JOURNAL OF BACTERIOLOGY, May 1992, p. 3395–3398
0021-9193/92/103395-04$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 174, No. 10

Sequence and Characterization of *pvuIIR*, the *Pvu*II Endonuclease Gene, and of *pvuIIC*, Its Regulatory Gene

TAO TAO† AND ROBERT M. BLUMENTHAL*

Department of Microbiology, Medical College of Ohio, P.O. Box 10008, Toledo, Ohio 43699-0008

Received 20 February 1992/Accepted 9 March 1992

An open reading frame partially overlaps *pvuIIR*, and genetic evidence implies that this open reading frame, named *pvuIIC*, specifies a positive regulator of *pvuIIR* (T. Tao, J. C. Bourne, and R. M. Blumenthal, J. Bacteriol. 173:1367–1375, 1991). Inducible constructs of *pvuIIC* produced a protein of the expected size. The site of C · *Pvu*II action appears to lie within *pvuIIC* itself; thus, *pvuIIC* may be a self-contained regulatory cassette.

Many type II restriction modification systems (RMS2s) are coded for by mobile genetic elements and must be regulated on entering a new host cell to allow methylation of host DNA before restriction activity appears. These regulatory mechanisms are still not well understood.

The *Pvu*II RMS2 from *Proteus vulgaris* recognizes the sequence 5′-CAGCTG-3′ (10), and its genes were previously cloned in this laboratory (4). The *Pvu*II DNA methyltransferase (M · *Pvu*II) generates N^4-methylcytosine at the internal C (4, 7), the restriction endonuclease (R · *Pvu*II) cleaves between the central two bases (10), and the genes *pvuIIM* and *pvuIIR* are divergently oriented with respect to one another (4). We previously reported the sequence of *pvuIIM* (21). Another laboratory reported the sequence of *pvuIIR* (1), although without corroborating evidence. We had independently determined the sequence of *pvuIIR* and agree substantially with the previously reported version.

During that determination, we discovered *pvuIIC*, a third open reading frame (ORF) between *pvuIIR* and *pvuIIM* (20). Its product, C · *Pvu*II, appears to be a *trans*-acting factor that stimulates expression of *pvuIIR* up to 10^5-fold and belongs to a newly discovered family of very similar helix-turn-helix proteins (6, 20). We report here the sequence of *pvuIIC*, along with evidence that *pvuIIC* does produce a protein of the expected size and that its site of action on *pvuIIR* is probably within *pvuIIC* itself.

Base sequence of *pvuIICR*. A 1,167-bp region was sequenced by using the dideoxy method, synthetic primers, and a double-stranded template known to produce active restriction endonuclease (2, 4, 17). We also isolated IS*1* insertions by cloning *pvuIIR* in the absence of a protective methyltransferase and sequenced them with two IS*1*-specific primers. Throughout the coding region, both strands were sequenced (Fig. 1A) and each set of reactions was repeated four to six times. Our sequence differed at 22 positions from that reported by Anthanasiadis et al. (1; indicated in our GenBank entry), although none of these differences lie within the major ORFs. The N-terminal amino acids of R · *PvuII* were determined (3). The Met had been removed in vivo, but the next 11 residues match the sequence prediction exactly (Fig. 1B).

The sequence predicts a 157-amino-acid 18.3-kDa polypeptide, in good agreement with our previous estimate of 20 kDa from in vitro expression of *pvuIIR* (4). Thus, R · *Pvu*II is the smallest restriction endonuclease found so far; the next smallest is *Mun*I, with 202 amino acids (23, 24). A partially purified extract of a *pvuIIR*-overexpressing construct, applied to a calibrated gel filtration column, yielded a single activity peak corresponding to a Stokes radius of 2.8 nm and a mass of about 40.8 kDa (data not shown). This value is 2.2 times that calculated from the primary sequence and implies that the active form of R · *Pvu*II is a homodimer, like most other RMS2 restriction endonucleases (5, 13, 14, 18).

The *pvuIIR* sequence was joined to that of the adjacent *pvuIIM* gene, which we had determined earlier (21), and has been deposited with GenBank. The joined sequences revealed *pvuIIC* between the *pvuIIR* and *pvuIIM* genes (20). The sequence of *pvuIIC* is shown in Fig. 1B, and it predicts that the *pvuIIC*-encoded polypeptide is 84 amino acids long, has a molecular weight of 9,403 and a denatured pI of 9.28, and contains no Phe, Trp, or Cys.

Inducible overexpression of *pvuIIC*. To demonstrate that a *pvuIIC*-specified protein is actually made, we put the *pvuIIC* gene under the control of a T7 promoter as shown in Fig. 2. A cotransformant of *Escherichia coli* JM107MA2 (4), carrying plasmids p*Pvu*C1.2/T7pBS and pGP1-2, was grown at 30°C. Plasmid pGP1-2 carries the gene for T7 RNA polymerase under the control of the thermolabile λ *c*I *ts*857 repressor (19). In the late log phase, the culture was shifted to 42°C and rifampin was added after 20 min. Time point aliquots were labeled for 5 min with a ^{35}S-labeled Met-Cys mixture, chased with unlabeled Met, and processed for electrophoresis. For the induced culture, equal trichloroacetic acid-precipitable counts were resolved on a sodium dodecyl sulfate (SDS)–20% polyacrylamide gel modified to resolve small proteins (11). A labeled band corresponding to a molecular weight of 8,900 ± 900 (Fig. 3, lanes A to D) was induced, and this band was not prominent in an uninduced culture (lane E). Cotransformants of pGP1-2 and parental vector pBS(−) did not produce this band (data not shown).

The induced protein accumulated to substantial levels, even in the absence of rifampin. In a separate experiment (data not shown), JM107MA2(p*Pvu*C1.2/T7pBS, pGP1-2) was grown at 30°C in M9 medium to the mid-log phase. The culture was then shifted to 42°C without rifampin, and 30 min after induction ^{35}S-labeled Met-Cys was added to the culture for continuous labeling. Aliquots were removed at times up to 22 h later and then chased for 5 min with excess unlabeled Met. Autoradiographs of the resulting SDS gels were

* Corresponding author.
† Present address: Department of Pharmacology, Case Western Reserve University School of Medicine, Cleveland, OH 44106.

Downloaded from http://jb.asm.org/ on November 22, 2019 by guest

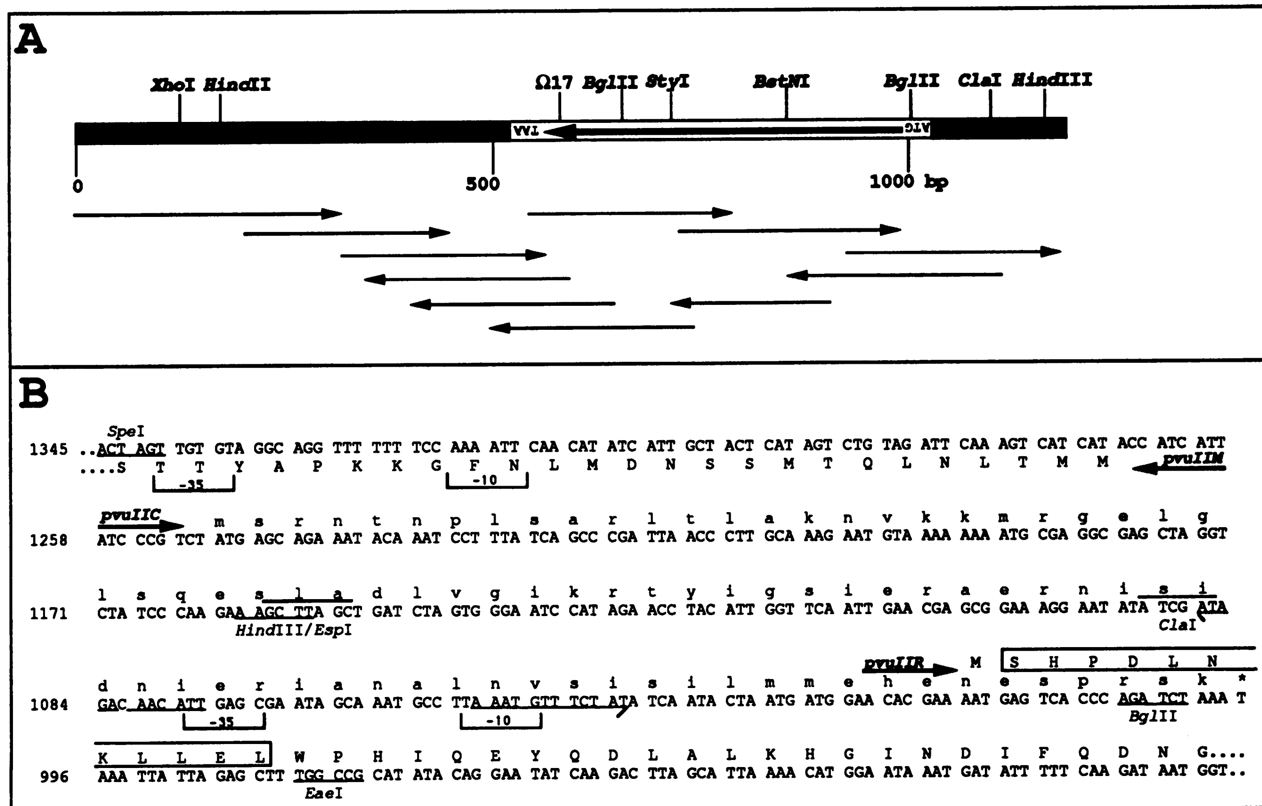


FIG. 1. (A) Strategy for sequence determination of the *pvuIIR* gene. Synthetic oligonucleotide primers were used as shown, together with plasmid p*PvuRM*3.4 (4) and its derivatives. Each arrow shows the direction and length of the sequence determined by using one primer. Ω17 refers to an IS*1* insertion mutant, p*PvuR*1.2::IS*1*Ω17, from which some of the sequence was derived by using IS*1*-specific primers. The boundaries of *pvuIIR* are indicated. The numbering begins beyond the 3′ end of *pvuIIR*, so as to be codirectional with the divergently oriented and previously published *pvuIIM* sequence (21). (B) Sequence of the *pvuIIC* region. The entire 2,973-bp sequence of this RMS2 has been deposited with GenBank (see text). Only the *pvuIIC* region is shown here. Translations of the *pvuIIR*, *pvuIIM*, and *pvuIIC* ORFs are shown above the DNA sequence. The first 11 amino acids specified by *pvuIIR* (boxed) have been verified by N-terminal amino acid sequencing (3; the Met is removed in vivo). An uncorroborated version of the sequence was published recently, from base 24 to the *Spe*I site in *pvuIIM* at position 1,345 (1). The potential promoters were sought by comparison with the *E. coli* σ^{70} consensus (15). An inverted repeat in the *pvuIIR* candidate promoter is indicated (half arrows).

scanned and digitized as previously described (22). A band at the expected position for C · *Pvu*II peaked after 3 h at 10.8% of the total lane autoradiographic density and was still 7.4% of the total at 22 h. This band was undetectable in an uninduced culture labeled for the same times. These results suggest that the putative *pvuIIC* product is relatively stable in vivo, at least under the conditions we used.

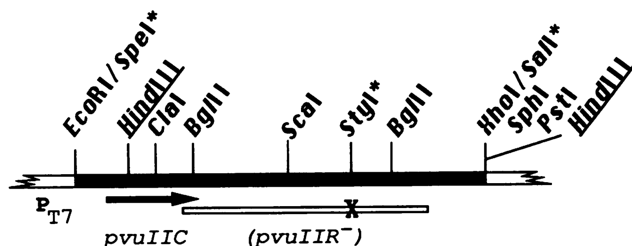


FIG. 2. Construction of *pvuIIC*-overproducing plasmid p*PvuC*1.2/T7pBS. An *Eco*RV-*Xho*I fragment from plasmid p*PvuRM*3.4 (4) was ligated into *Sal*I-*Sma*I-digested pBS(−) (Stratagene, La Jolla, Calif.). The *pvuIIR* gene was inactivated by opening, filling in, and religating a unique *Sty*I site. The *pvuIIM* gene was then deleted by *Eco*RI and *Spe*I double digestion, filling in, and religation. This brought *pvuIIC* closer to and put it under control of the T7 promoter. Plasmid p*PvuC*ΔH3/T7pBS, a derivative of p*PvuC*1.2/T7pBS, was made by deletion between the two *Hin*dIII sites (underlined).

In vitro expression of *pvuIIC*. We sought to express *pvuIIC* in vitro, both to confirm the in vivo result and to test the native expression sequences. As can be seen by comparing Fig. 1B and 2, p*PvuC*1.2/T7pBS retains the putative native expression sequences downstream of the T7 promoter. An *E. coli* S30 fraction (Promega, Madison, Wis.) was used with 25 μCi of a ^{35}S-labeled Met-Cys mixture. The Promega protocol was followed, except for two reactions which were supplemented with 10 U of T7 RNA polymerase (GIBCO-BRL). Equal amounts of labeled protein from each reaction were resolved on an SDS–20% polyacrylamide gel as described above (11), and the autoradiograph is shown in Fig. 3, lanes F to N. The expected protein was made, but it was detectable only when expressed from the T7 promoter (lane M). To confirm that the labeled band was coded for by the *pvuIIC* gene, we used p*PvuC*ΔH3/T7pBS, a derivative of p*PvuC*1.2/T7pBS deleted between the two *Hin*dIII sites (Fig. 2) which should yield a 3.9-kDa polypeptide. When this construct was used, we saw a major protein band of the expected mass (3.5 ± 0.2 kDa; Fig. 3, lane N).

Downloaded from http://jb.asm.org/ on November 22, 2019 by guest

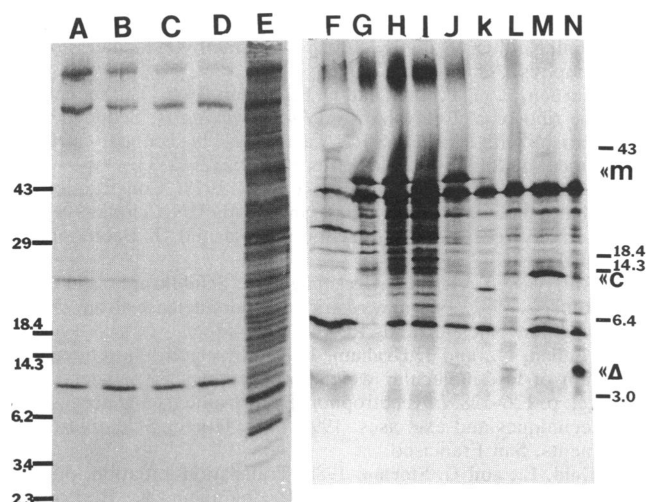
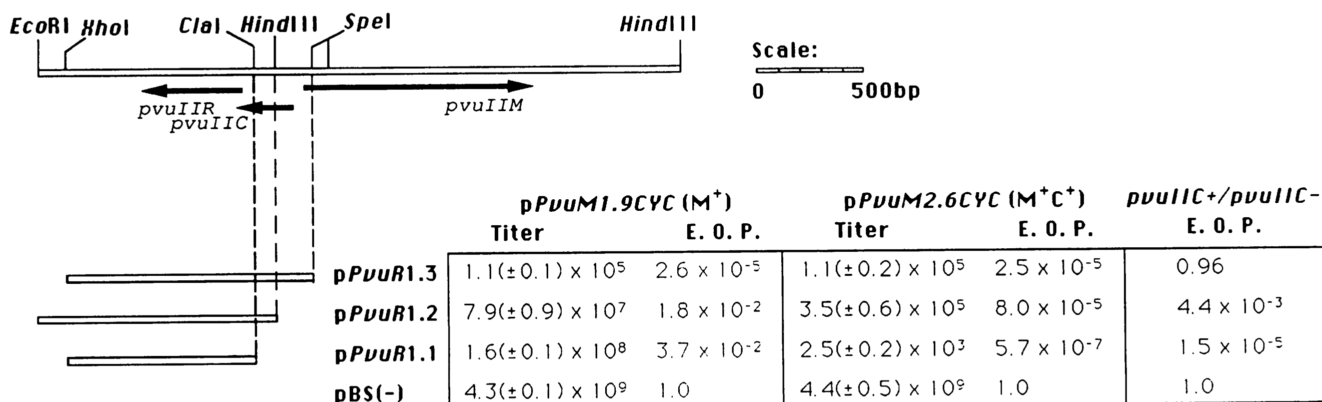


FIG. 3. Expression of *pvuIIC* in vivo and in vitro. Lanes A to E are from in vivo expression assays. Strain JM107MA2 carrying both p*PvuC*1.2/T7pBS and pGP1-2 was grown at 30°C in M9 glucose minimal medium containing 120 μg of carbenicillin per ml and 15 μg of kanamycin per ml (16). When the culture reached a Klett density of 94, it was shifted to 42°C. Rifampin was added to a concentration of 10 μg/ml 20 min after induction, and 5-ml aliquots were removed at 20 (A), 50 (B), 80 (C), and 110 (D) min after the temperature shift. These aliquots were pulse-labeled for 5 min with 60 μCi of a ^{35}S-labeled Met-Cys mixture (Dupont, Boston, Mass.) and then chased with unlabeled Met at a final concentration of 30 μg/ml for 5 min (C · *Pvu*II contains no Cys). The cells were chilled on ice, collected by centrifugation, suspended in 200 μl of H_2O, and sonicated for 90 s at 150 W. The samples were boiled for 2 min in sample buffer (final concentrations: 80 mM Tris · HCl [pH 6.8], 80 mM dithiothreitol, 2% [wt/vol] SDS, 10% [vol/vol] glycerol, 0.02% bromphenol blue), and for the induced culture equal amounts of labeled extract (trichloroacetic acid-precipitable counts) were resolved on an SDS–20% polyacrylamide gel modified to resolve small proteins (11). Lane E shows the labeling of an aliquot from the parallel, unshifted control culture. The bars at the left indicate migration of the molecular mass standards (kilodaltons). Lanes F to N are from in vitro expression assays. A 3-μg sample of plasmid DNA was used as the template for transcription and translation as specified in the text. The plasmids used were pBR322 (F), p*PvuM*1.9 (G), p*PvuM*1.9p_{tac}2 (H), p*PvuM*Δ*Spe*p_{tac}2 (I), p*PvuRM*3.4 (J), pBS(−) (K), p*PvuC*1.2/T7pBS (L and M), and p*PvuC*ΔH3/T7pBS (N). For lanes M and N, the reactions were supplemented with 10 U of T7 RNA polymerase. Equal amounts of incorporated label were resolved on an SDS–20% polyacrylamide gel. Double arrows indicate the positions of M · *Pvu*II (m), C · *Pvu*II (c), and C · *Pvu*IIΔH3 (Δ), and the bars indicate molecular mass markers (kilodaltons).

This confirms the in vivo finding that *pvuIIC* specifies a protein but suggests that the native expression sequences are weak, at least in *E. coli*. We searched upstream of *pvuIIC* for potential promoters and calculated their strength by the method of Mulligan et al. (15). The candidate indicated in Fig. 1B has a remarkably high score, comparable to that for the λ p_R promoter (15). However, the translational initiator of *pvuIIC* has only limited similarity to either the *E. coli* or the *P. vulgaris* consensus (8, 12).

Where does C · *Pvu*II protein act? Our previous work suggested that C · *Pvu*II acts on *pvuIIR* expression either just before or within *pvuIIC* itself (20). We sought to refine that work. Specifically, a plasmid containing *pvuIIR* with various lengths of upstream DNA was put into cells together with a second, compatible plasmid. As a positive control, the largest *pvuIIR* plasmid (p*PvuR*1.3) contained *pvuIIC*$^+$, and expression of *pvuIIR* is not dependent on the second plasmid in this case. As a negative control, vector pBS(−) was used. The second plasmid contained either *pvuIIM*$^+$*C*$^+$, to provide active methyltransferase and C · *Pvu*II under all conditions, or *pvuIIM*$^+$*C*, to ensure that the effects seen were due to the C · *Pvu*II being supplied in *trans*. For each pair of plasmids, expression of *pvuIIR* was measured by in vivo restriction of bacteriophage λ *vir*. The plasmid constructs and the results obtained are shown in Fig. 4. The *pvuIIC*-dependent stimulation of *pvuIIR* expression occurred even when all of the DNA 5′ to the *Cla*I site was deleted (p*PvuR*1.1). The *Cla*I



	pPvuM1.9CYC (M$^+$)		pPvuM2.6CYC (M$^+$C$^+$)		pvuIIC+/pvuIIC−
	Titer	E. O. P.	Titer	E. O. P.	E. O. P.
pPvuR1.3	1.1(±0.1) x 10^5	2.6 x 10^{-5}	1.1(±0.2) x 10^5	2.5 x 10^{-5}	0.96
pPvuR1.2	7.9(±0.9) x 10^7	1.8 x 10^{-2}	3.5(±0.6) x 10^5	8.0 x 10^{-5}	4.4 x 10^{-3}
pPvuR1.1	1.6(±0.1) x 10^8	3.7 x 10^{-2}	2.5(±0.2) x 10^3	5.7 x 10^{-7}	1.5 x 10^{-5}
pBS(-)	4.3(±0.1) x 10^9	1.0	4.4(±0.5) x 10^9	1.0	1.0

FIG. 4. Mapping of the C · *Pvu*II site of action. A series of deletion constructs was made to map the site of action of C · *Pvu*II. In all of these constructs, the coding region of *pvuIIR* was intact; only the length of the sequence upstream of it was altered. These plasmids, shown to the left of the table, were introduced into strain JM107MA2 together with p*ACYC*184 (9) construct p*PvuM*1.9CYC or p*PvuMR*$^-$2.6CYC, and transformants were used for λ phage restriction assays. A λ *vir* phage stock was serially diluted in SM (0.1 M NaCl, 10 mM $MgSO_4$, 50 mM Tris · HCl [pH 7.5]) and plated in triplicate onto JM107MA2 carrying the different plasmid combinations. The final titers (in PFU per milliliter) are shown as means, with standard errors in parentheses, and the efficiencies of plating (E. O. P.) are normalized to that of nonrestricting control pBS. The *pvuIIC*-dependent effects on restriction were assessed by dividing the p*PvuMR*$^-$2.6CYC (C$^+$) efficiency of plating by the p*PvuM*1.9CYC (C$^-$) efficiency of plating for each row.

Downloaded from http://jb.asm.org/ on November 22, 2019 by guest

site is only 70 bp upstream of the *pvuIIR* ORF, and this region includes a potential promoter (Fig. 1B). As the *pvuIIC* ORF runs six codons into the *pvuIIR* ORF, this entire 70-bp region lies within *pvuIIC*. Our working hypothesis is that C · *Pvu*II acts as a transcriptional activator just upstream of the *pvuIIR* ORF, although it is still possible that C · *Pvu*II acts within *pvuIIR*, for example, as an antiterminator.

Genes similar to *pvuIIC* appear to be widespread (20, 24). The C genes could be recruited more easily into new RMS2s if they did not have to change their specificity to match each new RMS2, so it is noteworthy that the *pvuIIMCR* genes lack any occurrences of 5'-CAGCTG-3' and thus C · *Pvu*II must not bind the same sequence as M · *Pvu*II and R · *Pvu*II. The C genes would also be particularly mobile if, as *pvuIIC* may, they carry a C-controlled promoter within their coding sequences. This property suggests that genes such as *pvuIIC* could be regulatory cassettes, capable of positively controlling any gene to which they insert immediately upstream.

Nucleotide sequence accession numbers. The nucleotide sequence for the entire *Pvu*II restriction-modification system, including the sequence in Fig. 1, has been assigned GenBank accession no. M77223. The sequences reported in references 1 and 21 have been assigned EMBL accession no. X52681 and GenBank accession no. X13778, respectively.

We thank A. S. Bhagwat and J. C. Dunbar of Wayne State University and P. F. Lehmann of the Medical College of Ohio for suggestions and comments on the manuscript. We also thank J. S. Benner of New England BioLabs, Inc., for sharing unpublished data and C. Vincenz of the University of Michigan for help with image analysis.

This report is based on work supported by the National Science Foundation under grant DMB-8818673.

REFERENCES

1. **Anthanasiadis, A., M. Gregoriu, D. Thamos, M. Kokkinidis, and J. Papamatheakis.** 1990. Complete nucleotide sequence of the *Pvu*II restriction enzyme gene from *Proteus vulgaris*. Nucleic Acids Res. **18:**6434.
2. **Bartlett, J. A., R. K. Gaillard, Jr., and W. K. Joklik.** 1986. Sequencing of supercoiled plasmid DNA. BioTechniques **4:**208–210.
3. **Benner, J. S., L. Dorner, L. Greenough, D. Heiter, L. Hornstra, and W. Jack. (New England BioLabs, Inc.).** 1990. Unpublished data.
4. **Blumenthal, R. M., S. A. Gregory, and J. S. Cooperider.** 1985. Cloning of a restriction-modification system from *Proteus vulgaris* and its use in analyzing a methylase-sensitive phenotype in *Escherichia coli*. J. Bacteriol. **164:**501–509.
5. **Bougueleret, L., M. Schwarzstein, A. Tsugita, and M. Zabeau.** 1984. Characterization of the genes coding for the *Eco*RV restriction and modification system of *Escherichia coli*. Nucleic Acids Res. **12:**3659–3675.
6. **Brooks, J. E., P. D. Nathan, D. Landry, L. A. Sznyter, P. Waite-Rees, C. L. Ives, L. S. Moran, B. E. Slatko, and J. S. Benner.** 1991. Characterization of the cloned *Bam*HI restriction modification system: its nucleotide sequence, properties of the MTase, and expression in heterologous hosts. Nucleic Acids Res. **19:**841–850.
7. **Butkus, V., S. Klimasauskas, L. Petrauskiene, Z. Maneliene, A. Lebionka, and A. A. Janulaitis.** 1987. Interaction of *Alu*I, *Cfr*6I and *Pvu*II restriction-modification enzymes with substrates containing either N^4-methylcytosine or 5-methylcytosine. Biochim. Biophys. Acta **909:**201–207.
8. **Carbon, P., J. P. Ebel, and C. Ehresman.** 1981. The sequence of the ribosomal 16S RNA from *Proteus vulgaris*. Sequence comparison with *E. coli* 16S RNA and its use in secondary structure model building. Nucleic Acids Res. **9:**2325–2333.
9. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:**1141–1156.
10. **Gingeras, T. R., L. Greenough, I. Schildkraut, and R. J. Roberts.** 1981. Two new restriction endonucleases from *Proteus vulgaris*. Nucleic Acids Res. **9:**4525–4536.
11. **Giulian, G., and J. Graham.** 1990. The electrophoretic separation of low molecular weight polypeptides in polyacrylamide gel, p. 126–128. *In* Electrophoresis instruments and accessories, techniques and exercises. 1990–1991. Hoefer Scientific Instruments, San Francisco.
12. **Gold, L., and G. Stormo.** 1988. Translation initiation, p. 1302–1307. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
13. **Greene, P. J., M. Gupta, H. W. Boyer, W. E. Brown, and J. M. Rosenberg.** 1981. Sequence analysis of the DNA encoding the *Eco*RI endonuclease and methylase. J. Biol. Chem. **256:**2143–2153.
14. **Modrich, P., and R. J. Roberts.** 1982. Type-II restriction and modification enzymes, p. 109–154. *In* S. M. Linn and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1983. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12:**789–800.
16. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., p. A.1–A.6. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory, N.Y.
17. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
18. **Sullivan, K. M., and J. R. Saunders.** 1989. Nucleotide sequence and genetic organization of the *Ngo*PII restriction-modification system of *Neisseria gonorrhoeae*. Mol. Gen. Genet. **216:**380–387.
19. **Tabor, S., and C. C. Richardson.** 1985. A bacterial phage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
20. **Tao, T., J. C. Bourne, and R. M. Blumenthal.** 1991. A family of regulatory genes associated with type II restriction-modification systems. J. Bacteriol. **173:**1367–1375.
21. **Tao, T., J. Walter, K. Brennan, M. Cotterman, and R. M. Blumenthal.** 1989. Sequence, internal homology and high-level expression of gene for a DNA-(cytosine-*N4*)-methyltransferase, M · *Pvu*II. Nucleic Acids Res. **17:**4161–4175.
22. **Vincenz, C., J. Fronk, G. A. Tank, and J. P. Langmore.** 1991. Nucleoprotein hybridization: a method for isolating active and inactive genes as chromatin. Nucleic Acids Res. **19:**1325–1336.
23. **Wilson, G. G.** 1991. Organization of restriction-modification systems. Nucleic Acids Res. **19:**2539–2566.
24. **Wilson, G. G., and N. E. Murray.** 1991. Restriction and modification systems. Annu. Rev. Genet. **25:**585–627.

Downloaded from http://jb.asm.org/ on November 22, 2019 by guest