THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 26, Issue of June 30, pp. 20077–20083, 2000

Printed in U.S.A.

A Novel 3-Methyladenine DNA Glycosylase from *Helicobacter pylori* Defines a New Class within the Endonuclease III Family of Base Excision Repair Glycosylases*

Received for publication, February 9, 2000, and in revised form, April 11, 2000
Published, JBC Papers in Press, April 20, 2000, DOI 10.1074/jbc.M001071200

Eyleen J. O'Rourke‡§¶, Catherine Chevalier‖, Serge Boiteux‡, Agnès Labigne‖, Luis Ielpi§‡‡, and J. Pablo Radicella‡§§**

From the ‡ Département de Radiobiologie et Radiopathologie, Commissariat à l'Energie Atomique, UMR217 CEA/CNRS, BP6, 92265 Fontenay-aux-Roses, France, the §Instituto de Investigaciones Bioquímicas Fundación Campomar, Facultad de Ciencias Exactas y Naturales-UBA and the Consejo Nacional de Investigaciones Científicas y Técnicas, 1405 Buenos Aires, Argentina, and the ‖Unité de Pathogénie Bactérienne des Muqueuses, Institut Pasteur, 75724 Paris, France

The cloning, purification, and characterization of MagIII, a 3-methyladenine DNA glycosylase from *Helicobacter pylori*, is presented in this paper. Sequence analysis of the genome of this pathogen failed to identify open reading frames potentially coding for proteins with a 3-methyladenine DNA glycosylase activity. The putative product of the HP602 open reading frame, reported as an endonuclease III, shares extensive amino acid sequence homology with some bacterial members of this family and has the canonic active site helix-hairpin-helix-GPD motif. Surprisingly, this predicted *H. pylori* endonuclease III encodes a 25,220-Da protein able to release 3-methyladenine, but not oxidized bases, from modified DNA. MagIII has no abasic site lyase activity and displays the substrate specificity of the 3-methyladenine-DNA glycosylase type I of *Escherichia coli* (Tag) because it is not able to recognize 7-methylguanine or hypoxanthine as substrates. The expression of the *magIII* open reading frame in null 3-methyladenine glycosylase *E. coli* (*tag alkA*) restores to this mutant partial resistance to alkylating agents. MagIII-deficient *H. pylori* cells show an alkylation-sensitive phenotype. *H. pylori* wild type cells exposed to alkylating agents present an adaptive response by inducing the expression of *magIII*. MagIII is thus a novel bacterial member of the endonuclease III family, which displays biochemical properties not described for any of the members of this group until now.

Living organisms have developed a variety of strategies for protecting their genetic information. DNA, the repository of this information in all cells, is constantly exposed to chemical challenges (1). A variety of reactive intracellular and environmental compounds can induce modifications in the DNA bases. If left uncorrected, these lesions may cause mutations or impair DNA replication. In the case of alkylating agents, such as the cellular methyl donor *S*-adenosylmethionine or genotoxic compounds such as dimethylsulfate (DMS),[1] methylmethanesulfonate (MMS), or *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG), they can react nonenzymatically with DNA to produce 7-methylguanine (7-MeG) and 3-methyladenine (3-MeA) among other alkylated bases (2). In particular, 3-MeA posses a serious problem to the cell because it blocks DNA synthesis (3–6). To eliminate this kind of modified base, most cells possess monofunctional DNA glycosylases that specifically recognize and excise the methylated bases from DNA by cleavage of the C1′-N glycosylic bond. This first step of base excision repair is followed by the repair of the resulting abasic (AP) site by excision/replacement DNA synthesis.

At least four different classes of 3-MeA glycosylases have been described, based on their sequence and substrate specificities (7). All of these are low molecular weight enzymes, and no associated AP lyase activity has been attributed to any of them. In *Escherichia coli*, the two first classes are represented. Class I is defined by the product of the *tag* gene that codes for a constitutive 3-MeA DNA glycosylase (8, 9). Tag is quite specific for the removal of 3-MeA, although 3-MeG would also be excised with a much lower efficiency *in vitro* (10). Class II is typified by the AlkA enzyme of *E. coli*. There is no amino acid sequence similarity between class I and II 3-MeA DNA glycosylases. AlkA has a much broader substrate specificity, being able to release not only 3-MeA but also 7-MeG, 3-MeG, 7-MeA, and O^2-methylpyrimidines (11). In the last few years, it has been found that AlkA has the ability to act on a variety of modified bases other than the classical alkylation damaged. Among those, the *alkA* product can release oxidized bases from DNA (12), as well as deamination products of adenine such as hypoxanthine (13) or ethenopurines (14, 15). Berdal *et al*. (16) have shown that AlkA is even able to release normal bases from DNA. Another difference with Tag resides in the fact that *alkA* is induced by treatment of the cells with alkylating agents (7). This induction occurs through activation of the *alkA* promoter

* This work was supported in France by the Commissariat à l'Energie Atomique, CNRS, and Electricite de France and in Argentina by a fellowship from Fundación Antorchas (to E. J. O.), grants from the Laboratorios Bagó SA, and Argentine Federal Government Grant AN-PCT PICT97-00419) (to L. I.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Recipient of a doctoral fellowship from the University of Buenos Aires (Fondo Para el Mejoramiento de la Calidad Universitaria).

** Member of Unité INSERM U389.

‡‡ Research Career Investigator of the Consejo Nacional de Investigaciones Científicas y Técnicas.

§§ To whom correspondence should be addressed. Tel.: 33-1-46-54-88-57; Fax: 33-1-46-54-88-59; E-mail: jpradicella@cea.fr.

[1] The abbreviations used are: DMS, dimethylsulfate; MMS, methylmethanesulfonate; MNNG, *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine; MeG, methylguanine; MeA, methyladenine; AP, abasic; Hx, hypoxanthine; HhH, helix-hairpin-helix; EndoIII, endonuclease III; ORF, open reading frame; IPTG, isopropyl-β-D-thiogalactopyranoside; PAGE, polyacrylamide gel electrophoresis; HPLC, high pressure liquid chromatography; GST, glutathione *S*-transferase; PCR, polymerase chain reaction.

TABLE I
Bacterial strains used in this work

Strain	Relevant genotype	Source or reference
E. coli		
BL21	*F ompT hsdS$_B$($r_B^-m_B^-$) gal dcm*	Novagen
GC4801	AB 1157, X::Tn5	4
GC4803	AB 1157, X::Tn5, *alkA1, tagA1*	4
H. pylori		
13/5	Wild type	Laboratory collection
ADM1	Wild type	Laboratory collection
HAS141	Wild type	Laboratory collection
HAS141 *magIII*	*magIII*	This work
26695	Wild type	24

by the Ada protein. Class III is typified by the mammalian methylpurine DNA glycosylases. These enzymes are biochemically quite similar to AlkA but differ markedly in their primary sequence. The recent determination of the crystal structures for the human 3-MeA glycosylase AAG (17) and the bacterial AlkA (18, 19) confirmed that class II and class III enzymes are structurally unrelated, despite the fact that they share a broad substrate specificity. Indeed, for its structure, AlkA belongs to the superfamily of DNA glycosylases, the archetype of which is the EndoIII from *E. coli,* an *N*-glycosylase that excises oxidized pyrimidines from DNA (18). This structural correspondence is striking in view of the limited homology between the members of this family, restricted to the helix-hairpin-helix (HhH) motif present opposite a catalytic aspartic acid in the active site. A fourth class of 3-MeA glycosylases has been proposed after the identification of the MpgII enzyme from *Thermotoga maritima* (20). This enzyme, as in the case of AlkA, belongs to the superfamily of the EndoIII, as defined by the HhH aspartic motif (21, 22), is able to release not only 3-MeA but also 7-MeG from *N*-methyl-*N*-nitrosourea-treated DNA; and lacks an AP endonuclease activity. Despite those similarities, the lack of sequence homology to AlkA, other than the HhH aspartic motif, has prompted the authors to define the class IV of 3-MeA DNA glycosylases (20).

In our search for new members of the EndoIII family in *Helicobacter pylori,* we identified an open reading frame (ORF), HP602, that although labeled as EndoIII (23, 24), codes for an enzyme that does not recognize oxidized bases or AP sites as substrates. The lack of potential methylpurine DNA glycosylase-encoding genes in the genomic sequence of this pathogen, together with the fact that the protein coded by this ORF has the HhH aspartic motif but lacks the lysine at the active site and the canonical 4Fe-4S cluster characteristic of endonuclease III, prompted us to investigate the possibility that ORF HP602 could be the gene for a methylpurine DNA glycosylase. We present here data showing that the enzyme coded by ORF HP602 is a functional 3-MeA DNA glycosylase, MagIII, inducible at the transcriptional level by MNNG. During the preparation of this manuscript and based on its sequence, the product of this ORF was attributed to the MpgII group (20). Although it presents predicted structural similarities with members of the HhH aspartic motif superfamily also displayed by AlkA and has an overall sequence similarity to the members of the MpgII class, this 3-MeA DNA glycosylase has a more restricted substrate specificity, comparable to the one characteristic of Tag. These characteristics suggest that this enzyme from *H. pylori* defines a novel class of 3-MeA DNA glycosylases.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Media, and Plasmids—-Strains used in this study are listed in Table I. *E. coli* strains were grown in Luria-Bertani (LB) broth or agar supplemented with ampicillin (200 μg/ml) or kanamycin (30 μg/ml); Minimal A salts were prepared as described by Miller (25). *H. pylori* strains were grown on blood brucella agar plates at 37 °C in a 5% CO_2 and 95% humidity atmosphere. Disruption mutant *H. pylori* strains were cultured with 20 μg/ml kanamycin. pGEX-4T-1 and pET29a(+) were obtained from Amersham Pharmacia Biotech and Novagen, respectively.

Materials and Reference Compounds—All chemicals were, unless otherwise stated, from Sigma. [^{32}P]ATP (3000Ci/mmol), [^{3}H]dimethyl sulfate (3.8 Ci/mmol), and Nensorb-20 nucleic acid purification cartridges were from New England Nuclear. T4 polynucleotide kinase, T4 DNA ligase, and restriction enzymes were purchased from New England Biolabs.

pGHP602-I or pGHP602-V Construction—The *magIII* gene with its own start codon was synthetized by PCR from two independent *H. pylori* isolates (13/5 and ADM1 strains) by using oligonucleotide primers deduced from the HP602 ORF. The amplified DNA was cloned in pGEX-4T-1 previously digested with restriction enzymes *Bam*HI and *Pst*I. The resulting plasmids, termed pGHP602-I and pGHP602-V, were expressed in *E. coli* as a fusion protein consisting of MagIII with glutathione *S*-transferase (GST) at its N terminus. In each case, the nucleotide sequences of the cloned genes were confirmed by dideoxynucleotide sequencing.

Sensitivity to Methylmethane Sulfonate—Plasmids pGHP602-I and pGHP602-V and the vector were independently introduced by electroporation into *E. coli* GC4803 (4). Overexpression of the proteins coded by the plasmids after induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was confirmed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Survival curves were made with individual transformants. Independent transformants and the wild type strain GC4801 were incubated for 3 h with IPTG to allow gene expression. They were subsequently challenged with 0, 5, 10, 25, or 50 mM MMS in minimal A salts for 20 min at 37 °C and spread, after appropriate dilution, on LB plates to calculate cell viability. For *H. pylori* strains, evaluation of the resistance to the cytotoxic effect of the MMS was conducted by determining the survival rate of the wild type and mutant strains following a 20-min exposure in liquid to different concentrations of MMS (10–100 mM), spreading on blood agar plates, and colony counting.

Overproduction and Purification of 3-Methyladenine DNA Glycosylase III—Both alleles of *magIII* gene (GIG and GVG) were subcloned from pGHP602-I or pGHP602-V into pET29a+ (Novagen, Inc.) using *Bam*HI and *Pst*I restriction sites. The resulting plasmids were named pSHP602-I and pSHP602-V, respectively. *E. coli* BL21 cells were electrotransformed with plasmids pSHP602I, pSHP602V, and pET29a+. The expression of the S-tag-MagIII fusions in *E. coli* was confirmed by S-tag Western blot according to the manufacturer's instructions. Cells were used to inoculate 500 ml of LB medium supplemented with kanamycin and were incubated at 37 °C with shacking until the A_{600} reached 1. IPTG was added to a final concentration of 1 mM, and growth was continued overnight at 18 °C with shaking. Cells were harvested by centrifugation at 4000 × *g,* and MagIII was purified following the instructions of the S-tag thrombin kit (Novagen, Inc.) except that 50 mM KCl, 1 mM dithiothreitol, and 5% glycerol were added to all buffers. The thrombin eluted fractions were concentrated by Centricon-30 (Amicon) and loaded onto a fast protein liquid chromatography HR10/30 Superdex 75 gel filtration/size exclusion column (Amersham Pharmacia Biotech) equilibrated with 50 mM HEPES, pH 7.5, 100 mM KCl, 5 mM dithiothreitol, and 5% glycerol. The column was eluted with 30 ml of the same buffer (0.7 ml/min), and 500-μl fractions were collected. Protein elution profile was followed with a UV monitor. Under these conditions, purified MagIII eluted at 12 min. Protein concentrations were determined by the Bradford protein assay method (26).

DNA Substrates—The 34-mer oligodeoxyribonucleotides used in this study have the following sequence: 5′-GGCTTCATCGTTGTC*X*CAGACCTGGTGGATACCG-3′, where *X* represents uracil, hypoxanthine (Hx),

or dihyrothymine residue. Except for the latter, a kind gift from Dr. J. Cadet (Commissariat à l'Energie Atomique, Grenoble, France), these oligonucleotides and their complements with each of the four bases opposite the lesion in the duplex were purchased from Oligo Express (Grenoble, France). To generate AP sites, the uracil-containing duplex was incubated in the presence of uracil DNA glycosylase (27). The [^{3}H]DMS-poly(dG-dC)·poly(dG-dC) and [H^3]DMS-DNA were prepared as described (28, 29).

Enzymatic Activity Assays—For the cleavage of lesion-containing DNA duplexes, the oligonucleotide carrying the modified base was ^{32}P-labeled at the 5′-end and annealed to its complementary strand as described (27). In a standard reaction (final volume, 10 μl), 50 fmol of labeled duplex were incubated in the reaction buffer (25 mM Tris-HCl, pH 7.6, 2 mM Na_2EDTA, 50 mM NaCl) with the indicated protein fraction at 37 °C. For Hx- and dihydrothymine-containing oligos, reactions were stopped by the addition of 0.2 N NaOH. In the case of oligos harboring an AP site, 10 mM $NaBH_4$ was added to stabilize the substrate. After addition of 6 μl of formamide dye, the products were separated by 7 M urea 20% PAGE. Gels were analyzed by autoradiography.

3-Methylpurine DNA glycosylase activities were determined by the method of Karran *et al.* (30). The reaction mixture (50 μl) contained 50 mM HEPES-KOH (pH 7.5), 5 mM dithiothreitol, 1 mM Na_2EDTA, 100 mM KCl, and 2000 or 15,000 cpm of [^{3}H]DMS-DNA or [^{3}H]DMS-poly(dG-dC)·poly(dG-dC) (2000 cpm/pmol). The reaction mixture was incubated at 37 °C for 30 min, and then 75 μl of stop solution containing 1 mg/ml bovine serum albumin, 1 mg/ml salmon testes DNA, and 0.5 M NaCl were added, followed by the addition of 500 μl of cold ethanol. The reactions were kept at −20 °C for 15 min and centrifuged, and the supernatants were taken (ethanol soluble fraction). The radioactivity in the soluble fraction was determined in a liquid scintillation counter. To analyze the products, the reaction mixtures were separated using high pressure liquid chromatography (HPLC) (4, 6). The internal nonradioactive markers 3-MeA and 7-MeG were added to the samples and detected by their UV absorbance. The radioactivity in the fractions was determined in a liquid scintillation counter.

Construction of an H. pylori Mutant Deficient in 3-Methyladenine DNA Glycosylase—The HP602 ORF was amplified by PCR from the genomic DNA from strain 26695 (24) using the forward 5′-CAUCAUCAUTGTGTTGGATAGTTTTGAGAT (602-F) and the reverse 5′-CUACUACUAAAATCAAAGTTTTAATTCCAA (602-R) oligonucleotides. The 650-base pair DNA product was cloned in the pILL570 derivative plasmid (3.6 kilobase) following a treatment with Uracil DNA Glycosylase (Roche Molecular Biochemicals), producing 3′-protruding ends on both the amplicon and the vector (31). The cloned ORF was disrupted in *E. coli* by insertion of a transposable element (MiniTn*3*-Km) (32) into the recombinant plasmid. Twenty-four independent disruptions of the pILL570-HP602 plasmid were pooled and purified using MIDI columns (Qiagen). The pooled plasmids were introduced by natural transformation (33) into *H. pylori* strain HAS141 (34). Chromosomal DNA from six individual kanamycin resistant transformants were purified (QIAamp tissue kit, Qiagen) and used as template for PCR using the 602-F primer together with Km-1 (5′-CGGTATAATCTTACCTATCACCTCA) or Km-2 (5′-TTTGACTTACTGGGGATCAAGCCTG), two divergent primers flanking the kanamycin resistance gene of the kanamycin cassette (35). Each PCR fragment was directly sequenced with primer Km-3 (5′-GATCCTTTTTGATAATCTCATGACC, complementary to the end of the MiniTn*3*-Km), allowing us to determine the orientation of the transposon as well as the precise site of insertion of the transposable element within the HP602 ORF. One mutant with an insertion at nucleotide +272 of the HP602 ORF was further characterized (mutant designated through the manuscript as HAS141 *magIII*).

Semiquantitative Reverse Transcription-PCR—Two-day-old *H. pylori* plates were independently challenged with 50 μg/ml MNNG, 5 μg/ml MMS, or 0.5 μg/ml DMS. After 0 min, 20 min, 1 h, 2 h, or 4 h, the cells were harvested, rinsed twice, and frozen in liquid nitrogen. The cell pellets were stored at −70 °C. RNA extraction from frozen *H. pylori* HAS141 was performed with guanidinium thiocyanate-phenol chloroform (Trizol, Life Technologies, Inc.). Total RNA was dissolved in 20 μl of diethyl pyrocarbonate-treated water and incubated at 37 °C for 30 min with 1 unit/μl RNase-free DNase (Promega). DNA absence was confirmed by PCR. cDNA was synthetized with 2 pmol of reverse specific primers for 16 S RNA and *magIII*, added to 1 μg of each RNA sample. The mixtures were incubated at 70 °C for 10 min, placed on ice for 2 min, and incubated with Superscript II reverse transcriptase (Life Technologies, Inc.) according to the manufacturer's recommendations. Reverse primers were as follows: for 16 S RNA, 5′-CTGGAGAGACTAAGCCCTCC-3′; for *magIII*, 5′-TTTTTTCTGCAGAAGTTTTAAT-3′. Relative quantification of mRNA levels was performed in the exponential phase of amplification as described previously (36). Briefly, PCR amplification of target template and of 16 S RNA as external control were run in parallel. DNA amplification was carried out in 1× PCR buffer supplemented with 200 μM dNTPs, 2 μM each of 5′- and 3′-specific primers, 1 unit of *Taq* polymerase (Roche Molecular Biochemicals) in a final volume of 50 μl. The exponential phase of amplification was determined by carrying out the reaction for 16–38 cycles using a fixed quantity of cDNA. The number of cycles was set for *magIII* to 26 sequential cycles at 95 °C for 30 s, 48 °C for 45 s, 72 °C for 40 s, and for 16 S RNA to 20 cycles at 95 °C for 30 s and 60 °C for 30 s. In this phase of amplification, the amounts of the amplicons are proportional to the initial amounts of templates. The cDNA was serially diluted in water from 1:5 (10 ng of RNA) to 1:15625 (3.2 pg of RNA). 20 μl of *magIII* and 16 S RNA PCR products amplified in parallel from the same serial dilutions were loaded on agarose gel (1% agarose, 0.05% ethidium bromide). A negative control in which cDNA was omitted was run in parallel with each experiment.

RESULTS

Identification of a New Member of the Endonuclease III Glycosylase Family in H. pylori—The availability of the complete genomic sequence of *H. pylori* allowed us to identify several putative members of the EndoIII family of DNA glycosylases involved in the repair of modified bases. Two open reading frames were annotated as homologs of the EndoIII from *E. coli,* HP602 and HP585 (24). To analyze whether these two ORFs coded for *bona fide* homologs of the EndoIII and whether they corresponded to redundant activities, the putative coding sequences from two independently isolated strains (13/5 and ADM1) were amplified by PCR and cloned in *E. coli* vectors that allow their expression as fusions to the C terminus of the GST protein. For HP585 the purification of the GST fusion protein confirmed that this ORF codes for an active EndoIII-like enzyme (data not shown). Sequencing of the plasmids harboring the HP602 ORFs established that the sequences differed between the two isolates (Fig. 1). Comparison of the coded proteins and those of the two strains of which the genomes are completely sequenced shows that for any pair, there are four or five amino acid differences among the 218 residues. In particular, for the two newly sequenced ORFs, a polymorphism was found, GIG *versus* GVG, at what has been determined to be the active site for the members of this family (22) (HhH motif, Fig. 1). As a consequence, the analysis of both clones (herein denominated 602-I and 602-V) was carried out in parallel. HP602 also presented regions (Fig. 1) structurally homologous to the EndoIII. The three-dimensional PSSM web server was searched for HP602 structurally related proteins. The strongest structural similarities were found with respect to EndoIII (E value, 2.7×10^{-3}) and AlkA (E value, 2.8×10^{-2}) crystals (37, 38). In the predicted secondary structure of the HP602 product it is possible to define, with a high level of confidence (≥50%) the helices αC to αM of EndoIII and AlkA, which constitute the regions at which fold and topology are shared by both enzymes (18) (Fig. 1). Moreover, identical amino acids are clustered over these conserved secondary structures. However, HP602 product lacks the pattern of cysteines giving rise to the iron-sulfur cluster present in EndoIII and MpgII proteins. No activity was detected on DNA substrates containing dihydrothymine for the product of the GST fused to either of the HP602 ORF variants (data not shown). Furthermore, neither of these constructs was able to complement the mutator phenotype of an *E. coli* strain deficient in both EndoIII and EndoVIII DNA glycosylases (data not shown). We therefore examined whether these ORFs could code for a function related to the other member of the EndoIII family, AlkA, which also shares a similar overall folding pattern with EndoIII, as well as the HhH aspartic motif (Fig. 1).

Complementation of an E. coli alkA tag Mutant—To examine

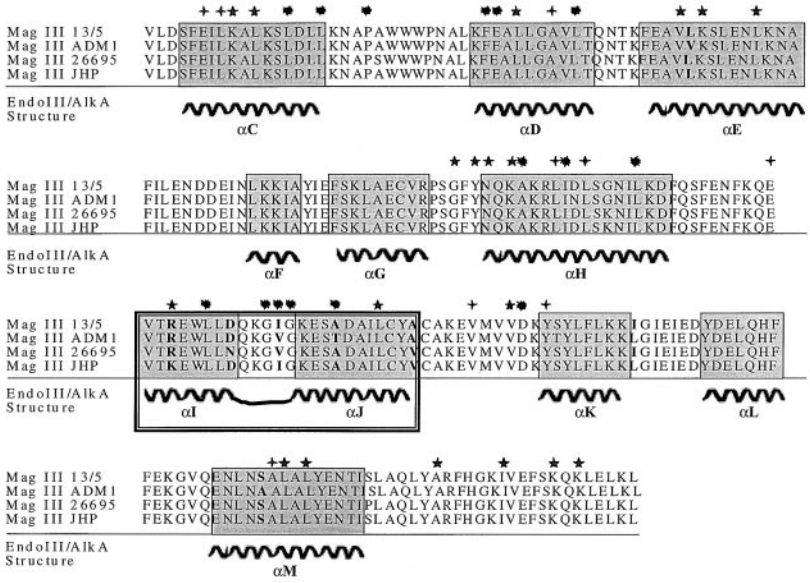

FIG. 1. **MagIII (HP602) sequence analysis.** The alignment of the HP602 coded protein from four different isolates is presented. 13/5 and ADM1 were obtained in this work, and *26695* and *JHP* correspond to the isolates of which the genomes were completely sequenced (23, 24). The predicted secondary structures shared with the crystallographic data obtained for EndoIII and AlkA are shown beneath the alignment; residues involved in these structures are *gray-shaded*. The putative HhH motif is *double-boxed*. Amino acids marked with a *star* are identical to EndoIII. Positions marked with a *diamond* are identical to AlkA. Amino acids marked with an *asterisk* are identical to both enzymes. *Boldface letters* highlight intraspecies changes. These data were generated by the three-dimensional PSSM web server, by the Biomolecular Modelling Laboratory, Imperial Cancer Research Fund (38), and according to the Labahn *et al*. (18) EndoIII/AlkA comparison.

whether *H. pylori* HP602 can substitute for *E. coli* AlkA and/or Tag in the repair of damaged DNA, HP602/GIG or HP602/GVG fused to GST were expressed in *E. coli* GC4803 (4), a strain phenotypically sensitive to alkylating agents due to null mutations in *tag* and *alkA*. To confirm that the proteins were expressed, SDS-PAGE analyses of crude *E. coli* extracts were performed (data not shown). Independent transformants for each plasmid and the isogenic wild type strain GC4801 were incubated for 3 h with IPTG to allow gene expression and were subsequently challenged with 0, 5, 10, 25, or 50 mM MMS before spreading on LB plates to calculate the survival fraction. Fig. 2 shows that both HP602-I and HP602-V products are able to partially restore the wild type resistance to MMS in the *alkA tag* double mutant. Although it represents only the 10–15% of the DNA alkylation products, 3-MeA is considered to be the major cause for the cytotoxicity induced by alkylating agents (7). This result indicates that ORF HP602 is likely to code for a protein with a 3-MeA DNA glycosylase activity.

Substrate Specificity of HP602 Gene Product—To further investigate the characteristics of these DNA glycosylases, the crude extracts of *E. coli* GC4803 harboring the pGHP602-I or pGHP602-V plasmid or the vector were assayed for their ability to liberate tritiated bases from [^{3}H]dimethyl sulfate-treated DNA or poly(dG-dC)·poly(dG-dC) (Table II). Both enzymes were able to release methylpurines from [^{3}H]DMS-DNA but not from [^{3}H]DMS-poly(dG-dC)·poly(dG-dC). This difference in the substrate specificity suggested that MagIII could not recognize 7-MeG, the major alkylation product. We therefore named the corresponding gene *magIII* for 3-MeA DNA glycosylase.

In order to unambiguously characterize the substrate specificities, HP602-I and HP602-V were subcloned into pET29a+ (Novagen, Inc.) to yield pSHP602-I and pSHP602-V, respectively. The plasmids encode the fusion proteins S·tag-MagIII(I) and S·tag-MagIII(V), respectively. This allowed the isolation of the fusion peptide and subsequent cleavage and purification of the native protein. The SDS-PAGE protein profile, after induction with 1 mM IPTG, showed a prominent band of about 26.5 kDa that was present in cells carrying pSHP602-I and pSHP602-V but absent when pET29a+ was present in the same strain. After purification of the fusion protein, the thrombin-eluted fractions were concentrated and loaded onto a Superdex 75 column. Both MagIII proteins were purified to yield

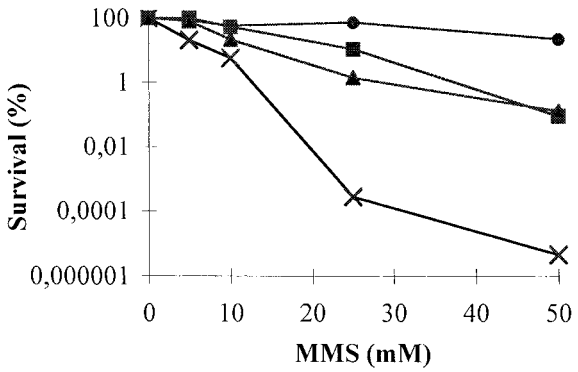

FIG. 2. **Sensitivity to MMS exposure of a *tag alkA E. coli* mutant strain expressing ORF HP602.** Survival curves for *E. coli* GC4803 (*tag alkA*) harboring pGEX-4T1 (×), pGHP602-V (■), or pGHP602-I (▲) and the parental wild type strain (●) incubated for 20 min with the indicated amounts of MMS in liquid medium.

TABLE II
Methylpurine DNA glycosylase activity in crude lysates of E. coli 4803 (alkA tag) harboring pGEX-4T1, pGHP602-I, or pGHP602-V plasmids or purified AlkA

Enzymatic fraction	DNA	
	[^{3}H]DMS-polydG-dC	[^{3}H]DMS-treated
	cpm liberated / μg protein	
E. coli 4803 pGEX-4T1	29	0
E. coli 4803 pGHP6021	45	257[a]
E. coli 4803 pGHP602V	41	243[a]
AlkA	1169	1758

[a] This value corresponds to the total content of 3-MeA available as substrate.

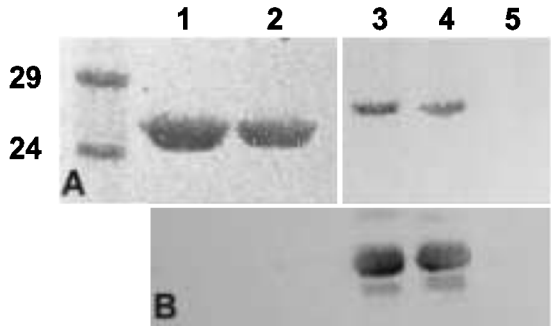

FIG. 3. **SDS-PAGE and Western blot analysis of S-tag and pure MagIII.** A, red Ponceau-stained polyvinylidene fluoride membrane (of a 10% SDS-PAGE). Molecular size markers are indicated. *Lane 1,* 15 μg of MagIII/GIG; *lane 2,* 15 μg of MagIII/GVG; *lane 3,* induced BL21 pGHP602-I; *lane 4,* induced BL21 pGHP602-V; *lane 5,* induced BL21 pET29a+. *B,* S-tag-directed Western blot of the membrane shown in *A*.

a single Coomassie Blue-stained band migrating in SDS-PAGE with a molecular mass of 25 kDa, which corresponds to the full-length protein coded by the ORF HP602 (Fig. 3). The capacities of both purified forms of MagIII to release specific modified bases from [^{3}H]DMS-DNA were examined using HPLC. Fig. 4 shows that MagIII is able to release 3-MeA but not 7-MeG from alkylated DNA, which is consistent with the lack of activity on [^{3}H]DMS-poly(dG-dC)·poly(dG-dC). This was also apparent after 4 h of reaction. Under conditions where the amounts of enzyme were normalized for their capacity to excise 3-MeA with respect to AlkA, no excision of 7-MeG was detected.

Because AlkA can also release Hx from DNA (13), a duplex oligonucleotide harboring a unique Hx residue at a defined position was challenged with the purified MagIII and subsequently treated with 0.2 N NaOH to reveal the AP site resulting from the glycosylase activity. Fig. 5*A* shows that the Hx residue is not released by MagIII at detectable rates from the modified DNA. Similar experiments were carried out with oligonucleotides harboring oxidized bases as dihydrothymine or 8-oxoguanine, and no glycosylase activity was detected for these substrates (data not shown). To confirm the sequence prediction that MagIII does not have an AP lyase activity, a ^{32}P 5′-end-labeled oligonucleotide with a single abasic site opposite a cytosine was used as a substrate. No detectable cleavage of the oligonucleotide was observed after incubation with either of the MagIII variants (Fig. 5*B*).

Disruption of the magIII from H. pylori—As mentioned above, the examination of the complete genomic sequences of *H. pylori* fails to reveal ORFs potentially coding for either Tag-like or AlkA-like enzymes. To analyze whether the MagIII is indeed functional *in vivo* and confers resistance to alkylating agents, a disruption mutant was generated in *H. pylori* by insertion of a kanamycin transposon into the ORF HP602 of strain HAS141. The disruption of this gene was confirmed by sequencing of the genomic DNA. HAS141 and HAS141*magIII* were then treated with increasing concentrations of MMS to study the cytotoxic effect of the alkylating agent. Fig. 6 shows survival reduction of the mutant strain by concentrations of MMS that do not affect the survival of the parental strain. This result demonstrates that in *H. pylori,* MagIII is the main defense against the presence of 3-MeA generated by alkylating agents on DNA.

Adaptive Response of H. pylori to Alkylating Agents—The possibility of an adaptive response to the challenge by alkylating agents was explored by analyzing the levels of *magIII* transcripts in wild type *H. pylori* after treatment with sublethal doses of alkylating agents as DMS, MMS, or MNNG. As shown in Fig. 7, semiquantitative reverse transcription-PCR experiments show that exposure to 50 μg/ml MNNG induced at

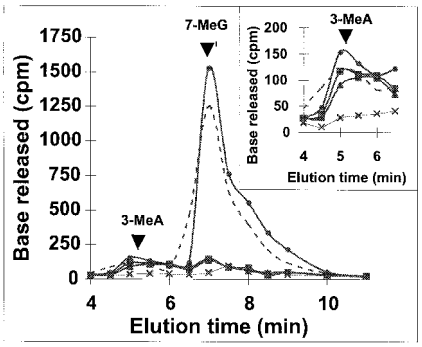

FIG. 4. **Reverse phase HPLC of methylated bases released from [^{3}H]dimethyl sulfate-treated DNA.** Incubations with MagIII/GIG (▲), MagIII/GVG (■), AlkA (●), or denatured enzyme (×) were carried out for 30 min at 37 °C. Radioactivity measured in ethanol supernatants from the heat denaturing of the substrate is indicated by the *dotted line*.

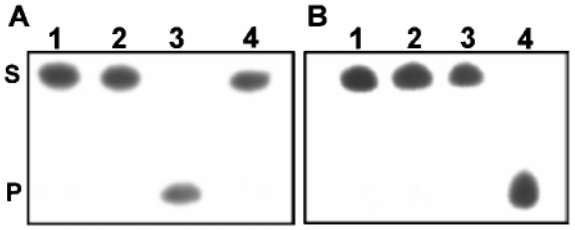

FIG. 5. **Enzymatic activities on oligonucleotide substrates.** *A,* repair of hypoxanthine by MagIII. *Lane 1,* 10 μg of MagIII/GIG; *lane 2,* 10 μg of MagIII/GVG; *lane 3,* 1 μg of pure AlkA; *lane 4,* reaction carried out with denatured MagIII as negative control. *B,* AP lyase activity of MagIII. *Lane 1,* 5 μg of MagIII/GIG; *lane 2,* 5 μg of MagIII/GVG; *lane 3,* reaction carried out with denatured MagIII as negative control; *lane 4,* 2 μg of EndoIII. *S* and *P* indicate substrate and product, respectively.

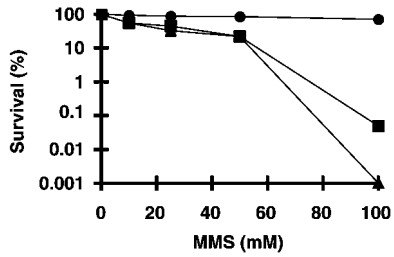

FIG. 6. **Sensitivity of *H. pylori* to MMS.** Survival curves for *H. pylori* HAS141/*magIII* (clone 5) (▲), HAS141/*magIII* (clone 7) (■), and the parental wild type strain (●) incubated for 20 min with increased amounts of MMS in liquid medium. Data are representative of two independent experiments.

least a 5-fold increase of the levels of *magIII* transcripts. Experiments using 1:4 serial dilutions revealed that the increase of *magIII* mRNA induced by exposure to MNNG is between 16- and 24-fold. There was no induction of *magIII* mRNA after treatment of the cells with either DMS or MMS.

DISCUSSION

H. pylori is a Gram-negative bacterium that is associated with several human pathologies (39). Indeed, gastric inflammation and peptic ulcer (40, 41), as well as predisposition to gastric cancer and non-Hodgkin lymphomas of the stomach (42, 43), have been linked to infection by this pathogen. Almost half of the population of the world harbors an *H. pylori* infection. There is considerable evidence that the bacterial genotype is an important factor in determining the nature and severity of the disease induced by the infection (44, 45). The availability of the complete genomic sequences of two strains (23, 24) allows not only the comparison of the genomes to other organisms sequences but also a preview of the intraspecies variations in

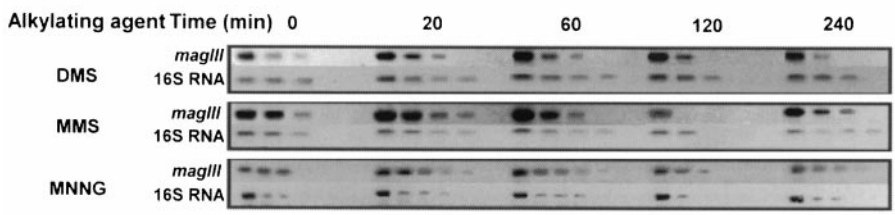

FIG. 7. **Semiquantitative reverse transcription-PCR analysis of *magIII* expression in *H. pylori* exposed to alkylating agents.** The cells were challenged with 0.5 μg/ml DMS, 5 μg/ml MMS, or 50 μg/ml MNNG for 0, 20, 60, 120, or 240 min. Lanes for each time point correspond to 1:5 serial dilutiions. *MagIII lanes* correspond to 50, 10, 2, 0.4, and 0.08 ng of cDNA template for all samples. *16S RNA lanes* show the PCR products obtained from 10, 2, 0.4, 0.08, and 0.016 ng for DMS and MMS challenge. An extra dilution (3.2 pg) for the MNNG experiment is presented. Gels are representative of two independent experiments.

their coding sequences. As an example, Fig. 1 shows that in the *magIII* gene, there are single amino acid variations between isolates at seven positions. The genetic variability of *H. pylori* has been the focus of many investigations (39, 46). The study of natural isolates suggests that the genetic diversity of *H. pylori* exceeds the level of diversity recorded in virtually all other bacterial species studied. It has been concluded that the presence of ORFs with the capacity to code for DNA repair enzymes preclude point mutations as a possible cause of the genetic diversity. However, the inference of a gene function by sequence homology can lead to misinterpretation. This is the case especially when the homologies invoked are not to proteins for which experimental evidence of their function exists but rather to proteins for which the function has been inferred from their homology to other sequences. Moreover, in the case of *H. pylori,* the presence of some genes involved in particular DNA repair pathways has been identified without finding the other functions for the same pathway. In the case of the mismatch repair system, a MutS homolog has been predicted, but no ORF encoding a MutL-like protein has been identified. From the genome analysis, several enzymes from the base excision repair pathway also seem to be missing; there are no Fpg homologs (although there is a putative MutY coding sequence), and no ORF has been predicted as coding for the repair of alkylated bases by base excision repair. The description in this paper of a 3-MeA DNA glycosylase from *H. pylori* underscores the limitations of negative predictions about the presence of functions based on sequence comparisons.

MagIII was originally attributed an endonuclease III-like function because of the presence of the GI/VG motif and the aspartic acid at position 150 that are the signature of the superfamily of DNA glycosylases resembling EndoIII (22, 47, 48). However, it lacks the 4Fe-4S cluster characteristic of all the EndoIII proteins described. It is shown here that MagIII is not a functional EndoIII because it does not recognize oxidized pyrimidines as substrates and does not have an AP lyase activity. We have established that MagIII is, however, a monofunctional 3-MeA glycosylase belonging to the EndoIII family of DNA glycosylases. Although the *E. coli* class II 3-MeA glycosylase AlkA is also a member of this superfamily, no significant primary sequence similarity is found between these proteins. Moreover, they differ significantly in their substrate specificities. While this work was in progress, Begley *et al.* (20) described yet another 3-MeA DNA glycosylase, MpgII from *T. maritima,* defining what they called class IV. They also suggested that the putative protein coded by ORF HP602 (MagIII) from *H. pylori* could belong to class IV of the 3-MeA DNA glycosylases by sequence similarity (28% identity). The most conserved structure is the HhH (G/P)$X_{18\text{-}21}$D motif, the presence of which suggests that MagIII may use a common mechanism of base excision by nucleophilic activation. However, MagIII shows an important difference in lacking the iron-sulfur cluster, which has been implicated in DNA phosphate backbone recognition. At the sequence level, the 28% identity is not much higher than the level of identity between MpgII and MutY or EndoIII (25 or 20%, respectively). Indeed, MpgII shares a greater homology with other members of its subclass, such as the enzyme from *Methanococcus jannaschi* (43% identity), that lack the iron-sulfur cluster. MagIII substrate specificity is also different. Unlike MpgII and AlkA, MagIII does not excise 7-MeG. Moreover MagIII does not recognize Hx as substrate. Taking into account the alkylated lesions tested, MagIII seems to recognize almost exclusively 3-MeA, defining then a substrate specificity closer to the class I 3-MeA glycosylases such as Tag (7, 49). We therefore suggest that MagIII defines a novel class of 3-MeA glycosylases. It will be interesting to study whether the other putative MpgII described by Begley *et al.* (20) are MpgII-like or MagIII-like enzymes. This will shed light on the sequence features responsible for the substrate specificities.

At the physiological level, *magIII,* as *alkA* in *E. coli,* is transcriptionally induced by exposure of the cells with MNNG. This represents an enhanced protection against the cytotoxic effects of alkylating agents. However, no homolog for the *E. coli ada* gene has been reported from the sequenced genomes. The O^6-methylguanine alkyl transferase deduced from the sequence of ORF HP676 lacks the signal molecule found associated to it in other bacterial species (7). Extensive data base searches using the Ada A sequence from *Bacillus subtilis* or the conserved regions of all the Ada-like activators (50) failed to detect *H. pylori* primary sequences with homology to them. The mechanisms and the genes required for this adaptive response in *H. pylori* remain to be found.

The construction of an *H. pylori* strain with a mutant allele of *magIII* showed that MagIII is functional *in vivo* and seems to be essential for the protection of the bacterial genome from the effects of alkylating agents. The sensitivity displayed by the mutants suggests that there are no efficient backup systems for MagIII in *H. pylori*. The presence of MagIII allows *H. pylori* to protect itself from the cytotoxic effects of 3-MeA adducts produced on DNA by the action of endogenous or environmental alkylating agents. However, MagIII will not excise the most abundant product generated on DNA by alkylating agents, 7-MeG. It has been suggested that this adduct, although innocuous for cell survival, could have a mutagenic effect (51, 52). This could potentially contribute to the high genetic variability found within *H. pylori.*

Acknowledgments—The support of the ECOS-Sud program (Université René Descartes and SECYT) for the development of this collaboration between French and Argentinian laboratories is acknowledged. We thank Claudine Dhérin for the preparation of the alkylated substrates and for sharing her technical skills. We are grateful to Dr. Rodolfo Ugalde and Licenciado Cristian Danna for their assistance and for providing materials necessary for semiquantitative reverse transcription-PCR experiments. We also thank Marta Bravo for technical assistance on FPLC purification.

REFERENCES

1. Lindahl, T. (1993) *Nature* **362,** 709–715
2. Rydberg, B., and Lindahl, T. (1982) *EMBO J.* **1,** 211–216
3. Karran, P., Lindahl, T., Ofsteng, I., Evensen, G. B., and Seeberg, E. (1980) *J. Mol. Biol.* **140,** 101–127
4. Boiteux, S., Huisman, O., and Laval, J. (1984) *EMBO J.* **3,** 2569–2573
5. Larson, K., Sahm, J., Shenkar, R., and Strauss, B. (1985) *Mutat. Res.* **150,** 77–84
6. O'Connor, T. R., Boiteux, S., and Laval, J. (1988) *Nucleic Acids Res.* **16,** 5879–5894
7. Seeberg, E., and Berdal, K. G. (1997) in *Base Excision Repair of DNA Damage* (Hickson, I. D., ed) pp. 151–168, Landes Bioscience, Austin, TX
8. Riazuddin, S., and Lindahl, T. (1978) *Biochemistry* **17,** 2110–2118
9. Thomas, L., Yang, C. H., and Goldthwait, D. A. (1982) *Biochemistry* **21,** 1162–1169
10. Bjelland, S., Bjoras, M., and Seeberg, E. (1993) *Nucleic Acids Res.* **21,** 2045–2049
11. McCarthy, T. V., Karran, P., and Lindahl, T. (1984) *EMBO J.* **3,** 545–550
12. Bjelland, S., Birkeland, N. K., Benneche, T., Volden, G., and Seeberg, E. (1994) *J. Biol. Chem.* **269,** 30489–30495
13. Saparbaev, M., and Laval, J. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5873–5877
14. Matijasevic, Z., Stering, A., Niu, T. Q., Austin-Ritchie, P., and Ludlum, D. B. (1996) *Carcinogenesis* **17,** 2249–2252
15. Saparbaev, M., Kleibl, K., and Laval, J. (1995) *Nucleic Acids Res.* **23,** 3750–3755
16. Berdal, K. G., Johansen, R. F., and Seeberg, E. (1998) *EMBO J.* **17,** 363–367
17. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) *Cell* **95,** 249–258
18. Labahn, J., Scharer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L., and Ellenberger, T. E. (1996) *Cell* **86,** 321–329
19. Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yasumura, K., Tomita, K., Ihara, K., Fujii, Y., Nakabeppu, Y., Sekiguchi, M., and Fujii, S. (1996) *Cell* **86,** 311–319
20. Begley, T. J., Haas, B. J., Noel, J., Shekhtman, A., Williams, W. A., and Cunningham, R. P. (1999) *Curr. Biol.* **9,** 653–656
21. Kuo, C. F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P., and Tainer, J. A. (1992) *Science* **258,** 434–440
22. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) *EMBO J.* **14,** 4108–4120
23. Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., deJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F., and Trust, T. J. (1999) *Nature* **397,** 176–180
24. Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L.. M., Lee, N., Adams, M. D., Venter, J. C., *et al.* (1997) *Nature* **388,** 539–547
25. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
27. Girard, P. M., Guibourt, N., and Boiteux, S. (1997) *Nucleic Acids Res.* **25,** 3204–3211
28. Boiteux, S., Belleney, J., Roques, B. P., and Laval, J. (1984) *Nucleic Acids Res.* **12,** 5429–5439
29. Laval, J. (1977) *Nature* **269,** 829–832
30. Karran, P., Hjelmgren, T., and Lindahl, T. (1982) *Nature* **296,** 770–773
31. Rashtchian, A. (1995) *Curr. Opin. Biotechnol.* **6,** 30–36
32. Labigne, A. (1997) in *Methods in Molecular Medicine: Helicobacter pylori Protocols* (Clayton, C. L., and Mobley, H. L. T., eds) pp. 153–163, Humana Press, Totowa, NJ
33. Wang, Y., Roos, K. P., and Taylor, D. E. (1993) *J. Gen. Microbiol.* **39,** 2485–2493
34. Janvier, B., Grignon, B., Audibert, C., Pezennec, L., and Fauchere, J. L. (1999) *FEMS Immunol. Med. Microbiol.* **24,** 27–33
35. Trieu-Cuot, P., Gerbaud, G., Lambert, T., and Courvalin, P. (1985) *EMBO J.* **4,** 3583–3587
36. Murphy, L. D., Herzog, C. E., Rudick, J. B., Fojo, A. T., and Bates, S. E. (1990) *Biochemistry* **29,** 10351–10356
37. Kelley, L. A., Maccallum, R., and Sternberg, M. J. E. (1999) in *RECOMB99: Proceedings of the Third Annual Conference on Computational Molecular Biology* (Sorin, I., Pevzner, P., and Waterman, M., eds) pp. 218–225, Association for Computing Machinery, New York
38. Fischer, D., Barret, C., Bryson, K., Elofsson, A., Godzik, A., Jones, D., Karplus, K. J., Kelley, L. A., MacCallum, R. M., Pawowski, K., Rost, B., Rychlewski, L., and Sternberg, M. (1999) *Proteins* **Suppl 3,** 209–217
39. Marais, A., Monteiro, L., and Megraud, F. (1999) *Curr. Top. Microbiol. Immunol.* **241,** 103–122
40. Marshall, B. J., and Warren, J. R. (1984) *Lancet* **1,** 1311–1315
41. Megraud, F., and Lamouliatte, H. (1992) *Dig. Dis. Sci.* **37,** 769–772
42. Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A. B., Warnke, R. A., Jellum, E., Orentreich, N., Vogelman, J. H., and Friedman, G. D. (1994) *N. Engl. J. Med.* **330,** 1267–1271
43. Forman, D., Webb, P., and Parsonnet, J. (1994) *Lancet* **343,** 243–244
44. Mobley, H. L. (1997) *Gastroenterology* **113,** Suppl. 6, S21–S28
45. Mobley, H. L. (1996) *Am. J. Med.* **100,** 2S–11S
46. Marais, A., Mendz, G. L., Hazell, S. L., and Megraud, F. (1999) *Microbiol. Mol. Biol. Rev.* **63,** 642–674
47. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdine, G. L. (1996) *Curr. Biol.* **6,** 968–980
48. Doherty, A. J., Serpell, L. C., and Ponting, C. P. (1996) *Nucleic Acids Res.* **24,** 2488–2497
49. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E., and Samson, L. D. (1999) *BioEssays* **21,** 668–676
50. Morohoshi, F., Hayashi, K., and Munakata, N. (1990) *Nucleic Acids Res.* **18,** 5473–5480
51. Xiao, W., and Samson, L. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 2117–2121
52. Elder, R. H., Jansen, J. G., Weeks, R. J., Willington, M. A., Deans, B., Watson, A. J., Mynett, K. J., Bailey, J. A., Cooper, D. P., Rafferty, J. A., Heeran, M. C., Wijnhoven, S. W., van Zeeland, A. A., and Margison, G. P. (1998) *Mol. Cell. Biol.* **18,** 5828–5837