An Operon of *Bacillus subtilis* Motility Genes Transcribed by the σ^D Form of RNA Polymerase

DANIEL B. MIREL, VERONICA M. LUSTRE, AND MICHAEL J. CHAMBERLIN*

Division of Biochemistry and Molecular Biology, 401 Barker Hall, University of California, Berkeley, California 94720

Received 30 September 1991/Accepted 24 April 1992

Two genes controlling motility functions in *Bacillus subtilis* were identified by DNA sequence analysis of a chromosomal fragment containing a strong promoter for σ^D RNA polymerase. Previous studies had shown that this σ^D -dependent promoter controls synthesis of a 1.6-kb transcript in vivo and in vitro. Sequence analysis revealed that the 1.6-kb transcript contains two open reading frames coding for protein sequences homologous to the *Escherichia coli* *motA* and *motB* gene products, respectively, and ends in a rho-independent termination site. Direct evidence linking these genes to motility functions in *B. subtilis* was obtained by precise localization by polymerase chain reaction of Tn917 transposon insertion mutations of *Mot*[−] strains, isolated by Zuberi et al. (A. R. Zuberi, C. Ying, H. M. Parker, and G. W. Ordal, *J. Bacteriol.* 172:6841–6848, 1990), to within this *mot* operon. Replacement of each wild-type gene by in-frame deletion mutations yielded strains possessing paralyzed flagella and confirmed that both *motA* and *motB* are required for the motility of *B. subtilis*. These current findings support our earlier suggestions that σ^D in *B. subtilis* plays a central role in the control of gene expression for flagellar assembly, chemotaxis, and motility functions. σ^E , the enteric homolog of σ^D , controls similar functions in *E. coli* and *Salmonella typhimurium*, and these factors appear to be representative of a family of factors implicated in flagellar synthesis in many bacterial species, which we propose to designate the σ^{28} family.

The regulation of bacterial gene expression often occurs at the first step in transcription, involving promoter recognition by RNA polymerase. The specificity of this interaction is determined by proteins known as sigma (σ) factors that confer recognition of distinct promoter sequences. The gram-positive bacterium *Bacillus subtilis* possesses many alternative sigma factors and employs them in the temporal and developmental regulation of specific sets of genes or regulons.

B. subtilis σ^D is a secondary sigma factor that controls a regulon of genes expressed in exponential and early post-exponential growth (15, 16, 33, 35) (see reference 22 for a review). This factor was originally identified as a sigma factor of 28 kDa (σ^{28}) that conferred recognition of a unique promoter specificity on *B. subtilis* RNA polymerase (47). Subsequently, the σ^D promoter consensus sequence was found to be CTAAA...N₁₆...CCGATAT for *B. subtilis* promoters (17, 42). *B. subtilis* cells bearing a disruption of the σ^D structural gene are viable but have a nonflagellated (Fla[−]) phenotype, suggesting a role for σ^D in the transcription of genes affecting flagellar synthesis or assembly (24).

This finding led to the recognition that operons of the flagellar assembly-chemotaxis-motility (*fla/che/mot*) regulon of enteric bacteria are preceded by this same consensus and led to the prediction that these operons might be transcribed by an enteric homolog of *B. subtilis* σ^D (5, 23). This conjecture was confirmed with the identification and isolation of the *Escherichia coli* homolog, designated σ^E (3), now known to be the product of the *flaA* gene (37). Factors having this same promoter specificity are implicated in flagellar gene expression in a wide variety of bacterial species (22) and appear to form a family of sigma factors with related promoter specificities.

While a large number of genes involved in *fla/che/mot* functions had been identified and mapped in enteric bacteria, only a few of the corresponding *B. subtilis* genes were known, and this made it more difficult to determine the role of σ^D in their expression. Consequently, we set out to sequence a number of *B. subtilis* DNA fragments, each of which is known to contain a strong σ^D promoter (denoted a P_D), in the hope of obtaining information about the genes controlled by σ^D . This approach led to the identification of the *B. subtilis* *hag* gene, encoding the flagellar filament structural protein flagellin, as a monocistronic transcription unit transcribed only by σ^D RNA polymerase (35). As we show in this report, this approach has now led to the identification of two genes transcribed in a σ^D -dependent operon that are required for the motility of *B. subtilis*. These two *mot* genes are homologous in structure and function to the *motA* and *motB* genes of enteric bacteria.

MATERIALS AND METHODS

Bacterial strains, media, growth, and transformation. The *E. coli* host for growth of recombinant plasmid and M13 DNA was either strain TG-2 [$\Delta(lac-proAB)$ *supE thi hsdD5* $\Delta(srl-recA)306::Tn10(Tet^r)$ EcoK[−] F⁺ *traD36 proAB lacI^rZ* Δ M15] or DH5 α [F[−] ϕ 80*dlacZ* Δ M15 $\Delta(lacZYA-argF)$ *lacU169 recA1 endA1 hsdR17 supE44 thi-1 gyrA relA1* λ^-]. *E. coli* was grown in Luria broth medium (32) supplemented with ampicillin at 50 μ g/ml when appropriate. Transformation of *E. coli* was performed by standard published procedures (32).

The *B. subtilis* strains used in this study are listed in Table 1. Strains were grown in Penassay broth and on tryptone blood agar base plates. Transformation of strain CB25 by plasmid or chromosomal DNA was performed by standard procedures (2). Cm^r strains were maintained with 5 μ g of chloramphenicol per ml, and Ery^r strains were maintained

* Corresponding author.

TABLE 1. *B. subtilis* strains used in this study^a

Strain name	Relevant genotype
CB25	
CB147	<i>motAΔ12</i>
CB148	<i>motBΔ21</i>
OI1085	
OI2282	Tn917Ω1831
OI2283	Tn917Ω1833
OI2298	Tn917Ω1835
OI2300	Tn917Ω1847
OI2301	Tn917Ω1848
OI2302	Tn917Ω1851
OI2357	Tn917Ω1817

^a The CB strains are derived from the parent CB25 (originally JH642; Mot⁺ Che⁺ *trpC2 pheA1*) and were constructed as described in the text. The OI strains are derived from strain OI1085 (Mot⁺ Che⁺ *hisH2 metC trpF7*) and were described previously (54). The OI strains were kindly provided by G. Ordal.

with 1 µg of erythromycin per ml and 25 µg of lincomycin per ml.

Motility and flagellation assays. The motility of strains was assessed primarily by wet-mount microscopy of a sample of a culture grown for 6 h at 37°C in SPI medium (2), which consists of Spizizen salts, 0.02% Casamino Acids, 0.1% yeast extract, 0.5% glucose, 0.1 mM MnSO₄, and 1 µM FeSO₄, supplemented with auxotrophic requirements to 50 µg/ml. Swarm plates consisted of 1% Bacto Tryptone, 0.5% NaCl, and 0.27% Bacto Agar. Swarming was measured by the rate of increase of the swarm radius over time at 37°C, after inoculating a fresh colony into the swarm plate. Agglutination by anti-flagellin polyclonal antibody was performed as described previously (54). This antibody was provided by G. Ordal.

DNA manipulations. Small quantities of plasmid DNA were prepared by the alkaline lysis method of Birnboim and Doly (6). Large quantities of DNA were prepared by a modification of this procedure involving an additional LiCl precipitation. Restriction, ligation, and related manipulations of plasmid DNA were performed by published procedures (32). DNA fragments were isolated from polyacrylamide gels via overnight elution into buffer consisting of 2 mM Tris-HCl (pH 8.0)–0.2 mM EDTA and from agarose gels by treatment with NaI and Glassmilk with a GeneClean kit (Bio 101, Inc.). *B. subtilis* chromosomal DNA was prepared by a modification of published procedures (4).

DNA sequencing and analysis. A 2.74-kb segment of the 3.2-kb total insert from the original plasmid pCD4322 (47) was subcloned in two fragments, one a 1.5-kb *Hind*III fragment in pBR322, giving the plasmid pMG201 (17), and the other a 1.2-kb *Hind*III-*Pst*I fragment in the vector pBS-KS⁺ (Stratagene), giving the plasmid pVL202 (see Fig. 1). The sequence of both strands of the 2,736 bp of pCD4322 contained on these two subclones was obtained by the chain termination method of Sanger et al. (40) and is presented in Fig. 2. Sequencing reactions were carried out with Sequenase 2.0 (United States Biochemical) as recommended by the manufacturer. Either single-stranded DNA subcloned into M13mp18 and mp19 (52) or supercoiled, double-stranded plasmid DNA (10) was used as the template. Nested deletions generated by digestion with exonuclease III and S1 nuclease were performed by modified versions of published procedures (4, 25, 32). Primers for sequencing were the universal M13 primer (21) and (23–25)-mer oligonucleotides synthesized on a Biosearch 8750 DNA synthe-

sizer (MilliGen/Biosearch). Some oligonucleotides were purified by polyacrylamide gel electrophoresis followed by elution and ethanol precipitation, while the remainder were used in sequencing reactions without this purification procedure. DNA sequences were compiled, analyzed, and searched against GenBank and EMBL data banks by using IntelliGenetics programs (8, 9).

Southern blotting and hybridization analysis. Hybridization analysis of DNA restriction fragments was performed as described by Maniatis et al. (32) and by Southern (43). ³²P-labeled DNA fragments used as probes were prepared by nick translation of the appropriate purified DNA with [α -³²P]dCTP (Dupont/NEN Research Products or Amersham). The labeled DNA was separated from unincorporated nucleotides by passage through a Sephadex G-50 (Pharmacia Fine Chemicals) spin column. Hybridization conditions and nitrocellulose filter washing procedures were performed as previously described (35).

PCR. The polymerase chain reaction (PCR) was performed according to methods described by Innis et al. (26). Ten nanograms of *B. subtilis* chromosomal DNA and 20 pmol each of two oligonucleotide primers were used in a 100-µl reaction mixture containing all four deoxynucleoside triphosphates at 0.2 mM each, 10 mM Tris (pH 8.2), 30 mM KCl, 1.5 mM MgCl₂, 0.1 mg of gelatin per ml, and 2.5 U of AmpliTaq DNA Polymerase (Perkin-Elmer Cetus). The reaction was overlaid with washed light mineral oil, heated to 94°C for 5 min, and then subjected to 25 cycles in a Thermal Cycler (Perkin-Elmer Cetus). Each cycle consisted of 1 min of denaturation at 94°C, 2 min of annealing at 40°C, and 3 min of extension at 72°C. Subsequently, the mineral oil was removed by chloroform extraction and the reaction products were analyzed by gel electrophoresis.

Construction of in-frame deletion mutations. The creation of in-frame deletions of each open reading frame (ORF) was achieved by limited exonucleolytic cleavage by *Bal* 31 exonuclease of plasmid DNA linearized at a single restriction site located within the ORF.

Bidirectional deletion of *motA* was performed at the *Hinc*II site located within that ORF (see Fig. 1). The 1.5-kb *Hinc*III fragment of pMG102 had previously been subcloned into the pUC18-based integrational vector pJM102 (24), giving the plasmid pEAK2 (42) (see Fig. 1). The polylinker *Hinc*II site of pEAK2 was removed by digestion with *Sal*I, filling-in with Klenow fragment DNA polymerase, and self-ligation, to create the intermediate plasmid pEAK2-S. Next, pEAK2-S was digested at the unique *Hinc*II site, treated with exonuclease *Bal* 31 at 37°C under conditions such that several hundred base pairs were removed (4), and then self-ligated and transformed into *E. coli*. The extent of *Bal* 31 digestion was assessed by sequencing a number of resulting plasmid clones. One such plasmid, pMOTΔ12, possesses a deletion of 336 bp (from bases 1,022 to 1,357 inclusive) or 112 codons of the *motA* ORF (see Fig. 3).

To aid in the recombination of the insert DNA flanking the deletion, several other plasmids containing the entire operon were constructed. pDM203 and pMOTΔ12B were created by ligating the *Hind*III inserts of pEAK2-S and of pMOTΔ12, respectively, into *Hind*III-digested pVL202. Restriction mapping and sequencing confirmed that pDM203 and pMOTΔ12B contained the correct *Hind*III junction and possessed the entire transcription unit shown in Fig. 1. Finally, the integrational vector pMOTΔ12B-INT was created by isolating the full insert of pMOTΔ12B as an *Xba*I-*Xho*I fragment, filling-in with Klenow fragment DNA polymerase, and ligating into *Sma*I-digested pJM102.

To create the *motB* ORF in-frame deletion, pVL202 was digested with *Stu*I at its unique site within *motB* (see Fig. 1), treated with *Bal* 31 nuclease as described above, self-ligated, and transformed into *E. coli*. The resulting plasmids were sequenced, and candidates were identified. One deletion was subcloned into pJM102 as a *Hind*III-*Pst*I fragment. This final plasmid, pMOTBA21, possesses a deletion of 420 bp (from bases 1,807 to 2,226 inclusive) or 140 codons of the *motB* ORF (see Fig. 3).

Introduction of in-frame deletions into *B. subtilis*. Creation of *B. subtilis* strains bearing in-frame deletions in *motA* or *motB* was accomplished by the gene replacement technique first used for the alkaline and neutral protease genes (45, 51). In short, a nonreplicating plasmid containing a gene deletion is integrated at the gene locus by Campbell-type, single-crossover recombination, such that both a wild-type and a deleted copy of the gene are inserted as well as selectable antibiotic resistance and vector sequences. In some cases, a double-crossover event can occur so that two deleted copies of the gene of interest replace the wild-type version. Subsequently, the replacement of the wild-type gene is effected by homologous recombination of the directly repeated plasmid sequences, with the concomitant loss of the antibiotic resistance marker, as described previously (45).

To create strains bearing the *mot* gene deletions, *B. subtilis* CB25 was transformed to Cm^r by integration of pMOTA Δ 12B-INT or pMOTBA21 plasmid DNA. A Mot^- Cm^r transformant of pMOTA Δ 12B-INT, shown by PCR and Southern hybridization analysis to possess two deleted copies of the *motA* gene (*motA* Δ 12 alleles), and a Mot^+ Cm^r transformant of pMOTBA21 containing both a wild-type and a deleted *motB* gene were used as the strains to be reverted to Cm^s .

Reversion of the integrant strains from Cm^r to Cm^s was achieved by growth and passage for several days in Penassay broth without chloramphenicol and screening for Cm^s colonies on replica plates. Cm^s revertants were found and tested for motility. The nonmotile strains CB147 and CB148 were shown to possess only the in-frame deleted allele, *motA* Δ 12 or *motB* Δ 21, respectively, by PCR and Southern hybridization analysis.

Nucleotide sequence accession number. The DNA sequence presented in Fig. 2 has been submitted to GenBank under the accession number M77238.

RESULTS AND DISCUSSION

Nucleotide sequence analysis of the *B. subtilis* *mot* region. We have previously described the isolation of a collection of plasmids containing *B. subtilis* DNA fragments bearing strong in vitro promoters for *B. subtilis* σ^D holoenzyme (16, 47). One plasmid, pCD4322, was found to contain a strong P_D , designated P_{D-2} , reading into a relatively efficient in vitro transcription terminator, giving a transcript of about 1.6 kb. Northern (RNA) blot analysis showed that this transcription unit gave a single mRNA of the same size in vivo (15). Transcription from P_{D-2} initiates from a common site in vivo and in vitro and is entirely σ^D dependent (15). The P_{D-2} transcription unit was mapped to a region near the *spo0E* locus at 120 degrees on the *B. subtilis* chromosome (14).

The region of the *B. subtilis* chromosome (2,736 bp) containing this transcription unit was sequenced from a *Hind*III site 0.8 kb upstream of the promoter through a *Pst*I site 0.3 kb downstream of the terminator (Fig. 1). Both DNA strands were sequenced by the Sanger method, subcloning fragments into M13mp18, M13mp19, or pBS-KS⁺. The

nucleotide sequence is presented in Fig. 2. The P_{D-2} promoter region and transcription start site were previously mapped and sequenced (17) and are noted in Fig. 2. A sequence characteristic of a typical rho-independent terminator (7) is found approximately 1,640 nucleotides downstream of the promoter; this distance is similar to that determined for the transcription unit in vivo and in vitro (15). This termination signal had been shown to function efficiently in vitro (47) and confirms its assignment as such. The terminator is noteworthy in that the RNA stem formed is not GC rich, although the sequence of 7 U residues following the RNA hairpin is characteristic of efficient rho-independent terminators (50).

Within this transcription unit are found two successive, long ORFs (Fig. 2). Each has a strong consensus to the gram-positive Shine-Dalgarno sequence coupled with an AUG translational start codon in the correct position (19), and hence each is a likely protein-coding sequence. The deduced amino acid sequences are homologous to *E. coli* *mot* proteins (Fig. 3); the first reading frame corresponds to the *motA* gene product of *E. coli*, and the second corresponds to *motB*. The *E. coli* *motA* and *motB* gene products act as integral membrane proteins required for the operation of the flagellar motor (11, 13, 44, 48, 49). Comparison shows that approximately 27% of the amino acids are identical for the *E. coli* and *B. subtilis* homologs, and the homology increases to 39% for plausible conserved substitutions (Fig. 3). The degree of homology seen for these *mot* gene products is similar to that seen for other flagellar assembly and chemotaxis genes when *B. subtilis* and enteric counterparts are compared (1, 53).

Predicted protein structural features also suggest that these *B. subtilis* *Mot* proteins may be homologs of *E. coli* *MotA* and *MotB*. Hydropathy profiles, determined by the method of Kyte and Doolittle (30), indicate that the distributions of potentially hydrophilic or hydrophobic regions along the primary sequence of the respective *Mot* proteins from *E. coli* and *B. subtilis* are very similar (data not shown). For example, the two *MotB* sequences both contain a single long stretch of hydrophobic residues near the amino terminus of the peptide; for *E. coli* *MotB*, this region is believed to span the inner membrane (11, 44). Other similar features between the *MotB* peptides include a very basic amino terminus and a proline-rich carboxy terminus (44).

Figure 2 illustrates that the *motA* and *motB* coding sequences overlap extensively, suggesting that the two proteins may be translationally coupled. In *E. coli*, these two genes are the first two cistrons of the *mocha* operon which also contains the *cheA* and *cheW* genes (13, 41, 44). In *B. subtilis*, the *cheA* and *cheW* genes are found upstream of the *sigD* gene at the promoter-distal end of a long operon of flagellar and chemotaxis genes (33).

Mapping of transposons from Mot^- strains. The sequence homology between the *MotA* and *MotB* proteins of *E. coli* and the two σ^D -dependent gene products of *B. subtilis* suggests a role for these gene products in motility. Zuberi et al. (54) recently reported the identification of six Tn917*lacZ* insertions (some of which showed *sigD*-dependent *lacZ* expression), mapping near *spo0E*, that conferred a Mot^- phenotype. Our results, taken with the mapping information (14), strongly suggested that these insertions would map in the *motAB* operon we had identified.

We first used Southern blot analysis to map roughly the sites of these insertions relative to our *B. subtilis* *mot* operon. We were able to locate the sites of insertion of all six transposons to DNA contained on pMG201 and pVL202

FIG. 1. Physical map of the *B. subtilis* *motAB* operon. The segment of the *B. subtilis* chromosome known to contain the σ^D -dependent promoter P_{D-2} and transcription terminator T, subcloned on plasmid pCD4322, is shown as a thick horizontal line; several restriction sites are indicated. The locations of the two ORFs encoding MotA and MotB are shown as shaded boxes. The pCD4322 subclones pMG201 and pVL202, which were used for DNA sequencing, are shown below the original isolate to indicate the segments of the original DNA they possess. The inserts of other plasmids used in this study are also shown and are described in more detail in the text. For the plasmids containing in-frame deletions (pMOT Δ 12, pMOT Δ 12B, pMOT Δ 12B-INT, and pMOTB Δ 21), the deleted region is shown as a bubble; the number of codons deleted is indicated within the bubble. The locations of the six insertions of Tn917*lacZ* (symbolized by triangles) into the chromosome which result in a Mot⁻ phenotype (54) are indicated by the point of contact of the lines connecting the triangles to the map of the chromosomal DNA. The small black arrow within the triangle indicates the direction of transcription and translation of the *E. coli lacZ* gene present on the transposons; this gene served to determine the orientation of insertion of the transposons, as described in the text. The transposon insertions are labeled as described previously (54).

(Fig. 1), all on an approximately 6.6-kb *EcoRI-EcoRI* fragment of the chromosome. Zuberi et al. (54) originally estimated this fragment to be 5.7 kb on the basis of the size of transposon-bearing DNA hybridizing to transposon sequences, whereas we have used *mot* DNAs as hybridization probes of wild-type chromosomal DNA. Zuberi et al. (54) also reported that Tn917 Ω 1835 had integrated into a 2.2-kb *EcoRI* fragment of the chromosome; our data suggest that, in fact, this transposon lies on the same 6.6-kb *EcoRI* fragment as the other Tn917 insertions. We are currently uncertain as to the cause of this discrepancy.

Zuberi et al. mention that although a seventh transposon insertion, Tn917 Ω 1831, was found to lie in an *EcoRI* fragment of the same size as those described above and mapped to the same region of the chromosome, it had inserted into a different, *sigD*-independent transcription unit (54). We have found that this transposon does not lie on the chromosome on a segment spanned by pMG201 or pVL202 and thus does not disrupt the *motAB* operon (data not shown). Strain OI2282, which carries Tn917 Ω 1831, is in fact motile but shows defects in chemotaxis to aspartate and mannitol (54). Therefore, the genetic locus disrupted by Tn917 Ω 1831, called *cheX*, may contain new and previously unidentified

genes affecting chemotaxis in *B. subtilis* and deserves further study.

We used the PCR method to localize more precisely the sites of insertion of these Tn917*lacZ* transposons. We designed an oligonucleotide primer which was complementary to sequences in the *E. coli lacZ* gene contained on this transposon, whose 5' end was 436 bp from the *erm*-proximal end of Tn917*lacZ* (38). PCR experiments using this primer in combination with one of the several primers used for sequencing were performed to amplify the DNA spanning the junction of the chromosome and the *erm*-proximal end of the transposon. Analysis of the sizes of these amplified fragments coupled with information about the sequencing primer used allowed us to determine both the sites and the orientations of insertion of the six transposons, as depicted in Fig. 1. The two transposons that had been shown to display *sigD*-dependent *lacZ* expression, namely Tn917 Ω 1833 and Tn917 Ω 1817 (54), are in fact oriented so that the *lacZ* gene is transcribed from P_{D-2} .

An insertion of Tn917*lacZ* at the *motAB* locus could cause a Mot⁻ phenotype in several ways. The insertion Tn917 Ω 1851 is within the promoter region and thus most likely disrupts the transcription of the entire operon. In the

1 AAGCTTCAAAGACACTGTGATCATCATGACAAGTAATGCGGGTCTGGTGAGAAACAAACGAAAGTCGGTTTCCAATCAGATGACAGTGTTCATCGAAGAA

101 CAAACATTAATTGATTCACTGAGCATGTTCTTTAAACCTGAGTTCCTCAACCGTTTTGACAGCATTATTGAGTTCGGCTCATTGGAAAAAGAACATCTTG

201 TCAAATCGTCAGCCTTCTTCTTGAGAACTTGAAGAAACATTGGCGGAACGGGGCATTAGCTTGAATGTGACAGATGAAGCGAAAGAAAAAATCGCTGA

301 GCTGGGTACCAACCTTCATTTCGGTGCACGTCCGCTTAGAAGAACCATCCAAGAATGGGTTGAGGATGAAATGACCGATCTGCTGCTTGATAATGGCGAG

401 ATCACAAGTTTTACGTGATTTTAGAAGATGATAAAATCAAAGTGCAGCAAAATAACAATCAGCGGTTTCTTTTAGGAAGCCGCTTTTTTTTATACTT

501 TCATACTCTTGACTTAGAAATTACCGTCCCTTCTCTGAAGCGCTACTTCCCTTCATCTTTACACTTTTTTAAGGAGGGATGGAACATGTTTATATACA

601 GTGAATATGTATTTCTACTATTATGAATACCATCTGAAGGACAGCTTGCAAGCAAGCTGCTAATTTTCAGTAGACCCCGGGCTTTTCAATTTTGGAGA

-35 PD-2 -10

701 GCCCGTTTTTGTATTCTCCCTTCTATGCCTCTTCTTGAATATTTTCATGAACAAATTCACAATGTCCCTAAAGTTCGGGGACCAAAACCGATATTA

+1 RBS motA

----- M D K T S L I G I I L A F V A L S V G M V

801 ACCATAGACAAGCTAGTAAAAAGGATTTGGTGAACACTATGGATAAACTTCGTTAATCGGTATTATTCTTGCTTTTGTGGCATTGAGCGTCGGGATGG

L K G V S F S A L A N P A A I L I I I A G T I S A V V I A F P T K

901 TTCTGAAAGGCGTCAGTTTCAGCGCCCTTGCAACCCCGCTGCCATTTAATTATTATCGCCGGGACAATCTCAGCAGTCGTTATTGCGTTCCCAACAAA

E I K K V P T L F R V L F K E N K O L T I E E L I P M F S E W A Q

1001 AGAAATTAAGGATGCCTACGCTGTTTCGAGTGTATTATTAAGGAAATAAACAGCTCACAATAGAGGAATCATTCTATGTTCTCTGAATGGGCTCAG

L A R R E G L L A L E A S I E D V D D A F L K N G L S M A V D G Q S

1101 CTTCGACGCCCGAAGGTCTGCTTGTCTTAGAAGCAAGCATTGAGGATGTAGATGATGCTTTCTTGAAAAACGGGCTCAGCATGGCTGTGACGGGCAAA

A E F I R D I M T E E V E A M E D R H Q A G A A I F T Q A G T Y A

1201 GCGCGAATTATTAAGAGATATTATGACAGAGGAAGTCGAAGCAATGGAGGACAGGCACCAAGCAGGAGCCGCTATTTTACACAAGCAGGAACGTACGC

P T L G V L G A V I G L I A A L S H M D N T D E L G H A I S A A F

1301 TCCGACACTTGGAGTACTCGGCGCTGTAATCGGGCTGATTGCCGCTCTCTCTCATATGGATAACACAGATGAGCTTGGACACGCCATCAGTGTCTGCTTT

V A T L L G I F T G Y V L W H P F A N K L K R K S K Q E V K L R E V

1401 GTTGCCACACTTCTCGGTATCTTTACAGGTATGTGTTATGGCATCTTTTCGAAATAAATTAACGAAATCAAAACAGGAAGTAAACTGCGTGAGG

M I E G V L S V L E G Q A P K V I E Q K L L M Y L P A K D R L K F

1501 TCATGATTGAAGGTGTTTATCCGTTTGAAGGACAAGCACCGAAAGTCATCGAACAAAGCTTTAATGTATCTTCTCGAAGGACCGCTTGAAT

RBS motB

----- M A R K K K K K H E D E H V D E S W L V P Y A D I L

1601 TGCAGACAAGGAGAGCGCAAAATGGCGAGAAAAAGAAAGAAGCATGAGGACGAGCACGTTGATGAATCATGGCTCGTCTCTTACGCCGACATCTT

T L L L A L F I V L Y A S S S I D A A K F Q M L S K S F N E V F T

1701 TACTCTTCTCTGGCATTGTTTATTGTGCTGTACGCGAGCAGCTCGATTGACGCGCTAAGTTTCAAATGCTCTCAAATCATTTAATGAAGTTTTTACA

G G T G V L D Y S S V T P P E N E S D G I D E V K K E K E K E K N

1801 GCGGAACCGGTGTACTGGACTACTCCAGTGTCACTCTCTGAAACGAGTCAGACGGCATCGATGAAGTGAAGGAAAAAGAAAGAGAAAGAGAAAA

K K E K E K A A D Q E E L E N V K S Q V E K F I K D K K L E H Q L

1901 ACAAGAAAGAAAAAGAAAGCAGCTGATCAAGAAGAACTTGAAATGTGAAGACCGGTTGAAAGTTTCATCAAGATAAAAGCTGGAACATCAGCT

E T K M T S E G L L I T I K D S I F F D S G K A T I R K E D V P L

2001 GGAGACGAAGATGACTAGTGAAGGCCTTCTGATTACGATTAAGACAGCATCTTCTTCGATTCTGGAAAAGCGACCATCCGTAAGGAAGATGTGCCGCTT

A K E I S N L L V I N P P R N I I I S G H T D N M P I K N S E F Q S

2101 GCAAAAGAGATTTCAAATCTTCTGTGATTAAACCGCAAGAAATATCATTATCAGCGGACATCTGATAATATGCCAATTAATAATCTGAATTCCAAT

N W H L S V M R A V N F M G L L I E N P K L D A K V F S A K G Y G

2201 CAAACTGGCATTAAAGCGTCATGAGAGCGGTAACCTTTATGGGGCTCTGATTGAAAAACCCCAAGCTCGATGCAAAAGTGTTCAGCGCAAGGGTTATGG

E Y K P V A S N K T A E G R S K N R R V E V L I L P R G A A E T N

2301 CGAGTATAAACCGGTGGCTTCCAATAAACTGCGGAAGGCCGAAGCAAAAAACCGCGGGTTGAAGTTCTCATTTTGGCGAGAGGCGCAGCGGAAACAAAT

E K . >>>>>>> <<<<<<<<

2401 GAAAAATAGCAAAAAAGGAGCCTTGTGACATATCAGGCTTCCTTTTTTACTTATTATGAACTGTTTTGCGTTTTAATCTACTCTTTTTTAACATTTCG

2501 GTGCACATCTCTACCTCTTCTTCGCTTAATACGCTGATCATTTTCATCAATTCAGCAGCGTATCAGGGAAAAATACGTTCAAAAGACCGATGCCTTTTT

2601 CAGTAATTTGCGCAATGTCACTCGTCTGTCTGTAGGAGACGCTTCCGAATGAGAAGTTCTTTTTGCTCCAGCTTATCTACAACGTATGTGATCGTCCC

2701 GCTAGCCAAGAGAATTTTATCTCCTATCTGCTGCAG

FIG. 2. Nucleotide sequence of the *B. subtilis* *motAB* operon. The 2,736 nucleotides of pCD4322 that are contained on pMG201 and pVL202, starting from the left *Hind*III site and proceeding to the right *Pst*I site (in Fig. 1), are shown. The features of the σ^D -dependent promoter PD-2 known to transcribe this operon (the -35 and -10 promoter consensus regions and the transcription start site [+1]) are indicated by a single line above the nucleotide sequence. The putative translation initiation signals (RBS) for the *motA* and *motB* ORFs are designated by a double line. A region of dyad symmetry followed by a T-run, which is likely to serve as the transcription termination signal, is shown by the > and < symbols. The amino acid sequence of the conceptual translations of the *motA* and *motB* reading frames are shown above the nucleotide sequence, in the one-letter code.

In-frame deletions of *motA* and *motB* result in a loss of motility. To unambiguously assign functions to the two ORFs independently, it was necessary to create mutations in each that did not affect the expression of the other. For this reason, we chose to genetically engineer in-frame deletions of coding sequences internal to each ORF and use them to replace the wild-type versions of each gene. An in-frame deletion of a large central portion of one ORF should yield a

nonfunctional protein product while not interfering with the expression of the other gene in the operon.

We used the gene replacement technique developed for *B. subtilis* (45, 51) to create the strains CB147 and CB148, which contain the in-frame deletion alleles *motA* Δ 12 and *motB* Δ 21, respectively (Fig. 3). The *motA* Δ 12 allele consists of a deletion of 112 codons from the center of the *motA* ORF, and *motB* Δ 21 consists of a deletion of 140 codons from the center of the *motB* ORF; these deletions are likely to be loss-of-function mutations. CB147 and CB148 were observed to be completely nonmotile as assessed by light microscopy and swarm plates. However, these strains did agglutinate in the presence of anti-flagellin antibody, indicating that the strains were in fact flagellated. The Mot⁻ Fla⁺ phenotype of these strains is the same as that observed by Zuberi et al. for the transposon insertion mutants (54). Hence, the presence of either the *motA* Δ 12 or the *motB* Δ 21 allele results in a cell with paralyzed flagella and indicates that the products of the *motA* and *motB* genes are required for the motility of *B. subtilis*.

B. subtilis σ^D is representative of a family of σ factors having similar functions in many diverse bacterial species. The *B. subtilis* σ^D factor and enteric σ^F factors are homologous in promoter recognition specificity and also in protein structure (3, 37). Both factors are implicated in the transcription of the corresponding genes for flagellin (*hag*) and for motility proteins (*motA* and *motB*). In *E. coli* and *S. typhimurium*, these flagellin and motility genes are members of the class 3 set of *fla/che/mot* genes, in that their expression is dependent on both the *fliA* gene product (encoding σ^F) and the expression of the class 2 genes which encode the flagellar hook and basal body components (for a more detailed discussion of this genetic hierarchy, see references 23, 27, 29, and 31). The flagellin genes of *B. subtilis* (*hag*) and of *E. coli* (now called *fliC*) have been shown to be dependent in vivo on σ^D and σ^F , respectively (20, 35). Transcription of *B. subtilis* *motA* and *motB* is entirely σ^D dependent in *B. subtilis* (15, 17). However, there is no direct evidence that in *E. coli* these *mot* genes are transcribed from a *fliA*-dependent promoter in vitro or in vivo. A σ^{70} promoter consensus was originally proposed to direct transcription of this operon (13). The σ^F promoter consensus sequence and transcriptional start site proposed to control transcription of the *mocha* operon lie within the coding sequence of the upstream *fliHDC* operon and are separated from the translational start of *motA* by a putative transcription termination signal for the *fliHDC* operon (5, 13, 23). It is clear that differences in *fla/che/mot* gene expression and regulation exist between the enteric bacteria and *B. subtilis*.

On the other hand, the similarities in *fla/che/mot* gene expression among bacterial species have been very informative. The conservation of both recognition specificity and cellular function among such evolutionarily distant bacteria suggested that other bacterial species might have related sigma factors controlling flagellar functions (for a review, see reference 22). Indeed, recent reports confirm that flagellar structural genes in *Campylobacter* spp. (18, 36), *Pseudomonas aeruginosa* (46), and *Rhizobium meliloti* (39) are transcribed from promoters having the *B. subtilis* σ^D consensus and that there are similar sequences upstream of genes for outer membrane proteins in *Neisseria gonorrhoeae* (12).

The nomenclature used to designate secondary sigma factors is currently confusing; such factors in *B. subtilis* are designated σ^A , σ^B , and σ^D , etc., while in *E. coli* they are designated RpoH and RpoN, etc. We propose that the related family of σ factors discussed herein be referred to as

the σ^{28} family; σ^{28} was the original designation of *B. subtilis* σ^D (47). This nomenclature is in keeping with that used for the σ^{54} family of σ factors (28), which are also widely distributed among prokaryotes.

ACKNOWLEDGMENTS

We thank Robert Macnab for assistance in the identification of the *motB* homolog.

This work was supported by Public Health Service research grant GM12010 and training grant GM07232 from the National Institute of General Medical Sciences.

REFERENCES

- Albertini, A. M., T. Caramori, W. D. Crabb, F. Scoffone, and A. Galizzi. 1991. The *flaA* locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions, and an ATPase-like polypeptide. *J. Bacteriol.* 173:3573-3579.
- Anagnostopoulos, C., and J. Spizizen. 1961. Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 81:741-746.
- Arnosti, D. N., and M. J. Chamberlin. 1989. Secondary σ factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 86:830-834.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl. 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
- Bartlett, D. H., B. B. Frantz, and P. Matsumura. 1988. Flagellar transcriptional activators FlbB and FlaI: gene sequences and 5' consensus sequences of operons under FlbB and FlaI control. *J. Bacteriol.* 170:1575-1581.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.
- Brendel, V., and E. N. Trifonov. 1984. A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res.* 12:4411-4427.
- Brutlag, D. L., J. Clayton, P. Friedland, and L. H. Kedes. 1982. SEQ: a nucleotide sequence analysis and recombination system. *Nucleic Acids Res.* 10:279-294.
- Brutlag, D. L., J. P. Dautricourt, S. Maulik, and J. Relph. 1990. Improved sensitivity of biological sequence database searches. *Comput. Appl. Biosci.* 6:237-245.
- Chen, E. Y., and P. H. Seeberg. 1985. Laboratory methods. Supercil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* 4:165-170.
- Chun, S. Y., and J. S. Parkinson. 1988. Bacterial motility: membrane topology of the *Escherichia coli* MotB protein. *Science* 239:276-278.
- Connell, T. D., W. J. Black, T. H. Kawula, D. S. Barritt, J. A. Dempsey, K. Kverneland, Jr., A. Stephenson, B. S. Schepart, G. L. Murphy, and J. G. Cannon. 1988. Recombination among protein II genes of *Neisseria gonorrhoeae* generates new coding sequences and increases structural variability in the protein II family. *Mol. Microbiol.* 2:227-236.
- Dean, G. E., R. M. Macnab, J. Stader, P. Matsumura, and C. Burks. 1984. Gene sequence and predicted amino acid sequence of the *motA* protein, a membrane-associated protein required for flagellar rotation in *Escherichia coli*. *J. Bacteriol.* 159:991-999.
- Ferrari, F. A., E. Ferrari, and J. A. Hoch. 1982. Chromosomal location of a *Bacillus subtilis* DNA fragment uniquely transcribed by sigma-28-containing RNA polymerase. *J. Bacteriol.* 152:780-785.
- Gilman, M. Z., and M. J. Chamberlin. 1983. Developmental and genetic regulation of *Bacillus subtilis* genes transcribed by σ^{28} -RNA polymerase. *Cell* 35:285-293.
- Gilman, M. Z., J. S. Glenn, V. L. Singer, and M. J. Chamberlin. 1984. Isolation of sigma-28-specific promoters from *Bacillus subtilis* DNA. *Gene* 32:11-20.
- Gilman, M. Z., J. L. Wiggs, and M. J. Chamberlin. 1981. Nucleotide sequences of two *Bacillus subtilis* promoters used by *Bacillus subtilis* sigma-28 RNA polymerase. *Nucleic Acids Res.* 9:5991-6000.

18. Guerry, P., S. M. Logan, S. Thornton, and T. J. Trust. 1990. Genomic organization and expression of *Campylobacter* flagellin genes. *J. Bacteriol.* **172**:1853–1860.
19. 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*, vol. 2. Academic Press, Inc., New York.
20. Hanafusa, T., A. Sakai, A. Tominaga, and M. Enomoto. 1989. Isolation and characterization of *Escherichia coli* *hag* operator mutants whose *hag48* expression has become repressible by a *Salmonella* H1 repressor. *Mol. Gen. Genet.* **216**:44–50.
21. Heidecker, G., J. Messing, and B. Gronenborn. 1980. A versatile primer for DNA sequencing in the M13mp2 cloning system. *Gene* **10**:69–73.
22. Helmann, J. D. 1991. Alternative sigma factors and the regulation of flagellar gene expression. *Mol. Microbiol.* **5**:2875–2882.
23. Helmann, J. D., and M. J. Chamberlin. 1987. DNA sequence analysis suggests that expression of flagellar and chemotaxis genes in *Escherichia coli* and *Salmonella typhimurium* is controlled by an alternative σ factor. *Proc. Natl. Acad. Sci. USA* **84**:6422–6424.
24. Helmann, J. D., L. M. Márquez, and M. J. Chamberlin. 1988. Cloning, sequencing, and disruption of the *Bacillus subtilis* σ^{28} gene. *J. Bacteriol.* **170**:1568–1574.
25. Henikoff, S. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* **155**:156–165.
26. Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White. 1990. PCR protocols. Academic Press, San Diego, Calif.
27. Komeda, Y. 1986. Transcriptional control of flagellar genes in *Escherichia coli* K-12. *J. Bacteriol.* **168**:1315–1318.
28. Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* **53**:367–376.
29. Kutsukake, K., Y. Ohya, and T. Iino. 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J. Bacteriol.* **172**:741–747.
30. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105–132.
31. Macnab, R. M. 1990. Genetics, structure, and assembly of the bacterial flagellum, p. 77–106. In J. P. Armitage and J. M. Lackie (ed.), *Biology of the chemotactic response*, vol. 46. Cambridge University Press, Cambridge.
32. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. Márquez, L. M., J. D. Helmann, E. Ferrari, H. M. Parker, G. W. Ordal, and M. J. Chamberlin. 1990. Studies of σ^D -dependent functions in *Bacillus subtilis*. *J. Bacteriol.* **172**:3435–3443.
34. McLachlan, A. D. 1972. Repeating sequences and gene duplication in proteins. *J. Mol. Biol.* **64**:417–437.
35. Mirel, D. B., and M. J. Chamberlin. 1989. The *Bacillus subtilis* flagellin gene (*hag*) is transcribed by the σ^{28} form of RNA polymerase. *J. Bacteriol.* **171**:3095–3101.
36. Nuijten, P. J. M., F. J. A. M. van Asten, W. Gastra, and B. A. M. van der Zeijst. 1990. Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. *J. Biol. Chem.* **265**:17798–17804.
37. Ohnishi, K., K. Kutsukake, H. Suzuki, and T. Iino. 1990. Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. *Mol. Gen. Genet.* **221**:139–147.
38. Perkins, J. B., and P. J. Youngman. 1986. Construction and properties of Tn917-lac, a transposon derivative that mediates transcriptional gene fusions in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* **83**:140–144.
39. Pleier, E., and R. Schmitt. 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. *J. Bacteriol.* **171**:1467–1475.
40. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
41. Silverman, M., and M. Simon. 1976. Operon controlling motility and chemotaxis in *E. coli*. *Nature (London)* **264**:577–580.
42. Singer, V. L. 1987. Ph.D. thesis. University of California, Berkeley.
43. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
44. Stader, J., P. Matsumura, D. Vacante, G. E. Dean, and R. M. Macnab. 1986. Nucleotide sequence of the *Escherichia coli* *motB* gene and site-limited incorporation of its product into the cytoplasmic membrane. *J. Bacteriol.* **166**:244–252.
45. Stahl, M. L., and E. Ferrari. 1984. Replacement of the *Bacillus subtilis* subtilisin gene with an in vitro-derived deletion mutation. *J. Bacteriol.* **158**:411–418.
46. Totten, P. A., and S. Lory. 1990. Characterization of the type A flagellin gene from *Pseudomonas aeruginosa* PAK. *J. Bacteriol.* **172**:7188–7199.
47. Wiggs, J. L., M. Z. Gilman, and M. J. Chamberlin. 1981. Heterogeneity of RNA polymerase in *Bacillus subtilis*: evidence for an additional σ factor in vegetative cells. *Proc. Natl. Acad. Sci. USA* **78**:2762–2766.
48. Wilson, M. L., and R. M. Macnab. 1988. Overproduction of the MotA protein of *Escherichia coli* and estimation of its wild-type level. *J. Bacteriol.* **170**:588–597.
49. Wilson, M. L., and R. M. Macnab. 1990. Co-overproduction and localization of the *Escherichia coli* motility proteins MotA and MotB. *J. Bacteriol.* **172**:3932–3939.
50. Yager, T. D., and P. H. von Hippel. 1987. Transcript elongation and termination in *Escherichia coli*, p. 1241–1275. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
51. Yang, M. Y., E. Ferrari, and D. J. Henner. 1984. Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an in vitro-derived deletion mutation. *J. Bacteriol.* **160**:15–21.
52. Yanisch-Perron, C., 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.
53. Zuberi, A. R., C. Ying, D. S. Bischoff, and G. W. Ordal. 1991. Gene-protein relationships in the flagellar hook-basal body complex of *Bacillus subtilis*: sequences of the *flgB*, *flgC*, *flgG*, *fliE*, and *fliF* genes. *Gene* **101**:23–31.
54. Zuberi, A. R., C. Ying, H. M. Parker, and G. W. Ordal. 1990. Transposon Tn917lacZ mutagenesis of *Bacillus subtilis*: identification of two new loci required for motility and chemotaxis. *J. Bacteriol.* **172**:6841–6848.