



# BioTechnology

An Indian Journal

FULL PAPER

BTALJ, 12(4), 2016 [175-181]

## Cloning and characterization of a new site specific methyl-directed DNA endonuclease EcoBLI recognizing 5'-G(5mC)NGC-3'/3'-CGN(5mC)G-5'

D.A.Gonchar\*, V.A.Chernukhin, M.A.Abdurashitov, E.V.Kileva, V.S.Dedkov, N.A.Mikhnenkova, E.N.Lomakovskaya, S.G.Udalyeva, S.Kh. Degtyarev  
SibEnzyme Ltd., 2/12, Timakova street, Novosibirsk, 630117, (RUSSIA)  
Danila A.Gonchar, Ph.D; SibEnzyme Ltd., 2/12, Timakova street, Novosibirsk, 630117, (RUSSIA)  
E-mail: gonchar@sibenzyme.ru

### ABSTRACT

A gene coding BisI, site specific 5mC-directed DNA endonuclease recognizing DNA sequence 5'-G(5mC)NGC-3'/3'-CGN(5mC)G-5', was recently identified in the sequenced genome of the strain-producer *Bacillus subtilis* T30. In this work we have undertaken a search of *bisI* gene homologues among the sequenced genomes of enterobacteria. DNA analysis has revealed a small group of highly homologous ORFs with unknown function including one ORF in DNA of well-known strain *E.coli*. This ORF WP 001276099.1 from *E.coli* BL21 (DE3) was amplified and cloned. An obtained recombinant strain *E.coli* pEcoBLI produces MD-endonuclease named EcoBLI. The new enzyme has the same substrate specificity as BisI MD-endonuclease. Thus, ORF WP 001276099.1 from *E.coli* BL21 (DE3) encodes site-specific 5mC-directed DNA-endonuclease EcoBLI recognizing and cleaving DNA sequence as indicated by arrows 5'-G(5mC)<sup>^</sup>NGC-3'/3'-CGN<sup>^</sup>(5mC)G-5<sup>^</sup>.

© 2016 Trade Science Inc. - INDIA

### KEYWORDS

Methyl-directed DNA endonuclease;  
MD-endonuclease;  
Epigenetics;  
Methylome analysis;  
DNA endonuclease gene cloning;  
*Escherichia coli* BL21 (DE3).

### INTRODUCTION

Site specific methyl-directed DNA endonucleases (MD-endonucleases) recognize and cleave DNA sequences with 5-methylcytosines (5mC), and don't cleave unmethylated DNA. Over the past 10 years about 15 prototypes of these enzymes with different recognition sites have been described including BisI<sup>[1]</sup>. BisI recognizes and cleaves DNA sequence 5'-G(5mC)<sup>^</sup>NGC-3'/3'-CGN<sup>^</sup>(5mC)G-5<sup>^</sup> as

indicated by arrows. Recently a genome of *Bacillus subtilis* T30 - producer of BisI has been determined and a gene coding BisI has been identified<sup>[2]</sup>.

BisI and other MD-endonucleases were found in bacteria of different taxonomic groups, but mainly in *Microbacteriaceae* and *Bacillaceae* families. We couldn't detect the MD-endonuclease activity in cell lysates of *Enterobacteriaceae* strains, due to either an absence or extremely low enzymatic activity in bacteria of this group. In this work we have made an

## FULL PAPER

attempt to find MD-endonuclease in *Enterobacteriaceae* strain not by biochemical methods, but based on the *Enterobacteriaceae* genome sequences analysis.

Nowadays, many genomes of enterobacteria are published. This allows us to find a *BisI* similar enzyme by a search of homologues based on the known amino acid sequence of *BisI*. We have used a program PSI-BLAST<sup>[3]</sup> to carry out a homology search by method of successive iterations, at first finding firstly the most similar sequences and then the remote homologues with a low degree of similarity and identity.

Here we describe a search and cloning *ecoBli* gene from bacterial genome of well-known strain *E. coli* BL21 (DE3) with subsequent determination of *EcoBli* endonuclease substrate specificity.

### EXPERIMENTAL

Searching for distant amino acid homologues in *Enterobacteriaceae* was carried out with using the PSI-BLAST program<sup>[3]</sup>.

Multiple alignment of nucleotide and amino acid sequences was carried out with using of Vector NTI Suite 7<sup>[4]</sup>.

Amplification and cloning of the gene encoding MD-endonuclease from *E. coli* BL21 (DE3) strain

Amplification of *BisI* homologue from DNA of *E. coli* BL21 (DE3) laboratory strain (the Sequence Identifier in the GenBank database is ACT43858) was carried out using the following primers:

**Esp-1:** 5'-CCCCATATGAGTGCACGTGAAGCATATC-3'

**Esp-2:** 5'-CGCGGATCCTTAGGGATTACTGACTGAACTCTTC-3'.

These primers contain recognition sites for *FauNDI* and *BamHI* ENases (underlined) to insert the PCR fragment into a plasmid vector.

The resulting PCR fragment was inserted into *pMTL22* plasmid<sup>[5]</sup> on *FauNDI* and *BglIII* sites with *T4* DNA Ligase treatment by standard cloning procedure<sup>[6]</sup>. The resulting plasmid was called *pEcoBli*. After transformation of *E. coli* ER2267 laboratory strain with *pEcoBli* plasmid the resulting recombinant clones were seeded on Petri dishes with LB

agar supplemented with ampicillin (50 µg/mL). The clones were grown overnight at 37°C and then sub-cultured onto separate plates with ampicillin (100 µg/mL) and again incubated overnight for further analysis. Crude lysates of obtained clones were tested for the presence of 5mC-directed DNA endonuclease activity. For further work one of the clones of *E. coli* *pElmI* that gave the best results on the hydrolysis of methylated substrate plasmid was selected.

The analysis of enzymatic activity in *E. coli* recombinant clones with *pElmI* was performed by adding the lysate of bacteria into 20 µL of reaction mixture containing 1 µg of *pFsp4HI3* plasmid DNA<sup>[7]</sup> previously linearized with *DriI* restriction enzyme (*pFsp4HI3/DriI*) in SE-buffer «B» (10 mM Tris-HCl, (pH 7.6 at 25°C), 10 mM MgCl<sub>2</sub>, 1 mM DTT) and incubating the reaction mixture for 1 hour at 37°C. Bacterial lysate was obtained by suspending ~ 0.5-1 µL of biomass grown on a Petri dish in Lysis Buffer (10 mM Tris-HCl, pH 8.5, 0.1 mg/mL Lysozyme, 0.5 M NaCl, 1 mM EDTA, 1 g/L Triton X-100). The products of digestion were determined by electrophoresis in 10 g/L agarose gel.

### The strain producer growth

The recombinant *E. coli* strain producer transformed with *pElmI* was grown in the flasks containing 200 mL of LB broth (10 g/L Tryptone, 5 g/L Yeast extract, 5 g/L NaCl, pH 7.5) with 100 µg/mL ampicillin and cultured in a shaker at 30°C and stirring at 150 rev/min to later log phase, then induced with 0.5 mM IPTG for 4 hours. The cells were precipitated by centrifugation at 5,000 rpm for 30 min at 4°C in the centrifuge J2-21 ("Beckman", USA), rotor JA-10. The cells pellet was stored at -20°C.

### Enzyme purification

*EcoBli* enzyme preparation was purified by one chromatographic step on Heparin-agarose ("Sigma", USA). All purification steps were carried out at 0-4°C. Frozen cell paste was thawed with 50 mL of Buffer A (10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 7 mM 2-mercaptoethanol) containing 1 mM Phenylmethylsulfonyl fluoride (PMSF), 0.1 mg/mL Lysozyme and disrupted by sonication on the ultra-

sound desintegrator Soniprep 150 (MSE, UK) using five 1 min impulses with 1 min intervals. Cell debris was removed by centrifugation at 15,000 rpm for 30 min in the centrifuge J2-21 ("Beckman", USA), rotor JA-20. The supernatant was loaded onto Heparin-agarose column (20 mL) equilibrated with Buffer A with 0.2 M NaCl and washed with two volumes of Buffer A with 0.2 M NaCl. Adsorbed material was eluted with 500 mL of linear gradient (0.2-1 M NaCl) in Buffer A. The fractions containing a peak of EcoBLI activity were collected and dialyzed against 300 mL of Buffer A with 550 g/L Glycerol for 16 hours. The purified enzyme preparation was stored at -20°C. Target enzyme activity assay in chromatographic fractions was performed by addition of 2 µL aliquots from the fractions to 20 µL of reaction mixture containing pFsp4HI3/DriI in SE-buffer «B» for 1 hour at 37°C. The products of digestion were determined by electrophoresis in 10 g/L agarose gel.

### Sequencing

Determination of the nucleotide sequence was

carried out by Sanger-method on the automated sequencer ABI 3130xI Genetic Analyzer ("Applied Biosystems", USA) according to manufacturer's recommendations.

Enzymes preparations, DNA, deoxynucleoside triphosphates and synthetic oligonucleotides, as well as molecular weight markers "1 kb DNA Ladder" used in the work, were from "SibEnzyme Ltd." (Russia). For dilution of enzyme preparations the SE Buffer «B100» (10 mM Tris-HCl, (pH 7.6 at 25°C), 50 mM KCl, 0.1 mM EDTA, 200 µg/mL BSA, 1 mM DTT, 500 g/L Glycerol) was used.

## RESULTS AND DISCUSSION

### BisI MD-endonuclease distant homologues search in databases of enterobacteria protein sequences

A search of the amino acid sequences homologues of MD-endonuclease BisI (Genbank AJW87312)<sup>[2]</sup> among *Enterobacteriaceae* putative proteins was made with PSI-BLAST program. After two rounds of iterations fifty amino acid sequences with unknown functions from enterobacte-

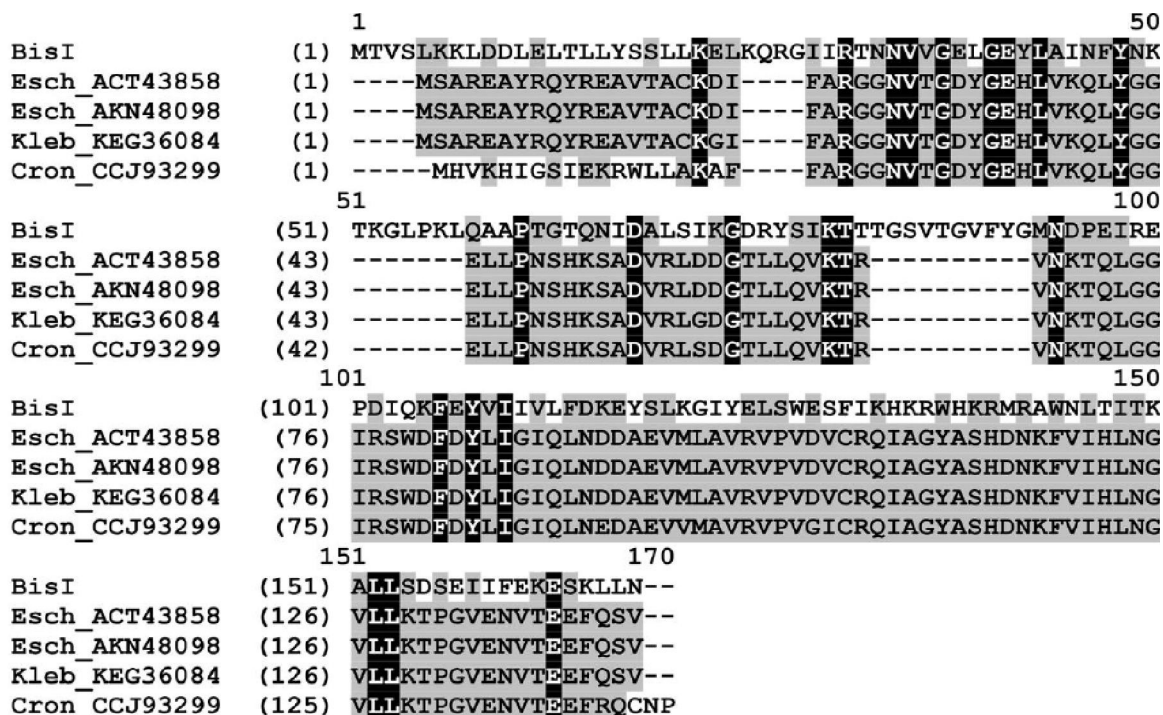


Figure 1 : Multiple amino acid sequence alignment of BisI with high homologous proteins: Legends of amino acid sequences are given according to GenBank bacterial strains numbers: Amino acids that are identical in all protein sequences presented marked in white on a black background: Amino acids that are identical or similar on physical and chemical properties in the majority of homologous sequences marked in black on a gray background

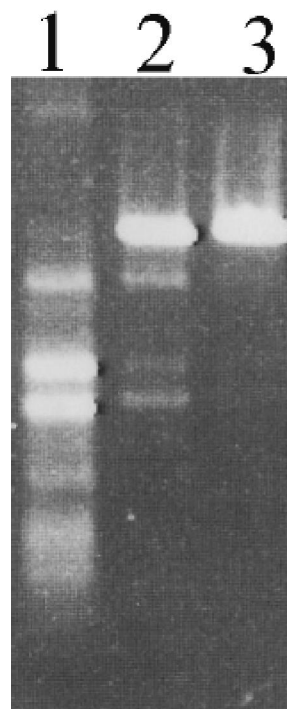
## FULL PAPER

ria strains displaying a homology (~30-50% similarity and ~ 10-12% identity) to the BisI protein sequence were found. Among 50 amino acid homologues we have revealed four very similar hypothetical proteins 143-144 amino acid residues in length belonging to *Escherichia* (ACT43858 and AKN48098), *Cronobacter* (CCJ93299) and *Klebsiella* (KEG36084). The amino acid sequences alignment of BisI MD-endonuclease and a group of four homologous putative proteins is presented in Figure 1.

The nucleotide sequences of corresponding genes have been extracted from Genbank database and compared. A comparison of DNA structure of these genes reveals 93-99% identity of the sequences. Surprisingly, *Escherichia* (ACT43858) turned out to be a well known laboratory strain *Escherichia coli* BL21 (DE3) used in many laboratories for genetic engineering experiments. In our experiments we couldn't isolate any MD-endonuclease from biomass of *Escherichia coli* BL21 (DE3) cells. We have cloned open reading frame (ORF) WP\_001276099 from *Escherichia coli* BL21 (DE3) to determine this ORF function. Structure of Esp-1 and Esp-2 primers complemented to 5' and 3' ends of the ORF WP\_001276099.1 is provided in the methods.

### Amplification and gene cloning of EcoBLI MD-endonuclease

The target ORF (WP\_001276099.1) was successfully amplified from *Escherichia coli* BL21 (DE3) DNA. After elution from agarose gel, the amplified fragment 432 bp in length was treated with FauNDI and BamHI ENases and ligated with pMTL22 vector, previously digested with FauNDI and BglII. After transformation of *E. coli* ER2267 cells with a ligation mixture a recombinant *E. coli* strain was obtained. Determination of MD-endonuclease activity in the cell lysate of this strain was carried out with pFsp4HI3 plasmid DNA carrying methylated sites 5'-G(5mC)NGC-3'/3'-CGN(5mC)G-5'<sup>[7]</sup>, previously linearized with DriI (pFsp4HI3/DriI). DriI ENase cleaves the plasmid at the unique site 5'-GACNNNNGTC-3'. The result of this experiment is presented in Figure 2. Unlike the original *E. coli* BL21 (DE3) the recombinant strain produced MD-endonuclease called



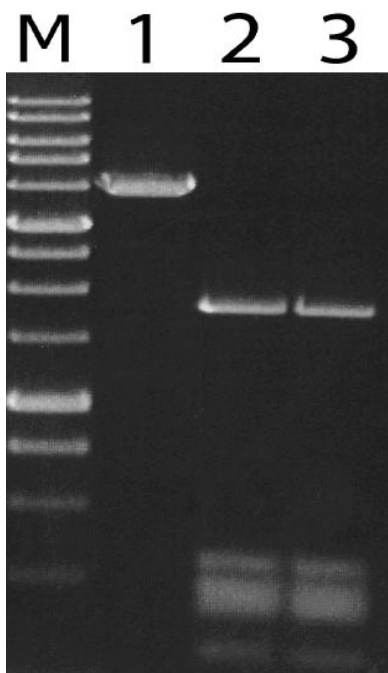
**Figure 2 : Site-specific DNA endonuclease activity assay in the lysate of *E. coli* pEcoBLI recombinant strain on pFsp4HI3/DriI: Electrophoresis in 10 g/L agarose gel: Lanes: 1, pFsp4HI3/DriI + 3  $\mu$ L of *E. coli* pEcoBLI lysate; 2, pFsp4HI3/DriI + 1  $\mu$ L of *E. coli* pEcoBLI lysate; 3, original pFsp4HI3/DriI DNA**

EcoBLI. A recombinant strain was named *E. coli* pEcoBLI. From the above described experiments we can conclude that ORF WP\_001276099.1 is coding MD-endonuclease EcoBLI but this ORF is silencing in original *E. coli* BL21 (DE3) strain.

EcoBLI enzyme preparation was isolated by Heparin-agarose chromatography as described in EXPERIMENTAL. Target enzyme was eluted at 0.9 M NaCl, concentrated by dialysis and stored at -20°C. The yield of enzyme was 11 mL from 6,5 g of strain producer biomass.

### Determination of EcoBLI substrate specificity

To determine a substrate specificity of the new enzyme we have made the treatment of pFsp4HI3/DriI DNA with EcoBLI and control BisI (0.5 units of activity in  $\mu$ L, u.a./ $\mu$ L) enzyme preparations. This plasmid comprises the gene coding Fsp4HI DNA Methyltransferase (MTase) resulting in the first cytosines occur to be methylated in all sequences 5'-GCNGC-3' on both strands<sup>[8]</sup>. pFsp4HI3 also contains three hypermethylated DNA regions due to the



**Figure 3 : Analysis of EcoBli site-specificity on pFsp4HI3/DriI DNA: Electrophoresis in 10 g/L agarose gel: Lanes: M, 1 kb DNA Ladder; 1, pFsp4HI3/DriI DNA; 2, pFsp4HI3/DriI + EcoBli; 3, pFsp4HI3/DriI + BisI**

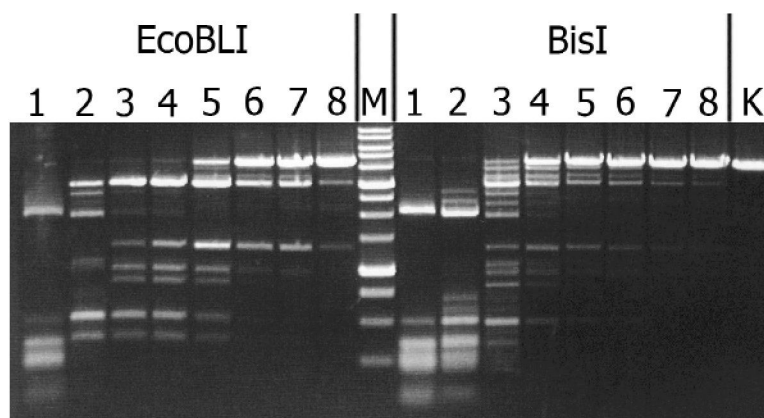
presence of two or three overlapped Fsp4HI MTase recognition sites. The resulting data of this DNA cleavage by 2 μL of EcoBli and BisI are shown in Figure 3.

This figure shows that 2 μL of EcoBli cleaves pFsp4HI3/DriI (lane 2) forming DNA fragments of the same lengths that BisI does (lane 3). Thus EcoBli as well as BisI recognizes and cleaves the sequence 5'-GCNGC-3' when containing at least two 5-methylcytosines and has specific activity.

Figure 4 demonstrates the results of comparative digestion of pFsp4HI3/DriI DNA with diluted EcoBli and BisI enzyme preparations. The substrate DNA was incubated with aliquots of EcoBli and BisI consistently diluted in 2 times (from 2 up to 1/64 μL) with the Buffer "B100".

As can be seen from this figure the results of pFsp4HI3/DriI digestion with EcoBli and BisI are identical when enzyme's quantities are maximal (2 μL of the enzyme with concentration 0.5 u.a./μL), that is EcoBli can digest BisI recognition site containing two 5-methylcytosines: 5'-G(5mC)NGC-3'/3'-CGN(5mC)G-5'. But at the same time the pictures of hydrolysis with diluted enzyme preparations have certain differences. BisI proved to have an equal efficiency of hydrolysis its recognition site containing two, three or four 5-methylcytosines<sup>[1]</sup> whereas EcoBli has the definite preferences.

As was mentioned above pFsp4HI3 plasmid comprises two sequences 5'-GCNGCNGC-3' and one sequence 5'-GCNGCNGCNGC-3' which contain two and three overlapped Fsp4HI MTase recognition sites respectively<sup>[8]</sup>. So as a result of Fsp4HI methylation the sequence 5'-GCNGCNGC-3'/3'-CGNCGNCG-5' comprises two 5'-GCNGC-3'/3'-CGNCG-5' sites with three 5-methylcytosines and the sequence 5'-GCNGCNGCNGC-3'/3'-CGNCGNCGNCG-5' comprises one central 5'-GCNGC-3'/3'-CGNCG-5' site with four 5-methylcytosines and two sites with three 5-methylcytosines. The analysis of the pFsp4HI3 pri-



**Figure 4 : Comparison of pFsp4HI3/DriI DNA cleavage with diluted EcoBli and BisI enzyme preparations: Electrophoresis in 10 g/L agarose gel: Lanes: 1, 2 μL of the enzyme; 2, 1 μL of the enzyme; 3, 1/2 μL of the enzyme; 4, 1/4 μL of the enzyme; 5, 1/8 μL of the enzyme; 6, 1/16 μL of the enzyme; 7, 1/32 μL of the enzyme; 8, 1/64 μL of the enzyme; M, 1 kb DNA Ladder; K, untreated pFsp4HI3/DriI DNA**

## FULL PAPER

mary structure conducted by using the Vector NTI Suite 7 showed that the resulting cleavage pFsp4HI3/DriI at only three hypermethylated sites leads to appearance of ~ 3000 ~ 490 (double fragment) and ~ 340 bp DNA fragments. As it is seen from Figure 4 these fragments were well visualized using 1-1/8  $\mu$ L of EcoBLI (lanes 2-5, left) as opposed to BisI. In the case of BisI these fragments are visualized worse – only with using 1/2-1/4  $\mu$ L of the enzyme (lanes 3-4, right). These fragments disappear completely in the case of DNA substrate treatment with 2  $\mu$ L of EcoBLI or 2 and 1  $\mu$ L of BisI.

These data suggest that unlike BisI new MD-endonuclease EcoBLI cleaves the variants of the sequence 5'-GCNGC-3'/3'-CGNCG-5' with three or four 5-methylcytosines better if compared with the case of methylation only two cytosines.

As expected EcoBLI like BisI didn't cut DNA substrates methylated with following MTases: HpaII (5'-C(5mC)GG-3'), MspI (5'-(5mC)CGG-3') and M1.HgaI+M2.HgaI 5'-GA(5mC)GC-3'/5'-G(5mC)GTC-3') (data not shown).

Therefore a substrate specificity of the new enzyme EcoBLI from *E.coli* BL21 (DE3) is similar to one of MD-endonuclease BisI from *B.subtilis* T30.

### Determination of DNA position hydrolysis for EcoBLI.

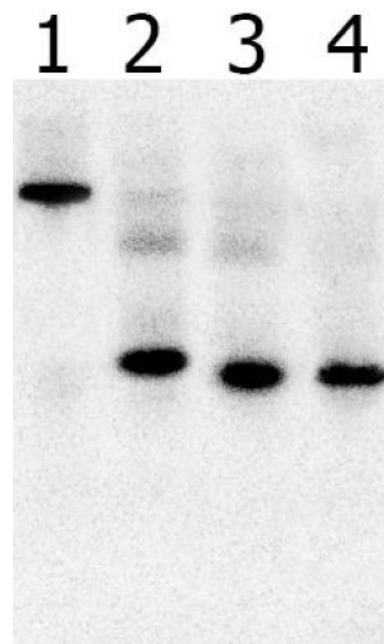
Determination of DNA position hydrolysis for EcoBLI was carried out by comparing of EcoBLI, Pkr and BisI digestion products of oligonucleotide duplex D1/D2, formed from synthesized oligonucleotides D1 and D2 (estimated EcoBLI recognition sequence is underlined):

**D1:** 5'-GAGTTTAG(5mC)GG(5mC)TATCGATCC-3'

**D2:** 5'-GGATCGATAG(5mC)CG(5mC)TAAACTC-3'.

Figure 5 presents the autoradiograph of the labeled duplex \*D1/D2 digested with the enzymes after the electrophoresis in 200 g/L polyacrylamide gel with 7M urea.

Figure 5 shows that DNA fragments formed by PkrI and EcoBLI digestion (lanes 2 and 3, respectively) of D1\*/D2 duplex have different electrophoretic mobilities suggesting that these enzymes have different hydrolysis positions relatively to the recognition site. At the same time the electrophoretic



**Figure 5 : Determination of DNA hydrolysis position for EcoBLI MD-endonuclease on oligonucleotide duplex D1/D2: The symbol “\*” denotes the labeled chain: Electrophoresis in 200 g/L polyacrylamide gel with 7M urea: Lanes: 1, undigested D1\*/D2 duplex; 2, D1\*/D2 + PkrI; 3, D1\*/D2 + EcoBLI; 4, D1\*/D2 + BisI.**

mobilities of DNA fragments formed by BisI and EcoBLI are identical (lanes 3 and 4, respectively). Thus, EcoBLI and BisI have the same hydrolysis positions with respect to the recognition sequence 5'-GCNGC-3'.

## CONCLUSIONS

A search of *bisI* gene homologues among the sequenced genomes of enterobacteria has been made and ORF WP001276099.1 from a well-known strain *E.coli* BL21 (DE3) was amplified and cloned. An obtained recombinant strain *E.coli* pEcoBLI produces MD-endonuclease named EcoBLI with the same substrate specificity as BisI MD-endonuclease. EcoBLI has a recognition sequence 5'-G(5mC)<sup>^</sup>NGC-3'/3'-CGN<sup>^</sup>(5mC)G-5' and cleaves it before the central nucleotide “N” in both strands of DNA forming single-nucleotide 5'-protruding ends. We identified ORF WP001276099.1 from the strain *E.coli* BL21 (DE3) as a silenced gene coding MD-endonuclease.

We believe that EcoBLI as well as BisI MD-endonucleases can be used in epigenetics for detec-

tion and analysis of methylation pattern in mammalian DNA, which usually contains a considerable number of 5-methylcytosine bases. Such type of DNA methylation is known to play a significant role in the regulation of genes expression<sup>[9]</sup>.

### ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education and Science of Russian Federation under the Agreement № 14.576.21.0075 from 06.11.2014, (the unique identifier RFMEFI57614X0075), concluded in the framework of the Federal Target Program “Research and development on priority directions of scientific-technological complex of Russia for 2014-2020 years “.

### DECLARATION

The manuscript is original and is not published or communicated for publication elsewhere either in part or full

Danila Aleksandrovich Gonchar, Ph.D.



### REFERENCES

- [1] E.V.Chmuzh, Yu.G.Kashirina, Yu.E.Tomilova, N.V.Mezentseva, V.S.Dedkov, D.A.Gonchar, M.A.Abdurashitov, S.Kh.Degtyarev; *Biotechnology in Russia*, <http://elibrary.ru/item.asp?id=9441053>, **3**, 35 (2005).
- [2] S.Y.Xu, M.Boitano, T.A.Clark, T.Vincze, A.Fomenkov, S.Kumar, PH-M.Too, D.A.Gonchar, S.Kh.Degtyarev, R.J.Roberts; *Genome Announc*, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4424299>, **3**, e00395 (2015).
- [3] S.F.Altschul, T.L.Madden, A.A.Schaffer, J.Zhang, Z.Zhang, W.Miller, D.J.Lipman; *Nucleic Acids Res.*, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC146917>, **25**, 3389 (1997).
- [4] G.Lu, E.N.Moriyama; *Brief. Bioinform.*, <http://bib.oxfordjournals.org/content/5/4/378.long>, **5**, 378 (2004).
- [5] S.P.Chambers, S.E.Prior, D.A.Barstow, N.P.Minton; *Gene*, <http://www.ncbi.nlm.nih.gov/pubmed/2851488>, **68**, 139 (1988).
- [6] J.Sambrook, E.F.Fritsch, T.Manias; *Molecular Cloning, A laboratory manual 2<sup>nd</sup> Edition*, Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press., <http://trove.nla.gov.au/version/180145381>, (1989).
- [7] V.A.Chernukhin, E.V.Chmuzh, J.E.Tomilova, T.N.Nayakshina, D.A.Gonchar, V.S.Dedkov, S.Kh.Degtyarev; *Bulletin of biotechnology and physical and chemical biology named after Ovchinnikov (Russian)*, **3(2)**, 13 (2007). <http://elibrary.ru/item.asp?id=12837363>
- [8] E.V.Chmuzh, Yu.G.Kashirina, Yu.E.Tomilova, V.A.Chernukhin, S.S.Okhapkina, D.A.Gonchar, V.S.Dedkov, M.A.Abdurashitov, S.Kh.Degtyarev; *Molecular Biology*, <http://dx.doi.org/10.1134/S0026893307010062>, **41(1)**, 37 (2007).
- [9] J.F.Costello, C.Plass; *J. Med. Genet.*, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1734882>, **38**, 285 (2001).