JOURNAL OF BACTERIOLOGY, Sept. 1988, p. 3843–3846
0021-9193/88/093843-04$02.00/0

Vol. 170, No. 9

Identification of a *celE*-Binding Protein and Its Potential Role in Induction 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

***Thermomonospora fusca* cellulase E_5 is encoded by the *celE* gene. This gene appears to be regulated at the transcriptional level by both induction and repression, and three putative closely linked promoters have been located by S1 mapping. To study its regulatory mechanism, a gel retardation assay was used to identify a protein in *T. fusca* cell extracts that interacted specifically with the DNA fragment containing the *celE* promoters. It was found that the binding activity appeared only when cellulase synthesis was induced, and it therefore resembled an activator protein involved in cellulase induction. DNase I footprinting identified the target sequence for this protein as a 21-base-pair sequence downstream from the putative *celE* promoters. The level of this protein was measured in two cellulase constitutive mutants, and the results suggest a complex control for *celE* induction.**

Synthesis of cellulases in *Thermomonospora fusca* is coordinately regulated by two independent controls, induction and repression (6). Further study of the *T. fusca celE* gene (coding for cellulase E_5) suggested that it was transcriptionally regulated and identified the locations of three potential *celE* promoters (7).

To understand the molecular mechanism of *celE* gene regulation, it is necessary to identify the *cis*- and *trans*-acting elements mediating regulation. Here we report the identification of a *trans*-acting element for the *celE* gene and the site at which it binds. A gel retardation assay was used to identify and measure the level of a protein binding near the *celE* promoter, which may have a positive effect on cellulase induction. DNase I footprinting was used to find the target sequence for this protein. A preliminary model for the regulation of cellulase induction is proposed based on these studies.

MATERIALS AND METHODS

Preparation of cell extracts. *T. fusca* was grown in minimal medium (6) until mid-log phase. For the gel retardation assay, the culture was harvested, washed two times with and resuspended in lysis buffer, and then lysed with a French press as described (7). The cell lysate was centrifuged at 3,000 × *g* for 5 min, and the protein concentration of the supernatant was measured (7).

For DNase I footprinting, the lysate was prepared as above except that the lysis buffer used contained 100 mM sodium phosphate, 20 mM Tris (pH 7.5), and 2.5 mM EDTA. Nucleic acids were precipitated by adding streptomycin sulfate to 2%, followed by a 15-min incubation at 0°C and then a 5-min centrifugation at 4,000 × *g*. The supernatant was mixed with an equal volume of autoclaved hydroxylapatite suspension (Clarkson Chemical Co.) (equilibrated with lysis buffer at 0°C), incubated at 0°C for 15 min, and centrifuged at 3,000 × *g* for 5 min. Treated supernatant can be stored frozen at −20°C without loss of activity. This treatment was found to remove about 70% of the soluble intracellular protein with a minimal loss of *celE* promoter-binding activity.

Gel retardation assay. Plasmid pD318 was cut with *Sal*I, *Xho*I, and *Pvu*II, and the 3′ ends were labeled with 100 μCi of [α-^{32}P]dCTP (800 Ci/mmol, 10 μCi/μl) by using 20 U of reverse transcriptase (Molecular Genetic Resources) at 37°C for 15 min in 50 mM Tris (pH 8.3)–40 mM KCl–6 mM $MgCl_2$, with 50 μM each dATP, dGTP, and dTTP. The labeled fragments were separated from the nucleotides by gel filtration. Each binding reaction mix contained 5 to 20 μg of labeled DNA fragments and cell extract containing 1 to 3 μg of total protein. The binding reaction was carried out in 15 to 20 μl of 150 mM KCl–10 mM $MgCl_2$–40 mM Tris (pH 8.0)–0.1 mM EDTA–0.1 mM dithiothreitol–10 μg of poly(dG-dC) per ml–12.5% glycerol. After a 10-min incubation at 37 or 55°C, the sample was separated on a 0.7% agarose gel and run in a buffer containing 1.08% Tris, 0.55% boric acid, and 0.093% EDTA (pH 8.3). After electrophoresis, the gel was dried in a 50°C oven overnight on a nonporous support and autoradiographed.

DNase I footprinting. Plasmid pD318 was cut with *Nco*I dephosphorylated with 10 U of bacterial alkaline phosphatase at 65°C for 1 h, phenol extracted, and ethanol precipitated. The 5′ ends were labeled with 80 μCi of [γ-^{32}P]ATP (2,000 Ci/mmol, 10 mCi/ml) and 10 U of polynucleotide kinase (New England Biolab) at 65°C for 1 h. The DNA was ethanol precipitated and then cut with *Sal*I. All of the above reactions were carried out under the standard conditions recommended by the suppliers. The *Sal*I-*Nco*I fragment was isolated by electroblotting onto DEAE-nitrocellulose from an agarose gel, eluted, and precipitated with 1 ml of ethanol at −70°C. The dried pellet (containing ≥10^7 cpm of radioactivity) was dissolved in 45 μl of TE buffer (0.01 M Tris, 1 mM EDTA [pH 7.5]) containing 25 μg of carrier tRNA, and 7 μl of the solution was used for each sequencing reaction (8), while 3.3 μl was used for each footprinting reaction. The protein-DNA binding reaction was carried out as for the gel retardation assay except that the concentration of glycerol was only 9%. After 10 min of incubation at 37°C, the samples were digested with 1 μl of DNase I (0.75 mg/ml, prepared as described in Lin and Wilson [7]) at 25°C for 80 s. The reactions were stopped by adding 2 μl of 2% sodium dodecyl sulfate (SDS) and chilling in dry ice. The samples were then

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

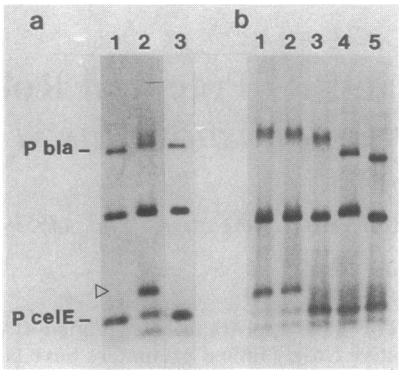

FIG. 1. *celE* promoter-binding activity. pD318 fragments were incubated with an extract of a cellobiose-grown *T. fusca* culture (3 μg of protein), electrophoresed, and autoradiographed as described in the text. The arrowhead shows the position of the shifted band. (a) No protein (lane 1); extract (lane 2); extract incubated for 10 min at 25°C with protease K (lane 3). (b) Extracts were added to the binding reaction mix after being heated for 10 min at 0°C (lane 1), 25°C (lane 2), 55°C (lane 3), 100°C (1 min) (lane 4), or 25°C plus 1% SDS (lane 5).

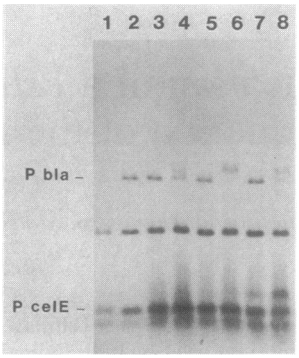

FIG. 2. Level of *celE* promoter-binding activity in *T. fusca* YX. The samples used in the binding assay were: 1 μg of *E. coli* RNA polymerase containing the σ^{70} subunit (lane 1), no protein (lane 2), and cell extracts of *T. fusca* YX grown on glucose (lanes 3 and 4), xylose (lanes 5 and 6), and cellobiose (lanes 7 and 8). The amounts of extract protein were 1 μg for lanes 3, 5, and 7 and 3 μg for lanes 4, 6, and 8.

adjusted to 0.3 M sodium acetate (pH 5.8), extracted with phenol and ether, and then ethanol-precipitated two times. The dried pellets were dissolved in 2 μl of sequencing dye, loaded on an 8% polyacrylamide–7 M urea gel along with the sequencing reaction samples, electrophoresed, and autoradiographed.

RESULTS

***celE* promoter-binding protein.** An extract prepared from cellobiose-grown *T. fusca* YX was used to look for protein binding to the *celE* promoter. In Fig. 1a, it can be seen that some of the fragment containing the *celE* promoter and some of the fragment containing several promoters from pBR322 (including P_{bla}, $P_{RNA\text{-}1}$, P_P, and P_4 [15]) were retarded. The affinity of the binding factors was promoter specific, because the other two fragments lacking promoters were not retarded. The *celE* promoter-containing fragment appeared as a sharp band with reduced mobility after binding, whereas the fragment containing the pBR322 promoters was retarded as a smear which, on close inspection, consisted of a number of weak discrete bands.

The binding activities for both fragments were completely inactivated by incubating the cell extract with *T. fusca* culture supernatant for a few minutes before binding. *T. fusca* culture supernatant contains an active protease, and partially purified *T. fusca* major protease and protease K were even more efficient inactivators of the binding factor(s). The binding factor (s) was also inactivated by incubating the cell extract at 55°C or in SDS, and inactivation increased at higher temperatures (Fig. 1b). The factor binding to the *celE* promoter appeared to be more sensitive to both heating to 55°C and protease (data not shown) than the factor binding to the pBR322 promoters. These results suggest that these promoter-binding factors are different proteins.

The *celE* promoter-binding activity could be significantly stabilized by binding to its target sequence, since preincubating the cell extract with E_5 DNA before the shift to 55°C slowed down the inactivation (data not shown). This may be the result of a change in the protein structure caused by the DNA-protein interaction, but it is also possible that the inactivation was caused by a low level of protease, the target site of which was protected by bound DNA.

It was also found that *Escherichia coli* RNA polymerase containing σ^{70} subunit did not recognize the *celE* promoter sequence under identical conditions (Fig. 2, lane 1) even though it bound to the *bla* promoter fragment. The fact that the *bla* promoter band was not visible in this lane could be because it was not able to enter the gel or because it was shifted to many different locations depending on how many polymerase molecules were bound and no band was intense enough to be visible. The molecular mass of the binding protein has not been determined, but it is probably larger than 30 kilodaltons (kDa) because it was retained by a filtration membrane with a cutoff at 30 kDa.

Level of the *celE* promoter-binding protein. The protein binding to the *celE* promoter could be an RNA polymerase or a regulatory protein. These possibilities might be distinguished by measuring the level of this binding protein under various growth conditions. For this purpose, cell extracts prepared from different cultures were used in the gel retardation assay. The results (Fig. 2) showed that only cellobiose-induced cells contained this binding protein, while the level of the protein binding to the pBR322 promoters was nearly the same in all extracts. The addition of cellobiose to glucose-grown extracts did not cause binding (data not shown). These results suggest that the *celE* promoter-binding protein might act as an activator for the *celE* gene.

If this hypothesis is correct, the level of the *celE* promoter-binding protein might increase in parallel with the cellulase level in constitutive mutants. This was tested by measuring the binding protein levels in the cellulase constitutive mutants CC-1 and CC-2 (Fig. 3). These results showed that the binding protein level indeed increased as expected in the partially constitutive strain CC-1 so that it became detectable in the absence of the inducer cellobiose. However, it was undetectable in the fully constitutive strain CC-2 in cells grown in either the presence or absence of cellobiose. There is no simple explanation for this observation, even though a negative control model for cellulase induction is likely (see Discussion for details).

DNA target sequence for the inducer protein. DNase I footprinting is the most widely used method for detecting the target sequence of a regulatory protein, and it was used to identify the target sequence of the *celE*-binding protein.

Partially purified cell extracts of *T. fusca* YX were tested for the *celE*-binding protein by the gel retardation assay, and

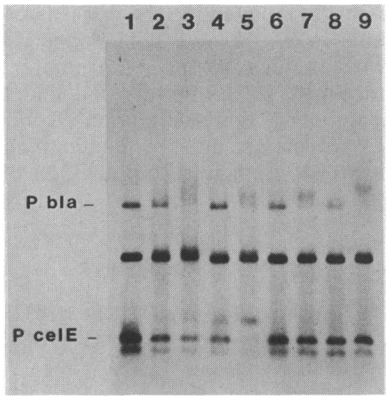

FIG. 3. Level of *celE* promoter-binding activity in *T. fusca* CC-1 and CC-2. The samples used in the binding assay were: no protein (lane 1), extracts of the CC-1 strain grown on glucose (lanes 2 and 3) or cellobiose (lanes 4 and 5), and extracts of the CC-2 strain grown on glucose (lanes 6 and 7) or cellobiose (lanes 8 and 9). The amounts of total cell extract protein were 1.5 μg for lanes 2, 4, 6, and 8 and 3 μg for lanes 3, 5, 7, and 9.

no loss in binding activity occurred during purification (data not shown). Attempts at further purification with conventional methods always resulted in a loss of binding activity. The treated cell extract was used in a DNase I footprinting assay as described in Materials and Methods. The results (Fig. 4) showed that a 21-base-pair (bp) DNA sequence was

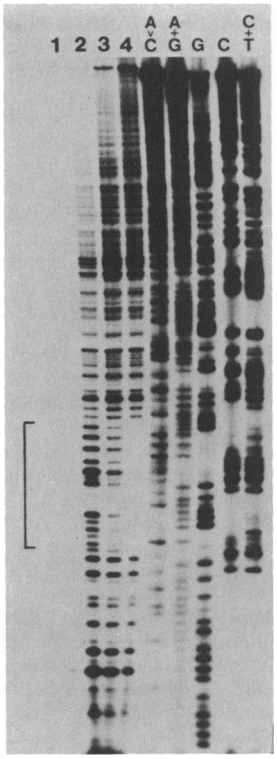

FIG. 4. DNase I footprinting of the target sequence of the *celE* promoter-binding protein. Footprinting was carried out as described in the text. Lanes: No protein (lane 1), treated cell extracts of *T. fusca* YX grown on glucose (lane 2), cellobiose (lanes 3 and 4). The amount of total cell extract protein was 25 μg (lanes 2 and 4) or 15 μg (lane 3). The protected region is indicated by the vertical bar.

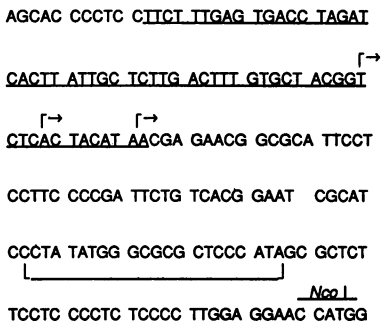

FIG. 5. Binding sequence for the *celE* gene activator protein. The DNA sequence protected by the *celE* promoter-binding protein is indicated by the box. The 5′ ends of the *celE* mRNAs are indicated by arrows.

specifically protected by the cell extract from the induced culture. The extract from noninduced cells at the same protein level did not specifically protect any DNA sequence in the *Sal*I-*Nco*I fragment under identical conditions, even though it reduced the rate of DNase I digestion of all sites, as did the induced sample. This indicates that the binding protein detected by the footprinting assay was the same one detected by the gel retardation assay. The sequence of the protected site is shown in Fig. 5.

DISCUSSION

A *celE* promoter-binding activity was detected by a DNA gel retardation assay, and its sensitivity to heat, SDS, and protease suggests that this binding activity is associated with a protein. *T. fusca* extracts also caused retardation of a segment of pBR322 DNA containing several known *E. coli* promoters. However, this activity differed from the *celE*-binding protein in its sensitivity to protease and heating and in its regulation, so that different proteins appear to be shifting each DNA fragment.

The level of the promoter-binding activity depended on the growth conditions. It was only detected in cells induced for cellulase, suggesting that it is an activator protein involved in cellulase induction. Whether such induction-dependent binding also occurs in vivo and, if so, whether it is caused by de novo protein synthesis remain to be tested.

Some gene activator proteins are known to require cofactors for their function, i.e., the *E. coli araC* protein (3) and the CAP protein (11), but we did not find any effect of cellobiose, the probable inducer of cellulase synthesis in *T. fusca*, on binding.

The levels of this binding activity were also measured in constitutive mutants. The result obtained with the CC-1 strain is consistent with its proposed function in induction, but the result with the CC-2 strain suggests a more complex mechanism for induction. A speculative model is that the *celE* gene is regulated by negative control and the binding protein inactivates the repressor. In the CC-2 mutant, both the negative control and the positive control would have to have been inactivated to give constitutive synthesis. The only simple explanation for this is that the binding protein is related to the unknown negative control factor, either at the protein level (e.g., like the *E. coli araC* protein [5]) or at the level of regulation of its gene.

The unusual behavior of strain CC-2 strain raises the possibility that it is not derived from *T. fusca* YX. However, this was ruled out by the following observations. All the

major cellulases are present in the CC-2 strain in a ratio similar to that in the wild-type strain as shown by immunoinhibition (6), and the *celE* gene is present in the CC-2 mutant with a restriction map indistinguishable from that in the wild-type strain, according to genomic Southern hybridization data (data not shown).

The results of DNase I footprinting show that the inducer protein specifically protected a 21-bp DNA sequence located downstream from the putative promoters (Fig. 5). This sequence is not homologous to any known procaryotic regulatory sequence. Although a downstream location for a regulatory protein recognition sequence has been found in both the *E. coli lac* (10) and *gal* (4) operons, most activator protein-binding sites (e.g., cyclic AMP receptor protein and λ *c*I and *c*II) are found to be upstream from the transcription initiation site (9).

The downstream location raises the possibility that this protein might exert its function by interacting with *celE* mRNA. This possibility is unlikely, because the only kind of protein in procaryotes that is known to bind to both DNA and RNA (double stranded or single stranded) is the DNA-binding protein II family (2). DNA-binding protein II is quite different from regulatory proteins because it interacts with DNA by non-sequence-specific electrostatic binding, which is required by its proposed function in maintaining genome structure, and binding is totally eliminated at high ionic strength. Another possibility is that the *cel* gene is regulated by attenuation and the binding protein blocks premature termination. However, there is no open reading frame in the 5′ leader sequence.

Neither the gel retardation assay nor the footprinting assay showed evidence for binding of RNA polymerase at the putative promoter or binding of the proposed negative control factor in the extract of noninduced cells at any location tested so far. This allows us specifically to detect the inducer protein, but it raises the question of why the other expected binding proteins were not observed. It is highly unlikely that RNA polymerase, which should be more abundant than the inducer protein, was completely inactivated in cell extracts.

The inability to detect the other expected promoter-binding proteins may be caused by the relatively high ionic strength of the binding buffer. It was shown that the potential of formation of a protein-nucleic acid complex depends on the ionic strength, as predicted by thermodynamic theories (1, 12, 14, 16, 17). Under high-salt conditions, the sequence-nonspecific binding may completely vanish while sequence-specific binding is also weakened, as the electrostatic interactions do not contribute to the binding energy in the presence of salt. This could explain our observations if, and only if, the inducer protein can form a more stable complex than the other *celE* promoter-binding proteins.

Alternatively, this phenomenon may result from a lack of some unknown essential factors in the cell extract or the DNA molecule tested, such as DNA superhelicity, which has been discussed (7).

Although it is not known what kind of RNA polymerase is responsible for transcribing the *celE* gene, this enzyme cannot be similar to the σ^{70} enzyme in *E. coli*, as the DNA sequence of the putative promoters and the region further upstream are not homologous to the consensus sequence of the σ^{70} promoter in *E. coli* (9, 13). The observation that the σ^{70} enzyme fails to recognize the *celE* promoter even though it recognizes the pBR322 promoters under identical conditions seems to confirm this assumption. Nonetheless, this unknown RNA polymerase must be related to one of those in *E. coli* because of the efficient transcription of the cloned *celE* gene in *E. coli* (7).

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. **Frankel, A. D., G. K. Ackers, and H. O. Smith.** 1985. Measurement of DNA-protein equilibria using gel chromatography. Biochemistry **24**:3049–3054.
2. **Geider, K., and H. Hoffmann-Berling.** 1981. Proteins controlling the helical structure of DNA. Annu. Rev. Biochem. **50**:233–260.
3. **Hahn, S., W. Hendrickson, and R. Schleif.** 1986. Transcription of *Escherichia coli ara* in vitro. J. Mol. Biol. **188**:355–367.
4. **Irani, M. H., L. Orosz, and S. Adhya.** 1983. A control element within a structural gene: the *gal* operon of *Escherichia coli*. Cell **32**:783–788.
5. **Lee, N.** 1978. Molecular aspects of *ara* regulation, p. 389–410. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. **Lin, E., and D. B. Wilson.** 1987. Regulation of β-1,4-endogluconase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. **53**:1352–1357.
7. **Lin, E., and D. B. Wilson.** 1988. Transcription of the *celE* gene in *Thermomonospora fusca*. J. Bacteriol. **170**:3838–3842.
8. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–552.
9. **McClure, W. R.** 1985. Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. **54**:171–204.
10. **Miller, J. H.** 1978. The *lac*I gene: its role in *lac* operon control and its use as a genetics system, p. 31–88. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Pabo, C., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
12. **Revzin, A., and P. H. von Hippel.** 1977. Direct measurement of association constants for the binding of *Escherichia coli lac* repressor to non-operator DNA. Biochemistry **16**:4769–4776.
13. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
14. **Rouviere-Yaniv, J., M. Yaniv, and J.-E. Germond.** 1979. *E. coli* DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. Cell **17**:265–274.
15. **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.
16. **Tanaka, I., K. Appelt, J. Dijk, S. W. White, and K. S. Wilson.** 1984. 3-Å resolution structure of a protein with histone-like properties in prokaryotes. Nature (London) **310**:376–381.
17. **Winter, R. B., and P. H. von Hippel.** 1981. Diffusion driven mechanisms of protein translocations on nucleic acids. Biochemistry **20**:6948–6960.