Cloning and Sequence Analysis of *himA/hup_2* gene of *Mycoplasma gallisepticum*

To amplify the *hup_2* gene from *M. gallisepticum* genome, the PCR primers (forward, NdeI site is underlined: atgcacatatgtttattatggcaaaaatcaaatc; reverse, BamHI site is underlined: atgcggatccctatttgtgcgaatctac) were designed based on the *hup_2* gene sequence obtained from the annotated genome data of *M. gallisepticum*. *M. gallisepticum* genomic DNA was used as a template for the PCR [conditions: 98^{o}C, 10 s denaturation; 55^{o}C, 10 s annealing; 72^{o}C, 30 s extension; 30 cycles, Taq polymerase (Lytech LTD, Russia). The PCR product was digested with NdeI and BamHI, ligated with NdeI-BamHI digested pET15b vector (Novagen), and transformed into E. coli DH5α cells. The nucleotide sequences were determined for both strands using an ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit and ABI Prism 3100 sequencer (Applied Biosciences, U.S.A.) and was found the same for five clones. The DNA and amino acid sequence data analyses were then performed using Vector NTI program (Invitrogen, USA). The recombinant plasmid was named pET15b-hin.

Expression and Purification of his-tagged hup_2 from *E. coli*

For the expression and purification of the protein, the pET15b-hin plasmid was transformed into the *E. coli* strain B834 (DE3) (Life Technologies, USA). The *E. coli* B834(DE3)/ pET15b-hin was then incubated in 300 ml of a Luria-Bertani medium containing 100 µg/ml ampicillin at 37^{o}C for optical density at 600 nm ~0.8, IPTG was added to final concentration of 0.1mM to induce expression, and the cells were incubated for additional 3 h. Cells were then disrupted by sonication (conditions: pulse-on 1 s, pulse-off 1 s, 4℃, 3 min) and the expressed hup_2 was purified using HisTrap

histidine tagged proteins purification kit (Amersham Biosciences, USA) according to the manufacturer's protocol with minor modifications. Protein concentration was measured using the Bradford method (Bio-Rad Laboratories, U.S.A.).

Analysis of the purified his-tagged protein

The protein purity was estimated by SDS-gel as > 95%, protein migrates as. 16 kDa polypeptide (Fig. S1A). Sequencing of constructed pET15b-hin reveals the following nucleotide sequence of transcribed insert in hup_2 expressing vector pET15b-hin

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatgtttattatggcaaaaat
caaatcattaagtgctgctgaatatcttaaagaaatggcagacgaaactaacattaaggttcaagatatccgtttagttgtta
cttctttacaaaaagtattagctaaagaactagctactactggtgaagtaagattatttgatattggtaagttcaaattagttaca
actaaacctcgaactggaatcaaccctaaaaccaaacaaaagattcagatcccagcagggaagaaaatcaaactaac
tgtttcaaagatcttaactgatgcagtagattcgcacaaatag

underlined sequence is from the vector. The predicted protein sequence is

MGSSHHHHHHSSGLVPRGSHMFIMAKIKSLSAAEYLKEMADETNIKVQDIRLVVTSLQK
VLAKELATTGEVRLFDIGKFKLVTTKPRTGINPKTKQKIQIPAGKKIKLTVSKILTDAVDSH
K

with molecular weight of 13 517. The underlined sequence belongs for the his-tag.

To verify the purity and the amino acid composition, the MALDI mass-spectrometric analysis was carried out for the purified protein without proteolytic digestion (Fig. S1B). The mass-spectrum reveals three major peaks of 4463, 6695, and 13 386 Da corresponding to a single polypeptide with charges 3, 2 , and 1 e, respectively. The measured MW of the polypeptide is 13 386 Da. This MW is smaller than that computed for the sequence by the mass of one methionine residue. We assumed, therefore, that the first Met of the polypeptide is absent. Indeed, in this case its calculated MW is 13

386, in perfect agreement with the MS result. No traces of *E. coli* HUs of MW 9 226 (HU-beta) and 9 535 (HU-alpha) as well as other proteins are seen.

Panels A and B of Fig. S2 compare the gel mobility of two representative DNA substrates complexed with three proteins, namely, the hig-tagged mgHU, the untagged protein in *M. gallisepticum* extracts, and *E. coli* HU. As expected, the apparent MW of his-mgHU is somewhat larger than the other two. It is known that *E. coli* HU binds to DNA as a dimer with the total MW about 19 kDa. The pattern in Fig. S2B suggests that mgHU also binds dsDNA as a dimer (with MW about 27 kDa).

PT-PCR

RNA was isolated with TRIzol® reagent (Invitrogen) according to the manufacturer manual and treated with DNaseI. mRNA was copied to cDNA by reverse transcriptase using random primers. RT-PCR was performed using the Brilliant SYBR Green. Cycling was performed on the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems). The primers used are listed in Table S1. *hup_1* and *hup_2* gene expression was compared with the expression of the house-keeping gene *gap* (encoding Glyceraldehyde-3-phosphate dehydrogenase). For the latter, the Ct value is 22.1 cycles, for *hup_2* Ct is 22.3 cycles, while for *hup_1* Ct is 25.4. Thus, both hup genes of *M. gallisepticum* are highly expressed, while the expression of hup_2 (mgHU that recognizes DNA mismatches) is 8-fold higher than hup_1 expression.

Table S1. Primers used for PCR

Gene	Direction	Sequence (5'-3')
himA/hup1 (MGA_0869)	Forward	AGGCGGAATTCTAGTTTGTG
	Reverse	ATTATCGCTGAATGTACTGGAG
himA/hup2 (MGA_0061)	Forward	ATTTGTGCGAATCTACTGCA
	Reverse	AATGGCAGACGAAACTAACC
gap (MGA_1186)	Forward	GATGTAGAAGTTGTAGCCGT
	Reverse	ACTACTTTCTTAGCACCACCT

Legends to Supplementary figures.

Figure S1. Purification of the his-tagged mgHU expressed in *E. coli*. Panel A. Coomassie stained SDS-gel, lanes 1 and 2, *E. coli* lysate before and after 3 hrs IPTG induction, respectively; lane HU, purified recombinant mgHU; lane M, MW markers (in kDa). Panel B. MALDI-MS analysis of intact recombinant mgHU. The three peaks correspond a 13 366 Da polypeptide with electric charges of 3, 2, and 1 e. No traces of other proteins are seen.

Figure S2. Comparison of DNA binding profiles observed for the refined his-tagged mgHU, purified *E. coli* HU (ecoHU) and the cell-free extracts of *M. gallisepticum* (mg extract). Panel A compares the gel mobilities of the CC mismatch DNA substrate complexed with the two protein preparations (150 mM NaCl). Panel B compares the gel mobilities of the dsDNA substrate complexed with the two proteins (36 mM NaCl).

Figure S3. The binding pattern of mgHU to DNA bulges of different length and sequence. The *M. gallisepticum* extract (0.25 µl per lane) was incubated with a series of labeled dsDNA constructs (15 nM) carrying inserts of seven or three bases as indicated

below the lanes. X3 and Y3 denote TCC inserted at different positions, V3 and W3 denote GAA inserted at different positions. For comparison, the last three lanes show the mgHU binding to substrates with single substitution mismatches. All binding assays were carried out in 150 mM NaCl.

Figure S4. The binding of gene products hup_1 and hup_2 (mgHU) with the DNA bulge A7. Different concentrations (indicated at the bottom in μM) of the two recombinant proteins were incubated with DNA in 50 mM NaCl and the the resulting DNA-protein complexes were gel-separated in 100 mM Tris-borate. The notation is similar to Figure 1.

Figure S5. Analysis of RT-PCR products on ethidium bromide stained agarose gels. Lanes 2, 4, and 6 display the results for *himA/hup_1* (261 bp), *himA/hup_2* (254 bp), and *gap* (278 bp), respectively. Lanes 1, 3, and 5 represent negative controls (no RT). Lane M shows a DNA ladder (100 bp plus).