Nucleic Acids Research, 1992, Vol. 20, No. 24 6517–6523

*Bsu*BI—an isospecific restriction and modification system of *Pst*I: characterization of the *Bsu*BI genes and enzymes

Guo-Liang Xu, Wolfgang Kapfer, Jörn Walter+ and Thomas A.Trautner*
Max-Planck-Institut für Molekulare Genetik, Ihnestr. 73, D-1000 Berlin 33, Germany

Received September 29, 1992; Accepted November 23, 1992 GenBank accession no. L01541

ABSTRACT

The enzymes of the *Bacillus subtilis* *Bsu*BI restriction/modification (R/M) system recognize the target sequence 5'CTGCAG. The genes of the *Bsu*BI R/M system have been cloned and sequenced and their products have been characterized following overexpression and purification. The gene of the *Bsu*BI DNA methyltransferase (M.*Bsu*BI) consists of 1503 bp, encoding a protein of 501 amino acids with a calculated M_r of 57.2 kD. The gene of the restriction endonuclease (R.*Bsu*BI), comprising 948 bp, codes for a protein of 316 amino acids with a predicted M_r of 36.2 kD. M.*Bsu*BI modifies the adenine (A) residue of the *Bsu*BI target site, thus representing the first A-N6-DNA methyltransferase identified in *B. subtilis*. Like R.*Pst*I, R.*Bsu*BI cleaves between the A residue and the 3' terminal G of the target site. Both enzymes of the *Bsu*BI R/M system are, therefore, functionally identical with those of the *Pst*I R/M system, encoded by the Gram negative species *Providencia stuartii*. This functional equivalence coincides with a pronounced similarity of the *Bsu*BI/*Pst*I DNA methyltransferases (41% amino acid identity) and restriction endonucleases (46% amino acid identity). Since the genes are also very similar (58% nucleotide identity), the *Bsu*BI and *Pst*I R/M systems apparently have a common evolutionary origin. In spite of the sequence conservation the gene organization is strikingly different in the two R/M systems. While the genes of the *Pst*I R/M system are separated and transcribed divergently, the genes of the *Bsu*BI R/M system are transcribed in the same direction, with the 3' end of the M gene overlapping the 5' end of the R gene by 17 bp.

INTRODUCTION

Restriction and modification (R/M) systems occur in a wide range of procaryotes. They represent interesting model systems to study sequence specific protein-DNA interaction as well as molecular evolution (1,2). To date the genes of more than hundred R/M systems have been cloned (3).

In this paper we report on the cloning and sequencing of the genes of the *Bsu*BI R/M system, one of six R/M systems found in the Gram positive bacterium *B.subtilis* (1,2,4). We also carried out a biochemical characterization of the gene products, in particular with respect to the cleavage and modification specificities of the restriction endonuclease (ENase) and DNA-methyltransferase (MTase), respectively. Our studies were motivated by the fact that the *Bsu*BI system has the same sequence specificity as the well characterized *Pst*I R/M system (5,6), found in the Gram negative bacterium *P.stuartii*. We were interested to know how the two systems, encoded by phylogenetically distant bacterial species, would compare at the gene and protein level. Beyond this general aspect concerning the evolution of R/M systems, the comparison between *Bsu*BI and *Pst*I was of special interest. Within the common target 5'CTGCAG, M.*Pst*I modification is mediated through methylation of the A residue to produce N6-methyladenine (N6-mA) (6). The position of M.*Bsu*BI mediated methylation had not yet been determined. It, therefore, remained unresolved whether M.*Bsu*BI would be a C5-MTase, as all other phage and host encoded MTases identified so far in *B.subtilis* (1), or, like M.*Pst*I, an A-N6-MTase. An answer to this question seemed of particular interest, as R.*Pst*I can be blocked by both A-N6- and C5-modification of the 5'CTGCAG target sequence (7). Even if the restriction enzymes of the *Pst*I and *Bsu*BI systems would be related, different types of MTases could, hence, be operative in the two systems.

MATERIALS AND METHODS

Bacterial strains, plasmids, phages, media

B.subtilis ISB8 (*hsr*BI$^+$, *hsm*BI$^+$, *hsr*MI$^-$, *hsm*MI$^+$) was provided by Shibata (4). *E.coli* NM675 (e14$^-$, (*mrr-hsd*RMS-*mcr*BC)) (8) and *B.subtilis* MI112 (*hsr*MI$^-$, *hsm*MI$^-$, *hsr*BI$^-$, *hsm*BI$^-$, *recE4*) (9) were used for maintenance of plasmids encoding the *Bsu*BI R/M genes. Plasmid pOU71 (10) was used for cloning of the M.*Bsu*BI and R.*Bsu*BI genes. pBR328 (11), pGB2 (12) and the *E.coli*/*B.subtilis* shuttle plasmid pHP13 (13) were used for subcloning. Plasmid pMS119 E/H (13) served as an expression vector for the M.*Bsu*BI and R.*Bsu*BI enzymes. Phages ϕ105c (15) and λ_{vir} were used for *in vivo* restriction analyses.

* To whom correspondence should be addressed

+ Present address: Department of Molecular Embryology, IAPGR, CRS, Babraham, Cambridge CB2 4AT, UK

B.subtilis strains were grown in TY-medium (16) with addition of 5 μg chloramphenicol/ml, if required. *E.coli* strains were grown in L-medium (17), supplemented with 50 to 100 μg ampicillin/ml, 70 μg spectinomycin/ml or 50 μg erythromycin/ml, if required.

Enzymes and chemicals

Restriction and modification enzymes, snake venom phosphodiesterase, calf intestine phosphatase (CIP), T4 DNA ligase and DNase I were purchased from Boehringer, Mannheim (FRG) and New England Biolabs (USA). *N*6-methyl-2'-deoxyadenosine (N6-mdA) and 5-methyl-2'-deoxycytidine (5mdC) were obtained from Sigma (FRG). [α-^{35}S]dATP, [γ-^{32}P]ATP and [^{3}H-methyl]-S-adenosylmethionine (SAM) as methyl group donor were from Amersham (UK).

Preparation of a genomic DNA library

Chromosomal DNA of *B.subtilis* ISB8 was partially digested with *Sau*3AI. DNA fragments generated were inserted into the *Bam*HI site of pOU71. The ligation mixture was transformed in competent *E.coli* NM675 cells. About 10^4 ampicillin-resistant (Ap^R) colonies were scraped from the plates and resuspended in 50 ml L-medium. 150 ml L-medium were inoculated with the resuspended bacteria, and after overnight growth, plasmid DNA was isolated according to standard procedures (18). Selection for recombinant plasmids mediating M.*Bsu*BI expression was performed by incubation of about 5S0256T mg plasmid DNA with a high excess of R.*Pst*I, and subsequent retransformation of the NM675 with the R.*Pst*I treated DNA.

Methylation activity assay

The methylation assay *in vitro* was performed by incubation of *M.lysodeikticus* DNA with 5 μl crude extracts of NM675 cells harbouring the M.*Bsu*BI gene on different plasmids (Fig. 2 and 5). Reactions were carried out for 1 h at 37°C in a 50 μl volume with 50 mM Tris−HCl (pH 7.5), 10 mM EDTA, 5 mM 2-mercaptoethanol and [^{3}H-methyl] SAM (1 μCi). The methylation activity was measured according to Günthert *et al.* (19).

Restriction activity assay

R.*Bsu*BI specific restriction activity of *B. subtilis* ISB8 and MI112 [pBH64] was measured according to Trautner et al. (20) by determining the plating efficiency of phage φ105c. The R.*Bsu*BI restriction activity of *E.coli* cells carrying the R.*Bsu*BI gene on different pMS119 E/H or pBR328 derivatives (Fig. 2) was monitored by determining the plating efficiency of phage λ_{vir} as described (21)

Restriction activity *in vitro* was assayed by incubation of 2 μg λ-DNA with 1 μl crude extract of NM675 [pMS1.9, pGB3.6] or 2 μl of a selected purification fraction in a final volume of 20 μl containing 50 mM Tris−HCl (pH 7.5), 0.2 mM EDTA, 5mM $MgCl_2$, 5mM 2-mercaptoethanol for 1 h at 37°C. The cleavage products were analyzed on 1% agarose gels. One unit of restriction activity was defined as the amount of enzyme required for the complete digestion of 1 μg λ-DNA in one hour.

Primer extension experiments

For identification of the transcriptional starting points of the *Bsu*BI MTase and ENase genes, total RNA of *B.subtilis* and *E.coli* was prepared as described (22,18). Synthetic oligonucleotides complementary to the nonsense DNA strand of the *Bsu*BI MTase gene at nt 240−267 and the ENase gene at nt 1592−1620/1762−1789 were 5' end labelled with [γ-32P] ATP and hybridized with the RNA probes under appropriate conditions. The primer extension reactions were carried out for 1 h at 52°C using 4 units AMV reverse transcriptase (Promega, USA). The products were analyzed on 8% sequencing gels.

Purification of the R.*Bsu*BI and M.*Bsu*BI enzymes

For overproduction and partial purification of M.*Bsu*BI, a 50 ml culture of *E.coli* NM675 [pMS3.1] induced for 2.5 h with 1mM IPTG was used. The cells were pelleted and disrupted with a French press. The crude extract was treated with an ice cold saturated ammonium sulfate solution. The MTase was partially purified by two chromatography steps using a DEAE Sepharose (Pharmacia) and DNA Cellulose (Pharmacia) columns according to Kapfer et al. (23).

For the overproduction of R.*Bsu*BI, a 200 ml culture of *E.coli* NM675 [pMS1.9/pBG3.6] grown to OD_{600} = 0.8 was induced with 1mM IPTG for 4 h. The whole purification procedure was further carried out as described (23).

Determination of the M.*Bsu*BI methylation position

Two synthetic oligonucleotides 5'CCTCTACTGCAGG and 5'AGGCCTGCAGTAG (the M.*Bsu*BI recognition sequence is underlined) were phosphorylated with T4 polynucleotide kinase. The complementary oligonucleotides were annealed under appropriate conditions resulting in DNA duplexes with cohesive ends. The DNA was ligated with T4 DNA ligase to generate DNA concatemers with multiple 5'CTGCAG sites. The DNA was methylated *in vitro* with a crude extract of *E.coli* NM675 [pMS3.1] containing the overexpressed M.*Bsu*BI enzyme using [^{3}H-methyl] SAM under standard methylation conditions (see above). Control reactions were performed with crude extracts of *E.coli* HB101 expressing the 5'GGCC C5-MTase M.*Bsp*RI from pKK223-3 (24) or the commercially supplied A-N6-MTase M.*Pst*I (5 units). The methylated DNA substrates were hydrolyzed gradually with DNase I, phosphodiesterase, CIP to nucleosides (25). The degradation products were separated by paper chromatography (26). The chromatograms were analyzed either under UV light or after treatment with enhancer spray (NEN, USA) by autoradiography on Kodak X-OMAT films exposed for 3 days.

Determination of the cleavage site of R.*Bsu*BI ENase

The restriction cleavage site of R.*Bsu*BI enzyme was determined according to Brown et al. (27). Standard dideoxy sequencing reactions (28) were carried out using 1 μg M13mp18 single strand DNA as template and the M13 17-mer sequencing primer (−40) obtained from USB (USA). An aliquot (1 μl) of the sequencing reaction 'A' was stopped by incubation of the reaction mixture at 70°C for 10 min to inactivate the sequenase. The reaction mixture was then treated with an excess (about 10 units) of either the purified R.*Bsu*BI or the R.*Pst*I ENase in a 20 μl volume under standard restriction buffer conditions (see above) at 37°C for 1 h. Prior to further analysis the cleavage products were diluted 10 times in sequencing loading buffer and 2.5 μl of each mixture were loaded onto an 8% denaturing polyacrylamide gel.

Other methods

For DNA sequencing, restriction fragments of interest were cloned into M13mp18 and M13mp19. Nucleotide sequences of

both DNA strands were determined by the dideoxy chain termination method (28).

Protein concentrations were measured according to the method of Bradford (29) using a BioRad Kit (FRG) with BSA as a protein standard. The N-terminal amino acid determinations of the *Bsu*BI enzymes were performed as described before (23), with the specific protein bands isolated by SDS-polyacrylamide gels.

Sequence comparisons were performed on a VAX using the UWGCG7 software (30).

RESULTS

Cloning and expression of the *Bsu*BI MTase and ENase genes

The isolation of the composite genes of the *Bsu*BI system made use of selective recombinant plasmids containing fragments of *B.subtilis* ISB8 DNA, providing resistance to R.*Pst*I mediated by M.*Bsu*BI (31). From the R.*Pst*I digested genomic library of *B.subtilis* ISB8 in the *E.coli* plasmid pOU71 we isolated three identical plasmids, including pBJ64, which was used for further studies. These plasmids contained a 4.8 kb insert of ISB8 DNA and were resistant to R.*Pst*I, indicating that they carried and expressed the M.*Bsu*BI gene. Further reduction of the 4.8 kb insert and parallel measurement of the M.*Bsu*BI activity, placed the M.*Bsu*BI gene approximately in the center of the 4.8 kb ISB8 DNA fragment (Fig. 1A). No R.*Bsu*BI restriction activity was found in *E.coli* NM675 cells with pBJ64 and its noninducible derivatives. This was due to the absence of expression of the ENase gene in the heterologous *E.coli* host. When the 3.6 kb *Hin*dIII DNA fragment from pBJ64 was cloned into the shuttle vector pHP13 yielding pHH64 (Fig. 1A), *B.subtilis* MI112 cells

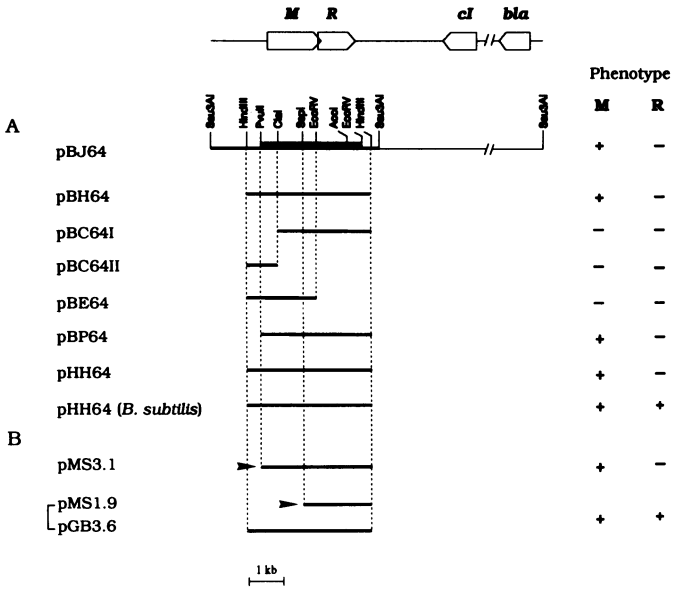

Figure 1. Schematic map of plasmid pBJ64 and its derivatives. Heavy horizontal lines represent chromosomal DNA fragments of ISB8 cloned in various plasmids. Arrows indicate the position and orientation of genes of pOU71 and of M.*Bsu*BI (*M*) and R.*Bsu*BI (*R*). The restriction and methylation activities of *E.coli*/*B. subtilis* cells harbouring the plasmids drawn are shown in the right panel (- denotes deficient; + denotes proficient). The sequenced region between the *Pvu*II site and one of the *Eco*RV sites is indicated by cross hatching. **A.** Plasmids constructed to identify the *Bsu*BI genes. **B.** Plasmids constructed to overexpress the *Bsu*BI genes. Filled arrowheads represent the *tac* promoter from pMS119 E/H and the direction of transcription.

with this plasmid restricted the growth of ϕ105c thousandfold, whereas the same plasmid in *E.coli* did not affect the growth of λ_{vir} phages.

Determination of the nucleotide sequences of the *Bsu*BI R/M genes and the deduced amino acid sequences

The nucleotide sequence of both strands of the 2.9 kb *Pvu*II-*Eco*RV DNA fragment obtained from pBJ64 (Fig. 1A) was determined. Two identically oriented large ORFs, overlapping by 17 bp, could be detected (Fig. 2). On the basis of the phenotypes of cells carrying different portions of the cloned *B.subtilis* ISB8 DNA (Fig. 1A) and the presence of the *Cla*I and *Eco*RV sites in this frame, the first ORF could be assigned to M.*Bsu*BI. The second one encoded R.*Bsu*BI. This assignment was confirmed by inserting into this frame two base pairs in the unique *Acc*I site, which abolished the capacity to specify restriction activity.

The ORF encoding M.*Bsu*BI consists of 1503 bp with the translational start codon, ATG, at nucleotide (nt) 228 and the stop codon, TAG, at nt 1731 (Fig. 2). The ORF encodes a protein of 501 aa with a calculated M_r of 57.2 kD. The deduced amino acid sequence contains the sequence motif GxGxG, a variant of the proposed SAM binding site characteristic for all types of DNA MTases (32, 33) and the motif NPPY which suggests that the enzyme is an A-N6-MTase (34, 35). Seven nt upstream of the start codon a putative S/D sequence, 5'TAAGGAG, could be identified. The assignment of the translational start of the M.*Bsu*BI gene was verified by microsequencing of the N-terminus of the M.*Bsu*BI protein recovered from the DNA Cellulose fraction (see Fig. 4A, track 4). Except for the initial methionine, which is missing, the amino acid sequence obtained (T-Q-I-L-E-X-V-D-K-X-R-L) corresponds to that predicted for the N-terminus of M.*Bsu*BI. For the *Bsu*BI MTase gene a putative -10 box at nt 185–190 was found (Fig. 2), which resembles the consensus sequence of the -10 region of *E.coli* and the *B.subtilis* vegetative promoters (36). No canonical -35 region could be identified.

There are two potential translational start codons for the ORF representing the R.*Bsu*BI gene. One is the ATG codon at nt 1728, the other is the unusual start codon GTG at nt 1714. Only the latter codon is preceded at the unconventional distance of 12 bp to the translational start site by a prototype S/D sequence 5'GG-AGG. This is the start codon utilized since the N-terminus of the R.*Bsu*BI enzyme (Fig. 4B, track 11) had the amino acid sequence M-T-E-G-M-H-S-N-X-V-K. Thus, the R.*Bsu*BI gene consists of 948 bp, with the capacity to encode a protein of 316 amino acids with a predicted M_r of 36.2 kD.

Identification of the transcriptional start point of the M.*Bsu*BI gene

The gene arrangement of the *Bsu*BI system suggests that both genes constitute an operon with a common promoter in front of the MTase gene. The transcriptional start point of the M.*Bsu*BI MTase gene was determined by primer extension using a 28mer oligonucleotide complementary to the coding strand of the M.*Bsu*BI gene (nt 240–267, see Fig. 2). A single cDNA reverse transcript was obtained using the total RNA isolated from *B.subtilis* ISB8 as a template (Fig. 3). The transcriptional start point was mapped at the C at nt 198, 30 bp upstream from the translational start codon of the MTase gene. A single reverse transcript was also detected when the same oligonucleotide was used to prime transcription of RNA isolated from *E.coli* cells

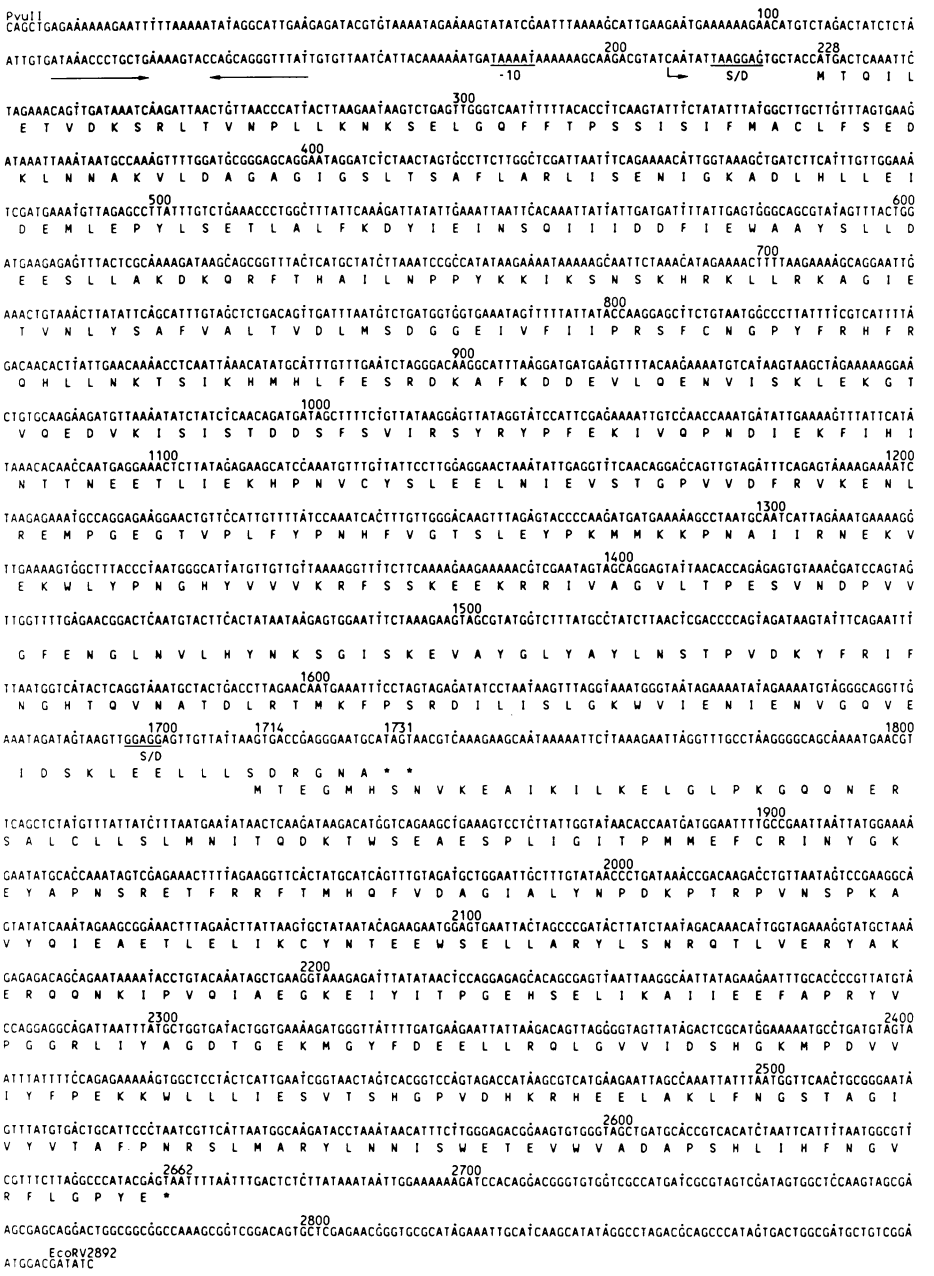

Figure 2. Nucleotide sequence of the *Bsu*BI MTase and ENase genes and deduced amino acid sequences of the corresponding enzymes. The proposed −10 box of the M.*Bsu*BI gene is indicated. The transcriptional start point of the M.*Bsu*BI gene is denoted with a short arrow at nt position 198. The putative S/D sequences of both *Bsu*BI genes are underlined. The translation stop codons are shown as asterisks. An inverted repeat of 14 bp in the 5′ region of the *Bsu*BI MTase gene is indicated by arrows.

containing the entire *Bsu*BI system on plasmid pBH64. The cDNA transcript had the same mobility as that from *B.subtilis* RNA, indicating that MTase transcription initiates at identical positions in *B.subtilis* and *E.coli* (Fig. 3). Analogous experiments with respect to the ENase gene were carried out with either of two oligonucleotides complementary to nt 1762−1789 and 1592−1620, respectively (see Fig. 2). No reverse transcript was detectable, suggesting that the initiation of transcription of the R.*Bsu*BI gene started far upstream from its translational start point.

Overproduction and purification of the *Bsu*BI enzymes

For protein purification, the 3.1 kb *Pvu*II-*Hin*dIII DNA fragment harbouring both the M.*Bsu*BI and the R.*Bsu*BI ENase genes was brought in the expression vector pMS119 E/H under the transcriptional control of the *tac* promoter (pMS3.1; Fig. 1B).

After IPTG induction, crude extracts of *E.coli* [pMS3.1], subjected to SDS-PAGE, contained only one additional protein band with an estimated M_r of 60 kD when compared to the control extract (Fig. 4A). The intensity of this band increased over a period of 4 hrs, parallel to an increasing methylation

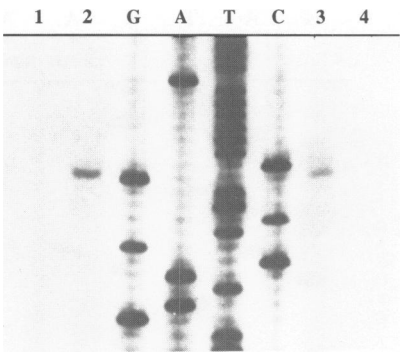

Figure 3. Primer extension analysis of the M.*Bsu*BI gene to determine the transcriptional start point. The reactions were carried out using a 5′ end-labelled oligonucleotide as primer with RNA isolated from (lane 1) *B.subtilis* YB886 (*hsr*BI$^-$, *hsm*BI$^-$) (44), (lane 2) *B.subtilis* ISB8 (*hsr*BI$^+$,*hsm*BI$^+$), (lane 3) *E.coli* NM675 [pBH64], (lane 4) *E.coli* NM675 [pBR328]. The adjacent DNA sequence ladders were produced using the same primer complementary to ssDNA of a M13mp18 derivative.

activity in the crude extracts (data not shown). The M_r of this protein band corresponded well to that predicted for the M.*Bsu*BI enzyme from sequencing. The identity of this band with M.*Bsu*BI was later confirmed by microsequencing of the protein partially purified by two chromatographic steps (Fig. 4A, lane 4).

Although the insert contained in pMS3.1 carries both genes of the *Bsu*BI R/M system (Fig. 1B), restriction activity was detected neither *in vitro* nor *in vivo*. In pMS3.1 the R.*Bsu*BI gene is separated from the *tac* promoter by the M.*Bsu*BI gene. Assuming that this arrangement might contribute to the failure to express R.*Bsu*BI from pMS3.1, we brought the R.*Bsu*BI gene in greater proximity to the *tac* promoter by inserting into pMS119 E/H a 1.9 kb *Ssp*I-*Hin*dIII DNA fragment, carrying only the intact R.*Bsu*BI gene (pMS1.9; Fig. 1B). This construct was introduced into NM675 cells containing a recombinant derivative, pGB3.6, of the compatible plasmid pGB2, which specified M.*Bsu*BI activity. After IPTG induction, *E.coli* NM675 [pMS1.9/pGB3.6] produced one additional protein band following SDS-PAGE, which was absent in the control, with an apparent M_r of about 35 kD. This M_r agrees well with that calculated for the deduced amino acid sequence of R.*Bsu*BI (36.2 kD). The intensity of this protein band correlated also with an increasing restriction activity determined *in vitro* and *in vivo*. Four chromatographic steps were carried out to purify the R.*Bsu*BI enzyme to apparent homogeneity (Fig. 4B). Purification yielded about 20 μg R.*Bsu*BI obtained from about 1.0 g cells.

The subunit structure of the active form of the R.*Bsu*BI enzyme was determined under native conditions by gel filtration on a G-100 Sephadex column (23). The active R.*Bsu*BI enzyme eluted at a volume which corresponds to a M_r of about 70 kD corresponding to the double M_r of the monomer, indicating that this enzyme like other ENases works as a dimer (data not shown).

Determination of the methylation and cleavage specificity of the *Bsu*BI enzymes

Previous results (4) suggested that enzymes of both the *Bsu*BI and *Pst*I R/M systems recognize the same sequence, 5′CTGC-AG. These analyses, however, did not provide information on whether the positions at which methylation or cleavage occurs within this sequence are identical for both systems.

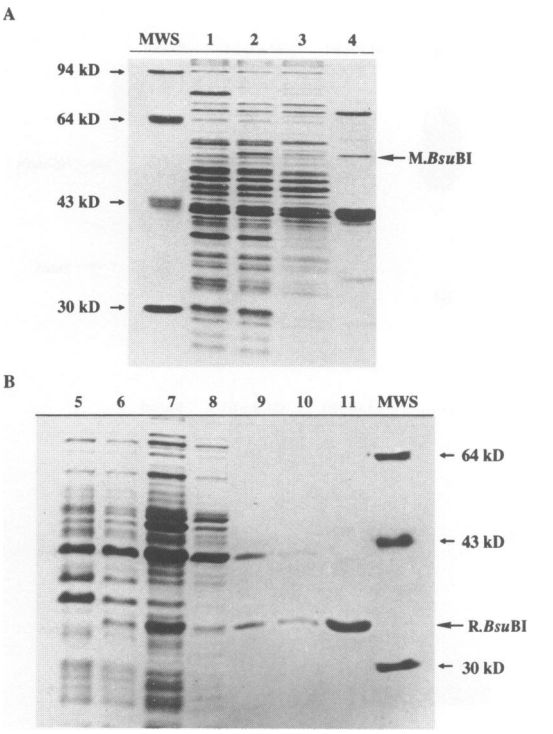

Figure 4. 0.1% SDS-15% polyacrylamide gel stained with Coomassie brilliant blue of crude cell extracts and fractions of the *Bsu*BI MTase and ENase during purification. MWS: molecular weight standard: Phosphorylase b (94 kD), BSA (64kD), ovalbumin (43 kD) and carbonic anhydrase (30 kD). **A.** M.*Bsu*BI: Lane 1 and 2: crude extracts of not induced and IPTG induced *E.coli* NM 675 [pMS3.1]; Lane 3: crude extract after DEAE Sepharose, and after DNA Cellulose chromatography (lane 4). **B.** R.*Bsu*BI: Lane 5: crude extract of IPTG induced NM675 [pMS119 E/H, pGB3.6] (control); Lane 6: crude extract of NM675 [pGB3.6, pMS1.9]; following lanes: fractions of the latter extract after different purification steps: (7) ammonium sulfate precipitation, (8) DEAE Sepharose, (9) S Sepharose, (10) DNA Cellulose and (11) gel filtration G-100.

To determine the base methylated by M.*Bsu*BI, synthetic DNA with multiple *Bsu*BI recognition sites (see Materials and Methods) was modified *in vitro* with M.*Bsu*BI using [^{3}H-methyl] SAM as a methyl group donor. The modified DNA was degraded to single nucleosides which were subsequently separated by paper chromatography (Fig. 5A) and analyzed by autoradiography (Fig. 5B). This showed that only *N*6-methyl-2′-deoxyadenosine is formed in the methylation reaction catalyzed by M.*Bsu*BI, identifying M.*Bsu*BI as a A-N6-MTase.

The unique *Bsu*BI/*Pst*I restriction site within the polylinker of M13 mp18 DNA, 35 bp away from the M13 (−40) sequencing primer site, was used to determine the cleavage specificity of the R.*Bsu*BI ENase by digestion of the products of a standard Sanger sequencing reaction, with R.*Bsu*BI. One specific cleavage product was observed which comigrated with an 'A' band of the corresponding sequence ladder (Fig. 6), and therefore terminated at the A residue. Thus, R.*Bsu*BI cleaves between A and G within 5′CTGCAG. Both with respect to the cleavage and methylation specificity *Bsu*BI and *Pst*I are isospecific R/M systems (6).

DISCUSSION

The *B.subtilis Bsu*BI R/M system has been found to be isospecific with the *Pst*I R/M system of *P.stuartii* (6). In addition to their sensitivity to A-methylation mediated by the genuine

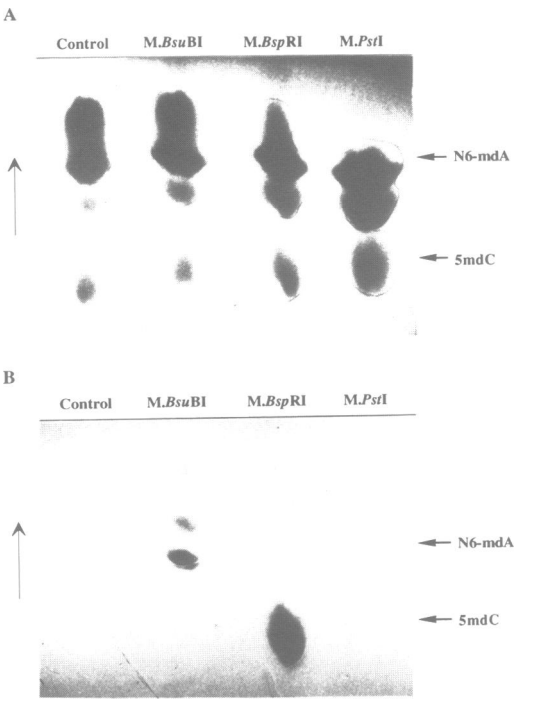

Figure 5. Determination of the nucleotide modified by M.*Bsu*BI. Synthetic DNA was methylated *in vitro* with either a crude extract of *E.coli* NM675 [pMS119 E/H] (control), of *E.coli* NM675 [pMS3.1] containing M.*Bsu*BI (M.*Bsu*BI), of *E. coli* HB101 containing the C5-MTase M.*Bsp*RI (M.*Bsp*RI) or the commercial *Pst*I MTase (M.*Pst*I). Unlabelled standard *N*6-methyl-2′-deoxyadenosine (N6-mdA) and 5-methyl-2′-deoxycytidine (5mdC) were added to the degradation products to provide UV absorption markers. **A.** Photography of separated N6-mdA and 5mdC on Whatman 3MM paper chromatogram observed under UV light. **B.** Autoradiography of the chromatogram. The locations of N6-mdA and 5mdC are indicated. Arrows show the running direction of the solvent.

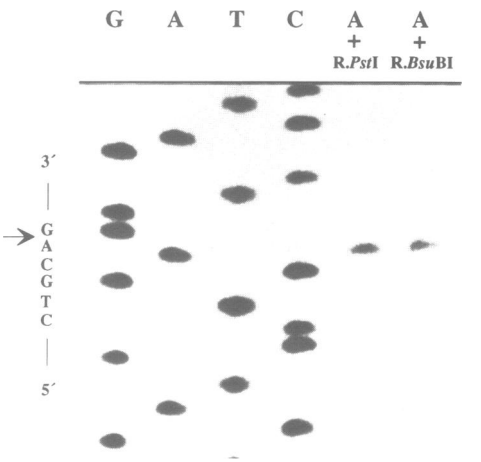

Figure 6. Determination of R.*Bsu*BI cleavage specificity. Lanes: G, A, T and C are the sequence ladders through the *Bsu*BI recognition site. Aliquots of the 'A' sequencing reaction was taken and digested either with R.*Bsu*BI or R.*Pst*I (control) before subjected on the 8% polyacrylamide sequencing gel.

M.*Bsu*BI/M.*Pst*I MTases, both restriction enzymes are also sensitive to C5 methylation of the target site. This could readily be shown with plasmid pϕB121 (37) which carries the gene for M.ϕ3TI, a multispecific MTase which modifies *Fnu*4HI and *Hae*III recognition target sites (38). pϕB121 contains two *Bsu*BI/*Pst*I sites both of which overlap with the *Fnu*4HI target sequence 5′GCNGC:

5′GCTGCAGC, 5′CTGCAGC.

pϕB121 DNA was neither cleavable at such sites by R.*Bsu*BI nor by R.*Pst*I (unpublished observations). This observation raises the open evolutionary question concerning the choice of modification associated with a given restriction activity. C5-methylation is the only type of modification that had been observed so far in *B.subtilis* (1). In spite of the fact that C5-methylation within the *Bsu*BI target would cause resistance to R.*Bsu*BI cleavage, modification in the *Bsu*BI R/M system is also provided by an A-N6-MTase, the first to be identified in *B.subtilis* (1).

In line with the functional similarity between the *Bsu*BI and *Pst*I R/M systems, both pairs of MTases and ENases show striking amino acid sequence similarity over their entire length (aa identities amount to 41.1% and 46.2, respectively). The similarities between the two MTase sequences include the two motifs GxGxG and NPPY, which are of presumable functional importance (35). In addition to these two conserved elements the primary sequence of M.*Bsu*BI contains a region of amino acids (181−195) strongly resembling 'block III' previously identified in seven other A-N6-MTases including M.*Pst*I (39, 40), whose target sites share the common sequence 5′TNNA. As this block is characteristic for such MTases it has been assumed to be involved in the recognition mechanism of the related target sequences of these enzymes (39, 40). The *Bsu*BI and *Pst*I R/M systems represent one of the rare cases, where also two ENases of isospecific R/M systems with related MTases have extensive sequence similarity (41, 42, 43). Such similarity was also also reflected at the gene level (58% nt identity for both genes), suggesting that the two systems originate from a common ancestor. Among isospecific R/M systems which imply A-N6-modification the only comparable relatedness was observed among the enzyme pairs of the *Taq*I and *Tth*HB8I systems (44) which are, however, encoded by two closely related Gram negative species.

Transcription of the two genes of the *Pst*I R/M system must involve the synthesis of two monocistronic mRNAs, permitting transcriptional control to coordinate the expression of both genes (6). Although it was not possible to perform Northern blot analysis with the transcript(s) of the *Bsu*BI genes (most likely due to its/their extreme instability), data obtained from the primer extension experiments identified a transcriptional initiation site only in the 5′ region of the MTase gene, strongly suggesting the joined genes are transcribed together as a dicistronic mRNA. In this configuration a coordinated expression of the R/M genes could also be provided at the transcriptional level.

The specific arrangement of the *Bsu*BI R/M genes might be of relevance for the observation that *E.coli* strains carrying recombinant plasmids harbouring both genes only expressed MTase but no ENase activity. The assumption of an arrangement specific effect underlying this observation follows from the finding that no expression of ENase activity in *E.coli* occured when both *Bsu*BI genes were brought under the control of the *tac* promoter, while expression of the ENase gene was readily detected when the MTase gene was removed (Fig. 1B). Experiments performed with the *E.coli*/*B.subtilis* shuttle plasmid

pHH64 carrying both genes further indicated that the gene arrangement dependent failure in the expression of ENase activity in *E.coli* was specific for the heterologous host. In *B.subtilis* both genes were expressed concommitantly. Further experiments have to clarify the significance of the gene arrangement and the possible role of a host factor in the control of the expression of the *Bsu*BI R/M genes.

The arrangement of the *Bsu*BI R/M genes described is strikingly different from that of the closely related genes of the *Pst*I R/M system which are transcribed divergently (6). The differences in the gene arrangements suggest that the two genes of each R/M system may have become joined in the two organisms in independent,stepwise events. Alternatively, different gene arrangements may have evolved following joint transfer of the ancestral genes into the different hosts. A long evolutionary history of the present *Pst*I and *Bsu*BI R/M systems is suggested by the fact that the nt sequences immediately bordering the R/M genes are different in *B.subtilis* and *P.stuarti* (data not shown).

ACKNOWLEDGEMENTS

We thank M. Noyer-Weidner for useful discussions and critical help during the preparation of the manuscript and J.Willert for advice and support during protein purification. We are grateful to V.Kruft for performance of the protein sequencing. G.-L.Xu is a recipient of a predoctoral fellowship granted within the framework of scientific cooperation between Max-Planck-Gesellschaft and Academia Sinica. J. Walter was supported by Deutsche Forschungsgemeinschaft (SFB 344).

REFERENCES

1. Trautner,T.A. and Noyer-Weidner,M. (1993) in *Bacillus subtilis*, the Model Gram-positive Bacterium: Physiology, Biochemistry and Molecular Genetics. (J. A. Hoch, R. Losick and A. L. Sonenshein eds) American Society for Microbiology, Washington, D.C. (in press).
2. Noyer-Weidner, M. and Trautner,T.A. (1992) in DNA Methylation: Biological Significance. (J. P. Jost and H. P. Saluz eds) Birkhäuser Verlag, Basel (in press) .
3. Wilson,G.G. (1991) Nucl. Acids Res., **19**, 2539–2566.
4. Shibata,T., Ikawa,S., Komatsu,Y., Ando,T. and Saito,H. (1979) J. Bacteriol., **139**, 308–310.
5. Walder,R.Y., Hartley,J.L., Donelson,J.E. and Walder,J.A. (1981) Proc. Natl.Acad. Sci., **78**, 1503–1507.
6. Walder,R.Y., Walder,J.A. and Donelson,J.E. (1984) J. Biol. Chem., **259**, 8015–8026.
7. Dobritsa,A.P. and Dobritsa,S.V. (1980) Gene, **10**, 105–112.
8. Woodcook,D.M., Crowthers,P.J., Doherty,J., Jefferson,S., DeCruz,E., Noyer-Weidner,M., Smith,S.S., Michael,M.Z. and Graham,M.W. (1989) Nucl. Acids Res., **17**, 3469–3478.
9. Tanaka,T. (1979) Mol. Gen. Genet., **175**, 235–237.
10. Larsen,J.E.L., Gerdes,K., Light,J. and Molin,S. (1984) Gene, **28**, 45–54.
11. Bolivar,F., Rodriguez,R., Greene,P.J., Betlach,M.C., Heyneker,H.L., Boyer,H.W., Crosa,J.H. and Falkow,S. (1977) Gene, **2**, 95–113.
12. Churchwald,G., Belin,D. and Nagamine,Y. (1984) Gene, **31**, 165–171.
13. Haima,P., Bron,S. and Venema,G. (1987) Mol. Gen. Genet., **209**, 335–342.
14. Fürste,J.P., Pansegrau,W., Frank,R., Blöcker,H., Scholz,P., Bagdasarian,M. and Lanka, E. (1986) Gene, **48**, 119–131.
15. Shibata,T. and Ando,T. (1974) Mol. Gen. Genet., **131**, 275–280.
16. Biswal,N., Kleinschmidt,A.K., Spatz,H.Ch. and Trautner,T.A. (1967) Mol. Gen. Genet., **100**, 39–55.
17. Luria,J. and Burrous,J.W. (1957) J. Bacteriol., **74**, 461–476.
18. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. Günthert,U., Freund,M. and Trautner,T.A. (1981) J. Biol. Chem., **256**, 9340–9345.
20. Trautner,T.A., Pawlek,B., Bron,S. and Anagnostopoulos,C. (1974) Mol. Gen. Genet., **131**, 181–191.
21. Davies,R.W., Botstein,D. and Roth,J.R. (1980) in Advanced Bacterial Genetics. A Manual for Genetic Engineering. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. Chai,S., Bravo,A., Lüder,G., Nedlin,A., Trautner,T.A. and Alonso,J.C. (1992) J. Mol. Biol., **224**, 87–102.
23. Kapfer,W., Walter,J. and Trautner,T.A. (1991) Nucl. Acids Res., **19**, 6457–6463.
24. Posfai,G., Kiss,A. and Venetianer,P. (1986) Gene, **50**, 63–67.
25. Woodcock,D.M., Crowther,P.J., Simmons,D.L. and Cooper,I.A. (1984) Biochim. Biophys. Acta., **783**, 227–233.
26. Harper,J.E., Miceli,S.M., Roberts,R.J. and Manley,J.L. (1990) Nucl. Acids Res., **18**, 5735–5741.
27. Brown,N.L. and Smith,M. (1980) Methods Enzymol., **65**, 391–404.
28. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA, **74**, 5463–5467.
29. Bradford,M. (1976) Anal. Biochem., **72**, 248–254.
30. Devereux,J., Haeberli,P. and Smithies,O. (1984) Nucl. Acids Res., **12**, 387–395.
31. Szomolanyi,E., Kiss,A. and Venetianer, P. (1980) Gene **10**, 219–225
32. Lauster,R., Trautner,T.A. and Noyer-Weidner,M. (1989) J. Mol. Biol., **206**, 305–312.
33. Posfai,J., Bhagwat,A.S., Posfai,G. and Roberts,R.J. (1989) Nucl. Acids Res., **17**, 2421–2435.
34. Chandrasegaran,S. and Smith,H. O. (1988) in Structure and Expression: From Proteins to Ribosomes (Sarma, R.H. & Sarma, M.H., eds) Adenine Press, NY., Vol. 1, pp 149–156.
35. Klimasauskas,S., Timinskas,A., Menkevicius,S., Butkiene,D., Butkus,V. and Janulaitis,A. (1989) Nucl. Acids Res., **17**, 9823- 9832.
36. Moran Jr.,C.P., Lang,N., LeGrice,S.F.J., Lee,G., Stephens,M., Sonenshein, A.L., Pero,J. and Losick,R. (1982) Mol. Gen. Genet., **186**, 339–346.
37. Günthert,U., Reiners,L. and Lauster,R. (1986) Gene **41**, 261–267.
38. Noyer-Weidner,M., Jentsch,S., Pawlek,B., Günthert,U. and Trautner,T.A.(1983) J. Virol., **46**, 446–453.
39. Ito,H., Shimato,H., Sadaoka,A., Kotani,H., Kimizuka,F. and Kato,I. (1992) Nucl. Acids Res., **20**, 705–709.
40. Ito,H., Sadaoka,A., Kotani,H., Hiraoka,N. and Nakamura,T. (1990) Nucl. Acids Res., **18**, 3903–3911.
41. Wilson,G.G. and Murray,N.E. (1991) Annu. Rev. Genet., **25**, 585–627.
42. Erdmann,D., Düsterhöft,A. and Kröger,M. (1991) Eur. J. Biochem. **202**, 1247–1256.
43. Nölling,J. and de Vos,W.M. (1992) J. Bacteriol. **174**, 5719–5726.
44. Barany,F., Danzitz,M., Zebala,J. and Mayer,A. (1992) Gene, **112**, 3–12.
45. Yasbin, R. E., Fields, P. I. and Andersen. B. J. (1980) Gene, **12**, 155–159.