

Proteomic Analyses of Nucleoid-Associated Proteins in *Escherichia coli*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus*

Ryosuke L. Ohniwa*, Yuri Ushijima, Shinji Saito, Kazuya Morikawa

Institute of Basic Medical Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tennodai, Tsukuba, Ibaraki, Japan

Abstract

Background: The bacterial nucleoid contains several hundred kinds of nucleoid-associated proteins (NAPs), which play critical roles in genome functions such as transcription and replication. Several NAPs, such as Hu and H-NS in *Escherichia coli*, have so far been identified.

Methodology/Principal Findings: Log- and stationary-phase cells of *E. coli*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus* were lysed in spermidine solutions. Nucleoids were collected by sucrose gradient centrifugation, and their protein constituents analyzed by liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS). Over 200 proteins were identified in each species. Envelope and soluble protein fractions were also identified. By using these data sets, we obtained lists of contaminant-subtracted proteins enriched in the nucleoid fractions (csNAP lists). The lists do not cover all of the NAPs, but included Hu regardless of the growth phases and species. In addition, the csNAP lists of each species suggested that the bacterial nucleoid is equipped with the species-specific set of global regulators, oxidation-reduction enzymes, and fatty acid synthases. This implies bacteria individually developed nucleoid associated proteins toward obtaining similar characteristics.

Conclusions/Significance: Ours is the first study to reveal hundreds of NAPs in the bacterial nucleoid, and the obtained data set enabled us to overview some important features of the nucleoid. Several implications obtained from the present proteomic study may make it a landmark for the future functional and evolutionary study of the bacterial nucleoid.

Citation: Ohniwa RL, Ushijima Y, Saito S, Morikawa K (2011) Proteomic Analyses of Nucleoid-Associated Proteins in *Escherichia coli*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus*. PLoS ONE 6(4): e19172. doi:10.1371/journal.pone.0019172

Editor: Sue Cotterill, St. Georges University of London, United Kingdom

Received: September 16, 2010; **Accepted:** March 29, 2011; **Published:** April 26, 2011

Copyright: © 2011 Ohniwa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a grant provided by The Ichiro Kanehara Foundation (to K.M.), by a Grant-in-Aid for Young Scientists of the Japan Society for the Promotion of Science (to R.L.O.), and by a Collaborative Research (B) grant from the National Institute of Genetics, Japan (to R.L.O.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ohniwa@md.tsukuba.ac.jp

Introduction

The genomes of all living organisms are packed in cells with proteins involved in various cellular processes such as transcription and replication. Most bacteria have circular genomes of various sizes (*Staphylococcus aureus*: 1 mm; *Bacillus subtilis*: 1.4 mm; *Escherichia coli*: 1.6 mm; *Pseudomonas aeruginosa*: 2.1 mm; per genome [estimated according to a base pair size of 0.34 nm]), which are packed into cells of a few micrometers in the form of a “nucleoid” [1,2].

Over 300 protein species are expected to be associated with the nucleoids isolated from *E. coli* and *B. subtilis*, although most of them have yet to be identified [3–6]. In the case of *E. coli*, several proteins have been identified as major nucleoid-associated proteins (NAPs): heat-unstable nucleoid protein (HU), integration host factor (IHF) (Hu paralogue), histone-like nucleoid structuring protein (H-NS), factor for inversion stimulation (Fis), host factor for RNA phage Q β replication (Hfq), suppressor of T4 *td* mutant phenotype A (StpA) (H-NS homologue), and DNA-binding protein from starved cells (Dps) [7,8]. These proteins occupy wide portions of genomic DNA [9,10] and are involved in a series of genome functions, such as

transcription (Hu, IHF, H-NS, StpA, and Fis) [11–15], translation (Hu, HNS, StpA, and Hfq) [16–19], replication (HU, IHF, Fis, and Dps) [20–23], DNA protection (Dps) [24,25], and DNA packing (Hu, H-NS, Fis, and Dps) [26–30].

These NAPs are not quantitatively static throughout growth [31]. Hu and Fis are abundant in the log phase but decrease toward the stationary phase. Conversely, IHF and Dps are induced toward the stationary phase and become the major components. Under anaerobic conditions, DNA-binding protein in anaerobic conditions (DAN) becomes the most abundant NAP [32]. In addition, comparative genomic analysis revealed that these *E. coli* NAPs are not commonly conserved in the bacterial kingdom [33,34] (see Table 1). Fis, H-NS, and StpA are present only in the Gammaproteobacteria. About half of all bacteria lack Hfq. Chlamydia and some of the Proteobacteria, Actinobacteria, and Firmicutes lack Dps. Even the most broadly conserved NAP, Hu or its homologue IHF, is absent in *Leptospira interrogans* and *Corynebacterium diphtheriae*. Thus, the features of the major NAPs of *E. coli* suggest the necessity of a comparative study of NAPs to gain insight into the general/specific characteristics of the bacterial nucleoid.

Table 1. Major NAPs Identified in This Study.

		<i>hupA</i>	<i>hupB</i>	<i>ihfA</i>	<i>ihfB</i>	<i>hns</i>	<i>stpA</i>	<i>fis</i>	<i>hfq</i>	<i>dps</i>	
log phase	<i>E. coli</i>	Protein amount*	55,000		10,000	20,000	25,000	60,000	57,000	7,500	
	<i>E. coli</i>	Nucleoid fraction ⁺	47.63	3.3	0.35	-	1.45	3.89	0.36	0.84	0.46
		Envelop fraction ⁺	-	-	-	-	-	0.36	-	0.21	
		Top fraction ⁺	-	-	-	0.56	-	-	0.84	0.46	
	<i>P. aeruginosa</i>	Nucleoid fraction ⁺	0.42	1.09	-	-	×	0.34	1.09	-	
		Envelop fraction ⁺	-	-	-	-	×	-	-	-	
		Top fraction ⁺	-	-	-	-	×	-	-	-	
	<i>B. subtilis</i>	Nucleoid fraction ⁺	-	0.98	×	×	×	×	×	-	
		Envelop fraction ⁺	-	-	×	×	×	×	×	-	
		Top fraction ⁺	-	-	×	×	×	×	×	-	
	<i>S. aureus</i>	Nucleoid fraction ⁺	-	0.42	×	×	×	×	×	-	
		Envelop fraction ⁺	-	-	×	×	×	×	×	-	
		Top fraction ⁺	-	-	×	×	×	×	×	-	
stationary phase	<i>E. coli</i>	Protein amount*	15,000		28,000	6,500	9,000	0	17,500	160,000	
	<i>E. coli</i>	Nucleoid fraction ⁺	0.42	-	0.35	0.89	0.96	0.25	-	-	0.46
		Envelop fraction ⁺	-	-	-	-	-	-	-	0.21	
		Top fraction ⁺	-	-	-	-	-	-	-	0.76	
	<i>S. aureus</i>	Nucleoid fraction ⁺	-	1.85	×	×	×	×	×	-	
		Envelop fraction ⁺	-	-	×	×	×	×	×	-	
		Top fraction ⁺	-	-	×	×	×	×	×	-	

*The amount of molecules were determined according to Azam et al, 1999 (molecules/cell).

+emPAI values.

- Not detected; × Absence of gene.

doi:10.1371/journal.pone.0019172.t001

The biochemical methods for purifying the cell membrane and cell wall fractions have been established [35,36], and the total proteins in such purified fractions have been identified by mass spectrometry techniques (reviewed in [37]). However, difficulty in isolating nucleoids remains. The isolated nucleoid under physiological salt conditions always includes contaminants derived from the cell membrane, cell wall, and cytosol [4,5,38–43]. Treatment with high salt and/or RNase can disassociate the contaminated proteins, but the number of proteins identified in this manner was limited [44–46].

In this study, we identified the proteins in the isolated spermidine nucleoids of *E. coli*, *P. aeruginosa*, *B. subtilis*, and *S. aureus* by liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS). Nucleoids isolated in a spermidine solution with mild ionic strength retain the most proteins directly/indirectly bound to the genomic DNA [5,39]. The comparison of proteins detected in the nucleoid and envelope and soluble fractions suggested some characteristics of the bacterial nucleoid. We discuss several implications yielded by the obtained NAP data sets regarding the possible functional and evolutionary aspects of the bacterial nucleoid.

Results

Identification of NAPs in *E. coli*

We isolated spermidine nucleoids from *E. coli* cells grown in aerobic conditions (see Figure S1 for growth curves and sampling points). DNA content was monitored for each sucrose gradient fraction (Figure 1A, 1B, 1D and 1E), and the one with the highest DNA content was further analyzed as the nucleoid fraction. In the

case of the log phase, multiple peaks were sometimes observed, probably owing to the viscous characteristics of the nucleoid (Figure 1A-1B). It might be because of the heterogeneous density of nucleoids: e.g. existence of the stationary-phase type nucleoid (Figure 1E). However, analysis by SDS-PAGE showed that the protein signal patterns of fractions 1 and 2 were indistinguishable (Figure 1C). The upper fraction (fraction 1) was analyzed by LC-MS/MS. We also analyzed proteins in the envelope and top fractions, as described in the Materials and Methods section (Figure 1C and 1F). The numbers of the identified proteins and peptides are summarized in Figure 2A to 2C. The full list of identified proteins with the number of corresponding peptides is provided in the supplementary tables (Table S1, S2, and S3).

The number of the detected peptides was correlated with the number of identified proteins (Figure 2B). This implied that the identified protein sets with lower numbers of peptides were not sufficient to cover all the proteins in the given fractions. By using *in silico* simulation, we estimated the coverage rates, which represent the extent of the identified proteins out of the total proteins in the fractions (Figure 2D, see Materials and Methods). Here, we applied 3 different models representing the amounts of individual proteins in the fractions (Figure S2). The constant model represents the uniform distribution of protein amounts, and the linear model represents the linear decrease of the distribution. The simplified canonical law (SCL) model has been reported to be consistent with the relative expression of proteins in prokaryotic cells [47]. In the case of the nucleoid fractions of the log phase, the coverage rates of the identified protein species according to these models were estimated as 0.63 to 0.93, and the rates of protein amounts covered by the identified proteins as 0.88 to 0.96. In the

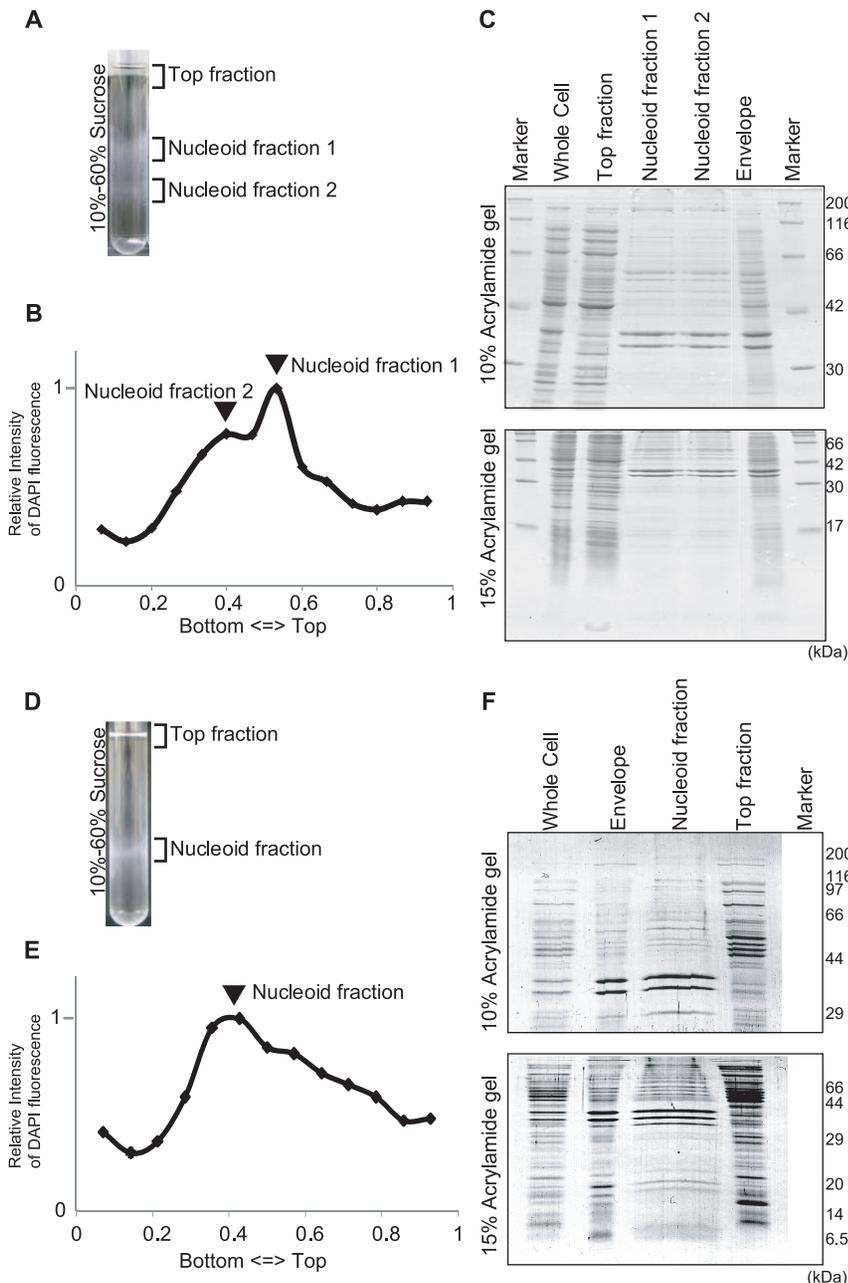


Figure 1. *E. coli* nucleoid isolation. Isolation of spermidine nucleoids in the log (A to C) and stationary (D to F) phases. (A) (D) The lysed cells were fractionated by sucrose-gradient centrifugation with a 10%-to-60% sucrose gradient. (B) (E) The relative DNA amount in the sucrose gradient detected by DAPI fluorescence. (C) (F) SDS-PAGE of the whole-cell lysates, the top fraction of the sucrose gradient, the nucleoid fractions, and the envelope fraction. The gels were stained with Coomassie Brilliant Blue (CBB). The pattern of the envelope in the log phase was similar to that reported by Lai et al [89].

doi:10.1371/journal.pone.0019172.g001

case of the stationary phase, the rates of the species and amounts were 0.82 to 0.94 and 0.92 to 0.97, respectively. Thus, it can be expected that over 60% of the protein species that occupy at least 80% of the protein amounts out of the total proteins in the nucleoid fractions were identified in *E. coli*. The rates of the envelope and top fractions were 0.43 to 0.96 for the protein species, and 0.71 to 0.98 for the protein amounts. These data suggest that we successfully identified at least 70% of protein amounts although there may remain various unidentified protein

species in each fraction. In Table S4, we summarized the potential number of protein species in each fraction and the numbers of the peptides required for identifying all of them. The number of peptides required to identify all the proteins in each fraction were estimated based on the SCL model (Table S4).

Major NAPs were identified in the nucleoid fractions (Table 1). Hu was detected in the nucleoid fraction in both the log and stationary phases, consistent with the Western blot analysis against Hu (Figure 3). In the nucleoid fraction of the log phase, Hu (*hupA*

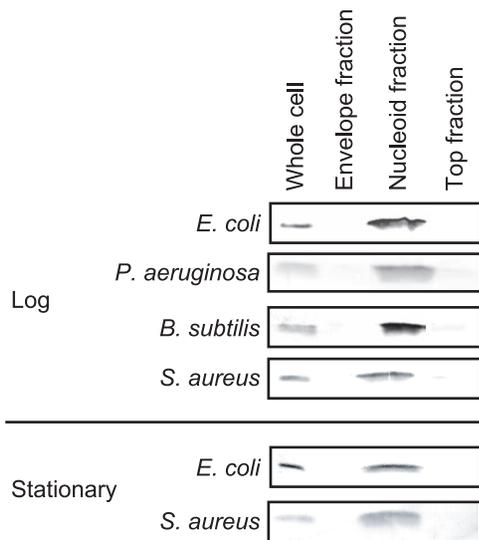


Figure 3. Western blots against Hu. Five micrograms of proteins from the whole-cell lysates, the nucleoid fractions, the envelope fractions, and the top fractions were separated by SDS-PAGE, and Hu was detected by Western blotting. ‘Log’ and ‘Stationary’ represent the log phase and stationary phase, respectively.
doi:10.1371/journal.pone.0019172.g003

In addition to the major NAPs, various DNA and RNA binding proteins were identified in the nucleoid fractions (Table 2). Twenty-six transcription factors were identified in the log phase, and three in the stationary phase. RNA polymerases and ribosomal proteins, which have been identified in spermidine nucleoids [5], were also identified in both phases. Other proteins related to translation (*tsf*, *infC* etc.), replication (*seqA*, *topA* etc.), and DNA repair (*mutS*, *wvrA* etc.) were also included. Non-DNA binding proteins should be included in the isolated nucleoids. For example, it is known that transcription elongation factors Rho (*rho*) and NusA (*nusA*) interact with RNA polymerase, and, indeed, our list included them.

The nucleoid fractions were enriched in the envelope and cytosolic proteins (Figure 4, Table S2, and Table S3). When the proteins in the nucleoid fractions were sorted by emPAI value, 18 out of the top 40 proteins were envelope and cytosolic proteins in the log phase (Figure 4A). These include F₀F₁-ATPase (coded by *atpA*, *atpD*, etc.), porins (*ompA*, *ompC*, etc.), flagellin (*fljC*, etc.), a chaperone (*groEL*), and metabolic enzymes (*tnaA*, *nuoC*, etc.). In the stationary phase, 31 out of the 40 proteins were envelope or cytosolic proteins including F₀F₁-ATPase, porins, flagellin, a chaperone, and metabolic enzymes (Figure 4B). Most of these proteins were also identified in the envelope and/or top fractions with high emPAI values, suggesting that these proteins in the nucleoid fractions are contaminated largely owing to the difficulty of nucleoid isolation (see Discussion).

NAPs in *P. aeruginosa*, *B. subtilis*, and *S. aureus*

We investigated the NAPs of *P. aeruginosa*, *B. subtilis*, and *S. aureus*. *P. aeruginosa* is Gram-negative and, like *E. coli*, belongs to the Gammaproteobacteria. *B. subtilis* and *S. aureus* are both Gram-positive and belong to the Firmicutes/Bacillales. The number of major *E. coli* NAP genes varies depending on the species (Table 1). *P. aeruginosa* possesses *hupA*, *hupB*, *ihfA*, *ihfB*, *fis*, *hfq*, and *dps*, but not *hns* and *stpA* (but possesses *mvaT* and *mvaU* as functional counterparts of *hns* and *stpA* [49]). *B. subtilis* possesses only *hupB*

(annotated as *hbs*), its homologues, *yonN*, and *dps*. *S. aureus* possesses *hupB* (*hu*) and *dps* (*mrgA*) but lacks *fis*, *hns*, *stpA*, and *hfq*.

Proteins in the spermidine-nucleoid fractions, as well as in the envelope and top fractions, were isolated (Figure 5) and identified by LC-MS/MS (Figure 2, Figure S3, Tables S1 and S5, S6, S7, S8). Over 200 proteins were identified in each nucleoid fraction. The coverage rates of the proteins in the nucleoid fractions were 0.60 to 0.92 as protein species and 0.73 to 0.96 as protein amounts (Figure 2D), suggesting that major part, but not all, of the nucleoid proteins were identified (Table S4). Hu was exclusively detected in the nucleoid fractions of all species tested. This was consistent with the Western blots against Hu (Figure 3). In the case of *P. aeruginosa*, Hu (*hupA* and *hupB*), Fis (*fis*), and Hfq (*hfq*) were detected in the nucleoid fraction, but IHF (*ihfA* and *ihfB*), Dps (*dps*), MvaT (*mvaT*), and MvaU (*mvaU*) were not. In the case of *B. subtilis* and *S. aureus*, HU (*hbs*, *yonN*, and *hu*) was identified, but Dps (*dps* and *mrgA*) was not. DNA and RNA binding proteins other than ribosomal proteins are listed in Table 2. In all species tested, as in *E. coli*, many ribosomal proteins, envelope proteins, and cytosolic proteins were identified in the nucleoid fractions (Table S1 and Table S5, S6, S7, S8).

Data analysis: contaminant-subtracted NAPs (csNAPs)

The isolated nucleoids contained various envelope and cytosolic proteins. This was the expected result, as described in the Introduction section. To deduce reasonable considerations in the Discussion section, we here define “contaminant-subtracted NAPs (csNAPs)” as follows:

csNAPs = “Proteins detected only in the nucleoid fraction” + “Proteins calculated to be relatively abundant in the nucleoid fraction.”

- “Proteins detected only in the nucleoid fraction”: Proteins detected only in the nucleoid fraction in a given condition.
- “Proteins calculated to be relatively abundant in the nucleoid fraction”: The number of peptides detected by LC-MS/MS is a good benchmark to investigate the fraction in which the target protein is dominantly present. If the number of peptides of a certain protein identified in the nucleoid fraction is larger than that of the other fractions, the protein is thought to be abundant in the nucleoid [50]. For instance, in the log phase of *E. coli*, 7 Fis peptides were detected in the nucleoid fraction, and 2 in the envelope fractions, suggesting that Fis was more abundant in the nucleoid fraction. Here, we need to pay attention to the total number of peptides detected by the LC-MS/MS because more peptides should be detected if more sample is loaded for the LC-MS/MS [51]. In the case of the log phase of *E. coli*, 7148 peptides from 401 proteins were detected in the nucleoid fraction, and 13657 peptides from 334 proteins in the envelope fraction (Figure 2). Therefore, the relative abundances of Fis peptides in the log phase of *E. coli* were 9.8×10^{-4} ($7/7148$) and 1.5×10^{-4} ($2/13657$) in the nucleoid fraction and the top fraction, respectively, and the ratio was 6.7 ($9.8 \times 10^{-4} / 1.5 \times 10^{-4}$). We arbitrarily selected proteins with a ratio higher than 3 as csNAPs.

According to the above-mentioned criterion, 164, 66, 98, and 92 proteins were selected as csNAPs from the log phases of *E. coli*, *P. aeruginosa*, *B. subtilis*, and *S. aureus*, respectively (Figure 2, Figure 6, 7, 8, and Table S9, S10, S11, S12). From the stationary phase, 76 and 141 proteins were selected from *E. coli* and *S. aureus*, respectively (Figure 2, Figure 6, 7, 8, and Table S13-S14).

Various major envelope and cytosolic proteins were excluded in this operation. For instance, in the log phase of *E. coli*, all the

Table 2. DNA and RNA Binding Proteins Identified in the Nucleoid Fractions (except major NAPs).

Category	Phase	Species	Genes
Transcription factors	log phase	<i>E. coli</i>	<u>arcA*</u> , <u>ascG</u> , <u>chbR*</u> , <u>crl*</u> , <u>crp*</u> , <u>cspE*</u> , <u>fruR*</u> , <u>fur*</u> , <u>idnR*</u> , <u>lact*</u> , <u>lldR*</u> , <u>lrp*</u> , <u>malT*</u> , <u>osmE*</u> , <u>ompR*</u> , <u>phoU*</u> , <u>pspB*</u> , <u>purR*</u> , <u>putA*</u> , <u>rtcR*</u> , <u>srlR*</u> , <u>ybaD</u> , <u>yciT*</u> , <u>ydeW</u> , <u>ydfH*</u> , <u>yheO*</u>
		<i>P. aeruginosa</i>	<u>crp</u> , <u>fabG</u> , <u>yebK</u> , <u>yhbY</u>
		<i>B. subtilis</i>	<u>degA</u> , <u>gabR</u> , <u>gntR</u> , <u>kdgR</u> , <u>rsfA</u> , <u>spo0J</u> , <u>xylR</u> , <u>ybbB</u> , <u>ydiP</u> , <u>yhdQ</u> , <u>yplP</u>
	stationary phase	<i>S. aureus</i>	<u>codY</u> , <u>fabG</u> , <u>graR</u> , <u>rex</u> , <u>rot</u> , <u>sarA</u> , <u>sarH1</u> , <u>sarR</u> , <u>spxA</u> , <u>srrA</u> , <u>vraR</u>
		<i>E. coli</i>	<u>osmE*</u> , <u>lysR*</u> , <u>putA*</u>
		<i>S. aureus</i>	<u>ahrC</u> , <u>codY</u> , <u>fabG</u> , <u>graR</u> , <u>mgrA</u> , <u>nreC</u> , <u>pyrR</u> , <u>rocA</u> , <u>saeR</u> , <u>sarA</u> , <u>sarR</u> , <u>sarH1</u> , <u>sarV</u> , <u>sarZ</u> , <u>srrA</u> , <u>tcaR</u> , <u>vraR</u> , <u>vicR</u>
Proteins involved in transcription, translation, replication, and DNA repair	log phase	<i>E. coli</i>	<u>deaD</u> , <u>fusA</u> , <u>treA</u> , <u>hrpA</u> , <u>hdsS</u> , <u>infB</u> , <u>infC</u> , <u>insB</u> , <u>insL</u> , <u>mfd</u> , <u>mutS</u> , <u>nusA</u> , <u>nusG</u> , <u>parC*</u> , <u>pcnB</u> , <u>pnp</u> , <u>polA*</u> , <u>rdgC</u> , <u>rhlB</u> , <u>rho</u> , <u>rimM</u> , <u>rimB</u> , <u>rmuC</u> , <u>rnb</u> , <u>rne</u> , <u>rpoA*</u> , <u>rpoB*</u> , <u>rpoC*</u> , <u>rpoD*</u> , <u>rusA</u> , <u>seqA*</u> , <u>srmB</u> , <u>topA*</u> , <u>tsf</u> , <u>tufB</u> , <u>unvrA*</u> , <u>xthA</u> , <u>yejK</u> , <u>yjgJ</u>
		<i>P. aeruginosa</i>	<u>efp</u> , <u>fusA</u> , <u>gyrA</u> , <u>gyrB</u> , <u>infB</u> , <u>infC</u> , <u>mutS</u> , <u>nusG</u> , <u>parC</u> , <u>parE</u> , <u>pnp</u> , <u>rdgC</u> , <u>recA</u> , <u>rhlB</u> , <u>rho</u> , <u>rne</u> , <u>rpoA</u> , <u>rpoB</u> , <u>rpoC</u> , <u>rpoD</u> , <u>ssb</u> , <u>topA</u> , <u>tsf</u> , <u>tufA</u> , <u>uvrA</u> , <u>yjgK</u>
		<i>B. subtilis</i>	<u>fusA</u> , <u>gidB</u> , <u>gyrA</u> , <u>gyrB</u> , <u>ihfB</u> , <u>ihfC</u> , <u>mutL</u> , <u>mutS</u> , <u>nusA</u> , <u>pnpA</u> , <u>polA</u> , <u>polC</u> , <u>rpoA</u> , <u>rpoB</u> , <u>rpoC</u> , <u>sigF</u> , <u>smc</u> , <u>tsf</u> , <u>tuf</u> , <u>uvrA</u> , <u>uvrC</u> , <u>ydbR</u> , <u>yhaM</u> , <u>yhcR</u> , <u>yirY</u> , <u>yqfR</u> , <u>yqjW</u>
	stationary phase	<i>S. aureus</i>	<u>fus</u> , <u>efp</u> , <u>tsf</u> , <u>tufA</u> , <u>end4</u> , <u>ermA</u> , <u>infA</u> , <u>nusG</u> , <u>pnpA</u> , <u>recA</u> , <u>rcn</u> , <u>rnh3</u> , <u>rpoA</u> , <u>rpoB</u> , <u>rpoC</u> , <u>rpoE</u> , <u>uvrC</u> , <u>xerD</u>
		<i>E. coli</i>	<u>hrpA</u> , <u>rdgC</u> , <u>rpoA*</u> , <u>rpoB*</u> , <u>rpoC*</u> , <u>rpoD*</u> , <u>rpoZ*</u> , <u>rusA</u> , <u>ruvA</u> , <u>topA*</u> , <u>tsf</u> , <u>tufB</u> , <u>uvrA*</u> , <u>yjgK</u>
		<i>S. aureus</i>	<u>lig</u> , <u>dnaN</u> , <u>fus</u> , <u>efp</u> , <u>tsf</u> , <u>tufA</u> , <u>gyrB</u> , <u>hdsR</u> , <u>infA</u> , <u>infB</u> , <u>infC</u> , <u>mfd</u> , <u>nusG</u> , <u>parC</u> , <u>parE</u> , <u>pnpA</u> , <u>rcn</u> , <u>rnj1</u> , <u>rnj2</u> , <u>rpoA</u> , <u>rpoB</u> , <u>rpoC</u> , <u>rpoE</u> , <u>gidB</u> , <u>ruvA</u> , <u>ssb</u> , <u>topA</u> , <u>Y1885</u>

Underlined genes represent the proteins selected as contaminant-subtracted NAPs (csNAPs).

*Genes reported as DNA-binding proteins in EcoSal [1].

doi:10.1371/journal.pone.0019172.t002

porins (*ompA*, *ompC*, *ompF*, *ompN*, *ompT*, and *ompX*), most of the flagellar components (*fliC*, *fliF*, *fliL*, and *fliM* in both phases, but not *fliO* in the log phase), and many of the F₀F₁-ATPase subunits (*atpA*, *atpD*, *atpE*, *atpF*, and *atpG* in both phases, but not *atpB* and *atpH* in the log phase), chaperones (*groEL*, *groES*, *tig*, etc.), and metabolic enzymes (*tnaA*, *nucC*, etc.) were excluded.

In contrast, the major NAPs, such as Hu (*hupA* and *hupB*), IHF (*himA* and *himB*), Fis (*fis*), and StpA (*stpA*), were judged to be csNAPs. H-NS (*hns*) and Hfq (*hfq*) were not included in the csNAPs in the log phase because they were also detected with high peptide numbers in the envelope and/or top fraction.

Discussion

The csNAPs were selected as those proteins that were relatively abundant in the nucleoid fraction as compared with in the envelope and top fractions. It should be noted that the list of csNAPs is an incomplete one. Indeed, various proteins known to be included in the nucleoid (eg, ribosomal proteins, RNA polymerases, and some major NAPs such as HNS and Hfq) were eliminated from the csNAPs list. Nevertheless, it is reasonable to expect that the proteins selected as csNAPs are indeed involved in the functions of the bacterial nucleoid. Below, we discuss the csNAPs to gain insight into some of the characteristics of the bacterial nucleoid.

Characteristics of csNAPs

When csNAPs of *E. coli* in the log phase are sorted according to the emPAI values that roughly represent the amount of proteins in the mixture [48], various DNA and/or RNA binding proteins, such as Hu (*hupA*, *hupB*) and StpA (*stpA*), are included in the top 10 *E. coli* csNAPs (Figure 6A). These proteins are

global transcription/translation regulators that facilitate the response to various environmental changes [12–14]. In addition, other global regulators, such as CspE (*cspE*) [52], Lrp (*lrp*) [53], Fis (*fis*), and IHF (*ihfA*), were ranked in the top 20 *E. coli* csNAPs. In the stationary phase, global regulators including HNS (*hns*), IHF (*ihfA*, *ihfB*), Hu (*hupA*), and StpA (*stpA*) were involved in the top 20 (Figure 6B). Global regulators were also abundant in the csNAPs of *P. aeruginosa* (Figure 7A): The top 20 csNAPs of this bacterium included 5 global regulators, Hu (*hupA*, *hupB*), Hfq (*hfq*), Fis (*fis*), and HexR (*hexR*) [54,55]. Thus, the csNAPs in the Gammaproteobacteria comprise a plentiful amount of global regulators. Other than Hu, such global regulator genes are not genetically conserved in *B. subtilis* and *S. aureus* (Firmicutes/Bacillales). Recently, staphylococcal accessory regulator A (SarA) and its homologues in *S. aureus* were proposed as global regulators [56]. Our csNAP list of *S. aureus* contains SarA (*sarA*) and its homologues SarR (*sarR*) and SarH1 (*sarH*) in both the log and stationary phases (Figure 8). In addition, the stationary-phase csNAPs include SarV (SA2062) and SarZ (SA2174).

Many functionally unknown proteins (y-genes) were also included with high emPAI values in the lists, especially in the *B. subtilis* list (Figure 7B). The prediction of DNA/RNA binding abilities showed that several y-genes had strong potential to bind DNA and/or RNA (Figure 7B, Figure S4, and Materials and Methods). There were 10 y-genes in the top 30 csNAPs of *B. subtilis*, and all of them encoded predicted DNA-binding (*yfII*, *yneK*, and *ydhF*) or RNA-binding (*yneK*, *yhbF*, and *ydhF*) proteins. It is possible that these genes encode novel global regulators.

An additional feature of csNAPs is the abundance of enzymes required for stress responses. For instance, in the log phase of *S. aureus*, the top 30 csNAPs included 3 reductase-like superoxide dismutases, *sodA*, alkyl hydroperoxide reductase (*ahpC*), and

A *E. coli* in log phase

ID	gene	Annotation	Nucleoid fraction		Envelope fraction		Top fraction	
			emPAI	pep_num	emPAI	pep_num	emPAI	pep_num
ecj:JW3964	<i>hupA</i>	DNA-binding protein HU-alpha	47.63	23				
ecj:JW3271	<i>rplX</i>	50S ribosomal protein L24	7.25	11			14.08	18
ecj:JW2203	<i>ompC</i>	Outer membrane protein C	6.75	68	19.63	135	0.19	7
ecj:JW1908	<i>fljC</i>	Flagellin	4.03	49	4.79	84	0.52	15
ecj:JW2644	<i>stpA</i>	DNA-binding protein stpA	3.89	13				
ecj:JW3264	<i>rpmD</i>	50S ribosomal protein L30	3.44	7	0.64	6		
ecj:JW0430	<i>hupB</i>	DNA-binding protein HU-beta	3.3	9				
ecj:JW0940	<i>ompA</i>	Outer membrane protein A	3.22	50	13.71	103		
ecj:JW0554	<i>ompT</i>	Protease 7	3.1	24	2.7	24		
ecj:JW1633	<i>slyB</i>	Outer membrane lipoprotein slyB	2.79	11	0.95	16		
ecj:JW3712	<i>atpA</i>	ATP synthase subunit alpha	2.68	39	4.8	76	0.58	10
ecj:JW3714	<i>atpF</i>	ATP synthase subunit b	2.36	15	5.15	22		
ecj:JW1667	<i>lpp</i>	Major outer membrane lipoprotein	2.32	23	15.48	23		
ecj:JW0912	<i>ompF</i>	Outer membrane protein F	2.27	33	7.15	53	0.1	2
ecj:JW3710	<i>atpD</i>	ATP synthase subunit beta	2.15	42	8.84	104		
ecj:JW3281	<i>rplD</i>	50S ribosomal protein L4	2.06	17	3.21	25	0.38	6
ecj:JW3282	<i>rplC</i>	50S ribosomal protein L3	2.04	16	1.21	9	0.89	9
ecj:JW4161	<i>rplI</i>	50S ribosomal protein L9	2.02	15	20.98	47	3.69	34
ecj:JW3152	<i>rpmA</i>	50S ribosomal protein L27	1.99	4			0.44	6
ecj:JW3303	<i>rpsG</i>	30S ribosomal protein S7	1.87	15	1.02	15	1.87	18
ecj:JW3265	<i>rpsE</i>	30S ribosomal protein S5	1.71	12	43.17	50	1.22	14
ecj:JW3268	<i>rpsH</i>	30S ribosomal protein S8	1.65	18	0.63	18	1.65	7
ecj:JW2462	<i>nlpB</i>	Lipoprotein 34	1.64	29	0.97	15		
ecj:JW0799	<i>ompX</i>	Outer membrane protein X	1.56	9	1.56	13		
ecj:JW3946	<i>rplK</i>	50S ribosomal protein L11	1.54	13	1.54	14	2.21	11
ecj:JW3258	<i>rpsD</i>	30S ribosomal protein S4	1.46	22	0.82	13	0.16	8
ecj:JW1225	<i>hns</i>	DNA-binding protein H-NS	1.45	9			0.56	4
ecj:JW3153	<i>rplU</i>	50S ribosomal protein L21	1.43	9	2.27	7	1.43	8
ecj:JW2640	<i>ygaU</i>	Uncharacterized protein ygaU	1.38	11				
ecj:JW3947	<i>rplA</i>	50S ribosomal protein L1	1.36	17	2.14	38	0.54	7
ecj:JW0165	<i>tsf</i>	Elongation factor Ts	1.26	22	0.6	19	0.79	14
ecj:JW4103	<i>groL</i>	60 kDa chaperonin	1.26	40	0.99	49	0.46	21
ecj:JW5448	<i>ygdI</i>	Uncharacterized lipoprotein ygdI	1.23	3	1.23	5		
ecj:JW2587	<i>rplS</i>	50S ribosomal protein L19	1.2	8	0.69	7	1.2	9
ecj:JW3711	<i>atpG</i>	ATP synthase gamma chain	1.2	20	0.25	20		
ecj:JW3272	<i>rplN</i>	50S ribosomal protein L14	1.14	10	1.14	13	1.75	10
ecj:JW3686	<i>tnaA</i>	Tryptophanase	1.11	32	0.84	34	1.26	21
ecj:JW3304	<i>rpsL</i>	30S ribosomal protein S12	1.1	6	0.28	6	0.64	8
ecj:JW5375	<i>nuoC</i>	NADH-quinone oxidoreductase subunit C/D	1.08	40	1.44	47	0.23	12
ecj:JW2590	<i>rnsP</i>	30S ribosomal protein S16	1.07	7	1.07	7	1.99	6

B *E. coli* in stationary phase

ID	gene	Annotation	Nucleoid fraction		Envelope fraction		Top fraction	
			emPAI	pep_num	emPAI	pep_num	emPAI	pep_num
ecj:JW0940	<i>ompA</i>	Outer membrane protein A	7.27	88	10.03	93	0.21	12
ecj:JW2203	<i>ompC</i>	Outer membrane protein C	6.75	66	3.54	50	0.31	10
ecj:JW1667	<i>lpp</i>	Major outer membrane lipoprotein	6.4	20	15.48	24	2.32	4
ecj:JW2640	<i>ygaU</i>	Uncharacterized protein ygaU	2.66	12	0.91	18		
ecj:JW0912	<i>ompF</i>	Outer membrane protein F	2.27	23	1.07	17		
ecj:JW0397	<i>yajC</i>	UPF0092 membrane protein yajC	1.37	5				
ecj:JW3710	<i>atpD</i>	ATP synthase subunit beta	1.36	38	1.92	39	0.15	16
ecj:JW3257	<i>rpoA</i>	DNA-directed RNA polymerase subunit alpha	1.18	19	0.22	15	0.1	5
ecj:JW0554	<i>ompT</i>	Protease 7	1.02	18	0.65	12		
ecj:JW5539	<i>yhcB</i>	Putative cytochrome d ubiquinol oxidase subunit	1	9	0.26	7		
ecj:JW1225	<i>hns</i>	DNA-binding protein H-NS	0.96	10				
ecj:JW4108	<i>ecnB</i>	Entericidin B	0.9	2	0.9	2	0.9	3
ecj:JW0895	<i>ihfB</i>	Integration host factor subunit beta	0.89	11				
ecj:JW3714	<i>atpF</i>	ATP synthase subunit b	0.83	13	0.83	8	0.22	8
ecj:JW0799	<i>ompX</i>	Outer membrane protein X	0.76	8				
ecj:JW1728	<i>osmE</i>	Osmotically-inducible lipoprotein E	0.75	7	1.32	16		
ecj:JW1908	<i>fljC</i>	Flagellin	0.75	20	0.23	16	0.52	18
ecj:JW4132	<i>hflK</i>	Protein hflK	0.74	18	0.61	25		
ecj:JW0849	<i>ybjP</i>	Uncharacterized lipoprotein ybjP	0.73	9	0.44	8		
ecj:JW3713	<i>atpH</i>	ATP synthase subunit delta	0.72	11	0.44	9		
ecj:JW3160	<i>yrbD</i>	Uncharacterized protein yrbD	0.71	7	0.2	4		
ecj:JW4110	<i>blc</i>	Outer membrane lipoprotein blc	0.7	7			0.19	9
ecj:JW2276	<i>nuoI</i>	NADH-quinone oxidoreductase subunit I	0.65	13	0.4	14		
ecj:JW3474	<i>slp</i>	Outer membrane protein slp	0.65	5	0.4	6	0.18	5
ecj:JW3304	<i>rpsL</i>	30S ribosomal protein S12	0.64	6	0.28	9	0.64	7
ecj:JW3946	<i>rplK</i>	50S ribosomal protein L11	0.59	9	1.01	11	0.59	14
ecj:JW4158	<i>rpsF</i>	30S ribosomal protein S6	0.58	10	0.58	10	1.43	14
ecj:JW1200	<i>lolB</i>	Outer-membrane lipoprotein lolB	0.57	11	0.16	7		
ecj:JW0713	<i>sdhA</i>	Succinate dehydrogenase flavoprotein subunit	0.56	30	0.74	35		
ecj:JW1633	<i>slyB</i>	Outer membrane lipoprotein slyB	0.56	7	0.95	8		
ecj:JW4161	<i>rplI</i>	50S ribosomal protein L9	0.55	4	0.25	18	0.94	22
ecj:JW3200	<i>rplM</i>	50S ribosomal protein L13	0.55	12	0.24	7	0.55	8
ecj:JW4103	<i>groL</i>	60 kDa chaperonin	0.55	39	0.37	28	0.76	44
ecj:JW3864	<i>fdoH</i>	Formate dehydrogenase-O iron-sulfur subunit	0.52	10	0.37	9		
ecj:JW2283	<i>nuoA</i>	NADH-quinone oxidoreductase subunit A	0.52	5	0.52	7		
ecj:JW0422	<i>cyoA</i>	Ubiquinol oxidase subunit 2	0.51	22	0.36	22		
ecj:JW1606	<i>ydgA</i>	Protein ydgA	0.49	17	0.3	22	0.14	12
ecj:JW0714	<i>sdhB</i>	Succinate dehydrogenase iron-sulfur subunit	0.48	17	0.3	15		
ecj:JW3351	<i>damX</i>	Protein damX	0.48	9	0.26	9		
ecj:JW2500	<i>vfgA</i>	Cytoskeleton protein rodZ	0.48	8	0.1	3		

Figure 4. Proteins identified in the nucleoid fractions of *E. coli*. The listed proteins were the top 40 proteins sorted by the empAI values. The colors of the letters indicate the categories of the proteins (red: major NAPs; green: transcription factors; dark blue: DNA/RNA binding proteins involved in transcription, translation, replication, and DNA repair; light blue: ribosomal proteins; black: cytosolic-type proteins; and orange: envelope-type proteins). The major NAPs, transcription factors, and DNA/RNA binding proteins were classified according to EcoSal [1] and their annotations by KEGG [86]. The ribosomal proteins were based on gene annotations. The residual proteins were classified into cytosolic-type and envelope-type proteins according to the prediction of their intracellular localization. The localizations of *E. coli* proteins were predicted by EchoLOCATION [90]. Those of *P. aeruginosa*, *B. subtilis*, and *S. aureus* were predicted by PSORTb [91]. Underlined genes were reported as DNA binding proteins in EcoSal. pep_num represents the number of peptides detected from each protein.
doi:10.1371/journal.pone.0019172.g004

thioredoxine reductase (*trxA*), which have been reported to play important roles in coping with oxidative stress-responsive elements [57,58]. In all the species, the various enzymes involved in the oxidative stress responses were selected as csNAPs. A NAP having such property has been found in the nucleoid of the plant plastid. Sulfite reductase (SiR), the enzyme that catalyzes the reduction of sulfite to sulfide in the sulfur assimilation pathway (see review in [59]), has been identified as a DNA-binding protein of the plastid nucleoid [60]. Recently, it has been proposed that SiR is also involved in oxidative stress resistance [61]. Although the present study only suggested the possibility of association/interaction of these enzymes with the nucleoid, it is possible that the bacterial genomic DNA is, in general, protected by these enzymes from reactive oxygen species, which we term the ‘armor hypothesis’ (see also Future Perspective in Discussion).

Constitutive csNAPs through growth phases

The csNAPs in the log- and stationary-phase lists were different (Figure 9). In the list for *E. coli*, 4.3% of the csNAPs (10 out of 230 csNAPs [164 csNAPs in the log phase + 76 csNAPs in the stationary phase – 10 csNAPs present in both the log and stationary phases]) including Hu (*hupA*), StpA (*stpA*), and IHF (*ihfA*) were common between the log and stationary phases. In *S. aureus*, 15.9% of the csNAPs (32 out of 201 [92 in the log phase + 141 in the stationary phase – 32 present in both the log and stationary phases]) were common csNAPs throughout growth. One of the reasons for such a limited number of constitutive csNAPs could be the incomplete coverage of the proteins. Estimation of the real numbers of csNAPs on the basis of the coverage rates showed that

the number of *E. coli* csNAPs may be 88 to 225 in the log phase and 62 to 85 in the stationary phase, while those of *S. aureus* may be 92 to 130 in the log phase and 127 to 169 in the stationary phase (Table S15). If all of the additional csNAPs overlap between the log and stationary phases, 34.8% of the csNAPs will be common in *E. coli* (80 out of 230 [225 in the log phase + 85 in the stationary phase – 80 present in both phases]). Accordingly, 45.3% of *S. aureus* csNAPs might be common (91 out of 201). Although it is unlikely that all the additional csNAPs overlapped, these data suggest that additional common csNAPs are constitutively present on the nucleoid.

Even considering the unidentified csNAPs, over 50% of csNAPs were expected not to be common. One reason would be the on/off of their expression. In *E. coli* and *S. aureus*, 66.5% (153 out of 230) and 36.3% (73 out of 201) csNAPs, respectively, exhibited log- or stationary-phase specific expression. Another reason seems to be the growth-dependent changes in localization of the csNAPs. For instance, *E. coli* HNS was detected both in the nucleoid and top fractions in the log phase (thereby not selected as a csNAP), but only in the nucleoid fraction in the stationary phase. *S. aureus* superoxide dismutase (SodA) appeared in the nucleoid fraction only in the log phase and was detected in all the fractions in the stationary phase. The percentage of csNAPs that exhibited such growth-dependent changes was 28.7% (66 out of 230) in *E. coli* and 47.8% (96 out of 201) in *S. aureus*.

Little is known about how NAPs change in the process of growing. It should be noted that the growth-dependent structural change of the nucleoid is induced in *E. coli*, but not in *S. aureus* [28,33,34,62]. The increase in the amount of Dps and the

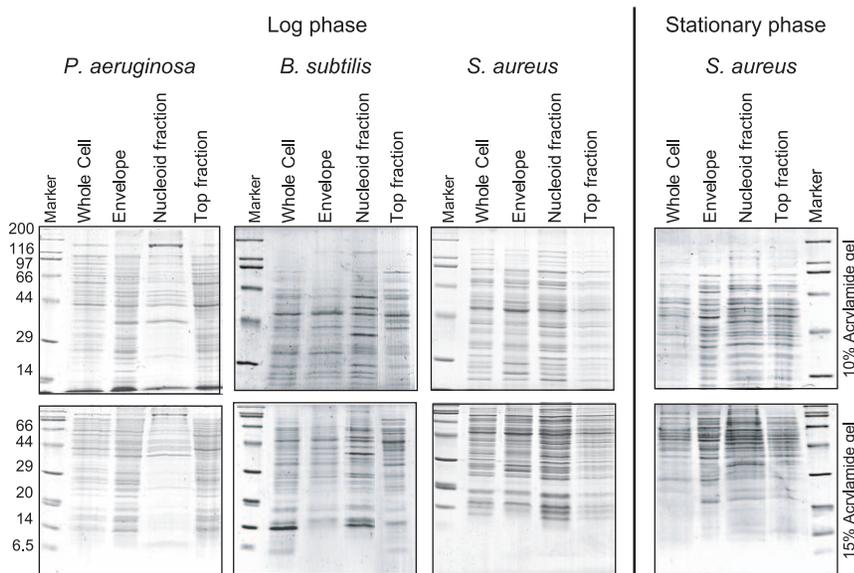


Figure 5. Proteins in the nucleoid, top, and envelope fractions of *S. aureus*, *P. aeruginosa*, and *B. subtilis*. SDS-PAGE analysis of the whole-cell lysates, envelope fractions, nucleoid fractions, and top fractions of the sucrose gradient assay. The gels were stained with CBB.
doi:10.1371/journal.pone.0019172.g005

A

<i>E. coli</i> in log phase							
ID	gene	Annotation	emPAI	pep_num	DBS	RBS	
ecj:JW3964	<i>hupA</i>	DNA-binding protein HU-alpha	47.63	23	13.3	17.8	
ecj:JW2644	<i>stpA</i>	DNA-binding protein stpA	3.89	13	14.2	23.9	
ecj:JW0430	<i>hupB</i>	DNA-binding protein HU-beta	3.3	9	13.3	18.9	
ecj:JW2462	<i>nlpB</i>	Lipoprotein 34	1.64	29	14.0	13.1	
ecj:JW2640	<i>ygaU</i>	Uncharacterized protein ygaU	1.38	11	6.7	6.7	
ecj:JW3261	<i>rpmJ</i>	50S ribosomal protein L36	0.98	2	21.1	31.6	
ecj:JW2283	<i>nuoA</i>	NADH-quinone oxidoreductase subunit A	0.88	5	14.3	15.0	
ecj:JW0419	<i>cyoD</i>	Cytochrome o ubiquinol oxidase protein cyoD	0.76	3	8.3	5.5	
ecj:JW0618	<i>cspE</i>	Cold shock-like protein cspE	0.56	3	5.8	11.6	
ecj:JW1079	<i>fabG</i>	3-oxoacyl-[acyl-carrier-protein] reductase	0.52	11	7.4	9.4	
ecj:JW3584	<i>secB</i>	Protein-export protein secB	0.49	7	5.2	1.3	
ecj:JW5178	<i>ymgG</i>	Uncharacterized protein ymgG	0.47	3	13.8	14.9	
ecj:JW1298	<i>pspB</i>	Phage shock protein B	0.46	2	13.5	17.6	
ecj:JW0872	<i>ltp</i>	Leucine-responsive regulatory protein	0.45	3	9.8	11.0	
ecj:JW0674	<i>seqA</i>	Protein seqA	0.41	9	9.9	13.3	
ecj:JW1078	<i>fabD</i>	Malonyl CoA-acyl carrier protein transacylase	0.39	12	4.2	2.3	
ecj:JW1837	<i>yebG</i>	Uncharacterized protein yebG	0.38	4	8.3	6.3	
ecj:JW3229	<i>fis</i>	DNA-binding protein fis	0.36	7	15.3	13.3	
ecj:JW1702	<i>ihfA</i>	Integration host factor subunit alpha	0.35	7	19.2	26.3	
ecj:JW5930	<i>yihA</i>	Probable GTP-binding protein engB	0.35	4	10.0	15.2	
ecj:JW5927	<i>ygiB</i>	UPF0441 protein ygiB	0.35	6	25.1	25.6	
ecj:JW5856	<i>trxA</i>	Thioredoxin-1	0.33	4	3.7	8.3	
ecj:JW0628	<i>rlpA</i>	Rare lipoprotein A	0.33	7	13.8	16.3	
ecj:JW1989	<i>yeeX</i>	UPF0265 protein yeeX	0.31	6	12.8	17.4	
ecj:JW0166	<i>pyrH</i>	Uridylate kinase	0.31	17	5.4	5.4	
ecj:JW5316	<i>fliO</i>	Flagellar protein fliO	0.31	5	9.9	9.1	
ecj:JW5296	<i>yoaC</i>	Uncharacterized protein yoaC	0.3	8	8.4	7.6	
ecj:JW4002	<i>dgkA</i>	Diacylglycerol kinase	0.3	2	5.7	7.4	
ecj:JW5755	<i>yjgF</i>	UPF0076 protein yjgF	0.29	2	6.3	6.3	
ecj:JW1063	<i>flaE</i>	Flagellar hook protein flaE	0.29	7	9.5	13.4	

B

<i>E. coli</i> in stationary phase							
ID	gene	Annotation	emPAI	pep_num	DBS	RBS	
ecj:JW0397	<i>yajC</i>	UPF0092 membrane protein yajC	1.37	5	9.1	15.5	
ecj:JW1225	<i>hns</i>	DNA-binding protein H-NS	0.96	10	15.3	24.8	
ecj:JW0895	<i>ihfB</i>	Integration host factor subunit beta	0.89	11	11.7	18.1	
ecj:JW0799	<i>ompX</i>	Outer membrane protein X	0.76	8	17.5	23.4	
ecj:JW2500	<i>rodZ</i>	Cytoskeleton protein rodZ	0.48	8	14.2	18.4	
ecj:JW3964	<i>hupA</i>	DNA-binding protein HU-alpha	0.42	2	13.3	17.8	
ecj:JW3624	<i>rhoZ</i>	DNA-directed RNA polymerase subunit omega	0.39	4	7.7	9.9	
ecj:JW3027	<i>htrG</i>	Uncharacterized protein ygiM	0.36	7	10.2	10.2	
ecj:JW1702	<i>ihfA</i>	Integration host factor subunit alpha	0.35	9	19.2	26.3	
ecj:JW5875	<i>nuoB</i>	NADH-quinone oxidoreductase subunit B	0.32	14	9.1	12.7	
ecj:JW0736	<i>ybgS</i>	Uncharacterized protein ybgS	0.3	2	11.1	19.8	
ecj:JW1462	<i>narY</i>	Respiratory nitrate reductase 2 beta chain	0.27	24	8.0	8.4	
ecj:JW0192	<i>rcsF</i>	Protein rcsF	0.27	6	20.1	22.4	
ecj:JW3252	<i>mscL</i>	Large-conductance mechanosensitive channel	0.26	7	2.9	6.6	
ecj:JW2644	<i>stpA</i>	DNA-binding protein stpA	0.25	5	14.2	23.9	
ecj:JW3586	<i>yibN</i>	Uncharacterized protein yibN	0.25	4	8.4	9.1	
ecj:JW0429	<i>lon</i>	ATP-dependent protease La	0.23	36	8.5	11.2	
ecj:JW2971	<i>yqhA</i>	UPF0114 protein yqhA	0.21	6	3.0	3.7	
ecj:JW3219	<i>mreC</i>	Rod shape-determining protein mreC	0.2	15	9.5	13.1	
ecj:JW3119	<i>yraP</i>	Uncharacterized protein yraP	0.19	4	11.0	15.2	
ecj:JW1541	<i>nohA</i>	Putative prophage Qin DNA-packaging protein NU1 homolog	0.18	8	10.6	8.5	
ecj:JW1850	<i>ruvA</i>	Holliday junction ATP-dependent DNA helicase ruvA	0.17	3	6.4	9.4	
ecj:JW2140	<i>foiE</i>	GTP cyclohydrolase 1	0.15	10	10.8	9.9	
ecj:JW5162	<i>lolD</i>	Lipoprotein-releasing system ATP-binding protein lolD	0.15	10	9.9	15.0	
ecj:JW0286	<i>yagY</i>	Uncharacterized protein matC	0.15	3	12.6	17.6	
ecj:JW4114	<i>frdB</i>	Fumarate reductase iron-sulfur subunit	0.14	4	6.1	12.3	
ecj:JW2577	<i>yfiO</i>	UPF0169 lipoprotein yfiO	0.14	9	9.4	9.4	
ecj:JW5867	<i>aidB</i>	Protein aidB	0.13	13	6.3	8.5	
ecj:JW3690	<i>yieE</i>	Uncharacterized protein yieE	0.13	9	10.3	10.7	
ecj:JW2406	<i>cysZ</i>	Protein cysZ	0.13	5	5.5	4.7	
ecj:JW0193	<i>metQ</i>	D-methionine-binding lipoprotein metQ	0.13	4	5.5	6.6	

Figure 6. csNAPs of *E. coli*. The listed proteins were the top 30 csNAPs sorted by the emPAI values. (A) log phase and (B) stationary phase. The meanings of the colors, letters, and underlines are the same as those for Figure 4. The proteins with a yellow background are the oxidation-reduction enzymes. DBS and RBS represent the number of DNA-binding sites and RNA-binding sites per amino acid predicted by BindN [88], respectively. Values over 10 indicate high possibilities to bind to DNA and/or RNA (see materials and methods). doi:10.1371/journal.pone.0019172.g006

A

<i>P. aeruginosa</i> in log phase						
ID	gene	Annotation	emPAI	pep_num	DBS	RBS
pae:PA3263	<i>rdgC</i>	Recombination-associated protein rdgC	1.56	17	7.5	9.2
pae:PA4238	<i>rpoA</i>	DNA-directed RNA polymerase subunit alpha	1.18	23	6.0	9.0
pae:PA4944	<i>hfq</i>	Protein hfq	1.09	5	11.0	11.0
pae:PA1804	<i>hupB</i>	DNA-binding protein HU-beta	1.09	3	11.1	20.0
pae:PA4254	<i>rpsQ</i>	30S ribosomal protein S17	0.95	2	18.2	19.3
pae:PA4433	<i>rpIM</i>	50S ribosomal protein L13	0.92	5	18.3	21.8
pae:PA4753	<i>PA4753</i>	Probable RNA-binding protein PA4753	0.8	4	8.7	15.4
pae:PA4053	<i>ribH</i>	6 7-dimethyl-8-ribityllumazine synthase	0.53	3	6.3	6.3
pae:PA0969	<i>tolQ</i>	Protein tolQ	0.52	5	6.9	8.7
pae:PA3831	<i>pepA</i>	Cytosol aminopeptidase	0.51	13	8.3	10.3
pae:PA5348	<i>hupA</i>	DNA-binding protein HU-alpha	0.42	2	17.8	22.2
pae:PA3184	<i>hexR</i>	HTH-type transcriptional regulator hexR	0.41	5	10.2	7.4
pae:PA3653	<i>frr</i>	Ribosome-recycling factor	0.41	2	11.4	14.6
pae:PA4762	<i>grpE</i>	Protein grpE	0.41	3	4.3	5.4
pae:PA2743	<i>infC</i>	Translation initiation factor IF-3	0.4	5	9.3	13.7
pae:PA3152	<i>hisH</i>	Imidazole glycerol phosphate synthase subunit hisH 2	0.36	3	5.9	8.4
pae:PA4853	<i>fis</i>	Putative fis-like DNA-binding protein	0.34	2	15.9	15.9
pae:PA5172	<i>arcB</i>	Glutamate-1-semialdehyde 2 1-aminomutase	0.32	7	7.1	10.1
pae:PA0867	<i>mltC</i>	Membrane-bound lysozyme inhibitor of C-type lysozyme	0.28	2	7.1	11.0
pae:PA4234	<i>uvrA</i>	UvrABC system protein A	0.27	13	10.4	14.5
pae:PA3011	<i>topA</i>	DNA topoisomerase 1	0.25	12	14.1	17.6
pae:PA5568	<i>oxaA</i>	Inner membrane protein oxaA	0.25	5	9.9	11.8
pae:PA0523	<i>norC</i>	Nitric oxide reductase subunit C	0.24	2	6.8	8.2
pae:PA5119	<i>glnA</i>	Glutamine synthetase	0.23	7	8.3	10.2
pae:PA0402	<i>pyrB</i>	Aspartate carbamoyltransferase	0.22	5	9.9	11.4
pae:PA3849	<i>ndpA</i>	Nucleoid-associated protein ndpA	0.21	8	9.6	10.2
pae:PA1084	<i>flgI</i>	Flagellar P-ring protein	0.21	4	6.8	10.0
pae:PA3626	<i>truD</i>	tRNA pseudouridine synthase D	0.2	3	9.3	10.7
pae:PA1074	<i>braC</i>	Leucine-isoleucine-valine-threonine-and alanine-binding protein	0.2	3	7.5	8.3
pae:PA4765	<i>omlA</i>	Outer membrane lipoprotein omlA	0.2	4	9.1	16.5

B

<i>B. subtilis</i> in log phase						
ID	gene	Annotation	emPAI	pep_num	DBS	RBS
bsu:BSU22790	<i>hbs</i>	DNA-binding protein HU 1	0.98	5	12.0	20.7
bsu:BSU07670	<i>yfil</i>	Uncharacterized protein yfil	0.74	5	11.8	7.8
bsu:BSU35870	<i>ywtC</i>	Uncharacterized protein ywtC precursor	0.64	4	1.8	1.8
bsu:BSU21940	<i>degR</i>	Regulatory protein degR	0.58	4	8.3	8.3
bsu:BSU33630	<i>secG</i>	Probable protein-export membrane protein secG	0.51	2	6.6	9.2
bsu:BSU15920	<i>acpP</i>	Acyl carrier protein	0.48	3	0.0	0.0
bsu:BSU25550	<i>rpsT</i>	30S ribosomal protein S20	0.42	8	19.3	33.0
bsu:BSU21050	<i>yonN</i>	SPBc2 prophage-derived DNA-binding protein HU 2	0.41	7	12.0	21.7
bsu:BSU38590	<i>licB</i>	Lichenan-specific phosphotransferase enzyme IIB component	0.36	11	5.9	8.8
bsu:BSU28780	<i>araD</i>	L-ribulose-5-phosphate 4-epimerase	0.32	5	7.0	7.0
bsu:BSU07550	<i>yflT</i>	General stress protein 17M	0.3	7	0.9	7.8
bsu:BSU24330	<i>yqhY</i>	Uncharacterized protein yqhY	0.26	5	1.5	5.2
bsu:BSU14990	<i>yneK</i>	Regulatory protein ylbF	0.23	6	10.1	14.8
bsu:BSU17960	<i>yjBF</i>	Uncharacterized protein yneK	0.23	7	2.8	10.6
bsu:BSU15890	<i>plsX</i>	Fatty acid/phospholipid synthesis protein plsX	0.22	15	11.1	8.7
bsu:BSU13540	<i>ogt</i>	Methylated-DNA--protein-cysteine methyltransferase	0.21	4	10.9	11.5
bsu:BSU24420	<i>spoIIAB</i>	Stage III sporulation protein AB	0.2	8	7.0	11.1
bsu:BSU31990	<i>dhbC</i>	Isochorismate synthase dhbC	0.18	9	7.3	9.8
bsu:BSU01930	<i>ybcS</i>	Uncharacterized protein skfC precursor	0.18	8	9.7	7.6
bsu:BSU00070	<i>gyrA</i>	DNA gyrase subunit A	0.17	29	9.9	12.9
bsu:BSU27930	<i>ylxF</i>	Sporulation initiation phosphotransferase B	0.17	4	5.2	5.7
bsu:BSU16260	<i>spo0B</i>	FlaA locus 22.9 kDa protein	0.17	9	10.3	22.5
bsu:BSU06180	<i>pspA</i>	Phage shock protein A homolog	0.15	8	7.5	11.0
bsu:BSU24130	<i>prpD</i>	Protein mmgE	0.14	8	7.4	8.7
bsu:BSU23920	<i>yqjD</i>	Putative propionyl-CoA carboxylase beta chain	0.14	15	9.5	15.2
bsu:BSU14570	<i>ykyA</i>	Uncharacterized protein ykyA	0.14	10	9.3	9.7
bsu:BSU25710	<i>cwlH</i>	N-acetylmuramoyl-L-alanine amidase cwlH precursor	0.14	8	17.2	23.2
bsu:BSU23960	<i>yqiZ</i>	Arginine transport ATP-binding protein artM	0.14	10	13.0	15.1
bsu:BSU05730	<i>ydhF</i>	Uncharacterized protein ydhF precursor	0.14	6	11.4	17.8
bsu:BSU37620	<i>rsfA</i>	Prespore-specific transcriptional regulator rsfA	0.13	8	9.3	11.2
bsu:BSU40050	<i>gntR</i>	Gluconate operon transcriptional repressor	0.13	7	5.8	6.2
bsu:BSU23450	<i>sigF</i>	RNA polymerase sigma-F factor	0.13	2	5.9	10.2
bsu:BSU30800	<i>menB</i>	Naphthoate synthase	0.13	8	7.4	10.7
bsu:BSU30850	<i>ytDA</i>	Putative UTP--glucose-1-phosphate uridylyltransferase	0.13	4	5.5	5.5
bsu:BSU02920	<i>vceF</i>	Uncharacterized membrane protein vceF	0.13	5	4.3	5.1

Figure 7. csNAPs of *P. aeruginosa* and *B. subtilis*. The listed proteins were the top 30 csNAPs sorted by the emPAI values. (A) log phase of *P. aeruginosa* and (B) log phase of *B. subtilis*. The meanings of the colors, letters, and underlines are the same as those for Figure 4. The proteins with a yellow background are the oxidation-reduction enzymes. DBS and RBS represent the number of DNA-binding sites and RNA-binding sites as described in Figure 6.

doi:10.1371/journal.pone.0019172.g007

A

<i>S. aureus</i> in log phase							
ID	gene	Annotation	emPAI	pep_num	DBS	RBS	
sau:SA0944	<i>phdB</i>	Pyruvate dehydrogenase E1 component subunit beta	2.39	27	6.5	6.8	
sau:SA1414	<i>rpsT</i>	30S ribosomal protein S20	2.06	4	20.5	28.9	
sau:SA2033	<i>rplF</i>	50S ribosomal protein L6	1.91	21	14.6	21.3	
sau:SA0723	<i>clpP</i>	ATP-dependent Clp protease proteolytic subunit	1.67	11	7.2	8.7	
sau:SA0504	<i>rpsG</i>	30S ribosomal protein S7	1.2	15	17.9	14.1	
sau:SA1382	<i>sodA</i>	Superoxide dismutase [Mn/Fe] 1	1.18	9	5.0	7.0	
sau:SA0729	<i>tpi</i>	Triosephosphate isomerase	1.18	11	6.7	8.3	
sau:SA1663	SA1663	UPF0342 protein SA1663	1.17	5	0.0	1.8	
sau:SA0366	<i>ahpC</i>	Alkyl hydroperoxide reductase subunit C	0.95	5	4.8	3.7	
sau:SA0456	<i>spoVG</i>	Putative septation protein spoVG	0.83	3	11.1	9.3	
sau:SA2036	<i>rplX</i>	50S ribosomal protein L24	0.81	2	6.7	25.7	
sau:SA1073	<i>fabD</i>	Malonyl CoA-acyl carrier protein transacylase	0.7	11	2.9	5.2	
sau:SA1930	<i>rpoE</i>	Probable DNA-directed RNA polymerase subunit delta	0.66	6	0.0	0.6	
sau:SA1113	<i>rbfA</i>	Ribosome-binding factor A	0.66	4	9.5	9.5	
sau:SA2312	<i>ddh</i>	D-lactate dehydrogenase	0.63	13	6.0	4.8	
sau:SA1404	<i>rpsU</i>	30S ribosomal protein S21	0.62	2	41.4	46.6	
sau:SA0856	<i>spxA</i>	Regulatory protein spx	0.56	2	17.6	15.3	
sau:SA2029	<i>rplO</i>	50S ribosomal protein L15	0.56	8	15.8	30.1	
sau:SA1901	<i>fabZ</i>	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase	0.53	5	5.5	7.5	
sau:SA0719	<i>trxB</i>	Thioredoxin reductase	0.53	6	6.1	6.4	
sau:SA2039	<i>rpoZ</i>	50S ribosomal protein L29	0.51	6	17.4	17.4	
sau:SA2026	<i>infA</i>	Translation initiation factor IF-1	0.49	4	15.3	9.7	
sau:SA0245	<i>ispD</i>	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase 2	0.49	8	6.3	7.9	
sau:SA0918	<i>purC</i>	Phosphoribosylaminoimidazole-succinocarboxamide synthase	0.49	8	6.8	10.7	
sau:SA0941	SA0941	UPF0356 protein SA0941	0.46	3	5.4	8.9	
sau:SA0354	<i>rpsR</i>	30S ribosomal protein S18	0.43	4	30.0	28.8	
sau:SA1653	<i>trap</i>	Signal transduction protein TRAP	0.43	6	15.0	10.2	
sau:SA1305	<i>hu</i>	DNA-binding protein HU	0.42	3	13.3	18.9	
sau:SA1359	<i>efp</i>	Elongation factor P	0.41	3	10.8	13.0	
sau:SA0942	<i>pdf1</i>	Peptide deformylase	0.41	7	6.6	6.6	

B

<i>S. aureus</i> in stationary phase							
ID	gene	Annotation	emPAI	pep_num	DBS	RBS	
sau:SA0992	<i>trxA</i>	Thioredoxin	4.91	11	0.0	1.9	
sau:SA0295	SA0295	30 kDa neutral phosphatase (Fragment)	3.92	2	14.9	22.0	
sau:SA0873	SA0873	UPF0477 protein SA0873	2.57	8	11.8	11.2	
sau:SA1178	SA1178	UPF0154 protein SSP1415	2.14	15	11.3	13.8	
sau:SA1305	<i>hu</i>	DNA-binding protein HU	1.85	8	13.3	18.9	
sau:SA1663	SA1663	UPF0342 protein SA1663	1.82	8	0.0	1.8	
sau:SA2043	<i>rpsS</i>	30S ribosomal protein S19	1.6	7	17.4	17.4	
sau:SA1067	<i>rpmB</i>	50S ribosomal protein L28	1.55	3	41.9	50.0	
sau:SA0456	<i>spoVG</i>	Putative septation protein spoVG	1.47	7	11.1	9.3	
sau:SA1909	<i>atpF</i>	ATP synthase subunit b	1.46	10	3.5	9.8	
sau:SA1709	<i>ftn</i>	Ferritin	1.45	10	2.4	4.8	
sau:SA0160	<i>isdI</i>	Heme-degrading monooxygenase isdI	1.24	3	8.3	3.7	
sau:SA2062	<i>sarV</i>	HTH-type transcriptional regulator sarV	1.1	5	12.1	6.0	
sau:SA0760	<i>gcsH</i>	Glycine cleavage system H protein	1.09	7	1.6	1.6	
sau:SA0108	<i>sarH1</i>	HTH-type transcriptional regulator sarS	1.04	12	6.0	6.8	
sau:SA078	<i>rpmJ</i>	50S ribosomal protein L36	1.01	3	27.0	29.7	
sau:SA1904	<i>atpC</i>	ATP synthase epsilon chain	1.01	4	3.7	12.7	
sau:SA0032	<i>bleO</i>	Bleomycin resistance protein	0.99	12	4.5	5.2	
sau:SA0494	<i>nusG</i>	Transcription antitermination protein nusG	0.98	9	9.3	12.6	
sau:SA0478	<i>pdxT</i>	Glutamine amidotransferase subunit pdxT	0.97	5	5.4	5.4	
sau:SA2038	<i>rpmC</i>	30S ribosomal protein S17	0.94	4	17.4	17.4	
sau:SA0245	<i>ispD</i>	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase 2	0.94	11	6.3	7.9	
sau:SA1901	<i>fabZ</i>	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase	0.9	4	5.5	7.5	
sau:SA1256	<i>msrB</i>	Peptide methionine sulfoxide reductase msrB	0.89	3	5.6	9.2	
sau:SA0128	<i>sodM</i>	Superoxide dismutase [Mn/Fe] 2	0.85	8	6.5	6.0	
sau:SA1019	SA1019	Uncharacterized N-acetyltransferase SA1019	0.85	3	3.4	4.1	
sau:SA0437	SA0437	UPF0133 protein SAB0428	0.81	12	5.7	3.8	
sau:SA1074	<i>fabG</i>	3-oxoacyl-[acyl-carrier-protein] reductase	0.72	9	9.3	9.8	
sau:SA2431	<i>isaB</i>	Immunodominant staphylococcal antigen B	0.72	9	18.9	22.3	
sau:SA1041	<i>pvrR</i>	Bifunctional protein pvrR	0.7	4	7.8	10.6	

Figure 8. csNAPs of *S. aureus*. The listed proteins were the top 30 csNAPs sorted by the emPAI values. (A) log phase and (B) stationary phase. The meanings of the colors, letters, and underlines are the same as those for Figure 4. The proteins with a yellow background are the oxidation-reduction enzymes. DBS and RBS represent the number of DNA-binding sites and RNA-binding sites as described in Figure 6. doi:10.1371/journal.pone.0019172.g008

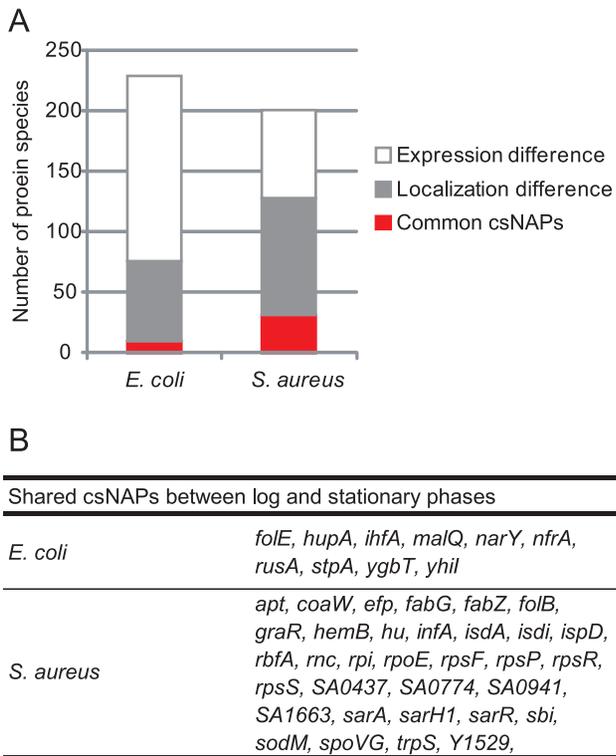


Figure 9. Common csNAPs between the log and stationary phases. (A) Classification of csNAPs Common csNAPs are csNAPs common to the log and stationary phases. Localization difference indicates the proteins which were not classified as csNAPs in either the log or the stationary phase but were detected in the envelope, and/or top fractions in the other phase. Expression difference represents the proteins which were classified as csNAPs in either the log or the stationary phase but not detected in any fractions in the other phase. (B) Common csNAPs between the log and stationary phases in *E. coli* and *S. aureus*. doi:10.1371/journal.pone.0019172.g009

decrease in Fis cause nucleoid condensation in the stationary phase in *E. coli*. On the other hand, the Dps orthologue, MrgA, is hardly expressed throughout the normal growth of *S. aureus*, and its nucleoid does not alter its apparent structure toward the stationary phase. The number of constitutive csNAPs that we could detect was smaller in *E. coli* than in *S. aureus*. This might imply the correlation between such structural change and the exchange of the NAP constituents.

Even under such growth-dependent changes, several global regulators and oxidation-reduction enzymes were constitutively present. In *E. coli*, Hu (*hupA*), IHF (*ihfA*), and StpA (*stpA*) were listed as global regulators, and nitrate reductase (*narY*) was listed as an oxidation-reduction enzyme. In *S. aureus*, the global regulators Hu (*hu*), SarA (*sarA*), SarR (*sarR*), and SarH (*sarH*) and the oxidation-reduction enzyme SodM (*sodM*) was constitutively present as csNAPs. These results suggest that global regulators and oxidation-reduction enzymes have important roles in the nucleoid regardless of their growths.

csNAPs shared by *E. coli*, *P. aeruginosa*, *B. subtilis*, and *S. aureus*

Comparison of csNAPs among species showed that Hu- β (coded by *hupB* in *E. coli*, *hupB* in *P. aeruginosa*, *hbs* in *B. subtilis*, and *hu* in *S. aureus*) was common to the 4 species. *E. coli* and *P. aeruginosa* share 5

additional csNAPs, whereas *B. subtilis* and *S. aureus* share 2 additional csNAPs (Figure 10). This low number of common csNAPs seems due both to ‘gene-level difference (the lack of the orthologous genes)’ and to ‘protein-level difference (expression or localization difference)’. *E. coli* and *P. aeruginosa*, for example, did not share 33.9% of the csNAPs genes as orthologues (76 out of 224 [164 *E. coli* csNAPs + 66 *P. aeruginosa* csNAPs – 6 common csNAPs between these species]). Among the residual 148 csNAPs, whose genes are present in both species, 66.2% (98 out of 148) csNAPs were not detected in either *E. coli* or *P. aeruginosa*. While 30.4% (45 out of 148) of the csNAPs were detected in both species, they were not selected as csNAPs in either. *B. subtilis* and *S. aureus* showed a similar pattern. Although the degree of ‘gene-level difference’ increased in the distantly related species (41.8% to 66.9%), the

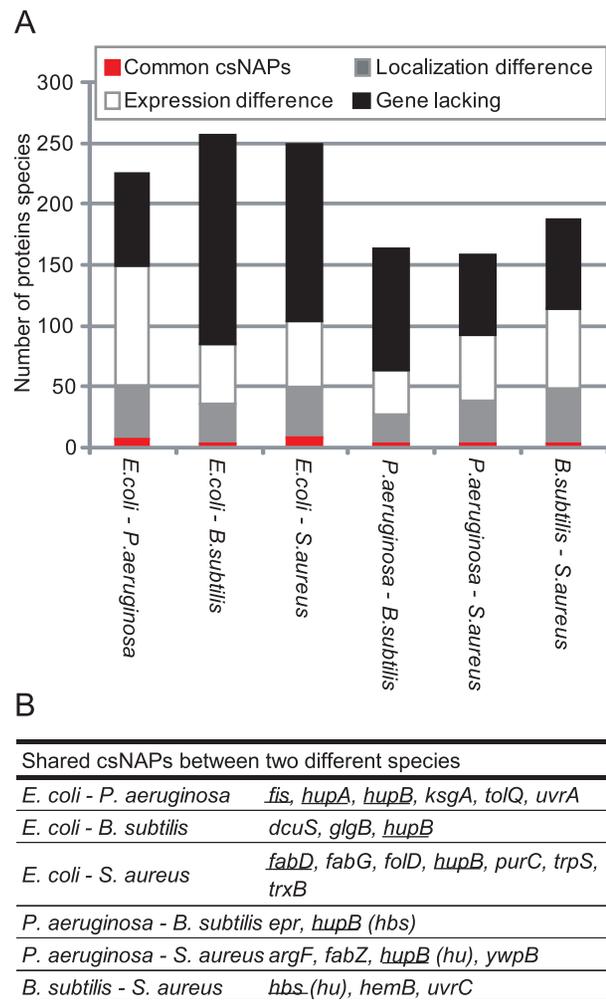


Figure 10. Common csNAPs in *E. coli*, *P. aeruginosa*, *B. subtilis*, and *S. aureus*. (A) Classification of csNAPs. Common csNAPs represent csNAPs common to 2 species. Localization difference means the proteins that were not classified as csNAPs in either 2 species but were detected in the envelope and/or top fractions in the other species. Expression difference indicates the proteins that were classified as csNAPs in either species but were not detected in any fractions in the other species. Gene lacking represents the proteins that were csNAPs in either species but for which the corresponding genes were not present in the other species’ genomes. (B) csNAPs common to 2 species in the log phase. Underlined genes were ranked in the top 30 emPAI values in both compared species. doi:10.1371/journal.pone.0019172.g010

'protein-level difference' could still explain approximately half of the cross-species difference of the csNAPs (Figure 10A). As discussed in the previous section, it is possible that additional common csNAPs remain undetected (Table S15). However, the amount of those additional csNAPs in the cell should be small. Accordingly, the main csNAPs seem to differ among the species, suggesting that as each species evolved, it developed its own proteins associated with the nucleoid.

When the common csNAPs were sorted by emPAI values, Hu (*hupA*, *hupB*, *hbs*, *hu*), Fis (*fis*), and FabD (*fabD*) were included in the top 30 in certain pairs of species (Figure 10B). Hu gained high emPAI values in all the species. Given that it is crucial in various species [63–65], Hu should play a critical role in the bacterial kingdom. The abundance of Fis in *E. coli* and *P. aeruginosa* implies its importance in Gammaproteobacteria. It is interesting that FabD, also known as malonyl CoA-acyl carrier protein transacylase, was abundantly present in *E. coli* and *S. aureus*. In both species, *fabG*, which encodes 3-oxoacyl-[acyl-carrier-protein] reductase (FabG) and forms an operon with *fabD* [66], was also common. These genes are involved in fatty acid biosynthesis [67]. Between *P. aeruginosa* and *S. aureus*, FabZ, which is (3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase and also involved in fatty acid synthesis, was a common csNAP. The prediction of their localization and DNA/RNA binding ability showed that these proteins are cytoplasmic proteins with less potential to bind DNA and RNA (Figure 6A and 6E). It might be possible that fatty acid synthesis occurs near the genomic DNA in those bacteria.

NAPs not listed in the csNAPs: contaminant or genuine NAP?

Many proteins that have been reported to exist in the nucleoid were not included in the list of csNAPs. The involvement of the real NAPs in the top fraction might be due to disassociation from the NAPs during the purification procedure or might reflect their dynamic association with or disassociation from the nucleoid. Indeed, Hfq was reported to exist in both the cytoplasm and the nucleoid [68]. It might also be possible that high expression levels of NAPs lead to their higher accumulation in the cytoplasm owing to their saturation in the nucleoid.

Regarding the cell envelopes, the outer membranes of Gram-negative bacteria are resistant to nonionic detergents such as Brij-58 that was used in this study [5,36]. It is likely that the outer membrane proteins detected in *E. coli* and *P. aeruginosa*, such as flagellin and porins, were contaminated owing to the insufficient solubilization of the outer membrane. The cell walls of Gram-positive bacteria are relatively thicker than those of Gram-negative bacteria. In particular, *S. aureus* possesses a thick cell wall that is resistant to lysozyme [69]. We used lysostaphin [70] to disrupt the cell wall, but it is possible that the cell wall fragments remained, in which case, cell-surface proteins such as immunodominant antigen A (coded by *isaA*) [71] and protein A (coded by *spa*) [72] in *S. aureus* would be the contaminants.

The inner membranes of *E. coli* are readily solubilized by Brij-58 [36], and indeed various inner membrane proteins, such as TonB [73], TolR [74], TatAB [75], and SecE [76], were detected only in the envelope fraction (Table S1). Nevertheless, some inner membrane proteins, such as the methyl-accepting chemotaxis proteins Tar and Tsr [77] were included in our nucleoid fractions (but not selected as csNAPs). These proteins may play some role in the nucleoid characteristics that require, for instance, interaction with the membrane. Similarly, Portalier and Worcel previously reported that several inner membrane proteins remained in the *E. coli* nucleoid even after treatment with sarkosyl and 1M NaCl [5]. In the present study, cytosolic portions of F₀F₁-ATPase (F₁ β

[*atpD*] and F₁ γ [*atpG*]) were detected in the nucleoid and envelope fractions with high emPAI values under all the conditions tested. These remained in the nucleoid even after high salt treatment [45].

Recently, cell envelope-attached proteins, such as MreB, have been reported to interact with RNA polymerases and elongation factor EF-Tu [78,79]. In *B. subtilis*, these proteins were indeed detected in both the nucleoid and envelope fractions. In *E. coli*, MreB (*mreB*) itself was not detected in the nucleoid fractions, but RpoZ (*rpoZ*), which is also a cytoskeleton protein and interacts with MreB [80], was detected in both the nucleoid and the envelope fractions. Thus, cytoskeleton proteins are also possible candidates for the linker of the nucleoid and cell envelope.

Future Perspectives

Through the analyses of csNAPs in this study, we showed that global regulators, oxidation-reduction enzymes, and fatty acid synthases were enriched in the nucleoids. The distinct evolutionary origins of those proteins imply that bacteria have individually developed nucleoid-associated proteins in order to obtain similar characteristics.

While it was reasonable that various global regulators were in the nucleoid, it was surprising that oxidation-reduction enzymes and fatty acid synthase were enriched in the nucleoid because they have been believed to work in the cytosol and/or the envelope. The enrichment of oxidation-reduction enzymes facilitates our proposal of a new hypothesis - the armor hypothesis, which postulates the proteins in the nucleoid defuse the oxidative stress elements that challenge genomic DNA.

The presence of csNAPs involved in fatty acid synthesis implies certain relationships between the cellular membrane and the nucleoid. Phospholipids have been reported to regulate DNA replication via direct interaction with the replication initiator protein DnaA [81–83]. On the contrary, mutations in the replication machinery proteins such as DnaA, SeqA (negative modulator of initiation of replication), and Dam (DNA adenine methyltransferase) change the phospholipid constituents [84,85], suggesting the presence of a bidirectional regulatory system to maintain the genomic DNA and cellular membrane. Fatty acid synthases in the nucleoid might also be involved in such a crosstalk system.

Although the current study only suggested the presence of the oxidation-reduction enzymes and fatty acid synthases in the nucleoid, it would be fascinating in future studies to focus on nucleoid characteristics such as the 'armor hypothesis' and 'nucleoid-membrane crosstalk.'

Materials and Methods

Bacterial strains and growth conditions

Glycerol stocks of *E. coli* K-12 W3110, *P. aeruginosa* PAO1, and *B. subtilis* 168 were inoculated in LB medium and cultured at 37°C with constant shaking (180 rpm; Bioshaker BR-15, TAITEC) for 18 h. Twenty-five microliters of the saturated culture was inoculated in 25 or 50 mL of fresh LB medium and cultured at 37°C with constant shaking (180 rpm) until the OD₆₀₀ reached 0.7 (log phase) (Figure S1). The cell density was determined by measuring absorbance at 600 nm. Glycerol stocks of *S. aureus* N315 were inoculated in Brain Heart Infusion (BHI) medium (Difco, Detroit, MI, USA) and cultured at 37°C with constant shaking (180 rpm) for 18 h. Two hundred fifty microliters of the saturated culture was inoculated in 25 mL of fresh BHI and cultured at 37°C, with constant shaking, to an appropriate cell density (OD₆₀₀ = 0.7). The cultures in the stationary phases were collected

12 to 14 h after the inoculations, and cells collected from 2 mL culture were used for the subsequent studies.

Isolation of nucleoids

Cultures (25 mL for *E. coli*, *P. aeruginosa*, and *S. aureus* and 50 mL for *B. subtilis*) were centrifuged at 8000 g for 10 min at 4°C. Cell pellets were suspended in 0.5 mL ice-cold Buffer A (10 mM Tris-HCl [pH 8.2], 100 mM NaCl, and 20% sucrose) followed by the addition of 0.1 mL ice-cold Buffer B (100 mM Tris-HCl [pH 8.2], 50 mM EDTA, 0.6 mg/mL lysozyme [plus 100 µg/mL lysostaphin in the case of *S. aureus*]). The mixtures were incubated for the appropriate time at the appropriate temperature according to previous reports (*E. coli* [4] and *B. subtilis* [3]) or the preliminary experiments to understand the minimum requirements for the incubation conditions: in detail, 1 to 3 min on ice for the log phase of *E. coli*, 5 min at room temperature (RT) for the stationary phase of *E. coli*, 5 min on ice for the log phase of *P. aeruginosa*, 15 min on ice for the log phase of *B. subtilis*, and 15 min at RT for the log and stationary phases of *S. aureus*). Then, 0.5 mL ice-cold Buffer C (10 mM Tris-HCl [pH 8.2], 10 mM EDTA, 10 mM spermidine, 1% Brij-58, and 0.4% deoxycholate) was added, and the mixtures were incubated for the appropriate time (3–5 min at RT for the log phase of *E. coli*, 15 min at RT for the stationary phase of *E. coli*, 10 min at RT for *P. aeruginosa*, 20 min at RT for *B. subtilis*, and 30 min at RT for the log and stationary phases of *S. aureus*). The lysed cell suspensions were loaded onto linear sucrose density gradients containing 10 mM Tris-HCl (pH 8.2) and 100 mM NaCl (10%–60% for *E. coli*, *P. aeruginosa*, and *B. subtilis*, and 10%–30% for *S. aureus*) and centrifuged at 10,000 rpm with a Beckmann SW 40 Ti rotor (20 min for *E. coli*, 50 min for *P. aeruginosa* and *S. aureus*, and 40 min for *B. subtilis*). The DNA concentration in each fraction was quantified by DAPI fluorescence signals [4].

Preparation of envelope fractions

Cellular envelopes were purified according to Zimmerman's method with several modifications [4]. The cultures (25 mL for *E. coli*, *P. aeruginosa*, and *S. aureus* and 50 mL for *B. subtilis*) were centrifuged at 8000 g for 10 min. Cell pellets were suspended in 0.5 mL ice-cold Buffer A followed by the addition of 0.1 mL ice-cold Buffer B (plus lysostaphin [25 µg/mL final concentration] in the case of *S. aureus*). The mixtures were incubated for the appropriate time (2 min at RT for *E. coli* and *P. aeruginosa* and 5 min at RT for *B. subtilis* and *S. aureus*). After adding PMSF (1 mM final concentration), the solutions were sonicated in ice water until they became clear. The debris was removed by centrifugation at 1200 g for 20 min at 4°C. The supernatants were collected, and 5 µg RNase, 10 U DNase, and MgCl₂ (40 mM final) added. After 60-min incubation at 37°C, the envelope fractions were collected as pellets by centrifugation at 20,000 g for 60 min at 4°C.

LC-MS/MS

Each lane of the Coomassie Brilliant Blue (CBB)-stained SDS-PAGE gels (8.5×6 cm) was cut into 10 sequential slices. Proteins in each gel slice were destained in 50% acetonitrile/50 mM ammonium bicarbonate (ABB), deoxidized by 10 mM DTT in 100 mM ABB, and alkylated by 55 mM iodoacetamide in 100 mM ABB. After being washed in 100 mM ABB and then 50% acetonitrile and 50 mM ABB, the gel slices were dried thoroughly in a MicroVac (Tomy Digital Biology, Tokyo, Japan). The proteins were then digested with 15 ng/µL trypsin (Trypsin Gold; [Promega, Madison, WI, USA]) in 50 mM ABB for 8 to 12 h at 37°C. Tryptic peptides were extracted by sonication in 50% acetonitrile/0.1% trifluoroacetic acid (TFA), and the supernatants collected. Again, the peptides were extracted by sonication in 75%

acetonitrile/0.1% TFA and collected as supernatants. The samples were dried using the MicroVac and suspended in 2% acetonitrile/0.1% TFA. After being filtered by C-TIP (AMR Technology, Albany, NY, USA), the samples were analyzed by LC-MS/MS.

Reverse-phase nano-LC-MS/MS was performed using a Paradigm MS4 system (Michrom BioResources, Auburn, CA, USA) coupled to an LXQ (Thermo Scientific). The digested peptides were separated on an HPLC column (0.1×150 mm, 3 µm Magic C18AQ; Michrom BioResource) using a linear gradient of 6.4% to 41.6% acetonitrile in 0.1% formic acid for 20 min at a flow rate of 500 nL/min and detected by the ion trap in the 450–1800 m/z range following the supplier's recommendations. Mass spectra were acquired in the positive-ion mode with automated data-dependent MS/MS on the 3 most intense ions from the precursor MS scans.

Protein identification was performed using a Mascot Server (Matrix Science). Protein identifications were obtained by processing the experimental data against the SwissProt bacteria subset database (Release 57.4, June 16, 2009). The search parameters were as follows: trypsin was used as the cutting enzyme, mass tolerance for the monoisotopic peptide window was set to ±2 Da, the MS/MS tolerance window was set to ±1 Da, and 1 missed cleavage was allowed. Cysteine carbamidomethyl modification and oxidized methionine were chosen as the variable modifications. The criteria of positive identification were set as follows: identification of at least 2 peptides with more than 7 amino acids, and a significant threshold of $P < 0.05$.

Estimation of coverage rates

The coverage rates of proteins identified in each fraction were calculated by *in silico* simulation. We first created hypothetical protein sets (X-axis in Figure S2 B–D) and estimated the expected number of proteins identified by random sampling with a given number of the peptides (Y-axis in Figure S2 B–D). These plots make it possible to expect the actual number of the protein in the cognate samples. In detail, the steps of the procedure were as follows:

Step 1: Creation of hypothetical protein sets

We randomly collected 1 to 1500 protein sequences without overlaps from the genomic database of *E. coli*, *P. aeruginosa*, *B. subtilis*, or *S. aureus* [86]. The relative protein amounts (P_r) in the hypothetical protein sets were given as follows: here we applied 3 different models: the constant model, the linear model, and the simplified canonical law (SCL) model (Figure S2A).

$P_r = 1/N$ constant model

$P_r = (N - r + 1) / \sum_{r=1}^N r$ linear model

$P_r = (r + \rho)^{-1/\theta} / \sum_{r=1}^N (r + \rho)^{-1/\theta}$ SCL model,

where r is the rank of individual protein species sorted by their amounts, N is the number of the collected protein species (1–1500 in this study), and ρ and θ are the parameters to determine the individual protein amounts. In the constant model, the amounts of individual proteins were equal. In the linear model, the protein amounts linearly decreased. SCL was reported to fit well with the expression pattern in prokaryotic cells [47]. ρ and θ were set as 5.17 and 0.58, respectively, because Ramsden and Vohradsky showed that *E. coli* followed these values [47]. The ranks of the individual proteins were randomly determined. The number of the least protein was set as 1. Each of the collected protein sequences was replicated for the times according to the above statistical models. The summation of the number of replicated proteins was defined as the total amount of proteins. (We symbolize this value as M in the following steps.)

Step 2: Creation of theoretical peptide sets

The theoretical peptide sets that were theoretically produced by the digestion of trypsin were created from the hypothetical protein sets. Since the lengths of our detected peptides were between 6 and 45, we discarded the peptides with lengths out of this range.

Step 3: Estimation of the expected number of proteins identified in the hypothetical protein sets

Random sampling of the peptides from the theoretical peptide set was performed. Here, the number of the sampling cycles was set as the number of the experimentally identified peptides (eg, the log phase of the *E. coli* nucleoid as 7148). Then, the number of the selected protein species was counted (we counted only the proteins that were hit by more than 2 peptides according to our criteria for the protein identification). The amount of the selected proteins was set according to the number of the individual protein sequences determined in Step 1.

Step 4: Repetition of Steps 1 to 3 for calculation of the averages

We repeated Step 1 to 3 5 times and calculated the averages of the number of protein species (N_e) and the amount of proteins (M_e). Examples of the plots of N , M , N_e , and M_e are shown in Figure S2 B to F.

Step 5: Estimation of the coverage rates

The coverage rates of protein species (R_N) and amounts (R_M) were calculated as follows:

$$R_N = N_{es}/N_s$$

$$R_M = M_{es}/M_s$$

where N_{es} is the number of protein species identified in the real sample and N_s is the value of N when N_e gains N_{es} . M_{es} and M_s are the values of M_e and M , respectively, when N_e gains N_{es} . Examples of the R_N and R_M values are shown in Figure S2 G and H.

The number of peptides required for identifying all the potential proteins in the sample was estimated based on the above criterion of SCL model. We fixed the number of proteins in Step 1, and increased the numbers of peptides (with 1,000 intervals) to achieve the coverage rate of 1. The average of 5 trials is shown in Table S4. The above analyses were performed wholly by means of the Perl script.

Comparative genomic analyses

The orthologue relationships were determined by best-best hit analyses using the FASTA package [87] on a desk-top computer. The pairs of proteins that showed a 'best-best' relation and a gain of less than 0.0001 of e-values were determined to be orthologous pairs. Total protein sequences identified in *E. coli* W3110, *P. aeruginosa* PAO1, *B. subtilis* 168, and *S. aureus* N315 were downloaded from KEGG on June 1, 2009 [86] and used for the best-best hit analyses.

The prediction of DNA/RNA binding abilities

The DNA/RNA binding sites of the csNAPs were predicted by BindN [88]. The criterion for the search is 'the predicted DNA/RNA binding residues with expected specificity equal to 90%.' We estimated the percentages of DNA/RNA binding residues in a protein and set 10% as the proteins having high DNA/RNA binding ability. This criterion was based on the following observations (Figure S4): The investigation of the number of DNA/RNA binding sites of Hu, IHF, HNS, StpA, Fis, and Hfq

and of ribosomal proteins in *E. coli* showed that over 10% of the residues in each protein were predicted to be DNA/RNA binding sites. The distribution of the rates of the DNA/RNA binding sites of the csNAPs showed that 2 normal distributions appeared whose peaks were around 7.5% and 14% and that the gulf of the 2 distributions was around 10%. In contrast, the distribution of the protein detected only in the envelope and/or top fractions showed that only 1 normal distribution appeared whose peak was around 7.5%. These results support the notion that proteins whose predicted DNA/RNA binding sites are over 10% have high potential to bind to DNA/RNA.

The number of potential csNAPs

The number of potential csNAPs (N_p) was estimated based on the coverage rates of proteins and the number of subtracted proteins in the selection of csNAPs. N_p was calculated as follows:

$$N_p = N_n/R_n - (N_{ne}/R_n R_e + N_{nt}/R_n R_t - N_{net}/R_n R_e R_t)$$

where N_n is the number of proteins identified in the nucleoid fractions, N_{ne} is the number of 'overlapped proteins' between the nucleoid and envelope fractions ('overlapped proteins' represent the proteins subtracted in the process of the csNAPs selection), and N_{nt} is the number of 'overlapped proteins' between the nucleoid and top fractions. N_{net} is the number of 'overlapped proteins' in all 3 fractions. R_n , R_e and R_t represent the coverage rates of proteins in the nucleoid, envelope, and top fractions, respectively. The highest number of N_p was obtained when R_n was the lowest and R_e and R_t were the highest. The minimum N_p was obtained vice versa.

Supporting Information

Figure S1 Sampling points of bacterial cells. The log phase cultures were collected at $OD_{600} = 0.7$, and the stationary phase cultures were collected 12 to 14 h after inoculation (arrows). (TIF)

Figure S2 The estimation of the coverage rates of the proteins in the samples. (A) The probability distribution of protein amounts according to 3 different models: constant model (blue line), linear model (red line), and SCL model (green line). (B, C, D) The number of protein species expected to be identified in the hypothetical protein set with the constant model (B), the linear model (C), and the SCL model (D). The x axis represents the number of protein species in the hypothetical protein set, and the y axis, the expected number of protein species detected in the hypothetical protein sets. (E, F) The protein amounts expected to be identified in the hypothetical protein set with the SCL model (E), and the linear model (F). The x axis represents the number of protein species, and the y axis, the protein amounts expected to be identified. The solid line indicates the total amounts of proteins in hypothetical protein sets. (G, H) The coverage rates of protein species (G) and amounts (H) estimated with the SCL model. The x axis represents the number of protein species identified, and the y axis, the coverage rates of protein species (G) and protein amounts (H), respectively. The pink circle, square, and triangle indicate the points corresponding to the experimental results. (EPS)

Figure S3 Nucleoid isolation of *S. aureus*, *P. aeruginosa*, and *B. subtilis*. The nucleoid isolations of the log phases of *S. aureus* (A, B), the stationary phase of *S. aureus* (C, D), the log phase of *P. aeruginosa* (E, F), and the log phase of *B. subtilis* (G, H). The spermidine nucleoids were fractionated by sucrose-gradient

centrifugation with a 10%-to-30% (60%) gradient (A, C, E, G). The fractions containing genomic DNA were identified by DAPI fluorescence (B, D, F, H).

(TIF)

Figure S4 The distribution of the percentages of DNA/RNA binding sites in proteins. (A) The number of DNA/RNA binding sites in Hu, IHF, Fis, HNS, StpA, and Hfq in *E. coli*. Percent (%) represents the percentage of DNA (or RNA) binding amino acid in each protein. (B) The distributions of the percentages of the DNA/RNA binding sites of the csNAPs and the proteins that appeared only in the envelope and/or top fractions (env_top_specific) in *E. coli*.

(TIF)

Table S1 Full list of the proteins identified in this study.
(XLS)

Table S2 Proteins identified in the nucleoid fraction of the log phase of *E. coli*.

(XLS)

Table S3 Proteins identified in the nucleoid fraction of the stationary phase of *E. coli*.

(XLS)

Table S4 Number of peptides required to cover all the potential proteins in samples.

(XLS)

Table S5 Proteins identified in the nucleoid fraction of the log phase of *P. aeruginosa*.

(XLS)

Table S6 Proteins identified in the nucleoid fraction of the log phase of *B. subtilis*.

(XLS)

Table S7 Proteins identified in the nucleoid fraction of the log phase of *S. aureus*.

(XLS)

Table S8 Proteins identified in the nucleoid fraction of the stationary phase of *S. aureus*.

(XLS)

Table S9 csNAPs of *E. coli* in the log phase.

(XLS)

Table S10 csNAPs of *E. coli* in the stationary phase.

(XLS)

Table S11 csNAPs of *P. aeruginosa* in the log phase.

(XLS)

Table S12 csNAPs of *B. subtilis* in the log phase.

(XLS)

Table S13 csNAPs of *S. aureus* in the log phase.

(XLS)

Table S14 csNAPs of *S. aureus* in the stationary phase.

(XLS)

Table S15 Potential number of csNAPs.

(XLS)

Acknowledgments

We thank Dr. Kunio Takeyasu for his valuable suggestions and for providing us with Hu antibody. The LC-MS/MS analyses were supported by the "Nanotechnology Network Project" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. In particular, Dr. Taro Takemura of the National Institute for Material Sciences kindly helped us with handling the LC-MS/MS devices. We also thank Dr. Flaminia for her helpful proof reading.

Author Contributions

Conceived and designed the experiments: RLO. Performed the experiments: RLO YU. Analyzed the data: RLO YU SS KM. Contributed reagents/materials/analysis tools: RLO SS KM. Wrote the paper: RLO KM.

References

- Ishihama A (2009) EcoSal—*Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. The nucleoid: an Overview Böck A, Curtiss R, III, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, Squires CL, Ussery D, eds. Washington, DC: ASM Press.
- Robinow C, Kellenberger E (1994) The bacterial nucleoid revisited. *Microbiol Rev* 58: 211–232.
- Guillen N, Le Hegaret F, Fleury AM, Hirschbein L (1978) Folded chromosomes of vegetative *Bacillus subtilis*: composition and properties. *Nucleic Acids Res* 5: 475–489.
- Murphy LD, Zimmerman SB (1997) Isolation and characterization of spermidine nucleoids from *Escherichia coli*. *J Struct Biol* 119: 321–335.
- Portelier R, Worcel A (1976) Association of the folded chromosome with the cell envelope of *E. coli*: characterization of the proteins at the DNA-membrane attachment site. *Cell* 8: 245–255.
- Yamazaki T, Yahagi S, Nakamura K, Yamane K (1999) Depletion of *Bacillus subtilis* histone-like protein, HBSu, causes defective protein translocation and induces upregulation of small cytoplasmic RNA. *Biochem Biophys Res Commun* 258: 211–214.
- Azam TA, 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.
- Drlca K, Rouviere-Yaniv J (1987) Histone-like proteins of bacteria. *Microbiol Rev* 51: 301–319.
- Grainger DC, Hurd D, Goldberg MD, Busby SJ (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34: 4642–4652.
- Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N (2006) *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res* 13: 141–153.
- Altuvia S, Almiron M, Huismann G, Kolter R, Storz G (1994) The dps promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol* 13: 265–272.
- Lucchini S, McDermott P, Thompson A, Hinton JC (2009) The H-NS-like protein StpA represses the RpoS (sigma38) regulon during exponential growth of *Salmonella Typhimurium*. *Mol Microbiol*.
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, et al. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313: 236–238.
- Oberto J, Nabti S, Jooste V, Mignot H, Rouviere-Yaniv J (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS ONE* 4: e4367.
- Weinstein-Fischer D, Elgrably-Weiss M, Altuvia S (2000) *Escherichia coli* response to hydrogen peroxide: a role for DNA supercoiling, topoisomerase I and Fis. *Mol Microbiol* 35: 1413–1420.
- Aronsson H, Jarvis P (2002) A simple method for isolating import-competent Arabidopsis chloroplasts. *FEBS Lett* 529: 215–220.
- Balandina A, Claret L, Hengge-Aronis R, Rouviere-Yaniv J (2001) The *Escherichia coli* histone-like protein HU regulates rpoS translation. *Mol Microbiol* 39: 1069–1079.
- Brescia CC, Kaw MK, Sledjeski DD (2004) The DNA binding protein H-NS binds to and alters the stability of RNA in vitro and in vivo. *J Mol Biol* 339: 505–514.
- Mayer O, Waldsich C, Grossberger R, Schroeder R (2002) Folding of the td pre-RNA with the help of the RNA chaperone StpA. *Biochem Soc Trans* 30: 1175–1180.
- Chodavarapu S, Gomez R, Vicente M, Kaguni JM (2008) *Escherichia coli* Dps interacts with DnaA protein to impede initiation: a model of adaptive mutation. *Mol Microbiol* 67: 1331–1346.
- Filutowicz M, Ross W, Wild J, Gourse RL (1992) Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J Bacteriol* 174: 398–407.
- Hwang DS, Kornberg A (1992) Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J Biol Chem* 267: 23083–23086.
- Polaczek P (1990) Bending of the origin of replication of *E. coli* by binding of IHF at a specific site. *New Biol* 2: 265–271.

24. Almirón M, Link AJ, Furlong D, Kolter R (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev* 6: 2646–2654.
25. Nair S, Finkel SE (2004) Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* 186: 4192–4198.
26. Dame RT, Noom MC, Wuite GJ (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* 444: 387–390.
27. Grant RA, Filman DJ, Finkel SE, Kolter R, Hogle JM (1998) The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol* 5: 294–303.
28. Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, et al. (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *Embo J* 25: 5591–5602.
29. Rouviere-Yaniv J, Yaniv M, Germond JE (1979) *E. coli* DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA. *Cell* 17: 265–274.
30. Swinger KK, Lemberg KM, Zhang Y, Rice PA (2003) Flexible DNA bending in HU-DNA cocystal structures. *Embo J* 22: 3749–3760.
31. Talukder AA, Iwata A, Nishimura A, Ueda S, Ishihama A (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181: 6361–6370.
32. Teramoto J, Yoshimura SH, Takeyasu K, Ishihama A (2010) A novel nucleoid protein of *Escherichia coli* induced under anaerobic growth conditions. *Nucleic Acids Res* 38: 3605–3618.
33. Kim J, Yoshimura SH, Hizume K, Ohniwa RL, Ishihama A, et al. (2004) Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res* 32: 1982–1992.
34. Takeyasu K, Kim J, Ohniwa RL, Kobori T, Inose Y, et al. (2004) Genome architecture studied by nanoscale imaging: analyses among bacterial phyla and their implication to eukaryotic genome folding. *Cytogenet Genome Res* 107: 38–48.
35. Osborn MJ, Gander JE, Parisi E, Carson J (1972) Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Isolation and characterization of cytoplasmic and outer membrane. *J Biol Chem* 247: 3962–3972.
36. Schnaitman CA (1970) Protein composition of the cell wall and cytoplasmic membrane of *Escherichia coli*. *J Bacteriol* 104: 890–901.
37. Weiner JH, Li L (2008) Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim Biophys Acta* 1778: 1698–1713.
38. Cunha S, Odijk T, Suleymanoglu E, Woldringh CL (2001) Isolation of the *Escherichia coli* nucleoid. *Biochimie* 83: 149–154.
39. Kornberg T, Lockwood A, Worcel A (1974) Replication of the *Escherichia coli* chromosome with a soluble enzyme system. *Proc Natl Acad Sci U S A* 71: 3189–3193.
40. Maternan EC, Van Gool AP (1978) Nucleoid release from *Escherichia coli* cells. *J Bacteriol* 133: 878–883.
41. Pettijohn DE, Clarkson K, Kossman CR, Stonington OG (1970) Synthesis of ribosomal RNA on a protein-DNA complex isolated from bacteria: a comparison of ribosomal RNA synthesis in vitro and in vivo. *J Mol Biol* 52: 281–300.
42. Pettijohn DE, Stonington OG, Kossman CR (1970) Chain termination of ribosomal RNA synthesis in vitro. *Nature* 228: 235–239.
43. Stonington OG, Pettijohn DE (1971) The folded genome of *Escherichia coli* isolated in a protein-DNA-RNA complex. *Proc Natl Acad Sci U S A* 68: 6–9.
44. Murphy LD, Zimmerman SB (1995) Condensation and cohesion of lambda DNA in cell extracts and other media: implications for the structure and function of DNA in prokaryotes. *Biophys Chem* 57: 71–92.
45. Zimmerman SB (2006) Cooperative transitions of isolated *Escherichia coli* nucleoids: implications for the nucleoid as a cellular phase. *J Struct Biol* 153: 160–175.
46. Alberts BM, Amodio FJ, Jenkins M, Gutmann ED, Ferris FL (1968) Studies with DNA-cellulose chromatography. I. DNA-binding proteins from *Escherichia coli*. *Cold Spring Harb Symp Quant Biol* 33: 289–305.
47. Ramsden JJ, Vohradsky J (1998) Zipf-like behavior in prokaryotic protein expression. *Physical Review E* 58: 7777–7780.
48. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, et al. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 4: 1265–1272.
49. Vallet-Gely I, Donovan KE, Fang R, Joung JK, Dove SL (2005) Repression of phase-variable cup gene expression by H-NS-like proteins in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 102: 11082–11087.
50. Washburn MP, Wolters D, Yates JR, 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19: 242–247.
51. Liu H, Sadygov RG, Yates JR, 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76: 4193–4201.
52. Phadtare S, Tadigotla V, Shin WH, Sengupta A, Severinov K (2006) Analysis of *Escherichia coli* global gene expression profiles in response to overexpression and deletion of CspC and CspE. *J Bacteriol* 188: 2521–2527.
53. Tani TH, Khodursky A, Blumenthal RM, Brown PO, Matthews RG (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A* 99: 13471–13476.
54. del Castillo T, Duque E, Ramos JL (2008) A set of activators and repressors control peripheral glucose pathways in *Pseudomonas putida* to yield a common central intermediate. *J Bacteriol* 190: 2331–2339.
55. Kim J, Jeon CO, Park W (2008) Dual regulation of *zwf-1* by both 2-keto-3-deoxy-6-phosphogluconate and oxidative stress in *Pseudomonas putida*. *Microbiology* 154: 3905–3916.
56. Fujimoto DF, Higginbotham RH, Sterba KM, Maleki SJ, Segall AM, et al. (2009) *Staphylococcus aureus* SarA is a regulatory protein responsive to redox and pH that can support bacteriophage lambda integrase-mediated excision/recombination. *Mol Microbiol* 74: 1445–1458.
57. Karavolos MH, Horsburgh MJ, Ingham E, Foster SJ (2003) Role and regulation of the superoxide dismutases of *Staphylococcus aureus*. *Microbiology* 149: 2749–2758.
58. Vlamis-Gardikas A (2008) The multiple functions of the thiol-based electron flow pathways of *Escherichia coli*: Eternal concepts revisited. *Biochim Biophys Acta* 1780: 1170–1200.
59. Mendoza-Cozatl D, Loza-Tavera H, Hernandez-Navarro A, Moreno-Sanchez R (2005) Sulfur assimilation and glutathione metabolism under cadmium stress in yeast, protists and plants. *FEMS Microbiol Rev* 29: 653–671.
60. Sekine K, Fujiwara M, Nakayama M, Takao T, Hase T, et al. (2007) DNA binding and partial nucleoid localization of the chloroplast stromal enzyme ferredoxin:sulfite reductase. *FEBS J* 274: 2054–2069.
61. Nakamura M, Kuramata M, Kasugai I, Abe M, Youssefian S (2009) Increased thiol biosynthesis of transgenic poplar expressing a wheat O-acetylserine(thiol) lyase enhances resistance to hydrogen sulfide and sulfur dioxide toxicity. *Plant Cell Rep* 28: 313–323.
62. Morikawa K, Ohniwa RL, Kim J, Maruyama A, Ohta T, et al. (2006) Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. *Genes Cells* 11: 409–423.
63. Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, et al. (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* 10: 291.
64. Kano Y, Imamoto F (1990) Requirement of integration host factor (IHF) for growth of *Escherichia coli* deficient in HU protein. *Gene* 89: 133–137.
65. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, et al. (2003) Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A* 100: 4678–4683.
66. Podkovyrov SM, Larson TJ (1996) Identification of promoter and stringent regulation of transcription of the *fabH*, *fabD* and *fabG* genes encoding fatty acid biosynthetic enzymes of *Escherichia coli*. *Nucleic Acids Res* 24: 1747–1752.
67. Fujita Y, Matsuoka H, Hirooka K (2007) Regulation of fatty acid metabolism in bacteria. *Mol Microbiol* 66: 829–839.
68. Kajitani M, Kato A, Wada A, Inokuchi Y, Ishihama A (1994) Regulation of the *Escherichia coli* *hfq* gene encoding the host factor for phage Q beta. *J Bacteriol* 176: 531–534.
69. Wooley RE, Blue JL (1975) *In-vitro* effect of edta-tris-lysozyme solutions on selected pathogenic bacteria. *J Med Microbiol* 8: 189–194.
70. Schindler CA, Schurhard VT (1964) Lysostaphin: A New Bacteriolytic Agent for the *Staphylococcus*. *Proc Natl Acad Sci U S A* 51: 414–421.
71. Sakata N, Terakubo S, Mukai T (2005) Subcellular location of the soluble lytic transglycosylase homologue in *Staphylococcus aureus*. *Curr Microbiol* 50: 47–51.
72. Forsgren A, Sjoquist J (1966) "Protein A" from *S. aureus*. I. Pseudo-immune reaction with human gamma-globulin. *J Immunol* 97: 822–827.
73. Letain TE, Postle K (1997) TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Mol Microbiol* 24: 271–283.
74. Cascales E, Lloubes R, Sturgis JN (2001) The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. *Mol Microbiol* 42: 795–807.
75. Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, et al. (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93: 93–101.
76. Hanada M, Nishiyama K, Tokuda H (1996) SecY plays a critical role in protein translocation in the absence of the proton motive force as well as at low temperature. *FEBS Lett* 381: 25–28.
77. Hazelbauer GL, Falke JJ, Parkinson JS (2008) Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33: 9–19.
78. Defeu Soufo HJ, Reimold C, Linne U, Knust T, Gescher J, et al. (2010) Bacterial translation elongation factor EF-Tu interacts and colocalizes with actin-like MreB protein. *Proc Natl Acad Sci U S A* 107: 3163–3168.
79. Kruse T, Blagov B, Lobner-Olesen A, Wachi M, Sasaki K, et al. (2006) Actin homolog MreB and RNA polymerase interact and are both required for chromosome segregation in *Escherichia coli*. *Genes Dev* 20: 113–124.
80. van den Ent F, Johnson CM, Persons L, de Boer P, Lowe J (2010) Bacterial actin MreB assembles in complex with cell shape protein RodZ. *Embo J* 29: 1081–1090.
81. Yung BY, Kornberg A (1988) Membrane attachment activates *dnaA* protein, the initiation protein of chromosome replication in *Escherichia coli*. *Proc Natl Acad Sci U S A* 85: 7202–7205.
82. Ichihashi N, Kurokawa K, Matsuo M, Kaito C, Sekimizu K (2003) Inhibitory effects of basic or neutral phospholipid on acidic phospholipid-mediated dissociation of adenine nucleotide bound to DnaA protein, the initiator of chromosomal DNA replication. *J Biol Chem* 278: 28778–28786.

83. Sekimizu K, Kornberg A (1988) Cardiolipin activation of dnaA protein, the initiation protein of replication in *Escherichia coli*. *J Biol Chem* 263: 7131–7135.
84. Suzuki E, Kondo T, Makise M, Mima S, Sakamoto K, et al. (1998) Alteration in the contents of unsaturated fatty acids in dnaA mutants of *Escherichia coli*. *Mol Microbiol* 28: 95–102.
85. Daghfous D, Chatti A, Marzouk B, Landoulsi A (2006) Phospholipid changes in *seqA* and *dam* mutants of *Escherichia coli*. *C R Biol* 329: 271–276.
86. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27–30.
87. Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* 85: 2444–2448.
88. Wang L, Brown SJ (2006) BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. *Nucleic Acids Res* 34: W243–248.
89. Lai EM, Nair U, Phadke ND, Maddock JR (2004) Proteomic screening and identification of differentially distributed membrane proteins in *Escherichia coli*. *Mol Microbiol* 52: 1029–1044.
90. Horler RS, Butcher A, Papangelopoulos N, Ashton PD, Thomas GH (2009) EchoLOCATION: an in silico analysis of the subcellular locations of *Escherichia coli* proteins and comparison with experimentally derived locations. *Bioinformatics* 25: 163–166.
91. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, et al. (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26: 1608–1615.