JOURNAL OF BACTERIOLOGY, Apr. 1999, p. 2261–2266 Vol. 181, No. 7
0021-9193/99/$04.00+0

Mutational Analysis of the *vacA* Promoter Provides Insight into Gene Transcription in *Helicobacter pylori*

MARK H. FORSYTH[1] AND TIMOTHY L. COVER[1,2]*

Departments of Medicine and Microbiology and Immunology, Vanderbilt University School of Medicine,[1] and Veterans Affairs Medical Center,[2] Nashville, Tennessee

Received 29 September 1998/Accepted 20 January 1999

Analysis of 12 *Helicobacter pylori* promoters indicates the existence of a consensus −10 hexamer (TAtaaT) but little conservation of −35 sequences. In this study, mutations in either the *H. pylori vacA* −10 region or the −35 region resulted in decreased *vacA* transcription and suggested that an extended −10 motif is utilized. Thus, despite the lack of a −35 consensus sequence for *H. pylori* promoters, the −35 region plays a functional role in *vacA* transcription.

Helicobacter pylori bacteria are curved gram-negative organisms that persistently colonize the gastric epithelium of humans and other primates. Gastric mucosal inflammation occurs in all *H. pylori*-infected persons and usually does not cause any symptoms. However, *H. pylori* infection is a risk factor for the development of peptic ulcer disease and gastric adenocarcinoma (7).

At present, very little is known about the basic processes of gene transcription and transcriptional regulation in *H. pylori*. Our current knowledge of gene transcription in prokaryotes is based primarily on extensive studies that have been done with *Escherichia coli* and related organisms (5, 14, 15). The DNA-dependent RNA polymerase holoenzyme of *E. coli* is composed of a core enzyme with an $\alpha_2\beta\beta'$ structure, along with one of several possible σ subunits. The core enzyme is capable of RNA synthesis and nonspecific DNA binding, and the individual σ factors mediate specific binding of the holoenzyme to promoter elements (for reviews, see references 5, 14, and 15). The general σ factor used for the transcription of most genes in *E. coli* is σ^{70} (RpoD) (12), which binds to two hexameric DNA motifs centered approximately 10 and 35 bp upstream from transcriptional start points (TSP). Consensus DNA sequences for both of these hexamers (TATAAT and TTGACA) have been determined in *E. coli*, and many other genera of prokaryotes utilize similar sequences to facilitate binding of RNA polymerase (10, 11, 18, 27).

Analysis of consensus promoter sequences and RpoD in *H. pylori*. To identify consensus promoter sequences in *H. pylori*, we analyzed 11 different genes for which primer extension data were available (Fig. 1). A consensus −10 hexamer (TAtaaT), which closely resembles the −10 consensus sequence in *E. coli* (TATAAT), was identified. However, we were unable to identify an obvious −35 consensus sequence for this set of *H. pylori* genes.

The absence of a consensus *H. pylori* −35 sequence led us to speculate that there may be structural differences between *H. pylori* RpoD (22) and orthologous proteins in other bacterial species. To investigate this possibility, we aligned the DNA binding domains of RpoD from *H. pylori* 26695 (26) with the corresponding domains of RpoD from 10 other bacterial genera. Alignment of the various 2.4 domains, responsible for binding to −10 hexamers, indicates that the *H. pylori* RpoD sequence differs from the consensus sequence at four positions (Fig. 2A). More striking is the degree to which divergence has occurred in the 4.2 domain of *H. pylori* RpoD, which typically mediates binding to −35 promoter elements (Fig. 2B). The high degree of degeneracy in domain 4.2 of *H. pylori* RpoD is consistent with the presence of −35 promoter elements in *H. pylori* that are quite different from those found in *E. coli*.

Mutational analysis of the *H. pylori vacA* promoter. To experimentally investigate the promoter sites required for binding of *H. pylori* RNA polymerase, we selected *vacA* (encoding a vacuolating cytotoxin) (3) as a model and introduced a series

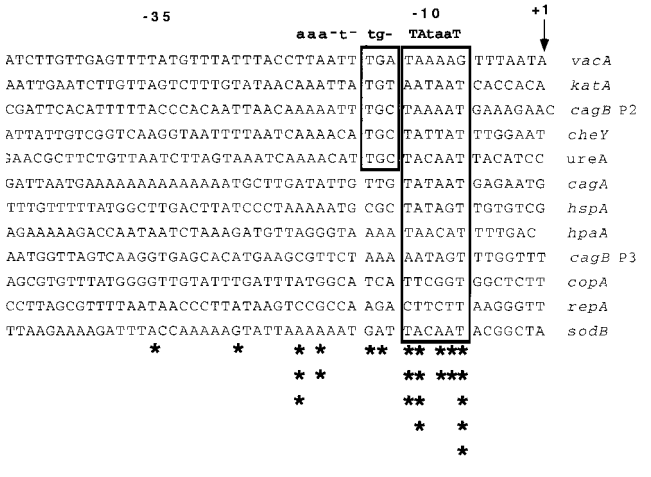

FIG. 1. Alignment of putative *H. pylori* promoter sequences. Putative promoter sequences deduced from primer extension analyses of 11 different *H. pylori* genes were aligned based on their TSP, designated as +1. Sequences similar to the consensus *E. coli* −10 hexamer (TATAAT) are boxed. The consensus sequence above the alignment shows highly conserved positions (capital letters) and weakly conserved positions (lowercase letters). In a subset of these promoter sequences, there is a −15/−13 TGN motif (indicated with a box) that is similar to a corresponding motif in *E. coli* extended −10 hexamers (1, 17). Asterisks denote the degree of conservation at each position among these promoters, as follows: *****, 11 of 12; ****, 10 of 12; ***, 9 of 12; **, 8 of 12; and *, 7 of 12. Genes analyzed are *vacA* (vacuolating toxin), *katA* (catalase), *cagA* and *cagB* (genes located in the *cag* pathogenicity island), *cheY* (chemotaxis response regulator), *ureA* (urease), *hspA* (GroES heat shock protein), *hpaA* (flagellar sheath protein), *copA* (copper-transporting ATPase), *repA* (plasmid replication), and *sodB* (superoxide dismutase) (2, 8, 9, 13, 16, 19, 20, 23, 24, 25).

* Corresponding author. Mailing address: Division of Infectious Diseases, A3310 Medical Center North, Vanderbilt University School of Medicine, Nashville TN 37232-2605. Phone: (615) 322-2035. Fax: (615) 343-6160.

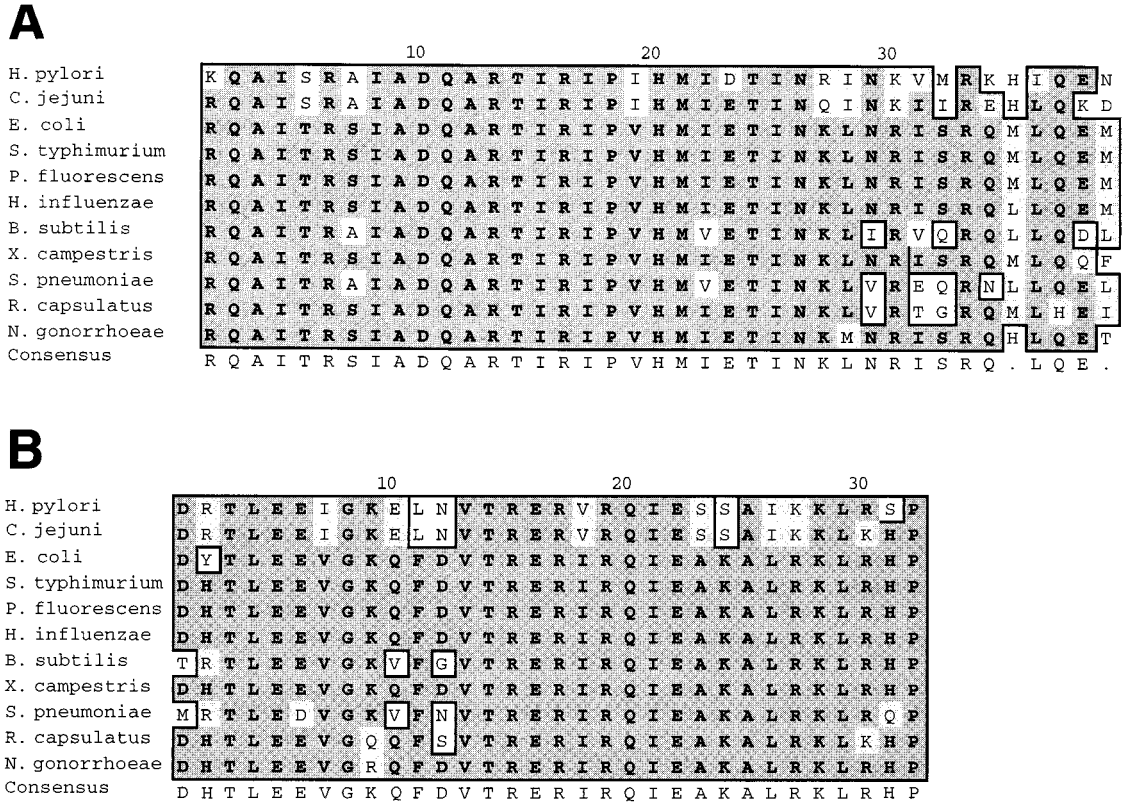

FIG. 2. Alignment of −10 and −35 binding domains of bacterial RpoD proteins. Domains of RpoD responsible for binding −10 and −35 promoter elements (domains 2.4 and 4.2, shown in panels A and B, respectively) were aligned with the ClustalW algorithm. Domain 2.5, which mediates binding to extended −10 sequences (1), is found at positions 22 to 39 in panel A. Conservative substitutions are shown in light gray boxes, and nonconservative amino acid substitutions are in white boxes. A consensus sequence for each domain is shown below each alignment. The Swiss-Prot accession numbers are as follows: *H. pylori*, P55993; *Campylobacter jejuni*, AJ 002379; *E. coli*, P00579; *Salmonella typhimurium*, P07336; *Pseudomonas fluorescens*, P52326; *Haemophilus influenzae*, P43766; *Bacillus subtilis*, P06224; *Xanthomonas campestris*, U82763; *Streptococcus pneumoniae*, Y11463; *Rhodobacter capsulatus*, P46400; and *Neisseria gonorrhoeae*, P52325.

of mutations into the promoter region of this chromosomal gene. Previous mapping of the 5′ end of the *vacA* transcript in *H. pylori* by primer extension analysis has demonstrated that there is a single conserved TSP located 119 nucleotides upstream of the AUG start codon, which suggests that *vacA* transcription is initiated by a single promoter (8, 21). Alignment of the *vacA* −10 regions of 12 different *H. pylori* strains (Fig. 3A) reveals a TAAAAA consensus sequence, which matches the consensus *E. coli* −10 hexamer (TATAAT) at four of six positions. Alignment of the *vacA* −35 regions reveals a consensus sequence (TTTATG) that matches the consensus *E. coli* −35 hexamer (TTGACA) at three of six positions (Fig. 3B).

pCTB2CAT (8), which contains 567 bp of *cysS*, the *cysS-vacA* intergenic region, 274 bp of *vacA* coding sequence from *H. pylori* 60190, and a chloramphenicol acetyltransferase (CAT) gene inserted at the 3′ terminus of *cysS*, was used as the template for all site-directed mutagenesis reactions. Inverse PCR for substitution mutagenesis of the −10 and −35 sequences and deletion of −108/−67, −66/−40, and −37/−32 sequences was performed by previously described methods (8). Briefly, oppositely oriented primers with *Bgl*II restriction sites incorporated at their 5′ ends were used to generate substitution mutations at the sites of the −10 and the −35 hexamers. After completion of 12 cycles of thermal cycling, template DNA was eliminated by *Dpn*I (New England Biolabs [NEB], Beverly, Mass.) digestion, and PCR products were then digested with *Bgl*II (NEB), purified by phenol-chloroform extraction and ethanol precipitation, and recircularized with T4 DNA ligase (NEB). To generate deletion mutations, oppositely oriented primers were designed such that the 5′ nucleotides of each of the two primers defined the region to be deleted. PCR products amplified with these primers were end polished with *Pfu* DNA polymerase (Stratagene) and recircularized with T4 DNA ligase. Religated plasmids (Table 1) were transformed into *E. coli* DH5α. Point mutations in the −10 region as well as substitution and insertion mutations in the −10 to −35 spacer region were introduced into pCB2CAT by using the GeneEditor system (Promega), a positive selection method, following the manufacturer's protocols. Mutation-bearing plasmids were introduced into the chromosome of *H. pylori* 60190 VX-1 (a reporter strain that contains a *vacA*::*xylE* transcriptional fusion) by natural transformation and allelic exchange, as described previously (4, 8). Mutants were selected on brucella agar plates containing kanamycin (25 μg/ml) and chloramphenicol (10 μg/ml). As a control, pCTB2CAT was introduced into *H. pylori* 60190 VX-1 to generate strain 60190 VX-1 CAT control (8).

To determine whether the putative *vacA* −10 hexamer is essential for transcription, we substituted the sequence AGA TCT for the native −10 *vacA* sequence (TAAAAG) found in *H. pylori* 60190. Introduction of this mutation into the chromosome of *H. pylori vacA* reporter strain 60190 VX-1 resulted in a mutant strain, 60190 VPC1/VX-1, with about 15-fold less

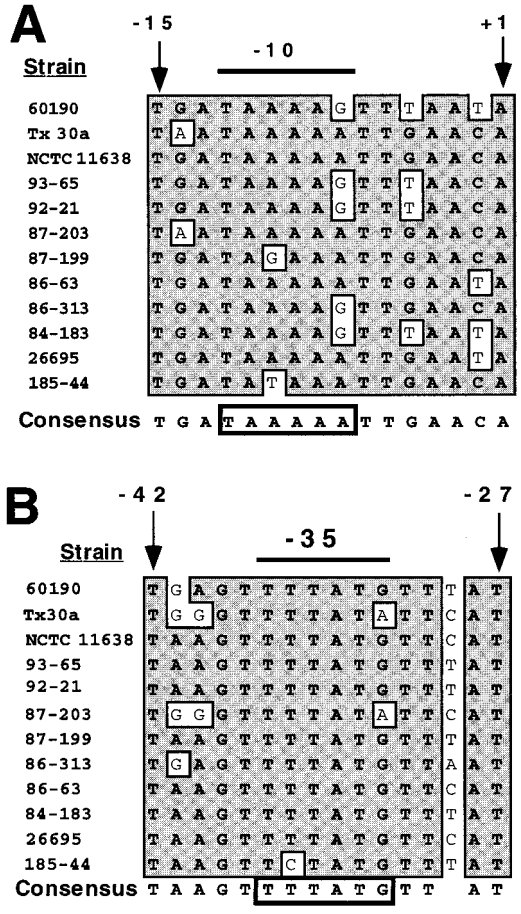

FIG. 3. Alignment of *vacA* −10 and −35 promoter elements. The putative *vacA* −10 (A) and −35 (B) promoter elements from 12 different *H. pylori* strains were aligned and compared. Solid bars indicate the predicted sites of RNA polymerase contact, and consensus sequences are indicated below the alignments.

XylE activity than the control strain (60190 VX-1 CAT control) (Fig. 4). This indicates that the native −10 sequence is essential for *vacA* transcription, as expected.

Five of the promoters shown in Fig. 1, including *vacA*, contain a TGN motif located at positions −15 to −13. To test whether this region might be involved in *vacA* transcription, point mutations were introduced at the −15, −14, −13, and −12 positions of the *vacA* promoter region in the chromosome of *H. pylori* 60190 VX-1 (Fig. 4). A transversion mutation at −15 (T to A) had no effect on the level of *vacA* transcription, whereas a transversion mutation at −14 (G to T) reduced *vacA* transcription more than twofold. Mutation of the −13 nucleotide had no effect on *vacA* transcription, but a mutation at the −12 position (T to G) resulted in a fourfold decrease. Taken together, these data indicate that the −10 region of the *vacA* promoter is essential for *vacA* transcription; they also suggest that the −14 position, located outside the consensus −10 hexamer, may contribute to *H. pylori* RNA polymerase binding.

To determine whether sequences in the −35 region play a functional role in *vacA* transcription, we deleted six nucleotides in this region (TTTATG) from the chromosome of *H. pylori* 60190 VX-1 (yielding strain 60190 VPC6/VX-1) (Fig. 5). Deletion of this region (−37 to −32) resulted in an approximately sixfold decrease in XylE activity. We also constructed a second mutant strain in which the native −37/−32 (TTTATG) sequence was replaced with a heterologous 6-bp sequence (AG ATCT), thereby altering six positions within the −35 region. The XylE activity of the −37/−32 substitution mutant (60190 VPC7/VX-1) was about fourfold less than the activity of the control strain (Fig. 5), which suggests that this region is indeed involved in binding RNA polymerase.

In most σ^{70} promoters, the region between the −35 and −10 hexamers does not play a sequence-specific role in mRNA initiation, but it is important for the maintenance of the proper spacing between the two RNA polymerase contact sites (6). To test, by an alternate approach, whether specific −35 sequences are important in *vacA* transcription, we introduced two separate mutations into the spacer region of the *vacA* promoter. First, a 5-bp heterologous substitution was made from −23 to −19 (CCTTA to GATCT). XylE specific activity in this mutant strain (60190 VPC10/VX-1) was similar to that of the control strain (Fig. 5), which indicates that specific sequences are not required in this spacer region. Next, a 10-bp heterologous insertion was introduced into the spacer region between −24 and −23, so that the native −35 and −10 sequences remained intact on the appropriate faces of the DNA helix. In this mutant strain (60190 VPC11/VX-1), the original −35 and −10 sites are now further apart and contact with RpoD at both sites simultaneously would be unlikely. XylE levels in mutant strain 60190 VPC11/VX-1 were markedly reduced, which provides further evidence that the native −35 sequence is involved in *H. pylori* RNA polymerase binding.

To examine the possibility that *vacA* transcription may depend on sequences upstream from the −35 hexamer, we deleted a 27-bp region from −66 to −40, as well as a 42-bp region from −108 to −67, and introduced each of these deletions into the chromosome of *H. pylori* 60190 VX-1. These sequences span the entire region between the *vacA* promoter and the upstream gene (*cysS*). XylE levels in these deletion mutants, under standard in vitro culture conditions, were essentially the same as in the control strain (60190 VX-1 CAT control) (Fig. 5). This demonstrates that all sequences necessary for basal levels of *vacA* transcription lie between −39 and −7.

The construction of various mutations in the *H. pylori vacA* promoter region allows us to reach several conclusions regarding transcription of this gene. First, as expected, at least a portion of the predicted −10 hexamer (TAAAAG) is essential for *vacA* transcription. Second, mutation of the −14 position (the G position of the TGN motif) results in a significant decrease in *vacA* transcription. This suggests that *H. pylori* RpoD may utilize an extended −10 promoter, perhaps analogous to a set of *E. coli* σ^{70} promoters that contain a TGN motif immediately upstream of the canonical −10 hexamer (1, 17, 23). Moreover, we speculate that the −14 *vacA* mutation in this study might underestimate the importance of this position, as it could lead to an alternative promoter (TATAAA). Finally, despite a lack of consensus −35 sequences among *H. pylori* promoters, mutation of this region results in decreased levels of *vacA* transcription. Thus, transcription of *H. pylori vacA* seems to involve the use of both an extended −10 sequence and a −35 binding site. More than half of the *H. pylori* promoters examined in Fig. 1 have a G nucleotide at position −14, and therefore, we speculate that this nucleotide might also be important in the transcription of other genes. It will be important in future studies to experimentally examine the potential use of −35 sequences in transcription of multiple *H. pylori* genes and also to determine whether extended −10 sequences are commonly utilized.

The lack of conservation among *H. pylori* promoters outside the −10 hexamer suggests that there may be considerable

TABLE 1. Plasmids used in this study

Designation	Genotype[b]	Source or reference
pCTB2CAT[a]	551 bp of 3′ *cysS*, CAT, P_{vacA}, 330 bp of 5′ *vacA*	8
pVacSUB(−10)	P_{vacA}, −12/−7 substitution (TAAAAG to AGATCT)	This study
pVacPRM(−15)	P_{vacA}, −15 substitution (T to A)	This study
pVacPRM(−14)	P_{vacA}, −14 substitution (G to T)	This study
pVacPRM(−13)	P_{vacA}, −13 substitution (A to G)	This study
pVacPRM(−12)	P_{vacA}, −12 substitution (T to G)	This study
pVacSUB(−23/−19)	P_{vacA}, −23 to −19 substitution (CCTTA to GATCT)	This study
pVacINS(−24/−23)	P_{vacA}, −24 to −23 insertion (GATCTTTTTT)	This study
pVacSUB(−37/−32)	P_{vacA}, −37 to −32 substitution (TTTATG to AGATCT)	This study
pVacDEL(−37/−32)	P_{vacA}, Δ−37/−32	This study
pVacDEL(−66/−40)	P_{vacA}, Δ−66/−40	This study
pVacDEL(−109/−67)	P_{vacA}, Δ−109/−67	This study

[a] pCTB2CAT was used as a template to generate all subsequent mutant plasmids listed.
[b] P_{vacA}, *vacA* promoter. CAT was from *Campylobacter coli*.

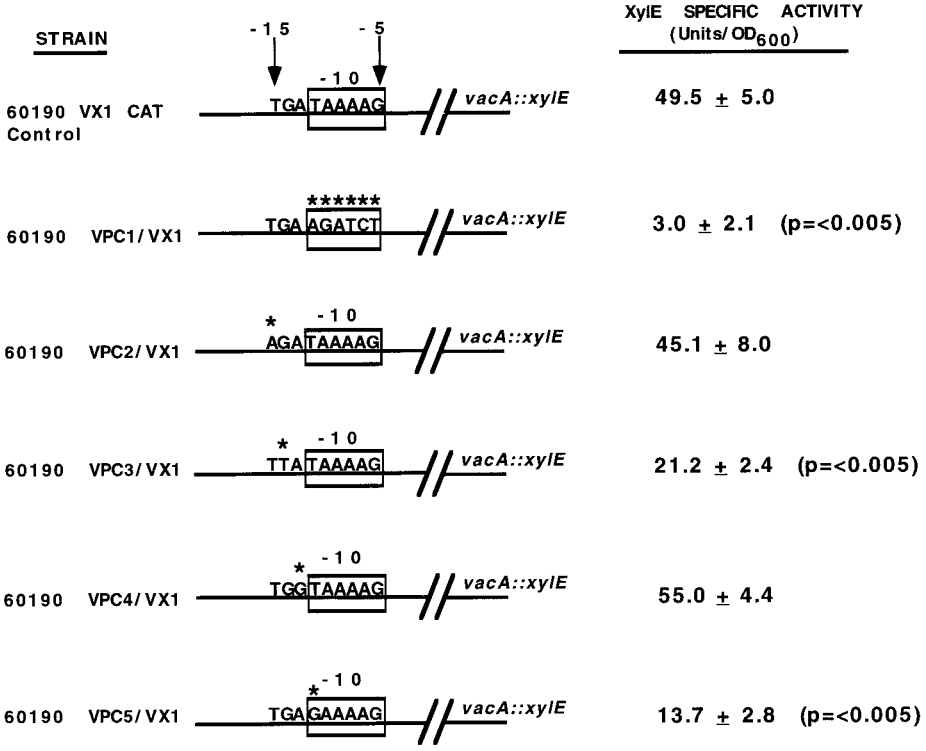

FIG. 4. Mutations in the *vacA* −10 promoter element. Mutations were introduced into the *vacA* −10 region of a reporter strain (*H. pylori* 60190 VX-1), which contains a *vacA*::*xylE* transcriptional fusion. Strain 60190 VX-1 CAT control is a *vacA*::*xylE* reporter strain that contains a CAT gene at the 3′ end of *cysS* but no changes in *vacA* promoter sequences (8). Strain 60190 VPC1/VX-1 contains a 6-bp substitution for the native −10 *vacA* hexamer. Other strains contain single nucleotide substitutions at the −15, −14, −13, and −12 positions of the *vacA* promoter. Sites of nucleotide substitutions are indicated by asterisks. The XylE specific activity of each strain is shown on the right. These values represent the means ± standard deviations for triplicate independent assays. *P* values represent comparisons with the XylE specific activity of the control strain (60190 VX-1 CAT control). OD_{600}, optical density at 600 nm.

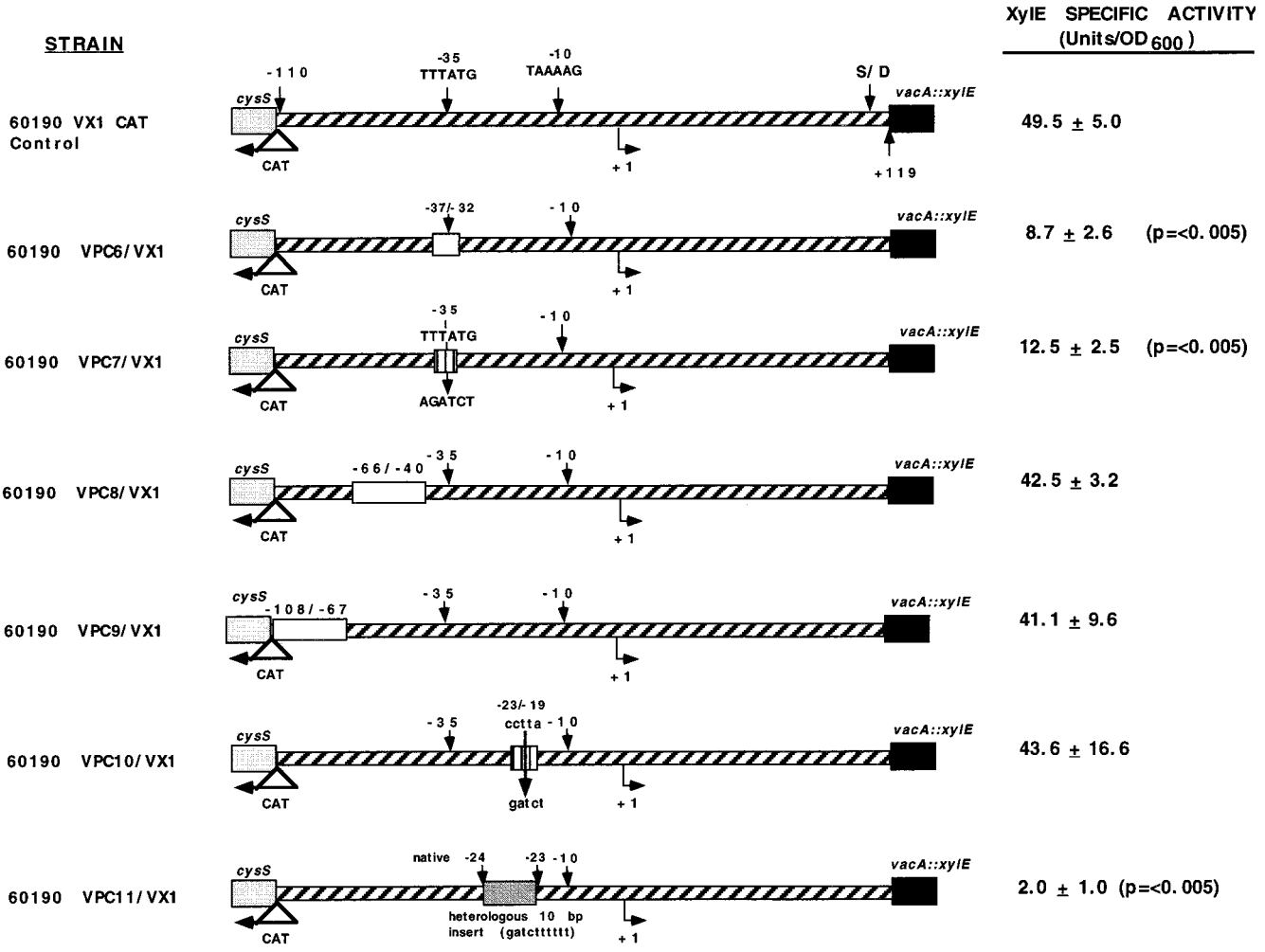

FIG. 5. Roles of the −35 hexamer, −10/−35 spacing, and the −108 to −40 region in *vacA* transcription. Various deletion mutations (open boxes), substitution mutations (vertically hatched boxes), or an insertion mutation were introduced into the *vacA* promoter region of an *H. pylori* reporter strain that contains a *vacA*::*xylE* transcriptional fusion (60190 VX-1). The *vacA* TSP is represented as +1. The putative *vacA* −35 sequence (TTTATG) and the −10 sequence (TAAAAG) are indicated with arrows, and upstream *cysS* and CAT genes are also indicated. Strain 60190 VX-1 CAT control is a *vacA*::*xylE* reporter strain that contains a CAT gene at the 3′ end of *cysS*, but no changes in *vacA* promoter sequences (8). The XylE specific activity of each strain is shown on the right. These values represent the means ± standard deviations for triplicate independent assays. *P* values represent comparisons with the XylE specific activity of the control strain (60190 VX-1 CAT control). OD_{600}, optical density at 600 nm.

variation in the avidity of RpoD binding to different promoters. We speculate that this may represent a mechanism for determining the levels at which individual genes are constitutively transcribed. Further study of these interactions may provide insight into the fundamental process of gene transcription in *H. pylori* and may be relevant to understanding how this organism persists in the human stomach for many decades.

This work was supported by grant AI 39657 from the National Institutes of Health and by the Medical Research Service of the Department of Veterans Affairs. Sequencing facilities used in this study are supported by NIH grant CA68485.

REFERENCES

1. **Barne, K. A., J. A. Bown, S. J. W. Busby, and S. D. Minchin.** 1997. Region 2.5 of the *Escherichia coli* RNA polymerase σ70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. EMBO J. **16:**4034–4040.
2. **Beier, D., G. Spohn, R. Rappuoli, and V. Scarlato.** 1997. Identification and characterization of an operon of *Helicobacter pylori* that is involved in motility and stress adaptation. J. Bacteriol. **179:**4676–4683.
3. **Cover, T. L.** 1996. The vacuolating cytotoxin of *Helicobacter pylori*. Mol. Microbiol. **20:**241–246.
4. **Cover, T. L., M. K. R. Tummuru, P. Cao, S. A. Thompson, and M. J. Blaser.** 1994. Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains. J. Biol. Chem. **269:**10566–10573.
5. **deHaseth, P. L., M. L. Zupancic, and M. T. Record, Jr.** 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. J. Bacteriol. **180:**3019–3025.
6. **Domboski, A. J., B. D. Johnson, M. Lonetto, and C. A. Gross.** 1996. The sigma subunit of *Escherichia coli* RNA polymerase senses promoter spacing. Proc. Natl. Acad. Sci. USA **93:**8858–8862.
7. **Dunn, B. E., H. Cohen, and M. J. Blaser.** 1997. *Helicobacter pylori*. Clin. Microbiol. Rev. **10:**720–741.
8. **Forsyth, M. H., J. C. Atherton, M. J. Blaser, and T. L. Cover.** 1998. Heterogeneity in levels of vacuolating cytotoxin gene (*vacA*) transcription among *Helicobacter pylori* strains. Infect. Immun. **66:**3088–3094.
9. **Ge, Z., and D. E. Taylor.** 1996. *Helicobacter pylori* genes *hpcopA* and *hpcopP* constitute a *cop* operon involved in copper export. FEMS Microbiol. Lett. **145:**181–188.
10. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15:**2343–2361.
11. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2254.
12. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57:**839–872.
13. **Heuermann, D., and R. Haas.** 1995. Genetic organization of a small cryptic plasmid of *Helicobacter pylori*. Gene **165:**17–24.

14. **Ishihama, A.** 1981. Subunit assembly of *Escherichia coli* RNA polymerase. Adv. Biophys. **14:**1–35.
15. **Ishihama, A.** 1990. Molecular assembly and functional modulation of *Escherichia coli* RNA polymerase. Adv. Biophys. **26:**19–31.
16. **Jones, A. C., R. P. H. Logan, S. Foynes, A. Cockayne, B. W. Wren, and C. W. Penn.** 1997. A flagellar sheath protein of *Helicobacter pylori* is identical to HpaA, a putative *N*-acetylneuraminyllactose-binding hemagglutinin, but is not an adhesin for AGS cells. J. Bacteriol. **179:**5643–5647.
17. **Kumar, A., R. A. Malloch, N. Fujia, D. A. Smillie, A. Ishihama, and R. S. Hayward.** 1993. The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. J. Mol. Biol. **232:**406–418.
18. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843–3849.
19. **Odenbreit, S., B. Wieland, and R. Haas.** 1996. Cloning and genetic characterization of *Helicobacter pylori* catalase and construction of a catalase-deficient mutant strain. J. Bacteriol. **178:**6960–6967.
20. **Pesci, E. C., and C. L. Pickett.** 1994. Genetic organization and enzymatic activity of a superoxide dismutase from the microaerophilic human pathogen, *Helicobacter pylori*. Gene **143:**111–116.
21. **Schmitt, W., and R. Haas.** 1994. Genetic analysis of the *Helicobacter pylori* vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein. Mol. Microbiol. **12:**307–319.
22. **Solnick, J. V., L. M. Hansen, and M. Syvanen.** 1997. The major sigma factor (RpoD) from *Helicobacter pylori* and other gram-negative bacteria shows an enhanced rate of divergence. J. Bacteriol. **179:**6196–6200.
23. **Spohn, G., D. Beier, R. Rappuoli, and V. Scarlato.** 1997. Transcriptional analysis of the divergent *cagAB* genes encoded by the pathogenicity island of *Helicobacter pylori*. Mol. Microbiol. **26:**361–372.
24. **Spohn, G., and V. Scarlato.** 1999. Motility of *Helicobacter pylori* is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. J. Bacteriol. **181:**593–599.
25. **Suerbaum, S., J. M. Thiberge, I. Kansau, R. L. Ferrero, and A. Labigne.** 1994. *Helicobacter pylori* hspA-hspB heat-shock cluster: nucleotide sequence, expression, putative function and immunogenicity. Mol. Microbiol. **14:**959–974.
26. **Tomb, J. F., O. White, A. R. Kerlavage, et al.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539–547.
27. **Wösten, M. M. S. M., M. Boeve, M. G. A. Koot, A. C. van Nuenen, and B. A. M. van der Zeijst.** 1998. Identification of *Campylobacter jejuni* promoter sequences. J. Bacteriol. **180:**594–599.