



TẠO DÒNG GENE MÃ HÓA NATTOKINASE TỪ CHỦNG *BACILLUS SUBTILIS* N05

Nguyễn Thị Anh Thu^{1,2*}, Nguyễn Trần Mễ Khuê¹

¹ Trường Đại học Khoa học, Đại học Huế

² Trường Đại học Y Dược Huế

Tóm tắt: Natto là một loại thực phẩm chức năng có giá trị đã được chứng minh là giúp phòng các bệnh tim mạch, tai biến. Hoạt chất chính trong natto là một subtilisin có tên gọi nattokinase (EC 3.4.21.62). Nattokinase do *Bacillus subtilis* natto, một loại vi khuẩn có lợi có mặt trong nhiều thực phẩm lên men tiết ra. Nattokinase hiện nay được sản xuất bằng phương pháp lên men lẫn công nghệ DNA tái tổ hợp. Trong nghiên cứu này chúng tôi báo cáo việc tạo dòng gene mã hóa nattokinase từ một chủng *Bacillus subtilis* N05 có hoạt tính protease mạnh được phân lập từ chế phẩm natto Nhật Bản nhằm phục vụ cho việc sản xuất nattokinase bằng công nghệ DNA tái tổ hợp. Gen mã hóa (N05) cho nattokinase là một đoạn DNA có kích thước 1147 nucleotide (Genbank accession No: KU341115) mã hóa cho một protein trưởng thành có kích thước xấp xỉ 27,7 kDa.

Từ khóa: *Bacillus subtilis*, nattokinase, aprN, tạo dòng

1 Đặt vấn đề

Đột quy, bao gồm nhồi máu não và nhồi máu cơ tim được xác định là nguyên nhân gây tử vong hàng đầu trên thế giới trong số các bệnh không truyền nhiễm, với 17 triệu ca tử vong vào năm 2012 [7]. Tại Việt Nam hiện nay đang có xu hướng trẻ hóa đối với bệnh nhân bị tai biến và tỉ lệ tử vong do tai biến mạch máu não lên đến 50% [1]. Cơ chế của đột quy đã được xác định là do sự hình thành các cục máu đông tại những mạch máu hẹp ở não và tim, khiến cho máu không thể đến các mô do mạch máu bị tắc nghẽn cung cấp, đồng thời gây ra nguy cơ vỡ mạch máu. Khi mạch máu bị vỡ, máu sẽ tràn vào không gian xung quanh, gây tăng áp lực tại các mô này, khiến cho đột quy càng nghiêm trọng hơn [3]. Cơ thể có cơ chế để phân giải các cục máu đông này thông qua hệ thống phân hủy fibrin, trong đó có các thành phần chủ yếu như plasminogen (khi hoạt hóa trở thành plasmin, tác nhân chính gây phân hủy fibrin), các chất hoạt hóa plasminogen (PA) và chất ức chế phân hủy fibrin. Nhiều protease serine có khả năng hoạt hóa plasminogen thành plasmin, do đó gián tiếp có khả năng làm tan cục máu đông. Cơ thể đạt được sự cân bằng giữa các yếu tố phân hủy fibrin, là plasmin, plasminogen, PA và các chất gây hình thành fibrin hoặc ức chế phân hủy fibrin để cân bằng giữa đông máu và chảy máu. Trong một số trường hợp các chất PA ngoại lai như streptokinase, staphylokinase và

* Liên hệ: thu.dhyhue@gmail.com

APSAC có thể được đưa vào cơ thể để hỗ trợ cho PA nội sinh nhằm làm tan các cục máu đông, tuy nhiên việc sử dụng các PA nói trên thường đi kèm một số nguy cơ [6].

Nattokinase (EC 3.4.21.62) là một loại protease serine có mặt trong thực phẩm lên men truyền thống của Nhật Bản gọi là natto. Vi sinh vật chính trong men sản xuất ra nattokinase là *Bacillus subtilis natto* (*B. subtilis natto*). Enzyme này được sản xuất ở dạng bất hoạt chứa trình tự dẫn signal peptide (30 axit amin), trình tự propeptide ức chế protease (70 nối với đầu N của chuỗi polypeptide trưởng thành). Trình tự propeptide có vai trò quan trọng đối với việc biểu hiện nattokinase tái tổ hợp [4].

Đã có nhiều nghiên cứu chứng minh khả năng cải thiện nguy cơ tai biến của nattokinase khi sử dụng trên người. Những nghiên cứu này cũng chỉ ra rằng nattokinase ít gây hiệu ứng phụ trên người và có thể sử dụng như một loại thuốc bổ [5]. Trong thực tế, thuốc có tên gọi nattoşpes [1] đã được bán rộng rãi trên thị trường Việt Nam. Tuy nhiên giá thành của loại thuốc này còn quá cao so với thu nhập bình quân của người Việt.

Công nghệ DNA tái tổ hợp là giải pháp hữu hiệu để thu nhận được một lượng protein lớn có độ tinh khiết cao. Xét thấy nhu cầu sử dụng nattokinase của người dân Việt Nam rất cao, việc sản xuất nattokinase với giá thành rẻ là một giải pháp tốt để giải quyết nhu cầu cấp bách này. Trong nghiên cứu này chúng tôi báo cáo việc tạo dòng gene mã hóa nattokinase từ chủng *B. subtilis* N05 đã được phân lập từ chế phẩm natto Nhật Bản bằng phương pháp PCR.

2 Vật liệu và phương pháp nghiên cứu

2.1 Vật liệu nghiên cứu

Là chủng *B. subtilis* N05 đã được phân lập và cất giữ ở -80°C trong phòng thí nghiệm Bộ môn Công nghệ Sinh học, Trường Đại học Khoa học, Đại học Huế.

2.2 Phương pháp nghiên cứu

Tách chiết DNA tổng số từ *B. subtilis* N05

Chủng *B. subtilis* N05 được phân lập từ chế phẩm natto Nhật Bản trước đây được nuôi cấy trong môi trường Luria Bertani (LB) (1% tryptone, 0,5% Yeast extract, 1% NaCl) qua đêm với tốc độ lắc 185 vòng/phút ở 37°C . Sinh khối được thu nhận bằng cách ly tâm 10.000 vòng/phút trong 5 phút. DNA tổng số được thu nhận bằng phương pháp CTAB (3% cetyltrimethyl ammonium bromide, 28% NaCl, 4% EDTA, 100 mM Tris-HCl, 0,2% mercaptoethanol, 3%

polyvinylpyrrolidone, pH=8) [8]. DNA tổng số được kiểm tra trên gel agarose 0,8% trong đệm TAE.

Thiết kế môi đặc hiệu để tạo dòng gene aprN mã hóa cho nattokinase

Trình tự dùng để thiết kế môi tạo dòng gene aprN được thu nhận từ cơ sở dữ liệu NCBI nucleotide với từ khóa tìm kiếm aprN complete CDS AND *B. subtilis*[organism]. Một trong số các trình tự này (mã số truy cập) được chọn làm khuôn để thiết kế cặp môi đặc hiệu. Để đảm bảo tính đặc hiệu của cặp môi, chúng tôi sử dụng chương trình Primer BLAST trên NCBI, với khuôn là trình tự gene AF368283 và kiểm tra tính đặc hiệu trên toàn bộ genome của các chủng *B. subtilis*. Để tạo điều kiện thuận lợi cho việc tạo dòng gene aprN (là gene mã hóa cho nattokinase) vào vector biểu hiện, trình tự cắt hạn chế của *Bam*HI (GGATCC) và *Xma*I (CCCGGG) lần lượt được thiết kế vào đuôi 5' của môi xuôi và môi ngược.

Điện di DNA

Mẫu DNA được phân tích trên gel agarose có nồng độ khác nhau (từ 0,8 - 1,0%) trong đệm Tris Acetic acid EDTA (TAE). Kết quả được phát hiện bằng cách nhuộm bằng thuốc nhuộm Ethidium bromide (EtBr) và đọc trên máy đọc UV transilluminator.

Tạo dòng gene aprN bằng phương pháp PCR

Hỗn hợp phản ứng PCR gồm 40 ng DNA tổng số, môi đặc hiệu mỗi loại 10 pmol và 10 ul 2xPCR Master mix (Fermentas) với tổng thể tích là 20 ul. Chu trình khuếch đại gồm biến tính 95°C trong 5 phút, và 30 chu kỳ với biến tính 94°C trong 30 giây, gắn môi 55°C trong 30 giây và kéo dài 72°C trong 1 phút. Cuối cùng phản ứng được kết thúc bằng bước kéo dài 72°C trong 10 phút và làm giảm dần nhiệt độ tới 4°C.

Sản phẩm PCR được tinh sạch bằng bộ kit GeneJET Gel Extraction Kit (Thermo Scientific) và thực hiện phản ứng gắn đuôi A. Thành phần phản ứng gắn đuôi A gồm toàn bộ lượng sản phẩm PCR thu nhận được bổ sung với 10 mM dNTP và 1 U Taq Polymerase. Tổng thể tích phản ứng là 50 ul. Phản ứng được thực hiện trong máy luân nhiệt ở 72°C trong 10 phút. Sản phẩm được tinh sạch bằng bộ kit GeneJET Gel Extraction Kit.

Để đọc trình tự gene, sản phẩm gắn đuôi A nói trên được nối vào vector pCR® 2.1 (Fermentas) với thành phần phản ứng gồm 2 ul 10x buffer, 1 ul 5u/ul ligase T4, 40 ng vector pCR® 2.1, 40 ng sản phẩm PCR gắn đuôi A. Phản ứng được ủ trong 4h ở 22°C, sau đó đem biến nạp vào chủng *E.coli* Top10 bằng phương pháp sốc nhiệt như mô tả[2]. Chủng *E.coli* sau khi biến

nạp được trải lên đĩa LB chứa 50 ug/ml Ampicillin có bổ sung IPTG và X-gal. Thể biến nạp được dựa trên màu trắng xanh của khuẩn lạc mọc lên trên đĩa sau 18 giờ nuôi cấy ở 37°C.

Xác định trình tự sản phẩm PCR

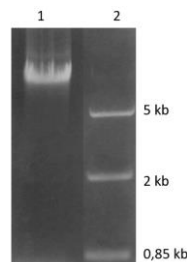
Khuẩn lạc trắng được thu nhận, kiểm tra sự hiện diện của insert bằng phản ứng PCR với cặp mồi đặc hiệu nói trên, và sau đó đem nuôi cấy để tách plasmid. Plasmid pCR® 2.1 tái tổ hợp được tách bằng bộ kit miniprep GeneJET Plasmid Miniprep (Thermoscientific), cắt bằng enzyme EcoRI để kiểm tra sự hiện diện của insert, và sau đó gửi đi giải trình tự bằng cặp mồi xuôi và ngược của M13 có sẵn trên vector.

Kết quả giải trình tự được kiểm tra bằng phần mềm FinchTV nhằm loại bỏ những base đọc không tốt, sau đó được chuyển đổi sang định dạng fasta. Trình tự insert được nối lại từ hai trình tự đọc xuôi và đọc ngược bằng phần mềm Contigexpress trong bộ phần mềm Vector NTI. Tiếp đến, trình tự vector được loại bỏ bằng cách xử lý trình tự thu nhận được ở trên bằng phần mềm Vecscreen. Cuối cùng trình tự insert được so sánh với trình tự lý thuyết bằng phương pháp giống cột 2 trình tự, sử dụng phần mềm lalign (http://www.ch.embnet.org/software/LALIGN_form.html).

3 Kết quả nghiên cứu và thảo luận

3.1 Tách chiết DNA tổng số

Kết quả điện di DNA tổng số của chủng *B. subtilis* N05 chuẩn bị bằng phương pháp CTAB cho thấy sản phẩm DNA tổng số tách được đảm bảo chất lượng và lượng cho các thí nghiệm tiếp theo (Hình 1). Băng DNA sáng, rõ và không có vệt phía dưới, chứng tỏ DNA tổng số không bị đứt gãy và đủ sạch.



Hình 1. DNA tổng số tách chiết từ chủng *B. subtilis* N05. Đường chạy 1: DNA tổng số, 3 ul mẫu được tải lên giếng số 1; đường chạy 2: thang chuẩn DNA (5 kb DNA ladder); mẫu được phân tích trên gel agarose 0,8% trong đệm TAE.

3.2 Thiết kế mồi đặc hiệu cho gene arpN

Kết quả thiết kế mồi bằng phần mềm primer BLAST cho ra nhiều cặp mồi đặc hiệu. Tất cả các cặp mồi này đều thỏa mãn điều kiện khuyếch đại một sản phẩm duy nhất trong genome của tất cả các chủng *B. subtilis*.

Trình tự cặp mồi mà chúng tôi thu nhận được là:

Mồi xuôi (Forward primer):

5'-GGATCCTTCAGCAACAAGTCTGC-3'

Mồi ngược (Reverse primer):

5'-CCCGGGTTATTGTGCAGCTGCTT-3'

trong đó trình tự gạch dưới là trình tự cắt của các enzyme cắt hạn chế *Bam*HI và *Xma*I được thêm vào mồi đặc hiệu sau khi được thiết kế.

3.3 Tạo dòng gene arpN bằng PCR

Kết quả PCR được phân tích trên gel agarose 1% trong đệm TAE. Sau khi nhuộm gel bằng EtBr và quan sát dưới máy đọc UV, kết quả cho thấy có một sản phẩm duy nhất có kích thước khoảng trên 1 kb (Hình 2).

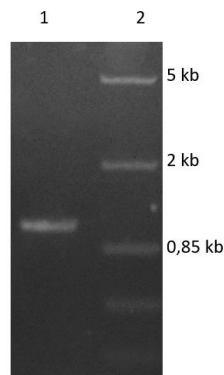


Fig. 2. Hình 2. Kết quả điện di sản phẩm PCR bằng cặp mồi đặc hiệu nói trên với khuôn là DNA tổng số từ *B. subtilis* N05. 3 ul/20 ul thể tích phản ứng được phân tích trên gel agarose 1% trong đệm TAE. Đường chạy 1: sản phẩm PCR; đường chạy 2: thang chuẩn DNA (5 kb DNA ladder).

3.4 Tạo dòng sản phẩm PCR vào pCR® 2.1 và xác định trình tự

Sản phẩm PCR được tạo dòng vào vector pCR® 2.1 (kết quả không trình bày) và gửi đi giải trình tự. Sau khi xử lý kết quả giải trình tự như đã mô tả ở phần phương pháp, trình tự thu nhận được được đem so sánh trên cơ sở dữ liệu ngân hàng gene GenBank bằng BLAST và giống cột với trình tự trên ngân hàng GenBank AF368283. Kết quả cho thấy gene thu nhận được chính là gene aprN và giống nhau đến 99,9% với trình tự gene có mã truy cập KJ174339 (Hình 4) về trình tự nucleotide và giống nhau 100% về trình tự axit amin. Điều này chứng tỏ gene aprN đã được tạo dòng thành công và sẵn sàng cho việc biểu hiện. Trình tự aprN của chúng tôi đã được đăng ký trên GenBank với mã truy cập KU341115.

```

10  20  30  40  50  60
AF3682 GTGAGAAGCAAAAAATTGTGGATCAGCTTGTGTGTTGCGTTAACGTTAATCTTTACGATG
.....
KU3411 GTGAGAAGCAAAAAATTGTGGATCAGCTTGTGTGTTGCGTTAACGTTAATCGTTACGATG
10  20  30  40  50  60

70  80  90  100  110  120
AF3682 GCGTTCAGCAACATGTCTGCGCAGGCTGCCGGAAAAAGCAGTACAGAAAAGAAATACATT
.....
KU3411 GCGTTCAGCAACATGTCTGCGCAGGCTGCCGGAAAAAGCAGTACAGAAAAGAAATACATT
70  80  90  100  110  120

130  140  150  160  170  180
AF3682 GTCGGATTTAAGCAGACAATGAGTGCCATGAGTTCGCCAAGAAAAAGGATGTTATTTCT
.....
KU3411 GTCGGATTTAAGCAGACAATGAGTGCCATGAGTTCGCCAAGAAAAAGGATGTTATTTCT
130  140  150  160  170  180

190  200  210  220  230  240
AF3682 GAAAAAGGCGGAAAGTTCAAAAAGCAATTTAAGTATGTTAACGCGGCCGAGCAACATTG
.....
KU3411 GAAAAAGGCGGAAAGTTCAAAAAGCAATTTAAGTATGTTAACGCGGCCGAGCAACATTG
190  200  210  220  230  240

250  260  270  280  290  300
AF3682 GATGAAAAAGCTGTAAAAGAATTGAAAAAGATCCGAGCGTTGCATATGTGGAAGAAGAT
.....
KU3411 GATGAAAAAGCTGTAAAAGAATTGAAAAAGATCCGAGCGTTGCATATGTGGAAGAAGAT
250  260  270  280  290  300

310  320  330  340  350  360
AF3682 CATATTGCACATGAATATGCGCAATCTGTTTCCTTATGGCATTCTCAAATTAAGCGCCG
.....
KU3411 CATATTGCACATGAATATGCGCAATCTGTTTCCTTATGGCATTCTCAAATTAAGCGCCG
310  320  330  340  350  360

370  380  390  400  410  420

```

AF3682 GCTCTTCACTCTCAAGGCTACACAGGCTCTAACGTAAGTAGCTGTTATCGACAGCGGA
.....
KU3411 GCTCTTCACTCTCAAGGCTACACAGGCTCTAACGTAAGTAGCTGTTATCGACAGCGGA
370 380 390 400 410 420

430 440 450 460 470 480
AF3682 ATTGACTCTTCTCATCCTGACTTAAACGTCAGAGGCGGAGCAAGCTTCGTTCTTCTGAA
.....
KU3411 ATTGACTCTTCTCATCCTGACTTAAACGTCAGAGGCGGAGCAAGCTTCGTTCTTCTGAA
430 440 450 460 470 480

490 500 510 520 530 540
AF3682 ACAAACCCATACCAGGACGGCAGTTCTCACGGTACGCATGTCGCCGGTACGATTGCCGCT
.....
KU3411 ACAAACCCATACCAGGACGGCAGTTCTCACGGTACGCATGTCGCCGGTACGATTGCCGCT
490 500 510 520 530 540

550 560 570 580 590 600
AF3682 CTTAATAACTCAATCGGTGTTCTGGGCGTAGCGCCAAGCGCATCATTATATGCAGTAAAA
.....
KU3411 CTTAATAACTCAATCGGTGTTCTGGGCGTAGCGCCAAGCGCATCATTATATGCAGTAAAA
550 560 570 580 590 600

610 620 630 640 650 660
AF3682 GTGCTTGATTCAACAGGAAGCGGCCAATATAGCTGGATTATTAACGGCATTGAGTGGGCC
.....
KU3411 GTGCTTGATTCAACAGGAAGCGGCCAATATAGCTGGATTATTAACGGCATTGAGTGGGCC
610 620 630 640 650 660

670 680 690 700 710 720
AF3682 ATTTCCAACAATATGGATGTTATCAACATGAGCCTTGGCGGACCTACTGGTTCTACAGCG
.....
KU3411 ATTTCCAACAATATGGATGTTATCAACATGAGCCTTGGCGGACCTACTGGTTCTACAGCG
670 680 690 700 710 720

730 740 750 760 770 780
AF3682 CTGAAAACAGTAGTTGATAAAGCGGTTTCCAGCGGTATCGTCGTTGCTGCCGCAGCCGGA
.....
KU3411 CTGAAAACAGTAGTTGATAAAGCGGTTTCCAGCGGTATCGTCGTTGCTGCCGCAGCCGGA
730 740 750 760 770 780

790 800 810 820 830 840
AF3682 AACGAAGGTTTCATCCGGAAGCACAAGCACAGTCGGCTACCCTGCAAAAATATCCTTCTACT
.....
KU3411 AACGAAGGTTTCATCCGGAAGCACAAGCACAGTCGGCTACCCTGCAAAAATATCCTTCTACT
790 800 810 820 830 840
850 860 870 880 890 900
AF3682 ATTGCAGTAGGTGCGGTAAACAGCAGCAACCAAAGAGCTTCATTCTCCAGCGTAGGTTCT
.....
KU3411 ATTGCAGTAGGTGCGGTAAACAGCAGCAACCAAAGAGCTTCATTCTCCAGCGTAGGTTCT
850 860 870 880 890 900

```

          910   920   930   940   950   960
AF3682 GAGCTTGATGTAATGGCTCCTGGCGTGTCCATCCAAAGCACACTTCTGGAGGCACTTAC
.....
KU3411 GAGCTTGATGTAATGGCTCCTGGCGTGTCCATCCAAAGCACACTTCTGGAGGCACTTAC
          910   920   930   940   950   960

          970   980   990   1000  1010  1020
AF3682 GCGCCTTATAACGGAACGTCCATGGCGACTCCTCACGTTGCCGGAGCAGCAGCGCTAATT
.....
KU3411 GCGCCTTATAACGGAACGTCCATGGCGACTCCTCACGTTGCCGGAGCAGCAGCGCTAATT
          970   980   990   1000  1010  1020
          1030  1040  1050  1060  1070  1080
AF3682 CTTTCTAAGCACCCGACTTGGACAAACGCGCAAGTCCGTGATCGTTTAGAAAGCACTGCA
.....
KU3411 CTTTCTAAGCACCCGACTTGGACAAACGCGCAAGTCCGTGATCGTTTAGAAAGCACTGCA
          1030  1040  1050  1060  1070  1080

          1090  1100  1110  1120  1130  1140
AF3682 ACATATCTTGAAACTCTTCTACTATGGAAAAGGGTAAATCAACGTACAAGCAGCTGCA
.....
KU3411 ACATATCTTGAAACTCTTCTACTATGGAAAAGGGTAAATCAACGTACAAGCAGCTGCA
          1090  1100  1110  1120  1130  1140

AF3682 CAATAA-
.....
KU3411 CAATAAA

```

Hình 3. Kết quả sắp giống cột hai trình tự của gene aprN từ chủng *B. subtilis* N05 (KU341115) với một trình tự aprN đã được công bố trên GenBank (AF368283) được sử dụng để thiết kế cặp mồi đặc hiệu. Hộp nhỏ đánh dấu vị trí sai khác nucleotide giữa hai trình tự.

4 Kết luận

Trong công trình này chúng tôi báo cáo việc tạo dòng thành công gene aprN từ chủng *B. subtilis* N05 phân lập từ chế phẩm natto Nhật Bản, cũng là một nguồn khá phổ biến cho việc phân lập *B. subtilis*. Sử dụng phần mềm thích hợp, các mồi đặc hiệu được thiết kế để khuếch đại một sản phẩm duy nhất từ genome của *B. subtilis* N05. Kết quả phân tích trình tự cho thấy gene khuếch đại chính là aprN. Gene này đã được đăng ký trên GenBank với mã số truy cập là KU341115. Nó có trình tự giống đến 99% với phần lớn các gene aprN trên GenBank ở mức nucleotide và giống 100% ở mức axit amin.

Tài liệu tham khảo

1. Hạnh, T. (2017). "Nhiều 9X đột quy: Chó chủ quan nếu đau đầu dai dẳng." Retrieved 25/7/, 2017, from <http://vietnamnet.vn/vn/suc-khoe/cac-loai-benh/9x-dot-quy-tai-bien-mach-mau-nao-ngay-cang-nhieu-366680.html>.

2. Christine E. Seidman, Kevin Struhl, Jen Sheen and a. T. Jessen (1997). Introduction of Plasmid DNA into Cells. Current protocol in molecular biology. F. M. Ausubel. USA, John Wiley & Sons: 1.8.1-1.8.10.
3. Hademenos, G. J. and T. F. Massoud (1997). "Biophysical mechanisms of stroke." Stroke 28(10): 2067-2077.
4. Ikemura, H., H. Takagi and M. Inouye (1987). "Requirement of pro-sequence for the production of active subtilisin E in Escherichia coli." J Biol Chem 262(16): 7859-7864.
5. Suzuki, Y., K. Kondo, Y. Matsumoto, B. Q. Zhao, K. Otsuguro, T. Maeda, Y. Tsukamoto, T. Urano and K. Umemura (2003). "Dietary supplementation of fermented soybean, natto, suppresses intimal thickening and modulates the lysis of mural thrombi after endothelial injury in rat femoral artery." Life Sci 73(10): 1289-1298.
6. Weitz, J. I. (1990). "Mechanism of action of the thrombolytic agents." Baillieres Clin Haematol 3(3): 583-599.
7. WHO (2014). Global status report on noncommunicable diseases 2014. W. H. Organization. Switzerland: 9.
8. Wilson, K. (1997). Preparation of Genomic DNA from Bacteria. Current Protocols in Molecular Biology F. M. Ausubel. USA, John Wiley & Sons: 2.4.1-2.4.5.

CLONING OF A GENE ENCODING NATTOKINASE FROM *BACILLUS SUBTILIS* N05

Nguyen Thi Anh Thu^{1,2*}, Nguyen Tran Me Khue¹

¹ Truong Dai hoc Khoa hoc, Dai hoc Hue

² Truong Dai hoc Y Duoc Hue

Abstract: Natto is a kind of functional food that is highly valued in Asia, and has been proved to significantly reduce the risk of cardiovascular accidents as well as cerebral stroke. The main bioactive component in natto is a kind of subtilisin called nattokinase (EC 3.4.21.62). Nattokinase is produced by various *Bacillus subtilis natto* strains, a beneficial bacterium that is present in many fermented foods. Currently nattokinase can be produced by traditional fermentation method or by recombinant DNA technology. In this research we report the cloning of a gene encoding nattokinase from *Bacillus subtilis* strain N05 that has previously been isolated from natto products of Japan. The encoding gene (N05, Genbank accession No: KU341115) is 1147 bp long, encodes for a 27.7 kDa mature protein.

Keywords: *Bacillus subtilis*, nattokinase, aprN, cloning