồng nhất” đề cập đến mức độ tương ứng giữa các trình tự axit amin hoặc các trình tự nucleotit và có thể được biểu thị bởi phần trăm. Theo sáng chế, trình tự tương đồng có hoạt tính giống hoặc tương tự với trình tự axit amin hoặc trình tự nucleotit đã cho có thể được chỉ định theo thuật ngữ “% tương đồng” hoặc “% đồng nhất”. Ví dụ, tương đồng có thể được xác nhận bằng cách sử dụng phần mềm tiêu chuẩn, cụ thể là BLAST 2.0, để tính toán các tham số như là điểm số, đồng nhất, và tương tự hoặc bằng cách so sánh các trình tự bằng phép lai southern trong điều kiện nghiêm ngặt đã xác định. Các điều kiện lai phù hợp đã xác định có thể nằm trong phạm vi trong cùng lĩnh vực kỹ thuật và có thể được xác định bằng phương pháp đã biết bởi người có hiểu biết trung bình trong cùng lĩnh vực kỹ thuật (ví dụ, J. Sambrook và cộng sự, *Nhân dòng phân tử, hướng dẫn thí nghiệm*, Ấn bản lần 2, Phòng thí nghiệm Cold Spring Harbor, Cold Spring Harbor, New York, 1989; F.M. Ausubel và cộng sự, *Các quy trình hiện tại trong sinh học phân tử*, John Wiley & Sons, Inc., New York). Thuật ngữ “các điều kiện nghiêm ngặt” đề cập đến các điều kiện cho phép lai đặc hiệu giữa các polynucleotit. Ví dụ, các điều kiện như vậy được mô tả chi tiết trong tài liệu tham khảo (ví dụ, J. Sambrook và cộng sự).

2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza theo sáng chế có thể bao gồm các polynucleotit mã hóa các protein chứa

các trình tự axit amin được biểu thị bởi các SEQ ID NO được mô tả ở trên, hoặc có sự tương đồng hoặc đồng nhất tối thiểu 80%, tối thiểu 85%, tối thiểu 90%, tối thiểu 91%, tối thiểu 92%, tối thiểu 93%, tối thiểu 94%, tối thiểu 95%, tối thiểu 96%, tối thiểu 97%, tối thiểu 98%, tối thiểu 99% với các trình tự axit amin được biểu thị bởi các SEQ ID NO được mô tả ở trên, miễn là các polynucleotit có hoạt tính sinh học giống hệt hoặc tương ứng với từng enzym.

Ngoài ra, khi xét đến các codon được ưu tiên trong vi sinh vật, tại đó protein được biểu hiện, do sự thoái biến codon, các biến đổi có thể được thực hiện trong vùng mã hóa của trình tự nucleotit trong phạm vi không làm thay đổi trình tự axit amin của protein được biểu hiện từ vùng mã hóa. Do đó, bất kỳ polynucleotit có trình tự mã hóa từng protein enzym có thể được bao gồm mà không bị giới hạn.

Ngoài ra, bất kỳ trình tự nào mã hóa protein có hoạt tính của các protein enzym 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, hoặc transketolaza bằng cách lai với bất kỳ đầu dò nào có thể được điều chế từ các trình tự gen đã biết (ví dụ, các trình tự bổ sung cho tất cả hoặc một phần trình tự polynucleotit ở trên) trong các điều kiện nghiêm ngặt, có thể được bao gồm mà không bị giới hạn.

Thuật ngữ “các điều kiện nghiêm ngặt” đề cập đến các điều kiện cho phép lai đặc hiệu giữa các polynucleotit. Các điều kiện được mô tả cụ thể trong tài liệu tham khảo (ví dụ, J Sambrook và cộng sự, supra). Ví dụ, điều kiện lai giữa các gen có tính tương đồng hoặc đồng nhất cao, độ tương đồng hoặc đồng nhất tối thiểu 40%, tốt hơn là tối thiểu 90%, tốt hơn là tối thiểu 95%, tốt hơn nữa là tối thiểu 97%, và tốt nhất là tối thiểu 99%, không cho phép lai giữa các gen có độ tương đồng hoặc đồng nhất thấp hơn các độ tương đồng hoặc đồng nhất được mô tả ở trên; hoặc điều kiện rửa thông thường cho lai southern, tức là, rửa một lần, tốt hơn là hai hoặc ba lần ở nồng độ muối và nhiệt độ tương ứng với 60°C, 1×SSC, 0,1% SDS, tốt hơn nữa là 60°C, 0,1×SSC, 0,1% SDS, và tốt nhất là 68°C, 0,1×SSC, 0,1% SDS, có thể được liệt kê.

Quá trình lai đòi hỏi hai polynucleotit phải có các trình tự bổ sung, mặc dù sự không phù hợp giữa các bazơ có thể xảy ra tùy thuộc vào mức độ nghiêm ngặt của phép lai. Thuật ngữ “bổ sung” được sử dụng để mô tả mối quan hệ giữa các bazơ nucleotit có thể lai với nhau. Ví dụ, đối với ADN, adenosin được bổ sung cho thymin, và xytoxin được bổ sung cho guanin. Theo đó, sáng chế cũng có thể bao gồm các đoạn polynucleotit

được phân lập bổ sung cho toàn bộ trình tự cũng như các trình tự polynucleotit giống nhau về cơ bản.

Cụ thể là, các polynucleotit có tính tương đồng hoặc đồng nhất có thể được phát hiện bằng cách sử dụng các điều kiện lai bao gồm bước lai ở giá trị T_m là 55°C và sử dụng các điều kiện được mô tả ở trên. Ngoài ra, giá trị T_m có thể là 60°C , 63°C , hoặc 65°C , nhưng nhiệt độ không bị giới hạn ở đây và có thể được điều chỉnh phù hợp bởi người có hiểu biết trung bình trong cùng lĩnh vực tùy theo mục đích.

Sự nghiêm ngặt phù hợp cho phép lai các polynucleotit phụ thuộc vào độ dài và mức độ bổ sung của các polynucleotit, và biến số đã biết đến trong cùng lĩnh vực (xem Sambrook và cộng sự, supra, 9.50 đến 9.51 và 11.7 đến 11.8). Như được sử dụng trong sáng chế, thuật ngữ “tăng cường hoạt tính” có nghĩa là hoạt tính của protein enzym được đưa vào, hoặc hoạt tính được tăng cường so với hoạt tính nội sinh của nó hoặc hoạt tính trước khi biến đổi mà vi sinh vật sở hữu. “Đưa vào” hoạt tính có nghĩa là vi sinh vật biểu hiện tự nhiên hoặc nhân tạo hoạt tính protein cụ thể mà ban đầu không có trong vi sinh vật đó. Cụ thể là, vi sinh vật có hoạt tính tăng cường của protein enzym đề cập đến vi sinh vật trong đó hoạt tính của protein enzym được tăng cường so với vi sinh vật loại hoang dại tự nhiên hoặc vi sinh vật không bị biến đổi. Tăng cường hoạt tính có thể bao gồm, ví dụ, cả việc tăng cường hoạt tính bằng cách đưa 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và/hoặc transketolaza ngoại sinh vào vi sinh vật; hoặc tăng cường hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và/hoặc transketolaza nội sinh.

Cụ thể là, theo sáng chế, phương pháp để tăng cường hoạt tính có thể bao gồm:

- (1) phương pháp tăng số lượng bản sao của các polynucleotit mã hóa các enzym;
- (2) phương pháp biến đổi trình tự kiểm soát biểu hiện để tăng sự biểu hiện các polynucleotit;
- (3) phương pháp biến đổi các trình tự polynucleotit trên nhiễm sắc thể để tăng cường các hoạt tính của các enzym; hoặc
- (4) phương pháp biến đổi để tăng cường bằng cách kết hợp các phương pháp (1) đến (3), v.v. ở trên, nhưng sáng chế không bị giới hạn ở các phương pháp này.

Phương pháp (1) tăng số lượng bản sao của các polynucleotit có thể được thực hiện ở dạng polynucleotit được liên kết hoạt động với vector hoặc bằng cách chèn polynucleotit vào nhiễm sắc thể của tế bào chủ, nhưng phương pháp không bị giới hạn cụ thể ở đây. Ngoài ra, để thay thế, số lượng bản sao có thể được tăng lên bằng cách đưa vào polynucleotit ngoại sinh biểu hiện hoạt tính của enzym hoặc polynucleotit được biến đổi tối ưu hóa codon của polynucleotit ở trên vào tế bào chủ. Polynucleotit ngoại sinh có thể được sử dụng mà không bị giới hạn về nguồn gốc hoặc trình tự của nó miễn là polynucleotit biểu hiện hoạt tính đồng nhất hoặc tương tự với enzym. Việc đưa vào có thể được thực hiện bởi người có hiểu biết trung bình trong cùng lĩnh vực bằng cách lựa chọn phù hợp phương pháp biến nạp đã biết, và hoạt tính của enzym có thể được tăng cường sao cho polynucleotit được đưa vào được biểu hiện trong tế bào chủ để sản xuất enzym.

Tiếp theo, phương pháp (2) biến đổi trình tự kiểm soát biểu hiện để tăng sự biểu hiện các polynucleotit có thể được thực hiện bằng cách tạo ra biến đổi trong trình tự bằng cách xóa, chèn, thay thế bảo thủ hoặc không bảo thủ, hoặc kết hợp chúng để tăng cường hơn nữa hoạt tính của trình tự kiểm soát biểu hiện; hoặc bằng cách thay thế trình tự axit nucleic bằng trình tự axit nucleic có hoạt tính mạnh hơn, nhưng phương pháp không bị giới hạn cụ thể ở đây. Trình tự kiểm soát biểu hiện có thể bao gồm trình tự khởi động, trình tự điều khiển, trình tự mã hóa vị trí liên kết ribosom, các trình tự kiểm soát sự kết thúc phiên mã và dịch mã, v.v., nhưng trình tự kiểm soát biểu hiện không bị giới hạn cụ thể ở đây.

Cụ thể là, trình tự khởi động dị hợp mạnh có thể được liên kết với vùng ngược dòng của đơn vị biểu hiện polynucleotit thay vì trình tự khởi động ban đầu, và các ví dụ của trình tự khởi động mạnh có thể bao gồm trình tự khởi động CJ7, trình tự khởi động lysCP1, trình tự khởi động EF-Tu, trình tự khởi động groEL, trình tự khởi động aceA hoặc aceB, v.v. Cụ thể hơn là, đơn vị biểu hiện polynucleotit có thể được liên kết hoạt động với trình tự khởi động có nguồn gốc từ *Corynebacterium* như trình tự khởi động lysCP1 (công bố đơn quốc tế số WO 2009/096689), trình tự khởi động CJ7 (công bố đơn quốc tế số WO 2006/065095), trình tự khởi động SPL (patent Hàn Quốc số KR 10-1783170 B), hoặc trình tự khởi động o2 (patent Hàn Quốc số KR 10-1632642 B) để cải thiện tốc độ biểu hiện của polynucleotit mã hóa enzym, nhưng phương pháp không bị giới hạn ở đây.

Ngoài ra, phương pháp (3) biến đổi trình tự polynucleotit trên nhiễm sắc thể có thể được thực hiện bằng cách tạo ra biến đổi trong trình tự kiểm soát biểu hiện bằng cách xóa, chèn, thay thế bảo thủ hoặc không bảo thủ, hoặc kết hợp chúng để tăng cường hơn nữa hoạt tính của trình tự polynucleotit; hoặc bằng cách thay thế trình tự axit nucleic với trình tự polynucleotit được cải thiện có hoạt tính mạnh hơn, nhưng phương pháp không bị giới hạn cụ thể ở đây.

Cuối cùng, phương pháp (4) biến đổi để tăng cường bằng cách kết hợp các phương pháp (1) đến (3) có thể được thực hiện bằng cách áp dụng một hoặc nhiều phương pháp cùng nhau trong số các phương pháp: phương pháp tăng số lượng bản sao của các polynucleotit mã hóa các enzym; phương pháp biến đổi trình tự kiểm soát biểu hiện để tăng biểu hiện; và phương pháp biến đổi các trình tự polynucleotit trên nhiễm sắc thể, hoặc phương pháp biến đổi các polynucleotit ngoại sinh biểu hiện hoạt tính của enzym hoặc polynucleotit biến đổi được tối ưu hóa codon của chúng.

Các polynucleotit có thể được mô tả là các gen khi chúng là nhóm các polynucleotit có khả năng hoạt động. Theo sáng chế, các polynucleotit và các gen có thể được sử dụng thay thế cho nhau, và các trình tự polynucleotit và các trình tự nucleotit có thể được sử dụng thay thế cho nhau.

Như được sử dụng trong sáng chế, thuật ngữ “vector” đề cập đến cấu trúc ADN bao gồm trình tự nucleotit của polynucleotit mã hóa protein đích, trong đó protein đích được liên kết hoạt động với trình tự kiểm soát phù hợp sao cho nó có thể biểu hiện trong vật chủ phù hợp. Trình tự kiểm soát có thể bao gồm trình tự khởi động có khả năng khởi đầu phiên mã, bất kỳ trình tự điều khiển để kiểm soát phiên mã, trình tự mã hóa vị trí liên kết mARN ribosom phù hợp, và các trình tự để kiểm soát sự kết thúc phiên mã và dịch mã. Vector sau khi được biến nạp vào tế bào chủ phù hợp, có thể được sao chép hoặc hoạt động không phụ thuộc bộ gen của vật chủ, hoặc có thể được tích hợp vào bộ gen của chính vật chủ.

Vector được sử dụng trong sáng chế không bị giới hạn cụ thể miễn là vector có thể được sao chép trong tế bào chủ, và bất kỳ vector được biết đến trong kỹ thuật có thể được sử dụng. Các ví dụ về các vector thông thường có thể bao gồm plasmit, cosmit, virus, và thể thực khuẩn tự nhiên hoặc tái tổ hợp. Ví dụ, pWE15, M13, MBL3, MBL4, IXII, ASHII, APII, t10, t11, Charon4A, và Charon21A, v.v. có thể được sử dụng làm vector

phage hoặc vector cosmit; và pBR, pUC, pBluescriptII, pGEM, pTZ, pCL, và pET, v.v. có thể được sử dụng làm vector plasmit. Cụ thể là, các vector như pDZ, pACYC177, pACYC184, pCL, pECCG117, pUC19, pBR322, pMW118, pCC1BAC, pSKH, pRS-413, pRS-414, và pRS-415, v.v. có thể được sử dụng, nhưng các vector không bị giới hạn ở đây.

Các vector được sử dụng trong sáng chế không bị giới hạn cụ thể, nhưng bất kỳ vector biểu hiện đã biết nào có thể được sử dụng. Ngoài ra, polynucleotit mã hóa protein đích trong nhiễm sắc thể có thể được chèn vào nhiễm sắc thể với vector để chèn nhiễm sắc thể trong tế bào. Chèn polynucleotit vào nhiễm sắc thể có thể được thực hiện bằng cách sử dụng bất kỳ phương pháp được biết đến trong kỹ thuật (ví dụ, bằng cách tái tổ hợp tương đồng), nhưng phương pháp chèn không bị giới hạn ở đây. Chỉ thị chọn lọc để xác nhận việc chèn vào nhiễm sắc thể có thể có thể được thêm vào. Chỉ thị chọn lọc được sử dụng để lựa chọn các tế bào được biến nạp vào vector (tức là, để xác nhận có hay không phân tử axit nucleic đích được chèn) và các dấu hiệu có khả năng cung cấp các kiểu hình có thể lựa chọn (ví dụ, kháng thuốc, dinh dưỡng thụ động, kháng các tác nhân gây độc tế bào, và biểu hiện của các protein bề mặt) có thể được sử dụng. Trong những trường hợp các tác nhân chọn lọc được xử lý, chỉ các tế bào có khả năng biểu hiện các chỉ thị chọn lọc mới có thể tồn tại hoặc biểu hiện các đặc điểm kiểu hình khác, và do đó được các tế bào biến nạp có thể dễ dàng được chọn lọc.

Như được sử dụng trong sáng chế, thuật ngữ “biến nạp” đề cập đến sự đưa vào vector chứa polynucleotit mã hóa protein đích vào tế bào chủ để protein được mã hóa bởi polynucleotit có thể biểu hiện trong tế bào chủ. Polynucleotit được biến nạp có thể không bị giới hạn cụ thể miễn là nó có thể biểu hiện trong tế bào chủ, không phụ thuộc vào polynucleotit được biến nạp được chèn vào nhiễm sắc thể tế bào chủ và nằm tại đó hoặc nằm ngoài nhiễm sắc thể. Ngoài ra, polynucleotit bao gồm ADN và ARN mã hóa protein đích. Miễn là polynucleotit có thể được đưa vào tế bào chủ và biểu hiện trong đó, không quan trọng polynucleotit được đưa vào ở dạng nào. Ví dụ, polynucleotit có thể là được đưa vào tế bào chủ dưới dạng catxet biểu hiện, là cấu trúc gen bao gồm tất cả các yếu tố cần thiết để tự biểu hiện. Catxet biểu hiện có thể bao gồm trình tự khởi động, thông thường được liên kết hoạt động với polynucleotit, tín hiệu kết thúc phiên mã, vị trí liên kết ribosom, và tín hiệu kết thúc dịch mã. Catxet biểu hiện có thể là vector biểu hiện có khả năng tự sao chép. Ngoài ra, polynucleotit có thể là được đưa vào tế bào

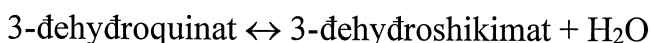
chủ như polynucleotit và để liên kết với trình tự cần thiết được biểu hiện trong tế bào chủ, nhưng sáng chế không giới hạn ở đây. Các phương pháp biến nạp bao gồm bất kỳ phương pháp để đưa axit nucleic vào tế bào, và có thể được thực hiện bằng cách chọn kỹ thuật tiêu chuẩn phù hợp được biết đến trong kỹ thuật tùy thuộc vào tế bào chủ. Ví dụ, các phương pháp biến nạp có thể bao gồm kỹ thuật xung điện, kết tủa canxi phosphat (CaPO_4), kết tủa canxi clorua (CaCl_2), vi tiêm, phương pháp polyetylen glycol (PEG), phương pháp EAE-dextran, phương pháp liposom cation, phương pháp liti axetat-DMSO, v.v., nhưng các phương pháp không bị giới hạn ở đây.

Ngoài ra, như được sử dụng trong sáng chế, thuật ngữ “liên kết hoạt động” có nghĩa là liên kết chức năng giữa trình tự khởi động để khởi đầu và làm trung gian phiên mã của polynucleotit mã hóa protein đích theo sáng chế và trình tự polynucleotit. Liên kết hoạt động có thể được tạo ra bằng cách sử dụng kỹ thuật tái tổ hợp di truyền được biết đến trong kỹ thuật, và cắt và nối ADN tại vị trí đặc hiệu có thể được tạo ra bằng cách sử dụng enzym giới hạn, ligaza, v.v., được biết đến trong kỹ thuật, nhưng sáng chế không bị giới hạn ở đây.

Ở vi sinh vật theo sáng chế, hoạt tính của 3-đehydroquinat dehydrataza có thể còn bị bất hoạt.

Như được sử dụng trong sáng chế, thuật ngữ “3-đehydroquinat dehydrataza” đề cập đến enzym xúc tác cho phản ứng thuận nghịch trong sơ đồ phản ứng bên dưới, và cụ thể là, nó có thể chuyển đổi 3-đehydroquinat thành 3-đehydroshikimat, nhưng sáng chế không giới hạn ở đây.

Sơ đồ phản ứng



3-đehydroquinat dehydrataza không bị giới hạn về nguồn gốc hoặc trình tự của nó vì có những trường hợp các protein thể hiện hoạt tính của 3-đehydroquinat dehydrataza khác nhau về các trình tự axit amin của chúng tùy thuộc vào các loài vi sinh vật hoặc vi sinh vật. Cụ thể là, 3-đehydroquinat dehydrataza có thể là protein chứa trình tự axit amin được biểu thị bởi SEQ ID NO: 90, nhưng sáng chế không giới hạn ở đây. Ngoài ra, 3-đehydroquinat dehydrataza có thể bao gồm trình tự axit amin được biểu thị bởi SEQ ID NO: 90 hoặc trình tự axit amin có tính tương đồng hoặc đồng nhất ít nhất 80%, 90%, 95%, 96%, 97%, 98%, hoặc 99% với SEQ ID NO: 90. Ngoài ra, rõ ràng là bất kỳ trình

tự axit amin nào bị xóa, biến đổi, thay thế, hoặc thêm một phần của trình tự cũng có thể được bao gồm trong phạm vi của sáng chế, miễn là trình tự axit amin có tính tương đồng hoặc đồng nhất được mô tả ở trên và có hoạt tính sinh học tương đồng hoặc tương ứng với protein ở trên.

Như được sử dụng trong sáng chế, thuật ngữ “bất hoạt” đề cập đến trường hợp trong đó hoạt tính của protein enzym bị suy yếu so với hoạt tính nội sinh của protein enzym hoặc hoạt tính của nó trước khi biến đổi mà vi sinh vật sở hữu ban đầu; trường hợp trong đó protein hoàn toàn không được biểu hiện; hoặc trường hợp trong đó protein được biểu hiện nhưng không có hoạt tính. Bất hoạt là khái niệm bao gồm: trường hợp trong đó hoạt tính của bản thân enzym bị suy yếu so với hoạt tính nội sinh của nó được sở hữu bởi vi sinh vật do sự biến đổi của polynucleotit mã hóa enzym, v.v. hoặc hoạt tính bị loại bỏ; trường hợp mức độ hoạt động nội bào tổng thể của enzym thấp hơn so với mức độ hoạt động nội bào của vi sinh vật hoang dại hoặc hoạt tính bị loại bỏ, do sự ức chế biểu hiện hoặc dịch mã của gen mã hóa enzym, v.v.; trường hợp trong đó toàn bộ hoặc một phần gen mã hóa enzym bị xóa; và sự kết hợp chúng, nhưng bất hoạt không bị giới hạn ở đây. Tức là, vi sinh vật trong đó hoạt tính của enzym bị bất hoạt đề cập đến vi sinh vật trong đó hoạt tính của protein enzym yếu hơn so với hoạt tính của vi sinh vật loại hoang dại tự nhiên hoặc vi sinh vật không bị biến đổi hoặc trong đó hoạt tính bị loại bỏ.

Sự bất hoạt hoạt tính của enzym có thể đạt được bằng cách áp dụng các phương pháp khác nhau được biết đến trong kỹ thuật. Các ví dụ của các phương pháp trên có thể bao gồm: 1) phương pháp xóa toàn bộ hoặc một phần gen mã hóa enzym trên nhiễm sắc thể; 2) phương pháp biến đổi trình tự kiểm soát biểu hiện để giảm sự biểu hiện gen mã hóa protein trên nhiễm sắc thể; 3) phương pháp biến đổi trình tự của gen mã hóa protein trên nhiễm sắc thể để hoạt tính của protein bị loại bỏ hoặc suy yếu; 4) phương pháp đưa vào oligonucleotit đối mã (ví dụ, ARN đối mã), liên kết bổ sung với bản phiên mã của gen mã hóa protein trên nhiễm sắc thể; 5) phương pháp làm cho quá trình gắn ribosom không thể xảy ra bởi cấu trúc thứ cấp được tạo thành bằng cách bổ sung trình tự, bổ sung cho trình tự Shine-Dalgarno (SD), ở đầu trước của trình tự SD của gen mã hóa protein trên nhiễm sắc thể; 6) phương pháp kỹ thuật phiên mã ngược (reverse transcription engineering: RTE), trong đó trình tự khởi động được phiên mã ngược được bổ sung vào đầu 3' của khung đọc mở (khung đọc mở: ORF) của trình tự polynucleotit mã hóa protein,

v.v.; và bất hoạt có thể đạt được bằng cách kết hợp các phương pháp nêu trên, nhưng các phương pháp không bị giới hạn cụ thể.

Phương pháp xóa toàn bộ hoặc một phần gen mã hóa enzym trên nhiễm sắc thể có thể được thực hiện bằng cách thay thế polynucleotit mã hóa protein đích nội sinh trong nhiễm sắc thể với polynucleotit hoặc gen chỉ thị có trình tự nucleotit bị xóa một phần bằng cách sử dụng vector để chèn nhiễm sắc thể. Ví dụ về phương pháp để xóa toàn bộ hoặc một phần polynucleotit, phương pháp xóa polynucleotit bằng cách tái tổ hợp tương đồng có thể được sử dụng, nhưng phương pháp không bị giới hạn ở đây.

Phương pháp biến đổi trình tự kiểm soát biểu hiện có thể được thực hiện bằng cách tạo ra biến đổi của trình tự axit nucleic trong trình tự kiểm soát biểu hiện thông qua xóa, chèn, thay thế bảo thủ hoặc không bảo thủ, hoặc kết hợp của chúng để làm suy yếu hơn nữa hoạt tính của trình tự kiểm soát biểu hiện; hoặc bằng cách thay thế trình tự axit nucleic có hoạt tính yếu hơn. Trình tự kiểm soát biểu hiện có thể bao gồm trình tự khởi động, trình tự điều khiển, trình tự mã hóa vị trí liên kết ribosom, và trình tự để điều hòa phiên mã và dịch mã, nhưng sáng chế không giới hạn ở đây.

Phương pháp biến đổi trình tự gen trên nhiễm sắc thể có thể được thực hiện bằng cách tạo ra biến đổi trong trình tự gen bằng cách xóa, chèn, thay thế bảo thủ hoặc không bảo thủ, hoặc kết hợp của chúng để là suy yếu hơn nữa hoạt tính của enzym; hoặc bằng cách thay thế trình tự gen với trình tự gen được biến đổi để có hoạt tính yếu hơn hoặc trình tự gen được biến đổi để không có hoạt tính, nhưng phương pháp không bị giới hạn ở đây.

Như được sử dụng trong sáng chế thuật ngữ “phần”, có thể khác nhau tùy thuộc vào các loại polynucleotit, cụ thể có thể là 1 nucleotit đến 300 nucleotit, tốt hơn là 1 nucleotit đến 100 nucleotit, và tốt hơn nữa là 1 nucleotit đến 50 nucleotit, nhưng không bị giới hạn cụ thể ở đây.

Trong vi sinh vật theo sáng chế, hoạt tính của 3-dehydroquinate synthase protein có thể được tăng cường hơn nữa so với vi sinh vật không bị biến đổi.

3-dehydroquinate synthase đề cập đến enzym xúc tác cho phản ứng thuận nghịch được biểu thị bởi sơ đồ phản ứng dưới đây, và cụ thể có thể tổng hợp 3-dehydroquinate (3-DHQ), nhưng sáng chế không giới hạn ở đây.

Sơ đồ phản ứng



Như được sử dụng trong sáng chế, thuật ngữ “axit amin dạng mycosporin (MAA)” đề cập đến các hợp chất vòng hấp thụ các tia cực tím (UV). Theo sáng chế, axit amin dạng mycosporin không bị giới hạn miễn là nó có thể hấp thụ các tia UV, và cụ thể là, nó có thể là hợp chất có vòng trung tâm là xyclohexanon hoặc xyclohexenimin; hoặc có thể là hợp chất trong đó các chất khác nhau (ví dụ, các axit amin, v.v.) được liên kết với vòng trung tâm, và cụ thể hơn là, nó có thể là mycosporin-2-glyxin, palythanol, axit palythenic, đeoxygadusol, mycosporin-metylamin-threonin, mycosporin-glyxin-valin, palythin, asterina-330, shinorin, porphyra-334, euhalothece-362, mycosporin-glyxin, mycosporin-ornithin, mycosporin-lyxin, mycosporin-glutamic axit-glyxin, mycosporin-metylamin-serin, mycosporin-aurin, palythen, palythin-serin, palythin-serin-sulfat, palythanol, usujiren, hoặc kết hợp của chúng.

Theo sáng chế, thuật ngữ axit amin dạng mycosporin có thể được sử dụng thay thế cho MAA và MAAs.

Như được sử dụng trong sáng chế, thuật ngữ “vi sinh vật sản sinh axit amin dạng mycosporin (MAA)” có thể đề cập đến vi sinh vật chứa gen của enzym tham gia vào quá trình sinh tổng hợp MAA hoặc cụm các gen này, hoặc vi sinh vật trong đó cụm gen được đưa vào hoặc tăng cường. Ngoài ra, như được sử dụng trong sáng chế, thuật ngữ “cụm gen sinh tổng hợp axit amin dạng mycosporin (MAA)” có thể đề cập đến nhóm các gen mã hóa các enzym tham gia vào quá trình tổng hợp MAA, và cụ thể là, có thể bao gồm gen sinh tổng hợp MAA, gen của enzym có hoạt tính gắn gốc axit amin bổ sung vào MAA, hoặc cụm gen ở trên. Gen sinh tổng hợp MAA bao gồm cả các gen ngoại sinh và/hoặc các gen nội sinh của vi sinh vật, miễn là vi sinh vật bao gồm gen có thể sản sinh MAA. Các gen ngoại sinh có thể tương đồng và/hoặc không đồng nhất.

Các loài vi sinh vật mà gen sinh tổng hợp MAA được tạo ra không bị giới hạn, miễn là các vi sinh vật có thể cùng sản sinh các enzym tham gia vào quá trình tổng hợp MAA và do đó có thể sản sinh MAA. Cụ thể là, các loài vi sinh vật mà từ đó gen sinh tổng hợp MAA được tạo ra có thể có nguồn gốc là vi khuẩn lam (ví dụ, *Anabaena variabilis*, *Nostoc punctiforme*, *Nodularia spumigena*, *Cyanothece* sp. PCC 7424, *Lyngbya* sp. PCC 8106, *Microcystis aeruginosa*, *Microcoleus chthonoplastes*,

Cyanothece sp. ATCC 51142, *Crocospaera watsonii*, *Cyanothece* sp. CCY 0110, *Cylindrospermum stagnale* sp. PCC 7417, *Aphanothece halophytica*, hoặc *Trichodesmium erythraeum*); nấm (ví dụ, *Magnaporthe oryzae*, *Pyrenophora tritici-repentis*, *Aspergillus clavatus*, *Nectria haematococca*, *Aspergillus nidulans*, *Gibberella zeae*, *Verticillium albo-atrum*, *Botryotinia fuckeliana*, hoặc *Phaeosphaeria nodorum*); hoặc *Nematostella vectensis*; *Heterocapsa triquetra*, *Oxyrrhis marina*, *Karlodinium micrum*, *Actinosynnema mirum*, v.v., nhưng không bị giới hạn ở đây.

Theo phương án này, vi sinh vật theo sáng chế sản sinh MAA bao gồm gen sinh tổng hợp MAA hoặc cụm gen của chúng. Cụ thể là, trong vi sinh vật, cụm gen sinh tổng hợp MAA có thể được đưa vào hoặc các hoạt tính của các protein được mã hóa bởi các gen có thể được tăng cường so với các hoạt tính nội sinh hoặc các hoạt tính trước khi biến đổi, nhưng vi sinh vật không bị giới hạn ở đây.

Ngoài ra, mặc dù gen sinh tổng hợp MAA không bị giới hạn trong các tên gọi của các enzym hoặc các vi sinh vật mà các gen tạo ra, miễn là các vi sinh vật có thể sản sinh MAA, gen sinh tổng hợp MAA có thể bao gồm gen mã hóa protein enzym có hoạt tính tương đồng và/hoặc tương tự với một hoặc nhiều, cụ thể là, ít nhất một, ít nhất hai, ít nhất ba, hoặc tất cả các protein enzym được chọn từ nhóm bao gồm 2-đemetyl 4-đeoxygađusol synthaza, O-metyltransferaza, và C-N ligaza.

Ví dụ, 2-đemetyl 4-đeoxygađusol synthaza là enzym chuyển đổi sedoheptuloza-7-phosphat thành 2-đemetyl-4-đeoxygađusol. O-metyltransferaza là enzym chuyển đổi 2-đemetyl-4-đeoxygađusol thành 4-đeoxygađusol, và glyxyl hóa 4-đeoxygađusol được xúc tác bởi C-N ligaza.

Ngoài ra, vi sinh vật sản sinh MAAs có thể bao gồm gen của enzym có hoạt tính gắn gốc axit amin bổ sung vào MAA, hoặc cụm gen. Mặc dù tên của gen hoặc cụm gen ở trên không bị giới hạn bởi các enzym hoặc các vi sinh vật mà các gen được tạo ra, miễn là các vi sinh vật sản sinh MAAs có thể sản sinh MAAs với hai hoặc nhiều gốc axit amin được gắn vào. Các vi sinh vật sản sinh MAAs có thể bao gồm gen mã hóa enzym protein có hoạt tính tương đồng và/hoặc tương tự với một hoặc nhiều, cụ thể là, ít nhất một, ít nhất hai, ít nhất ba, hoặc tất cả enzym các protein, được chọn từ nhóm bao gồm non-ribosomal peptit synthetaza (NRPS), enzym dạng non-ribosomal peptit synthetaza (enzym dạng NRPS), và D-alanin D-alanin ligaza (D-Ala D-Ala ligaza;

DDL).

Một số MAAs chứa gốc axit amin thứ hai trong mycosporin-glyxin. Một hoặc nhiều enzym được chọn từ nhóm bao gồm non-ribosomal peptid synthetaza, enzym dạng non-ribosomal peptid synthetaza, và D-Ala D-Ala ligaza, có thể gắn gốc axit amin thứ hai vào mycosporin-glyxin.

Theo phương án này, vi sinh vật sản sinh MAAs có thể bao gồm, mà không bị giới hạn về các tên enzym hoặc các loài vi sinh vật tạo ra các gen sinh tổng hợp MAA, các enzym miễn là chúng có hoạt tính có khả năng gắn axit amin thứ hai vào mycosporin-glyxin, như trong non-ribosomal peptid synthetaza, enzym dạng non-ribosomal peptid synthetaza, và D-Ala D-Ala ligaza.

Ví dụ, enzym dạng non-ribosomal peptid synthetaza (Ava_3855) trong *Anabaena variabilis* hoặc D-Ala D-Ala ligaza (NpF5597) trong *Nostoc punctiforme* có thể gắn gốc serin vào mycosporin-glyxin để tạo thành shinorin. Trong ví dụ khác, mycosporin-2-glyxin có thể được tạo thành bằng cách gắn gốc glyxin thứ hai bằng D-Ala D-Ala ligaza homolog (Ap_3855) trong *Aphanothece halophytica*. Tương tự, trong *Actinosynnema mirum*, serin hoặc alanin có thể được gắn bởi D-Ala D-Ala ligaza và do đó có thể tạo thành shinorin hoặc mycosporin-glyxin-alanin. Vi sinh vật theo phương án của sáng chế có thể lựa chọn và bao gồm các enzym phù hợp để sản xuất MAAs mong muốn trong các enzym được mô tả ở trên hoặc các enzym có các hoạt tính tương đồng và/hoặc tương tự với các enzym được mô tả ở trên.

2-demetyl 4-deoxygadusol synthaza, O-metyltransferaza, C-N ligaza, non-ribosomal peptid synthetaza, enzym dạng non-ribosomal peptid synthetaza, và/hoặc D-Ala D-Ala ligaza có thể được sử dụng trong sáng chế mà không bị giới hạn ở các loài vi sinh vật mà các enzym được tạo ra, miễn là các enzym được biết là có thể thực hiện các chức năng và vai trò giống và/hoặc tương tự với các enzym được mô tả ở trên, và phạm vi các giá trị tương đồng hoặc đồng nhất của chúng cũng không bị giới hạn. Ví dụ, MylA, MylB, MylD, MylE, và MylC của *C. stagnale* PCC 7417 có tính tương đồng với 2-demetyl 4-deoxygadusol synthaza, O-metyltransferaza, C-N ligaza, và D-Ala D-Ala ligaza có nguồn gốc từ *Anabaena variabilis* và *Nostoc punctiforme*, và mức độ giống nhau giữa chúng nằm trong khoảng từ 61% đến 88% (Ứng dụng và vi sinh môi trường nuôi cấy 2016, 82(20): 6167-6173; Tạp chí vi khuẩn học 2011, 193(21), 5923-5928).

Tức là, các enzym có thể được sử dụng trong sáng chế không bị giới hạn cụ thể ở các loài mà các enzym được tạo ra, trình tự tương đồng, hoặc trình tự đồng nhất, miễn là chúng được biết là thể hiện các chức năng và hiệu quả giống và/hoặc tương tự. Ngoài ra, các tài liệu non-patent được mô tả trong các tài liệu kỹ thuật đã có được kết hợp ở đây bằng cách tham khảo toàn bộ.

Ngoài ra, gen sinh tổng hợp MAA có thể là polynucleotit mã hóa protein chứa trình tự axit amin được biểu thị bởi SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, hoặc SEQ ID NO: 122, nhưng gen sinh tổng hợp MAA không bị giới hạn ở đây.

Ngoài ra, gen sinh tổng hợp MAA có thể bao gồm trình tự nucleotit mã hóa protein chứa trình tự axit amin có 50%, 60%, hoặc tối thiểu 70%, tốt hơn là tối thiểu 80%, tốt hơn là tối thiểu 90%, tốt hơn nữa là tối thiểu 95%, và tốt nhất là tối thiểu 99% tương đồng hoặc đồng nhất với trình tự axit amin được biểu thị bởi SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, hoặc SEQ ID NO: 122, và có thể bao gồm nhưng không bị giới hạn ở trình tự nucleotit mã hóa protein nằm ngoài phạm vi tương đồng hoặc đồng nhất ở trên, miễn là vi sinh vật có thể sản sinh MMA. Cụ thể là, gen sinh tổng hợp MAA có thể bao gồm trình tự nucleotit được biểu thị bởi SEQ ID NO: 102, SEQ ID NO: 103, SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 107, SEQ ID NO: 108, hoặc SEQ ID NO: 109, nhưng sáng chế không giới hạn ở đây.

Ngoài ra, rõ ràng là bất kỳ trình tự axit amin bị xóa, biến đổi, thay thế, hoặc thêm một phần của trình tự cũng có thể nằm trong phạm vi của sáng chế, miễn là trình tự axit amin có tính tương đồng hoặc đồng nhất với các trình tự ở trên và về cơ bản có hoạt tính sinh học tương đồng hoặc tương ứng với các protein được biểu thị bởi các SEQ ID NO được mô tả ở trên.

Ngoài ra, khi xem xét các codon được ưu tiên trong vi sinh vật, nơi protein được biểu hiện, do sự thoái biến codon, các biến đổi khác nhau có thể được thực hiện trong vùng mã hóa của trình tự nucleotit trong phạm vi không làm thay đổi trình tự axit amin của protein được biểu hiện từ vùng mã hóa. Do đó, đề cập đến gen sinh tổng hợp MAA, bất kỳ trình tự nucleotit có thể được bao gồm trong sáng chế mà không bị giới hạn miễn là trình tự nucleotit mã hóa protein tham gia vào quá trình tổng hợp MAA.

Ngoài ra, bất kỳ trình tự nào mã hóa protein tham gia vào quá trình tổng hợp MAA, bằng cách lai với bất kỳ đầu dò nào có thể được điều chế từ các trình tự gen đã biết (ví dụ, các trình tự bổ sung cho tất cả hoặc một phần trình tự polynucleotit) trong các điều kiện nghiêm ngặt, có thể được bao gồm trong sáng chế mà không bị giới hạn.

Theo phương án này, vi sinh vật sản sinh MAA có thể bao gồm các gen sinh tổng hợp MAA có nguồn gốc khác nhau.

Theo sáng chế, tăng cường hoạt tính protein và/hoặc sự đưa vào các gen có thể được thực hiện đồng thời, tuần tự, và ngược lại theo bất kể trình tự nào.

Vi sinh vật sản sinh MAA có thể sản sinh MAA bằng cách chứa cụm gen sinh tổng hợp MAA, và ngoài ra, có thể là vi sinh vật trong đó khả năng sản sinh MAA được tăng lên bằng cách tăng cường hoạt tính của một hoặc nhiều protein được chọn từ nhóm bao gồm 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza. Ngoài ra, vi sinh vật theo sáng chế không bị giới hạn miễn là nó là vi sinh vật trong đó khả năng sản sinh MAA được tăng lên bằng cách tăng cường hoạt tính của một hoặc nhiều protein được chọn từ nhóm bao gồm 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza. Cụ thể là, vi sinh vật theo sáng chế có thể là vi sinh vật thuộc chi *Corynebacterium*, vi sinh vật thuộc chi *Escherichia*, hoặc nấm men.

Vi sinh vật thuộc chi *Corynebacterium* có thể, cụ thể là, *Corynebacterium glutamicum*, *Corynebacterium ammoniagenes*, *Brevibacterium lactofermentum*, *Brevibacterium flavum*, *Corynebacterium thermoaminogenes*, *Corynebacterium efficiens*, v.v., và cụ thể hơn nữa có thể là *Corynebacterium glutamicum*, nhưng vi sinh vật không bị giới hạn ở đây.

Vi sinh vật thuộc chi *Escherichia* có thể, cụ thể là, *Escherichia albertii*, *Escherichia coli*, *Escherichia fergusonii*, *Escherichia hermannii*, *Escherichia vulneris*, v.v., và tốt hơn nữa là có thể là *Escherichia coli*, nhưng vi sinh vật không bị giới hạn ở đây.

Nấm men có thể, cụ thể là, *Saccharomycotina* và *Taphrinomycotina* thuộc ngành *Ascomycota*, *Agaricomycotina* thuộc ngành *Basidiomycota*, vi sinh vật thuộc ngành *Pucciniomycotina*, v.v., tốt hơn nữa là vi sinh vật thuộc chi *Saccharomyces*, vi sinh vật thuộc chi *SchizoSaccharomyces*, vi sinh vật thuộc chi *Phaffia*, vi sinh vật thuộc chi

Kluyveromyces, vi sinh vật thuộc chi *Pichia*, và vi sinh vật thuộc chi *Candida*, và tốt nhất là *Saccharomyces cerevisiae*, nhưng vi sinh vật không bị giới hạn ở đây.

Khía cách khác của sáng chế đề xuất phương pháp sản xuất axit amin dạng mycosporin, bao gồm nuôi cấy vi sinh vật theo sáng chế trong môi trường nuôi cấy; và thu hồi axit amin dạng mycosporin (MAA) từ vi sinh vật được nuôi cấy hoặc môi trường nuôi cấy.

“Vi sinh vật” và “axit amin dạng mycosporin (MAA)” như được mô tả ở trên.

Như được sử dụng trong sáng chế, thuật ngữ “nuôi cấy” có nghĩa là vi sinh vật được sinh trưởng trong các điều kiện môi trường được kiểm soát phù hợp. Quá trình nuôi cấy theo sáng chế có thể được thực hiện trong môi trường nuôi cấy phù hợp và nuôi cấy trong các điều kiện được biết đến trong kỹ thuật. Quá trình nuôi cấy có thể dễ dàng được điều chỉnh để được sử dụng bởi người có hiểu biết trung bình trong cùng lĩnh vực theo vi sinh vật được chọn. Bước nuôi cấy vi sinh vật có thể được thực hiện trong nuôi cấy theo mẻ, nuôi cấy liên tục, nuôi cấy theo mẻ bổ sung cơ chất, v.v. đã được biết đến trong kỹ thuật, nhưng bước nuôi cấy vi sinh vật không bị giới hạn cụ thể ở đây. Môi trường và các điều kiện nuôi cấy khác được sử dụng để nuôi cấy vi sinh vật theo sáng chế không bị giới hạn cụ thể, nhưng bất kỳ môi trường thông thường được sử dụng trong nuôi cấy vi sinh vật có thể được sử dụng. Cụ thể là, vi sinh vật theo sáng chế có thể được nuôi cấy trong các điều kiện hiếu khí trong môi trường thông thường chứa nguồn cacbon, nguồn nitơ, nguồn phospho, hợp chất vô cơ, axit amin, và/hoặc vitamin phù hợp, v.v. trong khi điều chỉnh nhiệt độ, độ pH, v.v. Cụ thể là, độ pH có thể được điều chỉnh bằng cách sử dụng hợp chất bazơ (ví dụ, natri hydroxit, kali hydroxit, hoặc amoniac) hoặc hợp chất axit (ví dụ, axit phosphoric hoặc axit sulfuric) để thu được độ pH tối ưu (ví dụ, độ pH 5 đến 9, tốt hơn là độ pH 6 đến 8, và tốt nhất là độ pH 6,8), nhưng phương pháp điều chỉnh độ pH không bị giới hạn ở đây. Ngoài ra, oxy hoặc khí chứa oxy có thể được bơm vào môi trường nuôi cấy để duy trì trạng thái hiếu khí của môi trường nuôi cấy; hoặc khí nitơ, hydrogên, hoặc cacbon đioxit có thể được bơm vào môi trường nuôi cấy mà không cần bơm khí để duy trì tình trạng kỵ khí hoặc vi hiếu khí của môi trường nuôi cấy, nhưng khí không bị giới hạn ở đây. Ngoài ra, nhiệt độ môi trường nuôi cấy có thể được duy trì ở 20°C đến 45°C, và tốt hơn là 25°C đến 40°C, và nuôi cấy có thể được thực hiện trong khoảng từ 10 giờ đến 160 giờ, nhưng sáng chế không bị giới hạn ở đây. Ngoài ra, trong quá trình nuôi cấy, chất chống tạo bọt (ví dụ, polyglycol este axit béo)

có thể được bổ sung vào để ngăn tạo bọt, nhưng sáng chế không giới hạn ở đây.

Ngoài ra, nguồn cacbon được sử dụng trong môi trường để nuôi cấy, sacarit và các carbohydrat (ví dụ, glucoza, sucroza, lactoza, fructoza, maltoza, ri đường, tinh bột, và xenluloza), các dầu và các chất béo (ví dụ, dầu đậu nành, dầu hướng dương, dầu đậu phộng, và dầu dừa), các axit béo (ví dụ, axit palmitic, axit stearic, và axit linoleic), các rượu (ví dụ, glyxerol và etanol), các axit hữu cơ (ví dụ, axit axetic), v.v. có thể được sử dụng riêng lẻ hoặc kết hợp, nhưng nguồn cacbon không bị giới hạn ở đây. Nguồn nitơ, hợp chất hữu cơ chứa nitơ (ví dụ, pepton, chiết xuất nấm men, nước thịt, dịch chiết mạch nha, rượu ngô, bột đậu nành, và urê), và hợp chất vô cơ (ví dụ, amoni sulfat, amoni clorua, amoni phosphat, amoni cacbonat, và amoni nitrat), v.v. có thể được sử dụng riêng lẻ hoặc kết hợp, nhưng nguồn nitơ không bị giới hạn ở đây. Nguồn phospho, kali dihydro phosphat, đikali hydro phosphat, các muối chứa natri tương ứng của chúng có thể được sử dụng riêng lẻ hoặc kết hợp, nhưng nguồn phospho không bị giới hạn ở đây. Ngoài ra, môi trường có thể chứa các nguyên liệu thúc đẩy tăng trưởng thiết yếu, như muối kim loại (ví dụ, magie sulfat hoặc sắt sulfat), các axit amin, và các vitamin.

MAAs được sản xuất bằng nuôi cấy có thể được tiết vào trong môi trường hoặc giữ lại trong các tế bào.

Như được sử dụng trong sáng chế, thuật ngữ “môi trường” đề cập đến môi trường nuôi cấy để nuôi cấy vi sinh vật theo sáng chế và/hoặc sản phẩm thu được sau nuôi cấy. Môi trường là khái niệm bao gồm cả dạng chứa vi sinh vật và dạng trong đó vi sinh vật được loại bỏ bằng ly tâm, lọc v.v., từ dịch nuôi cấy chứa vi sinh vật.

Trong bước thu hồi MAAs được sản xuất trong bước nuôi cấy theo sáng chế ở trên, MAAs đích có thể được thu thập từ dung dịch nuôi cấy sử dụng phương pháp phù hợp được biết đến trong kỹ thuật theo phương pháp nuôi cấy. Ví dụ, ly tâm, lọc, sắc ký trao đổi ion, kết tinh, HPLC-High Performance Liquid Chromatography (sắc ký lỏng hiệu năng cao), v.v. có thể được sử dụng, và MAAs đích có thể được thu hồi từ vi sinh vật được nuôi cấy hoặc môi trường nuôi cấy sử dụng phương pháp phù hợp được biết đến trong kỹ thuật. Bước thu hồi MAAs còn có thể bao gồm bước tách và/hoặc bước tinh chế.

Khía cách khác của sáng chế là đề xuất việc sử dụng vi sinh vật theo sáng chế để sản xuất axit amin dạng mycosporin (MMA).

“Vi sinh vật” và “axit amin dạng mycosporin” như được mô tả ở trên.

Ví dụ thực hiện sáng chế

Sau đây, sáng chế sẽ được mô tả chi tiết hơn thông qua các ví dụ. Tuy nhiên, các ví dụ này được đưa ra chỉ nhằm mục đích minh họa, và phạm vi sáng chế không bị giới hạn bởi các ví dụ này.

Chuẩn bị các vi sinh vật tái tổ hợp sản xuất MAAs có nguồn gốc từ *E.coli* và sản xuất MAAs sử dụng vi sinh vật tái tổ hợp này.

Ví dụ 1: Chuẩn bị các chủng trong đó hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza được tăng cường

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza được tăng cường đã được chuẩn bị. Cụ thể là, gen *aroG* (2-dehydro-3-deoxyphosphoheptonat aldolaza được biểu thị bởi các SEQ ID NO: 1 và SEQ ID NO: 2) được đưa thêm vào dựa trên chủng *E.coli* W3110. Các mạch khuôn và các đoạn mồi được sử dụng để điều chế các plasmit được thể hiện trong Bảng 1 bên dưới.

Bảng 1

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn mồi được sử dụng (Mồi xuôi, mồi ngược)
fhuA arm 1	ADN bộ gen W3110	SEQ ID NO: 3, SEQ ID NO: 4
fhuA arm 2	ADN bộ gen W3110	SEQ ID NO: 5, SEQ ID NO: 6
Pn_aroG	ADN bộ gen W3110	SEQ ID NO: 7, SEQ ID NO: 8
Ptrc	pECCG117 Ptrc_GFP	SEQ ID NO: 9, SEQ ID NO: 10
Pcj1 (KR10-620092B)	pECCG117 Pcj1_GFP	SEQ ID NO: 11, SEQ ID NO: 12
aroG	ADN bộ gen W3110	SEQ ID NO: 13, SEQ ID NO: 8

Sau khi khuếch đại các đoạn gen thông qua PCR bằng cách sử dụng các mạch khuôn và các đoạn mồi ở trên, các đoạn đã khuếch đại được gắn vào vector pSKH trong đó các đoạn gen fhuA arm 1 và fhuA arm 2 được xử lý bằng các enzym giới hạn *Bam*H1-*Spe*I, sử dụng kit tách dòng In-Fusion^R HD (Clontech), và vector được điều chế được đặt tên là pSKH-ΔfhuA. Do fhuA gen bị xóa, sự lây truyền thể thực khuẩn của *E.coli* bị ức chế.

Đoạn gen Pn_aroG được gắn vào vector pSKH-ΔfhuA được cắt bằng các enzym

giới hạn *Spe1-EcoRV*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Sau quá trình phân hủy các đoạn gen *Ptrc* và *Pcj1* được biết đến là các trình tự khởi động được tăng cường, bằng các enzym giới hạn *Spe1-Nde1* và sự phân hủy đoạn gen *aroG* bằng các enzym giới hạn *Nde1-EcoRV*, các đoạn gen *Ptrc* và *aroG* hoặc *Pcj1* (patent Hàn Quốc số 10-620092) và các đoạn gen *aroG* tương ứng được gắn vào vector pSKH- Δ fhuA được cắt bằng các enzym giới hạn *Spe1-EcoRV*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Các vector đã điều chế được đặt tên là pSKH- Δ fhuA-Pn-*aroG*, pSKH- Δ fhuA-*Ptrc-aroG*, và pSKH- Δ fhuA-*Pcj1-aroG*, tương ứng.

Các vector trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của quá trình tách dòng và trình tự gen trong từng vector, và sau đó, được biến nạp vào từng chủng *E. coli* W3110 loại hoang dại bằng kỹ thuật xung điện. Từng gen đã biến nạp được đưa vào nhiễm sắc thể bằng cách tái tổ hợp sơ cấp (trao đổi chéo), và vùng plasmit được cắt bỏ khỏi nhiễm sắc thể thông qua tái tổ hợp thứ cấp (trao đổi chéo). Đối với từng chủng *E. coli* đã biến nạp, trong đó quá trình tái tổ hợp thứ cấp được hoàn thành, sự đưa vào gen *aroG* được xác nhận bằng PCR bằng cách sử dụng các đoạn môi được biểu thị bởi các SEQ ID NO: 14 (môi xuôi) và SEQ ID NO: 8 (môi ngược).

Ví dụ 2: Điều chế gen sinh tổng hợp shinorin có nguồn gốc vi tảo từ vector biểu hiện quá mức

Cụm gen gốc *variabilis* để sinh tổng hợp shinorin bao gồm bốn gen (*Ava_ABCD*), mã hóa 2-đemetyl 4-đeoxygadusol synthaza, O-metyltransferaza, C-N ligaza, và non-ribosomal peptit synthetaza. Cụm gen để sinh tổng hợp shinorin được xác định bằng cách sử dụng ADN bộ gen của *A. variabilis* ATCC29413. Vector chứa gen sinh tổng hợp shinorin có nguồn gốc từ *A. variabilis* ATCC29413 được điều chế bằng cách sử dụng vector pECCG117-*Pcj1*-GFP-terminator. Tên của vector biểu hiện gen sinh tổng hợp shinorin và mạch khuôn và các đoạn môi tương ứng để điều chế vector được thể hiện trong Bảng 2 bên dưới.

Bảng 2

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn môi được sử dụng (Môi xuôi, môi ngược)
<i>Ava_ABCD</i>	ADN bộ gen của <i>A. variabilis</i> ATCC29413	SEQ ID NO: 15, SEQ ID NO: 16

Các đoạn gen thu được bằng cách sử dụng mạch khuôn và các đoạn mồi ở trên, và từng đoạn gen được gắn vào vector pECCG 117_*Pcj1*_GFP_terminator, được xử lý bằng các enzym giới hạn *EcoRV*-*XbaI*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Vector đã điều chế được đặt tên là pECCG117_*Pcj1*_Ava_*ABCD*, và sự thành công của quá trình tách dòng và trình tự gen của vector được xác nhận bằng cách giải trình tự. Trình tự nucleotit của gen *Ava*_ABCD được biểu diễn bởi SEQ ID NO: 17.

Ví dụ 3: Đánh giá khả năng sản sinh shinorin của chủng trong đó hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza được tăng cường

Plasmid pECCG117_*Pcj1*_Ava_*ABCD* đã điều chế trong Ví dụ 2 được đưa vào từng chủng đã chuẩn bị trong Ví dụ 1, trong đó gen *aroG* được tăng cường, và chủng W3110 loại hoang dại bằng kỹ thuật xung điện, và từng chủng đã biến nạp được cấy trải trên môi trường rắn LB. Các chủng được nuôi cấy ở nhiệt độ 37°C qua đêm trong tủ ẩm, và sau đó một vòng que cấy bạch kim được cấy chuyển vào 25 mL môi trường chuẩn độ [thành phần môi trường: glucoza 40 g/L, KH₂PO₄ 0,3 g/L, K₂HPO₄ 0,6 g/L, (NH₄)₂SO₄ 15 g/L, MgSO₄·7H₂O 1 g/L, NaCl 2,5 g/L, natri xitrat 1,2 g/L, chiết xuất nấm men 2,5 g/L, canxi cacbonat 40 g/L: độ pH 7,0], được ủ ở nhiệt độ 37°C trong tủ ẩm với tốc độ quay 200 vòng/phút trong 48 giờ. Các chủng thu được được phân tích bằng HPLC (Waters Corp.) và các kết quả như được thể hiện trong Bảng 3 bên dưới.

Bảng 3

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
W3110/pECCG117_ <i>Pcj1</i> _Ava_ <i>ABCD</i>	19,4	348
W3110Δ <i>fhuA</i> ::Pn- <i>aroG</i> /pECCG117_ <i>Pcj1</i> _Ava_ <i>ABCD</i>	19,6	418
W3110Δ <i>fhuA</i> :: <i>Ptrc</i> - <i>aroG</i> /pECCG117_ <i>Pcj1</i> _Ava_ <i>ABCD</i>	20,3	589
W3110Δ <i>fhuA</i> :: <i>Pcj1</i> - <i>aroG</i> /pECCG117_ <i>Pcj1</i> _Ava_ <i>ABCD</i>	20,5	674

Như thể hiện trong bảng 3 ở trên, nồng độ của shinorin đã sản sinh trong chủng (W3110Δ*fhuA*::Pn-*aroG*/pECCG117_*Pcj1*_Ava_*ABCD*), ở đó gen *aroG* được tăng cường, đã tăng lên khoảng 20% so với nhóm đối chứng. Cụ thể là, trong trường hợp các chủng mà gen *aroG* được tăng cường bằng cách tăng cường trình tự khởi động (tức là, W3110Δ*fhuA*::*Ptrc*-*aroG*/pECCG117_*Pcj1*_Ava_*ABCD* và W3110Δ*fhuA*::*Pcj1*-

aroG/pECCG117_Pcj1_Ava_ABCD), nồng độ shinorin đã tăng lên 69% và 94%, tương ứng.

Ví dụ 4: Chuẩn bị chủng trong đó hoạt tính của phosphoenolpyruvat syntetaza được tăng cường

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của phosphoenolpyruvat syntetaza được tăng cường đã được chuẩn bị. Cụ thể là, gen *pps* (phosphoenolpyruvat syntetaza được biểu thị bởi SEQ ID NO: 18 và SEQ ID NO 19) được đưa thêm vào dựa trên chủng *E.coli* W3110. Mạch khuôn và các đoạn môi đã sử dụng để điều chế plasmid được thể hiện trong Bảng 4 bên dưới.

Bảng 4

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn môi được sử dụng (Môi xuôi, môi ngược)
Pn- <i>pps</i>	ADN bộ gen W3110	SEQ ID NO: 20, SEQ ID NO: 21
<i>pps</i>	ADN bộ gen W3110	SEQ ID NO: 22, SEQ ID NO: 21

Sau khi khuếch đại các đoạn gen bằng cách sử dụng mạch khuôn và các đoạn môi ở trên, đoạn gen Pn-*pps* được gắn vào vector pSKH- Δ fhuA, được cắt bằng enzym giới hạn *Spe1-EcoRV* các, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Sau quá trình phân hủy các đoạn gen *Ptrc* và *Pcj1* đã điều chế trong Ví dụ 1 bằng các enzym giới hạn *Spe1-Nde1* và phân hủy đoạn gen *pps* bằng các enzym giới hạn *Nde1-EcoRV*, các đoạn gen *Ptrc* và *pps* hoặc các đoạn gen *Pcj1* và *pps* tương ứng được gắn vào vector pSKH- Δ fhuA, được cắt bằng các enzym giới hạn *Spe1-EcoRV*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Các vector đã điều chế được đặt tên là pSKH- Δ fhuA-Pn-*pps*, pSKH- Δ fhuA-*Ptrc-pps*, và pSKH- Δ fhuA-*Pcj1-pps*, tương ứng.

Các vector ở trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của tách dòng và trình tự gen trong từng vector, và sau đó, được biến nạp vào từng chủng *E.coli* W3110 loại hoang dại bằng kỹ thuật xung điện. Từng gen đã biến nạp được đưa vào nhiễm sắc thể bằng cách tái tổ hợp sơ cấp (trao đổi chéo), và vùng plasmid được cắt bỏ khỏi nhiễm sắc thể thông qua tái tổ hợp thứ cấp (trao đổi chéo). Đối với từng chủng *E.coli* đã biến nạp, trong đó tái tổ hợp thứ cấp được hoàn thành, sự đưa vào gen *pps* được xác nhận bằng PCR bằng cách sử dụng các đoạn môi được biểu thị bởi SEQ ID NO: 14 (môi xuôi) và SEQ ID NO 21 (môi ngược).

Ví dụ 5: Đánh giá khả năng sản sinh shinorin của chủng trong đó hoạt tính của phosphoenolpyruvat syntetaza được tăng cường

Plasmid pECCG117_*Pcj1_Ava_ABCD* đã điều chế trong Ví dụ 2 được đưa vào từng chủng đã chuẩn bị trong Ví dụ 4, trong đó gen *pps* được đưa vào, và chủng W3110 loại hoang dại bằng kỹ thuật xung điện, và từng chủng đã biến nạp được cấy trải trên môi trường rắn LB. Các chủng được nuôi cấy ở nhiệt độ 37°C qua đêm trong tủ ẩm, và một vòng que cấy bạch kim được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 3, được ủ ở nhiệt độ 37°C trong tủ ẩm ở tốc độ 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 5 bên dưới.

Bảng 5

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
W3110/pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,4	348
W3110Δ <i>fhuA::Pn-pps</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	21,3	494
W3110Δ <i>fhuA::Ptrc-pps</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	20,8	511
W3110Δ <i>fhuA::Pcj1-pps</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	21,4	556

Như thể hiện trên bảng 5 ở trên, nồng độ của shinorin đã sản sinh trong chủng mà gen *pps* được tăng cường đã tăng lên 41%, và trong trường hợp trong đó hoạt tính của nó được tăng cường bằng cách thay thế với trình tự khởi động mạnh, nồng độ shinorin đã tăng lên đến 60% so với nhóm đối chứng.

Ví dụ 6: Chuẩn bị chủng trong đó hoạt tính của transketolaza I/II được tăng cường

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của transketolaza được tăng cường đã được chuẩn bị. Cụ thể là, dựa trên chủng *E.coli* W3110, gen *tktA* (transketolaza được biểu thị bởi SEQ ID NO: 23 và SEQ ID NO 24) được đưa vào đó. Mạch khuôn và các đoạn mồi đã sử dụng trong quá trình điều chế các plasmid được thể hiện trong Bảng 6 bên dưới.

Bảng 6

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn mồi được sử dụng (Mồi xuôi, mồi ngược)
--------------	-----------------------------	---

Pn- <i>tktA</i>	ADN bộ gen W3110	SEQ ID NO: 25, SEQ ID NO: 26
<i>tktA</i>	ADN bộ gen W3110	SEQ ID NO: 27, SEQ ID NO: 26

Sau khi khuếch đại gen các đoạn thông qua PCR bằng cách sử dụng mạch khuôn và các đoạn mồi ở trên, đoạn gen Pn-*tktA* được gắn vào vector pSKH- Δ fhuA, được cắt bằng các enzym giới hạn *Spe*1-*Eco*RV, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Sau quá trình phân hủy các đoạn gen *Ptrc* và *Pcj1* được điều chế trong ví dụ 1 bằng các enzym giới hạn *Spe*1, *Nde*1 và phân hủy đoạn gen *tktA* bằng các enzym giới hạn *Nde*1-*Eco*RV, các đoạn gen *Ptrc* và *tktA* hoặc các đoạn gen *Pcj1* và *tktA* tương ứng được gắn vào vector pSKH- Δ fhuA, được cắt bằng các enzym giới hạn *Spe*1-*Eco*RV, bằng cách sử dụng kit tách dòng In-Fusion^R HD (Clontech). Các vector đã điều chế được đặt tên là pSKH- Δ fhuA-Pn-*tktA*, pSKH- Δ fhuA-*Ptrc-tktA*, và pSKH- Δ fhuA-*Pcj1-tktA*, tương ứng.

Các vector ở trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của tách dòng và trình tự gen trong từng vector, và sau đó, được biến nạp vào từng chủng *E.coli* W3110 loại hoang dại bằng kỹ thuật xung điện. Từng gen đã biến nạp được đưa vào nhiễm sắc thể bằng cách tái tổ hợp sơ cấp (trao đổi chéo), và vùng plasmit được cắt bỏ khỏi nhiễm sắc thể thông qua tái tổ hợp thứ cấp (trao đổi chéo). Đối với từng chủng *E.coli* đã biến nạp, trong đó tái tổ hợp thứ cấp được hoàn thành, sự đưa vào gen *tktA* được xác nhận bằng PCR sử dụng các đoạn mồi được biểu thị bởi SEQ ID NO: 14 (mồi xuôi) và SEQ ID NO 26 (mồi ngược).

Ví dụ 7: Đánh giá khả năng sản sinh shinorin ở chủng trong đó hoạt tính của transketolaza được tăng cường

Plasmit pECCG117-*Pcj1_Ava_ABCD* đã điều chế trong ví dụ 2 được đưa vào từng các chủng đã chuẩn bị trong ví dụ 6, trong đó gen *tktA* được đưa vào, và chủng W3110 loại hoang dại bằng kỹ thuật xung điện, và từng chủng đã biến nạp được cấy trên môi trường rắn LB. Các chủng được nuôi cấy ở nhiệt độ 37°C qua đêm trong tủ ẩm, và một vòng que cấy bạch kim được nuôi cấy qua đêm của từng chủng được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 3, được ủ ở nhiệt độ 37°C trong tủ ẩm với tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 7 bên dưới.

Bảng 7

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
W3110/pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,4	348
W3110Δ <i>fhuA</i> ::Pn- <i>tktA</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,5	364
W3110Δ <i>fhuA</i> :: <i>Ptrc-tktA</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,3	447
W3110Δ <i>fhuA</i> :: <i>Pcj1-tktA</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,5	461

Như thể hiện trên bảng 7 ở trên, nồng độ của shinorin được sản sinh trong chủng mà gen *tktA* được tăng cường đã tăng lên 4,5%, và trong trường hợp trong đó hoạt tính của nó được tăng cường bằng cách thay thế bằng trình tự khởi động mạnh, nồng độ shinorin đã tăng lên tới 32% so với nhóm đối chứng.

Ví dụ 8: Chuẩn bị chủng trong đó các hoạt tính của 2-đehydro-3-đeoxyphosphoheptonat aldolaza/phosphoenolpyruvat syntetaza/transketolaza được tăng cường

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của từng 2-đehydro-3-đeoxyphosphoheptonat aldolaza/phosphoenolpyruvat syntetaza/transketolaza được tăng cường đã được chuẩn bị. Cụ thể là, dựa trên chủng *E.coli* W3110, từng gen *aroG*, gen *pps*, và gen *tktA* được đưa thêm vào trong đó. Các mạch khuôn và các đoạn mồi được sử dụng trong quá trình điều chế các plasmid được thể hiện trong Bảng 8 bên dưới.

Bảng 8

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn mồi được sử dụng (Mồi xuôi, mồi ngược)
<i>Pcj1-aroG</i>	pSKH-Δ <i>fhuA</i> - <i>Pcj1-aroG</i>	SEQ ID NO: 11, SEQ ID NO: 28
<i>Pcj1-pps</i>	pSKH-Δ <i>fhuA</i> - <i>Pcj1-pps</i>	SEQ ID NO: 29, SEQ ID NO: 30
<i>Pcj1-tktA</i>	pSKH-Δ <i>fhuA</i> - <i>Pcj1-tktA</i>	SEQ ID NO: 31, SEQ ID NO: 26

Sau khi khuếch đại gen các đoạn sử dụng các mạch khuôn và các đoạn mồi ở trên, từng đoạn gen của chúng được gắn vào vector pSKH-Δ*fhuA*, được cắt bằng các enzym giới hạn *SpeI*-*EcoRV*, sử dụng In-Fusion^R HD kit tách dòng (Clontech). Vector được điều chế được đặt tên là pSKH-Δ*fhuA*-*Pcj1-aroG*-*Pcj1-ppsA*-*Pcj1-tktA*.

Vector ở trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công

của quá trình tách dòng và trình tự các gen trong vector, và sau đó, được biến nạp vào chủng *E.coli* W3110 loại hoang dại bằng kỹ thuật xung điện. Các gen đã biến nạp được đưa vào nhiễm sắc thể bằng cách tái tổ hợp sơ cấp (trao đổi chéo), và vùng plasmid được cắt bỏ khỏi nhiễm sắc thể thông qua tái tổ hợp thứ cấp (trao đổi chéo). Về các chủng *E.coli* đã biến nạp, trong đó tái tổ hợp thứ cấp được hoàn thành, sự đưa vào các gen *aroG*, *pps*, và *tktA* được xác nhận bằng PCR sử dụng các đoạn mồi được biểu thị bởi các SEQ ID NO: 14 (mồi xuôi) và SEQ ID NO: 26 (mồi ngược).

Ví dụ 9: Đánh giá khả năng sản sinh shinorin của chủng trong đó các hoạt tính của 2-đehydro-3-đeoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza được tăng cường

Plasmid pECCG117_*Pcj1_Ava_ABCD* đã điều chế trong ví dụ 2 được đưa vào từng chủng đã chuẩn bị trong ví dụ 8, trong đó *aroG*, *pps*, và *tktA* các gen được đưa vào, và chủng W3110 loại hoang dại bằng kỹ thuật xung điện, và từng chủng đã biến nạp được cấy trải trên môi trường rắn LB. Các chủng được nuôi cấy ở nhiệt độ 37°C qua đêm trong tủ ấm, và một vòng que cấy bạch kim được nuôi cấy qua đêm của từng chủng được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 3, được ủ ở nhiệt độ 37°C trong tủ ấm ở tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 9 bên dưới.

Bảng 9

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
W3110/pECCG117_ <i>Pcj1_Ava_ABCD</i>	19,4	348
W3110Δ <i>fhuA::Pcj1-aroG-Pcj1-pps-Pcj1-tktA</i> / pECCG117_ <i>Pcj1_Ava_ABCD</i>	18,2	1.279

Như thể hiện trong bảng 9 ở trên, nồng độ của shinorin đã sản sinh trong chủng mà ba loại gen (tức là, *aroG*, *pps*, và *tktA*) được kết hợp và tăng cường đã tăng lên 267% so với nhóm đối chứng. Đây là một kết quả bất ngờ cho thấy sự cải thiện ngoài mong đợi so với tổng các mức tăng tối đa thu được bằng cách thay thế trình tự khởi động của từng gen với trình tự khởi động mạnh. Có nghĩa là, xác nhận rằng khi ba gen (tức là, *aroG*, *pps*, và *tktA*) được kết hợp, có thể sản sinh shinorin với nồng độ cao hơn.

Ví dụ 10: Chuẩn bị chủng trong đó hoạt tính của 3-đehydroquinat dehydrataza bị bất hoạt

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của 3-dehydroquinate dehydratase (*aroD*) bị bất hoạt được chuẩn bị.

Cụ thể là, gen kháng chloramphenicol của plasmid pKD3 được sử dụng làm chỉ thị chèn gen, và catxet xóa *aroD*, bao gồm một phần của gen *aroD* và gen kháng chloramphenicol của plasmid pKD3, được điều chế bằng PCR sử dụng các đoạn môi được biểu thị bởi SEQ ID NO: 32 (môi xuôi) và SEQ ID NO: 33 (môi ngược). Các tế bào khả biến được điều chế bằng cách biến nạp chủng *E.coli* W3110 loại hoang dại và chủng đã chuẩn bị trong ví dụ 8, trong đó các gen *aroG*, *pps*, và *tklA* được đưa vào, với plasmid pKD46 bao gồm gen tái tổ hợp lamda đỏ, tiếp theo là tạo ra sự biểu hiện gen tương ứng bằng cách sử dụng arabinosa. Sau khi đưa catxet xóa *aroD* vào các tế bào khả biến bằng kỹ thuật xung điện, thu được các tế bào khả biến được cấy trên môi trường rắn LB chứa 30 mg/L chloramphenicol. Chủng thu được được phản ứng PCR sử dụng các đoạn môi được biểu thị bởi SEQ ID NO: 34 (môi xuôi) và SEQ ID NO: 35 (môi ngược), và gen xóa *aroD* được xác nhận bằng cách quan sát đoạn 1.300 bp được khuếch đại.

Ví dụ 11: Đánh giá khả năng sản sinh shinorin của chủng trong đó 3-dehydroquinate dehydratase bị bất hoạt

Plasmid pECCG117 *Pcj1_Ava_ABCD* đã điều chế trong ví dụ 2 được đưa vào chủng đã chuẩn bị trong ví dụ 10, trong đó gen *aroD* bị xóa, bằng kỹ thuật xung điện, và chủng thu được được cấy trên môi trường rắn LB. Chủng được nuôi cấy ở nhiệt độ 37°C qua đêm trong tủ ấm, và một vòng que cấy bạch kim được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 3, được ủ ở nhiệt độ 37°C trong tủ ấm ở tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 10 bên dưới.

Bảng 10

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
W3110ΔfhuA:: <i>Pcj1-aroG-Pcj1-ppsA-Pcj1-tklA</i> / pECCG117_ <i>PCJ1_Ava_ABCD</i>	18,6	1.248
W3110Δ <i>aroD</i> ΔfhuA:: <i>Pcj1-aroG-Pcj1-ppsA-Pcj1-tklA</i> / pECCG117_ <i>PCJ1_Ava_ABCD</i>	17,3	2.077

Như thể hiện trên bảng 10 ở trên, nồng độ của shinorin đã sản sinh trong chủng, ở đó gen *aroD* bị xóa thêm, đã tăng lên 66% so với chủng sản xuất shinorin tại đó các gen *aroG*, *pps*, và *tklA* được tăng cường. Chủng W3110ΔfhuA::*Pcj1-aroG-Pcj1-ppsA-*

Pcj1-tktA/pECCG117_PCJ1_Ava_ABCD, là chủng trong đó các gen *aroG*, *pps*, và *tktA* được tăng cường, được đặt tên là CB06-0020, và được ký gửi tại Bảo tàng giống chuẩn Vi sinh vật Hàn Quốc theo Hiệp ước Budapest vào ngày 14 tháng 02 năm 2018, với số đăng ký KCCM12224P.

Chuẩn bị vi sinh vật tái tổ hợp sản sinh MAAs dựa trên *Corynebacterium glutamicum* và sản xuất MAAs bằng cách sử dụng vi sinh vật này

Ví dụ 12: Điều chế vectơ trong đó hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza được tăng cường và đánh giá khả năng sản sinh shinorin của nó

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó hoạt tính của 2-dehydro-3-deoxyphosphoheptonat aldolaza được tăng cường đã được điều chế. Cụ thể là, *aroG* gen (2-dehydro-3-deoxyphosphoheptonat aldolaza được biểu thị bởi SEQ ID NO: 36 và SEQ ID NO: 37) được đưa thêm vào dựa trên chủng *Corynebacterium glutamicum* ATCC13032. Mạch khuôn và các đoạn mồi được sử dụng để điều chế plasmit được thể hiện trong Bảng 11 bên dưới.

Bảng 11

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn mồi được sử dụng (Mồi xuôi, mồi ngược)
Pn-cgl <i>aroG</i>	c.gl 13032 ADN bộ gen	SEQ ID NO: 38, SEQ ID NO: 39
<i>Pcj7</i> -cgl <i>aroG</i>	c.gl 13032 ADN bộ gen	SEQ ID NO: 40, SEQ ID NO: 41

Sau khi thu được các đoạn gen bằng cách sử dụng mạch khuôn và các đoạn mồi ở trên, từng đoạn gen được gắn vào các vectơ pECCG 117 và pECCG 117_*Pcj7*_GFP_terminator (patent Hàn Quốc số 10-620092, p117-cj7-gfp), được xử lý bằng các enzym giới hạn *EcoRV/XbaI*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Các vectơ được điều chế được đặt tên tương ứng là pECCG117_Pn_cgl *aroG* và pECCG117_*Pcj7*_cgl *aroG*. Các vectơ trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của tách dòng và trình tự gen trong từng vectơ.

Trước tiên, vì vi sinh vật thuộc chi *Corynebacterium* không thể sản sinh shinorin, chủng trong đó đường sinh tổng hợp shinorin được đưa vào được điều chế. Cụ thể là, gen *Ava_ABCD* được phản ứng PCR sử dụng pECCG117_*Ptrc_Ava_ABCD* làm mạch khuôn cùng với cặp mồi được biểu thị bởi các SEQ ID NO: 42 (mồi xuôi) và SEQ ID

NO: 43 (mỗi ngược). pDZTn_Ava_ABCD được điều chế bằng cách gắn đoạn PCR khoảng 7 kb vào vector pDZTn (công bố đơn quốc tế số WO 2009-125992A), được xử lý bằng enzym giới hạn *NdeI*, sử dụng kit tách dòng In-Fusion^R HD (Clontech). Sau đó, đoạn trình tự khởi động O2 (patent Hàn Quốc số 10-1632642) được phản ứng PCR sử dụng cặp mồi được biểu thị bởi SEQ ID NO: 44 (mỗi xuôi) và SEQ ID NO: 45 (mỗi ngược), và được gắn vào pDZTn_Ava_ABCD, được xử lý bằng enzym giới hạn *NdeI*, sử dụng kit tách dòng In-Fusion^R HD (Clontech), điều chế pDZTn_PO2_Ava_ABCD.

Plasmid tái tổ hợp được biến nạp vào ATCC13032 loại hoang dại bằng kỹ thuật xung điện (van der Rest và cộng sự 1999), và plasmid được đưa vào nhiễm sắc thể bằng cách tái tổ hợp sơ cấp (trao đổi chéo), và vùng plasmid được cắt bỏ khỏi nhiễm sắc thể thông qua tái tổ hợp thứ cấp (trao đổi chéo).

Đối với từng chủng *Corynebacterium glutamicum* đã biến nạp, trong đó tái tổ hợp thứ cấp được hoàn thành, sự đưa vào gen *Ava_ABCD* được xác nhận bằng PCR bằng cách sử dụng cặp mồi đặc hiệu cho gen được biểu thị bởi SEQ ID NO: 42 (mỗi xuôi) và SEQ ID NO: 43 (mỗi ngược). Chủng được điều chế được đặt tên là *Corynebacterium glutamicum* 13032 ΔN1021PO2_Ava_ABCD.

Từng vector pECCG117_Pn_cgl *aroG* và pECCG117_Pcj7_cgl *aroG* được biến nạp vào chủng *Corynebacterium glutamicum* 13032 ΔN1021_PO2_Ava_ABCD bằng kỹ thuật xung điện.

Các chủng được điều chế ở trên và nhóm đối chứng *Corynebacterium glutamicum* ATCC13032 (c.gl 13032) được nuôi cấy qua đêm trong môi trường rắn BHIS chứa kanamycin, và một vòng que cấy bạch kim được nuôi cấy vào 25 mL môi trường chuẩn độ [thành phần môi trường: glucoza 40 g/L, KH₂PO₄ 1 g/L, (NH₄)₂SO₄ 10 g/L, MgSO₄·7H₂O 5g/L, NaCl 5 g/L, chiết xuất nấm men 5 g/L, canxi cacbonat 30 g/L: độ pH 7,0], được ủ ở nhiệt độ 37°C trong tủ ấm với tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 12 bên dưới.

Bảng 12

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
c.gl 13032	72,1	-
c.gl 13032ΔN1021_PO2_Ava_ABCD	71,5	180

c.gl 13032ΔN1021_PO2_Ava_ABCD/ pECCG117_Pn_cgl aroG	69,6	250
c.gl 13032ΔN1021_PO2_Ava_ABCD/ pECCG117_Pcj7_cgl aroG	71,9	324

Như thể hiện trong bảng 12 ở trên, khi mức độ biểu hiện *aroG* được tăng lên ở chủng chứa gen sinh tổng hợp shinorin, nồng độ của shinorin đã tăng lên 39%. Cụ thể là, khi trình tự khởi động được tăng cường, xác nhận rằng nồng độ shinorin được cải thiện lên tới 80%.

Ví dụ 13: Điều chế vector trong đó các hoạt tính của phosphoenolpyruvat syntetaza/transketolaza được tăng cường và đánh giá khả năng sản sinh shinorin của nó

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *Corynebacterium glutamicum* trong đó hoạt tính của *tkt* hoặc *pps* được tăng cường đã được điều chế. Cụ thể là, *tkt* (transketolaza được biểu thị bởi SEQ ID NO: 95 và SEQ ID NO: 96) hoặc *pps* (phosphoenolpyruvat syntetaza được biểu thị bởi SEQ ID NOS: 97 và SEQ ID NO: 98) được đưa thêm vào dựa trên chủng *Corynebacterium glutamicum* ATCC13032. Mạch khuôn và các đoạn nối được sử dụng để điều chế plasmit được thể hiện trong Bảng 13 bên dưới.

Bảng 13

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn nối được sử dụng (Mối xuôi, mối ngược)
Pn-cgl <i>tkt</i>	ADN bộ gen c.gl 13032	SEQ ID NO: 46, SEQ ID NO: 47
<i>Pcj7</i> -cgl <i>tkt</i>	ADN bộ gen c.gl 13032	SEQ ID NO: 48, SEQ ID NO: 49
<i>Ptrc</i> -cgl <i>pps</i>	ADN bộ gen c.gl 13032	SEQ ID NO: 50, SEQ ID NO: 51
<i>Pcj7</i> -cgl <i>pps</i>	ADN bộ gen c.gl 13032	SEQ ID NO: 52, SEQ ID NO: 51

Các vector được điều chế bằng cách gắn gen các đoạn, thu được thông qua công nghệ PCR trong đó mạch khuôn phù hợp với sự kết hợp với các đoạn nối, với các vector pECCG117, pECCG117_*Ptrc*_GFP_terminator và pECCG 117_*Pcj7*_GFP_terminator, được xử lý bằng các enzym giới hạn *EcoRV/XbaI*, sử dụng kit tách dòng In-Fusion[®] HD (Clontech). Các vector đã điều chế được đặt tên tương ứng là pECCG117-Pn-*tkt*/pECCG117-*Pcj7*-*tkt* và pECCG117-*Ptrc*-*pps*/pECCG117-*Pcj7*-*pps*. Các vector trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của tách dòng và trình tự gen trong từng vector và sau đó được biến nạp vào chủng *Corynebacterium*

glutamicum 13032 Δ N1021_PO2_Ava_ABCD bằng kỹ thuật xung điện. Từng chủng được nuôi cấy trong môi trường rắn BHIS chứa kanamycin qua đêm và vòng que cấy bạch kim của chúng được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 12, được ủ ở nhiệt độ 37°C trong tủ ấm với tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 14 bên dưới.

Bảng 14

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
c.gl 13032 Δ N1021_PO2_Ava_ABCD	73	175
c.gl 13032 Δ N1021_PO2_Ava_ABCD/ pECCG117_Pn_cgl <i>tkl</i>	72	211
c.gl 13032 Δ N1021_PO2_Ava_ABCD/ pECCG117- <i>Pcj7</i> -cgl <i>tkl</i>	71,5	275
c.gl 13032 Δ N1021_PO2_Ava_ABCD/ pECCG117- <i>Ptrc</i> -cgl <i>pps</i>	70,9	298
c.gl 13032 Δ N1021_PO2_Ava_ABCD/ pECCG117- <i>Pcj7</i> -cgl <i>pps</i>	70,2	302

Như thể hiện trên bảng 14 ở trên, xác nhận rằng khi gen *tkl* hoặc gen *pps* được tăng cường, sự sản sinh shinorin được cải thiện tới 57% hoặc 72%, tương ứng.

Ví dụ 14: Chuẩn bị chủng trong đó các hoạt tính của 2-đehydro-3-deoxyphosphoheptonat aldolaza/phosphoenolpyruvat syntetaza/transketolaza được tăng cường và đánh giá các chủng này

Để tăng khả năng sản sinh MAAs của vi sinh vật, các chủng *E.coli* trong đó các hoạt tính của các gen *aroG*, *pps*, và *tkl* được tăng cường đã được chuẩn bị, và để xác nhận sự có mặt MAAs được sản xuất với hàm lượng cao hơn, 3-đehydroquinat dehydrataza (*aroD*) tiếp tục bị bất hoạt. Cụ thể là, để tăng cường các gen *aroG*, *pps*, và *tkl*, plasmit pDZ- Δ *aroD*-*Pcj7*-*aroG*-*Pcj7*-*pps*-*Pcj7*-*tklA* được điều chế. Các mạch khuôn và các đoạn môi được sử dụng để điều chế plasmit pDZ- Δ *aroD*-*Pcj7*-*aroG*-*Pcj7*-*pps*-*Pcj7*-*tklA* được thể hiện trong Bảng 15 bên dưới.

Bảng 15

Sản phẩm PCR	Các mạch khuôn được sử dụng	Các đoạn môi được sử dụng (Môi xuôi, môi ngược)
<i>Pcj7</i> - <i>aroG</i>	pECCG117_ <i>Pcj7</i> _cgl <i>aroG</i>	SEQ ID NO: 53, SEQ ID NO: 54

<i>Pcj7-tkt</i>	pECCG117- <i>Pcj7-cgl tkt</i>	SEQ ID NO: 55, SEQ ID NO: 56
<i>Pcj7-pps</i>	pECCG117- <i>Pcj7-cgl pps</i>	SEQ ID NO: 57, SEQ ID NO: 58

Trước tiên, để chuẩn bị chủng trong đó gen *aroD* (được biểu thị bởi SEQ ID NO: 89 và SEQ ID NO: 90) của *Corynebacterium glutamicum* bị xóa, plasmit pDZ- Δ *aroD* trong đó khung đọc mở của gen *aroD* bị xóa bên trong được chuẩn bị. Việc xóa gen bên trong của plasmit pDZ- Δ *aroD* đạt được bằng cách thực hiện PCR chéo sử dụng ADN bộ gen của chủng *Corynebacterium glutamicum* ATCC 13032 làm mạch khuôn cùng với SEQ ID NO: 91 và SEQ ID NO: 92, và SEQ ID NO: 93 và SEQ ID NO: 94 làm các cặp mồi xuôi và ngược, sau đó đưa các đoạn gen thu được vào vector pDZ.

Sau đó, từng đoạn gen của các gen *aroG*, *pps*, và *tkt* được khuếch đại thông qua PCR sử dụng các mạch khuôn và các đoạn mồi như thể hiện trong Bảng 15 ở trên, và sau đó được đưa vào vector pDZ- Δ *aroD*, được cắt bằng enzym giới hạn *SpeI*, tương ứng. Hai loại vector ở trên được xác nhận bằng cách giải trình tự liên quan đến sự thành công của tách dòng và trình tự gen trong từng vector, và sau đó, được biến nạp vào chủng *Corynebacterium glutamicum* 13032 Δ N1021 _PO2_Ava_ABCD bằng kỹ thuật xung điện. Từng chủng được nuôi cấy trong môi trường rắn chứa kanamycin qua đêm và vòng que cấy bạch kim được nuôi cấy qua đêm của từng chủng được cấy chuyển vào 25 mL môi trường chuẩn độ của Ví dụ 12, được ủ ở nhiệt độ 37°C trong tủ ấm với tốc độ quay 200 vòng/phút trong 48 giờ. Các kết quả như được thể hiện trong Bảng 16 bên dưới.

Bảng 16

Tên chủng	OD (600 nm)	Nồng độ Shinorin (mg/L)
c.gl 13032 Δ N1021_PO2_Ava_ABCD_ Δ <i>aroD</i>	73	425
c.gl 13032 Δ N1021_PO2_Ava_ABCD_ Δ <i>aroD</i> , <i>Pcj7_aroG</i> , <i>Pcj7_tkt</i> , <i>Pcj7_pps</i>	69	531

Như thể hiện trên bảng 16 ở trên, nồng độ của shinorin đã sản sinh trong chủng có ba loại gen (*aroG*, *pps*, và *tktA*) được tăng cường đã tăng lên khoảng 25%. Đã xác nhận rằng ngay cả trong chủng, mà khả năng sản sinh shinorin đã tăng lên thông qua xóa gen *aroD*, shinorin có thể được sản xuất ở nồng độ cao bằng cách kết hợp ba loại gen. Ngoài ra, có thể giải thích rằng khi gen *aroD* tiếp tục bị bất hoạt trong chủng mà ba loại gen được kết hợp, shinorin có thể còn được sản xuất ở nồng độ cao hơn nữa.

Chuẩn bị vi sinh vật tái tổ hợp sản sinh MAAs dựa vào nấm men và sản xuất MAAs

bằng cách sử dụng vi sinh vật này

Ví dụ 15: Chuẩn bị chủng *Saccharomyces cerevisiae* (*S. cerevisiae*) sản sinh shinorin

Để sử dụng chủng *Saccharomyces cerevisiae* (*S. cerevisiae*) làm chủng sản xuất shinorin, gen sinh tổng hợp shinorin có nguồn gốc từ *A. variabilis* ATCC29413 được đưa vào vector biểu hiện nấm men. Các gen *Ava_A* và *Ava_B* được chèn vào vector pRS-413 sử dụng trình tự khởi động GPD. Cụ thể là, các vùng pGPD-*Ava_A* và pGPD-*Ava_B* được gắn bằng cách sử dụng PCR chồng lắp. Các vector và các sản phẩm PCR được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế vector pRS-413-pGPD-*Ava_A*-pGPD-*Ava_B*.

Sau đó, các gen *Ava_A* và *Ava_B* được chèn vào vector pRS-414 sử dụng trình tự khởi động GPD. Cụ thể là, các vùng pGPD-*Ava_C* và pGPD-*Ava_D* được gắn bằng cách sử dụng PCR chồng lắp, và sau đó, các vector và các sản phẩm PCR được xử lý bằng *Bam*HI và *Sal*I và được gắn bằng cách sử dụng T4 ligaza để điều chế vector pRS-414-pGPD-*Ava_C*-pGPD-*Ava_D*. Các đoạn môi và mạch khuôn ADN được sử dụng để điều chế vector được thể hiện trong Bảng 17 bên dưới.

Bảng 17

Sản phẩm PCR	Các mạch khuôn được sử dụng	SEQ ID NO (Môi xuôi, môi ngược)
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 59, SEQ ID NO: 60
<i>Ava_A</i>	gADN <i>A. variabilis</i> ATCC29413	SEQ ID NO: 61, SEQ ID NO: 62
pGPD- <i>Ava_A</i>	Sản phẩm PCR	SEQ ID NO: 59, SEQ ID NO: 62
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 63, SEQ ID NO: 64
<i>Ava_B</i>	gADN <i>A. variabilis</i> ATCC29413	SEQ ID NO: 65, SEQ ID NO: 66
pGPD- <i>Ava_B</i>	Sản phẩm PCR	SEQ ID NO: 63, SEQ ID NO: 66
pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i>	Sản phẩm PCR	SEQ ID NO: 59, SEQ ID NO: 66
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 67,

		SEQ ID NO: 68
Ava_C	gADN <i>A. variabilis</i> ATCC29413	SEQ ID NO: 69, SEQ ID NO: 70
pGPD-Ava_C	Sản phẩm PCR	SEQ ID NO: 67, SEQ ID NO: 70
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 71, SEQ ID NO: 72
Ava_D	gADN <i>A. variabilis</i> ATCC29413	SEQ ID NO: 73, SEQ ID NO: 74
pGPD-Ava_D	Sản phẩm PCR	SEQ ID NO: 71, SEQ ID NO: 74
pGPD-Ava_C-pGPD-Ava_D	Sản phẩm PCR	SEQ ID NO: 67, SEQ ID NO: 74

Vector pRS-413-pGPD-*Ava_A*-pGPD-*Ava_B* và vector pRS-414-pGPD-*Ava_C*-pGPD-*Ava_D* được đưa vào chủng *Saccharomyces cerevisiae* CEN.PK-1D (*S. cerevisiae* CEN.PK-1D) thông qua biến nạp lithi axetat, và sau đó sự hiện diện của quá trình sản xuất shinorin được xác nhận. Chủng được cấy trải trên môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp và His (tức là, các chỉ thị dinh dưỡng thụ động) được loại trừ, và được nuôi cấy qua đêm ở nhiệt độ 30°C trong tủ ấm. Một vòng que cấy bạch kim của chủng được nuôi cấy qua đêm trong môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp và His được loại trừ, được cấy chuyển vào 25 mL môi trường chuẩn độ của Bảng 18, và sau đó được nuôi cấy ở nhiệt độ 30°C trong tủ ấm với tốc độ quay 150 vòng/phút trong 24 giờ. Các kết quả như được thể hiện trong Bảng 19 bên dưới.

Bảng 18

Thành phần	Nồng độ sử dụng (g/L)
Cơ sở nitơ nấm men (YNB) không chứa các axit amin	6,7
Các hỗn hợp axit amin (không có Leuxin, Tryptophan, Histidin, Uraxil)	2
Glucoza	20

Bảng 19

Tên chủng	24 Giờ		Nồng độ Shinorin (mg/L)
	OD (600nm)	Sacarit dư	
CEN.PK-1D WT + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> + pRS414-pGPD- <i>Ava_C</i> -	11,2	0	331

pGPD-Ava_D			
------------	--	--	--

Kết quả của thí nghiệm đã xác nhận rằng chủng *Saccharomyces cerevisiae* loại hoang dại, không sản xuất shinorin, sản xuất được 331 mg/L shinorin nhờ đưa vào gen sinh tổng hợp shinorin.

Ví dụ 16: Tăng lượng shinorin sản xuất thông qua tăng cường *TKL1* (transketolaza) của *Saccharomyces cerevisiae*

Để tăng khả năng sản sinh MAAs, chủng *Saccharomyces cerevisiae* trong đó hoạt tính của *TKL1* được tăng cường đã được chuẩn bị. Để đạt được mục đích này, sự biểu hiện gen *TKL1* được tăng cường bằng cách tách dòng gen *TKL1* (SEQ ID NO: 110 và SEQ ID NO: 123) thành các vector pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF.

Trình tự khởi động GPD chứa trong vector pRS-415-pGPD là trình tự khởi động của glyxeraldehyt-3-phosphat dehydrogenaza (*GAPDH*) và các gen isozym 3 (TDH3), và nó bao gồm trình tự từ -674 bp đến -1 bp từ codon khởi đầu của ORF của gen THD3.

Trình tự khởi động ADH chứa trong vector pRS-415-pADH là trình tự khởi động của gen alcohol dehydrogenaza (*ADH1*), và nó bao gồm trình tự từ -1.500 bp đến -1 bp từ codon khởi đầu của ORF của gen *ADH1*.

Trình tự khởi động TEF chứa trong vector pRS-415-pTEF là trình tự khởi động của gen EF-1 alpha (*TEF1*) kéo dài dịch mã, và nó bao gồm trình tự từ -500 bp đến -1 bp từ codon khởi đầu của ORF của gen *TEF1*.

Cụ thể là, gen *TEF1* được phản ứng PCR sử dụng các đoạn mỗi của Bảng 20 bên dưới, và các sản phẩm PCR và các vector pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế các vector pRS-415-pGPD-*TKL1*, pRS-415-pADH-*TKL1*, và pRS-415-pTEF-*TKL1*.

Bảng 20

Sản phẩm PCR	Các mạch khuôn được sử dụng	SEQ ID NO (Mỗi xuôi, mỗi ngược)
<i>TKL1</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 75, SEQ ID NO: 76

Sau đó, plasmit để sinh tổng hợp shinorin được điều chế trong ví dụ 15 được đưa vào chủng *Saccharomyces cerevisiae* CEN.PK-1D cùng với pRS-415-pGPD-*TKL1*,

pRS-415-pADH-*TKL1*, và pRS-415-pTEF-*TKL1*, và từng chủng thu được được cấy trải trên môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, và được nuôi cấy qua đêm ở nhiệt độ 30°C trong tủ ẩm. Một vòng que cấy bạch kim của chủng được nuôi cấy qua đêm trong môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, được cấy chuyển vào 25 mL môi trường chuẩn độ, và sau đó được nuôi cấy ở nhiệt độ 30°C trong tủ ẩm với tốc độ quay 150 vòng/phút trong 24 giờ. Các kết quả như được thể hiện trong Bảng 21 bên dưới.

Bảng 21

Tên chủng	24 Giờ		Nồng độ Shinorin (mg/L)
	OD (600 nm)	Sacarit dư	
CEN.PK-1D WT + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	11,2	0	327
CEN.PK-1D(pADH- <i>TKL1</i>) + pRS413- pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414- pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	13	0	370
CEN.PK-1D(pTEF- <i>TKL1</i>) + pRS413- pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414- pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,1	0	420
CEN.PK-1D(pGPD- <i>TKL1</i>) + pRS413- pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414- pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,5	0	610

Như thể hiện trên bảng 21 ở trên, đã xác nhận rằng lượng sản xuất shinorin đã tăng lên trong chủng, trong đó gen biểu hiện *TKL1* được tăng cường bằng cách sử dụng trình tự khởi động GPD, so với chủng WT. Ngoài ra, cũng xác nhận rằng khi cường độ trình tự khởi động tăng lên (tức là, pGPD > pTEF > pADH), lượng sản xuất shinorin đã tăng lên.

Ví dụ 17: Tăng lượng sản xuất shinorin thông qua tăng cường *ARO4* (3-đeoxy-D-arabino-heptulosonat-7-phosphat (DAHP) synthaza) của *Saccharomyces cerevisiae*

Để tăng khả năng sản sinh MAAs, các chủng *Saccharomyces cerevisiae* trong đó hoạt tính của *ARO4* được tăng cường đã được chuẩn bị. Để đạt được mục đích này, cách được sử dụng trong đó sự biểu hiện *ARO4* gen được tăng cường bằng cách tách dòng gen *ARO4* (được biểu thị bởi SEQ ID NO: 111 và SEQ ID NO: 124) vào các vectơ pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF. Cụ thể là, gen *ARO4* được phản ứng

PCR sử dụng các đoạn môi của Bảng 22 bên dưới, và các sản phẩm PCR của các vector *ARO4* và pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế các vector pRS-415-pGPD-*ARO4*, pRS-415-pADH-*ARO4*, và pRS-415-pTEF-*ARO4*.

Bảng 22

Sản phẩm PCR	Các mạch khuôn được sử dụng	SEQ ID NO (Môi xuôi, môi ngược)
<i>ARO4</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 77, SEQ ID NO: 78

Sau đó, plasmit để sinh tổng hợp shinorin được điều chế trong ví dụ 15 được đưa vào chủng *Saccharomyces cerevisiae* CEN.PK-1D cùng với pRS-415-pGPD-*ARO4*, pRS-415-pADH-*ARO4*, và pRS-415-pTEF-*ARO4*, và từng chủng thu được được cấy trải trên môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, và được nuôi cấy qua đêm ở nhiệt độ 30°C trong tủ ẩm. Một vòng que cấy bạch kim của chủng được nuôi cấy qua đêm trong môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, được cấy chuyển vào 25 mL môi trường chuẩn độ, và sau đó được nuôi cấy ở nhiệt độ 30°C trong tủ ẩm ở tốc độ quay 150 vòng/phút trong 24 giờ. Các kết quả như được thể hiện trong Bảng 23 bên dưới.

Bảng 23

Tên chủng	24 Giờ		Nồng độ Shinorin (mg/L)
	OD (600nm)	Sacarit dư	
CEN.PK-1D WT + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	11,2	0	310
CEN.PK-1D(pADH- <i>ARO4</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	13	0	415
CEN.PK-1D(pTEF- <i>ARO4</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,1	0	610
CEN.PK-1D(pGPD- <i>ARO4</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,5	0	890

Như thể hiện trên bảng 23 ở trên, đã xác nhận rằng lượng sản xuất shinorin đã

tăng lên đến 187% trong chủng, trong đó gen *ARO4* biểu hiện được tăng cường bằng cách sử dụng trình tự khởi động GPD, so với chủng WT.

Ví dụ 18: Tăng lượng sản xuất shinorin thông qua tăng cường phosphoenolpyruvat syntetaza (*pps*) của *Saccharomyces cerevisiae*

Để tăng khả năng sản sinh MAAs, các chủng *Saccharomyces cerevisiae* trong đó hoạt tính của *pps* được tăng cường đã được chuẩn bị. Để đạt được mục đích này bằng cách sử dụng trong đó sự biểu hiện gen *pps* được tăng cường bằng cách tách dòng gen *pps* thành các vector pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF và sự biểu hiện gen *pps* được tăng cường.

Cụ thể là, gen *pps* được phản ứng PCR bằng cách sử dụng các đoạn mồi của Bảng 24 bên dưới, và các sản phẩm PCR của *pps* và các vector pRS-415-pGPD, pRS-415-pADH, và pRS-415-pTEF được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế các vector pRS-415-pGPD-*pps*, pRS-415-pADH-*pps*, và pRS-415-pTEF-*pps*.

Bảng 24

Sản phẩm PCR	Các mạch khuôn được sử dụng	SEQ ID NO (Mồi xuôi, mồi ngược)
<i>pps</i>	gADN <i>E.coli</i> MG1655	SEQ ID NO: 79, SEQ ID NO: 80

Sau đó, plasmid để sinh tổng hợp shinorin được điều chế trong ví dụ 15 được đưa vào chủng *Saccharomyces cerevisiae* CEN.PK-1D cùng với pRS-415-pGPD-*pps*, pRS-415-pADH-*pps*, và pRS-415-pTEF-*pps*, và từng chủng thu được được cấy trải trên môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, và được nuôi cấy qua đêm ở nhiệt độ 30°C trong tủ ấm. Một vòng que cấy bạch kim của chủng được nuôi cấy qua đêm trong môi trường rắn tổng hợp hoàn thiện (SC), trong đó Trp, Ura, và His được loại bỏ, được cấy chuyển vào 25 mL môi trường chuẩn độ, và sau đó được nuôi cấy ở nhiệt độ 30°C trong tủ ấm với tốc độ quay 150 vòng/phút trong 24 giờ. Các kết quả như được thể hiện trong Bảng 25 bên dưới.

Bảng 25

Tên chủng	24 Giờ		Nồng độ Shinorin (mg/L)
	OD (600nm)	Sacarit dư	

CEN.PK-1D WT + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	11,2	0	340
CEN.PK-1D(pADH- <i>pps</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	13	0	375
CEN.PK-1D(pTEF- <i>pps</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,1	0	410
CEN.PK-1D(pGPD- <i>pps</i>) + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	12,5	0	580

Như thể hiện trên bảng 25 ở trên, đã xác nhận rằng lượng sản xuất shinorin đã tăng lên 70% trong chủng, trong đó gen *pps* được biểu hiện quá mức, so với chủng WT. Ngoài ra, cũng đã xác nhận rằng cường độ của trình tự khởi động tăng lên (tức là, pGPD > pTEF > pADH), lượng sản xuất shinorin tăng lên.

Ví dụ 19: Tăng lượng sản xuất shinorin thông qua tăng cường *TKL1*, tăng cường *ARO4*, và sự đưa vào gen *pps* trong chủng *Saccharomyces cerevisiae*

Dựa trên các kết quả của các Ví dụ 16, 17, và 18, các gen *TKL1*, *ARO4*, và *pps* (*E.coli*) được lựa chọn làm các nhân tố hiệu quả có ảnh hưởng đến sinh tổng hợp shinorin ở *Saccharomyces cerevisiae*, và nỗ lực được thực hiện để tăng sinh tổng hợp shinorin thông qua tăng cường đồng thời cả ba loại gen. Để đưa vào ba loại gen, các vectơ pRS-415-pGPD-*TKL1*-pGPD-*ARO4*, và pRS-416-pGPD-*pps* được điều chế. Cụ thể là, sau khi các vùng pGPD-*TKL1* và pGPD-*ARO4* được kết nối bằng PCR chồng lấp, các vectơ và các sản phẩm PCRs được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế vectơ pRS-415-pGPD-*TKL1*-pGPD-*ARO4*.

Sau đó, gen *pps* có nguồn gốc từ *E.coli* được phản ứng PCR. Các sản phẩm PCR của gen *pps* và vectơ pRS-416-pGPD được xử lý bằng các enzym giới hạn *Bam*HI và *Sal*I, và sau đó được gắn bằng cách sử dụng T4 ligaza để điều chế vectơ pRS-416-pGPD-*pps*. Các đoạn mồi và các mạch khuôn ADN được sử dụng để điều chế các vectơ được thể hiện trong Bảng 26 bên dưới.

Bảng 26

Sản phẩm PCR	Các mạch khuôn được sử dụng	SEQ ID NO (Mỗi xuôi, mỗi ngược)
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 81, SEQ ID NO: 82
<i>TKL1</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 83, SEQ ID NO: 84
pGPD- <i>TKL1</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 81, SEQ ID NO: 84
pGPD	gADN <i>S. cerevisiae</i>	SEQ ID NO: 85, SEQ ID NO: 86
<i>ARO4</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 87, SEQ ID NO: 88
pGPD- <i>ARO4</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 85, SEQ ID NO: 88
pGPD- <i>TKL1</i> -pGPD- <i>ARO4</i>	gADN <i>S. cerevisiae</i>	SEQ ID NO: 81, SEQ ID NO: 88
<i>ppsA</i>	gADN <i>E.coli</i> MG1655	SEQ ID NO: 79, SEQ ID NO: 80

Sau đó, plasmit để sinh tổng hợp shinorin được điều chế trong ví dụ 15 được đưa vào chủng *Saccharomyces cerevisiae* CEN.PK-1D cùng với các vector pRS-415-pGPD-*TKL1*-pGPD-*ARO4* và pRS-416-pGPD-*pps*, và từng chủng thu được được cấy trải trên môi trường rắn tổng hợp hoàn thiện (SC), trong đó Leu, Trp, Ura, và His được loại bỏ, và được nuôi cấy qua đêm ở nhiệt độ 30°C trong tủ ẩm. Một vòng que cấy bạch kim của chủng được nuôi cấy qua đêm trong môi trường rắn tổng hợp hoàn thiện (SC), trong đó Leu, Trp, Ura, và His được loại bỏ, được cấy chuyển vào 25 mL môi trường chuẩn độ, và sau đó được nuôi cấy ở nhiệt độ 30°C trong tủ ẩm với tốc độ quay 150 vòng/phút trong 24 giờ. Các kết quả như được thể hiện trong Bảng 27 bên dưới.

Bảng 27

Tên chủng	24 Giờ		Nồng độ Shinorin (mg/L)
	OD (600nm)	Sacarit dư	
CEN.PK-1D WT + pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i>	11,2	0	333
CEN.PK-1D WT+ pRS413-pGPD- <i>Ava_A</i> -pGPD- <i>Ava_B</i> , pRS414-pGPD- <i>Ava_C</i> -pGPD- <i>Ava_D</i> , p415-pGPD- <i>TKL1</i> -pGPD- <i>ARO4</i> , p416-pGPD- <i>pps</i>	13,3	0	1.100

Như thể hiện trên bảng 27 ở trên, đã xác nhận rằng lượng sản xuất shinorin đã tăng lên đáng kể tới 230% ở chủng, trong đó ba loại gen hiệu quả (tức là, *pps*, *TKL1*, và *ARO4*) được biểu hiện quá mức, so với chủng WT.

Trong bản mô tả, các phần mô tả chi tiết về các nội dung có thể được nhận biết và

suy ra đầy đủ bởi người có hiểu biết trung bình trong cùng lĩnh vực kỹ thuật bị lược bỏ. Ngoài các phương án cụ thể được mô tả trong bản mô tả, có thể có nhiều sửa đổi khác nhau mà không làm thay đổi nguyên lý kỹ thuật hoặc các cấu trúc thiết yếu của sáng chế. Do đó, sáng chế có thể được thực hiện theo cách khác từ những phần mô tả cụ thể và được lấy làm ví dụ trong bản mô tả này, có thể được hiểu rõ bởi người có hiểu biết trung bình trong cùng lĩnh vực kỹ thuật.

YÊU CẦU BẢO HỘ

1. Vi sinh vật sản sinh axit amin dạng mycosporin, trong đó hoạt tính của ít nhất một protein được chọn từ nhóm bao gồm 2-dehydro-3-deoxyphosphoheptonat aldolaza, phosphoenolpyruvat syntetaza, và transketolaza được tăng cường.
2. Vi sinh vật theo điểm 1, trong đó vi sinh vật còn bao gồm cụm gen sinh tổng hợp axit amin dạng mycosporin.
3. Vi sinh vật theo điểm 2, trong đó cụm gen sinh tổng hợp axit amin dạng mycosporin bao gồm gen mã hóa ít nhất một protein được chọn từ nhóm bao gồm 2-demetyl 4-deoxygadusol synthaza, O-metyltransferaza, và C-N ligaza.
4. Vi sinh vật theo điểm 2, trong đó cụm gen sinh tổng hợp axit amin dạng mycosporin bao gồm gen mã hóa ít nhất một protein được chọn từ nhóm bao gồm non-ribosomal peptit synthetaza, enzym dạng non-ribosomal peptit synthetaza (enzym dạng NRPS), và D-alanin D-alanin ligaza (D-Ala D-Ala ligaza).
5. Vi sinh vật theo điểm 1, trong đó vi sinh vật là vi sinh vật thuộc chi *Corynebacterium* hoặc vi sinh vật thuộc chi *Escherichia*, hoặc nấm men.
6. Vi sinh vật theo điểm 1, trong đó axit amin dạng mycosporin là ít nhất một loại được chọn từ nhóm bao gồm mycosporin-2-glyxin, palythanol, axit palythenic, deoxygadusol, mycosporin-metylamin-threonin, mycosporin-glyxin-valin, palythin, asterina-330, shinorin, porphyra-334, euhalothece-362, mycosporin-glyxin, mycosporin-ornithin, mycosporin-lyxin, mycosporin-glutamic axit-glyxin, mycosporin-metylamin-serin, mycosporin-taurin, palythen, palythin-serin, palythin-serin-sulfat, palythanol, và usujirene.
7. Phương pháp sản xuất axit amin dạng mycosporin bao gồm:

nuôi cấy vi sinh vật theo điểm bất kỳ từ 1 đến 6 trong môi trường nuôi cấy; và

thu hồi axit amin dạng mycosporin từ vi sinh vật được nuôi cấy hoặc môi trường nuôi cấy.

DANH MỤC TRÌNH TỰ

<110> CJ CheilJedang Corporation

<120> VI SINH VẬT SẢN SINH AXIT AMIN DẠNG MYCOSPORIN VÀ PHƯƠNG PHÁP SẢN XUẤT AXIT AMIN DẠNG MYCOSPORIN BẰNG VI SINH VẬT

<130> OPA18547-PCT

<150> KR10-2018-0022185

<151> 2018-02-23

<160> 124

<170> KoPatentIn 3.0

<210> 1

<211> 1053

<212> ADN

<213> *Escherichia coli* W3110

<400> 1

atgaattatc agaacgacga tttacgcac aaagaaatca aagagttact tcctcctgtc 60

gcattgctgg aaaaattccc cgctactgaa aatgccgcga atacggttgc ccatgcccga 120

aaagcgatcc ataagatcct gaaaggtaat gatgatcgcc tgttggttgt gattggccca 180

tgtcgaattc atgatcctgt cgcggcaaaa gagtatgcca ctgccttgct ggcgctgcgt 240

gaagagctga aagatgagct ggaaatcgta atgcgcgtct attttgaaaa gccgcgtacc 300

acggtgggct ggaaagggct gattaacgat ccgcataatg ataataagctt ccagatcaac 360

gacggtctgc gtatagcccg taaattgctg cttgatatta acgacagcgg tctgccagcg 420

gcaggtgagt ttctcgatat gatcacccca caatatctcg ctgacctgat gagctggggc 480

gcaattggcg cacgtaccac cgaatcgag gtgcaccgcg aactggcatc agggctttct 540

tgtccggtcg gcttcaaaaa tggcaccgac ggtacgatta aagtggctat cgatgccatt 600

aatgccgccg gtgcgccgca ctgcttcctg tccgtaacga aatgggggca ttcggcgatt 660

gtgaatacca gcggtaacgg cgattgccat atcattctgc gcggcggtaa agagcctaac 720

tacagcgcga agcacgttgc tgaagtgaag gaagggtga acaaagcagg cctgccagca 780

caggatgatga tcgatttcag ccatgctaac tcgtccaaac aattcaaaaa gcagatggat 840

gtttgtgctg acgtttgcca gcagattgcc ggtggcgaaa aggccattat tggcgtgatg 900

gtggaaagcc atctgggtga aggcaatcag agcctcgaga gcggggagcc gctggcctac 960

ggtaagagca tcaccgatgc ctgcatcggc tgggaagata ccgatgctct gttacgtcaa 1020

ctggcgcaatg cagtaaaagc gcgtcgcggg taa 1053

<210> 2

<211> 350

<212> PRT

<213> *Escherichia coli* W3110

<400> 2

Met Asn Tyr Gln Asn Asp Asp Leu Arg Ile Lys Glu Ile Lys Glu Leu
 1 5 10 15
 Leu Pro Pro Val Ala Leu Leu Glu Lys Phe Pro Ala Thr Glu Asn Ala
 20 25 30
 Ala Asn Thr Val Ala His Ala Arg Lys Ala Ile His Lys Ile Leu Lys
 35 40 45
 Gly Asn Asp Asp Arg Leu Leu Val Val Ile Gly Pro Cys Ser Ile His
 50 55 60
 Asp Pro Val Ala Ala Lys Glu Tyr Ala Thr Arg Leu Leu Ala Leu Arg
 65 70 75 80
 Glu Glu Leu Lys Asp Glu Leu Glu Ile Val Met Arg Val Tyr Phe Glu
 85 90 95
 Lys Pro Arg Thr Thr Val Gly Trp Lys Gly Leu Ile Asn Asp Pro His
 100 105 110
 Met Asp Asn Ser Phe Gln Ile Asn Asp Gly Leu Arg Ile Ala Arg Lys
 115 120 125
 Leu Leu Leu Asp Ile Asn Asp Ser Gly Leu Pro Ala Ala Gly Glu Phe
 130 135 140
 Leu Asp Met Ile Thr Pro Gln Tyr Leu Ala Asp Leu Met Ser Trp Gly
 145 150 155 160
 Ala Ile Gly Ala Arg Thr Thr Glu Ser Gln Val His Arg Glu Leu Ala
 165 170 175
 Ser Gly Leu Ser Cys Pro Val Gly Phe Lys Asn Gly Thr Asp Gly Thr
 180 185 190
 Ile Lys Val Ala Ile Asp Ala Ile Asn Ala Ala Gly Ala Pro His Cys
 195 200 205
 Phe Leu Ser Val Thr Lys Trp Gly His Ser Ala Ile Val Asn Thr Ser
 210 215 220
 Gly Asn Gly Asp Cys His Ile Ile Leu Arg Gly Gly Lys Glu Pro Asn
 225 230 235 240
 Tyr Ser Ala Lys His Val Ala Glu Val Lys Glu Gly Leu Asn Lys Ala
 245 250 255
 Gly Leu Pro Ala Gln Val Met Ile Asp Phe Ser His Ala Asn Ser Ser
 260 265 270
 Lys Gln Phe Lys Lys Gln Met Asp Val Cys Ala Asp Val Cys Gln Gln
 275 280 285
 Ile Ala Gly Gly Glu Lys Ala Ile Ile Gly Val Met Val Glu Ser His
 290 295 300
 Leu Val Glu Gly Asn Gln Ser Leu Glu Ser Gly Glu Pro Leu Ala Tyr
 305 310 315 320
 Gly Lys Ser Ile Thr Asp Ala Cys Ile Gly Trp Glu Asp Thr Asp Ala
 325 330 335
 Leu Leu Arg Gln Leu Ala Asn Ala Val Lys Ala Arg Arg Gly

	340	345	350
<210>	3		
<211>	49		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	fhuA arm 1_F		
<400>	3		
	ccaagcgacg cccaacctgc catcatggcg cgttcacaaa ctgctcagc		49
<210>	4		
<211>	40		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	fhuA arm 1_R		
<400>	4		
	tatcctcgag actagtgagc tcctttcagc ggttcggtgg		40
<210>	5		
<211>	40		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	fhuA arm 2_F		
<400>	5		
	ctagtctoga ggatatcgac cacaccctgc tgaccggtgt		40
<210>	6		
<211>	48		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	fhuA arm 2_R		
<400>	6		
	atagaataacc aattggcatg ctttttagaa acggaaggtt gcggttgc		48
<210>	7		
<211>	34		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pn_aroG_F		
<400>	7		

aaaggagctc actagtagga tgctcctggt atgg

34

<210> 8
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pn_aroG_R

<400> 8
 cagggtgtgg tcgatatctt acccgcgacg cgcttt

36

<210> 9
 <211> 34
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ptrc_F

<400> 9
 aaaggagctc actagtccgc ttgctgcaac tctc

34

<210> 10
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ptrc_R

<400> 10
 tggacgatac tcatatgttt cctgtgtgaa attgttatcc

40

<210> 11
 <211> 35
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcj1_F

<400> 11
 aaaggagctc actagtaccg cgggcttatt ccatt

35

<210> 12
 <211> 41
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcj1_R

<400>	12		
		atctatagtc tgcatatgtt aatctcctag attgggtttc a	41
<210>	13		
<211>	32		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	aroG_F		
<400>	13		
		taggagatta acatatgaat tatcagaacg ac	32
<210>	14		
<211>	18		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	certification primer		
<400>	14		
		ggaacgctca gattgcgt	18
<210>	15		
<211>	45		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ava_ABCD_F		
<400>	15		
		acaatttcac acaggaaaga tatcatgagt atcgtccaag caaag	45
<210>	16		
<211>	45		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ava_ABCD_R		
<400>	16		
		ctcatccgcc aaaacagctc tagattatga attattttcc agaca	45
<210>	17		
<211>	6461		
<212>	ADN		
<213>	Anabaena variabilis ATCC29413		
<400>	17		
		atgagtatcg tccaagcaaa gtttgaagct aaggaaacat cttttcatgt agaaggttac	60

gaaaagattg agtatgattt ggtgtatgta gatggtattt ttgaaatcca gaattctgca	120
ctagcagatg tatatcaagg ttttggacga tgcttggcga ttgtagatgc taacgtcagt	180
cggttgtatg gtaatcaaatt tcaggcatat ttccagtatt atggtataga actgaggcta	240
tttcctatta ccattactga accagataag actattcaaa ctttcgagag agttatagat	300
gtctttgcag atttcaaatt agtccgcaaa gaaccagtat tagtcgtggg tggcggttta	360
attacagatg ttgtcggctt tgcttgttct acatatcgtc gcagcagcaa ttacatccgc	420
attcctacta cattgattgg attaatgat gccagtgtag caattaaggt agcagttaat	480
catcgcaaac tgaaaaaccg tttgggtgct tatcatgctt ctcgcaaagt attttttagat	540
ttctccttgt tgcgtaactc ccctacagac caagtacgta acgggatggc ggaattggta	600
aaaatcgctg tagtagcgca tcaagaagtt tttgaattgt tggagaagta cggcgaagaa	660
ttactacgta ctcatTTTTGG caatatagat gcaactccag agattaaaga aatagcccat	720
cgTTTgactt acaaagctat ccataagatg ttggaattgg aagttcccaa cctgcatgag	780
ttagacctag atagggtgat tgcttacggt cacacttggg gtcccacctt ggaacttgcg	840
cctcgtctac ccatgttcca cggacacgcc gttaatgtag atatggcttt ctcggaacg	900
atcgccgccc gtagaggata tattacaatt gcagaacgcg atcgtatttt aggattaatg	960
agtcgcgttg gtctatccct cgaccatccc atgttggata tagatatTTT gtggcgtggt	1020
actgaatcta tcacattaac tcgtgatggg ttgttaagag ctgctatgcc aaaaccatt	1080
ggtgattgtg tcttcgtcaa tgacctgaca agagaagaat tagcagccgc attagctgac	1140
cacaaagaac tttgtaccag ttatccccgt ggtgggtgaag gtgtggatgt gtatccccgt	1200
tatcaaaaag aattaatcgg gagtgTTaaa taatgacttt tttgaattca aaatgcaaaa	1260
tactccacgg atacactgcg cgagcgcggt agcatttctg ttcgcggagc gtcccgtagg	1320
gaaagagaag gctacgcaaa taatcggaca ctaattgtct ttaattttga attttgaatt	1380
ttgaattttg aattggagcg aagcgacttg acaaatgtga ttgtccaacc aacagctaga	1440
cctgttacac cattgggaat tttaaccaag cagttagaag ccatagtcca agaggTTaag	1500
caacatccag atttacctgg ggaattgata gcaaacatcc atcaggcttg gcgtttagcc	1560
gcaggtatag acccttattt ggaagaatgc accactccag aatctcctga actcgctgca	1620
ttggcaaaaa ccacagccac cgaagcctgg ggagaacact tccacggagg tacaaccgtc	1680
cgctcctctag aacaagagat gctttctggg catatcgaag gacaaacctt aaagatgttt	1740
gttcacatga ccaaagctaa aaaagtctta gaaattggga tgtttaccgg ttattcggcg	1800
ctggcgatgg cggaagcatt accagaggat ggactgcttg tggcttGTga agttgaccct	1860
tacgcggcgg aaattggaca gaaagccttt caacaatctc cccacgggtg aaagattcgt	1920
gtggaattgg atgcagcctt agcaactctt gataagttag cagaagctgg ggagtctttt	1980

gacttgggtat ttatcgacgc agataaaaaa gagtatgtag cctatttttca caagttgcta	2040
ggtagcagtt tgtagcacc agatggcttt atttgtgtag ataacacctt attacaaggg	2100
gaagtttatc taccagcaga ggaacgtagc gtcaatgggtg aagcgatcgc gcaatttaat	2160
catacagtag ctatagaccc ccgtgtagaa caggttttgt tgccgttgcg agatggttta	2220
acaattatcc gcagaataca accttaattg tccaatcgac tatggcaciaa tcccttcccc	2280
tttcttccgc acctgctaca ccgtctcttc cttcccagac gaaaatagcc gcaattatcc	2340
aaaatatctg cactttggct ttgttattac tagcattgcc cattaatgcc accattgttt	2400
ttatatcctt gttagtcttc cgaccgcaaa aggtcaaagc agcaaacccc caaaccattc	2460
ttatcagtgg cggtaagatg accaaagctt tacaactagc aagggtcattc cacgcggctg	2520
gacatagagt tgtcttgggtg gaaacccata aatactgggt gactgggtcat cgtttttccc	2580
aagcagtgga taagttttac acagtccccg cccccagga caatccccaa gcttacattc	2640
aggctttgggt agatatcgtc aaacaagaaa acatcgatgt ttatatccc gtcaccagtc	2700
cagtgggtag ctactacgac tcattagcca aaccagagtt atcccattat tgcgaaagtgt	2760
ttcactttga cgcagatatt acccaaagtgt tggatgataa atttgcgttg acacaaaaag	2820
cgcgatcgtt tggtttatca gtacccaaat cctttaaaat tacctcacca gaacaagtca	2880
tcaacttcga tttttctgga gagacacgta aatacatcct caaaagcatt ccctacgact	2940
cagtgcggcg gttggactta accaaactcc cctgtgctac tccagaggaa acagcagcat	3000
tcgtcagaag tttgccaatt actcccgaaa aaccgtggat tatgcaggaa tttatccccg	3060
gtaaggaatt ctgcacccat agcaccgttc ggaatgggga actcagactg cattgctgtt	3120
gcgaatcttc agccttccaa gttaattatg agaatgtaaa taaccgcga attaccgaat	3180
gggtacagca ttttgtcaag gaactgaaac tgacaggaca gatttccttt gactttatcc	3240
aagccgaaga cggaacagtt tacgccatcg agtgaaccc ccgcacacat tcagcaatta	3300
ccacatttta cgaccacccc caggtagcag aagcgtactt gagtcaagca ccgacgactg	3360
aaaccataca accactaacg acaagcaagc ctacctattg gacttatcac gaagtttggc	3420
gtttaactgg tatccgttct ttcacccagt tgcaaagatg gctggggaat atttggcgcg	3480
ggactgatgc gatttatcag ccagatgacc cttaccgtt tttgatggta catcattggc	3540
aaattcccct actgttattg aataatttgc gtctcttaa aggttggacg cggatagatt	3600
tcaatattgg gaagttgggtg gaattggggg gagattagtt tttaaacgca gagggacgct	3660
gaggttagcg cagcgaaaag ttctggagga gggtttccct ccgtaggaaa cttttcaaga	3720
gagagggacg cggagtgtgt tttctctgcg tctctgctg agaaattttt tattattgag	3780
caaagttaga agatatgcag actatagatt ttaatatctg taagttactt gtagagtgga	3840
acgcgacca cagagattat gatctttccc agagtttaca tgaactaatt gtagctcaag	3900

tagaacgaac acctgaggcg atcgctgtca cctttgacaa gcaacaacta acttatcaag	3960
aactaaatca taaagcaaac cagctaggac attatattaca aacattagga gtccagccag	4020
aaaccctggt aggcgtttgt ttagaacgtt ccttagaaat ggttatctgt cttttaggaa	4080
tcctcaaagc tgggggtgct tatgttccta ttgaccctga atatcctcaa gaacgcatag	4140
cttatatgct agaagattct caggtgaagg tactactaac tcaagaaaaa ttactcaatc	4200
aaattcccca ccatcaagca caaactatct gtgtagatag ggaatgggag aaaatttcca	4260
cacaagctaa taccaatccc aaaagtaata taaaaacgga taatcttgct tatgtaattt	4320
acacctctgg ttccactggt aaacccaaaag gtgcaatgaa caccacaaaa ggtatctgta	4380
atcgcttatt gtggatgcag gaagcttatt aaatcgattc cacagatagc attttacaaa	4440
aaaccccoctt tagttttgat gtttccgttt gggagttctt ttggacttta ttaactggcg	4500
cacgtttggt aatagccaaa ccaggcggac ataaagatag tgcttacctc atcgatttaa	4560
ttactcaaga acaaatcact acgttgcatt ttgtcccctc aatgctgcaa gtgtttttac	4620
aaaatcgcca tgtaagcaaa tgcagctctc taaaaagagt tatttgtagc ggtgaagctt	4680
tatctataga ttacaaaaat agatttttcc agcatttgca atgtgaatta cataacctct	4740
atggcccgac agaagcagca attgatgtca cattttggca atgtagaaaa gatagtaatt	4800
taaagagtgt acctattggt cgtcccattg ctaatactca aatttatatt cttgatgccg	4860
atttacaacc agtaaataatt ggtgtcactg gtgaaattta tattggtggt gtaggggttg	4920
ctcgtggtta tttgaataaa gaagaattga ccaaagaaaa atttattatt aatccctttc	4980
ccaattctga gtttaagcga ctttataaaa caggtgattt agctcgttat ttacccgatg	5040
gaaatattga atatcttggt agaacagatt atcaagtaaa aattcggggt tatagaattg	5100
aaattggcga gattgaaaat gttttatctt cacaccaca agtcagagaa gctgtagtca	5160
tagcgcggga tgataacgt caagaaaaac aaatcatcgc ttatattacc tataactcca	5220
tcaaacctca gcttgataat ctgcgtgatt tcctaaaagc aaggctacct gattttatga	5280
ttccagccgc ttttgtgatg ctggagcatc ttcttttaac tcccagtggg aaagtagacc	5340
gtaaggcatt acctaagcct gatttattta attatagtga acataattcc tatgtagcgc	5400
ctcggaatga agttgaagaa aaattagtag aaatctgggc gaatattctg catttaccta	5460
aagtaggtgt gacagaaaac tttttcgcta ttggtggtaa ttccctcaaa gctctacatt	5520
taatttctca aattgaagag ttatttgcta aagagatata cttagcaaca cttttaacaa	5580
atccagtaat tgcagattta gccaagggtta ttcaagcaaa caaccaaata cataattcac	5640
ccctagttcc aattcaacca caaggtaagc agcagccttt cttttgtata catcctgctg	5700
gtggtcatgt tttatgctat tttaaactcg cacaatatat aggaactgac caaccatttt	5760
atggccttaca agctcaagga ttttatggag atgaagcacc cttgacgcga gttgaagata	5820

tggctagtct ctacgtcaaa actattagag aatttcaacc ccaagggcct tatcgtgtcg 5880
 gggggtggtc atttggtgga gtcgtagctt atgaagtagc acagcagtta catagacaag 5940
 gacaagaagt atctttacta gcaatattag attcttacgt accgattctg ctggataaac 6000
 aaaaacccat tgatgacgtt tatttagttg gtgttctctc cagagttttt ggcggtatgt 6060
 ttggtcaaga taatctagtc acacctgaag aaatagaaaa tttaactgta gaagaaaaaa 6120
 ttaattacat cattgataaa gcacggagcg ctagaatatt cccgcctggg tagaacgtc 6180
 aaaataatcg ccgatttctt gatgttttgg tgggaacttt aaaagcaact tattcctata 6240
 taagacaacc atatccagga aaagtcactg tatttcgagc cagggaaaaa catattatgg 6300
 ctctgaccc gacctagtt tgggtagaat tattttctgt aatggcggct caagaaatta 6360
 agattattga tgtccctgga aaccattatt cgtttgttct agaaccccat gtacaggttt 6420
 tagcacagcg tttaacaagat tgtctggaaa ataattcata a 6461

<210> 18
 <211> 2379
 <212> ADN
 <213> *Escherichia coli* W3110

<400> 18
 atgtccaaca atggctcgtc accgctgggtg ctttgggtata accaactcgg catgaatgat 60
 gtagacaggg ttggggggcaa aaatgcctcc ctgggtgaaa tgattactaa tctttccgga 120
 atgggtgttt ccgttccgaa tggtttcgcc acaaccgccg acgcgtttaa ccagtttctg 180
 gaccaaagcg gcgtaaacca gcgcatttat gaactgctgg ataaaacgga tattgacgat 240
 gttactcagc ttgcgaaaagc gggcgcgcaa atccgccagt ggattatcga cactcccttc 300
 cagcctgagc tggaaaacgc catccgcgaa gcctatgcac agctttccgc cgatgacgaa 360
 aacgcctctt ttgcggtgcg ctctccgcc accgcagaag atatgccgga cgcttctttt 420
 gccggtcagc aggaaacctt cctcaacgtt cagggttttg acgccgttct cgtggcagtg 480
 aaacatgtat ttgcttctct gttaaacgat cgcgccatct cttatcgtgt gcaccaggg 540
 tacgatcacc gtggtgtggc gctctccgcc ggtgttcaac ggatgggtgcg ctctgacctc 600
 gcatcatctg gcgtgatgtt ctccattgat accgaatccg gctttgacca ggtgggtgtt 660
 atcacttccg catggggcct tggtagatg gtcgtgcagg gtgcggttaa cccggatgag 720
 ttttacgtgc ataaaccgac actggcggcg aatcgcccgg ctatcgtgcg ccgcaccatg 780
 gggtcgaaaa aaatccgcat ggtttacgcg ccgaccagg agcacggcaa gcagggttaa 840
 atcgaagacg taccgcagga acagcgtgac atcttctcgc tgaccaacga agaagtgcag 900
 gaactggcaa aacaggccgt acaaattgag aaacactacg gtcgcccgat ggatattgag 960
 tgggcgaaa atggccacac cggtaaactg ttcatgtgc aggcgcgtcc ggaaaccgtg 1020
 cgctcacgcg gtcaggtcat ggagcgttat acgctgcatt cacagggtaa gattatcgcc 1080

```

gaaggccgtg ctatcgggtca tcgcatcggg gcggggtccgg tgaaagtcac ccatgacatc      1140
agcgaaatga accgcatcga acctggcgac gtgctgggtta ctgacatgac cgacccggac      1200
tggaaccga tcatgaagaa agcatctgcc atcgtcacca accgtggcgg tcgtacctgt      1260
cacgcggcga tcatcgctcg tgaactgggc attccggcgg tagtgggctg tggagatgca      1320
acagaacgga tgaaagacgg tgagaacgtc actgtttctt gtgccgaagg tgataccggt      1380
tacgtctatg cggagttgct ggaatttagc gtgaaaagct ccagcgtaga aacgatgccg      1440
gatctgccgt tgaaagtgat gatgaacgtc ggtaaccggg accgtgcttt cgacttcgcc      1500
tgccctaccga acgaaggcgt gggccttgcg cgtctggaat ttatcatcaa ccgtatgatt      1560
ggcgctccacc cacgcgcact gottgagttt gacgatcagg aaccgcagtt gcaaaacgaa      1620
atccgcgaga tgatgaaagg ttttgattct ccgcgtgaat ttacgttgg tcgtctgact      1680
gaagggatcg cgacgctggg tgccgcgttt tatccgaagc gcgtcattgt ccgtctctct      1740
gattttaaat cgaacgaata tgccaacctg gtcggtgggtg agcgttacga gccagatgaa      1800
gagaaccgga tgctcggctt ccgtggcgcg ggccgctatg tttccgacag cttccgcgac      1860
tgtttcgcgc tggagtgtga agcagtgaag cgtgtgcgca acgacatggg actgaccaac      1920
gttgagatca tgatcccggt cgtgcgtacc gtagatcagg cgaaagcggg gggtgaagaa      1980
ctggcgcgctc aggggctgaa acgtggcgag aacgggctga aaatcatcat gatgtgtgaa      2040
atcccggtcca acgccttgct ggccgagcag ttcctcgaat atttcgacgg cttctcaatt      2100
ggctcaaacg atatgacgca gctggcgctc ggtctggacc gtgactccgg cgtgggtgtct      2160
gaattgttcg atgagcgcaa cgatgcgggtg aaagcactgc tgtcgatggc tatccgtgcc      2220
gcgaagaaac agggcaaata tgtcgggatt tgccggtcagg gtccgtccga ccacgaagac      2280
tttgccgcat gggtgatgga agaggggatc gatagcctgt ctctgaaccc ggacaccgtg      2340
gtgcaaacct ggttaagcct ggctgaactg aagaaataa      2379

```

```

<210>      19
<211>      792
<212>      PRT
<213>      Escherichia coli W3110

```

```

<400>      19
Met Ser Asn Asn Gly Ser Ser Pro Leu Val Leu Trp Tyr Asn Gln Leu
  1              5              10              15

Gly Met Asn Asp Val Asp Arg Val Gly Gly Lys Asn Ala Ser Leu Gly
          20              25              30

Glu Met Ile Thr Asn Leu Ser Gly Met Gly Val Ser Val Pro Asn Gly
          35              40              45

Phe Ala Thr Thr Ala Asp Ala Phe Asn Gln Phe Leu Asp Gln Ser Gly
          50              55              60

```

Val Asn Gln Arg Ile Tyr Glu Leu Leu Asp Lys Thr Asp Ile Asp Asp
 65 70 75 80
 Val Thr Gln Leu Ala Lys Ala Gly Ala Gln Ile Arg Gln Trp Ile Ile
 85 90 95
 Asp Thr Pro Phe Gln Pro Glu Leu Glu Asn Ala Ile Arg Glu Ala Tyr
 100 105 110
 Ala Gln Leu Ser Ala Asp Asp Glu Asn Ala Ser Phe Ala Val Arg Ser
 115 120 125
 Ser Ala Thr Ala Glu Asp Met Pro Asp Ala Ser Phe Ala Gly Gln Gln
 130 135 140
 Glu Thr Phe Leu Asn Val Gln Gly Phe Asp Ala Val Leu Val Ala Val
 145 150 155 160
 Lys His Val Phe Ala Ser Leu Phe Asn Asp Arg Ala Ile Ser Tyr Arg
 165 170 175
 Val His Gln Gly Tyr Asp His Arg Gly Val Ala Leu Ser Ala Gly Val
 180 185 190
 Gln Arg Met Val Arg Ser Asp Leu Ala Ser Ser Gly Val Met Phe Ser
 195 200 205
 Ile Asp Thr Glu Ser Gly Phe Asp Gln Val Val Phe Ile Thr Ser Ala
 210 215 220
 Trp Gly Leu Gly Glu Met Val Val Gln Gly Ala Val Asn Pro Asp Glu
 225 230 235 240
 Phe Tyr Val His Lys Pro Thr Leu Ala Ala Asn Arg Pro Ala Ile Val
 245 250 255
 Arg Arg Thr Met Gly Ser Lys Lys Ile Arg Met Val Tyr Ala Pro Thr
 260 265 270
 Gln Glu His Gly Lys Gln Val Lys Ile Glu Asp Val Pro Gln Glu Gln
 275 280 285
 Arg Asp Ile Phe Ser Leu Thr Asn Glu Glu Val Gln Glu Leu Ala Lys
 290 295 300
 Gln Ala Val Gln Ile Glu Lys His Tyr Gly Arg Pro Met Asp Ile Glu
 305 310 315 320
 Trp Ala Lys Asp Gly His Thr Gly Lys Leu Phe Ile Val Gln Ala Arg
 325 330 335
 Pro Glu Thr Val Arg Ser Arg Gly Gln Val Met Glu Arg Tyr Thr Leu
 340 345 350
 His Ser Gln Gly Lys Ile Ile Ala Glu Gly Arg Ala Ile Gly His Arg
 355 360 365
 Ile Gly Ala Gly Pro Val Lys Val Ile His Asp Ile Ser Glu Met Asn
 370 375 380
 Arg Ile Glu Pro Gly Asp Val Leu Val Thr Asp Met Thr Asp Pro Asp
 385 390 395 400
 Trp Glu Pro Ile Met Lys Lys Ala Ser Ala Ile Val Thr Asn Arg Gly

				405						410					415				
Gly	Arg	Thr	Cys	His	Ala	Ala	Ile	Ile	Ala	Arg	Glu	Leu	Gly	Ile	Pro				
			420					425					430						
Ala	Val	Val	Gly	Cys	Gly	Asp	Ala	Thr	Glu	Arg	Met	Lys	Asp	Gly	Glu				
		435					440					445							
Asn	Val	Thr	Val	Ser	Cys	Ala	Glu	Gly	Asp	Thr	Gly	Tyr	Val	Tyr	Ala				
	450					455					460								
Glu	Leu	Leu	Glu	Phe	Ser	Val	Lys	Ser	Ser	Ser	Val	Glu	Thr	Met	Pro				
465					470					475					480				
Asp	Leu	Pro	Leu	Lys	Val	Met	Met	Asn	Val	Gly	Asn	Pro	Asp	Arg	Ala				
				485					490						495				
Phe	Asp	Phe	Ala	Cys	Leu	Pro	Asn	Glu	Gly	Val	Gly	Leu	Ala	Arg	Leu				
			500					505					510						
Glu	Phe	Ile	Ile	Asn	Arg	Met	Ile	Gly	Val	His	Pro	Arg	Ala	Leu	Leu				
		515					520					525							
Glu	Phe	Asp	Asp	Gln	Glu	Pro	Gln	Leu	Gln	Asn	Glu	Ile	Arg	Glu	Met				
		530				535					540								
Met	Lys	Gly	Phe	Asp	Ser	Pro	Arg	Glu	Phe	Tyr	Val	Gly	Arg	Leu	Thr				
545					550					555					560				
Glu	Gly	Ile	Ala	Thr	Leu	Gly	Ala	Ala	Phe	Tyr	Pro	Lys	Arg	Val	Ile				
				565					570						575				
Val	Arg	Leu	Ser	Asp	Phe	Lys	Ser	Asn	Glu	Tyr	Ala	Asn	Leu	Val	Gly				
			580					585					590						
Gly	Glu	Arg	Tyr	Glu	Pro	Asp	Glu	Glu	Asn	Pro	Met	Leu	Gly	Phe	Arg				
		595					600					605							
Gly	Ala	Gly	Arg	Tyr	Val	Ser	Asp	Ser	Phe	Arg	Asp	Cys	Phe	Ala	Leu				
	610					615					620								
Glu	Cys	Glu	Ala	Val	Lys	Arg	Val	Arg	Asn	Asp	Met	Gly	Leu	Thr	Asn				
625					630					635					640				
Val	Glu	Ile	Met	Ile	Pro	Phe	Val	Arg	Thr	Val	Asp	Gln	Ala	Lys	Ala				
				645					650					655					
Val	Val	Glu	Glu	Leu	Ala	Arg	Gln	Gly	Leu	Lys	Arg	Gly	Glu	Asn	Gly				
			660					665					670						
Leu	Lys	Ile	Ile	Met	Met	Cys	Glu	Ile	Pro	Ser	Asn	Ala	Leu	Leu	Ala				
		675					680					685							
Glu	Gln	Phe	Leu	Glu	Tyr	Phe	Asp	Gly	Phe	Ser	Ile	Gly	Ser	Asn	Asp				
	690					695					700								
Met	Thr	Gln	Leu	Ala	Leu	Gly	Leu	Asp	Arg	Asp	Ser	Gly	Val	Val	Ser				
705					710					715					720				
Glu	Leu	Phe	Asp	Glu	Arg	Asn	Asp	Ala	Val	Lys	Ala	Leu	Leu	Ser	Met				
				725				730						735					
Ala	Ile	Arg	Ala	Ala	Lys	Lys	Gln	Gly	Lys	Tyr	Val	Gly	Ile	Cys	Gly				
			740					745					750						

Gln Gly Pro Ser Asp His Glu Asp Phe Ala Ala Trp Leu Met Glu Glu
755 760 765

Gly Ile Asp Ser Leu Ser Leu Asn Pro Asp Thr Val Val Gln Thr Trp
770 775 780

Leu Ser Leu Ala Glu Leu Lys Lys
785 790

<210> 20
<211> 34
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pn-pps_F

<400> 20
aaaggagctc actagttttg ttcttcccgt gatg

34

<210> 21
<211> 36
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pn-pps_R

<400> 21
caggggtgtgg tcgatatctt atttcttcag ttcagc

36

<210> 22
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> pps_F

<400> 22
taggagatta acatatgtcc aacaatgggt cg

32

<210> 23
<211> 1992
<212> ADN
<213> *Escherichia coli* W3110

<400> 23
atgtcctcac gtaaagagct tgccaatgct attcgtgcgc tgagcatgga cgcagtacag 60
aaagccaaat ccggtcaccc gggtgcccct atgggtatgg ctgacattgc cgaagtcctg 120
tggcgtgatt tcctgaaaca caaccgcag aatccgtcct gggctgaccg tgaccgcttc 180
gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct caccggttac 240

gatctgccga tggaagaact gaaaaacttc cgtcagctgc actctaaaac tccgggtcac 300
ccggaagtgg gttacaccgc tgggtgtgaa accaccaccg gtccgctggg tcaggggtatt 360
gccaacgcag tcggtatggc gattgcagaa aaaacgctgg cggcgcagtt taaccgtccg 420
ggccacgaca ttgtcgacca ctacacctac gccttcatgg gcgacggctg catgatggaa 480
ggcatctccc acgaagtttg ctctctggcg ggtacgctga agctgggtaa actgattgca 540
ttctacgatg acaacggtat ttctatcgat ggtcacgttg aaggctgggt caccgacgac 600
accgcaatgc gtttcgaagc ttacggctgg cacgttattc gcgacatcga cggtcatgac 660
gcggcatcta tcaaacgcgc agtagaagaa gcgcgcgcag tgactgacaa accttccctg 720
ctgatgtgca aaaccatcat cggtttcggt tccccgaaca aagccggtag ccacgactcc 780
cacggtgcgc cgctgggcga cgctgaaatt gccctgacct gcgaacaact gggctggaaa 840
tatgcgccgt tcgaaatccc gtctgaaatc tatgctcagt gggatgcgaa agaagcaggc 900
caggcgaaag aatccgcagt gaacgagaaa ttcgctgctt acgcgaaagc ttatccgcag 960
gaagccgctg aatttaccgc ccgtatgaaa ggcgaaatgc cgtctgactt cgacgctaaa 1020
gcgaaagagt tcatcgctaa actgcaggct aatccggcga aaatcgccag ccgtaaagcg 1080
tctcagaatg ctatcgaaagc gttcgggtccg ctggttgccg aattcctcgg cggttctgct 1140
gacctggcgc cgtctaacct gacctgtgg tctggttcta aagcaatcaa cgaagatgct 1200
gcgggtaact acatccacta cgggtgttcgc gagttcggta tgaccgcgat tgctaacggt 1260
atctccctgc acggtggctt cctgcogtac acctccacct tccatgatgtt cgtggaatac 1320
gcacgtaacg ccgtacgtat ggctgcgctg atgaaacagc gtcagggtgat ggtttacacc 1380
cacgactcca tcggtctggg cgaagacggc ccgactcacc agccggttga gcaggctcgt 1440
tctctgcgcg taaccccgaa catgtctaca tggcgtccgt gtgaccaggt tgaatccgcg 1500
gtcgcgtgga aatacgggtg tgagcgtcag gacggcccga ccgactgat cctctcccgt 1560
cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc gcgcggtggt 1620
tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac cggttcagaa 1680
gttgaactgg ctggtgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg 1740
gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta 1800
ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtac 1860
aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcgggtga atctgctccg 1920
gcagagctgc tgtttgaaga gttcggcttc actgttgata acgttggtgc gaaagcaaaa 1980
gaactgctgt aa 1992

<210> 24
<211> 663
<212> PRT

<213> *Escherichia coli* W3110

<400> 24

```

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
 1          5          10          15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
          20          25          30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
          35          40          45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
          50          55          60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
 65          70          75          80

Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
          85          90          95

Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
          100          105          110

Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
          115          120          125

Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
          130          135          140

Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
          145          150          155          160

Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
          165          170          175

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

Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
          195          200          205

Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile
          210          215          220

Lys Arg Ala Val Glu Glu Ala Arg Ala Val Thr Asp Lys Pro Ser Leu
          225          230          235          240

Leu Met Cys Lys Thr Ile Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly
          245          250          255

Thr His Asp Ser His Gly Ala Pro Leu Gly Asp Ala Glu Ile Ala Leu
          260          265          270

Thr Arg Glu Gln Leu Gly Trp Lys Tyr Ala Pro Phe Glu Ile Pro Ser
          275          280          285

Glu Ile Tyr Ala Gln Trp Asp Ala Lys Glu Ala Gly Gln Ala Lys Glu
          290          295          300

Ser Ala Trp Asn Glu Lys Phe Ala Ala Tyr Ala Lys Ala Tyr Pro Gln
          305          310          315          320

Glu Ala Ala Glu Phe Thr Arg Arg Met Lys Gly Glu Met Pro Ser Asp

```

				325						330					335	
Phe	Asp	Ala	Lys	Ala	Lys	Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	
			340					345					350			
Ala	Lys	Ile	Ala	Ser	Arg	Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	
		355					360					365				
Gly	Pro	Leu	Leu	Pro	Glu	Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	
	370					375					380					
Ser	Asn	Leu	Thr	Leu	Trp	Ser	Gly	Ser	Lys	Ala	Ile	Asn	Glu	Asp	Ala	
385					390					395					400	
Ala	Gly	Asn	Tyr	Ile	His	Tyr	Gly	Val	Arg	Glu	Phe	Gly	Met	Thr	Ala	
				405					410					415		
Ile	Ala	Asn	Gly	Ile	Ser	Leu	His	Gly	Gly	Phe	Leu	Pro	Tyr	Thr	Ser	
			420					425					430			
Thr	Phe	Leu	Met	Phe	Val	Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala	
		435					440					445				
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile	
	450					455					460					
Gly	Leu	Gly	Glu	Asp	Gly	Pro	Thr	His	Gln	Pro	Val	Glu	Gln	Val	Ala	
465					470					475					480	
Ser	Leu	Arg	Val	Thr	Pro	Asn	Met	Ser	Thr	Trp	Arg	Pro	Cys	Asp	Gln	
				485					490					495		
Val	Glu	Ser	Ala	Val	Ala	Trp	Lys	Tyr	Gly	Val	Glu	Arg	Gln	Asp	Gly	
			500					505					510			
Pro	Thr	Ala	Leu	Ile	Leu	Ser	Arg	Gln	Asn	Leu	Ala	Gln	Gln	Glu	Arg	
		515					520					525				
Thr	Glu	Glu	Gln	Leu	Ala	Asn	Ile	Ala	Arg	Gly	Gly	Tyr	Val	Leu	Lys	
	530					535					540					
Asp	Cys	Ala	Gly	Gln	Pro	Glu	Leu	Ile	Phe	Ile	Ala	Thr	Gly	Ser	Glu	
545					550				555						560	
Val	Glu	Leu	Ala	Val	Ala	Ala	Tyr	Glu	Lys	Leu	Thr	Ala	Glu	Gly	Val	
				565					570					575		
Lys	Ala	Arg	Val	Val	Ser	Met	Pro	Ser	Thr	Asp	Ala	Phe	Asp	Lys	Gln	
			580					585					590			
Asp	Ala	Ala	Tyr	Arg	Glu	Ser	Val	Leu	Pro	Lys	Ala	Val	Thr	Ala	Arg	
	595						600					605				
Val	Ala	Val	Glu	Ala	Gly	Ile	Ala	Asp	Tyr	Trp	Tyr	Lys	Tyr	Val	Gly	
	610					615					620					
Leu	Asn	Gly	Ala	Ile	Val	Gly	Met	Thr	Thr	Phe	Gly	Glu	Ser	Ala	Pro	
625					630					635					640	
Ala	Glu	Leu	Leu	Phe	Glu	Glu	Phe	Gly	Phe	Thr	Val	Asp	Asn	Val	Val	
				645					650					655		
Ala	Lys	Ala	Lys	Glu	Leu	Leu										
				660												

<210> 25
 <211> 34
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pn-ktA_F

<400> 25
 aaaggagctc actagtgttt tctcttccgc acaa

34

<210> 26
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pn-ktA_R

<400> 26
 cagggtgtgg tcgatatctt acagcagttc ttttgc

36

<210> 27
 <211> 32
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> tktA_F

<400> 27
 taggagatta acatatgtcc tcacgtaaag ag

32

<210> 28
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcjl-aroG_R

<400> 28
 cccgcggttt acccgcgacg cgcttt

26

<210> 29
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcjl-pps_F

<400> 29

gcgggtaaac cgcgggctta ttccat

26

<210> 30
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcjl-pps_R

<400> 30
 cccgcggttt atttcttcag ttcagc

26

<210> 31
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pcjl-ktA_F

<400> 31
 agaaataaac cgcgggctta ttccat

26

<210> 32
 <211> 71
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> aroD deletion cassette_F

<400> 32
 tcatgggggtt cgggtgcctga caggctgacc gcgtgcagaa agggtaaaaa gctggagctg
 cttcgaagtt c

60

71

<210> 33
 <211> 79
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> aroD deletion cassette_R

<400> 33
 atatattttt tagttcggcg gggaggggtgt tcccgccgaa atattattgc gccatgggcc
 atatgaatat cctccttag

60

79

<210> 34
 <211> 23
 <212> ADN
 <213> Trình tự nhân tạo

<220>		
<223>	aroD deletion_F	
<400>	34	
caaagatttc	cctctggaat atg	23
<210>	35	
<211>	21	
<212>	ADN	
<213>	Trình tự nhân tạo	
<220>		
<223>	aroD deletion_R	
<400>	35	
cagatgtgat	tttccctacg c	21
<210>	36	
<211>	1389	
<212>	ADN	
<213>	<i>Corynebacterium glutamicum</i> ATCC13032	
<400>	36	
gtgagttgga	cagttgatat ccctaaagaa gttctccctg atttgccacc attgccagaa	60
ggcatgcagc	agcagttcga ggacaccatt tcccgtgacg ctaagcagca acctacgtgg	120
gatcgtgcac	aggcagaaaa cgtgcgcaag atccttgagt cggttcctcc aatcgttggt	180
gcccctgagg	tacttgagct gaagcagaag cttgctgatg ttgccaacgg taaggccttc	240
ctcttgacagg	gtggtgactg tgcggaaact ttcgagtcaa aactgagcc gcacattcgc	300
gccaacgtaa	agactctgct gcagatggca gttgttttga cctacggtgc atccactcct	360
gtgatcaaga	tggctcgtat tgctggtcag tacgcaaagc ctcgctcttc tgatctggat	420
ggaaatggtc	tgccaaacta ccgtggcgat atcgtcaacg gtgtggaggc aacccagag	480
gctcgtcgcc	acgatcctgc ccgcatgac cgtgcttacg ctaacgcttc tgctgcgatg	540
aacttggtgc	gcgcgctcac cagctctggc accgctgac tttaccgtct cagcgagtgg	600
aaccgcgagt	tcgttgcgaa ctccccagct ggtgcacgct acgaggctct tgctcgtgag	660
atcgactccg	gtctgcgctt catggaagca tgtggcgtgt ccgatgagtc cctgcgtgct	720
gcagatatct	actgctccca cgaggctttg ctggtggatt acgagcgttc catgctgcgt	780
cttgcaaccg	atgaggaagg caacgaggaa ctttacgac tttcagctca ccagctgtgg	840
atcggcgagc	gcacccgtgg catggatgat ttccatgtga acttcgcac catgatctct	900
aacccaatcg	gcatcaagat tggtcctgg atcaccctg aagaggctgt tgcatacgct	960
gacaagctcg	atccgaactt cgagcctggc cgtttgacca tcgttgctcg catggggccac	1020
gacaaggttc	gctccgtact tcctgggtgt atccaggctg ttgaggcatc cggacacaag	1080
gttatTTTggc	agtcggatcc gatgcacggc aacactttca ccgcatccaa tggctacaag	1140

acccgctcact tcgacaaggt tatcgatgag gtccagggct tcttcgaggt ccaccgcgca 1200
 ttgggcaccc acccagggcg aatccacatt gagttcactg gtgaagatgt caccgagtgc 1260
 ctcggtggcg ctgaagacat caccgatgtt gatctgccag gccgctacga gtccgcatgc 1320
 gatcctcgcc tgaacactca gcagtctttg gagttggctt tcctcgttgc agaaatgctg 1380
 cgtaactaa 1389

<210> 37
 <211> 462
 <212> PRT
 <213> *Corynebacterium glutamicum* ATCC13032

<400> 37
 Met Ser Trp Thr Val Asp Ile Pro Lys Glu Val Leu Pro Asp Leu Pro
 1 5 10 15
 Pro Leu Pro Glu Gly Met Gln Gln Gln Phe Glu Asp Thr Ile Ser Arg
 20 25 30
 Asp Ala Lys Gln Gln Pro Thr Trp Asp Arg Ala Gln Ala Glu Asn Val
 35 40 45
 Arg Lys Ile Leu Glu Ser Val Pro Pro Ile Val Val Ala Pro Glu Val
 50 55 60
 Leu Glu Leu Lys Gln Lys Leu Ala Asp Val Ala Asn Gly Lys Ala Phe
 65 70 75 80
 Leu Leu Gln Gly Gly Asp Cys Ala Glu Thr Phe Glu Ser Asn Thr Glu
 85 90 95
 Pro His Ile Arg Ala Asn Val Lys Thr Leu Leu Gln Met Ala Val Val
 100 105 110
 Leu Thr Tyr Gly Ala Ser Thr Pro Val Ile Lys Met Ala Arg Ile Ala
 115 120 125
 Gly Gln Tyr Ala Lys Pro Arg Ser Ser Asp Leu Asp Gly Asn Gly Leu
 130 135 140
 Pro Asn Tyr Arg Gly Asp Ile Val Asn Gly Val Glu Ala Thr Pro Glu
 145 150 155 160
 Ala Arg Arg His Asp Pro Ala Arg Met Ile Arg Ala Tyr Ala Asn Ala
 165 170 175
 Ser Ala Ala Met Asn Leu Val Arg Ala Leu Thr Ser Ser Gly Thr Ala
 180 185 190
 Asp Leu Tyr Arg Leu Ser Glu Trp Asn Arg Glu Phe Val Ala Asn Ser
 195 200 205
 Pro Ala Gly Ala Arg Tyr Glu Ala Leu Ala Arg Glu Ile Asp Ser Gly
 210 215 220
 Leu Arg Phe Met Glu Ala Cys Gly Val Ser Asp Glu Ser Leu Arg Ala
 225 230 235 240
 Ala Asp Ile Tyr Cys Ser His Glu Ala Leu Leu Val Asp Tyr Glu Arg

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

<210> 38
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pn-cgl aroG_F

<400> 38
 ggtatcgata agcttgattc gatctgctcc cattcc

36

<210> 39
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Pn-cgl aroG_R

<400>	39		
		cggtggcggc cgctctagtt agttacgcag catttc	36
<210>	40		
<211>	39		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-cgl aroG_F		
<400>	40		
		cgaaaggaaa cactcgatgt gagttggaca gttgatatc	39
<210>	41		
<211>	37		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-cgl aroG_R		
<400>	41		
		ccgccaacac agctctagtt tagttacgca gcatttc	37
<210>	42		
<211>	60		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ava_ABCD_F		
<400>	42		
		ccaaacacca acaaaaggct ctacccatat gatgagtatc gtccaagcaa agtttgaagc	60
<210>	43		
<211>	60		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ava_ABCD_R		
<400>	43		
		agagggtgac gctagtcaga actagtttat gaattatattt ccagacaatc ttgtaaacgc	60
<210>	44		
<211>	54		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	O2 promoter_F		

<400>	44		
		ccaaacacca acaaaaggct ctacccatat gcaataatcg tgaattttgg cagc	54
<210>	45		
<211>	53		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	02 promoter_R		
<400>	45		
		caaactttgc ttggacgata ctcattgttt tgatctcctc caataatcta tgc	53
<210>	46		
<211>	37		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pn-cgl tkt_F		
<400>	46		
		ggtatcgata agcttgataa gattgatcac acctttg	37
<210>	47		
<211>	36		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pn-cgl tkt_R		
<400>	47		
		cgggtggcggc cgctctagtt aaccgttaat ggagtc	36
<210>	48		
<211>	36		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-cgl tkt_F		
<400>	48		
		cgaagggaaa cactcgattt gaccaccttg acgctg	36
<210>	49		
<211>	37		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-cgl tkt_R		

<400>	49		
	ccgccaaaac agctctagtt taaccgttaa tggagtc		37
<210>	50		
<211>	36		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ptrc-cgl pps_F		
<400>	50		
	tcacacagga aagatatcat gaccaacagt ttgaac		36
<210>	51		
<211>	39		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Ptrc-cgl pps_R		
<400>	51		
	ccgccaaaac agctctagtt tacttcgtgc cggtcattg		39
<210>	52		
<211>	36		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-cgl pps_F		
<400>	52		
	cgaaaggaaa cactcgatat gaccaacagt ttgaac		36
<210>	53		
<211>	37		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			
<223>	Pcj7-aroG_F		
<400>	53		
	ttagaaatgg aacctaaaag aaacatccca gcgctac		37
<210>	54		
<211>	36		
<212>	ADN		
<213>	Trình tự nhân tạo		
<220>			

<223> Pcj7-aroG_R

<400> 54
tagcgctggg atgtttcttt agttacgcag catttc

36

<210> 55
<211> 36
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pcj7-tkf_F

<400> 55
gaaatgctgc gtaactaaag aaacatccca gcgcta

36

<210> 56
<211> 37
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pcj7-tkf_R

<400> 56
gtagcgctgg gatgtttctt taaccgttaa tggagtc

37

<210> 57
<211> 37
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pcj7-pps_F

<400> 57
gactccatta acggttaaag aaacatccca gcgctac

37

<210> 58
<211> 39
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Pcj7-pps_R

<400> 58
ggagaatact gtcgttcatt tacttcgtgc cggtcattg

39

<210> 59
<211> 26
<212> ADN
<213> Trình tự nhân tạo

<220>
 <223> pGPD_F

<400> 59
 ggatccacag tttattcctg gcatcc 26

<210> 60
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_R

<400> 60
 ttgcttgga cgatactcat ttgtttgttt atgtgtgttt 40

<210> 61
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ava_A_F

<400> 61
 aaacacacat aaacaaacaa atgagtatcg tccaagcaaa 40

<210> 62
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ava_A_R

<400> 62
 ggatgccagg aataaactgt ttatttaaca ctcccgatta 40

<210> 63
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_F

<400> 63
 taatcgggag tgttaaataa acagtttatt cctggcatcc 40

<210> 64
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_R

<400> 64
 tggacaatca catttgtcaa ttgtttgttt atgtgtgttt 40

<210> 65
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ava_B_F

<400> 65
 aaacacacat aaacaaacaa ttgacaaatg tgattgtcca 40

<210> 66
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Ava_B_R

<400> 66
 gtcgacttaa gggtgtattc tgcgga 26

<210> 67
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_F

<400> 67
 ggatccacag tttattcctg gcatcc 26

<210> 68
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_R

<400> 68
 aaagagattg attgtgccat ttgtttgttt atgtgtgttt 40

<210> 69
 <211> 40
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Ava_C_F

<400> 69

aaacacacat aaacaaacaa atggcacaat caatctcttt

40

<210> 70

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Ava_C_R

<400> 70

ggatgccagg aataaactgt ttagtcgccc cctaattcca

40

<210> 71

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> pGPD_F

<400> 71

tggaattagg gggcgactaa acagtttatt cctggcatcc

40

<210> 72

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> pGPD_R

<400> 72

aggatattaa gtactggcat ttgtttgttt atgtgtgttt

40

<210> 73

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Ava_D_F

<400> 73

aaacacacat aaacaaacaa atgccagttac ttaatatcct

40

<210> 74

<211> 26

<212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Ava_D_R

 <400> 74
 gtcgactcaa ttttgtaaca cctttt

 <210> 75
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> TKL1_F

 <400> 75
 ggatccatga ctcaattcac tgacat

 <210> 76
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> TKL1_R

 <400> 76
 gtcgacttag aaagcttttt tcaaag

 <210> 77
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ARO4_F

 <400> 77
 ggatccatga gtgaatctcc aatggt

 <210> 78
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ARO4_R

 <400> 78
 gtcgacctat ttcttggttaa cttctc

 <210> 79

26

26

26

26

26

<211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pps_F

<400> 79
 ggatccatgt ccaacaatgg ctcgtc

26

<210> 80
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pps_R

<400> 80
 gtcgacttat ttcttcagtt cagcca

26

<210> 81
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_F

<400> 81
 ggatccacag tttattcctg gcatcc

26

<210> 82
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_R

<400> 82
 atgtcagtgga attgagtcatt ttgtttgttt atgtgtgttt

40

<210> 83
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> TKL1_F

<400> 83
 aaacacacat aaacaaacaa atgactcaat tcactgacat

40

<210> 84
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> TKL1_R

<400> 84
 ggatgccagg aataaactgt ttagaaagct tttttcaaag

40

<210> 85
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_F

<400> 85
 ctttgaaaaa agcttttctaa acagtttatt cctggcatcc

40

<210> 86
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> pGPD_R

<400> 86
 aacattggag attcactcat ttgtttgttt atgtgtgttt

40

<210> 87
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ARO4_F

<400> 87
 aaacacacat aaacaaacaa atgagtgaat ctccaatgtt

40

<210> 88
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ARO4_R

<400> 88
 gtcgacctat ttcttggttaa cttctc

26

<210> 89
 <211> 438
 <212> ADN
 <213> *Corynebacterium glutamicum* ATCC13032

<400> 89
 atgcctggaa aaattctcct cctcaacggc ccaaacctga acatgctggg caaacgcgag 60
 cctgacattt acggacacga caccttggaa gacgtcgtcg cgctggcaac cgctgaggct 120
 gcgaagcacg gccttgaggt tgaggcgtcg cagagcaatc acgaaggtga gctaatacgat 180
 gcgctgcaca acgctcgcgg caccacatc ggttgctgta ttaaccccg cggcctgact 240
 cacacttcgg tggcgctttt ggatgctgtg aaggcgtctg agcttcctac cgttgagggtg 300
 cacatttcca atccgcatgc ccgtgaagag ttccgccacc attcttatcat ttccctcgcc 360
 gcggtctccg ttatcgctgg cgctggcatc cagggttacc gtttcgcggt cgatatcctg 420
 gcaaattctca aaaagtag 438

<210> 90
 <211> 145
 <212> PRT
 <213> *Corynebacterium glutamicum* ATCC13032

<400> 90
 Met Pro Gly Lys Ile Leu Leu Leu Asn Gly Pro Asn Leu Asn Met Leu
 1 5 10 15
 Gly Lys Arg Glu Pro Asp Ile Tyr Gly His Asp Thr Leu Glu Asp Val
 20 25 30
 Val Ala Leu Ala Thr Ala Glu Ala Ala Lys His Gly Leu Glu Val Glu
 35 40 45
 Ala Leu Gln Ser Asn His Glu Gly Glu Leu Ile Asp Ala Leu His Asn
 50 55 60
 Ala Arg Gly Thr His Ile Gly Cys Val Ile Asn Pro Gly Gly Leu Thr
 65 70 75 80
 His Thr Ser Val Ala Leu Leu Asp Ala Val Lys Ala Ser Glu Leu Pro
 85 90 95
 Thr Val Glu Val His Ile Ser Asn Pro His Ala Arg Glu Glu Phe Arg
 100 105 110
 His His Ser Tyr Ile Ser Leu Ala Ala Val Ser Val Ile Ala Gly Ala
 115 120 125
 Gly Ile Gln Gly Tyr Arg Phe Ala Val Asp Ile Leu Ala Asn Leu Lys
 130 135 140
 Lys
 145

<210> 91
 <211> 32
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> aroD_F

<400> 91

gatcggatcc cactggacgt ttgggtgaga cc

32

<210> 92

<211> 37

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> aroD_R

<400> 92

gatcggatcc actagtttta ggttccattt ctaattg

37

<210> 93

<211> 29

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> aroD_F

<400> 93

gatcactagt atgaacgaca gtattctcc

29

<210> 94

<211> 30

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> aroD_D

<400> 94

gatcaagctt gttacatcct gacgttggtg

30

<210> 95

<211> 2103

<212> ADN

<213> *Corynebacterium glutamicum* ATCC13032

<400> 95

ttgaccacct tgacgctgtc acctgaactt caggcgctca ctgtacgcaa ttaccctctt

60

gattggtccg atgtggacac caaggctgta gacactgttc gtgtcctcgc tgcagacgct

120

gtagaaaact gtggctccgg ccaccaggc accgcaatga gcctgggtcc ccttgcatat

180

accttgtagc agcgggttat gaacgtatg ccacaggaca ccaactgggc aggccgtgac

240

cgcttcgttc tttcttggtg cactcctct ttgaccagc acatccagc ttacttgggt

300

ggattcggcc ttgagatgga tgacctgaag gctctgcgca cctgggattc cttgacccca	360
ggacaccctg agtaccgcca caccaagggc gttgagatca ccactggccc tcttggccag	420
ggtcttgcac ctgcagttgg tatggccatg gctgctcgtc gtgagcgtgg cctattcgac	480
ccaaccgctg ctgagggcga atccccattc gaccaccaca tctacgtcat tgcttctgat	540
ggtgacctgc aggaaggtgt cacctctgag gcatcctcca tcgctggcac ccagcagctg	600
ggcaacctca tcgtgttctg ggatgacaac cgcattctcca tcgaagacaa cactgagatc	660
gctttcaacg aggacgttgt tgctcgttac aaggcttacg gctggcagac cattgaggtt	720
gaggctggcg aggacgttgc agcaatcgaa gctgcagtgg ctgaggctaa gaaggacacc	780
aagcgaccta ccttcatccg cgttcgcacc atcatcggct tcccagctcc aactatgatg	840
aacaccggtg ctgtgcacgg tgctgctctt ggcgagctg aggttgcagc aaccaagact	900
gagcttggat tcgatcctga ggctcacttc gcgatcgacg atgaggttat cgctcacacc	960
cgctccctcg cagagcgcgc tgcacagaag aaggctgcat ggcaggtcaa gttcgatgag	1020
tgggcagctg ccaaccctga gaacaaggct ctgttcgatc gctgaactc ccgtgagctt	1080
ccagcgggct acgctgacga gctcccaaca tgggatgcag atgagaaggg cgtcgcaact	1140
cgtaaggctt ccgaggtgc acttcaggca ctgggcaaga cccttcctga gctgtggggc	1200
ggttccgctg acctcgcagg ttccaacaac accgtgatca agggctcccc ttccttcggc	1260
cctgagtcca tctccaccga gacctggtct gctgagcctt acggccgtaa cctgcacttc	1320
ggtatccgtg agcacgctat gggatccatc ctcaacggca tttccctcca cggcggcacc	1380
cgcccatagc gcggaacctt cctcatcttc tccgactaca tgcgtcctgc agttcgtctt	1440
gcagctctca tggagaccga cgcttactac gtctggaccc acgactccat cggctctgggc	1500
gaagatggcc caaccaccca gcctgttgaa accttggctg cactgcgcgc catcccaggt	1560
ctgtccgtcc tgcgtcctgc agatgcgaac gagaccgcc aggcttgggc tgcagcactt	1620
gagtacaagg aaggccctaa gggctcttgca ctgaccgcc agaacgttcc tgttctggaa	1680
ggcaccaagg agaaggctgc tgaaggcgtt cgccgcggtg gctacgtcct ggttgagggc	1740
tccaaggaaa cccagatgt gatcctcatg ggctccggct ccgaggttca gcttgagtt	1800
aacgtgcga aggtcttga agctgagggc gttgcagctc gcgttgtttc cgttccttgc	1860
atggatttgt tccaggagca ggacgcagag tacatcgagt ccgttctgcc tgcagctgtg	1920
accgtcgtg tgtctgttga agctggcatc gcaatgcctt ggtaccgctt cttgggcacc	1980
cagggccgtg ctgtctccct tgagcacttc ggtgcttctg cggattacca gaccctgttt	2040
gagaagttcg gcatcaccac cgatgcagtc gtggcagcgg ccaaggactc cattaacggt	2100
taa	2103

<210> 96
 <211> 700
 <212> PRT
 <213> *Corynebacterium glutamicum* ATCC13032

 <400> 96
 Met Thr Thr Leu Thr Leu Ser Pro Glu Leu Gln Ala Leu Thr Val Arg
 1 5 10 15
 Asn Tyr Pro Ser Asp Trp Ser Asp Val Asp Thr Lys Ala Val Asp Thr
 20 25 30
 Val Arg Val Leu Ala Ala Asp Ala Val Glu Asn Cys Gly Ser Gly His
 35 40 45
 Pro Gly Thr Ala Met Ser Leu Ala Pro Leu Ala Tyr Thr Leu Tyr Gln
 50 55 60
 Arg Val Met Asn Val Asp Pro Gln Asp Thr Asn Trp Ala Gly Arg Asp
 65 70 75 80
 Arg Phe Val Leu Ser Cys Gly His Ser Ser Leu Thr Gln Tyr Ile Gln
 85 90 95
 Leu Tyr Leu Gly Gly Phe Gly Leu Glu Met Asp Asp Leu Lys Ala Leu
 100 105 110
 Arg Thr Trp Asp Ser Leu Thr Pro Gly His Pro Glu Tyr Arg His Thr
 115 120 125
 Lys Gly Val Glu Ile Thr Thr Gly Pro Leu Gly Gln Gly Leu Ala Ser
 130 135 140
 Ala Val Gly Met Ala Met Ala Ala Arg Arg Glu Arg Gly Leu Phe Asp
 145 150 155 160
 Pro Thr Ala Ala Glu Gly Glu Ser Pro Phe Asp His His Ile Tyr Val
 165 170 175
 Ile Ala Ser Asp Gly Asp Leu Gln Glu Gly Val Thr Ser Glu Ala Ser
 180 185 190
 Ser Ile Ala Gly Thr Gln Gln Leu Gly Asn Leu Ile Val Phe Trp Asp
 195 200 205
 Asp Asn Arg Ile Ser Ile Glu Asp Asn Thr Glu Ile Ala Phe Asn Glu
 210 215 220
 Asp Val Val Ala Arg Tyr Lys Ala Tyr Gly Trp Gln Thr Ile Glu Val
 225 230 235 240
 Glu Ala Gly Glu Asp Val Ala Ala Ile Glu Ala Ala Val Ala Glu Ala
 245 250 255
 Lys Lys Asp Thr Lys Arg Pro Thr Phe Ile Arg Val Arg Thr Ile Ile
 260 265 270
 Gly Phe Pro Ala Pro Thr Met Met Asn Thr Gly Ala Val His Gly Ala
 275 280 285
 Ala Leu Gly Ala Ala Glu Val Ala Ala Thr Lys Thr Glu Leu Gly Phe
 290 295 300
 Asp Pro Glu Ala His Phe Ala Ile Asp Asp Glu Val Ile Ala His Thr

305		310		315		320
Arg Ser Leu Ala Glu Arg Ala Ala Gln Lys Lys Ala Ala Trp Gln Val	325		330		335	
Lys Phe Asp Glu Trp Ala Ala Ala Asn Pro Glu Asn Lys Ala Leu Phe	340		345		350	
Asp Arg Leu Asn Ser Arg Glu Leu Pro Ala Gly Tyr Ala Asp Glu Leu	355		360		365	
Pro Thr Trp Asp Ala Asp Glu Lys Gly Val Ala Thr Arg Lys Ala Ser	370		375		380	
Glu Ala Ala Leu Gln Ala Leu Gly Lys Thr Leu Pro Glu Leu Trp Gly	385		390		395	400
Gly Ser Ala Asp Leu Ala Gly Ser Asn Asn Thr Val Ile Lys Gly Ser	405		410		415	
Pro Ser Phe Gly Pro Glu Ser Ile Ser Thr Glu Thr Trp Ser Ala Glu	420		425		430	
Pro Tyr Gly Arg Asn Leu His Phe Gly Ile Arg Glu His Ala Met Gly	435		440		445	
Ser Ile Leu Asn Gly Ile Ser Leu His Gly Gly Thr Arg Pro Tyr Gly	450		455		460	
Gly Thr Phe Leu Ile Phe Ser Asp Tyr Met Arg Pro Ala Val Arg Leu	465		470		475	480
Ala Ala Leu Met Glu Thr Asp Ala Tyr Tyr Val Trp Thr His Asp Ser	485		490		495	
Ile Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Thr Leu	500		505		510	
Ala Ala Leu Arg Ala Ile Pro Gly Leu Ser Val Leu Arg Pro Ala Asp	515		520		525	
Ala Asn Glu Thr Ala Gln Ala Trp Ala Ala Ala Leu Glu Tyr Lys Glu	530		535		540	
Gly Pro Lys Gly Leu Ala Leu Thr Arg Gln Asn Val Pro Val Leu Glu	545		550		555	560
Gly Thr Lys Glu Lys Ala Ala Glu Gly Val Arg Arg Gly Gly Tyr Val	565		570		575	
Leu Val Glu Gly Ser Lys Glu Thr Pro Asp Val Ile Leu Met Gly Ser	580		585		590	
Gly Ser Glu Val Gln Leu Ala Val Asn Ala Ala Lys Ala Leu Glu Ala	595		600		605	
Glu Gly Val Ala Ala Arg Val Val Ser Val Pro Cys Met Asp Trp Phe	610		615		620	
Gln Glu Gln Asp Ala Glu Tyr Ile Glu Ser Val Leu Pro Ala Ala Val	625		630		635	640
Thr Ala Arg Val Ser Val Glu Ala Gly Ile Ala Met Pro Trp Tyr Arg	645		650		655	

Phe Leu Gly Thr Gln Gly Arg Ala Val Ser Leu Glu His Phe Gly Ala
660 665 670

Ser Ala Asp Tyr Gln Thr Leu Phe Glu Lys Phe Gly Ile Thr Thr Asp
675 680 685

Ala Val Val Ala Ala Ala Lys Asp Ser Ile Asn Gly
690 695 700

<210> 97
<211> 1095
<212> ADN
<213> *Corynebacterium glutamicum* ATCC13032

<400> 97
atgaccaaca gtttgaacat cccgtttgtc cagcgcttcg atgaaggcct ggatcctggt 60
ctagaagtac tcggtggcaa gggcgcttca ctagtcacca tgacagatgc tggaatgccc 120
gttccacctg gatttgtggt cactactgcc agctttgatg aattcatccg tgaagcaggg 180
gttgctgaac acatcgataa attcctaaac gatctcgatg cagaagatgt taaggaagtg 240
gatcgagttt ctgcgatcat ccgcgatgag ctgtgcagtc ttgacgttcc agagaatgct 300
cgtttcgcag tgcaccaggc ttatcgcgat ctcatggaac gatgcggtgg cgacgtcccg 360
gttgctgtcc ggtcatcggc cactgccgaa gatctgcccg atgcttcctt cgcagggcaa 420
caggacacct atctgtggca agtcggtttg agcgtgtca ctgaacacat ccgtaaatgc 480
tgggcttcgc tgttcacttc ccgtgccatt atctaccgtc tgaaaaacaa catccccaat 540
gagggcctct ccatggcggg agttgttcaa aaaatgggtca actctcgtgt cgcaggcgtg 600
gcaatcacta tgaatccttc caacggcgac cgctcgaaga tcaccatcga ttctcatgg 660
gggtgttggtg aaatggtggt ctgaggtgaa gtgacaccag acaatatctt gctggacaag 720
atcacgctgc aggttgtctc cgaacacatt ggaagcaaac acgctgaact catccccgat 780
gccaccagtg gaagcctcgt ggaaaagccc gttgatgaag aacgcgcaaa ccgccgcagt 840
ctgactgatg aggaaatgct cgctgtggca caaatggcta agcgtgcaga aaaacactac 900
aagtgcccac aagatatcga atgggcgctg gacgctgatc tgccagatgg agaaaacctt 960
ctgttattgc aatcccggcc ggaaactatc cactccaacg gtgtgaagaa ggaaacccca 1020
actccgcagg ctgcaaaaac cataggcacc ttcgatttca gctcaatcac cgtcgcaatg 1080
accggcacga agtaa 1095

<210> 98
<211> 364
<212> PRT
<213> *Corynebacterium glutamicum* ATCC13032

<400> 98
Met Thr Asn Ser Leu Asn Ile Pro Phe Val Gln Arg Phe Asp Glu Gly
1 5 10 15

Leu Asp Pro Val Leu Glu Val Leu Gly Gly Lys Gly Ala Ser Leu Val
 20 25 30
 Thr Met Thr Asp Ala Gly Met Pro Val Pro Pro Gly Phe Val Val Thr
 35 40 45
 Thr Ala Ser Phe Asp Glu Phe Ile Arg Glu Ala Gly Val Ala Glu His
 50 55 60
 Ile Asp Lys Phe Leu Asn Asp Leu Asp Ala Glu Asp Val Lys Glu Val
 65 70 75 80
 Asp Arg Val Ser Ala Ile Ile Arg Asp Glu Leu Cys Ser Leu Asp Val
 85 90 95
 Pro Glu Asn Ala Arg Phe Ala Val His Gln Ala Tyr Arg Asp Leu Met
 100 105 110
 Glu Arg Cys Gly Gly Asp Val Pro Val Ala Val Arg Ser Ser Ala Thr
 115 120 125
 Ala Glu Asp Leu Pro Asp Ala Ser Phe Ala Gly Gln Gln Asp Thr Tyr
 130 135 140
 Leu Trp Gln Val Gly Leu Ser Ala Val Thr Glu His Ile Arg Lys Cys
 145 150 155 160
 Trp Ala Ser Leu Phe Thr Ser Arg Ala Ile Ile Tyr Arg Leu Lys Asn
 165 170 175
 Asn Ile Pro Asn Glu Gly Leu Ser Met Ala Val Val Val Gln Lys Met
 180 185 190
 Val Asn Ser Arg Val Ala Gly Val Ala Ile Thr Met Asn Pro Ser Asn
 195 200 205
 Gly Asp Arg Ser Lys Ile Thr Ile Asp Ser Ser Trp Gly Val Gly Glu
 210 215 220
 Met Val Val Ser Gly Glu Val Thr Pro Asp Asn Ile Leu Leu Asp Lys
 225 230 235 240
 Ile Thr Leu Gln Val Val Ser Glu His Ile Gly Ser Lys His Ala Glu
 245 250 255
 Leu Ile Pro Asp Ala Thr Ser Gly Ser Leu Val Glu Lys Pro Val Asp
 260 265 270
 Glu Glu Arg Ala Asn Arg Arg Ser Leu Thr Asp Glu Glu Met Leu Ala
 275 280 285
 Val Ala Gln Met Ala Lys Arg Ala Glu Lys His Tyr Lys Cys Pro Gln
 290 295 300
 Asp Ile Glu Trp Ala Leu Asp Ala Asp Leu Pro Asp Gly Glu Asn Leu
 305 310 315 320
 Leu Leu Leu Gln Ser Arg Pro Glu Thr Ile His Ser Asn Gly Val Lys
 325 330 335
 Lys Glu Thr Pro Thr Pro Gln Ala Ala Lys Thr Ile Gly Thr Phe Asp
 340 345 350

Phe Ser Ser Ile Thr Val Ala Met Thr Gly Thr Lys
 355 360

<210> 99
 <211> 3926
 <212> ADN
 <213> *Saccharomyces cerevisiae*

<400> 99
 ttggtcatgc ttatgaagct atactaacc cacaagcatt acatggtgaa tgtgtgtcca 60
 ttggtatggt taaagaggcg gaattatccc gttatttcgg tattctctcc cctacccaag 120
 ttgcacgtct atccaagatt ttggttgcct acgggttgcc tgtttcgcct gatgagaaat 180
 ggtttaaaga gctaacctta cataagaaaa caccattgga tatcttattg aagaaaatga 240
 gtattgacaa gaaaaacgag gggtccaaaa agaagggtgt cattttagaa agtattggta 300
 agtgctatgg tgactccgct caatttggtt gcgatgaaga cctgagattt attctaacag 360
 atgaaaccct cgtttacccc ttcaaggaca tccttgctga tcaacagaaa gttgttatcc 420
 ccctggttc taagtccatc tccaatcgtg ctttaattct tgctgccctc ggtgaaggtc 480
 aatgtaaaat caagaactta ttacattctg atgatactaa acatatgtta accgctgttc 540
 atgaattgaa aggtgctacg atatcatggg aagataatgg tgagacggta gtggtggaag 600
 gacatggtgg ttccacattg tcagcttggtg ctgacccctt atatctaggt aatgcaggta 660
 ctgcatctag atttttgact tccttggtg ccttggtcaa ttctacttca agccaaaagt 720
 atatcgtttt aactggtaac gcaagaatgc aacaagacc aattgctcct ttggtcgatt 780
 ctttgctgct taatgggtact aaaattgagt acttgaataa tgaagggttc ctgccaatca 840
 aagtttatac tgattcggta ttcaaagggt gtagaattga attagctgct acagtttctt 900
 ctcagtacgt atcctctatc ttgatgtgtg ccccatacgc tgaagaacct gtaactttgg 960
 ctcttggttg tggttaagcca atctctaaat tgtacgtcga tatgacaata aaaatgatgg 1020
 aaaaattcgg tatcaatggt gaaacttcta ctacagaacc ttacacttat tatattccaa 1080
 agggacatta tattaacca tcagaatacg tcattgaaag tgatgcctca agtgctacat 1140
 accattggc cttcgcgcga atgactggta ctaccgtaac gggtccaaac attggttttg 1200
 agtcgttaca aggtgatgcc agatttgcaa gagatgtctt gaaacctatg ggttgtaaaa 1260
 taactcaaac ggcaacttca actactgttt cgggtcctcc tgtaggtact ttaaagccat 1320
 taaaacatgt tgatatggag ccaatgactg atgcgttctt aactgcatgt gttgttgccg 1380
 ctatttcgca cgacagtgat ccaaattctg caaatacaac caccattgaa ggtattgcaa 1440
 accagcgtgt caaagagtgt aacagaattt tggccatggc tacagagctc gccaaatttg 1500
 gcgtcaaaac tacagaatta ccagatggta ttcaagtcca tggtttaaac tcgataaaag 1560
 atttgaagggt tccttccgac tcttctggac ctgtcgggtg atgcacatat gatgatcatc 1620

gtgtggccat gagtttctcg cttcttgcag gaatggtaaa ttctcaaaat gaacgtgacg	1680
aagttgctaa tcctgtaaga atacttgaaa gacattgtac tggtaaaacc tggcctggct	1740
ggtgggatgt gttacattcc gaactaggtg ccaaattaga tggcgcagaa ccttttagagt	1800
gcacatccaa aaagaactca aagaaaagcg ttgtcattat tggcatgaga gcagctggca	1860
aaactactat aagtaaattg tgcgcatccg ctctgggtta caaattagtt gacctagacg	1920
agctgtttga gcaacagcat aacaatcaaa gtgttaaaca atttgttgtg gagaacgggt	1980
gggagaagtt ccgtgaggaa gaaacaagaa ttttcaagga agttattcaa aattacggcg	2040
atgatggata tgttttctca acaggtggcg gtattgttga aagcgctgag tctagaaaag	2100
ccttaaaaga ttttgccctca tcaggtggat acgttttaca cttacatagg gatattgagg	2160
agacaattgt ctttttacia agtgatcctt caagacctgc ctatgtggaa gaaattcgtg	2220
aagtttgaa cagaaggag ggggtgtata aagaatgctc aaatttctct ttctttgctc	2280
ctcattgctc cgcagaagct gagttccaag ctctaagaag atcgtttagt aagtacattg	2340
caaccattac aggtgtcaga gaaatagaaa ttccaagcgg aagatctgcc tttgtgtgtt	2400
taacctttga tgacttaact gaacaaactg agaatttgac tccaatctgt tatggttgtg	2460
aggctgtaga ggtcagagta gaccatttgg ctaattactc tgctgatttc gtgagtaaac	2520
agttatctat attgcgtaaa gccactgaca gtattcctat ctttttact gtgcgaacca	2580
tgaagcaagg tggcaacttt cctgatgaag agttcaaaac cttgagagag ctatacgata	2640
ttgccttgaa gaatggtgtt gaattccttg acttagaact aactttacct actgatatcc	2700
aatatgaggt tattaacaaa aggggcaaca ccaagatcat tggttcccat catgacttcc	2760
aaggattata ctctgggac gacgctgaat gggaaaacag attcaatcaa gcgttaactc	2820
ttgatgtgga tgttgtaaaa tttgtgggtg cggtgttaa tttcgaagat aatttgagac	2880
tggaacactt tagggatata cacaagaata agcctttaat tgcagttaat atgacttcta	2940
aaggtagcat ttctcgtgtt ttgaataatg ttttaacacc tgtgacatca gatttattgc	3000
ctaactccgc tgcccctggc caattgacag tagcacaaat taacaagatg tatacatcta	3060
tgggaggtat cgagcctaag gaactgtttg ttgttggaag gccaatggc cactctagat	3120
cgccaatttt acataaact ggctatgaaa ttttaggttt acctcacaag ttcgataaat	3180
ttgaaactga atccgcacaa ttggtgaaag aaaaactttt ggacggaaac aagaactttg	3240
gcggtgctgc agtcacaatt cctctgaaat tagatataat gcagtacatg gatgaattga	3300
ctgatgctgc taaagttatt ggtgctgtaa acacagttat accattgggt aacaagaagt	3360
ttaaggggtga taataccgac tggtaggtta tccgtaatgc ctttaattaac aatggcgttc	3420
ccgaatatgt tggtcatacc gctggtttgg ttatcggctg aggtggcact tctagagccg	3480
ccctttacgc cttgcacagt ttaggttgca aaaagatctt cataatcaac aggacaactt	3540

cgaaattgaa gccattaata gagtcacttc catctgaatt caacattatt ggaatagagt 3600
 ccactaaatc tatagaagag attaaggaac acgttggcgt tgctgtcagc tgtgtaccag 3660
 ccgacaaaacc attagatgac gaacttttaa gtaagctgga gagattcctt gtgaaagggtg 3720
 cccatgctgc ttttgtacca accttatttg aagccgcata caaaccaagc gttactcccg 3780
 ttatgacaat ttcacaagac aaatatcaat ggcacgttgt ccctggatca caaatgttag 3840
 tacaccaagg tgtagctcag tttgaaaagt ggacaggatt caagggccct ttcaaggcca 3900
 tttttgatgc cgttacgaaa gagtag 3926

<210> 100
 <211> 2247
 <212> ADN
 <213> Anabaena variabilis ATCC29413

<400> 100
 atgagtatcg tccaagcaaa gtttgaagct aaggaaacat cttttcatgt agaaggttac 60
 gaaaagattg agtatgattt ggtgtatgta gatggtatth ttgaaatcca gaattctgca 120
 ctagcagatg tatatcaagg ttttggacga tgcttggoga ttgtagatgc taacgtcagt 180
 cggttgtatg gtaatcaaat tcaggcatat ttocagtatt atggtataga actgaggcta 240
 tttcctatta ccattactga accagataag actattcaaa ctttcgagag agttatagat 300
 gtctttgcag atttcaaatt agtccgcaaa gaaccagtat tagtcgtggg tggcggttta 360
 attacagatg ttgtcggcct tgcttgttct acatatcgtc gcagcagcaa ttacatccgc 420
 attcctacta cattgattgg attaatgat gccagtgtag caattaagggt agcagttaat 480
 catcgcaaac tgaaaaacog tttgggtgct tatcatgctt ctcgcaaagt attttttagat 540
 ttctccttgt tgcgtactct ccctacagac caagtacgta acgggatggc ggaattggta 600
 aaaatcgctg tagtagcgca tcaagaagtt tttgaattgt tggagaagta cggcgaagaa 660
 ttactacgta ctcatthttg caatatagat gcaactccag agattaaaga aatagcccat 720
 cgthttgactt acaaagctat ccataagatg ttggaattgg aagttcccaa cctgcatgag 780
 ttagacctag ataggggtgat tgcttacggt cacacttggg gtcccacctt ggaacttgcg 840
 cctcgtctac ccattgtcca cggacacgcc gttaatgtag atatggctth ctcggaacg 900
 atcgccgccc gtagaggata tattacaatt gcagaacgcg atcgtattht aggattaatg 960
 agtcgcgttg gtctatccct cgaccatccc atgttgata tagatatht gtggcgtggt 1020
 actgaatcta tcacattaac tcgtgatggt ttgttaagag ctgctatgcc aaaaccatt 1080
 ggtgattgtg tcttcgtcaa tgacctgaca agagaagaat tagcagccgc attagctgac 1140
 cacaaagaac tttgtaccag ttatccccgt ggtggtgaag gtgtggatgt gtatccgtht 1200
 tatcaaaaag aattaatcgg gagtgthaaa taatgactth tttgaattca aaatgcaaaa 1260
 tactccacgg atacactgcg cgagcgcggt agcathtctg ttcgcggagc gtcccgtagg 1320

gaaagagaag gctacgcaaa taatcggaca ctaattgtct ttaattttga attttgaatt 1380
 ttgaattttg aattggagcg aagcgacttg acaaagtga ttgtccaacc aacagctaga 1440
 cctgttacac cattgggaat tttaaccaag cagttagaag ccatagtcca agaggttaag 1500
 caacatccag atttacctgg ggaattgata gcaaacatcc atcaggcttg gcgttttagcc 1560
 gcaggatatag acccttattt ggaagaatgc accactccag aatctcctga actcgtgca 1620
 ttggcaaaaa ccacagccac cgaagcctgg ggagaacact tccacggagg tacaaccgtc 1680
 cgtcctctag aacaagagat gctttctggt catatcgaag gacaaacctt aaagatgttt 1740
 gttcacatga ccaaagctaa aaaagtctta gaaattggga tgtttaccgg ttattcggcg 1800
 ctggcgatgg cggaagcatt accagaggat ggactgcttg tggcttgtga agttgaccct 1860
 tacggggcgg aaattggaca gaaagccttt caacaatctc cccacggtgg aaagattcgt 1920
 gtggaattgg atgcagcctt agcaactctt gataagttag cagaagctgg ggagtctttt 1980
 gacttggtat ttatcgacgc agataaaaaa gagtatgtag cctattttca caagttgcta 2040
 ggtagcagtt tgtagcacc agatggcttt atttgtgtag ataacacctt attacaaggg 2100
 gaagtttata taccagcaga ggaacgtagc gtcaatgggtg aagcgatcgc gcaatttaat 2160
 catacagtag ctatagaccc ccgtgtagaa caggttttgt tgccgttgcg agatggttta 2220
 acaattatcc gcagaataca accttaa 2247

<210> 101
 <211> 3532
 <212> ADN
 <213> Anabaena variabilis ATCC29413

<400> 101
 tccagaattc tgcactagca gatgtatatc aaggtttttg acgatgcttg gcgattgtag 60
 atgctaacgt cagtcgggtg tatggtaatc aaattcaggc atatttccag tattatggta 120
 tagaactgag gctatttcct attaccatta ctgaaccaga taagactatt caaactttcg 180
 agagagttat agatgtcttt gcagatttca aattagtccg caaagaacca gtattagtgcg 240
 tgggtggcgg tttaattaca gatgttgctg gctttgcttg ttctacatat cgtcgcagca 300
 gcaattacat ccgcattcct actacattga ttggattaat tgatgccagt gtagcaatta 360
 aggtagcagt taatcatcgc aaactgaaaa accgtttggg tgcttatcat gcttctcgca 420
 aagtattttt agatttctcc ttgttgcgta ctctccctac agaccaagta cgtaacggga 480
 tggcggaatt ggtaaaaatc gctgtagtag cgcatcaaga agtttttgaa ttgttggaga 540
 agtacggcga agaattacta cgtactcatt ttggcaatat agatgcaact ccagagatta 600
 aagaaatagc ccatcgtttg acttaciaaag ctatccataa gatgttggaa ttggaagtcc 660
 ccaacctgca tgagtttagac ctagataggg tgattgctta cggtcacact tggagtccca 720

ccttggaact tgcgcctcgt ctacccatgt tccacggaca cgccgttaat gtagatatgg	780
ctttctcggc aacgatcgcc gcccgtagag gatataattac aattgcagaa cgcgatcgta	840
ttttaggatt aatgagtcgc gttggtctat ccctcgacca tcccatgttg gatatagata	900
ttttgtggcg tggtagtgaa tctatcacat taactcgtga tggtttgta agagctgcta	960
tgccaaaacc cattggtgat tgtgtcttcg tcaatgacct gacaagagaa gaattagcag	1020
ccgcattagc tgaccacaaa gaactttgta ccagttatcc ccgtggtggg gaagggtgtgg	1080
atgtgtatcc cgtttatcaa aaagaattaa tcgggagtgt taaataatga cttttttgaa	1140
ttcaaaatgc aaaatactcc acggatacac tgcgcgagcg cggtagcatt tctgttcgcg	1200
gagcgtcccg tagggaaaga gaaggctacg caaataatcg gacactaatt gtctttaatt	1260
ttgaattttg aattttgaat tttgaattgg agcgaagcga cttgacaaat gtgattgtcc	1320
aaccaacagc tagacctgtt acaccattgg gaattttaac caagcagtta gaagccatag	1380
tccaagaggt taagcaacat ccagatttac ctggggaatt gatagcaaac atccatcagg	1440
cttggcgttt agccgcaggt atagaccctt atttggaaga atgcaccact ccagaatctc	1500
ctgaactcgc tgcattggca aaaaccacag ccaccgaagc ctggggagaa cacttccacg	1560
gaggtacaac cgtccgtcct ctagaacaag agatgctttc tggtagatgc gaaggacaaa	1620
ccttaaagat gtttgttcac atgaccaaag ctaaaaaagt cttagaaatt gggatgttta	1680
ccggttattc ggcgctggcg atggcggaag cattaccaga ggatggactg cttgtggctt	1740
gtgaagttga cccttacgcg gcggaaattg gacagaaagc ctttcaacaa tctccccacg	1800
gtggaaagat tcgtgtggaa ttggatgcag ccttagcaac tcttgataag ttagcagaag	1860
ctggggagtc ttttgacttg gtatttatcg acgcagataa aaaagagtat gtagcctatt	1920
ttcacaagtt gctaggtagc agtttgttag caccagatgg ctttatattgt gtagataaca	1980
ccttattaca aggggaagtt tatctaccag cagaggaacg tagcgtcaat ggtgaagcga	2040
tgcgcgaatt taatcataca gtagctatag acccccgtgt agaacagggt ttgttgccgt	2100
tgcgagatgg ttttaacaatt atccgcagaa tacaacctta attgtccaat cgactatggc	2160
acaatccctt cccctttctt ccgcacctgc tacaccgtct cttccttccc agacgaaaat	2220
agccgcaatt atccaaaata tctgcacttt ggctttgtta ttactagcat tgcccattaa	2280
tgccaccatt gtttttatat ccttgttagt cttccgaccg caaaagggtca aagcagcaaa	2340
cccccaaacc attcttatca gtggcggtaa gatgaccaa gctttacaac tagcaaggtc	2400
attccacgcg gctggacata gagttgtctt ggtggaaacc cataaatact ggttgactgg	2460
tcatcgtttt tcccaagcag tggataagtt ttacacagtc cccgcacccc aggacaatcc	2520
ccaagcttac attcaggctt tggtagatat cgtcaaacaa gaaaacatcg atgtttatat	2580
tcccgtcacc agtccagtgg gtagctacta cgactcatta gccaaaccag agttatccca	2640

ttattgcgaa gtgtttcact ttgacgcaga tattacccaa atgttgatg ataaatttgc 2700
 gttgacacaa aaagcgcgat cgcttggttt atcagtaccc aaatccttta aaattacctc 2760
 accagaacaa gtcacaaact tcgatttttc tggagagaca cgtaaataca tcctcaaaag 2820
 cattccctac gactcagtgc ggcggttgga cttaaccaa ctcccctgtg ctactccaga 2880
 ggaaacagca gcattcgtca gaagtttgcc aattactccc gaaaaaccgt ggattatgca 2940
 ggaatttatc cccggttaag aattctgcac ccatagcacc gttcggaatg gggaactcag 3000
 actgcattgc tgttgccaat cttcagcctt ccaagttaat tatgagaatg taaataaccc 3060
 gcaaattacc gaatgggtac agcattttgt caaggaactg aaactgacag gacagatttc 3120
 ctttgacttt atccaagccg aagacggaac agtttacgcc atcagagtga acccccgcac 3180
 acattcagca attaccacat ttacgacca ccccaggta gcagaagcgt acttgagtca 3240
 agcaccgacg actgaaacca tacaaccact aacgacaagc aagcctacct attggactta 3300
 tcacgaagtt tggcggttaa ctggtatccg ttctttcacc cagttgcaaa gatggctggg 3360
 gaatatattg cgcgggactg atgcgattta tcagccagat gaccocctac cgtttttgat 3420
 ggtacatcat tggcaaattc ccctactggt attgaataat ttgcgtcgtc ttaaagggtg 3480
 gacgcggata gatttcaata ttgggaagtt ggtggaattg gggggagatt ag 3532

<210> 102
 <211> 1233
 <212> ADN
 <213> *Anabaena variabilis* ATCC29413

<400> 102
 atgagtatcg tccaagcaaa gtttgaagct aaggaaacat cttttcatgt agaaggttac 60
 gaaaagattg agtatgattt ggtgtatgta gatggatttt ttgaaatcca gaattctgca 120
 ctacgagatg tatatcaagg ttttggacga tgcttggcga ttgtagatgc taacgtcagt 180
 cggttgatg gtaatcaaat tcaggcatat ttccagtatt atggtataga actgaggcta 240
 tttcctatta ccattactga accagataag actattcaaa ctttcgagag agttatagat 300
 gtctttgcag atttcaaatt agtccgcaaa gaaccagtat tagtcgtggg tggcggttta 360
 attacagatg ttgtcggctt tgcttgttct acatatcgtc gcagcagcaa ttacatccgc 420
 attcctacta cattgattgg attaattgat gccagtgtag caattaaggt agcagttaat 480
 catcgcaaac tgaaaaaccg tttgggtgct tatcatgctt ctcgcaaagt attttttagat 540
 ttctccttgt tgcgtactct ccctacagac caagtacgta acgggatggc ggaattggta 600
 aaaatcgtg tagtagcgca tcaagaagtt tttgaattgt tggagaagta cggcgaagaa 660
 ttactacgta ctatttttg caatatagat gcaactccag agattaaaga aatagcccat 720
 cgtttgactt acaaagctat ccataagatg ttggaattgg aagttcccaa cctgcatgag 780
 ttagacctag ataggggtgat tgcttacggt cacacttgga gtccacactt ggaacttgcg 840

cctcgtctac ccatgttcca cggacacgcc gttaatgtag atatggcttt ctcggaacg 900
 atcgccgccc gtagaggata tattacaatt gcagaacgcg atcgtatattt aggattaatg 960
 agtcgcgttg gtctatccct cgaccatccc atgttgata tagatatttt gtggcgtggg 1020
 actgaatcta tcacattaac tcgtgatggg ttgttaagag ctgctatgcc aaaaccatt 1080
 ggtgattgtg tcttcgtcaa tgacctgaca agagaagaat tagcagccgc attagctgac 1140
 cacaaagaac tttgtaccag ttatccccgt ggtggtgaag gtgtggatgt gtatcccggt 1200
 tatcaaaaag aattaatcgg gagtgtaaa taa 1233

<210> 103
 <211> 840
 <212> ADN
 <213> *Anabaena variabilis* ATCC29413

<400> 103
 ttgacaaatg tgattgtcca accaacagct agacctgtta caccattggg aattttaacc 60
 aagcagttag aagccatagt ccaagagggt aagcaacatc cagatttacc tggggaattg 120
 atagcaaaca tccatcaggc ttggcggtta gcgcaggta tagaccotta tttggaagaa 180
 tgcaccactc cagaatctcc tgaactcgtc gcattggcaa aaaccacagc caccgaagcc 240
 tggggagaa acttccacgg aggtacaacc gtccgtcctc tagaacaaga gatgctttct 300
 ggtcatatcg aaggacaaac cttaaagatg tttgttcaca tgaccaaagc taaaaaagtc 360
 ttagaaattg ggatgtttac cggttattcg gcgctggcga tggcggaagc attaccagag 420
 gatggactgc ttgtggcttg tgaagttgac ccttacgcgg cggaattgg acagaaagcc 480
 tttcaacaat ctccccacgg tggaaagatt cgtgtggaat tggatgcagc cttagcaact 540
 cttgataagt tagcagaagc tggggagtct tttgacttgg tatttatcga cgcagataaa 600
 aaagagtatg tagcctatct tcacaagttg ctaggtagca gtttgtagc accagatggc 660
 tttatttgtg tagataacac cttattacaa ggggaagttt atctaccagc agaggaacgt 720
 agcgtcaatg gtgaagcgat cgcgcaattt aatcatagc tagctataga ccccggtgta 780
 gaacagggtt tgttgccgtt gcgagatggg ttaacaatta tccgcagaat acaaccttaa 840

<210> 104
 <211> 1377
 <212> ADN
 <213> *Anabaena variabilis* ATCC29413

<400> 104
 atggcacaat cccttccccct ttcttccgca cctgctacac cgtctcttcc tcccagacg 60
 aaaatagccg caattatcca aaatatctgc actttggctt tgttattact agcattgccc 120
 attaatgcca ccattgtttt tatatccttg ttagtcttcc gaccgcaaaa ggtcaaagca 180
 gcaaaccccc aaaccattct tatcagtggc ggtaagatga ccaaagcttt acaactagca 240

aggtcattcc acgcggtctg acatagagtt gtcttggtgg aaaccataa atactgggtg 300
 actgggtcatc gtttttccca agcagtggat aagttttaca cagtccccgc accccaggac 360
 aatccccaaag cttacattca ggcttttgga gatatcgtca aacaagaaaa catcgatggt 420
 tatattccccg tcaccagtcc agtgggtagc tactacgact cattagccaa accagagtta 480
 tcccattatt gcgaagtgtt tcactttgac gcagatatta cccaaatggt ggatgataaa 540
 tttgcgttga cacaaaaagc gcgatcgctt ggtttatcag tacccaaata ctttaaaatt 600
 acctcaccag aacaagtcac caacttcgat ttttctggag agacacgtaa atacatctc 660
 aaaagcattc cctacgactc agtgcggcgg ttggacttaa ccaaactccc ctgtgctact 720
 ccagaggaaa cagcagcatt cgtcagaagt ttgccaatta ctccgaaaa accgtggatt 780
 atgcaggaat ttatccccg taaggaatto tgcaccata gcaccgttcg gaatggggaa 840
 ctcagactgc attgctgttg cgaatcttca gccttccaag ttaattatga gaatgtaaat 900
 aaccgcgaaa ttaccgaatg ggtacagcat tttgtcaagg aactgaaact gacaggacag 960
 atttcctttg actttatcca agccgaagac ggaacagttt acgccatcga gtgtaacccc 1020
 cgcacacatt cagcaattac cacatcttac gaccaccccc aggtagcaga agcgtacttg 1080
 agtcaagcac cgacgactga aaccatacaa cactaacga caagcaagcc tacctattgg 1140
 acttatcacg aagtttggcg tttaactggt atccgttctt tcaccagtt gcaaagatgg 1200
 ctggggaata tttggcgcg gactgatgcg atttatcagc cagatgaccc cttaccgttt 1260
 ttgatggtac atcattggca aattccccta ctgttattga ataatttgcg tcgtcttaaa 1320
 ggttggacgc ggatagattt caatattggg aagttgggtg aattgggggg agattag 1377

<210> 105
 <211> 2358
 <212> ADN
 <213> *Anabaena variabilis* ATCC29413

<400> 105
 gtccctattg accctgaata tcctcaagaa cgcatagctt atatgctaga agattctcag 60
 gtgaagggtac tactaactca agaaaaatta ctcaatcaaa ttccccacca tcaagcacia 120
 actatctgtg tagataggga atgggagaaa atttcacac aagctaatac caatcccaaa 180
 agtaatataa aaacggataa tcttgcttat gtaatttaca cctctgggtc cactggtaaa 240
 ccaaaggtg caatgaacac ccacaaagg atctgtaatc gcttattgtg gatgcaggaa 300
 gcttatcaaa tcgattccac agatagcatt ttacaaaaaa ccccctttag ttttgatggt 360
 tccgtttggg agttcttttg gactttatta actggcgcac gtttggtaat agccaaacca 420
 ggcggacata aagatagtgc ttacctcatc gatttaatta ctcaagaaca aatcactacg 480
 ttgcattttg tcccctcaat gctgcaagtg tttttacaaa atcgccatgt aagcaaatgc 540

agctctctaa aaagagttat ttgtagcggg gaagctttat ctatagattt acaaaataga	600
tttttccagc atttgcaatg tgaattacat aacctctatg gcccgcaga agcagcaatt	660
gatgtcacat tttggcaatg tagaaaagat agtaatttaa agagtgtacc tattgggtcgt	720
cccattgcta atactcaaat ttatatctt gatgccgatt tacaaccagt aaatattggg	780
gtcactgggtg aaatttatat tgggtgggtgta ggggttgctc gtgggttattt gaataaagaa	840
gaattgacca aagaaaaaatt tattattaat ccctttccca attctgagtt taagcgactt	900
tataaaacag gtgatttagc tcgttattta cccgatggaa atattgaata tcttggtaga	960
acagattatc aagtaaaaat tcgggggttat agaattgaaa ttggcgagat tgaaaatggt	1020
ttatcttcac acccacaagt cagagaagct gtagtcatag cgcgggatga taacgctcaa	1080
gaaaaacaaa tcatcgctta tattacctat aactccatca aacctcagct tgataatctg	1140
cgtgatattcc taaaagcaag gctacctgat tttatgattc cagccgcttt tgtgatgctg	1200
gagcatcttc cttaactcc cagtggtaaa gtagaccgta aggcattacc taagcctgat	1260
ttatttaatt atagtgaaca taattcctat gtagcgctc ggaatgaagt tgaagaaaaa	1320
ttagtacaaa tctgggtcgaa tattctgcat ttacctaaag taggtgtgac agaaaacttt	1380
ttcgctattg gtggtaattc cctcaaagct ctacatttaa tttctcaaat tgaagagtta	1440
tttgctaaag agatatcctt agcaacactt ttaacaaatc cagtaattgc agatttagcc	1500
aaggttattc aagcaaacaa ccaaattccat aattcacccc tagttccaat tcaaccacaa	1560
ggtaagcagc agcctttctt ttgtatacat cctgctgggtg gtcattgttt atgctatttt	1620
aaactgcac aatatatagg aactgaccaa ccattttatg gcttacaagc tcaaggattt	1680
tatggagatg aagcaccctt gacgcgagtt gaagatatgg ctagtctcta cgtcaaaact	1740
attagagaat ttcaacccca agggccttat cgtgtcgggg ggtggtcatt tgggtggagtc	1800
gtagcttatg aagtagcaca gcagttacat agacaaggac aagaagtatc ttactagca	1860
atattagatt cttacgtacc gattctgctg gataaacaaa aaccattga tgacgtttat	1920
ttagttgggtg ttctctocag agtttttggc ggtatgtttg gtcaagataa tctagtcaca	1980
cctgaagaaa tagaaaattt aactgtagaa gaaaaaatta attacatcat tgataaagca	2040
cggagcgcta gaatattccc gcctgggtgta gaacgtcaaa ataatcgccg tattcttgat	2100
gttttgggtg gaactttaaa agcaacttat tcctatataa gacaaccata tccaggaaaa	2160
gtcactgtat ttcgagccag ggaaaaacat attatggctc ctgaccgcac cttagtttgg	2220
gtagaattat tttctgtaat ggcggctcaa gaaattaaga ttattgatgt ccctggaaac	2280
cattattcgt ttgttctaga acccatgta caggttttag cacagcgttt acaagattgt	2340
ctggaaaata attcataa	2358

<211> 1233
 <212> ADN
 <213> Nostoc punctiforme ATCC29133

<400> 106
 atgagtaatg ttcaagcatc gtttgaagca acggaagctg aattccgcgt ggaaggttac 60
 gaaaaaattg agtttagtct tgtctatgta aatgggtgcat ttgatatcag taacagagaa 120
 attgcagaca gctatgagaa gtttggccgt tgtctgactg tgattgatgc taatgtcaac 180
 agattatatg gcaagcaaatt caagtcatat tttagacact atgggtattga tctgaccgta 240
 gttcccatg tgattactga gcctactaaa acccttgcaa cctttgagaa aattggtgat 300
 gctttttctg acttttggtt aatccgcaag gaaccagtat tagtagtggg tgggtggttta 360
 accactgatg tagctgggtt tgcgtgtgct gcttaccgtc gtaagagtaa ctatattcgg 420
 gttccgacaa cgctgattgg tctgattgat gcagggtgtag cgattaagggt tgcagtcaac 480
 catcgcaagt taaaaaatcg cttgggtgca tatcatgctc ctttgaaagt catcctcgat 540
 ttctccttct tgcaaacatt accaacggct caagttcgtat atgggatggc agagtgtggtc 600
 aaaattgctg ttgtggcgaa ctctgaagtc tttgaattgt tgtatgaata tggcgaagag 660
 ttgctttcca ctcaacttgg atatgtgaat ggtacaaagg aactgaaagc gatcgccacat 720
 aaactcaatt acgaggcaat aaaaactatg ctggagttgg aaactccaaa cttgcatgag 780
 ttagacctcg atcgcgtcat tgccacgggt cacacttgga gtccgacctt agaattagct 840
 ccgatgatac cgttgtttca cggtcatgcc gtcaatatag acatggcttt gtctgcaact 900
 attgcggaac gacgtggcta cattacatca ggagaacgcg atcgcatctt gagcttgatg 960
 agtcgtatag gtttatcaat cgatcatcct ctactagatg gcgatttgct ctggtatgct 1020
 acccaatcta ttagcttgac gcgagacggg aaacaacgcg cagctatgcc taaaccatt 1080
 ggtgagtgtc tctttgtcaa cgatttcacc cgtgaagaac tagatgcagc tttagctgaa 1140
 cacaaacgtc tgtgtgctac ataccctcgt ggtggagatg gcattgacgc ttacatagaa 1200
 actcaagaag aatccaaact attgggagtg tga 1233

<210> 107
 <211> 834
 <212> ADN
 <213> Nostoc punctiforme ATCC29133

<400> 107
 atgaccagta ttttaggacg agataccgca agaccaataa cgccacatag cattctggta 60
 gcacagctac aaaaaaccct cagaatggca gaggaagta atattccttc agagatactg 120
 acttctctgc gccaaagggtt gcaattagca gcagggttag atccctatct ggatgattgc 180
 actacccttg aatcgaccgc attgacagca ctagcgcaga agacaagcat tgaagactgg 240
 agtaaacgct tcagtgatgg tgaacagtg cgtcaattag agcaagaaat gctctcagga 300

catcttgaag gacaaacact aaagatgttt gtgcatatca ctaaggctaa aagcatccta	360
gaagtgggaa tgttcacagg atattcagct ttggcaatgg cagaggcggt accagatgat	420
gggcgactga ttgcttgtga agtagactcc tatgtggccg agtttgctca aacttgcttt	480
caagagtctc cccacggccg caagattggt gtagaagtgg cacctgcact agagacgctg	540
cacaagctgg tggctaataa agaatccttt gatctgatct tcattgatgc ggataaaaag	600
gagtatatag aatacttcca aattatcttg gatagccatt tactagctcc cgacggatta	660
atctgtgtag ataatacttt gttgcaagga caagtttacc tgccatcaga acagcgtaca	720
gccaatggtg aagcgatcgc tcaattcaac cgcattgtcg ccgcagatcc tcgtgtagag	780
caagttctgt taccatacag agatggtata accctgatta gacgcttgggt ataa	834
<210> 108	
<211> 1086	
<212> ADN	
<213> Nostoc punctiforme ATCC29133	
<400> 108	
tactggttat cagggcatag attctcaaat tctgtgagtc gtttttatac agttcctgca	60
ccacaagacg acccagaagg ctatacccaa gcgctattgg aaattgtcaa acgagagaag	120
attgacgttt atgtaccogt atgcagccct gtagctagtt attacgactc tttggcaaag	180
tctgcactat cagaatattg tgaggttttt cactttgatg ctgatataac caagatgctg	240
gatgataaat ttgcctttac cgatcgggcg cgatcgcttg gtttatcagc cccgaaatct	300
tttaaaatta ccgatcctga acaagttatc aacttcgatt ttagtaaaga gacgcgcaaa	360
tatattctta agagtatttc ttacgactca gttcgccgct taaatttaac caaacttcct	420
tgtgataccc cagaagagac agcagcgttt gtcaagagtt taccatcag cccagaaaaa	480
ccttggatta tgcaagaatt tattcctggg aaagaattat gcacccatag cacagtccga	540
gacggcgaat taaggttgca ttgctgttca aattcttcag cgtttcagat taactatgaa	600
aatgtcgaaa atccccaaat tcaagaatgg gtacaacatt tcgtcaaaaag tttacggctg	660
actggacaaa tatctcttga ctttatccaa gctgaagatg gtacagctta tgccattgaa	720
tgtaatcctc gcacccattc ggcgatcaca atgttctaca atcaccagc tgttgcagaa	780
gcctatcttg gtaaaactcc tctagctgca cctttggaac ctcttcgaga tagcaagccc	840
acttactgga tatatcacga aatctggcga ttgactggga ttcgctcttg acaacaatta	900
caaacttggg ttgggagatt agtcagaggt acagatgcc aattatcgct ggacgatccg	960
ataccatttt taactttgca ccattggcag attactttac ttttgctaca aaatttgcaa	1020
cgactcaaag gttgggtaaa gatcgatttt aatatcggt aactcgtgga attagggggc	1080
gactaa	1086

<210> 109
 <211> 1047
 <212> ADN
 <213> Nostoc punctiforme ATCC29133

<400> 109
 atgccagttac ttaatatcct tcatttagtt gggctctgcac acgataagtt ttactgtgat 60
 ttatcacgtc tttatgccca agactgttta gctgcaacag cagatccatc gctttataac 120
 tttcaaattg catatatcac acccgatcgg cagtggcgat ttcctgactc tctcagtcga 180
 gaagatattg ctcttaccaa accgattcct gtgtttgatg ccatacaatt tctaacaggc 240
 caaaacattg acatgatgtt accacaaatg ttttgtattc ctggaatgac tcagtaccgt 300
 gccctattcg atctgctcaa gatcccttat ataggaaata cccagatat tatggcgatc 360
 gcggcccaca aagccagagc caaagcaatt gtcgaagcag caggggtaaa agtgcctcgt 420
 ggagaattgc ttcgcccaagg agatattcca acaattacac ctccagcagt cgtcaaacct 480
 gtaagttctg acaactcttt aggagtagtc ttagttaaag atgtgactga atatgatgct 540
 gccttaaaga aagcatttga atatgcttcg gaggtcatcg tagaagcatt catcgaactt 600
 ggtcgagaag tcagatgcgg catcattgta aaagacggtg agctaataagg tttacccctt 660
 gaagagtatc tggtagacct acacgataaa cctatccgta actatgctga taaactccaa 720
 caaactgacg atggcgactt gcatttgact gctaaagata atatcaaggc ttggatttta 780
 gaccctaacg acccaatcac ccaaaagggt cagcaagtgg ctaaaagggt tcatcaggct 840
 ttgggttgct gccactacag tttatttgac ttccgaatcg atccaaaggg acaaccttgg 900
 ttcttagaag ctggattata ttgttctttt gcccccaaaa gtgtgatttc ttctatggcg 960
 aaagcagccg gaatccctct aaatgattta ttaataaccg ctattaatga aacattgggt 1020
 agtaataaaa aggtgttaca aaattga 1047

<210> 110
 <211> 1626
 <212> ADN
 <213> *Saccharomyces cerevisiae* CEN.PK-1D

<400> 110
 aacaagccgg gctttacctt gtctgacaac tacacctatg ttttcttggg tgacggttgt 60
 ttgcaagaag gtatttcttc agaagcttcc tccttggctg gtcatttgaa attgggtaac 120
 ttgattgcca tctacgatga caacaagatc actatcgatg gtgctaccag tatctcatc 180
 gatgaagatg ttgctaagag atacgaagcc tacgggttggg aagttttgta cgtagaaaat 240
 ggtaacgaag atctagccgg tattgccaag gctattgctc aagctaagtt atccaaggac 300
 aaaccaactt tgatcaaaaat gaccacaacc attgggttacg gttccttgca tgccggctct 360
 cactctgtgc acggtgcccc attgaaagca gatgatgtta aacaactaaa gagcaaattc 420
 ggtttcaacc cagacaagtc ctttgttggt ccacaagaag tttacgacca ctaccaaag 480

acaatttttaa agccaggtgt cgaagccaac aacaagtgga acaagttggt cagcgaatac 540
 caaaagaaat tcccagaatt aggtgctgaa ttggctagaa gattgagcgg ccaactaccc 600
 gcaaattggg aatctaagtt gccaaactac accgccaagg actctgccgt ggccactaga 660
 aaattatcag aaactgttct tgaggatggt tacaatcaat tgccagagtt gattgggtggt 720
 tctgccgatt taacaccttc taacttgacc agatggaagg aagcccttga cttccaacct 780
 ctttcttccg gttcaggtaa ctactctggt agatacatta ggtacggtat tagagaacac 840
 gctatgggtg ccataatgaa cggtatttca gctttcgggtg ccaactacaa accatacgggt 900
 ggtactttct tgaacttcgt ttcttatgct gctgggtgccg ttagattgtc cgctttgtct 960
 ggccaccag ttatttgggt tgctacacat gactctatcg gtgtcgggtga agatgggtcca 1020
 acacatcaac ctattgaaac tttagcacac ttcagatccc taccaaacat tcaagtttgg 1080
 agaccagctg atggtaacga agtttctgcc gcctacaaga actctttaga atccaagcat 1140
 actccaagta tcattgcttt gtccagacaa aacttgccac aattggaagg tagctctatt 1200
 gaaagcgctt ctaaggggtg ttacgtacta caagatgttg ctaaccaga tattatttta 1260
 gtggctactg gttccgaagt gtctttgagt gttgaagctg ctaagacttt ggccgcaaag 1320
 aacatcaagg ctctgtttgt ttctctacca gatttcttca cttttgacaa acaacccta 1380
 gaatacagac tatcagtctt accagacaac gttccaatca tgtctgttga agttttggct 1440
 accacatggt ggggcaaata cgctcatcaa tccttcggta ttgacagatt tgggtgcctcc 1500
 ggtaaggcac cagaagtctt caagttcttc ggtttcaccc cagaagggtg tgctgaaaga 1560
 gctcaaaaga ccattgcatt ctataagggg gacaagctaa tttctccttt gaaaaaagct 1620
 ttctaa 1626

<210> 111
 <211> 1113
 <212> ADN
 <213> *Saccharomyces cerevisiae* CEN.PK-1D

<400> 111
 atgagtgaat ctccaatggt cgctgccaac ggcatgccaa aggtaaatca aggtgctgaa 60
 gaagatgtca gaattttagg ttacgacca ttagcttctc cagctctcct tcaagtgcaa 120
 atcccagcca caccaacttc tttggaaact gccaaagagag gtagaagaga agctatagat 180
 attattaccg gtaaagacga cagagttctt gtcattgtcg gtccttggtc catccatgat 240
 ctagaagccg ctcaagaata cgctttgaga ttaaagaaat tgtcagatga attaaaagg 300
 gatttatcca tcattatgag agcatacttg gagaagccaa gaacaaccgt cggctggaaa 360
 ggtctaatta atgaccctga tgtaacaac actttcaaca tcaacaaggg tttgcaatcc 420
 gctagacaat tgtttgtaa cttgacaaat atcggtttgc caattggttc tgaaatgctt 480

gataccattt ctcctcaata cttggctgat ttggtctcct tcggtgccat tgggtgccaga 540
 accaccgaat ctcaactgca cagagaattg gcctccggtt tgtctttccc agttgggttc 600
 aagaacggta ccgatgggtac cttaaagtgt gctgtggatg cttgtcaagc cgctgctcat 660
 tctcaccatt tcatgggtgt tactaagcat ggtgttgctg ctatcaccac tactaaggggt 720
 aacgaacact gcttcgttat tctaagaggt ggtaaaaagg gtaccaacta cgacgctaag 780
 tccgttgcag aagctaaggc tcaattgcct gccggttcca acggtctaata gattgactac 840
 tctcacggta actccaataa ggatttcaga aaccaaccaa aggtcaatga cgttggttgt 900
 gagcaaatcg ctaacgggtga aaacgccatt accggtgtca tgattgaatc aaacatcaac 960
 gaaggttaacc aaggcatccc agccgaagggt aaagccggct tgaaatatgg tgtttccatc 1020
 actgatgctt gtataggttg ggaaactact gaagacgtct tgaggaaatt ggctgctgct 1080
 gtcagacaaa gaagagaagt taacaagaaa tag 1113

<210> 112
 <211> 1588
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 112
 Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15
 Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30
 Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45
 Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60
 Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80
 Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95
 Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110
 Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125
 Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140
 Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160
 Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175
 Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

```

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
195                               200                               205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
210                               215                               220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile
225                               230                               235                               240

Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys
245                               250                               255

Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg
260                               265                               270

Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile
275                               280                               285

Leu Thr Pro Gln Ala Leu His Gly Glu Cys Val Ser Ile Gly Met Val
290                               295                               300

Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln
305                               310                               315                               320

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
325                               330                               335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
340                               345                               350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
355                               360                               365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
370                               375                               380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
385                               390                               395                               400

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
405                               410                               415

Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
420                               425                               430

Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
435                               440                               445

His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
450                               455                               460

Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
465                               470                               475                               480

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
485                               490                               495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
500                               505                               510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
515                               520                               525

```

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540
 Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560
 Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575
 Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590
 Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605
 Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620
 Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640
 Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655
 Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
 660 665 670
 Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
 675 680 685
 Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
 690 695 700
 Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
 705 710 715 720
 Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
 725 730 735
 Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
 740 745 750
 Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
 755 760 765
 Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
 770 775 780
 Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
 785 790 795 800
 Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
 805 810 815
 Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830
 Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845
 Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860
 Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu

865		870		875		880
Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met						
		885		890		895
Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu						
		900		905		910
Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn						
		915		920		925
Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe						
		930		935		940
Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly						
		945		950		955
Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala						
		965		970		975
Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val						
		980		985		990
Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser						
		995		1000		1005
Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp Asn						
		1010		1015		1020
Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe Phe Ala						
		1025		1030		1035
Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg Arg Ser Phe						
		1045		1050		1055
Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu Ile Glu Ile Pro						
		1060		1065		1070
Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe Asp Asp Leu Thr Glu						
		1075		1080		1085
Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr Gly Cys Glu Ala Val Glu						
		1090		1095		1100
Val Arg Val Asp His Leu Ala Asn Tyr Ser Ala Asp Phe Val Ser Lys						
		1105		1110		1115
Gln Leu Ser Ile Leu Arg Lys Ala Thr Asp Ser Ile Pro Ile Ile Phe						
		1125		1130		1135
Thr Val Arg Thr Met Lys Gln Gly Gly Asn Phe Pro Asp Glu Glu Phe						
		1140		1145		1150
Lys Thr Leu Arg Glu Leu Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu						
		1155		1160		1165
Phe Leu Asp Leu Glu Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val						
		1170		1175		1180
Ile Asn Lys Arg Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe						
		1185		1190		1195
Gln Gly Leu Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn						
		1205		1210		1215

Gln Ala Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala
 1220 1225 1230
 Val Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245
 Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser Ile
 1250 1255 1260
 Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp Leu Leu
 1265 1270 1275 1280
 Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln Ile Asn Lys
 1285 1290 1295
 Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu Leu Phe Val Val
 1300 1305 1310
 Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile Leu His Asn Thr Gly
 1315 1320 1325
 Tyr Glu Ile Leu Gly Leu Pro His Lys Phe Asp Lys Phe Glu Thr Glu
 1330 1335 1340
 Ser Ala Gln Leu Val Lys Glu Lys Leu Leu Asp Gly Asn Lys Asn Phe
 1345 1350 1355 1360
 Gly Gly Ala Ala Val Thr Ile Pro Leu Lys Leu Asp Ile Met Gln Tyr
 1365 1370 1375
 Met Asp Glu Leu Thr Asp Ala Ala Lys Val Ile Gly Ala Val Asn Thr
 1380 1385 1390
 Val Ile Pro Leu Gly Asn Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp
 1395 1400 1405
 Leu Gly Ile Arg Asn Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val
 1410 1415 1420
 Gly His Thr Ala Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala
 1425 1430 1435 1440
 Ala Leu Tyr Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile
 1445 1450 1455
 Asn Arg Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser
 1460 1465 1470
 Glu Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485
 Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys Pro
 1490 1495 1500
 Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val Lys Gly
 1505 1510 1515 1520
 Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala Tyr Lys Pro
 1525 1530 1535
 Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys Tyr Gln Trp His
 1540 1545 1550

Val Val Pro Gly Ser Gln Met Leu Val His Gln Gly Val Ala Gln Phe
 1555 1560 1565

Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe Lys Ala Ile Phe Asp Ala
 1570 1575 1580

Val Thr Lys Glu
 1585

<210> 113
 <211> 820
 <212> PRT
 <213> *Anabaena variabilis* ATCC29413

<400> 113
 Met Ser Ile Val Gln Ala Lys Phe Glu Ala Lys Glu Thr Ser Phe His
 1 5 10 15

Val Glu Gly Tyr Glu Lys Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly
 20 25 30

Ile Phe Glu Ile Gln Asn Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe
 35 40 45

Gly Arg Cys Leu Ala Ile Val Asp Ala Asn Val Ser Arg Leu Tyr Gly
 50 55 60

Asn Gln Ile Gln Ala Tyr Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu
 65 70 75 80

Phe Pro Ile Thr Ile Thr Glu Pro Asp Lys Thr Ile Gln Thr Phe Glu
 85 90 95

Arg Val Ile Asp Val Phe Ala Asp Phe Lys Leu Val Arg Lys Glu Pro
 100 105 110

Val Leu Val Val Gly Gly Gly Leu Ile Thr Asp Val Val Gly Phe Ala
 115 120 125

Cys Ser Thr Tyr Arg Arg Ser Ser Asn Tyr Ile Arg Ile Pro Thr Thr
 130 135 140

Leu Ile Gly Leu Ile Asp Ala Ser Val Ala Ile Lys Val Ala Val Asn
 145 150 155 160

His Arg Lys Leu Lys Asn Arg Leu Gly Ala Tyr His Ala Ser Arg Lys
 165 170 175

Val Phe Leu Asp Phe Ser Leu Leu Arg Thr Leu Pro Thr Asp Gln Val
 180 185 190

Arg Asn Gly Met Ala Glu Leu Val Lys Ile Ala Val Val Ala His Gln
 195 200 205

Glu Val Phe Glu Leu Leu Glu Lys Tyr Gly Glu Glu Leu Leu Arg Thr
 210 215 220

His Phe Gly Asn Ile Asp Ala Thr Pro Glu Ile Lys Glu Ile Ala His
 225 230 235 240

Arg Leu Thr Tyr Lys Ala Ile His Lys Met Leu Glu Leu Glu Val Pro
 245 250 255

```

Asn Leu His Glu Leu Asp Leu Asp Arg Val Ile Ala Tyr Gly His Thr
260                               265                      270

Trp Ser Pro Thr Leu Glu Leu Ala Pro Arg Leu Pro Met Phe His Gly
275                               280                      285

His Ala Val Asn Val Asp Met Ala Phe Ser Ala Thr Ile Ala Ala Arg
290                               295                      300

Arg Gly Tyr Ile Thr Ile Ala Glu Arg Asp Arg Ile Leu Gly Leu Met
305                               310                      315                      320

Ser Arg Val Gly Leu Ser Leu Asp His Pro Met Leu Asp Ile Asp Ile
325                               330                      335

Leu Trp Arg Gly Thr Glu Ser Ile Thr Leu Thr Arg Asp Gly Leu Leu
340                               345                      350

Arg Ala Ala Met Pro Lys Pro Ile Gly Asp Cys Val Phe Val Asn Asp
355                               360                      365

Leu Thr Arg Glu Glu Leu Ala Ala Ala Leu Ala Asp His Lys Glu Leu
370                               375                      380

Cys Thr Ser Tyr Pro Arg Gly Gly Glu Gly Val Asp Val Tyr Pro Val
385                               390                      395                      400

Tyr Gln Lys Glu Leu Ile Gly Ser Val Lys Met Ser Ile Val Gln Ala
405                               410                      415

Lys Phe Glu Ala Lys Glu Thr Ser Phe His Val Glu Gly Tyr Glu Lys
420                               425                      430

Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly Ile Phe Glu Ile Gln Asn
435                               440                      445

Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe Gly Arg Cys Leu Ala Ile
450                               455                      460

Val Asp Ala Asn Val Ser Arg Leu Tyr Gly Asn Gln Ile Gln Ala Tyr
465                               470                      475                      480

Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu Phe Pro Ile Thr Ile Thr
485                               490                      495

Glu Pro Asp Lys Thr Ile Gln Thr Phe Glu Arg Val Ile Asp Val Phe
500                               505                      510

Ala Asp Phe Lys Leu Val Arg Lys Glu Pro Val Leu Val Val Gly Gly
515                               520                      525

Gly Leu Ile Thr Asp Val Val Gly Phe Ala Cys Ser Thr Tyr Arg Arg
530                               535                      540

Ser Ser Asn Tyr Ile Arg Ile Pro Thr Thr Leu Ile Gly Leu Ile Asp
545                               550                      555                      560

Ala Ser Val Ala Ile Lys Val Ala Val Asn His Arg Lys Leu Lys Asn
565                               570                      575

Arg Leu Gly Ala Tyr His Ala Ser Arg Lys Val Phe Leu Asp Phe Ser
580                               585                      590

Leu Leu Arg Thr Leu Pro Thr Asp Gln Val Arg Asn Gly Met Ala Glu

```

595 600 605
 Leu Val Lys Ile Ala Val Val Ala His Gln Glu Val Phe Glu Leu Leu
 610 615 620
 Glu Lys Tyr Gly Glu Glu Leu Leu Arg Thr His Phe Gly Asn Ile Asp
 625 630 635 640
 Ala Thr Pro Glu Ile Lys Glu Ile Ala His Arg Leu Thr Tyr Lys Ala
 645 650 655
 Ile His Lys Met Leu Glu Leu Glu Val Pro Asn Leu His Glu Leu Asp
 660 665 670
 Leu Asp Arg Val Ile Ala Tyr Gly His Thr Trp Ser Pro Thr Leu Glu
 675 680 685
 Leu Ala Pro Arg Leu Pro Met Phe His Gly His Ala Val Asn Val Asp
 690 695 700
 Met Ala Phe Ser Ala Thr Ile Ala Ala Arg Arg Gly Tyr Ile Thr Ile
 705 710 715 720
 Ala Glu Arg Asp Arg Ile Leu Gly Leu Met Ser Arg Val Gly Leu Ser
 725 730 735
 Leu Asp His Pro Met Leu Asp Ile Asp Ile Leu Trp Arg Gly Thr Glu
 740 745 750
 Ser Ile Thr Leu Thr Arg Asp Gly Leu Leu Arg Ala Ala Met Pro Lys
 755 760 765
 Pro Ile Gly Asp Cys Val Phe Val Asn Asp Leu Thr Arg Glu Glu Leu
 770 775 780
 Ala Ala Ala Leu Ala Asp His Lys Glu Leu Cys Thr Ser Tyr Pro Arg
 785 790 795 800
 Gly Gly Glu Gly Val Asp Val Tyr Pro Val Tyr Gln Lys Glu Leu Ile
 805 810 815
 Gly Ser Val Lys
 820

 <210> 114
 <211> 1099
 <212> PRT
 <213> Anabaena variabilis ATCC29413

 <400> 114
 Met Ser Ile Val Gln Ala Lys Phe Glu Ala Lys Glu Thr Ser Phe His
 1 5 10 15
 Val Glu Gly Tyr Glu Lys Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly
 20 25 30
 Ile Phe Glu Ile Gln Asn Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe
 35 40 45
 Gly Arg Cys Leu Ala Ile Val Asp Ala Asn Val Ser Arg Leu Tyr Gly
 50 55 60
 Asn Gln Ile Gln Ala Tyr Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu

65		70		75		80									
Phe	Pro	Ile	Thr	Ile	Thr	Glu	Pro	Asp	Lys	Thr	Ile	Gln	Thr	Phe	Glu
				85					90					95	
Arg	Val	Ile	Asp	Val	Phe	Ala	Asp	Phe	Lys	Leu	Val	Arg	Lys	Glu	Pro
			100					105					110		
Val	Leu	Val	Val	Gly	Gly	Gly	Leu	Ile	Thr	Asp	Val	Val	Gly	Phe	Ala
		115					120					125			
Cys	Ser	Thr	Tyr	Arg	Arg	Ser	Ser	Asn	Tyr	Ile	Arg	Ile	Pro	Thr	Thr
	130					135					140				
Leu	Ile	Gly	Leu	Ile	Asp	Ala	Ser	Val	Ala	Ile	Lys	Val	Ala	Val	Asn
145					150					155					160
His	Arg	Lys	Leu	Lys	Asn	Arg	Leu	Gly	Ala	Tyr	His	Ala	Ser	Arg	Lys
				165					170					175	
Val	Phe	Leu	Asp	Phe	Ser	Leu	Leu	Arg	Thr	Leu	Pro	Thr	Asp	Gln	Val
			180					185					190		
Arg	Asn	Gly	Met	Ala	Glu	Leu	Val	Lys	Ile	Ala	Val	Val	Ala	His	Gln
		195					200					205			
Glu	Val	Phe	Glu	Leu	Leu	Glu	Lys	Tyr	Gly	Glu	Glu	Leu	Leu	Arg	Thr
	210					215					220				
His	Phe	Gly	Asn	Ile	Asp	Ala	Thr	Pro	Glu	Ile	Lys	Glu	Ile	Ala	His
225					230					235					240
Arg	Leu	Thr	Tyr	Lys	Ala	Ile	His	Lys	Met	Leu	Glu	Leu	Glu	Val	Pro
				245					250					255	
Asn	Leu	His	Glu	Leu	Asp	Leu	Asp	Arg	Val	Ile	Ala	Tyr	Gly	His	Thr
			260					265					270		
Trp	Ser	Pro	Thr	Leu	Glu	Leu	Ala	Pro	Arg	Leu	Pro	Met	Phe	His	Gly
		275					280					285			
His	Ala	Val	Asn	Val	Asp	Met	Ala	Phe	Ser	Ala	Thr	Ile	Ala	Ala	Arg
	290					295					300				
Arg	Gly	Tyr	Ile	Thr	Ile	Ala	Glu	Arg	Asp	Arg	Ile	Leu	Gly	Leu	Met
305					310					315					320
Ser	Arg	Val	Gly	Leu	Ser	Leu	Asp	His	Pro	Met	Leu	Asp	Ile	Asp	Ile
				325					330					335	
Leu	Trp	Arg	Gly	Thr	Glu	Ser	Ile	Thr	Leu	Thr	Arg	Asp	Gly	Leu	Leu
			340					345					350		
Arg	Ala	Ala	Met	Pro	Lys	Pro	Ile	Gly	Asp	Cys	Val	Phe	Val	Asn	Asp
		355					360					365			
Leu	Thr	Arg	Glu	Glu	Leu	Ala	Ala	Ala	Leu	Ala	Asp	His	Lys	Glu	Leu
		370				375					380				
Cys	Thr	Ser	Tyr	Pro	Arg	Gly	Gly	Glu	Gly	Val	Asp	Val	Tyr	Pro	Val
385					390					395					400
Tyr	Gln	Lys	Glu	Leu	Ile	Gly	Ser	Val	Lys	Met	Ser	Ile	Val	Gln	Ala
				405					410					415	

Lys Phe Glu Ala Lys Glu Thr Ser Phe His Val Glu Gly Tyr Glu Lys
 420 425 430
 Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly Ile Phe Glu Ile Gln Asn
 435 440 445
 Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe Gly Arg Cys Leu Ala Ile
 450 455 460
 Val Asp Ala Asn Val Ser Arg Leu Tyr Gly Asn Gln Ile Gln Ala Tyr
 465 470 475 480
 Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu Phe Pro Ile Thr Ile Thr
 485 490 495
 Glu Pro Asp Lys Thr Ile Gln Thr Phe Glu Arg Val Ile Asp Val Phe
 500 505 510
 Ala Asp Phe Lys Leu Val Arg Lys Glu Pro Val Leu Val Val Gly Gly
 515 520 525
 Gly Leu Ile Thr Asp Val Val Gly Phe Ala Cys Ser Thr Tyr Arg Arg
 530 535 540
 Ser Ser Asn Tyr Ile Arg Ile Pro Thr Thr Leu Ile Gly Leu Ile Asp
 545 550 555 560
 Ala Ser Val Ala Ile Lys Val Ala Val Asn His Arg Lys Leu Lys Asn
 565 570 575
 Arg Leu Gly Ala Tyr His Ala Ser Arg Lys Val Phe Leu Asp Phe Ser
 580 585 590
 Leu Leu Arg Thr Leu Pro Thr Asp Gln Val Arg Asn Gly Met Ala Glu
 595 600 605
 Leu Val Lys Ile Ala Val Val Ala His Gln Glu Val Phe Glu Leu Leu
 610 615 620
 Glu Lys Tyr Gly Glu Glu Leu Leu Arg Thr His Phe Gly Asn Ile Asp
 625 630 635 640
 Ala Thr Pro Glu Ile Lys Glu Ile Ala His Arg Leu Thr Tyr Lys Ala
 645 650 655
 Ile His Lys Met Leu Glu Leu Glu Val Pro Asn Leu His Glu Leu Asp
 660 665 670
 Leu Asp Arg Val Ile Ala Tyr Gly His Thr Trp Ser Pro Thr Leu Glu
 675 680 685
 Leu Ala Pro Arg Leu Pro Met Phe His Gly His Ala Val Asn Val Asp
 690 695 700
 Met Ala Phe Ser Ala Thr Ile Ala Ala Arg Arg Gly Tyr Ile Thr Ile
 705 710 715 720
 Ala Glu Arg Asp Arg Ile Leu Gly Leu Met Ser Arg Val Gly Leu Ser
 725 730 735
 Leu Asp His Pro Met Leu Asp Ile Asp Ile Leu Trp Arg Gly Thr Glu
 740 745 750

Ser Ile Thr Leu Thr Arg Asp Gly Leu Leu Arg Ala Ala Met Pro Lys
 755 760 765
 Pro Ile Gly Asp Cys Val Phe Val Asn Asp Leu Thr Arg Glu Glu Leu
 770 775 780
 Ala Ala Ala Leu Ala Asp His Lys Glu Leu Cys Thr Ser Tyr Pro Arg
 785 790 795 800
 Gly Gly Glu Gly Val Asp Val Tyr Pro Val Tyr Gln Lys Glu Leu Ile
 805 810 815
 Gly Ser Val Lys Met Thr Asn Val Ile Val Gln Pro Thr Ala Arg Pro
 820 825 830
 Val Thr Pro Leu Gly Ile Leu Thr Lys Gln Leu Glu Ala Ile Val Gln
 835 840 845
 Glu Val Lys Gln His Pro Asp Leu Pro Gly Glu Leu Ile Ala Asn Ile
 850 855 860
 His Gln Ala Trp Arg Leu Ala Ala Gly Ile Asp Pro Tyr Leu Glu Glu
 865 870 875 880
 Cys Thr Thr Pro Glu Ser Pro Glu Leu Ala Ala Leu Ala Lys Thr Thr
 885 890 895
 Ala Thr Glu Ala Trp Gly Glu His Phe His Gly Gly Thr Thr Val Arg
 900 905 910
 Pro Leu Glu Gln Glu Met Leu Ser Gly His Ile Glu Gly Gln Thr Leu
 915 920 925
 Lys Met Phe Val His Met Thr Lys Ala Lys Lys Val Leu Glu Ile Gly
 930 935 940
 Met Phe Thr Gly Tyr Ser Ala Leu Ala Met Ala Glu Ala Leu Pro Glu
 945 950 955 960
 Asp Gly Leu Leu Val Ala Cys Glu Val Asp Pro Tyr Ala Ala Glu Ile
 965 970 975
 Gly Gln Lys Ala Phe Gln Gln Ser Pro His Gly Gly Lys Ile Arg Val
 980 985 990
 Glu Leu Asp Ala Ala Leu Ala Thr Leu Asp Lys Leu Ala Glu Ala Gly
 995 1000 1005
 Glu Ser Phe Asp Leu Val Phe Ile Asp Ala Asp Lys Lys Glu Tyr Val
 1010 1015 1020
 Ala Tyr Phe His Lys Leu Leu Gly Ser Ser Leu Leu Ala Pro Asp Gly
 1025 1030 1035 1040
 Phe Ile Cys Val Asp Asn Thr Leu Leu Gln Gly Glu Val Tyr Leu Pro
 1045 1050 1055
 Ala Glu Glu Arg Ser Val Asn Gly Glu Ala Ile Ala Gln Phe Asn His
 1060 1065 1070
 Thr Val Ala Ile Asp Pro Arg Val Glu Gln Val Leu Leu Pro Leu Arg
 1075 1080 1085
 Asp Gly Leu Thr Ile Ile Arg Arg Ile Gln Pro

1090

1095

<210> 115
 <211> 410
 <212> PRT
 <213> Anabaena variabilis ATCC29413

<400> 115

Met Ser Ile Val Gln Ala Lys Phe Glu Ala Lys Glu Thr Ser Phe His
 1 5 10 15
 Val Glu Gly Tyr Glu Lys Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly
 20 25 30
 Ile Phe Glu Ile Gln Asn Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe
 35 40 45
 Gly Arg Cys Leu Ala Ile Val Asp Ala Asn Val Ser Arg Leu Tyr Gly
 50 55 60
 Asn Gln Ile Gln Ala Tyr Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu
 65 70 75 80
 Phe Pro Ile Thr Ile Thr Glu Pro Asp Lys Thr Ile Gln Thr Phe Glu
 85 90 95
 Arg Val Ile Asp Val Phe Ala Asp Phe Lys Leu Val Arg Lys Glu Pro
 100 105 110
 Val Leu Val Val Gly Gly Gly Leu Ile Thr Asp Val Val Gly Phe Ala
 115 120 125
 Cys Ser Thr Tyr Arg Arg Ser Ser Asn Tyr Ile Arg Ile Pro Thr Thr
 130 135 140
 Leu Ile Gly Leu Ile Asp Ala Ser Val Ala Ile Lys Val Ala Val Asn
 145 150 155 160
 His Arg Lys Leu Lys Asn Arg Leu Gly Ala Tyr His Ala Ser Arg Lys
 165 170 175
 Val Phe Leu Asp Phe Ser Leu Leu Arg Thr Leu Pro Thr Asp Gln Val
 180 185 190
 Arg Asn Gly Met Ala Glu Leu Val Lys Ile Ala Val Val Ala His Gln
 195 200 205
 Glu Val Phe Glu Leu Leu Glu Lys Tyr Gly Glu Glu Leu Leu Arg Thr
 210 215 220
 His Phe Gly Asn Ile Asp Ala Thr Pro Glu Ile Lys Glu Ile Ala His
 225 230 235 240
 Arg Leu Thr Tyr Lys Ala Ile His Lys Met Leu Glu Leu Glu Val Pro
 245 250 255
 Asn Leu His Glu Leu Asp Leu Asp Arg Val Ile Ala Tyr Gly His Thr
 260 265 270
 Trp Ser Pro Thr Leu Glu Leu Ala Pro Arg Leu Pro Met Phe His Gly
 275 280 285
 His Ala Val Asn Val Asp Met Ala Phe Ser Ala Thr Ile Ala Ala Arg

```

290                295                300
Arg Gly Tyr Ile Thr Ile Ala Glu Arg Asp Arg Ile Leu Gly Leu Met
305                310                315                320

Ser Arg Val Gly Leu Ser Leu Asp His Pro Met Leu Asp Ile Asp Ile
325                330                335

Leu Trp Arg Gly Thr Glu Ser Ile Thr Leu Thr Arg Asp Gly Leu Leu
340                345                350

Arg Ala Ala Met Pro Lys Pro Ile Gly Asp Cys Val Phe Val Asn Asp
355                360                365

Leu Thr Arg Glu Glu Leu Ala Ala Ala Leu Ala Asp His Lys Glu Leu
370                375                380

Cys Thr Ser Tyr Pro Arg Gly Gly Glu Gly Val Asp Val Tyr Pro Val
385                390                395                400

Tyr Gln Lys Glu Leu Ile Gly Ser Val Lys
405                410

<210>      116
<211>      410
<212>      PRT
<213>      Anabaena variabilis ATCC29413

<400>      116
Met Ser Ile Val Gln Ala Lys Phe Glu Ala Lys Glu Thr Ser Phe His
 1                5                10                15

Val Glu Gly Tyr Glu Lys Ile Glu Tyr Asp Leu Val Tyr Val Asp Gly
20                25                30

Ile Phe Glu Ile Gln Asn Ser Ala Leu Ala Asp Val Tyr Gln Gly Phe
35                40                45

Gly Arg Cys Leu Ala Ile Val Asp Ala Asn Val Ser Arg Leu Tyr Gly
50                55                60

Asn Gln Ile Gln Ala Tyr Phe Gln Tyr Tyr Gly Ile Glu Leu Arg Leu
65                70                75                80

Phe Pro Ile Thr Ile Thr Glu Pro Asp Lys Thr Ile Gln Thr Phe Glu
85                90                95

Arg Val Ile Asp Val Phe Ala Asp Phe Lys Leu Val Arg Lys Glu Pro
100               105               110

Val Leu Val Val Gly Gly Gly Leu Ile Thr Asp Val Val Gly Phe Ala
115               120               125

Cys Ser Thr Tyr Arg Arg Ser Ser Asn Tyr Ile Arg Ile Pro Thr Thr
130               135               140

Leu Ile Gly Leu Ile Asp Ala Ser Val Ala Ile Lys Val Ala Val Asn
145               150               155               160

His Arg Lys Leu Lys Asn Arg Leu Gly Ala Tyr His Ala Ser Arg Lys
165               170               175

Val Phe Leu Asp Phe Ser Leu Leu Arg Thr Leu Pro Thr Asp Gln Val

```

180	185	190
Arg Asn Gly Met Ala Glu Leu Val Lys Ile Ala Val Val Ala His Gln		
195	200	205
Glu Val Phe Glu Leu Leu Glu Lys Tyr Gly Glu Glu Leu Leu Arg Thr		
210	215	220
His Phe Gly Asn Ile Asp Ala Thr Pro Glu Ile Lys Glu Ile Ala His		
225	230	235
Arg Leu Thr Tyr Lys Ala Ile His Lys Met Leu Glu Leu Glu Val Pro		
245	250	255
Asn Leu His Glu Leu Asp Leu Asp Arg Val Ile Ala Tyr Gly His Thr		
260	265	270
Trp Ser Pro Thr Leu Glu Leu Ala Pro Arg Leu Pro Met Phe His Gly		
275	280	285
His Ala Val Asn Val Asp Met Ala Phe Ser Ala Thr Ile Ala Ala Arg		
290	295	300
Arg Gly Tyr Ile Thr Ile Ala Glu Arg Asp Arg Ile Leu Gly Leu Met		
305	310	315
Ser Arg Val Gly Leu Ser Leu Asp His Pro Met Leu Asp Ile Asp Ile		
325	330	335
Leu Trp Arg Gly Thr Glu Ser Ile Thr Leu Thr Arg Asp Gly Leu Leu		
340	345	350
Arg Ala Ala Met Pro Lys Pro Ile Gly Asp Cys Val Phe Val Asn Asp		
355	360	365
Leu Thr Arg Glu Glu Leu Ala Ala Ala Leu Ala Asp His Lys Glu Leu		
370	375	380
Cys Thr Ser Tyr Pro Arg Gly Gly Glu Gly Val Asp Val Tyr Pro Val		
385	390	395
Tyr Gln Lys Glu Leu Ile Gly Ser Val Lys		
405	410	
<210>	117	
<211>	279	
<212>	PRT	
<213>	Anabaena variabilis ATCC29413	
<400>	117	
Met Thr Asn Val Ile Val Gln Pro Thr Ala Arg Pro Val Thr Pro Leu		
1	5	10
Gly Ile Leu Thr Lys Gln Leu Glu Ala Ile Val Gln Glu Val Lys Gln		
20	25	30
His Pro Asp Leu Pro Gly Glu Leu Ile Ala Asn Ile His Gln Ala Trp		
35	40	45
Arg Leu Ala Ala Gly Ile Asp Pro Tyr Leu Glu Glu Cys Thr Thr Pro		
50	55	60
Glu Ser Pro Glu Leu Ala Ala Leu Ala Lys Thr Thr Ala Thr Glu Ala		

<210>	118
<211>	888
<212>	PRT
<213>	Anabaena variabilis ATCC29413

85										90					95				
Leu	Lys	Ala	Gly	Gly	Ala	Tyr	Val	Pro	Ile	Asp	Pro	Glu	Tyr	Pro	Gln				
			100					105					110						
Glu	Arg	Ile	Ala	Tyr	Met	Leu	Glu	Asp	Ser	Gln	Val	Lys	Val	Leu	Leu				
		115					120					125							
Thr	Gln	Glu	Lys	Leu	Leu	Asn	Gln	Ile	Pro	His	His	Gln	Ala	Gln	Thr				
		130				135						140							
Ile	Cys	Val	Asp	Arg	Glu	Trp	Glu	Lys	Ile	Ser	Thr	Gln	Ala	Asn	Thr				
145					150					155					160				
Asn	Pro	Lys	Ser	Asn	Ile	Lys	Thr	Asp	Asn	Leu	Ala	Tyr	Val	Ile	Tyr				
				165					170					175					
Thr	Ser	Gly	Ser	Thr	Gly	Lys	Pro	Lys	Gly	Ala	Met	Asn	Thr	His	Lys				
			180					185					190						
Gly	Ile	Cys	Asn	Arg	Leu	Leu	Trp	Met	Gln	Glu	Ala	Tyr	Gln	Ile	Asp				
		195					200					205							
Ser	Thr	Asp	Ser	Ile	Leu	Gln	Lys	Thr	Pro	Phe	Ser	Phe	Asp	Val	Ser				
		210				215					220								
Val	Trp	Glu	Phe	Phe	Trp	Thr	Leu	Leu	Thr	Gly	Ala	Arg	Leu	Val	Ile				
225					230					235					240				
Ala	Lys	Pro	Gly	Gly	His	Lys	Asp	Ser	Ala	Tyr	Leu	Ile	Asp	Leu	Ile				
				245					250					255					
Thr	Gln	Glu	Gln	Ile	Thr	Thr	Leu	His	Phe	Val	Pro	Ser	Met	Leu	Gln				
			260					265					270						
Val	Phe	Leu	Gln	Asn	Arg	His	Val	Ser	Lys	Cys	Ser	Ser	Leu	Lys	Arg				
		275					280					285							
Val	Ile	Cys	Ser	Gly	Glu	Ala	Leu	Ser	Ile	Asp	Leu	Gln	Asn	Arg	Phe				
		290				295					300								
Phe	Gln	His	Leu	Gln	Cys	Glu	Leu	His	Asn	Leu	Tyr	Gly	Pro	Thr	Glu				
305					310					315					320				
Ala	Ala	Ile	Asp	Val	Thr	Phe	Trp	Gln	Cys	Arg	Lys	Asp	Ser	Asn	Leu				
				325					330					335					
Lys	Ser	Val	Pro	Ile	Gly	Arg	Pro	Ile	Ala	Asn	Thr	Gln	Ile	Tyr	Ile				
			340					345					350						
Leu	Asp	Ala	Asp	Leu	Gln	Pro	Val	Asn	Ile	Gly	Val	Thr	Gly	Glu	Ile				
		355					360					365							
Tyr	Ile	Gly	Gly	Val	Gly	Val	Ala	Arg	Gly	Tyr	Leu	Asn	Lys	Glu	Glu				
		370				375					380								
Leu	Thr	Lys	Glu	Lys	Phe	Ile	Ile	Asn	Pro	Phe	Pro	Asn	Ser	Glu	Phe				
385					390					395					400				
Lys	Arg	Leu	Tyr	Lys	Thr	Gly	Asp	Leu	Ala	Arg	Tyr	Leu	Pro	Asp	Gly				
				405					410					415					
Asn	Ile	Glu	Tyr	Leu	Gly	Arg	Thr	Asp	Tyr	Gln	Val	Lys	Ile	Arg	Gly				
			420					425					430						

Tyr Arg Ile Glu Ile Gly Glu Ile Glu Asn Val Leu Ser Ser His Pro
 435 440 445
 Gln Val Arg Glu Ala Val Val Ile Ala Arg Asp Asp Asn Ala Gln Glu
 450 455 460
 Lys Gln Ile Ile Ala Tyr Ile Thr Tyr Asn Ser Ile Lys Pro Gln Leu
 465 470 475 480
 Asp Asn Leu Arg Asp Phe Leu Lys Ala Arg Leu Pro Asp Phe Met Ile
 485 490 495
 Pro Ala Ala Phe Val Met Leu Glu His Leu Pro Leu Thr Pro Ser Gly
 500 505 510
 Lys Val Asp Arg Lys Ala Leu Pro Lys Pro Asp Leu Phe Asn Tyr Ser
 515 520 525
 Glu His Asn Ser Tyr Val Ala Pro Arg Asn Glu Val Glu Glu Lys Leu
 530 535 540
 Val Gln Ile Trp Ser Asn Ile Leu His Leu Pro Lys Val Gly Val Thr
 545 550 555 560
 Glu Asn Phe Phe Ala Ile Gly Gly Asn Ser Leu Lys Ala Leu His Leu
 565 570 575
 Ile Ser Gln Ile Glu Glu Leu Phe Ala Lys Glu Ile Ser Leu Ala Thr
 580 585 590
 Leu Leu Thr Asn Pro Val Ile Ala Asp Leu Ala Lys Val Ile Gln Ala
 595 600 605
 Asn Asn Gln Ile His Asn Ser Pro Leu Val Pro Ile Gln Pro Gln Gly
 610 615 620
 Lys Gln Gln Pro Phe Phe Cys Ile His Pro Ala Gly Gly His Val Leu
 625 630 635 640
 Cys Tyr Phe Lys Leu Ala Gln Tyr Ile Gly Thr Asp Gln Pro Phe Tyr
 645 650 655
 Gly Leu Gln Ala Gln Gly Phe Tyr Gly Asp Glu Ala Pro Leu Thr Arg
 660 665 670
 Val Glu Asp Met Ala Ser Leu Tyr Val Lys Thr Ile Arg Glu Phe Gln
 675 680 685
 Pro Gln Gly Pro Tyr Arg Val Gly Gly Trp Ser Phe Gly Gly Val Val
 690 695 700
 Ala Tyr Glu Val Ala Gln Gln Leu His Arg Gln Gly Gln Glu Val Ser
 705 710 715 720
 Leu Leu Ala Ile Leu Asp Ser Tyr Val Pro Ile Leu Leu Asp Lys Gln
 725 730 735
 Lys Pro Ile Asp Asp Val Tyr Leu Val Gly Val Leu Ser Arg Val Phe
 740 745 750
 Gly Gly Met Phe Gly Gln Asp Asn Leu Val Thr Pro Glu Glu Ile Glu
 755 760 765

Asn Leu Thr Val Glu Glu Lys Ile Asn Tyr Ile Ile Asp Lys Ala Arg
 770 775 780
 Ser Ala Arg Ile Phe Pro Pro Gly Val Glu Arg Gln Asn Asn Arg Arg
 785 790 795 800
 Ile Leu Asp Val Leu Val Gly Thr Leu Lys Ala Thr Tyr Ser Tyr Ile
 805 810 815
 Arg Gln Pro Tyr Pro Gly Lys Val Thr Val Phe Arg Ala Arg Glu Lys
 820 825 830
 His Ile Met Ala Pro Asp Pro Thr Leu Val Trp Val Glu Leu Phe Ser
 835 840 845
 Val Met Ala Ala Gln Glu Ile Lys Ile Ile Asp Val Pro Gly Asn His
 850 855 860
 Tyr Ser Phe Val Leu Glu Pro His Val Gln Val Leu Ala Gln Arg Leu
 865 870 875 880
 Gln Asp Cys Leu Glu Asn Asn Ser
 885
 <210> 119
 <211> 410
 <212> PRT
 <213> Nostoc punctiforme ATCC 29133
 <400> 119
 Met Ser Asn Val Gln Ala Ser Phe Glu Ala Thr Glu Ala Glu Phe Arg
 1 5 10 15
 Val Glu Gly Tyr Glu Lys Ile Glu Phe Ser Leu Val Tyr Val Asn Gly
 20 25 30
 Ala Phe Asp Ile Ser Asn Arg Glu Ile Ala Asp Ser Tyr Glu Lys Phe
 35 40 45
 Gly Arg Cys Leu Thr Val Ile Asp Ala Asn Val Asn Arg Leu Tyr Gly
 50 55 60
 Lys Gln Ile Lys Ser Tyr Phe Arg His Tyr Gly Ile Asp Leu Thr Val
 65 70 75 80
 Val Pro Ile Val Ile Thr Glu Pro Thr Lys Thr Leu Ala Thr Phe Glu
 85 90 95
 Lys Ile Val Asp Ala Phe Ser Asp Phe Gly Leu Ile Arg Lys Glu Pro
 100 105 110
 Val Leu Val Val Gly Gly Gly Leu Thr Thr Asp Val Ala Gly Phe Ala
 115 120 125
 Cys Ala Ala Tyr Arg Arg Lys Ser Asn Tyr Ile Arg Val Pro Thr Thr
 130 135 140
 Leu Ile Gly Leu Ile Asp Ala Gly Val Ala Ile Lys Val Ala Val Asn
 145 150 155 160
 His Arg Lys Leu Lys Asn Arg Leu Gly Ala Tyr His Ala Pro Leu Lys
 165 170 175

Val Ile Leu Asp Phe Ser Phe Leu Gln Thr Leu Pro Thr Ala Gln Val
 180 185 190
 Arg Asn Gly Met Ala Glu Leu Val Lys Ile Ala Val Val Ala Asn Ser
 195 200 205
 Glu Val Phe Glu Leu Leu Tyr Glu Tyr Gly Glu Glu Leu Leu Ser Thr
 210 215 220
 His Phe Gly Tyr Val Asn Gly Thr Lys Glu Leu Lys Ala Ile Ala His
 225 230 235 240
 Lys Leu Asn Tyr Glu Ala Ile Lys Thr Met Leu Glu Leu Glu Thr Pro
 245 250 255
 Asn Leu His Glu Leu Asp Leu Asp Arg Val Ile Ala Tyr Gly His Thr
 260 265 270
 Trp Ser Pro Thr Leu Glu Leu Ala Pro Met Ile Pro Leu Phe His Gly
 275 280 285
 His Ala Val Asn Ile Asp Met Ala Leu Ser Ala Thr Ile Ala Ala Arg
 290 295 300
 Arg Gly Tyr Ile Thr Ser Gly Glu Arg Asp Arg Ile Leu Ser Leu Met
 305 310 315 320
 Ser Arg Ile Gly Leu Ser Ile Asp His Pro Leu Leu Asp Gly Asp Leu
 325 330 335
 Leu Trp Tyr Ala Thr Gln Ser Ile Ser Leu Thr Arg Asp Gly Lys Gln
 340 345 350
 Arg Ala Ala Met Pro Lys Pro Ile Gly Glu Cys Phe Phe Val Asn Asp
 355 360 365
 Phe Thr Arg Glu Glu Leu Asp Ala Ala Leu Ala Glu His Lys Arg Leu
 370 375 380
 Cys Ala Thr Tyr Pro Arg Gly Gly Asp Gly Ile Asp Ala Tyr Ile Glu
 385 390 395 400
 Thr Gln Glu Glu Ser Lys Leu Leu Gly Val
 405 410

<210> 120
 <211> 277
 <212> PRT
 <213> Nostoc punctiforme ATCC 29133

<400> 120
 Met Thr Ser Ile Leu Gly Arg Asp Thr Ala Arg Pro Ile Thr Pro His
 1 5 10 15
 Ser Ile Leu Val Ala Gln Leu Gln Lys Thr Leu Arg Met Ala Glu Glu
 20 25 30
 Ser Asn Ile Pro Ser Glu Ile Leu Thr Ser Leu Arg Gln Gly Leu Gln
 35 40 45
 Leu Ala Ala Gly Leu Asp Pro Tyr Leu Asp Asp Cys Thr Thr Pro Glu
 50 55 60

Ser Thr Ala Leu Thr Ala Leu Ala Gln Lys Thr Ser Ile Glu Asp Trp
 65 70 75 80
 Ser Lys Arg Phe Ser Asp Gly Glu Thr Val Arg Gln Leu Glu Gln Glu
 85 90 95
 Met Leu Ser Gly His Leu Glu Gly Gln Thr Leu Lys Met Phe Val His
 100 105 110
 Ile Thr Lys Ala Lys Ser Ile Leu Glu Val Gly Met Phe Thr Gly Tyr
 115 120 125
 Ser Ala Leu Ala Met Ala Glu Ala Leu Pro Asp Asp Gly Arg Leu Ile
 130 135 140
 Ala Cys Glu Val Asp Ser Tyr Val Ala Glu Phe Ala Gln Thr Cys Phe
 145 150 155 160
 Gln Glu Ser Pro His Gly Arg Lys Ile Val Val Glu Val Ala Pro Ala
 165 170 175
 Leu Glu Thr Leu His Lys Leu Val Ala Lys Lys Glu Ser Phe Asp Leu
 180 185 190
 Ile Phe Ile Asp Ala Asp Lys Lys Glu Tyr Ile Glu Tyr Phe Gln Ile
 195 200 205
 Ile Leu Asp Ser His Leu Leu Ala Pro Asp Gly Leu Ile Cys Val Asp
 210 215 220
 Asn Thr Leu Leu Gln Gly Gln Val Tyr Leu Pro Ser Glu Gln Arg Thr
 225 230 235 240
 Ala Asn Gly Glu Ala Ile Ala Gln Phe Asn Arg Ile Val Ala Ala Asp
 245 250 255
 Pro Arg Val Glu Gln Val Leu Leu Pro Ile Arg Asp Gly Ile Thr Leu
 260 265 270
 Ile Arg Arg Leu Val
 275
 <210> 121
 <211> 461
 <212> PRT
 <213> Nostoc punctiforme ATCC 29133
 <400> 121
 Met Ala Gln Ser Ile Ser Leu Ser Leu Pro Gln Ser Thr Thr Pro Ser
 1 5 10 15
 Lys Gly Val Arg Leu Lys Ile Ala Ala Leu Leu Lys Thr Ile Gly Thr
 20 25 30
 Leu Ile Leu Leu Leu Ile Ala Leu Pro Leu Asn Ala Leu Ile Val Leu
 35 40 45
 Ile Ser Leu Met Cys Arg Pro Phe Thr Lys Lys Pro Ala Val Ala Thr
 50 55 60
 His Pro Gln Asn Ile Leu Val Ser Gly Gly Lys Met Thr Lys Ala Leu
 65 70 75 80

Gln Leu Ala Arg Ser Phe His Ala Ala Gly His Arg Val Ile Leu Ile
 85 90 95
 Glu Gly His Lys Tyr Trp Leu Ser Gly His Arg Phe Ser Asn Ser Val
 100 105 110
 Ser Arg Phe Tyr Thr Val Pro Ala Pro Gln Asp Asp Pro Glu Gly Tyr
 115 120 125
 Thr Gln Ala Leu Leu Glu Ile Val Lys Arg Glu Lys Ile Asp Val Tyr
 130 135 140
 Val Pro Val Cys Ser Pro Val Ala Ser Tyr Tyr Asp Ser Leu Ala Lys
 145 150 155 160
 Ser Ala Leu Ser Glu Tyr Cys Glu Val Phe His Phe Asp Ala Asp Ile
 165 170 175
 Thr Lys Met Leu Asp Asp Lys Phe Ala Phe Thr Asp Arg Ala Arg Ser
 180 185 190
 Leu Gly Leu Ser Ala Pro Lys Ser Phe Lys Ile Thr Asp Pro Glu Gln
 195 200 205
 Val Ile Asn Phe Asp Phe Ser Lys Glu Thr Arg Lys Tyr Ile Leu Lys
 210 215 220
 Ser Ile Ser Tyr Asp Ser Val Arg Arg Leu Asn Leu Thr Lys Leu Pro
 225 230 235 240
 Cys Asp Thr Pro Glu Glu Thr Ala Ala Phe Val Lys Ser Leu Pro Ile
 245 250 255
 Ser Pro Glu Lys Pro Trp Ile Met Gln Glu Phe Ile Pro Gly Lys Glu
 260 265 270
 Leu Cys Thr His Ser Thr Val Arg Asp Gly Glu Leu Arg Leu His Cys
 275 280 285
 Cys Ser Asn Ser Ser Ala Phe Gln Ile Asn Tyr Glu Asn Val Glu Asn
 290 295 300
 Pro Gln Ile Gln Glu Trp Val Gln His Phe Val Lys Ser Leu Arg Leu
 305 310 315 320
 Thr Gly Gln Ile Ser Leu Asp Phe Ile Gln Ala Glu Asp Gly Thr Ala
 325 330 335
 Tyr Ala Ile Glu Cys Asn Pro Arg Thr His Ser Ala Ile Thr Met Phe
 340 345 350
 Tyr Asn His Pro Gly Val Ala Glu Ala Tyr Leu Gly Lys Thr Pro Leu
 355 360 365
 Ala Ala Pro Leu Glu Pro Leu Ala Asp Ser Lys Pro Thr Tyr Trp Ile
 370 375 380
 Tyr His Glu Ile Trp Arg Leu Thr Gly Ile Arg Ser Gly Gln Gln Leu
 385 390 395 400
 Gln Thr Trp Phe Gly Arg Leu Val Arg Gly Thr Asp Ala Ile Tyr Arg
 405 410 415
 Leu Asp Asp Pro Ile Pro Phe Leu Thr Leu His His Trp Gln Ile Thr

420 425 430
 Leu Leu Leu Leu Gln Asn Leu Gln Arg Leu Lys Gly Trp Val Lys Ile
 435 440 445
 Asp Phe Asn Ile Gly Lys Leu Val Glu Leu Gly Gly Asp
 450 455 460

 <210> 122
 <211> 348
 <212> PRT
 <213> Nostoc punctiforme ATCC 29133

 <400> 122
 Met Pro Val Leu Asn Ile Leu His Leu Val Gly Ser Ala His Asp Lys
 1 5 10 15
 Phe Tyr Cys Asp Leu Ser Arg Leu Tyr Ala Gln Asp Cys Leu Ala Ala
 20 25 30
 Thr Ala Asp Pro Ser Leu Tyr Asn Phe Gln Ile Ala Tyr Ile Thr Pro
 35 40 45
 Asp Arg Gln Trp Arg Phe Pro Asp Ser Leu Ser Arg Glu Asp Ile Ala
 50 55 60
 Leu Thr Lys Pro Ile Pro Val Phe Asp Ala Ile Gln Phe Leu Thr Gly
 65 70 75 80
 Gln Asn Ile Asp Met Met Leu Pro Gln Met Phe Cys Ile Pro Gly Met
 85 90 95
 Thr Gln Tyr Arg Ala Leu Phe Asp Leu Leu Lys Ile Pro Tyr Ile Gly
 100 105 110
 Asn Thr Pro Asp Ile Met Ala Ile Ala Ala His Lys Ala Arg Ala Lys
 115 120 125
 Ala Ile Val Glu Ala Ala Gly Val Lys Val Pro Arg Gly Glu Leu Leu
 130 135 140
 Arg Gln Gly Asp Ile Pro Thr Ile Thr Pro Pro Ala Val Val Lys Pro
 145 150 155 160
 Val Ser Ser Asp Asn Ser Leu Gly Val Val Leu Val Lys Asp Val Thr
 165 170 175
 Glu Tyr Asp Ala Ala Leu Lys Lys Ala Phe Glu Tyr Ala Ser Glu Val
 180 185 190
 Ile Val Glu Ala Phe Ile Glu Leu Gly Arg Glu Val Arg Cys Gly Ile
 195 200 205
 Ile Val Lys Asp Gly Glu Leu Ile Gly Leu Pro Leu Glu Glu Tyr Leu
 210 215 220
 Val Asp Pro His Asp Lys Pro Ile Arg Asn Tyr Ala Asp Lys Leu Gln
 225 230 235 240
 Gln Thr Asp Asp Gly Asp Leu His Leu Thr Ala Lys Asp Asn Ile Lys
 245 250 255
 Ala Trp Ile Leu Asp Pro Asn Asp Pro Ile Thr Gln Lys Val Gln Gln

260 265 270
 Val Ala Lys Arg Cys His Gln Ala Leu Gly Cys Arg His Tyr Ser Leu
 275 280 285
 Phe Asp Phe Arg Ile Asp Pro Lys Gly Gln Pro Trp Phe Leu Glu Ala
 290 295 300
 Gly Leu Tyr Cys Ser Phe Ala Pro Lys Ser Val Ile Ser Ser Met Ala
 305 310 315 320
 Lys Ala Ala Gly Ile Pro Leu Asn Asp Leu Leu Ile Thr Ala Ile Asn
 325 330 335
 Glu Thr Leu Gly Ser Asn Lys Lys Val Leu Gln Asn
 340 345

 <210> 123
 <211> 680
 <212> PRT
 <213> *Saccharomyces cerevisiae* CEN.PK-1D

 <400> 123
 Met Thr Gln Phe Thr Asp Ile Asp Lys Leu Ala Val Ser Thr Ile Arg
 1 5 10 15
 Ile Leu Ala Val Asp Thr Val Ser Lys Ala Asn Ser Gly His Pro Gly
 20 25 30
 Ala Pro Leu Gly Met Ala Pro Ala Ala His Val Leu Trp Ser Gln Met
 35 40 45
 Arg Met Asn Pro Thr Asn Pro Asp Trp Ile Asn Arg Asp Arg Phe Val
 50 55 60
 Leu Ser Asn Gly His Ala Val Ala Leu Leu Tyr Ser Met Leu His Leu
 65 70 75 80
 Thr Gly Tyr Asp Leu Ser Ile Glu Asp Leu Lys Gln Phe Arg Gln Leu
 85 90 95
 Gly Ser Arg Thr Pro Gly His Pro Glu Phe Glu Leu Pro Gly Val Glu
 100 105 110
 Val Thr Thr Gly Pro Leu Gly Gln Gly Ile Ser Asn Ala Val Gly Met
 115 120 125
 Ala Met Ala Gln Ala Asn Leu Ala Ala Thr Tyr Asn Lys Pro Gly Phe
 130 135 140
 Thr Leu Ser Asp Asn Tyr Thr Tyr Val Phe Leu Gly Asp Gly Cys Leu
 145 150 155 160
 Gln Glu Gly Ile Ser Ser Glu Ala Ser Ser Leu Ala Gly His Leu Lys
 165 170 175
 Leu Gly Asn Leu Ile Ala Ile Tyr Asp Asp Asn Lys Ile Thr Ile Asp
 180 185 190
 Gly Ala Thr Ser Ile Ser Phe Asp Glu Asp Val Ala Lys Arg Tyr Glu
 195 200 205
 Ala Tyr Gly Trp Glu Val Leu Tyr Val Glu Asn Gly Asn Glu Asp Leu

210					215					220				
Ala Gly Ile Ala Lys	Ala Ile Ala Gln Ala Lys	Leu Ser Lys Asp Lys	225	230	235	240								
Pro Thr Leu Ile Lys	Met Thr Thr Thr Ile Gly Tyr Gly Ser Leu His		245	250	255									
Ala Gly Ser His Ser Val His Gly	Ala Pro Leu Lys Ala Asp Asp Val		260	265	270									
Lys Gln Leu Lys Ser Lys Phe Gly	Phe Asn Pro Asp Lys Ser Phe Val		275	280	285									
Val Pro Gln Glu Val Tyr Asp His Tyr Gln Lys Thr Ile Leu Lys Pro			290	295	300									
Gly Val Glu Ala Asn Asn Lys Trp Asn Lys Leu Phe Ser Glu Tyr Gln			305	310	315									
Lys Lys Phe Pro Glu Leu Gly Ala Glu Leu Ala Arg Arg Leu Ser Gly			325	330	335									
Gln Leu Pro Ala Asn Trp Glu Ser Lys Leu Pro Thr Tyr Thr Ala Lys			340	345	350									
Asp Ser Ala Val Ala Thr Arg Lys Leu Ser Glu Thr Val Leu Glu Asp			355	360	365									
Val Tyr Asn Gln Leu Pro Glu Leu Ile Gly Gly Ser Ala Asp Leu Thr			370	375	380									
Pro Ser Asn Leu Thr Arg Trp Lys Glu Ala Leu Asp Phe Gln Pro Pro			385	390	395									
Ser Ser Gly Ser Gly Asn Tyr Ser Gly Arg Tyr Ile Arg Tyr Gly Ile			405	410	415									
Arg Glu His Ala Met Gly Ala Ile Met Asn Gly Ile Ser Ala Phe Gly			420	425	430									
Ala Asn Tyr Lys Pro Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr			435	440	445									
Ala Ala Gly Ala Val Arg Leu Ser Ala Leu Ser Gly His Pro Val Ile			450	455	460									
Trp Val Ala Thr His Asp Ser Ile Gly Val Gly Glu Asp Gly Pro Thr			465	470	475									
His Gln Pro Ile Glu Thr Leu Ala His Phe Arg Ser Leu Pro Asn Ile			485	490	495									
Gln Val Trp Arg Pro Ala Asp Gly Asn Glu Val Ser Ala Ala Tyr Lys			500	505	510									
Asn Ser Leu Glu Ser Lys His Thr Pro Ser Ile Ile Ala Leu Ser Arg			515	520	525									
Gln Asn Leu Pro Gln Leu Glu Gly Ser Ser Ile Glu Ser Ala Ser Lys			530	535	540									
Gly Gly Tyr Val Leu Gln Asp Val Ala Asn Pro Asp Ile Ile Leu Val			545	550	555									

Ala Thr Gly Ser Glu Val Ser Leu Ser Val Glu Ala Ala Lys Thr Leu
565 570 575

Ala Ala Lys Asn Ile Lys Ala Arg Val Val Ser Leu Pro Asp Phe Phe
580 585 590

Thr Phe Asp Lys Gln Pro Leu Glu Tyr Arg Leu Ser Val Leu Pro Asp
595 600 605

Asn Val Pro Ile Met Ser Val Glu Val Leu Ala Thr Thr Cys Trp Gly
610 615 620

Lys Tyr Ala His Gln Ser Phe Gly Ile Asp Arg Phe Gly Ala Ser Gly
625 630 635 640

Lys Ala Pro Glu Val Phe Lys Phe Phe Gly Phe Thr Pro Glu Gly Val
645 650 655

Ala Glu Arg Ala Gln Lys Thr Ile Ala Phe Tyr Lys Gly Asp Lys Leu
660 665 670

Ile Ser Pro Leu Lys Lys Ala Phe
675 680

<210> 124

<211> 370

<212> PRT

<213> *Saccharomyces cerevisiae* CEN.PK-1D

<400> 124

Met Ser Glu Ser Pro Met Phe Ala Ala Asn Gly Met Pro Lys Val Asn
1 5 10 15

Gln Gly Ala Glu Glu Asp Val Arg Ile Leu Gly Tyr Asp Pro Leu Ala
20 25 30

Ser Pro Ala Leu Leu Gln Val Gln Ile Pro Ala Thr Pro Thr Ser Leu
35 40 45

Glu Thr Ala Lys Arg Gly Arg Arg Glu Ala Ile Asp Ile Ile Thr Gly
50 55 60

Lys Asp Asp Arg Val Leu Val Ile Val Gly Pro Cys Ser Ile His Asp
65 70 75 80

Leu Glu Ala Ala Gln Glu Tyr Ala Leu Arg Leu Lys Lys Leu Ser Asp
85 90 95

Glu Leu Lys Gly Asp Leu Ser Ile Ile Met Arg Ala Tyr Leu Glu Lys
100 105 110

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

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

Phe Val Asn Leu Thr Asn Ile Gly Leu Pro Ile Gly Ser Glu Met Leu
145 150 155 160

Asp Thr Ile Ser Pro Gln Tyr Leu Ala Asp Leu Val Ser Phe Gly Ala
165 170 175

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