

**THIẾT KẾ PROBE ĐỂ KHAI THÁC GEN MÃ HÓA
PECTINESTERASE TỪ DỮ LIỆU GIẢI TRÌNH TỰ DNA METAGENOME VÀ
ĐỒNG BIỂU HIỆN GEN *GPECS1* CỦA VI KHUẨN TRONG DẠ CỎ CỦA ĐÊ
VỚI CHAPERONE pG-KJE8 TRONG *Escherichia coli***

**Nguyễn Khánh Hoàng Việt^{1,2,3}, Đỗ Thị Huyền^{1,3}, Lê Tùng Lâm¹,
Phùng Thị Lan¹, Nguyễn Thủy Tiên^{1,4}, Phùng Thu Nguyệt¹, Trương Nam Hải^{1,3*}**

¹Viện Công nghệ sinh học, Viện Hàn lâm KH & CN Việt Nam

²Viện Công nghệ mới, Viện Khoa học và Công nghệ quân sự

³Học viện Khoa học và Công nghệ, Viện Hàn lâm KH & CN Việt Nam

⁴Trường Đại học Bách khoa Hà Nội

TÓM TẮT: Bài báo giới thiệu các bước xây dựng và sử dụng probe để khai thác gen mã hóa pectinesterase từ dữ liệu giải trình tự DNA metagenome bằng công cụ giải trình tự gen thế hệ mới. Probe được sử dụng để khai thác, lựa chọn gen mã hóa cho pectinesterase từ dữ liệu giải trình tự DNA metagenome của vi khuẩn trong dạ cỏ của dê và từ đó chọn một trình tự để biểu hiện trong *E. coli*. Theo phân loại của CAZy, pectinesterase thuộc họ carbohydrate esterases CE8 là enzyme có nhiều ứng dụng trong công nghiệp chế biến thực phẩm, xử lý môi trường, chế biến thức ăn chăn nuôi và y dược học. Dựa trên 3 trình tự thuộc họ CE8 tìm kiếm được từ cơ sở dữ liệu, chúng tôi đã xây dựng được một probe có độ dài 367 amino acid, trong đó có 200 amino acid hoàn toàn giống nhau, 72 vị trí giống nhau ở đa số các trình tự và 47 vị trí giống nhau ở một số trình tự. Sử dụng probe được thiết kế, chúng tôi đã lọc được 4 trình tự mã hóa cho pectinesterase từ dữ liệu giải trình tự DNA metagenome của vi khuẩn trong dạ cỏ của dê. Kết quả ước đoán cấu trúc không gian bằng Phyre2 đã chọn được duy nhất một trình tự (mã số 46301) có độ tin cậy, độ bao phủ lần lượt là 100% và 90% với pectinesterase. Gen đã được đặt tổng hợp và đưa vào vector pET22b(+) tại vị trí *NcoI*, *XhoI* để đồng biểu hiện với chaperone từ vector pG-KJE8 trong *E. coli*. Pectinesterase được biểu hiện ở dạng tan và có hoạt tính sinh học thủy phân cơ chất pectin. Kết quả này đã chứng minh việc sử dụng probe trong khai thác gen là khả thi.

Từ khóa: *E. coli*, chaperone pG-KJE8, *gpecs1*, pectinesterase, probe.

MỞ ĐẦU

Pectinesterase (EC 3.1.1.11) thuộc nhóm enzyme thủy phân pectinase, xúc tác khử ester hóa nhóm methoxyl của pectin tạo thành axit pectic và methanol. Enzyme này hoạt động đặc hiệu với nhóm methyleste của axit galacturonic nằm bên cạnh axit galacturonic không bị este hóa (Pooja et al., 2015). Pectinesterase được sử dụng phổ biến trong ngành công nghiệp sản xuất thực phẩm, chế biến thức ăn chăn nuôi, chế tạo chất tẩy rửa, ly trích các dược liệu và hỗ trợ thẩm tích bột giấy trong ngành công nghiệp sản xuất giấy (Pooja et al., 2015). Đặc biệt trong công nghiệp thực phẩm, pectinesterase được sử dụng để làm trong dịch nước quả, rau củ tăng hiệu suất ép, đảm bảo quá trình lọc diễn ra tốt hơn (Pooja et al., 2015). Pectinesterase có khả năng hoạt động ở pH cao (pH 8) có nhiều tính

ứng dụng thực tiễn đã được tìm thấy ở một số loài vi sinh vật như vi khuẩn *Erwinia chrysanthemi*, nấm *Trichoderma reesei* và ở loài thực vật *Lycopersicon esculentum* (Duvetter et al., 2006; Markovič et al., 1984). Việc sử dụng pectinesterase chịu nhiệt, chịu kiềm cho quá trình tiền xử lý và thủy phân pectin đã mang lại lợi ích lớn cho ngành công nghiệp sản xuất giấy, nước ép quả và một số công nghiệp khác (Kashyap et al., 2001; Chung & Jong, 2015).

Trong nghiên cứu này, chúng tôi trình bày các bước để xây dựng được probe cho pectinesterase và dựa trên probe này để có thể khai thác một cách nhanh chóng gen từ dữ liệu giải trình tự DNA metagenome của vi sinh vật trong dạ cỏ của dê. Trên thực tế, ngân hàng NCBI chứa phần lớn các trình tự pectinesterase

không được chứng minh hoạt tính bằng thực nghiệm. Để đảm bảo trình tự thu thập đáng tin cậy, chúng tôi chỉ lựa chọn các trình tự amino acid đã được nghiên cứu kỹ tính chất; so sánh độ tương đồng của nhiều trình tự để xây dựng được probe cho tìm kiếm tất cả các gen tiềm năng mã hóa cho pectinesterase trong DNA metagenome. Probe sẽ là cơ sở quan trọng cho việc lựa chọn nhanh các gen để đưa vào thực nghiệm với khả năng thành công cao và tiết kiệm thời gian khai thác gen từ DNA metagenome (Mitsuhashi et al., 1994). Gen sau khi được khai thác sẽ được tổng hợp nhân tạo và đưa vào vector biểu hiện. Nghiên cứu trước đây cho thấy, khi gen được biểu hiện không có sự hỗ trợ của chaperon, hầu hết enzyme tổng hợp ra nằm ở pha không tan (Nguyễn Khánh Hoàng Việt và nnk., 2017). Để cải thiện trạng thái này, chúng tôi đã biểu hiện thành công protein ở dạng tan bằng việc biểu hiện đồng thời gen *gpecs1* mã hóa cho pectinesterase với chaperone từ vector pG-KJE8 trong *Escherichia coli*. Kết quả này có ý nghĩa thực tiễn, là cơ sở khoa học cho các nghiên cứu xây dựng probe để khai thác nhanh gen đích và đồng biểu hiện với chaperone nhằm thu được protein dưới dạng tan.

VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Dữ liệu metagenome DNA gồm 164.644 khung đọc mở (ORF) của vi sinh vật trong dạ cỏ của dê được giải trình tự bằng hệ thống giải trình tự HiSeq Illuminia (BGI, Hồng Kông) và được chú giải chức năng dựa trên dữ liệu KEGG (Do et al., 2017). Gen *gpecs1* có kích thước 975 bp được khai thác từ dữ liệu DNA metagenome của vi sinh vật trong dạ cỏ của dê và được đưa vào vector pET22b(+) tại điểm cắt của enzyme giới hạn *NcoI* và *XhoI* để tạo thành pET22-*gpecs1*. Plasmid chaperone pG-KJE8 của hãng TaKaRa (Nhật Bản) được đưa vào cùng biểu hiện để biểu hiện kèm với gen *gpecs1*. Chủng *E. coli* Rosetta 1 được dùng làm chủng biểu hiện. Tất cả các hóa chất khác có xuất xứ từ hãng Merck (CHLB Đức).

Xác định các họ GH có chứa pectinesterase theo CAZy

CAZy (Carbohydrate-Active enzymes;

<http://www.cazy.org>) cung cấp số liệu trực tuyến và cập nhật liên tục các dữ liệu của GenBank về chuỗi thông tin của gần 340.000 enzyme tham gia chuyển hóa carbohydrate được phân vào 15 họ CE với các đặc điểm nhận biết khác nhau (Cantarel et al., 2009; Lombard et al., 2014). Dựa trên dữ liệu này, các trình tự pectinesterase đã nghiên cứu tính chất thuộc họ CE8 được thu thập, tổng hợp thành bảng; trong đó tích hợp trình tự với đặc điểm enzyme như khả năng chịu nhiệt, chịu kiềm/axit, cấu trúc không gian, trung tâm hoạt động pectinesterase.

Xây dựng probe và tìm giá trị tham chiếu

Các trình tự thu thập được ở trên sẽ được so sánh tương đồng bằng ClustalW-PBIL. Kết quả so sánh của phần mềm này sẽ cho ra một trình tự được cho là bảo tồn cao nhất (Ký hiệu Prim.cons), trong đó có các trình tự được đánh dấu về mức độ bảo tồn đặc thù cho họ enzyme. Dựa trên kết quả của Prim.cons các vị trí giống nhau hoặc tương đối giống nhau sẽ ưu tiên lựa chọn để làm probe và các trình tự trùng sẽ được loại bỏ. Probe này sẽ được so sánh lại với các trình tự đã dùng để xây dựng nên probe bằng BLASTP. Giá trị tham chiếu được xác định là giá trị E, điểm tối đa, độ bao phủ và mức độ tương đồng thấp nhất mà probe bắt được với trình tự.

Lựa chọn nhanh và ước đoán cấu trúc bậc ba

Sử dụng BLASTP để so sánh probe pectinesterase CE8 với trình tự amino acid của các ORF thuộc metagenome của vi sinh vật trong dạ cỏ của dê. Các trình tự có giá trị điểm tối đa, mức độ bao phủ, độ tương đồng cao hơn giá trị tham chiếu sẽ được lựa chọn. Việc chọn lựa có thể kèm theo việc đánh giá qua chỉ số alkaline để ưu tiên chọn được trình tự có khả năng chịu kiềm. Sau đó, các trình tự này được so sánh lại với kết quả chú giải gen dựa trên dữ liệu KEGG và CAZy. Cấu trúc bậc ba của phân tử được ước đoán dựa trên các trình tự và với các protein khung đã được nghiên cứu về cấu trúc bằng phần mềm Phyre2.

Biểu hiện gen *gpecs1* với protein chaperone

Chủng biểu hiện chứa cả hai plasmid pG-KJE8 và pET22-*gpecs1* được nuôi cấy qua đêm trong môi trường Luria Broth có ampicillin

nồng độ 100 µg/ml và chloramphenicol nồng độ 20 µg/ml (LBAC) ở 37°C, 200 vòng/phút. Sau đó, chuyển 1% dịch tế bào nuôi cấy sang môi trường LBAC mới ở nhiệt độ 25°C, nuôi trong khoảng 1 giờ để OD₆₀₀ đạt 0,1; bổ sung L-arabinose 0,5 mg/ml và tetracycline 5 ng/ml, tiếp tục nuôi cấy trong cùng điều kiện đến khi OD₆₀₀ = 0,6 – 0,8; cảm ứng IPTG nồng độ 0,1 mM. Sau 5 giờ cảm ứng, mật độ tế bào được đo lại; tế bào được thu lại từ dịch nuôi cấy bằng phương pháp ly tâm rồi hòa vào nước để đạt OD₆₀₀ = 10. Xử lý mẫu và điện di kiểm tra protein trên gel SDS-PAGE 12,6%.

Kiểm tra sơ bộ hoạt tính pectinesterase

Mẫu tế bào OD₆₀₀ = 10 được siêu âm để thu protein pha tan dùng cho thử nghiệm hoạt tính sơ bộ theo phương pháp phát hiện nhanh hoạt tính phân huỷ pectin của Ronald (Ronald, 1978). Để rút ngắn thời gian thực hiện, phương pháp được thử nghiệm trực tiếp trong ống eppendorf thay vì trên đĩa thạch. Tổng thể tích của phản ứng là 575 µl/mẫu, trong đó mẫu thí nghiệm (PE) gồm có: (1) Cơ chất pectin 2% với thể tích là 200 µl; (2) PBS 30X, pH = 7,4 với thể tích là 35 µl; (3) bromothymol blue 5% được bổ sung với thể tích 40 µl, quan sát sự chuyển màu của mẫu, nếu thấy xuất hiện màu xanh lá cây là pH trung tính (pH = 7-7,5); (4) sau đó bổ sung pectinesterase với thể tích 300 µl. Mẫu đối chứng âm 1 (ĐC 1) với các thành phần tương tự mẫu PE nhưng không bổ sung pectinesterase mà thay bằng nước deion khử

trùng với thể tích tương đương; mẫu đối chứng âm 2 (ĐC 2) được bổ sung enzyme với thể tích 300 µl/mẫu nhưng sau khi quá trình ủ kết thúc. Quan sát sự chuyển màu của các mẫu khi quá trình ủ sau 4 giờ, ở 37°C kết thúc.

KẾT QUẢ VÀ THẢO LUẬN

Khai thác trình tự amino acid của pectinesterase đã được nghiên cứu đặc tính để xây dựng probe

Trên cơ sở dữ liệu của CAZy, pectinesterase (3.1.1.11) được phân vào họ CE8 có mô hình cấu trúc không gian (β)-helix.

Số liệu DNA metagenome chủ yếu từ vi khuẩn nên để đảm bảo kết quả đáng tin cậy, chúng tôi chỉ thu thập các trình tự amino acid của enzyme có nguồn gốc từ vi khuẩn. Chúng tôi đã tìm kiếm được 3 trình tự thỏa mãn các yêu cầu đề ra từ dữ liệu các trình tự mã hóa cho pectinesterase đã được nghiên cứu tính chất (bảng 1).

Kết quả phân tích sau khi sử dụng 3 trình tự thuộc họ CE8 bằng phần mềm ClustalW – PBIL (hình 1) cho thấy, khi so sánh 3 trình tự amino acid thu thập được có 200 amino acid hoàn toàn giống nhau, 72 vị trí giống nhau ở đa số các trình tự và có 47 vị trí giống nhau ở một số trình tự, còn lại là khác nhau. Probe của họ CE8 bao gồm có 367 amino acid đã được xây dựng chủ yếu dựa trên các trình tự bảo tồn cao được lựa chọn này (hình 2).

Bảng 1. Tổng hợp dữ liệu đã được nghiên cứu chi tiết về pectinesterase

STT	Mã số trong GENBANK	Vi khuẩn	Số amino acid
1	CAA68628.1	<i>Dickeya chrysanthemi</i>	361
2	P0C1A9.1	<i>Erwinia chrysanthemi</i> strain 3937	366
3	PDB: 3UW0_A	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> 8081	364

```

CAA686281    MLKTISGTLALSLLIAASVHQQAATTYNVAVVSKSSSDGKTFKTIADAIASAPAGSTPFV
P0C1A91      MLKTISGTLALSLLIAASVHQQAATTYNVAVVSKSSSDGKTFKTIADAIASAPAGSTPFV
PDBxxx2     MPINLLGKTLWLGLISFAVLGTVNAAQYNVAVVSTTPQGFDEFSSINAALKSAPKDDTPFI
*   .: * .      :*: :* :   *: ***** :.:* .:* * :  *** ..***:
Prim.cons.   MLKTISGTLALSLLIAASVHQQAATTYNVAVVSKSSSDGKTFKTIADAIASAPAGSTPFV
CAA686281    ILIKNGVYNERLTI TRNLLLKGESRNGAVIAAATAAGTLKSDGSKWGTAGSSTITISAK
P0C1A91      ILIKNGVYNERLTI TRNLLHLKGESRNGAVIAAATAAGTLKSDGSKWGTAGSSTITISAK
PDBxxx2     IFLKNGVYTERLEVARSHVTLKGENRDGTVIGANTAAGMLNPQGEKWGTSGSSTVLVNP
*: :*****.*** ::*.: : *****.*:***.* ***** *.:*.*****:*****: ..*
    
```


Khai thác trình tự mã hóa cho pectinesterase bằng probe từ dữ liệu DNA metagenome

Dựa trên dữ liệu KEGG, công ty BGI đã chú giải 80 trình tự có hoạt tính pectinesterase và cả 80 trình tự này từ dữ liệu DNA metagenome của vi khuẩn trong dạ cỏ của dê đều được nhận biết bằng probe đã thiết kế. Tuy nhiên, không có trình tự nào có tổng điểm

tối thiểu là 418, độ bao phủ và độ tương đồng tối thiểu lần lượt trên 98% và 57%. Hơn nữa, việc tìm kiếm enzyme cần dựa thêm vào chỉ số alkaline để đánh giá mức độ chịu kiềm. Chúng tôi khảo sát giá trị alkaline và nhận thấy có 4 trình tự phù hợp, có chỉ số alkaline cao nhất được ưu tiên lựa chọn (bảng 3).

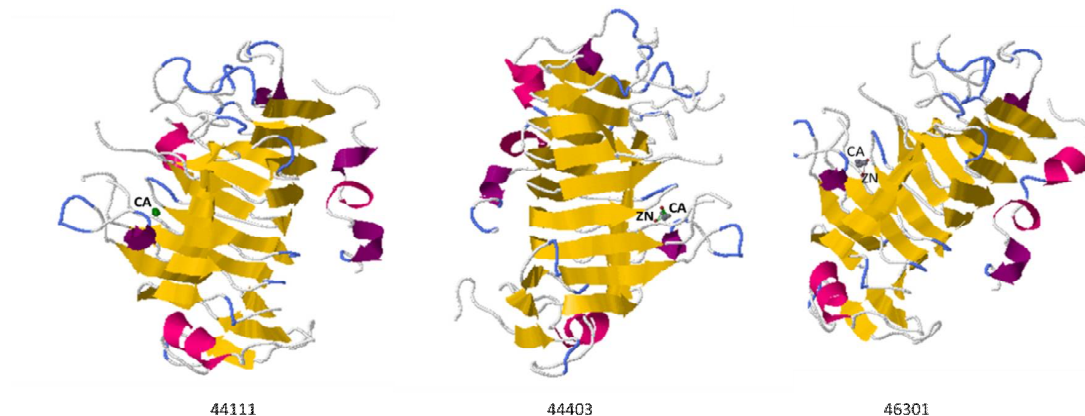
Bảng 3. Khảo sát giá trị alkaline của các trình tự phù hợp về so sánh độ bao phủ

Mã trình tự	Nhiệt độ (°C)	pH	
		Enzyme axit	Enzyme kiềm
44111	> 65	0,021	0,979
43901	55-65	0,030	0,970
44403	< 55	0,033	0,967
46301	55-65	0,059	0,941

Khảo sát cấu trúc của các trình tự pectinesterase được khai thác bằng probe

Để chắc chắn hơn, chúng tôi tiến hành khảo sát cấu trúc không gian của enzyme mã hóa bởi gen của 4 trình tự có chỉ số alkaline cao nêu trên. Vì các trình tự khai thác đều có độ tương đồng thấp với gen trên ngân hàng gen dựa trên BLASTP nên chúng tôi đã sử dụng phần mềm Phyre2 để ước đoán cấu trúc không gian. Kết

quả cho thấy chỉ có 3 trình tự là 44111, 44403 và 46301 xuất hiện phối tử trong cấu trúc không gian (hình 3). Trong đó, trình tự 46301 có độ tin cậy (100%), độ bao phủ (90%) cao nhất trong 3 trình tự được lựa chọn và có phối tử CA, ZN liên quan đến hoạt tính sinh học với pectinesterase của khuôn 1gq8a_ theo ước đoán của Phyre2.



Hình 3. Cấu trúc bậc ba của các pectinesterase dựa trên khuôn (template) 1gq8a_

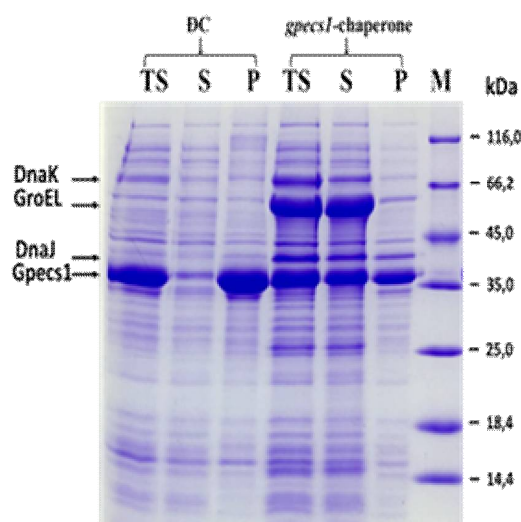
Biểu hiện gen plasmid tái tổ hợp pET22-gpecs1 trong *E. coli* Rosetta 1

Gen lựa chọn sau đó được tổng hợp nhân tạo, đưa vào vector biểu hiện; tuy nhiên protein được biểu hiện trong *E. coli* nằm hoàn toàn ở pha tủa. Để cải thiện tính tan của protein ngoài

các phương pháp tối ưu điều kiện biểu hiện như nhiệt độ, môi trường nuôi cấy và nồng độ chất cảm ứng IPTG, việc biểu hiện protein kèm với một chaperone làm tăng lượng protein biểu hiện đúng cấu trúc trong bộ máy vi khuẩn (Martínez et al., 2010) như DnaK và GroESL giúp ngăn

cản sự hình thành protein dạng không tan và giúp protein cuộn xoắn đúng cấu trúc.

Trong thí nghiệm này, chúng tôi tiến hành biểu hiện kèm plasmid tái tổ hợp pET22-*gpecs1* cùng với plasmid chaperone pG-KJE8 (chứa protein GroEL và GroES, DnaK và DnaJ) trong chủng tế bào *E. coli* Rosetta 1. Kết quả cho thấy, protein chaperone GroEL, DnaK, DnaJ và protein Gpecs1 đều được biểu hiện ở mức cao, đúng kích thước lần lượt khoảng 70 kDa, 60 kDa và 40 kDa, cùng với protein Gpecs1 có kích thước khoảng 37 kDa. Như vậy, protein Gpecs1 được biểu hiện kèm với chaperone đã làm tăng tính tan của protein Gpecs1 lên đáng kể, đạt khoảng trên 50% lượng enzyme được tổng hợp (hình 4).



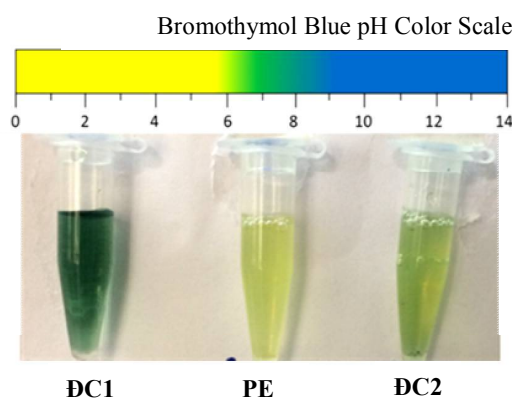
Hình 4. Protein Gpecs1 biểu hiện kèm chaperone pG-KJE8 trong chủng *E. coli* Rosetta 1

M: protein chuẩn (Fermentas); DC: protein không biểu hiện kèm chaperone; *gpecs1*-Chaperone: protein Gpecs1 biểu hiện kèm chaperone pG-KJE8; TS: protein tổng số; S: protein pha tan; P: protein pha không tan.

Kiểm tra sơ bộ hoạt tính của pectinesterase

Sau khi biểu hiện thành công protein Gpecs1 ở dạng tan, chúng tôi tiến hành kiểm tra hoạt tính của enzyme bằng chất chỉ thị pH bromothymol blue. Kết quả thử hoạt tính cho thấy mẫu pectinesterase ở dạng tan sẽ chuyển sang màu vàng đặc trưng của dung dịch chứa

pectinic axit khi ủ với bromothymol blue ở 37°C trong 4 giờ. Trong khi mẫu đối chứng 1 (ĐC1) có cùng thành phần nhưng thay enzyme bằng nước deion, không xuất hiện màu vàng mà vẫn giữ nguyên màu đặc trưng của dung dịch có pH hơi kiềm. Mẫu đối chứng 2 (ĐC2) trong cùng điều kiện nhưng enzyme được bổ sung sau quá trình ủ, xuất hiện màu vàng xanh; điều đó chứng tỏ trong điều kiện nhiệt độ 37°C và được ủ sau 4 giờ để phản ứng xảy ra, lượng axit sinh ra cao hơn nhiều so với mẫu ĐC2 không được ủ. Như vậy, chúng tôi kết luận pectinesterase được biểu hiện trong *E. coli* tái tổ hợp có hoạt tính sinh học.



Hình 5. Kết quả kiểm tra sơ bộ hoạt tính của pectinesterase (PE) và thang màu chỉ thị pH bromthymol blue

ĐC1: mẫu đối chứng âm không chứa enzyme PE; PE: mẫu chứa enzyme PE; ĐC2: mẫu chứa enzyme PE bổ sung sau khi ủ.

KẾT LUẬN

Trong công trình nghiên cứu này, chúng tôi đã xây dựng probe để lựa chọn được một gen mã hóa cho pectinesterase từ dữ liệu giải trình tự metagenome của vi sinh vật trong dạ cỏ của dê. Trình tự này đã được kiểm chứng lại về cấu trúc không gian bằng phần mềm Phyre2. Dựa trên kết quả khai thác, gen được lựa chọn đã được tổng hợp và biểu hiện đồng thời với chaperone pG-KJE8 trong *E. coli* để tăng khả năng tan của protein biểu hiện. Kết quả biểu hiện thành công khi thu được protein Gpecs1 ở dạng tan và có hoạt tính sinh học.

Lời cảm ơn: Công trình được thực hiện bằng nguồn kinh phí của Đề tài độc lập “Nghiên cứu metagenome của một số hệ sinh thái mini tiềm năng nhằm khai thác các gen mới mã hóa hệ enzyme chuyển hóa hiệu quả lignocelluloses” mã số ĐTĐLNCN.15/14 và trang thiết bị của phòng Thí nghiệm trọng điểm Công nghệ gen.

TÀI LIỆU THAM KHẢO

- Cantarel B. L., Coutinho P. M., Rancurel C., Bernard T., Lombard V., Henrissat B., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.*, 37 (Database issue): 233-238.
- Chung H. K., Jong J. C., 2015. Characterization of the intact form of *Thermotoga maritima* pectinase TmPecN expressed in *Escherichia coli*. *J. Appl. Biol. Chem.*, 58 (2): 97-100.
- Do T. H., Le N. G., Dao T. K., Nguyen T. M. P., Le T. L., Luu H. L., Nguyen K. H. V., Nguyen V. L., Le L. A., Phung T. N., Straalen N. M., Roelofs D., Truong N. H., 2018. Metagenomic insights into lignocellulose-degrading genes through Illumina-based de novo sequencing of the microbiome in Vietnamese native goats rumen. *J. Gen Appl. Microbiol.* DOI: 10.2323/jgam.2017.08.004. [Epub ahead of print].
- Duvetter T., Fraeye I., Sila D. N., Verlent I., Smout C., Hendrickx M., Van L. A., 2006. Mode of de-esterification of alkaline and acidic pectin methyl esterases at different pH conditions. *J. Agric. Food Chem.*, 54 (20): 7825-7831.
- Kashyap D. R., Vohra P. K., Chopra S., Tewari R., 2001. Applications of pectinases in the commercial sector: a review. *Bioresour Technol.*, 77(3): 215-227.
- Lombard V., Golaconda R. H., Drula E., Coutinho P. M., Henrissat B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.*, 42 (Database issue): 490-495.
- Martínez A. M., García F. E., Ferrer M. N., Rinas U., Villaverde A., 2010. Side effects of chaperone gene co-expression in recombinant protein production. *Microb Cell Factories.*, DOI: 10.1186/1475-2859-9-64.
- Markovič O., Kohn R., 1984. Mode of pectin deesterification by *Trichoderma reesei* pectinesterase. *Experientia*, 40(8): 842-843.
- Mitsuhashi M., Cooper A., Ogura M., Shinagawa T., Yano K., Hosokawa T., 1994. Oligonucleotide probe design: a new approach. *Nature*, 367(6465): 759-761.
- Nguyễn Khánh Hoàng Việt, Lê Tùng Lâm, Phùng Thị Lan, Đỗ Thị Huyền, Trương Nam Hải, 2017. Nghiên cứu biểu hiện GPECS1 mã hóa pectinesterase khai thác từ dữ liệu giải trình tự DNA metagenome vi khuẩn dạ cỏ dê trong tế bào *Escherichia coli*, sử dụng vector pET22b(+). *Tạp chí Y học Việt Nam*, 468 (số đặc biệt): 197-203.
- Pooja K., Manmohit K., Reena G., 2015. Pectin methylesterases: a review. *J Bioprocess Biotech.*, 5(5): 1-7.
- Ronald E. Z., 1978. A rapid assay for pectinesterase activity which can be used as a prescreen for pectinesterase inhibitors. *Anal Biochem.*, 85(s1): 219-223.

PROBE DESIGN FOR EXPLOITING GENE ENCODING PECTINESTERASE FROM DNA METAGENOME DATA OF BACTERIA IN GOAT RUMEN AND CO-EXPRESSION OF *GPECS1* GEN WITH CHAPERONE pG-KJE8 IN *Escherichia coli*

**Nguyen Khanh Hoang Viet^{1,2,3}, Do Thi Huyen^{1,3}, Le Tung Lam¹,
Phung Thi Lan¹, Nguyen Thuy Tien^{1,4}, Phung Thu Nguyet¹, Truong Nam Hai^{1,3*}**

¹Institute of Biotechnology, VAST

²Institute of New Technology, Academy of Military Science and Technology

³Graduate University of Science and Technology, VAST

⁴Hanoi University of Science and Technology

SUMMARY

This article introduces the steps of constructing and using probe to exploit the gene encoding pectinesterase from metagenome DNA sequencing data by next generation gene sequencing tools. Probe was used to exploit and select the gene encoding for pectinesterase from the metagenome DNA sequences of bacteria in goat rumen and thereby select a sequence to express in *E. coli*. According to the CAZy classification system, pectinesterase belongs to the family of carbohydrates esterases CE8 is an enzyme that has many applications in the food processing industry, environmental treatment, animal feed processing and medicine. As the results, 3 sequences of CE8 was retrieved from CAZy database and one probe was designed, this probe length was 367 amino acids contained all the conserved amino acid residues: 200 conserved residues in all sequence, 72 residues similar in almost sequences and residues conserved in many sequences and homologus; choosed highest alkalinity index. Using the probe designed, we filtered four coding sequences for pectinesterase from metagenome DNA sequencing data of bacteria in goat rumen. Spatial structure estimation with Phyre2 has only one sequencing (code 46301) with 100% sequence identity and 90% query coverage with pectinesterase. A artificial gene were synthesized and inserted into the vector pET22b (+) at the *NcoI*, *XhoI* to co-express with chaperone pG-KJE8 in *E. coli*. The recombinant pectinesterase enzyme is expressed in soluble form and has a pectin substrate biodegradation activity. The results demonstrate that using probe for gene extraction is feasible.

Keywords: Chaperone pG-KJE8, *E. coli*, *gpecs1*, pectinesterase, probe.

Citation: Nguyen Khanh Hoang Viet, Do Thi Huyen, Le Tung Lam, Phung Thi Lan, Nguyen Thuy Tien, Phung Thu Nguyet, Truong Nam Hai, 2018. Probe design for exploiting gene encoding pectinesterase from DNA metagenome data of bacteria in goat rumen and co-expression of *gpecs1* gen with chaperone pG-KJE8 in *Escherichia coli*. Tap chi Sinh hoc, 40(1): 84-91. DOI: 10.15625/0866-7160/v40n1.10917.

*Corresponding author: thuy_tt@hnue.edu.vn

Received 19 October 2017, accepted 12 January 2018