JOURNAL OF BACTERIOLOGY, Sept. 1988, p. 3838–3842
0021-9193/88/093838-05$02.00/0

Vol. 170, No. 9

# Transcription of the *celE* Gene in *Thermomonospora fusca*

ERSHEN LIN† AND DAVID B. WILSON*

*Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853*

Received 8 February 1988/Accepted 3 June 1988

**The steady-state level of *celE* mRNA (coding for cellulase $E_5$) in *Thermomonospora fusca* YX was measured by Northern (RNA blot) hybridization under conditions causing induction or repression of cellulase synthesis. A good correlation was found between the mRNA level and the level of cellulase $E_5$, suggesting that the *T. fusca celE* gene is regulated at the level of mRNA and, most likely, at the level of transcription. The 5′ and 3′ ends of the *celE* gene transcription unit were determined by S1 mapping with single-stranded DNA probes. These results showed that there were three species of *celE* mRNA in *T. fusca* YX with closely spaced 5′ ends and identical 3′ ends. The size of each mRNA was about 1.5 kilobases, from both the Northern and S1 data. This size is only slightly longer than that required to code for the 45-kilodalton $E_5$ protein. In *Escherichia coli* D318 (*celE*), the 5′ ends of the *celE* mRNAs are identical to those in *T. fusca*, but the 3′ ends are located ca. 300 base pairs upstream of the *T. fusca* 3′ end. The region where the putative *celE* promoters were located had some interesting features, including a 60-base-pair A+T-rich sequence and sequences resembling $\sigma^{60}$ promoters.**

In cellulolytic bacteria and fungi, cellulase synthesis is usually regulated by induction and repression (4, 7, 9, 13, 15, 16, 26, 28). The study of the regulation of cellulase synthesis and its molecular mechanism is complicated by the multiplicity of cellulase genes and products as well as the lack of basic information about cellulolytic microorganisms. Nothing was known about the molecular mechanisms of cellulase gene regulation until a recent study with *Cellulomonas fimi* showed a positive correlation between the level of the mRNA for two cellulase genes and the overall cellulase level (10).

We have previously provided evidence that both induction and repression cause coordinated regulation of the cellulase genes in *Thermomonospora fusca* YX (17). Here we report on the expression of the *T. fusca* cellulase $E_5$ gene (*celE* gene) (5). The levels of *celE* mRNA were measured in *T. fusca* YX, and the 5′ ends and the 3′ ends of *celE* mRNA isolated from both *T. fusca* and *Escherichia coli* were determined by S1 mapping.

## MATERIALS AND METHODS

**Bacterial and phage strains.** *E. coli* D318 is strain HB101 transformed by plasmid pD318. Plasmid pD318 (Fig. 1) was described previously (5). Phage M13mp10 and the host *E. coli* strain JM101 (kindly supplied by R. Wu) were described by Messing (20).

**Preparation of the Northern hybridization probe.** pD318 plasmid DNA was cut with the restriction enzymes *Sma*I and *Stu*I. The mixture was fractionated by electrophoresis on a 0.7% agarose gel, and the small fragment (0.37 kilobase [kb]) was electroblotted onto DEAE-cellulose membrane NA45 (Schleicher & Schuell) and eluted with 1.0 M NaCl–20 mM Tris (pH 8.0)–0.1 mM EDTA at 55°C, as suggested by the supplier. After five *n*-butanol extractions and two ethanol precipitations, the fragment was labeled with [α-$^{32}$P]dCTP (800 Ci/mmol) by nick translation (19) to give a specific radioactivity of $\geq 3 \times 10^7$ cpm/μg of DNA.

**RNA sample preparation.** Bacterial cultures in early to mid-log phase were centrifuged at 4,000 × *g* for 5 min at 4°C, resuspended in 1/10 the original volume of lysis buffer (200 mM NaCl, 20 mM Tris [pH 7.5], 5 mM EDTA) plus 0.5% sodium dodecyl sulfate (SDS) containing 1% diethyl pyrocarbonate, at 0°C, and lysed at 4°C in a French press sterilized by repeated SDS washing, and then RNA was isolated by the procedure given in Smith and Calvo (24). The final RNA pellet was lyophilized and dissolved in $H_2O$, and the RNA concentration was determined by the orcinol reaction.

**Northern hybridization.** RNA (10 μg) in 5 to 10 μl of $H_2O$ was mixed with 2 volumes of 75% formamide–3.3 M formaldehyde–15 mM sodium acetate (pH 7.4)–0.75 mM EDTA. The sample was denatured by heating at 65°C for 10 min, chilled in ice-water, mixed with 5× loading dye (0.825% bromophenol blue in 0.5% SDS, 25% glycerol, 20 mM EDTA), and loaded on a 0.8% agarose gel containing 10 mM sodium acetate (pH 7.8) and 1.2 M formaldehyde. The gel was run at 20 to 25 V in 1.1 M formaldehyde–10 mM sodium phosphate (pH 7.4) with buffer circulation. The RNA was transferred from the gel to a nitrocellulose membrane and hybridized to the DNA probe mentioned earlier by the method described by Southern (25) with the following modifications. Prehybridization was carried out at 40°C for 4 h, and hybridization was carried out at 40°C for 12 to 15 h. The nitrocellulose membrane was washed two times with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 40°C for 15 min and then washed two times with 0.1× SSPE–0.1% SDS; the first time was at room temperature for 15 min, and the second time was at 68°C for 10 min. The membrane was dried and exposed to Kodak XAR-5 film.

**M13 subcloning of the *celE* gene.** The 5′ end of the *celE* gene was subcloned into M13mp10 by inserting the *Eco*RI-*Sma*I fragment into m13mp10 DNA cut with *Eco*RI and *Sma*I. The 3′ end of the *celE* gene was subcloned into M13mp10 DNA by inserting the *Stu*I-*Sal*I fragment from plasmid pD318 into phage DNA cut with *Sal*I and *Sma*I as described by Messing (20). The plaques were screened for the inserts by plaque hybridization with the fragments being cloned, labeled by nick translation with [α-$^{32}$P]dCTP, as probes. Positive clones were purified and used for preparation of single-stranded DNA molecules.

**Synthesis of S1 mapping probes.** Single-stranded DNA

* Corresponding author.
† Present address: Department of Anatomy and Cell Biology, State University of New York, Health Science Center, Brooklyn, NY 11203.

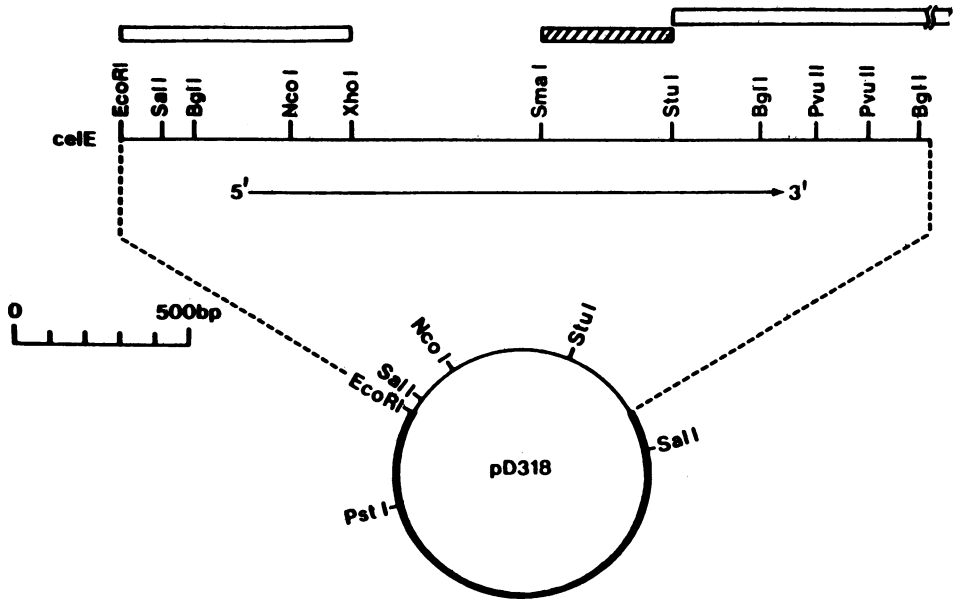


FIG. 1. *T. fusca celE* gene. Plasmid pD318 contains the *celE* gene. The DNA from plasmid pBR322 is represented by the thick line. Above the restriction map of the *celE* gene, the hatched bar represents the probe used for Northern hybridization, and the open bars represent the probes used for S1 mapping.

probes were synthesized by the primer extension procedure (20). The primer for the 3'-end probe was oligonucleotide 1212 (New England Biolab), and the primer for the 5'-end probe was the denatured *Xho*I-*Nco*I fragment from plasmid pD318.

**S1 mapping.** The method developed by Berk and Sharp (3) was used with some modifications. RNA (15 μg) was ethanol precipitated in a siliconized Eppendorf tube, centrifuged, lyophilized, dissolved in 20 μl of 80% formamide–40 mM PIPES [piperazine-*N*,*N*'-bis(2-ethanesulfonic acid), pH 6.4], 0.4 M NaCl, and 1 mM EDTA. It was denatured by heating at 67°C for 7 to 10 min. Because of its high G+C content, the single-stranded DNA probe in TE buffer (0.01 M Tris, 1 mM EDTA [pH 8.0]) was boiled for 3 min and chilled to 0°C. The RNA sample was transferred from 67 to 50°C, and the probe (≤2 μl per sample) was immediately added, mixed, and incubated at 50°C for 2.5 to 3 h and then at 68°C for 10 min. The sample was transferred to 56°C, and 0.4 ml of S1 nuclease (700 U/ml) in 280 mM NaCl–30 mM sodium acetate [pH 4.4]–4.5 mM zinc acetate–20 μg of denatured calf thymus DNA per ml was quickly added, mixed, chilled in ice, and then incubated at 37°C for 1 h. To stop the reaction, 30 μl of 2.5 M ammonium acetate–50 mM EDTA and 30 μl of carrier tRNA (1 μg/ml) were added. The DNA was precipitated with 0.55 ml of isopropanol at −70°C for 15 min and centrifuged for 15 min. The pellet was dried and dissolved in 90% formamide–0.05% bromophenol blue–0.05% xylene cyanol. The sample was boiled for 4 min, chilled to 0°C, separated on a 5% polyacrylamide–7 M urea gel, and autoradiographed.

## RESULTS

**Measurement of *celE* mRNA.** Northern hybridization was used to determine the steady-state level of *celE* in *T. fusca* YX cells grown on different carbon sources. The RNA samples were prepared from log-phase cultures, in which cellulase synthesis is highest. The results are shown in Fig. 2a. *E. coli* HB101 was used as the negative control, and no hybridization signal was detected; an RNA sample from *E. coli* D318 was used as a positive control, and it gave a 1.2-kb signal. RNA samples from *T. fusca* YX gave hybridization signals of different intensity depending on the carbon source, but the size of the mRNA was consistently 1.5 kb. The amount of *celE* mRNA present in each sample was estimated by scanning the autoradiogram with a densitometer, and the values are given in Table 1 along with the level of cellulase $E_5$ measured under identical growth conditions. There was a clear positive correlation between the relative level of cellulase $E_5$ and the level of *celE* mRNA in cells grown on different carbon sources. This shows that both the induction and the repression controls regulate the level of cellulase $E_5$ mRNA, which suggests transcriptional regulation of the *celE* gene.

The *celE* mRNA level was also measured during transient repression (17), and the results are shown in Fig. 2b. The

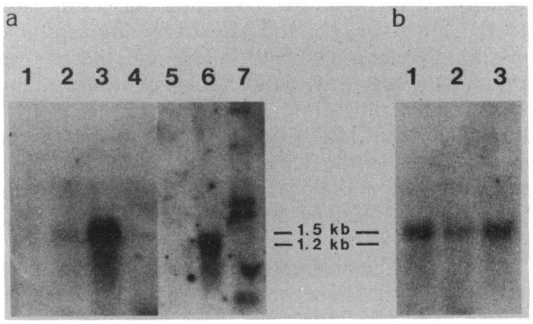


FIG. 2. *celE* mRNA levels in *T. fusca* YX and transformed *E. coli* cells. The RNA samples used were isolated from the following cultures: (a) *T. fusca* YX grown on glucose (lane 1), cellobiose (lane 2), cellulose (lane 3), and maltose (lane 4), *E. coli* HB101 (lane 5) and D318 (lane 6), and molecular weight markers (lane 7); (b) *T. fusca* YX induced by cellulose with no repressor (lane 1), with glucose as repressor (lane 2), and with maltose as repressor (lane 3).

TABLE 1. *celE* mRNA levels in *T. fusca* YX cultures[a]

| Carbon source | Relative level of *celE* mRNA | Relative level of cellulase $E_5$ | $E_5$ sp act (IU/mg) |
|---|---|---|---|
| Glucose | 0.015 | 0.006 | 0.055 |
| Maltose | 0.02 | 0.01 | 0.097 |
| Cellobiose | 0.11 | 0.10 | 0.99 |
| Cellulose | 1 | 1 | 9.8 |

[a] Cellulase $E_5$ level and activity were measured by the immunoinhibition method (17).

levels of *celE* mRNA calculated by densitometry relative to that with no repressor were 0.57 and 0.75 for glucose and maltose, respectively; the relative cellulase synthesis rate measured under identical conditions was 0.33 and 0.69, respectively. Again, there was a good correlation between the level of mRNA and the rate of overall cellulase synthesis; glucose, a strong repressor, reduced both the level of mRNA and the rate of cellulase synthesis more effectively than the weak repressor maltose.

**Mapping of the 5′ and 3′ ends of *celE* mRNA.** To determine the location of the *celE* promoter and terminator, the 5′ end and the 3′ end of the *celE* mRNA were determined by S1 mapping. The restriction fragments of plasmid pD318 corresponding to the 5′ end and the 3′ end of the *celE* gene were subcloned into M13mp10 phage, and single-stranded DNA probes were synthesized as described in Materials and Methods. Because *T. fusca* mRNA is highly unstable, the hybridization temperature was lowered to 50°C to avoid RNA degradation. It was then shifted to 68°C to denature nonspecific duplexes and finally shifted to 56°C before S1 nuclease was added to avoid enzyme inactivation. These modifications were found to be necessary in preliminary experiments.

In the 5′-end mapping, the RNA sample from *T. fusca* YX gave three protected fragments; 303 ± 2 (standard error), 299 ± 2, and 292 ± 2 nucleotides long, indicating that three closely spaced 5′ ends were located about 303, 299, and 292 base pairs (bp) upstream from the *Xho*I site (Fig. 3a). The RNA sample from *E. coli* D318 produced protected fragments identical to those of *T. fusca* RNA. Some additional signals were seen and were probably caused by nonspecific initiations in *E. coli* or readthrough from pBR322 promoters. The RNA sample from strain HB101 gave no signal, as expected.

In the 3′-end mapping (Fig. 3b), *T. fusca* YX RNA gave a signal of 326 ± 2 nucleotides, indicating that all the mRNAs ended at the same site, located ca. 326 bp downstream from the *Stu*I site. The length of *celE* mRNA calculated from the S1 mapping data was ca. 1.54 kb, consistent with the result from Northern hybridization. The RNA samples from the *E. coli* strains D318 and HB101 gave no signal with the probe, as expected, because the 1.2-kb *celE* mRNA in strain D318 should end very near the *Stu*I site and therefore would not hybridize with the 3′-end mapping probe, which starts at this site.

The DNA sequence around the 5′ ends of the mRNAs was determined (Fig. 4). There was a 61-bp region immediately upstream from the 5′ ends that had a relatively high A+T content (61%) compared with the average value of 41% for the whole noncoding region shown. If the A+T-rich region is not counted, the average A+T content was only about 35%. In addition, the DNA sequence upstream from the putative initiation sites resembled the so-called $\sigma^{60}$ promoters present in some bacteria, including species of *Escherichia* (22),

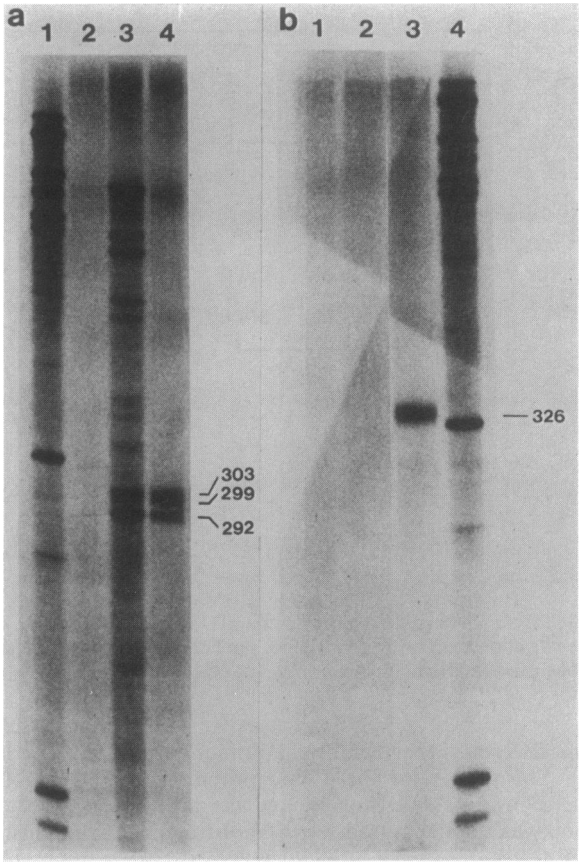


FIG. 3. S1 mapping of the 5′ and 3′ ends of *celE* mRNA. (a) 5′-end mapping. The RNA samples were from *E. coli* HB101 (lane 2), *E. coli* D318 (lane 3), and *T. fusca* YX (lane 4); molecular weight markers (lane 1). (b) 3′-end mapping. The RNA samples were from *E. coli* HB101 (lane 1), *E. coli* D318 (lane 2), and *T. fusca* YX (lane 3); molecular weight markers (lane 4). The sizes of the bands indicated on the right were calculated from the molecular weight standards and have an error of ±2 bp. The molecular weight markers were prepared by cutting λ DNA with *Hin*dIII and *Eco*RV and labeling with reverse transcriptase. The marker closest to the three 5′-end fragments is 317 bp long.

*Salmonella* (11), *Caulobacter* (Mullin et al., personal communication), *Klebsiella*, and *Rhizobium* (1), most of which are regulated by nitrogen metabolism. The distances between these sequences and the corresponding putative initiation sites were 2 bp shorter than has been found for the other promoters, which could result from an error in the S1 mapping. The sequence following the ATG start codon encoded a potential signal sequence followed by the N-terminal sequence of cellulase $E_5$. There was a 106-bp noncoding sequence present at the 5′ end of the smallest *celE* mRNA and a 117-bp noncoding sequence in the largest *celE* mRNA. There was a 5-bp-long *Streptomyces lividans* ribosome-binding site present just before the ATG start codon and an 8-bp *E. coli* ribosome-binding site starting 2 bases downstream from the initiation codon.

## DISCUSSION

In this study, we show that the level of *celE* mRNA and the rate of cellulase synthesis are correlated, suggesting that the *T. fusca celE* gene is probably regulated at the level of

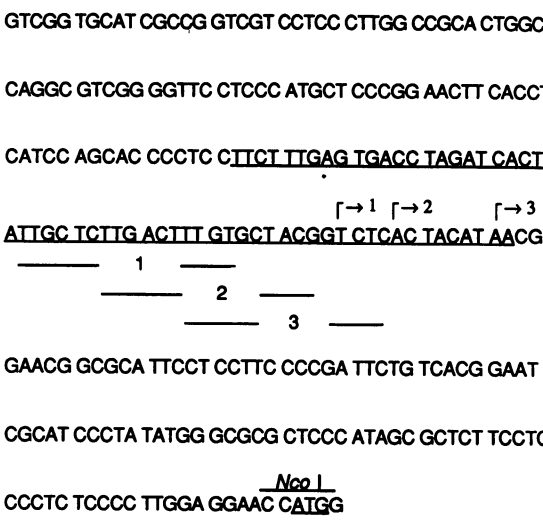


FIG. 4. DNA sequence of the *celE* promoter. The promoter region was sequenced by E. Jung (unpublished result). The 5′ ends of the *celE* mRNAs are indicated by arrows. The A+T-rich region is underlined. The DNA sequences homologous to the −24 and −12 regions of the $\sigma^{60}$ promoter are indicated by horizontal bars, with the numbers corresponding to the three putative initiation sites. The ATG start codon for the *celE* gene is also underlined.

transcription. As all the major *T. fusca* cellulases are coordinately regulated (17), this conclusion could be true for the synthesis of the other cellulases in this organism. A study with *C. fimi* (10) has shown that the levels of the mRNAs produced from two cellulase genes change in parallel when the cells are grown on different carbon sources and that these mRNA levels correlate with the total cellulase level, but it has not been shown that all the cellulases are coordinately regulated in *C. fimi*.

The putative promoters of the *celE* gene are located in a unique A+T-rich region which is a feature of many procaryotic promoters. Among typical examples are the *E. coli* $rnnp_1$ promoter (6), a *Bacillus subtilis* promoter (21), the *Klebsiella pneumoniae nifA* promoter (27), and the *C. thermocellum celA* promoter (2).

Interestingly, a recent study in which an in vivo system was used to select from strong promoters from T5 phage resulted in a group of promoters, all with an A+T-rich stretch about 43 bases upstream from the start of transcription, around position −43, as the most significant feature (8). Some in vitro studies also suggest the importance of an upstream high-A+T sequence in various procaryotic promoters (8, 12, 21). The transition temperature of the promoter may be influenced by upstream DNA sequences that do not overlap with the region where melting-in occurs during formation of the open complex (12, 14). The influence of the upstream sequence on the promoter strength may be explained by the observation that transient unwinding occurs in A+T-rich regions in superhelical DNA (23). If the same phenomenon occurs in vivo, it may facilitate the melting-in process by trapping negative superhelicity and consequently reducing the kinetic barrier for the formation of the open complex.

The DNA sequence of the putative *celE* promoters gives no significant homology when compared with cellulase promoter sequences from *C. fimi* (10) and *C. thermocellum* (2).

To our knowledge, the transcription system of *T. fusca* has not been studied, but our results show that in both *E. coli* and *T. fusca* the *celE* mRNAs have the same 5′ end. If these 5′ ends are not processing products, the transcription systems in *E. coli* and *T. fusca* must have some similarities. The level of *celE* mRNA in *E. coli* D318 is similar to the level of *celE* mRNA in induced *T. fusca*, yet the level of $E_5$ activity in strain D318 is less than 1% of the value in induced *T. fusca*. At this time we do not know the reason for the low level of cellulase activity; however, the major species of *celE* mRNA in *E. coli* D318 is shorter than that in *T. fusca* due to the difference in their 3′ ends and probably cannot code for the entire $E_5$ molecule.

## ACKNOWLEDGMENTS

This work was supported by grant DMB-8318432 from the National Science Foundation and grant DE-F602-84ER13233 from the Department of Energy.

## LITERATURE CITED

1. **Ausubel, F. M.** 1984. Regulation of nitrogen fixation genes. Cell **37**:5–6.
2. **Béguin, P., M. Rocancourt, M.-C. Chebrou, and J.-P. Aubert.** 1986. Mapping of mRNA encoding endoglucanase A from *Clostridium thermocellum*. Mol. Gen. Genet. **202**:251–254.
3. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
4. **Breuil, C., and D. J. Kushner.** 1976. Cellulase induction and the use of cellulose as a preferred growth substrate by *Cellvibro gilvus*. Can. J. Microbiol. **22**:1776–1781.
5. **Collmer, A., and D. B. Wilson.** 1983. Cloning and expression of a *Thermomonospora* YX endocellulase gene by *E. coli*. Bio/Technology **1**:594–601.
6. **Csordás-Tóth, É., I. Boros, and P. Venetianer.** 1979. Structure of the promoter region for the *rrnB* gene in *Escherichia coli*. Nucleic Acids Res. **7**:2189–2197.
7. **Eriksson, K.-E.** 1981. Cellulases of fungi. Basic Life Sci. **18**:19–32.
8. **Gentz, R., and H. Bujard.** 1985. Promoters recognized by *Escherichia coli* RNA polymerase selected by function: highly efficient promoters from bacteriophage T5. J. Bacteriol. **164**:70–77.
9. **Gong, C-S., and G. T. Tsao.** 1979. Cellulase and biosynthesis regulation. Annu. Rep. Ferment. Processes **3**:111–140.
10. **Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1987. Regulation, initiation, and termination of the *cenA* and *cex* transcripts of *Cellulomonas finii*. J. Bacteriol. **169**: 646–653.
11. **Hirschman, J., P.-K. Wong, K. Sei, J. Keener, and S. Kustu.** 1985. Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria active *glnA* transcription *in vitro*: evidence that the *ntrA* product is an α factor. Proc. Natl. Acad. Sci. USA **82**: 7525–7529.
12. **Horn, G. T., and R. D. Wells.** 1981. The leftward promoter of bacteriophage λ: structure, biological activity, and influence by adjacent regions. J. Biol. Chem. **256**:2003–2009.
13. **Kauri, T., and D. J. Kushner.** 1985. Role of contact in bacterial degradation of cellulose. FEMS Microbiol. Ecol. **31**:301–306.
14. **Kirkegaard, K., H. Buc, A. Spassky, and J. C. Wang.** 1983. Mapping of single-stranded regions in duplex DNA at the sequence level: single-stranded-specific cytosine methylation in RNA polymerase-promoter complexes. Proc. Natl. Acad. Sci. USA **80**:2544–2548.
15. **Kolankaya, N.** 1980. Control of cellulase synthesis in *Cellumonas flavigena* Hacettepe. Bull. Nat. Sci. Eng. **9**:1–10.
16. **Langsford, M. L., N. R. Gilkers, W. W. Wakarchuk, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1984. The cellulase system of *Cellulomonas fimi*. J. Gen. Microbiol. **130**: 1367–1376.
17. **Lin, E., and D. B. Wilson.** 1987. Regulation of β-1,4-endoglucanase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. **53**:1352–1357.

18. **Lin, E., and D. B. Wilson.** 1988. Identification of a *celE*-binding protein and its potential role in induction of the *celE* gene in *Thermomonospora fusca*. J. Bacteriol. **170**:3843–3846.
19. **Maniatis, T., E. F. Fritsch, and J. C. Sambrook.** 1982. Molecular cloning, p. 109–112. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
21. **Moran, C. P., Jr., N. Lang, C. D. B. Banner, W. G. Haldenwang, and R. Losick.** 1981. Promoter for a developmentally regulated gene in *Bacillus subtilis*. Cell **25**:783–791.
22. **Reitzer, L. J., and B. Magasanik.** 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45**:785–792.
23. **Sheffin, L. G., and O. Kowalski.** 1985. Altered DNA conformations detected by mung bean nuclease occur in promoter and terminator regions of supercoiled pBR322 DNA. Nucleic Acids Res. **13**:6137–6154.
24. **Smith, D. R., and J. M. Calvo.** 1979. Regulation of dihydrofolate reductase synthesis in *Escherichia coli*. Mol. Gen. Genet. **175**: 31–38.
25. **Southern, E. M.** 1975. Gel electrophoresis of restriction fragments. Methods Enzymol. **68**:152–176.
26. **Stewart, B. J., and J. M. Leatherwood.** 1976. Depressed synthesis of cellulase by *Cellulomonas*. J. Bacteriol. **128**:609–615.
27. **Sundaresan, V., J. D. G. Jones, D. W. Ow, and F. M. Ausubel.** 1983. *Klebsiella pneumoniac nifA* product activates the *Rhizobium meliloti* nitrogenase promoter. Nature (London) **301**:728–731.
28. **Vladut-Taylor, M., T. Kauri, and D. J. Kushner.** 1986. Effects of cellulose on growth, enzyme production, and ultrastructure of a *Cellulomonas* species. Arch. Microbiol. **144**:191–195.