

NGHIÊN CỨU BIỂU HIỆN INSULIN NGƯỜI Ở *E. coli*

BÙI THỊ HUYỀN, LÊ VĂN PHÙNG, PHAN VĂN CHI

1. ĐẶT VẤN ĐỀ

Insulin là một hormone quan trọng có vai trò điều hòa đường huyết do các tế bào beta trong đảo Langerhans của tuyến tụy sản xuất [9]. Hàm lượng đường (glucose) trong máu là nguồn năng lượng thiết yếu cho cơ thể, nếu không duy trì ở mức bình thường có thể gây ra những căn bệnh nguy hiểm [3]. Khi đường huyết tăng có thể gây ra sự bài tiết đường qua nước tiểu, dẫn đến mất/thiếu hụt glucose, hiện tượng này còn gọi là bệnh đái tháo đường [7, 5]. Nguyên nhân là do sự thiếu hụt insulin hoặc vì một lí do nào đó cơ thể không thể sử dụng được insulin dẫn đến đường huyết tăng cao và gây đái tháo đường [8]. Bệnh đái tháo đường được chia làm hai loại là type I và type II [10]. Đái tháo đường type I là loại bệnh phụ thuộc insulin, thường gặp ở người trẻ tuổi (< 35 tuổi), lứa tuổi hay gặp nhất là 10 - 16 tuổi. Đây là một dạng bệnh nặng trong đó các tế bào của tuyến tụy có nhiệm vụ tiết insulin bị khiếm khuyết nên cơ thể không có insulin để sử dụng. Nếu không điều trị bằng cách tiêm insulin, bệnh nhân sẽ hôn mê và tử vong [8]. Đái tháo đường type II là loại đái tháo đường không phụ thuộc insulin, bệnh thường gặp ở người > 40 tuổi, người béo phì, trong đó cơ thể vẫn sản xuất insulin nhưng vì một lí do nào đó không thể sử dụng được nó [4].

Về cấu trúc, ở người phân tử insulin bao gồm hai chuỗi polypeptide: chuỗi A gồm 21 axit amin và chuỗi B gồm 30 axit amin [1, 2]. Hai cầu nối disulfide (vị trí A7 - B7, và A20 - B19) đồng hóa trị như hai sợi dây buộc chặt hai chuỗi với nhau, trong chuỗi A còn có một cầu nối disulfide giữa gốc A6 và A11 đảm bảo phân tử khá ổn định trong không gian. Đặc biệt là cả 3 cầu nối này đều bất biến ở tất cả các dạng insulin của động vật.

Hiện nay, tình hình phát triển bệnh đái tháo đường ở Việt Nam cũng như trên thế giới đang có xu hướng tăng nhanh. Insulin là hormon đầu tiên được tổng hợp nhân tạo, được sản xuất bằng công nghệ DNA tái tổ hợp phục vụ điều trị bệnh này [14]. Năm 2005, nhu cầu insulin dùng trong trị bệnh đái tháo đường ước tính khoảng 4 đến 5 tấn và dự kiến năm 2010 là 16 tấn [12]. Nhu cầu về insulin của thế giới vượt qua con số vài tấn/năm và vì thế nguồn cung cấp insulin cho trị bệnh tiểu đường đang ngày càng thiếu hụt [13]. Trong bài báo này, chúng tôi trình bày một số kết quả nghiên cứu biểu hiện và xác định insulin người từ vector pET 32c (+).

2. VẬT LIỆU VÀ PHƯƠNG PHÁP

2.1. Vật liệu

Nguồn gen mã hóa cho insulin chuỗi A và B đã được dòng hóa trong vector pET14bInsA/B do nhóm của PGS.TS Lê Văn Phùng (Đại học Y Hà Nội) cung cấp.

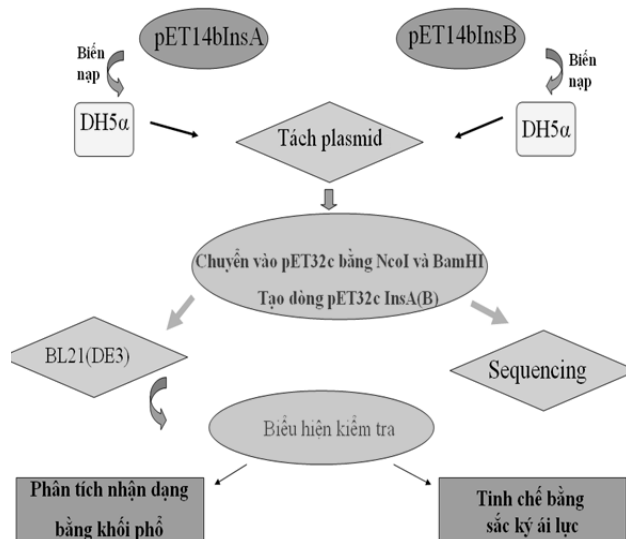
Các hóa chất sử dụng cho điện di SDS-PAGE từ BioRad; Acetonitrile (ACN) từ J.T Barker (Mỹ); acid formic (FA), trifluoroacetate (TFA) mua từ Fluka (Thụy Sĩ); dithiothreitol (DTT),

iodoacetamide (IAA), ammonium bicarbonate (NH₄HCO₃), trypsin từ Sigma-Aldrich (Mỹ) và một số hóa chất thông dụng dùng trong sinh học phân tử.

2.2. Phương pháp nghiên cứu

2.2.1. Thiết kế vector biểu hiện cho hai chuỗi A và B insulin

Gen mã hóa cho chuỗi A của insulin gồm 63 cặp nucleotide (tương ứng với 21 amino acid) đã được dòng hóa vào vector pET-14b có chứa điểm cắt của hai enzyme giới hạn *NdeI* và *BamHI* ở hai đầu gen và đoạn His-Tag (6xHis) ở đầu N của chuỗi A. Gen mã hóa cho chuỗi B gồm 90 cặp nucleotide (tương ứng với 30 amino acid) cũng được thiết kế tương tự. Do kích thước mỗi gen rất nhỏ (60 và 90 bp) nên protein dạng dung hợp His-Tag chain A và B biểu hiện rất khó khăn và hầu hết tồn tại ở trạng thái không tan. Để khắc phục nhược điểm này, các đoạn gen lần lượt được chuyển vào vector pET-32c(+) bằng hai enzyme giới hạn *NcoI* và *BamHI* tạo ra protein dạng dung hợp mới Trx.Tag-His.Tag-Thrombin- S.Tag-Eterokinase-Chain A/B-His.Tag. Theo tính toán lí thuyết dung hợp protein mới có kích thước 22 và 23 kDa. Cấu trúc này làm tăng khả năng hòa tan nhờ nối ghép với Thioredoxin (Trx) và tăng kích thước phân tử, tạo điều kiện thuận lợi cho các quá trình tinh chế tiếp theo (hình 1).



Hình 1. Mô hình quy trình thiết kế, biểu hiện và xác định insulin

Quy trình thí nghiệm được mô tả tóm tắt như sau:

Plasmid tái tổ hợp pET14bInsA/B có chứa gen mã hóa cho chuỗi A/B của insulin và vector biểu hiện pET 32c(+) được xử lí lần lượt bằng hai enzym giới hạn *NcoI* và *BamHI*. Phản ứng được thực hiện ở 37°C trong 3 giờ. Điện di sản phẩm thu được trên gel agarose 1,5% sử dụng kit thời gel QIAGEN để thu lại băng DNA mã hóa cho chuỗi A, B. Sau đó đoạn gen mã này nối ghép với vector pET 32c(+) đã mở vòng với sự có mặt của enzym nối T₄ ligase, phản ứng được thực hiện ở 22°C trong khoảng thời gian 16 giờ. Sản phẩm phản ứng được biến nạp vào tế bào khả biến DH5α, được nuôi cấy, tách chiết plasmid. Kết quả của quá trình biến nạp được kiểm tra lại bằng phản ứng cắt bởi enzym giới hạn *NdeI*, khẳng định đoạn gen mong muốn đã được đưa vào vectơ.

2.2.2. Đọc trình tự gen

Trình tự DNA của hai đoạn gen mã hóa cho hai chuỗi A và B của insulin được tiến hành xác định trên máy đọc trình tự tự động ABI PRISM® 3100 Avant Genetic Analyzer (Applied Biosystems).

2.2.3. Biểu hiện insulin chuỗi A và B dưới dạng dung hợp

Vector tái tổ hợp (sau khi đã kiểm tra bằng phản ứng cắt với *NdeI*) được biến nạp vào chủng *E. coli* BL21 (DE3) tạo thành dòng tế bào BL21 (DE3)/pET32cInsA/B với dung hợp protein mới Trx.Tag-His.Tag-Thrombin- S.Tag-Eterokinase-Chain A/B- His.Tag. Dòng tế bào này được nuôi cấy trong môi trường LB lỏng có bổ sung ampicilin (nồng độ cuối cùng là 100µg/ml), lắc 200 vòng/phút ở 37°C. Khi giá trị OD_{600 nm} của dịch nuôi cấy đạt 0,5, bổ sung IPTG với nồng độ cuối cùng là 1 mM. Điện di SDS-PAGE trên gel polyacrylamid 12,6% để kiểm tra kết quả biểu hiện.

2.2.4. Xác định chuỗi A, B insulin bằng phương pháp khối phổ

Băng protein insulin tái tổ hợp (chuỗi A và B dung hợp) trên bản điện di SDS-PAGE được cắt ra và xử lý với dung dịch rửa (50 mM NH₄HCO₃, pH 8,0, 50% ACN). Sau đó, các mảnh gel được khử bằng dung dịch DTT 5 mM ở 56°C, trong 1 giờ và alkyl hóa bằng dung dịch IAA 5 mM trong tối, 1 giờ ở nhiệt độ phòng. Trypsin (Sigma) được thêm vào với tỉ lệ enzym : cơ chất là 1 : 50, ở 37°C, qua đêm.

Hỗn hợp peptide sau thủy phân được phân tách trên hệ sắc kí lỏng nano (1DnanoLC) với cột ngược pha RP C18 và phân tích trên hệ máy khối phổ QSTAR® XL với sự hỗ trợ của phần mềm Analyst QS để ghi phổ MS và MS/MS. Phần mềm Mascot v1.8 dựa trên thuật toán Mowse [6] được sử dụng để nhận diện protein trên cơ sở dữ liệu NCBI và Swiss-Prot theo quy trình đã được mô tả [11].

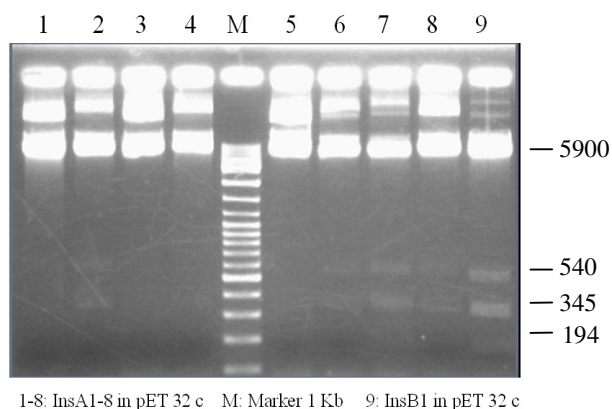
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kết quả cắt plasmid tái tổ hợp pET32cInsA/B bằng *NdeI*

Do chuỗi A (InsA) và chuỗi B (InsB) của insulin được đưa vào pET-14b bởi hai enzyme *NdeI* và *BamHI* nên quá trình chuyển đoạn DNA có chứa gen A/B vào vector pET-32c sẽ thành công nếu trong vector tái tổ hợp mới có thêm 1 điểm cắt của *NdeI*. Trong vector pET-32c đã có sẵn 2 điểm cắt của *NdeI* tại vị trí 396 và 691, đoạn DNA được chèn vào từ vị trí 198 (*NcoI*) đến vị trí 212 (*BamHI*) làm tăng thêm 1 điểm cắt cho *NdeI*. Theo tính toán lí thuyết thì khi xử lý vector tái tổ hợp mới bởi *NdeI* sẽ cho ra 3 băng với kích thước tương ứng là 194 bp, 345 bp và vector gốc 5900 bp.

Kết quả của phản ứng cắt plasmid tái tổ hợp bởi *NdeI* được thể hiện trên hình 2 cho thấy, sau khi xử lý bằng enzym *NdeI*, trong số các plasmid 1 - 8 (InsA trong pET-32c) thì chỉ có dòng 2, 6, 7, 8 và 9 (InsB trong pET-32c) đều cho ra 4 băng có kích thước tương đương nhau. Băng trên cùng là vector gốc pET-32c có kích thước khoảng 5900 bp, tiếp đó là hai băng 194 bp và 345bp phù hợp với tính toán. Trên hình 2 còn xuất hiện thêm một băng 540 bp, có thể đây là băng DNA tổng của 2 băng 194 bp và 345 bp do chưa được cắt hoàn toàn. Như vậy đoạn DNA mong muốn đã được chèn vào vector đích. Phân tích kết quả xác định trình tự nucleotide gen mã

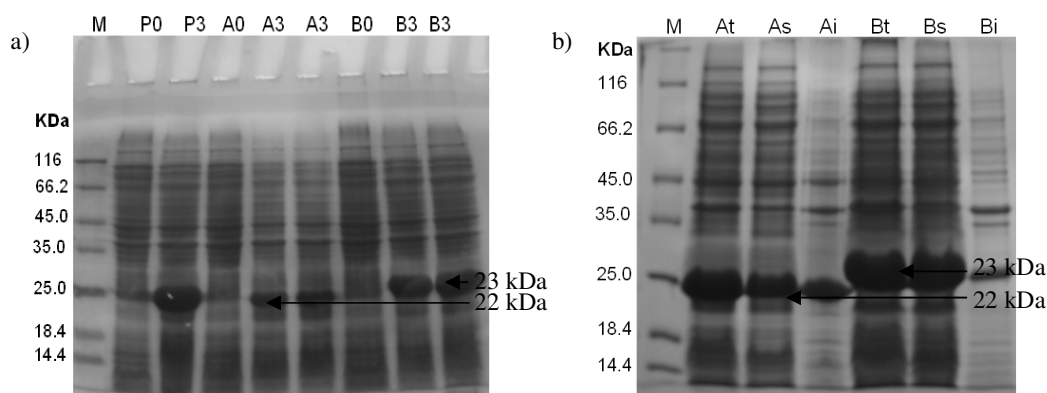
hóa cho chuỗi A và B của insulin trên máy đọc trình tự tự động ABI PRISM® 3100 Avant Genetic Analyzer cho thấy có sự tương đồng 100% với trình tự chuẩn chứng tỏ vector đã được thiết kế thành công (kết quả không trình bày ở đây).



Hình 2. Kết quả cắt plasmid tái tổ hợp pET32cInsA/B bằng *Nde*I

3.2. Biểu hiện chuỗi A và B của insulin dạng dung hợp ở *E. coli*

Kết quả biểu hiện chuỗi A và B của insulin dưới dạng dung hợp được kiểm tra bằng điện di SDS-PAGE trên hình 3a cho thấy, trong dịch chiết tổng số của các dòng tế bào thu lại ở thời điểm 3h sau cảm ứng (A3, B3) có xuất hiện một băng protein mới so với mẫu dịch chiết tế bào không được cảm ứng bởi IPTG (A0 và B0). So với thang M chuẩn, sau 3 giờ cảm ứng P3 xuất hiện 1 băng protein có kích thước khoảng 23 kDa, A3 cũng xuất hiện một băng có kích thước khoảng 22,0 kDa (do đoạn DNA 120 bp có chứa gen mã hóa cho chuỗi A chèn vào) và B3 xuất hiện băng protein mới khoảng 23 kDa. Kết quả này phù hợp với tính toán lí thuyết. Như vậy hệ biểu hiện pET 32c mang tổ hợp gen InsA (InsB) đã hoạt động ổn định để biểu hiện protein lai mới.



Hình 3. Điện di kiểm tra biểu hiện dung hợp protein chuỗi A/B trên gel polyacrylamide 12,6%

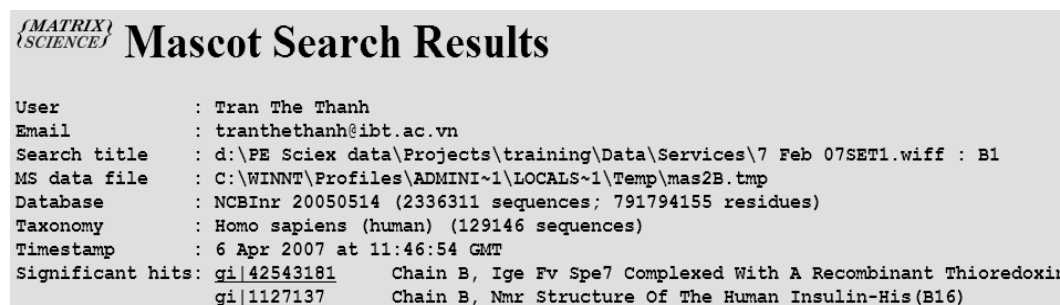
a) Biểu hiện các dòng tế bào mang gen A/B; b) Kiểm tra độ tan của các dòng tế bào đã biểu hiện

M: Marker SM 0431 (Fermentas); P0: dòng tế bào chứa vector nguyên thể trước khi cảm ứng;
A0, B0: các dòng tế bào chứa gen A/B trước khi cảm ứng; P3, A3 và B3: các dòng tế bào sau khi cảm ứng;
At, Bt: dịch tế bào tổng số; As, Bs: dịch tan; Ai, Bi: phần không tan

Chúng tôi tiếp tục khảo sát tính tan của protein mới (hình 3b). Mặc dù biểu hiện protein dưới dạng thể vùi cũng có một số ưu điểm nhất định nhưng protein tạo ra thường bị mất hoạt tính sinh học và quá trình tái tạo lại cấu trúc tương đối phức tạp và tốn kém. Hơn nữa để có thể tinh sạch được protein dễ dàng thì cần phải biểu hiện protein dưới dạng hòa tan. Trong thí nghiệm này các dòng tế bào được cảm ứng ở 22°C khoảng 6 giờ khi OD₆₀₀ của dịch tế bào đạt 0,6 -0,8. Kết quả thể hiện trên hình 3b cho thấy protein ở dạng tan (As và Bs) chiếm tỷ lệ rất lớn và lượng protein tồn tại ở trạng thái không tan (Ai và Bi) là không đáng kể so với protein tổng số (At và Bt).

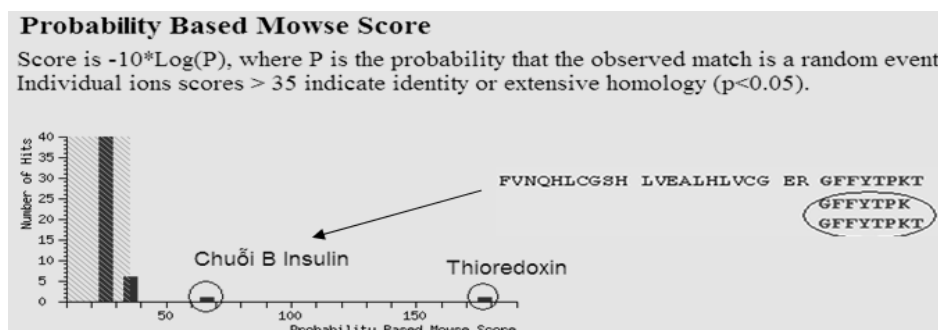
3.3. Nhận diện insulin tái tổ hợp

Bằng phương pháp 1DnanoLC-ESI-MS/MS, phổ khối của các peptide được tiến hành nhận diện trên cơ sở NCBI nr với sự hỗ trợ của phần mềm Mascot v1.8. Điểm đáng chú ý ở đây là chuỗi A với 21 amino acid không có điểm cắt (sau K và R) của trypsin nên không nhận diện được các mảnh peptide ở chuỗi A mà chỉ nhận diện được các mảnh peptide của thioredoxin- Trx đồng biểu hiện (kết quả không trình bày ở đây). Trong khi đó, chuỗi B có hai điểm cắt của trypsin nên kết quả nhận diện được minh họa như hình 4.



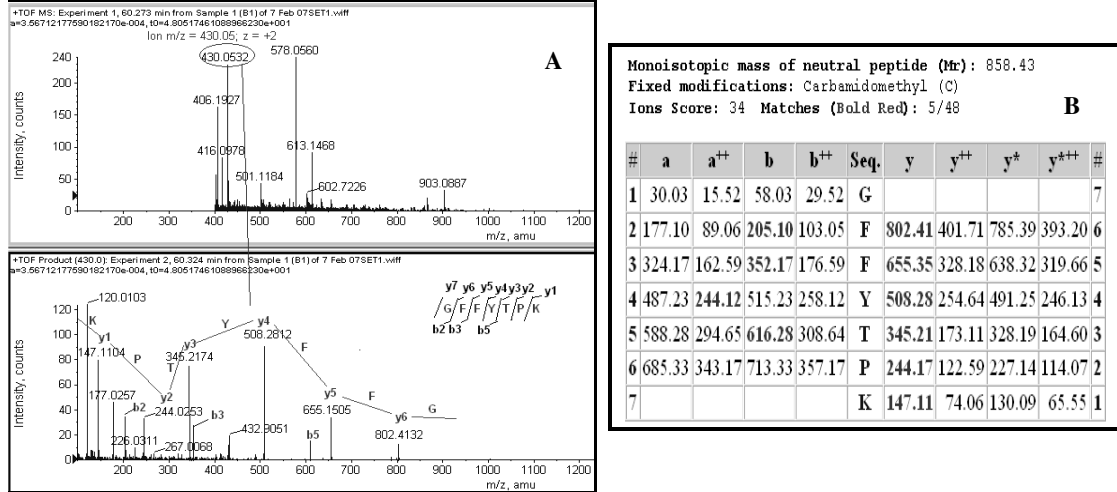
Hình 4. Kết quả nhận diện insulin chuỗi B dạng dung hợp với thioredoxin

Kết quả trên cho thấy, khi phân tích nhận dạng bằng khối phổ trong băng biểu hiện (khoảng 23 kDa) có hai protein được phát hiện là thioredoxin và chuỗi B của insulin có gắn đuôi His, kết quả này phù hợp với thiết kế vector lí thuyết. Điểm số (score) của thioredoxin và chuỗi B insulin lần lượt là 177 và 65 (hình 5).



Hình 5. Điểm số nhận diện protein theo thuật toán Mowse

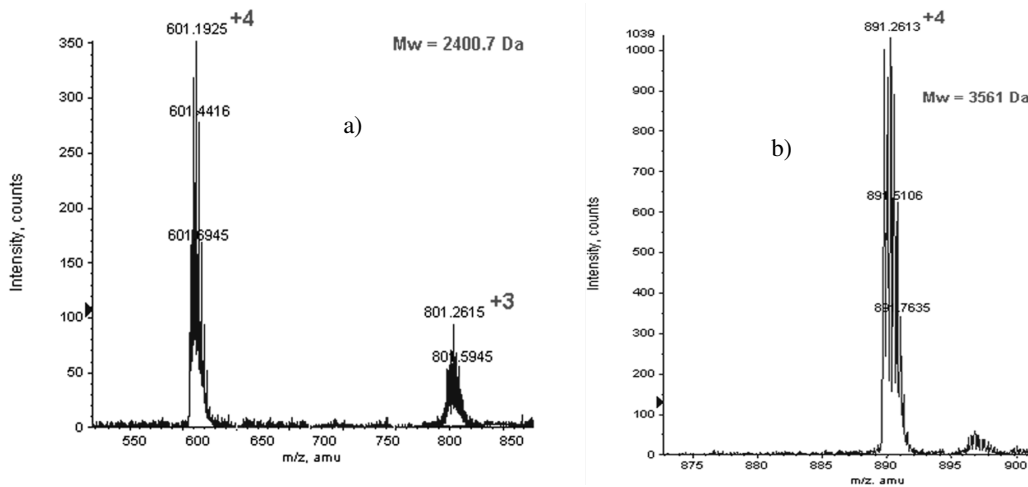
Trình tự amino acid chuỗi B insulin cho thấy chỉ có hai điểm cắt (tại vị trí B22 và B29) khi thủy phân bằng trypsin, tạo thành hai mảnh peptide là GFFYTPK và GFFYTPKT. Khi phân tích khối phổ trình tự của hai mảnh này đều được phát hiện (màu đỏ đậm).



Hình 6. A. Phổ khối của ion m/z = 430,05 amu; B. Giải trình tự peptide GFFYTPK

Phân tích phổ khối trình tự mảnh ion peptide GFFYTPK, chúng tôi nhận thấy mảnh này có giá trị m/z = 430,05 amu, z = +2 (m = 858,43 Da). Quá trình phân mảnh và giải trình tự ion peptide GFFYTPK được minh họa như hình 6.

3.4. Xác định khối lượng của insulin chuỗi A và B



Hình 7. Xác định khối lượng phân tử của insulin chuỗi A (a) và chuỗi B (b) bằng phương pháp ESI-Q-TOF

Chuỗi A và B của insulin dưới dạng dung hợp được xử lí với enterokinase nhằm loại bỏ thioredoxin và đuôi histidine. Sau khi xử lí, hai chuỗi A và B của insulin được xác định khối lượng bằng hệ thống khối phổ ESI-Q-TOF (Electrospray-Quadruple-Time of Flight) (Hình 4). Phổ TOF-MS cho thấy chuỗi A và B lần lượt có kích thước là 2400.7 Da và 3561 Da, kết quả này phù hợp với lí thuyết do chuỗi A có 21 aminoacid và chuỗi B có 31 amino acid và trong tế bào *E. coli* nhóm SH của cystein ở dạng tự do. Như vậy, với phương pháp này chúng tôi đã khẳng định được sự biểu hiện của insulin tái tổ hợp (chuỗi A, B), và xác định được khối lượng của hai chuỗi riêng biệt.

4. KẾT LUẬN

Chuỗi A và B của insulin tái tổ hợp đã được biểu hiện trong tế bào *E. coli* chủng BL21(DE3). Các kết quả xác định bằng khối phổ đã khẳng định chuỗi A và B của insulin tái tổ hợp đã được biểu hiện phù hợp với tính toán lí thuyết và thiết kế ban đầu.

Lời cảm ơn. Công trình được thực hiện tại Phòng Thí nghiệm Trọng điểm Công nghệ Gen, Viện Công nghệ Sinh học, Viện Khoa học và Công nghệ Việt Nam.

TÀI LIỆU THAM KHẢO

1. Bentley G. E., Dodson G. D., Hodgkin D., and Mercola D. - Structure of insulin in 4-zinc insulin, *Nature* **261**(5556) (1976) 166-168.
2. Dunn M. F. - Zinc-ligand interactions modulate assembly and stability of the insulin hexamer - a review, *BioMetals* **18** (4) (2005) 295–303.
3. Farmer A., Wade A., Goyder E. - Impact of self monitoring of blood glucose in the management of patients with non-insulin treated diabetes: open parallel group randomised trial, *BMJ* **335** (2007) 132.
4. Harris R., Donahue K., Rathore S. S., Frame P., Woolf S. H., Lohr K. N. - Screening adults for type 2 diabetes: a review of the evidence for the U.S. Preventive Services Task Force, *Ann. Intern. Med.* **138** (3) (2003) 215–29.
5. Louise G. J. - Diabetes, Atlanta: Centres for Disease Control, 2007.
6. Pappin D. J., Hojrup P., Bleasby A. J. - Rapid identification of proteins by peptide-mass fingerprinting, *Curr. Biol.* **3** (6) (1993) 327-332.
7. Pugliese G., Zanuso S., Alessi E., Simonelli P., Fallucca S., Fallucca F., Balducci S. - Self glucose monitoring and physical exercise in diabetes, *Diabetes Metab. Res. Rev.* **25** (2009) 11-17.
8. Rother K. I. - Diabetes Treatment — Bridging the Divide, *N. Engl J. Med.* **356** (15) (2007) 1499–1501.
9. Steiner D. F. and James D. E. - Cellular and molecular biology of the beta cell, *Diabetologia* **35** (2) (1992) S41-S48.
10. Tierney L. M., McPhee S. J., Papadakis M. A. - Current medical Diagnosis & Treatment, International edition, New York: Lange Medical Books/McGraw-Hill. pp. 1203–15. ISBN 0-07-137688-7, 2002.

11. Tran T. T., Nguyen T. M. P., Nguyen B. N., Phan V. C. - Changes of serum glycoproteins in lung cancer patients, *Journal of Proteomics and Bioinformatics JPB* **1** (1) (2008) 15-21.
12. www.annals.org/cgi/pmidlookup?view=long&pmid=12558362. Retrieved on 2008-07-19.
13. www.cdc.gov/nccdphp/publications/aag/ddt.htm. Retrieved on 2007-09-14
14. Xu B., S. Q. Hu, Y. C. Chu, S. Wang, R. Y. Wang, S. H. Nakagawa, P. G. Katsoyannis and M. Weiss - Diabetes-associated mutations in insulin identify invariant receptor contacts, *Diabetes* **53** (6) (2004) 1599-602.

SUMMARY

STUDYING ON EXPRESSION AND IDENTIFICATION OF HUMAN INSULIN IN *E. coli*

Insulin is a hormone that produced by beta cells in the pancreas and regulated the glucose blood levels. Briefly, our strategy for the gene construction and expression is described as follow. pET14bInsA vector contained gene A that encoding for the A peptide chain and pET14bInsB vector contained gene B that encoding for the B peptide chain. These vectors and the expression pET-32c(+) vector were digested by *NcoI* and *BamHI*. The released insulin A/B fragment and the opened plasmid pET-32c(+) were purified and ligated by T4 DNA ligase at 22°C, overnight. The resulting recombinant plasmid pET32cInsA/B was transformed into the *E. coli* BL21(DE3) strain. The expression of the fusion protein was observed by induction with 1 mM IPTG and analysed by SDS-PAGE. Two new protein bands with the size of approximately 22 kDa (equivalent the size of fusion A) and 23 kDa (the size of fusion B) were then excised and digested with trypsin overnight. The results of identifying by NanoLC-MS/MS (Nano Liquid Chromatography – Tandem Mass Spectrometry) showed that the recombinant protein was truly the fusion A(B)-Trx that contained different peptides from chain A/B and thioredoxin. This result is an evidence to confirm the successful construction and expression of chain A/B of insulin in the fusion form in *E. coli* with pET-32c(+) vector.

Key words. Insulin; pET-32 c(+); *E. coli* BL21 (DE3); NanoLC-MS/MS

Địa chỉ:

Nhận bài ngày 12 tháng 8 năm 2008

Bùi Thị Huyền, Phan Văn Chi
Viện Công nghệ sinh học, Viện KH&CN VN.
Lê Văn Phùng,
Đại học Y Hà Nội.