



(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)



1-0035109

(51)⁷ C12P 7/26; C12P 7/66; C12P 17/06

(13) B

(21) 1-2018-00067

(22) 10/06/2016

(86) PCT/EP2016/063242 10/06/2016

(87) WO 2016/198564 15/12/2016

(30) 15171375.7 10/06/2015 EP

(45) 27/03/2023 420

(43) 25/07/2018 364A

(73) 1. DANMARKS TEKNISKE UNIVERSITET (DK)

Anker Engelunds Vej 1, Bygning 101A, 2800 Kgs Lyngby, Denmark

2. KØBENHAVNS UNIVERSITET (DK)

Nørregade 10, 1165 København K, Denmark

3. Oterra A/S (DK)

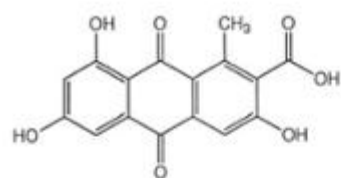
Agern Alle 24, DK-2970 Hørsholm, Denmark

(72) MØLLER, Birger Lindberg (DK); MADSEN, Bjørn (DK); STÆRK, Dan (DK); OKKELS, Finn Thyge (DK); ANDERSEN-RANBERG, Johan (DK); KONGSTAD, Kenneth Thermann (DK); BINDERUP, Kim (DK); BENNEDSEN, Mads (DK); NAFISI, Majse (DK); KHORSAND-JAMAL, Paiman (DK); KANNANGARA, Rubini Maya (DK); MORTENSEN, Uffe Hasbro (DK); THRANE, Ulf (DK); FRANDSEN, Rasmus John Normand (DK).

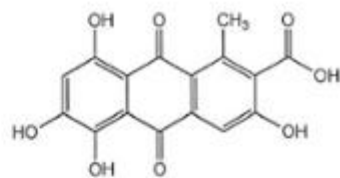
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) PHƯƠNG PHÁP SẢN XUẤT HỢP CHẤT THƠM THU ĐƯỢC TỪ OCTAKETIT

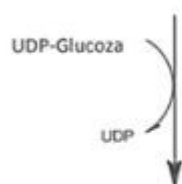
(57) Sáng chế đề cập đến phương pháp sản xuất hợp chất thơm thu được từ octaketit cần quan tâm (ví dụ, axit carminic), trong đó phương pháp này bao gồm (I): biểu hiện khác loại gen polyketit syntaza (polyketide synthase: PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (octaketide synthase: OKS) để thu được octaketit không khử *in vivo* trong tế bào chủ tái tổ hợp và (II): chuyển hóa *in vivo* octaketit không khử ở bước (I) thành hợp chất thơm C₁₄-C₃₄ cần quan tâm (ví dụ, axit carminic).



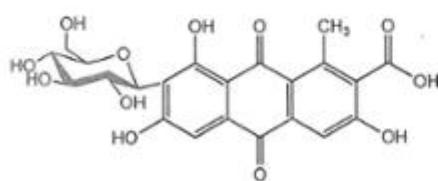
Axit flavokermesic



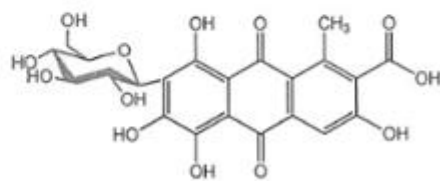
Axit kermesic



DcUGT2



DCII (Axit flavokermesic-glucoside)



Axit carminic

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất hợp chất thơm thu được từ octaketit cần quan tâm (ví dụ, axit carminic), trong đó phương pháp này bao gồm (I): biểu hiện khác loại gen polyketit syntaza (polyketide synthase: PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (octaketide synthase: OKS) để thu được octaketit không khử *in vivo* trong tế bào chủ tái tổ hợp và (II): chuyển hóa *in vivo* octaketit không khử ở bước (I) thành hợp chất thơm C₁₄-C₃₄ cần quan tâm (ví dụ, axit carminic).

Tình trạng kỹ thuật của sáng chế

Axit carminic màu tự nhiên là một trong các chất tạo màu được sử dụng thường xuyên nhất trong thực phẩm, y học, mỹ phẩm và sản phẩm dệt.

Axit carminic là chất tạo màu, mà có thể được chiết từ cơ thể côn trùng cái của *Dactylopius coccus costa* (tên khác *Coccus cacti* L.). Các côn trùng này sống trên *Nopalea coccinellifera*, *Opuntia fidus indica* và các cây khác của họ Cactaceae được trồng, ví dụ, ở các vùng sa mạc của Mexico, Trung và Nam Mỹ và quần đảo Canary. Tùy theo độ pH, chất tạo màu này có thể là màu trong phổ từ màu da cam phía trên màu đỏ đến màu tím và thường được biết là cánh kiến đỏ hoặc màu cánh kiến đỏ. Chất tạo màu đỏ Carmine được sử dụng rộng rãi trong thực phẩm và đồ uống.

Như đã biết trong lĩnh vực, *Porphyrophora polonica* cũng sản xuất axit carminic và được trồng để sản xuất axit carminic, ví dụ, ở Ba lan.

Trong sản xuất công nghiệp có liên quan hiện nay, axit carminic được thu hoạch bằng cách chiết từ xác khô của côn trùng bằng nước hoặc cồn.

Để cố gắng giải quyết vấn đề về các thay đổi không mong muốn và dao động giá - US5424421 (European Colour, công bố năm 1995) mô tả quy trình tổng hợp hóa học axit carminic bằng con đường tổng hợp bao gồm các hợp chất trung gian khác nhau.

Như đã thảo luận, ví dụ, trong WO2006/056585A1 (Chr. Hansen A/S), trong quá trình chiết axit carminic từ côn trùng dựa trên nước, lượng protein côn trùng cũng được giải phóng từ côn trùng và sẽ nằm trong dịch chiết màu và đã được báo cáo rằng protein côn trùng cánh kiến đỏ có thể tạo ra một số vấn đề liên quan đến dị ứng. Trong

WO2006/056585A1 quy trình đặc biệt để làm giảm lượng protein côn trùng từ dịch chiết côn trùng được mô tả. Tuy nhiên, chế phẩm/sản phẩm màu thành phẩm của WO2006/056585A1 vẫn sẽ chứa một số lượng protein của côn trùng *Dactylopius coccus costa*.

Cấu trúc của axit carminic được thể hiện trên Fig.1 trong bản mô tả này. Như có thể thấy từ hình vẽ, cấu trúc này được gọi C-glucosit (tức là trong đó glucoza được nối/liên hợp với aglucon bằng liên kết cacbon-cacbon).

Như được thể hiện trên Fig.1 trong bản mô tả này, sự thủy phân axit C-glucosit carminic có thể tạo ra glucoza và axit aglucon kermesic (KA).

Con đường sinh tổng hợp *in vivo* của axit carminic ở côn trùng (*Dactylopius coccus*) hiện vẫn chưa được mô tả chi tiết. Theo đó, dựa trên giải pháp đã biết, người có hiểu biết trung bình trong lĩnh vực kỹ thuật không biết được hợp chất nào là aglucon trong quá trình sản xuất sinh tổng hợp *in vivo* của axit carminic ở *Dactylopius coccus*.

Việc phân tích *Dactylopius coccus* đã chỉ ra rằng có rất nhiều loại hợp chất liên quan đến axit carminic có mặt trong dịch chiết từ *Dactylopius coccus* và nhiều trong số các hợp chất này nói chung có thể được glucosyl hóa trong quá trình sản xuất sinh tổng hợp *in vivo* của axit carminic.

Ví dụ, trong tài liệu của Stathopoulou et al. (Analytica Chimica Acta 804 (2013) 264– 272) sáu anthraquinon mới được mô tả trong dịch chiết từ *D. coccus* và bất kỳ trong số sáu anthraquinon mới này (xem, ví dụ, fig.1 của tài liệu) nói chung có thể là phân tử mà được glucosyl hóa trong quá trình sản xuất sinh tổng hợp *in vivo* của axit carminic ở *Dactylopius coccus*.

Ngoài ra, như đã biết trong lĩnh vực kỹ thuật này hợp chất được glucosyl hóa bậc một đã tạo thành trong quá trình sản xuất sinh tổng hợp *in vivo* của sản phẩm cuối glucosit có thể là hợp chất trung gian không ổn định mà sẽ không xác định được trong dịch chiết được phân lập từ *Dactylopius coccus*, ví dụ, như được phân tích ở tài liệu nêu trên của Stathopoulou et al.

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, dựa trên tài liệu kỹ thuật đã biết, có thể suy đoán rằng hợp chất được glucosyl hóa bậc một thích hợp trong quá trình sản xuất sinh tổng hợp *in vivo* của axit carminic ở

Dactylopius coccus có thể, ví dụ, là hợp chất polyketit trung gian không ổn định với khoảng cùng số lượng nguyên tử cacbon với, ví dụ, axit flavokermesic.

ADN theo sáng chế hoặc trình tự axit amin của glycosyltransferaza tham gia vào con đường sinh tổng hợp ở côn trùng (*Dactylopius coccus*) *in vivo* của axit carminic không được mô tả rõ ràng trong tài liệu kỹ thuật đã biết.

Như đã biết trong lĩnh vực, đối với côn trùng tích lũy hóa chất khối lượng phân tử thấp, các gen của con đường sinh tổng hợp thích hợp đôi khi không có mặt ở bộ gen của côn trùng này.

Ví dụ, một số côn trùng hấp thụ glycosit từ cây chúng ăn – ví dụ xem tài liệu của Zagrobelny et al (Cyanogenic glucosides and plant-insect interactions; Phytochemistry. 2004 Feb;65(3):293-306) hoặc tài liệu của Geuder et al (Journal of Chemical Ecology, Vol. 23, No. 5, 1997).

Côn trùng *Dactylopius coccus* ăn cây xương rồng và có thể là côn trùng *Dactylopius coccus* (như các côn trùng khác) hấp thụ glycosit thích hợp từ cây xương rồng chúng ăn vào.

Như đã biết trong lĩnh vực, đối với côn trùng tích lũy glycosit khối lượng phân tử thấp, gen của con đường sinh tổng hợp thích hợp đôi khi được tìm thấy ở các vi sinh vật sống trong côn trùng, ví dụ xem tài liệu của Genta et al. (Potential role for gut microbiota in cell wall digestion and glucoside detoxification in *Tenebrio molitor* larvae), Journal of Insect Physiology 52 (2006) 593–601.

Do đó, dựa vào tài liệu kỹ thuật đã biết người có hiểu biết trung bình trong lĩnh vực kỹ thuật không thể biết được liệu bộ gen của *Dactylopius coccus* có thật sự chứa gen mã hóa glycosyltransferaza tham gia vào con đường sinh tổng hợp *in vivo* dẫn tới axit carminic hay không.

Các polyketit được tổng hợp bởi nhóm các enzym mà thường được gọi là polyketit syntaza (polyketide synthase: PKS). Tất cả các PKS cùng có khả năng xúc tác cho phản ứng dung hợp dựa trên sự ngưng tụ Claisen của các nhóm axyl bằng cách tạo liên kết cacbon-cacbon có giải phóng cacbon dioxit. Phản ứng này được xúc tác bởi miền beta-ketosyntaza (KS). Ngoài miền/vị trí hoạt động này, quá trình tổng hợp cũng có thể phụ thuộc vào, nhưng không loại trừ, hoạt động của Axyl-Chất mang-Protein (ACP), Axyl-transferaza (AT), Chất khởi động-Axyl-Transferaza (SAT), khuôn sản phẩm (Product template: PT), ThioEsteraza (TE), Yếu tố chiều dài mạch

(CLF, còn được biết là KS β), Claisen XYClaza (CYC), Ketoreductaza (KR), dehydrataza (DH), enoyl reductaza (ER) và C-metyl transferaza (Cmet). Các cơ chất để tổng hợp polyketit thường được phân loại thành các đơn vị khởi động và kéo dài, trong đó đơn vị khởi động, bao gồm nhưng không giới hạn ở axetyl-CoA là đơn vị được bổ sung đầu tiên của mạch polyketit phát triển; và các đơn vị kéo dài, ví dụ nhưng không loại trừ malonyl-CoA, tất cả đều là các đơn vị cacbon-cacbon được bổ sung sau đó.

Ở mức trình tự bậc một (trình tự axit amin), mức cấu trúc bậc hai (gấp nếp cục bộ), mức cấu trúc bậc ba (gấp nếp toàn bộ) và mức cấu trúc bậc bốn (tương tác protein-protein), các PKS thể hiện tính đa dạng rất lớn, và do đó được chia nhỏ thành các kiểu khác nhau.

Các hệ PKS typ I thường được tìm thấy ở nấm sợi và vi khuẩn, trong đó chúng chịu trách nhiệm về sự tạo thành polyketit thơm, thơm đa vòng và được khử.

Các thành viên của PKS typ I có một vài vị trí hoạt động trên cùng một mạch polypeptit và mỗi enzym có thể xúc tác cho phản ứng ngưng tụ được lặp lại của các đơn vị hai cacbon. Tập hợp tối thiểu các miền ở PKS typ I bao gồm KS, AT và ACP. PKS typ I còn được chia nhỏ thành PKS môđun và PKS lắp, trong đó PKS lắp chỉ có một bản sao duy nhất của mỗi kiểu vị trí hoạt động và tái sử dụng chúng lặp lại đến khi mạch polyketit phát triển đạt đến chiều dài định trước. PKS lắp typ I mà tạo thành hợp chất thơm và thơm đa vòng thường dựa vào miền PT và CYC để gấp nếp trực tiếp mạch polyketit không được khử đã tạo thành. PKS môđun chứa một vài bản sao của cùng các vị trí hoạt động, chúng được tổ chức thành các trình tự lặp lại của các vị trí hoạt động mà được gọi là môđun, mỗi môđun chịu trách nhiệm cho việc bổ sung và biến đổi đơn vị ketit đơn. Mỗi vị trí hoạt động ở từng môđun chỉ được sử dụng một lần trong quá trình tổng hợp polyketit đơn. PKS lắp typ I thường được tìm thấy ở nấm, trong khi PKS môđun typ I thường được tìm thấy ở vi khuẩn.

Các hệ PKS typ II chịu trách nhiệm tạo thành hợp chất thơm và thơm đa vòng ở vi khuẩn.

PKS typ II là các phức hợp protein trong đó các enzym đơn lẻ tương tác để tạo thành enzym PKS chức năng. Các enzym đơn lẻ bao gồm các hoạt tính cho KS, CLF và ACP.

Kiểu PKS này được đặc trưng bởi việc tạo bởi nhiều enzym khác nhau mà tạo thành phức hợp protein, mà được gọi chung là PKS hoạt động. PKS typ II tạo thành polyketit không được khử mà gấp nếp tự phát thành hợp chất thơm/vòng phức tạp. Tuy nhiên, ở các hệ vi khuẩn sự gấp nếp của khung polyketit thường được trợ giúp/định hướng bởi các lớp enzym khác nhau, hoạt động in trans (không phụ thuộc vào enzym PKS) để thúc đẩy gấp nếp không tự phát. Các lớp enzym liên quan được gọi là aromataza và xyclaza. Quá trình sinh tổng hợp của một hợp chất thơm đa vòng trong các hệ này thường bao gồm hoạt động liên tục của nhiều aromataza/xyclaza khác nhau. Các aromataza và xyclaza có thể được chia thành hai nhóm dựa trên kiểu cơ chất mà chúng hoạt động trên đó: trong đó nhóm đầu tiên chỉ hoạt động trên mạch polyketit thẳng và xúc tác tạo nhóm thơm/vòng đầu tiên, nhóm enzym thứ hai chỉ chấp nhận các cơ chất bao gồm nhóm thơm hoặc vòng (= sản phẩm từ nhóm đầu tiên của aromataza/xyclaza). Đã chứng minh được là không thể biểu hiện chức năng các hệ PKS typ II ở vật chủ sản xuất thích hợp (*E. coli*, *Bacillus*, nấm men), có thể do đây là các phức hợp nhiều enzym đòi hỏi mức biểu hiện được cân bằng, và có thể dựa vào các yếu tố chưa biết.

PKS typ III thường chỉ chứa miền KS, mà trong tài liệu có thể, ví dụ, được gọi là miền KASIII hoặc Chalcone syntaza mà hoạt động không phụ thuộc vào miền ACP. PKS typ III từ vi khuẩn, thực vật và nấm đã được mô tả.

PKS typ III đã được biết từ lâu ở thực vật, tại đó chúng chịu trách nhiệm tạo thành các hợp chất như flavonoid (chất màu/chất chống oxy hóa) và stilben, được tìm thấy ở nhiều loài thực vật khác nhau. Các sản phẩm của PKS typ III thường gấp nếp tự phát thành hợp chất thơm/vòng phức tạp.

Tài liệu của Yu et al. (2012) xem xét lại về Polyketit syntaza typ III trong quá trình sinh tổng hợp sản phẩm tự nhiên. Tài liệu của Yu et al. (2012) bộc lộ: “PKS typ III là các enzym tự chủ mà tạo thành homodime. Vị trí hoạt động duy nhất của chúng trong mỗi monome xúc tác phản ứng mở, mở rộng, và tạo vòng lặp lại để tạo thành sản phẩm polyketit. Mặc dù cấu trúc của chúng đơn giản, PKS typ III sản xuất mảng rộng các hợp chất như chalcon, pyron, acridon, phloroglucinol, stilben, và lipid resorcinolic. Trong những năm gần đây, PKS typ III đã thu hút nhiều sự chú ý do các sản phẩm đa dạng của chúng, sự phân bố rộng, các cấu trúc tương đối đơn giản, và khả năng thao tác di truyền dễ dàng. Trong tài liệu này, chúng tôi sẽ thảo luận một cách có

hệ thống PKS typ III từ thực vật, vi khuẩn, và nấm cũng như tiến bộ gần đây trong nghiên cứu PKS typ III.”

Nói ngắn gọn, dựa vào tài liệu kỹ thuật đã biết, người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ biết PKS cụ thể cần quan tâm là PKS typ I, typ II hay typ III.

Ngoài việc phân loại dựa vào cấu trúc và chức năng protein của các hệ thống PKS, phân loại theo cách khác là dựa vào mức biến đổi được tìm thấy ở sản phẩm polyketit cuối cùng. Lưu ý rằng các biến đổi này có thể được đưa vào bởi chính PKS hoặc bởi các enzym sau khi hoạt động. Trong sơ đồ phân loại này, các sản phẩm được chia thành hai nhóm: (I) polyketit không khử và (II) polyketit khử. Loại không được khử được đặc trưng bởi sự có mặt của nhóm keton trong ketit ($-\text{CH}_2\text{-CO}-$), có nguồn gốc từ các đơn vị khởi động hoặc kéo dài, dưới dạng keton hoặc ở dạng liên kết đôi trong nhóm thơm. Trong polyketit khử, keton đơn hoặc tất cả các keton được khử thành nhóm rượu ($-\text{CH}_2\text{-CHOH}-$) bằng miền/enzym KR, hoặc tiếp tục thành nhóm alken ($-\text{C}=\text{C}-$) bằng miền/enzym DH, hoặc thậm chí còn tiếp tục thành nhóm alkan ($-\text{CH}_2\text{-CH}_2-$) bằng miền/enzym ER. Dựa vào các tính chất hóa học này của sản phẩm tạo thành, các PKS liên quan được phân loại là PKS không khử hoặc PKS khử.

Việc gấp nếp mạch polyketit tạo thành vào các cấu trúc phức tạp với các môtip vòng thường là quy trình được xúc tác và hướng dẫn bởi enzym sau PKS. Các enzym chịu trách nhiệm thuộc một số họ enzym khác nhau, thường là các aromataza và/hoặc xyclaza. PKS lặp typ I ở nấm là đặc biệt do sắp đặt miền PT chịu trách nhiệm tạo vòng thơm trong khi miền CYC chịu trách nhiệm về giải phóng sản phẩm kết hợp với tạo vòng thơm. Aromataza và xyclaza hoạt động trên polyketit đã được mô tả từ các hệ vi khuẩn và thực vật. Ngoài ra, một số ví dụ đã có, trong đó việc gấp nếp polyketit là quy trình tự phát, ví dụ flavonoit ở thực vật.

PKS đã được phân lập và được phân tích về chức năng từ vi khuẩn, nấm và thực vật. Tuy nhiên, chưa có PKS có nguồn gốc động vật nào được mô tả, và quy trình tổng hợp polyketit ở côn trùng trong một vài trường hợp có liên quan đến hoạt động trao đổi chất của vi khuẩn nội cộng sinh.

Tài liệu của Tang, Y. et al. (2004) mô tả rằng sự biểu hiện ở chủng vi khuẩn *Streptomyces coelicolor* CH999, chứa đột biến mất đoạn nhiễm sắc thể ảnh hưởng đến toàn bộ cụm gen Act chịu trách nhiệm về sự sinh tổng hợp actinorhodin. PKS nhỏ (*Act*

PKS = Act_KS, Act_CLF và Act_ACP), thuộc PKS typ II, sản sinh axit flavokermesic (FK) (được gọi là TMAC ở tài liệu về vi khuẩn) khi kết hợp với sự biểu hiện khác loại của ZhuI aromataza/xyclaza và ZhuJ xyclaza từ cụm gen *zhu* trong *Streptomyces* sp. R1128.

Do đó, tài liệu này mô tả việc đưa tái tổ hợp gen *Streptomyces* PKS vào tế bào chủ *Streptomyces*, vì vậy PKS không thuộc giống khác với tế bào chủ.

Fig.2 trong bản mô tả này là fig.5A đã được thể hiện của tài liệu Tang, Y. et al (2004). Như có thể được thấy trên hình vẽ này và như được mô tả thêm trong tài liệu, Act PKS (được gọi là octaketit syntaza (OKS)) tạo ra octaketit không khử và octaketit này, nhờ ZhuI aromataza/xyclaza và ZhuJ xyclaza, được chuyển hóa thành axit flavokermesic (FK) (được gọi là TMAC). Các hợp chất SEK4 và SEK4B cũng được sản xuất tự phát (cấu trúc được thể hiện trên Fig.2 có thể được gọi ở đây là các sản phẩm song song).

Trong cây *Aloe arborescens*, các PKS xác định đã được chỉ ra là sản xuất các polyketit có các chiều dài khác nhau bao gồm octaketit, ví dụ xem Mizuuchi et al (2009) trong đó trên fig.1 được minh họa rằng octaketit syntaza (OKS) được gọi là PKS4 và PKS5 có thể, bằng cách sử dụng malonyl-CoA làm các đơn vị kéo dài, tạo ra octaketit không khử. Các hợp chất song song SEK4 và SEK4B cũng được tạo thành tự phát.

Cây *Hypericum perforatum* (St. John's wort) còn chứa octaketit syntaza, ví dụ xem Karppinen et al (2008), trong đó mô tả rằng PKS gọi là HpPKS2 được biểu hiện ở *E. coli*, tiếp theo là tinh chế và phân tích hóa sinh *in vitro* enzym này. Fig.1 của tài liệu này minh họa rằng PKS gọi là HpPKS2 tạo ra octaketit không khử (sử dụng axetyl-CoA làm đơn vị khởi động và malonyl-CoA làm các đơn vị kéo dài) và sản phẩm song song SEK4 và SEK4B được tạo thành tự phát.

Tài liệu của Yu et al. (2012) xem xét lại về Polyketit syntaza typ III trong sinh tổng hợp sản phẩm tự nhiên; tài liệu đọc trên trang 293: “Các PKS typ III khác nhau đã được thao tác di truyền vào *E. coli* để tạo ra polyketit mới. Việc sản xuất curcuminoid đặc hiệu ở thực vật đã được tái tạo trong *E. coli* bằng cách đồng biểu hiện CUS với phenylalanin ammonia-lyaza từ *Rhodotorula rubra* và 4-coumarat:CoA ligaza (4CL) từ *Lithospermum erythrorhizon*”. Như đã giải thích trong tài liệu, PKS

được gọi là “CUS” tổng hợp diketit-CoA và do đó CUS không phải là octaketit syntaza.

Tài liệu *Jadhav et al* (2014) mô tả rằng PKS hexaketit typ III từ *Plumbago zeylanica* (PzPKS) được tách dòng và biểu hiện ở các cây thuốc lá để nghiên cứu liệu các cây thuốc lá chuyển gen biểu hiện PzPKS có tổng hợp được polyketit quan trọng về mặt dược phẩm, plumbagin hay không.

Trong các tài liệu liên quan đến PKS nêu trên không có tài liệu nào thảo luận về sản xuất axit carminic.

Không bị giới hạn bởi lý thuyết, tin là giải pháp đã biết không mô tả rằng các PKS octaketit syntaza (OKS) typ III theo sáng chế có thể hoạt động *in vivo* ở tế bào chủ sản xuất khác loại thuộc giống khác, ví dụ OKS typ III thực vật có thể tạo ra octaketit không khử *in vivo* ở tế bào chủ sản xuất khác loại, ví dụ như tế bào chủ sản xuất tái tổ hợp *Aspergillus*.

Đơn yêu cầu cấp bằng sáng chế PCT/EP2014/078540 đã được nộp ngày 18 tháng 12 năm 2014. Vào ngày nộp đơn/ưu tiên của đơn yêu cầu cấp bằng sáng chế PCT/EP2014/078540 không được công bố. Tài liệu này mô tả glycosyltransferaza (GT) được phân lập từ côn trùng *Dactylopius coccus costa* mà có khả năng: (I): liên hợp glucoza với axit flavokermesic (FK); và/hoặc (II): liên hợp glucoza với axit kermesic (KA) và sử dụng GT này, ví dụ, để tạo axit carminic.

PCT/EP2014/078540 không mô tả trực tiếp và rõ ràng các octaketit không khử và/hoặc polyketit syntaza (PKS) được thảo luận ở đây.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất con đường sinh tổng hợp thích hợp có thể dẫn đến axit carminic.

Ưu điểm của việc đề xuất con đường sinh tổng hợp thích hợp như được mô tả trong bản mô tả này là nó mở ra khả năng sản xuất khác loại (ví dụ, ở *Aspergillus* hoặc nấm men) axit carminic mà không cần tạo dịch chiết từ côn trùng và do đó có thể tạo ra chế phẩm/sản phẩm màu axit carminic không chứa protein từ côn trùng *Dactylopius coccus costa* không mong muốn.

Một phần của giải pháp này đề cập đến việc các tác giả sáng chế đã xác định dịch chiết *Dactylopius coccus* (bao gồm dịch chiết của sinh vật nội cộng sinh có mặt trong *Dactylopius coccus*) có hoạt tính glycosyltransferaza GT theo sáng chế. Như đã

thảo luận ở đây, các tác giả sáng chế phân tích GT và xác định rằng nó có khả năng: (I): liên hợp glucoza với axit flavokermesic (FK); và/hoặc (II): liên hợp glucoza với axit kermesic (KA). Do đó, GT này có thể được sử dụng, ví dụ, để tạo ra axit carminic.

Trình tự polynucleotit mã hóa *Dactylopius coccus costa* glycosyltransferaza mới được phân lập/tách dòng được mô tả ở đây được thể hiện trong SEQ ID NO: 1 ở đây và trình tự axit amin polypeptit được thể hiện trong SEQ ID NO: 2 ở đây.

Enzym glycosyltransferaza theo sáng chế của SEQ ID NO: 2 có thể được gọi là “DcUGT2” hoặc đơn giản là “DcUGT” trong bản mô tả này.

Fig.1 biểu diễn bằng sơ đồ hoạt tính glycosyltransferaza theo sáng chế của DcUGT glycosyltransferaza được phân lập/tách dòng được mô tả ở đây của SEQ ID NO:2, như được minh họa trên hình vẽ này, enzym này đã được phát hiện là có thể liên hợp glucoza với aglycon axit flavokermesic (FK) và axit kermesic (KA).

Dựa vào tài liệu kỹ thuật đã biết, người có hiểu biết trung bình trong lĩnh vực kỹ thuật không biết được hợp chất nào là hợp chất được glucosyl hóa bậc một trong quá trình sản xuất sinh tổng hợp axit carminic *in vivo* ở *Dactylopius coccus*.

Các tác giả sáng chế chứng minh rằng *Dactylopius coccus* chứa GT có thể C-glycosyl hóa axit flavokermesic (FK) và/hoặc axit kermesic (KA). Do đó, và không bị giới hạn bởi lý thuyết, có thể tin rằng axit flavokermesic (FK) và/hoặc axit kermesic (KA) là các aglycon thích hợp, ví dụ, để sinh tổng hợp khác loại *in vivo*, ví dụ, axit carminic.

Dựa trên kiến thức này, có thể tin rằng axit flavokermesic (FK) và/hoặc axit kermesic (KA) là các aglycon thích hợp, ví dụ, để sinh tổng hợp khác loại *in vivo*, ví dụ, axit carminic, các tác giả sáng chế đã xác định được rằng các glycosyltransferaza theo sáng chế cũng có thể được xác định ở các cây *Aloe*, các cây *Haworthia* và cây *Sorghum* hoặc cây lúa.

Nhờ việc xác định được các aglycon có thể thích hợp, các tác giả sáng chế có thể quay trở lại và cố gắng xác định các hợp chất trung gian trước thích hợp mà có thể thích hợp để sinh tổng hợp *in vivo* các aglycon này (ví dụ FK/KA).

Như đã thảo luận trong ví dụ thử nghiệm 6 trong bản mô tả này, các tác giả sáng chế đã biểu hiện tái tổ hợp polyketit syntaza (PKS) octaketit syntaza (OKS) typ III thực vật (của *Aloe arborescens* (AaOKS) và/hoặc *Hypericum perforatum* (HpPKS2)) trong *Aspergillus nidulans* và xác định được rằng OKS thực sự hoạt động

trong vật chủ khác loại, có octaketit không khử được tạo ra *in vivo* trong tế bào *Aspergillus nidulans*. Các octaketit không khử là không ổn định và việc xác định rằng các octaketit không khử xuất hiện *in vivo* được chứng minh bằng sự tích tụ của các sản phẩm song song/thoái biến SEK4 và SEK4B (xem Fig.2 trong bản mô tả này và các giải pháp đã biết đã thảo luận trên đây + Ví dụ 6 trong bản mô tả này).

Như đã thảo luận trong phần kết luận của ví dụ thử nghiệm 6 trong bản mô tả này, sự biểu hiện của PKS typ III thực vật (*HpPKS2* hoặc *AaOKS*) dẫn đến sự sản sinh các hợp chất khác nhau bao gồm các hợp chất SEK4, SEK4B và axit flavokermesic (FK) *in vivo* ở *Aspergillus nidulans*. Vì ở ví dụ 6 này, xyclaza và/hoặc aromataza khác loại không được gắn xen vào chủng *Aspergillus* và hợp chất FK được nhận diện, nó chỉ ra rằng chủng *Aspergillus* có thể chứa xyclaza và/hoặc aromataza tương đồng có khả năng chuyển hóa octaketit không khử thành hợp chất FK *in vivo*.

Như đã thảo luận trong phần kết luận của ví dụ thử nghiệm 8 trong bản mô tả này, các tác giả sáng chế đã tạo ra sự đồng biểu hiện của PKS thực vật khác loại (*AaOKS*) và glycosyltransferaza thực vật khác loại (*DcUGT2*) ở *Aspergillus nidulans* và dẫn đến sự sản sinh *in vivo* axit carminic (CA) và DcII.

Theo công trình thực nghiệm của các tác giả sáng chế liên quan đến sự biểu hiện/sản xuất khác loại ở cây *Nicotiana benthamiana* (họ hàng gần của cây thuốc lá), như đã thảo luận trong phần kết luận của ví dụ thử nghiệm 11 trong bản mô tả này, các kết quả của ví dụ 11 này chứng minh rằng:

(i): Gen PKS *AaOKS* thực vật của *Aloe arborescens* được biểu hiện tạm thời ở *N. benthamiana* và *in vivo*, điều này dẫn đến sự tạo thành SEK4 và SEK4B, chứng minh rằng *AaOKS* có thể hoạt động như octaketit syntaza hoạt động *in vivo* ở *N. benthamiana*;

(ii): Vì axit flavokermesic (FK) anthron hoặc FK không được quan sát thấy khi một mình *AaOKS* được xâm nhiễm (agroinfiltrated), *N. benthamiana* có thể thiếu enzym nội sinh để chuyển hóa thêm octaketit không khử thành các hợp chất này.

(iii): Các gen *Streptomyces* sp. R1128 xyclaza, *ZhuI* và *ZhuJ*, được xâm nhiễm đồng thời (co-agroinfiltrated) (tức là, được đồng biểu hiện *in vivo*) với *AaOKS* và sự sản xuất/tích lũy *in vivo* các hợp chất khác nhau bao gồm FK được quan sát thấy. Do đó, sự biểu hiện khác loại của gen *Streptomyces* R1128 xyclaza dẫn đến sự sản xuất *in vivo* các hợp chất khác nhau bao gồm FK;

(iv): Sự sản xuất *in vivo* DcII và axit carminic (CA) được phát hiện khi *DcUGT2* được đồng biểu hiện với *AaOKS*, *ZhuI* và *ZhuJ* ở *N. benthamiana*.

Do đó, các tác giả sáng chế đã chứng minh:

(I): Gen polyketit syntaza (PKS) tương tự typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) (ví dụ, như, *Aloe arborescens* (*AaOKS*) và/hoặc *Hypericum perforatum* (*HpPKS2*)) và trong đó OKS là của giống khác với tế bào chủ (ví dụ, như, tế bào chủ thực vật *Nicotiana benthamiana* hoặc *Aspergillus nidulans* nấm) có khả năng chuyển hóa đơn vị khởi động và kéo dài thích hợp thành octaketit không khử trong các điều kiện sinh trưởng thích hợp và sau đó sản xuất *in vivo* octaketit không khử; và

(II): Có thể chuyển hóa *in vivo*, trong tế bào chủ tái tổ hợp đang sinh trưởng, octaketit không khử ở bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm (ví dụ, axit flavokermesic (FK) hoặc axit kermesic (KA)), trong đó hợp chất aglycon thơm cần quan tâm này không phải là SEK4 và/hoặc SEK4B (tức là, SEK4 và/hoặc SEK4B thu được tự phát từ octaketit không khử và trong bản mô tả này có thể được gọi là các sản phẩm song song);

(III): Nếu tế bào chủ tái tổ hợp còn chứa gen glycosyltransferaza mã hóa glycosyltransferaza sau đó hợp chất aglycon thơm cần quan tâm đã được tạo ra ở bước (II) có thể được glycosyl hoá *in vivo* bằng glycosyltransferaza để sản xuất hợp chất glycosit thơm cần quan tâm C₁₄-C₃₄ (ví dụ axit carminic (CA) hoặc DcII).

Được tin rằng bước (II) nêu trên tự nó có thể được xem là đóng góp mới, có ý nghĩa của sáng chế cho lĩnh vực này, vì dựa trên tài liệu kỹ thuật đã biết người có hiểu biết trung bình trong lĩnh vực kỹ thuật không thể biết chắc chắn rằng octaketit không khử được tạo ra trong bước (I) trên thực tế có thể có sẵn "một cách tự do" *in vivo* trong tế bào chủ tái tổ hợp thuộc giống khác để thực sự được chuyển thành hợp chất aglycon thơm khác cần quan tâm (ví dụ, axit flavokermesic (FK)) và không chỉ các hợp chất SEK4 và/hoặc SEK4B được sản xuất tự phát.

Như đã thảo luận trong bản mô tả này, có thể tốt hơn, nếu sự chuyển hóa *in vivo* của octaketit không khử được thực hiện nhờ sự tham gia của các aromataza/xyclaza hoạt động *in trans* (không phụ thuộc enzym PKS).

Như đã thảo luận ở trên, công trình thử nghiệm của các tác giả sáng chế như đã thảo luận trong các ví dụ thử nghiệm trong bản mô tả này đã chứng minh rằng có thể

sản xuất khác loại *in vivo*, ví dụ, axit carminic (CA) của côn trùng *Dactylopius costa* ở các tế bào chủ sản xuất tái tổ hợp khác như tế bào nấm *Aspergillus nidulans* hoặc tế bào thực vật *Nicotiana benthamiana* (họ hàng gần của cây thuốc lá).

Như đã thảo luận ở trên, hiện nay chỉ có thể có được axit carminic (CA) bằng cách phân lập trực tiếp từ cơ thể côn trùng *Dactylopius* hoặc bằng quy trình tổng hợp hóa học axit carminic bằng con đường bao gồm các hợp chất trung gian khác nhau (US5,424,421).

Do đó, có thể được xem là đóng góp lớn nhờ chứng minh được rằng có thể sản xuất khác loại *in vivo* axit carminic (CA), ví dụ, ở tế bào nấm *Aspergillus* hoặc tế bào thực vật *Nicotiana*.

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, dựa vào thông tin kỹ thuật mới ở đây, không có lý do để tin rằng không thể sản xuất *in vivo* bằng tế bào tái tổ hợp khác loại các hợp chất thu được từ octaketit khác nhau cần quan tâm – ví dụ như, hợp chất thơm C₁₄-C₃₄ cần quan tâm, trong đó ví dụ về hợp chất thơm C₁₄-C₃₄ có thể là Mutactin (ví dụ, xem Fig.2 trong tài liệu Tang, Y. et al (2004) đã thảo luận ở trên); Emodin, Hyperixin hoặc Pseudohyperixin (ví dụ, xem Fig.1 trong tài liệu Karppinen et al (2008) đã thảo luận ở trên); hoặc Barbaloin (ví dụ, xem Fig.1 và 4 trong tài liệu Mizuuchi et al (2009) đã thảo luận ở trên); hoặc axit carminic (CA).

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, ở bước (II) của phương pháp theo khía cạnh đầu tiên như được mô tả ở đây, các loại đặc hiệu của các hợp chất đặc hiệu aglycon thơm C₁₄-C₃₄ được tạo ra *in vivo* cần quan tâm (ví dụ như, axit flavokermesic (FK) hoặc axit kermesic (KA)) thường sẽ phụ thuộc vào enzym sau PKS (ví dụ, aromataza và/hoặc xyclaza tương đồng hoặc khác loại) có mặt trong tế bào chủ tái tổ hợp được sử dụng.

Tài liệu kỹ thuật đã biết mô tả nhiều aromataza và/hoặc xyclaza khác nhau mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể sử dụng để tạo ra aglycon thơm C₁₄-C₃₄ đặc hiệu cần quan tâm ở bước (II) của phương pháp theo khía cạnh thứ nhất như được mô tả ở đây.

Dựa vào kết quả được thảo luận ở đây, được thể hiện trong Fig.3 là ví dụ về mô hình cho con đường sinh tổng hợp thích hợp mà có thể dẫn đến axit carminic.

Nói ngắn gọn, ví dụ về mô hình con đường sinh tổng hợp bao gồm các bước sau:

(i): sử dụng PKS octaketit syntaza typ III để tạo ra octaketit không khử;

(ii): chuyển hóa octaketit không khử bằng tác dụng *in trans* (không phụ thuộc enzym PKS) của aromataza/xyclaza thành axit flavokermesic (FK) và/hoặc axit kermesic (KA); và

(iii): (1): axit kermesic (KA) aglucon được glucosyl hóa để thu được axit carminic CA; hoặc

(2): FK được glucosyl hóa để thu được DcII, hợp chất này được hydroxyl hóa tiếp để thu được axit carminic (CA).

Như đã thảo luận ở trên và không bị giới hạn bởi lý thuyết, tin rằng giải pháp đã biết không mô tả là PKS octaketit syntaza (OKS) typ III theo sáng chế có thể hoạt động *in vivo* ở tế bào chủ sản xuất khác loại thuộc giống khác, ví dụ OKS typ III thực vật có thể tạo ra octaketit không khử *in vivo* ở tế bào chủ sản xuất khác loại, ví dụ như, tế bào chủ sản xuất tái tổ hợp *Aspergillus*.

Như đã thảo luận ở trên, trong giải pháp đã biết, được biết rằng octaketit không khử có thể được chuyển hóa *in vivo* thành các hợp chất thơm thu được từ octaketit khác nhau cần quan tâm, loại hợp chất thơm đặc hiệu cần quan tâm thường sẽ phụ thuộc vào các enzym sau PKS đặc hiệu (ví dụ, aromataza và/hoặc xyclaza) có mặt *in vivo* trong tế bào chủ.

Ví dụ, Fig.1 của tài liệu Karppinen et al (2008) đã thảo luận ở trên thể hiện rằng octaketit không khử có thể được chuyển hóa *in vivo* trong *H. perforatum* thành hợp chất thơm Emodin anthron và/hoặc Emodin (Anthraquinon).

Fig.1 của tài liệu Mizuuchi et al. (2009) đã thảo luận ở trên thể hiện rằng octaketit không khử có thể được chuyển hóa *in vivo* trong *A. arborescens* thành hợp chất octaketit anthron barbaloin thơm.

Do đó, khía cạnh thứ nhất của sáng chế đề cập đến phương pháp sản xuất hợp chất thơm thu được từ octaketit cần quan tâm, trong đó phương pháp này bao gồm các bước sau:

(I): cho tiếp xúc *in vivo* trong tế bào chủ tái tổ hợp chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ:

(i): đơn vị khởi động và đơn vị kéo dài thích hợp với OKS đã được đưa vào bằng cách tái tổ hợp có khả năng chuyển hóa đơn vị khởi động và kéo dài thành

octaketit không khử trong các điều kiện thích hợp mà trong đó octaketit không khử được sản xuất ra; và

(II): chuyển hóa *in vivo*, trong tế bào chủ tái tổ hợp, octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm, trong đó hợp chất aglycon thơm cần quan tâm không phải là SEK4 và/hoặc SEK4B; và

(III): tùy ý, tế bào chủ tái tổ hợp còn chứa gen glycosyltransferaza mã hóa glycosyltransferaza và hợp chất aglycon thơm đã sản xuất ở bước (II) cần quan tâm được glycosyl hoá *in vivo* bởi glycosyltransferaza để sản xuất ra hợp chất glycosit thơm C₁₄-C₃₄ cần quan tâm; và

(IV): phân lập hợp chất aglycon thơm cần quan tâm của bước (II) và/hoặc phân lập hợp chất glycosit thơm cần quan tâm của bước (III) để thu được chế phẩm, trong đó chế phẩm này chứa ít hơn 1% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp.

Như đã thảo luận ở trên, các tác giả sáng chế đã biểu hiện tái tổ hợp polyketit syntaza (PKS) octaketit syntaza (OKS) typ III thực vật (*Aloe arborescens* và/hoặc *Hypericum perforatum*) ở *Aspergillus nidulans* và xác định rằng OKS thực sự hoạt động trong vật chủ khác loại. Do đó, ví dụ về tế bào chủ ở bước (I) có thể là *Aspergillus nidulans* và ví dụ về PKS/OKS typ III của bước (I) có thể là OKS từ *Aloe arborescens* và/hoặc *Hypericum perforatum*.

Như đã thảo luận ở trên và như đã biết trong lĩnh vực, octaketit không khử có thể được chuyển hóa tự phát *in vivo* thành hợp chất SEK4 và SEK4B (các cấu trúc được thể hiện trên Fig.2 có thể được gọi là sản phẩm song song).

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, việc xác định SEK4 và/hoặc SEK4B *in vivo* trong tế bào chủ tái tổ hợp chứng minh rằng octaketit không khử có mặt *in vivo*.

Do đó và như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, octaketit syntaza (OKS) typ III của bước (I) có thể được định nghĩa là OKS, mà có khả năng sản xuất *in vitro* các sản phẩm song song thu được từ octaketit SEK4 và/hoặc SEK4B.

Thử nghiệm cho phân tích *in vitro* này được mô tả trong ví dụ thử nghiệm 9 trong bản mô tả này.

Như đã thảo luận ở trên, thuật ngữ polyketit syntaza (PKS) typ III đã được biết rõ bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật và do đó người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể xác định theo cách thông thường là PKS cần quan tâm có được hiểu là PKS typ III và do đó không là PKS typ I hoặc typ II hay không.

Như đã thảo luận ở trên, trong giải pháp đã biết, được biết rằng octaketit không khử có thể được chuyển hóa *in vivo* thành các hợp chất thơm thu được từ octaketit khác nhau cần quan tâm và việc này sẽ thường phụ thuộc vào enzym sau PKS đặc hiệu (ví dụ, aromataza và/hoặc xyclaza) có mặt *in vivo* trong tế bào chủ.

Do đó và như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, sự chuyển hóa ở bước (II) của octaketit không khử thành hợp chất aglycon thơm C_{14} - C_{34} cần quan tâm thường được thực hiện nhờ sự tham gia của enzym sau PKS thích hợp (ví dụ, aromataza và/hoặc xyclaza) có mặt *in vivo* trong tế bào chủ.

Hợp chất aglycon thơm cần quan tâm của bước (II), về nguyên tắc, có thể là hợp chất aglycon thơm C_{14} - C_{34} cần quan tâm bất kỳ trong bản mô tả này, ví dụ như, hợp chất anthraquinon thơm hoặc hợp chất anthron thơm (chi tiết hơn xem bên dưới).

Như đã thảo luận ở trên - các tác giả sáng chế đã xác định là axit flavokermesic (FK) có mặt trong tế bào *Aspergillus nidulans* chứa OKS typ III đã được đưa vào bằng cách tái tổ hợp từ *Aloe arborescens* và/hoặc *Hypericum perforatum*.

Axit flavokermesic (FK) và axit kermesic (KA) đều có 16 nguyên tử cacbon (C), - chúng được hiểu là các ví dụ về hợp chất aglycon thơm C_{16} và là các ví dụ về hợp chất aglycon thơm C_{14} - C_{34} cần quan tâm ở bước (II).

Như đã thảo luận trong bản mô tả này, nếu hợp chất aglycon thơm cần quan tâm được phân lập ở bước (IV) là FK, hợp chất này có thể được chuyển hóa *in vitro*, nhờ phản ứng glucosyl hóa, thành DcII và sau đó thành axit carminic (CA) (ví dụ, xem Fig.3 trong bản mô tả này).

Như đã thảo luận trong bản mô tả này, ví dụ thay thế tùy ý khác có thể là tế bào chủ còn biểu hiện *in vivo* glycosyltransferaza theo sáng chế và DcII và/hoặc axit carminic (CA) được sản xuất *in vivo* trong tế bào chủ này trong bước (III). Đây có thể được xem là ví dụ về bước (III) tùy ý của khía cạnh thứ nhất.

Các ví dụ thử nghiệm ở đây thể hiện các ví dụ được ưu tiên của phương án thay thế tùy ý này.

DcII và/hoặc axit carminic (CA) đều là glucosit và phần aglycon (tức là, không bao gồm glucoza) của các hợp chất này có 16 nguyên tử cacbon (C).

Do đó, cả hợp chất DcII và/hoặc axit carminic (CA) đều được hiểu ở đây là các ví dụ về hợp chất glycosit thơm C_{16} và là các ví dụ về hợp chất glycosit thơm C_{14} - C_{34} cần quan tâm trong bước (III) tùy ý.

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, ở ngữ cảnh này, bước phân lập (IV) chủ yếu liên quan đến bước phân lập/tinh chế hợp chất thơm cần quan tâm từ vật liệu tế bào chủ tái tổ hợp (hoặc môi trường nuôi cấy/sinh trưởng nếu tế bào chủ này là, ví dụ, tế bào nấm/vi sinh vật).

Các tác giả sáng chế đã xác định rằng axit flavokermesic (FK) có mặt trong tế bào *Aspergillus nidulans* chứa OKS typ III đã được đưa vào bằng cách tái tổ hợp từ *Aloe arborescens* và/hoặc *Hypericum perforatum*.

Bất ngờ là giải pháp đã biết không mô tả rằng octaketit syntaza (OKS) PKS typ III theo sáng chế có thể hoạt động *in vivo* trong tế bào chủ sản xuất khác loại thuộc giống khác, ví dụ OKS typ III thực vật có thể tạo ra octaketit không khử *in vivo* trong tế bào chủ sản xuất khác loại, ví dụ như, tế bào chủ sản xuất tái tổ hợp *Aspergillus*.

Như đã thảo luận ở trên, giải pháp kỹ thuật đã biết mô tả nhiều ví dụ về thử nghiệm *in vitro* trên các enzym OKS khác nhau. Tuy nhiên, không có giải pháp đã biết nào mô tả là OKS hoạt động *in vivo* trong tế bào chủ thuộc giống khác.

Không bị giới hạn bởi lý thuyết, OKS typ III có thể hoạt động trong vật chủ khác loại khi OKS đang được bảo vệ *in vivo* kháng lại sự thoái biến (ví dụ, qua sự tạo thành đơn vị trao đổi chất/phức hợp với protein khác). Cơ chế thoái biến này không xuất hiện *in vitro*.

Việc OKS hoạt động trong tế bào thực vật *Aspergillus nidulans* và *Nicotiana* chứng minh rằng tế bào thực vật có khả năng bảo vệ OKS được đưa vào bằng cách tái tổ hợp khác loại.

Tin rằng điều này có thể được xem là vấn đề chung, tức là các hệ bảo vệ này cũng sẽ có mặt trong nhiều loại tế bào chủ khác (ví dụ như, nấm men, cây thuốc lá v.v).

Ngoài ra, sản phẩm octaketit của OKS bị thoái biến tự phát thành SEK4 và/hoặc SEK4B, hợp chất này có thể ức chế enzym ở các nồng độ *in vivo*. Khi enzym chuyển

hóa octaketit có mặt, PKS sẽ không bị ức chế bởi SEK4 và/hoặc SEK4B và vẫn còn hoạt động.

Tóm lại, phương pháp thích hợp để sản xuất OKS hoạt động là biểu hiện nó cùng với enzym chuyển hóa octaketit và/hoặc tạo các phức hợp với các enzym. Các enzym này có thể có mặt tự nhiên trong các tế bào biểu hiện OKS đã được đưa vào hoặc có thể được đưa vào bằng cách tái tổ hợp.

Như được thể hiện ở đây, việc axit flavokermesic (FK) có mặt trong tế bào *Aspergillus nidulans* chứa OKS typ III đã được đưa vào bằng cách tái tổ hợp từ *Aloe arborescens* và/hoặc *Hypericum perforatum* chỉ ra rằng tế bào *Aspergillus nidulans* chứa các enzym (ví dụ, aromataza/xyclaza) chuyển hóa octaketit không khử đã được sản xuất.

Không bị giới hạn bởi lý thuyết, tin rằng điều này có thể được xem là vấn đề chung, tức là xyclaza và/hoặc aromataza này cũng sẽ có mặt trong nhiều loại tế bào chủ khác (ví dụ như nấm men, cây thuốc lá v.v).

Định nghĩa

Tất cả các định nghĩa của thuật ngữ trong bản mô tả này đều phù hợp với định nghĩa được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật trong bản mô tả này.

Thuật ngữ “aglycon” dùng để chỉ phần không phải carbohydrat của dạng được glycosyl hóa tương ứng của aglycon. Khi đường là glucoza, aglycon có thể được gọi là aglucon. Ngoài ra, khi đường là glucoza, thuật ngữ được glycosyl hóa có thể được sử dụng thay vì được glycosyl hóa.

Khi aglycon được glycosyl hoá ở nhóm hydroxy, thường tạo ra liên kết *O*-glycosidic, tức là *O*-Glycosit (hoặc *O*-Glucosit nếu đường là glucoza).

Khi aglycon được glycosyl hóa bằng liên kết cacbon-cacbon, nó có thể được gọi ở đây là liên kết *C*-glycosidic, tức là *C*-Glycosit (hoặc *C*-Glucosit nếu đường là glucoza).

Thuật ngữ “glycosit” dùng để chỉ hợp chất, mà khi thủy phân có thể tạo ra đường và phần còn lại không phải đường (aglycon), ví dụ, các glucosit có thể tạo ra glucoza, galactosit có thể tạo ra galactoza.

Thuật ngữ “glycosyltransferaza” dùng để chỉ enzym có khả năng liên hợp đường được hoạt hóa bởi nucleotit với hợp chất (ví dụ, hợp chất aglycon). Đường, ví

dự, có thể là isome D và L của galactosa, glucosamin, *N*-axetylglusamin, xyloza, axit glucuronic, rhamnoza, arabinoza, mannoza hoặc glucoza. Theo cách khác, đường có thể là dẫn xuất carbohydrat, ví dụ như, inositol, olivoza, rhodinoza và v.v. có sẵn dưới dạng nucleotit diphosphat. Ngoài ra, đường có thể, ví dụ, là monosaccarit, disaccarit hoặc trisaccarit. Trong trường hợp oligo- và polysaccarit, các đường được liên kết từng phân tử một, ví dụ, bằng cách sử dụng một hoặc một số glycosyltransferaza. Nếu đường là glucoza, glycosyltransferaza có thể được gọi là glucosyltransferaza.

Khi glycosyltransferaza liên hợp đường được hoạt hóa bởi nucleotit với hợp chất qua liên kết *C*-glycosidic, nó có thể được gọi ở đây là *C*-glycosyltransferaza.

Khi glycosyltransferaza liên hợp đường với aglycon qua liên kết *O*-glycosidic, nó có thể được gọi ở đây là *O*-glycosyltransferaza.

Thuật ngữ “lai” liên quan đến polynucleotit mà lai trong các điều kiện có độ nghiêm ngặt ít nhất là trung bình với (i) các nucleotit từ 1 đến 1548 của SEQ ID NO:1 hoặc (ii) chuỗi bổ sung của (i) đề cập đến trình tự nucleotit lai với đoạn dò axit nucleic được gắn nhãn tương ứng với trình tự nucleotit được thể hiện trong SEQ ID NO:1 hoặc chuỗi bổ sung của nó trong các điều kiện có độ nghiêm ngặt trung bình đến rất cao. Phân tử mà đoạn dò axit nucleic lai với nó trong các điều kiện này có thể được phát hiện bằng cách sử dụng, ví dụ, màng tia X.

Các điều kiện lai nghiêm ngặt theo sáng chế là các điều kiện nghiêm ngặt mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật thường sẽ hiểu là thích hợp, tức là, cho các điều kiện nghiêm ngặt trung bình, các điều kiện mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ hiểu là các điều kiện nghiêm ngặt trung bình. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật biết các điều kiện lai nghiêm ngặt thích hợp, ví dụ xem (J. Sambrook, E.F. Fritsch, và T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, xuất bản lần thứ 2, Cold Spring Harbor, New York).

Theo lĩnh vực này, đối với đoạn dò dài có độ dài ít nhất 100 nucleotit, các điều kiện nghiêm ngặt rất thấp đến rất cao được định nghĩa là lai sơ bộ và lai ở 42°C trong 5X SSPE, SDS 0,3%, 200 µg/ml ADN tinh trùng cá hồi được cắt và làm biến tính, và 25% formamit đối với độ nghiêm ngặt rất thấp và thấp, 35% formamit đối với độ nghiêm ngặt trung bình và trung bình cao, hoặc 50% formamit đối với độ nghiêm ngặt cao và rất cao, theo quy trình thấm Southern tiêu chuẩn trong từ 12 đến 24 giờ theo cách tối ưu.

Đối với các đoạn dò dài có chiều dài ít nhất 100 nucleotit, vật liệu mang cuối cùng được rửa ba lần mỗi lần 15 phút sử dụng 2X SSC, SDS 0,2% tốt hơn, nếu ít nhất ở 45°C (độ nghiêm ngặt rất thấp), tốt hơn nữa, nếu ít nhất ở 50°C (độ nghiêm ngặt thấp), tốt hơn nữa, nếu ít nhất ở 55°C (độ nghiêm ngặt trung bình), tốt hơn nữa, nếu ít nhất ở 60°C (độ nghiêm ngặt trung bình-cai), thậm chí tốt hơn nữa, nếu ít nhất ở 65°C (độ nghiêm ngặt cao), và tốt nhất, nếu ít nhất ở 70°C (độ nghiêm ngặt rất cao).

Thuật ngữ “*in vitro*” (tiếng Latin: trong ống thủy tinh) đề cập đến các nghiên cứu được tiến hành bằng cách sử dụng các thành phần của sinh vật đã được phân lập khỏi môi trường sinh học xung quanh thông thường của chúng để cho phép phân tích chi tiết hơn hoặc thuận tiện hơn là có thể được thực hiện với toàn bộ sinh vật. Thông thường, các thí nghiệm này thường được gọi là “thí nghiệm ống nghiệm”. Ngược lại, các nghiên cứu *in vivo* là các nghiên cứu được thực hiện với các sinh vật sống ở trạng thái nguyên vẹn bình thường.

Thuật ngữ “*in vivo*” (tiếng Latin là “trong cơ thể sống”) đề cập đến việc thử nghiệm sử dụng sinh vật sống, nguyên vẹn trái ngược với một phần sinh vật hoặc sinh vật đã chết, hoặc môi trường được kiểm soát *in vitro* (“trong ống thủy tinh”, ví dụ, trong ống nghiệm hoặc đĩa petri).

Thuật ngữ “polynucleotit được phân lập” chủ yếu đề cập đến polynucleotit được phân lập khỏi môi trường tự nhiên của nó. Nói cách khác là chế phẩm polynucleotit này gần như không chứa vật liệu polynucleotit khác mà được kết hợp với nó trong tự nhiên. Trình tự polynucleotit mã hóa glycosyltransferaza mới đã phân lập/tách dòng được mô tả ở đây được thể hiện trong SEQ ID NO:1 và nó được phân lập từ côn trùng *Dactylopius coccus*. Do đó, như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, thuật ngữ “polynucleotit được phân lập” như được sử dụng trong bản mô tả này không bao gồm polynucleotit của SEQ ID NO:1 vì nó có mặt tự nhiên trong bộ gen của *Dactylopius coccus*. Thuật ngữ “polynucleotit được phân lập” chủ yếu đề cập đến polynucleotit được phân lập ở dạng thích hợp để sử dụng trong các hệ sản xuất protein được thao tác di truyền. Vì vậy, polynucleotit được phân lập chứa nhiều nhất là 10%, tốt hơn, nếu nhiều nhất là 8%, tốt hơn nữa, nếu nhiều nhất là 6%, tốt hơn nữa, nếu nhiều nhất là 5%, tốt hơn nữa, nếu nhiều nhất là 4%, tốt hơn nữa, nếu nhiều nhất là 3%, thậm chí tốt hơn nữa, nếu nhiều nhất là 2%, tốt nhất, nếu nhiều nhất là 1%, và thậm chí tốt nhất, nếu nhiều nhất là 0,5% theo khối lượng là vật

liệu polynucleotit khác mà kết hợp tự nhiên với nó. Thuật ngữ “polynucleotit được phân lập” có thể được gọi theo cách khác là “polynucleotit được tách dòng”.

Thuật ngữ “polypeptit được phân lập” chủ yếu đề cập đến polypeptit được phân lập khỏi môi trường tự nhiên của nó. Glycosyltransferaza polypeptit mới được mô tả ở đây như được thể hiện trong SEQ ID NO: 2 ở đây được phân lập từ côn trùng *Dactylopius coccus*. Do đó, như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, thuật ngữ “polypeptit được phân lập” như được sử dụng ở đây không bao gồm glycosyltransferaza polypeptit của SEQ ID NO:2 vì nó có mặt tự nhiên trong bộ gen của *Dactylopius coccus*. Thuật ngữ “polypeptit được phân lập” ở đây thể hiện chế phẩm polypeptit chứa nhiều nhất là 10%, tốt hơn, nếu nhiều nhất là 8%, tốt hơn nữa, nếu nhiều nhất là 6%, tốt hơn nữa, nếu nhiều nhất là 5%, tốt hơn nữa, nếu nhiều nhất là 4%, nhiều nhất là 3%, thậm chí tốt hơn nữa, nếu nhiều nhất là 2%, tốt nhất, nếu nhiều nhất là 1%, và thậm chí tốt nhất, nếu nhiều nhất là 0,5% theo khối lượng là vật liệu polypeptit khác mà kết hợp tự nhiên với nó, như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, thuật ngữ “vật liệu polypeptit khác mà kết hợp tự nhiên với nó” có thể, liên quan đến glycosyltransferaza polypeptit mới được mô tả ở đây như được thể hiện trong SEQ ID NO: 2, được thấy liên quan đến vật liệu polypeptit khác mà kết hợp tự nhiên với nó trong *Dactylopius coccus*. Trong một số trường hợp, có thể tốt hơn, nếu “polypeptit được phân lập” dùng để chỉ polypeptit có độ tinh khiết ít nhất 20%, tốt hơn nếu, tinh khiết ít nhất 40%, tốt hơn nữa, nếu tinh khiết ít nhất 60%, thậm chí còn tốt hơn nữa, nếu tinh khiết ít nhất 80%, tốt nhất, nếu tinh khiết ít nhất là 90%, và thậm chí tốt nhất, nếu tinh khiết ít nhất là 95%, như được xác định bởi SDS-PAGE.

Thuật ngữ “octaketit không khử” biểu thị octaketit không khử, trong đó loại không khử được đặc trưng bởi sự có mặt của nhóm keton ban đầu trong ketit ($-\text{CH}_2-\text{CO}-$), có nguồn gốc từ các đơn vị khởi động hoặc kéo dài, dưới dạng keton hoặc ở dạng liên kết đôi trong nhóm thơm. Trong polyketit đã khử, một hoặc tất cả keton đã được khử thành nhóm rượu ($-\text{CH}_2-\text{CHOH}-$), ví dụ, bởi miền/enzym KR, hoặc tiếp tục thành nhóm alken ($-\text{C}=\text{C}-$), ví dụ, bởi miền/enzym DH, hoặc còn tiếp tục thành nhóm alkan ($-\text{CH}_2-\text{CH}_2-$), ví dụ, bởi miền/enzym ER. Dựa vào các tính chất hóa học này của sản phẩm tạo thành, các PKS liên quan được phân loại là PKS không khử hoặc PKS khử.

Thuật ngữ “PKS không khử” hoặc “polyketit syntaza không khử” biểu thị PKS mà không khử nhóm keton trong mạch polyketit đã tạo thành. Việc thiếu các nhóm khử có thể do (I) thiếu vị trí hoạt động keto-reductaza (KR) cần thiết trong enzym; và/hoặc (II) thiếu các enzym biến đổi có khả năng xúc tác cho phản ứng khử keto.

Thuật ngữ "cấu trúc axit nucleic" như được sử dụng ở đây dùng để chỉ phân tử axit nucleic, chuỗi đơn hoặc kép, được phân lập từ gen xuất hiện tự nhiên hoặc được biến đổi để chứa các đoạn axit nucleic theo cách mà không tồn tại trong tự nhiên. Thuật ngữ cấu trúc axit nucleic đồng nghĩa với thuật ngữ “catxet biểu hiện” khi cấu trúc axit nucleic chứa các trình tự điều khiển cần thiết để biểu hiện trình tự mã hóa theo sáng chế. Như đã biết trong lĩnh vực, các trình tự điều khiển bao gồm tất cả các thành phần, cần thiết hoặc thuận lợi để biểu hiện polynucleotit mã hóa polypeptit theo sáng chế. Mỗi trình tự điều khiển có thể là tự nhiên hoặc ngoại sinh đối với trình tự nucleotit mã hóa polypeptit. Trình tự điều khiển này bao gồm, nhưng không giới hạn ở, gen chỉ huy, trình tự polyadenyl hóa, trình tự tiền peptit, gen khởi đầu, trình tự peptit tín hiệu, và đoạn kết thúc phiên mã. Ở mức tối thiểu, trình tự điều khiển bao gồm gen khởi đầu, và tín hiệu dừng phiên mã và dịch mã. Các trình tự điều khiển có thể được cung cấp các đoạn liên kết với mục đích đưa vào các vị trí giới hạn đặc hiệu tạo điều kiện thuận lợi cho việc nối trình tự điều khiển với vùng mã hóa của trình tự nucleotit mã hóa polypeptit.

Thuật ngữ “octaketit” (tiếng Hy Lạp là “tám”) biểu thị chuỗi polyketit bao gồm tám đơn vị ketit, có nghĩa là khung polyketit bao gồm 16 nguyên tử cacbon. Thuật ngữ “ketit” dùng để chỉ đơn vị $\text{-CH}_2\text{-CO-}$ hoặc dạng biến đổi của nhóm này.

Thuật ngữ "vector biểu hiện tái tổ hợp" đề cập đến vector biểu hiện tái tổ hợp chứa polynucleotit theo sáng chế, gen khởi đầu, và tín hiệu dừng phiên mã và dịch mã. Các axit nucleic và trình tự điều khiển khác nhau được mô tả ở trên có thể được liên kết với nhau để tạo ra vector biểu hiện tái tổ hợp mà có thể bao gồm một hoặc nhiều vị trí giới hạn thuận tiện để cho phép chèn hoặc thay thế trình tự nucleotit mã hóa polypeptit ở các vị trí này.

Thuật ngữ “tế bào chủ tái tổ hợp” trong bản mô tả này nên được hiểu theo lĩnh vực kỹ thuật. Như đã biết trong lĩnh vực này, phân tử polynucleotit tái tổ hợp (ví dụ, ADN) là các phân tử polynucleotit (ví dụ, ADN) được tạo ra bằng các phương pháp tái tổ hợp di truyền (như tách dòng phân tử) trong phòng thí nghiệm để gom các vật liệu

di truyền từ nhiều nguồn lại với nhau, tạo ra trình tự mà sẽ không được tìm thấy trong cơ thể sinh học theo cách khác. Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, tế bào chủ tái tổ hợp chứa phân tử polynucleotit tái tổ hợp (ví dụ, ADN) và tế bào chủ tái tổ hợp này do đó sẽ không được hiểu là bao gồm tế bào kiểu đại tự nhiên như vậy, ví dụ, như tế bào *Dactylopius coccus* kiểu đại tự nhiên.

Thuật ngữ “độ đồng nhất về trình tự” đề cập đến mối liên hệ giữa hai trình tự axit amin hoặc giữa hai trình tự nucleotit.

Với mục đích của sáng chế, mức độ đồng nhất trình tự giữa hai trình tự axit amin được xác định bằng cách sử dụng thuật toán Needleman-Wunsch (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) như được thực hiện trong chương trình Needle của gói phần mềm EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), tốt hơn nếu là bản 3.0.0 hoặc bản gần hơn. Các thông số tùy ý được sử dụng là điểm phạt mở khoảng trống bằng 10, điểm phạt mở rộng khoảng trống bằng 0,5, và ma trận thay thế EBLOSUM62 (bản EMBOSS của BLOSUM62). Dữ liệu đầu ra của Needle được gán là “độ đồng nhất dài nhất” (thu được bằng cách sử dụng tùy chọn nobrief) được sử dụng làm tỷ lệ phần trăm đồng nhất và được tính như sau:

$$(\text{Các gốc giống nhau} \times 100) / (\text{Chiều dài đoạn sắp thẳng hàng} - \text{Tổng số khoảng trống trong đoạn sắp thẳng hàng})$$

Với mục đích của sáng chế, mức độ đồng nhất trình tự giữa hai trình tự nucleotit được xác định bằng cách sử dụng thuật toán Needleman-Wunsch (Needleman and Wunsch, 1970, đề cập trên đây) như được thực hiện trong chương trình Needle của gói phần mềm EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, đề cập trên đây), tốt hơn nếu là bản 3.0.0 hoặc bản gần hơn. Các thông số tùy ý được sử dụng là điểm phạt mở khoảng trống bằng 10, điểm phạt mở rộng khoảng trống bằng 0,5, và ma trận thay thế EDNAFULL (bản EMBOSS của NCBI NUC4.4). Dữ liệu đầu ra của Needle được gán là “tỷ lệ đồng nhất dài nhất” (thu được bằng cách sử dụng tùy chọn nobrief) được sử dụng làm tỷ lệ phần trăm đồng nhất và được tính như sau:

$$(\text{Deoxyribonucleotit giống nhau} \times 100) / (\text{Chiều dài đoạn sắp thẳng hàng} - \text{Tổng số khoảng trống trong đoạn sắp thẳng hàng})$$

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, đối với cả “độ đồng nhất trình tự giữa hai trình tự nucleotit” và “độ đồng nhất trình tự giữa hai trình tự axit amin”, thuật ngữ “chiều dài đoạn sắp thẳng hàng” nên được hiểu là chiều dài thực tế của đoạn sắp thẳng hàng giữa hai trình tự cần được so sánh về mức độ đồng nhất trình tự.

Ví dụ, nếu trình tự tham chiếu là SEQ ID cụ thể gồm, ví dụ, 100 axit amin và trình tự còn lại là trình tự đồng nhất có ít hơn 25 axit amin ở một đầu (tức là, trình tự còn lại này có chiều dài bằng 75 axit amin) thì “chiều dài đoạn sắp thẳng hàng” sẽ bằng 75 axit amin và tỷ lệ phần trăm đồng nhất sẽ bằng 100%.

Ví dụ khác là, nếu trình tự tham chiếu là SEQ ID cụ thể gồm, ví dụ, 100 axit amin và trình tự còn lại là trình tự đồng nhất có thêm 25 axit amin ở một đầu (tức là, trình tự còn lại này có chiều dài bằng 125 axit amin), thì “chiều dài đoạn sắp thẳng hàng” bằng 100 axit amin và tỷ lệ phần trăm đồng nhất sẽ bằng 100%.

Thuật ngữ “Polyketit syntaza (PKS) typ III”, như đã thảo luận ở đây, là thuật ngữ đã được biết rõ bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật và người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ biết PKS đặc hiệu cần quan tâm có phải là PKS typ III hay không. Như đã thảo luận trong tài liệu đánh giá của *Yu et al* (2012), các PKS typ III là các enzym tự chủ tạo thành homodime. Vị trí hoạt động duy nhất của chúng trong mỗi monome xúc tác cho việc nối và kéo dài để tạo thành sản phẩm polyketit.

Các phương án theo sáng chế được mô tả dưới đây, chỉ bằng các ví dụ.

Mô tả vắn tắt hình vẽ

Fig. 1: Thể hiện bằng sơ đồ hoạt tính glycosyltransferaza theo sáng chế của glycosyltransferaza mới đã được phân lập/tách dòng được mô tả ở đây của SEQ ID NO:2. Như đã minh họa trên hình vẽ này, đã phát hiện ra là có thể liên hợp glucoza với aglycon axit flavokermesic (FK) và axit kermesic (KA).

Fig. 2: Thể hiện Fig. 5A của tài liệu *Tang, Y. et al* (2004). Như có thể được thấy trên hình vẽ này và như được mô tả thêm trong tài liệu, *Act* PKS (được gọi là octaketit syntaza (OKS)) tạo ra octaketit không khử và octaketit này, nhờ *ZhuI* aromataza/*xyclaza* và *ZhuJ* *xyclaza*, được chuyển hóa thành axit flavokermesic (FK) (được gọi là TMAC). Các hợp chất SEK4 và SEK4B cũng được tạo thành tự phát (các cấu trúc được thể hiện trên Fig. 2 có thể được gọi là các sản phẩm song song).

Fig. 3: Thể hiện ví dụ về mô hình cho con đường sinh tổng hợp thích hợp có thể dẫn đến axit carminic.

Fig. 4: Thể hiện khung 9,10-anthraquinon; khung anthron và cấu trúc hợp chất FK antron.

Mô tả chi tiết sáng chế

Tế bào chủ tái tổ hợp – ví dụ, bước (I)

Dựa vào, ví dụ, thông tin trình tự được bộc lộ trong bản mô tả này, người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể tạo ra tế bào chủ tái tổ hợp theo sáng chế bằng các phương pháp thông thường. Ví dụ, dựa trên giải pháp đã biết người có hiểu biết trung bình trong lĩnh vực kỹ thuật biết được rất nhiều tế bào chủ tái tổ hợp thích hợp khác nhau mà trong nhiều năm đã được sử dụng làm tế bào chủ tái tổ hợp, ví dụ, để biểu hiện các polypeptit khác nhau cần quan tâm.

Tế bào chủ tái tổ hợp trong phương pháp theo khía cạnh thứ nhất có thể là tế bào chủ tái tổ hợp đang sinh trưởng hoặc, ví dụ, trong pha ổn định.

Tốt hơn là, tế bào chủ tái tổ hợp trong phương pháp theo khía cạnh thứ nhất là tế bào chủ tái tổ hợp đang sinh trưởng và bước (I) và bước (II) theo khía cạnh thứ nhất là:

(I): trong tế bào chủ tái tổ hợp đang sinh trưởng chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ, cho tiếp xúc *in vivo*:

(i): đơn vị khởi động và đơn vị kéo dài thích hợp với OKS đã được đưa vào bằng cách tái tổ hợp có khả năng chuyển hóa đơn vị khởi động và kéo dài thành octaketit không khử trong các điều kiện sinh trưởng thích hợp, trong đó octaketit không khử được sản xuất *in vivo*; và

(II): chuyển hóa *in vivo*, trong tế bào chủ tái tổ hợp đang sinh trưởng, octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm, trong đó hợp chất aglycon thơm cần quan tâm không phải là SEK4 và/hoặc SEK4B.

Tế bào chủ tái tổ hợp có thể là tế bào thích hợp bất kỳ như tế bào nhân chuẩn bất kỳ [ví dụ, tế bào động vật có vú (ví dụ như, tế bào trứng chuột đồng Trung Quốc (CHO)) hoặc tế bào thực vật] hoặc tế bào tiền nhân bất kỳ.

Có thể tốt hơn, nếu tế bào chủ tái tổ hợp là tế bào thực vật, tốt hơn, nếu trong đó tế bào thực vật là tế bào thực vật được chọn từ nhóm bao gồm: *Nicotiana* sp. (ví dụ, tế bào *Nicotiana benthamiana*); cây đại hoàng, kiều mạch, *Hypericum* và *Aloe* sp..

Tốt hơn, nếu tế bào thực vật là *Nicotiana* sp., tốt hơn nữa, nếu tế bào thực vật là *Nicotiana benthamiana*.

Tế bào chủ tái tổ hợp có thể là tế bào được chọn từ nhóm bao gồm tế bào nấm sợi và tế bào vi sinh vật.

Nấm sợi bao gồm tất cả dạng có sợi của phân ngành Eumycota và Oomycota (như được xác định bởi Hawksworth et al. , 1995, nêu trên). Nấm sợi được đặc trưng bởi khuẩn ty dinh dưỡng chứa chitin, xenluloza, glucan, chitosan, mannan, và các polysaccarit phức tạp khác. Sự sinh trưởng sinh dưỡng là nhờ sự kéo dài sợi nấm và sự dị hóa cacbon là hiệu quả bắt buộc. Ngược lại, sự sinh trưởng sinh dưỡng bởi nấm men như *Saccharomyces cerevisiae* là bằng cách nảy mầm tản đơn bào và sự dị hóa cacbon có thể là lên men.

Có thể tốt hơn, nếu tế bào nấm sợi là tế bào của loài được chọn từ nhóm bao gồm, nhưng không giới hạn ở, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, và *Trichoderma* hoặc dạng sinh sản hữu tính hoặc đồng nghĩa của chúng.

Tế bào *Aspergillus* được ưu tiên là *Aspergillus niger*, *Aspergillus oryzae* hoặc *Aspergillus nidulans*.

Tế bào *Aspergillus* được ưu tiên nhất là *Aspergillus nidulans*.

Tế bào vi sinh vật được ưu tiên theo sáng chế là tế bào vi sinh vật được chọn từ nhóm bao gồm tế bào nấm men và tế bào tiên nhân.

Tế bào nấm men được ưu tiên là tế bào nấm men thuộc ngành được chọn từ nhóm bao gồm Ascomycetes, Basidiomycetes và nấm bất toàn. Tốt hơn là tế bào nấm men thuộc ngành Ascomycetes.

Tế bào nấm men Ascomycetes được ưu tiên thuộc giống được chọn từ nhóm bao gồm *Ashbya*, *Botryosphaeria*, *Debaryomyces*, *Hansenula*, *Kluyveromyces*, *Lipomyces*, *Saccharomyces*, như *Saccharomyces cerevisiae*, *Pichia* và *Schizosaccharomyces*.

Tế bào nấm men được ưu tiên là tế bào nấm men thuộc giống được chọn từ nhóm bao gồm *Saccharomyces*, như *Saccharomyces cerevisiae* và *Pichia*.

Tế bào tiền nhân được ưu tiên được chọn từ nhóm bao gồm: *Bacillus*, *Streptomyces*, *Corynebacterium*, *Pseudomonas*, vi khuẩn axit lactic và tế bào *E. coli*.

Tế bào *Bacillus* được ưu tiên là *B. subtilis*, *Bacillus amyloliquefaciens* hoặc *Bacillus licheniformis*.

Tế bào *Streptomyces* được ưu tiên là *Streptomyces setonii* hoặc *Streptomyces coelicolor*.

Tế bào *Corynebacterium* được ưu tiên là *Corynebacterium glutamicum*.

Tế bào *Pseudomonas* được ưu tiên là *Pseudomonas putida* hoặc *Pseudomonas fluorescens*.

Gen polyketit syntaza (PKS) mã hóa octaketit syntaza (OKS) – ví dụ bước (I)

Bước (I) theo khía cạnh thứ nhất đề cập đến “gen polyketit syntaza (PKS) typ III mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ”.

Như đã thảo luận ở trên và như đã biết trong lĩnh vực - octaketit không khử có thể được chuyển hóa tự phát *in vivo* thành hợp chất SEK4 và SEK4B (các cấu trúc được thể hiện trên Fig.2 trong bản mô tả này – có thể được gọi ở đây là các sản phẩm song song).

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này – việc nhận diện SEK4 và/hoặc SEK4B *in vivo* trong tế bào chủ tái tổ hợp chứng minh rằng octaketit không khử có mặt *in vivo*.

Do đó, octaketit syntaza (OKS) typ III của bước (I) có thể được định nghĩa là OKS, có khả năng sản xuất *in vitro* các sản phẩm song song thu được từ octaketit SEK4 và/hoặc SEK4B.

Thử nghiệm cho phân tích *in vitro* này được mô tả, ví dụ, trong ví dụ thử nghiệm 9 trong bản mô tả này.

Như đã thảo luận ở trên, người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ biết PKS đặc hiệu cần quan tâm có phải là PKS typ III hay không.

Như đã thảo luận ở trên, trong giải pháp kỹ thuật đã biết, một số PKS typ III khác nhau từ thực vật, vi khuẩn và/hoặc nấm đã được biết.

Do đó, có thể được ưu tiên là gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ thực vật, vi khuẩn hoặc nấm.

Tốt hơn là, gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ thực vật.

Nếu gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ thực vật, tốt hơn nếu thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe* spp. (tốt hơn là *Aloe arborescens*), *Hypericum* spp. (tốt hơn là *Hypericum perforatum*), cây đại hoàng, kiều mạch và *Hawortia* spp..

Tốt hơn là thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe* spp. và *Hypericum* spp..

Tốt hơn nữa là, thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe arborescens* và *Hypericum perforatum*.

Nếu gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ vi khuẩn, tốt hơn nếu vi khuẩn này là vi khuẩn được chọn từ nhóm bao gồm: *Streptomyces* spp. (tốt hơn là *Streptomyces coelicolor*).

Nếu gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ nấm, tốt hơn nếu nấm này là nấm được chọn từ nhóm bao gồm: *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, và *Trichoderma*.

Aspergillus được ưu tiên là *Aspergillus niger*, *Aspergillus oryzae* hoặc *Aspergillus nidulans*.

Aspergillus được ưu tiên nhất là *Aspergillus nidulans*.

Như đã biết trong lĩnh vực này, PKS typ I và/hoặc PKS typ II có thể được gây đột biến để loại bỏ các phần tử thích hợp (ví dụ, vị trí hoạt động) để chuyển nó thành hợp chất mà về mặt chức năng được xem là PKS typ III bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. PKS, mà về mặt chức năng được xem là PKS typ III bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, được hiểu ở đây là PKS Typ III.

Như đã thảo luận ở trên và trong các ví dụ thử nghiệm ở đây, các tác giả sáng chế đã biểu hiện tái tổ hợp polyketit syntaza (PKS) octaketit syntaza (OKS) typ III thực vật (*Aloe arborescens* và/hoặc *Hypericum perforatum*) trong *Aspergillus nidulans* và đã xác định rằng OKS thực sự hoạt động trong vật chủ khác loại, tức là octaketit không khử được tạo ra *in vivo* trong tế bào *Aspergillus nidulans*.

Như đã thảo luận ở đây, axit flavokermesic (FK) cũng được xác định trong tế bào *Aspergillus nidulans*.

Các trình tự OKS *Aloe arborescens* sẵn có công khai được thể hiện ở đây với số SED ID sau đây:

Ở đây được gọi là *AaOKS*:SEQ ID NO:6 [Số ID ngân hàng gen AY567707 (nucleotit seq.) và SEQ ID NO:7: [Số ID ngân hàng gen AAT48709 (polypeptit seq)];

Ở đây được gọi là *AaOKS2*(PKS4):SEQ ID NO:8 [Số ID ngân hàng gen FJ536166] (nucleotit seq.) và SEQ ID NO:9 [Số ID ngân hàng gen ARC19997] (polypeptit seq);

Ở đây được gọi là *AaOKS3*(PKS5):SEQ ID NO:10 [Số ID ngân hàng gen FJ536167] (nucleotit seq.) và SEQ ID NO:11 [Số ID ngân hàng gen ARC19998] (polypeptit seq);

Ở đây được gọi là *AaPKS2_A207G*: SEQ ID NO:12 (nucleotit seq.) và SEQ ID NO:13: (polypeptit seq).

Trình tự axit amin OKS *Hypericum perforatum* sẵn có công khai được thể hiện ở đây với số SED ID sau đây:

Được gọi ở đây là *HpPKS2*: SEQ ID NO:14 [Số ID ngân hàng gen HQ529467] (nucleotit seq.) và SEQ ID NO:15: [Số ID ngân hàng gen AEE69029] (polypeptit seq);

Mức độ đồng nhất (%) giữa các PKS typ III ở mức axit amin:

	AaOKS	AaPKS2	AaOKS3	HpPKS2	AaPKS2_A207G
AaOKS	-	96,28%	89,88%	43,07%	91,32%
AaOKS2		-	90,12%	43,81%	91,07%
AaOKS3			-	45,07%	85,93%
HpPKS2				-	43,56%
AaPKS2_A207G					-

Có thể được ưu tiên là octaketit syntaza (OKS) ở bước (I) theo khía cạnh thứ nhất là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:7.

Có thể được ưu tiên là octaketit syntaza (OKS) ở bước (I) theo khía cạnh thứ nhất là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:9.

Có thể được ưu tiên là octaketit syntaza (OKS) ở bước (I) theo khía cạnh thứ nhất là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất

80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 405 của SEQ ID NO:11.

Có thể được ưu tiên là octaketit syntaza (OKS) ở bước (I) theo khía cạnh thứ nhất là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:13.

Có thể được ưu tiên là octaketit syntaza (OKS) ở bước (I) theo khía cạnh thứ nhất là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 393 của SEQ ID NO:15.

Bước (I) - chất khác

Như đã thảo luận ở trên, bước (I) (i) theo khía cạnh thứ nhất là: “các đơn vị khởi động và các đơn vị kéo dài thích hợp (tiếp xúc) với OKS đã được đưa vào bằng cách tái tổ hợp có khả năng chuyển hóa các đơn vị khởi động và kéo dài thành octaketit không khử trong các điều kiện thích hợp trong đó octaketit không khử được sản xuất *in vivo*”.

Như đã thảo luận ở trên, các đơn vị khởi động và các đơn vị kéo dài thích hợp đã được biết trong lĩnh vực.

Theo lĩnh vực kỹ thuật, các đơn vị khởi động thích hợp, ví dụ, có thể là axetyl-CoA hoặc malonyl-CoA và các đơn vị kéo dài thích hợp, ví dụ, có thể là malonyl-CoA.

Việc xác định các điều kiện sinh trưởng thích hợp trong đó octaketit không khử được sản xuất *in vivo* như đã biết bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật là công việc thông thường đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật, các điều kiện sinh trưởng thích hợp này thường sẽ phụ thuộc vào tế bào chủ tái tổ hợp được sử dụng đặc hiệu.

Tế bào chủ tái tổ hợp và OKS thuộc giống khác nhau

Như đã thảo luận ở trên, bước (I) theo khía cạnh thứ nhất là: “trong tế bào chủ tái tổ hợp chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ, cho tiếp xúc *in vivo*”

Tốt nhất ở đây là sử dụng tế bào chủ tái tổ hợp được ưu tiên như đã thảo luận trong bản mô tả này và OKS được ưu tiên thuộc giống khác như đã thảo luận ở đây.

Ví dụ, phương án được ưu tiên ở đây đề cập đến phương án, trong đó:

- tế bào chủ tái tổ hợp là tế bào chủ được chọn từ nhóm bao gồm: *Aspergillus* (tốt hơn là *Aspergillus nidulans*) và *Nicotiana* sp. (tốt hơn là *Nicotiana benthamiana*); và

- OKS thuộc giống khác với tế bào chủ là OKS được chọn từ nhóm bao gồm: OKS từ *Aloe* spp. (tốt hơn là *Aloe arborescens*) và *Hypericum* spp. (tốt hơn là *Hypericum perforatum*).

Tốt hơn là, octaketit syntaza (OKS) từ *Aloe* spp. là:

- OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:7; hoặc

- OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:9; hoặc

- OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 405 của SEQ ID NO:11; hoặc

OKS chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:13.

Tốt hơn là, octaketit syntaza (OKS) từ *Hypericum* spp. là OKS chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 393 của SEQ ID NO:15.

Hợp chất aglycon thơm cần quan tâm – bước (II) theo khía cạnh thứ nhất

Như đã thảo luận ở trên, bước (II) theo khía cạnh thứ nhất là: “chuyển hóa *in vivo* trong tế bào chủ tái tổ hợp octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm, trong đó hợp chất aglycon thơm cần quan tâm không phải là SEK4 và/hoặc SEK4B”.

Theo phương án được ưu tiên, hợp chất aglycon thơm cần quan tâm là hợp chất aglycon thơm C₁₄-C₁₈ cần quan tâm.

Tốt hơn là hợp chất aglycon thơm cần quan tâm là ít nhất là một hợp chất được chọn từ nhóm bao gồm: hợp chất anthraquinon và hợp chất anthron.

Trong nhóm nêu trên, được ưu tiên là hợp chất aglycon thơm cần quan tâm của bước (II) là hợp chất anthraquinon.

Theo lĩnh vực kỹ thuật, anthraquinon (còn được biết là anthraquinonoit) là nhóm hợp chất phenolic dựa trên khung 9,10-anthraquinon (xem Fig.4 trong bản mô tả này).

Ví dụ và như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật, axit flavokermesic (FK) và axit kermesic (KA) là các ví dụ về anthraquinon.

Anthron là nhóm hợp chất phenolic dựa trên khung anthron (xem fig.4 trong bản mô tả này).

Fig.4 thể hiện cấu trúc hợp chất FK antron— tức là ví dụ về hợp chất anthron theo sáng chế.

Tốt hơn là, hợp chất aglycon thơm cần quan tâm là hợp chất aglycon thơm C₁₆ cần quan tâm.

Theo phương án được ưu tiên, hợp chất aglycon thơm C₁₆ cần quan tâm là axit flavokermesic (FK) hoặc axit kermesic (KA).

Như đã thảo luận ở trên, sự chuyển hóa trong bước (II) của octaketit không khử thành hợp chất aglycon thơm C₁₄-C₃₄ (tốt hơn là C₁₄-C₁₈) cần quan tâm thường được thực hiện bằng sự tham gia của các enzym sau PKS thích hợp (ví dụ, aromataza và/hoặc xyclaza) có mặt *in vivo* trong tế bào chủ.

Do đó, có thể là tốt hơn, nếu sự chuyển hóa *in vivo* trong bước (II) của phương pháp theo khía cạnh thứ nhất của octaketit không khử thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm được thực hiện nhờ sự tham gia của ít nhất một aromataza/xyclaza.

Tốt hơn là, nó được thực hiện nhờ sự tham gia của ít nhất một aromataza/xyclaza hoạt động *in trans* (không phụ thuộc enzym PKS).

Có thể là tốt hơn, nếu bước (II) của phương pháp theo khía cạnh thứ nhất là:

(II): chuyển hóa *in vivo* trong tế bào chủ tái tổ hợp octaketit không khử của bước (I) thành hợp chất aglycon thơm cần quan tâm, trong đó sự chuyển hóa *in vivo*

của octaketit không khử thành hợp chất aglycon thơm cần quan tâm được thực hiện nhờ sự tham gia của ít nhất một aromataza/xyclaza hoạt động *in trans* (không phụ thuộc enzym PKS) và trong đó aromataza/xyclaza là từ nấm hoặc vi khuẩn và khác loại với tế bào chủ tái tổ hợp và từ giống khác với PKS.

SEQ ID NO: 17 trong bản mô tả này là trình tự amino sẵn có công khai của *Streptomyces* ZhuI aromataza/xyclaza (Số truy cập ngân hàng gen AAG30197) và SEQ ID NO: 19 trong bản mô tả này là trình tự amino sẵn có công khai của *Streptomyces* ZhuJ aromataza/xyclaza (Số truy cập ngân hàng gen AAG30196) (xem ví dụ tài liệu được thảo luận ở đây của Tang, Y. et al (2004)).

Như đã thảo luận ở dưới, các aromataza/xyclaza này được sử dụng thành công trong các ví dụ thử nghiệm ở đây.

Có thể là tốt hơn, nếu aromataza/xyclaza là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 169 của SEQ ID NO:17.

Có thể là tốt hơn, nếu aromataza/xyclaza là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 256 của SEQ ID NO:19.

Hợp chất glycosit thơm cần quan tâm – bước (III) theo khía cạnh thứ nhất

Như đã thảo luận ở trên, tùy ý bước (III) theo khía cạnh thứ nhất là: “tế bào chủ tái tổ hợp cũng chứa gen glycosyltransferaza mã hóa glycosyltransferaza và hợp chất aglycon thơm cần quan tâm được sản xuất trong bước (II) được glycosyl hóa *in vivo* bằng glycosyltransferaza để sản xuất hợp chất glycosit thơm C₁₄-C₃₄ cần quan tâm”.

Có thể là tốt hơn, nếu bước (III) này được thực hiện (tức là, không tùy ý).

Theo phương án được ưu tiên, hợp chất glycosit thơm cần quan tâm là hợp chất glycosit thơm C₁₄-C₁₈ cần quan tâm.

Như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này – hợp chất aglycon thơm được ưu tiên cần quan tâm của bước (II) có thể được chuyển hóa thành hợp chất glycosit thơm được ưu tiên cần quan tâm tương ứng ở bước (III).

Tốt hơn là hợp chất glycosit thơm cần quan tâm là ít nhất là một hợp chất được chọn từ nhóm bao gồm: hợp chất anthraquinon, hợp chất anthron.

Trong nhóm nêu trên, tốt hơn nếu hợp chất glycosit thơm cần quan tâm của bước (III) là hợp chất anthraquinon.

Ví dụ và như đã hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật - DcII và axit carminic (CA) là các ví dụ về anthraquinon.

Tốt hơn là, hợp chất glycosit thơm cần quan tâm là hợp chất glycosit thơm C₁₆ cần quan tâm.

Theo phương án được ưu tiên, hợp chất glycosit thơm C₁₆ cần quan tâm là glycosit của axit flavokermesic hoặc glycosit của axit kermesic.

Theo phương án được ưu tiên, glycosit của axit flavokermesic là DcII.

Theo phương án được ưu tiên, glycosit của axit kermesic là axit carminic (CA).

Glycosyltransferaza (GT) của bước (III) có thể, ví dụ, là GT được mô tả ở đây từ *Dactylopius coccus*.

Theo cách khác, nó có thể, ví dụ, là các glycosyltransferaza được thảo luận ở đây từ cây *Aloe*, cây *Haworthia*, cây *Sorghum* và/hoặc lúa.

Tốt hơn là, glycosyltransferaza là glycosyltransferaza polypeptit có khả năng:

(I): liên hợp glucoza được hoạt hóa bởi nucleotit với axit flavokermesic (FK); và/hoặc

(II): liên hợp glucoza được hoạt hóa bởi nucleotit với axit kermesic (KA).

Trình tự polynucleotit mã hóa *Dactylopius coccus costa* glycosyltransferaza mới được phân lập/tách dòng được mô tả ở đây được thể hiện trong SEQ ID NO: 1 ở đây và trình tự axit amin polypeptit được thể hiện trong SEQ ID NO: 2 ở đây.

Enzym glycosyltransferaza theo sáng chế của SEQ ID NO: 2 có thể được gọi ở đây là “DcUGT2”.

Do đó, có thể được ưu tiên là glycosyltransferaza là glycosyltransferaza polypeptit có khả năng:

(I): liên hợp glucoza được hoạt hóa bởi nucleotit với axit flavokermesic (FK); và/hoặc

(II): liên hợp glucoza được hoạt hóa bởi nucleotit với axit kermesic (KA);

và trong đó glycosyltransferaza polypeptit là ít nhất một polypeptit được chọn từ nhóm bao gồm:

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 70% với các axit amin từ 1 đến 515 của SEQ ID NO:2;

(b) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 70% với các axit amin từ 1 đến 468 của SEQ ID NO:2;

(c) polypeptit mà được mã hóa bởi polynucleotit lai trong các điều kiện nghiêm ngặt ít nhất là trung bình với (i) các nucleotit từ 1 đến 1548 của SEQ ID NO:1 hoặc (ii) chuỗi bổ sung của (i); và

(d) mảnh của các axit amin từ 1 đến 515 của SEQ ID NO:2, có hoạt tính glycosyltransferaza như được mô tả ở (I) và/hoặc (II).

Phương án được ưu tiên đề cập đến phương án, trong đó glycosyltransferaza polypeptit là:

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 80% với các axit amin từ 1 đến 515 của SEQ ID NO:2; tốt hơn nữa là

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 90% với các axit amin từ 1 đến 515 của SEQ ID NO:2; thậm chí tốt hơn nữa là

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 95% với các axit amin từ 1 đến 515 của SEQ ID NO:2; và tốt nhất là

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 98% với các axit amin từ 1 đến 515 của SEQ ID NO:2.

Có thể được ưu tiên là glycosyltransferaza polypeptit theo khía cạnh thứ nhất là polypeptit chứa trình tự axit amin với các axit amin từ 1 đến 515 của SEQ ID NO:2.

Như đã thảo luận trong bản mô tả này, số liệu/kết quả đã xác định của ví dụ thử nghiệm 4 thể hiện rằng các enzym GT theo sáng chế có thể được nhận diện, ví dụ, trong *Sorghum* và cây lúa.

Trình tự polypeptit *Sorghum* (số ID ngân hàng gen: AAF17077.1) được thể hiện dưới dạng SEQ ID NO: 4 ở đây.

Trình tự polypeptit của lúa (số ID ngân hàng gen: CAQ77160.1) được thể hiện dưới dạng SEQ ID NO: 5 ở đây.

Có thể thích hợp là glycosyltransferaza là glycosyltransferaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 492 của SEQ ID NO:4.

Có thể thích hợp là glycosyltransferaza là glycosyltransferaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 471 của SEQ ID NO:5.

Theo cách khác, các glycosyltransferaza cũng có thể là glycosyltransferaza từ cây *Haworthia*.

Phân lập hợp chất thơm cần quan tâm – bước (IV)

Như đã thảo luận ở trên, bước (IV) theo khía cạnh thứ nhất là: “phân lập hợp chất aglycon thơm cần quan tâm của bước (II) và/hoặc phân lập hợp chất glycosit thơm cần quan tâm của bước (III) để thu được chế phẩm, trong đó chế phẩm này chứa ít hơn 1% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp.

Trong ngữ cảnh này, người có hiểu biết trung bình trong lĩnh vực kỹ thuật thường biết cách phân lập hợp chất thơm cần quan tâm đặc hiệu theo sáng chế liên quan đến bước (IV), tức là, không cần phải thảo luận chi tiết hơn về vấn đề này.

Theo phương án được ưu tiên của phương pháp như được mô tả ở đây, trong đó chế phẩm đã phân lập ở bước (IV) chứa hợp chất glycosit thơm cần quan tâm và trong đó hợp chất glycosit thơm cần quan tâm là glycosit của axit flavokermesic (tốt hơn là DcII) và/hoặc glycosit của axit kermesic (tốt hơn là axit carminic (CA)).

Có thể tốt hơn, nếu có chế phẩm tương đối tinh khiết, do đó có thể là tốt hơn, nếu chế phẩm này chứa ít nhất 10% khối lượng/khối lượng chất khô của hợp chất thơm cần quan tâm và trong đó chế phẩm này chứa ít hơn 0,5% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp; hoặc chế phẩm này chứa ít nhất 50% khối lượng/khối lượng chất khô của hợp chất thơm cần quan tâm và trong đó chế phẩm này chứa ít hơn 0,1% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp.

Nếu trong bước (IV) thu được chế phẩm chứa hợp chất aglycon thơm cần quan tâm đã được phân lập (ví dụ, axit flavokermesic (FK) hoặc axit kermesic (KA)), phương pháp theo khía cạnh thứ nhất có thể bao gồm bước bổ sung:

(IVa): glycosyl hóa hợp chất aglycon thơm cần quan tâm để sản xuất hợp chất glycosit thơm cần quan tâm.

Bước này có thể được thực hiện bằng quy trình tổng hợp hóa học theo lĩnh vực này.

Theo cách khác, nó có thể được thực hiện bằng cách sử dụng glycosyltransferaza có khả năng glycosyl hóa aglycon trong các điều kiện thích hợp, trong đó aglycon glycosit được sản xuất ra.

Một phương án của sáng chế đề cập là hợp chất aglycon thơm cần quan tâm của chế phẩm được phân lập của bước (IV) là axit flavokermesic (FK) và/hoặc axit kermesic (KA) và phương pháp theo khía cạnh thứ nhất bao gồm bước bổ sung là:

(IVa) cho tiếp xúc:

(a1): axit flavokermesic (FK) với glycosyltransferaza có khả năng glycosyl hóa axit flavokermesic trong các điều kiện thích hợp, trong đó glycosit của axit flavokermesic được sản xuất ra; và/hoặc

(a2): axit kermesic (KA) với glycosyltransferaza có khả năng glycosyl hóa axit kermesic trong các điều kiện thích hợp, trong đó glycosit của axit kermesic được sản xuất ra.

Theo phương án được ưu tiên, glycosit của axit flavokermesic là DcII.

Theo phương án được ưu tiên, glycosit của axit kermesic là axit carminic (CA).

Như đã thảo luận ở đây, glycosyltransferaza (GT) của bước này có thể, ví dụ, là GT được mô tả ở đây từ *Dactylopius coccus*.

Theo cách khác, nó có thể là, ví dụ, glycosyltransferaza được thảo luận ở đây từ cây *Aloe*, cây *Haworthia*, *Sorghum* và/hoặc cây lúa.

Các khía cạnh và phương án được ưu tiên-trong ngôn ngữ/dạng yêu cầu bảo hộ

(Các) khía cạnh được mô tả ở đây và các phương án được ưu tiên của chúng có thể được trình bày trong ngôn ngữ/dạng yêu cầu bảo hộ. Việc này được thực hiện dưới đây đối với một số khía cạnh được mô tả ở đây và các phương án được ưu tiên của chúng.

1. Phương pháp sản xuất hợp chất thơm thu được từ octaketit cần quan tâm, trong đó phương pháp này bao gồm các bước sau:

(I): trong tế bào chủ tái tổ hợp chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ, cho tiếp xúc *in vivo*:

(i): đơn vị khởi động và đơn vị kéo dài thích hợp với OKS đã được đưa vào bằng cách tái tổ hợp có khả năng chuyển hóa đơn vị khởi động và kéo dài thành octaketit không khử trong các điều kiện thích hợp mà trong đó octaketit không khử được sản xuất ra; và

(II): chuyển hóa *in vivo*, trong tế bào chủ tái tổ hợp, octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm, trong đó hợp chất aglycon thơm cần quan tâm không phải là SEK4 và/hoặc SEK4B; và

(III): tùy ý, tế bào chủ tái tổ hợp còn chứa gen glycosyltransferaza mã hóa glycosyltransferaza và hợp chất aglycon thơm đã sản xuất ở bước (II) cần quan tâm được glycosyl hoá *in vivo* bởi glycosyltransferaza để sản xuất ra hợp chất glycosit thơm C₁₄-C₃₄ cần quan tâm; và

(IV): phân lập hợp chất aglycon thơm cần quan tâm của bước (II) và/hoặc phân lập hợp chất glycosit thơm cần quan tâm của bước (III) để thu được chế phẩm, trong đó chế phẩm này chứa ít hơn 1% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp.

2. Phương pháp theo điểm 1, trong đó tế bào chủ tái tổ hợp là tế bào thực vật, tế bào nấm sợi, tế bào nấm men hoặc tế bào tiền nhân.

3. Phương pháp theo điểm 2, trong đó tế bào chủ tái tổ hợp là tế bào thực vật.

4. Phương pháp theo điểm 3, trong đó tế bào thực vật là tế bào thực vật được chọn từ nhóm bao gồm: *Nicotiana sp.* (ví dụ, tế bào *Nicotiana benthamiana*); cây đại hoàng, kiều mạch, *Hypericum* và *Aloe sp.*

5. Phương pháp theo điểm 4, trong đó tế bào thực vật là *Nicotiana sp.*

6. Phương pháp theo điểm 5, trong đó tế bào thực vật là *Nicotiana benthamiana*.

7. Phương pháp theo điểm 2, trong đó tế bào chủ tái tổ hợp là tế bào nấm sợi.

8. Phương pháp theo điểm 7, trong đó tế bào nấm sợi là *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, hoặc *Trichoderma*.

9. Phương pháp theo điểm 8, trong đó tế bào nấm sợi là tế bào *Aspergillus*.

10. Phương pháp theo điểm 9, trong đó tế bào *Aspergillus* là *Aspergillus niger*, *Aspergillus oryzae* hoặc *Aspergillus nidulans*.

11. Phương pháp theo điểm 10, trong đó tế bào *Aspergillus* là *Aspergillus nidulans*.

12. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ thực vật, vi khuẩn hoặc nấm.

13. Phương pháp theo điểm 12, trong đó gen polyketit syntaza (PKS) typ III của bước (I) là gen PKS từ thực vật.

14. Phương pháp theo điểm 13, trong đó thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe spp.* (tốt hơn là *Aloe arborescens*), *Hypericum spp.* (tốt hơn là *Hypericum perforatum*), cây đại hoàng, kiều mạch và *Hawortia spp.*

15. Phương pháp theo điểm 14, trong đó thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe spp.* và *Hypericum spp.*

16. Phương pháp theo điểm 15, trong đó thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe arborescens* và *Hypericum perforatum*.

17. Phương pháp theo điểm 13, trong đó octaketit syntaza (OKS) ở bước (I) theo điểm 1 là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:7.

18. Phương pháp theo điểm 13, trong đó octaketit syntaza (OKS) ở bước (I) theo điểm 1 là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:9.

19. Phương pháp theo điểm 13, trong đó octaketit syntaza (OKS) ở bước (I) theo điểm 1 là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 405 của SEQ ID NO:11.

20. Phương pháp theo điểm 13, trong đó octaketit syntaza (OKS) ở bước (I) theo điểm 1 là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 403 của SEQ ID NO:13.

21. Phương pháp theo điểm 13, trong đó octaketit syntaza (OKS) ở bước (I) theo điểm 1 là OKS chứa trình tự axit amin có độ đồng nhất ít nhất là 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 393 của SEQ ID NO:15.

22. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó đơn vị khởi động thích hợp là axetyl-CoA và/hoặc malonyl-CoA.

23. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó các đơn vị kéo dài thích hợp là malonyl-CoA.

24. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó:

- tế bào chủ tái tổ hợp là tế bào chủ được chọn từ nhóm bao gồm: *Aspergillus* (tốt hơn là *Aspergillus nidulans*) và *Nicotiana sp.* (tốt hơn là *Nicotiana benthamiana*); và

- OKS thuộc giống khác với tế bào chủ là OKS được chọn từ nhóm bao gồm: OKS từ *Aloe spp.* (tốt hơn là *Aloe arborescens*) và *Hypericum spp.* (tốt hơn là *Hypericum perforatum*).

25. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó hợp chất aglycon thơm cần quan tâm là hợp chất aglycon thơm C₁₄-C₁₈ cần quan tâm.

26. Phương pháp theo điểm 25, trong đó hợp chất aglycon thơm cần quan tâm là ít nhất một hợp chất được chọn từ nhóm bao gồm: hợp chất anthraquinon và hợp chất anthron.

27. Phương pháp theo điểm 26, trong đó hợp chất aglycon thơm cần quan tâm là anthraquinon và anthraquinon là axit flavokermesic (FK) hoặc axit kermesic (KA).

28. Phương pháp theo điểm 25, trong đó hợp chất aglycon thơm cần quan tâm là hợp chất aglycon thơm C₁₆ cần quan tâm.

29. Phương pháp theo điểm 28, trong đó hợp chất aglycon thơm C₁₆ cần quan tâm là axit flavokermesic (FK) hoặc axit kermesic (KA).

30. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó sự chuyển hóa *in vivo* trong bước (II) theo điểm 1 của octaketit không khử thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm được thực hiện nhờ sự tham gia của ít nhất một aromataza/xyclaza.

31. Phương pháp theo điểm 30, trong đó sự chuyển hóa *in vivo* trong bước (II) theo điểm 1 của octaketit không khử thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm được thực hiện nhờ sự tham gia của ít nhất một aromataza/xyclaza hoạt động *in trans* (không phụ thuộc enzym PKS).

32. Phương pháp theo điểm 31,

trong đó aromataza/xyclaza này là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm

chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 169 của SEQ ID NO:17 (*Streptomyces* ZhuI); và/hoặc

trong đó aromataza/xyclaza này là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 256 của SEQ ID NO:19 (*Streptomyces* ZhuJ).

33. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó bước (III) được thực hiện (nghĩa là, không tùy ý) – nghĩa là, trong đó bước (III) là: tế bào chủ tái tổ hợp còn chứa gen glycosyltransferaza mã hóa glycosyltransferaza và hợp chất aglycon thơm cần quan tâm được sản xuất trong bước (II) được glycosyl hóa *in vivo* bởi glycosyltransferaza để sản xuất hợp chất glycosit thơm C₁₄-C₃₄ cần quan tâm.

34. Phương pháp theo điểm 33, trong đó hợp chất glycosit thơm cần quan tâm là hợp chất glycosit thơm C₁₄-C₁₈ cần quan tâm.

35. Phương pháp theo điểm 34, trong đó hợp chất glycosit thơm cần quan tâm là hợp chất glycosit thơm C₁₆ cần quan tâm.

36. Phương pháp theo điểm 35, trong đó hợp chất glycosit thơm C₁₆ cần quan tâm là glycosit của axit flavokermesic hoặc glycosit của axit kermesic.

37. Phương pháp theo điểm 36, trong đó hợp chất glycosit thơm C₁₆ cần quan tâm là glycosit của axit flavokermesic và glycosit của axit flavokermesic là DcII.

38. Phương pháp theo điểm 36, trong đó hợp chất glycosit thơm C₁₆ cần quan tâm là glycosit của axit kermesic và glycosit của axit kermesic là axit carminic (CA).

39. Phương pháp theo điểm bất kỳ trong các điểm từ 33 đến 38, trong đó glycosyltransferaza (GT) của bước (III) là GT từ *Dactylopius coccus*, GT từ các cây *Aloe*, GT từ các cây *Haworthia*, GT từ cây *Sorghum* hoặc GT từ cây lúa.

40. Phương pháp theo điểm 36, trong đó glycosyltransferaza (GT) là glycosyltransferaza polypeptit có khả năng:

(I): liên hợp glucoza được hoạt hóa bởi nucleotit với axit flavokermesic (FK); và/hoặc

(II): liên hợp glucoza được hoạt hóa bởi nucleotit với axit kermesic (KA).

41. Phương pháp theo điểm bất kỳ trong các điểm từ 33 đến 40, trong đó glycosyltransferaza là glycosyltransferaza polypeptit có khả năng:

(I): liên hợp glucoza được hoạt hóa bởi nucleotit với axit flavokermesic (FK); và/hoặc

(II): liên hợp glucoza được hoạt hóa bởi nucleotit với axit kermesic (KA);

và trong đó glycosyltransferaza polypeptit là ít nhất một polypeptit được chọn từ nhóm bao gồm:

(a) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 70% với các axit amin từ 1 đến 515 của SEQ ID NO:2;

(b) polypeptit chứa trình tự axit amin có độ đồng nhất ít nhất 70% với các axit amin từ 1 đến 468 của SEQ ID NO:2;

(c) polypeptit mà được mã hóa bởi polynucleotit lai trong các điều kiện nghiêm ngặt ít nhất là trung bình với (i) các nucleotit từ 1 đến 1548 của SEQ ID NO:1 hoặc (ii) chuỗi bổ sung của (i); và

(d) mảnh của các axit amin từ 1 đến 515 của SEQ ID NO:2, có hoạt tính glycosyltransferaza như được mô tả ở (I) và/hoặc (II).

42. Phương pháp theo điểm bất kỳ trong các điểm từ 33 đến 40, trong đó glycosyltransferaza là glycosyltransferaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 492 của SEQ ID NO:4.

43. Phương pháp theo điểm bất kỳ trong các điểm từ 33 đến 40, trong đó glycosyltransferaza là glycosyltransferaza chứa trình tự axit amin có độ đồng nhất ít nhất 70% (tốt hơn là ít nhất 80%, tốt hơn nữa là ít nhất 90% và thậm chí tốt hơn nữa là ít nhất 98%) với các axit amin từ 1 đến 471 của SEQ ID NO:5.

44. Phương pháp theo điểm bất kỳ trong các điểm từ 33 đến 43, trong đó chế phẩm đã phân lập ở bước (IV) chứa hợp chất glycosit thơm cần quan tâm và trong đó hợp chất glycosit thơm cần quan tâm là glycosit của axit flavokermesic (tốt hơn là DcII) và/hoặc glycosit của axit kermesic (tốt hơn là axit carminic (CA)).

45. Phương pháp theo điểm 44, trong đó hợp chất glycosit thơm cần quan tâm là glycosit của axit flavokermesic và glycosit của axit flavokermesic là DcII.

46. Phương pháp theo điểm 44, trong đó hợp chất glycosit thơm cần quan tâm là glycosit của axit kermesic và glycosit của axit kermesic là axit carminic (CA).

47. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó chế phẩm được phân lập trong bước (IV) theo điểm 1 chứa ít nhất 10% khối lượng/khối

lượng chất khô của hợp chất thơm cần quan tâm và trong đó chế phẩm này chứa ít hơn 0,5% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp.

48. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó trong bước (IV), thu được chế phẩm chứa hợp chất aglycon thơm được phân lập cần quan tâm (ví dụ, axit flavokermesic (FK) hoặc axit kermesic (KA)) và phương pháp này bao gồm thêm bước:

(IVa): glycosyl hóa hợp chất aglycon thơm cần quan tâm để sản xuất hợp chất glycosit thơm cần quan tâm.

49. Phương pháp theo điểm 48, trong đó bước glycosyl hóa hợp chất aglycon thơm cần quan tâm được thực hiện bằng cách sử dụng glycosyltransferaza có khả năng glycosyl hóa aglycon trong các điều kiện thích hợp, trong đó aglycon glycosit được sản xuất ra.

50. Phương pháp theo điểm 49, trong đó hợp chất aglycon thơm cần quan tâm của chế phẩm được phân lập của bước (IV) là axit flavokermesic (FK) và/hoặc axit kermesic (KA) và phương pháp theo khía cạnh thứ nhất bao gồm thêm bước:

(IVa): cho tiếp xúc:

(a1): axit flavokermesic (FK) với glycosyltransferaza có khả năng glycosyl hóa axit flavokermesic trong các điều kiện thích hợp, trong đó glycosit của axit flavokermesic được sản xuất ra; và/hoặc

(a2): axit kermesic (KA) với glycosyltransferaza có khả năng glycosyl hóa axit kermesic trong các điều kiện thích hợp, trong đó glycosit của axit kermesic được sản xuất ra.

51. Phương pháp theo điểm bất kỳ trong các điểm từ 48 đến 50, trong đó glycosit của axit flavokermesic là DcII.

52. Phương pháp theo điểm bất kỳ trong các điểm từ 48 đến 50, trong đó glycosit của axit kermesic là axit carminic (CA).

53. Phương pháp theo điểm bất kỳ trong các điểm trước, trong đó tế bào chủ tái tổ hợp theo điểm 1 là tế bào chủ tái tổ hợp đang sinh trưởng và bước (I) và bước (II) theo điểm 1 là:

(I): trong tế bào chủ tái tổ hợp đang sinh trưởng chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS) và trong đó OKS thuộc giống khác với tế bào chủ, cho tiếp xúc *in vivo*:

(i): đơn vị khởi động và đơn vị kéo dài thích hợp với OKS đã được đưa vào bằng cách tái tổ hợp có khả năng chuyển hóa đơn vị khởi động và kéo dài thành octaketit không khử trong các điều kiện sinh trưởng thích hợp, trong đó octaketit không khử được sản xuất *in vivo*; và

(II): chuyển hóa *in vivo*, trong tế bào chủ tái tổ hợp đang sinh trưởng, octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ cần quan tâm, trong đó hợp chất aglycon thơm cần quan tâm không phải là SEK4 và/hoặc SEK4B.

Ví dụ thực hiện sáng chế

Thông tin về trình tự thích hợp có thể được tìm thấy trong tài liệu Danh mục trình tự được cung cấp trong bản mô tả này

Ví dụ 1 – Tách dòng *D. coccus* GT và thử nghiệm hoạt tính FK và KA của nó

Vật liệu và phương pháp

Tinh chế ADN và mARN

Dactylopius coccus đông lạnh tươi (thu được từ Lanzarote). *Porphyrophora polonica* đông lạnh tươi thu được từ Ba Lan. Côn trùng đông lạnh được nghiền thành bột trong nitơ lỏng và ADN/ARN được tinh chế: ADN được tinh chế bằng cách sử dụng bộ kit DNeasy Blood & Tissue (Qiagen), theo phương pháp chuẩn của nhà cung cấp. ARN được tinh chế bằng cách sử dụng bộ kit mini RNeasy (Qiagen) theo phương pháp chuẩn của nhà cung cấp.

mARN của tế bào nhân chuẩn được chuyển thành cADN bằng cách sử dụng bộ kit RT² Easy First Strand (Qiagen) theo phương pháp chuẩn của nhà cung cấp bằng cách sử dụng kỹ thuật môi poly-dT của phản ứng transcriptaza ngược.

Giải trình tự ADN và ARN:

ADN và cADN được gửi đi để giải trình tự tại BGI (Shenzen, Trung Quốc) để giải trình tự bằng cách sử dụng công nghệ Illumina đầu cùng tạo cặp 100 bp theo phương pháp chuẩn của Illumina ở độ quét khoảng 60-100X và kết quả theo định dạng dữ liệu fastq.

Phân tích trình tự ADN và ARN/cADN:

Trình tự fastq thu được của ADN và ARN/cADN được nhập vào Genomic Workbench phiên bản 5.4 (CLC-bio, Århus, Đan Mạch) và được lắp ráp bằng cách sử dụng thuật toán lắp ráp *de novo* thành các contig. Tập kết quả được xuất ra ở định dạng FASTA.

Tệp FASTA về ARN/cADN sau đó được nhập vào IOGMA v. 10 (Genostar, Grenoble, Pháp) và các gen giả định được xác định bằng cách sử dụng “thiết bị dò gen của tế bào tiền nhân dựa trên mô hình Markov-Matrix ẩn.

Các gen giả định được chú thích bằng cách sử dụng BLAST (basic local alignment sequence tool: công cụ sắp xếp thẳng hàng trình tự khu trú cơ bản) dựa vào ngân hàng gen (NCBI) bằng cách sử dụng cả trình tự nucleotit làm trình tự protein được dịch mã. Các gen giả định cũng được chú thích bằng cách so sánh độ tương đồng với cơ sở dữ liệu PFAM của họ protein.

Tạo các đoạn protein từ *D. coccus*

Ba gam côn trùng *D. coccus* tươi được đồng nhất hóa trong 120mL chất đệm phân lập [sucroza 350mM, Trixin 20mM (độ pH = 7,9), NaCl 10mM, DTT 5mM, PMSF 1mM) chứa 0,3g polyvinylpolypyrrolidon. Dịch đồng nhất được lọc qua vải nylon (mắt lưới 22µm) và được ly tâm trong (10 phút, 10000 x g ở nhiệt độ 4°C). Phần dịch nổi trên bề mặt được ly tâm (1 giờ, 105000xg, ở nhiệt độ 4°C), thu được phần protein hòa tan và phần protein liên kết màng. Phần protein hòa tan được cô thành 1mL và được trao đổi chất đệm với Trixin 20mM (độ pH = 7,9), DTT 5mM bằng cách sử dụng thiết bị Amicon ultrafugation-3K (Millipore). Pelet protein liên kết màng được rửa 3 lần bằng cách tạo huyền phù lại pelet này trong 60mL Trixin 20mM (độ pH = 7,9), DTT 5mM bằng cách sử dụng chổi quét sơn lông chồn sau đó là ly tâm lại. Pelet protein liên kết màng cuối cùng được tạo huyền phù lại trong 1mL Trixin 20mM (độ pH = 7,9), DTT 5mM. Phần protein hòa tan và phần protein liên kết màng được phân tích về hoạt tính glycosyl hóa.

Tinh chế hoạt tính GT đặc hiệu axit flavokermesic/axit kermesic từ protein màng *D. coccus*

Phần protein liên kết màng được phân lập từ 3g côn trùng *D. coccus* tươi được hòa tan bằng cách bổ sung Triton x-100 1% (thể tích/thể tích) (dạng được khử) và khuấy nhẹ trong 1,5 giờ trong môi trường lạnh. Dung dịch được xử lý Triton x-100 được ly tâm (1 giờ, 105000xg, ở nhiệt độ 4°C) và phần dịch nổi trên bề mặt được phân lập và được phủ lên cột được đóng gói với 2mL dòng chảy Q-sepharose Fast (Pharmacia). Cột được rửa trong 4mL chất đệm A [Trixin 20mM (độ pH = 7,9), Triton x-100 0,1% (thể tích/thể tích) (dạng được khử), NaCl 50mM] và protein được rửa giải với Trixin 20mM (độ pH = 7,9), Triton x-100 0,1% (thể tích/thể tích) (dạng được khử)

bằng cách sử dụng gradien NaCl ngắt quãng từ 100mM-500mM, (với số gia 50mM). Các phần 0,5ml được gom, được khử muối, được phân tích bằng SDS-PAGE và được theo dõi về hoạt tính glucosyl hóa bằng cách sử dụng thử nghiệm enzym glucosyl hóa đánh dấu phóng xạ được mô tả. Phần chứa hoạt tính GT đặc hiệu axit flavokermesic/axit kermesic được làm giàu được đưa vào phân tích vết khối lượng peptit.

Thử nghiệm enzym và phát hiện sản phẩm glucosit

Việc glucosyl hóa axit flavokermesic và axit kermesic được thực hiện trong hỗn hợp thử nghiệm gồm 60 μ L chứa Trixin 20mM (độ pH = 7,9), 3,3 μ M UDP[14C]glucoza và 20 μ L dịch chiết protein (protein liên kết màng hoặc protein hòa tan). Hỗn hợp phản ứng được ủ trong 0,5 giờ ở nhiệt độ 30°C và được kết thúc bằng cách bổ sung 180 μ L metanol. Các mẫu được ly tâm ở 16000xg trong 5 phút ở nhiệt độ 4°C và phần dịch nổi trên bề mặt được chấm trên đĩa TLC (đĩa silica gel 60 F254; Merck). Sản phẩm thử nghiệm được tách trong diclometan:metanol:axit formic (7:2:2, theo thể tích). Sản phẩm đã được đánh dấu phóng xạ được quan sát bằng mắt thường bằng cách sử dụng STORM 840 PhosphorImager (Molecular Dynamics, <http://www.moleculardynamics.com>).

Sự biểu hiện của DcUGT2, DcUGT4 và DcUGT5 được tối ưu hóa codon trong *S. cerevisiae*

Dạng được tối ưu hóa codon tổng hợp của DcUGT2 và hai trình tự GT giả định khác từ transcriptome (tập hợp tất cả các gen được biểu hiện) của *Dactylopius coccus* được gọi là DcUGT4 và DcUGT5 để biểu hiện nấm men được mua từ GenScript với các vị trí attL tái tổ hợp gateway bên sườn. Các đoạn tổng hợp được sử dụng làm khuôn PCR với các đoạn mồi đặc hiệu để tạo ra các dạng được gắn đuôi StrepII đầu cùng C tương ứng. Sáu cấu trúc gen (đoạn có đuôi và không có đuôi) được tách dòng vào plasmit đích gateway pYES-DEST52 (Invitrogen) bằng cách sử dụng hỗn hợp enzym LR clonaseII. Sáu cấu trúc plasmit pYES-DEST52 được biến nạp riêng vào chủng nấm men Invsc1 (Invitrogen) và các thể biến nạp dương tính được xác nhận bằng PCR. Việc sản xuất protein khác loại được thực hiện theo hướng dẫn của vector Gateway pYES-DEST52 (Invitrogen). Việc sản xuất protein đuôi StrepII khác loại được xác nhận bằng kỹ thuật thâm tách Western bằng cách sử dụng kháng thể kháng Strep. Phần protein liên kết màng được tạo ra từ thể biến nạp nấm men đã được xác

nhận như được mô tả trong (D. Pompon, B. Louerat, A. Bronine, P. Urban, Yeast expression of animal and plant P450s in optimized redox environments, *Methods Enzymol.* 272 (1996) 51–64) và được sàng lọc về hoạt tính glucosyl hóa đối với axit flavokermesic/axit kermesic. Trình tự được tối ưu hóa của nấm men được thể hiện trong SEQ ID NO: 3 ở đây.

LC-MS-MS

LC/MS được thực hiện trên Agilent Q-TOF với hệ thống HPLC sau:

Cột Kinetix 2,6 μ XB-C18 100A (100 x 4,60mm, Phenomenex). Dung môi A (900ml nước khử ion, 100ml metanol và 50ml axit formic). Dung môi B (700ml metanol, 300ml nước khử ion và 50ml axit formic).

Tốc độ dòng chảy 0,8ml/phút. 35°C.

Gradient rửa giải. 0-1 phút 100% A. Gradient tuyến tính đến 83% A 3 phút. Gradient tuyến tính đến 63% A 6 phút, gradient tuyến tính đến 45% A 9 phút, gradient tuyến tính đến 27% A 12 phút, gradient tuyến tính đến 10% A 15 phút, gradient tuyến tính đến 3% A 17 phút, gradient tuyến tính đến 2% A 19 phút, gradient tuyến tính đến 0% A 20 phút, 0% A 22 phút, gradient tuyến tính đến 100% A 25 phút.

Thời gian lưu là 7,6 phút đối với axit carminic, 7,8 phút đối với DC II, 13,7 phút đối với axit flavokermesic và 13,9 phút đối với axit kermesic.

Kết quả:

Khả năng glycosyl hóa axit flavokermesic/axit kermesic bằng cách sử dụng C14-UDP-glucosa làm cơ chất được phát hiện ra ở côn trùng *D. coccus* đã được đồng nhất hóa. Hoạt tính này được thể hiện là liên kết màng và hoạt tính này được tinh chế và protein đã tinh chế được đưa vào phân tích protein ở quy mô lớn. Đã được chỉ ra là hoạt tính enzym là do polypeptit có trình tự tương ứng với gen ứng viên DcUGT2 của các tác giả sáng chế.

Như đã thảo luận ở trên, enzym glycosyltransferaza theo sáng chế của SEQ ID NO: 2 có thể được gọi ở đây là “DcUGT2”.

Trình tự axit amin của DcUGT2 cho thấy độ đồng nhất ít hơn 45% với glycosyl transferaza đã biết bất kỳ.

Biết được rằng việc tách dòng trình tự kiểu đại vào nấm men không tạo ra hoạt tính enzym thích hợp, các tác giả sáng chế đã thiết kế lại trình tự nucleotit của DcUGT2 thành trình tự mã hóa cùng một polypeptit nhưng sử dụng các codon

nucleotit được tối ưu hóa cho *S. cerevisiae*, quy trình này được gọi là tối ưu hóa codon (trình tự được tối ưu hóa *S. cerevisiae* được thể hiện là SEQ ID No. 3 trong bản mô tả này). Tiếp theo, trình tự được tối ưu hóa codon của DcUGT2 được tách dòng và biểu hiện trong nấm men. Chủng nấm men khác loại chứa hoạt tính enzym liên kết màng có khả năng glucosyl hóa axit kermesic và axit flavokermesic.

Sau khi thu được dữ liệu vết khối lượng peptit từ phần protein *Dactylopius coccus* được làm giàu hoạt tính GT đối với axit flavokermesic/axit kermesic, các tác giả sáng chế so khớp các khối lượng peptit với tập dữ liệu transcriptome và xác định ba UGT giả định (DcUGT2, DcUGT4 và DcUGT5).

Sự biểu hiện khác loại của ba ứng viên trong nấm men cho thấy rằng chỉ một trong các UGT này, cụ thể là DcUGT2 chịu trách nhiệm về hoạt tính glucosyl hóa quan sát được đối với axit flavokermesic/axit kermesic trong phần protein *D. coccus*.

Việc xử lý viscozyme của glucosit được đánh dấu phóng xạ C-14 đã tạo ra, cho thấy rằng nó chống lại phản ứng thủy phân, ngoài ra còn gợi ý là DcUGT2 là C-GT, chịu trách nhiệm sản xuất DCII và axit carminic.

LC-MS-MS cho thấy sự tạo thành các sản phẩm với thời gian lưu, phổ, phân tử khối và mô hình thoái biến phân tử lần lượt giống như DCII và axit carminic.

Kết luận

Kết quả của Ví dụ 1 này đã chứng minh rằng việc phân lập/tách dòng enzym glycosyltransferaza theo sáng chế của SEQ ID NO: 2, mà được gọi ở đây là “DcUGT2” hoặc đơn giản là “DcUGT” không phải là một nhiệm vụ dễ dàng.

Ví dụ, các trình tự gen đã được xác định của bộ gen và transcriptome của côn trùng *D. coccus* được phân tích về độ tương đồng với các trình tự C-glycosyltransferaza đã biết công khai theo sáng chế và kết quả là âm tính theo nghĩa là không có trình tự gen đã được xác định nào của bộ gen/transcriptome cho thấy độ tương đồng đáng kể với các trình tự C-glycosyltransferaza đã biết công khai theo sáng chế.

Tuy nhiên, ngay cả khi phân tích độ tương đồng trình tự tin sinh học có thể được cho là chỉ ra rằng bộ gen của *Dactylopius coccus* không chứa gen mã hóa glycosyltransferaza theo sáng chế – các tác giả sáng chế tiếp tục nghiên cứu vấn đề đó và các tác giả sáng chế xác định được dịch chiết *Dactylopius coccus* (bao gồm dịch chiết của tế bào nội cộng sinh có mặt trong *D. coccus*) có hoạt tính GT theo sáng chế

và bằng cách kết hợp bước tinh chế và thử nghiệm theo sáng chế, các tác giả sáng chế cuối cùng cũng có thể có được phân/chế phẩm tương đối tinh khiết mà từ đó có thể thu được một vài trình tự axit amin một phần của các ứng viên enzym GT có thể có.

Các tác giả sáng chế thử nghiệm hoạt tính của glycosyltransferaza mới được phân lập/tách dòng được mô tả ở đây của SEQ ID NO:2 (DcUGT2) và phát hiện ra rằng có thể liên hợp glucoza với các aglycon axit flavokermesic (FK) và axit kermesic (KA) – xem Fig.1 trong bản mô tả này.

Ví dụ 2 thử nghiệm hoạt tính KA GT của UrdGT2 đã biết trước đó

Như đã thảo luận ở trên, UrdGT2 được mô tả trong tài liệu Baig et al (Angew Chem Int Ed Engl. 2006 Nov 27;45(46):7842-6).

Như đã thảo luận ở trên, tài liệu này mô tả rằng UrdGT2 có khả năng glycosyl hóa các phân tử aglycon khác nhau mà có thể được coi là tương tự về mặt cấu trúc với aglycon axit kermesic (KA) và axit flavokermesic (FK) theo sáng chế.

Dạng tổng hợp được tối ưu codon của UrdGT2 để biểu hiện *E. coli* được tách dòng và được biểu hiện tái tổ hợp trong *E. coli*. Dịch chiết protein thô hòa tan chứa UrdGT2 tái tổ hợp thu được, tức là, dịch chiết chứa UrdGT2

Hoạt tính GT của UrdGT2 được phân tích *in vitro* bằng cách sử dụng UDP-glucoza hoặc TDP-glucoza làm chất cho đường và FA/KA làm cơ chất aglycon. Không phát hiện thấy hoạt tính đối với các aglycon này, tức là, không có hoạt tính GT theo sáng chế được xác định liên quan đến các aglycon này.

Tuy nhiên, đã xác nhận được là UrdGT2 tái tổ hợp có hoạt tính, như được chứng minh bởi sự tạo thành *in vitro* glucosit được đánh dấu phóng xạ C14 thu được từ quá trình glucosyl hóa hợp chất không được xác định trong dịch chiết *E. coli* thô.

Ví dụ 3 Hoạt tính GT ở cây *Aloe* và cây *Haworthia*

Phân lập và thử nghiệm hoạt tính GT từ *Aloe*

- 1) Cây được rửa sạch hạt đất và được tách thành: A) Rễ, B) Mô lá xanh và C) vật liệu gel từ lá
- 2) 5g mô được làm đông lạnh ngay lập tức trong nitơ lỏng và được nghiền thành bột mịn trong cối lạnh bằng chày.
- 3) 20mL chất đệm chiết lạnh [Trixin 20mM-HCl, NaCl 10mM, DTT 5mM, PMSF 1mM, độ pH = 7,9] chứa chất ức chế proteaza đầy đủ không có

EDTA (Roche), proteamin sulfat 0,1% (khối lượng/thể tích) và 0,5g PVPP được bổ sung vào bột và được khuấy.

4) Dịch đồng nhất được khuấy nhẹ ở nhiệt độ 4°C trong 10 phút và sau đó được ly tâm ở 12000 xg ở nhiệt độ 4°C trong 5 phút.

5) Phần dịch nổi trên bề mặt được phân lập và 1mL proteamin sulfat 2% (khối lượng/thể tích) trong Trixin 20mM-HCl, độ pH = 7,9 được bổ sung từng giọt vào trong 2 phút ở nhiệt độ 4°C trong điều kiện khuấy liên tục.

6) Phần dịch nổi trên bề mặt được lọc qua 2 miếng lưới nylon. Phần dịch nổi trên bề mặt đã được lọc sau đó được ly tâm ở 12000 xg ở nhiệt độ 4°C trong 5 phút.

7) Phần dịch nổi trên bề mặt được phân lập và được siêu ly tâm ở 110000 xg ở nhiệt độ 4°C trong 1 giờ.

8) Phần protein hòa tan (phần dịch nổi trên bề mặt) được phân lập và được trao đổi chất đệm 5 lần với Trixin 20mM-HCl, độ pH = 7,9 chứa DTT 5mM bằng cách sử dụng thiết bị lọc siêu ly tâm Amicon -3K (Millipore)

9) 20µL dịch chiết protein hòa tan được ủ trong tổng thể tích phản ứng bằng 60µL chứa UDP-glucosa (nồng độ cuối cùng 1,25mM) và FK (nồng độ cuối cùng 50µM), KA (nồng độ cuối cùng 50µM) hoặc MeO-FK/EtO-FK (nồng độ cuối cùng 50µM/50µM) trong 2 giờ ở nhiệt độ 30°C, lắc ở 650 vòng/phút.

10) Phản ứng enzym được kết thúc bằng 180µL metanol lạnh và được lọc qua thiết bị lọc 0,45 micron và được đưa vào phân tích HPLC-MS.

Giá trị m/z [M-H] ----- Protein hòa tan <i>Aloe</i>	475 m/z [M-H] ⁻ FK-monoglc	491 m/z [M-H] ⁻ KA-monoglc	489 m/z [M-H] ⁻ MeOFK- monoglc	503 m/z [M-H] ⁻ EtOFK-monoglc
Lá	3,73	3,71	5,81	6,63
Gel				
Rễ		3,71		

Bảng 1. Glucosit được tạo thành trong thử nghiệm glucosyl hóa *in vitro* bằng cách sử dụng dịch chiết enzym từ *Aloe*.

Dịch chiết enzym thô hòa tan của ba mô *Aloe*, vật liệu lá xanh (Lá), vật liệu gel từ lá (Gel) và rễ được thử nghiệm về hoạt tính glucosyl hóa đối với axit flavokermesic (FK), axit kermesic (KA), metyl este của axit flavokermesic (MeOFK) và etyl este của axit flavokermesic (EtOFK). Các ký tự số tương ứng với thời gian lưu (phút) sau khi tách bằng HPLC-MS các glucosid mới được tạo thành *in vitro* (Bảng 1).

Giá trị m/z 475 và 491 lần lượt là các giá trị m/z giống như thu được đối với DcII và CA, được hòa tan trong các dung dịch tương tự. Cả hai giá trị m/z đều lớn hơn 162 (giá trị m/z của glucoza trong glucosid) so với giá trị m/z của FK và KA chỉ ra rằng gốc glucoza từ UDP-glucoza trong chất đệm phản ứng đã được chuyển đến aglycon bởi GT trong dịch chiết. Giá trị m/z $[M-H]^-$ 489 và 503 cũng lần lượt lớn hơn 162 so với các giá trị m/z thu được với MeOFK và EtOFK, chỉ ra rằng đơn vị glucoza đã được bổ sung vào cả MeOFK và EtOFK bởi GT có mặt trong dịch chiết.

Phân lập và thử nghiệm hoạt tính GT từ *Haworthia limifolia*

Quy trình này là như được mô tả đối với *Aloe* nhưng mô cây được phân tích như sau: A) Mô lá xanh, B) Vật liệu gel từ lá, C) Mô gốc (phần màu hồng giữa rễ và thân) và D) Mô rễ.

Dịch chiết enzym thô hòa tan của bốn mô *Haworthia limifolia*, vật liệu lá xanh (Lá), vật liệu gel từ lá (Gel), mô màu hồng giữa rễ và thân (Gốc) và rễ được thử nghiệm về hoạt tính glucosyl hóa đối với axit flavokermesic (FK), axit kermesic (KA), metyl este của axit flavokermesic (MeOFK) và etyl este của axit flavokermesic (EtOFK). Các ký tự số tương ứng với thời gian lưu (phút) sau khi tách bằng HPLC-MS các glucosid mới được tạo thành *in vitro* (Bảng 2).

Giá trị m/z $[M-H]^-$	475 m/z $[M-H]^-$ FK-monogluc	491 m/z $[M-H]^-$ KA-monogluc	489 m/z $[M-H]^-$ MeOFK-monogluc	503 m/z $[M-H]^-$ EtOFK-monogluc

Protein hòa tan <i>Haworthia</i>				
Lá	3,73	3,71	5,81	6,63
Gel				
Gốc	3,73	3,71	5,81	6,63
Rễ	3,73	3,71	5,81	6,63

Bảng 2. Glucosid được tạo thành trong thử nghiệm glucosyl hóa *in vitro* bằng cách sử dụng dịch chiết enzym từ *Haworthia limifolia*.

Giá trị m/z 475 và 491 lần lượt là các giá trị m/z giống như thu được đối với DcII và CA, được hòa tan trong các dung dịch tương tự. Cả hai giá trị m/z đều lớn hơn 162 (giá trị m/z của glucoza trong glucosit) so với các giá trị m/z của FK và KA chỉ ra rằng gốc glucoza từ UDP-glucoza trong chất đệm phản ứng đã được chuyển đến aglycon bởi GT trong dịch chiết. Giá trị m/z $[M-H]$ 489 và 503 cũng lần lượt lớn hơn 162 so với các giá trị m/z thu được với MeOFK và EtOFK, chỉ ra rằng đơn vị glucoza đã được bổ sung vào cả MeOFK và EtOFK bởi GT có mặt trong dịch chiết.

Kết luận

Các kết quả của ví dụ này chỉ ra rằng enzym glycosyltransferaza theo sáng chế (GT) có thể được xác định trong cây *Aloe* và cây *Haworthia*.

Nói cách khác, cây *Aloe* và cây *Haworthia* chứa glycosyltransferaza có khả năng glycosyl hóa axit flavokermesic để tạo ra glycosit của axit flavokermesic; và/hoặc có khả năng glycosyl hóa axit kermesic để tạo ra glycosit của axit kermesic.

Ví dụ 4 hoạt tính GT ở cây *Sorghum* và cây lúa

Như đã biết trong lĩnh vực này, cây *Sorghum* và cây lúa chứa glycosyltransferaza.

Như đã biết trong lĩnh vực này, một số glycosyltransferaza của cây *Sorghum* và cây lúa có thể glycosyl hóa hợp chất aglycon phân tử lượng thấp.

Glycosyltransferaza được mô tả trong lĩnh vực này từ cây *Sorghum* và cây lúa có độ đồng nhất ít hơn 70% với các axit amin từ 1 đến 515 của SEQ ID NO:2 như được bộc lộ ở đây.

Chưa được biết trong lĩnh vực này là các glycosyltransferaza của cây *Sorghum* và/hoặc cây lúa có phải là các glycosyltransferaza theo sáng chế hay không, tức là glycosyltransferaza có khả năng glycosyl hóa axit flavokermesic để tạo ra glycosit của axit flavokermesic; và/hoặc có khả năng glycosyl hóa axit kermesic để tạo ra glycosit của axit kermesic.

Các glycosyltransferaza đã biết từ cây *Sorghum* (*Sorghum bicolor*), SbUGT85B1, với Số ID ngân hàng gen AF199453.1 (nucleotit seq.) / AAF17077.1 (polypeptit seq) và cây lúa (*Oryza sativa*), OsCGT, với Số ID ngân hàng gen FM179712.1 (nucleotit seq.) / CAQ77160.1 (polypeptit seq) được biểu hiện trong *E.coli* chủng Xjb và dịch chiết protein *E.coli* thô được tạo ra và được thử nghiệm về

hoạt tính glucosyl hóa đối với các cơ chất axit kermesic và axit flavokermesic như được mô tả bởi Kannangara *et al.* (2011) và Augustin *et al.* (2012).

Fig.2 trong PCT/EP2014/078540 (ngày nộp đơn PCT là 18/12/2014) thể hiện các phân tích LC-MS của các sản phẩm được glucosyl hóa đã tạo thành trong thử nghiệm chứa dịch dung giải thô của *E.coli* chủng Xjb biểu hiện SbUGT85B1 hoặc OsCGT, UDP-glucoza và axit flavokermesic (FK) hoặc axit kermesic (KA). Dịch chiết thô từ *E.coli* chủng Xjb được sử dụng làm đối chứng âm trong thử nghiệm.

Các KA glycosit được xác định (491 m/z [M-H] – giá trị m/z[M-H] của CA) đối với cả hai glycosyltransferaza và FK glycosit được xác định (475 m/z [M-H] - giá trị m/z[M-H] của DcII) đối với OsCGT.

Kết luận

Kết quả của ví dụ này đã chứng minh rằng các enzym glycosyltransferaza theo sáng chế (GT) có thể được xác định trong cây *Sorghum* và/hoặc cây lúa.

Nói cách khác, cây *Sorghum* và/hoặc cây lúa chứa glycosyltransferaza mà có khả năng glucosyl hóa axit flavokermesic để tạo ra glycosit của axit flavokermesic; và/hoặc có khả năng glucosyl hóa axit kermesic để tạo ra glycosit của axit kermesic.

Ví dụ 5 Sử dụng gen GT nội sinh hoặc hoạt tính GT nội sinh

Như đã biết trong lĩnh vực này, các glycosyltransferaza có thể glucosyl hóa hợp chất phân tử lượng thấp có mặt trong nhiều sinh vật khác nhau. Phương pháp cho glycosyltransferaza của tế bào sinh vật tiếp xúc với hợp chất phân tử lượng thấp là đưa vào một hoặc nhiều gen điều khiển sự sinh tổng hợp của hợp chất phân tử lượng thấp và nhờ đó cho phép các tế bào này glucosyl hóa hợp chất phân tử lượng thấp này. Hợp chất phân tử lượng thấp này có thể là, ví dụ, axit flavokermesic hoặc axit kermesic hoặc các dạng biến đổi của các phân tử này.

Một hoặc nhiều gen điều khiển sự sinh tổng hợp của axit flavokermesic hoặc axit kermesic hoặc dạng được biến đổi của các phân tử này được đưa vào sinh vật chứa glycosyltransferaza, ví dụ, cây thuốc lá, *Nicotiana benthamiana*.

Khi gen/các gen được biểu hiện tạm thời theo các phương pháp được mô tả trong (D'Aoust et al. (Methods Mol Biol 483 (2009): 41-50) trong ví dụ, mô cây, hợp chất hoặc các hợp chất phân tử lượng thấp được tạo ra. Các tế bào biểu hiện ổn định gen/các gen được tạo ra và được chọn theo các phương pháp được mô tả trong Gelvin (Microbiol Mol Biol Rev 67(1) (2003): 16-37)).

Trong các tế bào chứa gen/các gen được biểu hiện ổn định và/hoặc được biểu hiện tạm thời, hợp chất phân tử lượng thấp trở thành tiếp xúc với glycosyltransferaza nội sinh, dẫn đến tạo ra một hoặc nhiều glycosit của axit flavokermesic, axit kermesic hoặc các dạng được biến đổi của các phân tử này.

Sự có mặt của các glycosit đã được chứng minh bằng phương pháp chiết và phương pháp phân tích được mô tả trong Ví dụ 3.

Các mẫu được tạo ra cho LC/MS bằng phương pháp chiết được mô tả bởi (Rauwald and Sigler (Phytochemical Analysis 5 (1994):266-270).

Kết luận

Các kết quả của ví dụ này chứng minh rằng các glycosyltransferaza nội sinh có mặt trong tế bào của sinh vật tái tổ hợp có thể được sử dụng để chuyển hóa axit flavokermesic, axit kermesic hoặc các dạng được biến đổi của các phân tử này thành glycosit khi gen/các gen điều khiển sự sinh tổng hợp của các aglycon được đưa vào sinh vật đó.

Nói cách khác, việc đưa vào gen hoặc các gen điều khiển sự sinh tổng hợp của axit flavokermesic, axit kermesic, các dạng được biến đổi của các phân tử này, hoặc các hợp chất phân tử lượng thấp có liên quan là phương pháp cho hợp chất phân tử lượng thấp tiếp xúc với các glycosyltransferaza và vì thế là phương pháp sản xuất các glycosit của axit flavokermesic, axit kermesic hoặc các dạng được biến đổi của các hợp chất này.

Ví dụ 6: Biểu hiện ổn định PKS typ III trong *Aspergillus nidulans*

Vật liệu và phương pháp

Môi trường và dung dịch

Tất cả các dung dịch được điều chế bằng Milli-Q H₂O và được vô trùng ở nhiệt độ 121°C trong 20 phút.

Dung dịch nguyên tố vi lượng: Đối với 1 L trộn 0,4 g CuSO₄·5H₂O; 0,04 g Na₂B₄O₇·10H₂O; 0,8 g FeSO₄·7H₂O; 0,8 g MnSO₄·2H₂O; 0,8 g Na₂MoO₄·2H₂O; và 8,0 g ZnSO₄·7H₂O.

Hỗn hợp khoáng chất (50x): Đối với 1 L dung dịch trộn 26 g KCl; 26 g MgSO₄·7H₂O; và 76 g KH₂PO₄; và 50mL Dung dịch nguyên tố vi lượng.

Dung dịch muối nitrat 20X: Đối với 1 L dung dịch: hòa tan 120 g NaNO₃, 10,4 g KCl, 10,4 g MgSO₄* H₂O, 30,4 g KH₂PO₄ trong nước Milli-Q.

Dung dịch nguyên tố vi lượng (1 L dung dịch gốc): Đối với 1 L dung dịch: hòa tan 0,4 g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0,04 g $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$, 0,8 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,8 g $\text{MnSO}_4 \cdot 2\text{H}_2\text{O}$, 0,8 g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 8 g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, đến 1 L nước Milli-Q.

Thiamin 1%: Nồng độ cuối cùng 0,001%.

D-glucosa 20% (khối lượng/thể tích): Đối với 1 L dung dịch 200 g D-glucosa trong nước Milli-Q

Chất đệm tạo tế bào trần *Aspergillus* (APB): Nồng độ cuối cùng MgSO_4 1,1M và chất đệm Na-phosphat 10mM. Độ pH được điều chỉnh bằng NaOH 2N đến 5,8.

Chất đệm biến nạp *Aspergillus* (ATB): Nồng độ cuối cùng: Sorbitol 1,2 M; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 50mM; Tris 20mM; và KCl 0,6 M. Độ pH được điều chỉnh bằng HCl 2N đến 7,2.

PCT: Nồng độ cuối cùng: 50% khối lượng/thể tích PEG 8000 (4000, 6000 và PEG khác cũng có thể được sử dụng); CaCl_2 50mM; Tris 20mM; và KCl 0,6M. Độ pH được điều chỉnh bằng HCl 2N đến 7,5.

Môi trường tối thiểu (MM): Đối với 1L dung dịch: 1ml nguyên tố vi lượng; 50ml muối nitrat (1M); 50ml glucosa 20% khối lượng/thể tích; 1mL Thiamin; 20g thạch Agar (So.Bi gel). Đối với MM dạng lỏng thì thạch Agar không được bổ sung vào.

Thành phần bổ sung vào môi trường: Các thành phần bổ sung được bổ sung vào các môi trường khác nhau nếu cần, dựa vào kiểu gen của chủng *Aspergillus nidulans*, với các lượng sau: Arginin 4 mM, uraxil 10 mM và uridin 10 mM. Để chọn lọc ngược chỉ thị *Afp_{pyrG}*, axit 5-florotic (5-FOA) được bổ sung đến nồng độ bằng 1,3mg/mL.

Môi trường biến nạp rắn (TM): Đối với 1L dung dịch: 1ml nguyên tố vi lượng; 50ml muối nitrat (1M); Sucroza 171,15g; 1mL Thiamin; 20g thạch Agar (So.Bi gel). Đối với MM dạng lỏng thì thạch Agar không được bổ sung vào.

Biến nạp *Aspergillus nidulans*

Các tế bào trần *Aspergillus nidulans* được tạo ra theo phương pháp chuẩn của các giải pháp đã biết, được tóm tắt ngắn gọn ở đây: môi trường lớn *Aspergillus nidulans* được cho nảy mầm qua đêm, và sinh khối thu được được thu hoạch bằng cách lọc qua thiết bị lọc Miracloth (Merck Millipore). Thể sợi được tạo huyền phù lại trong 10ml chất đệm APB chứa 40mg Glucanex/ml (Novozymes A/S). Hỗn hợp này được ủ trên thiết bị lắc ở nhiệt độ 37°C với 150 vòng/phút trong 3 giờ. APB được bổ

sung vào để thu được tổng thể tích là 40ml. Lớp phủ 5ml ATB 50% và nước MQ 50% được phủ lên và các ống được ly tâm ở 3000 RCF, 16°C, trong 12 phút tạo ra hệ hai pha có tế bào trần trong kỳ trung gian. Các tế bào trần được rửa bằng cách sử dụng 40mL ATB và ly tâm ở 3000 RCF, 16°C, trong 12 phút. Pelet thu được được tạo huyền phù lại trong 1ml ATB.

Các tế bào trần tạo thành được sử dụng cho các thử nghiệm biến nạp di truyền ở các phần phân ước bằng 50µl được trộn với 1,5-5µg ADN và 150µL PCT. Hỗn hợp biến nạp được ủ trong 10 phút. Hỗn hợp này sau đó được bổ sung 250µL ATB và được chuyển vào đĩa biến nạp với các chất bổ sung cần thiết được chuyển bằng que trộn Drigalski.

Sau khi ủ trong 4 ngày ở nhiệt độ 37°C, thể biến nạp *Aspergillus nidulans* tạo thành được phân lập và được nuôi cấy cấp hai trên các đĩa thạch riêng với chế độ chọn lọc thích hợp.

Việc tích hợp hướng đích của catxet biểu hiện được phân tích bằng PCR sử dụng các đoạn mồi gốc được sử dụng để khuếch đại gen cần biểu hiện, sau đó là các cặp đoạn mồi với một đoạn mồi gắn mồi bên trong đoạn xen vào và một đoạn mồi trong bộ gen xung quanh. Trong trường hợp các chủng dựa trên AfpyrG, chỉ thị chọn lọc được loại bỏ, sau khi biến nạp, bằng cách chọn lọc ngược trên đĩa chứa 5-FOA và tái tổ hợp tương đồng giữa các đoạn lặp có hướng ngăn bao quanh gen chỉ thị AfpyrG trong catxet biểu hiện này. Sau khi loại bỏ gen chỉ thị, chủng này được xác nhận một lần nữa bằng PCR chẩn đoán. Việc loại bỏ gen chỉ thị chọn lọc cho phép chu kỳ biến nạp tiếp theo với gen đích mới, sử dụng cùng một gen chỉ thị chọn lọc như được sử dụng trong quá trình biến nạp thứ nhất.

Tạo cấu trúc chủng chủ *Aspergillus nidulans*

Chủng *Aspergillus nidulans* được sử dụng được liệt kê trong Bảng 3 dưới đây và Bảng 4 tóm tắt các locus/gen đã biến đổi. Việc làm mất đoạn (hoặc thay thế) gen đích đạt được bằng cách tạo cấu trúc catxet hướng đích gen, bao gồm gen chỉ thị chọn lọc tái sinh được bao quanh bởi hai trình tự khoảng 1500 bp giống hệt với các trình tự bao quanh locut mà nên được thay thế trong bộ gen. Cấu trúc hướng đích để thay thế các gen *wA* và *yA* được tạo cấu trúc bằng phương pháp dựa trên PCR chỉ thị phân tách được mô tả trong (Nielsen M.L., Albertsen L., Lettier G., Nielsen J.B., Mortensen M.H., 2006. Efficient PCR-based gene targeting with a recyclable marker

for *Aspergillus nidulans*. Fungal Genetics and Biology, Vol. 43:54-64). Đầu tiên, gen *wA* được hướng đích để làm khuyết và các trình tự hướng đích được yêu cầu được khuếch đại từ ADN bộ gen bằng cách sử dụng các đoạn mồi: AN*wA*-dl-Up-F (5'-GGAAGAAGGTCGCATACCA-3') được kết hợp với AN*wA*-dl-Up-Rad (5'-gatccccgggaattgccatgGATCAGGAGAAGGAGAGTCAAG-3') và AN*wA*-dl-Dw-Fad (5'-aattccagctgaccaccatgGGCGAAAAGGCAAAGGAGC-3') được kết hợp với AN*wA*-dl-Dw-R (5'-GCTAGAAAAGGCAAGGGAGG-3'). Hai đoạn gen chỉ thị này được khuếch đại bằng cách kết hợp đoạn mồi M1 (5'-catggcaattccccggggatc-GCCGGCAATTCTTTTATAGGTAGC-3') được kết hợp với đoạn mồi M2 (5'-CCAGAAGCAGTACACGGC-3') và đoạn mồi M3 (5'-GTTGTCTGCTTGCGCTTCTTC-3') với đoạn mồi M4 (5'-catggtggtcagctggaatTCCTCCGCCATTTCTTATTCCC-3'). Sau khi khuếch đại PCR các đoạn hướng đích gen và đoạn ADN chỉ thị, các đoạn này được dung hợp bằng PCR, như được mô tả trong Nielsen *et al.* 2006, các đoạn ADN này được tinh chế gel và được biến nạp vào chủng nhận *Aspergillus nidulans* NID1 (được mô tả trong Nielsen J.B, Nielsen M.L., Mortensen U.H, 2008, Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. Fungal Genetics and Biology, Vol. 45:165-170). Gen chỉ thị trong chủng tạo thành được loại bỏ bằng cách chọn lọc ngược trên các đĩa 5-FOA để xác định các thể biến nạp, trong đó gen chỉ thị *afpyrG* đã tạo vòng ngoài một cách tự phát. Việc làm khuyết hướng đích gen *yA* được thực hiện như được mô tả đối với gen *wA*, nhưng sử dụng hai cặp đoạn mồi sau đây để khuếch đại các trình tự hướng đích: Del-*yA*-5'-F (5'-GTGGGTTGAACCGCTTACTCAG-3') được kết hợp với Del-*yA*-5'-R (5'-gatccccgggaattgccatg-CCCGGAGGAATCAAAATGACGC-3') và Del-*yA*-3'-F (5'-aattccagctgaccaccatgGTTTGGGATTCTTAGGTGAGCTC-3') được kết hợp với Del-*yA*-3'-R (5'-CCTCCCTGGCGTATACACAAAC-3'). Chủng không chứa gen chỉ thị *afpyrG* tạo thành được gọi là NID598 trong phần mô tả tiếp theo.

Việc làm khuyết hướng đích asperthecin PKS (*aptA*) được thực hiện trong nền tảng NID598, sử dụng chiến lược thử nghiệm được mô tả cho việc làm khuyết gen *yA* và *wA*. Hai đoạn ADN hướng đích được khuếch đại PCR bằng cách sử dụng đoạn mồi AnAptA-UP-F (5'-GCTCGAGCTTGCCAGCC-3') được kết hợp với AnAptA-UP-R (5'-gatccccgggaattgccatg-GCTGGTGTGGGACACACG-3') và AnAptA-Dw-F (5'-

aattccagctgaccaccatgGCTTGGAAATCAGTATAGCTTTCTG -3') được kết hợp với AnAptA-Dw-R (5'- GCTTGTGGTCTGTCTGAATCG -3'). Cấu trúc hướng đích được tinh chế gel sau đó được biến nạp vào chủng NID598 không chứa gen chỉ thị, tạo ra chủng NID930.

Cụm *Apt*, cụm *mdp* và cụm *stc* lần lượt được làm khuyết trong nền tảng NID598. Đối với quy trình này, các cấu trúc hướng đích được yêu cầu được tạo cấu trúc bằng kỹ thuật tách dòng chất phản ứng cắt bỏ đặc hiệu uraxil định hướng (directional Uracil-Specific Excision Reagent Cloning - USER) của các đoạn hướng đích tương ứng vào vector P1 (P6), như được mô tả trong (Hansen B.G., Salomonsen B., Nielsen M.T., Nielsen J.B., Hansen N.B., Nielsen K.F., Regueira T.B., Nielsen J., Patil K.R., Mortensen U.H. 2011. Versatile enzym expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketit synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. Vol. 77(9):3044-51). Để làm khuyết cụm *Apt*, các đoạn hướng đích được khuếch đại với các đoạn mỗi cụm ANAPT-DI-Up-FU (5'-GGGTTTAAAdUGAGGAGCAGAGGATGCGG-3') được kết hợp với cụm ANAPT-DI-Up-RU (5'-GGACTTAAAdUGTAGTGGTGGTGGTGGTGG-3') và cụm ANAPT-DI-Dw-FU (5'-GGCATTAAdUCGCGTGGAATTTGGAAGAGAG-3') được kết hợp với cụm ANAPT-DI-Dw-RU (5'-GGTCTTAAAdUGTGCTCGGGGACGTGAAAG-3'). Mỗi đoạn mỗi được sử dụng bao gồm 2-deoxyuridin (dU) dựa vào để cho phép tạo ra đoạn nhô 3'. Các đoạn PCR tạo thành được tinh chế gel và được tách dòng định hướng bằng kỹ thuật tách dòng USER vào vector P1(P6) được cắt PacI/Nt.BbvCI. Plasmid tạo thành được cắt SmaI để giải phóng đoạn hướng đích, bao gồm gen chỉ thị *Afp_{pyr}G*. Plasmid đã cắt được biến nạp vào chủng NID598 không chứa gen chỉ thị. Gen chỉ thị *Afp_{pyr}G* sau đó được loại bỏ bằng kỹ thuật chọn lọc ngược 5-FOA, như được mô tả đối với *wA* nêu trên, để tạo ra chủng để làm khuyết cụm *mdp*. Việc làm khuyết được hướng đích cụm *mdp* và *stc* được thực hiện như được mô tả đối với cụm *apt*. Để tạo cấu trúc trình tự hướng đích cho cụm *mdp* được khuếch đại PCR bằng cách sử dụng đoạn mỗi cụm ANMDP-DI-Up-FU (5'-GGGTTTAAAdUGGTCGTCTGTCAAGGAGTTG-3') được kết hợp với đoạn mỗi cụm ANMDP-DI-Up-RU (5'-GGACTTAAAdUGCAGTGCTGTATATGGGTCTTG-3') và đoạn mỗi cụm ANMDP-DI-Dw-FU (5'-GGCATTAAdUCGCGTGGAATTTGGAAGAGAG-3') được kết hợp với cụm ANAPT-DI-Dw-RU (5'-GGTCTTAAAdUGTGCTCGGGGACGTGAAAG-3').

GGCATTAAdUGAGTTTGTGAGATGTTTCAGGATGG-3') kết hợp với đoạn mỗi cụm ANMDP-DI-Dw-RU (5'-GGTCTTAAdUGAGGTGAAGGACACAGCG-3'). Ngoài ra, trình tự hướng đích cho cụm *stc* được khuếch đại PCR với các đoạn mỗi; cụm ANSTC-DI-Up-FU (5'-GGGTTTAAAdUCGCAGAGACTAGGACACAAGTG-3') được kết hợp với cụm ANSTC-DI-Up-RU (5'-GGACTTAAdUGCGGCGATCTGTGGTAGAG-3') và cụm ANSTC-DI-Dw-FU (5'-GGCATTAAdUGCCAGCATATTCAAACCCAGTC-3') được kết hợp với cụm ANSTC-DI-Dw-RU (5'-GGTCTTAAdUCACACAACCAACCTCCGATC-3'). Chúng tạo thành có đột biến khuyết cụm *apt*, *mdp* và *stc* được gọi là NID_SMA.

Bảng 3. Các chủng *Aspergillus nidulans* khác nhau được sử dụng trong nghiên cứu

Tên chủng	Kiểu gen
NID1	<i>nkuAΔ argB2, pyrG89, veA1</i>
NID598	<i>nkuAΔ, argB2, pyrG89, veA1, wAΔ, yAΔ</i>
NID930	<i>nkuAΔ, argB2, pyrG89, veA1, wAΔ, yAΔ, aptAΔ</i>
NID_SMA	<i>nkuAΔ, argB2, pyrG89, veA1, wAΔ, yAΔ, aptΔ, mdpΔ, stcΔ</i>

Bảng 4. Các gen bị ảnh hưởng

Tên gen và số truy cập trong cơ sở dữ liệu-bộ gen- <i>Aspergillus</i>	
yA = AN6635	Sinh tổng hợp chất màu nang bào tử (Laccase)
wA = AN8209	Sinh tổng hợp chất màu nang bào tử (PKS)
aptA = AN6000	Asperthecin PKS
apt = aptA (AN6000) đến aptC (AN6002)	Cụm gen Asperthecin
mdp = mdpL (AN10023) đến mdpA (AN10021)	Cụm gen Monodictyphenone/emondin
stc = stcW (AN7804) đến stcA (AN7825)	Cụm gen Sterigmatocystin

Biểu hiện PKS typ III trong *Aspergillus nidulans*

Các dạng được tối ưu hóa codon tổng hợp (tổng hợp de novo) của *HpPKS2*, *AaOKS* được tạo ra để biểu hiện nấm men, và các trình tự có sử dụng codon tự nhiên được mua từ Genscript. Các gen được tối ưu hóa codon được ký hiệu với hậu tố 'ScOpt', ví dụ *HpPKS2-ScOpt*, trong khi các gen có sử dụng codon gốc từ vật chủ tự

nhiên được ký hiệu với hậu tố ‘Orig’, ví dụ, *HpPKS2-Orig*. Các đoạn ADN tổng hợp được sử dụng làm khuôn PCR cho các phản ứng PCR với các đoạn mồi đặc hiệu (IDT) cho các trình tự mã hóa của các gen bao gồm codon dừng. Các đoạn mồi được sử dụng có đoạn nhô 5’ phù hợp với việc tách dòng Chất phản ứng cắt bỏ đặc hiệu uraxil (USER™) của các amplicon PCR tạo thành vào các vector biểu hiện để kết hợp hướng đích các catxet biểu hiện vào một trong bảy vị trí có thể có trong bộ gen *Aspergillus nidulans*. Các plasmit có đặc trưng là gen chỉ thị chọn lọc của nấm (*Aspergillus fumigatus* *AfpYrG*), vị trí tách dòng USER (AsiSI và Nb.btsI) và được gắn bên sườn bởi hai trình tự ADN 1-1,5 kb *Aspergillus nidulans* (đặt tên là Ngược và Xuôi) để cho phép kết hợp vào bộ gen *Aspergillus nidulans* bằng cách tái tổ hợp tương đồng hướng đích (Hansen, B. G., Salomonsen, B., Nielsen, M. T., Nielsen, J. B., Hansen, N. B., Nielsen, K. F., ... Mortensen, U. H. (2011)). Hệ đặc trưng và biểu hiện enzyme toàn diện đối với *Aspergillus nidulans*, với gen polyketit syntaza của *Penicillium brevicompactum* từ cụm gen axit mycophenolic làm trường hợp kiểm thử. *Applied and Environmental Microbiology*, 77(9), 3044–3051. <http://doi.org/10.1128/AEM.01768-10>). Các amplicon PCR được khuếch đại bằng cách sử dụng PfuX7 ADN polymeraza và các đoạn ADN tạo thành được tinh chế gel. Các vector nhận được tạo ra để tách dòng USER™ bằng cách tiêu hóa nó với AsiSI và Nb.btsI qua đêm, sau đó tinh chế gel. Các amplicon PCR được tách dòng định hướng vào vector nhận bằng cách tách dòng USER™, kết hợp với biến nạp vào *E. coli*. Các tế bào DH5α được rửa đông trên đá, và hỗn hợp USER™ được bổ sung vào 50 µL tế bào (1×10^8 cfu/µg pUC19). Hỗn hợp được đặt trên đá trong 10 phút, và sốc nhiệt trong 90 giây ở nhiệt độ 45°C. Chuyển vào đá một lần nữa và ủ trong 5 phút. Các tế bào được dàn trên đĩa LB với kháng sinh chọn lọc và được ủ ở nhiệt độ 37°C qua đêm. Các khuẩn lạc được xác nhận bằng PCR và các thể biến nạp chính xác được nuôi trong môi trường LB lỏng với kháng sinh chọn lọc qua đêm để tinh chế plasmit (Taylor, R. G., Walker, D. C., & McInnes, R. R. (1993). Chúng vật chủ *E. coli* ảnh hưởng đáng kể đến chất lượng các chế phẩm ADN plasmit ở quy mô nhỏ được sử dụng để giải trình tự. *Nucleic Acids Research*, 21(7), 1677–8. Được truy hồi từ <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=309390&tool=pmcentrez&rendertype=abstract>).

Các đoạn môi được sử dụng cho quy trình tạo cấu trúc, trong đó dU là 2-deoxyuridin:

An_HpPKS2-ScOpt-F	5'-AGAGCGAdUATGGGTTTCCTTAGACAACGGTTC
An_HpPKS2-ScOpt-R	5'-TCTGCGAdUTCACAAAGGAACACTTCTCAAAACC
An_AaOKS-ScOpt-F	5'-AGAGCGAdUATGAGTTCCTCTCCAACGCTTCC
An_AaOKS-ScOpt-R	5'-TCTGCGAdUTCACATGAGAGGCAGGCTGTG
An_HpPKS2-Orig-F	5'-AGAGCGAdUGGGTTCCCTTGACAATGGTT
An_HpPKS2-Orig-R	5'-TCTGCGAdUTTAGAGAGGCACACTTCGGAGAAC
An_AaOKS-Orig-F	5'-AGAGCGAdUATGAGTAGTTTATCAAATGCCAGTC
An_AaOKS-Orig-R	5'-TCTGCGAdUTCACATCAATGGCAAGGAA

Các plasmit biểu hiện đã được xác nhận được tiêu hóa bằng *SwaI* (với gen chỉ thị *Afp_{pyr}G*), để giải phóng catxet biểu hiện được gắn bên sườn bởi các trình tự hướng đích xuôi và ngược. Catxet này được đưa vào tế bào trần *Aspergillus nidulans* như được xác định ở trên.

Các plasmit biểu hiện sau đây và được tạo cấu trúc:

pIS53(pyrG)::HpPKS2_ScOpt	HpPKS2_ScOpt trong locut IS53
pIS53(pyrG)::HpPKS2_Orig	HpPKS2_Orig trong locut IS53
pIS53(pyrG)::AaOKS_ScOpt	AaOKS_ScOpt trong locut IS53
pIS53(pyrG)::AaOKS_Orig	AaOKS_Orig trong locut IS53

Các catxet biểu hiện của từng plasmit được giải phóng bằng cách tiêu hóa *SwaI* (với gen chỉ thị *Afp_{pyr}G*) và được biến nạp tiếp vào tế bào trần *Aspergillus nidulans*.

Sau khi tạo cấu trúc của chủng cụ thể, gen chỉ thị *Afp_{pyr}G* đã sử dụng được loại bỏ bằng cách chọn lọc ngược 5-FOA.

Phân tích hóa học các chủng *Aspergillus nidulans* đã tạo cấu trúc

Để đánh giá tác dụng của các gen đã được đưa vào đối với sự chuyển hóa của chủng *Aspergillus nidulans* đã được tạo cấu trúc, chủng này được nuôi trên MM, với các chất bổ sung thích hợp, trong 5 đến 14 ngày ở nhiệt độ 37°C. Các chất chuyển hóa được chiết bằng quy trình chiết vi mô được mô tả bởi Smedsgaard (Smedsgaard, J. (1997). Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *Journal of Chromatography A*, 760(2), 264–270. Truy hồi từ <http://linkinghub.elsevier.com/retrieve/pii/S0021967396008035>). Tóm tắt: Mười nút tròn 0,4-0,6mm được cắt từ đĩa nuôi cấy bằng cách sử dụng ‘máy khoan

xoắn' và được đặt trong lọ thủy tinh 8ml chứa 2mL dung dịch chiết làm từ metanol và axit formic 1% (thể tích/thể tích). Các ống được nghiền bằng sóng siêu âm trong 60 phút trong bể nước. Phần dịch nổi trên bề mặt được chuyển vào lọ mới và được làm bay hơi trong dòng khí nitơ và nhiệt (30°C), và pelet tạo thành được hòa tan trong 100 μ L metanol cấp độ HPLC và axit formic 1% (Sigma-Aldrich) bằng cách ủ dung dịch trong 20 phút trong bể siêu âm. Dung dịch chiết thu được sau đó được lọc qua thiết bị lọc dạng xi lanh PTFE 0,45 μ m, 15mm Premium (Agilent Technologies) vào các lọ HPLC. Các chất chuyển hóa đã chiết được phân tích bằng cách sử dụng Agilent 1200 HPLC được lắp với máy đo phổ khối Bruker micrOTOF-Q II được trang bị nguồn ion hóa tia điện. Việc tách sắc ký đạt được trên cột Luna C₁₈ (2) (150 $\times$ 4,6mm, 3 μ m, 100 Å, Phenomenex) được duy trì ở nhiệt độ 40°C. Dung dịch nước rửa giải (A) chứa nước/axetonitril (95:5, thể tích/thể tích) và dung dịch rửa giải hữu cơ (B) chứa nước/axetonitril (5:95, thể tích/thể tích); cả hai đều được axit hóa bằng axit formic 0,1%. Tùy theo mục đích của phân tích, hai mẫu rửa giải gradien tuyến tính khác nhau được sử dụng:

Phương pháp 1: 0 phút, 0% B; 20 phút, 100% B; 22 phút 100% B; 24 phút, 0% B. Tốc độ dòng chảy được duy trì ở 0,8mL/phút và cân bằng 7 phút được sử dụng.

Phương pháp 2: 0 phút, 0% B; 30 phút, 100% B; 33 phút 100% B; 35 phút, 0% B. Tốc độ dòng chảy được duy trì ở 0,5mL/phút và cân bằng 10 phút được sử dụng.

Các mẫu được phân tích ở cả chế độ ion hóa dương và âm. Tập hợp dữ liệu thu được được phân tích bằng cách sử dụng phần mềm MasHunter (Agilent technologies) sau đây.:

Các sắc ký đồ đỉnh cơ sở (Base Peak Chromatograms: BPC) thu được được kiểm tra để xác định các thay đổi ở quá trình chuyển hóa của nấm. Trong khi các sắc ký đồ ion đã tách (Extracted Ion Chromatograms: EIC) cho các khối lượng phù hợp với các hợp chất trung gian mong đợi trong con đường sinh tổng hợp được tạo ra để dễ dàng xác nhận các đỉnh thích hợp.

Theo cách khác, quá trình sản xuất các chất chuyển hóa được phân tích trong môi trường nuôi cấy MM lỏng. Đối với quy trình này, bình vô trùng Erlenmeyer 500mL chứa 100mL môi trường tối thiểu lỏng, có các chất bổ sung thích hợp, được cấy bằng dung dịch chứa bào tử dính lớn từ các chủng cần được phân tích. Các bình được ủ trong 1-15 ngày ở nhiệt độ 37°C trong thiết bị ủ lắc với 150 vòng/phút. Thẻ sợi

được rót qua Miracloth (Merck Millipore) và được tách khỏi môi trường nuôi cấy. Thể sợi được làm đông lạnh nhanh bằng nitơ lỏng. Môi trường được lọc vô trùng qua FPE-214-250 JET BIOFIL®. Môi trường vô trùng này sau đó được chuyển vào lọ 40mL để phân tích hóa học.

Chiết

Canh thang lên men (40mL) được làm bay hơi đến khô trong chân không và được chiết bằng etanol + axit formic 10% (20mL) bằng phương pháp sóng âm (2 giờ) ở nhiệt độ 60°C. Dung môi được gạn và được làm khô trong chân không trước khi được hòa tan trong 1,5mL etanol + axit formic 10% bằng cách sử dụng kỹ thuật sóng âm. Phần dịch nổi trên bề mặt được làm khô trong máy cô mẫu Speedvac và được hòa tan trong 180 µL metanol 90% + axit formic 10%, được ly tâm và được phân tích bằng HPLC-MS/MS

Kết quả:

Phân tích các chất chuyển hóa đã tạo thành bằng UHPLC-HRMS của các chủng chứa *HpPKS2* hoặc *AaOKS* khác loại cho thấy sự biểu hiện của PKS typ III (*HpPKS2* hoặc *AaOKS*) dẫn đến việc tạo ra các hợp chất khác nhau bao gồm hợp chất SEK4, SEK4B và FK. Sự biểu hiện ở chủng khuyết cụm SMA tạo ra tỷ lệ cao hơn của FK và SEK4/SEK4B.

Kết luận:

Sự biểu hiện của PKS typ III thực vật (*HpPKS2* hoặc *AaOKS*) dẫn đến việc tạo ra các hợp chất khác nhau bao gồm hợp chất SEK4, SEK4B, FK anthron, Mutactin và FK *in vivo* trong *Aspergillus nidulans*. Sự biểu hiện ở chủng có các cụm gen asperthecin, sterigmatocystin và monodictyphenon/emodin đã được làm khuyết có tác động tích cực đến việc sản xuất chất chuyển hóa.

Vì trong Ví dụ 6 này không có các xyclaza và/hoặc aromataza khác loại được gắn xen vào chủng *Aspergillus* và hợp chất FK được xác định – điều đó chỉ ra rằng chủng *Aspergillus* có thể chứa các xyclaza và/hoặc aromataza đồng nhất có khả năng chuyển hóa *in vivo* octaketit không khử thành hợp chất FK.

Ví dụ 7: Sự biểu hiện ổn định các PKS typ III và các xyclaza và/hoặc aromataza trong *Aspergillus nidulans*

Vật liệu và phương pháp

Sự biểu hiện của xyclaza và aromataza vi khuẩn trong *Aspergillus nidulans*

Dạng được tối ưu hóa codon tổng hợp của *ZhuI*, *ZhuJ* cho sự biểu hiện *Aspergillus*, *S. cerevisiae* và *E. coli* được mua từ Genscript. Tương tự như được mô tả đối với PKS typ III, các gen mã hóa xyclaza và aromataza được tách dòng vào vector biểu hiện hướng đích các vị trí cụ thể trong bộ gen *Aspergillus*, khác với vị trí được sử dụng để biểu hiện của PKS typ III (xem Ví dụ 6). Nói cách khác, các gen *ZhuI* và *ZhuJ* được tách dòng vào chủng *Aspergillus* của Ví dụ 5 mà đã chứa gen PKS typ III khác loại (*HpPKS2*, *AaOKS*).

Các đoạn mồi được sử dụng cho quy trình tạo cấu trúc, trong đó dU là 2-deoxyuridin:

An_ZhuI_EcOpt-F 5'- AGATATACCA dUGCGTCATGTTGAACATACCGT
 An_ZhuI_EcOpt-R 5'- ATGGCTGC dUTTATGCGGTA ACTGTACCAACACCA
 An_ZhuJ_EcOpt-F 5'- ATATACATA dUGAGCGGTCGTAAAACCTTT
 An_ZhuJ_EcOpt-R 5'- ATATCCAAT dUTTAATCCTCTTCTTCTTGTTC
 An_ZhuI_ScOpt-F 5'-AGAGCGA dUGAGACACGTTGAACACA
 An_ZhuI_ScOpt-R 5'-TCTGCGA dUTTATGCAGTTACGGTACCA
 An_ZhuJ_ScOpt-F 5'-AGAGCGAUGTCCGGTAGAAAGACCTT
 An_ZhuJ_ScOpt-R 5'-TCTGCGA dUTTAATCTTCTTCTTCTTGTTCG

Trình tự mã hóa được khuếch đại PCR được tách dòng vào các vector để kết hợp hướng đích trong bộ gen *Aspergillus nidulans*: *ZhuI* trong IS80 và *ZhuJ* trong IS82. Các catxet biểu hiện đã được xác nhận được biến nạp vào *Aspergillus nidulans*, bằng cách kết hợp hướng đích tuần tự từng catxet biểu hiện, như được mô tả trong Ví dụ 6. Gen chỉ thị *Afp_{yrG}* được loại bỏ sau mỗi chu kỳ biến nạp để cho phép chu kỳ biến nạp tiếp theo đưa catxet biểu hiện bổ sung vào.

Phân tích hóa học được thực hiện như được mô tả trong Ví dụ 6.

Kết quả:

Phân tích các chất chuyển hóa đã tạo thành bằng UHPLC-HRMS cho thấy rằng sự biểu hiện của *ZhuI* và *ZhuJ* không ảnh hưởng đáng kể đến lượng chất chuyển hóa được tạo ra (bao gồm hợp chất SEK4, SEK4B và FK), tức là lượng được tạo ra bởi các chủng của Ví dụ 7 này (chứa PKS (*HpPKS2* hoặc *AaOKS*) + *ZhuI* và *ZhuJ*) là tương tự như lượng được tạo ra bởi các chủng của Ví dụ 6 (chỉ chứa PKS (*HpPKS2* hoặc *AaOKS*)).

Kết luận:

Kết quả cho thấy rằng sự biểu hiện của ZhuI và ZhuJ được tối ưu hóa cho sự biểu hiện trong *E. coli* không ảnh hưởng đáng kể đến các chất chuyển hóa được tạo ra trong *Aspergillus nidulans*.

Ví dụ 8: Sự biểu hiện khác loại của các PKS typ III và *Dactylopius coccus* C-glycosyltransferaza trong *Aspergillus nidulans*:

Vật liệu và phương pháp

Việc tạo cấu trúc của các plasmit, biến nạp di truyền của *Aspergillus nidulans* và phân tích hóa học được thực hiện như được mô tả trong Ví dụ 6.

Tạo cấu trúc các vector cho sự biểu hiện của *Dactylopius coccus* C-glycosyltransferaza trong *Aspergillus nidulans*

Gen *DcUGT2* được tách dòng/gắn xen vào chủng PKS (*AaOKS*) *Aspergillus nidulans* của Ví dụ 6.

Gen *DcUGT2* có chiều dài đầy đủ từ *Dactylopius coccus* được tối ưu hóa codon cho sự biểu hiện trong *Aspergillus nidulans* và được mua từ GenScript dưới dạng ADN tổng hợp. Để khuếch đại dựa trên PCR trình tự mã hóa *DcUGT2*, các đoạn môi sau được sử dụng: DcUGT2_AnOpt-F: 5'-AGAGCGAdUATGGAGTTTCGCTTGCTTATCCT và DcUGT2_AnOpt-R: 5'-TCTGCGAdUTTAATTCTTCTTCAACTTTTCCGACTTAG. Amplicon PCR thu được được tách dòng vào vector biểu hiện *Aspergillus nidulans* như được mô tả trong Ví dụ 6. Vector biểu hiện đã sử dụng hướng đích vị trí IS52 trong bộ gen của *Aspergillus nidulans*.

Kết quả:

Phân tích các chất chuyển hóa đã được tạo ra của các chủng *Aspergillus nidulans* chứa *AaOKS* và *DcUGT2* khác loại cho thấy rằng sự biểu hiện dẫn đến việc tạo ra các hợp chất khác nhau bao gồm hợp chất CA, DCII, KA. Các hợp chất FK, SEK4 và SEK4b cũng được nhận ra, và sự biểu hiện đồng thời của *DcUGT2* tạo ra hiệu suất cao hơn của các hợp chất được đề cập trước đó.

Kết luận

Sự biểu hiện đồng thời của PKS (*AaOKS*) và glycosyltransferaza (*DcUGT2*) khác loại trong *Aspergillus nidulans* dẫn đến việc tạo ra CA, DCII, KA. Các hợp chất FK, SEK4 và SEK4b cũng được nhận ra, và sự biểu hiện đồng thời của *DcUGT2* tạo ra hiệu suất cao hơn của các hợp chất được đề cập trước đó.

Ví dụ 9: Sự biểu hiện của AaOKS trong *N. benthamiana* và trong thử nghiệm hoạt tính *in vitro*

Sự biểu hiện của AaOKS trong *N. benthamiana*

AaOKS (Số truy cập ngân hàng gen AY567707) (Abe I, Oguro S, Utsumi Y, Sano Y, Noguchi H (2005b). Engineered biosynthesis of plant polyketides: chain length control in an octaketide-producing plant type III polyketide synthase. *J Am Chem Soc* 127: 12709-12716.; Abe I, Utsumi Y, Oguro S, Morita H, Sano Y, Noguchi H (2005a). A plant type III polyketide synthase that produces pentaketide chromone. *J Am Chem Soc* 127:1362-1363) được tổng hợp bởi Genscript và được khuếch đại bằng các đoạn mồi chứa đoạn nhô ra USER và đuôi his đầu cùng C để tách dòng vào vector pEAQ tương thích với USER (Sainsbury F, Thuenemann EC, Lomonosoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol J* 7: 682-693) với phương pháp USER (Nour-Eldin HH, Hansen BG, Nørholm MH, Jensen JK, Halkier BA (2006). Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res* 34: e122.) để tạo ra plasmit pEAQ-AaOKS thu được. Plasmit này được nhân lên trong tế bào *E. coli* TOP10 và được xác nhận bằng cách giải trình tự. Plasmit này được biến nạp vào *Agrobacterium tumefaciens* bằng kỹ thuật xung điện và các dòng vô tính dương được chọn trên thạch Luria-Bertani (LB) được bổ sung rifampixin (12,5 µg/mL), kanamycin (50 µg/mL), và ampicilin (25 µg/mL) sau đó là PCR để xác nhận sự có mặt của AaOKS.

Để biểu hiện, *A. tumefaciens* mang pEAQ-AaOKS hoặc pEAQ được nuôi cấy qua đêm trong LB lỏng được bổ sung rifampixin (12,5 µg/mL), kanamycin (50 µg/mL) và ampicilin (25 µg/mL). Các tế bào được tạo pelet bằng cách ly tâm và tạo huyền phù lại trong môi trường chuyển nạp (infiltration) (10 mM MES, 100 µM acetosyringone, 10 mM MgCl₂) đến OD₆₀₀ cuối cùng: 0,5 và được ủ ở nhiệt độ phòng trong 1 giờ. *A. tumefaciens* được chuyển nạp vào lá của cây *N. benthamiana* 3-4 tuần tuổi bằng xi lanh không kim và các cây này được trồng 5 ngày nữa trong nhà kính.

Chiết protein hòa tan từ *N. benthamiana*

300mg vật liệu thực vật được nghiền trong cối lạnh bằng chày và bổ sung 0,5g PVPP và 10mL chất đệm chiết lạnh (natri phosphat 50mM, sucroza 400mM, DTT 4mM, độ pH = 7,2 chứa chất ức chế proteinaza hoàn chỉnh (Roche)). Dịch đồng nhất

được lọc qua lưới nylon vào ống ly tâm lạnh và được ly tâm ở 10.000 xg ở nhiệt độ 4°C trong 15 phút. Phần dịch nổi trên bề mặt được phân lập và được ly tâm ở 100.000 xg ở nhiệt độ 4°C trong 60 phút. Phần hòa tan (phần dịch nổi trên bề mặt) được phân lập. Việc tạo ra AaOKS đuôi his được xác nhận bằng kỹ thuật thấm tách Western sử dụng kháng thể kháng his. AaOKS đuôi his được tinh chế từ phần hòa tan bằng cách sử dụng cột His Spin Trap (GE health care biosciences) theo hướng dẫn của nhà sản xuất.

Thử nghiệm polyketit syntaza

Dịch chiết enzym thô hòa tan của lá cây *N. benthamiana* và his-AaOKS được tinh chế ái lực được thử nghiệm về hoạt tính polyketit syntaza. Thử nghiệm tiêu chuẩn chứa 100µl protein hòa tan hoặc 20µl his-AaOKS được tinh chế ái lực trong tổng thể tích phản ứng là 250µl chứa natri phosphat 50mM, sucroza 500mM, DTT 1mM, malonyl-CoA 40µM, axetyl-coA 20µM, độ pH = 7. Thử nghiệm tiêu chuẩn được bổ sung [2-¹⁴C]-malonyl-CoA (1,8mCi/mmol) đối với phân tích TLC hoặc [¹³C]-malonyl-CoA đối với phân tích cấu trúc. Các hỗn hợp phản ứng được ủ ở nhiệt độ 30°C trong 90 phút, được kết thúc bằng cách bổ sung 25µl HCl 20% và được chiết hai lần bằng 500µl etyl axetat, axit axetic 1% (khối lượng/thể tích). Pha hữu cơ được làm khô trong chân không và được tạo huyền phù lại trong 20µl metanol, axit axetic 1% (khối lượng/thể tích) và được phân tích bằng TLC hoặc LC-MS.

Kết quả - Hoạt tính enzym *in vitro* của AaOKS:

Dịch chiết enzym thô từ *N. benthamiana* biểu hiện AaOKS được thử nghiệm về hoạt tính octaketit syntaza bằng cách sử dụng malonyl-CoA làm cơ chất khởi động. Khi ¹⁴C-malonyl-CoA được ủ với dịch chiết enzym thô hoặc AaOKS được tinh chế ái lực thì cơ chất này được chuyển hóa, như được quan sát thấy bởi sự xuất hiện của vạch hoạt tính phóng xạ trên sắc ký lớp mỏng, trong khi đó không có sản phẩm nào được quan sát thấy khi sử dụng dịch chiết enzym thô từ cây thuốc lá được chuyển nạp một mình plasmit. Để làm rõ về cấu trúc, ¹³C-malonyl-coA được sử dụng làm cơ chất và các sản phẩm đã đánh dấu được xác định bằng LC-MS là SEK4 và SEK4B. Không có sự khác biệt về sự tạo thành sản phẩm khi AaOKS được gắn đuôi với đuôi his đầu cùng C và điều này cho thấy rằng đuôi HIS không can thiệp vào hoạt tính enzym. Kết quả này cho thấy rằng AaOKS được biểu hiện khác loại trong *N. benthamiana* có hoạt

tính octaketit syntaza và enzym được tinh chế ái lực sản sinh *in vitro* các sản phẩm song song thu được từ octaketit là SEK4 và SEK4B.

Kết luận:

Kết quả này cho thấy rằng *AaOKS* được biểu hiện khác loại trong *N. benthamiana* có hoạt tính octaketit syntaza và enzym được tinh chế ái lực sản sinh *in vitro* các sản phẩm song song thu được từ octaketit là SEK4 và SEK4B.

Ví dụ 10: Tách dòng các dạng được cắt cụt của *DcUGT2* trong *S. cerevisiae* và thử nghiệm hoạt tính *in vitro* của chúng đối với FK và KA

Vật liệu và phương pháp

Sự biểu hiện của các cấu trúc gen *DcUGT2* được cắt cụt trong *S. cerevisiae*

Dạng được cắt cụt của *DcUGT2* được tạo ra thiếu cả trình tự tín hiệu và phần neo màng (*ASP_DcUGT2AMD-Strep*) hoặc một mình phần neo màng (*DcUGT2AMD-Strep*). Các đoạn mã hóa *ASP_DcUGT2AMD-Strep* và *DcUGT2AMD-Strep* được khuếch đại riêng rẽ từ plasmit *pYES-DEST52-DcUGT2-Strep* bằng cách sử dụng các đoạn mồi PCR đặc hiệu để kết hợp đuôi StrepII đầu cùng C (xem bảng dưới đây). Vị trí tái tổ hợp Gateway, attB1 và attB2, được đưa vào các đoạn đã được tạo ra trong PCR sau đây bằng cách sử dụng đoạn mồi thuận: 5'-ggggacaagttgtacaaaaagcaggct-3' và đoạn mồi đảo: 5'-ggggaccactttgtacaagaaagctgggt-3'. *ASP_DcUGT2AMD-Strep* và *DcUGT2AMD-Strep* được gắn bên sườn bởi các vị trí attB được tách dòng vào plasmit *pDONR207* (Invitrogen) và được chuyển vào plasmit đích, *pYES-DEST52* (Invitrogen) bằng cách sử dụng hệ thống công nghệ Gateway. Hai cấu trúc plasmit *pYES-DEST52* được biến nạp riêng rẽ vào chủng nấm men *Invsc1* (Invitrogen) và các thể biến nạp dương tính được xác nhận bằng PCR. Việc sản xuất protein khác loại được thực hiện theo hướng dẫn của plasmit gateway *pYES-DEST52* (Invitrogen). Việc sản xuất protein đuôi StrepII khác loại được xác nhận bằng kỹ thuật thẩm tách Western sử dụng kháng thể kháng Strep. Phần liên kết màng và phần protein hòa tan được tạo ra từ các thể biến nạp nấm men đã được xác nhận như được mô tả trong (Pompon, D., Louerat, B., Bronine, A., Urban, P. (1996). Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* 272:51–64) và được sàng lọc về hoạt tính glucosyl hóa đối với axit flavokermesic/axit kermesic như được mô tả trên đây.

Ví dụ 11: Sản xuất in planta axit carminic và DcII bằng cách biểu hiện khác loại các gen mã hóa PKS, enzym xyglaza và DcUGT có chiều dài đầy đủ trong *Nicotiana benthamiana*.

Sự biểu hiện tạm thời của các cấu trúc gen trong *Nicotiana benthamiana*

Đoạn ADN tổng hợp mã hóa *ZhuI* (Số truy cập ngân hàng gen: AAG30197) và *ZhuJ* (Số truy cập ngân hàng gen: AAG30196) được tối ưu hóa codon cho sự biểu hiện *N. benthamiana* và *AaOKS* (Số truy cập ngân hàng gen: AY567707) được mua từ Genscript. Tất cả các đoạn tổng hợp và plasmid *pYES-DEST52-DcUGT2-Strep* được mô tả ở đây được sử dụng làm khuôn PCR với các đoạn mồi chứa deoxyuraxil (dU) tương thích (xem Bảng 5 dưới đây) để tạo ra các cấu trúc được tách dòng vào *pEAQ-HT-USER* bằng công nghệ USER. Dạng *DcUGT2* được cắt cụt, *DcUGT2ΔMD-Strep*, được chuyển từ plasmid *pDONR207* (Invitrogen) vào plasmid đích, *pEAQ-HT-DEST1* (Sainsbury, F., Saxena, P., Geisler, K., Osbourn, A., Lomonossoff, G.P. (2012). Using a Virus-Derived System to Manipulate Plant Natural Product Biosynthetic Pathways. *Methods Enzymol.* 517:185–202), sử dụng hệ thống công nghệ Gateway. Tất cả các cấu trúc plasmid *pEAQ-HT* đều được biến nạp vào chủng *Agrobacterium tumefaciens*, *AGL-1* và được chuyển nạp vào lá của cây *N. benthamiana* như được mô tả trong (Bach, S.S., Bassard, J.É., Andersen-Ranberg, J., Møldrup, M.E., Simonsen, H.T., Hamberger, B. (2014). High-Throughput Testing of Terpenoid Biosynthesis Candidate Genes Using Transient Expression in *Nicotiana benthamiana*. In M Rodríguez Concepción, ed, *Plant Isoprenoids, Methods in Molecular Biology*, Vol. 1153. Humana Press, New York).

Bảng 5. Các trình tự đoạn mồi để khuếch đại các cấu trúc gen khác nhau

Các đoạn gen	Trình tự đoạn mồi	
AaOKS	Thuận	5'-GGCTTAA/dU/ATGAGTTCACCTCTCCAACGCTTCCCATC-3'
	Đảo	5'-GGTTTAA/dU/TTACATGAGAGGCAGGCTGTGGAGAAGGATAGT-3'
ZhuI	Thuận	5'-GGCTTAA/dU/ATGAGGCATGTGCGAGCAT-3'
	Đảo	5'-GGTTTAA/dU/TTATGCCGTGACAGTTCCGACAC-3'
ZhuJ	Thuận	5'-GGCTTAA/dU/ATGTCCGGACGTAAGACG-3'
	Đảo	5'-GGTTTAA/dU/TTAATCTTCCTCCTCCTGTTCAA-3'
DcUGT2	Thuận	5'-GGCTTAA/dU/ATGGAATTCAGATTGTTGATATTGGCCT-3'

-Strep	Đảo	5'-GGTTTAA/dU/TTATTTTTCGAATTGTGGATGAGACCAAGCAGA-3'
DcUGT2 ΔMD- Strep	Thuận (attB1)	5'-ggggacaagtttgtaaaaaagcaggct-3'
	Đảo	5'- TTATTTTTCGAATTGTGGATGAGACCAAGCAGAGTGCAAAAAGG CACCTGCAGT-3'

Chiết chất chuyển hóa và phân tích LC-MS/MS

Các chất chuyển hóa được chiết từ các đĩa ($\varnothing=3\text{cm}$) chứa lá cây *N. benthamiana* đã được chuyển nạp vào tế bào. Các đĩa chứa lá, được cắt bằng máy khoan xoắn, được làm đông lạnh nhanh trong nitơ lỏng. 0,5mL chất đệm chiết (methanol 85% (thể tích/thể tích), axit formic 0,1% (thể tích/thể tích)), được làm cân bằng đến nhiệt độ 50°C, được bổ sung vào từng đĩa chứa lá đông lạnh sau đó ủ trong 1 giờ ở nhiệt độ 50°C, khuấy ở 600 vòng/phút. Phần dịch nổi trên bề mặt được phân lập và cho đi qua đĩa lọc Multiscreen_{HTS} HV 0,45 μm (Merck Milipore). Phần dịch nổi trên bề mặt đã được lọc này được đưa vào phân tích LC-MS/MS mà được thực hiện trên Agilent 1200 HPLC kết hợp với thiết bị đo phổ khối Bruker micrOTOF-Q II được trang bị nguồn ion hóa tia điện. Tách sắc ký thu được trên cột Luna C₁₈(2) (150 × 4,6mm, 3 μm , 100Å, Phenomenex) được duy trì ở nhiệt độ 40°C. Dung dịch rửa giải (A) bao gồm nước/axetonitril (95:5, thể tích/thể tích) và dung dịch rửa giải hữu cơ (B) bao gồm nước/axetonitril (5:95, thể tích/thể tích); cả hai dung dịch này được axit hóa bằng axit formic 0,1%. Tùy theo mục đích của phân tích, hai mẫu rửa giải gradien tuyến tính khác nhau được sử dụng:

Phương pháp 1: 0 phút, 0% B; 30 phút, 100% B; 33 phút 100% B; 35 phút, 0% B. Tốc độ dòng chảy được duy trì ở 0,5mL/phút và cân bằng 10 phút được sử dụng. Thời gian lưu là 15,5 phút đối với dcII, 15,6 phút đối với axit carminic (CA), 17,1 phút đối với flavokermesic-*O*-Glc 2 (FK-*O*-Glc 2), 17,2 phút đối với SEK4, 17,7 phút đối với SEK4B, 22,7 phút đối với axit flavokermesic (FK) và 22,7 phút đối với axit kermesic (KA).

Phương pháp 2: 0 phút, 10% B; 25 phút, 20% B; 27 phút 100% B; 35 phút, 100% B; 36 phút, 10% B. Tốc độ dòng chảy được duy trì ở 0,5mL/phút và cân bằng 10 phút được sử dụng.

Thời gian lưu là 16,1 phút đối với flavokermesic-*O*-Glc 1 (FK-*O*-Glc 1), 17,0 phút đối với dcII, 18,2 phút đối với axit carminic (CA), 24,1 phút đối với SEK4, 25,5 phút đối với flavokermesic-*O*-Glc 2 (FK-*O*-Glc 2), 26,8 phút đối với SEK4B, 35,5 phút đối với axit flavokermesic (FK) và 36,0 phút đối với axit kermesic (KA).

Các *O*-glucosit được xác định bằng kỹ thuật xử lý viscozyme L và bằng hao hụt phân tử trung tính 162 Da trong phổ MS/MS.

Kết quả:

AaOKS của *Aloe arborescens* đã được mô tả trước đó là octaketit syntaza thuộc nhóm enzym PKS typ III. Để nghiên cứu khả năng sử dụng enzym PKS typ III này để tạo ra tiền chất octaketit cần thiết cho việc sản xuất axit carminic, gen *AaOKS* được biểu hiện tạm thời trong *N. benthamiana*. Gen *AaOKS*, khi so sánh với đối chứng plasmit rỗng, dẫn đến việc tạo ra SEK4 và SEK4B sau khi chuyển nạp vào tế bào lá của *N. benthamiana*. Điều này chứng minh rằng AaOKS có thể hoạt động như octaketit syntaza hoạt tính *in vivo* trong *N. benthamiana*. Ngoài ra, vì không có axit flavokermesic (FK) anthron hoặc FK nào có thể được quan sát thấy khi *AaOKS* được chuyển nạp một mình vào tế bào, *N. benthamiana* có thể thiếu enzym nội sinh để chuyển hóa thêm SEK4 và SEK4B thành các hợp chất này. Vì vậy, các gen *xylC*, *ZhuI* và *ZhuJ*, được chuyển nạp đồng thời với *AaOKS*. *ZhuI* và *ZhuJ* có nguồn gốc từ con đường sinh tổng hợp kháng sinh R1128 và đã được kết hợp *in vivo* từ trước với PKS tối thiểu actinorhodin trong *Streptomyces coelicolor* để tạo ra FK (Tang, Y., Lee, T.S., Khosla, C. (2004). Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. PLOS Biol. 2(2):E31). Trong *N. benthamiana*, khi *ZhuJ* được chuyển nạp đồng thời vào tế bào cùng với *AaOKS*, quan sát thấy sự tích tụ FK và axit flavokermesic-*O*-glucosit, FK-*O*-Glc 2 và đây không phải là trường hợp khi *AaOKS* được chuyển nạp đồng thời vào tế bào cùng với *ZhuI*. Việc tạo ra FK-*O*-Glc 2 gợi ý rằng một hoặc một vài (các) UGT của *N. benthamiana* nội sinh có khả năng sử dụng một cách hiệu quả FK làm cơ chất. Ngoài ra, việc chuyển nạp đồng thời vào tế bào *ZhuI* cùng với *AaOKS* làm giảm sự sản sinh ra SEK4B bằng cách thúc đẩy việc tạo thành SEK4 mà phù hợp với kết quả trước đó là *ZhuI* điều khiển sự tạo vòng C7-C12 ban đầu (thay vì tạo vòng C10-C15 mà dẫn đến việc tạo ra SEK4B) của octaketit mạch thẳng (Ames, B.D., Lee, M.Y., Moody, C., Zhang, W., Tang, Y., Tsai, S.C. (2011). Structural and biochemical characterization of

ZhuI aromatase/xyclase from the R1128 polyketide pathway. Biochemistry. 39: 8392-8406). Việc tạo vòng C7-C12 ban đầu này cũng là cần thiết cho sự sinh tổng hợp mạch chính anthraquinon của FK/CA và khi *ZhuI* được chuyển nạp đồng thời vào tế bào cùng với *AaOKS* và *ZhuJ*, tổng lượng gộp của FK và axit flavokermesic-*O*-glucosit thực sự tăng lên rõ rệt so với khi nó không có mặt.

Việc tạo ra DCII và CA *in planta* được phát hiện khi *DcUGT2*_Strep có chiều dài đầy đủ được biểu hiện đồng thời cùng với *AaOKS*, *ZhuI* và *ZhuJ* trong *N. benthamiana*. Việc tạo ra CA chỉ được phát hiện ra ở các lượng nhỏ khi so với nồng độ DCII. Việc tạo ra CA này có thể do hoạt tính monooxygenaza nội sinh không đặc hiệu trong *N. benthamiana* mà có khả năng hydroxyl hóa FK thành axit kermesic (KA) hoặc DCII thành CA. Tuy nhiên, đáng lưu ý rằng không thể phát hiện được KA trong nghiên cứu chuyển nạp bất kỳ, chỉ ra rằng hợp chất này được tạo ra ở các mức thấp hơn giới hạn phát hiện của các tác giả sáng chế hoặc nó có thể có độc tính cao đối với *N. benthamiana* và do đó được chuyển hóa nhanh chóng thành sản phẩm chưa biết.

Kết luận:

Các kết quả của ví dụ này chứng minh:

(i): Gen *AaOKS* từ *Aloe arborescens* được biểu hiện tạm thời trong *N. benthamiana* và điều này dẫn đến việc tạo ra SEK4 và SEK4B trên *in vivo*, chứng minh rằng *AaOKS* có thể hoạt động như là octaketit syntaza hoạt tính *in vivo* trong *N. benthamiana*;

(ii): Vì không quan sát thấy axit flavokermesic (FK) anthron hoặc FK khi *AaOKS* được chuyển nạp một mình vào tế bào, *N. benthamiana* có thể thiếu enzym nội sinh để chuyển hóa thêm SEK4 và SEK4B thành các hợp chất này.

(iii): Các gen xyclaza *ZhuI* và *ZhuJ*, từ *Streptomyces* sp. R1128, được chuyển nạp đồng thời vào tế bào cùng với *AaOKS* và khi *ZhuJ* được chuyển nạp đồng thời vào tế bào cùng với *AaOKS*, quan sát thấy sự tích tụ của FK và axit flavokermesic-*O*-glucosit, FK-*O*-Glc 2. Khi *ZhuI* được chuyển nạp đồng thời vào tế bào cùng với *AaOKS* và *ZhuJ*, tổng lượng gộp của FK và axit flavokermesic-*O*-glucosit thực sự tăng lên đáng kể so với khi nó không có mặt. Do đó, sự biểu hiện khác loại của các gen *Streptomyces* sp xyclaza dẫn đến việc tạo ra *in vivo* các hợp chất khác nhau bao gồm FK;

(iv): Việc tạo ra DCII và CA *in vivo* được phát hiện khi *DcUGT2_Strep* có chiều dài đầy đủ được biểu hiện đồng thời cùng với *AaOKS*, *ZhuI* và *ZhuJ* in *N. benthamiana*.

Ví dụ 12: Sự biểu hiện khác loại dạng được cắt cụt của DcUGT2 và thử nghiệm hoạt tính *in planta* của nó trong *N. benthamiana*.

Để nghiên cứu protein DcUGT2 có thể được hòa tan và vẫn giữ được hoạt tính glucosyl hóa đối với KA và FK hay không, phần neo màng được làm khuyết. DcUGT2 được dự đoán, thông qua tin sinh học, được đặt trong ER có đầu cùng C được neo trong màng ER và vị trí hoạt động đầu cùng N đối diện với lumen. Do đó suy đoán được là protein này có thể được *N*-glycosyl hóa hay không và, nếu như vậy thì biến đổi sau dịch mã này có thể là cần thiết cho hoạt tính của enzym DcUGT2. Phân tích dự đoán tin sinh học *N*-glycosyl hóa được thực hiện, xác định 3 vị trí *N*-glycosyl hóa giả định. Để nghiên cứu việc glycosyl hóa và việc hướng đích ER có vai trò đối với hoạt tính của enzym DcUGT2 hay không không, dạng DcUGT2 đã được cắt cụt được tạo ra, trong đó peptit tín hiệu còn nguyên vẹn nhưng phần neo màng bị làm khuyết (DcUGT2 Δ MD-Strep). Protein DcUGT2 Δ MD-Strep này có hoạt tính *in vivo* và tạo ra DCII khi được biểu hiện đồng thời với *AaOKS* và *ZhuJ* trong *N. benthamiana*. Sự sản xuất DCII ở các cây này là thấp hơn so với các mức DCII ở các cây được chuyển nạp đồng thời vào tế bào cùng với *DcUGT2*. Điều này có thể ám chỉ rằng, mặc dù protein DcUGT2 Δ MD-Strep có hoạt tính, mức hoạt tính enzym có thể bị giảm đi do việc làm khuyết phần neo màng. Việc DcUGT2 Δ MD-Strep vẫn có một số hoạt tính glucosyl hóa đối với FK và/hoặc KA còn chỉ ra rằng việc *N*-glycosyl hóa và (hoặc) hướng đích ER của protein này có thể là thiết yếu đối với hoạt tính của enzym. Ngoài ra, được chỉ ra rằng DcUGT2 có chiều dài đầy đủ dường như cạnh tranh một cách hiệu quả với hoạt tính FK *O*-glucosyl hóa của *N. benthamiana*, từ đó làm giảm đáng kể lượng gộp của axit flavokermesic-*O*-Glc (FK-*O*-Glc 1 và 2) và điều này không được quan sát thấy đối với dạng DcUGT2 Δ MD-Strep được cắt cụt.

Kết luận

Các kết quả của Ví dụ này đã chứng minh rằng *DcUGT2 Δ MD-Strep* (thiếu phần neo màng - các axit amin từ 1 đến 468 của SEQ ID NO:2) có hoạt tính *in vivo* ở cây *N. benthamiana*.

Hoạt tính của *DcUGT2AMD-Strep* (thiếu phần neo màng) là thấp hơn so với enzym DcUGT2 có chiều dài đầy đủ – tuy nhiên, việc nó có hoạt tính có thể chỉ ra rằng nó có thể đặc biệt hữu ích cho việc sản xuất khác loại trong, ví dụ, sinh vật tiền nhân.

TÀI LIỆU THAM KHẢO

- 1: US5424421 (European Colour, công bố năm 1995)
- 2: WO2006/056585A1 (Chr. Hansen A/S)
- 3: Stathopoulou et al (Analytica Chimica Acta 804 (2013) 264–272)
- 4: Zagrobelny et al (Cyanogenic glucosides and plant-insect interactions; Phytochemistry. 2004 Feb;65(3):293-306)
- 5: Geuder et al (Journal of Chemical Ecology, Vol. 23, No. 5, 1997)
- 6: Genta et al, (Potential role for gut microbiota in cell wall digestion and glucoside detoxification in *Tenebrio molitor* larvae), Journal of Insect Physiology 52 (2006) 593–601
- 7: WO2004/111254A1 (Poalis A/S)
- 8: Gutmann et al (Pure Appl. Chem, 2013-07-09)
- 9: Baig et al (Angew Chem Int Ed Engl. 2006 Nov 27;45(46):7842-6)
- 10: Radomska-Pandya A, Bratton SM, Redinbo MR, Miley MJ. Drug Metab Rev. 2010 Feb;42(1):133-44)
- 11: Plant Physiology, November 2008, Vol. 148, pp. 1295–1308
- 12: Esben Halkjaer Hansen et al. Phytochemistry 70(4): 473-482
- 13: Tang, Y., Lee, T. S., and Khosla, C. (2004) Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. PLoS Biol. 2, e31
14. Mizuuchi Y, Shi SP, Wanibuchi K, Kojima A, Morita H, Noguchi H, Abe I. Novel type III polyketide synthases from *Aloe arborescens*. FEBS J. 2009 Apr;276(8):2391-2401.
15. Karppinen K, Hokkanen J, Mattila S, Neubauer P, Hohtola A. Octaketide-producing type III polyketide synthase from *Hypericum perforatum* is expressed in dark glands accumulating hypericins. FEBS J. 2008 Sep;275(17):4329-4342.
16. Dayu Yu, Fuchao Xu, Jia Zeng, Jixun Zhan. Type III Polyketide Synthases in Natural Product Biosynthesis. IUBMB Life, 64(4): 285–295, April 2012
17. Supriya Jadhav, Prasad Phapale, Hirekodathakallu V. Thulasiram, Sujata Bhargava. Polyketide synthesis in tobacco plants transformed with a *Plumbago zeylanica* type III hexaketide synthase. Phytochemistry 98 (2014) 92–100.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất hợp chất thơm thu được từ octaketit, trong đó phương pháp này bao gồm các bước sau:

(I): cho tiếp xúc *in vivo* trong tế bào chủ tái tổ hợp chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS), trong đó OKS thuộc giống khác với tế bào chủ:

(i): đơn vị khởi động và đơn vị kéo dài với OKS nêu trên sao cho đơn vị khởi động và đơn vị kéo dài này chuyển hóa thành octaketit không khử;

(II): chuyển hóa *in vivo* trong tế bào chủ tái tổ hợp octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄ nhờ ít nhất một aromataza/xyclaza hoạt động *in trans*, trong đó hợp chất aglycon thơm cần quan tâm này không phải là SEK4 và/hoặc SEK4B;

(III): tế bào chủ tái tổ hợp còn chứa gen glycosyltransferaza mã hóa glycosyltransferaza có độ đồng nhất trình tự ít nhất là 90% với SEQ ID NO: 2 hoặc các axit amin từ 1 đến 468 của SEQ ID NO: 2, mà glycosyl hóa hợp chất aglycon thơm được sản xuất ở bước (II) thành hợp chất glycosit thơm C₁₄-C₃₄; và

(IV): phân lập hợp chất glycosit thơm của bước (III) để thu được chế phẩm, trong đó chế phẩm này chứa ít hơn 1% khối lượng/khối lượng chất khô của vật liệu tế bào chủ tái tổ hợp và trong đó tế bào chủ tái tổ hợp là tế bào thực vật, tế bào nấm sợi hoặc tế bào nấm men.

2. Phương pháp theo điểm 1, trong đó tế bào chủ tái tổ hợp theo điểm 1 là tế bào chủ tái tổ hợp đang sinh trưởng và bước (I) và bước (II) bao gồm:

(I): cho tiếp xúc *in vivo* trong tế bào chủ tái tổ hợp đang sinh trưởng chứa gen polyketit syntaza (PKS) typ III đã được đưa vào bằng cách tái tổ hợp mã hóa octaketit syntaza (OKS), trong đó OKS thuộc giống khác với tế bào chủ:

(i): đơn vị khởi động và đơn vị kéo dài với OKS nêu trên sao cho đơn vị khởi động và đơn vị kéo dài này chuyển hóa thành octaketit không khử; và

(II): chuyển hóa *in vivo* trong tế bào chủ tái tổ hợp đang sinh trưởng, octaketit không khử của bước (I) thành hợp chất aglycon thơm C₁₄-C₃₄, trong đó hợp chất aglycon thơm này không phải là SEK4 và/hoặc SEK4B.

3. Phương pháp theo điểm bất kỳ trong các điểm nêu trên, trong đó tế bào thực vật là *Nicotiana* sp..
4. Phương pháp theo điểm bất kỳ trong các điểm từ 1 đến 2, trong đó tế bào nấm sợi là tế bào *Aspergillus*.
5. Phương pháp theo điểm bất kỳ trong các điểm từ 1 đến 2, trong đó tế bào nấm men là *S. cerevisiae*.
6. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó gen polyketit syntaza (PKS) Typ III của bước (I) là gen PKS từ thực vật và trong đó thực vật này là thực vật được chọn từ nhóm bao gồm: *Aloe* spp. và *Hypericum* spp..
7. Phương pháp theo điểm 6, trong đó octaketit syntaza (OKS) của bước (I) là OKS chứa trình tự axit amin, có độ đồng nhất trình tự ít nhất là 70% với các axit amin từ 1 đến 403 của SEQ ID NO:7 (*AaOKS*).
8. Phương pháp theo điểm 6, trong đó octaketit syntaza (OKS) của bước (I) là OKS chứa trình tự axit amin, có độ đồng nhất trình tự ít nhất là 70% với các axit amin từ 1 đến 393 của SEQ ID NO:15 (*HpPKS2*).
9. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó đơn vị khởi động là axetyl-CoA và/hoặc malonyl-CoA.
10. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó:
 - tế bào chủ tái tổ hợp là tế bào chủ được chọn từ nhóm bao gồm: *Aspergillus* và *Nicotiana* sp.; và
 - OKS thuộc giống khác với tế bào chủ là OKS được chọn từ nhóm bao gồm: OKS từ *Aloe* spp. và *Hypericum* spp..
11. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó hợp chất aglycon thơm là hợp chất aglycon thơm C₁₆.
12. Phương pháp theo điểm 11, trong đó hợp chất aglycon thơm C₁₆ là axit flavokermesic (FK) hoặc axit kermesic (KA).
13. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên,
 - trong đó aromataza/xyclaza là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất trình tự ít nhất là 70% với các axit amin từ 1 đến 169 của SEQ ID NO:17 (*Streptomyces* ZhuI); và/hoặc

trong đó aromataza/xyclaza là aromataza/xyclaza chứa trình tự axit amin có độ đồng nhất trình tự ít nhất là 70% với các axit amin từ 1 đến 256 của SEQ ID NO:19 (*Streptomyces* ZhuJ).

14. Phương pháp theo điểm bất kỳ trong số các điểm nêu trên, trong đó hợp chất glycosit thơm là hợp chất glycosit thơm C₁₆ và trong đó hợp chất glycosit thơm C₁₆ là glycosit của axit flavokermesic hoặc glycosit của axit kermesic.

15. Phương pháp theo điểm 14, trong đó hợp chất glycosit thơm C₁₆ là glycosit của axit flavokermesic và glycosit của axit flavokermesic là DcII.

16. Phương pháp theo điểm 14, trong đó hợp chất glycosit thơm C₁₆ là glycosit của axit kermesic và glycosit của axit kermesic là axit carminic (CA).

Fig.1

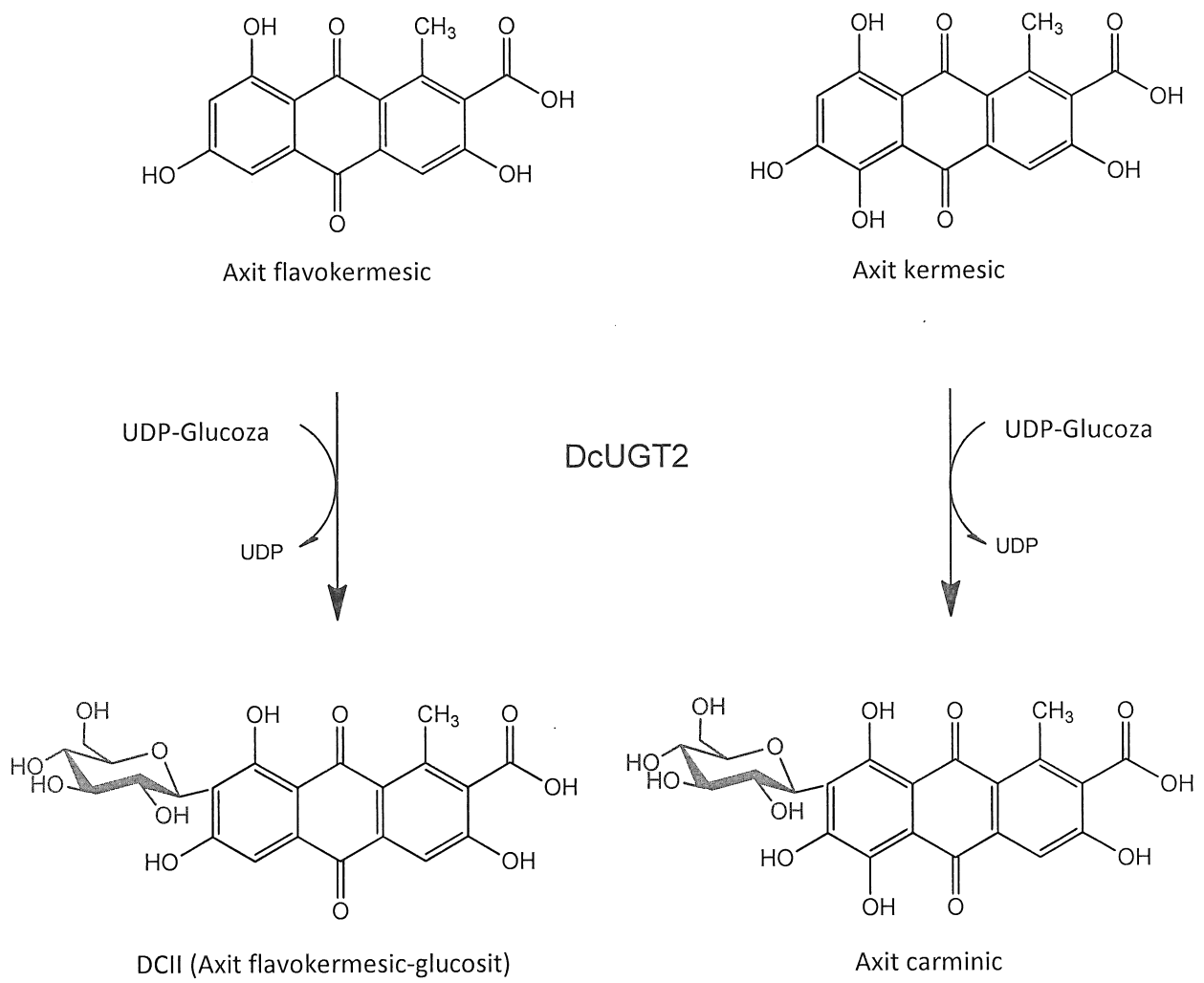


Fig.2

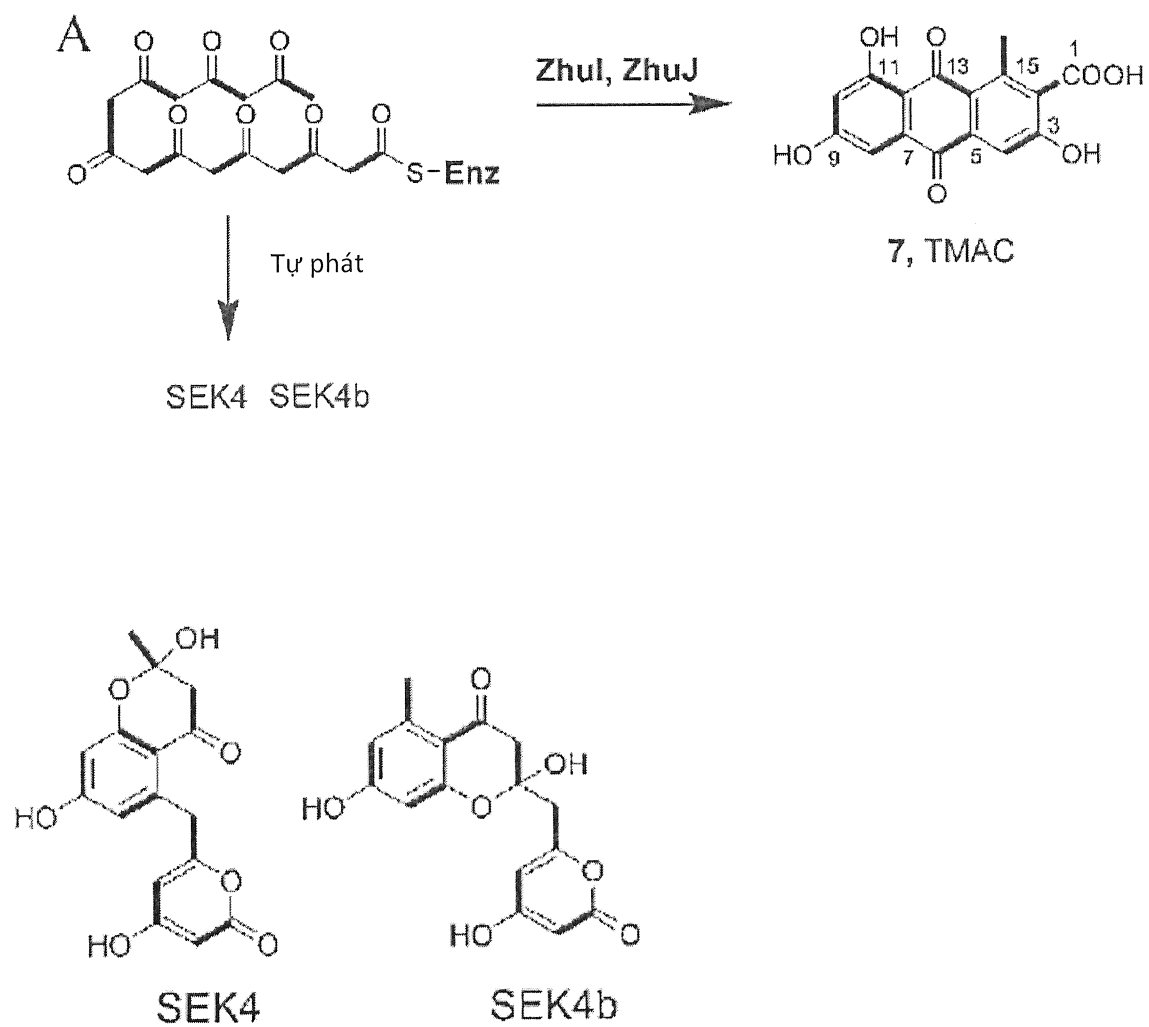


Fig.3

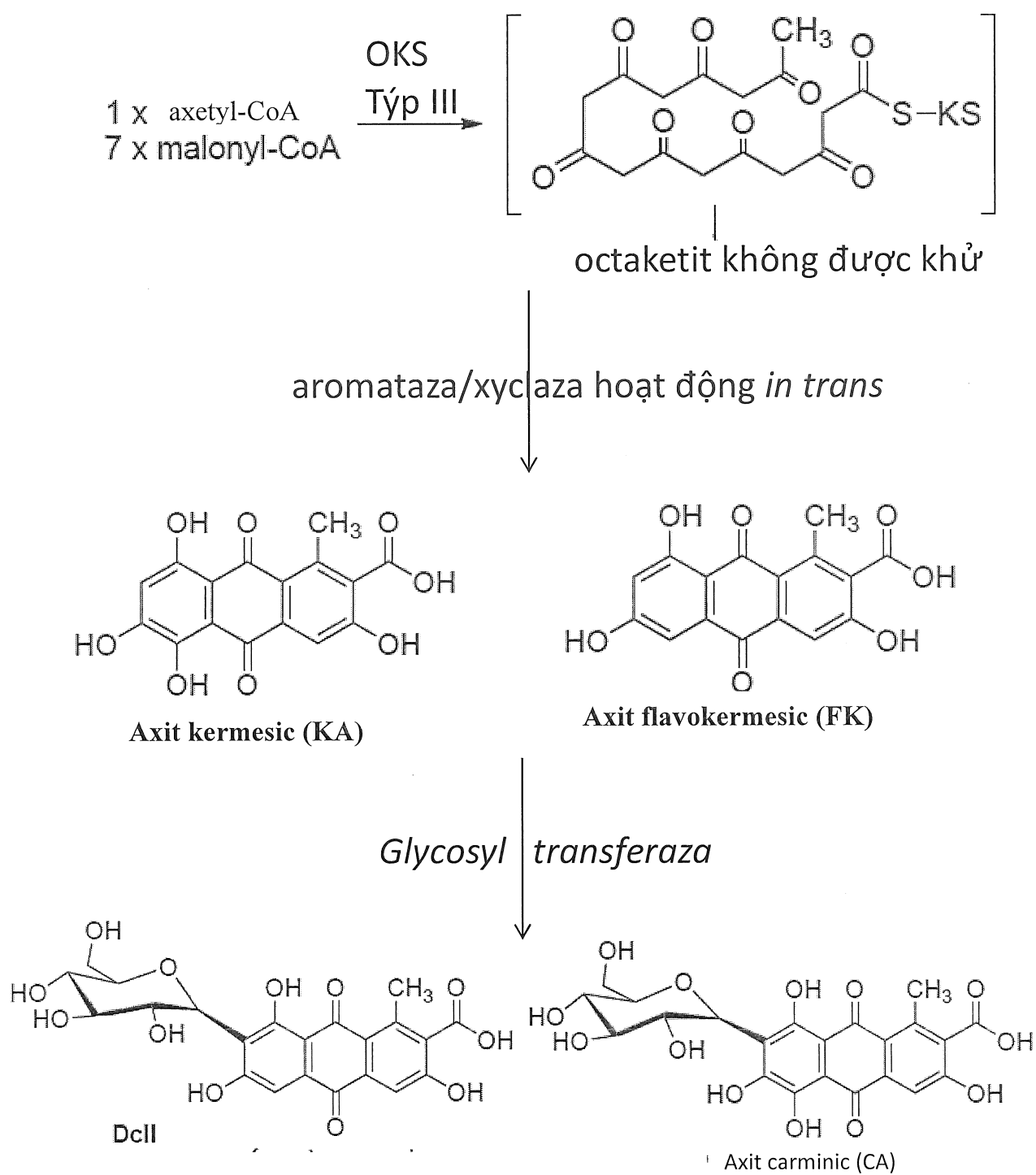
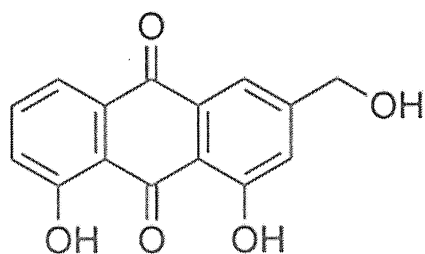
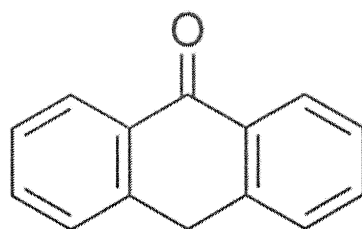
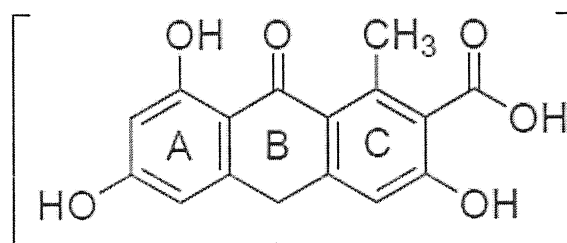


Fig.4

9,10-anthraquinon skeleton



Anthron



FK antron

DANH MỤC TRÌNH TỰ

<110> Københavns Universitet
 Danmarks Tekniske Universitet

<120> Phương pháp sản xuất hợp chất thơm thu được từ octaketit

<130> 56189EP01

<160> 89

<170> PatentIn phiên bản 3.5

<210> 1
 <211> 1548
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<220>
 <221> CDS
 <222> (1)..(1548)

<400> 1
 atg gaa ttt cgt tta cta atc ctg gct ctt ttt tct gta ctt atg agt
 48
 Met Glu Phe Arg Leu Leu Ile Leu Ala Leu Phe Ser Val Leu Met Ser
 1 5 10 15
 act tca aac gga gca gaa att tta gct ctt ttc cct att cac ggt atc
 96
 Thr Ser Asn Gly Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile
 20 25 30
 agt aat tat aat gtt gct gaa gca ctg ctg aag acc tta gct aac cgg
 144
 Ser Asn Tyr Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg
 35 40 45
 ggt cat aat gtt aca gtt gtc aca tct ttt cct caa aaa aaa cct gta
 192
 Gly His Asn Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val
 50 55 60
 cct aat ttg tac gaa att gac gta tct gga gct aaa ggc ttg gct act
 240
 Pro Asn Leu Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr
 65 70 75 80

```

aat tca ata cat ttt gaa aga tta caa acg att att caa gat gta aaa
288
Asn Ser Ile His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys
      85                      90                      95

tcg aac ttt aag aac atg gta cga ctt agc aga aca tac tgt gag att
336
Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
      100                      105                      110

atg ttt tct gat ccg agg gtt ttg aac att cga gac aag aaa ttc gat
384
Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
      115                      120                      125

ctc gta ata aac gcc gta ttt ggc agt gac tgc gat gcc gga ttc gca
432
Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
      130                      135                      140

tgg aaa agt caa gct cca ttg att tca att ctc aat gct aga cat act
480
Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
      145                      150                      155                      160

cct tgg gcc cta cac aga atg gga aat cca tca aat cca gcg tat atg
528
Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
      165                      170                      175

cct gtc att cat tct aga ttt cct gta aaa atg aat ttc ttc caa aga
576
Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
      180                      185                      190

atg ata aat acg ggt tgg cat ttg tat ttt ctg tac atg tac ttt tat
624
Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
      195                      200                      205

tat ggt aat gga gaa gat gcc aac aaa atg gcg aga aaa ttt ttt ggc
672
Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
      210                      215                      220

aac gac atg ccc gac ata aat gaa atg gtt ttt aat aca tct tta tta
720
Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
      225                      230                      235                      240

ttc gta aat act cac ttt tcg gtt gat atg cca tat cct ttg gtt cca
768
Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
      245                      250                      255

```

```

aac tgc att gaa ata gga gga ata cat gta aaa gag cca caa cca ctg
816
Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu
      260                      265                      270

cct ttg gaa ata caa aaa ttc atg gac gaa gca gaa cat ggg gtc att
864
Pro Leu Glu Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile
      275                      280                      285

ttc ttc acg cta gga tca atg gtg cgt act tcc acg ttt cca aat caa
912
Phe Phe Thr Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln
      290                      295                      300

act att caa gca ttt aag gaa gct ttt gcc gaa tta cct caa aga gtc
960
Thr Ile Gln Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val
305                      310                      315                      320

tta tgg aag ttt gag aat gaa aat gag gat atg cca tca aat gta ctc
1008
Leu Trp Lys Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu
      325                      330                      335

ata agg aaa tgg ttt cca caa aat gat ata ttc ggt cat aag aat atc
1056
Ile Arg Lys Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile
      340                      345                      350

aaa gca ttc att agt cac ggt gga aat tct gga gct ctg gag gct gtt
1104
Lys Ala Phe Ile Ser His Gly Gly Asn Ser Gly Ala Leu Glu Ala Val
      355                      360                      365

cat ttc gga gta ccg ata att gga att cct tta ttc tac gat cag tac
1152
His Phe Gly Val Pro Ile Ile Gly Ile Pro Leu Phe Tyr Asp Gln Tyr
      370                      375                      380

agg aat att ttg agt ttc gtt aaa gaa ggt gtt gcc gtt ctt ttg gat
1200
Arg Asn Ile Leu Ser Phe Val Lys Glu Gly Val Ala Val Leu Leu Asp
385                      390                      395                      400

gtg aat gat ctg acg aaa gat aat att tta tct tct gtc agg act gtt
1248
Val Asn Asp Leu Thr Lys Asp Asn Ile Leu Ser Ser Val Arg Thr Val
      405                      410                      415

gtt aat gat aag agt tac tca gaa cgt atg aaa gca ttg tca caa cta
1296
Val Asn Asp Lys Ser Tyr Ser Glu Arg Met Lys Ala Leu Ser Gln Leu

```

420 425 430
 ttc cga gat cga cca atg agt cct ctt gac aca gct gtt tac tgg aca
 1344
 Phe Arg Asp Arg Pro Met Ser Pro Leu Asp Thr Ala Val Tyr Trp Thr
 435 440 445
 gaa tat gtc atc cgc cat aga gga gcc cat cac ctc aag acc gct ggc
 1392
 Glu Tyr Val Ile Arg His Arg Gly Ala His His Leu Lys Thr Ala Gly
 450 455 460
 gca ttt ttg cat tgg tat cag tat tta ctt ttg gac gtt att acc ttc
 1440
 Ala Phe Leu His Trp Tyr Gln Tyr Leu Leu Leu Asp Val Ile Thr Phe
 465 470 475 480
 tta tta gtc aca ttc tgc gct ttt tgt ttt att gtg aaa tat ata tgt
 1488
 Leu Leu Val Thr Phe Cys Ala Phe Cys Phe Ile Val Lys Tyr Ile Cys
 485 490 495
 aaa gct ctc att cat cat tat tgg agc agt tcg aaa tct gaa aag ttg
 1536
 Lys Ala Leu Ile His His Tyr Trp Ser Ser Ser Lys Ser Glu Lys Leu
 500 505 510
 aaa aaa aat taa
 1548
 Lys Lys Asn
 515

 <210> 2
 <211> 515
 <212> PRT
 <213> Trình tự nhân tạo

 <220>
 <223> Cấu trúc Tổng hợp

 <400> 2
 Met Glu Phe Arg Leu Leu Ile Leu Ala Leu Phe Ser Val Leu Met Ser
 1 5 10 15
 Thr Ser Asn Gly Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile
 20 25 30
 Ser Asn Tyr Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg
 35 40 45

Gly His Asn Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val
 50 55 60

Pro Asn Leu Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr
 65 70 75 80

Asn Ser Ile His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys
 85 90 95

Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
 100 105 110

Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
 115 120 125

Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
 130 135 140

Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
 145 150 155 160

Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
 165 170 175

Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
 180 185 190

Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
 195 200 205

Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
 210 215 220

Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
 225 230 235 240

Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
 245 250 255

Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu

			260					265					270			
Pro	Leu	Glu	Ile	Gln	Lys	Phe	Met	Asp	Glu	Ala	Glu	His	Gly	Val	Ile	
		275					280					285				
Phe	Phe	Thr	Leu	Gly	Ser	Met	Val	Arg	Thr	Ser	Thr	Phe	Pro	Asn	Gln	
	290					295					300					
Thr	Ile	Gln	Ala	Phe	Lys	Glu	Ala	Phe	Ala	Glu	Leu	Pro	Gln	Arg	Val	
305					310					315					320	
Leu	Trp	Lys	Phe	Glu	Asn	Glu	Asn	Glu	Asp	Met	Pro	Ser	Asn	Val	Leu	
				325					330					335		
Ile	Arg	Lys	Trp	Phe	Pro	Gln	Asn	Asp	Ile	Phe	Gly	His	Lys	Asn	Ile	
			340					345					350			
Lys	Ala	Phe	Ile	Ser	His	Gly	Gly	Asn	Ser	Gly	Ala	Leu	Glu	Ala	Val	
		355					360					365				
His	Phe	Gly	Val	Pro	Ile	Ile	Gly	Ile	Pro	Leu	Phe	Tyr	Asp	Gln	Tyr	
	370					375					380					
Arg	Asn	Ile	Leu	Ser	Phe	Val	Lys	Glu	Gly	Val	Ala	Val	Leu	Leu	Asp	
385					390					395					400	
Val	Asn	Asp	Leu	Thr	Lys	Asp	Asn	Ile	Leu	Ser	Ser	Val	Arg	Thr	Val	
				405					410					415		
Val	Asn	Asp	Lys	Ser	Tyr	Ser	Glu	Arg	Met	Lys	Ala	Leu	Ser	Gln	Leu	
			420					425					430			
Phe	Arg	Asp	Arg	Pro	Met	Ser	Pro	Leu	Asp	Thr	Ala	Val	Tyr	Trp	Thr	
		435					440					445				
Glu	Tyr	Val	Ile	Arg	His	Arg	Gly	Ala	His	His	Leu	Lys	Thr	Ala	Gly	
	450					455					460					
Ala	Phe	Leu	His	Trp	Tyr	Gln	Tyr	Leu	Leu	Leu	Asp	Val	Ile	Thr	Phe	
465					470					475					480	

Leu Leu Val Thr Phe Cys Ala Phe Cys Phe Ile Val Lys Tyr Ile Cys
 485 490 495

Lys Ala Leu Ile His His Tyr Trp Ser Ser Ser Lys Ser Glu Lys Leu
 500 505 510

Lys Lys Asn
 515

<210> 3
 <211> 1548
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự được tối ưu hóa ở nấm men

<400> 3
 atggaattca gattgttgat attggccttg ttctccgtat tgatgtctac ctctaattggt
 60

gccgaaatct tggctttatt ccctattcat ggtatatcta actacaacgt agctgaagca
 120

ttgttgaaga ctttggctaa cagaggtcac aacggtaccg ttgtaacttc atttccacaa
 180

aagaaaccag ttcctaattt gtacgaaatt gatgtatcag gtgcaaaggg tttagccaca
 240

aactccatcc atttcgaaag attgcaaacc atcatccaag atgtcaagag taacttcaag
 300

aacatgggta gattgtctag aacatactgt gaaatcatgt tctcagaccc aagagttttg
 360

aacatcagag ataaaaagtt tgacttggtt ataaacgccg tattcggttc agattgcgac
 420

gctgggtttg catggaaaag tcaagctcct ttaatatcta tottgaatgc cagacatata
 480

ccatgggctt tgcacagaat gggtaatcct tccaaccag catatatgcc tgtaatccat
 540

agtagattcc cagtcaagat gaatttcttt caaagaatga taaacaccgg ttggcactta
 600

tactttttgt acatgtactt ctactacggt aatgggtgaag atgctaacaa aatggcaaga
 660

aagtttttcg gtaatgatat gcctgacata aacgaaatgg tttttaacac ctcottgttg
720

ttcgtaaaca ctcatttcag tgtcgatatg ccataccott tagtcccaaa ctgtatcgaa
780

atcgggtgga tccatgttaa ggaaccacaa ctttgccat tggaaatcca aaagtttatg
840

gatgaagcag aacatgggtg aatctttttc acottgggta gtatggtcag aacttotaca
900

ttccctaato aaactattca agccttttaa gaagccttcg ctgaattacc acaaagagtt
960

ttgtggaagt tcgaaaacga aaacgaagat atgccttcca acgttttgat cagaaagtgg
1020

ttcccacaaa acgacatctt cggtcataag aacatcaagg ctttcatttc acacgggtgg
1080

aattccggtg ctttggaagc tgtccatttc ggtgttccta tcataggtat cccattgttt
1140

tatgatcaat acagaaacat cttgtctttc gttaaagaag gtgtagctgt cttgttggat
1200

gtaaacgact taactaagga taacatcttg tcttcagtta gaacagtcgt taacgacaag
1260

tcataactcg aaagaatgaa ggcattgtct caattgttta gagatagacc tatgtcacca
1320

ttagacacag ctgtttattg gaocgaatac gtaattagac atagaggtgc acatcactta
1380

aaaactgcag gtgccttttt gcactggat caatacttgt tgttggatgt catcacattt
1440

ttgttgggta cattctgtgc attctgcttc atcgtaaagt acatctgcaa ggccttaato
1500

catcactact ggtccagttc taaatctgaa aagttgaaaa agaattaa
1548

<210> 4

<211> 492

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 4

Met Gly Ser Asn Ala Pro Pro Pro Pro Thr Pro His Val Val Leu Val
 1 5 10 15

Pro Phe Pro Gly Gln Gly His Val Ala Pro Leu Met Gln Leu Ala Arg
 20 25 30

Leu Leu His Ala Arg Gly Ala Arg Val Thr Phe Val Tyr Thr Gln Tyr
 35 40 45

Asn Tyr Arg Arg Leu Leu Arg Ala Lys Gly Glu Ala Ala Val Arg Pro
 50 55 60

Pro Ala Thr Ser Ser Ala Arg Phe Arg Ile Glu Val Ile Asp Asp Gly
 65 70 75 80

Leu Ser Leu Ser Val Pro Gln Asn Asp Val Gly Gly Leu Val Asp Ser
 85 90 95

Leu Arg Lys Asn Cys Leu His Pro Phe Arg Ala Leu Leu Arg Arg Leu
 100 105 110

Gly Gln Glu Val Glu Gly Gln Asp Ala Pro Pro Val Thr Cys Val Val
 115 120 125

Gly Asp Val Val Met Thr Phe Ala Ala Ala Ala Ala Arg Glu Ala Gly
 130 135 140

Ile Pro Glu Val Gln Phe Phe Thr Ala Ser Ala Cys Gly Leu Leu Gly
 145 150 155 160

Tyr Leu His Tyr Gly Glu Leu Val Glu Arg Gly Leu Val Pro Phe Arg
 165 170 175

Asp Ala Ser Leu Leu Ala Asp Asp Asp Tyr Leu Asp Thr Pro Leu Glu
 180 185 190

Trp Val Pro Gly Met Ser His Met Arg Leu Arg Asp Met Pro Thr Phe
 195 200 205

Cys Arg Thr Thr Asp Pro Asp Asp Val Met Val Ser Ala Thr Leu Gln
 210 215 220

Gln Met Glu Ser Ala Ala Gly Ser Lys Ala Leu Ile Leu Asn Thr Leu
 225 230 235 240

Tyr Glu Leu Glu Lys Asp Val Val Asp Ala Leu Ala Ala Phe Phe Pro
 245 250 255

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

Ser Ala Ser Ala Gly Leu Ala Ala Met Asp Ile Ser Ile Trp Gln Glu
 275 280 285

Asp Thr Arg Cys Leu Ser Trp Leu Asp Gly Lys Pro Ala Gly Ser Val
 290 295 300

Val Tyr Val Asn Phe Gly Ser Met Ala Val Met Thr Ala Ala Gln Ala
 305 310 315 320

Arg Glu Phe Ala Leu Gly Leu Ala Ser Cys Gly Ser Pro Phe Leu Trp
 325 330 335

Val Lys Arg Pro Asp Val Val Glu Gly Glu Glu Val Leu Leu Pro Glu
 340 345 350

Ala Leu Leu Asp Glu Val Ala Arg Gly Arg Gly Leu Val Val Pro Trp
 355 360 365

Cys Pro Gln Ala Ala Val Leu Lys His Ala Ala Val Gly Leu Phe Val
 370 375 380

Ser His Cys Gly Trp Asn Ser Leu Leu Glu Ala Thr Ala Ala Gly Gln
 385 390 395 400

Pro Val Leu Ala Trp Pro Cys His Gly Glu Gln Thr Thr Asn Cys Arg
 405 410 415

Gln Leu Cys Glu Val Trp Gly Asn Gly Ala Gln Leu Pro Arg Glu Val

420 425 430
 Glu Ser Gly Ala Val Ala Arg Leu Val Arg Glu Met Met Val Gly Asp
 435 440 445
 Leu Gly Lys Glu Lys Arg Ala Lys Ala Ala Glu Trp Lys Ala Ala Ala
 450 455 460
 Glu Ala Ala Ala Arg Lys Gly Gly Ala Ser Trp Arg Asn Val Glu Arg
 465 470 475 480
 Val Val Asn Asp Leu Leu Leu Val Gly Gly Lys Gln
 485 490
 <210> 5
 <211> 471
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> ADN tổng hợp
 <400> 5
 Met Pro Ser Ser Gly Asp Ala Ala Gly Arg Arg Pro His Val Val Leu
 1 5 10 15
 Ile Pro Ser Ala Gly Met Gly His Leu Val Pro Phe Gly Arg Leu Ala
 20 25 30
 Val Ala Leu Ser Ser Gly His Gly Cys Asp Val Ser Leu Val Thr Val
 35 40 45
 Leu Pro Thr Val Ser Thr Ala Glu Ser Lys His Leu Asp Ala Leu Phe
 50 55 60
 Asp Ala Phe Pro Ala Val Arg Arg Leu Asp Phe Glu Leu Ala Pro Phe
 65 70 75 80
 Asp Ala Ser Glu Phe Pro Gly Ala Asp Pro Phe Phe Leu Arg Phe Glu
 85 90 95
 Ala Met Arg Arg Ser Ala Pro Leu Leu Gly Pro Leu Leu Thr Gly Ala

			100					105					110			
Gly	Ala	Ser	Ala	Leu	Ala	Thr	Asp	Ile	Ala	Leu	Thr	Ser	Val	Val	Ile	
		115					120					125				
Pro	Val	Ala	Lys	Glu	Gln	Gly	Leu	Pro	Cys	His	Ile	Leu	Phe	Thr	Ala	
	130					135					140					
Ser	Ala	Ala	Met	Leu	Ser	Leu	Cys	Ala	Tyr	Phe	Pro	Thr	Tyr	Leu	Asp	
145					150					155					160	
Ala	Asn	Ala	Gly	Gly	Gly	Gly	Gly	Val	Gly	Asp	Val	Asp	Ile	Pro	Gly	
				165					170					175		
Val	Tyr	Arg	Ile	Pro	Lys	Ala	Ser	Ile	Pro	Gln	Ala	Leu	His	Asp	Pro	
			180					185					190			
Asn	His	Leu	Phe	Thr	Arg	Gln	Phe	Val	Ala	Asn	Gly	Arg	Ser	Leu	Thr	
		195					200					205				
Ser	Ala	Ala	Gly	Ile	Leu	Val	Asn	Thr	Phe	Asp	Ala	Leu	Glu	Pro	Glu	
	210					215					220					
Ala	Val	Ala	Ala	Leu	Gln	Gln	Gly	Lys	Val	Ala	Ser	Gly	Phe	Pro	Pro	
225					230					235					240	
Val	Phe	Ala	Val	Gly	Pro	Leu	Leu	Pro	Ala	Ser	Asn	Gln	Ala	Lys	Asp	
				245					250					255		
Pro	Gln	Ala	Asn	Tyr	Met	Glu	Trp	Leu	Asp	Ala	Gln	Pro	Ala	Arg	Ser	
			260					265					270			
Val	Val	Tyr	Val	Ser	Phe	Gly	Ser	Arg	Lys	Ala	Ile	Ser	Arg	Glu	Gln	
		275					280					285				
Leu	Arg	Glu	Leu	Ala	Ala	Gly	Leu	Glu	Gly	Ser	Gly	His	Arg	Phe	Leu	
	290					295					300					
Trp	Val	Val	Lys	Ser	Thr	Val	Val	Asp	Arg	Asp	Asp	Ala	Ala	Glu	Leu	
305					310					315					320	

Gly Glu Leu Leu Asp Glu Gly Phe Leu Glu Arg Val Glu Lys Arg Gly
 325 330 335

Leu Val Thr Lys Ala Trp Val Asp Gln Glu Glu Val Leu Lys His Glu
 340 345 350

Ser Val Ala Leu Phe Val Ser His Cys Gly Trp Asn Ser Val Thr Glu
 355 360 365

Ala Ala Ala Ser Gly Val Pro Val Leu Ala Leu Pro Arg Phe Gly Asp
 370 375 380

Gln Arg Val Asn Ser Gly Val Val Ala Arg Ala Gly Leu Gly Val Trp
 385 390 395 400

Ala Asp Thr Trp Ser Trp Glu Gly Glu Ala Gly Val Ile Gly Ala Glu
 405 410 415

Glu Ile Ser Glu Lys Val Lys Ala Ala Met Ala Asp Glu Ala Leu Arg
 420 425 430

Met Lys Ala Ala Ser Leu Ala Glu Ala Ala Ala Lys Ala Val Ala Gly
 435 440 445

Gly Gly Ser Ser His Arg Cys Leu Ala Glu Phe Ala Arg Leu Cys Gln
 450 455 460

Gly Gly Thr Cys Arg Thr Asn
 465 470

<210> 6
 <211> 1212
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 6
 atgagttcac tctccaacgc ttcccatctg atggaggatg tgcagggcat ccggaaggcc
 60

cagagagccg atggcacggc caccgtcatg gccatcggaa cagctcacco toctoatato
120

tttcctcagg acacctacgc tgactttctac ttccgcgcc acaacagcga gcacaaggto
180

gagctcaaga agaagttcga tcgcatctgc aaaaagacaa tgataggcaa gcgctacttc
240

aactacgacg aggagttctt gaagaaatat cccaatatca cttcattcga tgagcccagc
300

ctcaacgacc gccaggacat ttgtgtccct ggggtgccag ccctgggagc cgaagcagct
360

gtgaaagcca tcgcggaatg gggacgcccc aagtctgaga ttactcatct cgtgttctgc
420

acctcctgcg gtgtcgacat gccacgcgc gacttccagt gcgccaagct ccttggcctc
480

cgcaccaatg tcaacaagta ctgcgtctac atgcaaggat gctatgctgg tggcacccgc
540

atgcggtatg ccaaggatct ggccgagaa aaccgtggtg ctggtgtttt ggtggtgtgt
600

gcggagctca ccataatcg gcttcgaggc cctaattgagt cccatctcga caacgccatc
660

ggaaattctc ttttcggaga tggagctgcc gcgttgatcg tcgggtcggg ccccatcatc
720

ggtgtcgaga agcccatggt cgagatcgtg tgtgccaagc aaactgtgat cccaaacagc
780

gaagacgtta tccatctcca catgagagag gcaggtctga tgttctacat gagcaaggac
840

agtcccgaga ccatctccaa taacgtagag gcttgccctg ttgatgtgtt caagtctgtg
900

gggatgactc ctcccagga ctggaactct ctcttctgga tccctcacco cgggtggtcgc
960

gccatccttg atcaagttga ggccaagctg aagcttcgtc ctgagaagtt ccgtgcgact
1020

cgaaccgtgc tctgggattg cggtaacatg gtcagtgctg gtgtgctcta catattggat
1080

gagatgagaa gaaaatccgc tgatgaagga ctagagacct acggagaggg actagagtgg
1140

ggtgtcttgc ttggatttgg accaggggatg accgttgaaa ctatccttct ccacagcctg
1200

cctctcatgt ga
1212

<210> 7
<211> 403
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 7

Met Ser Ser Leu Ser Asn Ala Ser His Leu Met Glu Asp Val Gln Gly
1 5 10 15

Ile Arg Lys Ala Gln Arg Ala Asp Gly Thr Ala Thr Val Met Ala Ile
20 25 30

Gly Thr Ala His Pro Pro His Ile Phe Pro Gln Asp Thr Tyr Ala Asp
35 40 45

Phe Tyr Phe Arg Ala Thr Asn Ser Glu His Lys Val Glu Leu Lys Lys
50 55 60

Lys Phe Asp Arg Ile Cys Lys Lys Thr Met Ile Gly Lys Arg Tyr Phe
65 70 75 80

Asn Tyr Asp Glu Glu Phe Leu Lys Lys Tyr Pro Asn Ile Thr Ser Phe
85 90 95

Asp Glu Pro Ser Leu Asn Asp Arg Gln Asp Ile Cys Val Pro Gly Val
100 105 110

Pro Ala Leu Gly Ala Glu Ala Ala Val Lys Ala Ile Ala Glu Trp Gly
115 120 125

Arg Pro Lys Ser Glu Ile Thr His Leu Val Phe Cys Thr Ser Cys Gly
130 135 140

Val Asp Met Pro Ser Ala Asp Phe Gln Cys Ala Lys Leu Leu Gly Leu

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

Glu Gly Leu Glu Thr Tyr Gly Glu Gly Leu Glu Trp Gly Val Leu Leu
 370 375 380

Gly Phe Gly Pro Gly Met Thr Val Glu Thr Ile Leu Leu His Ser Leu
 385 390 395 400

Pro Leu Met

<210> 8
 <211> 1212
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 8
 atgggttcac tctccaacta ctgccagtg atggaggatg tgcaggccat ccgaaaggcc
 60

caaaaagcag atggaaccgc aacagtgatg gccatcggaa cagctcaccc tcctcatatc
 120

tttcctcagg acacctacgc cgacttctac ttccgcgcca ccaacagcga gcacaaagtc
 180

gagctcaaga agaaattcga tcgtatctgc aaaaagacaa tgataggcaa gcgctacttc
 240

aattacgatg aggagttcct caagaagtat cccaacatta cctcattcga cgagcccagc
 300

ctcaacgacc gccaggatat ttgcgtccct ggggtgccgg ccctgggagc cgaagcagct
 360

gtcaaagcca tcgctgaatg gggacgtoca aagtctgaga tcaactcatct cgtcttctgc
 420

acctcctgcg gtgtcgacat gcttagcgcc gaattccagt gcgccaagct cctcggcctc
 480

cgcaccaatg tcaacaagta ttgcgtctac atgcagggat gctatgctgg cggcacagtc
 540

atgcggtacg ccaaggatct cgccgagaac aaccgtggtg ctcggtttct agtgggtgtgc
 600

gccgagctca ccatcatcgg gcttcgcgga ccaaatgagt cccatctcga caacgccatc
 660

ggcaactccc ttttcggaga cggagctgct gcgctgatcg tcgggtcaga ccccatcatt
720

ggtgtcgaga ggcctatggt cgagatcgtg tgcgcaaagc agaccgtgat cccaaacagt
780

gaagatgtta tccatctcca catgagggag gcgggtctaa tgttctacat gagcaaggac
840

agccccgaga ccatctccaa caatgtagag gcatgccttg tcgatgtggt caagtcggtg
900

gggatgactc ctcccagga ctggaactct ctcttctgga tccctcacc cggcggtcga
960

gctatcctcg accaggttga ggccaggctt aagcttcgtc ccgagaagtt cggcgcgact
1020

cgaactgtgc tctgggattg cggaaacatg gtgagcgcgt gtgttctcta cattttggat
1080

gagatgagaa gaaaatctgt tgccgacgga ctagcaacct acggagaggg gctggagtgg
1140

ggtgtcttgc ttggtttcgg accagggatg accgttgaaa ctatccttct ccacagcctg
1200

ccccctgtgt aa
1212

<210> 9
<211> 403
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 9

Met	Gly	Ser	Leu	Ser	Asn	Tyr	Ser	Pro	Val	Met	Glu	Asp	Val	Gln	Ala
1				5					10					15	

Ile	Arg	Lys	Ala	Gln	Lys	Ala	Asp	Gly	Thr	Ala	Thr	Val	Met	Ala	Ile
			20					25					30		

Gly	Thr	Ala	His	Pro	Pro	His	Ile	Phe	Pro	Gln	Asp	Thr	Tyr	Ala	Asp
		35					40					45			

Phe Tyr Phe Arg Ala Thr Asn Ser Glu His Lys Val Glu Leu Lys Lys
 50 55 60

Lys Phe Asp Arg Ile Cys Lys Lys Thr Met Ile Gly Lys Arg Tyr Phe
 65 70 75 80

Asn Tyr Asp Glu Glu Phe Leu Lys Lys Tyr Pro Asn Ile Thr Ser Phe
 85 90 95

Asp Glu Pro Ser Leu Asn Asp Arg Gln Asp Ile Cys Val Pro Gly Val
 100 105 110

Pro Ala Leu Gly Ala Glu Ala Ala Val Lys Ala Ile Ala Glu Trp Gly
 115 120 125

Arg Pro Lys Ser Glu Ile Thr His Leu Val Phe Cys Thr Ser Cys Gly
 130 135 140

Val Asp Met Pro Ser Ala Asp Phe Gln Cys Ala Lys Leu Leu Gly Leu
 145 150 155 160

Arg Thr Asn Val Asn Lys Tyr Cys Val Tyr Met Gln Gly Cys Tyr Ala
 165 170 175

Gly Gly Thr Val Met Arg Tyr Ala Lys Asp Leu Ala Glu Asn Asn Arg
 180 185 190

Gly Ala Arg Val Leu Val Val Cys Ala Glu Leu Thr Ile Ile Gly Leu
 195 200 205

Arg Gly Pro Asn Glu Ser His Leu Asp Asn Ala Ile Gly Asn Ser Leu
 210 215 220

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

Gly Val Glu Arg Pro Met Phe Glu Ile Val Cys Ala Lys Gln Thr Val
 245 250 255

Ile Pro Asn Ser Glu Asp Val Ile His Leu His Met Arg Glu Ala Gly
 260 265 270

Leu Met Phe Tyr Met Ser Lys Asp Ser Pro Glu Thr Ile Ser Asn Asn
 275 280 285

Val Glu Ala Cys Leu Val Asp Val Phe Lys Ser Val Gly Met Thr Pro
 290 295 300

Pro Glu Asp Trp Asn Ser Leu Phe Trp Ile Pro His Pro Gly Gly Arg
 305 310 315 320

Ala Ile Leu Asp Gln Val Glu Ala Arg Leu Lys Leu Arg Pro Glu Lys
 325 330 335

Phe Gly Ala Thr Arg Thr Val Leu Trp Asp Cys Gly Asn Met Val Ser
 340 345 350

Ala Cys Val Leu Tyr Ile Leu Asp Glu Met Arg Arg Lys Ser Val Ala
 355 360 365

Asp Gly Leu Ala Thr Tyr Gly Glu Gly Leu Glu Trp Gly Val Leu Leu
 370 375 380

Gly Phe Gly Pro Gly Met Thr Val Glu Thr Ile Leu Leu His Ser Leu
 385 390 395 400

Pro Pro Val

<210> 10
 <211> 1218
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 10
 atgggttcga tcgccgagtc ttcaccactg atgagtaggg agaattgtgga gggcatcaga
 60

aaagcgcaga gagctgaggg aaccgcaact gtgatggcca tcggaactgc tcacctccc
 120

catatctttc ctcaggacac ctacgcagac ttctacttcc ggcgccacaa cagcgagcac
180

aaagttgagc tcaagaagaa gttcgaccga atctgcaaaa agacaatgat tggcaaacgc
240

tacttcaact acgacgagga gttcctcaag aagtacccaa acatcacatc cttcgacgag
300

cccagcctga acgaccgcca ggacatctgc gtccccggag tccccgcctt gggtaaggag
360

gccgctctca aagccatcga ggaatggggg caacctctgt ccaagatcac ccatctcgtc
420

ttctgcacct cctgcggcgt cgacatgcc agcgccgatt tccagctcgc caagctcctc
480

ggactcaaca ccaacgtcaa caagtactgc gtctacatgc agggctgcta cgccggcggc
540

accgtcctcc gctacgcaa ggacctcgcc gagaacaacc gcggctcccg cgtcctcgtc
600

gtctgcgccg agctcaccat catcggcctc cgcggcccaa acgagtccca cctcgacaac
660

gccatcgga actccctctt cggcgacggg gcgcgcgcgc tcacgtggg cgccgatcct
720

attgttgga tagagaagcc catcttcgag atcgtctcgc caaagcagac cgtcatcccc
780

gacagcgagg acgtcatcca cctccacctc cgcgaggccg gcctcatggt ctacatgagc
840

aaggacagcc ccgagaccat ctccaacaac gtcgagggct gcctcgtcga catcttcaag
900

tccgtcggca tgaccccgcc cgccgactgg aactccctct tctggatccc ccacccggc
960

ggccgagcca tcctcgacga ggtcgaggcc aggtcaagc tccgcccga gaagtttaga
1020

gcaaccaggc acgtgctctg ggagtacggg aacatggtca gcgcatgct tctctacata
1080

ctggacgaga tgaggaacaa gtccgcagcc gacggattgg ggacctacgg agaaggactc
1140

gaatggggcg tcttgctcgg ttctggacct ggaatgaccg tcgagaccat cctcctccac
1200

agcctgcctc ctgtctaa
1218

<210> 11
<211> 405
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 11

Met Gly Ser Ile Ala Glu Ser Ser Pro Leu Met Ser Arg Glu Asn Val
1 5 10 15

Glu Gly Ile Arg Lys Ala Gln Arg Ala Glu Gly Thr Ala Thr Val Met
20 25 30

Ala Ile Gly Thr Ala His Pro Pro His Ile Phe Pro Gln Asp Thr Tyr
35 40 45

Ala Asp Phe Tyr Phe Arg Ala Thr Asn Ser Glu His Lys Val Glu Leu
50 55 60

Lys Lys Lys Phe Asp Arg Ile Cys Lys Lys Thr Met Ile Gly Lys Arg
65 70 75 80

Tyr Phe Asn Tyr Asp Glu Glu Phe Leu Lys Lys Tyr Pro Asn Ile Thr
85 90 95

Ser Phe Asp Glu Pro Ser Leu Asn Asp Arg Gln Asp Ile Cys Val Pro
100 105 110

Gly Val Pro Ala Leu Gly Lys Glu Ala Ala Leu Lys Ala Ile Glu Glu
115 120 125

Trp Gly Gln Pro Leu Ser Lys Ile Thr His Leu Val Phe Cys Thr Ser
130 135 140

Cys Gly Val Asp Met Pro Ser Ala Asp Phe Gln Leu Ala Lys Leu Leu
145 150 155 160

Gly Leu Asn Thr Asn Val Asn Lys Tyr Cys Val Tyr Met Gln Gly Cys
 165 170 175

Tyr Ala Gly Gly Thr Val Leu Arg Tyr Ala Lys Asp Leu Ala Glu Asn
 180 185 190

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

Gly Leu Arg Gly Pro Asn Glu Ser His Leu Asp Asn Ala Ile Gly Asn
 210 215 220

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

Ile Val Gly Ile Glu Lys Pro Ile Phe Glu Ile Val Cys Ala Lys Gln
 245 250 255

Thr Val Ile Pro Asp Ser Glu Asp Val Ile His Leu His Leu Arg Glu
 260 265 270

Ala Gly Leu Met Phe Tyr Met Ser Lys Asp Ser Pro Glu Thr Ile Ser
 275 280 285

Asn Asn Val Glu Gly Cys Leu Val Asp Ile Phe Lys Ser Val Gly Met
 290 295 300

Thr Pro Pro Ala Asp Trp Asn Ser Leu Phe Trp Ile Pro His Pro Gly
 305 310 315 320

Gly Arg Ala Ile Leu Asp Glu Val Glu Ala Arg Leu Lys Leu Arg Pro
 325 330 335

Glu Lys Phe Arg Ala Thr Arg His Val Leu Trp Glu Tyr Gly Asn Met
 340 345 350

Val Ser Ala Cys Val Leu Tyr Ile Leu Asp Glu Met Arg Asn Lys Ser
 355 360 365

Ala Ala Asp Gly Leu Gly Thr Tyr Gly Glu Gly Leu Glu Trp Gly Val
 370 375 380

Leu Leu Gly Phe Gly Pro Gly Met Thr Val Glu Thr Ile Leu Leu His
 385 390 395 400

Ser Leu Pro Pro Val
 405

<210> 12
 <211> 1212
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 12
 atgggtagtt tatctgactc cactccattg atgaaggatg tccaaggtat tagaaaagca
 60

caaaaagccg acggtacagc aacagttatg gccattggta ccgctcatcc acctcacata
 120

atctctcaag attcatacgc agacttctac ttcagagtta caaactcaga acataaggct
 180

gaattaaaaa agaaattcga tagaatctgt aagaaaacca tgataggcaa gagatacttc
 240

aacttcgatg aagaattctt gaagaaatac ccaaacataa catccttcga caaacctagt
 300

ttaaacgata gacatgacat ttgcatacca ggtgttcctg ctttgggtgc agaagctgca
 360

gtcaaggcaa tagaagaatg gggtagacca aagtccgaaa taacacactt ggtattttgt
 420

acctccggtg gtgttgatat gcctagtgtc gacttccaat gcgcaaagtt gttagggttg
 480

agaaccaacg tcaacaagta ctgtatctac atgcaagggt gctacgctgg tggtagctgt
 540

atgagatacg caaaagattt ggccgaaaat aacagagggt ccagagtatt aatgggttgt
 600

gctgaattga ctatcattgg tttaagaggt ccaaatgatt ctcatataga caatgccatc
 660

ggtaactcat tgtttgggtg tgggtgcgct gcattaattg tcggttcaga cccaataatc
 720

ggtgtagaaa agcctatggt cgaaattgtc tgtgctaaac aaactgtaat acctaattcc
780

gaagaagtta tccatttgca cttgagagaa agtggtttga tgttctacat gactaaggat
840

tctgccgcta caatatcaaa caacatcgaa gcttgcttgg ttgatgtctt taaatctggt
900

ggtatgacac cacctgaaga ctggaattca ttgttctgga tcccacatcc tgggtggtaga
960

gcaatttttag atcaagttga agccaagttg aaattaagac cagaaaagtt ttccgctact
1020

agaacagtat tatgggatta tggtaacatg atcagtgcac gtgttttgta catcttggat
1080

gaaatgagaa gaaagtctgc agccgaaggt ttggaaacat acggtgaagg tttagaatgg
1140

ggtgttttgt taggtttcgg tcctggatat accatcgaaa ctattttggt acactctttg
1200

ccacctgttt aa
1212

<210> 13
<211> 403
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 13

Met	Gly	Ser	Leu	Ser	Asp	Ser	Thr	Pro	Leu	Met	Lys	Asp	Val	Gln	Gly
1				5					10					15	

Ile	Arg	Lys	Ala	Gln	Lys	Ala	Asp	Gly	Thr	Ala	Thr	Val	Met	Ala	Ile
			20					25					30		

Gly	Thr	Ala	His	Pro	Pro	His	Ile	Ile	Ser	Gln	Asp	Ser	Tyr	Ala	Asp
		35					40					45			

Phe	Tyr	Phe	Arg	Val	Thr	Asn	Ser	Glu	His	Lys	Val	Glu	Leu	Lys	Lys
	50					55					60				

Lys Phe Asp Arg Ile Cys Lys Lys Thr Met Ile Gly Lys Arg Tyr Phe
65 70 75 80

Asn Phe Asp Glu Glu Phe Leu Lys Lys Tyr Pro Asn Ile Thr Ser Phe
85 90 95

Asp Lys Pro Ser Leu Asn Asp Arg His Asp Ile Cys Ile Pro Gly Val
100 105 110

Pro Ala Leu Gly Ala Glu Ala Ala Val Lys Ala Ile Glu Glu Trp Gly
115 120 125

Arg Pro Lys Ser Glu Ile Thr His Leu Val Phe Cys Thr Ser Gly Gly
130 135 140

Val Asp Met Pro Ser Ala Asp Phe Gln Cys Ala Lys Leu Leu Gly Leu
145 150 155 160

Arg Thr Asn Val Asn Lys Tyr Cys Ile Tyr Met Gln Gly Cys Tyr Ala
165 170 175

Gly Gly Thr Val Met Arg Tyr Ala Lys Asp Leu Ala Glu Asn Asn Arg
180 185 190

Gly Ala Arg Val Leu Met Val Cys Ala Glu Leu Thr Ile Ile Gly Leu
195 200 205

Arg Gly Pro Asn Asp Ser His Ile Asp Asn Ala Ile Gly Asn Ser Leu
210 215 220

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

Gly Val Glu Lys Pro Met Phe Glu Ile Val Cys Ala Lys Gln Thr Val
245 250 255

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

Leu Met Phe Tyr Met Thr Lys Asp Ser Ala Ala Thr Ile Ser Asn Asn

275 280 285
 Ile Glu Ala Cys Leu Val Asp Val Phe Lys Ser Val Gly Met Thr Pro
 290 295 300
 Pro Glu Asp Trp Asn Ser Leu Phe Trp Ile Pro His Pro Gly Gly Arg
 305 310 315 320
 Ala Ile Leu Asp Gln Val Glu Ala Lys Leu Lys Leu Arg Pro Glu Lys
 325 330 335
 Phe Ser Ala Thr Arg Thr Val Leu Trp Asp Tyr Gly Asn Met Ile Ser
 340 345 350
 Ala Cys Val Leu Tyr Ile Leu Asp Glu Met Arg Arg Lys Ser Ala Ala
 355 360 365
 Glu Gly Leu Glu Thr Tyr Gly Glu Gly Leu Glu Trp Gly Val Leu Leu
 370 375 380
 Gly Phe Gly Pro Gly Met Thr Ile Glu Thr Ile Leu Leu His Ser Leu
 385 390 395 400
 Pro Pro Val

 <210> 14
 <211> 1182
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ADN tổng hợp

 <400> 14
 atgggttccc ttgacaatgg ttcagctaga attaacaacc agaaatctaa tgggttggct
 60
 tcaatttttg ccattggaac tgcacttccg ccgatttgca ttaagcaaga tgactatcct
 120
 gattactact tccgagtcac caaaagcgac cacaagacgc aactgaaaga gaagtttcgt
 180

cgcatctgtg aaaagtcagg agtgacaaaa cgatacacag tactaaccga agacatgac
240

aaggagaacg agaacataat aacctacaag gctccgtcac tggatgctcg ccaagcgac
300

ctacacaagg agacacccaa gctcgccata gaagcagcct tgaagaccat ccaagaatgg
360

ggccaacccg tctctaagat caccacactg ttcttttgct cctcctctgg cggctgctat
420

cttccgagct ccgattttca gatcgctaag gcactcggcc tcgagccgac cgtccagagg
480

tccatgggtgt ttctcatgg atgctatgct gccagttctg gcctgcgttt ggccaaggac
540

attgcagaga acaacaaaga tgcacgcgtg ctgggtgggtg gctgcgagtt gatgggtgctg
600

agcttccatg caccatcgga ggacgcgac ggaatgctaa taggtcatgc catcttcggc
660

gatggagcgg cctgcgcaat tgtaggagca gaccggggc ctacggagcg cccaatattc
720

gagctagtga agggcggaca ggtgatcgtc ccagacacgg aagactgtct gggaggggtg
780

gtgatggaga tgggatggat ctacgatctc aacaagcgcc ttctcaagc cctagccgac
840

aacatcctcg gagccctaga tgacaccctg aggctgacag gtaaaaggga tgacctcaat
900

ggccttttct acgtgctcca cccgggtggg cgggccatca tcgacctgct tgaggagaag
960

cttgagctaa caaaggacaa gctcgagagt agccgtcgtg tgctcagcaa ctatggcaac
1020

atgtggggcc ctgcgctagt gttcacgctc gacgagatga ggaggaagtc aaaggaggac
1080

aacgccacca cactggtgg cgggtccgag ctcggcctga tgatggcggt tggacctggc
1140

ctcaccaccg agatcatggt tctccgaagt gtgcctctct aa
1182

<210> 15
<211> 393

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 15

Met Gly Ser Leu Asp Asn Gly Ser Ala Arg Ile Asn Asn Gln Lys Ser
 1 5 10 15

Asn Gly Leu Ala Ser Ile Leu Ala Ile Gly Thr Ala Leu Pro Pro Ile
 20 25 30

Cys Ile Lys Gln Asp Asp Tyr Pro Asp Tyr Tyr Phe Arg Val Thr Lys
 35 40 45

Ser Asp His Lys Thr Gln Leu Lys Glu Lys Phe Arg Arg Ile Cys Glu
 50 55 60

Lys Ser Gly Val Thr Lys Arg Tyr Thr Val Leu Thr Glu Asp Met Ile
 65 70 75 80

Lys Glu Asn Glu Asn Ile Ile Thr Tyr Lys Ala Pro Ser Leu Asp Ala
 85 90 95

Arg Gln Ala Ile Leu His Lys Glu Thr Pro Lys Leu Ala Ile Glu Ala
 100 105 110

Ala Leu Lys Thr Ile Gln Glu Trp Gly Gln Pro Val Ser Lys Ile Thr
 115 120 125

His Leu Phe Phe Cys Ser Ser Ser Gly Gly Cys Tyr Leu Pro Ser Ser
 130 135 140

Asp Phe Gln Ile Ala Lys Ala Leu Gly Leu Glu Pro Thr Val Gln Arg
 145 150 155 160

Ser Met Val Phe Pro His Gly Cys Tyr Ala Ala Ser Ser Gly Leu Arg
 165 170 175

Leu Ala Lys Asp Ile Ala Glu Asn Asn Lys Asp Ala Arg Val Leu Val
 180 185 190

Val Cys Cys Glu Leu Met Val Ser Ser Phe His Ala Pro Ser Glu Asp
 195 200 205

Ala Ile Gly Met Leu Ile Gly His Ala Ile Phe Gly Asp Gly Ala Ala
 210 215 220

Cys Ala Ile Val Gly Ala Asp Pro Gly Pro Thr Glu Arg Pro Ile Phe
 225 230 235 240

Glu Leu Val Lys Gly Gly Gln Val Ile Val Pro Asp Thr Glu Asp Cys
 245 250 255

Leu Gly Gly Trp Val Met Glu Met Gly Trp Ile Tyr Asp Leu Asn Lys
 260 265 270

Arg Leu Pro Gln Ala Leu Ala Asp Asn Ile Leu Gly Ala Leu Asp Asp
 275 280 285

Thr Leu Arg Leu Thr Gly Lys Arg Asp Asp Leu Asn Gly Leu Phe Tyr
 290 295 300

Val Leu His Pro Gly Gly Arg Ala Ile Ile Asp Leu Leu Glu Glu Lys
 305 310 315 320

Leu Glu Leu Thr Lys Asp Lys Leu Glu Ser Ser Arg Arg Val Leu Ser
 325 330 335

Asn Tyr Gly Asn Met Trp Gly Pro Ala Leu Val Phe Thr Leu Asp Glu
 340 345 350

Met Arg Arg Lys Ser Lys Glu Asp Asn Ala Thr Thr Thr Gly Gly Gly
 355 360 365

Ser Glu Leu Gly Leu Met Met Ala Phe Gly Pro Gly Leu Thr Thr Glu
 370 375 380

Ile Met Val Leu Arg Ser Val Pro Leu
 385 390

<210> 16
 <211> 510
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 16
 atgagacacg ttgaacacac agttaccgtt gccgctccag cagatttggt ttgggaagtt
 60
 ttagccgatg tattaggtta cgcagatata ttcccaccta cagaaaaggt tgaaatcttg
 120
 gaagaaggtc aaggttacca agttgtaaga ttgcatgtag atgtcgccgg tgaaattaat
 180
 acttgacat ctagaagaga tttggaccca gctagaagag ttatcgcata cagacaatta
 240
 gaaacagcac ctattgtagg tcatatgtct ggtgaatgga gagcctttac cttggatgct
 300
 gaaagaacac aattggtctt aaccacgac ttcgttacta gagctgcagg tgacgacggt
 360
 ttggttgctg gtaaattaac tccagatgaa gcaagagaaa tgttagaagc cgtcgttgaa
 420
 agaaattccg tagcagactt gaacgccgtc ttaggtgaag ctgaaagaag agttagagcc
 480
 gctggtggtg ttggtaccgt aactgcataa
 510

<210> 17
 <211> 169
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 17

Met	Arg	His	Val	Glu	His	Thr	Val	Thr	Val	Ala	Ala	Pro	Ala	Asp	Leu
1				5				10						15	

Val	Trp	Glu	Val	Leu	Ala	Asp	Val	Leu	Gly	Tyr	Ala	Asp	Ile	Phe	Pro
			20					25					30		

Pro Thr Glu Lys Val Glu Ile Leu Glu Glu Gly Gln Gly Tyr Gln Val
 35 40 45

Val Arg Leu His Val Asp Val Ala Gly Glu Ile Asn Thr Trp Thr Ser
 50 55 60

Arg Arg Asp Leu Asp Pro Ala Arg Arg Val Ile Ala Tyr Arg Gln Leu
 65 70 75 80

Glu Thr Ala Pro Ile Val Gly His Met Ser Gly Glu Trp Arg Ala Phe
 85 90 95

Thr Leu Asp Ala Glu Arg Thr Gln Leu Val Leu Thr His Asp Phe Val
 100 105 110

Thr Arg Ala Ala Gly Asp Asp Gly Leu Val Ala Gly Lys Leu Thr Pro
 115 120 125

Asp Glu Ala Arg Glu Met Leu Glu Ala Val Val Glu Arg Asn Ser Val
 130 135 140

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

Ala Gly Gly Val Gly Thr Val Thr Ala
 165

<210> 18

<211> 771

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 18

atgtccggta gaaagacctt tttagatttg tcatttgcca ccagagatac accatccgaa
 60

gccactccag tcgttggtga tttgttggat catgtcactg gtgccacagt tttgggttta
 120

tccccagaag attttctga cggtatggca atcagtaacg aaaccgttac tttgactaca
 180

catacaggta cccacatgga tgctccatta cattatgggc ctttgtctgg tgggtgtacca
240

gcaaaatcaa tcgaccaagt cccattagaa tgggtgttacg gtccctgggtg tagattggat
300

gttagacacg taccagctgg tgacgggtatc actgtagacc atttgaatgc tgcattggat
360

gccgctgaac acgacttggc tcctgggtgac attgttatgt tatggacagg tgctgatgca
420

ttgtggggta ctagagaata tttgtctaca tttccagggt tgaccggtaa aggtactcaa
480

ttcttagttg aagcagggtg caagggttatt ggtatagacg cctgggggtt ggatagacca
540

atggcagcca tgattgaaga atatagaaga accgggtgaca aagggtgcatt atggcctgcc
600

catgtttatg gtagaacaag agaatacttg caattggaaa agttgaacaa cttaggtgct
660

ttgccagggtg caactgggtta cgacatatca tgctttcctg tagccgtcgc tggtagagg
720

gccggttgga ccagagttgt agctgttttc gaacaagaag aagaagatta a
771

<210> 19

<211> 256

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 19

Met	Ser	Gly	Arg	Lys	Thr	Phe	Leu	Asp	Leu	Ser	Phe	Ala	Thr	Arg	Asp
1				5				10						15	

Thr	Pro	Ser	Glu	Ala	Thr	Pro	Val	Val	Val	Asp	Leu	Leu	Asp	His	Val
			20					25					30		

Thr	Gly	Ala	Thr	Val	Leu	Gly	Leu	Ser	Pro	Glu	Asp	Phe	Pro	Asp	Gly
		35					40						45		

Met Ala Ile Ser Asn Glu Thr Val Thr Leu Thr Thr His Thr Gly Thr
 50 55 60

His Met Asp Ala Pro Leu His Tyr Gly Pro Leu Ser Gly Gly Val Pro
 65 70 75 80

Ala Lys Ser Ile Asp Gln Val Pro Leu Glu Trp Cys Tyr Gly Pro Gly
 85 90 95

Val Arg Leu Asp Val Arg His Val Pro Ala Gly Asp Gly Ile Thr Val
 100 105 110

Asp His Leu Asn Ala Ala Leu Asp Ala Ala Glu His Asp Leu Ala Pro
 115 120 125

Gly Asp Ile Val Met Leu Trp Thr Gly Ala Asp Ala Leu Trp Gly Thr
 130 135 140

Arg Glu Tyr Leu Ser Thr Phe Pro Gly Leu Thr Gly Lys Gly Thr Gln
 145 150 155 160

Phe Leu Val Glu Ala Gly Val Lys Val Ile Gly Ile Asp Ala Trp Gly
 165 170 175

Leu Asp Arg Pro Met Ala Ala Met Ile Glu Glu Tyr Arg Arg Thr Gly
 180 185 190

Asp Lys Gly Ala Leu Trp Pro Ala His Val Tyr Gly Arg Thr Arg Glu
 195 200 205

Tyr Leu Gln Leu Glu Lys Leu Asn Asn Leu Gly Ala Leu Pro Gly Ala
 210 215 220

Thr Gly Tyr Asp Ile Ser Cys Phe Pro Val Ala Val Ala Gly Thr Gly
 225 230 235 240

Ala Gly Trp Thr Arg Val Val Ala Val Phe Glu Gln Glu Glu Glu Asp
 245 250 255

<210> 20
 <211> 1212

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 20

atgagtagtt tatcaaatgc cagtcactta atggaagatg ttcaaggtat cagaaaagcc
60

caaagagccg acggtacagc cacagtcattg gctatcggta ccgcacatcc acctcacatt
120

tttcacaaag atacatatgc tgactttttac ttccagagcaa ccaattccga acataagggt
180

gaattgaaaa agaaattcga tagaatctgt aagaaaacaa tgatcggcaa gagatacttc
240

aactacgacg aagaattctt aaagaaatac ccaaacatca cctccttcga tgaacctagt
300

ttgaacgata gacaagacat ttgcgtccca ggtgtacctg cattaggtgc cgaagctgca
360

gttaaggcca ttgctgaatg gggtagacca aaatctgaaa taaccactt ggttttctgt
420

acttcttgcg gtgtcgatat gccttcagca gacttccaat gtgccaagtt gttgggtttg
480

agaactaacg taaacaagta ctgtgtttac atgcaagggt gctatgctgg tggtaacgtt
540

atgagatacg ccaagatttt agctgaaaat aacagagggt ccagagtctt ggttgtctgt
600

gctgaattga ctattatagg tttaagaggt ccaaatgaat cccatttgga taatgcaatc
660

ggtaacagtt tatttggtga cgggtgcgct gcattgattg taggttctga tccaatcatt
720

ggtgttgaaa agcctatggt cgaaattgtc tgtgctaagc aaacagtaat cccaaactca
780

gaagacgtta tccatttgca catgagagaa gcaggtttaa tgttttacat gtctaaggat
840

tcacctgaaa ccatttcaaa taacgtcgaa gcttgcttgg tagacgtttt taaatctgtt
900

ggtagactc cacctgaaga ttggaattca ttattctgga taccacatcc tgggtggtaga
960

gcaatcttgg atcaagttga agccaagttg aaattaagac ctgaaaagtt cagagctact
1020

agaacagttt tatgggattg tggtaacatg gtttccgcat gcgtcttgta catattggat
1080

gaaatgagaa gaaagagtgc cgacgaaggt ttggaaacat acggtgaagg tttagaatgg
1140

ggtgttttgt taggttttgg tcctgggatg actgtagaaa ctatcttatt gcattccttg
1200

ccattgatgt ga
1212

<210> 21
<211> 1212
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 21
atgggttcat tatcaaacta ttcacctgtt atggaagatg ttcaagcaat cagaaaggct
60

caaaaagcag atggtaccgc tactgttatg gctatcggta ctgcacatcc acctcacatt
120

tttcacaag atacttatgc cgacttttac ttcagagcta caaactccga acataagggt
180

gaattgaaaa agaaattcga tagaatctgt aagaaaacta tgatcggcaa gagatacttc
240

aactacgacg aagaattctt aaagaaatac ccaaacatca catccttcga tgaacctagt
300

ttgaacgata gacaagacat ctgcgtccca ggtgtacctg cattaggtgc cgaagctgca
360

gttaaggcca ttgctgaatg gggtagacca aaatctgaaa taacacactt ggttttctgt
420

acctcttgcg gtgtcgatat gccttcagct gacttccaat gtgcaaagtt gttggggttg
480

agaaccaacg taaacaagta ctgtgtttac atgcaagggt gctatgcagg tgggtactgta
540

atgagatacg ccaaggattt agctgaaaat aacagagggtg caagagtttt gggtgtctgt
600

gccgaattga ccattatagg tttaagaggt ccaaataaat cccatttgga taatgctatc
660

ggtaacagtt tatttggtga cgggtgcccgt gcattgattg taggttctga tccaatcatt
720

gggtgtgaaa gacctatgtt cgaaattgtc tgtgcaaagc aaacagtaat cccaaactca
780

gaagacgtta tccatttgca catgagagaa gctgggttaa tgttttacat gtctaaggat
840

tcacctgaaa caatctcaaa caacgttgaa gcttgcttgg tagacgtttt taaatctgtc
900

ggtatgaccc cacctgaaga ttggaattca ttattctgga taccacatcc tgggtgtaga
960

gctatcttgg atcaagtcga agcaagattg aagttaagac cagaaaaatt tggtgccact
1020

agaacagttt tatgggattg tggtaacatg gtttccgctt gcgtcttgta catattggat
1080

gaaatgagaa gaaagagtgt tgcagacggt ttggccacat acggtgaagg tttagaatgg
1140

gggtgtcttgt taggtttcgg tccctggtatg accgtagaaa ctattttgtt acactcttta
1200

ccacctgttt aa
1212

<210> 22
<211> 1218
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 22
atgggtagta tcgocgaaag tagtccattg atgtccagag aaaacgtcga aggtataaga
60

aaggcacaaa gagccgaagg tacagccacc gttatggcaa taggtactgc ccatccacct
120

cacatctttc cacaagatac ttatgcagac ttttacttca gagccacaaa ctcagaacat
180

aaggttgaat taaaaaagaa attcgataga atttgtaaga aaactatgat cggcaagaga
240

tacttcaact acgacgaaga attcttaaag aaatacccaa acatcacatc cttcgatgaa
300

cctagtttga atgatagaca agacatttgc gtaccaggtg ttcctgcott gggcaaggaa
360

gctgcattaa aagctatcga agaattgggt caaccattgt ccaagattac acacttagta
420

ttttgtacct cttgcggtgt tgatatgcct tcagccgact tccaattggc taagttggtg
480

ggtttgaaca ccaacgttaa caagtactgt gtctacatgc aaggttgcta tgcaggtggt
540

actgttttga gatacgctaa agatttagca gaaaataaca gaggttctag agtcttggtt
600

gtctgtgccg aattgactat tataggttta agagggtcaa atgaatccca tttggataac
660

gctattggtg acagtttggt tgggtgacggt gccgctgcat tgatagtcgg tgctgatcca
720

atcgtaggta ttgaaaagcc tatattcgaa atcgtttggtg caaaacaaac agtcatccca
780

gattctgaag acgttattca tttgcacttg agagaagcag gtttgatggt ctacatgtct
840

aaggattcac ctgaaacaat ctcaaacaac gtagaagggt gcttggttga catctttaa
900

tctgttggtg tgacccacc tgctgattgg aattcattat tctggattcc acatcctggt
960

ggtagagcta tattggatga agttgaagca agattgaagt tgagaccaga aaagttcaga
1020

gcaaccagac acgttttatg ggaatatggt aacatggtct ccgcctgtgt attgtacata
1080

ttggatgaaa tgagaaacaa aagtgccgct gacggtttgg gtacatacgg tgaaggttta
1140

gaatggggtg tcttgttagg tttcggctct ggtatgactg tagaaacaat attgttgcat
1200

tctttgccac ctgtttta
1218

<210> 23
 <211> 1180
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 23
 gggttcctta gacaacggtt ccgccagaat caacaatcaa aagtccaacg gtttagcatc
 60
 aatccttagca atcgggtacag ccttgccacc tatatgtatc aagcaagatg actaccaga
 120
 ttactacttc agagttacaa agtcagacca taagacccaa ttgaaggaaa agttcagaag
 180
 aatatgcgaa aaatccggtg ttacaaagag atacactgtc ttgacagaag atatgatcaa
 240
 ggaaaacgaa aacataatca cctacaaggc cccaagttta gatgcaagac aagccatttt
 300
 gcataaagaa actcctaagt tagcaattga agctgcattg aaaacaatac aagaatgggg
 360
 tcaaccagtt tctaagatca cacacttatt tttctgttct tcatccggtg gttgctattt
 420
 gcctagttct gatttcctaaa tagctaaggc attgggttta gaaccaactg tccaaagatc
 480
 aatggtattc cctcacggtt gttacgccgc ttcacccggt ttgagattgg ctaaggatat
 540
 cgcagaaaac aacaaggacg ctagagtatt agttgtctgt tgcgaattga tggtttcttc
 600
 ttttcatgca ccatccgaag atgccattgg catgttaata ggtcacgcta tcttcggtga
 660
 tgggtgcagcc tgtgccattg ttggtgctga ccaggtcct accgaaagac caatctttga
 720
 attagtcaaa ggtggtcaag taattgttcc tgatactgaa gactgcttgg gtggttgggt
 780
 tatggaaatg ggttggtatc acgatttgaa taagagattg ccacaagcct tggctgacaa
 840

catcttgggt gcttttagatg acaccttgag attaactggc aagagagatg acttgaatgg
900

tttgttttat gttttgcac ctaggtggtag agcaatcac gatttggttg aagaaaagtt
960

ggaattgaca aaggataagt tagaatcac cagaagagtc ttgtcaaatt acggtaacat
1020

gtggggtcct gcttttagtat tcactttgga tgaaatgaga agaaagagta aggaagacaa
1080

cgcaactaca accggtggtg gttctgaatt gggtttaatg atggcttttg gtccaggttt
1140

gactaccgaa ataatggttt tgagaagtgt tcctttgtga
1180

<210> 24

<211> 510

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 24

atgaggcatg tcgagcatac ggtaactgta gctgcacctg ctgatttggt gtgggaagtt
60

cttgctgatg tgctgggcta tgctgacatt ttccctccga ctgaaaaggt tgaaatcctt
120

gaagaaggac agggttatca agtagttaga ctccacgtag atgttgcagg tgagatcaat
180

acatggacgt ccagacgtga tttggatcct gcaagaagag tgatagctta taggcagtta
240

gaaacagctc caattgttgg gcatatgtct ggcgagtgga gagcttttac acttgatgca
300

gaacggactc aacttgtttt aacacatgac ttctgacta gagcagcagg tgatgacggg
360

cttgttgccg gtaagcttac ccagatgag gccagagaaa tgctcgaggc cgtagtggag
420

aggaactcag ttgctgattt gaacgctgtc ttgggcgaag ccgagcgtcg gggtcgagca
480

gctgggggtg tcggaactgt cacggcatga
510

<210> 25
 <211> 510
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 25
 atgctgcacg tcgaacacac cgtcaccgta gcggccccgg ccgacctggt ctgggaggtg
 60
 ctggccgatg tgctcggcta cgccgacatc ttcccgcga ccgagaaggt cgagatcctc
 120
 gaggagggtc agggctacca ggtcgtccgc ctgcacgtgg acgtcgccgg cgagatcaac
 180
 acctggacct cgcgccgcga cctggaccgg gcccgccgg tcacgccta ccggcagttg
 240
 gagacggccc cgatcgtcgg gcacatgagc ggcgagtggc gcgccttcac gctggacgcc
 300
 gagcgcaccc aactggtgct caccacgac ttctgtgacgc gcgcggccgg cgacgacggc
 360
 ctggtggccg gcaagctcac ccccgacgag gcgcgcgaga tgctggaggc ggtcgtcgag
 420
 cgcaacagcg tcgccgacct gaacgcggtg ctccggcgagg ccgagcgctc ggtgcgcgcg
 480
 gccggcgggc tcgggacggt gaccgcgtga
 510

<210> 26
 <211> 771
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 26
 atgtccggac gtaagacgtt tttggatttg tctttcgcaa ccagggatac accatcagaa
 60
 gctacacctg tgggtggtcga ccttcttgat catgttactg gagcaaccgt tctaggtctc
 120

tcaccagagg actttcctga tggcatggcc atttctaacg agaccgttac attaacgact
180

cacactggga ctcatatgga cgcacctcta cactatgggc ctctatcagg cgggggtgcct
240

gctaagagta ttgatcaagt tcctcttgaa tgggtgttatg gccctggagt gagacttgat
300

gtcaggcatg tgcccgcagg ggatggcata actgttgatc atttgaacgc cgctttggat
360

gctgcagagc acgatcttgc accaggcgat attgtgatgc tttggactgg ggcagatgct
420

ctgtggggta caagagaata tttatcaact tttcctgggc tcaactggtaa ggggactcaa
480

tttcttgtgg aagcaggagt taaagtgatt ggaatagatg cttgggggttt ggataggcca
540

atggcagcca tgatcgaaga atatcggagg actgggtgata aggggtgcact ctggccggca
600

catgtatacg gcagaacgcg agaatatcta caacttgaaa agcttaataa cctcggagct
660

ctcccgggcg caaccggtta cgatatatca tgttttccag tggccgttgc cggaactggc
720

gcaggctgga ctctgttagt cgcagttttt gaacaggagg aggaagattg a
771

<210> 27

<211> 771

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 27

gtgagcggcc gcaagacgtt cctcgacctg agcttcgcga cccgggacac cccagcgag
60

gccaccccg tctctctga cctgctcgac cacgtgaccg gcgcgacggt gctcgggctg
120

agccccgagg acttccccga cggcatggcc atctcgaacg agacggtcac cctcaccacg
180

cacaccggca cgcacatgga cgcgccgctg cactacggcc cgctcagcgg cggcgtgccg
240

gccaagtcga tcgaccaggt cccgctggag tgggtgctacg gcccgggcgt gcgcctggac
300

gtgcggcacg tgccggccgg cgacggcatc accgtcgacc acctcaacgc cgcgctcgac
360

gccgccgagc acgacctggc gcccgggcgc atcgtcatgt tgtggaccgg cgcggacgcc
420

ctgtggggca cccgcgagta cctgtccacc ttccccggac tgaccggcaa gggaacgcag
480

ttcctggctg aggcggggcgt gaaggatgac ggcatcgacg cgtggggcct ggaccggccg
540

atggccgcga tgatcgagga gtaccggcgt accggcgaca agggcgccct gtggccggcg
600

cacgtctacg ggcgcacccg cgaatacctc cagttggaga agctcaacaa cctggggcgcg
660

ctgccccgcg ccaccggata cgacatcagt tgcctccccg tcgccgtcgc gggcaccggt
720

gcgggctgga cccgggtggg cgccgtgttc gagcaggaag aggaggacta g
771

<210> 28

<211> 1578

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 28

atggaattca gattgttgat attggccttg ttctccgtat tgatgtctac ctctaattgt
60

gccgaaatct tggctttatt ccctattcat ggtatatcta actacaacgt agctgaagca
120

ttgttgaaga ctttggctaa cagaggtcac aacgttaccg ttgtaacttc atttccacaa
180

aagaaaccag ttcttaattt gtacgaaatt gatgtatcag gtgcaaaggg tttagccaca
240

aactccatcc atttcgaaag attgcaaacc atcatccaag atgtcaagag taacttcaag
300

aacatgggta gattgtctag aacatactgt gaaatcatgt tctcagaccc aagagttttg
360

aacatcagag ataaaaagtt tgacttggtt ataaacgccg tattcgggtc agattgcgac
420

gctgggtttg catggaaaag tcaagctcct ttaatatcta tcttgaatgc cagacataca
480

ccatgggctt tgcacagaat gggtaatcct tccaaccag catatatgcc tgtaatccat
540

agtagattcc cagtcaagat gaatttcttt caaagaatga taaacaccgg ttggcactta
600

tactttttgt acatgtactt ctactacggg aatgggtgaag atgctaaca aatggcaaga
660

aagtttttcg gtaatgatat gcctgacata aacgaaatgg tttttaacac ctccctgttg
720

ttcgtaaaca ctcatctcag tgtcgatatg ccataccctt tagtcccaaa ctgtatcgaa
780

atcgggtgga tccatgttaa ggaaccacaa cctttgccat tggaaatcca aaagtttatg
840

gatgaagcag aacatgggtg aatctttttc accttgggta gtatgggtcag aacttctaca
900

ttccctaata aaactattca agccttttaa gaagccttcg ctgaattacc acaaagagtt
960

ttgtggaagt tcgaaaacga aaacgaagat atgccttcca acgttttgat cagaaagtgg
1020

ttcccacaaa acgacatctt cggtcataag aacatcaagg ctttcatttc acacggtggt
1080

aattccggtg ccttggaagc tgtccatttc ggtgttccta tcataggtat cccattgttt
1140

tatgatcaat acagaaacat cttgtctttc gttaaagaag gtgtagctgt cttgttggtat
1200

gtaaacgact taactaagga taacatcttg tottcagtta gaacagtcgt taacgacaag
1260

tcataactccg aaagaatgaa ggcattgtct caattgttta gagatagacc tatgtcacca
1320

ttagacacag ctgtttattg gaccgaatac gtaattagac atagagggtgc acatcactta
1380

aaaactgcag gtgccttttt gcactgggtat caatacttgt tggtggatgt catcacattt
1440

ttggttggtta cattctgtgc attctgcttc atcggttaagt acatctgcaa ggccttaatc
1500

catcactact ggtccagttc taaatctgaa aagttgaaaa agaattctgc ttggtotcat
1560

ccacaattcg aaaaataa
1578

<210> 29
<211> 525
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 29

Met Glu Phe Arg Leu Leu Ile Leu Ala Leu Phe Ser Val Leu Met Ser
1 5 10 15

Thr Ser Asn Gly Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile
20 25 30

Ser Asn Tyr Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg
35 40 45

Gly His Asn Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val
50 55 60

Pro Asn Leu Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr
65 70 75 80

Asn Ser Ile His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys
85 90 95

Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
100 105 110

Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
115 120 125

Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
 130 135 140

Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
 145 150 155 160

Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
 165 170 175

Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
 180 185 190

Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
 195 200 205

Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
 210 215 220

Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
 225 230 235 240

Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
 245 250 255

Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu
 260 265 270

Pro Leu Glu Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile
 275 280 285

Phe Phe Thr Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln
 290 295 300

Thr Ile Gln Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val
 305 310 315 320

Leu Trp Lys Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu
 325 330 335

Ile Arg Lys Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile

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

<210>	30
<211>	1437
<212>	ADN
<213>	Trình tự nhân tạo
<220>	
<223>	ADN tổng hợp

<400> 30

atggaattca gattgttgat attggccttg ttctccgtat tgatgtctac ctctaattggt
60

gccgaaatct tggctttatt ccctattcat ggtatatcta actacaacgt agctgaagca
120

ttgttgaaga ctttggctaa cagaggtcac aacgttaccg ttgtaacttc atttccacaa
180

aagaaaccag ttcctaattt gtacgaaatt gatgtatcag gtgcaaaggg tttagccaca
240

aactccatcc atttcgaaag attgcaaacc atcatccaag atgtcaagag taacttcaag
300

aacatgggta gattgtctag aacatactgt gaaatcatgt tctcagaccc aagagttttg
360

aacatcagag ataaaaagtt tgacttggtt ataaacgccg tattcggttc agattgcgac
420

gctgggtttg catggaaaag tcaagctcct ttaatatacta tcttgaatgc cagacataca
480

ccatgggctt tgcacagaat gggtaatcct tccaaccag catatatgcc tgtaatccat
540

agtagattcc cagtcaagat gaatttcttt caaagaatga taaacaccgg ttggcactta
600

tactttttgt acatgtactt ctactacggg aatgggtgaag atgctaacaa aatggcaaga
660

aagtttttcg gtaatgatat gcoctgacata aacgaaatgg tttttaacac ctccctgttg
720

ttcgtaaaca ctcatctcag tgcgatatg ccataaccct tagtcccaaa ctgtatcgaa
780

atcgggtgga tccatgttaa ggaaccacaa cctttgccat tggaaatcca aaagtttatg
840

gatgaagcag aacatgggtg aatctttttc accttgggta gtatggtcag aacttctaca
900

ttccctaata aaactattca agcctttaaa gaagccttcg ctgaattacc acaaagagtt
960

ttgtggaagt tcgaaaacga aaacgaagat atgccttcca acgttttgat cagaaagtgg
1020

ttcccacaaa acgacatctt cggtcataag aacatcaagg ctttcatttc acacgggtgg
1080

aattccgggtg ccttggaagc tgtccatttc ggtgttccta tcataggtat cccattgttt
1140

tatgatcaat acagaaacat cttgtctttc gttaaagaag gtgtagctgt cttgttggtat
1200

gtaaacgact taactaagga taacatcttg tcttcagtta gaacagtcgt taacgacaag
1260

tcatactccg aaagaatgaa ggcattgtct caattgttta gagatagacc tatgtcacca
1320

ttagacacag ctgtttattg gaccgaatac gtaattagac atagaggtgc acatcactta
1380

aaaactgcag gtgccttttt gcactctgct tgggtctcatc cacaattcga aaaataa
1437

<210> 31

<211> 478

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 31

Met	Glu	Phe	Arg	Leu	Leu	Ile	Leu	Ala	Leu	Phe	Ser	Val	Leu	Met	Ser
1				5					10					15	

Thr	Ser	Asn	Gly	Ala	Glu	Ile	Leu	Ala	Leu	Phe	Pro	Ile	His	Gly	Ile
			20					25					30		

Ser	Asn	Tyr	Asn	Val	Ala	Glu	Ala	Leu	Leu	Lys	Thr	Leu	Ala	Asn	Arg
			35				40					45			

Gly	His	Asn	Val	Thr	Val	Val	Thr	Ser	Phe	Pro	Gln	Lys	Lys	Pro	Val
	50					55				60					

Pro	Asn	Leu	Tyr	Glu	Ile	Asp	Val	Ser	Gly	Ala	Lys	Gly	Leu	Ala	Thr
65					70					75					80

Asn	Ser	Ile	His	Phe	Glu	Arg	Leu	Gln	Thr	Ile	Ile	Gln	Asp	Val	Lys
				85					90					95	

Ser Asn Phe Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile
 100 105 110

Met Phe Ser Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp
 115 120 125

Leu Val Ile Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala
 130 135 140

Trp Lys Ser Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr
 145 150 155 160

Pro Trp Ala Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met
 165 170 175

Pro Val Ile His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg
 180 185 190

Met Ile Asn Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr
 195 200 205

Tyr Gly Asn Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly
 210 215 220

Asn Asp Met Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu
 225 230 235 240

Phe Val Asn Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro
 245 250 255

Asn Cys Ile Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu
 260 265 270

Pro Leu Glu Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile
 275 280 285

Phe Phe Thr Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln
 290 295 300

Thr Ile Gln Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val
305 310 315 320

Leu Trp Lys Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu
325 330 335

Ile Arg Lys Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile
340 345 350

Lys Ala Phe Ile Ser His Gly Gly Asn Ser Gly Ala Leu Glu Ala Val
355 360 365

His Phe Gly Val Pro Ile Ile Gly Ile Pro Leu Phe Tyr Asp Gln Tyr
370 375 380

Arg Asn Ile Leu Ser Phe Val Lys Glu Gly Val Ala Val Leu Leu Asp
385 390 395 400

Val Asn Asp Leu Thr Lys Asp Asn Ile Leu Ser Ser Val Arg Thr Val
405 410 415

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

Phe Arg Asp Arg Pro Met Ser Pro Leu Asp Thr Ala Val Tyr Trp Thr
435 440 445

Glu Tyr Val Ile Arg His Arg Gly Ala His His Leu Lys Thr Ala Gly
450 455 460

Ala Phe Leu His Ser Ala Trp Ser His Pro Gln Phe Glu Lys
465 470 475

<210> 32
<211> 1380
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 32

atggccgaaa tcttggtttt attccctatt catggtatat ctaactacaa cgtagctgaa
 60
 gcattgttga agactttggc taacagaggt cacaacgtta ccgttgtaac ttcatttcca
 120
 caaaagaaac cagttcctaa tttgtacgaa attgatgtat caggtgcaaa gggtttagcc
 180
 acaaactcca tccatttcga aagattgcaa accatcatcc aagatgtcaa gagtaacttc
 240
 aagaacatgg ttagattgtc tagaacatac tgtgaaatca tgttctcaga cccaagagtt
 300
 ttgaacatca gagataaaaa gtttgacttg gttataaacg ccgtattcgg ttcagattgc
 360
 gacgctggtt ttgcatggaa aagtcaagct cctttaatat ctatcttgaa tgccagacat
 420
 acaccatggg ctttgcacag aatgggtaat ccttccaacc cagcatatat gcctgtaatc
 480
 catagtagat tcccagtcaa gatgaatttc tttcaaagaa tgataaacac cggttggcac
 540
 ttatactttt tgtacatgta cttctactac ggtaatggtg aagatgctaa caaaatggca
 600
 agaaagtfff tcggtaatga tatgcctgac ataaacgaaa tggtttttaa cacctccttg
 660
 ttgttcgtaa aactcattt cagtgtogat atgccatacc ctttagtccc aaactgtatc
 720
 gaaatcgggtg gtatccatgt taaggaacca caacctttgc cattggaaat ccaaaagttt
 780
 atggatgaag cagaacatgg tgtaatcttt ttcaccttgg gtagtatggt cagaacttct
 840
 acattcccta atcaaactat tcaagccttt aaagaagcct tcgctgaatt accacaaaga
 900
 gttttgtgga agttcgaaaa cgaaaacgaa gatatgcctt ccaacgtttt gatcagaaag
 960
 tggttcccac aaaacgacat cttcgggtcat aagaacatca aggctttcat ttcacacggt
 1020
 ggtaattccg gtgccttgga agctgtccat ttcggtgttc ctatcatagg tatcccattg
 1080

ttttatgatc aatacagaaa catcttgtct ttcgttaaag aaggtgtagc tgtcttggtg
1140

gatgtaaacg acttaactaa ggataacatc ttgtcttcag ttagaacagt cgtaaacgac
1200

aagtcatact ccgaaagaat gaaggcattg totcaattgt ttagagatag acctatgtca
1260

ccattagaca cagctgttta ttggaccgaa tacgtaatta gacatagagg tgcacatcac
1320

ttaaaaactg caggtgcctt ttgcaactct gcttggtctc atccacaatt cgaaaaataa
1380

<210> 33

<211> 459

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 33

Met Ala Glu Ile Leu Ala Leu Phe Pro Ile His Gly Ile Ser Asn Tyr
1 5 10 15

Asn Val Ala Glu Ala Leu Leu Lys Thr Leu Ala Asn Arg Gly His Asn
20 25 30

Val Thr Val Val Thr Ser Phe Pro Gln Lys Lys Pro Val Pro Asn Leu
35 40 45

Tyr Glu Ile Asp Val Ser Gly Ala Lys Gly Leu Ala Thr Asn Ser Ile
50 55 60

His Phe Glu Arg Leu Gln Thr Ile Ile Gln Asp Val Lys Ser Asn Phe
65 70 75 80

Lys Asn Met Val Arg Leu Ser Arg Thr Tyr Cys Glu Ile Met Phe Ser
85 90 95

Asp Pro Arg Val Leu Asn Ile Arg Asp Lys Lys Phe Asp Leu Val Ile
100 105 110

Asn Ala Val Phe Gly Ser Asp Cys Asp Ala Gly Phe Ala Trp Lys Ser
 115 120 125

Gln Ala Pro Leu Ile Ser Ile Leu Asn Ala Arg His Thr Pro Trp Ala
 130 135 140

Leu His Arg Met Gly Asn Pro Ser Asn Pro Ala Tyr Met Pro Val Ile
 145 150 155 160

His Ser Arg Phe Pro Val Lys Met Asn Phe Phe Gln Arg Met Ile Asn
 165 170 175

Thr Gly Trp His Leu Tyr Phe Leu Tyr Met Tyr Phe Tyr Tyr Gly Asn
 180 185 190

Gly Glu Asp Ala Asn Lys Met Ala Arg Lys Phe Phe Gly Asn Asp Met
 195 200 205

Pro Asp Ile Asn Glu Met Val Phe Asn Thr Ser Leu Leu Phe Val Asn
 210 215 220

Thr His Phe Ser Val Asp Met Pro Tyr Pro Leu Val Pro Asn Cys Ile
 225 230 235 240

Glu Ile Gly Gly Ile His Val Lys Glu Pro Gln Pro Leu Pro Leu Glu
 245 250 255

Ile Gln Lys Phe Met Asp Glu Ala Glu His Gly Val Ile Phe Phe Thr
 260 265 270

Leu Gly Ser Met Val Arg Thr Ser Thr Phe Pro Asn Gln Thr Ile Gln
 275 280 285

Ala Phe Lys Glu Ala Phe Ala Glu Leu Pro Gln Arg Val Leu Trp Lys
 290 295 300

Phe Glu Asn Glu Asn Glu Asp Met Pro Ser Asn Val Leu Ile Arg Lys
 305 310 315 320

Trp Phe Pro Gln Asn Asp Ile Phe Gly His Lys Asn Ile Lys Ala Phe
 325 330 335

Ile Ser His Gly Gly Asn Ser Gly Ala Leu Glu Ala Val His Phe Gly
 340 345 350

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

Leu Ser Phe Val Lys Glu Gly Val Ala Val Leu Leu Asp Val Asn Asp
 370 375 380

Leu Thr Lys Asp Asn Ile Leu Ser Ser Val Arg Thr Val Val Asn Asp
 385 390 395 400

Lys Ser Tyr Ser Glu Arg Met Lys Ala Leu Ser Gln Leu Phe Arg Asp
 405 410 415

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

Ile Arg His Arg Gly Ala His His Leu Lys Thr Ala Gly Ala Phe Leu
 435 440 445

His Ser Ala Trp Ser His Pro Gln Phe Glu Lys
 450 455

<210> 34
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 34
 ggaagaaggt cgcatacca
 19

<210> 35
 <211> 42
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 35
 gatccccggg aattgccatg gatcaggaga aggagagtca ag
 42

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

<220>
 <223> ADN tổng hợp

<400> 36
 aattccagct gaccaccatg ggcgaaaagg caaaggagc
 39

<210> 37
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 37
 gctagaaaag gcaagggagg
 20

<210> 38
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 38
 catggcaatt cccggggatc gccggcaatt ctttttaggt agc
 43

<210> 39
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 39

ccagaagcag tacacggc
18

<210> 40
<211> 21
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 40
gttggtctgct tgcgcttctt c
21

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

<220>
<223> ADN tổng hợp

<400> 41
catggtggtc agctggaatt cctccgccat ttcttattcc c
41

<210> 42
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 42
gtgggttgaa ccgcttactc ag
22

<210> 43
<211> 42
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 43
gatccccggg aattgccatg cccggaggaa tcaaaatgac gc
42

<210> 44
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 44
 aattccagct gaccacatg gtttgggatt cttaggtgag ctc
 43

<210> 45
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 45
 cctccctggc gtatacaca ac
 22

<210> 46
 <211> 17
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 46
 gctcgagctt gccagcc
 17

<210> 47
 <211> 39
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 47
 gatccccggg aattgccatg gctggtgttg ggacacacg
 39

<210> 48
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 48
 aattccagct gaccaccatg gcttggaat cagtatagct ttctg
 45

<210> 49
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 49
 gcttgtgggc tgtctgaatc g
 21

<210> 50
 <211> 27
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 50
 ggggtttaatg aggagcagag gatgcgg
 27

<210> 51
 <211> 27
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 51
 ggacttaatg tagtgggtgt gctgggtg
 27

<210> 52
 <211> 30

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 52
 ggcattaatc gcgtggaatt tggaagagag
 30

<210> 53
 <211> 28
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 53
 ggtcttaatg tgctcgggga cgtgaaag
 28

<210> 54
 <211> 29
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 54
 gggtttaatg gtcgtctgtc aaggagttg
 29

<210> 55
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 55
 ggacttaatg cagtgtgtga tatgggtctt g
 31

<210> 56
 <211> 33
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 56

ggcattaatg agtttgtgag atgttcagga tgg
33

<210> 57

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 57

ggtcttaatg aggtgaagga cacagcg
27

<210> 58

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 58

gggtttaatc gcagagacta ggacacaagt g
31

<210> 59

<211> 28

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 59

ggacttaatg cggcgatctg tggtagag
28

<210> 60

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 60
ggcattaatg ccagcatatt caaacccagt c
31

<210> 61
<211> 29
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 61
ggctttaatc acacaaccaa cctccgata
29

<210> 62
<211> 31
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 62
agagcgatat gggttcotta gacaacggtt c
31

<210> 63
<211> 33
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 63
tctgcgattc acaaaggaac acttctcaaa acc
33

<210> 64
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 64
agagcgatat gagttcactc tccaacgctt cc
32

<210> 65
<211> 29
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 65
tctgcgattc acatgagagg caggctgtg
29

<210> 66
<211> 28
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 66
agagcgatgg gttcccttga caatggtt
28

<210> 67
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 67
tctgcgattt agagaggcac acttcggaga ac
32

<210> 68
<211> 33
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 68

agagcgatat gagtagttta tcaaagcca gtc
33

<210> 69
<211> 27
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 69
tctgcgattc acatcaatgg caaggaa
27

<210> 70
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 70
agatatacca tgcgtcatgt tgaacatacc gt
32

<210> 71
<211> 34
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 71
atggctgott tatgcggtaa ctgtaccaac acca
34

<210> 72
<211> 29
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> ADN tổng hợp

<400> 72
atatacatat gagcggctgt aaaaccttt
29

<210> 73
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 73
 atatccaatt ttaatcctct tcttcttggt c
 31

<210> 74
 <211> 25
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 74
 agagcgatga gacacgttga acaca
 25

<210> 75
 <211> 27
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 75
 tctgcgattt atgcagttac ggtacca
 27

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

<220>
 <223> ADN tổng hợp

<400> 76
 agagcgatgt ccggtagaaa gacctt
 26

<210> 77
 <211> 30
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 77
 tctgcgattt aatcttcttc ttcttggtcg
 30

<210> 78
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 78
 agagcgatat ggagtttcgc ttgcttatcc t
 31

<210> 79
 <211> 37
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 79
 tctgcgattt aattcttctt caacttttcc gacttag
 37

<210> 80
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 80
 ggcttaatat gagttcactc tccaacgctt cccatc
 36

<210> 81
 <211> 41

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 81
 ggtttaattt acatgagagg caggctgtgg agaaggatag t
 41

<210> 82
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 82
 ggcttaatat gaggcacgtc gagcat
 26

<210> 83
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 83
 ggtttaattt atgccgtgac agttccgaca c
 31

<210> 84
 <211> 26
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ADN tổng hợp

<400> 84
 ggcttaatat gtccggacgt aagacg
 26

<210> 85
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 85

ggttttaattt aatcttcctc ctctgttca a
31

<210> 86

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 86

ggcttaatat ggaattcaga ttgttgatat tggcct
36

<210> 87

<211> 41

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 87

ggttttaattt atttttcgaa ttgtggatga gaccaagcag a
41

<210> 88

<211> 29

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 88

ggggacaagt ttgtacaaaa aagcaggct
29

<210> 89

<211> 54

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ADN tổng hợp

<400> 89

ttatatttttcg aattgtggat gagaccaagc agagtgc aaa aaggcacctg cagt
54