Article

AT-Hook-Like Motif Is the Common DNA Binding Element of H-NS Family Proteins: Insights from the Solution Structure of *E. coli* H-NS$^{85–137}$/DNA Complex

Jiayu Wang [1,2], Fan Liu [1,3], Bo Duan [4,*], and Bin Xia [1,2,3,*]

[1] Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China
[2] College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
[3] School of Life Sciences, Peking University, Beijing 100871, China
[4] Suzhou Key Laboratory of Pathogen Bioscience and Anti-Infective Medicine, MOE Key Laboratory of Geriatric Diseases and Immunology, Institute of Molecular Enzymology, School of Biology and Basic Medical Sciences, Soochow University, Suzhou 215123, China
* Correspondence: duanbo@suda.edu.cn (B.D.); binxia@pku.edu.cn (B.X.)
Received: 17 August 2026; Revised: 27 October 2026; Accepted: 9 November 2026; Published: 20 November 2026

Abstract: H-NS is a major xenogeneic silencer that preferentially targets AT-rich DNA and plays central roles in bacterial chromosome organization and transcriptional regulation. Despite extensive biochemical and genetic studies, the structural basis underlying its DNA recognition remains incompletely understood. Here we report the solution structure of *E. coli* H-NS residues 85–137 (H-NS$^{85–137}$) in complex with a AT-rich DNA. The structure confirms that the conserved AT-hook-like motif inserts into the DNA minor groove and forms multiple hydrogen bonds with AT base pairs. Strikingly, the side chain of linker residue R90 also penetrates the minor groove adjacent to residue Q112, establishing direct hydrogen bonds with DNA bases and expanding the protein/DNA interface. The linker residue R93 bridges the linker region and AT-hook-like motif through a hydrogen-bond network centered on residue Q112. Mutational and NMR analyses demonstrate that these interactions cooperatively contribute to DNA binding, with the AT-hook-like motif serving as the primary recognition element and the linker providing auxiliary contacts. MD simulations reveal that H-NS$^{85–137}$ binding widens and shallows the minor groove, increases local roll and promotes DNA bending, providing a structural basis for its preference of flexible AT-rich sequences containing TpA steps. Finally, comparative sequence analysis of H-NS family proteins identifies three major structural types and 29 AT-hook-like core tripeptide variants, revealing that the AT-hook-like motif serves as a common DNA recognition element across the family. This work provides a structural framework for understanding H-NS mediated xenogeneic silencing, post-translational regulation, and the functional diversification of H-NS family proteins.

Keywords: xenogeneic silencing; nucleoid-associated proteins; biomolecular NMR; DNA binding; transcription repression

1. Introduction

The histone-like nucleoid-structuring protein (H-NS) is one of the most thoroughly characterized nucleoid-associated proteins (NAPs) in bacteria, and it plays dual roles in chromosome organization and global transcriptional regulation [1,2]. Although H-NS contributes to nucleoid organization by compacting genomic DNA, its primary physiological identity lies in its function as a premier xenogeneic silencer that can selectively suppress foreign genes acquired through horizontal gene transfer (HGT) [3–5]. H-NS preferentially recognizes AT-rich DNA, a characteristic feature of foreign genomic islands, plasmids, prophages, and other mobile genetic elements [6]. By binding these regions, H-NS suppresses transcription initiated from promoters, thereby preventing inappropriate gene expression and maintaining transcriptional homeostasis. Through this xenogeneic silencing mechanism, H-NS regulates a wide range of biological processes, including stress responses, environmental adaptation, metabolism, antimicrobial resistance, and bacterial virulence, in many bacterial pathogens, including *Salmonella enterica*, *Shigella flexneri*, enteropathogenic and enterohemorrhagic *Escherichia coli*, *Vibrio cholerae*, and *Acinetobacter baumannii* [7–11]. Consequently, H-NS has emerged as a central regulator linking bacterial genome evolution, environmental sensing, and pathogenesis [12,13].

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

H-NS consists of an N-terminal oligomerization domain (NTD), a flexible linker, and a C-terminal DNA binding domain (CTD). The N-terminal domain of H-NS plays a central role in xenogeneic silencing by controlling protein self-association and the formation of higher-order nucleoprotein complexes through multivalent interaction with DNA. Structural and biochemical studies have shown that the N-terminal domain has two independent dimeric interfaces which serves the basic dimeric building blocks required for continuous oligomerization [14]. Solution structure of the C-terminal domain revealed a small compact fold containing a loop with a conformation mimicking the AT-hook motif of eukaryotic HMGA1 [15,16]. HADDOCK model of the C-terminal domain and DNA complex generated based on NMR titration results showed that H-NS binds AT-rich DNA mainly through a QGR (residues Q112-G113-R114) core tripeptide of the AT-hook-like motif inserting into the minor-groove. The side-chains of Q112 and R114 extend in opposite directions along the groove, allowing the protein to interact with an extended AT-rich DNA segment. Protein binding microarray study and NMR titration results showed that H-NS prefers binding AT-rich DNA with multiple TpA steps [15]. This AT-hook-like motif conformation is conserved in Mycobacterial xenogeneic silencer Lsr2 and H-NS family protein Bv3V from *Burkholderia vietnamiensis*, both have an RGR core tripeptide. Alternative core tripeptide sequences (RGR, NGR, IGR and VGR) have also been identified in plasmid-borne homologs (HppX) of H-NS from *Enterobacterales* bacteria [17].

Later studies reported that the unstructured linker region (residues 81–94) connecting the NTD and CTD of *E. coli* H-NS plays an important role in its DNA binding and gene silencing function [18]. Particularly, the two positively charged residues R90 and R93, along with two of the three Lys residues (K83, K87, and K89), are required for the proper function of H-NS. It was further proposed that H-NS initially engages DNA non-specifically via reasonably high affinity ($K_d \sim 60$ nM) electrostatic interactions with basic residues in the linker region, and the linker region was even suggested to be the main determinant of the DNA binding affinity of H-NS, instead of the C-terminal DNA binding domain [19]. An NMR and EMSA-based study also demonstrated the importance of residues R90 and R93 in the DNA binding of *Salmonella* H-NS [20]. Recently, several MD simulation studies have been carried out to study the interaction between H-NS DNA binding domain and DNA [21–23]. It was also suggested that the interaction of the AT-hook-like motif with DNA is facilitated by residue R93, along with four Lys residues [21]. Although the two Arg residues in the linker region of *E. coli* H-NS are highly conserved among its homologous proteins with the QGR core tripeptide, they are not conserved in other H-NS family proteins, such as Bv3F-alike proteins with an RGR core tripeptide [13].

Recently, increasing evidences revealed that H-NS undergoes diverse post-translational modifications (PTMs), including phosphorylation, methylation, deamidation, oxidation, acetylation, succinylation, and 2-hydroxyisobutyrylation [24,25]. Depending on the specific sites within distinct domains of H-NS, PTM can modulate its oligomerization or DNA binding ability, and thus regulate its function. Dilweg & Dame have identified 29 PTMs of *E. coli* H-NS by summarizing bacterial proteomic studies results, in which 26 of them are within the DNA binding domain [24], and some of them have been shown to regulate the function of H-NS. In *E. coli*, the acetylation has been found for 4 of total 5 Lys residues (K96, K107, K120, and K128) within the C-terminal DNA binding domain, along with two in the linker region (K83 and K87) [26]. Acetylation mimic mutants H-NS, with these Lys residues mutated individually into Gln, all failed to inhibit T3/T6SS expression. Especially, the acetylation of K120 impairs DNA binding ability of H-NS, so does K120Q and K120A mutations. The 2-hydroxyisobutyrylation has been identified for the other residue K121, which causes reduced DNA binding ability of H-NS, and releases the silencing of acid-tolerance genes to promote survival of *E. coli* in acidic conditions [27]. In addition, it was reported that a phosphorylation mimic mutant T106E of *E. coli* H-NS no longer suppresses the expression of SPI-1 genes. It is proposed that PTMs of H-NS impair the gene silencing ability of H-NS and serve as a new counter-silencing mechanism [25].

A major obstacle in understanding H-NS-mediated xenogeneic silencing has been the absence of an experimentally determined structure of the H-NS in complex with DNA. Although the previous docking model revealed the role of the conserved AT-hook-like motif in DNA binding, it failed to explain the contributions of the adjacent linker region, particularly the highly conserved arginine residues, or the functional significance of several post-translational modification sites, all of which have been implicated in modulating DNA binding. Here, we report the structure of the DNA-bound complex of residues 85–137 of *E. coli* H-NS, including the DNA binding domain and the linker region. The structure reveals that, in addition to the hydrogen bonds with DNA bases by the QGR core tripeptide of the AT-hook-like motif, the side-chain guanidinium group of linker residue R90 directly inserts into the DNA minor groove and forms a hydrogen bond with DNA base next to residue Q112. The DNA binding affinity is also enhanced by hydrogen bonds from backbone NH groups and electrostatic interaction between Lys residues and DNA phosphate groups. Furthermore, we have classified the DNA binding domains of H-NS family proteins into three types according to characteristics in the primary sequences of the DNA binding domain, and identified 29 core tripeptide sequence variants. The conserved AT-hook-like motif serves as a

common structural element for DNA recognition in all three types of H-NS family proteins, while the additional DNA base-contacting residue R90 of *E. coli* H-NS is only unique to Type I H-NS proteins with QGR/K core tripeptides.

2. Materials and Methods

2.1. Sample Preparations

DNA fragments encoding residues 85–137 of the *E. coli* H-NS was synthesized and cloned into a pET-21a expression vector (Novagen), between the *Nde* I and *Xho* I restriction sites, and thus an 8-residue (LEHHHHHH) His_6-tag was introduced to the C-terminus. The plasmids were then transformed into *Escherichia coli* BL21 (DE3) competent cells for protein expression. The procedures of protein expression and purification have been described in detail in the Supplementary Information of a previous publication [28]. In short, after cell lysis, the His6-tagged protein was first purified with a Ni-NTA affinity column (Qiagen), followed by size exclusion chromatography using a Superdex 75 Gel Filtration Column (GE) in 50 mM sodium phosphate buffer at pH 5.6, with 50 mM NaCl.

Point mutations of H-$NS^{85–137}$ were generated using the site-directed mutagenesis kit (SBS Genetech, Beijing, China). Expression and purification of the mutant proteins followed the same procedures.

For the preparation of DNA samples, equimolar amounts of synthesized complementary oligonucleotides were heated to 94 °C for 5 min, and then annealed by slowly cooling down to room temperature.

2.2. NMR Titration Experiments

All ^{15}N-labeled WT H-$NS^{85–137}$ and its mutant NMR samples contained 0.1 mM uniformly ^{15}N-labeled protein in 450 μL of 50 mM sodium phosphate buffer (pH 7.0) with 50 mM NaCl, with 20 μL D_2O added. DNA molecules were added to protein samples at different DNA/protein molar ratios. 2D ^{1}H-^{15}N SOFAST-HMQC or 2D ^{1}H-^{15}N HSQC NMR experiments were performed on a 600 MHz Bruker Ascend™ NMR spectrometer, at 298 K.

Dissociation constant K_d was determined for individual NH signal by fitting the combined NH chemical shift changes ($\Delta\delta_{comb} = [\Delta\delta_H^2 + (\Delta\delta_N/6.5)^2]^{1/2}$) from NMR titration experiment data vs. DNA/protein molecular ratios ($R_{D/P}$) to the following equation:

$$\Delta\delta_{comb} = \Delta\delta_{max} \{(K_d/C_P + k\cdot R_{D/P} + 1) - [(K_d/C_P + k\cdot R_{D/P} + 1)^2 - 4k\cdot R_{D/P}]^{1/2}\}/2$$

where $\Delta\delta_H$ and $\Delta\delta_N$ are the ^{1}H and ^{15}N chemical shift changes, respectively; $\Delta\delta_{max}$ is the NH signal chemical shift difference between free and bound states of protein, C_P is the total protein concentration, and k is a correction factor for $R_{D/P}$. The reported K_d is the average of K_d values determined from multiple NH signals, along with the standard error.

2.3. NMR Data Collections for Structure Calculation

For resonance assignments and structure calculation, all NMR samples contained 0.5 mM ^{15}N/^{13}C -labeled or ^{15}N -labeled H-$NS^{85–137}$ and 5 mM non-labeled double-stranded Top1 DNA, in 50 mM sodium phosphate buffer (pH 5.6) with 50 mM NaCl for the collection of NMR data.

The following NMR experiments were performed for protein backbone and side-chain resonance assignments on a 500 MHz Bruker AVANCE spectrometer [29]: 2D ^{1}H-^{15}N HSQC, 2D ^{1}H-^{13}C HSQC, 3D HNCACB, 3D CBCA(CO)NH, 3D HNCA, 3D HN(CO)CA, 3D (H)CCH-COSY, 3D (H)CCH-TOCSY, 3D H(C)CH-COSY, and 3D H(C)CH-TOCSY. 2D F1, F2-^{15}N/^{13}C-filtered ^{1}H-^{1}H NOESY experiments with a mixing time of 150 ms or 300 ms were performed on a 950 MHz Bruker AVANCE spectrometer for the assignments of DNA protons and the acquisition of intra-DNA distance restraints. 3D ^{1}H-^{15}N-edited NOESY-HSQC, 3D ^{1}H-^{13}C-edited NOESY-HSQC, 2D ^{13}C-edited ^{1}H-^{1}H NOESY optimized for aromatic resonances, 3D F1-^{15}N/^{13}C filtered, F2-^{15}N edited NOESY-HSQC, and 3D F1-^{15}N/^{13}C filtered, F2-^{13}C edited NOESY-HSQC experiments with a mixing time of 150 ms were also performed on the 950 MHz Bruker AVANCE spectrometer for the acquisition of intra-protein and intermolecular distance restraints. All NMR experiments were performed at 298 K. The raw NMR data were processed with NMRPipe [30] and the spectra were visualized with NMRView [31]. Proton chemical shifts were referenced directly to internal DSS. ^{15}N and ^{13}C chemical shifts were referenced indirectly to DSS.

2.4. Solution Structure Determination

The H-$NS^{85–137}$/DNA complex was calculated based on the intra-DNA, intra-protein, and intermolecular NOE signals, accompanied by protein dihedral angle restraints obtained using TALOS-N [32], and empirical torsion angle restraints and Watson-Crick restraints of B-form DNA [33].

The detailed procedure for structure calculation was described previously [34]. Briefly, 20 initial structures of H-NS$^{85–137}$ were generated based on the free DNA binding domain structure of *Salmonella* H-NS (PDB ID 2L93). Distance restraints for the protein were first generated based on the initial structures using SANE [35], and were used for structure calculation with the DYANA module of CYANA [36], and the resulting structures were used in the next round of SANE analysis. The iterative SANE-DYANA cycle was carried out iteratively until no distance violation larger than 0.2 Å existed. Then, the coordinates of 100 protein conformers were arbitrarily combined with 100 coordinates of Top-1 DNA, which were calculated using AMBER 14 with distance restraints derived from the 2D F1, F2-^{15}N/^{13}C-filtered NOESY spectrum and empirical torsion angle restraints and Waston-Crick restraints of B-form DNA according to the AMBER tutorial [37]. These 100 protein and DNA coordinates were subjected to further complex structure calculation and refinement using AMBER14, together with intermolecular NOE restraints obtained from the 3D F1-^{15}N/^{13}C filtered, F2-^{13}C edited NOESY-HSQC and 3D ^{1}H-^{13}C-edited NOESY-HSQC spectra. Structure calculation finished when all NOE signals had been properly assigned and no distance restraints violations over 0.2 Å or dihedral angle restraints violations over 5° existed. The 20 conformers with the lowest AMBER energies were selected for representation.

2.5. Molecular Dynamic Simulations

All-atom molecular dynamics (MD) simulations were performed using the GPU-accelerated pmemd.cuda module of AMBER [37]. Initial system coordinates were generated with the tleap module of AmberTools21 [38]. The protein, DNA, and water molecules were modeled using the Amber ff14SB, Parmbsc1, and TIP3P force fields, respectively. Each system was placed in a cubic periodic water box, with a minimum distance of 10 Å maintained between the solute surface and the box boundary. Counterions (Na^+ or Cl^-) were added according to the net charge of each system to achieve electroneutrality.

Before the production runs, each solvated system underwent a multistep energy-minimization and equilibration protocol. First, the solvent was energy-minimized while positional restraints were applied to the solute, followed by a short NPT equilibration to relax the solvent. The entire system was then energy-minimized without positional restraints. The system was subsequently heated from 0 to 298 K over 500 ps under constant-volume (NVT) conditions. During heating, positional restraints with a harmonic force constant of 10 kcal mol^{-1} Å^{-2} were applied to the DNA or protein/DNA complex. The temperature was regulated using the Berendsen weak-coupling thermostat with a coupling constant of 0.5 ps. The system was then equilibrated for 500 ps under constant-pressure and constant-temperature (NPT) conditions at 298 K, with the same positional restraints maintained. During this stage, temperature was controlled using a Langevin thermostat with a collision frequency of 2.0 ps^{-1}, while the pressure was maintained at 1 bar using isotropic pressure scaling with a pressure-relaxation time of 1.0 ps.

Following equilibration, all positional restraints were removed, and production MD simulations were performed for 100 ns under NPT conditions at 298 K and 1 bar. Temperature regulation was maintained using a Langevin thermostat with a collision frequency of 2.0 ps^{-1}, and pressure was controlled using isotropic pressure scaling with a pressure-relaxation time of 2.0 ps. The random seed for the Langevin thermostat was generated automatically using ig = −1. Periodic boundary conditions were applied throughout the simulations. Long-range electrostatic interactions were treated using the particle-mesh Ewald (PME) method, with a 10 Å cutoff for short-range nonbonded interactions. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm, allowing an integration time step of 2 fs.

A single 100 ns production MD trajectory was generated for H-NS$^{85–137}$/Top1 DNA complex and Top1 DNA alone. Structural and dynamical properties were analyzed over the equilibrated portion of each trajectory. After the MD simulation, the equilibrated portion of each trajectory were analyzed using the Cpptraj module of AmberTools. Frames were extracted every 0.1 ns and DNA geometric parameters were obtained from analysis using Curves+ [39]. Structural parameters were analyzed to compare the DNA geometries sampled in the free and H-NS-bound states. The resulting values characterize the conformational ensembles sampled in the respective trajectories.

2.6. Primary Sequence Analysis and Motif Classification

Sequence homologs were identified by iterative online BLAST- and PHI-BLAST-based searches (https://blast.ncbi.nlm.nih.gov/Blast.cgi, accessed on 1 August 2025).

The sequences of residues 85–137 of *E. coli* H-NS and residues 71–112 of *B. vietnamiensis* Bv3F were used as initial query sequences in NCBI Protein BLAST to search against the NCBI RefSeq Select proteins database (refseq_select) with default parameters (maximum number of target sequences set 5000). BLAST hits from the two queries were pooled together and proteins between 80 and 300 residues in length were selected. Sequences

containing the WxGxxxP motif were further retained only if at least 70 residues preceded the motif and at least 10 residues followed it.

The retained sequences were grouped according to different core tripeptide sequences with a center Gly residue between the conserved Trp and Pro residues of the WxGxxxP motif. For example, the corresponding motifs are WxGQGRxP in *E. coli* H-NS and WxGRGRxP in Bv3F. Each was treated as an independent sequence group. A representative protein sequence from each WxGxxxP motif group with unique core tripeptide was selected and used as the subsequent query for further PHI-BLAST searches against the NCBI non-redundant protein database (nr) with the defined search pattern (e.g., WxGQGRxP for H-NS and WxGRGRxP for Bv3F). The maximum number of target sequences was set to 20,000, and all remaining parameters were set as default values. Only hits exhibiting 50–100% query coverage were retained.

The resulting homologs typically shared at least 30% sequence identity with the query. Taxonomic analyses were performed directly on the complete PHI-BLAST output before downstream filtering to minimize potential sampling bias. Again, protein sequences shorter than 80 amino acids or longer than 300 amino acids were discarded. Sequences lacking the expected motif or containing fewer than 70 residues upstream or fewer than 10 residues downstream of the motif were excluded. All the remaining protein sequences were annotated as H-NS family proteins, and the residues more than 25 residues upstream of the WxGxxxP motif were trimmed from the sequences to retain the region corresponding to the DNA binding domain.

Multiple sequence alignments for each WxGxxxP motif group were performed using Clustal Omega with default parameters [40]. Alignments were exported in FASTA format, and sequence conservation logos were generated using WebLogo3 (see Supplementary data for sequences used in generating logos) [41].

3. Results

3.1. Solution Structure of E. coli H-NS$^{85–137}$ and DNA Complex

We have previously reported that NH signals of the AT-hook-like motif of *Salmonella* H-NS CTD (residues 91–137) gradually disappeared during DNA titration, which prevented us from determining the complex structure [15]. To address this limitation, we have carried out NMR sample optimizations, in terms of protein construct, DNA sequence, buffer condition and DNA/protein ratio, etc. We have expressed residues 85–137 of *E. coli* H-NS (H-NS$^{85–137}$), which includes both the C-terminal DNA binding domain and the linker region that contains four of the suggested essential positively charged residues. A palindromic double-stranded DNA molecule d(CGAATATATTCG)$_2$ (Top1-DNA) was used, which contains the AATATATT sequence with the highest 8-mer E-score in the PBM data [15]. Backbone NMR resonances were assigned for free H-NS$^{85–137}$ (Supplementary Figure S1A), and NMR titration (Supplementary Figures S1B and S2A) revealed that its dissociation constant (K_d) is 2.0 μM for binding Top1-DNA, e.g., the DNA binding affinity of H-NS$^{85–137}$ is over 4 times higher than the previously reported *Salmonella* H-NS CTD (K_d ~9 μM). Also, H-NS$^{85–137}$ shows higher DNA binding affinity at acidic pH. During the DNA titration, some of the NH signals of the AT-hook-like motif first become weaker or even disappeared, but reappeared at a DNA/protein ratio of 1.5, and the NH signals stop change after the ratio reached 2. However, a higher DNA/protein ratio (>5) is required to get high-quality NMR spectrum of the DNA in the complex, along with NMR spectra for detecting intermolecular NOEs. With the optimized sample condition, we were able to obtain high quality NMR data suitable for the structure determination of the complex.

The backbone and side-chain resonance assignments of H-NS$^{85–137}$ in complex with Top1-DNA are nearly complete, and 57 backbone NH signals of the total 60 non-proline residues were observed, except for residues G85, G111, and the first Met (Figure 1A). Resonances from all Top-1 DNA residues were assigned (Figure 1B). The solution structure of the H-NS$^{85–137}$ and Top1 complex was determined using 2264 intra-protein NOE restraints, 870 intra-DNA NOE restraints, and 141 inter-molecular NOE restraints (Supplementary Figure S2B). The ensemble of the 20 conformers with the lowest energy is shown in Figure 1C, with a root-mean-square deviation (RMSD) of 0.71 Å, calculated for the backbone heavy atoms of DNA and protein secondary structure regions. 91.7% of the protein main chain dihedral angles are distributed within the most favored regions of the Ramachandran plot, with no dihedral angle in the disallowed regions. The statistics for restraints and structures are summarized in Table 1.

In the complex structure, residues 95–135 of H-NS$^{85–137}$ are folded with a two-stranded anti-parallel β-sheet (β1, residues 97–100; β2, residues 106–109), together with an α-helix (H1, residues 117–125) and a 3$_{10}$-helix (H2, residues 130–133) (Figure 1D). There are three loops connecting these secondary structure elements: L1 (residues 101–105), L2 (residues 110–116), and L3 (residues 126–129). The hydrophobic core of H-NS$^{85–137}$ is primarily composed of the side-chains of residues K96, Y97, Y99, W109, P116, V118, I119, A122, M123, L130, F133, and L134, with an additional small hydrophobic patch (A95, T108, and I135) further stabilizing the local conformation

of the C-terminal tail (Supplementary Figure S3A). The N-terminal region of H-NS$^{85–137}$ (e.g., the linker residues 85–94) is unstructured.

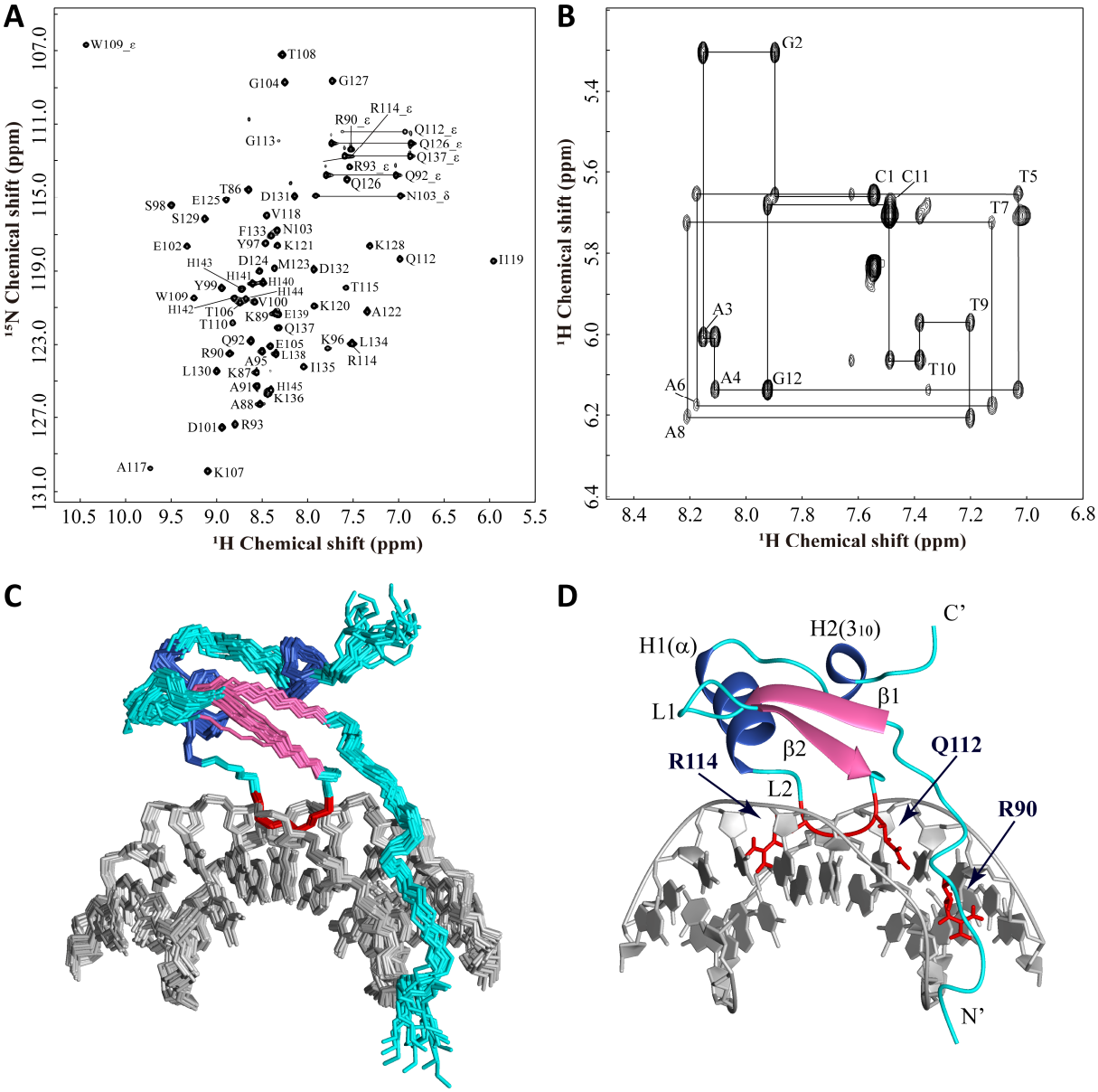

Figure 1. Chemical shift assignments and structure of the complex of *E. coli* H-NS$^{85–137}$ and Top1-DNA. (**A**) Chemical shift assignment of the 2D ^{1}H-^{15}N HSQC spectrum H-NS$^{85–137}$ in complex with Top1-DNA. (**B**). Chemical shift assignments in the H1′-H6/8 fingerprint region of the 2D F1, F2-^{15}N/^{13}C filtered ^{1}H-^{1}H NOESY spectrum of Top1-DNA in the complex. (**C**) Backbone ensemble of the 20 lowest-energy structures of the complex. (**D**) Ribbon diagram of the lowest-energy representative solution structure of the complex. Helices and β-strands in protein secondary structure regions are respectively shown in blue and magenta, the region inserted into the DNA minor groove is shown in red, and the remaining parts are shown in cyan. DNA is shown in light gray.

Consistent with the previous docking model [15], the QGR core tripeptide residues of the AT-hook-like motif at the L2 loop of H-NS$^{85–137}$ are inserted into the minor groove of the DNA and form multiple hydrogen bonds with A/T bases, with the side-chains of Q112 and R114 extending in opposite directions. The side-chain ε-NH_2 group of residue Q112 forms a hydrogen bond with the N3 atom from the base of DNA residue A_{20}. The main chain NH group of residue R114 forms a hydrogen bond with the O2 atom from the base of T_7, while its side-chain guanidinium group forms hydrogen bonds with the O2 atoms of T_{17} and T_9, as well as the O4′ atoms from the sugars of A_{18} and T_9. Although residue G113 is inserted into the minor groove, it does not form hydrogen bond with DNA bases. Instead, its NH group hydrogen-bonds with the O4′ atom of A_{20} (Figure 2A). Interestingly, the side-chain of linker residue R90 is also inserted into the DNA minor groove, adjacent to the side-chain of residue Q112, and one of the terminal amines simultaneously forms two hydrogen bonds with the N3 atom of DNA residue

A_4 and the O4′ atom of C_{23} (Figure 2A). Collectively, the "R-QGR" four residues make contacts with DNA inside the minor groove, occupying a region that covers six base pairs ($A_4T_5A_6T_7A_8T_9$) (Figure 2C). In addition, H-NS$^{85–137}$ forms four hydrogen bonds with the DNA phosphate groups through backbone NH groups of residues K87, R90, A91 and A117 (Figure 2A,C). Among them, the NH signal of A117 displays the largest chemical shift perturbation upon DNA binding (Supplementary Figure S2A), which may imply that its hydrogen bond with DNA, right above residue R114, should be very important for maintaining the proper DNA binding conformation.

Table 1. Restraints and structure statistics for the solution structure of H-NS$^{85–137}$ and Top1 DNA complex.

Structural Restraints	
Restraints for protein	
NOE restraints	2264
Intra-residues	714
Inter-residues	876
Sequential	408
Non-sequential	468
Ambiguous	674
Dihedral angle restraints	88
φ angle	55
ψ angle	33
Chirality restraints	224
ω angle	61
Chirality	163
Restraints for DNA	
NOE restraints	870
Hydrogen bonds	56
Sugar pucker	24
Backbone dihedral angle	160
Intermolecular NOE restraints	141
Structure Statistics	
Distance restraints violations (>0.2 Å)	0
Dihedral angle restraints violations (>5°)	0
Average pairwise RMSD (Å)	
Protein	
All heavy atoms	1.01
Backbone heavy atoms for secondary structures	0.39
DNA	
All DNA heavy atoms	0.52
Complex	
Backbone heavy atoms for protein secondary structures and DNA	0.71
PROCHECK	
Most favored regions (%)	91.7
Additionally allowed regions (%)	8.1
Generously allowed regions (%)	0.2
Disallowed regions (%)	0.0

It is noteworthy that the guanidinium group of residue R93 forms two hydrogen bonds with the carbonyl O atom of residue T110, while the hydroxyl O atom of residue T110 is hydrogen-bonded by the backbone NH group of residue Q112, and the backbone carbonyl O atom of residue Q112 is hydrogen-bonded by the side-chain NH group of residue W109 (Figure 2A). This hydrogen bond network centered on Q112 should be important for maintaining the conformation of the AT-hook-like motif, as well as the relative position of the linker regions. As a result, the guanidinium group of residue R93 is located right above residue Q112 of the AT-hook-like motif, and it also forms two hydrogen bonds with the DNA backbone phosphate of T_{21} (Figure 2A,C).

The binding surface of H-NS$^{85–137}$ to Top1-DNA is highly positively charged (Supplementary Figure S3B). In addition to residues R90, R93, and R114, there are three Lys residues (K87, K89, and K120) that interact electrostatically with DNA phosphate groups (Figure 2B,C), defined as the Lys side-chain terminal N^{ζ} atom within 6 Å of a DNA phosphate oxygen atom in at least a third of the 20 conformers. Two of the Lys residues are located at the N-terminal tail, i.e., the linker region of H-NS. Together, these electrostatic interactions and hydrogen bonds should collectively secure the complex conformation and enhance the binding affinity of H-NS$^{85–137}$ to the DNA synergistically.

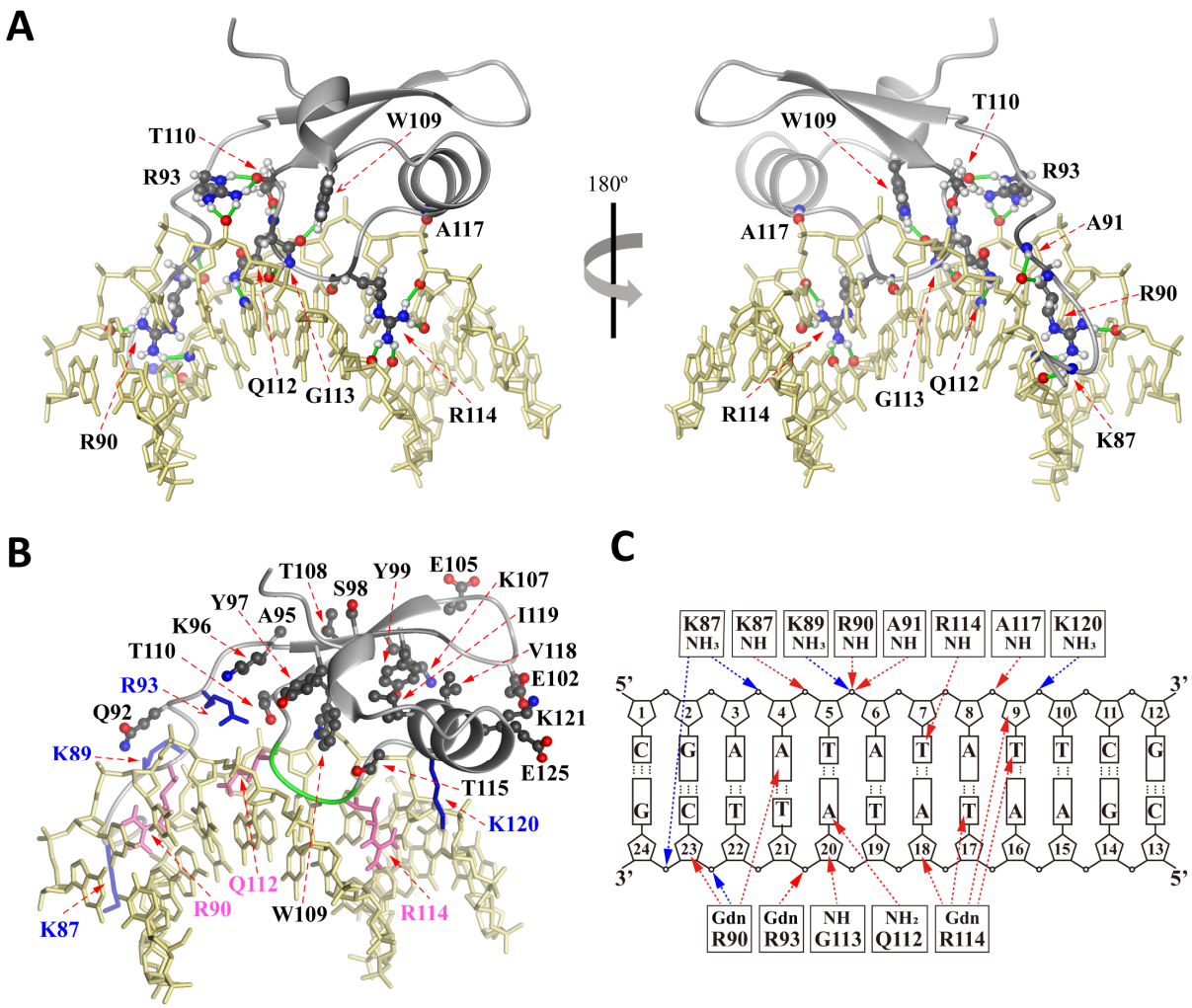

Figure 2. DNA recognition mechanism of *E. coli* H-NS$^{85–137}$. (**A**) Interactions between H-NS$^{85–137}$ and Top1-DNA. Intermolecular hydrogen bonds, as well as those that stabilize the conformation of the AT-hook-like structure, are presented in green sticks. (**B**) The QGR loop is colored green. The side-chains of Q112, R114, and R90, which are inserted into DNA minor groove, are colored pink. The side-chains of K87, K89, R93, and K120, which exhibit electrostatic interactions with DNA phosphate groups, are colored blue. The side-chains of the residues with over 0.1 ppm of combined NH chemical shift differences between free H-N$^{S85–137}$ and that with 5-fold of Top1 DNA are shown in ball and sticks. (**C**) Schematic representation of the protein-DNA interactions. Red lines with arrow head indicate hydrogen bonds, and blue ones indicate electrostatic intersections.

Comparison of the ^{1}H-^{15}N heteronuclear NOE values of free and DNA-bound H-NS$^{85–137}$ revealed that DNA binding significantly stabilizes the N-terminal tail and L2 loop of H-NS$^{85–137}$, while it slightly destabilizes the L1 loop that is not directly involved in DNA interaction (Supplementary Figure S4A). There are three negatively charged residues (D101, E102, and E105) on L1 loop, which may have electrostatic interactions with nearby positively charged residues K107 and K121 and thus stabilize the conformation of L1 loop in the free form. Upon DNA binding, the negative charge of DNA may disrupt these intra-protein salt bridges through electrostatic repulsion, thereby destabilizing the L1 loop (Supplementary Figure S4B). This also explains why residue E102, which is not at the binding interface, exhibits substantial chemical shift perturbation upon DNA binding (Supplementary Figures S1B and S2A).

We also compared the structure of DNA-bound *E. coli* H-NS$^{85–137}$ with the previously reported structure of the DNA binding domain of *Salmonella* H-NS (residues 91–137, PDB ID 2L93), which differ only by five residues (Supplementary Figure S5A). The major difference lies at the L1 loop connecting the two β-strands. There exists a β-bulge at the β2 strand of *Salmonella* H-NS$^{91–137}$, with both NH groups of residues T106 and K107 hydrogen-bonded to the carbonyl group of residue V100 on the β1 strand (Supplementary Figure S5B). In contrast, no β-bulge exists in the complex structure. This structural difference is unlikely caused by the five varied residues or DNA binding, as all these residues are relatively distant from the β-sheet and the DNA binding interface lies on the opposite face of the domain. In the solution structure of *Salmonella* H-NS$^{91–137}$, the β-bulge results in atom

T106-H^N in close proximity to V100-H^α and K107-H^N (≤2.8 Å), while the corresponding distances in the *E. coli* H-$NS^{85–137}$/Top1-DNA complex are over 4.4 Å. Examining the 3D ^1H-^{15}N NOESY-HSQC spectra of *E. coli* H-$NS^{85–137}$ in both free and DNA-bound states, no NOE signals characteristic for the β-bulge were found (Supplementary Figure S5C). Thus, no β-bulge was detected in free or DNA-bound *E. coli* H-$NS^{85–137}$ under the present solution conditions. However, the conformation of the AT-hook-like motif is largely unaffected, and residues 109–126 are well aligned between the two structures.

3.2. Mutagenesis of Key Residues in H-$NS^{85–137}$

Mutagenesis studies were carried out for residues in direct contact with DNA in the complex structure, along with those with large NH chemical shift perturbations during DNA titration. NMR titration experiments with Top1-DNA were conducted to evaluate the DNA binding affinities of these mutants (Figure 3A).

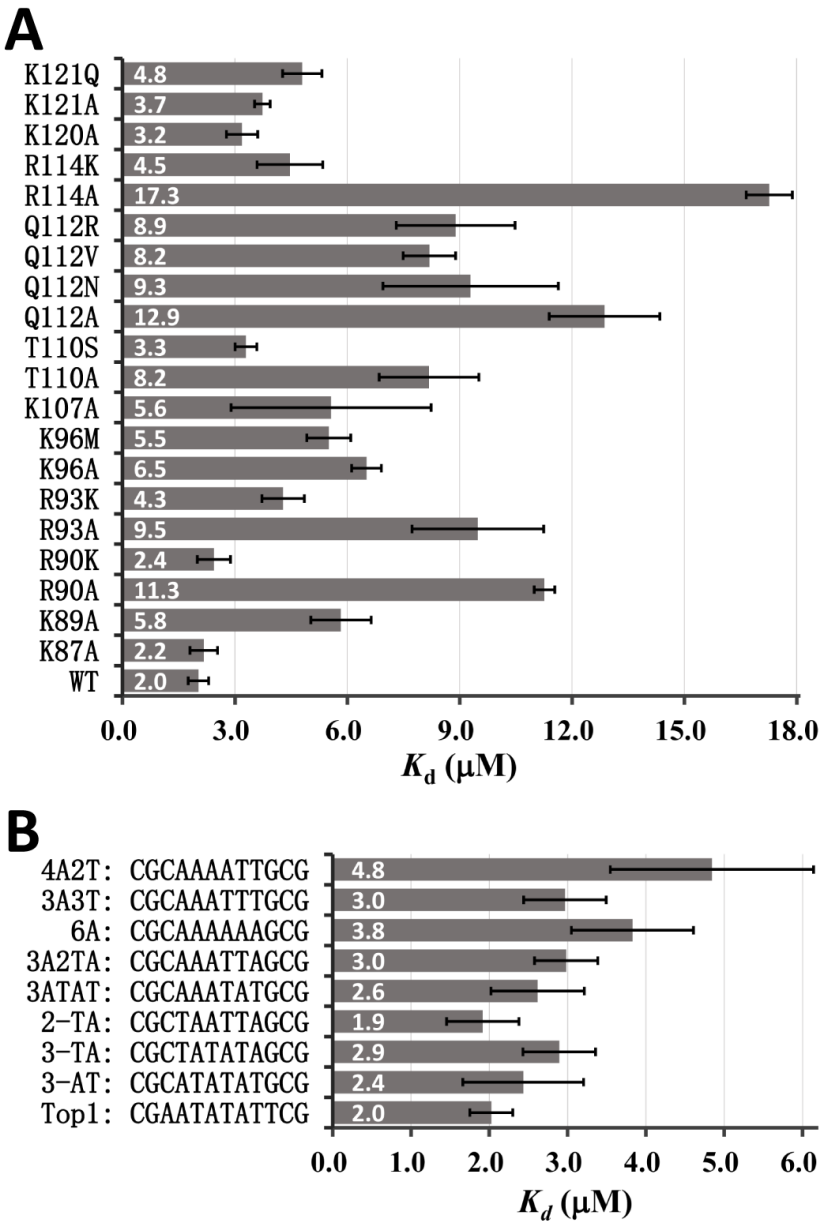

Figure 3. (**A**) Comparison of dissociation constants (K_d) for WT H-$NS^{85–137}$ and mutants for binding to Top1 DNA. (**B**) Comparison of dissociation constants (K_d) for WT H-NS^{85-137} and different AT-rich DNA sequences. K_d values are determined from NMR titration data, error bars represent standard errors based on fitting results from multiple NH peaks.

We first examined the three residues R90, Q112 and R114, which form hydrogen bonds with DNA bases in the minor groove. The R114A mutation results in the largest decrease in DNA binding affinity (K_d 17.3 μM) among all mutants (Figure 4, Supplementary Figures S6A and S7), compared to the DNA binding K_d of 2.0 μM for the

WT protein. The mutant R114K also displays a more than 2-fold reduction in DNA binding affinity (K_d 4.5 μM) (Supplementary Figures S6B and S7), likely because the side-chain amino group of Lys is incapable of forming as many hydrogen bonds as the guanidinium group of Arg. The affinity of mutant Q112A (Figure 4, Supplementary Figure S8A) towards Top1 DNA is ~6.5 times weaker (K_d 12.9 μM) than that of the WT protein, consistent with the observation that the side-chain of Q112 forms a hydrogen bond with DNA base. As alternative core tripeptide sequences (RGR, NGR, IGR and VGR) have been reported for H-NS homolog proteins [17], we also generated mutants Q112R, Q112N, and Q112V, which all show over 4-fold reduced DNA binding affinities (Figure 3A, Supplementary Figures S8 and S9), indicating that the Gln at this position is crucial for *E. coli* H-NS to bind DNA. The mutant R90A (Figure 4, Supplementary Figures S7C and S11A) also shows significantly reduced DNA binding affinity (K_d 11.3 μM), comparable to those of Q112 mutants. The DNA binding K_d (2.4 μM) of R90K mutant is about the same as WT protein, and the NMR spectra of DNA-bound mutant R90K and WT protein are largely overlapped (Supplementary Figures S7D and S11B), consistent with that the guanidinium group of residue R90 forms a hydrogen bond with DNA base, which could be easily replaced by the side-chain amino group of Lys.

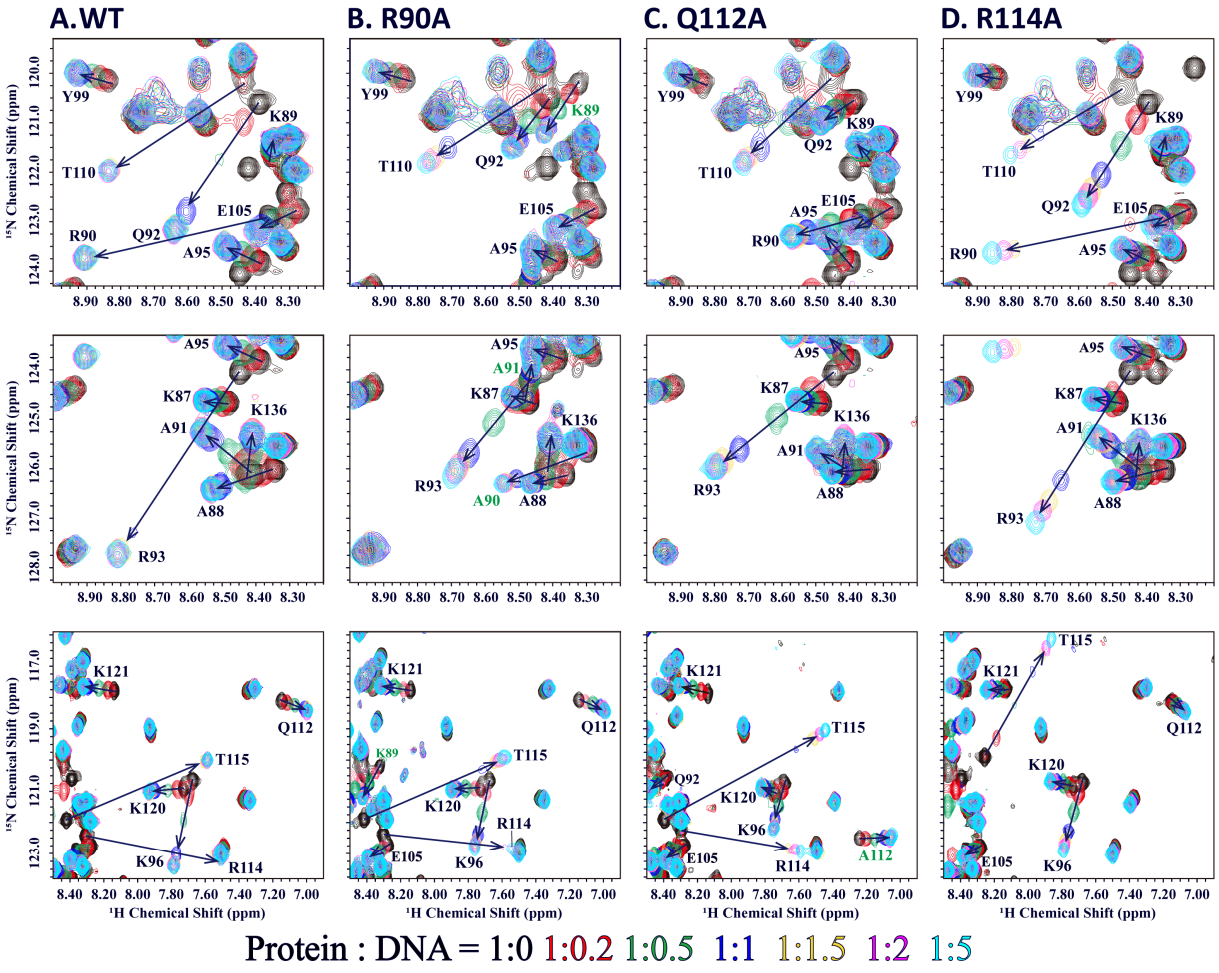

Figure 4. Comparison of selected regions of 2D ^{1}H-^{15}N HSQC spectra of Top1 DNA titration for WT H-NS$^{85–137}$ (**A**), mutants R90A (**B**), Q112A (**C**), and R114A (**D**). Traces of NH signal shift during titration are indicated by lines with arrow head, with residue type and number labeled. Tentative assignments for mutant signals are labeled in green. Color code for different protein/DNA ratios is shown at the bottom of the figure.

It appears that the R90A mutation mainly interferes with the binding of the linker residues to DNA, while the interactions of the AT-hook-like motif residues with DNA are not much affected, as evidenced by that the NH signal positions of residues T108, W109, Q112, R114, T115, and V118 of the mutant R90A in DNA-bound state are essentially the same as in WT protein (Figure 4, Supplementary Figure S11A). On the contrary, the Q112 mutations have a great impact on the interaction of the linker region with DNA, as the magnitudes of NH chemical shift perturbation of linker residues R90 and Q92 upon DNA binding are reduced dramatically, while those of the AT-hook-like motif residues are less affected, such as residues T110 and T115 (Figure 4, Supplementary Figures S8 and S9). For comparison, R114A mutation results in relatively uniform reduced chemical shift perturbation scales for NH signals of AT-hook motif and linker residues (Figure 4, Supplementary Figure S6A). It is worth noting that all these mutations

primarily affect NH signals from nearby residues in the free protein (Supplementary Figures S7–S10). Together, these data indicate that the interaction between the linker residues and DNA, especially residue R90, should be highly dependent on the interaction of the AT-hook-like motif with DNA.

The guanidinium group of residue R93 is involved in multiple hydrogen bonds within the protein and with DNA (Figure 2A). The K_d of mutant R93A to Top1-DNA is 9.5 μM, nearly five times higher than that of WT protein, whereas it is 4.3 μM for mutant R93K (Figure 3A). NH signals of residues T108, W109, T110, and G113 in the 2D ^{1}H-^{15}N HSQC spectra of free proteins are affected by R93A mutation, but not by R93K (Supplementary Figures S12A and S12B), in line with that residue R93 forms a hydrogen bond with the carbonyl group of residue T110 (Figure 2A). Consistently, the T110A mutation apparently destabilizes the protein conformation and causes many NH signals to shift or even disappear in the 2D ^{1}H-^{15}N HSQC spectrum (Supplementary Figure S12C), whereas only a few NH signals from residues near T110 are affected by the T110S mutation (Supplementary Figure S12D), as the hydroxyl group of residue T110 is hydrogen-bonded by the NH group of residue Q112. These results indicate that the hydrogen bond network centered on residue Q112 is not only important for maintaining the AT-hook-like motif conformation, but also critical for the overall folding of H-NS$^{85–137}$.

Similar to the case of Q112A mutant, the DNA-induced NH chemical shift perturbation magnitudes for linker residues 88–92 are dramatically reduced in mutants R93A and T110A, while those of the AT-hook motif residues are only slightly affected by R93A mutation and NH signals of residues T110, Q112, and T115 are shifted even further than WT protein (Figure 5, Supplementary Figures S13A and S14A). The T110A mutation results in changes in the directions of the NH signal titration traces for residues Q112, R114, and T115, possibly due to a change of the conformation of the AT-hook-like motif in mutant T110A (Figure 5, Supplementary Figure S14A). Again, these data suggest that the engagement between linker residues and DNA depends on the interaction between AT-hook-like motif and DNA. It is likely that the hydrogen bond of residue R93 with T110 plays a role to bring the linker residues closer to the AT-hook-like motif, facilitating residue R90 inserting into the minor groove of the DNA next to residue Q112 of the AT-hook-like motif.

For the three Lys residues that may have electrostatic interactions with DNA phosphate groups, individually mutating Lys to Ala does not affect the overall structure of H-NS (Supplementary Figure S15). It appears that the K87A mutation (K_d 2.2 μM) almost has no effect on the DNA binding affinity of H-NS, likely because the location of K87 is stabilized by its backbone NH that hydrogen-bonds to a DNA phosphate group, and its side-chain is positioned near the phosphate as a result. In the DNA-bound state, the NH signals of the L2 loop residues of the mutant K87A completely overlap with those of WT in the 2D ^{1}H-^{15}N HSQC spectra, although some signals are weaker (Supplementary Figure S16). Both the K89A and K120A mutants show reduced DNA binding affinities, with K_d values of 5.8 μM and 3.2 μM, respectively, indicating that the two Lys residues function to enhance DNA binding.

It has been reported that post-translational modifications of residues K96, K107, and K121 can affect the DNA binding ability of H-NS, even though they do not reside on the DNA binding interface. The K96A mutation results in global NH signal changes (Supplementary Figure S18A), indicating that residue K96 should be important for the overall folding of H-NS$^{85–137}$. Consistently, its side-chain is relatively well defined in the structure ensemble, and is likely involved in hydrophobic interactions with the side-chains of nearby residues Y97 and L134. Thus, we also generated a K96M mutant, which indeed has fewer NH signals displaying chemical shift changes (Supplementary Figure S18B). In the complex structure ensemble, the side-chain of residue K96 is positioned above the guanidinium group of residue R93 (Figure 2B), with the distances between C^{ϵ} atom of residue K96 and C^{ζ} atom of residue R93 in the range of 4.9–7.2 Å. It is possible that the electrostatic repulsion between the positively charged residues K96 and R93 also plays a role in defining the conformation of the AT-hook-like motif, which provides an explanation for the observation that the NH signal of residue K96 shows very large chemical shift perturbation upon DNA binding (Supplementary Figures S1B and S2A). In accordance, both the K96A and K96M mutants have reduced DNA binding affinities, with K_d values of 6.5 and 5.5 μM, respectively.

Both the K107A and K121A mutations result in chemical shift changes for many NH signals, some of which are weakened or even missing (Supplementary Figures S15D and S18C), indicating that both residues K107 and K121 should play structural roles in maintaining the H-NS conformation and its stability. In the structure, side-chains of residues K107 and D101 are within distances for electrostatic interaction, while K121 can electrostatically interact with residues E102 and E125 (Figure 2B). We have previously shown that the D101A mutation also destabilizes the folding of the DNA binding domain of H-NS [15]. Both mutants K107A and K121A show reduced DNA binding affinities, with K_d values of 5.6 μM and 3.7 μM, respectively. As it was reported that a lysine 2-hydroxyisobutyrylation mimic mutant K121Q of H-NS almost completely lost its DNA binding ability [27], we also generated a K121Q mutant of H-NS$^{85–137}$, and its affinity toward Top1 DNA (K_d 4.7 μM) is quite similar to that of K121A (Supplementary Figures S18D and S20).

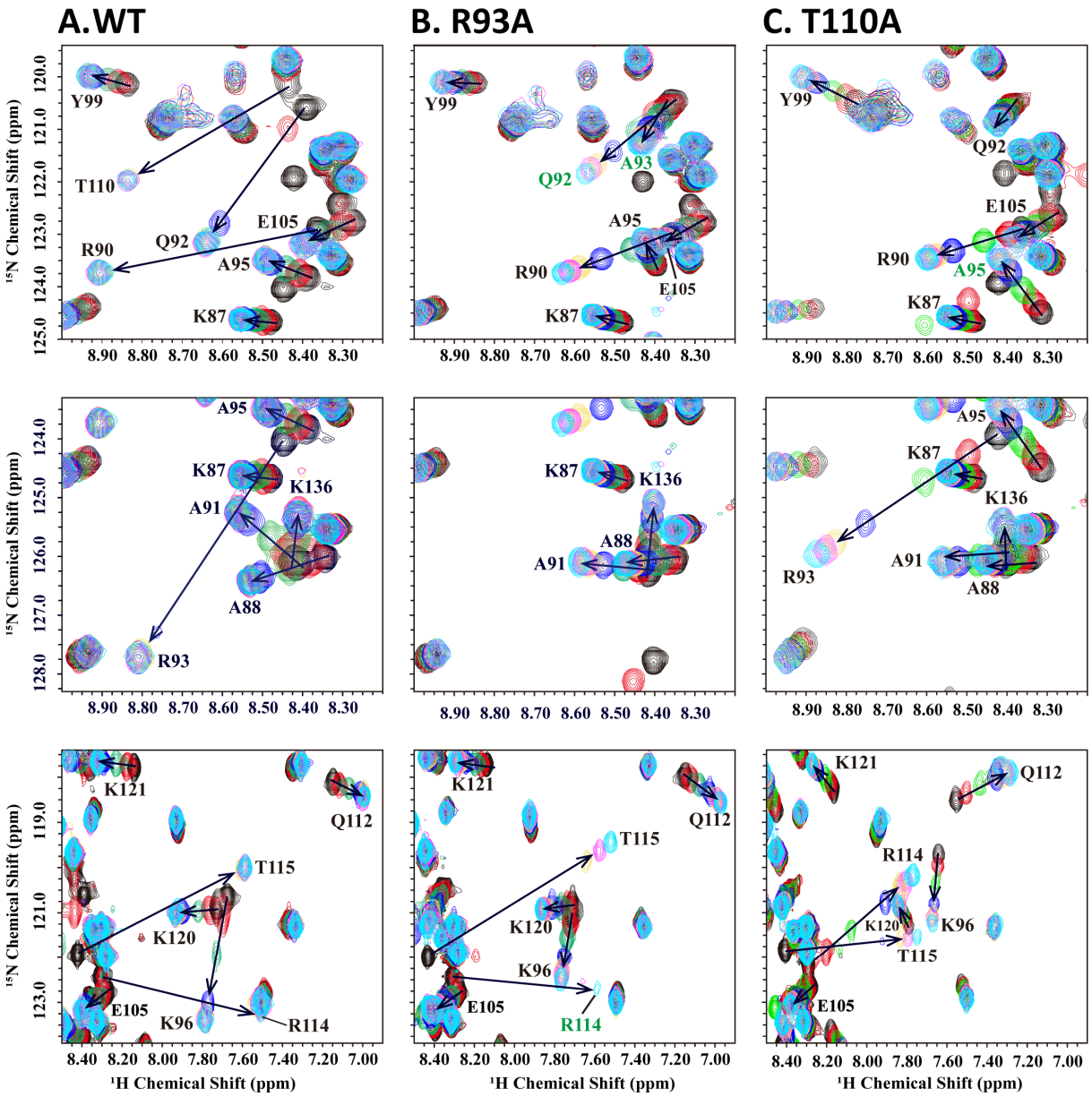

Figure 5. Comparison of selected regions of 2D ^{1}H-^{15}N HSQC spectra of Top1 DNA titration for WT H-NS$^{85–137}$ (**A**), mutants R93A (**B**), and T110A (**C**). Traces of NH signal shift during titration are indicated by lines with arrow head, with residue type and number labeled. Tentative assignments for mutant signals are labeled in green. Color code for different protein/DNA ratios is shown at the bottom of the figure.

3.3. DNA Deformation upon H-NS$^{85–137}$ Binding in MD Simulation

To assess the effects of H-NS binding on the DNA conformation, we have carried out 100 ns single-trajectory molecular dynamics simulations for the H-NS$^{85–137}$/Top1-DNA complex structure and the structure of Top1-DNA alone as a model for free DNA. DNA geometry parameters of the simulation trajectories were analyzed with Curves+. Upon the binding of H-NS$^{85–137}$, the RMSD values of Top1-DNA structure in the trajectory are lower than those of free DNA (Supplementary Figure S21), meaning that the binding of H-NS$^{85–137}$ stabilizes the overall DNA structure.

Comparing the distributions of DNA geometric parameters of the two MD trajectories, the most obvious DNA conformation change due to H-NS$^{85–137}$ binding is the enlargement of minor groove width for the three base pairs A_6-T_7-A_8, with the median minor groove width at T_7 increased to 9.1 Å from 5.4 Å in free DNA (Figure 6A). H-NS$^{85–137}$ binding also results in a shallower minor groove for the three base pairs, with the average minor groove depth of the T_7 base pair showing the largest decrease of over 1.5 Å (Figure 6B). Consistently, there are overall negative shifts of the base-pair plane x-displacement and positive shift of base-pair plane inclination values for T_5

to T_9 base-pairs in the H-NS$^{85–137}$ bound DNA (Figure 7A,B), suggesting that the base pairs are moved toward the minor groove [42], due to hydrogen bonds formed between H-NS and DNA bases inside the minor groove. As expected, the corresponding major groove depths are thus deeper in H-NS-bound DNA, while the major groove widths for the center 6 base pairs are more uniform (Figure 6C,D). For the two TpA steps T_5-A_6 and T_7-A_8, the ranges for base pair step shift of the H-NS$^{85–137}$-bound state are notably smaller than those in free DNA (Figure 7C), indicating that H-NS binding stabilizes the more flexible TpA steps of DNA, while the average base pair step roll values are increased by over 10° and 5°, respectively (Figure 7D). Increased roll is indicative of enhanced bending between adjacent base pairs and may represent a major contributor to the DNA curvature induced by H-NS binding. Indeed, the local axis bend angles are all higher in H-NS$^{85–137}$-bound DNA (Figure 7E), whereas the average total bend is increased to 19.2° from 14.2° in free DNA (Figure 7F). Consistently, Top1 DNA also displays an average total bend of 18.8°, in the 20 experimentally determined conformers of the complex. Together, these data show that H-NS binding remodels DNA architecture by inducing sharp local Roll variations and enhanced overall DNA bending, accompanied by increased minor groove width and DNA curvature.

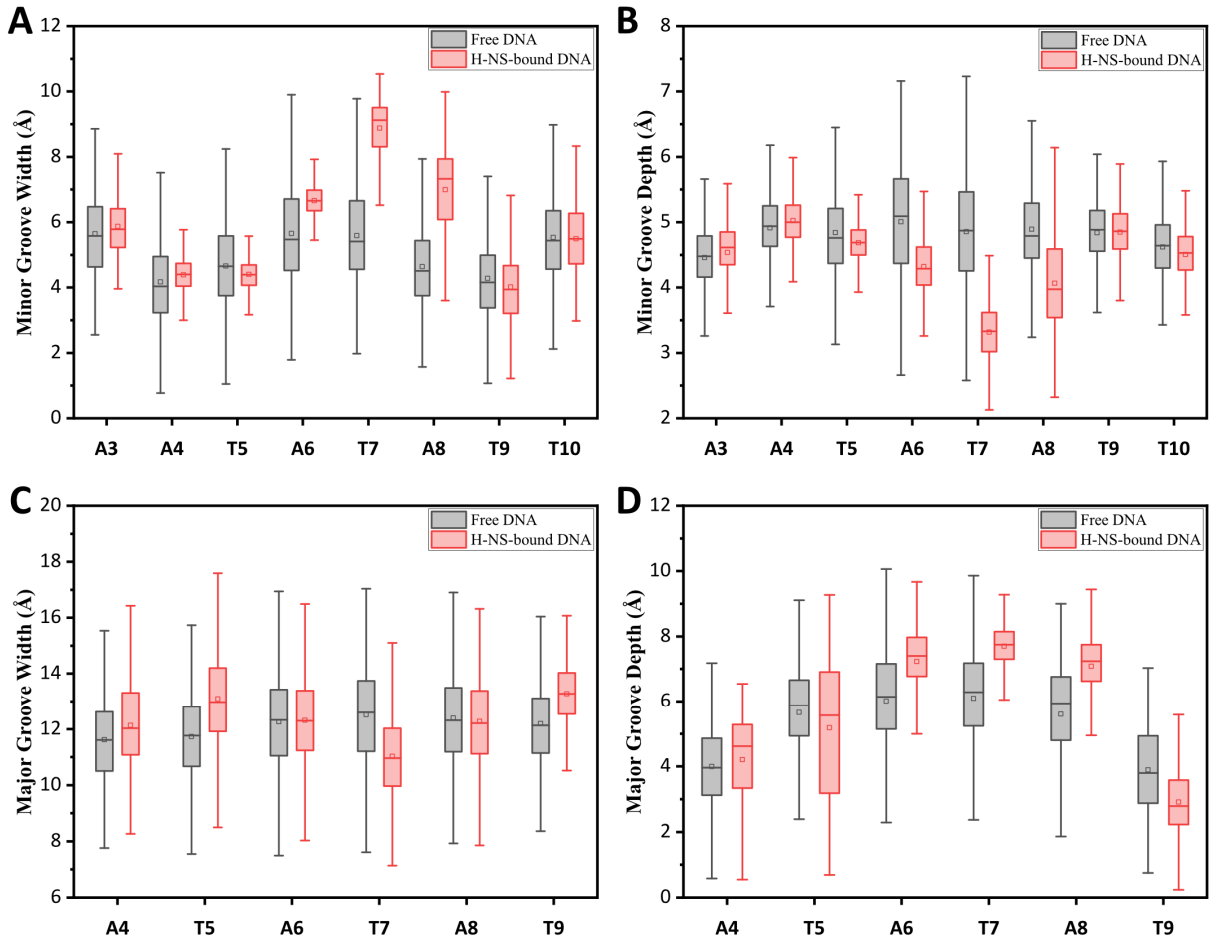

Figure 6. Comparison of distributions of DNA minor groove width (**A**), minor groove depth (**B**), major groove width (**C**), and major groove depth (**D**) values of the selected 1000 frames from MD simulations of the structure of H-NS$^{85–137}$ in complex with Top1 DNA and Top1 DNA alone. In the box plot, the box represents the interquartile range (IQR), with the horizontal line inside the box indicating the median value and the small open squares denoting the mean values. The whiskers extend to the farthest data points within 1.5 times the IQR.

While all the hydrogen bonds in our determined complex structure were captured in the simulation trajectories, MD simulation also revealed that the residues inserted into the minor groove can form different hydrogen bonds with the DNA bases. For example, the guanidinium group of R90 could form hydrogen bonds with the O2 atom of T_{22} instead of A4, the guanidinium group of R93 could form a hydrogen bond with the phosphate group of T_{22}, and the guanidinium group of R114 could form hydrogen bonds with the N3 atom of A_{18}. However, both the solution structure and the MD simulation results indicate that H-NS mainly contacts the central 6-bp AT sequence of Top1-DNA (5′-CGAATATATTCG-3′). We thus did NMR titration of 3AT-DNA (5′-CGCATATATGCG-3′) with 6-bp AT sequence, and its affinity (K_d 2.4 μM) is quite similar to that of the Top1-

DNA (K_d 2.1 μM), where 2D ^{1}H-^{15}N HSQC spectra are almost identical for H-NS bound to either DNA (Supplementary Figure S22A). We then examined the binding affinity of H-NS for other 6-bp consecutive AT DNA sequences with NMR titration. Generally, H-NS show higher binding affinity for sequences with TpA step(s) than A-tract DNA, with 4A2T-DNA (5′-CGCAAAATTGCG-3′; K_d 4.8 μM) and 6A (5′-CGCAAAAAAGCG-3′; K_d 3.8 μM) DNA showing lowest affinities, where 2-TA (5′-CGCTAATTAGCG-3′; K_d 1.8 μM) sequence has the highest affinity (Figure 3B, Supplementary Figures S22–S25).

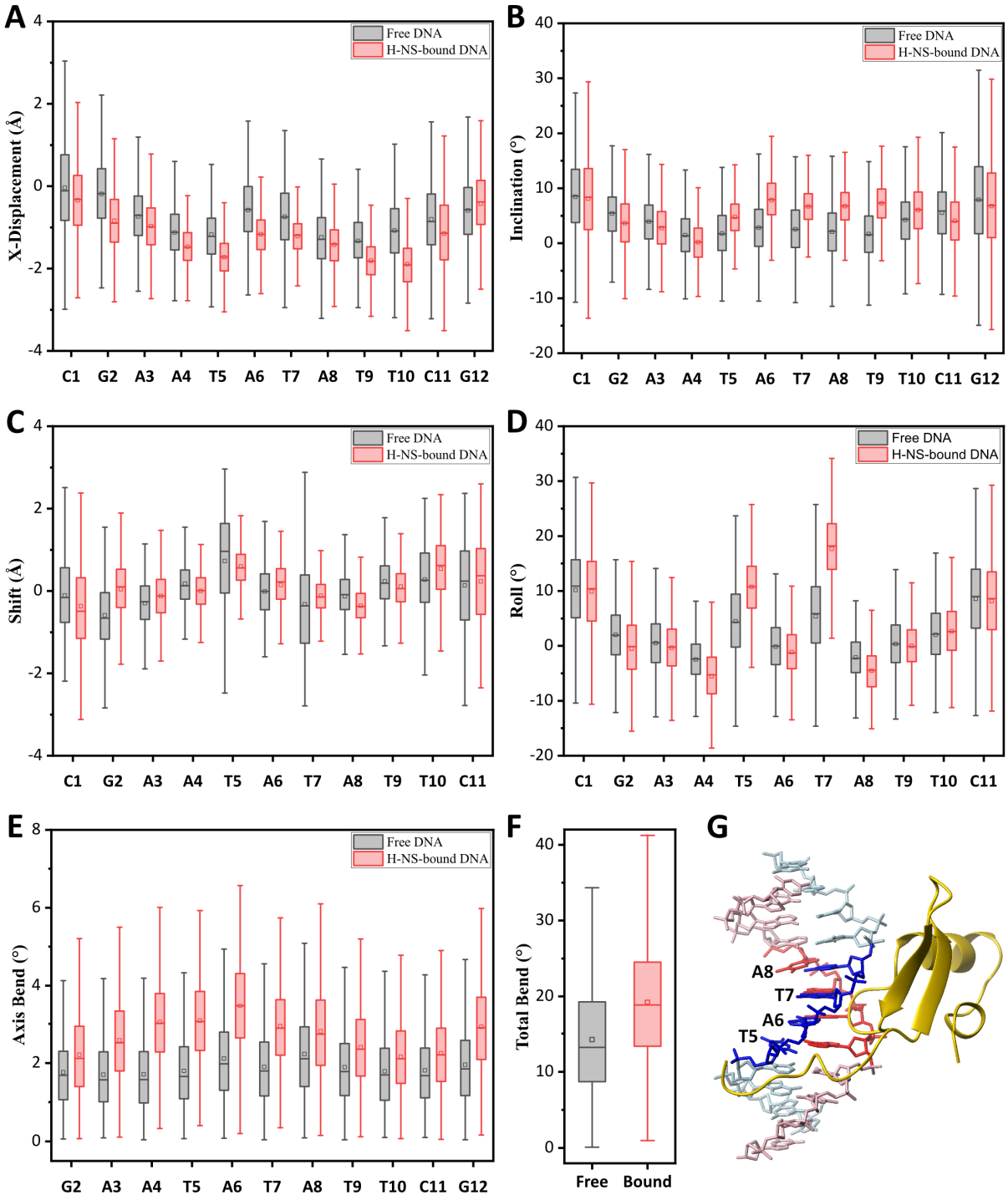

Figure 7. Comparison of distributions of DNA geometric parameters x-displacement (**A**), inclination (**B**), shift (**C**), roll (**D**), axis bend (**E**), and total bend (**F**) values of the selected 1000 frames from MD simulations of the structure of H-NS$^{85–137}$ in complex with Top1 DNA and Top1 DNA alone. (**G**) Representative structure showing H-NS$^{85–137}$ bound DNA conformation, with residues T_5, A_6, T_7, and A_8 highlighted in blue, and the corresponding base-paired residues in red. In the box plot, the box represents the interquartile range (IQR), with the horizontal line inside the box indicating the median value and the small open squares denoting the mean values. The whiskers extend to the farthest data points within 1.5 times the IQR.

3.4. Classification of H-NS Family Proteins

We have previously reported that the Bv3F-like proteins, with the RGR core tripeptide in the AT-hook like motif, can be classified as a new group of xenogeneic silencers, as they have some unique features in the primary sequences that are not found in canonical H-NS with the QGR core tripeptide [13]. Here, we have systematically analyzed the primary sequences of proteins annotated as H-NS family proteins with sequence homology to the DNA binding domain sequence of H-NS or Bv3F, in the NCBI protein database (RefSeq Select proteins, refseq_select). The results show that there exist other variants of the core tripeptide, such as NGK, RGK, RGP, and VGR, in addition to QGR and RGR. A total of over 31,400 sequences annotated as H-NS family proteins were used for the analysis, among which over 99% are from the *Pseudomonadota* (*Proteobacteria)* phylum.

We first grouped H-NS family proteins based on the individual core tripeptide signature of the AT-hook-like motif. Comparison of the primary sequences revealed that all H-NS family proteins have a conserved 8-residue sequence W(T/S)G(Q/R/z)G(R/K/z)xP for the AT-hook-like motif (Figure 8A), in which the first Trp and the last Pro residues are invariant, along with two Gly residues in between. We propose that this 8-residue sequence be defined as the AT-hook-like motif for H-NS family proteins. Two additional conserved sequence features termed "KY" and "FLI" signatures are identified for the DNA binding domain of H-NS family proteins. The "KY" signature refers to two consecutive conserved Lys (or Arg) and Tyr (or Phe) residues preceding the AT-hook-like motif, corresponding to residues K96 and Y97 in *E. coli* H-NS. The "FLI" signature refers to the three residues corresponding to C-terminal residues F133, L134, and I135 in *E. coli* H-NS, which could also be replaced by other hydrophobic residue types in other H-NS family proteins. "KY" and "FLI" signatures thus define the boundaries of the folded DNA binding domain of H-NS family proteins. According to the spacings of "KY" and "FLI" signatures relative to the 8-residue AT-hook-like motif, we have classified the H-NS family proteins into three types, among the H-NS-family sequences retrieved (Figure 8).

Type I H-NS is represented by the canonical *E. coli* H-NS with "QGR" core tripeptide, and its "KY" and "FLI" signatures are separated from the AT-hook-like motif (Type I logo residues 26–33) by ~11 and ~16 residues, respectively. As in the DNA binding domain of Type I H-NS, the second residue following the first α-helix is an invariant Gly residue (G44 in the sequence logo of Figure 8A), and the corresponding residue in *E. coli* H-NS allows a sharp turn after the extended α-helix so that the C-terminal "FLI" signature can fold back to complete the hydrophobic packing against the "KY" signature, given the exceptional conformational freedom of the Gly residue (Figure 8B). As a result, the backbone carbonyl group of the "KY" signature residue K13 hydrogen-bonds to the NH group of the "FLI" signature residue I52, which also engages in side-chain contact with residue A12. Residue L51 from the "FLI" signature of Type I logo makes hydrophobic interaction with both "KY" signature residues K13 and Y14, while residue F50 in Type I logo is involved in core packing with residues I36 and A39 in the α-helix, while residue F50 of the "FLI" signature directly participates in the hydrophobic core packing (herein, residue numbers in this section refer to those in the sequence logos of Figure 8A, for comparison). Therefore, both "KY" and "FLI" signatures are important for the structural packing and the proper conformation of the DNA binding domain of Type I H-NS, as revealed by the representative complex structure of *E. coli* H-NS$^{85–137}$.

Type II H-NS is composed of Bv3F-like proteins, and the spacings of "KY" and "FLI" signatures from the AT-hook-like motif (Type II logo residues 26–33) are 8 or 9 residues, respectively (Figure 8A). Due to the short spacings, a one-turn 3_{10} helix replaces the long α-helix of Type I H-NS and both β-strands only have two residues instead of four, as in the structure of Bv3F DNA binding domain (PDB ID 2L92) (Figure 8B). Type II H-NS contains a characteristically conserved Trp residue (residue W35 in the Type II logo, or residue W95 in 2L92) following the AT-hook-like motif, but it lacks the conserved Gly residue between the AT-hook-like motif and the "FLI" signature in Type I H-NS. Similar to the case in Type I H-NS, the carbonyl group of Type II logo residues K16 (K76 in 2L92) of "KY" signature is hydrogen-bonded by the NH group of residue I46 (I105 in 2L92), whose side-chain also contacts residue P15 and R18 (P75 and R78 in 2L92). The side-chain of residue L45 from the "FLI" signature in the Type II logo (L104 in 2L92) also hydrophobically interacts with both K16 and Y17 (Y77 in 2L92) of the "KY" signature, while Type II logo residue F44 (F103 in 2L92) participates in the hydrophobic core through π–π stacking with W35 on the first 3_{10}-helix.

Type III H-NS has both the conserved Gly residue of Type I and Trp residue of Type II between the AT-hook-like motif and "FLI" signature. The spacings of "KY" and "FLI" signatures from the AT-hook-like motif (Type III logo residues 26–33) are respectively ~9 and ~15 residues, slightly shorter than those of Type I H-NS (Figure 8A). To analyze the structural features, we generated a predicted representative structural model of Type III H-NS with AlphaFold 3, using the sequence (residues 53–103) of H-NS family protein WP_095357136.1 from *Acinetobacter baumannii,* with an RGK core tripeptide sequence (Figure 8B). The folded DNA binding domain was predicted with high confidence (mean pLDDT = 93.6; pTM = 0.69), whereas the positively charged N-

terminal linker showed substantially lower confidence (mean pLDDT = 72.79), consistent with a non-conserved flexible linker. The predicted Type III H-NS structure also has a long α-helix following the AT-hook-like motif as Type I does, but the two β-strands are the same as those in Type II H-NS. The packing patterns of residues from "KY" and "FLI" signatures should resemble those in Type II H-NS, and the NH group of residue I51 of the "FLI" signature in Type III logo is also hydrogen-bonded to the carbonyl group of residue K15 from the "KY" signature. To be noted, both Type II and Type III H-NS carry many positively charged residues within the linker region preceding the "KY" signature of the DNA binding domain, but their positions are not conserved, and it is likely that they contribute to the DNA binding through non-specific electrostatic interaction.

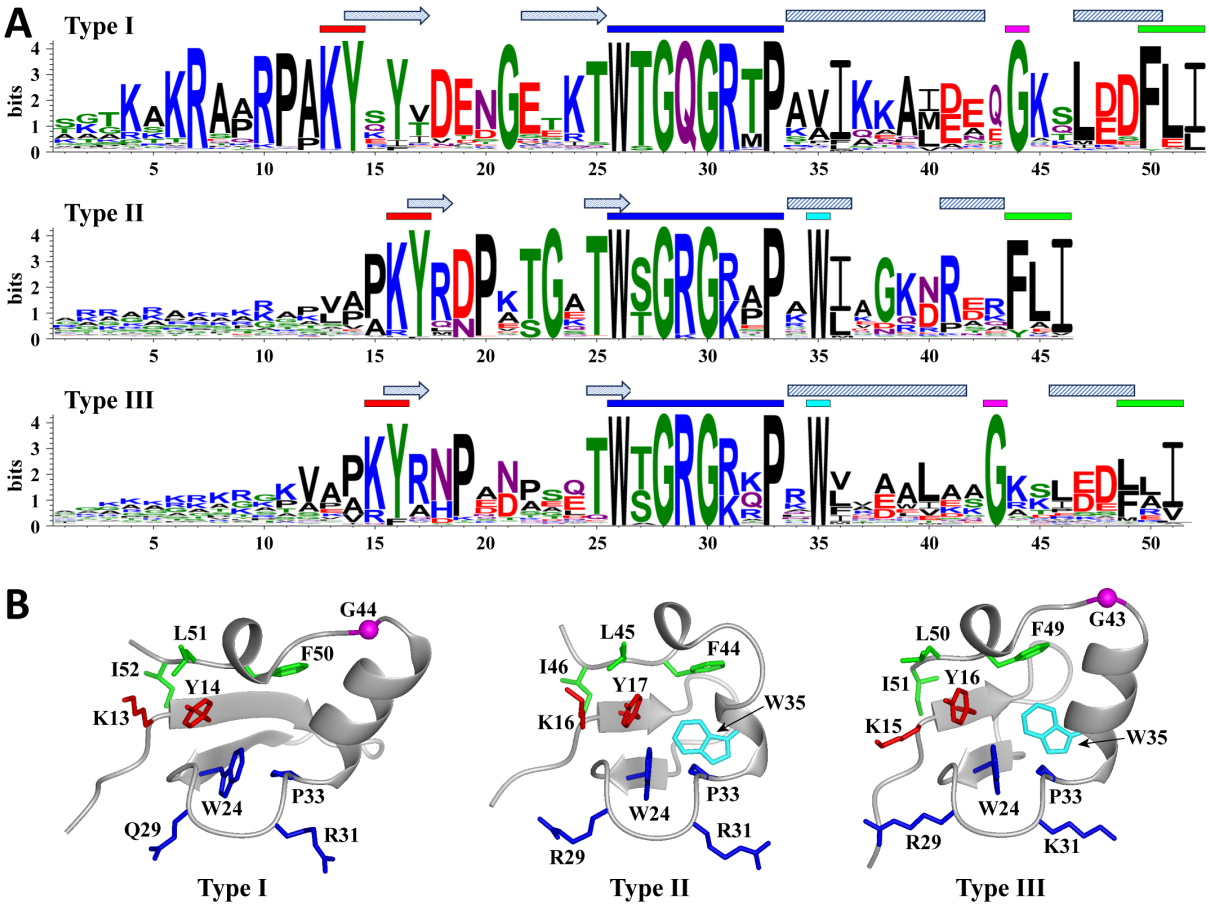

Figure 8. Three types of H-NS family proteins based on comparative analsis of the DNA binding domain sequences. (**A**) Representative sequence logos for the three types of H-NS family proteins based on analysis of the DNA binding domain sequences. AT-hook-like motif is indicated by blue bar above, where "KY" and "FLI" signatures are indicated by red and green bars above, respectively. Conserved Trp and Gly residues are indicated by cyan and magenta bars above, respectively. Secondary structure elements are indicated by shaded arrows (β-strands) and shaded boxes (helices). (**B**) Representative DNA binding domain structures of Type I (*E. coli* H-NS, this study), Type II (*B. vietnamiensis* Bv3F, PDB ID 2L92) and Type III (predicted by AlphaFold 3 with *A. baumannii* WP_095357136.1 sequence) H-NS family proteins. Side-chains of residues of "KY" and "FLI" signatures, as well as in AT-hook-like motif, are shown. Conserved Gly and Trp residues are indicated with balls at C^{α} positions. The color codes are the same as in (**A**). The labeled residues numbers refer to those in the corresponding logos in (**A**).

It is interesting to notice that 99.5% of the ~12,400 Type I H-NS sequences we analyzed are from bacteria of the *Gammaproteobacteria* class, with the most abundant ones from genera such as *Escherichia*, *Aeromonas*, *Klebsiella*, *Salmonella*, *Vibrio*, *Enterobacter*, *Pantoea*, *Citrobacter*, *Photobacterium*, and *Shewanella*, etc. For Type II H-NS, 96.6% of the over 7800 sequences are from bacteria of the *Betaproteobacteria* class, including those from genera *Burkholderia*, *Paraburkholderia*, *Cupriavidus*, *Ralstonia*, *Caballeronia*, *Comamonas*, *Delftia*, and *Acidovorax*, etc. However, among the ~11,200 Type III H-NS sequences, 35.0% are from the *Alphaproteobacteria* class (genera *Bradyrhizobium*, *Paracoccus*, *Paracoccaceae*, *Ruegeria*, and *Leisingera*, etc.), 28.0% from the *Betaproteobacteria* class (genera *Massilia*, *Janthinobacterium*, *Achromobacter*, *Cupriavidus*,

Rubrivivax, and *Rhodoferax*, etc.), and 30.1% from the *Gammaproteobacteria* class (genera *Acinetobacter*, *Stenotrophomonas*, *Xanthomonas*, *Rhodanobacter*, *Dyella*, and *Psychrobacter*, etc.).

The three types of H-NS have characteristic sequences for the DNA base contacting the three-residue core tripeptide signatures. Most Type I H-NS proteins have the canonical QGR core tripeptide signature, but there are also a small number of sequences with variant signatures, including QGK, NGK, NGR, IGR, VGR and RGR (Figure 9). It appears that the conserved arginine residue at logo position 7 (R7) in Type I H-NS proteins bearing the QGR or QGK core tripeptide signature, corresponding to the minor groove contacting residue R90 in *E. coli* H-NS, is absent from Type I H-NS proteins with other core tripeptide signatures, including RGR, NGR, IGR, and VGR. Our mutagenesis experiments showed that all three substitutions at Q112 (Q112R, Q112N, and Q112V) result in a greater than fourfold reduction in DNA binding affinity. These findings raise the possibility that R90 in *E. coli* H-NS is only compatible structurally or functionally with the QGR/QGK core tripeptide signatures, but not with RGR, NGR, IGR, or VGR signatures.

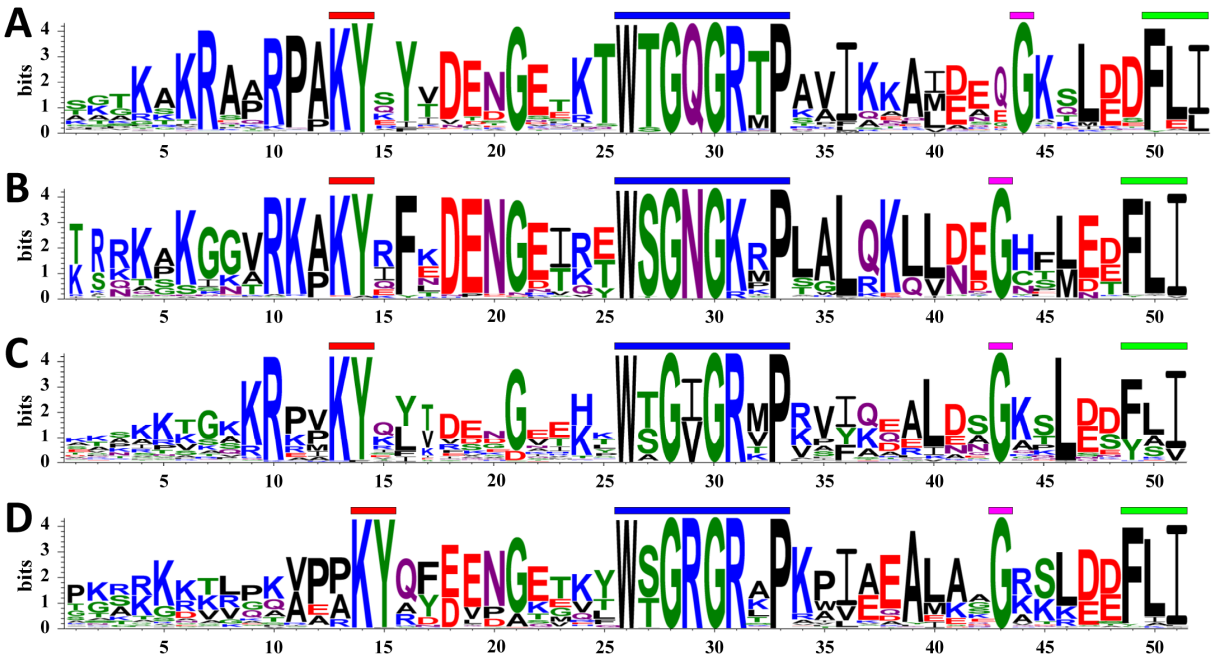

Figure 9. Sequence logos of the DNA binding domains of Type I H-NS family proteins with different core tripeptide sequences: (**A**) QGR (1613) and QGK (27); (**B**) NGK (152) and NGR (12); (**C**) IGR (149) and VGR (148); (**D**) RGR (207). AT-hook-like motif is indicated by blue shaded box above, where "KY" and "FLI" signatures are indicated by red and green shaded boxes above, respectively. The conserved Gly residue is indicated by magenta bars above. The number in the parentheses following each core tripeptide indicates the number of non-redundant sequences used in generating the logos.

Most of the Type II and Type III H-NS proteins have the tripeptide signatures of RGR or RGK (Figures 10 and 11), and there also exist variants including KGR, KGK, KGA, KGP, RGN, RGP, HGR, HGK, AGR, FGK, FGR, and LGR for Type II, and RGA, RGI, RGL, RGM, RGV, RGQ, MGK, MGR, and VGK for Type III. A total of 29 tripeptide signature sequences were identified, all of which have at least one positively charged Arg or Lys on either side of the invariant central Gly residue. Except for RGR and RGK, which are shared by different types of H-NS, all other core signature sequences are unique to a specific type.

Unexpectedly, we have also found some sequences with "HAR" signature (Supplementary Figure S26), in which the central Gly is replaced by Ala, and the His residue could be substituted by Trp or Arg. For these sequences, a highly conserved proline appears at the penultimate position of the AT-hook-like motif. In addition, although most H-NS family proteins only have one DNA binding domain, there are some proteins harboring two DNA binding domains in tandem. The two domains are separated by a non-conserved, positively charged linker of more than 20 residues, which is predicted to be a flexible loop by AlphaFold 3. All these sequences share the characteristic features of Type II H-NS.

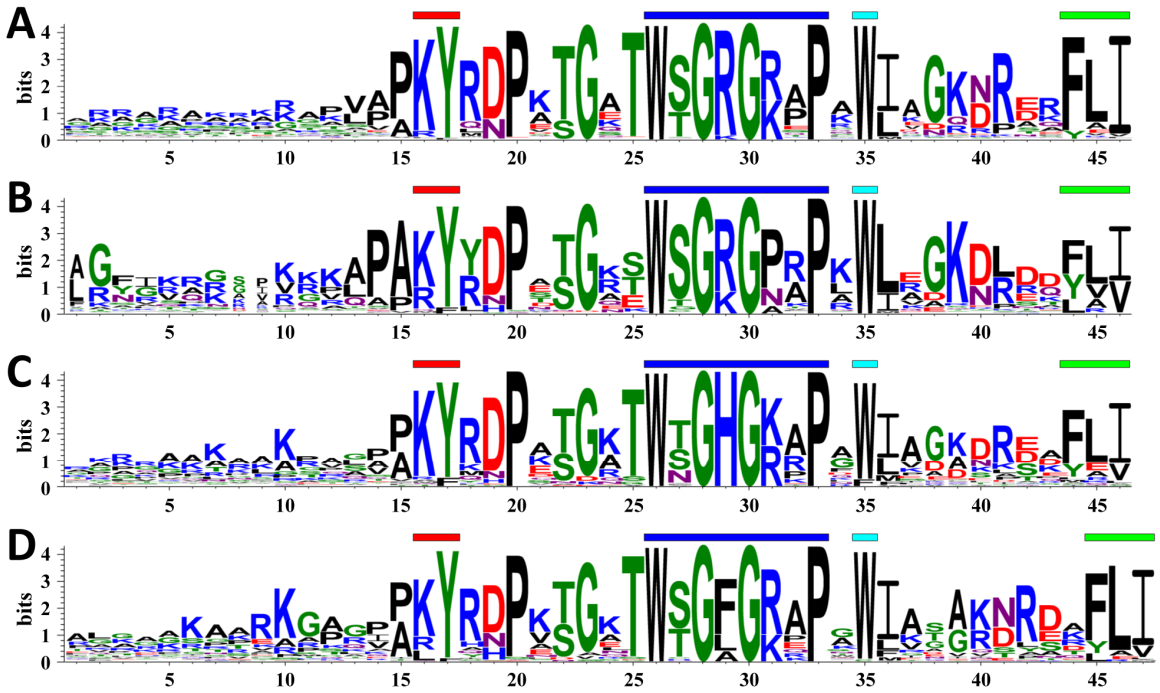

Figure 10. Sequence logos of the DNA binding domains of Type II H-NS family proteins with different core tripeptide sequences: (**A**) RGR (1815), RGK (1505), KGR (115), and KGK (36); (**B**) RGP (264), RGN (101), KGP (67), and KGA (45); (**C**) HGK (31) and HGR (27); (**D**) FGR (32), FGK (25), LGR (17), and AGR (11). AT-hook-like motif is indicated by blue shaded box above, where "KY" and "FLI" signatures are indicated by red and green shaded boxes above, respectively. The conserved Trp residues is indicated by cyan bars above, respectively. The number in the parentheses following each core tripeptide indicates the number of non-redundant sequences used in generating the logos.

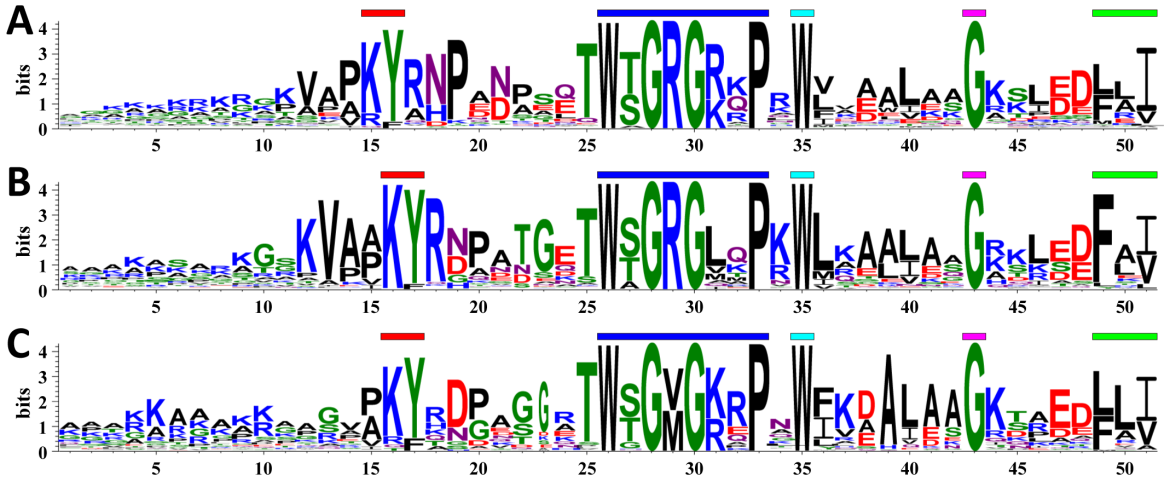

Figure 11. Sequence logos of the DNA binding domains of Type III H-NS family proteins with different core tripeptide sequences: (**A**) RGR (3894) and RGK (2032); (**B**) RGL (855), RGV (151), RGM (137), RGA (71), RGQ (69), and RGI (60); (**C**) VGK (78), MGK (58), and MGR (14). AT-hook-like motif is indicated by blue shaded box above, where "KY" and "FLI" signatures are indicated by red and green shaded boxes above, respectively. Conserved Trp and Gly residues are indicated by cyan and magenta bars above, respectively. The number in the parentheses following each core tripeptide indicates the number of non-redundant sequences used in generating the logos.

4. Discussion

The molecular basis by which H-NS recognizes AT-rich DNA has been investigated for more than two decades. Previous structural studies established that the conserved AT-hook-like motif is responsible for minor-groove recognition, whereas biochemical and genetic studies suggested that the adjacent linker region also contributes substantially to DNA binding and gene silencing. However, in the absence of an experimentally determined H-NS/DNA complex structure, the structural basis for the interaction between H-NS and DNA has remained unclear.

The present structure of H-NS$^{85–137}$ in complex with Top1-DNA confirms the central role of the conserved AT-hook-like motif in DNA recognition and further suggests that the linker region contributes to DNA binding through structurally coordinated interactions with both DNA and the AT-hook-like motif. In the complex structure, residues Q112 and R114 of the AT-hook-like motif form multiple hydrogen bonds with DNA bases. Importantly, the structure reveals that the linker region is not merely an electrostatic extension of the DNA binding domain, but forms an integral part of the DNA recognition interface. The complex structure reveals that the side-chain of residue R90 inserts into the minor groove adjacent to residue Q112, and establishes direct hydrogen-bonding interaction with DNA bases, expanding the DNA-base-contacting region beyond the canonical AT-hook-like motif. These hydrogen bonds formed with N3 atom of adenine or O2 atom of thymidine clearly explain H-NS's preference for AT-rich DNA over GC-rich DNA, as the guanine NH_2 group positioned in the minor groove would prevent H-NS from properly forming hydrogen bonds with DNA bases.

In the complex structure, the guanidinium group of residue R93 is hydrogen bonded to residue T110 of the AT-hook-like motif and directly participates in the hydrogen bond network centered on residue Q112, which maintains and stabilizes the AT-hook-like motif. Both the R93A and Q112A mutations dramatically reduce the DNA-induced NH chemical shift perturbations of linker residues, but not those of the AT-hook-like motif residues. In contrast, the R90A mutation almost does not affect the NH chemical shift patterns of the AT-hook-like motif residues. These observations suggest that the interaction of the linker residue R90 with DNA depends on proper engagement of the AT-hook-like motif, and is facilitated by the hydrogen bond between residue R93 and T110, whereas disruption of linker contacts alone does not substantially alter the binding mode of the AT-hook-like motif.

Full-length H-NS binds DNA as an oligomer, in a multivalent interaction manner. It has been shown that full-length H-NS and its monomeric DNA binding domain (residues 90–137) display the same DNA binding specificity based on the primary binding sites within the promoter region of *virF* detected by footprinting experiments [43], suggesting that its DNA recognition mechanism should be relatively independent of the N-terminal oligomerization domain. The DNA contacting features revealed by the complex structure of H-NS$^{85–137}$ with Top1-DNA are consistent with existing *in vivo* functional analyses and *in vitro* DNA binding assay results for full-length H-NS. It was reported that point mutations R90H, R90C, and R93C of *E. coli* H-NS resulted in de-repression of the *proVWX* operon from a systematic mutational analysis study [44], and H-NS mutant R93H failed to repress *proU* and *fimB* expression [45]. Mutants R90G and R93G of *E. coli* H-NS showed lower DNA affinities than the WT protein from EMSA analysis [20]. Recently, McLelland et al. reported an *hns* hypomorph strain of *Salmonella enterica* serovar Typhimurium, with T110A mutation, and it was demonstrated that the T110A mutation results in reduced DNA binding affinity for full-length H-NS, which increases the proportion of virulence-expressing cells, resulting in enhanced epithelial cell infection *in vitro* [46].

The convergence of our structural and mutational data offers a plausible explanation for the functional significance of several PTMs or mutations. Several lysine residues, including K96, K107, and K121, are modified by acetylation or 2-hydroxyisobutyrylation *in vivo*, and these modifications have been reported to impair DNA binding and relieve xenogeneic silencing. In the present structure, all three residues contribute to maintaining the structural integrity of the DNA binding domain through electrostatic interactions. Consistent with these structural observations, mutations at these positions perturb the NMR spectra and reduce DNA binding affinity despite the absence of direct protein-DNA contacts. The phosphorylation of T106 or the T106E mutation is likely to affect the conformational stability of the L1 loop by introducing a negative charge, as residue T106 is right beside the highly negatively charged L1 loop. With the complex structure, we can also easily predict the outcomes of some of the other PTMs without functional annotation. Phosphorylation of residue Y99 is predicted to weaken the DNA binding affinity of H-NS, as the hydroxyl O atom is within 5.3 Å of the phosphate O atom of DNA residue T_9. Residue N103 resides in loop L1 and its NH2 group forms a hydrogen bond with the carboxyl group of residue D101, which stabilizes the loop. Deamidation of residue N103 could result in partial unfolding of the DNA binding domain structure by eliminating the hydrogen bond, while simultaneously introducing one more negative charge to this already highly negatively charged loop. Acetylation or succinylation of residue K128 neutralizes its positive charge and abolishes its electrostatic interaction with D132, which may destabilize the DNA binding domain and thereby reduce DNA binding affinity. Phosphorylation of residue S129 adds a negative charge that could perturb the hydrogen bond of its hydroxyl group formed by the NH group of residue D131 due to electrostatic repulsion, potentially destabilizing the packing of the 3_{10} helix.

Although the differences in DNA geometric parameters between free DNA and H-NS-bound DNA may be specific to the sampled trajectories, the most pronounced H-NS binding induced structural alterations occur within the central six-base-pair region, e.g., where H-NS$^{85–137}$ binds in the complex structure (Figure 7G). In the MD simulation trajectories, H-NS binding is accompanied by localized changes in DNA geometry, characterized by increased minor groove width, reduced minor groove depth, altered base-pair displacement and inclination,

increased roll values, and enhanced bending. This pronounced minor groove expansion may provide a physical rationale for the preferential binding of H-NS to AT-rich sequences containing TpA steps rather than rigid A-tracts. TpA step exhibits inherent conformational flexibility and has the most positive bendability, and having more TpA steps makes DNA more bendable [47], making them conformationally predisposed to accommodate the DNA deformation upon H-NS insertion. In contrast, opening the rigid, narrow minor groove of A-tract DNA demands a higher energetic penalty, which may contribute to the lower binding affinities observed for A-tract sequences. A previous MD study based on the free H-NS DNA binding domain structure also reported that H-NS binding induces a structural adaptation in the AT-rich DNA, including minor groove widening and increased bending, as well as localized B-to-A conformation transition [23]. Indeed, the directions of x-displacement (more negative) and inclination (more positive) value changes induced by H-NS binding in our MD simulation results are characteristic of a shift towards an A-like conformation [48].

H-NS has long been known to prefer binding intrinsically curved DNA [49], which is consistent with our observations that H-NS binding increases the total bend of Top1 DNA, along with axis bend values for all 12 base pairs. It was found that sequence-dependent DNA intrinsic curvature (high bendability) is the key determinant in positioning plectonemes of supercoiled DNA, and analysis of prokaryotic genomes indicated that plectonemes localize directly upstream of promoters [50]. In recent years, more and more studies indicate that the function of H-NS depends on the interplay between the 3D structure of chromosomal DNA and gene transcription [51]. As an oligomer, full-length H-NS binds DNA in two modes, either DNA-H-NS-DNA bridges or H-NS-DNA filaments. It was shown that the osmo-responsive *proVWX* operon of *E. coli* is switched on and off in response to osmolarity, by the disruption and restoration of two H-NS-mediated bridges [52]. A recent study suggested that H-NS preferentially binds negatively supercoiled DNA and bridge the opposite arms of the plectoneme, while positive supercoils generated by an upstream gene transcription could diffuse to and relax the H-NS-bound negative supercoil, causing H-NS to dissociate from DNA and thus alleviating its transcription repression of *hilD* gene in *Salmonella* [53]. We have also shown that H-NS displays much higher transcription suppression ability on negatively supercoiled plasmid DNA template than linearized DNA template, using the quantitative *in vitro* transcription assay QRIVTA [54]. It is possible that negative supercoiling not only promotes H-NS-mediated DNA bridging but also renders AT-rich DNA locally prone to adopting a bend conformation that favors H-NS binding.

The primary sequence analysis extends our structural study of *E. coli* H-NS$^{85–137}$ to the broader H-NS protein family. Despite considerable sequence diversity, the DNA binding domain of all H-NS family proteins contains a conserved AT-hook-like motif, together with characteristic “KY” and “FLI” sequence signatures that define the boundaries of the folded DNA binding domain. Based on these conserved structural features, H-NS family proteins are classified into three major types. The experimentally determined structure presented here represents the canonical Type I H-NS. This structure, together with our mutagenesis result, help explain why the additional DNA base-contacting residue R90 is conserved primarily in QGR-containing proteins of Type I H-NS, but absent from the others, as well as Type II and Type III proteins. It shows again that the AT-hook-like motif is the determinant for H-NS to bind the AT-rich DNA minor groove, while the interaction of the guanidium group of residue R90 with DNA bases plays an auxiliary role in *E. coli* H-NS. Taken together, all H-NS family proteins employ a common mechanism for recognizing AT-rich DNA, in which the AT-hook-like motif provides the central structural element for minor-groove binding.

The three types of H-NS identified in this study therefore are variations of a conserved structural scaffold, and the variations may influence the local detail conformation and dynamics of the AT-hook-like motif or the surrounding structural environment, thereby modulating DNA binding properties while maintaining the overall recognition mechanism. Within this conserved structural framework, the AT-hook-like motif exhibits notable sequence diversity, particularly in the core tripeptide region. Our analysis identified nearly thirty core tripeptide sequence signatures, with QGR, RGR, and RGK representing the major variants. Importantly, the core tripeptide variations are not randomly distributed but are associated with different types of H-NS family proteins, suggesting that sequence diversification of the core tripeptide has occurred together with broader structural adaptation. A recent study reported that Sfx, belonging to the Type I H-NS family with an RGR core tripeptide, encoded by the model IncX2 plasmid R6K, can repress R6K conjugation, but chromosomal H-NS cannot. Mutating the RGR of Sfx into QGR dramatically decreased its conjugation repression activity by more than 50-fold [55]. Interestingly, a QGR to RGR (Q112R) mutation enables H-NS to display ~10-fold repression of the conjugation of R6K plasmids lacking *sfx*, even though our results indicate that the Q112R mutation reduces the DNA binding affinity of H-NS$^{85–137}$ by more than fourfold. It has been speculated that it is the DNA binding preference of H-NS, rather than the DNA binding affinity, prevents H-NS from suppressing R6K conjugation. As the xenogeneic silencer Lsr2, containing an AT-hook-like motif with an RGR core tripeptide, prefers to bind A-tract DNA but not TpA steps [56]. Therefore, it is possible that the diversity of the H-NS core tripeptide is related to target gene selection in different bacteria.

H-NS family proteins are restricted to the *Pseudomonadota* (*Proteobacteria*) phylum of bacteria. The phylogenetic distributions of three types of H-NS—Type I predominantly in the *Gammaproteobacteria* class, Type II in the *Betaproteobacteria* class, and Type III spread across the *Alpha-*, *Beta-*, and *Gammaproteobacteria* classes—suggest that these sequence and structural features have been maintained over evolutionary time and are not simply the result of random sequence drift. The present taxonomic distribution is compatible with lineage-associated diversification, but does not by itself establish the order of origin of the three groups. However, the functional significance of the classification and distribution of H-NS family proteins, along with the sequence diversity of the core tripeptide, particularly whether these characteristic features are responsible for differences in target gene selection or resulted from adaptation to, or co-evolution with, corresponding bacterial genomes, remains to be investigated.

5. Conclusions

In summary, the solution structure of the *E. coli* H-NS$^{85–137}$/DNA complex establishes the AT-hook-like motif as the central determinant for AT-rich DNA recognition and reveals that the adjacent linker contributes directly to DNA binding through coordinated interactions involving R90 and R93. Sequence analysis further shows that, despite substantial diversity among H-NS family proteins, a conserved AT-hook-like motif is shared across three types of DNA binding domains, supporting a common mechanism for minor-groove recognition.

Supplementary Materials: The following supporting information can be downloaded at: https://media.sciltp.com/articles/others/2611201344123083/HM-26080130-SI.pdf.

Author Contributions: B.X. and B.D.: conceived project; J.W. and F.L.: conducted experiments and analyzed data; J.W., B.D. and B.X.: wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by National Natural Science Foundation of China [32371260 and 32571389 to B.X.], and by financial support from Beijing National Laboratory for Molecular Sciences.

Institutional Review Board Statement: No animal and human participants were involved in this study.

Informed Consent Statement: Not applicable.

Data Availability Statement: The solution structure of H-NS$^{85–137}$ and Top1 DNA complex has been deposited at The RCSB Protein Data Bank (www.rcsb.org) under accession number 25NZ. All chemical shift assignments were deposited at the BioMagResBank (www.bmrb.wisc.edu) under BMRB ID 36843.

Acknowledgments: All NMR experiments were performed at the Beijing NMR Center and the NMR facility of National Center for Protein Sciences at Peking University. We thank Hongwei Li and Xiaogang Niu for their assistance in NMR data collection.

Conflicts of Interest: The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies: No AI tools were utilized for this paper.

References

1. Fang, F.C.; Rimsky, S. New insights into transcriptional regulation by H-NS. *Curr. Opin. Microbiol.* **2008**, *11*, 113–120. https://doi.org/10.1016/j.mib.2008.02.011.
2. Ishihama, A.; Shimada, T. Hierarchy of transcription factor network in *Escherichia coli* K-12: H-NS-mediated silencing and Anti-silencing by global regulators. *FEMS Microbiol. Rev.* **2021**, *45*, fuab032. https://doi.org/10.1093/femsre/fuab032.
3. Navarre, W.W.; McClelland, M.; Libby, S.J.; Fang, F.C. Silencing of xenogeneic DNA by H-NS—Facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev.* **2007**, *21*, 1456–1471. https://doi.org/10.1101/gad.1543107.
4. Ali, S.S.; Xia, B.; Liu, J.; Navarre, W.W. Silencing of foreign DNA in bacteria. *Curr. Opin. Microbiol.* **2012**, *15*, 175–181. https://doi.org/10.1016/j.mib.2011.12.014.
5. Will, W.R.; Navarre, W.W.; Fang, F.C. Integrated circuits: How transcriptional silencing and counter-silencing facilitate bacterial evolution. *Curr. Opin. Microbiol.* **2015**, *23*, 8–13. https://doi.org/10.1016/j.mib.2014.10.005.
6. Dorman, C.J. H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid* **2014**, *75*, 1–11. https://doi.org/10.1016/j.plasmid.2014.06.004.
7. Vera, M.A.G.; Brune, A. Hand over that gun: Lateral gene transfer provides an amoeba with a bacterial weapon. *Environ. Microbiol.* **2017**, *19*, 847–848. https://doi.org/10.1111/1462-2920.13680.
8. Picker, M.A.; Wing, H.J. H-NS, Its Family Members and Their Regulation of Virulence Genes in Shigella Species. *Genes* **2016**, *7*, 112. https://doi.org/10.3390/genes7120112.
9. Ayala, J.C.; Silva, A.J.; Benitez, J.A. H-NS: An overarching regulator of the Vibrio cholerae life cycle. *Res. Microbiol.* **2017**, *168*, 16–25. https://doi.org/10.1016/j.resmic.2016.07.007.

10. Platenkamp, A.; Mellies, J.L. Environment Controls LEE Regulation in Enteropathogenic Escherichia coli. *Front. Microbiol.* **2018**, *9*, 1694. https://doi.org/10.3389/fmicb.2018.01694.
11. Cooper, C.; Legood, S.; Wheat, R.L.; Forrest, D.; Sharma, P.; Haycocks, J.R.J.; Grainger, D.C. H-NS is a bacterial transposon capture protein. *Nat. Commun.* **2024**, *15*, 7137. https://doi.org/10.1038/s41467-024-51407-5.
12. Singh, K.; Milstein, J.N.; Navarre, W.W. Xenogeneic Silencing and Its Impact on Bacterial Genomes. *Annu. Rev. Microbiol.* **2016**, *70*, 199–213. https://doi.org/10.1146/annurev-micro-102215-095301.
13. Duan, B.; Ding, P.; Navarre, W.W.; Liu, J.; Xia, B. Xenogeneic Silencing and Bacterial Genome Evolution: Mechanisms for DNA Recognition Imply Multifaceted Roles of Xenogeneic Silencers. *Mol. Biol. Evol.* **2021**, *38*, 4135–4148. https://doi.org/10.1093/molbev/msab136.
14. Arold, S.T.; Leonard, P.G.; Parkinson, G.N.; Ladbury, J.E. H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 15728–15732. https://doi.org/10.1073/pnas.1006966107.
15. Gordon, B.R.G.; Li, Y.; Cote, A.; Weirauch, M.T.; Ding, P.; Hughes, T.R.; Navarre, W.W.; Xia, B.; Liu, J. Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10690–10695. https://doi.org/10.1073/pnas.1102544108.
16. Fonfría-Subirós, E.; Acosta-Reyes, F.; Saperas, N.; Pous, J.; Subirana, J.A.; Campos, J.L. Crystal Structure of a Complex of DNA with One AT-Hook of HMGA1. *PLoS ONE* **2012**, *7*, e37120. https://doi.org/10.1371/journal.pone.0037120.
17. Fitzgerald, S.; Kary, S.C.; Alshabib, E.Y.; MacKenzie, K.D.; Stoebel, D.M.; Chao, T.C.; Cameron, A.D.S. Redefining the H-NS protein family: A diversity of specialized core and accessory forms exhibit hierarchical transcriptional network integration. *Nucleic Acids Res.* **2020**, *48*, 10184–10198. https://doi.org/10.1093/nar/gkaa709.
18. Gao, Y.; Foo, Y.H.; Winardhi, R.S.; Tang, Q.; Yan, J.; Kenney, L.J. Charged residues in the H-NS linker drive DNA binding and gene silencing in single cells. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12560–12565. https://doi.org/10.1073/pnas.1716721114.
19. Gulvady, R.; Gao, Y.; Kenney, L.J.; Yan, J. A single molecule analysis of H-NS uncouples DNA binding affinity from DNA specificity. *Nucleic Acids Res.* **2018**, *46*, 10216–10224. https://doi.org/10.1093/nar/gky826.
20. Fernández-de-Alba, C.; Berrow, N.S.; Garcia-Castellanos, R.; García, J.; Pons, M. On the Origin of the Selectivity of Plasmidic H-NS towards Horizontally Acquired DNA: Linking H-NS Oligomerization and Cooperative DNA Binding. *J. Mol. Biol.* **2013**, *425*, 2347–2358. https://doi.org/10.1016/j.jmb.2013.03.006.
21. Riccardi, E.; van Mastbergen, E.C.; Navarre, W.W.; Vreede, J. Predicting the mechanism and rate of H-NS binding to AT-rich DNA. *PLoS Comput. Biol.* **2019**, *15*, e1006845. https://doi.org/10.1371/journal.pcbi.1006845.
22. van Heesch, T.; Bolhuis, P.G.; Vreede, J. Decoding dissociation of sequence-specific protein–DNA complexes with non-equilibrium simulations. *Nucleic Acids Res.* **2023**, *51*, 12150–12160. https://doi.org/10.1093/nar/gkad1014.
23. van Heesch, T.; Sharma, S.; van Erp, B.; Ortíz, A.P.D.A.; Dame, R.T.; Vreede, J.; Gavvala, K. Adaptation of DNA to Protein Binding Revealed by Spectroscopy and Molecular Simulation. *J. Phys. Chem. B* **2025**, *129*, 5653–5663. https://doi.org/10.1021/acs.jpcb.5c00189.
24. Dilweg, I.W.; Dame, R.T. Post-translational modification of nucleoid-associated proteins: An extra layer of functional modulation in bacteria? *Biochem. Soc. Trans.* **2018**, *46*, 1381–1392. https://doi.org/10.1042/bst20180488.
25. Liu, Y.; Wang, X. Post-translational modifications of the nucleoid protein H-NS: Sites, mechanisms, and regulatory cues. *FEMS Microbiol. Rev.* **2025**, *49*, fuaf045. https://doi.org/10.1093/femsre/fuaf045.
26. Liu, Y.; Zhou, M.; Bu, Y.; Qin, L.; Zhang, Y.; Shao, S.; Wang, Q. Lysine acetylation regulates the AT-rich DNA possession ability of H-NS. *Nucleic Acids Res.* **2024**, *52*, 1645–1660. https://doi.org/10.1093/nar/gkad1172.
27. Dong, H.; Zhao, Y.; Bi, C.; Han, Y.; Zhang, J.; Bai, X.; Zhai, G.; Zhang, H.; Tian, S.; Hu, D.; et al. TmcA functions as a lysine 2-hydroxyisobutyryltransferase to regulate transcription. *Nat. Chem. Biol.* **2022**, *18*, 142–151. https://doi.org/10.1038/s41589-021-00906-3.
28. Wang, J.; Wei, H.; Xia, B. c-di-GMP does not bind H-NS, nor inhibits H-NS binding DNA. *Nat. Commun.* **2025**, *16*, 5287. https://doi.org/10.1038/s41467-025-60689-2.
29. Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93–158. https://doi.org/10.1016/s0079-6565(98)00025-9.
30. Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293. https://doi.org/10.1007/bf00197809.
31. Johnson, B.A.; Blevins, R.A. NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **1994**, *4*, 603–614. https://doi.org/10.1007/bf00404272.
32. Shen, Y.; Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **2013**, *56*, 227–241. https://doi.org/10.1007/s10858-013-9741-y.
33. Spring-Connell, A.M.; Evich, M.; Germann, M.W. NMR Structure Determination for Oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* **2018**, *72*, 7–28. https://doi.org/10.1002/cpnc.48.

34. Duan, B.; Ding, P.; Hughes, T.R.; Navarre, W.W.; Liu, J.; Xia, B. How bacterial xenogeneic silencer rok distinguishes foreign from self DNA in its resident genome. *Nucleic Acids Res.* **2018**, *46*, 10514–10529. https://doi.org/10.1093/nar/gky836.
35. Duggan, B.M.; Legge, G.B.; Dyson, H.J.; Wright, P.E. SANE (Structure Assisted NOE Evaluation): An automated model-based approach for NOE assignment. *J. Biomol. NMR* **2001**, *19*, 321–329. https://doi.org/10.1023/a:1011227824104.
36. Güntert, P.; Mumenthaler, C.; Wüthrich, K. Torsion angle dynamics for NMR structure calculation with the new program Dyana. *J. Mol. Biol.* **1997**, *273*, 283–298. https://doi.org/10.1006/jmbi.1997.1284.
37. Case, D.A.; Cerutti, D.S.; Cruzeiro, V.W.D.; Darden, T.A.; Duke, R.E.; Ghazimirsaeed, M.; Giambaşu, G.M.; Giese, T.J.; Götz, A.W.; Harris, J.A.; et al. Recent Developments in Amber Biomolecular Simulations. *J. Chem. Inf. Model.* **2025**, *65*, 7835–7843. https://doi.org/10.1021/acs.jcim.5c01063.
38. Case, D.A.; Aktulga, H.M.; Belfon, K.; Cerutti, D.S.; Cisneros, G.A.; Cruzeiro, V.W.D.; Forouzesh, N.; Giese, T.J.; Götz, A.W.; Gohlke, H.; et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191. https://doi.org/10.1021/acs.jcim.3c01153.
39. Blanchet, C.; Pasi, M.; Zakrzewska, K.; Lavery, R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.* **2011**, *39*, W68–W73. https://doi.org/10.1093/nar/gkr316.
40. Madeira, F.; Madhusoodanan, N.; Lee, J.; Eusebi, A.; Niewielska, A.; Tivey, A.R.N.; Lopez, R.; Butcher, S. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **2024**, *52*, W521–W525. https://doi.org/10.1093/nar/gkae241.
41. Crooks, G.E.; Hon, G.; Chandonia, J.M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. https://doi.org/10.1101/gr.849004.
42. Chaurasiya, K.R.; van der Valk, R.; Henneman, B.; Dame, R.T. Vizualizing the Mechanism of H-NS Gene Regulation. *Biophys. J.* **2018**, *114*, 439a. https://doi.org/10.1016/j.bpj.2017.11.2430.
43. Ulissi, U.; Fabbretti, A.; Sette, M.; Giuliodori, A.M.; Spurio, R. Time-resolved assembly of a nucleoprotein complex between Shigella flexneri virF promoter and its transcriptional repressor H-NS. *Nucleic Acids Res.* **2014**, *42*, 13039–13050. https://doi.org/10.1093/nar/gku1052.
44. Ueguchi, C.; Suzuki, T.; Yoshida, T.; Tanaka, K.I.; Mizuno, T. Systematic Mutational Analysis Revealing the Functional Domain Organization of Nucleoid Protein H-NS. *J. Mol. Biol.* **1996**, *263*, 149–162. https://doi.org/10.1006/jmbi.1996.0566.
45. Donato, G.M.; Kawula, T.H. Phenotypic Analysis of Random *hns* Mutations Differentiate DNA-Binding Activity from Properties of *fimA* Promoter Inversion Modulation and Bacterial Motility. *J. Bacteriol.* **1999**, *181*, 941–948. https://doi.org/10.1128/jb.181.3.941-948.1999.
46. McLelland, L.; Spratt, M.; Karki, D.; Marand, A.; Burt, M.; Zhao, J.; Lane, K. Perturbing H-NS function reveals roles in restricting virulence heterogeneity and pathogen adaptation. *PLoS Pathog.* **2026**, *22*, e1014488. https://doi.org/10.1371/journal.ppat.1014488.
47. Basu, A.; Bobrovnikov, D.G.; Cieza, B.; Arcon, J.P.; Qureshi, Z.; Orozco, M.; Ha, T. Deciphering the mechanical code of the genome and epigenome. *Nat. Struct. Mol. Biol.* **2022**, *29*, 1178–1187. https://doi.org/10.1038/s41594-022-00877-6.
48. Dickerson, R.E.; Ng, H.L. DNA structure from A to B. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6986–6988. https://doi.org/10.1073/pnas.141238898.
49. Tupper, A.E.; Owen-Hughes, T.A.; Ussery, D.W.; Santos, D.S.; Ferguson, D.J.; Sidebotham, J.M.; Hinton, J.C.; Higgins, C.F. The chromatin-associated protein H-NS alters DNA topology *in vitro*. *EMBO J.* **1994**, *13*, 258–268. https://doi.org/10.1002/j.1460-2075.1994.tb06256.x.
50. Kim, S.H.; Ganji, M.; Kim, E.; van der Torre, J.; Abbondanzieri, E.; Dekker, C. DNA sequence encodes the position of DNA supercoils. *eLife* **2018**, *7*, e36557. https://doi.org/10.7554/elife.36557.
51. Rashid, F.Z.M.; Dame, R.T. 2024: A "nucleoid space" odyssey featuring H-NS. *BioEssays* **2024**, *46*, e2400098. https://doi.org/10.1002/bies.202400098.
52. Rashid, F.Z.M.; Crémazy, F.G.E.; Hofmann, A.; Forrest, D.; Grainger, D.C.; Heermann, D.W.; Dame, R.T. The environmentally-regulated interplay between local three-dimensional chromatin organisation and transcription of proVWX in *E. coli*. *Nat. Commun.* **2023**, *14*, 7478. https://doi.org/10.1038/s41467-023-43322-y.
53. Figueroa-Bossi, N.; Fernández-Fernández, R.; Kerboriou, P.; Bouloc, P.; Casadesús, J.; Sánchez-Romero, M.A.; Bossi, L. Transcription-driven DNA supercoiling counteracts H-NS-mediated gene silencing in bacterial chromatin. *Nat. Commun.* **2024**, *15*, 2787. https://doi.org/10.1038/s41467-024-47114-w.
54. Liu, F.; Xu, J.; Hu, X.; Duan, B.; Xia, B. Quantitative real-time *in vitro* transcription assay (QRIVTA) for transcriptional regulation studies. *Protein Cell* **2025**, *16*, 484–490. https://doi.org/10.1093/procel/pwae054.
55. Wang, A.; Cordova, M.; Navarre, W.W. Evolutionary and functional divergence of Sfx, a plasmid-encoded H-NS homolog, underlies the regulation of IncX plasmid conjugation. *mBio* **2025**, *16*, e02089-24. https://doi.org/10.1128/mbio.02089-24.
56. Gordon, B.R.G.; Li, Y.; Wang, L.; Sintsova, A.; van Bakel, H.; Tian, S.; Navarre, W.W.; Xia, B.; Liu, J. Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5154–5159. https://doi.org/10.1073/pnas.0913551107. Correction in *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18741.