JOURNAL OF BACTERIOLOGY, Dec. 2002, p. 6615–6623 Vol. 184, No. 23
0021-9193/02/$04.00+0 DOI: 10.1128/JB.184.23.6615–6623.2002

Transcriptional Phase Variation of a Type III Restriction-Modification System in *Helicobacter pylori*

Nicolette de Vries,[1,2] Dirk Duinsbergen,[3] Ernst J. Kuipers,[1] Raymond G. J. Pot,[1] Patricia Wiesenekker,[3] Charles W. Penn,[4] Arnoud H. M. van Vliet,[1] Christina M. J. E. Vandenbroucke-Grauls,[3] and Johannes G. Kusters[1]*

Department of Gastroenterology and Hepatology, Erasmus MC—University Medical Center Rotterdam, Rotterdam,[1] and Departments of Gastroenterology[2] and Medical Microbiology and Infection Control,[3] VU University Medical Center, Amsterdam, The Netherlands, and School of Biosciences, University of Birmingham, Birmingham, United Kingdom[4]

Received 16 May 2002/Accepted 4 September 2002

Phase variation is important in bacterial pathogenesis, since it generates antigenic variation for the evasion of immune responses and provides a strategy for quick adaptation to environmental changes. In this study, a *Helicobacter pylori* clone, designated MOD525, was identified that displayed phase-variable *lacZ* expression. The clone contained a transcriptional *lacZ* fusion in a putative type III DNA methyltransferase gene (*mod*, a homolog of the gene JHP1296 of strain J99), organized in an operon-like structure with a putative type III restriction endonuclease gene (*res*, a homolog of the gene JHP1297), located directly upstream of it. This putative type III restriction-modification system was common in *H. pylori*, as it was present in 15 out of 16 clinical isolates. Phase variation of the *mod* gene occurred at the transcriptional level both in clone MOD525 and in the parental *H. pylori* strain 1061. Further analysis showed that the *res* gene also displayed transcriptional phase variation and that it was cotranscribed with the *mod* gene. A homopolymeric cytosine tract (C tract) was present in the 5′ coding region of the *res* gene. Length variation of this C tract caused the *res* open reading frame (ORF) to shift in and out of frame, switching the *res* gene on and off at the translational level. Surprisingly, the presence of an intact *res* ORF was positively correlated with active transcription of the downstream *mod* gene. Moreover, the C tract was required for the occurrence of transcriptional phase variation. Our finding that translation and transcription are linked during phase variation through slipped-strand mispairing is new for *H. pylori*.

Helicobacter pylori is a gastric pathogen that infects more than half of the human population (12). Colonization with this bacterium is a major cause of gastric and duodenal ulcers and is associated with the development of gastric atrophy and adenocarcinoma (8). Unless treated, colonization usually persists for life, which indicates that *H. pylori* is well adapted to the gastric environment.

Phase variation is an adaptive process involving the frequent, random, and reversible on-and-off switching of a gene. It occurs in a variety of bacterial species and plays an important role in bacterial pathogenesis and virulence (17, 45). Phase variation generates phenotypic variation in a bacterial population that allows bacteria to evade immune responses and to adapt efficiently to environmental changes. In this context, it is not surprising that phase-variable bacterial structures, such as flagella (34), pili (6), fimbriae (14), capsular structures (23), outer membrane proteins (38), and lipopolysaccharide biosynthesis genes (42), often have a function in the interaction with the host environment or are involved in virulence. However, other classes of genes with no established role in host-pathogen interaction, such as restriction and modification (RM) genes, also may display phase-variable expression (10, 13). Phase variation occurs at either the transcriptional level or the translational level (6, 42). Several molecular mechanisms can mediate phase variation at the transcriptional level, including promoter inversion (34), methylation of promoter sequences (6), and homopolymeric DNA tracts in the promoter (38). Slipped-strand mispairing is the most common mechanism of translational phase variation. Slippage of the DNA polymerase at a nucleotide repeat present in the coding region of a gene causes reversible frameshift mutations. This introduces a premature stop codon directly downstream of the nucleotide repeat which interrupts translation but not transcription, resulting in a truncated protein. In *H. pylori*, several genes display phase-variable expression. These include lipopolysaccharide synthesis genes (3, 4, 41), the *hopZ* gene, encoding a porin possibly involved in adhesion (29), the *oipA* gene, encoding an outer membrane proinflammatory protein (44), *pldA*, encoding a phospholipase A involved in outer membrane phospholipid composition (36), and *fliP*, encoding a flagellar basal body protein (21). All display translational phase variation through slipped-strand mispairing. So far, phase variation at the transcriptional level has not been described for *H. pylori*.

H. pylori possesses an unusual abundance of RM systems (27, 37). Several of these systems contain nucleotide repeats, which are thought to mediate phase variation through slipped-strand mispairing (2, 33, 37), although no experimental data

* Corresponding author. Mailing address: Department of Gastroenterology and Hepatology, room L-457, Erasmus MC—University Medical Center Rotterdam, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands. Phone: 31 10 4632982. Fax: 31 10 4632793. E-mail: kusters@mdl.azr.nl.

TABLE 1. Primers

Primer[a]	Molecule (bp position)[b]	5′→3′ sequence[c]	Use
BW4F	*lacZ* in pBW	CGTCAGTATCGGCGGAAT	Inverse PCR
BW2R	*lacZ* in pBW	GCAGCTCCTGCATCAATC	Inverse PCR
lacZ-RT-F	*lacZ* in pBW	CGACATTGGCGTAAGTGAAG	RT-PCR *lacZ* gene
lacZ-RT-3R	*lacZ* in pBW	AGTTTACCCGCTCTGCTACC	RT-PCR *lacZ* gene
1298-22F	RM locus (1)	TCGGATCCTTGTCAAGCCGTAAGA	Construction (RES29 and RES60)
1298-154F	RM locus (675)	AACGGATCCAAAGCCCTAGCCTCTTGTAA	Construction (RES284), PCR (clinical isolates)
1297-11R	RM locus (1063)	TCTGGATCCGGGTAACTTTGAGTGAATG	Construction (RES29)
1297-40R	RM locus (1094)	TCTGGATCCGTGATCTTTTATTGCTGATTG	Construction (RES60)
1297-77R	RM locus (1318)	AAGGGATCCTAGCCTTGCTTGTAGCATTC	Construction (RES284), PCR (clinical isolates)
1297PE-240R	RM locus (1064)	GGGGTAACTTTGAGTGAATG	Primer extension *res* gene
1297PE-433R	RM locus (1261)	CCGGTTGCCATTTCAAACAT	Primer extension *res* gene
1297RT-321F	RM locus (1336)	TTGTGAACAGCACCAGCATT	RT-PCR *res* gene
1297RT-607R	RM locus (1660)	TCACTAGCCTCACTATCATT	RT-PCR *res* gene
1297RT-2350F	RM locus (3365)	AGGCTTCACTTGTTATATGG	RT-PCR overlap *res-mod*
1296RT-71R	RM locus (3664)	ATCAGCTGTTAGGATTAAGG	RT-PCR overlap *res-mod*
1296RT-F	RM locus (3587)	CCACTCATAACCCTACTAGATA	RT-PCR *mod* gene
1296RT-R	RM locus (3946)	CTCGTTAGATTTGGTGTTGT	RT-PCR *mod* gene
katA-RT-F	*katA*/HP0875[d]	CTCAAACCAATTTGCCTAAC	RT-PCR *katA* gene
katA-RT-R	*katA*/HP0875	AAACGGATGGAATCGATACT	RT-PCR *katA* gene

[a] F and R in primer names indicate forward and reverse orientations, respectively, with respect to the target gene.
[b] Base pair positions that refer to the RM locus are as indicated in Fig. 2.
[c] *Bam*HI sites in 5′ extension of primers are underlined.
[d] Gene annotation number for *katA* (Tomb et al. [37]).

are available that support this. RM systems generally encode a restriction endonuclease, which cleaves DNA at specific recognition sites, and a DNA methyltransferase, which protects DNA from cleavage by methylation (22, 26). It has been stated that RM systems are tools in cellular defense, protecting bacteria against invading phages and foreign DNA from other sources (22). DNA fragmentation, stimulating the formation of recombinants, may be another function, and it has even been suggested that RM systems are selfish, mobile elements (22). In addition, DNA methyltransferases have been implied in the regulation of virulence genes (18). RM systems are classified as type I, II, or III on the basis of their composition and cofactor requirements, the nature of their target sequence, and the position of the DNA cleavage site. Type I systems are the most complex, comprising a three-subunit enzyme which is responsible for DNA recognition and catalyzes both restriction and modification. Its activity requires the cofactors *S*-adenosyl-L-methionine (AdoMet), ATP, and Mg^{2+}, and DNA cleavage can take place at variable sites several hundred base pairs away from the recognition site (26). Type II RM systems are the simplest, most common, and most studied of the RM systems. They consist of separate restriction and modification enzymes which function independently of each other, requiring only Mg^{2+} as a cofactor (26). Type III RM systems consist of a complex of the methyltransferase (the *mod* gene product) and the restriction enzyme (the *res* gene product) and require AdoMet, ATP, and Mg^{2+} as cofactors. The methyltransferase provides DNA recognition for both restriction and modification and catalyzes modification independently of the restriction enzyme. DNA cleavage requires a complex of both the subunits (19) and occurs at a specific site approximately 25 bp to one side of the recognition site (24). Type III methyltransferases typically contain two conserved motifs involved in AdoMet binding and catalysis (31). The presence of conserved DEAD box helicase motifs, involved in the ATP-dependent reactions of the enzyme, is characteristic of type III restriction enzymes (31).

In this study, we describe the identification and characterization of the molecular mechanism of transcriptional phase variation of a putative type III restriction-modification system for *H. pylori*.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The *H. pylori* strain 1061 was used as the parental strain for mutagenesis (15). Sixteen clinical *H. pylori* strains (92-1040, 92-1041, 92-1069, 92-1112, 92-213, 93-451, 93-720, 94-45, 94-841, 9A, 31C, 94-A, 93-178, 93-214, 93-216, and 93-236) were described previously (39). A previously described *H. pylori* strain 1061 library of 250 recombinant clones was used which contained genomic transcriptional *lacZ* reporter gene fusions generated by random integration of the *H. pylori* suicide vector pBW (11). *H. pylori* strains and pBW mutant derivatives were routinely cultured on Dent plates (Columbia agar [Oxoid, Basingstoke, United Kingdom], 7% lysed horse blood, Dent *H. pylori*-selective supplement [Oxoid]) at 37°C under microaerobic conditions (5% O_2, 10% CO_2, 85% N_2). *Escherichia coli* strains ER1793 (New England Biolabs, Beverly, Mass.) and DH5α MCR (Gibco BRL Life Technologies, Breda, The Netherlands) were used as host strains for cloning and were grown on Luria-Bertani media at 37°C. All media were supplemented with 100 μg of ampicillin, 30 μg of chloramphenicol or 20 μg of kanamycin per ml when appropriate.

Analysis of *lacZ* expression and switch rate by blue-white staining. The rate at which *lacZ* expression in the *H. pylori* recombinant clones switched on and off was determined by selecting single blue (on) or white (off) colonies after blue-white staining with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (Gibco BRL), as described previously (11). Selected colonies were subcultured to patches on agar once for 24 h and subsequently plated out to obtain approximately 100 to 200 single colonies per plate. After blue-white staining, the colonies that had switched to the opposite color were counted and the total number of colonies on each plate was determined. The switch rate was calculated as the proportion of switched colonies among the total number of colonies on a plate. For determining the ratio between the numbers of blue and white colonies in an equilibrium state, individual blue or white colonies were subcultured for approximately 250 generations, and the numbers of blue and white colonies were determined.

Recombinant DNA techniques. Natural transformation of *H. pylori* strain 1061, selection of kanamycin-resistant *H. pylori* recombinant clones, and inverse PCR

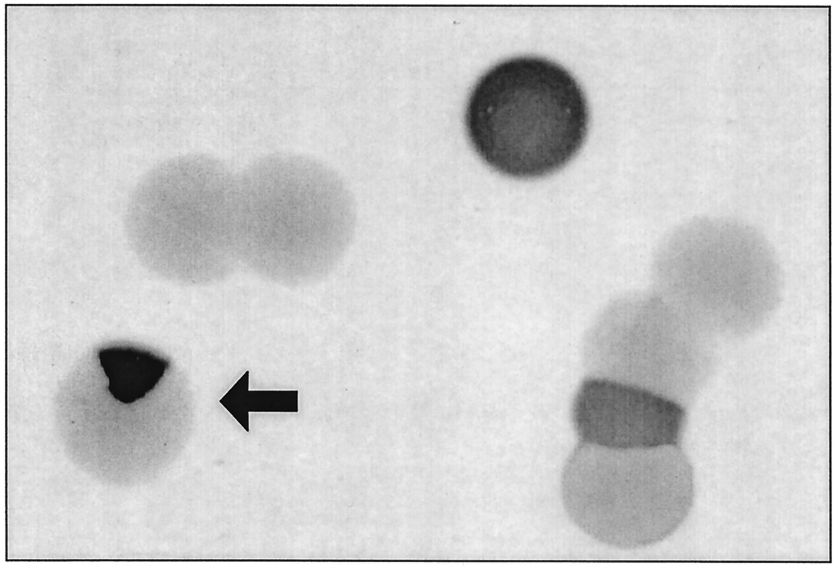

FIG. 1. Blue, white, and sectored colonies of *H. pylori* clone MOD525 as observed after subculturing of single white and blue colonies, which display phase-variable *lacZ* expression. Blue colonies are dark grey, white colonies are light grey, and a sectored colony is indicated by an arrow.

with primers BW4F and BW2R (Table 1) for determination of genomic pBW vector insertion points were carried out as described previously (5, 11). All DNA manipulations were carried out according to standard protocols (32). Restriction and modifying enzymes were purchased from Promega (Madison, Wis.). PCR was performed using the PCR Core system (Promega) and primers (Isogen Bioscience bv, Maarssen, The Netherlands) as listed in Table 1. Sequence analysis was performed directly on PCR products.

DNA and RNA isolation of single *H. pylori* colonies. A single *H. pylori* colony was suspended in 50 μl of phosphate-buffered saline (pH 7.4). For chromosomal DNA isolation, 25 μl of this suspension was boiled for 10 min, and after centrifugation, chromosomal DNA was extracted from the supernatant with the QIAEX II gel extraction kit (Qiagen, Hilden, Germany). For total RNA isolation, the remaining 25 μl of the suspension was spread out in a small patch on a Dent plate and cultured for 16 h. Bacteria were harvested in 1 ml of phosphate-buffered saline (pH 7.4) and washed once, and the optical density of the bacterial suspension was measured at 600 nm (OD_{600}), with an optical density at 600 nm of 1 corresponding to 5×10^8 CFU per ml. Total RNA was isolated from approximately 10^9 *H. pylori* cells with Trizol (Gibco BRL Life Technologies) and dissolved in distilled water according to the manufacturer's instructions. RNA was then treated with 10 U of RQ1 DNase I (Promega), ethanol precipitated, and extracted again with 250 μl of Trizol to remove any residual DNA contamination.

Primer extension. To map the transcription start site of the *res* gene of *H. pylori* strain 1061, primer extension was performed as described previously (9, 20), using primers 1297PE-240R and 1297PE-433R (Table 1). Nucleotide sequence reactions were performed with the same primers, using the Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham) and the RES284-promoter fragment (see below) cloned in the $pGEMT_{easy}$ vector (Promega).

RT-PCR. Reverse transcription (RT)-PCR was performed on total RNA with specific primers based on the *res* gene, the *mod* gene, an overlapping region of the *res* and *mod* genes, the *lacZ* gene, and the *katA* gene (Table 1), yielding amplicons of 325, 360, 300, 317, and 377 bp, respectively. Two to five micrograms of total RNA and 25 pmol of each reverse primer were mixed, and stepwise primer annealing was carried out for 2 min at 70°C, 1 min at 65°C, 1 min at 60°C, 1 min at 55°C, and 1 min at 45°C, followed by incubation for 5 min on ice. The RT reaction was performed with 5 U of avian myeloblastosis virus reverse transcriptase (Promega) for 30 min at 42°C, followed by five cycles of 1 min at 50°C, 1 min at 53°C, and 1 min at 56°C. Of the cDNA, 2 μl was used in separate standard PCRs of 25 μl for each primer combination. Ten microliters of the PCR products was analyzed on a 1.5% agarose gel. To confirm that similar amounts of total RNA were used in the individual reactions, RT-PCR with primers based on the housekeeping gene *katA*, encoding the catalase enzyme (28), was carried out in parallel. It was previously demonstrated that the *katA* gene is constitutively transcribed under the growth conditions used in the present study (11) and was therefore considered to be a valid control for RT-PCR. To exclude the presence of residual DNA, for each RNA sample the complete RT-PCR procedure was also carried out without adding RT.

Construction of *H. pylori* recombinant clones containing promoter-*lacZ* fusions. Fragments of 644 to 1094 bp, containing the 5′ end of the upstream gene (homolog of HP1406 [37] and JHP1298 [2]), the promoter region, and the 5′ end of the *res* gene, were amplified by PCR using primers with a 5′ extension containing a *Bam*HI restriction site (Table 1). PCR fragments were cloned in the $pGEM\text{-}T_{easy}$ vector (Promega) and, as described previously (11), subcloned into the unique *Bgl*II site upstream of the promoterless *lacZ* reporter gene in the pBW vector, resulting in the suicide plasmids pRES284, pRES60, and pRES29. These plasmids were transformed into *H. pylori* strain 1061, and subsequent selection on kanamycin-containing agar plates resulted in the *H. pylori* recombinant clones RES284, RES60, and RES29, with a genomic pBW vector insertion flanked by two copies of the cloned fragment (see Fig. 3). The orientation of the PCR fragment in the pBW plasmids and correct genomic integration of the constructs were verified by PCR followed by sequence analysis.

Nucleotide sequence accession number. The DNA sequence of the *res* and *mod* locus, containing an intact *res* open reading frame (ORF) with a C tract of 14 cytosines, has been deposited in the GenBank sequence database under accession number AF536178.

RESULTS

Identification of an *H. pylori* clone showing phase-variable *lacZ* expression. In a previous screening for regulated *H. pylori* genes of an *H. pylori* strain 1061 library with random genomic transcriptional *lacZ* reporter gene fusions (11), a clone was identified showing blue, white, and blue-white sectored colonies (Fig. 1). Subculturing of a sectored colony yielded single blue, white, and sectored colonies. The progeny of the individual blue and white colonies again were blue, white, and sectored, which strongly indicated that this *H. pylori* clone, designated MOD525, displays phase-variable *lacZ* expression.

If phase variation were involved, random on-and-off switching of *lacZ* expression would eventually lead to an equilibrium situation where the ratio between the numbers of white and blue colonies depends on the ratio between the rates at which transcription is switched on and off. With clone MOD525, the

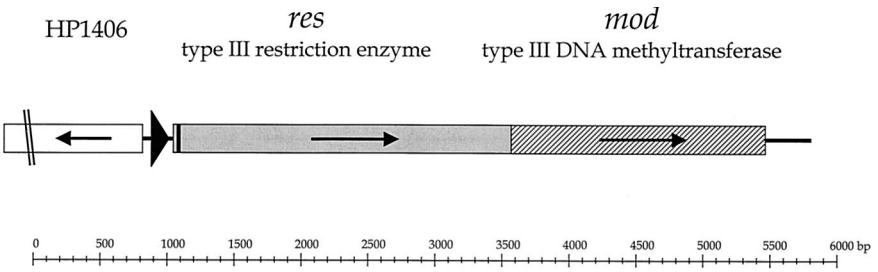

FIG. 2. Genomic organization of the type III RM locus in the wild-type *H. pylori* strain 1061. The complete *res* gene (grey) and *mod* gene (hatched) and the 5′ end of the upstream gene (no fill) are shown. At 27 bp downstream from the translational start site at position 1035 of the *res* gene is a C tract (black box). A gene encoding a putative biotin synthetase (HP1406) is located 200 bp upstream of the *res* gene in the opposite transcriptional orientation. The relative base pair positions are indicated on a line at the bottom. The black triangle represents the promoter region of the *res* gene, and small black arrows indicate the transcriptional orientation of the genes.

rate of switching from off to on was 0.0075 (12 colonies out of 1,603 switched from white to blue). The switch rate from on to off was 3.5-fold higher, namely, 0.027 (13 colonies out of 483 switched from blue to white). Due to the 3.5-fold difference in switch rates, an equilibrium population is expected to have a proportion of 22% of the bacteria containing a switched-on *lacZ* gene and 78% with a switched-off *lacZ* gene. Indeed, in the equilibrium state that evolved either from white or from blue colonies, the proportion of blue colonies was 24% and that of white colonies was 76% in a total of 4,800 colonies. It was concluded that *H. pylori* clone MOD525 contained a transcriptional *lacZ* fusion in a genetic locus that was subject to reversible on-and-off switching, i.e., phase variation.

A putative type III DNA methyltransferase gene mediates phase-variable *lacZ* expression. To identify the *H. pylori* locus that was responsible for the phase-variable transcription of the *lacZ* reporter gene, the chromosomal DNA sequences flanking the integrated pBW vector in *H. pylori* clone MOD525 were determined by inverse PCR and sequence analysis. The *lacZ* fusion was located in an ORF of 1,860 bp (Fig. 2 and 3A), with a nucleotide identity of 92% to the JHP1296 gene encoding a putative type III DNA methyltransferase (*mod*) in *H. pylori* strain J99 (2). An ORF of 2,547 bp was present directly upstream of the *mod* gene and showed 91% identity to the JHP1297 gene (*res*), which encodes a putative type III restriction endonuclease in *H. pylori* strain J99. At 200 bp upstream of the *res* gene, a putative biotin synthetase-encoding gene of 849 bp (HP1406; JHP1298) was present in the opposite transcriptional orientation (2, 37). The *res* and *mod* ORFs show an overlap of 7 bp and are organized in an operon-like structure with a GC content of 35%. No clear homologs are present in the genome of *H. pylori* strain 26695 (37). The conserved motifs involved in AdoMet binding, catalysis, and ATP-dependent reactions, which are characteristic for all type III RM systems (31), are present in the putative translated products of the *mod* and the *res* genes. It is therefore likely that the RM system identified in this study is a type III RM system.

The transcription start site of the *res* gene was mapped by primer extension with the primer 1297PE-240R, and this yielded a weak but reproducible band (Table 1; Fig. 4A). In contrast, primer 1297PE-433R, which was located 198 bp further downstream (Table 1), did not yield detectable bands (data not shown). The putative transcription start site of the *res* gene was thus located 21 nucleotides upstream of the translational start site of the *res* gene (Fig. 4B).

Phase variation of the *mod* gene occurs at the transcriptional level. To demonstrate that the observed phase variation occurred at the transcriptional level, RT-PCRs based on the *mod* and *lacZ* genes were performed with total RNA derived from blue and from white colonies of clone MOD525. RT-PCRs with the *katA* gene carried out on all RNA samples yielded bands of the same intensity, indicating that similar quantities of RNA were used in each reaction (Fig. 5). Strong RT-PCR bands for *mod* and *lacZ* were detected only in blue MOD525 colonies, whereas either weak or no bands were observed when RNA of white colonies was used (Fig. 5A). To determine whether the *H. pylori* wild-type strain also displayed transcriptional phase variation of the *mod* gene, RT-PCRs were performed with RNA from 20 single colonies of the *H. pylori* parental strain 1061 (Fig. 5B). In two colonies no mRNA of *mod* was detectable. In all other colonies the intensities of the bands ranged from very weak to strong, whereas the *katA* control always yielded bands of similar intensities (Fig. 5).

The *res* gene also displays transcriptional phase variation. To investigate whether the transcription of the upstream *res* gene correlated with *mod* transcription, an *H. pylori* recombinant clone, designated RES284, was created. The *H. pylori* suicide plasmid pRES284 was integrated into the *H. pylori* strain 1061 chromosome, resulting in a genomic transcriptional *lacZ* fusion to a DNA fragment containing the *res* promoter region and the 5′ end of the *res* gene (Fig. 3B). Single kanamycin-resistant colonies of clone RES284 were stained by blue-white staining, and these were blue, white, and sectored, similar to those of the original clone MOD525. Reversible on-and-off switching of *lacZ* expression was verified by the finding that blue colonies yielded both blue and white colonies, and conversely, white colonies switched to yield both blue and white colonies.

The *res* and *mod* genes are cotranscribed. To determine whether the *res* and *mod* genes are coordinately transcribed, RT-PCRs were performed with both genes on RNA isolated from four white colonies and four blue colonies of clone MOD525. Strong RT-PCR bands were detected for the *mod* and *res* genes for all blue colonies, whereas only weak or no bands were observed for both genes for white colonies (Fig. 6). To investigate whether this coregulation of the *res* and *mod*

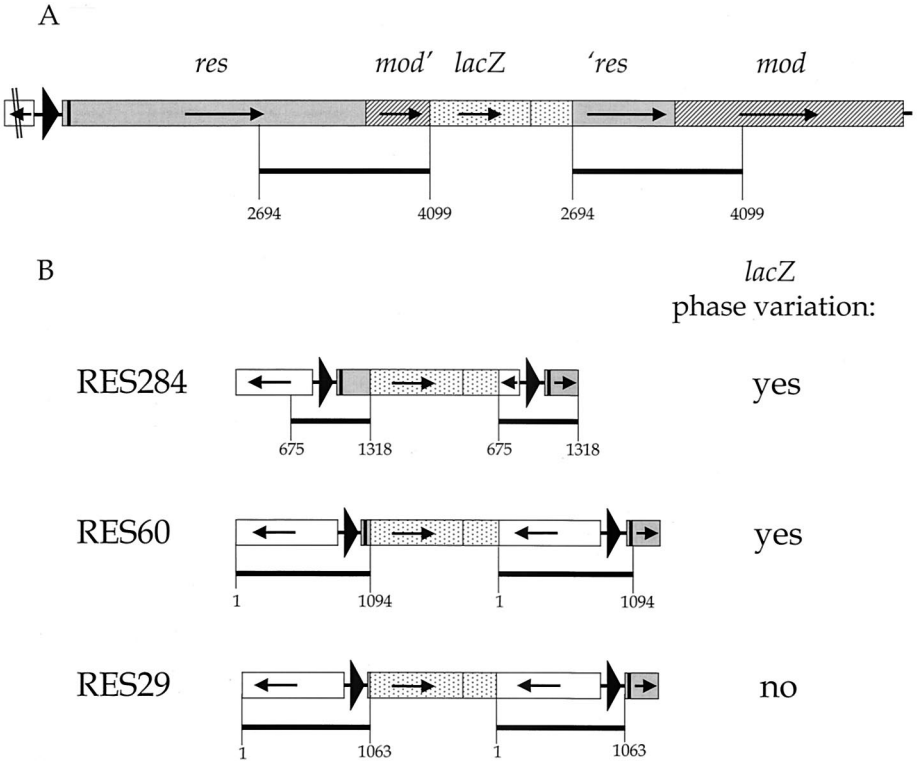

FIG. 3. Genomic organization of the type III RM locus in the *H. pylori* strain 1061 clones MOD525 (A) and RES284, RES60, and RES29 (B). Numbers in the names of the clones refer to the distance between the pBW insertion point and the translational start codon of the *mod* or *res* gene. (A) The pBW integration in clone MOD525 starts at base pair position 4099, 525 bp downstream from the translational start site of the *mod* gene. A 1,406-bp fragment containing the 3′ end of the *res* gene (grey) and part of the 5′ end of the *mod* gene (hatched) was duplicated upon integration of the pBW construct into the genome via single homologous recombination. (B) Clones RES284, RES60, and RES29 contain a genomic transcriptional *lacZ* fusion to one copy of the duplicated DNA fragment containing the 5′ end of the HP1406 gene (no fill) and the promoter region and 5′ end of the *res* gene (grey). The second copy of this fragment is downstream of the pBW vector integration and is part of an intact RM locus (not fully shown here). For an explanation of symbols see also the legend to Fig. 2; dotted areas represent the integrated pBW vector with the *lacZ* gene (not to scale), and black bars refer to duplicated DNA fragments, with the numbers at the borders indicating base pair positions corresponding to those in Fig. 2.

genes was due to the presence of a bicistronic transcript, cDNA was synthesized by RT with primer 1296RT-R1, located 372 bp downstream from the *mod* translational start at bp position 3575 (Table 1). This cDNA was then used as a template for a PCR with a region overlapping the *res* gene with 210 bp and the *mod* gene with 90 bp (Fig. 6). This resulted in strong RT-PCR bands for blue MOD525 colonies but weak or absent bands for white colonies, suggesting that the *res* and *mod* genes were transcribed on the same mRNA.

The length of the C tract in the *res* gene correlates with transcriptional phase variation. Interestingly, the *res* gene, located upstream of the *mod* gene, contains a tract of cytosines (C tract) in its 5′ end, 27 bp downstream of the *res* translational start site (Fig. 2 and 4B). Slipped-strand mispairing at C tracts is known to cause frameshift mutations and phase variation in other genes (3). Therefore, from 9 individual blue colonies and 16 white colonies of *H. pylori* clone MOD525, the lengths of the C tracts in the *res* gene were determined (Table 2). In 6 of the 25 colonies a C tract of 14 cytosines was present, resulting in an intact *res* ORF, the translation of which would lead to the production of a full-length *res* protein. Five of these colonies (83%) were blue, having *mod* transcription switched on. In 19 colonies the C-tract length was 12, 13, or 15 cytosines, which led to a premature stop codon directly downstream of the C tract (Table 2). Fifteen of these colonies (79%) were white, having *mod* transcription switched off. An imperfect correlation was observed between the presence of an intact *res* ORF, i.e., a putatively translated *res* gene, and blue colonies, i.e., *mod* transcription being switched on ($P = 0.025$, Fisher's exact test).

C tract mediates the on-and-off switching of transcription. To provide further evidence for the role of the C tract in the on-and-off switching of transcription, two additional *H. pylori* recombinant clones, RES29 and RES60, were created, using the *H. pylori* suicide plasmids pRES29 and pRES60 (Fig. 3B). Clone RES60, which contained a *lacZ* fusion to a fragment that included the C tract, showed reversible on-and-off switching of *lacZ* expression, yielding blue, white, and sectored colonies upon repeated subculturing. This phenotype was identical to that of clones MOD525 and RES284. In contrast, clone RES29, which contained a *lacZ* fusion to the fragment that lacked the C tract, showed constitutive *lacZ* expression, yielding colonies that were always blue, also after repeated subculturing.

The RM locus is present in most *H. pylori* clinical strains.

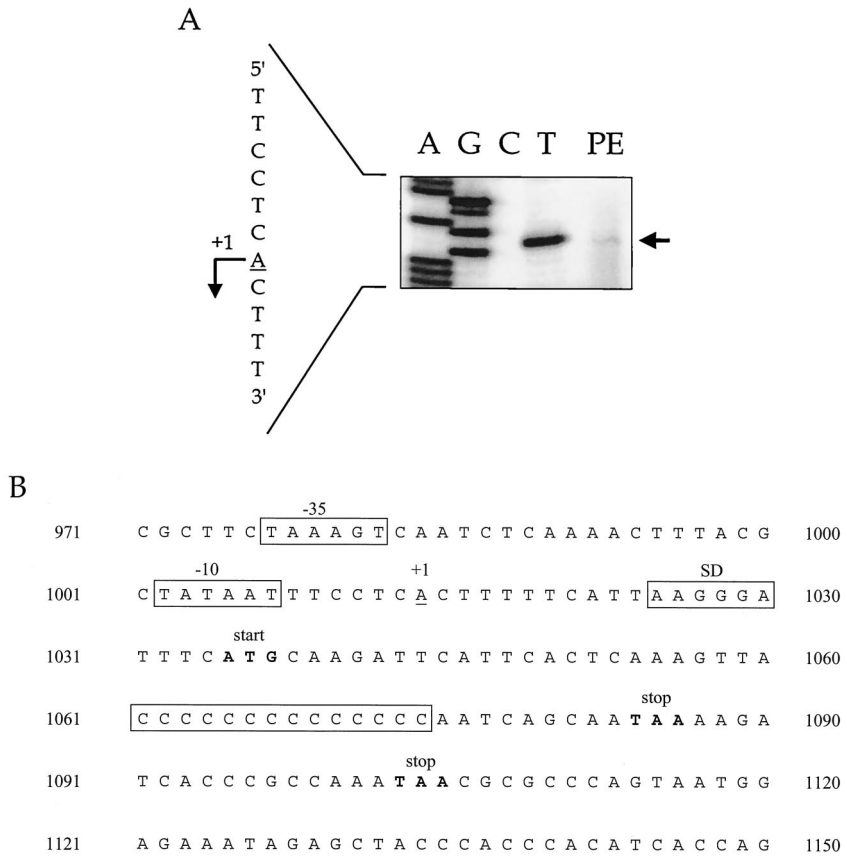

FIG. 4. Mapping of the transcription start site of the *res* gene by primer extension. (A) Primer extension performed with primer 1297PE-240R results in a weak but reproducible band, shown in lane PE. The sequence reactions performed with the same primer are shown in lanes A, G, C, and T and represent the noncoding DNA strand. At the left-hand side of the panel, the transcription start site is indicated with a bent arrow on the corresponding coding DNA strand. (B) Promoter and 5′ region of the *res* gene are shown, including a C tract of 14 Cs, to yield an intact *res* ORF. The transcription start site (+1) is underlined; the −10 and −35 boxes, the putative Shine-Dalgarno sequence (SD), and the C tract are boxed; the translational start codon at position 1035 and premature stop codons, at position 1084 when there are 13 Cs and at position 1103 when there are 12 or 15 Cs, are in boldface.

To investigate the presence of the RM locus in *H. pylori* clinical strains, a fragment including the 5′ end of the upstream gene (HP1406) and the promoter and 5′ end of the *res* gene was amplified by PCR from 16 clinical *H. pylori* strains (39). For 15 strains, a PCR fragment of the predicted size was amplified, whereas one strain, strain 9A, did not yield a product (data not shown). This suggests that the *res-mod* locus is widely distributed in *H. pylori*.

DISCUSSION

Little is known about the mechanisms that allow *H. pylori* to adapt so efficiently to the gastric environment, enabling it to establish a lifelong colonization. Phase variation is one strategy that could be important in this adaptation, since it allows a small part of the bacterial population to express genes that are required only when conditions change rapidly or unexpectedly. Several of the many RM systems of *H. pylori* are thought to display phase variation (33). The presence at critical time points in the infection of a small portion of the *H. pylori* population expressing RM genes may provide a selective advantage by facilitating DNA uptake and transformation or, in contrast, by preventing uptake and integration of foreign DNA. In the present study, we demonstrate that one of the putative *H. pylori* type III RM systems shows phase variation both at the translational level by slipped-strand mispairing at a C tract and at the transcriptional level.

Both the *res* gene and the *mod* gene showed transcriptional phase variation. In addition, the transcription of the *res* and *mod* genes seemed to be coregulated, since mRNA of both genes was clearly detected in blue colonies of the MOD525 clone, whereas in white colonies it was absent or detected at only very low levels. A restriction enzyme and its matching DNA methyltransferase are likely to be expressed coordinately, since these enzymes usually operate in a cooperative manner (26). In particular, this is expected for type III RM enzymes, since their restriction activity requires the formation of a complex between the restriction enzyme and the DNA methyltransferase (19). Consequently, RM genes are often transcribed in an operon by a mutual promoter or are regulated by a common regulator (7). We failed to detect any transcripts by Northern hybridization analysis, possibly due to the low transcription levels of the *res* and *mod* genes and long, unstable transcripts (not shown). However, RT-PCR with a

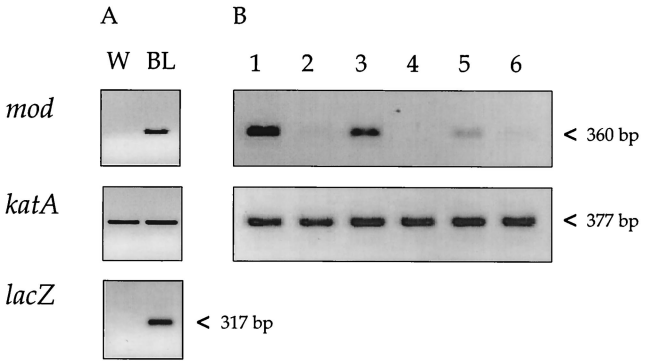

FIG. 5. Phase variation of the *mod* gene is at the transcriptional level. (A) RT-PCR products for the *mod* (360 bp), *katA* (377 bp, control), and *lacZ* (317 bp) genes, synthesized from total RNA derived from a white (W) and blue (BL) colony of *H. pylori* clone MOD525. (B) RT-PCR of single colonies of the *H. pylori* 1061 parental strain. RT-PCR results for the *mod* and *katA* genes (control) are shown from 6 of the 20 colonies.

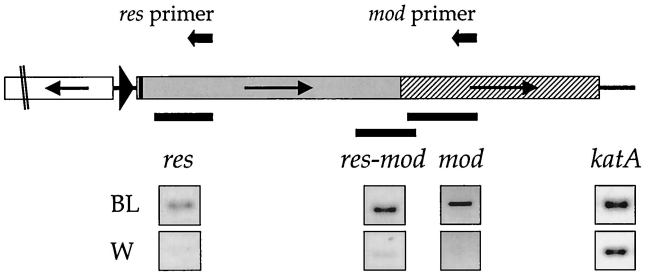

FIG. 6. Cotranscription of the *res* and *mod* genes shown by RT-PCR. RNA from a blue MOD525 colony (BL) yields positive PCRs for the *res* gene, the *mod* gene, and the overlapping region, whereas RNA from a white colony (W) gives no PCR signal or a weak PCR signal. The locations of the amplified fragments of the *res* gene (325 bp), the *mod* gene (360 bp), and a region overlapping the *res* and *mod* genes (*res-mod*, 300 bp) are indicated by black bars. The *katA* gene was used as a control. The reverse primers 1297RT-607R (*res* primer) and 1296RT-R (*mod* primer), used for the synthesis of the cDNA template, are indicated by black arrows above the respective genes.

region overlapping the genes strongly suggests that the genes are transcribed on a bicistronic mRNA and are organized in an operon.

The RNA from some white colonies of the MOD525 clone yielded weak RT-PCR bands instead of the expected negative signal (see Fig. 6). During the subculturing of colonies prior to RNA isolation, bacteria will inevitably switch, which leads to a mixed population of cells in the on and off status. The weak RT-PCR signal observed with RNA from white colonies probably represents a low proportion of switched-on bacteria in a population of predominantly white cells. Single colonies of the *lacZ* clones were either blue, white, or sectored, and no visible differences in the intensity of blue staining were observed. Therefore, it seems unlikely that intermediate transcription levels caused the variable RT-PCR signal. The status for colonies of the wild-type *H. pylori* strain 1061 is unknown, and selected colonies could be composed of a mixed cell population. This probably explains why colonies of the wild-type strain yielded RT-PCR signals that ranged from negative to strong.

The presence of an intact *res* ORF, as inferred from the length of the C tract, was positively correlated with active transcription of the downstream *mod* gene in clone MOD525 (see Table 2). This indicates that the translation of the *res* gene is linked to the transcription of the RM genes. The correlation between C-tract length and *lacZ* expression was imperfect, however. This may be due to PCR errors and sequencing artifacts caused by the presence of the C tract. Alternatively, an imperfect correlation may result from being unable to differentiate blue colonies with a small white sector from fully blue colonies. DNA isolated from these colonies would also contain DNA with a C tract of a length other than 14 Cs. When the C tract is shorter than 14 Cs, it might have a competitive advantage in the PCR due to the depletion of dCTP, in particular. Although the majority of the DNA template contains a tract of 14 Cs, this would lead to amplification of a fragment with a C tract associated with switched-off expression.

A homopolymeric nucleotide tract present in an ORF and related to phase variation usually acts at the translational level through slipped-strand mispairing. Interestingly, the presence of the C tract located in the 5′ region of the *res* gene not only affected translation but also was essential for phase-variable transcription of the *res* gene. When a homopolymeric nucleotide tract is located in a promoter, slippage at this tract can modulate transcription levels by changing the spacing of the −10 and −35 boxes of the promoter (38). However, a *lacZ* reporter gene fusion to a *res* gene fragment lacking the C tract was constitutively transcribed, suggesting that the promoter driving *res* transcription was located upstream of the C tract. Indeed, at 21 nucleotides upstream of the translational start of the *res* gene, a transcription start site was present, preceded by a −10 (TATAAT) box and a −35 (TAAAGT) box that were highly similar to the *H. pylori* consensus σ^{80} promoter elements, TATAAT and TTAAGC (40). No alternative transcription start site or translational start codon for the *res* gene, associated with a Shine-Dalgarno sequence, could be identified in the region downstream of the C tract. This indicated that the C tract indeed was located in the *res* ORF and not in the promoter region.

A possible explanation for the effect of C-tract length on transcription of the *res* gene could be an increased instability of the untranslated mRNA, present after a frameshift mutation at the C tract. Alternatively, the addition or deletion of a cytosine

TABLE 2. The presence of an intact and disrupted *res* ORF, as determined from variable C-tract lengths of 25 individual blue and white colonies of clone MOD525, correlates with *mod* transcription being switched on and off, respectively, as determined from the *lacZ* expression ($P = 0.025$, Fisher's exact test)

lacZ expression (*mod* transcription)	No. of colonies with *res* ORF			
	Intact, 14 Cs[c]	Disrupted[d]		
		12 Cs	13 Cs	15 Cs
On[a]	5	0	2	2
Off[b]	1	1	2	12

[a] Blue colonies ($n = 9$).
[b] White colonies ($n = 16$).
[c] Total of six colonies.
[d] Total of 19 colonies.

in the C tract may change the bending of the DNA or lead to formation of terminator structures in the RNA, affecting the transcription process (25, 30, 35). A third possibility may involve transcription termination at intragenic Rho-dependent terminator sites, where the presence of a premature stop codon leaves a region of the mRNA free of ribosomes, allowing binding by the Rho protein (16, 23, 43). Thus far, only one study with *Neisseria meningitidis* has shown that this mechanism could also be involved in phase variation (23). A homolog of Rho (HP0550) is present in the genome sequence of *H. pylori* (37) and in the *res* gene of *H. pylori* strain 1061; slipped-strand mispairing at the C tract introduces premature stop codons, which presumably stops the translation. Furthermore, the cytosine-guanine content of the region downstream of the C tract ranges between 1.5 and 2, suggesting that Rho-dependent termination sites may be present (1). However, no experimental evidence on the function of Rho in *H. pylori* is available, and Rho-dependent premature transcription termination has not been described for *H. pylori*.

In the context of efficient use of resources, it would be expected that bacteria prevent the synthesis of an untranslated transcript. In the present study, the *res* gene displayed phase variation by slipped-strand mispairing at a C tract, which presumably switches *res* translation on and off. The transcription of the *res* and *mod* genes is switched on and off accordingly, a finding that fits well in the above theory. Translational coupling of transcription, as well as translational phase variation through slipped-strand mispairing, is well studied in many bacteria. However, our finding that translation and transcription are linked during the process of phase variation is new for *H. pylori* and has been reported only once before, for another bacterial species (23). The *H. pylori* genome has several genes that contain repeats in the 5′ region, which may be implied in phase variation through slipped-strand mispairing (33). It is not unlikely that for some of these genes, as well as phase-variable genes of other bacteria, their transcription is linked to translation in order to limit the waste of resources through the synthesis of unused transcripts.

ACKNOWLEDGMENTS

This study was financially supported by a grant from the Research Stimulation Fund of the Vrije Universiteit, Amsterdam, The Netherlands (USF), to E.J.K. Part of the work was financed by a fellowship of the Federation of European Microbiological Societies (FEMS) to N.D.V.

We thank Ben Appelmelk for useful discussions and advice and Ferry Namavar for providing the clinical *H. pylori* strains.

REFERENCES

1. **Alifano, P., F. Rivellini, D. Limauro, C. B. Bruni, and M. S. Carlomagno.** 1991. A consensus motif common to all Rho-dependent prokaryotic transcription terminators. Cell **64:**553–563.
2. **Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176–180.
3. **Appelmelk, B. J., S. L. Martin, M. A. Monteiro, C. A. Clayton, A. A. McColm, P. Zheng, T. Verboom, J. J. Maaskant, D. H. van den Eijnden, C. H. Hokke, M. B. Perry, C. M. Vandenbroucke-Grauls, and J. G. Kusters.** 1999. Phase variation in *Helicobacter pylori* lipopolysaccharide due to changes in the lengths of poly(C) tracts in α3-fucosyltransferase genes. Infect. Immun. **67:** 5361–5366.
4. **Appelmelk, B. J., M. C. Martino, E. Veenhof, M. A. Monteiro, J. J. Maaskant, R. Negrini, F. Lindh, M. Perry, G. Del Giudice, and C. M. J. E. Vandenbroucke-Grauls.** 2000. Phase variation in H type I and Lewis a epitopes of *Helicobacter pylori* lipopolysaccharide. Infect. Immun. **68:**5928–5932.
5. **Bijlsma, J. J., C. M. Vandenbroucke-Grauls, S. H. Phadnis, and J. G. Kusters.** 1999. Identification of virulence genes of *Helicobacter pylori* by random insertion mutagenesis. Infect. Immun. **67:**2433–2440.
6. **Blyn, L. B., B. A. Braaten, and D. A. Low.** 1990. Regulation of *pap* pilin phase variation by a mechanism involving differential *dam* methylation states. EMBO J. **9:**4045–4054.
7. **Butler, D., and G. F. Fitzgerald.** 2001. Transcriptional analysis and regulation of expression of the *Scr*FI restriction-modification system of *Lactococcus lactis* subsp. *cremoris* UC503. J. Bacteriol. **183:**4668–4673.
8. **Cover, T. L., and M. J. Blaser.** 1992. *Helicobacter pylori* and gastroduodenal disease. Annu. Rev. Med. **43:**135–145.
9. **Davies, B. J., N. de Vries, S. G. Rijpkema, A. H. M. van Vliet, and C. W. Penn.** 2002. Transcriptional and mutational analysis of the *Helicobacter pylori* urease promoter. FEMS Microbiol. Lett. **213:**27–32.
10. **De Bolle, X., C. D. Bayliss, D. Field, T. van de Ven, N. J. Saunders, D. W. Hood, and E. R. Moxon.** 2000. The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase variation rate of a gene with homology to type III DNA methyltransferases. Mol. Microbiol. **35:**211–222.
11. **de Vries, N., E. J. Kuipers, N. E. Kramer, A. H. M. van Vliet, J. J. E. Bijlsma, M. Kist, S. Bereswill, C. M. J. E. Vandenbroucke-Grauls, and J. G. Kusters.** 2001. Identification of environmental stress-regulated genes in *Helicobacter pylori* by a *lacZ* reporter gene fusion system. Helicobacter **6:**300–309.
12. **Dunn, B. E., H. Cohen, and M. J. Blaser.** 1997. *Helicobacter pylori*. Clin. Microbiol. Rev. **10:**720–741.
13. **Dybvig, K., R. Sitaraman, and C. T. French.** 1998. A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. Proc. Natl. Acad. Sci. USA **95:**13923–13928.
14. **Freitag, C. S., J. M. Abraham, J. R. Clements, and B. I. Eisenstein.** 1985. Genetic analysis of the phase variation control of expression of type 1 fimbriae in *Escherichia coli*. J. Bacteriol. **162:**668–675.
15. **Goodwin, A., D. Kersulyte, G. Sisson, V. Z. Veldhuyzen, D. E. Berg, and P. S. Hoffman.** 1998. Metronidazole resistance in *Helicobacter pylori* is due to null mutations in a gene (*rdxA*) that encodes an oxygen-insensitive NADPH nitroreductase. Mol. Microbiol. **28:**383–393.
16. **Govantes, F., and E. Santero.** 1996. Transcription termination within the regulatory *nifLA* operon of *Klebsiella pneumoniae*. Mol. Gen. Genet. **250:** 447–454.
17. **Gumulak-Smith, J., A. Teachman, A. H. Tu, J. W. Simecka, J. R. Lindsey, and K. Dybvig.** 2001. Variations in the surface proteins and restriction enzyme systems of *Mycoplasma pulmonis* in the respiratory tract of infected rats. Mol. Microbiol. **40:**1037–1044.
18. **Heithoff, D. M., R. L. Sinsheimer, D. A. Low, and M. J. Mahan.** 1999. An essential role for DNA adenine methylation in bacterial virulence. Science **284:**967–970.
19. **Humbelin, M., B. Suri, D. N. Rao, D. P. Hornby, H. Eberle, T. Pripfl, S. Kenel, and T. A. Bickle.** 1988. Type III DNA restriction and modification systems *Eco*P1 and *Eco*P15. Nucleotide sequence of the *Eco*P1 operon, the *Eco*P15 mod gene and some *Eco*P1 *mod* mutants. J. Mol. Biol. **200:**23–29.
20. **Jones, A. C., R. P. 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.
21. **Josenhans, C., K. A. Eaton, T. Thevenot, and S. Suerbaum.** 2000. Switching of flagellar motility in *Helicobacter pylori* by reversible length variation of a short homopolymeric sequence repeat in *fliP*, a gene encoding a basal body protein. Infect. Immun. **68:**4598–4603.
22. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. **29:**3742–3756.
23. **Lavitola, A., C. Bucci, P. Salvatore, G. Maresca, C. B. Bruni, and P. Alifano.** 1999. Intracistronic transcription termination in polysialyltransferase gene (*siaD*) affects phase variation in *Neisseria meningitidis*. Mol. Microbiol. **33:** 119–127.
24. **Meisel, A., T. A. Bickle, D. H. Kruger, and C. Schroeder.** 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. Nature **355:**467–469.
25. **Merino, E., P. Babitzke, and C. Yanofsky.** 1995. *trp* RNA-binding attenuation protein (TRAP)-*trp* leader RNA interactions mediate translational as well as transcriptional regulation of the *Bacillus subtilis trp* operon. J. Bacteriol. **177:**6362–6370.
26. **Murray, N. E.** 2002. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology **148:**3–20.
27. **Nobusato, A., I. Uchiyama, S. Ohashi, and I. Kobayashi.** 2000. Insertion with long target duplication: a mechanism for gene mobility suggested from comparison of two related bacterial genomes. Gene **259:**99–108.
28. **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.

29. **Peck, B., M. Ortkamp, K. D. Diehl, E. Hundt, and B. Knapp.** 1999. Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*. Nucleic Acids Res. **27:**3325–3333.
30. **Perez-Martin, J., F. Rojo, and V. de Lorenzo.** 1994. Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. Microbiol. Rev. **58:**268–290.
31. **Saha, S., I. Ahmad, Y. V. Reddy, V. Krishnamurthy, and D. N. Rao.** 1998. Functional analysis of conserved motifs in type III restriction-modification enzymes. Biol. Chem. **379:**511–517.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Saunders, N. J., J. F. Peden, D. W. Hood, and E. R. Moxon.** 1998. Simple sequence repeats in the *Helicobacter pylori* genome. Mol. Microbiol. **27:** 1091–1098.
34. **Scott, T. N., and M. I. Simon.** 1982. Genetic analysis of the mechanism of the *Salmonella* phase variation site specific recombination system. Mol. Gen. Genet. **188:**313–321.
35. **Sheridan, S. D., M. L. Opel, and G. W. Hatfield.** 2001. Activation and repression of transcription initiation by a distant DNA structural transition. Mol. Microbiol. **40:**684–690.
36. **Tannaes, T., N. Dekker, G. Bukholm, J. J. Bijlsma, and B. J. Appelmelk.** 2001. Phase variation in the *Helicobacter pylori* phospholipase A gene and its role in acid adaptation. Infect. Immun. **69:**7334–7340.
37. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, et al.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539–547.
38. **van der Ende, A., C. T. Hopman, S. Zaat, B. B. Essink, B. Berkhout, and J. Dankert.** 1995. Variable expression of class 1 outer membrane protein in *Neisseria meningitidis* is caused by variation in the spacing between the −10 and −35 regions of the promoter. J. Bacteriol. **177:**2475–2480.
39. **van Doorn, N. E., F. Namavar, J. G. Kusters, E. P. van Rees, E. J. Kuipers, and J. de Graaff.** 1998. Genomic DNA fingerprinting of clinical isolates of *Helicobacter pylori* by REP-PCR and restriction fragment end-labelling. FEMS Microbiol. Lett. **160:**145–150.
40. **Vanet, A., L. Marsan, A. Labigne, and M. F. Sagot.** 2000. Inferring regulatory elements from a whole genome. An analysis of *Helicobacter pylori* sigma(80) family of promoter signals. J. Mol. Biol. **297:**335–353.
41. **Wang, G., D. A. Rasko, R. Sherburne, and D. E. Taylor.** 1999. Molecular genetic basis for the variable expression of Lewis Y antigen in *Helicobacter pylori*: analysis of the alpha (1,2) fucosyltransferase gene. Mol. Microbiol. **31:**1265–1274.
42. **Weiser, J. N., and N. Pan.** 1998. Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide. Mol. Microbiol. **30:**767–775.
43. **Yakhnin, H., J. E. Babiarz, A. V. Yakhnin, and P. Babitzke.** 2001. Expression of the *Bacillus subtilis trpEDCFBA* operon is influenced by translational coupling and Rho termination factor. J. Bacteriol. **183:**5918–5926.
44. **Yamaoka, Y., D. H. Kwon, and D. Y. Graham.** 2000. A M_r 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. Proc. Natl. Acad. Sci. USA **97:**7533–7538.
45. **Zhao, H., X. Li, D. E. Johnson, I. Blomfield, and H. L. Mobley.** 1997. In vivo phase variation of MR/P fimbrial gene expression in *Proteus mirabilis* infecting the urinary tract. Mol. Microbiol. **23:**1009–1019.