JOURNAL OF BACTERIOLOGY, Apr. 1988, p. 1622–1630
0021-9193/88/041622-09$02.00/0

Vol. 170, No. 4

Transcriptional Organization of the *Escherichia coli* Hemolysin Genes

RODNEY A. WELCH* AND SHAHAIREEN PELLETT

Department of Medical Microbiology, University of Wisconsin Medical School, 1300 University Avenue, Madison, Wisconsin 53706

Received 6 October 1987/Accepted 19 January 1988

The transcriptional organization of the *Escherichia coli* hemolysin genes (*hlyCABD*) encoded by pSF4000 was examined. The use of different hemolysin gene-specific radiolabeled probes in blots containing isolated in vivo RNA revealed 4.0-kilobase *hlyCA* and 8.0-kilobase *hlyCABD* transcripts. The treatment of cells with rifampin just before RNA isolation showed the half-lives of these mRNAs to be 10.2 and 4.4 min, respectively. The 5′ ends of the *hly* transcripts were 462 and 464 nucleotides from the putative initiation codon of *hlyC* based on a primer extension method of RNA mapping. Deletion analysis of pSF4000 combined with quantification of the hemolysin structural protein HlyA by immunoblotting confirmed that major control of HlyA expression occurs within a 168-base-pair *Pst*I fragment located 433 base pairs upstream of the start of *hlyC*. A second recombinant plasmid, pANN202-312, encoding an *E. coli* hemolysin of different origin expressed 6-fold less total HlyA and 50-fold less extracellular HlyA than pSF4000 in identical cell backgrounds. The pANN202-312 recombinant had a different *hly* promoter, with the *hly* mRNA beginning 264 nucleotides upstream from the start of *hlyC*. We showed by RNA blotting that cells harboring pANN202-312 compared with pSF4000 have similar steady-state levels of the *hlyCA* transcript but they lack a consistently detectable *hlyCABD* transcript. We propose that one reason for the disparate levels of extracellular hemolysin produced by hemolytic *E. coli* is dissimilar levels of mRNA encoding in part the transport genes *hlyB* and *hlyD*.

The *Escherichia coli* hemolysin is a significant contributor to the virulence of recombinant *E. coli* strains in different animal models of extraintestinal *E. coli* disease (9, 31). Recent epidemiological evidence indicates that the *E. coli* hemolysin represents one branch of a gene family (MEPP-toxins) which is disseminated among other members of the *Enterobacteriaceae* as well as *Pasteurella haemolytica* (17, 18, 30). The MEPP-toxins are not only interesting to study from a biomedical viewpoint, but because the *E. coli* hemolysin is a secreted extracellular protein (5), it has potential biotechnological impact as a gene fusion system that would facilitate purification of valuable recombinant products.

Four genes (*hlyC*, *−A*, *−B*, and *−D*) are needed to produce the beta-hemolytic phenotype (6, 24, 29). HlyC activates the hemolysin structural gene product HlyA in an unknown way (23). The secretion of HlyA is unusual because it does not involve cleavage of an amino-terminal leader sequence (5). HlyB and HlyD act to transport HlyA across the cell envelope and release it from the cell surface (29). All four genes are encoded sequentially on the same DNA strand (6, 12). Candidate promoters for *E. coli* hemolysins are identified by inspection of the DNA sequences upstream of *hlyC* (6, 12, 14). Genetic evidence suggests that the genes are probably not part of one large operon but that *hlyCAB* compose one transcript, with the most distal gene, *hlyD*, transcribed separately (29). The analysis of the DNA sequence and the lack of apparent polarity effects on *hlyD* favor this view of the transcriptional organization of the hemolysin (6). However, the presence of a potential *rho*-independent terminator between *hlyA* and *hlyB* and a close match to the consensus sequence for *E. coli* promoters immediately upstream of *hlyB* leads to some confusion about the transcriptional organization (6).

In this report, we provide physical evidence that the previous operon models for the *E. coli* hemolysin are not entirely correct. There are two hemolysin-specific transcripts: one is an abundant *hlyCA* species and the other encodes all four genes, making it a single, large operon. In addition, we localize the 5′ ends of *hly* mRNA to the DNA sequences of two hemolysins of different origin. On the basis of those data, we provide evidence for the evolution of different promoter elements for the *E. coli* hemolysins.

MATERIALS AND METHODS

Bacterial strains and plasmid construction. The strain designations and origins of bacteria and recombinant plasmids used in this study are listed in Table 1. DNA fragment isolation and recombinant methodology used in the construction of new plasmids was performed as previously described (6, 33).

The original hemolysin recombinant plasmid, pSF4000, possesses on either side of the hemolysin determinant approximately 3 kilobases (kb) of DNA of unknown sequence and function (6, 33). For this and future studies, we constructed, in a three-stage process, a new hemolysin recombinant plasmid, pWAM04, in which the entire DNA sequence of the plasmid is known. In brief, the pSF4000 *Hin*dIII D fragment (base pairs [bp] 6979 to 8206 [6]) was inserted into the *Hin*dIII site of pUC19, resulting in pWAM572. This plasmid was then doubly digested with *Apa*I (a unique site in the *Hin*dIII insert and in pSF4000) and *Sal*I (a unique site in pWAM572 present in the pUC19 multiple cloning site [6]). The larger of the two resulting fragments was ligated to the 10-kb pSF4000 *Sal*I-*Apa*I fragment covering the *hlyCAB* region. This construction (pWAM575) was then digested with *Sal*I and *Sst*II (a unique site within pWAM575 present at bp 110 of the hemolysin sequence). The noncomplementary ends of the largest fragment were filled in with the Klenow fragment of DNA

* Corresponding author.

TABLE 1. Bacterial strains and recombinant plasmids

Strain	Relevant characteristics	Source or reference
DH1	K-12 strain, *recA*, nonhemolytic	D. Walker
JM83lac	K-12 strain, background for pUC19 cloning	M. Howe
J198	Human fecal isolate, nonhemolytic	31
J96	Human urinary tract isolate, source of pSF4000 hemolysin determinant	31
WAF260	J198(pANN202-312)	31
WAF270	J198(pSF4000)	31
WAM331	J198(pWAF222)	32
WAM356	JM83 *lac*(pWAM356) (pSF4000 *Sma*I B fragment[a] cloned into pUC9 in orientation A)	This study
WAM357	JM83 *lac*(pWAM357) (pSF4000 *Sma*I B fragment[a] cloned into pUC9 in the opposite orientation)	This study
WAM572	JM83 *lac*(pWAM 572) (pSF4000 *Hin*dIII D fragment cloned into pUC19)	This study
WAM575	JM83 *lac*(pWAM575) (pSF4000 *Sal*I-*Apa*I cloned into pWAM572)	This study
WAM582	DH1(pSF4000)	This study
WAM589	DH1(pWAM04)	This study
WAM738	DH1(pSF4000) Δ*Sal*I-*Hpa*I	6
WAM739	DH1(pSF4000) Δ*Sal*I-*Bst*EII	6
WAM742	DH1(pANN202-312)	W. Goebel
WAM750	DH1(pWAM04) Δ*Pst*I-H	This study
WAM751	DH1(pWAF222)	32
WAM752	DH1(pSF4000) Δ*Pst*I-D	32
WAM753	WAM742(pWAM356)	This study
WAM754	WAM742(pWAM357)	This study

[a] pSF4000 *Ava*I D and *Sma*I B fragments are analogous.

polymerase I, and the blunt ends of the fragment were ligated to yield pWAM04.

The construction of pWAM04 also enabled us to delete the small 168-bp *Pst*I fragment from the hemolysin promoter region (the *Pst*I H fragment of pSF4000, see Fig. 1). This was accomplished by isolating the two pWAM04 *Bam*HI fragments, digesting the smaller *Bam*HI fragment with *Pst*I, and then adding these digested fragments back to the large *Bam*HI fragment. The DNA mixture was ligated and transformed into strain DH1 with selection on ampicillin-sheep erythrocyte agar. Plasmid DNA from transformants was screened for the loss of the small 168-bp *Pst*I fragment, and a plasmid, pWAM750, harboring the appropriate pWAM04 *Pst*I deletion was isolated.

Media and reagents. Bacteria were grown in L broth (21), and when necessary to ensure maintenance of recombinant plasmids, the cultures were supplemented with either chloramphenicol (10 μg/ml) or ampicillin (100 μg/ml). Salts, buffer reagents, and antibiotics were purchased from either Sigma Chemical Co. (St. Louis, Mo.) or U.S. Biochemicals (Cleveland, Ohio).

RNA blotting. RNA was prepared by the method of Hagblom et al. (10) with the exception that cells were harvested from late-log-phase broth cultures (optical density at 600 nm of 0.9 to 1.0). The RNA was size fractionated by suspending 20-μg RNA samples in glyoxal and subjecting them to agarose gel electrophoresis by the method of Maniatis et al. (19). RNA size markers purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) were included in each gel. The separated RNA was transferred to nitrocellulose (Schleicher & Schuell, Inc., Keene, N.H.) by blotting (19). Two different types of radiolabeled probes were generated for detection of hemolysin-specific transcripts. Isolated restriction endonuclease fragments of pSF4000 were labeled with ^{32}P by nick translation (25). Strand-specific ^{32}P-labeled RNA probes were generated in vitro with SP6 and T7 RNA polymerases with *hly* gene-specific recombinants cloned into the Gemini vectors (Promega Biotec, Madison, Wis.) (20). Isolated pSF4000 restriction endonuclease fragments (*Pst*I-C and *Eco*RI-D) were cloned into pGEM1 and pGEM2, and radiolabeled transcripts were produced by the protocol provided by the manufacturer. Probes were hybridized to blots overnight in 40% formamide–1% sodium dodecyl sulfate–1 M NaCl–10% dextran sulfate at 42°C. The blots were washed in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1% sodium dodecyl sulfate at 60°C, dried, and subjected to autoradiography.

The stability of specific transcripts was studied by incubating the cells with rifampin (0.3 mg/ml) by the method of Belasco et al. (1) and then immediately lysing 10-ml culture samples at specific time intervals and preparing the RNA as described above. Comparisons of the intensity of autoradiographic signals were made on a Zeineh Soft Laser scanning densitometer, and integration of peak areas was determined with software provided by the densitometer manufacturer on a dedicated Apple IIE computer. The decay rate was calculated by a determination of the slope of −ln(band intensity) versus time (28).

Localization of 5′ termini of *hly* mRNA. Low-resolution mapping of the 5′ ends of *hly* mRNA species was accomplished by two different methods. Low-resolution S1 nuclease protection experiments were performed by the method of Burton et al. (4). Isolated pSF4000 restriction endonuclease fragments had their 5′ termini labeled with ^{32}P by using T4 polynucleotide kinase, and the labeled fragments were heat denatured and hybridized to in vivo RNA prepared from WAF270 or WAF260. After 3 h of hybridization at 51°C, the samples were treated with 16 U of S1 nuclease (Sigma) for 1.5 h at 45°C and then the treated samples were subjected to either 1% agarose or 10% acrylamide gel electrophoresis. The gels were subjected to autoradiography, and either 5′-labeled bacteriophage lambda *Hin*dIII or pBR322 *Hae*II fragments served as molecular size markers. The orientation of transcripts was determined by secondary restriction endonuclease digestion of the radiolabeled fragments at asymmetrical sites before hybridization to the RNA. A primer extension method of mapping 5′ termini of

mRNA based on the technique of Kassavetis and Geiduschek (15) was also used. Synthetic oligonucleotides (primer 1, ATACATGCCCAAGAACCTCT, reverse complement of bp 815 to 835 [6]; primer 2, ATCTTCTGTATTCAGCA GAA, reverse complement of bp 579 to 600 [6]; primer 3, AGACTCAAAAAACATCAAGC, reverse complement of bp 625 to 644 [12]) representing the reverse complements of different hemolysin mRNA regions were radiolabeled at their 5′ ends with [γ-^{32}P]ATP and T4 polynucleotide kinase. They were then hybridized to 3 μg of in vivo RNA preparations per μl. Hybridizations were performed in 0.05 M Tris hydrochloride (pH 8.3)–0.06 M NaCl–0.01 M dithiothreitol–0.001 M EDTA. The RNA-primer mixture was heated at 90°C for 3 min and then cooled to 45°C. Avian myeloblastosis virus reverse transcriptase (0.3 to 0.6 U/μl; Boehringer Mannheim Biochemicals, Indianapolis, Ind.) was then added in 0.08 M Tris hydrochloride (pH 8.3)–0.1 M NaCl–0.016 M dithiothreitol–0.01 M magnesium (acetate)$_2$ along with 0.66 mM of each deoxynucleoside triphosphate, and the mixture was incubated for 30 min at 45°C. Reactions were chased for 30 min by adding a mix of reverse transcriptase, deoxynucleoside triphosphates, magnesium (acetate)$_2$, and the hybridization buffer. The polymerization reaction was stopped by drying down samples, which were then suspended in 100% deionized formamide containing xylene cyanole and bromophenol blue. The samples were then subjected to 7 M urea–10% acrylamide gel electrophoresis. Initial low-resolution mapping required running the samples against the radiolabeled pBR322 *Hae*II size markers. Once low-resolution localization had been performed, the primer extension method was utilized for fine resolution of the 5′ ends by using a synthetic oligonucleotide primer that would be predicted to hybridize between 100 to 300 nucleotides from the apparent 5′ termini. The reactions were done as described above, but the size of the newly synthesized DNA strand was determined by comparing its migration with several lanes of a known DNA sequence synthesized by the Klenow fragment in an M13-based dideoxy sequencing reaction (26, 34).

Immunoblotting. The expression of HlyA was monitored by immunoblotting by our earlier methods (30). The relative expression of HlyA among different strains was quantified by using a scanning densitometer and comparing areas of HlyA optical density peaks in tracings of the immunoblots. To ensure appropriate comparisons, the material from each strain was subjected to serial dilutions and the dilutions were immunoblotted. The diluted samples were then scanned, and a linear response of peak area versus HlyA concentration was established.

RESULTS

The DNA sequence and genetic analysis of the *E. coli* hemolysin have not provided adequate evidence of its transcriptional organization nor the location of its promoter(s). Therefore, we performed experiments intended to physically characterize hemolysin-specific in vivo mRNAs. These experiments would identify the size of different hemolysin gene-specific mRNAs and the sequence identity of their 5′ end(s). We show in Fig. 1 that gene probes for *hlyCA* hybridized to two different size species of RNA. These were estimated to be 4.0 and 8.0 kb in length based on their comparative electrophoretic migration with RNA species of known size. The discrete length of the two mRNAs is not readily apparent in Fig. 1 but can be easily observed in Fig. 2. DNA gene probes for either *hlyB* (data not shown) or *hlyD*

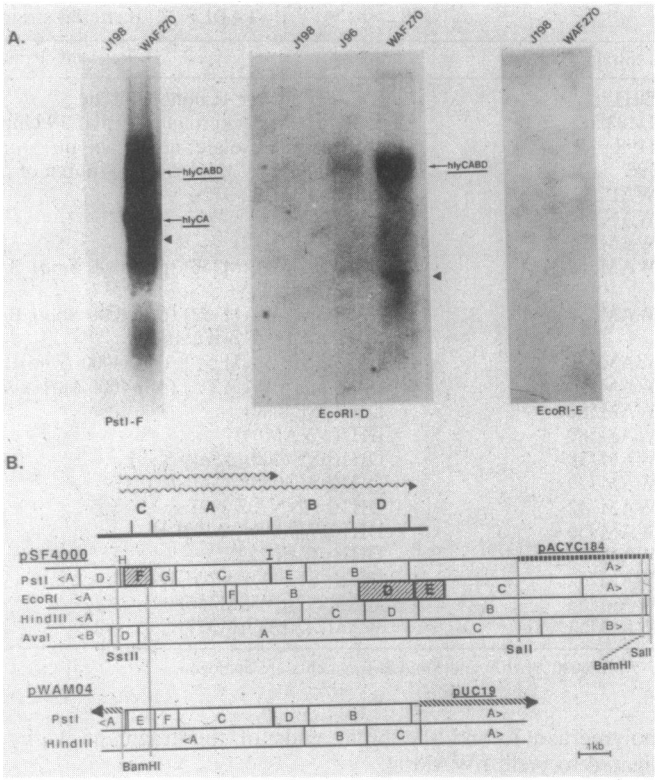

FIG. 1. RNA blotting analysis of the *E. coli* hemolysin. (A) Total cellular RNA (20 μg) from different *E. coli* strains was suspended in glyoxal and fractionated by 1% agarose gel electrophoresis. The gels were blotted to nitrocellulose, and the RNA was hybridized to different ^{32}P-radiolabeled probes (listed at the bottom of each blot). The strains examined are presented across the top [WAF270 is J198(pSF4000)]. The identity of the hemolysin transcripts is presented to the right of the blot pictures. The transcripts in the blot probed with the *Pst*I F fragment appear as a smear which probably represents intermediates in synthesis and degradation. A picture which better demonstrates their discrete length and size estimation can be seen in RNA turnover experiment presented in Fig. 2. The unlabeled arrowheads along the right edges of the gels indicate where the the smear yields artifactual bands around the regions of rRNA migration, a phenomenon observed by others in blots of procaryotic RNA (22). (B) Restriction endonuclease fragment maps of pSF4000 and pWAM04. The probes used in the RNA blots are indicated by a shading of the relevant fragment. A summary of the transcriptional organization of the hemolysin is indicated by the wavy lines above the hemolysin gene map.

(Fig. 1A) hybridized to only the 8.0-kb RNA species. An mRNA species similar in size to the pSF4000-encoded *hlyCABD* mRNA species was also apparent in cellular RNA isolated from J96 (Fig. 1A, lane 4), which is the original clinical isolate that was the source of the pSF4000 hemolysin. The 8.0-kb mRNA species was not detected in RNA blots probed with the pSF4000 *Eco*RI E DNA fragment (Fig. 1A). This DNA fragment is located downstream of *hlyD* and begins 91 bp from the *hlyD* stop codon (6). RNA blots probed with radiolabeled strand-specific RNA probes (pSF4000 *Pst*I-C and *Eco*RI-D cloned in pGEM1 and pGEM2) established that the encoding strand for the two RNA species is identical and coincident to that predicted by DNA sequence analysis.

Reports of relatively more stable mRNAs for secreted gene products (3, 13) led us to determine the in vivo half-life of the two hemolysin transcripts after rifampin treatment.

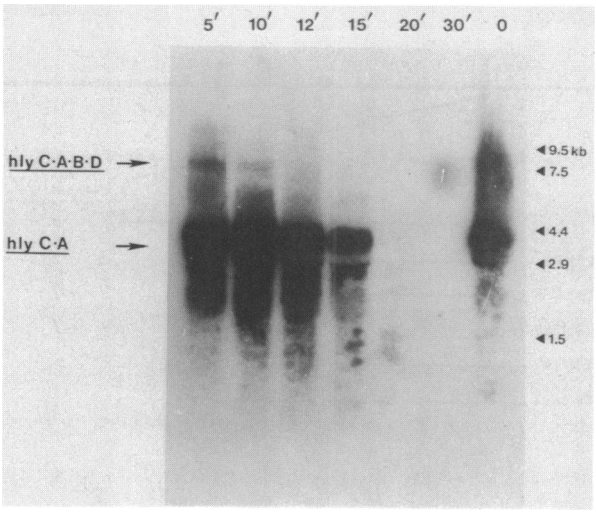

FIG. 2. Decay of full-length *hlyCA* and *hlyCABD* mRNA. Total cellular RNA isolated at various time intervals after rifampin treatment of WAF270 was suspended in glyoxal and fractionated by 1% agarose gel electrophoresis. The time intervals in minutes are indicated across the top of the gel. The gels were blotted to nitrocellulose, and the RNA was hybridized to a ^{32}P-radiolabeled *Pst*I F DNA fragment covering *hlyCA* as described in the legend to Fig. 1. The identities of the two transcripts are noted to the left of the figure. The migration of RNA molecular size markers is shown along the right side.

Densitometer tracings of autoradiographs which were similar but less exposed than that shown in Fig. 2 indicated that the peak area for the *hlyCA* species was reduced to one-half that of the steady-state level 10.2 min after the block in transcription initiation. Similar analysis of the *hlyCABD* species showed a half-life of 4.4 min.

We previously presented genetic evidence based on deletions of pSF4000 that the likely promoter controlling transcription of at least the *hlyCA* genes was upstream of a *Bst*EII site that is 301 bp upstream of the putative *hlyC* start codon (6). Preliminary low-resolution S1 nuclease protection experiments indicated that the 5′ end of mRNA encoded within the 937-bp *Sma*I fragment of pSF4000 was greater than 450 nucleotides upstream of the *hlyC* start (data not shown). In addition, low-resolution primer extension mapping with an amino-terminal HlyC complementary synthetic oligonucleotide (primer 1) showed that the 5′ end of the longest DNA strand synthesized was greater than 450 bp from the beginning of *hlyC* (Fig. 3). The precise localization of the 5′ end of the mRNA was performed by using a synthetic oligonucleotide representing the complement of the DNA sequence sense strand from nucleotides 580 to 600 (primer 2) (6). An example of those results is presented in Fig. 3B in which the 5′ end of the mRNA template results in DNA strands 266 and 268 bases in length, indicating that mRNA begins at nucleotides 332 and 334. This region would have been eliminated in the pSF4000 deletion encompassing both the *Pst*I F and H fragments (pWAF193) which we described previously as well as in the pWAM04 *Pst*I deletion derivative pWAM750 (see Fig. 4). The hemolytic phenotypes encoded by pWAF193 and pWAM750 were reduced

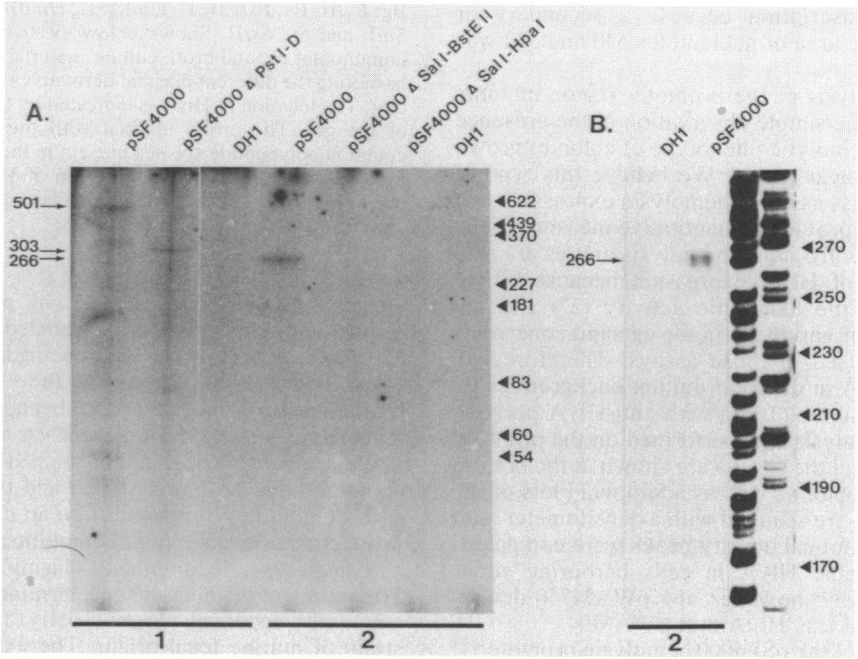

FIG. 3. Reverse transcriptase mapping analysis of the *E. coli hly* mRNA. Radiolabeled oligonucleotides (primers 1 and 2, described in the text) were hybridized to RNA preparations, and primer extension was performed with avian myeloblastosis virus reverse transcriptase. The products were subjected to electrophoresis in either 1% agarose (low resolution) or 10% acrylamide denaturing gels, and separated radiolabeled DNAs were identified by autoradiography. (A) Low-resolution mapping of the transcripts, using primers 1 and 2. The identities of the primer and RNA used in each reaction are listed across the bottom and top, respectively, of the gel. DH1 is the common host background. The longest and prominent DNA species are indicated to the left of the photograph by arrows labeled 501, 303, and 266. The migration of radiolabeled pBR322 *Hae*II fragments is indicated by arrowheads along the right-hand margin. (B) High-resolution mapping of the *hly* mRNA 5′ end, using primer 2. A portion of a known DNA sequence ladder is presented to the right and the base numbers of that sequence are listed along the right edge.

compared with those of the parental pSF4000- and pWAM04-containing clones. This suggested that either a positive controlling element or the primary promoter was altered by the *Pst*I-F-H or *Pst*I-H deletions. The strains harboring the pSF4000 deletion derivatives with endpoints closer to *hlyC*, the *Bst*EII or *Hpa*I sites (respectively, pWAM739 and pWAM738; see Fig. 4), had no detectable hemolytic activity on sheep erythrocyte agar plates (6). Therefore, we performed primer extension mapping of the 5′ ends of *hly* RNA encoded by the pSF4000 deletion derivatives. The 5′ end of mRNA encoded by pWAM04 was identical to that directed by pSF4000 (data not shown). We demonstrated that pWAF193 encodes in WAM582 an mRNA species which in primer extension analysis gives a DNA strand 303 bases long (Fig. 3). Fine-resolution analysis indicated that there are two 5′ ends for this mRNA, which would begin at nucleotides 530 and 532 (data not shown). We could not detect a signal when using RNA prepared from the WAM738 and WAM739 backgrounds (Fig. 3A, pSF4000 Δ *Sal*I-*Bst*EII and pSF4000 Δ *Sal*I-*Hpa*I). In addition, no RNA species were detected in RNA blots with an *hlyCA* DNA probe hybridized to cellular RNA from either of these two strains.

We present in Fig. 4 a physical map covering the hemolysin promoter region. This figure is a significant revision of a previously published figure (6). We now include more physical details and new phenotypic data from additional deletions of the region. The pWAM04 recombinant demonstrates that expression of the hemolytic phenotype is not controlled by sequences upstream of the *Sst*II site at nucleotide 110. Although the apparent primary transcriptional start site lies between the two *Pst*I sites at nucleotides 195 and 363, a deletion of that 168-bp fragment (pWAM750) did not eliminate *hly* transcription because a secondary or cryptic promoter in the area of nucleotides 530 and 532 was utilized.

The phenotypic analysis of the promoter region mutants presented in Fig. 4 is the simple observation of the presence or absence of a beta-hemolytic phenotype of colonies grown on sheep erythrocyte agar plates. We believe this type of analysis provides a gross assay of hemolysin expression and secretion but does not provide a quantitative measure of the synthesis of HlyA. In vitro liquid hemolysis assays are also inadequate measures of HlyA expression because of the marked instability of the hemolytic activity (27) and the unavoidable problem of variations in the age and concentration of erythrocytes used in those assays. Therefore, we chose to quantify HlyA in different mutant backgrounds by Western blotting (immunoblotting) with anti-HlyA polyclonal antiserum. This analysis was performed on the pSF4000 deletion derivatives, and the results are shown at the bottom of Fig. 4. This immunoblot as well as additional blots of the total culture material were scanned with a densitometer, and the areas of the HlyA optical density peaks were compared. There was no detectable HlyA in cells harboring either pWAM738 or pWAM739; however, the pWAM750 derivative produced eightfold less HlyA that pWAM04.

The DNA sequence of the pSF4000 hemolysin promoter(s) is shown in Fig. 5A. The promoter assignment is based on the mapping of the 5′ ends of *hly* mRNA as well as the deletion data presented in Fig. 3 and 4. There is some ambiguity to the assignment of the −10 and −35 regions because the primer extension method of mapping has reproducibly yielded two DNA molecules where the mRNA would begin at nucleotides 332 and 334. We chose to present the promoters as a pair of −10 and −35 regions which have

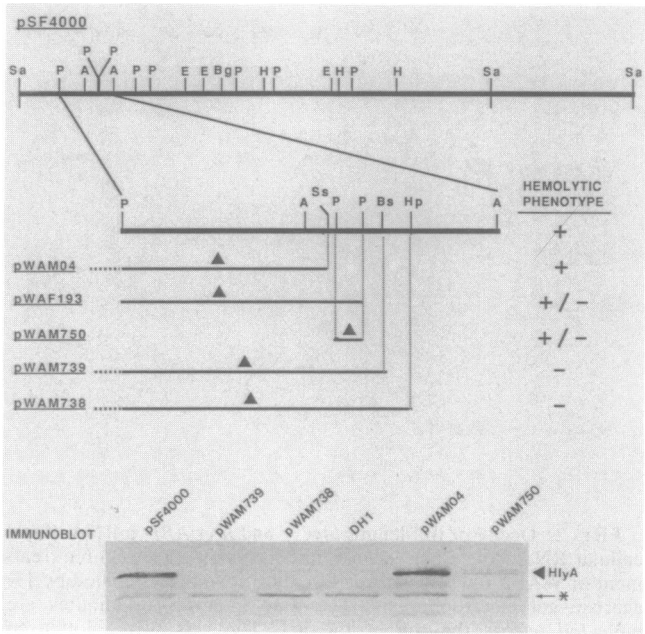

FIG. 4. Deletion map and phenotype analysis of the *E. coli* hemolysin promoter region. At the top of the figure is a restriction endonuclease fragment map of pSF4000. The region encompassing the pSF4000 *Pst*I D and *Ava*I D fragments is expanded in the central part of the figure. Below the map are the designations, deletion spans (▲), and hemolytic phenotypes for different in vitro constructs described in the text. The + and +/− phenotype signs indicate that colonies have either an identical beta-hemolytic zone or a reduced zone size compared with that of the colonies harboring pSF4000. The abbreviations of restriction endonuclease sites are: A, *Ava*I; Bg, *Bgl*II; Bs, *Bst*EII; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; P, *Pst*I; Sa, *Sal*I; and Ss, *Sst*II. Shown below the maps is a photograph of an immunoblot of total broth culture material from *E. coli* DH1 strains harboring the different plasmid derivatives listed along the top of the blot. The location of HlyA is indicated by the arrowhead to the right of the gel. The arrow marked with the asterisk (*) indicates a common polypeptide species present in the DH1 background that is used to normalize the quantification of HlyA. Additional adsorptions of the antiserum with DH1 cellular material lead to the reduced appearance of this polypeptide.

elements conforming to the *E. coli* promoter consensus of sequences (11). The lower bar labeled −10 best matches the −10 consensus; however, its spacing to the mRNA 5′ ends is 4 and 6 bases away. Therefore, the −10 sequence indicated by the upper bar is also shown because it would be exactly 7 nucleotides away from the mRNA start at nucleotide 332.

We previously showed that the pSF4000 *Ava*I D fragment containing the promoter region and the 5′-terminal third of *hlyC* confers in *cis* configuration an up expression effect on production of extracellular hemolytic activity by a second *E. coli* hemolysin recombinant plasmid, pANN202-312 (32). The source of the hemolysin determinant in this instance was a naturally occurring plasmid, pHly152, present in an *E. coli* strain of murine fecal origin. The expression effect occurs when the fragment is inserted in frame at the common *Ava*I site of pANN202-312, yielding a plasmid called pWAF222 (32). This led to the conjecture that *E. coli* hemolysins differ in levels of hemolysin production because of differences in the strength of promoters (32). Preliminary data involving the comparison of DNA sequences upstream of *hlyC* for pSF4000 and pANN202-312 indicate 30% homology in this region compared with 96% homology of the *hlyC* genes (S.

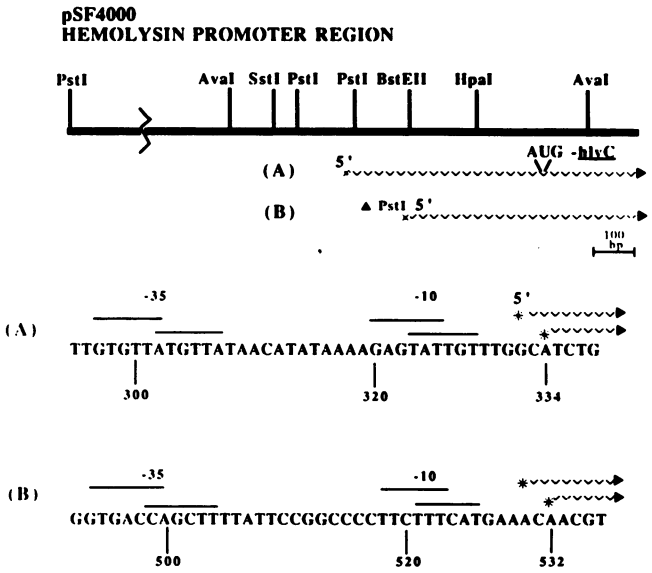

FIG. 5. Promoter region of pSF4000-encoded hemolysin. Shown at the top of the figure is a physical map of the promoter region for *hly* mRNA. Below the restriction endonuclease fragment map are the DNA sequences for the pSF4000 hemolysin promoter (A) and the secondary promoter (B) present in the pSF4000 *PstI* deletion derivatives (pWAM752; Fig. 3 and 4). The bars labeled −35 and −10 above the DNA sequence refer to the location of the promoters. The native 5′ ends of the *hly* mRNA occur at nucleotides 336 and 334 (indicated by the asterisks). The 5′ ends of the *hly* mRNA encoded by pWAM752 are at nucleotides 534 and 532 and are marked by asterisks in the lower DNA sequence (B).

Pellett, T. Karakash, L. Segal, and R. Welch, Program Abstr. 24th Intersci. Conf. Antimicrob. Agents Chemother., abstr. no. 199, 1984). The low DNA sequence homology in the area upstream of *hlyC* supported the contention that there were probably differences in the strength of hemolysin promoters. However, despite the general pattern of DNA sequence divergence, there are small regions of sequence similarity which resemble the −10 and −35 consensus sequences for *E. coli* promoters. Hess et al. (12) proposed that this is evidence of common promoters for the *E. coli* hemolysins. To address this issue, we compared the in vivo synthesis of *hlyCA* and *hlyCABD* mRNAs by cells harboring pSF4000, pANN202-312, and pWAF222. Shown to the left in Fig. 6 are RNA blots containing WAF260 RNA from two independent preparations. The first result to be noted is the lack of a con-sistently detectable *hlyCABD* transcript in the pANN202-312-containing cells. Second, by densitometry of less exposed autoradiographs, the intensity of the WAF260 *hlyCA* signal was only two- to fourfold less than that of WAF270. The striking difference in the amount of stability of *hlyCABD* is most evident in the RNA blots shown in Fig. 6B. At the 5-min point of post rifampin treatment, *hlyCABD* was still readily apparent in WAF270 and WAM331 backgrounds, but no signal was observable for WAF260.

Juarez et al. (14) determined the DNA sequence of the pANN202-312 *hlyC* gene and a portion of its upstream region. It was reported that the likely promoter was 15 nucleotides upstream of the *hlyC* start codon based on the analysis of promoter region subclones in a *gal*K transcription fusion vector. As mentioned above, Hess et al. (12) presented a comparative DNA sequence analysis of the pANN202-312 promoter region with several other *E. coli* hemolysins of chromosomal origin. A region of limited DNA

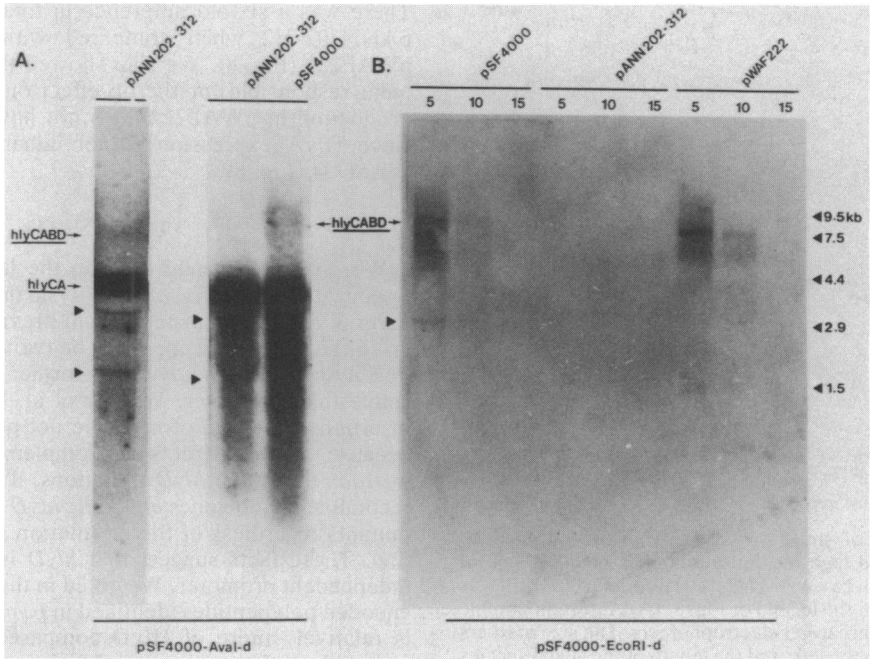

FIG. 6. RNA blotting of pANN202-312-, pSF4000-, and pWAF222-encoded hemolysins. Total cellular RNA from the J198 background harboring the plasmids listed along the top of the gels was fractionated, blotted, and hybridized to the probes noted at the bottom of the gels. The hemolysin transcripts are designated by labeled arrows. (A) The first and second lanes contain independent preparations of WAF260 RNA. (B) RNA turnover experiment in which the time of RNA harvest post rifampin treatment (5, 10, or 15 min) is listed at the top of each lane. The migration of RNA size reference standards is shown by labeled arrowheads along the right margin. The unlabeled arrowheads indicate 23S and 16S RNA-induced banding artifacts (Fig. 1).

homology exists among the different hemolysins which is, respectively, 206 and 216 nucleotides upstream of the pANN202-312 and pSF4000 start of *hlyC*. It was proposed that these similarities are evidence of a common promoter for *E. coli* hemolysins (12). We performed experiments to localize the in vivo 5′ ends of mRNA encoded by pANN202-312. Low-resolution S1 nuclease protection (data not shown) and primer extension experiments indicated that the *hly* mRNA had a 5′ end which was greater than 250 nucleotides from the *hlyC* start codon (Fig. 7A). We determined the precise location of the 5′ end (Fig. 7B) to be at nucleotides 530 and 532 in the sequence presented by Hess and coworkers (12). This shows that the pSF4000 and pANN202-312 hemolysin determinants utilize different promoter sequences. The DNA sequence of the pANN202-312 *hly* promoter region is presented at the bottom of Fig. 7.

We previously demonstrated that cells harboring either pSF4000 or pWAF222 secrete at least 50-fold more HlyA into the culture medium than cells with pANN202-312 (5). That observation was based on the detection of Coomassie- or silver-stained HlyA in electrophoresis gels containing precipitates of late-log-phase culture supernatants. We were

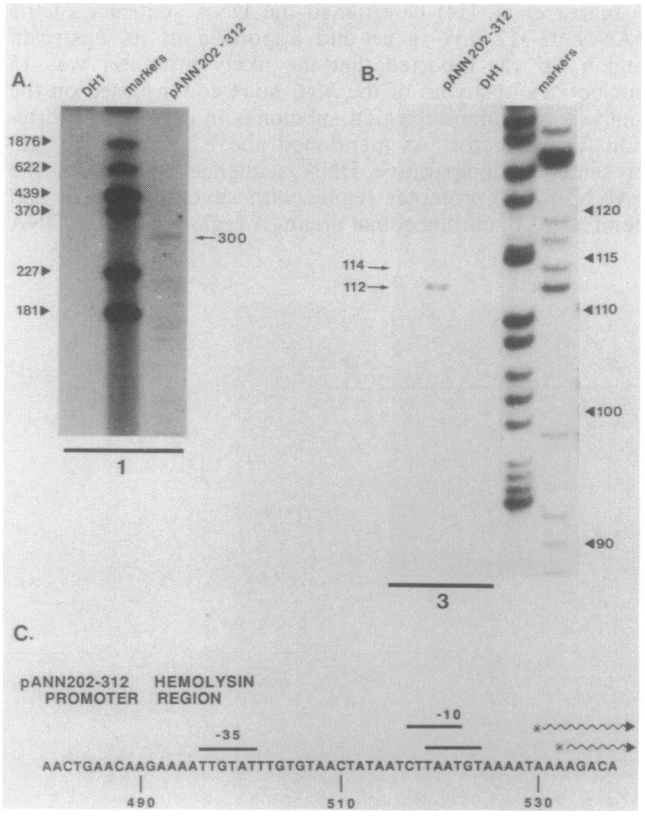

FIG. 7. Localization of the 5′ end of pANN202-312-encoded *hlyCA* mRNA. Low- and high-resolution mapping of the 5′ end of *hlyCA* mRNA produced by WAM742 [DH1(pANN202-312)] was performed as described in the legend of Fig. 3. (A) Autoradiographic signal resolved by 1% agarose gel electrophoresis. The size markers (nucleotides) are listed to the left, and the length of the signal species is shown to the right. The precise mapping of the 5′ end is shown in panel B. (C) DNA sequence of the promoter region for the pANN202-312-encoded *hlyCA*. The bars indicate the likely −35 and −10 regions of the promoter based on the mRNA 5′ ends indicated by the asterisks. The numbering of the DNA sequence is that of Hess et al. (12).

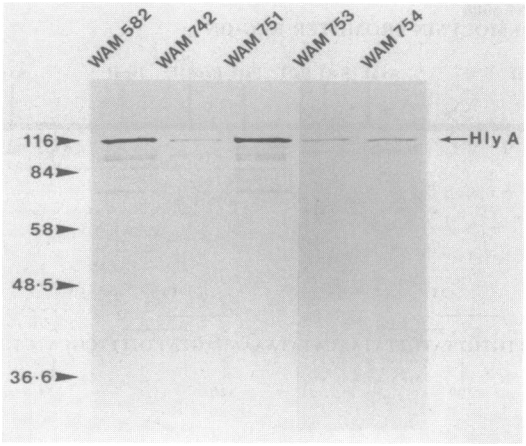

FIG. 8. HlyA immunoblotting. Late-log-phase culture material (50 μl) was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis and then immunoblotted as previously described (30). The different hemolytic strains used are listed across the top, and the presence of the 110-kilodalton HlyA polypeptide is indicated by the arrow to the right. DH1 is the common host background. The strain designations and plasmid(s) involved are: WAM582, pSF4000; WAM742, pANN202-312; WAM751, pWAF 222; WAM753, pANN202-312 and pWAM356; WAM754, pANN 202-312 and pWAM357. The migration of reference markers (kilodaltons) is shown to the left of the photograph.

curious to find out whether our observed differences in the *hly* transcripts coincided with either a reduced amount of total HlyA expressed by pANN202-312- compared with pSF4000-containing cells or an inability to secrete HlyA as efficiently. This was addressed by simply comparing the total amount of HlyA detectable in total broth cultures of pANN202-312-, pSF4000-, and pWAF222-containing cells. There was a sixfold difference in total HlyA expressed by pANN202-312 when compared with either pSF4000 or pWAF222 (Fig. 8). We also showed that the pSF4000 fragment responsible for the up effect on pANN202-312 HlyA production in pWAF222 does not have this effect in *trans* since HlyA expression is not enhanced in WAM753 or WAM754 (Fig. 8).

DISCUSSION

We present physical data on the in vivo transcriptional organization of the *E. coli* hemolysin operon, which differs in several significant respects from previously proposed models (6, 12, 29). There appear to be two transcripts composing *hlyCA* and *hlyCABD*, with the former transcript being more stable than the latter. Wagner et al. (29) showed that Tn*5* insertions upstream of *hlyD* are not polar on its expression because these constructs can complement hemolytic activity in *trans* to known *hlyD* mutations. They also showed that recombinant subclones encoding *hlyD* can complement *hlyD* mutants regardless of the orientation of the fragment insert (29). These facts suggest that *hlyD* is under control of an independent promoter. We noted in the analysis of pSF4000-encoded polypeptides identified in *E. coli* minicells that there is relatively more of HlyD compared with the upstream-encoded HlyB species (6). This appeared to be evidence supporting the absence of transcriptional polarity and strengthened the contention that *hlyD* is transcribed independently. However, the prediction of the *hly* transcriptional organization was equivocal because we showed that *hlyB* recombinant subclones produce HlyB independent of the

orientation of the DNA fragment insert (6). Additionally, the identification of a *rho*-independent terminatorlike RNA sequence between *hlyA* and *hlyB* (6) suggests that *hlyB* is not encoded on the same mRNA species as *hlyA*. Therefore, the in vivo evidence of *hlyCA* and *hlyCABD* transcripts presented here must be reconciled with the earlier genetic and physical data.

Although Tn*5* is strongly polar on downstream genes, Berg et al. (2) estimate that one-third of all Tn*5* insertions in *lacZ* or *lacY* express distal genes because of a Tn*5*-associated promoter. The complementation observed by Wagner et al. (29) of *hlyD* mutations by Tn*5* insertions did not involve quantification of any gene product, only the presence or absence of cell-free hemolytic activity. Therefore, it is unclear whether Tn*5* was causing a polar effect or whether some degree of Tn*5*-directed transcription was permitting expression of *hlyD*. The expression of *hlyC* and *hlyA* alone without *hlyB* and *hlyD* in a recombinant subclone is lethal (7). Therefore, it may also be feasible that the Tn*5* insertions in the 3'-terminal region of *hlyA* or *hlyB* which are polar on *hlyB* and *hlyD* are not available for analysis. We believe that these possibilities in conjunction with the physical evidence of an *hlyCABD*-specific transcript sufficiently weaken the earlier genetic arguments for a unique *hlyD* transcript. We recognize, however, that the *hly* transcription pattern for pANN202-312 may not be similar to that of the pSF4000 hemolysin. We could only occasionally detect the *hlyCABD* RNA and could not detect an mRNA species specific for either *hlyB* or *hlyD* alone from pANN202-312-containing cells. Nonetheless, it is apparent that the *hlyCA*, *hlyCABD* operon structure is operative in the original J96 chromosomal hemolysin determinant. The higher-level production of pSF4000-encoded HlyD compared with that of HlyB in *E. coli* minicells (6) is speculated to be controlled at a posttranscriptional stage. For example, it is known that the putative ribosome-binding site for *hlyB* is within several bases of a large stem-loop structure (6) which is predicted to occur in the *hlyCABD* mRNA. This may hinder HlyB expression by inhibition of ribosome readthrough from *hlyA* or de novo binding of ribosomes at the start of *hlyB*.

The hemolysin transcripts are members of an unusual class of stable procaryotic mRNAs with half-lives of greater than 2 to 3 min. Both et al. (3) and Hirashima et al. (13) observed that the transcripts encoding extracellular proteins and outer membrane proteins are more stable than those of cytoplasmic proteins. It was proposed the mRNAs of secreted gene products survive longer within the cell because they are protected and translated at membrane-bound ribosomes (8). Belasco and co-workers (1) demonstrated that the enhanced stability of *ompA* mRNA is a 5'-specific feature.

A second unusual feature of the hemolysin transcripts is the long 5' leader sequence. The 462-nucleotide distance between the 5' end and the putative *hlyC* start codon for the pSF4000 *hly* mRNA is the longest RNA leader of procaryotes described to date (16). The previous longest leader sequence was 170 bases for the *crp* mRNA (16). The −10 and −35 promoter regions corresponding to the hemolysin mRNA 5' end do not have an unusual sequence composition when compared with other known *E. coli* promoters (11). The two different *Pst*I deletions at the 5' end of the leader region (WAM750 and WAM752) establish that despite the existence of several candidate promoters closer to the *hlyC* translational start (6), the expression of HlyA is normally dependent on the presence of the promoter contained in the 168-bp *Pst*I fragment. We do not have any evidence that the alternative mRNA 5' end observed in the WAM750 background is present in the native *hly* transcripts. We acknowledge that it remains to be established that the two *hly* transcripts have identical 5' termini. The alternative possibility exists that the two mRNA species do not represent the products of independent transcription events but that, as suggested by the work from several laboratories, the *hlyCA* is a processed form of the longer *hlyCABD* molecule (4, 22). Given the relative abundance and stability of the *hlyCA* species compared with *hlyCABD*, it is our conclusion that 5' ends at nucleotides 462 and 460 most likely represent the beginning of the *hlyCA* mRNA.

The lack of an *hlyCABD* signal in the RNA blots with the pSF4000 *Eco*RI-F probe indicate that the *hlyCABD* 3' end must occur between nucleotides 8037 (the *hlyD* termination codon) and 8130 (the left-hand end of the *Eco*RI fragment). We propose that the likely candidate for the *hlyCABD* transcription terminator is a *rho*-independent-like potential stem-loop structure beginning at nucleotide 8047 and ending with an unpaired sequence of uridylate residues at nucleotide 8095.

We also addressed what may be controlling factors responsible for the wide differences in levels of HlyA secreted by different hemolytic *E. coli* isolates. Our results involving the detection of *hly* transcripts produced by different recombinants suggest that there is a clear difference in the steady-state level of *hlyCABD*. The insertion of the 5' encoding region from pSF4000 into pANN202-312 to give pWAF222 results in the increase in the *hlyCABD* mRNA produced. The observation that the pSF4000 *Sma*I B fragment present in pWAM356 or pWAM357 does not increase HlyA expression in *trans* to pANN202–312 suggests that this region does not encode a diffusible positive effector. We are now examining the possibility that the *Pst*I H region provides increased mRNA stability or a stronger promoter for the *hlyCABD* operon. It is likely that one or both of these possibilities could account for observations involving pWAF222.

ACKNOWLEDGMENTS

We thank Per Hagblom and Maggie So for their protocols on RNA isolation and primer extension mapping. We thank F. de la Cruz, F. Dailey, and P. Goode for their constructive comments.

The oligonucleotides used in this study were synthesized by the Protein/DNA Sequence/Synthesis Facility of the University of Wisconsin Biotechnology Center, which is supported by instrumentation grants from the National Institutes of Health. This work was supported by Public Health Service grant AI-20323 from the National Institutes of Health. Rodney A. Welch is a Pew Scholar in the Biomedical Sciences.

LITERATURE CITED

1. **Belasco, J., G. Nilsson, A. vonGabian, and S. Cohen.** 1986. The stability of *E. coli* gene transcripts is dependent on determinants localized to specific mRNA segments. Cell **46**:245–251.
2. **Berg, D., A. Weiss, and L. Crossland.** 1980. Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **142**:439–446.
3. **Both, G., J. McInnes, J. Hanlon, B. May, and W. Elliot.** 1972. Evidence for an accumulation of messenger RNA specific for extracellular proteases and its relevance to the mechanism of enzyme secretion in bacteria. J. Mol. Biol. **67**:199–217.
4. **Burton, Z., C. Gross, K. Watanabe, and R. Burgess.** 1983. The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in *E. coli* K12. Cell **32**:335–349.
5. **Felmlee, T., S. Pellett, E. Y. Lee, and R. A. Welch.** 1985. *Escherichia coli* hemolysin is released extracellularly without cleavage of a signal peptide. J. Bacteriol. **163**:88–93.
6. **Felmlee, T., S. Pellett, and R. A. Welch.** 1985. Nucleotide

sequence of an *Escherichia coli* chromosomal hemolysin. J. Bacteriol. **163**:94–105.
7. **Goebel, W., and J. Hedgpeth.** 1982. Cloning and characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*. J. Bacteriol. **151**:1290–1298.
8. **Green, P., and M. Inouye.** 1984. Roles of the 5′ leader region of the *ompA* mRNA. J. Mol. Biol. **176**:431–442.
9. **Hacker, J., C. Hughes, H. Hof, and W. Goebel.** 1983. Cloned hemolysin genes from *Escherichia coli* that cause urinary tract infection determine different levels of toxicity in mice. Infect. Immun. **42**:57–63.
10. **Hagblom, P., E. Segal, E. Billyard, and M. So.** 1985. Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. Nature (London) **315**:156–158.
11. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
12. **Hess, J., W. Wels, M. Vogel, and W. Goebel.** 1986. Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence. FEMS Microbiol. Lett. **34**:1–11.
13. **Hirashima, A., G. Childs, and M. Inouye.** 1973. Differential inhibitory effects of antibiotics on the biosynthesis of envelope proteins of *Escherichia coli*. J. Mol. Biol. **79**:373–389.
14. **Juarez, A., C. Hughes, M. Vogel, and W. Goebel.** 1984. Expression and regulation of the plasmid-encoded hemolysin determinant of *Escherichia coli*. Mol. Gen. Genet. **197**:196–203.
15. **Kassavetis, G., and E. P. Geiduschek.** 1982. Bacteriophage T4 late promoters: mapping 5′ ends of T4 gene 23 mRNAs. EMBO J. **1**:107–114.
16. **Kennell, D.** 1986. The instability of messenger RNA in bacteria, p. 101. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworths, Stoneham, Mass.
17. **Koranakis, V., M. Cross, B. Senior, E. Koranakis, and C. Hughes.** 1987. The secreted hemolysins of *Proteus mirabilis*, *Proteus vulgaris*, and *Morganella morganii* are genetically related to each other and to the alpha-hemolysin of *Escherichia coli*. J. Bacteriol. **169**:1509–1515.
18. **Lo, R. Y. C., C. Strathdee, and P. Shewen.** 1987. Nucleotide sequence of the leucotoxin genes of *Pasteurella haemolytica* A1. Infect. Immun. **55**:1987–1996.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Melton, D., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
21. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **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.
23. **Nicaud, J. M., N. Mackman, L. Gray, and I. B. Holland.** 1985. Characterization of HlyC and mechanism of activation and secretion of hemolysin from *E. coli* 2001. FEBS Lett. **187**:339–344.
24. **Noegel, A., U. Rdest, W. Springer, and W. Goebel.** 1979. Plasmid cistrons controlling synthesis and secretion of the exotoxin α-hemolysin of *Escherichia coli*. Mol. Gen. Genet. **175**:343–350.
25. **Rigby, P., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
27. **Smith, H. W.** 1963. The haemolysins of *Escherichia coli*. J. Pathol. Bacteriol. **85**:197–211.
28. **vonGabian, A., J. Belasco, J. Schottel, A. C. Y. Chang, and S. Cohen.** 1983. Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. Proc. Natl. Acad. Sci. USA **80**:653–657.
29. **Wagner, W., M. Vogel, and W. Goebel.** 1983. Transport of hemolysin across the outer membrane of *Escherichia coli* requires two functions. J. Bacteriol. **154**:200–210.
30. **Welch, R. A.** 1987. Identification of two different hemolysin determinants in uropathogenic *Proteus* isolates. Infect. Immun. **55**:2183–2190.
31. **Welch, R. A., E. P. Dellinger, B. Minshew, and S. Falkow.** 1981. Haemolysin contributes to virulence of extraintestinal *Escherichia coli* infections. Nature (London) **294**:665–667.
32. **Welch, R. A., and S. Falkow.** 1984. Characterization of *Escherichia coli* hemolysin conferring quantitative differences in virulence. Infect. Immun. **43**:156–160.
33. **Welch, R. A., R. Hull, and S. Falkow.** 1983. Molecular cloning and physical characterization of a chromosomal hemolysin from *Escherichia coli*. Infect. Immun. **42**:178–186.
34. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and puC19 vectors. Gene **33**:103–119.