Protein Expression and Purification 60 (2008) 97–102

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Protein expression and purification of human Zbtb7A in *Pichia pastoris* via gene codon optimization and synthesis

Hui Wang [a], Qingling Wang [a], Feifan Zhang [a], Yahong Huang [a], Yulian Ji [a], Yayi Hou [a,b,*]

[a] *Immunology and Reproduction Biology Lab, Medical School & State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, PR China*
[b] *Jiangsu Key Laboratory of Molecular Medicine, Nanjing University, Room 2011, Mengminwei Building, 22# Hankou Road, Nanjing, Jiangsu 210093, PR China*

ARTICLE INFO

Article history:
Received 28 July 2007
and in revised form 24 February 2008
Available online 30 March 2008

Keywords:
Human Zbtb7A
Codon optimization
P. pastoris
Expression
Purification

ABSTRACT

Human Zbtb7A was proved to be an important molecular switch in oncogenesis. However, it is difficult to obtain its protein expression in prokaryotic system, due to high G + C content and rare codons in *zbtb7a* gene. Therefore, to further research the function and application of this protein, we optimized its coding sequence according to the codon bias of *Pichia pastoris*, synthesized the sequence with two-step PCR and confirmed the accuracy by DNA sequencing. The assembled fragment was introduced into *P. pastoris* expression vector pPIC9K and the resultant plasmid pPIC9K–*zbtb7a*–his_6 was transformed into the *P. pastoris* strain GS115 by electroporation. The products of the transformants induced by methanol were analyzed by 10% SDS–PAGE and identified by Western Blot assay. The expression conditions of the selected transformant were optimized. Additionally, a two-step purification protocol was applied to purify the recombinant protein. The results showed that the synthetic coding sequence of human Zbtb7A was successfully obtained and inserted into pPIC9K vector. Human Zbtb7A protein was expressed in *P. pastoris* and identified by western blot. The optimal conditions for its expression in *P. pastoris* were under a final concentration of 1% methanol and a time-course of 4d. Through the two-step purification, Zbtb7A protein was purified in high purity and its production reached up to as high as 18 mg/L. These results indicated that an effective procedure for expressing and purifying human Zbtb7A in *P. pastoris* was established.

Introduction

Human Zbtb7A (zinc finger and BTB domain containing 7A) encoded by *zbtb7a* gene is also known as LRF [1], OCZF [2], FBI-1 [3] and Pokemon [4]. It was originally identified as a protein that binds specifically to a HIV-1 promoter element [5]. Zbtb7A plays a critical role in cellular transformation mediated by oncogenes and tumorigenesis. It was found that the protein was overexpressed in a subset of T-cell lymphomas, diffuse large B-cell lymphoma (DLBCL[1]), follicular lymphoma (FL) and non-Hodgkin's lymphoma (NHL) [4,6]. Recently it was also found that Zbtb7A could oppose the Notch signaling pathway, which suppresses malignant human B cell [6]. These indicate that Zbtb7A is potentially a therapeutic target for human cancer therapy. However, Zbtb7A protein is not easily obtained so as to limit further research to the protein. Previously, we had tried some *Escherichia coli* expression vectors for expressing this protein in prokaryotic system. We failed to obtain visible expression, although we optimized many factors, which might influence the expression such as growth media, induction time and temperature, preferential codons for host strain, and so on.

Pichia pastoris, a methylotrophic yeast, can utilize methanol as sole carbon source for its growth and propagation and is an effective host in the production of heterologous protein [7,8]. In order to obtain the protein of human Zbtb7A for further research of its function and application, in the present study, we performed, for the first time, the expression of human Zbtb7A in the *P. pastoris* yeast GS115 by optimizing the coding sequence of Zbtb7A according to the codon bias of *P. pastoris*.

Materials and experiments

Materials

Escherichia coli DH5α, *P. pastoris* expression vector pPIC9K and *Pichia* host strains GS115 (his^-mut^+) were stored by our laboratory. PrimeSTAR™ DNA polymerase, DNA marker, T4 DNA ligase, pMD18-T simple vector, restriction enzyme SnaBI, SalI and EcoRI were purchased from Takara Biotechnology Co., Ltd. Protein molecular weight marker was obtained from Ferments USA. PNGase F was purchased from New England Biolabs Inc. Yeast extract, tryptone power, yeast nitrogen base (YNB) and D-sorbitol were pur-

* Corresponding author. Address: Jiangsu Key Laboratory of Molecular Medicine, Nanjing University, Room 2011, Mengminwei Building, 22# Hankou Road, Nanjing, Jiangsu 210093, PR China. Fax: +86 25 83686441.
E-mail address: yayihou@niu.edu.cn (Y. Hou).

[1] *Abbreviations used:* DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; NHL, non-Hodgkin's lymphoma; YNB, yeast nitrogen base; ECL, enhanced chemiluminescence; MW, molecular weight.

1046-5928/$ - see front matter Crown copyright © 2008 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.03.017

chased from Bio Basic Inc. Q-Sepharose Fast Flow was purchased from Amersham Biosciences. His·Bind Fractogel (M) was obtained from Novagen. All the other chemicals were of analytical grade.

Synthesis of optimized coding sequence of human Zbtb7A and construction of recombinant vector pPIC9K–zbtb7a

In order to synthesize coding sequence of human Zbtb7A, 66 optimized oligonucleotides were designed based on the protein sequence of human Zbtb7A (GenBank Accession No. AAH84568) according to the codon bias of *P. pastoris* (http://www.kazusa.or.jp/codon). The optimized oligonucleotides were synthesized by Shanghai Invitrogen Biotechnology Co., Ltd (Table 1). One micro liter of each optimized oligonucleotide at a concentration of 1 μg/μl was mixed together and diluted with water to final concentration at 10 ng/μl for each optimized oligonucleotide.

The optimized coding sequence was assembled by two-step PCR [9,10]. Firstly, to obtain the template of the optimized coding sequence of human Zbtb7A, a first-step PCR containing 1 μl of the mixer, 5 μl of 10× PrimeSTAR Buffer, 4 μl of 2.5 mM dNTPs, 0.5 μl PrimeSTAR™ DNA polymerase and 39.5 μl of water was performed. The first-step PCR began with a step of denaturation (98 °C for 1 min), subsequently followed by 30 cycles of denaturation (98 °C for 15 s), annealing (60 °C for 15 s), and extension (72 °C for 30 s), and finally followed by an incubation step (72 °C for 5 min). Secondly, for target fragment amplification, 1 μl of the product resulted from the first-step PCR was used for the second-step PCR with forward primer (5′ TTTACGTAGCAGGTGGAGTTGATGGTCCTA 3′) and reverse primer (5′ TTTGAATTCTTATTAATGATGATGATGATGGTGGG CCAATCCAGCAGT3 ′). A SnaBI restriction site was inserted into the forward primer, while an EcoRI restriction site linked with two stop codons (TAATAA) and a His·tag coding sequence was inserted into the reverse primer. The same protocol as the first-step PCR was conducted. After A-Tailing and agarose gel DNA fragment recovery, the target product of the second-step PCR was cloned into PMD 18-T simple vector generating plasmid named PMD18T–*zbtb7a–his*$_6$. The plasmid was sequenced by Shanghai Invitrogen Biotechnology Co., Ltd. After the sequence was confirmed to be correct, the target fragment was digested from PMD18T–*zbtb7a–his*$_6$ by restriction enzymes of SnaBI and EcoRI and then inserted into pPIC9K so as to yield recombinant vector pPIC9K–*zbtb7a–his*$_6$.

Transformation of P. pastoris and selection of recombinant clones

Five micrograms of recombinant vector pPIC9K–*zbtb7a–his*$_6$ or parent vector pPIC9K was linearized by Sal I and purified for transformation. Transformation of the *P. pastoris* strain GS115 (His^-Mut^+) was performed using the electroporation method according to the instruction of Multi-copy *Pichia* Expression Kit (version F). To analyze the copy number of the clones, about 500 transformants were recovered and replica-plated onto YPD medium (1% yeast extract, 2% tryptone power, 2% dextrose, 2% agar) containing 0.25, 0.5, 0.75 and 1.5 mg/ml G418 (Amresco). The Mut^+ and Mut^s phenotypes of the transformants were evaluated by spotting them on minimal medium agar plates (1.34% yeast nitrogen base with ammonium sulfate without amino acids (YNB) (BBI), 4×10^{-5}% D-biotin, 1.5% agar) containing either 2% dextrose or 0.5% methanol as the primary carbon source. The clone screening was performed using PCR analysis with *zbtb7a* specific primer (5′ TTTACGTAGCAGGTGGAGTTGATGGTCCTA 3′) and 3′AOX1 primer (5′ GCAAATGGCATTCTGACATCC 3′).

Extracellular expression of human Zbtb7A in P. pastoris

The selected recombinant strains (His^-Mut^+) integrated with pPIC9K (GS115/pPIC9K) or pPIC9K–*zbtb7a–his*$_6$ (GS115/pPIC9K–*zbtb7a–his*$_6$) were cultured in 5 mL BMGY media (1% yeast extract, 2% tryptone power, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, 1% glycerol) at 28 °C until OD_{600} reached approximately 6. To induce expression of the recombinant protein, the strains were harvested by centrifugation at 2000*g* for 5 min at room temperature and then resuspended to approximately 1 at OD_{600} in BMMY (1% yeast extract, 2% tryptone power, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, 0.5% methanol). One hundred percent methanol was added to a final concentration of 0.5% every 24 h to maintain the induction. At

Table 1
Optimized oligonucleotides for synthesis of coding sequence of human Zbtb7A

Number	Oligonucleotides
1	TTTACGTAGCAGGTGGAGTTGATGGTCCTA
2	ACTGGAATGATCAGGAAATGGGATTCCGATAGGACCATCAACTCCACCTGCTAC
3	TCGGAATCCCATTTCCTGATCATTCCAGTGACATTCTGTCAGGCCTGAATGAACAACG
4	GACAACATCACACAACAGACCTTGGGTACGTTGTTCATTCAGGCCTGACAGAATGTC
5	TACCCAAGGTCTGTTGTGTGATGTTGTCATTCTGGTTGAAGGTCGTGAATTTCCAACTC
6	TGAGAACAGGCAGCCAAGACACTTCTATGAGTTGGAAATTCACGACCTTCAACCAGAAT
7	ATAGAAGTGTCTTGGCTGCCTGTTCTCAATATTTCAAGAAGCTGTTCACCAGTGGTGC
8	TCTCGTACACGTTCTGTTGATCGACAACTGCACCACTGGTGAACAGCTTCTTGAAATAT
9	AGTTGTCGATCAACAGAACGTGTACGAGATCGACTTCGTTAGTGCCGAAGCCTTG
10	GCGGTATAGGCGAAATCCATCAGAGCAGTCAAGGCTTCGGCACTAACGAAGTCGA
11	ACTGCTCTGATGGATTTCGCCTATACCGCCACCTTGACCGTTAGTACAGCCAATGTC
12	CAAACGGGCTGCTGACAGGATATCACCGACATTGGCTGTACTAACGGTCAAGGTG
13	GGTGATATCCTGTCAGCAGCCCGTTTGCTGGAAATTCCAGCCGTTAGTCATGTCT
14	CAAGATCTGACGATCCAGCAAGTCAGCACAGACATGACTAACGGCTGGAATTTCCAG
15	GTGCTGACTTGCTGGATCGTCAGATCTTGGCAGCCGATGCTGGTGCCGAT
16	GTCAACCAGGTCCAGTTGACCTGCATCGGCACCAGCATCGGCTGC
17	GCAGGTCAACTGGACCTGGTTGACCAAATCGATCAACGTAACTTGCTGAGAGCC
18	TTACTCTGGAAGAACTCCAGGTACTCCTTGGCTCTCAGCAAGTTACGTTGATCGATTTG
19	AAGGAGTACCTGGAGTTCTTCCAGAGTAATCCAATGAACTCCTTGCCTCCAGC
20	GAAACTTGCAGCAGCTGCGGCAGCTGCTGGAGGCAAGGAGTTCATTGGA
21	AGCTGCCGCAGCTGCTGCAAGTTTCCCTTGGTCTGCCTTTGGTGCCAG
22	GCTTCTTTGGTAGCATCCAGATCGTCATCACTGGCACCAAAGGCAGACCAAGG
23	TGATGACGATCTGGATGCTACCAAAGAAGCCGTTGCAGCCGCTGTTGCCG
24	GTCCAATCCATTACAATCACCAGCAGCAACTGCGGCAACAGCGGCTGCAACG
25	CAGTTGCTGCTGGTGATTGTAATGGATTGGACTTCTATGGACCTGGTCCACCTGCC
26	ATCGCCATCACCTGTAGGTGGACGTTCGGCAGGTGGACCAGGTCCATAGAA
27	GAACGTCCACCTACAGGTGATGGCGATGAAGGTGACTCTAATCCAGGTTTGTGGC
28	CGGTAGGTGCATCTTCATCACGTTCTGGCCACAAACCTGGATTAGAGTCACCTTC
29	CAGAACGTGATGAAGATGCACCTACCGGTGGACTGTTTCCACCTCCAGTTGC
30	CCATTCTGGGTTGCAGCTGGAGGTGCAACTGGAGGTGGAAACAGTCCAC
31	ACCTCCAGCTGCAACCCAGAATGGTCATTATGGTAGAGGCGGTGAAGAGGAAG
32	GGTGCAGCTTCAGACAAACTGGCAGCTTCCTCTTCACCGCCTCTACCATAATGA
33	CTGCCAGTTTGTCTGAAGCTGCACCTGAACCAGGTGATAGTCCTGGATTCCTGT
34	GTCTTCACCTTCGGCTGCACCTGACAGGAATCCAGGACTATCACCTGGTTCA
35	CAGGTGCAGCCGAAGGTGAAGACGGCGATGGTCCAGATGTTGATGGTTTG
36	CATTTGCTGCAACAGGGTACTGGCTGCCAAACCATCAACATCTGGACCATCGCC
37	GCAGCCAGTACCCTGTTGCAGCAAATGATGTCCAGTGTTGGTCGTGCAGGAG
38	CTACTTTCTTCATCAGAATCACCAGCTGCGGCTCCTGCACGACCAACACTGGACAT
39	CCGCAGCTGGTGATTCTGATGAAGAAAGTAGAGCCGACGATAAGGGTGTCATGGAC
40	CATCATGGGCACCAGAGAAATATTTCAGATAATAGTCCATGACACCCTTATCGTCGGCT
41	TACTACCTGAAATATTTCTCTGGTGCCCATGATGGCGATGTCTACCCTGCATGGAGTC
42	GGCTTTAGCACCGATCTTCTTCTCAACTTTCTGACTCCATGCAGGGTAGACATCGC
43	AGAAAGTTGAGAAGAAGATCCGTGCTAAAGCCTTCCAGAAGTGTCCAATCTGTGAGAAG
44	CTAGGCAACTTACCTGCACCTTGAATGACCTTCTCACAGATTGGACACTTCTGGAA
45	GTCATTCAAGGTGCAGGTAAGTTGCCTAGACACATTCGAACCCATACAGGTGAGAAGC
46	CGAACCTTGCAGATGTTGCACTCGTAAGGCTTCTCACCTGTATGGGTTCGAATGTGT
47	CTTACGAGTGCAACATCTGCAAGGTTCGTTTCACCAGACAAGACAAGCTGAAAGTCCAC
48	GGTTTCTCACCGGTATGCTTACGCATGTGGACTTTCAGCTTGTCTTGTCTGGTGAAA
49	ATGCGTAAGCATACCGGTGAGAAACCATATTTGTGTCAACAGTGTGGTGCTGCATTTG
50	GTGGTTCTTCAGGTCGTAGTTGTGGGCAAATGCAGCACCACACTGTTGACACAAATAT
51	CCCACAACTACGACCTGAAGAACCACATGCGTGTTCATACTGGTTTGAGACCTTATCA
52	CGAAGGTCTTGCAGCAGGAATCACATTGATAAGGTCTCAAACCAGTATGAACACGCAT
53	ATGTGATTCCTGCTGCAAGACCTTCGTCCGAAGTGATCATCTGCATCGTCACTTG
54	GGAACACCGTTACAGCCGTCTTTCTTCAAGTGACGATGCAGATGATCACTTCGGA
55	AAGAAAGACGGCTGTAACGGTGTTCCATCACGTAGAGGTCGTAAACCTAGAGTTCG
56	GGACTAGGATCTGGTGCTCCACCACGAACTCTAGGTTTACGACCTCTACGTGAT
57	TGGTGGAGCACCAGATCCTAGTCCAGGTGCTACAGCCACTCCTGGTG
58	CTGGACTAGAAGGTTGGGCTGGTGCACCAGGAGTGGCTGTAGCACCT
59	CACCAGCCCAACCTTCTAGTCCAGATGCCCGTAGGAATGGCCAAGAG
60	CTTCGTCTTCATCTTCGTCTTTGAAGTGCTTCTCTTGGCCATTCCTACGGGCAT
61	AAGCACTTCAAAGACGAAGATGAAGACGAAGATGTTGCCTCTCCTGATGGTCTGG
62	CACCAGCTCCTGCAACATTCAAACGACCCAGACCATCAGGAGAGGCAACAT
63	GTCGTTTGAATGTTGCAGGAGCTGGTGGAGGTGGAGATAGTGGTGGAGGTC
64	GAAGTTACCATCGGTTGCAGCACCTGGACCTCCACCACTATCTCCACCTC
65	CAGGTGCTGCAACCGATGGTAACTTCACTGCTGGATTGGCCCACCATC
66	TTTGAATTCTTATTAATGATGATGATGATGGTGGGCCAATCCAGCAGT