



(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)



1-0028052

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

(13) B

(21) 1-2015-03397

(22) 15/09/2015

(30) 2014-189007 17/09/2014 JP

(45) 25/04/2021 397

(43) 25/03/2016 336A

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

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

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

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

(54) ENZYM β -XYLOSIDAZA CHỊU NHIỆT, HỖN HỢP GLYCOSIT HYĐROLAZA CHỨA ENZYM NÀY, VÀ PHƯƠNG PHÁP SẢN XUẤT SẢN PHẨM PHÂN GIẢI LIGNOXENLULOZA

(57) Sáng chế đề cập đến enzym β -xylosidaza chịu nhiệt, 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 enzym β -xylosidaza 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), và lignin. 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 để sản xuất nhiều sản phẩm 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, 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, cấu trúc của hemixenluloza lại thay đổi phụ thuộc vào loài thực vật, ví dụ xylan là thành phần chính có trong thực vật gỗ cứng và thực vật thân thảo. Quá trình thủy phân xylan cần có xylanaza (endo-1,4- β -xylanaza, EC 3.2.1.8) và enzym β -xylosidaza (EC 3.2.1.37). Enzym β -xylosidaza là một enzym tham gia vào quá trình thủy phân oligosacarit được tạo ra trong quá trình thủy phân hemixenluloza bằng xylanaza để tạo ra monosacarit.

Trong quy trình thông thường để biến đổi lignoxenluloza thành etanol, cơ chất ban đầu với nồng độ chất rắn cao nằm trong khoảng từ 30% đến 60% đã được thử nghiệm nhằm đạt được hiệu suất năng lượng cao hơn. Tuy nhiên, khi nồng độ chất rắn trong cơ chất cao, thì độ nhớt của hỗn dịch sinh khối đặc cũng lớn và phản ứng thủy phân lignoxenluloza hầu như không thể xảy ra. Do đó, bằng cách sử dụng enzym chịu nhiệt và thực hiện phản ứng thủy phân ở nhiệt độ 80°C hoặc cao hơn, thì tốc độ phản ứng thủy phân có thể được nâng cao, và độ nhớt của hỗn dịch sinh khối đặc cũng được giảm xuống, nhờ vậy có thể rút ngắn được thời gian thủy phân và giảm lượng enzym cần sử dụng. Do đó, trong số các glycosit hydrolaza khác nhau, cần 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, và xác định trình tự ADN của chúng, sau đó biểu hiện chúng bằng *E. coli*, nấm sợi hoặc các sinh vật tương tự. Ví dụ, Tài liệu sáng chế 1 đề cập đến enzym β -xylosidaza có nguồn gốc từ nấm sợi. Tài liệu sáng chế 2 đề cập đến enzym β -xylosidaza có nguồn gốc từ nấm sợi *Aspergillus oryzae* có hoạt tính enzym ở nhiệt độ 30°C. Tài liệu sáng chế 3 đề cập đến enzym β -xylosidaza có nguồn gốc từ *Alicyclobacillus acidocaldarius* có hoạt tính enzym ở độ pH = 5,5 hoặc thấp hơn và nhiệt độ 50°C hoặc cao hơn. Tài liệu sáng chế 4 đề cập đến enzym β -xylosidaza có nguồn gốc từ *Acremonium cellulolyticus* có hoạt tính enzym ở nhiệt độ 45°C. Ngoài ra, Tài liệu phi sáng chế 1 đến 6 đề cập đến enzym β -xylosidaza được phân lập từ vi khuẩn và nấm sợi đặc trưng có nhiệt độ tối ưu khoảng 60°C..

Tài liệu viện dẫn

Tài liệu sáng chế 1: Bản dịch tiếng Nhật đã công bố số Hei 11-507837 của Đơn PCT.

Tài liệu sáng chế 2: Đơn Công bố Patent Nhật Bản số Hei 11-313683.

Tài liệu sáng chế 3: Bản dịch tiếng Nhật đã công bố số 2011-523346 của Đơn PCT.

Tài liệu sáng chế 4: Đơn Công bố Patent Nhật Bản số 2013-59272.

Tài liệu phi sáng chế 1: Kormelink *et al.*, *Journal of Biotechnology*, 1993, vol. 27, p. 249-265.

Tài liệu phi sáng chế 2: Herrmann *et al.*, *Biochemical Journal*, 1997, vol. 321, p. 375-381.

Tài liệu phi sáng chế 3: Kitamoto *et al.*, *Applied and Environmental Microbiology*, 1999, vol. 65, p. 20-24.

Tài liệu phi sáng chế 4: La Grange *et al.*, *Applied and Environmental Microbiology*, 2001, vol. 67, p. 5512-5519.

Tài liệu phi sáng chế 5: Shao *et al.*, *Applied and Environmental Microbiology*, 2011, vol. 77, p. 719-726.

Tài liệu phi sáng chế 6: Morais *et al.*, *Journal of Biological Chemistry*, 2012, vol. 287, p. 9213-9221.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề cập đến enzym β -xylosidaza chịu nhiệt thể hệ mới có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit (sau đây gọi tắt là PNPX) ít nhất ở nhiệt độ bằng 80°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 enzym β -xylosidaza chịu nhiệt có trình tự axit amin mới bằng cách chiết xuất 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.

Cụ thể, enzym β -xylosidaza chịu nhiệt, polynucleotit, vector biểu hiện, thể biến nạp, phương pháp sản xuất enzym β -xylosidaza 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.

Phương án 1: Enzym β -xylosidaza chịu nhiệt, chứa vùng xúc tác β -xylosidaza gồm:

(A) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5,

(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.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0, hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng bằng 80% hoặc cao hơn so với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0.

Phương án 2: Enzym β -xylosidaza chịu nhiệt theo phương án 1, trong đó enzym này còn có hoạt tính β -glucosidaza.

Phương án 3: Polynucleotit, chứa miền mã hóa vùng xúc tác β -xylosidaza gồm:

(a) trình tự nucleotit mã hóa polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5,

(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.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0,

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng bằng 80% hoặc cao hơn so với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0,

(d) trình tự nucleotit có độ tương đồng bằng 80% hoặc cao hơn so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6, và mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt

độ bằng 80°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.2, SEQ ID NO.4 hoặc SEQ ID NO.6 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0.

Phương án 4: Polynucleotit theo phương án 3, trong đó polypeptit này còn có hoạt tính β -glucosidaza.

Phương án 5: Vector biểu hiện, chứa polynucleotit theo phương án 3 hoặc 4, trong đó vector này có khả năng biểu hiện polypeptit có hoạt tính enzym β -xylosidaza ở tế bào vật chủ.

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

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

Phương án 8: Phương pháp sản xuất enzym β -xylosidaza chịu nhiệt, bao gồm bước tạo enzym β -xylosidaza chịu nhiệt trong thể biến nạp theo phương án 6 hoặc 7.

Phương án 9: Hỗn hợp glycosit hydrolaza, chứa enzym β -xylosidaza chịu nhiệt theo phương án 1 hoặc 2, enzym β -xylosidaza chịu nhiệt mã hóa bởi polynucleotit theo phương án 3 hoặc 4, hoặc enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 8, và ít nhất một glycosit hydrolaza khác.

Phương án 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, hemixenluloza và lignin tiếp xúc với enzym β -xylosidaza chịu nhiệt theo phương án 1 hoặc 2, enzym β -xylosidaza chịu nhiệt mã hóa bởi polynucleotit theo phương án 3 hoặc 4, thể biến nạp theo phương án 6 hoặc 7, enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 8, hoặc hỗn hợp glycosit hydrolaza theo phương án 9.

Enzym β -xylosidaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất

p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0. Do đó, enzym β -xylosidaza chịu nhiệt này thích hợp để thủy phân nguyên liệu lignoxenluloza ở nhiệt độ cao, ví dụ enzym β -xylosidaza chịu nhiệt này thích hợp để thủy phân nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza.

Ngoài ra, enzym β -xylosidaza chịu nhiệt theo sáng chế cũng thích hợp để thủy phân hợp chất có liên kết β -xylosidic ở nhiệt độ cao, đặc biệt là nguyên liệu chứa oligosacarit có liên kết β -xylosidic.

Mặt khác, enzym β -xylosidaza chịu nhiệt theo sáng chế cũng thích hợp để thủy phân hợp chất có liên kết β -glycosidic ở nhiệt độ cao, đặc biệt là nguyên liệu chứa oligosacarit có liên kết β -glycosidic ở nhiệt độ cao.

Ngoài ra, polynucleotit theo sáng chế, vector biểu hiện chứa polynucleotit này và thể biến nạp chứa vector này cũng thích hợp để sản xuất enzym β -xylosidaza 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 sản phẩm tách dòng gen AR19M-311-2 và trình tự axit amin của enzym β -xylosidaza GH3 (SEQ ID NO.10) có nguồn gốc từ *Dictyoglomus thermophilum*.

Fig.2 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin (SEQ ID NO.1) của khung đọc mở AR19M-311, trình tự axit amin (SEQ ID NO.3) của protein AR19M-311-2, và trình tự axit amin (SEQ ID NO.5) của protein AR19M-311-11.

Fig.3A là hình ảnh thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein AR19M-311-2 thu được bằng cách biểu hiện gen AR19M-311-2 ở *E. coli* trong ví dụ 1 (hình ảnh bên trái).

Fig.3B là hình ảnh thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein AR19M-311-11 thu được bằng cách biểu hiện gen AR19M-311-11 ở *E. coli* (hình ảnh bên phải) trong ví dụ 1.

Fig.4 là đồ thị thể hiện kết quả đo hoạt tính thủy phân các cơ chất khác nhau của

protein AR19M-311-2 được biểu hiện ở *E. coli* trong ví dụ 1.

Fig.5 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở độ pH = 4,0) của protein AR19M-311-2 được biểu hiện ở *E. coli* ở các nhiệt độ khác nhau trong ví dụ 1.

Fig.6 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở nhiệt độ 70°C) của protein AR19M-311-2 được biểu hiện ở *E. coli* ở các độ pH khác nhau trong ví dụ 1.

Fig.7A là đồ thị thể hiện sự thay đổi cường độ huỳnh quang của thuốc nhuộm màu cam SYPRO do sự biến tính nhiệt của protein AR19M-311-2 được biểu hiện ở *E. coli* trong ví dụ 1.

Fig.7B là đồ thị thể hiện đạo hàm bậc nhất “ $-d(\text{Cường độ huỳnh quang})/dt$ ” của cường độ huỳnh quang của thuốc nhuộm màu cam SYPRO do sự biến tính nhiệt của protein AR19M-311-2 được biểu hiện ở *E. coli* trong ví dụ 1.

Mô tả chi tiết sáng chế

Enzym β -xylosidaza 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 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 ít 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, nhờ sự phát triển của máy giải trình tự thế hệ mới cho phép giải trình tự được hàng tỷ cặp bazơ, nên đã có thể giải trình tự toàn bộ hệ 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 xuất, trong đó ADN hệ 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. Phương pháp này góp phần rút ngắn thời gian giải trình tự bộ gen của vi sinh vật khó

nuôi cấy.

Như được thể hiện trong ví dụ 1, các tác giả sáng chế đã tổng hợp được ADN hệ 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 406 khung đọc mở (Open Reading Frame - ORF) mã hóa các trình tự axit amin tương đồng với các β -glucosidaza hoặc enzym β -xylosidaza đã biết (tức là, các trình tự axit amin có trị số kỳ vọng (trị số E) nhỏ hơn $1e^{-20}$). Đối với mỗi khung đọc mở có chiều dài đầy đủ, mà từ đó vùng xúc tác β -glucosidaza hoặc vùng xúc tác β -xylosidaza được xác định, đoạn môi được thiết kế dựa trên thông tin trình tự nucleotit của khung đọc mở này, 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 đã tách dòng PCR được biến nạp vào *E. coli*, và các protein mã hóa bởi các trình tự nucleotit này được biểu hiện và sàng lọc chức năng bằng cách đo hoạt tính phân giải PNPX. Cuối cùng, thu được hai enzym β -xylosidaza chịu nhiệt có hoạt tính phân giải PNPX là AR19M-311-2 và AR19M-311-11 từ các khung đọc mở này. Trình tự axit amin của AR19M-311-2 và trình tự nucleotit mã hóa trình tự axit amin của AR19M-311-2 lần lượt như nêu trong SEQ ID NO.3 và SEQ ID NO.4, trình tự axit amin của AR19M-311-11 và trình tự nucleotit mã hóa trình tự axit amin của AR19M-311-11 lần lượt như nêu trong SEQ ID NO.5 và SEQ ID NO.6.

Như được thể hiện trong Ví dụ 1, AR19M-311-2 có hoạt tính thủy phân mạnh PNPX, và enzym này cũng có hoạt tính thủy phân p-nitrophenyl- β -D-glucopyranosit (PNPG). Trái lại, AR19M-311-2 hầu như không có hoạt tính phân giải Avicel tương tự như liên hợp axit phosphoric (sau đây gọi tắt là PSA), xenluloza tinh thể Avicel, carboxymetyl xenluloza (sau đây gọi tắt là CMC), laminarin chứa β -1,3 glucan và β -1,6 glucan, lichenan chứa β -1,3 glucan và β -1,4 glucan, và xylan. Từ kết quả phân tích độ đặc hiệu cơ chất này chứng tỏ rằng AR19M-311-2 là glycosit hydrolaza có hoạt tính enzym β -xylosidaza.

Thuật ngữ “hoạt tính enzym β -xylosidaza” được dùng trong bản mô tả để chỉ hoạt

tính thủy phân cơ chất PNPX.

Ngoài ra, thuật ngữ “hoạt tính β -glucosidaza” được dùng trong bản mô tả để chỉ hoạt tính thủy phân cơ chất PNPG.

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

Do đó, thuật ngữ “có hoạt tính enzym β -xylosidaza” được dùng trong bản mô tả để chỉ hoạt tính thủy phân ít nhất một cơ chất PNPX và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử hoặc cường độ phản ứng tạo màu của cơ chất PNPX đã thủy phân so với mẫu đối chứng âm.

Tương tự, thuật ngữ “có hoạt tính β -glucosidaza” được dùng trong bản mô tả để chỉ hoạt tính thủy phân ít nhất một cơ chất PNPG và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử hoặc cường độ phản ứng tạo màu của cơ chất PNPG đã thủy phân so với mẫu đối chứng âm.

Ngoài ra, thuật ngữ “có hoạt tính enzym β -xylosidaza” cũng được dùng trong bản mô tả để chỉ hoạt tính thủy phân ít nhất một hợp chất có liên kết β -xylosidic, và tốt hơn nếu hợp chất này là cơ chất chứa oligosacarit có liên kết β -xylosidic và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử hoặc cường độ phản ứng tạo màu của cơ chất đã thủy phân so với mẫu đối chứng âm.

Ngoài ra, thuật ngữ “có hoạt tính enzym β -xylosidaza” cũng được dùng trong bản mô tả để chỉ hoạt tính thủy phân ít nhất một hợp chất có liên kết β -glucosidic, và tốt hơn nếu hợp chất này là cơ chất chứa oligosacarit có liên kết β -glucosidic và tạo ra sự khác biệt đáng kể về hàm lượng đầu khử hoặc cường độ phản ứng tạo màu của cơ chất đã thủy phân so với mẫu đối chứng âm.

Kết quả tra cứu trình tự axit amin của AR19M-311-2 bằng cơ sở dữ liệu trình tự axit amin đã biết cho thấy trình tự axit amin của AR19M-311-2 có độ tương đồng cao nhất với trình tự axit amin của enzym β -xylosidaza GH3 (SEQ ID NO.10) (mã truy cập

Ngân hàng Gen: YP_002250567.1) có nguồn gốc từ *Dictyoglomus thermophilum*, và bằng 70%. Các phân tích độ đặc hiệu cơ chất, và độ tương đồng của trình tự axit amin này với các protein đã biết cho thấy AR19M-311-2 là enzym β -xylosidaza GH3 thế hệ mới.

AR19M-311-2 có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit (tức là hoạt tính enzym β -xylosidaza) ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0. Cụ thể, như được thể hiện trong phần 10 của ví dụ 1, ở độ pH bằng 4,0, AR19M-311-2 có hoạt tính enzym β -xylosidaza ở khoảng nhiệt độ rộng nằm trong khoảng từ 30°C đến 90°C. Hoạt tính enzym β -xylosidaza của AR19M-311-2 được biểu hiện bởi E.coli tăng khi nhiệt độ tăng nằm trong khoảng từ 30°C đến 80°C, nhưng sau đó thì giảm khi nhiệt độ tăng nằm trong khoảng từ 80°C đến 100°C. Ngoài ra, hoạt tính phân giải PNPX của AR19M-311-2 có nhiệt độ tối ưu ($T_{\text{tối ưu}}$) = 80°C và nhiệt độ nóng chảy T_m = 82,7°C, do đó enzym β -xylosidaza chịu nhiệt AR19M-311-2 có độ ổn định nhiệt cao hơn các enzym β -xylosidaza đã biết. Độ pH tối ưu cho hoạt tính phân giải PNPX của AR19M-311-2 là pH = 4,0, tuy nhiên enzym này có hoạt tính enzym β -xylosidaza ở khoảng pH rộng nằm trong khoảng từ 3,5 đến 8,0.

Nhìn chung, ở protein có một số hoạt tính sinh học, một 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 AR19M-311-2, một 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 nó, bao gồm hoạt tính enzym β -xylosidaza.

Enzym β -xylosidaza theo sáng chế là glycosit hydrolaza chịu nhiệt chứa vùng xúc tác β -xylosidaza gồm polypeptit bất kỳ trong số các polypeptit từ (A) đến (C) sau:

(A) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5,

(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.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0, hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng bằng 80% hoặc cao hơn so với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0.

Thuật ngữ “polypeptit trong đó có axit amin được loại bỏ” được dùng trong bản mô tả để chỉ đoạn axit amin mà tạo ra polypeptit bị mất (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.1, SEQ ID NO.3 hoặc SEQ ID NO.5 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.

Trong polypeptit (C), không có giới hạn cụ thể về độ tương đồng với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5 miễn là nằm trong khoảng từ 80% đến 100%, tốt hơn nếu nằm trong khoảng từ 85% đến 100%, tốt hơn nữa nếu nằm trong khoảng từ 90% đến 100%, còn tốt hơn nữa nếu nằm trong khoảng từ 95% đến 100%, và tốt nhất nữa nếu nằm trong khoảng từ 98% đến 100%.

Độ tương đồng trình tự (độ tương đồng) giữa cặp trình tự axit amin được xác định bằng cách đặt hai trình tự axit amin cạnh nhau sao cho các axit amin tương ứng là giống nhau, đồng thời khoảng trống được chèn vào các vị trí xuất hiện đột biến chèn thêm và loại bỏ, 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 AR19M-311-2, 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 được tổng hợp nhân tạo dựa trên polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5 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 p-nitrophenyl- β -D-xylopyranosit (hoạt tính enzym β -xylosidaza) ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0. Do đó, khi enzym chứa vùng xúc tác β -xylosidaza là polypeptit bất kỳ từ (A) đến (C), thì có thể thu được enzym β -xylosidaza chịu nhiệt.

Enzym β -xylosidaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất PNPG. Enzym β -xylosidaza chịu nhiệt này cũng có hoạt tính thủy phân cơ chất β -glucan, oligosacarit hoặc các cơ chất khác. Ví dụ về các cơ chất khác của enzym β -xylosidaza chịu nhiệt theo sáng chế bao gồm PNPG, 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 -1,4 như lichenan; xenluloza tinh thể, như Avicel, xenluloza vi tinh thể có nguồn gốc từ vi khuẩn (BMCC) và giấy lọc, xenluloza phi tinh thể như Avicel tương ứng liên hợp axit phosphoric (PSA), CMC, glucan chứa liên kết β -1,4, oligosacarit chứa liên kết β -1,4 như xenlobioza; 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. Ngoài PNPG, thì tốt hơn nếu enzym β -xylosidaza chịu nhiệt theo sáng chế cũng có hoạt tính thủy phân ít nhất một cơ chất được chọn từ nhóm bao gồm PNPG, 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, và p-nitrophenyl- β -D-galactopyranosit, và tốt hơn nữa nếu enzym β -xylosidaza chịu nhiệt theo sáng chế cũng có hoạt tính thủy phân cơ chất PNPG.

Tốt hơn nếu enzym β -xylosidaza chịu nhiệt theo sáng chế có hoạt tính phân giải cơ chất PNPX (hoạt tính enzym β -xylosidaza) ít nhất ở độ pH bằng 4,0 và nhiệt độ nằm trong khoảng từ 70°C đến 85°C, tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 60°C đến 85°C, còn tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 30°C đến 90°C. Tốt hơn nếu nhiệt độ tối ưu của enzym β -xylosidaza chịu nhiệt theo sáng chế nằm trong khoảng từ 70°C đến 85°C và tốt hơn nữa nằm trong khoảng từ 75°C đến 85°C.

Mặc dù thay đổi phụ thuộc vào nhiệt độ phản ứng, nhưng độ pH tối ưu của enzym β -xylosidaza chịu nhiệt theo sáng chế nằm trong khoảng từ 4,0 đến 6,0. Tốt hơn nếu enzym β -xylosidaza chịu nhiệt theo sáng chế có hoạt tính enzym β -xylosidaza ít nhất ở độ pH nằm trong khoảng từ 4,0 đến 8,0.

Enzym β -xylosidaza chịu nhiệt theo sáng chế cũng có thể có hoạt tính glycosit hydrolaza khác với hoạt tính enzym β -xylosidaza. Ví dụ về các hoạt tính glycosit hydrolaza này bao gồm hoạt tính endoglucanaza, hoạt tính xylanaza, hoạt tính β -glucosidaza và hoạt tính xenlobiohydrolaza. Ngoài hoạt tính enzym β -xylosidaza, tốt hơn nếu enzym β -xylosidaza theo sáng chế cũng có hoạt tính β -glucosidaza.

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

Khi enzym β -xylosidaza chịu nhiệt theo sáng chế chứa một hoặc nhiều vùng khác với vùng xúc tác β -xylosidaza, thì tốt hơn nếu enzym này chứa vùng Fibronectin loại III.

Vùng Fibronectin loại III có thể nằm trước (ở đầu N) hoặc nằm sau (ở đầu C) vùng xúc tác β -xylosidaza. Ngoài ra, vùng Fibronectin loại III và vùng xúc tác β -xylosidaza có thể được liên kết trực tiếp với nhau hoặc thông qua phần tử liên kết có chiều dài thích hợp. Tốt hơn nếu enzym β -xylosidaza chịu nhiệt theo sáng chế chứa vùng Fibronectin loại III nằm trước hoặc nằm sau vùng xúc tác β -xylosidaza thông qua phần tử liên kết, và tốt hơn nữa nếu enzym β -xylosidaza chịu nhiệt theo sáng chế chứa vùng Fibronectin loại

III nằm sau vùng xúc tác β -xylosidaza thông qua phân tử liên kết.

Ngoài ra, enzym β -xylosidaza chịu nhiệt theo sáng chế có thể cũng chứa peptit tín hiệu có khả năng 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 N hoặc đầu C. Ví dụ về các peptit tín hiệu này bao gồm peptit tín hiệu vận chuyển không bào, peptit tín hiệu duy trì mạng lưới nội chất, peptit tín hiệu vận chuyển nhân, hoặc peptit tín hiệu bài tiết. Ví dụ cụ thể về peptit tín hiệu duy trì mạng lưới nội chất bao gồm peptit tín hiệu chứa trình tự axit amin HDEL. Khi chứa peptit tín hiệu ở đầu N hoặc đầu C, thì enzym β -xylosidaza chịu nhiệt theo sáng chế 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, enzym β -xylosidaza chịu nhiệt theo sáng chế cũng có thể chứa các trình tự đánh dấu khác nhau ở đầu N hoặc đầu 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 HA (hemagglutinin), 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 theo sáng chế.

Polynucleotit mã hóa enzym β -xylosidaza chịu nhiệt theo sáng chế

Polynucleotit nêu trên mã hóa enzym β -xylosidaza chịu nhiệt theo sáng chế. Enzym β -xylosidaza chịu nhiệt theo sáng chế có thể được sản xuất bằng hệ biểu hiện của vật chủ được tạo ra bằng cách chèn vector biểu hiện chứa polynucleotit này vào vật chủ.

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

(a) trình tự nucleotit mã hóa polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5,

(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.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0,

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng bằng 80% hoặc cao hơn so với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3 hoặc SEQ ID NO.5, và có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0,

(d) trình tự nucleotit có độ tương đồng bằng 80% hoặc cao hơn so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6, và mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°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.2, SEQ ID NO.4 hoặc SEQ ID NO.6 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°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 mất (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 thuật ngữ “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 $\times$ (thành phần của dung dịch SSC 20 $\times$: natri clorua 3M, dung dịch axit xitric 0,3M, độ pH bằng 7,0), dung dịch Denhardt 5 $\times$ (thành phần dung dịch Denhardt 100 $\times$: 2% theo khối lượng albumin huyết thanh bò, 2% theo khối lượng ficoll, 2% theo khối lượng polyvinylpyrrolidon), 0,5% theo 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 $\times$ chứa 0,1% theo khối lượng SDS được sử dụng làm dung dịch đệm rửa, và tốt hơn nữa nếu dung dịch SSC 0,1 $\times$

chứa 0,1% theo 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.2, SEQ ID NO.4 hoặc SEQ ID NO.6, 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.2, SEQ ID NO.4 hoặc SEQ ID NO.6. 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.2, SEQ ID NO.4 hoặc SEQ ID NO.6 có thể được tổng hợp hóa học dựa trên thông tin về trình tự nucleotit hoặc polynucleotit này có thể được tổng hợp từ tự nhiên bằng kỹ thuật tái tổ hợp gen dưới dạng gen mã hóa AR19M-311-2 có chiều dài đầy đủ (cũng được gọi là “gen AR19M-311-2” hoặc “sản phẩm tách dòng gen AR19M-311-2”) hoặc một đoạn gen của nó chứa vùng xúc tác β -xylosidaza (ví dụ, vùng mã hóa đoạn peptit chứa 325 gốc axit amin từ threonin (T) ở vị trí số 18 đến alanin (A) ở vị trí số 342 trên trình tự SEQ ID NO.3, vùng mã hóa đoạn peptit chứa 236 gốc axit amin từ isoleucin (I) ở vị trí số 381 đến threonin (T) ở vị trí số 616 của trình tự SEQ ID NO.3, hoặc vùng mã hóa đoạn peptit chứa 599 gốc axit amin từ threonin (T) ở vị trí số 18 đến threonin (T) ở vị trí số 616 của trình tự SEQ ID NO.3). Gen AR19M-311-2 có chiều dài đầy đủ hoặc một đoạn gen của nó 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 này, cùng với đoạn mồi xuôi và đoạn mồi ngược được thiết kế dựa trên trình tự nucleoit như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6. ADN bổ sung (ADNbs) được tổng hợp bằng phản ứng phiên mã ngược từ khuôn là ARN thông tin thu được từ mẫu. 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.2, SEQ ID NO.4 hoặc SEQ ID NO.6 miễn là nằm trong khoảng từ 80% đến 100%, tốt hơn nếu nằm trong khoảng từ 85% đến 100%, tốt hơn nữa nếu nằm trong khoảng từ 90% đến 100%, và còn tốt hơn nữa 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 cách đặt hai hai trình tự nucleotit cạnh nhau sao cho các trình tự nucleotit tương ứng là giống nhau, đồng thời khoảng trống được chèn vào ở các vị trí xuất hiện đột biến chèn thêm hoặc loại bỏ; 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 chứa trình tự nucleotit (b), (c), hoặc (d) có thể được tổng hợp nhân tạo bằng cách loại bỏ, thay thế, hoặc bổ sung một hoặc nhiều nucleotit trong polynucleotit có trình tự như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6. 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 AR19M-311-2 hoặc một đoạn trình tự của nó. Gen tương đồng của gen AR19M-311-2 có thể được tạo ra 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 β -xylosidaza, hoặc ngoài miền mã hóa vùng xúc tác β -xylosidaza, thì polynucleotit theo sáng chế cũng có thể chứa một hoặc nhiều miền mã hóa đoạn gắn kết xenluloza, phần tử liên kết, các petít tín hiệu khác nhau, các trình tự đánh dấu khác nhau, hoặc các trình tự tương tự.

Vector biểu hiện theo sáng chế

Vector nêu trên chứa polynucleotit theo sáng chế và có khả năng biểu hiện polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- β -D-xylopyranosit ít nhất ở nhiệt độ bằng 80°C và độ pH bằng 4,0 ở tế bào vật chủ. Cụ thể, vector biểu hiện này chứa polynucleotit mã hóa enzym β -xylosidaza chịu nhiệt theo sáng chế. Đặc biệt, vector biểu hiện này cần được kết hợp với cat-xet biểu hiện chứa, theo thứ tự xuôi dòng từ trước ra sau, ADN có trình tự vùng khởi đầu (promoter), 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, theo thứ tự xuôi dòng từ trước ra sau, ADN có trình tự vùng khởi đầu, ADN có trình tự để kết hợp ADN ngoại sinh, 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 nữa, ngoài polynucleotit theo sáng chế, thì vector biểu hiện theo sáng chế còn chứa gen kháng thuốc hoặc gen tương tự. Điều này cho phép dễ dàng sàng lọc được tế bào vật chủ đã chuyển gen và tế bào vật chủ chưa chuyển gen bằng vector này.

Ví dụ về các gen kháng thuốc bao gồm gen kháng kanamycin, gen kháng hygromycin, và gen kháng bialaphos.

Thể biến nạp theo sáng chế

Thể biến nạp nêu trên chứa vector biểu hiện theo sáng chế. Enzym β -xylosidaza chịu nhiệt theo sáng chế có thể được biểu hiện ở thể 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, thể biến nạp theo sáng chế là *E. coli*, nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú, hoặc tế bào thực vật chứa vector biểu hiện theo sáng chế.

Enzym β -xylosidaza chịu nhiệt theo sáng chế có thể được sản xuất một cách đơn giản trên quy mô lớn bằng cách nuôi cấy thể biến nạp *E. coli* này. Mặt khác, do protein bị glycosyl hóa ở tế bào nhân thật, nên enzym β -xylosidaza chịu nhiệt được tạo ra bằng tế bào nhân thật sẽ có độ ổn định nhiệt cao hơn enzym β -xylosidaza chịu nhiệt được tạo ra bằng 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 dùng súng gia tốc hạt, phương pháp xung điện, phương pháp chuyển gen nhờ polyetylen glycol. Khi vật chủ là tế bào thực vật, tốt hơn nếu thể biến nạp theo sáng chế được sản xuất bằng phương pháp dùng súng gia tốc 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 giống như phương pháp nuôi cấy vật chủ chưa chuyển gen.

Phương pháp sản xuất enzym β -xylosidaza chịu nhiệt theo sáng chế

Phương pháp nêu trên là phương pháp để sản xuất enzym β -xylosidaza 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 đã được biến nạp nằm sau vùng khởi đầu không có khả năng điều hòa thời gian biểu hiện hoặc các yếu tố tương tự, thì enzym β -xylosidaza 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 enzym β -xylosidaza 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.

Enzym β -xylosidaza chịu nhiệt được sản xuất bằng thể biến nạp có thể được sử dụng ở trạng thái nằm trong thể biến nạp, hoặc có thể được chiết xuất và tinh chế từ thể biến nạp.

Không có giới hạn cụ thể về phương pháp chiết xuất và tinh chế enzym β -xylosidaza chịu nhiệt từ thể biến nạp miễn là phương pháp này không làm giảm hoạt tính của enzym β -xylosidaza chịu nhiệt, 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 enzym β -xylosidaza 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. Enzym β -xylosidaza 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 enzym β -xylosidaza 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 enzym β -xylosidaza 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 dịch nuôi cấy thu được. Hơn nữa, khi enzym β -xylosidaza chịu nhiệt theo sáng chế có trình tự đánh dấu như trình tự đánh dấu His, thì enzym β -xylosidaza 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 enzym β -xylosidaza chịu nhiệt theo sáng chế bao gồm các bước nuôi cấy thể biến nạp theo sáng chế, sản xuất enzym β -xylosidaza chịu nhiệt trong thể biến nạp này, sau đó là bước chiết xuất và tinh chế enzym β -xylosidaza chịu nhiệt từ thể biến nạp (khi cần).

Hỗn hợp glycosit hydrolaza theo sáng chế

Hỗn hợp glycosit hydrolaza nêu trên chứa enzym β -xylosidaza chịu nhiệt theo sáng chế, hoặc enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym β -xylosidaza chịu nhiệt theo sáng chế, và ít nhất một glycosit hydrolaza khác. Enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym β -xylosidaza 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 và 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 nguyên liệu lignoxenluloza.

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

Tốt hơn nếu glycosit hydrolaza khác chứa trong hỗn hợp glycosit hydrolaza theo sáng chế là enzym chịu nhiệt có hoạt tính ít nhất ở nhiệt độ bằng 85°C, và tốt hơn nữa nếu glycosit hydrolaza này là enzym 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 (tức là, các enzym có nhiệt độ tối ưu cho hoạt tính enzym hoặc nhiệt độ nóng chảy cho protein enzym bằng 70°C hoặc cao hơn), thì có thể phân giải hiệu quả nguyên liệu 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 nguyên liệu lignoxenluloza, thì có thể thực hiện phản ứng thủy phân lignoxenluloza ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C. 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, bao gồm bước thủy phân oligosacarit được tạo ra bằng cách thủy phân hemixenluloza với xylanaza

hoặc oligosacarit được tạo ra bằng cách thủy phân xenluloza với xenlobiohydrolaza, thành monosacarit với enzym β -xylosidaza chịu nhiệt theo sáng chế.

Thuật ngữ “sản phẩm phân giải lignoxenluloza” được dùng trong bản mô tả để chỉ sản phẩm được tạo ra bằng cách phân cắt liên kết β -xylosidic hoặc liên kết β -glycosidic trong nguyên liệu lignoxenluloza, ví dụ nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, và tốt hơn nếu nguyên liệu chứa oligosacarit có liên kết β -xylosidic hoặc oligosacarit có liên kết β -glycosidic. Ví dụ cụ thể về các sản phẩm phân giải lignoxenluloza bao gồm monosacarit, như glucoza, xyloza và các đường tương tự.

Cụ thể, sản phẩm phân giải hemixenluloza hoặc xenluloza, và đặc biệt là sản phẩm được tạo ra bằng cách phân cắt liên kết β -xylosidic hoặc liên kết β -glycosidic, được sản xuất bằng cách cho nguyên liệu lignoxenluloza, ví dụ nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, và tốt hơn nếu nguyên liệu chứa oligosacarit có liên kết β -xylosidic hoặc oligosacarit có liên kết β -glycosidic tiếp xúc với enzym β -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym β -xylosidaza chịu nhiệt theo sáng chế, hoặc hỗn hợp glycosit hydrolaza theo sáng chế.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế, bao gồm bước tạo sản phẩm (ví dụ, monosacarit, như glucoza, xyloza và các đường tương tự) chủ yếu bằng cách phân cắt vị trí gắn kết β -xylosit của hemixenluloza, và tốt hơn nếu oligosacarit có liên kết β -xylosidic, hoặc bằng cách phân cắt vị trí gắn kết β -glycosit của xenluloza, tốt hơn nếu oligosacarit có liên kết β -glycosidic, bằng cách cho nguyên liệu lignoxenluloza, ví dụ nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, và tốt hơn nếu nguyên liệu chứa oligosacarit có liên kết β -xylosidic hoặc oligosacarit có liên kết β -glycosidic tiếp xúc với enzym β -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym β -xylosidaza chịu nhiệt theo sáng chế.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế, bao gồm bước tạo sản phẩm (ví dụ, monosacarit, như

glucoza, xyloza và các đường tương tự) bằng cách phân cắt vị trí gắn kết β -glycosit của hemixenluloza hoặc xenluloza, bằng cách cho nguyên liệu lignoxenluloza, ví dụ nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza tiếp xúc với hỗn hợp glycosit hydrolaza theo sáng chế.

Không có giới hạn cụ thể về nguyên liệu lignoxenluloza, ví dụ nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, và tốt hơn nếu nguyên liệu chứa oligosacarit có liên kết β -xylosidic hoặc oligosacarit có liên kết β -glycosidic, miễn là nó chứa hemixenluloza hoặc xenluloza, hoặc tốt hơn nếu oligosacarit có liên kết β -xylosidic hoặc oligosacarit có liên kết β -glycosidic. Ví dụ về nguyên liệu lignoxenluloza bao gồm sinh khối xenluloza như cỏ dại, chất thải nông nghiệp, giấy thải, hoặc nguyên liệu tương tự. Trước khi tiếp xúc với enzym β -xylosidaza chịu nhiệt theo sáng chế, tốt hơn nếu nguyên liệu lignoxenluloza được xử lý cơ học, 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 đậm thích hợp, hoặc dung dịch tương tự.

Tức là, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế có thể còn bao gồm bước trong đó nguyên liệu lignoxenluloza được xử lý cơ học, xử lý hóa học, hoặc ngâm hoặc phân giải vào dung dịch đậm thích hợp trước khi tiếp xúc với enzym β -xylosidaza chịu nhiệt theo sáng chế

Tốt hơn nếu phản ứng thủy phân nguyên liệu lignoxenluloza bởi enzym β -xylosidaza chịu nhiệt theo sáng chế được thực hiện ở điều kiện mà ở đó enzym này có hoạt tính thủy phân xenlulooligosacarit. Tốt hơn nếu phản ứng thủy phân nguyên liệu lignoxenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 60°C đến 90°C và độ pH nằm trong khoảng từ 5,0 đến 9,0, và tốt hơn nữa nếu phản ứng thủy phân nguyên liệu lignoxenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C và độ pH nằm trong khoảng từ 5,0 đến 9,0 và còn tốt hơn nữa nếu phản ứng thủy phân nguyên liệu lignoxenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C và độ pH nằm trong khoảng từ 6,0 đến 8,5. Thời gian phản ứng thủy phân nguyên liệu lignoxenluloza 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 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 có thể được thực hiện trong thời gian từ 1 giờ đến 100 giờ.

Để thủy phân nguyên liệu lignoxenluloza, tốt hơn nữa nếu sử dụng đồng thời hoặc riêng rẽ ít nhất một glycosit hydrolaza khác ngoài enzym β -xylosidaza chịu nhiệt theo sáng chế. Glycosit hydrolaza này có thể giống như glycosit hydrolaza chứa trong hỗn hợp glycosit hydrolaza theo sáng chế, và tốt hơn nếu glycosit hydrolaza này là enzym 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 enzym β -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc enzym β -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym β -xylosidaza 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 enzym β -xylosidaza 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 enzym β -xylosidaza chịu nhiệt thể hệ mới theo sáng chế có hoạt tính ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C, 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 này được giải trình tự.

Các mẫu đất ở khu vực suối nước nóng có độ pH từ trung tính đến kiềm yếu được thu nhận từ 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 (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 (ISOIL Large for Beads ver. 2, do NIPPON GENE Co., Ltd. sản xuất) từ 10g mỗi mẫu đất thu được ở các khu vực suối nước nóng. ADN chiết tách được được giải trình tự bằng máy GS FLX Titan 454 do Roche Diagnostics Ltd. sản xuất và máy HiSeq 2000 do Illumina, Inc sản xuất. ADN chiết tách được (5µg) và sản phẩm khuếch đại bằng kit khuếch đại ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare, Inc. sản xuất) được giải trình tự lần lượt bằng máy GS FLX Titan 454 và máy HiSeq 2000. Khi giải trình tự bằng máy HiSeq 2000, thư viện ADN và các chất phản ứng được đổ vào tế bào chảy bằng hệ thống cBot do Illumina, Inc. sản xuất để tự động tạo ra cụm gen có trình tự giống hệt với phân tử ADN duy nhất trong tế bào chảy dòng chảy. Các đầu bắt cặp có chiều dài bằng 101bp được giải trình tự bằng máy HiSeq 2000 để thu được dữ liệu trình tự hệ gen môi trường.

Kết quả giải trình tự ADN hệ gen môi trường của mẫu đất suối nước nóng AR19 bằng máy GS FLX Titan 454 thu được các đoạn trình tự đọc có chiều dài trung bình bằng 396bp, tổng số các đoạn trình tự đọc bằng 2.766.332, và tổng chiều dài bộ gen đã được giải trình tự bằng 1.106.243.280bp và kết quả giải trình tự ADN hệ gen môi trường bằng máy HiSeq 2000 thu được các đoạn trình tự đọc có đầu bắt cặp với chiều dài trung bình bằng 92,65bp, tổng số các đoạn trình tự đọc bằng 894.238.096, và tổng chiều dài bộ gen đã được giải trình tự bằng 83.112.168.755bp, tức là tổng cộng chiều dài bộ gen đã được giải trình tự bằng khoảng 84,2Gbp.

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.

Các trình tự nucleotit đọc được bằng máy GS FLX Titan 454 và máy HiSeq 2000 được sàng lọc chất lượng và lắp ráp lại một lần nữa bằng hệ thống CLC Genomics Workbench (phiên bản 5.5.1) do CLC bio A/S sản xuất. Sau khi sàng lọc chất lượng, tổng chiều dài các đoạn trình tự đọc bằng máy GS FLX Titan 454 bằng 2.766.328bp, và tổng chiều dài đoạn trình tự đọc của dữ liệu trình tự nucleotit thu được bằng máy HiSeq 2000 bằng 81.323.692.563bp. Sau khi lắp ráp, tổng số lượng các đoạn tiếp giáp có chiều dài

bằng 500bp hoặc lớn hơn bằng 967.925, và tổng chiều dài các đoạn tiếp giáp này bằng 419.787.603bp, và đoạn tiếp giáp lớn nhất có chiều dài bằng 287.641bp.

3. Quá trình xác định các khung đọc mở (Khung đọc mở - ORF) của enzym 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 mở (các vùng gen) được xác định từ các trình tự tiếp giáp thu được ở phần 2 bằng phần mềm Metagene (Noguchi *et al.*, *DNA Research*, 2008, 15(6)) (tùy chọn Metagene: -m). Gen mã hóa glycosit hydrolaza từ các trình tự khung đọc mở thu được được xác định bằng phần mềm BLASTP (phiên bản 2.2.18) và đối chiếu với 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 mở mục tiêu thu được là gen mã hóa glycosit hydrolaza. Các trình tự nucleotit thu được bao gồm các gen mã hóa glycosit hydrolaza như xenlulaza, endohemixenlulaza, 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 nêu trên được phân loại chức năng bằng phần mềm cơ sở dữ liệu về trình tự vùng chức năng của protein 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). Cụ thể, mỗi trình tự nucleotit này được phân loại thành các họ glycosit hydrolaza (GH) dựa trên độ tương đồng của chúng với cơ sở dữ liệu về trình tự vùng chức năng của protein Pfam bằng thuật toán tra cứu độ tương đồng trình tự protein 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 cụ thể ở trình tự)). Các trình tự nucleotit 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 dữ liệu trình tự hệ gen môi trường AR19 có chiều dài bằng 84,2Gbp, 602.589 khung đọc mở được xác định bằng phần mềm Metagene, và tổng số lượng khung đọc mở có chiều dài đầy đủ bằng 251.146.

406 khung đọc mở mục tiêu được xác định là trình tự β -glucosidaza hoặc enzym β -xylosidaza bằng phần mềm tra cứu độ tương đồng BLASTP. Trong số 406 khung đọc mở này, 168 khung đọc mở được xác định là gen mã hóa β -glucosidaza hoặc enzym β -xylosidaza bằng phần mềm Pfam HMMs, trong khi đó 238 khung đọc mở còn lại (gồm 92 khung đọc mở có chiều dài đầy đủ và 146 khung đọc mở có chiều dài không đầy đủ) chứa nhỏ hơn 70% trình tự vùng xúc tác GH, hoặc không phải là trình tự mục tiêu trên cơ sở dữ liệu Pfam.

Kết quả phân loại họ GH của 168 khung đọc mở được xác định là gen mã hóa β -glucosidaza hoặc enzym β -xylosidaza được thể hiện trong Bảng 1. Các trình tự chứa nhỏ hơn 70% vùng xúc tác GH và các trình tự không tương đồng với trình tự bất kỳ trên cơ sở dữ liệu Pfam được phân loại thành họ GH chưa biết. Như được thể hiện trong Bảng 1, trong số các khung đọc mở được xác định là gen mã hóa β -glucosidaza hoặc enzym β -xylosidaza thu được từ hệ gen môi trường AR19, thì 19 khung đọc mở có chiều dài đầy đủ thuộc họ GH1, 57 khung đọc mở có chiều dài đầy đủ thuộc họ GH3, 13 khung đọc mở có chiều dài đầy đủ thuộc họ GH31, và 13 khung đọc mở có chiều dài đầy đủ thuộc họ GH43. Các đoạn môi được thiết kế cho toàn bộ các khung đọc mở có chiều dài đầy đủ mà đã được xác định là gen mã hóa β -glucosidaza hoặc enzym β -xylosidaza, 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 enzym β -xylosidaza được phân lập từ khung đọc mở AR19M-311 thuộc họ GH3 và có trình tự của gen enzym β -xylosidaza.

Bảng 1

Hệ gen môi trường AR19	Phân loại họ GH của gen β -glucosidaza hoặc enzym β -xylosidaza					
	GH1	GH3	GH31	GH43	GH khác	Tổng số

Các khung đọc mở có chiều dài đầy đủ	19	57	13	13	3	105
Các khung đọc mở có chiều dài không đầy đủ	3	52	4	4	0	63
Tổng số khung đọc mở	22	109	17	17	3	168

5. Khung đọc mở AR19M-311

Khung đọc mở AR19M-311 mã hóa polypeptit (SEQ ID NO.1) chứa 751 gốc axit amin và có chiều dài đầy đủ (SEQ ID NO.2), polypeptit này bắt đầu từ metionin 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 polypeptit mã hóa bởi khung đọc mở AR19M-311 là protein đa vùng chứa vùng xúc tác của glycosit hydrolaza GH3 gồm 325 gốc axit amin từ threonin (T) ở vị trí số 18 đến alanin (A) ở vị trí số 342, vùng xúc tác của glycosit hydrolaza GH3 gồm 236 gốc axit amin từ isoleucin (I) ở vị trí số 381 đến threonin (T) ở vị trí số 616; và vùng Fibronectin loại III gồm 70 gốc axit amin từ axit glutamic (E) ở vị trí số 651 đến serin (S) ở vị trí số 720. Kết quả phân tích bằng phần mềm xác định trình tự tín hiệu SignalP 4.1, cho thấy trình tự axit amin từ metionin (M) ở vị trí số 1 đóng vai trò là bộ ba mã hóa khởi đầu đến axit amin ở vị trí số 17 không mã hóa peptit tín hiệu bài tiết và có chức năng chưa biết.

6. Quá trình tách dòng gen.

Phản ứng PCR được thực hiện bằng cách sử dụng khuôn là ADN thu được từ đất ở khu vực suối nước nóng đã được khuếch đại bằng kit khuếch đại ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare, Inc. sản xuất), và đoạn mồi xuôi có trình tự nucleotit như nêu trong SEQ ID NO.9 (5'-CACCATGGAAGAAAGATGGTTACAAAG-3' được thiết kế bằng cách bổ sung 4 nucleotit (CACC) vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.7, trong đó 4 nucleotit CACC được bổ sung ở đầu 5' này là trình tự để chèn vào vectơ) và đoạn mồi ngược có trình tự nucleotit như nêu trong SEQ ID NO.8 (5'-TTAAGGTTCTATAATTACCTCGCTAG-3'). Trình tự nucleotit như nêu trong SEQ ID NO.7 là trình tự nucleotit tương đồng với đoạn trình tự chứa các nucleotit ở vị trí số 1 đến 23 của trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6. Hơn nữa, trình tự nucleotit như nêu trong SEQ ID NO.8 là trình tự nucleotit bổ sung với đoạn trình tự chứa các nucleotit ở vị trí số 2.231 đến 2.256 của trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4 hoặc SEQ ID NO.6. 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 Inc. 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 và nuôi cấy trong môi trường LB lỏng 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 ADN 3730 do Life Technologies Corporation sản xuất.

Kết quả thu được hai sản phẩm tách dòng gen AR19M-311-2 và AR19M-311-11 từ khung đọc mở AR19M-311 bằng quá trình tách dòng PCR. Trình tự nucleotit (SEQ ID NO.4) của sản phẩm tách dòng gen enzym β -xylosidaza tiềm năng AR19M-311-2 (sau đây gọi tắt là gen AR19M-311-2) và trình tự nucleotit (SEQ ID NO.6) của sản phẩm tách dòng gen enzym β -xylosidaza tiềm năng AR19M-311-11 (sau đây gọi tắt là gen AR19M-311-11) có chiều dài đều bằng 2.256bp tương đồng với trình tự nucleotit của khung đọc mở AR19M-311 (SEQ ID NO.2), và nhưng khác trình tự SEQ ID NO.2 lần lượt 19 và 20 nucleotit. Sự khác biệt về các nucleotit trên gen AR19M-311-2, gen AR19M-311-11 và khung đọc mở AR19M-311 được thể hiện trên Bảng 2. Sự khác biệt về trình tự nucleotit của khung đọc mở AR19M-311, gen AR19M-311-2, và gen AR19M-311-11 là do sự lắp ráp lỗi của máy giải trình hoặc dạng đa hình đơn nucleotit của các gen này.

Bảng 2

Nucleotit ở vị trí số	AR19M-311	AR19M-311-2	AR19M-311-11
438	T	T	C
560	G	A	A
561	A	G	G
564	A	G	A
567	A	T	T
570	T	A	A
579	A	C	C
581	T	C	C
606	T	A	A
609	T	G	G
612	G	A	A
615	A	T	T
617	C	G	G
618	A	T	T
625	T	C	C
630	T	C	C

672	C	C	A
932	C	T	T
1064	T	C	T
1086	T	T	C
1107	C	C	T
1224	T	C	C
1903	G	A	G

Trình tự axit amin (SEQ ID NO.3) của polypeptit (AR19M-311-2) mã hóa bởi gen AR19M-311-2 và trình tự axit amin (SEQ ID NO.5) của polypeptit (AR19M-311-11) mã hóa bởi gen AR19M-311-11 khác biệt với trình tự axit (SEQ ID NO.1) mã hóa bởi khung đọc mở AR19M-311 lần lượt ở 6 gốc axit amin và 4 gốc axit amin. Sự khác biệt về các gốc axit amin trên protein mã hóa bởi gen AR19M-311-2, gen AR19M-311-11 và khung đọc mở AR19M-311 được thể hiện trên Bảng 3. Ngoài ra, Fig.2 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin của khung đọc mở AR19M-311, protein AR19M-311-2, và protein AR19M-311-11. Trên Fig.2, ký tự “•” thể hiện các gốc axit amin giống hệt nhau, trong khi các gốc axit amin khác biệt được thể hiện bằng các chữ cái màu trắng trên nền đen. Trình tự axit amin của protein AR19M-311-2, và protein AR19M-311-11 chỉ khác nhau ở hai gốc axit amin, trong đó gốc axit amin ở vị trí số 355 và vị trí số 635 của protein AR19M-311-2 lần lượt là alanin (A) và isoleucin (I), trong khi đó các gốc axit amin ở vị trí số 355 và vị trí số 635 của protein AR19M-311-11 đều là valin (V).

Bảng 3

Axit amin ở vị trí số	AR19M-311	AR19M-311-2	AR19M-311-11
187	R	K	K
194	V	A	A
206	A	G	G
311	T	I	I
355	V	A	V
635	V	I	V

Fig.1 là hình vẽ thể hiện sự sắp thẳng hàng các trình tự axit amin của vùng xúc tác β -xylosidaza của sản phẩm tách dòng gen AR19M-311-2 và trình tự axit amin của enzym β -xylosidaza GH3 có nguồn gốc từ vi khuẩn *Dictyoglomus thermophilum*. Trên Fig.1, các axit amin được thể hiện bằng các chữ cái màu trắng trên nền đen là các gốc axit amin giống hệt nhau trong hai trình tự axit amin này, các gốc axit amin được thể hiện bằng các chữ cái màu đen trên nền xám được đánh bóng là các gốc axit amin gần giống nhau trên các trình tự axit amin này và ký tự “-” thể hiện các đoạn gạch bỏ (khoảng trống). Tính cả các khoảng trống, AR19M-311-2 có độ tương đồng bằng 70% so với

enzym β -xylosidaza GH3 của *Dictyoglomus thermophilum*.

7. Quá trình biểu hiện và tinh chế protein enzym β -xylosidaza.

Sau khi giải trình tự, các plasmit chứa gen đích (gen AR19M-311-2 hoặc gen AR19M-311-11) được biến nạp vào *E. coli* để biểu hiện protein bằng phương pháp sốc nhiệt. Chúng *E. coli* BL21 Star (DE3) có sẵn trong kit biểu hiện Champion pET Directional TOPO (do Life Technologies Inc. sản xuất) được sử dụng làm tế bào vật chủ để chuyển gen. Cụ thể, *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 để biểu hiện protein đích. 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 dùng phân cắt 5 phút được lặp lại từ 7 đến 8 lần bằng thiết bị Astrason 3000 (do MISONIX, Inc. sản xuất), nhờ đó thu được dịch chiết thô *E. coli* tái tổ hợp gen chứa protein đích. Dịch chiết thô chứa *E. coli* tái tổ hợp gen này được lọc qua màng lọc (đường kính lỗ lọc $\phi = 0,45\mu\text{m}$, do EMD Millipore Corporation sản xuất), và dịch lọc thu được được sử dụng làm dịch nuôi cấy đồng nhất *E. coli* tái tổ hợp gen.

Dịch nuôi cấy đồng nhất *E. coli* tái tổ hợp gen chứa gen AR19M-311-2 và gen AR19M-311-11 được nạp vào cột trao đổi ion HiTrap Q HP (do GE Healthcare, Inc. sản xuất) được làm cân bằng với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 50% trong dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 1M NaCl bằng hệ thống sắc ký lỏng AKTA (do GE Healthcare, Inc. sản xuất). Các phân đoạn có hoạt tính enzym β -xylosidaza được trộn toàn bộ, và 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 Biotech SA sản xuất). Các phân đoạn có hoạt tính enzym β -xylosidaza sau khi trao đổi dung dịch được nạp vào cột phân tách tương tác kỵ nước HiTrap Phenyl HP (do GE Healthcare, Inc. sản xuất) được làm cân bằng với cùng dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), và các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 100%

trong dung dịch đệm Tris-HCl 50mM (độ pH = 8,0). Các phân đoạn có hoạt tính enzym β -xylosidaza 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 bổ sung vào cột lọc gel Hiload 26/60 superdex 200 pg (do GE Healthcare, Inc. 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.

Tiếp theo, đối với gen AR19M-311-2, các phân đoạn có hoạt tính enzym β -xylosidaza được trộn toàn bộ, trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), và cô bằng màng siêu lọc ly tâm VIVASPIN 20, sau đó các protein tiếp tục được phân đoạn bằng cột trao đổi ion HiTrap Q HP theo cùng cách thức nêu trên. Tiếp theo, đối với cả hai gen AR19M-311-2 và AR19M-311-11, các phân đoạn có hoạt tính enzym β -xylosidaza được trộn toàn bộ, trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) và cô để thu được enzym tinh khiết AR19M-311-2 có nồng độ cuối cùng bằng khoảng 3,77mg/mL và enzym tinh khiết AR19M-311-11 có nồng độ cuối cùng bằng khoảng 35,6mg/mL.

Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và các enzym tinh khiết này được phân tích điện di trên gel SDS-PAGE (gel SDS-polyacrylamit). Đối với mỗi gen, dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (5 μ g) và enzym tinh khiết (0,5g) được trộn với dung dịch đệm mẫu đã được cô 4 lần chứa 2-mercaptoetanol (do Wako Pure Chemical Industries, Ltd. sản xuất) để điều chỉnh nồng độ dung dịch đệm đến 1 lần, sau đó gia nhiệt ở nhiệt độ 95°C trong 4 phút, tiếp theo điện di trên 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 bằng màu hệ thống ChemiDoc (do Bio-Rad Laboratories, Inc. sản xuất).

Fig.3A hoặc Fig.3B là hình ảnh 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 AR19M-311-2 (Fig.3A) hoặc gen AR19M-311-11 (Fig.3B) và các enzym tinh khiết được tinh chế từ các dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen này. Trên Fig.3, làn 1 thể hiện mẫu điện di của gen chỉ thị phân tử protein, làn 2 thể hiện mẫu điện

di của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen, và làn 3 thể hiện mẫu điện di của enzym tinh khiết. Kết quả phân tích cho thấy đối với cả hai gen AR19M-311-2 và AR19M-311-11, thì trong dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (làn 2) có một băng đậm có trọng lượng phân tử gần bằng trọng lượng phân tử của polypeptit có trình tự axit amin lần lượt nêu trong SEQ ID NO.3 hoặc SEQ ID NO.5, và trong mỗi enzym tinh khiết này (làn 3) có một băng duy nhất tương ứng với băng đậm này (được thể hiện bằng mũi tên trên Fig.3).

8. Đo hoạt tính enzym β -xylosidaza bằng cơ chất PNPX (hoạt tính thủy phân PNPX)

Để so sánh hoạt tính đặc riêng của enzym β -xylosidaza AR19M-311-2 và AR19M-311-11 tinh khiết, thì hoạt tính thủy phân PNPX (hoạt tính enzym β -xylosidaza) của mỗi enzym này được đánh giá. Đối với mỗi enzym, dung dịch enzym tinh khiết được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 bằng nước đến nồng độ bằng 10ng/ μ L được sử dụng.

PNPX được sử dụng làm cơ chất để đo hoạt tính enzym β -xylosidaza. Dung dịch cơ chất được điều chế bằng cách hòa tan PNPX (do Sigma-Aldrich Co. LLC. sản xuất) trong nước ở nồng độ 3,4mM (sau đây được gọi tắt là “dung dịch nước PNPX 3,4mM”). Dung dịch nước PNPX 3,4mM được điều chế bằng phương pháp nêu trên được sử dụng trong toàn bộ các thử nghiệm sau.

Cụ thể, dung dịch chứa 100 μ L dung dịch nước PNPX 3,4mM, 50 μ L dung dịch đệm axetat 200mM (độ pH bằng 4,0), 20 μ L dung dịch enzym tinh khiết (10ng/ μ L), và 30 μ L nước tinh khiết được để phản ứng ở nhiệt độ 70°C trong 20 phút. Trong toàn bộ các phép đo, dung dịch hỗn hợp được điều chế bằng cách bổ sung nước tinh khiết thay cho dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen hoặc enzym tinh khiết và để phản ứng ở cùng điều kiện được sử dụng làm mẫu đối chứng. Hơn nữa, dung dịch cơ chất và enzym (dung dịch enzym tinh khiết) được giữ riêng biệt ở nhiệt độ phản ứng trong 5 phút trước khi trộn với nhau để hoạt hóa phản ứng. Trong mỗi trường hợp, sau khi kết thúc phản ứng, phản ứng được dừng bằng cách bổ sung dung dịch natri carbonat 0,2M vào hỗn hợp phản ứng với cùng thể tích, sau đó hỗn hợp thu được được ly tâm để 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, sau đó hàm lượng p-nitrophenol tạo ra bởi quá trình thủy phân enzym được đo bằng cách so sánh độ sai lệch với mẫu đối chứng. Hoạt tính enzym để tạo ra 1 μ mol 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 3 lần để thu được trị số trung bình và độ lệch chuẩn.

Kết quả đo cho thấy cả AR19M-311-2 lẫn AR19M-311-11 đều có hoạt tính enzym β -xylosidaza (hoạt tính thủy phân PNPX). Hoạt tính riêng của AR19M-311-2 và AR19M-311-11 ở nhiệt độ 70°C lần lượt bằng 28,1U/mg, và 25,6U/mg. Trong các thử nghiệm dưới đây, AR19M-311-2 được sử dụng do nó có hoạt tính riêng cao hơn.

9. Độ đặc hiệu cơ chất của AR19M-311-2

Hoạt tính thủy phân các cơ chất xenluloza và hemixenluloza khác nhau của AR19M-311-2 được đo. Trong phép đo này, dung dịch enzym tinh khiết được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với nước đến nồng độ bằng 10ng/ μ L được sử dụng. Ngoài ra, PSA, bột Avicel, 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, Inc. sản xuất), và laminarin (có nguồn gốc từ *Laminaria digitata*, 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. 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 dung dịch phản ứng chứa 50 μ L dung dịch đệm acetat 200mM (độ pH bằng 4,0), 20 μ L dung dịch enzym tinh khiết (10ng/ μ L) và 30 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 (dung dịch nước chứa 1% theo khối

lượng PSA, bột Avicel, CMC, xylan, lichenan hoặc laminarin, hoặc dung dịch nước chứa 3,4mM PNPX hoặc PNPG) vào và ủ hỗn hợp thu được ở nhiệt độ 70°C trong 20 phút (hoặc 2 giờ khi cơ chất là bột Avicel). Khi PSA, bột Avicel, hoặc xylan được sử dụng làm cơ chất, thì trong quá trình phản ứng, hỗn hợp phản ứng này được khuấy bằng máy khuấy nhiệt ở tốc độ 1.400 vòng/phút (do Eppendorf AG sản xuất) để tránh kết tủa các cơ chất không tan.

Khi PNPG hoặc PNPX được sử dụng làm cơ chất, thì độ hấp thụ ở bước sóng 420nm của phần dịch nổi của hỗn hợp sau phản ứng, hàm lượng p-nitrophenol được tạo ra bằng quá trình thủy phân, và hoạt tính riêng (U/mg) được đo theo cùng cách thức như nêu trong phần 8 để đo hoạt tính thủy phân PNPX của AR19M-311-2. Khi PSA, bột Avicel, CMC, xylan, lichenan hoặc laminarin được sử dụng làm cơ chất, thì sau khi kết thúc phản ứng thuốc thử axit 3,5-dinitrosalicylic (dung dịch DNS) được bổ sung vào hỗn hợp phản ứng với cùng thể tích, và 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 nhiệt độ 25°C, và tốc độ 15.000g trong 5 phút để 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 (khi xylan được sử dụng làm cơ chất, thì hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn xyloza), và 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 1μmol 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).

Mỗi phép đo được lặp lại 3 lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.4. Các kết quả này cho thấy AR19M-311-2 có hoạt tính thủy phân mạnh PNPX và cũng có hoạt tính thủy phân PNPG. Trái lại, AR19M-311-2 hầu như không có hoạt tính phân giải các cơ chất còn lại.

10. Đánh giá mức độ phụ thuộc vào nhiệt độ và độ pH của hoạt tính enzym β -xylosidaza bằng cơ chất PNPX.

Mức độ phụ thuộc vào nhiệt độ và độ pH của hoạt tính thủy phân PNPX của

AR19M-311-2 được đánh giá. Trong phép đo này, dung dịch enzym tinh khiết (10ng/ μ L) được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với nước được sử dụng.

Mức độ phụ thuộc vào độ pH của hoạt tính thủy phân PNPX của enzym tinh khiết AR19M-311-2 được đo theo cùng cách thức như nêu trong phần 8, chỉ khác là cho hỗn hợp chứa 100 μ L dung dịch nước PNPX 3,4mM, 50 μ L dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3,0 đến 8,0), 50 μ L dung dịch đệm axetat (độ pH nằm trong khoảng từ 3,5 đến 6,0), hoặc 50 μ L dung dịch đệm phosphat (độ pH nằm trong khoảng từ 6,0 đến 8,0), 30 μ L nước tinh khiết và 20 μ L dung dịch enzym tinh khiết (10ng/ μ L) được cho phản ứng ở nhiệt độ 70°C trong 20 phút, và hàm lượng p-nitrophenol được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PNPX (U/mg) được đo ở mỗi độ pH.

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PNPX của enzym tinh khiết AR19M-311-2 được đo theo cùng cách thức như nêu trong phần 8, chỉ khác là phản ứng được thực hiện ở nhiệt độ 30°C, 40°C, 50°C, 60°C, 65°C, 70°C, 75°C, 80°C, 85°C, 90°C, 95°C, và 100°C, và hàm lượng p-nitrophenol được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PNPX (U/mg) được đo ở mỗi nhiệt độ.

Mỗi phép đo được lặp lại 3 lần để thu được trị số trung bình và độ lệch chuẩn. Các kết quả đo được thể hiện trên Fig.5 và Fig.6. Fig.5 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở độ pH bằng 4,0) của enzym tinh khiết AR19M-311-2 ở các nhiệt độ khác nhau, trong đó nhiệt độ được thể hiện trên trục hoành và Fig.6 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở nhiệt độ 70°C) của enzym tinh khiết AR19M-311-2 ở các độ pH khác nhau, trong đó độ pH được thể hiện trên trục hoành. Trên Fig.6 ký hiệu “MB” thể hiện kết quả đo thu được khi sử dụng dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3,0 đến 8,0), ký hiệu “SAB” thể hiện kết quả đo thu được khi sử dụng dung dịch đệm axetat (độ pH nằm trong khoảng từ 3,5 đến 6,0), và ký hiệu “PB” thể hiện kết quả đo thu được khi sử dụng dung dịch đệm phosphat (độ pH nằm trong khoảng từ 6,0 đến 8,0). 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.

Ở độ pH = 4,0, enzym tinh khiết AR19M-311-2 có hoạt tính thủy phân mạnh

PNPX ở nhiệt độ nằm trong khoảng từ 70°C đến 85°C (Fig.5). Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) ở đó enzym này có hoạt tính mạnh nhất bằng 80°C ở độ pH bằng 4,0. Hoạt tính thủy phân PNPX của enzym tinh khiết AR19M-311-2 bị giảm nhanh khi nhiệt độ phản ứng enzym bằng 90°C hoặc cao hơn ở độ pH = 4,0.

Ngoài ra, enzym tinh khiết AR19M-311-2 có hoạt tính thủy phân PNPX ở độ pH nằm trong khoảng từ 4,0 đến 8,0 ở nhiệt độ 70°C (Fig.6). Độ pH tối ưu của enzym tinh khiết AR19M-311-2 ở nhiệt độ 70°C là 4,19 (các trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym).

11. Đánh giá độ ổn định nhiệt của enzym β -xylosidaza bằng phương pháp phân tích nhiệt cắt lớp (Differential Scanning fluorimetry -DSF).

Phương pháp phân tích nhiệt cắt lớp (DSF) là phương pháp đo nhiệt độ biến tính của protein bằng thuốc nhuộm huỳnh quang và máy PCR định lượng, và có thể áp dụng cho nhiều protein. Thuốc nhuộm huỳnh quang được sử dụng để phân tích nhiệt cắt lớp, như thuốc nhuộm màu cam SYPRO, là chất phát huỳnh quang trong điều kiện không phân cực khi gắn kết vào vị trí kỵ nước, và sự phát huỳnh quang bị ức chế trong điều kiện phân cực khi được hòa tan trong nước. Nhìn chung, cấu trúc của protein bị tháo xoắn ở nhiệt độ biến tính, và các vùng kỵ nước của protein bị lộ ra khỏi bề mặt protein. Khi thuốc nhuộm màu cam SYPRO gắn kết với các vùng kỵ nước bị lộ ra bởi ánh sáng kích thích có bước sóng nằm trong khoảng từ 470nm đến 480nm, thì thuốc nhuộm này phát ra ánh sáng huỳnh quang có cực đại ở bước sóng khoảng 595nm. Bằng cách tăng dần nhiệt độ của dung dịch protein trong khoảng thời gian nhất định, và đo cường độ huỳnh quang tương ứng thì có thể xác định được nhiệt độ biến tính của protein (tức là nhiệt độ làm thay đổi cường độ huỳnh quang).

Trong phép đo này, dung dịch enzym tinh khiết (1mg/mL) được điều chế bằng cách pha loãng enzym tinh khiết thu được ở phần 7 với nước được sử dụng

Cụ thể, 2 μ L thuốc nhuộm màu cam SYPRO đã pha loãng 100 lần (do Life Technologies Inc. sản xuất Inc.), 1 μ L enzym tinh khiết có nồng độ bằng 1mg/mL, 5 μ L dung dịch đệm axetat 200mM (pH = 4,0) và 12 μ L nước tinh khiết được bổ sung vào các giếng của đĩa PCR 96-giếng (đĩa PCR 96-giếng MLL-9651, do Bio-Rad Laboratories, Inc.

sản xuất) sao cho thể tích của mỗi giếng bằng 20 μ L. Đĩa PCR được đậy kín bằng các nắp đậy có 8 mũ chụp (do Bio-Rad Laboratories, Inc. sản xuất), nhiệt độ của các giếng được tăng từ 30°C đến 100°C mỗi lần 0,2°C bằng máy PCR định lượng (Hệ thống PCR định lượng CFX96 Touch, do Bio-Rad Laboratories, Inc. sản xuất), và sau 10 giây kể từ khi đạt đến nhiệt độ mục tiêu, cường độ huỳnh quang của các giếng tương ứng được định lượng đồng thời. Thuốc nhuộm màu cam SYPRO được kích thích bằng đi-ốt phát quang (LED) có bước sóng ánh sáng phát ra nằm trong khoảng từ 450nm đến 490nm, và ánh sáng được phát xạ từ thuốc nhuộm màu cam SYPRO được cho qua bộ lọc thông dải ở bước sóng nằm trong khoảng từ 560nm đến 580nm, cường độ huỳnh quang được định lượng bằng máy ảnh CCD, và sự thay đổi về cường độ huỳnh quang được vẽ đồ thị dưới dạng hàm số của nhiệt độ. Dữ liệu được phân tích bằng phần mềm CFX Manager (do Bio-Rad Laboratories, Inc. sản xuất) được tích hợp trên máy PCR định lượng. Mỗi phép đo được lặp lại 3 lần.

Fig.7A và Fig.7B thể hiện sự thay đổi về cường độ huỳnh quang của thuốc nhuộm màu cam SYPRO được định lượng bằng phương pháp DSF và gây ra bởi sự biến tính nhiệt của protein enzym AR19M-311-2. Fig.7A thể hiện dữ liệu thực đo được, và Fig.7B thể hiện đạo hàm bậc nhất “ $-d(\text{Cường độ huỳnh quang})/dT$ ” của đồ thị thay đổi cường độ huỳnh quang trên Fig.7A. Nhiệt độ biến tính (nhiệt độ nóng chảy; trị số T_m) được xác định là trị số cực đại của đạo hàm bậc nhất “ $-d(\text{Cường độ huỳnh quang})/dT$ ” được thể hiện trên trục Y của Fig.7B. Khi độ pH = 4,0, đạo hàm bậc nhất cường độ huỳnh quang của AR19M-311-2 có hai cực đại lần lượt ở $71,5^\circ\text{C} \pm 0,1^\circ\text{C}$ ($n = 3$) và $82,7^\circ\text{C} \pm 0,1^\circ\text{C}$ ($n = 3$). Dựa trên kết quả đo nhiệt độ tối ưu thực tế cho hoạt tính thủy phân PNPX của AR19M-311-2 bằng 80°C , thì nhiệt độ cực đại thứ nhất bằng $71,5^\circ\text{C}$ là do protein enzym AR19M-311-2 bị biến đổi một phần cấu trúc, nhưng hoạt tính enzym của nó không bị giảm. Nhiệt độ cực đại thứ hai là gần bằng nhiệt độ tối ưu $T_{\text{tối ưu}} = 80^\circ\text{C}$ cho hoạt tính enzym, và kết quả này cho thấy nhiệt độ nóng chảy của protein enzym AR19M-311-2 bằng $82,7^\circ\text{C}$.

YÊU CẦU BẢO HỘ

1. Enzym β -xylosidaza chịu nhiệt bao gồm polypeptit cấu thành từ trình tự axit amin như nêu trong SEQ ID NO.1 và ít nhất một vùng được chọn từ nhóm bao gồm vùng tương tự Fibronectin loại III, gốc liên kết, peptit tín hiệu và trình tự đánh dấu.
2. Enzym β -xylosidaza chịu nhiệt theo điểm 1, trong đó enzym này còn có hoạt tính β -glucosidaza.
3. Hỗn hợp glycosit hydrolaza chứa enzym β -xylosidaza chịu nhiệt theo điểm 1 và ít nhất một glycosit hydrolaza khác.
4. 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, hemixenluloza và lignin tiếp xúc với enzym β -xylosidaza chịu nhiệt theo điểm 1.
5. 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, hemixenluloza và lignin tiếp xúc với hỗn hợp glycosit hydrolaza theo điểm 3.

DANH MỤC TRÌNH TỰ

<110> HONDA MOTOR CO., LTD.
 <120> ENZYM BETA-XYLOSIDAZA CHỊU NHIỆT, HỖN HỢP GLYCOSIT HYDROLAZA CHỨA ENZYM NÀY, VÀ PHƯƠNG
 PHÁP SẢN XUẤT SẢN PHẨM PHÂN GIẢI LIGNOXENLULOZA
 <130> H114098901
 <160> 10
 <170> PatentIn version 3.1
 <210> 1
 <211> 751
 <212> PRT
 <213> Trình tự chưa biết
 <220>
 <223> Trình tự axit amin được suy ra của khung đọc mở AR19M-311
 <400> 1
 Met Glu Glu Arg Trp Leu Gln Arg Arg Val Glu Glu Leu Leu Ser Lys
 1 5 10 15
 Met Thr Leu Glu Glu Lys Ile Ala Gln Leu Gly Ser Ile Pro Ser Gly
 20 25 30
 Lys Leu Val Glu Asn Gly Lys Phe Ser Arg Glu Lys Ala Lys Glu Leu
 35 40 45
 Leu Lys Asn Gly Ile Gly Gln Ile Thr Arg Val Ala Gly Tyr Ala Glu
 50 55 60
 Arg Glu Pro Glu Glu Ser Ile Glu Leu Ile Asn Glu Ile Gln Arg Phe
 65 70 75 80
 Leu Lys Glu Glu Thr Arg Leu Gly Ile Pro Ala Ile Ile His Glu Glu
 85 90 95
 Cys Leu Ser Gly Val Met Thr Lys Gly Ala Thr Thr Phe Pro Gln Ala
 100 105 110
 Ile Gly Met Ala Ser Thr Phe Glu Pro Asp Asp Ile Gln Arg Met Thr
 115 120 125
 Ser Ile Ile Arg Lys Glu Met Lys Ala Phe Gly Val His Gln Gly Leu
 130 135 140
 Ser Pro Val Leu Asp Ile Pro Arg Asp Pro Arg Trp Gly Arg Thr Glu
 145 150 155 160
 Glu Thr Phe Gly Glu Asp Pro Tyr Leu Val Ser Lys Met Ala Glu Ser
 165 170 175
 Tyr Ile Lys Gly Leu Gln Gly Glu Asp Trp Arg Glu Gly Ile Ile Ala
 180 185 190
 Thr Val Lys His Phe Thr Ala Tyr Gly Ile Ser Glu Gly Ala Arg Asn
 195 200 205
 Leu Gly Pro Ala Arg Val Ser Glu Arg Glu Leu Arg Glu Val Phe Leu
 210 215 220
 Phe Pro Phe Glu Val Ala Ile Arg Lys Ala Asn Ala Gly Ser Val Met
 225 230 235 240
 Asn Ala Tyr His Glu Ile Asp Gly Val Pro Cys Ala Ser Ser Lys Phe
 245 250 255
 Leu Leu Thr Lys Ile Leu Arg Glu Glu Trp Gly Phe Lys Gly Phe Val
 260 265 270
 Val Ser Asp Tyr Ser Ala Ile Glu Met Leu His Thr Phe His Lys Val
 275 280 285
 Ala Lys Asp Leu Lys Thr Ala Ala Ile Lys Ala Leu Glu Ala Gly Ile
 290 295 300
 Glu Ile Glu Leu Pro Glu Thr Lys Cys Tyr Gly Glu Pro Leu Leu Ser
 305 310 315 320
 Ala Val Lys Glu Gly Lys Val Ser Val Ser Val Ile Asp Thr Ala Val
 325 330 335
 Ala Arg Val Leu Arg Ala Lys Ile Leu Leu Gly Leu Leu Asp Asp Ile
 340 345 350
 Ile Tyr Val Asp Pro Ser Lys Ile Arg Ala Val Leu Asp Asn Pro Glu
 355 360 365
 His Arg Ala Phe Ala Arg Glu Leu Ala Arg Lys Ser Ile Val Leu Leu
 370 375 380
 Lys Asn Asp Gly Ile Leu Pro Ile Ser Lys Gly Val Lys Thr Ile Ala
 385 390 395 400
 Val Ile Gly Pro Ser Ala Asp Ser Thr Lys Asn Leu His Gly Asp Tyr
 405 410 415
 Ser Tyr Thr Ser His Ile Ala Gly Val Ala Asp Gly Val Arg Thr Val
 420 425 430
 Thr Val Leu Glu Gly Ile Lys Asn Lys Val Ser Ser Gly Thr Thr Val
 435 440 445
 Leu Tyr Ala Lys Gly Cys Glu Leu Ser Asp Glu Ser Arg Glu Gly Phe
 450 455 460
 Lys Glu Ala Leu Asp Ile Ala Ser Arg Ser Asp Val Ile Ile Ala Val
 465 470 475 480
 Met Gly Glu Asn Ser Gly Leu Phe Lys Arg Gly Ile Ser Gly Glu Gly

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

<210> 2
 <211> 2256
 <212> ADN
 <213> Trình tự chưa biết
 <220>
 <223> Khung đọc mở AR19M-311
 <400> 2

atggaagaaa	gatgggttaca	aagaagagtt	gaagagttat	tatcgaagat	gactcttgaa	60
gaaaaaatag	cccaacttgg	ttcaatacca	tcgggttaa	tagtagaaaa	tggaattt	120
tctagagaaa	aagccaaaga	acttttgaaa	aatggtatag	gacaaataac	tagagtagca	180
ggatacgccg	agagagagcc	agaagaatcg	atagaactta	taaatgagat	tcagagattt	240
ctaaaagagg	agactagact	aggtataccg	gcaataattc	acgaagagtg	tccttcgggt	300
gtcatgacca	aaggagcaac	aacatttoca	caggctatag	ggatggctag	cacatttgag	360
ccagacgata	ttcaaagaat	gacttccatt	ataagaaaag	agatgaaggc	atttgggtgt	420
catcaagggc	tctctcctgt	ccttgatatt	cctagggacc	ctagatggg	aaggacagaa	480
gaaacatttg	gagaagaccc	atatcttgta	tcaaagatgg	cagaaagtta	tataaaagg	540
cttcaggcag	agaattggag	agaaggaatt	attgctacag	ttaaacattt	tacagcttac	600
ggtattttctg	agggagcaag	aaatttgggt	cctgccagag	tatctgaaag	ggaattaaga	660
gaggtattcc	tctttccatt	tgaagtcgca	ataagaaaag	ccaatgctgg	ttctgtaatg	720
aatgcctatc	atgagataga	tgggggtccca	tgtgcgtcct	ccaagttttt	actgacaaaa	780
atccttagag	agaattgggg	attcaaagga	tttgtagtct	cagactattc	agctatagag	840
atgtttacata	cattccacaa	agtagctaaa	gacctaaaaga	ctgcagctat	aaaagcccta	900
gaggcaggaa	ttgaaataga	attaccagag	acaaaatggt	atggtgaacc	acttttatca	960
gcagtcaaa	aaggaaagg	atctgtctct	gttatagata	ccgctgtagc	tagagtctta	1020
agggcacaag	tactcctagg	actatttagat	gatattatct	acgtagatcc	aagtaagata	1080
agagctgtct	tagataaccc	agagcacagg	gcatttgcca	gagaactagc	tagaaaatct	1140
attgttttat	taaagaatga	tggaaacta	cccataagta	aagggtgtaa	aaccattgca	1200
gttataggtc	cgagtgcaga	tagtacaaaa	aatctacatg	gagactatag	ttataacctca	1260
catatagcag	gagtagcgga	cggtgtaagg	acagtaacag	tggttagagg	tatcaaaaa	1320
aaagtgtcct	caggaaccac	tgtcctctat	gctaagggtc	gtgaacttag	tgatgaatct	1380
agagaaggat	ttaaagaggc	attagatata	gctagtagat	cagatgtcat	tatagcgggtg	1440
atgggagaaa	atagcggact	atttaagaga	ggaatctcag	gagaaggtaa	tgacaggata	1500
gatctaaaa	ttccagggtg	tcaagaggaa	cttctaaagg	ctttaaagga	agtgggtaaa	1560
ccaattgtgc	tagtttctagt	taattggaag	cccctctcta	taaagtggga	gaaagaaaa	1620
atcccagcta	ttctagaagt	atggtatcca	ggggaagaag	gaggaaacgc	aatagcagat	1680
gtaatatattg	gagattataa	cccaggaggg	aagttaccaa	tctcattccc	aaaggatgtt	1740
ggacaaaatac	cagtctacta	taatagaaag	ccatctgcct	ttagtgaata	cttaacaaca	1800
gatactaaac	ccttctttoc	atttggacat	ggcttgagtt	atacaacctt	tgagtattca	1860
gaactaaaaa	taatcccaga	aaatgtcatg	cccgggggat	atgtagacat	aagctttaa	1920
gtaagaaata	caggcaatat	agacggagat	gaagtagtgc	aactttatat	acatgatgaa	1980
tgggcaagtg	tagaaagacc	aataaaagaa	ctaaaagggt	ttaaagggat	tcacctaaaa	2040
gcaagagaag	aaaagaagg	caccttcaga	ttatttacag	accaggttagc	attttatgat	2100
gaagtgatgc	gatttgtggt	agaagcaggg	acatttgaag	ttatggtagg	gtcctcctca	2160
gaagatataa	gattaacagg	taagtttgag	gtcttagaga	ccaaggttat	aacaaggat	2220

agaaaatttg ctagcgaggt aattatagaa ccttaa

2256

<210> 3

<211> 751

<212> PRT

<213> Trình tự chưa biết

<220>

<223> AR19M-311-2

<400> 3

```

Met Glu Glu Arg Trp Leu Gln Arg Arg Val Glu Glu Leu Leu Ser Lys
 1           5           10           15
Met Thr Leu Glu Glu Lys Ile Ala Gln Leu Gly Ser Ile Pro Ser Gly
          20           25           30
Lys Leu Val Glu Asn Gly Lys Phe Ser Arg Glu Lys Ala Lys Glu Leu
        35           40           45
Leu Lys Asn Gly Ile Gly Gln Ile Thr Arg Val Ala Gly Tyr Ala Glu
        50           55           60
Arg Glu Pro Glu Glu Ser Ile Glu Leu Ile Asn Glu Ile Gln Arg Phe
        65           70           75           80
Leu Lys Glu Glu Thr Arg Leu Gly Ile Pro Ala Ile Ile His Glu Glu
          85           90           95
Cys Leu Ser Gly Val Met Thr Lys Gly Ala Thr Thr Phe Pro Gln Ala
        100          105          110
Ile Gly Met Ala Ser Thr Phe Glu Pro Asp Asp Ile Gln Arg Met Thr
        115          120          125
Ser Ile Ile Arg Lys Glu Met Lys Ala Phe Gly Val His Gln Gly Leu
        130          135          140
Ser Pro Val Leu Asp Ile Pro Arg Asp Pro Arg Trp Gly Arg Thr Glu
        145          150          155          160
Glu Thr Phe Gly Glu Asp Pro Tyr Leu Val Ser Lys Met Ala Glu Ser
        165          170          175
Tyr Ile Lys Gly Leu Gln Gly Glu Asp Trp Lys Glu Gly Ile Ile Ala
        180          185          190
Thr Ala Lys His Phe Thr Ala Tyr Gly Ile Ser Glu Gly Gly Arg Asn
        195          200          205
Leu Gly Pro Ala Arg Val Ser Glu Arg Glu Leu Arg Glu Val Phe Leu
        210          215          220
Phe Pro Phe Glu Val Ala Ile Arg Lys Ala Asn Ala Gly Ser Val Met
        225          230          235          240
Asn Ala Tyr His Glu Ile Asp Gly Val Pro Cys Ala Ser Ser Lys Phe
        245          250          255
Leu Leu Thr Lys Ile Leu Arg Glu Glu Trp Gly Phe Lys Gly Phe Val
        260          265          270
Val Ser Asp Tyr Ser Ala Ile Glu Met Leu His Thr Phe His Lys Val
        275          280          285
Ala Lys Asp Leu Lys Thr Ala Ala Ile Lys Ala Leu Glu Ala Gly Ile
        290          295          300
Glu Ile Glu Leu Pro Glu Ile Lys Cys Tyr Gly Glu Pro Leu Leu Ser
        305          310          315          320
Ala Val Lys Glu Gly Lys Val Ser Val Ser Val Ile Asp Thr Ala Val
        325          330          335
Ala Arg Val Leu Arg Ala Lys Ile Leu Leu Gly Leu Leu Asp Asp Ile
        340          345          350
Ile Tyr Ala Asp Pro Ser Lys Ile Arg Ala Val Leu Asp Asn Pro Glu
        355          360          365
His Arg Ala Phe Ala Arg Glu Leu Ala Arg Lys Ser Ile Val Leu Leu
        370          375          380
Lys Asn Asp Gly Ile Leu Pro Ile Ser Lys Gly Val Lys Thr Ile Ala
        385          390          395          400
Val Ile Gly Pro Ser Ala Asp Ser Thr Lys Asn Leu His Gly Asp Tyr
        405          410          415
Ser Tyr Thr Ser His Ile Ala Gly Val Ala Asp Gly Val Arg Thr Val
        420          425          430
Thr Val Leu Glu Gly Ile Lys Asn Lys Val Ser Ser Gly Thr Thr Val
        435          440          445
Leu Tyr Ala Lys Gly Cys Glu Leu Ser Asp Glu Ser Arg Glu Gly Phe
        450          455          460
Lys Glu Ala Leu Asp Ile Ala Ser Arg Ser Asp Val Ile Ile Ala Val
        465          470          475          480
Met Gly Glu Asn Ser Gly Leu Phe Lys Arg Gly Ile Ser Gly Glu Gly
        485          490          495
Asn Asp Arg Ile Asp Leu Lys Leu Pro Gly Val Gln Glu Glu Leu Leu
        500          505          510
Lys Ala Leu Lys Glu Val Gly Lys Pro Ile Val Leu Val Leu Val Asn
        515          520          525
Gly Arg Pro Leu Ser Ile Lys Trp Glu Lys Glu Asn Ile Pro Ala Ile
        530          535          540
Leu Glu Val Trp Tyr Pro Gly Glu Glu Gly Gly Asn Ala Ile Ala Asp

```

545 550 555 560
 Val Ile Phe Gly Asp Tyr Asn Pro Gly Gly Lys Leu Pro Ile Ser Phe
 565 570 575
 Pro Lys Asp Val Gly Gln Ile Pro Val Tyr Tyr Asn Arg Lys Pro Ser
 580 585 590
 Ala Phe Ser Glu Tyr Leu Thr Thr Asp Thr Lys Pro Leu Phe Pro Phe
 595 600 605
 Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Glu Leu Lys Ile
 610 615 620
 Ile Pro Glu Asn Val Met Pro Gly Gly Tyr Ile Asp Ile Ser Phe Lys
 625 630 635 640
 Val Arg Asn Thr Gly Asn Ile Asp Gly Asp Glu Val Val Gln Leu Tyr
 645 650 655
 Ile His Asp Glu Trp Ala Ser Val Glu Arg Pro Ile Lys Glu Leu Lys
 660 665 670
 Gly Phe Lys Arg Ile His Leu Lys Ala Arg Glu Glu Lys Lys Val Thr
 675 680 685
 Phe Arg Leu Phe Thr Asp Gln Leu Ala Phe Tyr Asp Glu Val Met Arg
 690 695 700
 Phe Val Val Glu Ala Gly Thr Phe Glu Val Met Val Gly Ser Ser Ser
 705 710 715 720
 Glu Asp Ile Arg Leu Thr Gly Lys Phe Glu Val Leu Glu Thr Lys Val
 725 730 735
 Ile Thr Lys Asp Arg Lys Phe Ala Ser Glu Val Ile Ile Glu Pro
 740 745 750
 <210> 4
 <211> 2256
 <212> ADN
 <213> Trình tự chưa biết
 <220>
 <223> AR19M-311-2
 <400> 4
 atggaagaaa gatgggttaca aagaagagtt gaagagttat tatcgaagat gactcttgaa 60
 gaaaaaatag cccaacttgg ttcaatacca tcgggttaaac tagtagaaaa tggaaaattt 120
 tctagagaaa aagccaaaaga acttttgaaa aatggtatag gacaaataac tagagtagca 180
 ggatacgccg agagagagcc agaagaatcg atagaactta taaatgagat tcagagattt 240
 ctaaaagagg agactagact aggtataccg gcaataattc acgaagagtg tctttcgggt 300
 gtcatgacca aaggagcaac aacatttcca caggctatag ggatggctag cacatttgag 360
 ccagacgata ttcaaaagaat gacttccatt ataagaaaag agatgaaggc atttggtgtt 420
 catcaagggc tctctcctgt ccttgatatt cctagggacc ctatagggg aaggacagaa 480
 gaaacatttg gagaagaccc atatcttgta tcaaagatgg cagaaagtta tataaaagg 540
 gttcaggggg aagatttgaa ggagggtata attgctaccg ctaaacattt tacagcttac 600
 ggtatatcgg aaggtggtag aaatctgggc cctgccagag tatctgaaag ggaattaaga 660
 gaggtattcc tctttccatt tgaagtcgca ataagaaaag ccaatgctgg ttctgtaatg 720
 aatgcctatc atgagataga tgggggtccca tgtgcgtcct ccaagttttt actgacaaaa 780
 atccttagag aagaatgggg attcaaagga tttgtagtct cagactattc agctatagag 840
 atgttacata cattccacaa agtagctaaa gacctaaaga ctgcagctat aaaagcccta 900
 gaggcaggaa ttgaaataga attaccagag ataaaaatg attggtgaacc acttttattca 960
 gcagtcaaaag aaggaaaagg atctgtctct gttatagata ccgctgtagc tagagtctta 1020
 agggcaaaga tactcctagg actattagat gataattatc acgcagatcc aagtaagata 1080
 agagctgtct tagataaacc agagcacagg gcatttgcca gagaactagc tagaaaatct 1140
 attgttttat taaagaatga tggaaacta cccataagta aaggggtaaa aaccattgca 1200
 gttataggtc cgagtgcaga tagcacaaaa aatctacatg gagactatag ttataacctca 1260
 catatagcag gtagtagcga cggtgtaagg acagtaacag tggttagagg tatcaaaaaat 1320
 aaagtgtcct caggaaccac tgtcctctat gctaagggct gtgaacttag tgatgaatct 1380
 agagaaggat ttaaagaggc attagatata gctagtagat cagatgtcat tatagcgggtg 1440
 atgggagaaa atagcggact atttaagaga ggaatctcag gagaaggtaa tgacaggata 1500
 gatctaaaac ttccaggggt tcaagaggaa cttctaaagg ttttaaagga agtgggtaaa 1560
 ccaattgtgc tagttctagt taatggaagg cccctctcta taaagtggga gaaagaaaaat 1620
 atcccagcta ttctagaagt atggtatcca ggggaagaag gaggaacgc aatagcagat 1680
 gtaatatatt gagattataa cccaggaggg aagttaccaa tctcattccc aaaggatgtt 1740
 ggacaaaatac cagtctacta taatagaaag ccatctgcct ttagtgaata cttacaaca 1800
 gatactaaac ccttgtttcc atttggacat ggcttgagtt atacaacctt tgagtattca 1860
 gaactaaaaa taatcccaga aaatgtcatg cccgggggat atatagacat aagctttaaa 1920
 gtaagaaata caggcaatat agacggagat gaagtagtgc aactttatat acatgatgaa 1980
 tgggcaagtg tagaaagacc aataaaagaa ctaaaagggt ttaagaggat tcacctaaaa 2040
 gcaagagaag aaaagaagggt caccctcaga ttatttacag accagttagc attttatgat 2100
 gaagtgatgc gatttgtggt agaagcaggg acatttgaag ttatggtagg gtcctcctca 2160
 gaagatataa gattaacagg taagtttgag gtcttagaga ccaaggttat aacaaaggat 2220
 agaaaatttg ctacgcagggt aattatagaa ccttaa 2256
 <210> 5
 <211> 751
 <212> PRT
 <213> Trình tự chưa biết
 <220>
 <223> AR19M-311-11
 <400> 5

```

Met Glu Glu Arg Trp Leu Gln Arg Arg Val Glu Glu Leu Leu Ser Lys
 1           5           10           15
Met Thr Leu Glu Glu Lys Ile Ala Gln Leu Gly Ser Ile Pro Ser Gly
          20           25           30
Lys Leu Val Glu Asn Gly Lys Phe Ser Arg Glu Lys Ala Lys Glu Leu
          35           40           45
Leu Lys Asn Gly Ile Gly Gln Ile Thr Arg Val Ala Gly Tyr Ala Glu
          50           55           60
Arg Glu Pro Glu Glu Ser Ile Glu Leu Ile Asn Glu Ile Gln Arg Phe
 65           70           75           80
Leu Lys Glu Glu Thr Arg Leu Gly Ile Pro Ala Ile Ile His Glu Glu
          85           90           95
Cys Leu Ser Gly Val Met Thr Lys Gly Ala Thr Thr Phe Pro Gln Ala
          100          105          110
Ile Gly Met Ala Ser Thr Phe Glu Pro Asp Asp Ile Gln Arg Met Thr
          115          120          125
Ser Ile Ile Arg Lys Glu Met Lys Ala Phe Gly Val His Gln Gly Leu
          130          135          140
Ser Pro Val Leu Asp Ile Pro Arg Asp Pro Arg Trp Gly Arg Thr Glu
          145          150          155          160
Glu Thr Phe Gly Glu Asp Pro Tyr Leu Val Ser Lys Met Ala Glu Ser
          165          170          175
Tyr Ile Lys Gly Leu Gln Gly Glu Asp Trp Lys Glu Gly Ile Ile Ala
          180          185          190
Thr Ala Lys His Phe Thr Ala Tyr Gly Ile Ser Glu Gly Gly Arg Asn
          195          200          205
Leu Gly Pro Ala Arg Val Ser Glu Arg Glu Leu Arg Glu Val Phe Leu
          210          215          220
Phe Pro Phe Glu Val Ala Ile Arg Lys Ala Asn Ala Gly Ser Val Met
          225          230          235          240
Asn Ala Tyr His Glu Ile Asp Gly Val Pro Cys Ala Ser Ser Lys Phe
          245          250          255
Leu Leu Thr Lys Ile Leu Arg Glu Glu Trp Gly Phe Lys Gly Phe Val
          260          265          270
Val Ser Asp Tyr Ser Ala Ile Glu Met Leu His Thr Phe His Lys Val
          275          280          285
Ala Lys Asp Leu Lys Thr Ala Ala Ile Lys Ala Leu Glu Ala Gly Ile
          290          295          300
Glu Ile Glu Leu Pro Glu Ile Lys Cys Tyr Gly Glu Pro Leu Leu Ser
          305          310          315          320
Ala Val Lys Glu Gly Lys Val Ser Val Ser Val Ile Asp Thr Ala Val
          325          330          335
Ala Arg Val Leu Arg Ala Lys Ile Leu Leu Gly Leu Leu Asp Asp Ile
          340          345          350
Ile Tyr Val Asp Pro Ser Lys Ile Arg Ala Val Leu Asp Asn Pro Glu
          355          360          365
His Arg Ala Phe Ala Arg Glu Leu Ala Arg Lys Ser Ile Val Leu Leu
          370          375          380
Lys Asn Asp Gly Ile Leu Pro Ile Ser Lys Gly Val Lys Thr Ile Ala
          385          390          395          400
Val Ile Gly Pro Ser Ala Asp Ser Thr Lys Asn Leu His Gly Asp Tyr
          405          410          415
Ser Tyr Thr Ser His Ile Ala Gly Val Ala Asp Gly Val Arg Thr Val
          420          425          430
Thr Val Leu Glu Gly Ile Lys Asn Lys Val Ser Ser Gly Thr Thr Val
          435          440          445
Leu Tyr Ala Lys Gly Cys Glu Leu Ser Asp Glu Ser Arg Glu Gly Phe
          450          455          460
Lys Glu Ala Leu Asp Ile Ala Ser Arg Ser Asp Val Ile Ile Ala Val
          465          470          475          480
Met Gly Glu Asn Ser Gly Leu Phe Lys Arg Gly Ile Ser Gly Glu Gly
          485          490          495
Asn Asp Arg Ile Asp Leu Lys Leu Pro Gly Val Gln Glu Glu Leu Leu
          500          505          510
Lys Ala Leu Lys Glu Val Gly Lys Pro Ile Val Leu Val Leu Val Asn
          515          520          525
Gly Arg Pro Leu Ser Ile Lys Trp Glu Lys Glu Asn Ile Pro Ala Ile
          530          535          540
Leu Glu Val Trp Tyr Pro Gly Glu Glu Gly Gly Asn Ala Ile Ala Asp
          545          550          555          560
Val Ile Phe Gly Asp Tyr Asn Pro Gly Gly Lys Leu Pro Ile Ser Phe
          565          570          575
Pro Lys Asp Val Gly Gln Ile Pro Val Tyr Tyr Asn Arg Lys Pro Ser
          580          585          590
Ala Phe Ser Glu Tyr Leu Thr Thr Asp Thr Lys Pro Leu Phe Pro Phe
          595          600          605
Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Glu Leu Lys Ile

```

```

        610                615                620
Ile Pro Glu Asn Val Met Pro Gly Gly Tyr Val Asp Ile Ser Phe Lys
625                630                635                640
Val Arg Asn Thr Gly Asn Ile Asp Gly Asp Glu Val Val Gln Leu Tyr
        645                650                655
Ile His Asp Glu Trp Ala Ser Val Glu Arg Pro Ile Lys Glu Leu Lys
        660                665                670
Gly Phe Lys Arg Ile His Leu Lys Ala Arg Glu Glu Lys Lys Val Thr
        675                680                685
Phe Arg Leu Phe Thr Asp Gln Leu Ala Phe Tyr Asp Glu Val Met Arg
        690                695                700
Phe Val Val Glu Ala Gly Thr Phe Glu Val Met Val Gly Ser Ser Ser
705                710                715                720
Glu Asp Ile Arg Leu Thr Gly Lys Phe Glu Val Leu Glu Thr Lys Val
        725                730                735
Ile Thr Lys Asp Arg Lys Phe Ala Ser Glu Val Ile Ile Glu Pro
        740                745                750

<210> 6
<211> 2256
<212> ADN
<213> Trình tự chưa biết
<220>
<223> AR19M-311-11
<400> 6
atggaagaaa gatggttaca aagaagagtt gaagagttat tatcgaagat gactcttgaa      60
gaaaaaatag cccaacttgg ttcaatacca tcgggtaaac tagtagaaaa tggaaaaattt      120
tctagagaaa aagccaaaga acttttgaaa aatggtatag gacaaataac tagagtagca      180
ggatacgcgc agagagagcc agaagaatcg atagaactta taaatgagat tcagagattt      240
ctaaaagagg agactagact aggtataccg gcaataattc acgaagagtg tctttcgggt      300
gtcatgacca aaggagcaac aacatttcca caggctatag ggatggctag cacatttgag      360
ccagacgata ttcaaagaat gacttccatt ataagaaaag agatgaaggc atttggtgtt      420
catcaagggc tctctcccg tcttgatatt cctagggacc ctatagggg aaggacagaa      480
gaaacatttg gagaagaccc atatcttgta tcaaagatgg cagaaagtta tataaaaggg      540
cttcaggggg aagattggaa ggaaggtata attgctaccg ctaaactttt tacagcttac      600
ggtatatcgg aaggtggtag aaatctgggc cctgccagag tatctgaaag ggaattaaga      660
gaggtattcc tatttccatt tgaagtcgca ataagaaaag ccaatgctgg ttctgtaatg      720
aatgcctatc atgagataga tggggtccca tgtgcgtctt ccaagttttt actgacaaaa      780
atccttagag aagaatgggg attcaaagga tttgtagtct cagactattc agctatagag      840
atgttacata cattccacaa agtagctaaa gacctaaaag ctgcagctat aaaagcccta      900
gaggcaggaa ttgaaataga attaccagag ataaaatggt atggtgaacc acttttatca      960
gcagtcaaaag aaggaaaggt atctgtctct gttagataga cgcgtgtagc tagagtctta      1020
agggcaaaga tactcctagg actattagat gatattatct acgtagatcc aagtaagata      1080
agagccgtct tagataaccc agagcatagg gcatttgcca gagaactagc tagaaaaatct      1140
attgttttat taaagaatga tgggaatacta cccataagta aaggggtaaa aaccattgca      1200
gttataggtc cgagtcgaga tagcacaaaa aatctacatg gagactatag ttataacctca      1260
catatagcag gagtagcgga cgggtgaagg acagtaacag tgtttagagg tatcaaaaaat      1320
aaagtgtcct caggaaccac tgtcctctat gctaaagggt gtgaacttag tgatgaatct      1380
agagaaggat ttaaagaggg attagatata gctagtagat cagatgtcat tatagcgggtg      1440
atggggaaaa atagcggact atttaagaga ggaatctcag gagaaggtaa tgacaggata      1500
gatctaaaaa ttccaggggt tcaagaggaa cttctaaagg ctttaaagga agtgggtaaa      1560
ccaattgtgc tagttctagt taatggaagg cccctctcta taaagtggga gaaagaaaaat      1620
atcccagcta ttctagaagt atggtatcca ggggaagaag gaggaacgc aatagcagat      1680
gtaatatatt gagattataa cccaggaggg aagttaccaa tctcattccc aaaggatgtt      1740
ggacaaatac cagtctacta taatgaaagg ccatctgcct ttagtgaata cttacaaca      1800
gatactaaac ccttgtttcc atttgacat ggcttgagtt atacaacctt tgagtattca      1860
gaactaaaaa taatcccaga aaatgtcatg cccgggggat atgtagacat aagctttaaa      1920
gtaagaaata caggcaatat agacggagat gaagtagtgc aactttatat acatgatgaa      1980
tgggcaagtg tagaaagacc aataaaagaa ctaaaagggt ttaagaggat tcacctaaaa      2040
gcaagagaag aaaagaagct caccctcaga ttatttacag accagttagc attttatgat      2100
gaagtgatgc gatttgggtt agaagcaggg acatttgaag ttatggtagg gtcctcctca      2160
gaagatataa gattaacagg taagtttgag gtcttagaga ccaaggttat aacaaaggat      2220
agaaaatttg ctacgcagggt aattatagaa ccttaa
2256
<210> 7
<211> 23
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.
<400> 7
atggaagaaa gatggttaca aag
23

<210> 8
<211> 26
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.

```

<400> 8
ttaaaggttct ataattacct cgctag

26

<210> 9
<211> 27
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.
<400> 9

caccatggaa gaaagatggt tacaaag

27

<210> 10
<211> 762
<212> PRT
<213> Dictyoglomus thermophilum
<220>
<223> GH3 xylosidaza
<400> 10
Met Glu Glu Lys Glu Leu Ser Lys Lys Val Lys Asp Leu Ile Ala Lys
1 5 10 15
Met Thr Leu Glu Glu Lys Ile Ala Gln Leu Gln Ala Val Tyr Gly Lys
20 25 30
Asp Leu Val Asp Glu Asn Gly Asn Phe Ser Glu Glu Lys Ala Glu Lys
35 40 45
Leu Leu Lys Asn Gly Ile Gly Gln Ile Ser Arg Val Ala Gly Glu Arg
50 55 60
Gly Val Ser Pro Glu Lys Ala Val Glu Leu Ala Asn Lys Ile Gln Lys
65 70 75 80
Phe Leu Lys Glu Lys Thr Arg Leu Gly Ile Pro Ala Ile Ile His Glu
85 90 95
Glu Cys Leu Ser Gly Phe Met Ala Gln Gly Ala Thr Val Phe Pro Gln
100 105 110
Ala Ile Gly Met Ala Ser Thr Phe Glu Pro Glu Leu Ile Arg Arg Val
115 120 125
Ser Asp Val Ile Arg Gln His Met Lys Ala Ala Asn Val His Gln Gly
130 135 140
Leu Ser Pro Val Leu Asp Ile Pro Arg Asp Pro Arg Trp Gly Arg Thr
145 150 155 160
Glu Glu Thr Phe Gly Glu Asp Pro Tyr Leu Val Ser Arg Met Ala Thr
165 170 175
Glu Tyr Val Lys Gly Leu Gln Gly Glu Asp Trp Arg Glu Gly Ile Val
180 185 190
Ala Thr Val Lys His Phe Thr Ala Tyr Gly Ile Ser Glu Gly Ala Arg
195 200 205
Asn Leu Gly Pro Ala Lys Val Gly Glu Arg Glu Leu Arg Glu Val Phe
210 215 220
Leu Phe Pro Phe Glu Val Ala Ile Lys Glu Gly Gln Ala Gly Ser Leu
225 230 235 240
Met Asn Ala Tyr His Glu Ile Asp Gly Val Pro Cys Ala Ser Ser Lys
245 250 255
Phe Leu Leu Thr Lys Ile Leu Arg Trp Glu Trp Gly Phe Lys Gly Tyr
260 265 270
Val Val Ser Asp Tyr Ile Ala Val Arg Met Leu Glu Asn Phe His Lys
275 280 285
Val Ala Arg Asp Ala Lys Glu Ala Ala Val Leu Ala Leu Glu Ala Gly
290 295 300
Ile Asp Ile Glu Leu Pro Ser Val Asp Cys Tyr Gly Glu Pro Leu Ile
305 310 315 320
Gln Ala Val Lys Glu Gly Leu Ile Ser Glu Glu Val Ile Asn Ala Ser
325 330 335
Val Glu Arg Val Leu Arg Ala Lys Phe Met Leu Gly Leu Phe Asp Asp
340 345 350
Asn Leu Glu Lys Asp Pro Lys Lys Val Tyr Glu Val Phe Asp Lys Pro
355 360 365
Glu Phe Arg Asp Leu Ser Arg Glu Val Ala Arg Arg Ser Ile Val Leu
370 375 380
Leu Lys Asn Asp Gly Thr Leu Pro Leu Ser Lys Asn Leu Lys Lys Val
385 390 395 400
Ala Val Ile Gly Pro Asn Ala Asp Asn Pro Arg Asn Leu His Gly Asp
405 410 415
Tyr Ser Tyr Thr Ala His Ile Pro Ser Ile Ala Glu Gly Leu Glu Gly
420 425 430
Val Lys Val Glu Glu Lys Cys Val Val Arg Thr Val Ser Ile Leu Glu
435 440 445
Gly Ile Arg Asn Lys Val Ser Pro Glu Thr Glu Val Leu Tyr Ala Lys
450 455 460
Gly Cys Asp Ile Ile Ser Asp Ser Lys Asp Gly Phe Ala Glu Ala Ile

```

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

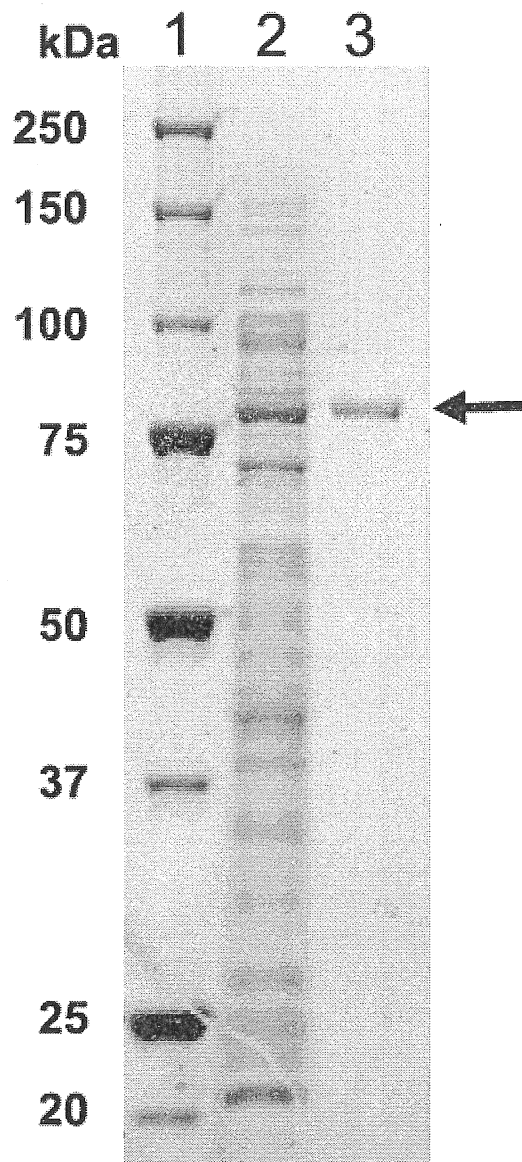
```

Fig.1

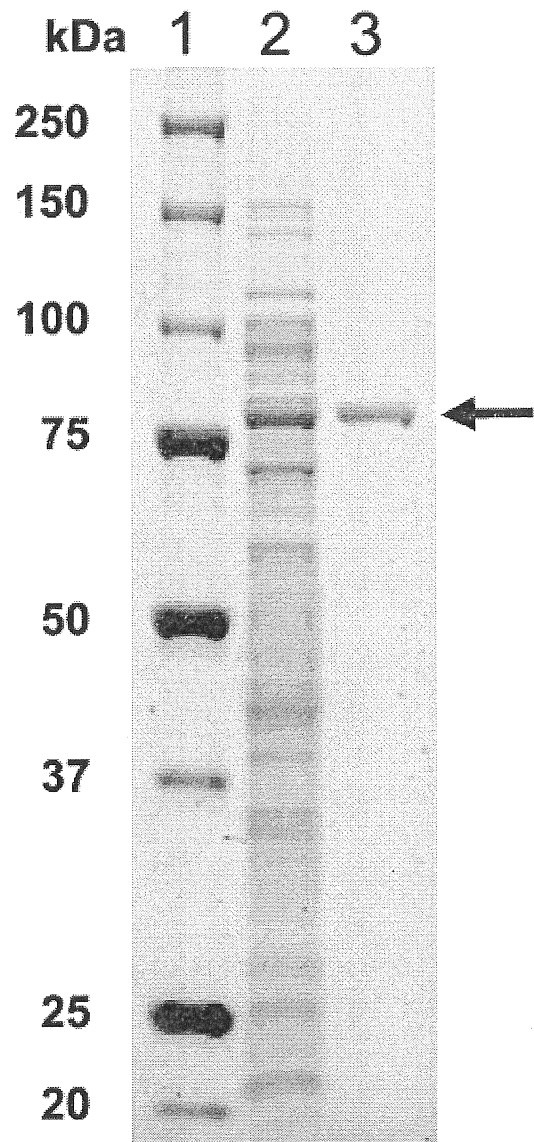
AR19W-311-2	Dictyoglomus thermophilum	1	MEERWQRRVEELISKMTLEEKTAOLGSPSGKLVE-NGKFSREKAKELLEKNGIGQITRVAGYAEREPEESIELI	74
AR19W-311-2	Dictyoglomus thermophilum	1	MEEEKLSKKVKDLIAKMTLEEKTAOLQAVGKDLVDENGNFSEEKAEKLLKNGIGQISRVAGERGVSPKAVELA	75
AR19W-311-2	Dictyoglomus thermophilum	75	NEIQRFLEKEETRLGIPATIIHEECISGMITKGATTEPOAIGMASTFEEDDITQRTSILIRKEMKAFGVHQGLSPVLD	149
AR19W-311-2	Dictyoglomus thermophilum	76	NKIQKFLKEKTRLGIPATIIHEECISGMFACGATVFPOAIGMASTFEPELIRRVSDVIRQIMKAANVHQGLSPVLD	150
AR19W-311-2	Dictyoglomus thermophilum	150	IPRDPRWGRTEETFGEDPYLVSKMAESYIKKLGQEDWKEGIIATIAKHFTAYGISSEGGRNILGPARYSERELREVFL	224
AR19W-311-2	Dictyoglomus thermophilum	151	IPRDPRWGRTEETFGEDPYLVSRMATEYVKKLGQEDWREGIVATVKKHFTAYGISSEGARNILGPAKVGERELREVFL	225
AR19W-311-2	Dictyoglomus thermophilum	225	FPFEVATRKANAGSYMAYHEIDGVPCCASSKFLTKILREEWGFKGFVWSDYSATIEMLHTFHKVAKDILKTAATKA	299
AR19W-311-2	Dictyoglomus thermophilum	226	FPFEVATKKGQAGSLMAYHEIDGVPCCASSKFLTKILRMENGHKGVWSDYIAVRMLENFHKVARDAKFAAVLA	300
AR19W-311-2	Dictyoglomus thermophilum	300	LGAGIIEIETPEIKCYGEPLISAVKEGKVSVDITAVARVLRAKTILGLLDDITVADPSKIRAMLDNPEIRAFAR	374
AR19W-311-2	Dictyoglomus thermophilum	301	LGAGIITELPSVDICYGEPLIQAVKEGLISEEVTVASVERVLRAKFMILGLFDDNLEKDPKKVYEVFDKPEFRDLSR	375
AR19W-311-2	Dictyoglomus thermophilum	375	ELARKSIVLLKNDGILPIISKGVKTIAVIGPSADSTKNLHGDYSYTSIHAGVADG-----VRTVTVLEGI	438
AR19W-311-2	Dictyoglomus thermophilum	376	EVARKSIVLLKNDGILPIISKNLKAVIGPNADNPNLHGDYSYTAHIPSIAEGL EGVKVEEKCVRVTSILEGI	450
AR19W-311-2	Dictyoglomus thermophilum	439	KNKVSSGTTVLVYAKGCELSDESREGFKEALDIASRSDVITIAVMGENSGLFKRGISGEGNDRIDLKIPGVQEEELK	513
AR19W-311-2	Dictyoglomus thermophilum	451	RNKVSPETEVLYAKGCDIISDSKDGFAEATEYAK EADVITIAVMGEEESGLFHRGISGEGNDRITILELFGVQRDLK	525
AR19W-311-2	Dictyoglomus thermophilum	514	ALKEVGKPTVLVLNGRPLSTIKWEKENIPAILLEVNPGEEGGNATADVIFGDYNPGGKLPISEFPKDVGOIPVYVN	588
AR19W-311-2	Dictyoglomus thermophilum	526	ELHKLGPVLVLINGRPOALKWEHENNATLEAVNPGEEGGNADVIFGDYNPSGKLPISEFPVATGQIPVYVN	600
AR19W-311-2	Dictyoglomus thermophilum	589	RKPSAFSEYITDTTKPLPFEGHGLSYTTFEYSLEKIIPENVMPGGYDIDISFKVRNTGNIDGGEWQLYTHIDEMAS	663
AR19W-311-2	Dictyoglomus thermophilum	601	RKPSAFSEYIDESA KPLPFEGHGLSYTTFEYSLEKISPEKVNLSLEKVEISFTIKNTGNRDGGEWQLYTHIDEMAS	675
AR19W-311-2	Dictyoglomus thermophilum	664	VERPIKELKGFKRIHLKAREEKKVTERLITDQLAFYDEVMRFWWEAGTFEVMVGSSSEDIRLITGKFEVLETKVIT	738
AR19W-311-2	Dictyoglomus thermophilum	676	LERPVKELKGFKKLYLKPGE SKRVTFILTYEQLAFYDEFMRFIVEKGVFEVMIGSSSEDIRLNGIFEVLETKVIT	750
AR19W-311-2	Dictyoglomus thermophilum	739	KDRKFAS E VITIEP	751
AR19W-311-2	Dictyoglomus thermophilum	751	EKRKFAS V KVE	762

Fig.2

AR19M -311	1	MEERWLQRRVEELL	80
AR19M -311-2	1		80
AR19M -311-11	1		80
AR19M -311	81	LKEETRLGIPALIIHEECLSGVMTKGATTFPQAIGMASTFEPPDDIQRMTSIIRKEMKAFGVHQGLSPVLDIPRDPRWGRTE	160
AR19M -311-2	81		160
AR19M -311-11	81		160
AR19M -311	161	ETFGEDPYLVSKMAESYIKGLQGDMAREGIIATVKKHFTAYGISGEGARNLGPARGSERELREVFLFPFEVAIRKANAGSVM	240
AR19M -311-2	161		240
AR19M -311-11	161		240
AR19M -311	241	NAYHEIDGVPCASSKFLLTkILREEWGFKGFVVSdysaIEMLHTFHkVAKDLKTAAlKALEAGIEIElPEIKCYGEPLLS	320
AR19M -311-2	241		320
AR19M -311-11	241		320
AR19M -311	321	AVKEGKVSVDtAVARVLRAKILLGLDDIIVDPSKIRAVLDNPEHRAFARELARKSIVLLKNDGILPISKGVKTIA	400
AR19M -311-2	321		400
AR19M -311-11	321		400
AR19M -311	401	VIGPSADSTKNLHGdysYtSHIAGVADGVRtVtVLEGIKNKVSSGtTVLYAKGCeLSDESREGfKEALDIASRSdVtIAV	480
AR19M -311-2	401		480
AR19M -311-11	401		480
AR19M -311	481	MGENSGLFKRGISGEGNDRIDLKLPGVQFEllKALKEVGKPTVlLVNGRPLSIKWEKENIPAILVWYPGEEGNAIAD	560
AR19M -311-2	481		560
AR19M -311-11	481		560
AR19M -311	561	VIFGDYNPGGKLPIsFPKdVgQIPVYNNRKPSAFSEYLTtTDTKPLFPFGHGLSYtTFEYSELKIIPENVMPPGGYVDISFK	640
AR19M -311-2	561		640
AR19M -311-11	561		640
AR19M -311	641	VRNTGNIDGDEVVQLYtHDEWASVERPIKELKGfKRiHLKAREEKKVtFRLFDQLAFYDEVMRfVVEAGtFEVMVGSSS	720
AR19M -311-2	641		720
AR19M -311-11	641		720
AR19M -311	721	EDIRLTGKFVLETKVITKDRKFASEVtIEP	751
AR19M -311-2	721		751
AR19M -311-11	721		751

Fig.3A

AR19M-311-2

Fig.3B

AR19M-311-11

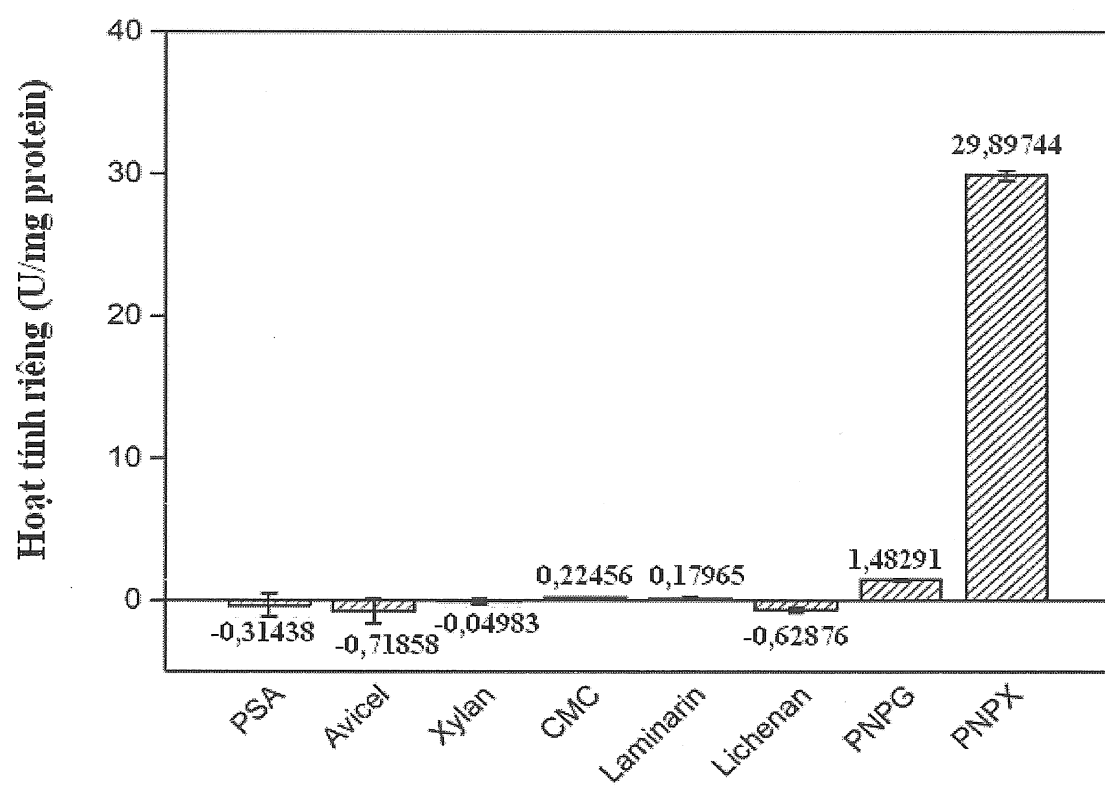
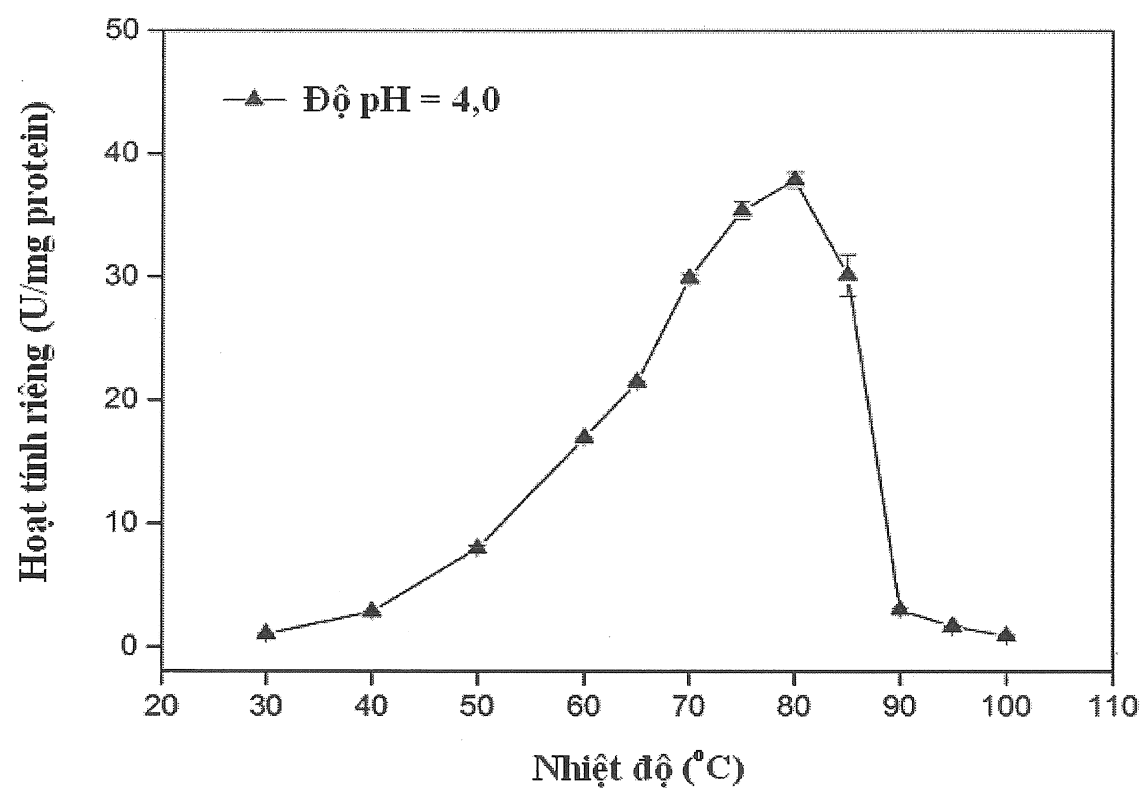
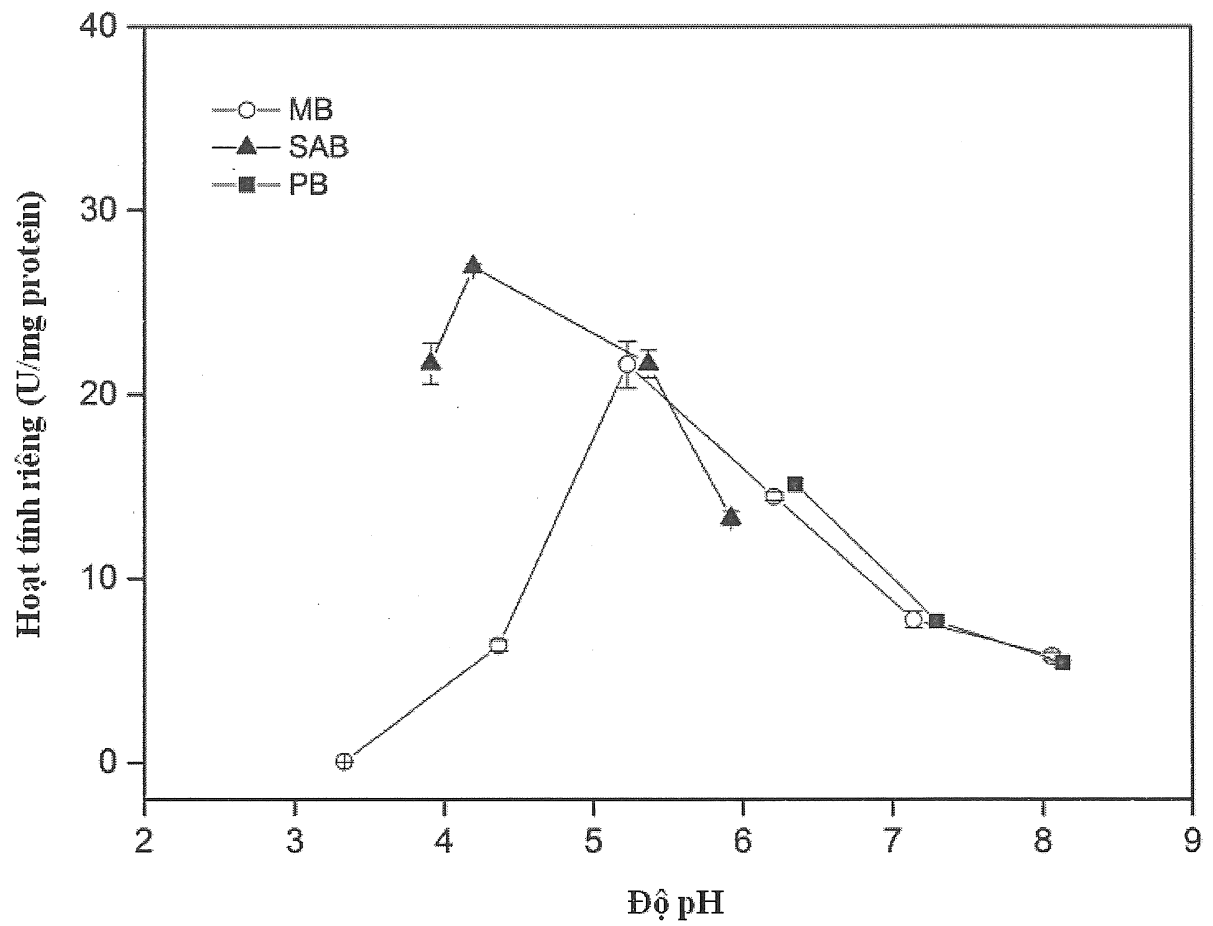
Fig.4**Fig.5**

Fig.6



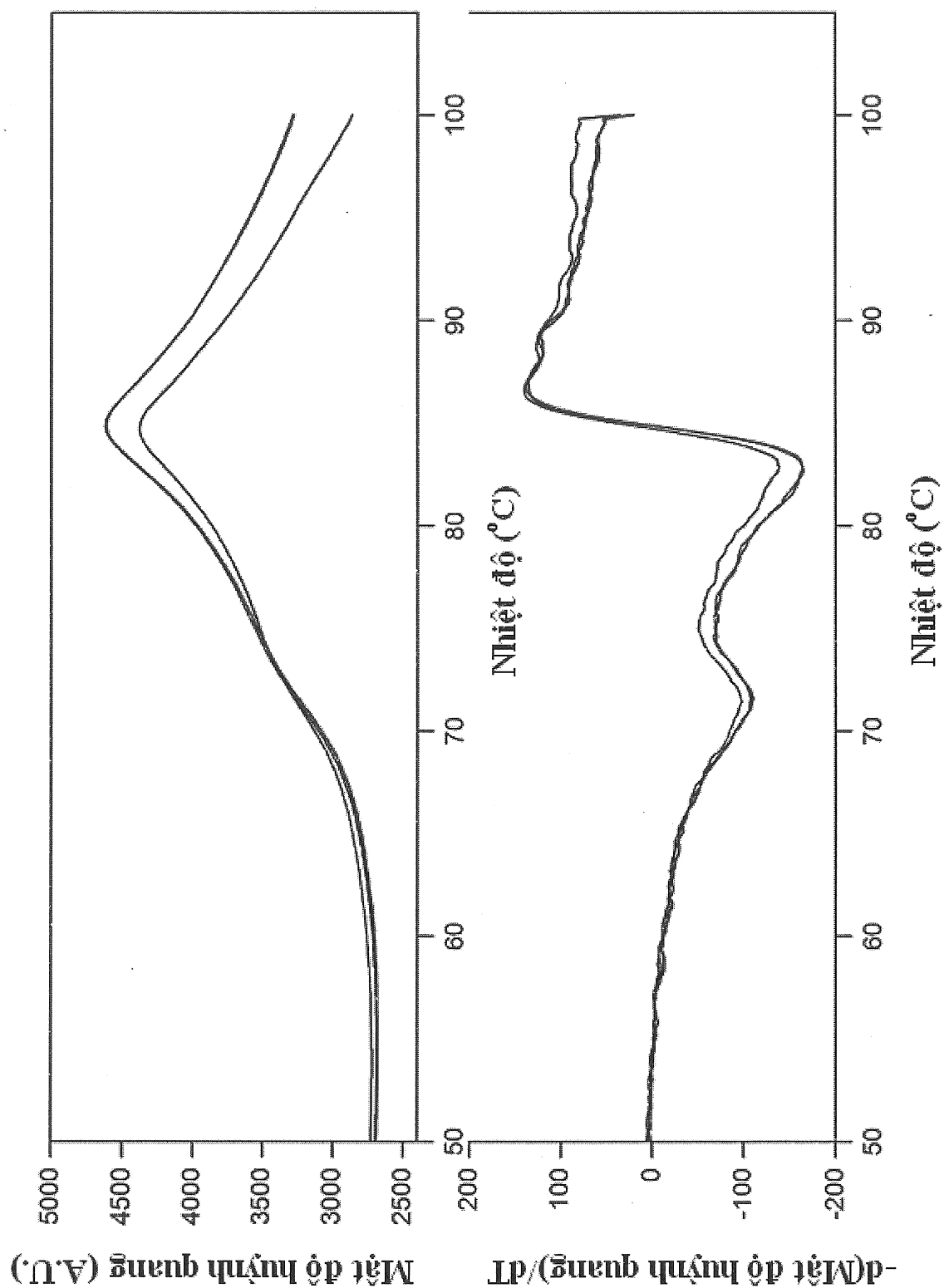


Fig. 7A

Fig. 7B