Published in final edited form as:

DNA Repair (Amst). 2017 September ; 57: 161–170. doi:10.1016/j.dnarep.2017.07.004.

Molecular basis for the functions of a bacterial MutS2 in DNA repair and recombination

Ge Wang and Robert J. Maier*

Department of Microbiology, University of Georgia, Athens, Georgia 30602

Abstract

Bacterial MutS2 proteins, consisting of functional domains for ATPase, DNA-binding, and nuclease activities, play roles in DNA recombination and repair. Here we observe a mechanism for generating MutS2 expression diversity in the human pathogen *Helicobacter pylori*, and identify a unique MutS2 domain responsible for specific DNA-binding. *H. pylori* strains differ in mutS2 expression due to variations in the DNA upstream sequence containing short sequence repeats. Based on Western blots, mutS2 in some strains appears to be co-translated with the upstream gene, but in other strains (e.g. UA948) such translational coupling does not occur. Accordingly, strain UA948 had phenotypes similar to its mutS2 derivative, whereas expression of MutS2 at a separate locus in UA948 (the genetically complemented strain) displayed a lower mutation rate and lower transformation frequency than did mutS2. A series of truncated HpMutS2 proteins were purified and tested for their specific abilities to bind 8-oxoG-containing DNA (GO:C) and Holiday Junction structures (HJ). The specific DNA binding domain was localized to an area adjacent to the Smr nuclease domain, and it encompasses 30-amino-acid-residues containing a “KPPKNKFKPPK” motif. Gel shift assays and competition assays supported that a truncated version of HpMutS2-C12 (~12 kDa protein containing the specific DNA-binding domain) has much greater capacity to bind to HJ or GO:C DNA than to normal double stranded DNA. By studying the in vivo roles of the separate domains of HpMutS2, we observed that the truncated versions were unable to complement the mutS2 strain, suggesting the requirement for coordinated function of all the domains in vivo.

Keywords

Helicobacter pylori; MutS2; oxidative DNA damage; DNA recombination; specific DNA binding; translational coupling

*Corresponding author: Robert J. Maier, Department of Microbiology, 815 Biological Sciences Building, University of Georgia, Athens, GA 30602, rmaier@uga.edu, Tel: 706 542 2323, Fax: 706 542 2674.

Conflict of interest

The authors declare no conflict of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The human gastric pathogen *Helicobacter pylori* is highly adapted to its ecologic niche and displays exceptional genetic variability and intra-species diversity [1]. *H. pylori* survives on the surface of the stomach lining for the lifetime of its host and causes a chronic inflammatory response [2]. Several lines of evidence suggest that *H. pylori* are exposed to oxidative damage soon after infection [3–5]. To survive in the host, *H. pylori* has a diverse repertoire of oxidant detoxification enzymes (e.g. superoxide dismutase, catalase, and peroxiredoxins) [6] as well as efficient DNA repair systems including base excision repair, nucleotide excision repair, and recombinational repair [7–9]. *H. pylori* lacks a DNA mismatch repair system, but has a MutS homolog that belongs to the MutS2 family [10].

Bacterial MutS proteins are divided into two families, MutS1 and MutS2. MutS1 is normally a part of the post-replication mismatch repair system along with MutH and MutL [11], while MutS2 is not functional in conventional DNA mismatch repair [12–14]. While *Helicobacters* and related bacteria (*Campylobacters* and *Wolinella succinogenes*) harbor only MutS2 but not MutS1, many other bacteria contain both MutS1 and MutS2 (e.g. *Deinococcus radiodurans*, *Bacillus subtilis*, *Thermus thermophilus*) [15]. Studies in the past decade revealed the functions of MutS2 in DNA recombination and repair. MutS2 in *H. pylori* as well as in some other bacteria was shown to have anti-recombination activity, playing a role in maintaining genome integrity by suppressing homologous and homeologous DNA recombination [13, 16–18]. However, *B. subtilis* MutS2 was recently shown to function in promoting homologous recombination [19]. The other major function of MutS2 is repair of oxidative DNA damage, which was first demonstrated in *H. pylori* [12], and subsequently revealed in *T. thermophilus* and *D. radiodurans* [20, 21].

MutS1 family proteins contain three major functional domains: an N-terminal mismatch recognition domain, a central DNA association domain and a C-terminal ATPase domain [11]. MutS2 lacks the N-terminal domain responsible for mismatch recognition identified in the MutS1 proteins [22]. The sequence homology between MutS1 and MutS2 proteins is mainly present at the ATPase domain, and MutS2 belongs to a highly conserved family of ABC transporter ATPases [15]. Recently, the role of the ATP binding and hydrolysis activities of HpMutS2 was elucidated [23]. Compared to MutS1, MutS2 proteins have an additional C-terminal extension (~200 residues). The C-terminal ~80 residues in MutS2 proteins constitute the highly conserved Smr domain (small *mutS*-related protein) [24]. In organisms that lack MutS2, such as *E. coli*, there exist a small stand-alone Smr protein [25]. Conversely, organisms containing mutS2 (such as *H. pylori*) usually do not have an independent *smr* gene [16]. The nuclease activity of the Smr domain in *T. thermophilus* MutS2 was extensively studied [17, 26], and a recent *H. pylori* MutS2 study revealed a nuclease site (LDLK motif) at the enzyme's N-terminus in addition to the Smr domain [27].

H. pylori exhibits the highest level of genetic diversity known among bacteria, due to a demonstrated high mutation rate and high frequency of DNA recombination [28, 29]. The high genetic variability provides *H. pylori* with an adaptive advantage in the gastric environment of the human host [30]. In this study, we analyzed the *mutS2* gene sequences from different *H. pylori* strains, and found a highly variable sequence containing short

sequence repeats in the 5'-end of the *mutS2* genes. Sequence variation at this region could control the ability of the *mutS2* to be co-translated with its upstream gene. Western blot analysis confirmed that *mutS2* in some strains are co-translated with the upstream gene, while in other strains such translational coupling does not occur. Those strains that do not express *MutS2* have a high mutation rate and have a higher capacity for DNA recombination. Herein we uncovered a novel mechanism (phase variation together with translational coupling) responsible for varied expression of *HpMutS2* in different strains, contributing in part for diversity generation.

In accordance with its two major functions, *H. pylori* *MutS2* is able to bind to DNA structures mimicking recombination intermediates and the double stranded DNA containing 8-oxoG lesions [12, 13]. However, the specific domain of *HpMutS2* responsible for binding to these anomalous DNA structures remains elusive. In the second part of this study, we focused on domain analysis of *HpMutS2* to identify such a specific DNA binding domain (SDB). Using a series of truncated *HpMutS2* constructs in gel shift assays, we first mapped SDB to the linker region between ATPase domain and Smr domain. Further, SDB was defined as a sequence of 30-amino-acid-residues containing a “KPPKNKFKPPK” motif. In comparing our study to that of *T. thermophilus* *MutS2* [17], here we defined a unique specific DNA binding domain in *HpMutS2*. Finally, we examined the in vivo roles of the separate domains of *HpMutS2* and discussed the differences between in vitro and in vivo activities of *HpMutS2* domains. This study reveals a molecular basis for the functions of *HpMutS2* in DNA recombination and repair.

2. Materials and Methods

2.1. *H. pylori* growth conditions and assays for sensitivity to oxidative stress

H. pylori was cultured on Brucella agar (BA) or Brain-heart-infusion/yeast extract (BHI-YE) agar plates supplemented with 10% defibrinated sheep blood or 5% fetal bovine serum. Cultures of *H. pylori* were grown microaerobically at 37°C in a 5% CO₂ incubator under continuously controlled levels of oxygen (4% partial pressure). For paper disk sensitivity assay, sterile filter paper disks (7.5 mm in diameter) were applied to BA plates that had been streaked for confluent growth of *H. pylori* strains. Ten µl of the oxidizing agent were applied to each disk. After the plates were incubated under a 4% partial pressure O₂ atmosphere (contains 5% CO₂, and 91% N₂) for 48 h, the clear zones surrounding the disks (the distances from the edge of the disk to the end of the clear zone) were measured. For air survival assays, grown *H. pylori* cells were suspended in phosphate-buffered saline (PBS) and incubated at 37°C under normal atmospheric conditions. Samples were removed at various times, serially diluted, and spread on BA plates. Colony counts were recorded after 3 days of incubation under the 4% O₂ condition.

2.2. Construction of *H. pylori mutS2* strains

A 2.52 kb fragment containing the *H. pylori mutS2* gene was PCR amplified from genomic DNA of strain 26695 and cloned into pGEM-T vector to generate pGEM-*mutS2*. Subsequently, a chloramphenicol acetyl transferase (CAT) cassette was inserted within the *mutS2* sequence of pGEM-*mutS2* at the *HindIII* sites. The disrupted *mutS2* gene was then

introduced into *H. pylori* by natural transformation *via* allelic exchange and chloramphenicol-resistant colonies were isolated. The 26695 *mutS2* strain was constructed previously [12] and the UA948 *mutS2* strain was constructed in this study. The disruption of the gene within the genome was confirmed by PCR, showing an increase in the expected size of the PCR product and by direct sequencing of the PCR fragment.

2.3. Construction of *H. pylori* strains for complementation

The complemented strains were constructed by inserting different versions (complete or truncated) of HpMutS2 gene in the *rdxA* locus of the *mutS2* chromosome. Disruption of *rdxA* results in metronidazole resistance that is used for selection of transformants. PCR products corresponding to the upstream sequence of the *rdxA* gene (~300 bp, including the promoter, RBS, and translation start codon ATG), the *mutS2* gene sequence (full-length or truncations), and the downstream sequence of the *rdxA* gene (~300 bp) were amplified in three separate PCRs and then stitched together in subsequent PCR (overlapping PCR). In the final PCR product, the *rdxA* gene was exactly replaced by the *mutS2* gene (full-length or truncated versions). The final PCR product was used to transform the *mutS2* strain with subsequent selection for metronidazole (16 µg/ml) resistant colonies.

2.4. Overexpression and purification of *H. pylori* MutS2 and its variant proteins

Different versions of *mutS* gene were amplified by PCR (for primers, see Table 1) from strain 26695 (or UA948) and cloned into pET-21a vector. The recombinant plasmid was transformed into *E. coli* BL21 Rossetta strain. The transformant *E. coli* cells were grown at 37°C to an OD₆₀₀ of 0.5, and expression of the MutS2 variant protein was induced by adding 0.8 mM isopropyl-β-D-thiogalactopyranoside (IPTG) followed by a 3 hr incubation of transformant cells. The MutS2 variant protein was purified using a nickel-nitrilotriacetic acid (Ni-NTA) affinity column (Qiagen) following vendor's instructions and protein expression was analyzed by SDS polyacrylamide gel electrophoresis.

2.5. Protein gel electrophoresis and Western blotting

Bacteria were harvested from the plates and re-suspended in PBS containing 20 mM sodium phosphate (pH 8.0) and 150 mM NaCl. After one wash with the buffer, cells were re-suspended in the same buffer and then broken by two passages through a French pressure cell at 138,000 kPa (SLM instruments, Inc.). The cell lysates were obtained by centrifugation (8,000 rpm for 10 min) and the supernatant was transferred to a clean tube. The protein concentration of cell extracts was determined by the Bradford protein assay (Bio-Rad). Seven micrograms of cell extract was mixed with the SDS buffer and incubated at 90°C for five minutes. Proteins were then separated on the 12.5% SDS-PAGE gel by electrophoresis for 1.5 hours at 100 V. The gels were either stained with Coomassie blue to visualize total proteins or subjected to Western blotting to identify the MutS2 protein. For Western blot, the proteins on the gel were electro-transferred onto a nitrocellulose membrane. The membrane was then incubated with anti-MutS2 antiserum (1:500 dilution) followed by incubation with secondary goat-anti-rabbit-IgG conjugated with alkaline phosphatase (Bio-Rad, 1:1000 dilution). Polyclonal anti-MutS2 antiserum against the purified HpMutS2 protein was produced in rabbits by Antagene, Inc., using their

conventional antiserum production protocols. Anti-UreA antiserum (rabbit polyclonal, Santa Cruz Biotechnology, Dallas, TX) was used as an internal control for Western blotting.

2.6. Detecting protein-DNA binding

Normal oligonucleotides as well as the oligonucleotides containing 8-oxoG lesions were synthesized by Integrated DNA Technologies (Coralville, IA). Some oligonucleotides were 5'-end labelled using [γ - ^{32}P] ATP and T4 polynucleotide kinase, and then purified by ethanol precipitation. dsDNA substrates were prepared by annealing molar equivalents of unlabeled or ^{32}P -labeled oligonucleotides with appropriate unlabeled complementary oligonucleotides as follows: heating at 90°C for 5 min followed by slow cooling overnight to room temperature. Residual single-stranded DNA was purified away from dsDNA using benzoylated naphthoylated DEAE-cellulose (Sigma). dsDNA was then concentrated using a YM-10 microcon centrifugal filter (Amicon Bioseparations). The HJ structure was created by annealing one ^{32}P -labeled or unlabeled strand with three unlabeled strands.

Protein-DNA binding assays were performed by incubating each ^{32}P labelled DNA substrate (10 nM) with purified HpMutS2 variant protein at the indicated concentrations, and 50 $\mu\text{g}/\text{ml}$ sonicated calf thymus DNA in binding buffer (25 mM Tris-acetate, pH 8.0, 10 mM KCl, 8 mM Mg-acetate, 1 mM dithiothreitol and 3.5% [w/v] PEG8000) at 30°C for 10 min. For the competition assays, different concentrations of unlabeled DNA substrates were added in the binding reaction. After adding the gel-loading dye, the reaction mixture was subjected to electrophoresis on a 5% native polyacrylamide gel. The gel was subsequently dried, and analyzed using a PhosphorImager and ImageQuant software. Alternatively, some protein-DNA binding assays were performed on agarose gels (1.5%) with unlabeled DNA substrates (1 μM). The agarose gels were visualized after staining with ethidium bromide. Another assay for protein-DNA binding was the filter retention assay [13]. The binding reaction mixtures of ^{32}P labelled DNA substrate (10 nM) with purified HpMutS2 variant proteins (100 nM) were loaded onto a nitrocellulose Hybond-C membrane (Amersham) previously treated with 0.4M KOH to avoid DNA-binding. Membranes were then washed and air-dried. Protein-bound DNA was determined by quantifying the radioactivity bound to the nitrocellulose, using a PhosphorImager.

2.7. Determination of mutation rate

H. pylori cells were grown in the 4% O_2 -containing atmosphere in BHI-YE broth to late-log phase (viable cell number of about 10^9 cells/ml). For determination of mutation frequency under oxidative stress conditions, cell cultures were exposed to air for 4 h at 37°C, while this step was omitted for the control experiments (without air-exposure). The cell culture was diluted 1:10000 in BHI-YE broth (to approximately 10^4 to 10^5 cells/ml) and divided into 20 independent 0.5-ml aliquots. These aliquots were then each allowed to grow (at 4% O_2) for 2 days to obtain parallel, independent cultures. The number of rifampicin resistant mutants that emerged in each culture was determined by plating the entire culture on BHI-YE agar plates containing 20 $\mu\text{g}/\text{ml}$ rifampicin. The total number of cells was determined by plating an appropriate dilution of three cultures on nonselective medium. Colonies on both selective and nonselective plates were counted after incubation for 4 days. The mutation rate was calculated by the Lea-Coulson method of median as described [31, 32].

2.8. Assay for DNA transformation frequency

The donor DNA used in this study included a 330 bp PCR fragment of *H. pylori rpoB* gene fragment containing a site-specific mutation (at the center of the fragment) conferring rifampicin resistance. *H. pylori* strains were grown on BA plates to late log-phase, and the cells were suspended in PBS at a concentration of $\sim 10^8$ /ml (recipient cells for transformation). A sample (30 μ l) of the suspension was mixed with 100 ng of donor DNA and spotted onto a BA plate. After incubation for 18 h under the microaerobic condition (4% partial pressure O_2) at 37°C, the transformation mixture was harvested and suspended in 1 ml of PBS. 100 μ l portions of the suspension (or appropriate dilutions) were plated onto either BA plates or BA plates containing selective antibiotic (20 μ g/ml rifampicin). The plates were incubated for 4 days under the microaerobic condition at 37°C, and the numbers of colonies were counted. The transformation frequency was determined by the number of resistant colonies divided by the total number of CFU. The frequency of transformation is expressed as the number of transformants per 10^5 recipient cells.

3. Results and Discussion

3.1. HpMutS2 expression in vitro and in vivo

Based on analysis of the *mutS2* gene sequences from different *H. pylori* strains, there exists a highly variable sequence in the extreme 5'-end of the *mutS2* genes, as shown in Fig. 1A. These highly variable sequences contain apparent "(C)nAATGAA" repeats, where (C)n denotes simple polyC repeats. While there are two "(C)nAATGAA" repeat units in the SS1 *mutS2*, the UA948 *mutS2* appears to contain three repeat units. In addition, the number of C's in the polyC repeat is 5 or 4, respectively, in the two strains. Table 2 summarizes the DNA sequence variations at HP0622-MutS2 junctions among *H. pylori* strains for which the sequence information are available. The number of the "(C)nAATGAA" repeat units varies with 2 or 3; and the number of C's in the polyC repeat varies at a range of 2 to 7.

For the reference strain 26695, the TTG (labeled with *** in Fig. 1A) was assigned as the start codon, which is located 35 nucleotides upstream of the stop codon of HP0622 [33]. Due to the variable length of this region, the hypothetical TTG start codon annotated in strain 26695 is not in the coding frame of the *mutS2* in some other strains (e.g. UA948). We propose that the ATG (underlined in Fig. 1A) is the start codon of *mutS2*, which has overlap with the stop codon of HP0622 in many strains (Table 2, and two examples SS1 and 26695 in Fig. 1A). There is no apparent Shine-Dalgarno sequence (ribosome binding site, RBS) in front of the ATG start codon. Thus, *mutS2* could be co-translated with HP0622, obviating the need for RBS to re-initiate the translation of *mutS2*. Such a "translational coupling" was observed and well described in many prokaryotic operons and it is defined as the interdependence of translation efficiency of neighboring genes encoded by the same polycistronic mRNA [34–36]. In the case of *E. coli* trpEDCBA operon, for example, the efficient translation of trpD is dependent on trpE due to overlap of the trpE stop codon with the trpD start codon; whereas the translational coupling is not operating between trpC and trpB with their 11 base pair intercistronic sequence [34]. It is assumed that in translational coupling the same ribosome that reads the upstream cistron reinitiates at the second frame [35].

We used the ATG (underlined in Fig. 1A) as the start codon for overexpression and purification of HpMutS2 in our previous work [12] as well as for this work. When the HpMutS2 gene sequence (including the coding sequence and additional sequence upstream of it) was cloned into pGEMT vector with T7 promoter but no RBS, the HpMutS2 protein was not expressed in *E. coli* (data not shown). In contrast, HpMutS2 protein was overexpressed and purified (See Fig. 1B, lane 1) from *E. coli* using plasmid pET21 that contains T7 promoter and a RBS at the appropriate position.

To examine the MutS2 expression in different *H. pylori* strains, Western blot analysis was performed (Fig. 1B). MutS2 was expressed in strain 26695 (lane 2) and SS1 (lane 6), but not in UA948 (lane 4). The strain UA948 result was similar to strain 26695 *mutS2* (lane 3). When the *mutS2* gene was introduced into *rdxA* locus (under control of *rdxA* promoter including the transcriptional and translational signals required for RdxA expression) in UA948 *mutS2* cells, the MutS2 was expressed (lane 5); the expression level was similar to that in the wild type strain 26695 or SS1 (lanes 2 and 6). These results support the hypothesis that translational coupling is operating between Hp0622 and MutS2 in the strains like 26695 and SS1. Genes that are involved in the same biochemical or physiological pathway are often grouped into an operon for coordination of their expression level. While MutS2 has been shown to function in DNA metabolism (recombination and repair), the function of the upstream gene HP0622 is unknown. Further upstream is gene HP0623, encoding a ligase enzyme MurC that is involved in the synthesis of peptidoglycan [37].

The translational coupling of HP0622-MutS2 is not operating in strain UA948; this is likely due to a 16-bp space between the HP0622 stop codon and the MutS2 start codon (Fig. 1A). As shown in Table 2, in many *H. pylori* strains there exist different length (4 to 19 bp) of space between the HP0622 stop codon and the MutS2 start codon due to sequence variations. The molecular mechanism regulating sequence variation at HP0622-MutS2 junctions is apparently slipped-strand mispairing mediated by short sequence repeats [38]. *H. pylori* was shown to have extensive, nonrandomly distributed repetitive chromosomal sequences, contributing to the variation within individual hosts [39]. Some known examples include presence of homopolymeric tracts in genes of fucosyltransferases that are involved in biosynthesis of Lewis antigens, enabling *H. pylori* to produce variable lipopolysaccharide epitopes on its surface [40]. The gene for a DNA glycosylase MutY contains a homopolymeric tract of eight adenines; slipped-strand mispairing during DNA replication can lead to frameshifts; this mechanism would produce a non-functional MutY and a mutator phenotype [41]. Here we show another example: sequence variation mediated by short sequence repeats together with translational coupling would be expected to provide a mechanism for diversification of MutS2 expression among *H. pylori* strains.

3.2. MutS2-related phenotypes of the strain UA948

UA948 is an *H. pylori* strain that was used for characterization of an $\alpha(1,3/4)$ fucosyltransferase responsible for Lewis antigen synthesis [42]. As there is no MutS2 expression in UA948, we were interested in examining its MutS2-related phenotypes. Our previous work indicated that strain SS1 *mutS2* was more sensitive than wild type cells to oxidative stress induced by agents such as H₂O₂, paraquat or oxygen [12]. However, when

similar assays were performed to compare wild type strains of 26695, SS1, and UA948, no significant difference was observed between the strains (data not shown). This could be explained by the abundance of diverse antioxidant systems among *H. pylori* strains [6]; the loss of MutS2 on oxidative stress resistance may be masked by the existence of other antioxidant enzymes, such as catalase, SodB, AhpC, NapA and a handful of DNA repair enzymes.

Next, we characterized the two observable phenotypes, mutation rate under oxygen stress conditions, and transformation frequency; strain UA948 WT, UA948 *mutS2*, and the complemented strain in which the MutS2 is expressed under the *rdxA* promoter were compared. Under normal laboratory conditions for *H. pylori* culture, no significant difference in spontaneous Rif^R mutation rates were observed between the strains (data not shown), which is in agreement with a previous observation [12]. However, under the condition of O₂ stress (exposure of the cells to air for 4 hours), the mutation rate for the complemented strain was about 8 fold lower than for the WT and *mutS2* strain (Table 3). The results further support that MutS2 is not expressed in the UA948 WT strain, and that upon expression of MutS2, as in the complemented strain, oxidative DNA damage is repaired. Similarly, both UA948 WT and *mutS2* strain have a relatively higher transformation frequency, while the MutS2 expression in the complemented strain lowered transformation frequency about 6 fold (Table 3). This is in agreement with the anti-recombination function of HpMutS2 [13].

All *H. pylori* strains contain the *mutS2* gene, but sequence variation mediated by short sequence repeats together with translational coupling could provide mechanisms for regulating MutS2 expression level among the strains. Due to MutS2's role in repairing oxidative DNA damage, the strains with low MutS2 expression would be expected to have a mutator phenotype under oxidative stress conditions. Reduced MutS2 (as well as MutY) expression may account for the high mutation frequency observed for certain natural *H. pylori* isolates [28] and for a mutation burst induced by the inflammatory response during the initial, acute *H. pylori* infection in humans and primates [43]. In addition, due to loss of MutS2 anti-recombination function, such strains would have a higher capacity for natural transformation. Thus, regulation of MutS2 expression levels would contribute to phenotypic diversity and adaptation of *H. pylori* in the dynamic gastric niche of human hosts.

3.3. Specific DNA binding domain of HpMutS2

HpMutS2 is composed of three known functional domains: (1) N-terminal DNA association domain, (2) intermediate ATPase/Dimerization domain, and (3) C-terminal nuclease Smr domain (Fig. 2A). There is a linker region (~120 amino acid residues) with unknown function between the ATPase domain and the Smr domain. To identify the domain responsible for binding to specific DNA substrates (anomalous DNA structures), truncated HpMutS2 gene constructs were prepared by PCR, and corresponding proteins were purified as described in Table 1 and Fig. 2A. W85 is the whole length HpMutS2 protein with a molecular mass of 85 kDa. N31 contains only the N-terminal "DNA Association" domain; N62 contains both "DNA Association" and ATPase domains; and C23 corresponds to the C-terminal region of MutS2 including Smr domain. It should be noted that the numbering for

amino acid residues of HpMutS2 and thus the molecular weights of the variant proteins (in Table 1 and Fig. 2A) are different from that reported by the previous studies of another group [23, 27], as we used the “ATG” start codon identified in Fig. 1.

Two representative types of DNA substrates, 8-oxoG containing dsDNA and Holiday Junction structures, were used to test the binding ability of HpMutS2 variant proteins. Formation of specific DNA-protein complex was examined by gel shift assay (Fig. 2B and C). As expected, the complete HpMutS2 protein (W85) was able to bind to GO:C containing DNA, but both N31 and N62 showed extremely weak binding ability. Surprisingly, C23 displayed the strongest activity for binding to GO:C (Fig. 2B). A similar binding pattern for all constructs was observed when HJ was used as DNA substrate (Fig. 2C). As the Smr domain is known for the nuclease activity, these results indicated that the specific DNA-binding domain localizes to the linker region between ATPase domain and Smr domain. This finding is in contrast to that observed in *T. thermophilus* MutS2 where the specific DNA binding domain is located in the “DNA association” domain [17, 26]. Thus, HpMutS2 appears to have a unique specific DNA recognition domain.

3.4. Specific DNA recognition motif

To further define the specific DNA recognition domain, we construct additional truncated MutS2 versions containing different lengths of the C-terminal region (Table 1 and Fig. 3A). C8 is expected to be the minimal size containing the Smr (nuclease) domain; C9 contains an additional 5 residues (IPKPK); C10 contains a further additional 14 residues including 5 lysine residues and comprising two KPPK motifs; and C12 contains a further additional 10 residues including 3 lysine residues. These truncated MutS2 variant proteins (C8, C9, C10, C12, C23) were overexpressed in *E. coli* and purified to near homogeneity as analyzed on SDS-PAGE (Fig. 3B). As this series of MutS2 truncations (C8 through C23) lack the ATPase/Dimerization domain, we assume they are present as a monomer in solution. Similarly truncated HpMutS2 protein in the study of Damke et al [27], HpexSmr (equivalent to our C23) or HpSmr (equivalent to our C8) was shown to be a monomer. We also unsuccessfully attempted to generate a construct that contains only the putative specific DNA-binding domain but lacks the Smr domain (namely C23 minus C8), suggesting such a construct is unstable.

Specific DNA binding abilities of these proteins were analyzed by both gel shift and filter retention assays. Fig. 3C is a representative gel shift assay showing the binding to HJ DNA substrate. As expected, C8 (Smr domain) lacked detectable specific binding, and C9 showed an extremely weak binding capacity. The three longer constructs (C10, C12, C23) showed extensive binding; among them C12 appears to have the strongest binding ability. Quantification of the specific DNA binding ability of each HpMutS2 truncated variant protein to the two DNA substrates (GO:C or HJ) was determined using the filter retention assay (Fig. 3D). Generally, these proteins bind to HJ with a slightly higher affinity than to GO:C. Significant specific binding to both DNA substrates was observed for C10, C12, and C23; the C12 version exhibited the strongest binding ability.

Based on these results, we propose the specific DNA binding domain (SDB in Fig. 3A) is the 30 amino acid residue (640–669) area that is an N-terminal extension of Smr domain. In

this region, there are ten lysine residues including a sequence motif “KPPKNKFKPPK” that may be responsible for recognizing/binding to anomalous DNA structures. The study of HpMutS2 by Damke et al [27] focused on the nuclease activities, but also examined DNA binding properties of HpMutS2 and variants. According to their data, all the proteins shown to have HJ binding ability appeared to contain the region that we now identify as the specific DNA binding domain; so there is no contradiction between their results and our present results. On the other hand, the study of *T. thermophilus* MutS2 [17] indicated that the N-terminal DNA association domain mainly contributes to the branched DNA recognition, although the ttSmr₆₂₁ (the Smr domain plus ~40 AA residues N-terminal extension, similar to our HpMutS-C12) is also capable of recognizing the branched DNA structure with a relatively lower affinity. Based on alignment of MutS2 sequences from various bacterial species (available in NCBI database), the specific DNA binding domain identified in this work, particularly the “KPPKNKFKPPK” motif, is only present in *H. pylori* and in *Helicobacter acinonychis*, but not in any other species (Fig. 3E).

3.5. Specific vs. non-specific DNA binding

In accordance with its two major functions (anti-recombination and oxidative DNA repair), HpMutS2 is able to bind to both the HJ structure and GO-containing DNA. GO-containing DNA may structurally appear as a small loop structure; together with recombination intermediates, they are all DNA anomalous structures. To clarify the specific binding to DNA anomalous structures vs. non-specific binding to normal dsDNA, we further examined the binding ability of HpMutS2-C12. Normal dsDNA, GO:C-containing DNA, and HJ DNA substrates were incubated with different amounts of HpMutS2-C12, and the formation of protein-DNA complexes were analyzed by gel shift assays (Fig. 4B). In the presence of 0.3 or 0.6 μM MutS2-C12, no binding to normal dsDNA was observed. Only with very high concentration of MutS2-C12 (1.2 μM), there appeared to be a small amount of binding; this is non-specific DNA binding. In contrast, significant binding to GO:C or HJ DNA was readily observed at the lower concentration (0.3 μM) of HpMutS2-C12; and the formation of specific DNA-protein complexes increased greatly when higher concentrations of HpMutS2-C12 were present. With 1.2 μM HpMutS2-C12, all the HJ DNA was bound to the protein (Fig. 4B). The same experiment was performed for HpMutS2-N31 (Fig. 4A). For any of the DNA substrates tested, no significant binding was observed when a lower concentration (0.3 or 0.6 μM) of MutS2-N31 was used. Only with a high concentration of MutS2-N31 (1.2 μM), there appeared to be a significant binding. When the same experiment was performed for HpMutS2-N62, the same result as for HpMutS2-N31 was obtained (not shown). These results suggest that the N-terminal DNA association domain of HpMutS2 is not responsible for specific binding to anomalous DNA structures, but may have the capacity for non-specific DNA binding.

The relative abilities of the C-12 version to recognize different DNA substrates were further examined with a competition assay. 100 nM HpMutS2-C12 was incubated with 10 nM ³²P-labeled GO:C DNA substrate and increasing amount of unlabeled G:C, GO:C, or HJ competitor DNAs. The normal dsDNA showed a very weak ability to compete for binding with GO:C, while the GO:C and HJ substrates were strong competitors. Based on these results, the relative affinity of HpMutS2-C12 for these substrates was G:C $\ll$ GO:C < HJ.

This specific DNA-binding pattern of HpMutS2-C12 is similar to that observed for the whole length HpMutS2 in the previous study [12]. The results provided additional evidence that HpMutS2-C12 contains the specific DNA-binding domain.

3.6. In vivo function of HpMutS2 and its variants

The biochemical assays above revealed that the 30 amino acid residues at the N-terminal extension of Smr domain is the specific DNA-binding domain of HpMutS2. Thus, HpMutS2-C12 or HpMutS2-C23 contains both the specific DNA-binding activity and the Smr nuclease activity, while HpMutS2-N62 contains the DNA association domain (including the nuclease motif LDLK) and the ATPase activity. To study whether these truncated versions of HpMutS2 encompass functional portions of HpMutS2 roles in vivo, these truncated proteins (expressed under control of *rdxA* promoter) were used to complement the 26695 *mutS2* strain; the mutation rates under oxygen stress and transformation frequencies were determined and compared (Table 4). As expected, deletion of the *mutS2* resulted in ~7-fold increase of mutation rate under an O₂ stress condition, while the whole length HpMutS2-W85 is able to complement the function of the MutS2 to bring the mutation rate back to the WT (low) level. In contrast, the 26695 *mutS2* strain expressing any of the truncated proteins (N62, C23, or C12) had a high level of mutation rate similar to the 26695 *mutS2* strain, suggesting that these truncated proteins are unable to complement the function of the MutS2 in vivo.

To determine transformation frequencies, a 330 bp PCR fragment of the *H. pylori rpoB* gene containing a site-specific mutation (conferring rifampicin resistance) was used as donor DNA for natural transformation. The 26695 WT strain showed a frequency of 5.8×10^{-5} . As expected, loss of MutS2 anti-recombination function resulted in a 6-fold increase of transformation frequency. The whole length HpMutS2-W85, but not any of the truncated forms (N62, C23, or C12), was able to complement this function in vivo (Table 4). These results are similar to those observed by Damke et al. [27] who suggested that the presence of both nuclease sites (N-terminal LDLK motif and C-terminal Smr domain) is essential in vivo for the efficient function of HpMutS2. On the other hand, the results may also indicate that the coordinated action of each domain is required for complete function of HpMutS2 in vivo, although each separate domain showed in vitro activities of ATPase, DNA-binding, and nuclease.

The DNA association domain in MutS1 was shown to form a major portion of a sliding clamp, with positive residues forming loose hydrophilic associations with DNA [44]. We hypothesize that the N-terminal “DNA association” domain of HpMutS2 is mainly responsible for non-specific DNA-binding and sliding on the genome to dynamically recognize anomalous DNA structures. When HpMutS2 encounters an anomalous DNA structure, the “specific DNA-binding” domain will recognize and bind to it, and the nuclease activity of Smr on site follows immediately. Both non-specific and specific DNA-binding activities as well as nuclease activities could be modulated by conformational changes within HpMutS2 driven by ATPase activity [23].

What could be the biological meaning for HpMutS2 having two functions: (1) suppression of DNA recombination and (2) repair of oxidative DNA damage? The observation that the

same domain of HpMutS2 is responsible for specific binding to 8-oxoG containing dsDNA as well as the recombination intermediates (HJ) suggests that HpMutS2 may play a role in coordinating DNA repair pathways. In physiological conditions (i.e. in vivo), *H. pylori* is thought to frequently suffer oxidative stress, leading to DNA damage [6, 9]. Accumulated oxidative DNA damage will eventually result in DNA double strand breaks that require recombination function for repair [45]. Compared to base excision repairs, DNA recombinational repair systems are much more complex and could be mutagenic due to homeologous recombination or cause rearrangement due to recombination with direct repeat sequences [45]. The base excision repair pathway is the most important mechanism for the repair of oxidative DNA damage in human-adapted pathogens [46]. Thus, in low O₂ stress conditions, HpMutS2 may continuously function as an enzyme for repairing 8-oxoG containing DNA, while suppressing DNA recombination. Under a severe stress condition when large amounts of DSBs are formed, DNA recombinational repair will be favored [47, 48].

4. Conclusion

Our results indicate that phase variation and translational coupling are the underlying mechanisms responsible for varied expression of MutS2 in different *H. pylori* strains, contributing to the high genetic variability observed in this bacterial species. As an example, we showed that the MutS2 is not expressed in a natural *H. pylori* isolate UA948, and this strain displays a high mutation rate and a high capacity for DNA recombination. With a series of DNA binding assays, a specific DNA-binding domain of HpMutS2 was identified; it is localized to an area adjacent to the Smr nuclease domain, and it encompasses 30-amino-acid-residues containing a “KPPKNKFKPPK” motif. This specific DNA-binding domain is responsible for specific binding to both 8-oxoG containing dsDNA and Holiday junction structures, providing the basis for the functions of HpMutS2 in both oxidative DNA damage repair and DNA recombination. Although each separate domain of HpMutS2 showed in vitro activities of ATPase, DNA-binding, and nuclease, the truncated versions were unable to complement the mutS2 strain, suggesting that the coordinated action of all domains are required for complete function of HpMutS2 in vivo.

Acknowledgments

This work was supported by NIH grant R21AI078096 and The University of Georgia Foundation. The authors are grateful to Leja Lo for technical assistance.

References

1. Suerbaum S. Genetic variability within *Helicobacter pylori*. *International journal of medical microbiology: IJMM*. 2000; 290:175–181. [PubMed: 11045922]
2. Kusters JG, van Vliet AH, Kuipers EJ. Pathogenesis of *Helicobacter pylori* infection. *Clinical microbiology reviews*. 2006; 19:449–490. [PubMed: 16847081]
3. Algood HM, Cover TL. *Helicobacter pylori* persistence: an overview of interactions between *H. pylori* and host immune defenses. *Clinical microbiology reviews*. 2006; 19:597–613. [PubMed: 17041136]
4. Chaturvedi R, Cheng Y, Asim M, Bussiere FI, Xu H, Gobert AP, Hacker A, Casero RA Jr, Wilson KT. Induction of polyamine oxidase 1 by *Helicobacter pylori* causes macrophage apoptosis by

hydrogen peroxide release and mitochondrial membrane depolarization. The Journal of biological chemistry. 2004; 279:40161–40173. [PubMed: 15247269]

5. Allen LA, Beecher BR, Lynch JT, Rohner OV, Wittine LM. Helicobacter pylori disrupts NADPH oxidase targeting in human neutrophils to induce extracellular superoxide release. J Immunol. 2005; 174:3658–3667. [PubMed: 15749904]
6. Wang G, Alamuri P, Maier RJ. The diverse antioxidant systems of Helicobacter pylori. Molecular microbiology. 2006; 61:847–860. [PubMed: 16879643]
7. Dorer MS, Sessler TH, Salama NR. Recombination and DNA repair in Helicobacter pylori. Annual review of microbiology. 2011; 65:329–348.
8. Wang, G., Maier, RJ. A Recombination Puzzle Solved: Role for New DNA Repair Systems in Helicobacter pylori Diversity/Persistence. In: Kruman, I., editor. DNA repair. InTech; 2011. p. 3-20.
9. O'Rourke EJ, Chevalier C, Pinto AV, Thiberge JM, Ielpi L, Labigne A, Radicella JP. Pathogen DNA as target for host-generated oxidative stress: role for repair of bacterial DNA damage in Helicobacter pylori colonization. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:2789–2794. [PubMed: 12601164]
10. Eisen JA. A phylogenomic study of the MutS family of proteins. Nucleic acids research. 1998; 26:4291–4300. [PubMed: 9722651]
11. Hsieh P. Molecular mechanisms of DNA mismatch repair. Mutat Res. 2001; 486:71–87. [PubMed: 11425513]
12. Wang G, Alamuri P, Humayun MZ, Taylor DE, Maier RJ. The Helicobacter pylori MutS protein confers protection from oxidative DNA damage. Molecular microbiology. 2005; 58:166–176. [PubMed: 16164556]
13. Pinto AV, Mathieu A, Marsin S, Veaute X, Ielpi L, Labigne A, Radicella JP. Suppression of homologous and homeologous recombination by the bacterial MutS2 protein. Molecular cell. 2005; 17:113–120. [PubMed: 15629722]
14. Fukui K, Masui R, Kuramitsu S. Thermus thermophilus MutS2, a MutS paralogue, possesses an endonuclease activity promoted by MutL. Journal of biochemistry. 2004; 135:375–384. [PubMed: 15113836]
15. Sachadyn P. Conservation and diversity of MutS proteins. Mutat Res. 2010; 694:20–30. [PubMed: 20833188]
16. Kang J, Huang S, Blaser MJ. Structural and functional divergence of MutS2 from bacterial MutS1 and eukaryotic MSH4-MSH5 homologs. Journal of bacteriology. 2005; 187:3528–3537. [PubMed: 15866941]
17. Fukui K, Nakagawa N, Kitamura Y, Nishida Y, Masui R, Kuramitsu S. Crystal structure of MutS2 endonuclease domain and the mechanism of homologous recombination suppression. The Journal of biological chemistry. 2008; 283:33417–33427. [PubMed: 18838375]
18. Jeong E, Jo H, Kim TG, Ban C. Characterization of multi-functional properties and conformational analysis of MutS2 from Thermotoga maritima MSB8. PloS one. 2012; 7:e34529. [PubMed: 22545085]
19. Burby PE, Simmons LA. MutS2 Promotes Homologous Recombination in Bacillus subtilis. Journal of bacteriology. 2017; 199
20. Fukui K, Wakamatsu T, Agari Y, Masui R, Kuramitsu S. Inactivation of the DNA repair genes mutS, mutL or the anti-recombination gene mutS2 leads to activation of vitamin B1 biosynthesis genes. PloS one. 2011; 6:e19053. [PubMed: 21552516]
21. Zhang H, Xu Q, Lu M, Xu X, Wang Y, Wang L, Zhao Y, Hua Y. Structural and functional studies of MutS2 from Deinococcus radiodurans. DNA repair. 2014; 21:111–119. [PubMed: 24811920]
22. Sixma TK. DNA mismatch repair: MutS structures bound to mismatches. Curr Opin Struct Biol. 2001; 11:47–52. [PubMed: 11179891]
23. Damke PP, Dhanaraju R, Marsin S, Radicella JP, Rao DN. Mutations in the nucleotide binding and hydrolysis domains of Helicobacter pylori MutS2 lead to altered biochemical activities and inactivation of its in vivo function. BMC microbiology. 2016; 16:14. [PubMed: 26843368]
24. Moreira D, Philippe H. Smr: a bacterial and eukaryotic homologue of the C-terminal region of the MutS2 family. Trends Biochem Sci. 1999; 24:298–300. [PubMed: 10431172]

25. Gui WJ, Qu QH, Chen YY, Wang M, Zhang XE, Bi LJ, Jiang T. Crystal structure of YdaL, a stand-alone small MutS-related protein from *Escherichia coli*. *Journal of structural biology*. 2011; 174:282–289. [PubMed: 21276852]
26. Fukui K, Kosaka H, Kuramitsu S, Masui R. Nuclease activity of the MutS homologue MutS2 from *Thermus thermophilus* is confined to the Smr domain. *Nucleic acids research*. 2007; 35:850–860. [PubMed: 17215294]
27. Damke PP, Dhanaraju R, Marsin S, Radicella JP, Rao DN. The nuclease activities of both the Smr domain and an additional LDLK motif are required for an efficient anti-recombination function of *Helicobacter pylori* MutS2. *Molecular microbiology*. 2015; 96:1240–1256. [PubMed: 25800579]
28. Bjorkholm B, Sjolund M, Falk PG, Berg OG, Engstrand L, Andersson DI. Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori*. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98:14607–14612. [PubMed: 11717398]
29. Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M. Free recombination within *Helicobacter pylori*. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95:12619–12624. [PubMed: 9770535]
30. Kang J, Blaser MJ. Bacterial populations as perfect gases: genomic integrity and diversification tensions in *Helicobacter pylori*. *Nature reviews. Microbiology*. 2006; 4:826–836. [PubMed: 17041630]
31. Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods*. 2000; 20:4–17. [PubMed: 10610800]
32. Wang G, Wilson TJM, Jiang Q, Taylor DE. Spontaneous Mutations That Confer Antibiotic Resistance in *Helicobacter pylori*. *Antimicrob Agents Chemother*. 2001; 45:727–733. [PubMed: 11181351]
33. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, et al. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*. 1997; 388:539–547. [PubMed: 9252185]
34. Oppenheim DS, Yanofsky C. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. *Genetics*. 1980; 95:785–795. [PubMed: 6162715]
35. Spanjaard RA, van Duin J. Translational reinitiation in the presence and absence of a Shine and Dalgarno sequence. *Nucleic acids research*. 1989; 17:5501–5507. [PubMed: 2668889]
36. Levin-Karp A, Barenholz U, Bareia T, Dayagi M, Zelcbuch L, Antonovsky N, Noor E, Milo R. Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol*. 2013; 2:327–336. [PubMed: 23654261]
37. Smith CA. Structure, function and dynamics in the mur family of bacterial cell wall ligases. *J Mol Biol*. 2006; 362:640–655. [PubMed: 16934839]
38. van Belkum A, Scherer S, van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and molecular biology reviews: MMBR*. 1998; 62:275–293. [PubMed: 9618442]
39. Aras RA, Kang J, Tschumi AI, Harasaki Y, Blaser MJ. Extensive repetitive DNA facilitates prokaryotic genome plasticity. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100:13579–13584. [PubMed: 14593200]
40. Wang G, Ge Z, Rasko DA, Taylor DE. Lewis antigens in *Helicobacter pylori*: biosynthesis and phase variation. *Molecular microbiology*. 2000; 36:1187–1196. [PubMed: 10931272]
41. Huang S, Kang J, Blaser MJ. Antimutator role of the DNA glycosylase mutY gene in *Helicobacter pylori*. *Journal of bacteriology*. 2006; 188:6224–6234. [PubMed: 16923889]
42. Rasko DA, Wang G, Palcic MM, Taylor DE. Cloning and characterization of the alpha(1,3/4) fucosyltransferase of *Helicobacter pylori*. *The Journal of biological chemistry*. 2000; 275:4988–4994. [PubMed: 10671538]
43. Linz B, Windsor HM, McGraw JJ, Hansen LM, Gajewski JP, Tomsho LP, Hake CM, Solnick JV, Schuster SC, Marshall BJ. A mutation burst during the acute phase of *Helicobacter pylori* infection in humans and rhesus macaques. *Nat Commun*. 2014; 5:4165. [PubMed: 24924186]
44. Obmolova G, Ban C, Hsieh P, Yang W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*. 2000; 407:703–710. [PubMed: 11048710]

45. Cromie GA, Connelly JC, Leach DR. Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol Cell*. 2001; 8:1163–1174. [PubMed: 11779493]
46. van der Veen S, Tang CM. The BER necessities: the repair of DNA damage in human-adapted bacterial pathogens. *Nature reviews Microbiology*. 2015; 13:83–94. [PubMed: 25578955]
47. Wang G, Maier RJ. Critical role of RecN in recombinational DNA repair and survival of *Helicobacter pylori*. *Infection and immunity*. 2008; 76:153–160. [PubMed: 17954726]
48. Dorer MS, Fero J, Salama NR. DNA damage triggers genetic exchange in *Helicobacter pylori*. *PLoS pathogens*. 2010; 6:e1001026. [PubMed: 20686662]

Fig. 1. Expression of *mutS2* genes in different *H. pylori* strains

A: Gene organization and the variable sequences of the *mutS2* locus in the *H. pylori* genome. Thick arrows represent the orientation and extent of ORFs, for which the HP numbers (referring to the 26695 genome sequence) and the gene names are given. The nucleotide sequences between HP0622 and *mutS2* from different strains are aligned. The TTG labeled with *** at the beginning of the sequences are the translation start codon of *mutS2* annotated for the strain 26695. The stop codons of HP0622 are overlined, and the start codons of *mutS2* that we proposed are underlined. Hyphens are introduced for sequence alignment (in the middle region) and for highlighting the coding frame of the *mutS2* genes (at the right end). **B:** Western blot analysis for MutS2 in *H. pylori* strains. Shown is a representative result from three independent experiments. *H. pylori* cells were grown in an atmosphere containing 4% O₂ to late log phase. Seven mg of each crude cell extract was resolved on a SDS-12.5% polyacrylamide gel followed by transfer onto a nitrocellulose membrane, and the proteins were detected using anti-MutS2 antiserum and anti-UreA antiserum (as an internal loading control). Lane 1: Purified MutS2 protein; Lane 2: 26695 WT cells; Lane 3: 26695 *mutS2* cells; Lane 4: UA948 WT cells; Lane 5: UA948 *mutS2* + P_{rdxA}*mutS2* cells; Lane 6: SS1 WT cells.

Fig. 2. Identification of specific DNA binding domain of HpMutS2

A: Schematic representation of HpMutS2 and truncated constructs. Shown on the top are known functional domains and a linker region with unknown function. Different truncated HpMutS2 gene constructs were prepared by PCR and proteins were purified as described in Table 1. The names of proteins and their corresponding amino acid numbers are shown. **B and C:** Formation of specific MutS2-DNA complex analyzed by gel shift assay. HpMutS2 variant proteins (1 µg/ml) were incubated with 10 nM ³²P-labelled DNA substrates (GO:C or Holiday Junction). The protein-DNA complexes and the free DNA substrates are marked. “NO” represents the control with no protein.

Fig. 3. Specific DNA binding ability of HpMutS2 truncated variants

A: Schematic representation of HpMutS2 domains and truncated constructs. SDB represents the newly identified Specific DNA-Binding domain, and the corresponding amino acid sequence motif responsible for specific DNA-binding is shown at bottom. **B:** HpMutS2 truncated variant proteins (C8, C9, C10, C12, C23) were overexpressed in *E. coli* and purified to near homogeneity as analyzed on a 20% SDS-PAGE. **C:** A mobility shift assay. One μM HJ DNA substrate was incubated with 1 μM of each HpMutS2 truncated variant protein followed by analysis on a 1.5% agarose gel. The bands of protein-DNA complex and the free DNA substrates are marked. "NO" represents the control with no protein. **D:** Quantification of the specific DNA binding ability of each HpMutS2 truncated variant protein determined by the filter retention assay. ^{32}P labelled DNA substrate (GO:C or HJ, 10 nM) was incubated with different proteins (100 nM). Protein-DNA complexes retained on the filter were quantified with a PhosphorImager. Error bars indicate the standard deviation of three independent experiments. **E:** Alignment of MutS2 sequences (at the identified SDB domain in HpMutS2) from various bacterial species (retrieved from NCBI database, with accession numbers in parentheses). The sequence motif "LDLRG" is one of the most

conserved region among Smr sequences, wherein the aspartic residue is involved in the endonuclease activity.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4. Specific vs non-specific DNA binding ability

A & B: Detection of protein-DNA complex formation analyzed by gel shift assay. One μ M DNA substrates (normal dsDNA, GO:C-containing DNA, or Holiday Junction) was incubated with HpMutS2-N31 (**A**) or HpMutS2-C12 protein (**B**) at indicated concentrations followed by analysis on 1.5% agarose gel. The bands of protein-DNA complex (C) and the free DNA substrates (F2, dsDNA; F4, four stranded DNA) are marked. **C:** Competition assay. 100 nM HpMutS2-C12 was incubated with 10 nM ^{32} P-labeled GO:C DNA substrate and varying concentrations (10, 100, 500 nM) of unlabeled G:C, GO:C, or HJ competitor DNAs, and then the amount of protein-DNA complex was determined using PAGE gel shift assays. The data (% GO:C bound) are averages of three experiments with standard deviations.

Table 1

Construction of Hp MutS2 variant (truncation) proteins:

Name	Primer Sequences (5'–3') ^a used for PCR		Length (AA residues)	MW (kDa)
	Forward	Reverse		
W85	ATGATGAATAATAATAAT	CAATTTAACGATTTT	749	85.35
N31	ATGATGAATAATAATAAT	TTTAGCGAAATTAAG	275	31.46
N62	ATGATGAATAATAATAAT	TTCTTTTGTTTTTC	547	62.56
C23	ATGGAAAAACAAAAGAA	CAATTTAACGATTTT	208	23.59
C12	TTAAAAGCGCATGGG	CAATTTAACGATTTT	110	12.04
C10	CAAAAACCCCTAAA	CAATTTAACGATTTT	99	10.85
C9	ATTCCTAACCTAAA	CAATTTAACGATTTT	85	9.23
C8	GAAGCGAGCTTGCGC	CAATTTAACGATTTT	80	8.66

^a Actual primers include additional nucleotide sequences at 5' end that match the pET21a sequence at the specific cloning sites.

Table 2DNA sequence variations at HP0622-MutS2 junctions among *H. pylori* strains ^a

Strains	# of CnAATGAA repeat units ^b	# of C's ^c within the CnAATGAA repeats	Space (bp) ^d between HP0622 stop and MutS2 start codon
26695	1	3	0
Rif2	1	3	0
2017	1	3	0
908	1	3	0
GAMchJs124i	1	2	0
J99	1	2	0
J166	1	3	0
B8	1	3	0
Cuz20	1	4	0
PNG84A	1	3	0
HUP-B14	1	4	0
Sat464	1	4	0
K26A1	1	3	0
ausabrJ05	1	2	0
SouthAfrica20	1	3	0
MKM1	1	4	0
HPAG1	2	4, 6	0
BM013A, BM013B	2	4, 5	0
UA1182	2	5, 5	0
SS1	2	5, 5	0
I1639	2	4, 6	0
India7	3	5, 4, 4	0
G27	2	5, 4	-2
SJM180	2	5, 4	-2
UA802	1	3	4
Oki898, Oki112, Oki422	1	3	4
MKM5	1	4	6
F17	1	4	6
NY40	1	2	6
Gambia94/24	1	2	6
Shi112	1	4	6
Shi169, Shi417, Shi470	2	4, 3	6
UA1195	2	4, 4	6
F55	2	4, 4	6
UM037	2	5, 4	6
ELS37	2	4, 3	7

Strains	# of CnAATGAA repeat units ^b	# of C's ^c within the CnAATGAA repeats	Space (bp) ^d between HP0622 stop and MutS2 start codon
SNT49	2	5, 5	7
P12	2	5, 5	7
Aklavik86	2	4, 6	8
GAMchJs114i	1	3	12
SouthAfrica7	1	3	15
CC33C	2	4, 3	14
PeCan18	2	4, 3	14
29CaP	2	4, 3	14
UA948	3	4, 4, 4	16
F24, F63	3	4, 5, 4	17
Oki128	2	7, 4	18
Oki102	3	4, 6, 4	18
MKM6	3	4, 6, 4	18
Hp238	3	4, 6, 4	18
F21, F38, F67, F70, F90,	3	4, 6, 4	18
F209			
ML1, ML3	3	4, 6, 4	18
UM032, UM066, UM298,	3	4, 6, 4	18
UM299			
51	3	4, 6, 4	18
MKF3	3	4, 7, 4	19
OK310	3	4, 7, 4	19

^aThe sequences are retrieved mainly from a public database (NCBI website, genome sequences). DNA sequences covering the region of HP0622-MutS2 in a few UA strains were determined by us, and are not available in public databases.

^b“CnAATGAA” repeat sequence may not be perfect in all strains. If the number is 1, that means there is no repeat (for example, see the sequence of strain 26695 in Fig. 1A).

^cThe number of C's within each “CnAATGAA” repeat unit are given. For the cases that the number of C's is less than 4, they may not be considered as polymeric repeats subjected to slipped strand mispairing mechanism.

^dThe number 0 means “overlap” between HP0622 stop and MutS2 start codon. A negative number means that MutS2 start codon is in front of HP0622 stop codon. According to this Table, in about 40% of strains the HP0622 stop codon and MutS2 start codon are overlapped. There is a space of <10 bp between the two codons in about 30% of strains, and a space of >10 bp in another 30% of strains.

Table 3

MutS2-related phenotypes of the strain UA948

Strain phenotype	Mutation rate under O ₂ stress ^a	Transformation frequency ^b
UA948 WT	126 ± 19	22.4 ± 12.6
UA948 mutS2	135 ± 23	24.1 ± 13.2
UA948 mutS2 + P _{rdxA} -MutS2	16.1 ± 5.8	3.8 ± 2.7

^a Cell suspensions were exposed to air for 4 hours before plating to assess mutation rate. Mutation rates were determined as mutation to rifampicin resistance per cell division (x 10⁸). Data were calculated from an experimentally determined median number of mutants from 20 independent cultures with standard deviations.

^b The transformation frequencies are presented as the number of transformants (rifampicin resistant colonies) per 10⁵ recipient cells. Data are means ± standard errors from three independent determinations.

Table 4

Mutation rates and transformation frequencies

Strain phenotype	Mutation rate under O ₂ stress ^a	Transformation frequency ^b
26695 WT	15.2 ± 3.2	5.8 ± 3.6
26695 mutS2	113 ± 28	35.7 ± 15.4
26695 mutS2 + P _{rdxA} -W85	12.5 ± 3.8	4.3 ± 1.7
26695 mutS2 + P _{rdxA} -N62	105 ± 31	32.6 ± 9.5
26695 mutS2 + P _{rdxA} -C23	98 ± 22	28.7 ± 13.3
26695 mutS2 + P _{rdxA} -C12	93 ± 26	30.3 ± 12.7

^a and ^b Same as the legend of Table 3.