



(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)



1-0026973

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

(13) B

(21) 1-2015-02722

(22) 27/07/2015

(30) 2014-158650 04/08/2014 JP

(45) 25/01/2021 394

(43) 25/02/2016 335A

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

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

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

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

(54) ENZYM BETA-XYLOSIDAZA ( $\beta$ -XYLOSIDAZA) CHỊU NHIỆT

(57) Sáng chế đề xuất enzym beta-xylosidaza ( $\beta$ -xylosidaza) chịu nhiệt chứa polypeptit phân lập 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.

### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến enzym beta-xylosidaza ( $\beta$ -xylosidaza) chịu nhiệt, 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.

### **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 này, quá trình thủy phân xenluloza thường cần ba loại enzym: endoglucanaza glucosit hydrolaza (endo-1,4- $\beta$ -D-glucanaza, EC 3.2.1.4), exo-xenlobiohydrolaza (1,4- $\beta$ -xenlobiosidaza hoặc xenlobiohydrolaza, EC 3.2.1.91, EC 3.2.1.176), và  $\beta$ -glucosidaza (EC 3.2.1.21). Mặt khác, mặc dù cấu trúc của hemixenluloza có thể thay đổi phụ thuộc vào loại thực vật, ví dụ, xylan là thành phần chính có trong thực vật lá rộng, thực vật thân thảo và các thực vật tương tự. Để thủy phân xylan, thì cần có xylanaza (endo-1,4- $\beta$ -xylanaza, EC 3.2.1.8) và  $\beta$ -xylosidaza (EC 3.2.1.37).  $\beta$ -xylosidaza là 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 có nồng độ chất rắn cao nằm trong khoảng từ 30% đến 60% đã được dùng để đạt

được hiệu quả năng lượng cao hơn và sử dụng ít nước hơn. Tuy nhiên, khi nồng độ chất rắn trong cơ chất cao, thì phản ứng enzym thủy phân lignoxenluloza hầu như không thể xảy ra do độ nhớt cao của dung dịch sinh khối đã thủy phân cao. Do đó, bằng cách thực hiện phản ứng thủy phân ở nhiệt độ 80°C hoặc cao hơn và sử dụng enzym chịu nhiệt, thì ngoài việc tăng tốc độ phản ứng thủy phân, độ nhớt của dung dịch sinh khối đã thủy phân cũng được giảm xuống, nhờ vậy có thể rút ngắn được thời gian thủy phân và giảm lượng enzym cần sử dụng. Do đó, có nhu cầu phát triển các enzym có độ ổn định nhiệt cao

Nhiều glycosit hydrolaza chịu nhiệt đã được thu nhận bằng cách phân lập và xác định các vi sinh vật ưa nhiệt sống trong môi trường có nhiệt độ cao, tách dòng các gen từ các vi sinh vật đã được phân lập và nuôi cấy này, và xác định trình tự ADN của chúng, sau đó biểu hiện chúng bằng *E. coli*, nấm sợi và các sinh vật tương tự. Ví dụ,  $\beta$ -xylosidaza có nguồn gốc từ nấm sợi và  $\beta$ -xylosidaza có nguồn gốc từ nấm sợi *Aspergillus oryzae* có hoạt tính enzym ở nhiệt độ 30°C lần lượt được bộc lộ trong Tài liệu sáng chế 1 và Tài liệu sáng chế 2.  $\beta$ -xylosidaza có nguồn gốc từ *Alicyclobacillus acidocaldarius* có hoạt tính enzym ở nhiệt độ 50°C hoặc cao hơn và độ pH = 5,5 hoặc thấp hơn được bộc lộ trong Tài liệu sáng chế 3.  $\beta$ -xylosidaza có nguồn gốc từ *Acremonium cellulolyticus* có hoạt tính enzym ở nhiệt độ 45°C được bộc lộ trong Tài liệu sáng chế 4. Ngoài ra,  $\beta$ -xylosidaza được phân lập từ một số vi khuẩn và nấm sợi có nhiều độ tối ưu khoảng 60°C được bộc lộ trong Tài liệu phi sáng chế 1 đến 6.

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  $\beta$ -xylosidaza chịu nhiệt thể hệ mới có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0, 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  $\beta$ -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.

Enzym  $\beta$ -xylosidaza chịu nhiệt, polynucleotit, vector biểu hiện, thể biến nạp, phương pháp sản xuất enzym  $\beta$ -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.

1. Enzym  $\beta$ -xylosidaza chịu nhiệt chứa vùng xúc tác  $\beta$ -xylosidaza bao gồm:

(A) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1;

(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, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0; hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 80% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

2. Enzym  $\beta$ -xylosidaza chịu nhiệt theo phương án 1, trong đó enzym này còn có hoạt tính  $\beta$ -glucosidaza.

3. Polynucleotit chứa miền mã hóa vùng xúc tác  $\beta$ -xylosidaza bao gồm:

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

(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, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0;

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 80% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0;

(d) trình tự nucleotit có độ tương đồng ít nhất bằng 80% so với trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3, và mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0; hoặc

(e) trình tự nucleotit của polynucleotit mà lai hóa với polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

4. Polynucleotit theo phương án 3, trong đó polypeptit này cũng có hoạt tính  $\beta$ -glucosidaza.

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

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

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

8. Phương pháp sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt bao gồm bước sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt trong thể biến nạp theo phương án 6 hoặc 7.

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

10. Phương pháp sản xuất sản phẩm phân giải lignoxenluloza bao gồm bước cho nguyên liệu lignoxenluloza chứa xenluloza, hemixenluloza và lignin tiếp xúc với enzym  $\beta$ -xylosidaza chịu nhiệt theo phương án 1 hoặc 2, enzym  $\beta$ -xylosidaza chịu nhiệt mã hóa bởi polynucleotit theo phương án 3 hoặc 4, thể biến nạp theo phương án 6 hoặc 7, enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 8, hoặc hỗn hợp glycosit hydrolaza theo phương án 9.

Enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit (sau đây được gọi tắt là PNPX) ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0. Do đó, enzym  $\beta$ -xylosidaza chịu nhiệt này thích hợp để thủy phân nguyên liệu lignoxenluloza ở nhiệt độ cao.

Ngoài ra, enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế cũng thích hợp để thủy phân nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic ở nhiệt độ cao.

Mặt khác, enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế cũng thích hợp để thủy phân oligosacarit có liên kết  $\beta$ -xylosidic hoặc nguyên liệu chứa oligosacarit có liên kết  $\beta$ -glycosidic ở nhiệt độ cao.

Nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa hemixenluloza bằng xylanaza. Nguyên liệu chứa oligosacarit có liên kết  $\beta$ -glycosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa xenluloza bằng xenlobiohydrolaza.

Ngoài ra, polynucleotit, vectơ biểu hiện chứa polynucleotit này và thể biến nạp chứa vectơ này theo sáng chế cũng thích hợp để sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế.

### Mô tả vắn tắt hình vẽ

Fig.1 là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin (SEQ ID NO.1) của vùng xúc tác hoạt tính  $\beta$ -xylosidaza của sản phẩm tách dòng gen AR19M-346-18 và trình tự axit amin của  $\beta$ -xylosidaza GH3 (SEQ ID NO.7) có nguồn gốc từ *Fervidobacterium gondwanense*.

Fig.2 là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein AR19M-346-18 thu được bằng cách biểu hiện gen AR19M-346-18 ở *E. coli* trong ví dụ 1.

Fig.3 là biểu đồ thể hiện kết quả đo hoạt tính thủy phân mỗi cơ chất của protein AR19M-346-18 được biểu hiện ở *E. coli* trong ví dụ 1.

Fig.4 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở độ pH = 6,0) của protein AR19M-346-18 được biểu hiện ở *E. coli* ở các nhiệt độ tương ứng trong ví dụ 1.

Fig.5 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở nhiệt độ 70°C) của protein AR19M-346-18 được biểu hiện ở *E. coli* ở các độ pH tương ứng trong ví dụ 1.

Fig.6 là đồ thị thể hiện sự thay đổi mật độ huỳnh quang của thuốc nhuộm màu cam SYPRO do sự biến tính nhiệt của protein AR19M-346-18 được biểu hiện ở *E. coli* trong ví dụ 1.

## Mô tả chi tiết sáng chế

### Enzym $\beta$ -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 khoảng 99% vi sinh vật sống trong môi trường như đất là chưa được biết đến. Đặc biệt, rất khó nuôi cấy được vi sinh vật sống trong môi trường có nhiệt độ cao, và chỉ 0,1% hoặc 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 tiếp theo cho phép giải trình tự được hàng tỷ cặp bazơ, nên đã có thể giải trình tự bộ gen của vi sinh vật chứa trong đất và các môi trường tương tự. Bằng cách sử dụng kỹ thuật phân tích này, phương pháp phân tích hệ gen môi trường đã được đề xuất, trong đó ADN bộ gen của quần thể vi sinh vật được tổng hợp từ mẫu môi trường như đất, bộ gen của quần thể vi sinh vật có thành phần không đồng nhất được giải trình tự trực tiếp và toàn bộ, và dữ liệu trình tự được lắp ráp bằng máy tính song song, để thiết kế lại trình tự bộ gen của vi sinh vật. Điều này góp phần rút ngắn thời gian giải trình tự bộ gen của vi sinh vật khó nuôi cấy.

Như được thể hiện ví dụ 1, các tác giả sáng chế đã tổng hợp được ADN bộ gen (ADN hệ gen môi trường) của quần thể vi sinh vật sống trong đất có nhiệt độ cao thu được ở khu vực suối nước nóng (ví dụ, mẫu nước thu được ở khu vực suối nước nóng có nhiệt độ nằm trong khoảng từ 58°C đến 78°C chứa đất, bùn, thảm vi sinh vật, màng sinh học và các loại tương tự), và thực hiện giải trình tự ADN hệ gen môi trường này. Kết quả là thu được 405 khung đọc mã mở (Open Reading Frame - ORF) mã hóa các trình tự axit amin tương đồng với các  $\beta$ -glucosidaza hoặc  $\beta$ -xylosidaza đã biết (ví dụ, các trình tự axit amin có độ tương đồng bằng 20% hoặc cao hơn và trị số Kỳ vọng (tức là trị số E) nhỏ hơn  $1e^{-20}$ ). Trong số các khung đọc mã mở này, các đoạn mã được thiết kế dựa trên thông tin trình tự nucleotit của 167 khung đọc mã mở có chiều dài đầy đủ trong đó sự có mặt của vùng xúc tác  $\beta$ -glucosidaza hoặc vùng xúc tác  $\beta$ -xylosidaza có thể được xác định, các gen tiềm năng được tách dòng từ ADN hệ gen môi trường của đất có nhiệt độ cao ở khu vực suối nước nóng bằng phương pháp PCR. Các ADN được tách dòng PCR được biến nạp vào *E. coli*, và các protein mã hóa bởi các trình tự nucleotit này được biểu

hiện. Các protein này được sàng lọc chức năng bằng cách đo hoạt tính phân giải PNPX. Cuối cùng, enzym  $\beta$ -xylosidaza chịu nhiệt có hoạt tính phân giải PNPX (sau đây được gọi tắt là “AR19M-346-18”) thu được từ các khung đọc mã mở này. Trình tự axit amin của AR19M-346-18 và trình tự nucleotit mã hóa trình tự axit amin của AR19M-346-18 lần lượt như nêu trong SEQ ID NO.1 và SEQ ID NO.3.

Như được thể hiện trong Ví dụ 1, AR19M-346-18 có hoạt tính thủy phân cả PNPX lẫn p-nitrophenyl- $\beta$ -D-glucopyranosit (sau đây gọi tắt là PNPG). Trái lại, AR19M-346-18 hầu như không có hoạt tính phân giải Avicel trương nở liên hợp axit phosphoric (phosphoric acid swollen Avicel - PSA), carboxymetyl xenluloza (CMC), laminarin chứa  $\beta$ -1,3 glucan và  $\beta$ -1,6 glucan, lichenan chứa  $\beta$ -1,3 glucan và  $\beta$ -1,4 glucan, và xenluloza tinh thể Avicel. Từ kết quả phân tích độ đặc hiệu cơ chất này chứng tỏ rằng AR19M-346-18 là glycosit hydrolaza có hoạt tính  $\beta$ -xylosidaza.

Thuật ngữ “hoạt tính  $\beta$ -xylosidaza” được dùng trong bản mô tả để chỉ hoạt tính thủy phân cơ chất PNPX.

Ngoài ra, “hoạt tính  $\beta$ -glucosidaza” được dùng trong bản mô tả để chỉ hoạt tính thủy phân cơ chất PNPG.

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

Kết quả tra cứu trình tự axit amin của AR19M-346-18 bằng cơ sở dữ liệu trình tự axit amin đã biết cho thấy trình tự axit amin của AR19M-346-18 có độ tương đồng cao nhất với trình tự axit amin của  $\beta$ -xylosidaza GH3 (SEQ ID NO.7) (mã truy cập Ngân hàng Gen: AFY97406.1) có nguồn gốc từ *Fervidobacterium gondwanense*, và bằng 76%.

Các phân tích độ đặc hiệu cơ chất, và phân tích độ tương đồng của trình tự axit amin này với các protein đã biết cho thấy AR19M-346-18 là  $\beta$ -xylosidaza GH3 thế hệ mới.

AR19M-346-18 có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit (tức là hoạt tính  $\beta$ -xylosidaza) ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0. Cụ thể, như được thể hiện trong phần 10 của ví dụ 1, AR19M-346-18 có hoạt tính  $\beta$ -xylosidaza ở nhiệt độ nằm trong khoảng từ 50°C đến 95°C, và có hoạt tính  $\beta$ -xylosidaza mạnh ở nhiệt độ nằm trong khoảng từ 70°C đến 90°C. Cụ thể hơn, hoạt tính  $\beta$ -xylosidaza của AR19M-346-18 tăng khi nhiệt độ tăng nằm trong khoảng từ 50°C đến 85°C, và giảm nhanh khi nhiệt độ trên 90°C.

Nhìn chung, ở protein có một số hoạt tính sinh học, một hoặc hai hoặc nhiều axit amin có thể được loại bỏ, thay thế, hoặc bổ sung, mà không làm giảm hoạt tính sinh học của protein này. Nói cách khác, trong AR19M-346-18, một hoặc hai hoặc nhiều axit amin cũng có thể được loại bỏ, thay thế, hoặc bổ sung mà không làm giảm hoạt tính thủy phân glycosit, bao gồm hoạt tính  $\beta$ -xylosidaza.

Enzym  $\beta$ -xylosidaza theo sáng chế là glycosit hydrolaza chịu nhiệt chứa vùng xúc tác  $\beta$ -xylosidaza bao 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 (tức là AR19M-346-18);

(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, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0; hoặc

(C) polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 80% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

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

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

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

Trong polypeptit (B), tốt hơn nếu số lượng axit amin được loại bỏ, thay thế, hoặc bổ sung trên trình tự axit amin như nêu trong SEQ ID NO.1 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 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%, còn tốt hơn nữa nếu nằm trong khoảng từ 98% đến 100%.

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

Các polypeptit (B) và (C) có thể được thiết kế nhân tạo hoặc cũng có thể là thể tương đồng của AR19M-346-18 và các biến thể tương tự, hoặc các đoạn protein của chúng.

Các polypeptit (A) đến (C) có thể lần lượt được tổng hợp hóa học dựa trên trình tự axit amin, hoặc cũng có thể được sản xuất bởi hệ biểu hiện protein sử dụng polynucleotit theo sáng chế. Ngoài ra, các polypeptit (B) và (C) cũng có thể lần lượt được tổng hợp nhân

tạo dựa trên polypeptit có trình tự axit amin như nêu trong SEQ ID NO.1 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- $\beta$ -D-xylopyranosit (hoạt tính  $\beta$ -xylosidaza) ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0. Do đó, Khi enzym chứa vùng xúc tác  $\beta$ -xylosidaza là polypeptit bất kỳ từ (A) đến (C), thì có thể thu được  $\beta$ -xylosidaza chịu nhiệt.

Enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất PNPX. Enzym  $\beta$ -xylosidaza chịu nhiệt này cũng có hoạt tính thủy phân cơ chất  $\beta$ -glucan hoặc oligosacarit khác. Ví dụ về các cơ chất khác của enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế bao gồm PNPG, p-nitrophenyl- $\alpha$ -L-arabinofuranosit, p-nitrophenyl- $\alpha$ -L-arabinopyranosit, p-nitrophenyl- $\beta$ -L-arabinopyranosit, p-nitrophenyl- $\beta$ -D-mannopyranosit, p-nitrophenyl- $\alpha$ -D-galactopyranosit, p-nitrophenyl- $\beta$ -D-galactopyranosit; glucan chứa liên kết  $\beta$ -1,3 và  $\beta$ -1,4 như lichenan; xylan; glucan chứa liên kết  $\beta$ -1,4 bao gồm 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 trương nở liên hợp axit phosphoric (phosphoric acid swollen Avicel - PSA), và CMC; oligosacarit chứa liên kết  $\beta$ -1,4 bao gồm xenlobioza; glucan chứa liên kết  $\beta$ -1,3 và  $\beta$ -1,6 như laminarin; glucan chứa liên kết  $\beta$ -1,3; glucan chứa liên kết  $\beta$ -1,6 và oligosacarit chứa liên kết  $\beta$ -1,6 như gentiobioza.

Ngoài PNPX, thì enzym  $\beta$ -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- $\alpha$ -L-arabinofuranosit, p-nitrophenyl- $\alpha$ -L-arabinopyranosit, p-nitrophenyl- $\beta$ -L-arabinopyranosit, p-nitrophenyl- $\beta$ -D-mannopyranosit, p-nitrophenyl- $\alpha$ -D-galactopyranosit, và p-nitrophenyl- $\beta$ -D-galactopyranosit, và đặc biệt là enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất PNPG.

Tốt hơn nếu enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính phân giải PNPX (hoạt tính  $\beta$ -xylosidaza) ít nhất ở độ pH bằng 6,0 và nhiệt độ nằm trong khoảng từ 70°C đến 90°C, tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 60°C đến 95°C, còn tốt hơn nữa nếu ở nhiệt độ nằm trong khoảng từ 50°C đến 95°C. Tốt hơn nếu nhiệt độ tối ưu

của enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế nằm trong khoảng từ 70°C đến 90°C và tốt hơn nữa nằm trong khoảng từ 80°C đến 90°C.

Mặc dù thay đổi phụ thuộc vào nhiệt độ phản ứng, nhưng độ pH tối ưu của enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế nằm trong khoảng từ 5,0 đến 7,0. Tốt hơn nếu enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính  $\beta$ -xylosidaza ít nhất ở độ pH nằm trong khoảng từ 4,5 đến 8,0, và tốt hơn nữa nếu enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có hoạt tính phân giải PNPX ở độ pH nằm trong khoảng từ 4,0 đến 8,0.

Enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế cũng có thể có hoạt tính thủy phân glycosit khác với hoạt tính  $\beta$ -xylosidaza. Ví dụ về các hoạt tính thủy phân glycosit khác bao gồm hoạt tính endoglucanaza, hoạt tính xylanaza, hoạt tính  $\beta$ -glucosidaza và hoạt tính xenlobiohydrolaza.

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

Khi enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế chứa vùng khác với vùng xúc tác  $\beta$ -xylosidaza, thì tốt hơn nếu enzym này chứa vùng tương tự Fibronectin loại III.

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

Ngoài ra, enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có thể cũng chứa peptit tín hiệu cho phép định vị enzym này vào vùng đặc hiệu ở tế bào, hoặc peptit tín hiệu tạo ra

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

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

Polynucleotit mã hóa enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế

Polynucleotit nêu trên mã hóa enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế. Enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế có thể được sản xuất bằng hệ biểu hiện của vật chủ thu được 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  $\beta$ -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.

(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, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng ít nhất bằng 80% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

(d) trình tự nucleotit có độ tương đồng ít nhất bằng 80% so với trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3, và mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

(e) trình tự nucleotit của polynucleotit mà lai hóa với polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3 trong điều kiện nghiêm ngặt, và là trình tự nucleotit mã hóa polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0.

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

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

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

Ví dụ về “các điều kiện nghiêm ngặt” theo sáng chế bao gồm phương pháp được mô tả trong tài liệu: *Molecular Cloning - A Laboratory Manual Third Edition* (Sambrook *et al.*, Cold Spring Harbor Laboratory Press). Ví dụ, các điều kiện trong dung dịch đệm lai hóa SSC 6 $\times$  (thành phần của dung dịch SSC 20 $\times$ : 3M natri clorua, 0,3M dung dịch axit xitric, độ pH bằng 7,0), dung dịch Denhardt 5 $\times$  (thành phần dung dịch Denhardt 100 $\times$ : 2% khối lượng albumin huyết thanh bò, 2% khối lượng ficoll, 2% khối lượng polyvinylpyrrolidon), 0,5% khối lượng SDS, 0,1mg/mL ADN tinh trùng cá hồi, và 50% formamit, ở nhiệt độ nằm trong khoảng từ 42°C đến 70°C trong vài giờ đến qua đêm. Sau khi ủ, tốt hơn nếu dung dịch SSC 1 $\times$  chứa 0,1% 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% 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, trình tự nucleotit như nêu trong SEQ ID NO.3, 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 hoặc SEQ ID NO.3. Lưu ý rằng, các bộ ba mã hóa này có thể được cải biến bằng kỹ thuật cải biến trình tự gen hoặc tổng hợp gen nhân tạo đã biết.

Polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3 có thể được tổng hợp hóa học dựa trên thông tin về trình tự nucleotit hoặc polynucleotit này có thể được tổng hợp dưới dạng gen mã hóa AR19M-346-18 có chiều dài đầy đủ (cũng được gọi là “gen AR19M-346-18”) hoặc một đoạn gen của nó chứa vùng xúc tác  $\beta$ -xylosidaza (tức là vùng mã hóa đoạn peptit chứa 325 gốc axit amin từ threonin ở vị trí số 23 đến valin ở vị trí số 347, vùng mã hóa đoạn peptit chứa 260 gốc axit amin từ isoleucin ở vị trí số 379 đến threonin ở vị trí số 638, hoặc vùng mã hóa đoạn peptit chứa 616 gốc axit amin từ threonin ở vị trí số 23 đến threonin ở vị trí số 638 trên SEQ ID NO.1) từ tự nhiên bằng kỹ thuật tái tổ hợp gen. Gen AR19M-346-18 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, cũng như đoạn mỗi xuôi và đoạn mỗi ngược được thiết kế dựa trên trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3. ADN bổ sung (ADNbs) được tổng hợp bằng phản ứng phiên mã ngược từ khuôn là ARN thông tin thu được từ mẫu. Lưu ý rằng, tốt hơn nếu mẫu để thu nhận khuôn axit nucleic là mẫu thu được từ môi trường có nhiệt độ cao, như đất ở khu vực suối nước nóng.

Không có giới hạn cụ thể về độ tương đồng của trình tự nucleotit (d) so với trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3 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%.

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

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

Ví dụ, polynucleotit chứa trình tự nucleotit (b), (c), hoặc (d) có thể lần lượt được tổng hợp nhân tạo bằng cách loại bỏ, thay thế, hoặc bổ sung một hoặc hai hoặc nhiều nucleotit trong polynucleotit có trình tự như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3. 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-346-18 hoặc một đoạn trình tự của nó. Gen tương đồng của gen AR19M-346-18 có thể thu được bằng kỹ thuật gen tái tổ hợp dùng để thu nhận gen tương đồng của gen có trình tự nucleotit đã biết.

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

#### Vector biểu hiện theo sáng chế

Vector nêu trên là vector chứa polynucleotit theo sáng chế và có khả năng biểu hiện polypeptit có hoạt tính thủy phân cơ chất p-nitrophenyl- $\beta$ -D-xylopyranosit ít nhất ở nhiệt độ bằng 85°C và độ pH bằng 6,0 ở tế bào vật chủ. Cụ thể, vector này là vector biểu hiện chứa polynucleotit mã hóa enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế. Đặc biệt, vector biểu hiện cần được kết hợp với cat-xet biểu hiện chứa, từ đoạn nằm trước, ADN có trình tự vùng khởi đầu, polynucleotit theo sáng chế, và ADN có trình tự vùng kết thúc. Lưu ý rằng, polynucleotit theo sáng chế có thể được biến nạp vào vector biểu hiện này bằng kỹ thuật tái tổ hợp gen đã biết. Kit tổng hợp vector biểu hiện có bán trên thị trường cũng có thể được sử dụng để biến nạp polynucleotit vào vector biểu hiện.

Thuật ngữ “vector biểu hiện” được dùng trong bản mô tả để chỉ vector chứa, từ đoạn nằm trước, ADN có trình tự vùng khởi đầu, ADN có trình tự để kết hợp 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 tích với vật chủ có thể được sử dụng làm vector biểu hiện theo sáng chế.

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

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

#### Thế biến nạp theo sáng chế

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

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

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

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

#### Phương pháp sản xuất enzym $\beta$ -xylosidaza chịu nhiệt theo sáng chế

Phương pháp sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế chế là phương pháp để sản xuất enzym  $\beta$ -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 nằm sau vùng khởi đầu mà không có khả năng điều hòa thời gian biểu hiện hoặc các yếu tố tương tự, enzym  $\beta$ -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  $\beta$ -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  $\beta$ -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/tinh chế từ thể biến nạp.

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

#### Hỗn hợp glycosit hydrolaza theo sáng chế

Hỗn hợp glycosit hydrolaza nêu trên chứa enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, hoặc enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, và ít nhất một hoặc nhiều loại glycosit hydrolaza khác.  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -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 hoặc tinh chế từ thể biến nạp này. Khi hỗn hợp glycosit hydrolaza theo sáng chế được sử dụng để thủy phân polysacarit, thì có thể phân giải hiệu quả lignoxenluloza.

Không có giới hạn cụ thể về glycosit hydrolaza khác với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế miễn là nó có hoạt tính thủy phân lignoxenluloza. Ví dụ về glycosit hydrolaza khác bao gồm hemixenlulaza như xylanaza, xenlobiohydrolaza,  $\beta$ -glucosidaza, endoglucanaza, hoặc enzym tương tự. Tốt hơn nếu ngoài enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, thì hỗn hợp glycosit hydrolaza theo sáng chế còn chứa ít nhất một glycosit hydrolaza (tức là hemixenlulaza và endoglucanaza) và tốt hơn nữa nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa cả hai glycosit hydrolaza này (tức là chứa cả hemixenlulaza

lẫn endoglucanaza). Đặc biệt, tốt hơn nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa ít nhất một glycosit hydrolaza được chọn từ nhóm bao gồm xylanaza,  $\beta$ -xylosidaza, xenlobiohydrolaza, và endoglucanaza; và tốt hơn nữa nếu hỗn hợp glycosit hydrolaza theo sáng chế chứa toàn bộ các glycosit hydrolaza (tức là xylanaza,  $\beta$ -xylosidaza, xenlobiohydrolaza, và endoglucanaza).

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

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

Phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế 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  $\beta$ -xylosidaza chịu nhiệt theo sáng chế.

Ví dụ về sản phẩm phân giải lignoxenluloza bao gồm monosacarit, như glucoza và xyloza.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương pháp, trong đó nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic được thủy phân với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế để thu được sản phẩm phân giải lignoxenluloza chứa sản phẩm phân giải của oligosacarit này.

Theo một khía cạnh khác, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương pháp, trong đó nguyên liệu chứa oligosacarit có liên kết  $\beta$ -glycosidic được thủy phân với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế để thu được sản phẩm phân giải lignoxenluloza chứa sản phẩm phân giải của oligosacarit này.

Nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa hemixenluloza với xylanaza. Nguyên liệu chứa oligosacarit có liên kết  $\beta$ -glycosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa xenluloza với xenlobiohydrolaza. Cụ thể hơn, phương pháp sản xuất sản phẩm phân giải lignoxenluloza theo sáng chế là phương pháp bao gồm bước sản xuất sản phẩm phân giải xenluloza hoặc hemixenluloza bằng cách cho nguyên liệu lignoxenluloza, cụ thể hơn, nguyên liệu lignoxenluloza chứa hemixenluloza hoặc xenluloza, tiếp xúc với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -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 cho nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic tiếp xúc với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -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 cho nguyên liệu chứa oligosacarit có liên kết  $\beta$ -glycosidic tiếp xúc với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, hoặc hỗn hợp glycosit hydrolaza theo sáng chế.

Nguyên liệu chứa oligosacarit có liên kết  $\beta$ -xylosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa hemixenluloza với xylanaza, và nguyên liệu chứa

oligosacarit có liên kết  $\beta$ -glycosidic có thể được điều chế bằng cách thủy phân lignoxenluloza chứa xenluloza với xenlobiohydrolaza.

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

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 chứa hemixenluloza hoặc xenluloza được xử lý vật lý, xử lý hóa học, hoặc ngâm hoặc phân giải vào dung dịch đệm thích hợp trước khi tiếp xúc với enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế.

Tốt hơn nếu phản ứng thủy phân hemixenluloza bởi enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế được thực hiện ở điều kiện mà ở đó enzym này có hoạt tính thủy phân xenlooligosacarit. Tốt hơn nếu phản ứng thủy phân hemixenluloza 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 hemixenluloza 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 hemixenluloza 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 hemixenluloza 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 hemixenluloza có thể được thực hiện trong thời gian từ 10 phút đến 100 giờ và khi phân giải sinh khối chứa xenluloza, thì phản ứng thủy phân hemixenluloza có thể được thực hiện trong thời gian từ 1 giờ đến 100 giờ.

Để thủy phân lignoxenluloza, tốt hơn nữa nếu sử dụng ít nhất một hoặc nhiều loại glycosit hydrolaza khác ngoài enzym  $\beta$ -xylosidaza chịu nhiệt theo sáng chế. Glycosit hydrolaza khác có thể giống như glycosit hydrolaza mà có thể chứa trong hỗn hợp glycosit hydrolaza nêu trên, và tốt hơn nếu enzym này là glycosit hydrolaza chịu

hiệt 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  $\beta$ -xylosidaza chịu nhiệt theo sáng chế, thể biến nạp theo sáng chế, hoặc enzym  $\beta$ -xylosidaza chịu nhiệt được sản xuất bằng phương pháp sản xuất enzym  $\beta$ -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  $\beta$ -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  $\beta$ -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 được giải trình tự.

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

ADN được chiết tách bằng kit chiết tách ADN từ 10g mỗi mẫu đất thu được ở các khu vực suối nước nóng (ISOIL Large for Beads ver. 2, do NIPPON GENE Co., Ltd. sản xuất). ADN chiết tách được được giải trình tự bằng máy giải trình tự GS FLX Titan 454 do Roche Diagnostics K.K sản xuất và máy giải trình tự HiSeq 2000 do Illumina, Inc sản xuất. ADN chiết tách được và sản phẩm (5µg) được khuếch đại bằng kit khuếch đại

ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare sản xuất) được giải trình tự lần lượt bằng máy giải trình tự GS FLX Titan 454 và máy giải trình tự HiSeq 2000. Trong quá trình giải trình tự bằng máy giải trình tự HiSeq 2000, thư viện ADN và các chất phản ứng được đổ vào công đo dòng 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 công đo dòng chảy. Các đầu bắt cặp có chiều dài bằng 101bp được giải trình tự bằng máy giải trình tự HiSeq 2000 để thu được dữ liệu trình tự hệ gen môi trường.

Quá trình giải trình tự ADN hệ gen môi trường này được hiện bằng cách sử dụng mẫu đất suối nước nóng AR19. Kết quả là thu được trình tự bộ gen có chiều dài đoạn đọc 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 bằng máy giải trình tự GS FLX Titan 454 và chiều dài đoạn trình tự đọc trung bình bằng 92,65bp ở các đoạn trình tự đọc có đầu bắt cặp, 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 bằng máy giải trình tự HiSeq 2000, 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.

Trình tự nucleotit thu được bằng máy giải trình tự GS FLX Titan 454 và máy giải trình tự 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 sản xuất. Sau khi sàng lọc chất lượng, tổng chiều dài đoạn trình tự đọc bằng máy giải trình tự 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 giải trình tự HiSeq 2000 bằng 81.323.692.563bp. Sau khi lắp ráp, tổng số lượng các đoạn tiếp giáp có chiều dài lớn hơn 500bp bằng 967.925, và tổng chiều dài bằng 419.787.603bp. Trong số các đoạn tiếp giáp này, đoạn tiếp giáp lớn nhất có chiều dài bằng 287.641bp.

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

Các trình tự có mã phân loại enzym 3.2.1.4 (xenlulaza), 3.2.1.21 ( $\beta$ -glucosidaza), 3.2.1.37 ( $\beta$ -xylosidaza), 3.2.1.91 (xenluloza 1,4- $\beta$ -xenlobiosidaza), và 3.2.1.8 (endo 1,4- $\beta$ -xylanaza) được tải về từ cơ sở dữ liệu UniProt (<http://www.uniprot.org/>, ngày truy cập:

09/12/2011), và cơ sở dữ liệu về các gen mã hóa glycosit hydrolaza này được thiết kế. Khung đọc mã mở (các vùng gen) được phát hiện từ các trình tự tiếp giáp thu được ở phần 2 bằng phần mềm Metagene (Noguchi *et al.*, *DNA Research*, 2008, 15(6)) (tùy chọn Metagene: -m). Gen mã hóa glycosit hydrolaza từ trình tự ORF phát hiện được được xác định bằng phần mềm BLASTP (phiên bản 2.2.18) và tham khảo cơ sở dữ liệu nêu trên. Điều kiện tùy chọn của phần mềm BLASTP được cài đặt như sau: “trình tự truy vấn = sai”, “trị số kỳ vọng (E) <  $1e^{-20}$ ”, “trị số mặc định - trị số để phát hiện khoảng trống = -1, trị số để mở rộng khoảng trống = -1, trị số sắp thẳng hàng được tạo khoảng trống X = 0, trị số lý thuyết cho các trình tự mục tiêu mở rộng = 0, cỡ chữ = mặc định”, và các trình tự khung đọc mã mở mục tiêu được thu thập là gen mã hóa glycosit hydrolaza. Các trình tự nucleotit thu được bao gồm các gen mã hóa glycosit hydrolaza như 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 pfam HMMs (Pfam phiên bản 23.0 và HMMER phiên bản 2.3 Finn *et al.*, *Nucleic Acids Research Database*, 2010, Issue 38, p.D211-222) dưới dạng cơ sở dữ liệu về trình tự vùng chức năng của protein. Cụ thể, mỗi trình tự nucleotit này được phân loại thành họ glycosit hydrolaza (GH) dựa trên độ tương đồng của chúng với cơ sở dữ liệu vùng Pfam bằng thuật toán tra cứu độ tương đồng trình tự HMMER (Durbin *et al.*, “*The theory behind profile HMMs. Biological sequence analysis: probabilistic models of protein and nucleic acids*”, 1998, Cambridge University Press.; hmmpfam (phiên bản 2.3.2), trị số kỳ vọng  $E < 1e^{-5}$ ; cơ sở dữ liệu = Pfam\_fs (mô hình có thể được sử dụng để phát hiện các đoạn của các vùng đại diện ở trình tự)). Lưu ý rằng biến thể chứa 70% trình tự của vùng xúc tác GH hoặc cao hơn được tính là các gen mã hóa enzym thuộc các họ tương ứng.

Từ dữ liệu trình tự hệ gen môi trường AR19 có chiều dài 84,2Gbp, 602.589 khung đọc mã mở được xác định bằng phần mềm Metagene, và tổng số lượng khung đọc mã mở có chiều dài đầy đủ bằng 251.146.

406 khung đọc mã mở mục tiêu được xác định là trình tự  $\beta$ -glucosidaza hoặc  $\beta$ -xylosidaza bằng phần mềm tra cứu độ tương đồng BLASTP. Trong số 406 khung đọc mã

mở này, 168 khung đọc mã mở được xác định là gen mã hóa  $\beta$ -glucosidaza hoặc  $\beta$ -xylosidaza bằng phần mềm pfam HMMs, trong khi đó 238 khung đọc mã mở chứa trình tự vùng xúc tác GH nhỏ hơn 70%, hoặc không tương đồng với trình tự bất kỳ trên cơ sở dữ liệu pfam.

Kết quả phân loại họ GH của 168 khung đọc mã mở được xác định là gen mã hóa  $\beta$ -glucosidaza hoặc  $\beta$ -xylosidaza được thể hiện trong Bảng 1. Các trình tự chứa vùng xúc tác GH nhỏ hơn 70% 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 mã mở được xác định là gen mã hóa  $\beta$ -glucosidaza hoặc  $\beta$ -xylosidaza, 19 khung đọc mã mở có chiều dài đầy đủ thuộc họ GH1, 57 khung đọc mã mở có chiều dài đầy đủ thuộc họ GH3, 13 khung đọc mã mở có chiều dài đầy đủ thuộc họ GH31, và 13 khung đọc mã mở có chiều dài đầy đủ thuộc họ GH43 thu được từ hệ gen môi trường AR19. Các đoạn môi được thiết kế cho toàn bộ các khung đọc mã mở này đã được xác định là gen mã hóa  $\beta$ -glucosidaza hoặc  $\beta$ -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  $\beta$ -xylosidaza được phân lập từ khung đọc mã mở AR19M-346 thuộc họ GH3 và có trình tự của gen  $\beta$ -xylosidaza.

Bảng 1

| Hệ gen môi trường<br>AR19                     | Phân loại họ GH của các gen $\beta$ -glucosidaza hoặc $\beta$ -xylosidaza |     |      |      |         |         |
|-----------------------------------------------|---------------------------------------------------------------------------|-----|------|------|---------|---------|
|                                               | GH1                                                                       | GH3 | GH31 | GH43 | GH khác | Tổng số |
| Các khung đọc mã mở có chiều dài đầy đủ       | 19                                                                        | 57  | 13   | 13   | 3       | 105     |
| Các khung đọc mã mở có chiều dài không đầy đủ | 3                                                                         | 52  | 4    | 4    | 0       | 63      |
| Tổng số khung đọc mã mở                       | 22                                                                        | 109 | 17   | 17   | 3       | 168     |

#### 5. Khung đọc mã mở AR19M-346

Khung đọc mã mở AR19M-346 mã hóa polypeptit (SEQ ID NO.1) chứa 773 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 mã mở AR19M-346 là protein đa vùng chứa đầu tận cùng N của vùng xúc tác của glycosit hydrolaza GH3 chứa 325 gốc axit amin từ threonin (T) ở vị trí số 23 đến

valin (V) ở vị trí số 347, đầu tận cùng C của vùng xúc tác của glycosit hydrolaza GH3 chứa 260 gốc axit amin từ isoleucin (I) ở vị trí số 379 đến threonin (T) ở vị trí số 638; và vùng tương tự Fibronectin loại III chứa 70 gốc axit amin từ axit glutamic (E) ở vị trí số 673 đến serin (S) ở vị trí số 742. Kết quả phân tích bằng phần mềm dự đoán 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 22 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 mà đã được khuếch đại bằng kit khuếch đại ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare sản xuất), và đoạn môi xuôi có trình tự nucleotit như nêu trong SEQ ID NO.6 (5'-CACCATGGAGATTTATAAGGATCCA-3' được thiết kế bằng cách bổ sung 4 nucleotit (CACC) vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.4. 4 nucleotit CACC được bổ sung ở đầu 5' là trình tự để chèn vào vector) và đoạn môi ngược có trình tự nucleotit như nêu trong SEQ ID NO.5 (5'-TTATATAATTTTAACCTCGGTAAAGAATATCCTTT-3'). Trình tự nucleotit như nêu trong SEQ ID NO.4 là trình tự nucleotit tương đồng với đoạn trình tự chứa các nucleotit ở vị trí số 1 đến 21 của trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3. Hơn nữa, trình tự nucleotit như nêu trong SEQ ID NO.5 là trình tự nucleotit bổ sung với đoạn trình tự chứa các nucleotit ở vị trí số 2.288 đến 2.322 của trình tự nucleotit như nêu trong SEQ ID NO.2 hoặc SEQ ID NO.3. Các sản phẩm PCR đã được khuếch đại được chèn vào vector pET101/D-TOPO bằng kit biểu hiện Champion pET Directional TOPO (do Life Technologies sản xuất), và biến nạp vào chủng One Shot TOP10. Các dòng đã biểu hiện được chọn lọc bằng phương pháp PCR trực tiếp từ khuẩn lạc, sau đó nuôi cấy trong môi trường lỏng LB chứa ampicilin với nồng độ bằng 100mg/L ở nhiệt độ 37°C và tốc độ 200 vòng/phút trong 17 đến 20 giờ, tiếp theo các plasmit được tổng hợp bằng kit miniprep (hệ thống tinh chế ADN Wizard plus SV Minipreps, do Promega sản xuất). Các plasmit đã tổng hợp được giải trình tự bằng máy giải trình tự ADN 3730 do Life Technologies sản xuất.

Kết quả thu được hai sản phẩm tách dòng gen AR19M-346-18 và AR19M-346-43 từ khung đọc mã mở AR19M-346 bằng quá trình tách dòng PCR. Trình tự nucleotit của

sản phẩm tách dòng gen AR19M-346-18 (hoặc SEQ ID NO.3) là gen  $\beta$ -xylosidaza tiềm năng có chiều dài bằng 2.322bp tương đồng với trình tự nucleotit của khung đọc mã mở AR19M-346 (SEQ ID NO.2), và chỉ khác trình tự SEQ ID NO.2 một nucleotit. Tức là, nucleotit ở vị trí số 1.653 của sản phẩm tách dòng gen AR19M-346-18 (SEQ ID NO.3) là xytosin (C), trong khi đó nucleotit ở vị trí số 1.653 của khung đọc mã mở AR19M-346 (SEQ ID NO.2) là thymine (T). Tuy nhiên, không có sự khác biệt giữa các trình tự axit amin (SEQ ID NO.1) mã hóa bởi các khung đọc mã mở này.

Polypeptit (AR19M-346-18) mã hóa bởi gen  $\beta$ -xylosidaza tiềm năng AR19M-346-18 (sau đây được gọi tắt là “gen AR19M-346-18”) chứa 773 gốc axit amin, và có trình tự axit amin tương đồng với trình tự axit amin của polypeptit mã hóa bởi khung đọc mã mở AR19M-346 chứa 773 gốc axit amin (SEQ ID NO.1). AR19M-346-18 là protein đa vùng chứa đầu tận cùng N của vùng xúc tác của glycosit hydrolaza GH3 chứa 325 gốc axit amin từ threonine ở vị trí số 23 đến valine ở vị trí số 347 (T23-V347), đầu tận cùng C của vùng xúc tác của glycosit hydrolaza GH3 chứa 260 gốc axit amin từ isoleucine ở vị trí số 379 đến threonine ở vị trí số 638 (I379-T638); và vùng tương tự Fibronectin loại III chứa 70 gốc axit amin từ axit glutamic ở vị trí số 673 đến serine ở vị trí số 742 (E673-S742).

Fig.1 là hình vẽ thể hiện sự sắp thẳng hàng các trình tự axit amin của vùng xúc tác  $\beta$ -xylosidaza của sản phẩm tách dòng gen AR19M-346-18 và trình tự axit amin của  $\beta$ -xylosidaza GH3 (mã truy cập Ngân hàng gen: AFY97406.1) có nguồn gốc từ vi khuẩn *Fervidobacterium gondwanense* thuộc ngành *Thermotogae*. Trên Fig.1, các axit amin nằm sau có màu đen/trắng thể hiện các gốc axit amin giống hệt nhau trong toàn bộ các trình tự axit amin này, và các gốc axit amin được đánh bóng thể hiện 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). Vùng xúc tác  $\beta$ -xylosidaza của sản phẩm tách dòng gen AR19M-346-18 chứa khoảng trống có độ tương đồng bằng 76% so với  $\beta$ -xylosidaza GH3 của *Fervidobacterium gondwanense*.

## 7. Quá trình biểu hiện và tinh chế protein $\beta$ -xylosidaza.

Sau khi giải trình tự, các plasmit chứa gen đích được biến nạp vào *E. coli* để biểu hiện protein bằng phương pháp sốc nhiệt. Chủng BL21 Star (DE3) có sẵn trong kit biểu hiện Champion pET Directional TOPO (do Life Technologies sản xuất) được sử dụng

làm tế bào chủ để chuyển gen. *E. coli* chứa gen đích được cấy vào môi trường LB chứa ampicilin với hàm lượng bằng 100mg/L và nuôi cấy cho đến khi mật độ quang ở bước sóng 600nm nằm trong khoảng từ 0,2 đến 0,8, sau đó IPTG (isopropyl- $\beta$ -D(-)-thiogalactopyranosit) được bổ sung vào, và nuôi cấy tiếp trong 5 đến 20 giờ. Kết quả là quá trình biểu hiện protein đích được cảm ứng. Sau khi nuôi cấy, *E. coli* được thu nhận bằng cách ly tâm, và dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) có thể tích bằng 1/10 thể tích dịch nuôi cấy được bổ sung vào để tạo hỗn dịch vi khuẩn. Sau đó, quy trình phân cắt trong 5 phút và rung lắc trong 5 phút được lặp lại từ 7 đến 8 lần bằng thiết bị Astrason 3000 (do Misonix, Inc. sản xuất). Kết quả là thu được dịch chiết thô chứa *E. coli* tái tổ hợp gen chứa protein đích. Dịch chiết thô chứa *E. coli* tái tổ hợp gen được lọc qua màng lọc (đường kính lỗ lọc  $\phi = 0,45\mu\text{m}$ , do Millipore sản xuất), và dịch lọc thu được được sử dụng làm dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen.

Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen được nạp vào cột trao đổi ion HiTrap Q HP (do GE Healthcare 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% với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 1M NaCl bằng hệ thống sắc ký lỏng áp suất từ trung bình đến cao AKTA (do GE Healthcare sản xuất). Các phân đoạn có hoạt tính  $\beta$ -xylosidaza được trộn toàn bộ, sau đó trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 750 mM amoni sulfat bằng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất). Các phân đoạn có hoạt tính  $\beta$ -xylosidaza sau khi trao đổi dung dịch được nạp vào cột phân tách tương tác kỵ nước HiTrap Phenyl HP (do GE Healthcare sản xuất) được làm cân bằng với cùng dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 100% với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0). Các phân đoạn có hoạt tính  $\beta$ -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 sản xuất) được làm cân bằng với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 150mM NaCl, và được phân đoạn bằng cùng dung dịch đệm có thể tích gấp 1 đến 1,5 lần thể tích của cột ở tốc độ dòng nằm trong khoảng từ 2mL/phút đến 3mL/phút. Các phân đoạn có hoạt tính  $\beta$ -xylosidaza được trộn toàn bộ, sau đó trao đổi với dung dịch đệm axit phosphoric 1mM (độ pH = 6,8) và cô bằng màng siêu lọc ly tâm VIVASPIN 20, sau đó nạp vào cột

hydroapatit CHT5-1 (do Bio-Rad Laboratories, Inc. sản xuất) được làm cân bằng với cùng dung dịch đệm, các protein được phân đoạn với gradient nồng độ nằm trong khoảng từ 0% đến 100% với dung dịch đệm axit phosphoric 400mM (độ pH = 6,8). Các phân đoạn có hoạt tính  $\beta$ -xylosidaza được trộn toàn bộ, sau đó trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) và cô. Kết quả là thu được enzym tinh khiết có nồng độ cuối cùng bằng khoảng 3,5mg/mL.

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

Fig.2 thể hiện kết quả phân tích điện di trên gel SDS-PAGE của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen được điều chế từ *E. coli* đã chuyển gen AR19M-346-18 và enzym tinh khiết được tinh chế từ dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen này. Làn 1 là gen chỉ thị phân tử protein, làn 2 và làn 3 lần lượt thể hiện các mẫu điện di của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen, hoặc enzym tinh khiết. Kết quả phân tích cho thấy trong dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (làn 2) có một băng đậm có trọng lượng phân tử gần bằng trọng lượng phân tử của polypeptit được xác định từ trình tự axit amin (SEQ ID NO.1), và trong enzym tinh khiết (làn 3) có một băng duy nhất tương ứng với băng nêu trên (được thể hiện bằng mũi tên trên Fig.2).

#### 8. Đo hoạt tính $\beta$ -xylosidaza bằng cơ chất PNPX (hoạt tính thủy phân PNPX)

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

Hoạt tính thủy phân PNPX của protein enzym (AR19M-346-18) mã hóa bởi gen AR19M-346-18 được đo. Cụ thể hơn, hỗn hợp chứa 100 $\mu$ L dung dịch nước PNPX 3,4mM, 50 $\mu$ L dung dịch đệm axit axetic 200mM (độ pH bằng 6,0), 20 $\mu$ L dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen thu được ở phần 7 hoặc dung dịch enzym tinh khiết (10ng/ $\mu$ L) được điều chế bằng cách pha loãng enzym tinh khiết (3,5mg/mL) với nước, và 30 $\mu$ L nước tinh khiết được để phản ứng ở nhiệt độ nằm trong khoảng từ 30°C đến 100°C trong 20 phút. Trong toàn bộ các phép đo, 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 phản ứng ở cùng điều kiện được sử dụng làm mẫu đối chứng. Hơn nữa, dung dịch cơ chất và enzym (dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen hoặc enzym tinh khiết) được giữ lần lượt và riêng biệt ở nhiệt độ phản ứng được duy trì trong 5 phút, sau đó trộn. Khoảng thời gian này được thiết lập để kích hoạt phản ứng. Trong quá trình phản ứng, các hỗn hợp này được khuấy bằng máy khuấy nhiệt (do Eppendorf AG sản xuất). Sau khi kết thúc phản ứng, phản ứng được dừng bằng cách bổ sung dung dịch natri carbonat 0,2M vào mỗi hỗn hợp với cùng thể tích kết hợp khuấy trộn, sau đó ly tâm. Kết quả là thu được phần dịch nổi. Độ hấp thụ ở bước sóng 420nm được đo bằng máy quang phổ, và hàm lượng p-nitrophenol trong phần dịch nổi được đo bằng đường chuẩn p-nitrophenol. Hàm lượng p-nitrophenol tạo ra bởi quá trình thủy phân được đo bằng cách so sánh độ sai lệch với mẫu đối chứng. Hoạt tính enzym để tạo ra 1 $\mu$ 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). Ngoài ra, 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 dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen lẫn enzym tinh khiết đều có hoạt tính  $\beta$ -xylosidaza (hoạt tính thủy phân PNPX).

#### 9. Độ đặc hiệu cơ chất của AR19M-346-18

Hoạt tính thủy phân các cơ chất xenluloza và hemixenluloza khác nhau của protein enzym (AR19M-346-18) mã hóa bởi gen AR19M-346-18 được đo. Trong phép đo này, dung dịch enzym tinh khiết được điều chế bằng cách pha loãng enzym tinh khiết (nồng độ bằng 3,5mg/mL) thu được ở phần 7 thành 10ng/ $\mu$ L với nước được sử dụng. Ngoài ra, PNPX (do Sigma-Aldrich Co. LLC. sản xuất), PNPG (do Sigma-Aldrich Co. LLC. sản xuất), 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, LLC. sản xuất), và laminarin (có nguồn gốc từ *Laminaria digitata*, do Sigma-Aldrich Co. LLC. sản xuất) được sử dụng làm các cơ chất.

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

Cụ thể hơn, trước tiên phản ứng enzym được thực hiện bằng cách ủ sơ bộ hỗn hợp phản ứng chứa 50 $\mu$ L dung dịch đệm axit axetic 200mM (độ pH bằng 6,0), 20 $\mu$ L dung dịch enzym tinh khiết (10ng/ $\mu$ L) và 30 $\mu$ L nước tinh khiết ở nhiệt độ 70°C trong 5 phút, tiếp theo bổ sung thêm 100 $\mu$ L mỗi dung dịch cơ chất (dung dịch nước chứa 1% khối lượng PSA, bột Avicel, CMC, xylan, lichenan và laminarin, và dung dịch nước chứa 3,4mM PNPX và PNPG) vào và ủ hỗn hợp thu được ở nhiệt độ 70°C trong 20 phút (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 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 được đo, và hoạt tính riêng (U/mg) được đo theo cùng cách thức như nêu trong phần 8. 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-đinitrosalixylic (dung dịch DNS) được bổ sung vào. 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. Kết quả là thu được phần dịch nổi. Độ hấp thụ ở bước sóng 540nm được đo bằng máy quang phổ, và hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn glucoza (khi xylan được sử dụng làm cơ chất, thì hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn xyloza). Lượng đường khử tạo ra bởi quá trình thủy phân được đo bằng cách so sánh độ sai lệch với từ mẫu đối chứng. Hoạt tính enzym để tạo ra 1 $\mu$ mol đường khử trong 1 phút được tính bằng

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

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.3. Các kết quả này cho thấy AR19M-346-18 có hoạt tính thủy phân mạnh PNPX và cũng có hoạt tính phân giải PNPG. Trái lại, AR19M-346-18 hầu như không có hoạt tính phân giải các cơ chất khác.

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

Mức độ phụ thuộc vào nhiệt độ và độ pH của hoạt tính thủy phân PNPX của protein enzyme (AR19M-346-18) mã hóa bởi gen AR19M-346-18 được đánh giá. Trong phép đo này, dung dịch enzyme tinh khiết (10ng/ $\mu$ L) được điều chế bằng cách pha loãng enzyme tinh khiết (nồng độ bằng 3,5mg/mL) thu được ở phần 7 với nước được sử dụng.

Mức độ phụ thuộc vào độ pH của hoạt tính thủy phân PNPX của enzyme tinh khiết AR19M-346-18 được đo theo cùng cách thức như nêu trong phần 8, chỉ khác là cho hỗn hợp chứa 100 $\mu$ L dung dịch nước PNPX 3,4mM, 50 $\mu$ L dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3 đến 8), 50 $\mu$ L dung dịch đệm axit axetic (độ pH nằm trong khoảng từ 3,5 đến 6), hoặc 50 $\mu$ L dung dịch đệm axit phosphoric (độ pH nằm trong khoảng từ 6 đến 8), 30 $\mu$ L nước tinh khiết và 20 $\mu$ L dung dịch enzyme tinh khiết (10ng/ $\mu$ L) phản ứng ở nhiệt độ 70°C trong 20 phút, trong đó hàm lượng p-nitrophenol được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PNPX (U/mg) được đo.

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PNPX của enzyme tinh khiết AR19M-346-18 được đo theo cùng cách thức như nêu trong phần 8, chỉ khác là phản ứng được thực hiện ở nhiệt độ bằng 30°C, 40°C, 50°C, 60°C, 65°C, 70°C, 75°C, 80°C, 85°C, 90°C, 95°C, hoặc 100°C, trong đó hàm lượng p-nitrophenol được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PNPX (U/mg) được đo.

Các kết quả đo được thể hiện trên Fig.4 và Fig.5. Fig.4 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở độ pH bằng 6,0) của enzyme tinh khiết AR19M-346-18 ở các nhiệt độ tương ứng, trong đó nhiệt độ được thể hiện trên trục hoành. Fig.5 là đồ thị thể hiện kết quả đo hoạt tính thủy phân PNPX (ở nhiệt độ 70°C) của enzyme tinh khiết

AR19M-346-18 ở các độ pH tương ứng, trong đó độ pH được thể hiện trên trục hoành. Trên Fig.5 ký hiệu “MB” thể hiện kết quả đo thu được khi dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3 đến 8) được sử dụng, ký hiệu “SAB” thể hiện kết quả đo thu được khi dung dịch đệm axit axetic (độ pH nằm trong khoảng từ 3,5 đến 6) được sử dụng, và ký hiệu “PB” thể hiện kết quả đo thu được khi dung dịch đệm axit phosphoric (độ pH nằm trong khoảng từ 6 đến 8) được sử dụng. Độ pH được vẽ đồ thị bằng các trị số đo thực của hỗn hợp dung dịch chứa cơ chất, dung dịch đệm, và enzym.

Enzym tinh khiết AR19M-346-18 có hoạt tính thủy phân mạnh PNPX ở nhiệt độ nằm trong khoảng từ 80°C đến 90°C (Fig.4). Nhiệt độ tối ưu ( $T_{\text{tối ưu}}$ ) ở đó enzym này có hoạt tính mạnh nhất bằng 85°C ở độ pH bằng 6,0. Hoạt tính thủy phân PNPX của enzym tinh khiết AR19M-346-18 bị giảm nhanh khi nhiệt độ phản ứng enzym bằng 95°C hoặc cao hơn.

Ngoài ra, enzym tinh khiết AR19M-346-18 có hoạt tính thủy phân PNPX ở độ pH nằm trong khoảng từ 4,0 đến 8,0 ở nhiệt độ 70°C. Cụ thể, enzym tinh khiết AR19M-346-18 có hoạt tính thủy phân PNPX mạnh nhất ở nhiệt độ 70°C và độ pH = 6,0.

11. Đánh giá độ ổn định nhiệt của  $\beta$ -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 mật độ huỳnh quang tương ứng thì có thể xác định được nhiệt độ biến tính của protein (tức là nhiệt độ thay đổi mật độ huỳnh quang).

Trong phép đo này, dung dịch enzym tinh khiết (1mg/mL) được điều chế bằng cách pha loãng enzym tinh khiết (nồng độ bằng 3,5mg/mL) thu được ở phần 7 với nước được sử dụng

Cụ thể hơn, 2 $\mu$ L thuốc nhuộm màu cam SYPRO đã pha loãng 100 lần (do Life Technologies sản xuất Inc.), 1 $\mu$ L enzym tinh khiết có nồng độ bằng 1 mg/mL, 5 $\mu$ L dung dịch đệm axit phosphoric 200mM (pH 6,0) và 12 $\mu$ 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  $\mu$ 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 nhiệt độ đạt đến nhiệt độ mục tiêu, mật 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, mật độ huỳnh quang được định lượng bằng máy ảnh CCD, và sự thay đổi về mật 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. Fig.6 thể hiện sự thay đổi về mật 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-346-18. Phần đồ thị bên trên của Fig.6 thể hiện dữ liệu thực đo được, và phần đồ thị bên dưới của Fig.6 thể hiện đạo hàm bậc nhất “ $-d(\text{mật độ huỳnh quang})/dT$ ” của đồ thị thay đổi mật độ huỳnh quang của phần đồ thị bên trên của Fig.6. 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 “ $-d(\text{mật độ huỳnh quang})/dT$ ” được thể hiện trên trục Y của phần đồ thị bên dưới trên Fig.6 của đồ thị mật độ huỳnh quang dưới dạng hàm số của nhiệt độ. Nhiệt độ biến tính ( $T_m$ ) của protein enzym AR19M-346-18 được xác định từ đạo hàm bậc nhất của đồ thị mật độ huỳnh quang bằng  $91,9^\circ\text{C} \pm 0,1^\circ\text{C}$  ( $n = 3$ ) khi độ pH bằng 6,0, gần bằng nhiệt độ tối ưu  $T_{\text{tối ưu}} = 85^\circ\text{C}$  của enzym này định lượng được ở phần đo hoạt tính thủy phân PNPX.

**YÊU CẦU BẢO HỘ**

1. Enzym beta-xylosidaza ( $\beta$ -xylosidaza) chịu nhiệt chứa polypeptit phân lập 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.

Fig.1

AR19M-346-18 1 MEIYKOPSAPVILRVEDLLSRNITLEEKV FOLGSIWYSYLLNEDGSFDECKAYELKYGIGOTIRPGGATNFEPERAAEFNKIOTREFL 87  
F. gondwanense 1 MEIYKOSKPPHELRVEDLLSRNITLEEKVSOLGSIWYSYOLLDENGWFDECKAFELKDGIGOTISRPGGATNFOPEEVAQGFNDKNVQKFL 87

AR19M-346-18 88 IENTRLGIPAMHHEECLTGYNGLGGTSFPVPPIAMASTWEPEL INKAASVIRDEIRTVGAHQGLAPVLDVVRDPRWGRVEETFGESPY 174  
F. gondwanense 88 IENTRLGIPAMHHEECLTGYNGLNGSNMFPVPPIAMASTWEPELVKEVTKVIRQEURNNGIHQGLAPVLDVARDPRWGRVEETFGESPY 174

AR19M-346-18 175 LVATNGCAYINGLOGNDLRUGVIATAKHFGVGYGASEGGRNWAPTNIIPRELREVF LPFEAAVKISKIGSVMNSYSEIDGVPVAASE 261  
F. gondwanense 175 LVASNGCAYVEGLOGEDLKDGVIATTKHFVGYSASEGGRNWAPTNIIPRELREIFLPFEAAVKVAKVGSVMNSYSEIDGVPVAAAR 261

AR19M-346-18 262 ELTRDVLKENGFDGIWSDYESIALMEYHKIAETKADAALADASIDVELPKIDCYKHVELVRDGKISERLIDESVRRILKVK 348  
F. gondwanense 262 ELTRDVLKENGFDGIWSDYFSUKLMEHHKARDKAEAAKYALEAGIDVELPNTDCYAHVELDLVKSGLPEKLDQIVRRILKVK 348

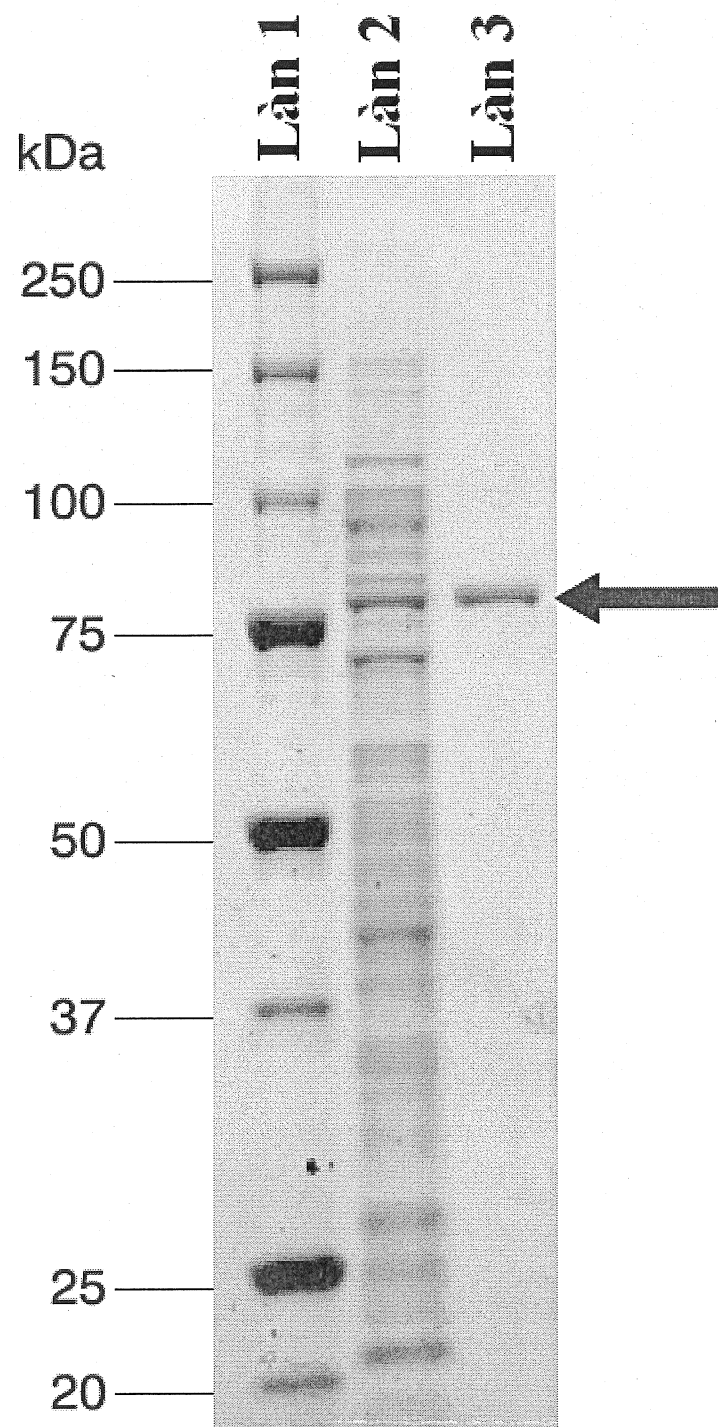
AR19M-346-18 349 FLGLFEDNPYVRPS--KIVKNSGLALEIARKSIVLLKNDGILPLKNEKVALIGPNAESVRNMLGDYMYLSHISWMLNNENFNAP 433  
F. gondwanense 349 FKLGLFDKPYVEPSKAKVKNLEALEVARKSIVLLKNDGILPLKDKKVALIGPNAADVRNMLGDYMYLAHIKIMLENVNLAFDAP 435

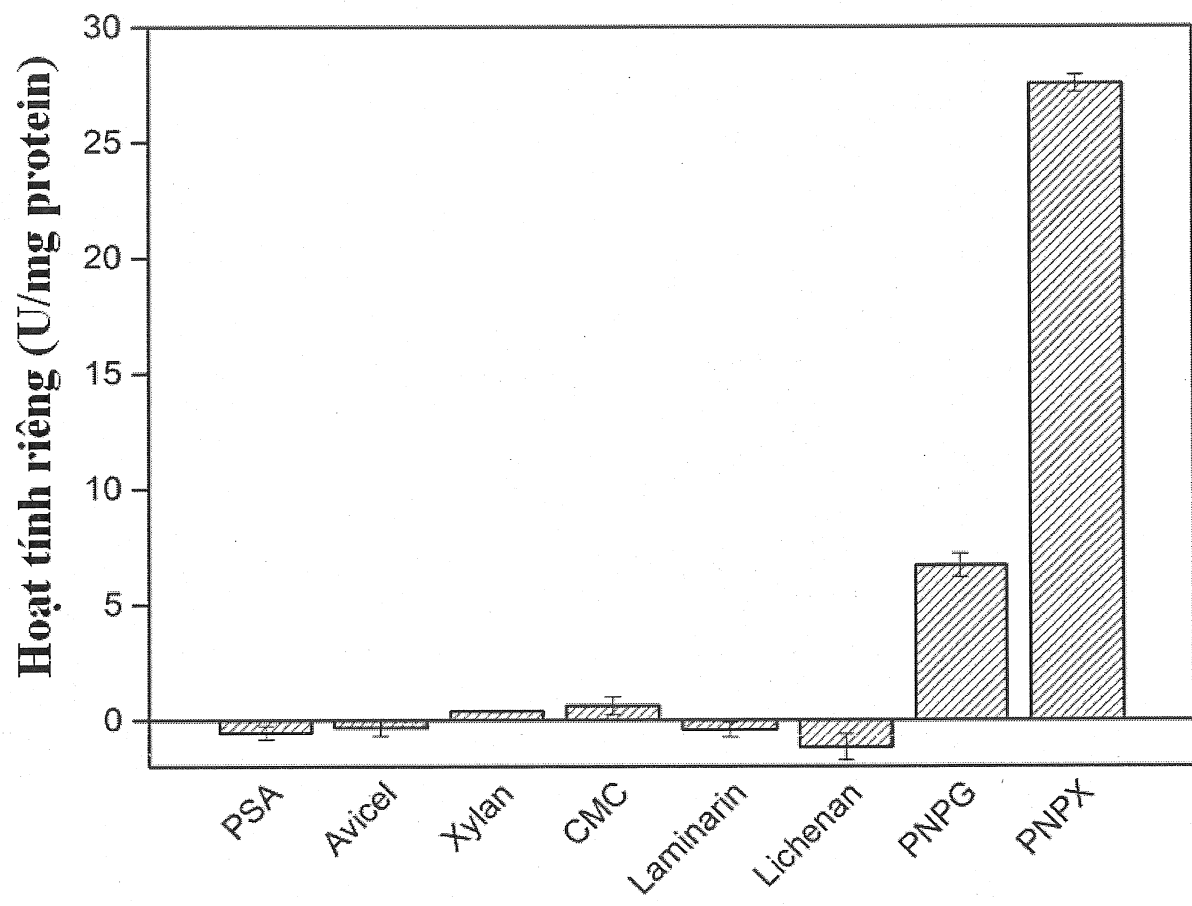
AR19M-346-18 434 KENLSGVKESVEKNMSITKKTIFDEEGIKVIYAKGGDVLGTSTEGFSKALEAVEKCDVAIVVVGDRSGLTKDCTTGESRDTSTLK 520  
F. gondwanense 436 KENLSVVKKSVEESMNKIKSIEMLKEESVQETIYAKGGDVLGDSKEGNEALKAVENS DVAIVVVGDRSGLTNDCTTGESRDSANLK 522

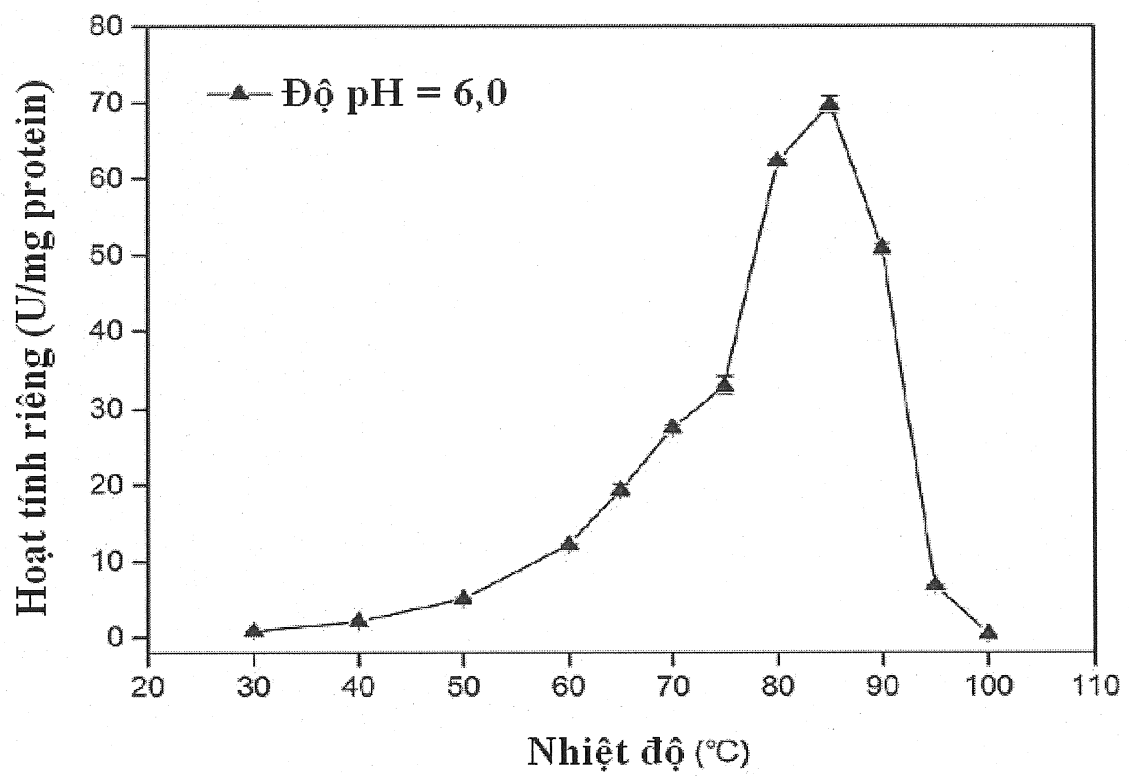
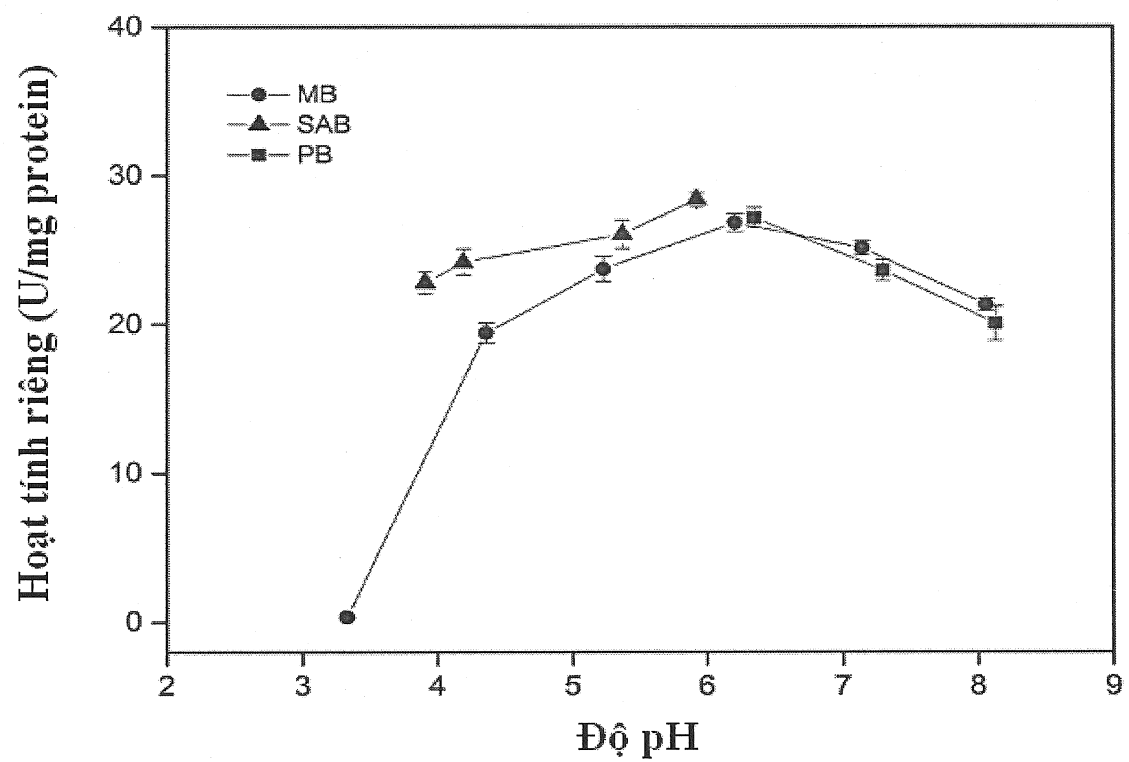
AR19M-346-18 521 LPGVQEEELTDAVSNVGKPVVVLISGRPYSLAKVVDKVSALVQAWLPGEAGAEAFDVLYGKYNPSGKLPUTIPRSVGOIPLFHYFK 607  
F. gondwanense 523 LPGVQEEELTIEVSKVGKPVVVALINGRPYSLTRVVDKVSALVEAWLPGEIGAKAIVDVLYGKYNPSGKLPUTIPRSAGOIPLFHYFK 609

AR19M-346-18 608 PSGGRSSWHGDYVSESVKPLFEFGFLSYLTTEYSDELITPSRVSGIGTVEISLNVNSGEYGEENVQLYIGREYASVTRPVKELK 694  
F. gondwanense 610 PSGGRSSWHGDYVDESVPKLPFEGHGLSYLTNEDYSGLEISPSKVPNAGSVEISLYVENIGVEGEENVQLYIGRECASVTRPVKELK 696

AR19M-346-18 695 GFAKVGKPKGEKRKVFKLHTEQLAFYGIIDUKLIGEPGVYNNIGSSSEDIRLRGTFSDGEEKIKVPNERIFFTEVKII\* 713  
F. gondwanense 697 GFAKVNPKPKGEKRKVFNHLDVLAFYGRDUKLCQEPGVYNNIGSSSDIRLRKGSFEVDGIBREVFEIDRVFEETKVYTF 715

**Fig.2**

**Fig.3**

**Fig.4****Fig.5**

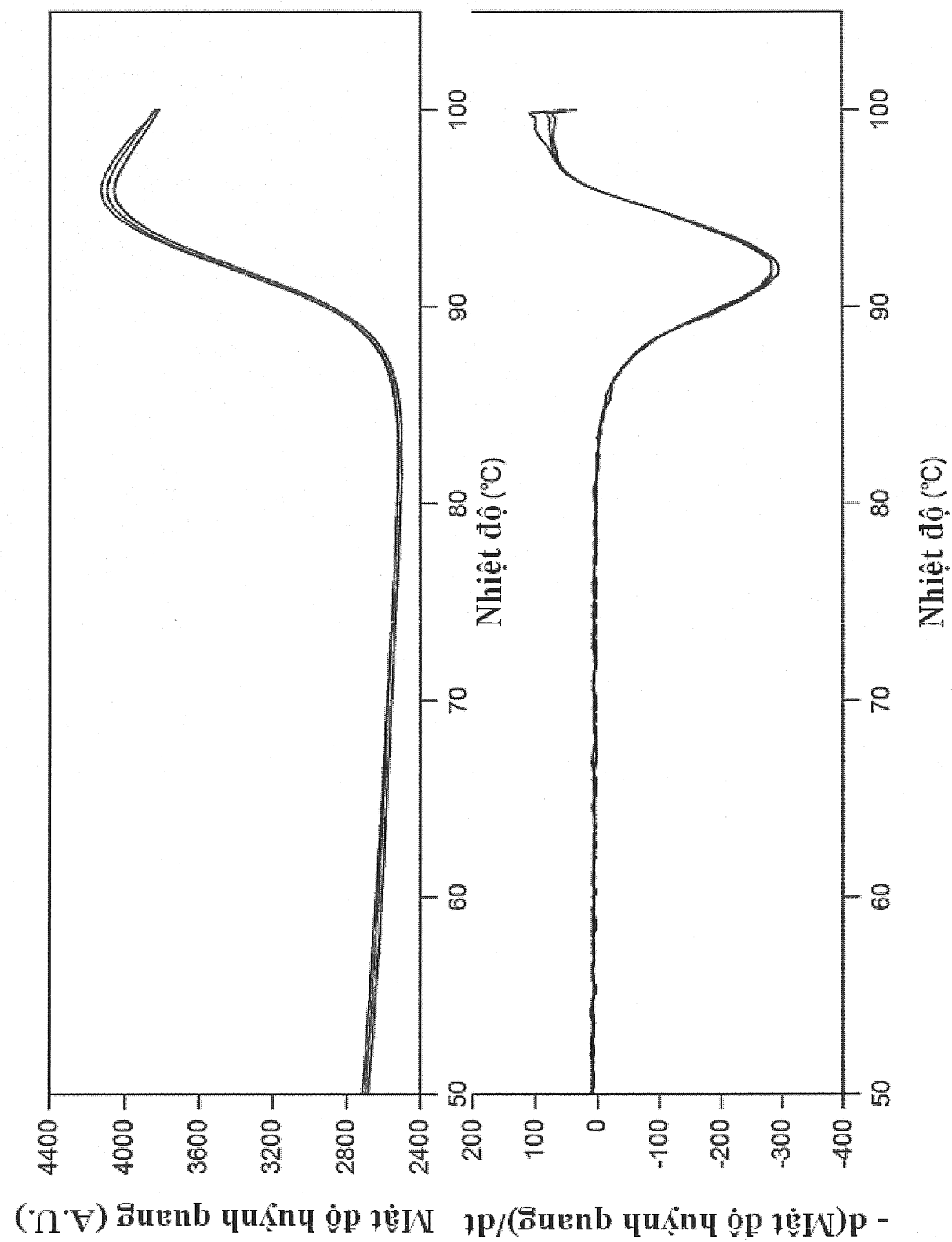


Fig.6

## DANH MỤC TRÌNH TỰ

<110> HONDA MOTOR CO., LTD.  
 <120> ENZYM BETA-XYLOSIDAZA ( $\beta$ -XYLOSIDAZA) CHỊU NHIỆT  
 <130> H114049701  
 <160> 7  
 <170> PatentIn version 3.1  
 <210> 1  
 <211> 773  
 <212> PRT  
 <213> Trình tự chưa biết  
 <220>  
 <223> AR19M-346-18  
 <400> 1  
 Met Glu Ile Tyr Lys Asp Pro Ser Ala Pro Val Ile Leu Arg Val Glu  
 1 5 10 15  
 Asp Leu Leu Ser Arg Met Thr Leu Glu Glu Lys Val Phe Gln Leu Gly  
 20 25 30  
 Ser Ile Trp Ser Tyr Glu Leu Leu Asn Glu Asp Gly Ser Phe Asp Glu  
 35 40 45  
 Gln Lys Ala Tyr Glu Leu Leu Lys Tyr Gly Ile Gly Gln Ile Thr Arg  
 50 55 60  
 Pro Gly Gly Ala Thr Asn Phe Glu Pro Glu Arg Ala Ala Glu Phe Ser  
 65 70 75 80  
 Asn Lys Ile Gln Arg Phe Leu Ile Glu Asn Thr Arg Leu Gly Ile Pro  
 85 90 95  
 Ala Ile Met His Glu Glu Cys Leu Thr Gly Tyr Met Gly Leu Gly Gly  
 100 105 110  
 Thr Ser Phe Pro Val Pro Ile Ala Met Ala Ser Thr Trp Glu Pro Glu  
 115 120 125  
 Leu Ile Asn Lys Ala Ala Ser Val Ile Arg Asp Glu Leu Arg Thr Val  
 130 135 140  
 Gly Ala His Gln Gly Leu Ala Pro Val Leu Asp Val Val Arg Asp Pro  
 145 150 155 160  
 Arg Trp Gly Arg Val Glu Glu Thr Phe Gly Glu Ser Pro Tyr Leu Val  
 165 170 175  
 Ala Thr Met Gly Cys Ala Tyr Ile Asn Gly Leu Gln Gly Asn Asp Leu  
 180 185 190  
 Arg Asn Gly Val Ile Ala Thr Ala Lys His Phe Val Gly Tyr Gly Ala  
 195 200 205  
 Ser Glu Gly Gly Arg Asn Trp Ala Pro Thr Asn Ile Pro Pro Arg Glu  
 210 215 220  
 Leu Arg Glu Val Phe Leu Leu Pro Phe Glu Ala Ala Val Lys Ile Ser  
 225 230 235 240  
 Lys Ile Gly Ser Val Met Asn Ser Tyr Ser Glu Ile Asp Gly Val Pro  
 245 250 255  
 Val Ala Ala Ser Glu Glu Leu Ile Arg Asp Val Leu Arg Lys Glu Trp  
 260 265 270  
 Gly Phe Asp Gly Ile Val Val Ser Asp Tyr Phe Ser Ile Ala Leu Leu  
 275 280 285  
 Tyr Glu Tyr His Lys Ile Ala Glu Thr Lys Ala Gln Ala Ala Lys Leu  
 290 295 300  
 Ala Leu Gln Ala Ser Ile Asp Val Glu Leu Pro Lys Ile Asp Cys Tyr  
 305 310 315 320  
 Lys His Leu Val Glu Leu Val Arg Asp Gly Lys Ile Ser Glu Arg Leu  
 325 330 335  
 Ile Asp Glu Ser Val Arg Arg Ile Leu Lys Val Lys Phe Leu Leu Gly  
 340 345 350  
 Leu Phe Asp Asn Pro Tyr Val Arg Pro Ser Lys Ile Val Lys Asn Ser  
 355 360 365  
 Gly Leu Ala Leu Glu Ile Ala Arg Lys Ser Ile Val Leu Leu Lys Asn  
 370 375 380  
 Asp Gly Ile Leu Pro Leu Lys Asn Glu Met Lys Val Ala Leu Ile Gly  
 385 390 395 400  
 Pro Asn Ala Glu Ser Val Arg Asn Met Leu Gly Asp Tyr Met Tyr Leu  
 405 410 415  
 Ser His Ile Ser Val Met Leu Glu Asn Ile Asn Glu Asn Phe Asn Ala  
 420 425 430  
 Pro Lys Phe Asn Leu Ser Gly Val Lys Glu Ser Val Glu Lys Asn Met  
 435 440 445

```

515          520          525
Ile Asp Ala Val Ser Asn Val Gly Lys Pro Val Ile Val Val Leu Ile
530          535          540
Ser Gly Arg Pro Tyr Ser Leu Ala Lys Val Val Asp Lys Val Ser Ala
545          550          555          560
Ile Val Gln Ala Trp Leu Pro Gly Glu Ala Gly Ala Glu Ala Ile Phe
565          570          575
Asp Val Leu Tyr Gly Lys Tyr Asn Pro Ser Gly Lys Leu Pro Ile Thr
580          585          590
Ile Pro Arg Ser Val Gly Gln Ile Pro Leu Phe His Tyr Phe Lys Pro
595          600          605
Ser Gly Gly Arg Ser Ser Trp His Gly Asp Tyr Val Ser Glu Ser Val
610          615          620
Lys Pro Leu Phe Glu Phe Gly Phe Gly Leu Ser Tyr Thr Thr Phe Glu
625          630          635          640
Tyr Ser Asp Leu Glu Ile Thr Pro Ser Arg Val Ser Gly Ile Gly Thr
645          650          655
Val Glu Ile Ser Leu Asn Val Glu Asn Ser Gly Glu Leu Tyr Gly Glu
660          665          670
Glu Val Val Gln Leu Tyr Ile Gly Arg Glu Tyr Ala Ser Val Thr Arg
675          680          685
Pro Val Lys Glu Leu Lys Gly Phe Ala Lys Val Gly Ile Lys Pro Gly
690          695          700
Glu Lys Arg Lys Val Val Phe Lys Leu His Thr Glu Gln Leu Ala Phe
705          710          715          720
Tyr Gly Ile Asp Met Lys Leu Cys Ile Glu Pro Gly Val Tyr Asn Val
725          730          735
Met Ile Gly Ser Ser Ser Glu Asp Ile Arg Leu Arg Gly Thr Phe Ser
740          745          750
Ile Asp Gly Glu Lys Ile Lys Val Pro Asn Glu Arg Ile Phe Phe Thr
755          760          765
Glu Val Lys Ile Ile
770

```

```

<210> 2
<211> 2322
<212> ADN
<213> Trình tự chưa biết
<220>
<223> Khung đọc mã mở AR19M-346
<400> 2

```

```

atggagattt ataaggatcc aagcgcgcc a gtgattttgc gagttgaaga tttgctttct 60
agaatgactt tggaagaaaa agtttttcag ttgggttcaa tttggagcta tgaactactg 120
aatgaagatg gaagtttcga cgagcaaaaa gcatatgagc ttttgaaata cggcataggt 180
caaattacta ggccaggcgg tgctacaaa ttcgaacctg aacgcgctgc agaatttagc 240
aataaaatcc agaggtttct cattgaaaac acaagacttg gaataacctg tataatgcac 300
gaagaatgct tgactgggta catgggactt ggtggtacaa gttttccggt accaattgct 360
atggcaagta cttgggaacc agaactgatt aacaaggctg cgtctgtcat tcgagatgaa 420
ctcagaacag ttggtgctca ccaaggcctt gcaccggtac ttgatgttgt tcgggatccc 480
agatggggaa ggttagaaga aacttttgga gagtctccgt atcttggtgg aacgatggga 540
tgcgcttata tcaacggcct acaaggaaat gacctaaaga acggagttat tgcaactgct 600
aaacattttg ttggttacgg agcatccgaa ggcggaagga actgggcacc gacgaatata 660
ccacctcgtg aattaagaag ggtgtttctc ttaccgtttg aggcgcgtgt gaagatttca 720
aaaaatcggt cagtgatgaa ttctgtacagt gaaatcgacg gagtcccagt agctgcttct 780
gaggaactca tcgcgatgt gcttagaaa gaatgggggt ttgatgggtat agttgtttcc 840
gattattttt ccatagcact tttgtatgag tatcataaaa ttgctgagac caaagcacia 900
gcagccaagc tcgcgcttca agcctctatc gatgttgaa taccgaagat tgactgttac 960
aagcatctcg tagagttggt gagggatggc aaaatttcgg aaaggctgat agatgaatcg 1020
gtgagaagga tcctaaaggt taagttcctg cttggattat ttgataatcc ttacgttcga 1080
ccgtcgaaga tcgttaagaa ttcttgacta gctctcgaaa tcgctagaaa atctatatgt 1140
ctgcttaaaa acgatggtat tcttccgcta aaaaacgaaa tgaaagtgtg gctgatagga 1200
cctaattcgg aaagtgttag aaatatgtta ggcgattata tgtacctttc tcacatcagt 1260
gtgatgcttg aaaatataaa tgagaatttc aatgctccga aattcaatct ctcgtggtgt 1320
aaagaatcgg tagagaaaaa tatgagtatg attaagagtt tgaaaactat attcgatgaa 1380
gagggcataa aatacatcta cgctaagggt tgcgatgttc tgggaacgtc taccgaagga 1440
ttcagcaaa g cattagaagc ggttgaaaaa tgtgacgtcg cagttgttgt ggtcgggtgac 1500
aggtcaggtt tgacaaaaa ttgcactaca ggggagtc aa gggatacatc tactttaaaa 1560
ctaccgggtg ttcaggaaga gttgattgac gcagtttcaa atgtgggtgaa gcctgtgata 1620
gttggtttga taagtggcag accatattct cttgctaaag ttgtcgataa agtgtcgggt 1680
atagttcagg catggttacc tggagaagcg ggagcagaag cgatatttga tgttctgtat 1740
gggaagtata atccatctgg aaagcttcc taaacaattc cgagaagtgt tggacagatt 1800
ccgctcttcc attattttaa accatctggt ggcagatcaa gctggcatg tgattatgtt 1860
tccgagagtg tcaaacact ttttgagttt ggttttggac tatcgtaac accttttgag 1920
tacagcgatt tagaaatcac accatcgaga gtaagtggta tcgggacagt tgaaatttct 1980
ttaaatgtcg agaacagcg agagctgtac ggtgaggaag ttgttcaact ctacatagga 2040
agagaatacg cttcagtcag acgcctgtg aaaggtttaa agggatttgc aaaagtgtgt 2100
ataaaaccgg gagaaaagag gaaggttgtt ttcaaattgc acacagaaca gttggcgttt 2160
tatggaatag atatgaagct atgtatcgaa ccaggtgtgt acaatgtcat gattggtagt 2220
tcgtctgaag atatacgact ccgtggaacg ttttcaatag atggtgagaa aataaaagtt 2280
ccaaatgaaa ggatattctt taccgaggtt aaaattatat aa 2322
<210> 3

```

```

<211> 2322
<212> ADN
<213> Trình tự chưa biết
<220>
<223> Gen AR19M-346-18
<400> 3
atggagattt ataaggatcc aagcgcgccg gtgatttttc gagttgaaga tttgctttct 60
agaatgactt tgggaagaaaa agtttttcag ttgggttcaa tttggagcta tgaactactg 120
aatgaagatg gaagtttcga cgagcaaaaa gcatatgagc ttttgaaata cggcataggt 180
caaattacta ggccaggcgg tgctacaaat ttccaacctg aacgcgctgc agaatttagc 240
aataaaatcc agaggtttct cattgaaaac acaagacttg gaatacctgc tataatgcac 300
gaagaatgct tgactgggta catgggactt ggtggtacaa gttttccggt accaattgct 360
atggcaagta cttgggaacc agaactgatt aacaaggctg cgtctgtcat tcgagatgaa 420
ctcagaacag ttggtgctca ccaaggcctt gcaccggtac ttgatgttgc tcgggatccc 480
agatggggaa gggtagaaga aacttttgga gagtctccgt atcttgtggc aacgatggga 540
tgcgcttata tcaacggcct acaaggaaat gacctaagaa acggagttat tgcaactgct 600
aaacattttg ttggttacgg agcatccgaa ggcggaagga actgggcacc gacgaatata 660
ccacctcggt aattaagaga ggtgtttctc ttaccgtttg aggcgctgt gaagatttca 720
aaaatcggtt cagtgtgaa ttctgacagt gaaatcgacg gagtcccagt agctgcttct 780
gaggaaactc tccgcgatgt gcttagaagg gaatgggggt ttgatggtat agttgtttcc 840
gattattttt ccatagcact ttgtatgag tatcataaaa ttgctgagac caaagcacia 900
gcagccaagc tcgcgcttca agcctctatc gatgttgaat taccgaagat tgactgttac 960
aagcatctcg tagagttggt gagggatggc aaaatttcgg aaaggctgat agatgaatcg 1020
gtgagaagga tcgtaaaggt taagttcctg cttggattat ttgataatcc ttacgttcga 1080
ccgtcgaaga tcgttaagaa ttctggacta gctctgaaa tcgctagaaa atctatagtt 1140
ctgcttaaaa acgatgggtat tcttccgcta aaaaacgaaa tgaaagttgc gctgatagga 1200
cctaattcgg aaagtgttag aaatatgtta ggcgattata tgtacctttc tcacatcagt 1260
gtgatgcttg aaaaataaaa tgagaatttc aatgctccga aattcaatct ctctggtggt 1320
aaagaatcgg tagagaaaaa tatgagtatg attaagagtt tgaaaactat attcgatgaa 1380
gagggcataa aatacatcta cgctaagggt tgcgatgttc tgggaacgtc taccgaagga 1440
ttcagcaaa g cattagaagc ggttgaaaaa tgtgacgtcg cagttgttgc ggtcggtgac 1500
aggtcaggtt tgacaaaaga ttgcactaca ggggagtcga gggatacatc tactttaaaa 1560
ctaccggtg ttccaggaaga gttgattgac gcagtttcaa atgtgggtaa gcctgtgata 1620
gttggtttga taagttgacg accatattct ctgcgtaaag ttgtcgataa agtgtcggct 1680
atagttcagg catggttacc tggagaagcg ggagcagaag cgatatattga tgttctgtat 1740
gggaagtata atccatctgg aaagcttcct ataacaattc cgagaagtgt tggacagatt 1800
ccgctcttcc attattttta accatctggt ggcagatcaa gctggcatgg tgattatggt 1860
tccgagagtg tcaaacctact ttttgagttt ggttttgac tatcgtacac aacctttgag 1920
tacagcgatt tagaaatcac accatcgaga gtaagtggtt tcgggacagt tgaaatttct 1980
ttaaattgct agaacagcgg agagctgtac ggtgaggaag ttgttcaact ctacatagga 2040
agagaatacg cttcagtcac acgccctgtg aaagagttta agggatttgc aaaagttggt 2100
ataaaacggg gaaaaaagag gaaggttgtt ttcaaattgc acacagaaca gttggcggtt 2160
tatggaatag atatgaagct atgtatcgaa ccaggtgtgt acaatgtcat gattggtagt 2220
tcgtctgaag atatacgact ccgtggaacg ttttcaatag atggtgagaa aataaaagtt 2280
ccaaatgaaa ggatattctt taccgaggtt aaaattatat aa 2322
<210> 4
<211> 21
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.
<400> 4
atggagattt ataaggatcc a 21
<210> 5
<211> 35
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.
<400> 5
ttatataatt ttaacctcgg taaagaatat ccttt 35
<210> 6
<211> 25
<212> ADN
<213> Trình tự nhân tạo
<220>
<223> Mô tả trình tự nhân tạo: đoạn mỗi.
<400> 6
caccatggag atttataagg atcca 25
<210> 7
<211> 775
<212> PRT
<213> Fervidobacterium gondwanense
<220>
<223> GH3 xylosidaza
<400> 7
Met Glu Ile Tyr Lys Asp Ser Ser Lys Pro Ile Glu Leu Arg Val Glu
1 5 10 15
Asp Leu Leu Ser Arg Met Thr Leu Glu Glu Lys Val Ser Gln Leu Gly
20 25 30

```

```

Ser Val Trp Ser Tyr Gln Leu Leu Asp Glu Asn Gly Asn Phe Asp Glu
   35           40           45
Gly Lys Ala Phe Glu Leu Leu Lys Asp Gly Ile Gly Gln Ile Ser Arg
   50           55           60
Pro Gly Gly Ala Thr Asn Phe Gln Pro Glu Glu Val Ala Gln Phe Asp
   65           70           75           80
Asn Lys Val Gln Lys Phe Leu Ile Glu Asn Thr Arg Leu Gly Ile Pro
           85           90           95
Ala Leu Met His Glu Glu Cys Leu Thr Gly Tyr Met Gly Leu Asn Gly
           100          105          110
Ser Asn Phe Pro Val Pro Ile Ala Met Ala Ser Thr Trp Glu Pro Glu
           115          120          125
Leu Val Lys Glu Val Thr Lys Val Ile Arg Gln Glu Met Arg Asn Met
           130          135          140
Gly Ile His Gln Gly Leu Ala Pro Val Leu Asp Val Ala Arg Asp Pro
           145          150          155          160
Arg Trp Gly Arg Val Glu Glu Thr Phe Gly Glu Ser Pro Tyr Leu Val
           165          170          175
Ala Ser Met Gly Cys Ala Tyr Val Glu Gly Leu Gln Gly Glu Asp Leu
           180          185          190
Lys Asp Gly Val Ile Ala Thr Thr Lys His Phe Val Gly Tyr Ser Ala
           195          200          205
Ser Glu Gly Gly Arg Asn Trp Ala Pro Thr Asn Ile Pro Pro Arg Glu
           210          215          220
Leu Arg Glu Ile Phe Met Phe Pro Phe Glu Ala Ala Val Lys Val Ala
           225          230          235          240
Lys Val Gly Ser Val Met Asn Ser Tyr Ser Glu Ile Asp Gly Val Pro
           245          250          255
Leu Ala Ala Ser Arg Glu Leu Leu Thr Asp Val Leu Arg Lys Glu Trp
           260          265          270
Gly Phe Asp Gly Leu Val Val Ser Asp Tyr Phe Ser Val Lys Leu Ile
           275          280          285
His Glu His His Lys Leu Ala Arg Asp Lys Ala Glu Ala Ala Lys Tyr
           290          295          300
Ala Leu Glu Ala Gly Ile Asp Val Glu Leu Pro Asn Thr Asp Cys Tyr
           305          310          315          320
Ala His Val Leu Asp Leu Val Lys Ser Gly Val Ile Pro Glu Lys Leu
           325          330          335
Leu Asp Gln Thr Val Arg Arg Ile Leu Lys Met Lys Phe Lys Leu Gly
           340          345          350
Leu Phe Asp Lys Pro Tyr Val Glu Pro Ser Lys Ala Lys Val Val Lys
           355          360          365
Asn Thr Glu Leu Ala Leu Glu Val Ala Arg Lys Ser Ile Val Leu Leu
           370          375          380
Lys Asn Asp Gly Ile Leu Pro Leu Lys Lys Asp Met Lys Val Ala Leu
           385          390          395          400
Ile Gly Pro Asn Ala Ala Asp Val Arg Asn Met Leu Gly Asp Tyr Met
           405          410          415
Tyr Leu Ala His Ile Lys Ile Met Leu Glu Asn Val Asn Leu Ala Phe
           420          425          430
Asp Ala Pro Lys Phe Asn Leu Ser Ser Val Lys Lys Ser Val Glu Glu
           435          440          445
Ser Met Asn Lys Ile Lys Ser Ile Glu Met Leu Leu Lys Glu Glu Ser
           450          455          460
Val Gln Phe Thr Tyr Ala Lys Gly Cys Asp Val Leu Gly Asp Ser Lys
           465          470          475          480
Glu Gly Phe Asn Glu Ala Leu Lys Ala Val Glu Asn Ser Asp Val Ala
           485          490          495
Ile Val Val Val Gly Asp Arg Ser Gly Leu Thr Met Asp Cys Thr Thr
           500          505          510
Gly Glu Ser Arg Asp Ser Ala Asn Leu Lys Leu Pro Gly Val Gln Glu
           515          520          525
Glu Leu Ile Ile Glu Val Ser Lys Val Gly Lys Pro Val Val Leu Ala
           530          535          540
Leu Leu Asn Gly Arg Pro Tyr Ser Leu Thr Arg Val Val Asp Lys Val
           545          550          555          560
Ser Ala Ile Val Glu Ala Trp Leu Pro Gly Glu Ile Gly Ala Lys Ala
           565          570          575
Ile Val Asp Val Leu Tyr Gly Lys Val Asn Pro Ser Gly Lys Leu Pro
           580          585          590
Met Thr Phe Pro Arg Ser Ala Gly Gln Ile Pro Leu Phe His Tyr Phe
           595          600          605
Lys Pro Ser Gly Gly Arg Ser Ser Trp His Gly Asp Tyr Val Asp Glu
           610          615          620
Ser Val Lys Pro Leu Phe Pro Phe Gly His Gly Leu Ser Tyr Thr Asn
           625          630          635          640
Phe Asp Tyr Ser Gly Leu Glu Ile Ser Pro Ser Lys Val Pro Met Ala
           645          650          655
Gly Ser Val Glu Ile Ser Leu Tyr Val Glu Asn Thr Gly Glu Val Glu
           660          665          670

```

Gly Glu Glu Val Val Gln Leu Tyr Ile Gly Arg Glu Cys Ala Ser Val  
 675 680 685  
 Thr Arg Pro Val Lys Glu Leu Lys Gly Phe Ala Lys Val Asn Leu Lys  
 690 695 700  
 Pro Gly Glu Lys Arg Lys Val Leu Phe Asn Leu His Thr Asp Val Leu  
 705 710 715 720  
 Ala Phe Tyr Gly Arg Asp Met Lys Leu Cys Val Glu Pro Gly Val Tyr  
 725 730 735  
 Asn Val Met Ile Gly Ser Ser Ser Asp Asp Ile Arg Leu Lys Gly Ser  
 740 745 750  
 Phe Glu Val Asp Gly Met Arg Arg Glu Val Phe Glu Asp Arg Val Phe  
 755 760 765  
 Phe Thr Lys Val Tyr Thr Phe  
 770 775