

NGHIÊN CỨU BIỂU HIỆN GEN *eg9* MÃ HÓA ENDOGLUCANASE GH8 CÓ NGUỒN GỐC TỪ DỮ LIỆU GIẢI TRÌNH TỰ DNA ĐA HỆ GEN CỦA VI KHUẨN TRONG MÙN XUNG QUANH NẤM MỤC TRẮNG THỦY PHÂN GỠ TRONG TẾ BÀO *ESCHERICHIA COLI*

Đỗ Thị Huyền^{1,2}, Lê Thu Hoài^{1,3}, Nguyễn Hải Đăng¹, Nguyễn Thị Quý¹, Trương Nam Hải^{1,2,✉}

¹Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

³Học viện Nông nghiệp Việt Nam

✉Người chịu trách nhiệm liên lạc. E-mail: tnhai@ibt.ac.vn

Ngày nhận bài: 02.11.2020

Ngày nhận đăng: 02.3.2021

TÓM TẮT

Endoglucanase là một loại enzyme quan trọng tham gia vào quá trình thủy phân cellulose và có thể được sử dụng trong nhiều lĩnh vực công nghiệp. Gen (mã số GL0183420 gọi tắt là *eg9*) có kích thước 1398 bp được khai thác từ dữ liệu giải trình tự DNA metagenome của vi khuẩn trong mùn xung quanh khu nấm mục trắng thủy phân gỗ, mã hóa cho endoglucanase họ GH8. Đầu 5' của gen có 126 nucleotide tận cùng mã hóa cho tín hiệu tiết ở vi khuẩn gram âm và 1269 nucleotide tiếp theo mã hóa endoglucanase GH8. Trong nghiên cứu này, gen *eg9* (không chứa trình tự mã hóa tín hiệu tiết) đã được tối ưu mã bộ ba cho biểu hiện gen ở *E. coli*, sau đó được đặt tổng hợp nhân tạo và chuyển vào vector pET22b(+) để tiến hành biểu hiện gen trong các chủng *E. coli*. Qua khảo sát 5 chủng *E. coli* biểu hiện bao gồm Rosetta, JM109, Origami, SoluBL21, C43, EG9 biểu hiện tốt nhất trong chủng Rosetta và C43. Tuy nhiên, ở nhiệt độ nuôi cấy 30°C trong môi trường LB, hầu hết EG9 được biểu hiện ở pha không tan trong đó Rosetta là chủng có khả năng biểu hiện EG9 tốt nhất. Enzyme EG9 có thể được biểu hiện tốt trong chủng Rosetta ở các thời điểm cảm ứng khi OD600 của tế bào dao động từ 0,2 đến 1. Sau khi giảm nhiệt độ nuôi cấy xuống 25°C, 50% EG9 đã được biểu hiện ở pha tan và có hoạt tính thủy phân tốt cơ chất CMC tạo vòng trong trên đĩa thạch, trong đó môi trường PE là môi trường thích hợp nhất cho chủng Rosetta tái tổ hợp mang gen *eg9* sinh trưởng và tổng hợp endoglucanase EG9. EG9 được biểu hiện tốt ở các nồng độ chất cảm ứng IPTG khác nhau từ 0,05 mM đến 0,9 mM. Kết quả này mang lại tiềm năng lớn cho việc sản xuất EG9 tái tổ hợp phục vụ nghiên cứu tính chất của enzyme.

Từ khóa: Biểu hiện gen, *eg9*, endoglucanase *gh8*, *Escherichia coli*, mùn

MỞ ĐẦU

Cellulase là nhóm enzyme phân cắt liên kết β -(1,4) glycoside giữa các gốc đường trên sợi cellulose (Davies *et al.*, 1993) để giải phóng glucose cho vi sinh vật sinh trưởng phát triển. Nguồn đường này còn có giá trị trong lên men sản xuất nhiều sản phẩm quan trọng như bioethanol. Cellulase được chia thành ba nhóm chính đó là endoglucanase (EC 3.2.1.4), exoglucanase/cellobiohydrolase (EC 3.2.1.176) và β -glucosidase (EC 3.2.1.21) (Bhat, Bhat 1997). Trong 3 loại enzyme trên thì endoglucanase là enzyme quan trọng nhất phân cắt ngẫu nhiên sợi cellulose để beta-glucosidase và cellobiohydrolase phân cắt tiếp thành glucose. Đây là enzyme được sử dụng trong nhiều lĩnh vực công nghiệp (Lin, Fu, Huang 2016).

Theo phân loại của CAZy, dựa trên sự tương đồng của trình tự axit amin, thành phần cấu tạo và cấu trúc vùng xúc tác, cellulase được xếp vào nhóm glycoside hydrolase (GH) là enzyme xúc tác cho phản ứng thủy phân các liên kết glycoside có trong các đường phức. Glycoside hydrolase họ 8 (GH8) bao gồm endoglucanase, licheninase, chitosanase và xylanase là các enzyme cần thiết cho sự phân hủy polysaccharide (Ontañon *et al.*, 2019). Nhóm enzyme này có vai trò chuyển hóa cellulose không hòa tan thành cello-oligosaccharide (COS). Đây là nhóm mạch đường ngắn có tiềm năng ứng dụng trong ngành công nghiệp thức ăn và thực phẩm. Chúng có khả năng tăng cường sự phát triển của vi khuẩn sinh axit lactic, và chúng được sử dụng như prebiotics (Karnaouri *et al.*, 2019).

Trong nghiên cứu này để tìm kiếm được các endoglucanase mới có hoạt tính tốt, chúng tôi đã lựa chọn một trình tự từ dữ liệu giải trình tự DNA metagenome của vi khuẩn trong mùn xung quanh khu nấm mục trắng thủy phân gỗ, mã hóa cho endoglucanase GH8 để nghiên cứu biểu hiện, kiểm tra hoạt tính để hướng tới thu được enzyme cho nghiên cứu tính chất, xem xét khả năng ứng dụng.

VẬT LIỆU VÀ PHƯƠNG PHÁP

Vật liệu

Gen *eg9* mã số GL0183420 (1398 bp) được khai thác từ dữ liệu giải trình tự DNA metagenome của vi khuẩn xung quanh khu nấm mục trắng thủy phân gỗ. Trong đó 126 nucleotide đầu 5' mã hóa cho tín hiệu tiết ở vi khuẩn gram âm và 1269 nucleotide tiếp theo mã hóa endoglucanase họ GH8 có 423 amino acid, có khối lượng phân tử theo tính toán lý thuyết khoảng 49 kDa. Dựa vào kết quả BLAST, gen có nguồn gốc từ vi khuẩn *Paraburkholderia*.

Các chủng *E. coli* dùng cho biểu hiện gen bao gồm chủng Rosetta, JM109, Origami, SoluBL21, C43 có trong bộ sưu tập chủng biểu hiện của phòng Kỹ thuật di truyền.

Thiết kế vector biểu hiện gen *eg9*

Trình tự gen *eg9* mã hóa cho enzyme trưởng thành được kiểm tra mức độ phù hợp mã bộ ba với mã bộ ba của *E. coli* bằng phần mềm tìm kiếm mã hiếm của Genscript (Mỹ). Sau đó gen *eg9* được tối ưu mã bộ ba cho phù hợp với mã bộ ba của chủng biểu hiện *E. coli*, được tổng hợp nhân tạo tại công ty Genscript và chuyển vào vector pET22b(+) tại vị trí *NcoI-XhoI* để tạo thành vector pET22-*eg9*. Trình tự gen *eg9* trong vector tái tổ hợp đã được kiểm tra bằng giải trình tự gen, và kiểm tra khung đọc.

Biểu hiện EG9 trong các chủng vi khuẩn *E. coli*

Để so sánh và tuyển chọn các dòng vi khuẩn có khả năng biểu hiện protein cao nhất, chúng tôi tiến hành biến nạp DNA plasmid pET22-*eg9* vào 5 chủng *E. coli* khả biến bao gồm Rosetta, JM109, Origami, SoluBL21, C43 theo phương pháp sốc nhiệt của Sambrook và Russell (Sambrook, Russell 2001). Các khuẩn lạc tái tổ hợp mang vector pET22-*eg9* được nuôi cấy riêng rẽ trong môi trường LBA (1% trypton, 1% NaCl, 0,5% cao nấm men, có bổ sung ampicillin 100 µg/ml) ở 37°C, 170 vòng/phút qua đêm. Sau đó, dịch tế bào nuôi cấy được chuyển sang môi trường LBA mới với tỷ lệ tiếp giống là 2%. Tế bào được nuôi cấy ở cùng điều kiện đến khi OD₆₀₀

đạt khoảng 0,6-0,8 thì tiến hành bổ sung IPTG với nồng độ cuối cùng 0,5 mM để cảm ứng sinh tổng hợp enzyme ngoại lai. Tế bào sau cảm ứng được nuôi ở 25°C, 170 vòng/phút trong 4-5 giờ. Sau thời gian nuôi cấy, tế bào được thu lại bằng ly tâm 8000 vòng/phút trong 5 phút và huyền phù vào nước sao cho OD₆₀₀=10 và giữ tế bào ở -20°C cho đến khi phân tích protein. Nhiệt độ cảm ứng, nồng độ chất cảm ứng (0, 0,05; 0,1; 0,3; 0,6; 0,9 mM IPTG) thích hợp cho biểu hiện gen cũng được khảo sát. Các môi trường bao gồm LB (1% tryptone, 0,5% cao nấm men, 1% NaCl), SB (3,2% peptone, 0,5% NaCl, 2% cao nấm men), PE (1% cao nấm men, 2% peptone), TB (1,2% peptone, 2,4% cao nấm men, 72 mM K₂HPO₄, 17 mM KH₂PO₄, 0,4% glycerol) và (5) TBD (1,2% peptone, 2,4% cao nấm men, 72 mM K₂HPO₄, 17 mM KH₂PO₄, 0,24% glucose) cũng được khảo sát để tìm ra môi trường thích hợp cho biểu hiện gen trong chủng Rosetta mang pET22-*eg9*.

Để đánh giá khả năng biểu hiện cũng như khả năng tan của protein tái tổ hợp, tế bào biểu hiện từ trong -20°C có OD₆₀₀=10 được làm tan nhanh ở 40°C, sau đó được phá vỡ bằng siêu âm với chu kỳ 10 giây siêu âm và 20 giây nghỉ trong tổng khoảng 30 đến 40 xung ở tần số 20 kHz. Dịch siêu âm được ly tâm 13.000 vòng/phút trong 10 phút để tách protein pha tan và pha không tan. Pha tan là phần dịch nổi được thu lại, còn phần cặn là pha không tan được hòa trở lại trong nước cất vô trùng và điện di trên gel polyacrylamide 12,6% (Laemmli 1970) để kiểm tra so sánh với mẫu protein tổng số, xác định khả năng tan, tủa của enzyme.

Khảo sát sơ bộ hoạt tính enzyme trong dịch nuôi cấy

Dịch protein tổng số, protein pha tan, pha tủa của các lần nuôi cấy được tiến hành kiểm tra hoạt tính endoglucanase trên đĩa thạch chứa 1% CMC, 1,4% agar trong đệm PBS pH6. Các khoanh giấy Whatman có kích thước 5 mm được đặt lên đĩa thạch. Sau đó nhỏ 20 µl mẫu hoặc 10 µl cellulase (0,5 U/ml cellulase, Sigma) lên từng khoanh giấy và ủ đĩa thạch qua đêm ở 37°C. Đĩa sau khi lấy ra được nhuộm với đỏ Congo 0,1% trong khoảng 1 giờ và rửa lại bằng NaCl 1M cho tới khi vòng sáng trong xung quanh đối chứng dương xuất hiện.

KẾT QUẢ VÀ THẢO LUẬN

Cải biến mã bộ ba gen *eg9* và thiết kế vector biểu hiện gen

Gen mã hóa enzyme EG9 hoàn thiện (không chứa trình tự tín hiệu tiết gồm 1269 nucleotide) đã được

kiểm tra mức độ phù hợp mã bộ ba cho việc biểu hiện gen trong tế bào *E. coli*. Kết quả cho thấy gen có 70,84% GC, chỉ số phù hợp mã bộ ba (CAI: codon adaptation index) đạt 0,75 và phần trăm mã hiếm là 6%. Với kết quả này, tỷ lệ % GC của gen là hơi cao cho việc biểu hiện gen trong *E. coli*. Bên cạnh đó, 6% mã hiếm cũng sẽ ảnh hưởng tới mức độ dịch mã của gen. Vì vậy, gen đã được tối ưu mã bộ ba để tăng chỉ

số phù hợp lên 0,92, loại bỏ 100% mã hiếm và tỷ lệ % GC giảm xuống còn 66,88%. Trình tự gen gốc và gen tối ưu mã bộ ba được mô tả ở Hình 1.

Gen sau đó được tổng hợp hóa học, chuyển vào vector pET22b(+) để tạo vector pET22-*eg9*. Gen *eg9* cải biến trong vector đã được kiểm tra lại bằng giải trình tự. Kết quả cho thấy gen đã được đưa vào đúng khung đọc trong vector pET22b(+).

```

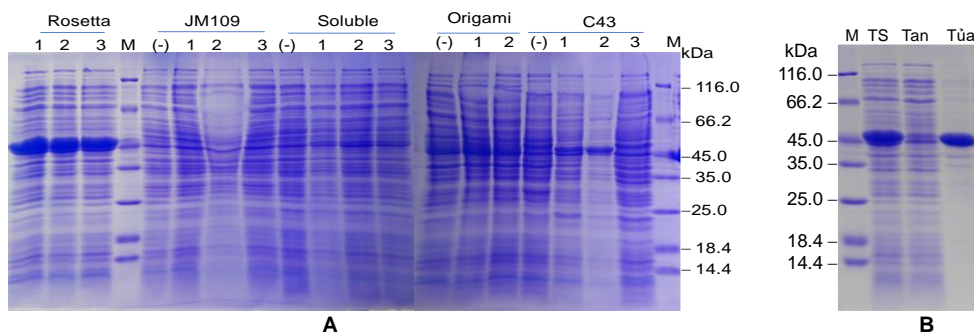
GEN GỐC   GCA ACC GCT ACT AAA CCT GCC GCG CAG CCC GCT GGC CGG GCG GCT GCC GCC GCC TCT GCC GCG CCG GCT TGC GAC GAC TGG CGC AGC TAT 90
CẢI BIẾN ..G ... ..G ..C ..G ..G ..T ... ..A ..G ..G ..T ..T ... ..G ..G ..G ..G AGC ..G ... ..G ... ..T ... ..T ... ..C
GEN GỐC   CGC AAC TTC GTC ACC CGC TTC GTG CAG CAG GAC GGC CGC GTC GTC GAC TTC TCG ACG CCG CAG CAG CAG ACC ACG TCC GAA GGG CAG TCG 180
CẢI BIẾN ..T ... ..G ... ..G ... ..T ..T ..T ... ..A ... ..T ..T ..G ..T ..T ... ..G ... ..G ... ..G ... ..G ... ..G ... ..G
GEN GỐC   TAT GCG ATG TTT TTC GCG CTG GTC GCC AAC GAC CGC GCG AGC TTC GAG CCG CTG CTG CAC TGG ACC CGC GCG AAT CTG TCC GCA GGC CGC 270
CẢI BIẾN ... ..T ... ..C ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   TTC GAT GCG AAC GAC CTG AAG CTG CCC GCA TGG CAG TGG GGC CGC AAG CCC GAC GGC TCG TAC GGC GTG CTC GAT CCG AAC TCG GCG TCC 360
CẢI BIẾN ..T ... ..C ... ..T ... ..G ... ..G ... ..G ... ..G ... ..T ..T ..T ..A ..G ..T ..T AGC ... ..T ... ..G ... ..G ... ..G
GEN GỐC   GAT GCC GAT CTG TGG ATC GCG TAC GAC CTG TTC GAG GCC GGC CGG CTA TGG CAC GAG CCG TCG TAC ACG CAG CTC GCG TGG GCG CTG ATC 450
CẢI BIẾN ..C ... ..G ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   ACG CAG ATC GCG AAG CAG GAA GTG AGC ACG CTC GAC GGC CTG GGG CCG ATG CTG TTG CCG GGG CCG CAA GGG TTT CCG AAC GGC GGC ACG 540
CẢI BIẾN ..C ... ..A ... ..T ... ..A ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   ACG GCG CTG AAC CCC AGC TAT CTG CCA TTG CCG CTG CTG GCG GCG CTC GCG GCC GAA GCG CCG GGT CCC TGG GCG AGC ATC GCG GCG 630
CẢI BIẾN ..C ..T ... ..G ... ..C ... ..G ..C ... ..T ... ..T ... ..G ... ..G ... ..G ... ..G ... ..G ... ..G ... ..G ... ..G
GEN GỐC   AAC GCG TAT ACG CTC GTC AGG ACC ACC GCG CCG GCG GGC TTC GCG CCG GAC TGG GCC GCG TGG GCG GAC GGC CGC TTT ATC GTC GAC CCG 720
CẢI BIẾN ... ..T ... ..C ..T ... ..G ..G ..C ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   AAG AAC GGC GAC GTC GGC AGT TAC GAC GCG ATT GCG GTC TAT CTG TGG GCT GGC ATC ACC GCG CCC GCC GAC CCG CTG GCC AAG CCG TGG 810
CẢI BIẾN ... ..T ... ..T ..T ..T ... ..C ... ..T ..T ... ..G ..T ..T ... ..G ..G ... ..G ... ..G ... ..G ... ..G ... ..A
GEN GỐC   CTC GCG GCG CTG GAC GGC ATG CGT GCG CAG ATC GCG CAA AGC GGC TAT CCG CCG GAG CGC GTC GCG ACC ACC AGC GGC GCG GCG CAG GGC 900
CẢI BIẾN ..G ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   GAA GCG CCG CTC GCG TTC TGG GCG GCG CTG CCG TAT TTC CCG GCG CTG AAC GAC ACG GCG GCG AGT CTC GCG CAA ACG CAT CTG 990
CẢI BIẾN ... ..T ... ..G ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   GCC GCG CTC GAC GCG GCG TCT GCG AGC GCC GCG GCG GCG GCA GGC GCG GCC GCG CCG CTG GGC ACA TCC GCG GGC GCA CCG GGC GCC ACG 1080
CẢI BIẾN ..G ... ..G ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   TCC GCA GGC ACA TCT CTA TCG ATA GGC AAG TCT GCA GAC ACC CCC GGC GCG ACT CCC GCA CCG GTC TAC GAC GAA GTG CTG ATG CTG 1170
CẢI BIẾN ..G ... ..G ... ..C AGC ..G AGC ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T ... ..T
GEN GỐC   TTC GCG ACC GGC TTC GCG GAC GGC CCG TAC CAT TTC GAC GAC ACC GGC CCG CTC GTG CCG CGT TGG GAG AAC TCA TGC CAA ACC GCC CAC 1260
CẢI BIẾN ... ..T ... ..T ..G ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T ..T
GEN GỐC   GCG GCG TCG 1269
CẢI BIẾN ... ..T AGC
    
```

Hình 1. Trình tự gen *eg9* gốc và trình tự cải biến mã hóa cho endoglucanase GH8.

Biểu hiện enzyme EG9 ở các chủng *E. coli* khác nhau

Plasmid pET22-*eg9* đã được biến nạp vào 5 chủng *E. coli* để tiến hành biểu hiện gen. Kết quả trên Hình 2 cho thấy so với các dòng đối chứng không được cảm ứng, EG9 có kích thước khoảng 49 kDa đã được biểu hiện rõ nhất ở chủng Rosetta và chủng C43, biểu hiện thấp trong các chủng Origami, SoluBL21. Chủng *E. coli* Rosetta là chủng được ưa chuộng để biểu hiện các loại cellulase tái tổ hợp như

cellobiohydrolase (*PEcbh*) từ *P. eryngii* (Romruen, Bangyeekhun, 2016), endoglucanase (JqCel5A) từ *Jonesia quinghaiensis* (Lin *et al.*, 2016), endoglucanase (Cel5A), endo, exoglucanase (Cel9A) kiếm từ *Cellulomonas bogoriensis* (Li *et al.*, 2020). Trong đó cellobiohydrolase biểu hiện trong chủng Rosetta có hoạt tính tốt hơn chủng BL21 (DE3) (Romruen, Bangyeekhun, 2016).



Hình 2. Phân tích protein tổng số của 5 chủng biểu hiện Rosetta, JM109, SoluBL21, Origami, C43 mang pET22-*eg9* (A) và protein pha tổng số (TS), pha tan, pha tủa được biểu hiện trong cùng Rosetta (B). M: Protein chuẩn (Fermentas); 1-3: Các dòng khác nhau trong 1 chủng biểu hiện; (-) dòng 1 không được cảm ứng sinh tổng hợp EG9 bằng IPTG.

Enzyme EG9 được biểu hiện ở chủng Rosetta chủ yếu dạng không tan (đường chạy số 3) và ít thấy xuất hiện bằng protein ở pha tan (đường chạy 2). Ngoài sử dụng nước để siêu âm phá tế bào, chúng tôi cũng đã sử dụng đệm PBS có các pH khác nhau (4, 5, 6, 7, 8, 9), đệm Tris có pH khác nhau để siêu âm hòa tan EG9 nhưng khả năng tan của EG9 không được cải thiện.

Thông thường, enzyme được biểu hiện ở pha protein tan thường có cấu trúc bậc 3 bậc 4 chính xác nên có hoạt tính sinh học. Để tăng khả năng biểu hiện protein ở dạng tan, có hoạt tính sinh học, chúng tôi tiến hành khảo sát các điều kiện biểu hiện như môi trường nuôi cấy, nồng độ chất cảm ứng, thời điểm cảm ứng.

Ảnh hưởng của thời điểm cảm ứng đến sự biểu hiện protein EG9

Thời điểm cảm ứng đóng vai trò quan trọng trong biểu hiện nhiều loại protein tái tổ hợp. Trong nghiên cứu này, chúng tôi đã tiến hành nuôi cấy cảm ứng chủng Rosetta mang gen *eg9* ở các thời điểm OD₆₀₀ đạt 0,2; 0,3; 0,5; 0,6; 0,8; 1,0 và nuôi cấy lắc 170 vòng/phút ở 30°C trong 4 giờ. Kết quả cho thấy sinh khối mẫu không cảm ứng tính theo OD₆₀₀ đạt 2,2 trong khi đó tất cả mẫu cảm ứng ở các thời điểm khác nhau giao động từ 1,3 đến 1,5 và không có sự sai khác có ý nghĩa thống kê. Ở tất cả các thời điểm cảm ứng, tế bào đều sinh tổng hợp enzyme EG9 với hàm lượng tương đương nhau được đánh giá tương đối bằng điện di SDS-PAGE và hầu hết enzyme tái tổ hợp đều nằm ở pha không tan (Kết quả không được trình bày). Như vậy có thể nói EG9 được sinh ra không ảnh hưởng tới sự sinh trưởng, phát triển của tế bào và thời điểm cảm ứng tế bào sinh tổng hợp EG9 là khi OD₆₀₀ dưới 1.

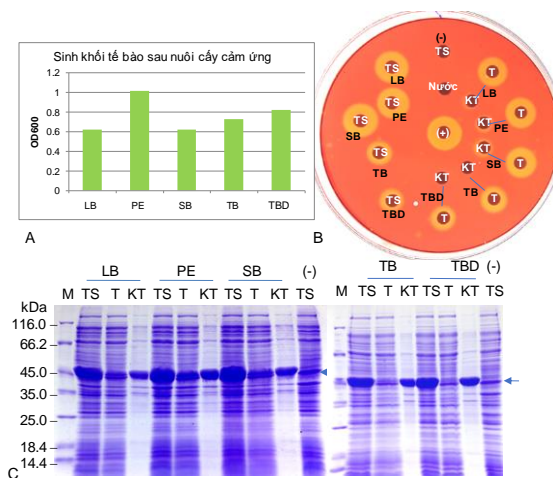
Nghiên cứu lựa chọn môi trường thích hợp cho sinh tổng hợp EG9

Trong nghiên cứu này, EG9 được nghiên cứu biểu hiện trên 5 loại môi trường với các thành phần khác nhau (LB, PE, SB, TB và TBD). Thí nghiệm được tiến hành trong cùng một điều kiện cảm ứng với 0,5 mM IPTG, lắc 170 vòng/phút, ở 25°C trong thời gian 4 giờ.

Kết quả (Hình 3A) cho thấy môi trường PE cho sinh khối tế bào cao nhất sau đến TBD rồi đến môi trường TB. Môi trường PE là môi trường có chứa nguồn N phong phú trong đó khá nghèo nguồn carbon. Như vậy có thể thấy chủng Rosetta mang gen mã hóa *eg9* ít sử dụng đường cho sinh trưởng.

Kết quả điện di protein trên hình 3 cho thấy protein EG9 (~ 49 kDa) biểu hiện ở cả 5 môi trường nghiên cứu. Đặc biệt EG9 đã được nhìn thấy biểu hiện ở pha tan trong môi trường LB, PE và SB. Mặc dù ở môi trường TB và TB cải biến cũng xuất hiện các băng vạch đậm tương đương kích thước EG9 nhưng chủ yếu lại nằm ở pha không tan (pha tủa).

Kết quả kiểm tra hoạt tính thủy phân CMC trên đĩa thạch cho thấy, các dòng đối chứng không có khả năng phân hủy CMC trong khi đó chủng tái tổ hợp mang gen *eg9* được nuôi cấy ở cả 5 môi trường đều thể hiện hoạt tính thủy phân CMC tạo vòng sáng xung quanh khoanh giấy thấm dịch. Trong đó ở pha tổng số, hoạt tính EG9 thể hiện mạnh nhất khi tế bào được nuôi cấy trên môi trường LB, PE và SB. Hầu hết các protein biểu hiện ở pha không tan không có hoạt tính sinh học hoặc có nhưng rất yếu. Chúng tỏ rất có thể khi được biểu hiện ở dạng không tan, cấu trúc bậc 3 của enzyme không được hoàn thiện dẫn đến enzyme không có hoạt tính sinh học. Trong khi đó, tất cả pha tan của tế bào ở các môi trường nuôi cấy đều có hoạt tính tốt. Thậm chí, một phần rất nhỏ EG9 được biểu hiện trong TB và TBD cũng thể hiện hoạt tính sinh học.



Hình 3. Phân tích kết quả nuôi cấy chủng *E. coli* Rosetta biểu hiện EG9 ở các môi trường khác nhau. A. Sinh khối tế bào tại thời điểm thu mẫu; B. Kiểm tra hoạt tính endoglucanase ở các pha protein tổng số (TS), tan (T), không tan (KT); C: Phân tích protein được tổng hợp ở các môi trường khác nhau bằng SDS-PAGE. (-) Dòng tế bào Rosetta mang pET22b(+); (+): cellulase (Sigma); M: Protein chuẩn (Fermentas). Mũi tên chỉ vị trí của EG9.

Môi trường LB là môi trường được sử dụng khá phổ biến để biểu hiện thành công nhiều loại enzyme trong *E. coli* trong đó có cellulase (Li *et al.*, 2020, Lin *et al.*, 2016, Romruen, Bangyeekhun, 2016) và ít

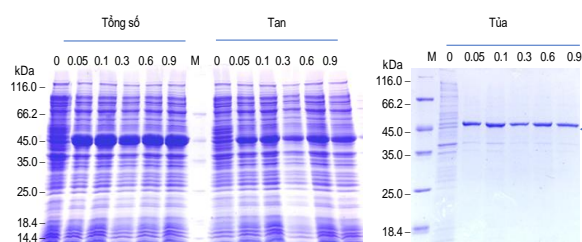
nghiên cứu khảo sát môi trường tăng tổng hợp enzyme này trong *E. coli* đặc biệt là đối với môi trường đơn giản như PE. Nghiên cứu của Fu và đồng tác giả (2020) đã khảo sát nhiều loại môi trường khác nhau bao gồm MX, LBBSMG, TB-GN, LBBM, LB, LBBNM, TB, LBBNG, SOB để biểu hiện feruloyl esterase từ *Burkholderia pyrrocinia*, tác giả cũng nhận được enzyme biểu hiện tốt ở môi trường LB và TB (Fu *et al.*, 2020).

Nhìn tổng thể có thể thấy trong nghiên cứu này môi trường PE vừa cho sinh khối tế bào cao, vừa giúp enzyme được biểu hiện ở pha tan ở 25°C và enzyme lại thể hiện hoạt tính tốt. Do vậy, trong thí nghiệm này chúng tôi đã lựa chọn môi trường PE cho biểu hiện enzyme EG9 và lựa chọn nhiệt độ 25°C để biểu hiện gen.

Ảnh hưởng của nồng độ IPTG đến sự biểu hiện protein EG9

Nồng độ IPTG được sử dụng cho cảm ứng sinh tổng hợp EG9 được khảo sát là 0,05; 0,1; 0,3; 0,6 và 0,9 mM. Điều kiện nuôi cấy sinh tổng hợp EG9 là trong môi trường PE với nhiệt độ cảm ứng 25°C, lắc 170 vòng/phút trong 5 giờ.

Kết quả so sánh mật độ tế bào OD₆₀₀ từ các mẫu cảm ứng cho thấy mẫu không được cảm ứng cho sinh khối cao nhất. Tuy nhiên, khi được cảm ứng ở các nồng độ IPTG tăng dần, sinh khối tế bào thu được có thấp hơn so với đối chứng nhưng sự sai khác này không đáng kể (Kết quả không trình bày). Do vậy có thể nói việc cảm ứng sinh tổng hợp EG9 không gây ảnh hưởng tới sự sinh trưởng của chúng.



Hình 4. Phân tích kết quả nuôi cấy chủng *E. coli* Rosetta biểu hiện EG9 ở các nồng độ chất cảm ứng khác nhau. M: Protein chuẩn (Fermentas)

Kết quả kiểm tra sự biểu hiện của protein EG9 bằng phương pháp điện di SDS-PAGE cho thấy ở các mẫu protein tổng số và protein không tan thu nhận từ tế bào *E. coli* Rosetta1 mang pET22-*eg9* được cảm ứng biểu hiện bởi IPTG đã sinh tổng hợp

protein EG9 ở ngay cả nồng độ chất cảm ứng IPTG thấp 0,05 mM. Feruloyl esterase của *Burkholderia pyrrocinia* cũng được biểu hiện ở nồng độ IPTG 0,5 mM (Fu *et al.*, 2020). Nhìn trên bản điện di có thể thấy lượng enzyme tái tổ hợp được biểu hiện ở các nồng độ IPTG khác nhau là không khác nhau. Trong đó, khoảng một nửa EG9 được biểu hiện ở pha tan. Trong nghiên cứu này, chúng tôi chọn nồng độ chất cảm ứng IPTG 0,1 mM để biểu hiện EG9 ở các thí nghiệm sau này. Enzyme cũng được biểu hiện ở 20°C nhưng lượng protein pha tan thu được không được cải thiện. Vì vậy sau khi lựa chọn các điều kiện cho biểu hiện gen, chúng tôi thấy EG9 được biểu hiện tốt nhất với khoảng 50% enzyme ở dạng tan khi tế bào *E. coli* Rosetta mang gen *eg9* được nuôi cấy trong môi trường PE, với nồng độ chất cảm ứng IPTG 0,1 mM, ở 25°C.

KẾT LUẬN

Gen *eg9* từ dữ liệu giải trình tự metagenome của vi khuẩn trong mùn xung quanh khu nấm mục trắng thủy phân gỗ mã hóa endoglucanase GH8 đã được biểu hiện thành công trong *E. coli*. Trong môi trường PE có 0,1 mM IPTG, tế bào *E. coli* Rosetta mang gen *eg9* đã sinh tổng hợp khoảng 50% EG9 ở dạng tan, có hoạt tính sinh học thủy cơ chất CMC.

Lời cảm ơn: Công trình được thực hiện bằng nguồn kinh phí của đề tài Nghị định thư Việt-Đức mã số NĐT.50.GER/18 do GS. TS. Trương Nam Hải chủ nhiệm và trang thiết bị của Phòng thí nghiệm trọng điểm Công nghệ gen, viện Công nghệ sinh học, viện Hàn lâm Khoa học và Công nghệ Việt Nam.

TÀI LIỆU THAM KHẢO

- Bhat MK, Bhat S (1997) Cellulose degrading enzymes and their potential industrial applications. *Biotechnol Adv* 15: 583–620.
- Davies GJ, Dodson GG, Hubbard RE, Tolley SP, Dauter Z, Wilson KS, Hjort C, Mikkelsen JM, Rasmussen G, Schülein M (1993) Structure and function of endoglucanase V. *Nature* 365: 362–364.
- Davies GJ, Henrissat B (2002) Structural enzymology of carbohydrate-active enzymes: implications for the post-genomic era. *Biochem Soc Trans* 30: 291–297.
- Fu Z, Fan G, Zhu Y, Teng C, Li H, Liu Q, Yang R, Li X (2020) Soluble expression of a novel feruloyl esterase from *Burkholderia pyrrocinia* B1213 in *Escherichia coli* and optimization of production conditions. *Biotechnol Biotechnol Equip* 34: 732–746.

- Karnaouri A, Matsakas L, Krikigianni E, Rova U, Christakopoulos P (2019) Valorization of waste forest biomass toward the production of cello-oligosaccharides with potential prebiotic activity by utilizing customized enzyme cocktails. *Biotechnol Biofuels* 12.
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.
- Li F, Dong J, Lv X, Wen Y, Chen S (2020) Recombinant expression and characterization of two glycoside hydrolases from extreme alkaliphilic bacterium *Cellulomonas bogoriensis* 69B4 T. *AMB Express* 10: 44.
- Lin L, Fu C, Huang W (2016) Improving the activity of the endoglucanase, Cel8M from *Escherichia coli* by error-prone PCR. *Enzyme Microb Technol* 86: 52–58.
- Lin L, Liu X, Zhou Y, Guan L, Jiajia H, Huang W (2016) A novel pH-stable, endoglucanase (JqCel5A) isolated from a salt-lake microorganism, *Jonesia quinghaiensis*. *Electron J Biotechnol* 19: 56–62.
- Ontañón OM, Ghio S, Marrero Díaz de Villegas R, Garrido MM, Talia PM, Fehér C, Campos E (2019) A thermostable GH8 endoglucanase of *Enterobacter* sp. R1 is suitable for β -glucan deconstruction. *Food Chem* 298: 124999.
- Romruen U, Bangyeekhun E (2016) Cloning and expression of the cellulase gene from the King oyster mushroom, *Pleurotus eryngii*. *Sci Eng Health Stud* 22–30.
- Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual*. CSHL Press.

EXPRESSION OF *eg9* GENE ENCODING ENDOGLUCANASE GH8 DERIVED FROM METAGENOMIC DNA DATA OF BACTERIA IN HUMUS AROUND WHITE ROT FUNGI DEGRADING WOODS, IN *ESCHERICHIA COLI*

Do Thi Huyen^{1,2}, Le Thu Hoai^{1,3}, Nguyen Hai Dang¹, Nguyen Thi Quy¹, Truong Nam Hai^{1,2}

¹Institute of Biotechnology (IBT), Vietnam Academy of Science and Technology (VAST)

²Graduate University of Science and Technology (GUST), Vietnam Academy of Science and Technology (VAST)

³Vietnam National University of Agriculture

SUMMARY

Endoglucanase is an important enzyme participating in the hydrolysis of cellulose into oligosaccharides or glucoses can be used in a variety of industrial fields. A gene (code GL0183420 designated as *eg9*) of 1398 bp encoding for endoglucanase family GH8 was identified from metagenomic DNA data of bacteria in the humus surrounding wood hydrolyzed by white rot fungi. The 126 terminal nucleotides at 5' terminal of the gene encode for a signal peptide in gram-negative bacteria and the next 1269 nucleotides encode the endoglucanase GH8. In this study, the *eg9* gene (lack of sequence encoding the signal peptid) coding for the mature endoglucanase was codon optimized for expression in *E. coli*, artificial synthesized, then inserted into pET22b(+) for gene expression in *E. coli* strains. Through a survey of 5 expressive strains, EG9 was overexpressed in Rosetta and C43 strains. However, at 30°C in LB medium, most of EG9 were expressed in the insoluble fraction, in which Rosetta was the best strain for synthesis of EG9. The enzyme EG9 can be overexpressed in Rosetta strain at a large range of OD₆₀₀ from 0.2 to 1. After reducing the culture temperature to 25°C, 50% of EG9 was expressed in soluble fraction and exhibited good hydrolysis activity in CMC substrate on agar plate. The investigation of media showed that PE was the most suitable medium for the recombinant Rosetta strain grow and synthesize endoglucanase EG9. EG9 was well expressed at various IPTG concentrations from 0.05 to 0.9 mM. These results offer great potential for production of recombinant EG9 for enzyme properties research.

Keywords: *eg9*, endoglucanase gh8, *Escherichia coli*, gene expression, humus.