Nucleotide sequence of the BsuRI restriction-modification system

Antal Kiss, György Posfai+, Christopher C.Keller, Pal Venetianer+ and Richard J. Roberts

Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, NY 11724, USA, and +Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, P.O. Box 521, Szeged, Hungary

Received 20 August 1985; Accepted 27 August 1985

ABSTRACT

The genes of the 5'-GGCC specific BsuRI restriction-modification system of Bacillus subtilis have been cloned and expressed in E. coli and their nucleotide sequence has been determined. The restriction and modification genes code for polypeptides with calculated molecular weights of 66,314 and 49,642, respectively. Both enzymes are coded by the same DNA strand. The restriction gene is upstream of the methylase gene and the coding regions are separated by 780 bp. Analysis of the RNA transcripts by S1-nuclease mapping indicates that the restriction and modification genes are transcribed from different promoters. Comparison of the amino acid sequences revealed no homology between the BsuRI restriction and modification enzymes. There are, however, regions of homology between the BsuRI methylase and two other GGCC specific modification enzymes, the BspRI and SPR methylases.

INTRODUCTION

Type II restriction and modification enzymes are promising as model systems for the study of sequence-specific DNA-protein interactions. The development of molecular cloning, DNA sequencing and in vitro mutagenesis techniques, together with more classical biochemical methods, provide the tools that are necessary to study the detailed molecular mechanism of sequence-specific DNA recognition. One of the attractions of the restriction and modification enzymes is that for most recognition sequences different enzymes recognizing the same sequence are available. Comparison of these enzymes may help to elucidate some general rules for DNA sequence recognition. Although a number of type II restriction and modification genes have been cloned, only the EcoRI (1,2), HhaI (3), PstI (4) and EcoRV (5) systems have been sequenced. In addition to this, the nucleotide sequences of the E. coli dam (6), BspRI (7), SPR (8,9) and T4 dam (10) methylases have also been determined.

The restriction-modification system BsuRI of B. subtilis R was discovered and genetically mapped on the bacterial chromosome by T. Trautner and coworkers (11). The BsuRI nuclease (12) cleaves the recognition sequence
5'-GGCC between the central G and C (13) and the modification enzyme methylates the inner C (14). The BsuRI restriction and modification enzymes have been purified to homogeneity and characterized biochemically (15-17). We reported the cloning and expression in E. coli of the gene that codes for the BsuRI methylase (18). Here we report the cloning of the BsuRI endonuclease gene and the nucleotide sequence of the complete BsuRI system.

MATERIALS AND METHODS

Strains and media

Bacillus subtilis R (11) was provided by Dr. T. Trautner. The E. coli strains HB101 (19) and RR1 (20) were used as hosts in plasmid cloning experiments and JM107 (21) was used as host for cloning in phage M13. B. subtilis and E. coli were grown in LB medium (22), at 37 °C.

Restriction endonucleases were either prepared in the Biol. Res. Cent. of the Hung. Acad. Sci. according to published protocols (23) or were purchased from New England Biolabs. DNA polymerase I large (Klenow) fragment was from BRL, S1 nuclease from Boeringer Mannheim and polynucleotide kinase from New England Biolabs. Synthetic oligonucleotides were prepared by M. Zoller using an Applied Biosystem Synthesizer. All other chemicals were reagent grade commercial products.

Preparation of plasmids, transformation of E. coli, restriction mapping, agarose gel electrophoresis and subcloning of DNA fragments were done by standard procedures (22).

Cloning the BsuRI endonuclease gene

DNA purified from B. subtilis R (18) was digested with SphI and SalI and ligated to pBR322 (20) cleaved with the same enzymes. The ligated DNA was transformed in E. coli RR1. AmpR transformants (approximately 8,000 TetS recombinants) were grown to saturation in 200 ml LB containing 100 μg/ml ampicillin and the cells were used for preparation of plasmid DNA. This plasmid DNA was digested with HaeIII and transformed in E. coli RR1. AmpR transformants were selected. HaeIII (an isoschizomer of BsuRI) could be used to select for the BsuRI methylase gene because the BsuRI-specific methylation protects the DNA against HaeIII cleavage (14).

Detection of BsuRI endonuclease activity in the clones

Cells from 100 ml saturated cultures of the E. coli clones carrying the cloned BsuRI genes were sedimented by centrifugation, washed in 20 mM
Nucleic Acids Research

TrisHCl pH 8.0, resuspended in 1 ml of 50 mM TrisHCl pH 8.0, 5 mM 2-mercaptoethanol, 0.1 mM EDTA and disrupted by sonication. After adding NaCl to 1M, the homogenate was centrifuged at 15,000 g for 1 hr at 4 °C. The supernatant was purified further by gel filtration on a 2.5 x 43 cm Bio-Gel A-0.5m column in 10 mM TrisHCl pH 8.0, 10 mM 2-mercaptoethanol and 1 M NaCl. 1.5 μl aliquots of the fractions were assayed for BsuRI endonuclease activity in 25 μl reaction mixtures containing 1 μg pBR322 DNA, 10 mM TrisHCl pH 7.5, 10 mM MgCl₂, 50 mM NaCl and 1 mM dithiothreitol. The samples were incubated for 1 hour at 37 °C and analyzed on 1% agarose gels.

Measurement of in vivo restriction

Nonmodified and modified λvir and φ80 c phages were prepared by growing the phage on either E. coli HB101 (pBR322) or E. coli HB101 carrying the plasmid pSU1 which codes for the BsuRI methylase. In vivo restriction was measured by determining the restriction ratio i.e. the ratio between the titers of a nonmodified phage on the host investigated and on a nonrestricting host. The plating efficiency was determined by standard methods (24).

Determination of the nucleotide sequence

Part of the DNA sequence (nucleotides 1-2904) was determined by the chain termination method (25) using M13 clones as templates, the other part (nucleotides 2616-4253) by the chemical cleavage method (26).

Chain termination method: Specific fragments were cloned in the M13 vectors mp18 and mp19 (27) and sequenced (28) using the synthetic oligonucleotide 5'-GTAAAACGACGGCCAGT as universal primer and ³⁵S-α-dATP as label (29). The sequencing products were run on 40 x 20 x 0.04 cm 6 or 8% polyacrylamide gels containing 8 M urea in 100 mM Tris-borate pH 8.3. After the run the gels were fixed in 10% methanol, 10% acetic acid for 10 min, dried and exposed to autoradiographic film at room temperature for 1 - 2 days. Sequence data from the autoradiograms were entered directly into the computer using a digitizing tablet (30).

Chemical cleavage method: Labeling of restriction fragments by polynucleotide kinase, sequencing reactions and gel electrophoresis were carried out as described in (26).

Mapping of the transcripts

Preparation of single-stranded probes: Two synthetic oligonucleotides, An3 and An4, complementary to nucleotides 314-334 and 2837-2857 (Fig.3.) were used. Following synthesis, oligonucleotides were fractionated on a 16% polyacrylamide gel containing 7 M urea and, after elution, were purified
further on a small Sep-Pak column (31). 0.2 pmol of An3 and An4 labeled with $^{32}$P at the 5'-end were annealed in 10 μl containing 10 mM TrisHCl pH 7.5, 10 mM MgCl$_2$ to approximately 1 μg of the single-stranded M13 templates 20/1 and 9/6, respectively. The M13 clone 20/1 contains the 494 bp EcoRI-BglII fragment covering the 5'-end of the restriction gene (nucleotides 1-494) and clone 9/6 contains the 434 bp EcoRI-BglII fragment covering the 5'-end of the modification gene (nucleotides 2466-2899), both cloned between the EcoRI and BamHI sites in the vector mp18.

After annealing, 7 μl of 10 mM TrisHCl pH 7.9 and 6 μl of a solution containing 0.5 mM of the four deoxynucleotide triphosphates were added and the primers were extended with 1 μl (3 U) of DNA polymerase Klenow fragment for 20 min at room temperature. The reaction was stopped by the addition of 40 mM EDTA pH 8.0, the DNA was extracted with phenol/chloroform, twice with chloroform and precipitated with ethanol. The precipitated DNA was dissolved and digested with EcoRI, then extracted with phenol and chloroform and precipitated with ethanol. The terminally labeled 329 nucleotide (probe R) and 387 nucleotide (probe M) long single-stranded fragments (Fig. 3.) were isolated (26) from a 6% polyacrylamide gel containing 8 M urea.

$S_1$ mapping: 10-14 μg total bacterial RNA isolated (32) from B. subtilis R or E. coli HB101 carrying pSU13 was hybridized to probes R or M (3-10 x $10^3$ cpm) in 50% formamide, 4 x SSC at 37 °C for 14 hrs then the hybridization mixture (15 μl) was added to 300 μl $S_1$ buffer (0.045 M Na-acetate pH 4.6, 0.28 M NaCl, 0.005 M ZnSO$_4$) and digested with 5000 U (Boeringer) of $S_1$ nuclease for 1 hour at 30 °C. After adding 10 μl of 0.5 M EDTA pH 8.0 and 10 μg tRNA the reaction mixtures were extracted with phenol and chloroform, then the $S_1$-resistant hybrids were precipitated with ethanol. The fragments protected from $S_1$-digestion were analyzed by running them side-by-side, on 6% polyacrylamide/8 M urea gels, with sequencing products obtained from sequencing reactions on templates 20/1 and 9/6 using the kinase-labeled An3 and An4 oligonucleotides as sequencing primers. Positions of the $S_1$-resistant fragments determined the transcription initiation sites.

Computer analysis of the sequence

Handling and analysis of nucleic acid sequence data were done with computer programs developed at the Cold Spring Harbor Laboratory (30) and at the Biological Research Center of the Hungarian Academy of Sciences (J.
RESULTS

Cloning of the R.BsuRI gene

Clones pSU1 and pSU11, two overlapping clones isolated independently, cover an approximately 4.7 kb region of B. subtilis DNA (18). Both clones express the BsuRI methylase but they do not code for the BsuRI endonuclease (18). We knew from genetic data (11) that the BsuRI genes were linked. Therefore, it seemed possible that the endonuclease and methylase genes could be cloned together on a larger fragment. The 1.3 kb HindIII-BgIII fragment containing the 5'-end of the methylase gene and the 5'-flanking sequences (Fig. 1) was isolated from pSU1 and used as hybridization probe to identify, by Southern blotting, larger fragments of the B. subtilis chromosomal DNA which would carry the methylase gene and would extend into the neighbouring regions (data not shown). One such fragment was a 9.5 kb SphI-Sall fragment. Cloning in E. coli RR1 of this fragment was carried out using selection for the methylase as described in Materials and Methods. The resulting clone (pSU12) carried the 9.5 kb SphI-Sall fragment.

DNA isolated from this clone was resistant to HaeIII showing that it carried and expressed the BsuRI methylase gene. (BsuRI methylation protects the DNA against HaeIII cleavage; ref. 14). To test whether the plasmid also coded for the restriction enzyme, we prepared a cell-free extract from the clone and fractionated it on a Bio-Gel A-0.5m column. Some of the fractions showed BsuRI activity indicating the presence and expression of the BsuRI endonuclease gene (Fig. 2).

Colonies of the clone RR1(pSU12) were slightly heterogeneous in size on LB agar plate, therefore the plasmid pSU12 from the original isolate was transformed in E. coli HB101. HB101(pSU12) always gave homogeneous cultures. The observed difference between the RR1 and HB101 clones was not further investigated. In the rest of the work HB101 was used as host.

To test whether the clone showed in vivo restriction, the phages λvir and φ80c grown either on the nonmodifying host E. coli HB101(pBR322) or on the modifying host E. coli HB101(pSU1) were used to infect the E. coli clone carrying pSU12. The modified phage plated with an e.o.p. of 1, whereas the non-modified phage was restricted (e.o.p.: 10^-4 - 10^-5, data not shown).

It was known from previous data (18) that the methylase gene was close to the SalI site of the insert in pSU12 (Fig. 1). To localize the restriction
gene, shorter subfragments (Fig. 1) of the original 9.5 kb SphI-SalI insert were cloned in pBR322 and tested for restriction. The shortest derivative still showing in vivo restriction was pSU19 (Fig. 1).

**Nucleotide sequence of the genes**

We determined the sequence of 4253 nucleotides starting at the nearest EcoRI site upstream of the endonuclease gene (Fig. 1 and 3). The sequence contains only two large open reading frames, both on the same strand: the first (nucleotides 297-2027) contains 576 amino acids, while the second (2808-4118) contains 436 amino acids (Fig. 3). On the basis of the restriction-modification phenotype of the clones carrying different portions of
Fig. 2. Endonuclease assay of fractions obtained from Bio-Gel chromatography of a cell-free extract prepared from the E. coli clone carrying pSU12. Panel A: Bio-Gel fractions. Panel B: Time course of digestion. pBR322 DNA was digested for different lengths of time with 1/4 of fraction 10 shown in panel A. Lanes HaeIII: pBR322 DNA digested with HaeIII.

the cloned B. subtilis DNA (Fig. 1), the first reading frame was assigned to the endonuclease and the second to the methylase. The two reading frames are separated by 780 bp. This intergenic region is longer than any other so far found in type II restriction-modification systems (1-5). The endonuclease gene codes for a 66,314 dalton protein and the methylase gene for a 49,642 dalton protein.

The sequencing led to the unexpected observation that the BsuRI methylase coded by pSU1 and by its derivative pSU15 (18, Fig. 1) is not the complete protein. These plasmids lack the sequence downstream of the HindIII site at position 4077 (Fig. 3), and which contains the extreme 3'-end of the
Fig. 3. Nucleotide sequence of the BsuRI genes and the deduced amino acid sequence of the proteins. Shine-Dalgarno sequences are indicated by closed circles under the sequence, transcriptional initiation sites by arrow-heads and the self-complementary sequence of the putative transcriptional termination site by horizontal arrows. The -10 sequences of the promoters are boxed. Regions complementary to probes R and M used in the transcription mapping are underlined with a solid line (corresponding to oligonucleotide An3 and An4) or with a dashed line (corresponding to the part of the probes made by primer extension).
Fig. 4. S₁-nuclease mapping of the BsuRI transcripts. Gel electrophoresis of fragments of the restriction gene specific (R) and modification gene specific (M) probes protected from S₁ digestion by RNA isolated from E. coli (pSU13) or B. subtilis R. 20/1 and 9/6: sequencing products obtained using template 20/1 and 9/6.

methylase gene. In these constructs the truncated methylase gene is linked to pBR322 via this HindIII site, in the same orientation as the Tet⁺ gene. Examination of the sequence around the junction shows that translation of the
methylase ends at an in-phase UAA codon immediately after the *HindIII* site:

\[
\text{*HindIII* stop} \\
\text{methylase.....GAA CGC AAG CTT TAA TGC.....pBR322}
\]

Thus both pSU1 and pSU15 code for a functional *BsuRI* methylase which lacks the 11 amino acids from the C-terminus of the wild-type protein.

### Transcription of the genes

As the two genes are in the same orientation, in principle they could be transcribed as a single mRNA, starting from the endonuclease promoter. However, this seemed unlikely since the methylase gene was active in several recombinant plasmids where the endonuclease gene was not present (Fig. 1.).

The transcriptional initiation points were determined by S1 mapping using RNA isolated from *B. subtilis* and *E. coli* HB101(pSU13) cells. Fig. 4. shows that there are separate initiation sites for the two genes. The transcription of the endonuclease starts with U and A at positions 235 and 236, respectively, and the methylase transcript starts with A at 2779. It can also be seen that the major initiation sites are identical in *B. subtilis* and *E. coli*. The sequences preceding the initiation sites were searched for promoter-like structures. *B. subtilis* is known to have several different forms of RNA polymerase which can use different promoter sequences (35). The consensus sequence of the promoters recognized by the major form of RNA polymerase present in vegetative cells (G55) is identical with the consensus sequence for *E. coli* promoters: TTGACA for the -35 and TATAAT for the -10 region (36,37). Examination of the sequence preceding the initiation sites of the *BsuRI* nuclease and methylase transcripts reveals the presence of hexanucleotides (Fig. 3) showing homology with the canonical -10 structure. The promoters of the *BsuRI* genes seem, however, atypical because they lack an appropriately spaced -35 region. The sequences in the promoter regions do not resemble promoters utilized by minor forms of *B. subtilis* RNA polymerase (35).

The transcriptional termination of the *BsuRI* genes has not been investigated experimentally. In *E. coli* (36) and probably also in *B. subtilis* (38) transcriptional terminators are characterized by a self-complementary structure followed by several T residues. We can find a similar sequence downstream of the methylase stop codon (Fig. 3), nucleotides 4140-4150 and 4164-4174 and we would propose that this potential stem-and-loop structure probably functions as a transcriptional terminator. No similar terminator-like...
sequence can be found downstream of the endonuclease gene (Fig. 3.).

To test whether the endonuclease gene is expressed in the absence of the methylase we tried to delete the methylase gene from the plasmid pSU13. This was done by digesting the plasmid with BamHI (Fig. 1B) and then ligating it at low DNA concentration. Aliquots of the ligation mixture were used to transform HB101 and HB101(pSU184-11) cells. pSU184-11 is a pACYC184 derivative carrying the functional BsuRI methylase gene on the 3.7 kb EcoRI fragment originally cloned in pSU11 (Fig. 1). Equal amounts of ligated DNA yielded approximately 100 times fewer transformants in HB101 than in HB101(pSU184-11).

Restriction analysis of six clones obtained in the HB101 host revealed that all retained the 2.4 kb BamHI fragment coding for the methylase. Two clones contained this BamHI fragment in the original orientation, (i.e. they were identical with pSU13), four clones contained it in the opposite orientation. Inverting the methylase gene did not seem to influence the r-m phenotype: the DNA was fully resistant to HaeIII and the clone showed the same level of restriction in vivo as pSU13. As expected from the difference in the transformation efficiencies, most of the Tc^R Amp^R clones obtained with the HB101(pSU184-11) host did not contain the 2.4 kb BamHI fragment; the BsuRI methylase gene on the compatible replicon compensated for the deleted gene of pSU13. Phenotypically (methylation of DNA and in vivo restriction) these clones were indistinguishable from pSU13.

From these experiments we conclude that expression of the endonuclease does not require the presence of active methylase and the endonuclease is lethal without the methylase. Although we did not try to quantitate the level of methylase in those derivatives where the methylase gene has been turned around or put on another compatible replicon, these observations confirm the conclusion drawn from transcript mapping, that the two genes do not constitute an operon. They are transcribed independently.

Translational signals

The AUG start codons of the BsuRI genes are preceded by sequences showing complementarity with the 3'-end of B. subtilis 16 S RNA (ref. 39, Fig. 3). The ribosomal binding site of the methylase allows for a much stronger Shine-Dalgarno interaction (calculated free energy -18.8 kcal/mol; ref. 40) than that of the nuclease (-9.4 kcal/mol).

Base composition and codon usage

The base composition of the BsuRI genes is characterized by a high A + T content: 68.4% for the nuclease and 61.9% for the methylase. This is
higher than the average A + T content of *B. subtilis* DNA (57%, ref. 41). The base composition is reflected in the codon usage: A and U are strongly preferred nucleotides in the third position or, whenever possible, in the first position of codons (Table 1).

**Comparison of protein sequences**

In addition to the *BsuRI* enzymes, the amino acid sequences of two other proteins recognizing the tetranucleotide GGCC, are known. These are the
Fig. 5. Sequence comparisons between the M.BsuRI, M.BspRI, M.SPR and R.BsuRI enzymes. Dot matrix outputs were generated by a computer program (DIAGON, ref. 32) which utilizes Dayhoff's similarity scores between amino acids. Diagonal stretches represent regions of homology.

BspRI methylase of Bacillus sphaericus (7) and the SPR methylase of the B. subtilis phage SPR (8,9). The M.BspRI and M.SPR enzymes were found to share partial sequence homology (8,9). We performed a computer search to test whether there is any sequence similarity between the BsuRI endonuclease and methylase or between the BspRI and SPR methylases and the BsuRI enzymes. No significant homology was detected between the BsuRI endonuclease and any of the three methylases. There are, however, regions of homology between the BsuRI methylase and the two other modification enzymes (Fig. 5). The similarity of the amino acid sequences is especially strong between the BsuRI and BspRI methylases, the homology extends over almost the whole molecule (Fig. 5 and 6). Weaker, but significant homology exists between the SPR methylase and the two other methylases. These sequence similarities were found in two blocks (Fig. 6). In these regions many amino acids are conserved in all three enzymes. The enzymes are also very similar in size, the BspRI methylase consists of 424, the BsuRI methylase 436 and the SPR methylase of 439 amino acids.
DISCUSSION

The potential lethality of r-m systems is of primary concern in all attempts aimed at cloning these genes. Although several complete r-m systems have been cloned (1-5 and G. Wilson, pers. comm.), in at least as many other cases the cloning attempts have failed. Cloning of the *BsuRI* genes has been easily accomplished in the host *E. coli* RR1 and the system was found to be stable in *E. coli* HB101. On the other hand, we were unable to transform several other *E. coli* strains (e.g. DH1) even with plasmids coding only for the methylase. We found that plasmids coding for the *BspRI* or *SPR* methylases behaved in the same way and the inability to transform certain *E. coli* strains was due to the methylase function. Other investigators have made similar observations with genes coding for different modification enzymes that methylate cytosine (R. Blumenthal, T. Trautner and G. Wilson, pers. comm.).

In addition to the *BsuRI* system, the organization of five other type II r-m systems (*EcoRI*, *HhaI*, *PstI*, *EcoRV* and *PaeR7*) has been determined and very different gene arrangements have been found (1-5 and J. Brooks, pers. comm.). The arrangement of the *BsuRI* genes is similar to the *EcoRI* genes in
the sense that they are tandemly arranged and the restriction gene is upstream of the methylase gene, but there is considerable difference in the length of the intergenic region (21 bp between the EcoRI genes versus 780 bp between the *BsuRI* genes).

One intriguing observation about cloned r-m systems is that plasmids coding for both enzymes transform *E. coli* with a frequency similar to that of other plasmids of similar size. Clearly, some mechanism must exist which ensures that expression of the endonuclease is delayed compared to the methylase. Sequential transcription (3), difference between the synthesis rates of the nuclease and the methylase, time needed for the nuclease to assemble in active dimer form (4) and inhibition of the nuclease translation by a potential higher structure of the nuclease mRNA (5) have been suggested as control mechanisms. In the *BsuRI* system the endonuclease gene is upstream of the methylase gene, therefore sequential transcription, as was suggested for the *HhaII* system (3), cannot play a role in regulating the expression of the nuclease. The *BsuRI* nuclease is thought to be a monomeric enzyme (15) so dimerization cannot be a regulating factor. A search of the sequence, that is likely to correspond to the *BsuRI* endonuclease mRNA, for self-complementary regions failed to reveal any that could form a secondary structure similar to that predicted for the *EcoRV* endonuclease mRNA (5). At present the only feature of the *BsuRI* system that might suggest a control mechanism is the difference in the ribosomal binding sites. The Shine-Dalgarno sequence preceding the methylase start codon is much stronger than that preceding the nuclease start codon and this difference might be important in the regulation of the relative amounts of methylase and endonuclease.

The calculated molecular weight of the endonuclease (66,314) is in good agreement with the value observed experimentally (68,000; ref. 15). The evaluation of the molecular weight calculated for the methylase (49,642) is more problematic. Gunthert et. al. (16,17) found two methylases associated with the *BsuRI* system. The enzymes differed slightly in their enzymological characteristics and in their molecular weights. Depending upon the method used these were 37-43 kd for M.*BsuRI*a and 39-43 kd for M.*BsuRI*b. It was not determined whether the two enzymes differ as a result of post-translational modification or whether they are coded by different genes (16, 17). Considering the similarity of the two enzymes and the genetic data available for the *BsuRI* system (11) the first possibility seems more likely. We assume that the gene we have characterized in this paper codes for a precursor that would undergo post-translational processing to give rise to the two enzyme
forms. The same mechanism was suggested to explain the existence of two methylases in the HpaII system (42). Obviously, further studies are needed to understand this phenomenon and the functioning of the BsuRI system in general, but the availability of the cloned genes and their sequence should prove very useful for future work.

Our main purpose with the study of restriction-modification systems is to learn how these enzymes recognize a specific sequence in the DNA. The approach we have taken is to clone and sequence genes of enzymes recognizing the same DNA sequence. It seemed likely that comparison of these proteins may help find the common structural determinants of the sequence-specific interaction. At present there are four GGCC-specific enzymes (BspRI, BsuRI and SPR methylases and the BsuRI nuclease) for which the protein sequence (derived from the DNA sequence) is known (7,8,9 and this paper).

Comparison at the amino acid level of the BsuRI methylase and nuclease did not reveal any significant degree of homology. Investigation of two other RM systems (EcoRI, PstI) led to similar conclusions (1,2,4). Similarly, in a computer search of the amino acid sequences published for the HhaI and EcoRV enzymes (3,5) we failed to detect any homology between the endonuclease and methylase of the same system. Although it is possible that analysis of the three-dimensional structure may reveal common structural elements in the nuclease and methylase belonging to the same system, it now seems more likely that the lack of amino acid sequence homology reflects the different nature of the molecular mechanisms by which the nuclease and the methylase interact with their target sequence. Enzymological studies support this conclusion (43).

Comparison of the BspRI and SPR methylases detected partial homology between the amino acid sequences (8,9). In this paper we show that there is a similar degree of homology between the SPR and BsuRI methylases and much higher homology between the BspRI and BsuRI methylases. We interpret the homologies detected as an indication of evolutionary relatedness. It remains to be seen, however, whether these homologies are related to the enzymatic function.

ACKNOWLEDGEMENTS

We thank Dr. T. Trautner for the strain B. subtilis R, Dr. M. Zoller for synthesis of the oligonucleotides, Dr. J. Posfai for doing some of the computer work, Drs. T. Trautner, R. Blumenthal, G. Wilson and J. Brooks for
communicating their results before publication, Dr. Ashok Bhagwat for his helpful comments on the manuscript and E. Csorba for the skillful technical assistance. Part of this work was supported by NSF grant DMB 8217553. A. K. was supported under a collaborative agreement between Cold Spring Harbor Laboratory and Exxon Research and Engineering.

REFERENCES