



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

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

(11)



1-0026935

(51)⁷ C12N 15/09; C12N 9/42; C12N 1/19

(13) B

(21) 1-2015-02795

(22) 31/07/2015

(30) 2014-158651 04/08/2014 JP

(45) 25/01/2021 394

(43) 25/02/2016 335A

(73) HONDA MOTOR CO., LTD. (JP)

1-1, Minami-Aoyama 2-chome, Minato-ku, Tokyo, 107-8556, Japan

(72) Migiwa SUDA (JP); Jiro OKUMA (JP); Yoshitsugu HIROSE (JP); Asuka YAMAGUCHI (JP); Yasuhiro KONDO (JP); Tomohiko KATO (JP); Daisuke SHIBATA (JP).

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) ENDOGLUCANAZA SIÊU CHỊU NHIỆT, POLYNUCLEOTIT MÃ HÓA ENZYM NÀY, VECTƠ BIỂU HIỆN ENZYM NÀY, THỂ BIẾN NẠP CHỨA VECTƠ NÀY, PHƯƠNG PHÁP SẢN XUẤT ENZYM NÀY, HỖN HỢP GLYCOSIT HYĐROLAZA CHỨA ENZYM NÀY, VÀ PHƯƠNG PHÁP SẢN XUẤT SẢN PHẨM PHÂN GIẢI LIGNOXENLULOZA

(57) Sáng chế đề cập đến endoglucanaza siêu chịu nhiệt, polynucleotit mã hóa enzym này, vectơ biểu hiện enzym này, thể biến nạp chứa vectơ này, phương pháp sản xuất enzym này, hỗn hợp glycosit hydrolaza chứa enzym này, và phương pháp sản xuất sản phẩm phân giải lignoxenluloza bằng enzym này.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến endoglucanaza siêu chịu nhiệt, polynucleotit mã hóa enzym này, vectơ biểu hiện enzym này, thể biến nạp chứa vectơ này, và phương pháp sản xuất sản phẩm phân giải lignoxenluloza bằng enzym này.

Tình trạng kỹ thuật của sáng chế

Trong những năm gần đây, việc phát triển các nguồn năng lượng thay thế dầu mỏ ngày càng trở nên quan trọng, do các vấn đề về môi trường như trái đất nóng lên và ô nhiễm bầu khí quyển, cũng như quan ngại liên quan đến việc cung cấp năng lượng cho các phương tiện vận tải. Sinh khối thực vật là nguồn năng lượng tái tạo phong phú nhất trên trái đất, được mong đợi là nguồn năng lượng thay thế cho dầu mỏ. Lignoxenluloza là thành phần chính của sinh khối thực vật và bao gồm các polysacarit như xenluloza, hemixenluloza (bao gồm xylan, arabinan và mannan), lignin và các pectin khác. Các polysacarit này được thủy phân thành các monosacarit như glucoza và xyloza bởi nhiều glycosit hydrolaza và được sử dụng làm nhiên liệu sinh học hoặc nguyên liệu hóa học.

Lignoxenluloza là hợp chất có cấu trúc rất phức tạp, nên rất khó có thể phân giải hoặc thủy phân được bằng một enzym duy nhất. Do đó, trong số các polysacarit này, quá trình thủy phân xenluloza thường cần ba loại enzym glycosit hydrolaza bao gồm endoglucanaza (endo-1,4- β -D-glucanaza, EC 3.2.1.4), exo-xenlobiohydrolaza (1,4- β -xenlobiosidaza hoặc xenlobiohydrolaza, EC 3.2.1.91, EC 3.2.1.176), và β -glucosidaza (EC 3.2.1.21). Mặt khác, hemixenluloza chứa xylan, arabinan, manan và các chất tương tự, và thành phần của nó có thể thay đổi phụ thuộc vào loài thực vật. Ví dụ, xylan là thành phần chính có trong thực vật lá rộng, thực vật thân thảo và các thực vật tương tự. Để thủy phân xylan, thì cần có xylanaza (endo-1,4- β -xylanaza, EC 3.2.1.8) và β -xylosidaza (EC 3.2.1.37).

Trong quy trình thông thường biến đổi lignoxenluloza thành etanol, cơ chất ban đầu có nồng độ chất rắn cao nằm trong khoảng từ 30% đến 60% đã được dùng để đạt được hiệu quả năng lượng cao hơn và sử dụng ít nước hơn. Tuy nhiên, khi nồng độ chất rắn trong cơ chất cao, thì phản ứng enzym thủy phân lignoxenluloza hầu như không thể xảy ra do độ nhớt cao của dung dịch sinh khối đã thủy phân. Do đó, bằng cách thực hiện phản ứng thủy phân ở nhiệt độ 80°C hoặc cao hơn và sử dụng enzym chịu nhiệt, thì ngoài việc tăng tốc độ phản ứng thủy phân, độ nhớt của dung dịch sinh khối đã thủy phân cũng được giảm xuống, nhờ vậy có thể rút ngắn được thời gian thủy phân và giảm lượng enzym cần sử dụng. Do đó, có nhu cầu phát triển các enzym có độ ổn định nhiệt cao.

Nhiều glycosit hydrolaza chịu nhiệt đã được thu nhận bằng cách phân lập và xác định các vi sinh vật ưa nhiệt sống trong môi trường có nhiệt độ cao, tách dòng các gen từ các vi sinh vật đã được phân lập và nuôi cấy này và giải trình tự ADN của chúng, sau đó biểu hiện chúng bằng *E. coli*, nấm sợi và các sinh vật tương tự. Để thủy phân lignoxenluloza ở nhiệt độ cao, như sợi xenluloza, vỏ cây và giấy, nhiều enzym chịu nhiệt có thể được sử dụng, đặc biệt là cho đến nay đã phân lập được từ vi khuẩn ưa nhiệt, nấm sợi, vi khuẩn cổ và sinh vật tương tự, các endoglucanaza cần để thủy phân xenluloza (ví dụ tham khảo Tài liệu Sáng chế 1 đến 4, và tài liệu phi Sáng chế 1 và 2). Ngoài ra, các nỗ lực cũng đã được thực hiện để cải thiện tiếp hoạt tính riêng và tính bền nhiệt, ví dụ bằng cách sử dụng các thể đột biến của sinh vật chủ hoặc cải biến một phần trình tự axit amin của các enzym này (ví dụ tham khảo Tài liệu phi Sáng chế 3 và 4). Tuy nhiên, hầu hết các enzym này có nhiệt độ tối ưu nằm trong khoảng từ 60°C đến 80°C, và cần có độ bền nhiệt cao hơn. Mặt khác, endoglucanaza siêu chịu nhiệt có nhiệt độ tối ưu cao hơn 90°C đã được phân lập (ví dụ tham khảo Tài liệu Sáng chế 5 đến 7, và tài liệu phi Sáng chế 1 và 5). Endoglucanaza có nguồn gốc từ *Sulfolobus solfataricus* có nhiệt độ tối ưu bằng 95°C được bộc lộ trong Tài liệu phi Sáng chế 5, tuy nhiên hoạt tính riêng của nó ở nhiệt độ 90°C không cao và chỉ bằng 5,5U/mg protein. Enzym này cũng có hoạt tính xylanaza, tuy nhiên hoạt tính riêng xylanaza của nó ở nhiệt độ 90°C cũng thấp và chỉ bằng 4,0U/mg protein.

Tài liệu viện dẫn

Tài liệu Sáng chế 1: Đơn Công bố Patent Mỹ 2003/0054539.

Tài liệu Sáng chế 2: Đơn Công bố Patent Mỹ 2005/0070003.

Tài liệu Sáng chế 3: Đơn Công bố Patent Mỹ 2003/0129723.

Tài liệu Sáng chế 4: Đơn Công bố Patent Mỹ 2005/0215450.

Tài liệu Sáng chế 5: Đơn Công bố Patent Nhật Bản 2003-210182.

Tài liệu Sáng chế 6: Đơn Công bố Patent Nhật Bản 2014-27928.

Tài liệu Sáng chế 7: Đơn Công bố Patent Mỹ 2002/0102699.

Tài liệu phi Sáng chế 1: Ando *et al.*, *Applied and Environmental Microbiology*, 2002, vol. 68, p. 430-433.

Tài liệu phi Sáng chế 2: Shi *et al.*, *Bioresource Technology* 2013, vol. 142, p338-344.

Tài liệu phi Sáng chế 3: Zhang *et al.*, *Journal of Biotechnology*, 1997, vol. 57, p. 101-113.

Tài liệu phi Sáng chế 4: Telke *et al.*, *PLoS One*, 2013, vol. 8, p. e65727.

Tài liệu phi Sáng chế 5: Maurelli *et al.*, *Extremophiles*, 2008, vol. 12, p. 689-700.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề cập đến endoglucanaza siêu chịu nhiệt thể mới có hoạt tính thủy phân cơ chất carboxymetyl xenluloza (sau đây gọi tắt là CMC) ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0, polynucleotit mã hóa enzym này, vector biểu hiện enzym này, thể biến nạp chứa vector này, và phương pháp sản xuất sản phẩm phân giải lignoxenluloza bằng enzym này.

Để giải quyết các vấn đề nêu trên, các tác giả sáng chế đã thu nhận thành công endoglucanaza siêu chịu nhiệt có trình tự axit amin mới bằng cách chiết tách ADN trực tiếp từ đất có nhiệt độ cao ở khu vực suối nước nóng và thực hiện giải trình tự bộ gen của hệ vi sinh vật khó nuôi cấy lấy từ môi trường này và hoàn thành sáng chế này.

Endoglucanaza siêu chịu nhiệt, polynucleotit, vector biểu hiện, thể biến nạp, phương pháp sản xuất endoglucanaza siêu chịu nhiệt, hỗn hợp glycosit hydrolaza, và phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế được mô tả trong các phương án từ 1 đến 10 dưới đây.

1. Endoglucanaza siêu chịu nhiệt chứa vùng xúc tác endoglucanaza gồm:

(A) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11;

(B) polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung ít nhất một axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0; hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 70% so với trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

2. Endoglucanaza siêu chịu nhiệt theo phương án 1, trong đó enzym này cũng có hoạt tính xylanaza ở nhiệt độ bằng 110°C và độ pH bằng 7,0.

3. Polynucleotit chứa miền mã hóa vùng xúc tác endoglucanaza gồm:

(a) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11;

(b) trình tự nucleotit mã hóa polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung ít nhất một axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0;

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 70% so với trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0;

(d) trình tự nucleotit có độ tương đồng ít nhất bằng 70% so với trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12, và mã hóa polypeptit có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0; hoặc

(e) trình tự nucleotit của polynucleotit mà lai hóa với polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

4. Polynucleotit theo phương án 3, trong đó polypeptit này cũng có hoạt tính xylanaza ở nhiệt độ bằng 110°C và độ pH bằng 7,0.

5. Vector biểu hiện chứa polynucleotit theo phương án 3 hoặc 4, và có khả năng biểu hiện polypeptit có hoạt tính endoglucanaza ở tế bào vật chủ.

6. Thể biến nạp chứa vector biểu hiện theo phương án 5.

7. Thể biến nạp theo phương án 6, trong đó thể biến nạp này là vi sinh vật nhân thật.

8. Phương pháp sản xuất endoglucanaza siêu chịu nhiệt bao gồm bước sản sinh endoglucanaza siêu chịu nhiệt trong thể biến nạp theo phương án 6 hoặc 7.

9. Hỗn hợp glycosit hydrolaza chứa endoglucanaza siêu chịu nhiệt theo phương án 1 hoặc 2, endoglucanaza siêu chịu nhiệt mã hóa bởi polynucleotit theo phương án 3 hoặc 4, hoặc endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp theo phương án 8, và ít nhất một hoặc nhiều loại glycosit hydrolaza khác.

10. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với endoglucanaza siêu chịu nhiệt theo phương án 1 hoặc 2, endoglucanaza siêu chịu nhiệt mã hóa bởi polynucleotit theo phương án 3 hoặc 4, thể biến nạp theo phương án 6 hoặc 7, endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp theo phương án 8, hoặc hỗn hợp glycosit hydrolaza theo phương án 9.

Endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0. Do đó, endoglucanaza siêu chịu nhiệt này thích hợp để thủy phân nguyên liệu lignoxenluloza chứa xenluloza ở nhiệt độ cao.

Ngoài ra, polynucleotit, vectơ biểu hiện chứa polynucleotit này và thể biến nạp chứa vectơ này theo sáng chế cũng thích hợp để sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin (SEQ ID NO.1) của polypeptit (sau đây gọi tắt là AR15G-90) mã hóa bởi khung đọc mã mở AR15G-90 và trình tự axit amin của glycosit hydrolaza GH12 (SEQ ID NO.8) có nguồn gốc từ *Ignisphaera aggregans* DSM17230.

Fig.2 là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein AR15G-90-3 được tạo ra bằng cách biểu hiện gen AR15G-90-3 ở *E. coli* trong ví dụ 1.

Fig.3 là biểu đồ thể hiện kết quả đo hoạt tính thủy phân mỗi cơ chất của protein AR15G-90-3 được biểu hiện ở *E. coli* trong ví dụ 1.

Fig.4 là đồ thị thể hiện kết quả đo hoạt tính thủy phân CMC (ở độ pH = 4,0) của protein AR15G-90-3 được biểu hiện ở *E. coli* khi không có mặt ion canxi (“AR15G-90-3” trên Fig.4) và khi có mặt ion canxi (“AR15G-90-3 + 0,5mM Ca²⁺” trên Fig.4) và hoạt tính thủy phân CMC (ở độ pH = 4,0) của endoglucanaza TrEG (*Trichoderma reesei* EG2) và endoglucanaza TmEG (*Thermotoga maritima* EG) ở các nhiệt độ tương ứng trong ví dụ 1.

Fig.5 là đồ thị thể hiện kết quả đo hoạt tính thủy phân CMC (ở nhiệt độ 70°C) của protein AR15G-90-3 được biểu hiện ở *E. coli* ở các độ pH tương ứng trong ví dụ 1.

Fig.6 là đồ thị thể hiện kết quả đo hoạt tính xylanaza (ở độ pH = 7,0) của protein AR15G-90-3 được biểu hiện ở *E. coli* ở các nhiệt độ tương ứng trong ví dụ 1.

Fig.7 là đồ thị thể hiện kết quả đo hoạt tính xylanaza (ở nhiệt độ 70°C) của protein

AR15G-90-3 được biểu hiện ở *E. coli* ở các độ pH tương ứng trong ví dụ 1.

Fig.8 là đồ thị thể hiện kết quả đo hoạt tính thủy phân CMC của protein AR15G-90-3 được biểu hiện ở *E. coli* trong ví dụ 1 (hoạt tính tương đối (%), với giả định rằng hoạt tính của lô không xử lý (không ủ sơ bộ) bằng 100%).

Fig.9 là đồ thị thể hiện kết quả đo hoạt tính xylanaza của protein AR15G-90-3 được biểu hiện ở *E. coli* trong ví dụ 1 (hoạt tính tương đối (%), với giả định rằng hoạt tính của lô không xử lý (không ủ sơ bộ) bằng 100%).

Fig.10 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin (SEQ ID NO.1) của polypeptit (AR15G-90) mã hóa bởi khung đọc mã mở AR15G-90, trình tự axit amin (SEQ ID NO.9) của polypeptit (sau đây gọi tắt là AR15G-90B) mã hóa bởi khung đọc mã mở AR15G-90B và trình tự axit amin của glycosit hydrolaza GH12 (SEQ ID NO.8) có nguồn gốc từ *Ignisphaera aggregans* DSM17230.

Fig.11 là đồ thị thể hiện kết quả đo hoạt tính thủy phân CMC của protein AR15G-90-B-15 được biểu hiện ở *E. coli* ở các nhiệt độ tương ứng trong ví dụ 2 (hoạt tính tương đối (%)) giả định rằng hoạt tính ở nhiệt độ 90°C bằng 100%).

Mô tả chi tiết sáng chế

Endoglucanaza siêu chịu nhiệt theo sáng chế

Nhiều vi sinh vật, bao gồm nấm sợi, vi khuẩn, và vi khuẩn cổ rất khó được nuôi cấy và khoảng 99% vi sinh vật sống trong môi trường như đất là chưa được biết đến. Đặc biệt, rất khó nuôi cấy được vi sinh vật sống trong môi trường có nhiệt độ cao, và chỉ 0,1% hoặc thấp hơn vi sinh vật sống trong đất đã được phân lập và nuôi cấy bằng các kỹ thuật đã biết. Việc khó nuôi cấy vi sinh vật sống trong đất có nhiệt độ cao là một yếu tố cản trở sự phát triển các enzym chịu nhiệt.

Gần đây, do sự phát triển của máy giải trình tự thế hệ tiếp theo cho phép giải trình tự được hàng tỷ cặp bazơ, nên có thể giải trình tự bộ gen của vi sinh vật chứa trong đất và các môi trường tương tự. Bằng cách sử dụng kỹ thuật phân tích này, phương pháp phân tích hệ gen môi trường đã được đề cập đến, trong đó ADN bộ gen của quần thể vi sinh vật được tổng hợp từ mẫu môi trường như đất, bộ gen của quần thể vi sinh vật có thành

phần không đồng nhất được giải trình tự trực tiếp và toàn bộ, và dữ liệu trình tự được lắp ráp bằng máy tính song song, để thiết kế lại trình tự bộ gen của vi sinh vật. Điều này góp phần rút ngắn thời gian giải trình tự bộ gen của vi sinh vật khó nuôi cấy.

Như được thể hiện ví dụ 1, các tác giả sáng chế đã tổng hợp được ADN bộ gen (ADN hệ gen môi trường) của quần thể vi sinh vật sống trong đất có nhiệt độ cao thu được ở khu vực suối nước nóng (ví dụ mẫu nước thu được ở khu vực suối nước nóng có nhiệt độ nằm trong khoảng từ 58°C đến 78°C chứa đất, bùn, thảm vi sinh vật, màng sinh học và các loại tương tự), và thực hiện giải trình tự ADN hệ gen môi trường này. Kết quả là thu được các khung đọc mã mở (Open Reading Frame - ORF) mã hóa các trình tự axit amin tương đồng với các endoglucanaza đã biết (ví dụ các trình tự axit amin có độ tương đồng bằng 20% hoặc cao hơn và trị số Kỳ vọng (tức là trị số E) nhỏ hơn $1e^{-20}$). Trong số các khung đọc mã mở này, các đoạn môi được thiết kế dựa trên thông tin trình tự nucleotit của 106 khung đọc mã mở, trong đó sự có mặt của vùng xúc tác endoglucanaza có thể được xác định, và các gen tiềm năng được tách dòng từ ADN hệ gen môi trường của đất có nhiệt độ cao ở khu vực suối nước nóng bằng phương pháp PCR. Các ADN được tách dòng PCR được biến nạp vào *E. coli*, và các protein mã hóa bởi các trình tự nucleotit này được biểu hiện. Các protein này được sàng lọc chức năng bằng cách đo hoạt tính thủy phân CMC. Cuối cùng, endoglucanaza siêu chịu nhiệt có hoạt tính endoglucanaza (sau đây được gọi tắt là “AR15G-90-3”) thu được từ các khung đọc mã mở này. Trình tự axit amin của AR15G-90-3 và trình tự nucleotit mã hóa trình tự axit amin của AR15G-90-3 lần lượt như nêu trong SEQ ID NO.3 và SEQ ID NO.4.

AR15G-90-3 là một đoạn protein không có đầu tận cùng N, và như được thể hiện trong Ví dụ 2, AR15G-90B-15 là protein có chiều dài đầy đủ của AR15G-90-3. Trình tự axit amin của AR15G-90B-15 và trình tự nucleotit mã hóa trình tự axit amin của AR15G-90B-15 lần lượt như nêu trong SEQ ID NO.11 và SEQ ID NO.12.

Như được thể hiện trong phần 9 của Ví dụ 1, AR15G-90-3 có hoạt tính thủy phân rất mạnh lichenan, và cũng có hoạt tính thủy phân mạnh CMC, Avicel tương nơ liên hợp axit phosphoric (sau đây gọi tắt là PSA)-là xenluloza phi tinh thể, và xylan, và cũng có hoạt tính thủy phân xenluloza tinh thể Avicel, laminarin, và p-nitrophenyl- β -D-xenlobiosit (sau đây gọi tắt là PNPC). Trái lại, AR15G-90-3 hầu như không có hoạt tính

phân giải p-nitrophenyl- β -D-glucopyranosit (sau đây gọi tắt là PNPG) và p-nitrophenyl- β -D-xylopyranosit (sau đây gọi tắt là PNPX).

Nói cách khác, AR15G-90-3 là endoglucanaza siêu chịu nhiệt có độ đặc hiệu cơ chất cao với các hợp chất có liên kết β -1,4 (ví dụ CMC, PSA, Avicel, PNPC, và các hợp chất tương tự) và các hợp chất có liên kết β -1,3 và β -1,4 (ví dụ lichenan, và các hợp chất tương tự), và cũng có hoạt tính xylanaza.

Kết quả tra cứu trình tự axit amin của AR15G-90-3 bằng cơ sở dữ liệu trình tự axit amin đã biết cho thấy trình tự axit amin của AR15G-90-3 có độ tương đồng cao nhất với trình tự axit amin của glycosit hydrolaza GH12 (mã truy cập Ngân hàng Gen: ADM27292.1) (SEQ ID NO.8) có nguồn gốc từ *Ignisphaera aggregans* DSM17230 thuộc ngành *Crenarchaeota*, và bằng 56% đối với chiều dài đầy đủ và 64% đối với vùng xúc tác GH12. Các phân tích độ đặc hiệu cơ chất, và phân tích độ tương đồng của trình tự axit amin này với các protein đã biết cho thấy AR15G-90-3 là endoglucanaza GH12 thế hệ mới.

AR15G-90-3 có hoạt tính thủy phân cơ chất CMC ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0. Cụ thể, như được thể hiện trong phần 10 và 11 của ví dụ 1, AR15G-90-3 có hoạt tính endoglucanaza ở nhiệt độ nằm trong khoảng từ 60°C đến 160°C, và độ pH nằm trong khoảng từ 3,0 đến 8,0, có hoạt tính endoglucanaza mạnh ở nhiệt độ nằm trong khoảng từ 80°C đến 120°C và độ pH nằm trong khoảng từ 3,0 đến 8,0, và có hoạt tính endoglucanaza rất mạnh ở nhiệt độ nằm trong khoảng từ 90°C đến 110°C và độ pH nằm trong khoảng từ 4 đến 6.

Cụ thể hơn, hoạt tính thủy phân CMC của AR15G-90-3 ở độ pH bằng 4,0 tăng khi nhiệt độ tăng nằm trong khoảng từ 60°C đến 90°C, mạnh và ổn định khi nhiệt độ nằm trong khoảng từ 90°C đến 120°C và giảm từ từ ở nhiệt độ 120°C hoặc cao hơn.

Hơn nữa, ngoài hoạt tính endoglucanaza, AR15G-90-3 cũng có hoạt tính xylanaza ở nhiệt độ 110°C và độ pH bằng 7,0. Cụ thể, như được thể hiện trong phần 12 và 13 của Ví dụ 1, AR15G-90-3 có hoạt tính xylanaza ở nhiệt độ nằm trong khoảng từ 40°C đến 140°C và độ pH nằm trong khoảng từ 5,0 đến 10,0, có hoạt tính xylanaza mạnh ở nhiệt độ nằm trong khoảng từ 80°C đến 120°C và độ pH nằm trong khoảng từ 6,0 đến 9,5, và

có hoạt tính xylanaza rất mạnh ở nhiệt độ nằm trong khoảng từ 90°C đến 115°C và độ pH nằm trong khoảng từ 6,5 đến 9,0.

Cụ thể hơn, hoạt tính xylanaza của AR15G-90-3 ở độ pH bằng 7,0 đối với cơ chất xylan tăng khi nhiệt độ tăng nằm trong khoảng từ 40°C đến 110°C, và giảm nhanh khi nhiệt độ trên 110°C.

Thuật ngữ “có hoạt tính” hoặc “biểu hiện hoạt tính” được dùng trong bản mô tả để chỉ hoạt tính phân giải ít nhất một cơ chất và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử được thủy phân của cơ chất hoặc cường độ phản ứng tạo màu khi so với mẫu đối chứng âm.

Do đó, thuật ngữ “có hoạt tính endoglucanaza” được dùng trong bản mô tả để chỉ hoạt tính phân giải ít nhất một cơ chất được chọn từ nhóm bao gồm các hợp chất có liên kết β -1,4 (ví dụ CMC, PSA, Avicel, PNPC, và các hợp chất tương tự) và các hợp chất có liên kết β -1,3 và β -1,4 (ví dụ lichenan, và các hợp chất tương tự) và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử được thủy phân của cơ chất hoặc cường độ phản ứng tạo màu khi so với mẫu đối chứng âm.

Thuật ngữ “có hoạt tính xylanaza” được dùng trong bản mô tả để chỉ hoạt tính phân giải ít nhất một cơ chất xylan và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử được thủy phân của xylan hoặc cường độ phản ứng tạo màu khi so với mẫu đối chứng âm.

Nhìn chung, ở protein có một số hoạt tính sinh học, một hoặc hai hoặc nhiều axit amin có thể được loại bỏ, thay thế, hoặc bổ sung, mà không làm giảm hoạt tính sinh học của protein này. Nói cách khác, trong AR15G-90-3, một hoặc hai hoặc nhiều axit amin cũng có thể được loại bỏ, thay thế, hoặc bổ sung mà không làm giảm hoạt tính thủy phân glycosit của enzym này.

Endoglucanaza theo sáng chế là glycosit hydrolaza siêu chịu nhiệt chứa vùng xúc tác endoglucanaza gồm polypeptit bất kỳ trong số các polypeptit từ (A) đến (C) sau.

(A) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11 (tức là AR15G-90-3);

(B) polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ

sung ít nhất một axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0; hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 70% so với trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

Thuật ngữ “polypeptit trong đó axit amin được loại bỏ” được dùng trong bản mô tả để chỉ đoạn axit amin mà tạo ra polypeptit bị khuyết đoạn (loại bỏ).

Thuật ngữ “polypeptit trong đó axit amin được thế” được dùng trong bản mô tả để chỉ axit amin mà tạo ra polypeptit được thay thế bằng axit amin khác.

Thuật ngữ “polypeptit trong đó axit amin được bổ sung” được dùng trong bản mô tả để chỉ axit amin mới được chèn vào polypeptit.

Trong polypeptit (B), tốt hơn nếu số lượng axit amin được loại bỏ, thay thế, hoặc bổ sung trên trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11 nằm trong khoảng từ 1 đến 20, tốt hơn nữa nếu nằm trong khoảng từ 1 đến 10, và còn tốt hơn nữa nếu nằm trong khoảng từ 1 đến 5.

Không có giới hạn cụ thể về độ tương đồng của polypeptit (C) với trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11 miễn là nằm trong khoảng từ 70% đến 100%, tốt hơn nếu nằm trong khoảng từ 75% đến 100%, tốt hơn nữa nếu nằm trong khoảng từ 80% đến 100%, còn tốt hơn nữa nếu nằm trong khoảng từ 85% đến 100%, còn tốt hơn nữa nếu nằm trong khoảng từ 90% đến 100% và tốt nhất nếu nằm trong khoảng từ 95% đến 100%.

Độ tương đồng trình tự (độ tương đồng) giữa các trình tự axit amin được xác định bằng phương pháp sau: ở trạng thái, trong đó khoảng trống được tạo ra ở các vị trí xuất hiện đột biến chèn thêm hoặc loại bỏ, hai trình tự axit amin được đặt cạnh nhau sao cho các axit amin tương ứng là giống nhau; và tỷ lệ axit amin giống nhau so với tổng số trình tự axit amin trừ đi khoảng trống trên trình tự sắp thẳng hàng được xác định là độ tương đồng trình tự. Độ tương đồng giữa các trình tự axit amin có thể được xác định bằng nhiều

phần mềm tra cứu độ tương đồng đã biết trong lĩnh vực này. Theo sáng chế, độ tương đồng của các trình tự axit amin được xác định dựa trên trình tự sắp thẳng hàng thu được bằng phần mềm tra cứu độ tương đồng BLASTP.

Các polypeptit (B) và (C) có thể được thiết kế nhân tạo hoặc cũng có thể là thể tương đồng của AR15G-90-3 và các biến thể tương tự, hoặc các đoạn protein của chúng.

Các polypeptit (A) đến (C) có thể lần lượt được tổng hợp hóa học dựa trên trình tự axit amin, hoặc cũng có thể được sản xuất bởi hệ biểu hiện protein sử dụng polynucleotit theo sáng chế. Ngoài ra, các polypeptit (B) và (C) cũng có thể lần lượt được tổng hợp nhân tạo dựa trên polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11 bằng kỹ thuật tái tổ hợp gen để chèn thêm đột biến axit amin.

Các polypeptit (A) đến (C) có hoạt tính thủy phân cơ chất CMC ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0. Do đó, khi enzym chứa vùng xúc tác endoglucanaza là polypeptit bất kỳ từ (A) đến (C), thì có thể thu được endoglucanaza siêu chịu nhiệt.

Endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân ít nhất một cơ chất được chọn từ nhóm bao gồm hợp chất có liên kết β -1,3 và β -1,4; hợp chất có liên kết β -1,4; và xylan.

Ví dụ về hợp chất có liên kết β -1,3 và β -1,4 bao gồm lichenan, β -glucan, và các hợp chất tương tự. Ví dụ về hợp chất có liên kết β -1,4 bao gồm xenluloza tinh thể như Avicel, xenluloza vi tinh thể của vi khuẩn (BMCC), giấy lọc, CMC, PSA, xenlobioza, và các hợp chất tương tự.

Tốt hơn nếu ngoài việc có hoạt tính thủy phân CMC và xylan, thì endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân ít nhất một cơ chất được chọn từ nhóm bao gồm lichenan, β -glucan, PSA, Avicel, PNPC, và xyloglucan; tốt hơn nữa nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân ít nhất một cơ chất được chọn từ nhóm bao gồm lichenan, PSA, và xyloglucan; và còn tốt hơn nữa nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân ít nhất một cơ chất được chọn từ nhóm bao gồm CMC, xylan, và lichenan.

Ngoài các cơ chất nêu trên, endoglucanaza siêu chịu nhiệt theo sáng chế cũng có

thể có hoạt tính thủy phân cơ chất β -glucan khác. Ví dụ về các cơ chất khác của endoglucanaza siêu chịu nhiệt theo sáng chế bao gồm PNPX, p-nitrophenyl- β -D-glucopyranosit, p-nitrophenyl- α -L-arabinofuranosit, p-nitrophenyl- α -L-arabinopyranosit, p-nitrophenyl- β -L-arabinopyranosit, p-nitrophenyl- β -D-mannopyranosit, p-nitrophenyl- α -D-galactopyranosit, p-nitrophenyl- β -D-galactopyranosit; glucan chứa liên kết β -1,3 và β -1,6 như laminarin; glucan chứa liên kết β -1,3; glucan chứa liên kết β -1,6 và oligosacarit chứa liên kết β -1,6 như gentiobioza.

Tốt hơn nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính thủy phân CMC ít nhất ở độ pH bằng 4,0 và nhiệt độ nằm trong khoảng từ 90°C đến 110°C, tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 80°C đến 120°C, và còn tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 60°C đến 160°C. Tốt hơn nếu nhiệt độ tối ưu của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 90°C đến 120°C ở độ pH bằng 4,0.

Mặc dù thay đổi phụ thuộc vào nhiệt độ phản ứng và cơ chất, nhưng độ pH tối ưu của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 4,0 đến 5,0. Ví dụ, đối với cơ chất CMC, độ pH tối ưu của endoglucanaza siêu chịu nhiệt theo sáng chế ở nhiệt độ 70°C bằng 4,4. Tốt hơn nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính endoglucanaza ít nhất ở độ pH nằm trong khoảng từ 4,0 đến 6,0, và tốt hơn nữa nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính endoglucanaza ở độ pH nằm trong khoảng từ 3,0 đến 8,0.

Tốt hơn nếu endoglucanaza siêu chịu nhiệt theo sáng chế cũng có hoạt tính xylanaza ít nhất ở độ pH bằng 7,0 và nhiệt độ nằm trong khoảng từ 90°C đến 115°C, tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 80°C đến 120°C, và còn tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 40°C đến 140°C. Tốt hơn nếu nhiệt độ tối ưu cho hoạt tính xylanaza của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 90°C đến 120°C ở độ pH bằng 7,0, và tốt hơn nữa nếu nhiệt độ tối ưu cho hoạt tính xylanaza của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 100°C đến 115°C.

Mặc dù thay đổi phụ thuộc vào nhiệt độ phản ứng và các yếu tố tương tự, nhưng độ pH tối ưu cho hoạt tính xylanaza của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 6,5 đến 9,0, và tốt hơn nếu độ pH tối ưu cho hoạt tính xylanaza của endoglucanaza siêu chịu nhiệt theo sáng chế nằm trong khoảng từ 7,5 đến 8,5. Tốt hơn

nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính xylanaza ít nhất ở độ pH nằm trong khoảng từ 6,5 đến 9,0, và tốt hơn nữa nếu endoglucanaza siêu chịu nhiệt theo sáng chế có hoạt tính xylanaza ở độ pH nằm trong khoảng từ 5,0 đến 10,0.

Endoglucanaza siêu chịu nhiệt theo sáng chế cũng có thể có hoạt tính thủy phân glycosit khác với hoạt tính endoglucanaza và hoạt tính xylanaza. Ví dụ về các hoạt tính thủy phân glycosit khác bao gồm hoạt tính β -xylosidaza, hoạt tính β -glucosidaza, hoạt tính xenlobiohydrolaza, hoặc các hoạt tính tương tự.

Endoglucanaza siêu chịu nhiệt theo sáng chế có thể chỉ chứa vùng xúc tác endoglucanaza gồm polypeptit bất kỳ từ (A) đến (C), hoặc có thể chứa các vùng khác. Ví dụ về các vùng khác bao gồm vùng khác với vùng xúc tác enzym của glycosit hydrolaza đã biết. Ví dụ, endoglucanaza siêu chịu nhiệt theo sáng chế cũng bao gồm các enzym thu được bằng cách thay thế vùng xúc tác enzym của glycosit hydrolaza đã biết bằng các polypeptit từ (A) đến (C).

Khi endoglucanaza siêu chịu nhiệt theo sáng chế chứa vùng khác với vùng xúc tác endoglucanaza, thì tốt hơn nếu enzym này chứa đoạn gắn kết xenluloza. Đoạn gắn kết xenluloza có thể nằm trước (đầu tận cùng N) hoặc nằm sau (đầu tận cùng C) vùng xúc tác endoglucanaza. Ngoài ra, đoạn gắn kết xenluloza và vùng xúc tác endoglucanaza có thể được liên kết trực tiếp với nhau hoặc thông qua gốc liên kết có chiều dài thích hợp. Tốt hơn nếu endoglucanaza siêu chịu nhiệt theo sáng chế chứa đoạn gắn kết xenluloza nằm trước hoặc nằm sau vùng xúc tác endoglucanaza thông qua gốc liên kết, tốt hơn nữa nếu endoglucanaza siêu chịu nhiệt theo sáng chế chứa đoạn gắn kết xenluloza nằm sau vùng xúc tác endoglucanaza thông qua gốc liên kết.

Đoạn gắn kết xenluloza chứa trong endoglucanaza siêu chịu nhiệt theo sáng chế có thể thích hợp khi nó là vùng có khả năng gắn kết với xenluloza, ví dụ vùng có khả năng gắn kết với PSA hoặc Avicel tinh thể. Không có giới hạn cụ thể về trình tự axit amin của đoạn gắn kết xenluloza theo sáng chế. Ví dụ, đoạn gắn kết xenluloza của protein đã biết hoặc đoạn cải biến thích hợp của nó có thể được sử dụng làm đoạn gắn kết xenluloza theo sáng chế. Ngoài ra, khi endoglucanaza siêu chịu nhiệt theo sáng chế chứa vùng xúc tác endoglucanaza và đoạn gắn kết xenluloza, tốt hơn nếu các vùng này được liên kết với nhau thông qua gốc liên kết. Không có giới hạn cụ thể về trình tự axit amin, chiều dài, và

các đặc tính tương tự của gốc liên kết này

Ngoài ra, endoglucanaza siêu chịu nhiệt theo sáng chế có thể cũng chứa peptit tín hiệu cho phép định vị enzym này vào vùng đặc hiệu ở tế bào, hoặc peptit tín hiệu tạo ra tác dụng bài tiết enzym này ra ngoài tế bào ở đầu tận cùng N hoặc đầu tận cùng C. Ví dụ về peptit tín hiệu bao gồm peptit tín hiệu vận chuyển không bào, peptit tín hiệu duy trì mạng lưới nội chất, peptit tín hiệu vận chuyển nhân, peptit tín hiệu bài tiết, hoặc các peptit tín hiệu tương tự. Ví dụ về peptit tín hiệu duy trì mạng lưới nội chất bao gồm peptit tín hiệu chứa trình tự axit amin HDEL, hoặc trình tự axit amin tương tự. Khi chứa peptit tín hiệu ở đầu tận cùng N hoặc đầu tận cùng C, thì endoglucanaza siêu chịu nhiệt được biểu hiện ở thể biến nạp theo sáng chế có thể được bài tiết ra ngoài tế bào hoặc được định vị trong mạng lưới nội chất hoặc vùng bên trong tương tự của tế bào.

Ngoài ra, endoglucanaza siêu chịu nhiệt theo sáng chế cũng có thể chứa các trình tự đánh dấu khác nhau ở đầu tận cùng N hoặc đầu tận cùng C để cho phép tinh chế dễ dàng và hiệu quả khi sản xuất enzym này bằng hệ biểu hiện. Các trình tự đánh dấu His, trình tự đánh dấu hemagglutinin (HA), trình tự đánh dấu Myc, và trình tự đánh dấu Flag để biểu hiện và tinh chế protein tái tổ hợp có thể được sử dụng làm trình tự đánh dấu.

Polynucleotit mã hóa endoglucanaza siêu chịu nhiệt theo sáng chế

Polynucleotit nêu trên mã hóa endoglucanaza siêu chịu nhiệt theo sáng chế. Endoglucanaza siêu chịu nhiệt theo sáng chế có thể được sản xuất bằng hệ biểu hiện của vật chủ thu được bằng cách chèn vectơ biểu hiện chứa polynucleotit này vào vật chủ.

Cụ thể hơn, polynucleotit theo sáng chế chứa miền mã hóa vùng xúc tác endoglucanaza, gồm trình tự nucleotit bất kỳ trong số các trình tự nucleotit từ (a) đến (e) sau:.

(a) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11.

(b) trình tự nucleotit mã hóa polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung ít nhất một axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl

xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 70% so với trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11, và có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

(d) trình tự nucleotit có độ tương đồng ít nhất bằng 70% so với trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12, và mã hóa polypeptit có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

(e) trình tự nucleotit của polynucleotit mà lai hóa với polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0.

Thuật ngữ “polynucleotit trong đó nucleotit được loại bỏ” được dùng trong bản mô tả để chỉ đoạn nucleotit mà tạo ra polynucleotit bị khuyết đoạn (loại bỏ).

Thuật ngữ “polynucleotit trong đó nucleotit được thay thế” được dùng trong bản mô tả để chỉ nucleotit mà tạo ra polynucleotit được thay bằng nucleotit khác.

Thuật ngữ “polynucleotit trong đó nucleotit được bổ sung” được dùng trong bản mô tả để chỉ nucleotit mới được chèn vào polynucleotit.

Ví dụ về “các điều kiện nghiêm ngặt” theo sáng chế bao gồm phương pháp được mô tả trong tài liệu: *Molecular Cloning - A Laboratory Manual Third Edition (Sambrook et al., Cold Spring Harbor Laboratory Press)*. Ví dụ, các điều kiện trong dung dịch đệm lai hóa SSC 6× (thành phần của dung dịch SSC 20×: 3M natri clorua, 0,3M dung dịch axit xitric, độ pH bằng 7,0), dung dịch Denhardt 5x (thành phần dung dịch Denhardt 100x: 2% khối lượng albumin huyết thanh bò, 2% khối lượng ficoll, 2% khối lượng polyvinylpyrrolidon), 0,5% khối lượng SDS, 0,1mg/mL ADN tinh trùng cá hồi, và 50% formamit, ở nhiệt độ nằm trong khoảng từ 42°C đến 70°C trong vài giờ đến qua đêm. Sau khi ủ, tốt hơn nếu dung dịch SSC 1× chứa 0,1% khối lượng SDS được sử dụng làm dung dịch đệm rửa, và tốt hơn nữa nếu dung dịch SSC 0,1× chứa 0,1% khối lượng SDS được

sử dụng làm dung dịch đệm rửa.

Từ các trình tự nucleotit từ (a) đến (e), có thể chọn lọc được bộ ba mã hóa thoái biến có tần xuất sử dụng cao ở vật chủ. Ví dụ, trình tự nucleotit (a) có thể là trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12, hoặc trình tự nucleotit thu được bằng cách cải biến để có bộ ba mã hóa có tần xuất sử dụng cao ở vật chủ mà không làm biến đổi trình tự axit amin cần mã hóa bởi trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12. Lưu ý rằng, các bộ ba mã hóa này có thể được cải biến bằng kỹ thuật cải biến trình tự gen hoặc tổng hợp gen nhân tạo đã biết.

Polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12 có thể được tổng hợp hóa học dựa trên thông tin về trình tự nucleotit hoặc polynucleotit này có thể được tổng hợp dưới dạng gen mã hóa AR15G-90-3 có chiều dài đầy đủ (cũng được gọi là “gen AR15G-90-3” hoặc “sản phẩm tách dòng gen AR15G-90-3”) hoặc gen mã hóa AR15G-90B-15 có chiều dài đầy đủ (cũng được gọi là “gen AR15G-90B-15” hoặc “sản phẩm tách dòng gen AR15G-90B-15”) hoặc một đoạn gen của chúng chứa vùng xúc tác endoglucanaza (trong trường hợp của gen AR15G-90-3, polynucleotit này có thể được tổng hợp dưới dạng vùng mã hóa đoạn peptit chứa 171 gốc axit amin từ glyxin (G) ở vị trí số 101 đến leucin (L) ở vị trí số 271 trên SEQ ID NO.3) từ tự nhiên bằng kỹ thuật tái tổ hợp gen. Gen AR15G-90-3 có chiều dài đầy đủ hoặc gen AR15G-90B-15 có chiều dài đầy đủ hoặc một đoạn gen của chúng có thể được tổng hợp bằng cách thu nhận mẫu chứa các vi sinh vật từ tự nhiên; và thực hiện phương pháp PCR với khuôn là ADN bộ gen thu được từ mẫu, cũng như đoạn mỗi xuôi và đoạn mỗi ngược được thiết kế dựa trên trình tự nucleoit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12 bằng phương pháp thông thường. ADN bổ sung được tổng hợp bằng phản ứng phiên mã ngược từ khuôn là ARN thông tin thu được từ mẫu. Lưu ý rằng, tốt hơn nếu mẫu để thu nhận khuôn axit nucleic là mẫu thu được từ môi trường có nhiệt độ cao, như đất ở khu vực suối nước nóng.

Không có giới hạn cụ thể về độ tương đồng của trình tự nucleotit (d) so với trình tự nucleoit như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12 miễn là nằm trong khoảng từ 70% đến 100%, tốt hơn nếu nằm trong khoảng từ 75% đến 100%, tốt hơn nữa nếu nằm trong khoảng từ 80% đến 100%, còn tốt hơn nữa nếu nằm trong khoảng từ 85% đến 100,

còn tốt hơn nữa nếu nằm trong khoảng từ 90% đến 100 và tốt nhất nếu nằm trong khoảng từ 95% đến 100%.

Độ tương đồng trình tự (độ tương đồng) giữa các trình tự nucleotit được xác định bằng phương pháp sau: ở trạng thái trong đó khoảng trống được tạo ra ở các vị trí xuất hiện đột biến chèn thêm hoặc loại bỏ, hai trình tự nucleotit được đặt cạnh nhau sao cho các trình tự nucleotit tương ứng là giống nhau; và tỷ lệ các trình tự nucleotit giống nhau so với tổng số trình tự nucleotit trừ đi khoảng trống trên trình tự sắp thẳng hàng được xác định là độ tương đồng trình tự. Độ tương đồng trình tự giữa các trình tự nucleotit có thể được xác định bằng nhiều phần mềm tra cứu độ tương đồng đã biết trong lĩnh vực này. Theo sáng chế, độ tương đồng của các trình tự nucleotit được xác định dựa trên trình tự sắp thẳng hàng thu được bằng phần mềm tra cứu độ tương đồng BLASTN.

Ví dụ, polynucleotit có trình tự nucleotit (b), (c), hoặc (d) có thể lần lượt được tổng hợp nhân tạo bằng cách loại bỏ, thay thế, hoặc bổ sung một hoặc hai hoặc nhiều nucleotit trong polynucleotit có trình tự như nêu trong SEQ ID NO.4 hoặc SEQ ID NO.12. Ngoài ra, trình tự nucleotit (b), (c), hoặc (d) cũng có thể là trình tự có chiều dài đầy đủ của gen tương đồng với gen AR15G-90-3 (hoặc gen AR15G-90B-15) hoặc một đoạn trình tự của nó. Gen tương đồng với gen AR15G-90-3 (hoặc gen AR15G-90B-15) có thể thu được bằng kỹ thuật gen tái tổ hợp dùng để thu nhận gen tương đồng của gen có trình tự nucleotit đã biết.

Polynucleotit theo sáng chế có thể chỉ chứa miền mã hóa vùng xúc tác endoglucanaza, hoặc ngoài miền mã hóa vùng xúc tác endoglucanaza, thì polynucleotit theo sáng chế cũng có thể chứa miền mã hóa đoạn gắn kết xenluloza, gốc liên kết, các petit tín hiệu khác nhau, các trình tự đánh dấu khác nhau, hoặc các trình tự tương tự.

Vectơ biểu hiện theo sáng chế

Vectơ nêu trên chứa polynucleotit theo sáng chế và có khả năng biểu hiện polypeptit có hoạt tính thủy phân cơ chất carboxymetyl xenluloza ít nhất ở nhiệt độ bằng 110°C và độ pH bằng 4,0 ở tế bào vật chủ. Cụ thể, vectơ biểu hiện này chứa polynucleotit mã hóa endoglucanaza siêu chịu nhiệt theo sáng chế. Đặc biệt, vectơ biểu hiện cần được kết hợp với cat-xet biểu hiện chứa, từ đoạn nằm trước, ADN có trình tự vùng khởi đầu,

polynucleotit theo sáng chế, và ADN có trình tự vùng kết thúc. Polynucleotit theo sáng chế có thể được biến nạp vào vector biểu hiện này bằng kỹ thuật tái tổ hợp gen đã biết. Kit tổng hợp vector biểu hiện có bán trên thị trường cũng có thể được sử dụng để biến nạp polynucleotit vào vector biểu hiện.

Thuật ngữ “vector biểu hiện” được dùng trong bản mô tả để chỉ vector chứa, từ đoạn nằm trước, ADN có trình tự vùng khởi đầu, ADN có trình tự để kết hợp AND ngoại lai, và ADN có trình tự vùng kết thúc.

Vector biểu hiện theo sáng chế có thể được chèn vào tế bào nhân sơ như *E. coli* hoặc tế bào nhân thật như nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú, hoặc tế bào thực vật. Vector bất kỳ tương thích với vật chủ có thể được sử dụng làm vector biểu hiện theo sáng chế.

Tốt hơn nếu, vector biểu hiện theo sáng chế không chỉ chứa polynucleotit theo sáng chế, mà còn chứa gen kháng thuốc hoặc các gen tương tự. Điều này cho phép dễ dàng chọn lọc được tế bào vật chủ đã chuyển gen và tế bào vật chủ chưa chuyển gen bằng vector này.

Ví dụ về gen kháng thuốc bao gồm gen kháng kanamycin, gen kháng hygromycin, gen kháng bialaphos, hoặc các gen tương tự.

Thế biến nạp theo sáng chế

Thế biến nạp nêu trên chứa vector biểu hiện theo sáng chế. Endoglucanaza siêu chịu nhiệt theo sáng chế có thể được biểu hiện ở thế biến nạp này. Vật chủ để chèn vector biểu hiện theo sáng chế có thể là tế bào nhân sơ như *E. coli*, hoặc tế bào nhân thật như nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú hoặc tế bào thực vật. Tức là, thế biến nạp theo sáng chế là *E. coli*, nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú, tế bào thực vật, hoặc tế bào tương tự chứa vector biểu hiện theo sáng chế.

Endoglucanaza siêu chịu nhiệt theo sáng chế có thể được sản xuất một cách đơn giản với hiệu suất cao bằng cách nuôi cấy thế biến nạp *E. coli* này. Mặt khác, do protein bị thủy phân ở tế bào nhân thật, nên endoglucanaza siêu chịu nhiệt có độ ổn định nhiệt cao hơn có thể được sản xuất bằng cách sử dụng thế biến nạp là tế bào nhân thật chứ

không sử dụng thể biến nạp là tế bào nhân sơ

Không có giới hạn cụ thể về phương pháp sản xuất thể biến nạp theo sáng chế bằng cách sử dụng vector biểu hiện và thể biến nạp theo sáng chế có thể được sản xuất bằng phương pháp thông thường. Ví dụ về các phương pháp sản xuất thể biến nạp bao gồm phương pháp sốc nhiệt, phương pháp chuyển gen nhờ *Agrobacterium*, phương pháp súng hạt, phương pháp xung điện, phương pháp chuyển gen nhờ polyetylen glycol, và các phương pháp tương tự. Khi vật chủ là tế bào thực vật, tốt hơn nếu thể biến nạp theo sáng chế được sản xuất bằng phương pháp súng hạt hoặc phương pháp chuyển gen nhờ *Agrobacterium*.

Khi tế bào nhân sơ, nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú hoặc các tế bào tương tự được sử dụng làm vật chủ, thì thể biến nạp thu được có thể được nuôi cấy bằng phương pháp thông thường giống như phương pháp nuôi cấy vật chủ chưa chuyển gen.

Phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế

Phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế chế là phương pháp để sản xuất endoglucanaza siêu chịu nhiệt trong thể biến nạp theo sáng chế. Khi nuôi cấy thể biến nạp được sản xuất bằng cách sử dụng vector biểu hiện chứa polynucleotit theo sáng chế nằm sau vùng khởi đầu mà không có khả năng điều hòa thời gian biểu hiện hoặc các yếu tố tương tự, thì endoglucanaza siêu chịu nhiệt theo sáng chế được biểu hiện liên tục. Mặt khác, trong thể biến nạp được sản xuất bằng cách sử dụng vùng khởi đầu cảm ứng biểu hiện để cảm ứng biểu hiện bằng hợp chất đặc hiệu, điều kiện nhiệt độ đặc hiệu, hoặc các yếu tố tương tự, thì endoglucanaza siêu chịu nhiệt được biểu hiện trong thể biến nạp này bằng cách nuôi cấy thể biến nạp và thực hiện cảm ứng phù hợp với điều kiện cảm ứng biểu hiện tương ứng

Endoglucanaza siêu chịu nhiệt được sản xuất bằng thể biến nạp có thể được sử dụng ở trạng thái vẫn nằm trong thể biến nạp, hoặc có thể được chiết xuất/tinh chế từ thể biến nạp.

Không có giới hạn cụ thể về phương pháp chiết xuất hoặc tinh chế endoglucanaza siêu chịu nhiệt từ thể biến nạp miễn là phương pháp này không làm giảm hoạt tính thủy

phân glycosit của endoglucanaza siêu chịu nhiệt này, và phương pháp thông thường để chiết xuất polypeptit từ tế bào hoặc mô sinh học có thể được sử dụng. Ví dụ về các phương pháp này bao gồm phương pháp, trong đó thể biến nạp được ngâm trong dung dịch đệm thích hợp để chiết xuất endoglucanaza siêu chịu nhiệt, sau đó dịch chiết và bã rắn được phân tách. Tốt hơn nếu dung dịch đệm chứa chất tăng độ hòa tan như chất hoạt động bề mặt. Khi thể biến nạp là thực vật, thì thể biến nạp có thể được băm hoặc nghiền trước khi ngâm trong dung dịch đệm. Hơn nữa, phương pháp phân tách rắn-lỏng đã biết như phương pháp lọc, phương pháp lọc nén, hoặc phương pháp ly tâm có thể được sử dụng để phân tách dịch chiết và bã rắn, hoặc thể biến nạp được ngâm trong dung dịch đệm có thể được ép. Endoglucanaza siêu chịu nhiệt trong dịch chiết có thể được tinh chế bằng phương pháp đã biết như phương pháp kết tinh muối, phương pháp siêu lọc, hoặc phương pháp sắc ký.

Khi endoglucanaza siêu chịu nhiệt theo sáng chế được biểu hiện ở trạng thái có peptit tín hiệu bài tiết nằm trong thể biến nạp, thì dung dịch chứa endoglucanaza siêu chịu nhiệt có thể được thu nhận dễ dàng bằng cách nuôi cấy thể biến nạp, sau đó thu gom phần dịch nổi nuôi cấy bằng cách loại bỏ thể biến nạp ra khỏi sản phẩm nuôi cấy thu được. Hơn nữa, khi endoglucanaza siêu chịu nhiệt theo sáng chế có trình tự đánh dấu như trình tự đánh dấu His, thì endoglucanaza siêu chịu nhiệt trong dịch chiết hoặc dịch nổi nuôi cấy có thể được tinh chế dễ dàng bằng phương pháp sắc ký ái lực sử dụng trình tự đánh dấu này.

Nói cách khác, phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế bao gồm bước nuôi cấy thể biến nạp theo sáng chế để sản xuất endoglucanaza siêu chịu nhiệt, sau đó là bước chiết xuất và tinh chế endoglucanaza siêu chịu nhiệt từ thể biến nạp (khi cần).

Hỗn hợp glycosit hydrolaza theo sáng chế

Hỗn hợp glycosit hydrolaza nêu trên chứa endoglucanaza siêu chịu nhiệt theo sáng chế, hoặc endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế, và ít nhất một hoặc nhiều loại glycosit hydrolaza khác. Endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế có thể được sử dụng ở trạng thái nằm trong

thể biến nạp theo sáng chế hoặc có thể được chiết xuất hoặc tinh chế từ thể biến nạp này. Khi hỗn hợp glycosit hydrolaza theo sáng chế được sử dụng để thủy phân polysacarit, thì có thể phân giải hiệu quả lignoxenluloza.

Không có giới hạn cụ thể về glycosit hydrolaza khác với endoglucanaza siêu chịu nhiệt theo sáng chế miễn là nó có hoạt tính thủy phân lignoxenluloza. Ví dụ về glycosit hydrolaza khác bao gồm hemixenlulaza như xylanaza, β -xylosidaza, xenlobiohydrolaza, β -glucosidaza, endoglucanaza, hoặc các enzym tương tự. Tốt hơn nếu ngoài endoglucanaza siêu chịu nhiệt theo sáng chế, thì hỗn hợp glycosit hydrolaza theo sáng chế còn chứa ít nhất một glycosit hydrolaza (tức là hemixenlulaza hoặc endoglucanaza) và tốt hơn nữa nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa cả hai glycosit hydrolaza này (tức là chứa cả hemixenlulaza lẫn endoglucanaza). Đặc biệt, ngoài endoglucanaza siêu chịu nhiệt theo sáng chế, tốt hơn nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa ít nhất một glycosit hydrolaza được chọn từ nhóm bao gồm xylanaza, β -xylosidaza, xenlobiohydrolaza, và β -glucosidaza; và tốt hơn nữa nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa toàn bộ các glycosit hydrolaza (tức là xylanaza, β -xylosidaza, xenlobiohydrolaza, và β -glucosidaza).

Tốt hơn nếu glycosit hydrolaza khác chứa trong hỗn hợp glycosit hydrolaza theo sáng chế là glycosit hydrolaza chịu nhiệt có hoạt tính ít nhất ở nhiệt độ bằng 85°C, và tốt hơn nữa nếu enzym này là glycosit hydrolaza chịu nhiệt có hoạt tính ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C. Khi hỗn hợp glycosit hydrolaza theo sáng chế chứa toàn bộ các enzym chịu nhiệt (ví dụ nhiệt độ tối ưu của hoạt tính enzym hoặc nhiệt độ biến tính của protein enzym bằng 70°C hoặc cao hơn), thì có thể phân giải hiệu quả lignoxenluloza bằng hỗn hợp glycosit hydrolaza này ở điều kiện nhiệt độ cao. Cụ thể, khi hỗn hợp glycosit hydrolaza theo sáng chế chỉ chứa glycosit hydrolaza chịu nhiệt được sử dụng để thủy phân lignoxenluloza, thì có thể thực hiện phản ứng thủy phân lignoxenluloza ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C (quá trình thủy phân ở nhiệt độ cao). Đối với quá trình thủy phân ở nhiệt độ cao này, thì lượng enzym, thời gian thủy phân và chi phí cho phản ứng thủy phân có thể giảm.

Phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế

Phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương

pháp để thu được sản phẩm bằng cách thủy phân nguyên liệu lignoxenluloza chứa xenluloza để tạo ra oligosacarit bằng endoglucanaza siêu chịu nhiệt theo sáng chế. Cụ thể, sản phẩm phân giải lignoxenluloza chứa sản phẩm phân giải hemixenluloza hoặc xenluloza được sản xuất bằng cách cho nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza tiếp xúc với endoglucanaza siêu chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế, hoặc hỗn hợp glycosit hydrolaza theo sáng chế.

Thuật ngữ “sản phẩm phân giải hemixenluloza hoặc xenluloza” được dùng trong bản mô tả để chỉ sản phẩm được tạo ra bằng cách phân cắt liên kết glycosidic của hemixenluloza hoặc xenluloza.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương pháp, trong đó nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được cho tiếp xúc với endoglucanaza siêu chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế để thu được sản phẩm chủ yếu được tạo ra bằng cách phân cắt liên kết glycosidic của xenluloza.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương pháp, trong đó nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được cho tiếp xúc với hỗn hợp glycosit hydrolaza theo sáng chế để thu được sản phẩm được tạo ra bằng cách phân cắt liên kết glycosidic của hemixenluloza hoặc xenluloza.

Không có giới hạn cụ thể về nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, miễn là nó chứa hemixenluloza hoặc xenluloza. Ví dụ về nguyên liệu lignoxenluloza bao gồm sinh khối xenluloza như cỏ dại, chất thải nông nghiệp, giấy thải, hoặc các nguyên liệu tương tự. Trước khi tiếp xúc với endoglucanaza siêu chịu nhiệt theo sáng chế, tốt hơn nếu nguyên liệu chứa hemixenluloza hoặc xenluloza được xử lý vật lý, như nghiền hoặc băm nhỏ, xử lý hóa học bằng axit, kiềm, hoặc hợp chất tương tự, hoặc ngâm hoặc phân giải vào dung dịch đệm thích hợp, hoặc dung dịch tương tự.

Tốt hơn nếu phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza bởi endoglucanaza siêu chịu nhiệt theo sáng chế được thực hiện ở điều kiện mà ở đó enzym này có hoạt tính endoglucanaza, và tốt hơn nếu phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza bởi endoglucanaza siêu chịu nhiệt theo sáng chế được thực hiện ở điều kiện mà ở đó enzym này có hoạt tính endoglucanaza và hoạt tính xylanaza. Tốt hơn nếu phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 60°C đến 160°C và độ pH nằm trong khoảng từ 3,0 đến 8,0, và tốt hơn nữa nếu phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 80°C đến 120°C và độ pH nằm trong khoảng từ 5,0 đến 8,0 và còn tốt hơn nữa nếu phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 90°C đến 110°C và độ pH nằm trong khoảng từ 4,0 đến 7,0. Thời gian phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza được điều chỉnh thích hợp theo loại nguyên liệu được sử dụng, phương pháp xử lý sơ bộ nguyên liệu, lượng nguyên liệu, hoặc các yếu tố tương tự. Ví dụ, phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza có thể được thực hiện trong thời gian từ 10 phút đến 100 giờ và khi phân giải sinh khối xenluloza, thì phản ứng thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza có thể được thực hiện trong thời gian từ 1 giờ đến 100 giờ.

Để thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, cũng tốt hơn nếu sử dụng ít nhất một hoặc nhiều loại glycosit hydrolaza khác ngoài endoglucanaza siêu chịu nhiệt theo sáng chế. Glycosit hydrolaza khác có thể giống như glycosit hydrolaza chứa trong hỗn hợp glycosit hydrolaza nêu trên, và tốt hơn nếu enzym này là glycosit hydrolaza chịu nhiệt có hoạt tính ít nhất ở nhiệt độ bằng 85°C, và tốt hơn nếu ở nhiệt độ ít nhất nằm trong khoảng từ 70°C đến 90°C. Ngoài ra, theo một khía cạnh phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế sử dụng endoglucanaza siêu chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp sản xuất endoglucanaza siêu chịu nhiệt theo sáng chế, và theo một khía cạnh khác phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế sử dụng hỗn hợp glycosit hydrolaza theo

sáng chế.

Ví dụ thực hiện sáng chế

Sáng chế sẽ được mô tả chi tiết thông các ví dụ dưới đây. Tuy nhiên, sáng chế không bị giới hạn bởi các ví dụ này.

Ví dụ 1

Quy trình tách dòng endoglucanaza siêu chịu nhiệt thể hệ mới từ đất ở khu vực suối nước nóng.

1. Chiết tách ADN từ đất ở khu vực suối nước nóng và giải trình tự bộ gen

Để giải trình tự gen của endoglucanaza siêu chịu nhiệt theo sáng chế, ADN từ đất ở khu vực suối nước nóng có độ pH từ trung tính đến kiềm yếu được thu nhận và ADN hệ gen môi trường của quần thể vi sinh vật trong đất được giải trình tự.

Mẫu nước chứa đất, bùn và sinh khối được thu nhận từ 5 khu vực suối nước nóng ở 3 vùng của Nhật Bản được sử dụng làm mẫu đất ở khu vực suối nước nóng có độ pH từ trung tính đến kiềm yếu (các mẫu ADN hệ gen môi trường N2, AR19, AR15, OJ1, và H1). Các mẫu đất ở khu vực suối nước nóng này có nhiệt độ nằm trong khoảng từ 58°C đến 78°C và độ pH nằm trong khoảng từ 7,2 đến 8,0 ở thời điểm lấy mẫu.

ADN được chiết tách bằng kit chiết tách ADN từ 10g mỗi mẫu đất thu được ở các khu vực suối nước nóng (ISOIL Large for Beads ver. 2, do NIPPON GENE Co., Ltd. sản xuất). ADN chiết tách được (5µg) được giải trình tự bằng máy giải trình tự GS FLX Titan 454 do Roche Diagnostics K.K sản xuất. Lượng ADN còn lại được sử dụng để tách dòng PCR gen endoglucanaza.

Quá trình giải trình tự ADN hệ gen môi trường này được hiện bằng cách sử dụng mẫu đất suối nước nóng AR15. Kết quả là thu được trình tự bộ gen có chiều dài đoạn đọc trung bình bằng 370bp, tổng số các đoạn trình tự đọc bằng 5.419.406, và tổng chiều dài bộ gen đã được giải trình tự bằng 2.007.725.040bp.

2. Quá trình lắp ráp và thống kê dữ liệu hệ gen môi trường của mẫu đất thu được ở khu vực suối nước nóng.

Liên quan đến trình tự nucleotit đã được đọc bằng máy giải trình tự GS FLX Titan 454, dữ liệu đầu ra (tệp tin sff) được xử lý đệm bằng phần mềm PyroBayes (Quinlan *et al.*, *Nature Methods*, **2008**, Vol. 5, p.179-81.) để thu được các tệp trình tự và tệp trình tự chính xác định dạng FASTA. Sau khi cắt bỏ các đầu tận cùng của chúng để tăng độ chính xác, các trình tự đọc thu được được lắp ráp bằng phần mềm Newbler phiên bản 2.3 được tích hợp với máy giải trình tự 454 Life Science. Quá trình lắp ráp được thực hiện trong các điều kiện: “độ trùng lặp tối thiểu chấp nhận được (mi) = 0,9”, “tùy chọn: - lớn (đối với các bộ gen có kích cỡ lớn hoặc phức tạp, tăng tốc độ lắp ráp, nhưng làm giảm độ chính xác)”.

Tổng chiều dài của các đoạn tiếp giáp mà đã được lắp ráp thành 100bp hoặc cao hơn bằng 118.600.846bp. Bộ dữ liệu này được sử dụng để phân tích gen mã hóa xenlulaza. Trong tổng số 5.419.406 đoạn trình tự đọc, thì 4.805.640 đoạn trình tự đọc được lắp ráp thành 103.508 đoạn tiếp giáp có chiều dài trung bình bằng 1.146bp hoặc dài hơn. Trong số các đoạn tiếp giáp này, đoạn tiếp giáp lớn nhất có chiều dài bằng 151.585bp.

3. Quá trình phát hiện khung đọc mã mở (Open Reading Frame - ORF) của β -xylosidaza.

Các trình tự có mã phân loại enzym 3.2.1.4 (xenlulaza), 3.2.1.21 (β -glucosidaza), 3.2.1.37 (β -xylosidaza), 3.2.1.91 (xenluloza 1,4- β -xenlobiosidaza), và 3.2.1.8 (endo 1,4- β -xylanaza) được tải về từ cơ sở dữ liệu UniProt (<http://www.uniprot.org/>, ngày truy cập: 09/12/2011), và cơ sở dữ liệu về các gen mã hóa glycosit hydrolaza này được thiết kế. Khung đọc mã mở (các vùng gen) được phát hiện từ các trình tự tiếp giáp thu được ở phần 2 bằng phần mềm Orphelia (Hoff *et al.*, *Nucleic Acids Research*, **2009**, 37 (Web Server issue: W101-W105, tùy chọn Orphelia: mặc định (kiểu = Net 700, độ trùng lặp tối đa = 60)). Gen mã hóa glycosit hydrolaza từ trình tự khung đọc mã mở phát hiện được được xác định bằng phần mềm BLASTP (phiên bản 2.2.18) và tham khảo cơ sở dữ liệu nêu trên. Điều kiện tùy chọn của phần mềm BLASTP được cài đặt như sau: “trình tự truy vấn = sai”, “trị số kỳ vọng (E) < $1e^{-20}$ ”, “trị số mặc định - trị số để phát hiện khoảng trống = -1, trị số để mở rộng khoảng trống = -1, trị số sắp thẳng hàng được tạo khoảng trống X = 0, trị số lý thuyết cho các trình tự mục tiêu mở

rộng = 0, cỡ chữ = mặc định”, và các trình tự khung đọc mã mở mục tiêu được thu thập là gen mã hóa glycosit hydrolaza. Các trình tự nucleotit thu được bao gồm các gen mã hóa glycosit hydrolaza như xenulaza, endohemixenulaza, và các enzym cắt mạch nhánh.

4. Quá trình phân loại các gen thành các họ glycosit hydrolaza.

Các trình tự nucleotit thu được ở phần 3 được phân loại chức năng bằng phần mềm pfam HMMs (Pfam phiên bản 23.0 và HMMER phiên bản 2.3 Finn *et al.*, *Nucleic Acids Research Database*, 2010, Issue 38, p.D211-222) dưới dạng cơ sở dữ liệu về trình tự vùng chức năng của protein. Cụ thể, mỗi trình tự nucleotit này được phân loại thành họ glycosit hydrolaza (GH) dựa trên độ tương đồng của chúng với cơ sở dữ liệu vùng Pfam bằng thuật toán tra cứu độ tương đồng trình tự HMMER (Durbin *et al.*, “*The theory behind profile HMMs. Biological sequence analysis: probabilistic models of protein and nucleic acids*”, 1998, Cambridge University Press.; hmmpfam (phiên bản 2.3.2), trị số kỳ vọng $E < 1e^{-5}$; cơ sở dữ liệu = Pfam_fs (mô hình có thể được sử dụng để phát hiện các đoạn của các vùng đại diện ở trình tự)). Lưu ý rằng biến thể chứa 70% trình tự của vùng xúc tác GH hoặc cao hơn được tính là các gen mã hóa enzym thuộc các họ tương ứng.

Từ quá trình tra cứu độ tương đồng bằng phần mềm BLASTP dựa trên dữ liệu trình tự hệ gen môi trường AR15, 106 khung đọc mã mở được xác định là gen mã hóa endoglucanaza. Kết quả phân loại họ GH của các khung đọc mã mở này được thể hiện trong Bảng 1. Như được thể hiện trong Bảng 1, thu được 13 khung đọc mã mở có chiều dài đầy đủ của gen endoglucanaza thuộc họ GH5, 4 khung đọc mã mở có chiều dài đầy đủ của gen endoglucanaza thuộc họ GH9, và 4 khung đọc mã mở có chiều dài đầy đủ của gen endoglucanaza thuộc họ GH12 từ hệ gen môi trường AR15. Các đoạn mỗi được thiết kế cho toàn bộ các khung đọc mã mở này đã được xác định là gen endoglucanaza, và các gen này được tách dòng từ ADN thu được từ đất ở khu vực suối nước nóng bằng phương pháp PCR. Kết quả là gen endoglucanaza được phân lập từ khung đọc mã mở AR15G-90 thuộc họ GH12 và có trình tự của gen endoglucanaza.

Bảng 1

Hệ gen môi trường AR19	Phân loại họ GH của các gen β -glucosidaza hoặc β -xylosidaza					
	GH5	GH9	GH12	GH48	GH khác	Tổng số
Các khung đọc mã mở có chiều dài đầy đủ	13	4	4	0	40	61
Các khung đọc mã mở có chiều dài không đầy đủ	3	3	1	1	37	45
Tổng số khung đọc mã mở	16	7	5	1	77	106

5. Khung đọc mã mở AR15G-90

Khung đọc mã mở AR15G-90 mã hóa polypeptit chứa 283 gốc axit amin và có chiều dài không đầy đủ (SEQ ID NO.1), polypeptit này bắt đầu từ valin (V) ở vị trí số 1, và đầu 3' của trình tự nucleotit mã hóa polypeptit này được kết thúc bằng bộ ba mã hóa kết thúc. Kết quả phân tích độ tương đồng trình tự cho thấy rằng trong polypeptit mã hóa bởi khung đọc mã mở AR15G-90, đoạn peptit chứa 171 gốc axit amin từ glyxin (G) ở vị trí số 101 đến leucin (L) ở vị trí số 271 được xác định là vùng xúc tác của glycosit hydrolaza GH12. Trình tự axit amin đã biết có độ tương đồng cao nhất với trình tự axit amin mã hóa bởi khung đọc mã mở AR15G-90 là glycosit hydrolaza GH12 có nguồn gốc từ *Ignisphaera aggregans* DSM17230 thuộc ngành *Crenarchaeota* (mã truy cập Ngân hàng Gen: ADM27292.1). Do độ tương đồng giữa hai trình tự axit amin này cho toàn bộ chiều dài bằng 56% và 64% cho vùng xúc tác GH12 như được xác định bằng thuật toán ClustalW, nên khung đọc mã mở AR15G-90 được xác định là trình tự thể hệ mới.

Fig.1 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin của polypeptit mã hóa bởi khung đọc mã mở AR15G-90 và trình tự axit amin của glycosit hydrolaza GH12 (SEQ ID NO.8) có nguồn gốc từ *Ignisphaera aggregans* DSM17230. Trên Fig.1, các axit amin nằm sau màu đen/trắng thể hiện các gốc axit amin giống hệt nhau trong toàn bộ các trình tự axit amin này, và ký tự “-” thể hiện các đoạn gạch bỏ (khoảng trống)

6. Quá trình tách dòng gen từ khung đọc mã mở AR15G-90.

Phản ứng PCR được thực hiện bằng cách sử dụng khuôn là ADN thu được từ đất ở khu vực suối nước nóng mà đã được khuếch đại bằng kit khuếch đại ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare sản xuất), và đoạn môi xuôi có trình tự nucleotit như nêu trong SEQ ID NO.7 (5'-

CACCATGGTGACCATCACGCCGAGTACA-3' được thiết kế bằng cách bổ sung 4 nucleotit (CACC) và bộ ba mã hóa khởi đầu (ATG) vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.5. 4 nucleotit CACC được bổ sung ở đầu 5' là trình tự để chèn vào vector) và đoạn mỗi ngược có trình tự nucleotit như nêu trong SEQ ID NO.6 (5'-CTACCTCAGTGTTTTACCTGGC-3'). Trình tự nucleotit như nêu trong SEQ ID NO.5 là trình tự nucleotit tương đồng với đoạn trình tự chứa các nucleotit ở vị trí số 1 đến 21 của trình tự nucleotit như nêu trong SEQ ID NO.2. Hơn nữa, trình tự nucleotit như nêu trong SEQ ID NO.6 là trình tự nucleotit bổ sung với đoạn trình tự chứa các nucleotit ở vị trí số 831 đến 852 của trình tự nucleotit như nêu trong SEQ ID NO.2. Các sản phẩm PCR đã được khuếch đại được chèn vào vector pET101/D-TOPO bằng kit biểu hiện Champion pET Directional TOPO (do Life Technologies sản xuất), và biến nạp vào chủng One Shot TOP10. Các dòng đã biểu hiện được chọn lọc bằng phương pháp PCR trực tiếp từ khuẩn lạc, sau đó nuôi cấy trong môi trường lỏng LB chứa ampicilin với nồng độ bằng 100mg/L ở nhiệt độ 37°C và tốc độ 200 vòng/phút trong 17 đến 20 giờ, tiếp theo các plasmit được tổng hợp bằng kit miniprep (hệ thống tinh chế ADN Wizard plus SV Minipreps, do Promega sản xuất). Các plasmit đã tổng hợp được giải trình tự bằng máy giải trình tự ADN 3730 do Life Technologies sản xuất.

Kết quả thu được 5 sản phẩm tách dòng gen AR15G-90-2, AR15G-90-3, AR15G-90-9, AR15G-90-12, và AR15G-90-13 từ khung đọc mã mở AR15G-90 bằng quá trình tách dòng PCR. Trình tự nucleotit của gen endoglucanaza tiềm năng AR15G-90-3 (SEQ ID NO.4) có chiều dài 852bp tương đồng với trình tự nucleotit của khung đọc mã mở AR15G-90 (SEQ ID NO.2), và chỉ khác trình tự SEQ ID NO.2 bởi 4 nucleotit. Tức là, nucleotit ở vị trí số 93 trên khung đọc mã mở AR15G-90 là thymin (T), trong khi đó nucleotit ở vị trí số 93 trên khung đọc mã mở AR15G-90-3 là xytosin (C); nucleotit ở vị trí số 335 trên khung đọc mã mở AR15G-90 là guanin (G), trong khi đó nucleotit ở vị trí số 335 trên khung đọc mã mở AR15G-90-3 là adenin (A); nucleotit ở vị trí số 652 trên khung đọc mã mở AR15G-90 là guanin (G), trong khi đó nucleotit ở vị trí số 652 trên khung đọc mã mở AR15G-90-3 là adenin (A); và nucleotit ở vị trí số 699 trên khung đọc mã mở AR15G-90 là guanin (G), trong khi đó nucleotit ở vị trí số 699 trên khung đọc mã mở AR15G-90-3 là adenin (A). Tất nhiên, sự khác biệt về các nucleotit ở hai vị trí cũng dẫn đến sự khác biệt về các axit amin, và trình tự axit amin của khung đọc mã mở

AR15G-90 (SEQ ID NO.1) và trình tự axit amin của gen endoglucanaza tiềm năng AR15G-90-3 (SEQ ID NO.3) khác nhau bởi 2 gốc axit amin. Tức là, gốc axit amin ở vị trí số 112 trên khung đọc mã mở AR15G-90 là arginin (R), trong khi đó gốc axit amin ở vị trí số 112 trên khung đọc mã mở AR15G-90-3 là lysin (K); và gốc axit amin ở vị trí số 218 trên khung đọc mã mở AR15G-90 là valin (V), trong khi đó gốc axit amin ở vị trí số 218 trên khung đọc mã mở AR15G-90-3 là isoleucin (I).

7. Quá trình biểu hiện gen AR15G-90-3 và tinh chế protein enzym.

Protein được tổng hợp bằng cách bổ sung trình tự đánh dấu histidin vào đầu tận cùng C được sử dụng để tinh chế protein enzym. Trình tự được bổ sung trình tự đánh dấu histidin được tổng hợp bằng cách chèn gen endoglucanaza tiềm năng AR15G-90-3 đã được loại bỏ bộ ba mã hóa kết thúc vào vector pET101/D-TOPO. Các plasmid chứa gen đích được bổ sung trình tự đánh dấu histidin được biến nạp vào *E. coli* để biểu hiện protein bằng phương pháp sốc nhiệt. Chủng Rosetta-gamiB (DE3) pLysS (do Merck sản xuất) được sử dụng làm tế bào chủ để chuyển gen. *E. coli* chứa gen đích được cấy vào môi trường LB chứa ampicilin với hàm lượng bằng 100mg/L và nuôi cấy cho đến khi mật độ quang ở bước sóng 600nm nằm trong khoảng từ 0,2 đến 0,8, sau đó IPTG (isopropyl- β -D(-)-thiogalactopyranosit) được bổ sung vào, và nuôi cấy tiếp trong 5 đến 20 giờ. Kết quả là quá trình biểu hiện protein đích được cảm ứng. Sau khi nuôi cấy, *E. coli* được thu nhận bằng cách ly tâm, và dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) có thể tích bằng 1/10 thể tích dịch nuôi cấy được bổ sung vào để tạo hỗn dịch vi khuẩn. Sau đó, quy trình phân cắt trong 5 phút và rung lắc trong 5 phút được lặp lại từ 7 đến 8 lần bằng thiết bị Astrason 3000 (do Misonix, Inc. sản xuất). Kết quả là thu được dịch chiết thô chứa *E. coli* tái tổ hợp gen chứa protein đích. Dịch chiết thô chứa *E. coli* tái tổ hợp gen được lọc qua màng lọc (đường kính lỗ lọc $\phi = 0,45\mu\text{m}$, do Millipore sản xuất), và dịch lọc thu được được sử dụng làm dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen.

NaCl được bổ sung vào dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen đến nồng độ cuối cùng bằng 500mM, và hỗn hợp thu được được nạp vào cột trao đổi ion HisTrap FF (do GE Healthcare sản xuất) được làm cân bằng với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 500mM NaCl, các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 100% với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0)

chứa 500mM NaCl và 500mM imidazol bằng hệ thống sắc ký lỏng AKTA (do GE Healthcare sản xuất).

Các phân đoạn có hoạt tính thủy phân CMC được trộn toàn bộ, sau đó trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 750mM amoni sulfat bằng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất). Các phân đoạn có hoạt tính thủy phân CMC sau khi trao đổi dung dịch được nạp vào cột phân tách tương tác kỵ nước HiTrap Phenyl HP (do GE Healthcare sản xuất) được làm cân bằng với cùng dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 750mM amoni sulfat, các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 100% với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0). Các phân đoạn có hoạt tính thủy phân CMC được trộn toàn bộ, sau đó cô bằng màng siêu lọc ly tâm VIVASPIN 20 cho đến khi thể tích dung dịch bằng khoảng 8mL. Mẫu đã cô được nạp vào cột lọc gel Hiload 26/60 superdex 200 pg (do GE Healthcare sản xuất) được làm cân bằng với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 150mM NaCl, và được phân đoạn bằng cùng dung dịch đệm có thể tích gấp 1 đến 1,5 lần thể tích của cột ở tốc độ dòng nằm trong khoảng từ 2mL/phút đến 3mL/phút. Các phân đoạn có hoạt tính thủy phân CMC được trộn toàn bộ, sau đó trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) và cô. Kết quả là thu được enzym tinh khiết có nồng độ cuối cùng bằng khoảng 1,0mg/mL.

Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và enzym tinh khiết này được kiểm chứng bằng phân tích điện di trên gel SDS-PAGE (điện di trên gel SDS-polyacrylamit). Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (5µg) và enzym tinh khiết (0,5µg) lần lượt được trộn với dung dịch đã được cô 4 lần chứa 2-mercaptoetanol chứa dung dịch đệm mẫu (do Wako Pure Chemical Industries, Ltd. sản xuất) sao cho nồng độ dung dịch đệm bằng 1 lần, sau đó gia nhiệt ở nhiệt độ 95°C trong 4 phút, tiếp theo điện di bằng gel không chứa thuốc nhuộm Criterion TGX 10% (do Bio-Rad Laboratories, Inc. sản xuất). Sau khi chạy điện di, các băng protein được hiện màu bằng hệ thống ChemiDoc (do Bio-Rad Laboratories, Inc. sản xuất).

Fig.2 thể hiện kết quả phân tích điện di trên gel SDS-PAGE của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen được điều chế từ *E. coli* đã chuyển gen AR15G-90-3 và enzym tinh khiết được tinh chế từ dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen này.

Làn 1 là gen chỉ thị phân tử protein, làn 2 và làn 3 lần lượt thể hiện các mẫu điện di của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen, và enzym tinh khiết. Kết quả phân tích cho thấy trong dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (làn 2) có một băng có trọng lượng phân tử khoảng 35,6kDa được xác định từ trình tự axit amin (SEQ ID NO.3) và trình tự đánh dấu histidin của vector pET101/D-TOPO. Trong enzym tinh khiết (làn 3), mặc dù có hai băng mờ được phát hiện cùng với một băng đậm tương ứng với băng nêu trên, nhưng đã kiểm chứng được rằng protein này chính là protein enzym AR15G-90-3 (được thể hiện bằng mũi tên trên Fig.2).

8. Đo hoạt tính thủy phân CMC của AR15G-90-3.

Trước tiên, hoạt tính thủy phân CMC của protein enzym (AR15G-90-3) mã hóa bởi gen AR15G-90-3 được đánh giá bằng cách sử dụng cơ chất là CMC (carboxymethyl xenluloza, do Sigma-Aldrich Co. LLC. sản xuất).

Hoạt tính thủy phân CMC của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen thu được ở phần 7 hoặc mẫu enzym thu được ở giữa quá trình tinh chế được đo bằng cách cho hỗn hợp chứa 100 μ L dung dịch nước CMC 1% khối lượng, 50 μ L dung dịch đệm axit axetic 200mM (độ pH = 4,0), và 50 μ L dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen hoặc mẫu enzym thu được ở giữa quá trình tinh chế, phản ứng ở nhiệt độ nằm trong khoảng từ 40°C đến 90°C trong 10 đến 15 phút.

Trong toàn bộ các phép đo, hỗn hợp được điều chế bằng cách bổ sung dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) thay cho dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen hoặc enzym tinh khiết và phản ứng ở cùng điều kiện được sử dụng làm mẫu đối chứng. Hơn nữa, dung dịch cơ chất và enzym được giữ lần lượt và riêng biệt ở nhiệt độ phản ứng được duy trì trong 5 phút, sau đó trộn. Khoảng thời gian này được thiết lập để kích hoạt phản ứng. Sau khi kết thúc phản ứng, thuốc thử axit 3,5-đinitrosalixylic (dung dịch DNS) được bổ sung vào ở thể tích bằng nhau. Các hỗn hợp thu được được gia nhiệt ở nhiệt độ 100°C trong 5 phút, làm nguội trên nước đá trong 5 phút, sau đó ly tâm ở tốc độ 17.400g trong 5 phút. Kết quả là thu được phần dịch nổi. Độ hấp thụ ở bước sóng 540nm được đo bằng máy quang phổ, và hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn glucoza. Lượng đường khử tạo ra bởi quá trình thủy phân được đo bằng cách so sánh độ sai lệch với mẫu đối chứng. Hoạt tính enzym để tạo ra 1 μ mol

đường khử trong 1 phút được tính bằng 1U, và trị số thu được bằng cách chia hoạt tính này cho khối lượng protein được tính bằng hoạt tính riêng (U/mg).

9. Độ đặc hiệu cơ chất của AR15G-90-3.

Hoạt tính thủy phân các cơ chất xenluloza và hemixenluloza khác nhau của protein enzym (AR15G-90-3) mã hóa bởi gen AR15G-90-3 được đo. Trong phép đo này, dung dịch enzym tinh khiết (0,2mg/mL) thu được ở phần 7 được sử dụng. Ngoài ra, PSA, bột Avicel (bột xenluloza vi tinh thể, do Merck sản xuất), CMC (do Sigma-Aldrich Co. LLC. sản xuất), xylan (có nguồn gốc từ gỗ sồi, do Sigma-Aldrich Co. LLC. sản xuất), lichenan (do MP Biomedicals, LLC. sản xuất), laminarin (có nguồn gốc từ *Laminaria digitata*, do Sigma-Aldrich Co. LLC. sản xuất), PNPC (do Sigma-Aldrich Co. LLC. sản xuất), PNPX (do Sigma-Aldrich Co. LLC. sản xuất), và PNPG (do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm các cơ chất.

PSA được điều chế bằng cách hòa tan từ từ bột Avicel (bột xenluloza vi tinh thể, do Merck sản xuất) với dung dịch axit phosphoric, tiếp theo kết tủa Avicel bằng cách bổ sung nước tinh khiết vô trùng, sau đó rửa hỗn hợp này cho đến khi độ pH bằng 5 hoặc cao hơn. Lưu ý rằng PSA được sử dụng trong toàn bộ các thử nghiệm sau được điều chế bằng phương pháp nêu trên.

Cụ thể hơn, trước tiên phản ứng enzym được thực hiện bằng cách ủ sơ bộ hỗn hợp phản ứng chứa 50 μ L dung dịch đệm axit axetic 200mM (độ pH = 4,0) (dung dịch đệm McIlvaine (độ pH = 7,0) khi cơ chất là xylan), 5 μ L dung dịch enzym tinh khiết (0,2mg/mL) và 45 μ L nước tinh khiết ở nhiệt độ 70°C trong 5 phút, tiếp theo bổ sung thêm 100 μ L mỗi dung dịch cơ chất có nhiệt độ được duy trì theo cùng cách thức (dung dịch nước chứa 1% khối lượng PSA, bột Avicel, CMC, xylan, lichenan và laminarin, và dung dịch nước chứa 3,4mM PNPG, PNPX và PNPC) vào và ủ hỗn hợp thu được ở nhiệt độ 70°C trong 20 phút. Trong quá trình phản ứng, các hỗn hợp này được khuấy bằng máy khuấy nhiệt ở tốc độ 1.400 vòng/phút (do Eppendorf sản xuất) để tránh kết tủa các cơ chất không tan. Trong toàn bộ các phép đo, hỗn hợp được điều chế bằng cách bổ sung dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) thay cho dung dịch enzym tinh khiết và phản ứng ở cùng điều kiện được sử dụng làm mẫu đối chứng.

Khi PSA, bột Avicel, CMC, xylan, lichenan hoặc laminarin được sử dụng làm cơ chất, thì lượng đường khử tạo ra bởi quá trình thủy phân, và hoạt tính riêng (U/mg) được đo theo cùng cách thức nêu trong phần 8, trong đó hoạt tính thủy phân CMC của AR15G-90-3 được đo. Tuy nhiên, khi xylan được sử dụng làm cơ chất, thì hàm lượng đường khử được đo bằng đường chuẩn xyloza.

Khi PNPG, PNPX hoặc PNPC được sử dụng làm cơ chất, thì sau khi kết thúc phản ứng, dung dịch nước natri carbonat 200mM được bổ sung vào ở thể tích bằng nhau, và các hỗn hợp thu được được ly tâm trong 5 phút. Kết quả là thu được phần dịch nổi. Độ hấp thụ ở bước sóng 420nm được đo bằng máy quang phổ, và hàm lượng p-nitrophenol trong phần dịch nổi được đo bằng đường chuẩn p-nitrophenol. Lượng p-nitrophenol tạo ra bởi quá trình thủy phân được đo bằng cách so sánh độ sai lệch với mẫu đối chứng. Hoạt tính enzym để tạo ra 1µmol p-nitrophenol trong 1 phút được tính bằng 1U, và trị số thu được bằng cách chia hoạt tính này cho khối lượng protein được tính bằng hoạt tính riêng (U/mg).

Mỗi phép đo được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.3. Các kết quả này cho thấy AR15G-90-3 có hoạt tính thủy phân CMC, xylan, PSA, và lichenan, tuy nhiên AR15G-90-3 có hoạt tính thủy phân rất yếu Avicel, laminarin và PNPC, và hầu như không có hoạt tính thủy phân PNPG và PNPX.

10. Đánh giá mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân CMC của AR15G-90-3.

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân CMC của protein enzym AR15G-90-3 được đánh giá. Trong phép đo này, dung dịch enzym tinh khiết (0,13mg/mL) được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với dung dịch đệm Tris-HCl 50mM (độ pH =8,0) được sử dụng. Ngoài ra, CMC (carboxymetyl xenluloza, do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm cơ chất.

Cụ thể, phép đo này được thực hiện bằng cách cho hỗn hợp chứa 50µL dung dịch đệm axit axetic 200mM (độ pH = 4,0), 40µL dung dịch đệm Tris-HCl 50mM (độ pH =

8,0), 10 μ L dung dịch enzym tinh khiết (0,13mg/mL), và 100 μ L dung dịch nước CMC 1% khối lượng phản ứng ở nhiệt độ bằng 60°C, 70°C, 80°C, 90°C, 100°C, 110°C, 120°C, 130°C, 140°C, 150°C, hoặc 160°C trong 15 phút. Khi đo hoạt tính ở nhiệt độ 100°C hoặc cao hơn, thì bình thủy tinh, nút cao su, và giấy nhôm (do Nichiden-Rika Glass Co., Ltd. sản xuất) được sử dụng. Sau khi kết thúc phản ứng, lượng đường khử tạo ra bởi quá trình thủy phân được đo theo cùng cách thức nêu trong phần 8 để xác định hoạt tính thủy phân CMC (hoạt tính riêng, U/mg). Ngoài ra, hoạt tính thủy phân CMC được đo theo cùng cách thức bằng cách bổ sung 0,5mM ion canxi vào dung dịch phản ứng.

Ngoài ra, hoạt tính thủy phân CMC của endoglucanaza TrEG (*Trichoderma reesei* EG2; do Megazyme sản xuất) và endoglucanaza TmEG (*Thermotoga maritima* EG; do Megazyme sản xuất) cũng được đo theo cùng cách thức. Dung dịch đệm McIlvaine 200mM (độ pH = 4,0) và dung dịch đệm McIlvaine 200mM (độ pH = 5,0) lần lượt được sử dụng cho endoglucanaza TrEG và endoglucanaza TmEG, thay cho dung dịch đệm axit axetic 200mM, và để phản ứng ở nhiệt độ 30°C, 40°C, 50°C, 60°C, 70°C, 80°C, 90°C, hoặc 100°C.

Mỗi phép đo được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.4. Enzym AR15G-90-3 có hoạt tính phân giải CMC ở nhiệt độ nằm trong khoảng từ 60°C đến 160°C. Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) ở đó enzym này có hoạt tính mạnh nhất bằng 100°C ở độ pH bằng 4,0. Ngoài ra, khi 0,5mM ion canxi được bổ sung vào dung dịch phản ứng enzym-cơ chất, thì $T_{\text{tối ưu}}$ của enzym này tăng lên đến 110°C, và hoạt tính riêng của nó ở các nhiệt độ tương ứng cũng tăng. Trong khi đó, $T_{\text{tối ưu}}$ của endoglucanaza TrEG và endoglucanaza TmEG lần lượt bằng 60°C và 80°C. Nói cách khác, enzym AR15G-90-3 có nhiệt độ tối ưu cao hơn nhiệt độ tối ưu của endoglucanaza có nguồn gốc từ vi khuẩn ưa nhiệt *T. maritima* khoảng 20°C hoặc cao hơn.

Ngoài ra, hoạt tính phân giải CMC của enzym AR15G-90-3 ở toàn bộ các khoảng nhiệt độ cũng được cải thiện bằng cách bổ sung ion canxi. Cụ thể, ở nhiệt độ 120°C hoặc cao hơn, thì hoạt tính phân giải CMC của enzym AR15G-90-3 giảm từ từ khi có mặt ion canxi, và hoạt tính ở nhiệt độ 140°C bằng khoảng 40% hoạt tính ở nhiệt độ tối ưu khi vắng mặt ion canxi, trong khi đó hoạt tính ở nhiệt độ 160°C vẫn bằng 60% hoạt tính tối ưu khi có mặt ion canxi. Hiệu quả này thu được là do ion canxi có tác dụng làm ổn định

protein enzym.

11. Đánh giá mức độ phụ thuộc vào độ pH của hoạt tính thủy phân CMC của AR15G-90-3.

Mức độ phụ thuộc vào độ pH của hoạt tính thủy phân CMC của protein enzym AR15G-90-3 được đánh giá. Trong phép đo này, dung dịch enzym tinh khiết (0,13mg/mL) được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với dung dịch đệm Tris-HCl 50mM (độ pH =8,0) được sử dụng. Ngoài ra, CMC (carboxymetyl xenluloza, do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm cơ chất.

Cụ thể, phép đo này được thực hiện bằng cách cho hỗn hợp chứa 50 μ L dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3,0 đến 8,0), 30 μ L nước tinh khiết, 20 μ L dung dịch enzym tinh khiết (0,13mg/mL), và 100 μ L dung dịch nước CMC 1% khối lượng, phản ứng ở nhiệt độ 70°C trong 15 phút. Sau khi kết thúc phản ứng, lượng đường khử tạo ra bởi quá trình thủy phân được đo theo cùng cách thức nêu trong phần 8 để xác định hoạt tính thủy phân CMC (hoạt tính riêng, U/mg).

Mỗi phép đo được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.5. Độ pH được vẽ đồ thị bằng các trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym.

Enzym AR15G-90-3 có hoạt tính thủy phân CMC ở độ pH nằm trong khoảng từ 3,0 đến 8,0.

Độ pH tối ưu của AR15G-90-3 bằng 4,4 ở nhiệt độ 70°C (trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym).

12. Đánh giá mức độ phụ thuộc vào nhiệt độ của hoạt tính xylanaza của AR15G-90-3.

Do enzym AR15G-90-3 cũng có hoạt tính phân giải xylan (hoạt tính xylanaza), nên mức độ phụ thuộc vào nhiệt độ của hoạt tính xylanaza của protein enzym AR15G-90-3 cũng được đánh giá. Trong phép đo này, dung dịch enzym tinh khiết (0,13mg/mL) được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với dung dịch đệm

Tris-HCl 50mM (độ pH = 8,0) được sử dụng. Ngoài ra, xylan (có nguồn gốc từ gỗ sồi, do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm cơ chất.

Cụ thể, phép đo này được thực hiện bằng cách cho hỗn hợp chứa 50 μ L dung dịch đệm McIlvaine (độ pH = 7,0), 20 μ L dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), 30 μ L dung dịch enzyme tinh khiết (0,13mg/mL), và 100 μ L dung dịch nước xylan 1% khối lượng phản ứng ở nhiệt độ bằng 40°C, 50°C, 60°C, 70°C, 80°C, 90°C, 100°C, 110°C, 120°C, 130°C, hoặc 140°C trong 15 phút. Khi đo hoạt tính ở nhiệt độ 100°C hoặc cao hơn, thì bình thủy tinh, nút cao su, và giấy nhôm (do Nichiden-Rika Glass Co., Ltd. sản xuất) được sử dụng. Sau khi kết thúc phản ứng, lượng đường khử tạo ra bởi quá trình thủy phân được đo theo cùng cách thức nêu trong phần 8 để xác định hoạt tính xylanaza (hoạt tính riêng, U/mg), chỉ khác là đường chuẩn xyloza được sử dụng.

Mỗi phép đo được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.6. Enzyme AR15G-90-3 có hoạt tính xylanaza ở nhiệt độ nằm trong khoảng từ 40°C đến 140°C. Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) ở đó enzyme này có hoạt tính mạnh nhất bằng 110°C ở độ pH bằng 7,0.

13. Đánh giá mức độ phụ thuộc vào độ pH của hoạt tính xylanaza của AR15G-90-3.

Mức độ phụ thuộc vào độ pH của hoạt tính xylanaza của protein enzyme AR15G-90-3 được đánh giá. Trong phép đo này, dung dịch enzyme tinh khiết (0,10mg/mL) được điều chế bằng cách pha loãng enzyme tinh khiết thu được ở phần 7 với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) được sử dụng. Ngoài ra, CMC (carboxymetyl xenluloza, do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm cơ chất.

Cụ thể, phép đo này được thực hiện bằng cách cho hỗn hợp chứa 25 μ L dung dịch đệm Britton-Robinson (độ pH nằm trong khoảng từ 5,0 đến 10,0), 15 μ L nước tinh khiết, 10 μ L dung dịch enzyme tinh khiết (0,10mg/mL), và 50 μ L dung dịch nước xylan 1% khối lượng, phản ứng ở nhiệt độ 70°C trong 15 phút. Sau khi kết thúc phản ứng, lượng đường khử tạo ra bởi quá trình thủy phân được đo theo cùng cách thức nêu trong phần 12 để xác định hoạt tính xylanaza (hoạt tính riêng, U/mg).

Mỗi phép đo được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn.

Các kết quả đo được thể hiện trên Fig.7. Độ pH được vẽ đồ thị bằng các trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym. Enzym AR15G-90-3 có hoạt tính xylanaza ở độ pH nằm trong khoảng từ 5,0 đến 10,0. Độ pH tối ưu của AR15G-90-3 bằng 8,6 ở nhiệt độ 70°C (trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym).

14. Đánh giá độ ổn định nhiệt của AR15G-90-3 bằng cơ chất CMC và xylan.

Độ ổn định nhiệt của protein cũng được đánh giá thông qua nhiệt độ biến tính hoặc nhiệt độ nóng chảy, T_m . Nhiệt độ ủ sơ bộ ở đó hoạt tính enzym giảm 50% so với hoạt tính của lô không xử lý ủ sơ bộ được tính là nhiệt độ T_m của protein, và có thể được xác định bằng cách đo hoạt tính enzym. Nhiệt độ biến tính T_m của enzym AR15G-90-3 được đo bằng phương pháp này.

Cụ thể, phản ứng enzym được thực hiện sau khi ủ sơ bộ ở nhiệt độ nằm trong khoảng từ 50°C đến 130°C trong 30 phút, trong điều kiện không chứa cơ chất. Thành phần của dung dịch phản ứng enzym là giống như thành phần của dung dịch phản ứng nêu trong phần 10, khi CMC được sử dụng làm cơ chất, và giống như thành phần của dung dịch phản ứng nêu trong phần 12, khi xylan được sử dụng làm cơ chất, và quá trình phản ứng enzym được thực hiện trong 15 phút ở nhiệt độ 70°C. Độ ổn định nhiệt của enzym AR15G-90-3 được đánh giá bằng cách đo nhiệt độ T_{50} ở đó hoạt tính thủy phân CMC hoặc hoạt tính thủy phân xylan giảm 50%. Mỗi dữ liệu được tính gần đúng bằng cách sử dụng hàm logistic, và nhiệt độ ở đó đường cong gần đúng đạt được hoạt tính tương đối bằng 50% được xác định là T_{50} (tức là trị số T_m).

Kết quả đo hoạt tính thủy phân CMC và xylan lần lượt được thể hiện trên Fig.8 và Fig.9. Hoạt tính enzym được thể hiện dưới dạng trị số tương đối (hoạt tính tương đối, %), với giả định rằng hoạt tính của lô không xử lý ủ sơ bộ bằng 100%. Hoạt tính thủy phân CMC của AR15G-90-3 có nhiệt độ T_{50} bằng 108,6°C, và hoạt tính thủy phân xylan của AR15G-90-3 có nhiệt độ T_{50} bằng 107,6°C.

Ví dụ 2

Dựa trên dữ liệu trình tự bộ gen thu được từ mẫu đất ở khu vực suối nước nóng AR15, quá trình lắp ráp được thực hiện một lần nữa bằng phần mềm lắp ráp thế hệ mới

để thu được trình tự có chiều đầy đủ của gen AR15G-90.

1. Quá trình tái lắp ráp dữ liệu hệ gen môi trường thu được từ mẫu đất ở khu vực suối nước nóng.

Trình tự nucleotit đọc được bằng máy giải trình tự 454 ở phần 1 của Ví dụ 1 được lắp ráp bằng phần mềm Newbler phiên bản 2.7 được tích hợp trên máy giải trình tự 454 Life Sciences. Điều kiện lắp ráp được thiết lập giống như trong phần 2 của Ví dụ 1.

2. Quá trình phát hiện các khung đọc mã mở của endoglucanaza.

Các khung đọc mã mở được phát hiện từ các trình tự tiếp giáp thu được từ phần 2 theo cùng cách thức như nêu trong phần 3 của Ví dụ 1, chỉ khác là phần mềm Metagene (Noguchi et al., DNA Research, **2008**, 15(6)) được sử dụng.

3. Khung đọc mã mở AR15G-90B.

Quá trình phân loại các gen thành các họ glycosit hydrolaza được thực hiện theo cùng cách thức như nêu trong phần 4 của Ví dụ 1, và thu được khung đọc mã mở AR15G-90B dài hơn khung đọc mã mở AR15G-90 khoảng 55 gốc axit amin ở đầu tận cùng N bằng cách sử dụng phần mềm tra cứu độ tương đồng BLASTP và HMMER. Khung đọc mã mở AR15G-90B mã hóa polypeptit chứa 338 gốc axit amin và có chiều dài đầy đủ (SEQ ID NO.9), và gốc axit amin ở vị trí số 1 trên polypeptit này là metionin (M), và đầu 3' của trình tự nucleotit mã polypeptit này được kết thúc bằng bộ ba mã hóa kết thúc. Trong polypeptit nêu trên, ngoại trừ đoạn trình tự từ gốc axit amin ở vị trí số 56 là valin (V), còn đoạn trình tự còn lại là giống như trình tự của AR15G-90. Kết quả phân tích bằng phần mềm dự đoán trình tự peptit tín hiệu SignalP 4.1 cho thấy không xác định được trình tự peptit tín hiệu. Trình tự axit amin đã biết có độ tương đồng cao nhất với trình tự axit amin của polypeptit mã hóa bởi khung đọc mã mở AR15G-90B là glycosit hydrolaza GH12 (mã truy cập ngân hàng gen: ADM27292.1) có nguồn gốc từ *Ignisphaera aggregans* DSM17230 thuộc ngành *Crenarchaeota*. Do độ tương đồng giữa hai trình tự axit amin này cho toàn bộ chiều dài bằng 55% như được xác định bằng thuật toán ClustalW, nên khung đọc mã mở AR15G-90B được xác định là trình tự thể hệ mới.

Fig.10 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin (SEQ ID NO.9) của

polypeptit (AR15G-90B) mã hóa bởi khung đọc mã mở AR15G-90B, trình tự axit amin (SEQ ID NO.1) của polypeptit (AR15G-90) mã hóa bởi khung đọc mã mở AR15G-90, và trình tự axit amin của glycosit hydrolaza GH12 (SEQ ID NO.8) có nguồn gốc từ *Ignisphaera aggregans* DSM17230. Trên Fig.10, các axit amin nằm sau đen/trắng thể hiện các gốc axit amin giống hệt nhau trong toàn bộ các trình tự axit amin này, và ký tự “-” thể hiện các đoạn gạch bỏ (khoảng trống).

4. Quá trình tách dòng gen từ khung đọc mã mở AR15G-90B.

Quá trình tách dòng gen được thực hiện theo cùng cách thức như nêu trong phần 6 của Ví dụ 1, chỉ khác là đoạn môi xuôi có trình tự nucleotit như nêu trong SEQ ID NO.14 (5'-CACCATGAGTAGAAAGACAGCTGTTTACATAGCTATAGC-3' được thiết kế bằng cách bổ sung 4 nucleotit (CACC) vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13. 4 nucleotit CACC được bổ sung vào đầu 5' là trình tự để chèn vào vector) được sử dụng.

Sản phẩm tách dòng gen AR15G-90B-15 thu được từ khung đọc mã mở AR15G-90B bằng quá trình tách dòng PCR. Trình tự nucleotit của gen endoglucanaza tiềm năng AR15G-90B-15 (SEQ ID NO.12) có chiều dài bằng 1.017 bp tương đồng với trình tự nucleotit của khung đọc mã mở AR15G-90B (SEQ ID NO.10), và chỉ khác trình tự nucleotit của khung đọc mã mở AR15G-90B bởi 13 nucleotit (Bảng 2). Tất nhiên, sự khác biệt về các nucleotit ở ba vị trí cũng dẫn đến sự khác biệt về các axit amin, và trình tự axit amin của khung đọc mã mở AR15G-90B (SEQ ID NO.9) và trình tự axit amin của gen endoglucanaza tiềm năng AR15G-90B-15 (SEQ ID NO.11) khác nhau bởi 3 gốc axit amin. Tức là, gốc axit amin ở vị trí số 122 trên khung đọc mã mở AR15G-90B là serin (S), trong khi đó gốc axit amin ở vị trí số 122 trên khung đọc mã mở AR15G-90B-15 là asparagin (N); gốc axit amin ở vị trí số 273 trên khung đọc mã mở AR15G-90B là valin (V), trong khi đó gốc axit amin ở vị trí số 273 trên khung đọc mã mở AR15G-90B-15 là isoleucin (I), và gốc axit amin ở vị trí số 314 trên khung đọc mã mở AR15G-90B là metionin (M), trong khi đó gốc axit amin ở vị trí số 314 trên khung đọc mã mở AR15G-90B-15 là threonin (T).

Bảng 2

Nucleotit số	AR15G-90B	AR15G-90B-15
42	A	G
126	A	T
192	T	A
258	T	C
365	G	A
438	G	T
571	C	T
630	A	G
816	G	A
867	G	A
876	C	T
885	C	T
941	T	C

5. Quá trình biểu hiện gen và đo hoạt tính thủy phân CMC của AR15G-90B-15.

Gen *AR15G-90B-15* được biểu hiện bằng *E. coli* theo cùng cách thức như nêu trong phần 7 và phần 8 của Ví dụ 1, và hoạt tính endoglucanaza của protein enzym (AR15G-90B-15) mã hóa bởi gen AR15G-90B-15 được đo. Hoạt tính thủy phân CMC của AR15G-90B-15 được đo bằng cách thực hiện phản ứng trong 1 giờ ở nhiệt độ 50°C, 70°C, hoặc 90°C, bằng cách sử dụng phần dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen. Phép đo này được lặp lại ba lần để thu được trị số trung bình và độ lệch chuẩn. Lượng đường khử cao nhất được tính là 100%, và lượng đường khử tạo ra bởi quá trình thủy phân CMC (độ pH = 4,0) bởi protein AR15G-90B-15 ở các nhiệt độ tương ứng được xác định là hoạt tính tương đối (%), và được vẽ đồ thị dưới dạng hoạt tính thủy phân CMC tương đối (Fig.11). AR15G-90B-15 có hoạt tính thủy phân CMC ở mỗi nhiệt độ, và khi đo ba lần, thì nhiệt độ ở đó enzym này có hoạt tính cao nhất là 90°C.

YÊU CẦU BẢO HỘ

1. Endoglucanaza siêu chịu nhiệt tái tổ hợp phân lập chứa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11; và ít nhất một vùng được chọn từ nhóm bao gồm đoạn gắn kết xenluloza, gốc liên kết, peptit tín hiệu và trình tự đánh dấu.
2. Polynucleotit tái tổ hợp phân lập chứa trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.3 hoặc SEQ ID NO.11; và miền mã hóa ít nhất một vùng được chọn từ nhóm bao gồm đoạn gắn kết xenluloza, gốc liên kết, peptit tín hiệu và trình tự đánh dấu.
3. Vector biểu hiện chứa polynucleotit theo điểm 2, và có khả năng biểu hiện polypeptit có hoạt tính endoglucanaza ở tế bào vật chủ.
4. Thẻ biến nạp chứa vector biểu hiện theo điểm 3.
5. Thẻ biến nạp theo điểm 4, trong đó thẻ biến nạp này là vi sinh vật nhân thật.
6. Phương pháp sản xuất endoglucanaza siêu chịu nhiệt bao gồm bước sản sinh endoglucanaza siêu chịu nhiệt trong thẻ biến nạp theo điểm 4.
7. Hỗn hợp glycosit hydrolaza chứa endoglucanaza siêu chịu nhiệt theo điểm 1 và ít nhất một hoặc nhiều loại glycosit hydrolaza khác.
8. Hỗn hợp glycosit hydrolaza chứa endoglucanaza siêu chịu nhiệt mã hóa bởi polynucleotit theo điểm 2 và ít nhất một hoặc nhiều loại glycosit hydrolaza khác.
9. Hỗn hợp glycosit hydrolaza chứa endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp theo điểm 6 và ít nhất một hoặc nhiều loại glycosit hydrolaza khác.
10. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với endoglucanaza siêu chịu nhiệt theo điểm 1.
11. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza

tiếp xúc với endoglucanaza siêu chịu nhiệt mã hóa bởi polynucleotit theo điểm 2.

12. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với thể biến nạp theo điểm 4.

13. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với endoglucanaza siêu chịu nhiệt được sản xuất bằng phương pháp theo điểm 6.

14. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với hỗn hợp glycosit hydrolaza theo điểm 7.

15. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với hỗn hợp glycosit hydrolaza theo điểm 8.

16. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước tạo ra sản phẩm phân giải lignoxenluloza bằng cách cho nguyên liệu lignoxenluloza chứa xenluloza tiếp xúc với hỗn hợp glycosit hydrolaza theo điểm 9.

Fig.1

AR15G-90	0	-----	0
I. aggregans	1	MAVSGLQIIWILVIALALGIYIVIGGPGTKTIVTETITYTETVPRTIIQQETKTITTI	60
AR15G-90	1	-----VTITPSITTPQQLRAITLR	18
I. aggregans	61	IPTTIEGIVRETTTETIQYTTIVISIAVTMTSILSTVYQIVSQTIYTTITITSPMSIILE	120
AR15G-90	19	YPGNSS--SPVDLNSDGVYDVLLLEINPNWMEKADGFQTIWDLINGVIKVSDDLANVYPP	76
I. aggregans	121	YPKDLSTVKTLTDINGDGISDVKVAINPNWMRSAQGYQKMIINLSTRSIKFIISNLTNVSPA	180
AR15G-90	77	NWANGYPEIITIGRKPWTYYVNGLGVFPMRVRDARPFVISFYICVEHLDPSTMNFIAAD	136
I. aggregans	181	EWINGYPEIYIGRKPWDYSVNGFGVAFPMKISELKPFWSFYICIDRLDPSTMNFIAAD	240
AR15G-90	137	AWIVRESTAMSLGMPPAKGDLEIMVWLFSQNLNPAGRKIGEEVWPISINGSVIEAVFEVW	196
I. aggregans	241	AWIVRESTIARNAGASPSQGDIEIMVWLFNQNLQPAQSRIGEEILPIVINGTKYYKTFEVW	300
AR15G-90	197	RHDSVENGGWQYIAFRPKMWVRCGYIAYNPIDFVIAASKYAT-FDISDYLLGWEIGTE	255
I. aggregans	301	KMNSVPWGGWDYIAFKPKDWSIRCGSVWYDPTQFVKALSKYTDGDISNYLLDWEIGTE	360
AR15G-90	256	WGTMTSNGLAKFSWFLKDLEIQPGKTLR	283
I. aggregans	361	WGSRTSNGIAIFENTIRDFAIPGIEIK	388

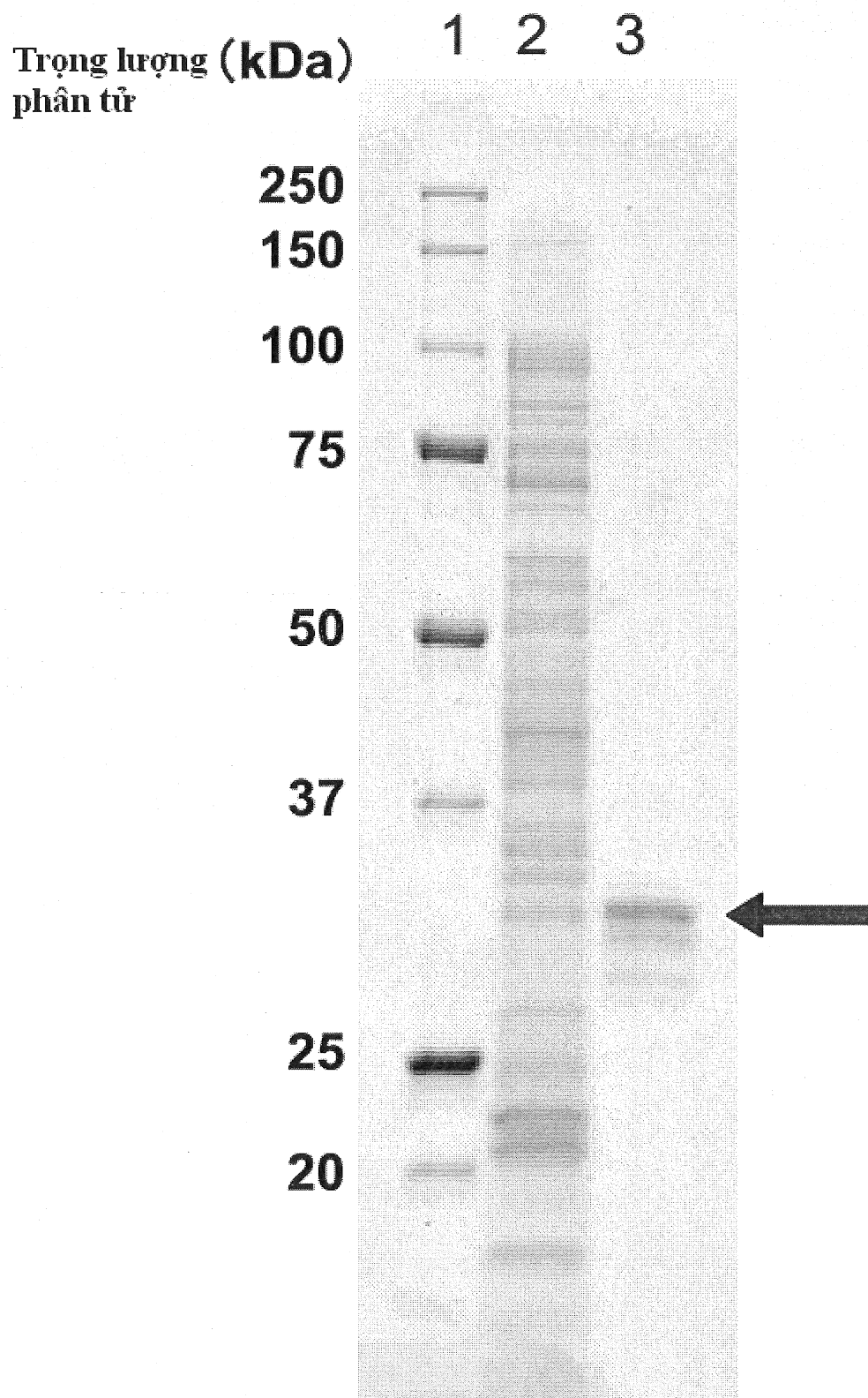
Fig.2

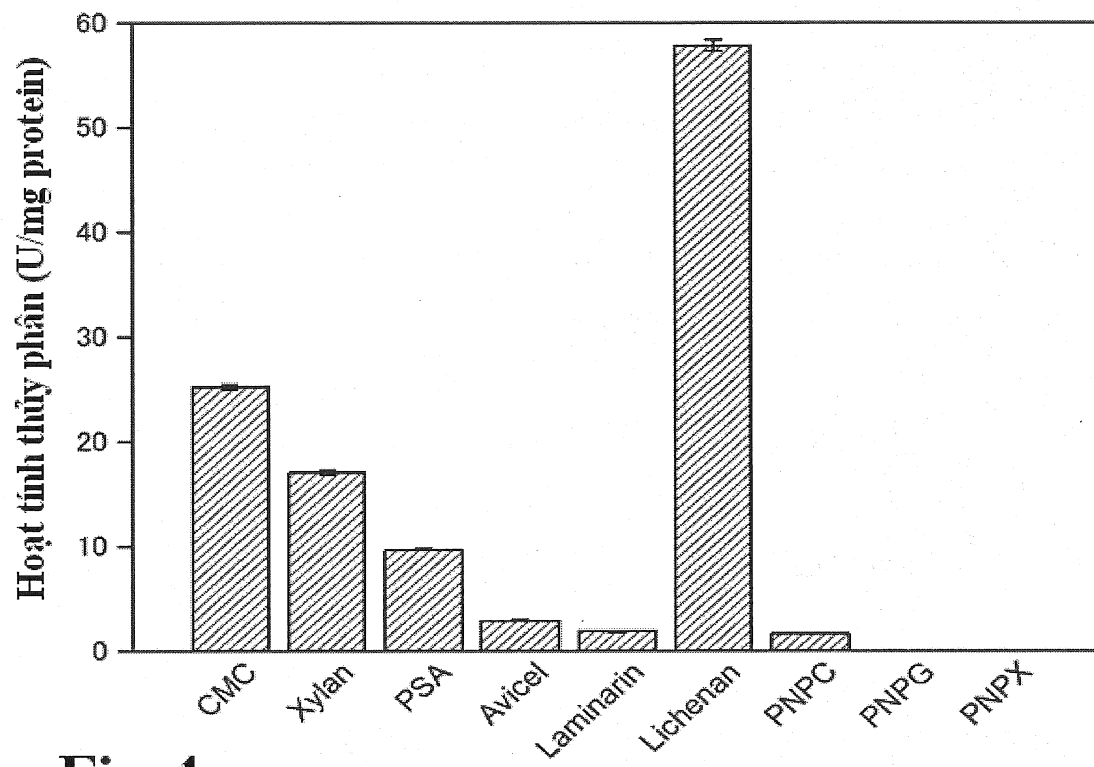
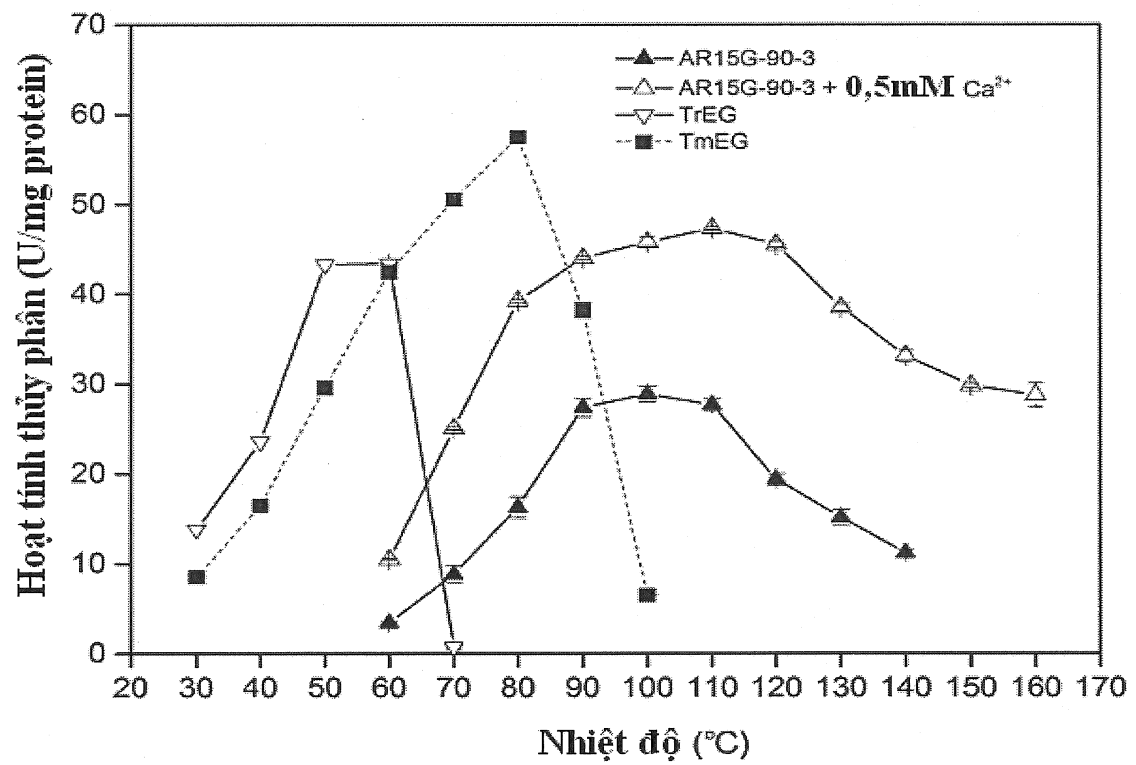
Fig.3**Fig.4**

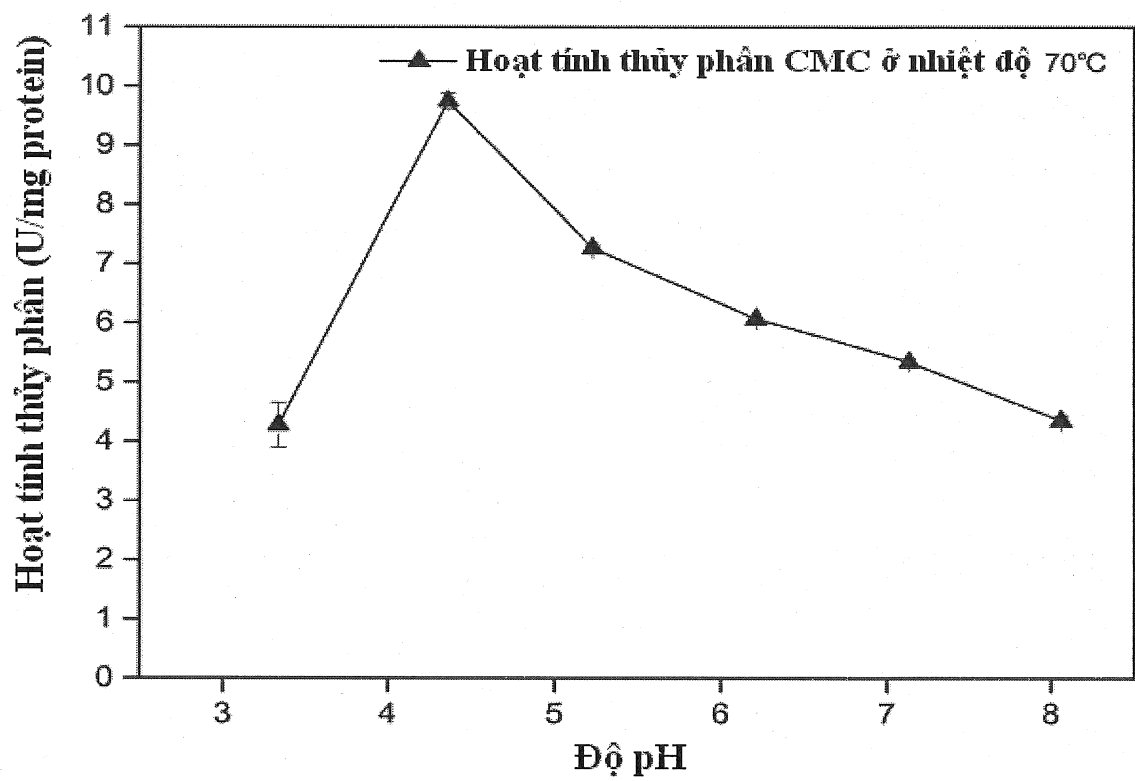
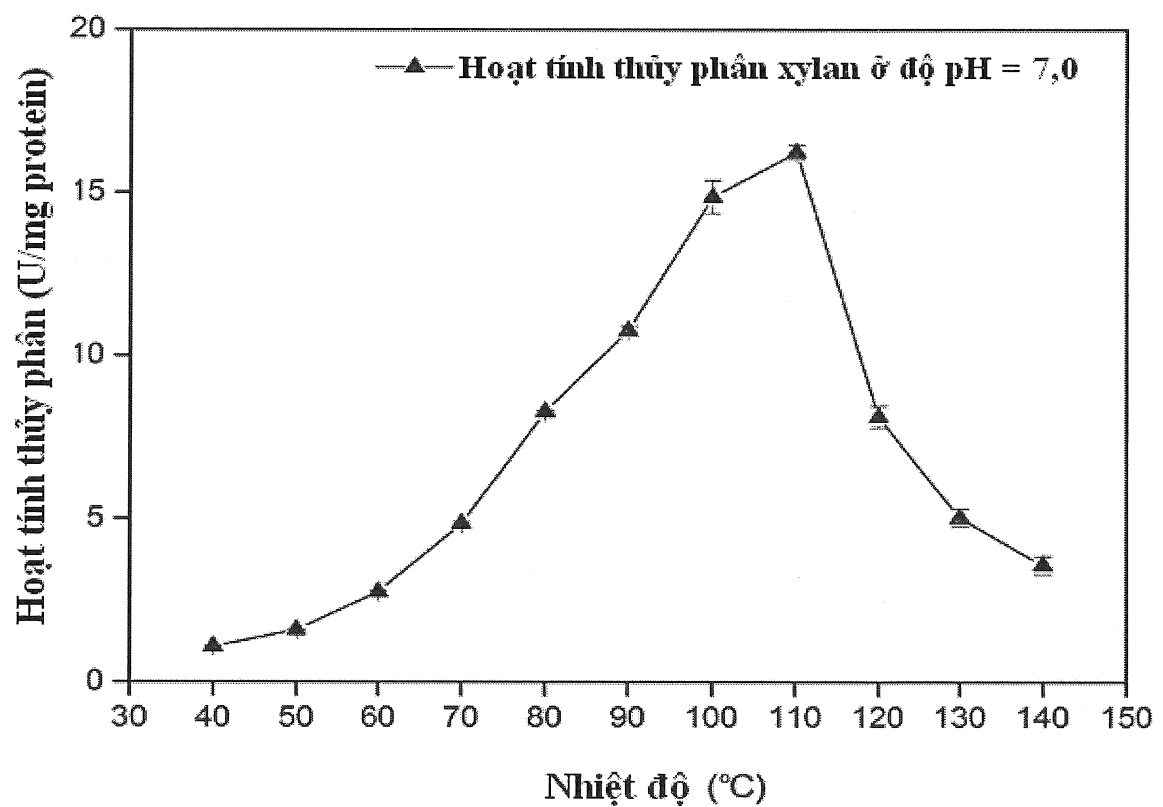
Fig.5**Fig.6**

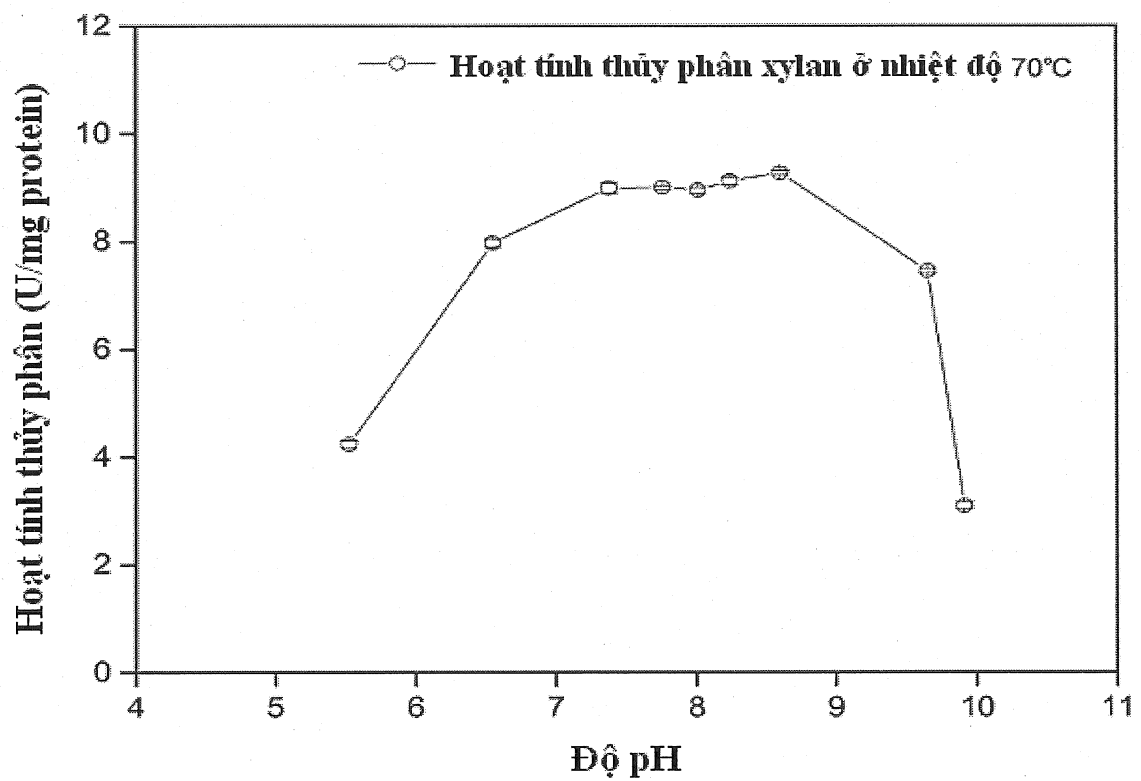
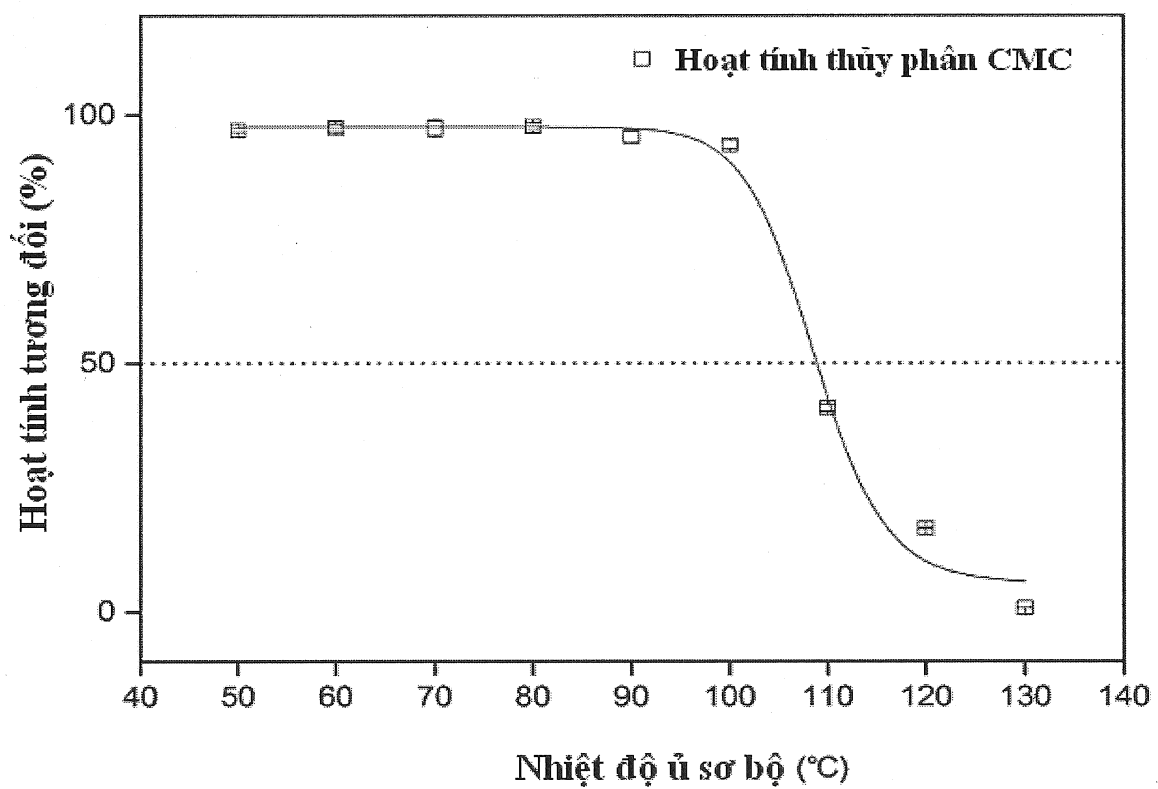
Fig.7**Fig.8**

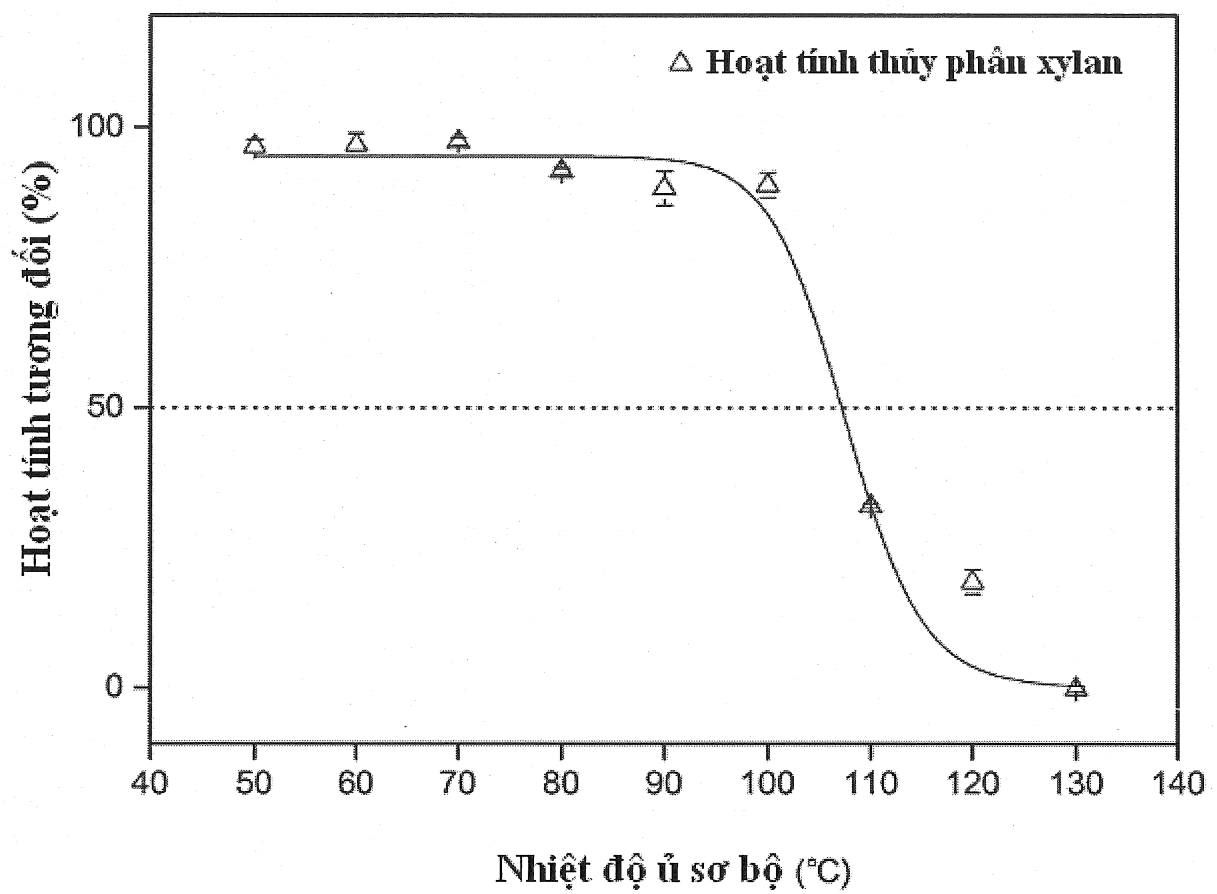
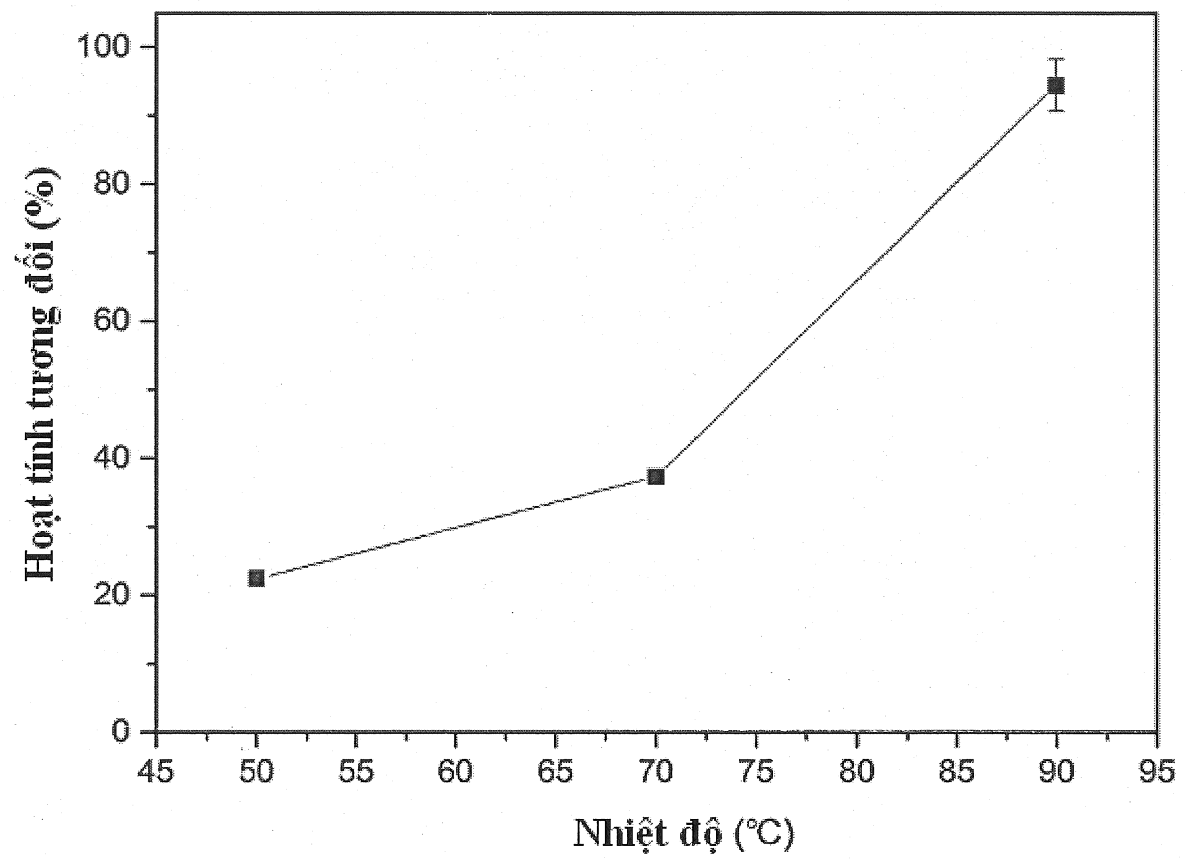
Fig.9

Fig.10

AR15G-90B 1 ---MS---RKTA**VVIAIAIA**LT**IVIGFLAG**-----**YQGVSR**-----**VEIVTIT** 37
 AR15G-90 0 -----
 I. aggregans 1 MA**VSG**LQI**IVVILVIA**IA**LALGIVIVIG**PGTKTIVTETIT**Y**TE**IVPR**IIQ**QETK**IT**ITTI** 60
 AR15G-90B 38 **VTT**-----**TAITTFYS**IT**ITIT**IT**IT**-----**PSTTPQQLRAITLR** 73
 AR15G-90 1 -----**VTT**-----**PSTTPQQLRAITLR** 18
 I. aggregans 61 **IP**TT**IEGIVRE**TT**ETIQY**IT**IVTSIA**TV**MTSIL**ST**VQVTSQ**IT**YTTIT**IT**SPMS**IT**LE** 120
 AR15G-90B 74 **Y**PGNSS--**SPVD**LS**DGVYDV**LL**LEIN**PN**MEKADG**FQ**TIWDL**ING**VIK**VSS**DLAN**VYPP 131
 AR15G-90 19 **Y**PGNSS--**SPVD**LS**DGVYDV**LL**LEIN**PN**MEKADG**FQ**TIWDL**ING**VIK**VSS**DLAN**VYPP 76
 I. aggregans 121 **Y**PK**DL**ST**VK**TL**DING**D**GISDV**K**VAIN**PN**MRSA**Q**GVQ**K**MIIN**L**STR**SI**KFI**SL**TNV**SPA 180
 AR15G-90B 132 **N**WANG**Y**PE**II**IG**RK**P**WT**TY**YNG**L**GV**EF**PM**R**VRD**AR**PF**IS**FY**IC**VE**HL**DPS**M**NF**NI**AA**D 191
 AR15G-90 77 **N**WANG**Y**PE**II**IG**RK**P**WT**TY**YNG**L**GV**EF**PM**R**VRD**AR**PF**IS**FY**IC**VE**HL**DPS**M**NF**NI**AA**D 136
 I. aggregans 181 **E**W**NG**Y**PE**IV**IG**R**K**P**WD**YS**YNG**F**GA**F**PM**K**ISE**L**K**PE**W**S**FY**IC**ID**RL**DPS**M**NF**NI**AA**D 240
 AR15G-90B 192 **A**W**TV**RE**SI**AM**SL**G**MP**PA**KGD**LE**IM**W**WLF**S**Q**N**LN**PAG**R**K**IG**EE**V**VP**IS**ING**S**VI**E**AV**FE**W 251
 AR15G-90 137 **A**W**TV**RE**SI**AM**SL**G**MP**PA**KGD**LE**IM**W**WLF**S**Q**N**LN**PAG**R**K**IG**EE**V**VP**IS**ING**S**VI**E**AV**FE**W 196
 I. aggregans 241 **A**W**TV**RE**SI**AR**NA**G**AS**PS**QGD**TE**IM**W**WLF**N**Q**N**L**Q**P**AG**S**R**IG**EE**IL**P**IV**ING**T**K**Y**Y**K**T**FE**W 300
 AR15G-90B 252 **R**HDS**V**EW**GG**W**Q**Y**IA**FR**PK**M**W**VR**CG**Y**TA**YN**PID**FT**A**ASK**YAT**-**FD**IS**DY**Y**LL**G**WE**IG**TE** 310
 AR15G-90 197 **R**HDS**V**EW**GG**W**Q**Y**IA**FR**PK**M**W**VR**CG**Y**TA**YN**PID**FT**A**ASK**YAT**-**FD**IS**DY**Y**LL**G**WE**IG**TE** 255
 I. aggregans 301 **K**M**N**SV**P**W**GG**W**D**Y**IA**FK**PK**D**W**SI**RC**GS**W**Y**DP**T**Q**F**V**K**AL**SK**Y**T**D**G**ID**IS**W**Y**LL**D**WE**IG**TE** 360
 AR15G-90B 311 **W**G**T**M**T**S**N**G**L**A**K**F**S**W**FL**K**D**LE**IQ**P**G**K**T**LR 338
 AR15G-90 256 **W**G**T**M**T**S**N**G**L**A**K**F**S**W**FL**K**D**LE**IQ**P**G**K**T**LR 283
 I. aggregans 361 **W**G**S**R**T**S**N**G**IA**I**F**E**W**T**IR**D**F**MA**IP**G**IE**IK 388

Fig.11

DANH MỤC TRÌNH TỰ

<110> HONDA MOTOR CO., LTD.
 <120> ENDOGLUCANAZA SIÊU CHỊU NHIỆT, POLYNUCLEOTIT MÃ HÓA ENZYM NÀY, VECTO BIỂU HIỆN ENZYM NÀY, THỂ BIẾN NẠP CHỨA VECTO NÀY, PHƯƠNG PHÁP SẢN XUẤT ENZYM NÀY, HỖN HỢP GLYCOSIT HYDROLAZA CHỨA ENZYM NÀY, VÀ PHƯƠNG PHÁP SẢN XUẤT SẢN PHẨM PHÂN GIẢI LIGNOXENLULOZA
 <130> H114098801
 <160> 14
 <170> PatentIn version 3.1
 <210> 1
 <211> 283
 <212> PRT
 <213> Trình tự chưa biết
 <220>
 <223> Trình tự axit amin mã hóa bởi khung đọc mã mở AR15G-90
 <400> 1
 Val Thr Ile Thr Pro Ser Thr Thr Pro Gln Gln Leu Arg Ala Ile Thr
 1 5 10 15
 Leu Arg Tyr Pro Gly Asn Ser Ser Ser Pro Val Asp Leu Asn Ser Asp
 20 25 30
 Gly Val Tyr Asp Val Leu Leu Glu Ile Asn Pro Trp Asn Met Glu Lys
 35 40 45
 Ala Asp Gly Phe Gln Thr Ile Val Val Asp Leu Ile Asn Gly Val Ile
 50 55 60
 Lys Val Ser Ser Asp Leu Ala Asn Val Tyr Pro Pro Asn Trp Ala Asn
 65 70 75 80
 Gly Tyr Pro Glu Ile Ile Ile Gly Arg Lys Pro Trp Thr Thr Tyr Tyr
 85 90 95
 Tyr Asn Gly Leu Gly Val Glu Phe Pro Met Arg Val Arg Asp Ala Arg
 100 105 110
 Pro Phe Val Ile Ser Phe Tyr Ile Cys Val Glu His Leu Asp Pro Ser
 115 120 125
 Met Asn Phe Asn Ile Ala Ala Asp Ala Trp Ile Val Arg Glu Ser Ile
 130 135 140
 Ala Met Ser Leu Gly Met Pro Pro Ala Lys Gly Asp Leu Glu Ile Met
 145 150 155 160
 Val Trp Leu Phe Ser Gln Asn Leu Asn Pro Ala Gly Arg Lys Ile Gly
 165 170 175
 Glu Glu Val Val Pro Ile Ser Ile Asn Gly Ser Val Ile Glu Ala Val
 180 185 190
 Phe Glu Val Trp Arg His Asp Ser Val Glu Trp Gly Gly Trp Gln Tyr
 195 200 205
 Ile Ala Phe Arg Pro Lys Met Trp Gly Val Arg Cys Gly Tyr Ile Ala
 210 215 220
 Tyr Asn Pro Ile Asp Phe Val Thr Ala Ala Ser Lys Tyr Ala Thr Phe
 225 230 235 240
 Asp Ile Ser Asp Tyr Tyr Leu Leu Gly Trp Glu Ile Gly Thr Glu Trp
 245 250 255
 Gly Thr Met Thr Ser Asn Gly Leu Ala Lys Phe Ser Trp Phe Leu Lys
 260 265 270
 Asp Leu Glu Ile Gln Pro Gly Lys Thr Leu Arg
 275 280
 <210> 2
 <211> 852
 <212> ADN
 <213> Trình tự chưa biết
 <220> □@
 <223> Khung đọc mã mở AR15G-90
 <400> 2
 gtgaccatca cgccgagtag aacacctcaa cagctcagag ctataacctt gagatacccc 60
 ggcaactctt ctgcgcctgt agatctcaac agtgatgggtg tttagcatgt tcttttagag 120
 ataaacccgt ggaatatgga gaaagccgac ggatttcaga caatagttgt agacctaatt 180
 aacgggtgta ttaaggtgag ctctgatcta gccaatgtat acccaccgaa ctgggctaatt 240
 ggctaccctg agataatcat cggtagaaag ccgtggacta cgtactacta caacggctta 300
 ggcgtagaat ttcctatgag agtacgcgac gctagacctt ttgtaatctc gttctatata 360
 tgcgtagagc acctagaccc atcgatgaac ttcaacatag cggctgacgc ctggattgtt 420
 agagaaagca ttgcgatgag cttgggtatg ccgcctgcta aaggagatct agagataatg 480
 gtgtggctct tcagccaaaa ccttaaccca gctggttagga aaatcgggga ggaagtagta 540
 cccatatcta tcaacgggag tggtatagaa gctgttttcg aggtctggag acatgatagc 600
 gttgagtggg gtgggtggca gtatatagcg tttaggccta agatgtgggg ggtaggtgc 660
 ggctatatag cgtacaaccc tatagacttc gtgactgcgg cgtctaagta cgctacattc 720
 gatatatctg actactacct cttaggctgg gagataggga ctgagtgggg cactatgaca 780
 tctaattggtc tggctaagtt ctcgtggttt ttgaaagatc tcgagattca gccaggtaaa 840
 aactgaggt ag 852
 <210> 3
 <211> 283
 <212> PRT
 <213> Trình tự chưa biết

<220>
 <223> AR15G-90-3
 <400> 3
 Val Thr Ile Thr Pro Ser Thr Thr Pro Gln Gln Leu Arg Ala Ile Thr
 1 5 10 15
 Leu Arg Tyr Pro Gly Asn Ser Ser Ser Pro Val Asp Leu Asn Ser Asp
 20 25 30
 Gly Val Tyr Asp Val Leu Leu Glu Ile Asn Pro Trp Asn Met Glu Lys
 35 40 45
 Ala Asp Gly Phe Gln Thr Ile Val Val Asp Leu Ile Asn Gly Val Ile
 50 55 60
 Lys Val Ser Ser Asp Leu Ala Asn Val Tyr Pro Pro Asn Trp Ala Asn
 65 70 75 80
 Gly Tyr Pro Glu Ile Ile Ile Gly Arg Lys Pro Trp Thr Thr Tyr Tyr
 85 90 95
 Tyr Asn Gly Leu Gly Val Glu Phe Pro Met Arg Val Arg Asp Ala Lys
 100 105 110
 Pro Phe Val Ile Ser Phe Tyr Ile Cys Val Glu His Leu Asp Pro Ser
 115 120 125
 Met Asn Phe Asn Ile Ala Ala Asp Ala Trp Ile Val Arg Glu Ser Ile
 130 135 140
 Ala Met Ser Leu Gly Met Pro Pro Ala Lys Gly Asp Leu Glu Ile Met
 145 150 155 160
 Val Trp Leu Phe Ser Gln Asn Leu Asn Pro Ala Gly Arg Lys Ile Gly
 165 170 175
 Glu Glu Val Val Pro Ile Ser Ile Asn Gly Ser Val Ile Glu Ala Val
 180 185 190
 Phe Glu Val Trp Arg His Asp Ser Val Glu Trp Gly Gly Trp Gln Tyr
 195 200 205
 Ile Ala Phe Arg Pro Lys Met Trp Gly Ile Arg Cys Gly Tyr Ile Ala
 210 215 220
 Tyr Asn Pro Ile Asp Phe Val Thr Ala Ala Ser Lys Tyr Ala Thr Phe
 225 230 235 240
 Asp Ile Ser Asp Tyr Tyr Leu Leu Gly Trp Glu Ile Gly Thr Glu Trp
 245 250 255
 Gly Thr Met Thr Ser Asn Gly Leu Ala Lys Phe Ser Trp Phe Leu Lys
 260 265 270
 Asp Leu Glu Ile Gln Pro Gly Lys Thr Leu Arg
 275 280

<210> 4
 <211> 852
 <212> ADN
 <213> Trình tự chưa biết

<220>
 <223> AR15G-90-3
 <400> 4
 gtgaccatca cgccgagtag aacacctcaa cagctcagag ctataacctt gagatacccc 60
 ggcaactcctt cttcgccctgt agatctcaac agcgatgggtg tttacgatgt tcttttagag 120
 ataaaccggt ggaatatgga gaaagccgac ggatttcaga caatagttgt agacctaatt 180
 aacgggtgta ttaaggtgag ctctgatcta gccaatgtat acccaccgaa ctgggctaatt 240
 ggctaccctg agataatcat cggtagaaag ccgtgggacta cgtactacta caacggctta 300
 ggcgtagaat ttcctatgag agtacgcgac gctaaacctt ttgtaatctc gttctatata 360
 tgcgtagagc acctagaccc atcgatgaac ttcaacatag cggctgacgc ctggattgtt 420
 agagaaagca ttgcgatgag cttgggtatg ccgcctgcta aaggagatct agagataatg 480
 gtgtggtctt tcagccaaaa ccttaaccca gctggttaga aaatcgggga ggaagtagta 540
 cccatatcta tcaacgggag tggttatagaa gctgttttcg aggtctggag acatgatagc 600
 gttgagtggg gtgggtggca gtatatagcg tttaggccta agatgtgggg gattaggtgc 660
 ggctatatag cgtacaaccc tatagacttc gtgactgcag cgtctaagta cgctacattc 720
 gatatatctg actactacct cttaggctgg gagataggga ctgagtgggg cactatgaca 780
 tctaattggtc tggctaagtt ctcgtggttt ttgaaagatc tcgagattca gccaggtaaa 840
 acactgaggt ag 852

<210> 5
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Mô tả trình tự nhân tạo: đoạn mỗi.
 <400>

gtgaccatca cgccgagtag a 21

<210> 6
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Mô tả trình tự nhân tạo: đoạn mỗi.
 <400> 6

ctacctcagtggtt ttacctggc

19

<210> 7
 <211> 28
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Mô tả trình tự nhân tạo: đoạn mỗi.
 <400> 7
 caccatggtg accatcacgc cgagtaca 28
 <210> 8
 <211> 388
 <212> PRT
 <213> Ignisphaera aggregans DSM17230
 <220>
 <223> Họ glycosit hydrolaza 12
 <400> 8
 Met Ala Val Ser Gly Leu Gln Ile Ile Val Val Ile Leu Val Ile Ala
 1 5 10 15
 Leu Leu Ala Leu Gly Ile Tyr Ile Val Ile Gly Gly Pro Gly Thr Lys
 20 25 30
 Thr Ile Val Thr Glu Thr Ile Thr Tyr Thr Glu Thr Val Pro Arg Thr
 35 40 45
 Ile Ile Gln Gln Glu Thr Lys Thr Ile Thr Thr Ile Ile Pro Thr Thr
 50 55 60
 Ile Glu Gly Ile Val Arg Glu Thr Thr Thr Glu Thr Gln Tyr Thr Thr
 65 70 75 80
 Val Thr Ser Ile Ala Thr Val Thr Met Thr Ser Ile Leu Thr Ser Thr
 85 90 95
 Val Tyr Gln Thr Val Ser Gln Thr Ile Tyr Thr Thr Thr Ile Thr
 100 105 110
 Ser Pro Met Ser Ile Ile Leu Glu Tyr Pro Lys Asp Leu Ser Thr Val
 115 120 125
 Lys Thr Leu Asp Ile Asn Gly Asp Gly Ile Ser Asp Val Lys Val Ala
 130 135 140
 Ile Asn Pro Trp Asn Met Arg Ser Ala Gln Gly Tyr Gln Lys Met Ile
 145 150 155 160
 Ile Asn Leu Ser Thr Arg Ser Ile Lys Phe Ile Ser Asn Leu Thr Asn
 165 170 175
 Val Ser Pro Ala Glu Trp Thr Asn Gly Tyr Pro Glu Ile Tyr Ile Gly
 180 185 190
 Arg Lys Pro Trp Asp Tyr Ser Tyr Val Asn Gly Phe Gly Val Ala Phe
 195 200 205
 Pro Met Lys Ile Ser Glu Leu Lys Pro Phe Val Val Ser Phe Tyr Ile
 210 215 220
 Cys Ile Asp Arg Leu Asp Pro Ser Met Asn Phe Asn Ile Ala Ala Asp
 225 230 235 240
 Ala Trp Ile Val Arg Glu Ser Ile Ala Arg Asn Ala Gly Ala Ser Pro
 245 250 255
 Ser Gln Gly Asp Ile Glu Ile Met Val Trp Leu Phe Asn Gln Asn Leu
 260 265 270
 Gln Pro Ala Gly Ser Arg Ile Gly Glu Glu Ile Leu Pro Ile Val Ile
 275 280 285
 Asn Gly Thr Lys Tyr Tyr Lys Thr Phe Glu Val Trp Lys Met Asn Ser
 290 295 300
 Val Pro Trp Gly Gly Trp Asp Tyr Ile Ala Phe Lys Pro Lys Asp Trp
 305 310 315 320
 Ser Ile Arg Cys Gly Ser Val Val Tyr Asp Pro Thr Gln Phe Val Lys
 325 330 335
 Ala Leu Ser Lys Tyr Thr Asp Gly Ile Asp Ile Ser Asn Tyr Tyr Leu
 340 345 350
 Leu Asp Trp Glu Ile Gly Thr Glu Trp Gly Ser Arg Thr Ser Asn Gly
 355 360 365
 Ile Ala Ile Phe Glu Trp Thr Ile Arg Asp Phe Met Ala Ile Pro Gly
 370 375 380
 Ile Glu Ile Lys
 385
 <210> 9
 <211> 338
 <212> PRT
 <213> Trình tự chưa biết
 <220>
 <223> Trình tự axit amin mã hóa bởi khung đọc mã mở AR15G-90B
 <400> 9
 Met Ser Arg Lys Thr Ala Val Tyr Ile Ala Ile Ala Leu Leu Thr Leu
 1 5 10 15
 Val Ile Gly Phe Leu Ala Gly Tyr Gln Tyr Gly Val Ser Arg Val Glu
 20 25 30
 Thr Val Thr Thr Thr Val Thr Thr Ala Thr Thr Thr Phe Tyr Ser
 35 40 45
 Thr Thr Thr Ile Thr Thr Thr Val Thr Ile Thr Pro Ser Thr Thr Pro

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

Leu Arg

<210> 10

<211> 1017

<212> ADN

<213> Trình tự chưa biết

<220> ☐@

<223> Khung đọc mã mở AR15G-90B

<400> 10

atgagtagaa agacagctgt ttacatagct atagctctgc taacccttgt aataggcttt 60
 ctacgaggtt accagtagcg tgtagcaga gtagagactg taaccacaac tgttacaaca 120
 actgcaacca caacgtttta tagcactaca actataacca ctactgtgac catcacgccc 180
 agtacaacac ctcaacagct cagagctata acctgagat accccggcaa ctcttcttcg 240
 cctgtagatc tcaacagtga tgggtgtttac gatgttcttt tagagataaa cccgtggaat 300
 atggagaaag ccgacggatt tcagacaata gttgtagacc taattaacgg tgttattaag 360
 gtgagctctg atctagccaa tgtataccca ccgaactggg ctaatggcta ccctgagata 420
 atcatcggtg gaaagccgtg gactacgtac tactacaacg gcttaggcgt agaatttcct 480
 atgagagtag cgcagcgtag accttttgta atctcgttct atatatgctg agagcaccta 540
 gacccatcga tgaacttcaa catagcggct gacgcctgga ttgttagaga aagcattgctg 600
 atgagcttgg gtatgccgcc tgctaaagga gatctagaga taatgggtgtg gctcttcagc 660
 caaaacctta acccagctgg taggaaaatc ggggaggaag tagtaccat atctatcaac 720
 gggagtgtta tagaagctgt ttctgaggtc tggagacatg atagcgttga gtggggtggg 780
 tggcagtata tagcgtttag gcctaagatg tgggggggta ggtgcggcta tatagcgtac 840
 aaccctatag acttcgtgac tgcggcgtct aagtacgcta cattcgatat atctgactac 900
 tacctcttag gctgggagat agggactgag tggggcacta tgacatctaa tggctcggct 960
 aagttctcgt ggtttttgaa agatctcgag attcagccag gtaaaacact gaggttag 1017

<210> 11

<211> 338

<212> PRT

<213> Trình tự chưa biết

<220>

<223> AR15G-90B-15

<400> 11

Met Ser Arg Lys Thr Ala Val Tyr Ile Ala Ile Ala Leu Leu Thr Leu
 1 5 10 15
 Val Ile Gly Phe Leu Ala Gly Tyr Gln Tyr Gly Val Ser Arg Val Glu
 20 25 30
 Thr Val Thr Thr Thr Val Thr Thr Thr Ala Thr Thr Thr Phe Tyr Ser
 35 40 45
 Thr Thr Thr Ile Thr Thr Thr Val Thr Ile Thr Pro Ser Thr Thr Pro
 50 55 60
 Gln Gln Leu Arg Ala Ile Thr Leu Arg Tyr Pro Gly Asn Ser Ser Ser
 65 70 75 80
 Pro Val Asp Leu Asn Ser Asp Gly Val Tyr Asp Val Leu Leu Glu Ile
 85 90 95

Asn Pro Trp Asn Met Glu Lys Ala Asp Gly Phe Gln Thr Ile Val Val
 100 105 110
 Asp Leu Ile Asn Gly Val Ile Lys Val Asn Ser Asp Leu Ala Asn Val
 115 120 125
 Tyr Pro Pro Asn Trp Ala Asn Gly Tyr Pro Glu Ile Ile Ile Gly Arg
 130 135 140
 Lys Pro Trp Thr Thr Tyr Tyr Tyr Asn Gly Leu Gly Val Glu Phe Pro
 145 150 155 160
 Met Arg Val Arg Asp Ala Arg Pro Phe Val Ile Ser Phe Tyr Ile Cys
 165 170 175
 Val Glu His Leu Asp Pro Ser Met Asn Phe Asn Ile Ala Ala Asp Ala
 180 185 190
 Trp Ile Val Arg Glu Ser Ile Ala Met Ser Leu Gly Met Pro Pro Ala
 195 200 205
 Lys Gly Asp Leu Glu Ile Met Val Trp Leu Phe Ser Gln Asn Leu Asn
 210 215 220
 Pro Ala Gly Arg Lys Ile Gly Glu Glu Val Val Pro Ile Ser Ile Asn
 225 230 235 240
 Gly Ser Val Ile Glu Ala Val Phe Glu Val Trp Arg His Asp Ser Val
 245 250 255
 Glu Trp Gly Gly Trp Gln Tyr Ile Ala Phe Arg Pro Lys Met Trp Gly
 260 265 270
 Ile Arg Cys Gly Tyr Ile Ala Tyr Asn Pro Ile Asp Phe Val Thr Ala
 275 280 285
 Ala Ser Lys Tyr Ala Thr Phe Asp Ile Ser Asp Tyr Tyr Leu Leu Gly
 290 295 300
 Trp Glu Ile Gly Thr Glu Trp Gly Thr Thr Thr Ser Asn Gly Leu Ala
 305 310 315 320
 Lys Phe Ser Trp Phe Leu Lys Asp Leu Glu Ile Gln Pro Gly Lys Thr
 325 330 335

Leu Arg

<210> 12

<211> 1017

<212> ADN

<213> Trình tự chưa biết

<220> □@

<223> AR15G-90B-15

<400> 12

atgagtagaa agacagctgt ttacatagct atagctctgc tgacccttgt aataggcttt	60
ctagcaggtt accagtagcg tgtagcaga gtagagactg taaccacaac tggtacaaca	120
actgctacca caacgtttta tagcactaca actataacca ctactgtgac catcacgccg	180
agtacaacac cacaacagct cagagctata accttgagat accccggcaa ctcttcttcg	240
cctgtagatc tcaacagcga tgggtgtttac gatgttcttt tagagataaa cccgtggaat	300
atggagaaa cgcagcgatt tcagacaata gttgtagacc taattaacgg tggtattaag	360
gtgaactctg atctagccaa tgtataccca ccgaactggg ctaatggcta ccctgagata	420
atcatcggtg gaaagccttg gactacgtac tactacaacg gcttaggcgt agaatttcct	480
atgagagtac gcgacgctag accttttgta atctcgttct atatatgcgt agagcaccta	540
gacccatcga tgaacttcaa tatagcggct gacgcctgga ttggttagaga aagcattgcg	600
atgagcttgg gtatgccgccc tgctaaaggg gatctagaga taatgggtgt gctcttcagc	660
caaaacctta acccagctgg taggaaaatc ggggaggaag tagtaccat atctatcaac	720
gggagtgtta tagaagctgt ttctgaggtc tggagacatg atagcgttga gtgggggtgg	780
tggcagtata tagcgttttag gcctaagatg tgggggatta ggtgcggcta tatagcgtac	840
aaccctatag acttcgtgac tgcggcatct aagtatgcta catttgatat atctgactac	900
tacctcttag gctgggagat agggactgag tggggcacta cgacatctaa tggctctggct	960
aagttctcgt ggtttttgaa agatctcgag attcagccag gtaaaacact gaggtag	1017

<210> 13

<211> 35

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn mỗi.

<400> 13

atgagtagaa agacagctgt ttacatagct atagc

35

<210> 14

<211> 39

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn mỗi.

<400> 14

caccatgagt agaaagacag ctgtttacat agctatagc

39