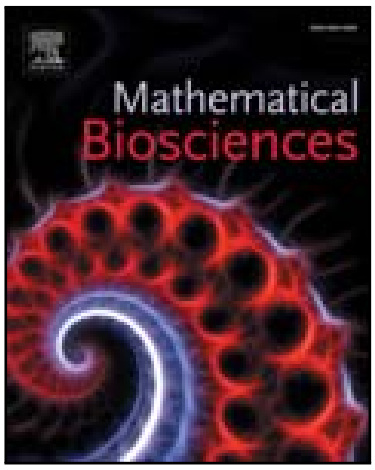

Physical properties of DNA may direct the binding of nucleoid-associated proteins along the E. coli genome

Guoqing Liu , Qin Ma , Ying Xu

PII: S0025-5564(17)30574-6
DOI: 10.1016/j.mbs.2018.03.026
Reference: MBS 8059

To appear in: *Mathematical Biosciences*

Received date: 20 October 2017
Revised date: 22 February 2018
Accepted date: 28 March 2018

Please cite this article as: Guoqing Liu , Qin Ma , Ying Xu , Physical properties of DNA may direct the binding of nucleoid-associated proteins along the E. coli genome, *Mathematical Biosciences* (2018), doi: 10.1016/j.mbs.2018.03.026

Highlights

- We characterized the binding sites of several important NAPs in E. coli K12 using two physical descriptors, intrinsic DNA curvature and flexibility.
- The binding sites of Fis, IHF, and FNR are shown to be intrinsically curved, while H-NS and Dps show no preference to specific curvature.
- The binding sites of Fis, IHF, Dps and FNR are flexible.
- The intrinsic curvature and flexibility at the binding sites of Fis and IHF are found to be coupled with the sequence specificity required in their binding, while those for Dps and FNR are independent of sequence specificity.

Physical properties of DNA may direct the binding of nucleoid-associated proteins along the *E. coli* genome

Guoqing Liu[1,2,*], Qin Ma[2,4], Ying Xu[2,3,*]

[1] School of Life Science and Technology, Inner Mongolia University of Science and Technology, Baotou, 014010, China

[2] Computational Systems Biology Laboratory, Department of Biochemistry and Molecular Biology, and Institute of Bioinformatics, the University of Georgia, Athens, GA 30602, USA.

[3] College of Computer Science and Technology, Jilin University, Changchun, 130012, China

[4] Bioinformatics and Mathematical Biosciences Lab, Department of Agronomy, Horticulture and Plant Science, South Dakota State University, SD, 57007, USA.

[*]Correspondence: gqliu1010@163.com; xyn@uga.edu

Abstract: Nucleoid-associated proteins (NAPs) play important roles in both chromosome packaging and gene regulation in bacteria. The underlying mechanisms, however, remain elusive particularly for how NAPs contribute to chromosome packaging. We report here a characterization of the binding sites for several major NAPs in *E. coli*, namely H-NS, IHF, Fis, Dps and a non-NAP protein, FNR, in terms of the physical properties of their binding DNA. Our study shows that (i) as compared with flanking regions, the binding sites for IHF, Fis and FNR tend to have high intrinsic curvature, while no characterized pattern of intrinsic curvature distribution around those of H-NS and Dps; (ii) all the binding sites analyzed in this study except those of H-NS are characterized by high structural flexibility; (iii) the intrinsic curvature and flexibility at the binding sites for Fis and IHF are found to be coupled with the sequence specificity required in their binding, while the physical properties of the binding regions for both Dps and FNR are independent of sequence specificity. Our data suggest that physical properties of DNA sequence may contribute to binding of NAPs and mediate genome packaging and transcriptional regulation of the downstream genes. Our results should be informative for prediction of NAPs binding sites and understanding of the bacterial chromosome packaging.

Keywords: nucleoid-associated protein; chromosome packaging; DNA intrinsic curvature; DNA flexibility

INTRODUCTION

Nucleoid-associated proteins (NAPs) play important roles in gene regulation and chromosome packaging in bacteria [1-4]. Several NAPs, such as H-NS, HU, IHF, Fis, Dps and StpA have been identified and studied [5]. These proteins are believed to play two types of functional roles in chromosomal folding: bending and bridging. HU and IHF have been found to act as DNA benders [6] while H-NS was shown to be, at least in some cases, involved in the formation of DNA duplex by bridging, resulting in DNA loops where RNA polymerase can be trapped and gene transcription is then interrupted [7,10]. Fis, which is an abundantly expressed protein in the exponential growth phase, can also bend DNA [11,12]. In addition, a stationary phase-specific protein, Dps, was shown to play a role in genome condensation [12,13].

Generally, the functions of an NAP may not be limited to one type. For example, H-NS has been found to either stiffen or bridge DNA molecules [8,9], bringing about diverse influences on DNA compaction and gene regulation. In addition, H-NS is known to regulate gene expression in response to changes in the intra- and extracellular environment, such as pH and temperature [14,15]. It is also likely that the concentrations of certain NAPs, which can alter the kinetics of their binding to DNA target sites, affect their regulatory effects [16].

While increasingly more data such as those collected from ChIP-seq and high-resolution imaging have revealed the importance of NAPs in the packaging of bacterial chromosomes [5,17-20], the detailed mechanisms, particularly in terms of if and how the physical properties of DNA molecules may play important roles in such a process, remain elusive. For example, *in vivo* studies have found that H-NS proteins form a few large compact clusters along the *E. coli* genomic DNA [5], which might be driven by oligomerization of the proteins. Specifically, the authors proposed that H-NS may act as a global regulator in the packing of the chromosome by linking two distant loci on the DNA. However, as shown in a subsequent study [17], the H-NS clustering might be the result of the added fluorescent proteins, whose dimerization can induce aggregation of the target proteins, although H-NS has been previously found to be able to oligomerize without the fluorescent protein. Furthermore, a chromosome conformation analysis showed no evidence of binding site clustering for NAPs such as Fis, IHF and H-NS [21]. The authors of the study proposed that the folded chromosome of *E. coli* is determined by DNA replication and transcription machineries [21] rather than NAPs such as H-NS. Based on co-expression analyses at a genome scale, Ma *et al*. proposed that the folded structure of the

chromosome may adopt a few distinct conformations consistent with a few general physiological conditions; and for each such condition, the folded conformation is driven by minimization of the total energy needed to unfold the chromosome structure to enable transcription of genes that need to be activated under the current condition, presumably assisted by NAPs [22]. Liu *et al*. has recently proposed a model for circular genome packaging based on a uniform dualtoroidal-spool conformation formed by circular plasmid DNA and claimed that the formation of H-NS clusters in *E. coli* cells during cell cycle is in close agreement with their model [23].

Knowing the possible roles of NAPs in chromosome packaging and gene regulation, it is imperative to take a deeper look at how NAPs bind with DNA in general. It has been established that multiple factors affect NAPs-DNA binding. One is the *cis* factors of DNA, such as DNA intrinsic curvature and sequence-specificity. For example, it is possible that some NAPs prefer to bind genomic regions with high curvature [24], which provides a structural scaffold for NAP binding. In addition, some studies suggested that NAPs binding sites have conserved motifs [25,26], suggesting sequence-level specificity in NAPs binding. Another is the *trans* factors, such as the binding competency of the proteins (e.g. transcription factors and DNA polymerases) to the DNA. While substantial progress has been made in term of our knowledge of the NAPs binding mechanism, a comprehensive re-evaluation of their binding sites in terms of sequence specificity and physical properties is required for improved understanding of NAP-DNA binding.

In this report, we present two physical descriptors, intrinsic DNA curvature and flexibility, to characterize the binding sites for several important NAPs in *E. coli K12*, and discuss their roles in NAPs binding to DNA.

DATA AND METHODS

Data

We have analyzed all the DNA binding sites of three NAPs (H-NS, Fis and IHF) and one non-NAP DNA-binding protein FNR given in the RegulonDB database [27]. Among the known NAPs, HU and StpA are not analyzed here due to the small sample sizes (<10) of their binding sites in RegulonDB. FNR is a transcription factor with no known functional role in genome packaging, and hence included in this study as a comparison. In addition to their binding sites in RegulonDB, we have also analyzed the binding regions for H-NS, Fis, IHF, Dps and FNR derived from relevant ChIP-chip and ChIP-seq data [16,25,28], which are classified into two equal-sized groups: strong *vs*. weak binding sites

according to the binding intensities measured by peak heights in the ChIP data. Table 1 lists the proteins along with the numbers of their binding sites. Two versions of the complete genome of *E. coli* K-12 MG1655 (version U00096.2 and U00096.3) available at Genbank (http://www.ncbi.nlm.nih.gov/) have been used to retrieve sequences of the binding sites. The current version (U00096.3) differs from U00096.2 in genome size and coordinates of genomic elements, and therefore one should be cautious when retrieving sequences of binding sites from the genome according to their coordinates. The genomic coordinates of Dps-binding regions [28] used in this study correspond to the genome version U00096.3, and all other coordinates refer to U00096.2.

Intrinsic curvature of DNA

DNA curvature is sequence-dependent and changes with the thermal energy of the environment. The ensemble average of the curvature of a DNA segment is given as [29]:

$$< C >= C_0 + < \chi > \qquad (1)$$

where C_0 represents the intrinsic curvature and χ is for the dynamic fluctuation, with $< \chi >$ = 0. The intrinsic curvature is the statistical average of the curvature for a given sequence or alternatively, the time average of the DNA superstructure.

DNA bending is caused largely by base-step parameters roll and tilt, the contributions of which to global curvature of the sequence are modulated by cumulative twist [30]. For example, as illustrated in Figure S1, in case of DNA bending in the xz plane toward minor groove at the central base-pair, opening of the roll defined between the central base-pair and its neighbors toward the major groove would facilitate the bending. The contribution of roll to the curvature can readily be seen, assuming twist is zero (Figure S1B). The intrinsic curvature of a DNA segment, *L* bps in size, centered at position *i* in the sequence is defined as previously formulated [30-32]:

$$C_i = \sum_{i=-\frac{L-1}{2}}^{\frac{L-1}{2}} \left[\rho_i \cos \Omega_i + \tau_i \sin \Omega_i \right] \qquad (2)$$

where ρ_i and τ_i are, respectively, the roll and tilt angles at equilibrium and Ω_i is the cumulative helical twist at dinucleotide step *i* relative to the center of the sequence [30-32]. Average intrinsic curvature of the (naked) DNA sequence per base-pair step is

$$< C_i >= \frac{C_i}{L-1} \quad (3)$$

The equilibrium roll and tilt values used in Eq.(2) are taken from literature [29] and a constant twist value of 34.8° for all dinucleotide steps, which is the mean of twists for naked DNA [30], is used in cumulative twist calculation. This set of roll and tilt values were first obtained by minimizing the conformational energy of the different dinucleotide steps and later refined to optimize the agreement with experimental data [33]. Successful applications of the set of roll and tilt values, as well as DNA rigidity parameters described later in this work in many areas (e.g. predicting experimentally-determined curvature of DNA tracts, gel electrophoretic retardations of duplex oligonucleotides, etc) confirmed their reliability [33].

It is noteworthy that the intrinsic curvature denoted by Eq.(2) is similar to the real part of the complex number previously used to calculate the intrinsic curvature [29], but differs in the cumulative helical twist calculation. We calculated the cumulative twist from sequence center toward both ends of the sequence, and therefore the Eq.(2) measures the curvature of DNA sequence in a plane (e.g. the xz plane in Figure S1), which is determined jointly by helical axis and one being orthogonal to the helical axis pointing in the direction of major groove at the central base-pair (see Ref. [30,32] for details). As a result, Eq.(2) may not necessarily capture the maximum intrinsic curvature of the sequence in 3-dimensional space. To approximate the maximum intrinsic curvature, we have used the following procedure: (1) calculate the intrinsic curvature at each position in the sequence using a sliding window of 31 bps; (2) select local maximal values using a second sliding window of 14 bps within the first; (3) fit the local maxima along the sequence using a cubic spline, and then calculate the maximum intrinsic curvature at each nucleotide position using the spline. The first window size, 31, is selected based on following consideration: a window size larger than 31 tends to lose the local intrinsic curvature information for proteins whose binding sites are about 20 bps long, and a size much smaller than 31 tends to make the data too noisy. The second window size is so selected to capture one local maximum within each cycle (~10 bps) of the double helical structure. Both numbers are selected through trial and errors guided by the above consideration, to give the best estimation.

DNA flexibility

DNA flexibility is also a sequence-dependent property and inversely correlated with the sequence rigidity, where the rigidity parameters were retrieved from [29], which measure the normalized melting temperatures for dinucleotides. The average flexibility of a DNA segment can be assessed by using the inverse of the rigidity averaged over all dinucleotides in the sequence.

Binding motif identification

For a particular set of binding regions, binding motifs were identified using MEME [34]. In the implementation of MEME, motif discovery mode is "Discriminative mode", in which the complete genome of *E. coli* is used as a control, and each input sequence is expected to contain at most one occurrence of each motif; The most enriched motif is identified by setting the number of motifs expected to find to 1; All other options are used as default.

RESULTS AND DISCUSSION

Intrinsic curvature of NAP binding sites

To test whether NAPs tend to bind with genomic regions having a specific range of curvature levels, e.g., high intrinsic curvature, we have aligned the binding sites for each NAP (Table 1) at their central positions, and obtained intrinsic curvature profiles for the regions encompassing the binding sites (Figure 1). It is obvious from the figure that all the four target proteins analyzed here prefer to bind to intrinsically curved DNA regions, regardless whether they need sequence specificity in their binding. In addition, some differences can be seen from the data: (1) Fis appears to prefer high curvature with very little fluctuation at the binding regions; (2) the preference of H-NS for high intrinsic curvature is not as clear as the other DNA-binding proteins.

We have also analyzed the binding regions for H-NS, Fis, IHF, Dps and FNR as revealed by genome-scale ChIP-chip and ChIP-seq data (Table 1). Our results are consistent with those observed above: the binding regions for IHF, Fis and FNR have high intrinsic curvature compared with the flanking regions (Figure 2) while H-NS and Dps shows little preference for high intrinsic curvature at its binding sites. Interestingly, the binding regions for FNR show higher levels of intrinsic curvature compared with those of the NAPs.

To further test whether the binding of these proteins to DNA may be directed by intrinsic curvature, we then compared the intrinsic curvature between the binding regions with high binding affinity (strong binding sites) and those with low binding affinity (weak binding sites). For Fis, the intrinsic

curvature at the central binding regions, 101 bps in size, is significantly higher for strong binding sites than weak ones (t-test: p-value < 0.0001) (Figure 3A), indicating that the intrinsic curvature may positively contribute to Fis' binding affinity. In contrast, the binding affinity for H-NS, IHF, Dps and FNR does not seem to depend on the level of intrinsic curvature (Figure 3B-E), suggesting that other factors may play a role in determining their binding affinity (see next section).

Is there any relationship between intrinsic curvature and binding motif enrichment in the binding sites? To test this, we have identified the most enriched binding motif (or best motif) in the 101-bp regions aligned at their summits as described in [25] using MEME (see Methods for implementation details), and then examined the curvature change once we remove the motif from the binding regions, the two flanking regions surrounding the motif being concatenated. For Fis, a highly conserved motif has been found to occur in the majority (93% for early exponential phase and 98% for mid exponential phase) of their binding sites, and the removal of such binding motifs indeed leads to a remarkable decrease in the intrinsic curvature (Figure 4 A-B), suggesting that the high intrinsic curvature at the Fis binding sites is largely attributable to the highly enriched binding motifs. For IHF, 37% of their binding sites are enriched with a motif and a slight decrease in intrinsic curvature is observed when the binding motif is removed. One exception is FNR, where the intrinsic curvature of binding sites is independent of the sequence-level conservation (Figure 4D).

To further test if the difference in peak height of ChIP-seq data results from the relative binding affinities of Fis to individual motifs, we have calculated the correlation between the peak height of ChIP-seq data of Fis and the similarity between individual binding motifs and their consensus sequence, where the consensus sequence has been derived as described in [16] from a position weight matrix, which has been constructed from an alignment among the binding motifs of Fis. The similarity is computed using PatSer [35]. We noted that the similarity level positively correlates with the peak height (Figure 5), suggesting that intrinsic curvature coupled with sequence specificity plays a positive role in the binding of Fis.

DNA flexibility at NAP binding sites

Flexibility is another physical property of DNA sequence that may affect the binding affinity with proteins. Generally, more flexible regions tend to wrap more easily around binding proteins. To see how DNA flexibility affects the binding of proteins of interest, we plotted the DNA rigidity profile around the protein-binding sites in the Figure 6. We can see: the binding sites for H-NS, Fis, IHF and

FNR, retrieved from regulonDB, show a remarkable reduction in structural rigidity at the binding sites, regardless if they are involved in genome packing or not. However, a more detailed analysis of the binding regions derived from ChIP-seq data reveals rather complex relations between the flexibility of the binding sites and binding affinity (Figure 7).

Specifically, we noted: (i) strong binding regions for Fis tend to have reduced rigidity, while weak binding regions have no clear preference to a specific range of DNA flexibility (Figure 7A); (ii) H-NS the overall level of DNA rigidity around H-NS binding sites is as low as that for other proteins, but the rigidity is slightly elevated at the summit with highest binding intensity compared to flanking regions (Figure 7B); (iii) there seems to be no substantial difference in DNA rigidity between the strong and weak binding regions for H-NS (Figure 7B); and (iv) IHF, Dps and FNR prefer binding regions with high degree of flexibility and this preference is stronger in strong binding sites than in weak ones (Figure 7 C-E). By removing the most enriched binding motifs for Fis and IHF from the central binding regions of 101 bps, we found a sharp increase in DNA rigidity, suggesting the motif-dependent high flexibility (or low rigidity) (Figure 8). In comparison, the low level of DNA rigidity at the binding sites for FNR remains remarkable when the most enriched binding motifs were removed. Given that an identified motif for Dps-binding sites occurs only in a small fraction of the binding sites (29/451) and the binding sites containing the motif have a relatively high level of rigidity (Figure 8D), it is evident that the reduced rigidity at the Dps-binding sites shown in Figure 7D is not caused by this motif.

DISCUSSION

How do the DNA properties like intrinsic curvature and flexibility revealed in this study affect the binding of NAPs and favor a particular type of DNA-binding mode of NAPs ? Single-molecule experiments have shown that Fis are able to bend DNA and condense the genome by juxtaposition of remotes DNA sites together [11,12]; IHF was reported to non-specifically binds to DNA and is able to result in various organizations of DNA such as local bending and global DNA condensation [36,37], depending on its multiple distinct binding modes in response to solution conditions. Therefore, the preference of Fis and IHF for intrinsically curved and flexible DNA discovered in this study is suggestive of its positive role in bending the DNA, as genomic regions with higher curvature and flexibility are generally easy to bend.

The results for H-NS in this study are probably related to its DNA-binding modes. H-NS has two distinct DNA-binding modes: DNA-stiffening mode and DNA-bridging mode [7-9]. In DNA-stiffening

mode, H-NS polymerizes along DNA to form an extended rigid nucleoprotein filament [9,38]. In addition to the polymerization effect, a tight binding between H-NS and DNA via strong interaction bonds may also be required to form a rigid nucleoprotein filament. Despite a slight increase in DNA rigidity at the summits of binding regions for H-NS, the overall flexibility of binding sites for H-NS is high (e.g. average of the rigidity values at position -200 to 200 in Figure 7B is as low as the minima for other proteins), which might assist the tight binding. The little dependence of H-NS binding on the intrinsic curvature of DNA suggests that curvature does not play an important role in the binding of H-NS to DNA, regardless of its two DNA-binding modes. Because of the presence of H-NS polymerization along the packaged genome of *E.coli*, we think that at least a part of the binding regions for H-NS may be the result of H-NS polymerization and cannot represent true binding preference. Moreover, another binding mode of H-NS, DNA-bridging mode, may also need different physical properties at the binding sites for H-NS. Therefore, we believe that discriminated identification and analysis of binding sites for the two binding modes in future are able to provide a deeper insight into how DNA physical properties facilitate the binding of H-NS.

Dps is the most abundant NAP during the early stationary growth phase and was shown to participate in genome condensation and gene regulation, presumably in a cross-talk manner with other proteins like IHF [39]. Previous studies have shown that Dps binds to DNA without apparent sequence specificity [39,40]. A recent work based on ChIP-seq experiments have reported a non-random distribution of Dps binding sites across the *E.coli* genome in exponentially growing cells [28]. Our results based on this ChIP-seq data show that there is no extensively expressed motif in the 101-bp core regions centered at the summits of the binding sites for Dps, supporting the notion of sequence non-specificity in Dps binding. However, we found that the binding regions are more flexible compared with their flanking regions and the high level of flexibility has nothing to do with sequence motif. FNR shows a similar pattern in this regard. These suggest that high DNA flexibility is likely to be a target that acts as a "structure-specificity" in the binding of the proteins.

Overall, our main findings on how DNA intrinsic curvature and flexibility affect binding of nucleoid-associated proteins are consistent with expectations from the DNA-binding modes revealed in previous experiments.

To conclude, we investigated the intrinsic curvature and flexibility of binding sites for three NAPs: H-NS, Fis, IHF, Dps and one non-NAP protein, FNR in *E. coli*, and found that the binding sites for Fis,

IHF and FNR show high intrinsic curvature while H-NS and Dps show no preference to specific curvature level. The binding sites for all the proteins, except for H-NS, tend to be more flexible compared to the flanking regions. Overall, the intrinsic curvature and flexibility of the binding sites may represent key physical properties needed to facilitate high-affinity binding with proteins, in addition to sequence specificity.

ACKNOWLEDGMENT

This work was supported by grants from the National Natural Science Foundation (31660322, 61102162), Science Foundation for Excellent Youth Scholars of Inner Mongolia University of Science and Technology (2016YQL06).

REFERENCES

[1] C.J. Dorman. H-NS: a universal regulator for a dynamic genome. Nat. Rev. Microbiol., 2004, 2: 391-400.

[2] D.F. Browning, D.C. Grainger, and S.J. Busby. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. Curr. Opin. Microbiol., 2010, 13: 773-780.

[3] J. Johansson, C. Balsalobre, S.Y. Wang, J. Urbonaviciene, D.J. Jin, B. Sondén, and B.E. Uhlin. Nucleoid proteins stimulate stringently controlled bacterial promoters: a link between the cAMP-CRP and the (p)ppGpp regulons in Escherichia coli. Cell, 2000, 102: 475-485.

[4] C.J. Dorman. Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. Biochem. Soc. Trans., 2013, 41: 542-547.

[5] W. Wang, G.W. Li, C. Chen, X.S. Xie, and X. Zhuang. Chromosome organization by a nucleoid-associated protein in live bacteria. Science 2011, 333: 1445-1449.

[6] K.K. Swinger, and P.A. Rice. IHF and HU: flexible architects of bent DNA. Curr. Opin. Struct. Biol., 2004,14: 28-35.

[7] M.V. Kotlajich, D.R. Hron, B.A. Boudreau, Z. Sun, Y.L. Lyubchenko, and R. Landick. Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria. Elife，2015, 4, doi: 10.7554/eLife.04970.

[8] C.J. Lim, L.J. Kenney, and J. Yan. Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. Nucleic Acids Res., 2014, 42: 8369-8378.

[9] Y. Liu, H. Chen, L.J. Kenney, and J. Yan. A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. Genes Dev., 2010, 24: 339-344.

[10] R.T. Dame, C. Wyman, R. Wurm, R. Wagner, and N. Goosen. Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. J. Biol. Chem., 2002, 277: 2146-2150.

[11] D. Skoko, D. Yoo, H. Bai, B. Schnurr, J. Yan, S.M. McLeod, J.F. Marko, and R.C. Johnson. Mechanism of chromosome compaction and looping by the Escherichia coli nucleoid protein Fis. J. Mol. Biol., 2006, 364: 777-798.

[12] Y.T. Sato, S. Watanabe, T. Kenmotsu, M. Ichikawa, Y. Yoshikawa, J. Teramoto, T. Imanaka, A. Ishihama, and K. Yoshikawa. Structural change of DNA induced by nucleoid proteins: growth phase-specific Fis and stationary phase-specific Dps. Biophys. J. 2013, 105: 1037-1044.

[13] S.Y. Lee, C.J. Lim, P. Dröge, and J. Yan. Regulation of Bacterial DNA Packaging in Early Stationary Phase by Competitive DNA Binding of Dps and IHF. Scientific Reports, 2015, 5: 18146. doi:10.1038/srep18146

[14] T. Atlung, and H. Ingmer. H-NS: a modulator of environmentally regulated gene expression. Mol. Microbiol., 1997, 24: 7-17.

[15] S. Ono, M.D. Goldberg, T. Olsson, D. Esposito, J.C. Hinton, J.E. Ladbury. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem. J., 2005, 391: 203-213.

[16] K.S. Myers, H. Yan, I.M. Ong, D. Chung, K. Liang, F. Tran, S. Keleş, R. Landick, and P.J. Kiley. Genome-scale analysis of Escherichia coli FNR reveals complex features of transcription factor binding. PLoS Genet., 2013, 9: e1003565.

[17] S. Wang, J.R. Moffitt, G.T. Dempsey, X.S. Xie, and X. Zhuang. Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. Proc. Natl. Acad. Sci. U.S.A., 2014, 111: 8452-8457.

[18] J.K. Fisher, A. Bourniquel, G. Witz, B. Weiner, M. Prentiss, and N. Kleckner. Four-dimensional imaging of E. coli nucleoid organization and dynamics in living cells. Cell, 2013, 153: 882-895.

[19] N.J. Kuwada, B. Traxler, and P.A.Wiggins. Genome-scale quantitative characterization of bacterial protein localization dynamics throughout the cell cycle. Molecular Microbiology, 2015, 95: 64-79.

[20] C.J. Lim, Lee S.Y., L.J. Kenney, and J. Yan. Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. Sci. Rep., 2012, 2: 509

[21] C. Cagliero, R.S. Grand, M.B. Jones, D.J. Jin, and J.M. O'Sullivan. Genome conformation capture reveals that the Escherichia coli chromosome is organized by replication and transcription, Nucleic Acids Res., 2013, 41: 6058-6071.

[22] Q. Ma, Y. Yin, M.A. Schell, H. Zhang, G. Li, and Y. Xu. Computational analyses of transcriptomic data reveal the dynamic organization of the Escherichia coli chromosome under different conditions. Nucleic Acids Res., 2013, 41: 5594-5603.

[23] Y. Liu, P. Xie, P. Wang, M. Li, H. Li, W. Li, and S. Dou. A model for chromosome organization during the cell cycle in live E. coli. Sci. Rep., 2015, 5: 17133.

[24] R.T. Dame, C. Wyman, and N. Goosen. Structural basis for preferential binding of H-NS to curved DNA. Biochimie, 2001, 83: 231-234.

[25] C. Kahramanoglou, A.S. Seshasayee, A.I. Prieto, D. Ibberson, S. Schmidt, J. Zimmermann, V. Benes, G.M. Fraser, and N.M. Luscombe. Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic Acids Res., 2011, 39: 2073-2091.

[26] B. Cho, E.M. Knight, C.L. Barrett, and B.Ø. Palsson. Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. Genome Res., 2008, 18: 900-910.

[27] H. Salgado, M. Peralta-Gil, S. Gama-Castro, A. Santos-Zavaleta, L. Muñiz-Rascado, J.S. García-Sotelo, V. Weiss, H. Solano-Lira, I. Martínez-Flores, A. Medina-Rivera, et al.. RegulonDB (version 8.0): Omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Research, 2012, 41(Database issue): D203-213

[28] S.S. Antipov, M.N. Tutukina, E.V. Preobrazhenskaya, F.A. Kondrashov, M.V. Patrushev, S.V. Toshchakov, I. Dominova, U.S. Shvyreva, V.V. Vrublevskaya, O.S. Morenkov, N.A. Sukharicheva, V.V. Panyukov, and O.N. Ozoline. The nucleoid protein Dps binds genomic DNA of Escherichia coli in a non-random manner. PLoS One, 2017, 12: e0182800.

[29] A. Scipioni, C. Anselmi, G. Zuccheri, B. Samori, and P. De Santis. Sequence-dependent DNA curvature and flexibility from scanning force microscopy images. Biophys. J., 2002, 83: 2408-2418.

[30] F. Battistini, C.A. Hunter, E.J. Gardiner, and M.J. Packer. Structural mechanics of DNA wrapping in the nucleosome. J. Mol. Biol., 2010, 396: 264-279.

[31] J.Y. Wang, J. Wang, and G. Liu. Calculation of nucleosomal DNA deformation energy: its implication for nucleosome positioning. Chromosome Res., 2012, 20: 889-902.

[32] G. Liu, Y. Xing, H. Zhao, J. Wang, Y. Shang and L. Cai. A deformation energy-based model for predicting nucleosome dyads and occupancy. Sci. Rep., 2016, 6: 2413

[33] P. De Santis, S. Morosetti, and A. Scipioni. Prediction of nucleosome positioning in genomes: limits and perspectives of physical and bioinformatic approaches. J. Biomol. Struct. Dyn., 2010, 27, 747-764.

[34] T. Bailey, M. Boden, F.A. Buske, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research, 2009, 37(Web Server issue): W202-W208.

[35] G.Z. Hertz, and G.D. Stormo. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. Bioinformatics, 1999, 15: 563-577

[36] B.M. Ali, R. Amit, I. Braslavsky, A.B. Oppenheim, O. Gileadi, and J. Stavans. Compaction of single DNA molecules induced by binding of integration host factor (IHF). Proc. Natl. Acad, Sci. USA, 2001, 98: 10658-10663.

[37] J. Lin, H. Chen, P. Droge, and J. Yan. Physical organization of DNA by multiple non-specific DNA-binding modes of integration host factor (IHF). PLoS One, 2012, 7: e49885.

[38] R. Amit, A.B. Oppenheim, and J. Stavans, Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. Biophys. J., 2003, 84: 2467-2473.

[39] M. Almiron, A.J. Link, D. Furlong, and R. Kolter. A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. Genes Dev., 1992, 6: 2646-2654.

[40] P. Ceci, S. Cellai, E. Falvo, C. Rivetti, G.L. Rossi, and E. Chiancone. DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus, Nucl. Acids Res., 2004, 32: 5935-5944.

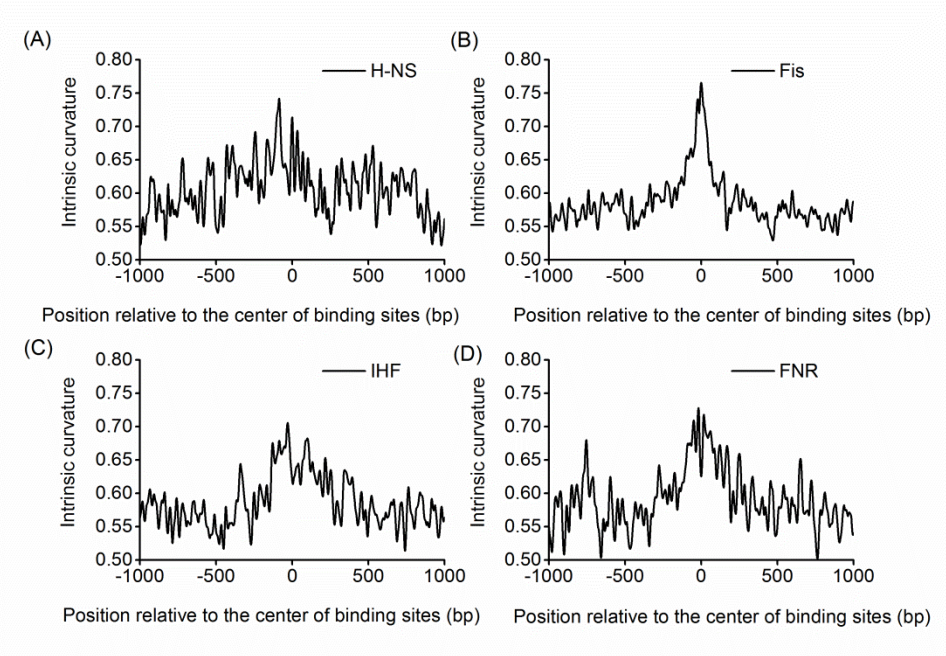

Figure 1. Intrinsic curvature around the binding sites from the RegulonDB database [27].

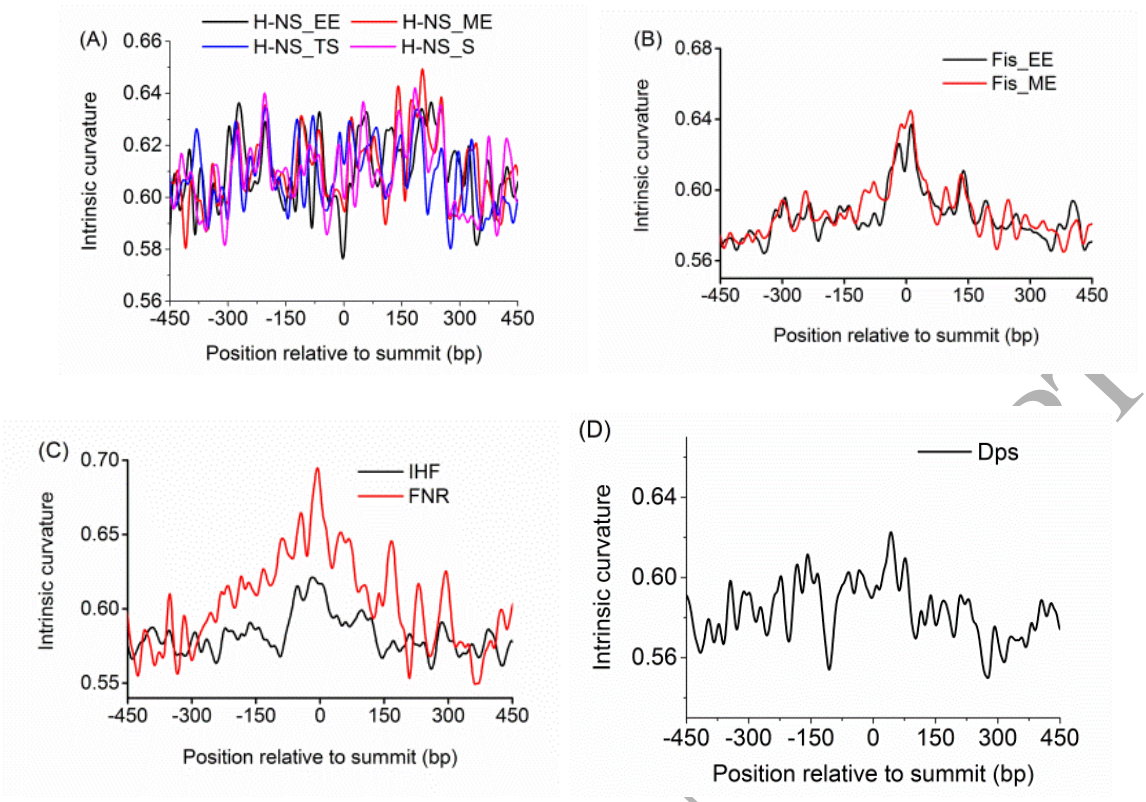

Figure 2. Intrinsic curvature around the binding sites derived from ChIP-chip and ChIP-seq data, where the summit is the position with the highest binding intensity from the ChIP-seq reads or ChIP-chip signal in each binding region [16,25,28]. EE, ME, S and TS are defined as in Table 1.

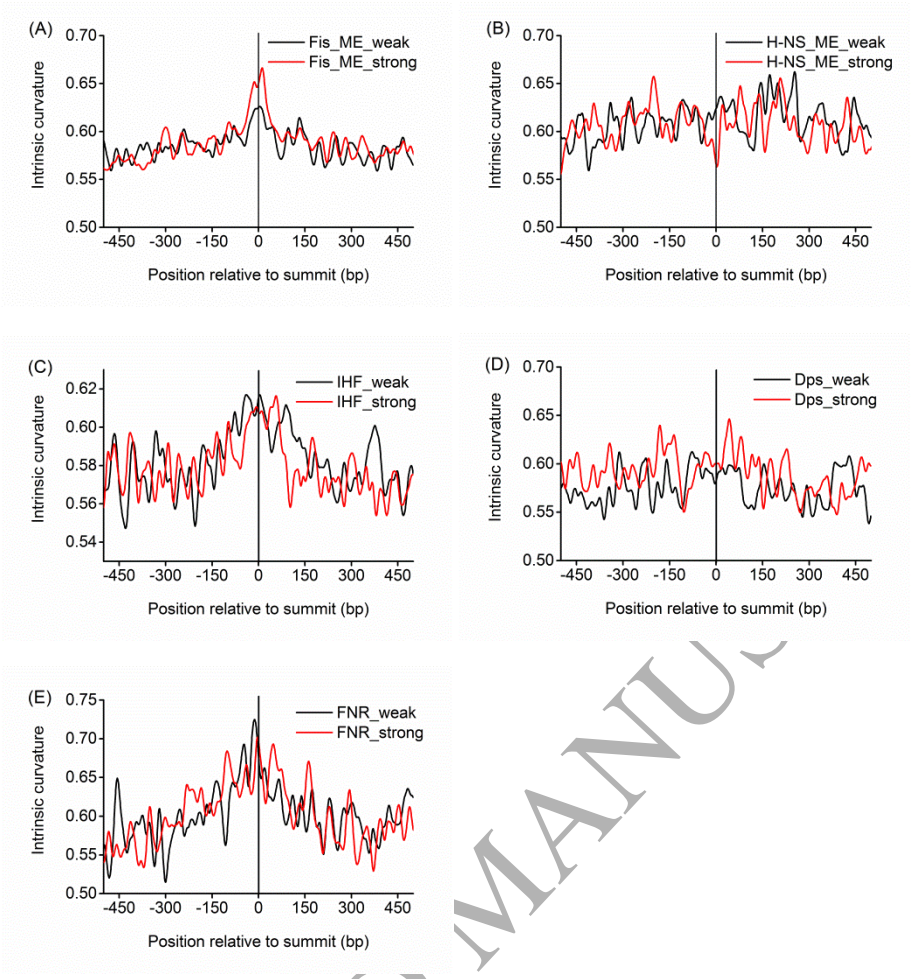

Figure 3. Intrinsic curvature around the binding sites, which are classified into two groups: those with high vs. low binding signals. For Fis and H-NS, results for the mid-exponential phase (ME) are shown in this plot, and similar results for other growth phases are shown in Figure S2.

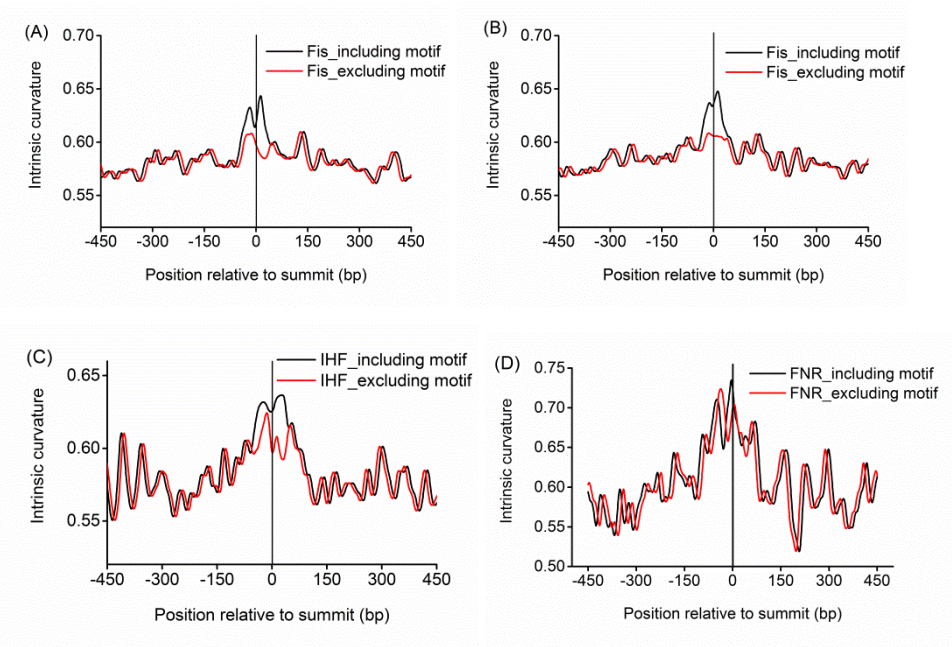

Figure 4. Change in intrinsic curvature after removal of the enriched binding motifs from the 101-bp binding regions aligned around their summits. Growth phases: (A) early exponential phase, and (B) mid-exponential phase.

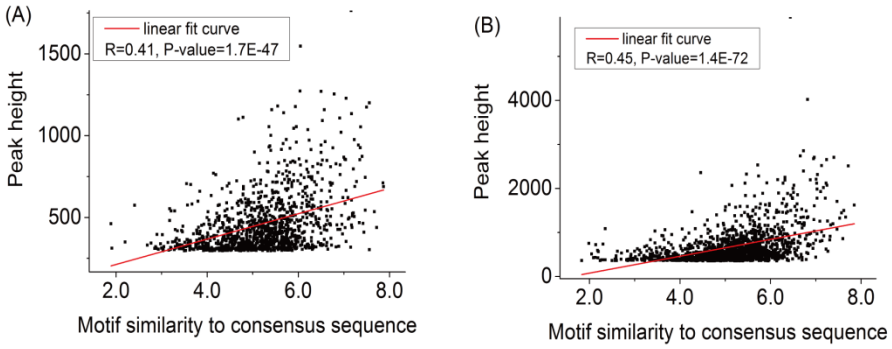

Figure 5. The peak heights of Fis's binding intensities positively correlate with the motif similarities to the consensus sequence. The binding motifs were identified in 101-bp regions aligned around the summits using MEME [34]. Growth phases: (A) early exponential phase, and (B) mid-exponential phase.

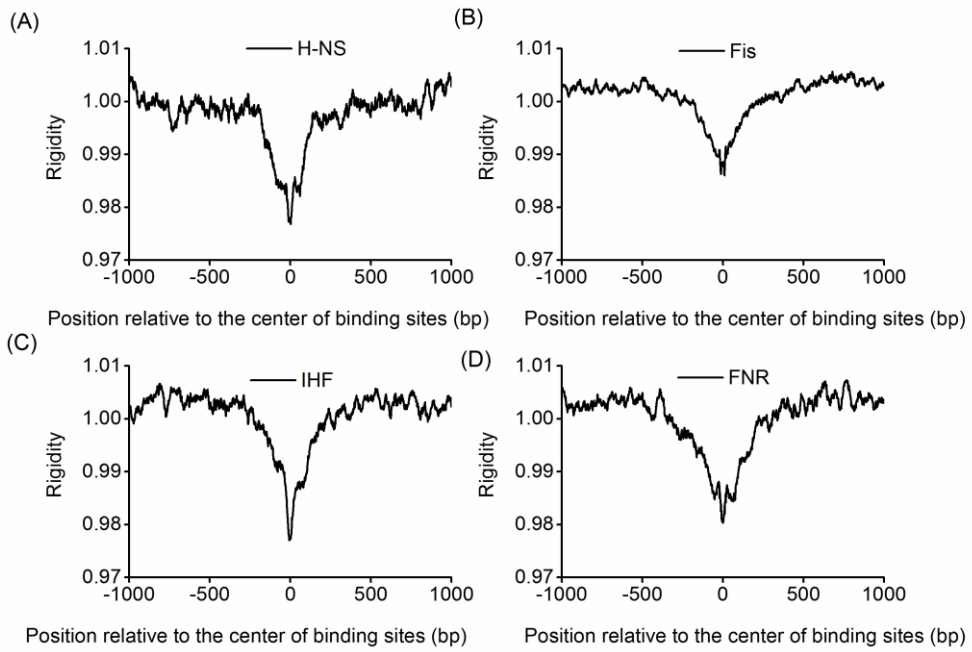

Figure 6. DNA rigidity around the binding sites from the RegulonDB database [27].

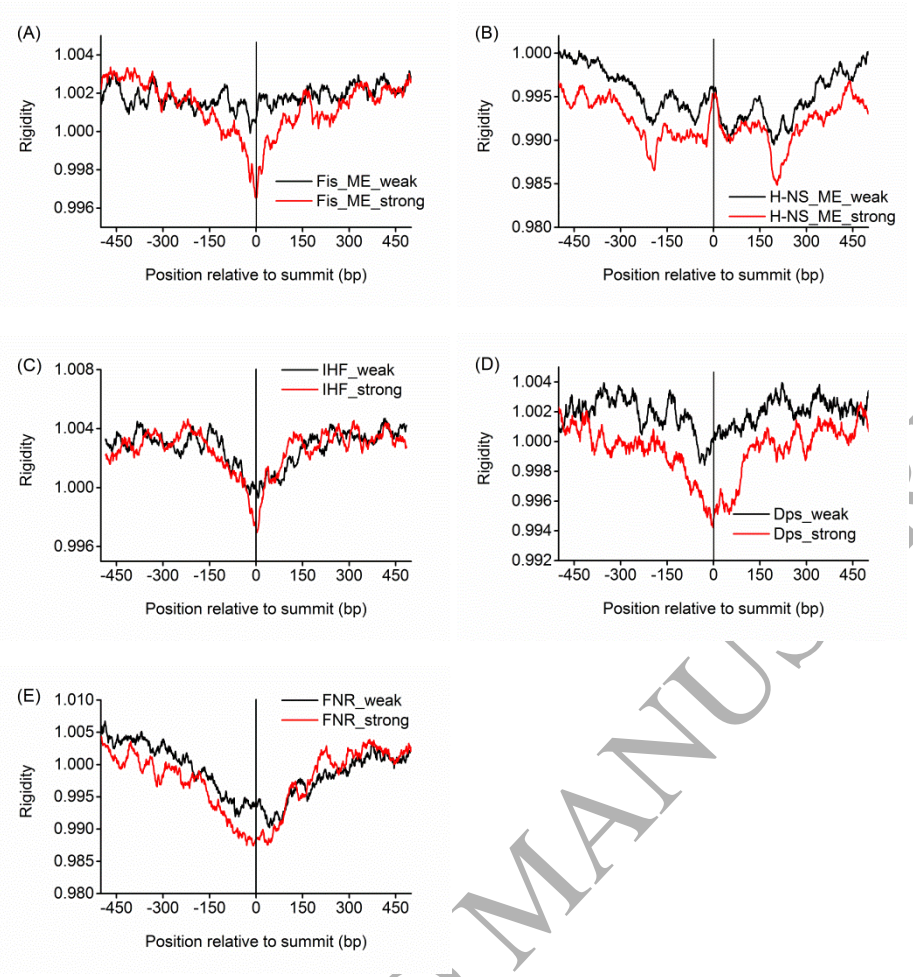

Figure 7. DNA rigidity around binding sites derived from ChIP data. ME denotes mid-exponential phase. For Fis and H-NS, similar results for other growth phases are shown in Figure S3.

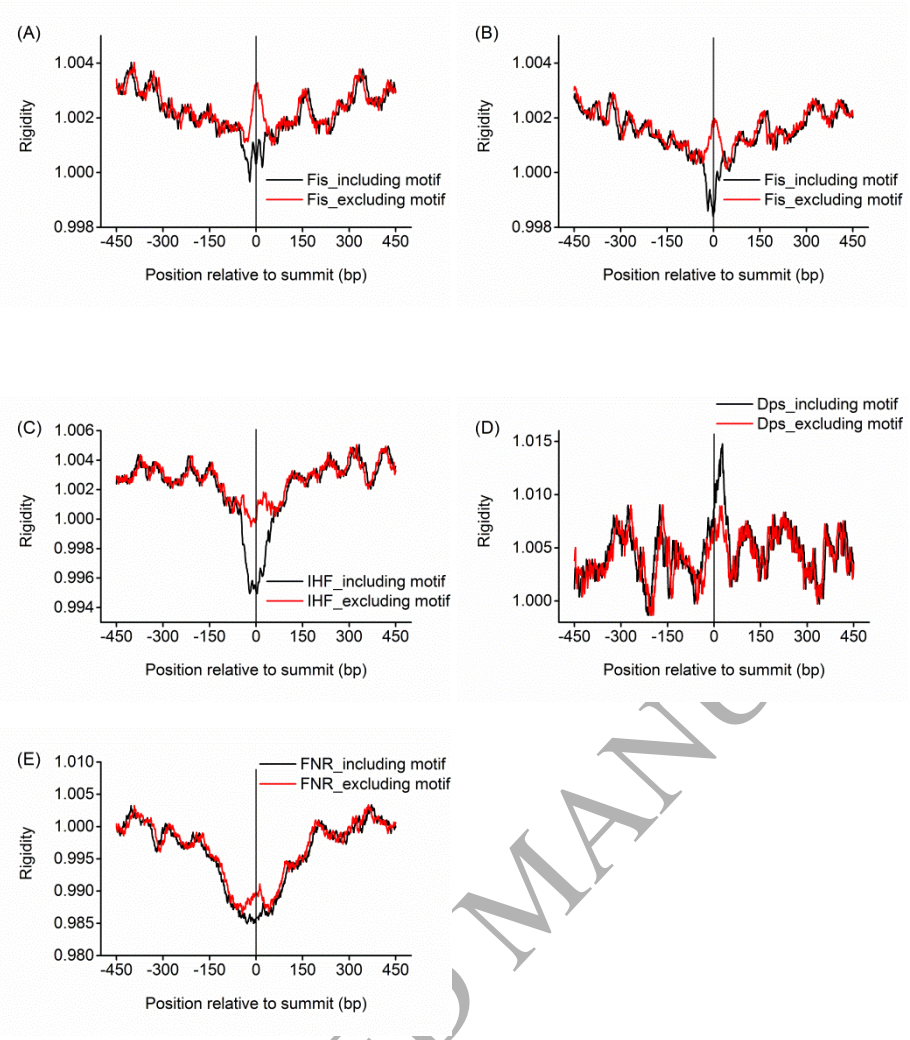

Figure 8. Change in DNA flexibility after exclusion of the most enriched binding motifs from the 101-bp regions aligned around the summits of the binding regions. (A) Early exponential phase. (B) Mid-exponential phase.

Table 1. Proteins and their binding sites analyzed.

Name	Type	Number of binding sites in RegulonDB[27]	Number of binding regions derived from ChIP data
H-NS	NAP	51	444 (EE), 458 (ME), 547 (TS), 537 (S)
Fis	NAP	222	1246 (EE), 1464 (ME)
IHF	NAP	98	1,020
Dps	NAP		451
FNR	non-NAP	87	208

Note: EE, ME, S and TS denote early, mid-exponential growth phase, stationary and transition-to-stationary phase, respectively; ChIP data for H-NS and Fis is from [25], for Dps is from [28], and for IHF and FNR is from [16].