JOURNAL OF BACTERIOLOGY, Feb. 1987, p. 771–778
0021-9193/87/020771-08$02.00/0

Vol. 169, No. 2

# Gene Encoding the 37,000-Dalton Minor Sigma Factor of *Bacillus subtilis* RNA Polymerase: Isolation, Nucleotide Sequence, Chromosomal Locus, and Cryptic Function

MARIAN L. DUNCAN, SUE S. KALMAN, SUSAN M. THOMAS, AND CHESTER W. PRICE*

*Department of Food Science and Technology, University of California, Davis, California 95616*

Received 29 August 1986/Accepted 17 November 1986

**We began an analysis of *rpoF*, the gene encoding the cryptic, 37,000-dalton minor sigma factor (sigma-37) of *Bacillus subtilis* RNA polymerase. Using antibody raised against sigma-37 holoenzyme to probe a λgt11 expression vector library, we isolated a 901-base-pair *Eco*RI fragment that expressed the COOH-terminal half of sigma-37 fused to *lacZ*. We used this fragment as a hybridization probe to isolate the entire *rpoF* gene and additional flanking sequences. Identity of the cloned gene was confirmed by the size and immunological reaction of its product expressed in *Escherichia coli* and, after DNA sequencing, by the homology of its predicted product (264 residues; 30,143 daltons) with other sigma factors. The DNA sequence also suggested that *rpoF* may lie in a gene cluster. Upstream of *rpoF* was an open reading frame that would encode a protein of 17,992 daltons; this frame overlapped the *rpoF*-coding sequence by 41 base pairs. Immediately following *rpoF* was a reading frame that would encode a protein of at least 20,000 daltons; expression of this region may be translationally coupled to that of *rpoF*. By plasmid integration and PBS1 transduction, we found the chromosomal locus of *rpoF* linked to *ddl* and *dal* at 40° on the *B. subtilis* map and near no known lesions affecting growth regulation or development. Further, an *rpoF* null mutation resulting from gene disruption had no effect on cell growth or sporulation in rich medium, suggesting that sigma-37 may partly control a regulon not directly involved in the sporulation process.**

Genetic analysis has begun to define the physiological roles of the multiple sigma factors which associate with *Bacillus subtilis* RNA polymerase core and which determine promoter recognition specificity in vitro (21). The major sigma factor (sigma-43; 43,000 daltons) is analogous to the *Escherichia coli* major sigma (8) and predominates during logarithmic growth (21). The minor form sigma-29 (13), which is the product of the *spoIIG* gene (32, 37), and the sigma-like *spoIIA* gene product (6, 7) both control some aspect of stationary-phase physiology essential to sporulation.

Less is known regarding the roles of the minor forms sigma-37 (14), sigma-32 (18), and sigma-28 (38). These are present during logarithmic growth and may have a vegetative function, but in the absence of authentic sigma-gene mutations their physiological roles remain unknown. We have therefore undertaken a genetic analysis to determine the role of sigma-37 during growth and sporulation and to determine which additional factors and metabolic signals govern its expression and activity. We report here the isolation, nucleotide sequence, and genetic mapping of the sigma-37 structural gene *rpoF*. During the course of these experiments, Binnie et al. (3) independently isolated the sigma-37 gene by using a different screening method. Our results agree with theirs in all major respects and provide additional data on *rpoF* map location and the possible genetic organization of the *rpoF* region.

## MATERIALS AND METHODS

**Bacterial strains.** We used *E. coli* Y1090 (41) as host for λgt11 (42), JM101 (40) as host for M13mp18 and M13mp19, and BNN45 (4) as transformation host for plasmid constructions. *B. subtilis* strains from the *Bacillus* Genetic Stock Center (BGSC), Ohio State University, Columbus, are described in the text.

**DNA and genetic methods.** Isolation of plasmid and lambda DNA, restriction endonuclease digestions, agarose gel electrophoresis, ligations, and transformations of *E. coli* strains were done as described by Davis et al. (4). DNA sequencing was done by the method of Sanger et al. (28) as previously described (34). For mapping the chromosomal locus of *rpoF*, we did plasmid integration, Southern blotting (31), transformation, and PBS1 transduction with *B. subtilis* strains as previously described (26).

**Antibody methods.** Sigma-37-containing holoenzyme ($E\sigma^{37}$) was purified from *B. subtilis* as described by Goldfarb et al. (9); this purified holoenzyme recognized the putative sigma-37 promoter of the subtilisin protease gene in vitro (39). Sigma-37 subunit was purified from the holoenzyme on sodium dodecyl sulfate-polyacrylamide gels (19). We made and characterized antibody from New Zealand White rabbits by using either $E\sigma^{37}$ (anti-$E\sigma^{37}$) or purified sigma-37 (anti-$\sigma^{37}$) as antigens, essentially as described for the alpha subunit (34). The Bio-Rad Express-Blot kit (Bio-Rad Laboratories, Richmond, Calif.) was used to detect antibody bound to plaque lifts (30) and to Western blots (36).

## RESULTS

**Isolation of λgt11 bacteriophages encoding sigma-37.** We screened two λgt11 libraries to isolate phages encoding a product recognized by anti-$E\sigma^{37}$ antibody. As described elsewhere (34), these libraries were constructed with *B. subtilis* chromosomal DNA cut either with *Eco*RI (*Eco*RI library) or with *Alu*I, *Hae*III, and *Rsa*I (random library). We found 18 positive signals among 31,000 phages of the *Eco*RI bank and 9 positive signals in 15,000 phages of the random bank.

Because the initial screening was done with antibody

* Corresponding author.

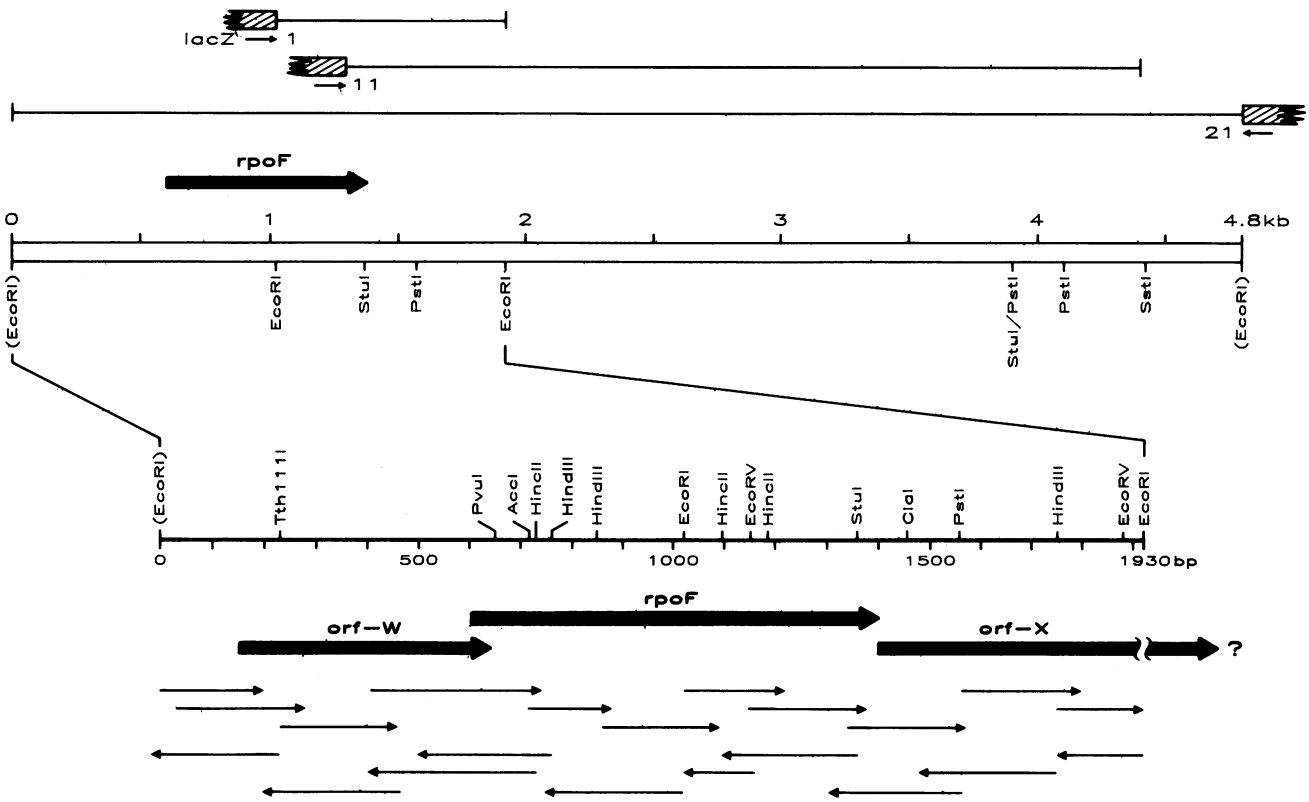



FIG. 1. Physical map of the *B. subtilis rpoF* region. The upper part is a map of the *rpoF* region derived from restriction analysis of the λgt11 recombinant phage inserts shown above the map. The flanking *Eco*RI sites (in parentheses) are from linkers inserted during library construction. Insert orientation is given by the hatched boxes and adjacent arrows, which show the location and direction of transcription of the *lacZ* gene carried on the right arm of λgt11 (42). The *rpoF*-coding sequence and direction of transcription are represented by the heavy arrow above the map. The lower part of the figure is an expanded map of the 1,029- and 901-bp *Eco*RI fragments carried by clone 21. The light arrows beneath the map show the strategy for sequencing both DNA strands. The coding region for *rpoF* and the two flanking open reading frames (ORF-W and ORF-X) are represented as heavy arrows.

raised against holoenzyme, we used epitope selection (30) to identify the clones that might encode sigma-37. Briefly, this technique involved using each plaque-purified clone to affinity purify a specific cross-reacting fraction from the antiholoenzyme antibody. We then assayed each fraction for reaction to the sigma-37 subunit by probing a Western blot of purified sigma-37 holoenzyme. This method identified one clone from the random bank and three from the *Eco*RI bank that likely carried part of the sigma-37 coding region (data not shown).

The restriction maps of the four clones showed that they fell into two classes. One class (represented by clone 1 in Fig. 1) comprised the three *Eco*RI clones, which all carried a 901-base-pair (bp) *Eco*RI fragment in the same orientation with respect to the *lacZ* gene of the vector. The other class was clone 11 from the random bank. Restriction analysis and subsequent DNA sequencing showed that clone 11 carried a 3,100-bp insert which overlapped one end of the 901-bp *Eco*RI fragment by 585 bp (Fig. 1).

**Identification of sigma-37 expressed in *E. coli*.** A Western blot of lysate proteins showed that both classes of clones likely encoded a β-galactosidase–sigma-37 fusion. In *E. coli*, clone 1 of the *Eco*RI class synthesized a 125,000-dalton protein that strongly bound both anti-E$\sigma^{37}$ and anti-$\sigma^{37}$ antibodies, whereas clone 11 of the random class encoded a protein of 115,000 daltons that reacted only with anti-E$\sigma^{37}$ (Fig. 2). The results suggest that clone 1 encoded an 11,000-dalton COOH-terminal fragment of sigma-37 fused to the 114,000-dalton $NH_2$-terminal fragment of *lacZ* and that clone 11 encoded only a 1,000-dalton COOH-terminal sigma fragment similarly fused. DNA sequencing later confirmed that clone 1 encoded the 122 COOH-terminal residues of sigma-37, whereas clone 11 encoded only 17 COOH-terminal residues. Thus, these 17 residues formed an epitope recognized by the anti-E$\sigma^{37}$ antibody but not by the anti-$\sigma^{37}$.

The Western blot results suggest that the remainder of the gene lay to the left of the 901-bp *Eco*RI fragment carried by clone 1 (Fig. 1). Consequently, we used this fragment as a hybridization probe to screen the random λgt11 library for clone 21, which carried additional flanking DNA (Fig. 1) and which expressed a protein with the mobility and antigenic properties of sigma-37 (Fig. 2). In contrast to sigma-29 (37), sigma-37 apparently lacks a larger precursor.

**Nucleotide sequence of the *rpoF* region.** We determined the DNA sequence of the 1,029- and 901-bp *Eco*RI fragments carried by clone 21 (Fig. 1) either by subcloning the appropriate fragments or by making a series of overlapping deletions within M13mp18 (40). Sequence was determined for both DNA strands by sequencing through all restriction sites and deletion endpoints used in the subcloning. To precisely map the beginning of fusion clones 1 and 11, we also sequenced the regions where their inserts abutted the λgt11 *lacZ* gene.

Based on the fusion data and the mass of sigma-37

estimated from gel mobility (Fig. 2), we expected the sigma-37-coding region would begin at about bp 500 on the 1,029-bp *Eco*RI fragment. Consistent with this estimate, the sequence had an open reading frame beginning at bp 601 and extending into the 901-bp *Eco*RI fragment in correct frame for fusion expression in clones 1 and 11 (Fig. 3).

Near the beginning of the frame, at nucleotides 587 through 594, was a sequence similar to other *B. subtilis* ribosomal binding sites (12), followed by two possible initiation codons, i.e., TTG at 601 through 603 and ATG at 607 through 609. We chose TTG because it was optimally spaced from the proposed ribosomal binding site (12). The protein product encoded by this frame had significant homology to other sigma factors in regions proposed to be important for core recognition and DNA binding (Fig. 4).

Thus, by the criteria of gel mobility, antigenicity, and homology to other sigma factors, we isolated the structural gene for sigma-37. Consistent with the accepted nomenclature for RNA polymerase genes (15), we designated the gene *rpoF*. As with the major sigma factor (8), the predicted size of sigma-37 (264 residues; 30,143 daltons) was smaller than that expected from its gel mobility.

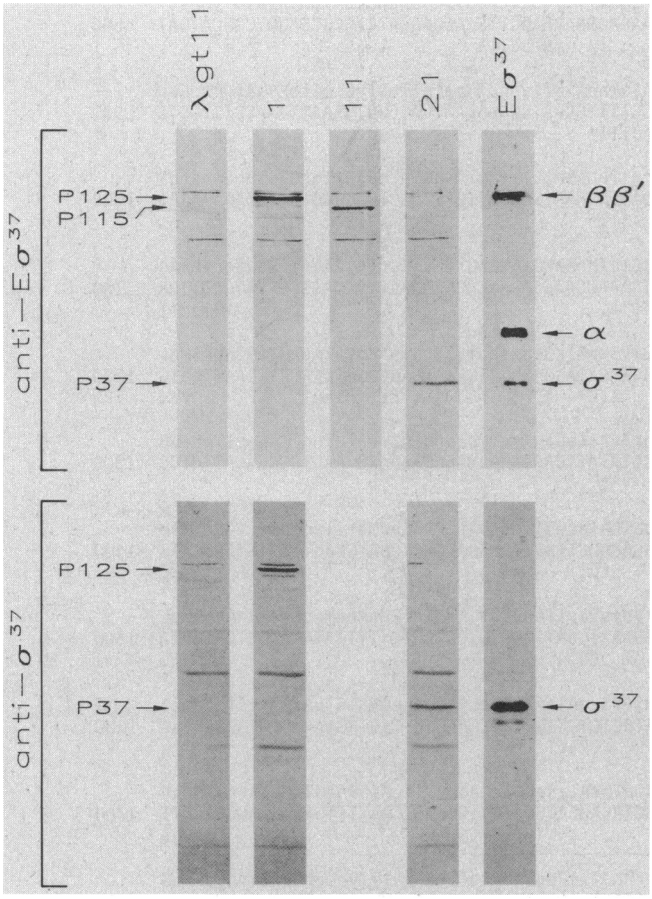



FIG. 2. Western blot of *B. subtilis* sigma-37 expressed in *E. coli*. Proteins from lysates of *E. coli* Y1090 infected with each of the three λgt11 clones and the λgt11 vector alone were separated on a 10% polyacrylamide gel (19) and electroeluted onto nitrocellulose. The blots were incubated with antibody raised against either sigma-37 RNA polymerase holoenzyme (anti-Eσ$^{37}$) or purified sigma-37 subunit alone (anti-σ$^{37}$). Clone 21 encoded a protein (P37) which had the same mobility as the sigma-37 standard in lane 5 and which was recognized by both antibodies. Clones 1 and 11 encoded fusion proteins of 125,000 (P125) and 115,000 (P115) daltons, respectively.

Additional features of the sequence suggest possible regulation and genetic organization of the *rpoF* region. Immediately preceding *rpoF* was a region of dyad symmetry at bp 523 through 572. A transcript through the region of symmetry has the potential to form a stem-loop structure (ΔG = −20.4 kcal [1 cal = 4.184 J] [35]) that may influence *rpoF* expression. Beginning 442 bp upstream of *rpoF* was an open reading frame (ORF-W) encoding a hypothetical protein of 160 residues (17,992 daltons), preceded by a possible ribosomal binding site and an ATG initiation codon (Fig. 3). This reading frame overlapped the *rpoF* coding sequence by 41 bp. We have not yet determined whether this reading frame is expressed.

Another possible ribosomal binding site lay within the 3′ end of *rpoF*, preceding at ATG initiation codon which began an open reading frame (ORF-X) that would encode a protein of at least 20,000 daltons (Fig. 3). Because this ribosomal binding site was quite weak (ΔG = −9.4 kcal) and because the termination codon of *rpoF* overlapped the initiation codon of ORF-X by 1 bp, expression of ORF-X may be translationally coupled to that of *rpoF*.

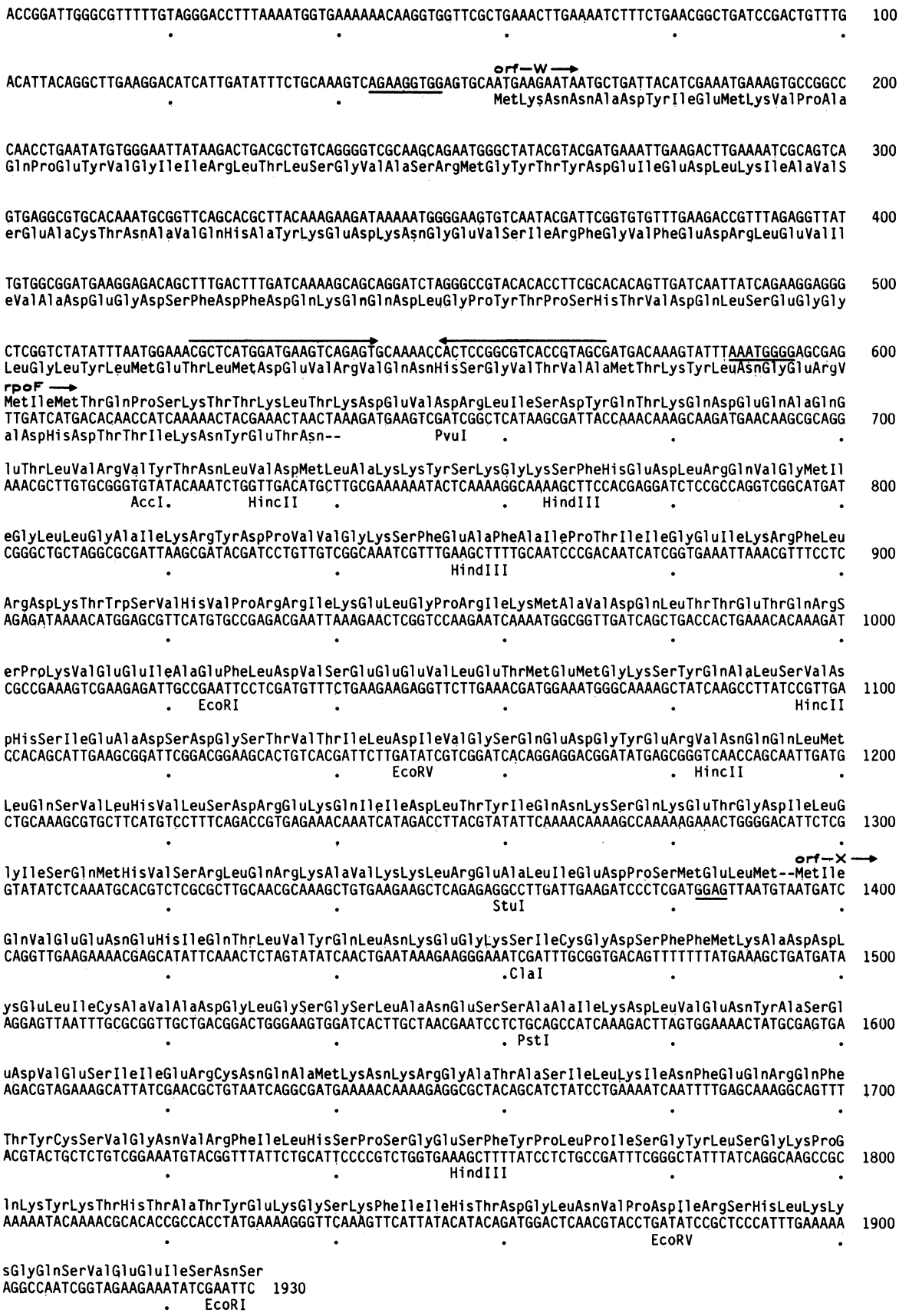
**Chromosomal locus of *rpoF*.** We mapped the chromosomal locus of *rpoF* to establish whether any known mutations lay either within the *rpoF*-coding sequence or within the flanking open reading frames. We moved the 2,500-bp *Eco*RI fragment from clone 11 into the pCP115 integrative mapping plasmid (26) and transformed strain PB72 (*furB1 trpC2*; BGSC 1A113) to Cm$^r$, thereby inserting a selectable drug resistance element (*cat*) adjacent to the chromosomal locus of *rpoF* (indicated by the symbol *rpoF*::*cat*). Southern blotting (31) confirmed the plasmid integration near *rpoF* (data not shown). We used the resulting Cm$^r$ strain PB73 (*rpoF*::*cat furB1 trpC2*) as donor in PBS1 transductional crosses with the kit strains of Dedonder et al. (5) as recipients.

With strain PB4 (*aroI906 dal-1 purB33 trpC2*; BGSC 1A4) as recipient, Cm$^r$, and therefore *rpoF*, was 90% linked to *dal* (86 Dal$^+$/95 Cm$^r$), 25% linked to *aroI* (24 Aro$^+$/95 Cm$^r$), and 4% linked to *purB* (4 Pur$^+$/95 Cm$^r$), placing *rpoF* at 40° on the *B. subtilis* map (25). We used one of the recombinants from this cross, strain PB74 (*aroI906 dal-1 rpoF*::*cat purB33 trpC2*), as donor in a three-factor transductional cross with PB75 (*ddl-1475 ilvA1 metB5 purA16*; BGSC 1A463) as recipient (Table 1). Marker order indicated by this cross was *ddl dal rpoF* (Fig. 5).

The linkage between *rpoF* and *dal* was confirmed by a transformational cross between the PB73 donor and PB4 recipient (Fig. 5). Because the transformational linkage data suggested that *rpoF* and *dal* were physically close, we also transformed PB4 with donor DNA from clone 21 and established that *dal* was not carried within the cloned region.

**An *rpoF* null mutation sporulates normally.** Because no known mutations lie within the *rpoF* locus, we made a null mutation to help establish sigma-37 function. We moved the 376-bp *Pvu*I-*Eco*RI fragment, which carried *rpoF*-coding information but which lacked the 5′ and 3′ ends of the gene (Fig. 1 and 3), to pCP115 and transformed *B. subtilis* PB2 (26) to Cm$^r$. Southern blotting (31) confirmed that Campbell integration of this plasmid disrupted *rpoF* (data not shown).

The null *rpoF* mutation caused no obvious phenotype. The growth rate, sporulation frequency, and spore heat resistance of the mutant and wild type were indistinguishable on Schaeffer 2× SG sporulation medium (26). The mutant also produced extracellular subtilisin protease activity during


orf-W
rpoF
orf-X
PvuI
AccI.
HincII
HindIII
EcoRI
EcoRV
StuI
ClaI
PstI

A.

| | | |
|---|---|---|
| Sigma-37 | 41-73 | tnLVdmLAKKYs.KGksf..HEDLrQvGmIGLLgAI |
| Sigma-70 | 384-416 | LRLVIsIAKKYtNRGLqf...lDLIQEGNIGLMKAV |
| Sigma-43 | 143-175 | LRLVVsIAKRYvgRGMlf...lDLIhEGNMGLMKAV |
| Sigma-32 | 58-90 | LRfVVhIARnYagyGLpq...aDLIQEGNIGLMKAV |
| Sigma-29 | 67-99 | LRLVVyIARKFeNtGIni...EDLIsiGtIGLIKAV |
| SPO1 gp28 | 28-63 | Litfaarqqmenngadtmms RQDLeQEGllkLydcw |
| SPO1 gp34 | 25-58 | LRksVy..KKFkDKmInqsdREDLMgtiDqifLqlV |
| T4 gp55 | 48-83 | MlIaegLsKRFnfsGytqswKQEMIaDGieAsIKGL |

| | | |
|---|---|---|
| Sigma-37 | 74-109 | kRYDPvvGksFeaFAiptIigeIkRflrDktwSVHV |
| Sigma-70 | 417-452 | DKFEyrRGyKFsTYATwWIRQaItRsIaDQaRTIRI |
| Sigma-43 | 176-211 | EKFDyrKGyKFsTYATwWIRQaItRaIaDQaRTIRI |
| Sigma-32 | 91-126 | rRFNPevGvRlvSFAvhWIKaeIheyVlrNwRiVKV |
| Sigma-29 | 100-134 | NtFNPeKkiKlaTYASrcIeNeIlmyLrrNnK.IRs |
| SPO1 gp28 | 64-98 | EKWcfkenkqmdeFGpiF.RksLfRkVkQsggTgRa |
| SPO1 gp34 | 59-94 | seYNPnRGvdFpyYikrmLelrtyHhItkylKrIng |
| T4 gp55 | 84-115 | hnFDetKyknphaYiT..Qa cfnafvQriKkeRk |

B.

| | | |
|---|---|---|
| Sigma-37 | 132-165 | QRSPKVEEIAEFLDVSEEEVLETMEMGKSYQALS |
| Sigma-70 | 475-506 | GREPTPEELAERMLMPEDKIRK..VLKIAKEPIS |
| Sigma-43 | 234-265 | GREPTPEEIAEDMDLTPEKVRE..ILKIAQEPVS |
| Sigma-32 | 149-181 | FNQDEVEMVARELGVTSKDVRE.MESRMAAQDMT |

α — T — α

| | | |
|---|---|---|
| Sigma-37 | 221-253 | IQNKSQKETGDILGISQMHVS.RLQRKAVKKLRE |
| Sigma-70 | 567-600 | NTDYTLEEVGKQFDVTRERIRQ.IEAKALRKLRH |
| Sigma-43 | 327-359 | GRTRTLEEVGKVFGVTRERIRQ.IEAKALRKLRH |
| Sigma-32 | 248-280 | DNKSTLQELADRYGVSAERVRQ.LEKNAMKKLRA |
| Sigma-29 | 200-233 | EEEKTQKDVADMMGISQSYIS.RLEKRIIKRLRK |
| SPO1 gp28 | 185-217 | DTTVRMKHIDQTLGISNKQYDSEL.KKFVKRLTI |
| SPO1 gp34 | 151-183 | IRKKTLQELAQEEGVPLDRLHARLYF.LIRKFEK |

FIG. 4. Comparison of the predicted amino acid sequence of *B. subtilis* sigma-37 to the proposed core and DNA-binding regions of other sigma factors. The sequence of sigma-37 from Fig. 3 is given here in single-letter code. The alignments and proposed functional domains of the other seven sequences are from Gribskov and Burgess (10). (A) Alignment in the proposed core binding regions. Chemically similar residues (AG, DENQ, FWY, ILMV, KHR, and ST) are uppercase when present in at least four sequences and boxed when present in six or more. Periods indicate gaps introduced to improve the alignment. (B) Alignment of proposed DNA-binding regions. The helix-turn-helix secondary structure predicted by Gribskov and Burgess is indicated between the aligned sequences; boxes enclose conserved regions found in other DNA-binding proteins (10). Four of the five cellular sigma factors, including sigma-37, have two helix-turn-helix regions; sigma-29 and the SPO1 phage factors have one.

postexponential growth, grew well on glucose-ammonia minimal plates (26), and grew well at 37 and 45°C. Thus, sigma-37 is not essential for growth or development under the conditions used.

However, similar disruptions of the region downstream of *rpoF* did produce growth and sporulation phenotypes. Integrating pCP115 via the 367-bp *Pst*I-*Eco*RI fragment contained within ORF-X (Fig. 1 and 3) caused a small-colony phenotype, but this strain proved unstable, reverting to wild-type colony size and retaining $Cm^r$. In addition to mutagenesis of single cistrons, plasmid integration can also suggest the transcriptional organization of a region by separating downstream genes from their promoter (24). In this regard, integrating pCP115 by using the larger, 901-bp *Eco*RI fragment from clone 21 (carrying the 3′ end of *rpoF* as well as the 5′ end of ORF-X; Fig. 1 and 3) also caused a small-colony phenotype which proved stable. Notably, this strain grew more slowly than wild type in 2× SG medium (doubling time of 33 versus 23 min) and was oligosporogenous. Thus, ORF-X (or a downstream gene) may be cotranscribed with *rpoF* and likely functions in both vegetative and sporulating cells.

FIG. 3. DNA sequence of the *rpoF* region. The sequence of the 1,029- and 901-bp *Eco*RI fragments begins with the A (nucleotide 1) immediately following the *Eco*RI linker of clone 21. Clone 1 begins at nucleotide 1025, and clone 11 begins at 1340. The predicted amino acid sequences for sigma-37 (*rpoF*) and the protein that would be encoded by ORF-X are given above the DNA sequence, whereas that for the protein that would be encoded by ORF-W is given below. Sequences for proposed ribosomal binding sites (12) for ORF-W ($\Delta G = -12.8$ kcal [35]), *rpoF* ($\Delta G = -13.6$ kcal), and ORF-X ($\Delta G = -9.4$ kcal) are underlined at nucleotides 144 through 152, 587 through 594, and 1383 through 1386, respectively. A transcript through the region of dyad symmetry upstream of *rpoF* (converging arrows between nucleotides 523 through 572) has the potential to form a stem-loop structure ($\Delta G = -20.4$ kcal) that may affect *rpoF* expression.

TABLE 1. Three-factor transductional cross to map the *rpoF* chromosomal locus[a]

| Selection | Recipient classes[b] | | | No. of recombinants | Order implied |
|---|---|---|---|---|---|
| | Ddl | Dal | Cm$^r$ | | |
| Ddl$^+$ | 1 | 1 | 1 | 75 | *ddl dal rpoF* |
| | 1 | 1 | 0 | 5 | |
| | 1 | 0 | 1 | 0 | |
| | 1 | 0 | 0 | 16 | |
| Cm$^r$ | 1 | 1 | 1 | 84 | *ddl dal rpoF* |
| | 0 | 1 | 1 | 7 | |
| | 1 | 0 | 1 | 0 | |
| | 0 | 0 | 1 | 5 | |

[a] Donor was PB74 *aroI906 dalI purB33 rpoF::cat trpC2*. Recipient was PB75 *ddl1475 ilvA1 metB5 purA16*.
[b] 1, Donor phenotype; 0, recipient phenotype.

## DISCUSSION

A genetic analysis is essential to establish the physiological roles, regulation, and functional interactions of the components of the complex *B. subtilis* transcriptase. Losick and Pero (21) have suggested that the minor sigma factors in association with the transcriptase partly control the sporulation process, and indeed, mutations within the *spoIIAC* and *spoIIG* genes, which encode sigmalike products, do affect development (6, 7, 32, 37).

In contrast, the minor sigma factors of the enteric transcriptase partly control global systems such as the heat shock (11, 20) and nitrogen (16, 17) regulons. Here we report the beginning of a genetic analysis of the 37,000-dalton minor sigma factor of *B. subtilis* RNA polymerase, sigma-37. In vitro, the sigma-37 holoenzyme is known to transcribe three genes, all of which are expressed only during the stationary growth phase: *aprA*, encoding the extracellular subtilisin protease (39), and the cryptic *spoVG* and *ctc* genes (14). However, sigma-37 is also found in logarithmically growing cells (14) and thus may have a vegetative function analogous to the enteric factors.

We isolated the sigma-37 gene from a λgt11 library by using an antibody probe and confirmed the identity of the gene product by its size, antigenicity, and homology with other sigma factors (Fig. 2 and 4). Consistent with the accepted nomenclature for RNA polymerase genes (15), we designated the sigma-37 gene *rpoF*. Using an oligonucleotide probe, Binnie et al. (3) independently isolated the sigma-37 structural gene. Our results agree in all major respects. They found an identical nucleotide sequence for the fragment they analyzed (between bp 477 and 1472 [Fig. 3]); they suggested a start of the coding region the same within 6 bp; they determined a similar chromosomal locus; and they found a cryptic phenotype for a null mutation in the sigma-37 gene, which they call *sigB*.

Because several mutations isolated for their developmental phenotype later proved to be alterations in sigma-factor genes (6, 7, 26, 32, 37), we mapped the chromosomal locus of *rpoF* (or *sigB*) to determine whether any known mutations lie within the sigma-37-coding sequence. By plasmid integration and PBS1 transduction, *rpoF* was closely linked to *dal* and *ddl* at 40° on the *B. subtilis* map, a region presently devoid of markers that affect development or growth regulation (25). However, the research groups of Haldenwang, Losick, and Moran have isolated and mapped to this region mutations which enhance *ctc* gene expression; these may lie within either *rpoF* (22) or the flanking reading frames we identified (Fig. 1). Because the function of the *ctc* gene itself remains cryptic, in vitro mutagenesis of *rpoF* provided another approach to establishing the physiological role of sigma-37. In this regard, our gene disruption experiments suggest that sigma-37 is not essential for either sporulation or subtilisin gene expression (39) and support the notion that sigma-37 controls a regulon not directly involved with development. The results of Binnie and his colleagues also support this view; their null sigma-37 mutation had no effect on sporulation (3) or transcription of *spoVG* (22), but it abolished transcription of *ctc* (22). But because an *rpoF* disruption would also prevent expression of distal genes cotranscribed with *rpoF*, a definitive analysis of sigma-37 function requires a better understanding of the genetic organization of the region.

Once the role of sigma-37 has been identified, the question becomes what factors and metabolic signals regulate expression and activity of sigma-37 itself. Our DNA sequence analysis suggests two features that may be important in *rpoF* regulation. First, we found a potential RNA secondary structure encoded immediately upstream of *rpoF* (Fig. 3) which may partly control *rpoF* expression. Second, we found possible open reading frames overlapping both the 5′ and 3′ ends of the *rpoF*-coding sequence (Fig. 1 and 3).

The initiation codon of the downstream frame, ORF-X, overlapped the first of the dual termination codons of *rpoF*, i.e., UAAUG. This arrangement is strikingly similar to genes known to be translationally coupled in the tryptophan operon of *E. coli* (23). Translational coupling is presumed to ensure equimolar synthesis of products that associate in vivo (2, 23, 29), although this is not invariably the case (27). We have no direct evidence that ORF-X is expressed, but plasmid integration in the ORF-X region produced slow-growing, oligosporogenous strains. Our results suggest that ORF-X (or a distal gene) is cotranscribed with *rpoF* and may function during both the vegetative and sporulation phases of growth. If ORF-X proves to be cotranscribed with *rpoF*, the absence of a small-colony phenotype in our *rpoF* gene

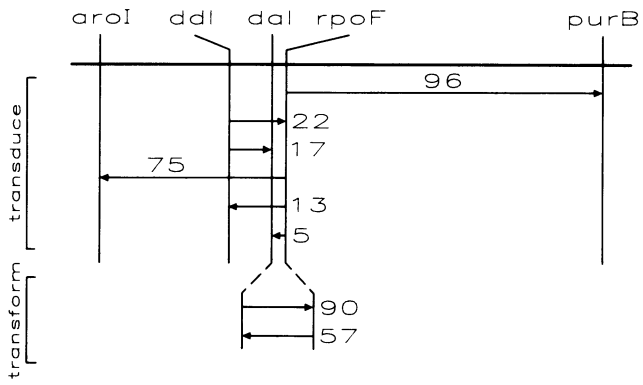



FIG. 5. Genetic map of the *B. subtilis rpoF* region. Genetic distances within the *rpoF-ddl* interval were calculated from the three-factor transductional cross in Table 1; those to *aroI* and *purB* were from the two-factor cross described in the text. Both are expressed as percent recombination (100% minus percent cotransduction). Distances between *rpoF* and *dal* from the two-factor transformational cross described in the text are expressed as 100% minus percent cotransformation. Arrows point from the selected marker to the unselected marker in each cross. For *rpoF*, selection and scoring were actually for the Cm$^r$ element of the pCP115 mapping plasmid (26) integrated adjacent to the *rpoF* chromosomal locus.

disruption is formally analogous to intergenic suppression and leads us to speculate that loss of the ORF-X product is deleterious only when *rpoF* is expressed. Establishing the possible functional relationship between sigma-37 and the ORF-X product will require additional genetic analysis.

Nor do we yet have evidence that the upstream ORF-W is expressed. This frame extends 41 bp into the $NH_2$-terminal coding region of *rpoF*. A precedent for such sequence overlap is found in the *B. subtilis* tryptophan operon, in which the *trpD*-coding region begins 29 bp before the *trpE* termination codon (1). Because the results of Schumperli et al. (29) show that even a 98-bp overlap does not adversely affect translation of a downstream gene, we cannot exclude the possibility that ORF-W encodes a functional product. S1 mapping and operon fusion experiments can address whether the *rpoF* transcriptional unit encompasses these adjacent reading frames. Should *rpoF* constitute part of an operon, such experiments will also help interpret the phenotypes of null and point mutations which may manifest transcriptional or translational polarity in vivo.

Finally, our sigma-37 sequence supports and extends the analysis of Gribskov and Burgess (10) and of Stragier et al. (33), who suggested on the basis of homology that certain sigma domains are important for core recognition and DNA binding. By using the cloned sigma-37 gene, hybrid sigma factors can now be constructed to experimentally define sigma domains crucial for promoter specificity.

## ACKNOWLEDGMENTS

We thank Richard Burgess and Richard Losick for communicating their results before publication, Stanley Zahler for helpful discussions on the genetic mapping, and Daniel Zeigler and Donald Dean of the *Bacillus* Genetic Stock Center for providing strains.

This research was supported by Public Health Service grant AI 22459 from the National Institute of Allergy and Infectious Diseases and by funds from the University of California.

## LITERATURE CITED

1. **Band, L., H. Shimotsu, and D. J. Henner.** 1984. Nucleotide sequence of the *Bacillus subtilis trpE* and *trpD* genes. Gene **27**:55–65.
2. **Baughman, G., and M. Nomura.** 1983. Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in *Escherichia coli*. Cell **34**:979–988.
3. **Binnie, C., M. Lampe, and R. Losick.** 1986. Gene encoding the $\sigma^{37}$ species of RNA polymerase $\sigma$ factor from *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **83**:5943–5947.
4. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics: a manual for genetic engineering. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Dedonder, R. A., J.-A. Lepesant, J. Lepesant-Kejzlarová, A. Billault, M. Steinmetz, and F. Kunst.** 1977. Construction of a kit of reference strains for rapid genetic mapping in *Bacillus subtilis* 168. Appl. Environ. Microbiol. **33**:989–993.
6. **Errington, J., P. Fort, and J. Mandelstam.** 1986. Duplicated sporulation genes in bacteria. FEBS Lett. **188**:184–188.
7. **Fort, P., and P. J. Piggot.** 1984. Nucleotide sequence of sporulation locus *spoIIA* in *Bacillus subtilis*. J. Gen. Microbiol. **130**:2147–2153.
8. **Gitt, M. A., L.-F. Wang, and R. H. Doi.** 1985. A strong sequence homology exists between the major RNA polymerase sigma factors of *Bacillus subtilis* and *Escherichia coli*. J. Biol. Chem. **260**:7178–7185.
9. **Goldfarb, D. S., S.-L. Wong, T. Kudo, and R. H. Doi.** 1983. A temporally regulated promoter from *Bacillus subtilis* is transcribed only by an RNA polymerase with a 37,000 dalton sigma factor. Mol. Gen. Genet. **191**:319–325.
10. **Gribskov, M., and R. R. Burgess.** 1986. Sigma factors from *E. coli*, *B. subtilis*, phage SPO1, and phage T4 are homologous proteins. Nucleic Acids Res. **14**:6745–6763.
11. **Grossman, A. D., J. W. Erickson, and C. Gross.** 1984. The *htpR* gene product of *E. coli* is a sigma factor for heat shock proteins. Cell **38**:383–390.
12. **Hager, P. W., and J. C. Rabinowitz.** 1985. Translational specificity in *Bacillus subtilis*, p. 1–32. *In* D. A. Dubnau (ed.), The molecular biology of the bacilli. Academic Press, Inc., N.Y.
13. **Haldenwang, W. G., N. Lang, and R. Losick.** 1981. A sporulation-induced sigma-like regulatory protein from *B. subtilis*. Cell **23**:615–624.
14. **Haldenwang, W. G., and R. Losick.** 1980. A novel RNA polymerase sigma factor from *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **77**:7000–7005.
15. **Hayward, R. S., and J. G. Scaife.** 1976. Systematic nomenclature for the RNA polymerase genes of prokaryotes. Nature (London) **260**:646–648.
16. **Hirschman, J., P.-K. Wong, K. Sei, J. Keener, and S. Kustu.** 1985. Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a $\sigma$ factor. Proc. Natl. Acad. Sci. USA **82**:7525–7529.
17. **Hunt, T. P., and B. Magasanik.** 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc. Natl. Acad. Sci. USA **82**:8453–8457.
18. **Johnson, W. C., C. P. Moran, and R. Losick.** 1983. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature (London) **302**:800–804.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Landick, R., V. Vaughn, E. T. Lau, R. A. VanBogelen, J. W. Erickson, and F. C. Neidhardt.** 1984. Nucleotide sequence of the heat shock regulatory gene of *E. coli* suggests its protein product may be a transcription factor. Cell **38**:175–182.
21. **Losick, R., and J. Pero.** 1981. Cascades of sigma factors. Cell **25**:582–584.
22. **Losick, R., P. Youngman, and P. J. Piggot.** 1986. Genetics of endospore formation in *Bacillus subtilis*. Annu. Rev. Genet. **20**:625–670.
23. **Oppenheim, D. S., and C. Yanofsky.** 1980. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. Genetics **95**:785–795.
24. **Piggot, P. J., C. A. M. Curtis, and H. de Lencastre.** 1984. Use of integrational plasmid vectors to demonstrate the polycistronic nature of a transcriptional unit (*spoIIA*) required for sporulation of *Bacillus subtilis*. J. Gen. Microbiol. **130**:2123–2136.
25. **Piggot, P. J., and J. A. Hoch.** 1985. Revised genetic linkage map of *Bacillus subtilis*. Microbiol. Rev. **49**:158–179.
26. **Price, C. W., and R. H. Doi.** 1985. Genetic mapping of *rpoD* implicates the major sigma factor of *Bacillus subtilis* RNA polymerase in sporulation initiation. Mol. Gen. Genet. **201**:88–96.
27. **Riggs, D., and S. Artz.** 1984. The *hisD-hisC* gene border of the *Salmonella typhimurium* histidine operon. Mol. Gen. Genet. **196**:526–529.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
29. **Schumperli, D., K. McKenney, D. A. Sobieski, and M. Rosenberg.** 1982. Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. Cell **30**:865–871.
30. **Snyder, M., and R. W. Davis.** 1985. Screening λgt11 expression libraries with antibody probes, p. 397–406. *In* T. Springer (ed.), Hybridoma technology in biosciences and medicine. Plenum Publishing Corp., New York.
31. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

32. **Stragier, P., J. Bouvier, C. Bonamy, and J. Szulmajster.** 1984. A developmental gene product of *Bacillus subtilis* homologous with the sigma factor of *Escherichia coli*. Nature (London) **312**:376–378.
33. **Stragier, P., C. Parsot, and J. Bouvier.** 1985. Two functional domains conserved in major and alternative sigma factors. FEBS Lett. **187**:11–15.
34. **Suh, J.-W., S. A. Boylan, and C. W. Price.** 1986. Gene for the alpha subunit of *Bacillus subtilis* RNA polymerase maps in the ribosomal protein gene cluster. J. Bacteriol. **168**:65–71.
35. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.
36. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
37. **Trempy, J. E., C. Bonamy, J. Szulmajster, and W. G. Haldenwang.** 1985. *Bacillus subtilis* sigma factor $\sigma^{29}$ is the product of the sporulation-essential gene *spoIIG*. Proc. Natl. Acad. Sci. USA **82**:4189–4192.
38. **Wiggs, J. L., M. Z. Gilman, and M. J. Chamberlin.** 1981. Heterogeneity of RNA polymerase in *Bacillus subtilis*: evidence for an additional sigma factor in vegetative cells. Proc. Natl. Acad. Sci. USA **78**:2762–2766.
39. **Wong, S.-L., C. W. Price, D. S. Goldfarb, and R. H. Doi.** 1984. The subtilisin E gene of *Bacillus subtilis* is transcribed from a $\sigma^{37}$ promoter *in vivo*. Proc. Natl. Acad. Sci. USA **81**:1184–1188.
40. **Yanisch-Perron, C. Y., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
41. **Young, R. A., and R. W. Davis.** 1983. Yeast RNA polymerase II genes: isolation with antibody probes. Science **222**:778–783.
42. **Young, R. A., and R. W. Davis.** 1983. Efficient isolation of genes using antibody probes. Proc. Natl. Acad. Sci. USA **80**: 1194–1198.