

## Multiplicity of Site-specific DNA Methyltransferases of the BstF5I Restriction-Modification System

Category: [The enzymes genes cloning and study](#)

Hits: 6129

L. N. Golikova, N. A. Netesova, V. V. Gutorov, P. A. Belavin, M. A. Abdurashitov, D. A. Gonchar and S. Kh. Degtyarev

Translated from MOLECULAR BIOLOGY (Russia) Vol. 34 No. 3, 443-447, 2000

*A fragment located downstream of the genes for DNA methyltransferases of Bacillus stearothermophilus F5 (M.BstF5I-1 and M.BstF5I-2) was sequenced. The fragment contains a gene for another methylase. M.BstF5I-3. structurally and functionally similar to the N-terminal domain of M.FokI. Thus, in contrast to other restriction-modification systems, the BstF5I system includes three methylases, two being homologous to the individual M.FokI domains.*

### INTRODUCTION

Restriction-modification systems are widely presented in microorganisms. In RM systems recognizing palindromic DNA sequences there is one methylase (M.) which modifies both DNA strands in the recognition site and thereby protects bacterial DNA from hydrolysis with cognate restriction endonuclease (R.). In nonpalindromic restriction-modification systems (type IIS) there are two methylases. each modifying only one DNA strand in the recognition sequence. Homologous RM systems FokI [1] and SstI [2] are exceptions: both of them has only one methylase. which consists of two D12 domains each methylating one strand in its recognition site [3]. The domains are of 300-350 amino acid residues, thus being comparable in length with other DNA methylases [4]. Hence the FokI, methylase gene may be assumed to result from fusion of two methylase genes [5].

We previously found Bacillus stearothermophilus strain F5 producing a restriction endonuclease BstF5I which, like endonucleases FokI and SstI, recognized 5'-GGATG-3'. However, the cleavage site was 2/0 for R.BstF5I, 9/13 for R.FokI and 10/14 for R.SstI [6].

We cloned the gene for M.BstF5I-I [7] and showed that this enzyme is a D<sub>21</sub> methylase modifying an adenine on the upper strand in the recognition site. Primary structure analysis showed that M.BSTF5I-1 is nonhomologous to M.FokI and M. SstI.

We also cloned and sequenced the bstF5IM-2 gene for the second DNA methyltransferase and showed that the gene is located downstream of the gene for M.BstF5I-I on the bacterial chromosome [8]. The protein encoded has conserved regions specific for adenine methylases of subclass D<sub>12</sub>. M.BstF5I-2 is homologous to the C-terminal domain of M.FokI and recognizes the same sites. Thus, the BstF5I

restriction-modification system was shown to include M.BstE5I-1 modifying the DNA upper strand and M.BstF5I-2 modifying the other strand in the recognition site [8].

Here we demonstrate that the BstF5I system includes the third methylase (M.BstF5I-3), which also modifies an adenine in the 5'-GGATG-3' sites, and analyze its primary structure.

## **EXPERIMENTAL**

We used *Escherichia coli* RRI (New England Biolabs, United States), restriction endonucleases, T4 DNA ligase, calf intestinal alkaline phosphatase (SibEnzyme, Novosibirsk), and S-adenosyl [3H]methionine (Amersham). Gene engineering manipulations followed the standard protocols [9]. *B. stearothermophilus* F5 genomic DNA was isolated from low-melting agarose using block insertion [10]. DNA sequencing was done by the Maxam-Gilbert method [11]. Amino acid sequences were aligned with the ALIGN program of the FASTA 2.0 package [12]: PCR primers were selected with the OLIGO program [13].



## RESULTS AND DISCUSSION

### Gene Cloning and Expression, Protein Isolation and Analysis

To obtain a genome library, *B. stearothermophilus* F5 DNA was partly digested with BstX2I (isoschizomer of XhoII). The fragments were ligated into the BamHI site of pMTL22 [14]. The resulting DNA was used to transform *E. coli* RR1. The plasmid pool was isolated and treated with BstF5I. Transformation of *E. coli* RR1 with the hydrolysis products yielded 120 clones. Of the 48 clones analyzed, eight carried plasmids with an insert of 2.2 kb resistant to R.BstF5I. The plasmid was designated pF5-32.

Sequencing showed that the BstX2I fragment cloned had a 822-nt overlap with the AluI fragment cloned earlier and containing the bstf5IM-2 gene [8] (Fig. 1a). The BstX2I fragment contained an open reading frame (ORF) of 1038 nt and the canonical Shine-Dalgarno signal sequence (AGGAG) located 10 nt upstream of its start codon.

Protection of pF5-32 from hydrolysis with R.BstF5I indicates that the ORF codes for a site-specific DNA methyltransferase. The ORF was designated bstF5IM-3.

The bstF5IM-3 gene was amplified from *B. stearothermophilus* F5 genomic DNA with primers 5'-CCCCATATGAGATATATCGGCAGCAA and 5'-CCCGGATCCTTCTTCTTCAATTAGAAAATTT and cloned into the BamHI-NdeI sites of expression vector pJW [15]. The vector was used to transform *E. coli*, protein synthesis was induced by heating, and the cells were lysed. As expected, disk electrophoresis revealed a recombinant 41.1 K protein. The protein was purified to homogeneity by chromatography on heparin-Sepharose, hydroxyapatite, and phosphocellulose P-II.

Analysis of the deduced amino acid sequence showed that M.BstF5I-3 had all conserved motifs characteristic of D12 methylases. Sequence alignment revealed 57.6% identity and 81.1% similarity with the N domain of M.FokI (Fig. 1b), which modifies an adenine in the recognition site 5'-GGATG-3' on the upper strand [3J]. Thus, M.BstF5I-3 is probably a functional analog of M.BstF5I-I [7]. To check this assumption, we studied the methylation of the 35-bp duplex

| 19 nt                                                       | 16 nt |
|-------------------------------------------------------------|-------|
| 5'-CAAGGATGCATATGACCAG↓GTCCAGCTAGCGGGTA-3'                  |       |
| 3'-GTT <u>CTAC</u> GTATACTGGT <b>CCAG</b> ↑GTCGATCGCCCAT-5' |       |
| 22 nt                                                       | 13 nt |

containing the recognition sites for R.BstF5I and R.Bme18I (G<sup>^</sup>GWCC, boldface; hydrolysis sites are shown with arrows; the size of the hydrolysis products is indicated) and for M.BstF5I-3 (underlined) [7]. The reaction mixture contained 0.5 μM M.BstF5I-3, 2.0 μM duplex, and 3.3 μM S-adenosyl methionine with 3H-labeled methyl group. Methylation was carried out at 50°C for 2 h. The product was cleaved with Bme18I (isoschizomer of Avail), and the resulting fragments were resolved by PAGE in 20% denaturing gel.

As expected, the upper strand with the 5'-GGATG-3' site was methylated preferentially (table), which

confirmed the functional similarity of M.BstF5I-3 to M.FokIN.

Nucleotide sequence analysis revealed a promoter upstream of *bstF5IM-I* [7] and Shine-Dalgarno sequences upstream of each gene. The fragment sequenced did not contain transcription termination sites. Thus, the genes probably form an operon and are transcribed to produce a single mRNA. The Shine-Dalgarno sequences provide for efficient in vivo translation of each gene.

This is consistent with the data that each individually cloned gene sufficiently to protect the corresponding plasmid from R.BstF5I, though only one DNA strand in the recognition site was methylated. Thus, the BstF5I restriction-modification system includes three methylases recognizing 5'-GGATG-3' and modifying an adenine in the upper or lower strand.

## Arrangement of the BstF5I Methylase Genes

Although M.BstF5I-I and M.BstF5I-3 both modify . adenine bases in the DNA upper strand and recognize 5'-GGATG-3'. they are nonhomologous and belong to different subclasses (D21 and D12, respectively). The other methylase, M.BstF5I-2, is highly homologous to M.FokIC and modifies an adenine in the other strand [8]. The arrangement of the three methylase genes in the *B. stearothermophilus* F5 genome is shown in Fig. 1a.

To rule out chromosomal DNA recombination in constructing genome libraries, a genomic DNA fragment was amplified with primers 5'-CCCAAAAGC-CTGTAAAGCT and 5'-ATCGCCTTCAAGTTC-CCTA complementary to the central region of *bstF5IM-I* and *bstF5IM-3*, respectively. The resulting fragment was of 1950 nt (expected size 1941 nt). Restriction enzyme analysis of the fragment confirmed the gene arrangement shown in Fig. 1a.

| Fragment size, nt | Incorporated [ <sup>3</sup> H], cpm |
|-------------------|-------------------------------------|
| 11                | 210                                 |
| 19                | 2590                                |
| 16                | 91                                  |
| 13                | 63                                  |

Identification of the DNA strand methylated by M.BstF5I-3

## Analysis of the Fourth ORF

Interestingly, an incomplete ORF4 was found immediately downstream of *bstF5IM-3* in the BstX2I fragment cloned (Fig. 1a). Region 145-212 of the deduced amino acid sequence contained methylase specific conserved motifs X, I, and II and was the most similar to the C-terminal fragment of D21 methylases (Fig. 2). Region 1-144 had no appreciable homology with any known protein.

Possibly, ORF4 is a chimeric gene resulting from fusion of an unknown gene with the 3' part of a gene for D21 methylase. Establishing the role of the protein encoded requires cloning and sequencing of the complete ORF4.

## Possible Causes of Methylase Multiplicity

The BstF5I restriction-modification system differs from other known systems, which include no more than two methylase genes. While M.BstF5I-2 and M.BstF5I-3 are homologous to the M.FokI domains, the order of the corresponding genes on the chromosome is opposite to that of the domains. Possibly, a DNA fragment coding for these enzymes had been transferred into the *B. stearothermophilus* F5 genome from another microorganism. Such a recombination event is well consistent with the presence of a spacer with a long tandem repeat between bstF5IM-I and bstF5IM 2[8] and the chimeric ORF4 downstream of bstF5MI-3. We assume that an ancestral operon

|             |     | CM X                                                                                            | CM I   |     |
|-------------|-----|-------------------------------------------------------------------------------------------------|--------|-----|
| M. DpnA     | 188 | SLTKKVEKWAGKHPT <b>Q</b> KPEYLLERIILASTKEGDYILD <b>P</b> FVGSGTTGVVA                            |        |     |
| M. MboIC    | 178 | PAVGSWEKTQGHPT <b>Q</b> KPLGLLSRIILSSTQKDDLIL <b>D</b> PFSGSGTTGIAG                             |        |     |
| M. HinfI    | 181 | KERIKDEEGKKAHST <b>Q</b> KPESLLYKVILSSSKPNDVVL <b>D</b> PFFGTGTTGAVA                            |        |     |
| M. BstF5I-1 | 178 | YPSESGGGFKRYHPT <b>Q</b> KPVKLI EYLIRIHSNENDIIL <b>D</b> PFMGSGTTGVAS                           |        |     |
| M. HindIII  | 222 | LKVRITSEDKGYHVA <b>Q</b> KPLN <b>L</b> MKLLT <b>D</b> LV <b>T</b> KEEQIVL <b>D</b> PFAGSGTTLLAA |        |     |
| M. HpaI     | 161 | SHVHYCNKNRKPHPT <b>Q</b> KPEALYERMILASSNEGDIVL <b>D</b> PFVGS <b>G</b> T <b>L</b> NFVC          |        |     |
| M. MboII    | 181 | KEKVN <b>G</b> KTVKLTHITPKPRDLIERIIRASSNPNDLV <b>L</b> DCFMGS <b>G</b> T <b>T</b> AI <b>V</b> A |        |     |
| ORF4        | 133 | GKEGYAHHIRKVHPS <b>P</b> KPPQLTKEIISFFTKENE <b>V</b> LDYFMGVGGTLLGA                             |        |     |
|             |     | * : * * : * : : * * * * *                                                                       |        |     |
|             |     | CM II                                                                                           | CM III |     |
| M. DpnA     |     | KRLGRRFIGIDAEKEYLKIARKRLEAENETN*                                                                |        | 268 |
| M. MboIC    |     | VLLDRNYIGIEQELEFLELSKRRYHEITPVLKNEFKQKIRKQISAI*                                                 |        | 273 |
| M. HinfI    |     | KALGRNYIGIEREQKYIDVAEKRLREIKPNPNDIELLSLEIKPPKV...                                               |        | 276 |
| M. BstF5I-1 |     | LNLNRNFIGFEINEEYVQIANERIKNIQLSLI*                                                               |        | 259 |
| M. HindIII  |     | KELNRHFIGYEKNNGIYNI <sup>AV</sup> NRLGIEKN <sup>NC</sup> CFYNKEKK*                              |        | 309 |
| M. HpaI     |     | KHLNRSGIGIDINKEYIEMAKERLDSEFN <b>G</b> FDSIDERMKRC <b>P</b> NDLS...                             |        | 256 |
| M. MboII    |     | KKLGRNFIGCDMNAEYVNQANFVLN <b>Q</b> LEIN*                                                        |        | 260 |
| ORF4        |     | SLCN <b>R</b> RAIGIDLN <b>G</b> FYIEKYKEAAKALNLD...                                             |        | 212 |
|             |     | . * ** : : ...                                                                                  |        |     |

**Fig. 2.** Amino acid sequence alignment of the C-terminal regions of D21 methylases and region 133-212 of the deduced sequence encoded by ORF4. Conserved motifs (CM) are indicated with upper lines; highly conserved residues are in bold. Bottom row, (\*) identity, (:) functional similarity, (·) functional compatibility of amino acid residues

included two genes for D21 methylases, one being bstF5IM-I. Of the other gene, the 5' portion was deleted and the 3' portion fused with a fragment of an unknown gene to produce ORF4.

It is unclear why the operon still preserves the genes for two enzymes (M.BstF5I-I and M.BstF5I-3) methylating the same strand in the recognition site. Possibly, the recombination event is evolutionary too recent for an excessive gene to have been eliminated. Another reason is that the activity of

R.BstF5I in *B. stearothermophilus* F5 cells is unusually high, two orders of magnitude higher than that of R.FokI in *Flavibacterium okeanokoites* (> 100.000 and <1000 units per gram biomass, respectively). Hence both genes must be preserved if either methylase alone does not suffice to protect the corresponding DNA strand. In addition, M.BstF5I-I and M.BstF5I-3 may be involved in cell processes other than DNA methylation.

Thus, the BstF5I restriction-modification system unique among such systems known to date.

## ACKNOWLEDGMENT

This work was supported by the Russian Foundation for Basic Research (project no. 98-04-49514).

## REFERENCES

1. Landry, D.// *Gene*, 1989, vol. 77, pp. 1-10.
2. Kita, K., Suisha, M., Kotani, H., et al.// *Nucleic Acids Res.*, 1992, vol. 20, pp. 4167-4172.
3. Susisaki, H. // *Biol. Chem.*, 1989. vol. 264, pp. 5757-5761.
4. Leisman, O.// *Eur. J. Biochem.*, 1997, vol. 251, pp. 899-906.
5. Looney, M.C., Moran, L.S., Jack, W.E., et al.// *Gene*, 1989, vol. 80, pp. 193-208.
6. Abdurashitov, M.A., Kileva, E.V., Shinkarenko, N.M. et al.// *Gene*, 1996, vol. 172, pp. 49-51.
7. Degtyarev, S.Kh., Netesova, N.A., Abdurashitov, M.A., and Shevchenko, A.V.// *Gene*, 1997, vol. 187, pp. 217-219.
8. Abdurashitov, M.A., Netesova, N.A., Golikova, L.N. et al.// *Mol. Biol.*, 2000, vol. 33, pp. 1-8.
9. Maniatis, T., Fritsch, E.F., and Sambrook, J.// *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor. N.Y.: Cold Spring Harbor Lab. Press, 1982.
10. *Genom Analysis: A Practical Approach*, Davies, K.E.. Ed., Oxford, 1988.
11. Maxam, J.M. and Gilbert, W.C.// *Pmc. Nad. Acad. Sci. USA*, 1987, vol. 74, pp. 7191-7203.
12. Pearson, W.R. and Lipman, D.J.// *Proc. Natl. Acad. Sci. USA*. 1988, vol. 85, pp. 2444-2448.
13. Rychlik, W. and Rhoads, R.E.// *Nucleic Acids Res.*, 1989. vol. 17, p. 8543.
14. Chamberts, S.P., Prior, S.E., Barstow, D.A., and Minton, N.P.// *Gene*, 1988, vol. 68, pp. 139-149.
15. Kossykh, V.G., Schlanman, S.L., and Hattman, S.// *J. Biol. Chem.*, 1995, vol. 270, pp. 14389-14393.