

Role and Mechanism of Action of C · PvuII, a Regulatory Protein Conserved among Restriction-Modification Systems

ROY M. VIJESURIER,^{1†} LEON CARLOCK,¹ ROBERT M. BLUMENTHAL,^{2*} AND JOAN C. DUNBAR¹

Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, Michigan 48201,¹
and Department of Microbiology and Immunology, Medical College of Ohio, Toledo, Ohio 43614-5806²

Received 29 April 1999/Accepted 27 October 1999

The PvuII restriction-modification system is a type II system, which means that its restriction endonuclease and modification methyltransferase are independently active proteins. The PvuII system is carried on a plasmid, and its movement into a new host cell is expected to be followed initially by expression of the methyltransferase gene alone so that the new host's DNA is protected before endonuclease activity appears. Previous studies have identified a regulatory gene (*pvuIIC*) between the divergently oriented genes for the restriction endonuclease (*pvuIIR*) and modification methyltransferase (*pvuIIM*), with *pvuIIC* in the same orientation as and partially overlapping *pvuIIR*. The product of *pvuIIC*, C · PvuII, was found to act in *trans* and to be required for expression of *pvuIIR*. In this study we demonstrate that premature expression of *pvuIIC* prevents establishment of the PvuII genes, consistent with the model that requiring C · PvuII for *pvuIIR* expression provides a timing delay essential for protection of the new host's DNA. We find that the opposing *pvuIIC* and *pvuIIM* transcripts overlap by over 60 nucleotides at their 5' ends, raising the possibility that their hybridization might play a regulatory role. We furthermore characterize the action of C · PvuII, demonstrating that it is a sequence-specific DNA-binding protein that binds to the *pvuIIC* promoter and stimulates transcription of both *pvuIIC* and *pvuIIR* into a polycistronic mRNA. The apparent location of C · PvuII binding, overlapping the –10 promoter hexamer and the *pvuIICR* transcriptional starting points, is highly unusual for transcriptional activators.

The bacterial type II restriction-modification systems include a DNA modification methyltransferase (MTase) and a restriction endonuclease (REase), both of which act independently on the same DNA sequence (65). The REase cleaves duplex DNA sequences in the absence of sequence-specific DNA modification by the MTase. These systems can defend bacterial cells against viral infection, although other functional roles have also been proposed (45). Restriction-modification systems have provided an important focus for studies of molecular recognition. Biochemical and crystallographic analyses are yielding significant insights into the mechanisms of sequence recognition and catalytic activity of these proteins (3, 18, 50, 66).

It is evident from their opposing roles that very careful control of the relative activities of the MTase and REase is critically important: too low a MTase/REase activity ratio would lead to cell death via autorestriction (22), while too high a ratio would fail to provide protection from invading viral DNA. Furthermore, many restriction-modification systems are carried on plasmids, and following transfer to a new host cell there must be a period during which MTase activity is present and REase activity is not, so as to protect the new host's DNA before endonuclease activity appears. Accordingly, restriction-modification systems must be temporally regulated while they are establishing themselves in new host cells.

The PvuII restriction-modification system is a type II system isolated from the gram-negative bacterium *Proteus vulgaris* (24). It was found to be carried by a small plasmid (11, 16), its

genes were cloned, and their nucleotide sequences were determined (7, 58, 60). The PvuII REase (R · PvuII) recognizes the sequence CAGCTG and cleaves the central GpC on both strands to yield blunt ends (24); this enzyme has been crystallographically characterized as an apoenzyme and in complex with its DNA substrate (8, 19, 30). The MTase (M · PvuII) recognizes the same CAGCTG sequence and modifies the internal cytosine (11), generating N4-methylcytosine (15). The MTase has also been characterized crystallographically (25).

The mechanisms underlying regulated expression of these enzymes still remain to be defined. A subset of restriction-modification systems produce a protein that has been shown to play an important role in regulation, though via unknown mechanisms. This protein has been named C (for controller), and its gene generally precedes and in some cases partially overlaps the REase gene. C protein was originally discovered in the BamHI (13) and PvuII (59) systems, and homologs were identified at that time in several other systems (59). These C proteins have not yet been structurally characterized, but their amino acid sequences reveal that they are very probably helix-turn-helix proteins similar to known activators and repressors of gene expression (67). The C proteins act in *trans* and are required for expression of the REase gene. Furthermore, there is some cross-complementation between the C genes from different restriction-modification systems (31, 36). New members of this family continue to be identified (5).

Sequence comparisons have identified a conserved DNA sequence element termed a "C box" immediately upstream of most C genes (48). The C box has been suggested, though not proven, to be a site of action for the C gene product. In this study we have directly investigated the effects of C · PvuII on transcription of the PvuII genes, and we demonstrate that it is a sequence-specific DNA binding protein that binds to the C box and stimulates transcription; surprisingly, we find that the C box overlaps the –10 promoter hexamer and transcription

* Corresponding author. Mailing address: Department of Microbiology and Immunology, Medical College of Ohio, Toledo, OH 43614-5806. Phone: (419) 383-5422. Fax: (419) 383-3002. E-mail: rblumenthal@mco.edu.

† Present address: Department of Pediatrics, Children's Hospital of Michigan, Detroit, MI 48201.

starting points for the activated promoter. We have also investigated the role of *C · PvuII* in the temporal regulation of the restriction-modification system and subsequent host cell survival.

MATERIALS AND METHODS

Bacterial strains and plasmids. The *P. vulgaris* strain used for RNA isolation was originally obtained from the American Type Culture Collection (ATCC 13315) and is the same strain from which the *PvuII* restriction-modification system was originally isolated (24). *Escherichia coli* HB101 was routinely used for cloning experiments; this strain is MerBC⁻ and thus permissive for the *PvuII* MTase (11, 46). Competent HB101 cells were obtained from Life Technologies.

pPvuRM3.4CYC, which contains the genes for the entire *PvuII* restriction-modification system, was generated from pPvuRM3.4 (11) by excision of the *EcoRV-EcoRI* fragment and subcloning into the *EcoRI* and *ScaI* sites of pACYC184 (17). Inserting this fragment inactivated the vector chloramphenicol acetyltransferase gene (*cat*) but left the tetracycline resistance gene intact. The pKK232-8 plasmid (14), which contains a promoterless *cat* gene, was obtained from Pharmacia Biotech. The pFLAG.2 expression vector was obtained from Kodak (but is currently distributed by Sigma).

Construction of *C · PvuII*^{FLAG} fusion proteins. The *pvuII*C gene was amplified as a 310-bp fragment from plasmid pPvuRM3.4 by using gene-specific primers (Macromolecular Structure Facility, Michigan State University). The forward primer (5'-CAT CAT TAT CAG ATC TAT GAG CAG AA) contains a 3-nucleotide (nt) mismatch that generated a *BglII* site (underlined) immediately upstream of the *pvuII*C initiation codon. The reverse primer (5'-GTC TTG ATA TTC CTG TAT) corresponds to DNA downstream of the 3' end of *pvuII*C where a native *BglII* site occurs. Template DNA was amplified with *Taq* DNA polymerase (Life Technologies), and the cycling parameters used were 94°C for 2 min; 35 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 3 min; and finally 72°C for 7 min. The PCR product was digested with *BglII* and fractionated on a 1.5% agarose gel, and the 248-bp product was purified by using the Wizard PCR purification kit (Promega). This fragment was subsequently cloned into the *BglII* site of the pFLAG.2 expression vector, and the resulting ligation mixture was used to transform competent HB101 cells. The pFLAG.2 vector adds an 8-amino-acid (aa) FLAG epitope tag (Asp Tyr Lys Asp Asp Asp Lys) to the amino-terminal end of the cloned protein, allowing immunoaffinity purification (12, 28). Transformants were selected by ampicillin resistance, and positive clones were detected by filter hybridization using the random primer-labeled PCR product as the probe. The DNA sequences of positive clones were determined in order to verify the orientation and fidelity of the inserted DNA.

Expression and extraction of *C · PvuII*^{FLAG} fusion proteins. To detect the active (*C · PvuII*^{FLAG}) and inactive (*C^{Leu} · PvuII*^{FLAG}) fusion proteins analytically, cells were grown at 37°C in Luria-Bertani (LB) medium to mid-log phase and isopropyl-β-D-thiogalactopyranoside (IPTG; Life Technologies) was added to a final concentration of 1 mM to induce expression of *C · PvuII*^{FLAG}. The cells were incubated a further 2 h and then sedimented by centrifugation at 5,000 × *g* for 10 min. Pellets were resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer and analyzed on SDS-17.5% polyacrylamide gels. After electrophoresis the samples were electroeluted (Bio-Rad Transblot electroeluter) onto nitrocellulose membranes for 1 h at 90 V. The primary and secondary antibodies for immunodetection were the mouse anti-FLAG.2 monoclonal antibody (Kodak) and a sheep antimouse antibody linked to horseradish peroxidase (Amersham). The immune complexes were visualized by using the ECL chemiluminescence detection kit (Amersham).

For larger preparations of *C · PvuII*^{FLAG}, 100-ml cultures were grown and induced as described in the preceding paragraph. The cells were sedimented and then resuspended in 7 ml of 20 mM HEPES (pH 7.5)-50 mM KCl-2 mM EDTA-0.2 mM dithiothreitol-5% (vol/vol) glycerol-1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were disrupted by sonication for 2 min in six 20-s pulses by using a Branson sonicator, and the resulting lysate was clarified by centrifugation at 20,000 × *g* for 1 h at 4°C. The supernatant was removed and assayed for the fusion protein by Western blot analyses as described above and for total protein concentration by the Bio-Rad protein assay.

Transformation assays. In cotransformation experiments, competent *E. coli* HB101 cells were simultaneously transformed by the pFLAG.2-*pvuII*C and pPvuRM3.4CYC plasmids or their derivatives. Fifty femtomoles of each plasmid was added to the cells, and the mixture was incubated at 42°C for 2 min. Following this heat shock, the cells were diluted into 5 ml of LB medium and incubated at 37°C for 1 h. The cells were pelleted, serially diluted, and plated on LB agar containing both ampicillin (25 µg/ml; selects for pFLAG plasmids) and tetracycline (25 µg/ml; selects for pACYC184 plasmids). The number of transformants was determined after incubation of the plates at 37°C overnight.

In sequential transformation assays, the pFLAG.2-*pvuII*C or pPvuRM3.4CYC plasmid was separately established in HB101 cells. These cells were rendered competent by treatment with calcium chloride and were then transformed with the alternate plasmids. The transformants were selected on double antibiotic plates as described above.

Gel mobility shift analysis. DNA binding assays were carried out by using duplex, synthetic oligonucleotide substrates. The complementary oligonucleo-

tides were annealed, 5'-end-labeled with T4 polynucleotide kinase (New England Biolabs), and purified as previously described (68). Experiments used crude extracts containing equivalent concentrations of *C · PvuII*^{FLAG} fusion proteins, as determined by Western blot analyses. The binding buffer contained 20 mM HEPES (pH 7.0)-50 mM potassium glutamate-0.5 mM EDTA-0.1 mM dithiothreitol; in some cases it also contained as much as 300 ng of poly(dI-dC). Protein samples were incubated with 50 nM DNA substrate for 1 h at 25°C in a total volume of 20 µl. Immediately after the addition of 2 µl of a loading solution (0.05% bromophenol blue in 10× binding buffer containing 20% glycerol), the samples were electrophoresed on 1-mm-thick 10% polyacrylamide gels (acrylamide to bisacrylamide, 29:1) in 20 mM HEPES (pH 7.4)-2 mM EDTA at a constant voltage of 10 V/cm. The protein-DNA complexes were detected by autoradiography of the dried gel.

Primer extension. To obtain total-cell RNA, 100-ml bacterial cultures were grown in LB medium to mid-log phase (*A*₅₅₀ = 0.5) and the cells were harvested by centrifugation at 5,000 × *g* for 15 min. RNA was extracted by using guanidine thiocyanate (20). All solutions were treated with diethylpyrocarbonate (Sigma) before use. The precipitated RNA was resuspended in H₂O, and the concentration of RNA was estimated from the absorbance at 260 nm. The integrity of the RNA sample was determined by electrophoresis for ≤3 h in agarose gels containing 1% formaldehyde. For the primer extension assays, 1 pmol of a ³²P-end-labeled oligonucleotide was initially annealed to 5 µg of RNA template at 70°C for 10 min. The samples were then cooled to 42°C for 2 min, after which 1 µl of SUPERScript RNase H⁻ reverse transcriptase (Life Technologies) was added. The extension reaction was allowed to proceed at 42°C for 30 min and was then terminated by addition of 10 µl of a solution containing 10 mM NaOH, 95% formamide, 0.05% bromophenol blue, and 0.05% xylene cyanol. The samples were boiled for 2 min before being loaded onto a denaturing, 8% polyacrylamide sequencing gel. Standards used on the sequencing gels were either size standards (*HaeIII*-digested φX174 DNA [New England Biolabs] that were 5' end labeled with ³²P) or sequencing reactions using the corresponding primers and the cloned genes as a template. Sequencing reactions were carried out according to the manufacturer's instructions by using the *f*mol Sequencing Kit from Promega. The oligonucleotides used were anti-R (5'-GTCCTGATATTCCTGTAT), anti-C1 (5'-TCGGGCTGATAAAGGATTT), anti-C2 (5'-GGGTCTATGTATAT AGGT), and anti-M (5'-ACTCATAGTCTGTAGATT).

Construction of promoter clones. Fragments to be assayed for promoter activity were PCR amplified with selected oligonucleotide primers and with plasmid pPvuRM3.4 as the template. The PCR products were purified from agarose gels by using a Wizard PCR purification kit (Promega), any nonblunt ends were filled in with the Klenow fragment of DNA polymerase (New England Biolabs), and these DNA products were ligated into the *SmaI*-digested plasmid pKK232-8. The ligation products were used to transform competent *E. coli* HB101 cells, and positive clones were identified by filter hybridization using a ³²P-random primer-labeled PCR product as the probe. Positive clones were sequenced to determine the orientation and fidelity of the insert.

Assay of promoter activity. Fragments inserted into plasmid pKK232-8 were assayed for promoter activity by using the FAST CAT Chloramphenicol Acetyltransferase Assay kit (Molecular Probes). Bacterial cell extracts were prepared and enzyme assays were carried out according to the manufacturer's instructions. The fluorescent substrate and reaction products were spotted onto Whatman TLC silica gel thin-layer chromatography plates and resolved with a chloroform-methanol (9:1 [vol/vol]) solvent mixture. The plates were illuminated with UV light and photographed with Polaroid T 55 film, and the photographic negative was analyzed on a Molecular Dynamics densitometer by using the associated image analysis software. The initial rate of the reaction was determined for each of the samples by serial dilution of the extracts, using incubation times yielding less than 50% substrate conversion.

The initial rates were also normalized to the concentration of β-lactamase in each of the extracts, in order to correct for possible variation in plasmid copy number, by using an assay adapted from ones described previously (32, 38). Crude extracts were prepared as described for the *C · PvuII*^{FLAG} fusion proteins. β-Lactamase activity was determined spectrophotometrically by using nitrocefin (Becton-Dickinson Microbiology Systems), a chromogenic cephalosporin with an absorption maximum at 482 nm following hydrolysis. The reactions were carried out at 37°C by incubating cell extracts (diluted 1:25 in Tris · HCl, pH 8.0) together with 0.1 mM nitrocefin and 0.1 M phosphate-buffered saline (80 mM Na₂HPO₄, 20 mM NaH₂PO₄, and 100 mM NaCl) in a total volume of 1 ml. Absorbance was monitored for 5 min after the addition of cell extracts. Reaction rates were calculated by linear regression of a plot of *A*₄₈₂ versus time, which was found to be linear for as long as 120 s.

RESULTS

***C · PvuII* mediates temporal control of the *PvuII* genes.** After the *PvuII* restriction-modification system moves into a new host cell, *pvuIIM* expression is expected to occur prior to significant expression of *pvuIIR* to avoid autorestriction. Our working model for this temporal control has been that *C · PvuII* accumulates and, after a significant amount of time has

TABLE 1. Effects of *pvuIC* on transformation by the *PvuII* restriction-modification system

Resident plasmid ^c	Successive transformations ^a			Plasmid 1	Cotransformations ^b		
	Transforming plasmid ^d				Plasmid 2 ^d	R ⁻ C ⁻ M ⁺	R ⁺ C ⁺ M ⁺
	R ⁺ C ⁺ M ⁺	C ⁻	C ⁺				
R ⁺ C ⁺ M ⁺	ND	0.80 ± 0.07	0.76 ± 0.06	C ⁻	1.53 ± 0.11	1.07 ± 0.22	
C ⁻	0.85 ± 0.08	ND	ND	C ⁺	1.09 ± 0.09	<0.05	
C ⁺	0.02 ± 0.02	ND	ND				

^a Cells containing the resident plasmid were made competent and transformed by the second plasmid. The two plasmids are, in each case, compatible and specify distinct antibiotic resistances.

^b Competent cells containing no plasmids were simultaneously transformed by two compatible plasmids specifying distinct antibiotic resistances.

^c The *PvuII* phenotypes of the plasmids used in each case are shown; R refers to the restriction endonuclease, M to the DNA methyltransferase, and C to the regulatory protein C · *PvuII*. The full designations of the plasmids used are pPvuRM3.4CYC (R⁺ C⁺ M⁺), pPvuM1.9CYC (R⁻ C⁻ M⁺), pFLAG.2-*pvuIC* Esp19 (C⁻), and pFLAG.2-*pvuIC* (C⁺).

^d Each result is the number of doubly resistant transformants relative to the number of doubly resistant transformants when the incoming plasmid (successive transformations) or plasmid 2 (cotransformations) is the vector control plasmid. ND, not determined. Average of triplicates ± standard error.

elapsed, reaches a level that permits it to activate the transcription of *pvuIIR* (59). A consequent prediction of this model is that the simultaneous introduction of *pvuIIM* and *pvuIIR* into a cell that is already expressing *pvuIC* should be lethal, due to premature C · *PvuII*-activated expression of *pvuIIR*, while over-expression of C · *PvuII* should be tolerated if it occurs after the restriction-modification system has become established in a host cell.

To test this model, we cloned *pvuIC* separately from the other *PvuII* genes and either pre-expressed *pvuIC* before introducing the intact *PvuII* restriction-modification system or cotransformed the intact system with a *pvuIC*-overexpressing plasmid. For these studies *pvuIC* was subcloned into pFLAG.2 and expressed as an epitope-tagged fusion protein under the control of the strong *Ptac* promoter. Preliminary Western blot analysis revealed significant expression of *pvuIC* even in the absence of IPTG induction (not shown), so these studies were carried out in the absence of IPTG. As a control, a mutant C · *PvuII*^{FLAG} was generated by making a 3-bp insertion into *pvuIC* as described previously (59); opening a unique *EspI* site, filling in the 5' extensions, and religating yields a functionally inactive protein with an extra Leu in the first helix of the predicted helix-turn-helix motif (C^{Leu} · *PvuII*^{FLAG}). The results are shown in Table 1.

In the cotransformation experiments, *E. coli* HB101 cells were simultaneously transformed with one of the pFLAG.2-*pvuIC* plasmids and a compatible pACYC184-derived plasmid carrying either the intact *PvuII* restriction-modification system (pPvuRM3.4CYC) or a deletion derivative that does not produce REase (pPvuM1.9CYC). We were unable to recover double transformants when active alleles for both *pvuIC* and *pvuIIR* were cotransformed.

In a complementary series of experiments, one of the plasmids (either pFLAG.2-*pvuIC* or pPvuRM3.4CYC) was established in cells prior to transformation by the second plasmid. The transformation efficiency was drastically reduced when active *pvuIC* was expressed prior to transformation with pPvuRM3.4CYC. In contrast, transformation by pFLAG-*pvuIC* of cells already carrying the intact restriction-modification system was as efficient as control transformations even when both alleles were active. These results (61) are consistent with those obtained since by others (36) and indicate that cell viability, for strains producing the *PvuII* restriction-modification system, is not particularly sensitive to elevated levels of C · *PvuII* once the system has become established.

The available data thus support a role for C · *PvuII* as a critical regulator of temporal expression during establishment of the *PvuII* restriction-modification system in a new host cell.

We next turned our attention to the question of how this regulation was achieved.

C · *PvuII* is a sequence-specific DNA-binding protein. The C proteins, including C · *PvuII*, act in *trans* to stimulate the expression of REase genes (5, 13, 31, 36, 59). This stimulation must be strong because, when *pvuIC* is inactivated, *pvuIIR* expression is so low that *pvuIIR*⁺ *pvuIIM* cells are viable (though mutants accumulate) (58). One possible basis for this stimulation is that the C proteins are strong transcriptional activators. This possibility, and the apparent helix-turn-helix motifs implied by their amino acid sequences, led us to test the ability of C · *PvuII* to bind DNA in a sequence-specific manner. In particular, we sought to test binding to the C box, a consensus sequence of unknown function upstream of C genes (including *pvuIC*) (48). Gel mobility shift assays were therefore used in preliminary analyses of DNA binding.

The substrates used for the binding assays were 22-bp duplex oligonucleotides containing the originally proposed *PvuII* C box sequence or a single-base mutant version thereof (Fig. 1A). A protein-DNA complex was evident when the C · *PvuII*^{FLAG} fusion protein was incubated with the native C box duplex, even in the presence of as much as 300 ng of poly(dI-dC) per reaction, but not when the binding assays involved the functionally inactive C^{Leu} · *PvuII*^{FLAG} protein (Fig. 1B). The oligonucleotide substrate bearing a single-base-pair substitution at a highly conserved position of the C box sequence was not bound under the conditions used. These results are consistent with three points. First, they demonstrate that C · *PvuII* is a sequence-specific DNA-binding protein. Second, they indicate that the C box is at least one target of C · *PvuII* binding. Third, they define one of the base pairs that plays a central role in C box recognition by C · *PvuII*. An analysis of the nature of the C box sequence, including previously unremarked symmetrical elements, is presented in the Discussion.

Transcriptional analysis of the *PvuII* restriction-modification system. The next step in testing whether C · *PvuII* is a transcriptional activator, and in understanding how such activation might give the observed pattern of gene expression, was to determine the location and C · *PvuII* responsiveness of the *PvuII* promoters. This involved two experimental approaches. First, primer extension analyses with reverse transcriptase were used to identify the transcriptional start sites for *pvuIIM*, *pvuIC*, and *pvuIIR*. In general the template RNA used in these studies was isolated from *P. vulgaris*, the native host for the *PvuII* system, grown to mid-log phase in a rich medium. Some experiments were also carried out with RNA isolated from an *E. coli* strain that carries a plasmid clone of the *PvuII* genes (pPvuRM3.4).

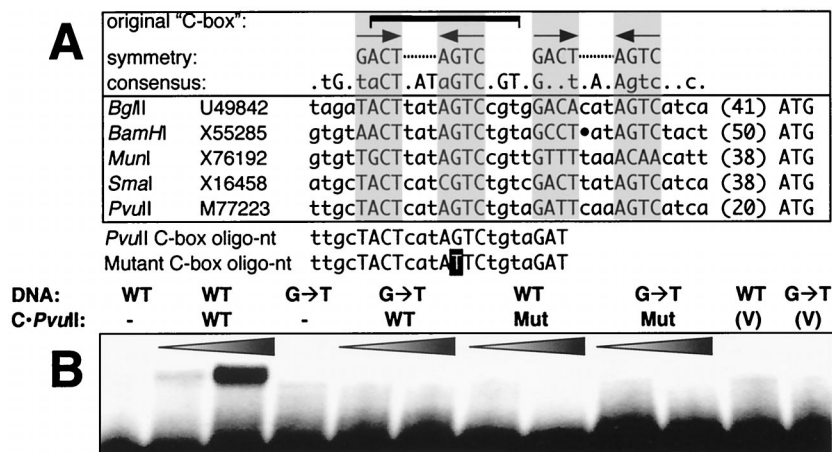


FIG. 1. Analysis of *C·PvuII* interaction with the C box. (A) C box regions from five restriction-modification systems are shown, along with the name of the source system and the GenBank accession number for the sequence. The ATG to the right represents the presumed initiation codon for the respective C genes (in no case has this been confirmed directly), preceded by the distance in nucleotides from the rest of the displayed sequence. Above the sequences are shown the range of the originally identified C box (48), the consensus for these five C boxes (capital letters represent fully conserved positions; lowercase letters indicate conservation in four of the five sequences), and a possible pair of symmetrical sequences identified in this study that might represent the actual binding targets of the homodimeric (and putative helix-turn-helix) C proteins (shaded boxes). Under these native sequences are sequences of two oligonucleotides used for gel mobility shift analyses; the T on a solid background indicates the position of the G→T alteration present in a tested mutant version of the oligonucleotide. (B) Gel mobility shift analysis with oligonucleotides comprising the originally defined C box. This analysis was carried out with a 50 nM concentration of the two (end-labeled, duplex) oligonucleotides whose sequences are given at the bottom of panel A. For the DNA, WT indicates use of the wild-type sequence as shown for *PvuII*, while G→T indicates use of a sequence altered at one conserved position. For *C·PvuII*, cell extracts containing the FLAG fusion proteins were used; WT indicates the wild-type C protein, while Mut refers to an inactive protein from a mutated gene in which an extra Leu codon has been inserted into the putative helix-turn-helix motif (59). V indicates use of control extracts from cells containing the pFLAG vector but no C gene. The wedges indicate increasing amounts of protein added.

In the second group of experiments, candidate segments of *PvuII* DNA were assayed for promoter activity. We cloned putative promoter regions for each of the genes upstream of the promoterless *cat* gene in plasmid pKK232-8, most often in both orientations. The *cat* gene in pKK232-8 is transcriptionally isolated from the rest of the plasmid by strong flanking bidirectional transcription terminators (14). The relative chloramphenicol acetyltransferase activity associated with each plasmid was then determined in the presence of a second, compatible plasmid that carried either the wild-type *pvuII*C gene or, as a control, the inactive mutant *pvuII*C gene bearing an extra Leu codon in the putative helix-turn-helix motif.

(i) **Transcription of *pvuIIM*.** The *PvuII* MTase gene appears to be associated with two promoters that yield two RNA transcripts differing by 39 nt, as reverse transcripts consistently included 82- and 121-nt products (Fig. 2A). As expected, both of the corresponding initiation sites are within the coding sequence of *pvuIIC* (Fig. 2B). Surprisingly, for what appear to be activator-independent promoters, the regions upstream of these transcript starts show only limited similarity to canonical (*E. coli*) promoter sequences (Fig. 3C), although the M121 promoter may have an extended -10 sequence (40). It seems unlikely that an alternative σ factor is involved, since *pvuIIM* must, under most or all growth conditions, be transcribed rapidly upon entry into a new host cell.

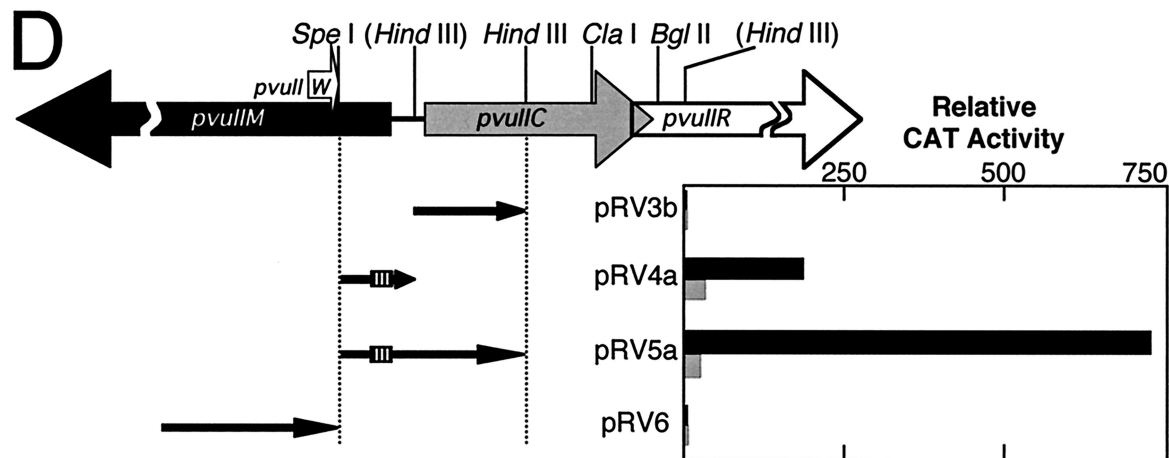
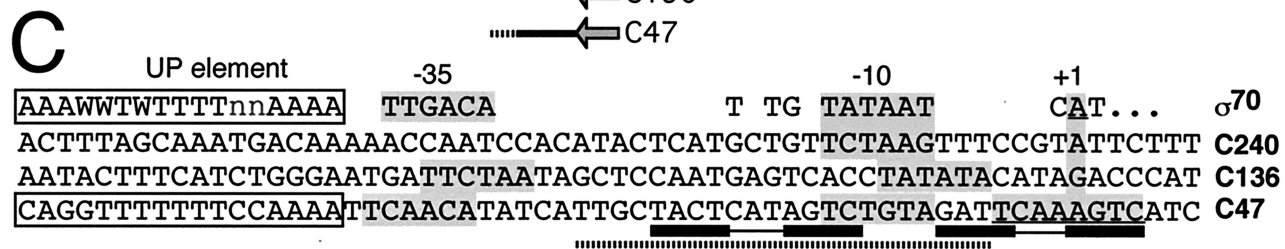
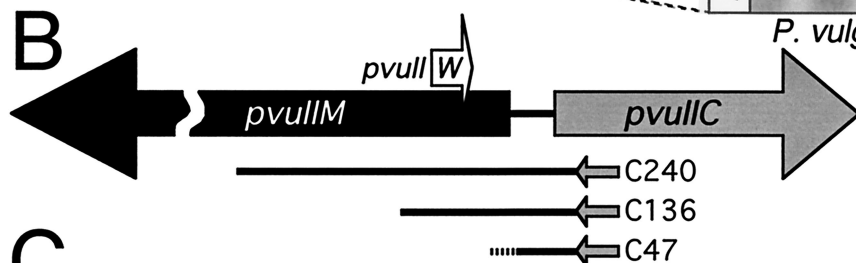
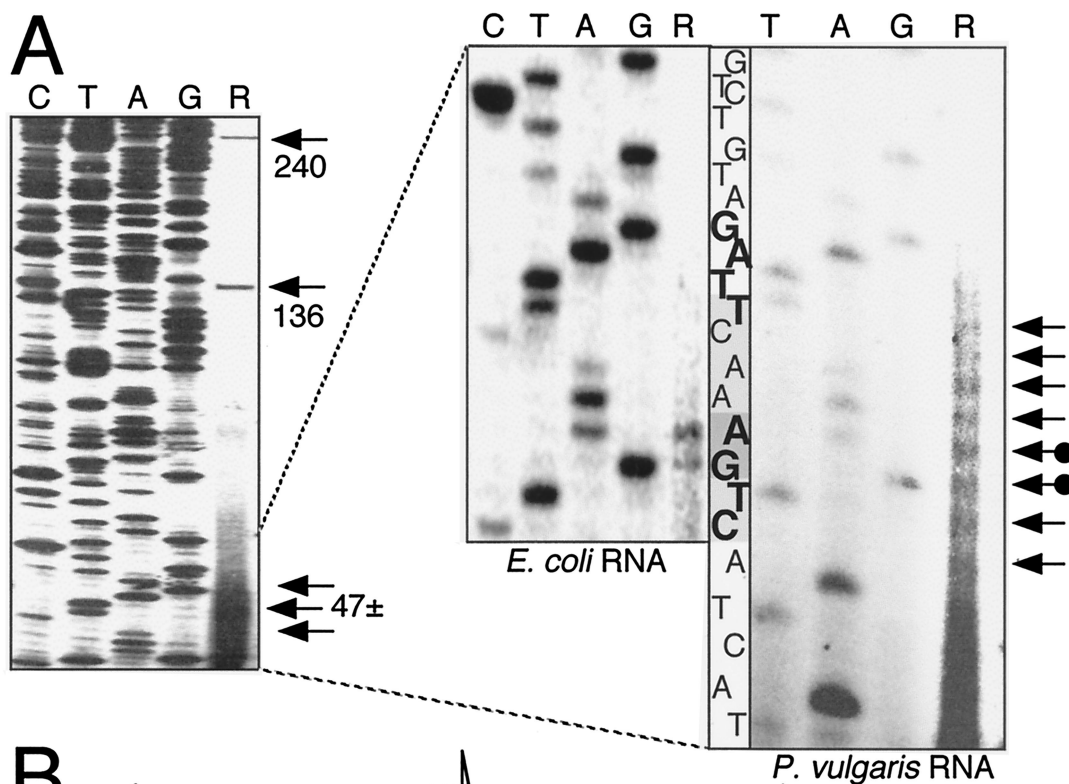
Several subclones were generated that contain potential *pvuIIM* promoters as implicated by the primer extension analysis. Plasmid pRV3a includes DNA up to 60 bp upstream of the M82 transcript starting point. In the absence of active *C·PvuII*, this plasmid yielded the highest CAT activity of all the transformants we analyzed (Fig. 2D); under this condition the pRV3a promoter activity was approximately fivefold greater than that of the putative C gene promoter clones (pRV4a and pRV5a [Fig. 3D]). Interestingly, adding more upstream DNA to the segment in pRV3a (yielding pRV2b) profoundly reduced promoter activity; there was no corresponding reduction

in β -lactamase activity from the vector *bla* gene (not shown), so this seems unlikely to be due to a change in plasmid copy number.

pRV1b also contains a potential *pvuIIM* promoter; this clone includes sequences upstream of the M121 MTase transcript only. pRV1b exhibited weak but significant promoter activity. A deletion between the *Cla*I and *Bgl*II sites that removes the putative RNA hairpin sequences (Fig. 4D) approximately doubled the CAT activity yielded by this plasmid (not shown). It is possible that the inverted repeat sequences are also responsible for the decreased activity of pRV2b relative to pRV3a, as both pRV1b and pRV2b have the same 5' end. This effect is somewhat surprising, as the inverted repeats are 85 bp upstream of the closer of the two putative *pvuIIM* -35 hexamers, and its basis is not yet known.

(ii) **Transcription of *pvuIIC*.** The extension products from primers complementary to *pvuIIC* mRNA indicate a cluster of adjacent transcription starts for *pvuIIC* (C47; Fig. 3A and B). The candidate TATA box (-10 hexamer) and appropriately spaced -35 region only weakly resemble the consensus sequences (Fig. 3C). Immediately upstream of the putative -35 hexamer is a strong match to the consensus for UP elements. An UP element is an α -subunit binding sequence that can increase promoter strength by an order of magnitude (23, 54). The UP consensus is matched at 12 of 15 positions, but in *pvuIIC* it is just 1 nt from the apparent -35 hexamer while the consensus spacing is 4 nt. The role, if any, played by this UP element is not yet clear, though UP elements appear to be present in some activated promoters such as *lacPI* (37).

When primer extension analyses were also carried out with RNA prepared from *E. coli* carrying the intact *PvuII* restriction-modification system, extension products analogous to those described above were similarly observed. However, rather than a cluster of multiple products, only two distinct transcripts were detected (Fig. 3A). This suggests that the clustered starts result from real 5' end variation in the template



(Fig. 4D; hairpin structure on left), and RNA hairpins can cause premature termination by reverse transcriptase. Higher annealing temperatures (up to 75°C) did not eliminate the shorter primer extension products, but the 90-nt product, which was the most abundant in reverse transcription runoff assays, was not detectable in S1 mapping analysis of *P. vulgaris* RNA (not shown).

A fourth *pvuIIR* primer extension product, ~320 nt long, was also detected. The corresponding transcript extends through the entire coding sequence of *pvuIIC* and predicts a cluster of RNA start sites 25 to 35 nt upstream of the initiator codon of *pvuIIC*. The R320 transcript appears to identify the same starting point as C47 (Fig. 3A and B). Thus, some very long transcripts may extend from upstream of *pvuIIR* through *pvuIIR* (C240 and R320).

Assays of promoter activity were used to assess the likelihood that the shorter *pvuIIR* transcripts reflect functional *pvuIIR*-specific promoters. CAT plasmid inserts that extended to approximately 140 bp upstream of the translational start site of *pvuIIR* (and which encompassed potential promoters for R90 and R140) exhibited very low levels of promoter activity (pRV1a; Fig. 4D). Promoter strength was also analyzed in a segment that added ~100 upstream base pairs to the segment in pRV1a and that would include a promoter for the observed 160-nt transcript (pRV2a); no significant CAT activity was detected with this clone either. (Note that pRV3b has been included in both Fig. 3D and 4D to facilitate comparison.) The CAT analyses therefore failed to define any sequences within 250 bp upstream of *pvuIIR* that possessed significant promoter activity whether or not C · PvuII was present. These observations strengthen the likelihood that the three shorter reverse transcripts detected in the primer extension analyses (R90, R140, and R165) represent premature termination by the reverse transcriptase rather than true transcriptional initiation sites. Other possible explanations are discussed below.

Do the *pvuIIM* promoters interfere with the *pvuIICR* promoter? In the native PvuII restriction-modification system, the *pvuIIM* promoters are opposed by the *pvuIIC* promoter. In theory, there could be mutual interference between the *pvuIIM* and *pvuIIC* promoters. CAT clones containing the C promoter (including the C box) and the M82 promoter suggest that this is not the case. Plasmids pRV5a and pRV5b represent both orientations of a *SpeI-HindIII* fragment. The CAT activity resulting from pRV5a increased ~25-fold in the presence of C · PvuII (Fig. 3D), but this large increase in opposing transcription was not associated with a significant change in CAT activity resulting from pRV5b (Fig. 2D).

DISCUSSION

In this study we have characterized the transcription units of the PvuII system as a step in understanding the regulatory mechanisms that control its potentially lethal genes.

General model. We previously proposed that the requirement for C proteins serves to generate a timing delay, allowing MTase to appear before REase in new host cells (59). The greatest basal promoter activity we observed in the PvuII system is associated with *pvuIIM* (Fig. 2D; pRV3a and pRV5b). This strength indicates efficient expression of the MTase gene and is consistent with the need for rapid protective DNA modification in a new host cell. In contrast, REase expression is probably delayed, as expected, because the polycistronic *pvuIICR* promoter is relatively weak in the absence of activation (Fig. 3D; pRV4a and pRV5a). The transformation studies described in Table 1 confirm that appropriate regulation of the polycistronic *pvuIICR* promoter is critical to host cell survival. For comparison, such transcription level delay of REase expression was not found in a transferable type I restriction-modification system (in which the REase is not an independently active protein), though some type of posttranscriptional control appeared to play the equivalent role (43, 44).

The results of the transcription assays are consistent with those of previous studies that implicate C proteins as stimulators of REase gene expression (5, 13, 31, 36, 48, 58, 59). However, the present study provides the first mechanistic evidence that C · PvuII is a DNA-binding protein that binds to the C box and that autogenous activation by C · PvuII of the polycistronic *pvuIICR* promoter contributes to the temporal regulation of *pvuIIR* expression. Similar evidence is being accumulated with respect to C · BamHI (A. Sohail, I. Ghosh, R. M. Fuentes, and J. E. Brooks, unpublished results).

C boxes. The C proteins probably contain helix-turn-helix motifs and appear to act as homodimers. Most dimeric helix-turn-helix proteins recognize symmetrical DNA sequences, and because some of the various C proteins can cross-complement, we searched the aligned C box regions for such symmetrical elements. As shown in Fig. 1A, we found two adjacent elements, the 5'-most of which comprises nearly all of the originally defined C box. In five of five cases examined, 2 occurrences of the dyad consensus sequence GACTN NNAGTC (where N is any nucleotide) were found, though only 1 of the 10 occurrences matched this sequence at all eight positions (the 3' repeat in the *SmaI* system). This mismatching may be designed to increase the amount of C protein needed to saturate the sites and to increase cooperativity in the binding of the two sites. An introduced AGTC→ATTC change in the 5' dyad abolished C · PvuII binding to a synthetic oligonucleotide (Fig. 1A and B). C · PvuII does bind to an oligonucleotide containing the double consensus dyad (not shown); we are currently determining the stoichiometry of this binding.

In addition to sequence, the spacing and polarity of these sites are conserved. On the sense strands (as defined by the downstream C genes), the central N in GACTNNNAGTC is without exception an adenine, and the central 2 nucleotides of 4 separating each pair of GACTNNNAGTC sites is without

FIG. 3. Transcription of *pvuIIC*, the gene for the PvuII regulatory protein. Except where indicated, panel descriptions correspond to those in the legend to Fig. 2. (A) Results of reverse transcriptase primer runoff assays. Total RNA from *P. vulgaris* (left and right panels) or from an *E. coli* strain carrying the PvuII genes on a moderate-copy-number vector (center panel) was used as a template for reverse transcriptase, with a ³²P-end-labeled primer complementary to the mRNA near the 5' end of the *pvuIIC* coding region. The two initiation points upstream of *pvuIIC* common to both *E. coli* and *P. vulgaris* RNA are indicated by arrows with filled circles at the ends. (B) Map positions of *pvuIIC* transcript starting points. The dotted ending to reverse transcript C47 indicates the stuttering start shown in panel A. (C) Sequences upstream of *pvuIIC* transcript starting points. These sequences are aligned with respect to the transcript starting points (numbered as +1; for C47 one of the starting points found in both *E. coli* and *P. vulgaris* is used). The dotted bar underneath C47 indicates the C · PvuII-shifted oligonucleotide (Fig. 1B), and the linked rectangles show the C-box-associated dyad repeats (Fig. 1A), while the open boxes at the left indicate a possible transcription-enhancing UP element and its consensus (23, 54). (D) Promoter activity associated with various DNA segments upstream of *pvuIIC*. Relative to Fig. 2D, these four segments are in the opposite orientation; thus any promoter activity detected would be rightward. The small hatched box represents the position of the C box. Each clone was assayed for CAT activity in the presence of either pFLAG.2-*pvuIIC* (solid bars) or pFLAG.2-*pvuIIC*^{Esp} (shaded bars). Note that the scale differs from that in Fig. 2D. Triplicate assays of each dual transformant were carried out, and the mean values are shown; standard errors ranged from 4 to 51% of the means (not shown; errors for samples with ≥50 relative activity units ranged from 4 to 33% of the means).

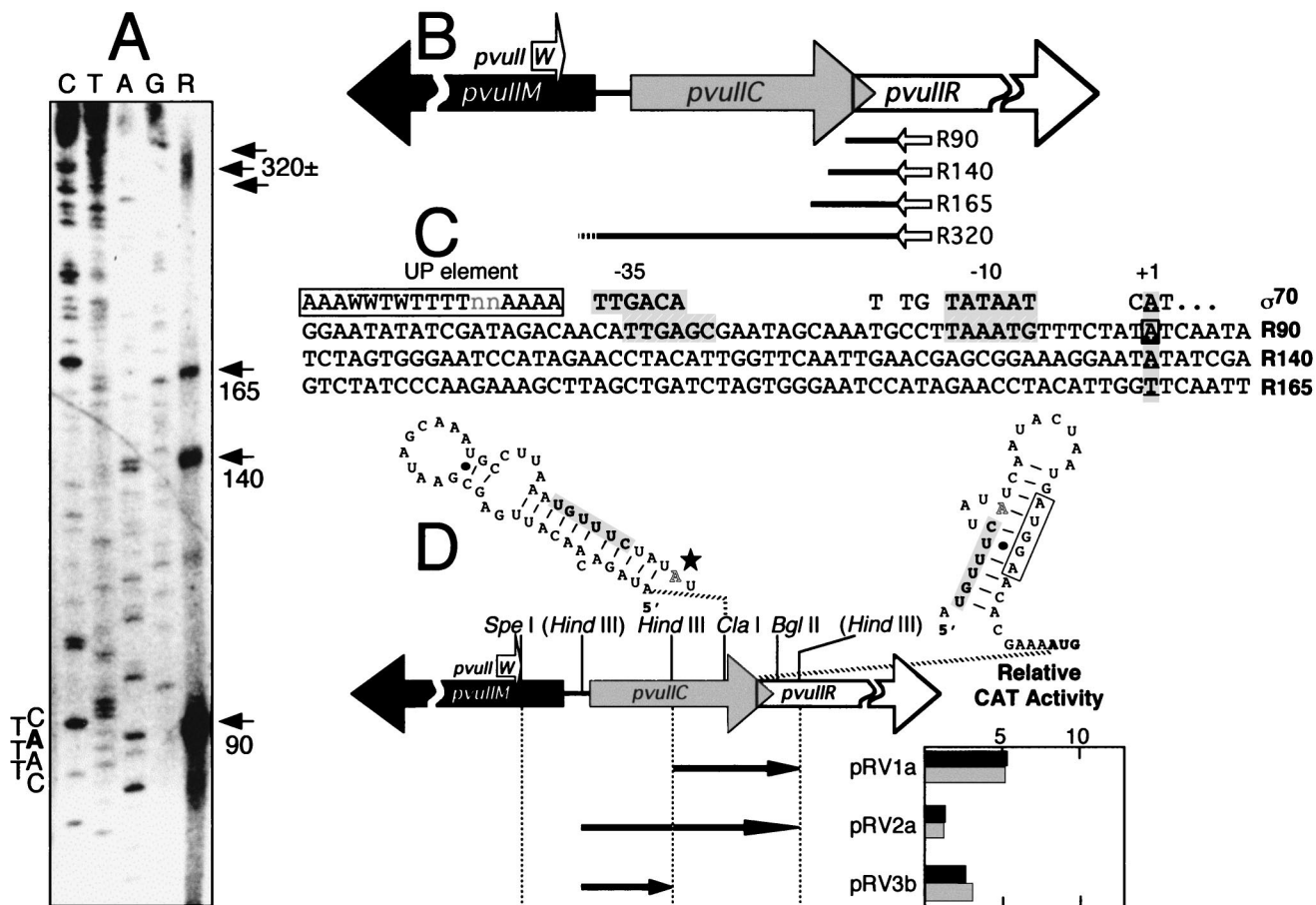


FIG. 4. Transcription of *pvuIIR*, the gene for the *PvuII* REase. Except where indicated, panel descriptions correspond to those in the legend to Fig. 2. (A) Results of reverse transcriptase primer runoff assays. Total RNA from *P. vulgaris* was used as the template for reverse transcriptase, with a ³²P-end-labeled primer complementary to the mRNA near the 5' end of the *pvuIIR* coding region. (B) Map positions of *pvuIIR* transcript starting points. The dotted ending to reverse transcript R320 corresponds to the stuttering start of transcript C47 (Fig. 3A and B). (C) Sequences upstream of *pvuIIR* transcript starting points. R320 is not shown because it is the same as C47 (Fig. 3B). Aside from R320, only R90 showed any obvious match to the σ^{70} promoter consensus (hatched shaded boxes), and there is reason to believe that this transcript "start" is artifactual. (D) Promoter activity associated with various DNA segments upstream of *pvuIIR*. The *HindIII* sites in parentheses are nonnative and were generated by site-directed mutagenesis. Relative to Fig. 2D, these three segments are in the opposite orientation; thus, any promoter activity detected would be rightward. Each clone was assayed for CAT activity in the presence of either pFLAG.2-*pvuIIC* (solid bars) or pFLAG.2-*pvuIIC*^{ESP} (shaded bars). Note that the scale differs from that in Fig. 3D; clone pRV3b is shown both here and in Fig. 3D to facilitate comparison. Triplicate assays of each dual transformant were carried out, and the mean values are shown; standard errors ranged from 10 to 39% of the means (not shown). Above the genetic map are the alternative hairpin structures predicted to form in the mRNA immediately upstream of *pvuIIR* (57); the shaded nucleotides are shared by the two structures, and the open box indicates the putative Shine-Dalgarno sequence for *pvuIIR*. A star (left hairpin) indicates the position of the first nucleotide in transcript R90.

exception GT (Fig. 1A). *Bam*HI deviates slightly from the spacing rules, missing 1 bp from the 3' repeat. It is not clear why the polarity is conserved in all cases. In at least some cases, however, the overlapping promoter poses sequence constraints that may require a particular polarity. If in fact each GACTN NNAGTC is bound by a C protein dimer, the center-to-center distance of 15 bp for the two sites means that the two dimers would occupy opposite faces of the double helix. This pattern of multiple binding sites, all with the same polarity, on two faces of the double helix is also seen with another transcriptional activator—the leucine-responsive regulatory protein (Lrp) of *E. coli* (56, 62, 64). Lrp, like the C proteins, is a small dimeric protein with a predicted helix-turn-helix motif that recognizes a symmetrical sequence with a central A/T triplet (21, 51, 56). The interactions with RNA polymerase are not understood for either Lrp or the C proteins.

Cotranscription of *pvuIIC* and *pvuIIR*. Polycistronic transcripts that include the REase gene have previously been re-

ported for some type II restriction-modification systems, such as the *EcoRI* and *SalI* systems (39, 52, 53). In both of these examples, however, the REase gene precedes the MTase gene and an internal promoter for the MTase is also present. The existence of a polycistronic message for the *PvuII* genes was revealed not only by the extended transcript observed in primer extension assays but also by assays of promoter activity. We were unable to detect any significant independent (monocistronic) *pvuIIR* promoter activity in any of the CAT plasmids, despite the appearance of reverse transcripts that implied that such a promoter might exist. As described earlier, these shorter *pvuIIR*-specific products could be due to premature termination in the reverse transcription reactions.

Two caveats must be mentioned here, however. First, it is possible that some promoters are active in *P. vulgaris* but not in *E. coli*. This would be consistent with indirect evidence that promoters can behave somewhat differently in *Proteus* and *Escherichia* (see, e.g., references 9 and 49), even though both

genera belong to the *Enterobacteriaceae*. The transcript maps were generated with RNA from both *P. vulgaris* and an *E. coli* strain bearing a plasmid clone of the *PvuII* genes, and only minor qualitative differences were seen. Interestingly, the yield of reverse transcripts was significantly lower with the *E. coli* than with the *P. vulgaris* RNA, despite the fact that pPvuRM3.4 is a pBR322-derived plasmid that has a copy number of 50 to 60 in rich medium (34), while the native *PvuII* plasmid pPvu1 appears to have a very low copy number (16). As equal amounts of total RNA were used, and as the primers and other conditions did not vary, this suggests that the *PvuII* genes are transcribed less efficiently (or that their transcripts are less stable) in *E. coli* than in *P. vulgaris*. It was only possible to carry out the CAT promoter assays in *E. coli*.

The second caveat is that the transcript mapping and CAT assays were carried out with strains in which the *PvuII* genes were already established. Some promoters may be expressed only transiently following entry into a new host cell, and we are currently testing this possibility. It thus remains possible that cryptic promoters for C240, C136, and R165 might be active under some conditions.

Relationship between promoters for *pvuIIM* and *pvuICR*. C protein-associated reduction in expression of the MTase gene has been reported in the case of *BamHI* (13), though we saw no such effect in the *PvuII* system. In both systems a pair of MTase promoters oppose the C gene promoters (Sohail et al., unpublished results), so C proteins might influence MTase gene transcription via interfering convergent transcription as in bacteriophage lambda (63). In the *PvuII* system, transcription of *pvuICR* can apparently increase 25-fold without noticeably decreasing transcription from the opposing *pvuIIM* promoters (Fig. 2D and 3D). This is consistent with the observation that when the *trp* and *lacUV5* promoters were placed in opposition to one another, in vitro transcriptional interference resulted only under abnormal conditions (low purine nucleoside triphosphate concentrations [29]). At least under some conditions, in vivo transcription from tandem promoters can lead to RNA polymerase collisions and termination by the trailing polymerase (42), though among the spacings tested this effect was seen only when the two promoters were 83 bp apart, and in *pvuIIM* the two transcript starts are separated by half that distance. It would be interesting to see if relative use of the two *pvuIIM* promoters changes during establishment, in analogy to the way growth conditions affect relative use of the two tandem promoters upstream of rRNA operons in *E. coli* (33).

A second possible interaction between the C and MTase genes involves 5'-end hybridization of the complementary *pvuIIM* and *pvuICR* transcripts. These transcripts are opposite in polarity and overlap by 62 or 101 nt (depending on which *pvuIIM* transcript is involved and using the 5'-most of the clustered starts of *pvuICR* mRNA). Unless one or both complementary transcripts are rapidly loaded with ribosomes, they would likely hybridize as they are being produced in close proximity to one another. Hybridization would occlude the entire *pvuIIC* translation initiation region, in analogy to the effect of *micF* RNA on *ompF* mRNA (2), and may dampen what would otherwise be a potentially explosive autogenous activation circuit. Translation of *pvuIIM* begins at alternate Met codons 39 nt apart, generating protein products that differ by 13 aa; when the cloned *PvuII* system is established in *E. coli*, more than 90% of the initiation is at the internal Met codon (11). Hybridization of the *pvuIIM* and *pvuICR* mRNAs would only occlude the upstream initiator, and the initiator at Met14 would be outside the double-stranded region. Thus, increased transcription of *pvuICR* may act as a switch between the two *pvuIIM* translation initiators and explain why the upstream

initiator is used only 5 to 10% of the time in cells with the established *PvuII* restriction-modification system.

Mechanism of activation. C · *PvuII* was initially proposed to be a transcriptional regulatory protein on the basis of amino acid sequence similarity to other prokaryotic activators and repressors, and due to the increase in R · *PvuII* activity in its presence (58, 59). The increased activity of the *pvuICR* promoter when C · *PvuII* is provided in *trans* strongly supports the proposed activator function (Fig. 3D), especially when combined with our observation that C · *PvuII* binds specifically to the C box DNA sequence (Fig. 1B).

Almost all characterized bacterial transcription activators bind upstream of their target promoter, though several overlap the -35 hexamer (41). For this reason it is surprising that the *pvuICR* transcripts begin (in both *E. coli* and *P. vulgaris*) immediately adjacent to the C box sequence (Fig. 3A), to which we have demonstrated that C · *PvuII* binds (Fig. 1). This observation is not dependent on our tentative assignment of -10 and -35 hexamers, it is not affected by the fact that we have not yet demonstrated where RNA polymerase binds, and it does not depend on the presence or absence of possible additional promoters closer to *pvuIIR*. If the symmetrical sequences identified in Fig. 1A are both bound by C · *PvuII*, as suggested by the large DNase I footprint observed by others with C · *BamHI* (Sohail et al., unpublished results), then C · *PvuII* binds to DNA that completely spans the transcription start sites.

Given the apparent location of the *pvuICR* -10 hexamer within the sequence that is bound by C · *PvuII* (Fig. 1 and 3C), it is possible that the C proteins are activating transcription via an unusual mechanism. Three activators, IlvY, SoxR, and MerR, are known to bind farther upstream between the -35 and -10 hexamers (without overlapping the -10 hexamer), and all three activate transcription by modulating the twist or bending of the DNA (4, 27, 47). Transcriptional activation by SoxR and MerR depends on the nonconsensus spacing of 19 nt between the -35 and -10 hexamers; if the spacing is reduced to 18 nt, which is the apparent spacing in the *pvuICR* promoter (Fig. 3C), then the promoters controlled by SoxR and MerR exhibit high basal rates of transcription. However, even 18-nt spacing has been associated with a substantial weakening of promoters, which can be overcome by negative supercoiling (6). In the case of IlvY, the ideal 17-nt hexamer spacing is present but there is a weak -35 hexamer; IlvY bends the promoter DNA and thus enhances RNA polymerase binding (47). There are some transcriptional activators that have binding sites both upstream and downstream of the promoter; examples include PhoP of *Bacillus subtilis* (35) and SpvR of *Salmonella* (55). The downstream binding site of SpvR covers +1 to +27 relative to the start of transcription; however, binding that spans the -10 hexamer and transcription start site has not, to our knowledge, previously been observed for a bacterial transcription activator.

We would like to raise one additional possibility for the mechanism of action of C · *PvuII* at the C box, which we consider unlikely but which our present data cannot rule out. That is the possibility that C · *PvuII* is acting as an antiterminator rather than as an activator. Our reverse priming experiments would not have detected transcripts that terminated close to the *pvuICR* promoter, and the CAT plasmids only measure net transcription emerging from the cloned DNA segment without distinguishing between activation and antitermination. However the shortest of the C · *PvuII*-stimulated inserts ends ~120 bp downstream of the transcription starting cluster, and there are no obvious (rho-independent) transcription terminators in that region.

Whatever mechanism is used by $C \cdot PvuII$ to stimulate *pvuII*CR transcription is probably also used by other C proteins. First, in all cases analyzed to date, the C protein stimulates expression of the REase gene, though the mechanism of stimulation has not been defined. Second, the C genes consistently occur upstream of and in the same orientation as the R genes. The *SmaI* REase gene, in particular, has also been shown to be transcribed as part of a polycistronic *smaI*CR message together with its upstream C gene (26). Third, the C genes from organisms as different as *Bacillus* and *Proteus* can cross-complement (31, 36).

C proteins as regulators. The autogenous regulation of *pvuII*C has implications for its genetic mobility. The C protein genes appear to represent readily moved regulatory modules, as they should positively regulate the expression of any gene to which they insert upstream. This feature, together with the ability to function with a variety of bacterial RNA polymerases (31, 36), may substantially widen the host range of the genes they control. It could also explain the observation that the various restriction-modification systems containing closely related C genes have essentially unrelated MTase and REase genes.

The observations described above provide a role and suggest mechanisms for the temporal regulation of the C-producing restriction-modification systems by C proteins. A question remains, however, as to whether there are additional roles for proteins such as $C \cdot PvuII$. A site of $C \cdot PvuII$ action was functionally mapped to within 70 bp of the *pvuII*R translational initiation codon (58) (downstream of the *Cl*A site shown in Fig. 4D). Experiments are under way to determine whether the observed $C \cdot PvuII$ activation of *pvuII*R expression arises from additional transcriptional or posttranscriptional events that are independent of the C box and the *pvuII*CR promoter.

ACKNOWLEDGMENTS

We thank Jessica L. Brust for technical support and Alexander J. Ninfa (University of Michigan) and Joan E. Brooks (Proteome, Inc.) for critical reading of the manuscript. We also thank Joan E. Brooks for sharing unpublished results.

This research was supported by the U.S. National Science Foundation under grants MCB-9205248 (to R.M.B. and J.C.D.) and MCB-9631137 and MCB-9904523 (to R.M.B.). R.M.V. received additional support from the Wayne State University School of Medicine.

REFERENCES

- Adams, G. M., and R. M. Blumenthal. 1995. Gene *pvuIIW*: a possible modulator of *PvuII* endonuclease subunit association. *Gene* **157**:193–199.
- Anderson, J., S. A. Forst, K. Zhao, M. Inouye, and N. Delias. 1989. The function of *micF* RNA. *micF* RNA is a major factor in the thermal regulation of OmpF protein in *Escherichia coli*. *J. Biol. Chem.* **264**:17961–17970.
- Anderson, J. E. 1993. Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.* **3**:24–30.
- Ansari, A. Z., J. E. Bradner, and T. V. O'Halloran. 1995. DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* **374**:371–375.
- Anton, B. P., D. F. Heiter, J. S. Benner, E. J. Hess, L. Greenough, L. S. Moran, B. E. Slatko, and J. E. Brooks. 1997. Cloning and characterization of the *Bgl*II restriction-modification system reveal a possible evolutionary footprint. *Gene* **187**:19–27.
- Aoyama, T., and M. Takamami. 1988. Supercoiling response of *E. coli* promoters with different spacer lengths. *Biochim. Biophys. Acta* **949**:311–317.
- Athanasiadis, A., M. Gregoriu, D. Thanos, M. Kokkinidis, and J. Papamatheakis. 1990. Complete nucleotide sequence of the *PvuII* restriction enzyme gene from *Proteus vulgaris*. *Nucleic Acids Res.* **18**:6434.
- Athanasiadis, A., M. Vlasi, D. Kotsifaki, P. A. Tucker, K. S. Wilson, and M. Kokkinidis. 1994. Crystal structure of *PvuII* endonuclease reveals extensive structural homologies to *EcoRV*. *Struct. Biol.* **1**:469–475.
- Baumberg, S., and M. Roberts. 1984. Anomalous expression of the *E. coli lac* operon in *Proteus mirabilis*. II. Effects of *lacI* and *lacP* mutations. *Mol. Gen. Genet.* **198**:166–171.
- Beach, M. B., and R. Osuna. 1998. Identification and characterization of the *fis* operon in enteric bacteria. *J. Bacteriol.* **180**:5932–5946.
- 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.
- Brizzard, B. L., R. G. Chubet, and D. L. Vizard. 1994. Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. *BioTechniques* **16**:730–735.
- Brooks, J. E., P. D. Nathan, D. Landry, L. A. Szyner, 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 methylase, and expression in heterologous hosts. *Nucleic Acids Res.* **19**:841–850.
- Brosius, J. 1984. Plasmid vectors for the selection of promoters. *Gene* **27**:151–160.
- Butkus, V., S. Klimasauskas, L. Petrauskienė, Z. Maneliene, A. Leblionka, and A. A. Janulaitis. 1987. Interaction of *AluI*, *Cfr*6I and *PvuII* restriction-modification enzymes with substrates containing either N4-methylcytosine or 5-methylcytosine. *Biochim. Biophys. Acta* **909**:201–207.
- Calvin-Koons, M. D., and R. M. Blumenthal. 1995. Characterization of *pPvu1*, the autonomous plasmid from *Proteus vulgaris* that carries the genes of the *PvuII* restriction-modification system. *Gene* **157**:73–79.
- Chang, A. C., 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.
- Cheng, X. 1995. Structure and function of DNA methyltransferases. *Annu. Rev. Biophys. Biomol. Struct.* **24**:293–318.
- Cheng, X., K. Balendiran, I. Schildkraut, and J. E. Anderson. 1995. Crystal structure of the *PvuII* restriction endonuclease. *Gene* **157**:139–140.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156–159.
- Cui, Y., Q. Wang, G. D. Stormo, and J. M. Calvo. 1995. A consensus sequence for binding of Lrp to DNA. *J. Bacteriol.* **177**:4872–4880.
- De Backer, O., and C. Colson. 1991. Transfer of the genes for the StyLTI restriction-modification system of *Salmonella typhimurium* to strains lacking modification ability results in death of the recipient cells and degradation of their DNA. *J. Bacteriol.* **173**:1328–1330.
- Estrem, S. T., T. Gaal, W. Ross, and R. L. Gourse. 1998. Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. USA* **95**:9761–9766.
- 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.
- Gong, W., M. O'Gara, R. M. Blumenthal, and X. Cheng. 1997. Structure of *PvuII* DNA-(cytosine N4) methyltransferase, an example of domain permutation and protein fold assignment. *Nucleic Acids Res.* **25**:2702–2715.
- Heidmann, S., W. Seifert, C. Kessler, and H. Domdey. 1989. Cloning, characterization and heterologous expression of the *SmaI* restriction-modification system. *Nucleic Acids Res.* **17**:9783–9796.
- Hidalgo, E., and B. Dimple. 1997. Spacing of promoter elements regulates the basal expression of the *soxS* gene and converts SoxR from a transcriptional activator into a repressor. *EMBO J.* **16**:1056–1065.
- Hopp, T. P., B. Gallis, and K. S. Prickett. 1996. Metal-binding properties of a calcium-dependent monoclonal antibody. *Mol. Immunol.* **33**:601–608.
- Horowitz, H., and T. Platt. 1982. Regulation of transcription from tandem and convergent promoters. *Nucleic Acids Res.* **10**:5447–5465.
- Horton, J. R., J. Bonventre, and X. Cheng. 1998. How is modification of the DNA substrate recognized by the *PvuII* restriction endonuclease? *Biol. Chem.* **379**:451–458.
- Ives, C. L., A. Sohail, and J. E. Brooks. 1995. The regulatory C proteins from different restriction-modification systems can cross-complement. *J. Bacteriol.* **177**:6313–6315.
- Klotsky, R. A., and I. Schwartz. 1987. Measurement of *cat* expression from growth-rate-regulated promoters employing beta-lactamase activity as an indicator of plasmid copy number. *Gene* **55**:141–146.
- Liebig, B., and R. Wagner. 1995. Effects of different growth conditions on the *in vivo* activity of the tandem *Escherichia coli* ribosomal RNA promoters P1 and P2. *Mol. Gen. Genet.* **249**:328–335.
- Lin-Chao, S., and H. Bremer. 1986. Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*. *Mol. Gen. Genet.* **203**:143–149.
- Liu, W., Y. Qi, and F. M. Hulett. 1998. Sites internal to the coding regions of *phoA* and *pstS* bind PhoP and are required for full promoter activity. *Mol. Microbiol.* **28**:119–130.
- Nakayama, Y., and I. Kobayashi. 1998. Restriction-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion. *Proc. Natl. Acad. Sci. USA* **95**:6442–6447.
- Noel, R. J., Jr., and W. S. Reznikoff. 1998. CAP, the –45 region, and RNA polymerase: three partners in transcription initiation at *lacP1* in *Escherichia coli*. *J. Mol. Biol.* **282**:495–504.
- O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler. 1972. Novel

- method for detection of β -lactamases by using a chromogenic cephalosporin substrate. *Antimicrob. Agents Chemother.* **1**:283–288.
39. O'Connor, C. D., and G. O. Humphreys. 1982. Expression of the *EcoRI* restriction-modification system and the construction of positive-selection cloning vectors. *Gene* **20**:219–229.
 40. O'Neill, M. C. 1989. *Escherichia coli* promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. *J. Biol. Chem.* **264**:5522–5530.
 41. Perez-Rueda, E., J. D. Gralla, and J. Collado-Vides. 1998. Genomic position analyses and the transcription machinery. *J. Mol. Biol.* **275**:165–170.
 42. Ponnambalam, S., and S. Busby. 1987. RNA polymerase molecules initiating transcription at tandem promoters can collide and cause premature transcription termination. *FEBS Lett.* **212**:21–27.
 43. Prakash-Cheng, A., S. S. Chung, and J.-I. Ryu. 1993. The expression and regulation of *hsd_K* genes after conjugative transfer. *Mol. Gen. Genet.* **241**:491–496.
 44. Prakash-Cheng, A., and J. Ryu. 1993. Delayed expression of in vivo restriction activity following conjugal transfer of *Escherichia coli* *hsd_K* (restriction-modification) genes. *J. Bacteriol.* **175**:4905–4906.
 45. Raleigh, E. A., and J. E. Brooks. 1998. Restriction modification systems: where they are and what they do, p. 78–92. In F. J. De Bruijn, J. R. Lupski, and G. M. Weinstock (ed.), *Bacterial genomes*. Chapman & Hall, New York, N.Y.
 46. Raleigh, E. A., N. E. Murray, H. Revel, R. M. Blumenthal, D. Westaway, A. D. Reith, P. W. J. Rigby, J. Elhai, and D. Hanahan. 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucleic Acids Res.* **15**:1563–1575.
 47. Rhee, K. Y., D. F. Senear, and G. W. Hatfield. 1998. Activation of gene expression by a ligand-induced conformational change of a protein-DNA complex. *J. Biol. Chem.* **273**:11257–11266.
 48. Rimseliene, R., R. Vaisvila, and A. Janulaitis. 1995. The *eco72IC* gene specifies a *trans*-acting factor which influences expression of both DNA methyltransferase and endonuclease from the *Eco72I* restriction-modification system. *Gene* **157**:217–219.
 49. Roberts, M., and S. Baumberg. 1984. Anomalous expression of the *E. coli lac* operon in *Proteus mirabilis*. I. Effects of *L8* and *L8 UV5*. *Mol. Gen. Genet.* **198**:159–165.
 50. Roberts, R. J., and X. Cheng. 1998. Base flipping. *Annu. Rev. Biochem.* **67**:181–198.
 51. Robison, K., A. M. McGuire, and G. M. Church. 1998. A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J. Mol. Biol.* **284**:241–254.
 52. Rodicio, M. R., and K. F. Chater. 1988. Cloning and expression of the *SalI* restriction-modification genes of *Streptomyces albus* G. *Mol. Gen. Genet.* **213**:346–353.
 53. Rodicio, M. R., T. Quinton-Jager, L. S. Moran, B. E. Slatko, and G. G. Wilson. 1994. Organization and sequence of the *SalI* restriction-modification system. *Gene* **151**:167–172.
 54. Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R. L. Gourse. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**:1407–1413.
 55. Sheehan, B. J., and C. J. Dorman. 1998. In vivo analysis of the interactions of the LysR-like regulator SpvR with the operator sequences of the *spvA* and *spvR* virulence genes of *Salmonella typhimurium*. *Mol. Microbiol.* **30**:91–105.
 56. Shultzaberger, R. K., and T. D. Schneider. 1999. Using sequence logos and information analysis of Lrp DNA binding sites to investigate discrepancies between natural selection and SELEX. *Nucleic Acids Res.* **27**:882–887.
 57. Tao, T. 1992. Doctoral dissertation. Medical College of Ohio, Toledo.
 58. Tao, T., and R. M. Blumenthal. 1992. Sequence and characterization of *pvuIIR*, the *PvuII* endonuclease gene, and of *pvuIIC*, its regulatory gene. *J. Bacteriol.* **174**:3395–3398.
 59. 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.
 60. Tao, T., J. Walter, K. J. Brennan, M. M. Cotterman, and R. M. Blumenthal. 1989. Sequence, internal homology and high-level expression of the gene for a DNA-(cytosine N4)-methyltransferase, M · *PvuII*. *Nucleic Acids Res.* **17**:4161–4175.
 61. Vijesurier, R. M. 1996. Doctoral dissertation. Wayne State University, Detroit, Mich.
 62. Wang, Q., M. Sacco, E. Ricca, C. T. Lago, M. DeFelice, and J. M. Calvo. 1993. Organization of Lrp-binding sites upstream of *ilvIH* in *Salmonella typhimurium*. *Mol. Microbiol.* **7**:883–891.
 63. Ward, D. F., and N. E. Murray. 1979. Convergent transcription in bacteriophage lambda: interference with gene expression. *J. Mol. Biol.* **133**:249–266.
 64. Wiese, D. E., II, B. R. Ernsting, R. M. Blumenthal, and R. G. Matthews. 1997. A nucleoprotein activation complex between the leucine-responsive regulatory protein and DNA upstream of the *gltBDF* operon in *Escherichia coli*. *J. Mol. Biol.* **270**:152–168.
 65. Wilson, G. G., and N. E. Murray. 1991. Restriction and modification systems. *Annu. Rev. Genet.* **25**:585–627.
 66. Winkler, F. K. 1992. Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.* **2**:93–99.
 67. Wintjens, R., and M. Rooman. 1996. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. *J. Mol. Biol.* **262**:294–313.
 68. Withers, B. E., and J. C. Dunbar. 1995. Sequence-specific DNA recognition by the *SmaI* endonuclease. *J. Biol. Chem.* **270**:6496–6504.