Volume 12 Number 23 1984 Nucleic Acids Research

Structure of the gene coding for the sequence-specific DNA-methyltransferase of the *B. subtilis* phage SPR

György Pósfai, Frank Baldauf+, Sára Erdei, János Pósfai[1], Pál Venetianer and Antal Kiss*

Institute of Biochemistry and [1]Institute of Biophysics, Biological Research Center of the Hungarian Academy of Sciences, Szeged, P.O.Box 521, Hungary

Received 16 May 1984; Revised and Accepted 18 October 1984

ABSTRACT

The nucleotide sequence of the gene coding for the 5'-GGCC and 5'-CCGG specific DNA methyltransferase of the *Bacillus subtilis* phage SPR was determined by the Maxam-Gilbert procedure. Transcriptional and translational signals of the sequence were assigned with the help of S[1] mapping and translation in *E. coli* minicells. The gene codes for a 49 kd polypeptide. The amino acid sequence of the SPR methylase shows regions of homology with the sequence of the 5'-GGCC-specific *Bsp*RI modification methylase.

INTRODUCTION

DNA methyltransferases offer attractive possibilities to investigate the problems of sequence-specific DNA-protein interaction. A necessary prerequisite of any such study is the detailed knowledge of the structure of the interacting protein. The rapid advance of recombinant DNA technology and DNA sequencing allowed the determination of the amino acid sequence of numerous proteins indirectly, that is by cloning and sequencing the respective genes. Using this approach, the structures of several sequence-specific DNA-methylases have been determined (1-5).

Some of the temperate phages of *Bacillus subtilis* carry genes coding for sequence-specific DNA methyltransferases (6). Phage SPR belongs to this group. SPR was previously thought to be identical with SP*β*, but later it was shown that the two phages are different (7). The methyltransferase activity of this phage protects DNA against cleavage by *Hpa*II, *Msp*I and *Bsu*RI (or *Bsp*RI) nucleases (8-10). The gene coding for these activities has been cloned and shown to be expressed in *E. coli* (10,11,11a). In this paper we report the DNA sequence of this gene, together with the assignment of the transcriptional and translational signals. We also compare the amino acid sequence of the SPR methylase (M.SPR) with that of the *Bsp*RI modification methylase (M.*Bsp*RI, ref. 4) and find significant homologies between the two sequences.

While this work was in progress, the SPR methylase gene sequence has also been determined by another group (12).

© IRL Press Limited, Oxford, England.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

MATERIALS AND METHODS

Strains, plasmids and media

All recombinant DNA work used *E. coli* HB101 host (13) and pBR322 vector (14). Plasmids pSU2 and pSU21 carry the functional SPR methylase gene on a 4.3 kb *Eco*RI and on a 2.1 kb *Eco*RI-*Pst*I fragment, respectively, cloned in pBR322 (10). Plasmid pSU23 contains the 3′ half of the gene on a 3 kb *Sal*I-*Eco*RI fragment in pBR322 (10). *E. coli* DS410 (15) was used for minicell experiments. Media and growth conditions were described previously (10).

Enzymes and chemicals

Most restriction endonucleases, T4 polynucleotide kinase and T4 DNA ligase were prepared in this laboratory. *Hin*fI was a gift of Dr. M. Hartmann (Jena). *Taq*I and BAL-31 were from New England Biolabs, bacterial alkaline phosphatase from Worthington and S_1 nuclease from Sigma. *Eco*RI linkers were from Collaborative Research and [γ-^{32}P] ATP was the product of the Isotope Institute, Budapest. All other chemicals were reagent grade commercial products.

Deletion mapping with BAL-31 nuclease

Plasmid pSU21 linearized either with *Eco*RI or *Pst*I was digested for various lengths of time with BAL-31 as recommended by the supplier. The progress of the reaction was monitored by gel electrophoresis. *Eco*RI linkers were ligated to the BAL-31 treated DNA, then the DNA was digested with *Eco*RI (16). After this the plasmids which had been linearized at and shortened from the *Pst*I site were ligated and used to transform *E. coli* HB101. Plasmids which had been linearized at the *Eco*RI site, were digested with *Pst*I, then mixed with *Eco*RI- and *Pst*I-digested pBR322, ligated and transformed into *E. coli*. Tc^R Ap^S transformants were selected and analysed for SPR-specific methylation by digesting their plasmids with *Bsp*RI (isoschizomer of *Bsu*RI), *Msp*I and *Hpa*II.

S_1 mapping, translation in minicells

RNA was isolated (17) from *E. coli* HB101 containing plasmid pSU2. The method of hybridization, S_1 digestion and subsequent electrophoresis in denaturing polyacrylamide gels followed the procedure described in (16). Translation in minicells was done as described previously (4).

DNA sequencing

DNA sequencing was done by the Maxam-Gilbert method (18) using G, A+G, C+T and C reactions. Electrophoresis was carried out in 0.2 mm x 250 mm x 400 mm 8% polyacrylamide gels containing 8.3 M urea.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

Computer analysis of the sequence

Data handling, reading frame search and analysis of homologies were performed by a FORTRAN program package in a PDP-compatible minicomputer.

RESULTS

Localization of the SPR methylase gene

The starting material was plasmid pSU21, a pBR322-derived recombinant plasmid containing a 2.1 kb *Eco*RI-*Pst*I fragment of SPR DNA (10). pSU21 carries the functional SPR methylase gene, as indicated by the resistance of the plasmid to *Bsp*RI, *Hpa*II and *Msp*I (10). We tried to localize the gene on the 2.1 kb insert by linearizing the plasmid with either *Eco*RI or *Pst*I and by generating deletions by BAL-31 digestion. The methylation of the resulting plasmids carrying deletions of different lengths was tested by digestion with *Bsp*RI, *Hpa*II and *Msp*I. As shown in Fig. 1, the deletions defined a 1.4 kb long region of DNA located (approximately) between nucleotides 550 and 1950 (numbering starts at the *Eco*RI site) which was sufficient to code for methylase activity. As the disappearance of the resistance to *Bsp*RI, *Hpa*II and *Msp*I was simultaneous, these experiments confirmed our earlier assumption (10) that a single enzyme was responsible for all observed methylation activities.

DNA sequence

First a detailed physical map of the 2.1 kb insert of pSU21 was

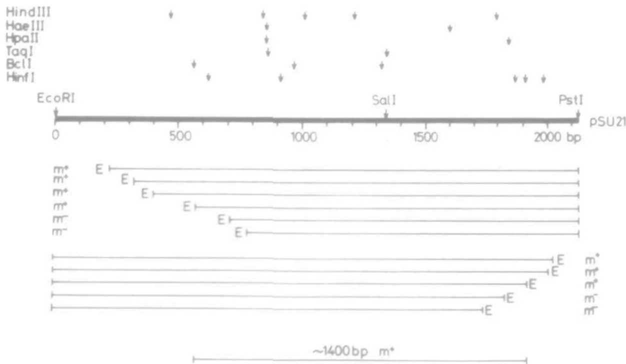


Fig. 1. Mapping of the SPR methylase gene by restriction enzymes and by BAL-31 deletions. Heavy line: the 2100 bp SPR DNA fragment cloned in plasmid pSU21. The cleavage sites of selected restriction enzymes are marked by arrows. Thin lines: cloned deletional derivatives of the insert of pSU21 which retain (m^+) or lose (m^-) the ability to express the methylase. E: *Eco*RI sites produced with *Eco*RI linkers. At the bottom the predicted shortest m^+ region is shown.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

```
289   GTTGCTTAAATTAAAATTTAGAAATCAAACACTGGAATTGCGTAGAAAAATGTCCTTACAATTACACCGTTTATTAGAA

368   GACATGACAGCTTGGTCATTTAATGAAAATGAAGCTCAAAAATATATTGATGAATATGAATTAGAGTTAGTTGAGAAGA

447   GTGTAGACAATAAATTTAAGCTTCAAATAGTAGATAAGAATAAAAATATTTGGGGCGAATTAAAACCCAAAGAGCCCAGG

526   GTACATTGGATATTTTGTACAGACTGAATAAATGATCAATTTTATTTAGAGTAAAAATAAAATATATGGAGGTTGTTTA

605   TTG GGT AAA CTA CGT GTA ATG AGT CTT TTT AGT GGG ATC GGT GCA TTT GAA GCT GCA CTA
      M   G   K   L   R   V   M   S   L   F   S   G   I   G   A   F   E   A   A   L

665   AGA AAC ATT GGG GTT GGT TAT GAG CTC GTT GGT TTT AGT GAG ATT GAT AAA TAT GCC GTC
      R   N   I   G   V   G   Y   E   L   V   G   F   S   E   I   D   K   Y   A   V

725   AAA TCT TTT TGT GCA ATT CAC AAC GTT GAT GAG CAA TTA AAT TTT GGA GAT GTA AGC AAG
      K   S   F   C   A   I   H   N   V   D   E   Q   L   N   F   G   D   V   S   K

785   ATT GAT AAG AAA AAA CTA CCT GAA TTT GAT CTT TTA GTT GGA GGA TCT CCT TGT CAA AGC
      I   D   K   K   K   L   P   E   F   D   L   L   V   G   G   S   P   C   Q   S

845   TTT AGT GTA GCC GGC CAT CGA AAG GGA TTT GAA GAT ACA AGA GGG ACA TTG TTT TTT CAA
      F   S   V   A   G   H   R   K   G   F   E   D   T   R   G   T   L   F   F   Q

905   TAC GTT GAG ACT CTT AAG GAA AAG CAA CCA AAG TTT TTT GTT TTT GAA AAT GTT AAA GGG
      Y   V   E   T   L   K   E   K   Q   P   K   F   F   V   F   E   N   V   K   G

965   TTG ATC AAC CAT GAT AAA GGA AAT ACA TTA AAT GTT ATG GCT GAA GCT TTC AGT GAA GTT
      L   I   N   H   D   K   G   N   T   L   N   V   M   A   E   A   F   S   E   V

1025  GGG TAC AGA ATT GAC CTA GAG CTG CTT AAT TCA AAA TTC TTT AAT GTT CCA CAA AAT AGG
      G   Y   R   I   D   L   E   L   L   N   S   K   F   F   N   V   P   Q   N   R

1085  GAG CGA CTT TAC ATA ATT GGA ATT AGA GAA GAT TTA ATT AAA AAT GAA GAA TGG TCT TTG
      E   R   L   Y   I   I   G   I   R   E   D   L   I   K   N   E   E   W   S   L

1145  GAT TTT AAA AGA AAG GAT ATA CTT CAA AAA GGG AAA CAG AGA TTG GTA GAA TTA GAT ATT
      D   F   K   R   K   D   I   L   Q   K   G   K   Q   R   L   V   E   L   D   I

1205  AAA AGC TTT AAT TTT AGA TGG ACA GCT CAA TCG GCT GCT ACG AAG AGG CTA AAA GAT TTA
      K   S   F   N   F   R   W   T   A   Q   S   A   A   T   K   R   L   K   D   L

1265  TTA GAA GAA TAC GTT GAT GAA AAG TAC TAC TTG AAT GAA GAT AAA ACA AAC AGT TTG ATC
      L   E   E   Y   V   D   E   K   Y   Y   L   N   E   D   K   T   N   S   L   I

1325  AAA GAG TTG TCT ACA AGT CGA CTT AAT GAA AAT CTT ACT GTT GAG CAA GTA GGT AAC ATT
      K   E   L   S   T   S   R   L   N   E   N   L   T   V   E   Q   V   G   N   I

1385  AAT CCC TCT GGT AAT GGA ATG AAT GGA AAT GTT TAT AAT TCA TCT GGA TTA AGC CCC ACA
      N   P   S   G   N   G   M   N   G   N   V   Y   N   S   S   G   L   S   P   T

1445  ATT ACC ACT AAT AAA GGA GAG GGA CTG AAA ATT GCA GTT GAG TAC TCC AGA AAA AGC GGC
      I   T   T   N   K   G   E   G   L   K   I   A   V   E   Y   S   R   K   S   G

1505  CTT GGA CGA GAA CTT GCT GTA TCT CAT ACG CTT TCT GCT TCT GAC TGG AGA GGA TTG AAT
      L   G   R   E   L   A   V   S   H   T   L   S   A   S   D   W   R   G   L   N

1565  AGG AAC CAA AAA CAA AAT GCA GTT GTT GAG GTA AGG CCA GTA TTA ACC CCA GAA AGG GGC
      R   N   Q   K   Q   N   A   V   V   E   V   R   P   V   L   T   P   E   R   G

1625  GAG AAG CGA CAA AAT GGA AGA AGA TTT AAA GAT GAC GGT GAA CCA GCA TTT ACA GTA AAC
      E   K   R   Q   N   G   R   R   F   K   D   D   G   E   P   A   F   T   V   N

1685  ACA ATT GAC AGA CAC GGG GTA GCG GTT GGA GAG TAT CCA AAA TAC AGA ATT AGA AGA TTA
      T   I   D   R   H   G   V   A   V   G   E   Y   P   K   Y   R   I   R   R   L

1745  ACA CCG TTA GAG TGC TTT AGG CTA CAG GCT TTT GAT GAC GAA GAT TTT GAA AAA GCT TTT
      T   P   L   E   C   F   R   L   Q   A   F   D   D   E   D   F   E   K   A   F

1805  GCT GCG GGA ATA AGT AAC TCA CAA TTA TAT AAG CAA GCC GGT AAT TCA ATT ACT GTA ACT
      A   A   G   I   S   N   S   Q   L   Y   K   Q   A   G   N   S   I   T   V   T

1865  GTG CTT GAG TCA ATA TTC AAG GAA TTA ATA CAT ACA TAC GTT AAT AAA GAA TCT GAA TAA
      V   L   E   S   I   F   K   E   L   I   H   T   Y   V   N   K   E   S   E

1925  AATTTGTCTTTTAAACAAATGTAAAATAAGAGGAGTGGATTGATTGGAAAGTTACCCTGAGTCCTTAAAAAGAGAGACA

2004  GAGGAGATTAAAGAGCGTGTTAGGAATGGAAATATCAAAGAAGACAGGATTAAAGAAATTGCAGAAACGACAGTTGAGT

2083  TTTTGAAATCAGAGGAGAAAAGACATAAATACTTTTCTGAAG
```


Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

established (19). In order to avoid any ambiguities in the DNA sequence both strands of the fragment were sequenced, and all restriction sites used in fragment preparation were sequenced by using overlapping fragments. The sequence shown on Fig. 2 contains only one large open reading frame corresponding to the region established by deletion mapping. This begins with an ATG at position 623 and ends with TAA at position 1922.

Transcription

Previous experiments from this laboratory (10) suggested that the cloned SPR methylase gene was expressed from its own promoter because it was active in both orientations. In contrast, other investigators (11a) reported that expression must be due to an external promoter. We tried to localize the promoter of the methylase gene by S_1 mapping. RNA isolated from *E. coli* HB101 (pSU2) was hybridized to the 857 bp *Eco*RI-*Hpa*II fragment encompassing the 5' end of the gene (Fig. 1 and 2). The result of this experiment is shown on Fig. 3A. The larger protected fragment is similar in lenght to the probe and probably indicates readthrough transcription from the vector. The shorter, 280 nucleotide long fragment indicates an initiation site at position 578. Examination of the sequence upstream from this initiation site reveals the presence of a promoter-like structure sharing homology with the *E. coli* and *B. subtilis* consensus -35 and -10 sequences (21,22).

```
                                                              578
M.SPR:        ATATTTTGTACAGACTGAATAAATGATCAATTTTATTTAGAGTAAAAA
consensus:          TTGACA         18 bp      TATAAT
```

Thus we think that this structure functions, at least in *E. coli*, as the promoter of the SPR methylase gene This finding confirms our previous observation that the SPR methylase is expressed in *E. coli* from its own promoter. Vector promoter(s) may, however, contribute to the transcription of the gene. In fact, we have found (unpublished observation) that the level of expression of the cloned SPR methylase was slightly dependent on the orientation of the gene in pBR322.

Translation

The SPR methylase consists of a single subunit with a molecular weight of

Fig. 2. Nucleotide sequence of the region encoding the SPR methylase and the predicted amino acid sequence of the methylase protein. Numbering starts at the last nucleotide of the *Eco*RI site of the pSU21 insert. Only part of the determined DNA sequence is shown. The suggested transcriptional start point is marked by an arrow and Shine-Dalgarno complementarity by asterisks. The proposed promoter sequence is boxed. Discrepancies between our sequence data and that of Bukh *et al.* (12) are underlined.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

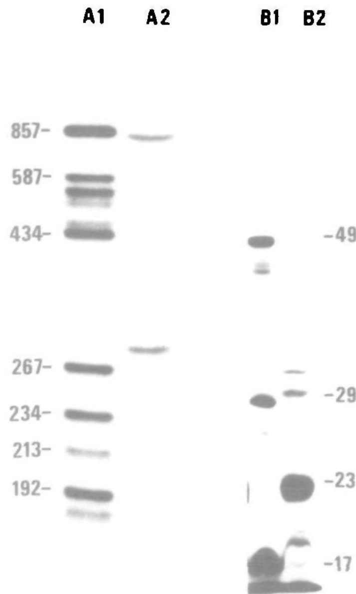


Fig. 3. S_1 mapping of the transcript and translation in minicells.
Panel A: Autoradiogram showing the result of the S_1 nuclease protection experiment.
A1: molecular weight markers. A2: DNA fragments protected by *in vivo* synthesized RNA. The 857 bp *Eco*RI-*Hpa*II fragment (Fig. 1 and 2) ^{32}P-labeled at the *Hpa*II end with polynucleotide kinase was used as probe. The length of the 280 nucleotide long protected fragment has been verified on a Maxam-Gilbert gel, where it was run along a sequence ladder (not shown).
Panel B: Autoradiogram of a 5-20% SDS-polyacrylamide gradient gel after electrophoresis of the ^{35}S-labeled proteins synthesized in *E. coli* minicells. B1: Minicells programmed by pSU21 (m^+). B2: Minicells programmed by pSU23 (m^-) which contains an extensive deletion of the methylase gene (10). Molecular weights of marker proteins are indicated (in kilodaltons).

47-49 kd (6 and U. Günthert, pers. comm). We could detect an approximately 49 kd protein in *E coli* minicells containing methylase-positive plasmids (Fig. 3). This strong protein band was never detected in minicells carrying plasmids (e.g. pSU23, Fig. 3) with inactive derivatives of the methylase gene.

The long open reading frame detected in the sequence begins with ATG at position 623. This methionine codon is, however, not preceded by a properly positioned Shine-Dalgarno sequence. In view of the fact that in *Bacillus* species the requirement for this sequence seems to be rather strict (23), we considered this assignment unlikely. In the region between the transcription initiation point at 578 and the ATG mentioned above a perfect Shine-Dalgarno

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

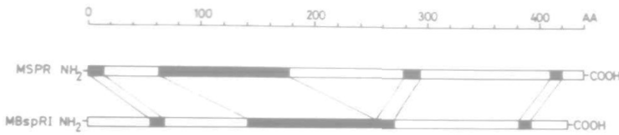


Fig. 4. Schematic comparison of the amino acid sequences of the SPR and the *Bsp*RI methyltransferases. Black bars represent regions of partial homology shown in detail in the text.

sequence can be detected at position 593-598. This is followed by an appropriately spaced (22) in-phase TTG codon (position 605). As UUG has been shown to function as an initiation codon in the case of the *B. subtilis* phage SP01 (22,23), we believe that this UUG is more likely to be the real translational start codon of the methylase gene. This would correspond to a calculated molecular weight of 49,826, in good agreement with the estimate based on electrophoretic mobility.

Protein structure

The amino acid sequence of the methylase gene, deduced from the DNA sequence, is shown in Fig. 2. This has been compared with the sequence of other sequence-specific DNA methyltransferases (M.*Eco*RI, M.*Eco*dam, M.*Bsp*RI and M.*Hha*II; ref. 1-5). Of these four enzymes, only M.*Bsp*RI shares recognition-specificity with the SPR methylase; both enzymes recognize and methylate GGCC. Although in previous DNA hybridization experiments we failed to detect homology between the genes of M.SPR and M.*Bsp*RI (10), the comparison of the amino acid sequences reveals a significant degree of homology between these two enzymes (see below and Fig. 4).

DISCUSSION

The SPR-coded DNA methylation was shown to affect 5'-GGCC and 5'-CCGG sequences (9). The experiments reported here prove that these activities reside in a single protein. The dual specificity contributes to the attraction of the SPR methylase as model-system for the study of sequence-specific DNA recognition.

The principal aim of this work was the determination of the amino acid sequence of the SPR methylase in order to compare it with the sequence of other methylases. The BAL-31 deletions unambiguously localized the gene to a 1.4 kb segment on a cloned fragment. Within this region a large open reading frame coding for a 49 kd protein was found. This protein is synthesized in *E. coli* minicells programmed with plasmids coding for the functional SPR methylase

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

but cannot be detected in minicells programmed with methylase plasmid derivatives.

At the DNA level, the sequence reported here differs from that reported by another group (12) at five points, two of which lead to differences at the protein level. The translational terminator of the open reading frame is at position 1922. The positioning of the transcriptional and translational initiation sequences is less straightforward. The S_1 mapping experiment localizes transcription initiation in the heterologous system to position 578. The structure upstream from the transcriptional initiation site, which we identified in the heterologous system as the promoter of the methylase gene, displays less conformity to the canonical -35 and -10 sequences than most *B. subtilis* promoters utilized by the σ^{55} RNA polymerase (22). It is possible, however, that the SPR methylase gene is transcribed in *B. subtilis* by another form of RNA polymerase. Buhk et al. (12) place the initiation point in the homologous system to position 591, thus it must be concluded that the point of transcription initiation is different in the homologous and heterologous systems.

On the basis of the translation experiments it is impossible to choose between the two possible translational start-points; both calculated molecular weights (49,219 and 49,826) can be reconciled with the observed 49 kd. Despite the fact that ATG is the generally preferred initiation signal we believe that in this case, in *B. subtilis*, the TTG at 605 is the actual start point for the following reasons:

1. It is preceded by an appropriately located good Shine-Dalgarno sequence: GGAGGT. No such sequence is present before the ATG at 623.

2. UUG has been shown to function as an initiation codon for wild-type genes in Gram-positive bacteria (22,23,24).

3. The amino acids coded between 605 and 623 show some homology with a domain of the *Bsp*RI methylase:

```
              60
BspRI :   ..SDKFNVLSLFCGAG..
SPR :       MGKLRVMSLFSGIG..
                   10
```

This homology is obvious at the DNA level as well:

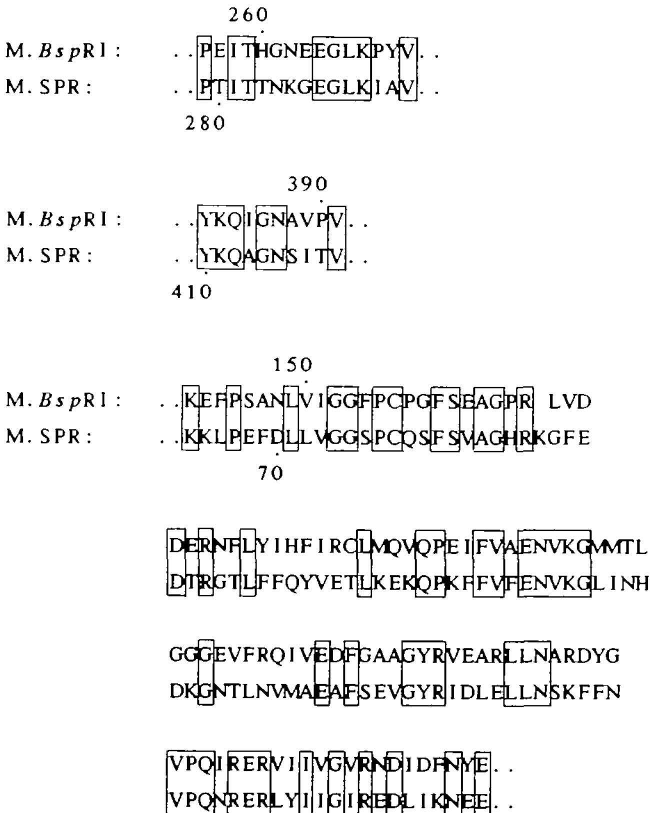
```
BspRI :   ..TCGGATAAGTTTAATGTTCTTTCTTTGTTCTGTGGAGCAGGA..
SPR :       TTGGGTAAACTACGTGTAATGAGTCTTTTTAGTGGGATCGGT..
```

The assignment of the translational start codon in *E. coli* is more difficult. Although it has been shown that in an *in vitro* system *E. coli* ribosomes can

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

utilize UUG as translational start signal (23,26), UUG has not been found as initiation codon in wild-type *E. coli* mRNAs (25). Furthermore, less extensive Shine-Dalgarno interaction is required to promote translation in Gram negative than in Gram positive bacteria (22,25). Due to lack of protein sequencing our data are not sufficient to decide whether translation of the SPR methylase in *E. coli* starts with TTG at position 605 or with ATG at 623.

Apart from the N-terminal domain discussed above, two other shorter and one rather long region of homology can be detected between the amino acid sequences of M.SPR and M.*Bsp*RI (Fig. 4).



Of course at present it would be premature to speculate about the role of any of these homologies. Nevertheless it seems likely that at least some of these regions might play a role in the recognition of the sequence GGCC by both enzymes. This might be a reasonable working hypothesis in the further investigation of the problem of sequence-specific protein-DNA interaction.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

ACKNOWLEDGEMENTS

Thanks are due to Dr. T.A. Trautner for communicating his results prior to publication, to Mrs. E. Csorba for valuable technical assistance, Dr. R.J. Roberts for comments on the manuscript and Mrs. J. Dellaporta for typing it.

+Permanent address: Department of Genetics, Section Biosciences, Martin-Luther-University, DDR-4020 Halle/S. Domplatz 1, GDR

*Present address: Cold Spring Harbor Laboratory, P.O.Box 100, Cold Spring Harbor, NY, USA

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024

REFERENCES

1. Greene, P.J., Gupta, M., Boyer, H.W., Brown, W.E. and Rosenberg, J.M. (1981) J. Biol. Chem. *256*, 2143-2153.
2. Newmann, A.K., Rubin, R.A., Kim, S. and Modrich, P. (1981) J. Biol. Chem. *256*, 2131-2139.
3. Brooks, J.E., Blumenthal, R.M. and Gingeras, T.R. (1983) Nucleic Acids Res. *11*, 837-851.
4. Pósfai, G., Kiss, A., Erdei, S., Pósfai, J. and Venetianer, P. (1983) J. Mol. Biol. *170*, 597-610.
5. Schoner, B., Kelly, S. and Smith, H.O. (1983) Gene *24*, 227-236.
6. Günthert, U. and Trautner, T.A. (1984) in *Current Topics in Microbiology and Immunology*, Trautner, T.A. ed., Vol. *108*, pp. 11-22, Springer Verlag, Berlin-Heidelberg-New York-Tokyo.
7. Noyer-Weidner, M., Jentsch, S., Pawlek, B., Günthert, U. and Trautner, T.A. (1983) J. Virol. *46*, 446-453.
8. Trautner, T.A., Pawlek, B., Günthert, U., Canosi, U., Jentsch, S. and Freund, M. (1980) Mol. Gen. Genet. *180*, 361-367.
9. Jentsch, S., Günthert, U. and Trautner, T.A. (1981) Nucleic Acids Res. *9*, 2753-2759.
10. Kiss, A. and Baldauf, F. (1983) Gene *21*, 111-119.
11. Behrens, B., Pawlek, B., Morelli, G. and Trautner, T.A. (1983) Mol. Gen. Genet. *189*, 10-16.
11a. Montenegro, M.A., Pawlek, B., Behrens, B. and Trautner, T.A. (1983) Mol. Gen. Genet. *189*, 17-20.
12. Buhk, H.J., Behrens, B., Tailor, R., Wilke, K., Prada, J.J., Günthert, U., Noyer-Weidner, M., Jentsch, S. and Trautner, T.A. Gene, in press.
13. Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. *41*, 459-472.
14. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Cross, J.H. and Falkow, S. (1977) Gene *2*, 95-113.
15. Dougan, G. and Sherratt, D. (1977) Mol. Gen. Genet. *151*, 151-160.
16. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
17. Brosius, J., Cate, R.L. and Perlmutter, A.P. (1982) J. Biol. Chem. *257*, 9205-9210.
18. Maxam, A.M. and Gilbert, W. (1980) in *Methods in Enzymology*, Grossman, L. and Moldave, K., ed., Vol. *65*, pp. 499-560, Part I. Academic Press.
19. Smith, H.O. and Birnstiel, M.L. (1976) Nucleic Acids Res. *3*, 2387-2398.
20. Stuber, D and Bujard, H. (1981) Proc. Natl. Acad. Sci. U.S.A. *78*, 167-171.

21. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. *13*, 319–353.
22. Moran, 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.
23. McLaughlin, J.R. Murray, C.L. and Rabinowitz, J.C. (1981) J. Biol. Chem. *256*, 11283–11291.
24. Harwood, C.R., Williams, D.M. and Lovett, P.S. (1983) Gene *24*, 163–169.
25. Gold, L., Pribnow, D., Schneider, T., Schinedling, S., Singer, B.S. and Stormo, G. (1981) Ann. Rev. Microbiol. *35*, 365–403.
26. McLaughlin, J.R., Murray, C.L. and Rabinowitz, J.C. (1981) J. Biol. Chem. *256*, 11273–11281.

Downloaded from https://academic.oup.com/nar/article/12/23/9039/2379400 by guest on 16 April 2024