MgaSpn and H-NS: Two Unrelated Global Regulators with Similar DNA-Binding Properties

Virtu Solano-Collado[1†], Mário Hüttener[2], Manuel Espinosa[1], Antonio Juárez[2,3*] and Alicia Bravo[1*]

[1] Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain, [2] Departament de Microbiologia, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain, [3] Institut de Bioenginyeria de Catalunya, Barcelona, Spain

Global regulators play an essential role in the adaptation of bacterial cells to specific niches. Bacterial pathogens thriving in the tissues and organs of their eukaryotic hosts are a well-studied example. Some of the proteins that recognize local DNA structures rather than specific nucleotide sequences act as global modulators in many bacteria, both Gram-negative and -positive. To this class of regulators belong the H-NS-like proteins, mainly identified in γ-Proteobacteria, and the MgaSpn-like proteins identified in Firmicutes. H-NS and MgaSpn from *Escherichia coli* and *Streptococcus pneumoniae*, respectively, neither have sequence similarity nor share structural domains. Nevertheless, they display common features in their interaction with DNA, namely: (i) they bind to DNA in a non-sequence-specific manner, (ii) they have a preference for intrinsically curved DNA regions, and (iii) they are able to form multimeric complexes on linear DNA. Using DNA fragments from the hemolysin operon regulatory region of the *E. coli* plasmid pHly152, we show in this work that MgaSpn is able to recognize particular regions on extended H-NS binding sites. Such regions are either located at or flanked by regions of potential bendability. Moreover, we show that the regulatory region of the pneumococcal *P1623B* promoter, which is recognized by MgaSpn, contains DNA motifs that are recognized by H-NS. These motifs are adjacent to regions of potential bendability. Our results suggest that both regulatory proteins recognize similar structural characteristics of DNA.

Keywords: global transcriptional regulators, nucleoid-associated proteins, Mga/AtxA family, protein-DNA interactions, DNA bendability

INTRODUCTION

Global modulators play key roles in the ability of bacterial cells to rapidly adapt to environmental fluctuations by adjusting their gene expression pattern. As a consequence, they enable the pathogenic bacteria to colonize and survive in different niches of their eukaryotic hosts. Whereas, some global modulators recognize specific DNA sequences, others exhibit a preference for particular DNA structures. Examples of the latter group are the H-NS-like proteins, mainly found in γ-Proteobacteria, and the MgaSpn-like proteins identified in Firmicutes.

In *Escherichia* and *Salmonella*, the DNA-binding properties of the nucleoid-associated protein H-NS (137 amino acids) have been studied in detail (for a review see Winardhi et al., 2015).

Like other nucleoid-associated proteins, H-NS is involved in both organization of the bacterial chromosome and regulation of gene expression. H-NS functions generally as a repressor or gene silencer, but it can also act indirectly as a transcriptional activator (Ko and Park, 2000). Many genes regulated by H-NS encode virulence determinants. H-NS consists of an N-terminal oligomerization domain and a C-terminal DNA-binding domain. In solution, H-NS is able to form higher-order oligomers. This ability correlates with its ability to form nucleoprotein filaments (Lim et al., 2012). Imaging studies have revealed that H-NS-like proteins are able to organize large DNA molecules into various conformations, including extended nucleoprotein filaments, hairpin-like large DNA bridges, and higher-order DNA condensations (Dame et al., 2000; Liu et al., 2010).

In vitro DNA binding experiments have shown that H-NS binds to DNA in a non-specific manner, although it has a strong preference for intrinsically curved AT-rich DNA regions. Furthermore, high-affinity DNA-binding sites for H-NS have been identified in AT-rich regions of the chromosomal DNA (Lang et al., 2007). Formation of H-NS nucleoprotein filaments from such high-affinity sites (nucleation sites) may lead to selective gene silencing either by inhibiting the binding of the RNA polymerase to the promoter region or by blocking RNA polymerase translocation (reviewed by Winardhi et al., 2015). Studies on the *LEE5* promoter supported an additional model for H-NS-mediated repression. In this model, H-NS spreading from a site located upstream of the *LEE5* promoter to a site located at the promoter would facilitate specific contacts between H-NS and the RNA polymerase (Shin et al., 2012).

H-NS is known to modulate the expression of the thermoregulated hemolysin (*hly*) operon (genes *hlyC*, *hlyA*, *hlyB*, and *hlyD*), which encodes the toxin α-hemolysin and additional gene products required for its activation and export. This toxin is produced by several uropathogenic *E. coli* strains. The *hly* operon of the *E. coli* plasmid pHly152 has been studied in detail. First, Vogel et al. (1988) identified an essential regulatory sequence located ∼2 kbp upstream of the *hly* operon, the so-called *hlyR* sequence. Between the *hlyR* sequence and the promoter region of the *hly* operon there is an IS*2* insertion element. Subsequently, Madrid et al. (2002) identified two extended H-NS binding sites upstream of the *hly* operon (see **Figure 1A**). One of them (site I; nucleotides 190 to 350 of plasmid pHly152) is located within the *hlyR* sequence. The second one (site II; nucleotides 2180 to 2330) overlaps the promoter region of the *hly* operon. A deletion analysis confirmed the relevance of site I for thermoregulation of the *hly* operon (Madrid et al., 2002).

The Mga*Spn* transcriptional regulator (493 amino acids) contributes to the virulence of *Streptococcus pneumoniae*. Some DNA-binding properties of Mga*Spn* resemble the ones reported for H-NS. Specifically, *in vitro* DNA binding studies (gel retardation, footprinting and electron microscopy) have shown that Mga*Spn* generates multimeric complexes on linear

---

Abbreviations: BSA, bovine serum albumin; EMSA, electrophoretic mobility shift assays; *hly*, hemolysin; PCR, polymerase chain reaction; T4 PNK, T4 polynucleotide kinase.

double-stranded DNA. Furthermore, Mga*Spn* binds to DNA with little or no sequence specificity, and shows a preference for DNA regions that contain a potential intrinsic curvature (Solano-Collado et al., 2013). Nevertheless, despite this fact, Mga*Spn* and H-NS are unrelated proteins. They neither have sequence similarity nor share structural domains. Mga*Spn* is a member of a new class of global response regulators known as the Mga/AtxA family, which includes the Mga, AtxA and MafR proteins from *S. pyogenes*, *Bacillus anthracis* and *Enterococcus faecalis*, respectively (Hondorp et al., 2013; Hammerstrom et al., 2015; Ruiz-Cruz et al., 2016). According to the Pfam database (Finn et al., 2016), Mga*Spn* has two putative N-terminal DNA-binding domains, the so-called HTH_Mga (residues 6 to 65) and Mga (residues 71 to 158) domains (Solano-Collado et al., 2012). These domains do not exhibit sequence similarity with the C-terminal DNA-binding domain of H-NS.

Mga*Spn* plays a significant role in both nasopharyngeal colonization and lung infection in mice (Hemsley et al., 2003). *In vivo* experiments showed that Mga*Spn* activates the pneumococcal *P1623B* promoter and, consequently, the transcription of a four-gene operon (*spr1623-spr1626*) of unknown function. This activation requires a 70-bp region (*PB* activation region) located upstream of the *P1623B* promoter (Solano-Collado et al., 2012) (see **Figure 1B**). Interestingly, Mga*Spn* recognizes the *PB* activation region as a primary binding site when it is located at internal position on a 222-bp DNA fragment, but not when it is positioned at one end of the DNA fragment (Solano-Collado et al., 2013). According to the bend.it program (Vlahovicek et al., 2003), the *PB* activation region contains a potential intrinsic curvature flanked by regions of bendability (**Figure 1B**).
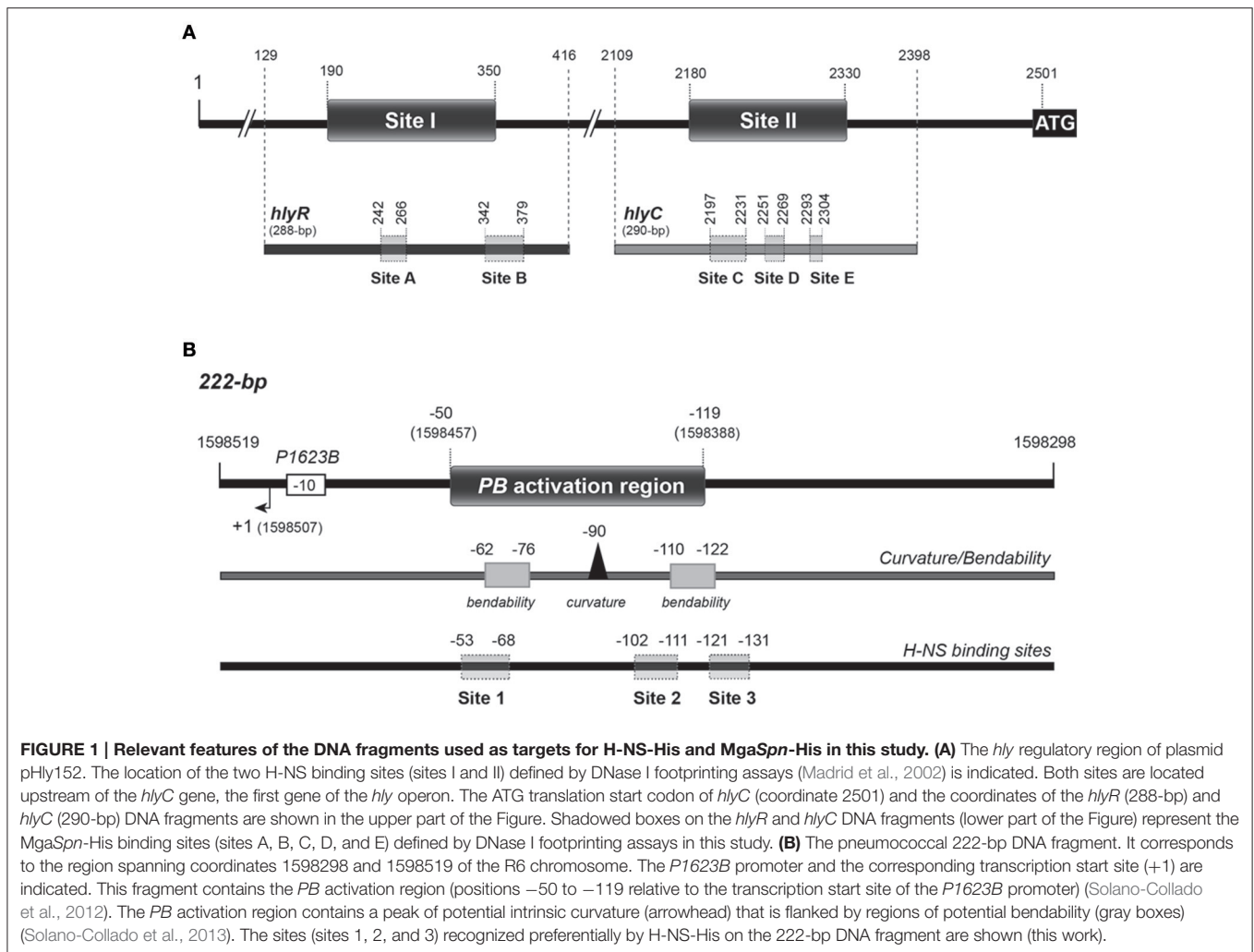
From the information mentioned above, it is apparent that H-NS and Mga*Spn* share some DNA-binding properties. In light of these observations, we hypothesized that both unrelated proteins might recognize similar characteristics of DNA. In this work we present evidence that supports this hypothesis. By gel retardation and DNase I footprinting assays, we show that Mga*Spn* is able to recognize particular regions on extended H-NS binding sites and vice versa.

MATERIALS AND METHODS

Polymerase Chain Reaction (PCR)

The Phusion High-Fidelity DNA polymerase (Thermo Scientific) was used. Reaction mixtures (50 μl) contained 5–20 ng of template DNA, 20 pmol of each primer, 200 μM each deoxynucleoside triphosphate (dNTP), and one unit of DNA polymerase. A initial denaturation step was performed at 98°C for 1 min, followed by 30 cycles that included the next steps: (i) denaturation at 98°C for 10 s, (ii) annealing of the primers around 55°C (depending on the primer melting temperature) for 20–30 s, and (iii) extension at 72°C for 20–40 (depending on the amplicon length). A final extension step was performed at 72°C for 10 min. PCR products were cleaned up with the QIAquick PCR purification kit (Qiagen).

FIGURE 1 | Relevant features of the DNA fragments used as targets for H-NS-His and Mga*Spn*-His in this study. (A) The *hly* regulatory region of plasmid pHly152. The location of the two H-NS binding sites (sites I and II) defined by DNase I footprinting assays (Madrid et al., 2002) is indicated. Both sites are located upstream of the *hlyC* gene, the first gene of the *hly* operon. The ATG translation start codon of *hlyC* (coordinate 2501) and the coordinates of the *hlyR* (288-bp) and *hlyC* (290-bp) DNA fragments are shown in the upper part of the Figure. Shadowed boxes on the *hlyR* and *hlyC* DNA fragments (lower part of the Figure) represent the Mga*Spn*-His binding sites (sites A, B, C, D, and E) defined by DNase I footprinting assays in this study. **(B)** The pneumococcal 222-bp DNA fragment. It corresponds to the region spanning coordinates 1598298 and 1598519 of the R6 chromosome. The *P1623B* promoter and the corresponding transcription start site (+1) are indicated. This fragment contains the *PB* activation region (positions −50 to −119 relative to the transcription start site of the *P1623B* promoter) (Solano-Collado et al., 2012). The *PB* activation region contains a peak of potential intrinsic curvature (arrowhead) that is flanked by regions of potential bendability (gray boxes) (Solano-Collado et al., 2013). The sites (sites 1, 2, and 3) recognized preferentially by H-NS-His on the 222-bp DNA fragment are shown (this work).

PCR-Amplification of DNA Regions

Oligonucleotides used for PCR amplifications are listed in **Table 1**. As DNA templates, chromosomal DNA from the pneumococcal R6 strain (Hoskins et al., 2001) and pANN202312R plasmid DNA from *E. coli* (Godessart et al., 1988) were used. Chromosomal DNA from *S. pneumoniae* was prepared as described previously (Ruiz-Cruz et al., 2010). For small-scale preparations of plasmid DNA, the High Pure Plasmid Isolation Kit (Roche Applied Science) was used. The 222-bp DNA region of the R6 chromosome (coordinates 1598298–1598519) was amplified using the 1622H and 1622I primers. The 288-bp *hlyR* DNA fragment (coordinates 129–416 in Madrid et al., 2002) was amplified using the hlyR-Fw and hlyR-Rev primers. The 290-bp *hlyC* DNA fragment (coordinates 2109–2398 in Madrid et al., 2002) was amplified using the hlyC-Fw and hlyC-Rev primers.

Radioactive Labeling of DNA Fragments

Oligonucleotides were radioactively labeled at the 5′-end using [γ-^{32}P]ATP (PerkinElmer) and T4 polynucleotide kinase (T4 PNK; New England Biolabs). Reactions (25 μl) contained 30

TABLE 1 | Oligonucleotides used in this work.

Name	Sequence 5′ to 3′
1622H	CGGATTAAACCTCTTGCAATTATACC
1622I	CAAATTCTTTAATTGTTGCTATTA
hlyR-Fw	GCCCACTGCATTGAATACTTAC
hlyR-Rev	CATACTTACCTACAGCTATAAG
hlyC-Fw	CAGATAAAGAAGAGTAGTTC
hlyC-Rev	CAAGTTTTTATTGCGCTGACTAAC
Hly Sal/Eco5	CAGACCACACCTGGAAAAAC
Hly Sal/Hind3	GGGCTTCACTGCGAAATTCA

pmol of oligonucleotide, 2.5 μl of 10 × kinase buffer (provided by the supplier), 50 pmol of [γ-^{32}P]ATP (3000 Ci/mmol, 10 mCi/ml) and 10 units of T4 PNK. After incubation at 37°C for 30 min, additional T4 PNK (10 units) was added. Reaction mixtures were then incubated at 37°C for 30 min. The enzyme was inactivated by incubation at 65°C for 20 min. Non-incorporated nucleotide was removed using Illustra MicroSpin^{TM} G-25 columns (GE Healthcare). The 5′-labeled oligonucleotides were used for

manual sequencing and for PCR amplification to obtain double-stranded DNA fragments labeled at either the coding or the non-coding strand.

Purification of Mga*Spn*-His and H-NS-His

Gene *mgaSpn* was engineered to encode a His-tagged Mga*Spn* protein (Mga*Spn*-His). This variant of Mga*Spn* carries six additional His residues at the C-terminal end. The procedure used to overproduce and purify Mga*Spn*-His was reported previously (Solano-Collado et al., 2012). Purified H-NS-His protein was obtained as described by Nieto et al. (1991).

Electrophoretic Mobility Shift Assays (EMSA)

Binding reactions (10 μl) contained 40 mM Tris-HCl, pH 7.6, 1 mM DTT, 0.4 mM EDTA, 1–2% glycerol, 50 mM NaCl, 10 mM MgCl$_2$, 500 μg/ml bovine serum albumin (BSA), 2 nM of 5′-labeled DNA and varying concentrations of Mga*Spn*-His (20 to 180 nM). Reactions were incubated at ambient temperature for 20 min. Free and bound DNA forms were separated on native polyacrylamide (5%) gels (Mini-PROTEAN system, Bio-Rad) using Tris-borate-EDTA, pH 8.3, buffer (TBE). Gels were run at 100 V and ambient temperature. Labeled DNA was visualized using a Fujifilm Image Analyzer FLA-3000 and quantified using the Quantity One software (Bio-Rad).

For competitive EMSA, a 407-bp DNA fragment from the 5′ untranslated region of the *hlyR* gene was generated by PCR using primers Hly Sal/Eco5 and Hly Sal/Hind3 (**Table 1**). For each reaction, 50 ng of the pneumococcal 222-bp DNA and 150 ng of the 407-bp DNA (competitor DNA) were mixed with increasing concentrations of H-NS-His in binding buffer (250 mM HEPES, pH 7.4, 350 mM KCl, 5 mM EDTA, 5 mM DTT, 500 μg/ml BSA, 25% glycerol) and incubated at 37°C for 30 min. Samples (20 μl) were loaded onto native polyacrylamide (5%) gels (TBE buffer). Bands were visualized using a Gel-doc system (Bio-Rad).

DNase I Footprinting Assays

In the case of H-NS-His, binding reactions (10 μl) contained 30 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 mM CaCl$_2$, 10 mM MgCl$_2$, 100 mM NaCl, 1% glycerol, 2 nM ^{32}P-labeled DNA and different concentrations of H-NS-His (2 nM to 500 nM). For Mga*Spn*-His, binding reactions (10 μl) contained 40 mM Tris-HCl, pH 7.6, 1.2 mM DTT, 0.2 mM EDTA, 1 mM CaCl$_2$, 10 mM MgCl$_2$, 50 mM NaCl, 1% glycerol, 500 μg/ml BSA, 2 nM ^{32}P-labeled DNA and different concentrations of Mga*Spn*-His (10 to 200 nM). In all cases, after 20 min at ambient temperature, 0.03 units of DNase I (Roche Applied Science) was added and the reaction proceeded for 5 min at the same temperature. DNase I digestion was stopped by adding 1 μl of 250 mM EDTA. Then, 4 μl of loading buffer (80% formamide, 1 mM EDTA, 10 mM NaOH, 0.1% bromophenol blue and 0.1% xylene cyanol) was added. After heating at 95°C for 5 min, samples were loaded onto 8 M urea-6% polyacrylamide gels. Dideoxy-mediated chain termination sequencing reactions were run in the same gel. Labeled products were visualized using a Fujifilm Image Analyser FLA-3000. The intensity of the bands was quantified using the Quantity One software (Bio-Rad).
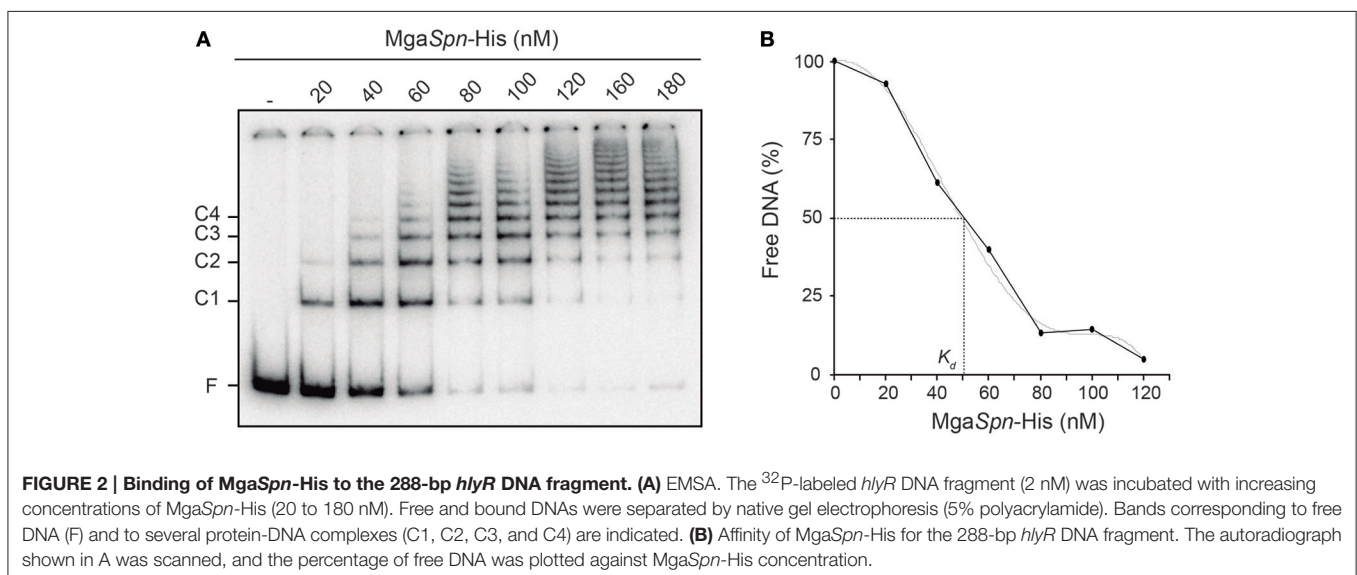
In silico Prediction of Intrinsic Curvature

The bendability/curvature propensity plots of the DNA fragments used in this study were calculated with the bend.it server (Vlahovicek et al., 2003; http://hydra.icgeb.trieste.it/dna/bend_it.html) as described previously (Solano-Collado et al., 2013).

RESULTS

Binding of Mga*Spn*-His to the *hlyR* Region of the *E. coli hly* Operon

In the *E. coli* plasmid pHly152, previous DNase I footprinting assays revealed the existence of two extended H-NS binding sites upstream of the *hly* operon (sites I and II in **Figure 1A**)

FIGURE 2 | Binding of Mga*Spn*-His to the 288-bp *hlyR* DNA fragment. (A) EMSA. The ^{32}P-labeled *hlyR* DNA fragment (2 nM) was incubated with increasing concentrations of Mga*Spn*-His (20 to 180 nM). Free and bound DNAs were separated by native gel electrophoresis (5% polyacrylamide). Bands corresponding to free DNA (F) and to several protein-DNA complexes (C1, C2, C3, and C4) are indicated. **(B)** Affinity of Mga*Spn*-His for the 288-bp *hlyR* DNA fragment. The autoradiograph shown in A was scanned, and the percentage of free DNA was plotted against Mga*Spn*-His concentration.
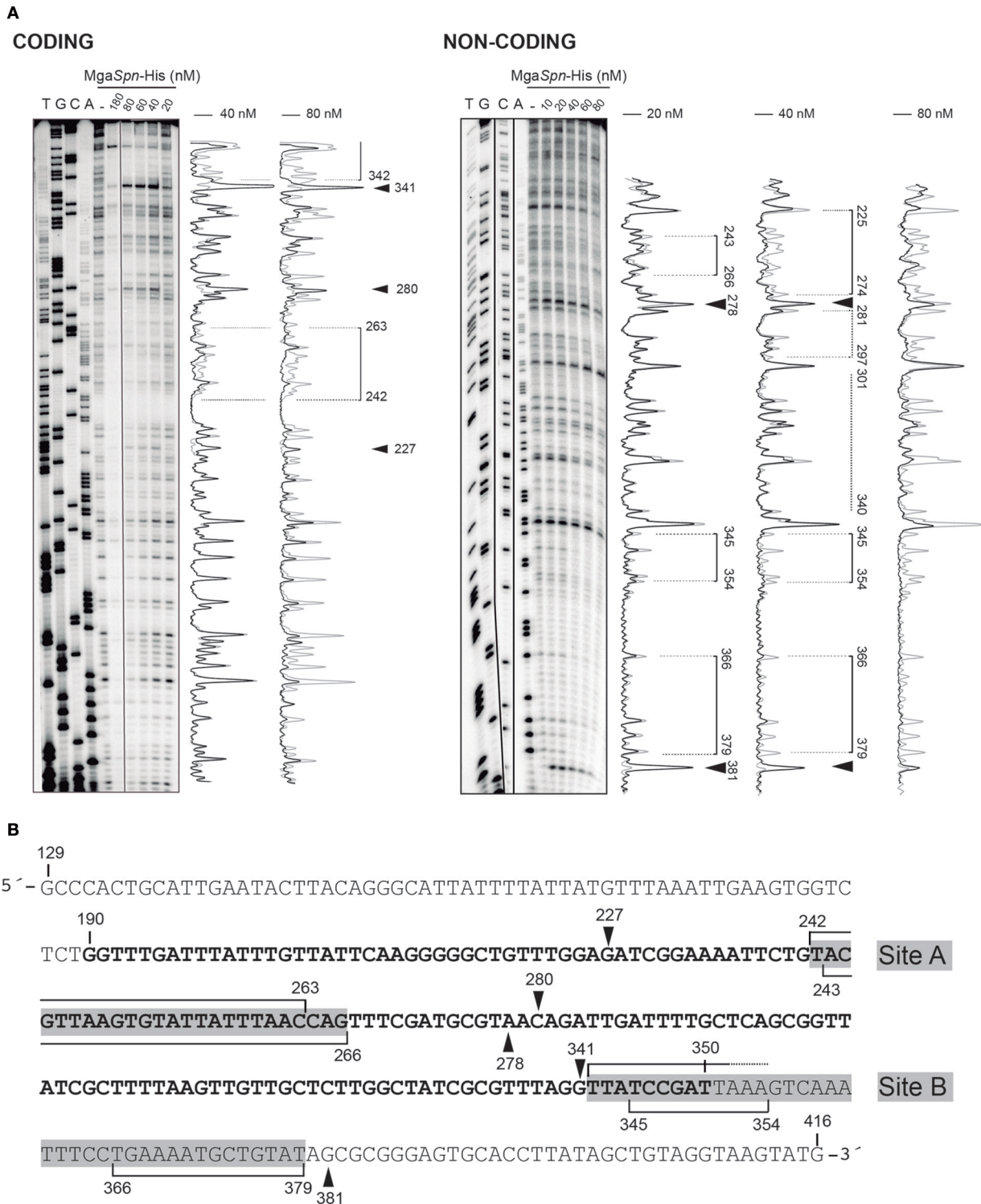
FIGURE 3 | DNase I footprints of complexes formed by MgaSpn-His on the 288-bp *hlyR* DNA fragment. (A) The *hlyR* DNA fragment (coordinates 129–416) was radioactively labeled at the 5′-end of either the coding or the non-coding strand. Labeled DNA (2 nM) was incubated with the indicated concentrations of MgaSpn-His and then digested with DNase I. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). Densitometer scans corresponding to free DNA (gray line) and DNA with protein (black line) are shown. Brackets represent the MgaSpn-His protected regions. Sites more sensitive to DNase I cleavage are indicated with arrowheads. **(B)** Nucleotide sequence of the *hlyR* DNA fragment. The H-NS binding site I (coordinates 190–350) is highlighted in bold. The two sites recognized by MgaSpn-His (sites A and B) are marked with gray boxes. MgaSpn-His protected regions on either the coding or the non-coding strand (brackets) as well as sites more sensitive to DNase I cleavage (arrowheads) are indicated.

(Madrid et al., 2002). One of them (site I; coordinates 190–350 of pHly152) is included within the so-called *hlyR* regulatory sequence (Vogel et al., 1988), and was shown to play a significant role in the thermoregulation of the *hly* operon (Madrid et al., 2002). To investigate whether the pneumococcal MgaSpn protein is able to recognize particular regions on the H-NS binding site I, we performed EMSA and DNase I footprinting experiments. We used a His-tagged version of MgaSpn (MgaSpn-His) and a 288-bp DNA fragment (here named *hlyR*; coordinates 129–416) that contains the site I (**Figure 1A**). The presence of a His-tag at the C-terminal end of MgaSpn does not affect its DNA-binding properties (Solano-Collado et al., 2012, 2013). For EMSA, radioactively labeled DNA was incubated with increasing concentrations of MgaSpn-His (**Figure 2A**). At 20 nM of MgaSpn-His, free DNA and two protein-DNA complexes were detected. However, in agreement with previous results (Solano-Collado et al., 2013), as the concentration of MgaSpn-His was increased, complexes of lower electrophoretic mobility appeared sequentially and complexes moving faster disappeared gradually, indicating that multiple protein units bind orderly on the same DNA molecule (formation of multimeric protein-DNA complexes).

By EMSA, we estimated the affinity of MgaSpn-His for the 288-bp *hlyR* DNA fragment (**Figure 2B**). Since MgaSpn-His generates multiple protein-DNA complexes, the protein concentration required to bind half the DNA was determined by measuring the decrease in free DNA rather than the increase in complexes, which gives an indication of the approximate magnitude of the dissociation constant, Kd (Carey, 1988). Such a concentration was about 50 nM. This value is similar to the apparent Kd of MgaSpn for the pneumococcal 222-bp DNA fragment that harbors the *PB* activation region (MgaSpn binding site) (see **Figure 1B**) (Solano-Collado et al., 2013).

The position of MgaSpn-His on the *hlyR* DNA fragment was further analyzed by DNase I footprinting assays (**Figure 3**). The DNA fragment was labeled either at the 5′-end of the coding strand or at the 5′-end of the non-coding strand. On the coding strand and at 40 nM of MgaSpn-His, protections against DNase I digestion were observed at a particular region (from position 242 to 263) (**Figure 3A**). Diminished cleavages were also observed from position 342 onwards (no resolution in the gel). Moreover, positions 227, 280, and 341 were more sensitive to DNase I cleavage. On the non-coding strand and at 20 nM of MgaSpn-His, diminished cleavages were observed from 243 to 266, from 345 to 354, and from 366 to 379 (**Figure 3A**). Additionally, the 278 and 381 positions were more sensitive to DNase I cleavage. These results indicated that MgaSpn-His recognizes preferentially two sites on the *hlyR* DNA fragment (**Figure 3B**). One of them (site A; positions 242 to 266) is located within the H-NS binding site I, whereas the other one (site B, positions 342 to 379) is adjacent to it. On both strands and at 80 nM of MgaSpn-His, regions protected against DNase I digestion were observed along the DNA fragment (**Figure 3A**), which is consistent with the pattern of protein-DNA complexes observed by EMSA (**Figure 2A**).

Figure 4A shows the bendability/curvature propensity plot of the 288-bp *hlyR* DNA fragment (coordinates 129–416) according to the bend.it program (Vlahovicek et al., 2003). The profile contains a peak of potential sequence-dependent curvature at position 238, just adjacent to the MgaSpn-His binding site A (positions 242-266). Its magnitude (9.7 degrees per helical turn) is within the values calculated for experimentally tested curved motifs (Gabrielian et al., 1997). Furthermore, the profile reveals that site A is flanked by regions of potential bendability (positions 215-222 and 271-278). The MgaSpn-His binding site B (positions 342-379) contains a peak of predicted curvature at position 364 (magnitude 14) which is also flanked by regions of potential bendability (positions 319-335 and 380-396). Thus, on the *hlyR* DNA fragment (see **Figure 1A**), MgaSpn-His binds preferentially to two sites (sites A and B) that are flanked by regions of potential bendability. Whereas, site A is located within the extended H-NS binding site I, site B is just adjacent to it.
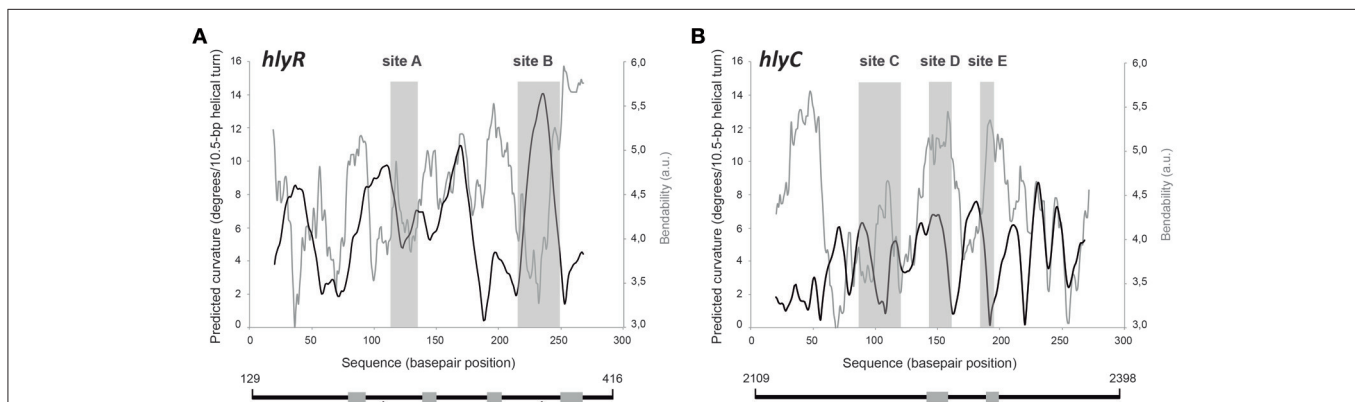

FIGURE 4 | Bendability/curvature propensity plots of the *hlyR* and *hlyC* DNA fragments according to the bend.it program (Vlahovicek et al., 2003). The sites recognized preferentially by MgaSpn-His on **(A)** the 288-bp *hlyR* DNA fragment (site A, site B), and **(B)** the 290-bp *hlyC* DNA fragment (site C, site D, and site E) are shown (gray rectangles). Lines below the plots represent the *hlyR* and *hlyC* DNA fragments. Coordinates of both DNA fragments are indicated. Gray boxes show regions of bendability that are either flanking or included into the MgaSpn-His binding sites. Arrowheads indicate peaks of sequence-dependent curvatures on the *hlyR* DNA fragment.
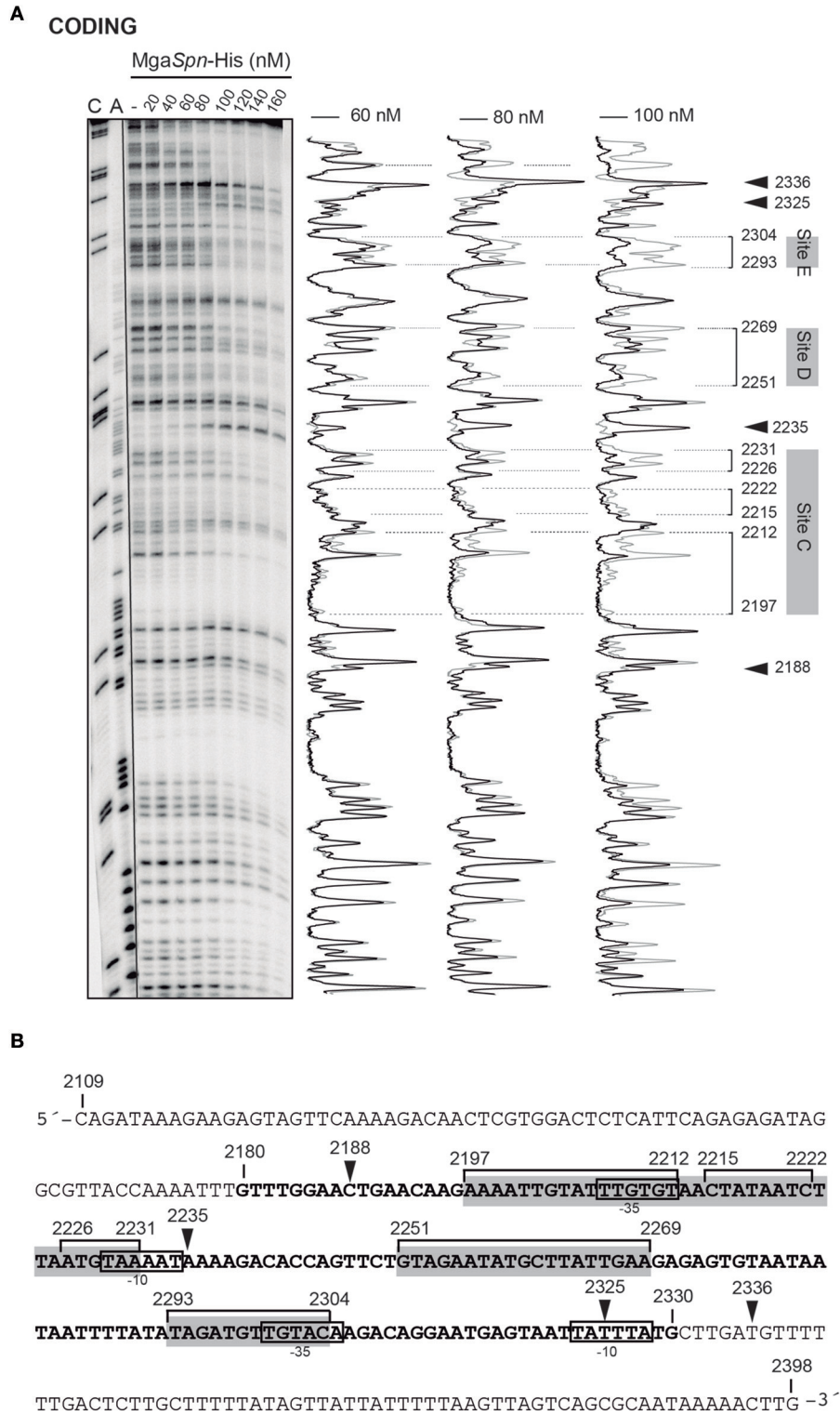
FIGURE 5 | DNase I footprints of complexes formed by Mga*Spn*-His on the 290-bp *hlyC* DNA fragment. (A) The *hlyC* DNA fragment (coordinates 2109–2398) was radioactively labeled at the 5′-end of the coding strand. Labeled DNA (2 nM) was incubated with the indicated concentrations of Mga*Spn*-His and then digested with DNase I. Dideoxy-mediated chain-termination sequencing reactions were run in the same gel (lanes A, C). Densitometer scans corresponding to free DNA (gray line) and DNA with protein (black line; 60, 80, and 100 nM) are shown. The three primary sites (site C, site D and site E) recognized by Mga*Spn*-His on

(Continued)

Binding of MgaSpn-His to the Promoter Region of the *E. coli hly* Operon

H-NS interacts not only with the site I of plasmid pHly152 but also with the site II (coordinates 2180–2330), which is located upstream of *hlyC*, the first gene of the *hly* operon (Madrid et al., 2002) (see **Figure 1A**). Site II includes two of the three promoters described for the *hly* operon (Koronakis and Hughes, 1988). To analyze whether MgaSpn-His recognizes particular regions on the H-NS binding site II, we used a 290-bp DNA fragment (here named *hlyC*; coordinates 2109–2398) that contains the site II at internal position (**Figure 1A**). Radioactively labeled DNA was incubated with increasing concentrations of MgaSpn-His. By EMSA, we found that MgaSpn-His also generates multimeric complexes on the *hlyC* DNA fragment (**Supplementary Figure 1**). The apparent Kd of MgaSpn-His for the *hlyC* DNA fragment was about 75 nM, slightly higher than for the *hlyR* DNA fragment (**Figure 2B**).

The position of MgaSpn-His on the *hlyC* DNA fragment was further analyzed by DNase I footprinting assays (**Figure 5A**). To this end, the 290-bp DNA fragment was labeled at the 5′-end of the coding strand. At 60 and 80 nM of MgaSpn-His, diminished DNase I cleavages were mainly observed from coordinate 2197 to 2231 (site C), from 2251 to 2269 (site D), and from 2293 to 2304 (site E). Moreover, positions 2188, 2235, 2325, and 2336 were more sensitive to DNase I digestion. At higher protein/DNA ratios, MgaSpn-mediated protections were observed along the *hlyC* DNA fragment (**Figure 5A**). Therefore, MgaSpn-His recognizes preferentially three regions within the extended H-NS binding site II (**Figure 5B**). Compared to the 288-bp *hlyR* DNA fragment (**Figure 4A**), the magnitude of the curvatures predicted in the 290-bp *hlyC* DNA fragment is slightly lower (<9 degrees per helical turn) (**Figure 4B**). Nevertheless, the MgaSpn-His binding sites D (2251-2269) and E (2293-2304) are located at regions of conspicuous bendability (positions 2248-2267 and 2296-2309, respectively).

Binding of H-NS-His to the Pneumococcal *PB* Activation Region

In vivo activation of the *P1623B* promoter requires a 70-bp region (*PB* activation region) located upstream of the promoter (from position −50 to −119) (**Figure 1B**). By DNase I footprinting experiments, we demonstrated previously that MgaSpn-His interacts with the *PB* activation region (positions −52 to −102) when it is located at internal position on a 222-bp DNA fragment (coordinates 1598298 to 1598519 of the pneumococcal R6 genome; see **Figure 1B**) (Solano-Collado et al., 2012). Similar results were obtained using an untagged form of the MgaSpn protein (Solano-Collado et al., 2013). In the present study, we analyzed whether protein H-NS-His is able to
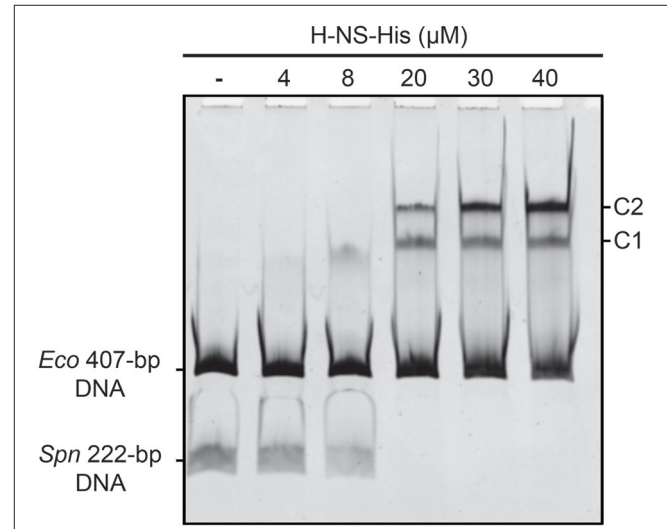
FIGURE 6 | Binding of H-NS-His to the pneumococcal 222-bp DNA fragment (competitive EMSA). The pneumococcal 222-bp DNA fragment (50 ng) was incubated with increasing concentrations of H-NS-His in the presence of competitor DNA (407-bp DNA fragment from the *E. coli* plasmid pHly152; 150 ng). Free and bound DNAs were separated by native gel electrophoresis (5% polyacrylamide). Bands corresponding to free DNAs and protein-DNA complexes (C1 and C2) are indicated.

bind to the pneumococcal 222-bp DNA fragment that contains the *PB* activation region. First, we performed a competitive gel retardation assay (**Figure 6**). The pneumococcal 222-bp DNA fragment was mixed with a 407-bp DNA fragment (competitor DNA) from the *E. coli* plasmid pHly152, and both DNAs were incubated with increasing concentrations of H-NS-His. The 407-bp DNA fragment was reported to lack preferential binding sites for H-NS (Madrid et al., 2002). As shown in **Figure 6**, protein H-NS-His showed a higher affinity for the pneumococcal 222-bp DNA fragment. Next, we used DNase I footprinting to identify the sites recognized by H-NS-His. The pneumococcal 222-bp DNA fragment was radioactively labeled at the 5′-end of the non-coding strand, and then it was incubated with increasing concentrations of H-NS-His (**Figure 7A**). At 10 nM of H-NS-His, changes in DNase I sensitivity (diminished cleavages) were observed from positions −53 to −68 (site 1), −102 to −111 (site 2), and −121 to −131 (site 3). Whereas, site 1 and site 2 are located within the *PB* activation region, site 3 is adjacent to it (**Figure 7B**). At 80 nM of protein, H-NS-His mediated protections were observed along the entire DNA fragment (**Figure 7A**). According to predictions of intrinsic DNA curvature (Solano-Collado et al., 2013), the pneumococcal 222-bp DNA fragment contains one peak of potential curvature (magnitude 9.5, position −90)
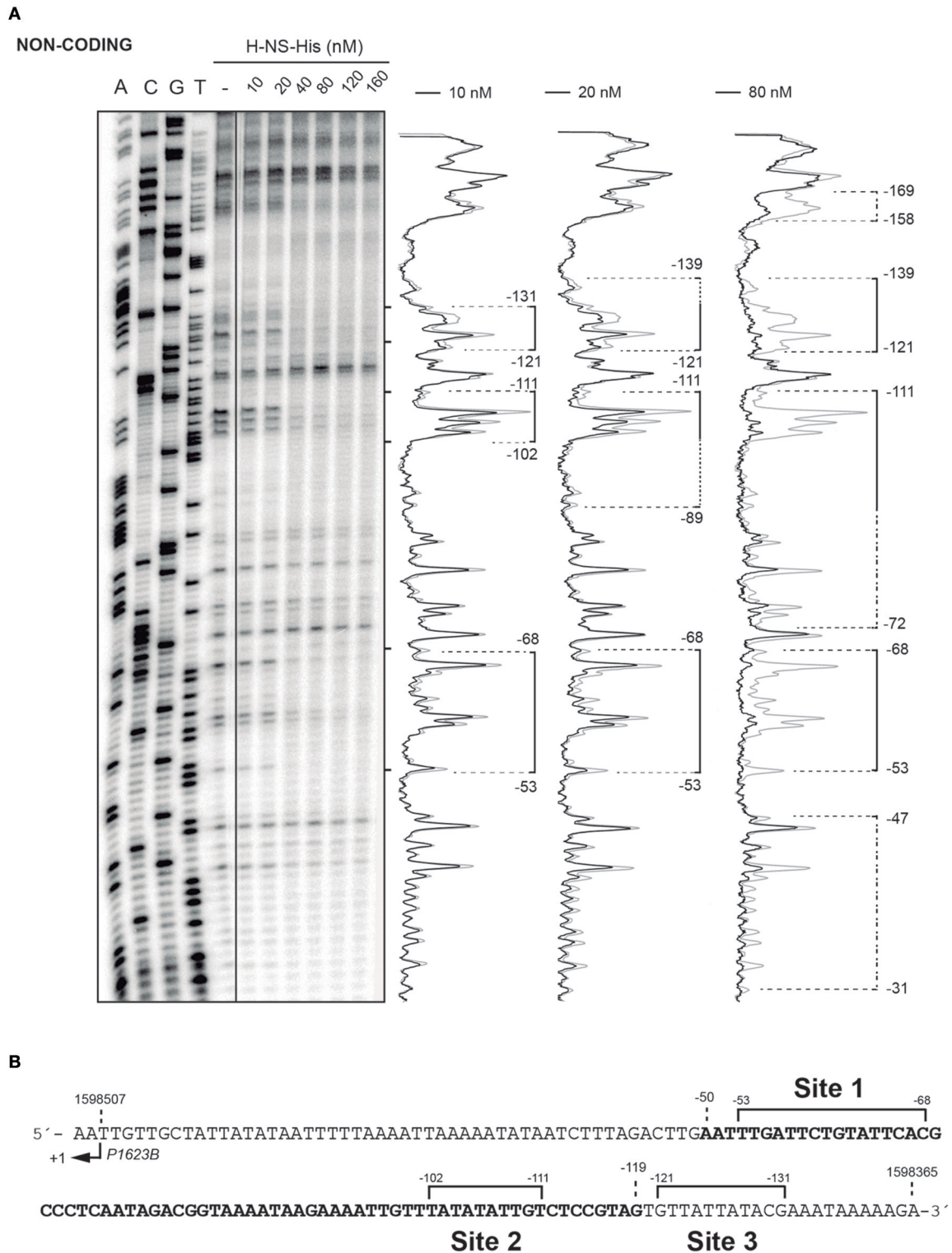
FIGURE 7 | DNase I footprints of complexes formed by H-NS-His on the 222-bp DNA fragment. (A) The 222-bp DNA fragment of the R6 chromosome (coordinates 1598298–1598519) was radioactively labeled at the 5′-end of the non-coding strand. Labeled DNA (2 nM) was incubated with increasing concentrations of H-NS-His and then digested with DNase I. Dideoxy-mediated chain-termination sequencing reactions were run in the same gel (lanes A, C, G, T). Densitometer scans corresponding to free DNA (gray line) and DNA with protein (black line; 10, 20, and 80 nM) are shown. Brackets indicate H-NS-His binding sites. Extension of

(Continued)

FIGURE 8 | Sites recognized by MgaSpn. **(A)** Nucleotide sequence alignment of the *PB* activation region and the *Pmga* promoter region of the pneumococcal R6 genome. Both MgaSpn binding sites were defined by hydroxyl radical footprinting assays (Solano-Collado et al., 2013). Identical nucleotides are highlighted in gray boxes. **(B)** Nucleotide sequence alignment of the regions recognized by MgaSpn-His on the *hly* regulatory region of the *E. coli* plasmid pHly152 (this work). These MgaSpn-His binding sites were defined by DNase I footprinting assays. Nucleotides of such sites that are present in the *PB* activation region and/or the *Pmga* promoter region are highlighted in black boxes. The A+T content of the different sites is indicated on the right. The global A+T content of the pneumococcal R6 genome is 60.3%. By EMSA, the minimum DNA size required for MgaSpn binding was shown to be between 20 and 26 bp (Solano-Collado et al., 2013).

within the *PB* activation region (**Figure 1B**). Furthermore, there are two regions of potential bendability (from −62 to −76 and from −110 to −122) flanking such a curvature. Hence, the three sites recognized by H-NS-His on the pneumococcal 222-bp DNA fragment are adjacent to regions of potential bendability.

DISCUSSION

Proteins MgaSpn and H-NS are very different in size. They neither exhibit sequence similarity nor share a common domain structure. MgaSpn is a member of an emerging class of global response regulators (the Mga/AtxA family) that contain phosphoenolpyruvate **p**hospho**t**ransferase **s**ystem (PTS) **r**egulation **d**omains (PRDs) (Hammerstrom et al., 2015). MgaSpn is predicted to have two N-terminal helix-turn-helix DNA-binding motifs, a central PRD and a C-terminal region with amino acid similarity to the PTS protein EIIB. On the other hand, proteins in the H-NS family consist of a coiled-coil N-terminal oligomerization domain and a C-terminal DNA-binding domain. Both domains are joined via an unstructured flexible linker (reviewed by Winardhi et al., 2015). Despite these differences between MgaSpn and H-NS, previous studies on the MgaSpn transcriptional regulator suggested that it shares certain DNA-binding properties with H-NS (Solano-Collado et al., 2013).

Based on *in vitro* DNA binding studies, we have proposed that MgaSpn regulates the expression of numerous genes by a mechanism that involves recognition of particular DNA conformations (Solano-Collado et al., 2013). By hydroxyl radical footprinting experiments, MgaSpn was

shown to bind to two regions of the *S. pneumoniae* R6 chromosome: the *PB* activation region (positions −60 to −99 of the *P1623B* promoter) and the *Pmga* promoter region (positions −23 to +21 of the *Pmga* promoter) (Solano-Collado et al., 2013). Both MgaSpn binding regions share the **GGT**(A/T)(A/T)**AAT**(A/C)(A/C)**GA**(A/T)**AATT** sequence element (**Figure 8A**). Moreover, they contain a potential intrinsic curvature flanked by regions of bendability (Solano-Collado et al., 2013). Results presented in this work support that MgaSpn recognizes structural features in its DNA targets. Specifically, DNase I footprinting experiments allowed us to identify five primary sites for MgaSpn on the *hly* operon regulatory region of the *E. coli* plasmid pHly152. All of them are located within or adjacent to extended H-NS binding sites (**Figure 1A**). Moreover, four out of the five MgaSpn binding sites (i) have a size between 19 and 38 bp, (ii) display a high A+T content (71.1–82.9%), (iii) share short DNA sequence motifs with the two pneumococcal MgaSpn binding sites (**Figure 8**), and (iv) are either located at or flanked by regions of potential bendability (**Figure 4**). Considering the global A+T content (60.3%) of the pneumococcal R6 genome, these results reinforce the conclusion that MgaSpn, like H-NS, has a preference for AT-rich DNA sites. Most likely, a preference for AT-rich DNA regions rather than for specific DNA sequences is a general feature of the global regulators that constitute the Mga/AtxA family. Sequence alignments of all established Mga binding regions revealed that they exhibit only 13.4% identity. Furthermore, a mutational analysis in some target promoters indicated that Mga binds to DNA in a promoter-specific manner (Hause and McIver, 2012). In the case of AtxA, sequence similarities in its target

promoters are not apparent, and it has been shown that the promoter regions of several target genes are intrinsically curved (Hadjifrangiskou and Koehler, 2008).

Proteins of the H-NS family have a high degree of sequence similarity in the DNA-binding domain, as it is the case of H-NS and Ler (Shindo et al., 1995; Cordeiro et al., 2011). The three-dimensional structure of a complex between the DNA-binding domain of Ler and a 15-mer DNA duplex has been solved (Cordeiro et al., 2011). This structure revealed that the DNA-binding domain of Ler does not participate in base-specific contacts but recognizes specific structural features in the DNA minor groove. Thus, Ler, and likely other members of the H-NS family, recognizes specific DNA shapes. By DNase I footprinting experiments, we have found that H-NS recognizes three particular sites on the regulatory region of the pneumococcal *P1623B* promoter (**Figure 1B**). The three sites, ranging in size from 10 to 16 bp, are adjacent to regions of potential bendability, which agrees with the preference of H-NS for AT-rich DNA regions. Moreover, two out of the three H-NS binding sites occur in the *PB* activation region, which includes an Mga*Spn* binding site. Hence, the regulatory region of the pneumococcal *P1623B* promoter contains structural motifs that are recognized by H-NS.

In conclusion, two categories of protein-DNA interactions, namely those in which the protein recognizes unique chemical signatures of the DNA bases (base readout), and those in which the protein recognizes a sequence-dependent DNA shape (shape readout) were defined by Rohs et al. (2010). Our present work suggests that two unrelated DNA-binding proteins from phylogenetically distant bacteria are able to recognize similar structural characteristics in their DNA targets. It is intriguing that unrelated bacterial species have evolved to encode proteins that seem to use a similar strategy to regulate the expression of a number of genes (silencing or activations). We take this as an indication of a successful strategy for proteins recognizing DNA regions that show intrinsic bendability/flexibility.

AUTHOR CONTRIBUTIONS

VS and MH performed laboratory work. VS, MH, ME, AJ, and AB designed the study, performed data analysis and wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmolb.2016.00060

Supplementary Figure 1 | Binding of Mga*Spn*-His to the 290-bp *hlyC* DNA fragment. (A) EMSA. The ^{32}P-labeled *hlyC* DNA fragment (2 nM) was incubated with increasing concentrations of Mga*Spn*-His (20 to 180 nM). Free and bound DNAs were separated by native gel electrophoresis (5% polyacrylamide). Bands corresponding to free DNA (F) and to several protein-DNA complexes (C1, C2, C3, and C4) are indicated. **(B)** Affinity of Mga*Spn*-His for the 290-bp *hlyC* DNA fragment. The autoradiograph shown in A was scanned, and the percentage of free DNA was plotted against Mga*Spn*-His concentration.

REFERENCES

Carey, J. (1988). Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. *Proc. Natl. Acad. Sci. U.S.A.* 85, 975–979. doi: 10.1073/pnas.85.4.975

Cordeiro, T. N., Schmidt, H., Madrid, C., Juárez, A., Bernadó, P., Griesinger, C., et al. (2011). Indirect DNA readout by an H-NS related protein: structure of the DNA complex of the C-terminal domain of Ler. *PLoS Pathog.* 7:e1002380. doi: 10.1371/journal.ppat.1002380

Dame, R. T., Wyman, C., and Goosen, N. (2000). H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* 28, 3504–3510. doi: 10.1093/nar/28.18.3504

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Gabrielian, A., Vlahovicek, K., and Pongor, S. (1997). Distribution of sequence-dependent curvature in genomic DNA sequences. *FEBS Lett.* 406, 69–74. doi: 10.1016/S0014-5793(97)00236-6

Godessart, N., Muñoa, F. J., Regué, M., and Juárez, A. (1988). Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in *Escherichia coli*. *J. Gen. Microbiol.* 134, 2779–2787.

Hadjifrangiskou, M., and Koehler, T. M. (2008). Intrinsic curvature associated with the coordinately regulated anthrax toxin gene promoters. *Microbiology* 154, 2501–2512. doi: 10.1099/mic.0.2007/016162-0

Hammerstrom, T. G., Horton, L. B., Swick, M. C., Joachimiak, A., Osipiuk, J., and Koehler, T. M. (2015). Crystal structure of *Bacillus anthracis* virulence regulator AtxA and effects of phosphorylated histidines on multimerization and activity. *Mol. Microbiol.* 95, 426–441. doi: 10.1111/mmi.12867

Hause, L. L., and McIver, K. S. (2012). Nucleotides critical for the interaction of the *Streptococcus pyogenes* Mga virulence regulator with Mga-regulated promoter sequences. *J. Bacteriol.* 194, 4904–4919. doi: 10.1128/JB.00809-12

Hemsley, C., Joyce, E., Hava, D. L., Kawale, A., and Camilli, A. (2003). MgrA, an orthologue of Mga, acts as a transcriptional repressor of the genes within the *rlrA* pathogenicity islet in *Streptococcus pneumoniae*. *J. Bacteriol.* 185, 6640–6647. doi: 10.1128/JB.185.22.6640-6647.2003

Hondorp, E. R., Hou, S. C., Hause, L. L., Gera, K., Lee, C.-E., and McIver, K. S. (2013). PTS phosphorylation of Mga modulates regulon expression and virulence in the group A streptococcus. *Mol. Microbiol.* 88, 1176–1193. doi: 10.1111/mmi.12250

Hoskins, J., Alborn, W. E. Jr., Arnold, J., Blaszczak, L. C., Burgett, S., Dehoff, B. S., et al. (2001). Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* 183, 5709–5717. doi: 10.1128/JB.183.19.5709-5717.2001

Ko, M., and Park, C. (2000). H-NS-dependent regulation of flagellar synthesis is mediated by a LysR family protein. *J. Bacteriol.* 182, 4670–4672. doi: 10.1128/JB.182.16.4670-4672.2000

Koronakis, V., and Hughes, C. (1988). Identification of the promoters directing *in vivo* expression of hemolysin genes in *Proteus vulgaris* and *Escherichia coli*. *Mol. Gen. Genet.* 213, 99–104. doi: 10.1007/BF00333404

Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C. O., et al. (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res.* 35, 6330–6337. doi: 10.1093/nar/gkm712

Lim, C. J., Lee, S. Y., Kenney, L. J., and Yan, J. (2012). Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. *Sci. Rep.* 2:509. doi: 10.1038/srep00509

Liu, Y., Chen, H., Kenney, L. J., and Yan, J. (2010). A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev.* 24, 339–344. doi: 10.1101/gad.1883510

Madrid, C., Nieto, J. M., Paytubi, S., Falconi, M., Gualerzi, C. O., and Juárez, A. (2002). Temperature- and H-NS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *J. Bacteriol.* 184, 5058–5066. doi: 10.1128/JB.184.18.5058-5066.2002

Nieto, J. M., Carmona, M., Bolland, S., Jubete, Y., De La Cruz, F., and Juárez, A. (1991). The *hha* gene modulates haemolysin expression in *Escherichia coli*. *Mol. Microbiol.* 5, 1285–1293. doi: 10.1111/j.1365-2958.1991.tb01902.x

Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., and Mann, R. S. (2010). Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* 79, 233–269. doi: 10.1146/annurev-biochem-060408-091030

Ruiz-Cruz, S., Espinosa, M., Goldmann, O., and Bravo, A. (2016). Global regulation of gene expression by the MafR protein of *Enterococcus faecalis*. *Front. Microbiol.* 6:1521. doi: 10.3389/fmicb.2015.01521

Ruiz-Cruz, S., Solano-Collado, V., Espinosa, M., and Bravo, A. (2010). Novel plasmid-based genetic tools for the study of promoters and terminators in *Streptococcus pneumoniae* and *Enterococcus faecalis*. *J. Microbiol. Methods* 83, 156–163. doi: 10.1016/j.mimet.2010.08.004

Shin, M., Lagda, A. C., Lee, J. W., Bhat, A., Rhee, J. H., Kim, J.-S., et al. (2012). Gene silencing by H-NS from distal DNA site. *Mol. Microbiol.* 86, 707–719. doi: 10.1111/mmi.12012

Shindo, H., Iwaki, T., Ieda, R., Kurumizaka, H., Ueguchi, C., Mizuno, T., et al. (1995). Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett.* 360, 125–131. doi: 10.1016/0014-5793(95)00079-O

Solano-Collado, V., Espinosa, M., and Bravo, A. (2012). Activator role of the pneumococcal Mga-like virulence transcriptional regulator. *J. Bacteriol.* 194, 4197–4207. doi: 10.1128/JB.00536-12

Solano-Collado, V., Lurz, R., Espinosa, M., and Bravo, A. (2013). The pneumococcal MgaSpn virulence transcriptional regulator generates multimeric complexes on linear double-stranded DNA. *Nucleic Acids Res.* 41, 6975–6991. doi: 10.1093/nar/gkt445

Vlahovicek, K., Kaján, L., and Pongor, S. (2003). DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res.* 31, 3686–3687. doi: 10.1093/nar/gkg559

Vogel, M., Hess, J., Then, I., Juárez, A., and Goebel, W. (1988). Characterization of a sequence (*hlyR*) which enhances synthesis and secretion of hemolysin in *Escherichia coli*. *Mol. Gen. Genet.* 212, 76–84. doi: 10.1007/BF00322447

Winardhi, R. S., Yan, J., and Kenney, L. J. (2015). H-NS regulates gene expression and compacts the nucleoid: insights from single-molecule experiments. *Biophys. J.* 109, 1321–1329. doi: 10.1016/j.bpj.2015.08.016