



Current Issues in Pharmacy and Medical Sciences

Formerly ANNALES UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA, SECTIO DDD, PHARMACIA

journal homepage: <http://www.curipms.umlub.pl/>



Obtaining the *tbf* gene which encodes immunodominant epitopes of pathogenic cholera strains

YULIA BAZARNOVA*^{ORCID}, TATIANA BOLOTNIKOVA, EKATERINA ARONOVA

Higher School of Biotechnology and Food Technology, Peter the Great Saint Petersburg State Polytechnical University

ARTICLE INFO

Received 06 November 2018
Accepted 08 February 2019

Keywords:

candidate cholera vaccine,
hybrid *tbf* gene,
Sanger sequencing,
protein design,
TBF recombinant protein,
immunodominant epitopes,
Vibrio cholerae.

ABSTRACT

We experimentally carried out the synthesis of the *tbf* gene that encodes highly-immunogenic epitopes of pathogenic cholera strains, through the cloning of the *tbf* gene in pGEM-T Easy plasmid. Moreover, we tested the hybrid gene sequence for absence of mutations, using the Sanger sequencing. We also calculated the nucleic sequence of the *tbf* gene. The obtained results have both scientific and practical significance.

INTRODUCTION

Cholera is an intestinal infection disease that rapidly dehydrates the human organism. It is caused by the bacterium of the *Vibrio cholerae* species, namely, toxigenic serogroup O1 and less commonly, O139. Cholera is characterized by a fecal-oral transmission route (more rarely human-to-human direct transmission) affecting the small bowel, and symptoms that include watery diarrhea, vomiting, rapid loss of water and electrolytes with various ranges of dehydration up to hypovolemic shock.

The cholera toxin (exotoxin) is responsible for the disease symptoms and its rapid evolution, and is a key factor of its pathogenicity.

As of today, scientists have come to a conclusion that oral immunization is the most efficient immune protection from cholera. Along with the developed attenuated and inactivated oral anti-cholera vaccines, recombinant vaccines based on epitopes of bacterial antigens occupy leadership positions for cholera treatment. These vaccines induce a pronounced immune response and depress the bacterial transcytosis, which is a key factor for the disease evolution.

The object under study in this work is the *tbf* gene that encodes immunodominant epitopes of proteins of pathogenic cholera strains TcpA and B(rBS), as well as an area which provides antigen penetration through the epithelium of the stomach wall.

The TcpA protein is a component of a toxin-coregulated adhesion pilus, and forms polymers from subunits of 20.5 kDa [1]. This protein is a factor in pathogen colonization [2]. Toxin-coregulated piles are important protective antigens, which possess well-pronounced protective properties [3].

The subunit of cholera exotoxin B (rBS) is a pentamer, each sub-subunit of which is presented as a polypeptide chain consisting of 103 amino acid residues [4]. The protein contains antigenic determinants [5], which induce neutralization antibodies.

The neonatal FcRn-receptor is a receptor of the Fc fragment that forms a heterodimer by non-covalent interaction of alpha-chain with beta-2 microglobulin [6]. At oxygen hydrolysis, the pH receptor bounds the segment that connects CH2- and CH3-domains of antibodies of the IgG class [7]. After antibody bonding, the receptors activate a cascade of reactions for pathogen elimination by antibody-dependent phagocytosis [8].

Thus, the vaccine, which contains purified immunogenic protein, is stable and safe, its chemical properties are well-studied, and it does not contain any additional proteins and nucleic acids that might cause undesirable effects in the host organism.

The aim of this work is to inform readers of a means of replicating the *tbf* gene that encodes immunodominant epitopes of pathogenic cholera strains.

The work was performed at Laboratory No. 6 "Genetic engineering of vaccines" of the State Research Institute of

* Corresponding author

e-mail: j.bazarnowa2012@yandex.ru

Highly Pure Biopreparations of the Russian Federal Medical-Biological Agency (FMBA), Saint-Petersburg, Russia.

MATERIALS AND METHODS

Objects under study

The gene, which encodes TBF recombinant protein, was obtained through chemical synthesis. Synthesis of sequence of the calculated gene that encodes the TBF protein, was performed by applying the polymerase chain reaction technique, using overlapping oligonucleotides. These oligonucleotides were synthesized by employing the DNA synthesizer ASM-800 (BIOSSET, Russia). The main requirements for primers were the following: length not more than 60 nucleotides, and hybridization segments not more than 20 nucleotides. Beyond the aforementioned, there should not be long segments with repeated G or C. In total, for the synthesis of the gene that encodes the TBF protein of 2085 nucleotide pairs, we used 77 primers. The synthesized sequence was obtained from agarose gel, by means of electrophoresis, and cloned in the plasmid vector pGEM-T Easy. We performed the blunt-end cloning. After sequencing utilizing the capillary sequencer Applied Biosystems 3500/3500xL Genetic Analyzer (Applied Biosystems, USA), the fragments were amplified in a thermocycler for amplification – C1000 ThermalCycler (Bio-Rad, USA). In the terminal gene segments, we included restriction sites XhoI and NdeI for further cloning in plasmid pET28a (+).

The vector map is presented in Figure 1.

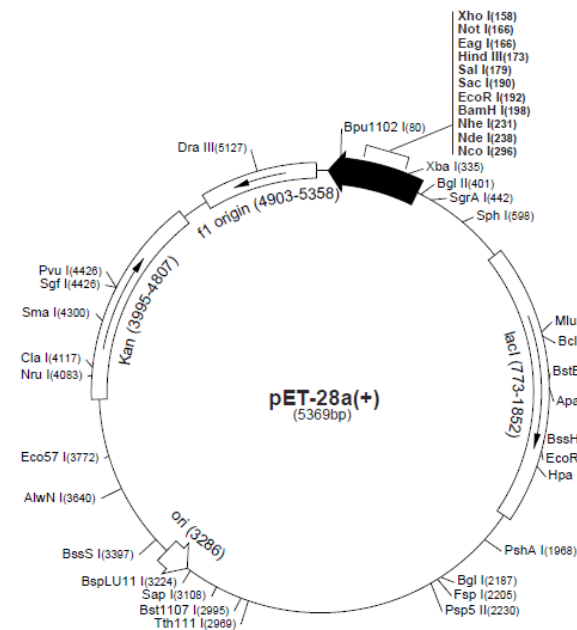


Figure 1. Circular pET28a (+) plasmid map [9]

The expression plasmid comprises the replication origin, promoter of T7 phage polymerase, lac-operator, kanamycin-resistance gene, start codon for translation of cloned fragments and a fragment which encodes polyhistidine located in the reading frame at the N-end of the sequence. Hence, any nucleic sequence cloned in the vector is expressed as a polyhistidine-tagged protein as a matter of convenience

for its further purification by means of immobilized metal affinity chromatography. For lac-operon operation, the plasmid contains a fragment that encodes the lactose repressor lacI.

The gene cloning in the vector was performed via the XhoI and NdeI restriction sites.

For the genetic engineering, we used *E. coli* DH10B/R (Gibco BRL, USA) cells with F-mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80dlacZ Δ M 15 Δ lacX74 deoR recA1 endA1 araD139 Δ (ara,leu)769 galUgalK λ - rpsLnpG genotype.

For the expression of the gene that encodes the TBF protein, we employed *E. coli* BL21 Star (DE3) cells with the F- ompThsdSB (rB-mB-) galdcM rne131 (DE3) genotype that contains in its genome, λ De3 lysogen and *rne131* mutation. The mutated gene *rne* (*rne131*) encodes reduced RNAase E, which decreases the intracellular destruction of mRNA. This results in the enhancement of its fermentation stability. Of note, lon- and ompT- mutations in protease genes allow the possibility to obtain large amounts of non-proteolyzed recombinant proteins.

Research methods

Synthesis of sequence of the calculated gene that encodes TBF protein, was performed by means of the polymerase chain reaction technique and the use of overlapping oligonucleotides. These oligonucleotides were synthesized by means of the DNA synthesizer ASM-800 (BIOSSET, Russia). The main requirements for the primers were the following: length not more than 60 nucleotides and hybridization segments not more than 20 nucleotides. Moreover, there should not be long segments with repeated G or C. In total, for the synthesis of gene that encodes the TBF protein of 2085 nucleotide pairs, we used 77 primers. The synthesized sequence was obtained from agarose gel, through electrophoresis, and cloned in the plasmid vector pGEM-T Easy. We performed the blunt-end cloning. After sequencing via the capillary sequencer Applied Biosystems 3500/3500xL Genetic Analyzer (Applied Biosystems, USA), the fragments were amplified in a thermocycler for amplification C1000 ThermalCycler (Bio-Rad, USA). In the terminal gene segments, we included restriction sites XhoI and NdeI for further cloning in plasmid pET28a(+).

For the design of TBF recombinant protein, we accessed NCBI databases (www.ncbi.nlm.nih.gov).

RESULTS

For the design of the TBF recombinant protein, we used the most protective epitopes of bacterial antigens: subunit of B cholera exotoxin, protein of cholera A pilus, and domain, which is a ligand to Fc-receptors at the stomach wall. The selected consensus segments, common for *V. cholerae* pathogenic strains, were analyzed for T and B cell epitopes. During the analysis, we employed the following software: Bepipred Linear Epitope Prediction 2.0 and IEDB Analysis Resource. The obtained amino acid sequence of TBF recombinant protein (695 amino acid residues) is presented in Figure 2.

Obtaining the *tbf* gene which encodes immunodominant epitopes of pathogenic cholera strains

```

10 20 30 40 50 60
MQLLKQLKFK KFVKEBHDKK TGQEGMTLLE VIIVLGMGVP VSAGVVTLAQ RAIDSQIMTK
70 80 90 100 110 120
AAQSLNSIQV ALTYTYRGLG NYPATADATA ASKLTSLGLV LGKISSDEAK NPFNGTMMNI
130 140 150 160 170 180
FSFPRNAAN KAF AISVDGL TQAQCKTLIT SVGDMFPYIA IKAGGAVALA DLGDFENSAA
190 200 210 220 230 240
AAETGVGVK SIAPASKNLD LTNITHVEKL CKGTAPFVVA FGNSSGGGGG GGMIKLKFVG
250 260 270 280 290 300
FFTLLSSAY AHGTPQNTID LCAEYHNTQI YTLNDKIFSY TESLAGKREM AIITFKNGAI
310 320 330 340 350 360
FQVEVPGSQH IDSQKKAIER MKDTRLRIAYL TEAKVEKLCV WNNKTPHAI AISMANGGGG
370 380 390 400 410 420
GGGGGASTKG PSVFPLAPSS KSTSGGTAAL GCLVKDYFPE PVTVSWNSGA LTSGVHTFPA
430 440 450 460 470 480
VLQSSGLYSL SSVVTVPSSS LGTQTYICNV NHKPSNTKVD KKVPEKSCDK THTCPPECPAP
490 500 510 520 530 540
ELLGGPSVFL FPKPKDKTLM ISRTEPEVTCV VVDVSHDEPE VKFNWYVDGV EVHNAKTKPR
550 560 570 580 590 600
EEQYNSYRV VSVLTVLHGD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQYVITL
610 620 630 640 650 660
PSRDELTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTIV
670 680 690
DKSRWQQGNV FSCSVMHEAL HNHYTQKSL SLPKG
    
```

Figure 2. TBF protein amino acid sequence

The nucleatic sequence that encodes the TBF protein (with length of 2085 nucleotide pairs) was obtained via calculation (Fig. 3).

```

10 20 30 40 50 60
ATGCAGCTGC TGAACACGCT GTTTAAAAAA AAATTGTGTA AAGAAGAACA TGATAAAAAA
70 80 90 100 110 120
ACCGGCCAGG AAGGCATGAC CCTGCTGGAA GTGATTATGT TGCTGGGCAT TATGGCCGTG
130 140 150 160 170 180
GTGAGCGCGG CGTGGTGTAC CCTGGCGCAG CGCGCGATTG ATAGCCAGAT TATGACCAAAA
190 200 210 220 230 240
GCGGCGCAGA GCCTGAACAG CATTACAGTG CCGCTGACCC AGACCTATCG CCGCCTGGCC
250 260 270 280 290 300
AACTATCCGG CGACCAGCGA TCGCAGCCGG CCGAGCAAAC TGACCAGCCG CCTGGTGAGC
310 320 330 340 350 360
CTGGGCAAAA TTAGCAGCGA TGAAGCGAAA AACCCGTTTA ACGGCACCAA CATGAACATT
370 380 390 400 410 420
TTTAGCTTTC CGCGCAACGC GCGCGCGAAC AAAGCGTTTG CGATTAGCGT GGATGGCCCTG
430 440 450 460 470 480
ACCGAGCCGC AGTGCAAAAC CCTGATTACC AGCGTGGGCG ATATGTTTCC GTATATTGCG
490 500 510 520 530 540
ATTAAGCGGG GCGGCGCGGT GCGGCTGGCG GATCTGGGCG ATTTTGAAAA CAGCGCGGCG
550 560 570 580 590 600
GCGGCGGAAA CCGGCGTGGG CGTGATTAAA AGCATTGGCC CGGCGAGCAA AAACCTGGAT
610 620 630 640 650 660
CTGACCAACA TTACCATGTT GAAAAAAGCT TGCAAAGGCA CCGCGCCGTT TGGCGTGCGG
670 680 690 700 710 720
TTTGCAACA CGGCGCGCGG CCGCGCGCGC GCGCGCATGA TTAACCTGTA ATTTGGCGTG
730 740 750 760 770 780
TTTTTTACCG TGCTGCTGAG CAGCGCGTAT GCGCATGCA CCCCAGAGAA CATTACCGAT
790 800 810 820 830 840
CTGTGCGCGG AATATCATA CACCCAGATT TATACCTGTA ACGATAAAAT TTTTAGCTAT
850 860 870 880 890 900
ACCGAARGCC TGGCGGGCAA ACGCAAAATG GCGATTATTA CCTTTAAAA CCGGCGCGAT
910 920 930 940 950 960
TTTCAGGTGG AAGTGCCGGG CAGCCAGCAT ATTGATAGCC AGAAAAAAGC GATTGAACGC
970 980 990 1000 1010 1020
ATGAAAAGATA CCCTGCGCAT TGCGTATCTG ACCGAAGCGA AAGTGAAAA ACTGTGCGTG
1030 1040 1050 1060 1070 1080
TGGAAACAACA AAACCCGCA TGCGATTGCG GCGATTAGCA TGGCGAARGC CCGGCGCGGC
1090 1100 1110 1120 1130 1140
GGCGGCGGCG GCGGCGGAG CACCAAAAGC CCGAGCGTGT TTCCGCTGGC GCGGCGGAGC
1150 1160 1170 1180 1190 1200
AAAAGCACCA GCGGCGGCGC CCGCGCGCTG GGTCGCTGG TGAAGATTA TTTTCCGGAA
1210 1220 1230 1240 1250 1260
CCGCTGACCG TGAGCTGAAA CAGCGCGCGC CTGACCAACG GCGTGCATAC CTTTCCGCGG
1270 1280 1290 1300 1310 1320
GTGCTGCAGA GCAGCGGCGT GTATAGCCTG AGCAGCGTGG TGACCGTGCC GAGCAGCAGC
1330 1340 1350 1360 1370 1380
CTGGGCACCC AGACCTATAT TTGCAACGTF AACCATAAAC CGAGCAACAC CAAAGTGGAT
    
```

```

1390 1400 1410 1420 1430 1440
AAAAAGTGG ACCGAAAAG CTGCGATAAA ACCCATACCT GCCCGCCGTG CCCGGCCCGG
1450 1460 1470 1480 1490 1500
GAACTGCTGG GCGGCCCCGAG CGTGTCTTCTG TTTCCGCCGA AACCGAAAAG TACCTGATG
1510 1520 1530 1540 1550 1560
ATTAGCCGCA CCCCAGGAGT GACCTGCGTG GTGGTGGATG TGAGCCATGA AGATCCGGAA
1570 1580 1590 1600 1610 1620
GTGAAATTTA ACTGGTATGT GGATGGCGTG GAAGTGCCATA ACGCGAAAAA CAAACCGCGC
1630 1640 1650 1660 1670 1680
GAAGAACAGT ATAACAGCAC CTATCCGCGTG GTGAGCGTGC TGACCGTGCT GCATCAGGAT
1690 1700 1710 1720 1730 1740
TGGCTGAACG GCAAAGAATA TAAATGCAAA GTGAGCAACA AAGCGCTGCC GCGCGCGATT
1750 1760 1770 1780 1790 1800
GAAAAACCA TTAGCAAAAG GAAAGGCCAG CCGCGCGAAC CGCAGGTGTA TACCTGCGG
1810 1820 1830 1840 1850 1860
CCGAGCCGCG ATGAAGTAC CAAAAACCA GTGAGCCGTA CCTGCCTGCT GAAAGCGTTT
1870 1880 1890 1900 1910 1920
TATCCGAGCG ATATTGCGGT GGAATGGGAA AGCAACGCC AGCCGGAATA CAACTATAAA
1930 1940 1950 1960 1970 1980
ACCACCCCGC CGTGTCTGGA TAGCGATGCG AGCTTTTTTC TGTATAGCAA ACTGACCGTG
1990 2000 2010 2020 2030 2040
GATAAAGCC GCTGGCAGCA GGGCAACGTT TTAGCTGCA GCGTATGCA TGAAGCGCTG
2050 2060 2070 2080
CATAACCAT ATACCCAGAA AAGCCTGAGC CTGAGCCCGG GCAA
    
```

Figure 3. *tbf* gene nucleatic sequence

DISCUSSION

In the study, we carried out the synthesis of a calculation model of *tbf* gene sequences with length of 2085 nucleotide pairs by means of the PCR technique and overlapping oligonucleotides that were synthesized via the DNA synthesizer ASM-800 (BIOSSET, Russia). The absence of mutations in the sequence of hybrid gene was indicated through applying the sequencing technique (Fig. 4).

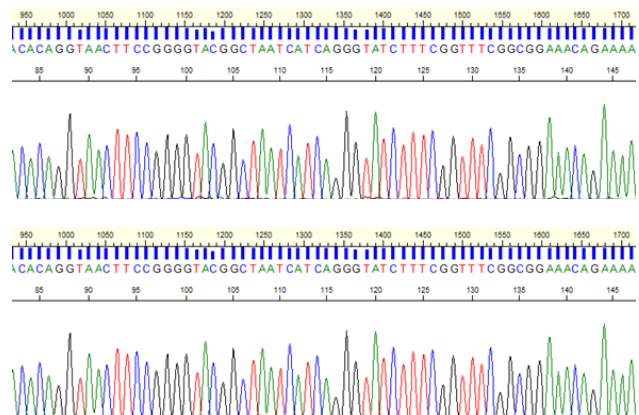


Figure 4. A *tbf* gene chromatogram fragment. Interpretation according to the direct-primer Sanger technique (TAATACGACTCACTATAGGG)

Sequencing of amplified DNA segments was performed according to the Sanger technique.

CONCLUSIONS

Our study was the first to calculate the sequence of nucleotide pairs forming part of *tbf* gene that encodes immunodominant epitopes of pathogenic cholera strains and to synthesized it.

We cloned the *tbf* gene in plasmid pGEM-T Easy and examined the hybrid gene sequence for the absence of mutations, using the Sanger sequencing technique.

The obtained results are of interest for further development of stable and safe recombinant vaccines that contains subunits of cholera toxin B (rBS), protein of cholera A piles (TcpA), as well as the domain that is a ligand to Fc-receptors at the stomach wall (FcL).

ORCID iDs

Julia Bazarnova  <https://orcid.org/0000-0001-9275-7913>

REFERENCES

1. Kiaie S, Abtahi H, Mosayebi G, Alikhani M, Pakzad I. Recombinant toxin-coregulated pilus A (TcpA) as a candidate subunit cholera vaccine. *Iran J Microbiol.* 2014;6:68-73.
2. Megli CJ, Yuen AS, Kolappan S, Richardson MR, Dharmasena MN. Crystal structure of the *Vibrio cholerae* colonization factor TcpF and identification of a functional immunogenic site. *Mol Biol.* 2011;409:146-58.
3. Liljeqvist S. Surface display of the cholera toxin B subunit on *Staphylococcus xylosum* and *Staphylococcus carnosus*. *Appl Environ Microbiol.* 1997;63:2481-8.
4. Sanchez J, Johansson S, Löwenadler B, Svennerholm A, Holmgren J. Recombinant cholera toxin B subunit and gene fusion proteins for oral vaccination. *Research in Microbiology.* 1990;141:971-9.
5. Praetor A, Jones RM, Wong WL, Hunziker W. Membrane-anchored human FcRn can oligomerize in the absence of IgG. *J. Mol. Biol.* 2002;321:277-84.
6. Schroeder HW, Cavacini L. Structure and function of immunoglobulins. *J Allergy Clin Immunol.* 2010;125(2,Suppl 2):S41-52. <https://doi.org/10.1016/j.jaci.2009.09.046>
7. Raghavan M, Bjorkman PJ. Fc receptors and their interactions with immunoglobulins. *Annu Rev Cell Dev Biol.* 1996;12:181-220.
8. Reed SM, Bayly WM, Sellon DC. *Equine internal medicine* (4 edition). St. Louis, Missouri: Elsevier;2018:1566.
9. Addgene: pET28a(+) www.addgene.org/20378/ .