JOURNAL OF BACTERIOLOGY, May 1997, p. 3095–3102 Vol. 179, No. 10
0021-9193/97/$04.00+0
Copyright © 1997, American Society for Microbiology

The RIB Element in the *goaG-pspF* Intergenic Region of *Escherichia coli*

GORAN JOVANOVIC AND PETER MODEL*

Laboratory of Genetics, The Rockefeller University, New York, New York 10021

Received 21 November 1996/Accepted 5 March 1997

The sequence (2,700 bp) between the *aldH* and *pspF* genes of *Escherichia coli* was determined. The *pspF* gene encodes a σ^{54} transcriptional activator of the phage shock protein (*psp*) operon (*pspA* to *pspE*). Downstream of the *pspF* transcribed region are two open reading frames (ORFs), *ordL* and *goaG*, convergently oriented with respect to *pspF*. These two ORFs, together with the adjacent *aldH* gene, may constitute a novel operon (*aldH-ordL-goaG*). The *goaG-pspF* intergenic region contains a complex extragenic mosaic element, RIB. The structure of this RIB element, which belongs to the BIME-1 family, is Y(REP1) > 16 < Z^1(REP2), where Y and Z^1 are palindromic units and the central 16 bases contain an L motif with an *ihf* consensus sequence. DNA fragments containing the L motif of the *psp* RIB element effectively bind integration host factor (IHF), while the Y palindromic unit (REP1) of the same RIB element binds DNA gyrase weakly. Computer prediction of the *pspF* mRNA secondary structure suggested that the transcribed stem-loop structures formed by the 3′-flanking region of the *pspF* transcript containing the RIB element can stabilize and protect *pspF* mRNA. Analysis of *pspF* steady-state mRNA levels showed that transcripts with an intact RIB element are much more abundant than those truncated at the 3′ end by deletion of either the entire RIB element or a single Z^1 sequence (REP2). Thus, the *pspF* 3′-flanking region containing the RIB element has an important role in the stabilization of the *pspF* transcript.

PspF is the constitutively transcribed and constitutively active σ^{54} transcriptional activator of the phage shock protein (*psp*) operon (*pspA* to *pspE*) of *Escherichia coli*, and its transcription is driven by one major (P1) and two minor (P2 and P3) σ^{70} promoters (27). Analysis of *pspF* transcription under noninducing and inducing conditions showed a low level of *pspF* mRNA in both (27). The DNA sequence downstream of the *pspF* gene contains an extragenic mosaic element, RIB (reiterative *ihf* BIME [bacterial interspersed mosaic element]) (27). This RIB consists of REP1 and REP2 sequences, oriented head to head and flanking a 16-bp-long DNA sequence that carries a putative binding site for integration host factor (IHF) (27). REPs are repetitive extragenic palindromes and have also been named PU (palindromic unit) sequences. A database search found over 150 sequences in *E. coli* and three in *Shigella* species that are highly homologous to all or part of the RIB element 3′ of *pspF* (27). Most such RIB elements are found in intercistronic regions or in the 3′ regions of transcriptional units (13). These elements have been proposed to be important for maintaining proper supercoiling between two convergently oriented, active transcription units (8, 29). DNA gyrase, IHF, and DNA polymerase I interact specifically with BIME DNA, suggesting a role for RIB elements in chromosomal folding and the potential formation of topological boundaries in the chromosome (5, 8, 14, 18, 19, 39, 44, 54). Because of the palindromic nature of the REP sequences and their ability to form stable stem-loop structures, some of these elements have been shown to be important for transcription termination (17, 20). The presence of BIMEs in transcribed RNA stabilizes the mRNA, probably due to protection of the RNA against a 3′-5′ exonuclease activity that can be explained by the ability of BIME (REP) RNA to form complex secondary structures (10, 24, 31, 37, 40, 50, 51, 55). Furthermore, binding of IHF to RIB might influence mRNA structure by altering the rate of dissociation of mRNA from the DNA template (8).

In this study, we sequenced and analyzed two novel open reading frames (ORFs), which we named *ordL* and *goaG*. The focus in this work was on the RIB element in the *goaG-pspF* intergenic region. To determine whether this extragenic mosaic element can really generate higher-order protein-DNA complexes, we examined the binding of IHF and DNA gyrase to fragments encompassing either all or parts of the RIB element. Finally, we explored the influence of the RIB-containing intergenic sequence on the formation of secondary structure and stability of the *pspF* mRNA.

MATERIALS AND METHODS

Bacterial strains and plasmids. *E. coli* K561 [HfrC λ^+ *relA1 spoT1* T_2^r (*ompF627 fadL701*) *lacI*^{q}] has been described previously (12, 52). Strain K1247 (PK2212) (Δ*kdpABC-5 thi-1* Δ*trkE* Tet^s *rha4 lacZ48 gal-33 malA35 zci-233::*Tn*10* λ^r λ^-) was kindly provided by P. Kuempel (Molecular, Cellular and Developmental Biology, Boulder, Colo.). Strain K1247 carries a deletion of *aldH-ordL-goaG-pspF-pspABCDE*. The deletion of Δ*aldH-pspE* was verified by using PCR and pairs of primers: G12-G11, G13-G11, G8-PSP500UP (described previously [27]), and dAsite (described previously [27])-JABR6 (26a). Strain K1247 is impaired for growth, but deletion of the putative *aldH-ordL-goaG* operon does not influence the induction and expression of the *psp* operon (in strain K1247, when transformed with plasmid pBRPS-1, which carries the entire *pspF-pspABCDE* sequences and 5′ truncated *goaG*, induction of the *psp* operon remains normal after stimulation with 10% ethanol treatment, hyperosmotic, or heat shock) (26a). 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-Bertani) or in minimal medium A supplemented with glucose (0.2%), required amino acids (30 μg/ml), and thiamine (1 μg/ml). When appropriate, antibiotics ampicillin (100 μg/ml) and tetracycline (15 μg/ml) were added. For detection of β-galactosidase activity, cultures were induced with a final concentration of 7 mM isopropyl-β-D-thiogalactopyranoside, and plates contained 0.15% X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside).

Plasmid pBRPS-1 was constructed by ligating a 4.5-kb *Eco*RI genomic fragment containing the complete *psp* operon and *pspF* gene to *Eco*RI-restricted pBR322 (53). Plasmids pMJ3 (carries a *pspF*877 mutation), pMJ5 (carries a *Bgl*II-*Stu*I fragment subcloned from pBRPS-1), and pMJ7 (carries a *Bgl*II-*Xmn*I fragment subcloned from pBRPS-1) were previously described (27). Plasmid pMJ19, used in primer extension experiments, carries the intact *pspF* regulatory sequence, the *pspF* gene, and 24 nucleotides (nt) of the immediate 3′-flanking region. To construct pMJ19, pBRPS-1 DNA was amplified by using the synthetic oligonucleotide primers JABR6 and REP1mut, which annealed to the 5′ *pspA*

* Corresponding author. Mailing address: Laboratory of Genetics, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-8648. Fax: (212) 327-7850. E-mail: model@rockvax.rockefeller.edu.

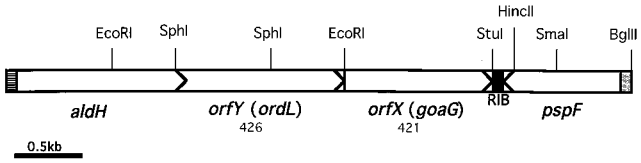


FIG. 1. Partial restriction map of the *aldH-orfY-orfX-pspF* region in *E. coli*. Black box, RIB sequence; shaded box, *pspF* promoter region; horizontally striped box, *aldH* upstream sequence; numbers, number of potentially encoded amino acids.

coding region and 24 nt downstream of the *pspF* coding sequence, respectively. The PCR product was cleaved with *Bgl*II and *Hin*dIII (created by REP1mut oligonucleotide), and the resulting 1.1-kb fragment was ligated into *Bam*HI- and *Hin*dIII-restricted pBR322. Plasmid pMJ18, used for sequencing, contains the 3′ sequence of *aldH*, all of *orfY* (*ordL*), and the 5′ sequence of *orfX* (*goaG*). To construct it, K561 chromosomal DNA was amplified by using a synthetic oligonucleotide primer (G11) that annealed to the 3′ sequence of *orfX* and a second primer (G12) covering the *Eco*RI site in the *aldH* coding region. The 2.2-kb PCR product was cleaved with *Eco*RI, and the resulting 1.8-kb fragment was ligated into pUC19 that had been digested with *Eco*RI.

Transformation of *E. coli* cells was performed as described previously (45).

DNA manipulations. For preparation of plasmid DNA and *E. coli* chromosomal DNA, electroelution and end labeling of DNA fragments, digestion with restriction enzymes, ligation, PCR amplification (*Pfu* polymerase; Stratagene), and other recombinant DNA techniques, well-established procedures were used (3, 45). The synthetic oligonucleotides (synthesized by Operon Technologies Inc., Alameda, Calif.) used for DNA sequencing, PCR amplification, and primer extension experiments were as follows: G5, 5′-CACAGGTCAGCAGGA-3′; G6, 5′-GATTGCACAGAAAAT-3′; G7, 5′-GTTCTTCACCACCGG-3′; G8, 5′-CCGGGCGGGCTTGGC-3′; G9, 5′-CGCCAGCGGGTTACC-3′; G10, 5′-CGGTCGTACGTATAT-3′; G11, 5′-GGTCAGCGCCATGGT-3′; G12, 5′-ATCGCCGATGAATTCTT-3′; G13, 5′-GGCGAAACGCAGCATGC-3′; G14, 5′-TCAATTTTCACGCCC-3′; G15, 5′-TCATCTAGCCGGACG-3′; GOJO30RIB, 5′-CATCCCGGGATGCGCAAT-3′; GOJO31, 5′-CCGCTGGATTTACGTGAG-3′; REP1mut, 5′-GGCAAGCTTGTATTGCTCAAC-3′; REP2mut, 5′-ATCAAGCTTGCAAACGACGTA-3′; JABR6, 5′-CGATGTCGGCAAAGCG-3′; and PE-F, 5′-CCTGTTCCAGCACTTCGAGAAAGCTGTTCG-3′.

The nucleotide sequences of *orfY* (*ordL*) and *orfX* (*goaG*) were determined for both strands by using single-stranded M13mp18 phage DNA. The 1.5-kb *Eco*RI-*Hin*cII DNA fragment from pBRPS-1 was ligated into *Eco*RI and *Hin*cII-restricted RfM13mp18; the G11-G13 PCR-generated fragment (0.6 kb) was cleaved with *Eco*RI and *Sph*I and subcloned into RfM13mp18 (digested with *Eco*RI and *Sph*I) or double-stranded DNA (plasmids pBRPS-1 and pMJ18) templates. DNA sequencing was performed by the dideoxynucleotide chain termination method (46). Sequencing reactions were carried out with a Sequenase sequencing kit (U.S. Biochemical Corp.), and primers were the universal and reverse M13 primers or specific synthetic oligonucleotides (G5-G11 and G13-G15). The sequencing products were resolved on denaturing 5 or 8% polyacrylamide gels.

DNA fragments used in gel mobility shift assay were constructed by using PCR amplification and the following pairs of primers: RIB fragment (333 bp), GOJO31 and GOJO30RIB oligonucleotides; ΔREP2 fragment (237 bp), GOJO31 and REP2mut oligonucleotides; ΔRIB fragment (177-bp), GOJO31 and REP1mut oligonucleotides.

Computer analyses. Nucleotide sequence and protein data were analyzed by using the Gen Info network BLAST server (BLASTN, BLASTP, and BLASTX) (2) and MACAW, with the PAM120 matrix provided by the National Center for Biotechnology Information (47). The hydropathy and secondary structure plots were obtained by using algorithms by Kyte and Doolittle (28) and Garnier et al. (16), respectively. The RNA and DNA secondary structure analyses were carried out by using Wisconsin Genetics Computer Group (GCG) program MFOLD (56).

Gel mobility shift assay. The RIB, ΔREP2, and ΔRIB PCR-generated fragments were end labeled with [γ-^{32}P]ATP (3,000 Ci mmol^{-1}; New England Nuclear, Boston, Mass.), using T4 polynucleotide kinase (New England Biolabs). The gel mobility shift assay has been described previously (41). Briefly, 2 ng of each fragment was incubated at 25°C with the desired concentrations of IHF (kindly provided by H. A. Nash, National Institute of Mental Health Bethesda, Md.) or/and DNA gyrase (kindly provided by M. Gellert, National Institutes of Health Bethesda, Md.) (GyrA and GyrB subunits were added in equimolar concentrations) in 50 mM Tris-HCl (pH 7.4)–70 mM KCl–10% glycerol–1 mM EDTA–1 mM β-mercaptoethanol–7 mM $MgCl_2$–200 μg of bovine serum albumin per ml (20 μl, final volume). Typically, a 100-fold weight excess of poly(dI-dC) (Pharmacia) was included in all reaction mixtures. After 20 min, 2 μl of 50% glycerol–0.1% xylene cyanol solution was added, and the mixture was loaded on a 5% nondenaturing acrylamide gel. Electrophoresis was done at 4°C in 0.5× Tris-borate-EDTA buffer at 14 V/cm for 2 h. The gels had been prerun for several hours to stabilize the conductivity.

Primer extension reaction. Primer extension analysis (45) was carried out by using a synthetic oligonucleotide (PE-F) complementary to nt 35 to 64 of the *pspF* coding sequence. The oligonucleotide was labeled at the 5′ end with [γ-^{32}P]ATP (3,000 Ci mmol^{-1}) and hybridized to 300 μg of total RNA isolated from *E. coli* K1247, K1247(pMJ7), K1247(pMJ5), K1247(pMJ19), and K1247 (pMJ3). RNA was extracted from late-log-phase cultures of *E. coli* grown in minimal medium at 37°C by the hot acid-phenol method (1). The quality of isolated total RNA used in primer extension reaction was verified by using a spectrophotometer (A_{260}/A_{280}) and by gel electrophoresis. 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 and a mixture with no RNA present were subjected to electrophoresis on a 5% polyacrylamide sequencing gel along with a nucleotide sequence generated with the same (unlabeled) primer.

Nucleotide sequence accession number. The nucleotide sequences of the *ordL* and *goaG* genes have been submitted to the GenBank data bank and assigned accession no. U38543.

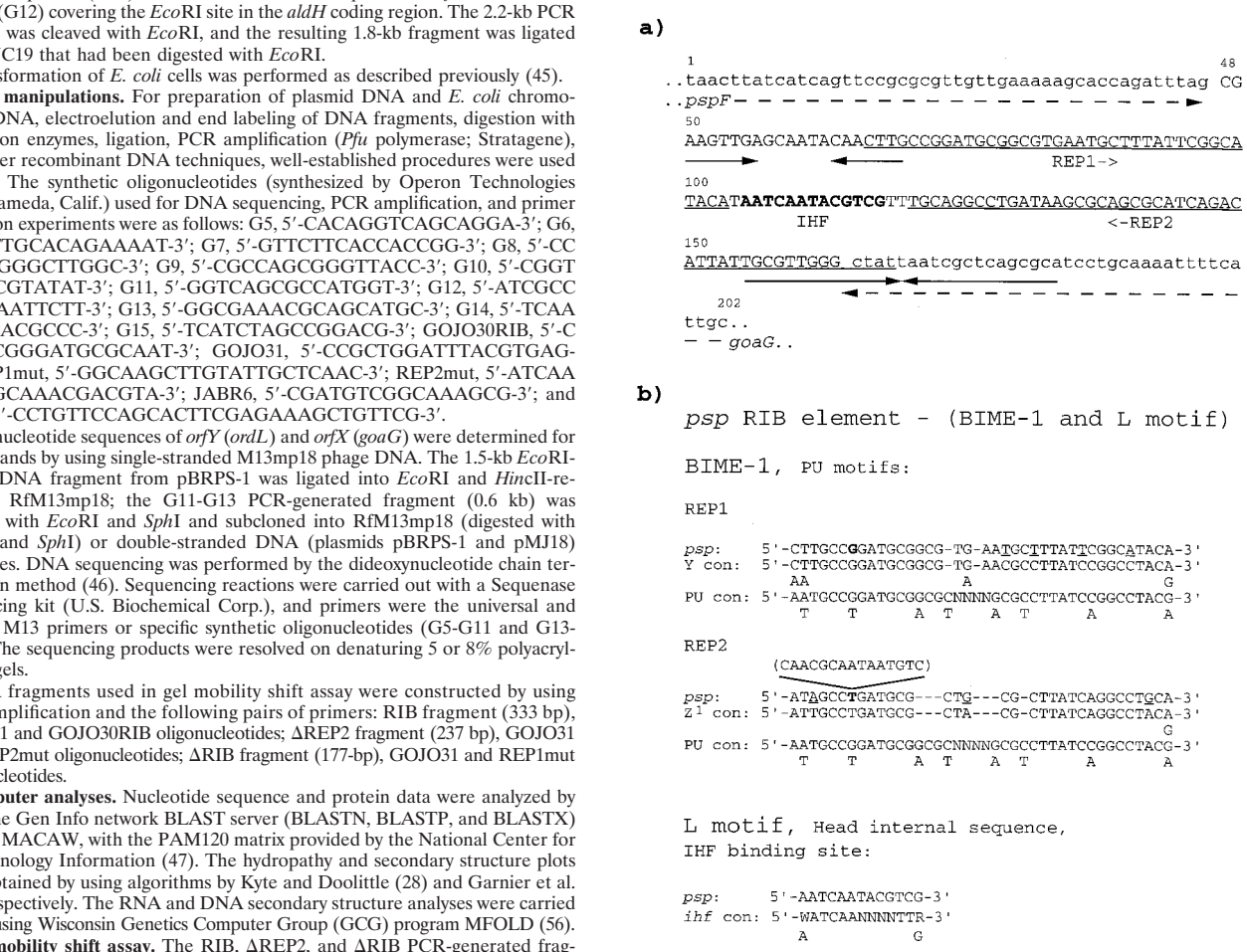


FIG. 2. *goaG-pspF* intergenic region contains the RIB element. (a) Nucleotide sequence of the *goaG-pspF* intergenic region. Dashed bold arrows, 3′ regions of *pspF* (1 to 47) and *goaG* (202 to 164); bold arrows, palindromic sequences 55-69 and 155-182; underlined nucleotides, REP1 and REP2; small arrows, 5′-3′ orientation of the REP sequences; bold nucleotides, IHF consensus sequence. (b) The *psp* RIB element belongs to the L-motif-containing BIME-1 family. BIME-1, PU motifs: alignment of the REP1 and REP2 sequences from the *goaG-pspF* intergenic region (*psp*) with the BIME-1 REPs (PU motifs), type Y, and type Z^1 canonical consensus sequences (PU con, Y con, and Z^1 con) (4). Bold nucleotides (position 7), nucleotides important for distinguishing the PU types; underlined nucleotides, mismatches with consensus sequences; nucleotides in parentheses, 15-bp insertion in the REP2 sequence. L motif, head internal sequence, IHF-binding site: alignment of the IHF consensus sequence of the *psp* RIB L motif (*psp*) with the IHF canonical consensus sequence (*ihf* con) (11, 15).

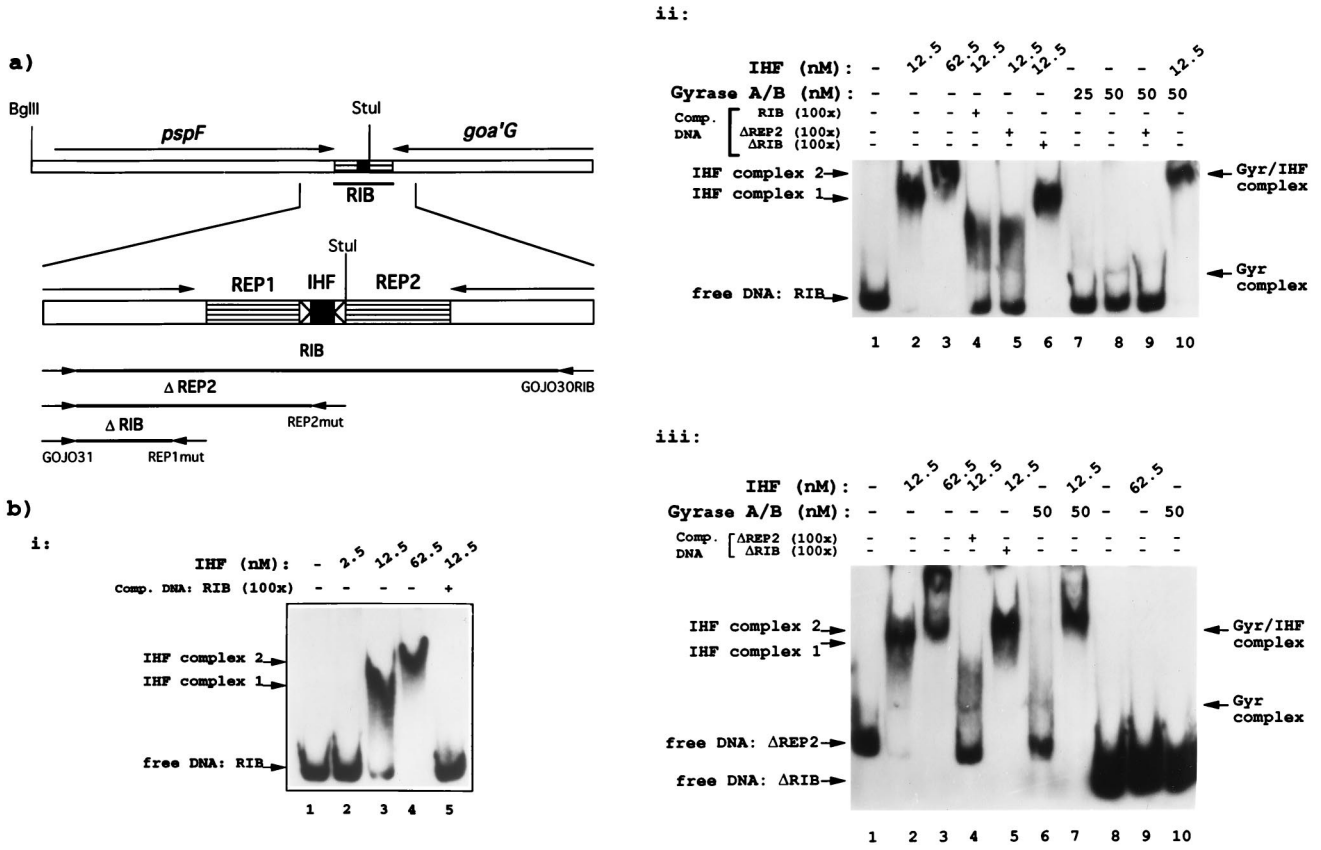


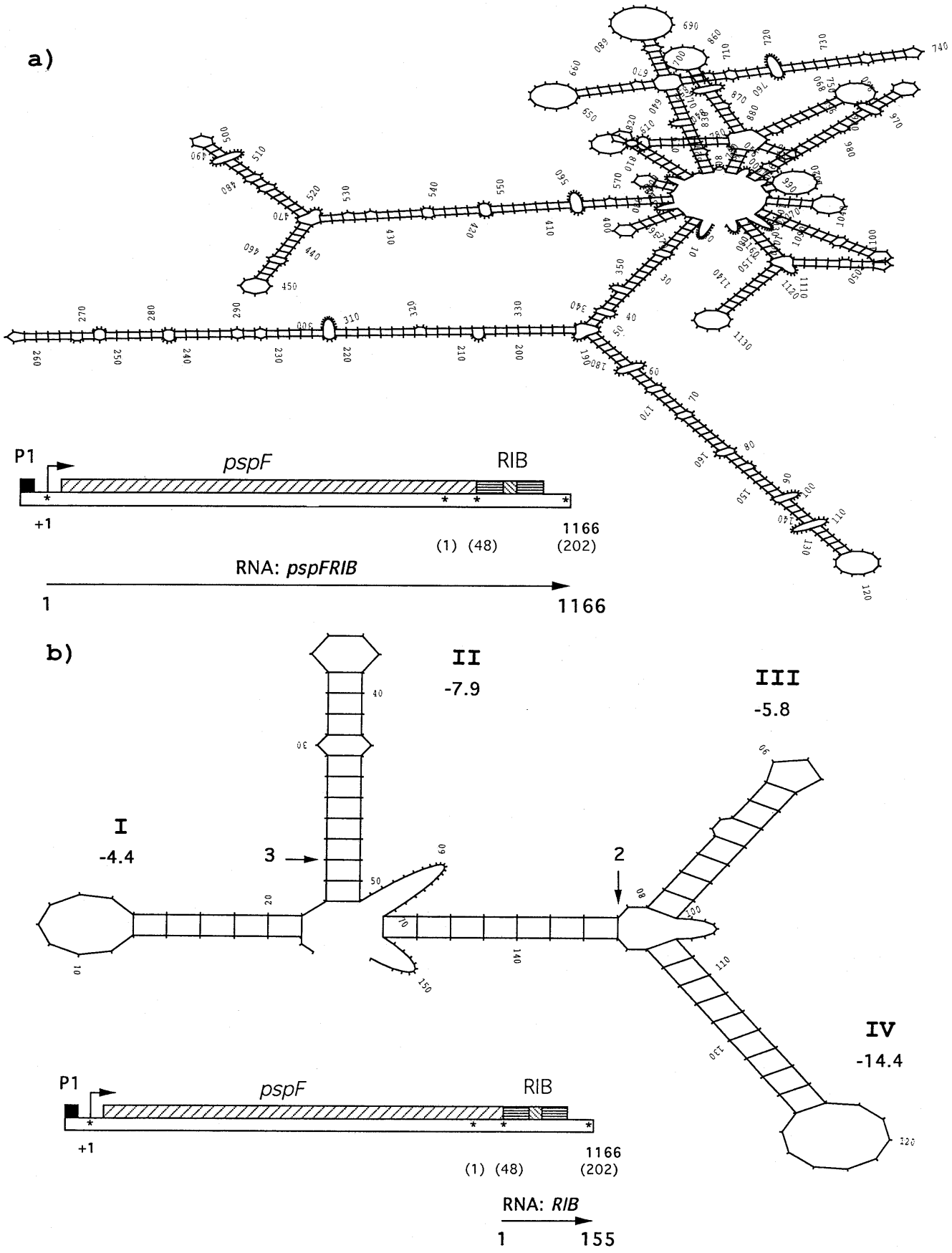
FIG. 3. IHF and DNA gyrase bind DNA in the *goaG-pspF* intergenic region. (a) Construction of the linear templates RIB, ΔREP2, and ΔRIB, used for gel mobility shift assays with IHF and DNA gyrase. Templates were generated by using PCR and primers GOJO31, GOJO30RIB, REP2mut, and REP1mut. Black box, IHF consensus sequence; blocks with horizontal lines, REP sequences. (b) Gel mobility shift assays using the IHF and DNA gyrase proteins and templates (2 ng) RIB, ΔREP2, and ΔRIB. (i) Gel mobility shift assay using IHF protein in the indicated molar concentrations and the 333-bp RIB fragment (2 ng) as a probe. Specific competitor (Comp.) DNA (RIB) was added in 100-fold molar excess (lane 5) prior to addition of the protein to the binding reaction. Gyr, DNA gyrase. (ii) Gel mobility shift assays using IHF (lanes 2 to 6), DNA gyrase (lanes 7 to 9), both IHF and DNA gyrase (lane 10) in the indicated molar concentrations, and RIB fragment (2 ng) as a probe. DNA gyrase subunits A and B were added in equimolar concentrations. Specific and nonspecific DNA competitors (lane 4, RIB; lanes 5 and 9, ΔREP2; and lane 6, ΔRIB) were added as described above. (iii) Gel mobility shift assays using IHF protein (lanes 2 to 5 and 9), DNA gyrase (lanes 6 and 10), both IHF and DNA gyrase (lane 7) in the indicated molar concentrations, and 237-bp ΔREP2 (2 ng) (lanes 1 to 7) and 177-bp ΔRIB (2 ng) (lanes 8 to 10) fragments as probes. Specific and nonspecific DNA competitors (lane 4, ΔREP2; and lane 5, ΔRIB) were added in molar excesses as described above.

RESULTS AND DISCUSSION

Nucleotide sequence of *orfY* and *orfX*. As shown previously, the region downstream of the *pspF* gene contains an extragenic mosaic element, RIB (27). Previously, 2.0 kb further downstream of the *pspF* gene, the position of the convergently oriented *aldH* gene and a partial ORF (754 nt) beginning 2 nt downstream from *aldH* had been determined (22), which could be part of a larger operon whose sequence might extend to the RIB element 3′ of *pspF*. Thus, the nucleotide sequence of the region between *aldH* and *pspF* was determined. We found two complete ORFs, 1,281-nt-long *orfY* and 1,269-nt-long *orfX*, both convergently oriented to the *pspF* gene (Fig. 1). The 5′ region of *orfY* contains the previously determined partial ORF that follows *aldH*. The predicted *orfX* start codon is 37 nt downstream of *orfY*. Both ORFs have putative ribosome-binding sites but not obvious upstream promoters. Hence, *aldH* together with these two genes may constitute an operon.

orfY codes for a 47.1-kDa (426-amino-acids) protein, which we have named OrdL. The hypothetical OrdL protein with an NAD/flavin adenine dinucleotide-binding motif shows partial homology to *E. coli* anaerobic glycerol-3-phosphate dehydrogenase and to human amine oxidase. We have named *orfY ordL*, for oxidoreductase L. *orfX* codes for a 44.7-kDa (421-amino-acid) protein with a putative pyridoxal phosphate-binding enzyme motif 1. Computer analyses of its hydrophilicity (28) and predicted secondary structure (16) showed several potential membrane-spanning domains in the C-terminal region. *orfX* has striking homology (50% identity) to γ-aminobutyric acid (GABA) aminotransferase of *E. coli* and *Mycobacterium leprae* and 46% identity to acetyl-ornithine aminotransferase of *E. coli* and *Bacillus subtilis*. Thus, we have called *orfX goaG*, for GABA or ornithine aminotransferase.

Although the functions of OrdL and GoaG are not known, it seems likely that they, together with AldH, could participate in some as yet ill-defined metabolic pathway involving oxidation-reduction and amination-deamination reactions. Enzymes with such activities are clustered in operons dedicated to the metabolism of arginine and ornithine in *E. coli*, such as the *gabDTP* cluster at 57 min and the *speABC* cluster at 66 min (6, 33, 38). A most puzzling observation is that the amino acid sequence of the activator of the σ^{54}-dependent *pspA-E* operon, PspF, is more similar to *E. coli* ornithine decarboxylase inhibitor (Az) than to any other sequence in the data bank (27). Az counteracts the activity of ornithine decarboxylase (23), and Canellakis et al. predicted it to have a second role as a σ^{54} transcriptional activator (9). Az, however, has an additional


a)
P1
pspF
RIB
+1
1166
(1) (48)
(202)
RNA: pspFRIB
1
1166
b)
I
-4.4
II
-7.9
III
-5.8
IV
-14.4
3
2
P1
pspF
RIB
+1
1166
(1) (48)
(202)
RNA: RIB
1
155

N-terminal domain which is absent from PspF. Ornithine decarboxylase generates the precursor of GABA, is essential for the utilization of ornithine and arginine, and is involved in the biosynthesis of polyamines (33, 49). An ancestral form of PspF may have been involved in sensing amino acids or their degradation products. Hence, the colocalization of *pspF* and the three genes of the putative *aldH-ordL-goaG* operon may reflect a movement of these genes as what was originally a functional unit.

The conditions under which the putative *aldH-ordL-goaG* operon might be transcriptionally active are uncertain. Northern blot analysis of RNA isolated from *E. coli* K561 grown under aerobic (37°C), anaerobic (37°C), or *psp* operon-inducing (f1 pIV induction at 37°C) conditions did not reveal an RNA that would code for *aldH*, *ordL*, and *goaG* (data not shown). However, the Northern blot analysis may not be sensitive enough to detect a very low level of transcription of these genes.

Analysis of the *goaG-pspF* intergenic region. The RIB element located in the *goaG-pspF* intergenic region (*psp* RIB element) consists of the PU sequences REP1 and REP2 flanking a canonical IHF-binding site (Fig. 2). GenBank release 82.0 contains more than 150 sequences that are highly homologous to the *psp* RIB element (27). This element belongs to the BIME family of mosaic combinations of several short-sequence motifs that are spread through the *E. coli* genome (13, 21, 39). The types of PUs can be distinguished according to the seventh position of the consensus (G for Y and T for Z type) (5, 20) (Fig. 2b). The REP1 sequence (positions 66 to 103 in Fig. 2a) resembles (with the exception of 3 nt) the Y type of PU, while the REP2 sequence (positions 120 to 167 in Fig. 2a) has the Z^1 PU-type consensus sequence (with the exception of 3 nt and with an insertion of 15 nt after position 6 (Fig. 2b). The REP1 and REP2 sequences are oriented head to head and separated by a 16-bp head internal sequence that contains the L motif with a consensus sequence for IHF binding (*ihf*) (Fig. 2). BIME-1 elements consist of two PUs, Y and Z^1, that flank an L motif (32 to 34 bp) which contains an IHF-binding site (*ihf*) sequence (4, 5, 15). The majority of RIB elements belong to the BIME-1 family (8, 13, 21) and are usually located after the last gene of an operon (5, 8, 39). The organization of the *psp* RIB element is like that of the BIME-1 subset of the BIME family except that its internal sequence (16 bp) is shorter.

In addition to the RIB element, palindromic sequences that we call the 50-69 palindrome and the 155-182 palindrome (positions 50 to 69 and 155 to 182 in Fig. 2a) are found. These palindromes partially overlap REP1 and REP2 (Fig. 2a). Together with the RIB element and IHF-binding site, they may contribute to the formation of higher-order DNA secondary structures and may participate in different functions involved in the organization of the bacterial chromatin and/or gene expression (4, 10, 25, 36, 37, 40, 55).

Binding of IHF and DNA gyrase to the *psp* RIB element. The RIB elements of *E. coli* specifically bind nucleoid-associated proteins (14, 18). It has been shown that DNA gyrase and DNA polymerase I bind to PU motifs in the BIME family, while IHF binds the *ihf* consensus sequence (YAANNNNTT GATW) in the center of the L motif (4, 5, 8, 11, 15, 19, 39, 54). We used a gel mobility shift assay to demonstrate that IHF and DNA gyrase also bind to the *psp* RIB element. As illustrated in Fig. 3a, we generated the following DNA fragments: RIB, which covers the entire *goaG-pspF* intergenic region; ΔRIB2, which lacks the REP2 sequence; and ΔRIB, a fragment that lacks the most of the *pspF* 3′-flanking sequence including the RIB element. Figure 3 shows that IHF binds specifically to the RIB and ΔREP2 fragments, which contain the L motif, but not with the ΔRIB fragment (Fig. 3b-iii, lane 9), which lacks the L motif. A second complex was detected at the higher IHF concentration (62.5 nM) (Fig. 3b-i, lane 4; Fig. 3b-ii, lane 3; Fig. 3b-iii, lane 2), suggesting the existence of an additional, low-affinity IHF-binding site in the REP1 sequence.

The gel retardation assay detected a weak interaction of DNA gyrase with the RIB (Fig. 3b-ii, lane 8) and ΔREP2 (Fig. 3b-iii, lane 6) templates but not the ΔRIB fragment (Fig. 3b-iii, lane 9). These results suggest that DNA gyrase binds to REP1 (Y PU type) sequence but not the REP2 (Z^1 PU type) sequence in which the Z^1 consensus is disrupted by a 15-bp insertion. When used together, IHF (12.5 nM) and DNA gyrase (50 nM) produced a modest supershift of RIB (Fig. 3b-ii, lane 10) and ΔREP2 (Fig. 3b-iii, lane 7) templates relative to the positions of the same fragments in the presence of IHF (12.5 nM) alone (Fig. 3b-ii and b-iii). The bulk of the template supershifts, indicating that in the presence of IHF, DNA gyrase binds REP1 more effectively. It has been shown that DNA gyrase binds preferentially to the Y type of PU motif, with lower affinity for the Z^2 type and about ninefold-lower affinity for the Z^1 type (4, 5, 54). Generally, DNA gyrase has much higher affinity for the BIME-2 (Y-Z^2) than for the BIME-1 (Y-Z^1) family, probably due to a cooperative binding of one DNA gyrase dimer simultaneously to Y and a Z^2. As shown here, the composition of PU motifs in *psp* RIB element corresponds to the BIME-1 family, where the REP1 sequence belongs to the Y PU type, while the imperfect REP2 sequence is a member of the Z^1 PU type. These facts could explain weak DNA gyrase binding to the REP1 sequence alone in the *psp* RIB element.

IHF binding induces a DNA bend of more than 160° (43). The function of IHF bound to the *ihf* consensus sequence in the *psp* RIB L motif could be to facilitate and stabilize the formation of a DNA gyrase-REP1 complex. The higher-order DNA structure due to binding of IHF and DNA gyrase to the *psp* RIB element could act as a topological boundary in the chromosome and/or could maintain the negative supercoiling and normal transcription of two convergently oriented transcriptional units (*pspF* and *aldH-ordL-goaG*).

In *E. coli*, the formation of positive supercoils is facilitated if the nascent RNA codes for a protein carrying a transmembrane domain which is cotranslationally inserted into the membrane, thereby acting as an anchor and further restricting the rotation of the transcriptional machinery (30, 32, 37, 42). Since most RIBs are found at the 3′ ends of transcription units encoding membrane-associated proteins (4, 8), it is intriguing

FIG. 4. Potential secondary structure of predicted *pspF* mRNA containing the 3′-flanking region. (a) A secondary structure model (GCG program MFOLD [56]) using the RNA sequence (*pspFRIB*; 1,166 nt) schematically presented at the bottom. (b) The secondary structure model of the *RIB* RNA sequence (155 nt). I to IV, stem-loop structures. Stem-loop I is formed by the 55-69 palindrome, stem-loop II is formed by the REP1 sequence, stem-loop III is formed by the REP2 sequence, and stem-loop IV is formed by a combination of REP2 and 155-182 palindrome sequences. The sequence is schematically presented at the bottom (3′-flanking region of *pspF*, including the RIB sequence). Numbers in parentheses correspond to nucleotide numbers in Fig. 2a. Calculated free energies of putative stem-loops I to IV (kilocalories per mol) are indicated near each stem-loop region. Arrows, deletion points in constructs 2 (deletes stem-loops III and IV) and 3 (deletes stem-loops II to IV) (Fig. 5a). Symbols on the diagrams at the bottom: black box, major *pspF* σ^{70} promoter (P1); broken arrow, *pspF* start of transcription (+1); hatched block, *pspF* gene; blocks with horizontal lines, REP sequences; hatched box, IHF-binding site; RIB, REP sequences and IHF-binding site of *psp* RIB element.

that the C-terminal region of the hypothetical GoaG protein contains several potential membrane-spanning domains. If under certain conditions both transcriptional units, *pspF* and *aldH-ordL-goaG*, were active, the RIB element is appropriately placed to relieve the unfavorable DNA topology created by cotranslational membrane insertion.

Potential secondary structure of the *pspF*-3′-flanking region RNA. Examination of optimal and suboptimal RNA secondary structure models generated by the Zuker algorithm (56) (GCG program MFOLD) by using the *pspF* transcript, which includes the 3′-flanking region as input, showed that the 1,166-nt RNA can be folded into several alternative conformations with similar stabilities (not shown). In all of the predicted structures, the 3′-flanking region of *pspF* formed the same stem-loop structures (positions 1011 to 1166) containing the RIB element and additional palindromic sequences (Fig. 4a). Similar analysis of the 155 nt (*RIB* in Fig. 4b) produced a single predicted structure identical to that generated from the entire *pspF* RNA transcript (*pspFRIB* in Fig. 4a). The predicted structure is composed of four stable stem-loop structures with a combined stability of -42.8 kcal mol^{-1} (Fig. 4b). Stem-loop I is formed by the 55-69 palindrome, stem-loop II is formed by the REP1 sequence, stem-loop III is formed by the REP2 sequence, and stem-loop IV is formed by a combination of the REP2 and 155-182 palindrome sequences (Fig. 4b). Although the prediction generated by this program is said to favor long-range interactions (48), this is not a concern here because the same structure was predicted by both the full-length transcript and the 3′-flanking region alone. Thus, the structure illustrated in Fig. 4b might exist in vivo.

The potential stem-loop structures that could be formed in the intergenic *pspF-goaG* region by the *psp* RIB element could be the transcriptional terminator for *pspF*. The transcribed stem-loops could also stabilize the 3′ end of *pspF* mRNA, as has been shown for other BIMEs (7, 10, 24, 25, 26, 35, 36, 37, 51).

The 3′-flanking region stabilizes *pspF* mRNA. The level of gene expression can be affected by differences in the rate of mRNA degradation. Hence, if the *psp* RIB element in concert with other palindromic sequences contributes to *pspF* mRNA stability in vivo, we should be able to detect a different, decreased level of 3′-truncated *pspF* transcripts. The in vivo steady-state level of *pspF* mRNA was measured by primer extension using plasmids that contained different portions of the 3′-flanking region of *pspF* (Fig. 5a). Construct 1 (pMJ7) carries the intact *pspF* gene and the entire 3′-flanking region (Fig. 5a). Construct 2 (pMJ5) carries the intact *pspF* gene and part of the 3′-flanking region but lacks REP2 and additional palindromic sequences (Fig. 5a). This deletion point corresponds to nt 125 in Fig. 2a and removes predicted stem-loops III and IV (Fig. 4b). Construct 3 (pMJ19) carries the *pspF* coding sequence but lacks most of the 3′-flanking region (Fig. 5a). The deletion point in this construct corresponds to nt 71 in Fig. 2a and removes stem-loops II to IV, leaving only stem-loop I intact (Fig. 4b). Construct 4 (pMJ3) carries the mutation *pspF*877 (27), which lacks the 3′ region of the coding sequence of the *pspF* gene (Fig. 5a).

Plasmids containing *pspF* genes with different 3′ ends (Fig. 5a) were introduced into *E. coli* K1247 (Δ*aldH-pspE*). Primer extension analysis of total RNA isolated from cells carrying these plasmids reveals strong differences in the amount of *pspF* transcripts (Fig. 5b). Removal of the most of the *pspF* 3′-flanking sequence (construct 4) and/or deletion of the putative stem-loop structures II to IV (construct 3) markedly decreased the steady-state level of *pspF* mRNA (Fig. 5b, lanes 3 and 4). The amount of *pspF* mRNA derived from the construct lacking

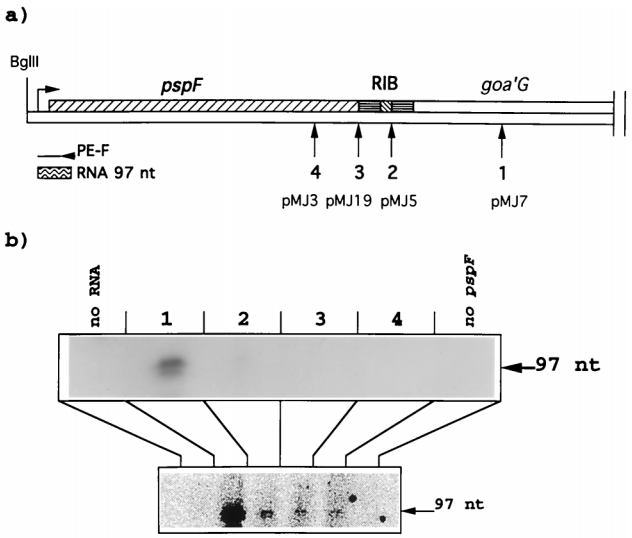


FIG. 5. The 3′-flanking region stabilizes the upstream *pspF* mRNA. (a) Construction of plasmids carrying different lengths of the 3′-flanking region of *pspF*. 1 to 4, deletion points in recombinant plasmids pMJ7, pMJ5, pMJ19, and pMJ3, respectively (see Materials and Methods). *pspF* gene (hatched block), 3′ sequence of *goaG* (shaded block), *psp* RIB element (horizontal line blocks, REP sequences; hatched box, IHF-binding site) are indicated. Bent arrow, *pspF* start of transcription corresponding to major P1 σ^{70} promoter; PE-F, primer used in primer extension reactions described below; waved block, 97-nt-long RNA product synthesized after primer extension reactions described below. (b) Primer extension analysis of the steady-state level of *pspF* mRNA in strains carrying the *pspF* transcriptional units with different 3′ ends. Primer extension reactions were carried out with primer PE-F (described above) and 300 μg of total RNAs isolated from strain K1247 [Δ(*aldH-ordL-goaG-pspF-pspABCDE*)] carrying the plasmids described above. No RNA, reaction mix containing the PE-F primer and no RNA; lane 1, PE-F-primed reaction with total RNA (A_{260}/A_{280} = 2.17) isolated from K1247(pMJ7); lane 2, PE-F-primed reaction with total RNA (A_{260}/A_{280} = 2.09) isolated from K1247(pMJ5); lane 3, PE-F-primed reaction with total RNA (A_{260}/A_{280} = 2.04) isolated from K1247(pMJ19); lane 4, PE-F-primed reaction with total RNA (A_{260}/A_{280} = 2.12) isolated from K1247(pMJ3); no *pspF*, PE-F-primed reaction with total RNA (A_{260}/A_{280} = 2.15) isolated from K1247. In the experiment using autoradiography, a signal was detected only in lane 1. Data shown at the bottom are from a print of a phosphorimager scan.

stem-loops III and IV (Fig. 5b, lane 2), predicted to be most stable (-14.4 kcal mol^{-1} in Fig. 4b) (construct 2), is slightly higher than that from constructs 3 and 4.

This primer extension analysis of *pspF* steady-state mRNA level showed that transcriptional units with different 3′ ends have different stabilities. The *pspF* transcript with an intact RIB element and additional palindromic sequences was stable, while 3′-truncated transcripts deleted for most of the 3′-flanking region (except for stem-loop I formed by the 50-69 palindrome), including the entire RIB element, or for a single REP2 sequence and the 155-182 palindromic sequence, were extremely unstable (Fig. 5). Therefore, the presence of an intact RIB and additional palindromic sequences stabilizes the *pspF* transcript.

Most mRNA degradation in *E. coli* is achieved by polynucleotide phosphorylase and RNase II, which remove ribonucleotides processively from the 3′ ends of mRNA molecules (25). Exoribonuclease-impeding factor, which binds to different stem-loop structures (10), may further impede exoribonuclease activity in vivo. The stem-loop structures at the 3′ end of the *pspF* transcript may block these enzymes as has been shown for *E. coli glyA*, *malEFG*, and *atp* transcripts (37, 40, 55). It has been shown that one REP sequence does not stabilize mRNA (10, 34, 37). Our experiments confirmed this observation since the construct with an intact REP1 and deleted REP2 sequence,

in combination with the 155-182 palindrome, had a much decreased steady-state level of *pspF* mRNA (Fig. 5b, line 2). These results demonstrate that full protection and stabilization of the *pspF* mRNA in vivo can be achieved only with an intact *psp* RIB element in combination with the 155-182 palindrome.

These results elucidate one of the possible roles of the RIB element in intergenic region of the putative *aldH-ordL-goaG* operon and the *pspF*.

ACKNOWLEDGMENTS

We thank Howard A. Nash (NIMH, Bethesda, Md.) for providing the IHF protein, Martin Gellert (NIH, Bethesda, Md.) for providing the DNA gyrase protein, Peter Kuempel (MCDB, Boulder, Colo.) for providing the PK2212 (K1247) strain, and Marjorie Russel and Norton D. Zinder for critical reading and helpful comments on 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.

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. **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.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.)**, 1989. Current protocols in molecular biology. Green Publishing Associates/Wiley Interscience, New York, N.Y.
4. **Bachellier, S., E. Gilson, M. Hofnung, and C. W. Hill.** 1996. Repeated Sequences, p. 2012–2040. *In* F. C. Neidhart et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
5. **Bachellier, S., W. Saurin, D. Perrin, M. Hofnung, and E. Gibson.** 1994. Structural and functional diversity among bacterial interspersed mosaic elements (BIMEs). Mol. Microbiol. **12:**61–70.
6. **Bartsch, K., A. van Johnn-Marteville, and A. Schulz.** 1990. Molecular analysis of two genes of the *Escherichia coli gab* cluster: nucleotide sequence of the glutamine:succinic semialdehyde gene (*gabT*) and characterization of the succinic semialdehyde dehydrogenase gene (*gabD*). J. Bacteriol. **172:**7035–7042.
7. **Belasko, J. G., J. T. Beatty, C. W. Adams, A. von Gabain, and S. Cohen.** 1985. Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic *rxc* A transcript. Cell **40:**171–181.
8. **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.
9. **Canellakis, E. S., A. A. Paterakis, S.-C. Huang, C. A. Panagiotidis, and D. A. Kyriakidis.** 1993. Identification, cloning, and nucleotide sequencing of the ornithine decarboxylase antizyme gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **90:**7129–7133.
10. **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.
11. **Craig, N. L., and H. A. Nash.** 1984. *E. coli* integration host factor binds to specific sites in DNA. Cell **39:**707–716.
12. **Davis, N. G., J. D. Boeke, and P. Model.** 1985. Fine structure of a membrane anchor domain. J. Mol. Biol. **181:**111–121.
13. **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.
14. **Echols, H.** 1986. Multiple DNA-protein interactions governing high-precision DNA transactions. Science **233:**1050–1056.
15. **Gamas, P., M. G. Chandler, P. Prentki, and D. J. Galas.** 1987. *Escherichia coli* integration host factor binds specifically to the ends of the insertion sequence IS*1* and to its major insertion hot-spot in pBR322. J. Mol. Biol. **195:**261–272.
16. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120:**97–120.
17. **Gilson, E., J.-M. Clement, D. Perrin, and M. Hofnung.** 1987. Palindromic units: a case of highly repetitive DNA sequences in bacteria. Trends Genet. **3:**226–230.
18. **Gilson, E., D. Perrin, J.-M. Clement, S. Szmelcman, E. Dassa, and M. Hofnung.** 1986. Palindromic units form *E. coli* as binding sites for a chromoid-associated protein. FEBS Lett. **206:**323–328.
19. **Gilson, E., D. Perrin, and M. Hofnung.** 1990. DNA polymerase I and a protein complex bind specifically to *E. coli* palindromic units highly repetitive DNA: implications for bacterial chromosome organization. Nucleic Acids Res. **18:**3941–3952.
20. **Gilson, E., J.-P. Rousset, J.-M. Clement, and M. Hofnung.** 1986. A subfamily of E. coli palindromic units implicated in transcription termination? Ann. Inst. Pasteur Microbiol. **137B:**259–270.
21. **Gilson, E., W. Saurin, D. Perrin, S. Bachellier, and M. Hofnung.** 1991. Palindromic units are part of a new bacterial interspersed mosaic element (BIME). Nucleic Acids Res. **19:**1375–1383.
22. **Heim, R., and E. E. Strehler.** 1991. Cloning an *Escherichia coli* gene encoding a protein remarkably similar to mammalian aldehyde dehydrogenases. Gene **99:**15–23.
23. **Heller, J. S., W. F. Fong, and E. S. Canellakis.** 1976. Induction of a protein inhibitor to ornithine decarboxylase by the end products of its reaction. Proc. Natl. Acad. Sci. USA **73:**1858–1862.
24. **Higgins, C. F., G. F.-L. Ames, W. M. Barnes, J.-M. Clement, and M. Hofnung.** 1982. A novel intercistronic regulatory element of prokaryotic operons. Nature **298:**760–762.
25. **Higgins, C. F., H. C. Causton, G. S. C. Dance, and E. A. Mudd.** 1993. The role of the 3′ end in mRNA stability and decay, p. 13–30. *In* J. Belasco and G. Brawerman (ed.), Control of messenger RNA stability. Academic Press, Inc., San Diego, Calif.
26. **Higgins, C. F., R. S. McClaren, and S. F. Newbury.** 1988. Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? Gene **72:**3–14.

26a. **Jovanovic, G., and P. Model.** Unpublished data.

27. **Jovanovic, G., L. Weiner, and P. Model.** 1996. Identification, nucleotide sequence, and characterization of PspF, the transcriptional activator of the *Escherichia coli* stress-induced *psp* operon. J. Bacteriol. **178:**1936–1945.
28. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
29. **Liu, L. F., and J. C. Wang.** 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA **84:**7024–7027.
30. **Lodge, J. K., T. Kazic, and D. E. Berg.** 1989. Formation of supercoiling domains in plasmid pBR322. J. Bacteriol. **171:**2181–2187.
31. **Lupski, J. R., and G. M. Weinstock.** 1992. Short, interspersed repetitive DNA sequences in prokaryotic genomes. J. Bacteriol. **174:**4525–4529.
32. **Lynch, A. S., and J. C. Wang.** 1993. Anchoring of DNA to the bacterial cytoplasmic membrane through cotranscriptional synthesis of polypeptides encoding membrane proteins or proteins for export: a mechanism of plasmid hypernegative supercoiling in mutants deficient in DNA topoisomerase I. J. Bacteriol. **175:**1645–1655.
33. **Maas, W. K.** 1972. Mapping of genes involved in the synthesis of spermidine in *E. coli*. Mol. Gen. Genet. **119:**1–9.
34. **Meyer, B. J., and J. L. Schottel.** 1992. Characterization of *cat* messenger RNA decay suggests that turnover occurs by endonucleolytic cleavage in a 3′ to 5′ direction. Mol. Microbiol. **6:**1095–1104.
35. **Mott, J. E., J. L. Galloway, and T. Platt.** 1985. Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3′ exonucleolytic processing after Rho-dependent termination. EMBO J. **4:**1887–1891.
36. **Newbury, S. F., N. H. Smith, and C. F. Higgins.** 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. Cell **51:**1131–1143.
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. **Niegemann, E., A. Schulz, K. Bartsch, and A. G. Hoechst.** 1993. Molecular organization of the *Escherichia coli gab* cluster: nucleotide sequence of the structural genes *gabD* and *gabP* and expression of the GABA permease gene. Arch. Microbiol. **160:**454–460.
39. **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.
40. **Plamann, M. D., and G. V. Stauffer.** 1990. *Escherichia coli glyA* mRNA decay: the role of 3′ secondary structure and the effects of the *pnp* and *rnb* mutations. Mol. Gen. Genet. **220:**301–306.
41. **Prentki, P., M. Chandler, and D. J. Galas.** 1987. *Escherichia coli* integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. EMBO J. **6:**2479–2487.
42. **Pruss, G. J., and K. Drlica.** 1986. Topoisomerase I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. Proc. Natl. Acad. Sci. USA **83:**8952–8956.
43. **Rice, P. A., S. Yang, K. Mizuuchi, and H. A. Nash.** 1996. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell **87:**1295–1306.
44. **Roberge, M., and S. M. Gasser.** 1992. DNA loops: structural and functional properties of scaffold-attached regions. Mol. Microbiol. **6:**419–423.
45. **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.
46. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.

47. **Schuler, G. D., S. F. Altschul, and D. J. Lipman.** 1991. A workbench for multiple alignment construction and analysis. Proteins **9:**180–190.
48. **Schulz, V. P., and W. S. Reznikoff.** 1990. *In vitro* secondary structure analysis of mRNA from *lacZ* translation initiation mutants. J. Mol. Biol. **211:**427–445.
49. **Shaibe, E., E. Metzer, and Y. S. Halpern.** 1985. Metabolic pathway for the utilization of L-arginine, L-ornithine, agmatine, and putrescine as nitrogen sources in *Escherichia coli* K-12. J. Bacteriol. **163:**933–937.
50. **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.
51. **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.
52. **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.
53. **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.
54. **Yang, Y., and G. F.-L. Ames.** 1988. DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. Proc. Natl. Acad. Sci. USA **85:**8850–8854.
55. **Ziemke, P., and J. E. McCarthy.** 1992. The control of mRNA stability in *Escherichia coli*: manipulation of the degradation pathway of the polycistronic *atp* mRNA. Biochim. Biophys. Acta **1130:**297–306.
56. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.