



(12) **BẢN MÔ TẢ GIẢI PHÁP HỮU ÍCH THUỘC BẰNG ĐỘC QUYỀN
GIẢI PHÁP HỮU ÍCH**

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

(11)



2-0004060

(51) **C12P 19/04; C12N 9/38; C12N 1/15; C12N**
2022.01 15/56

(13) **Y**

(21) 2-2024-00636

(22) 07/11/2019

(67) 1-2021-03048

(86) PCT/JP2019/043629 07/11/2019

(87) WO 2020/100706 22/05/2020

(30) 2018-212757 13/11/2018 JP

(45) 25/04/2025 445

(43) 27/09/2021 402A

(71) KABUSHIKI KAISHA YAKULT HONSHA (JP)

1-10-30, Kaigan, Minato-ku, Tokyo 1058660, Japan

(72) ISHIKAWA Eiji (JP); IKEDA Masakazu (JP); ANBE Minako (JP); HATANO Hiroshi (JP).

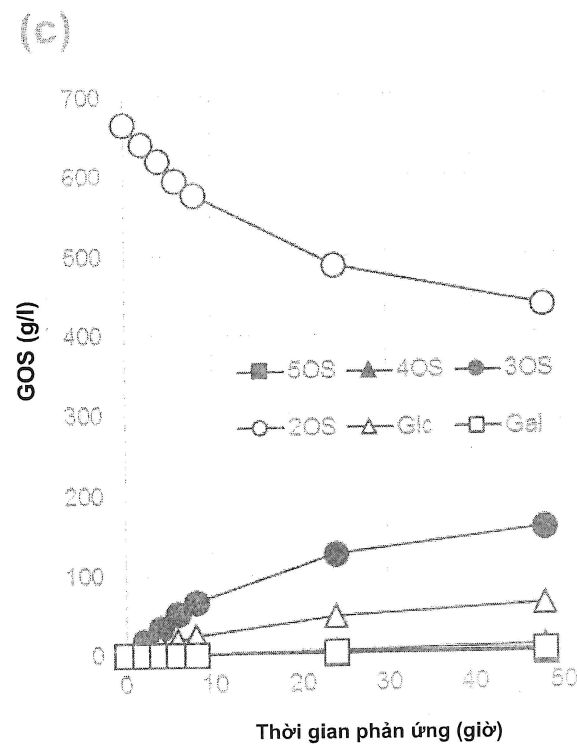
(74) Công ty Luật TNHH WINCO (WINCO LAW FIRM)

(54) **PHƯƠNG PHÁP SẢN XUẤT B-GALACTOSIDAZA ĐƯỢC TIẾT RA VÀ GEN B-
GALACTOSIDAZA**

(21) 2-2024-00636

(57) Giải pháp hữu ích đề cập đến phương pháp sản xuất β -galactosidaza được tiết ra, khác biệt ở chỗ, phương pháp này bao gồm bước gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra, và phương pháp sản xuất galactooligosacarit bằng cách sử dụng β -galactosidaza thu được bằng phương pháp nêu trên. Theo các phương pháp này, galactooligosacarit có thể được sản xuất dễ dàng.

Fig.14



Lĩnh vực kỹ thuật được đề cập

Giải pháp hữu ích đề cập đến phương pháp sản xuất β -galactosidaza được tiết ra, chất này được sử dụng dễ dàng trong quá trình sản xuất galactooligosacarit.

Tình trạng kỹ thuật của giải pháp hữu ích

β -galactosidaza đã được biết là có tác dụng xúc tác không chỉ đối với phản ứng thủy phân liên kết β -D-galactosit trong lactoza hoặc chất tương tự mà cả phản ứng trao đổi nhóm galactosyl, và được sử dụng trong quá trình sản xuất galactooligosacarit làm gia tăng chọn lọc giống *Bifidobacterium* trong ruột.

Người nộp đơn giải pháp hữu ích này đã thông báo về kỹ thuật sản xuất galactooligosacarit bằng cách sử dụng β -galactosidaza có nguồn gốc từ chủng đột biến của *Sporobolomyces singularis* sản sinh quá mức, đó là nấm men *Basidiomycetes* (tài liệu 1).

Tuy nhiên, do β -galactosidaza được sử dụng trong kỹ thuật này là protein không tiết (gắn kết với thành tế bào), nên cần sử dụng β -galactosidaza trong phản ứng dưới dạng dịch cô tế bào chứa các tế bào *Sporobolomyces singularis* để tạo ra β -galactosidaza.

Do dịch cô tế bào bao gồm các tế bào còn sống, dịch này dễ bị thoái biến, và ngoài ra, dịch này thu được chỉ bằng cách cô đặc các tế bào nên hoạt tính đặc hiệu của nó thấp và còn xuất hiện vấn đề là mảnh vụn tế bào phân tán vào dung dịch phản ứng galactooligosacarit, điều này dẫn đến làm tăng chi phí tinh chế hoặc yếu tố tương tự.

Tài liệu 1: JP-A-2006-223268

Bản chất kỹ thuật của giải pháp hữu ích

Một mục đích của giải pháp hữu ích là đề xuất phương pháp sản xuất β -galactosidaza để giải quyết các vấn đề nêu trên và chất này được sử dụng dễ dàng trong quá trình sản xuất galactooligosacarit.

Nhờ nhiều nghiên cứu chuyên sâu để giải quyết các vấn đề nêu trên, các tác giả giải pháp hữu ích đã phát hiện được rằng β -galactosidaza được tiết ra được tạo ra bằng cách gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae* và β -galactosidaza được tiết ra được sử dụng dễ dàng trong quá trình sản xuất galactooligosacarit, và do đó giải pháp hữu ích được hoàn thành.

Cụ thể, giải pháp hữu ích đề cập đến phương pháp sản xuất β -galactosidaza được tiết ra, khác biệt ở chỗ, phương pháp này bao gồm bước gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra.

Ngoài ra, giải pháp hữu ích đề cập đến gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes*, gen này có trình tự được thể hiện bởi các trình tự SEQ ID NO: 7, 13, hoặc 19.

Hơn nữa, giải pháp hữu ích đề cập đến thể biến nạp của nấm *Aspergillus oryzae*, khác biệt ở chỗ, gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* được gắn vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra.

Hơn nữa, giải pháp hữu ích đề cập đến phương pháp sản xuất galactooligosacarit, khác biệt ở chỗ, phương pháp này bao gồm bước cho β -galactosidaza thu được bằng phương pháp sản xuất β -galactosidaza nêu trên phản ứng trên cơ chất chứa ít nhất là lactoza.

Ngoài ra, giải pháp hữu ích đề cập đến β -galactosidaza được tiết ra thu được bằng cách gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae*, và nuôi cấy nấm *Aspergillus oryzae* tạo thành.

Hiệu quả đạt được của giải pháp hữu ích

Theo phương pháp sản xuất β -galactosidaza được tiết ra của giải pháp hữu ích, có thể thu được β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* dưới dạng β -galactosidaza được tiết ra.

Do đó, β -galactosidaza thu được bằng phương pháp sản xuất β -galactosidaza được tiết ra của giải pháp hữu ích có hoạt tính β -galactosidaza cao và cũng có tính chịu nhiệt cao, và ngoài ra, enzym này được tách và tinh chế dễ dàng, và được sử dụng dễ dàng trong quá trình sản xuất galactooligosacarit.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính bằng cách sử dụng chủng SsGal.

Fig.2 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính bằng cách sử dụng chủng SmGal.

Fig.3 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính bằng cách sử dụng chủng RmGal.

Fig.4 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính bằng cách sử dụng chủng SeGal.

Fig.5 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính (bất hoạt bằng nhiệt) bằng cách sử dụng chủng SsGal.

Fig.6 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính (bất hoạt bằng nhiệt) bằng cách sử dụng chủng SmGal.

Fig.7 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính (bất hoạt bằng nhiệt) bằng cách sử dụng chủng RmGal.

Fig.8 là hình ảnh thể hiện kết quả của phương pháp SDS-PAGE và kết quả xác định hoạt tính (bất hoạt bằng nhiệt) bằng cách sử dụng chủng SeGal.

Fig.9 là hình ảnh thể hiện kết quả của thử nghiệm xác định hoạt tính sau khi xử lý nhiệt bằng cách sử dụng chủng SeGal và chủng bố-mẹ.

Fig.10 là hình ảnh thể hiện thời gian phản ứng và nồng độ đường trong dung dịch trong quá trình sản xuất galactooligosacarit bằng cách sử dụng chủng SsGal.

Fig.11 là hình ảnh thể hiện thời gian phản ứng và nồng độ đường trong dung dịch trong quá trình sản xuất galactooligosacarit bằng cách sử dụng chủng SmGal.

Fig.12 là hình ảnh thể hiện thời gian phản ứng và nồng độ đường trong dung dịch trong quá trình sản xuất galactooligosacarit bằng cách sử dụng chủng SeGal.

Fig.13 là hình ảnh thể hiện thời gian phản ứng và nồng độ đường trong dung dịch trong quá trình sản xuất galactooligosacarit bằng cách sử dụng chủng SeGal (hình (a): ở nhiệt độ 70°C, hình (b): ở nhiệt độ 80°C).

Fig.14 là hình ảnh thể hiện thời gian phản ứng và nồng độ đường trong dung dịch trong quá trình sản xuất galactooligosacarit bằng cách sử dụng chủng SeGal (hình (c): ở nhiệt độ 90°C)

Mô tả chi tiết giải pháp hữu ích

Phương pháp sản xuất β -galactosidaza được tiết ra của giải pháp hữu ích (sau đây được gọi là “phương pháp sản xuất của giải pháp hữu ích”) là phương pháp bao gồm bước gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra.

Gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* được sử dụng trong phương pháp sản xuất của giải pháp hữu ích là gen mã hóa β -galactosidaza không tiết được sản sinh bởi nấm men *Basidiomycetes*. Ở đây, thuật ngữ “không tiết” nghĩa là có ái lực gắn kết với thành tế bào, đặc tính này có thể được khẳng định bằng cách nhuộm hoạt tính hoặc phương pháp tương tự.

Ngoài ra, không có giới hạn cụ thể về nấm men *Basidiomycetes* sản sinh ra β -galactosidaza không tiết, nhưng các ví dụ của nó bao gồm các nấm men *Basidiomycetes* thuộc về giống *Sporobolomyces* như *Sporobolomyces singularis*, giống *Sirobasidium* như *Sirobasidium magnum*, giống *Rhodotorula* như *Rhodotorula minuta*, giống *Sterigmatomyces* như *Sterigmatomyces elviae*, giống *Cryptococcus* như *Cryptococcus laurentii*, và giống tương tự. Trong số các nấm men *Basidiomycetes* này, nấm men thuộc về giống *Sporobolomyces* hoặc giống *Sterigmatomyces* là được ưu tiên, và giống *Sporobolomyces singularis* hoặc *Sterigmatomyces elviae* được ưu tiên hơn.

Ngoài ra, để làm gen mã hóa β -galactosidaza không tiết được sản sinh bởi nấm men *Basidiomycetes*, trước tiên, gen được tách dòng từ nấm men *Basidiomycetes* nêu trên sản sinh ra β -galactosidaza không tiết theo phương pháp thông thường như phản ứng PCR được ví dụ. Cần lưu ý rằng để làm gen nêu trên, gen thu được bằng cách tổng hợp chung theo vật chủ từ thông tin của gen thu được như đã mô tả trên đây là được ưu tiên.

Cụ thể, các gen sau đây được ví dụ. Cần lưu ý rằng trong các gen này, trình tự tín hiệu cũng được bao gồm.

- gen β -galactosidaza có nguồn gốc từ giống *Sporobolomyces singularis* gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 1 (trình tự ở vị trí từ 1 đến 57 trong trình tự này là trình tự tín hiệu)

- gen β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum* gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 7 (trình tự ở vị trí từ 1 đến 48 trong trình tự này là trình tự tín hiệu)

- gen β -galactosidaza có nguồn gốc từ giống *Rhodotorula minuta* gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 13 (trình tự ở vị trí từ 1 đến 57 trong trình tự này là trình tự tín hiệu)

- gen β -galactosidaza có nguồn gốc từ giống *Sterigmatomyces elviae* gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 19 (trình tự ở vị trí từ 1 đến 57 trong trình tự này là trình tự tín hiệu)

Ngoài ra, gen được ưu tiên của gen nêu trên là gen trong đó trình tự tín hiệu của mỗi nấm men *Basidiomycetes* được thay thế bằng trình tự tín hiệu của nấm *Aspergillus oryzae*. Các ví dụ về trình tự tín hiệu của nấm *Aspergillus oryzae* bao gồm trình tự tín hiệu tiết (tín hiệu TAA) của α -amylaza (Taka-amylase: TAA) của nấm *Aspergillus oryzae* (Okazaki, F., Aoki, J., Tabuchi, S., Tanaka, T., Ogino, C., and Kondo, A., Efficient heterologous expression and secretion in *Aspergillus oryzae* of a llama variable heavy-chain antibody fragment V(HH) against EGFR. *Appl Microbiol Biotechnol* 96, 81-88 (2012)), và trình tự tiết của lipaza của nấm *Rhizopus oryzae* (Hama, S., Tamalampudi, S., Shindo, N., Numata, T., Yamaji, H., Fukuda, H., and Kondo, A., Role of N-terminal 28-amino-acid region of *Rhizopus oryzae* lipase in directing proteins to secretory pathway of *Aspergillus oryzae*. *Appl Microbiol Biotechnol* 79, 1009-1018 (2008)). Sự thay thế trình tự tín hiệu này có thể được tiến hành theo phương pháp thông thường.

Để làm gen được ưu tiên trong số các gen β -galactosidaza mà khi đó trình tự tín hiệu của mỗi nấm men *Basidiomycetes* được thay thế bằng trình tự tín hiệu của nấm *Aspergillus oryzae*, các gen sau đây được ví dụ. Trình tự của các gen này là mỗi trình tự gồm trình tự tín hiệu tiết (tín hiệu TAA) của nấm *Aspergillus oryzae* và trình tự mã hóa β -galactosidaza tự nhiên.

- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 3 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 9 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 15 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 21 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)

Trong số các gen nêu trên, gen mà khi đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza không bị thay đổi là được ưu tiên. Để làm gen β -galactosidaza, các gen sau đây được ví dụ. Trình tự của các gen này là mỗi trình tự gồm trình tự tín hiệu tiết (tín hiệu TAA) của nấm *Aspergillus oryzae* và trình tự trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza không bị thay đổi.

- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 5 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 11 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 17 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)
- gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi trình tự SEQ ID NO: 23 (trình tự ở vị trí từ 1 đến 63 trong trình tự này là trình tự tín hiệu tiết)

Trong số các gen nêu trên, gen β -galactosidaza gồm trình tự bazơ được thể hiện bởi các trình tự SEQ ID NO: 5, 11, hoặc 23 là được ưu tiên.

Không có giới hạn cụ thể về nấm *Aspergillus oryzae* được sử dụng trong phương pháp sản xuất của giải pháp hữu ích và gen β -galactosidaza được gắn vào đó, nhưng ví dụ, chủng *Aspergillus oryzae* NS4 bị thiếu hụt gen ATP sulfurylaza (sC) và gen nitrat reductaza (niaD) (sản phẩm của Viện nghiên cứu quốc gia về sản xuất rượu bia (the National Research Institute of Brewing), 3-7-1, Kagamiyama, Higashihiroshima-shi, Hiroshima, 739-0046), chủng *Aspergillus oryzae* nia D300, *Aspergillus oryzae* RIB40, và *Aspergillus oryzae* ATCC 11488 được ví dụ. Trong số chúng, chủng *Aspergillus oryzae* NS4 là được ưu tiên.

Trong phương pháp sản xuất của giải pháp hữu ích, không có giới hạn cụ thể về phương pháp gắn gen nêu trên vào nấm *Aspergillus oryzae*, mà gen nêu trên chỉ cần được gắn vào, ví dụ, vectơ biểu hiện, bằng phương pháp thông thường. Không có giới hạn cụ thể về loại vectơ biểu hiện nhưng vectơ biểu hiện có nguồn gốc từ nấm *Aspergillus oryzae* là được ưu tiên, và cụ thể là vectơ biểu hiện ở mức cao chứa gen khởi đầu được cải thiện bằng cách sử dụng yếu tố cis (vùng III) tham gia vào quá trình điều hòa sự biểu hiện của gen amylaza (Improvement of the *Aspergillus oryzae* Enolase Promoter by the Introduction of cis-Element, Tsuboi, H. et al., Biosci. Biotechnol. Biochem., 69, 206-208 (2005)) và trình tự 5'UTR có hiệu quả dịch mã cao (Bằng độc quyền sáng chế Nhật Bản số 4413557) là được ưu tiên. Ngoài ra, trong vectơ này, gen kháng chất kháng sinh như ampicillin có thể được gắn vào để chọn lọc thể biến nạp, hoặc bằng biểu hiện ATP sulfurylaza hoặc sản phẩm tương tự có thể được gắn vào làm gen đánh dấu.

Vectơ biểu hiện nêu trên có thể được tạo ra dựa trên phương pháp đã được mô tả trong tài liệu nêu trên, hoặc ví dụ, có thể được tạo ra bằng cách sử dụng dịch vụ của công ty Contract Protein Expression Service of Ozeki Co., Ltd (4-9, Imazudezaike-cho,

Nishinomiya-shi, Hyogo, 663-8227).

Sau khi gắn gen vào vector biểu hiện, vector biểu hiện này được gắn vào nấm *Aspergillus oryzae* để biến nạp nấm *Aspergillus oryzae*. Không có giới hạn cụ thể về phương pháp biến nạp nấm *Aspergillus oryzae*, và ví dụ, quá trình biến nạp có thể được tiến hành bằng phương pháp thông thường như phương pháp PEG-thể nguyên sinh hoặc phương pháp gây xung điện. Sau khi biến nạp, các bước rửa, chọn lọc, thu hoạch tế bào, và bước tương tự có thể được thực hiện thích hợp theo phương pháp thông thường.

Theo cách này, có thể thu được thể biến nạp của nấm *Aspergillus oryzae* sản sinh ra β -galactosidaza được tiết ra bằng cách đưa gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae*. Bằng cách nuôi cấy thích hợp thể biến nạp trong môi trường DPY, môi trường CDD, hoặc môi trường tương tự, β -galactosidaza được tiết ra được sản sinh từ nấm *Aspergillus oryzae*.

Do β -galactosidaza thu được trên đây là protein tiết, và do đó, trong quá trình tinh chế, ví dụ, chỉ cần tách dung dịch nuôi cấy sau khi nuôi cấy bằng cách lọc, ly tâm hoặc phương pháp tương tự, và thu hoạch dịch nổi bề mặt. Ngoài ra, cũng có thể cô đặc dịch nổi bề mặt bằng cách sử dụng thiết bị siêu lọc hoặc thiết bị tương tự. β -galactosidaza có các đặc điểm là hoạt tính của β -galactosidaza cao, tính chịu nhiệt cũng cao, và hàm lượng tạp chất thấp.

Các ví dụ được ưu tiên của trình tự axit amin của β -galactosidaza được tiết ra này bao gồm các trình tự sau đây.

- β -galactosidaza có nguồn gốc từ giống *Sporobolomyces singularis* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 2 (trình tự ở vị trí từ 1 đến 575 trong trình tự này) (điều này cũng áp dụng với các trình tự axit amin được thể hiện bởi các trình tự SEQ ID NO: 4 và 6 (trình tự ở vị trí từ 1 đến 575 trong mỗi trình tự này))

- β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 8 (trình tự ở vị trí từ 1 đến 685 trong trình tự này) (điều này cũng áp dụng với các trình tự axit amin được thể hiện bởi các trình tự SEQ ID NO: 10 và 12 (trình tự ở vị trí từ 1 đến 685 trong mỗi trình tự này))

- β -galactosidaza có nguồn gốc từ giống *Rodotorula minuta* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 14 (trình tự ở vị trí từ 1 đến 581 trong trình tự này) (điều này cũng áp dụng với các trình tự axit amin được thể hiện bởi các trình tự SEQ ID NO: 16 và 18 (trình tự ở vị trí từ 1 đến 581 trong mỗi trình tự này))

· β -galactosidaza có nguồn gốc từ giống *Sterigmatomyces elviae* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 20 (trình tự ở vị trí từ 1 đến 581 trong trình tự này) (điều này cũng áp dụng với các trình tự axit amin được thể hiện bởi các trình tự SEQ ID NO: 22 và 24 (trình tự ở vị trí từ 1 đến 581 trong mỗi trình tự này))

Trong số các β -galactosidaza nêu trên, β -galactosidaza có nguồn gốc từ giống *Sporobolomyces singularis* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 2, và β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 8, hoặc β -galactosidaza có nguồn gốc từ giống *Sterigmatomyces elviae* gồm trình tự axit amin được thể hiện bởi trình tự SEQ ID NO: 20 là được ưu tiên.

β -galactosidaza có đặc tính của enzym được tiết bên ngoài tế bào, và còn có đặc tính về tính chịu nhiệt và tính ổn định khi bảo quản có lợi mà không làm giảm hoạt tính của β -galactosidaza ngay cả khi nó được bảo quản trong thời gian dài. Cần lưu ý rằng hoạt tính của β -galactosidaza có thể được khẳng định bằng phương pháp được mô tả trong các ví dụ dưới đây. Nói chung, để sản xuất galactooligosacarit có hiệu quả, nhiều β -galactosidaza đôi khi được sử dụng, tuy nhiên, β -galactosidaza thu được trên đây có thể tự sản sinh ra galactooligosacarit có hiệu quả.

β -galactosidaza thu được trên đây có thể được sử dụng, ví dụ, trong quá trình sản xuất galactooligosacarit bằng cách cho β -galactosidaza tác dụng lên cơ chất chứa ít nhất là lactoza theo cách giống như phương pháp sản xuất β -galactosidaza đã biết thông thường. Cần lưu ý rằng β -galactosidaza là protein tiết, và do đó, cũng không cần phải loại bỏ cụ thể các tế bào hoặc thành phần tương tự khi sản xuất galactooligosacarit.

Cụ thể, để cho β -galactosidaza thu được trên đây tác dụng lên cơ chất chứa ít nhất là lactoza, chỉ cần bổ sung β -galactosidaza vào cơ chất chứa ít nhất là lactoza và duy trì ở nhiệt độ định trước. Không có giới hạn cụ thể về lượng β -galactosidaza được bổ sung, nhưng lượng này nằm trong khoảng, ví dụ, từ 1 đến 50 U, và tốt hơn là nằm trong khoảng từ 5 đến 10 U so với 100g lactoza. Ngoài ra, không có giới hạn cụ thể về nhiệt độ mà khi đó β -galactosidaza được cho tác dụng lên cơ chất, nhưng nhiệt độ này nằm trong khoảng từ 30 đến 90°C, và tốt hơn là nằm trong khoảng từ 60 đến 90°C, và thời gian duy trì có thể chỉ được thiết lập, nếu cần. Sacarit cần được galactosyl hóa có thể được cho thêm vào cơ chất chứa ít nhất là lactoza. Không có giới hạn cụ thể về sacarit này, và các ví dụ của nó bao gồm galactoza, mannoza, riboza, xyloza, arabinoza, rhamnoza, N-axetylglucosamin,

α -methylmannosit, α -methylgalactosit, α -methylglucosit, 2-deoxyglucoza, và 2-deoxygalactoza.

Galactooligosacarit được sản xuất như đã mô tả trên đây chứa galactooligosacarit là pentasacarit hoặc sacarit thấp, cụ thể là galactooligosacarit chứa trisacarit với lượng lớn.

Ngoài ra, galactooligosacarit được sản xuất như đã mô tả trên đây có thể được sử dụng ngay nhưng có thể được tách và tinh chế bằng cách sử dụng phương pháp tinh chế thông thường. Không có giới hạn cụ thể về phương pháp tinh chế, nhưng cụ thể, quá trình tinh chế có thể được tiến hành bằng cách xử lý galactooligosacarit bằng các phương pháp sắc ký khác nhau như phương pháp sắc ký trao đổi ion, thẩm gel, than hoạt tính và sắc ký ái lực.

Galactooligosacarit thu được có thể được sử dụng trong các nguyên liệu thực phẩm, dược phẩm và chất phản ứng hữu ích.

Ví dụ thực hiện giải pháp hữu ích

Sau đây, giải pháp hữu ích sẽ được mô tả chi tiết cùng với các ví dụ, tuy nhiên, phạm vi của giải pháp hữu ích không bị giới hạn bởi các ví dụ này.

Các mã số lưu trữ của nấm men *Basidiomycetes* được sử dụng trong các ví dụ này là như sau.

· *Sporobolomyces singularis* ATCC 24193

· *Rodotorula minuta* CBS 319

· *Sterigmatomyces elviae* IFO 1843

· *Sirobasidium magnum* CBS 6803

ATCC: 10801 University Boulevard Manassas, VA 20110, Mỹ

CBS: Uppsalalaan 8, 3584 CT, Utrecht, Hà Lan

IFO: 2-17-85 Jusohonmachi, Yodogawa-ku, Osaka-shi 532-8686

Ví dụ 1

Thu được gen β -galactosidaza có nguồn gốc từ giống *Sporobolomyces Singularis*:

Gen β -galactosidaza có nguồn gốc từ giống *Sporobolomyces singularis* thu được theo phương pháp của các tài liệu đã biết (Ishikawa, E., Sakai, T., Ikemura, H., Matsumoto, K., and Abe, H., Identification, cloning, and characterization of a *Sporobolomyces singularis* beta-galactosidase-like enzyme involved in galacto- oligosaccharide production. J

Biosci Bioeng 99, 331-339 (2005)) (SEQ ID NO: 1). Gen này gồm trình tự tín hiệu và trình tự mã hóa β -galactosidaza. Trình tự (SEQ ID NO: 3) trong đó trình tự tín hiệu của gen này được thay thế bằng trình tự tín hiệu TAA của nấm *Aspergillus oryzae* thu được trên máy tính, và ngoài ra, trình tự (SEQ ID NO: 5) trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza thu được không bị thay đổi (SsGal). Gen SsGal được chèn vào GenScript, Inc. và được tổng hợp hoàn chỉnh.

Ví dụ 2

Thu được gen β -galactosidaza có nguồn gốc từ giống *Sirobasidium Magnum*:

Các đoạn mồi suy biến (Bảng 1) (SEQ ID NO: 25 - 29) được thiết kế từ vùng được bảo toàn, và các trình tự riêng phần được tách dòng bằng phản ứng RT-PCR trong tổng cộng sáu tổ hợp bằng cách sử dụng 2 loại đoạn mồi xuôi và 3 loại đoạn mồi ngược. Từ trình tự riêng phần, các quá trình khuếch đại RACE ở đầu 5' và RACE ở đầu 3' được thực hiện, nhờ đó thu được cADN có chiều dài đầy đủ.

Bảng 1

SEQ ID NO	Tên	Hướng	Trình tự bazơ của đoạn mồi suy biến
25	F1	xuôi	gccggcgcggtatthcargtngarggngcn
26	F2	xuôi	gtcaagaentggtyacnttyaaygarccn
27	R1	ngược	ctcggcccccaccraaytenswraartadat
28	R2	ngược	ccattcccarttrtcnacraanswcca
29	C-R70	ngược	gacgaggccnswrttcaytcraarttrtc

Dựa trên cADN có chiều dài đầy đủ thu được, gen β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum* (SEQ ID NO: 7) thu được bằng cách tương tự từ codon khởi đầu (ATG) trong vùng ngược dòng. Gen này gồm trình tự tín hiệu và trình tự mã hóa β -galactosidaza. Trình tự (SEQ ID NO: 9) trong đó trình tự tín hiệu của gen này được thay thế bằng trình tự tín hiệu TAA của nấm *Aspergillus oryzae* thu được trên máy tính, và ngoài ra, trình tự (SEQ ID NO: 11) trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza thu được không bị thay đổi (SmGal). Gen SmGal được chèn vào GenScript, Inc. và được tổng hợp hoàn chỉnh.

Ví dụ 3

Thu được gen β -galactosidaza có nguồn gốc từ giống *Rodotorula Minuta*:

Gen β -galactosidaza có nguồn gốc từ giống *Rodotorula minuta* (SEQ ID NO: 13) thu được theo cách giống như với gen β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum*. Gen này gồm trình tự tín hiệu và trình tự mã hóa β -galactosidaza. Trình tự (SEQ ID NO: 15) trong đó trình tự tín hiệu của gen này được thay thế bằng trình tự tín hiệu TAA của nấm *Aspergillus oryzae* thu được trên máy tính, và ngoài ra, trình tự (SEQ ID NO: 17) trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza thu được không bị thay đổi (RmGal). Gen RmGal được chèn vào GenScript, Inc. và được tổng hợp hoàn chỉnh.

Ví dụ 4

Thu được gen β -galactosidaza có nguồn gốc từ giống *Sterigmatomyces Elviae*:

Gen β -galactosidaza của giống *Sterigmatomyces elviae* (SEQ ID NO: 19) thu được theo cách giống như với gen β -galactosidaza có nguồn gốc từ giống *Sirobasidium magnum*. Gen này gồm trình tự tín hiệu và trình tự mã hóa β -galactosidaza. Trình tự (SEQ ID NO: 21) trong đó trình tự tín hiệu của gen này được thay thế bằng trình tự tín hiệu TAA của nấm *Aspergillus oryzae* thu được trên máy tính, và ngoài ra, trình tự (SEQ ID NO: 23) trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza thu được không bị thay đổi (SeGal). SeGal được chèn vào GenScript, Inc. và được tổng hợp hoàn chỉnh.

Ví dụ 5

Thu được thể biến nạp SsGal

SsGal thu được trong Ví dụ 1 được gửi đến công ty Contract Protein Expression Service of Ozeki Co., Ltd. (4-9, Imazudezaike-cho, Nishinomiya-shi, Hyogo, 663-8227), và được gắn vào vectơ biểu hiện.

Để làm vật chủ dùng để biến nạp, NS4 là chủng có nguồn gốc từ nấm *Aspergillus oryzae* và bị thiếu hụt gen nitrat reductaza (niaD) và gen ATP sulfurylaza (sC) (thu được từ National Research Institute of Brewing, 3-7-1, Kagamiyama, Higashihiroshima-shi, Hiroshima, 739-0046) được sử dụng, và vectơ biểu hiện được gắn vào vật chủ bằng phương pháp PEG-thể nguyên sinh, nhờ đó thu được thể biến nạp (chủng SsGal). Cần lưu ý rằng việc chọn lọc thể biến nạp được thực hiện trên cơ sở bổ sung tính trạng của sự thiếu hụt sC.

Ví dụ 6

Thu được thể biến nạp SmGal

Thu được vector biểu hiện đã chuyển nhiễm và thể biến nạp (chủng SmGal) theo cách giống như trong Ví dụ 5, ngoại trừ việc SmGal thu được trong Ví dụ 2 được sử dụng.

Ví dụ 7

Thu được thể biến nạp RmGal

Thu được vector biểu hiện đã chuyển nhiễm và thể biến nạp (chủng RmGal) theo cách giống như trong Ví dụ 5, ngoại trừ việc RmGal thu được trong Ví dụ 3 được sử dụng.

Ví dụ 8

Thu được thể biến nạp SeGal

Thu được vector biểu hiện đã chuyển nhiễm và thể biến nạp (chủng SeGal) theo cách giống như trong Ví dụ 5, ngoại trừ việc SeGal thu được trong Ví dụ 4 được sử dụng.

Ví dụ 9

Đánh giá khả năng sản sinh ra β -galactosidaza bởi thể biến nạp

(1) Đánh giá hoạt tính

Trong số các thể biến nạp tương ứng thu được trong các ví dụ từ 5 đến 8, chủng SsGal được nuôi cấy trong môi trường CDD (dextrin 2%, glucoza 0,2%, NH_4Cl 0,2%, KCl 0,002%, K_2HPO_4 0,001%, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,0005%, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ $2 \times 10^{-5}\%$, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ $1 \times 10^{-5}\%$, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ $1 \times 10^{-6}\%$, $\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$ $1 \times 10^{-6}\%$, AlCl_3 $1 \times 10^{-6}\%$, chất đệm MOPS-NaOH 200mM, độ pH = 7,0) ở nhiệt độ 30°C trong 144 giờ (thang tỷ lệ của bình Erlenmeyer theo thể tích 15ml/100ml). Chủng RmGal được nuôi cấy trong môi trường 2×DPY (dextrin 4%, hipolypepton 2%, chất chiết nấm men 2%, KH_2PO_4 1%, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,1%) ở nhiệt độ 30°C trong 144 giờ (thang tỷ lệ của bình lắc cổ dài theo thể tích 150ml/500ml). Chủng SmGal được nuôi cấy trong môi trường 2×DPY ở nhiệt độ 30°C trong 168 giờ (thang tỷ lệ của bình lắc cổ dài theo thể tích 150ml/500ml). Chủng SeGal được nuôi cấy trong môi trường DPY (dextrin 2%, hipolypepton 1%, chất chiết nấm men 1%, KH_2PO_4 0,5%, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,05%) ở nhiệt độ 30°C trong 168 giờ. Dịch nổi bề mặt nuôi cấy được thu hồi, và được trộn với lượng tương đương của 2×dung dịch đệm mẫu

(Tris 125mM -HCl (độ pH = 6,8), glyxerol 20%, bromophenol màu xanh da trời 0,01%, SDS 4%, DTT 200mM), và hỗn hợp tạo thành được xử lý ở nhiệt độ 100°C trong 10 phút và sau đó được phân tích bằng phương pháp SDS-PAGE (nhuộm CBB).

Ngoài ra, việc đánh giá hoạt tính được thực hiện theo phương pháp sau đây bằng cách sử dụng ONPG làm cơ chất. Dung dịch được điều chế bằng cách cho thêm 2-nitrophenyl- β -galactosit (ONPG) vào dung dịch đệm xitrat phosphat 50mM (độ pH = 4,0) để thu được nồng độ 12,5mM. 0,8ml dung dịch này được cho thêm 0,2ml dịch nổi bề mặt nuôi cấy chứa β -galactosidaza được pha loãng bằng dung dịch đệm xitrat phosphat 50mM (độ pH = 4,0) sao cho hệ số hấp thụ ở bước sóng 420nm nằm trong khoảng từ 0,2 đến 0,8 và để cho phản ứng xảy ra ở nhiệt độ 30°C trong 10 phút (dung dịch thử nghiệm). Sau khi làm ngừng phản ứng bằng cách cho thêm 4ml dung dịch natri cacbonat 0,25M, tiến hành ly tâm (ở tốc độ 3000 g, 10 phút), và lượng 2-nitrophenol được giải phóng có mặt trong dịch nổi bề mặt được xác định bằng cách xác định hệ số hấp thụ ở bước sóng 420nm bằng cách sử dụng phổ kế phân cực. Mặt khác, dung dịch thu được bằng cách cho dung dịch đệm xitrat phosphat 50mM (độ pH = 4,0) vào dung dịch 2-nitrophenyl- β -galactosit được sử dụng làm dung dịch không chứa chất phản ứng, và dung dịch thu được bằng cách trước tiên, cho dung dịch natri cacbonat vào, và sau đó cho thêm dịch nổi bề mặt nuôi cấy chứa β -galactosidaza vào đó và trộn lẫn, và đồng thời, làm ngừng phản ứng và tiến hành làm hiện màu được xác định làm dung dịch phản ứng ban đầu (mẫu mù). Một đơn vị (U) của hoạt tính enzym được xác định làm lượng enzym để giải phóng 1 μ mol 2-nitrophenol/phút trong các điều kiện và được tính toán theo công thức sau.

Công thức 1

Giá trị hoạt tính (U/ml) = $\left[\{(\text{hệ số hấp thụ của dung dịch thử nghiệm}) - (\text{hệ số hấp thụ dung dịch không chứa chất phản ứng}) - (\text{hệ số hấp thụ của mẫu mù})\} / \{0,91(\text{hệ số tắt của 2-nitrophenol}) \times 0,2 \times 10\} \right] \times \text{hệ số pha loãng}$

Kết quả của phương pháp phân tích SDS-PAGE và đánh giá hoạt tính được thể hiện trên các hình vẽ từ Fig.1 đến Fig.4. Bằng cách nhuộm CBB, đối với mỗi chủng trong số chủng RmGal, chủng SmGal, và chủng SeGal, không phát hiện được sự có mặt của dải cụ thể trong chủng bố-mẹ, và dải này được cho là β -galactosidaza của mỗi chủng. Đối với chủng SsGal, không quan sát được dải cụ thể trong môi trường DPY, nhưng khi chủng này được nuôi cấy trong môi trường CDD (độ pH = 7,0), không phát hiện được sự có mặt của

dải cụ thể trong chủng bố-mẹ, và dải này được cho là β -galactosidaza. Các điều kiện nuôi cấy trong đó năng suất tiết của mỗi β -galactosidaza trở nên cao được nghiên cứu. Kết quả là thu được hoạt tính lớn nhất trong các điều kiện mà chủng SsGal được nuôi cấy trong môi trường CDD (độ pH = 7,0) ở nhiệt độ 30°C trong 144 giờ, chủng RmGal được nuôi cấy trong môi trường 2×DPY ở nhiệt độ 30°C trong 144 giờ, chủng SmGal được nuôi cấy trong môi trường 2×DPY ở nhiệt độ 30°C trong 168 giờ, và chủng SeGal được nuôi cấy trong môi trường DPY ở nhiệt độ 30°C trong 168 giờ. Ngoài ra, năng suất của chủng SsGal, chủng RmGal, chủng SmGal, và chủng SeGal được xác định tương ứng bằng khoảng 200 mg/l, khoảng 200 mg/l, khoảng 200 mg/l, và khoảng 1 g/l, dựa trên sự tập trung của các dải của SDS-PAGE.

(2) Đánh giá số lượng bản sao

Ngoài ra, việc đánh giá số lượng bằng biểu hiện được gắn vào thể biến nạp được thực hiện bằng phương pháp PCR trong thời gian thực.

Từ các kết quả của phản ứng PCR, đánh giá được rằng mỗi chủng SsGal, chủng RmGal, và chủng SmGal là chủng mà trong đó một bản sao của băng biểu hiện được chèn vào, và chủng SeGal là chủng mà trong đó hai bản sao của băng biểu hiện được chèn vào.

(3) Thử nghiệm làm bất hoạt bằng nhiệt

1ml của mỗi dung dịch nuôi cấy của các thể biến nạp tương ứng và chủng bố-mẹ (chủng NS4) được nuôi cấy trong các điều kiện nuôi cấy như được mô tả trong mục (1) được ủ ở nhiệt độ 40°C, 50°C, 60°C, 70°C, hoặc 80°C trong 1 giờ, và tiến hành xác định hoạt tính enzym và phương pháp phân tích SDS-PAGE.

Kết quả của phương pháp SDS-PAGE và đánh giá hoạt tính được thể hiện trên các hình vẽ từ Fig.5 đến Fig.8. Chủng SsGal vẫn duy trì được hoạt tính của nó ở nhiệt độ lên đến 40°C, nhưng khi chủng SsGal được ủ ở nhiệt độ 50°C trong 1 giờ, hoạt tính bị giảm khoảng 70%, và hoạt tính bị mất đi ở nhiệt độ 70°C. Phát hiện được hoạt tính của chủng bố-mẹ được nuôi cấy trong cùng điều kiện là rất ít ở nhiệt độ lên đến 60°C, và hoạt tính bị mất đi ở nhiệt độ 70°C. Chủng RmGal vẫn duy trì được hoạt tính của nó ở nhiệt độ lên đến 50°C, nhưng khi chủng RmGal được ủ ở nhiệt độ 60°C trong 1 giờ, hoạt tính bị mất đi. Phát hiện được hoạt tính của chủng bố mẹ được nuôi cấy trong cùng điều kiện ở nhiệt độ lên đến 70°C, và hoạt tính bị mất đi ở nhiệt độ 80°C. Chủng SmGal vẫn duy trì được hoạt tính của nó ở nhiệt độ lên đến 50°C, nhưng khi chủng SmGal được ủ ở nhiệt độ 60°C trong 1 giờ,

hoạt tính bị giảm đi khoảng 20%, và hoạt tính bị mất đi ở nhiệt độ 80°C. Phát hiện được hoạt tính của chủng bố-mẹ được nuôi cấy trong cùng điều kiện ở nhiệt độ lên đến 70°C, và hoạt tính bị mất đi ở nhiệt độ 80°C. Chủng SeGal vẫn duy trì được hoạt tính của nó ở nhiệt độ lên đến 70°C. Khi chủng SeGal chủng được ủ ở nhiệt độ 80°C trong 1 giờ, hoạt tính bị giảm đi khoảng 97%. Phát hiện được hoạt tính của chủng bố-mẹ được nuôi cấy trong cùng điều kiện ở nhiệt độ lên đến 40°C, và hoạt tính bị mất đi ở nhiệt độ 50°C. Ngoài ra, đối với chủng SeGal, việc xử lý ở nhiệt độ 80°C được thực hiện trong thời gian ngắn (5 phút, 10 phút, và 20 phút) chứ không phải 1 giờ. Kết quả là hoạt tính bị giảm khoảng 37% khi thực hiện việc xử lý ở nhiệt độ 80°C trong 5 phút, và hoạt tính bị giảm khoảng 98% khi thực hiện việc xử lý trong 20 phút. Ngoài ra, theo kết quả đánh giá hoạt tính, tất cả các chủng đều có hoạt tính cao hơn hoạt tính của chủng bố-mẹ khi được xử lý ở nhiệt độ 40°C và 50°C.

Từ các kết quả trên đây, đã phát hiện được rằng chủng SeGal có thể duy trì được hoạt tính của nó ngay cả ở nhiệt độ cao.

Ví dụ 10

Loại bỏ các enzym tạp chất

Như được thể hiện trong Ví dụ 9, mục (3), đã phát hiện được rằng chủng SeGal có thể duy trì được hoạt tính của nó ngay cả ở nhiệt độ cao. Mặt khác, khi thử nghiệm làm bất hoạt bằng nhiệt được thực hiện theo cách giống như trong Ví dụ 9, mục (3) đối với enzyme tạp chất có nguồn gốc từ nấm *Aspergillus oryzae*, nghĩa là chủng bố-mẹ của chủng SeGal, các enzym tạp chất có thể được làm bất hoạt bằng cách xử lý nhiệt ở nhiệt độ 70°C. Do đó, đã phát hiện được rằng β -galactosidaza được sản sinh bởi chủng SeGal có thể được tinh chế bằng cách xử lý nhiệt (Fig.9).

Ví dụ 11

Sản xuất galactooligosacarit (1):

150ml dung dịch chứa lactoza 66% (trọng lượng/thể tích) được cho thêm mỗi dịch nổi bề mặt nuôi cấy của chủng SsGal, chủng SmGal và chủng SeGal thu được trong Ví dụ 9 với lượng tương ứng với 10 U và dung dịch này được để cho phản ứng ở nhiệt độ định trước trong khoảng thời gian định trước, nhờ đó galactooligosacarit được tạo ra. Thành phần đường và lượng của nó được xác định bằng phương pháp sắc ký lỏng hiệu năng cao.

Thời gian phản ứng và hàm lượng thành phần đường trong dung dịch được thể hiện trên các hình vẽ từ Fig.10 đến Fig.12 (Fig.10: chủng SsGal, Fig.11: chủng SmGal, Fig.12: chủng SeGal).

Từ các hình vẽ, đã phát hiện được rằng do lượng β -galactosidaza được sản sinh bởi chủng SsGal, chủng SmGal, và chủng SeGal, galactooligosacarit là trisacarit có thể được tạo ra chủ yếu từ lactoza.

Ngoài ra, khi galactooligosacarit được sản xuất bằng cách sử dụng β -galactosidaza do chủng SsGal tiết ra, hàm lượng galactooligosacarit bằng 56,0%, khi galactooligosacarit được sản xuất bằng cách sử dụng β -galactosidaza do chủng SmGal tiết ra, hàm lượng galactooligosacarit bằng 66,7%, và khi galactooligosacarit được sản xuất bằng cách sử dụng β -galactosidaza do chủng SeGal tiết ra, hàm lượng galactooligosacarit bằng 68,5%.

Cần lưu ý rằng các β -galactosidaza nêu trên là các protein tiết, và do đó không cần xử lý các tế bào sau khi tạo ra galactooligosacarit, và galactooligosacarit có thể được sản xuất theo cách hữu hiệu.

Ví dụ 12

Sản xuất galactooligosacarit (2):

150ml dung dịch chứa lactoza 66% (trọng lượng/thể tích) được cho thêm dịch nổi bề mặt nuôi cấy của chủng SeGal thu được trong Ví dụ 9 với lượng tương ứng với 1,0 U, và dung dịch này được để cho phản ứng ở nhiệt độ 70°C, 80°C, hoặc 90°C trong khoảng thời gian định trước, nhờ đó galactooligosacarit được tạo ra. Thành phần đường và lượng của nó được xác định bằng phương pháp sắc ký lỏng hiệu năng cao. Thời gian phản ứng và hàm lượng thành phần đường trong dung dịch được thể hiện trên Fig.13 (hình (a): ở nhiệt độ 70°C, hình (b): ở nhiệt độ 80°C) và Fig.14 (hình (c): ở nhiệt độ 90°C).

β -galactosidaza có nguồn gốc từ chủng SeGal có khả năng chịu nhiệt cao, và có thể đạt được sự tạo ra GOS ở nhiệt độ nằm trong khoảng từ 70 đến 90°C.

Khả năng ứng dụng trong công nghiệp

β -galactosidaza thu được bằng phương pháp sản xuất β -galactosidaza được tiết ra được tách và tinh chế dễ dàng, và có thể được sử dụng trong quá trình sản xuất galactooligosacarit.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất β -galactosidaza được tiết ra, khác biệt ở chỗ, phương pháp này bao gồm bước gắn gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra.
2. Phương pháp sản xuất β -galactosidaza được tiết ra theo điểm 1, trong đó nấm men *Basidiomycetes* thuộc về giống *Sporobolomyces*, giống *Sirobasidium*, giống *Rhodotorula*, hoặc giống *Sterigmatomyces*.
3. Phương pháp sản xuất β -galactosidaza theo điểm 1 hoặc 2, trong đó gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* gồm trình tự tín hiệu và trình tự mã hóa β -galactosidaza.
4. Phương pháp sản xuất β -galactosidaza theo điểm 3, trong đó trình tự tín hiệu của gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* là trình tự tín hiệu của nấm *Aspergillus oryzae*.
5. Phương pháp sản xuất β -galactosidaza theo điểm 3 hoặc 4, trong đó trình tự mã hóa β -galactosidaza là trình tự mà trong đó codon của trình tự mã hóa β -galactosidaza tự nhiên được thay đổi trong khoảng sao cho trình tự axit amin của β -galactosidaza không bị thay đổi.
6. Phương pháp sản xuất β -galactosidaza theo điểm 3, trong đó trình tự tín hiệu và trình tự mã hóa β -galactosidaza là trình tự được thể hiện bởi các trình tự SEQ ID NO: 1, 7, 13, hoặc 19.
7. Phương pháp sản xuất β -galactosidaza theo điểm 3 hoặc 4, trong đó trình tự tín hiệu và trình tự mã hóa β -galactosidaza là trình tự được thể hiện bởi các trình tự SEQ ID NO: 3, 9, 15, hoặc 21.
8. Phương pháp sản xuất β -galactosidaza theo điểm 5, trong đó trình tự tín hiệu và trình tự

mã hóa β -galactosidaza là trình tự được thể hiện bởi các trình tự SEQ ID NO: 5, 11, 17, hoặc 23.

9. Gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes*, trong đó gen này có trình tự được thể hiện bởi các trình tự SEQ ID NO: 7, 13, hoặc 19.
10. Thể biến nạp của nấm *Aspergillus oryzae*, khác biệt ở chỗ, gen β -galactosidaza không tiết có nguồn gốc từ nấm men *Basidiomycetes* được gắn vào nấm *Aspergillus oryzae* để sản xuất β -galactosidaza được tiết ra.
11. Phương pháp sản xuất galactooligosacarit, khác biệt ở chỗ, phương pháp này bao gồm bước cho β -galactosidaza thu được bằng phương pháp sản xuất β -galactosidaza theo điểm bất kỳ trong số các điểm từ 1 đến 8 tác dụng lên cơ chất chứa ít nhất là lactoza.

DANH MỤC TRÌNH TỰ

<110> Kabushiki Kaisha Yakult Honsha
 <120> Phương pháp sản xuất β -galactosidaza được tiết ra
 <130> PF-190014-WO
 <150> JP2018-212757
 <151> 2018-11-13
 <160> 29
 <170> PatentIn phiên bản 3.5
 <210> 1
 <211> 1785
 <212> ADN
 <213> *Sporobolomyces singularis*
 <220>
 <221> CDS
 <222> (1)..(1785)
 <223> Tác giả sáng chế: Ishikawa, Eiji; Ikeda, Masakazu; Anbe, Minako; Hatano, Hiroshi
 <220>
 <221> sig_peptit
 <222> (1)..(57)
 <220>
 <221> mat_peptit
 <222> (58)..(1782)
 <400> 1
 atg atg ctg cat gcg gca ctg ctc gtt gcg ctc ccc tgc gtg gtt ctt 48
 Met Met Leu His Ala Ala Leu Leu Val Ala Leu Pro Cys Val Val Leu
 -15 -10 -5
 gct cgt ccc gcc ggt gca gtt acc tac ccc ggt gcg att cca ctt agc 96
 Ala Arg Pro Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro Leu Ser
 -1 1 5 10
 ttg acc agc aat tac gag acg ccg agt ccg acc gcc atc ccc ctg gag 144
 Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro Leu Glu
 15 20 25
 ccg acc cca acg gcg acc gga acc gcc gaa ctt gat gcg ctc tgg aat 192
 Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu Trp Asn
 30 35 40 45
 ttg gtg gaa gca cag tac cct gtt cag acg gcg gct gtc acc acc ctg 240
 Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr Thr Leu
 50 55 60
 gtg acg gtg ccc gac gac tac aag ttt gaa gca gac cct cct tcc tat 288
 Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro Ser Tyr
 65 70 75
 gct ctt gct ggc tac gag aca tca gaa att gcc ggc ttg aag ttc ccg 336
 1

Ala	Leu	Ala	Gly	Tyr	Glu	Thr	Ser	Glu	Ile	Ala	Gly	Leu	Lys	Phe	Pro		
	80						85					90					
aag	ggg	ttc	aag	ttt	ggc	gtg	gcc	ggc	gcg	gct	att	caa	gtg	gaa	ggc	384	
Lys	Gly	Phe	Lys	Phe	Gly	Val	Ala	Gly	Ala	Ala	Ile	Gln	Val	Glu	Gly		
	95					100					105						
gca	gcg	aaa	gca	gag	gga	cga	ggc	cca	tcc	act	tgg	gat	tac	ttg	tgc	432	
Ala	Ala	Lys	Ala	Glu	Gly	Arg	Gly	Pro	Ser	Thr	Trp	Asp	Tyr	Leu	Cys		
110					115					120					125		
cac	cat	tac	gcg	tcc	aca	cag	tgc	aac	aac	tat	gat	cct	gac	att	acg	480	
His	His	Tyr	Ala	Ser	Thr	Gln	Cys	Asn	Asn	Tyr	Asp	Pro	Asp	Ile	Thr		
				130					135						140		
acg	aac	cat	tac	tac	ctt	tac	cct	ctt	gat	ttc	gcc	cgg	ctc	cag	cat	528	
Thr	Asn	His	Tyr	Tyr	Leu	Tyr	Pro	Leu	Asp	Phe	Ala	Arg	Leu	Gln	His		
			145					150					155				
cta	ggc	atc	aac	acg	tat	tcg	ttt	tca	atc	tcc	tgg	act	cgt	ata	tac	576	
Leu	Gly	Ile	Asn	Thr	Tyr	Ser	Phe	Ser	Ile	Ser	Trp	Thr	Arg	Ile	Tyr		
	160					165						170					
cct	ctg	ggg	gct	ggc	tac	gtt	aac	gaa	gcc	ggg	ttg	gcg	cat	tac	gac	624	
Pro	Leu	Gly	Ala	Gly	Tyr	Val	Asn	Glu	Ala	Gly	Leu	Ala	His	Tyr	Asp		
	175					180					185						
gcg	gta	atc	cac	tcg	gcc	aag	aag	tac	ggg	ctg	gag	cct	gtc	gga	aca	672	
Ala	Val	Ile	His	Ser	Ala	Lys	Lys	Tyr	Gly	Leu	Glu	Pro	Val	Gly	Thr		
190					195				200					205			
gta	ttt	cac	tgg	gac	acc	cct	ctc	agc	ctc	atg	ctc	aaa	tat	ggc	gcg	720	
Val	Phe	His	Trp	Asp	Thr	Pro	Leu	Ser	Leu	Met	Leu	Lys	Tyr	Gly	Ala		
				210				215						220			
tgg	caa	gat	acc	ggc	gac	cag	atc	gtt	aaa	gat	ttc	gtc	aca	tac	gcc	768	
Trp	Gln	Asp	Thr	Gly	Asp	Gln	Ile	Val	Lys	Asp	Phe	Val	Thr	Tyr	Ala		
			225					230					235				
acc	acc	gtc	ttc	aaa	cga	tac	ggg	aat	gaa	gtc	aag	acc	tgg	ttc	acg	816	
Thr	Thr	Val	Phe	Lys	Arg	Tyr	Gly	Asn	Glu	Val	Lys	Thr	Trp	Phe	Thr		
		240					245					250					
ttc	aat	gag	cct	cgc	gtg	ttc	tgt	tct	caa	aac	agt	ggc	ctt	ccc	tat	864	
Phe	Asn	Glu	Pro	Arg	Val	Phe	Cys	Ser	Gln	Asn	Ser	Gly	Leu	Pro	Tyr		
	255					260					265						
aac	ctc	acg	tat	cct	gag	gga	atc	aac	tca	act	tca	gcc	gtc	ttc	cgg	912	
Asn	Leu	Thr	Tyr	Pro	Glu	Gly	Ile	Asn	Ser	Thr	Ser	Ala	Val	Phe	Arg		
270					275					280					285		
tgt	act	tat	aac	gtc	ctg	aaa	gcc	cat	ggc	cac	gcg	gtt	aag	gtt	tac	960	
Cys	Thr	Tyr	Asn	Val	Leu	Lys	Ala	His	Gly	His	Ala	Val	Lys	Val	Tyr		
				290					295					300			
cgg	gat	ctc	gtt	gcc	agc	gga	acc	att	gct	gct	gga	gag	atc	ggc	ttc	1008	
Arg	Asp	Leu	Val	Ala	Ser	Gly	Thr	Ile	Ala	Ala	Gly	Glu	Ile	Gly	Phe		
			305					310					315				
aag	tcg	gac	gac	aac	tac	cca	atc	cca	gcg	cgg	ccc	gga	aac	gcg	gac	1056	
Lys	Ser	Asp	Asp	Asn	Tyr	Pro	Ile	Pro	Ala	Arg	Pro	Gly	Asn	Ala	Asp		

320	325	330	
gac gag gaa tcc gcc aaa cgt cac gaa gcg ttc cga atc gga atc ttt Asp Glu Glu Ser Ala Lys Arg His Glu Ala Phe Arg Ile Gly Ile Phe 335 340 345			1104
gcc cag cca gtt tac gga aac ggc gac tat cct gat gta gta aaa gag Ala Gln Pro Val Tyr Gly Asn Gly Asp Tyr Pro Asp Val Val Lys Glu 350 355 360 365			1152
acc gtt ggc gac atg ctg ccc gcc ctg acg gat gag gac aag ggc tac Thr Val Gly Asp Met Leu Pro Ala Leu Thr Asp Glu Asp Lys Gly Tyr 370 375 380			1200
atc aag ggc agc ggc gac atc ttc gcc att gac ggt tac cgg acc gat Ile Lys Gly Ser Gly Asp Ile Phe Ala Ile Asp Gly Tyr Arg Thr Asp 385 390 395			1248
atc tcg cat gcc gca ctg aat gga atc gcg aat tgc atc aga aac cag Ile Ser His Ala Ala Leu Asn Gly Ile Ala Asn Cys Ile Arg Asn Gln 400 405 410			1296
tcg gac cct aac tgg cct gtt tgc gag gaa ggg tct gac ccg ttc gcc Ser Asp Pro Asn Trp Pro Val Cys Glu Glu Gly Ser Asp Pro Phe Ala 415 420 425			1344
cac gta tac ccg tct ggt ttc gcc atc ggc cag tcc gcc gat ccg ctg His Val Tyr Pro Ser Gly Phe Ala Ile Gly Gln Ser Ala Asp Pro Leu 430 435 440 445			1392
tcg tca tgg ctc gtc aac tcc gcc cca ttt att cgc gac cag ctg aag Ser Ser Trp Leu Val Asn Ser Ala Pro Phe Ile Arg Asp Gln Leu Lys 450 455 460			1440
ttc ctc act caa acg tac ccg gca aag gga ggt att tac ttc agc gag Phe Leu Thr Gln Thr Tyr Pro Ala Lys Gly Gly Ile Tyr Phe Ser Glu 465 470 475			1488
ttt ggg tgg gcc gag gat gcg gag tac gac cgc cag ctg ttg tac caa Phe Gly Trp Ala Glu Asp Ala Glu Tyr Asp Arg Gln Leu Leu Tyr Gln 480 485 490			1536
atc acc tgg gac ggt ctt agg acc cag tat ctc act gac tac ctg tcc Ile Thr Trp Asp Gly Leu Arg Thr Gln Tyr Leu Thr Asp Tyr Leu Ser 495 500 505			1584
caa ctc ctg ctc gcc gtc cat aag gat ggg att aat ctt cgc ggc gcg Gln Leu Leu Leu Ala Val His Lys Asp Gly Ile Asn Leu Arg Gly Ala 510 515 520 525			1632
tta acc tgg agt ttc gtc gac aac tgg gaa tgg gga ctg ggg atg caa Leu Thr Trp Ser Phe Val Asp Asn Trp Glu Trp Gly Leu Gly Met Gln 530 535 540			1680
cag aaa ttc gga ttc cag ttt gtc aat cag tcg gac cca gat ctc acc Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp Leu Thr 545 550 555			1728
agg acc ttc aaa ctc tct gcg cac gct tac gct caa ttt ggt cgc aac Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly Arg Asn 560 565 570			1776

cac ctc tga
His Leu
575

1785

<210> 2
<211> 594
<212> PRT
<213> *Sporobolomyces singularis*

<400> 2

Met Met Leu His Ala Ala Leu Leu Val Ala Leu Pro Cys Val Val Leu
-15 -10 -5

Ala Arg Pro Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro Leu Ser
-1 1 5 10

Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro Leu Glu
15 20 25

Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu Trp Asn
30 35 40 45

Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr Thr Leu
50 55 60

Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro Ser Tyr
65 70 75

Ala Leu Ala Gly Tyr Glu Thr Ser Glu Ile Ala Gly Leu Lys Phe Pro
80 85 90

Lys Gly Phe Lys Phe Gly Val Ala Gly Ala Ala Ile Gln Val Glu Gly
95 100 105

Ala Ala Lys Ala Glu Gly Arg Gly Pro Ser Thr Trp Asp Tyr Leu Cys
110 115 120 125

His His Tyr Ala Ser Thr Gln Cys Asn Asn Tyr Asp Pro Asp Ile Thr
130 135 140

Thr Asn His Tyr Tyr Leu Tyr Pro Leu Asp Phe Ala Arg Leu Gln His
145 150 155

Leu Gly Ile Asn Thr Tyr Ser Phe Ser Ile Ser Trp Thr Arg Ile Tyr
160 165 170

Pro Leu Gly Ala Gly Tyr Val Asn Glu Ala Gly Leu Ala His Tyr Asp
 175 180 185

Ala Val Ile His Ser Ala Lys Lys Tyr Gly Leu Glu Pro Val Gly Thr
 190 195 200 205

Val Phe His Trp Asp Thr Pro Leu Ser Leu Met Leu Lys Tyr Gly Ala
 210 215 220

Trp Gln Asp Thr Gly Asp Gln Ile Val Lys Asp Phe Val Thr Tyr Ala
 225 230 235

Thr Thr Val Phe Lys Arg Tyr Gly Asn Glu Val Lys Thr Trp Phe Thr
 240 245 250

Phe Asn Glu Pro Arg Val Phe Cys Ser Gln Asn Ser Gly Leu Pro Tyr
 255 260 265

Asn Leu Thr Tyr Pro Glu Gly Ile Asn Ser Thr Ser Ala Val Phe Arg
 270 275 280 285

Cys Thr Tyr Asn Val Leu Lys Ala His Gly His Ala Val Lys Val Tyr
 290 295 300

Arg Asp Leu Val Ala Ser Gly Thr Ile Ala Ala Gly Glu Ile Gly Phe
 305 310 315

Lys Ser Asp Asp Asn Tyr Pro Ile Pro Ala Arg Pro Gly Asn Ala Asp
 320 325 330

Asp Glu Glu Ser Ala Lys Arg His Glu Ala Phe Arg Ile Gly Ile Phe
 335 340 345

Ala Gln Pro Val Tyr Gly Asn Gly Asp Tyr Pro Asp Val Val Lys Glu
 350 355 360 365

Thr Val Gly Asp Met Leu Pro Ala Leu Thr Asp Glu Asp Lys Gly Tyr
 370 375 380

Ile Lys Gly Ser Gly Asp Ile Phe Ala Ile Asp Gly Tyr Arg Thr Asp
 385 390 395

Ile Ser His Ala Ala Leu Asn Gly Ile Ala Asn Cys Ile Arg Asn Gln
 400 405 410

Ser Asp Pro Asn Trp Pro Val Cys Glu Glu Gly Ser Asp Pro Phe Ala
 5

415 420 425
 His Val Tyr Pro Ser Gly Phe Ala Ile Gly Gln Ser Ala Asp Pro Leu
 430 435 440 445
 Ser Ser Trp Leu Val Asn Ser Ala Pro Phe Ile Arg Asp Gln Leu Lys
 450 455 460
 Phe Leu Thr Gln Thr Tyr Pro Ala Lys Gly Gly Ile Tyr Phe Ser Glu
 465 470 475
 Phe Gly Trp Ala Glu Asp Ala Glu Tyr Asp Arg Gln Leu Leu Tyr Gln
 480 485 490
 Ile Thr Trp Asp Gly Leu Arg Thr Gln Tyr Leu Thr Asp Tyr Leu Ser
 495 500 505
 Gln Leu Leu Leu Ala Val His Lys Asp Gly Ile Asn Leu Arg Gly Ala
 510 515 520 525
 Leu Thr Trp Ser Phe Val Asp Asn Trp Glu Trp Gly Leu Gly Met Gln
 530 535 540
 Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp Leu Thr
 545 550 555
 Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly Arg Asn
 560 565 570
 His Leu
 575

<210> 3
 <211> 1791
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(1791)

<220>
 <221> sig_peptit
 <222> (1)..(63)

<220>

<221> mat_peptit
 <222> (64)..(1788)

<400> 3

atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg	48
Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala	
-20 -15 -10	
gca cct gct ttg gct gcc ggt gca gtt acc tac ccc ggt gcg att cca	96
Ala Pro Ala Leu Ala Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro	
-5 -1 1 5 10	
ctt agc ttg acc agc aat tac gag acg ccg agt ccg acc gcc atc ccc	144
Leu Ser Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro	
15 20 25	
ctg gag ccg acc cca acg gcg acc gga acc gcc gaa ctt gat gcg ctc	192
Leu Glu Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu	
30 35 40	
tgg aat ttg gtg gaa gca cag tac cct gtt cag acg gcg gct gtc acc	240
Trp Asn Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr	
45 50 55	
acc ctg gtg acg gtg ccc gac gac tac aag ttt gaa gca gac cct cct	288
Thr Leu Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro	
60 65 70 75	
tcc tat gct ctt gct ggc tac gag aca tca gaa att gcc ggc ttg aag	336
Ser Tyr Ala Leu Ala Gly Tyr Glu Thr Ser Glu Ile Ala Gly Leu Lys	
80 85 90	
ttc ccg aag ggg ttc aag ttt ggc gtg gcc ggc gcg gct att caa gtg	384
Phe Pro Lys Gly Phe Lys Phe Gly Val Ala Gly Ala Ala Ile Gln Val	
95 100 105	
gaa ggc gca gcg aaa gca gag gga cga ggc cca tcc act tgg gat tac	432
Glu Gly Ala Ala Lys Ala Glu Gly Arg Gly Pro Ser Thr Trp Asp Tyr	
110 115 120	
ttg tgc cac cat tac gcg tcc aca cag tgc aac aac tat gat cct gac	480
Leu Cys His His Tyr Ala Ser Thr Gln Cys Asn Asn Tyr Asp Pro Asp	
125 130 135	
att acg acg aac cat tac tac ctt tac cct ctt gat ttc gcc cgg ctc	528
Ile Thr Thr Asn His Tyr Tyr Leu Tyr Pro Leu Asp Phe Ala Arg Leu	
140 145 150 155	
cag cat cta ggc atc aac acg tat tcg ttt tca atc tcc tgg act cgt	576
Gln His Leu Gly Ile Asn Thr Tyr Ser Phe Ser Ile Ser Trp Thr Arg	
160 165 170	
ata tac cct ctg ggt gct ggc tac gtt aac gaa gcc ggt ttg gcg cat	624
Ile Tyr Pro Leu Gly Ala Gly Tyr Val Asn Glu Ala Gly Leu Ala His	
175 180 185	
tac gac gcg gta atc cac tcg gcc aag aag tac ggg ctg gag cct gtc	672
Tyr Asp Ala Val Ile His Ser Ala Lys Lys Tyr Gly Leu Glu Pro Val	
190 195 200	
gga aca gta ttt cac tgg gac acc cct ctc agc ctc atg ctc aaa tat	720

Gly Thr Val Phe His Trp Asp Thr Pro Leu Ser Leu Met Leu Lys Tyr	
205 210 215	
ggc gcg tgg caa gat acc ggc gac cag atc gtt aaa gat ttc gtc aca	768
Gly Ala Trp Gln Asp Thr Gly Asp Gln Ile Val Lys Asp Phe Val Thr	
220 225 230 235	
tac gcc acc acc gtc ttc aaa cga tac ggt aat gaa gtc aag acc tgg	816
Tyr Ala Thr Thr Val Phe Lys Arg Tyr Gly Asn Glu Val Lys Thr Trp	
240 245 250	
ttc acg ttc aat gag cct cgc gtg ttc tgt tct caa aac agt ggc ctt	864
Phe Thr Phe Asn Glu Pro Arg Val Phe Cys Ser Gln Asn Ser Gly Leu	
255 260 265	
ccc tat aac ctc acg tat cct gag gga atc aac tca act tca gcc gtc	912
Pro Tyr Asn Leu Thr Tyr Pro Glu Gly Ile Asn Ser Thr Ser Ala Val	
270 275 280	
ttc cgg tgt act tat aac gtc ctg aaa gcc cat ggc cac gcg gtt aag	960
Phe Arg Cys Thr Tyr Asn Val Leu Lys Ala His Gly His Ala Val Lys	
285 290 295	
gtt tac cgg gat ctc gtt gcc agc gga acc att gct gct gga gag atc	1008
Val Tyr Arg Asp Leu Val Ala Ser Gly Thr Ile Ala Ala Gly Glu Ile	
300 305 310 315	
ggc ttc aag tcg gac gac aac tac cca atc cca gcg cgg ccc gga aac	1056
Gly Phe Lys Ser Asp Asp Asn Tyr Pro Ile Pro Ala Arg Pro Gly Asn	
320 325 330	
gcg gac gac gag gaa tcc gcc aaa cgt cac gaa gcg ttc cga atc gga	1104
Ala Asp Asp Glu Glu Ser Ala Lys Arg His Glu Ala Phe Arg Ile Gly	
335 340 345	
atc ttt gcc cag cca gtt tac gga aac ggc gac tat cct gat gta gta	1152
Ile Phe Ala Gln Pro Val Tyr Gly Asn Gly Asp Tyr Pro Asp Val Val	
350 355 360	
aaa gag acc gtt ggc gac atg ctg ccc gcc ctg acg gat gag gac aag	1200
Lys Glu Thr Val Gly Asp Met Leu Pro Ala Leu Thr Asp Glu Asp Lys	
365 370 375	
ggc tac atc aag ggc agc ggc gac atc ttc gcc att gac ggt tac cgg	1248
Gly Tyr Ile Lys Gly Ser Gly Asp Ile Phe Ala Ile Asp Gly Tyr Arg	
380 385 390 395	
acc gat atc tcg cat gcc gca ctg aat gga atc gcg aat tgc atc aga	1296
Thr Asp Ile Ser His Ala Ala Leu Asn Gly Ile Ala Asn Cys Ile Arg	
400 405 410	
aac cag tcg gac cct aac tgg cct gtt tgc gag gaa ggg tct gac ccg	1344
Asn Gln Ser Asp Pro Asn Trp Pro Val Cys Glu Glu Gly Ser Asp Pro	
415 420 425	
ttc gcc cac gta tac ccg tct ggt ttc gcc atc ggc cag tcc gcc gat	1392
Phe Ala His Val Tyr Pro Ser Gly Phe Ala Ile Gly Gln Ser Ala Asp	
430 435 440	
ccg ctg tcg tca tgg ctc gtc aac tcc gcc cca ttt att cgc gac cag	1440
Pro Leu Ser Ser Trp Leu Val Asn Ser Ala Pro Phe Ile Arg Asp Gln	

445	450	455	
ctg aag ttc ctc act caa acg tac ccg gca aag gga ggt att tac ttc			1488
Leu Lys Phe Leu Thr Gln Thr Tyr Pro Ala Lys Gly Gly Ile Tyr Phe			
460	465	470	475
agc gag ttt ggg tgg gcc gag gat gcg gag tac gac cgc cag ctg ttg			1536
Ser Glu Phe Gly Trp Ala Glu Asp Ala Glu Tyr Asp Arg Gln Leu Leu			
	480	485	490
tac caa atc acc tgg gac ggt ctt agg acc cag tat ctc act gac tac			1584
Tyr Gln Ile Thr Trp Asp Gly Leu Arg Thr Gln Tyr Leu Thr Asp Tyr			
	495	500	505
ctg tcc caa ctc ctg ctc gcc gtc cat aag gat ggg att aat ctt cgc			1632
Leu Ser Gln Leu Leu Leu Ala Val His Lys Asp Gly Ile Asn Leu Arg			
	510	515	520
ggc gcg tta acc tgg agt ttc gtc gac aac tgg gaa tgg gga ctg ggg			1680
Gly Ala Leu Thr Trp Ser Phe Val Asp Asn Trp Glu Trp Gly Leu Gly			
	525	530	535
atg caa cag aaa ttc gga ttc cag ttt gtc aat cag tcg gac cca gat			1728
Met Gln Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp			
	540	545	550
ctc acc agg acc ttc aaa ctc tct gcg cac gct tac gct caa ttt ggt			1776
Leu Thr Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly			
	560	565	570
cgc aac cac ctc tga			1791
Arg Asn His Leu			
	575		
<210>	4		
<211>	596		
<212>	PRT		
<213>	Trình tự nhân tạo		
<220>			
<223>	Cấu trúc tổng hợp		
<400>	4		
Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala			
-20	-15	-10	
Ala Pro Ala Leu Ala Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro			
-5	-1 1	5	10
Leu Ser Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro			
15	20	25	
Leu Glu Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu			
30	35	40	

Trp Asn Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr
 45 50 55

Thr Leu Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro
 60 65 70 75

Ser Tyr Ala Leu Ala Gly Tyr Glu Thr Ser Glu Ile Ala Gly Leu Lys
 80 85 90

Phe Pro Lys Gly Phe Lys Phe Gly Val Ala Gly Ala Ala Ile Gln Val
 95 100 105

Glu Gly Ala Ala Lys Ala Glu Gly Arg Gly Pro Ser Thr Trp Asp Tyr
 110 115 120

Leu Cys His His Tyr Ala Ser Thr Gln Cys Asn Asn Tyr Asp Pro Asp
 125 130 135

Ile Thr Thr Asn His Tyr Tyr Leu Tyr Pro Leu Asp Phe Ala Arg Leu
 140 145 150 155

Gln His Leu Gly Ile Asn Thr Tyr Ser Phe Ser Ile Ser Trp Thr Arg
 160 165 170

Ile Tyr Pro Leu Gly Ala Gly Tyr Val Asn Glu Ala Gly Leu Ala His
 175 180 185

Tyr Asp Ala Val Ile His Ser Ala Lys Lys Tyr Gly Leu Glu Pro Val
 190 195 200

Gly Thr Val Phe His Trp Asp Thr Pro Leu Ser Leu Met Leu Lys Tyr
 205 210 215

Gly Ala Trp Gln Asp Thr Gly Asp Gln Ile Val Lys Asp Phe Val Thr
 220 225 230 235

Tyr Ala Thr Thr Val Phe Lys Arg Tyr Gly Asn Glu Val Lys Thr Trp
 240 245 250

Phe Thr Phe Asn Glu Pro Arg Val Phe Cys Ser Gln Asn Ser Gly Leu
 255 260 265

Pro Tyr Asn Leu Thr Tyr Pro Glu Gly Ile Asn Ser Thr Ser Ala Val
 270 275 280

Phe Arg Cys Thr Tyr Asn Val Leu Lys Ala His Gly His Ala Val Lys
 10

285	290	295
Val Tyr Arg Asp Leu	Val Ala Ser Gly Thr Ile	Ala Ala Gly Glu Ile
300	305	310 315
Gly Phe Lys Ser Asp	Asp Asn Tyr Pro Ile	Pro Ala Arg Pro Gly Asn
	320	325 330
Ala Asp Asp Glu Glu Ser	Ala Lys Arg His Glu	Ala Phe Arg Ile Gly
	335	340 345
Ile Phe Ala Gln Pro	Val Tyr Gly Asn Gly Asp Tyr	Pro Asp Val Val
	350	355 360
Lys Glu Thr Val Gly	Asp Met Leu Pro Ala Leu	Thr Asp Glu Asp Lys
	365	370 375
Gly Tyr Ile Lys Gly	Ser Gly Asp Ile Phe Ala	Ile Asp Gly Tyr Arg
380	385	390 395
Thr Asp Ile Ser His	Ala Ala Leu Asn Gly Ile	Ala Asn Cys Ile Arg
	400	405 410
Asn Gln Ser Asp Pro	Asn Trp Pro Val Cys Glu Glu	Gly Ser Asp Pro
	415	420 425
Phe Ala His Val Tyr	Pro Ser Gly Phe Ala Ile	Gly Gln Ser Ala Asp
	430	435 440
Pro Leu Ser Ser Trp	Leu Val Asn Ser Ala Pro	Phe Ile Arg Asp Gln
	445	450 455
Leu Lys Phe Leu Thr	Gln Thr Tyr Pro Ala Lys	Gly Gly Ile Tyr Phe
460	465	470 475
Ser Glu Phe Gly Trp	Ala Glu Asp Ala Glu Tyr	Asp Arg Gln Leu Leu
	480	485 490
Tyr Gln Ile Thr Trp	Asp Gly Leu Arg Thr Gln Tyr	Leu Thr Asp Tyr
	495	500 505
Leu Ser Gln Leu Leu	Leu Ala Val His Lys Asp Gly	Ile Asn Leu Arg
	510	515 520
Gly Ala Leu Thr Trp	Ser Phe Val Asp Asn Trp	Glu Trp Gly Leu Gly
	525	530 535

Met Gln Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp
 540 545 550 555

Leu Thr Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly
 560 565 570

Arg Asn His Leu
 575

<210> 5
 <211> 1791
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(1791)

<220>
 <221> sig_peptit
 <222> (1)..(63)

<220>
 <221> mat_peptit
 <222> (64)..(1788)

<400> 5
 atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg 48
 Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

gca cct gct ttg gct gct ggc gcc gtc aca tac ccc gga gct att cca 96
 Ala Pro Ala Leu Ala Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro
 -5 -1 1 5 10

ctg tcc ctc aca agc aac tac gag aca cct tct cct acc gcc att cct 144
 Leu Ser Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro
 15 20 25

ttg gag cca act ccg act gca act gga acc gca gaa ctg gat gcg ctc 192
 Leu Glu Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu
 30 35 40

tgg aac ctt gtc gaa gct cag tat ccc gtc caa acg gcc gca gtc acc 240
 Trp Asn Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr
 45 50 55

acc ctc gtc acc gtc cct gac gac tac aag ttc gag gcc gat cct cct 288
 Thr Leu Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro
 60 65 70 75

agc tat gca ctc gca ggt tac gag act tcc gag att gcg gga ctg aag 336
 12

Ser Tyr Ala Leu Ala Gly Tyr Glu Thr Ser Glu Ile Ala Gly Leu Lys	
80 85 90	
ttc ccc aaa ggt ttc aag ttc ggc gtt gct ggg gcc gcg att cag gtc	384
Phe Pro Lys Gly Phe Lys Phe Gly Val Ala Gly Ala Ile Gln Val	
95 100 105	
gaa ggc gct gct aaa gcg gaa gga cgt ggg ccc tct aca tgg gat tac	432
Glu Gly Ala Ala Lys Ala Glu Gly Arg Gly Pro Ser Thr Trp Asp Tyr	
110 115 120	
ctc tgt cat cac tat gcc agc act cag tgc aac aac tac gat ccc gat	480
Leu Cys His His Tyr Ala Ser Thr Gln Cys Asn Asn Tyr Asp Pro Asp	
125 130 135	
atc acc acc aac cat tac tac ctc tac ccc ctc gat ttc gcg cgt ctt	528
Ile Thr Thr Asn His Tyr Tyr Leu Tyr Pro Leu Asp Phe Ala Arg Leu	
140 145 150 155	
caa cat ctg ggc atc aac acc tac tcc ttt tcc att tcc tgg acc cga	576
Gln His Leu Gly Ile Asn Thr Tyr Ser Phe Ser Ile Ser Trp Thr Arg	
160 165 170	
atc tac cct ctc ggc gcc ggt tac gtc aac gag gcc ggc ttg gca cac	624
Ile Tyr Pro Leu Gly Ala Gly Tyr Val Asn Glu Ala Gly Leu Ala His	
175 180 185	
tac gat gct gtc att cac tcc gcc aag aag tac gga ttg gag cca gtt	672
Tyr Asp Ala Val Ile His Ser Ala Lys Lys Tyr Gly Leu Glu Pro Val	
190 195 200	
ggc acg gtc ttt cac tgg gac act cct ctg tcg ctc atg ctt aag tac	720
Gly Thr Val Phe His Trp Asp Thr Pro Leu Ser Leu Met Leu Lys Tyr	
205 210 215	
ggg gcg tgg cag gat act ggt gat cag atc gtc aag gac ttt gtg acg	768
Gly Ala Trp Gln Asp Thr Gly Asp Gln Ile Val Lys Asp Phe Val Thr	
220 225 230 235	
tat gcc acg acc gtt ttc aag cgc tat ggt aac gag gtc aag aca tgg	816
Tyr Ala Thr Thr Val Phe Lys Arg Tyr Gly Asn Glu Val Lys Thr Trp	
240 245 250	
ttc aca ttc aac gag cca cgt gtc ttc tgc tcc cag aat tcc ggg ctt	864
Phe Thr Phe Asn Glu Pro Arg Val Phe Cys Ser Gln Asn Ser Gly Leu	
255 260 265	
ccg tac aac ctg acc tat cct gaa ggc atc aac tct act tct gcg gtg	912
Pro Tyr Asn Leu Thr Tyr Pro Glu Gly Ile Asn Ser Thr Ser Ala Val	
270 275 280	
ttc cgt tgc acg tac aac gtg ctt aag gct cat ggt cat gct gtc aaa	960
Phe Arg Cys Thr Tyr Asn Val Leu Lys Ala His Gly His Ala Val Lys	
285 290 295	
gtg tat cga gat ctt gtg gca tcg ggt aca atc gct gcc ggc gag atc	1008
Val Tyr Arg Asp Leu Val Ala Ser Gly Thr Ile Ala Ala Gly Glu Ile	
300 305 310 315	
ggc ttc aag agc gac gac aac tac ccg atc ccg gct cgg cct ggt aat	1056
Gly Phe Lys Ser Asp Asp Asn Tyr Pro Ile Pro Ala Arg Pro Gly Asn	

320	325	330	
gcc gac gac gag gag tcg gcc aag cgc cac gaa gca ttt cga atc ggc Ala Asp Asp Glu Glu Ser Ala Lys Arg His Glu Ala Phe Arg Ile Gly 335 340 345			1104
atc ttc gcc cag cct gtg tat ggg aat ggt gac tat ccc gat gtg gtg Ile Phe Ala Gln Pro Val Tyr Gly Asn Gly Asp Tyr Pro Asp Val Val 350 355 360			1152
aag gag acc gtg ggc gac atg ctc ccc gcc ctt acc gat gag gac aaa Lys Glu Thr Val Gly Asp Met Leu Pro Ala Leu Thr Asp Glu Asp Lys 365 370 375			1200
ggt tac atc aag ggc tcg ggc gac atc ttc gcg att gac ggc tat cgg Gly Tyr Ile Lys Gly Ser Gly Asp Ile Phe Ala Ile Asp Gly Tyr Arg 380 385 390 395			1248
act gac atc tcg cac gcg gct ctg aat ggc atc gca aac tgc att cgc Thr Asp Ile Ser His Ala Ala Leu Asn Gly Ile Ala Asn Cys Ile Arg 400 405 410			1296
aat cag agc gac cct aac tgg ccg gtg tgt gaa gaa gga agc gat ccg Asn Gln Ser Asp Pro Asn Trp Pro Val Cys Glu Glu Gly Ser Asp Pro 415 420 425			1344
ttt gcc cat gtg tat ccc tct ggc ttt gca att gga caa agc gct gat Phe Ala His Val Tyr Pro Ser Gly Phe Ala Ile Gly Gln Ser Ala Asp 430 435 440			1392
cct ctg tct tct tgg ctc gtt aac tcc gct ccc ttc atc cga gat caa Pro Leu Ser Ser Trp Leu Val Asn Ser Ala Pro Phe Ile Arg Asp Gln 445 450 455			1440
ctg aag ttc ctg acg caa acc tat cct gca aaa ggc ggc atc tac ttc Leu Lys Phe Leu Thr Gln Thr Tyr Pro Ala Lys Gly Gly Ile Tyr Phe 460 465 470 475			1488
tcc gag ttt gga tgg gca gag gat gcg gaa tac gat cgg caa ctc ctt Ser Glu Phe Gly Trp Ala Glu Asp Ala Glu Tyr Asp Arg Gln Leu Leu 480 485 490			1536
tac cag atc acc tgg gac ggg ctt cgc act cag tat ctc acc gat tac Tyr Gln Ile Thr Trp Asp Gly Leu Arg Thr Gln Tyr Leu Thr Asp Tyr 495 500 505			1584
ctc tcc cag ctg ctg ttg gct gtg cat aag gat ggg atc aac ctt cgg Leu Ser Gln Leu Leu Leu Ala Val His Lys Asp Gly Ile Asn Leu Arg 510 515 520			1632
ggc gca ttg acc tgg tct ttc gtg gac aac tgg gaa tgg ggt ctg ggc Gly Ala Leu Thr Trp Ser Phe Val Asp Asn Trp Glu Trp Gly Leu Gly 525 530 535			1680
atg cag cag aag ttc gga ttc caa ttc gtc aat cag agc gat cca gac Met Gln Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp 540 545 550 555			1728
ttg acc cgc aca ttc aag ctc agc gct cat gcg tat gcc cag ttc ggg Leu Thr Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly 560 565 570			1776

cgc aat cac ctg taa
 Arg Asn His Leu
 575

1791

<210> 6
 <211> 596
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 6

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ala Gly Ala Val Thr Tyr Pro Gly Ala Ile Pro
 -5 -1 1 5 10

Leu Ser Leu Thr Ser Asn Tyr Glu Thr Pro Ser Pro Thr Ala Ile Pro
 15 20 25

Leu Glu Pro Thr Pro Thr Ala Thr Gly Thr Ala Glu Leu Asp Ala Leu
 30 35 40

Trp Asn Leu Val Glu Ala Gln Tyr Pro Val Gln Thr Ala Ala Val Thr
 45 50 55

Thr Leu Val Thr Val Pro Asp Asp Tyr Lys Phe Glu Ala Asp Pro Pro
 60 65 70 75

Ser Tyr Ala Leu Ala Gly Tyr Glu Thr Ser Glu Ile Ala Gly Leu Lys
 80 85 90

Phe Pro Lys Gly Phe Lys Phe Gly Val Ala Gly Ala Ala Ile Gln Val
 95 100 105

Glu Gly Ala Ala Lys Ala Glu Gly Arg Gly Pro Ser Thr Trp Asp Tyr
 110 115 120

Leu Cys His His Tyr Ala Ser Thr Gln Cys Asn Asn Tyr Asp Pro Asp
 125 130 135

Ile Thr Thr Asn His Tyr Tyr Leu Tyr Pro Leu Asp Phe Ala Arg Leu
 140 145 150 155

Gln His Leu Gly Ile Asn Thr Tyr Ser Phe Ser Ile Ser Trp Thr Arg
 15

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

Asn Gln Ser Asp Pro Asn Trp Pro Val Cys Glu Glu Gly Ser Asp Pro
 415 420 425

Phe Ala His Val Tyr Pro Ser Gly Phe Ala Ile Gly Gln Ser Ala Asp
 430 435 440

Pro Leu Ser Ser Trp Leu Val Asn Ser Ala Pro Phe Ile Arg Asp Gln
 445 450 455

Leu Lys Phe Leu Thr Gln Thr Tyr Pro Ala Lys Gly Gly Ile Tyr Phe
 460 465 470 475

Ser Glu Phe Gly Trp Ala Glu Asp Ala Glu Tyr Asp Arg Gln Leu Leu
 480 485 490

Tyr Gln Ile Thr Trp Asp Gly Leu Arg Thr Gln Tyr Leu Thr Asp Tyr
 495 500 505

Leu Ser Gln Leu Leu Leu Ala Val His Lys Asp Gly Ile Asn Leu Arg
 510 515 520

Gly Ala Leu Thr Trp Ser Phe Val Asp Asn Trp Glu Trp Gly Leu Gly
 525 530 535

Met Gln Gln Lys Phe Gly Phe Gln Phe Val Asn Gln Ser Asp Pro Asp
 540 545 550 555

Leu Thr Arg Thr Phe Lys Leu Ser Ala His Ala Tyr Ala Gln Phe Gly
 560 565 570

Arg Asn His Leu
 575

<210> 7
 <211> 2106
 <212> ADN
 <213> Sirobasidium magnum

<220>
 <221> CDS
 <222> (1)..(2106)

<220>
 <221> sig_peptit
 <222> (1)..(48)

<220>

<221> mat_peptit
 <222> (49)..(2103)

<400> 7

atg ttc aag ctt acc tct gtg ctg ttg ctg ctc ggt gca gct caa gca	48
Met Phe Lys Leu Thr Ser Val Leu Leu Leu Gly Ala Ala Gln Ala	
-15 -10 -5 -1	
gct gtt cta aac cct cgt caa gct ggc agc ggt aat tcc act gcc agc	96
Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly Asn Ser Thr Ala Ser	
1 5 10 15	
ggc tcg ata gcc ggc gat tcc act aga cca gcc acg aca tcc tcg gtc	144
Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala Thr Thr Ser Ser Val	
20 25 30	
gtc tca ccc tct gca gcc aga aac tcc act gcc gca gct act ggt aat	192
Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala Ala Ala Thr Gly Asn	
35 40 45	
gct tct cgc aat gct act gcg aca ggt act gcc gtc gct aca gcc act	240
Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala Val Ala Thr Ala Thr	
50 55 60	
ggc ggg gtt aca gca gcc acg tcc act gga atg gcg gtg act tcc cct	288
Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met Ala Val Thr Ser Pro	
65 70 75 80	
gcc cag gga gcc ggt acc gga gtc ggt acc gca gcc gct gct acg acg	336
Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala Ala Ala Thr Thr	
85 90 95	
act acc gcc acg cct agc caa tcc gac ttt gat aat tgg gtc ctc acc	384
Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp Asn Trp Val Leu Thr	
100 105 110	
agt gga ttg cct acc atc acc act tca ttg atc agt acc aat ccc gat	432
Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile Ser Thr Asn Pro Asp	
115 120 125	
gcc att act ccg act gcc agt act tca gga ccg aag cct acg gtc acg	480
Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro Lys Pro Thr Val Thr	
130 135 140	
ttc agc tcg tac tcg gac caa gag ctg gag aat ctc tgg gac gac ttt	528
Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn Leu Trp Asp Asp Phe	
145 150 155 160	
gtg gga caa gta caa caa cct cca ttc agc tat gtt cca gaa ccc caa	576
Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr Val Pro Glu Pro Gln	
165 170 175	
aac ccc tat cct ctg cca aac acc cca cca tcc ctc tat cca gac tgg	624
Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser Leu Tyr Pro Asp Trp	
180 185 190	
tac gtc aat tgc cct aca aag agt cta ccg ggg tac aaa ttc ccc aga	672
Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly Tyr Lys Phe Pro Arg	
195 200 205	
gga ttc ctg ttc ggc tgg gct aca gct gcg caa cag tgg gaa ggg gct	720
18	

Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln Gln Trp Glu Gly Ala	
210 215 220	
gtc aag gcg gat ggt aag ggt cct agt atc tgg gac tgg gca agt aga	768
Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp Asp Trp Ala Ser Arg	
225 230 235 240	
tac ccc ggc ttc atc gcg gac aac act act tct gat gtg gga gat ctg	816
Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser Asp Val Gly Asp Leu	
245 250 255	
gga tat tac cta tac aaa gaa gat atg gca cgc ctc gct gcg ttg gga	864
Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg Leu Ala Ala Leu Gly	
260 265 270	
gga aac gtc tac tct ttc tcc atc ttc tgg act cgt atc ctc ccc ttt	912
Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr Arg Ile Leu Pro Phe	
275 280 285	
gcg gtc caa gga tcc ccc gtg aac caa aag gga gta gac ttt tat cgg	960
Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly Val Asp Phe Tyr Arg	
290 295 300	
gac ttg atc gat tat tgc tgg agt ttg ggt atc gag cct gtc gtg aca	1008
Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile Glu Pro Val Val Thr	
305 310 315 320	
ctg ttc cac tgg gat aca cct tta gcg gtg caa ctg ctc tat gga gga	1056
Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln Leu Leu Tyr Gly Gly	
325 330 335	
ttc gca agt gac aag atc att gat gat tat gtc aat tat gcc gaa acg	1104
Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val Asn Tyr Ala Glu Thr	
340 345 350	
gtg ttc act gcc tat aat ggc tcg gtt cac aaa tgg atc acc ttc aac	1152
Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys Trp Ile Thr Phe Asn	
355 360 365	
gaa cca gta gta ttc tgc agc cag atg gct tct cct gtg aat tca aca	1200
Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser Pro Val Asn Ser Thr	
370 375 380	
ctg ccc gaa ggg ttg aac agc acc aca tac cca tac aca tgt agc tac	1248
Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro Tyr Thr Cys Ser Tyr	
385 390 395 400	
cat ctc acc ctg gct cac gcc aag acc gtc caa cga ttc aga gag ctc	1296
His Leu Thr Leu Ala His Ala Lys Thr Val Gln Arg Phe Arg Glu Leu	
405 410 415	
aac atc cag gga gag att gcg ctc aag tcg gac aac ttt aat ggt atc	1344
Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp Asn Phe Asn Gly Ile	
420 425 430	
cct tgg agg gaa ggg aat ccc gac gat gaa gaa gcc gtt gct agg cat	1392
Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu Ala Val Ala Arg His	
435 440 445	
tct gca tac cag att ggc atc ttt gcg gaa ccg ata tac aac act ggc	1440
Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro Ile Tyr Asn Thr Gly	

450	455	460	
gac tgg cca gaa ctg atc aag aac gat ctt gga ccc gac atc ttg ccc Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly Pro Asp Ile Leu Pro 465	470	475	480 1488
cga ttc acc gat gag cag atc cag atg atc aag ggt act gcc gac ttc Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys Gly Thr Ala Asp Phe 485	490	495	1536
ttt gcc att gat ggg tat cga gat ggc tgg gtc act gcc cca cct gct Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val Thr Ala Pro Pro Ala 500	505	510	1584
gga gtg cag gct tgc gtg gcc aat atc agt gat ccc ctc tgg cct gtg Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp Pro Leu Trp Pro Val 515	520	525	1632
tgc aat caa gtc aac ttc tac gac tct tct ccc gca ggt tgg gga atc Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro Ala Gly Trp Gly Ile 530	535	540	1680
gga gcg ttt ggt aat tgg cct acc act ccc tgg ctg caa aac act tgg Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp Leu Gln Asn Thr Trp 545	550	555	560 1728
caa ttt gtc cgg cca ttt ttg aaa gaa ttg act cag cag tac ccc acc Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr Gln Gln Tyr Pro Thr 565	570	575	1776
aaa ggt ggt atc tac ctc tcg gaa ttt ggc ttc tcc gaa cca ttc gag Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe Ser Glu Pro Phe Glu 580	585	590	1824
aac gag aaa aac ttc atc tac cag atc acg act gac ccg gga cgg gtg Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr Asp Pro Gly Arg Val 595	600	605	1872
gca tac ttt aac agt tac ctc ggt gaa gtg ctc ttg gcg atc aac gag Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu Leu Ala Ile Asn Glu 610	615	620	1920
gat gaa aca gat gtg aga ggg act ttt gga tgg agt ctt ttg gac aac Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp Ser Leu Leu Asp Asn 625	630	635	640 1968
ttt gag tgg aac tcg ggg ttg tcg act cgg ttc ggt gtc caa tat gtc Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe Gly Val Gln Tyr Val 645	650	655	2016
gat tac aac agt cct acg ctc gaa agg acg ttc aag cgc tct gcg atc Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe Lys Arg Ser Ala Ile 660	665	670	2064
gag atg agc cag ttc tgg aac act cat cgt tgc gag gac tag Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys Glu Asp 675	680	685	2106

<210> 8

<211> 701

<212> PRT

<213> Sirobasidium magnum

<400> 8

Met Phe Lys Leu Thr Ser Val Leu Leu Leu Leu Gly Ala Ala Gln Ala
 -15 -10 -5 -1

Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly Asn Ser Thr Ala Ser
 1 5 10 15

Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala Thr Thr Ser Ser Val
 20 25 30

Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala Ala Ala Thr Gly Asn
 35 40 45

Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala Val Ala Thr Ala Thr
 50 55 60

Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met Ala Val Thr Ser Pro
 65 70 75 80

Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala Ala Ala Ala Thr Thr
 85 90 95

Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp Asn Trp Val Leu Thr
 100 105 110

Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile Ser Thr Asn Pro Asp
 115 120 125

Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro Lys Pro Thr Val Thr
 130 135 140

Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn Leu Trp Asp Asp Phe
 145 150 155 160

Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr Val Pro Glu Pro Gln
 165 170 175

Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser Leu Tyr Pro Asp Trp
 180 185 190

Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly Tyr Lys Phe Pro Arg
 195 200 205

Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln Gln Trp Glu Gly Ala
 210 215 220

Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp Asp Trp Ala Ser Arg
 225 230 235 240

Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser Asp Val Gly Asp Leu
 245 250 255

Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg Leu Ala Ala Leu Gly
 260 265 270

Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr Arg Ile Leu Pro Phe
 275 280 285

Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly Val Asp Phe Tyr Arg
 290 295 300

Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile Glu Pro Val Val Thr
 305 310 315 320

Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln Leu Leu Tyr Gly Gly
 325 330 335

Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val Asn Tyr Ala Glu Thr
 340 345 350

Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys Trp Ile Thr Phe Asn
 355 360 365

Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser Pro Val Asn Ser Thr
 370 375 380

Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro Tyr Thr Cys Ser Tyr
 385 390 395 400

His Leu Thr Leu Ala His Ala Lys Thr Val Gln Arg Phe Arg Glu Leu
 405 410 415

Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp Asn Phe Asn Gly Ile
 420 425 430

Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu Ala Val Ala Arg His
 435 440 445

Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro Ile Tyr Asn Thr Gly
 22

```

450              455              460
Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly Pro Asp Ile Leu Pro
465              470              475              480
Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys Gly Thr Ala Asp Phe
485              490              495
Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val Thr Ala Pro Pro Ala
500              505              510
Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp Pro Leu Trp Pro Val
515              520              525
Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro Ala Gly Trp Gly Ile
530              535              540
Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp Leu Gln Asn Thr Trp
545              550              555              560
Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr Gln Gln Tyr Pro Thr
565              570              575
Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe Ser Glu Pro Phe Glu
580              585              590
Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr Asp Pro Gly Arg Val
595              600              605
Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu Leu Ala Ile Asn Glu
610              615              620
Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp Ser Leu Leu Asp Asn
625              630              635              640
Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe Gly Val Gln Tyr Val
645              650              655
Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe Lys Arg Ser Ala Ile
660              665              670
Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys Glu Asp
675              680              685

```

<210> 9
<211> 2121

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(2121)

<220>
 <221> sig_peptit
 <222> (1)..(63)

<220>
 <221> mat_peptit
 <222> (64)..(2118)

<400> 9
 atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg 48
 Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

gca cct gct ttg gct gct gtt cta aac cct cgt caa gct ggc agc ggt 96
 Ala Pro Ala Leu Ala Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly
 -5 -1 1 5 10

aat tcc act gcc agc ggc tcg ata gcc ggc gat tcc act aga cca gcc 144
 Asn Ser Thr Ala Ser Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala
 15 20 25

acg aca tcc tcg gtc gtc tca ccc tct gca gcc aga aac tcc act gcc 192
 Thr Thr Ser Ser Val Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala
 30 35 40

gca gct act ggt aat gct tct cgc aat gct act gcg aca ggt act gcc 240
 Ala Ala Thr Gly Asn Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala
 45 50 55

gtc gct aca gcc act ggc ggg gtt aca gca gcc acg tcc act gga atg 288
 Val Ala Thr Ala Thr Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met
 60 65 70 75

gcg gtg act tcc cct gcc cag gga gcc ggt acc gga gtc ggt acc gca 336
 Ala Val Thr Ser Pro Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala
 80 85 90

gcc gct gct acg acg act acc gcc acg cct agc caa tcc gac ttt gat 384
 Ala Ala Ala Thr Thr Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp
 95 100 105

aat tgg gtc ctc acc agt gga ttg cct acc atc acc act tca ttg atc 432
 Asn Trp Val Leu Thr Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile
 110 115 120

agt acc aat ccc gat gcc att act ccg act gcc agt act tca gga ccg 480
 Ser Thr Asn Pro Asp Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro
 125 130 135

aag cct acg gtc acg ttc agc tcg tac tcg gac caa gag ctg gag aat 528
 24

Lys Pro Thr Val Thr Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn	
140 145 150 155	
ctc tgg gac gac ttt gtg gga caa gta caa caa cct cca ttc agc tat	576
Leu Trp Asp Asp Phe Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr	
160 165 170	
gtt cca gaa ccc caa aac ccc tat cct ctg cca aac acc cca cca tcc	624
Val Pro Glu Pro Gln Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser	
175 180 185	
ctc tat cca gac tgg tac gtc aat tgc cct aca aag agt cta ccg ggg	672
Leu Tyr Pro Asp Trp Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly	
190 195 200	
tac aaa ttc ccc aga gga ttc ctg ttc ggc tgg gct aca gct gcg caa	720
Tyr Lys Phe Pro Arg Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln	
205 210 215	
cag tgg gaa ggg gct gtc aag gcg gat ggt aag ggt cct agt atc tgg	768
Gln Trp Glu Gly Ala Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp	
220 225 230 235	
gac tgg gca agt aga tac ccc ggc ttc atc gcg gac aac act act tct	816
Asp Trp Ala Ser Arg Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser	
240 245 250	
gat gtg gga gat ctg gga tat tac cta tac aaa gaa gat atg gca cgc	864
Asp Val Gly Asp Leu Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg	
255 260 265	
ctc gct gcg ttg gga gga aac gtc tac tct ttc tcc atc ttc tgg act	912
Leu Ala Ala Leu Gly Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr	
270 275 280	
cgt atc ctc ccc ttt gcg gtc caa gga tcc ccc gtg aac caa aag gga	960
Arg Ile Leu Pro Phe Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly	
285 290 295	
gta gac ttt tat cgg gac ttg atc gat tat tgc tgg agt ttg ggt atc	1008
Val Asp Phe Tyr Arg Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile	
300 305 310 315	
gag cct gtc gtg aca ctg ttc cac tgg gat aca cct tta gcg gtg caa	1056
Glu Pro Val Val Thr Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln	
320 325 330	
ctg ctc tat gga gga ttc gca agt gac aag atc att gat gat tat gtc	1104
Leu Leu Tyr Gly Gly Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val	
335 340 345	
aat tat gcc gaa acg gtg ttc act gcc tat aat ggc tcg gtt cac aaa	1152
Asn Tyr Ala Glu Thr Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys	
350 355 360	
tgg atc acc ttc aac gaa cca gta gta ttc tgc agc cag atg gct tct	1200
Trp Ile Thr Phe Asn Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser	
365 370 375	
cct gtg aat tca aca ctg ccc gaa ggg ttg aac agc acc aca tac cca	1248
Pro Val Asn Ser Thr Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro	
25	

380	385	390	395	
tac aca tgt agc tac cat ctc acc ctg gct cac gcc aag acc gtc caa				1296
Tyr Thr Cys Ser Tyr His Leu Thr Leu Ala His Ala Lys Thr Val Gln				
	400	405	410	
cga ttc aga gag ctc aac atc cag gga gag att gcg ctc aag tcg gac				1344
Arg Phe Arg Glu Leu Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp				
	415	420	425	
aac ttt aat ggt atc cct tgg agg gaa ggg aat ccc gac gat gaa gaa				1392
Asn Phe Asn Gly Ile Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu				
	430	435	440	
gcc gtt gct agg cat tct gca tac cag att ggc atc ttt gcg gaa ccg				1440
Ala Val Ala Arg His Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro				
	445	450	455	
ata tac aac act ggc gac tgg cca gaa ctg atc aag aac gat ctt gga				1488
Ile Tyr Asn Thr Gly Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly				
	460	465	470	475
ccc gac atc ttg ccc cga ttc acc gat gag cag atc cag atg atc aag				1536
Pro Asp Ile Leu Pro Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys				
	480	485	490	
ggt act gcc gac ttc ttt gcc att gat ggg tat cga gat ggc tgg gtc				1584
Gly Thr Ala Asp Phe Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val				
	495	500	505	
act gcc cca cct gct gga gtg cag gct tgc gtg gcc aat atc agt gat				1632
Thr Ala Pro Pro Ala Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp				
	510	515	520	
ccc ctc tgg cct gtg tgc aat caa gtc aac ttc tac gac tct tct ccc				1680
Pro Leu Trp Pro Val Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro				
	525	530	535	
gca ggt tgg gga atc gga gcg ttt ggt aat tgg cct acc act ccc tgg				1728
Ala Gly Trp Gly Ile Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp				
	540	545	550	555
ctg caa aac act tgg caa ttt gtc cgg cca ttt ttg aaa gaa ttg act				1776
Leu Gln Asn Thr Trp Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr				
	560	565	570	
cag cag tac ccc acc aaa ggt ggt atc tac ctc tcg gaa ttt ggc ttc				1824
Gln Gln Tyr Pro Thr Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe				
	575	580	585	
tcc gaa cca ttc gag aac gag aaa aac ttc atc tac cag atc acg act				1872
Ser Glu Pro Phe Glu Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr				
	590	595	600	
gac ccg gga cgg gtg gca tac ttt aac agt tac ctc ggt gaa gtg ctc				1920
Asp Pro Gly Arg Val Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu				
	605	610	615	
ttg gcg atc aac gag gat gaa aca gat gtg aga ggg act ttt gga tgg				1968
Leu Ala Ile Asn Glu Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp				
	620	625	630	635

agt ctt ttg gac aac ttt gag tgg aac tcg ggg ttg tcg act cgg ttc 2016
 Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe
 640 645 650

ggc gtc caa tat gtc gat tac aac agt cct acg ctc gaa agg acg ttc 2064
 Gly Val Gln Tyr Val Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe
 655 660 665

aag cgc tct gcg atc gag atg agc cag ttc tgg aac act cat cgt tgc 2112
 Lys Arg Ser Ala Ile Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys
 670 675 680

gag gac tag 2121
 Glu Asp
 685

<210> 10

<211> 706

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 10

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly
 -5 -1 1 5 10

Asn Ser Thr Ala Ser Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala
 15 20 25

Thr Thr Ser Ser Val Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala
 30 35 40

Ala Ala Thr Gly Asn Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala
 45 50 55

Val Ala Thr Ala Thr Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met
 60 65 70 75

Ala Val Thr Ser Pro Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala
 80 85 90

Ala Ala Ala Thr Thr Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp
 95 100 105

Asn Trp Val Leu Thr Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile
 27

110	115	120
Ser Thr Asn Pro Asp Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro		
125	130	135
Lys Pro Thr Val Thr Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn		
140	145	150
Leu Trp Asp Asp Phe Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr		
160	165	170
Val Pro Glu Pro Gln Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser		
175	180	185
Leu Tyr Pro Asp Trp Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly		
190	195	200
Tyr Lys Phe Pro Arg Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln		
205	210	215
Gln Trp Glu Gly Ala Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp		
220	225	230
Asp Trp Ala Ser Arg Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser		
240	245	250
Asp Val Gly Asp Leu Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg		
255	260	265
Leu Ala Ala Leu Gly Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr		
270	275	280
Arg Ile Leu Pro Phe Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly		
285	290	295
Val Asp Phe Tyr Arg Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile		
300	305	310
Glu Pro Val Val Thr Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln		
320	325	330
Leu Leu Tyr Gly Gly Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val		
335	340	345
Asn Tyr Ala Glu Thr Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys		
350	355	360

Trp Ile Thr Phe Asn Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser
365 370 375

Pro Val Asn Ser Thr Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro
380 385 390 395

Tyr Thr Cys Ser Tyr His Leu Thr Leu Ala His Ala Lys Thr Val Gln
400 405 410

Arg Phe Arg Glu Leu Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp
415 420 425

Asn Phe Asn Gly Ile Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu
430 435 440

Ala Val Ala Arg His Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro
445 450 455

Ile Tyr Asn Thr Gly Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly
460 465 470 475

Pro Asp Ile Leu Pro Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys
480 485 490

Gly Thr Ala Asp Phe Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val
495 500 505

Thr Ala Pro Pro Ala Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp
510 515 520

Pro Leu Trp Pro Val Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro
525 530 535

Ala Gly Trp Gly Ile Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp
540 545 550 555

Leu Gln Asn Thr Trp Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr
560 565 570

Gln Gln Tyr Pro Thr Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe
575 580 585

Ser Glu Pro Phe Glu Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr
590 595 600

Asp Pro Gly Arg Val Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu
605 610 615

Leu Ala Ile Asn Glu Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp
620 625 630 635

Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe
640 645 650

Gly Val Gln Tyr Val Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe
655 660 665

Lys Arg Ser Ala Ile Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys
670 675 680

Glu Asp
685

<210> 11
<211> 2121
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Beta-galactosidaza

<220>
<221> CDS
<222> (1)..(2121)

<220>
<221> sig_peptit
<222> (1)..(63)

<220>
<221> mat_peptit
<222> (64)..(2118)

<400> 11
atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg 48
Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
-20 -15 -10

gca cct gct ttg gct gcg gtg ctc aac cca cgc cag gct ggc tct gga 96
Ala Pro Ala Leu Ala Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly
-5 -1 1 5 10

aac tct act gcc tct ggg tcc att gct ggc gac tcg aca cgc ccg gct 144
Asn Ser Thr Ala Ser Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala
15 20 25

acg acc tct tcg gtt gtc tcc cct tcc gca gcc cga aac tcc act gcc 192
Thr Thr Ser Ser Val Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala
30

30	35	40	
gcc gcc acg ggt aat gct tcc cgc aac gct acg gcc acc gga aca gcc Ala Ala Thr Gly Asn Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala 45 50 55			240
gtt gcc act gcc acg ggc ggc gtt acg gcg gca acg tcc acg ggc atg Val Ala Thr Ala Thr Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met 60 65 70 75			288
gct gtt acg tct ccc gct cag ggt gcc gga acg ggt gtt ggc acc gcc Ala Val Thr Ser Pro Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala 80 85 90			336
gca gcc gca aca aca acg acg gca acg ccc tct caa tcc gac ttc gac Ala Ala Ala Thr Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp 95 100 105			384
aac tgg gtg ctc acg tcg ggg ctc ccc aca atc aca act agc ctg atc Asn Trp Val Leu Thr Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile 110 115 120			432
tcg aca aac cct gac gcc att acc cca acc gca tct acc tcg gga ccc Ser Thr Asn Pro Asp Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro 125 130 135			480
aaa cca acg gtt acc ttc tcc tcc tac tct gat cag gag ttg gag aat Lys Pro Thr Val Thr Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn 140 145 150 155			528
ctg tgg gac gac ttt gtc ggg cag gtt cag cag cct ccg ttc tcg tat Leu Trp Asp Asp Phe Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr 160 165 170			576
gtt cca gaa ccc caa aac ccg tac cca ttg ccc aac acc cca ccg tct Val Pro Glu Pro Gln Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser 175 180 185			624
ctc tac ccc gac tgg tac gtc aat tgt ccc acc aag agc ttg cct ggt Leu Tyr Pro Asp Trp Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly 190 195 200			672
tac aag ttc cct cga ggc ttc ctc ttc ggt tgg gcg acg gct gcg caa Tyr Lys Phe Pro Arg Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln 205 210 215			720
cag tgg gag ggg gct gtg aaa gct gac ggc aaa gga cca tcc atc tgg Gln Trp Glu Gly Ala Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp 220 225 230 235			768
gac tgg gcc tcc cga tat ccg ggc ttc atc gcc gat aac acc aca tcg Asp Trp Ala Ser Arg Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser 240 245 250			816
gac gtg ggc gac ttg ggc tac tac ttg tac aag gag gat atg gct cgt Asp Val Gly Asp Leu Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg 255 260 265			864
ctc gcc gcc ctc ggc ggc aat gtg tat tcc ttt tcc atc ttc tgg acg Leu Ala Ala Leu Gly Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr 270 275 280			912

cga att ctg ccg ttc gca gtg caa ggt tct ccg gtc aat cag aag ggt Arg Ile Leu Pro Phe Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly 285 290 295	960
ggt gac ttc tac cgg gac ctg atc gat tac tgt tgg agc ctt ggt atc Val Asp Phe Tyr Arg Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile 300 305 310 315	1008
gaa ccg gtc gtc aca ttg ttc cac tgg gat act ccg ttg gca gtg caa Glu Pro Val Val Thr Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln 320 325 330	1056
ctg ctc tat gga ggt ttc gcc tcg gac aag atc atc gac gac tac gtc Leu Leu Tyr Gly Gly Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val 335 340 345	1104
aac tac gcc gaa acc gtg ttc aca gcg tac aac ggt tcg gtc cat aag Asn Tyr Ala Glu Thr Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys 350 355 360	1152
tgg atc acc ttc aac gag cca gtt gtg ttc tgt tcc caa atg gca tcg Trp Ile Thr Phe Asn Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser 365 370 375	1200
cca gtg aat agc acg ctg cca gag ggg ttg aac tcg acc acc tac ccc Pro Val Asn Ser Thr Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro 380 385 390 395	1248
tat acg tgc tcg tac cat ctg acc ctt gcg cac gcg aag acc gtg caa Tyr Thr Cys Ser Tyr His Leu Thr Leu Ala His Ala Lys Thr Val Gln 400 405 410	1296
cgg ttc cgt gag ctc aac atc cag ggg gag att gcc ttg aag agc gat Arg Phe Arg Glu Leu Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp 415 420 425	1344
aac ttc aac ggc att ccg tgg cgg gag gga aac cct gat gat gaa gag Asn Phe Asn Gly Ile Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu 430 435 440	1392
gct gtc gcc cga cat tct gcc tat cag atc gga atc ttc gca gaa cct Ala Val Ala Arg His Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro 445 450 455	1440
atc tac aac act ggt gat tgg ccg gag ctg atc aag aac gat ctg ggg Ile Tyr Asn Thr Gly Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly 460 465 470 475	1488
ccg gat atc ttg ccc cgc ttc act gat gaa cag atc cag atg atc aag Pro Asp Ile Leu Pro Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys 480 485 490	1536
ggg acc gct gat ttc ttc gcg atc gat ggt tac cgt gac ggc tgg gtc Gly Thr Ala Asp Phe Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val 495 500 505	1584
aca gca cca cct gct ggt gtc caa gct tgc gtg gcg aac atc tct gat Thr Ala Pro Pro Ala Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp 510 515 520	1632

ccg ttg tgg ccc gtg tgt aac cag gtc aac ttc tac gat tcg agc ccg 1680
 Pro Leu Trp Pro Val Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro
 525 530 535

 gcc ggt tgg gga att ggc gca ttt ggt aac tgg cca act act ccg tgg 1728
 Ala Gly Trp Gly Ile Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp
 540 545 550 555

 ctt cag aac aca tgg cag ttt gtc cgc ccg ttt ctc aag gag ctt act 1776
 Leu Gln Asn Thr Trp Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr
 560 565 570

 cag caa tac cca acg aag ggg ggt atc tac ctc tcc gaa ttc ggt ttc 1824
 Gln Gln Tyr Pro Thr Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe
 575 580 585

 tcg gag ccc ttt gag aac gaa aag aac ttc atc tac caa atc aca aca 1872
 Ser Glu Pro Phe Glu Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr
 590 595 600

 gac ccc ggc cgg gtg gcc tac ttc aac agc tat ctg ggg gaa gtg ctc 1920
 Asp Pro Gly Arg Val Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu
 605 610 615

 ttg gcc atc aac gaa gat gag acg gac gtc cgc ggt act ttc gga tgg 1968
 Leu Ala Ile Asn Glu Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp
 620 625 630 635

 tct ctt ctc gac aac ttc gag tgg aac tct gga ctg agc aca cga ttc 2016
 Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe
 640 645 650

 ggg gtc cag tat gtg gac tac aac tcc cct acg ctt gaa cgg acc ttc 2064
 Gly Val Gln Tyr Val Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe
 655 660 665

 aag cgc tct gcg atc gaa atg tcc cag ttc tgg aac acg cat cgt tgc 2112
 Lys Arg Ser Ala Ile Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys
 670 675 680

 gag gac taa 2121
 Glu Asp
 685

<210> 12
 <211> 706
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 12

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ala Val Leu Asn Pro Arg Gln Ala Gly Ser Gly
 -5 -1 1 5 10

Asn Ser Thr Ala Ser Gly Ser Ile Ala Gly Asp Ser Thr Arg Pro Ala
 15 20 25
 Thr Thr Ser Ser Val Val Ser Pro Ser Ala Ala Arg Asn Ser Thr Ala
 30 35 40
 Ala Ala Thr Gly Asn Ala Ser Arg Asn Ala Thr Ala Thr Gly Thr Ala
 45 50 55
 Val Ala Thr Ala Thr Gly Gly Val Thr Ala Ala Thr Ser Thr Gly Met
 60 65 70 75
 Ala Val Thr Ser Pro Ala Gln Gly Ala Gly Thr Gly Val Gly Thr Ala
 80 85 90
 Ala Ala Ala Thr Thr Thr Thr Ala Thr Pro Ser Gln Ser Asp Phe Asp
 95 100 105
 Asn Trp Val Leu Thr Ser Gly Leu Pro Thr Ile Thr Thr Ser Leu Ile
 110 115 120
 Ser Thr Asn Pro Asp Ala Ile Thr Pro Thr Ala Ser Thr Ser Gly Pro
 125 130 135
 Lys Pro Thr Val Thr Phe Ser Ser Tyr Ser Asp Gln Glu Leu Glu Asn
 140 145 150 155
 Leu Trp Asp Asp Phe Val Gly Gln Val Gln Gln Pro Pro Phe Ser Tyr
 160 165 170
 Val Pro Glu Pro Gln Asn Pro Tyr Pro Leu Pro Asn Thr Pro Pro Ser
 175 180 185
 Leu Tyr Pro Asp Trp Tyr Val Asn Cys Pro Thr Lys Ser Leu Pro Gly
 190 195 200
 Tyr Lys Phe Pro Arg Gly Phe Leu Phe Gly Trp Ala Thr Ala Ala Gln
 205 210 215
 Gln Trp Glu Gly Ala Val Lys Ala Asp Gly Lys Gly Pro Ser Ile Trp
 220 225 230 235
 Asp Trp Ala Ser Arg Tyr Pro Gly Phe Ile Ala Asp Asn Thr Thr Ser
 240 245 250

Asp Val Gly Asp Leu Gly Tyr Tyr Leu Tyr Lys Glu Asp Met Ala Arg
 255 260 265

Leu Ala Ala Leu Gly Gly Asn Val Tyr Ser Phe Ser Ile Phe Trp Thr
 270 275 280

Arg Ile Leu Pro Phe Ala Val Gln Gly Ser Pro Val Asn Gln Lys Gly
 285 290 295

Val Asp Phe Tyr Arg Asp Leu Ile Asp Tyr Cys Trp Ser Leu Gly Ile
 300 305 310 315

Glu Pro Val Val Thr Leu Phe His Trp Asp Thr Pro Leu Ala Val Gln
 320 325 330

Leu Leu Tyr Gly Gly Phe Ala Ser Asp Lys Ile Ile Asp Asp Tyr Val
 335 340 345

Asn Tyr Ala Glu Thr Val Phe Thr Ala Tyr Asn Gly Ser Val His Lys
 350 355 360

Trp Ile Thr Phe Asn Glu Pro Val Val Phe Cys Ser Gln Met Ala Ser
 365 370 375

Pro Val Asn Ser Thr Leu Pro Glu Gly Leu Asn Ser Thr Thr Tyr Pro
 380 385 390 395

Tyr Thr Cys Ser Tyr His Leu Thr Leu Ala His Ala Lys Thr Val Gln
 400 405 410

Arg Phe Arg Glu Leu Asn Ile Gln Gly Glu Ile Ala Leu Lys Ser Asp
 415 420 425

Asn Phe Asn Gly Ile Pro Trp Arg Glu Gly Asn Pro Asp Asp Glu Glu
 430 435 440

Ala Val Ala Arg His Ser Ala Tyr Gln Ile Gly Ile Phe Ala Glu Pro
 445 450 455

Ile Tyr Asn Thr Gly Asp Trp Pro Glu Leu Ile Lys Asn Asp Leu Gly
 460 465 470 475

Pro Asp Ile Leu Pro Arg Phe Thr Asp Glu Gln Ile Gln Met Ile Lys
 480 485 490

Gly Thr Ala Asp Phe Phe Ala Ile Asp Gly Tyr Arg Asp Gly Trp Val
 495 500 505

Thr Ala Pro Pro Ala Gly Val Gln Ala Cys Val Ala Asn Ile Ser Asp
 510 515 520

Pro Leu Trp Pro Val Cys Asn Gln Val Asn Phe Tyr Asp Ser Ser Pro
 525 530 535

Ala Gly Trp Gly Ile Gly Ala Phe Gly Asn Trp Pro Thr Thr Pro Trp
 540 545 550 555

Leu Gln Asn Thr Trp Gln Phe Val Arg Pro Phe Leu Lys Glu Leu Thr
 560 565 570

Gln Gln Tyr Pro Thr Lys Gly Gly Ile Tyr Leu Ser Glu Phe Gly Phe
 575 580 585

Ser Glu Pro Phe Glu Asn Glu Lys Asn Phe Ile Tyr Gln Ile Thr Thr
 590 595 600

Asp Pro Gly Arg Val Ala Tyr Phe Asn Ser Tyr Leu Gly Glu Val Leu
 605 610 615

Leu Ala Ile Asn Glu Asp Glu Thr Asp Val Arg Gly Thr Phe Gly Trp
 620 625 630 635

Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Leu Ser Thr Arg Phe
 640 645 650

Gly Val Gln Tyr Val Asp Tyr Asn Ser Pro Thr Leu Glu Arg Thr Phe
 655 660 665

Lys Arg Ser Ala Ile Glu Met Ser Gln Phe Trp Asn Thr His Arg Cys
 670 675 680

Glu Asp
 685

<210> 13
 <211> 1803
 <212> ADN
 <213> Rodotorula minuta

<220>
 <221> CDS
 <222> (1)..(1803)

```

<220>
<221> sig_peptit
<222> (1)..(57)

<220>
<221> mat_peptit
<222> (58)..(1800)

<400> 13
atg gtc atc ctg cgc cac acg gta cta gcg gca gcg gtc gtc cag atc      48
Met Val Ile Leu Arg His Thr Val Leu Ala Ala Val Val Gln Ile
              -15              -10              -5

gcc ctc ggg gct ccg cag ttt agc ccc aca gtc aat acg gaa ggt atc      96
Ala Leu Gly Ala Pro Gln Phe Ser Pro Thr Val Asn Thr Glu Gly Ile
          -1  1              5              10

att cca att tcg gag cgc cct tcg cct aca caa tct tcg tcg tcc agc      144
Ile Pro Ile Ser Glu Arg Pro Ser Pro Thr Gln Ser Ser Ser Ser Ser
          15              20              25

gga ccc gtg att cca gct ggc tca tat gtc tcc gac ttt gat gcc agc      192
Gly Pro Val Ile Pro Ala Gly Ser Tyr Val Ser Asp Phe Asp Ala Ser
          30              35              40              45

gga ctt gca aac tta tgg agc caa gtc gaa gta gac ata cct gtc gag      240
Gly Leu Ala Asn Leu Trp Ser Gln Val Glu Val Asp Ile Pro Val Glu
              50              55              60

tct cca agg att tct gcc gta cct cct ctc aac aaa acg ttc agt gtt      288
Ser Pro Arg Ile Ser Ala Val Pro Pro Leu Asn Lys Thr Phe Ser Val
              65              70              75

ccg aaa act cca gtc ctg ccc aac tcc ctg cag gat cac ctt cca aaa      336
Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu Pro Lys
              80              85              90

gat gtc aaa gca ccc gaa ggc ttt gca tgg ggc gtg gcc tcc gtt gct      384
Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser Val Ala
          95              100              105

cag cag tac gag ggt gcc gtc aaa gca gat gga cga gga cca tct cat      432
Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro Ser His
          110              115              120              125

tgg gat ttc ctt tgc cat aga aac cct tcc agc tgc aca aac tac acc      480
Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn Tyr Thr
              130              135              140

agt gat atc act gat ctt ggc cgt tac tac tat aag aat gac ttg gca      528
Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp Leu Ala
              145              150              155

cga atg gct gcc atg ggt ata act cac tac tca ttt tca gtg agc tgg      576
Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val Ser Trp
              160              165              170

acc aga gtt gtg cca ttc ggc aag aag ggc agt cca gtc agc aac gaa      624
Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser Asn Glu
          175              180              185

```

ggc ctc gac tat tac gaa gat atc tgc aag acg gct ttg agc ttt gga 672
 Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser Phe Gly
 190 195 200 205

atc aag cct gtc att act tta ttc cac tgg gat act cct gcc aac tta 720
 Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala Asn Leu
 210 215 220

ctc ttc gaa tat gga ggt ttc ctc aac ggg aca atc gta gat gac tac 768
 Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp Asp Tyr
 225 230 235

tac tat tat gca gat ata gta ttc aga aga cta ggc aaa tat gcc gaa 816
 Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr Ala Glu
 240 245 250

acc ttc ttc acc ttc aat gag cct cgt gta tac tgc agc gag tat act 864
 Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr
 255 260 265

ggt cct ccg ttc gat gcc tat tac gaa cgc tat ggg ctg aat tct agc 912
 Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn Ser Ser
 270 275 280 285

act gca cca tat ccc tgc tcg tac aat ctg ctc cga gcc cat ggt gcc 960
 Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His Gly Ala
 290 295 300

gca gtc ggc agg tat cgt gcc ctg gtc aaa gaa ggc agc atc aaa tct 1008
 Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile Lys Ser
 305 310 315

ggt gaa atc gca ttc aaa aac gac gat agt tac cag cta cct caa aat 1056
 Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro Gln Asn
 320 325 330

cca gac tct gac gcc gac aag cga gct gca aaa cgc cac ttc gac ttc 1104
 Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe Asp Phe
 335 340 345

tac atc ggc ata ttc tcg caa ccc gta tac ggc aac ggc tac tat ccc 1152
 Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr Tyr Pro
 350 355 360 365

gaa acc gtc cgg aac aca att tca gag cgc ttc ctc cca gag ttc acc 1200
 Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu Phe Thr
 370 375 380

gca gcc gag cgc gaa caa att cag gga tca gcg gac ttc tac gcc atc 1248
 Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr Ala Ile
 385 390 395

gac ggt tat agg acg aat atc gca agc gct gct ccc aat ggt att gac 1296
 Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly Ile Asp
 400 405 410

gcg tgc ttg aga aac gct agt gat ccc aac tgg cca gtt tgc cag gac 1344
 Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys Gln Asp
 415 420 425

aac agt aac aca ggc caa tac gct act ctt gag gga ttc gca ttg gga 1392
 Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala Leu Gly
 430 435 440 445
 cct cct gca gat ccc aac gcg aat tgg cta tac aac act gca cct tac 1440
 Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala Pro Tyr
 450 455 460
 ctg cgt tac caa ttc aaa gtt ctg aaa gag aac ttc aat tac aag aag 1488
 Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr Lys Lys
 465 470 475
 atc tac ttg acg gaa ttc ggc ttt gca gaa cct ttt agc tat ttg cga 1536
 Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr Leu Arg
 480 485 490
 cag gat ctg tat gca ttg ctg tat gac act gat cgc act gca tat tat 1584
 Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala Tyr Tyr
 495 500 505
 caa gac tat cta gcg cag tgc atg ctg gct atc aaa gaa gat ggc atc 1632
 Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp Gly Ile
 510 515 520 525
 cct ctt gct ggt gtt ttc gca tgg tca ttt gtt gat aac ttc gaa tgg 1680
 Pro Leu Ala Gly Val Phe Ala Trp Ser Phe Val Asp Asn Phe Glu Trp
 530 535 540
 ggt tcc ggt ctt gag cag aga ttt gga atg caa tat gtc aac tac acc 1728
 Gly Ser Gly Leu Glu Gln Arg Phe Gly Met Gln Tyr Val Asn Tyr Thr
 545 550 555
 gat ccg gat ctc cca cga acc ttc aag ctt tct ttc ctg gca tat cgt 1776
 Asp Pro Asp Leu Pro Arg Thr Phe Lys Leu Ser Phe Leu Ala Tyr Arg
 560 565 570
 gat ttc atc aaa aac cac aaa aag tga 1803
 Asp Phe Ile Lys Asn His Lys Lys
 575 580

<210> 14
 <211> 600
 <212> PRT
 <213> Rodotorula minuta

<400> 14

Met Val Ile Leu Arg His Thr Val Leu Ala Ala Val Val Gln Ile
 -15 -10 -5

Ala Leu Gly Ala Pro Gln Phe Ser Pro Thr Val Asn Thr Glu Gly Ile
 -1 1 5 10

Ile Pro Ile Ser Glu Arg Pro Ser Pro Thr Gln Ser Ser Ser Ser
 15 20 25

Gly Pro Val Ile Pro Ala Gly Ser Tyr Val Ser Asp Phe Asp Ala Ser
 39

30	35	40	45
Gly Leu Ala Asn Leu Trp Ser Gln Val Glu Val Asp Ile Pro Val Glu	50	55	60
Ser Pro Arg Ile Ser Ala Val Pro Pro Leu Asn Lys Thr Phe Ser Val	65	70	75
Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu Pro Lys	80	85	90
Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser Val Ala	95	100	105
Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro Ser His	110	115	120
Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn Tyr Thr	130	135	140
Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp Leu Ala	145	150	155
Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val Ser Trp	160	165	170
Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser Asn Glu	175	180	185
Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser Phe Gly	190	195	200
Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala Asn Leu	210	215	220
Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp Asp Tyr	225	230	235
Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr Ala Glu	240	245	250
Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr	255	260	265
Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn Ser Ser	270	275	280
		40	285

Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His Gly Ala
 290 295 300

Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile Lys Ser
 305 310 315

Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro Gln Asn
 320 325 330

Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe Asp Phe
 335 340 345

Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr Tyr Pro
 350 355 360 365

Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu Phe Thr
 370 375 380

Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr Ala Ile
 385 390 395

Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly Ile Asp
 400 405 410

Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys Gln Asp
 415 420 425

Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala Leu Gly
 430 435 440 445

Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala Pro Tyr
 450 455 460

Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr Lys Lys
 465 470 475

Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr Leu Arg
 480 485 490

Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala Tyr Tyr
 495 500 505

Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp Gly Ile
 510 515 520 525

gtc gag tct cca agg att tct gcc gta cct cct ctc aac aaa acg ttc 288
Val Glu Ser Pro Arg Ile Ser Ala Val Pro Leu Asn Lys Thr Phe
42

60	65	70	75	
agt gtt ccg aaa act cca gtc ctg ccc aac tcc ctg cag gat cac ctt				336
Ser Val Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu	80	85	90	
cca aaa gat gtc aaa gca ccc gaa ggc ttt gca tgg ggc gtg gcc tcc				384
Pro Lys Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser	95	100	105	
gtt gct cag cag tac gag ggt gcc gtc aaa gca gat gga cga gga cca				432
Val Ala Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro	110	115	120	
tct cat tgg gat ttc ctt tgc cat aga aac cct tcc agc tgc aca aac				480
Ser His Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn	125	130	135	
tac acc agt gat atc act gat ctt ggc cgt tac tac tat aag aat gac				528
Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp	140	145	150	155
ttg gca cga atg gct gcc atg ggt ata act cac tac tca ttt tca gtg				576
Leu Ala Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val	160	165	170	
agc tgg acc aga gtt gtg cca ttc ggc aag aag ggc agt cca gtc agc				624
Ser Trp Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser	175	180	185	
aac gaa ggc ctc gac tat tac gaa gat atc tgc aag acg gct ttg agc				672
Asn Glu Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser	190	195	200	
ttt gga atc aag cct gtc att act tta ttc cac tgg gat act cct gcc				720
Phe Gly Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala	205	210	215	
aac tta ctc ttc gaa tat gga ggt ttc ctc aac ggg aca atc gta gat				768
Asn Leu Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp	220	225	230	235
gac tac tac tat tat gca gat ata gta ttc aga aga cta ggc aaa tat				816
Asp Tyr Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr	240	245	250	
gcc gaa acc ttc ttc acc ttc aat gag cct cgt gta tac tgc agc gag				864
Ala Glu Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu	255	260	265	
tat act ggt cct ccg ttc gat gcc tat tac gaa cgc tat ggg ctg aat				912
Tyr Thr Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn	270	275	280	
tct agc act gca cca tat ccc tgc tcg tac aat ctg ctc cga gcc cat				960
Ser Ser Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His	285	290	295	
ggt gcc gca gtc ggc agg tat cgt gcc ctg gtc aaa gaa ggc agc atc				1008
Gly Ala Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile	300	305	310	315

aaa tct ggt gaa atc gca ttc aaa aac gac gat agt tac cag cta cct Lys Ser Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro 320 325 330	1056
caa aat cca gac tct gac gcc gac aag cga gct gca aaa cgc cac ttc Gln Asn Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe 335 340 345	1104
gac ttc tac atc ggc ata ttc tcg caa ccc gta tac ggc aac ggc tac Asp Phe Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr 350 355 360	1152
tat ccc gaa acc gtc cgg aac aca att tca gag cgc ttc ctc cca gag Tyr Pro Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu 365 370 375	1200
ttc acc gca gcc gag cgc gaa caa att cag gga tca gcg gac ttc tac Phe Thr Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr 380 385 390 395	1248
gcc atc gac ggt tat agg acg aat atc gca agc gct gct ccc aat ggt Ala Ile Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly 400 405 410	1296
att gac gcg tgc ttg aga aac gct agt gat ccc aac tgg cca gtt tgc Ile Asp Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys 415 420 425	1344
cag gac aac agt aac aca ggc caa tac gct act ctt gag gga ttc gca Gln Asp Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala 430 435 440	1392
ttg gga cct cct gca gat ccc aac gcg aat tgg cta tac aac act gca Leu Gly Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala 445 450 455	1440
cct tac ctg cgt tac caa ttc aaa gtt ctg aaa gag aac ttc aat tac Pro Tyr Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr 460 465 470 475	1488
aag aag atc tac ttg acg gaa ttc ggc ttt gca gaa cct ttt agc tat Lys Lys Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr 480 485 490	1536
ttg cga cag gat ctg tat gca ttg ctg tat gac act gat cgc act gca Leu Arg Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala 495 500 505	1584
tat tat caa gac tat cta gcg cag tgc atg ctg gct atc aaa gaa gat Tyr Tyr Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp 510 515 520	1632
ggc atc cct ctt gct ggt gtt ttc gca tgg tca ttt gtt gat aac ttc Gly Ile Pro Leu Ala Gly Val Phe Ala Trp Ser Phe Val Asp Asn Phe 525 530 535	1680
gaa tgg ggt tcc ggt ctt gag cag aga ttt gga atg caa tat gtc aac Glu Trp Gly Ser Gly Leu Glu Gln Arg Phe Gly Met Gln Tyr Val Asn 540 545 550 555	1728

tac acc gat ccg gat ctc cca cga acc ttc aag ctt tct ttc ctg gca 1776
 Tyr Thr Asp Pro Asp Leu Pro Arg Thr Phe Lys Leu Ser Phe Leu Ala
 560 565 570

tat cgt gat ttc atc aaa aac cac aaa aag tga 1809
 Tyr Arg Asp Phe Ile Lys Asn His Lys Lys
 575 580

<210> 16
 <211> 602
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 16

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ala Pro Gln Phe Ser Pro Thr Val Asn Thr Glu
 -5 -1 1 5 10

Gly Ile Ile Pro Ile Ser Glu Arg Pro Ser Pro Thr Gln Ser Ser Ser
 15 20 25

Ser Ser Gly Pro Val Ile Pro Ala Gly Ser Tyr Val Ser Asp Phe Asp
 30 35 40

Ala Ser Gly Leu Ala Asn Leu Trp Ser Gln Val Glu Val Asp Ile Pro
 45 50 55

Val Glu Ser Pro Arg Ile Ser Ala Val Pro Pro Leu Asn Lys Thr Phe
 60 65 70 75

Ser Val Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu
 80 85 90

Pro Lys Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser
 95 100 105

Val Ala Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro
 110 115 120

Ser His Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn
 125 130 135

Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp
 140 145 150 155
 45

Leu Ala Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val
 160 165 170

Ser Trp Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser
 175 180 185

Asn Glu Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser
 190 195 200

Phe Gly Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala
 205 210 215

Asn Leu Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp
 220 225 230 235

Asp Tyr Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr
 240 245 250

Ala Glu Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu
 255 260 265

Tyr Thr Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn
 270 275 280

Ser Ser Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His
 285 290 295

Gly Ala Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile
 300 305 310 315

Lys Ser Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro
 320 325 330

Gln Asn Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe
 335 340 345

Asp Phe Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr
 350 355 360

Tyr Pro Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu
 365 370 375

Phe Thr Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr
 380 385 390 395

Ala Ile Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly
 400 405 410
 Ile Asp Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys
 415 420 425
 Gln Asp Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala
 430 435 440
 Leu Gly Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala
 445 450 455
 Pro Tyr Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr
 460 465 470 475
 Lys Lys Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr
 480 485 490
 Leu Arg Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala
 495 500 505
 Tyr Tyr Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp
 510 515 520
 Gly Ile Pro Leu Ala Gly Val Phe Ala Trp Ser Phe Val Asp Asn Phe
 525 530 535
 Glu Trp Gly Ser Gly Leu Glu Gln Arg Phe Gly Met Gln Tyr Val Asn
 540 545 550 555
 Tyr Thr Asp Pro Asp Leu Pro Arg Thr Phe Lys Leu Ser Phe Leu Ala
 560 565 570
 Tyr Arg Asp Phe Ile Lys Asn His Lys Lys
 575 580

<210> 17
 <211> 1809
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(1809)

```

<220>
<221> sig_peptit
<222> (1)..(63)

<220>
<221> mat_peptit
<222> (64)..(1806)

<400> 17
atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg      48
Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
      -20                      -15                      -10

gca cct gct ttg gct gcg cca cag ttc tcg cca aca gtc aac acc gaa      96
Ala Pro Ala Leu Ala Ala Pro Gln Phe Ser Pro Thr Val Asn Thr Glu
      -5                      -1 1                      5                      10

ggc atc atc ccc atc agc gag cgg cca agc ccg aca cag agc tcc tct      144
Gly Ile Ile Pro Ile Ser Glu Arg Pro Ser Pro Thr Gln Ser Ser Ser
                      15                      20                      25

tct tct ggg ccc gtg atc cca gcg ggt tcc tat gtc agc gac ttc gat      192
Ser Ser Gly Pro Val Ile Pro Ala Gly Ser Tyr Val Ser Asp Phe Asp
                      30                      35                      40

gcc agc ggg ttg gcg aac ctt tgg tcg cag gtt gaa gtg gac att ccg      240
Ala Ser Gly Leu Ala Asn Leu Trp Ser Gln Val Glu Val Asp Ile Pro
      45                      50                      55

gtt gag tct cca cgc att tcg gcc gtt cct ccg ctg aac aag acc ttc      288
Val Glu Ser Pro Arg Ile Ser Ala Val Pro Pro Leu Asn Lys Thr Phe
      60                      65                      70                      75

tcg gtt ccg aag act cca gtt ctc ccc aat tcg ctt caa gac cat ctg      336
Ser Val Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu
                      80                      85                      90

ccc aag gat gtg aaa gct cct gag ggc ttt gcc tgg gga gtg gcg tcc      384
Pro Lys Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser
                      95                      100                      105

gtg gca cag caa tac gag ggc gca gtc aaa gca gac gga cga ggc cct      432
Val Ala Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro
                      110                      115                      120

tct cac tgg gac ttc ctg tgc cat cgc aac ccc tcc tct tgt acc aac      480
Ser His Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn
                      125                      130                      135

tac acc tct gac atc act gat ctg ggc cgc tac tac tac aag aac gat      528
Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp
      140                      145                      150                      155

ctc gcc cgt atg gcc gca atg ggg atc act cac tac tcc ttt tcg gtg      576
Leu Ala Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val
                      160                      165                      170

tct tgg acc cga gtt gtc ccc ttc ggc aag aag ggt tcg cct gtt tcc      624
Ser Trp Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser
                      175                      180                      185

```

aac gag gga ttg gac tac tac gag gac atc tgc aag act gcc ctc tct Asn Glu Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser 190 195 200	672
ttc ggc atc aag ccc gtc atc acg ctg ttc cac tgg gat acc ccg gca Phe Gly Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala 205 210 215	720
aac ctc ctg ttt gag tat gga ggt ttc ctc aat ggc acg atc gtc gat Asn Leu Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp 220 225 230 235	768
gac tac tac tac tac gcc gac att gtc ttc cgg cgg ttg ggc aag tac Asp Tyr Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr 240 245 250	816
gcg gag acg ttc ttc acg ttc aac gag cca cgc gtc tac tgc agc gag Ala Glu Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu 255 260 265	864
tac act ggt cct ccc ttc gat gcc tac tac gag cga tac ggc ctc aac Tyr Thr Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn 270 275 280	912
tcc tcg act gcc cca tac cct tgt tcg tac aac ctt ttg cgc gct cat Ser Ser Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His 285 290 295	960
ggc gcc gct gtc ggg cga tac cgc gcc ctt gtc aag gag ggc tcc atc Gly Ala Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile 300 305 310 315	1008
aag tcc ggt gag att gcg ttc aag aac gat gac agc tat cag ctt cct Lys Ser Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro 320 325 330	1056
cag aat cct gat tcc gat gca gac aaa cgg gca gca aaa cgt cac ttt Gln Asn Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe 335 340 345	1104
gac ttc tac atc ggg atc ttc tcc caa ccg gtt tac gga aat ggg tac Asp Phe Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr 350 355 360	1152
tac ccg gaa aca gtg cgt aac aca atc tcc gaa cgc ttc ctg cct gag Tyr Pro Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu 365 370 375	1200
ttc acc gct gct gaa cgc gag cag att cag ggt agc gcc gac ttc tac Phe Thr Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr 380 385 390 395	1248
gcg atc gat gga tat cgg acc aac atc gca tct gca gcc ccg aat ggg Ala Ile Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly 400 405 410	1296
att gac gct tgc ttg cgt aac gct agc gac cct aac tgg ccg gtg tgc Ile Asp Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys 415 420 425	1344

caa gac aac tcg aac aca ggc cag tat gcc acc ctg gaa ggt ttt gca 1392
 Gln Asp Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala
 430 435 440
 ttg ggg cct ccc gca gac cca aat gcc aac tgg ctg tac aac acg gcg 1440
 Leu Gly Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala
 445 450 455
 ccg tat ctc cga tac caa ttc aag gtc ttg aag gag aac ttc aac tac 1488
 Pro Tyr Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr
 460 465 470 475
 aag aag atc tac ctg acc gaa ttc ggg ttt gct gag cct ttc tcc tat 1536
 Lys Lys Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr
 480 485 490
 ctg cgc caa gac ctg tat gcg ctt ctc tac gat act gat cga acc gct 1584
 Leu Arg Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala
 495 500 505
 tac tac cag gat tac ctg gcg cag tgc atg ctg gcc atc aag gag gac 1632
 Tyr Tyr Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp
 510 515 520
 ggt atc cca ttg gcc ggc gtc ttt gct tgg tcc ttc gtt gac aac ttc 1680
 Gly Ile Pro Leu Ala Gly Val Phe Ala Trp Ser Phe Val Asp Asn Phe
 525 530 535
 gag tgg ggt tcg gga ttg gaa cag cgc ttt ggc atg cag tat gtg aac 1728
 Glu Trp Gly Ser Gly Leu Glu Gln Arg Phe Gly Met Gln Tyr Val Asn
 540 545 550 555
 tac acc gac ccc gac ttg ccc cgc aca ttc aag ctc tcc ttt ctc gcc 1776
 Tyr Thr Asp Pro Asp Leu Pro Arg Thr Phe Lys Leu Ser Phe Leu Ala
 560 565 570
 tat cgg gac ttc atc aag aac cac aag aag taa 1809
 Tyr Arg Asp Phe Ile Lys Asn His Lys Lys
 575 580

<210> 18

<211> 602

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 18

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ala Pro Gln Phe Ser Pro Thr Val Asn Thr Glu
 -5 -1 1 5 10

Gly Ile Ile Pro Ile Ser Glu Arg Pro Ser Pro Thr Gln Ser Ser Ser
 15 20 25

50

Ser Ser Gly Pro Val Ile Pro Ala Gly Ser Tyr Val Ser Asp Phe Asp
 30 35 40

Ala Ser Gly Leu Ala Asn Leu Trp Ser Gln Val Glu Val Asp Ile Pro
 45 50 55

Val Glu Ser Pro Arg Ile Ser Ala Val Pro Pro Leu Asn Lys Thr Phe
 60 65 70 75

Ser Val Pro Lys Thr Pro Val Leu Pro Asn Ser Leu Gln Asp His Leu
 80 85 90

Pro Lys Asp Val Lys Ala Pro Glu Gly Phe Ala Trp Gly Val Ala Ser
 95 100 105

Val Ala Gln Gln Tyr Glu Gly Ala Val Lys Ala Asp Gly Arg Gly Pro
 110 115 120

Ser His Trp Asp Phe Leu Cys His Arg Asn Pro Ser Ser Cys Thr Asn
 125 130 135

Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg Tyr Tyr Tyr Lys Asn Asp
 140 145 150 155

Leu Ala Arg Met Ala Ala Met Gly Ile Thr His Tyr Ser Phe Ser Val
 160 165 170

Ser Trp Thr Arg Val Val Pro Phe Gly Lys Lys Gly Ser Pro Val Ser
 175 180 185

Asn Glu Gly Leu Asp Tyr Tyr Glu Asp Ile Cys Lys Thr Ala Leu Ser
 190 195 200

Phe Gly Ile Lys Pro Val Ile Thr Leu Phe His Trp Asp Thr Pro Ala
 205 210 215

Asn Leu Leu Phe Glu Tyr Gly Gly Phe Leu Asn Gly Thr Ile Val Asp
 220 225 230 235

Asp Tyr Tyr Tyr Tyr Ala Asp Ile Val Phe Arg Arg Leu Gly Lys Tyr
 240 245 250

Ala Glu Thr Phe Phe Thr Phe Asn Glu Pro Arg Val Tyr Cys Ser Glu
 255 260 265

Tyr Thr Gly Pro Pro Phe Asp Ala Tyr Tyr Glu Arg Tyr Gly Leu Asn
 270 275 280

Ser Ser Thr Ala Pro Tyr Pro Cys Ser Tyr Asn Leu Leu Arg Ala His
 285 290 295

Gly Ala Ala Val Gly Arg Tyr Arg Ala Leu Val Lys Glu Gly Ser Ile
 300 305 310 315

Lys Ser Gly Glu Ile Ala Phe Lys Asn Asp Asp Ser Tyr Gln Leu Pro
 320 325 330

Gln Asn Pro Asp Ser Asp Ala Asp Lys Arg Ala Ala Lys Arg His Phe
 335 340 345

Asp Phe Tyr Ile Gly Ile Phe Ser Gln Pro Val Tyr Gly Asn Gly Tyr
 350 355 360

Tyr Pro Glu Thr Val Arg Asn Thr Ile Ser Glu Arg Phe Leu Pro Glu
 365 370 375

Phe Thr Ala Ala Glu Arg Glu Gln Ile Gln Gly Ser Ala Asp Phe Tyr
 380 385 390 395

Ala Ile Asp Gly Tyr Arg Thr Asn Ile Ala Ser Ala Ala Pro Asn Gly
 400 405 410

Ile Asp Ala Cys Leu Arg Asn Ala Ser Asp Pro Asn Trp Pro Val Cys
 415 420 425

Gln Asp Asn Ser Asn Thr Gly Gln Tyr Ala Thr Leu Glu Gly Phe Ala
 430 435 440

Leu Gly Pro Pro Ala Asp Pro Asn Ala Asn Trp Leu Tyr Asn Thr Ala
 445 450 455

Pro Tyr Leu Arg Tyr Gln Phe Lys Val Leu Lys Glu Asn Phe Asn Tyr
 460 465 470 475

Lys Lys Ile Tyr Leu Thr Glu Phe Gly Phe Ala Glu Pro Phe Ser Tyr
 480 485 490

Leu Arg Gln Asp Leu Tyr Ala Leu Leu Tyr Asp Thr Asp Arg Thr Ala
 495 500 505

Tyr Tyr Gln Asp Tyr Leu Ala Gln Cys Met Leu Ala Ile Lys Glu Asp
510 515 520

Gly Ile Pro Leu Ala Gly Val Phe Ala Trp Ser Phe Val Asp Asn Phe
525 530 535

Glu Trp Gly Ser Gly Leu Glu Gln Arg Phe Gly Met Gln Tyr Val Asn
540 545 550 555

Tyr Thr Asp Pro Asp Leu Pro Arg Thr Phe Lys Leu Ser Phe Leu Ala
560 565 570

Tyr Arg Asp Phe Ile Lys Asn His Lys Lys
575 580

<210> 19
<211> 1785
<212> ADN
<213> Sterigmatomyces elviae

<220>
<221> CDS
<222> (1)..(1785)

<220>
<221> sig_peptit
<222> (1)..(57)

<220>
<221> mat_peptit
<222> (58)..(1782)

<400> 19
atg ctt gtc gga ctt gct ttg act gct ctg tta ggt gcc act cgt tat 48
Met Leu Val Gly Leu Ala Leu Thr Ala Leu Leu Gly Ala Thr Arg Tyr
-15 -10 -5

gtt ggc gca atc cct gct ttc cca atc act cca gat ttg gct ggt ggg 96
Val Gly Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala Gly Gly
-1 1 5 10

ctg gag tct gtg acc aac act cag acc tcg ctc cct tca gcg agc gct 144
Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala Ser Ala
15 20 25

gtg tcg tcg ccc tat aat caa gat gca ctc gac aag ctg tgg gct gag 192
Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp Ala Glu
30 35 40 45

gtc gaa aaa gac att cca gtc gag aca cca agc atc tcc agc gtt gtt 240
Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser Val Val
50 55 60

cca gta aac aac agc ttt gcg gtc ccc aaa acc cct act ctg ccc cga 288
Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu Pro Arg
53

65	70	75	
tct ctt cag gat cat gct acc agt ggc cgc aaa ttc ccc aaa ggc ttc			336
Ser Leu Gln Asp His Ala Thr	Ser Gly Arg Lys Phe	Pro Lys Gly Phe	
80	85	90	
aag ttt ggt gtc gcc acc gcc gat cag cag tat gaa ggt gcc gtc aag			384
Lys Phe Gly Val Ala Thr Ala	Asp Gln Gln Tyr Glu	Gly Ala Val Lys	
95	100	105	
gct gat ggc cgt ggc ccc tct cac tgg gat tac ctt tgc cat cgt ctc			432
Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His Arg Leu			
110	115	120	125
cca cag caa tgc aac aac tac acc tca gac atc act gac ctt ggt cgc			480
Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg			
130	135	140	
tac tat tat aag caa gat atc gca cga atc aag gcc atg gga gta aac			528
Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly Val Asn			
145	150	155	
act gta tca ctc acc ctt tct tgg tca cgt atc aag ccc ttc ggc acg			576
Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe Gly Thr			
160	165	170	
gcc gat agt cct gtc agc aaa gaa gga ctc caa ttt tac gat gac ttt			624
Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp Asp Phe			
175	180	185	
atc aac gag ctc atc gat aat ggc atc gaa cca gtc gtc acc ctg ttc			672
Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr Leu Phe			
190	195	200	205
cat tgg agt aca cca ctc aat ctg gtg ttc gaa tac ggg gcc ttc ctt			720
His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala Phe Leu			
210	215	220	
aat ggc agc tcg gtt gaa gat ttc gct agc tat gct aag ctt gtt ttt			768
Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu Val Phe			
225	230	235	
gag cat ttc ggt gac aga gta acc aca ttc ctt act ttc aac gag cct			816
Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn Glu Pro			
240	245	250	
cgt gta tac tgc tcc gaa tac act ggc gag cct ttt aat gat tat tgg			864
Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp Tyr Trp			
255	260	265	
cac ttt gga ggt ccc aac atc aat gcc acg acc gct ccc tat cct tgc			912
His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr Pro Cys			
270	275	280	285
acc tat aac att tta aaa gca cac gga cgt gcc gtt caa gaa tac aga			960
Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu Tyr Arg			
290	295	300	
gca ctg gtc aac agt gga aag atc aag aaa ggt gaa gtc gca att aaa			1008
Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala Ile Lys			
305	310	315	

aac gac gat agt tat ccc gtg cca gtc aac cca gac tcc gaa gcc gat Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu Ala Asp 320 325 330	1056
gta gaa gcg gcc aag cga cac ttc gat ttc tac att ggc att ttc agt Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile Phe Ser 335 340 345	1104
cag cct att tac ggt gat gga aag ttc cca gac acc gtt aga aac acc Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg Asn Thr 350 355 360 365	1152
atc tcc act gaa ttc ctg cca tac ctc acc gat gat gag aaa gcc atg Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys Ala Met 370 375 380	1200
att aaa gga agt ggt gac ttt ttc gcc ata gac gca tat cga acc aac Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg Thr Asn 385 390 395	1248
ctt gca aga gct gcg ccc aat ggc atc caa gcc tgc gta gct aac atc Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala Asn Ile 400 405 410	1296
tcg gat ccc aat tgg cct gta tgc caa gac aac agc cct gaa gga caa Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu Gly Gln 415 420 425	1344
tac caa acc atg gat ggc ttt gct ttc ggt ccc ccg gca gac ccc aac Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp Pro Asn 430 435 440 445	1392
gct gca tgg cta tat gac acc agc ttc aag ttg cgc tac cag ctt aag Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln Leu Lys 450 455 460	1440
aca ctc aaa gag gca ttc aac tat gac aag atc tac atc tca gag ttt Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser Glu Phe 465 470 475	1488
gga ttt gct cgg cct tac gaa tac ctc tac cct tac ggc ttc gac gtc Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe Asp Val 480 485 490	1536
ctg tac gac aca gac cgt gcc att tac tac caa gac tac atg gct gag Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met Ala Glu 495 500 505	1584
gcc ttg gat gcc att cat gac gac ggc att cct ctg gct ggt gtc ttt Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly Val Phe 510 515 520 525	1632
gct tgg tcc ttc gtt gac aat ttc gaa tgg gct tcc ggt ctt gaa cag Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu Glu Gln 530 535 540	1680
cga ttc ggc atg cag ttc gtg aac tac acg aca ctg gaa aga gag tac Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg Glu Tyr 545 550 555	1728

gaa gat taa 1785
Glu Asp
575

<210> 20
<211> 594
<212> PRT
<213> Sterigmatomyces elviae

<400> 20

Met Leu Val Gly Leu Ala Leu Thr Ala Leu Leu Gly Ala Thr Arg Tyr
-15 -10 -5

Val Gly Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala Gly Gly
-1 1 5 10

Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala Ser Ala
15 20 25

Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp Ala Glu
30 35 40 45

Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser Val Val
50 55 60

Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu Pro Arg
65 70 75

Ser Leu Gln Asp His Ala Thr Ser Gly Arg Lys Phe Pro Lys Gly Phe
80 85 90

Lys Phe Gly Val Ala Thr Ala Asp Gln Gln Tyr Glu Gly Ala Val Lys
95 100 105

Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His Arg Leu
110 115 120 125

Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu Gly Arg
130 135 140

Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly Val Asn
145 150 155

Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe Gly Thr
56

160	165	170
Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp Asp Phe		
175	180	185
Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr Leu Phe		
190	195	200
His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala Phe Leu		
	210	215
Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu Val Phe		
	225	230
Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn Glu Pro		
	240	245
Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp Tyr Trp		
	255	260
His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr Pro Cys		
270	275	280
Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu Tyr Arg		
	290	295
Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala Ile Lys		
	305	310
Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu Ala Asp		
	320	325
Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile Phe Ser		
	335	340
Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg Asn Thr		
350	355	360
Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys Ala Met		
	370	375
Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg Thr Asn		
	385	390
Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala Asn Ile		
	400	405

Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu Gly Gln
 415 420 425

Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp Pro Asn
 430 435 440 445

Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln Leu Lys
 450 455 460

Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser Glu Phe
 465 470 475

Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe Asp Val
 480 485 490

Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met Ala Glu
 495 500 505

Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly Val Phe
 510 515 520 525

Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu Glu Gln
 530 535 540

Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg Glu Tyr
 545 550 555

Lys Leu Ser Phe Leu Leu Tyr Arg Asp Phe Ile Glu Asn His Ser Cys
 560 565 570

Glu Asp
 575

<210> 21
 <211> 1791
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(1791)

<220>
 <221> sig_peptit

<222> (1)..(63)

<220>

<221> mat_peptit

<222> (64)..(1788)

<400> 21

atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg	48
Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala	
-20 -15 -10	
gca cct gct ttg gct atc cct gct ttc cca atc act cca gat ttg gct	96
Ala Pro Ala Leu Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala	
-5 -1 1 5 10	
ggg ggg ctg gag tct gtg acc aac act cag acc tcg ctc cct tca gcg	144
Gly Gly Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala	
15 20 25	
agc gct gtg tcg tcg ccc tat aat caa gat gca ctc gac aag ctg tgg	192
Ser Ala Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp	
30 35 40	
gct gag gtc gaa aaa gac att cca gtc gag aca cca agc atc tcc agc	240
Ala Glu Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser	
45 50 55	
ggt ggt cca gta aac aac agc ttt gcg gtc ccc aaa acc cct act ctg	288
Val Val Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu	
60 65 70 75	
ccc cga tct ctt cag gat cat gct acc agt ggc cgc aaa ttc ccc aaa	336
Pro Arg Ser Leu Gln Asp His Ala Thr Ser Gly Arg Lys Phe Pro Lys	
80 85 90	
ggc ttc aag ttt ggt gtc gcc acc gcc gat cag cag tat gaa ggt gcc	384
Gly Phe Lys Phe Gly Val Ala Thr Ala Asp Gln Gln Tyr Glu Gly Ala	
95 100 105	
gtc aag gct gat ggc cgt ggc ccc tct cac tgg gat tac ctt tgc cat	432
Val Lys Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His	
110 115 120	
cgt ctc cca cag caa tgc aac aac tac acc tca gac atc act gac ctt	480
Arg Leu Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu	
125 130 135	
ggt cgc tac tat tat aag caa gat atc gca cga atc aag gcc atg gga	528
Gly Arg Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly	
140 145 150 155	
gta aac act gta tca ctc acc ctt tct tgg tca cgt atc aag ccc ttc	576
Val Asn Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe	
160 165 170	
ggc acg gcc gat agt cct gtc agc aaa gaa gga ctc caa ttt tac gat	624
Gly Thr Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp	
175 180 185	
gac ttt atc aac gag ctc atc gat aat ggc atc gaa cca gtc gtc acc	672
Asp Phe Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr	
59	

190	195	200	
ctg ttc cat tgg agt aca cca ctc aat ctg gtg ttc gaa tac ggg gcc Leu Phe His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala 205 210 215			720
ttc ctt aat ggc agc tcg gtt gaa gat ttc gct agc tat gct aag ctt Phe Leu Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu 220 225 230 235			768
gtt ttt gag cat ttc ggt gac aga gta acc aca ttc ctt act ttc aac Val Phe Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn 240 245 250			816
gag cct cgt gta tac tgc tcc gaa tac act ggc gag cct ttt aat gat Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp 255 260 265			864
tat tgg cac ttt gga ggt ccc aac atc aat gcc acg acc gct ccc tat Tyr Trp His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr 270 275 280			912
cct tgc acc tat aac att tta aaa gca cac gga cgt gcc gtt caa gaa Pro Cys Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu 285 290 295			960
tac aga gca ctg gtc aac agt gga aag atc aag aaa ggt gaa gtc gca Tyr Arg Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala 300 305 310 315			1008
att aaa aac gac gat agt tat ccc gtg cca gtc aac cca gac tcc gaa Ile Lys Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu 320 325 330			1056
gcc gat gta gaa gcg gcc aag cga cac ttc gat ttc tac att ggc att Ala Asp Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile 335 340 345			1104
ttc agt cag cct att tac ggt gat gga aag ttc cca gac acc gtt aga Phe Ser Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg 350 355 360			1152
aac acc atc tcc act gaa ttc ctg cca tac ctc acc gat gat gag aaa Asn Thr Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys 365 370 375			1200
gcc atg att aaa gga agt ggt gac ttt ttc gcc ata gac gca tat cga Ala Met Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg 380 385 390 395			1248
acc aac ctt gca aga gct gcg ccc aat ggc atc caa gcc tgc gta gct Thr Asn Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala 400 405 410			1296
aac atc tcg gat ccc aat tgg cct gta tgc caa gac aac agc cct gaa Asn Ile Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu 415 420 425			1344
gga caa tac caa acc atg gat ggc ttt gct ttc ggt ccc ccg gca gac Gly Gln Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp 430 435 440			1392

ccc aac gct gca tgg cta tat gac acc agc ttc aag ttg cgc tac cag 1440
 Pro Asn Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln
 445 450 455
 ctt aag aca ctc aaa gag gca ttc aac tat gac aag atc tac atc tca 1488
 Leu Lys Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser
 460 465 470 475
 gag ttt gga ttt gct cgg cct tac gaa tac ctc tac cct tac ggc ttc 1536
 Glu Phe Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe
 480 485 490
 gac gtc ctg tac gac aca gac cgt gcc att tac tac caa gac tac atg 1584
 Asp Val Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met
 495 500 505
 gct gag gcc ttg gat gcc att cat gac gac ggc att cct ctg gct ggt 1632
 Ala Glu Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly
 510 515 520
 gtc ttt gct tgg tcc ttc gtt gac aat ttc gaa tgg gct tcc ggt ctt 1680
 Val Phe Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu
 525 530 535
 gaa cag cga ttc ggc atg cag ttc gtg aac tac acg aca ctg gaa aga 1728
 Glu Gln Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg
 540 545 550 555
 gag tac aag ctc tcc ttc ctg ctt tat cgt gac ttc att gaa aac cac 1776
 Glu Tyr Lys Leu Ser Phe Leu Leu Tyr Arg Asp Phe Ile Glu Asn His
 560 565 570
 agt tgc gaa gat taa 1791
 Ser Cys Glu Asp
 575

<210> 22
 <211> 596
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 22

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala
 -5 -1 1 5 10

Gly Gly Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala
 15 20 25

Ser Ala Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp
 61

30 35 40
 Ala Glu Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser
 45 50 55
 Val Val Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu
 60 65 70 75
 Pro Arg Ser Leu Gln Asp His Ala Thr Ser Gly Arg Lys Phe Pro Lys
 80 85 90
 Gly Phe Lys Phe Gly Val Ala Thr Ala Asp Gln Gln Tyr Glu Gly Ala
 95 100 105
 Val Lys Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His
 110 115 120
 Arg Leu Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu
 125 130 135
 Gly Arg Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly
 140 145 150 155
 Val Asn Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe
 160 165 170
 Gly Thr Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp
 175 180 185
 Asp Phe Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr
 190 195 200
 Leu Phe His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala
 205 210 215
 Phe Leu Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu
 220 225 230 235
 Val Phe Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn
 240 245 250
 Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp
 255 260 265
 Tyr Trp His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr
 270 275 280
 62

Pro Cys Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu
285 290 295

Tyr Arg Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala
300 305 310 315

Ile Lys Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu
320 325 330

Ala Asp Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile
335 340 345

Phe Ser Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg
350 355 360

Asn Thr Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys
365 370 375

Ala Met Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg
380 385 390 395

Thr Asn Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala
400 405 410

Asn Ile Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu
415 420 425

Gly Gln Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp
430 435 440

Pro Asn Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln
445 450 455

Leu Lys Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser
460 465 470 475

Glu Phe Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe
480 485 490

Asp Val Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met
495 500 505

Ala Glu Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly
510 515 520

Val Phe Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu
 525 530 535

Glu Gln Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg
 540 545 550 555

Glu Tyr Lys Leu Ser Phe Leu Leu Tyr Arg Asp Phe Ile Glu Asn His
 560 565 570

Ser Cys Glu Asp
 575

<210> 23
 <211> 1791
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Beta-galactosidaza

<220>
 <221> CDS
 <222> (1)..(1791)

<220>
 <221> sig_peptit
 <222> (1)..(63)

<220>
 <221> mat_peptit
 <222> (64)..(1788)

<400> 23
 atg atg gtc gcg tgg tgg tct cta ttt ctg tac ggc ctt cag gtc gcg 48
 Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

gca cct gct ttg gct att ccc gct ttc cca atc act ccc gat ttg gcc 96
 Ala Pro Ala Leu Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala
 -5 -1 1 5 10

gga ggc ctg gaa agc gtc aca aac acc caa acg tcc ctc cca tct gct 144
 Gly Gly Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala
 15 20 25

tcc gct gtt tcc agc ccg tac aac caa gac gct ctt gat aag ctg tgg 192
 Ser Ala Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp
 30 35 40

gca gag gtt gag aag gat atc ccc gtt gaa aca cca tcc atc agc tct 240
 Ala Glu Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser
 45 50 55

gtg gtg ccc gtg aac aac tcc ttc gcc gtg cct aag acg ccg acc ctc 288
 Val Val Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu
 64

60	65	70	75	
ccg cgc tcc ctg cag gat cat gct act tcg ggg cgg aag ttc cca aaa				336
Pro Arg Ser Leu Gln Asp His Ala Thr Ser Gly Arg Lys Phe Pro Lys	80	85	90	
ggg ttc aag ttc gga gtt gcc act gcg gac cag cag tac gaa ggc gcc				384
Gly Phe Lys Phe Gly Val Ala Thr Ala Asp Gln Gln Tyr Glu Gly Ala	95	100	105	
gtc aag gcc gac ggc cgc ggg ccg agc cac tgg gat tac ctg tgc cac				432
Val Lys Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His	110	115	120	
cgg ctg cct cag caa tgc aac aac tac acc tct gac atc act gac ttg				480
Arg Leu Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu	125	130	135	
gga cgt tac tac tac aag cag gat atc gcc cgg atc aag gcc atg ggc				528
Gly Arg Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly	140	145	150	155
gtg aat act gtc tcg ctt acc ctg agc tgg tcg cgc atc aag ccc ttt				576
Val Asn Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe	160	165	170	
ggc act gct gac tcg cct gtt tcg aag gag ggt ttg cag ttc tac gat				624
Gly Thr Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp	175	180	185	
gac ttc atc aac gaa ctc att gac aat gga atc gag cct gtt gtg aca				672
Asp Phe Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr	190	195	200	
ctc ttc cat tgg tct acc cct ctg aat ctg gtg ttt gag tac ggt gcc				720
Leu Phe His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala	205	210	215	
ttt ctc aac ggc agc agc gtg gaa gac ttt gcg tcc tat gca aag ctg				768
Phe Leu Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu	220	225	230	235
gtc ttc gaa cac ttt ggg gac cgc gtg aca acc ttc ctt acc ttc aac				816
Val Phe Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn	240	245	250	
gag ccg cgt gtg tac tgc agc gag tat acg ggc gaa ccc ttc aac gac				864
Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp	255	260	265	
tac tgg cat ttc ggt gga ccg aac atc aac gca acg acc gca ccc tat				912
Tyr Trp His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr	270	275	280	
cca tgc acg tac aac atc ctt aag gcg cat ggg cga gct gtt cag gag				960
Pro Cys Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu	285	290	295	
tat ccg gct ctg gtc aac tcg ggg aag atc aag aag ggt gaa gtt gca				1008
Tyr Arg Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala	300	305	310	315
		65		

atc aag aac gat gac tcc tac cct gtg ccc gtt aac cca gac tct gaa Ile Lys Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu 320 325 330	1056
gca gat gtt gaa gct gcg aaa cga cat ttc gac ttc tac atc ggc atc Ala Asp Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile 335 340 345	1104
ttc tcg cag ccc atc tat ggc gat ggc aag ttc cct gat acc gtc cgc Phe Ser Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg 350 355 360	1152
aac aca atc agc aca gaa ttc ctg cct tac ctc act gat gac gag aag Asn Thr Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys 365 370 375	1200
gcc atg atc aag ggt tcc ggc gac ttc ttc gca atc gat gct tac cgc Ala Met Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg 380 385 390 395	1248
acc aac ctt gct cga gct gct cca aac gga atc cag gcc tgt gtt gcc Thr Asn Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala 400 405 410	1296
aac atc tcg gat ccg aat tgg ccg gtc tgt caa gac aac agc cct gaa Asn Ile Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu 415 420 425	1344
ggg cag tat cag acg atg gat ggc ttt gcc ttt ggc cca cca gcg gac Gly Gln Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp 430 435 440	1392
cct aac gcg gcg tgg ctc tac gac acg agc ttc aag ctg cgg tat cag Pro Asn Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln 445 450 455	1440
ctt aag act ctc aaa gag gcc ttc aac tac gac aag atc tac atc agc Leu Lys Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser 460 465 470 475	1488
gaa ttc ggt ttc gcc cga ccc tat gag tat ctg tat ccc tat ggt ttt Glu Phe Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe 480 485 490	1536
gac gtt ctc tat gac aca gac cga gct atc tac tac cag gac tac atg Asp Val Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met 495 500 505	1584
gcc gag gca ctg gat gct atc cac gac gat ggt att cca ttg gcg ggt Ala Glu Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly 510 515 520	1632
gtg ttc gct tgg tcc ttc gtc gac aac ttc gaa tgg gcg tcc ggc ctt Val Phe Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu 525 530 535	1680
gag caa cgc ttt ggc atg cag ttc gtc aac tac aca acc ctc gag cgc Glu Gln Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg 540 545 550 555	1728

gag tac aag ctc tct ttc ctg ctg tat cgt gat ttc atc gag aat cat 1776
 Glu Tyr Lys Leu Ser Phe Leu Leu Tyr Arg Asp Phe Ile Glu Asn His
 560 565 570

agc tgc gag gat taa 1791
 Ser Cys Glu Asp
 575

<210> 24
 <211> 596
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 24

Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
 -20 -15 -10

Ala Pro Ala Leu Ala Ile Pro Ala Phe Pro Ile Thr Pro Asp Leu Ala
 -5 -1 1 5 10

Gly Gly Leu Glu Ser Val Thr Asn Thr Gln Thr Ser Leu Pro Ser Ala
 15 20 25

Ser Ala Val Ser Ser Pro Tyr Asn Gln Asp Ala Leu Asp Lys Leu Trp
 30 35 40

Ala Glu Val Glu Lys Asp Ile Pro Val Glu Thr Pro Ser Ile Ser Ser
 45 50 55

Val Val Pro Val Asn Asn Ser Phe Ala Val Pro Lys Thr Pro Thr Leu
 60 65 70 75

Pro Arg Ser Leu Gln Asp His Ala Thr Ser Gly Arg Lys Phe Pro Lys
 80 85 90

Gly Phe Lys Phe Gly Val Ala Thr Ala Asp Gln Gln Tyr Glu Gly Ala
 95 100 105

Val Lys Ala Asp Gly Arg Gly Pro Ser His Trp Asp Tyr Leu Cys His
 110 115 120

Arg Leu Pro Gln Gln Cys Asn Asn Tyr Thr Ser Asp Ile Thr Asp Leu
 125 130 135

Gly Arg Tyr Tyr Tyr Lys Gln Asp Ile Ala Arg Ile Lys Ala Met Gly
 140 145 150 155
 67

Val Asn Thr Val Ser Leu Thr Leu Ser Trp Ser Arg Ile Lys Pro Phe
 160 165 170

Gly Thr Ala Asp Ser Pro Val Ser Lys Glu Gly Leu Gln Phe Tyr Asp
 175 180 185

Asp Phe Ile Asn Glu Leu Ile Asp Asn Gly Ile Glu Pro Val Val Thr
 190 195 200

Leu Phe His Trp Ser Thr Pro Leu Asn Leu Val Phe Glu Tyr Gly Ala
 205 210 215

Phe Leu Asn Gly Ser Ser Val Glu Asp Phe Ala Ser Tyr Ala Lys Leu
 220 225 230 235

Val Phe Glu His Phe Gly Asp Arg Val Thr Thr Phe Leu Thr Phe Asn
 240 245 250

Glu Pro Arg Val Tyr Cys Ser Glu Tyr Thr Gly Glu Pro Phe Asn Asp
 255 260 265

Tyr Trp His Phe Gly Gly Pro Asn Ile Asn Ala Thr Thr Ala Pro Tyr
 270 275 280

Pro Cys Thr Tyr Asn Ile Leu Lys Ala His Gly Arg Ala Val Gln Glu
 285 290 295

Tyr Arg Ala Leu Val Asn Ser Gly Lys Ile Lys Lys Gly Glu Val Ala
 300 305 310 315

Ile Lys Asn Asp Asp Ser Tyr Pro Val Pro Val Asn Pro Asp Ser Glu
 320 325 330

Ala Asp Val Glu Ala Ala Lys Arg His Phe Asp Phe Tyr Ile Gly Ile
 335 340 345

Phe Ser Gln Pro Ile Tyr Gly Asp Gly Lys Phe Pro Asp Thr Val Arg
 350 355 360

Asn Thr Ile Ser Thr Glu Phe Leu Pro Tyr Leu Thr Asp Asp Glu Lys
 365 370 375

Ala Met Ile Lys Gly Ser Gly Asp Phe Phe Ala Ile Asp Ala Tyr Arg
 380 385 390 395

Thr Asn Leu Ala Arg Ala Ala Pro Asn Gly Ile Gln Ala Cys Val Ala
400 405 410

Asn Ile Ser Asp Pro Asn Trp Pro Val Cys Gln Asp Asn Ser Pro Glu
415 420 425

Gly Gln Tyr Gln Thr Met Asp Gly Phe Ala Phe Gly Pro Pro Ala Asp
430 435 440

Pro Asn Ala Ala Trp Leu Tyr Asp Thr Ser Phe Lys Leu Arg Tyr Gln
445 450 455

Leu Lys Thr Leu Lys Glu Ala Phe Asn Tyr Asp Lys Ile Tyr Ile Ser
460 465 470 475

Glu Phe Gly Phe Ala Arg Pro Tyr Glu Tyr Leu Tyr Pro Tyr Gly Phe
480 485 490

Asp Val Leu Tyr Asp Thr Asp Arg Ala Ile Tyr Tyr Gln Asp Tyr Met
495 500 505

Ala Glu Ala Leu Asp Ala Ile His Asp Asp Gly Ile Pro Leu Ala Gly
510 515 520

Val Phe Ala Trp Ser Phe Val Asp Asn Phe Glu Trp Ala Ser Gly Leu
525 530 535

Glu Gln Arg Phe Gly Met Gln Phe Val Asn Tyr Thr Thr Leu Glu Arg
540 545 550 555

Glu Tyr Lys Leu Ser Phe Leu Leu Tyr Arg Asp Phe Ile Glu Asn His
560 565 570

Ser Cys Glu Asp
575

<210> 25
<211> 30
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi F1

<220>
<221> đặc điểm khác
<222> (21)..(21)

<223> n là a, c, g, hoặc t

<220>

<221> đặc điểm khác

<222> (27)..(27)

<223> n là a, c, g, hoặc t

<220>

<221> đặc điểm khác

<222> (30)..(30)

<223> n là a, c, g, hoặc t

<400> 25

gccggcgcg ctathcargt ngangngcn

30

<210> 26

<211> 30

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mồi F2

<220>

<221> đặc điểm khác

<222> (9)..(9)

<223> n là a, c, g, hoặc t

<220>

<221> đặc điểm khác

<222> (18)..(18)

<223> n là a, c, g, hoặc t

<220>

<221> đặc điểm khác

<222> (30)..(30)

<223> n là a, c, g, hoặc t

<400> 26

gtcaagacnt ggttyacntt yaaygarccn

30

<210> 27

<211> 30

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mồi R1

<220>

<221> đặc điểm khác

<222> (19)..(19)

<223> n là a, c, g, hoặc t

<400> 27

ctcggcccac ccraaytcns wraartadat

30

<210> 28
 <211> 27
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mồi R2

<220>
 <221> đặc điểm khác
 <222> (16)..(16)
 <223> n là a, c, g, hoặc t

<220>
 <221> đặc điểm khác
 <222> (22)..(22)
 <223> n là a, c, g, hoặc t

<400> 28
 ccattcccar ttrtcnacra answcca

27

<210> 29
 <211> 30
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mồi C-R70

<220>
 <221> đặc điểm khác
 <222> (10)..(10)
 <223> n là a, c, g, hoặc t

<400> 29
 gacgaggccn swrttcayt craarttrtc

30

Fig.1

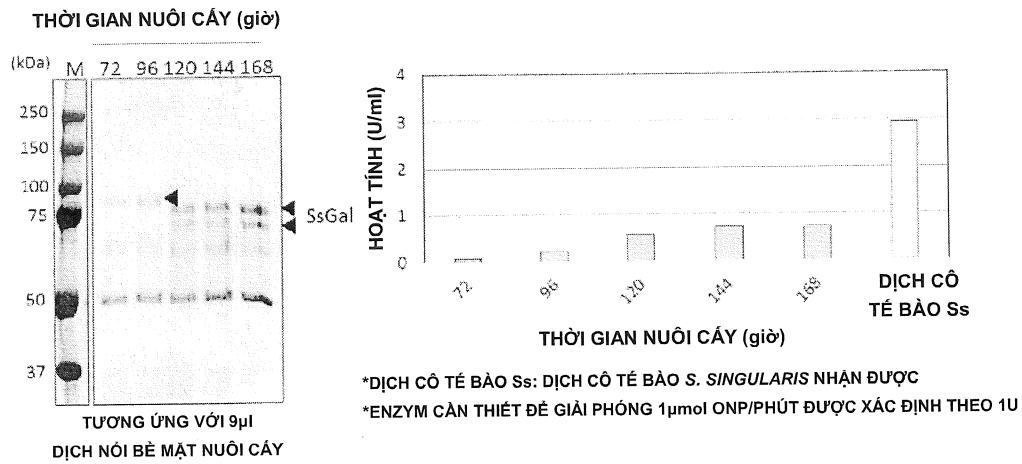


Fig.2

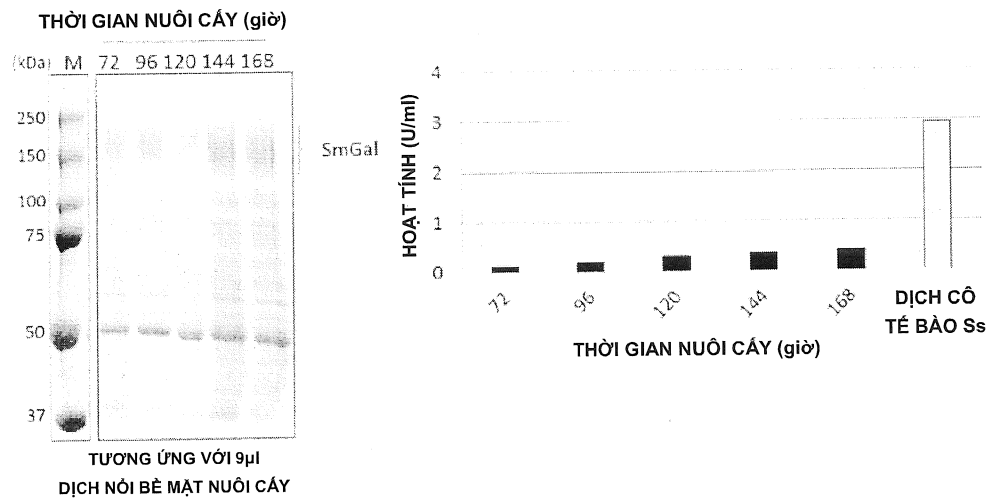


Fig.3

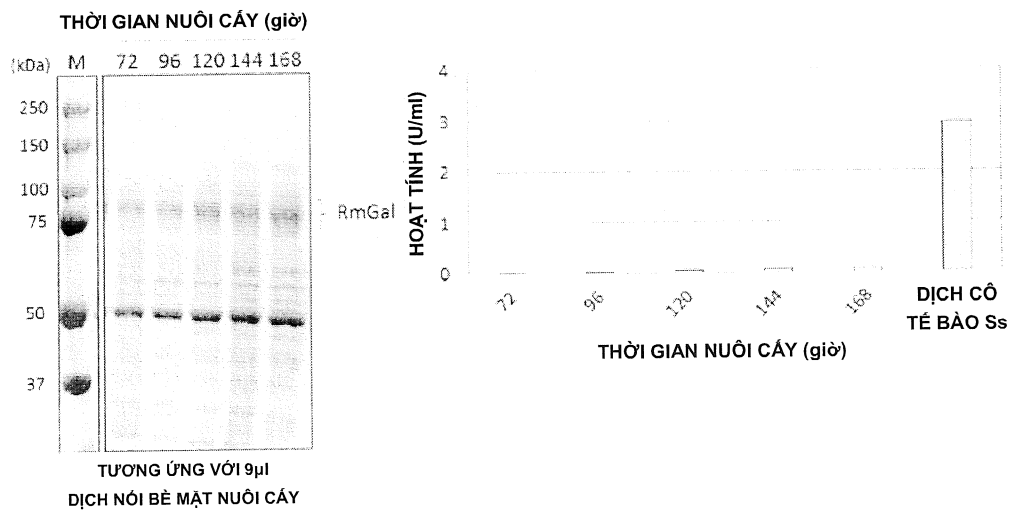


Fig.4

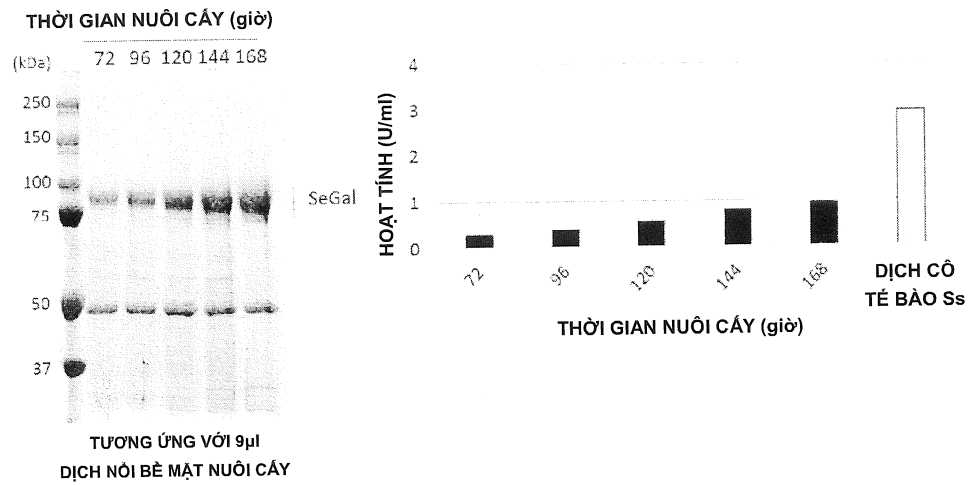


Fig.5

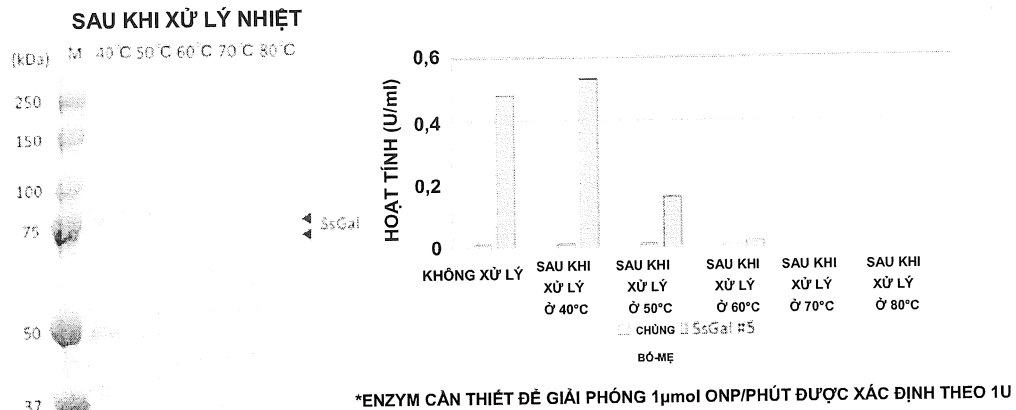
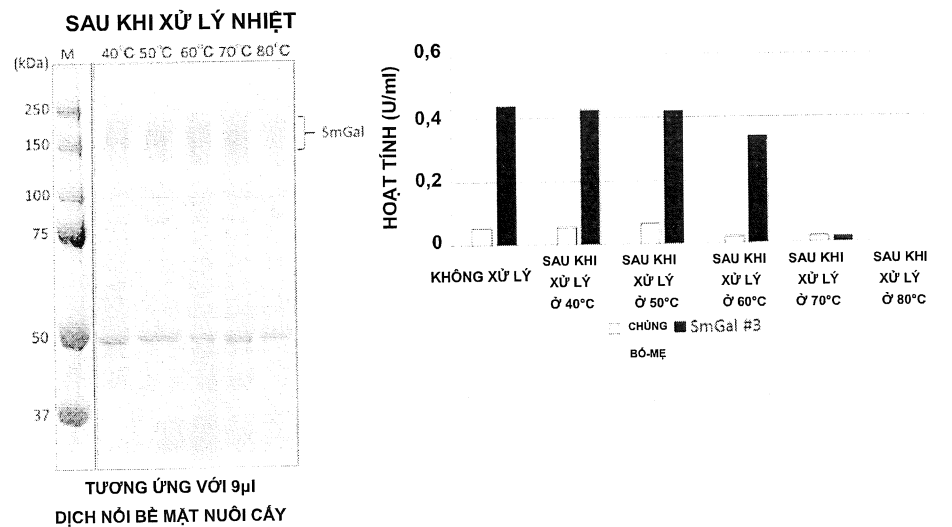


Fig.6



3/7

Fig.7

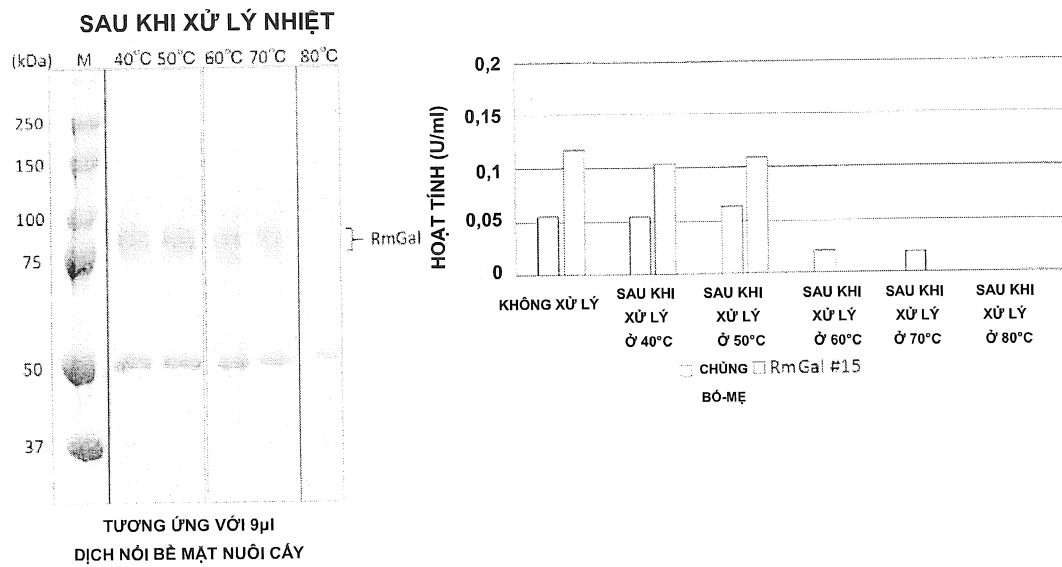
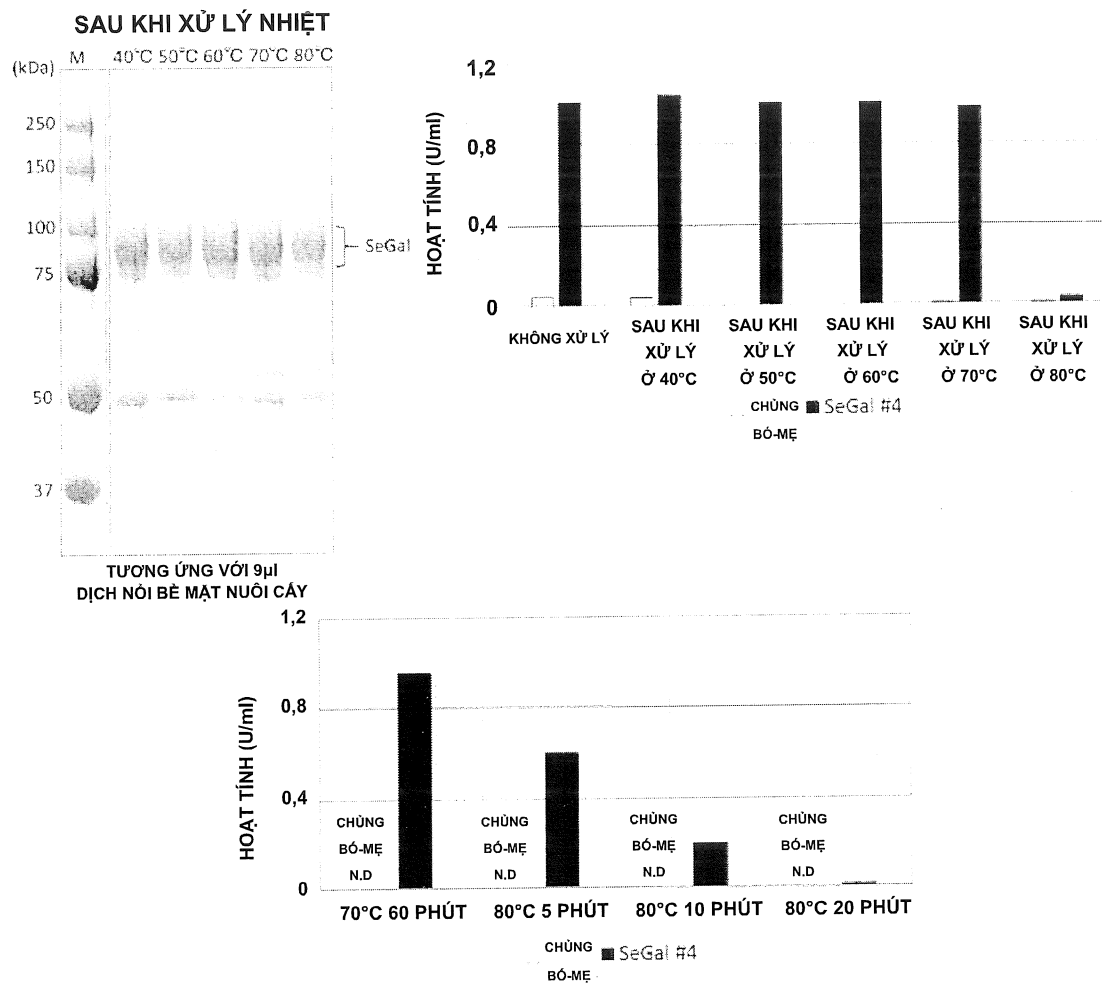


Fig.8



4/7

Fig.9

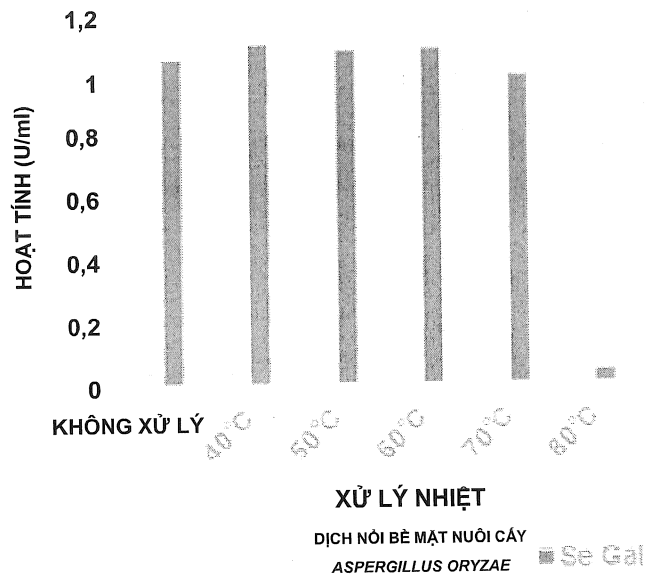
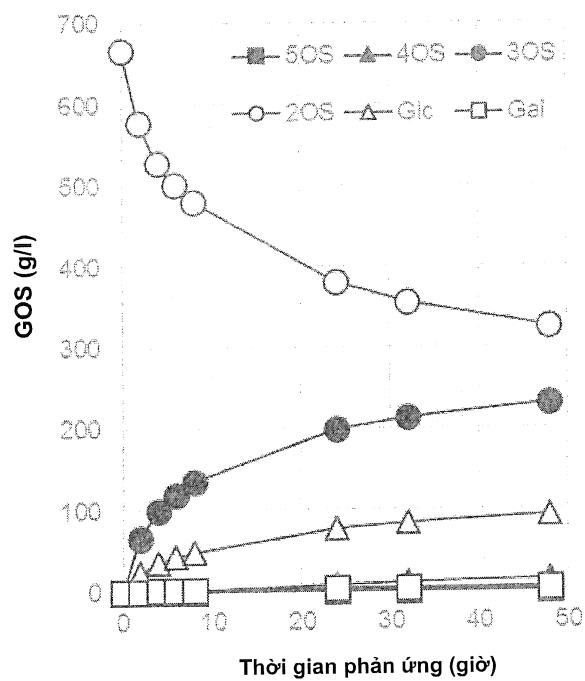


Fig.10



5/7

Fig.11

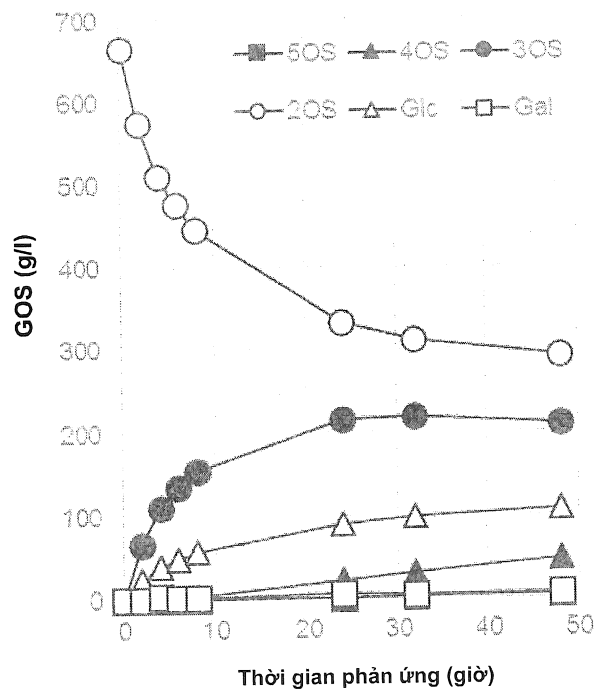


Fig.12

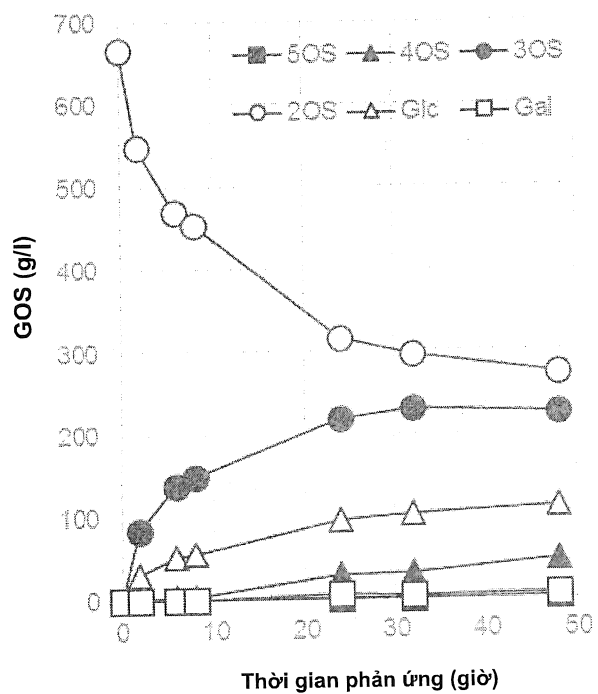


Fig.13

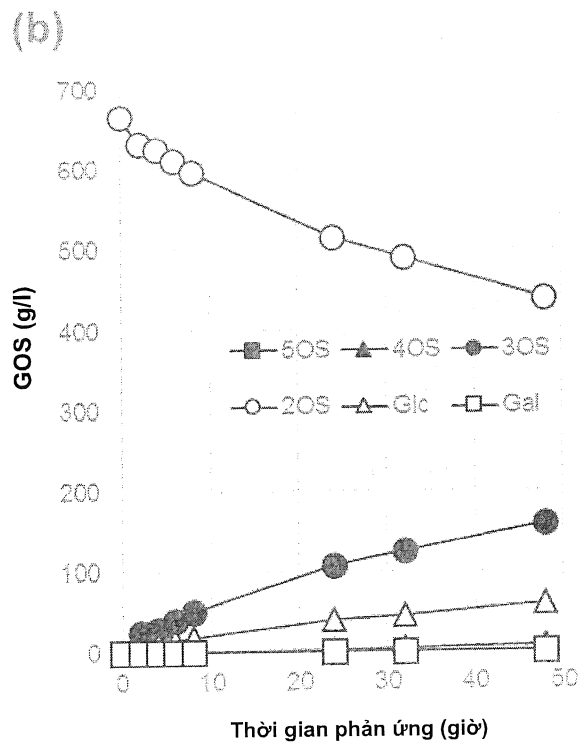
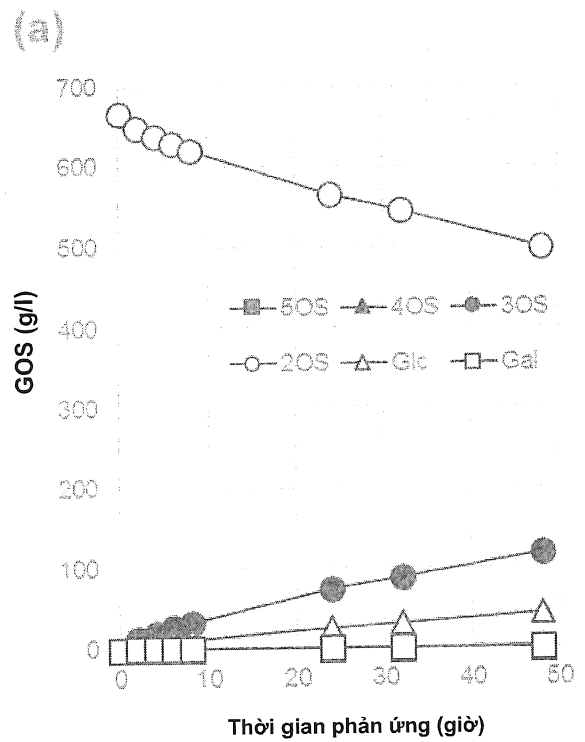


Fig.14

