Published online 7 August 2021

Nucleic Acids Research, 2021, Vol. 49, No. 16 ***9229–9245***
https://doi.org/10.1093/nar/gkab678

A phage-encoded nucleoid associated protein compacts both host and phage DNA and derepresses H-NS silencing

Bokyung Son[1,†], Jennifer Patterson-West[1,†], Melissa Arroyo-Mendoza[1], Revathy Ramachandran[2], James R. Iben[3], Jingen Zhu[4], Venigalla Rao[4], Emilios K. Dimitriadis[5] and Deborah M. Hinton[1,*]

[1]Gene Expression and Regulation Section, Laboratory of Biochemistry and Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA, [2]Laboratory of Biochemistry and Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, [3]Molecular Genomics Core, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA, [4]Department of Biology, The Catholic University of America, Washington, DC, USA and [5]Trans-NIH Shared Resource on Biomedical Engineering and Physical Science, National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Bethesda, MD, USA

Received September 10, 2020; Revised July 20, 2021; Editorial Decision July 21, 2021; Accepted August 06, 2021

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

ABSTRACT

Nucleoid Associated Proteins (NAPs) organize the bacterial chromosome within the nucleoid. The interaction of the NAP H-NS with DNA also represses specific host and xenogeneic genes. Previously, we showed that the bacteriophage T4 early protein MotB binds to DNA, co-purifies with H-NS/DNA, and improves phage fitness. Here we demonstrate using atomic force microscopy that MotB compacts the DNA with multiple MotB proteins at the center of the complex. These complexes differ from those observed with H-NS and other NAPs, but resemble those formed by the NAP-like proteins CbpA/Dps and yeast condensin. Fluorescent microscopy indicates that expression of *motB in vivo,* at levels like that during T4 infection, yields a significantly compacted nucleoid containing MotB and H-NS. *motB* overexpression dysregulates hundreds of host genes; ~70% are within the *hns* regulon. In infected cells overexpressing *motB*, 33 T4 late genes are expressed early, and the T4 early gene *repEB*, involved in replication initiation, is up ~5-fold. We postulate that MotB represents a phage-encoded NAP that aids infection in a previously unrecognized way. We speculate that MotB-induced compaction may generate more room for T4 replication/assembly and/or leads to beneficial global changes in host gene expression, including derepression of much of the *hns* regulon.

INTRODUCTION

Even in the well-studied bacteriophage T4, many genes encode products that have no homologs outside the phage world and whose functions have not yet been characterized (1). These genes represent a major portion of the 'dark matter' of the biological universe, providing a rich source of information about novel protein functions and structures (2). For lytic phages, such as T4, most of these uncharacterized proteins are early products, whose synthesis begins shortly after infection and which are thought to usurp host functions or combat host anti-viral activities. Though typically nonessential, these early genes can optimize infection and/or be required only under certain conditions.

The T4 nonessential early gene, *motB*, is highly toxic for *Escherichia coli* when expressed heterologously, suggesting that it might encode such a 'takeover' protein (3). Our previous work demonstrated that MotB binds both host and T4 DNA tightly and nonspecifically, and T4 phage with a *motB* knockdown yields a 2-fold lower burst than wild type (WT) T4 under laboratory conditions (3).

*To whom correspondence should be addressed. Tel: +1 301 496 9885; Fax: +1 301 402 0053; Email: dhinton@helix.nih.gov
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
Present addresses:
Jennifer Patterson-West, Office of Research Training and Development, National Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD, USA.
Revathy Ramachandran, Mohammed Bin Rashid University, Dubai Healthcare City, Dubai, UAE.

Published by Oxford University Press on behalf of Nucleic Acids Research 2021.
This work is written by (a) US Government employee(s) and is in the public domain in the US.

Unexpectedly, this work also showed that when bound to DNA, MotB co-purifies with the highly abundant *E. coli* histone-like protein H-NS and its less abundant paralog StpA (3), suggesting that MotB function might involve these host proteins. In the bacterial nucleoid, H-NS and StpA, along with other members of the Nucleoid Associated Proteins (NAPs) family, form higher-order nucleoprotein complexes with the host genomic DNA, organizing the DNA within the nucleoid and leading to transcriptional effects through regulation of specific genes (4–9). In particular, H-NS targets DNA with high AT-content, generating regions of 'stiffness' by linearly coating the DNA and regions with loops and bridges through H-NS protein-protein interactions (10–16). Proteins within the Hha/YdgT family present in enteric bacteria can also form heterocomplexes with H-NS, increasing this silencing [reviewed in (17)]. H-NS binding typically silences gene expression of the bound DNA by preventing RNA polymerase (RNAP) from binding or by sequestering bound RNAP (4,7). In addition, *in vitro* H-NS bridging increases transcriptional pausing and Rho-dependent transcription termination (11), which would also down-regulate gene expression. However, it is not clear yet whether this mechanism operates *in vivo*.

H-NS targeting of AT-rich sequences can also result in the repression of horizontally acquired pathogenicity islands and phage DNA, which are often more AT-rich than the 49% AT-containing *E. coli* genome (17–20). However, mobile elements and phages have mechanisms to thwart this repression. A family of H-NS derivatives, which lack the C-terminal DNA binding domain, but can still bind tightly to the H-NS central dimerization domain, are encoded by pathogenicity islands (10,16,21,22) while the T7 5.5 protein can also bind to H-NS (23,24). In these cases, the binding of these proteins to H-NS eliminates its ability to oligomerize and form higher-ordered complexes. In contrast, the T4 Arn protein is a DNA mimic that binds to the H-NS DNA binding domain, thus serving as a competitor for the phage genome (25). Thus, in these situations, the proteins encoded by the phage or mobile element can abrogate H-NS repression of the foreign DNA. However, in a T4 *motB* knockdown infection of *E. coli* B, there is only a slight delay in the expression of a few T4 late genes, and as infection proceeds, no significant effect on phage gene expression is observed (3). These results have suggested that shutting down H-NS repression of T4 gene expression might not be the primary function of MotB.

Here, we have used RNA-seq to determine how the presence of MotB affects host and phage gene expression in *E. coli* K12 and atomic force microscopy (AFM), fluorescent microscopy, and flow cytometry to investigate how MotB interacts with DNA and how this interaction is affected by H-NS. Our results indicate that a level of MotB that is similar to what is present during T4 infection compacts DNA both *in vitro* and *in vivo*, with multiple MotB proteins centered within the complex. The MotB/DNA complexes resemble those previously observed with the *E. coli* NAP-like proteins, DNA binding protein from starved cells (Dps) (26,27) and curved DNA-binding protein A (CbpA) (4) and with the yeast structural maintenance of chromosome (SMC) protein cohesin (28). *In vitro* the compaction formed in the presence of MotB is still observed in the presence of H-NS, indicating that MotB alters the types of structures that are made by H-NS alone. The overexpression of *motB* in *E. coli* K12 in the absence or presence of T4 infection dysregulates hundreds of host genes, including ∼70% of genes that are within the *hns* regulon. In addition, 33 T4 late genes are expressed earlier and expression of one T4 early gene *repEB*, which is involved in the initiation of replication (29–31), is up ∼5-fold at 5 min post-infection, a time when T4 replication commences. Our results suggest that MotB represents a phage-encoded NAP-like protein that can compact the host chromosome in the presence of the host NAPs. We postulate that this MotB-induced compaction early in infection could aid T4 by providing more room for subsequent phage replication and assembly and/or by the associated changes in host gene expression, including the dysregulation of H-NS repressed genes. We conclude that MotB represents a previously unrecognized way to help phage in host takeover.

MATERIALS AND METHODS

Strains

E. coli K12 strains TOP10F' (Invitrogen, Carlsbad, CA, USA), SX454 (9,32,33), and B834 (34) have been described. As indicated, LB media (Quality Biological, Gaithersburg, MD, USA) or MOPS EZ Rich defined media (Teknova, Hollister, CA, USA) were used, and unless otherwise noted, cells were grown at 37°C with shaking at 250 rpm.

DNA

The following plasmids have been previously described: pMLH07, a 6 kbp plasmid containing known H-NS binding sites (35); pNW129, a pACYC-based vector plasmid containing the P_{BAD} promoter followed by a multiple cloning site (36); pNW129-*motB* (referred to as p*motB*) in which a codon-optimized *motB* is located downstream of the inducible promoter P_{BAD} in pNW129 (3); pNW129-*motB-his* (referred to here as p*motB-his*), which is the same as p*motB*, except that *motB* has a C-terminal His_6-tag (3); pTXB1-HNS, an intein expression plasmid for H-NS production (3), and pUCBB-*eGFP* (enhanced green fluorescent protein, referred to here as p*eGFP*, plasmid # 32553, Addgene, Watertown, MA, USA (37)). pUCBB-*motBeGFP* (referred to here as p*motB-eGFP*) was constructed by cloning *motB* lacking its stop codon between the BgIII and NdeI sites immediately upstream of the *eGFP* gene in the vector p*eGFP*. This places *motB-eGFP* under the control of P_{BAD}. The plasmids pTE103-*motB-his* and pTE103-*motB-eGFP* consist of the T4 DNA from nucleotide 87 of *motB* through nucleotide 40 of the T4 downstream gene *cef* with the insertion of either a *his*$_6$ -TAA (stop codon) or *eGFP*-TAA at the C-terminus of *motB*, respectively, cloned into the vector pTE103 (38) between the BamHI and SalI sites. For both plasmids, the native stop codon of *motB* was omitted to fuse the C-terminus of *motB* to the tag or *eGFP*. GenScript (Piscataway Township, NJ, USA) performed gene synthesis, plasmid construction, and DNA sequencing of the constructed plasmids.

pET28b-*motB-his* and pET28b-*motB-eGFP*, kanamycin-resistant donor plasmids for the CRISPR-mediated con-

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

struction of T4-*motB-his* and T4-*motB-eGFP*, were constructed by PCR by first generating *motB-his* and *motB-eGFP* donor fragments using the plasmids pTE103-*motB-his* and pTE103-*motB-eGFP*, respectively, and the primers:

MotB-BglII-F: 5′TCGAGATCTGGAAGGGTTGTTTATACTCGCGC3′ and
MotB-XhoI-R: 5' GTGCTCGAGAGTGGAGCGATACGTTTCAGAAG3′.

(The BglII and XhoI recognition sequences are underlined.) The PCR products were digested with the restriction enzymes BglII and XhoI and cloned into pET28b (EMD Millipore, Burlington, MA, USA), which had been linearized with BglII and XhoI.

For the *CRISPR-Cpf1-motB* spacer plasmid construction, the spacer fragment was generated as follows: First, using 2X Phusion High-Fidelity PCR Master mix (ThermoFisher Scientific, Waltham, MA, USA), the CRISPR-Cpf1 backbone plasmid [referred to as CRISPR-LbCpf1/SpCas9 plasmid in (39)], and the primers listed below, two fragments were generated: a 300 bp left fragment using primers Cpf1-F and motB-ST-R and a 300 bp right fragment using primers motB-ST-F and Cpf1-R.

Primer Cpf1-F: 5′GTACGCACAGACATCGGTTAAACAC3′

Primer Cpf1-R: 5′AGAAAAAAAGGATCTCAAGAAGATC3′

Primer motB-ST-F: 5′(0:underline)ATGTTCGTCAAAAGACCTTTTTAATT(/0:underline)TCTACTAAGTGTAGATtgcatagagagcattgctatGAATTCTGTAAGCTTT3′ (The 26 bp complementary nucleotides are underlined, and a motB-spacer sequence is shown as lower-case.)

Primer motB-ST-R:
5′(0:underline)AATTAAAAAGGTCTTTTGACGAACAT(/0:underline)acaaagtcattcctATCTACACTTAGTAGAAATTAAAAAGGTCTTTTGACCTCGAG3′ (The 26 bp complementary nucleotides are underlined and the other motB-spacer sequence is shown as lower-case.)

The entire spacer fragment was then amplified by PCR in a reaction containing the purified left and right fragments (10 ng each, 50 µl reaction) without the primers to perform a template extension reaction (5–10 cycles), which allowed the overlapping regions of template to anneal and be extended by DNA polymerase. Then, the primers Cpf1-F and Cpf1-R were added to complete the generation of the entire spacer fragment. The purified *motB* spacer PCR product (572 bp) was digested with XhoI and EagI and cloned into the CRISPR-Cpf1 spacer plasmid that had been linearized with XhoI/EagI.

T4 genomic DNA (gDNA) was purified as previously described (3). To obtain the 1620 bp unmodified fragment P_{l8} used for AFM, PCR was performed using T4 gDNA, Pfu Turbo polymerase (Stratagene, San Diego, CA, USA), and the primers P_{l8} forward (CTTCTTCCCGCTGGAAATCT) and P_{l8} reverse (CCCGAATGCCTCATTAAATCTTCTGGGTC) (Integrated DNA Technologies, Coralville, IA, USA). These primers annealed to T4 gDNA such that the resulting fragment was composed of positions −143 to + 1477 relative to the transcription start site of the T4 late promoter P_{l8}. In the case of the modified DNA, PCR was performed with the same T4 gDNA and primers, but using AmpliTaq Gold 360 polymerase (Thermo Fisher Scientific, Waltham, MA, USA) and 5-hydroxymethyl-deoxycytidine triphosphate (5-hydroxymethyl-dCTP, Jena Bioscience, Jena, Germany) instead of dCTP, and previously defined parameters (40). The DNA products were purified using the Qiagen PCR clean-up kit (Qiagen, Hilden, Germany). The 5′-hydroxymethylcytosine (5-HmeC) modified DNA fragment was then glucosylated by incubation with uridine diphosphate glucose (New England Biolabs, Ipswich, MA, USA) and T4 β-glucosyltransferase (New England Biolabs) at 37°C overnight following the New England Biolabs protocol. The glucosylated DNA fragment (GHmeC P_{l8}) was electrophoresed on an E-Gel SizeSelect II 2% (w/v) Agarose Gel (Invitrogen, Carlsbad, CA, USA) and then further purified using the Qiagen PCR clean-up kit. The sizes of the unmodified P_{l8} and GHmeC P_{l8} were compared after electrophoresis on a 0.8% (w/v) agarose gel to confirm glucosylation.

CRISPR-mediated phages T4-*motB-his* and T4-*motB-eGFP* construction

T4-*motB-his* and T4-*motB-eGFP* phages were constructed using CRISPR-Cpf1 mediated editing and recombination as described previously (39,41). The spectinomycin-resistant CRISPR-Cpf1-*motB* spacer plasmid and the kanamycin-resistant donor plasmid, pET28b-*motB-his* or pET28b-*motB-eGFP* were co-transformed into the *E. coli* strain B834. Single plasmids, either the donor plasmid or the CRISPR-Cpf1-*motB* spacer plasmid, were transformed into B834 as controls. T4 phage was added to B834 and incubated for 7 min at 37°C. Then, 3.5 ml of 0.75% top agar containing 50 µg/ml spectinomycin and 50 µg/ml kanamycin was added. The infection mixture was poured onto an LB plate and incubated overnight at 37°C. Single plaques (Generation 1, G1) were picked using a sterile Pasteur pipet and transferred into a 1.5 ml Eppendorf tube containing 200 µl of Pi-Mg buffer (26 mM Na_2HPO_4, 68 mM NaCl, 22 mM KH_2PO_4, 1 mM $MgSO_4$, pH 7.5). The diluted G1 phage was used to infect *motB*-spacer-containing B834 cells to eliminate any parental phage background under CRISPR pressure. The resultant single G2 plaques were selected and used to infect B834 cells (without spacer or donor) to produce G3 phages. The purified single G3 plaques were then selected and transferred into 200 µl of Pi-Mg buffer to make 'zero stocks'. PCR was performed to confirm the presence of the *his* or *eGFP* insertion into the C-terminus of *motB*. One microliter of G3 phage was denatured at 94°C for 8 min and used as a template for PCR using Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA). The amplified DNA fragment was purified using QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) and the sequence was confirmed by DNA sequencing (Retrogene Inc., San Diego, CA, USA). A few drops of chloroform were added to the plaque-purified 'zero stocks' and the phage were stored at 4°C.

Plaque assays

Plaque size assays were performed as previously described (3). Single colonies of TOP10F' containing pNW129 (vector) or p*motB* were obtained after growth on 1.5% LB agar

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

plates containing 40 μg/ml kanamycin, 12 μg/ml tetracycline, and 0.5% (w/v) glucose. Overnight cultures, grown in LB containing 40 μg/ml kanamycin, 12 μg/ml tetracycline, and 0.025% (w/v) glucose, were then used to inoculate the same media but without glucose, and cultures were grown to an OD_{600} of ∼0.4. Protein synthesis was induced by the addition of 0.1% (*w/v*) arabinose (final concentration) for 20 min. Cells were collected on ice and then used for plating WT T4.

Protein purification

MotB containing a C-terminal His_6-tag (MotB-His) and H-NS were purified as previously described (3). After purification, MotB-His and H-NS were stored in MotB storage Buffer [50 mM Tris–HCl (pH 8.0), 0.1 mM ethylenediaminetetraacetic acid (EDTA), 50 mM NaCl, 0.01% (v/v) Triton X-100, 50% (v/v) glycerol, 0.1 mM dithiothreitol (DTT)] and H-NS storage Buffer [10 mM potassium phosphate (pH 7.5), 200 mM NaCl, 0.1 mM EDTA, 50% (v/v) glycerol], respectively, at −20°C.

Western blotting

E. coli TOP10F' or SX454 were grown to early log phase (OD_{600} ∼0.3), and then infected with T4-*motB-his* or T4-*motB-eGFP*, respectively, at a multiplicity of infection (MOI) of 10. For the TOP10F' cultures containing the p*motB-his* plasmid, cells were also grown to OD_{600} ∼0.3 and then as indicated, the synthesis of MotB-His was induced by the addition of arabinose, under the indicated conditions before T4 infection. At the indicated times, a 1 ml aliquot was centrifuged, and the pellet was resuspended in 1× Laemmli sample buffer (Bio-Rad, Hercules, CA, USA) so that 10 μl of sample was equivalent to an OD_{600} of 0.08. Proteins were separated by SDS-PAGE with different amounts of purified MotB-His protein as an amount marker. After electrophoresis, gels were dry blotted onto nitrocellulose membranes (iBlot 2 NC Ministacks, Invitrogen, Carlsbad, CA, USA) by the iBlot 2 transfer system (Invitrogen, 20 V for 7 min). The membranes were blocked with 3% nonfat milk in PBS plus 0.1% Tween-20 (PBS-T), washed with PBS-T, and then incubated with mouse anti-His-Tag IgG diluted 1:10 000 (Santa Cruz Biotechnology, Santa Cruz, CA, USA) or mouse anti-GFP IgG diluted 1:1000 (Santa Cruz Biotechnology) in PBS-T containing 0.3% nonfat milk at room temperature for 1 h, followed by three washes (5 min each) with PBS-T. The membrane was then incubated with goat anti-mouse IgG-HRP conjugated (1:10,000 for His detection or 1:20,000 for GFP detection, Biolegend, San Diego, CA, USA) in PBS-T containing 0.3% nonfat milk at room temperature for 1 h. After five washes (5 min each) with PBS-T, the membrane was developed using the Amersham ECL Plus Western Blotting Detection System (GE Healthcare, Chicago, IL, USA), and the signal was detected using the AI600 Chemiluminescent Imager (Amersham, Pistcataway, NJ, USA).

The level of MotB present in an infected cell was determined as follows using the TOP10F' cells infected with T4-*motB-his* for 5 or 10 min. From 3 independent Western analyses, the amount of MotB-His present in the 5 min T4 infection sample was estimated at 0.71 ± 0.29 pmol, which is equivalent to 4.3×10^{11} ($\pm 1.7 \times 10^{11}$) monomers of MotB. (The concentration of purified MotB-His used as the standard was determined using a Bradford protein assay with bovine serum albumin as the standard protein.) A 0.28 ml of culture at 3.9×10^7 cells/ml (1.09×10^7 cells total) was obtained for the Western analysis. Consequently, there were $\sim 3.9 \times 10^4$ MotB monomers ($\pm 1.6 \times 10^4$) per cell at 5 min of infection. With the *E. coli* chromosome size of 4.6×10^6 bp, this yields 1 monomer of MotB per ∼120 bp of DNA. The amount of MotB present in the 10 min T4 infection sample was estimated to be 4.8×10^4 ($\pm 9.2 \times 10^3$) MotB monomers per cell. Consequently, there is one monomer per ∼95 bp of DNA.

Atomic force microscopy

Imaging was performed on a Multimode-8 AFM with Nanoscope-V controller (Bruker-nano, Santa Barbara, CA, USA) using the 'tapping' mode imaging modality. Silicon probes with nominal stiffness of 2.8 nN/nm and resonance frequency of 75kHz (FESP-V2, Bruker-nano, Camarillo, CA, USA) were used throughout. All solutions were assembled in AFM buffer containing 10 mM Tris–HCl (pH 7.4), 50 mM KCl and 2 mM $MgCl_2$. Reactant concentrations were adjusted to ratios of one protein (MotB-His, H-NS or both) monomer per the indicated number of bp of DNA. For controls and protein samples, solutions also contained 1/5th volume of both H-NS and MotB storage buffers (final concentration of 20% glycerol). Components were mixed at room temperature and incubated for at least 1h before dilution to approximately 0.5 nM of DNA for AFM sample preparation. A portion from each such dilution (5μl) was deposited to a freshly cleaved mica substrate that had been modified using aminopropyl-silatrane (APS) solution as described previously (42). APS treatment renders the surface hydrophobic and slightly positively charged. This is known to facilitate DNA attachment to the mica substrate for AFM imaging. Images were acquired at high resolution (∼1 nm/pixel) and processed using the software provided with the AFM instrument (Nanoscope Analysis, v1.9).

Molecular volumes for MotB-His and H-NS were computed using ImageJ (NIH) and custom-built Matlab code (The MathWorks, Inc., Natick, MA, USA). Volume histograms and plots were constructed using Origin (Originlab Corp., Northampton, MA, USA). The expected molecular volumes for monomers were estimated assuming typical ∼35% volume hydration of the monomers in ambient air using the formula: $V_{exp} = M_{prot}(V_{pr} + d^*V_w)/N_a$, where M_{prot} is the known molecular mass of the protein monomer, V_{pr} and V_w are the partial specific volumes of the protein (typically, 0.73) and water (43), respectively, and N_a is Avogadro's number (6.022×10^{23}). The AFM molecular volumes of proteins were corrected for the finite probe size by fitting ellipsoids to the particle volumes above their half-height and using the volumes of the fitted ellipsoids as corrected estimates of the actual particle volumes.

Purification of total RNA

E. coli TOP10F' containing either pNW129 or p*motB* was streaked on 1.5% (w/v) LB plates containing 40 μg/ml

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

kanamycin, 12.5 μg/ml tetracycline, and 0.5% (w/v) glucose. Overnight cultures from single colonies were grown in LB containing 40 μg/ml kanamycin, 12.5 μg/ml tetracycline and 0.025% (w/v) glucose. The next morning inoculums were diluted to an OD_{600} of 0.1 with LB, cultures were grown to an OD_{600} of ∼0.3, and arabinose [final concentration 0.2% (w/v)] was added. At 20 min post-induction (OD_{600} ∼0.5), RNA was isolated using Method II as described (44). For infections using TOP10F' containing either pNW129 or p*motB*, at 20 min after induction of MotB synthesis, WT T4 was added to the cultures at an MOI of 10. Samples were taken at the indicated times post-infection, and RNA was isolated. In all cases, biological duplicates were obtained. Total RNA was analyzed on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) using the Agilent RNA 6000 Nano Kit to evaluate the quality of the sample.

RNA-seq

RNA-seq data were processed using 2 biological replicates as previously described (45). The cDNA library was prepared using a modified RNATagSeq workflow. Optimum fragmentation of the total RNA samples in this library was determined to be 3 min at 94°C in FastAP buffer (Thermo Fisher Scientific, Waltham, MA, USA). The cDNA library was run on a Bioanalyzer using the Agilent High Sensitivity DNA Kit (Agilent, Santa Clara, CA, USA) to evaluate the quality of the library. The concentration of the cDNA library was determined by qPCR using the KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, MA, USA) and CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). Sequencing was performed by the NIDDK Genomics Core facility using a MiSeq system with the single-end 50 bp Sequencing Kit (Illumina, San Diego, CA, USA). *E. coli* str. K-12 substr. MG1655 (NC_000913.3) was used as the reference genome for the host, and the bacteriophage T4 reference genome (NC_000866.4) was used for T4. The CRISPR1 and CRISPR2 loci in MG1655 were found by looking for the direct repeats located in NCBI reference sequence NC_000913.3 between genomic positions 2, 877,701 and 2,878,463 [genes *iap* (b2753) and *cas2* (b2754)] and between positions 2,904,014 and 2,904,407 [genes *ygcE* (b2776) and *ygcF* (b2777)], respectively. Differential expression between conditions was represented as fold difference (FD), and genes with a FD ≥2 or ≤0.5, an adjusted *P*-value ≤0.05, and mean reads ≥5 were considered significant. RNA-seq data is available in the National Center for Biotechnology Information (NCBI) database (GEO# GSE172467) and in Supplementary Tables S1–S3.

Visualization of the transcriptomics data into representative categories was performed using a modified version of the EcoCyc Omics Dashboard tool (ecocyc.org) as described (46). Each dataset was imported into an individual EcoCyc 'Smart Table' and the analysis was done using the Omics Dashboard Tool (47). A list of genes from each panel in the dashboard was downloaded and Supplementary Tables S1–S3 were made using pandas v1.2.5. These tables were also used to calculate the percentage of host genes that were affected by the expression of *motB* (Table 1). In addition, these tables were used for the calculation of percentages presented in Supplementary Figure S2, which was created using GraphPad Prism v8.1.1 (GraphPad Software, San Diego, CA, USA).

Fluorescence microscopy

TOP10F' containing pNW129 or p*motB* was streaked on 1.5% (w/v) LB plates containing 40 μg/ml kanamycin and 0.5% (w/v) glucose. TOP10F' containing p*eGFP* or p*motB-eGFP* was streaked on LB plates containing 100 μg/ml carbenicillin and 0.5% (w/v) glucose. SX454 containing p*eGFP* or p*motB-eGFP* was streaked on LB plates containing 100 μg/ml carbenicillin, 10 μg/ml chloramphenicol, and 0.5% (w/v) glucose. Overnight cultures were then started from a single colony in EZ Rich defined media containing 0.4% glycerol, the appropriate antibiotic(s), and 0.025% (w/v) glucose. The next morning the cultures were diluted in the same media but without glucose to an OD_{600} of ∼0.05 and then grown to an OD_{600} of ∼0.4. Arabinose (final concentration of 0.02% or 0.2%, as indicated) was added, and the culture was incubated for the indicated amount of time. Aliquots of cells (500 μl, OD_{600} of ∼0.6) were harvested by centrifugation at 12,000 × g and then stained using 0.1 ng/μl Hoechst 33342 in 1× PBS (Quality Biological, Gaithersburg, MD, USA) for 20 min at room temperature. Cells were harvested by centrifugation at 12,000 × g, the supernatant was carefully removed, and the cells were resuspended in 500 μl 1× PBS. A 5 μl aliquot was transferred to a slide and covered with a 1% agar disk containing EZ Rich defined media.

Images were collected with a Nikon TiE inverted microscope with Nikon 100×/1.4 Oil Plan Apo Ph3 DM lens, Lumencor sola light engine (Beaverton, OR, USA) and ImageEM EMCCD camera (Hamamatsu, Japan) and were analyzed using ImageJ with the plugins available through the FIJI distribution (http://fiji.sc/Downloads). Cell length was measured manually using the freehand line function in ImageJ for cells that had a signal corresponding to Hoechst or Hoechst and GFP for pNW129 and p*eGFP* samples, respectively. For the analysis of TOP10F' cells containing p*eGFP* or p*motB-eGFP*, the number of cells with 'No Hoechst or GFP', 'Hoechst, no GFP', 'GFP, no Hoechst', 'cytosolic GFP' and 'GFP localized to nucleoid' (co-localized with Hoechst) was manually counted using the counter plugin for ImageJ. For each condition, cell length was measured for a minimum of 50 cells.

For p*eGFP* samples using a 20 min induction of MotB using 0.2% arabinose (w/v), the percentage of cells stained with Hoechst and GFP ranged from 62% to 95%; for the cells induced with 0.02% arabinose, the percentage of cells stained with Hoechst and GFP ranged from 20% to 40% and from 80% to 95% for the 10 and 15 min inductions, respectively. For the GFP localization analysis with the 20 min induction at 0.2% arabinose (w/v), a minimum of 120 cells were counted; for the 10 and 15 min induction at 0.02% arabinose (w/v), 50 cells and 90 cells were counted for the 10 min and 15 min inductions, respectively. Three independent replicates were performed for all conditions.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

Table 1. Correlation between host genes upregulated by overexpression of *motB* and host genes previously identified as being part of the *hns* regulon

Ratio of (genes upregulated by *motB*)/(genes dysregulated by indicated condition)				
Comparison[a]	***Δhns*[b]**	***Δhns/stpA*[b]**	**T7 5.5[c]**	**C-truncated H-NS[c]**
motB/vector no infection	78% (134/172)	61% (356/583)	54% (127/237)	57% (175/309)
motB/vector 5 min infection	81% (140/172)	73% (423/583)	59% (139/237)	64% (198/309)
motB/vector 10 min infection	81% (140/172)	72% (417/583)	58% (137/237)	63% (195/309)

Ratio of (genes dysregulated by indicated condition)/(genes upregulated by *motB*)					
Comparison[a]	**Total[d]**	***Δhns*[b]**	***Δhns/stpA*[b]**	**T7 5.5[c]**	**C-truncated H-NS[c]**
motB/vector no infection	74% (400/539)	25% (134/539)	66% (356/539)	24% (127/539)	32% (175/539)
motB/vector 5 min infection	69% (475/687)	20% (140/687)	62% (423/687)	20% (139/687)	29% (198/687)
motB/vector 10 min infection	70% (468/672)	21% (140/672)	62% (417/672)	20% (137/672)	29% (195/672)

[a]Determined from the RNA-seq data (Supplementary Tables S1–S3) obtained after overexpression of *motB* (or the vector control) in TOP10F' cells without infection or 5 or 10 min post-infection with WT T4 as detailed in Materials and Methods.
[b]Using either the set of genes dysregulated by *Δhns* or *Δhns/stpA*, as indicated (Ueda *et al.*, 2013).
[c]Using either the set of genes dysregulated by the presence of the T7 5.5 protein or by a C-terminally truncated H-NS, as indicated (Ali *et al.*, 2011).
[d]Using the genes defined by all the conditions: *Δhns*, *Δhns/stpA*, T7 5.5 protein and C-terminally truncated H-NS.

Flow cytometry

Cultures of TOP10F' containing pNW129 or p*motB* were obtained as described above. At an OD_{600} of ∼0.3, 0.2% (w/v) arabinose (final concentration) was added, and cultures incubated for 1 h. Aliquots were removed for subsequent SDS-PAGE to monitor MotB production, and flow cytometry analyses were performed as described (48) with an Apogee-A40 flow cytometer (Apogee Flow Systems, UK) equipped with an Ar laser emitting at 405 nm. Cells were fixed with ice-cold ethanol (final concentration 70%) for more than 12 h and resuspended in a buffer containing 100 mM Tris–Cl (pH 7.4), 25 mM $MgCl_2$, 100 mM NaCl, mithramycin (50 μg/ml), and ethidium bromide (25 μg/ml), followed by incubation for 30 min at room temperature in the dark. Cell size and DNA content were determined by measuring forward scattering (at LS1 detector) and fluorescence intensity (at FL2 channel), respectively. All the solutions used were filtered through 0.22 μm filters.

RESULTS

The majority of *E. coli* host genes affected by expression of *motB* are within the *hns* regulon

Although a *motB* knockdown does not affect T4 gene expression (3), heterologous overexpression of *motB* results in larger WT T4 plaques, when using either the *E. coli* B strain BL21(DE3) (3) or the *E. coli* K12 strain TOP10F' (Supplementary Figure S1). In addition, this expression is toxic, resulting in cell death for BL21(DE3) and growth arrest for TOP10F', starting about 1 h after expression (3). To investigate how gene expression was affected by overexpression of *motB* in an *E. coli* K12 strain, we isolated RNA from TOP10F' 20 min after induction of MotB synthesis from p*motB* using 0.2% arabinose. At this time point, the cells expressing *motB* are still growing similarly to those containing the vector (3).

Previous work had indicated that during infection the level of MotB is high, and unlike many early proteins, this level does not diminish as infection proceeds (49,50). However, the amount of MotB in an infected cell has not previously been quantified. To compare the level of MotB in the infected cell to that in this overexpression system, we constructed a T4 phage containing *motB* with a C-terminal *his* tag. We then infected TOP10F' with T4-*motB-his* and estimated the level of MotB by Western analysis using purified MotB-His as the standard. From this analysis (Figure 1A, B), we found that the level of MotB during infection is indeed quite high with ∼4×10^4 MotB molecules/cell and ∼5×10^4 MotB molecules/cell at 5 and 10 min post-infection, respectively. This analysis estimated that the level of MotB present in the TOP10F' cells after *motB* overexpression was about 4-fold higher (Figure 1B).

Under these conditions of overexpression, 539 TOP10F' genes were significantly up-regulated with a FD ≥2, adjusted *P*-value ≤0.05, and mean reads ≥5 (Supplementary Table S1). Although it is formally possible that MotB production downregulated the other ∼3860 *E. coli* genes, the overall level of host proteins, as judged by SDS-PAGE, was similar with or without MotB production. This suggests that *motB* expression is not operating by this mechanism.

To investigate how these genes correlated with the *hns* regulon, we compared the genes to those previously reported to be dysregulated in a *Δhns* or a *Δhns/stpA* strain (51) or in the presence of the T7 5.5 protein or a C-terminally truncated H-NS (24), which are known to interfere with H-NS regulation. As seen in Table 1, there is a strong correlation between MotB up-regulated genes and genes affected by H-NS: 78% of the genes altered by the *Δhns* are upregulated by MotB and 74% of the genes, which were up-regulated by MotB, are present in at least one of the 4 categories that have been used to identify genes affected by *hns* (a deletion of *hns*, a double deletion of *hns/stpA*, or dysregulation by the presence of T7 5.5 protein or the C-terminally truncated H-NS).

Only three genes [*cueO*, which is involved in Cu homeostasis; *marC*, which encodes an inner membrane protein; and *yddW* (also known as *digH*), which encodes a peptidoglycan specific hydrolase] were significantly down-regulated (FD of ≤ 0.5, adjusted *P*-value ≤ 0.05, and mean reads ≥ 5) (Supplementary Table S1). None of these genes are known to be affected by H-NS, and to our knowledge, none are known to be involved in T4 biology.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

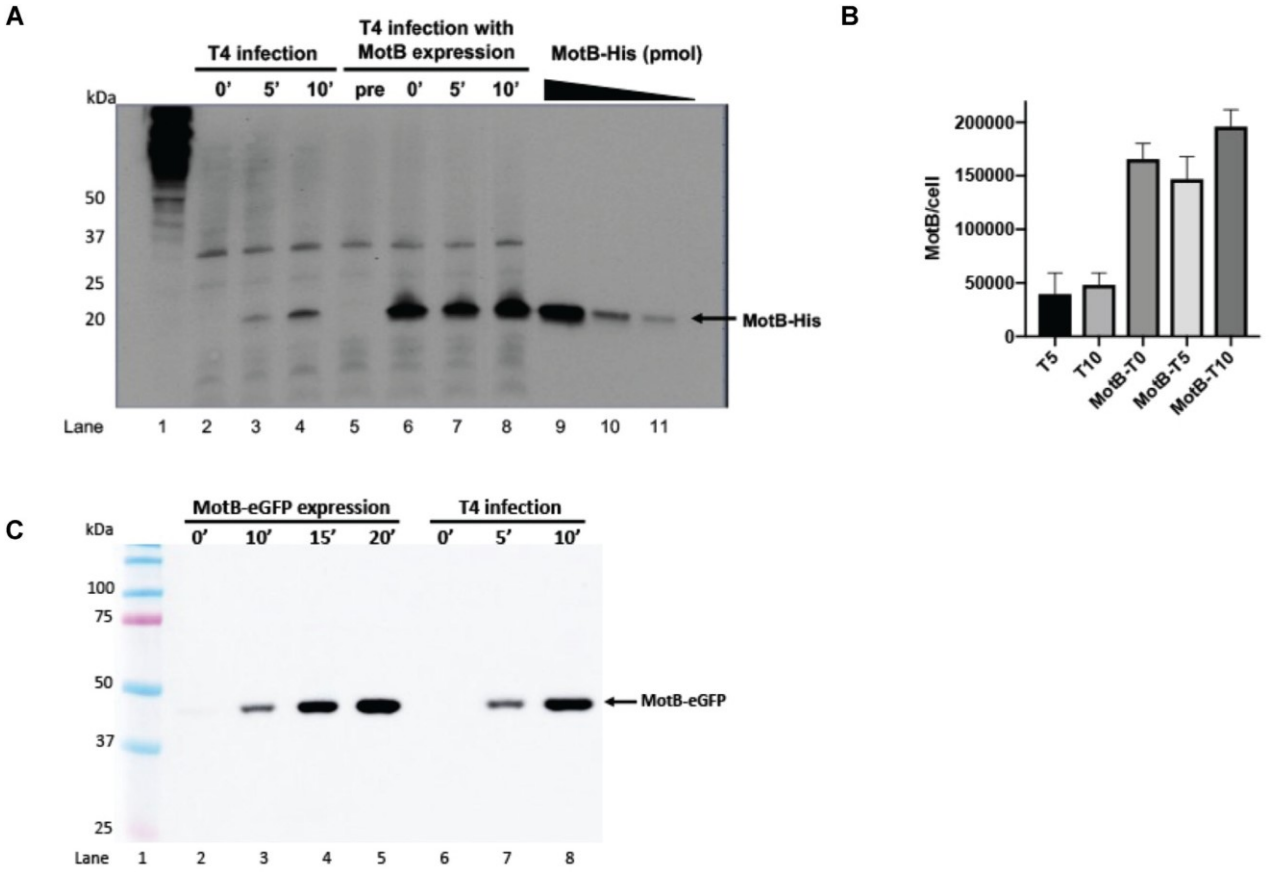

Figure 1. Quantification of MotB-His in TOP10F' and level of MotB-eGFP in SX454 after expression of *motB* compared to T4 infection. (**A**) Representative Western blot using anti-his antibody of SDS-PAGE gel shows: lane 1, molecular weight ladder with positions of the proteins of the indicated sizes (in kDa) marked; lanes 2–4, whole-cell extracts of TOP10F' at time 0, 5 and 10 min after infection with T4-*motB-his*, respectively; lane 5, whole-cell extracts of TOP10F' containing p*motB-his* pre-induction; lanes 6–8, whole-cell extracts at time 0, 5, and 10 min post-infection of TOP10F' containing p*motB-his* with T4-*motB-his* after a 20 min MotB induction; lanes 9-11, purified MotB-His protein at amounts of 3.2, 0.63 and 0.32 pmol, respectively. (**B**) Bar graph quantifying the amount of MotB-His/cell: T5, 5 min after infection with T4-*motB-his*; T10, 10 min after infection with T4-*motB-his*, MotB-T0, after 20 min induction of MotB synthesis using 0.2% arabinose; MotB-T5 and MotB-T10, cells were infected with T4-*motB-his* for 5 or 10 min, respectively. Error bars represent standard deviation ($n = 3$). (**C**) Representative Western blot using anti-GFP antibody of SDS-PAGE gel shows: lane 1, molecular weight ladder with the size of the proteins (in kDa) indicated; lanes 2-5, whole-cell extracts of SX454 containing p*motB-eGFP* preinduction of MotB synthesis or after induction with 0.02% arabinose for 10, 15, and 20 min, respectively; lanes 6-7, whole-cell extracts of SX454 0, 5 and 10 min post-infection with T4-*motB-eGFP*. Arrows indicate MotB-His (A) and MotB-eGFP (C).

To investigate if and how MotB affects host and T4 gene expression during phage infection of TOP10F', we performed RNA-seq on RNA isolated from the strain containing either pNW129 or p*motB*. MotB synthesis was induced the same way, but cells were infected with WT T4 for 5 or 10 min, times that correspond to high levels of middle and late T4 gene expression, respectively (52). In these cases, 687 and 672 host genes were up-regulated, while only 17 genes and 34 genes were down-regulated by overexpression of *motB* at 5 and 10 min post-infection, respectively (Supplementary Tables S2 and S3). Again, there was a strong correlation between MotB up-regulated genes and genes identified as being affected by H-NS. At either 5 or 10 min after infection, 81% of genes affected by a *b.*hns were up-regulated by MotB, and ∼70% of the genes up-regulated by MotB were present in at least one of the four categories that have been used to identify genes affected by *hns* (Table 1). Thus, when *motB* is overexpressed in TOP10F' either in the uninfected cell or during infection, the majority of affected genes involve those under H-NS regulation. Among the down-regulated genes, nine genes were affected at both times of infection. Only one, *ompF*, which was down at 10 min post-infection, is known to be affected by H-NS.

The effects of MotB on T4 gene expression were much less dramatic than those on the host (Supplementary Tables S2 and S3). At 10 min post-infection, no T4 genes were significantly affected. However, at 5 min, when late gene expression is just beginning, we observed higher expression of 33 (∼60%) of T4 late genes. In addition, one early gene *repEB*, which encodes a protein needed for efficient initiation at T4 origins of replication (29–31), was up-regulated ∼5-fold. Since both T4 replication and late gene expression begin around 5 min (52) and late transcription requires DNA synthesis (53), these results are consistent with the idea that a higher level of RepEB results in earlier replication that then leads to earlier synthesis of late RNA. However, there is no significant effect on the level of any late RNAs at 10 min. These results are in agreement with our

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

previous studies with the *E. coli* B strain BL21(DE3) (3), where we observed that a T4 *motB*^{am} infection resulted in delayed expression of a few T4 late genes at 5 min post-infection, but at 10 min, again no T4 genes were affected. However, in this analysis, we did not obtain reads from *repEB*.

To visualize the host transcriptomic data in more detail, we used a modified version of the 'Pathway Tools Omics Dashboard' (ecocyc.org) (Supplementary Figure S2) in which a series of panels present genes broadly related to cellular systems. We considered an effect of > 50% of a particular pathway as possibly significant. Given this criteria, in the uninfected cell, *motB* overexpression affected genes involved in two pathways: cell killing (2 out of 4, 50%) and pilus (17 out of 32, 53%). During T4 infection overexpression of *motB* also affected these particular pathways (cell killing, 2 out of 4, 50% at either 5 or 10 min after infection; pilus, 24 out of 32 at 5 min, 75%; 25 out of 32 at 10 min, 78%). In addition, two other pathways were up-regulated: adhesin (33 out of 57 at 5 min, 58%; 34 out of 57 at 10 min, 60%) and LPS metabolism (40 out of 80 at 5 min, 50%; 40 out of 80 at 10 min, 50%). It is known that T4 can use LPS as a receptor in addition to OmpC, (54,55), suggesting that upregulation of LPS might be related to T4 adsorption. However, it is unclear why affecting the other pathways might be advantageous for the phage, and in all these cases as well as in the other affected pathways, there was again a strong correlation between the affected genes and genes within the *hns* regulon (Supplementary Tables S1-S3). These results suggest that perhaps dysregulation of H-NS control might be the primary correlation here rather than the particular pathway itself. We conclude that MotB does not simply target a specific host operon or pathway, but instead results in global changes that predominantly include the *hns* regulon.

Finally, as H-NS paralogs have been implicated in altering CRISPR-Cas (56), we also investigated whether any genes involved in this system were affected by overexpression of *motB*. As seen in Supplementary Tables S1-S3, we observed a significant up-regulation of the entire *E. coli* K12 Type 1-E *cas* operon [*cas3*, *casABCDE, cas1* and *cas2*; (57–59)] at 5 min after infection (Supplementary Table S2) and up-regulation of all of these genes except *cas3* either in the absence of infection (Supplementary Table S1) or 10 min after infection (Supplementary Table S3). In addition, the CRISPR 1 locus [also known as CRISPR 2.1, located between *iap* (gene ID: b2753) and *cas2* (gene ID: b2754) (57,58)] was up-regulated after 10 min of infection (Supplementary Table S3). This *cas* operon is part of the H-NS regulon as it has previously been shown that the *cas* genes are repressed by H-NS (58,60). Furthermore, *casABCDE* is up-regulated in both the *b.hns* and *b.hns/stpA* strains, while *cas1* and *cas2* are up-regulated in the *b.hns/stpA* strain (51).

AFM shows that MotB compacts DNA in the presence and absence of H-NS

Previous work has shown that H-NS can bind to DNA in two modes, linear or bridged, and these modes are dependent on protein concentration and buffer conditions (14–16). In the linear binding mode, H-NS binds and oligomerizes along a segment of DNA, whereas in the bridged mode, H-NS binds two distant segments or separate pieces of DNA, 'zippering' them together. To investigate the binding by H-NS and MotB under our conditions, we used AFM, a technique that has been used previously to assess H-NS binding (5,10,11,13).

To first check our purified H-NS, we observed its binding to a 6 kbp plasmid (pMLH07) (35), which is known to contain H-NS binding sites (61). AFM revealed H-NS/DNA interactions that were consistent with both the linear and bridged forms, as seen in reports of H-NS binding to DNA (10,13) (Supplementary Figure S3A, white circles indicate binding that appears to be bridged or linear). The directionality of the DNA in these images arises from the direction of flow in the preparation of the samples. When deposited on the APS-treated mica, the supercoiled DNA alone exhibited some compaction, most likely from the Mg^{++} present in the buffer, but the effect of H-NS on the DNA is still observable. In addition, in other work (62) we have demonstrated that H-NS generates a DNase I protection pattern on *E. coli proV* DNA that is similar to what has been previously reported (63).

We then visualized the binding of MotB-His, H-NS, or both proteins together to a 1620 bp fragment that was either unmodified or modified by the presence of [5-glucosylated, hydroxymethyl cytosine (GHmeC)]. This modification was generated by using T4 13-glucosyltransferase to glucosylate DNA that had been synthesized using hydroxymethyl C. In the case of T4 DNA itself, the hydroxymethyl C's are modified by this transferase or by a T4-encoded a-glucosyl transferase, resulting in a mixture of stereoisomers (64). Consequently, we also tested WT T4 gDNA, which contains both the a and 13 forms of glucosylation. Our previous work (3) indicated that either MotB or H-NS bind unmodified, modified, and T4 DNA. There is no appreciable difference in DNA affinity depending on whether the modifications are present. However, the type of complexes were not investigated.

In these AFM experiments, we used a ratio of protein to DNA ranging from 1 monomer to ~80–160 bp of DNA for the linear fragments and one monomer to 100 bp or one monomer per 10 bp for T4 gDNA. Based on our quantification of the amount of MotB in an infected cell (Figure 1A, B), there are $\sim 4 \times 10^4$ MotB molecules/cell and $\sim 5 \times 10^4$ MotB molecules/cell at 5 and 10 min post-infection, respectively. Consequently, this would yield a ratio of MotB:DNA of 1 MotB monomer/120 bp DNA at 5 min and one MotB monomer/95 bp at 10 min. Given that over the course of the infection ~50% of the host DNA is degraded, but also at 5 min post-infection T4 DNA begins to be replicated, this is a rough estimate. Nonetheless, we conclude that the MotB/bp ratio used for the AFM experiments is in the range of that in a T4 infection.

In the AFM analysis, when no protein was present, the unmodified DNA appeared linear, as expected; despite the high density of DNA on the APS mica substrate, individual strands were discernible (Figure 2A). However, in the presence of MotB the DNA fragments became significantly compacted (Figure 2B). Unbound DNA fragments were also observed, suggesting that MotB binding was cooperative. Due to the extent of compaction, it was not possible

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

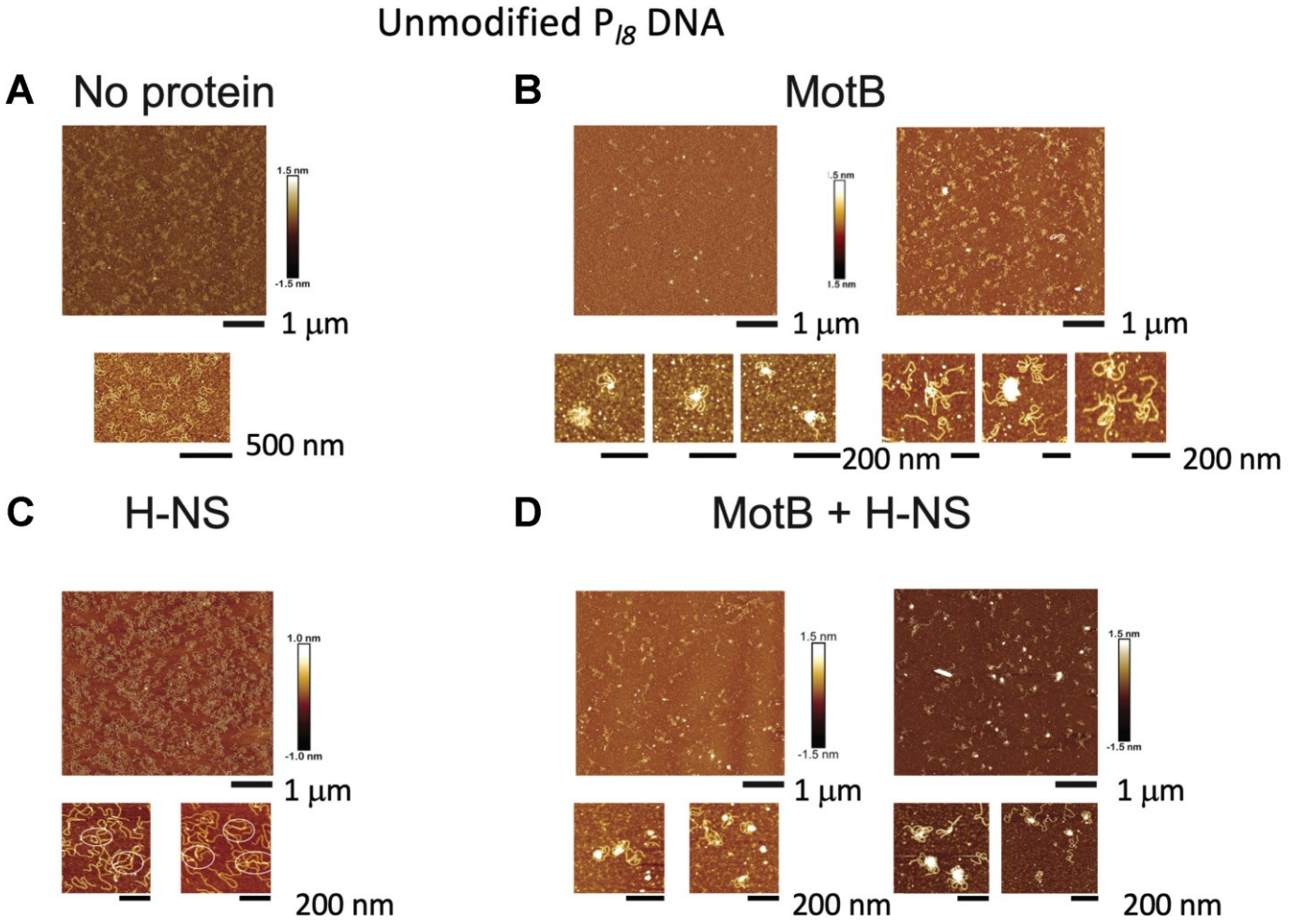

Figure 2. MotB compacts unmodified DNA in the presence and absence of H-NS. Representative AFM images of 1.6 kbp linear, unmodified P_{l8} DNA alone (**A**), with MotB (**B**), with H-NS (**C**), or with both proteins together (**D**). In each case, a larger field image is shown on top and various magnified samples are shown underneath; the height (*z*) scale is shown and lengths are indicated. In Panel C, white circles indicate regions of DNA that appear to be consistent with H-NS binding the DNA, as seen in previous reports (13,33). The ratio of protein monomer to DNA bp was 1:80–160 bp.

to quantify the number of DNA fragments present in each complex or the fraction of DNA bound by MotB. However, it was clear that these large complexes contained more than one DNA fragment and multiple MotB monomers.

The large, compacted complexes observed with MotB were quite distinct from complexes formed by H-NS (10,13) as well as AFM complexes that have been reported for other NAPs: Fis (65), HU (66), and IHF (67). However, they did resemble compacted DNA complexes that have previously been observed with the *E. coli* NAP-like proteins, DNA-binding protein from starved cells (Dps) (26,27) and curved DNA-binding protein A (CbpA), which are involved in nucleoid compaction during the stationary phase of *E. coli* growth (4), and with yeast structural maintenance of chromosome (SMC) protein cohesin (28). All of these proteins are involved in the compaction of chromosomal DNA.

As expected, in the presence of H-NS alone, we observed binding of the protein to the DNA. Several of these interactions (marked in Figure 2C) appeared to represent linear or bridged forms as have been seen in previous reports (10,13); however, we cannot definitively conclude that H-NS binding is indeed in those forms from this analysis alone. In addition, to confirm that the presence of H-NS did indeed result in more 'zippered' DNAs, we used images of DNA ± H-NS with similar DNA density (35−40 DNA molecules/µm^2), to measure lengths of bridged sections of parallel DNA stands, using the thickness of a single DNA chain as a threshold. We then graphed the data as violin plots with the size of the DNA vs. the 'population size', i.e., the number of occurrences in which zippered segments were of the size indicated on the y-axis (Supplementary Figure S3B). In the absence of H-NS, we could only find a few segments above the threshold, hence the low 'N' value for the analyses with DNA alone. However, in the presence of H-NS, we observed significantly more and longer zippered segments, consistent with H-NS-binding to the DNA.

When both MotB and H-NS were present, the large, compacted DNA-protein complexes, like those seen with MotB alone, were observed rather than the complexes seen with H-NS alone. Based on their size, some of these complexes appeared to contain multiple DNA fragments and multiple protein monomers (Figure 2D). However, the complexes

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

were heterogeneous. Some DNA fragments contained no detectible protein, others contained protein together with a moderately-sized node at the center of the compacted DNA, and others were found with protein within very large, compacted species. However, despite the heterogeneity, it was clear that in the presence of both MotB and H-NS the complexes resembled those formed in the presence of MotB alone rather than those formed by H-NS alone. It was not possible from this analysis to determine if MotB alone or both proteins were present together on the DNA. Similar results were obtained using either the modified DNA fragment (Figure 3A-D) or T4 WT gDNA (Figure 3E-H, Supplementary Figure S4A).

Our results indicated that both H-NS and MotB bound to both T4 gDNA and unmodified host DNA, forming complexes with similar appearances. Thus, the phage modification does not appear to affect the complexes formed by either protein. However, with either unmodified or modified DNA, MotB compacts the DNA in a manner that is distinct from that of H-NS, and this compaction does not appear to be altered, as can be discerned by AFM, in the presence of H-NS.

AFM analyses of MotB protein indicate 2 species, consistent with a mixture of monomers and dimers, and show no evidence of MotB aggregation

Previous work has used AFM analyses to estimate the presence of protein monomers and multimers in solution (27,68). With MotB alone, which is 17 kDa, AFM yielded 2 species, whose estimated sizes were consistent with a mixture of monomers and dimers Supplementary Figure S4B, top panel). Likewise, H-NS (15.5 kDa) also appeared to be in monomer and dimer forms (Supplementary Figure S4B, middle panel). Monomer, dimer, and higher oligomeric states for H-NS have been reported previously (69–72). Images of both proteins together gave similar results, although the range of sizes observed with MotB alone narrowed in the presence of H-NS (Supplementary Figure S4B, bottom panel). The AFM issue of finite probe shape convolution affects the shape of small protein particles and increases the probability that some particles that appear as dimers are merely the result of random landing of large numbers of molecules on the mica substrate. Consequently, we cannot definitively conclude from this analysis alone that these species truly represent monomers and dimers. However, it is clear that under our conditions neither protein alone or together forms large multimers or aggregates in the absence of DNA. Whether large multimers of MotB form under different conditions remains to be determined.

MotB co-localizes with compacted DNA *in vivo*

The AFM results suggested that MotB binding to DNA results in significant DNA compaction *in vitro*. To ask whether compaction is also seen *in vivo*, we first employed fluorescent microscopy using exponentially growing TOP10F'. To simultaneously observe MotB and the DNA, we generated a plasmid encoding *motB* with a C-terminal *eGFP* fusion (p*motB-eGFP*). Protein induction assays indicated that the MotB-eGFP fusion protein is also toxic when produced in the TOP10F' cells (Figure 4A). However, at times less than 1 h after induction with 0.2% arabinose, the growth of cells expressing *motB-eGFP* is similar to that of cells expressing the vector.

As expected in cells containing the vector, the DNA, as visualized by Hoechst staining, appeared to fill the nucleoid, which took up most of the cell, and the eGFP was present throughout the cytoplasm (Figure 4B, top). In contrast, in the presence of MotB-eGFP, the DNA was compacted, and MotB-eGFP localized with the DNA (Figure 4B, bottom), a result consistent with the idea that overexpression of *motB* results in compaction of the DNA *in vivo*. An analysis of hundreds of cells (Figure 4C) indicated that in nearly every cell in which MotB-eGFP was observed, it was found within the nucleoid.

To ensure that these results were not affected by the eGFP fusion to MotB, we also compared TOP10F' cells expressing *motB* without the fusion (p*motB*) to those with the vector control. As in the experiment of Figure 4, cells were analyzed 1 h after induction (Supplementary Figure S1B), and we observed compacted DNA in the presence of MotB (Supplementary Figure S1C). We also used the same induction conditions (20 min) that were used to obtain the RNA for the overexpression analyses described above and again we observed compacted nucleoids containing MotB (Supplementary Figure S1D).

MotB and H-NS co-localize with compacted DNA under biologically relevant conditions

The previous results indicated that under these conditions, overexpression of *motB* in TOP10F' resulted in significant compaction of the DNA. However, it was important to determine whether this effect was observed using a biologically relevant level of MotB. Thus, we constructed a T4 phage containing *motB* with the C-terminal *eGFP* tag. We then used the *E. coli* K12 strain SX454, which encodes a chromosomal copy of *hns* with a C-terminal *mCherry* fusion in order to localize H-NS and MotB within the cell. This strain has been used previously to observe the localization of H-NS within the nucleoid (9,32,33) and to show that the presence of the tagged H-NS does not alter the growth rate relative to the untagged strain (32).

We first determined the levels of MotB-eGFP produced in SX454 after various times of *motB* induction relative to the levels expressed during a T4 infection of this strain after 5 and 10 min, times which correspond to high levels of T4 middle and late gene expression, respectively (52) (Figure 1C). This analysis indicated that the level of MotB present after an induction of 10 min using an arabinose concentration of 0.02% was similar to that present in T4 infected cells after 5 min, while the level present after a 15 min induction was similar to that present in T4 infected cells after 10 min. Consequently, we performed fluorescence microscopy to localize MotB-eGFP, H-NS-mCherry, and the DNA in SX454 with levels of MotB similar to those after 5 min and 10 min of infection (Figure 5A and 5B, respectively).

In the absence of *motB* expression, eGFP was observed throughout the cell, while H-NS-mCherry protein was seen with the DNA throughout the nucleoid. This pattern is typical of what has been previously reported for H-NS *in vivo*

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

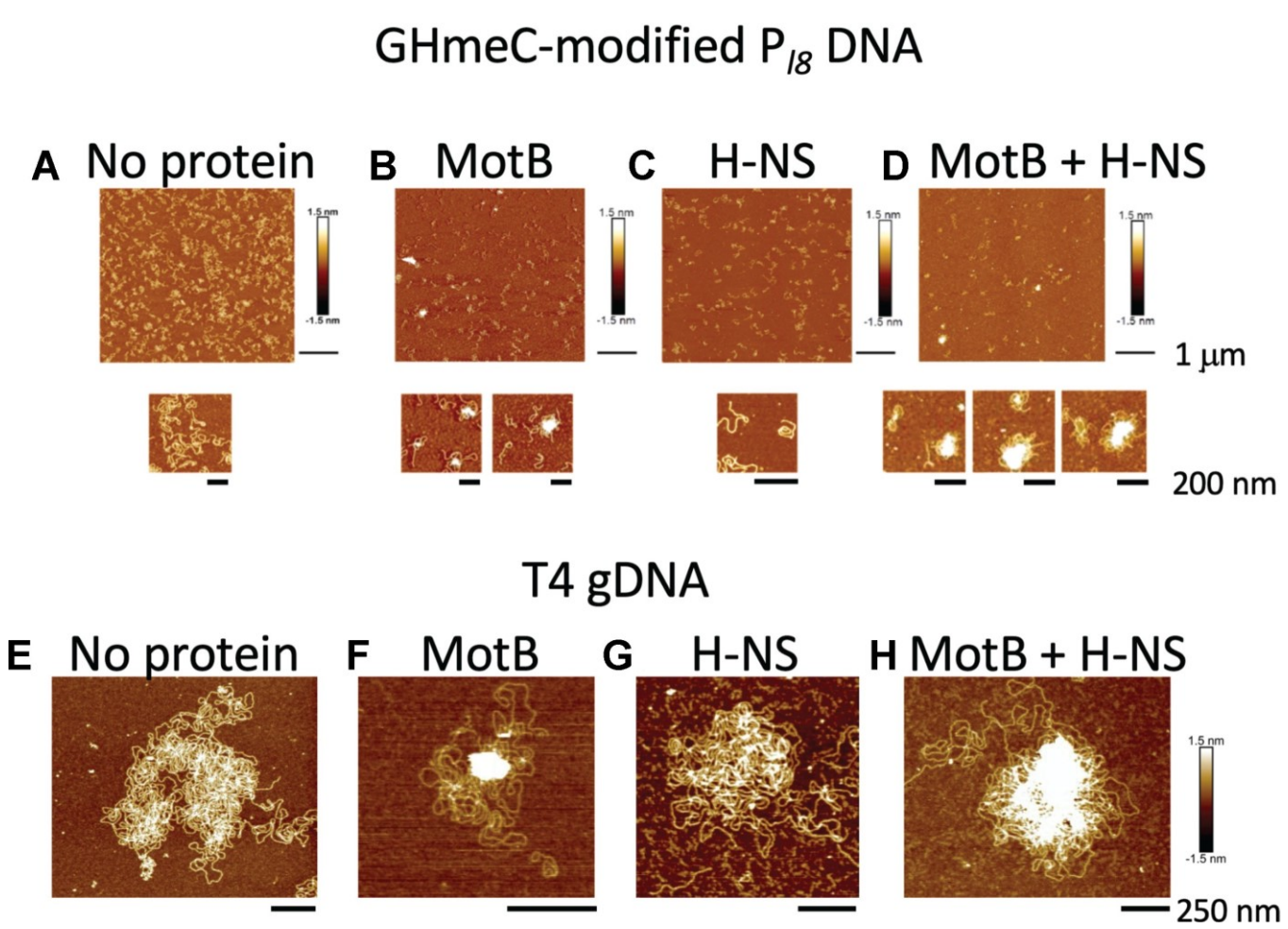

Figure 3. MotB compacts modified and T4 gDNA in the presence and absence of H-NS. Representative AFM images of GHmeC-modified 1.6 kbp linear P_{I8} DNA alone (**A**), with MotB (**B**), with H-NS (**C**), or with both proteins together (**D**) and T4 gDNA alone (**E**), with MotB (**F**), with H-NS (**G**), or with both proteins together (**H**). For A−D, a larger field image is shown on top and magnified samples are shown underneath; the height (*z*) scale is shown to the right of each image in A−D and to the right of H for all the images in E−H. Lengths are indicated. In the enlarged image in (C), the intensity of the signal was consistent with almost all of the DNA being coated with H-NS. For A−D, the ratio of protein monomer to DNA bp was 1:80−160 bp; for E-H, the ratio was ∼1:100.

and shows the binding/compaction performed by H-NS and other NAPs in the *E. coli* nucleoid (9,33,73). However, in the presence of MotB-eGFP produced after an induction of either 10 min (Figure 5A, large field Supplementary Figure S5A) or 15 min (Figure 5B, large field Supplementary Figure S5B), a different picture emerged. In many cells, the DNA was significantly more compacted and co-localized with both MotB-eGFP and H-NS-mCherry.

As more cells showed the DNA compaction with the 15 min induction of MotB synthesis (Figure 5B) than with the 10 min induction (Figure 5A), we also investigated how an even higher level of MotB would affect the DNA by inducing MotB production for 20 min with 10-fold more arabinose (0.2%). In this case, the results mirrored what was observed before except that DNA compaction was seen in every cell (Figure 6). Taken together, our results indicate that MotB compacts the DNA within the nucleoid and that MotB and H-NS co-localize with this DNA. The presence of H-NS with the DNA suggests that it remains within the nucleoid in the presence of MotB.

Expression of *motB* in TOP10F' cells results in significant cell lengthening

In addition to the DNA compaction observed after induction of MotB in TOP10F' and SX454 cells, fluorescent microscopy also indicated that the cells expressing high levels of *motB* or *motB-eGFP* appeared to be longer than those containing the vector or expressing *eGFP* alone (Figure 4D, Supplementary Figures S6A, S1E, F, S6D, E).

To investigate this further, we employed flow cytometry to determine the size of TOP10F' *E. coli* containing either the vector or the *motB* expression plasmid (p*motB*) 1 h after MotB induction. At this time, the level of MotB was checked by SDS-PAGE to confirm the presence of MotB within the cells containing p*motB* (Figure 7A). Light scattering histograms showed that the size distribution of TOP10F' *E. coli* cells expressing *motB* cells (purple filled) was shifted to the right compared to vector (black line) (Figure 7B), again indicating cell lengthening with MotB overexpression. Furthermore, cells expressing *motB*, which were

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

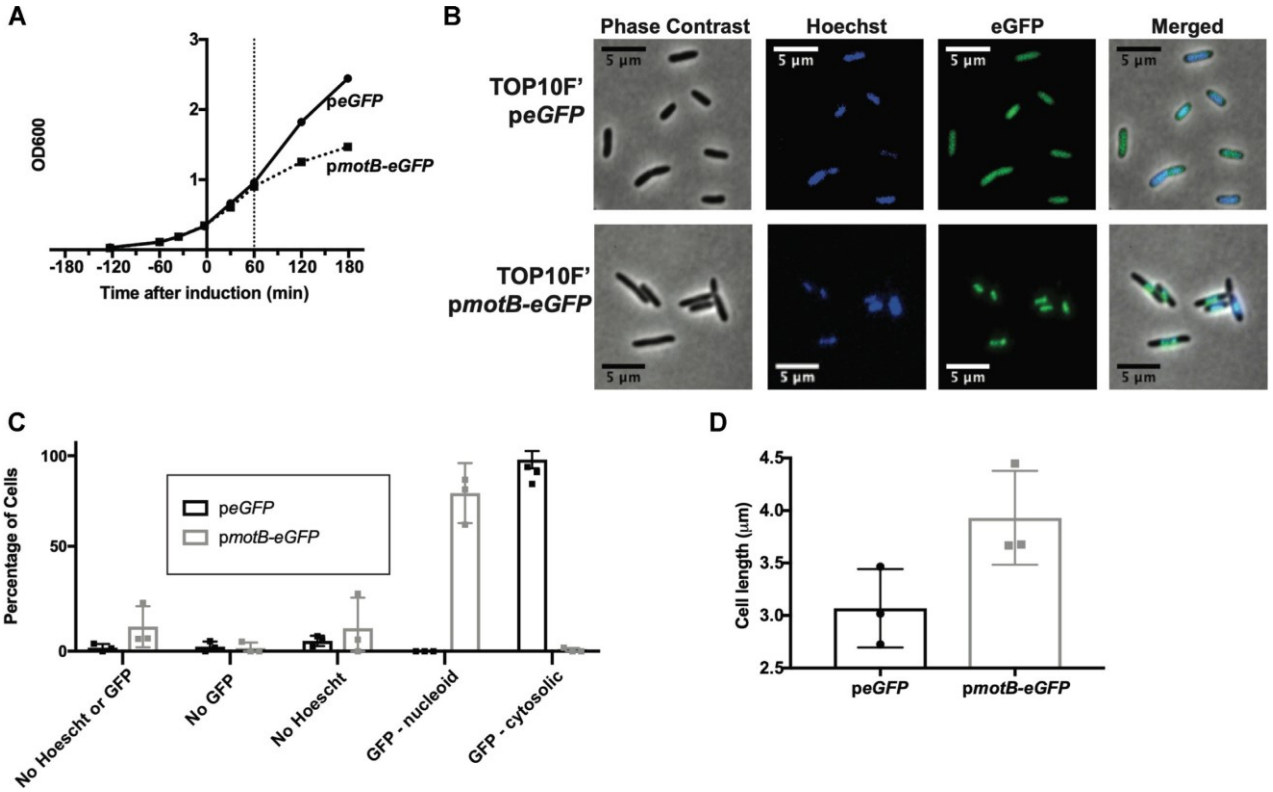

Figure 4. MotB-eGFP co-localizes with compacted DNA *in vivo*. (**A**) Representative growth curves for TOP10F' cells containing the vector (p*eGFP*, filled circle with solid line) or plasmid producing MotB-eGFP (p*motB-eGFP*, filled squares with dashed line). (**B**) Cell images obtained 1 h after induction of plasmids producing eGFP or MotB-eGFP. Cells were visualized by phase-contrast, DNA by blue Hoechst, and eGFP and MotB-eGFP by green fluorescence. (**C**) Quantification of eGFP or MotB-eGFP localization in the nucleoid or cytoplasm based on counting of a minimum of 120 cells. An average of 8.8% and 24.1% of cells did not contain signals in one or both channels for the eGFP control and MotB-eGFP cells, respectively (see Supplementary Figure S6B). (**D**) Bar graph quantifying the length of TOP10F' cells containing eGFP or MotB-eGFP at 1 h post induction indicated in panel A by the vertical dotted line. At least 90 cells containing a signal for both Hoechst and eGFP were measured for three independent replicates of each condition (see Supplementary Figure S6A).

increased in cell length, showed comparatively less DNA content than cells with the vector (Figure 7C red versus black lines, respectively) as measured by levels of ethidium bromide. These results suggest that the presence of MotB inhibits host DNA replication, which could explain the cell elongation. However, we cannot eliminate the possibility that the DNA compaction mediated by MotB inhibits ethidium bromide uptake, resulting in the appearance of less DNA.

DISCUSSION

Phage genes, particularly those expressed early in infection, provide a large resource of functions that are needed to take over the host. Thus, characterizing these functions can provide information about novel antibacterial strategies. In T4, expression of these genes begins immediately after infection and typically continues until about 5 min. Although some of these early genes encode products essential for the next (middle) phase of T4 gene expression, the majority are not required and presumably provide functions that contribute to host takeover or optimize the host for the ensuing infection.

The T4 MotB protein is an early product that is produced abundantly. Our analyses indicate that there are ∼40,000 MotB monomers per cell at 5 min after infection, and this amount does not diminish as infection proceeds. However, the biological function(s) of MotB has not been determined. Our previous work indicated that the presence of MotB improves phage fitness and that it binds tightly, but nonspecifically, to either the unmodified DNA of the host or the T4-modified DNA (3). It is highly conserved among *Tevenvirinae*, suggesting an important conserved function(s); our *in silico* analyses have predicted that MotB contains both a KOW domain and an OB-fold (62). Both of these motifs have been associated with DNA binding function.

In the presence of DNA, MotB copurifies with the *E. coli* NAPs, H-NS and its paralog StpA (3). Bacterial NAPs are responsible for the organization of genomic DNA through the formation of high-order complexes that condense DNA, impacting major processes including replication, recombination, repair, and transcription (4–8,32,73). NAPs include sequence-specific DNA binding proteins, e.g., CbpB, DnaA, Fis, and IHF, as well as sequence-nonspecific DNA binding proteins, e.g., H-NS, HU, StpA, and IHF when it binds in a non-sequence specific mode. The highly abun-

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

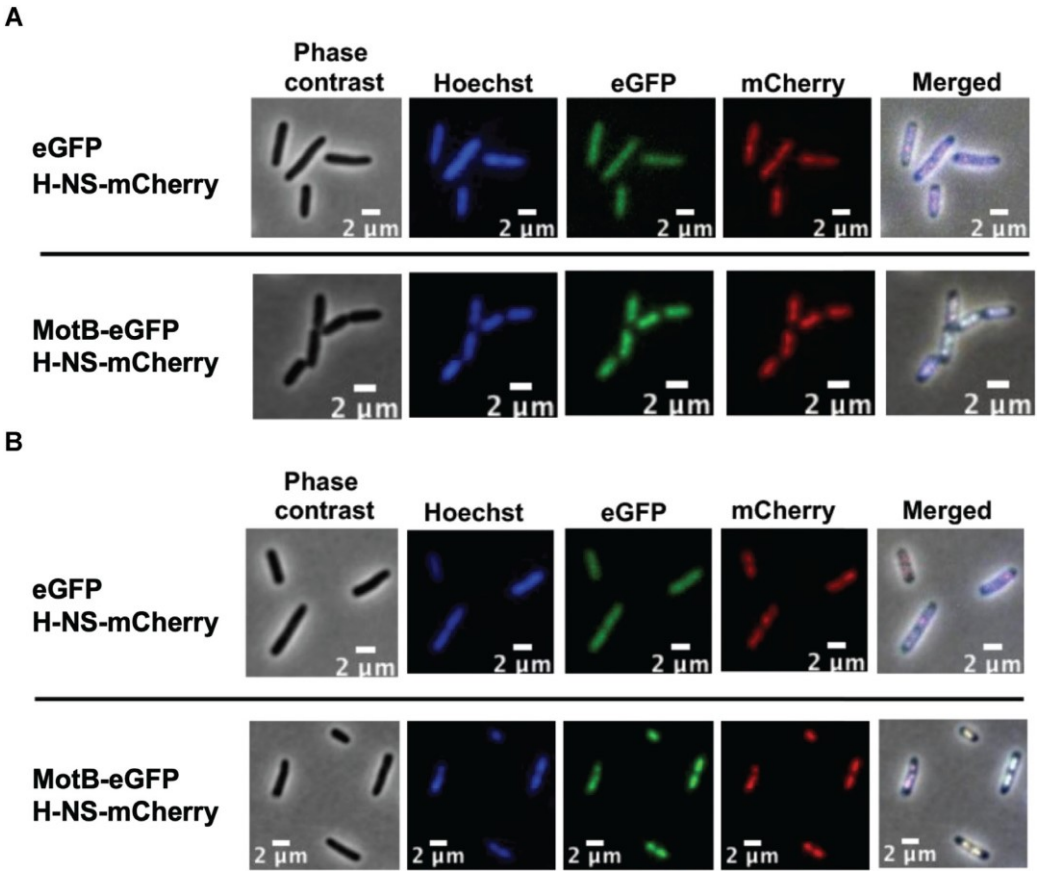

Figure 5. DNA compaction and co-localization of MotB-eGFP and H-NS-mCherry are observed *in vivo* with biologically relevant levels of MotB. Representative cell images show SX454 cells at (**A**) 10 min or (**B**) 15 min after induction of plasmids producing eGFP or MotB-eGFP by adding 0.02% arabinose, conditions that correspond to the level of MotB present in T4 infection of SX454 at 5 and 10 min, respectively. Panels from left to right show phase contrast, DNA stained with Hoechst (blue), eGFP or MotB-eGFP (green fluorescence), H-NS-mCherry (red fluorescence), and merged image of eGFP and mCherry.

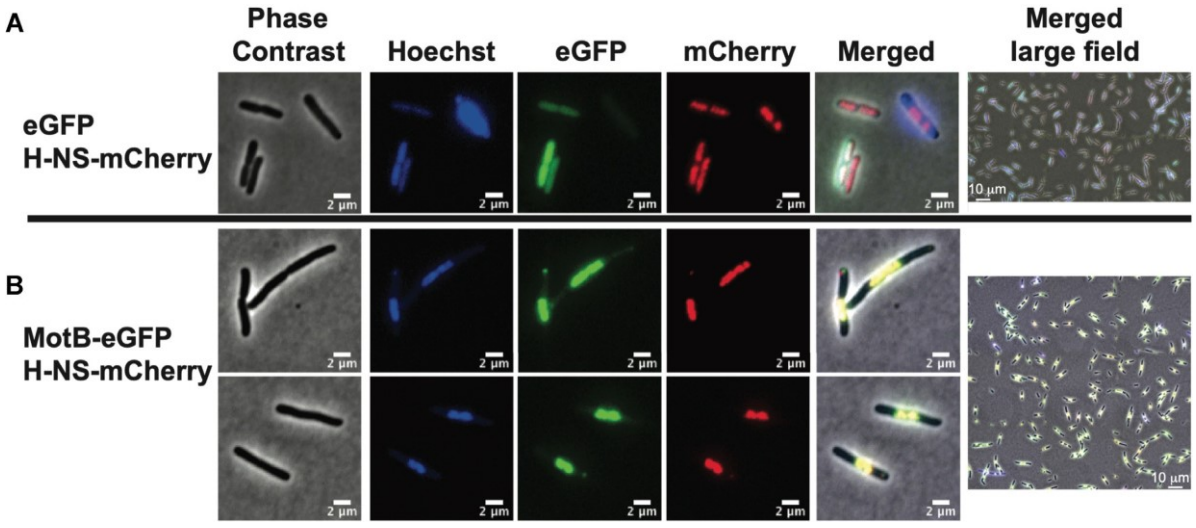

Figure 6. DNA compaction and co-localization of MotB-eGFP and H-NS-mCherry are observed with overexpression of *motB*. Representative cell images show SX454 cells containing chromosomal H-NS-mCherry and either the vector, p*eGFP* (**A**) or p*motB-eGFP* (**B**) after induction of protein synthesis for 1 hr with 0.2% arabinose. Panels from left to right show phase contrast, DNA stained with Hoechst (blue), eGFP or MotB-eGFP (green fluorescence), H-NS-mCherry (red fluorescence), merged image of eGFP and mCherry, and large field of a merged image.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

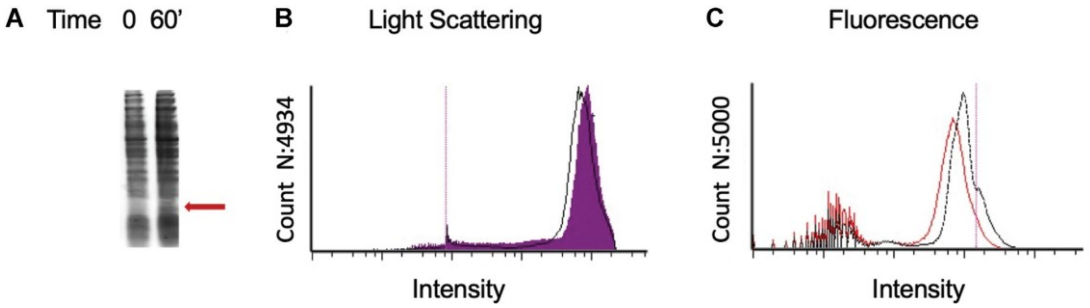

Figure 7. Flow Cytometry confirms cell lengthening in the presence of MotB and suggests that MotB inhibits replication. (**A**) SDS-PAGE gel showing the production of MotB at 60 min post-induction in TOP10F' containing p*motB*. The red arrow indicates the species corresponding to MotB. Flow cytometry analysis shows cell size distribution by light scattering (**B**) or DNA content by fluorescence (**C**), of TOP10F' cells containing the vector (black line in B and C) or containing p*motB* (filled purple in B, red line in C) at 60 min post-induction.

dant H-NS linearly coats the DNA and forms regions with loops and bridges, resulting in the global repression of many genes, particularly those with AT-rich sequences; the predominant type of binding varies depending on temperature, magnesium concentration, and salt (10–12).

As xenogeneic genes are often AT-rich, H-NS can serve to repress foreign genes acquired by horizontal transfer and has been implicated as a host defense mechanism against both lytic and lysogenic phages (74). However, it also silences specific host genes, and previous studies have shown that both host and phage proteins affect H-NS silencing (18–20). Proteins within the Hha/YdgT family can form heterocomplexes with H-NS, increasing H-NS repression [reviewed in (17)]. In contrast, H-NS derivatives that lack the C-terminal DNA binding domain (10,16,21,22) and the T7 5.5 protein (23,24) bind to H-NS to inhibit the formation of higher-order structures and relieve repression while T4 Arn protein, a structural DNA mimic, sequesters H-NS (25). Finally, DNA binding by the host HU and IHF proteins also appears to alter H-NS silencing at several loci, by changing the nucleoid structure (19), since both IHF (75) and HU (76,77) alter the physical organization of DNA. Both of these abundant proteins are involved in the nucleoid structure of the genomic DNA [reviewed in (73,78)] and interactions of HU and IHF with a large portion of the *E. coli* genome are known to globally affect gene expression (79). For example, the binding of IHF at specific loci, such as at the early promoter of phage Mu, alleviates H-NS repression (80).

Our results argue that MotB is a phage-encoded NAP-like protein that compacts genomic DNA, and like HU and IHF, this interaction with the DNA results in changes to the nucleoid itself with resulting alterations in gene expression. We have shown that MotB interacts with the DNA, resulting in compacted structures when observed *in vitro* by AFM (Figures 2 and 3) and in compacted nucleoids in live cells when observed by fluorescent microscopy (Figures 4–6). Importantly, this compaction is observed using levels of MotB that are similar to those during T4 infection. Taken together, these results suggest that MotB reorganizes the genomic DNA in the presence of H-NS as well as all of the other host NAPs. Global alterations in host gene expression accompany these changes in DNA compaction, but our transcriptomics pathways analyses do not identify a particular pathway that is targeted by MotB. Rather, the majority of affected host genes (Table 1, Supplementary Tables S1−S3) correlate with those that have been identified as part of the *hns* regulon (24,51). Thus, it seems likely that the global host expression changes that accompany *motB* overexpression reflect the changes within the nucleoid that affect DNA compaction and alter H-NS silencing. Our previous work indicated that MotB and H-NS (and StpA) copurify in the presence of DNA (3). However, whether there is a direct interaction between MotB and H-NS in the absence of DNA is not yet clear. Our attempts to observe a direct protein-protein interaction between MotB and H-NS in the absence of DNA by native protein gels or by pull-down assays have been inconclusive.

The gene expression changes associated with *motB* overexpression also include the up-regulation of the Type 1-E CRISPR locus present in *E.coli* that is known to be repressed by H-NS (57,60). This could reflect a counter-strategy by the host to thwart T4 infection. Such a phage/host arms race is common, as phage genes that aid in infection may also result in a counter defense from the host (81). However, because the *E. coli* CRISPR system is thought to be ineffective against phage (59,82), it is still unclear whether this *cas* up-regulation is particularly meaningful or simply reflects the overall derepression of the *hns* regulon.

Previously, other NAP-like proteins have been found within mobile genetic elements (83). In these cases, the proteins appear to function by locally altering NAP/DNA complexes, rather than by a dramatic change in the nucleoid itself. However, MotB appears to compact DNA in a different manner. It more resembles that of the NAP-like *E. coli* proteins Dps (26) and CbpA (27) and yeast cohesin (28). High levels of Dps and CbpA, which accumulate during stationary phase, result in nucleoid condensation and yield AFM images similar to what we have seen with MotB (26,27). Yeast cohesin generates protein−DNA clusters with multiple cohesin holocomplexes at the center of the DNA and DNA loops seen at the outer edges. Such AFM structures are also similar to those seen with MotB. As with MotB, cohesin compacted clusters are only seen in the presence of DNA (28), while Dps itself forms aggregates in the absence of DNA (26). However, unlike MotB, cohesin requires a minimum length of 3 kbp for cluster formation,

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

while we have shown that in gel shift assays MotB will bind to short DNA fragments (<50 bp) (3). Despite these differences, both the bacterial NAP-like proteins DpsA/CbpA and yeast cohesin are involved in the global organization of chromosomal DNA. The nucleoid compaction we observe with MotB *in vivo* suggests that it too serves this function.

We speculate that the synthesis of MotB as an early protein and its high level in the infected *E. coli* cell at 5 min will result primarily in compaction of host DNA at this time, since at 5 min, T4 DNA replication is just beginning. It seems likely that this could aid in T4 infection in multiple ways. First, it has been well-established that shortly after infection T4 DNases begin to degrade the *E. coli* genome, in preparation for the substrates needed for T4 replication. The degradation results in the eventual loss of ∼50% of the host DNA (84). In addition, heterologous expression of another T4 early protein, Ndd, has been shown to disrupt the host nucleoid, dispersing the host DNA throughout the cell (85). Consequently, we speculate that early in infection MotB may serve as a 'housekeeper', i.e., the early abundance of MotB could benefit T4 by 'sweeping' the host DNA back up, which could then provide more room for subsequent phage replication and assembly that commences 5 min post-infection. Second, it is also known that the host genome is degraded by a loose complex of T4 enzymes, which then channels the nucleotide substrates directly to the T4 replication machinery (86). Having a more compacted host genome could potentially provide a more efficient 'funnel' for this process. Finally, the MotB-mediated dysregulation of specific host genes and the overall depression of H-NS silencing of host DNA could also provide an extra benefit for the phage. However, these changes might also just be a by-product of the MotB compaction.

Later in infection, it seems likely that MotB may also serve a more direct role for the phage since we have found that MotB binds T4 gDNA similarly to the unmodified host DNA. It seems unlikely that this role is to provide general protection for the T4 DNA from H-NS repression, since the expression of T4 early/middle genes, except for *repEB*, is not affected by *motB* overexpression. However, the higher level of RNA for *repEB,* which encodes a protein involved in the initiation of DNA replication (29–31), suggests that MotB binding to T4 gDNA could affect the timing of phage replication. As late gene expression requires replication (53), this up-regulation of *repEB* would then explain early expression of many late RNAs at 5 min post-infection when *motB* is overexpressed. In addition, the abundance of MotB at 10 min, a time that coincides with high levels of phage replication/packaging, suggests that it might also have a role in these processes later in infection. Ongoing work centers on determining the effect of MotB during T4 infection.

DATA AVAILABILITY

RNA-seq data is available in the NCBI database (# GSE172467) and in Supplementary Tables S1-S3.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Leslie Knipling, Rawan Elaksher, and Abraham Correa-Medina for helpful discussions, the NIDDK genomics core for RNA sequencing, and Alec Goodman (Harvard University) for providing the *E. coli* strain SX454.

FUNDING

Intramural Research Program of the National Institutes of Health: National Institute of Diabetes and Digestive and Kidney Diseases (to J.P.-W., B.S, M. A.-M., D.M.H); National Institute of Biomedical Imaging and Bioengineering (to E.K.D.); Eunice Kennedy Shriver National Institute of Child Health and Human Development (to J.R.I.); National Cancer Institute (to R.R.); National Science Foundation [MCB-0923873]; National Institutes of Health [AI081726 to J.Z., V.R.]. Funding for open access charge: National Institutes of Health.

Conflict of interest statement. None declared.

REFERENCES

1. Miller,E.S., Kutter,E., Mosig,G., Arisaka,F., Kunisawa,T. and Ruger,W. (2003) Bacteriophage T4 genome. *Microbiol. Mol. Biol. Rev.*, **67**, 86–156.
2. Hatfull,G.F. and Hendrix,R.W. (2011) Bacteriophages and their genomes. *Curr. Opin. Virol.*, **1**, 298–303.
3. Patterson-West,J., Arroyo-Mendoza,M., Hsieh,M.L., Harrison,D., Walker,M.M., Knipling,L. and Hinton,D.M. (2018) The bacteriophage T4 MotB protein, a DNA-binding protein, improves phage fitness. *Viruses*, **10**, 343.
4. Dillon,S.C. and Dorman,C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.*, **8**, 185–195.
5. Ohniwa,R.L., Muchaku,H., Saito,S., Wada,C. and Morikawa,K. (2013) Atomic force microscopy analysis of the role of major DNA-binding proteins in organization of the nucleoid in *Escherichia coli*. *PLoS One*, **8**, e72954.
6. Song,D. and Loparo,J.J. (2015) Building bridges within the bacterial chromosome. *Trends Genet.*, **31**, 164–173.
7. Grainger,D.C. and Busby,S.J. (2008) Global regulators of transcription in *Escherichia coli*: mechanisms of action and methods for study. *Adv. Appl. Microbiol.*, **65**, 93–113.
8. Shen,B.A. and Landick,R. (2019) Transcription of bacterial chromatin. *J. Mol. Biol.*, **431**, 4040–4066.
9. Wang,S., Moffitt,J.R., Dempsey,G.T., Xie,X.S. and Zhuang,X. (2014) Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 8452–8457.
10. Liu,Y., Chen,H., Kenney,L.J. and Yan,J. (2010) A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev.*, **24**, 339–344.
11. Kotlajich,M.V., Hron,D.R., Boudreau,B.A., Sun,Z., Lyubchenko,Y.L. and Landick,R. (2015) Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria. *Elife*, **4**, e04970.
12. Winardhi,R.S., Yan,J. and Kenney,L.J. (2015) H-NS regulates gene expression and compacts the nucleoid: Insights from single-molecule experiments. *Biophys. J.*, **109**, 1321–1329.
13. Dame,R.T., Wyman,C. and Goosen,N. (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic. Acids. Res.*, **28**, 3504–3510.
14. Amit,R., Oppenheim,A.B. and Stavans,J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys. J.*, **84**, 2467–2473.
15. Dame,R.T., Noom,M.C. and Wuite,G.J. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, **444**, 387–390.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

16. van der Valk,R.A., Vreede,J., Qin,L., Moolenaar,G.F., Hofmann,A., Goosen,N. and Dame,R.T. (2017) Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity. *Elife*, **6**, e27369.
17. Navarre,W.W. (2016) The impact of gene silencing on horizontal gene transfer and bacterial evolution. *Adv. Microb. Physiol.*, **69**, 157–186.
18. Pfeifer,E., Hunnefeld,M., Popa,O. and Frunzke,J. (2019) Impact of xenogeneic silencing on phage-host interactions. *J. Mol. Biol.*, **431**, 4670–4683.
19. Navarre,W.W. (2010) In: Dame,R.T. and Dorman,C.J. (eds). *Bacterial Chromatin*. Springer, Dordrecht, Netherlands, pp. 251–322.
20. Stoebel,D.M., Free,A. and Dorman,C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology*, **154**, 2533–2545.
21. Williamson,H.S. and Free,A. (2005) A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol. Microbiol.*, **55**, 808–827.
22. Banos,R.C., Pons,J.I., Madrid,C. and Juarez,A. (2008) A global modulatory role for the Yersinia enterocolitica H-NS protein. *Microbiology (Reading)*, **154**, 1281–1289.
23. Liu,Q. and Richardson,C.C. (1993) Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 1761–1765.
24. Ali,S.S., Beckett,E., Bae,S.J. and Navarre,W.W. (2011) The 5.5 protein of phage T7 inhibits H-NS through interactions with the central oligomerization domain. *J. Bacteriol.*, **193**, 4881–4892.
25. Ho,C.H., Wang,H.C., Ko,T.P., Chang,Y.C. and Wang,A.H. (2014) The T4 phage DNA mimic protein Arn inhibits the DNA binding activity of the bacterial histone-like protein H-NS. *J. Biol. Chem.*, **289**, 27046–27054.
26. Ceci,P., Cellai,S., Falvo,E., Rivetti,C., Rossi,G.L. and Chiancone,E. (2004) DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic. Acids. Res.*, **32**, 5935–5944.
27. Cosgriff,S., Chintakayala,K., Chim,Y.T., Chen,X., Allen,S., Lovering,A.L. and Grainger,D.C. (2010) Dimerization and DNA-dependent aggregation of the Escherichia coli nucleoid protein and chaperone CbpA. *Mol. Microbiol.*, **77**, 1289–1300.
28. Ryu,J.K., Bouchoux,C., Liu,H.W., Kim,E., Minamino,M., de Groot,R., Katan,A.J., Bonato,A., Marenduzzo,D., Michieletto,D. *et al.* (2021) Bridging-induced phase separation induced by cohesin SMC protein complexes. *Sci. Adv.*, **7**, eabe5905.
29. Vaiskunaite,R., Miller,A., Davenport,L. and Mosig,G. (1999) Two new early bacteriophage T4 genes, repEA and repEB, that are important for DNA replication initiated from origin E. *J. Bacteriol.*, **181**, 7115–7125.
30. Brister,J.R. (2008) Origin activation requires both replicative and accessory helicases during T4 infection. *J. Mol. Biol.*, **377**, 1304–1313.
31. Kreuzer,K.N. and Brister,J.R. (2010) Initiation of bacteriophage T4 DNA replication and replication fork dynamics: a review in the Virology Journal series on bacteriophage T4 and its relatives. *Virol J*, **7**, 358.
32. Wang,W., Li,G.W., Chen,C., Xie,X.S. and Zhuang,X. (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*, **333**, 1445–1449.
33. Gao,Y., Foo,Y.H., Winardhi,R.S., Tang,Q., Yan,J. and Kenney,L.J. (2017) Charged residues in the H-NS linker drive DNA binding and gene silencing in single cells. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 12560–12565.
34. Wood,W.B. (1966) Host specificity of DNA produced by Escherichia coli: bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol.*, **16**, 118–133.
35. Hsieh,M.L., Hinton,D.M. and Waters,C.M. (2018) VpsR and cyclic di-GMP together drive transcription initiation to activate biofilm formation in Vibrio cholerae. *Nucleic Acids Res.*, **46**, 8876–8887.
36. Bonocora,R.P., Caignan,G., Woodrell,C., Werner,M.H. and Hinton,D.M. (2008) A basic/hydrophobic cleft of the T4 activator MotA interacts with the C-terminus of *E. coli* sigma70 to activate middle gene transcription. *Mol. Microbiol.*, **69**, 331–343.
37. Vick,J.E., Johnson,E.T., Choudhary,S., Bloch,S.E., Lopez-Gallego,F., Srivastava,P., Tikh,I.B., Wawrzyn,G.T. and Schmidt-Dannert,C. (2011) Optimized compatible set of BioBrick (TM) vectors for metabolic pathway engineering. *Appl Microbiol Biot*, **92**, 1275–1286.
38. Elliott,T. and Geiduschek,E.P. (1984) Defining a bacteriophage T4 late promoter: absence of a "-35" region. *Cell*, **36**, 211–219.
39. Zhu,J., Ananthaswamy,N., Jain,S., Batra,H., Tang,W.C., Lewry,D.A., Richards,M.L., David,S.A., Kilgore,P.B., Sha,J. *et al.* (2021) A universal bacteriophage T4 nanoparticle platform to design multiplex SARS-CoV-2 vaccine candidates by CRISPR engineering. bioRxiv doi: https://doi.org/10.1101/2021.01.19.427310, 20 January 2021, preprint: not peer reviewed.
40. Vlot,M., Houkes,J., Lochs,S.J.A., Swarts,D.C., Zheng,P., Kunne,T., Mohanraju,P., Anders,C., Jinek,M., van der Oost,J. *et al.* (2018) Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR-Cas effector complexes. *Nucleic Acids Res.*, **46**, 873–885.
41. Liu,Y., Dai,L., Dong,J., Chen,C., Zhu,J., Rao,V.B. and Tao,P. (2020) Covalent modifications of the bacteriophage genome confer a degree of resistance to bacterial CRISPR systems. *J. Virol.*, **94**, e01630-20.
42. Shlyakhtenko,L.S., Gall,A.A., Filonov,A., Cerovac,Z., Lushnikov,A. and Lyubchenko,Y.L. (2003) Silatrane-based surface chemistry for immobilization of DNA, protein-DNA complexes and other biological materials. *Ultramicroscopy*, **97**, 279–287.
43. Cantor,C.R. and Schimmel,P.R. (1980) In: *Biophysical Chemistry: Part II 'Techniques for the Study of Biological Structure and Function'*. WH Freeman and Co., Oxford.
44. Hinton,D.M. (1989) Transcript analyses of the uvsX-40-41 region of bacteriophage T4. Changes in the RNA as infection proceeds. *J. Biol. Chem.*, **264**, 14432–14439.
45. Patterson-West,J., James,T.D., Fernandez-Coll,L., Iben,J.R., Moon,K., Knipling,L., Cashel,M. and Hinton,D.M. (2018) The E. coli global regulator DksA reduces transcription during T4 infection. *Viruses*, **10**, 308.
46. Sanchez-Vazquez,P., Dewey,C.N., Kitten,N., Ross,W. and Gourse,R.L. (2019) Genome-wide effects on *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 8310–8319.
47. Keseler,I.M., Mackie,A., Santos-Zavaleta,A., Billington,R., Bonavides-Martinez,C., Caspi,R., Fulcher,C., Gama-Castro,S., Kothari,A., Krummenacker,M. *et al.* (2017) The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. *Nucleic Acids Res.*, **45**, D543–D550.
48. Jindal,S., Thampy,H., Day,P.J.R. and Kell,D.B. (2019) Very rapid flow cytometric assessment of antimicrobial susceptibility during the apparent lag phase of microbial (re)growth. *Microbiology*, **165**, 439–454.
49. Cowan,J., D'Acci,K., Guttman,B. and Kutter,E. (1994) In: Karam,J. (ed). *Molecular Biology of Bacteriophage T4*. American Society of Microbiology, Washington, DC, pp. 520–527.
50. Uzan,M., Daubentoncarafa,Y., Favre,R., Defranciscis,V. and Brody,E. (1985) The T4 mot protein functions as part of a pre-replicative DNA-protein complex. *J. Biol. Chem.*, **260**, 633–639.
51. Ueda,T., Takahashi,H., Uyar,E., Ishikawa,S., Ogasawara,N. and Oshima,T. (2013) Functions of the Hha and YdgT proteins in transcriptional silencing by the nucleoid proteins, H-NS and StpA, in *Escherichia coli*. *DNA Res.*, **20**, 263–271.
52. Rabussay,D. (1983) Phage-evoked changes in RNA polymerase. In: Mathews,C.K., Kutter,E.M., Mosig,G. and Berget,P.B. (eds). *Bacteriophage T4*. American Society of Microbiology, Washington, DC, pp. 167–173.
53. Geiduschek,E.P. and Kassavetis,G.A. (2010) Transcription of the T4 late genes. *Virol J*, **7**, 288.
54. Washizaki,A., Yonesaki,T. and Otsuka,Y. (2016) Characterization of the interactions between *Escherichia coli* receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. *Microbiologyopen*, **5**, 1003–1015.
55. Yu,F. and Mizushima,S. (1982) Roles of lipopolysaccharide and outer membrane protein OmpC of *Escherichia coli* K-12 in the receptor function for bacteriophage T4. *J. Bacteriol.*, **151**, 718–722.
56. Dorman,C.J. and Ni Bhriain,N. (2020) CRISPR-Cas, DNA supercoiling, and nucleoid-associated proteins. *Trends Microbiol.*, **28**, 19–27.
57. Diez-Villasenor,C., Almendros,C., Garcia-Martinez,J. and Mojica,F.J. (2010) Diversity of CRISPR loci in escherichia coli. *Microbiology (Reading)*, **156**, 1351–1361.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021

58. Xue,C. and Sashital,D.G. (2019) Mechanisms of Type I-E and I-F CRISPR-Cas systems in enterobacteriaceae. *EcoSal Plus*, **8**, https://doi.org/10.1128/ecosalplus.ESP-0008-2018.
59. Bozic,B., Repac,J. and Djordjevic,M. (2019) Endogenous gene regulation as a predicted main function of Type I-E CRISPR/Cas System in E. coli. *Molecules*, **24**, 784.
60. Pul,U., Wurm,R., Arslan,Z., Geissen,R., Hofmann,N. and Wagner,R. (2010) Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.*, **75**, 1495–1512.
61. Ayala,J.C., Wang,H., Silva,A.J. and Benitez,J.A. (2015) Repression by H-NS of genes required for the biosynthesis of the *Vibrio cholerae* biofilm matrix is modulated by the second messenger cyclic diguanylic acid. *Mol. Microbiol.*, **97**, 630–645.
62. Patterson-West,J., Tai,C.-H., Son,B., Hsieh,M.L., Iben,J.R. and Hinton,D.M. (2021) Overexpression of the bacteriophage T4 motB Gene Alters H-NS dependent repression of specific host DNA. *Viruses*, **13**, 84.
63. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of *Escherichia coli*. *J. Biol. Chem.*, **269**, 6578–6578.
64. Revel,H.R. (1983) DNA modification: glucosylation. In: Mathews,C.K., Kutter,E.M., Mosig,G. and Berget,P.B. (eds). *Bacteriophage T4*, American Society of MIcrobiology, Washington, DC, pp. 156–165.
65. Sato,Y.T., Watanabe,S., Kenmotsu,T., Ichikawa,M., Yoshikawa,Y., Teramoto,J., Imanaka,T., Ishihama,A. and Yoshikawa,K. (2013) Structural change of DNA induced by nucleoid proteins: growth phase-specific Fis and stationary phase-specific Dps. *Biophys. J.*, **105**, 1037–1044.
66. Nir,G., Lindner,M., Dietrich,H.R.C., Girshevitz,O., Vorgias,C.E. and Garini,Y. (2011) HU protein induces incoherent DNA persistence length. *Biophys. J.*, **100**, 784–790.
67. Seong,G.H., Kobatake,E., Miura,K., Nakazawa,A. and Aizawa,M. (2002) Direct atomic force microscopy visualization of integration host factor-induced DNA bending structure of the promoter regulatory region on the Pseudomonas TOL plasmid. *Biochem. Biophys. Res. Commun.*, **291**, 361–366.
68. Pleshakova,T.O., Bukharina,N.S., Archakov,A.I. and Ivanov,Y.D. (2018) Atomic force microscopy for protein detection and their physicosmall es, cyrillichemical characterization. *Int. J. Mol. Sci.*, **19**, 1142.
69. Bloch,V., Yang,Y., Margeat,E., Chavanieu,A., Auge,M.T., Robert,B., Arold,S., Rimsky,S. and Kochoyan,M. (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat. Struct. Biol.*, **10**, 212–218.
70. Spurio,R., Falconi,M., Brandi,A., Pon,C.L. and Gualerzi,C.O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J.*, **16**, 1795–1805.
71. Ueguchi,C., Seto,C., Suzuki,T. and Mizuno,T. (1997) Clarification of the dimerization domain and its functional significance for the Escherichia coli nucleoid protein H-NS. *J. Mol. Biol.*, **274**, 145–151.
72. Smyth,C.P., Lundback,T., Renzoni,D., Siligardi,G., Beavil,R., Layton,M., Sidebotham,J.M., Hinton,J.C., Driscoll,P.C., Higgins,C.F. *et al.* (2000) Oligomerization of the chromatin-structuring protein H-NS. *Mol. Microbiol.*, **36**, 962–972.
73. Azam,T.A. and Ishihama,A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.*, **274**, 33105–33113.
74. Navarre,W.W., McClelland,M., Libby,S.J. and Fang,F.C. (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev.*, **21**, 1456–1471.
75. Lin,J., Chen,H., Droge,P. and Yan,J. (2012) Physical organization of DNA by multiple non-specific DNA-binding modes of integration host factor (IHF). *PLoS One*, **7**, e49885.
76. van Noort,J., Verbrugge,S., Goosen,N., Dekker,C. and Dame,R.T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6969–6974.
77. Lioy,V.S., Cournac,A., Marbouty,M., Duigou,S., Mozziconacci,J., Espeli,O., Boccard,F. and Koszul,R. (2018) Multiscale structuring of the E. Coli chromosome by nucleoid-associated and condensin proteins. *Cell*, **172**, 771–783.
78. Verma,S.C., Qian,Z. and Adhya,S.L. (2019) Architecture of the escherichia coli nucleoid. *PLoS Genet.*, **15**, e1008456.
79. Prieto,A.I., Kahramanoglou,C., Ali,R.M., Fraser,G.M., Seshasayee,A.S. and Luscombe,N.M. (2012) 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.
80. van Ulsen,P., Hillebrand,M., Zulianello,L., van de Putte,P. and Goosen,N. (1996) Integration host factor alleviates the H-NS-mediated repression of the early promoter of bacteriophage Mu. *Mol. Microbiol.*, **21**, 567–578.
81. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, **577**, 327–336.
82. Touchon,M. and Rocha,E.P. (2010) The small, slow and specialized CRISPR and anti-CRISPR of Escherichia and Salmonella. *PLoS One*, **5**, e11126.
83. Flores-Rios,R., Quatrini,R. and Loyola,A. (2019) Endogenous and foreign nucleoid-associated proteins of bacteria: occurrence, interactions and effects on mobile genetic elements and host's biology. *Comput Struct Biotechnol J*, **17**, 746–756.
84. Warner,H.R., Snustad,P., Jorgensen,S.E. and Koerner,J.F. (1970) Isolation of bacteriophage T4 mutants defective in the ability to degrade host deoxyribonucleic acid. *J. Virol.*, **5**, 700–708.
85. Bouet,J.Y., Krisch,H.M. and Louarn,J.M. (1998) Ndd, the bacteriophage T4 protein that disrupts the *Escherichia coli* nucleoid, has a DNA binding activity. *J. Bacteriol.*, **180**, 5227–5230.
86. Mathews,C.K. (1993) The cell-bag of enzymes or network of channels? *J. Bacteriol.*, **175**, 6377–6381.

Downloaded from https://academic.oup.com/nar/article/49/16/9229/6345464 by guest on 12 October 2021