Non-coding RNAs binding to the nucleoid protein HU in *Escherichia coli*

Mirjana Macvanin¹, Rotem Edgar¹§, Feng Cui², Andrei Trostel¹, Victor Zhurkin² and Sankar Adhya¹#

¹Laboratory of Molecular Biology, ²Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-4264, USA

§ Current address: Epidemiology Division, The Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel

Running title: HU-binding RNAs in *E. coli*

Keywords: HU, nucleoid, RNA, RIP-Chip, repeat elements

# Corresponding author. E-mail address: adhya@nih.gov
Abstract

Some unidentified RNA molecules, together with the nucleoid protein HU, were suggested to be involved in the nucleoid structure of Escherichia coli. HU is a conserved protein known for its role in binding to DNA and maintaining negative supercoils in the latter. HU also binds to a few RNAs but the full spectrum of its binding targets in the cell is not known. To understand any interaction of HU with RNA in the nucleoid structure, we immunoprecipitated potential HU-RNA complexes from cells and examined bound RNAs by hybridization to whole-genome tiling arrays. We identified associations between HU and ten new intragenic and intergenic noncoding (nc) RNAs, two of which are homologous to the annotated Bacterial Interspersed Mosaic Elements (BIME) and boxC DNA repeat elements. We confirmed direct binding of HU to BIME RNA in vitro. We also studied the nucleoid shape of HU and two of the ncRNA mutants (nc1 and nc5) by transmission electron microscopy and showed that both HU and the two ncRNAs play a role in nucleoid morphology. We propose that at least two of the ncRNA species complex with HU and help the formation or maintenance of the architecture of the E. coli chromosome. We also observed binding of HU with rRNA and tRNA segments, a few sRNAs and a distinct small set of mRNAs although significance, if any, of these associations is not known.
Introduction

HU, a well conserved and abundant nucleoid protein in *E. coli*, is composed of two 9 kDa homologous subunits, α and β (9, 18, 40, 41, 48, 49). HU showed non-specific interactions with DNA with micromolar affinity (8, 52) and specific interactions with nicked, cruciform and kinked DNA with nanomolar affinities (5, 29, 30, 45, 52, 55, 60). HU also showed interactions with tRNAs, *dsr* and *rpoS* mRNA (4, 5).

It was suggested that RNA molecules, together with HU and other nucleoid proteins, are involved in maintaining *E. coli* chromosome structure although the identities of the RNAs have not been established (44). These and subsequent studies showed that 100-300 nt long RNA chains are part of the nucleoid (43, 44). Treatment of the nucleoid with RNase caused a decondensation with a dramatic decrease in the sedimentation constant of the isolated chromosome from 1600S to approximately 400-500S (44, 65). In this work, we followed the idea that some HU-RNA complex(es) are of importance in nucleoid structure with the aim of identifying the RNA(s) that bind to HU and then understanding the role of such complexes in the nucleoid.

We took the ribonomic approach, also referred to as RIP-Chip assay that involves immune-precipitation of ribonucleoprotein complexes (RNPs) with antibodies against specific proteins, extraction of RNA, and hybridization to microarrays (58). The assay helped global identification of putative endogenous RNA targets of specific proteins in several systems (11, 12, 24, 25, 32, 33, 50, 51, 53, 57, 66).
We report associations of HU with several RNA and their identification: a set of noncoding RNAs (ncRNAs), not previously reported, regions of rRNA and tRNA, a few sRNA, and a set of mRNAs. We further investigated some of the noncoding RNAs by confirming their transcription by Northern blots and by showing their direct binding to HU by Electrophoretic Mobility Shift Assay (EMSA). By Transmission Electron Microscopy (TEM) analysis, we showed that deletion of two ncRNAs (nc1 and nc5) lead to nucleoid de-condensation. We propose that at least two of the ncRNA play a role in the E. coli chromosome architecture.

Materials and Methods

Bacterial strains and growth conditions

All E. coli strains used in this study are derivatives of E. coli K-12 MG1655. In order to construct 8 myc tags fused to HUβ, two complementary primer sets, each with four myc tags, flanked by HindIII-EcoRI and EcoRI-BamHI were used (Supplemental Material, Table S1). Complementary primers were annealed, ligated to each other and cloned into the HindIII-BamHI cloning vector containing the hupB ORF. Finally, the 8 myc-tagged hupB was integrated into the chromosome in place of hupB (10). The protein expression profile, as monitored by Western blotting, of the 8 myc-tagged HUβ strains were identical to the wild type (data not shown). We tested the tagged HUβ functionality by testing two cellular functions as described in Results. For TEM analysis, we constructed deletions of nc1 or nc5 by recombineering method (10), introducing a ~1kb tetracycline resistance cassette which replaced the intergenic region between yagP and intF or yjdN and yjdM,
respectively. HU deletions were constructed by P1 transduction of hupA::Kan and hupB::Cm alleles into the wild-type, Δnc1 or Δnc5 background.

RIP-Chip assay

The RIP-Chip assay was adapted from Zhang et al. (66) and Hu et al. (22).

(a) Cell growth. Cells of the strain with eight myc tags fused to HUβ were grown in M63 medium supplemented with 0.2% glucose, 0.1% casaminoacids and 0.002% vitamin B1 at 37°C to exponential phase (OD600 = 0.4).

(b) Cell extracts. Over-night cultures, in M63 media + 0.2% glucose + 0.1% casamino acids + 0.0001% thiamine were diluted 1:100 into 50 ml and grown to OD600=0.8. Samples were centrifuged and resuspended in 400 ml lysis buffer (20mM Tris-HCl pH 8.0, 150mM KCl, 1mM MgCl2, 1mM DTT). After addition of 400 μl (about 0.6 g) glass beads (Sigma, Cat. # G1277) and 2μl RNasin, 10 cycles of alternate 30 sec vortexing and 5 sec incubations on ice were performed. Additional 800 μl lysis buffer were added and after a brief vortexing for 10 sec, the samples were spun down at 4°C for 10 min. 1 ml supernatant samples were transferred into new eppendorf tubes.

(c) Preparation of Ab-PAS beads. - 24 mg Protein A Sepharose CL-4B (PAS) (Amersham Biosciences, Cat # 17-0780-01) were swelled in 200 μl Net2 buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 0.05% Triton X-100). 20 μl α-myc antibody (Ab) (Sigma, M4439) were added and rotated at 4°C over-night. Ab-PAS was washed 5 times with Net2 buffer.
(d) **Immunoprecipitation.** - Immunoprecipitation was carried out according to the modified protocol of Wassarman and Storz (64). 200 μl cell extract, 200 μl Net2 buffer and 1 μl RNasin were added to the Ab-PAS pellet and rotated at 4°C for two hours. Samples were washed 5 times with 1.5 ml Net2 buffer.

(e) **Isolation of RNA.** - RNA fragments were purified using a QIAGEN RNeasy purification kit according to the manufacturer’s protocol. The RNA pellet was dissolved in DEPC H₂O. A single immunoprecipitation usually yields about 10 μg RNA (for 50ml). Since we omitted any crosslinking, there is a chance that RNAs are free to dissociate and bind HU in the cell lysate, and RNAs that may normally be sequestered in the cell and unavailable for binding to HU are potentially released in the lysis process and then bind to HU nonspecifically.

(f) **RNA hybridization and detection.** Tiling array chips on which the complete *Escherichia coli* MG1655 genome sequence is represented were used (*E. coli* Tab520346F; Affymetrix, Santa Clara, CA). The chips had 1,159,908 probes in 1.4cm x 1.4cm and a 25-mer probe every 8 bp in both strands of the entire *E. coli* genome. The probes overlapped by 4 bp with the opposite strand probes. Each 25-mer DNA probe in the tiling array chip is 8 bp apart from the next probe and designed to cover the whole *E. coli* genome. The intergenic regions were originally defined by Blattner et al. (6). The RNA samples were hybridized directly to the oligonucleotide arrays with no prior labeling or cDNA synthesis. The RNA samples were added to the hybridization solution (1X MES (100 mM MES, 1 M NaCl, 20 mM EDTA, 0.01% Tween 20, pH 6.6), 0.1 mg ml⁻¹ herring sperm DNA, 0.5 mg ml⁻¹ BSA and 50 pM of the
control Biotinoligonucleotide B2 (GeneChip® Expression Analysis Technical Manual), heated to 99°C for 5 minutes and then incubated at 45°C for an additional 5 minutes before being placed in the microarray cartridge. Hybridization was carried out at 45°C for 16 hours on a rotary mixer at 60 r.p.m. Following hybridization, the sample solution was removed and the array was washed as recommended in the technical manual (Affymetrix). Hybridization was assayed using antibodies specific to the RNA:DNA hybrid and the Digene HC ExpressArray™ Kit (Digene, Gaithersburg, MD) as described by Hu et al. (22). The antibody was a gift of Dr Stephen Leppla (National Institute of Allergy and Infectious Diseases, National Institutes of Health). 25 μl of an RNA:DNA polyclonal antibody (1.3 mg ml⁻¹) was resuspended in 475 μl of matrix solution, loaded on the array and incubated at 25°C for 20 minutes (22). After removal of the first antibody stain and 10 wash cycles, the array was incubated with the second antibody stain, containing 0.036 mg ml⁻¹ biotin rabbit anti-goat IgG and 0.4 mg ml⁻¹ rabbit IgG in 1X MES. The RNA:DNA hybrids were fluorescently labeled by incubating with 10 mg ml⁻¹ streptavidin-phycoerythrin (Molecular Probes, Eugene, OR) and 2 mg ml⁻¹ BSA in 1X MES. Hybridized, washed and stained tiling arrays were scanned using a Genechip Scanner 3000 (Affymetrix). Standardized signals, for each individual probe in the array, were generated using the MAT analysis software (27) at the University of Wisconsin-Madison, for Escherichia coli K-12 MG1655 version m56 (16).

RNA preparation and Northern blot experiments

Overnight cultures were grown at 37°C in M63 medium supplemented with 0.2% glucose, 0.1% casaminoacids and 0.002% vitamin B1. Strains were diluted 1:100 and
grown to mid-exponential phase at 37°C in M63 medium supplemented with 0.2% glucose, 0.1% casaminoacids and 0.002% vitamin B1. RNA was extracted using the hot phenol method (1). RNAs were detected by Northern blotting, as follows. Five μg of RNA from each sample was separated by 10% TBE-Urea polyacrylamide gel (Invitrogen, EC6875) that was pre-run for 30 minutes at 70V and subsequently electrophoresed at 60V for 2 hours in 1 X TBE. Serial dilutions of at least one sample were included on each gel. RNA was transferred onto a positively charged nylon Zeta-probe membrane (Bio-Rad) at 200 mA for 1 hour, in 0.5 X TBE. After UV cross-linking, membranes were incubated with Ultrahyb solution (Ambion) at 42°C for 30 minutes and then hybridized to the 5'-biotinylated probe (100 ng ml⁻¹ in Ultrahyb) for approximately 15 to 20 hours at 42°C. The sequences of 5'-biotinylated DNA probes were, for nc1: CGAAGGTCGTTGTCGACTCCTATTATCGGCACC and for nc5: GATAAGACGCGCAACGTCGTCGATCCAGGCAGTCCGCAC. The blots were washed, and labeled RNA was detected using a BrightStar BioDetect kit (Ambion), as recommended by the manufacturer.

**DNA Templates and RNA Synthesis**

The series of complementary ssDNA oligonucleotides containing the T7 promoter sequence (5'-TAATACGACTCACTATAGG-3') followed by a sequence which corresponds to nc5 was used (Supplemental Materials, Table S2). The dsDNAs were obtained by annealing the appropriate complementary oligonucleotides. Annealing reactions were carried out by incubating equal amounts of two complementary oligonucleotides (500 nM) for 30 seconds at 95°C in 20 mM Tris–HCl (pH 8.0), 200 mM...
NaCl, and then allowing them to cool slowly. Transcription reactions were carried out as described previously (38). Double stranded template DNA (2 nM) was preincubated at 37°C for 5 minutes in transcription buffer (50 µl) containing 20 mM Tris acetate (pH 7.8), 10 mM magnesium acetate, 200 mM potassium glutamate, 1 mM dithiothreitol, 0.8 U µl⁻¹ rRNasin (Promega), and 100 units of T7 RNA polymerase (USB; 200 units µl⁻¹). To initiate the reactions, nucleotides were added to final concentrations of 1 mM ATP, 0.1 mM GTP, 0.1 mM CTP, 0.01 mM UTP, and 5 µCi (α-3²P)UTP (MP Biomedicals; specific activity, 3,000 µCi mmol⁻¹, 10 µCi µl⁻¹). The reaction mixtures were incubated at 37°C for an additional 10 minutes, and then the reactions were terminated by the addition of 50 µl of loading dye (90% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue). After incubation at 90°C for 2 to 3 minutes, samples were chilled on ice. Sample aliquots of 10 µl were loaded onto an 8% sequencing gel and electrophoresed at 60 W.

Gel Mobility Shift Assay
Varying amounts of HU protein were incubated with (5'-3²P)-labeled RNAs for 15 minutes in high salt buffer (20 mm Tris-HCl, pH 7.5, 200 mm NaCl, 10% glycerol) at room temperature (5). Electrophoresis was carried out on 8% Polyacrylamide gel (29:1) buffered with 0.5 X Tris-borate buffer, pH 8.3. Gels were analyzed and quantified using a PhosphorImager with Molecular Dynamics software.

Transmission electron microscopy
Equal volumes of Escherichia coli cell culture and 2×fixative (8% formaldehyde and 4% glutaraldehyde in phosphate buffer saline, pH 7.4) were centrifuged at 200 × g for 15 min.
and processed for thin-section EM analysis. The pellet was rinsed thoroughly in sodium cacodylate buffer (0.1 M, pH 7.4) and postfixed in 1% osmium tetroxide in the same buffer for 1 h at room temperature. The pellet was dehydrated in a series of graded ethanol (e.g., 35, 50, 75, 95, and 100%) and 100% propylene oxide, infiltrated in an equal volume of 100% propylene oxide and epoxy resin overnight. The pellet was embedded in the pure epoxy resin and was cured at 55°C for 48 h. The cured block was thin sectioned in =50–60 nm sections with an ultramicrotome. The thin sections were mounted on a naked copper grid and stained in uranyl acetate and lead citrate solution to enhance the contrast. The grids were examined, and digital images were taken with an H7000 electron microscope (Hitachi, Tokyo) operated at 75 kV. We looked at 50 cells of each strain. Figures 5, 6 and 7 show two pictures of each strain. In each case, the two are representatives of 90% of the collected images for each strain.

Results

RIP-Chip identification of RNAs bound to HU

Exponentially growing Escherichia coli cells were shown to contain a mixture of homodimeric HUα2 and heterodimeric HUαβ but not homodimeric HUβ2 is detected (9). To identify the RNA targets of HUαβ heterodimer in exponentially growing cells, we constructed an E. coli strain in which the chromosomal hupB was replaced by the hupB gene with eight successive myc tags fused to its N-terminus. We confirmed that the tagged HUβ-myc protein was functional by two assays: (i) Mini-P1 plasmids cannot be maintained in the host cell unless either HUα or HUβ is present (42). We showed that, as
expected, mini-P1 plasmid cannot transform the *E. coli* strain that is genotypically Δ*hupA*Δ*hupB* whereas the same plasmid can be stably transformed in the Δ*hupA* hupB*~myc* host, suggesting that the HUβ*~myc* fusion protein is active (Table 1). (ii) Bacteriophage Mu does not form plaques on a Δ*hupA* Δ*hupB* *E. coli* lawn on agar plates (31). We found that Mu forms plaques on Δ*hupA* hupB*~myc* at almost normal efficiency compared to the wild type host suggesting that the fusion protein is biologically active (Table 1).

We chose to grow the cultures in minimal medium to allow for more genes to be expressed for RIP-Chip assays (56). It was previously reported that a larger number of genes have expression intensities above the background values when *E. coli* is grown under these conditions compared to rich medium. Although our minimal media contained casamino acids, the gene expression spectrum is still higher than in LB medium. After immunoprecipitation and RNA isolation, samples were identified by direct hybridization to the *E. coli* DNA tiling arrays, by using an antibody specific for RNA:DNA hybrids to detect hybridization on the DNA chips (22). As a negative control, we used RNA extracted from the parental strain in which no *myc* tag was present. Total RNA extracted from cells served as total signal. We used 0.8 μg of RNA in the immuno-precipitated experiment, 8 μg in the total RNA experiment, and 0.5 μg in the negative control experiment. Microarray analysis was done using a Model-based Analysis of Tiling-arrays (MAT) which models the baseline probe behavior by considering probe sequence and copy number on each array. It standardizes the probe value through the probe model, eliminating the need for sample normalization. MAT uses an innovative function to score
regions for enrichment, which allows for robust P value and false discovery rate calculations (27).

The RNA targets which co-immunoprecipitated with HU and identified by tiling arrays are given in Table 2 (accession number in GEO databank: GSE28565, http://www.ncbi.nlm.nih.gov/geo). We found segments of: (i) all tRNAs; (ii) 23S, 16S and 5S rRNAs; (iii) eleven mRNAs; and (iv) four sRNAs that immunoprecipitated with HUβ-myc. We additionally found ten RNA molecules transcribed from intergenic and intragenic regions; no RNA was previously reported to be transcribed from these regions. These RNAs were named nc1 to nc10 (where “nc” stands for noncoding).

We observed that HU is associated with rRNA and tRNAs. It is conceivable that HU binding to rRNA, tRNA and some mRNA is because of their known abundance and therefore is non-specific. Specific binding can be tested by the use of a HU mutant that does not bind RNA in the IP experiment. However due to lack of our knowledge about the RNA binding domain of HU, such mutants are not available. Nevertheless, the abundance scenario is unlikely because many other abundant mRNAs and sRNAs are not present in our IP experiments. However, our findings are in agreement with previous reports estimating that 40% of nucleoid-associated RNAs are segments of rRNA (43), and HU was previously shown to bind tRNA (5). It was reported that many tRNAs of *E. coli* contain stretches of sequences that are similar to those found in rRNAs (7). These similar sequences contain both loops and stems and are too frequent to be coincidental (7). It is possible that HU might recognize similar sequences and the corresponding
structural motifs in both RNA species. Previous studies of structural motifs important for
the HU binding revealed that HU recognizes a stem-loop structure (5).

Eleven mRNAs co-immunoprecipitated with HU, the majority of which code for
membrane-associated proteins (Table 2). Two representatives of HU-binding mRNAs,
\textit{lpp} and \textit{csp}A are shown in Figure 1, panel A and Figure S1, panel A (Supplemental
Material), respectively. It is possible that some of these mRNAs are also present in the
nucleoid since it was shown that nucleoid RNA, in addition to rRNAs, also contain some
mRNAs (19). The majority of membrane proteins listed in Table 2 have GC content
higher than 50%. As previously mentioned, HU binds specifically to the RNA fragment
containing the translational initiation region of \textit{rpoS} mRNA (4). However, we did not
detect immunoprecipitation of \textit{rpoS} mRNA with HU because we did not detect any \textit{rpoS}
mRNA in the total RNA under our conditions.

sRNAs are divided into two subgroups based on their function: the regulatory sRNAs that
act as regulators of gene translation, and the housekeeping sRNAs that affect different
aspects of cellular metabolism (61, 62). Our results show that HU binds to only four of
about hundred sRNAs: \textit{ssrA} (tmRNA), \textit{ssrS} (6S sRNA), \textit{ffs} (4.5S sRNA) and \textit{rnpB}, which
incidentally belong to the sRNAs that do not require the chaperone protein Hfq for their
function (66) (Table 2). Tilling array patterns of two of the HU-binding sRNAs, \textit{ssrA} and
\textit{ssrS}, are shown in Figure 1, panel B and Figure S1, panel B (Supplemental Materials).
These four “housekeeping” sRNAs are highly structured (63). In addition, consistent with
the observation that HU binds mRNA that are G/C-rich, the GC content of the
housekeeping sRNAs (ssrA 53%, ssrS 55%, ffs 62% and rnpB 62%) are also high. This is in contrast to the low GC content of the regulatory sRNAs (20).

Ten novel noncoding RNAs

A search for the E. coli sequences homologous to ncRNA fragments using NCBI BLAST (26) showed that a segment of nc1 has homology to the thrW tRNA gene. The fragment corresponding to nc1 is encoded in the intergenic region between yagP and intF (Figure 1, panel C). This intergenic region contains a pseudogene of thrW inserted at CP4-6 prophage att site, as annotated in EcoGene Database of Escherichia coli Sequence and Function (www.ecogene.org). An NCBI BLAST search revealed that nc2, nc5 (discussed later in more detail) and nc7 are homologous to repetitive palindromic extragenic DNA elements. We found that nc2 and nc7 are homologous to boxC DNA elements. On the E. coli K-12 chromosome, there are 22 extragenic regions containing either one or two boxCs, totaling 32 boxCs. boxCs contain 56-bp long imperfect palindromes, with a pyrimidine-rich 5' end (the tail) and a purine rich 3' end (2). The functional significance of boxC is not known. It was proposed that boxC represent potential sites for protein binding, but no protein has been so far identified to bind these DNA sequences (2). Our data suggest that at least two of the boxC elements are transcribed into RNA to which HU binds. nc2 RNA is encoded by the boxC element in the intergenic region between adk and hemH (Supplemental Material, Figure S1, panel C) whereas nc7 is encoded in the boxC element that is located between fabI and ycfD (Supplemental Material, Figure S1, panel D).
nc3, nc4, nc6, nc8 and nc9 are homologous to parts of the noncoding strands of narX, tdcB, yjiN, treF and dapB genes, respectively. Two representatives, nc3 and nc8, are shown in Figure S1, panels E and F (Supplemental Material). nc10 is complementary to the intergenic region between ymfL and ymfM, including parts of both ymfL and ymfM genes.

nc5 refers to the two highly similar immunoprecipitated fragments located in the intergenic region between phnA and phnB (Figure 1, panel D) that was previously annotated in The EcoGene Database of Escherichia coli Sequence and Function (www.ecogene.org) as the region that contains repetitive extragenic palindromic elements (REP). MAT analysis of RIP-Chip data (27) suggests that the two immunoprecipitated fragments are 88nt and 104nt-long, the first 88 of the 104nt-long RNA being identical in sequence to the individual 88nt RNA. BLAST search revealed that the segments of the E. coli genome that are homologous to nc5 represent BIME DNA elements (14).

We also verified the presence of nc1 and nc5 by Northern blot experiments with two 40 nt-long 5'-biotynilated DNA probes corresponding to nc1 and nc5 respectively. We observed an RNA band of approximately 90 nt length that hybridized with the probe specific for nc1 (Figure 2A). In the case of the nc5 probe (Figure 2B), we also detected a similar size RNA that matches the ~88 nt RNA detected by tiling array. Since estimating the length of RNA by MAT score is not precise, we believe that the length of nc5 RNA is approximately 90 nt as detected by Northern blot.
Electrophoretic Characterization of HU: RNA Complexes

We tested direct binding of HU to BIME nc5 RNA in vitro by electrophoresis mobility shift assay (Figure 3A). We used an 88nt long nc5 RNA at a constant amount (15 fmol) and increasing concentrations of HUαβ heterodimer (13–4400 nM). The results presented in Figure 3A established that HU formed a complex with the 88-mer RNA. The apparent Kd value for HU binding was about 450 nM (Figure 3B). BIMEs comprise combinations of several short sequence motifs called palindromic unit, PU (21) which are 40 bp-long imperfect palindromes found in the E. coli and S. typhimurium genomes. PUs are either found as single occurrences or in clusters, where they alternate in orientation and type (i.e., Y and Z2 according to classification suggested by Gilson and colleagues (3, 15). The M-fold analysis (67) of the nc5 RNA shows that it likely forms a BIME-like structure containing a PU Z2 hairpin (Figure 4). We found that the HU binds to the PU Z2 with a Kd around 600 nM (Figure S2, panel A, Supplemental Material). This value is similar to the Kd measured for HU binding to the whole nc5 RNA fragment (Figure 3B). Several bands of the HU:RNA complexes are observed, each possibly representing a unique number of HU molecules bound to these RNA fragments. Z2 RNA has two mismatches (Figure 4). HU was previously shown to bind strongly to DNA mismatches and bulges (28). Therefore, Z2 hairpin present in nc5 might be an important determinant of HU binding.

Transmission Electron Microscopy (TEM) Analysis of Nucleoid Structure

In order to establish whether ncRNAs identified in this work have effect on nucleoid structure, we investigated the morphology of nucleoids in cells with deletions of nc1, nc5
RNAs and HU by TEM (Figures 5, 6 and 7). In contrast to well-defined nucleoids of wild-type cells (Figure 5, panels A and B), nucleoids from HU-depleted cells show high degree of decondensation (Figure 5, panels C and D), which is in agreement with previous literature (23). Interestingly, deletions of either nc1 (Figure 6, panels A and B) or nc5 (Figure 7, panels A and B) have an impact on nucleoid organization. In both cases, nucleoids are less compacted than in the wild-type. In the strain with deletion of nc5, the nucleoid is decondensed but still retains shape (Figure 7A and B). Further deletion of HU in this genetic background leads to higher decompaction and loss of shape of the nucleoid (Figure 7, panels C and D). The strain with deletion of nc1 has a dispersed nucleoid with no defined shape (Figure 6, panels A and B). Further deletion of HU in the absence of nc1 does not seem to have any additional effect on nucleoid morphology (Figure 6, panels C and D). Taken together, these results suggest that nc1 and nc5 play a role in defining both the shape and compactation of the nucleoid in an unknown way. We are currently investigating the underlying mechanisms.

Discussion

The major purpose of the research described here is to identify the RNA species binding the nucleoid-associated protein HU. It has been suggested that in the nucleoid, segments of DNA (domains) are held together by RNA molecules and stabilized by architectural proteins such as HU (19, 34, 44, 54, 65). In this work, we emphasized two noncoding RNA, nc1 and nc5, because of their potential role in the E. coli nucleoid structure. nc1 originates from a threonine tRNA gene (E. coli DNA sequence co-ordinates 296361 –
interrupted by a prophage att site. \textit{nc}5 is homologous to the DNA repeat elements called BIME. BIME elements location on the \textit{E. coli} chromosome is non-random with most of the positions in the GC-rich genomic regions (red bars in Figure 8).

A previous report (13) proposed that all copies of BIME identified so far are either between two genes that are co-transcribed as part of a single operon or, alternatively, within the 3' untranslated region at the end of an operon. In all cases where the extent of transcription is known the BIME sequence falls within a transcribed region. However, some of BIME RNA, even if stable, may not interact with HU and thus were not present in our ncRNA collection.

We have determined the \textit{E. coli} transcriptome using DNA tiling array in exponentially growing cells of wild type MG1655 (data not shown). We found that of 264 BIMEs identified in \textit{E. coli} (http://www.pasteur.fr/recherche/unites/pmtg/repet/tableauBIMEcoli.html), only approximately 100 showed significant hybridization signals. Because of extensive sequence similarities between the BIME elements, we were unable to conclude whether all of those BIMEs are transcribed, or if only a few of them give transcripts which then hybridize on to the tiling array to other BIME locations because of sequence homology. Based on our results of the RIP-Chip assay, we assigned the BIME element(s) at location between \textit{phn}A and \textit{phn}B as the source of \textit{nc}5 RNA.

Deletion of both \textit{nc}5 and \textit{nc}1 RNAs showed effects on nucleoid structure. The nucleoid of the strain with the \textit{nc}5 deletion assumed a more open conformation (Figures 7A and B) and deletion of the HU protein further exacerbated nucleoid decompactation (Figures 7C...
and D). Hecht and Pettijohn suggested that certain unknown RNA-DNA interactions in the nucleoid restrain the rotation and extension of the DNA (19). The number of chromosomal domains was variously estimated to be a few hundreds, the upper limit of which is compatible with the number of BIME sequences occurring on the chromosome. ncRNAs may interact simultaneously with BIME DNA sequences around the chromosome and HU to form ternary complexes potentially playing the role of anchoring points, which define different chromosomal domains. Interestingly, deletion of nc1 leads to decondensation of the nucleoid regardless of the presence of HU in the cell (Figures 6 A-D). It is possible that both nc1 and nc5 RNAs participate in nucleoid structure by different mechanisms which are the subject of our future studies.

A recent publication on RNA-seq analysis in E. coli has confirmed the existence of all ncRNAs reported in this work in varying amounts, except that nc4 is expressed at a very low level in exponentially growing cells whereas nc1 and nc2 are present in large excess relative to the others (46). The same publication confirmed the presence of all the mRNA species that bind to HU in exponentially growing cells, except cspA.

Acknowledgements

We thank our colleagues in the laboratory for various helps and suggestions and Stephen Leppla for the generous gift of the RNA:DNA mouse monoclonal antibody. We also thank R. Raghavan and E. Groisman for the communication of their results to us. This research was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research.
References


identification of 22 bud-localized transcripts using DNA microarray analysis.


Nucleic Acids Res. 20:1553-1558.


Table 1. Function of HUβ-fusion protein as measured by efficiency of mini-P1 plasmid replication and bacteriophage Mu growth*

<table>
<thead>
<tr>
<th>Host strain</th>
<th>CFU of mini-P1 / µg DNA</th>
<th>PFU of phage µ/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.0 x 10^8</td>
<td>3.7 x 10^4</td>
</tr>
<tr>
<td>8myc-hupB</td>
<td>1.0 x 10^7</td>
<td>2.5 x 10^8</td>
</tr>
<tr>
<td>ΔhupA</td>
<td>4.0 x 10^7</td>
<td>2.2 x 10^8</td>
</tr>
<tr>
<td>ΔhupA, 8myc-hupB</td>
<td>2 x 10^5</td>
<td>4.4 x 10^7</td>
</tr>
<tr>
<td>ΔhupA, ΔhupB</td>
<td>&gt;1 x 10^3</td>
<td>&gt;1 x 10^2</td>
</tr>
</tbody>
</table>

* Mini-P1 plasmid maintenance was measured by transforming various competent E. coli cells with plasmid DNA and then counting the number of transformed cells (CFU) per unit of DNA. Efficiency of phage µ growth was estimated by measuring the number of plaques obtained by plating phage µ (PFU) on various hosts per given amount of phage particles.
Table 2. List of RNA fragments which co-immunoprecipitated with HU in RIP-Chip assays. Number of immunoprecipitated fragments and their coordinates on the *E. coli* genome are also listed. Cases where no physiological role is assigned to a given RNA species are marked as “unknown.”

<table>
<thead>
<tr>
<th>RNA species</th>
<th>Name</th>
<th>Physiological role</th>
<th>Genomic coordinates</th>
<th>Number of fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA</td>
<td></td>
<td>Protein synthesis</td>
<td></td>
<td>89</td>
</tr>
<tr>
<td>rRNA</td>
<td></td>
<td>Protein synthesis</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>sRNA</td>
<td><em>ffs</em> (4.5S)</td>
<td>Component of the signal recognition particle (SRP)</td>
<td>475672-475785</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>ssrA</em> (tmRNA)</td>
<td>Trans-translation</td>
<td>2753615-2753977</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>ssrS</em> (6S)</td>
<td>RNA-polymerase associated; transcription</td>
<td>3053970-3054220</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>rnpB</em></td>
<td>RNase P RNA component precursor</td>
<td>3268614-3268238</td>
<td>1</td>
</tr>
<tr>
<td>mRNA</td>
<td><em>rpsK</em></td>
<td>30S ribosomal protein S11</td>
<td>3440120-3439731</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>fliC</em></td>
<td>Flagellar filament structural protein (flagellin)</td>
<td>2001630-2000134</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>agaW</em></td>
<td>N-acetylgalactosamine-specific enzyme, IIC component of PTS</td>
<td>3278723-3279097</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>cspA</em></td>
<td>Major cold shock protein</td>
<td>3718072-3718284</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>lpp</em></td>
<td>Outer membrane murein lipoprotein</td>
<td>1755445-1755681</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>ompA</em></td>
<td>Outer membrane protein A</td>
<td>1018236-1019276</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>mrp</em></td>
<td>Antiporter inner membrane protein</td>
<td>2192190-2191081</td>
<td>1</td>
</tr>
<tr>
<td>Gene</td>
<td>Function</td>
<td>Start (Genomic Coordinates)</td>
<td>End (Genomic Coordinates)</td>
<td>Length</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------</td>
<td>----------------------------</td>
<td>---------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>flu (b2000)</td>
<td>CP4-44 prophage antigen 43, phase-variable biofilm</td>
<td>2069563-2072682</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ybhN</td>
<td>Conserved inner membrane protein</td>
<td>821721-820765</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ybhO</td>
<td>Cardiolipin synthase 2</td>
<td>822962-821721</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>yhhS</td>
<td>Predicted transporter</td>
<td>3609756-3608539</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noncoding</td>
<td>nc1 Homology to thrW</td>
<td>296361-296528</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc2 boxC repetitive element</td>
<td>497025-497169</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc3 Unknown</td>
<td>1276081-1276160</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc4 Unknown</td>
<td>3263241-3263360</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc5 BIME</td>
<td>4323897-4324000</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4324201-4324288</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nc6 Unknown</td>
<td>4564441-4564552</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc7 boxC repetitive element</td>
<td>1349125-1349315</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc8 Unknown</td>
<td>3668941-3669196</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc9 Unknown</td>
<td>28509-28620</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>nc10 Unknown</td>
<td>1202901-1203109</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 1. RIP-Chip assays: Representatives of immuneprecipitated RNAs; A) lpp B) ssrA, C) nc1, and D) nc5. “N”, stands for negative control (RNA from strain with no myc tag); IP, immunoprecipitated RNA; T, total RNA.
Figure 2. Northern blot analysis of nc1 and nc5 RNA. Wild type strain MG1655 was grown in M63 medium supplemented with 0.2% glucose, 0.1% casaminoacids and 0.002% vitamin B1 at 37°C and samples were removed at OD₆₀₀ of 0.5. RNA was extracted by hot phenol method, fractionated on a 10% TBE-urea gel, and transferred to a nylon membrane. 5’-biotinylated probes were used to detect (A) nc1 and (B) nc5. Ambion RNA Century markers were used to determine the size of RNA.
Figure 3. HU binding to nc5 RNA. (A) HU protein at the concentrations indicated (in nM) was mixed with 15 fmol nc5 RNA (88 nt-long) synthesized in an in vitro T7 polymerase reaction. The incubation of HU with nc5 RNA was performed under high salt conditions and analyzed by PAGE (see Materials and Methods). (B) Quantification of HU binding to nc5 RNA.
Figure 4. The predicted RNA fold of nc5 RNA. The secondary structure of nc5 RNA was predicted using Mfold Web Server (67). The hairpin which corresponds to PU Z2 motif found in BIME DNA (3, 15) is indicated. Positions of mismatches in Z2 hairpin are indicated with arrows.
Figure 5. Transmission electron microscopy (TEM) analysis of nucleoid morphology of: (A, B) wild-type *E. coli* MG1655; (C, D) deletion of *hupA* and *hupB* genes.
Figure 6. Transmission electron microscopy (TEM) analysis of nucleoid morphology of: (A, B) deletion of \( nc_1 \); (C, D) deletion of \( nc_1 \ hupA \ hupB \) genes.
Figure 7. Effects of nc5 on nucleoid morphology. Images of TEM analysis of: (A, B) deletion of nc5; (C, D) deletion of nc5 and hupA hupB genes.
Figure 8. Location of DNA sequences homologous to nc5 RNA. The circular *E. coli* chromosome. Red bars: 121 genomic sequences, BIMEs, having more than 50 nt identical to nc5 RNA (see Figure 1). The positions of the origin and the terminus of (6) replication are indicated (*oriC* and *terC* respectively).