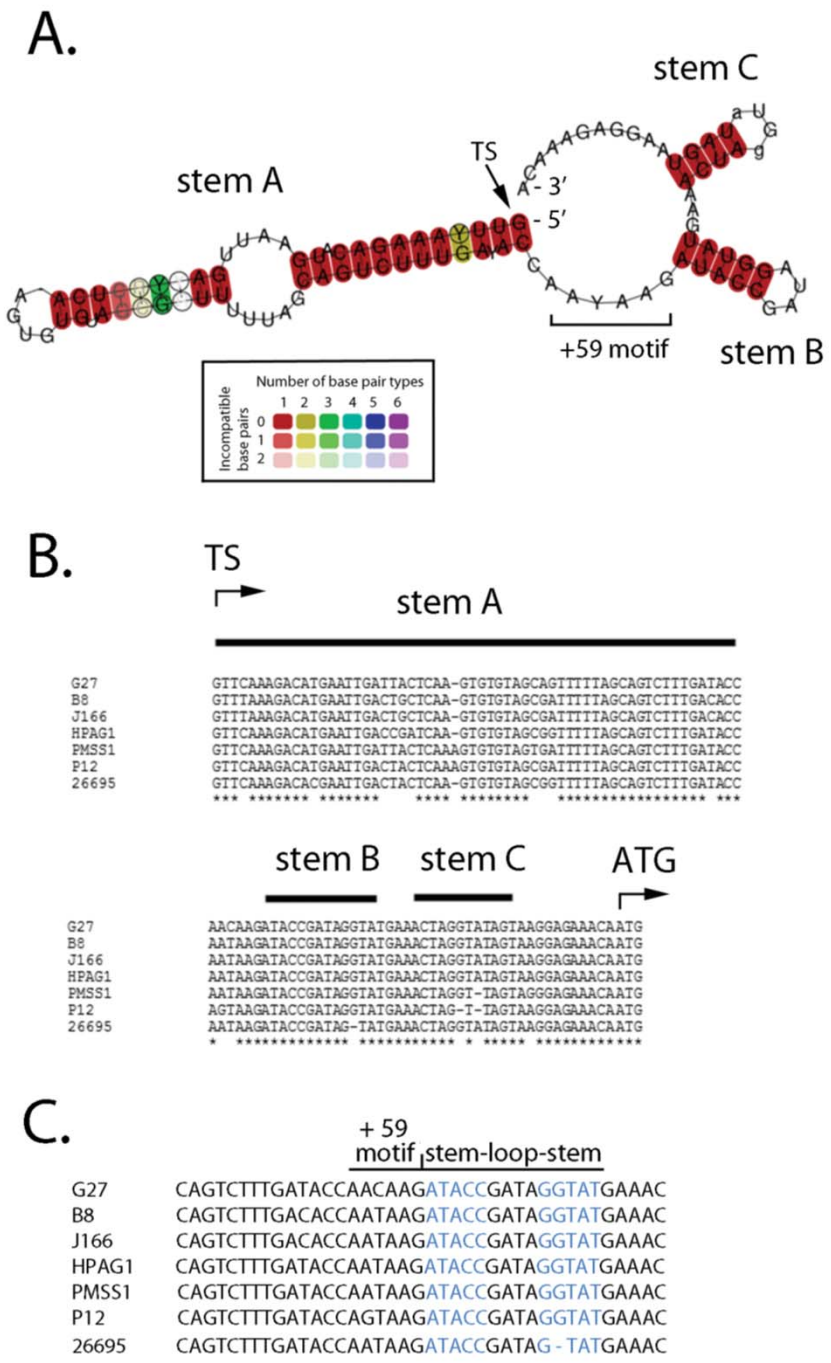

Figure S1. Analysis of the *cagA* 5'-UTR reveals three predicted stem-loop structures. (A) Consensus RNA secondary structure of the *cagA* 5'-UTR, based on analysis of 7 commonly used *H. pylori* strains that contain the strain-specific +59 motif (G27, B8, J166, HPAG1, PMSS1, P12, and 26695). The consensus secondary structure was generated by the Freiburg Tools online webserver LocARNA, a tool for alignment of multiple RNA sequences (Supplemental Reference 1), using default parameters. The RNA region analyzed begins at the *cagA* transcriptional start site (TS) and ends at the nucleotide before the ATG translational start site. Three predicted stem-loop structures (labeled stem A, stem B, and stem C) are shown. The +59 motif is also labeled. The color coding at each position reflects the sequence conservation of 6 compatible base pairing types (C-G, G-C, A-U, U-A, G-U or U-G) at the indicated positions among 7 *H. pylori* strains. For example, dark red indicates the presence of only 1 compatible base pair type in all 7 strains, while green indicates the presence of 3 different compatible base pairs at the indicated position. The presence of incompatible base pairs at the same position decreases color saturation. The color legend provides information on the number of incompatible and compatible base pairs. (B) ClustalW alignment of the nucleotides forming stem A, B and C in Panel A. (C) Nucleotides predicted to form stem B are shown in blue, and the adjacent +59 AATAAG motif is also labeled.

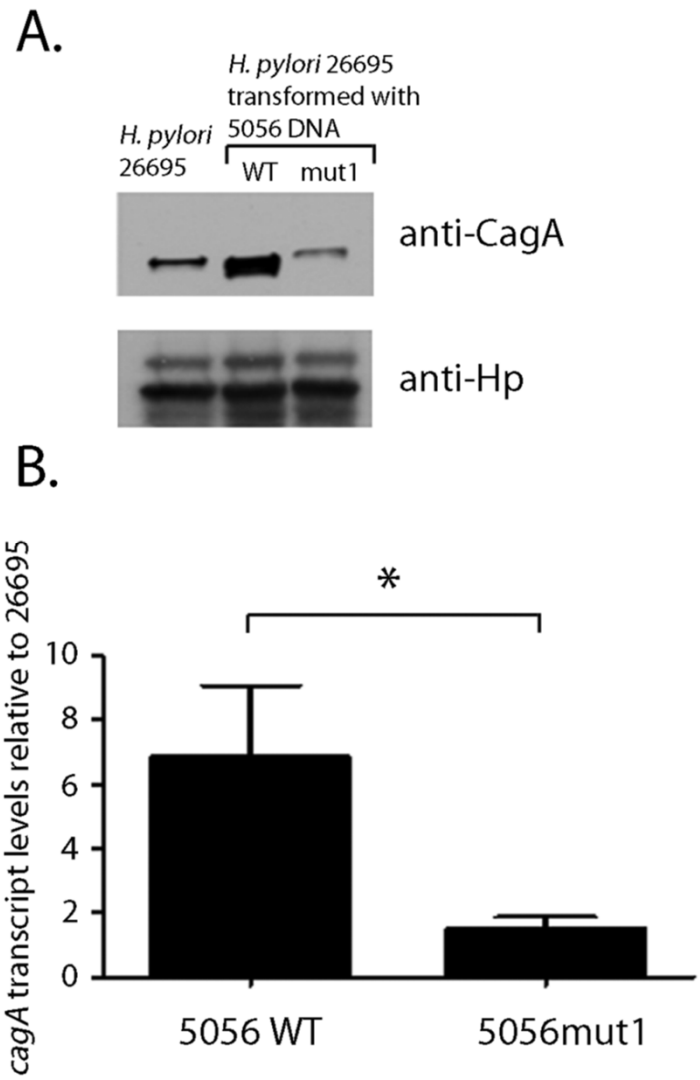

Figure S2. Effect of the mut1 mutation on *cagA* expression, analyzed *in H. pylori* strain 26695. (A) *H. pylori* 26695 strains harboring *cagA* 5' UTR sequences from PZ5056 (either WT or the mut1 sequence) were cultured as described in the Methods, and cell extracts (5 µg) were immunoblotted to detect CagA. Blots were then stripped with stripping buffer (Pierce), and reprobed with a polyclonal antiserum against *H. pylori* soluble proteins (anti-HP, 1:10,000) (Supplemental Reference 2). (B) *cagA* transcript levels in strain 26695 derivatives containing the indicated *cagA* UTR sequences were monitored by real-time PCR as described in the Methods. To quantify the relative expression of *cagA* in each strain, *cagA* specific signals were first normalized to 16S rRNA. The relative level of *cagA* expression (relative to WT strain 26695) was then calculated by dividing the normalized *cagA* value for each strain with that obtained for the WT *H. pylori* 26695 strain. All data points represent the results from analyses of 3 independent biological samples. The mean + SEM are shown. Statistical significance was analyzed with the Mann-Whitney test (* indicate $p<0.05$).

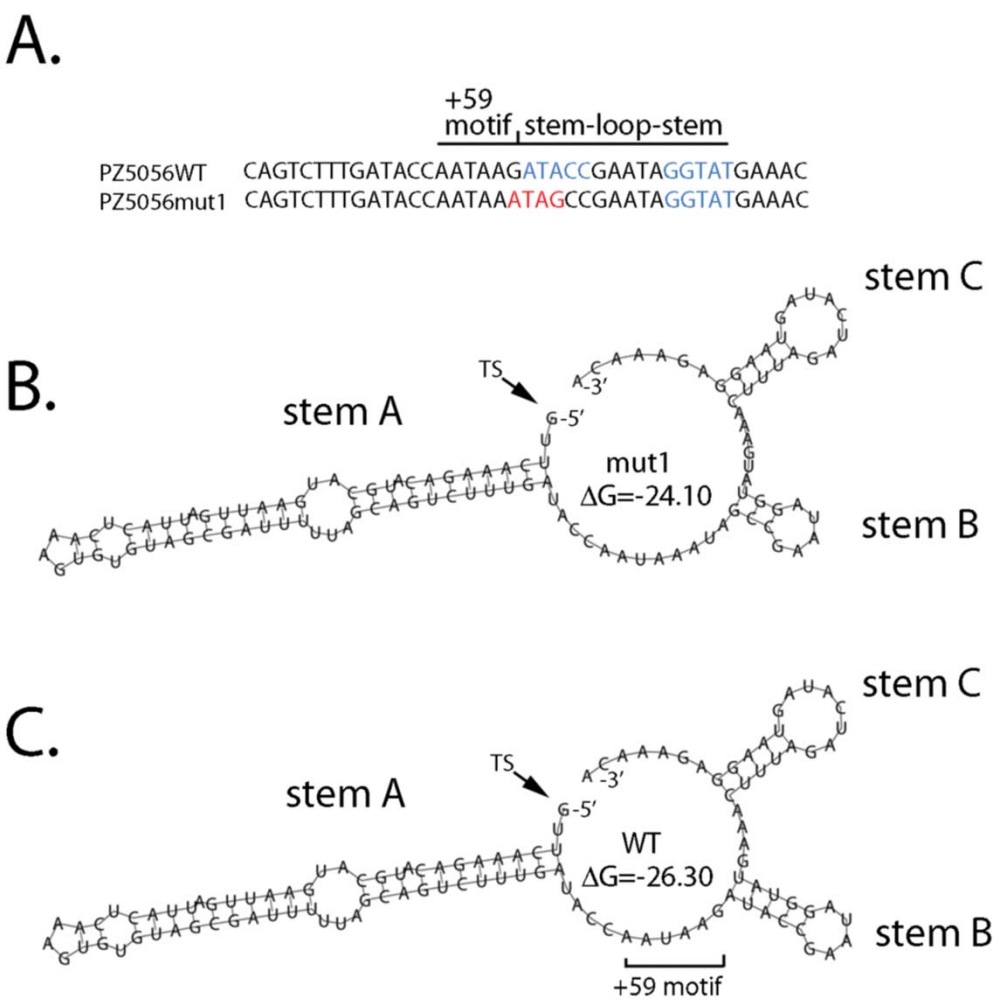

Figure S3. The mut1 mutation alters stem B in the *cagA* UTR. (A) Previous work showed that a mut1 mutation in the 5' UTR of *cagA* altered *cagA* expression (Supplemental Reference 3). The mut1 mutation (red font) alters several nucleotides that form a portion of stem B. The nucleotides forming the stem structure of stem-loop B in the *cagA* 5'-UTR of wild-type strain PZ5056 are shown in blue. (B) RNAfold (Supplemental Reference 4) was used to analyze the secondary structure of the PZ5056 *cagA* 5'-UTR containing the mut1 mutation. The region analyzed begins at the *cagA* transcriptional start site (TS) and ends at the nucleotide before the ATG translational start site. A minimum free energy (ΔG) value of -24.1 was predicted by RNAfold for the mut1 structure. (C) The predicted secondary structure of the *cagA* 5' UTR in WT strain PZ5056, which expresses high levels of CagA (Supplemental Reference 3). Three predicted stem-loop structures (labeled stem A, stem B and stem C) are similar to those predicted for the reference strains in Supplemental Figure S1. The length of stem B is different when comparing panels B and C. A minimum free energy (ΔG) value of -26.3 was predicted by RNAfold for the WT structure.

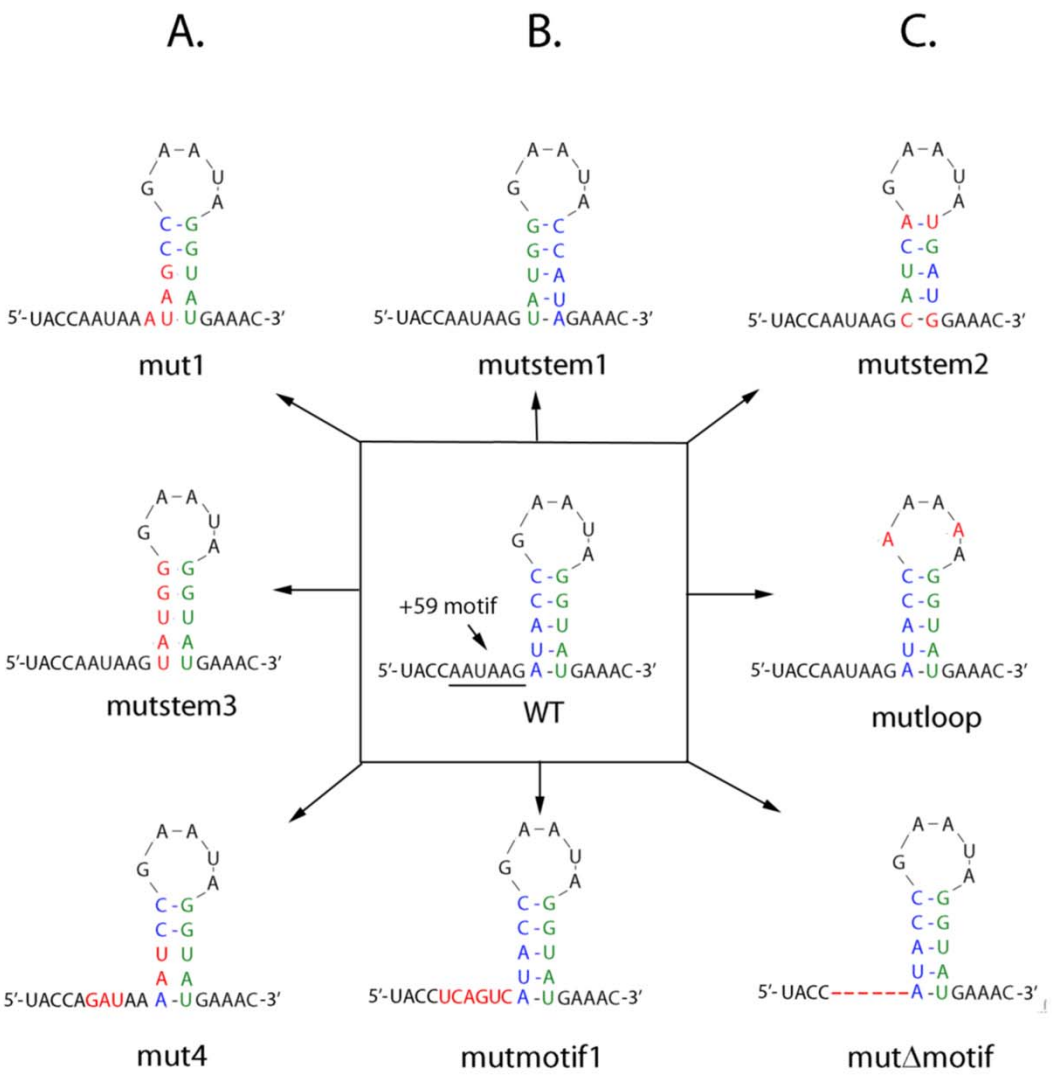

Figure S4. Schematic of stem loop B in the *cagA* 5′ UTR of *H. pylori* strain PZ5056, illustrating the mutations analyzed in this study. The mutations shown involve nucleotides changes to the +59 motif or stem B.

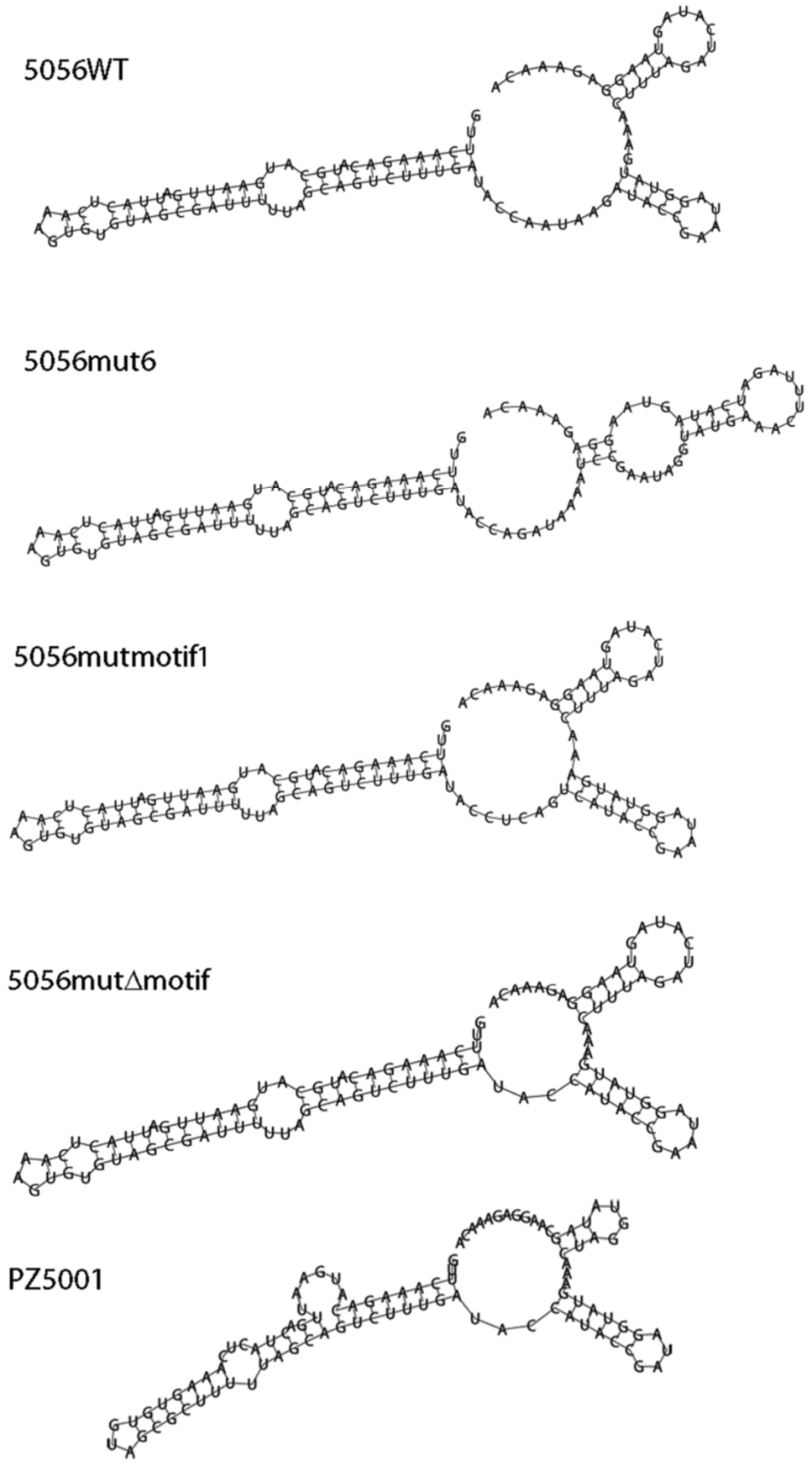

Figure S5. Predicted effects of the indicated mutations on secondary structure of the *cagA* 5' UTR in strain PZ5056. The predicted secondary structure of the *cagA* 5' UTR from wild-type strain PZ5001 is shown for comparison

Supplemental References

1. Smith, C., S. Heyne, A. S. Richter, S. Will, and R. Backofen. 2010. Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. Nucleic Acids Res 38:W373-377.
2. Cao, P., M. S. McClain, M. H. Forsyth, and T. L. Cover. 1998. Extracellular release of antigenic proteins by *Helicobacter pylori*. Infect Immun 66:2984-2986.
3. Loh, J. T., C. L. Shaffer, M. B. Piazuelo, L. E. Bravo, M. S. McClain, P. Correa, and T. L. Cover. 2011. Analysis of *cagA* in *Helicobacter pylori* strains from Colombian populations with contrasting gastric cancer risk reveals a biomarker for disease severity. Cancer Epidemiol Biomarkers Prev 20:2237-2249.
4. Gruber, A. R., R. Lorenz, S. H. Bernhart, R. Neubock, and I. L. Hofacker. 2008. The Vienna RNA websuite. Nucleic Acids Res 36:W70-74.