

PHÂN TÍCH TRÌNH TỰ GEN MÃ HÓA LUMBROKINASE VÀ TỐI ƯU CODON CHO HỆ THỐNG BIỂU HIỆN *E. COLI*

Đỗ Thị Tuyên, Lý Thị Bích Thủy, Chu Thị Hoa, Vũ Thị Bích Ngọc, **Quyền Đình Thi**

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

Ngày nhận bài: 28.8.2014

Ngày nhận đăng: 30.3.2015

TÓM TẮT

Biến chứng tắc nghẽn ở bệnh tim mạch là nguyên nhân chính gây tử vong, tai biến và do đó, sự tan cục huyết có thể ảnh hưởng tích cực đến kết quả điều trị các bệnh nguy hiểm tới tính mạng như nhồi máu cơ tim, tắc nghẽn mạch máu não, và biến chứng tắc nghẽn mạch máu. Các yếu tố tan huyết là những thể hoạt hóa plasminogen chuyển hóa plasminogen, proenzyme không hoạt hóa của hệ thống tan sợi fibrin trong máu thành enzyme plasmin dạng hoạt động. Plasmin làm tan fibrin của cục máu đông, nhưng cũng có thể phân hủy các thành phần bình thường của hệ thống cầm máu và cảm ứng trạng thái ly giải. Hiện nay, các thể tan cục huyết đã được phê chuẩn để sử dụng lâm sàng hoặc đang nghiên cứu lâm sàng ở các bệnh nhân nhồi máu cơ tim cấp đó là streptokinase, urokinase, rt-PA, APSAC, rscu-PA, staphylokinase (Sak) và lumbrokinase. Việc nghiên cứu sản xuất lumbrokinase tái tổ hợp từ *Eisenia fetida* để sử dụng làm thể tan cục huyết và chống tắc nghẽn là cần thiết đối với Việt Nam. Trong nghiên cứu này, chúng tôi tiến hành tối ưu codon gene mã hóa lumbrokinase LK1 từ loài giun *Eisenia fetida* được lấy từ ngân hàng Genbank với mã số AF433650 cho hệ biểu hiện *E. coli*. Kết quả đã chỉ ra, sau khi tối ưu, chỉ số CAI đã được cải thiện, đạt giá trị 0,86, phù hợp với hệ biểu hiện trên *E. coli*. Các codon đều có tần số sử dụng ở mức cao (trên 50%) và không còn codon hiếm. Gene mã hóa lumbrokinase sau khi được tối ưu hóa codon có chứa 747 bp, mã hóa cho 238 amino acid. Gene đã được tạo dòng trong vector pET22b(+) và biểu hiện trong *E. coli* BL21. Sau 6 giờ cảm ứng IPTG nồng độ 1mM, xuất hiện băng protein chiếm đến hơn 40% protein tổng số. Protein tinh sạch có kích thước khoảng 35 kDa. Hoạt tính thủy phân đĩa fibrin đạt 20 U/mg protein.

Từ khóa: Biểu hiện, *E. coli* BL21, Fibrinolytic, lumbrokinase, Plasminogen, tinh sạch

MỞ ĐẦU

Trong sản xuất protein tái tổ hợp, việc lựa chọn tế bào chủ đóng vai trò quan trọng. Các 'nhà máy' vi khuẩn thường được lựa chọn do chi phí thấp, sinh khối lớn và tốc độ sản xuất nhanh. *E. coli* là loài vật chủ vi khuẩn gram âm được ưa chuộng nhất trong sản xuất protein tái tổ hợp. Ngoài những lý do trên thì *E. coli* được lựa chọn là vì những thao tác kỹ thuật với *E. coli* khá dễ dàng và chuẩn hóa ở mọi phòng thí nghiệm. Do đó, *E. coli* được lựa chọn để nghiên cứu sản xuất LK tái tổ hợp.

Việc sản xuất protein tái tổ hợp thường cho năng suất thấp khi protein biểu hiện có nguồn gốc khác với tế bào chủ. Một trong những nguyên nhân quan trọng nhất là sự không tương thích về mã sử dụng của gene biểu hiện với khả năng mã của tế bào chủ. Mỗi amino acid có thể được mã hóa bởi nhiều hơn một codon và đối với mỗi loài sinh vật, hệ thống di truyền thường ưa thích một số codon này hơn những codon còn lại

trong tất cả 61 codon có khả năng mã hóa. Trong từng tế bào, số lượng của mỗi thành phần RNA vận chuyển có liên quan chặt chẽ đến khả năng mã của mRNA và có tính đặc trưng loài (Gustafsson *et al.*, 2004). Khi muốn sản xuất lượng lớn protein trong *E. coli* từ một mRNA khác biệt về khả năng mã của vật chủ, lượng sản phẩm protein tái tổ hợp được tổng hợp thường không lớn như những protein khác được biểu hiện trong cùng hệ thống do tế bào không đủ một hoặc một số tRNA nhất định. Việc thiếu tRNA này có thể làm dừng quá trình dịch mã giữa chừng hoặc làm lệch khung đọc dẫn đến thay đổi trình tự amino acid trong sản phẩm protein. Tối ưu hóa codon và tổng hợp gene cấu trúc sẽ nâng cao hiệu suất biểu hiện protein (Tiwari *et al.*, 2010 ; Liu *et al.*, 2012; Retnoningrum *et al.*, 2012; Tanaka *et al.*, 2012). Trong nghiên cứu này chúng tôi tiến hành phân tích trình tự gen mã hóa lumbrokinase và tối ưu codon cho hệ thống biểu hiện *E. coli*, từ đó làm nguyên liệu để nhân dòng cũng như thiết kế tạo chủng tái tổ hợp biểu hiện lumbrokinase trong *E. coli* đạt năng suất cao.

PHƯƠNG PHÁP NGHIÊN CỨU

Nguyên liệu

Trình tự lumbrokinase từ *E. fetida* được lấy từ ngân hàng Genbank với mã số AF433650 (Sugimoto, Nakajima 2001; Ge et al., 2005).

Chủng *E. coli* BL21 mang vector tái tổ hợp pELK thu nhận từ phòng Công nghệ sinh học enzyme, Viện Công nghệ sinh học.

Tối ưu hóa codon

Tối ưu hóa codon được thực hiện bởi Genscript (Mỹ), sử dụng chương trình OptimumGene™ algorithm để cân bằng tỷ lệ sử dụng codon phù hợp cho chủng biểu hiện và điều chỉnh hàm lượng GC. Quá trình cải biến này làm thay đổi codon theo hướng lựa chọn codon được chủng biểu hiện ưu tiên sử dụng nhất nhưng không làm thay đổi amino acid (Sharp, Li 1987; Jansen et al., 2003).

Phân tích sự phù hợp của gen tối ưu với chủng biểu hiện

Phân tích sự phù hợp của trình tự gen mã hóa lumbrokinase từ loài giun *E. foetida* với hệ biểu hiện *E. coli* dựa vào công cụ phân tích codon hiếm (Rare Codon Analysis Tool) của hãng Genscript. Sự phân tích dựa trên các chỉ tiêu chính sau: CAI (codon adaptation index - chỉ số phù hợp codon), hàm lượng GC, sự phân bố tỷ lệ phần trăm các codon. Trình tự gen sau khi được tối ưu hóa sẽ được đặt tổng hợp ở hãng Genscript.

Biểu hiện

Dòng *E. coli* BL21 mang vector tái tổ hợp pELK sau khi hoạt hóa được nuôi cấy qua đêm ở 37°C trong 5 ml môi trường LB có chứa 100 µg/ml ampicillin. Từ dịch nuôi cấy qua đêm cấy chuyển 1 ml sang bình 500 ml chứa 100 ml LB lỏng có bổ sung 100 µg/ml ampicillin, nuôi lắc 37°C trong 3 giờ để OD₆₀₀ đạt 0,5- 0,8. Bổ sung chất cảm ứng IPTG đến nồng độ cuối cùng là 1 mM và tiếp tục nuôi lắc 200 rpm ở 37°C để cảm ứng tế bào tổng hợp protein tái tổ hợp. Tế bào được thu sau 6 giờ nuôi cấy cảm ứng bằng ly tâm 8000 vòng/phút trong 10 phút

Tinh sạch LK tái tổ hợp

rLK được tinh sạch qua 2 bước: (1) tế bào được phá siêu âm trong đệm Tris-HCl, (2) protein dạng thể vùi được biến tính và tinh sạch bằng cột sắc ký ái lực Ni²⁺ của Invitrogen. Dịch biểu hiện rLK được ly tâm 8000 vòng/phút trong 10 phút để thu tế bào. Tế

bào sau đó được rửa bằng đệm Tris-HCl và phá siêu âm, ly tâm thu cặn. Bước này được thực hiện lặp lại 4 lần. Cặn thu được sau khi rửa được hòa tan trong dung dịch guanidine 6M và tinh sạch theo phương pháp Hybrid như hướng dẫn của hệ thống tinh sạch ProBond™ Nickel-Chelating Resin (Invitrogen).

Hoạt tính thủy phân fibrin trên đĩa thạch

Hoạt tính thủy phân fibrin được xác định dựa vào đường chuẩn plasmin: mỗi đĩa petri đường kính 10 cm được bổ sung 15 ml dung dịch fibrinogen tinh khiết trong NaCl 0,9% (nồng độ 0,2%) được ngưng kết bằng 50 µl dung dịch thrombin (đạt nồng độ cuối cùng là 25 UI đĩa). Chú ý đặt đĩa cho phẳng và đổ dung dịch cho thật đều. Để ở nhiệt độ phòng trong 1 giờ để fibrin được đông tụ hoàn toàn. Hút 10 µl dung dịch enzyme cách đều nhau lên bề mặt đĩa fibrin, ủ ở 37°C. Khi đông tụ, đĩa fibrin có màu trắng đục, do đó các vùng fibrin bị thủy phân trở nên trong suốt sẽ dễ dàng được quan sát thấy. Sau 5 giờ, đo đường kính và tính diện tích của vòng phản ứng (yêu cầu vòng phản ứng phải tròn). Mỗi mẫu được lặp lại ít nhất 3 lần để lấy giá trị trung bình. Hoạt tính thủy phân fibrin được xác định dựa vào đường chuẩn plasmin (Asturp, Mullertz, 1952).

Xử lý số liệu

Chương trình Edit-SeqDNASTar và BioEdit- Seq được dùng để phân tích, so sánh trình tự nucleotide và amino acid tương ứng.

KẾT QUẢ VÀ THẢO LUẬN

Phân tích mức độ phù hợp của gen mã hóa lumbrokinase

Kết quả phân tích mức độ phù hợp của trình tự gen *lk1* so với hệ biểu hiện *E. coli* cho thấy: Gen *lk1* trước cải biến có chỉ số CAI ở mức thấp, CAI = 0,69; các codon hiếm có tần số sử dụng dưới 30% còn nhiều (Hình 1A). Chỉ có 42% codon phân bố ở mức 91 – 100, có tới 8% các codon phân bố ở mức 31 – 40%, vẫn có 1% codon phân bố ở mức 11-20%, đây là các codon hiếm gây ảnh hưởng đến khả năng biểu hiện của gen trong cơ thể chủ (Hình 2A).

Sau khi tối ưu, chỉ số CAI đã được cải thiện, đạt giá trị 0,86- giá trị được cho là phù hợp với hệ biểu hiện *E. coli*, các codon đều có tần số sử dụng ở mức cao (trên 50%) và không còn codon hiếm (Hình 1B).

CAI là chỉ số sử dụng bộ codon chuẩn của các gen biểu hiện cao ở một loài để đánh giá sự phù hợp

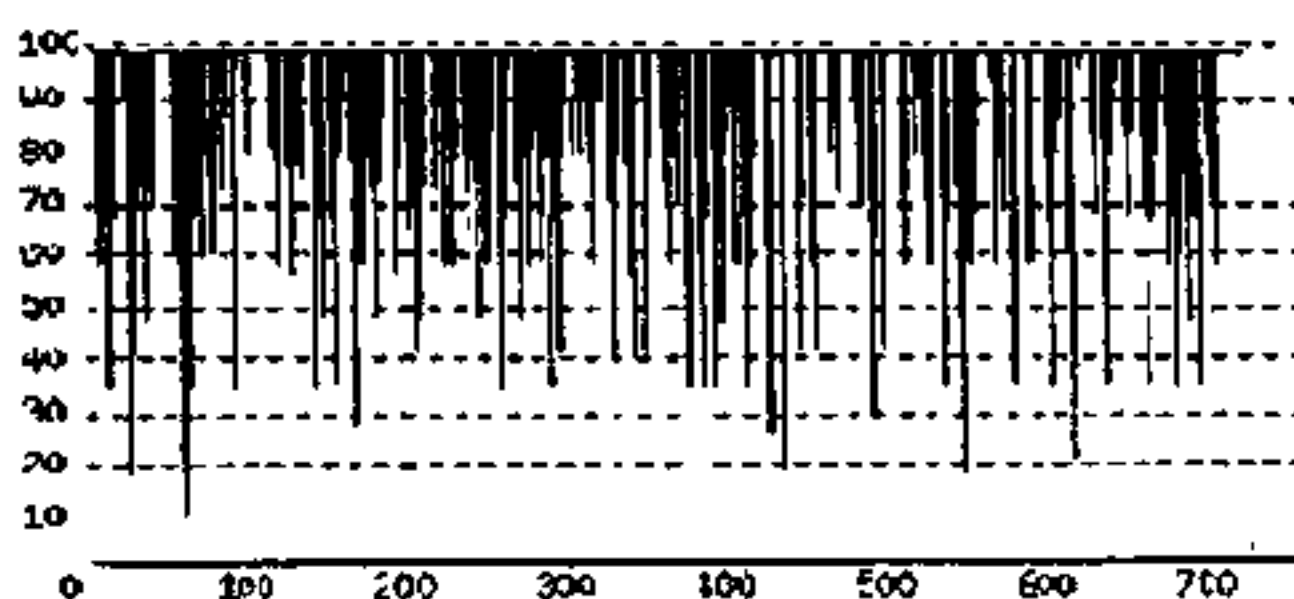
của mỗi codon của gen đang xét với loài biểu hiện đó và chỉ số phù hợp của cả một gen được tính từ tần số sử dụng của tất cả các codon trong gen đó. Chỉ số CAI đạt 1,0 được coi là lý tưởng đồng nghĩa gen đó có khả năng được biểu hiện cao ở hệ biểu hiện đang xét, chỉ số CAI càng thấp đồng nghĩa với khả năng biểu hiện gen càng thấp. Chỉ số CAI lớn hơn 0,8 được cho là phù hợp để biểu hiện trong tế bào chủ mong muốn (Sharp, Li 1987; Jansen *et al.*, 2003).

Đặc biệt kết quả phân tích trên hình 2B cho thấy: các codon phân bố ở mức 91-100 (mức tần số sử dụng cao nhất) tăng từ 42% trước cải biến lên 63%, mức thấp nhất sau khi cải biến là 51-60%. Lượng

GC phân bố đều hơn so với trước cải biến (Hình 2B).

Sự phân bố codon ở giá trị 100 là tần số sử dụng codon cao nhất trong hệ biểu hiện. Sự phân bố codon ở giá trị phần trăm càng thấp đồng nghĩa tần số sử dụng codon đó trong hệ biểu hiện càng thấp. Các codon phân bố ở các giá trị thấp hơn 30 là các codon hiếm đồng nghĩa với tác động cản trở sự biểu hiện gen. Các codon hiếm không chỉ liên quan đến mức độ biểu hiện protein thấp bởi hiếm có tRNA của hệ biểu hiện phù hợp nên gây dừng và lỗi dịch mã (Kane, 1995; Hayes *et al.*, 2002), ngoài ra còn có thể gây thay đổi khung đọc và gây kết hợp lỗi của amino acid (Kane, 1995; McNulty *et al.*, 2003).

Tần suất tương đối

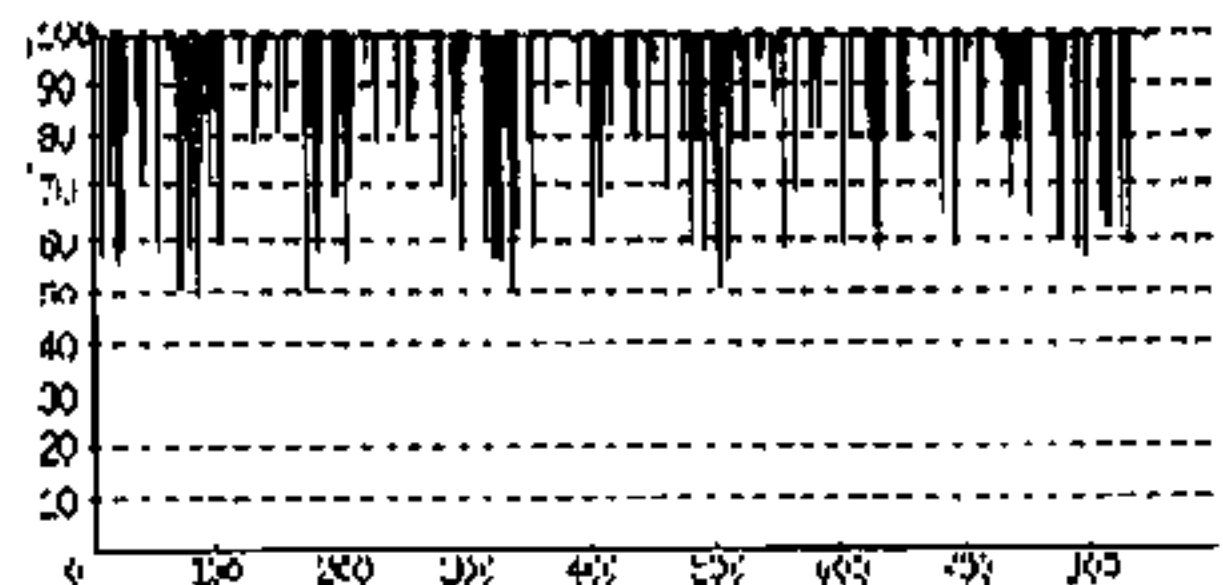


Vị trí tương đối của các codon

CAI = 0,69

A

Tần suất tương đối



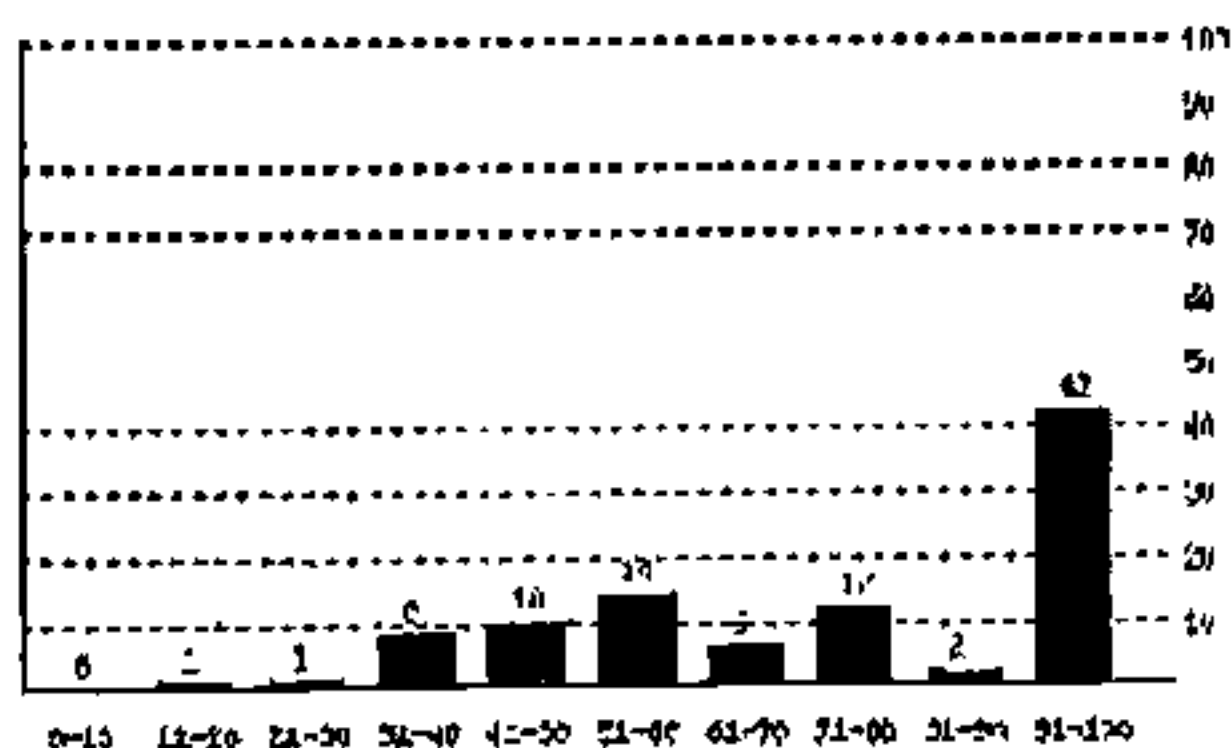
Vị trí tương đối của các codon

CAI = 0,86

B

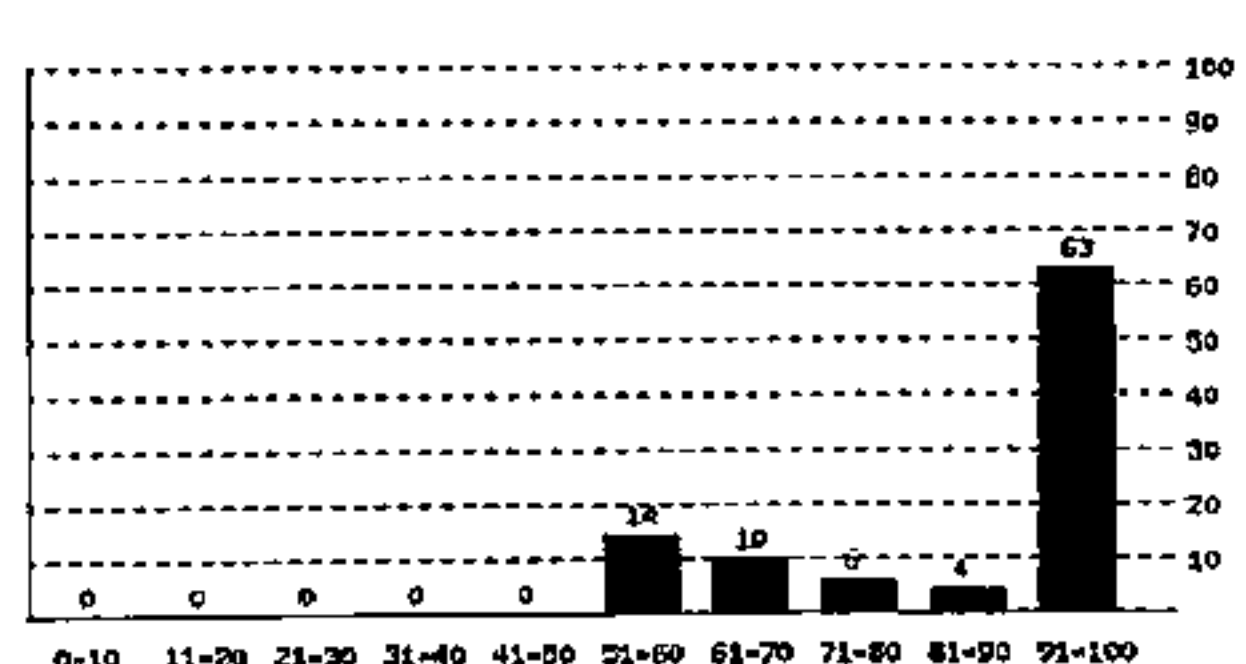
Hình 1. Phân bố tần suất sử dụng codon dọc theo chiều dài gen *lk1* trong hệ biểu hiện *E. coli*. (A) - trước khi tối ưu, (B) - sau khi tối ưu.

Tỷ lệ % các codon



A

Tỷ lệ % các codon



B

Hình 2. Sự phân bố tỷ lệ phần trăm các codon của gene *lk1* được sử dụng trong hệ biểu hiện *E.coli*. (A) - trước khi tối ưu, (B) - sau khi tối ưu.

Theo Oliveira *et al.*, 1993, hàm lượng GC cao và phân bố không đều dễ tạo cấu trúc thứ cấp, như cấu trúc kẹp tóc, cấu trúc thứ cấp được hình thành hay mất đi gần vùng mRNA không dịch mã và gần codon khởi đầu có tác động tới tới tốc độ suy thoái mRNA và ảnh hưởng tới khởi đầu dịch mã, ngoài ra còn gây khó khăn cho sự tương tác giữa mRNA với rRNA do đó ảnh hưởng tới hiệu quả biểu hiện protein ngoại lai. Khi phân tích hàm lượng GC của gen, nếu gen có các đỉnh phân bố ở ngoài khoảng 30-70% sẽ ảnh hưởng bất lợi đến hiệu quả phiên mã và

dịch mã (Oliveira *et al.*, 1993). Kết quả so sánh trước tới ưu hàm lượng GC trung bình đạt 53,74%, và sau hàm lượng đạt 51,3% (không dẫn hình), sự phân bố GC của gen trước cải biến và sau cải biến đều nằm trong khoảng cho phép biểu hiện tốt.

Như vậy, gen *lk1* sau khi cải biến đã thể hiện là phù hợp với khả năng sử dụng codon trong hệ biểu hiện *E.coli* (Hình 3). Sau khi dịch mã, trình tự amino acid suy từ gen sau cải biến và gen trước cải biến tương đồng 100%.

LK1_E.coli LK1opt_E.coli	ATTGTCGGAGGAATTGAAGCCAGACCATACGAGTTCCCATGGCAGGTGTCCGTCCGAAGG ATCGTTGGCGGTATTGAAGCCGTCCGTACGAATTTCCGTGGCAAGTTAGTGTCCGTCC ** ** ** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	AAGTCTTCCGATTCCCATTTCTGCGGAGGTAGCATCATCAACGATCGTTGGGTTGTCTGC AAAAGCTCTGATTCGATTCTGCGGCGGTAGCATCATCAACGATCGTTGGGTTGGTTTGC ** : ** ***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	GCTGCTCACTGCATGCAGGGAGAGAGCCCTGCCCTGGTTTCATTGGTCTGCTGGTGAGCAC GCGGCCCACTGTATGCAGGGCGAAAGCCCGGCACTGGTGTCTCTGGTCTGGGTGAACAT ** ** ***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	GATAGCAGCGCTGCGAGTACAGTACGTCAGACTCATGACGTTGACAGCATCTTCGTCCAC GATAGTTCCGCGAGCTTCGACCGTTCGTCAACGCATGATGTCGACAGCATTTTGTGCAC ***** : ***** : ***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	GAGGACTACAACGGAAATACCTTTGAGAACGACGTTTCTGTCAATCAAGACAGTTAACGCC GAAGATTATAACGGCAATACCTTCGAAACGACGTGTCTGTTATCAAAACGGTGAACGCG ** ** ***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	ATCGCCATCGACATCAACGATGGGCCAATCTGCGCTCCAGATCCAGCCAACGATTACGTC ATTGCCATCGATATTAATGACGGTCCGATTTCGCGCACCGGATCCGGCTAATGACTATGTC ** ***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	TACCGTAAGAGCCAGTGCTCCGGATGGGGAACATCAACTCAGGTGGAGTCTGCTGCCCC TACCGTAAATCACAATGTTCTGGGCTGGGGTACCATCAACAGCGGCGGTGTGTGCTGTCCG ***** : ***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	AACGTTCTGCGATATGTGACACTGAACGTCACAACCAACGCCTTCTGCGATGATATCTAC AATGTCCTGCGCTACGTGACCCCTGAACGTTACCACGAATGCATTTTGCATGACATTTAT ** ** ***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	AGCCCATATATACAATTACCAGEGACATGATCTGCGCCACGGACAACACCGGACAGAAC TCCCCGCTGTACACCATCACGTCAGATATGATTTCGCGCTACCGACAACACGGGCCAGAAT : ***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	GAGAGAGACTCTTGCCAGGGTGACTCTGGCGGCCCTCTGAGCGTCAAGGATGGCAACGGA GAACGCGATAGTTGTGAGGGTGACTCCGGCGGTCCGCTGTGAGTTAAGATGGCAATGGT ** :***** :***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	ATCTTCAGCCTCATTGGTATTGTGCTTGGGGAATCGGTTGCGCATCTGGCTATCCAGGA ATCTTCTCTCTGATCGGCATTGTGAGTTGGGCGATTGGTTGTGCTCCGGTTATCCGGGC ***** : ***** :***** :***** :***** :***** :*****
LK1_E.coli LK1opt_E.coli	GTCTACGCCCGCGTCGGATCCCAACTGGATGGATCACAGACATTATTACCAACAAC GTCTATGCTGCTGTGGGCGAGTCAACGGGCTGGATTACGGACATTATTACCAACAAC ***** :***** :***** :***** :***** :***** :*****

Hình 3. So sánh trình tự trước và sau cải biến của gene *lk1*, sử dụng công cụ so sánh ClustalOmega EMBL.

Biểu hiện lumbrokinase tái tổ hợp

Sau khi cấu trúc biểu hiện đã kiểm tra đúng khung đọc, vector tái tổ hợp pELK được biến nạp vào chủng biểu hiện để sinh tổng hợp protein tái tổ hợp. Để xác định protein tái tổ hợp rLK được thể hiện nội bào ở dạng hòa tan hay thể vùi, dịch lên

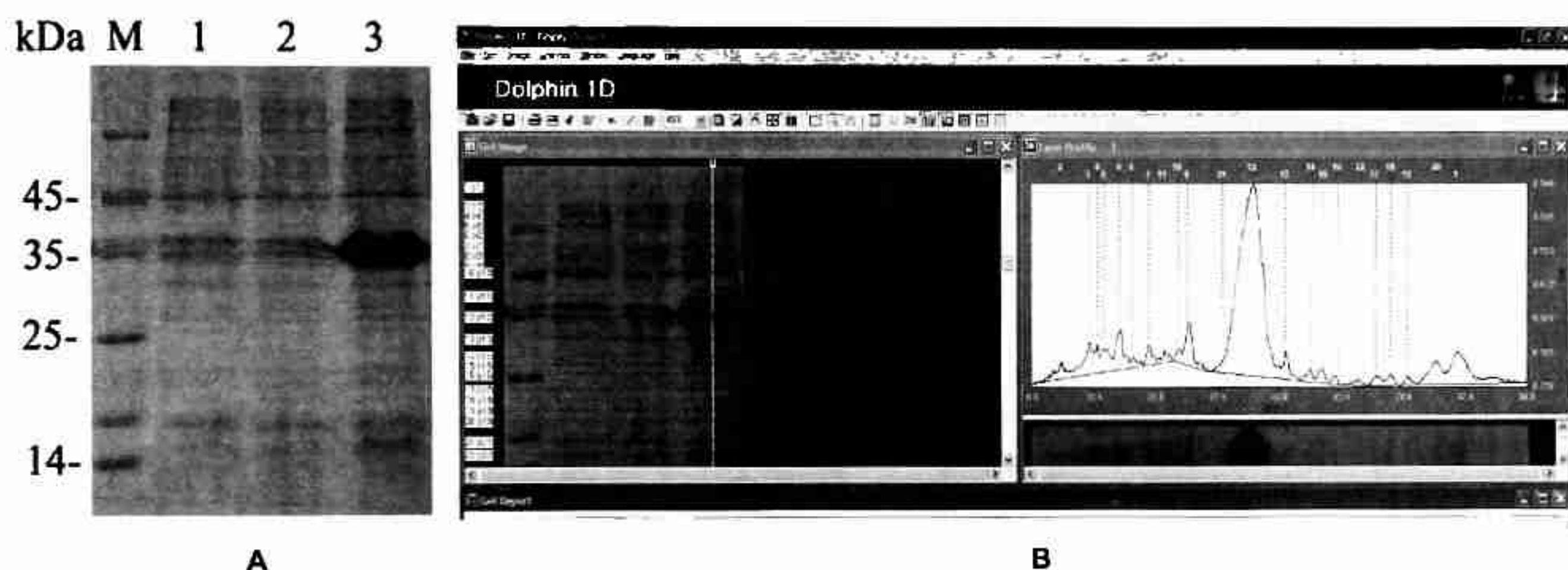
men của chủng tái tổ hợp *E. coli* BL21 sau 6 giờ cảm ứng IPTG 1 mM ly tâm loại dịch trong thu cặn tế bào. Phần cặn tế bào được xử lý bằng đệm Native Binding Buffer (NBB) (Invitrogen) bổ sung lysozyme, thu dịch trong và cặn để điện di kiểm tra trên SDS-PAGE. Kết quả trên hình 4 cho thấy, protein tái tổ hợp rLK được sinh tổng hợp chủ yếu ở

dạng thể vùi, căn tế bào mang gen thu tại thời điểm 6 giờ sau cảm ứng xuất hiện 1 băng đậm khác so với dịch nổi chùng không mang gen và dịch nổi của chùng có mang gen không được cảm ứng IPTG, kích thước khoảng 35 kDa. Điều này tương ứng với 1 số nghiên cứu về rLK đã công bố trên thế giới. Sử dụng phần mềm phân tích số liệu dolphin 1D cho thấy mức độ biểu hiện của protein tái tổ hợp rLK đạt hơn 40% protein tổng số.

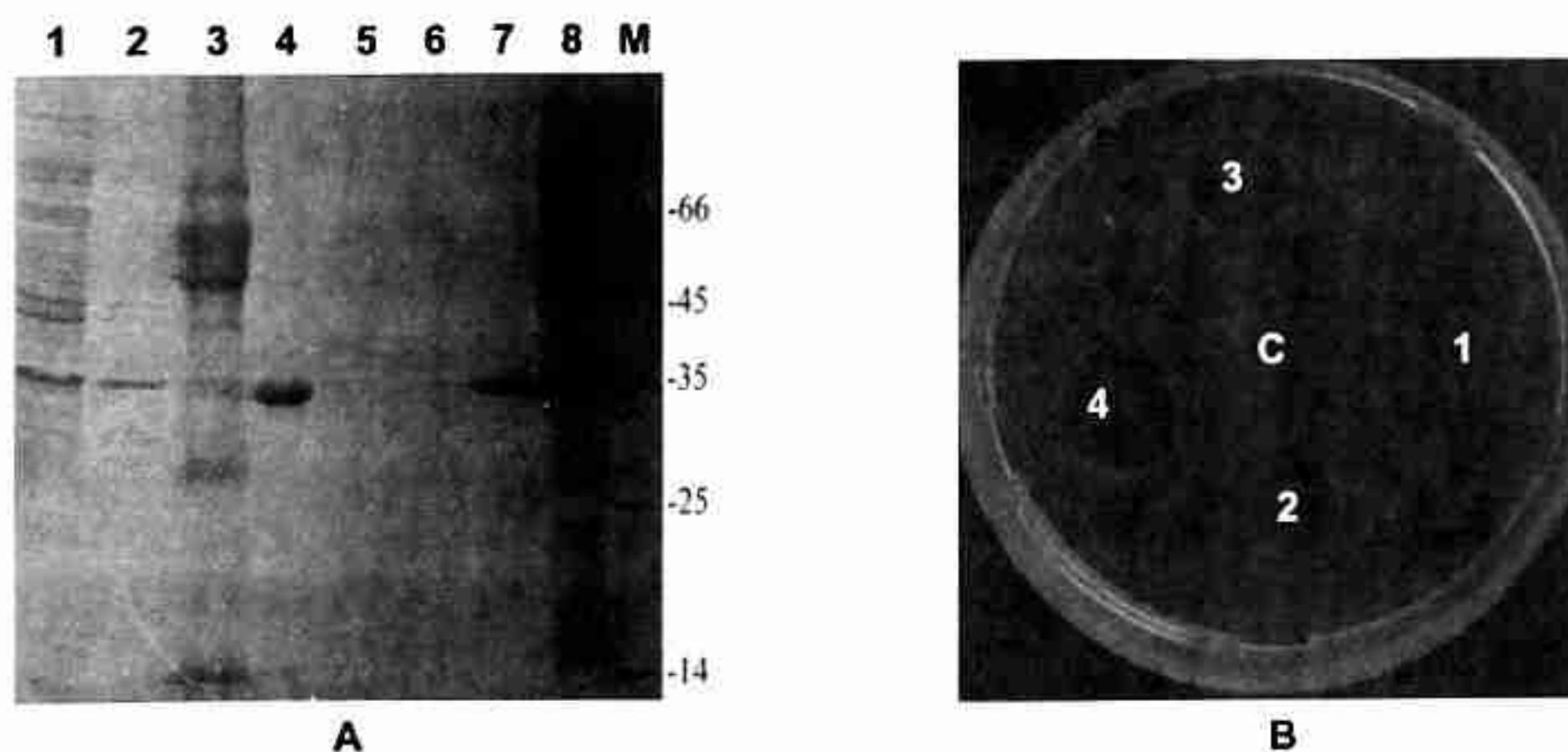
Xu hướng biểu hiện protein tái tổ hợp ở dạng thể vùi trong tế bào *E. coli* khá phổ biến. Kết quả nghiên cứu của chúng tôi cũng phù hợp với một số nghiên cứu trên thế giới. Dong *et al.*, (2004) đã nhân dòng và đánh giá tính chất của DNA tái tổ hợp mã hóa cho enzyme-3 (EFE-3) thủy phân fibrin từ giun đất *E. fetida* (GenBank: AY438622). Trong nghiên cứu này,

cDNA tái tổ hợp mã hóa cho EFE-3 được nhân dòng bằng RT-PCR. DNA tái tổ hợp có khung đọc mở (741 nucleotide) mã hóa một protein (247 amino acid). EFE-3 đã chỉ ra mức độ cao của các protease F-III-1, F-III-2 và bovine trypsin. EFE-3 tái tổ hợp đã được biểu hiện trong *E. coli* dưới dạng thể vùi và gen mã hóa dạng tự nhiên của EFE-3 được tiết ra trong môi trường tế bào COS-7. Hai sản phẩm tái tổ hợp này đều ở dưới dạng thể vùi và protease được tiết ra có hoạt tính thủy phân fibrin (Dong *et al.*, 2004).

Xu *et al.*, (2010) đã thiết kế và biểu hiện thành công đoạn peptide thuần thực của gen LK PI239 trong *E. coli*. Tác giả đã nhân dòng và đưa vào vector biểu hiện pET22b(-). Lumbrokinase tái tổ hợp đã được biểu hiện ở dạng thể vùi và có thể sử dụng để tinh sạch (Xu *et al.*, 2010)



Hình 4. (A) Điện di SDS-PAGE protein tổng số của mẫu biểu hiện rLK trong *E. coli* BL21(DE3)pLysE khi cảm ứng với 1 mM IPTG (3), không cảm ứng IPTG (2) và mẫu tế bào chỉ chứa vector pET22b(+) không mang gene *lk* (1); Thang protein chuẩn (M); (B) : Đánh giá mức độ biểu hiện protein bằng phần mềm Dolphin 1D.



Hình 5. (A) Phân tích protein trong quá trình siêu âm sơ chế dạng protein thể vùi trên SDS-PAGE (1,2,3,5,6: dịch nổi sau phá siêu âm lần 1, 2, 3, 4, 5, Giếng 4,7: cặn sau phá siêu âm lần 3 và 5, 8: protein tổng số); (B) Hoạt tính thủy phân đĩa fibrin (1: 10 µl protein tái tổ hợp ở dạng thể vùi; 2: 10 µl trypsin nồng độ 40 ng/µl; 3: 10 µl protein tái tổ hợp ở dạng thể vùi có bổ sung trypsin nồng độ 40 ng/µl; 4: 15 µl protein tái tổ hợp ở dạng thể vùi có bổ sung trypsin nồng độ 40 ng/µl; C: chuẩn plasmin).

Tinh sạch protein rLK1

Protein tổng số tiếp tục được hòa tan trong NBB và phá siêu âm 5 lần. Điện di kiểm tra dịch nổi và căn phá siêu âm, kết quả điện di thể hiện trên hình 5A.

Căn phá siêu âm lần 5 được hòa tan trong guanidin, và bổ sung enzyme trypsin để đánh giá hoạt tính thủy phân đĩa fibrin. Protein tái tổ hợp ở dạng thể vùi được xử lý bằng trypsin. Đây là enzyme thủy phân protein tạo ra các peptid có lysine (K) hoặc arginine (R) tại đầu C. Kết quả trên hình 6B cho thấy protein tái tổ hợp có hoạt tính thủy phân đĩa fibrin đạt 20 UI/mg protein. Điều này cũng phù hợp với nghiên cứu của Yan *et al.*, (2010) cho rằng lumbrakinase có thể được sử dụng làm thuốc uống vì nó có thể được hấp thụ từ ruột vào máu và hoạt hóa hệ thống thủy phân fibrin nội sinh (Yan *et al.*, 2010).

KẾT LUẬN

Từ những số liệu cho thấy, sau khi tối ưu, chỉ số CAI đã được cải thiện, đạt giá trị 0,86, phù hợp với hệ biểu hiện trên *E. coli*, các codon đều có tần số sử dụng ở mức cao (trên 50%) và không còn codon hiếm. Gen mã hóa lumbrakinase sau khi được tối ưu hóa codon chứa 747 bp, mã hóa cho 238 amino acid. Gen đã được tạo dòng trong vector pET22b(+) và biểu hiện trong *E. coli* BL21. Sau 6 giờ cảm ứng IPTG nồng độ 1 mM, xuất hiện băng protein chiếm đến hơn 40% protein tổng số. Protein tinh sạch có kích thước khoảng 35 kDa. Hoạt tính thủy phân đĩa fibrin đạt 20 U/mgprotein.

TÀI LIỆU THAM KHẢO

Asturf T, Mullertz S (1952) The fibrin plate method for estimating fibrinolytic activity. *Arch Biochem Biophys* 40: 346-351.

Dong GQ, Yuan XL, Shan YJ, Zhao ZH, Chen JP, Cong YW (2004) Molecular cloning and characterization of cDNA encoding fibrinolytic enzyme-3 from earthworm *Eisenia foetida*. *Acta Biochim Biophys Sin* (Shanghai) 36(4): 303-308.

Ge T, Sun ZJ, Fu SH, Liang GD (2005) Cloning of thrombolytic enzyme (lumbrakinase) from earthworm and its expression in the yeast *Pichia pastoris*. *Protein Expr Purif* 42(1): 20-28.

Hayes CS, Bose B, Sauer RT (2002) Stop codons preceded by rare arginine codons are efficient determinants of SsrA

tagging in *Escherichia coli*. *Proc Natl Acad Sci US* 99(6): 3440-3445.

Innerfield I (1960) Enzymes in clinical medicine. McGraw-Hill, New York.

Jansen R, Bussemaker HJ, Gerstein M (2003) Revisiting the codon adaptation index from a whole-genome perspective: analyzing the relationship between gene expression and codon occurrence in yeast using a variety of models. *Nucleic Acids Res* 31: 2242-2251.

Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Cur Opin Biotechnol* 6(5): 494-500.

Liu H, Li J, Du G, Zhou J, Chen J (2012) Enhanced production of α -cyclodextrin glycosyltransferase in *Escherichia coli* by systematic codon usage optimization. *J Ind Microbiol Biotechnol* 39(12): 1841-1849.

McNulty DE, Claffee BA, Huddleston MJ, Porter ML, Cavnar KM, Kane JF (2003) Mistranslational errors associated with the rare arginine codon CGG in *Escherichia coli*. *Protein Expr Purif* 27(2): 365-374.

Oliveira CC, van den Heuvel JJ, McCarthy JEG (1993) Inhibition of translational initiation in *Saccharomyces cerevisiae* by secondary structure: the roles of the stability and position of stem-loops in the mRNA leader. *Molecular Microbiology* 9(3): 521-532.

Retnoningrum DS, Pramesti HT, Santika PY, Valerius O, Asjarie S, Suciati T (2012) Codon optimization for high level expression of human bone morphogenetic protein-2 in *Escherichia coli*. *Protein Expr Purif* 84: 188-194.

Sharp PM, Li WH (1987) The codon Adaptation Index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15: 1281-1295.

Sugimoto M, Nakajima N (2001) Molecular cloning, sequencing and expression of cDNA encoding serine protease with fibrinolytic activity from earthworm. *Biosci Biotechnol Biochem* 65(7): 1575-1580.

Tanaka M, Tokuoka M, Shintani T, Gomi K (2012) Transcripts of a heterologous gene encoding mite allergen Der f 7 are stabilized by codon optimization in *Aspergillus oryzae*. *Appl Microbiol Biotechnol* 96(5): 1275-1282.

Tiwari A, Sankhyani A, Khanna N, Sinha S (2010) Enhanced periplasmic expression of high affinity humanized scFv against Hepatitis B surface antigen by codon optimization. *Protein Expr Purif* 74(2): 272-279.

Xu ZR, Yang YM, Gui QF, Zhang LN, Hu L (2010) Expression, purification, and characterization of recombinant lumbrakinase PL239 in *Escherichia coli*. *Protein Expr Purif* 69(2): 198-203.

Yan XM, Kim CH, Lee CK, Shin JS, Cho H, Sohn UR (2010) Intestinal absorption of fibrinolytic and proteolytic Lumbrokinase extracted from earthworm *Eisenia andrei*. *Korean J Physiol Pharmacol* 14: 71- 75.

SEQUENCE ANALYSIS OF LUMBROKINASE AND ITS CODON OPTIMIZATION FOR EXPRESSION IN *E. COLI* SYSTEM

Do Thi Tuyen[✉], Ly Thi Bich Thuy, Chu Thi Hoa, Vu Thi Bich Ngoc, Quyen Dinh Thi

Institute of Biotechnology, Vietnam Academy of Science and Technology

SUMMARY

Thrombotic complications of cardiovascular disease are a main cause of death and disability and, consequently, thrombolysis could favorably influence the outcome of such life-threatening diseases as myocardial infarction, cerebrovascular thrombosis, and venous thromboembolism. Thrombolytic agents are plasminogen activators that convert plasminogen, the inactive proenzyme of the fibrinolytic system in blood, to the proteolytic enzyme plasmin. Plasmin dissolves the fibrin of a blood clot, but may also degrade normal components of the hemostatic system and induce the so-called lytic state. Currently, thrombolytic agents are either approved for clinical use or under clinical investigation in patients with acute myocardial infarction: streptokinase, urokinase, rt-PA, APSAC, rscu-PA, staphylokinase and lumbrokinase (LK). The investigation on production of the recombinant lumbrokinase from *Eisenia foetida* for utilization as the thrombolytic and antithrombotic agent is essential for Vietnam. In this study, the gene encoding lumbrokinase from earthworm *Eisenia foetida* (GenBank Accession Nos: AF433650) was optimized codon for *E. coli* expression system. The results showed, after optimization, the CAI has improved and reached 0.86 of values, corresponding with the *E. coli* expression system. The codon frequencies were used in the high (above 50%) and no longer rare codons. The coding region of the *lkI* gene was 747 bp in lengths and encoded 238 amino acids. The gene was cloned in pET22b(+) vector and transformed and expressed into *E. coli* BL21. After induction with 1 mM IPTG, new protein bands accounted for approximately 40% of total protein and increased in induction period, reached the peak at 6 hrs after induction. The molecular mass of the purified rLK determined by SDS-PAGE was 35 kDa. Fibrinolytic activity of rLK reached 20 U/mgprotein.

Keywords: Expression, lumbrokinase; *E. coli* BL21; Fibrinolytic; Purification, Plasminogen

[✉] Author for correspondence: Tel: + 84-4-37568260; Fax: + 84-4-38363144; E-mail: dtuyen@ibt.ac.vn