



(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)



1-0026841

(51)⁷ C12N 15/09; C12N 1/15; C12N 1/19; (13) B
C12P 19/14; C12N 9/14; A01H 5/00;
C12N 1/21

(21) 1-2015-02708

(22) 27/03/2014

(86) PCT/JP2014/058798 27/03/2014

(87) WO2014/157492 02/10/2014

(30) PCT/JP2013/059028 27/03/2013 JP

(45) 25/12/2020 393

(43) 25/01/2016 334A

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

1-1, Minami-Aoyama 2-chome, Minato-ku, Tokyo 1078556, Japan

2. KAZUSA DNA RESEARCH INSTITUTE (JP)

2-6-7, Kazusa-kamatari, Kisarazu-shi, Chiba 2920818, Japan

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

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

(54) XENLOBIOHYĐROLAZA CHỊU NHIỆT, PHƯƠNG PHÁP SẢN XUẤT VÀ HỖN HỢP XENLULAZA CHỨA ENZYM NÀY

(57) Sáng chế đề cập đến xenlobiohydrolaza chịu nhiệt, polynucleotit mã hóa enzym này, phương pháp sản xuất polynucleotit này, vectơ biểu hiện enzym này, thể biến nạp chứa vectơ này, phương pháp sản xuất enzym này, hỗn hợp xenlulaza chứa enzym này, phương pháp sản xuất sản phẩm phân giải xenluloza và đoạn môi.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến xenlobiohydrolaza chịu nhiệt. Xenlobiohydrolaza là glycosit hydrolaza có hoạt tính thủy phân lignoxenluloza như xenluloza và hemixenluloza để tạo ra monosacarit. Cụ thể, sáng chế đề cập đến xenlobiohydrolaza chịu nhiệt thể hệ mới, 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 xenluloza bằng enzym này.

Tình trạng kỹ thuật của sáng chế

Gần đây, việc phát triển các nguồn năng lượng thay thế dầu mỏ ngày càng trở nên quan trọng, do quan ngại liên quan đến việc cung cấp năng lượng cho các phương tiện vận tải, như sự tăng giá dầu và nguy cơ cạn kiệt dầu khí trong tương lai gần, cũng như các vấn đề về môi trường như trái đất nóng lên và ô nhiễm bầu khí quyển. Sinh khối thực vật hoặc lignoxenluloza, là nguồn năng lượng tái tạo phong phú nhất trên trái đất, được mong đợi là nguồn năng lượng thay thế cho dầu mỏ. Các thành phần chính trong phần khô của sinh khối này bao gồm các polysacarit như xenluloza, hemixenluloza, và lignin. Ví dụ, các polysacarit được sử dụng làm nhiên liệu sinh học hoặc nguyên liệu để sản xuất nhiều sản phẩm hóa học, sau khi được thủy phân thành monosacarit như glucoza hoặc xyloza bởi các glycosit hydrolaza (cũng được gọi là xenlulaza).

Lignoxenluloza là hợp chất có cấu trúc rất phức tạp, nên rất khó có thể phân giải được bằng một enzym phân giải xenluloza duy nhất. Quá trình phân giải lignoxenluloza thành đường cần ít nhất ba loại enzym: endoglucanaza (xenlulaza hoặc endo-1,4- β -D-glucanaza, mã phân loại enzym số 3.2.1.4) phân cắt một cách ngẫu nhiên ở các vị trí bên trong mạch xenluloza, xenlobiohydrolaza (1,4- β -xenlobiosidaza hoặc xenlobiohydrolaza, mã phân loại enzym số 3.2.1.91) đóng vai trò như một exo-xenlulaza phân cắt ở các đầu khỉ hoặc các đầu không khỉ của mạch xenluloza và giải phóng sản

phẩm chính là xenlobioza, và β -glucosidaza (mã phân loại enzym số 3.2.121) thủy phân xenlobioza thành glucoza. Ngoài ra, cần có hỗn hợp chứa nhiều enzym bao gồm xylanaza (endo-1,4- β -xylanaza, mã phân loại enzym số 3.2.1.8) - là hemixenlulaza và các enzym phân giải thành tế bào thực vật khác.

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 hầu như không thể xảy ra do độ nhớt cao của các hỗn dịch cơ chất. Các enzym chịu nhiệt có ưu điểm ở chỗ cho phép sử dụng nồng độ cơ chất cao, do độ nhớt cơ chất giảm khi nhiệt độ tăng. Hơn nữa, nhiệt độ cao thường làm tăng tốc độ phản ứng xúc tác theo định luật van't Hoff Arrhenius, và tăng cường tính thấm của enzym và phân giải thành tế bào của nguyên liệu. Do đó, khi lignoxenluloza được thủy phân ở nhiệt độ cao hơn nhiệt độ thường bằng enzym chịu nhiệt, sẽ đạt được hiệu quả biến đổi sinh khối thành đường cao hơn, giảm lượng enzym cần sử dụng và rút ngắn thời gian thủy phân, nhờ đó giảm chi phí sử dụng enzym.

Giới hạn nhiệt độ để sinh trưởng của nấm sợi ưa nhiệt bằng khoảng 55°C, thấp hơn giới hạn nhiệt độ để sinh trưởng của vi khuẩn ưa nhiệt và vi khuẩn cổ siêu chịu nhiệt. Do đó, độ ổn định nhiệt của glycosit hydrolaza được biểu hiện hoặc bài tiết từ nấm sợi ưa nhiệt thường không cao. Cho đến nay, đã ghi nhận được rằng xenlobiohydrolaza CBH có nguồn gốc từ nấm sợi có độ ổn định nhiệt cao nhất là xenlobiohydrolaza CBHI và xenlobiohydrolaza CBHII của nấm sợi ưa nhiệt *Chaetomium thermophilum*, lần lượt có nhiệt độ tối ưu bằng 75°C và 70°C (ví dụ, tham khảo Tài liệu phi Sáng chế 1), và xenlobiohydrolaza CBHI có nguồn gốc từ *Thermoascus aurantiacus* có nhiệt độ tối ưu bằng 65°C (ví dụ, tham khảo Tài liệu phi Sáng chế 2). Độ ổn định nhiệt của xenlobiohydrolaza cũng có thể được cải thiện bằng cách thay thế một hoặc nhiều axit amin trong trình tự axit amin của xenlobiohydrolaza (ví dụ, tham khảo Tài liệu Sáng chế 1 hoặc Tài liệu Sáng chế 2). Tuy nhiên, độ ổn định nhiệt của xenlobiohydrolaza đột biến thu được không cao.

Mặt khác, đã phân lập và nuôi cấy được các loài sinh vật ưa nhiệt sinh trưởng ở nhiệt độ 55°C hoặc cao hơn trong môi trường khắc nghiệt như suối nước nóng, miệng phun thủy

nhiệt, mỡ dầu, hoặc mỡ kim loại, hoặc các loài siêu ưa nhiệt sinh trưởng ở nhiệt độ 80°C hoặc cao hơn. Hầu hết các enzym glycosit hydrolaza chịu nhiệt có nguồn gốc từ các vi khuẩn ưa nhiệt và vi khuẩn cổ siêu chịu nhiệt này đều có hoạt tính endoglucanaza, xylanaza, xylosidaza, hoặc glucosidaza. Liên quan đến các xenlobiohydrolaza có vai trò quan trọng nhất để thủy phân lignoxenluloza, thì chỉ một số xenlobiohydrolaza được phân lập từ ba loài vi khuẩn ưa nhiệt thuộc chi *Clostridium*, chi *Thermobifida*, và chi *Thermotoga*. Ví dụ, vi khuẩn kỵ khí ưa nhiệt *Clostridium thermocellum* chứa phức hệ enzym được gọi là xenlulosom có hoạt tính thủy phân lignoxenluloza mạnh, để bài tiết ra ngoài cơ thể vi khuẩn. Các phức hệ enzym xenlulosom chính là xenlobiohydrolaza, và ba biến thể của nó, tức là gen CelO thuộc họ GH5, và gen CbhA và gen CelK thuộc họ GH9, đã được phân lập. Toàn bộ các enzym này có nhiệt độ tối ưu ($T_{\text{tối ưu}}$) nằm trong khoảng từ 60°C đến 65°C (ví dụ, tham khảo Tài liệu phi Sáng chế 3 đến 5). Đã phân lập được hai loại gen xenlobiohydrolaza khác nhau từ xạ khuẩn ưa nhiệt *Thermobifida fusca* bao gồm gen E3 thuộc họ GH6 (ví dụ, tham khảo Tài liệu phi Sáng chế 6), và gen Cel48A thuộc họ GH48 (ví dụ, tham khảo Tài liệu phi Sáng chế 7). Các xenlobiohydrolaza này có độ ổn định nhiệt tương đối cao. Khoảng nhiệt độ mà tại đó các enzym này có hoạt tính bằng 50% hoạt tính tối đa nằm trong khoảng từ 40°C đến 60°C, và hoạt tính được duy trì ổn định ở nhiệt độ 55°C trong ít nhất 16 giờ. Tuy nhiên, hai xenlobiohydrolaza này không có hoạt tính ở nhiệt độ 70°C hoặc cao hơn. Khi xenluloza được thủy phân bằng hai xenlobiohydrolaza này, giới hạn nhiệt độ cho quá trình này nằm trong khoảng từ 60°C đến 65°C. Xenlobiohydrolaza có nguồn gốc từ vi khuẩn ưa nhiệt thuộc chi *Thermotoga* có độ ổn định nhiệt cao nhất, và có $T_{\text{tối ưu}}$ bằng 105°C và thời gian hoạt tính giảm 50% ($T_{\text{bán giảm}}$) bằng 70 phút ở nhiệt độ 108°C (ví dụ, tham khảo Tài liệu phi Sáng chế 8). Tuy nhiên, enzym này có độ đặc hiệu cơ chất tương tự endoglucanaza, và chỉ có hoạt tính phân giải xenluloza vô định hình và carboxymetyl xenluloza (CMC). Hơn nữa, do có hoạt tính thủy phân giấy lọc yếu, nên enzym này cũng không có khả năng thủy phân lignoxenluloza tinh thể.

Tài liệu viện dẫn

Tài liệu Sáng chế 1: Bản dịch tiếng Nhật đã công bố số 2006-515506 của Đơn PCT.

Tài liệu Sáng chế 2: Đơn Công bố Patent Nhật Bản số 2012-39967.

Tài liệu Sáng chế 3: Đơn Công bố Patent Nhật Bản số 2005-237233.

Tài liệu Sáng chế 4: Đơn Công bố Patent Nhật Bản số 2007-53994.

Tài liệu Sáng chế 5: Đơn Công bố Patent Nhật Bản số 2008-193953.

Tài liệu phi Sáng chế 1: Ganju *et al.*, Biochim. Biophys. Acta, 1989, vol. 993, p,266-274.

Tài liệu phi Sáng chế 2: Hong *et al.*, Appl Microbiol Biotechnol., 2003, vol. 63, p,42-50.

Tài liệu phi Sáng chế 3: Zverlov *et al.*, Microbiology, 2002, vol. 148, p,247-255.

Tài liệu phi Sáng chế 4: Zverlov *et al.*, Microbiology, 1997, vol. 143, p,3537-3542.

Tài liệu phi Sáng chế 5: Kataeva *et al.*, JouARNI of Bacteriology, 1999, vol. 181, p,5288-5295.

Tài liệu phi Sáng chế 6: Zhang *et al.*, Biochemistry, 1995, vol. 34, p,3386-3395.

Tài liệu phi Sáng chế 7: Irwin *et al.*, Eur J Biochem., 2000, vol. 267, p,4988-4997.

Tài liệu phi Sáng chế 8: Ruttersmith and Daniel, Biochemical JouARNI, 1991, vol. 277, p,887-890.

Tài liệu phi Sáng chế 9: Herai, S. *et al.*, Proc. Natl. Acad. Sci. USA, 2004, vol. 101, p. 14031-14035

Tài liệu phi Sáng chế 10: Ogino, S., Appl. Microbiol. Biotechnol., 2004, vol. 64, p. 823-828.

Tài liệu phi Sáng chế 11: Tamura, T. *et al.*, J. Environmental Biotechnol., 2007, vol. 7, p. 3-10.

Tài liệu phi Sáng chế 12: DN. Bolam *et al.*, Biochem. J., 1998, vol. 331, p. 775-781.

Tài liệu phi Sáng chế 13: N. DIN *et al.*, 1994, Proc. Natl. Acad. Sci. USA, vol. 91, p. 11383-11387.

Tài liệu phi Sáng chế 14: K. Riedel *et al.*, FEMS Microbiology Letters, 1998, vol.

164, p. 261-267.

Tài liệu phi Sáng chế 15: Mahadevan SA, Wi SG, Lee DS, Bae HJ: Site-directed mutagenesis và CBM engineering of Cel5A (*Thermotoga maritima*). FEMS Microbiol Lett., 2008, vol. 287, p. 205-211.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề cập đến xenlobiohydrolaza chịu nhiệt thể hệ mới có hoạt tính ít nhất ở điều kiện nhiệt độ bằng 75°C, 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 xenluloza bằng enzym này.

Để giải quyết các vấn đề nêu trên, các tác giả sáng chế đã thu nhận thành công xenlobiohydrolaza chịu nhiệt có các 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 vi sinh vật khó nuôi cấy lấy từ môi trường này và hoàn thành sáng chế này.

Xenlobiohydrolaza chịu nhiệt, polynucleotit, vector biểu hiện, thể biến nạp, phương pháp sản xuất xenlobiohydrolaza chịu nhiệt, hỗn hợp xenlulaza, phương pháp sản xuất sản phẩm phân giải xenluloza, phương pháp sản xuất polynucleotit và đoạn mồi theo sáng chế được mô tả trong các phương án từ 1 đến 17 dưới đây.

1. Xenlobiohydrolaza chịu nhiệt chứa vùng xúc tác xenlobiohydrolaza bao gồm:

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

(B) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

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

(D) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(E) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.5;

(F) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(G) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.7; hoặc

(H) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

2. Xenlobiohydrolaza chịu nhiệt theo phương án 1, trong đó enzym này còn chứa đoạn gắn kết xenluloza.

3. Polynucleotit chứa miễn mã hóa vùng xúc tác xenlobiohydrolaza bao gồm:

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

(b) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

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

(d) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(e) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.5;

(f) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng

nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(g) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.7;

(h) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5; hoặc

(i) trình tự nucleotit có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8, và mã hóa polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

4. Polynucleotit theo phương án 3, trong đó polynucleotit này còn chứa miền mã hóa đoạn gắn kết xenluloza.

5. Vector biểu hiện chứa polynucleotit theo phương án 3 hoặc 4, và có khả năng biểu hiện polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5 ở tế bào vật chủ.

6. Thẻ biến nạp chứa vector theo phương án 5.

7. Thẻ biến nạp theo phương án 6, trong đó thẻ biến nạp này là vi sinh vật nhân sơ.

8. Thẻ biến nạp theo phương án 6, trong đó thẻ biến nạp này là vi sinh vật nhân thật.

9. Thẻ biến nạp theo phương án 6, trong đó thẻ biến nạp này là thực vật.

10. Phương pháp sản xuất xenlobiohydrolaza chịu nhiệt bao gồm bước sản sinh xenlobiohydrolaza chịu nhiệt trong thẻ biến nạp theo phương án bất kỳ trong số các phương án từ 6 đến 9.

11. Hỗn hợp xenlulaza chứa xenlobiohydrolaza chịu nhiệt theo phương án 1,

xenlobiohydrolaza chịu nhiệt theo phương án 2, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 10, và ít nhất một hoặc nhiều loại xenlulaza khác.

12. Hỗn hợp xenlulaza theo phương án 11, trong đó ít nhất một hoặc nhiều loại xenlulaza khác được chọn từ nhóm bao gồm hemixenlulaza và endoglucanaza.

13. Phương pháp sản xuất sản phẩm phân giải xenluloza bao gồm bước tạo ra sản phẩm phân giải xenluloza bằng cách cho nguyên liệu chứa xenluloza tiếp xúc với xenlobiohydrolaza chịu nhiệt theo phương án 1, xenlobiohydrolaza chịu nhiệt theo phương án 2, thể biến nạp theo phương án bất kỳ trong số các phương án từ 6 đến 9, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 10.

14. Phương pháp theo phương án 13, trong đó nguyên liệu chứa xenluloza còn được cho tiếp xúc với ít nhất một hoặc nhiều loại xenlulaza khác.

15. Phương pháp theo phương án 14, trong đó ít nhất một hoặc nhiều loại xenlulaza khác được chọn từ nhóm bao gồm hemixenlulaza và endoglucanaza.

16. Phương pháp sản xuất polynucleotit mã hóa xenlobiohydrolaza chịu nhiệt bao gồm bước thực hiện phản ứng PCR bằng cách sử dụng khuôn là ADN thu được từ sinh vật hoặc sản phẩm phiên mã ngược của ARN thu được từ sinh vật, với đoạn môi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.12 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12, và đoạn môi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13, và thu nhận sản phẩm khuếch đại là polynucleotit chứa trình tự nucleotit mã hóa xenlobiohydrolaza chịu nhiệt.

17. Cặp đoạn môi bao gồm đoạn môi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.12, hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12 và đoạn môi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13, hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình vẽ thể hiện cây phả hệ trình tự axit amin mã hóa bởi 4 khung đọc mã mở (AR19G-166, AR19G-12 (OJ1-2), và OJ1-1) thuộc họ GH6 thu được từ phân tích hệ gen môi trường trong Ví dụ 1. Do có trình tự axit amin tương đồng 100% với gen AR19G-12, nên gen OJ1-2 giống hệt với gen AR19G-12 và là một đoạn trình tự của gen AR19G-12..

Fig.2A là hình vẽ thể hiện sự sắp thẳng hàng các trình tự axit amin mã hóa các khung đọc mã mở AR19G-166, AR19G-12, và OJ1-1, và trình tự axit amin của vùng xúc tác của glycosit hydrolaza GH6 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt độ trung bình *Herpetosiphon aurantiacus* DSM 785 thuộc ngành *Chloroflexi*, có độ tương đồng cao nhất với các trình tự axit amin này.

Fig.2B là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1, và trình tự axit amin của đoạn gắn kết xenluloza CBM3 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt *Caldibacillus cellulovorans*.

Fig.3A là sơ đồ thể hiện các polypeptit có trình tự axit amin mã hóa bởi các khung đọc mã mở AR19G-166, AR19G-12, và OJ1-1, và polypeptit mã hóa bởi trình tự nucleotit của glycosit hydrolaza GH6 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt độ trung bình *Herpetosiphon aurantiacus* DSM 785 thuộc ngành *Chloroflexi*, có độ tương đồng cao nhất với các trình tự axit amin này.

Fig.3B là sơ đồ thể hiện polypeptit có trình tự axit amin trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1 và đoạn gắn kết xenluloza CBM3 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt *Caldibacillus cellulovorans*.

Fig.4A là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE và Fig.4B là hình vẽ thể hiện kết quả phân tích thấm tách Western của protein AR19G-166-RA và protein AR19G-166-QV được biểu hiện bởi *E. coli* trong Ví dụ 1.

Fig.5A là hình vẽ thể hiện kết quả phân tích sản phẩm thủy phân PSA bởi protein AR19G-166-RA được biểu hiện bởi *E. coli*, bằng phương pháp sắc ký lỏng hiệu năng cao (HPLC), trong Ví dụ 1.

hình vẽ thể hiện kết quả phân tích thẩm tách Western của protein AR19G-166-RA và protein AR19G-166-QV được biểu hiện bởi *E. coli* trong Ví dụ 1.

Fig.5A là hình vẽ thể hiện kết quả phân tích sản phẩm thủy phân PSA bởi protein AR19G-166-RA được biểu hiện bởi *E. coli*, bằng phương pháp sắc ký lỏng hiệu năng cao (HPLC), trong Ví dụ 1.

Fig.5B là hình vẽ thể hiện kết quả phân tích sản phẩm thủy phân PSA bởi xenlobiohydrolaza TrCBHII GH6 có nguồn gốc từ nấm sợi *Trichoderma reesei*, bằng phương pháp sắc ký lỏng hiệu năng cao (HPLC), trong Ví dụ 1.

Fig.6A là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA của protein AR19G-166-RA được biểu hiện bởi *E. coli* đo ở các nhiệt độ tương ứng trong Ví dụ 1.

Fig.6B là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA của protein AR19G-166-QV được biểu hiện bởi *E. coli* đo ở các nhiệt độ tương ứng trong Ví dụ 1.

Fig.7A là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA của protein AR19G-166-RA được biểu hiện bởi *E. coli* đo ở các độ pH tương ứng trong Ví dụ 1.

Fig.7B là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA của protein AR19G-166-QV được biểu hiện bởi *E. coli* đo ở các độ pH tương ứng trong Ví dụ 1.

Fig.8A là đồ thị thể hiện kết quả đánh giá mức độ ảnh hưởng của thời gian ủ sơ bộ đến hoạt tính của protein AR19G-166-RA được biểu hiện bởi *E. coli* được đo trong Ví dụ 1.

Fig.8B là đồ thị thể hiện kết quả đánh giá mức độ ảnh hưởng của thời gian ủ sơ bộ đến hoạt tính của protein AR19G-166-QV được biểu hiện bởi *E. coli* được đo trong Ví dụ 1.

Fig.9A là đồ thị thể hiện kết quả đánh giá mức độ ảnh hưởng của nhiệt độ ủ sơ bộ đến hoạt tính của protein AR19G-166-RA được biểu hiện bởi *E. coli* được đo trong Ví dụ 1.

Fig.9B là đồ thị thể hiện kết quả đánh giá mức độ ảnh hưởng của nhiệt độ ủ sơ bộ đến hoạt tính của protein AR19G-166-QV được biểu hiện bởi *E. coli* được đo trong Ví dụ 1.

Fig.10 là hình vẽ thể hiện kết quả phân tích thẩm tách Western của dịch nổi nuôi cấy *Aspergillus* chuyển gen AR19G-166-RW và *Aspergillus* chuyển gen R19G-166-QW trong Ví dụ 2, và dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen AR19G-166-RA và gen AR19G-166-QV được sản xuất trong Ví dụ 1.

Fig.11A là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA (U/mg) của protein AR19G-166-RW và protein AR19G-166-QW được biểu hiện bởi *E. coli* trong Ví dụ 1, và hoạt tính thủy phân PSA (trị số tương đối (%)) của protein AR19G-166-RW và protein AR19G-166-QW được biểu hiện bởi *Aspergillus* trong Ví dụ 2, ở các nhiệt độ tương ứng.

Fig.11B là đồ thị thể hiện kết quả đánh giá hoạt tính thủy phân PSA (U/mg) của protein AR19G-166-RW và protein AR19G-166-QW được biểu hiện bởi *E. coli* trong Ví dụ 1, và hoạt tính thủy phân PSA (trị số tương đối (%)) của protein AR19G-166-RW và protein AR19G-166-QW được biểu hiện bởi *Aspergillus* trong Ví dụ 2, ở các độ pH tương ứng.

Fig.12A là hình vẽ thể hiện vector cat-xet pNtaGL và pNtaGLPL để chuyển gen vào lục lạp của cây thuốc lá, sử dụng để sản xuất cây thuốc lá chuyển gen lục lạp trong Ví dụ 3.

Fig.12B là hình vẽ thể hiện cat-xet pPXT và pPXTPL để chuyển gen vào lục lạp của cây thuốc lá, sử dụng để sản xuất cây thuốc lá chuyển gen lục lạp trong Ví dụ 3.

Fig.12C là hình vẽ thể hiện vector cat-xet pNtaGL-QV và pNtaGLPL-RA để chuyển gen vào lục lạp của cây thuốc lá, sử dụng để sản xuất cây thuốc lá chuyển gen lục lạp trong Ví dụ 3.

Fig.13A là hình vẽ thể hiện kết quả phân tích thẩm tách Southern của hai dòng cây thuốc lá chuyển gen lục lạp pNtaGL-QV (QV-2 và QV-17) và cây thuốc lá thể đại (WT(SR-1)) trong Ví dụ 3.

Fig.13B là hình vẽ thể hiện kết quả phân tích thẩm tách Southern của ba dòng cây thuốc lá chuyển gen lục lạp pNtaGLPL-RA (RA-6-2-1, RA-6-2-2, và RA-6-2-3) và cây thuốc lá thể đại (WT(SR-1)) trong Ví dụ 3.

Fig.13C là hình vẽ thể hiện kết quả phân tích thẩm tách Southern của cây thuốc lá

chuyển gen lục lạp pNtaGL và pNtaGLPL trong Ví dụ 3.

Fig.14A là hình ảnh thể hiện cây thuốc lá chuyển gen lục lạp AR19G-166-QV (thế hệ T₁), và cây thuốc lá chuyển gen lục lạp vector đối chứng pNtaGL (thế hệ T₁) ở giai đoạn ra hoa, trong Ví dụ 3.

Fig.14B là hình ảnh thể hiện cây thuốc lá chuyển gen lục lạp AR19G-166-RA (thế hệ T₁), và cây thuốc lá chuyển gen lục lạp vector đối chứng pNtaGLPL (thế hệ T₁) ở giai đoạn ra hoa, trong Ví dụ 3.

Fig.15A là biểu đồ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein hòa tan chiết xuất từ cây thuốc lá chuyển gen lục lạp AR19G-166-QV và cây thuốc lá chuyển gen lục lạp pNtaGL trong Ví dụ 3.

Fig.15B là hình vẽ thể hiện kết quả phân tích thấm tách Western của protein hòa tan chiết xuất từ cây thuốc lá chuyển gen lục lạp AR19G-166-QV và cây thuốc lá chuyển gen lục lạp pNtaGL trong Ví dụ 3.

Fig.15C là biểu đồ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein hòa tan chiết xuất từ cây thuốc lá chuyển gen lục lạp AR19G-166-RA và cây thuốc lá chuyển gen lục lạp pNtaGLPL trong Ví dụ 3.

Fig.15D là hình vẽ thể hiện kết quả phân tích thấm tách Western của protein hòa tan chiết xuất từ cây thuốc lá chuyển gen lục lạp AR19G-166-RA và cây thuốc lá chuyển gen lục lạp pNtaGLPL trong Ví dụ 3.

Fig.16A là đồ thị thể hiện hoạt tính thủy phân PSA của protein AR19G-166-QV được biểu hiện ở lục lạp của cây thuốc lá trong Ví dụ 3 tính theo hàm lượng đường khử được tạo ra (mM) ở các nhiệt độ khác nhau.

Fig.16B là đồ thị thể hiện hoạt tính thủy phân PSA của protein AR19G-166-RA được biểu hiện ở lục lạp của cây thuốc lá trong Ví dụ 3 tính theo hàm lượng đường khử được tạo ra (mM) ở các nhiệt độ khác nhau.

Fig.17 là hình vẽ thể hiện kết quả phân tích thấm tách Western của protein AR19G-166-RA và protein AR19G-166-QV được biểu hiện ở *Arabidopsis thaliana* trong Ví dụ 4.

Fig.18 là đồ thị thể hiện mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA của protein AR19G-166-RA và protein AR19G-166-QV được biểu hiện trong *Arabidopsis thaliana* trong Ví dụ 4 (tính theo đơn vị mM đường khử/20 phút).

Fig.19 là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của các dịch chiết không chứa tế bào chiết xuất từ xạ khuẩn tái tổ hợp gen AR19G-166-RA trong Ví dụ 5.

Fig.20 là đồ thị thể hiện mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân Avicel trong nở liên hợp axit phosphoric (Phosphoric acid-Swollen Avicel - PSA) của protein enzym (dịch chiết không chứa tế bào) mã hóa bởi gen AR19G-166-RA, được biểu hiện trong xạ khuẩn *Streptomyces lividans* trong Ví dụ 5.

Fig.21A là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE của protein enzym được tạo ra bằng cách biểu hiện gen AR19G-166-RA bổ sung CBM trong *E. coli* trong Ví dụ 6.

Fig.21B là hình vẽ thể hiện kết quả phân tích thấm tách Western của protein enzym được tạo ra bằng cách biểu hiện gen AR19G-166-RA bổ sung CBM trong *E. coli* trong Ví dụ 6.

Fig.22A là đồ thị thể hiện mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA (U/mg) của protein AR19G-166-RA bổ sung CBM được biểu hiện trong *E. coli* trong Ví dụ 6.

Fig.22B là đồ thị thể hiện mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân Avicel (U/mg) của protein AR19G-166-RA bổ sung CBM được biểu hiện trong *E. coli* trong Ví dụ 6.

Mô tả chi tiết sáng chế

Xenlobiohydrolaza chịu nhiệt theo sáng chế

Nhiều vi sinh vật bao gồm nấm sợi, vi khuẩn, và vi khuẩn cổ rất khó được nuôi cấy. Khoảng 99% nấm sống trong môi trường vi sinh vật như đất là chưa được biết đến. Cụ thể, rất khó nuôi cấy được vi sinh vật sống trong môi trường có nhiệt độ cao, và người ta cho rằng chỉ 0,1% hoặc ít hơn quần thể toàn bộ vi sinh vật sống trong đất đã được phân lập

và nuôi cấy bằng kỹ thuật nuôi cấy vi sinh vật hiện có. Việc khó nuôi cấy vi sinh vật sống trong đất có nhiệt độ cao là một yếu tố cản trở sự phát triển xenlobiohydrolaza 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 mẫu tương tự. Bằng cách sử dụng kỹ thuật phân tích này, phương pháp phân tích hệ gen môi trường đã được đề cập đến trong đó ADN bộ gen của quần thể vi sinh vật được tổng hợp từ mẫu môi trường như đất, bộ gen của quần thể vi sinh vật có thành phần không đồng nhất được giải trình tự trực tiếp và toàn bộ, và dữ liệu trình tự được lắp ráp bằng máy tính song song, để từ đó 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 trong Ví dụ 1, các tác giả sáng chế đã tổng hợp được ADN bộ gen (AND hệ gen môi trường) của quần thể vi sinh vật từ đất có nhiệt độ cao ở năm khu vực suối nước nóng của Nhật Bản, và thực hiện giải trình tự ADN hệ gen môi trường này. Kết quả là thu được 44 khung đọc mã mở (Open Reading Frame - ORF) có các trình tự axit amin tương đồng với các xenlobiohydrolaza đã biết. Các đoạn môi được thiết kế dựa trên trình tự nucleotit của các khung đọc mã mở này, và các gen tiềm năng được tách dòng từ ADN hệ gen môi trường của đất có nhiệt độ cao ở khu vực suối nước nóng bằng phương pháp PCR. Các ADN được tách dòng bằng 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 đánh giá hoạt tính phân giải Avicel trương nở liên hợp axit phosphoric (Phosphoric acid-Swollen Avicel - PSA) và hoạt tính phân giải carboxymetyl xenluloza (CMC). Cuối cùng, xenlobiohydrolaza chịu nhiệt có hoạt tính phân giải PSA thu được từ một ORF duy nhất.

Hai loại xenlobiohydrolaza chịu nhiệt, trong đó hai axit amin được thay thế thu được từ khung đọc mã mở quan tâm bằng phương pháp tách dòng PCR. Trong số hai loại kiểu gen này, một loại trong đó axit amin ở vị trí số 299 là arginin (R) và axit amin ở vị trí số 351 là alanin (A) được ký hiệu là AR19G-166-RA, và loại còn lại trong đó axit amin ở vị trí số 299 là glutamin (Q) và axit amin ở vị trí số 351 là valin (V) được ký hiệu là AR19G-166-QV. Trình tự nucleotit của AR19G-166-RA được thể hiện trong SEQ ID NO.2, và trình tự axit amin của AR19G-166-RA được thể hiện trong SEQ ID NO.1. Hơn nữa,

trình tự nucleotit của AR19G-166-QV được thể hiện trong SEQ ID NO.4, và trình tự axit amin của AR19G-166-QV được thể hiện trong SEQ ID NO.3. AR19G-166-RA và AR19G-166-QV thu được bằng phương pháp tách dòng PCR không chứa gốc metionin hoạt hóa. Do đó, các trình tự này gợi ý rằng có các đoạn gen chỉ có vùng xúc tác xenlobiohydrolaza của gen xenlobiohydrolaza, vốn đã có trong vi sinh vật chứa trong đất có nhiệt độ cao ở khu vực suối nước nóng.

Như được thể hiện trong Ví dụ 1 và các ví dụ tương tự, AR19G-166-RA và AR19G-166-QV có hoạt tính thủy phân mạnh PSA, và có hoạt tính phân giải lichenan chứa các liên kết glucan β -1,3 và β -1,4 và xenluloza tinh thể Avicel, mặc dù yếu hoạt tính này yếu. Trái lại, AR19G-166-RA và AR19G-166-QV hầu như không có hoạt tính phân giải CMC và laminarin chứa các liên kết glucan β -1,3 và β -1,6. Kết quả phân tích sản phẩm thủy phân PSA bởi enzyme AR19G-166-RA bằng phương pháp sắc ký lỏng hiệu năng cao (HPLC) cho thấy xenlobioza và lượng nhỏ xenlotrioza được tạo ra. Hơn nữa, kết quả tra cứu trình tự axit amin của AR19G-166-RA và AR19G-166-QV từ cơ sở dữ liệu trình tự axit amin đã biết cho thấy các trình tự axit amin này có độ tương đồng cao nhất với glycosit hydrolaza GH6 (SEQ ID NO.15) có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt độ trung bình *Herpetosiphon aurantiacus* DSM 785 thuộc ngành *Chloroflexi*, và chỉ bằng 66%. Các phân tích độ đặc hiệu cơ chất, phân tích sản phẩm thủy phân PSA bằng phương pháp HPLC, và phân tích độ tương đồng của trình tự axit amin này với xenlobiohydrolaza đã biết cho thấy biến thể AR19G-166-RA và AR19G-166-QV là xenlobiohydrolaza GH6 thế hệ mới.

Cả AR19G-166-RA và AR19G-166-QV đều có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5. Thực tế, như được thể hiện trong Ví dụ 1, cả AR19G-166-RA và AR19G-166-QV đều có hoạt tính xenlobiohydrolaza ở khoảng nhiệt độ rộng nằm trong khoảng từ 30°C đến 100°C. Tuy nhiên, khoảng nhiệt độ tối ưu của các xenlobiohydrolaza này khác nhau. Hoạt tính xenlobiohydrolaza của AR19G-166-RA được biểu hiện bởi *E. coli* có xu hướng tăng khi nhiệt độ tăng nằm trong khoảng từ 30°C đến 80°C, và hoạt tính xenlobiohydrolaza có xu hướng giảm khi nhiệt độ tăng nằm trong khoảng từ 80°C đến 100°C. Trái lại, hoạt tính xenlobiohydrolaza của AR19G-166-QV có xu hướng tăng khi nhiệt độ tăng nằm trong khoảng từ 30°C đến 70°C, đạt cực đại khoảng 70°C, và có xu hướng giảm khi nhiệt độ tăng nằm trong khoảng từ 70°C đến 100°C.

Thuật ngữ “hoạt tính xenlobiohydrolaza” được dùng trong bản mô tả để chỉ hoạt tính đối với ít nhất một cơ chất được chọn từ nhóm bao gồm glucan chứa liên kết β -1,3 và β -1,4 và xenluloza tinh thể, và Avicel tương tự liên hợp axit phosphoric, và cơ chất này được thủy phân từ đầu không khử để tạo ra xenlobioza.

Thuật ngữ “polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5” được dùng trong bản mô tả để chỉ hoạt tính xenlobiohydrolaza mạnh nhất ở nhiệt độ 75°C khi độ pH của dung dịch chứa polypeptit bằng 5,5. Nói cách khác, thậm chí khi dung dịch chứa polypeptit không có hoạt tính xenlobiohydrolaza ở độ pH không bằng 5,5 và nhiệt độ không bằng 75°C, nhưng dung dịch này có hoạt tính xenlobiohydrolaza ở độ pH bằng 5,5 và nhiệt độ 75°C, thì polypeptit cũng được coi là có hoạt tính xenlobiohydrolaza.

Tốt hơn nếu “xenlobiohydrolaza chịu nhiệt” theo sáng chế là enzyme có hoạt tính ở nhiệt độ nằm trong khoảng từ 55°C đến 80°C và độ pH nằm trong khoảng từ 3,5 đến 7,0, và tốt hơn nữa nếu là enzyme có hoạt tính xenlobiohydrolaza ở nhiệt độ nằm trong khoảng từ 70 đến 100°C và độ pH nằm trong khoảng từ 4,0 đến 6,0.

Enzyme AR19G-166-RW trong đó gốc axit amin ở vị trí số 351 được thế bằng tryptophan (W) trong AR19G-166-RA, và AR19G-166-QW trong đó gốc axit amin ở vị trí số 351 được thế bằng tryptophan (W) trong AR19G-166-QV cũng có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5, tương tự như AR19G-166-RA và AR19G-166-QV. Trình tự axit amin của AR19G-166-RW được thể hiện trong SEQ ID NO.5, và trình tự nucleotit mã hóa trình tự axit amin này được thể hiện trong SEQ ID NO.6. Trình tự axit amin của AR19G-166-QW được thể hiện trong SEQ ID NO.7, và trình tự nucleotit mã hóa trình tự axit amin này được thể hiện trong SEQ ID NO.8. Hoạt tính xenlobiohydrolaza của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *Aspergillus* được tăng khi nhiệt độ tăng nằm trong khoảng từ 30 đến 100°C, và có hoạt tính xenlobiohydrolaza mạnh nhất ở nhiệt độ 100°C.

Nhìn chung, ở protein có một số hoạt tính sinh học, một hoặc nhiều axit amin có thể được loại bỏ, thay thế, hoặc bổ sung, mà không làm giảm hoạt tính sinh học của protein này. Tức là, trong AR19G-166-RA, AR19G-166-QV, AR19G-166-RW, hoặc AR19G-166-QW, một hoặc nhiều axit amin cũng có thể được loại bỏ, thay thế, hoặc bổ sung mà không làm

giảm hoạt tính xenlobiohydrolaza của nó.

Sáng chế đề cập đến xenlobiohydrolaza chịu nhiệt chứa vùng xúc tác xenlobiohydrolaza, gồm polypeptit bất kỳ trong số các polypeptit từ (A) đến (L) sau:

(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 một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(C) polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(D) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.3;

(E) polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(F) polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(G) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.5;

(H) polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(I) polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(J) polypeptit có trình tự axit amin như nêu trong SEQ ID NO.7;

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

sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(L) polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

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 các polypeptit (B), (E), (H), và (K) nêu trên, tốt hơn nếu số lượng axit amin được loại bỏ, thay thế, hoặc bổ sung trên trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3, SEQ ID NO.5, hoặc SEQ ID NO.7 nằm trong khoảng từ 1 đến 20, tốt hơn nữa nếu nằm trong khoảng từ 1 đến 10, và còn tốt hơn nữa nếu nằm trong khoảng từ 1 đến 5. Không có giới hạn cụ thể về vị trí của axit amin cần được loại bỏ, thay thế hoặc bổ sung trên mỗi trình tự axit amin này, mặc dù tốt hơn nếu axit amin cần được loại bỏ, thay thế hoặc bổ sung ở vị trí số 299 là arginin.

Trong các polypeptit (C), (F), (I), và (L) nêu trên, không có giới hạn cụ thể về độ tương đồng với trình tự axit amin như nêu trong SEQ ID NO.1, SEQ ID NO.3, SEQ ID NO.5, hoặc SEQ ID NO.7 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 95% hoặc lớn hơn.

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), (C), (E), (F), (H), (I), (K), và (L) nêu trên có thể được thiết kế nhân tạo hoặc cũng có thể là thể tương đồng của AR19G-166-QV và các biến thể tương tự, hoặc các đoạn protein của chúng.

Các polypeptit (A) đến (L) nêu trên 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 khía cạnh thứ hai của sáng chế. Hơn nữa, các polypeptit (B), (C), (E), (F), (H), (I), (K), và (L) 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, SEQ ID NO.3, SEQ ID NO.5, hoặc SEQ ID NO.7, bằng kỹ thuật tái tổ hợp gen để chèn thêm đột biến axit amin.

Các polypeptit (A) đến (L) nêu trên có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5. Do đó, khi enzym chứa vùng xúc tác xenlobiohydrolaza là polypeptit bất kỳ từ (A) đến (L), thì có thể thu được xenlobiohydrolaza chịu nhiệt. Trong số này, tốt hơn nếu polypeptit bất kỳ từ (D) đến (L) được sử dụng làm vùng xúc tác xenlobiohydrolaza do chúng có hoạt tính xenlobiohydrolaza mạnh thậm chí ở nhiệt độ nằm trong khoảng từ 70°C đến 100°C.

Xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất Avicel trương nở liên hợp axit phosphoric (PSA). Xenlobiohydrolaza chịu nhiệt này cũng có hoạt tính thủy phân cơ chất β glucan khác. Ví dụ về các cơ chất β glucan khác bao gồm lichenan chứa liên kết β -1,3 và liên kết β -1,4, xenluloza tinh thể như Avicel, xenluloza vi tinh thể của vi khuẩn (Bacterial Microcrystalline Cellulose - BMCC) và giấy lọc, carboxymetylxenluloza (CMC), glucan chứa liên kết β -1,3 và liên kết β -1,6, glucan chứa liên kết β -1,3, glucan chứa liên kết β -1,6, xylan, và các cơ chất tương tự. Tốt hơn nếu xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất PSA và ít nhất một glucan chứa liên kết β -1,3 và liên kết β -1,4 và xenluloza tinh thể, và tốt hơn nữa nếu xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính thủy phân cơ chất PSA, glucan chứa liên kết β -1,3 và liên kết β -1,4, và xenluloza tinh thể.

Mặc dù thay đổi phụ thuộc vào nhiệt độ phản ứng, nhưng độ pH tối ưu của xenlobiohydrolaza chịu nhiệt theo sáng chế nằm trong khoảng từ 4,5 đến 6,0. Tốt hơn nếu xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính ở độ pH ít nhất nằm trong khoảng từ 4,5 đến 6,0, tốt hơn nữa nếu xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính ở độ pH nằm trong khoảng từ 4,0 đến 6,5, và còn tốt hơn nữa nếu xenlobiohydrolaza chịu nhiệt theo sáng chế có hoạt tính ở độ pH nằm trong khoảng từ 3,5 đến 7,0.

Xenlobiohydrolaza chịu nhiệt theo sáng chế có thể có hoạt tính thủy phân xenluloza khác với hoạt tính xenlobiohydrolaza. Ví dụ về các hoạt tính thủy phân xenluloza khác bao gồm hoạt tính endoglucanaza, hoạt tính xylanaza, hoạt tính β -glucosidaza và các hoạt tính tương tự.

Xenlobiohydrolaza chịu nhiệt theo sáng chế có thể chỉ chứa vùng xúc tác xenlobiohydrolaza bao gồm polypeptit bất kỳ từ (A) đến (L) nêu trên, hoặc có thể chứa các vùng khác. Ví dụ về các vùng khác bao gồm các vùng khác với vùng xúc tác của xenlobiohydrolaza đã biết. Ví dụ, xenlobiohydrolaza chịu nhiệt theo sáng chế bao gồm các enzym thu được bằng cách thay thế vùng xúc tác xenlobiohydrolaza của các xenlobiohydrolaza đã biết bằng các polypeptit từ (A) đến (L) nêu trên.

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

Đoạn gắn kết xenluloza chứa trong xenlobiohydrolaza chịu nhiệt theo sáng chế có thể thích hợp khi nó là vùng có khả năng gắn kết với xenluloza, ví dụ vùng có khả năng gắn kết với PSA hoặc Avicel tinh thể. Không có giới hạn cụ thể về trình tự axit amin của đoạn gắn kết xenluloza theo sáng chế. Ví dụ, đoạn gắn kết xenluloza của protein đã biết hoặc đoạn cải biến thích hợp của nó có thể được sử dụng làm đoạn gắn kết xenluloza theo sáng

chế. Theo sáng chế, tốt hơn nếu đoạn gắn kết xenluloza là polypeptit chứa 148 axit amin (T35-P182) từ threonin (T) ở vị trí số 35 đến prolin (P) ở vị trí số 182 của trình tự axit amin như nêu trong SEQ ID NO.11, hoặc polypeptit có trình tự axit amin trong đó một hoặc nhiều axit amin được loại bỏ, thay thế hoặc bổ sung vào polypeptit này, cũng như có khả năng gắn kết với xenluloza.

Khi xenlobiohydrolaza chịu nhiệt theo sáng chế chứa vùng xúc tác xenlobiohydrolaza và đoạn gắn kết xenluloza, tốt hơn nếu các vùng này được liên kết với nhau thông qua gốc liên kết. Không có giới hạn cụ thể về trình tự axit amin, chiều dài, và các đặc tính tương tự của gốc liên kết này. Ví dụ về các gốc liên kết này bao gồm polypeptit chứa 112 axit amin (S183-T294) từ serin (S) ở vị trí số 183 đến threonin (T) ở vị trí số 294 của trình tự axit amin như nêu trong SEQ ID NO.11, hoặc polypeptit có trình tự axit amin trong đó một hoặc nhiều axit amin được loại bỏ, thay thế hoặc bổ sung vào polypeptit này.

Lưu ý rằng, trình tự axit amin như nêu trong SEQ ID NO.11 là trình tự axit amin của polypeptit mã hóa bởi khung đọc mã mở OJ1-1 thu được từ ADN hệ gen môi trường của quần thể vi sinh vật trong đất có nhiệt độ cao ở khu vực suối nước nóng và gen OJ1-1-11 mới này thu được từ khung đọc mã mở quan tâm bằng phương pháp tách dòng PCR, sẽ được mô tả trong Ví dụ 1. Polypeptit chứa 148 axit amin (T35-P182) từ threonin ở vị trí số 35 đến prolin ở vị trí số 182 của OJ1-1-11 được cho là đoạn gắn kết xenluloza CBM3, và polypeptit chứa 112 axit amin (S183-T294) từ serin ở vị trí số 183 đến threonin ở vị trí số 294 được xác định là gốc liên kết.

Ngoài ra, xenlobiohydrolaza chịu nhiệt theo sáng chế có thể còn chứa peptit tín hiệu được định vị vào vùng đặc hiệu ở tế bào, hoặc peptit tín hiệu được bài tiết 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, và các peptit tín hiệu tương tự. Bằng cách bổ sung peptit tín hiệu vào đầu tận cùng N hoặc đầu tận cùng C, thì xenlobiohydrolaza chịu nhiệt theo sáng chế được biểu hiện trong thực vật chuyển gen có thể được định vị trong không bào, mạng lưới nội chất, hoặc các vùng tương tự của tế bào.

Không có giới hạn cụ thể về peptit tín hiệu vận chuyển không bào, miễn là peptit này có khả năng vận chuyển polypeptit vào không bào, và peptit tín hiệu vận chuyển không

bào đã biết có thể được sử dụng thích hợp. Ví dụ về peptit tín hiệu vận chuyển không bào bao gồm peptit tín hiệu của chất ức chế proteaza II ở khoai tây (ví dụ, tham khảo Wang *et al.*, *Transgenic Research*, 2005, vol,14, p167-178), hoặc các peptit tương tự. Hơn nữa, không có giới hạn cụ thể về peptit tín hiệu duy trì mạng lưới nội chất, miễn là peptit này có khả năng giữ polypeptit nằm trong mạng lưới nội chất, và peptit tín hiệu duy trì mạng lưới nội chất đã biết có thể được sử dụng thích hợp. Ví dụ về peptit tín hiệu duy trì mạng lưới nội chất bao gồm peptit tín hiệu chứa trình tự axit amin HDEL, hoặc các peptit tương tự.

Ngoài ra, khi xenlobiohydrolaza chịu nhiệt theo sáng chế được sản xuất bằng hệ biểu hiện, để có thể tinh chế enzym này một cách đơn giản, thì các trình tự đánh dấu khác nhau có thể được bổ sung vào đầu tận cùng N hoặc đầu tận cùng C. 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 xenlobiohydrolaza chịu nhiệt theo sáng chế

Polynucleotit nêu trên mã hóa xenlobiohydrolaza chịu nhiệt theo sáng chế. Xenlobiohydrolaza chịu nhiệt theo sáng chế có thể được tạo ra bằng hệ biểu hiện của vật chủ thu được bằng cách chèn vector biểu hiện chứa polynucleotit này vào vật chủ.

Cụ thể, polynucleotit theo sáng chế chứa miền mã hóa vùng xúc tác xenlobiohydrolaza, gồm trình tự nucleotit bất kỳ trong số các trình tự nucleotit từ (a) đến (n) 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 một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(c) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH

bằng 5,5;

(d) trình tự nucleotit mã hóa polypeptit có trình tự axit amin như nêu trong SEQ ID NO.3;

(e) trình tự nucleotit mã hóa polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(f) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(g) trình tự nucleotit mã hóa polypeptit có trình tự axit amin như nêu trong SEQ ID NO.5;

(h) trình tự nucleotit mã hóa polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(i) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(j) trình tự nucleotit mã hóa polypeptit có trình tự axit amin như nêu trong SEQ ID NO.7;

(k) trình tự nucleotit mã hóa polypeptit có trình tự axit amin được tạo ra bằng cách loại bỏ, thay thế hoặc bổ sung một hoặc nhiều gốc axit amin trên trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(l) trình tự nucleotit mã hóa polypeptit có trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

(m) Trình tự nucleotit có độ tương đồng nằm trong khoảng từ 80% đến 100% so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8, cũng như mã hóa polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

(n) Trình tự nucleotit của polynucleotit mà lai hóa với polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8 trong điều kiện nghiêm ngặt, cũng như trình tự nucleotit mã hóa polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

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

Từ các trình tự nucleotit từ (a) đến (l) nêu trên, 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 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. Tương tự, các trình tự nucleotit (d), (g), và (j) nêu trên cũng có thể lần lượt là các trình tự nucleotit như nêu trong SEQ ID NO.4, SEQ ID NO.6, và SEQ ID NO.8, hoặc các trình tự nucleotit, trong đó các bộ

ba mã hóa thoái biến ở các trình tự nucleoit này được cải biến thành các bộ ba mã hóa có tần xuất sử dụng cao ở vật chủ. Lưu ý rằng, các bộ ba mã hóa này có thể được cải biến bằng kỹ thuật tái tổ hợp gen đã biết.

Polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8 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 AR19G-166-RA, AR19G-166-QV có chiều dài đầy đủ và các gen tương tự (cũng được gọi là “gen AR19G-166”) hoặc một đoạn gen của nó chứa vùng xúc tác xenlobiohydrolaza từ tự nhiên bằng kỹ thuật tái tổ hợp gen. Gen AR19G-166 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 môi xuôi và đoạn môi ngược được thiết kế dựa trên trình tự nucleoit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8. 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 (m) nêu trên so với trình tự nucleoit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8 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. Cụ thể, ở trạng thái trong đó khoảng trống được tạo ra ở các vị trí xuất hiện đột biến chèn thêm hoặc loại bỏ, hai trình tự nucleotit được đặt cạnh nhau sao cho các trình tự nucleotit tương ứng là giống nhau; và tỷ lệ các trình tự nucleotit giống nhau so với tổng số trình tự nucleotit trừ đi khoảng trống trên trình tự sắp thẳng hàng được xác định là độ tương đồng trình tự. Độ tương đồng trình tự giữa các trình tự nucleotit có thể được xác định bằng nhiều phần mềm tra cứu độ tương đồng đã biết trong lĩnh vực này. Theo sáng chế, độ tương đồng của các trình tự nucleotit được xác định dựa trên trình tự sắp thẳng hàng thu được bằng phần mềm tra cứu độ tương đồng BLASTN.

Ví dụ, polynucleotit chứa trình tự nucleotit (b), (c), (e), (f), (h), (i), (k), (l), hoặc (m) nêu trên có thể lần lượt được tổng hợp nhân tạo bằng cách loại bỏ, thay thế, hoặc bổ sung một hoặc nhiều nucleotit trong polynucleotit có trình tự như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8. Hơn nữa, trình tự nucleotit (b), (c), (e), (f), (h), (i), (k), hoặc (l) cũng có thể là trình tự có chiều dài đầy đủ của gen tương đồng với gen AR19G-166 hoặc một đoạn trình tự của nó. Gen tương đồng của gen AR19G-166 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.

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.

Khi polynucleotit có trình tự nucleotit (b), (c), (e), (f), (h), (i), (k), (l) hoặc (m) nêu trên được tổng hợp nhân tạo, thì không có giới hạn cụ thể về số lượng nucleotit cần được loại bỏ, thay thế hoặc bổ sung trên trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6 hoặc SEQ ID NO.8, miễn là trình tự nucleotit của polynucleotit sau khi tổng hợp có độ tương đồng nằm trong khoảng từ 80% đến 100% so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6 hoặc SEQ ID NO.8, nhưng tốt hơn nếu số lượng nucleotit nằm trong khoảng từ 1 đến 256, tốt hơn nữa nếu nằm trong khoảng từ 1 đến 192, còn tốt hơn nữa nếu nằm trong khoảng từ 1 đến 128, và tốt nhất nếu nằm trong khoảng từ 1 đến 64.

Polynucleotit theo sáng chế có thể chỉ chứa miền mã hóa vùng xúc tác xenlobiohydrolaza, hoặc ngoài miền mã hóa vùng xúc tác xenlobiohydrolaza, thì polynucleotit theo sáng chế cũng có thể chứa miền mã hóa đoạn gắn kết xenluloza, gốc liên kết, các peptit 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 xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5 ở tế bào vật chủ. Cụ thể, vector này là vector biểu hiện chứa polynucleotit mã hóa xenlobiohydrolaza chịu nhiệt theo sáng chế.

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 lai, và ADN có trình tự vùng kết thúc.

Đặc biệt, vector biểu hiện cần được kết hợp dưới dạng 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, hoặc bằng kit tổng hợp vector biểu hiện có bán trên thị trường.

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

Ví dụ về vector biểu hiện theo sáng chế có thể được chèn vào tế bào thực vật là các vector nhị phân như pIG121 hoặc pIG121Hm. Ví dụ về vùng khởi đầu có thể sử dụng bao gồm vùng khởi đầu mã hóa nopalin synthaza, vùng khởi đầu 35S của virus gây bệnh khảm trên cây súp lơ (*cauliflower mosaic virus*), hoặc vùng khởi đầu tương tự. Hơn nữa, vùng kết thúc có thể sử dụng bao gồm vùng khởi kết thúc mã hóa nopalin synthaza, hoặc vùng kết thúc tương tự. Ngoài ra, vùng khởi đầu đặc hiệu mô hoặc vùng khởi đầu đặc hiệu cơ quan cũng có thể được sử dụng. Khi vùng khởi đầu đặc hiệu mô hoặc vùng khởi đầu đặc hiệu cơ quan được sử dụng, xenlobiohydrolaza chịu nhiệt có thể không được biểu hiện ở toàn bộ thực vật nhưng được biểu hiện chỉ ở các cơ quan hoặc mô đặc hiệu. Do đó, xenlobiohydrolaza chịu nhiệt được mong muốn chỉ biểu hiện ở bộ phận không ăn được của thực vật ăn được.

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 thực vật đã chuyển gen và thực vật 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.

Thể biến nạp theo sáng chế

Thể biến nạp nêu trên chứa vector biểu hiện theo sáng chế. Xenlobiohydrolaza chịu nhiệt theo sáng chế có thể được biểu hiện ở thể biến nạp này. Xenlobiohydrolaza đã biết trong lĩnh vực này chỉ được biểu hiện ở một số vật chủ, nói cách khác khó có thể biểu hiện được nhiều xenlobiohydrolaza ở các loài khác nhau. Trái lại, xenlobiohydrolaza chịu nhiệt theo sáng chế có thể được biểu hiện ở nhiều vật chủ như *E. coli*, xạ khuẩn, nấm men, nấm sợi, hoặc lục lạp của thực vật bậc cao.

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

Vật chủ để mang vector biểu hiện này có thể là tế bào nhân sơ như *E. coli* hoặc xạ khuẩn hoặc tế bào nhân thật như nấm men, nấm sợi, tế bào côn trùng, tế bào động vật có vú, và tế bào thực vật. Xenlobiohydrolaza chịu nhiệt theo sáng chế có thể được sản xuất một cách đơn giản với hiệu suất cao bằng cách nuôi cấy thể biến nạp *E. coli* hoặc xạ khuẩn này. Mặt khác, do protein bị glycosyl hóa ở tế bào nhân thật, nên xenlobiohydrolaza chịu nhiệt có độ ổn định nhiệt cao hơn có thể được sản xuất bằng cách sử dụng thể biến nạp là tế bào nhân thật chứ không sử dụng thể biến nạp là tế bào nhân sơ. Cụ thể, khi thể biến nạp là nấm sợi như *Aspergillus* hoặc vi sinh vật nhân thật (hoặc sinh vật nhân thật) như nấm men, xenlobiohydrolaza chịu nhiệt có độ ổn định nhiệt cao hơn có thể được sản xuất một cách đơn giản với hiệu suất cao. Hơn nữa, thực vật chuyển gen được chèn vector biểu hiện theo sáng chế cho phép canh tác trên quy mô lớn bằng các phương thức canh tác ngoài trời hoặc các phương thức canh tác tương tự cũng như thu được lượng xenlobiohydrolaza chịu nhiệt theo sáng chế tương đối lớn tính trên một thực vật. Hơn nữa, do thực vật chuyển gen thường chứa xenlobiohydrolaza chịu nhiệt bên trong thân thực vật, nên thực vật này thích

hợp để dùng làm nguồn sinh khối.

Tốt hơn nếu tế bào vật chủ dùng để biểu hiện xenlobiohydrolaza chịu nhiệt theo sáng chế là ít nhất một tế bào vật chủ được chọn từ nhóm bao gồm *E. coli*, nấm men, nấm sợi, xạ khuẩn và thực vật, tốt hơn nữa nếu ít nhất một tế bào vật chủ được chọn từ nhóm bao gồm nấm men, nấm sợi, xạ khuẩn và thực vật, và còn tốt hơn nữa nếu ít nhất một tế bào vật chủ là thực vật.

Khi tế bào nhân sơ, nấm men, nấm sợi, tế bào côn trùng, hoặc tế bào động vật có vú 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.

Khi thể biến nạp theo sáng chế là thực vật, thì tế bào thực vật, cơ quan thực vật hoặc mô thực vật có thể được sử dụng làm vật chủ. Thực vật chuyển gen có thể được tạo ra từ tế bào thực vật chuyển gen, mô sẹo, và các cơ quan tương tự bằng phương pháp nuôi cấy mô thực vật đã biết hoặc các phương pháp tương tự. Ví dụ, thực vật chuyển gen có thể được tạo ra bằng cách nuôi cấy tế bào thực vật chuyển gen trên môi trường tái biệt hóa không chứa hormon hoặc môi trường tương tự, sau đó trồng cây con thu được đã ra rễ vào đất hoặc môi trường tương tự và sinh trưởng.

Khi thể biến nạp theo sáng chế là thực vật, thì cat-xet biểu hiện xenlobiohydrolaza chịu nhiệt theo sáng chế thu được từ vectơ nêu trên có thể được kết hợp vào bộ gen của thực vật, mặc dù tốt hơn nếu cat-xet này được kết hợp vào bộ gen lục lạp. Trong thể biến nạp lục lạp, gen ngoại sinh được chèn vào được di truyền trong tế bào chất. Do đó, sự biểu hiện gen tái tổ hợp này với môi trường thông qua hạt phấn có thể được ngăn ngừa. Trong quá trình canh tác ngoài trời thực vật chuyển gen trên quy mô lớn, sự biểu hiện gen tái tổ hợp này với môi trường là vấn đề cần quan tâm. Tuy nhiên, thể biến nạp lục lạp có ưu điểm hơn thể biến nạp bộ gen dưới góc độ ngăn ngừa sự biểu hiện gen tái tổ hợp với môi trường.

Khi thể biến nạp theo sáng chế là thực vật, ngoài thực vật thu được trực tiếp từ quá trình chuyển gen, thì thể biến nạp này cũng bao gồm thực vật là thế hệ sau của thực vật chuyển gen này và biểu hiện xenlobiohydrolaza chịu nhiệt theo sáng chế theo cùng cách thức như thực vật chuyển gen này. Thuật ngữ “thực vật thế hệ sau” được dùng trong bản mô tả để chỉ thực vật thu được từ quá trình nảy mầm hạt giống của thực vật chuyển gen, thực vật thu được từ đoạn cành của nó, hoặc các bộ phận tương tự.

Không có giới hạn cụ thể về thực vật được sử dụng làm vật chủ, và thực vật này có thể là thực vật hai lá mầm, thực vật một lá mầm, dương xỉ, rêu, tảo hoặc vi tảo. Ví dụ, thực vật thuộc họ Cải (*Brassicaceae*), họ Hòa thảo (*Poaceae*), họ Cà (*Solanaceae*), họ Đậu (*Fabaceae*), họ Cúc (*Asteraceae*), họ Bìm bìm (*Convolvulaceae*), họ Thầu dầu (*Euphorbiaceae*), và các họ thực vật tương tự có thể được sử dụng. Thực vật thuộc họ Cà (*Solanaceae*), họ Cải (*Brassicaceae*), hoặc họ Hòa thảo (*Poaceae*) được ưu tiên do chúng thích hợp để chuyển gen nhờ *Agrobacterium*. Ví dụ về thực vật thuộc họ Cà (*Solanaceae*) bao gồm cây thuốc lá, cà tím, khoai tây, cà chua, ớt chuông, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Cải (*Brassicaceae*) bao gồm mù tạc (*Arabidopsis thaliana*), cây rau cải, cây rau tề, củ cải, bắp cải, cải ngựa Nhật Bản, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Hòa thảo (*Poaceae*) bao gồm lúa, ngô, lúa miến, lúa mì, lúa mạch, lúa mạch đen, kê, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Đậu (*Fabaceae*) bao gồm lạc, đậu xanh, đậu tương, đậu đen, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Cúc (*Asteraceae*) bao gồm cây ngưu bàng, cây ngải, cúc vạn thọ, hoa thanh cúc, hướng dương, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Bìm bìm (*Convolvulaceae*) bao gồm bìm bìm hoa hồng, cây hoa chuông, thực vật thuộc chi *Cuscuta*, bìm bìm hoa trắng, và các thực vật tương tự. Ví dụ về thực vật thuộc họ Thầu dầu (*Euphorbiaceae*) bao gồm đại kích, *E. sieboldiana*, *E. lasiocaula*, và các thực vật tương tự.

Khi thể biến nạp theo sáng chế là thực vật, thì tốt hơn nếu thể biến nạp này là thực vật một lá mầm, tốt hơn nữa nếu thể biến nạp này là thực vật thuộc họ Hòa thảo (*Poaceae*), và tốt nhất nếu thể biến nạp này là thực vật thuộc họ Hòa thảo (*Poaceae*) có lượng sinh khối lớn.

Phương pháp sản xuất xenlobiohydrolaza chịu nhiệt theo sáng chế

Phương pháp sản xuất xenlobiohydrolaza chịu nhiệt theo sáng chế là phương pháp để sản xuất xenlobiohydrolaza chịu nhiệt trong thể biến nạp theo phương án 4. Trong thể biến nạp được sản xuất bằng cách sử dụng vectơ biểu hiện chứa polynucleotit theo sáng chế nằm trước vùng khởi đầu mà không có khả năng điều hòa thời gian biểu hiện hoặc các yếu tố tương tự, thì xenlobiohydrolaza chịu nhiệt theo sáng chế được biểu hiện liên tục. Trái lại, 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, xenlobiohydrolaza chịu nhiệt được biểu hiện trong thể biến nạp này bằng cách

thực hiện cảm ứng phù hợp với điều kiện cảm ứng biểu hiện tương ứng

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

Không có giới hạn cụ thể về phương pháp chiết xuất/tinh chế xenlobiohydrolaza chịu nhiệt từ thể biến nạp miễn là phương pháp này không làm giảm hoạt tính của xenlobiohydrolaza 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 xenlobiohydrolaza 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. Xenlobiohydrolaza 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 xenlobiohydrolaza chịu nhiệt theo sáng chế được biểu hiện đồng thời, peptit tín hiệu bài tiết được giữ trong thể biến nạp, thì dung dịch chứa xenlobiohydrolaza chịu nhiệt có thể được thu nhận dễ dàng bằng cách nuôi cấy thể biến nạp, sau đó thu gom phân dịch nổi nuôi cấy bằng cách loại bỏ thể biến nạp ra khỏi sản phẩm nuôi cấy thu được. Hơn nữa, khi xenlobiohydrolaza chịu nhiệt theo sáng chế có trình tự đánh dấu như trình tự đánh dấu His, thì xenlobiohydrolaza 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 xenlobiohydrolaza chịu nhiệt theo sáng chế bao gồm bước biểu hiện xenlobiohydrolaza chịu nhiệt trong thể biến nạp theo phương án 4, và sau đó là bước chiết xuất và tinh chế xenlobiohydrolaza chịu nhiệt từ thể biến nạp.

Hỗn hợp xenlulaza theo sáng chế

Hỗn hợp xenlulaza theo phương án 6 chứa xenlobiohydrolaza chịu nhiệt theo

phương án 1, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 5, và ít nhất một hoặc nhiều loại xenlulaza khác. Xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 5 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 xenlulaza theo sáng chế được sử dụng để phân giải xenluloza, thì có thể phân giải hiệu quả lignoxenluloza.

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

Tốt hơn nếu xenlulaza khác chứa trong hỗn hợp xenlulaza theo sáng chế là xenlulaza bền nhiệt có hoạt tính ít nhất ở điều kiện nhiệt độ bằng 70°C, và tốt hơn nữa nếu enzym này là xenlulaza bền nhiệt có hoạt tính ở nhiệt độ ít nhất nằm trong khoảng từ 70°C đến 90°C. Khi hỗn hợp xenlulaza theo sáng chế chứa toàn bộ các enzym chịu nhiệt, thì có thể phân giải hiệu quả xenluloza bằng hỗn hợp xenlulaza này ở điều kiện nhiệt độ cao. Cụ thể, khi hỗn hợp xenlulaza theo sáng chế chỉ chứa xenlulaza 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. 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 xenluloza theo sáng chế

Phương pháp sản xuất sản phẩm phân giải xenluloza theo sáng chế là phương pháp để thu được sản phẩm bằng cách phân giải xenluloza bằng xenlobiohydrolaza chịu nhiệt theo sáng chế. Cụ thể, sản phẩm phân giải xenluloza được sản xuất bằng cách cho nguyên liệu chứa xenluloza tiếp xúc với xenlobiohydrolaza chịu nhiệt theo phương án 1, thể biến nạp theo phương án 4, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp

theo phương án 5.

Thuật ngữ “sản phẩm phân giải xenluloza” được dùng trong bản mô tả để chỉ xenlobioza.

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

Tốt hơn nếu phản ứng thủy phân xenluloza bởi xenlobiohydrolaza chịu nhiệt theo sáng chế được thực hiện ở điều kiện mà ở đó enzym này có hoạt tính xenlobiohydrolaza. Tốt hơn nếu phản ứng thủy phân xenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 55°C đến 80°C và độ pH nằm trong khoảng từ 3,5 đến 7,0, và tốt hơn nữa nếu phản ứng thủy phân xenluloza được thực hiện ở nhiệt độ nằm trong khoảng từ 70°C đến 100°C và độ pH nằm trong khoảng từ 4,0 đến 6,0. Thời gian phản ứng thủy phân xenluloza được điều chỉnh thích hợp theo loại nguyên liệu chứa xenluloza được sử dụng, phương pháp xử lý sơ bộ nguyên liệu, lượng nguyên liệu, và các yếu tố tương tự. Ví dụ, phản ứng thủy phân xenluloza có thể được thực hiện trong thời gian từ 10 phút đến 100 giờ. Khi phân giải sinh khối xenluloza, thì phản ứng thủy phân xenluloza có thể được thực hiện trong thời gian từ 1 giờ đến 100 giờ.

Để thủy phân xenluloza, thì ngoài xenlobiohydrolaza chịu nhiệt theo sáng chế tốt hơn nếu ít nhất một xenlulaza khác được sử dụng. Các xenlulaza khác chứa trong hỗn hợp xenlulaza nêu trên có thể được sử dụng, và tốt hơn nếu xenlulaza này là bền nhiệt và có hoạt tính ít nhất ở điều kiện nhiệt độ bằng 70°C và tốt hơn nếu ở nhiệt độ ít nhất nằm trong khoảng từ 70°C đến 100°C. Ngoài ra, trong phương pháp này, hỗn hợp xenlulaza theo phương án 6 có thể được sử dụng thay cho xenlobiohydrolaza chịu nhiệt theo phương án 1, thể biến nạp theo phương án 4, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo phương án 5.

Phương pháp sản xuất polynucleotit và đoạn mồi để sử dụng theo sáng chế

Phương pháp sản xuất polynucleotit mã hóa xenlobiohydrolaza chịu nhiệt theo sáng chế bao gồm bước thực hiện phản ứng PCR với khuôn là ADN thu được từ vi sinh vật hoặc sản phẩm phiên mã ngược của ARN thu được từ vi sinh vật, với đoạn mỗi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.12 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12 (đoạn mỗi theo phương án 9 của sáng chế), và đoạn mỗi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13 (đoạn mỗi theo phương án 10 của sáng chế), và thu nhận sản phẩm khuếch đại là polynucleotit có trình tự nucleotit mã hóa xenlobiohydrolaza chịu nhiệt.

Trình tự nucleotit như nêu trong SEQ ID NO.12 là trình tự tương đồng với đoạn trình tự chứa các nucleotit ở vị trí số từ 1 đến 22 của trình tự nucleotit như nêu trong SEQ ID NO.2. Hơn nữa, trình tự nucleotit như nêu trong SEQ ID NO.13 là trình tự bổ sung với đoạn trình tự chứa các nucleotit ở vị trí số từ 1263 đến 1284 của trình tự nucleotit như nêu trong SEQ ID NO.2. Do đó, polynucleotit mã hóa vùng xúc tác xenlobiohydrolaza của gen AR19G-166 (ví dụ, polynucleotit có trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8) có thể thu được dưới dạng sản phẩm khuếch đại bằng cách thực hiện phản ứng PCR với đoạn mỗi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.12 và đoạn mỗi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13 và khuôn là polynucleotit chứa trình tự nucleotit như nêu trong SEQ ID NO.2.

Trong phản ứng PCR, đầu 5' của đoạn mỗi cũng có thể chứa trình tự nucleotit hỗ trợ mà không tham gia vào quá trình lai hóa với khuôn. Ví dụ, bằng cách sử dụng đoạn mỗi xuôi chứa trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12 và đoạn mỗi ngược chứa trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13, polynucleotit trong đó một hoặc nhiều nucleotit thu được từ đoạn mỗi xuôi được bổ sung vào đầu 5' của miền mã hóa vùng xúc tác xenlobiohydrolaza của gen AR19G-166 và một hoặc nhiều nucleotit thu được từ đoạn mỗi ngược được bổ sung vào đầu 3' của nó có thể thu được. Ví dụ về trình tự nucleotit được bổ sung vào đầu 5' của mỗi đoạn mỗi bao gồm trình tự cần có để chèn sản phẩm khuếch đại vào vector biểu hiện, vị trí enzyme cắt giới hạn, trình tự nucleotit mã hóa trình tự đánh dấu,

trình tự nucleotit mã hóa peptit tín hiệu, hoặc trình tự tương tự. Ngoài ra, cũng tốt hơn nếu metionin hoạt hóa (ATG) được bổ sung vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12.

ADN được sử dụng làm khuôn trong phản ứng PCR là ADN thu được từ vi sinh vật hoặc sản phẩm phiên mã ngược (ADN bổ sung) của ARN thu được từ vi sinh vật. Sinh vật có thể là vi sinh vật được biến nạp plasmit, trong đó polynucleotit mã hóa vùng xúc tác xenlobiohydrolaza của gen AR19G-166 đã được kết hợp, thể biến nạp nêu trên, hoặc sinh vật chứa trong mẫu thu được từ tự nhiên. Khi ADN được sử dụng làm khuôn được tổng hợp từ mẫu thu được từ tự nhiên, tốt hơn nếu mẫu thu được môi trường có nhiệt độ cao như đất ở khu vực suối nước nóng.

Trong phương pháp sản xuất polynucleotit theo sáng chế, điều kiện phản ứng PCR và các điều kiện tương tự có thể được xác định thích hợp dựa trên loại polymeraza cần sử dụng và các thành phần tương tự bởi người có hiểu biết trung bình trong lĩnh vực này. Khi ADN bộ gen của gen AR19G-166 hoặc ADN bổ sung được tổng hợp từ ARN thông tin của gen chứa trong axit nucleic được sử dụng làm khuôn, polynucleotit mã hóa vùng xúc tác xenlobiohydrolaza của gen AR19G-166 có thể thu được dưới dạng sản phẩm khuếch đại từ phản ứng PCR.

Trình tự axit amin của vùng xúc tác xenlobiohydrolaza có tính bảo thủ cao giữa các gen tương đồng. Do đó, khi ADN bộ gen của gen tương đồng với gen AR19G-166 hoặc ADN bổ sung được tổng hợp từ ARN thông tin của gen tương đồng chứa trong axit nucleic được sử dụng làm khuôn, polynucleotit mã hóa vùng xúc tác xenlobiohydrolaza của gen tương đồng của gen AR19G-166 có thể thu được dưới dạng sản phẩm khuếch đại bằng cách sử dụng đoạn mồi xuôi và đoạn mồi ngược nêu trên và phương pháp sản xuất polynucleotit theo sáng chế.

Ngoài ra, khi ADN bộ gen của gen mà không phải là gen tương đồng với gen AR19G-166 nhưng có trình tự nucleotit tương tự với trình tự nucleotit của gen AR19G-166, hoặc ADN bổ sung được tổng hợp từ ARN thông tin của gen tương đồng chứa trong axit nucleic được sử dụng làm khuôn, thì polynucleotit mã hóa toàn bộ vùng xúc tác xenlobiohydrolaza hoặc một phần vùng xúc tác xenlobiohydrolaza của gen này có thể thu được dưới dạng sản phẩm khuếch đại bằng phương pháp sản xuất polynucleotit theo sáng

chế. Do đó, phương pháp sản xuất polynucleotit theo sáng chế cũng hữu ích để tách dòng xenlobiohydrolaza mới có trình tự axit amin tương tự với trình tự axit amin của gen AR19G-166.

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

Tách dòng xenlobiohydrolaza chịu nhiệt theo sáng chế từ đất ở khu vực suối nước nóng

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

Để giải trình tự gen mã hóa xenlobiohydrolaza chịu nhiệt (nhiệt độ tối ưu: 55°C hoặc cao hơn) và xenlobiohydrolaza chịu nhiệt (nhiệt độ tối ưu: 80°C hoặc cao hơn), 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 sống 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). Mỗi mẫu bộ gen (5µg) tạo ra lượng ADN bằng 10µg hoặc cao hơn được giải trình tự. Tức là, ADN chiết tách được được giải trình tự và ADN ribosom 16S được khuếch đại bằng máy giải trình tự GS FLX Titanium 454 do Roche Diagnostics sản xuất. Lượng ADN còn lại được sử dụng để tách dòng PCR gen xenlulaza. Mặt khác, các mẫu có lượng ADN thu được bằng 10µg hoặc nhỏ hơn được khuếch đại bằng kit khuếch đại ADN bộ gen (Kit khuếch đại ADN GenomiPhi V2, do GE Healthcare sản xuất) và sản phẩm khuếch đại thu được được giải trình tự AND bộ gen.

Quá trình giải trình tự ADN hệ gen môi trường này được hiện 3 đến 4 lần đối với mỗi mẫu đất ở khu vực suối nước nóng, tức là tổng cộng 19 lần. Kết quả là thu được trình tự bộ gen có chiều dài đoạn đọc trung bình bằng 394bp, tổng số các đoạn trình tự đọc bằng 26.295.463, và tổng chiều dài bộ gen đã được giải trình tự bằng 10,3Gbp.

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.

Thư viện trình tự được thiết kế từ ADN bộ gen được chiết tách từ mẫu đất thu được ở các khu vực suối nước nóng bằng máy giải trình tự Roche 454 GS FLX Titanium. Dữ liệu đầu ra (tệp tin sff) của máy giải trình tự Roche 454 được xử lý đệm bằng phần mềm PyroBayes (Quinlan *et al.*, *Nature Methods*, 2008, vol. 5, p. 179-81.) để thu được các tệp trình tự định dạng FASTA và tệp có chất lượng. Các đầu của trình tự đọc thu được được cắt bỏ để cải thiện chất lượng của chúng, và các trình tự đọc được lắp ráp bằng phần mềm Newbler phiên bản 2.3 hoặc 2.5.3 của máy giải trình tự 454 Life Sciences. Quá trình lắp ráp được thực hiện trong các điều kiện: “độ trùng lặp tối thiểu chấp nhận được (mi) = 0,9”, “tùy chọn: - lớn (đối với các bộ gen có kích cỡ lớn hoặc phức tạp, tăng tốc độ lắp ráp, nhưng làm giảm độ chính xác)”.

Tổng chiều dài của các đoạn trình tự đọc đã được xử lý tăng chất lượng và các đoạn tiếp giáp đã được lắp ráp có chiều dài bằng 100bp hoặc dài hơn bằng 2,5Gbp. Bộ dữ liệu này được sử dụng để phân tích gen mã hóa xenlulaza. Trong tổng số 26.294.193 đoạn trình tự đọc, 17.991.567 đoạn trình tự đọc được lắp ráp thành 595.602 đoạn tiếp giáp có chiều dài trung bình bằng 1kb hoặc dài hơn. Trong số các đoạn tiếp giáp này, đoạn tiếp giáp lớn nhất có chiều dài bằng 278.185bp.

Đối với các trình tự đã được lắp ráp, toàn bộ các đoạn tiếp giáp và các đoạn duy nhất được phân loại thành năm loài bao gồm vi khuẩn, vi khuẩn cổ, sinh vật nhân thật, virus, và nhóm không thuộc các loài này, tham khảo cơ sở dữ liệu KEGG (Kanehisa, *M. Science & Technology Japan*, 1996, No. 59, p. 34-38, <http://www.genome.jp/kegg/>, 11/5/2011). Trong số các trình tự đã được lắp ráp có chiều dài bằng 2,5Gbp (= tổng chiều dài các đoạn tiếp giáp + tổng chiều dài đoạn duy nhất), thì chiều dài trình tự của vi khuẩn bằng 258Mbp, chiều dài trình tự của vi khuẩn cổ bằng 27Mbp, chiều dài trình tự của sinh vật nhân thật bằng 193.561bp (0,008% tổng chiều dài của các trình tự đã lắp ráp), và chiều dài trình tự

của virus bằng 685.640bp (0,027% tổng chiều dài của các trình tự đã lắp ráp). Nguyên nhân không thu được nhiều các trình tự của sinh vật nhân thật là do nhiệt độ của hệ gen môi trường thu được từ đất ở khu vực suối nước nóng nằm trong khoảng từ 58°C đến 70°C, cao hơn giới hạn nhiệt độ sinh trưởng của sinh vật nhân thật như nấm sợi. Các kết quả này cho thấy cơ sở dữ liệu hệ gen môi trường chỉ chứa 11,3% trình tự ADN đã biết. Chiều dài trình tự của nhóm không thuộc các loài nêu trên bằng 2,2Gbp và chiếm 88,7% tổng chiều dài của các trình tự đã được lắp ráp. Các trình tự này là trình tự thể hệ mới thu được từ vi khuẩn, vi khuẩn cổ, hoặc sinh vật nhân thật bất kỳ. Kết quả này giúp khẳng định giả thuyết của Handelsman *et al.*, *Chem Biol.*, 1998, vol, 5, p.R245-R249 là hầu hết các ADN bộ gen tạo ra quần thể vi sinh vật ở một số môi trường cụ thể đã chưa được xác định rõ bằng các phương pháp nghiên cứu bộ gen vi sinh vật thông thường, tức là phương pháp nuôi cấy và phân lập vi sinh vật, sau đó giải trình tự ADN bộ gen bằng phương pháp Sanger.

3. Quá trình xác định khung đọc mã mở (Open Reading Frame - ORF) mã hóa xenlobiohydrolaza

Các trình tự có mã phân loại enzym 3.2.1.4 (xenluloza), 3.2.1.37 (β -xylosidaza), 3.2.1.91 (xenluloza 1,4- β -xenlobiosidaza), và 3.2.1.8 (endo 1,4- β -xylanaza) được tải về từ cơ sở dữ liệu UniProt (<http://www.uniprot.org/>, số truy cập: 13/04/2009), và cơ sở dữ liệu về các gen mã hóa glycosit hydrolaza được thiết kế. Các khung đọc mã mở được phát hiện từ các trình tự tiếp giáp thu được ở phần 2 bằng phần mềm Orphelia (Hoff *et al.*, *Nucleic Acids Research*, 2009, 37 (Web Server issue: W101-W105, tùy chọn Orphelia: mặc định (kiểu = Net 700, độ trùng lặp tối đa = 60), tùy chọn hệ gen môi trường: -m)) đối với các hệ gen môi trường AR15 và AR19 và phần mềm Metagene (Noguchi *et al.*, *DNA Research*, 2008, 15(6)) đối với các hệ gen môi trường H1, N2, và OJ1. Gen mã hóa glycosit hydrolaza được xác định từ trình tự ORF phát hiện được bằng phần mềm BLASTP (phiên bản 2.2.18) và tham khảo cơ sở dữ liệu cục bộ. 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 mục tiêu được thu thập là gen mã hóa glycosit hydrolaza.

Phần mềm Orphelia và Metagene không xử lý được sự dịch chuyển khung gây ra bởi các trình tự đọc lỗi hoặc các trình tự tương tự. Do đó, sự dịch chuyển khung được hiệu

chỉnh bằng phương pháp sau. Trước tiên, các đoạn tiếp giáp được cắt thành các đoạn có chiều dài 2kbp bằng cách xáo trộn các đoạn 1kbp. Do đó, các trình tự đã cắt được làm trùng lặp với các trình tự nằm trước và nằm sau đoạn có chiều dài 1kbp. Các trình tự tiếp giáp đã được cắt tương ứng được đối chiếu với cơ sở dữ liệu về các gen glycosit hydrolaza nêu trên ($E < 1e^{-20}$) để sàng lọc bằng phần mềm Blastx. Đối với các trình tự tiếp giáp mục tiêu, vùng mã hóa glycosit hydrolaza được lấy từ phần mềm Genewise (Wise2 package: <http://www.ebi.ac.uk/Tools/Wise2/>). Đồng thời, vùng mã hóa các trình tự có chiều dài 100bp hoặc ngắn hơn được loại bỏ. Đối với phần mềm Genewise, sự dịch chuyển khung trình tự được hiệu chỉnh bằng cách đối chiếu với các trình tự enzym trùng khớp với trình tự tiếp giáp mục tiêu trên cơ sở dữ liệu, và chèn hoặc loại bỏ các khoảng trống để đạt được độ sắp thẳng hàng tối đa.

Các glycosit hydrolaza, như xenlulaza, endohemixenlulaza, và các enzym cắt mạch nhánh thu được bằng phương pháp nêu trên được phân loại chức năng bằng phần mềm pfam HMMs (Pfam phiên bản 23.0 và HMMER v 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, các trình tự 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 sử dụng mô hình ẩn Markov (Durbin *et al.*, “*The theory behind profile HMMs. Biological sequence analysis: probabilistic models of protein and nucleic acids*”, 1998, Cambridge University Press.; hmmpfam (phiên bản 2.3.2), trị số kỳ vọng $E < 1e^{-5}$; cơ sở dữ liệu = Pfam_fs (mô hình có thể được sử dụng để phát hiện các đoạn của các vùng cụ thể ở trình tự)). Phân tích sàng lọc BLASTP được thực hiện và 44 khung đọc mã mở mục tiêu dưới dạng trình tự xenlobiohydrolaza được phân loại thành các họ GH.

4. Kiểm tra bộ ba mã hóa khởi đầu hiếm gặp thu được từ tín hiệu đầu ra của phần mềm Orphelia.

Phần mềm Orphelia không chỉ phát hiện được các khung đọc mã mở chứa ATG (metionin), mà còn phát hiện được các khung đọc mã mở có các bộ ba hiếm gặp GTG (valin), TTG (leucin), và ATA (isoleucin) là bộ ba mã hóa khởi đầu. Do đó, khi các trình tự tiếp giáp đã được lắp ráp không chứa khung đọc mã mở có chiều dài đầy đủ có ATG là bộ ba mã hóa khởi đầu, thì có thể xuất hiện lỗi, trong đó phần mềm Orphelia nhận dạng bộ ba hiếm gặp này là bộ ba mã hóa khởi đầu. Trong phần 3, trong số các khung đọc mã mở đầu

ra được xác định là có chiều dài đầy đủ từ phần mềm Orphelia, có 8 khung đọc mã mở có các bộ ba hiếm gặp GTG, TTG, và ATA là bộ ba mã hóa khởi đầu. Liên quan đến trình tự axit amin của các khung đọc mã mở đầu ra thu được từ phần mềm Genewise và các trình tự tiếp giáp chứa khung đọc mã mở, các khung đọc mã mở này được kiểm tra liệu đây là các trình tự có chiều dài đầy đủ có bộ ba hiếm gặp là bộ ba mã hóa khởi đầu hay là các trình tự lỗi đầu ra. Kết quả kiểm tra bằng phần mềm Orphelia cho thấy 8 khung đọc mã mở có các bộ ba hiếm gặp là bộ ba mã hóa khởi đầu này đều là các trình tự lỗi đầu ra, nói cách khác các trình tự này là các trình tự có chiều dài không đầy đủ.

Bảng 1

Hệ gen môi trường	Phân loại họ GH của các gen xylanaza					
	GH6	GH7	GH9	GH48	GH khác	Tổng cộng
AR19	2(0)	0	2(2)	5(1)	4(3)	13(6)
AR15	0	0	1(1)	2(1)	3(2)	6(4)
OJ1	2(0)	0	7(2)	2(1)	4(2)	15(5)
N2	0	0	5(3)	3(0)	2(2)	10(5)
H1	0	0	0	0	0	0
Tổng số khung đọc mã mở	4(0)	0(0)	15(8)	12(3)	13(9)	44(20)

Kết quả phân loại 44 khung đọc mã mở được xác định là gen mã hóa xenlobiohydrolaza thành các họ GH được thể hiện trong Bảng 1. Trong Bảng 1, chữ số trong các dấu ngoặc đơn thể hiện số lượng khung đọc mã mở có chiều dài đầy đủ có metionin là bộ ba mã hóa khởi đầu. Như được thể hiện trong Bảng 1, 2 khung đọc mã mở mã hóa xenlobiohydrolaza (AR19G-166 và AR19G-12) thuộc họ GH6 thu được từ hệ gen môi trường AR19, và 2 khung đọc mã mở mã hóa xenlobiohydrolaza (OJ1-1 và OJ1-2) thuộc cùng một họ thu được từ hệ gen môi trường OJ1, tức là tổng cộng 4 khung đọc mã mở. Trái lại, không thu được trình tự khung đọc mã mở thuộc họ GH7. 15 khung đọc mã mở thuộc họ GH9, và 12 khung đọc mã mở thuộc họ GH48 thu được. Tổng cộng thu được 13 khung đọc mã mở mã hóa xenlobiohydrolaza thuộc các họ GH khác (GH10, GH12, và GH26). Các đoạn môi được thiết kế cho toàn bộ các khung đọc mã mở này đã được xác định là gen mã hóa xenlobiohydrolaza bao gồm các trình tự có chiều dài không đầy đủ, và các gen này được tách dòng từ ADN hệ gen môi trường thu được từ đất ở khu vực suối nước nóng bằng phương pháp PCR.

Lưu ý rằng, các dung dịch xenlulaza dùng cho nhiên liệu sinh học có bán trên thị trường bao gồm Novozyme CELLIC (nhãn hiệu đã đăng ký) CTec2 (<http://www.bioenergy.novozymes.com/cellulosic-ethanol/>), và Genencor Accellerase

(nhãn hiệu đã đăng ký) TRIO (http://www.genencor.com/industries/biofuels/fuel_etanol_from_biomass_cellulosic_biofuels/), các sản phẩm này đều chứa các enzym được bài tiết bởi nấm phân hủy gỗ *Trichoderma reesei*. Các thành phần chính của glycosit hydrolaza (GH) được bài tiết từ nấm sợi này bao gồm xenlobiohydrolaza CBHI và xenlobiohydrolaza CBHII, lần lượt thuộc họ GH7 và họ GH6.

5. Khung đọc mã mở OJ1-1 và OJ1-2

Khung đọc mã mở OJ1-1 gồm 548 axit amin, và mã hóa enzym đa vùng có đoạn gắn kết xenluloza CBM3 (149bp)-gốc liên kết (111bp)-vùng xúc tác GH6. Tuy nhiên, nửa sau của vùng xúc tác này không chứa bộ ba mã hóa kết thúc và có chiều dài không đầy đủ. Hơn nữa, trình tự đoạn gắn kết xenluloza này là trình tự CBM3 mới có độ tương đồng trình tự axit amin bằng 63% so với đoạn gắn kết xenluloza CBM3 (SEQ ID NO.16) của xenlobiohydrolaza của vi khuẩn hiếu khí ưa nhiệt *Caldibacillus cellulosovorans* (Mã truy cập Ngân hàng gen: AAF22273.1). Vùng xúc tác xenlobiohydrolaza của OJ1-1 có độ tương đồng trình tự axit amin bằng 58% so với xenluloza 1,4- β -xenlobiosidaza của *Amycolatopsis mediterranei* U32 (Mã truy cập Ngân hàng gen: ADJ46954,1), và có độ tương đồng trình tự axit amin bằng 48% so với β -1,4-xenlobiohydrolaza của xạ khuẩn ưa nhiệt *Thermobifida fusca* YX, là enzym có hoạt tính xenlulaza mạnh (Mã truy cập Ngân hàng gen: AAA62211.1).

Một sản phẩm tách dòng gen (OJ1-1-11) thu được từ OJ1-1 bằng phương pháp tách dòng PCR. OJ1-1-11 mã hóa polypeptit chứa 548 axit amin, trong đó 32 axit amin (M1-A32) từ metionin (M) (vị trí số 1) của bộ ba mã hóa khởi đầu đến alanin ở vị trí số 32 là peptit tín hiệu bài tiết (peptit tín hiệu P 4.0), 148 axit amin (T35-P183) từ threonin ở vị trí số 35 đến prolin ở vị trí số 182 là đoạn gắn kết xenluloza CBM3, 112 axit amin (S183-T294) từ serin (S) ở vị trí số 183 đến threonin ở vị trí số 294 là gốc liên kết, và 254 axit amin từ histidin (H) ở vị trí số 295 đến kết thúc là một đoạn trình tự axit amin của vùng xúc tác xenlobiohydrolaza thuộc họ GH6. Tuy nhiên, trong phần 10 dưới đây, toàn bộ chiều dài của sản phẩm tách dòng gen này được biểu hiện bởi *E. coli* hoạt tính phân giải PSA và CMC được đánh giá, thì sản phẩm tách dòng gen này không có hoạt tính thủy phân cơ chất bất kỳ.

Khung đọc mã mở OJ1-2 gồm 247 axit amin, và mã hóa polypeptit chỉ chứa vùng xúc tác GH6. Do không có cả bộ ba mã hóa khởi đầu và bộ ba mã hóa kết thúc, và xenlobiohydrolaza thuộc họ GH6 thường chứa 400 axit amin hoặc cao hơn, nên OJ1-2 là trình tự không có chiều dài đầy đủ. Trình tự axit amin mã hóa bởi OJ1-2 là trình tự có độ tương đồng bằng 100% so với trình tự axit amin của AR19G-12, do đó gen OJ1-2 là giống hệt với AR19G-12, là một đoạn trình tự của AR19G-12. Vùng xúc tác của sản phẩm tách dòng gen thu được từ OJ1-2 bằng phương pháp tách dòng PCR được chèn vào vector chuyển gen, và biểu hiện bởi *E. coli*, không có hoạt tính thủy phân cơ chất PSA và CMC.

6. Khung đọc mã mở AR19G-166 và AR19G-12

Khung đọc mã mở AR19G-166 mã hóa polypeptit chứa of 474 axit amin (SEQ ID NO.9), nhưng không có chiều dài đầy đủ do không chứa bộ ba mã hóa khởi đầu, và chỉ chứa một đoạn trình tự gồm gốc liên kết và vùng xúc tác GH6. Vùng xúc tác GH6 của AR19G-166 có độ tương đồng trình tự axit amin bằng 66% so với glycosit hydrolaza có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt độ trung bình thuộc ngành *Chloroflexi*, *Herpetosiphon aurantiacus* DSM 785 (Mã truy cập Ngân hàng gen: ABX04776.1). Hai sản phẩm tách dòng gen (AR19G-166-RA và AR19G-166-QV) thu được từ AR19G-166 bằng phương pháp tách dòng PCR với đoạn môi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.14 (5'-CACCATGTTGGACAATCCATTCATCGGAG-3': 7 nucleotit (CACCATG) được bổ sung vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12. Trên trình tự đã bổ sung, ba nucleotit ATG ở đầu 3' là bộ ba mã hóa khởi đầu, và 4 nucleotit CACC ở đầu 5' là trình tự để chèn vào vector), và đoạn môi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13 (5'-TTAGGGTTGGATCGGCGGATAG-3'). AR19G-166-RA và AR19G-166-QV chỉ khác nhau ở hai axit amin ở vị trí số 299 và vị trí số 351. Trong AR19G-166-RA, axit amin ở vị trí số 299 là arginin và axit amin ở vị trí số 351 là alanin (SEQ ID NO.1). Trong AR19G-166-QV, axit amin ở vị trí số 299 là glutamin và axit amin ở vị trí số 351 là valin (SEQ ID NO.3).

Khung đọc mã mở AR19G-12 mã hóa polypeptit chứa 459 axit amin (SEQ ID NO.10), nhưng không có chiều dài đầy đủ do không chứa bộ ba mã hóa khởi đầu, và chỉ chứa một đoạn trình tự gồm gốc liên kết và vùng xúc tác GH6, tương tự như AR19G-166. Vùng xúc tác GH6 của AR19G-12 có độ tương đồng trình tự axit amin bằng 64% so với glycosit hydrolaza GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785 (Mã truy

cập Ngân hàng gen: ABX04776.1). Tuy nhiên, vùng xúc tác của sản phẩm tách dòng gen thu được từ AR19G-12 bằng phương pháp PCR được chèn vào vectơ chuyển gen, và biểu hiện bởi *E. coli*, không có hoạt tính thủy phân cơ chất PSA và CMC.

Phân tích phát sinh loài

Không giống như các gen được tách dòng từ cơ thể vi khuẩn được phân lập và nuôi cấy, nguồn gốc của các gen được tách dòng từ phân tích hệ gen môi trường là đã biết. Người ta không biết liệu 4 khung đọc mã mở (AR19G-166, AR19G-12 (OJ1-2), và OJ1-1) thuộc họ GH6 thu được từ hệ gen môi trường thu được từ đất có nhiệt độ cao có nguồn gốc từ sinh vật nhân sơ như vi khuẩn hoặc vi khuẩn cổ (vi khuẩn cổ đại), hay có nguồn gốc từ sinh vật nhân thật như nấm sợi hoặc nấm. Do đó, phân tích phát sinh loài được thực hiện bằng phương pháp sắp thẳng hàng nhiều trình tự axit amin của các vùng xúc tác và cây phát sinh loài để xác định nguồn gốc của các khung đọc mã mở này.

Fig.1 là hình vẽ thể hiện cây phát sinh loài của exo-glycosit hydrolaza GH6 (xenlobiohydrolaza, glycosit hydrolaza, exoglucanaza, và xenlobiosidaza). Các trình tự axit amin mã hóa bởi các khung đọc mã mở (AR19G-166, AR19G-12 (OJ1-2), và OJ1-1), và trình tự axit amin của vùng xúc tác của xạ khuẩn ưa nhiệt *Thermobifida fusca* YX Cel6B (Mã truy cập Ngân hàng gen: AAA62211,1), có hoạt tính phân giải xenluloza, được tra cứu độ tương đồng bằng phần mềm BLASTP từ Ngân hàng gen. Kết quả là thu được 21 loại trình tự exo-glycosit hydrolaza GH6. Tiếp theo, trình tự axit amin của vùng xúc tác của nấm sợi ưa nhiệt *Humicola insolens* Cel6A (PDB: 1VBW) cũng được tra cứu độ tương đồng bằng phần mềm BLASTP từ Ngân hàng gen. Kết quả là thu được 19 loại trình tự. 21 trình tự thu được từ vi khuẩn và 19 trình tự có nguồn gốc từ nấm sợi này, và các trình tự axit amin mã hóa bởi các khung đọc mã mở mục tiêu (AR19G-166, AR19G-12 (OJ1-2), và OJ1-1) bằng phương pháp tra cứu độ tương đồng được sắp thẳng hàng bằng phần mềm Geneious Pro 5.6.5 (Cost Matrix = Blosum80; Điểm tạo khoảng trống = 12; Điểm mở rộng khoảng trống = 3; loại sắp thẳng hàng = sắp thẳng hàng toàn bộ không có các khoảng trống ở đầu tận cùng). Tiếp theo, cây phát sinh loài được tạo ra bằng phương pháp liên kết loài liên kề. Endoglucanaza thuộc họ GH6 của *Thermobifida fusca* YX Cel6A (Mã truy cập Ngân hàng gen: AAC06388,1) được cài đặt là nhóm bên ngoài. Xác suất tự phát sinh được tính toán dựa trên 1.000 bản sao, và được thể hiện bằng tỷ lệ phần trăm ở mỗi nhánh của cây phát sinh loài. Trên Fig.1, đường gạch ngang ở phía dưới thể hiện khoảng cách di truyền (số

lượng vị trí/thay thế axit amin trung bình). Hơn nữa, danh pháp của các enzym này được thể hiện trong dấu ngoặc đơn, và ký hiệu “CBH” có nghĩa là xenlobiohydrolaza, và ký hiệu “GH” có nghĩa là glycosit hydrolaza.

Glycosit hydrolaza GH6 của vi khuẩn và nấm sợi được sử dụng làm mối liên hệ di truyền của cây phát sinh loài là như sau (chữ số trong các dấu ngoặc đơn thể hiện mã truy cập ngân hàng gen, Protein Data bank (PDB), hoặc EMBL-Bank). Glycosit hydrolaza GH6 của nấm sợi bao gồm: *Acremonium cellulolyticus* Y-94 xenlobiohydrolaza II (Mã truy cập Ngân hàng gen: BAA74458.1); *Agaricus bisporus* xenlobiohydrolaza (Mã truy cập Ngân hàng gen: AAA50608.1); *Aspergillus kawachii* IFO 4308 1,4-beta-D-glucan xenlobiohydrolaza C (Mã truy cập Ngân hàng gen: GAA89571.1); *Aspergillus niger* ATCC 1015 1,4-beta-D-glucan xenlobiohydrolaza C (Mã truy cập Ngân hàng gen: EHA25828.1); *Chaetomium thermophilum* xenlobiohydrolaza họ 6 (Mã truy cập Ngân hàng gen: AAW64927.1); *Colletotrichum higginsianum* glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: CCF33252.1); *Fomitiporia mediterranea* MF3/22 xenlulaza CEL6B (Mã truy cập Ngân hàng gen: EJD02201.1); *Glomerella graminicola* M1,001 glucosyl hydrolaza họ 6 (Mã truy cập Ngân hàng gen: EFQ25807.1); *Humicola insolens* Cel6A (PDB:1BVW); *Leptosphaeria maculans* JN3 xenlobiohydrolaza II(EMBL-Bank:CBX97039.1); *Magnaporthe grisea* 70-15 exoglucanaza 2 (Mã truy cập Ngân hàng gen: EHA57773.1); *Myceliophthora thermophila* ATCC 42464 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: AEO55787.1); *Neurospora crassa* OR74A exoglucanaza 2 (Mã truy cập Ngân hàng gen: EAA31534.1); *Penicillium decumbens* xenlobiohydrolaza II (Mã truy cập Ngân hàng gen: ADX86895.1); *Punctularia strigosozonata* HHB-11173 SS5 xenlobiohydrolaza II (Mã truy cập Ngân hàng gen: EIN07098.1); *Talaromyces emersonii* xenlobiohydrolaza II (Mã truy cập Ngân hàng gen: AAL33604,4); *Thielavia terrestris* NRRL 8126 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: AEO62210.1); *Trichoderma reesei* xenlobiohydrolaza II (Mã truy cập Ngân hàng gen: AAA34210.1); và *Verticillium dahliae* VdLs,17 exoglucanaza-6A (Mã truy cập Ngân hàng gen: EGY16046.1).

Glycosit hydrolaza GH6 của vi khuẩn bao gồm: *Acidothermus cellulolyticus* 11B glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: ABK52388.1); *Amycolatopsis mediterranei* U32 1,4-beta-xenlobiosidaza (Mã truy cập Ngân hàng gen: ADJ46954.1);

Cellulomonas fimi ATCC 484 1,4-beta-xenlobiohydrolaza (Mã truy cập Ngân hàng gen: AEE46055.1); *Cellvibrio japonicus* Ueda 107 xenlobiohydrolaza cel6A (Mã truy cập Ngân hàng gen: ACE85978.1); *Herpetosiphon aurantiacus* DSM 785 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: ABX04776.1); *Jonesia denitrificans* DSM 20603 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: ACV08399.1); *Ktedonobacter racemifer* DSM 44963 1,4-beta-xenlobiohydrolaza (Mã truy cập Ngân hàng gen: EFH85864.1); *Micromonospora lupini* str. Lupac 08 1,4-beta-xenlobiohydrolaza (Mã truy cập Ngân hàng gen: CCH20969.1); *Paenibacillus curdianolyticus* YK9 1,4-beta-xenlobiohydrolaza (Mã truy cập Ngân hàng gen: EFM08880.1); *Ralstonia solanacearum* Po82 xenlobiohydrolaza A (Mã truy cập Ngân hàng gen: AEG71050.1); *Salinispora arenicola* CNS-205 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: ABV99773.1); *Stackebrandtia nassauensis* DSM 44728 1,4-beta-xenlobiosidaza (Mã truy cập Ngân hàng gen: ADD42622.1); *Stigmatella aurantiaca* DW4/3-1 exoglucanaza A (Mã truy cập Ngân hàng gen: EAU67050.1); *Streptomyces avermitilis* MA-4680 1,4-beta-xenlobiosidaza (Mã truy cập Ngân hàng gen: BAC69564.1); *Teredinibacter turnerae* T7901 xenlobiohydrolaza (Mã truy cập Ngân hàng gen: ACR12723.1); *Thermobifida fusca* YX xenlobiohydrolaza Cel6B (Mã truy cập Ngân hàng gen: AAA62211.1); *Verrucosispora maris* AB-18-032 1,4-beta-xenlobiohydrolaza (Mã truy cập Ngân hàng gen: AEB46944.1); *Xanthomonas campestris* pv. *raphani* 756C exoglucanaza A (Mã truy cập Ngân hàng gen: AEL08359.1); *Xanthomonas oryzae* pv. *oryzae* KACC 10331 1,4-beta-xenlobiosidaza (Mã truy cập Ngân hàng gen: AAW77289.1); *Xylanimonas cellulolytica* DSM 15894 glucosit hydrolaza họ 6 (Mã truy cập Ngân hàng gen: ACZ30181.1); và *Xylella fastidiosa* Ann-1 xenlobiohydrolaza A (Mã truy cập Ngân hàng gen: EGO81204.1).

Exo-glycosit hydrolaza GH6 được phân loại thành hai nhánh có khoảng cách di truyền lớn bao gồm nhánh có nguồn gốc vi khuẩn và nhánh có nguồn gốc nấm sợi. Trên Fig.1, toàn bộ các nhánh bên dưới là các glycosit hydrolaza GH6 có nguồn gốc từ nấm sợi, trong khi đó toàn bộ các nhánh bên trên là các glycosit hydrolaza GH6 có nguồn gốc từ vi khuẩn. Toàn các khung đọc mã mở AR19G-166, AR19G-12, và OJ1-1 được phân loại vào nhóm vi khuẩn, và tạo thành một nhánh với glycosit hydrolaza GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785, xenlobiohydrolaza cel6A có nguồn gốc từ vi khuẩn gram (-) và phân giải gỗ *Cellvibrio japonicus*, và xenlobiohydrolaza có nguồn gốc từ vi khuẩn tự dưỡng γ *Teredinibacter turnerae*.

Như được thể hiện trong Fig.1, đã phát hiện thấy rằng 4 khung đọc mã mở (AR19G-166, AR19G-12 (OJ1-2), và OJ1-1) thuộc họ GH6 thu được từ phân tích hệ gen môi trường có khoảng cách di truyền xa với gen xenlobiohydrolaza có nguồn gốc từ nấm sợi ưa nhiệt *Acremonium cellulolyticus*, *Chaetomium thermophilum*, và nấm phân giải gỗ *Hypocerea jecorina* (*Trichoderma reesei*), và 4 khung đọc mã mở này có mối liên hệ di truyền gần với xenlobiohydrolaza có nguồn gốc từ xạ khuẩn ưa nhiệt *Thermobifida fusca* YX, và glycosit hydrolaza GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785. Do đó, 4 khung đọc mã mở thuộc họ GH6 này được xác định là các gen xenlobiohydrolaza có nguồn gốc từ vi khuẩn. OJ1-1 có đoạn gắn kết xenluloza CBM3 đặc trưng của vi khuẩn, góp phần khẳng định chính xác các kết quả này.

7. Phân tích sắp thẳng hàng trình tự axit amin

Glycosit hydrolaza GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785 có độ tương đồng trình tự cao so với AR19G-166-RA và AR19G-166-QV gồm 1.128 axit amin. Gen này là gen đa vùng bao gồm, bắt đầu với trình tự petit tín hiệu vận chuyển chứa 29 axit amin từ vị trí số 1 đến vị trí số 29, đoạn gắn kết xenluloza CBM2 chứa 100 axit amin từ vị trí số 37 đến vị trí số 136, vùng xúc tác GH6 chứa 370 axit amin từ vị trí số 241 đến vị trí số 611, và một vùng xúc tác GH6 khác chứa 370 axit amin từ vị trí số 713 đến vị trí số 1082. Cả hai vùng xúc tác GH6 này chứa 370 gốc axit amin như được phân tích bằng phần mềm Pfam, là ngắn hơn các vùng xúc tác GH6 có nguồn gốc vi khuẩn, ví dụ ngắn hơn vùng xúc tác có nguồn gốc từ *Thermobifida fusca* YX Cel6B chứa 423 gốc axit amin khoảng 50 gốc axit amin hoặc cao hơn. Do đó, gen này được xem là không chứa vùng xúc tác GH6 thực. Do đó, vùng xúc tác GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785 tương đồng với *Thermobifida fusca* YX Cel6B được tra cứu bằng phần mềm BLASTP, và kết quả được thể hiện trên Fig.2A, trong đó vùng xúc tác GH6 thứ nhất chứa 428 gốc axit amin valin (V) ở vị trí số 230 đến glutamin (Q) ở vị trí số 657.

Ngoài ra, mặc dù có chiều dài không đầy đủ, nhưng vùng xúc tác GH6 của OJ1-1 có độ tương đồng axit amin bằng 57% so với vùng xúc tác GH6 của AR19G-166, và có độ tương đồng axit amin chỉ bằng 25% so với TrCBHII có nguồn gốc từ nấm sợi *Trichoderma reesei*.

Trong khung đọc mã mở AR19G-166, 47 gốc axit amin nằm phía trước trình tự

vùng xúc tác GH6. Do có nhiều đoạn lặp prolin (P) và threonin (T), nên trình tự axit amin này được cho là một đoạn của gốc liên kết. Theo đó, tương tự như OJ1-1, AR19G-166 là gen đa vùng có đoạn gắn kết xenluloza CBM nằm trước vùng xúc tác xenlobiohydrolaza thông gốc liên kết.

Fig.2A là hình vẽ thể hiện sự sắp thẳng hàng các trình tự axit amin mã hóa bởi các khung đọc mã mở AR19G-166, AR19G-12, và OJ1-1, và glycosit hydrolaza GH6 có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785. Hơn nữa, Fig.2B là hình vẽ thể hiện sự sắp thẳng hàng trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1, và trình tự axit amin của đoạn gắn kết xenluloza CBM3 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt *Caldibacillus cellulovorans* (Mã truy cập Ngân hàng gen: AAF22273.1). Trên Fig.2A và Fig.2B, các axit amin nằm sau có màu đen/trắng thể hiện các vùng, trong đó gốc axit amin bảo thủ trong toàn bộ các trình tự axit amin này, và các axit amin được đánh bóng thể hiện các vùng, trong đó gốc axit amin bảo thủ ở hầu hết các trình tự axit amin này mặc dù có một số đột biến ở một số đoạn của các trình tự axit amin này.

Fig.3A là hình vẽ thể hiện các trình tự axit amin mã hóa bởi các khung đọc mã mở AR19G-166, AR19G-12, và OJ1-1 và gen CBH có nguồn gốc từ *Herpetosiphon aurantiacus* DSM 785. Hơn nữa, Fig.3B là hình vẽ thể hiện trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1 và trình tự axit amin của đoạn gắn kết xenluloza CBM3 có nguồn gốc từ vi khuẩn hiếu khí ưa nhiệt *Caldibacillus cellulovorans*. Trên Fig.3A và Fig.3B, ký hiệu “vùng (đoạn) xúc tác” và “gốc liên kết (đoạn liên kết)” lần lượt thể hiện chỉ một đoạn của mỗi vùng.

8. Quá trình tách dòng gen.

Gen xenlobiohydrolaza tiềm năng thu được bằng phương pháp tách dòng PCR được khuếch đại bằng phương pháp PCR bằng cách sử dụng khuôn là ADN thu được từ đất ở khu vực suối nước nóng đã khuếch đại bằng kit khuếch đại ADN bộ gen (kit khuếch đại ADN GenomiPhi V2, do GE Healthcare sản xuất). Các sản phẩm PCR đã được khuếch đại được chèn vào vectơ pET101/D-TOPO bằng kit biểu hiện Champion pET Directional TOPO do Invitrogen 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.

9. Quá trình biểu hiện và tinh chế protein enzym xenlobiohydrolaza.

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 TM pET Directional TOPO (do Invitrogen sản xuất) hoặc chủng Rosetta-gamiB (DE3) pLysS (do Merck sản xuất) được sử dụng làm tế bào chủ để chuyển gen. *E. coli* chứa gen đích được cấy vào môi trường LB chứa ampicilin với hàm lượng bằng 100mg/L và nuôi cấy cho đến khi mật độ quang ở bước sóng 600nm nằm trong khoảng từ 0,2 đến 0,8, sau đó IPTG (isopropyl- β -D(-)-thiogalactopyranosit) được bổ sung vào, và nuôi cấy tiếp trong 5 đến 20 giờ. Kết quả là quá trình biểu hiện protein đích được cảm ứng. Sau khi nuôi cấy, *E. coli* được thu nhận bằng cách ly tâm, và dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) có thể tích bằng 1/10 thể tích dịch nuôi cấy được bổ sung vào để tạo hỗn dịch vi khuẩn. Sau đó, quy trình phân cắt trong 30 giây và rung lắc trong 30 giây được lặp lại 10 lần bằng thiết bị BioRuptor UCD-200T (do Cosmo Bio Co. Ltd. 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 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 Tris-HCl 50mM (độ pH = 8,0) chứa 1M NaCl bằng hệ thống sắc ký lỏng AKTA (do GE Healthcare sản xuất). Các phân đoạn có hoạt tính xenlobiohydrolaza được trộn toàn bộ, sau đó trao đổi với dung dịch Tris-HCl 50mM (độ pH = 8,0) chứa 750mM amoni sulfat bằng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất). Các phân đoạn có hoạt tính xenlobiohydrolaza 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 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 Tris-HCl 50mM (độ pH = 8,0). Các phân đoạn có hoạt tính xenlobiohydrolaza 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 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 xenlobiohydrolaza được trộn toàn bộ, sau đó trao đổi với dung dịch Tris-HCl 50mM (độ pH = 8,0) và cô. Kết quả là thu được enzym tinh khiết có nồng độ cuối cùng bằng khoảng 1mg/mL.

Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và protein enzym xenlobiohydrolaza tinh khiết được kiểm chứng bằng phân tích điện di trên gel SDS-PAGE và thẩm tách Western.

Quá trình điện di trên gel SDS-PAGE của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và enzym tinh khiết này lần lượt được thực hiện bằng gel kích cỡ nhỏ với gradien nồng độ nằm trong khoảng từ 4% đến 20% và gel kích cỡ nhỏ 10% (do ATTO sản xuất). Phần dịch nổi nuôi cấy và enzym tinh khiết này được trộn với dung dịch Tris-SDS β ME (do Cosmo Bio Co., Ltd sản xuất.) ở tỷ lệ 1:1, sau đó gia nhiệt ở nhiệt độ 100°C trong 10 phút. dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (5 μ L) và enzym tinh khiết (0,5 μ L) lần lượt được chạy điện di. Sau khi chạy điện di, gel cố định được nhuộm với thuốc nhuộm màu xanh CBB R250 (do Merck sản xuất) để hiển thị các băng protein.

Liên quan đến phân tích thẩm tách Western của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và enzym tinh khiết, quá trình điện di trên gel SDS-PAGE được thực hiện bằng gel kích cỡ nhỏ 10% (do ATTO sản xuất), sau đó các protein được chuyển lên màng polyvinyliden florua (do ATTO sản xuất) bằng thiết bị Trans-Blot SD (do BioRad sản xuất). Các protein trên màng được cho phản ứng với kháng thể sơ cấp của thỏ đã được pha loãng 100 lần. Kháng thể sơ cấp của thỏ được sản xuất bằng cách tổng hợp polypeptit chứa 20 gốc axit amin từ vị trí số 384 đến 403 mã hóa bởi AR19G-166-QV (CDPNGQSRYSAYPTGALPN), và tinh chế huyết thanh từ thỏ đã được miễn dịch bằng phương pháp ái lực (do Operon Biotechnologies cung cấp). Sự gắn kết của kháng thể sơ cấp với protein được phát hiện bằng kit thẩm tách Western (do Pierce sản xuất), và tín hiệu

phát quang hóa học được phát hiện bằng thiết bị Ez-Capture MG (do ATTO sản xuất).

Fig.4A là hình vẽ thể hiện kết quả phân tích điện di trên gel SDS-PAGE và Fig.4B là hình vẽ thể hiện kết quả phân tích thấm tách Western của các protein enzym thu được bằng cách biểu hiện AR19G-166-RA và AR19G-166-QV trong *E. coli*. 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 AR19G-166-RA và AR19G-166-QV; làn 4 và làn 5 thể hiện mẫu điện di của protein enzym tinh khiết AR19G-166-RA và AR19G-166-QV.

Nhìn chung, gen xenlobiohydrolaza được biểu hiện rất yếu. Ví dụ, khi được biểu hiện bởi vật chủ là *E. coli*, thì các gen xenlobiohydrolaza được biểu hiện rất yếu cho dù gen này có nguồn gốc từ nấm sợi hay vi khuẩn. Tuy nhiên, cả hai sản phẩm tách dòng PCR AR19G-166-RA và AR19G-166-QV đều được biểu hiện mạnh trong *E. coli*. Kết quả phân tích điện di trên gel SDS-PAGE cho thấy trong dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen AR19G-166-RA và AR19G-166-QV có các băng đậm có trọng lượng phân tử bằng 46,7kDa (làn 2 và làn 3 trên Fig.4A) được xác định từ các trình tự axit amin (SEQ ID NO.1 và SEQ ID NO.3). Sau khi tinh chế các protein này, các băng đơn tương ứng với các băng nêu trên được phát hiện ở cả AR19G-166-RA lẫn AR19G-166-QV (làn 4 và làn 5 trên Fig.4A). Kết quả phân tích thấm tách Western bằng kháng thể của polypeptit chứa 20 gốc axit amin từ vị trí số 384 đến 403 của AR19G-166 cho thấy có các băng protein enzym duy nhất có trọng lượng phân tử bằng 46,7kDa được phát hiện ở cả dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (làn 2 và làn 3 trên Fig.4B) lẫn enzym tinh khiết (làn 4 và làn 5 trên Fig.4B).

10. Đo hoạt tính xenlobiohydrolaza bằng cơ chất PSA (hoạt tính thủy phân PSA)

Cơ chất Avicel trương nở liên hợp axit phosphoric (phosphoric acid-swollen Avicel - PSA) được sử dụng để đo hoạt tính xenlobiohydrolaza. PSA được điều chế bằng cách hòa tan bột Avicel (bột xenluloza vi tinh thể, do Merck sản xuất) vào dung dịch axit phosphoric; nước cất vô khuẩn được bổ sung vào để kết tủa các tinh thể, và các tinh thể này được rửa cho đến khi độ pH của chúng bằng 5,0 hoặc cao hơn. Toàn bộ PSA được sử dụng trong thử nghiệm dưới đây được điều chế theo phương pháp này.

Hoạt tính của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và mẫu enzym ở giai đoạn giữa của quá trình tinh chế được đo bằng cách cho hỗn hợp chứa 50μL dung dịch đậm

axit axetic 200mM chứa PSA với hàm lượng bằng 1% khối lượng (độ pH = 5,5) với 50 μ L dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen hoặc mẫu enzym ở giai đoạn giữa của quá trình tinh chế phản ứng ở nhiệt độ 30 đến 100°C trong 20 phút.

Trong toàn bộ các phép đo, hỗn hợp được điều chế bằng cách bổ sung dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) thay cho dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và phản ứng ở cùng điều kiện được sử dụng làm mẫu đối chứng. Hơn nữa, dung dịch cơ chất và enzym được giữ lần lượt và riêng biệt ở nhiệt độ phản ứng được duy trì trong 5 phút, sau đó trộn. Khoảng thời gian này được thiết lập để kích hoạt phản ứng. Trong quá trình phản ứng, các hỗn hợp này được khuấy bằng máy khuấy nhiệt Eppendorf (1.400 vòng/phút) để tránh kết tủa cơ chất không tan. Sau khi kết thúc phản ứng, thuốc thử axit 3,5-đinitrosalixylic (dung dịch DNS) được bổ sung vào mỗi hỗn hợp ở thể tích bằng nhau. Các hỗn hợp thu được được gia nhiệt ở nhiệt độ 100°C trong 5 phút, làm nguội trong 5 phút, sau đó ly tâm. Kết quả là thu được phần dịch nổi. Độ hấp thụ ở bước sóng 540nm được đo bằng máy quang phổ, và hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn glucoza. Lượng đường khử tạo ra bởi quá trình thủy phân được đo bằng cách so sánh độ sai lệch với mẫu đối chứng. Hoạt tính enzym để tạo ra 1 μ mol đường khử trong 1 phút được tính bằng 1U, và trị số thu được bằng cách chia hoạt tính này cho khối lượng protein được tính bằng hoạt tính riêng (U/mg).

Kết quả đo cho thấy trong số 44 xenlobiohydrolaza tiềm năng thu được bằng phương pháp tách dòng PCR, thì chỉ polypeptit mã hóa bởi khung đọc mã mở AR19G-166 của gen CBH thuộc họ GH6 là có hoạt tính thủy phân PSA. Cả hai sản phẩm tách dòng PCR thu được từ AR19G-166 (AR19G-166-RA và AR19G-166-QV) đều có hoạt tính thủy phân PSA.

11. Độ đặc hiệu cơ chất của xenlobiohydrolaza

Hoạt tính thủy phân các cơ chất xenluloza và hemixenluloza khác nhau của protein enzym AR19G-166-RA và AR19G-166-QV, mà đã được kiểm chứng là có hoạt tính thủy phân PSA, được đo. Trong phép đo này, các enzym tinh khiết (nồng độ cuối cùng bằng khoảng 1 mg/mL) thu được từ phần 9 được sử dụng.

Độ đặc hiệu cơ chất của xenlobiohydrolaza AR19G-166 được đo bằng cách sử dụng các cơ chất là PSA, bột Avicel, CMC (carboxymetyl xenluloza, do Sigma sản xuất), xylan

(có nguồn gốc từ gỗ sồi, do Sigma sản xuất), lichenan (do MP Biomedicals sản xuất), laminarin (có nguồn gốc từ *Laminaria digitata*, do Sigma sản xuất). Cụ thể, phép đo này được thực hiện bằng cách ủ sơ bộ hỗn hợp chứa 50 μ L dung dịch đệm axit phosphoric 200mM (độ pH = 5,5), 40 μ L nước tinh khiết và 10 μ L dung dịch enzyme tinh khiết ở nhiệt độ 50°C trong 5 phút, tiếp theo bổ sung thêm 100 μ L dung dịch chứa mỗi cơ chất với hàm lượng bằng 1% vào và cho hỗn hợp phản ứng ở nhiệt độ 50°C trong 20 phút (2 giờ khi cơ chất là bột Avicel). Lượng đường khử được tạo ra bằng quá trình thủy phân enzyme được đo, và hoạt tính riêng (U/mg) được đo theo cùng cách thức như nêu trong phần 10. Mỗi phép đo được lặp lại 3 lần để thu được trị số trung bình và độ lệch chuẩn. Hơn nữa, hoạt tính tương đối (%) của hoạt tính riêng đối với mỗi cơ chất đã được tính toán với giả định rằng hoạt tính riêng đối với PSA bằng 100%. Các kết quả đo được thể hiện trong Bảng 2.

Bảng 2. Độ đặc hiệu cơ chất của AR19G-166-RA và AR19G-166-QV

Cơ chất và nồng độ	AR19G-166-RA		AR19G-166-QV	
	Hoạt tính riêng (U/mg, trung bình $\pm$ độ lệch chuẩn)	Hoạt tính tương đối (%)	Hoạt tính riêng (U/mg, trung bình $\pm$ độ lệch chuẩn)	Hoạt tính tương đối (%)
0,5% PSA	1,81 $\pm$ 0,10	100%	1,39 $\pm$ 0,07	100%
0,5% Avicel	0,08 $\pm$ 0,05	4,4%	0,10 $\pm$ 0,03	7,2%
0,5% CMC	0,00 $\pm$ 0,01	0,0%	0,00 $\pm$ 0,00	0,0%
0,5% Laminarin	0,02 $\pm$ 0,01	1,1%	- 0,02 $\pm$ 0,01	-1,4%
0,5% Lichenan	0,27 $\pm$ 0,02	14,9%	0,26 $\pm$ 0,02	18,7%
0,5% Xylan	0,01 $\pm$ 0,02	0,6%	0,03 $\pm$ 0,02	2,1%

Kết quả đo cho thấy AR19G-166-RA và AR19G-166-QV có hoạt tính thủy phân mạnh PSA tan trong nước. Ngoài ra, AR19G-166-RA và AR19G-166-QV cũng có hoạt tính phân giải lichenan chứa liên kết glucan β -1,3 và β -1,4 và xenluloza tinh thể Avicel. Trái lại, AR19G-166-RA và AR19G-166-QV hầu như không có hoạt tính phân giải CMC, laminarin chứa liên kết glucan β -1,3 và β -1,6, và xylan. Kết quả phân tích độ đặc hiệu cơ chất cho thấy AR19G-166-RA và AR19G-166-QV có hoạt tính thủy phân xenluloza tinh thể Avicel mặc dù hoạt tính này yếu, cũng như không có hoạt tính phân giải xylan, chứng tỏ rằng AR19G-166-RA và AR19G-166-QV là xenlobiohydrolaza GH6.

12. Phân tích sản phẩm phân giải PSA bằng phương pháp HPLC.

Sản phẩm thủy phân cơ chất PSA với xenlobiohydrolaza AR19G-166-RA và TrCBHII có nguồn gốc từ nấm sợi phân giải gỗ *T. reesei* được phân tích bằng phương pháp sắc ký hiệu năng cao (HPLC). Phản ứng thủy phân PSA lần lượt được thực hiện với

AR19G-166-RA trong dung dịch đệm axit axetic 0,1M (độ pH = 5,5) ở nhiệt độ 70°C, và với TrCBHII trong dung dịch đệm axit axetic 0,1M (độ pH = 4,0) ở nhiệt độ 40°C trong 1 giờ và 24 giờ, sau đó dừng phản ứng bằng dung dịch natri carbonat 0,1M. Mỗi sản phẩm của phản ứng thủy phân được ly tâm ở nhiệt độ 4°C và tốc độ bằng 12.000 vòng/phút trong 10 phút, và phần dịch nổi được lọc qua màng lọc có kích cỡ lỗ lọc bằng 0,2 μ m, và phân tích bằng phương pháp HPLC. Thiết bị HPLC Alliance e2695 (do Waters sản xuất), và bộ phận phát hiện chỉ số khúc xạ (2414RI) được sử dụng để đo hàm lượng đường. Thiết bị HPLC này được kiểm soát và phân tích bằng phần mềm Empower phiên bản 3.0. Cột phân tích hydratcarbon HPLC có kích cỡ 300mm x 7,8mm (do Bio-Rad sản xuất), và dung môi là nước siêu tinh khiết được sử dụng. Mẫu thủy phân (10 μ L) được phân tích ở tốc độ dòng bằng 0,6mL/phút và nhiệt độ cột bằng 85°C. Đường chuẩn được tạo ra bằng các chất chuẩn (glucoza, xenlobioza, và xenlotrioza) được sử dụng để đo hàm lượng đường. Nồng độ tối đa của các mẫu này đo được bằng đường chuẩn là 0,2% khối lượng đối với glucoza, 0,4% khối lượng đối với xenlobioza, và 0,5% khối lượng đối với xenlotrioza. Sau khi điều chế các dung dịch pha loãng liên tiếp, đường chuẩn được tạo ra bằng 6 điểm nồng độ.

Kết quả phân tích HPLC được thể hiện trên Fig.5A và Fig.5B. Sản phẩm thủy phân PSA bằng AR19G-166-RA chứa chủ yếu xenlobioza (86,5%) với lượng nhỏ xenlotrioza (13,5%) sau 1 giờ, và 86,2% xenlobioza và 13,0% xenlotrioza, với lượng rất nhỏ glucoza (0,7%) sau 24 giờ (Fig.5A). Mặt khác, sản phẩm thủy phân PSA bằng TrCBHII chứa chủ yếu xenlobioza (87,5%) với lượng nhỏ xenlotrioza (12,5%), và 88,4% xenlobioza và 9,4% xenlotrioza với lượng rất nhỏ glucoza (2,2%) sau 24 giờ (Fig.5B). Các kết quả phân tích HPLC này chứng tỏ rằng AR19G-166 là xenlobiohydrolaza.

13. Đánh giá mức độ phụ thuộc vào nhiệt độ và độ pH của hoạt tính xenlobiohydrolaza.

Mức độ phụ thuộc vào nhiệt độ và độ pH của hoạt tính thủy phân PSA của AR19G-166-RA và AR19G-166-QV được đánh giá. Trong phép đo này, các enzym tinh khiết (nồng độ cuối cùng bằng khoảng 1mg/mL) thu được từ phần 9 được sử dụng.

Hoạt tính thủy phân PSA của enzym tinh khiết được đo theo cùng cách thức như nêu trong phần 10, chỉ khác là cho hỗn hợp chứa 100 μ L dung dịch nước PSA 1% khối lượng, 50 μ L dung dịch đệm McIlvaine (độ pH nằm trong khoảng từ 3 đến 8), 40 μ L nước tinh

khuyết, và 10 μ L enzym tinh khiết, phản ứng ở nhiệt độ 30°C, 40°C, 50°C, 60°C, 65°C, 70°C, 75°C, 80°C, 85°C, 90°C, hoặc 100°C trong 20 phút, trong đó lượng đường khử được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PSA (U/mg) được đo theo cùng cách thức.

Kết quả đo được thể hiện trên Fig.6 và Fig.7. Fig.6A và Fig.6B là đồ thị thể hiện kết quả đo hoạt tính thủy phân PSA ở các nhiệt độ tương ứng, trong đó trục hoành thể hiện nhiệt độ. Fig.7A và Fig.7B là đồ thị thể hiện kết quả đo hoạt tính thủy phân PSA ở các độ pH tương ứng, trong đó trục hoành thể hiện độ pH. 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 AR19G-166-RA có hoạt tính thủy phân PSA mạnh ở nhiệt độ nằm trong khoảng từ 60°C đến 80°C (Fig.6A). Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) mà ở đó enzym này có hoạt tính mạnh nhất bằng 70°C ở độ pH bằng 4,5 và 75°C ở độ pH nằm trong khoảng từ 5,0 đến 6,0. Khi phản ứng enzym được thực hiện ở nhiệt độ 85°C hoặc cao hơn, hoạt tính thủy phân PSA của enzym tinh khiết AR19G-166-RA bị giảm nhanh ở toàn bộ các khoảng pH. Trái lại, enzym tinh khiết AR19G-166-QV có hoạt tính thủy phân PSA thấp hơn AR19G-166-RA ở nhiệt độ 65°C hoặc cao hơn (Fig.6B). Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) mà ở đó enzym này có hoạt tính mạnh nhất bằng 70°C ở độ pH bằng 4,5 đến 5,5 và 75°C ở độ pH bằng 6,0. Hoạt tính thủy phân PSA của enzym tinh khiết AR19G-166-RA bị giảm nhanh ở toàn bộ các khoảng pH khi phản ứng enzym được thực hiện ở nhiệt độ 80°C hoặc cao hơn.

Hơn nữa, enzym tinh khiết AR19G-166-RA có hoạt tính thủy phân PSA mạnh nhất ở nhiệt độ nằm trong khoảng từ 65°C đến 80°C và độ pH nằm trong khoảng từ 4 đến 6 (Fig.7A). Độ pH tối ưu thay đổi phụ thuộc vào nhiệt độ phản ứng, và độ pH tối ưu bằng 4,6 (trị số đo thực) ở nhiệt độ nằm trong khoảng từ 60°C đến 65°C, độ pH tối ưu nằm trong khoảng từ 5,2 đến 5,3 (trị số đo thực) ở nhiệt độ 70 đến 75°C, và độ pH tối ưu bằng 5,8 (trị số đo thực) ở nhiệt độ 80°C. Enzym tinh khiết AR19G-166-RA có hoạt tính thủy phân PSA thấp ở độ pH nằm trong khoảng từ 3,2 đến 4,5 và độ pH nằm trong khoảng từ 7 đến 8. Mặt khác, tương tự như AR19G-166-RA, enzym tinh khiết AR19G-166-QV có hoạt tính thủy phân PSA mạnh nhất ở độ pH nằm trong khoảng từ 4 đến 6, và hoạt tính thủy phân PSA thấp ở độ pH nằm trong khoảng từ 3,2 đến 4,5 và độ pH nằm trong khoảng từ 7 đến 8 (Fig.7B).

14. Độ ổn định nhiệt của xenlobiohydrolaza chịu nhiệt theo sáng chế (thời gian hoạt tính enzym giảm 50%).

Để đánh giá độ ổn định nhiệt của AR19G-166-RA và AR19G-166-QV, quá trình ủ sơ bộ được thực hiện trong 20 phút đến 1440 phút, và hoạt tính thủy phân PSA của protein enzym đo ở các nhiệt độ tương ứng.

Trong phép đo này, các enzym tinh khiết (nồng độ cuối cùng bằng khoảng 1mg/mL) thu được từ phần 9 được sử dụng. Quá trình ủ sơ bộ mỗi enzym tinh khiết được thực hiện và bằng cách duy trì nhiệt độ của hỗn hợp (độ pH = 5,0) chứa 10 μ L enzym tinh khiết, 40 μ L nước tinh khiết, và 50 μ L dung dịch đệm axit axetic 200mM, ở nhiệt độ tương ứng nằm trong khoảng từ 50°C đến 80°C trong 0 phút, 20 phút, 40 phút, 60 phút, 120 phút, 240 phút, 480 phút, 960 phút, hoặc 1440 phút. Hoạt tính thủy phân PSA được đo theo cùng cách thức như nêu trong phần 10, chỉ khác là gia nhiệt riêng hỗn hợp đã ủ sơ bộ và dung dịch nước PSA 1% khối lượng tương ứng ở nhiệt độ 50°C trong 5 phút, sau đó bổ sung 100 μ L dung dịch nước PSA vào hỗn hợp, và để phản ứng trong 20 phút, trong đó lượng đường khử được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PSA (U/mg) được đo theo cùng cách thức.

Kết quả đo được thể hiện trên Fig.8A và Fig.8B. Hoạt tính enzym được thể hiện dưới dạng trị số tương đối (Hoạt tính tương đối %) với giả định rằng hoạt tính của lô không xử lý (không ủ sơ bộ) bằng 100%. Như được thể hiện đường gãy khúc và dốc xuống trên Fig.8A, thời gian ủ sơ bộ ở đó hoạt tính enzym giảm 50% so với hoạt tính của lô không xử lý được tính là thời gian bán giảm ($T_{\text{bán giảm}}$).

Khi nhiệt độ ủ sơ bộ nằm trong khoảng từ 50°C đến 60°C, enzym tinh khiết AR19G-166-RA không bị mất hoạt tính thủy phân PSA trong thời gian đo, và thời gian bán giảm $T_{\text{bán giảm}}$ bằng 24 giờ, là giới hạn trên của phép đo, hoặc lâu hơn. Ở nhiệt độ 70°C, thời gian bán giảm $T_{\text{bán giảm}}$ bằng 226 phút được xác định từ đường cong gần đúng dưới dạng hàm số mũ được thể hiện bằng đường in đậm trên Fig.8A. Ở nhiệt độ ủ sơ bộ bằng 80°C, hoạt tính enzym bị mất ngay lập tức (Fig.8A).

Mặt khác, độ ổn định nhiệt của enzym tinh khiết AR19G-166-QV thấp hơn nhiều so với AR19G-166-RA. Khi nhiệt độ ủ sơ bộ bằng 50°C, hoạt tính thủy phân PSA bị giảm

mạnh trong thời gian đo, và thời gian bán giảm $T_{\text{bán giảm}}$ bằng 24 giờ hoặc lâu hơn. Khi tăng nhiệt độ ủ sơ bộ, thì thời gian bán giảm $T_{\text{bán giảm}}$ của hoạt tính enzym ngắn, tức bằng 16 giờ ở nhiệt độ 60°C, và rất ngắn ở nhiệt độ 70°C, tức bằng 40 phút. Ở nhiệt độ ủ sơ bộ bằng 80°C, hoạt tính enzym bị mất ngay lập tức (Fig.8B).

15. Ảnh hưởng của ion kim loại hóa trị hai đến độ ổn định nhiệt của xenlobiohydrolaza chịu nhiệt theo sáng chế

Nhìn chung, các ion kim loại hóa trị hai có tác dụng làm ổn định cấu trúc của protein bằng cách gắn kết với protein, nhờ đó cải thiện được độ ổn định nhiệt. Mỗi ion kim loại hóa trị hai (nồng độ bằng 1mM) bao gồm canxi (Ca^{2+}), mangan (Mn^{2+}), coban (Co^{2+}), bari (Ba^{2+}), magie (Mg^{2+}), niken (Ni^{2+}), sắt (II) (Fe^{2+}), sắt (III) (Fe^{3+}), và kẽm (Zn^{2+}) được bổ sung vào dung dịch phản ứng enzym-cơ chất (PSA) để đánh giá ảnh hưởng đến hoạt tính thủy phân PSA của protein enzym. Trước tiên, hỗn hợp chứa 10 μL dung dịch enzym tinh khiết (nồng độ bằng 1mg/mL), 40 μL nước tinh khiết hoặc dung dịch nước chứa 5mM clorua của mỗi kim loại hóa trị hai, và 50 μL dung dịch đệm axit axetic 200mM (độ pH = 5,5) được ủ sơ bộ ở nhiệt độ 30°C trong 30 phút. Hoạt tính thủy phân PSA được đo theo cùng cách thức như nêu trong phần 10, chỉ khác là gia nhiệt riêng hỗn hợp đã ủ sơ bộ và dung dịch nước PSA 1% khối lượng tương ứng ở mỗi nhiệt độ nằm trong khoảng từ 30°C đến 100°C trong 5 phút, sau đó bổ sung 100 μL dung dịch nước PSA vào hỗn hợp, và để phản ứng ở mỗi nhiệt độ trong 20 phút, trong đó lượng đường khử được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PSA (U/mg) được đo theo cùng cách thức.

Kết quả đo cho thấy hoạt tính enzym của AR19G-166-RA và AR19G-166-QV tăng ở nhiệt độ 85°C hoặc cao hơn khi sử dụng ion canxi hoặc ion coban với nồng độ cuối cùng bằng 1mM, và hoạt tính enzym tăng đáng kể ở nhiệt độ nằm trong khoảng từ 60°C đến 90°C khi ion mangan (không thể hiện). Các ion kim loại hóa trị hai Ba^{2+} , Mg^{2+} , và Ni^{2+} không ảnh hưởng đến hoạt tính của enzym, và các ion kim loại hóa trị hai Fe^{2+} , Fe^{3+} , và Zn^{2+} không làm giảm hoạt tính của enzym.

16. Nhiệt độ nóng chảy của protein enzym xenlobiohydrolaza theo sáng chế

Độ ổn định nhiệt của protein cũng được đánh giá thông qua nhiệt độ biến tính hoặc nhiệt độ nóng chảy, T_m . Nhiệt độ ủ sơ bộ ở đó hoạt tính enzym giảm 50% so với hoạt tính

của lô không xử lý ủ sơ bộ được tính là nhiệt độ T_m của protein, và có thể được xác định bằng cách đo hoạt tính enzym, như được thể hiện đường gãy khúc và dốc xuống trên Fig.9A. Trị số T_m của AR19G-166-RA và AR19G-166-QV thu được bằng phương pháp này.

Hỗn hợp chứa 10 μL dung dịch enzym tinh khiết (nồng độ khoảng 1mg/mL), 40 μL nước tinh khiết dung dịch nước chứa 5mM clorua kim loại hóa trị hai (CaCl_2 hoặc MnCl_2), và 50 μL dung dịch đệm axit axetic 200mM (độ pH = 5,0) được ủ ở nhiệt độ 30°C trong 30 phút, sau đó ủ sơ bộ ở nhiệt độ 40°C đến 100°C trong 30 phút. Hoạt tính thủy phân PSA được đo theo cùng cách thức như nêu trong phần 10, chỉ khác là gia nhiệt riêng hỗn hợp đã ủ sơ bộ và dung dịch nước PSA 1% khối lượng tương ứng ở nhiệt độ 50°C trong 5 phút, sau đó bổ sung 100 μL dung dịch nước PSA vào hỗn hợp, và để phản ứng ở nhiệt độ 50°C trong 20 phút, trong đó lượng đường khử được tạo ra bằng quá trình thủy phân, và hoạt tính thủy phân PSA (U/mg) được đo theo cùng cách thức. Đối với mỗi nhiệt độ ủ sơ bộ, hoạt tính thủy phân PSA enzym này được đo ba lần, để thu được trị số trung bình và độ lệch chuẩn. Dữ liệu được chuẩn hóa bằng cách giả định rằng hoạt tính thủy phân trung bình ở vùng nhiệt độ thấp (nằm trong khoảng từ 40°C đến 65°C) và vùng nhiệt độ cao (nằm trong khoảng từ 90°C đến 100°C), ở đó hoạt tính thủy phân đã cân bằng lần lượt bằng 1 và 0.

Hoạt tính thủy phân PSA của protein AR19G-166-RA và protein AR19G-166-QV ở các nhiệt độ ủ sơ bộ tương ứng và các đường cong sigma gần đúng được thể hiện trên Fig.9A và Fig.9B. Từ đường cong gần đúng, trị số T_m của protein enzym (tức là, nhiệt độ ủ sơ bộ ở thời điểm khi hoạt tính chuẩn hóa = 0,5) được xác định. Trị số T_m của các protein enzym được xác định từ các đường cong gần đúng của dữ liệu về hoạt tính thủy phân PSA được thể hiện trong Bảng 3. Trong Bảng 3, ký hiệu “đối chứng” thể hiện kết quả phản ứng, trong đó nước tinh khiết được bổ sung thay cho dung dịch nước chứa 5mM clorua kim loại hóa trị hai. Trị số T_m của AR19G-166-RA bằng 76,4°C, trong khi đó trị số T_m của AR19G-166-QV bằng 68,5°C, tức là trị số T_m của AR19G-166-RA lớn hơn trị số T_m của AR19G-166-QV khoảng 8°C. Ngoài ra, trị số T_m của AR19G-166-RA và AR19G-166-QV lần lượt tăng 3,0°C và 4,1°C khi sử dụng ion mangan ở nồng độ bằng 1mM. Trị số T_m của AR19G-166-RA đã tăng nhẹ 0,7°C), trong khi đó trị số T_m của AR19G-166-QV đã tăng 3,1°C, khi sử dụng ion canxi ở nồng độ bằng 1mM.

Bảng 3. Nhiệt độ nóng chảy (T_m) của AR19G-166-RA và AR19G-166-QV

Enzym	T_m (°C)		
	Đối chứng	+ 1mM Ca^{2+}	+ 1mM Mn^{2+}
AR19G-166-RA	76,4	77,1	79,4
AR19G-166-QV	68,5	71,6	74,6

Ví dụ 2

Khi gen tùy ý được chèn vào vật chủ là sinh vật nhân thật như nấm sợi hoặc thực vật, thì nhiệt độ tối ưu của protein được biểu hiện thường tăng từ 5°C đến 10°C. Điều này là do các phản ứng biến đổi sau dịch mã để bổ sung đường vào protein, cũng được gọi là phản ứng glycosyl hóa. Các protein được glycosyl hóa trở nên bền nhiệt hơn. Gen AR19G-166 được chèn vào nấm sợi *Aspergillus oryzae*, để kiểm chứng ảnh hưởng của phản ứng glycosyl hóa đến độ ổn định nhiệt của protein enzym được mã hóa.

1. Sản xuất thể biến nạp *Aspergillus*

Thế đột biến AR19G-166-RW (trình tự axit amin: SEQ ID NO.5 và trình tự nucleotit: SEQ ID NO.6) và thế đột biến AR19G-166-QW (trình tự axit amin: SEQ ID NO.7 và trình tự nucleotit: SEQ ID NO.8) trong đó gốc axit amin ở vị trí số 351 của AR19G-166-RA và AR19G-166-QV được thay thế bằng tryptophan (W), được tạo ra. Cat-xet biểu hiện chứa mỗi thế đột biến này được chèn vào *Aspergillus*. Kết quả tạo ra được các thể biến nạp *Aspergillus* để biểu hiện AR19G-166-RW hoặc AR19G-166-QW.

Các thế đột biến gen thay thế axit amin AR19G-166 này được kết hợp vào vector chuyển gen để thiết kế vector biểu hiện. Vector này được chuyển nạp vào vật chủ *Aspergillus*. Thể biến nạp *Aspergillus* này được tạo ra bằng phương pháp biểu hiện protein do Ozeki Co., Ltd. cung cấp (<http://www.ozeki.co.jp/>). Thể biến nạp *Aspergillus* được tạo ra bằng phương pháp này là được tạo ra bằng phương pháp tái tổ hợp tương đồng, trong đó gen chèn được kết hợp vào vị trí đặc hiệu của nhiễm sắc thể. Do đó, mặc dù khả năng của nấm sợi thấp, nhưng khi được kết hợp vào nhiễm sắc thể, thì gen chèn được duy trì ổn định và protein mã hóa bởi gen chèn có thể được tạo ra ổn định. Hơn nữa, do peptit tín hiệu bài tiết được bổ sung, protein mã hóa bởi gen chèn được bài tiết từ cơ thể vi khuẩn vào dịch nuôi cấy.

2. Quá trình cô protein được tạo ra bởi thể biến nạp *Aspergillus*

Thế biến nạp *Aspergillus* thu được được nuôi cấy, sau đó phân dịch nổi nuôi cấy

được thu nhận. Phần dịch nổi nuôi cấy này được cô đến thể tích bằng 1/10 bằng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất), và trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0). Dung dịch này được sử dụng làm dịch nổi nuôi cấy cô đặc.

Mỗi dịch nổi nuôi cấy cô đặc của các *thể biến nạp Aspergillus* tương ứng (5 μ L) và dịch nổi nuôi cấy đồng nhất *E.coli* tái tổ hợp gen AR19G-166-RA và AR19G-166-QV (5 μ L) được điều chế trong Ví dụ 1 được phân tích thẩm tách Western. Phân tích thẩm tách Western được thực hiện theo cùng cách thức như nêu trong phần 9 của Ví dụ 1, và kết quả được thể hiện trong Fig.10. Từ phần dịch nổi nuôi cấy cô đặc của thể biến nạp *Aspergillus* chứa gen R19G-166-RW và thể biến nạp *Aspergillus* chứa gen AR19G-166-QW, quan sát thấy được các băng rất mờ tương ứng với các băng duy nhất của dịch nổi nuôi cấy đồng nhất *E.coli* tái tổ hợp gen AR19G-166 (làn 2 và làn 3 trên Fig.10) có trọng lượng phân tử bằng 46,7kD, và các băng đậm và rộng có trọng lượng phân tử nằm trong khoảng từ 50kDa đến 55kDa (làn 2 và làn 3 trên Fig.10). Theo phương thức này, trọng lượng phân tử quan sát được của hầu hết các protein mã hóa bởi gen AR19G-166 được biểu hiện bởi vật chủ *Aspergillus* đã tăng từ 3kDa đến 8kDa. Trọng lượng phân tử của các protein này tăng có là do phản ứng glycosyl hóa.

3. Quy trình sản xuất thể biến nạp *E. coli* chứa gen AR19G-166-RW và AR19G-166-QW và tinh chế protein enzym xenlobiohydrolaza

Các plasmit trong đó các thể đột biến AR19G-166-RW và AR19G-166-QW đã được chèn vào vectơ pET101/D-TOPO được tạo ra theo cùng cách thức như nêu trong phần 8 của Ví dụ 1. Thể biến nạp *E. coli* chứa các plasmit này để biểu hiện protein được tạo ra theo cùng cách thức như nêu trong phần 9 của Ví dụ 1. Do đó, các thể biến nạp *E. coli* biểu hiện cảm ứng. Sau khi nuôi cấy, *E. coli* được thu nhận bằng ly tâm. Dịch chiết thô của *E. coli* tái tổ hợp chứa protein đích được thu nhận và lọc. Kết quả là, thu được phần dịch nổi nuôi cấy đồng nhất *E. coli*. Quá trình phân đoạn và tinh chế được thực hiện bằng cột trao đổi ion, cột phân tách tương tác kỵ nước, và cột lọc gel, theo cùng cách thức như nêu trong phần 9 của Ví dụ 1. Kết quả là, thu được enzym tinh khiết có nồng độ cuối cùng bằng khoảng 1mg/mL.

4. Mức độ phụ thuộc vào nhiệt độ của hoạt tính xenlobiohydrolaza

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA của AR19G-166-RW

và AR19G-166-QW được tạo ra bởi thể biến nạp *Aspergillus* tái tổ hợp gen và thể biến nạp *E. coli* tái tổ hợp gen được đánh giá. Trong phép đo này, các dịch nổi nuôi cấy cô đặc thu được ở phần 2 nêu trên và các enzym tinh khiết thu được ở phần 3 được sử dụng.

Hoạt tính thủy phân PSA ở mỗi nhiệt độ được đo theo cùng cách thức như nêu trong phần 13 của Ví dụ 1, chỉ khác là sử dụng các dịch nổi nuôi cấy cô đặc thu được ở phần 2 nêu trên, và thiết lập độ pH của dung dịch phản ứng bằng 5,5, trong đó lượng đường khử được tạo ra bởi quá trình thủy phân enzym được đo theo cùng cách thức. Hoạt tính thủy phân PSA của các enzym tinh khiết của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *E. coli* được đo theo cùng cách thức như nêu trong phần 10 của Ví dụ 1, tức là hoạt tính enzym để tạo ra 1 μ mol đường khử trong 1 phút được tính bằng 1U, và trị số thu được bằng cách chia hoạt tính enzym cho khối lượng protein được tính bằng hoạt tính riêng (U/mg). Mặt khác, hoạt tính thủy phân PSA của AR19G-166-RW và AR19G-166-QW được tạo ra bởi thể biến nạp *Aspergillus* được đo dưới dạng hoạt tính tương đối (%) giả định rằng lượng đường khử được tạo ra bởi AR19G-166-RW có hoạt tính thủy phân mạnh nhất ở nhiệt độ 100°C bằng 100%.

Hoạt tính tương đối (%) của hoạt tính thủy phân PSA ở các nhiệt độ tương ứng được thể hiện trên Fig.11A. Trên Fig.11A, kết quả đo hoạt tính thủy phân PSA (U/mg) của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *E. coli* đã được đo trong phần 13 của Ví dụ 1 cũng được thể hiện (trên Fig.11A lần lượt được thể hiện bởi ký hiệu “RW bởi *E. coli*” và “QW bởi *E. coli*”). Nhiệt độ tối ưu ($T_{\text{tối ưu}}$) của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *E. coli* lần lượt bằng 80°C và 75°C. Mặt khác, nhiệt độ tối ưu của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *Aspergillus* (trên Fig.11A lần lượt được thể hiện bởi ký hiệu “RW bởi *A. oryzae*” và “QW bởi *A. oryzae*”) đều bằng 100°C hoặc cao hơn. Theo cách thức này, trọng lượng phân tử đã tăng khoảng 10%, và nhiệt độ tối ưu đã tăng từ 20°C đến 30°C, khi biểu hiện gen AR19G-166 bởi *Aspergillus*. AR19G-166 được biểu hiện bởi *Aspergillus* có nhiệt độ tối ưu tăng 30°C hoặc cao hơn và độ ổn định nhiệt cao hơn xenlobiohydrolaza thông thường có nguồn gốc từ nấm sợi ưa nhiệt.

Từ các kết quả này, có thể thu được enzym có hoạt tính xenlobiohydrolaza mạnh ngay ở cả nhiệt độ 80°C hoặc cao hơn, bằng cách biểu hiện xenlobiohydrolaza chịu nhiệt

theo sáng chế bằng hệ biểu hiện là sinh vật nhân thật như nấm sợi. Có thể sử dụng xenlobiohydrolaza chịu nhiệt có hoạt tính enzym mạnh ngay ở cả nhiệt độ nằm trong khoảng từ 80°C đến 100°C cùng với các xenlulaza chịu nhiệt khác, để thủy phân lignoxenluloza ở nhiệt độ 80°C hoặc cao hơn.

5. Mức độ phụ thuộc vào độ pH của hoạt tính xenlobiohydrolaza

Mức độ phụ thuộc vào độ pH của hoạt tính thủy phân PSA của AR19G-166-RW và AR19G-166-QW được tạo ra bởi thể biến nạp *Aspergillus* và thể biến nạp *E. coli* được đánh giá. Trong phép đo này, các dịch nổi nuôi cấy cô đặc thu được ở phần 2 nêu trên và các enzym tinh khiết thu được ở phần 3 được sử dụng.

Hoạt tính thủy phân PSA ở mỗi độ pH được đo theo cùng cách thức như nêu trong phần 13 của Ví dụ 1, chỉ khác là sử dụng các dịch nổi nuôi cấy cô đặc thu được ở phần 2 nêu trên, trong đó lượng đường khử được tạo ra bởi quá trình thủy phân enzym được đo theo cùng cách thức. AR19G-166-RW và AR19G-166-QW được tạo ra bởi thể biến nạp *E.coli* được cho thủy phân PSA lần lượt ở nhiệt độ 80°C và 75°C, và hoạt tính thủy phân PSA của chúng (U/mg) được đo. AR19G-166-RW và AR19G-166-QW được tạo ra bởi thể biến nạp *Aspergillus* được cho thủy phân PSA ở nhiệt độ của 80°C, và hoạt tính thủy phân PSA của chúng ở các độ pH tương ứng được đo dưới dạng hoạt tính tương đối (%) giả định rằng hoạt tính thủy phân PSA mạnh nhất ở độ pH bằng 5,2 bằng 100%.

Hoạt tính tương đối (%) của hoạt tính thủy phân PSA ở các độ pH tương ứng được thể hiện trên Fig.11B. Trên Fig.11B, kết quả đo hoạt tính thủy phân PSA (U/mg) của AR19G-166-RW và AR19G-166-QW được biểu hiện bởi *E. coli* đã được đo trong phần 13 của Ví dụ 1 cũng được thể hiện (trên Fig.11B lần lượt được thể hiện bởi ký hiệu “RW bởi *E. coli*” và “QW bởi *E. coli*”). Các enzym AR19G-166 được biểu hiện bởi *Aspergillus* (trên Fig.11B lần lượt được thể hiện bởi ký hiệu “RW bởi *A. oryzae*” và “QW bởi *A. oryzae*”) có khoảng pH đặc hiệu rộng hơn các enzym AR19G-166 được biểu hiện bởi *E. coli* (trên Fig.11B lần lượt được thể hiện bởi ký hiệu “RW bởi *E. coli*” và “QW bởi *E. coli*”). Độ pH ở đó hoạt tính của các enzym này bằng 50% hoạt tính tối đa nằm trong khoảng từ 3,3 đến 8,0 ở nhiệt độ 80°C. Mặt khác, độ pH ở đó hoạt tính của các enzym này bằng 50% hoạt tính tối đa nằm trong khoảng từ 4,5 đến 6,8 khi được biểu hiện bởi *E. coli*.

Ví dụ 3

Nhìn chung, protein mã hóa bởi gen tái tổ hợp được sản sinh ở vật chủ là thực vật có mức độ biểu hiện cao hơn hệ biểu hiện vi khuẩn như *E. coli* hoặc *Bacillus*, và hệ biểu hiện nấm sợi như *Aspergillus*. Cụ thể, thể biến nạp lục lập tích tụ protein ngoại sinh ở tỷ lệ nằm trong khoảng từ 5% đến 10% khối lượng tính theo tổng khối lượng protein hòa tan ở lá của thể biến nạp, và có thể tích tụ protein ngoại sinh ở tỷ lệ lên đến 40% khối lượng hoặc lớn hơn tính theo tổng khối lượng protein hòa tan ở lá của thể biến nạp. Mặt khác, thể biến nạp bộ gen nhân tích tụ protein ngoại sinh ở tỷ lệ bằng khoảng 1% khối lượng tính theo tổng khối lượng protein hòa tan ở mô thể biến nạp. Theo phương thức này, quá trình sản sinh protein bởi thực vật, và đặc biệt quá trình sản sinh protein ngoại sinh bởi thể biến nạp lục lập, có giá trị kinh tế cao hơn nấm sợi hoặc vi khuẩn, mà thường được sử dụng để sản xuất protein. Theo đó, cây thuốc lá chuyển gen lục lập được tạo ra bằng các chèn AR19G-166-RA và AR19G-166-QV vào bộ gen lục lập của cây thuốc lá.

1. Quy trình sản xuất cây thuốc lá chuyển gen lục lập

Vùng chèn gen vào lục lập, vùng điều hòa biểu hiện 5'/3', và gen chỉ thị chọn lọc và các gen tương tự để tạo ra thực vật chuyển gen lục lập được thiết kế theo phương pháp được mô tả trong tài liệu: Daniell *et al.*, *Methods in Molecular Biology*, 2005, Vol. 286, p. 111-138; Verma and Daniell, *Plant Physiol.*, 2007, Vol. 145, p. 1129-1143.

Trước tiên, các vector cat-xet pNtaGL và pNtaGLPL để chuyển gen vào lục lập của cây thuốc lá có cấu trúc được thể hiện trên Fig.12A được tạo ra. Các vector này là các vector để chèn gen đích bằng cách tái tổ hợp tương đồng vào vùng hai gen trnI-trnA ở trình tự lặp ngược chiều bên trong bộ gen lục lập của cây thuốc lá. Các vector này có cat-xet biểu hiện gen aadA (aminoglucosit 3'-adenylyltransferaza) dưới dạng gen chỉ thị chọn lọc mã hóa tính kháng spectinomycin, và vị trí chèn cat-xet biểu hiện (vị trí ClaI-BsiWI) cho gen đích nằm sau cat-xet biểu hiện gen aadA. Vùng khởi đầu gen ARN ribosom 16S có nguồn gốc từ cây thuốc lá (Prn) được chèn vào vector pNtaGL dưới dạng vùng điều hòa biểu hiện gen aadA, trong khi đó vector pNtaGLPL không chứa vùng khởi đầu, và quá trình biểu hiện gen aadA phụ thuộc vào vùng khởi đầu nội sinh nằm trước vùng tái tổ hợp tương đồng.

Fig.12B thể hiện các cat-xet để biểu hiện gen đích pPXT và pPXTPL. Cả hai cat-xet được thiết kế để chèn gen đích vào vị trí BamHI, và lần lượt có vùng khởi đầu Prn và trình

tự gen 10 có nguồn gốc từ thực khuẩn thể T7 (T7g10), hoặc chỉ có trình tự T7g10, ở đầu 5' của gen đích. Trình tự T7g10 được cho hữu hiệu để tích tụ nhiều protein ngoại sinh trong quá trình biểu hiện gen ngoại sinh ở lục lạp. Đầu 3'-UTR (TrbcL) của gen *rbcL* ở lục lạp của cây thuốc lá được định vị ở đầu 3' của gen đích ở cả hai cat-xet, và trong mỗi trường hợp, có thể được chèn vào vector cat-xet pNtaGL hoặc pNtaGLPL để chuyển gen lục lạp được thể hiện trên Fig.12A bằng cách sử dụng vị trí ClaI và vị trí BsiWI ở đầu đối diện của mỗi cat-xet.

Các sản phẩm tách dòng PCR AR19G-166-RA và AR19G-166-QV lần lượt được khuếch đại bằng các đoạn mồi chứa gốc liên kết BamHI, và các sản phẩm khuếch đại thu được được chèn vào vị trí BamHI của cat-xet biểu hiện pPXT hoặc pPXTPL. Cat-xet biểu hiện pPXT trong đó AR19G-166-QV đã được chèn được kết hợp vào pNtaGL bằng cách sử dụng vị trí ClaI và BsiWI, nhờ đó tạo ra gen chuyển lục lạp pNtaGL-QV có cấu trúc trong đó cat-xet biểu hiện gen *aadA* và cat-xet biểu hiện gen AR19G-166-QV đã được liên kết song song trong vùng hai gen *trnI-trnA*. Cat-xet biểu hiện pPXTPL trong đó AR19G-166-RA đã được chèn được kết hợp vào pNtaGLPL theo cùng cách thức, nhờ đó tạo ra gen chuyển lục lạp pNtaGLPL-RA. Fig.12C là hình vẽ thể hiện cấu trúc của các vector chuyển gen lục lạp pNtaGL-QV và pNtaGLPL-RA. Trên Fig.12C, “QV” tức là AR19G-166-QV và “RA” tức là AR19G-166-RA.

Quá trình chuyển gen vào lục lạp của cây thuốc lá được thực hiện theo phương pháp được mô tả trong tài liệu: Daniell *et al.*, *Methods in Molecular Biology*, **2005**, Vol. 286, p. 111-138. Cụ thể, các lá còn tươi của cây thuốc lá (*Nicotiana tabacum* cv. SR-1) đã được khử trùng và sinh trưởng trong môi trường đĩa thạch MS (môi trường Murashige and Skoog (MS) (độ pH = 5,8) chứa 30 g/L sucroza được hóa rắn bằng 3g/L gồm gellan) được cắt thành các mảnh có kích cỡ 0,5cm², và các mảnh lá đã được cắt được đặt vào đĩa Petri ở vị trí trung tâm đã chứa môi trường RMOP (Svab *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **1990**, Vol. 87, p. 8526-8530) sao cho các mặt trục của lá được tiếp xúc với môi trường. Sau khi nuôi cấy trong ngày với 16 giờ chiếu sáng và 8 giờ trong bóng tối, các vector đối chứng pNtaQV hoặc pNtaRA, và pNtaGL hoặc pNtaGLPL để phân tích so sánh, lần lượt được chèn vào bằng súng bắn gen (PDS-1000He, zz).

Khoảng 100mg các mảnh lá được thu nhận từ các chồi có khả năng chống chịu spectinomycin thu được từ ba hoặc bốn khử biệt hóa và tái sinh lặp lại, và các ADN được

chiết xuất từ các mảnh lá này. Bằng cách sử dụng khuôn là ADN này, phản ứng PCR được thực hiện bằng cách sử dụng các đoạn mồi được thiết kế để khuếch đại đặc hiệu các trình tự chèn, nhờ đó chọn lọc được các cá thể có mỗi gen đích được chèn vào vùng trnI-trnA của bộ gen lục lạp của cây thuốc lá.

Các cá thể được kiểm chứng là đã chứa gen AR19G-166-RA hoặc AR19G-166-QV và các vector đối chứng tương ứng được trồng vào môi trường MS chứa 500mg/L spectinomycin để kích thích ra rễ và phát triển thân cây. Các ADN được chiết xuất từ thực vật đã sinh trưởng đến chiều cao khoảng 10cm bằng Kit DNeasy (do QIAGEN sản xuất), và quá trình lai tương đồng của gen chèn được kiểm chứng bằng phương pháp thấm tách Southern. Trước tiên, mỗi ADN (1µg) được chiết xuất từ lá của cây thuốc lá chuyển gen lục lạp hoặc cây thuốc lá thể đại (WT (SR-1)) được phân cắt bằng enzym cắt giới hạn BglII. Các ADN đã phân cắt bằng enzym BglII được phân tích bằng phương pháp thấm tách Southern bằng cách sử dụng mẫu dò có trình tự nucleotit giống hết trình tự của vùng nằm trước đoạn trnI khoảng 1kb, và mẫu dò được phát hiện bằng phương pháp hóa quang bằng cách sử dụng kit AlkPhos Direct do GE Healthcare sản xuất. Theo lý thuyết, mẫu dò sẽ phát hiện các mảnh ADN có chiều dài khoảng 4,5kb đối với cây thuốc lá thể đại, các mảnh ADN có chiều dài khoảng 7,1kb đối với các cây thuốc lá chuyển gen lục lạp pNtaGL-QV và pNtaGLPL-RA, và các mảnh ADN có chiều dài khoảng 6,0kb hoặc 5,9kb đối với các cây thuốc lá chuyển gen lục lạp pNtaGL hoặc pNtaGLPL. Các kết quả phân tích thấm tách Southern của hai dòng cây thuốc lá chuyển gen lục lạp pNtaGL-QV (QV-2 và QV-17), và ba dòng cây thuốc lá chuyển gen lục lạp pNtaGLPL-RA (RA-6-2-1, RA-6-2-2 và RA-6-2-3), lần lượt được thể hiện trên Fig.13A và 13B. Ngoài ra, các kết quả phân tích thấm tách Southern của các cây thuốc lá chuyển gen lục lạp pNtaGL và pNtaGLPL được thể hiện trên Fig.13C. Trong toàn bộ các trường hợp, chỉ có băng có chiều dài 7,1kb, băng có chiều dài 6,0kb hoặc băng có chiều dài 5,9kb thu được từ gen lục lạp tái tổ hợp được phát hiện, và không phát hiện được băng có chiều dài khoảng 4,5kb thu được từ cây thuốc lá thể đại, do đó các dòng này được kiểm chứng là các cây thuốc lá chuyển gen đồng loài, trong đó toàn bộ các lục lạp được tái tổ hợp gen.

Các cây thuốc lá chuyển gen đồng loài được sinh trưởng tiếp, trồng vào chậu đất có kích cỡ 18cm khi chúng đã có khoảng 7 lá, sau đó chuyển vào nhà kính. Các cây thuốc lá chuyển gen lục lạp AR19G-166-QV và AR19G-166-RA không có hình thái bất thường

trong sinh trình tăng trưởng tiếp theo, và trong quá trình phát triển hạt. Phân tích kiểu hình của các cây thuốc lá thể hệ T_1 cho thấy mặc dù sinh trưởng hơi chậm hơn so với cây thuốc lá thể đại và cây thuốc lá chuyển gen đối chứng, nhưng các cây thuốc lá chuyển gen AR19G-166-QV và AR19G-166-RA không có hình thái bất thường, và lượng sinh khối trong giai đoạn ra hoa tương đương với các cây thuốc lá chuyển gen đối chứng. Fig.14A là hình vẽ thể hiện giai đoạn ra hoa của cây thuốc lá chuyển gen lục lạp AR19G-166-QV (thể hệ T_1) và cây thuốc lá chuyển gen lục lạp vector đối chứng pNtaGL (thể hệ T_1), và Fig.14B là hình vẽ thể hiện giai đoạn ra hoa của cây thuốc lá chuyển gen lục lạp AR19G-166-RA (thể hệ T_1) và cây thuốc lá chuyển gen lục lạp vector đối chứng pNtaGLPL (thể hệ T_1).

2. Quá trình chiết xuất protein enzym xenlobiohydrolaza từ cây thuốc lá chuyển gen lục lạp

Quá trình chiết xuất protein enzym xenlobiohydrolaza được thực hiện bằng cách chọn lọc một dòng của mỗi cây thuốc lá chuyển gen lục lạp AR19G-166-QV và AR19G-166-RA, sau đó thực hiện quá trình chiết xuất từ 3 cá thể của mỗi dòng theo cách thức dưới đây. Ba lá được cắt từ vùng tâm của cây thuốc lá chuyển gen lục lạp trong giai đoạn ra hoa, và 5 đến 10 mảnh lá (khoảng 100mg) được cắt từ mỗi lá. Các mảnh lá được đặt vào ống lấy mẫu có dung tích 2mL chứa 3 hạt wonfram có đường kính bằng 3mm (do QIAGEN sản xuất), và ống lấy mẫu này được đặt vào nitơ lỏng và bảo quản lạnh đông. Các mảnh lá lạnh đông được nghiền bằng máy nghiền trộn MM400 (do Retsch sản xuất) ở tần số 30 Hz trong 90 giây, và hỗn dịch được điều chế bằng cách bổ sung dung dịch đệm axit axetic 50mM (độ pH = 5,5) chứa 1% thể tích chất ức chế proteaza (do Sigma sản xuất) ở lượng gấp 10 lần khối lượng các mảnh lá. Sau khi trộn kỹ hỗn dịch này, quá trình phân tách ly tâm được thực hiện, để thu được dịch chiết chứa protein enzym hòa tan (AR19G-166-QV hoặc AR19G-166-RA). Trong trường hợp của AR19G-166-RA, dịch chiết được cô đến thể tích bằng 1/5 đến 1/10 bằng cách sử dụng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất). Ngoài ra, bằng cách sử dụng cùng quy trình, các dịch chiết protein hòa tan cũng được điều chế từ cây thuốc lá chuyển gen vector pNtaGL hoặc pNtaGLPL được sử dụng làm các mẫu đối chứng để phân tích so sánh.

Các dịch chiết protein hòa tan được kiểm chứng bằng phân tích điện di trên gel SDS-PAGE và phân tích thẩm tách Western. Quá trình điện di trên gel SDS và thẩm tách Western được thực hiện bằng cách sử dụng gel không chứa thuốc nhuộm Mini Protean

TGX (do Bio-Rad sản xuất). Mỗi dịch chiết và enzym tinh khiết được trộn với dung dịch Tris-SDS β -ME (do Cosmo Bio Co. Ltd. sản xuất) ở tỷ lệ 1:1, sau đó gia nhiệt ở nhiệt độ 100°C trong 10 phút. AR19G-166-QV (5 μ L), AR19G-166-RA (10 μ L) và mẫu đối chứng, và enzym tinh khiết (0,2 μ g) lần lượt được chạy điện di. Sau khi chạy điện di, các băng protein được phát hiện bằng thuốc nhuộm CBB. Phân tích thâm tách Western được thực hiện theo cùng cách thức nêu trong phần 9 của ví dụ 1.

Fig.15A đến Fig.15D thể hiện các kết quả phân tích điện di trên gel SDS-PAGE (Fig.15A) và phân tích thâm tách Western (Fig.15B) của các dịch chiết protein hòa tan thu được từ cây thuốc lá chuyển gen lục lạp AR19G-166-QV và cây thuốc lá chuyển gen lục lạp pNtaGL, và các kết quả phân tích điện di trên gel SDS-PAGE (Fig.15C) và phân tích thâm tách Western (Fig.15D) of các dịch chiết protein hòa tan thu được từ cây thuốc lá chuyển gen lục lạp AR19G-166-RA và cây thuốc lá chuyển gen lục lạp pNtaGLPL. Trong mỗi Fig.15A đến Fig.15D, làn 1 là gen chỉ thị phân tử protein, làn 2 là protein enzym tinh khiết, làn 3 đến 5 là các dịch chiết protein hòa tan thu được từ 3 cây thuốc lá chuyển gen lục lạp AR19G-166-QV hoặc AR19G-166-RA, và làn 6 đến 8 là các dịch chiết protein hòa tan thu được từ 3 cây thuốc lá chuyển gen lục lạp pNtaGL hoặc pNtaGLPL.

Các protein enzym AR19G-166-QV và AR19G-166-RA đều được biểu hiện trong lục lạp của cây thuốc lá. Trong phân tích điện di trên gel SDS-PAGE của các dịch chiết protein hòa tan của cây thuốc lá chuyển gen lục lạp AR19G-166-QV và AR19G-166-RA, các băng được phát hiện (trên làn 3 đến 5 của Fig.15A và Fig.15C) ở vị trí tương ứng với protein enzym tinh khiết tương ứng (làn 2 của Fig.15A và Fig.15C). Trái lại, các băng không được phát hiện ở các cây thuốc lá chuyển gen đối chứng chứa pNtaGL hoặc pNtaGLPL (làn 6 đến 8 của Fig.15A và Fig.15C).

Phân tích thâm tách Western được thực hiện bằng cách sử dụng kháng thể kháng polypeptit chứa 20 gốc axit amin từ vị trí số 384 đến 403 của AR19G-166, và trong các dịch chiết protein hòa tan của cả cây thuốc lá chuyển gen lục lạp AR19G-166-QV (làn 3 đến 5 của Fig.15B) lẫn cây thuốc lá chuyển gen lục lạp AR19G-166-RA (làn 3 đến 5 của Fig.15D), các băng được phát hiện ở vị trí tương ứng với protein enzym tinh khiết tương ứng (làn 2 của Fig.15B và Fig.15D). Ngoài ra, ở AR19G-166-QV và AR19G-166-RA, các băng cũng xuất hiện ở vị trí trọng lượng phân tử thấp hơn protein enzym. Điều này có thể do trộn protein enzym dẫn đến phân giải một phần protein enzym có trong cây thuốc lá,

hoặc protein enzym không được biểu hiện hoàn toàn. Trái lại, ở các cây thuốc lá đối chứng chuyển gen pNtaGL hoặc pNtaGLPL (lần 6 đến 8 của Fig.15B và Fig.15D), các băng hoàn toàn không được phát hiện.

3. Mức độ phụ thuộc vào nhiệt độ của hoạt tính xenlobiohydrolaza

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA của AR19G-166-QV và AR19G-166-RA được tạo ra bởi cây thuốc lá chuyển gen lục lạp được đánh giá. Trong phép đo này, dịch chiết protein hòa tan thu được nêu trên được sử dụng.

Hoạt tính thủy phân PSA ở mỗi nhiệt độ được đo bằng cách sử dụng dịch chiết protein hòa tan thu được nêu trên, theo cùng cách thức như nêu trong phần 10 của ví dụ 1, và lượng đường khử được tạo ra bởi quá trình thủy phân enzym được đo.

Hoạt tính thủy phân PSA ở các nhiệt độ khác nhau của protein AR19G-166-QV và protein AR19G-166-RA được biểu hiện trong lục lạp của cây thuốc lá lần lượt được thể hiện trên Fig.16A và Fig.16B. Trên Fig.16A và Fig.16B, hoạt tính enzym ở nhiệt độ nằm trong khoảng từ 30°C đến 100°C được xác định bằng lượng đường khử. Các kết quả cho thấy AR19G-166-QV và AR19G-166-RA được biểu hiện trong lục lạp của cây thuốc lá có cùng mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA so với AR19G-166-QV và AR19G-166-RA được biểu hiện trong *E. coli*, chứng tỏ chức năng protein được biểu hiện bình thường. Mặc dù các kết quả hơi khác biệt giữa các cây, AR19G-166-QV được biểu hiện ở lục lạp của cây thuốc lá (Fig.16A) có nhiệt độ tối ưu ($T_{\text{tối ưu}}$) nằm trong khoảng từ 70°C đến 75°C, và có độ ổn định nhiệt gần như tương đương với AR19G-166-QV được biểu hiện trong *E. coli*. Tương tự, mặc dù các kết quả hơi khác biệt giữa các cây, nhưng AR19G-166-RA được biểu hiện ở lục lạp của cây thuốc lá (Fig.16B) cũng có nhiệt độ tối ưu ($T_{\text{tối ưu}}$) nằm trong khoảng từ 75°C đến 80°C, và có độ ổn định nhiệt gần như tương đương với AR19G-166-RA được biểu hiện trong *E. coli*. Trái lại, các cây thuốc lá đối chứng chuyển gen pNtaGL hoặc pNtaGLPL, không có hoạt tính thủy phân PSA.

Ví dụ 4

Khi gen tùy ý được chèn vào vật chủ là sinh vật nhân thật như nấm sợi hoặc thực vật, thì nhiệt độ tối ưu của protein được biểu hiện thường tăng từ 5°C đến 10°C. Điều này là do các phản ứng biến đổi sau dịch mã để bổ sung đường vào protein, cũng được gọi là phản

ứng glycosyl hóa. Các protein được glycosyl hóa trở nên bền nhiệt hơn. Trên trình tự axit amin của gen AR19G-166-RA và AR19G-166-QV, cấu trúc glycosyl hóa liên kết N Asn-Xaa-Ser/Thr có ở 4 vị trí. Khi gen được biểu hiện bằng vật chủ là sinh vật nhân thật, thì phản ứng glycosyl hóa có thể xuất hiện ở các cấu trúc này, nhờ đó làm tăng độ ổn định nhiệt. Gen AR19G-166-RA và AR19G-166-QV được chèn vào cây mù tạc *Arabidopsis thaliana* thuộc họ Cải (*Brassicaceae*), để kiểm chứng ảnh hưởng của phản ứng glycosyl hóa đến độ ổn định nhiệt của protein enzym được mã hóa.

1. Quy trình sản xuất cây mù tạc chuyển gen (*Arabidopsis thaliana*)

PCR được thực hiện bằng cách sử dụng các khuôn là gen AR19G-166-RA và AR19G-166-QV, được chèn vào vector biểu hiện thực vật tích tụ không bào pIG121Bar. Bằng cách sử dụng phương pháp lạnh đông và rã đông, vector biểu hiện này được chèn vào *Agrobacterium tumefaciens*. Cụ thể, vector biểu hiện plasmid này (khoảng 1µg) được bổ sung vào chủng *Agrobacterium* EHA105 được hòa tan trong nước đá, sau đó khuấy nhẹ, dung dịch này được đông lạnh ngay lập tức bằng cách sử dụng nitơ lỏng. Sau đó, quá trình rã đông được thực hiện bằng cách làm ấm ở nhiệt độ 37°C trong 4 phút, 0,5mL môi trường SOC được bổ sung vào, và quá trình nuôi cấy được thực hiện ở nhiệt độ 28°C trong 1 đến 3 giờ. Tiếp theo, dịch nuôi cấy thu được được bổ sung vào môi trường thạch LB chứa 50mg/L kanamycin và 10mg/L PPT (phosphinothricin), và nuôi cấy ở vị trí thẳng đứng trong 2 ngày trong tủ nuôi cấy ở nhiệt độ 28°C để thu được *Agrobacterium* chuyển gen. Sau khi nuôi cấy dung dịch *Agrobacterium* chuyển gen, plasmid được chiết xuất và giải trình tự bằng máy giải trình tự ADN 3730 (do Life Technologies sản xuất).

Tiếp theo, cây Mù tạc đã được gieo trồng trong 2 tháng ở nhiệt độ 22°C với 24 giờ chiếu sáng và *Agrobacterium* chuyển gen đã được nuôi cấy bằng môi trường thạch LB chứa 50mg/L kanamycin và 10mg/L PPT được sử dụng để tạo ra cây Mù tạc chuyển gen.

Trước tiên, dịch nuôi cấy *Agrobacterium* có mật độ quang ở bước sóng 600nm bằng khoảng 1 được thu nhận và tạo hỗn dịch trong dung dịch chứa 5% sucroza và 0,05% Silwet. Sau đó, cây Mù tạc được nhúng vào hỗn dịch *Agrobacterium* trong vài giây để lây nhiễm các cây con. Sau giai đoạn trưởng thành của các cây con, các cây con được thu thập, và quá trình chọn lọc cây chuyển gen được thực hiện bằng cách sử dụng môi trường 1/2MS chứa 50mg/L kanamycin và 10mg/L PPT, để thu được 6 cây Mù tạc chuyển gen AR19G-166-RA

xenlobiohydrolaza, và 4 cây Mù tạc chuyển gen AR19G-166-QV xenlobiohydrolaza. Hai cây Mù tạc chuyển gen Arabi_RA4 và Arabi_QV4 được đo hoạt tính enzym xenlobiohydrolaza.

2. Chiết xuất và đo hoạt tính của protein enzym xenlobiohydrolaza được tạo ra bởi cây Mù tạc chuyển gen

Quá trình chiết xuất protein được thực hiện bằng cách nghiền 100 mg lá từ mỗi cây Mù tạc chuyển gen trong điều kiện nitơ lỏng bằng cối và chày, sau đó bổ sung dung dịch đệm axit axetics 20mM (độ pH = 5,5) chứa 1mL dung dịch chất ức chế proteaza 1% thể tích (do Sigma sản xuất-Aldrich), và trộn kỹ. Hỗn hợp thu được được chuyển vào ống dung tích 2mL và phân tách ly tâm trong 10 phút ở tốc độ 15.000 vòng/phút và nhiệt độ 4°C, và phần dịch nổi được thu gom làm dịch chiết enzym thô. Dịch chiết enzym thô được điều chế theo cùng cách thức bằng cách sử dụng cây Mù tạc thể đại được sử dụng làm mẫu đối chứng (lô không chuyển gen).

Đối với mỗi cây Mù tạc chuyển gen Arabi_RA4 và Arabi_QV4 và cây đối chứng (không chuyển gen), phân tích thẩm tách Western được thực hiện bằng cách sử dụng dịch chiết enzym thô để kiểm chứng biểu hiện gen tái tổ hợp xenlobiohydrolaza. Phân tích thẩm tách Western được thực hiện theo cùng cách thức nêu trong phần 9 của Ví dụ 1, và các kết quả được thể hiện trên Fig.17. Trên Fig.17, làn 1 là gen chỉ thị phân tử protein, làn 2 là dịch chiết enzym thô thu được từ lá của cây Mù tạc chuyển gen AR19G-166-RA (Arabi_RA4), làn 3 là dịch chiết enzym thô thu được từ lá của cây Mù tạc chuyển gen AR19G-166-QV (Arabi_QV4), và làn 4 là dịch chiết enzym thô thu được từ lá của cây Mù tạc thể đại (WT).

Dịch chiết enzym thô của cây Mù tạc chuyển gen AR19G-166-RA (Arabi_RA4) và cây Mù tạc chuyển gen AR19G-166-QV (Arabi_QV4) có các băng đậm lần lượt có trọng lượng phân tử 53,9kDa và 52,1kDa. Trọng lượng phân tử protein enzym của cây Mù tạc chuyển gen Arabi_RA4 hơi lớn hơn trọng lượng phân tử protein enzym của cây Mù tạc chuyển gen Arabi_QV4 (làn 2 và làn 3 của Fig.17). Trọng lượng phân tử của protein khi được biểu hiện trong *E. coli* bằng 46,7kDa, điều này chứng tỏ rằng trọng lượng phân tử của protein mã hóa bởi gen AR19G-166-RA và AR19G-166-QV được biểu hiện bằng vật chủ là cây Mù tạc đã lần lượt tăng 7,2kDa và 5,4kDa. Trọng lượng phân tử của protein tăng là do phản ứng glycosyl hóa, và sự khác biệt về trọng lượng phân tử của các protein enzym ở

các cây Mù tạc chuyển gen Arabi_RA4 và Arabi_QV4 là do mức độ glycosyl hóa khác nhau.

3. Mức độ phụ thuộc vào nhiệt độ của hoạt tính xenlobiohydrolaza của cây Mù tạc biểu hiện protein AR19G-166

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA của protein enzym AR19G-166-RA và protein enzym AR19G-166-QV được tạo ra bởi các cây Mù tạc chuyển gen nêu trên được đánh giá. Trong phép đo này, dịch chiết enzym thô thu được ở phần 2 nêu trên được sử dụng, và quá trình đánh giá được thực hiện theo cùng cách thức nêu trong phần 10 của ví dụ 1 để đo lượng đường khử được tạo ra bởi quá trình thủy phân enzym.

Hoạt tính thủy phân PSA đo được (lượng đường khử) ở các nhiệt độ đo khác nhau được thể hiện trên Fig.18. Trên Fig.18, hoạt tính xenlobiohydrolaza của cây Mù tạc thể đại (WT) được đo và vẽ đồ thị để so sánh.

Đối với protein enzym AR19G-166-RA được biểu hiện bằng vật chủ là cây Mù tạc, nhiệt độ ở đó hoạt tính enzym mạnh nhất là bằng 100°C, là giới hạn trên của phép đo, và giả định rằng nhiệt độ tối ưu bằng 100°C hoặc cao hơn. Mặt khác, nhiệt độ tối ưu của protein enzym AR19G-166-QV bằng 90°C. Theo phương thức này, khi gen AR19G-166 được biểu hiện trong hệ nhân gen của cây Mù tạc, thì trọng lượng phân tử tăng khoảng 10%, và nhiệt độ tối ưu tăng từ 15°C đến 30°C. AR19G-166-RA được biểu hiện bằng cây Mù tạc có độ ổn định nhiệt rất cao, và có nhiệt độ tối ưu cao hơn nhiệt độ tối ưu của xenlobiohydrolaza chịu nhiệt có nguồn gốc từ nấm sợi ưa nhiệt 30°C hoặc cao hơn.

Các kết quả này cho thấy rằng bằng cách biểu hiện trong hệ nhân gen của thực vật hoặc sinh vật nhân thật như nấm sợi, thì có thể thu được xenlobiohydrolaza chịu nhiệt có hoạt tính mạnh ngay ở cả nhiệt độ 80°C hoặc cao hơn có thể thu được. Lignoxenuloza có thể được thủy phân ở nhiệt độ 80°C hoặc cao hơn bằng xenlobiohydrolaza chịu nhiệt có hoạt tính mạnh ở nhiệt độ nằm trong khoảng từ 80°C đến 100°C theo sáng chế kết hợp với endoglucanaza, xylanaza hoặc β -glucosidaza chịu nhiệt tương đồng hoặc các enzym tương tự.

Ví dụ 5

Xạ khuẩn thuộc chi Streptomyces đã được sử dụng nhiều trong công nghiệp để sản

xuất các kháng sinh và hoạt chất sinh học hữu ích. Hệ thống biểu hiện trên quy mô lớn cho các gen ngoại sinh bằng các xạ khuẩn này hiện đang được phát triển, và được liệt kê trong Tài liệu Sáng chế 3, 4 và 5, và Tài liệu phi Sáng chế 9, 10 và 11. Cụ thể, do xạ khuẩn có bộ gen với hàm lượng GC cao, nên quá trình biểu hiện của gen có hàm lượng GC cao, mà gặp trở ngại khi biểu hiện trong *E. coli*, có thể dễ dàng được thực hiện (Tài liệu phi Sáng chế 11), và mức độ biểu hiện protein ngoại sinh có thể bằng 40% dịch chiết không chứa tế bào xạ khuẩn (Tài liệu phi Sáng chế 9). Quá trình biểu hiện protein ở xạ khuẩn đã được chèn gen AR19G-166-RA là thích hợp để sản xuất xenlobiohydrolaza chịu nhiệt theo sáng chế trên quy mô lớn và với chi phí thấp.

1. Quy trình sản xuất xạ khuẩn chứa gen AR19G-166-RA

Bằng cách sử dụng các gen AR19G-166-RA được tách dòng với khuôn là vector pET101/D-TOPO (do Life Technologies sản xuất), các gen này được chuyển nạp vào vector biểu hiện xạ khuẩn pHSA81 (Tài liệu Sáng chế 4) bằng phương pháp PCR, sau đó chuyển nhiễm vào chủng *Streptomyces lividans* TK24. Chủng *Streptomyces lividans* TK24 được John Innes Centre cung cấp (Norwich Research Park, Norwich, NR4 7UH, UK).

Quá trình chuyển nhiễm nêu trên được thực hiện bằng phương pháp dung hợp thể nguyên sinh bằng polyetylen được mô tả trong tài liệu “*Genetic manipulation of Streptomyces: a laboratory manual*”. Sau khi chuyển nhiễm, các dòng biểu hiện được chọn lọc bằng phương pháp PCR trực tiếp từ khuẩn lạc, tiếp theo nuôi cấy lắc trên môi trường YEME (chiết xuất nấm men: 0,3%, pepton tinh khiết: 0,5%, chiết xuất đại mạch: 0,3%, glucoza: 1%, sucroza: 34%, $MgCl_2$: 5 mM, glyxin: 0,5%), plasmit tái tổ hợp được chiết xuất, và giải trình tự bằng máy giải trình tự ADN 3730 (do Life Technologies sản xuất).

2. Quá trình biểu hiện protein AR19G-166-RA ở xạ khuẩn

Thế biến nạp thu được được cấy chuyển vào môi trường YEME chứa 5 μ g/mL thiostrepton, và nuôi cấy lắc ở nhiệt độ 28°C trong 5 ngày, sau đó thu nhận bằng phân tách ly tâm. Sau khi rửa xạ khuẩn này bằng dung dịch đệm Tris-HCl 50mM (độ pH = 8,0), dung dịch đệm Tris-HCl này được bổ sung vào với thể tích bằng 1/10 thể tích dịch nuôi cấy để tạo hỗn dịch xạ khuẩn. Sau đó, quy trình phân cắt trong 30 giây bằng thiết bị BioRuptor UCD-200T (do Cosmo Bio Co. Ltd. sản xuất) và 30 giây ngừng được lặp lại 10 lần, và sau đó ly tâm, phân tích điện di trên gel SDS-PAGE thực hiện bằng cách sử dụng phần dịch nổi

(dịch chiết không chứa tế bào). Các kết quả được thể hiện trên Fig.19. Làn 1 trên Fig.19 là gen chỉ thị phân tử protein, làn 2 là protein tinh khiết AR19G-166-RA (được thể hiện bằng mũi tên) được biểu hiện trong *E. coli*, và làn 3 là dịch chiết không chứa tế bào của xạ khuẩn *Streptomyces lividans* tái tổ hợp gen AR19G-166-RA. Trên Fig.19, protein đích đã được biểu hiện mạnh ở trọng lượng phân tử mong muốn bằng 46,7kDa (làn 3 trên Fig.19).

3. Đo hoạt tính của protein AR19G-166-RA được biểu hiện bằng xạ khuẩn

Hoạt tính xenlobiohydrolaza được đo bằng cách sử dụng dịch chiết không chứa tế bào xạ khuẩn tái tổ hợp gen AR19G-166-RA. Hoạt tính được đo bằng cách cho hỗn hợp chứa 50 μ L mẫu dịch chiết không chứa tế bào và 50 μ L dung dịch đệm axit axetic 200mM (độ pH = 5,5) chứa 1% khối lượng PSA phản ứng ở nhiệt độ 30 đến 100°C trong 20 phút.

Dung dịch cơ chất và dung dịch enzyme được giữ riêng biệt ở nhiệt độ phản ứng trong 5 phút, sau đó trộn với nhau để hoạt hóa phản ứng. Trong quá trình phản ứng, toàn bộ các dung dịch được khuấy bằng máy khuấy nhiệt do Eppendorf sản xuất (1.400 vòng/phút) để tránh kết tủa cơ chất không tan. Trong toàn bộ các phép đo, dung dịch hỗn hợp được điều chế bằng cách cho 50 μ L dung dịch đệm axit axetic 200mM (độ pH = 5,5) chứa 1% khối lượng PSA phản ứng ở mỗi nhiệt độ, sau đó bổ sung 50 μ L mẫu dịch chiết không chứa tế bào sau khi dừng phản ứng được sử dụng làm mẫu đối chứng. Sau khi kết thúc phản ứng, thuốc thử axit 3,5-đinitrosalixylic (dung dịch DNS) được bổ sung vào ở cùng thể tích, quá trình xử lý nhiệt được thực hiện ở nhiệt độ 100°C trong 5 phút, sau đó làm nguội trong 5 phút, hỗn hợp này được ly tâm để thu được phần dịch nổi. Độ hấp thụ của phần dịch nổi ở bước sóng 540nm được đo bằng máy quang phổ, hàm lượng đường khử trong phần dịch nổi được đo bằng đường chuẩn glucoza, và độ sai lệch từ trị số của mẫu đối chứng được sử dụng để đo lượng đường khử được tạo ra bởi quá trình thủy phân enzym. Mỗi phép đo được thực hiện với ba mẫu độc lập, và trị số trung bình và độ lệch chuẩn được xác định. Các kết quả được thể hiện trên Fig.20. Trên Fig.20, ba phép đo được thực hiện cho mỗi điểm dữ liệu, và trị số trung bình và độ lệch chuẩn được vẽ đồ thị theo nhiệt độ.

Hoạt tính enzym để tạo ra 1 μ mol đường khử trong 1 phút được tính là 1U, và trị số thu được bằng cách chia hoạt tính này cho khối lượng protein enzym được tính là hoạt tính riêng (U/mg), hoạt tính riêng ở nhiệt độ 80°C bằng 3,02. Như nêu trên, xenlobiohydrolaza AR19G-166-RA cũng được biểu hiện mạnh và có hoạt tính ở xạ khuẩn, và xạ khuẩn có thể

được sử dụng làm vật chủ chuyển gen theo sáng chế.

Ví dụ 6

Trong số các xenlulaza, có các enzym không chỉ chứa vùng xúc tác thủy phân xenluloza, mà còn chứa đoạn gắn kết xenluloza, cũng được gọi là đoạn gắn kết hydratcarbon (CBM – carbohydrate binding module). CBM bản thân nó không chỉ có hoạt tính phân giải mà còn có chức năng gắn kết xenluloza. Các chức năng đã biết của CBM bao gồm tăng mật độ vùng xúc tác đặc trưng cho CBM ở các vùng lân cận cơ chất bằng cách hấp thụ cơ chất không tan, nhờ đó tăng tốc độ phân giải xenluloza, và phân cắt liên kết hydro giữa các mạch xenluloza nhờ gắn kết CBM, nhờ đó phân giải các cấu trúc tinh thể (Tài liệu phi Sáng chế 12 và 13). Ngoài ra, khi CBM được loại bỏ ra khỏi CBH phân giải xenluloza tinh thể, thì mặc dù hoạt tính phân giải cơ chất không tan không thay đổi, hoạt tính phân giải và ái lực gắn kết với xenluloza tinh thể giảm mạnh, do đó CBM là vùng cần thiết cho các enzym phân giải xenluloza tinh thể (Tài liệu phi Sáng chế 14). Mặt khác, khi bổ sung CBM vào endoglucanaza mà ban đầu không chứa CBM, hoạt tính phân giải và ái lực gắn kết với xenluloza tinh thể có thể được cải thiện (Tài liệu phi Sáng chế 5).

Để tiếp tục cải thiện hoạt tính phân giải xenluloza của xenlobiohydrolaza chịu nhiệt theo sáng chế, thì gen chứa CBM được bổ sung vào AR19G-166-RA được tổng hợp, và ảnh hưởng của CBM đến hoạt tính phân giải xenluloza của protein enzym được mã hóa được đánh giá.

1. Quy trình sản xuất gen AR19G-166-RA được bổ sung CBM

Trình tự CBM và trình tự gốc liên kết được sử dụng là trình tự CBM và trình tự gốc liên kết của khung đọc mã mở OJ1-1 được thể hiện trong phần 5 của Ví dụ 1. Trình tự gốc liên kết của khung đọc mã mở OJ1-1 được bổ sung vào đầu tận cùng C của gen AR19G-166-R, bổ sung trình tự CBM3 của khung đọc mã mở OJ1-1 vào đầu tận cùng C. Trình tự nucleotit của toàn bộ gen AR19G-166-RA được bổ sung CBM được tối ưu hóa về bộ ba mã hóa có tần xuất sử dụng cao ở *E. coli* và tổng hợp nhân tạo. Trình tự axit amin của gen CBM3 và gốc liên kết được bổ sung như nêu trong SEQ ID NO.17 và trình tự nucleotit mã hóa nó như nêu trong SEQ ID NO.18, trong khi đó trình tự axit amin của toàn bộ gen AR19G-166-RA được bổ sung CBM được tối ưu hóa về bộ ba mã hóa có tần xuất sử dụng cao ở *E. coli* như nêu trong SEQ ID NO.19 và trình tự nucleotit mã hóa nó như nêu trong

SEQ ID NO.20. 3 nucleotit TAA từ vị trí số 808 đến vị trí số 810 của SEQ ID NO.18, và 3 nucleotit TAA từ vị trí số 2092 đến vị trí số 2094 của SEQ ID NO.20 đều là các bộ ba mã hóa kết thúc.

Gen được tổng hợp nêu trên được kết hợp vào vector biểu hiện pLEAD (do NIPPON GENE Co., Ltd. sản xuất), biến nạp vào chủng JM109, và giải trình tự bằng máy giải trình tự ADN 3730 do Life Technologies sản xuất.

2. Quá trình biểu hiện protein AR19G-166-RA được bổ sung CBM

Sau khi giải trình tự, sản phẩm tách dòng *E. coli* chứa plasmit có gen đích được cấy chuyển vào 5mL môi trường LB chứa 100mg/L ampicilin, và nuôi cấy lắc ở nhiệt độ 37°C trong 20 giờ. Sau khi nuôi cấy, *E.coli* được thu nhận bằng phân tách ly tâm, dung dịch đệm Tris-HCl 50mM (độ pH = 8) được bổ sung vào với thể tích bằng 1/10 thể tích dịch nuôi cấy để tạo hỗn dịch vi khuẩn. Sau đó, quy trình phân cắt trong 30 giây bằng thiết bị BioRuptor UCD-200T (do Cosmo Bio Co. Ltd. sản xuất) và 30 giây ngừng được lặp lại 10 lần, sau đó ly tâm để thu được phần dịch nổi (dịch chiết thô *E. coli*). Một phần dịch chiết thô *E. coli* được điện di trên gel SDS-PAGE, và protein đích được kiểm chứng là được biểu hiện ở trọng lượng phân tử mong muốn. Sau khi kiểm chứng biểu hiện protein, dịch nuôi cấy *E. coli* qua đêm ở nhiệt độ 37°C được sử dụng làm dịch nuôi cấy sơ bộ, và quá trình nuôi cấy chính được thực hiện trong môi trường LB gấp 100 lần chứa 100mg/L ampicilin.

3. Quá trình tinh chế protein AR19G-166-RA được bổ sung CBM

Sau khi nuôi cấy, *E. coli* được thu nhận bằng cách ly tâm, dung dịch đệm Tris-HCl 50mM (độ pH = 8) được bổ sung vào với thể tích bằng 1/10 thể tích dịch nuôi cấy để tạo hỗn dịch vi khuẩn. Sau đó, quy trình phân cắt trong 5 phút bằng thiết bị Astrason 3000 (do Misonix, Inc. sản xuất) và 5 phút ngừng được lặp lại từ 7 đến 8 lần để thu được dịch chiết thô *E. coli* tái tổ hợp gen chứa protein đích. Dịch chiết thô *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), và bằng hệ thống sắc ký lỏng áp suất từ trung bình đến cao AKTA (do

GE Healthcare sản xuất), các protein được phân đoạn ở gradien nồng độ nằm trong khoảng từ 0% đến 50% bằng cách sử dụng dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) chứa 1M of NaCl. Các phân đoạn có hoạt tính xenlobiohydrolaza được trộn với nhau, sau đó trao đổi dung dịch và cô với dung dịch đệm axit phosphoric 1mM (độ pH = 6,8) bằng màng siêu lọc ly tâm VIVASPIN 20 (do Sartorius stedim sản xuất), sau đó nạp vào cột hydroapatit CHT5-1 (do Bio-Rad sản xuất) được làm cân bằng với cùng dung dịch đệm để phân đoạn các protein ở gradien nồng độ nằm trong khoảng từ 0% đến 100% bằng cách sử dụng dung dịch đệm axit phosphoric 400mM (độ pH = 6,8). Các phân đoạn có hoạt tính xenlobiohydrolaza được trộn với nhau, và cô bằng màng siêu lọc ly tâm VIVASPIN 20 cho đến khi thể tích dung dịch bằng khoảng 8 mL. 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 150 mM of NaCl, và được phân đoạn bằng cách bổ sung lượng cùng dung dịch đệm bằng 1 đến 1,5 lần thể tích cột ở tốc độ dòng bằng 2 đến 3mL/phút. Các phân đoạn có hoạt tính xenlobiohydrolaza được trộn với nhau, sau đó trao đổi dung dịch và cô với dung dịch đệm axit phosphoric 50mM (độ pH = 6) bằng màng siêu lọc ly tâm VIVASPIN 20, sau đó nạp vào cột trao đổi ion HiTrap SP HP (do GE Healthcare sản xuất) được làm cân bằng với cùng dung dịch đệm để phân đoạn các protein ở gradien nồng độ nằm trong khoảng từ 0% đến 50% bằng cách sử dụng dung dịch đệm axit phosphoric 50mM (độ pH = 6) chứa 1M NaCl. Các phân đoạn có hoạt tính xenlobiohydrolaza được trộn với nhau, sau đó trao đổi với dung dịch đệm Tris-HCl 50mM (độ pH = 8,0) và cô để thu được enzym tinh khiết có nồng độ cuối cùng bằng khoảng 1mg/mL.

Dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và protein tinh khiết AR19G-166-RA được bổ sung CBM được kiểm chứng bằng phân tích điện di trên gel SDS-PAGE và phân tích thẩm tách Western. Chỉ khác là sử dụng gel không chứa thuốc nhuộm Mini Protean TGX (do Bio-Rad sản xuất), phân tích điện di trên gel SDS-PAGE và phân tích thẩm tách Western được thực hiện theo cùng cách thức nêu trong phần 9 của ví dụ 1, và 10 μ L dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen và 0,5 μ g enzym tinh khiết lần lượt được chạy điện di. Trong phân tích điện di trên gel SDS-PAGE, các băng protein được phát hiện bằng thiết bị Gel Doc EZ (do Bio-Rad sản xuất).

Fig.21A và Fig.21B thể hiện các kết quả phân tích điện di trên gel SDS-PAGE

(Fig.21A) và phân tích thẩm tách Western (Fig.21B) của protein enzym thu được bằng cách biểu hiện protein AR19G-166-RA được bổ sung CBM trong *E. coli*. Trên Fig.21A và Fig.21B, làn 1 là gen chỉ thị phân tử protein, làn 2 là dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen, làn 3 protein tinh khiết AR19G-166-RA được bổ sung CBM, và làn 4 là mẫu điện di của protein enzym xenlobiohydrolaza được tinh chế trong phần 9 của ví dụ 1.

Protein AR19G-166-RA được bổ sung CBM được biểu hiện trong *E. coli*. Trong phân tích điện di trên gel SDS-PAGE của dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen, băng có trọng lượng phân tử khoảng 74,5kDa được xác định từ trình tự axit amin (SEQ ID NO.19) (làn 2 của Fig.21A). Khi protein này được tinh chế, băng đậm tương ứng với băng nêu trên được quan sát thấy, và băng mờ có trọng lượng phân tử khoảng 60kDa cũng được quan sát thấy (làn 3 của Fig.21A). Phân tích thẩm tách Western được thực hiện bằng cách sử dụng kháng thể kháng polypeptit chứa 20 gốc axit amin từ vị trí số 384 đến 403 của AR19G-166, và trong dịch nổi nuôi cấy đồng nhất *E. coli* tái tổ hợp gen (làn 2 của Fig.21B), băng protein enzym có trọng lượng phân tử khoảng 74,5kDa được phát hiện. Ngoài ra, băng có trọng lượng phân tử khoảng 60kDa cũng được phát hiện, tuy nhiên điều này có thể do trộn protein enzym dẫn đến phân giải một phần protein enzym có trong tế bào *E. coli*, hoặc protein enzym không được biểu hiện hoàn toàn. Theo cùng cách thức, trong enzym tinh khiết (làn 3 của Fig.21B), băng protein enzym đậm có trọng lượng phân tử khoảng 74,5kDa được phát hiện và băng mờ có trọng lượng phân tử khoảng 60kDa cũng được phát hiện, nhưng thành phần protein chính được kiểm chứng là protein enzym nêu trên.

4. Hoạt tính của protein AR19G-166-RA được bổ sung CBM

Mức độ phụ thuộc vào nhiệt độ của hoạt tính thủy phân PSA và Avicel của protein AR19G-166-RA được bổ sung CBM được đánh giá. Trong phép đo này, enzym tinh khiết nêu trên (nồng độ cuối cùng bằng khoảng 1 mg/mL) được sử dụng.

Hoạt tính thủy phân PSA và Avicel của enzym tinh khiết được đo theo cùng cách thức như nêu trong phần 10 của ví dụ 1, chỉ khác là cho hỗn hợp chứa 100 μ L dung dịch nước cơ chất 1% khối lượng, 50 μ L dung dịch đệm axit axetic 200mM (độ pH = 5,5), 40 μ L nước tinh khiết hoặc dung dịch nước CaCl₂ 10mM, và 10 μ L enzym tinh khiết phản ứng trong 20 phút ở nhiệt độ 50°C, 60°C, 70°C, 75°C, 80°C, 85°C, 90°C hoặc 99°C, và lượng đường khử được tạo ra bởi quá trình thủy phân enzym được đo để đánh giá hoạt tính thủy

phân (U/mg).

Hoạt tính thủy phân PSA và Avicel của protein AR19G-166-RA được bổ sung CBM ở các nhiệt độ khác nhau lần lượt được thể hiện trên Fig.22A và Fig.22B. Các kết quả cho thấy CBM không ảnh hưởng đến hoạt tính thủy phân PSA khi có hoặc không có ion canxi, và thực tế hoạt tính thủy phân giảm tương đối so với dòng protein AR19G-166-RA (Fig.22A). Tuy nhiên, khi có mặt ion canxi hoạt tính phân giải Avicel đã tăng mạnh ở nhiệt độ nằm trong khoảng từ 5°C đến 85°C với sự có mặt của ion canxi (Fig.22B). Mặt khác, khi không có ion canxi, thì hoạt tính phân giải Avicel không tăng ở nhiệt độ 70°C hoặc cao hơn. Điều này chứng tỏ rằng ion canxi có tác dụng làm tăng độ ổn định nhiệt của CBM.

Các kết quả này cho thấy rằng bằng cách bổ sung CBM vào xenlobiohydrolaza chịu nhiệt theo sáng chế, hoạt tính thủy phân xenluloza tinh thể có thể tăng, ít nhất với sự có mặt của ion canxi.

Khả năng ứng dụng công nghiệp

Do có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5, nên xenlobiohydrolaza chịu nhiệt theo sáng chế thích hợp để thủy phân sinh khối chứa xenluloza ở nhiệt độ 75°C hoặc cao hơn. Do đó, xenlobiohydrolaza chịu nhiệt theo sáng chế, polynucleotit mã hóa enzym này, vectơ biểu hiện chứa polynucleotit này, và thể biến nạp chứa vectơ này có thể được ứng dụng trong lĩnh vực sản xuất năng lượng từ sinh khối xenluloza.

YÊU CẦU BẢO HỘ

1. Xenlobiohydrolaza chịu nhiệt chứa vùng xúc tác xenlobiohydrolaza bao gồm:

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

(B) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

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

(D) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(E) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.5;

(F) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(G) polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.7; hoặc

(H) polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 80% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

2. Xenlobiohydrolaza chịu nhiệt theo điểm 1, trong đó enzym này còn chứa đoạn gắn kết xenluloza.

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

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

(b) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ

ID NO.1, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

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

(d) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.3, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(e) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.5;

(f) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.5, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5;

(g) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin như nêu trong SEQ ID NO.7;

(h) trình tự nucleotit mã hóa polypeptit chứa trình tự axit amin có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự axit amin như nêu trong SEQ ID NO.7, và có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5; hoặc

(i) trình tự nucleotit có độ tương đồng nằm trong khoảng từ 90% đến nhỏ hơn 100% so với trình tự nucleotit như nêu trong SEQ ID NO.2, SEQ ID NO.4, SEQ ID NO.6, hoặc SEQ ID NO.8, và mã hóa polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5.

4. Polynucleotit theo điểm 3, trong đó polynucleotit này còn chứa miền mã hóa đoạn gắn kết xenluloza.

5. Vector biểu hiện chứa polynucleotit theo điểm 3 hoặc 4, và có khả năng biểu hiện

polypeptit có hoạt tính xenlobiohydrolaza ít nhất ở điều kiện nhiệt độ bằng 75°C và độ pH bằng 5,5 ở tế bào vật chủ.

6. Thẻ biến nạp chứa vector theo điểm 5.

7. Thẻ biến nạp theo điểm 6, trong đó thẻ biến nạp này là vi sinh vật nhân sơ.

8. Thẻ biến nạp theo điểm 6, trong đó thẻ biến nạp này là vi sinh vật nhân thật.

9. Thẻ biến nạp theo điểm 6, trong đó thẻ biến nạp này là thực vật.

10. Phương pháp sản xuất xenlobiohydrolaza chịu nhiệt bao gồm bước sản sinh xenlobiohydrolaza chịu nhiệt trong thẻ biến nạp theo điểm bất kỳ trong số các điểm từ 6 đến 9.

11. Hỗn hợp xenlulaza chứa xenlobiohydrolaza chịu nhiệt theo điểm 1, xenlobiohydrolaza chịu nhiệt theo điểm 2, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo điểm 10, và ít nhất một hoặc nhiều loại xenlulaza khác.

12. Hỗn hợp xenlulaza theo điểm 11, trong đó ít nhất một hoặc nhiều loại xenlulaza khác được chọn từ nhóm bao gồm hemixenlulaza và endoglucanaza.

13. Phương pháp sản xuất sản phẩm phân giải xenluloza bao gồm bước tạo ra sản phẩm phân giải xenluloza bằng cách cho nguyên liệu chứa xenluloza tiếp xúc với xenlobiohydrolaza chịu nhiệt theo điểm 1, xenlobiohydrolaza chịu nhiệt theo điểm 2, thẻ biến nạp theo điểm bất kỳ trong số các điểm từ 6 đến 9, hoặc xenlobiohydrolaza chịu nhiệt được sản xuất bằng phương pháp theo điểm 10.

14. Phương pháp theo điểm 13, trong đó nguyên liệu chứa xenluloza còn được cho tiếp xúc với ít nhất một hoặc nhiều loại xenlulaza khác.

15. Phương pháp theo điểm 14, trong đó ít nhất một hoặc nhiều loại xenlulaza khác được chọn từ nhóm bao gồm hemixenlulaza và endoglucanaza.

16. Phương pháp sản xuất polynucleotit mã hóa xenlobiohydrolaza chịu nhiệt bao gồm bước thực hiện phản ứng PCR bằng cách sử dụng khuôn là ADN thu được từ sinh vật hoặc sản phẩm phiên mã ngược của ARN thu được từ sinh vật, với đoạn mồi xuôi chứa trình tự

nucleotit như nêu trong SEQ ID NO.12 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12, và đoạn môi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13 hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13, và thu nhận sản phẩm khuếch đại là polynucleotit chứa trình tự nucleotit mã hóa xenlobiohydrolaza chịu nhiệt.

17. Cặp đoạn môi bao gồm đoạn môi xuôi chứa trình tự nucleotit như nêu trong SEQ ID NO.12, hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.12 và đoạn môi ngược chứa trình tự nucleotit như nêu trong SEQ ID NO.13, hoặc trình tự nucleotit được tạo ra bằng cách bổ sung một hoặc nhiều nucleotit vào đầu 5' của trình tự nucleotit như nêu trong SEQ ID NO.13.

DANH MỤC TRÌNH TỰ

<110> HONDA MOTOR CO., LTD.
 <110> KAZUSA DNA RESEARCH INSTITUTE
 <120> XENLOBIOHYDROLAZA CHỊU NHIỆT, POLYNUCLEOTIT MÃ HÓA ENZYM NÀY, PHƯƠNG PHÁP SẢN XUẤT POLYNUCLEOTIT NÀY, VECTO BIỂU HIỆN ENZYM NÀY, THỂ BIẾN NẠP CHỨA VECTO NÀY, PHƯƠNG PHÁP SẢN XUẤT ENZYM NÀY, HỖN HỢP XENLULAZA CHỨA ENZYM NÀY, PHƯƠNG PHÁP SẢN XUẤT SẢN PHẨM PHÂN GIẢI XENLULOZA VÀ CẶP ĐOẠN MÔI
 <130> H112165802
 <160> 20
 <170> PatentIn version 3.5
 <210> 1
 <211> 427
 <212> PRT
 <213> Trình tự chưa biết
 <220> Trình tự axit amin AR19G-166-RA
 <223> AR19G-166-RA
 <400> 1
 Leu Asp Asn Pro Phe Ile Gly Ala Ile Gly Tyr Val Asn Pro Asp Trp
 1 5 10 15
 Ala Thr Asn Val Ile Ser Gln Ala Asn Gln Thr Ala Asp Pro Thr Leu
 20 25 30
 Ala Ala Gln Met Arg Lys Val Ala Thr Tyr Ser Thr Ala Val Trp Leu
 35 40 45
 Asp Arg Ile Ala Ala Ile Thr Ala Gly Arg Gly Leu Arg Gly His Leu
 50 55 60
 Asp Glu Ala Leu Arg Gln Met Gln Gln Ala Gly Gln Pro Val Val Ile
 65 70 75 80
 Thr Leu Val Ile Tyr Asp Leu Pro Asn Arg Asp Cys Ser Ala Ala Ala
 85 90 95
 Ser Asn Gly Glu Leu Leu Val Ala Gln Asn Gly Leu Ala Arg Tyr Lys
 100 105 110
 Ala Glu Phe Ile Asp Pro Ile Val Ala Ile Leu Ser Asp Pro Arg Tyr
 115 120 125
 Ala Gly Leu Arg Ile Val Thr Ile Ile Glu Pro Asp Ser Leu Pro Asn
 130 135 140
 Leu Val Thr Asn Leu Ser Ile Pro Ala Cys Ala Glu Ala Gln Asn Ala
 145 150 155 160
 Tyr Ile Glu Gly Ile Arg Tyr Ala Val Asn Arg Leu Arg Thr Ile Pro
 165 170 175
 Asn Val Tyr Ile Tyr Leu Asp Ile Ala His Ser Gly Trp Leu Gly Trp
 180 185 190
 Asp Asn Asn Phe Asn Gly Ala Val Asn Leu Tyr Thr Gln Val Val Gln
 195 200 205
 Gly Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn Val
 210 215 220
 Ala Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu
 225 230 235 240
 Thr Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp Asn
 245 250 255
 Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala Phe
 260 265 270
 Ile Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr Ser
 275 280 285
 Arg Asn Gly Trp Gly Gly Cys Ser Tyr Gly Arg Cys Arg Pro Thr Gly
 290 295 300
 Pro Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser Arg
 305 310 315 320
 Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly
 325 330 335
 Ile Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Ala Tyr
 340 345 350
 Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly
 355 360 365
 Ile Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys Asp
 370 375 380
 Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu
 385 390 395 400
 Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu Ile
 405 410 415
 Leu Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro
 420 425
 <210> 2
 <211> 1284
 <212> ADN
 <213> Trình tự chưa biết
 <220> Trình tự nucleotit AR19G-166-RA
 <223> AR19G-166-RA

```

<400> 2
ttggacaatc cattcatcgg agccatcgga tacgtgaatc cggactgggc aaccaatgtg 60
atcagccaag cgaaccaaac ggctgatcca accttggcgg ctcaaatgcg taagggtggcc 120
acctactcca cagctgtctg gttggatcgt atcgccgcca tcaccgctgg ccgcggattg 180
cgcggggcatt tggatgaagc actacgcca atgcagcaag ctggccagcc ggtgtgtatc 240
acccttgtga tctatgatct gccaaatcga gattgtctct ctgctgcctc caatggcgaa 300
ttactggctg cccagaatgg actggcccg tcacaaagcg agttcatcga tcccatcgta 360
gccattctct cagatccccg atatgccggg ctacgcatcg tcaccatcat cgaaccggac 420
tccttaccga acctggtcac gaacctcagc atcccggcat gcgcagaagc tcagaatgca 480
tatatcgaag ggatccgcta tgcgggtaac cggctgcgga caattcccaa cgtctacatc 540
tatctggata tcgcccactc aggtctggtg ggctgggata ataactttaa tggggcagta 600
aatctctaca cccaagtgtg gcaagggaat gatcaagggt ttaacagtat cgaatggattc 660
atcaccaacg ttgccaaacta taccocctc gaagaaccct atttgcccga tcccaacctg 720
accatcgctg ggcaaccctg tgcgtctgcc agcttctacg aatggaacct ctactttgat 780
gagctggatt acgctctggc tctgcgcaac gcttttatcg gacgaggctt cccagcacc 840
attgggatgc tcatcgatac cagccgaaac ggggtgggag gatgcagcta tgggcgatgc 900
aggcccacgg gtcccagttc agataccagc agcgtgaatg cctacgtgga cggctcgcga 960
gtagaccgac gctatcatcg gggcaactgg tgtaaccagg cgggtgggat cggcgagcgc 1020
cctcaggccg caccgcggtc cggatcgcac gcctacgtgt gggtgaaacc acctggggag 1080
tcgatgggg tcagccaacc cgggatcgtc gatcccgacg atcccaacaa gaagttcgat 1140
cctatgtgtg atcccaacgg acagagccgg tataactccg catacccaac tggggctctg 1200
cccaacgcgc cccacgctgg cgggtgggtt cgcagcagat tcgaaatcct ggtgcggaac 1260
gcctatccgc cgatccaacc ctaa 1284

<210> 3
<211> 427
<212> PRT
<213> Trình tự chưa biết
<220> Trình tự axit amin AR19G-166-QV
<223> AR19G-166-QV
<400> 3
Leu Asp Asn Pro Phe Ile Gly Ala Ile Gly Tyr Val Asn Pro Asp Trp
1 5 10 15
Ala Thr Asn Val Ile Ser Gln Ala Asn Gln Thr Ala Asp Pro Thr Leu
20 25 30
Ala Ala Gln Met Arg Lys Val Ala Thr Tyr Ser Thr Ala Val Trp Leu
35 40 45
Asp Arg Ile Ala Ala Ile Thr Ala Gly Arg Gly Leu Arg Gly His Leu
50 55 60
Asp Glu Ala Leu Arg Gln Met Gln Gln Ala Gly Gln Pro Val Val Ile
65 70 75 80
Thr Leu Val Ile Tyr Asp Leu Pro Asn Arg Asp Cys Ser Ala Ala Ala
85 90 95
Ser Asn Gly Glu Leu Leu Val Ala Gln Asn Gly Leu Ala Arg Tyr Lys
100 105 110
Ala Glu Phe Ile Asp Pro Ile Val Ala Ile Leu Ser Asp Pro Arg Tyr
115 120 125
Ala Gly Leu Arg Ile Val Thr Ile Ile Glu Pro Asp Ser Leu Pro Asn
130 135 140
Leu Val Thr Asn Leu Ser Ile Pro Ala Cys Ala Glu Ala Gln Asn Ala
145 150 155 160
Tyr Ile Glu Gly Ile Arg Tyr Ala Val Asn Arg Leu Arg Thr Ile Pro
165 170 175
Asn Val Tyr Ile Tyr Leu Asp Ile Ala His Ser Gly Trp Leu Gly Trp
180 185 190
Asp Asn Asn Phe Asn Gly Ala Val Asn Leu Tyr Thr Gln Val Val Gln
195 200 205
Gly Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn Val
210 215 220
Ala Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu
225 230 235 240
Thr Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp Asn
245 250 255
Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala Phe
260 265 270
Ile Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr Ser
275 280 285
Arg Asn Gly Trp Gly Gly Cys Ser Tyr Gly Gln Cys Arg Pro Thr Gly
290 295 300
Pro Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser Arg
305 310 315 320
Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly
325 330 335
Ile Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Val Tyr
340 345 350
Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly
355 360 365
Ile Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys Asp
370 375 380
Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu

```


385 390 395 400
 Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu Ile
 405 410 415
 Leu Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro
 420 425
 <210> 4
 <211> 1284
 <212> ADN
 <213> Trình tự chưa biết
 <220> Trình tự nucleotit AR19G-166-QV
 <223> AR19G-166-QV
 <400> 4
 ttggacaatc cattcatcgg agccatcgga tacgtgaatc cggactgggc aaccaatgtg 60
 atcagccaag cgaaccaaac ggctgatcca accttggcgg ctcaaatgcg taaggtggcc 120
 acctactcca cagctgtctg gttggatcgt atcgccgccca tcaccgctgg ccgcgggattg 180
 cgcggggcatt tggatgaagc actacgccaa atgcagcaag ctggccagcg ggttgtgatc 240
 acccttgtga tctatgctat gccaaatcga gattgtctcg ctgctgcctc caatggcgaa 300
 ttacttggtcg cccagaatgg actggcccg ctaaaagcgg agttcatcga tcccatcgta 360
 gccattctct cagatccccc atatggccgg ctacgcacg tcaccatcat cgaaccggac 420
 tcctttaccca acctgggtcac gaacctcagc atccccggcat gcgcagaagc tcagaatgca 480
 tctatcgaag ggatccgcta tgcggtgaac cggctgcgga caattcccaa cgtctacatc 540
 tatctggata tcgcccactc aggctgggtt ggctgggata ataactttaa tggggcagta 600
 aatctctaca cccaagttgt gcaaggaatg gatcaagggt ttaacagtat cgatggattc 660
 atcaccaacg ttgccaacta tccccccctc gaagaaccct atttggccga tcccaacctg 720
 accatcgctg ggcaaccgtg tcgctctgcc agcttctacg aatgggaacc cgtcttggat 780
 gagctggatt acgctctggc tctgcgcaac gcttttatcg gacgaggctt cccagcacc 840
 attgggatgc tcatcgatac cagccgaaac ggttggggcg gatgcagcta tgggcaatgc 900
 agggccacgg gtcccagttc agataccagc agcgtgaatg cctactgtga cggctcgcga 960
 gtagaccgag gctatcatcg ggccaactgg tgttaaccag cgggtgggat cgcgcagcgc 1020
 cctcaggccg caccgcggtc cgggatcgac gtctactgtt ggggtgaaacc acctggggag 1080
 tccgatgggg tcagccaacc cgggatcgtc gatcccgacg atcccaacaa gaagttcgat 1140
 cctatgtgtg atcccaacgg acagagccgg tataactccg catacccaac tggggctctg 1200
 cccaacgcgc cccacgttgg gcggtggttc ccgcagcagt tcgaaatcct ggtgcggaac 1260
 gcctatccgc cgatecaacc ctaa 1284
 <210> 5
 <211> 427
 <212> PRT
 <213> Trình tự chưa biết
 <220> Trình tự axit amin AR19G-166-RW
 <223> AR19G-166-RW
 <400> 5
 Leu Asp Asn Pro Phe Ile Gly Ala Ile Gly Tyr Val Asn Pro Asp Trp
 1 5 10 15
 Ala Thr Asn Val Ile Ser Gln Ala Asn Gln Thr Ala Asp Pro Thr Leu
 20 25 30
 Ala Ala Gln Met Arg Lys Val Ala Thr Tyr Ser Thr Ala Val Trp Leu
 35 40 45
 Asp Arg Ile Ala Ala Ile Thr Ala Gly Arg Gly Leu Arg Gly His Leu
 50 55 60
 Asp Glu Ala Leu Arg Gln Met Gln Gln Ala Gly Gln Pro Val Val Ile
 65 70 75 80
 Thr Leu Val Ile Tyr Asp Leu Pro Asn Arg Asp Cys Ser Ala Ala Ala
 85 90 95
 Ser Asn Gly Glu Leu Leu Val Ala Gln Asn Gly Leu Ala Arg Tyr Lys
 100 105 110
 Ala Glu Phe Ile Asp Pro Ile Val Ala Ile Leu Ser Asp Pro Arg Tyr
 115 120 125
 Ala Gly Leu Arg Ile Val Thr Ile Ile Glu Pro Asp Ser Leu Pro Asn
 130 135 140
 Leu Val Thr Asn Leu Ser Ile Pro Ala Cys Ala Glu Ala Gln Asn Ala
 145 150 155 160
 Tyr Ile Glu Gly Ile Arg Tyr Ala Val Asn Arg Leu Arg Thr Ile Pro
 165 170 175
 Asn Val Tyr Ile Tyr Leu Asp Ile Ala His Ser Gly Trp Leu Gly Trp
 180 185 190
 Asp Asn Asn Phe Asn Gly Ala Val Asn Leu Tyr Thr Gln Val Val Gln
 195 200 205
 Gly Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn Val
 210 215 220
 Ala Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu
 225 230 235 240
 Thr Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp Asn
 245 250 255
 Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala Phe
 260 265 270
 Ile Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr Ser
 275 280 285
 Arg Asn Gly Trp Gly Gly Cys Ser Tyr Gly Arg Cys Arg Pro Thr Gly
 290 295 300

Pro Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser Arg
 305 310 315 320
 Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly
 325 330 335
 Ile Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Trp Tyr
 340 345 350
 Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly
 355 360 365
 Ile Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys Asp
 370 375 380
 Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu
 385 390 395 400
 Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu Ile
 405 410 415
 Leu Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro
 420 425

<210> 6
 <211> 1284
 <212> ADN
 <213> Trình tự chưa biết
 <220> Trình tự nucleotit AR19G-166-RW
 <223> AR19G-166-RW
 <400> 6

ttggacaatc cattcatcgg agccatcgga tacgtgaatc cggactgggc aaccaatgtg 60
 atcagccaaag cgaaccaaac ggctgatcca acctggcgg ctcaaagtgc taaggtggcc 120
 acctactcca cagctgtctg gttggatcgt atcgccgcca tcaccgctgg ccgcggttg 180
 cgcgggcatt tggatgaagc actacgcca atgcagcaag ctggccagcc ggttgatgac 240
 acccttggtga tctatgatct gccaaatcga gattgctctg ctgctgcctc caatggcgaa 300
 ttactggtcg ccagaaatgg actggcccgc tacaagcgg agttcatcga tcccatcgta 360
 gccattctct cagatccccg atatgcccgg ctacgcatcg tcaccatcat cgaaccggac 420
 tccttaccca acctggtcac gaacctcagc atcccggcat gcgcagaagc tcagaatgca 480
 tatatcgaag ggatccgcta tgcggtgaac cggctgcgga caattcccaa cgtctacatc 540
 tatctggata tcgcccactc aggtgtgttg ggctgggata ataactttaa tggggcagta 600
 aatctctaca ccaagtgtg gcaaggaatg gatcaagggt ttaacagtat cgaatggattc 660
 atcaccaaagc ttgccaaacta tccccccctc gaagaaccct atttgcccga tcccaacctg 720
 accatcgctg ggcaaccctg tgcgtctgcc agcttctacg aatggaacc cttactttgat 780
 gagctggatt acgctctggc tctgcgcaac gcttttatcg gacgaggctt cccagcacc 840
 attgggatgc tcatcgatac cagccgaaac ggggtggggc gatgcagcta tggcgatgc 900
 aggccacgg gtcccagttc agataccagc agcgtgaatg cctacgtgga cggctcgcga 960
 gtagaccgac gctatcatcg gggcaactgg tgtaaccagg cgggtgggat cggcgagcgc 1020
 cctcaggccg caccgcggtc cggtatcgac tggtagctgt gggtgaaacc acctggggag 1080
 tccgatgggg tcagccaacc cgggatcgtc gatcccgac atcccaacaa gaagtctgat 1140
 cctatgtgtg atcccaacgg acagagccgg tataactccg catacccaac tggggctctg 1200
 cccaacgcgc cccacgctgg gcggtggttc ccgcagcagt tcgaaatcct ggtgcggaac 1260
 gcctatccgc cgatccaacc ctaa 1284

<210> 7
 <211> 427
 <212> PRT
 <213> Trình tự chưa biết
 <220> Trình tự axit amin AR19G-166-QW
 <223> AR19G-166-QW
 <400> 7

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

Gly Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn Val
 210 215 220
 Ala Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu
 225 230 235 240
 Thr Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp Asn
 245 250 255
 Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala Phe
 260 265 270
 Ile Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr Ser
 275 280 285
 Arg Asn Gly Trp Gly Gly Cys Ser Tyr Gly Gln Cys Arg Pro Thr Gly
 290 295 300
 Pro Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser Arg
 305 310 315 320
 Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly
 325 330 335
 Ile Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Trp Tyr
 340 345 350
 Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly
 355 360 365
 Ile Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys Asp
 370 375 380
 Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu
 385 390 395 400
 Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu Ile
 405 410 415
 Leu Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro
 420 425

<210> 8

<211> 1284

<212> ADN

<213> Trình tự chưa biết

<220> Trình tự nucleotit AR19G-166-QW

<223> AR19G-166-QW

<400> 8

ttggacaatc cattcatcgg agccatcgga tacgtgaatc cggactgggc aaccaatgtg 60
 atcagccaag cgaaccaaac ggctgatcca acctggcgcg ctcaaatgcy taagggtggcc 120
 acctactcca cagctgtctg gttggatcgt atcgccgcca tcaccgctgg ccgcggtattg 180
 cgcgggcatt tggatgaagc actacgcca atgcagcaag ctggccagcc ggttgtgatc 240
 acccttgtga tctatgatct gccaaatcga gattgctctg ctgctgcctc caatggcgaa 300
 ttactggtcg cccagaatgg actggcccgc tacaagcgg agttcatcga tcccatcgta 360
 gccattctct cagatccccg atatgccggg ctacgcatcg tcaccatcat cgaaccggac 420
 tccttaccga acctggtcac gaacctcagc atcccggcat gcgcagaagc tcagaatgca 480
 tatatcgaag ggatccgcta tgcgtgaac cggctgcgga caattcccaa cgtctacatc 540
 tatctggata tcgcccactc aggcgtggtg ggctgggata ataactttaa tggggcagta 600
 aatctctaca cccaagttgt gcaaggaatg gatcaagggt ttaacagtat cgatggattc 660
 atcaccaacg ttgccaaacta tccccccctc gaagaaccct atttgcccga tcccaacctg 720
 accatcgctg ggcaaccctg tcgctctgcc agcttctacg aatggaacc ctactttgat 780
 gagctggatt acgctctggc tctgcgcaac gcttttatcg gacgaggctt cccagcacc 840
 attgggatgc tcatcgatac cagccgaaac ggggtggggcg gatgcagta tgggcaatgc 900
 aggccacggg gtcccagttc agataccagc agcgtgaatg cctacgtgga cggctcgcga 960
 gtagaccgac gctatcatcg gggcaactgg tgtaaccagg cgggtgggat cggcgagcgc 1020
 cctcaggccg caccgcggtc cggtatcgac tggtagctgt gggtgaaacc acctggggag 1080
 tcgcatgggg tcagccaacc cgggatcgtc gatcccgacg atcccaacaa gaagttcgat 1140
 cctatgtgtg atcccaacgg acagagccgg tataactccg catacccaac tggggctctg 1200
 cccaacgcgc cccacgctgg gcggtgggtc ccgcagcagt tcgaaatcct ggtgcggaac 1260
 gcctatccgc cgatccaacc cttaa 1284

<210> 9

<211> 474

<212> PRT

<213> Trình tự chưa biết

<220> Trình tự axit amin mã hóa bởi khung đọc mã mở AR19G-166

<223> Trình tự axit amin mã hóa bởi khung đọc mã mở AR19G-166

<400> 9

Pro Thr Arg Thr Pro Thr Pro Ala Ala Pro Thr Ala Thr Pro Thr Pro
 1 5 10 15
 Val Ala Pro Thr Ala Thr Pro Thr Arg Thr Pro Thr Pro Thr Leu Thr
 20 25 30
 Ser Thr Pro Gly Pro Thr Pro Thr Pro Pro Ser Gly Thr His Leu
 35 40 45
 Asp Asn Pro Phe Ile Gly Ala Ile Gly Tyr Val Asn Pro Asp Trp Ala
 50 55 60
 Thr Asn Val Ile Ser Gln Ala Asn Gln Thr Ala Asp Pro Thr Leu Ala
 65 70 75 80
 Ala Gln Met Arg Lys Val Ala Thr Tyr Ser Thr Ala Val Trp Leu Asp
 85 90 95
 Arg Ile Ala Ala Ile Thr Ala Gly Arg Gly Leu Arg Gly His Leu Asp
 100 105 110
 Glu Ala Leu Arg Gln Met Gln Gln Ala Gly Gln Pro Val Val Ile Thr

```

115      120      125
Leu Val Ile Tyr Asp Leu Pro Asn Arg Asp Cys Ser Ala Ala Ala Ser
130      135      140
Asn Gly Glu Leu Leu Val Ala Gln Asn Gly Leu Ala Arg Tyr Lys Ala
145      150      155      160
Glu Phe Ile Asp Pro Ile Val Ala Ile Leu Ser Asp Pro Arg Tyr Ala
165      170      175
Gly Leu Arg Ile Val Thr Ile Ile Glu Pro Asp Ser Leu Pro Asn Leu
180      185      190
Val Thr Asn Leu Ser Ile Pro Ala Cys Ala Glu Ala Gln Asn Ala Tyr
195      200      205
Ile Glu Gly Ile Arg Tyr Ala Val Asn Arg Leu Arg Thr Ile Pro Asn
210      215      220
Val Tyr Ile Tyr Leu Asp Ile Ala His Ser Gly Trp Leu Gly Trp Asp
225      230      235      240
Asn Asn Phe Asn Gly Ala Val Asn Leu Tyr Thr Gln Val Val Gln Gly
245      250      255
Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn Val Ala
260      265      270
Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu Thr
275      280      285
Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp Asn Pro
290      295      300
Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala Phe Ile
305      310      315      320
Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr Ser Arg
325      330      335
Asn Gly Trp Gly Gly Cys Ser Tyr Gly Gln Cys Arg Pro Thr Gly Pro
340      345      350
Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser Arg Val
355      360      365
Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly Ile
370      375      380
Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Ala Tyr Val
385      390      395      400
Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly Ile
405      410      415
Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys Asp Pro
420      425      430
Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu Pro
435      440      445
Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu Ile Leu
450      455      460
Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro
465      470
<210> 10
<211> 459
<212> PRT
<213> Trình tự chưa biết
<220> Trình tự axit amin mã hóa bởi khung đọc mã mở AR19G-12
<223> Trình tự axit amin mã hóa bởi khung đọc mã mở AR19G-12
<400> 10
Thr Ala Thr Pro Thr Arg Thr Pro Thr Pro Pro Ala Gln Pro Thr Ala
1      5      10      15
Thr Pro Thr Arg Thr Pro Thr Pro Thr Ala Val Ala Pro Thr Pro Val
20      25      30
Pro Gly Val His Leu Asp Asn Pro Phe Val Gly Ala Val Trp Tyr Val
35      40      45
Asn Pro Asp Trp Ala Ala Asn Val Leu Asn Gln Ala Gln Gln Thr Gly
50      55      60
Gly Thr Leu Gly Gln Arg Met Ala Ala Val Ala Gln Tyr Ser Thr Ala
65      70      75      80
Val Trp Leu Asp Arg Ile Gly Ala Ile Thr Glu Gly Arg Gly Leu Gln
85      90      95
Gly His Leu Asn Glu Ala Leu Ala Gln Gly Ala Asn Leu Ile Met Ile
100      105      110
Val Ile Tyr Asp Leu Pro Asn Arg Asp Cys Ala Ala Tyr Ala Ser Asn
115      120      125
Gly Glu Leu Leu Ile Ala Glu Asn Gly Leu Glu Arg Tyr Lys Arg Glu
130      135      140
Tyr Ile Asp Val Ile Tyr Asn Ile Leu Ala Gln Pro Gln Tyr Arg Asn
145      150      155      160
Leu Arg Ile Val Ala Val Ile Glu Pro Asp Ser Leu Pro Asn Leu Val
165      170      175
Thr Asn Leu Ser Glu Pro Ala Cys Ala Glu Ala Asn Ser Thr Gly Ala
180      185      190
Tyr Val Glu Gly Ile Arg Tyr Ala Val Asn Lys Leu His Gln Leu Pro
195      200      205
Asn Val Tyr Ile Tyr Leu Asp Ile Ala His Ser Gly Trp Leu Gly Trp

```

```

      210              215              220
Asp Asn Asn Phe Gln Gly Ala Val Asn Leu Tyr Thr Gln Val Val Gln
225              230              235              240
Gly Met Ala Asp Gly Phe Asn Ser Ile Asp Gly Phe Val Thr Asn Thr
      245              250              255
Ala Asn Tyr Thr Pro Val Glu Glu Pro Tyr Leu Pro Asp Pro Asn Leu
      260              265              270
Thr Ile Gly Gly Gln Pro Val Arg Ser Ala Asn Phe Tyr Glu Trp Asn
      275              280              285
Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Gln Ala Leu Arg Gln Ala Phe
      290              295              300
Ile Gly Arg Gly Phe Pro Ser Gly Ile Gly Met Leu Ile Asp Thr Ser
305              310              315              320
Arg Asn Gly Trp Gly Gly Ser Ser Tyr Gly Arg Ser Arg Pro Ser Gly
      325              330              335
Ala Ser Ser Asn Thr Thr Asn Val Asn Val Tyr Val Asp Glu Ser Arg
      340              345              350
Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly Gly
      355              360              365
Ile Gly Glu Arg Pro Arg Ala Asn Pro Ala Ala Gly Ile Asp Ala Tyr
      370              375              380
Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro Gly
385              390              395              400
Val Val Asp Pro Val Asp Pro Asn Lys Lys Phe Asp Ile Met Cys Asp
      405              410              415
Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala Leu
      420              425              430
Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Glu Gln Phe Glu Ile
      435              440              445
Leu Val Gln Asn Ala Tyr Pro Pro Ile Gln Pro
      450              455
<210> 11
<211> 548
<212> PRT
<213> Trình tự chưa biết
<220> Trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1
<223> Trình tự axit amin mã hóa bởi khung đọc mã mở OJ1-1
<400> 11
Met Ser Arg Ile Arg Trp Gln Ile Leu Leu Arg Ala Thr Leu Ile Val
 1              5              10              15
Ala Ile Ile Ala Ala Val Ala Leu Ile Ser Leu Gln Ala Arg Ala
      20              25              30
Gln Ser Thr Leu Lys Val Gln Tyr Arg Cys Ala Asp Thr Asn Ala Thr
      35              40              45
Gly Asn Gln Ile Lys Pro His Leu Asn Ile Val Asn Thr Gly Ser Ser
      50              55              60
Ala Val Ser Leu Thr Ala Leu Lys Ala Arg Tyr Tyr Tyr Thr Ile Asp
      65              70              75              80
Gly Glu Lys Ala Gln Ser Tyr Trp Cys Asp Tyr Ala Thr Ile Gly Cys
      85              90              95
Ser Asn Ile Thr Ala Ser Phe Val Lys Met Ala Thr Ala Val Ser Gly
      100              105              110
Ala Asp Tyr Tyr Leu Glu Val Gly Phe Thr Ser Gly Thr Leu Asn Ala
      115              120              125
Gly Ala Ser Thr Gly Glu Ile Gln Asn Arg Phe Ala Lys Ser Asp Trp
      130              135              140
Thr Asn Tyr Thr Gln Thr Gly Asp Tyr Ser Phe Asp Pro Ser Lys Thr
145              150              155              160
Ser Phe Ala Asp Trp Asn Lys Val Thr Leu Tyr Asn Asn Gly Thr Leu
      165              170              175
Val Trp Gly Val Glu Pro Ser Gly Ala Ser Ala Pro Thr Asn Thr Ser
      180              185              190
Ala Pro Ala Ala Thr Ala Thr Arg Thr Asn Thr Ala Val Thr Gly Pro
      195              200              205
Thr Asn Thr Ser Ala Pro Ala Ala Thr Ala Thr Arg Thr Asn Thr Pro
      210              215              220
Gly Gly Pro Thr Ala Thr Arg Thr Arg Thr Pro Thr Arg Thr Arg Thr
225              230              235              240
Pro Thr Arg Thr Ala Thr Gly Gln Val Ala Pro Thr Asn Thr Thr Ala
      245              250              255
Pro Ala Ala Thr Ala Thr Arg Thr Pro Thr Thr Gly Thr Gly Pro Thr
      260              265              270
Ala Thr Arg Thr Asn Thr Thr Val Ala Pro Thr Ala Thr Thr Gly Pro
      275              280              285
Thr Ala Ala Pro Gly Thr His Leu Asp Asn Pro Phe Val Gly Ala Lys
      290              295              300
Trp Tyr Ile Asn Pro Asp Trp Gln Gly Arg Thr Val Leu Asn Tyr Ser
305              310              315              320
Thr Ala Val Trp Met Asp Arg Ile Ala Ala Val Val Gly Gly Ser Gly

```

325 330 335
 Val Thr Arg Asn Leu Ala Gly His Leu Asp Ala Ala Leu Ser Gln Gly
 340 345 350
 Ala Asn Leu Ile Ile Ile Val Val Tyr Asp Leu Pro Asn Arg Asp Cys
 355 360 365
 Phe Ala Leu Ala Ser Asn Gly Glu Leu Lys Ile Ala Glu Asp Gly Phe
 370 375 380
 Asn Arg Tyr Lys Asn Glu Tyr Ile Glu Pro Met Ile Ala Thr Leu Asn
 385 390 395 400
 Gln Ala Lys Tyr Ala Asn Leu Arg Ile Val Ala Val Ile Glu Pro Asp
 405 410 415
 Ser Leu Pro Asn Leu Val Thr Asn Leu Ser Asp Pro Asp Cys Ala Gln
 420 425 430
 Ala Asn Gly Thr Gly Gly Tyr Lys Asp Gly Ile Val Tyr Ala Leu Asn
 435 440 445
 Arg Leu Asn Thr Met Thr Asn Val Tyr Pro Tyr Ile Asp Ile Ala His
 450 455 460
 Ser Gly Trp Leu Gly Trp Asp Ser Asn Phe Gly Pro Ala Ala Asp Leu
 465 470 475 480
 Phe Ala Asn Val Val Lys Ala Thr Thr Lys Gly Val Asn Ser Val Asp
 485 490 495
 Gly Phe Val Ser Asn Thr Ala Asn Tyr Thr Pro Leu Asp Glu Pro Phe
 500 505 510
 Leu Thr Asp Pro Asn Leu Thr Val Gly Gly Gln Gln Val Lys Ser Ala
 515 520 525
 Ser Phe Tyr Glu Trp Asn Pro Tyr Phe Asp Glu Arg Asp Tyr Val Thr
 530 535 540
 Ala Met Arg Asn

545

<210> 12

<211> 22

<212> ADN

<213> Trình tự nhân tạo

<220> Mô tả trình tự nhân tạo: đoạn mỗi xuôi

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

<400> 12

ttggacaatc cattcatcgg ag

22

<210> 13

<211> 22

<212> ADN

<213> Trình tự nhân tạo

<220> Mô tả trình tự nhân tạo: đoạn mỗi ngược

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

<400> 13

ttagggttgg atcggcggat ag

22

<210> 14

<211> 29

<212> ADN

<213> Trình tự nhân tạo

<220> Mô tả trình tự nhân tạo: đoạn mỗi xuôi

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

<400> 14

caccatgttg gacaatccat tcatcggag

29

<210> 15

<211> 428

<212> PRT

<213> Herpetosiphon aurantiacus DSM 785

<220> Vùng xúc tác của họ glucosit hydrolaza 6

<223> Vùng xúc tác của họ glucosit hydrolaza 6

<400> 15

Val Ala Asn Pro Phe Val Gly Ala Gln Gly Tyr Ile Asn Ser Glu Tyr
1 5 10 15

Ala Ala Arg Val Asn Ala Glu Ala Asn Ala Thr Gly Gly Thr Leu Gly
20 25 30

Ser Gln Met Arg Gln Val Ala Ser Tyr Pro Thr Ala Val Trp Leu Asp
35 40 45

Arg Ile Ala Ala Ile Ala Gly Ser Ser Glu Ser Met Gly Leu Arg Ala
50 55 60

His Leu Asp Ala Ala Leu Val Gln Gln Gln Thr Ser Gly Gln Gln Val
65 70 75 80

Ala Ile Ser Ile Val Val Tyr Asp Leu Pro Asn Arg Asp Cys Ala Ala
85 90 95

Leu Ala Ser Asn Gly Glu Leu Lys Ile Ser Glu Asn Gly Leu Asn Arg
100 105 110

Tyr Lys Thr Glu Tyr Ile Asp Pro Ile Ala Ala Ile Val Gly Glu Ser
115 120 125

Lys Tyr Ala Ser Leu Arg Ile Val Val Ile Leu Glu Pro Asp Ser Leu
130 135 140

Ser Asn Leu Val Thr Asn Ala Ser Ile Pro Ala Cys Ala Glu Ala Ile

```

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

```

```

<210> 16
<211> 150
<212> PRT
<213> Caldibacillus cellulovorans
<220> CBM3 của xenlobiohydrolaza
<223> CBM3 của xenlobiohydrolaza
<400> 16

```

```

Ser Leu Val Val Gln Tyr Arg Ala Ala Asp Thr Asn Ala Gly Asp Asn
 1              5              10              15
Gln Leu Lys Pro His Phe Arg Ile Val Asn Arg Gly Thr Ser Ser Val
              20              25              30
Pro Leu Ser Glu Leu Thr Ile Arg Tyr Trp Tyr Thr Val Asp Gly Asp
              35              40              45
Lys Pro Gln Val Phe Asn Cys Asp Trp Ala Gln Val Gly Cys Ser Asn
              50              55              60
Val Arg Gly Ser Phe Val Lys Leu Ser Thr Gly Arg Thr Gly Ala Asp
              65              70              75              80
Tyr Tyr Ile Glu Ile Thr Phe Thr Ser Gly Ala Gly Ser Leu Ala Ala
              85              90              95
Gly Ala Ser Ser Gly Asp Ile Gln Val Arg Ile Asn Lys Asn Asp Trp
              100              105              110
Thr Asn Tyr Asn Glu Ala Asn Asp Tyr Ser Tyr Asp Pro Thr Lys Thr
              115              120              125
Ser Phe Ala Asp Trp Asn Arg Val Thr Leu Tyr Arg Asn Gly Gln Leu
              130              135              140
Ile Trp Gly Val Glu Pro
145              150

```

```

<210> 17
<211> 269
<212> PRT
<213> Trình tự chưa biết
<220> Trình tự liên kết và trình tự axit amin CBM3
<223> Trình tự liên kết và CBM3
<400> 17

```

```

Ser Gly Ala Ser Ala Pro Thr Asn Thr Ser Ala Pro Ala Ala Thr Ala
 1              5              10              15
Thr Arg Thr Asn Thr Ala Val Thr Gly Pro Thr Asn Thr Ser Ala Pro
              20              25              30
Ala Ala Thr Ala Thr Arg Thr Asn Thr Pro Gly Gly Pro Thr Ala Thr
              35              40              45
Arg Thr Arg Thr Pro Thr Arg Thr Arg Thr Pro Thr Arg Thr Ala Thr
              50              55              60
Gly Gln Val Ala Pro Thr Asn Thr Thr Ala Pro Ala Ala Thr Ala Thr

```

```

65          70          75          80
Arg Thr Pro Thr Thr Gly Thr Gly Pro Thr Ala Thr Arg Thr Asn Thr
      85          90          95
Thr Val Ala Pro Thr Ala Thr Thr Gly Pro Thr Ala Ala Pro Gly Thr
      100          105          110
His Gln Ser Thr Leu Lys Val Gln Tyr Arg Cys Ala Asp Thr Asn Ala
      115          120          125
Thr Gly Asn Gln Ile Lys Pro His Leu Asn Ile Val Asn Thr Gly Ser
      130          135          140
Ser Ala Val Ser Leu Thr Ala Leu Lys Ala Arg Tyr Tyr Tyr Thr Ile
      145          150          155          160
Asp Gly Glu Lys Ala Gln Ser Tyr Trp Cys Asp Tyr Ala Thr Ile Gly
      165          170          175
Cys Ser Asn Ile Thr Ala Ser Phe Val Lys Met Ala Thr Ala Val Ser
      180          185          190
Gly Ala Asp Tyr Tyr Leu Glu Val Gly Phe Thr Ser Gly Thr Leu Asn
      195          200          205
Ala Gly Ala Ser Thr Gly Glu Ile Gln Asn Arg Phe Ala Lys Ser Asp
      210          215          220
Trp Thr Asn Tyr Thr Gln Thr Gly Asp Tyr Ser Phe Asp Pro Ser Lys
      225          230          235          240
Thr Ser Phe Ala Asp Trp Asn Lys Val Thr Leu Tyr Asn Asn Gly Thr
      245          250          255
Leu Val Trp Gly Val Glu Pro Ser Gly Ala Ser Ala Pro
      260          265

```

```

<210> 18
<211> 810
<212> ADN
<213> Trình tự chưa biết
<220> Trình tự liên kết và trình tự nucleotit CBM3
<223> Trình tự liên kết và CBM3
<400> 18

```

```

tccggtgccca gtgcgcccgc caatacgtct gccctgctg cgactgcaac gcgcaccaac 60
acagccgttta cgggtccgcac caacacgagt gcaccggccg ctacggcaac gcgtacgaac 120
actccgggtg gaccaaccgc aaccgcacg cgcacgcca cccggaccg tacgcccacg 180
cgcaccgcaa ccggtcaagt cgcaccaacg aataccactg cgccggcagc caccgccacg 240
cgcacaccga caactggag cggtcctacc gccacacgca ccaacaccac cgttgcacca 300
actgccacga ccggcccgcac cgctgcgccc ggcacacacc agtcaacgct gaaagtccag 360
tatcgttgcg ccgataccaa cgcaactggt aaccagatca aaccgcacct gaacattgtc 420
aacaccggta gttctgccgt atcgctcacc gctcttaaag cgcgctacta ctacaccatc 480
gatggcgaga aagcgcaatc gtattggtgc gactatgcaa cgattggctg cagcaatatc 540
accgcctcgt tcgtgaaaat ggcgacagca gtaagcggcg ccgactatta cttggaagtg 600
ggctttacca gcggtactct gaatgcagg gctagtacgg gtgaaatcca gaatcgcttt 660
gcgaaaagcg attggacgaa ctatacgag acaggcgatt acagcttcga tcccagcaaa 720
accagctttg cagactggaa caaagtaacc ctgtacaaca acggcacctt ggtttggggc 780
gtggaaccaa gtggcgcgag tgctccataa 810

```

```

<210> 19
<211> 697
<212> PRT
<213> Trình tự chưa biết
<220> Trình tự axit amin AR19G-166-RA được bổ sung CBM
<223> AR19G-166-RA được bổ sung CBM
<400> 19

```

```

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

```



```

      195              200              205
Gln Gly Met Asp Gln Gly Phe Asn Ser Ile Asp Gly Phe Ile Thr Asn
  210              215              220
Val Ala Asn Tyr Thr Pro Leu Glu Glu Pro Tyr Leu Pro Asp Pro Asn
  225              230              235              240
Leu Thr Ile Ala Gly Gln Pro Val Arg Ser Ala Ser Phe Tyr Glu Trp
      245              250              255
Asn Pro Tyr Phe Asp Glu Leu Asp Tyr Ala Leu Ala Leu Arg Asn Ala
      260              265              270
Phe Ile Gly Arg Gly Phe Pro Ser Thr Ile Gly Met Leu Ile Asp Thr
      275              280              285
Ser Arg Asn Gly Trp Gly Gly Cys Ser Tyr Gly Arg Cys Arg Pro Thr
      290              295              300
Gly Pro Ser Ser Asp Thr Ser Ser Val Asn Ala Tyr Val Asp Gly Ser
  305              310              315              320
Arg Val Asp Arg Arg Tyr His Arg Gly Asn Trp Cys Asn Gln Ala Gly
      325              330              335
Gly Ile Gly Glu Arg Pro Gln Ala Ala Pro Arg Ser Gly Ile Asp Ala
      340              345              350
Tyr Val Trp Val Lys Pro Pro Gly Glu Ser Asp Gly Val Ser Gln Pro
      355              360              365
Gly Ile Val Asp Pro Asp Asp Pro Asn Lys Lys Phe Asp Pro Met Cys
  370              375              380
Asp Pro Asn Gly Gln Ser Arg Tyr Asn Ser Ala Tyr Pro Thr Gly Ala
  385              390              395              400
Leu Pro Asn Ala Pro His Ala Gly Arg Trp Phe Pro Gln Gln Phe Glu
      405              410              415
Ile Leu Val Arg Asn Ala Tyr Pro Pro Ile Gln Pro Ser Gly Ala Ser
      420              425              430
Ala Pro Thr Asn Thr Ser Ala Pro Ala Ala Thr Ala Thr Arg Thr Asn
      435              440              445
Thr Ala Val Thr Gly Pro Thr Asn Thr Ser Ala Pro Ala Ala Thr Ala
      450              455              460
Thr Arg Thr Asn Thr Pro Gly Gly Pro Thr Ala Thr Arg Thr Arg Thr
  465              470              475              480
Pro Thr Arg Thr Arg Thr Pro Thr Arg Thr Ala Thr Gly Gln Val Ala
      485              490              495
Pro Thr Asn Thr Thr Ala Pro Ala Ala Thr Ala Thr Arg Thr Pro Thr
      500              505              510
Thr Gly Thr Gly Pro Thr Ala Thr Arg Thr Asn Thr Thr Val Ala Pro
      515              520              525
Thr Ala Thr Thr Gly Pro Thr Ala Ala Pro Gly Thr His Gln Ser Thr
      530              535              540
Leu Lys Val Gln Tyr Arg Cys Ala Asp Thr Asn Ala Thr Gly Asn Gln
  545              550              555              560
Ile Lys Pro His Leu Asn Ile Val Asn Thr Gly Ser Ser Ala Val Ser
      565              570              575
Leu Thr Ala Leu Lys Ala Arg Tyr Tyr Thr Ile Asp Gly Glu Lys
      580              585              590
Ala Gln Ser Tyr Trp Cys Asp Tyr Ala Thr Ile Gly Cys Ser Asn Ile
      595              600              605
Thr Ala Ser Phe Val Lys Met Ala Thr Ala Val Ser Gly Ala Asp Tyr
      610              615              620
Tyr Leu Glu Val Gly Phe Thr Ser Gly Thr Leu Asn Ala Gly Ala Ser
  625              630              635              640
Thr Gly Glu Ile Gln Asn Arg Phe Ala Lys Ser Asp Trp Thr Asn Tyr
      645              650              655
Thr Gln Thr Gly Asp Tyr Ser Phe Asp Pro Ser Lys Thr Ser Phe Ala
      660              665              670
Asp Trp Asn Lys Val Thr Leu Tyr Asn Asn Gly Thr Leu Val Trp Gly
      675              680              685
Val Glu Pro Ser Gly Ala Ser Ala Pro
      690              695
<210> 20
<211> 2094
<212> ADN
<213> Trình tự chưa biết
<220> Trình tự nucleotit AR19G-166-RA được bổ sung CBM
<223> AR19G-166-RA được bổ sung CBM
<400> 20
atgctggata acccgtttat cggcgcgatt ggctatgtta atccggattg ggcgacaaac 60
gtgattttctc aggcgaatca gacggccgat ccaactcttg cggcccaaat gcgcaaagtg 120
gccacctata gcactgctgt atggcttgat cgtattgcgg caattactgc tggtcgcgga 180
ttgcgcggcc atcttgatga agcattgcgt cagatgcaac aggcgggtca accggttggt 240
attacctggg tcatctacga tctgccgaat cgcgattggt cagcggcggc ctcgaatggt 300
gagtttgctg ttgccaaaa cgggttagct cgctataagg cggaattcat tgacctgatc 360
gttgcgattc tgtccgatcc acgctatgct ggtttacgga ttgttaccat catcgaaccg 420
gattcactcc ctaatctggt taccaacctc tcaatccctg cctgtgccga agcacagaac 480
gcatatattg aaggcattcg ttatgccgtg aatcgccctg gtaccattcc gaatgtgtac 540

```

atctatctgg	atattgcgca	tagcgggtgg	cttggatggg	acaataattt	caacgggtcg	600
gtgaacctgt	acactcaagt	cgtacagggg	atggatcagg	gcttcaacag	catcgatggt	660
tttatcacga	atgttgccaa	ttatacaccc	ttagaagagc	cctatctgcc	cgatcctaac	720
ctcactattg	cgggccaacc	ggtccgttct	gcttcctttt	acgaatggaa	tccgtatttt	780
gatgagctgg	attacgcgtt	agcgttacgg	aacgcgttca	ttggtcgcgg	ctttccgtcc	840
actatcggga	tgctgattga	caccagtcgc	aacgggtggg	gaggctgttc	ctatggacgg	900
tgccgtccga	caggcccttc	ctcagacacc	tcgtctgtga	acgcgtatgt	cgatggtagc	960
cgtgtggatc	gccgctatca	ccgtgggaat	tggtgcaatc	aggcgggcgg	tattggggaa	1020
cgtcctcaag	cagctccgcg	ttcggggatt	gacgcgtatg	tatgggtgaa	accaccgggt	1080
gaaagcgacg	gtgtctcgca	gcctggcatt	gtggaccggg	acgaccggaa	caagaagttt	1140
gaccctatgt	tgatccgaa	cggccaatcc	cgttacaatt	ctgcttacc	gaccggcgcg	1200
ctgccaaatg	cgccgcatgc	gggcccgttg	tttccgcagc	agttcgagat	cctgggtcgc	1260
aatgcctacc	ctccgattca	gccgtccggt	gccagtgcgc	cgaccaatac	gtctgccct	1320
gctgcgactg	caacgcgcac	caacacagcc	gttacgggtc	cgaccaacac	gagtgcaccg	1380
gccgctacgg	caacgcgtac	gaacactccg	ggtggaccaa	ccgcaaccgg	cacgcgcacg	1440
ccaaccggga	cccgtacgcc	cacgcgcacc	gcaaccggtc	aagtcgcacc	aacgaatacc	1500
actgcgcggg	cagccaccgc	cacgcgcaca	ccgacaactg	ggaccgggtc	taccgccaca	1560
cgcaccaaca	ccaccgttgc	accaactgcc	acgaccggcc	cgaccgctgc	gcccggcaca	1620
caccagtcaa	cgctgaaagt	ccagtatcgt	tgcccgata	ccaacgcaac	tggttaaccag	1680
atcaaaccgc	acctgaacat	tgtcaacacc	ggtagtctctg	ccgtatcgct	caccgctctt	1740
aaagcgcgct	actactacac	catcgatggc	gagaaagcgc	aatcgatttg	gtgcgactat	1800
gcaacgattg	gctgcagcaa	tatcacggcc	tcgttcgtga	aaatggcgac	agcagtaagc	1860
ggcgcggact	attacttgga	agtgggcttt	accagcggta	ctctgaatgc	aggggctagt	1920
acgggtgaaa	tccagaatcg	ctttgcgaaa	agcgattgga	cgaactatac	gcagacaggc	1980
gattacagct	tcgatcccag	caaaaccagc	tttgcagact	ggaacaaaagt	aaccctgtac	2040
aacaacggca	ccctggtttg	gggcgtggaa	ccaagtggcg	cgagtgtctc	ataa	2094

Fig.1

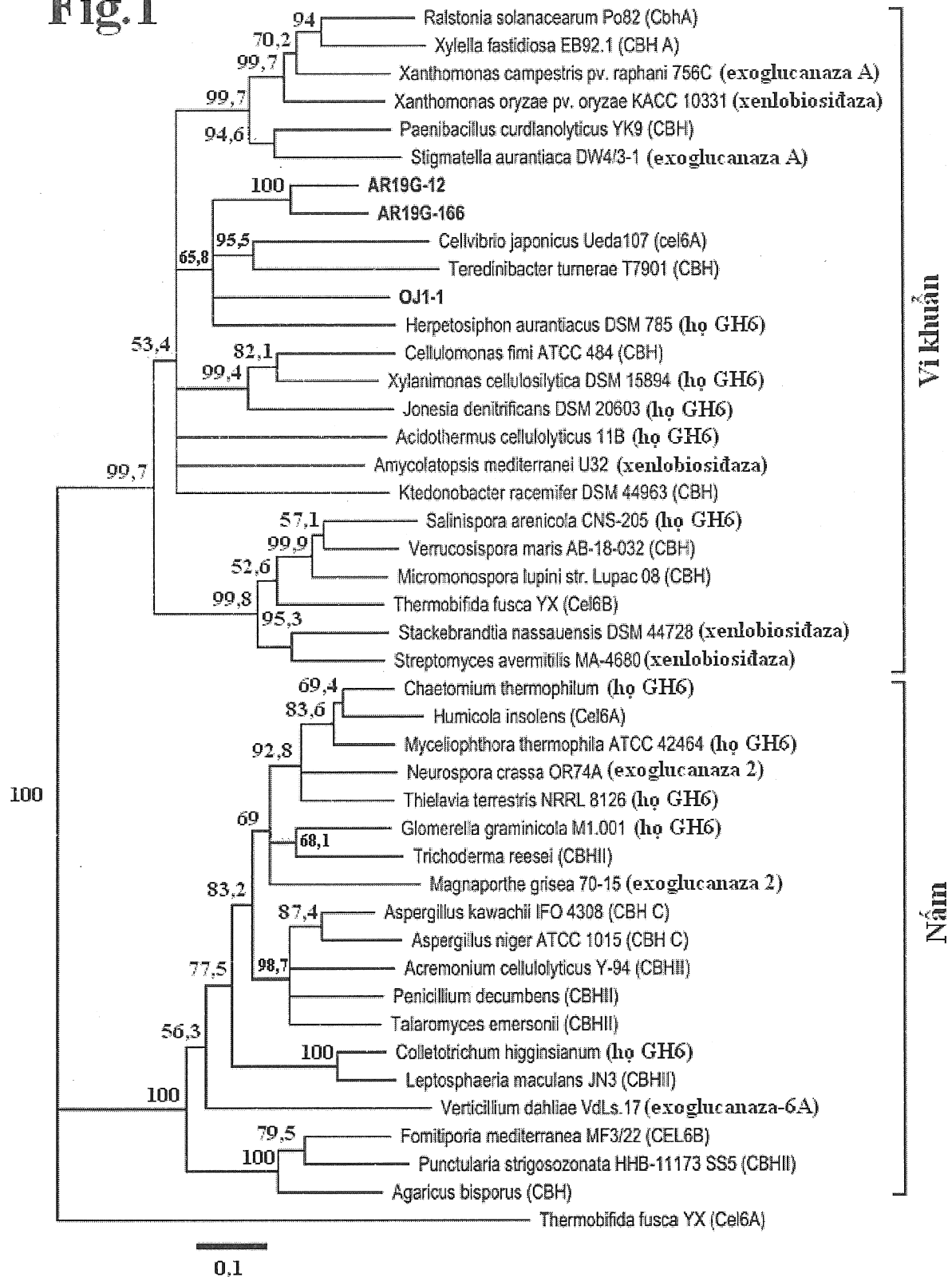


Fig.2A

OJ1-1	1	MSRIHQIILLRATLIVAIIAAAVALISLARAGSTLKVQYRCADTNATGNIKPHLIVNTGGSAVSLTALKARYYTI DEKAQSVWDVATIGGSHITASFVKVATANSADYVILEVGTSGTLWAGA	130
ARI98-12	0	-----	0
ARI98-166	0	-----	0
H. aurentiacus_CD	0	-----	0
OJ1-1	131	STGEIQNRFAKSDWNTYTGDSFDPKTSFADWKKVTLVANGTLVWGVEPSGASPTNLSAPAAATATRTNTATGPTNLSAPAAATATRTNTPGPTATRTNTPRTPTRTATGQVAPTNTTAPAA-	259
ARI98-12	0	-----	0
ARI98-166	1	-----PRTPTPAAP	10
H. aurentiacus_CD	0	-----	0
OJ1-1	260	TATRTP-TTGTGPTATRTNTTIVAPTATTGPTTAAGTHLDIPFGAAYINPDN-----QGRLLINSTAVNDRIAVW---GGSGVTRILAGHE-DAALSGGANLITLAVYDLPNR	366
ARI98-12	1	TATPTR-TPTPPAQPTATPTPTPTAVAPTTPGHHLDIPFGAAYINPDNANWLNQNDGTGG-TGGRMALADYSTAVNDRICAIT---EGRLGGHNEA-----AAGHANLITLAVYDLPNR	120
ARI98-166	11	TATPTVAPTATPTPTPTLSTPTPTTPPPSGTHLDIPFGAAYINPDNATNWSQNDGTADPTLAAGRHVATNSTAVNDRICAIT---AGRLRGHDEARQNDAGGPPVITLIVYDLPNR	137
H. aurentiacus_CD	1	-----VAIPFGAAGYINSEYAAKRWNAENATGG-TGGSRKQVASTAVNDRICAITAGSSSESHERAHDALNQDTSQGGVHSTAVYDLPNR	92
OJ1-1	367	DCFLASNGELKIAEDGFRYKTEIERNATLNOAKNARIVAVIPEDSIPLVNLTSPDCAHGTGGVQGIWALVREVTINVTPTD AHSGLGNDSEFSPADCFANVYKATNGVNSID	496
ARI98-12	121	DCMVASNGELTAENGFRYKTEISVYNILAQPCNRI VAVIPEDSIPLVNLTSPDCAEASTGAVIEIRYANVCHLPWVYTD AHSGLGNDSEFSGANLKTAVYDNDGFIISID	250
ARI98-166	138	DCSMASNGELVAONGFRYKTEIDPWILSDPFRAGERVITILEDSIPLVNLTSPDCAEAGN---ANIEGRYANDEETIPRPVYD AHSGLGNDSEFNGANLKTAVYDNDGCFISID	265
H. aurentiacus_CD	93	DCMLASNGELKISENGFRYKTEIDPMALNGESKVASERVWILEDSISNLTNVSIPACAEIISSEANVGGVAILGCVTSVYVTD AHSGLGNDSEFPAICLVOTPTGTTNGLVGID	222
OJ1-1	497	GFHSITANTPTDPTITDPNLTGG-DQVKSASVEYMPVFEEDYATAPN-----	548
ARI98-12	251	GFHTNTANTPTPEPTITDPNLTIGG-QPVRSANVEYMPVFEEDYACALDPAFEGGFPSSGIGNLTDSRNGHGSSVGRSPSGASSNTINWYIDESVYDPRYHRGNGWAGG-GIGERPRANPA	378
ARI98-166	266	GFHTNANTPTPEPTITDPNLTAG-QPVRSASVEYMPVFEEDYALALDPAFEGGFPSTIGNLTDSRNGHGSSVGGCRPTGFSSTSSNVAIYDGSYDPRYHRGNGWAGG-GIGERPPAAR	393
H. aurentiacus_CD	223	GFHSITANTPTPEPTITPNCETLGGNPVRSASVEYMPVFEEDYALGNACITAGPSPGIGNLTDSRNGWGTAT---PPTINWSS-SNLSLETIWNDSVLDRPFRGGGNGWAGG-GIGERTIAAPV	347
OJ1-1	548	-----	548
ARI98-12	379	AGIDAYVWVWPPGESDGVSGPQGVDDPDPAKKEDINCPDNGGSRVNSAVPTGALPNAPHAGRNPEQGEELVQNNVPPPIQP	459
ARI98-166	394	SGIDAYVWVWPPGESDGVSGPGLVDDPDPAKKEDIPDPCDNGGSRVNSAVPTGALPNAPHAGRNFPQOGEELVKNVATPPPIQP	474
H. aurentiacus_CD	348	SGIDAYVWVWPPGESDGVATAGVDDPDPAKDFDAPCDPNAQVRNTATPTTNALAGAPHAGRNFPSPGFMLVKNVATPPPISQ	428

Fig.2B

QJ1-1 1 KSRIRWOLLRALIYVAILAAKVALISLQRAQSTIKVQYRQADTNAIGCHKPHLVNIIIGSSWSEITALKARTYNTIDGFKQGSYVCDIATIGCSNLTASEKVAIVASGVYIEIGETSGI--LVA 128
 C. cel lulovorans_CBM3 1 -----SIVDYQADTAAEDQIKPHFRVWKGSSPLSELTIRVNTIDGKQGFNCDIAGGSMRSEFKLSTRTGADYIEITETSGAGSAA 96

QJ1-1 129 GASTGELQARFARSDVNTITQEDYS-DPKTSFADWIKVLEANGLENGVPSASAPTNTSAPAAIATRTNVTGPTNTSAPAAIATRTNTPGPTATRTPTRTPTRTATGGVAPNTTAPA 258
 C. cel lulovorans_CBM3 97 GASSGELQARINQDNTNYEANDSIDPKTSFADWIKVLEANGLENGVPSASAPTNTSAPAAIATRTNVTGPTNTSAPAAIATRTNTPGPTATRTPTRTPTRTATGGVAPNTTAPA 150

QJ1-1 259 ATATRTPTTGPTATRTNTTVAPTATTGPTAAPGTHLDNPFVGAKWYIPDQNGRTVLNVSTAVMDRIAANVGGSSNTRNLGCHLDKHLSDGKHLIIIVYDLPIRDCFLASNGELKLAEDGFNRVK 388
 C. cel lulovorans_CBM3 150 -----

QJ1-1 389 NEYIEPNIATLNGAKYANLRIVAVIEPDSLPNLVTNLSDPDCAQANGTGRYKDGIVYALNRLNTNTVYPPYIDIAHSGLGUDSNFEPADLEFANVYKATTKGWSVDGFVSNITANYTPIDEFLTDPNL 518
 C. cel lulovorans_CBM3 150 -----

QJ1-1 519 TVGGGQVKSASFYEMIPYFEDRDYVTANRN 548
 C. cel lulovorans_CBM3 150 -----

Fig.3A

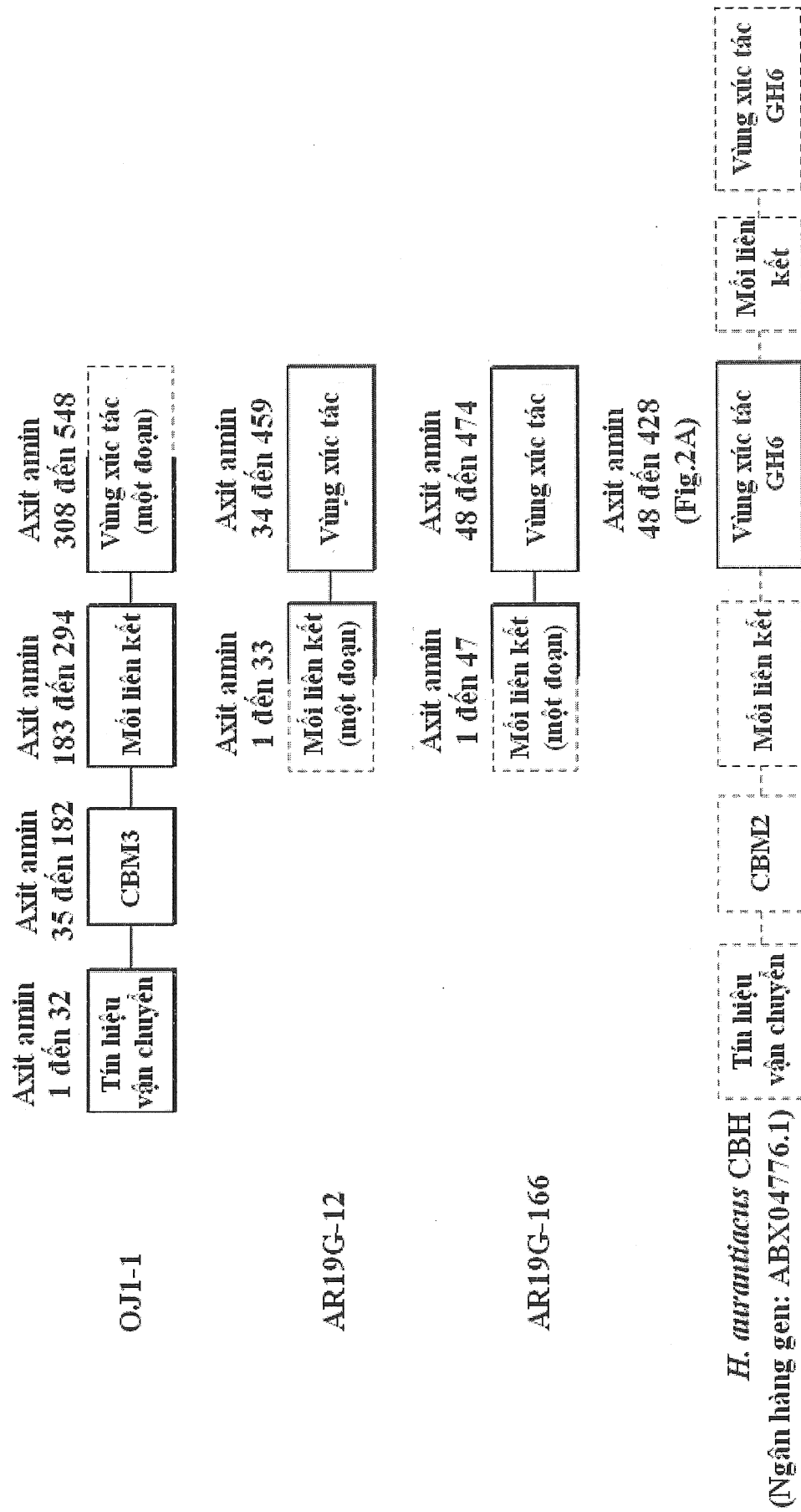


Fig.3B

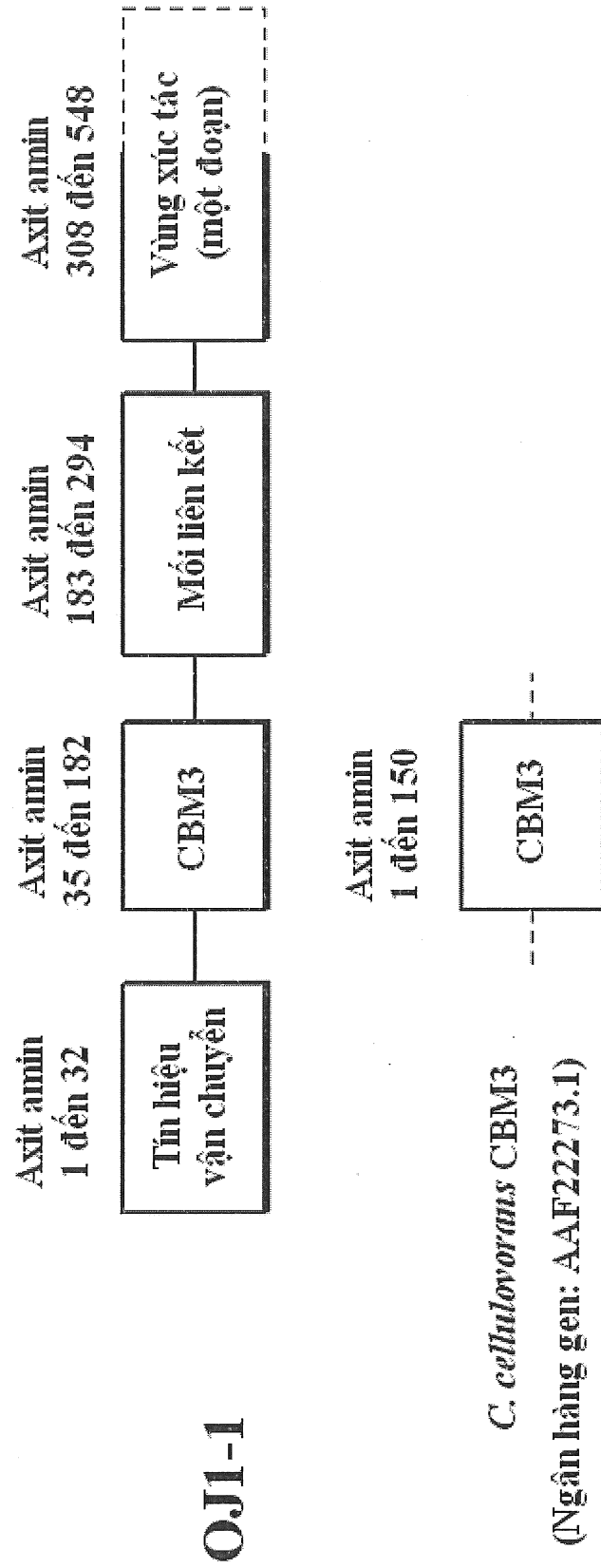


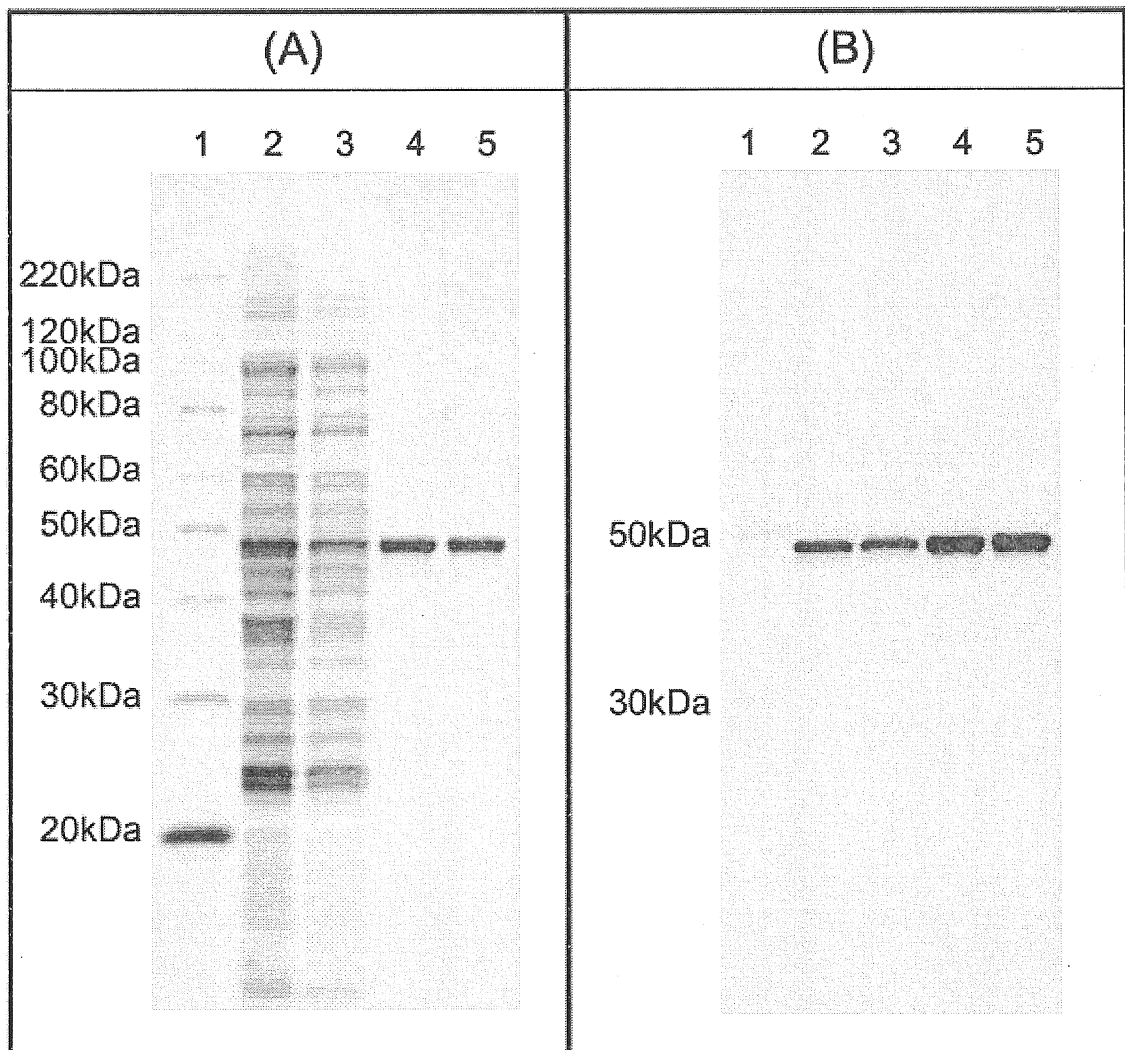
Fig.4

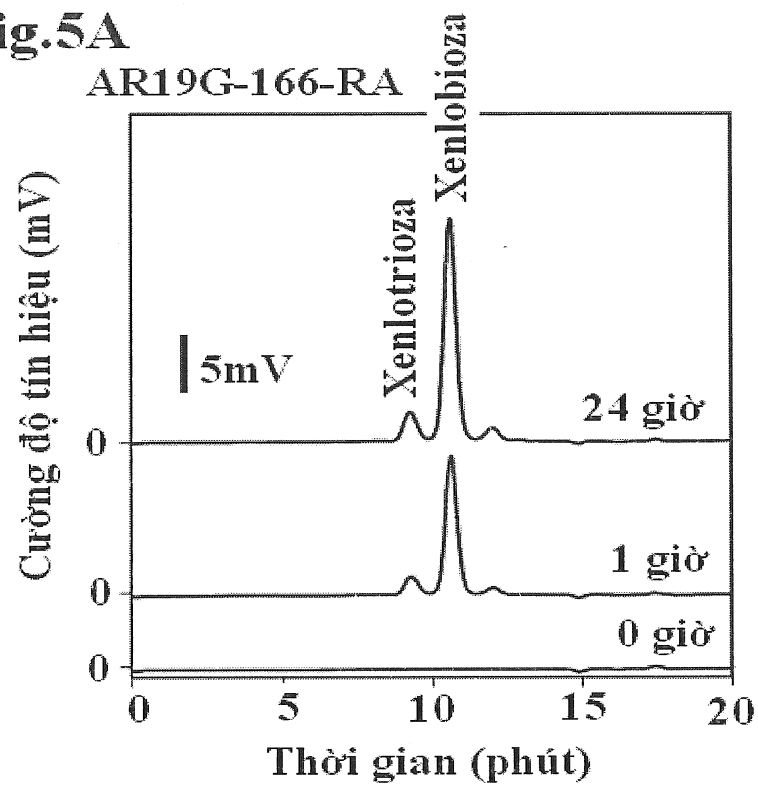
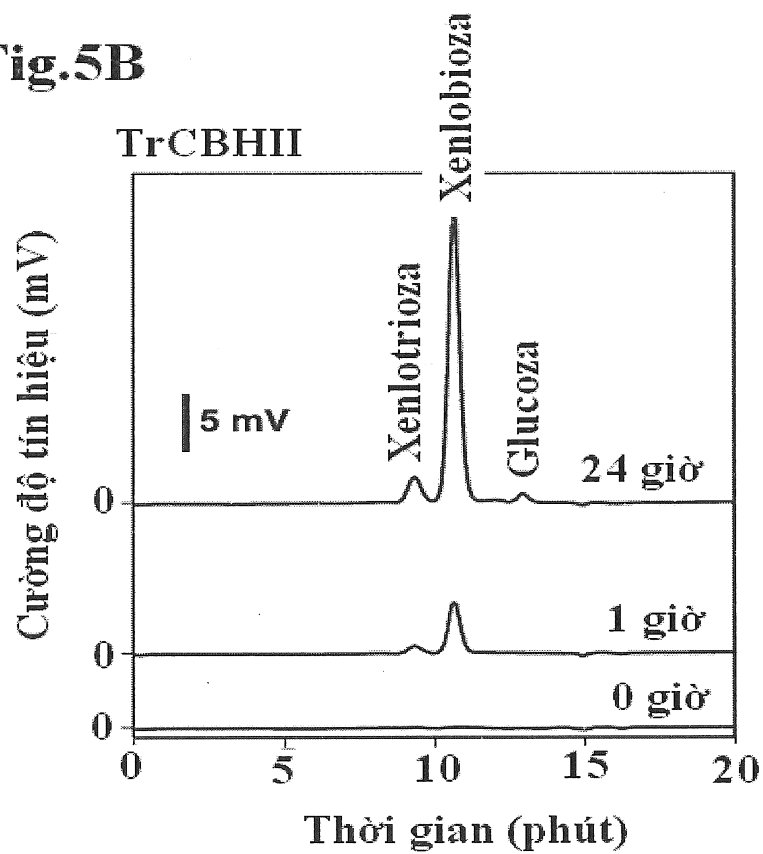
Fig.5A**Fig.5B**

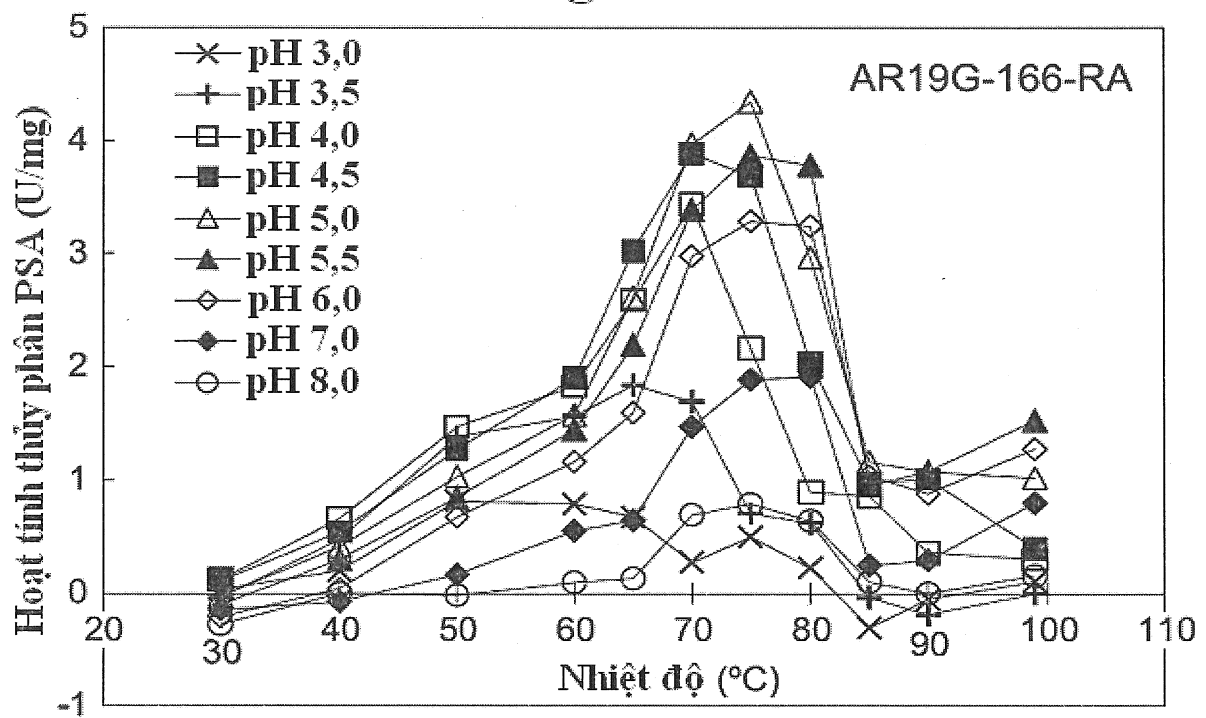
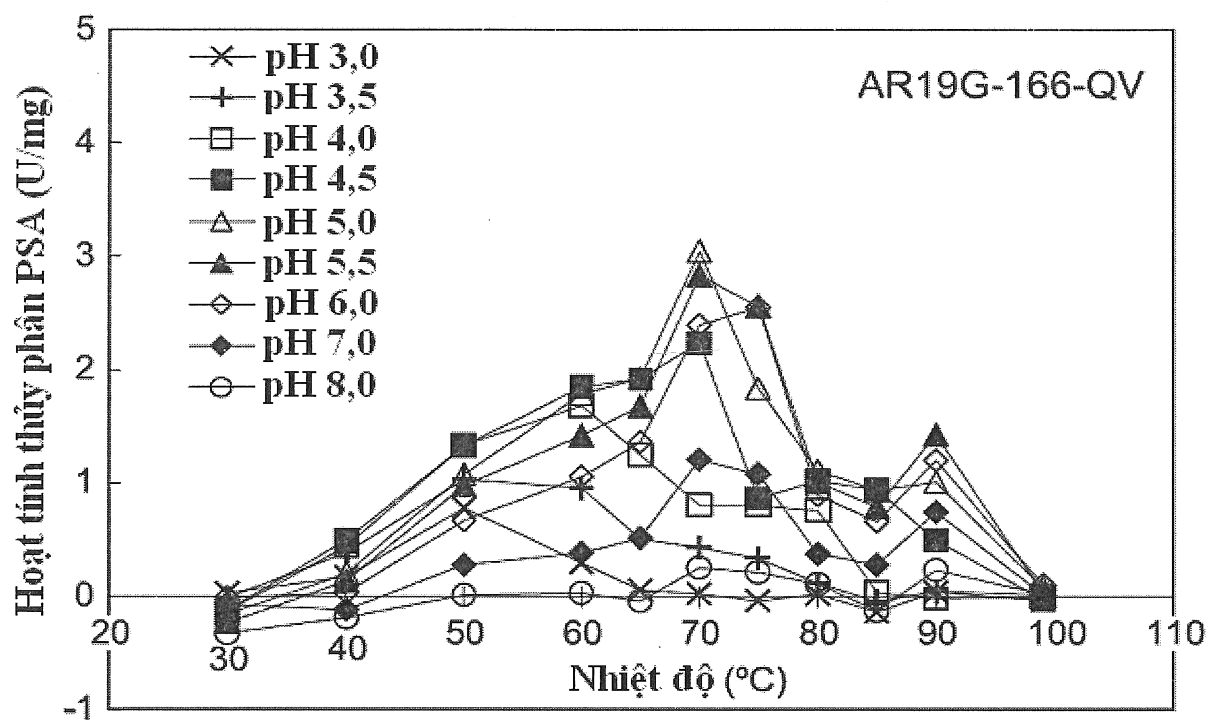
Fig.6A**Fig.6B**

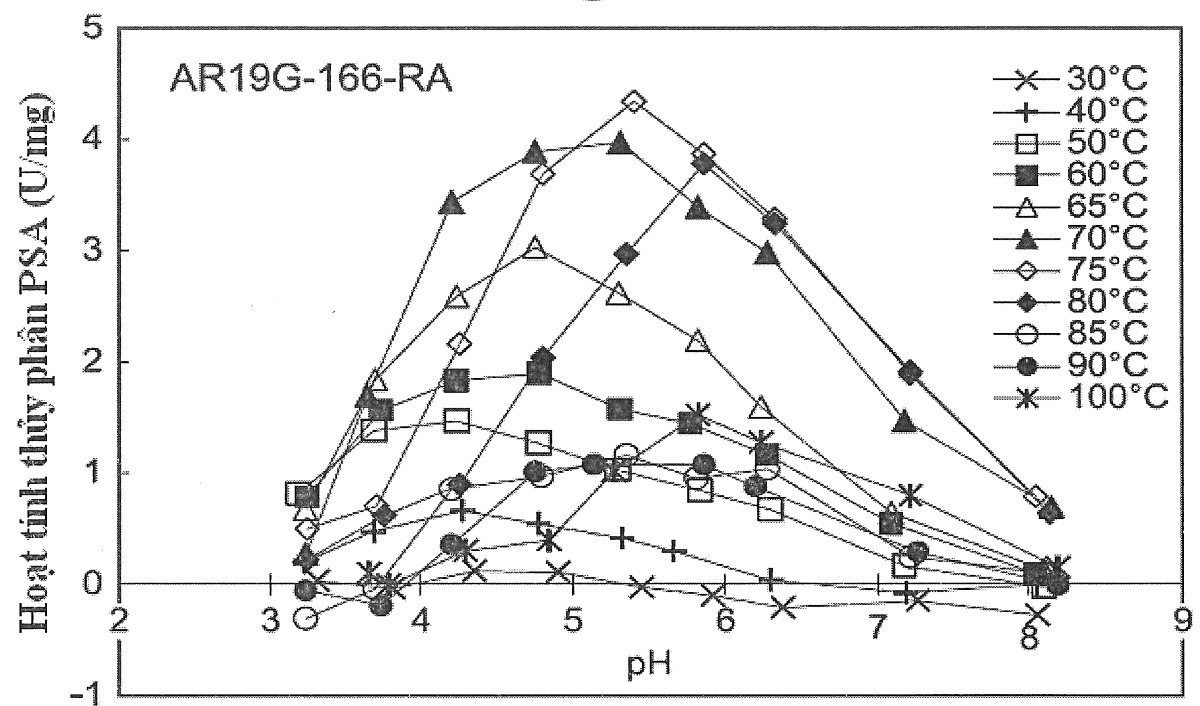
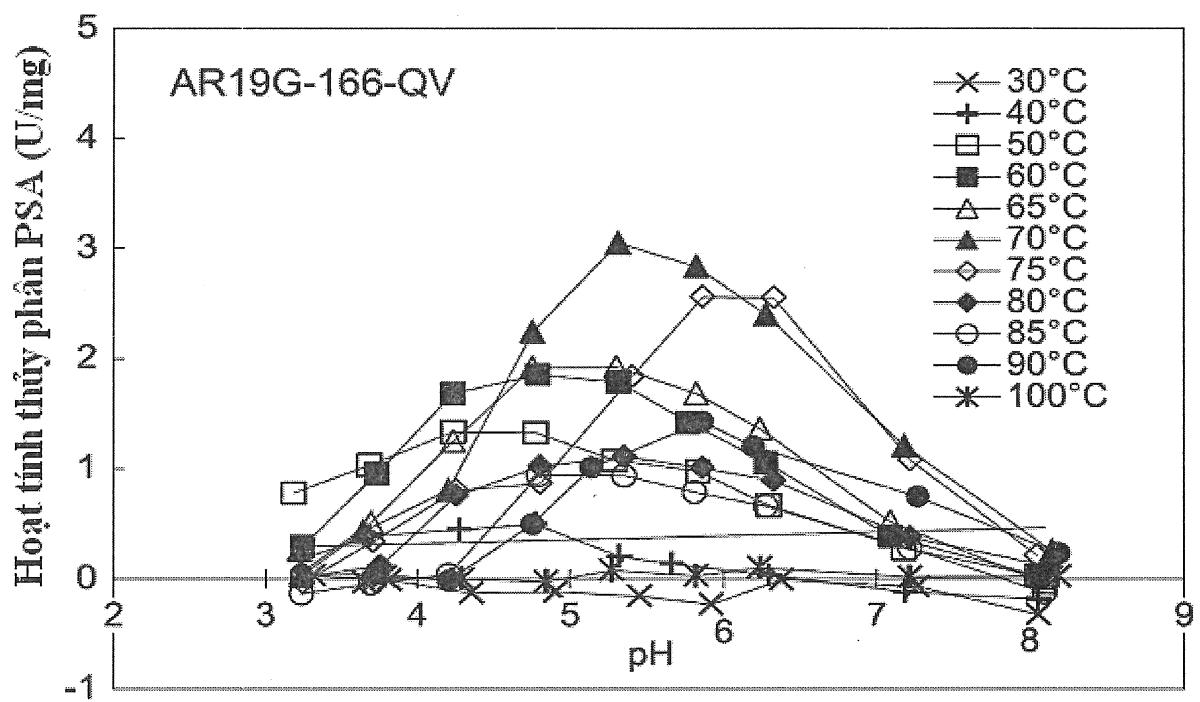
Fig.7A**Fig.7B**

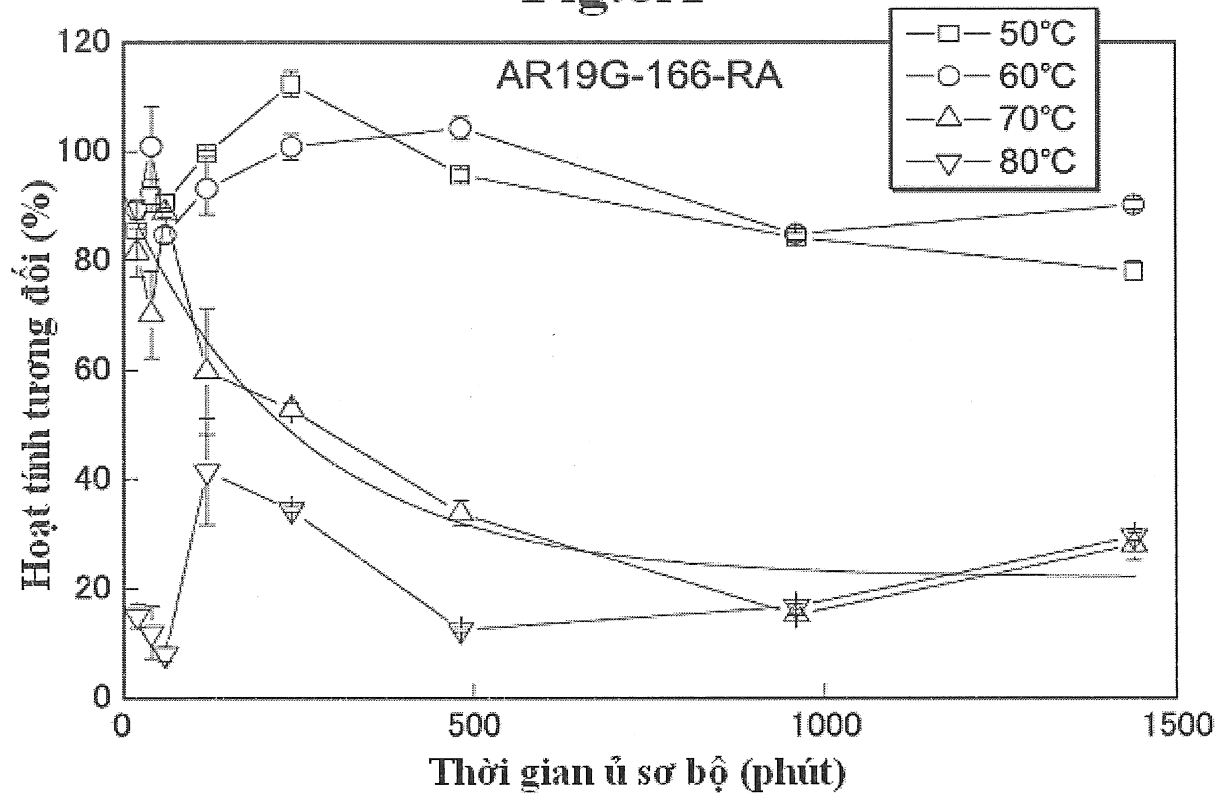
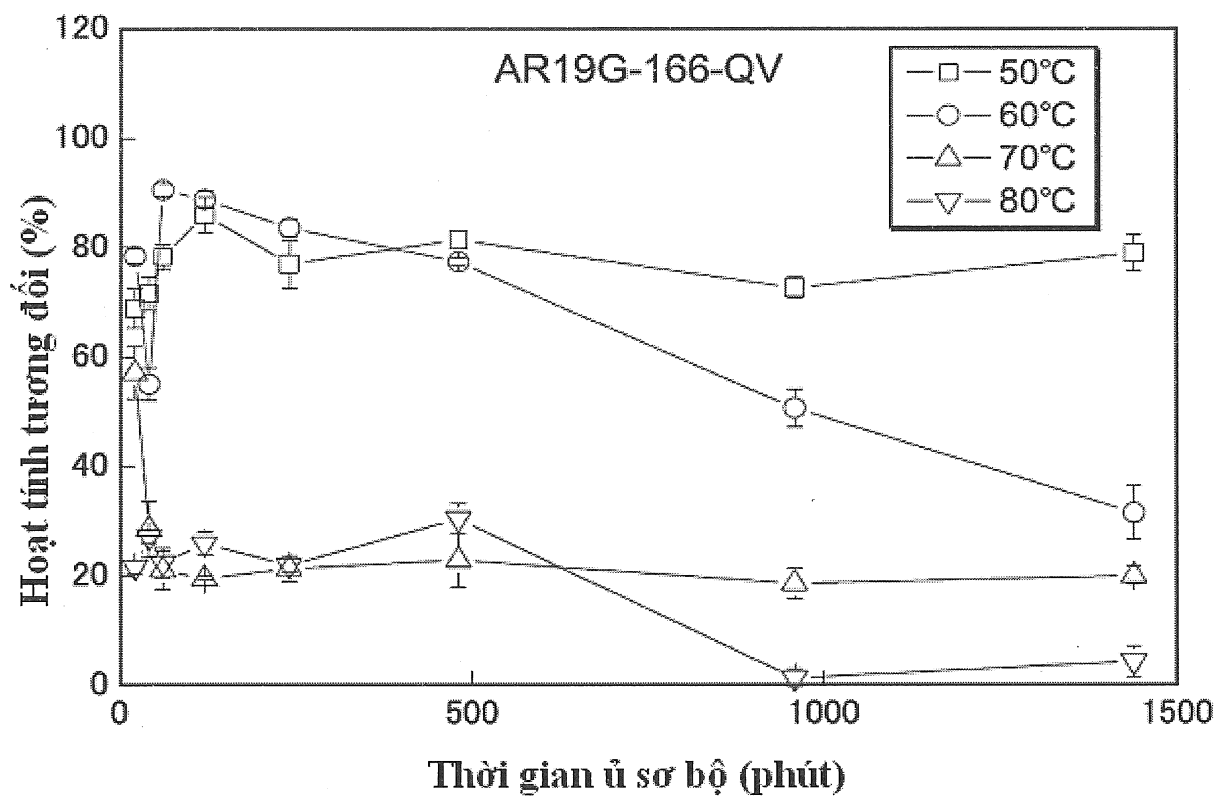
Fig.8A**Fig.8B**

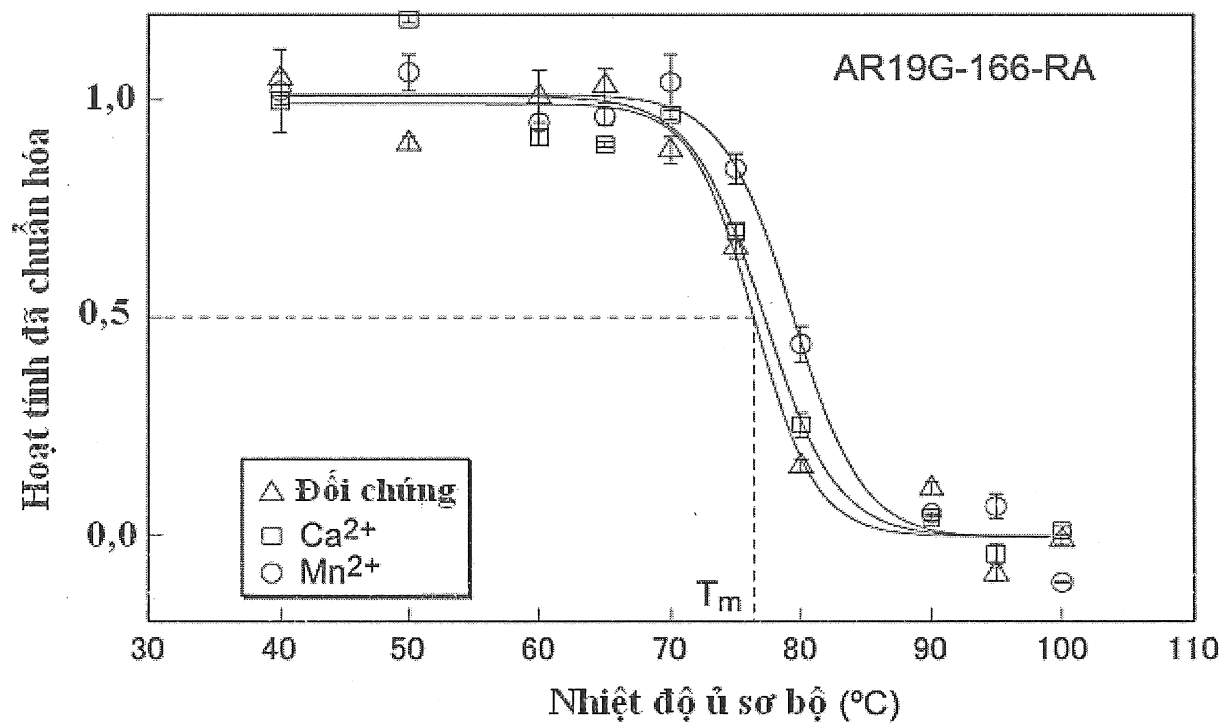
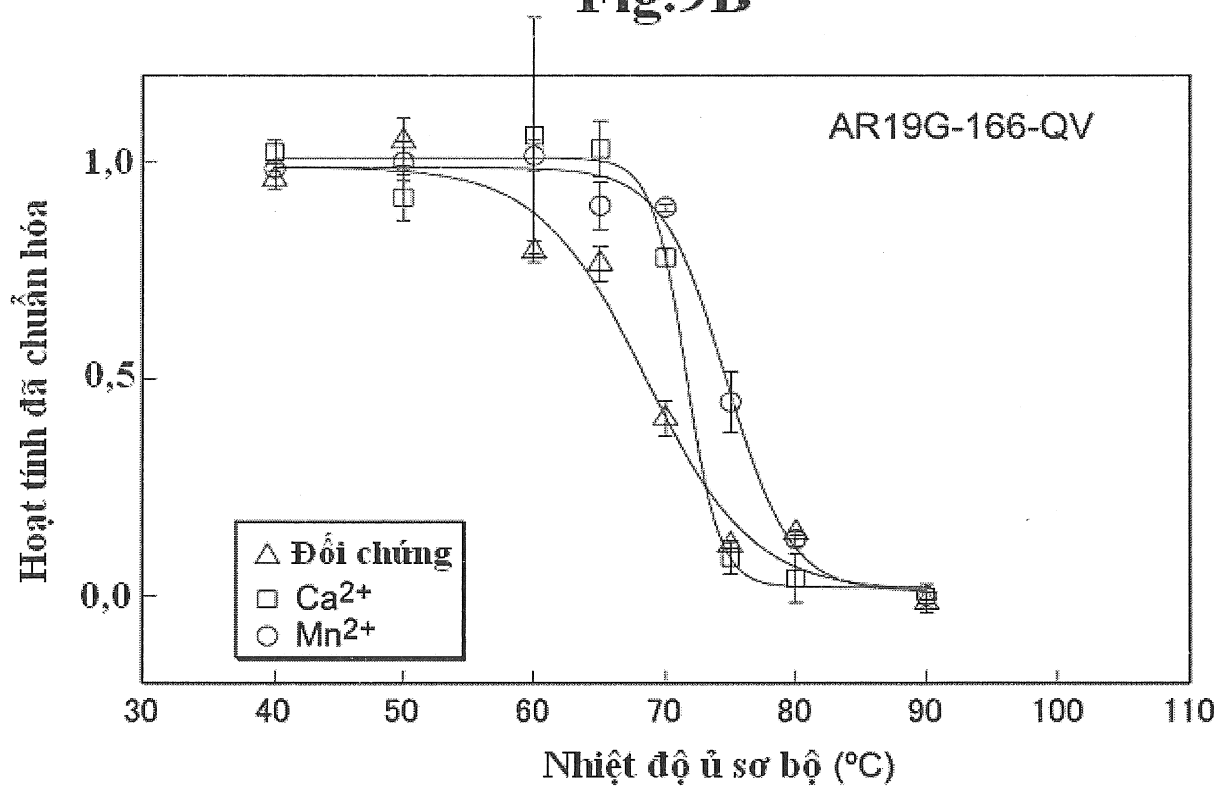
Fig.9A**Fig.9B**

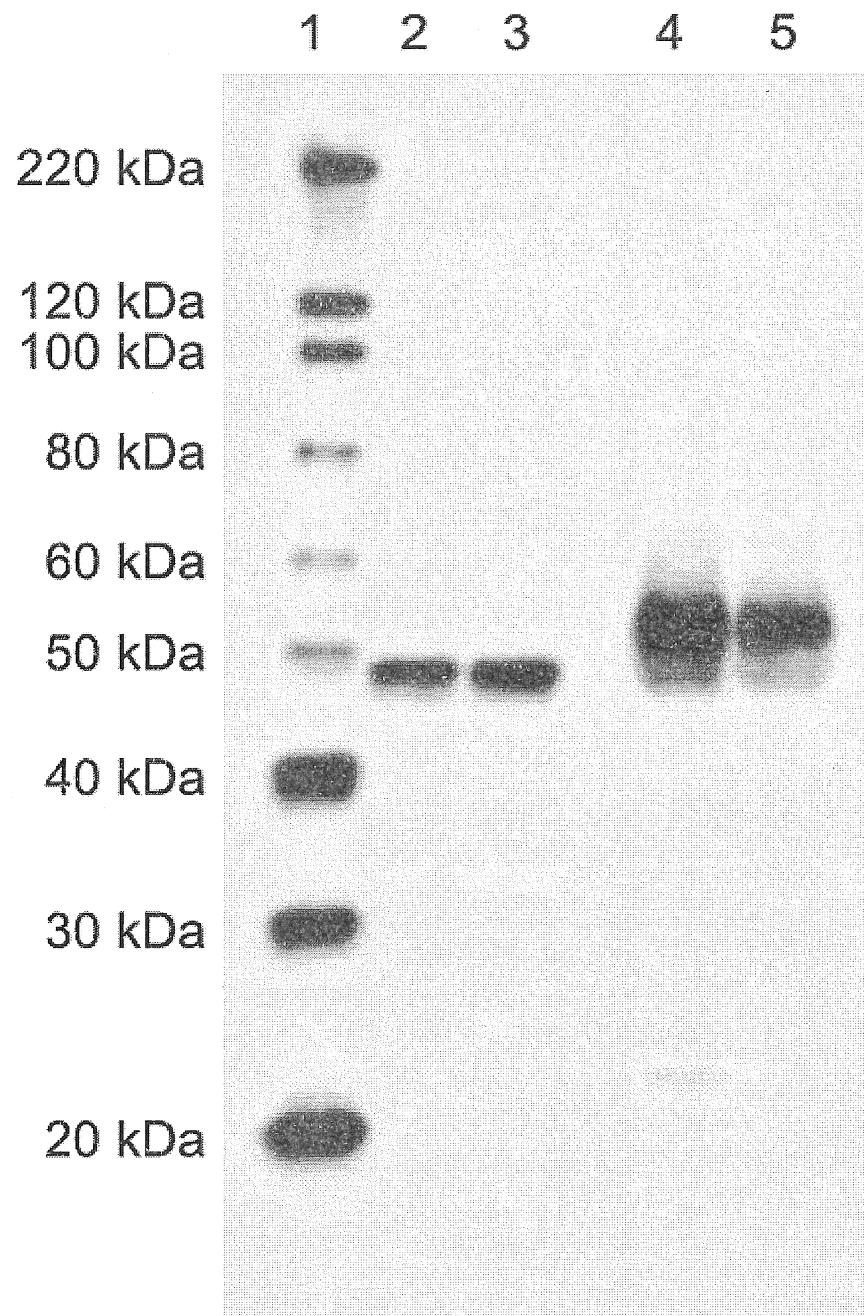
Fig.10

Fig.11A

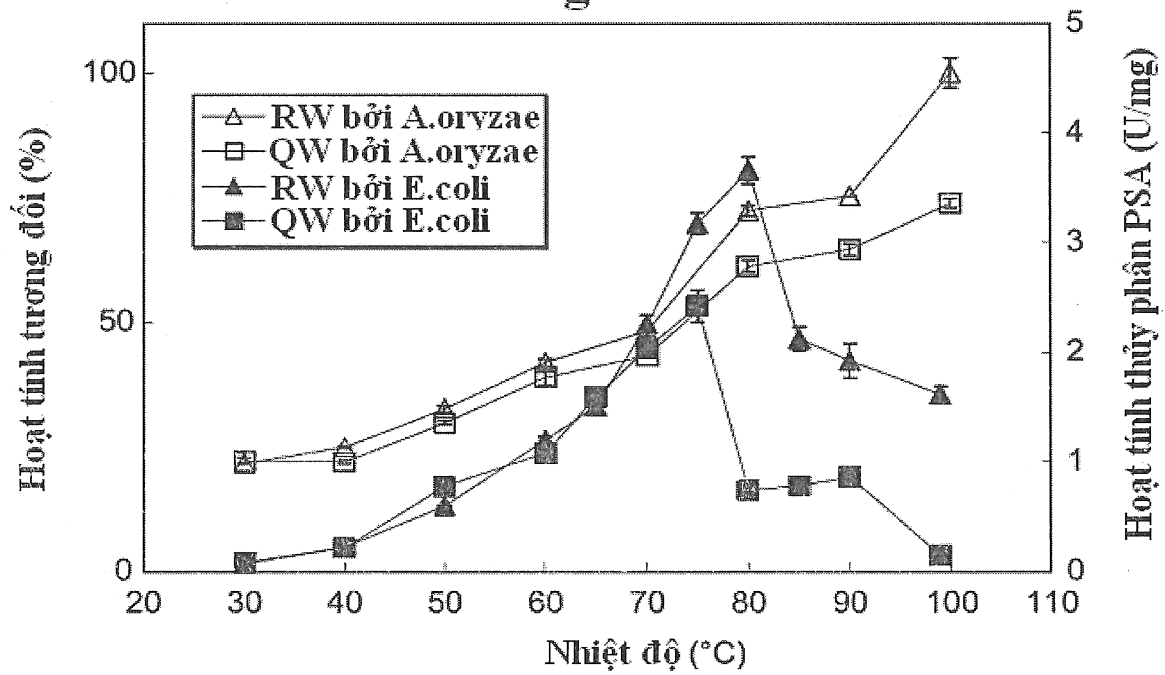


Fig.11B

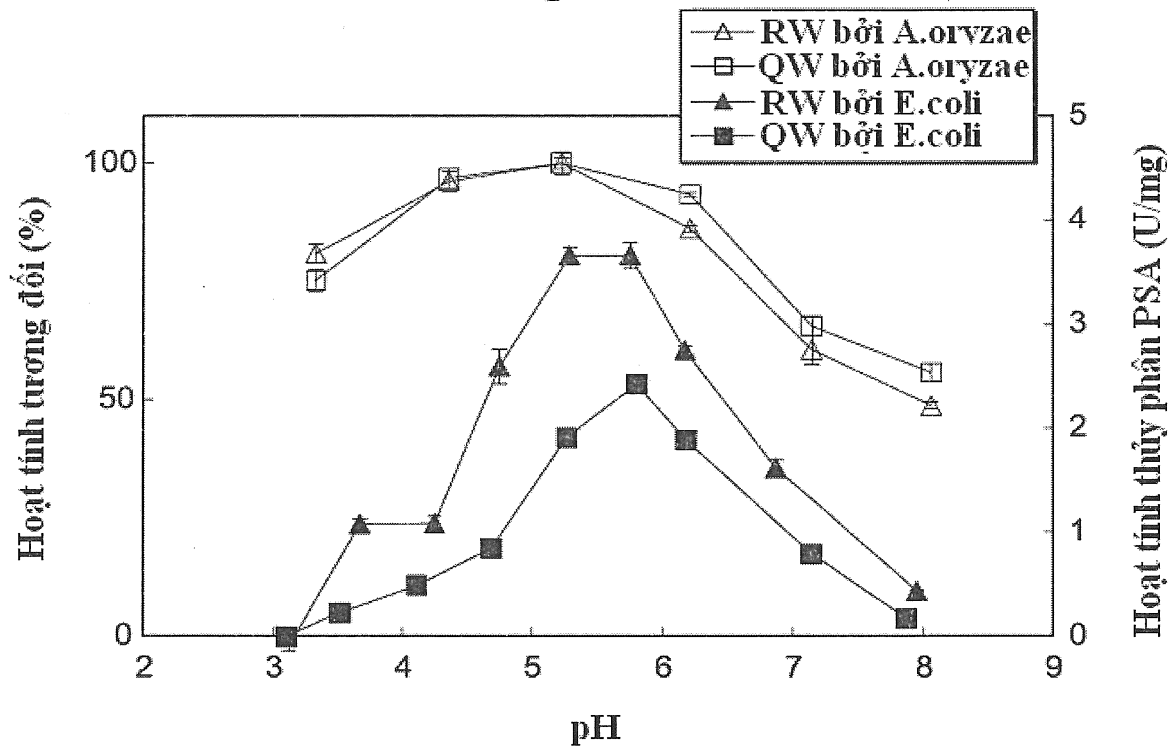


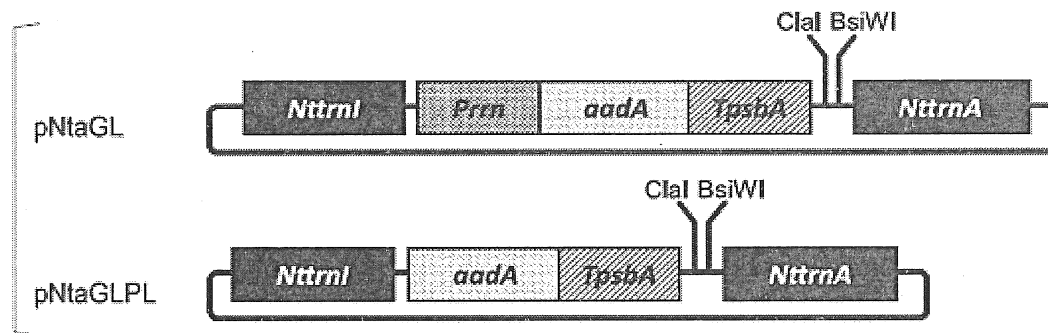
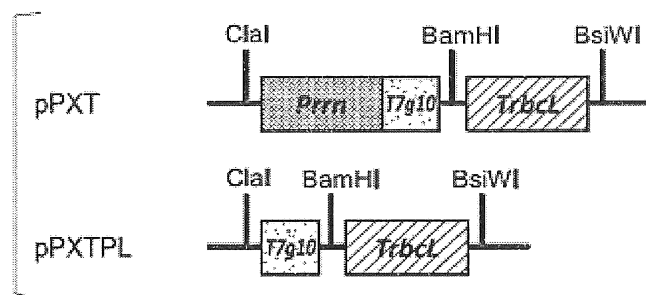
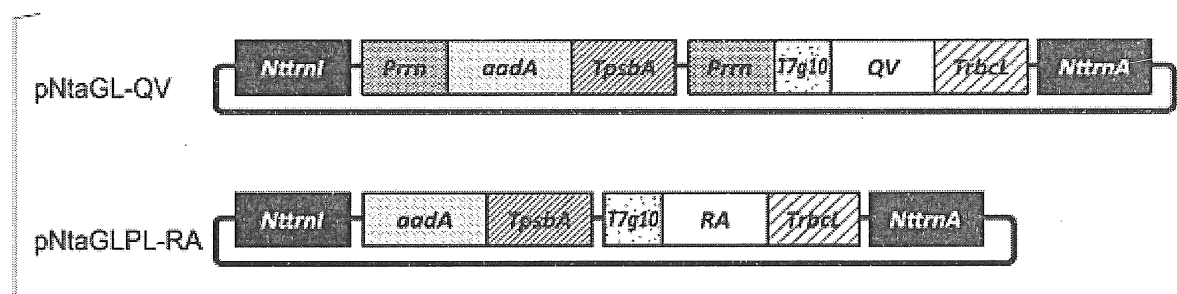
Fig.12A**Fig.12B****Fig.12C**

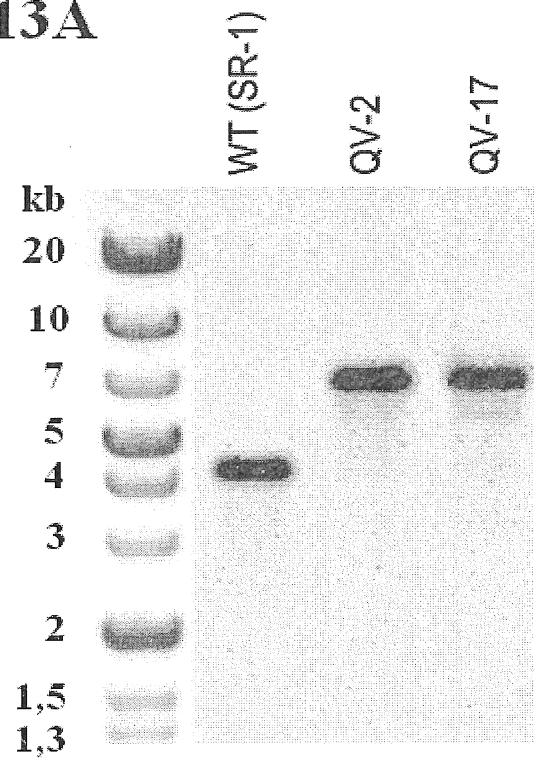
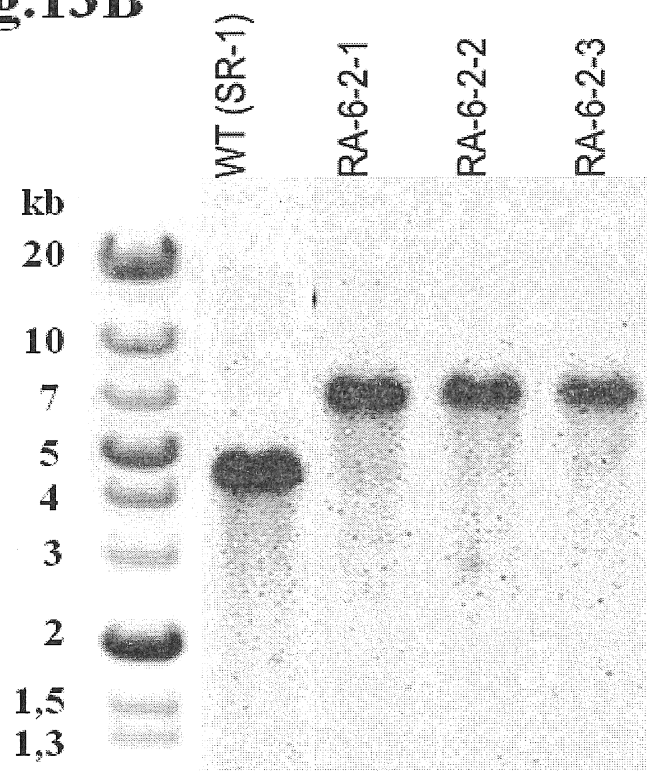
Fig.13A**Fig.13B**

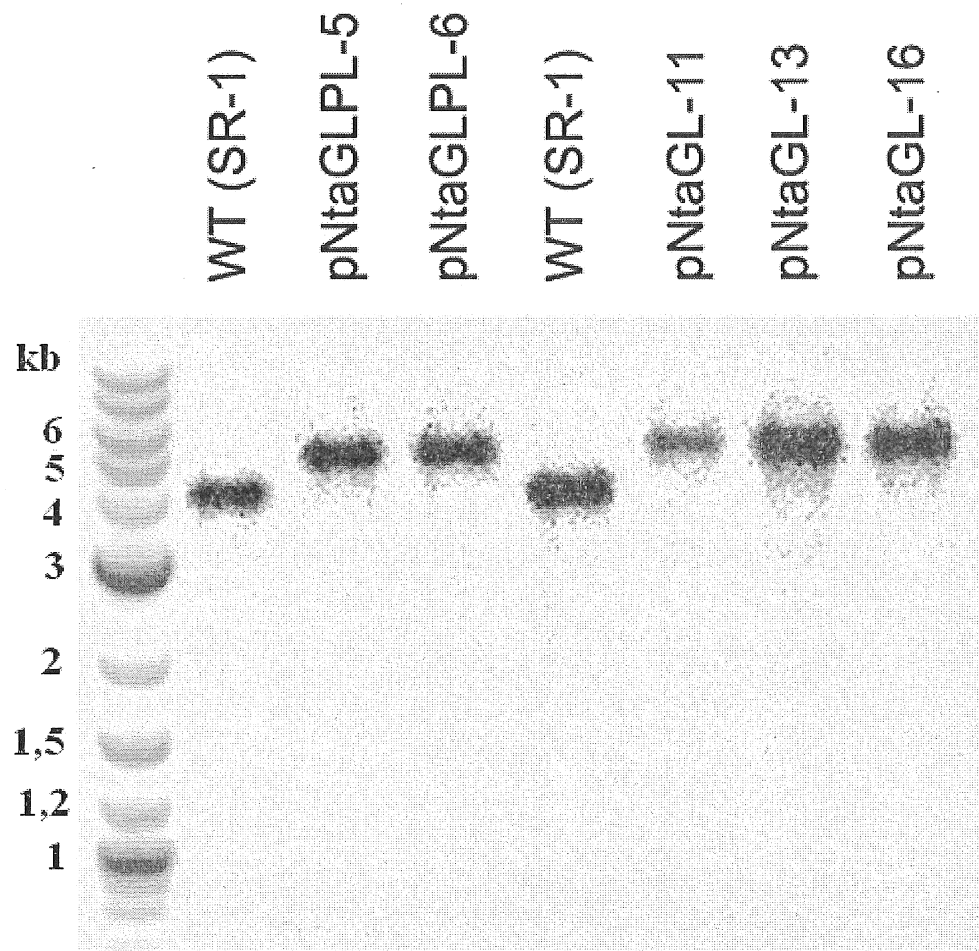
Fig.13C

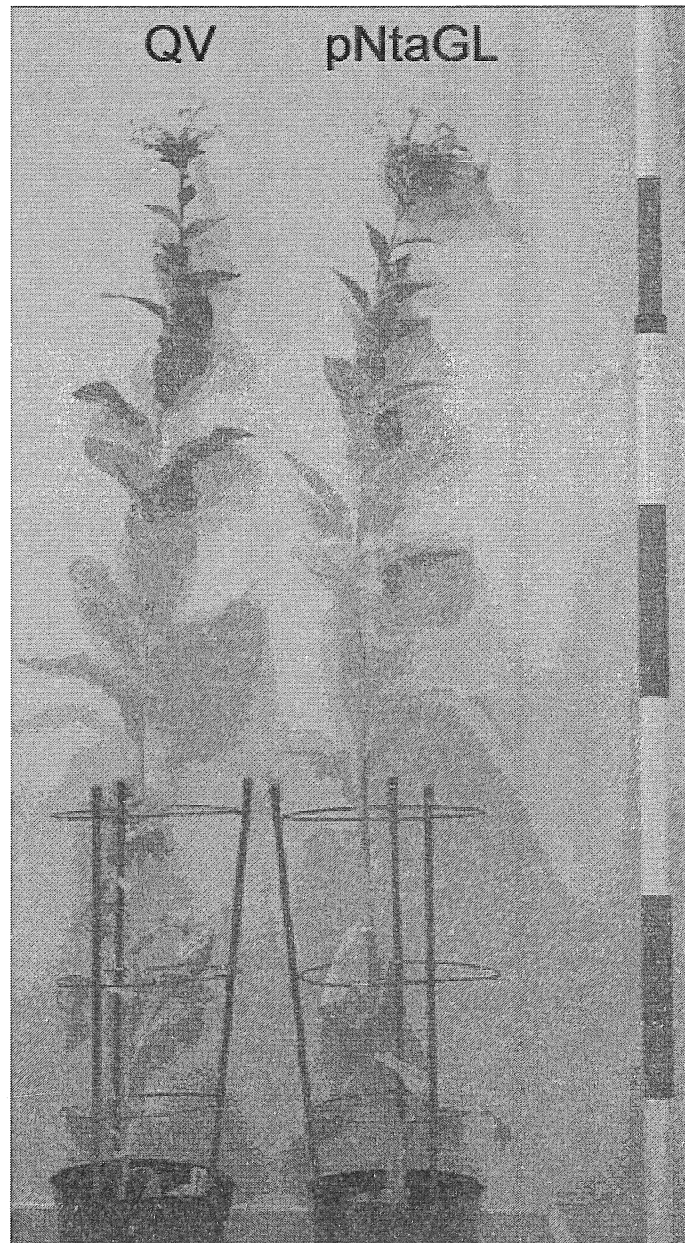
Fig.14A

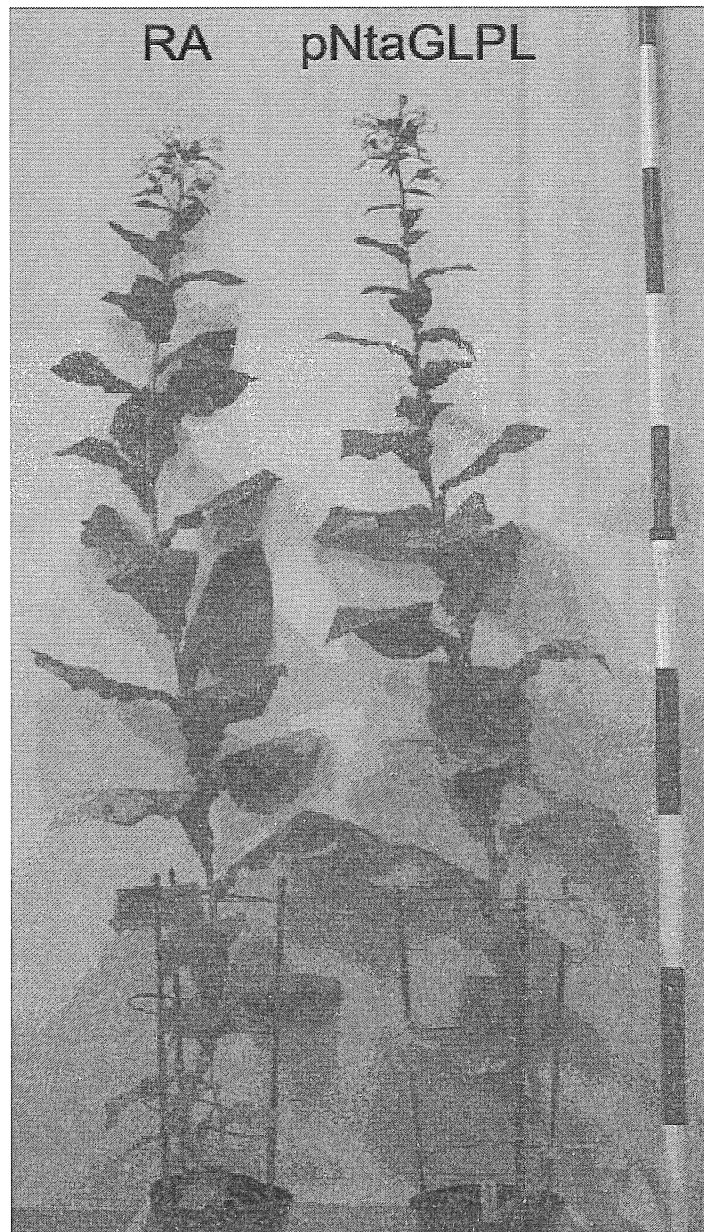
Fig.14B

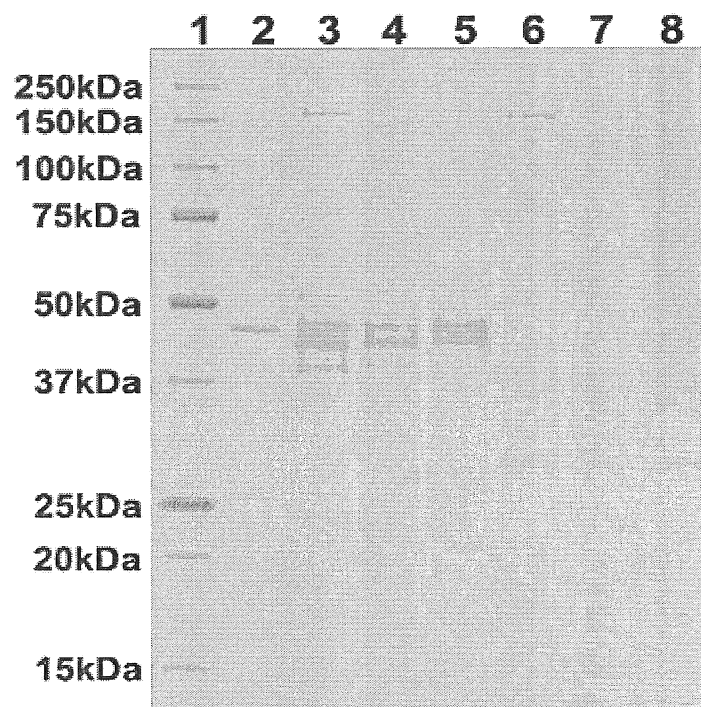
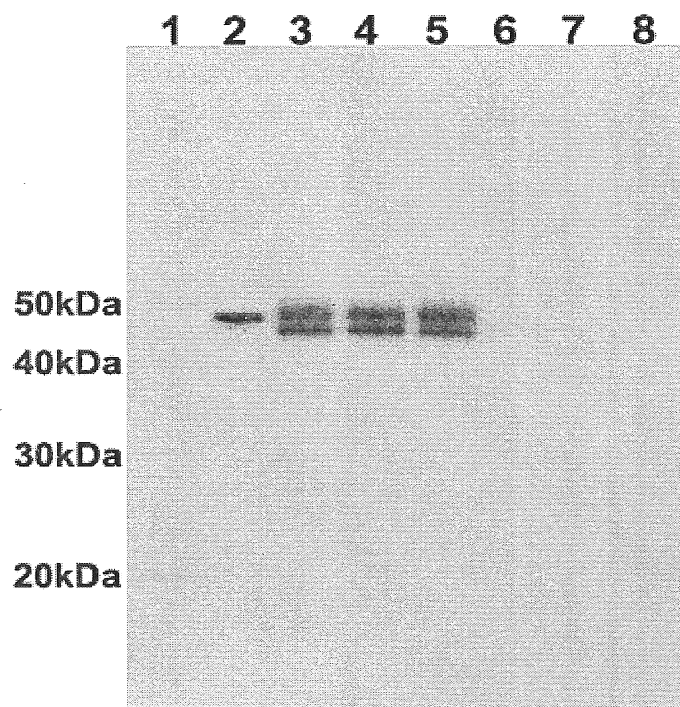
Fig.15A**Fig.15B**

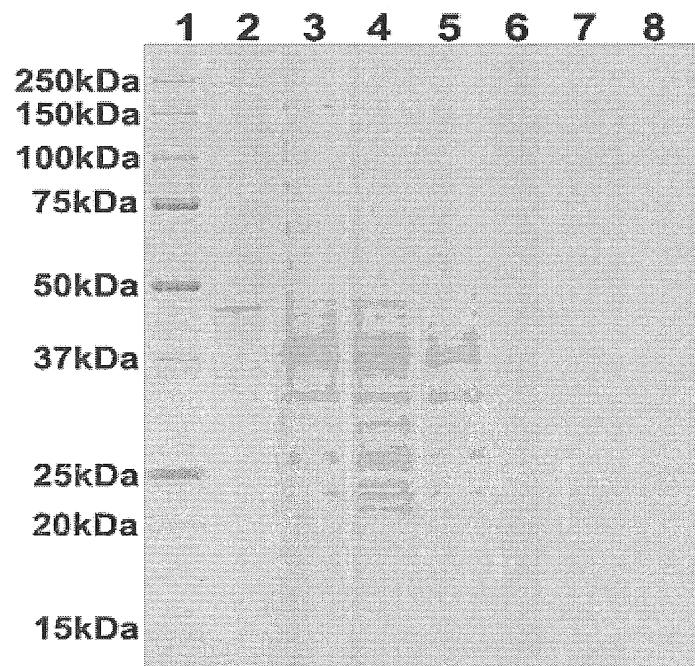
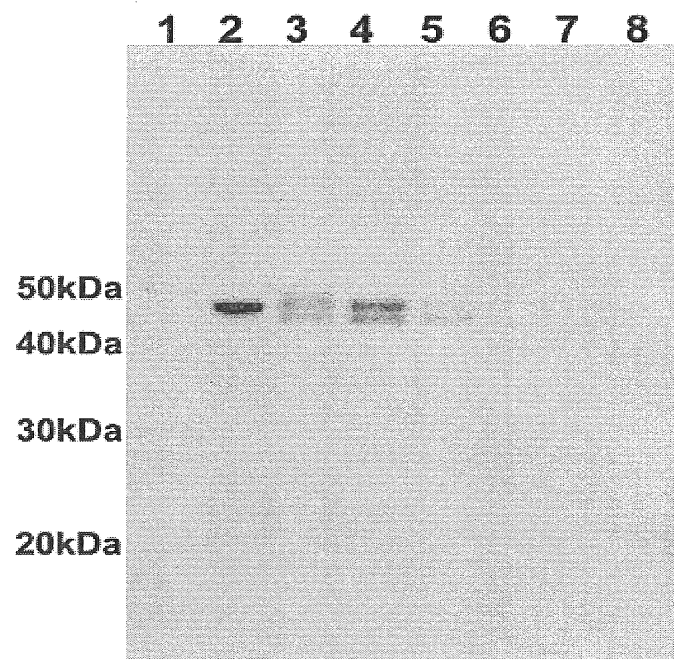
Fig.15C**Fig.15D**

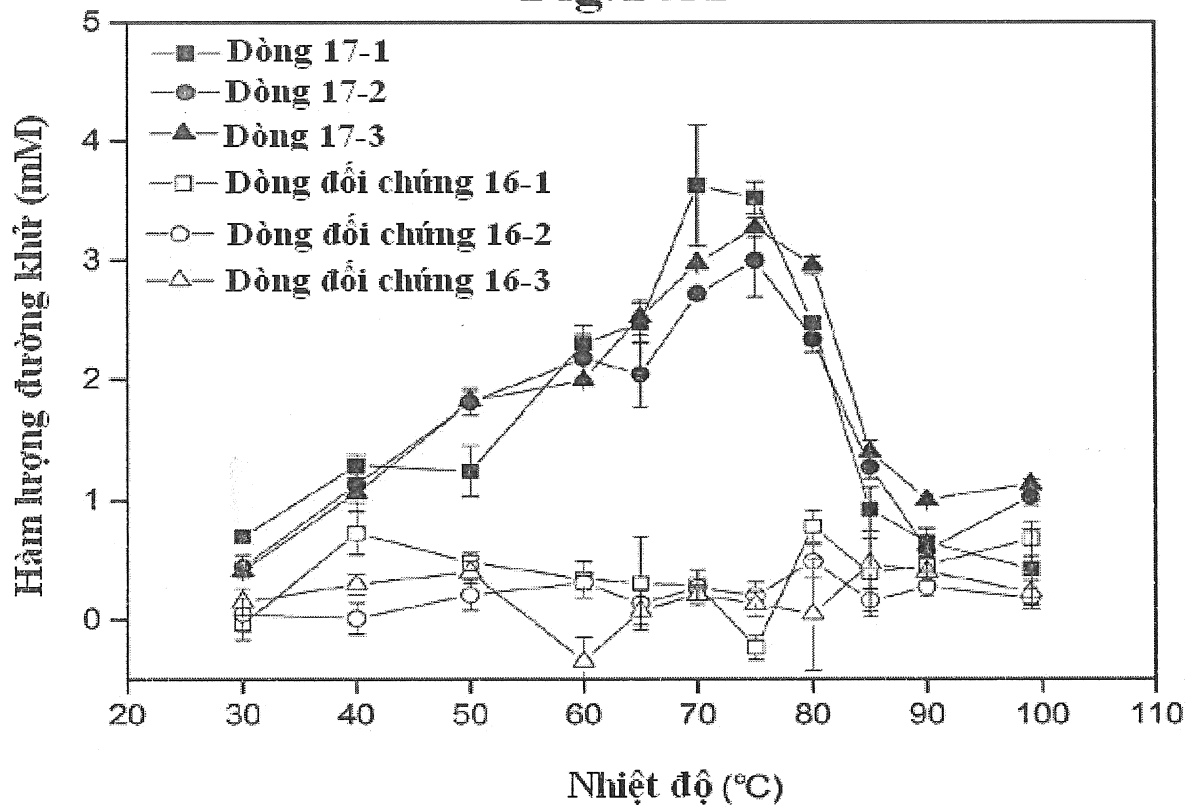
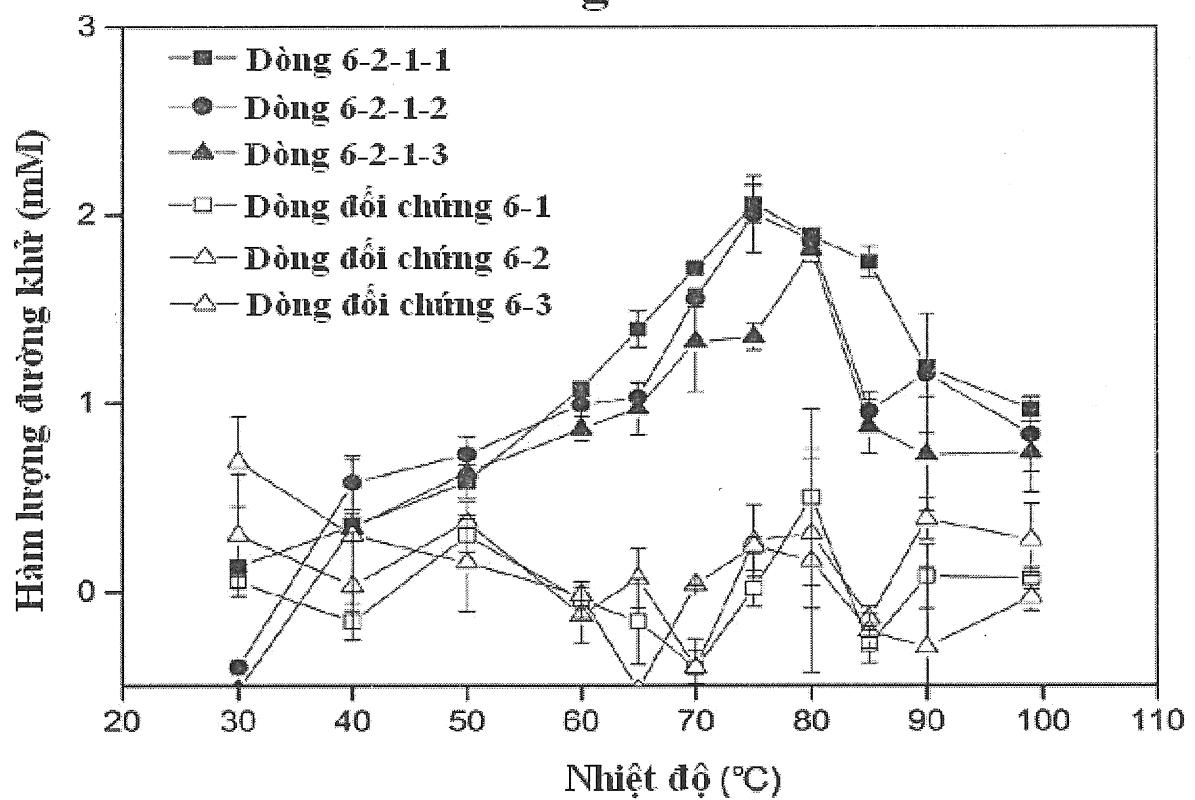
Fig.16A**Fig.16B**

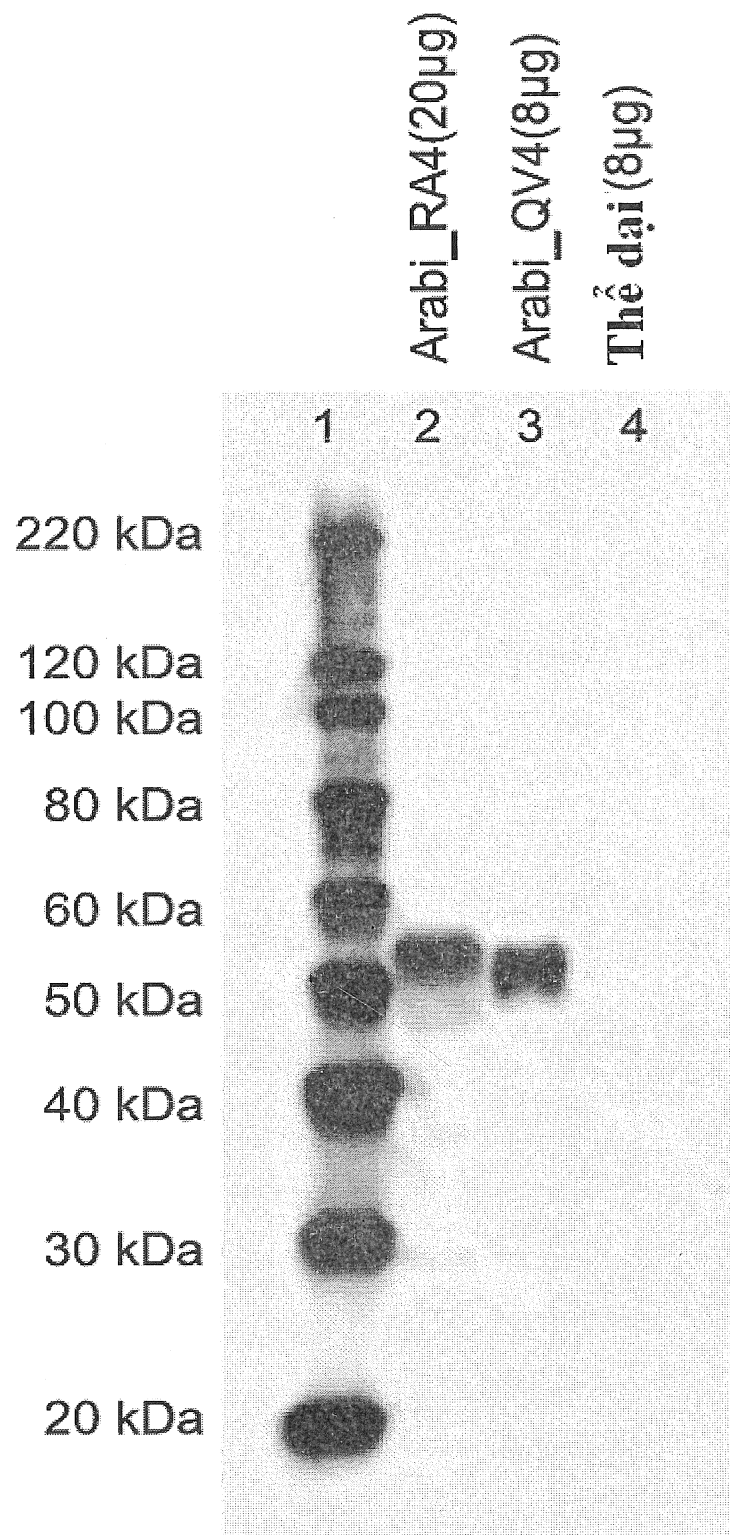
Fig.17

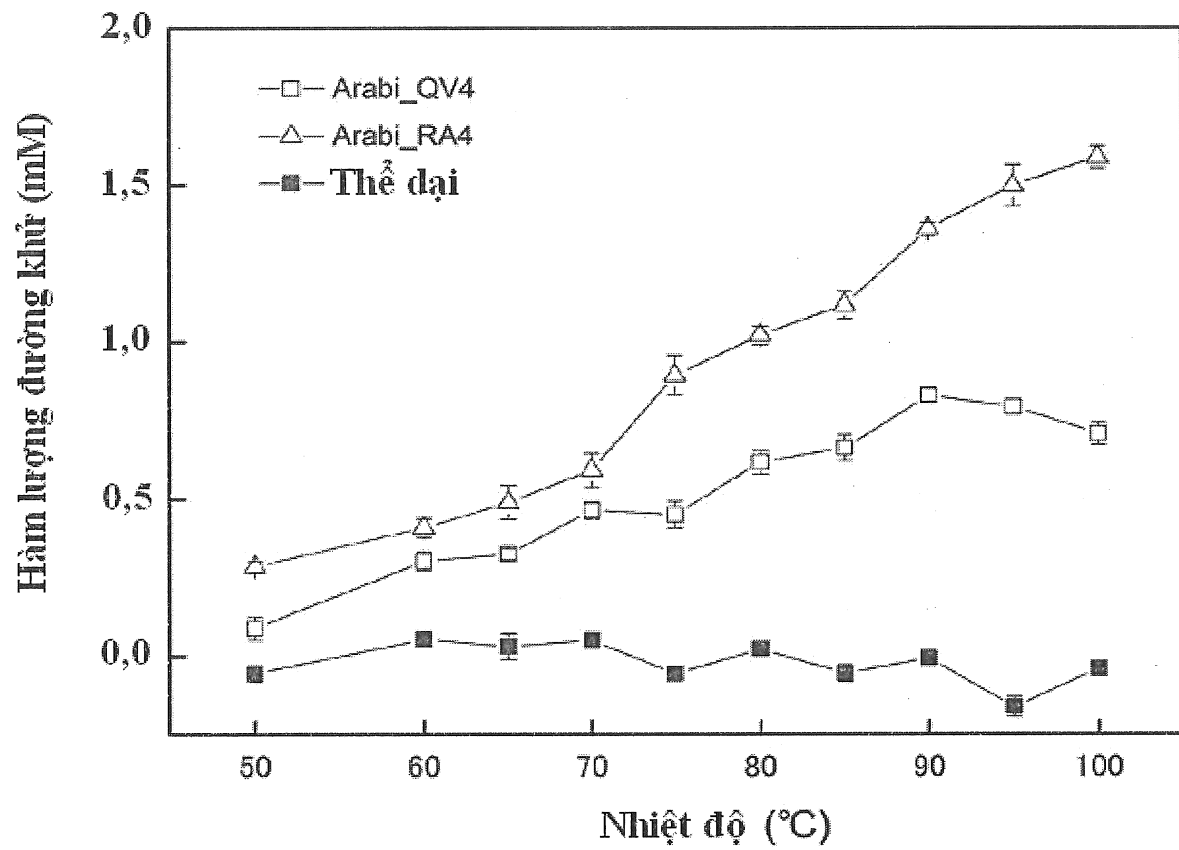
Fig.18

Fig.19

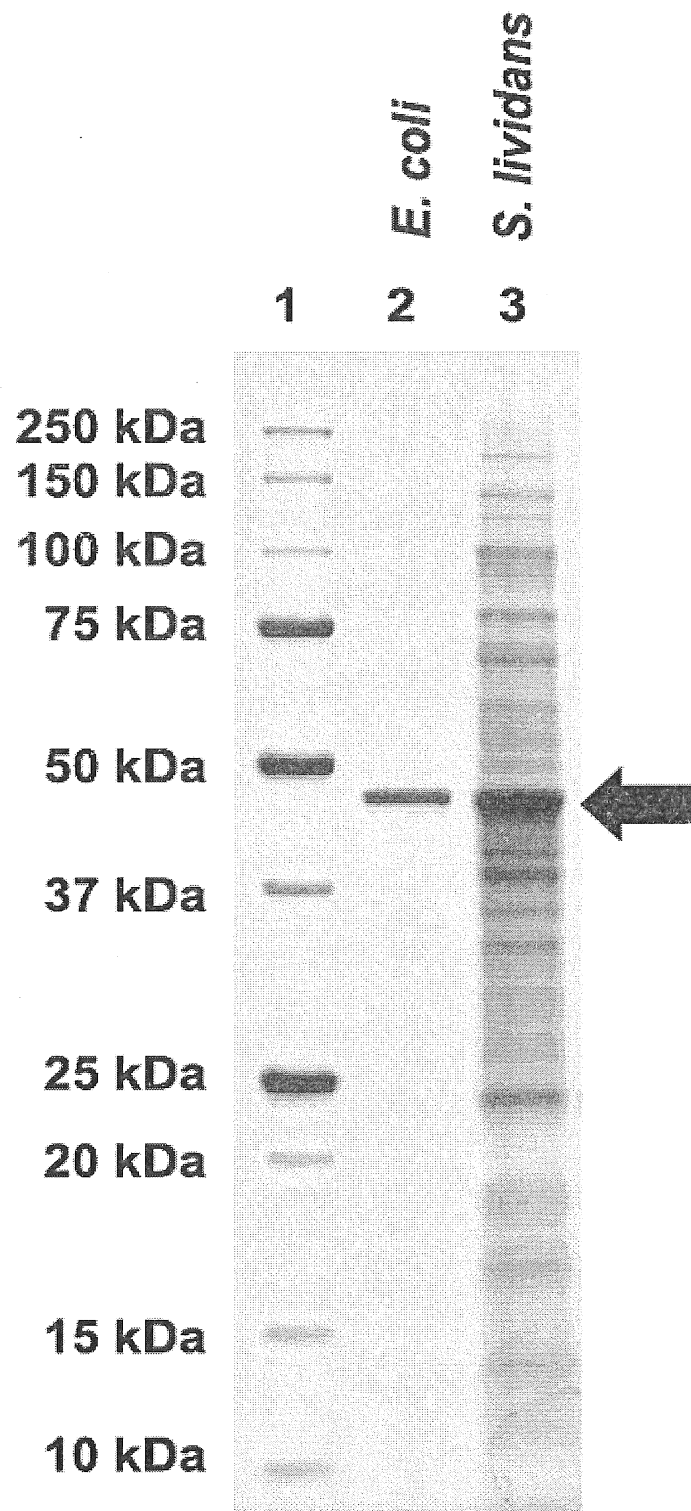


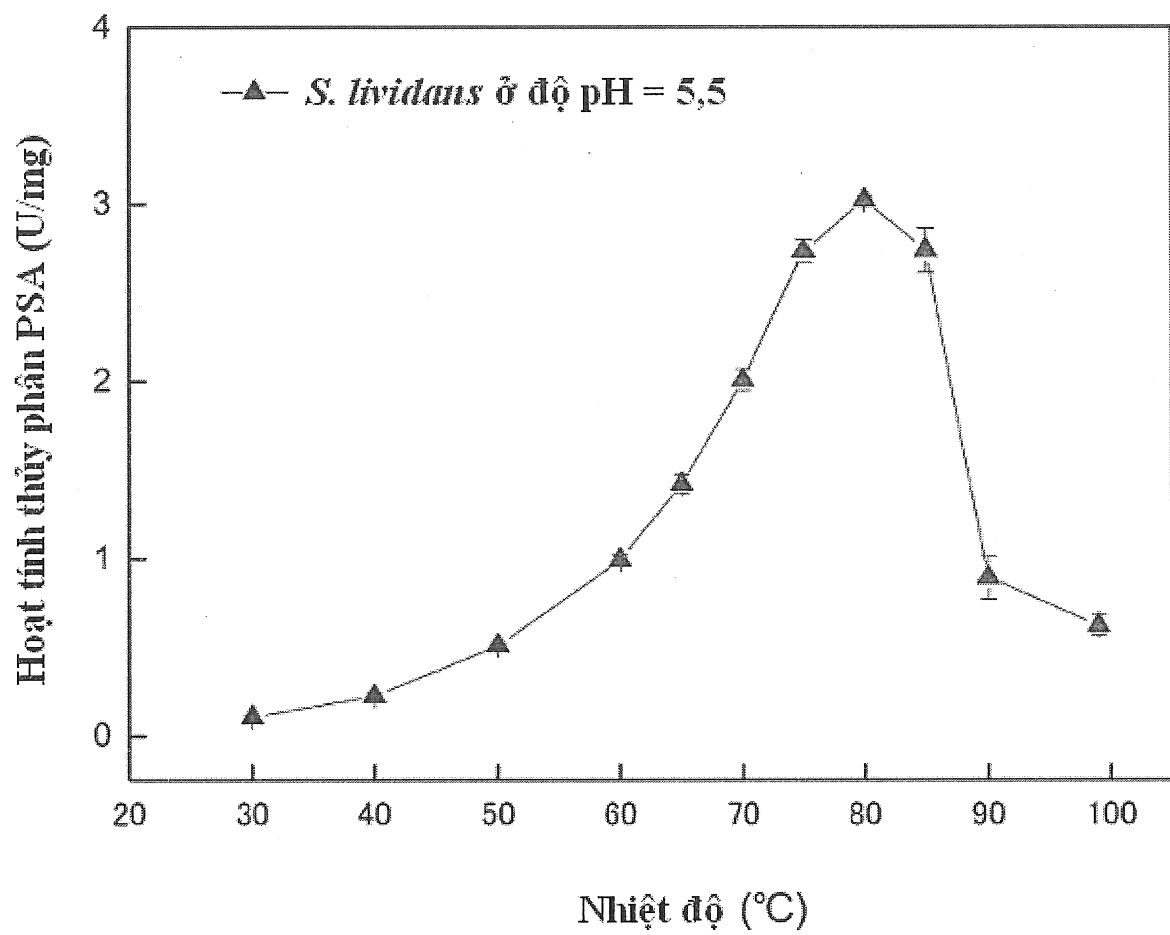
Fig.20

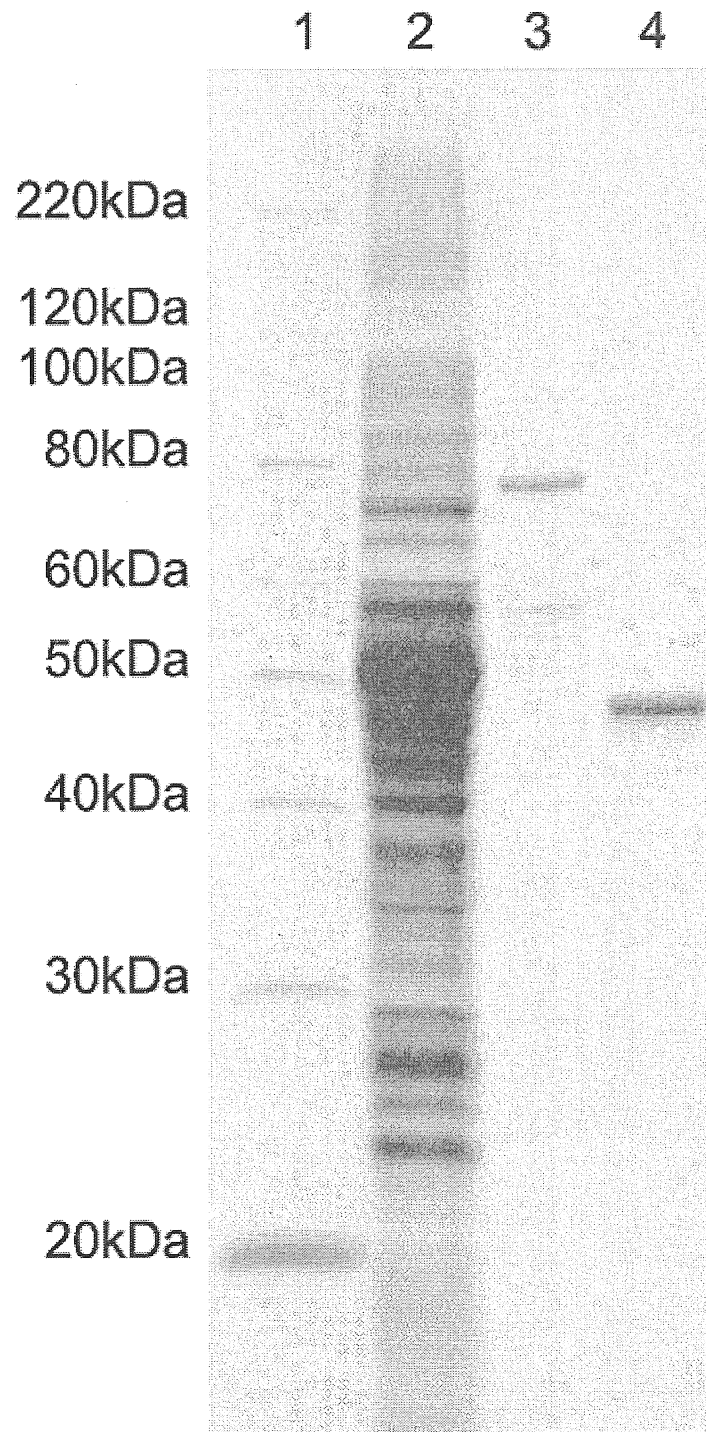
Fig.21A

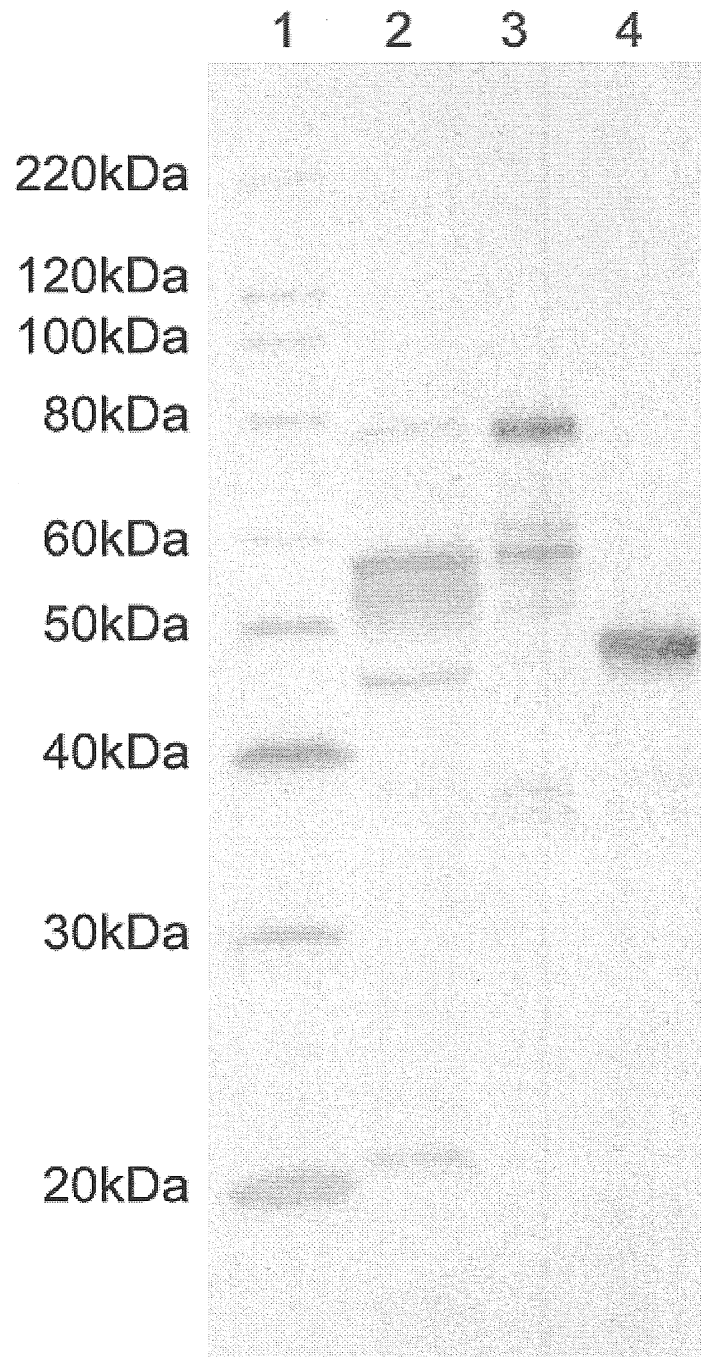
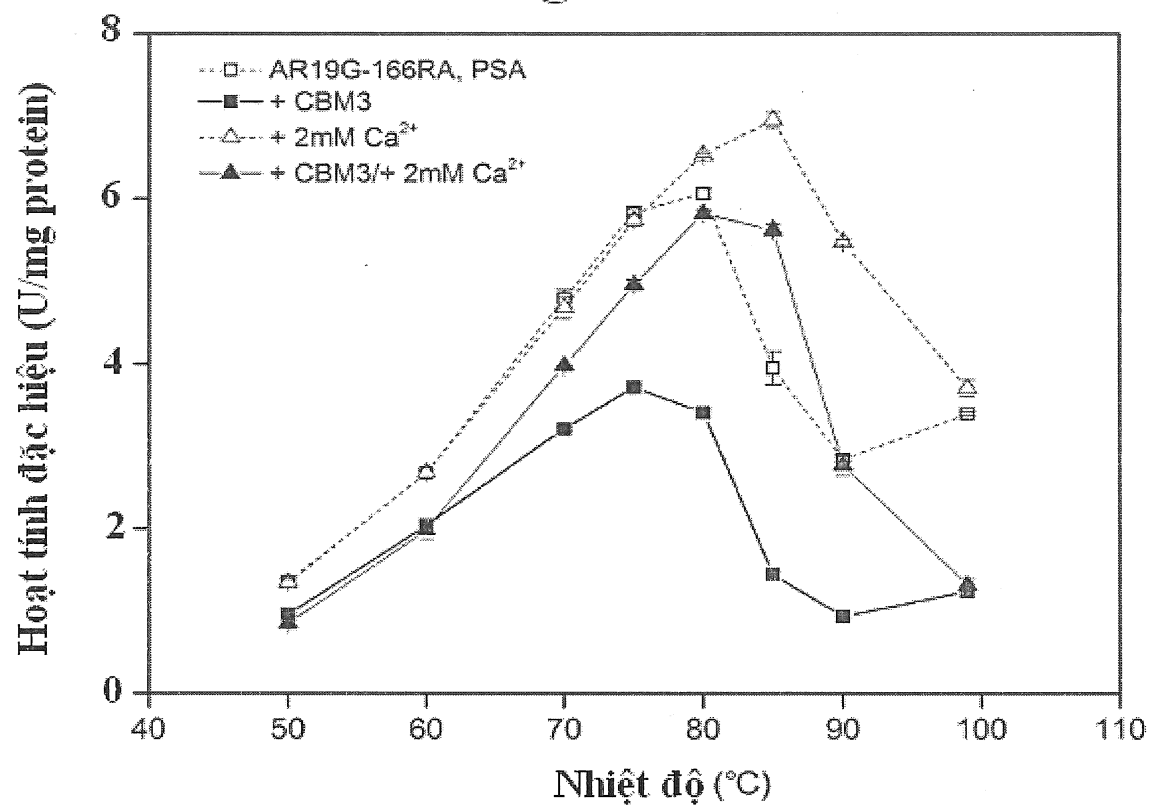
Fig.21B

Fig.22A**Fig.22B**