Volume 17 Number 3 1989 Nucleic Acids Research

Sigma-G RNA polymerase controls forespore-specific expression of the glucose dehydrogenase operon in *Bacillus subtilis*

Yoshihiro Nakatani*, Wayne L.Nicholson[1], Klaus-Dieter Neitzke, Peter Setlow[1] and Ernst Freese

Laboratory of Molecular Biology, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892 and [1]Department of Biochemistry, University of Connecticut Health Center, Farmington, CT 06032, USA

Received November 1, 1988; Revised and Accepted January 10, 1989

ABSTRACT

The gene encoding glucose dehydrogenase (*gdh*) is part of an operon whose expression is transcriptionally activated specifically in the developing forespore of *Bacillus subtilis* at stage III of sporulation. The *in vivo* startpoint of *gdh* transcription was determined using primer extension analysis. Deletion mapping and site-specific mutagenesis experiments indicated that the region responsible for regulated expression of *gdh in vivo* was limited to the "-35" and "-10" regions preceding the transcriptional start site. RNA polymerase containing σ^G ($E\sigma^G$) transcribed *gdh in vitro* with a start site identical to that found *in vivo*, and transcription of *gdh* by $E\sigma^G$ *in vitro* also did not require any specific sequences upstream from "-35" region. These results suggest that the appearance of $E\sigma^G$ in the forespore at stage III of sporulation is sufficient to cause temporal and compartment-specific expression of the *gdh* operon.

INTRODUCTION

When *Bacillus subtilis* cells initiate sporulation, they partition asymmetrically into a forespore compartment which develops into the endospore, and a mother cell compartment which lyses after maturation of the endospore. During sporulation, the expression of a number of genes is temporally regulated, and some are specific to one or the other compartment (For reviews, see ref. 1, 2). One such gene is *gdh*, the product of which, glucose dehydrogenase (GlcDH), apparently is synthesized only in the forespore compartment (3). The *gdh* gene has been cloned (4), and shown to be part of an operon which we call the *gdh* operon (5, 6; this study). The *gdh* operon consists of two open-reading frames, the first of which has an unknown function (called *orfX*), and the second of which is *gdh* (Figure 1). Site-specific mutagenesis studies in the control region (this study; 6), Nothern blot analysis (5), and transcription startpoint analysis (this study; 6) indicate that *orfX* and *gdh* are transcribed as a polycistronic message whose synthesis is regulated by an upstream control region.

Forespore-specific expression of the *gdh* operon could result from selective

repression of the operon in both the vegetative cell and mother cell. Such a mechanism is tenable because the developing forespore is separated from the mother cell by two membranes of opposite polarity, enabling each compartment to maintain a different composition of metabolites. Alternatively, expression of the *gdh* operon could be controlled by a transcription factor that determines the promoter specificity of RNA polymerase and is produced or activated only in the forespore. In this communication, we demonstrate that the nucleotides in the "-35" and "-10" regions preceding the transcription startpoint of the *gdh* operon determine its sporulation stage- and compartment-specific expression. Moreover, we show that this transcription *in vitro* depends on the newly discovered RNA polymerase, $E\sigma^G$ (RNA polymerase containing σ^G) (7,8). Evidence that *gdh* expression *in vivo* depends on $E\sigma^G$ will be published elsewhere.

MATERIALS AND METHODS

B.subtilis strains and growth conditions

B. subtilis strains used in this study are listed in Table 1. For assay of β-galactosidase or GlcDH, strains were grown in nutrient sporulation medium (NSMP)(11). For isolation of RNA polymerase from sporulating or vegetative cells, strains were grown in 2xSG (12) or 2xYT medium (13) respectively, with appropriate antibiotics: chloramphenicol (Cm), 5 μg/ml; kanamycin (Km), 10 μg/ml. Induction of the gene encoding σ^G [*sig*G or *spo*IIIG (8)] carried on plasmid pDG298 was by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. All cultures were incubated at 37 C with vigorous aeration.

Transformation of *B. subtilis*

The protoplast method (14) was used to introduce ligated DNA into *B. subtilis* strain 1A46. Otherwise, plasmid transformation was carried out using competent cells as described previously (10, 15).

Plasmid constructions

All molecular biology techniques were carried out essentially as described by Maniatis et al. (16), unless otherwise noted. Plasmid pEF1 (4, 17), which carries the entire *gdh* operon cloned as a 4 kb *Eco*RI fragment in pBR322, was the source of *gdh* DNA for all constructions outlined below. Plasmid pTV8 was provided by P. Youngman (18). Plasmid pGEM-4Z was obtained from Promega-Biotec; plasmid pTZ18 was from Pharmacia. Plasmid pDG298 (7) consists of the coding sequence of the *sig*G gene under the control of an IPTG-inducible promoter from plasmid pDG148 (19). Plasmid pJF751 (20) was provided by J. A. Hoch.

To integrate the *gdh-lacZ* fusion into the *gdh* locus of the *B. subtilis* chromosome, a 2.3 kb *Dra*I fragment (Fig. 1) was isolated from pEF1, and a 12-mer *Bam*HI linker (New

Table 1. *B. subtilis* strains used in this study.

Strain	Genotype or phenotype	Source(reference)
60015	*met*C, t*rp*C2	strain SB26 of Nester
62558	*rec*E4, *thr*-5, *trp*C2	1A46 of BGSC[a]
62542	*trp*C2, SPβdel2::Tn917	CU 2530 of Youngman,P.(9)
62546	*trp*C2, SPβdel2::Tn917, *gdh-lacZ* integrated at *gdh*	This work
62548	*trp*C2, SPβdel2::Tn917::*gdh-lacZ*	This work
62549	*trp*C2, SPβdel2::Tn917::promoter-less *lacZ*	This work
PS435	*trp*C2, *ssp*E-*lacZ*, Cm^R	Mason, J. M. (10)
PS614	*trp*C2, *spo*IIIC94, *ssp*E-*lacZ*, Cm^R	Sun, D. (7)
PS749	*phe*A1, *trp*C2, *spo*0A12, *ssp*A-*lacZ*, Cm^R, [pDG298; *spac-sig*G, Km^R][b]	This work
PS791	*leu, his, met, spo*0HΔ, *ssp*E-*lacZ*, Cm^R,[pDG298; *spac-sig*G, Km^R]	This work

[a]BGSC, Bacillus Genetic Stock Center
[b][ ], denote plasmids. These strains were constructed by transformation of the appropriate strains (10) to kanamycin resistant with plasmid pDG298 (7).

England Biolabs) was attached to both ends. After *Bam*HI digestion, the 2.3 kb fragment was subcloned into the *Bam*HI site of the promoter-less *E. coli lacZ*-containing plasmid pJF751 so that the first part of the *gdh* operon was fused in frame with *lacZ*. This plasmid, designated pYN151, was transformed into *B. subtilis* CU2530, and transformants were selected on plates containing Cm (15).

The *gdh-lacZ* fusion was introduced into the *B. subtilis* prophage SPβ locus by a modification of the method described by Zuber and Losick (21). A promoter-less *E. coli lacZ*-containing integration plasmid, pYN301, was constructed from pJF751 and plasmid pTV8 after modification of some restriction sites (Fig. 2A). Plasmid pYN301 was used to construct plasmids for integration of the *gdh-lacZ* fusion into the SPβ locus. One such derived plasmid was pYN331 in which the 1.2 kb *Sac*II-*Bam*HI fragment of pYN151 (derived from the right *Sac*II-*Dra*I fragment in Fig. 1) was

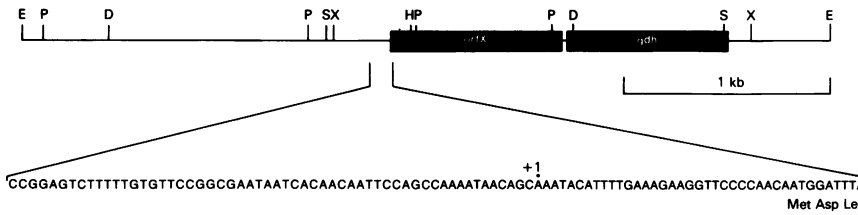


Figure 1. Restriction map and sequence of the upstream region of the *gdh* operon. Abbreviations: D, *Dra*I; E, *Eco*RI; H, *Hae*III (not all *Hae*III sites are shown); P, *Pvu*I; S, *Sac*I; X, *Xba*I.

incorporated into the *Sac*II-*Bam*HI area of pYN301. This plasmid was linearized with *Eco*RI, transformed into CU2530, and transformants were selected as described above.

To facilitate construction of deletion mutants, the 1.3 kb *Sac*II-*Pvu*II fragment of pYN331 (Fig. 2A), which contains the *gdh* promoter region, was subcloned into plasmid pBR322. The *Sca*I site of pBR322 was converted to a *Sac*II site using a synthetic linker, and the resulting plasmid pYN232 was used as follows. Plasmid pYN331 was digested with *Pvu*II (a *Pvu*II site is located 80 bp downstream from the junction of *gdh* and *lacZ*; see Fig. 2A) and *Pst*I linkers were attached to the ends. After digesting with *Pst*I and *Sac*II, a 1.3 kb *Pst*I-*Sac*II fragment was isolated and subcloned into *Pst*I-*Sac*II cut pYN232. The resulting plasmid pYN293 was used to generate the deletions which were subsequently reinserted into pYN301. To construct upstream deletion mutants, pYN293 was linearized with *Sac*II, digested with BAL-31 exonuclease, and ligated to a *Sac*II linker. After cutting with *Sac*II and *Bam*HI, the 0.9 to 1.0 kb fragments were isolated, ligated into *Sac*II-*Bam*HI-cleaved pYN301, and the resulting DNA was transformed into competent *E. coli* strain HB101. Upstream deletion mutants were subcloned as *Sal*I-*Bam*HI fragments into *Sal*I-*Bam*HI-cleaved pGEM-4Z for quantitative *in vitro* transcription analysis. To construct downstream deletion mutants, a 0.7 kb *Pvu*I fragment, which is located within the *orfX* gene of pEF1 (Fig. 1), was transiently replaced with an unrelated 0.5 kb *Pvu*I fragment which contained one *Kpn*I site 5 bp on the right of the left *Pvu*I site. The resulting plasmid was linearized at the unique *Kpn*I site, digested with BAL-31, and ligated to a *Pvu*I linker. After cleavage with *Pvu*I, DNA was recircularized by ligation at a low concentration (50 ng/ml), and transformed into *E. coli* HB101. Plasmids isolated from the mixture of transformants were digested with *Sac*II and *Bam*HI. The 0.3 to 0.4 kb fragments were isolated, ligated into *Sac*II-*Bam*HI-cut pYN301, and transformed into *E. coli* HB101. The endpoints of all deletions were determined by dideoxy-DNA-sequencing reactions (22). Upstream and downstream deletion plasmids were

linearized with *Eco*RI and transformed into *B. subtilis* CU2530. All mutants contained the ribosome binding site for *gdh-lacZ* fusion because that site is located downstream from the rightmost *Pvu*I site (Fig. 1).

RNA isolation and reverse transcriptase mapping

Nucleic acid was isolated from sporulating cells by the method of Penn et al. (23). Nucleic acid extracts were treated with RNase-free DNase (50 U/ ml, Promega-Biotec) in 100 mM PIPES-NaOH buffer (pH 7.5) containing 10 mM NaCl, 6 mM $MgCl_2$, and the RNase inhibitor RNasin (1000 U/ ml, Promega-Biotec) at 37 C for 30 min. RNA was collected by ethanol precipitation after phenol-chloroform extraction.

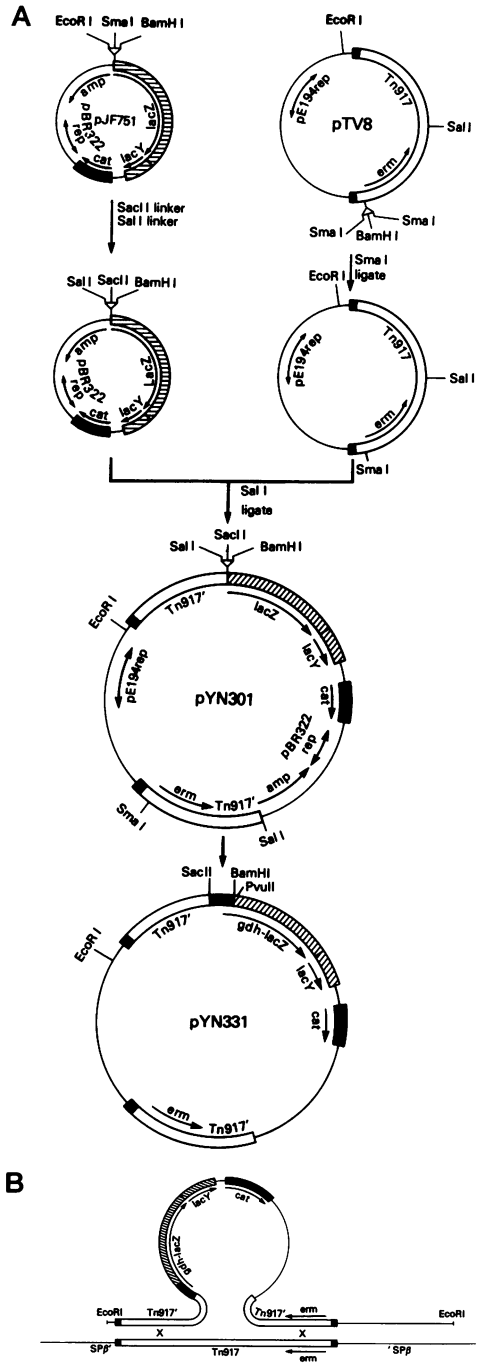
Two antisense-oligonucleotides, oligo-1 (positions +93 to +70: 5'-CACATTGAAGA-GAACAATGCTCCC-3') and oligo-2 (positions +353 to +330: 5' CGGAACACGATC-ACACCGAACAGC-3') were end-labelled using T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP. 1 x 10^6 cpm (100 fmol) of labelled oligonucleotide was hybridized with 5 µg of RNA in 10 µl of hybridization buffer (10 mM Tris-HCl [pH7.9], 1 mM EDTA, 0.25 M KCl) at 55 C for 60 min. After hybridization, the mixture was immediately added to 25 µl of the primer extension mixture (10 mM Tris-HCl [pH 8.7], 10 mM $MgCl_2$, 5 µM DTT, 0.4 mM dNTPs, RNasin [1000 U/ml, Promega-Biotec.], avian myeloblastosis virus reverse transcriptase [2000 U/ml, Promega-Biotec.]). The extension reaction was carried out at 42 C for 45 min. After ethanol precipitation, nucleic acid was dissolved in formamide dye mix and analyzed on a 6% polyacrylamide sequencing gel (24). The molecular weight marker was prepared by dideoxy-sequencing reactions on a single-stranded DNA of coliphage M13mp19 containing the sense strand of the 2.1 kb *Xba*I fragment of plasmid pEF1 (Fig. 1), using the same oligonucleotide as primer.

Isolation of RNA polymerase

RNA polymerase was isolated from sporulating cells of strain PS614 grown in 2xSG medium as described previously (7). Eσ^G from midexponential-phase *B. subtilis* strains PS749 and PS791 was isolated in a similar manner, except that cells grown in 2xYT medium were harvested 2 hr after induction with 1 mM IPTG (7).

***In vitro* transcription**

Runoff transcription of the cloned *gdh* gene was performed essentially as described by LeGrice et al. (25), with the following modifications: 2-4 µCi [α-^{32}P] UTP (800 Ci/mmol) and 2 µg (0.4-0.8 pmol) template DNA were used, and all reactions were started by addition of RNA polymerase. For dinucleotide priming reactions (26), nucleoside triphosphates were at 3 µM, and the dinucleotide was at 250 µM. Transcription reactions using RNA polymerase isolated from sporulating cells were labelled for 10 min at 37 C and then chased for 5 min at 37 C with 250 µM (final concentration) unlabelled UTP. Alternatively, transcription reactions using either RNA


A
EcoR I Sma I BamH I
amp
pBR322 rep
pJF751
lacZ
lacY
cat
SacI I linker
Sal I linker
Sal I SacI I BamH I
EcoR I
pE194rep
Tn917
pTV8
Sal I
erm
Sma I
BamH I
Sma I ligate
Sal I ligate
SacI I
Sal I
BamH I
Tn917'
pYN301
erm
Tn917'
BamH I
PvuII
gdh-lacZ
pYN331
B
gdh-lacZ
lacY
cat
EcoRI
Tn917'
X
SPβ'
Tn917
'SPβ

polymerase isolated from sporulating cells or $E\sigma^G$ isolated from IPTG-induced vegetative cells carrying pDG298 were carried out as follows. Reactions were incubated at 37 C for 1 or 2 min, and then heparin was added to 500 μg/ml to prevent reinitiation and to inhibit endogenous nuclease activity. Incubation with heparin for 2 or 4 min was followed by a 2 min chase with unlabelled UTP as described above, and reactions were terminated by ethanol precipitation. Air-dried samples were dissolved in gel-loading buffer (80% formamide, 1 mM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue), heated to 90 C for 1 min, quick-chilled on ice, and analyzed by autoradiography after electrophoresis through 6% polyacrylamide sequencing gels. Quantitation of specific transcripts was performed by their excision from the gels and counting in a scintillation counter.

β-galactosidase assay

For quantitative β-galactosidase assays cells, were harvested from 0.5 ml of culture, washed with cold 50 mM Tris-HCl buffer (pH 8.0), the pellets frozen in dry ice/ethanol and stored at -80 C overnight. Cells were resuspended in 0.9 ml of assay buffer containing 60 mM K_2HPO_4, 40 mM KH_2PO_4 100 mM NaCl, 1 mM PMSF, and 1 μg DNase I per ml. Either 0.1 ml of lysozyme (10 mg/ml in water) or water was added to the suspensions, which were incubated for 10 min at 37 C. β-galactosidase activity was measured using MUG (4-methylumbelliferyl-β-D-galactoside, 67 μg/ml) as a substrate (27). In all our quantitative assays with MUG, we subtracted a background of β-galactosidase activity (e.g., 0 MUG units at T_0 and 20 MUG units at T_7 by the lysozyme-MUG assay) found in the *B. subtilis* strain 62549 which contained a promoter-less *lacZ* gene in plasmid pYN301 integrated at the SPβ locus. One MUG unit was defined as one picomole of MUG hydrolyzed per min per A_{600}, where one A_{600} unit is that amount of bacteria which produces an OD_{600} of 1.0 if suspended in 1 ml. For the plate assay of β-galactosidase, cells were streaked onto NSMP plates and grown for 24 hr at 37 C. β-galactosidase was detected by spraying MUG on the plates as described by Youngman (27).

GlcDH assay

Cells harvested from 5 ml of culture were washed with 100 mM HEPES buffer (pH 6.5), the resulting pellets frozen in dry ice/ethanol and kept at -80 C overnight. Cells

Figure 2. Construction and integration of *gdh-lacZ* fusion plasmids. (A). Construction of plasmids pYN301 and pYN331. The *Eco*RI and *Sma*I sites of pJF751 were converted into *Sal*I and *Sac*II sites respectively, using synthetic linkers. The *Bam*HI site of pTV8 was deleted by digesting with *Sma*I and self-ligating. These two plasmids were combined at the *Sal*I sites to obtain pYN301. The 1.2 kb *Sac*II-*Dra*I fragment containing the *gdh* promoter was subcloned into pYN301at the *Sac*II/*Bam*HI sites to create pYN331 (for details, see Materials and Methods). (B). Likely mode for integration of the *gdh-lacZ* translational fusion into the *B. subtilis* chromosome.

were resuspended in 2 ml of 200 mM HEPES buffer (pH 6.5) containing 20% glycerol, and were disrupted by sonication on ice. Supernatants obtained by centrifugation at 14000 x g for 10 min at 4 C were assayed for GlcDH activity at 37 C as previously described (3). Protein was measured by the method of Bradford (28) using bovine serum albumin as a standard. One unit of GlcDH was defined as one nanomole of $NADP^+$ reduced per mg of protein per min.

RESULTS

Mapping of the *in vivo* transcription startpoint of the *gdh* operon

A ^{32}P end-labelled primer (oligo-1; see Materials and Methods) representing a part of the *orfX* sequence was hybridized with RNA extracted from *B. subtilis* at various times during development and extended by reverse transcriptase. Primer extension products were detected using RNA isolated at T_4 (Fig. 3, lane 7), but not at T_0 or T_2 (Fig. 3, lanes 5 and 6). This time of appearance of *gdh*-specific mRNA coincided with that of GlcDH activity (data not shown), in agreement with an earlier report that GlcDH synthesis is regulated at the transcriptional level (17). Two adjacent bands were detected by primer extension analysis with T_4 RNA (Fig. 3, lane 7), each ending in adjacent adenine residues. We assume that the upper major band represents the transcription startpoint.

The oligonucleotide used (oligo-1) started 93 nt from the transcription start site. To eliminate the possibility of another transcription start site within the *orfX* gene, we also used a 24-mer oligonucleotide (oligo-2) starting at position +353. A primer extension product of about 350 nt length was detected using T_4 RNA, but no additional band was found, that could represent a start site downstream from this area (data not shown). This result shows that there can not be another start site of RNA synthesis, at least within the 0.3 kb downstream from the transcription start site.

Construction of a translational fusion between *gdh* and *lacZ*

To study the regulation of *gdh* expression we constructed a translational fusion between *gdh* and *lacZ*, and introduced it into the *B. subtilis* prophage SPβ (see Materials and Methods). The fusion contained the 1.2 kb fragment of the *gdh* operon extending from the left *Sac*II site to the right *Dra*I site (Fig. 1). Plasmid pYN331 containing this *gdh-lacZ* translational fusion was linearized at the unique EcoRI site and introduced into *B. subtilis* strain CU2530, which contains a Tn917 insertion at the SPβ locus. Southern blot analysis of chromosomal DNA isolated from this strain showed that the *gdh-lacZ* fusion was integrated at the SPβ locus rather than at the *gdh* locus and that integration was consistent with the orientation presented schematically in Figure 2B (data not shown).

When evaluating the results, it should be noted that GlcDH normally is located

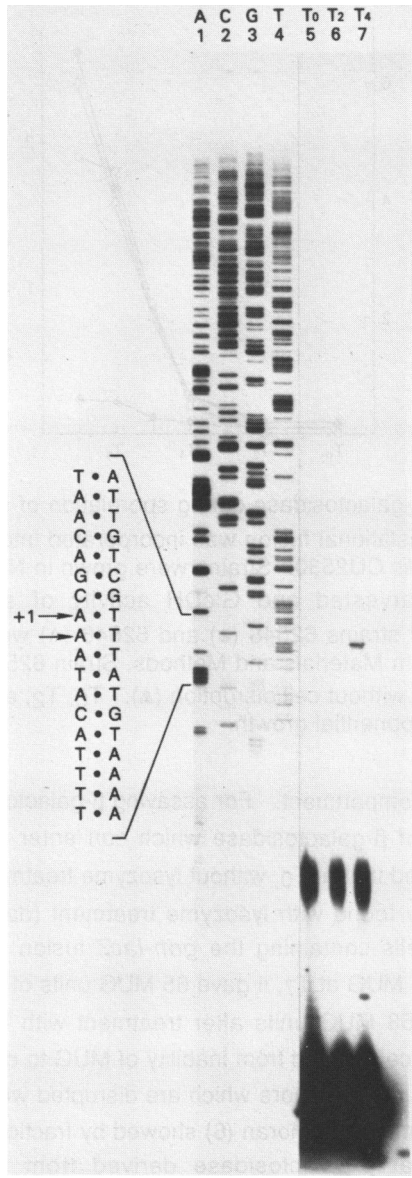


Figure 3. Reverse transcriptase mapping of the *in vivo* transcription startpoint of the *gdh* operon. RNA isolated from sporulating cells of strain 60015 at T_0 (lane 5), T_2 (lane 6), or T_4 (lane 7) was hybridized with the ^{32}P end-labelled 24-mer primer, the hybrids extended with reverse transcriptase and analyzed by gel electrophoresis (as described in Materials and Methods). Molecular weight markers (lanes 1-4) were obtained from dideoxynucleotide-sequencing reactions (as described in Materials and Methods).

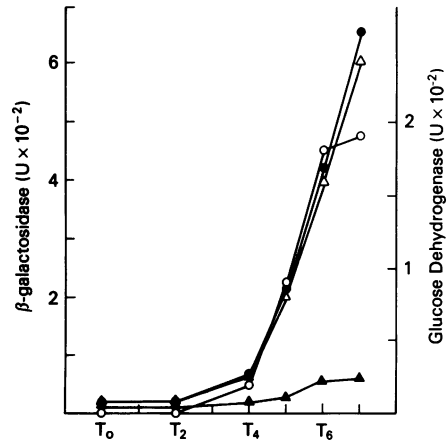


Figure 4. Synthesis of β-galactosidase during sporulation of strains carrying *gdh-lacZ* fusions. A *gdh-lacZ* translational fusion was incorporated into either the SPβ locus or the *gdh* locus of *B. subtilis* CU2530. Strains were grown in NSMP medium at 37 C; at intervals cells were harvested and GlcDH activity of strain 62548 (o) or β-galactosidase activitiy of strains 62548 (●) and 62546 (Δ) were determined after cell disruption as mentioned in Materials and Methods. Strain 62548 was also assayed for β-galactosidase by MUG without cell disruption (▲). T_0, T_2, etc. denote the number of hours after the end of exponential growth.

mainly in the forespore compartment. For assaying β-galactosidase, we used MUG, a hydrophobic substrate of β-galactosidase which can enter cells without any further treatment (27): we showed that at T_0, without lysozyme treatment, a *spo*VG-*lacZ* fusion gave 80% of the activity found with lysozyme treatment (data not shown). When a suspension of whole cells containing the *gdh-lacZ* fusion in the SPβ locus (strain 62548), was treated with MUG at T_7, it gave 65 MUG units of β-galactosidase activity. The same cells gave 658 MUG units after treatment with lysozyme (Fig. 4). We assume that this difference resulted from inability of MUG to diffuse rapidly through the envelope components of the forespore which are disrupted when cells are treated with lysozyme. Recently, Rather and Moran (6) showed by fractionation of mother cell and forespore materials that β-galactosidase derived from the *gdh-lacZ* fusion is predominantly detected in the forespore extract. Therefore, we believe that the large activity difference observed by the MUG assay between lysozyme treated and untreated cells provides a rapid method to judge compartmentalization of *lacZ* expression.

Essentially identical kinetics of β-galactosidase synthesis were observed during

sporulation of strains which carried a *gdh-lacZ* fusion integrated at the *gdh* locus (62546; this strain was constructed using the integration vector pYN151) or the SPβ locus (62548). Southern blot analysis of chromosomal DNA from strain 62546 indicated that integration of the *gdh-lacZ* fusion into the *gdh* locus occurred by a Campbell-type recombination event (data not shown); therefore, all *gdh* operon sequences upstream from the *gdh-lacZ* fusion should be present. The fact that the timing and extent of β-galactosidase production were the same whether the *gdh-lacZ* fusion was inserted into the SPβ locus or the *gdh* operon demonstrated that DNA sequences upstream from the left *Sac*II site or downstream from the right *Dra*I site (Fig. 1) were not needed to control *gdh* operon expression.

An upstream 35 bp region controls expression of the *gdh* operon

To determine the precise endpoints of the sequences which control *gdh* expression, a number of upstream (designated UP) and downstream (designated DN) deletion mutants of the *gdh-lacZ* fusion were constructed using the exonuclease BAL-31 (Fig. 5A; see Materials and Methods). For the UP mutants, the *Sac*II linker 5'-CCGCGGC-3' was attached to the endpoints, while for DN mutants, the PvuI linker 5'-GCGATCG-3' was attached to the endpoints. All mutants contained the ribosome binding site for *gdh-lacZ*, because that site is located downstream from the right *Pvu*I site (Fig. 5A). The mutants were analyzed by spraying MUG on colonies of sporulating cells. Upstream deletions extending to -35 bp from the transcription startpoint (UP-35) retained the capacity to produce β-galactosidase, whereas that expression was abolished in mutant UP-18 (Fig. 5A). Analysis of the downstream deletion mutants showed that DN-3 still expressed β-galactosidase activity whereas DN-15 showed no activity (Fig. 5A). [In the DN-3 mutant, the *Pvu*I linker supplies the GC that has been removed at the -3 and -2 positions.] Quantitative assays of cells grown in sporulation medium and treated with lysozyme indicated that the β-galactosidase activity of all mutants increased after T_4 (Figs. 5B and 5C), which indicates that the mutants were still under the same sporulation control as the parent strain. The quantitative differences of the active strains suggest that the deleted regions may have a minor effect on promoter activity. The results indicate that the regions upstream from -35 and between -1 and +828 are not necessary for sporulation-specific expression of *gdh*.

Interestingly, strain DN-3 had 2.5 times higher activity than its parent (Fig. 5C). Similar to the decreased activity observed in some of the other mutants, the increased activity could result from an effect of surrounding nucleotides on promoter activity. Alternatively, the change in the initiation area from the seqence AACAGCAAAT (underlined adenine corresponds to the original startpoint) into AACAGCGATCG may account for more effective utilization by RNA polymerase. The third possibility of an

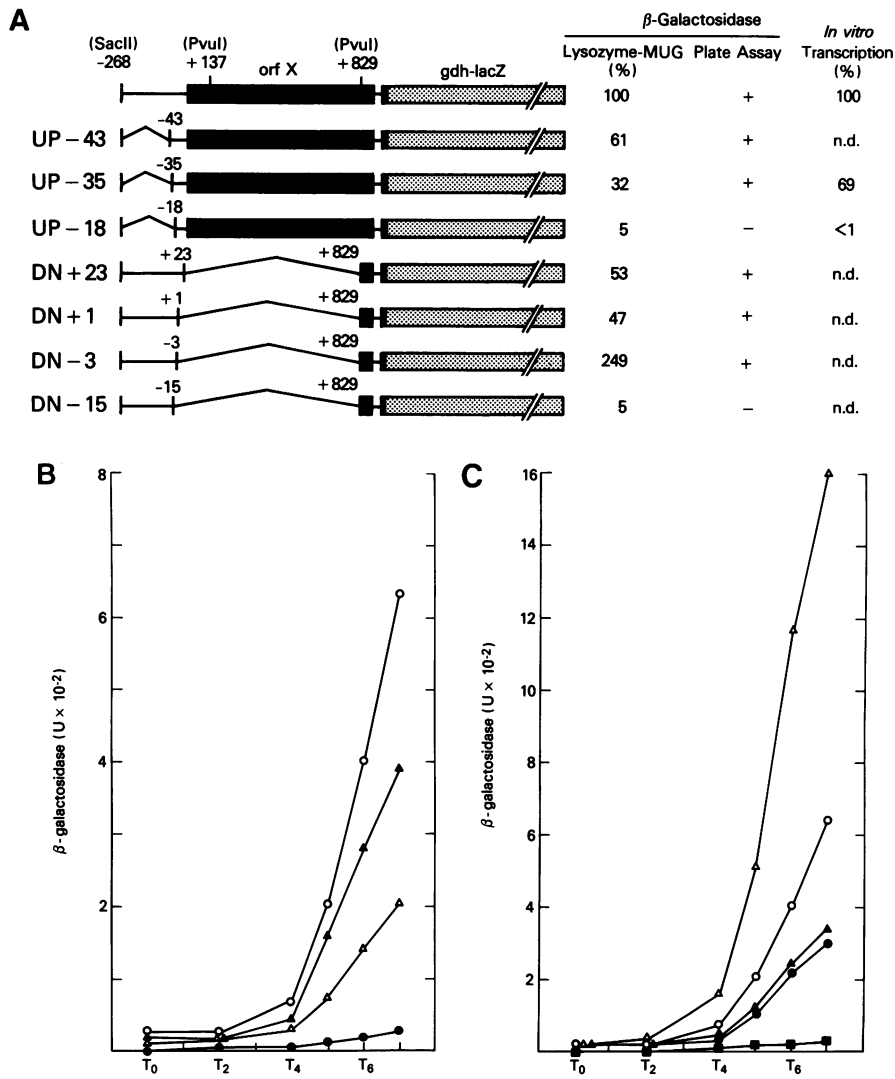


Figure 5. Analysis of β-galactosidase expression from various deletion mutations in *gdh-lacZ* fusion. (A) The *gdh* promoter region was deleted with BAL31 exonuclease from either upstream (UP) or downstream (DN); deletion mutations subcloned in plasmid pYN301 were introduced into the SPβ locus of *B. subtilis* CU2530. The endpoint of a deletion is indicated by the number of the last base still present in the remaining sequence. The transcription startpoint is numbered as +1. β-galactosidase activity was determined quantitatively by the lysozyme-MUG assay and qualitatively by spraying MUG on the surface of NSMP plates containing colonies of the appropriate strain. In the quantitative assay, activity in sporulating cells at T_7 is expressed as a

percentage of the activity in the non-deletion *B. subtilis* strain 62548. In the plate assay, activity is indicated as present (+) or absent (-). Quantitative *in vitro* transcription of *gdh* upstream deletion mutants by purified $E\sigma^G$ RNA polymerase was performed as described in Materials and Methods, and is expressed as a percentage of the non-deletion template as described above. (n.d.: indicates that this value was not determined) (B) β-galactosidase activities of UP mutants; non-deletion (o), UP-43 (▲), UP-35 (Δ), and UP-18 (●), during sporulation were determined quantitatively as described in the legend to Fig. 4 after cell disruption with lysozyme. (C) β-galactosidase activity of DN mutants: non-deletion (o), DN+23 (▲), DN+1 (●), DN-3 (Δ), and DN-15 (■), during sporulation were also determined the same way.

operator region being removed appears unlikely because the DN-3 mutant expressed β-galactosidase activity only after T4, i. e. its timing of expression was controlled just like the other mutants.

The spacer region between "-35" and "-10" is not essential for *gdh* expression

Since only a 35 bp region preceding the transcription start site is essential to control *gdh* expression during sporulation, it appears likely that this is the recognition sequence for RNA polymerase containing a specific sigma factor. In that case, only the bases in the "-35" and the "-10" regions must have the consensus sequence for sigma factor recognition. To test this possibility, we examined the effect of altering the bases in the spacer region between the "-35" and "-10" regions by using cassette mutagenesis to construct a mutant with an altered sequence from -26 to -13 (Fig. 6A). The base replacements were A for T and vice versa, and C for G and vice versa because these changes should produce the greatest possible chemical surface alteration. The resultant "spacer" mutant gave approximately 90% of the β-galactosidase activity of the parent (Fig. 6A). Furthermore, the *gdh-lacZ* fusion containing the "spacer" mutation was expressed only during sporulation and mainly in the forespore compartment as judged by the lysozyme-MUG assay (Fig. 6B). In contrast, a "-35" mutant, which has 3 altered bases at positions -35 to -33, produced no β-galactosidase activity (Fig. 6A). Apparently, only the "-35" and "-10" regions are required to control *gdh* operon expression.

$E\sigma^G$ transcribes *gdh in vitro*

In previous attempts to transcribe the *gdh* operon *in vitro*, Rather and Moran found that RNA polymerase containing σ^A, σ^B, σ^E, and σ^H did not give significant transcription (personal communication). Because the RNA polymerase isolated from stage III sporulating cells is known to contain a new forespore-specific σ–factor (7, 8), we used such an RNA polymerase preparation to test transcription of the *gdh* operon

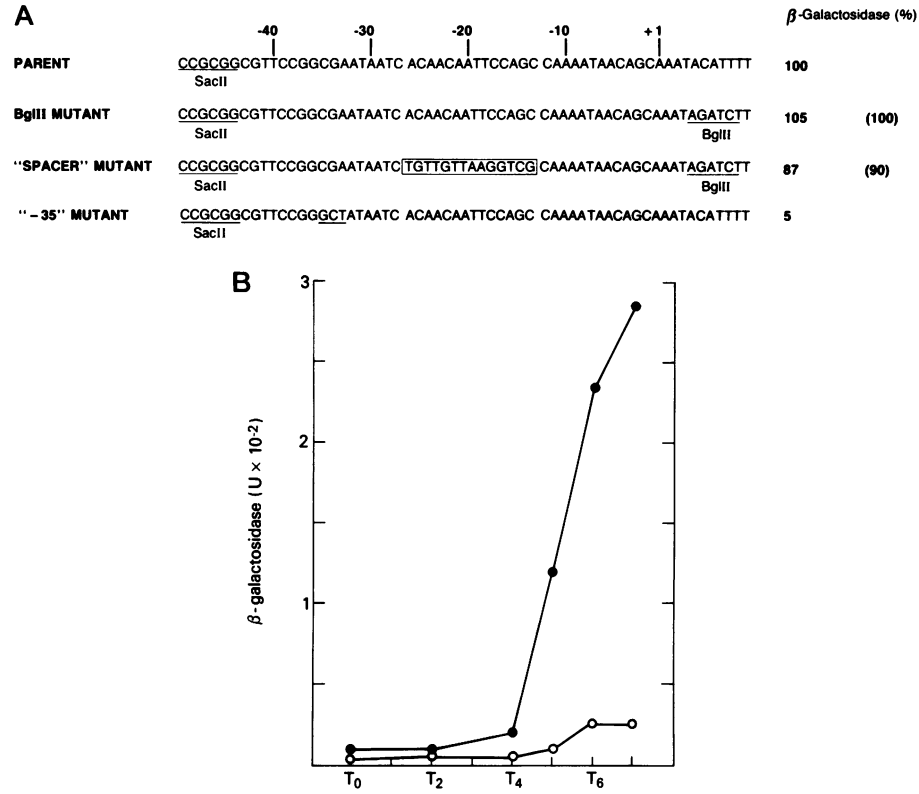


Figure 6. Analysis of β-galactosidase expression from mutants in the spacer region of *gdh-lacZ* fusion. (A) For construction of the "spacer" mutant, the 1 kb *Sal*I-*Bam*HI fragment [the *Sal*I site is in the pYN301 vector (Fig. 2A)] from upstream deletion mutant UP-43 (Fig. 5) was subcloned into plasmid pTZ18. A *Bgl*II site was introduced at +3 from the transcription startpoint of the UP-43 mutant by changing ACATTT to AGATCT by the method of Kunkel et al. (28). A synthetic double-stranded oligonucleotide carrying the spacer mutation was then inserted to replace the 51 bp *Sac*II-*Bgl*II fragment. The altered 1 kb *Sal*I-*Bam*HI fragments were recloned in plasmid pYN301 and integrated into the SPβ locus of strain CU2530. The "-35" mutant was created by changing triple bases in the "-35" region by the method of Kunkel et al. (28). β-galactosidase activity in T_7 cells was measured after lysozyme treatment, and is expressed as the percentage of the activity in either the UP-43 mutant or the strain in which only a *Bgl*II site had been created (values in parentheses). (B) β-galactosidase activity of the "spacer" mutant during sporulation was determined as mentioned in the legend to Fig. 4 after cell disruption with lysozyme (●) and without cell disruption (o).

in vitro. For runoff transcription we used *Hae*III or *Pvu*I, which cut plasmid pEF1 108 or 135 nt respectively, downstream from the *in vivo* transcription startpoint (Fig. 1). The sizes of the resulting runoff transcripts were in close agreement with the length of

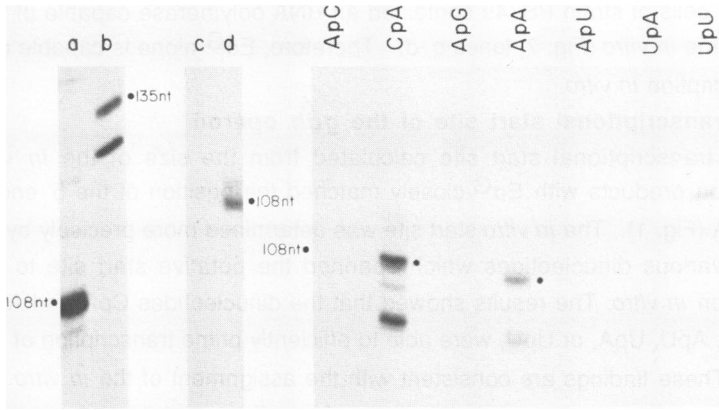


Figure 7. *In vitro* runoff transcription from the *gdh* promoter. *In vitro* transcription reactions were performed as described in Materials and Methods. Lanes (a) and (b): Runoff transcription to the *Hae*III site (lane a) or to the *Pvu*I site (lane b) of pEF1 (see Fig. 1) using RNA polymerase isolated from sporulating cells of strain PS614. Lanes (c) and (d): Runoff transcription with heparin to the *Hae*III site of pEF1 using RNA polymerase isolated from vegetative cells of strain PS435 (lane c) or from vegetative, IPTG-induced cells of strain PS749 (lane d). The remaining lanes are runoff transcripts to the *Hae*III site of pEF1 primed by the dinucleotides listed at the top of each lane, using RNA polymerase isolated from sporulating cells of strain PS614. Dinucleotide priming reactions were performed without heparin. The dots adjacent to the CpA and ApA lanes denote the position of the 108 nt marker.

transcripts expected from the location of the 5' end of *in vivo gdh* mRNA (Fig. 7, lanes a and b). The smaller predominant transcript seen with the *Pvu*I-cut template (Fig. 7, lane b) was seen only with RNA polymerase from sporulating cells. RNA polymerase from IPTG-induced cells of strain PS749 gave only a 135 nt transcript (data not shown). The smaller transcript probably arose from spurious transcription of vector sequences and did not reflect a secondary *gdh* transcript, as it was not observed in runoff transcription reactions of *Hae*III-cut template performed with heparin (Fig7, lane a).

RNA polymerase preparations isolated from sporulating *B. subtilis* at stage III contain both $E\sigma^G$ and $E\sigma^F$ (RNA polymerase containing σ^F), each of which can utilize σ^G-type promoters *in vitro*, although only $E\sigma^G$ is active on this class of promoters *in vivo* (7). To ensure that *in vitro* transcription from the *gdh* promoter was indeed a direct result of its utilization by $E\sigma^G$, runoff transcription was performed using RNA polymerase isolated either from vegetative cells of *B. subtilis* strain PS435 or from strain PS749 after induction with 1mM IPTG.

Vegetative cells of PS435 contain neither $E\sigma^G$ nor $E\sigma^F$, whereas IPTG-induced vegetative cells of PS749 contain only $E\sigma^G$ (7). Strikingly, only IPTG-induced

vegetative cells of strain PS749 contained an RNA polymerase capable of transcribing the *gdh* gene *in vitro* (Fig. 7, lanes c, d). Therefore, $E\sigma^G$ alone is capable of accurate *gdh* transcription *in vitro*.

In vitro transcriptional start site of the *gdh* operon

The transcriptional start site calculated from the size of the *in vitro* runoff transcription products with $E\sigma^G$ closely matched the position of the 5' end of *in vivo* *gdh* mRNA (Fig. 1). The *in vitro* start site was determined more precisely by testing the ability of various dinucleotides which spanned the putative start site to prime *gdh* transcription *in vitro*. The results showed that the dinucleotides CpA and ApA, but not ApC, ApG, ApU, UpA, or UpU, were able to efficiently prime transcription of *gdh in vitro* (Fig. 7). These findings are consistent with the assignment of the *in vitro* and *in vivo* transcription start sites based on the sizes of runoff transcripts and primer extension products, respectively (Fig. 1). [The dinucleotide priming reactions were performed without heparin, and that extra bands seen below 108 nt probably represent degradation products.]

Quantitative _in vitro_ transcription of *gdh* promoter deletion mutants

Plasmid DNA of the various deletion mutations approaching the *gdh* promoter from upstream were assayed by *in vitro* transcription using $E\sigma^G$ isolated from IPTG-induced vegetative cells of strain PS791 (Fig. 5A). A deletion extending to -35 which left the canonical "-35" region intact, reduced the level of *in vitro* transcription only 31%, while the deletion extending to -18 almost completely abolished *gdh* promoter activity *in vitro* (Fig. 5A). Thus, the level of *gdh* transcription *in vitro* from the deletion mutant templates was comparable to the *in vivo* levels of β-galactosidase produced by the corresponding chromosomal *gdh-lacZ* fusion strains. These data suggest that DNA immediately upstream from the *gdh* "-35" region is not essential for *gdh* promoter recognition by $E\sigma^G$ *in vitro*.

DISCUSSION

We have demonstrated that deletions from upstream of the *gdh* operon extending to 35 nucleotides preceding the transcription startpoint had no dramatic effect on *gdh* expression, whereas elimination of the "-35" region abolished *gdh* expression *in vivo*. Similarly, deletions from downstream of the transcription startpoint of the operon up to -3 had no dramatic effect, but elimination of the "-10" region abolished expression of the *gdh* operon *in vivo*. In contrast, sequence changes in the spacer region between the "-10" and the "-35" regions had almost no effect on *gdh* expression *in vivo*. This shows that the control region of the *gdh* operon is confined to a typical prokaryotic promoter, presumably recognized by a sigma factor of RNA polymerase.

We have presented our data assuming that *orfX* and *gdh* are part of operon controlled by the same "-35" and "-10" regions preceeding the transcription startpoint determined by the primer extension method. While it could be argued that this region is a promoter for *orfX* and there is another promoter for *gdh* in the *orfX*, the following results contradict this: (a) Primer extension from +353 showed only one startpoint, indicating there is no additional startpoint at least within 0.3 kb downstream from the startpoint. (b) Northern blots using a fragment within *gdh* and a 0.7 kb *PvuI* fragment within *orfX* (Fig. 1) as probes show only a 1.6 kb RNA band (5; our unpublished data). Moreover, this 1.6 kb RNA appears at the same time during sporulation with either probe. (c) The UP-18 deletion gives no β-galactosidase activity. (d) In our "-35" tripret mutation and the "-10" mutations of Rather and Moran (6), all β-galactosidase expression is abolished, although the sequence in the *orfX* has not been changed.

Recent studies have identified a new sigma factor, designated σ^G (7, 8). This factor directs the transcription both *in vivo* and *in vitro* of a class of sporulation genes (*ssp* genes) which are expressed only in the forespore during stage III of sporulation (7; and W.L. Nicholson, D. Sun, B. Setlow, and P. Setlow, unpublished results). A number of experiments have suggested that σ^G is also responsible for the forespore-specific expression of the *gdh* operon: (a) The *sig*G gene is itself expressed only in the forespore early in stage III of sporulation, shortly before GlcDH synthesis in the forespore (8); (b) GlcDH is not produced during sporulation of a strain carrying a *sig*G deletion mutation (7); and (c) induction of σ^G synthesis in vegetative cells of *B. subtilis* results in rapid synthesis of GlcDH (7). In this work we have further shown that purified $E\sigma^G$ transcribed the *gdh* operon efficiently *in vitro*, with a transcription start site identical to that seen *in vivo*, and with a dependence on bases upstream from the start site identical to that seen *in vivo*.

Transcriptional mapping experiments have revealed that the target genes transcribed by $E\sigma^G$, including *gdh*, contain homologous DNA sequences in the "-35" and "-10" regions of their respective transcriptional start sites, and which therefore define a new class of sporulation-specific promoters. The consensus sequences recognized by $E\sigma^G$ are TGAATA--17-18 bp--CATACTA (W.L. Nicholson, D. Sun, B. Setlow, and P. Setlow, unpublished results). Rather and Moran (6) have shown that single-base substitutions at positions -11 or -7 within the "-10" region abolish *gdh* promoter activity *in vivo*. Interestingly, these single-base alterations are situated at positions within the "-10" consensus sequence which contain invariant bases.

To date, temporal and compartmental control of expression of this class of forespore-specific genes, of which the *gdh* operon is a member, can be explained solely by the forespore-specific expression of *sig*G during stage III of sporulation.

Therefore, further elucidation of this particular subprogram in the overall developmental sequence of *B. subtilis* will require studying factors controlling *sig*G expression.

ACKNOWLEDGEMENTS

We thank Carl Banner for advice on deletion studies and for critical comments on the manuscript, Philip Youngman for helpful discussion on β-galactosidase assay, Soichi Ota for preparation of plasmids, Michael Brenner for critically reading the manuscript, and Devera Schoenberg for editing the manuscript.

*To whom correspondence should be addressed

REFERENCES

1. Piggot, P.J. and Coote, J. G. (1976) Bacteriol. Rev. **40**, 908-962.
2. Losick, R., Youngman, P., and Piggot, P. J. (1986) Ann. Rev. Genet. **20**, 625-669.
3. Fujita, Y., Ramaley, R., and Freese, E. (1977) J. Bacteriol. **132**, 282-293.
4. Vasantha, N., Uratani, B., Ramaley, R. F., and Freese, E. (1983) Proc. Natl. Acad. Sci. USA **83**, 785-789.
5. Uratani, B., Lampel, K. A., Lipsky, R. H., and Freese, E. (1985) In Hoch, J. A. and Setlow, P. (eds), Molecular Biology of Microbial Differentiation. American Society for Microbiology, Washington DC, pp. 71-76.
6. Rather, P. N. and Moran C. (1988) J. Bacteriol. **170**, 5086-5092
7. Sun, D., Stragier, P., and Setlow, P., Genes Dev. in press.
8. Karmazyn-Campelli, C., Bonamy, C., Savelli, B., and Stragier, P., Genes Dev. in press
9. Youngman, P.J., Perkins, J.B. and Losick, R. (1983) Proc. Natl. Acad. Sci. USA **80**, 2305-2309.
10. Mason, J.M., Hackett, R. H., and Setlow, P. (1988) J. Bacteriol. **170**, 239-244.
11. Vasantha, N. and Freese, E. (1980) J. Bacteriol. **144**, 1119-1125.
12. Leighton, T.J. and Doi, R. H. (1971) J. Biol. Chem. **246**, 3189-3195.
13. Messing, J. (1983) Methods Enzymol. **101**, 20-78.
14. Chang, S. and Cohen, S. N. (1979) Mol. Gen. Genet. **168**, 111-115.
15. Dubnau, D. and Davidoff-Abelson, R. (1971) J. Mol. Biol. **56**, 209-221.
16. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York.
17. Lampel, K.A., Uratani, B., Chaudry, G. R., Ramaley, R. F., and Rudikoff, S. (1986) J. Bacteriol. **166**, 238-243.
18. Youngman, P., Perkins, J. B., and Sandman, K. (1985) In Hoch, J. A. and Setlow, P. (eds), Molecular Biology of Microbial Differentiation, American Society for Microbiology, Washington DC, pp.47-54.
19. Yansura, D.G., and Henner, D. J. (1984) Proc. Natl. Acad. Sci. USA **81**, 439-443.
20. Ferrari, F., Trach, K., and Hoch, J. A. (1985) J. Bacteriol. **161**, 556-562.
21. Zuber, P. and Losick, R. (1983) Cell **35**, 275-283.
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467.

23. Penn, M. D., Thireos, G., and Greer, H. (1984) Mol. Cell. Biol. **4**, 520-528.
24. Maxam, A. and W. Gilbert. (1980) Methods Enzymol. **68**, 499-560.
25. LeGrice, S.F.J., Shih, C.-C., Whipple, F., and Sonenshein, A. L. (1986) Mol. Gen. Genet. **204**, 229-236.
26. Minkley, E.G., and Pribnow, D. (1973) J. Mol. Biol. **77**, 255-277.
27. Youngman, P. 1987. In Hardy, K. (ed), Plasmids: A Practical Approach, IRL Press, Oxford, pp. 79-103.
28. Bradford, M.M. (1976) Anal. Biochem. **72**, 248-254.
29. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. **154**, 498-511.