



OPEN

Structure-based inhibitors targeting the alpha-helical domain of the *Spiroplasma melliferum* histone-like HU protein

Yuliya K. Agapova^{1,7}✉, Dmitry A. Altukhov^{1,7}, Vladimir I. Timofeev², Victor S. Stroylov³, Vitaly S. Mityanov^{3,4}, Dmitry A. Korzhenevskiy¹, Anna V. Vlaskina¹, Eugenia V. Smirnova⁵, Eduard V. Bocharov^{5,6}✉ & Tatiana V. Rakitina⁵✉

Here we report bisphenol derivatives of fluorene (BDFs) as a new type of chemical probes targeting a histone-like HU protein, a global regulator of bacterial nucleoids, via its dimerization interface perturbation. BDFs were identified by virtual screening and molecular docking that targeted the core of DNA-binding β -saddle-like domain of the HU protein from *Spiroplasma melliferum*. However, NMR spectroscopy, complemented with molecular dynamics and site-directed mutagenesis, indicated that the actual site of the inhibitors' intervention consists of residues from the α -helical domain of one monomer and the side portion of the DNA-binding domain of another monomer. BDFs inhibited DNA-binding properties of HU proteins from mycoplasmas *S. melliferum*, *Mycoplasma gallicepticum* and *Escherichia coli* with half-maximum inhibitory concentrations in the range between 5 and 10 μ M. In addition, BDFs demonstrated antimicrobial activity against mycoplasma species, but not against *E. coli*, which is consistent with the compensatory role of other nucleoid-associated proteins in the higher bacteria. Further evaluation of antimicrobial effects of BDFs against various bacteria and viruses will reveal their pharmacological potential, and the allosteric inhibition mode reported here, which avoids direct competition for the binding site with DNA, should be considered in the development of small molecule inhibitors of nucleoid-associated proteins as well as other types of DNA-binding multimeric proteins.

In bacteria, proper assembly of the higher-order nucleoid structure requires the accessory of nucleoid associated proteins (NAPs) including leucine-responsive regulatory protein (LRP), factor for inversion stimulation (FIS), histone-like nucleoid structuring (H-NS), integration host factor (IHF) and histone-like HU protein (HU)¹. The most abundant protein among these global regulators of DNA topology² was named HU after "histone" and *Escherichia coli* strain U93 used in 1970s for bacterial nucleoid isolation³. HU is omnipresent in bacteria and was also found in plastid-bearing *Eukaryota*, *Euryarchaeota*, *Thaumarchaeota* and some viruses⁴. It is a small basic dimeric protein with a highly conserved spatial structure⁵, which consists of a compact α -helical domain and a DNA-binding domain containing a saddle-shaped β -sheet and protruding β -ribbon DNA-spanning arms, highly mobile in the absence of DNA^{6–9}. In contrast to the majority of *E. coli* DNA-binding proteins recognizing specific DNA sequences (e.g. Lac repressor¹⁰) and to its closest homolog IHF¹¹, HU proteins bind DNA in a

¹National Research Center "Kurchatov Institute", Akad. Kurchatova pl., 1, Moscow 123182, Russian Federation. ²Federal Scientific Research Center "Crystallography and Photonics", Russian Academy of Sciences, Leninskii Pr., 59, Moscow 119333, Russian Federation. ³N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninskii pr., 47, Moscow 119991, Russian Federation. ⁴D. Mendeleev, University of Chemical Technology of Russia, Miusskaya sq., 9, Moscow 125047, Russian Federation. ⁵Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow 117997, Russian Federation. ⁶Moscow Institute of Physics and Technology (State University), Institutskiy Per., 9, Dolgoprudny, Moscow Region 141701, Russian Federation. ⁷These authors contributed equally: Yuliya K. Agapova and Dmitry A. Altukhov. ✉email: agapova.jk@gmail.com; bon@nmr.ru; taniarakitina@yahoo.com

non-sequence-specific manner with an enhanced affinity to distorted DNA structures such as forks, junctions, nicks and overhangs^{12–16}.

As key NAPs, HU proteins are involved in DNA compaction and supercoiling as well as in the formation of higher order protein complexes, which regulate vital DNA functions including replication, recombination, repair, and transcription with consequent effect on the bacterial transcriptome and proteome^{17–21}. HU proteins have been shown to participate in the transcriptional regulation of the genes associated with virulence of *Salmonella enterica* serovar *Typhimurium*²², *Francisella tularensis*²³, and *Porphyromonas gingivalis*²⁴. HU proteins play a role in biofilm formation and overall pathogenesis of *Xanthomonas citri*²⁵ and *Cytophaga hutchinsonii*²⁶. In *E. coli*, HU proteins control expression of about 8% of genes mostly associated with adaptation and stress response²⁰. Deletion of both genes coding the HU heterodimer from the *E. coli* genome is not lethal if other NAPs are preserved²⁷. In some other bacteria, e.g. *Bacillus subtilis*²⁸, *Streptococcus pneumoniae*²⁹ and *Mycobacterium tuberculosis*³⁰, HU proteins are more essential for cell viability. In general, the importance of HU proteins for bacterial survival depends on the presence of other NAPs, which are able to compensate for the absence of HU. Thus, HU proteins are crucial for the bacteria where they are the only NAPs available, such as *Mycoplasma* genus (class Mollicutes), the smallest known microorganisms characterized by absence of the cell wall and reduced genome size^{31,32}. Mollicutes are parasites of higher Eukaryotes, frequently acting as causative agents of common infections and main contaminants of cell cultures and vaccines. Controlling mycoplasma infections is a serious challenge^{33,34}. The high importance of HU proteins for *Mycoplasma* survival makes them an attractive target for the development of antimycoplasmic agents.

Earlier, inhibition of DNA-binding properties of *M. tuberculosis* HU protein (MtbHU) was achieved using trans-stilbene derivatives, which also affected the growth of the corresponding bacteria, suggesting that MtbHU (as well as any other HU) might be considered as a target for pharmacological intervention³⁵. Recently, the same stilbene derivatives were shown to disrupt DNA-binding properties of a viral HU/IHF-homolog (pA104R) from African swine fever virus and inhibit the replication of the virus in primary porcine alveolar macrophages³⁶. We followed the design strategy proposed in³⁵ for the development of chemical inhibitors of HU protein from mycoplasma *Spiroplasma melliferum* (HUSpm). High-resolution spatial structure of HUSpm (PDB ID 5L8Z⁹) was used for virtual screening and docking, while its structure in solution (PDB ID 5OGU³⁷) was employed for the comparative analysis of conformational dynamics observed in the HUSpm dimer upon either independent or sequential addition of short double-stranded DNA and an inhibitor. We found out that bisphenol derivatives of fluorene (BDFs) inhibit DNA binding properties of three model HU proteins (HUSpm, HU α -homodimer from *E. coli* (HUEc) and HU from another mycoplasma species *Mycoplasma gallisepticum* (HUMgal)). BDFs also demonstrated antimicrobial activity against mycoplasma, but not against *E. coli*, which is consistent with a compensatory role of other NAPs in the higher bacteria. Surprisingly, high-resolution heteronuclear NMR spectroscopy complemented with molecular dynamics and site-directed mutagenesis revealed that BDFs do not target the site that had been described for the trans-stilbene derivatives in³⁵. Instead of the DNA-binding cleft, BDFs target the side portion of HUSpm that overlaps the dimer interface between the α -helical domain of one monomer and the β -saddle domain of another monomer causing a perturbation of the dimer interface which interferes allosterically with the natural conformational dynamics of the DNA-binding domain and the corresponding efficiency of DNA binding. We believe that this inhibition mode without direct competition for the binding site with natural HU proteins' ligand (DNA) is of high scientific and pharmacological interest. It is likely that other new HU protein inhibitors with similar properties will be discovered soon, since targeting the HU protein α -helical body may be less challenging than targeting a large and adaptive DNA-binding area.

Results

Identification of HUSpm–DNA interaction inhibitors among bisphenol derivatives of fluorene. In silico screening by molecular docking and filtering was performed to identify compounds interacting with the core region of the DNA-binding saddle-like domain of HU proteins (Fig. 1A) previously characterized in³⁵. High-resolution spatial structure of HUSpm (PDB ID 5L8Z⁹) was chosen as a target for virtual screening of one million compounds from the Vitas-M laboratory library collection (<https://vitasmlab.biz/>). The screening (see "Methods" for a description) identified a set of five compounds (Supplementary Table S1), further evaluation of which was carried out by an electromobility shift assay (EMSA) performed as described in⁵. In the assay, formation of DNA–protein complexes retards migration of dye-labeled DNA–duplexes during non-denaturing PAGE that gives a shifted fluorescent band. The presence of an inhibitor prevents complex formation and corresponding band shifting.

The first round of EMSA was performed using 100 μ M concentration of the compounds and a 48 bp oligonucleotide duplex (DS48) and demonstrated that a bisphenol derivative of fluorene (BDF), referred here as BDF1 (compound 5 in Supplementary Table S1), inhibits HUSpm–DNA binding (Supplementary Figure S1). Positioning of BDF1 in the core region of the HUSpm β -saddle and its polar contacts with amino acids of the DNA-binding domain are shown in Fig. 1B. The BDF1 inhibition curves (Fig. 1C) obtained in EMSA performed with two different oligonucleotide duplexes (DS48 and a 24 bp duplex carrying a single nucleotide insertion in one strand (DS24a)) showed very similar half-maximum inhibitory concentration (IC₅₀) values for HUSpm–DNA binding— $6.5 \pm 0.9 \mu$ M and $6.2 \pm 0.5 \mu$ M, respectively (Supplementary Table S1). The shorter duplex (DS24a) was routinely used in further experiments because of a sharper band of bound DNA at the physiological (150 μ M) salt concentration and better separation of bound and unbound DNA during PAGE, which facilitated quantification of EMSA results.

After successful evaluation of BDF1 as a HUSpm–DNA binding inhibitor, we obtained an additional set of six other BDFs found in the screening collection of Chemical Block Ltd (Moscow, Russia, <https://www.chemical-block.com>) (Table 1) and analyzed their effects on the HUSpm–DNA binding. Two compounds, BDF2 and BDF3,

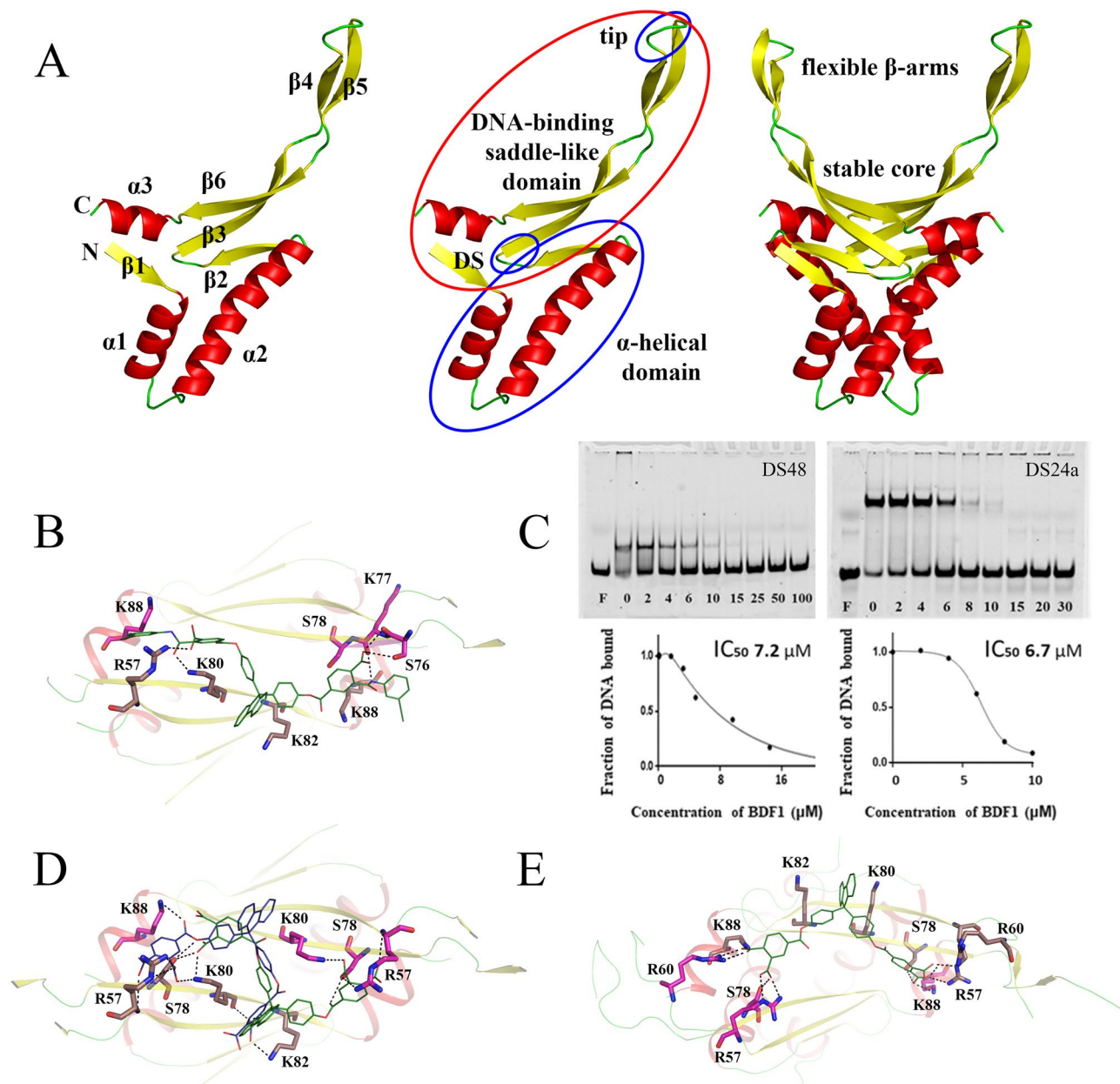


Figure 1. Description of HUSpm domain structure and discovery of HUSpm inhibitors. **(A)** Left to right: secondary structure elements, domains and their subdivisions are shown on cartoon models of HUSpm monomer and dimer. DS—dimerization signal (GFGKF), tip—DNA-intercalating loop (GRNP). **(B)** Model of BDF1 (compound 5 from Supplementary Table S1) binding to the core of the HUSpm DNA-binding domain and polar interactions between BDF1 and positively charged or polar residues of HUSpm. Residues from different monomers are in magenta and beige colour, respectively. **(C)** Inhibition of HUSpm-DNA binding by different concentrations of BDF1, corresponding inhibition curves and IC_{50} values. Two different oligonucleotide duplexes—48 bp DS48 and 24 bp DS24a carrying a single nucleotide insertion in one strand—were used in the assay with similar results. **(D)** Model of BDF4 (Supplementary Table S3) binding to the core of the HUSpm DNA-binding domain in two locations and polar interactions between BDF4 and positively charged or polar residues of HUSpm. Residues from different monomers are in magenta and beige colour, respectively. **(E)** BDF4 position and polar interactions after 1,000 ns MD simulation.

were identified as HUSpm inhibitors by EMSA performed with DS24a, although their IC_{50} values ($29 \pm 5 \mu\text{M}$ and $36 \pm 6 \mu\text{M}$, respectively) were significantly lower than that of BDF1 (Supplementary Table S1 and Figure S2). Comparison of BDF2 and BDF3 with BDF1 suggested that the carboxyl groups of BDF1 play an important role in the binding of the inhibitor to the HUSpm β -saddle core enriched in positively-charged amino acids.

Therefore, we decided to obtain three additional BDFs (BDF4, 5 and 6, Supplementary Table S3) using chemical synthesis. The synthesis scheme and the detailed protocols are provided in Supplementary Figure S3 and Supplementary methods, respectively. Given a high degree of similarity, BDF5 was used to replace BDF1, since

| Compounds | BDF4 | BDF5 | BDF6 |
|--|--|------------|------------|
| HU protein | IC ₅₀ , μM (average ± SD from three runs of EMSA) | | |
| HUSpm | 5.6 ± 0.8 | 6.3 ± 0.7 | 8.2 ± 1.1 |
| HUMgal | 5.0 ± 0.6 | 5.9 ± 0.8 | 8.1 ± 1.0 |
| HUEc | 7.5 ± 1.0 | 7.5 ± 0.9 | 10 ± 1.3 |
| Activities of BDF4 against HUSpm mutants, IC₅₀, μM (mean ± SD from three runs of EMSA) | | | |
| HUSpm_R57A | HUSpm_K80A | HUSpm_R88A | HUSpm_K35T |
| 6.3 ± 1.0 | 5.9 ± 0.9 | 6.3 ± 1.2 | 1.8 ± 0.3 |

Table 1. Inhibition of HU protein DNA-binding by synthesized BDFs.

the latter had been received from the Vitas-M laboratory collection in an insufficient quantity. BDF4 and BDF6 had lower molecular weights (their MWs were 734 and 646 Da vs. MWs 932 and 912 Da for BDF1 and BDF5, respectively) and contained four and two carboxyl groups, respectively. Docking of BDF4 in the core region of the DNA-binding saddle-like domain of HUSpm showed that twofold degenerate center-symmetrical binding poses (Fig. 1D) are possible ($\Delta G \sim -9$ kcal/mol), and the binding in both positions is similar to that observed for BDF1 (Fig. 1B). However, subsequent MD simulation revealed a more advantageous binding mode (shown in Fig. 1E). In this position, the fluorene was placed between side chains of K82 and K80 of one monomer thus forming pi-cation interactions, while carboxyl groups electrostatically interacted with side chains of the DNA-binding domains' basic and polar residues R57, R60, S78, and K88 from both monomers.

Bisphenol derivatives of fluorene inhibit DNA-binding properties of HU proteins from different bacteria as well as mycoplasma cell growth. For further evaluation of the three newly synthesized BDFs (BDF4, 4,4'-[4,4'-(9H-fluorene-9,9-diyl)-bis(4.1-phenylene)]-bis(oxy)-bis(oxomethylene)diphthalic acid; BDF5, 4,5'-[4,4'-(9H-fluorene-9,9-diyl)-bis(4.1-phenylene)]-bis(oxy)-bis(oxomethylene)-bis[2-(m-tolylcarbonyl)-benzoic acid]; BDF6, 3,3'-[4,4'-(9H-fluorene-9,9-diyl)-bis(4.1-phenylene)]-bis(oxy)-bis(oxomethylene)-dibenzoic acid) we employed two other HU proteins in addition to HUSpm. One protein was from mycoplasma *M. gallisepticum* (HUMgal) and another was HU α -homodimer from *E. coli* (HUEc). As we have reported earlier, the three chosen HU proteins belong to two distant HU clades and have differential specificity patterns towards various DNA structures⁵. Comparison of their amino acid sequences revealed 25% sequence identity for the HUSpm and HUMgal pair, 42% for the HUSpm and HUEc pair, and 23% for the HUMgal and HUEc pair. In addition, HUMgal has a six amino acid N-terminal extension and HUSpm has a two amino acid insertion between α -helices 1 and 2 (Supplementary Figure S4). To assess the effect of the three BDFs on DNA binding properties of the three HU proteins, increasing concentrations of inhibitors were mixed with preset DNA-protein complexes and EMSA was performed and quantified as described in the **Methods**. DS24a was used in all EMSA experiments.

As Fig. 2A shows, all compounds exhibited inhibitory effects against all target proteins with IC₅₀ values for DNA binding at micromolar (< 10 μM) concentration ranges (Table 1). In all cases, the observed order of compounds by effectivity of inhibition was consistent: BDF4 > BDF5 > BDF6. The difference between BDF4 and BDF5 (and its prototype BDF1) was not significant, while BDF6 showed slightly higher IC₅₀ values.

Since HU proteins are known to be essential for mycoplasma cell viability^{31, 32}, we assessed the effects of the inhibitors on *M. gallisepticum* and *S. melliferum* growth in culture. The minimal inhibitory concentrations (MIC) of BDF4, BDF5 and BDF6 for corresponding mycoplasmas were determined in the broth micro-dilution test³⁸. The growth of metabolizing organisms causes a change in the pH of the medium leading to a visible color change due to the presence of phenol red in the medium. When the growth of bacteria is inhibited, color change is not observed. The MIC is the lowest inhibitor concentration that prevents the color change at the time when the color change has developed in control (untreated) samples. The MICs were found to be in the range from 40 to 70 μM (Fig. 2B). BDF6 had the worst MIC as compared to either BDF4 or BDF5, whose MICs were not significantly different from each other (Fig. 2C).

No inhibitory effect on *E. coli* culture growth was observed (data not shown) which coincided with the report that *E. coli* cells with double knockout of HU heterodimer coding genes were viable and grew well in standard cultivation conditions²⁰. In addition, several other factors (e.g. intra-cellular permeability and stability) could contribute to the resistance of highly organized bacteria to HU inhibitors.

NMR spectroscopy complemented with molecular dynamics suggests that bisphenol derivatives of fluorene target the HUSpm alpha-helical domain. Earlier, we characterized the structural and motional features of HUSpm using nuclear magnetic resonance (NMR) spectroscopy; the backbone and side chain resonance assignments and the secondary structural features derived empirically from backbone chemical shifts were reported in³⁷. Since NMR can also be used to study the structure-activity relationships corresponding to specific inhibitors binding to target proteins, here we performed the backbone ¹⁵N-NMR relaxations experiments to characterize the motional features of HUSpm upon both simultaneous and separate addition of the inhibitor and DNA.

Among the three BDFs described in the previous charter, BDF4 was chosen for NMR study due to its high-est activity in all tests and water solubility. Both BDF5 and BDF6 had limited solubility; BDF5 also had higher

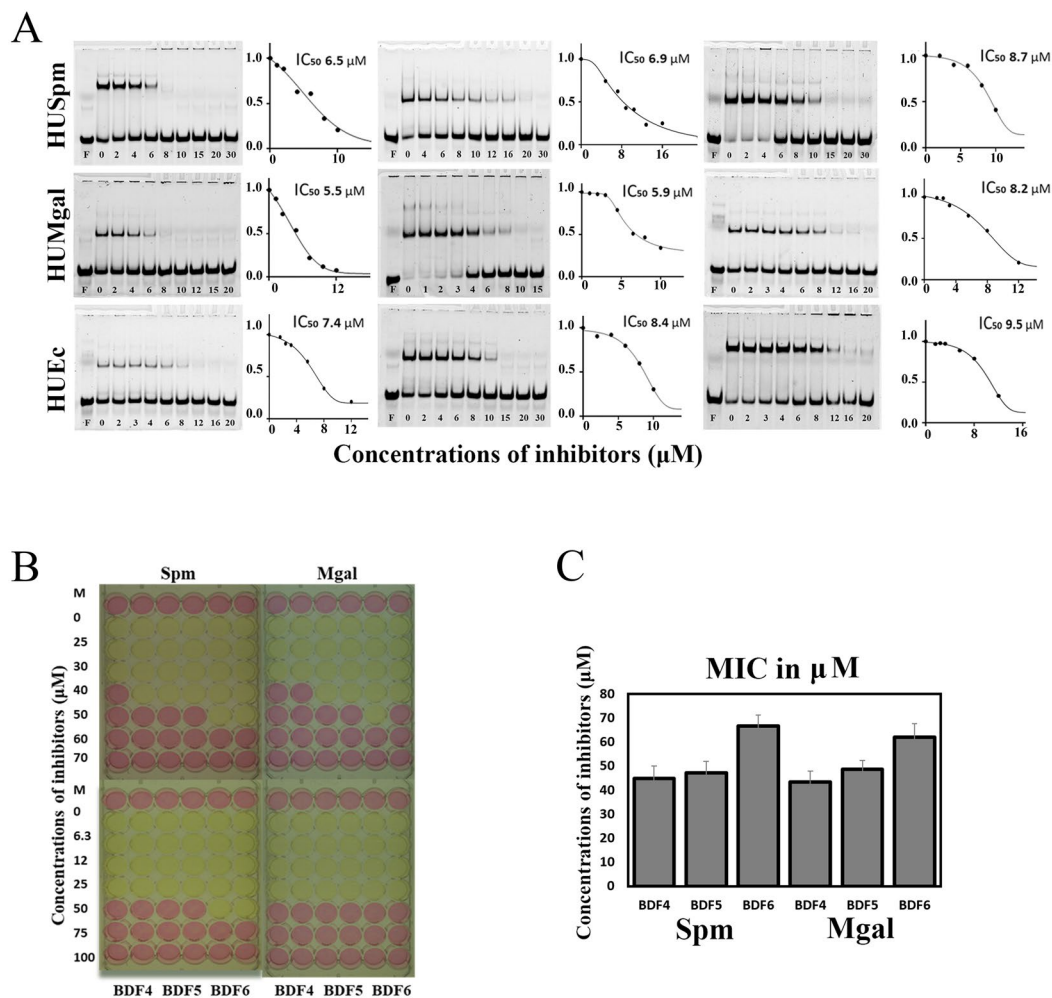


Figure 2. Effect of BDFs on DNA-binding of HU proteins from different bacteria and on mycoplasma cell growth. **(A)** Inhibition of HUSpm, HUMgal and HUEc DNA-binding determined in EMSA performed with DS24a duplex, corresponding inhibition curves and IC₅₀ values. **(B)** Effect of BDF4, BDF5 and BDF6 on *M. gallisepticum* and *S. melliferum* growth in culture. The yellow colour indicates the growth of bacteria, pink indicates no growth. MIC was defined as the minimum inhibitory concentration that prevented pink to yellow colour change. **(C)** Graphical representation of the combined results of three independent MIC determinations.

molecular weight, and BDF6 was less active. To confirm BDF4 binding to the β -saddle domain core of HUSpm, we performed titration of the ¹⁵N-labelled protein (both in its free form and in complex with the DS14 duplex (Supplementary Table S4)) with different amounts of the inhibitor. ¹⁵N-HUSpm preparation and conditions of NMR spectroscopy were established in the previous work³⁷. The interactions of HUSpm with the short canonical double-stranded DNA (DS14) were also subjected to the NMR study.

In all complexes (HUSpm + DS14, HUSpm + BDF4 and (HUSpm + DS14) + BDF4), significant and moderate amide shift perturbations were observed for some amide cross peaks of HUSpm (see Supplementary Figure S5). For all HUSpm residues with assigned backbone amide signals, quantifications of weighted chemical shift perturbations (CSPs) and normalized cross-peak intensity perturbations (CIPs) of the backbone amide groups are visualized by the graphical view (Figs. 3A,B, 4A–D) and the color-coded mapping on the HUSpm dimer ribbon structure (Figs. 3C,D, and 4E–H).

Increased CSP values are usually associated with local structural perturbations and changes in the local environment of the amide group of interest, whereas positive and negative CIP values, which result from the relative signal narrowing and broadening, respectively, can indicate changes in intramolecular local motions over a wide range of time scales from fast pico- to nanosecond movements to slow conformational transitions in a micro-millisecond time scale.

According to Fig. 3A,C, in case of HUSpm-DNA complex formation, the residues showing the most significant chemical shift changes belong to the C-terminal part of the protein and reside in the DNA-binding β -saddle-like domain including the α 3-helix (see Fig. 1A). This effect coincides with a model of HUSpm interaction with DS14 DNA taken from 1,000 ns MD simulations of the corresponding complex obtained by comparative

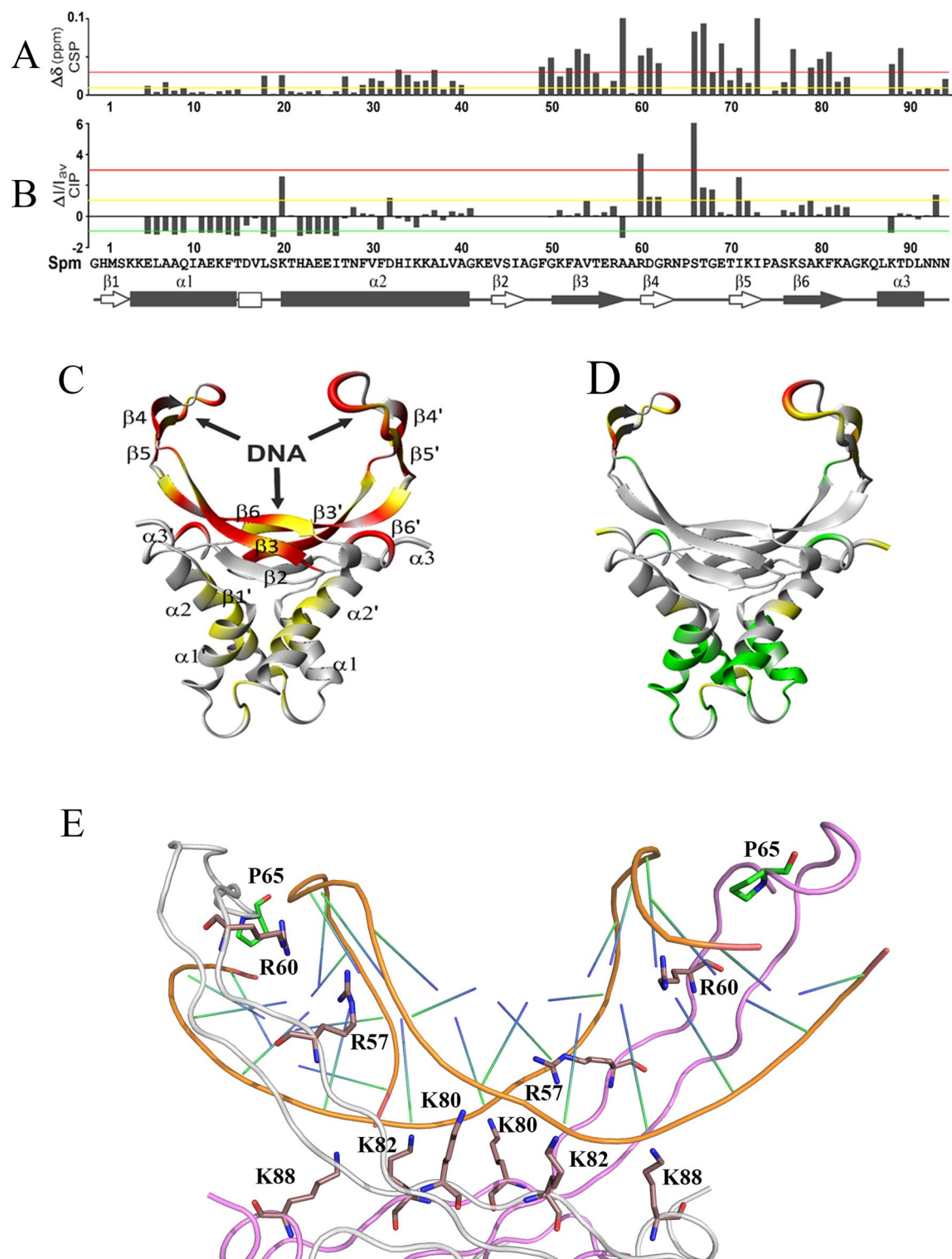


Figure 3. Monitoring of chemical shift perturbations (CSPs) and cross-peak intensity perturbations (CIPs) of backbone amide groups of the HUSpm dimer upon DNA-binding. Weighted CSPs (A,C) and normalized CIPs (B,D) of HUSpm residues after binding DS14 shown as diagrams (A,B) and their color-coded mapping on ribbon models of the HUSpm dimer (C,D). Green, yellow and red lines in (A) and (B) are cutoff levels for the color-coding used in (C) and (D). (E) Model of HUSpm interacting with a DS14 duplex after 1,000 ns MD simulation. Two monomers are shown with gray and magenta lines. Basic residues from both monomers interacting with DNA through their side chains are shown in beige sticks; Pro65 intercalating into the DNA minor groove is in green.

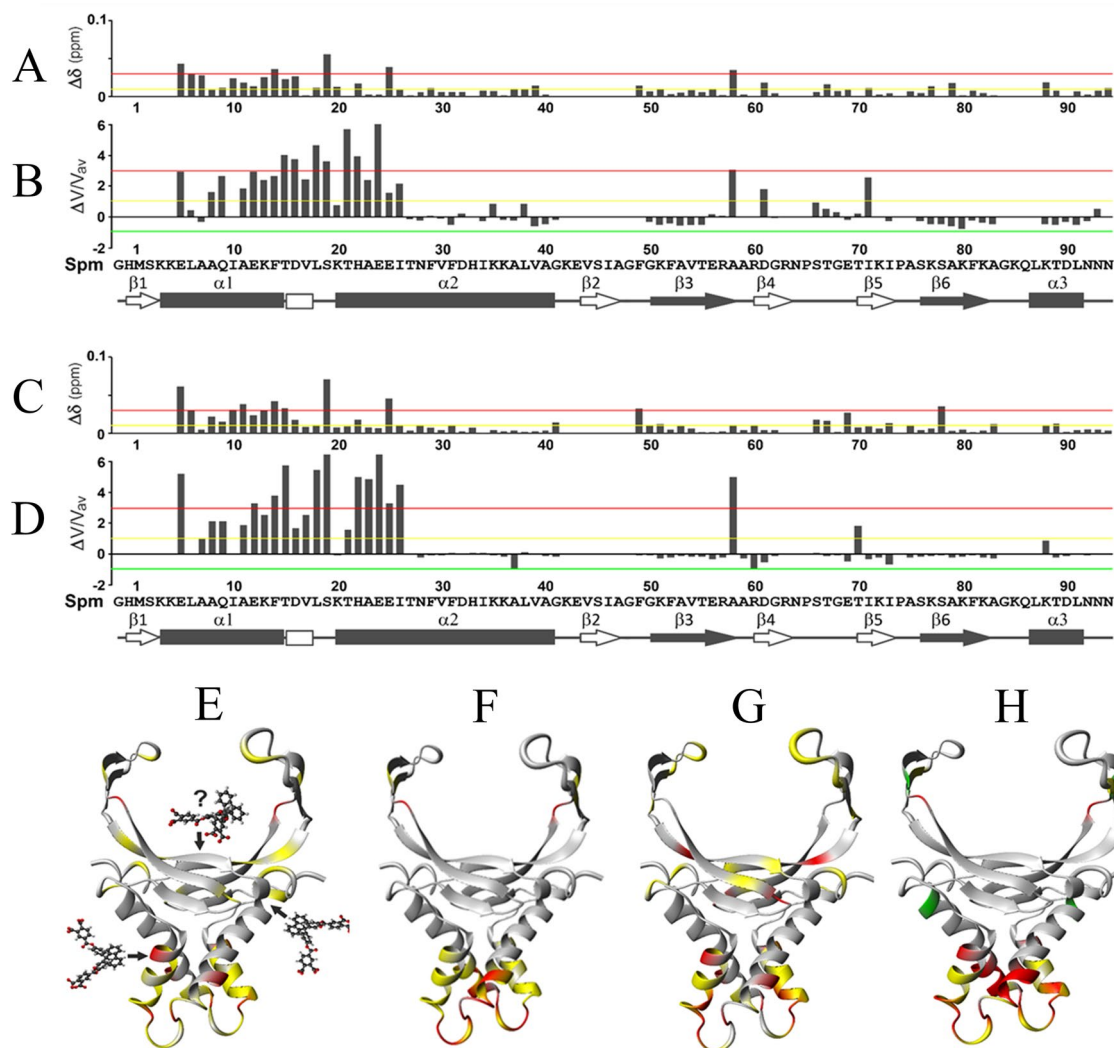


Figure 4. Monitoring of chemical shift perturbations (CSPs) and cross-peak intensity perturbations (CIPs) of backbone amide groups of the HUSpm dimer in its free form (A,B,E,F) and in complex with DS14 (C,D,G,H) upon BDF4 binding. Weighted CSPs (A,C,E,G) and normalized CIPs (B,D,F,H) of the HUSpm residues after binding BDF4 shown as diagrams (A–D) and their color-coded mapping on ribbon models of the HUSpm dimer (E–H). Green, yellow and red lines in (A–D) are cutoff levels for the color-coding used in (E–H).

modelling (Fig. 3E). A similar effect has been reported in the NMR study of the interaction of HU protein from *Helicobacter pylori* with 25 bp hairpin DNA, which was very similar in length to our 14 bp DNA-duplex³⁹.

In addition, as expected upon DNA-binding, significantly increased CIPs (> 1 and > 3 , yellow and red levels, respectively) due to NMR signal narrowing were observed for amide groups of residues belonging and adjacent to the $\beta 4$ - and $\beta 5$ -strands surrounding DNA-intercalating tips (Fig. 3B,D). This observation illustrates the transition of flexible DNA-binding arms from a free state characterized by conformational exchange and high-amplitude movement³⁷ to a DNA-binding mode characterized by rapid low-amplitude movement that facilitates the capture of DNA and intercalation of tips in the DNA minor groove. Thus, in this case, the narrowing of the signal is primarily caused by a decrease in conformational diversity with slow transitions in the micro- and millisecond range due to the formation of a DNA–protein complex.

Notably, moderate negative CIPs (< -1 , green level) due to NMR signal broadening were identified in the opposite, N-terminal helical part of the HUSpm dimer probably owing to the appearance of slow conformational transition (on the micro-millisecond time scale undetectable by the 1,000 ns MD simulation). Thus, both CSP and CIP distribution patterns indicate that an allosteric (in scissor-like manner) rearrangement of the protein structure and redistribution of its backbone motions occur in HUSpm dimer upon DNA binding.

The addition of BDF4 to HUSpm caused opposite effects compared to those observed upon DNA binding (Fig. 4). When the stoichiometry of inhibitor to HUSpm dimer (either in free form or in complex with DNA) was about 1:1, significant CSPs (between 0.01 and 0.03, yellow level) were observed mainly for residues from the N-terminal α -helical part of the protein located in the $\alpha 1$ -helix and in the N-terminal part of the $\alpha 2$ -helix, whereas CSPs were not significant for the majority of residues from the C-terminal DNA-binding domain (Fig. 4A,C,E,G). In both cases, the most significant CSPs (> 0.03 , red level) were observed for several residues

from the α 1- and α 2-helices, and for one and two residues of the DNA-binding domain in case of HUSpm + BDF4 and (HUSpm + DS14) + BDF4 complexes, respectively (Fig. 4A,C,E,G). These three residues are F49 from DS region of HUSpm (see Fig. 1A), A58 from the C-terminal end of the β 3-strand and S78 from the β 6-strand. The two former residues belong to the dimer interface and the latter to the stable core of DNA-binding domain. In this regard, we should note that residues with affected chemical shifts might be involved in the interaction with BDF4 either directly or indirectly by participating in the local or global conformational rearrangements in response to inhibitor binding.

Moreover, NMR signal intensity analysis revealed that more pronounced changes of CIPs (> 1 and > 3 , *yellow* and *red* levels, respectively) were also observed for residues from the α 1 and α 2-helices upon BDF4 interaction with both free HUSpm and a HUSpm + DS14 complex (Fig. 4B,D,E,H). These significant increases of CIPs indicated that residues from the α 1-helix and N-terminal part of the α 2-helix underwent fast pico- to nanosecond time scale motions resulting in local enhancement of backbone flexibility in the HUSpm α -helical domain.

Thus, both CSPs and CIPs distribution patterns revealed that interaction with BDF4 caused an opposite effect on structural and motional features of HUSpm compared to that occurring upon DS14 binding, which indicates low probability of BDF4 interactions with the core of the DNA-binding domain.

Since NMR-derived data contradicted with docking prediction about BDF4 interaction with the core of the DNA-binding domain of HUSpm, a new docking experiment was performed, in which the whole surface area of the HUSpm dimer was the target for BDF4 inhibitor; then the modelled complex was studied by Molecular Dynamics (Fig. 5). The available NMR data indicated the localization of the main conformational-dynamic changes in HUSpm, caused by the binding of the inhibitor, in the α -helical domain of the protein, but did not provide more detailed information about the specific localization site of the inhibitor at the recognition stage and in the final bound state. Therefore, docking and subsequent MD simulations were performed without using any NMR constraints.

As a result, new binding sites located on the opposite sides of the HUSpm dimer were found. Figure 5A illustrates the conformational dynamics of the HUSpm dimer with two BDF4 molecules (not shown in the Figure) bound to the new sites in the course of 1,000 ns MD simulation. Figure 5B–D shows the binding modes and amino acid residues interacting with the inhibitor at the beginning and end of the simulation. According to Fig. 5C,D, the new binding site consists of residues from α 1- and α 2-helices of one monomer and two loops adjacent to the core of the DNA-binding β -saddle (β 2-loop- β 3 and β 6-loop- α 3) of another monomer. Both the hydrophobic core and carboxyl groups of BDF4 were driving forces of the inhibitor binding, which coincides with the fact that other BDFs (Supplementary Table S2) are mainly inactive.

MD trajectories obtained for two models of HUSpm + BDF4 complex (which were dubbed top-site and side-site for the old and new model, respectively) were compared to those obtained for HUSpm and its complex with a DS14 DNA duplex. Figure 5E shows root mean square fluctuations (RMSF) values of C α -atoms for all HUSpm residues derived from calculated MD trajectories. Increased RMSF values indicate enhanced flexibility of the protein backbone. The enhanced flexibility of residues located at the N-terminal part of the protein on the α 1-helix and in the beginning of the α 2-helix was clearly visible at a nano-microsecond MD timescale of the side-site model (black line), and was in good agreement with results of NMR signal intensity analysis (Fig. 4B,F). Moreover, MD simulations of the side-site model demonstrated that such a mode of BDF4 binding causes a significant increase of the flexibility of DNA-binding arms and moderately increases the flexibility of the stable part of DNA-binding domain connecting the flexible arms, as compared to the free HUSpm dimer (Fig. 5A,E). These effects are opposite to those associated with DNA-binding (reduced RMSF values for residues from the DNA-binding domain), which reflects a decreased amplitude of motion of the DNA-binding arms.

Alanine substitution of the basic amino acid residues playing key roles in DNA binding did not affect HUSpm sensitivity to bisphenol derivatives of fluorene.

To confirm or definitively disprove the probability of BDF4 binding to the HUSpm DNA-binding surface (top-site model), mutagenesis of amino acid residues presumably playing key roles in this binding mode was employed. The residues R57, K80 and K88 from the DNA-binding β -cleft were chosen for alanine substitution because they form polar contacts with BDF4 carboxyl groups according to the top-site model (Fig. 1D,E). According to³⁵, alanine substitution of R55 (R57 in HUSpm) and K86 (K88 in HUSpm) led to reduced binding of both DNA and stilbene derivative inhibitors to HUMtb. Our previous analysis of conserved residues in the HU protein characteristic core³ showed that the residues R57 and K88, but not K80, were conserved among HU proteins (Supplementary Figure S4).

EMSA demonstrated that none of the three mutations in the β -saddle core significantly affected HUSpm sensitivity to BDF4 inhibitor (Fig. 6A, Table 1). In contrast, each of the mutant proteins bound various DNA structures with reduced affinities as compared to wild-type HUSpm (Fig. 6B,C and Table 2). Different DNA duplexes (Supplementary Table S4) including both canonical double stranded DNA (48) and distorted DNA structures of various length, carrying either nick (N) or an insertion of one and seven adenines in one DNA strand (A1, 24a and A7, respectively) were used in EMSA. To determine the affinities of each HUSpm mutant to different DNAs, the dissociation constants (k_d) of observed complexes were calculated as described in^{13,16}. Figure 6C shows that HUSpm_R57A has the weakest DNA binding (the increase of k_d values was between 51 and 88 folds for complexes with different DNAs over those of HUSpm), K88A and K80A mutations cause lesser effects (the increase of k_d values was between 25 and 42 folds and between 7 and 16 folds, respectively).

Since alanine substitutions of basic residues that reside in the HUSpm DNA-binding cleft and participate in DNA binding did not interfere with HUSpm interactions with BDF4, we analyzed effects of amino acid perturbations in the HUSpm dimer interface. Amino acid residues (F14, F29, F31 and K35) of HUSpm α -helical body contributing to the HUSpm dimer stabilization and maintaining its high thermal stability were identified earlier using structural analysis and site directed mutagenesis⁹. Phenylalanines are located in aromatic clusters

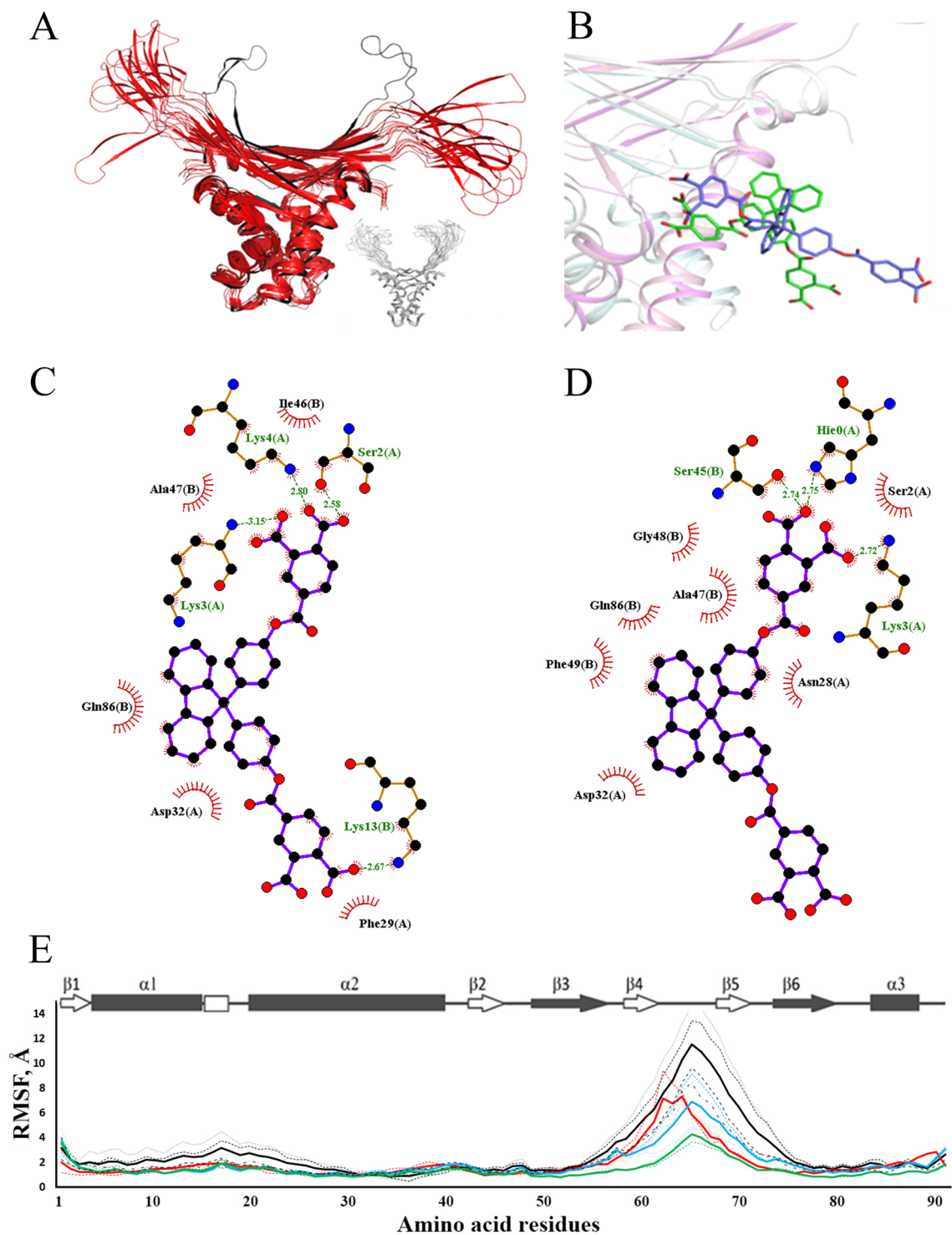


Figure 5. Side-site model of BDF4 binding to HUSpm and MD analysis of the corresponding complex. **(A)** Superposition of MD models (colored red) taken from the MD trajectory with an interval of 100 ns and an energy-minimized model of HUSpm (colored black) in complex with BDF4 bound to the side-site. Only the protein is shown to illustrate the increased movement of the DNA-binding domain. A similar picture obtained for free HUSpm (colored gray) has been inserted for comparison. **(B)** Position of BDF4 at the beginning (green colored inhibitor) and end (blue colored inhibitor) of MD simulation. The protein is shown as a cartoon model, one monomer is colored in magenta and another in gray. Only one BDF4 molecule is shown for clarity. **(C,D)** Polar and hydrophobic interactions of BDF4 molecule with HUSpm amino acid residues at the beginning **(C)** and end **(D)** of MD simulation. The diagrams were obtained with LigPlot software. **(E)** RMSF values of the Ca-atoms (in Å) derived from 1,000 ns MD trajectories obtained for a free HUSpm dimer (blue lines), HUSpm interacting with DS14 (green lines), and HUSpm interacting with BDF4, where the inhibitor targeted either the core of the DNA-binding domain (red lines) or overlaid the dimer interface between the α -helical domain of one monomer and the side portion of the DNA-binding domain of another monomer (black lines). Discontinuous and solid lines show the RMSF values obtained from independent MD experiments and the corresponding mean values, respectively. Secondary structure distribution derived from NMR data³⁷ is shown above.

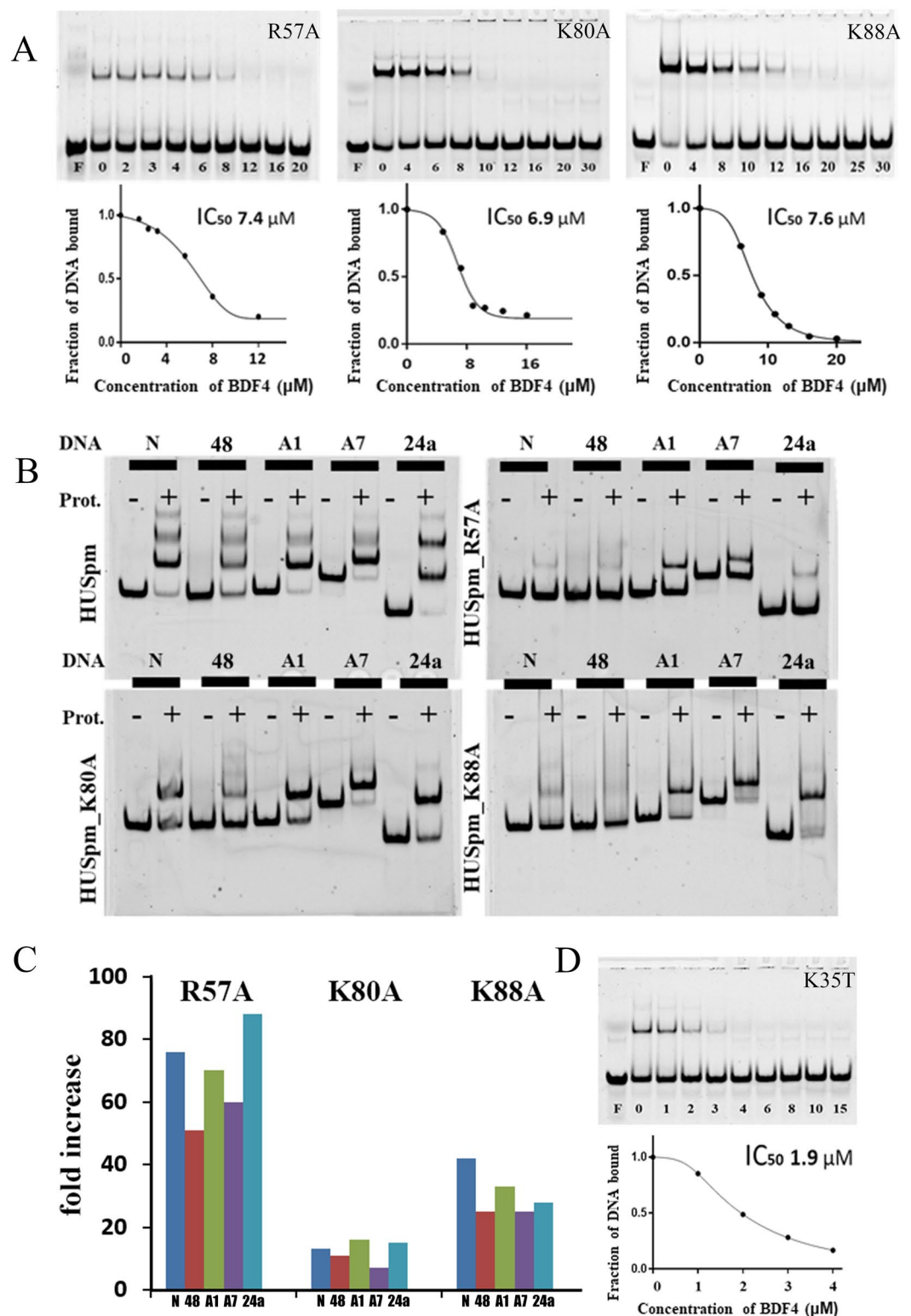


Figure 6. Disruptions in the DNA binding site did not influence the HUSpm sensitivity to BDF4. **(A)** Alanine substitutions of the positively charged amino acid residues located in the HUSpm β -saddle core does not affect HUSpm sensitivity to BDF4 in EMSA performed with the DS24a duplex; corresponding inhibition curves and IC_{50} values are shown. The mutant proteins were used in enhanced concentrations because of their low affinities to DNA: 100 nM for HUSpm_K80A and HUSpm_K88A, 200 nM for HUSpm_R57A. **(B)** Effects of the same mutations on the DNA binding properties of HUSpm. HUSpm (100 nM) and its mutant variants in concentrations indicated above were mixed with 5'-labelled DNA in a buffer containing 150 mM NaCl; the bound and free DNA were resolved in a polyacrylamide gel under non-denaturing conditions. DNA duplexes used are denoted at the top of the gel: N, nicked DNA; 48, DS48; A1, DS4a and A7, duplexes of different length, containing one or seven non-paired adenines in one strand. **(C)** Dissociation constants (k_d) of the corresponding complexes normalized to those of HUSpm wild type. **(D)** The K35T substitution increases efficacy of BDF4. Inhibition of the HUSpm_K35T-DNA binding by BDF4 determined in EMSA, corresponding inhibition curve and IC_{50} value.

| DNA | HUSpm wild type | HUSpm_R57A | HUSpm_K80A | HUSpm_K88A |
|---------------|-----------------|------------|------------|------------|
| Nick (N) | 5.2 ± 1.1 | 398 ± 45 | 69 ± 20 | 220 ± 40 |
| DS48 (48) | 18.6 ± 4.8 | 788 ± 100 | 163 ± 27 | 390 ± 55 |
| A1 bulge (A1) | 2.5 ± 0.6 | 175 ± 25 | 41 ± 8 | 83 ± 15 |
| A7 bulge (A7) | 2.0 ± 0.5 | 121 ± 15 | 13 ± 3 | 51 ± 8 |
| DS24a (24a) | 2.6 ± 0.6 | 230 ± 30 | 40 ± 10 | 84 ± 20 |

Table 2. Dissociation constants (K_d in nM, average ± SD, $n = 2$) for HUSpm wild type and HUSpm bearing mutations R57A, K80A and K88A binding to various DNA structures calculated from EMSA experiments (see Fig. 5B legend for a description of the DNA sequences).

strengthening the hydrophobic core of HUSpm dimer interface³⁶ and K35 participates in intermolecular hydrogen bond formation with G48 that connects the $\alpha 2$ -helix of one monomer with the loop between $\beta 2$ and $\beta 3$ strands of another monomer. This $\beta 2$ -loop- $\beta 3$ pattern appears to be one of the most conserved regions of HU proteins and is named the dimerization signal (DS) (see Fig. 1A) for its role in the HU dimer formation and stabilization⁸. Thus, previously obtained HUSpm mutants (F14A, F29A, F31A and K35T) were employed for inhibitor evaluation. The result of EMSA performed with the DS24a duplex and HUSpm_K35T is shown in Fig. 6D, wherein the inhibition curve indicates that K35T substitution enhances sensitivity to BDF4; since about threefold decrease of the IC_{50} value was observed as compared to wild type HUSpm (Table 1). Other mutations did not cause any significant effects (data not shown). It is conceivable that disruption of the intermolecular H-bond connecting DS of one monomer to $\alpha 2$ -helix of another monomer and the corresponding weakening of the dimer interface facilitates inhibitor action.

Discussion

The general idea of this work is based on a previous study of Bhowmick and colleagues who reported successful development of a new chemical probe perturbing the DNA-binding function of HU protein from *M. tuberculosis* (HUMtb)³⁵. The authors have performed very thorough analysis of the DNA-binding domain of HUMtb and suggested that the core or the base of the saddle-like DNA-binding domain (Fig. 1A) is the most “drugable” portion of HU proteins. Thus, following the methodology proposed in³⁵ we conducted virtual screening of one million compounds from the Vitas-M laboratory chemical library collection against a high-resolution spatial structure of HUSpm and identified HU protein inhibitors of a chemical nature drastically different from previously described trans-stilbene derivatives³⁵.

The main objects of this study, previously uncharacterized bisphenol derivatives of fluorene (BDF4, 5 and 6, see Supplementary Table S3), inhibit the DNA-binding properties of three tested HU proteins (Fig. 2A) and, as we strongly believe, should be active against any protein of this class. As we have shown, they also inhibit the growth of the simplest bacteria such as *Mycoplasma* spp. (Fig. 2B), where HU is the only NAP available; at the same time BDFs do not affect the growth of more complex bacteria, such as *E. coli*, which is consistent with the compensatory role of other NAPs in case of HU protein withdrawal. Notably, trans-stilbene derivatives inhibit growth of *M. tuberculosis*, whose genome has a remarkable under-representation of NAPs, but do not cause growth inhibition of *E. coli* and *M. smegmatis*³⁵.

Further analysis was focused on the protein-inhibitor interaction and was carried out using the most potent and soluble inhibitor (BDF4). In addition to the high-resolution crystal structure of HUSpm (PDB ID 5L8Z⁹) that was used as a target for the virtual screening, HUSpm solution structure (PDB ID 5OGU³⁷) obtained by means of NMR spectroscopy was used for a comparative analysis of the effects of BDF4 and DNA binding on the conformational dynamics of the target protein. A short canonical double-stranded DNA duplex (DS14) was used as model DNA in the NMR experiment. Analysis of the crystal structures (PDB ID 1P51, 1P71, 1P78, 2NP2 and 4QJU) obtained for HU proteins from cyanobacteria *Anabaena*, *Borrelia burgdorferi* and *Staphylococcus aureus*^{7,40,41} in complexes with DNA of various lengths indicated that such short DNA had to interact directly only with amino acid residues from the DNA-binding β -saddle domain including C-terminal $\alpha 3$ -helix. The same mode of interaction was observed in a NMR study of HU protein from *H. pylori* in complex with 25 bp harpin DNA³⁹. Thus, since BDF4 was designed and thought to recognize the core of the DNA-binding platform, chemical shift and cross-peak intensity perturbations in the NMR spectra of HUSpm titrated with BDF4 were supposed to be somehow similar to the CSPs and CIPs observed in case of the HUSpm + DS14 complex. In fact, addition of BDF4 to HUSpm caused the opposite effects (described in the Results) relatively to those expected for and observed in the NMR spectra of HUSpm after addition of DS14 (Figs. 3 and 4).

Since BDF4 specificity to the DNA binding site at the base of the β -saddle was not confirmed by NMR spectroscopy, we performed a new docking experiment, in which the whole surface area of HUSpm was the target for the inhibitor. Two symmetrical binding sites located on opposite side-surfaces of HUSpm were identified; each site overlapped a dimer interface between the α -helical domain of one monomer and the side portion of the DNA-binding domain of another monomer (Fig. 5B–D). Conformational dynamics of the HUSpm dimer observed during MD simulations of this model (Fig. 5E) coincided mostly with chemical shift and cross-peak intensity perturbations detected by NMR spectroscopy.

To distinguish finally between the initially proposed BDF4 binding site in the core of the β -saddle domain and the new site, we performed site-directed mutagenesis. It demonstrated that alanine substitutions of basic

DNA-interacting residues^{7,39–41} which reside in the core of DNA-binding domain and presumably participate in inhibitor binding (Fig. 1D,E), actually disturb DNA-binding, but do not affect the activity of BDF4 (Fig. 6A–C). This result confirmed the NMR-based conclusion about different locations of BDF4 and DS14 binding sites on the HUSpm dimer.

At the same time, K35T substitution, which weakened the HUSpm dimerization interface by disturbing the hydrogen bond between the β 2-loop- β 3 of one monomer with the α 2-helix of another monomer⁹, led to an increase of HUSpm sensitivity to BDF4 (Fig. 6D). It is conceivable that BDF4 binding to the site overlapping the aforementioned area of the dimer interface (Fig. 5C,D) initiated destructive changes in the dimerization interface. Moreover, according to the MD study, BDF4 binding to the side-surface of HUSpm significantly influences flexibility of the DNA-binding domain (Fig. 5E).

Thus, the inhibitor binding in the site located at the interface between the α -helical domain of one monomer and the DNA-binding domain of another monomer causes a conformational rearrangement of the latter. We strongly believe that this conformational rearrangement fixes the HU dimer in a defective DNA binding mode, in which the electrostatic interactions between DNA and protein are not supported by proper conformational adaptation of the DNA-binding domain (1), the strength of the DNA binding is lowered to a level not detectable in EMSA (2), and therefore the protein is not able to perform its biological function as the architect of the bacterial genome (3).

We suggest that any inhibitors targeting the DNA-binding surface may mimic and, to be precise, partially mimic the electrostatic part of DNA-HU protein interaction. This gives micromolar dissociation constants of protein-inhibitor complexes^{35,36} vs. nanomolar dissociation constants for HU protein-DNA complexes^{5,35,41} (for K_d values of the HUSpm-DNA complexes see Table 2). Thus, high potency of inhibitors targeting the DNA-binding surface in vivo may be questionable. Meanwhile, as we report here, there is another way to disrupt DNA-binding features of HU proteins without direct competition for the binding site between the inhibitor and native HU protein “ligand”, such as DNA. Both MD simulation of the protein-inhibitor complexes and NMR study performed here suggest that the α -helical domain of HU protein is “druggable” by so called “interface inhibitors”, which are able to disturb the DNA binding allosterically, and bisphenol derivatives of fluorene are the first examples of such inhibitors. We believe that new molecules with similar features are going to be identified soon, since targeting the interface area of the HU proteins’ α -helical body may be less challenging than targeting the large and adaptive DNA-binding surface.

Conclusion

Here we report bisphenol derivatives of fluorene (BDFs) as a new type of chemical probes targeting histone-like HU proteins, which are global regulators of bacterial nucleoids, through dimer interface perturbation. BDFs were found by virtual screening targeting the DNA-binding surface in the core of the β -saddle-like domain of HU protein from *S. melliferum*. They inhibited the DNA binding properties of three tested HU proteins from mycoplasmas *S. melliferum* and *M. gallicepticum*, and from enterobacteria *E. coli* with IC_{50} values within the range from 5 to 10 μ m. In addition, BDFs demonstrated antimicrobial activity against mycoplasmas, but not against *E. coli*, which was consistent with compensatory role of other NAPs in the higher bacteria. At the same time, NMR spectroscopy complemented with Molecular Dynamics and site-directed mutagenesis indicated that the dimer interface between the α -helical domain of one monomer and the β -saddle domain of another monomer is the actual site of inhibitor intervention. Thus, we propose that the allosteric inhibition mode reported here, which does not require direct competition with DNA for the binding site, should be considered during development of small molecule inhibitors of nucleoid-associated proteins as well as other type of DNA-binding multimeric proteins.

Methods

Compound library and virtual screening. Structures of compounds were downloaded from the chemical library collection of Vitas-M laboratory (Causeway Bay, Hong Kong) (<https://vitasmlab.biz/>). Virtual screening was performed using Lead Finder software package (MolTech LLC, Moscow, Russia⁴², the dock_filter protocol⁴³ was applied to filter the subset of ligands interacting simultaneously with both protein “arms” and the beta-saddle so as to screen for compounds with the best abilities to bind to HUSpm’s DNA-binding cleft. Ligands were processed with OpenBabel 2.3.2 software⁴⁴ to remove counterions, to generate physiological charge states and tautomers wherever applicable. Energy-minimized 3D structures for all compounds were generated and subjected to the molecular docking.

To prepare the target protein (PDB ID 5L8Z⁹) for docking, unresolved residues were restored using Modeller software⁴⁵. Hydrogen atoms were added to the HUSpm spatial structure using Build Model unit of the Lead Finder software package. Visualization of protein-inhibitor complexes and intermolecular interaction was performed using the VMD⁴⁶, LigPlot⁴⁷ and the PyMOL Molecular Graphics System, Version 1.9.0.0 (Schrödinger, LLC, New York, NY, USA).

Docking to the whole surface of HUSpm. Virtual screening and molecular docking were performed using the Lead Finder software package (MolTech LLC, Moscow, Russia) with built-in standard precision preset using the approach developed in the previous work⁴⁸. Searching for potential binding sites was performed in the whole surface of HUSpm as follows. A set of 50 energy grids evenly distributed around the protein surface was calculated, and then the test ligand was docked into each grid five times. Binding free energies calculated in each run were averaged in each energy grid. A potential binding site was considered to be found if an average calculated binding free energy for any given grid was significantly lower ($p < 0.05$) than the overall average.

The final compounds obtained by screening were re-docked using the extra precision Lead Finder preset to select the molecules with the most desirable sets of interactions.

Homology modelling and molecular dynamics calculations. Preparation of the HUSpm + DS14 protein-DNA complex was carried out using the COOT interactive graphics program⁴⁹. The structure of HU protein from *B. burgdorferi* complexed with DNA (PDB ID 2np2) was superimposed with the HUSpm model, and the coordinates of the DNA were transferred to the HUSpm model. After that, excess nucleotides were removed from the model, and the remaining nucleotides were replaced by those from the DS14 oligonucleotide duplex (Supplementary Table S4).

MD simulations were carried out for free HUSpm and its complexes with DS14 DNA (Supplementary Table S4) and BDF4 (Supplementary Table S3) using the GROMACS 2019.2 simulation package⁵⁰. OPLS-AA/L force field⁵¹ was used with a rectangular model box with a minimum distance from the solute molecule of 10 Å and the TIP3P water model. Sodium and chloride ions were added to the system for equalization of the total charge and adjustment of the ion strength to 0.15 M. A topology for BDF4 was created using the LigParGen service⁵². Charges were computed using the RESP procedure implemented in Antechamber⁵³ based on HF/6-31G QM-optimized geometry and electron density. The simulation protocol included 5,000 steps of steepest descent energy minimization, 100 ps NVT equilibration, and 1,000 ps NPT equilibration. The pressure and temperature of the system were controlled, where appropriate, with the Parinello-Rahman barostat and the V-rescale thermostat, respectively. All calculations were performed using the supercomputer of NRC “Kurchatov Institute”. The simulations were repeated two or three times with similar results.

Purification of the recombinant HU proteins. Cloning, expression and purification of the recombinant HU proteins have been described in our previous works for HUSpm⁵⁴, HUMgal⁵⁵, and HUEc⁵. Preparation of samples for heteronuclear NMR spectroscopy was carried out as described in^{37,56,57} with minor modifications. Briefly, ¹⁵N-labeled HUSpm was purified using Ni-NTA Superflow column (Qiagen, Hilden, Germany) and digested with TEV-protease (1 mg per 10 mg of protein). The 6xHisTev-tag and TEV-protease were removed by the second run of Ni-NTA affinity chromatography. After that, instead of size-exclusion chromatography, HUSpm was subjected to buffer exchange to 20 mM TrisHCl buffer pH 7.8 using PD10 desalting columns (GE Healthcare, Chicago, IL, USA) and loaded to the ResourceS cation exchange column (GE Healthcare, Chicago, IL, USA) equilibrated with the same buffer. The column was washed with the linear gradient of NaCl from 10 mM to 1 M. The major peak containing HUSpm was detected at NaCl concentration about 26%. This step prevented the proteolytic degradation previously observed during NMR spectroscopy. Finally, the protein solution was exchanged to 50 mM Na₂HPO₄/NaH₂PO₄ buffer containing 0.1 mM NaN₃ in 9/1 H₂O/D₂O (v/v) mixture with pH 7.2, and subsequently it was concentrated to 20 mg/ml (which corresponded to 1 mM of HUSpm dimer) using a 3 kDa cutoff centrifugal filter device (Millipore, Burlington, MA, USA).

Electrophoretic mobility shift assay (EMSA). The fluorescent-labeled oligonucleotide duplexes described in Supplementary Table S4 were prepared by annealing labeled oligonucleotides, 5'-HEX-D48 or 5'-HEX-D24a, and corresponding non-labeled oligonucleotides. All oligonucleotides were synthesized by Evrogen (Moscow, Russia). The annealing was performed by incubating the oligonucleotides (10–50 μM) for 3 min at 90 °C in 20 mM Tris-HCl (pH 8.0) and 200 mM NaCl, and then allowing them to cool slowly (~ 4 h) to 40 °C.

Binding of HU proteins to DNA was analyzed by electrophoretic mobility shift assays (EMSA) performed as described previously⁵. Different amounts of the HU proteins (30 nM HUSpm and HUSpm_K35T, 60 nM HUMgal and HUEc, 100 nM HUSpm_K80A and HUSpm_K88A, and 200 nM HUSpm_R57A) were incubated with DNA (10 nM) in 10 μl of the binding buffer (20 mM Tris-HCl pH 8.0, 7% glycerol, and 150 mM of NaCl) for 10 min. Then, the samples were loaded to pre-run (125 V, 30 min) 10% non-denaturing PAAG, buffered with 100 mM Tris-borate, and electrophoretically separated (125 V, 90 min). For 48 bp DNA duplexes, 8% PAAG was used. Gels were visualized using BIO RAD Faros FX Molecular Imager (532 nm EX, 605 nm BP), and quantification was performed using Quantity One software. The dissociation constants (K_d) were calculated as described earlier¹⁶.

To evaluate the effects of inhibitors, 1 μl volumes of serial dilutions of tested compounds or DMSO were added to the preformed protein-DNA complexes. After five minute incubation, the assay was carried out as described above. Inhibition curves and IC₅₀ values were obtained using GraphPad Prism 5 software. The figures showed representative results of an independent assay; the combined data from three independent experiments are shown in the tables.

Heteronuclear NMR spectroscopy. The heteronuclear NMR experiments⁵⁸ were recorded with uniformly ¹⁵N-labeled HUSpm (1 mM, dissolved in 50 mM Na₂HPO₄/NaH₂PO₄ buffer containing 0.1 mM NaN₃ in 95/5 H₂O/D₂O (v/v) mixture with pH 7.2) at 30 °C on Varian NMR-system 700 MHz spectrometer (Varian-Agilent, USA) equipped with ¹H/¹³C/¹⁵N gradient cryoprobe. The Resource Center of Molecular and Cell Biology of the Kurchatov complex of NBICS-technologies NRC “Kurchatov Institute” kindly provided access to the instrument. Two-dimensional heteronuclear ¹H/¹⁵N-HSQC (heteronuclear single quantum correlation) and ¹H/¹⁵N-TROSY (transverse relaxation-optimized spectroscopy, BEST version⁵⁹) spectra were acquired to monitor the HUSpm interactions with DNA and BDF4. The spectra were analyzed using Computer Aided Resonance Assignment (CARA)⁶⁰ based on the previously obtained HUSpm chemical shifts³⁷.

For studying the interaction of HUSpm with DNA, 14 bp oligonucleotide duplexes DS14 (Supplementary Table S4) were prepared as described above (see EMSA section) at the 5 mM concentration in the aqueous 50 mM phosphate buffer of pH 7.2. The complex of 0.5 mM ¹⁵N-labeled HUSpm with DS14 was prepared by mixing the protein with 20% molar excess of the DNA.

For studying the effects of inhibitor, BDF4 dissolved in DMSO (or DMSO only) was gradually added to ^{15}N -labeled HUSpm (0.5 mM) either in apo-form or in complex with DS14 with the molar ratios of the protein dimers or protein-DNA complexes were 1/1.2 and 1/2.0, after which the high protein precipitation was observed. The addition of the similar amount of DMSO to the apo-form or complex with DS14 did not result in any changes in the NMR spectra.

For quantification of the changes in the NMR spectra due to the interactions of HUSpm with DS14 and BDF4, the weighted chemical shift perturbations (CSPs) and normalized cross-peak intensity perturbations (CIPs) of the backbone amide groups were estimated at protein-dimer/DNA-duplex and protein-dimer/inhibitor molar ratios of 1/1.2 (where the protein precipitation was not observed). The CSPs for backbone amide $^{15}\text{N}^{15}\text{N}^{\text{H}}$ and $^1\text{H}^{\text{N}}$ resonances were calculated by the formula $\Delta\delta = [(\Delta^1\text{H}^{\text{N}})^2 + (\Delta^{15}\text{N}^{\text{H}}/5)^2]^{0.5}$, where $\Delta^1\text{H}^{\text{N}}$ and $\Delta^{15}\text{N}^{\text{H}}$ are changes in the chemical shifts of backbone amide ^1H and ^{15}N , respectively, upon either DS14 or BDF4 binding to HUSpm, or BDF4 binding to HUSpm + DS14 complex. The CIPs for HUSpm backbone amides were calculated by the formula $\Delta I/I_{\text{av}} = I/I_{\text{av}} - I_{\text{ref}}/I_{\text{ref,av}}$, where I and I_{ref} were the intensities of the amide cross-peaks in the pairs of the NMR spectra of (HUSpm + DS14 and HUSpm), (HUSpm + BDF4 and HUSpm) or (HUSpm + DS14 + BDF4 and HUSpm + DS14), which were normalized by the averaged “standard” cross-peak intensity I_{av} estimated along the region T27-K35 that undergoes minimal local flexibility of the protein backbone according to previous study³⁷.

Site-directed mutagenesis. Easy single-primer site-directed mutagenesis was performed as described in⁶¹. The synthetic oligonucleotide primers designed for the changing of amino acids (one primer for each mutant) and those designed for the selection of mutant clones are listed in the Supplementary Table S5. Eighteen cycles of PCR were performed with the template of the HUSpm-expressing plasmid using the Tersus Plus PCR kit (Evrogen, Moscow, Russia) according to the manufacturer’s recommendations. The PCR products were treated with DpnI endonuclease (Thermo Fisher Scientific, Massachusetts, United States), which digested the parental DNA template, and then transformed into *E. coli* Match1 competent cells. The mutant clones were selected by PCR performed directly on colonies using Taq DNA polymerase (Evrogen, Moscow, Russia) with check primers (Supplementary Table S5) and an appropriate T7 universal primer. Plasmid DNAs purified from mutant clones were sequenced to ensure the absence of random mutations associated with PCR. The expression and purification of mutant proteins were performed in the same manner as described for the wild type HUSpm. The purity of the mutants was estimated by SDS-PAGE with Coomassie staining, and the protein concentrations were measured using the Bicinchoninic Acid Protein Assay Kit (Sigma-Aldrich, St. Louis, USA).

Inhibition of mycoplasma growth. The *S. melliferum* KC3 strain was kindly provided by Prof. G. Wroblewski of the University of Rennes, France. The *M. gallisepticum* S6 strain was provided by Prof. S. N. Borkhsenius, Institute of Cytology, St. Petersburg, Russian Academy of Science. The bacteria were cultivated as described in⁶² without any specific aeration in covered flasks, for less than five passage. *M. gallisepticum* S6 was cultivated in a modified Edward’s liquid medium (tryptose (20 g/l), Tris (3 g/l), NaCl (5 g/l), KCl (1.3 g/l), yeast dialysate (5%), heat inactivated (56 °C–30 min) horse serum (6%), glucose (0.5%) and penicillin 300U/ml, pH 7.5) at 37 °C for 20 h after 1:100 dilution. *S. melliferum* was cultivated in modified SP4 medium (tryptone (10 g/l), peptone (10 g/l), BHI (2.5 g/l), NaCl (4.5 g/l), yeast extract (7 g/l), sorbitol (70 g/l), horse serum (10%), sucrose (1%), fructose (0.1%), glucose (0.1%) and penicillin 300U/ml, pH 7.6) at 30 °C for 20 h after 1:10 dilution. All reagents were from Sigma-Aldrich (St. Louis, MO, USA).

For the broth micro-dilution assay, decreasing concentrations of BDFs dissolved in DMSO (or DMSO alone as a control) were mixed with the predetermined number of mycoplasma cells ($10^4/\text{ml}$) in the 0.4 ml of standard broth medium containing pH indicator phenol red (red color at pH 7.5) in 48-well microplates. The plates were incubated until red color was changed to yellow (due to the pH decrease associated with lactate accumulation during mycoplasma growth) in the control (DMSO containing) wells. The MICs were considered as the lowest concentrations of BDFs that inhibited the color change at the time when it observed in the control wells. The experiments were repeated three times.

Received: 26 May 2020; Accepted: 25 August 2020

Published online: 15 September 2020

References

- Dillon, S. C. & Dorman, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**, 185–195 (2010).
- Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* **181**, 6361–6370 (1999).
- Rouviere-Yaniv, J. & Gros, F. Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*. *Proc. Natl. Acad. Sci.* **72**, 3428–3432 (1975).
- Grove, A. Functional evolution of bacterial histone-like HU proteins. *Curr. Issues Mol. Biol.* <https://doi.org/10.21775/cimb.013.001> (2011).
- Kamashev, D. *et al.* Comparison of histone-like HU protein DNA-binding properties and HU/IHF protein sequence alignment. *PLoS ONE* **12**, e0188037 (2017).
- White, S. W., Wilson, K. S., Appelt, K. & Tanaka, I. The high-resolution structure of DNA-binding protein HU from *Bacillus stearothermophilus*. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 801–809 (1999).
- Swinger, K. K., Lemberg, K. M., Zhang, Y. & Rice, P. A. Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J.* **22**, 3749–3760 (2003).
- Christodoulou, E., Rypniewski, W. R. & Vorgias, C. E. High-resolution X-ray structure of the DNA-binding protein HU from the hyper-thermophilic *Thermotoga maritima* and the determinants of its thermostability. *Extremophiles* **7**, 111–122 (2003).

9. Boyko, K. M. *et al.* Structural basis of the high thermal stability of the histone-like HU protein from the mollicute *Spiroplasma melliferum* KC3. *Sci. Rep.* **6**, 36366 (2016).
10. Kamashev, D. E., Esipova, N. G., Ebralidse, K. K. & Mirzabekov, A. D. Mechanism of Lac repressor switch-off: Orientation of the Lac repressor DNA-binding domain is reversed upon inducer binding. *FEBS Lett.* **375**, 27–30 (1995).
11. Dey, D., Nagaraja, V. & Ramakumar, S. Structural and evolutionary analyses reveal determinants of DNA binding specificities of nucleoid-associated proteins HU and IHF. *Mol. Phylogenet. Evol.* **107**, 356–366 (2017).
12. Kamashev, D., Balandina, A. & Rouviere-Yaniv, J. The binding motif recognized by HU on both nicked and cruciform DNA. *EMBO J.* **18**, 5434–5444 (1999).
13. Kamashev, D., Balandina, A., Mazur, A. K., Arimondo, P. B. & Rouviere-Yaniv, J. HU binds and folds single-stranded DNA. *Nucleic Acids Res.* **36**, 1026–1036 (2008).
14. Balandina, A., Kamashev, D. & Rouviere-Yaniv, J. The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. *J. Biol. Chem.* **277**, 27622–27628 (2002).
15. Kamashev, D. & Rouviere-Yaniv, J. The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *EMBO J.* **19**, 6527–6535 (2000).
16. Kamashev, D. *et al.* *Mycoplasma gallisepticum* produces a histone-like protein that recognizes base mismatches in DNA. *Biochemistry* **50**, 8692–8702 (2011).
17. Guo, F. & Adhya, S. Spiral structure of *Escherichia coli* HU α provides foundation for DNA supercoiling. *PNAS* **104**, 4309–4314 (2007).
18. Berger, M. *et al.* Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU. *EMBO Rep.* **11**, 59–64 (2010).
19. Prieto, A. I. *et al.* Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res.* **40**, 3524–3537 (2012).
20. Oberto, J., Nabti, S., Jooste, V., Mignot, H. & Rouviere-Yaniv, J. The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS ONE* **4**, e4367. <https://doi.org/10.1371/journal.pone.0004367> (2009).
21. Kamashev, D. E. *et al.* Proteome of HU-lacking *E. coli* studied by means of 2D gel electrophoresis. *Russ. J. Bioorganic Chem.* **45**, 366–373 (2019).
22. Mangan, M. W. *et al.* Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. *Microbiology* **157**, 1075–1087 (2011).
23. Stojkova, P. *et al.* HU protein is involved in intracellular growth and full virulence of *Francisella tularensis*. *Virulence* **9**, 754–770 (2018).
24. Priyadarshini, R. *et al.* The nucleoid-associated protein HU β affects global gene expression in *Porphyromonas gingivalis*. *Microbiology* **159**, 219–229 (2013).
25. Conforte, V. P. *et al.* The histone-like protein HupB influences biofilm formation and virulence in *Xanthomonas citri* ssp. *citri* through the regulation of flagellar biosynthesis. *Mol. Plant Pathol.* **20**, 589–598 (2018).
26. Guan, Z., Wang, Y., Gao, L., Zhang, W. & Lu, X. Effects of the histone-like protein HU on cellulose degradation and biofilm formation of *Cytophaga hutchinsonii*. *Appl. Microbiol. Biotechnol.* **102**, 6593–6611 (2018).
27. Huisman, O. *et al.* Multiple defects in *Escherichia coli* mutants lacking HU protein. *J. Bacteriol.* **171**, 3704–3712 (1989).
28. Micka, B. & Marahiel, M. A. The DNA-binding protein Hbsu is essential for normal growth and development in *Bacillus subtilis*. *Biochimie* **74**, 641–650 (1992).
29. Ferrándiz, M.-J., Carreño, D., Ayora, S. & de la Campa, A. G. HU of *Streptococcus pneumoniae* is essential for the preservation of DNA supercoiling. *Front. Microbiol.* **9**, 493 (2018).
30. Griffin, J. E. *et al.* High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**, e1002251 (2011).
31. Glass, J. I. *et al.* Essential genes of a minimal bacterium. *PNAS* **103**, 425–430 (2006).
32. Karr, J. R. *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401 (2012).
33. Waites, K. B., Lysnyansky, I. & Bebear, C. M. Emerging antimicrobial resistance in mycoplasmas of humans and animals. in *Mollicutes: Molecular Biology and Pathogenesis* (ed. Browning, G., Citti, C.) 289–322 (Caister Academic Press, 2014).
34. Chernova, O. A., Medvedeva, E. S., Mouzykantov, A. A., Baranova, N. B. & Chernov, V. M. Mycoplasmas and their antibiotic resistance: The problems and prospects in controlling infections. *Acta Naturae* **8**, 24–34 (2016).
35. Bhowmick, T. *et al.* Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. *Nat. Commun.* **5**, 4124 (2014).
36. Ruili, L. *et al.* The structural basis of African swine fever virus pA104R binding to DNA and its inhibition by stilbene derivatives. *PNAS* **117**, 11000–11009 (2020).
37. Timofeev, V. I. *et al.* Structural plasticity and thermal stability of the histone-like protein from *Spiroplasma melliferum* are due to phenylalanine insertions into the conservative scaffold. *J. Biomol. Struct. Dyn.* **36**, 4392–4404 (2018).
38. Whithear, K. G., Bowtell, D. D., Ghiocas, E. & Hughes, K. L. Evaluation and use of a micro-broth dilution procedure for testing sensitivity of fermentative avian mycoplasmas to antibiotics. *Avian Dis.* **27**, 937 (1983).
39. Jaiswal, N. *et al.* NMR elucidation of monomer-dimer transition and conformational heterogeneity in histone-like DNA binding protein of *Helicobacter pylori*. *Magn. Reson. Chem.* **56**, 285–299 (2018).
40. Mouw, K. W. & Rice, P. A. Shaping the *Borrelia burgdorferi* genome: Crystal structure and binding properties of the DNA-bending protein Hbb. *Mol. Microbiol.* **63**, 1319–1330 (2007).
41. Kim, D. H. *et al.* beta-Arm flexibility of HU from *Staphylococcus aureus* dictates the DNA-binding and recognition mechanism. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 3273–3289 (2014).
42. Stroganov, O. V., Novikov, F. N., Stroylov, V. S., Kulkov, V. & Chilov, G. G. Lead finder: An approach to improve accuracy of protein–ligand docking, binding energy estimation, and virtual screening. *J. Chem. Inf. Model.* **48**, 2371–2385 (2008).
43. Novikov, F. N., Stroylov, V. S., Stroganov, O. V. & Chilov, G. G. Improving performance of docking-based virtual screening by structural filtration. *J. Mol. Model.* **16**, 1223–1230 (2010).
44. O’Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).
45. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**(3), 779–815 (1993).
46. Humphrey, W., Dalke, A. & Schulten, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
47. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple ligand–protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).
48. Zanyatkin, I. *et al.* Inhibition of prion propagation by 3,4-dimethoxycinnamic acid. *Phytother. Res.* **31**, 1046–1055 (2017).
49. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
50. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
51. Jorgensen, W. L. & Tirado-Rives, J. Monte Carlo vs molecular dynamics for conformational sampling. *J. Phys. Chem.* **100**, 14508–14513 (1996).
52. Dodda, L. S., Cabeza de Vaca, I., Tirado-Rives, J. & Jorgensen, W. L. LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **45**, W331–W336 (2017).
53. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

54. Boyko, K. *et al.* Expression, purification, crystallization and preliminary X-ray crystallographic analysis of the histone-like HU protein from *Spiroplasma melliferum* KC3. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71**, 24–27 (2015).
55. Nikolaeva, A. Y. *et al.* Isolation, purification, crystallization, and preliminary X-ray diffraction study of the crystals of HU protein from *M. gallisepticum*. *Crystallogr. Rep.* **60**, 880–883 (2015).
56. Altukhov, D. A. *et al.* Preparation of the recombinant HU-proteins from *S. melliferum* and *M. gallisepticum* and of their complexes with DS-DNA for structural NMR experiments. *Moscow Univ. Chem. Bull.* **71**, 221–226 (2016).
57. Altukhov, D. A. *et al.* Enhanced conformational flexibility of the histone-like (HU) protein from *Mycoplasma gallisepticum*. *J. Biomol. Struct. Dyn.* **36**, 45–53 (2018).
58. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. III. & Skelton, N. J. Protein NMR spectroscopy, 2nd edition principles and practice.. *Curr. Biol.* <https://doi.org/10.1016/B978-012164491-8/50005-1> (2007).
59. Favier, A. & Brutscher, B. Recovering lost magnetization: Polarization enhancement in biomolecular NMR. *J. Biomol. NMR* **49**, 9–15 (2011).
60. Keller, R. *The computer aided resonance assignment tutorial* (Cantina Verlag, Goldau, 2004).
61. Mikhailova, A. G. *et al.* Activity modulation of the oligopeptidase B from *Serratia proteamaculans* by site-directed mutagenesis of amino acid residues surrounding catalytic triad histidine. *Biochimie* **139**, 125–136 (2017).
62. Vanyushkina, A. A., Fisunov, G. Y., Gorbachev, A. Y., Kamashev, D. E. & Govorun, V. M. Metabolomic analysis of three mollicute species. *PLoS ONE* **9**, e89312 (2014).

Acknowledgements

We thank Dr. Dmitri E. Kamashev for valuable advices about the DNA binding features of HU proteins. This research was funded by RFBR grant № 20-04-01001 to T.V.R., V.I.T, A.V.V, Y.K.A and D.A.A (study of HU proteins in complexes with DNA and inhibitors), Fundamental Research Program of the Presidium of RAS “Molecular and Cell Biology” (Grant # 0101-2014-0086) to E.V.S (in vivo experiments), and Russian Science Foundation (Grant # 20-64-46027) to E.V.B (in part of NMR data analysis). The work was partially supported by the Resource Center of Molecular and cell biology of NRC Kurchatov Institute to D.A.A and by the Ministry of Science and Higher Education within the State assignment FSRC “Crystallography and Photonics” RAS to V.I.T (in part of structural analysis).

Author contributions

V.S.S and V.I.T performed virtual screening and docking, homology modelling and MD simulations. V.S.M synthesised the inhibitors. D.A.K and A.V.V purified recombinant proteins. E.V.B and D.A.A performed NMR spectrometry. Y.K.A and E.V.S performed mutagenesis, EMSA and broth micro-dilution assay. Y.K.A, D.A.A, V.I.T, V.S.S, E.V.S, E.V.B, and T.V.R analysed and interpreted the data. Y.K.A, V.I.T, V.S.S, E.V.B, E.V.S, and T.V.R wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-72113-4>.

Correspondence and requests for materials should be addressed to Y.K.A., E.V.B. or T.V.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020