JOURNAL OF BACTERIOLOGY, Apr. 1996, p. 1936–1945 Vol. 178, No. 7
0021-9193/96/$04.00+0

Identification, Nucleotide Sequence, and Characterization of PspF, the Transcriptional Activator of the *Escherichia coli* Stress-Induced *psp* Operon

GORAN JOVANOVIC, LORIN WEINER,† AND PETER MODEL*

Rockefeller University, New York, New York 10021

Received 15 November 1995/Accepted 26 January 1996

The phage shock protein (*psp*) operon (*pspABCE*) of *Escherichia coli* is strongly induced in response to a variety of stressful conditions or agents such as filamentous phage infection, ethanol treatment, osmotic shock, heat shock, and prolonged incubation in stationary phase. Transcription of the *psp* operon is driven from a σ^{54} promoter and stimulated by integration host factor. We report here the identification of a transcriptional activator gene, designated *pspF*, which controls expression of the *psp* operon in *E. coli*. The *pspF* gene was identified by random miniTn*10-tet* transposon mutagenesis. Insertion of the transposon into the *pspF* gene abolished σ^{54}-dependent induction of the *psp* operon. The *pspF* gene is closely linked to the *psp* operon and is divergently transcribed from one major and two minor σ^{70} promoters. *pspF* encodes a 37-kDa protein which belongs to the enhancer-binding protein family of σ^{54} transcriptional activators. PspF contains a catalytic domain, which in other σ^{54} activators would be the central domain, and a C-terminal DNA-binding domain but entirely lacks an N-terminal regulatory domain and is constitutively active. The insertion mutant *pspF*::mTn*10-tet* (*pspF*$_{877}$) encodes a truncated protein (PspFΔHTH) that lacks the DNA-binding helix-turn-helix (HTH) motif. Although the central catalytic domain is intact, PspFΔHTH at physiological concentration cannot activate *psp* expression. In the absence of inducing stimuli, multicopy-plasmid-borne PspF or PspFΔHTH overcomes repression of the *psp* operon mediated by the negative regulator PspA.

The phage shock protein (*psp*) operon of *Escherichia coli* is specifically and continually induced by gene IV protein (pIV) of filamentous phages (8, 47). Psp genes are also synthesized transiently in response to several stresses such as heat shock, osmotic shock, and ethanol treatment (8). Inhibition of protein secretion or lipid biosynthesis, treatment with ionophores or free fatty acids, and prolonged stationary-phase incubation induce the *psp* operon as well (5, 11, 28, 59). Heat shock, osmotic shock, and ethanol treatment also induce the heat shock response in *E. coli* (33, 60). Although the *psp* operon is induced by many stimuli that also elicit the heat shock response, *psp* expression does not require the heat shock sigma factor, σ^{32} (9). Indeed, in σ^{32} mutants, *psp* expression is elevated during unstressed growth, and the response to heat or ethanol shock is prolonged (8, 9, 56). The function of the *psp* operon in *E. coli* physiology is not clear, but bacteria that lack the *psp* operon are less able to survive prolonged incubation in stationary phase at pH 9.0, show increased motility (58, 59), and exhibit a reduction in the efficiency of protein translocation (27a, 28).

The *psp* operon consists of four genes, *pspABCE*, and its expression is controlled principally at the transcriptional level (9, 28, 57). All *psp* transcription is driven by the alternative holoenzyme form of RNA polymerase (RNAP) containing the σ^{54} factor (σ^{54}-RNAP) (56). As has been previously shown, this polymerase, in most cases, forms a closed complex with the promoter and is unable to initiate transcription by itself (14, 29). Initiation of transcription requires the presence of a transcriptional activator that catalyzes the isomerization of the closed promoter complex to an open one (14, 29, 36, 38, 43). Typically, activator proteins bind to a specific upstream activation sequence (UAS) located 80 to 200 bp upstream of the σ^{54} core promoter (10, 14, 44). The function of the UAS is to tether the activator in the right position and to bring it in the vicinity of the promoter in order to increase the fidelity and efficiency of interaction between the σ^{54}-RNAP and the activator (13, 14, 18, 30, 35, 36, 44). Expression from the *psp* promoter requires a UAS (located between positions −125 and −76) which has the ability to activate at a distance (57). In many instances, interaction of the σ^{54}-RNAP bound to the promoter and the activator bound to the UAS is facilitated by DNA bending mediated by integration host factor (IHF), which binds to a site between the UAS and the core promoter region (14, 19, 21, 35, 42). Transcription of the *psp* operon is also stimulated by IHF in vivo (57). IHF binds to a site between the UAS and σ^{54}-RNAP consensus sequence of the *psp* operon (57).

In the absence of inducing stimuli, PspA is transcribed at a low level and negatively regulates *psp* expression (56). During induction, the *psp* genes are activated by the cooperative action of PspB and PspC (56). Mutations in *pspA* result in constitutive expression of the operon without any stressful conditions or agents (56). Hence, the function of PspB and PspC, which are positive regulators, is to antagonize repression mediated by PspA. PspA, -B, -C, and -E are not similar to any of the known two-component sensor-regulator proteins (2, 41, 53). Moreover, neither PspB nor -C, in the absence of PspA, is required for high-level expression from the *psp* σ^{54} promoter (56). Therefore, we postulated the existence of a putative σ^{54} transcriptional activator which could act through DNA binding and protein-protein interactions.

In this report, we describe the identification and molecular characterization of the *pspF* gene, which encodes the σ^{54} transcriptional activator of the *psp* operon and is a member of the

* Corresponding author. Mailing address: Rockefeller University, 1230 York Ave., New York, NY. 10021. Phone: (212) 327-8648. Fax (212) 327-7850.

† Present address: Department of Cell Biology, Harvard Medical School, 25 Shattuck St., Boston, MA 02115.

enhancer-binding protein (EBP) family. We have also isolated, mapped, and analyzed a *pspF*::mTn*10-tet* ($pspF_{877}$) chromosomal insertional mutation.

MATERIALS AND METHODS

Bacterial strains, phages, and plasmids. The *E. coli* strains K561 [HfrC λ^+ *relA1 spoT1* T_2^r (*ompF627 fadL701*) *lacI*^{q}], L57 [K561 *rpoN*::Tn*10* (Tc^r)], L108 [K561 *himD*Δ3::*cam* (Cm^r)], and MC1061 [F$^-$ *hsdR mcrB araD139* Δ(*araABC-leu*)*7697* Δ*lacI74 galV galK strA thi*] have been described previously (15, 56). K1472 [MC1061 Δ*pspABC*::*kan* (Kn^r)] and K1471 [MC1061 *rpoN*::Tn*10* (Tc^r)] were from our laboratory collection. C600 was kindly provided by Nancy Kleckner (Harvard University, Cambridge, Mass.). Strains K1522.1 to K1522.70 [K1472 mTn*10-tet* (Tc^r) {pLW38 [*bla* Φ(*pspA′-lac*)]}] were generated by consecutive transduction (λNK1323), random mTn*10-tet* mutagenesis, and transformation with plasmid pLW38. K1518 [K1472 *pspF*::mTn*10-tet* (Tc^r)], K1524 [MC1061 Δ*pspABC*::*kan pspF*::mTn*10-tet* (Kn^r Tc^r)], K1526 [MC1061 *pspF*::mTn*10-tet* (Tc^r)], K1528 [MC1061 mTn*10-tet* (Tc^r)], K1527 [K561 *pspF*::mTn*10-tet* (Tc^r)], and K1530 [MC1061 *himD*Δ3::*cam* (Cm^r)] were generated by P1 transduction. The *E. coli* host for growth of M13 DNA was strain XL1-blue {F′[*proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15* Tn*10* (Tc^r)]}.

The *E. coli* strains were grown at 37°C in rich medium (Luria broth) (34) or in minimal medium A (48). Minimal medium A was supplemented with glucose (0.2%), required amino acids (30 μg/ml), and thiamine (1 μg/ml). When appropriate, antibiotics were used at the following concentrations: ampicillin, 100 μg/ml; tetracycline, 15 μg/ml; kanamycin, 25 or 30 μg/ml; chloramphenicol, 25 μg/ml. For detection of β-galactosidase activity, cultures were induced with a final concentration of 7 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and plates contained 0.15% of X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside).

The bacteriophages used were λNK1323 (kindly provided by Nancy Kleckner), M13mp18, and $P1_{vir}$. Transductions were performed as previously described (34).

Plasmid pLW38 has a 0.6-kb *Bsm*I fragment (containing the *psp* promoter region 500 bp upstream from the *pspA* and the first 12 codons of *pspA*) fused in frame to codon 7 of the *lacZ* gene on plasmid pSKS107 (56). pBRPS-1 was constructed by ligating a 4.5-kb *Eco*RI genomic fragment containing the complete *psp* operon to *Eco*RI-restricted pBR322 (9). Plasmid pNAL200 (from our laboratory collection) contains f1 gene IV recloned from plasmid pJLB4 (7) digested with *Bam*HI and ligated into the *Bam*HI site of plasmid pGZ119EH in which the pIV gene is under the control of the IPTG-inducible *Taq* promoter (32). The plasmids pMJ1 and pMJ2 used for sequencing were made by ligating the 1.0-kb *Bgl*II-*Hin*cII fragment or a 1.4-kb *Eco*RI-*Hin*cII fragment (both derived from a 2.4-kb *Eco*RI-*Bgl*II fragment of pBRPS-1) into the *Bam*HI-*Hin*cII or *Eco*RI-*Bgl*II site of pUC19, respectively.

Plasmid pMJ3, used for sequencing and complementation experiments, contains the insertion mutant *pspF*::mTn*10-tet*. It includes the *pspF* promoter region, the first 877 bp of the *pspF* coding region, and 51 bp of the miniTn*10-tet* inverted repeat. To construct it, K1527 chromosomal DNA was amplified by using a synthetic oligonucleotide primer (IR1070Bam) that annealed to the miniTn*10-tet* inverted repeat (positions 35 to 55) and created a *Bam*HI site and a second primer (JABR6) upstream of *pspF* and in the coding region of the *pspA* gene (positions 16 to 31). The PCR product was cleaved with *Bam*HI and *Bgl*II, and the resulting 950-bp fragment was ligated into the *Bam*HI site of pBR322, creating pMJ3.

Additional plasmids used in the complementation analyses were pMJ4, pMJ5, pMJ6, and pMJ7. They contain DNA fragments subcloned from pBRPS-1. Plasmid pMJ6 carries the 2.4-kb *Eco*RI-*Bgl*II DNA fragment inserted into *Eco*RI-*Bam*HI-digested pBR322. Plasmid pMJ4 contains the 0.7-kb *Bgl*II-*Ava*I fragment inserted into *Bam*HI-*Ava*I-digested pBR322. Plasmid pMJ5 contains the 1.2-kb *Bgl*II-*Stu*I fragment inserted into *Bam*HI-*Eco*RV-digested pBR322. Plasmid pMJ7 contains a 1.4-kb *Bgl*II-*Xmn*I fragment inserted into pBR322 that had been digested with *Bam*HI and *Eco*RV.

Transformation of *E. coli* cells was performed as described previously in standard protocols by using the $CaCl_2$ procedure (48).

Mutagenesis. Random insertion mutagenesis was performed in the *E. coli* K-12 strain K1472 [Δ*pspABC*::*kan* (Kn^r) Δ*lac*] by infection with a derivative of λ phage, λNK1323 (27), which carries a Tn*10*-derived transposon, miniTn*10-tet*. Mutagenesis was done essentially as described previously (27), except that transductants were pooled after mutagenesis, transformed with toxic plasmid pLW38, and selected for growth on kanamycin (marker on the recipient strain), tetracycline (marker on the transposon), and ampicillin (marker on the plasmid) plates at 37°C. K1472 was used for the mutagenesis because in the absence of PspA, the *psp* negative regulator, it cannot maintain plasmid pLW38 in which the *lac* operon is under the control of the *psp* promoter region. K1472 isogenic strains, K1471 and K1530, that lack PspA and that have additional mutations in *rpoN* or *himD* can grow and maintain the otherwise toxic plasmid, pLW38 (see Results). Kn^r Tc^r Ap^r colonies were screened at 42°C on minimal medium plates supplemented with leucine alone or at 37°C with leucine and glutamine, in order to distinguish between IHF, σ^{54}, and other mutants (20, 24). A number of mutants that did not have an IHF or σ^{54} mutant phenotype were backcrossed into the parent strain by P1 transduction. Individual Tc^r Kn^r colonies were then tested for their ability to maintain pLW38. Those that maintained the plasmid (e.g., were Ap^r) were characterized further (see Results).

Analysis of *psp* induction. Expression of the *psp* operon was assayed without stimuli or following induction with the f1 gene IV protein (2 mM IPTG, 30 min), heat shock (48 to 50°C, 5 min), or osmotic shock (0.75 M NaCl, 10 min) or after ethanol treatment (10% ethyl alcohol, 30 min). Bacteria were grown to 1×10^8 to 4×10^8 cells per ml in DO salts (54) supplemented with 0.4% glucose, 5 μg of thiamine per ml, and 19 amino acids (0.2 mg/ml each; no methionine). Samples of each culture (0.2 ml) were pulse-labelled for 60 s with 20 μCi of [^{35}S]methionine (New England Nuclear; 1,000 Ci/mmol), precipitated with cold trichloroacetic acid (5%), resuspended in 4% sodium dodecyl sulfate, and immunoprecipitated (15) with anti-PspA serum and Protein A Sepharose beads (Pharmacia) (9).

DNA manipulations. Plasmid mini- and midipreparations, preparation of *E. coli* chromosomal DNA, Southern blot hybridization, PCR amplification (*Taq* polymerase; Perkin-Elmer Cetus), electroelution, and other recombinant DNA techniques were performed as previously described (48). The synthetic oligonucleotides (synthesized by Operon Technologies Inc., Alameda, Calif.) used for PCR amplification, DNA sequencing, and primer extension experiments were as follows: IR1070Bam, 5′-GGTGGATCCACCTTAACTTA-3′; JABR6, 5′-CGATGTCGGCAAAGCG-3′; IR1070, 5′-CCACCTTAACTTAATG-3′; dASITE, 5′-CTGGAATTCTGCTGTTGATT-3′; PSP500UP, 5′-GAACGTGCCGACGGCG-3′; G1, 5′-CGATAGCGTCTTCAG-3′; G2, 5′-TTGATGGCAGAATAC-3′; G3, 5′-TGGCAGTTGTACAAC-3′; G4, 5′-CGCGCTGACCTGCTC-3′; G5, 5′-CACAGGTCAGCAGGA-3′; G6, 5′-GATTGCACAGAAAAT-3′; and PE-F, 5′-CCTGTTCCAGCACTTCGAGAAAGCTGTTCG-3′.

The λNK1323 DNA was isolated by using a Qiagen>lambda< kit (Qiagen Inc.). A [α-^{32}P]ATP-labelled 1.65-kb DNA fragment (Tet/b), isolated from *Bam*HI-*Eco*RI-digested λNK1323 DNA, was used as a probe in Southern blot hybridization to identify the miniTn*10-tet* element. Similarly, probes 4 (a 1.1-kb *Hin*cII-*Bgl*II fragment from the 5′ region of the *psp* operon derived from plasmid pBRPS-1) and Kan (a 1-kb *Bam*HI fragment carrying the *kan* cassette) were used in Southern blot hybridizations to identify the *psp*$^+$ and Δ*pspABC* strains, respectively.

DNA sequencing and analysis. The nucleotide sequences of *pspF* and its 3′ downstream region were determined for both strands by using single-stranded M13mp18 phage DNA or double-stranded DNA (plasmids pBRPS-1, pMJ1, and pMJ2) templates. The nucleotide sequence of mutant gene *pspF*::miniTn*10-tet* was determined for the template strand by using plasmid pMJ3. DNA sequencing was performed by the dideoxynucleotide chain termination method (49). Sequencing reactions were carried out with the Sequenase sequencing kit (U.S. Biochemical Corp.), and primers were the universal and reverse M13 primers or specific synthetic oligonucleotides (G1-G6 and PSP500UP). The sequence of the miniTn*10-tet*–*pspF* junction from mutant K1527 (*psp*$^+$ *pspF*::mTn*10-tet*) was determined for plasmid pMJ3 by using primer IR1070 (nucleotides 33 to 48 from the inverted repeats on the miniTn*10-tet* transposon).

The sequencing products were resolved on denaturing 5 or 8% polyacrylamide gels.

Nucleotide sequence data were analyzed by using the Gen Info network BLAST server (BLASTN, BLASTP, and BLASTX) (3) and MACAW, with the PAM120 matrix provided by the National Center for Biotechnology Information (50).

Primer extension. Primer extension analysis (48) was carried out by using a synthetic oligonucleotide (PE-F) complementary to nucleotides 35 to 64 of the *pspF* coding sequence. The oligonucleotide was labelled at the 5′ end by using [γ-^{32}P]ATP and hybridized to 700 μg of total RNA isolated from *E. coli* K561 and K561(pNAL200) (after induction with IPTG in log phase; 5×10^8 cells per ml) or to 50 μg of total RNA isolated before and after heat shock (50°C, 5 min) of *E. coli* K561(pBRPS-1). RNA was extracted from late-log-phase cultures of *E. coli* grown in minimal medium at 37°C or after heat shock by the hot acid-phenol method (1). RNA samples were resuspended in RNase-free water and incubated at 37°C for 1 h in the presence of RNase-free DNase. The cDNA was extended at 37°C for 40 min with Moloney murine leukemia virus reverse transcriptase (Pharmacia Biotech). The reverse transcription products were subjected to electrophoresis on a 5% polyacrylamide sequencing gel along with a nucleotide sequence generated with the same (unlabelled) primer.

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL data bank with the accession number U38542.

RESULTS

Identification of the *pspF* genetic locus. In the absence of PspA, the negative regulator of the *psp* operon, the *psp* genes are expressed at high levels under normal growth conditions and without any inducing stimulus (56). Plasmid pLW38, in which the *lac* operon is under the control of the *psp* promoter region, cannot be introduced or maintained in a strain that carries a *pspABC* deletion. If, however, the strain also contains a mutation in *himD* (which encodes one of the subunits of

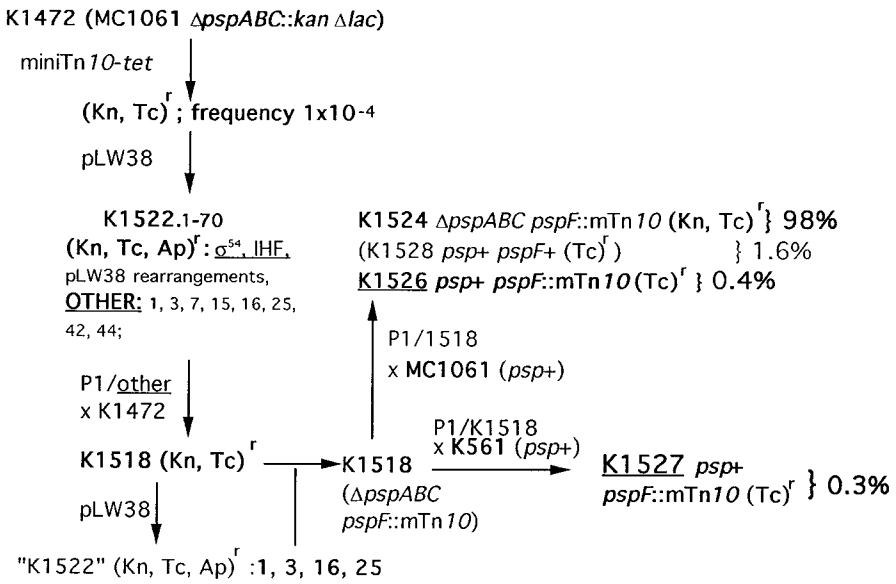

FIG. 1. Mini-transposon (miniTn*10-tet*) mutagenesis of strain K1472 and construction of the *psp*$^+$ *pspF*::miniTn*10-tet* mutant strains K1526 and K1527, by using consecutive P1 transductions (see the text). The percentages indicate the frequency with which a particular genotype was obtained.

IHF), the efficiency of transformation relative to that of the wild type is increased to 10^{-2}, and if the strain contains a mutation in *rpoN*, (which encodes σ^{54}), the relative efficiency is about 1. Since σ^{54} is absolutely required for *psp* transcription and IHF is strongly stimulatory, we concluded that the inability of Δ*pspABC* strains to maintain pLW38 is due to strong constitutive expression from the *psp* promoter. We assumed that such transcription reflects the action of a σ^{54} transcriptional activator that must be constitutively active and that strains containing a mutation in the putative activator would be able to maintain the otherwise toxic plasmid pLW38. Hence, we mutagenized strain K1472 [Δ*pspABC*::*kan* (Kn^r) Δ*lac*] by miniTn*10-tet* transposon mutagenesis (Fig. 1). Independent Tc^r Kn^r colonies (10^4) were pooled and transformed with pLW38, which confers ampicillin resistance. The transformants were selected on plates containing all three antibiotics. The 70 colonies obtained (K1522.1 to K1522.70) (Fig. 1) were screened for growth on minimal medium supplemented with leucine (at 42°C) or leucine and glutamine (at 37°C). Those that could not grow on leucine-containing plates were assumed to be IHF mutants (20). The Him (IHF) mutations cause a leucine-imposed, temperature-dependent auxotrophy and a requirement for isoleucine at 42°C (20). Leucine represses the synthesis of the enzymes required for isoleucine biosynthesis (AHAS) III, while expression of the other enzymes in the pathway requires active IHF (20). Those which required glutamine for growth were assumed to be σ^{54} mutants (24) (Fig. 1), since *rpoN* mutants of *E. coli* are glutamine auxotrophs (24). Of the 56 colonies that could grow on leucine and did not require glutamine, 8 in which pLW38 had no obvious rearrangements were chosen for further study (Fig. 1). The miniTn*10-tet* elements in these eight colonies were each transferred back to the parental strain (K1472) by P1 transduction to provide a clean background. The new recipients were again individually transformed with the toxic plasmid and tested for the ability to maintain the plasmid. Four independent recipients that passed this second screen (strains K1518.1, -3, -16, and -25; also referred to here as 1, 3, 16, and 25) (Fig. 1) are potential activator mutants. These Tc^r Δ*pspABC* strains were used as donors in another round of P1 transductions, this time into a recipient (MC1061) with an intact *psp* operon. The Tc^r transductants were screened for cotransduction of Δ*pspABC* (i.e., Kn^r marker). The cotransduction frequency, >98%, indicates that the potential activator mutants are closely linked to the *psp* genes (strain K1524) (Fig. 1). For each of these four candidates, 1, 3, 16, and 25 (data presented only for candidate 1), the transposition event and orientation of the inserted miniTn*10-tet* transposon were confirmed by a Southern blot of *Eco*RI-digested chromosomal DNA prepared from strains MC1061, K1472, K1518.1, K1524.1, K1526.1, and K1528 by using a [α-^{32}P]ATP-labelled probe, Tet/b (Fig. 2B). At the level of this structural analysis, the four isolates were indistinguishable. Further Southern blot hybridization analysis of these strains, using probes 4 and Kan (Fig. 2), and PCR analysis (Fig. 3) showed that the predominant miniTn*10-tet* insertion is present in a single copy on the bacterial chromosome and places it approximately 1.0 kb upstream from the *pspA* gene (Fig. 2; Fig. 3).

Southern (Fig. 2) and PCR (data not shown) analyses also showed that of the small number of transductants in which Kn^r was not cotransduced with Tc^r, only 20% had an insertion closely linked to the *psp* operon (strain K1526), and the remaining 80% had insertions positioned at some other place in the chromosome (strains K1528.3 and K1528.16) (Fig. 1).

To address whether the selected insertional mutants influence the induction of the *psp* operon, the level of expression of PspA was determined before and after a 48°C heat shock (Fig. 4A). Strains MC1061 (wild type), K1530 (*himD*), K1471 (*rpoN*), K1526 (miniTn*10-tet* insertion mutation closely linked to *psp*), and K1528 (miniTn*10-tet* insertional mutation not closely linked to *psp*) were labelled with [^{35}S]methionine, and the extracts were immunoprecipitated with anti-PspA serum. The miniTn*10-tet* insertion mutants closely linked to the *pspA* gene (K1526.1 and -3) (Fig. 4A, lanes 4 and 5) abolished induction of the *psp* operon as effectively as did the σ^{54} mutant (Fig. 4A, lane 3). The other two selected insertional mutants, K1526.16 and -25, had the same phenotype as did K1526.1 and -3 (data not shown). The mutants in which the transposon insertion was not closely linked to the *psp* operon were not defective in PspA induction (Fig. 4A, lanes 6 and 7). The

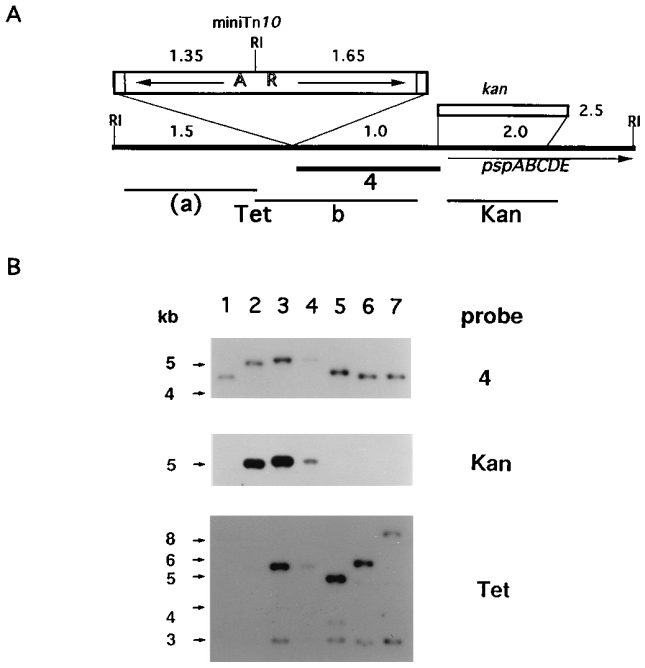

FIG. 2. Identification of the miniTn*10-tet* insertion. (A) Schematic representation of the chromosomal locus containing the *psp* operon, construction of the Δ*pspABC* by using the *kan* cassette (2-kb chromosomal fragment became 2.5 kb after deletion and insertion of *kan* cassette) (56), insertion point of the miniTn*10-tet* transposon, and probes used in Southern blot hybridization. A and R are *tet* genes. RI, *Eco*RI restriction site. Numbers identify the length of DNA in kilobases. Tet/b, 4, and Kan are probes used in hybridization. (B) Southern blots of chromosomal DNA digested with *Eco*RI and probed with [α-^{32}P]ATP-labelled probes 4, Kan, and Tet/b. Lane 1, MC1061 (*psp*$^+$); lane 2, K1472 (Δ*pspABC*); lane 3, K1518.1 (*pspF*::miniTn*10-tet* Δ*pspABC*); lane 4, K1524.1 (*pspF*::miniTn*10-tet* Δ*pspABC*); lane 5, K1526.1 (*psp*$^+$ *pspF*::miniTn*10-tet*); lane 6, K1528.16 (*psp*$^+$ miniTn*10-tet*); lane 7, K1528.3 (*psp*$^+$ miniTn*10-tet*). The lower band in the blot with probe Tet/b represents the residual hybridization with the 1.35-kb Tet/a DNA fragment retained after isolation of Tet/b (panel A). The 4.65-kb bands obtained after hybridization with probes 4 and Tet/b indicate an insertion of the miniTn*10-tet* transposon into the chromosome, closely linked to the *psp* operon in *psp*$^+$ strain MC1061 (lane 5). kb, DNA molecular weight standards (in kilobases).

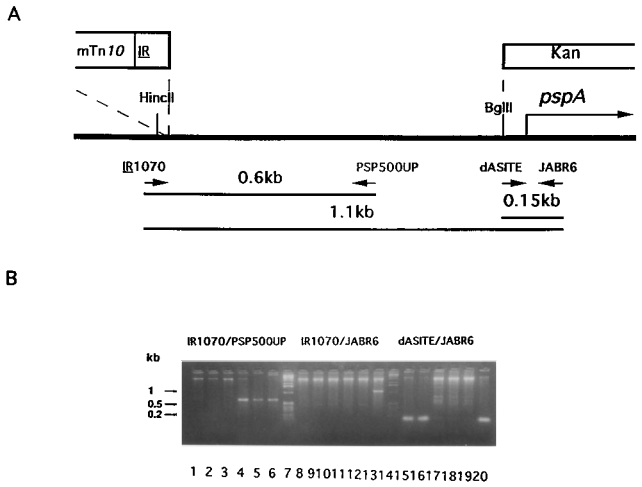

FIG. 3. Physical mapping of the miniTn*10-tet* transposon insertion by using PCR analysis. (A) Schematic representation of the chromosomal locus containing the upstream region of the *psp* operon, the 5′ region of the *pspA* gene, the insertion point of the miniTn*10-tet* transposon (mTn*10*), and the *kan* cassette (Kan). IR, 70-bp inverted repeat of the minitransposon, for which primer IR1070 is specific. Arrows, positions of the primers used in PCR. Bars, predicted DNA fragments which correspond to the positions of the particular pairs of primers: IR1070 and PSP500UP, IR1070 and JABR6, and dASITE and JABR6. (B) Identification of the miniTn*10-tet* insertion point by using the pairs of primers presented in panel A and PCR on isolated chromosomal DNA from strains MC1061 (*psp*$^+$) (lanes 1, 2, 8, 9, 15, and 16), K1472 (Δ*pspABC*) (lanes 3, 10, and 17), K1518.1 (*pspF*::miniTn*10-tet* Δ*pspABC*) (lanes 4, 11, and 18), K1524.1 (*pspF*::miniTn*10-tet* Δ*pspABC*) (lanes 5, 12, and 19), and K1526.1 (*psp*$^+$ *pspF*::miniTn*10-tet*) (lanes 6, 13, and 20). Lanes 7 and 14, 1-kb DNA ladder. Only in the *psp*$^+$ strain with a closely linked minitransposon insertion, K1526.1 (as well mutants -3, -16, and -25 [data not shown]), did the IR1070-JABR6 combination of primers amplify a DNA fragment of 1.1 kb (lane 13).

genetic locus identified by the miniTn*10-tet* in strain K1526 was named *pspF*.

Since much of the previous work on the *psp* operon was carried out in K561, we transferred the Tc^r marker from strain K1518.1 to strain K561 (*psp*$^+$), generating strain K1527 (*psp*$^+$ *pspF*::mTn*10-tet*) in which the miniTn*10-tet* insertion remained closely linked to the *psp* operon (Fig. 1).

psp induction in K561 and mutant K1527 in response to a number of stimuli is shown in Fig. 4B. The mutant strain does not express PspA either before or after induction with heat shock, osmotic shock, pIV f1 protein, or ethanol treatment (Fig. 4B).

Nucleotide sequence of the *pspF* gene. The *pspF* gene is present on plasmid pBRPS-1 which carries the entire *pspABCE* operon and 2.4 kb of upstream DNA, a portion of which had been previously sequenced (9). The nucleotide sequence of the entire *pspF* gene was determined for both strands after appropriate subcloning into M13mp18, pUC19, and pBR322 (see Materials and Methods). A 978-nucleotide open reading frame, divergently oriented from the *psp* operon, was identified (Fig. 5; Fig. 6). The ATG codon is preceded by a putative ribosome binding site (AAGTGGCGAA) at an appropriate distance (Fig. 6). The open reading frame is predicted to encode a protein of 325 amino acid residues with a calculated molecular mass of 37 kDa (Fig. 6). Overproduction of PspF showed that the *pspF* gene indeed encodes a protein, with a molecular mass of approximately 37 kDa (data not shown).

A comparison of the amino acid sequence with the GenBank database indicated that PspF is a member of the bacterial EBP family of σ^{54} transcriptional activators (Fig. 7). PspF differs from the other members of the EBP family by the apparent lack of an N-terminal regulatory domain. The central domain of PspF shows strong homology to the corresponding domain of known EBPs (36) (Fig. 7). The highly conserved regions in the central domain include sequences that are proposed to be important for interaction with σ^{54}, a glycine-rich region with a putative ATP-binding motif (GERGTGKE), and a region (QekLLRVie) which has similarity to segment 3 of adenylate kinase (36) (Fig. 6). Like other EBP regulators, the PspF C-terminal domain contains a helix-turn-helix (HTH) motif, characteristic of DNA-binding proteins (39) (Fig. 6). The HTH sequence is linked to the central domain by a 40-amino-acid segment.

The complete PspF protein is most similar to *E. coli* Az (ornithine decarboxylase inhibitor), NifA (*Azotobacter vinelandii*), HydG (*E. coli*), HydG (*Salmonella typhimurium*), NifA (*Rhizobium leguminosarum*), NifA (*Azospirillum lipoferum*), NifA (*Azospirillum brasilense*), and NifA (*Herbaspirillum seropedicae*) (Fig. 7). This suggests that PspF could belong to the Nif subgroup of the EBP family of response regulators. However, the HTH motif of PspF, especially the second recognition helix, differs significantly from that of other members of this EBP family (Fig. 6).

The 150-nucleotide sequence downstream from the *pspF*

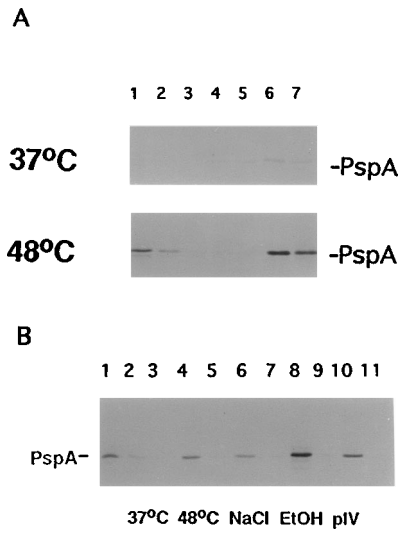

FIG. 4. *pspF*::miniTn*10-tet* mutation abolishes σ^{54}-dependent induction of the *psp* operon. (A) Immunoprecipitations using antiserum to PspA before and after heat shock induction. Lane 1, MC1061 (*psp*$^+$); lane 2, K1530 (*psp*$^+$ *himD*Δ3::*cam*); lane 3, K1471 (*psp*$^+$ *rpoN*::Tn*10*); lane 4, K1526.1 (*psp*$^+$ *pspF*::miniTn*10-tet*); lane 5, K1526.3 (*psp*$^+$ *pspF*::miniTn*10-tet*); lane 6, K1528.3 (*psp*$^+$ miniTn*10-tet*); lane 7, K1528.16 (*psp*$^+$ miniTn*10-tet*). (B) PspA synthesis after heat shock, osmotic shock, ethanol (EtOH) treatment, and f1 pIV induction of the *psp* operon. Lane 1, PspA marker band; lanes 2, 4, 6, 8, and 10, K561 (*psp*$^+$); lanes 3, 5, 7, 9, and 11, K1527 (*psp*$^+$ *pspF*::miniTn*10-tet*).

gene is partially or fully homologous to a set of 153 sequences found in intergenic regions or at transcriptionally 3′ positions in *E. coli* and in some *Shigella* chromosomes. These species-specific sequences represent a family of complex prokaryotic repetitive extragenic palindromes, called RIB (stands for reiterative IHF BIME [bacterial interspread mosaic elements]) (6, 16, 40, 51). The RIB sequence in the *pspF* 3′ flanking region is composed of two converging REP (stands for repetitive extragenic palindrome) sequences, REP1 and REP2, flanking a conserved segment which carries a static DNA bend and a consensus sequence for the binding of *E. coli* IHF (Fig. 6).

Transcription start site of the *pspF*. To determine the start site of *pspF* mRNA and, consequently, the location of the *pspF* promoter, primer extension reactions were performed with total RNA isolated from *E. coli* K561, K561(pNAL200), and K561(pBRPS-1) (as described in Materials and Methods). As a negative control, we used a reaction mixture without added RNA, containing only labelled PE-F primer. The results of the primer extension analysis are shown in Fig. 8. The major transcriptional start site was observed 33 bp upstream of the *pspF* initiation ATG codon; adjacent adenine and guanine residues (positions −93 and −94 [relative to the start site of *pspA* transcription] shown in Fig. 8B] were used as the major start

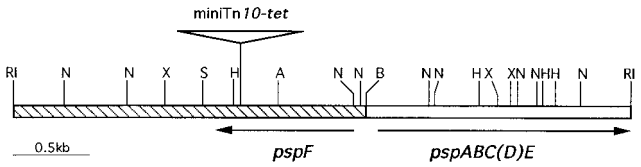

FIG. 5. The genetic organization and the partial restriction map of the *pspF* gene and the *pspABCE* operon. The position of the miniTn*10-tet* insertion is indicated. RI to RI, 4.5-kb *Eco*RI DNA fragment (cloned in pBRPS-1); hatched box, 2.4-kb *Eco*RI-*Bgl*II DNA fragment. RI, *Eco*RI; N, *Nla*III; X, *Xmn*I; S, *Stu*I; H, *Hin*cII; A, *Ava*I; B, *Bgl*II.

```
CTAACAGAGCGTTGATGTTGGCATTCACGATGTCGGCAAAGCGAGAAAAAATACCCAT AATGTTGTCCTCTTGATTTC
                                                    <--pspA
TGCGTTCAGGATTGTCCTGCTGAACTGTTAATACAATTTGCGTGCCAATTTTTTATCTTTTGATTTATAAAGATCTGA
                                                                          1/1
TTGAAGAATCAACAGCAACATGCCAGGATGAGTTAGCGAATTACACTAACAAGTGGCGAATTTCATC ATG GCA GAA
                                                        RBS               M   A   E
10/4
TAC AAA GAT AAT TTA CTT GGT GAG GCG AAC AGC TTT CTC GAA GTG CTG GAA CAG GTT TCG
 Y   K   D   N   L   L   G   E   A   N   S   F   L   E   V   L   E   Q   V   S
70/24
CAT CTC GCA CCG CTG GAC AAA CCG GTG CTC ATC ATC GGC GAA CGC GGC ACC GGT AAA GAG
 H   L   A   P   L   D   K   P   V   L   I   I   G   E   R   G   T   G   K   E
130/44
CTG ATT GCC AGC CGC CTG CAT TAT CTC TCC TCC CGT TGG CAA GGG CCG TTT ATT TCC CTT
 L   I   A   S   R   L   H   Y   L   S   S   R   W   Q   G   P   F   I   S   L
190/64
AAC TGC GCG GCG TTA AAT GAA AAT CTG CTG GAT TCC GAA CTG TTT GGT CAC GAA GCG GGG
 N   C   A   A   L   N   E   N   L   L   D   S   E   L   F   G   H   E   A   G
250/84
GCG TTT ACC GGT GCG CAA AAA CGT CAT CCA GGG AGA TTT GAA CGT GCC GAC GGC GGT ACG
 A   F   T   G   A   Q   K   R   H   P   G   R   F   E   R   A   D   G   G   T
310/104
CTA TTT CTT GAT GAA CTC GCT ACG GCA CCG ATG ATG GTG CAG GAG AAA TTA TTG CGC GTG
 L   F   L   D   E   L   A   T   A   P   M   M   V   (Q   E   K   L   L   R   V
370/124
ATT GAG TAC GGT GAA CTG GAG CGC GTT GGC GGC AGC CAA CCA TTG CAG GTG AAT GTG CGG
 I   E)  Y   G   E   L   E   R   V   G   G   S   Q   P   L   Q   V   N   V   R
430/144
TTG GTA TGC GCG ACG AAT GCC GAT CTC CCG GCG ATG GTC AAT GAA GGC ACT TTT CGC GCT
 L   V   C   A   T   N   A   D   L   P   A   M   V   N   E   G   T   F   R   A
490/164
GAC CTG CTC GAC CGA CTG GCT TTT GAT GTT GTA CAA CTG CCA CCA CTG CGC GAG CGC GAA
 D   L   L   D   R   L   A   F   D   V   V   Q   L   P   P   L   R   E   R   E
550/184
AGC GAC ATA ATG TTG ATG GCA GAA TAC TTT GCC ATC CAG ATG TGT CGG GAA ATC AAG CTG
 S   D   I   M   L   M   A   E   Y   F   A   I   Q   M   C   R   E   I   K   L
 610/204
CCT CTG TTC CCG GGG TTT ACG GAG CGC GCC AGA GAA ACA TTG CTG AAT TAT CGT TGG CCG
 P   L   F   P   G   F   T   E   R   A   R   E   T   L   L   N   Y   R   W   P
670/224
GGA AAT ATT CGT GAA TTG AAA AAC GTG GTG GAA CGT TCA GTG TAT CGC CAC GGC ACC AGC
 G   N   I   R   E   L   K   N   V   V   E   R   S   V   Y   R   H   G   T   S
730/244
GAT TAT CCG CTT GAT GAC ATC ATT ATT GAT CCC TTT AAA CGG CGT CCG CCT GAA GAC GCT
 D   Y   P   L   D   D   I   I   I   D   P   F   K   R   R   P   P   E   D   A
790/264
ATC GCC GTT TCA GAA ACC ACC TCG CTT CCA ACA CTG CCG CTG GAT TTA CGT GAG TTT CAG
 I   A   V   S   E   T   T   S   L   P   T   L   P   L   D   L   R   E   F   Q
850/284                                  *
ATG CAG CAG GAA AAA GAG TTG CTG CAA CTC AGT TTG CAA CAA GGG AAA TAT AAC CAG AAG
 M   Q   Q   E   K   E   L   L   Q   L   S   L   Q   Q   G   K   Y   N   Q   K
910/304
CGC GCG GCT GAA TTA CTG GGG TTA ACT TAT CAT CAG TTC CGC GCG TTG TTG AAA AAG CAC
 R   A   A   E   L   L   G   L   T   Y   H   Q   F   R   A   L   L   K   K   H
970/324
CAG ATT TAG CGAAGTTGAGCAATACAACTTGCCGGATGCGGCGTGAATGCTTTATTCGGCATACATAATCAATACG
 Q   I   -                                  REP1 ->                        IHF
TCGTTTGCAGGCCTGATAAGCGCAGCGCATCAGACATTATTGCGTTGGGGCTATTAATCGCTCAGCGCATCCTGCAAAAT
                  <- REP2
```

FIG. 6. Nucleotide sequence and deduced amino acid sequence of PspF. Nucleotide sequence of a fragment of *E. coli* chromosomal DNA containing the 5′ region of the *pspA* gene and the entire sequence of a novel *pspF* gene are shown. Nucleotide positions and relevant amino acid residues of the *pspF* gene and encoded PspF protein are indicated (1/1 to 978/325). The stop codon is indicated by a dash. The amino acid sequences of the predicted ATP-binding pocket and the HTH motif are underlined. The amino acid sequence with similarity to segment 3 of adenylate kinase is in parentheses. The asterisk indicates the insertion site of the miniTn*10-tet* transposon. The putative ribosome binding site (RBS), IHF consensus, and REP sequences (REP1 and REP2) are underlined. Arrows indicate the orientation of the REP sequences.

points of transcription. This location is correctly positioned relative to a σ^{70} promoter P1 (−10 box CAGGAT, positions −82 to −87; −35 box TTGAAG, positions −59 to −64) (Fig. 8B). Multiple weak transcriptional start sites were also observed; two of them correspond to σ^{70} promoters identified by sequence and designated P2 and P3 (Fig. 8).

Primer extension analysis also showed that the amount of PspF mRNA is unchanged after induction with f1 pIV carried on pNAL200 or after heat shock (Fig. 8A). Lanes 1 and 2 in Fig. 8A show the results of reactions using a high concentration of total RNA (700 μg) isolated from strains K561 and K561(pNAL200), in which *pspF* is present as a single chromosomal copy. In lanes 3 and 4, much less total RNA (50 μg) was used, because the *pspF* gene is present in multicopy on plasmid pBRPS-1 which also carries the entire *psp* operon. Hence, the signals in Fig. 8A shown in lanes 1 and 2 are not directly comparable to those in lanes 3 and 4.

***pspF*::mTn*10-tet* (*pspF*$_{877}$) encodes a PspFΔHTH protein without a DNA-binding motif.** Sequence analysis of the *pspF*–miniTn*10-tet* joint site from the insertional mutant K1527 (*psp*$^+$ *pspF*::mTn*10-tet*) (as described in Materials and Methods) showed that the mTn*10-tet* transposon was inserted be-

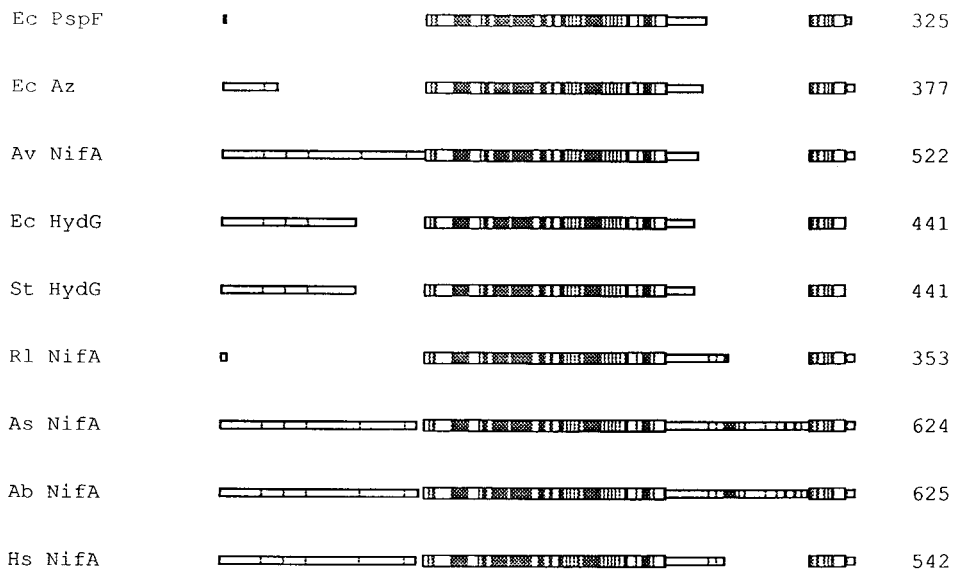

FIG. 7. Schematic representation of the alignment of the *E. coli* PspF protein with members of the σ^{54} bacterial EBP family (only proteins with the highest BLASTP scores). The sequences of nine EBPs were aligned by using the computer program MACAW provided by the National Center for Biotechnology Information. Small white blocks, 0 to 30% identity; bigger white blocks, 30 to 65% identity; shaded blocks, 65 to 100% identity. The numbers on the right indicate the number of amino acid residues. Ec, *E. coli*; Av, *A. vinelandii*; St, *S. typhimurium*; Rl, *R. leguminosarum*; As, *A. lipoferum*; Ab, *A. brasilense*; Hs, *H. seropedicae.*

tween nucleotides 877 and 878 (Fig. 9) of *pspF*, interrupting codon 293 (Leu) (Fig. 6 and 9). This insertion separates the central catalytic domain–C-terminal linker from the C-terminal DNA-binding HTH motif (Fig. 6; Fig. 9). The mutant gene ($pspF_{877}$) encodes a truncated, 304-amino-acid-residue (34.4-kDa) PspFΔHTH protein that lacks a DNA-binding motif (Fig. 9). Overexpression of the mutant gene $pspF_{877}$ gave rise to a protein of the predicted mobility (data not shown).

Complementation of the $pspF_{877}$ chromosomal mutation. A variety of plasmid constructs (Fig. 10A) were tested for their ability to complement the $pspF_{877}$ insertion mutant. The level of induction of the *psp* operon was monitored before and after heat shock or ethanol treatment by measuring the level of PspA expression with anti-PspA antibodies. Expression of wild-type PspF from a multicopy plasmid under its own promoter restores PspA synthesis in a strain with the $pspF_{877}$ mutation (Fig. 10B). Indeed, the level of synthesis of PspA is high even in the absence of an inducing stimulus, presumably because the higher concentration of activator overcomes the normal negative regulation of the operon. If the complementing plasmid (pMJ4) carries a *pspF* gene ($pspF_{618}$) with a 3′ deletion that extends into the region encoding the central catalytic domain of the protein, no complementation is observed (Fig. 10B). On the other hand, overproduction of the truncated protein PspFΔHTH (encoded by the $pspF_{877}$ gene) from plasmid pMJ3, which lacks the HTH motif, fully complemented the chromosomal $pspF_{877}$ mutation; as with wild-type *pspF*, overexpression of $pspF_{877}$ activates PspA expression even in the absence of an inducing stimulus (Fig. 10B).

Interestingly, induction of the *psp* operon before heat shock by overproduction of the HTH mutant protein PspFΔHTH is slightly better than that seen with wild-type PspF (Fig. 10B). However, after heat shock, strain K1527(pMJ3), which expresses the mutant protein, showed a decline in *psp* induction, while strain K1527, which expresses wild-type PspF protein (encoded by plasmid pMJ5, pMJ7, or pMJ6), displayed *psp* induction above pretreatment levels (Fig. 10B). Results obtained with ethanol treatment were basically the same (data not shown).

DISCUSSION

Transcription of the *pspABCE* operon in *E. coli* utilizes σ^{54}-RNAP, is facilitated by IHF, and is dependent on an upstream, *cis*-acting UAS (56, 57). Initiation of transcription employing the alternative sigma factor σ^{54}, in all cases hitherto studied, depends on the action of a specific transcriptional activator (14, 29, 36, 38, 43). We previously observed that in the absence of the PspA negative regulator and the PspB and PspC positive regulators, transcription from the *psp* promoter is constitutive (56). That observation implied the existence of a constitutively active transcriptional activator. In this study, we have identified, sequenced, and characterized a new *E. coli* σ^{54} transcriptional activator, PspF.

Insertion of miniTn*10-tet* into a single site in the *E. coli* chromosome (strain K1472) resulted in the loss of expression from the *psp* promoter. This chromosomal insertion which is closely linked to the *pspA* gene completely abolished σ^{54}-dependent induction of the *psp* operon in response to a variety of stimuli (expression of f1 pIV, ethanol treatment, and osmotic and heat shock). The gene into which the transposon inserted, designated *pspF*, is a 978-nucleotide open reading frame divergently oriented from the *pspABCE* operon and encodes a 37-kDa protein (PspF). From the deduced amino acid sequence of the *pspF* gene, we found a significant homology to the EBP family of transcriptional activators. During the course of this work, two reports appeared that identified a partial open reading frame 5′ of the *psp* operon that was similar to the central domain of the EBP transcriptional activators (36, 45) and represents the 5′ part of the *pspF* gene. A modular protein structure is a common feature of many σ^{54}-specific transcriptional activator proteins (2, 14, 36, 41). PspF protein has a two-domain structure. The central domain is homologous (52% identity) to the highly conserved central catalytic domain of EBPs which is important for interaction with σ^{54} (14, 36, 41, 53). A putative ATP-binding pocket (GERGTGKE) in PspF is found in the N-terminal part of the central domain. It has been proposed that binding and hydrolysis of a nucleoside triphos-

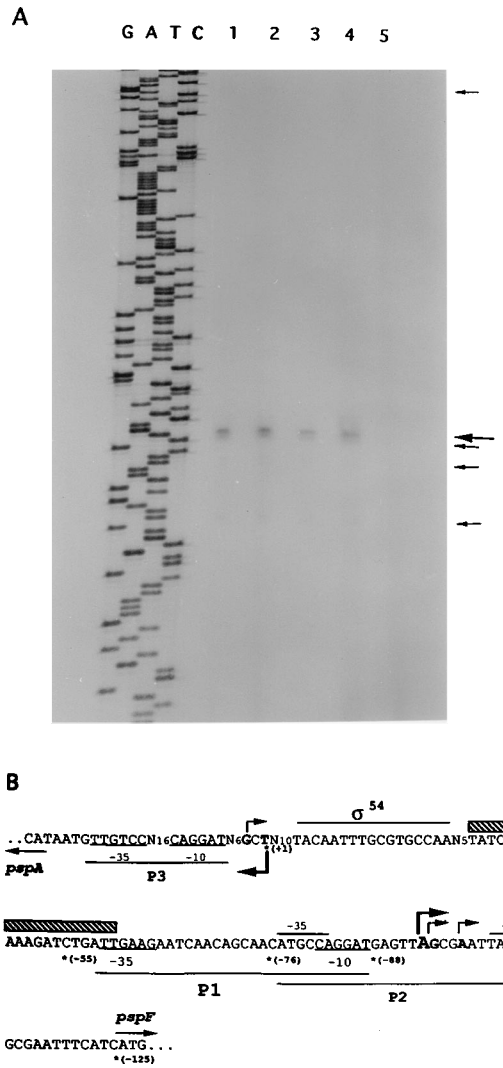

FIG. 8. Primer extension analysis of the transcriptional start site of the *pspF* gene. (A) Primer extension analysis of *pspF* transcription. Lane 1, K561; lane 2, K561(pNAL200); lanes 3 and 4, K561(pBRPS1); lane 5, primer [no RNA hybridized]. Total RNA was isolated from uninduced cultures (lanes 1 and 3) and from cultures induced with f1 pIV (2 mM IPTG, 30 min) (lane 2) or 50°C heat shock (lane 4). Lanes G, A, T, and C are products of sequencing reactions generated with the same oligonucleotide (PE-F) as that used for the primer extension. The arrows indicate the positions of the primer extension products. The major start site, A, is denoted by the thick arrow. (B) Sequence of the *pspF* and *pspABCE* operon promoter regulatory region. The DNA sequence of the *pspF* and *pspABCE* operon promoter regulatory region (*pspF* nontemplate strand) from positions +44 to −128 relative to the start site of *pspA* transcription (+1) is shown. The hatched box represents the IHF footprint site. The *pspABCE* operon σ^{54} promoter (57) is indicated. The major σ^{70} *pspF* promoter, P1 (−59 to −87) and minor promoters P2 (−77 to −105) and P3 (+37 to +10) are indicated. Putative −35 and −10 boxes are underlined. Asterisks and numbers in parentheses indicate the positions of nucleotides relative to the start site of *pspA* transcription (+1). Also indicated are the start sites (boldface letters) and direction of transcription of the *pspABCE* operon (+1, thick arrow) and *pspF* gene (A [adenine] and thick arrow, major transcription start site; small arrows, minor transcription start sites). The predicted position of the UAS is between nucleotides −76 and −125 (57).

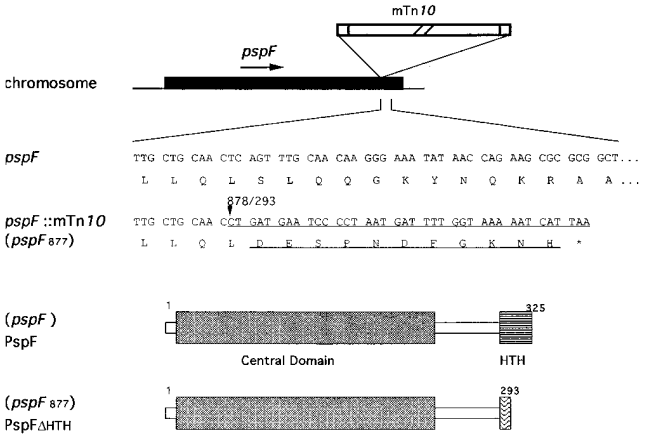

FIG. 9. The site of miniTn*10-tet* insertion on the *pspF* gene of K1527. The open and closed boxes represent the miniTn*10-tet* (mTn*10*) transposon and the *pspF* gene, respectively. The *pspF* nucleotide sequence indicates the sequence of the wild-type *pspF* gene, with the translated relevant region of the protein. The mutated *pspF* gene ($pspF_{877}$) with the C-terminal region of the deduced amino acid sequence is shown below. The miniTn*10-tet* transposon in K1527 (*pspF*::miniTn*10-tet*) was inserted (vertical arrow) after the C residue (nucleotide 877) of the codon for leucine (amino acid 293) in the *pspF* gene. The underlined nucleotide sequence represents the translated terminus (inverted repeat) of miniTn*10-tet*. Insertion of the transposon resulted in the addition of 11 amino acids (underlined amino acid sequence) to the truncated PspF protein, creating the PspFΔHTH protein (the scheme of the protein is at the bottom of the figure: shaded box, central domain; box with wavy lines, additional 11 amino acids). PspF, wild-type protein: shaded box, central domain; box with horizontal lines, HTH motif.

phate is required during transcriptional activation by σ^{54}-dependent activators (2, 14, 17, 36, 38, 41, 43, 53).

The COOH-terminal domain of PspF contains an HTH motif, characteristic of DNA-binding proteins (2, 14, 41, 43). All EBP σ^{54} transcriptional activators have a conserved HTH motif located within a COOH-terminal DNA-binding domain of variable length, and each subfamily of EBPs recognizes a specific enhancer-like element or UAS (10, 13, 14, 18, 38, 43, 44). Binding to these elements confers specificity to the initiation of transcription (18, 30, 36). The first helix in the PspF HTH motif has a significant degree of homology to the other σ^{54}-dependent activators, while the second helix (YHQFRALLK) of the putative DNA-binding motif is unique to the PspF protein. The second helix, named the recognition helix, is a specific feature of certain subfamilies of the activators (18, 30, 36). Since the PspF second helix does not resemble that of any of the others, it is probably a member of a novel subfamily. *psp* operon expression is dependent on a functional *pspF* product, and our previous experiments pointed out that the *psp* promoter region contains a UAS (located at nucleotides −76 to −125 upstream of the *pspA* transcription start site) that is required for *psp* induction and has the enhancer-like ability to activate at a distance (57). A DNA-binding activity that is dependent on the UAS was detected in crude protein extracts (57). In addition, *psp* expression is IHF dependent (57), and IHF binds to a site between the UAS and the σ^{54} recognition sequence (57). Hence, the interaction of the PspF activator and σ^{54}-RNAP could be facilitated by IHF-mediated DNA bending.

PspF differs from other members of the EBP family in the complete absence of an N-terminal domain. The native NifA protein of *R. leguminosarum* bv. trifolii and the *Pseudomonas syringae* HrpS protein have very short N-terminal domains, consisting of 12 and 74 amino acids, respectively (18, 22, 36). Furthermore, the N-terminal domain of NifA activators can be deleted without impairing the ability to activate *nif*-targeted promoters (4, 18). It has been proposed that the response regulator (activator) activity may be modulated through an intramolecular positive or negative interaction between the

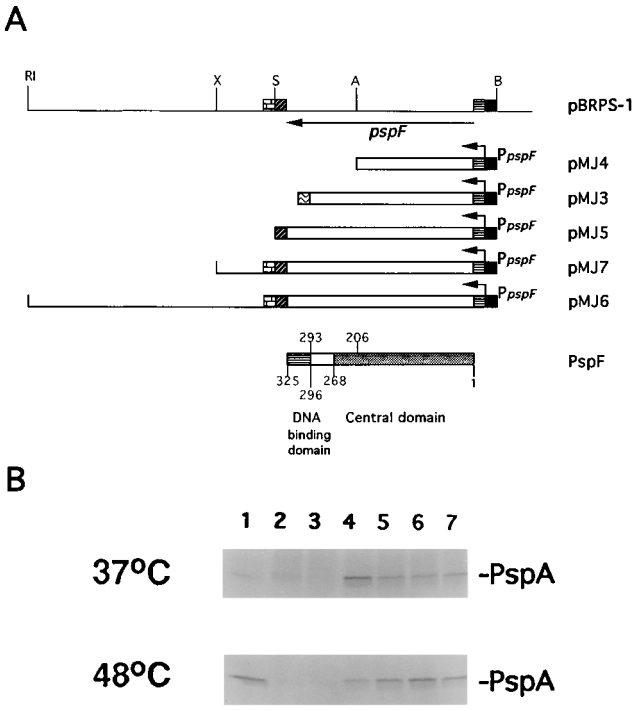

FIG. 10. Complementation of the $pspF_{877}$ chromosomal mutation. (A) Structures of the *pspF* gene on pMJ3 to pMJ7 subcloned from pBRPS-1 or as a PCR DNA fragment (see Material and Methods). Plasmids: closed box, P1 *pspF* promoter; arrow, major start site of *pspF* transcription; box with horizontal lines, predicted UAS position; open box, *pspF* sequence; box with wavy lines, addition of 11 amino acids from miniTn*10-tet* on truncated *pspF* ($pspF_{877}$); hatched box, REP1; bricks-box, REP2. RI, *Eco*RI; X, *Xmn*I; S, *Stu*I; A, *Ava*I; B, *Bgl*II. PspF protein (numbers are amino acid positions): 1 to 325, wild-type protein; 1 to 268, central domain; 268 to 325, DNA binding domain; 268 to 296, linker subdomain; 296 to 325, HTH domain; 1 to 206, pMJ4-encoded protein; 1 to 293 (304), pMJ3-encoded protein; 1 to 325, pMJ5-, pMJ6-, and pMJ7-encoded protein. (B) Synthesis of PspA before and 5 min after a temperature shift from 37 to 48°C (heat shock induction) screened by using [^{35}S]methionine-labelled total proteins immunoprecipitated with anti-PspA serum. The position of the PspA protein is indicated. Lanes 1, K561 (psp^+); lanes 2, K1527 (psp^+ $pspF_{877}$); lanes 3, K1527(pMJ4); lanes 4, K1527(pMJ3); lanes 5, K1527(pMJ5); lanes 6, K1527(pMJ7); lane 7, K1527(pMJ6).

N-terminal and other domains (18, 23). Since PspF lacks the regulatory N-terminal domain, this may explain why this activator is synthesized in a constitutively active form. A mutant *Rhizobium meliloti* DctD protein that lacks 142 amino acid residues from the N terminus ($DctD_{L143}$) is also constitutively active, and the CheB activator with a deleted N-terminal domain had elevated activity (23, 26, 31, 46).

The complementation of the chromosomal mutant $pspF_{877}$ by plasmid-borne wild-type PspF (pMJ5, pMJ6, and pMJ7), demonstrates that *pspF* consists of a simple transcription unit and represents a single complementation group. A *pspF* mutation which has a partially deleted central domain (C-terminal portion) ($pspF_{618}$) cannot complement the $pspF_{877}$ mutation. This is consistent with the observation that an intact functional central catalytic domain is required for activation of σ^{54}-regulated initiation of transcription (4, 25, 26). By contrast, plasmid pMJ3, which codes for a PspF protein that lacks the HTH motif (PspFΔHTH), was able to complement the identical chromosomal mutation. This suggests that the mutant protein is catalytically active but that when expressed at a low level from a single chromosomal gene is unable to stimulate transcription without the capability to bind to its UAS. On the other hand, the same mutant protein, when overproduced, does activate transcription, presumably by direct interaction with the σ^{54} holoenzyme. The central domain of NifA from *R. meliloti* and that of the homologous activator DctD from *R. leguminosarum* have been shown to activate transcription in vivo, as does a truncated form of NifA from *Klebsiella pneumoniae* that lacks its C-terminal domain (4, 25, 26). These results are in contrast to the case of NtrC, in which the central catalytic domain by itself is not active (39).

PspA has at least two roles. One is an unknown function in *E. coli* physiology, probably related to maintenance of the proton motive force (27a, 59), and the other is its role as a negative regulator. Functionally, the combination of PspF and PspA is the equivalent of a typical σ^{54} activator. As a negative regulator, PspA may provide the equivalent of the regulatory domain, found in many EBPs, while PspF provides the DNA binding and activation functions. In the absence of an inducing stimulus, it would be the ratio of PspA to PspF that determines the activity of the *pspABCE* promoter. In uninduced cells, the low but measurable level of PspA is enough to keep transcription at a low level. Overproduction of PspF can then overcome the negative regulation mediated by PspA. Exposure to an inducing stimulus probably results in the relief of the negative regulation provided by PspA, by an as yet not-understood mechanism. Although some inducing stimuli are more dependent on the presence of PspB and PspC than others, the common step needed to increase expression of the operon is probably this loss of PspA's repressive effect on PspF's function as a transcriptional activator. PspA could simply prevent interaction between PspF and σ^{54} by making a steric block or by influencing the DNA bending, or it could prevent the binding and proper positioning of the activator. It is also possible that PspA represses PspF activity indirectly, perhaps by eliminating an inducing signal or acting on another regulatory protein(s). We are carrying out experiments both in vivo and in vitro which should shed further light on the nature of the PspA-PspF interaction and, hence, also on whether PspA represents an "escaped" regulatory domain of PspF. From an evolutionary point of view, it would be interesting to know whether the prototypic σ^{54} activator consisted of a single protein, or whether like the PspA-PspF system, the regulatory element was separate from the activating function.

Primer extension showed that the PspF gene is transcribed from one major σ^{70} promoter (P1) and two minor σ^{70} promoters (P2 and P3). Two of them, P1 and P3, have the same sequence at the −10 box (CAGGAT). All three promoters have weak consensus sequences in comparison with canonical σ^{70} promoters, and according to primer extension analysis, transcription of *pspF* gene appears to be moderate or weak. The first two nucleotides in the −35 site of the P1 promoter overlap the IHF protected region. This particular IHF site has been reported to be constantly occupied in *E. coli* (55). The position of the major promoter P1 is at −59 to −87, and the position of the minor promoter P2 is at −77 to −105 (relative to the *pspA* transcription start site; Fig. 8B), while the predicted UAS is between nucleotides −76 and −125 (57) (Fig. 8B). Thus, the UAS to which PspF should bind is between the *pspF* P1 promoter and the translational start site of the gene (Fig. 8B). This implies a possible influence of PspF on its own transcription and suggests that *pspF* transcription may be autoregulated. In vitro studies testing this model are in progress.

Differences in the rate of mRNA degradation can determine the level of gene expression. A complex repetitive sequence (RIB), composed of two imperfect REP sequences and an IHF binding consensus in between, follows the *pspF* gene immediately further downstream. REP sequences in many *E. coli* intergenic or 3′ regions of transcriptional units are capable of

forming stem-loop structures that increase the stability of mRNA (37, 52). Stem-loop structures can protect upstream mRNA from degradation by impeding the processive activities of 3′-5′ exoribonucleases (37, 52). Recently, it was shown that EIF (stands for exoribonuclease impeding factor) in *E. coli* can bind to stem-loop structures (including REP sequences) and help to protect RNA from degradation (12). Whether *pspF* gene expression is modulated at the level of transcription initiation and elongation or postranscriptionally by stabilization of PspF mRNA remains to be determined. Nonetheless, the existence of the RIB sequence on the 3′ end of the PspF gene as well as the position of the UAS in a region that overlaps or is downstream from the major P1 and the minor P2 σ^{70} *pspF* promoters offer a number of possible routes for the regulation of PspF expression.

ACKNOWLEDGMENTS

We thank Nancy Kleckner for providing strains and λ phage containing the minitransposon, Simon Daefler for help with computer analysis, and Marjorie Russel for a critical reading of the manuscript.

This work was supported by NSF grant MCB 93-16625. G.J. was supported by a fellowship from the Beckman Fund for Biomedical Research. L.W. was supported by the Lucille P. Markey Charitable Trust (Miami, Fla.) and by training grant AIO7233 from the NIH.

REFERENCES

1. **Aiba, H., S. Adhya, and B. D. Crombrugghe.** 1981. Evidence for two *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:**11905–11910.
2. **Albright, L. M., E. Huala, and F. M. Ausubel.** 1989. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. Annu. Rev. Genet. **23:**311–336.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Berger, D. K., F. Narberhaus, H.-S. Lee, and S. Kustu.** 1995. In vitro studies of the domains of the nitrogen fixation regulatory protein NIFA. J. Bacteriol. **177:**191–199.
5. **Bergler, H., D. Abraham, H. Aschauer, and F. Turnowsky.** 1994. Inhibition of lipid biosynthesis induces the expression of the *pspA* gene. Microbiology **140:**1937–1944.
6. **Boccard, F., and P. Prentki.** 1993. Specific interaction of IHF with RIBs, a class of bacterial repetitive DNA elements located at the 3′ end of transcription units. EMBO J. **12:**5019–5027.
7. **Brissette, J. L., and M. Russel.** 1990. Secretion and membrane integration of a filamentous phage-encoded morphogenetic protein. J. Mol. Biol. **211:**565–580.
8. **Brissette, J. L., M. Russel, L. Weiner, and P. Model.** 1990. Phage shock protein, a stress protein of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **87:**862–866.
9. **Brissette, J. L., L. Weiner, T. L. Ripmaster, and P. Model.** 1991. Characterization and sequence of the *Escherichia coli* stress-induced *psp* operon. J. Mol. Biol. **220:**35–48.
10. **Buck, M., S. Miller, M. Drummond, and R. Dixon.** 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature (London) **320:**374–378.
11. **Carlson, J. H., and T. J. Silhavy.** 1993. Signal sequence processing is required for the assembly of LamB trimers in the outer membrane of *Escherichia coli*. J. Bacteriol. **175:**3327–3334.
12. **Causton, H., B. Py, R. S. McLaren, and C. F. Higgins.** 1994. mRNA degradation in *Escherichia coli*: a novel factor which impedes the exoribonucleolytic activity of PNPase at stem-loop structures. Mol. Microbiol. **14:**731–741.
13. **Chen, P., and L. J. Reitzer.** 1995. Active contribution of two domains to cooperative DNA binding of the enhancer-binding protein nitrogen regulator I (NtrC) of *Escherichia coli*: stimulation by phosphorylation and the binding of ATP. J. Bacteriol. **177:**2490–2496.
14. **Collado-Vides, J., B. Magasanik, and J. D. Galla.** 1991. Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. **55:**371–394.
15. **Davis, N. G., J. D. Boeke, and P. Model.** 1985. Fine structure of a membrane anchor domain. J. Mol. Biol. **181:**111–121.
16. **Dimri, G. P., K. E. Rudd, M. K. Morgan, H. Bayat, and G. F.-L. Ames.** 1992. Physical mapping of repetitive extragenic palindromic sequences in *Escherichia coli* and phylogenetic distribution among *Escherichia coli* strains and other enteric bacteria. J. Bacteriol. **174:**4583–4593.
17. **Eydmann, T., E. Soderback, T. Jones, S. Hill, S. Austin, and R. Dixon.** 1995. Transcriptional activation of the nitrogenase promoter in vitro: adenosine nucleotides are required for inhibition of NIFA activity by NIFL. J. Bacteriol. **177:**1186–1195.
18. **Fischer, H.-M.** 1994. Genetic regulation of nitrogen fixation in Rhizobia. Microbiol. Rev. **58:**352–386.
19. **Freundlich, M., N. Ramani, E. Mathew, A. Sirko, and P. Tsui.** 1992. The role of integration host factor in gene expression in *Escherichia coli*. Mol. Microbiol. **6:**2557–2563.
20. **Friedman, D. I., E. J. Olson, D. Carver, and M. Gellert.** 1984. Synergistic effect of *himA* and *gyrB* mutations: evidence that Him functions control expression of *ilv* and *xyl* genes. J. Bacteriol. **157:**484–489.
21. **Goosen, N., and P. V. D. Putte.** 1995. The regulation of transcription initiation by integration host factor. Mol. Microbiol. **16:**1–7.
22. **Grimm, C., W. Aufsatz, and N. J. Panopoulos.** 1995. The *hrpRS* locus of *Pseudomonas syringae* pv. *phaseolicola* constitutes a complex regulatory unit. Mol. Microbiol. **15:**155–165.
23. **Gu, B., J. H. Lee, T. R. Hoover, D. Scholl, and B. T. Nixon.** 1994. *Rhizobium meliloti* DctD, a σ^{54}-dependent transcriptional activator, may be negatively controlled by a subdomain in the C-terminal end of its two-component receiver module. Mol. Microbiol. **13:**51–66.
24. **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.
25. **Huala, E., and F. M. Ausubel.** 1989. The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti nifH* promoter. J. Bacteriol. **171:**3354–3365.
26. **Huala, E., J. Stigter, and F. M. Ausubel.** 1992. The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. J. Bacteriol. **174:**1428–1431.
27. **Kleckner, N., J. Bender, and S. Gottesman.** 1991. Uses of transposons with emphasis on Tn*10*. Methods Enzymol. **204:**139–180.

27a. **Kleerebezem, M., W. Crielaard, and J. Tommassen.** 1996. Involvement of stress protein PspA (phage shock protein A) of *Escherichia coli* in maintenance of the proton motive force under stress conditions. EMBO J. **15:**162–171.

28. **Kleerebezem, M., and J. Tommassen.** 1993. Expression of the *pspA* gene stimulates efficient protein export in *Escherichia coli*. Mol. Microbiol. **7:**947–956.
29. **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.
30. **Labes, M., V. Rastogi, R. Watson, and T. M. Finan.** 1993. Symbiotic nitrogen fixation by *nifA* deletion mutant of *Rhizobium meliloti*: the role of an unusual *ntrC* allele. J. Bacteriol. **175:**2662–2673.
31. **Lee, J. H., D. Scholl, B. T. Nixon, and T. R. Hoover.** 1994. Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of *Rhizobium meliloti* DCTD, a σ^{54}-dependent transcriptional activator. J. Biol. Chem. **269:**20401–20409.
32. **Lessl, M., D. Balzer, R. Lurz, V. L. Waters, D. G. Guiney, and E. Lanka.** 1992. Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region in *trans*. J. Bacteriol. **174:**2493–2500.
33. **Lindquist, S., and E. A. Craig.** 1988. The heat-shock proteins. Annu. Rev. Genet. **22:**631–677.
34. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Molina-Lopez, J. A., F. Govantes, and E. Santero.** 1994. Geometry of the process of transcription activation at the σ^{54}-dependent *nifH* promoter of *Klebsiella pneumoniae*. J. Biol. Chem. **269:**25419–25425.
36. **Morett, E., and L. Segovia.** 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175:**6067–6074.
37. **Newbury, S. F., N. H. Smith, E. C. Robinson, I. D. Hiles, and C. F. Higgins.** 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell **48:**297–310.
38. **Ninfa, A. J., L. J. Reitzer, and B. Magasanik.** 1987. Initiation of transcription at the bacterial *glnAp2* promoter by purified *E. coli* components is facilitated by enhancers. Cell **50:**1039–1046.
39. **North, A. K., K. E. Klose, K. M. Stedman, and S. Kustu.** 1993. Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. J. Bacteriol. **175:**4267–4273.
40. **Oppenheim, A. B., K. E. Rudd, I. Mendelson, and D. Teff.** 1993. Integration host factor binds to a unique class of complex repetitive extragenic DNA sequences in *Escherichia coli*. Mol. Microbiol. **10:**113–122.
41. **Parkinson, J. S., and E. C. Kofoid.** 1992. Communication modules in bacterial signaling proteins. Annu. Rev. Genet. **26:**71–112.
42. **Pérez-Martin, J., F. Rojo, and V. D. Lorenzo.** 1994. Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. Microbiol. Rev. **58:**268–290.
43. **Popham, D. L., D. Szeto, J. Keener, and S. Kustu.** 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science **243:**629–635.

44. **Reitzer, L. J., and B. Magasanik.** 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45:**785–792.
45. **Robinson, K., W. Gilbert, and G. M. Church.** 1994. Large scale bacterial gene discovery by similarity search. Nature Genet. **7:**205–214.
46. **Ronson, C. V.** 1988. Genetic regulation of C4-dicarboxylate transport in rhizobia, p. 547–551. *In* H. Bothe, F. J. deBruijn, and W. E. Newton (ed.), Nitrogen fixation: hundred years after. Gustav Fischer, Stuttgart, Germany.
47. **Russel, M., and B. Kazmierczak.** 1993. Analysis of the structure and subcellular location of filamentous phage pIV. J. Bacteriol. **175:**3998–4007.
48. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
49. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
50. **Schuler, G. D., S. F. Altschul, and D. J. Lipman.** 1991. A workbench for multiple alignment construction and analysis. Proteins **9:**180–190.
51. **Stern, M. J., G. F.-L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37:**1015–1026.
52. **Stern, M. J., E. Prossnitz, and G. F.-L. Ames.** 1988. Role of the intercistronic region in post-transcriptional regulation of gene expression in the histidine transport operon of *Salmonella typhimurium*: involvement of REP sequences. Mol. Microbiol. **2:**141–152.
53. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and the regulation of adaptive responses in bacteria. Microbiol. Rev. **53:**450–490.
54. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithase of Escherichia coli: partial purification and some properties. J. Biol. Chem. **218:**97–106.
55. **Wang, M. X., and G. M. Church.** 1992. A whole genome approach to *in vivo* DNA-protein interactions in *E. coli*. Nature (London) **360:**606–610.
56. **Weiner, L., J. L. Brissette, and P. Model.** 1991. Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on σ^{54} and modulated by positive and negative feedback mechanisms. Genes Dev. **5:**1912–1923.
57. **Weiner, L., J. L. Brissette, N. Ramani, and P. Model.** 1995. Analysis and *cis*-acting elements regulating the stress-induced phage shock protein operon. Nucleic Acids Res. **23:**2030–2036.
58. **Weiner, L., J. Dworkin, and P. Model.** Unpublished data.
59. **Weiner, L., and P. Model.** 1994. Role of an *Escherichia coli* stress-response operon in stationary-phase survival. Proc. Natl. Acad. Sci. USA **91:**2191–2195.
60. **Yura, T., H. Nagai, and H. Mori.** 1993. Regulation of the heat-shock response in bacteria. Annu. Rev. Microbiol. **47:**321–350.