Nucleic Acids Research, 2024, **52**, 7354–7366
https://doi.org/10.1093/nar/gkae459
Advance access publication date: 4 June 2024
Structural Biology

Structure of the *E. coli* nucleoid-associated protein YejK reveals a novel DNA binding clamp

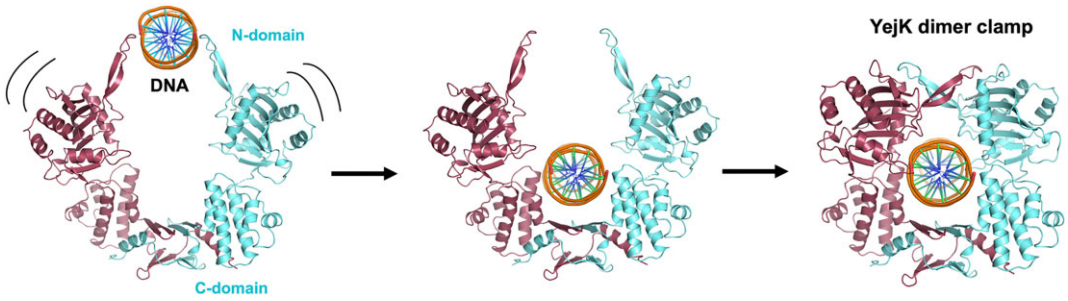
Maria A. Schumacher *, Rajiv R. Singh and Raul Salinas

Department of Biochemistry, 307 Research Dr., Box 3711, Duke University Medical Center, Durham, NC 27710, USA
*To whom correspondence should be addressed. Tel: +1 919 684 9468; Fax: +1 919 684 8885; Email: Maria.Schumacher@Duke.edu

Abstract

Nucleoid-associated proteins (NAPs) play central roles in bacterial chromosome organization and DNA processes. The *Escherichia coli* YejK protein is a highly abundant, yet poorly understood NAP. YejK proteins are conserved among Gram-negative bacteria but show no homology to any previously characterized DNA-binding protein. Hence, how YejK binds DNA is unknown. To gain insight into YejK structure and its DNA binding mechanism we performed biochemical and structural analyses on the *E. coli* YejK protein. Biochemical assays demonstrate that, unlike many NAPs, YejK does not show a preference for AT-rich DNA and binds non-sequence specifically. A crystal structure revealed YejK adopts a novel fold comprised of two domains. Strikingly, each of the domains harbors an extended arm that mediates dimerization, creating an asymmetric clamp with a 30 Å diameter pore. The lining of the pore is electropositive and mutagenesis combined with fluorescence polarization assays support DNA binding within the pore. Finally, our biochemical analyses on truncated YejK proteins suggest a mechanism for YejK clamp loading. Thus, these data reveal YejK contains a newly described DNA-binding motif that functions as a novel clamp.

Graphical abstract



Introduction

To fit within the confines of a bacterial cell, chromosomal DNA must be tightly organized while still enabling essential DNA processes. This organization is achieved by a combination of macromolecular crowding, DNA supercoiling and the binding of so-called nucleoid-associated proteins (NAPs) (1–19). NAPs have been shown to order the nucleoid DNA through several DNA-binding mechanisms (1–33). The NAPs in the model Gram-negative bacterium *Escherichia coli* are arguably the best understood. To date, approximately a dozen *E. coli* NAPs have been characterized (1,2). Among the most studied of these are integration host factor (IHF), histone-like protein (HU), suppressor of td phenotype A (StpA), histone-like nucleoid-structuring protein (H-NS) and factor for inversion stimulation (Fis) (4–19). These proteins also effect DNA processes such as transcription by directly interacting with promoter DNA or indirectly, through global modulation of chromosomal DNA conformation. Interestingly, several NAPs that impact DNA processes have been shown to bind or regulate topoisomerases (1,2,20–22,32,33). One example is Fis, which affects global supercoiling in an indirect manner by regulating the transcription of topoisomerases in a growth rate-dependent manner, thus altering DNA supercoiling at a global level (20–24). The *Caulobacter crescentus* GapR protein was recently shown to bind positive supercoiled DNA and recruit and stimulate the activity of gyrase and topoisomerase IV to relax positive supercoils, thus enabling DNA replication (32).

The *E. coli* NAP, YejK, which has also been referred to as NdpA, was first identified as a protein released from *E. coli* nucleoids after treatment with DNase I and later shown to interact with both gyrase and topoisomerase IV (33,34). Studies in the Marians lab showed that YejK impacted the activities of topoisomerase IV and gyrase in complex manners (33). Interestingly, YejK appeared to modulate DNA topology and untwist DNA upon binding. YejK has a very acidic pI of 4.4 for a DNA binding protein and, in particular, for a NAP, which typically are basic proteins. YejK contains 335 residues, which is also large for a NAP. However, data indicate that YejK is

Received: April 22, 2024. **Revised:** May 2, 2024. **Editorial Decision:** May 3, 2024. **Accepted:** May 15, 2024

present at high abundance in *E. coli* cells, at 24 000 copies per YejK monomer, which is similar to protein levels of well characterized *E. coli* NAPs, such as StpA, H-NS, IHF and CbpB (3–19).

Recent analyses have revealed an interesting feature of NAPs, which is that they are often encoded in phages and plasmids, with the number of encoded NAPs found often correlating with plasmid size (3,5–41). These findings suggest, that among other functions, NAPs may play roles in plasmid maintenance and host cell biology (36–41). Like several other NAPs, *yejK* encoding genes are present on several plasmids. In particular, *yejK* genes were identified on plasmids, p0908, pCAR1 and pQBR103, which are harbored in *Vibrio* sp. 0908, *Pseudomonas resinovorans* and *P. fluorescens* SBW25, respectively (35). In addition, the *A. sufurivorans* RW2 ortholog of *yejK* is located within a gene cluster flanked by integrase/transposase encoding genes and genes involved in mobile genetic element (MGE) transmission, suggesting that YejK might play a role in modulating the DNA supercoiling or the transfer of the corresponding MGE during its life cycle (36).

Thus, while YejK harbors some unusual properties for a NAP, it also has features consistent with it being categorized as a NAP. However, few studies have been published on YejK, limiting our understanding of this protein (33,34). Cellular studies revealed that *yejK E. coli* knockout strains showed a growth defect and bioinformatic analyses revealed that YejK homologs are highly conserved among Gram-negative bacteria, which suggest an important cellular role for YejK (33). YejK shows no sequence homology to any structurally characterized protein and contains no identifiable DNA-binding motif. Hence, how it may interact with DNA is unknown. To gain insight into the structure of YejK and its mechanism of DNA binding we performed biochemical and X-ray crystallographic analyses on the *E. coli* YejK protein. These studies revealed that YejK harbors a newly described fold consisting of two domains. Each domain contains an highly extended region that homodimerizes with another YejK subunit to create a clamp-like structure with dimensions suitable for DNA binding. Our biochemical data indicate that basic residues within the pore mediate DNA binding by YejK. Finally, the data also suggest a clamp loading mechanism for YejK.

Materials and methods

E. coli YejK purification

The gene encoding *E. coli* YejK was purchased from Genscript Corporation and subcloned into pET15b using NdeI and BamHI restriction sites. The resultant expressed protein had an N-terminal His-tag for purification (Piscataway, NJ, USA;http://www.genscript.com). C41(DE3) cells were transformed with this expression vector. Cells for protein expression were grown at 37°C in Luria-Bertani (LB) medium with 0.10 mg/ml ampicillin to an OD_{600} of 0.3–0.4, then induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) at 15°C overnight. Cells were harvested by centrifugation and then resuspended in 100 ml of Buffer A (50 mM Tris–HCl pH 7.5, 800 mM NaCl, 5% (v/v) glycerol, 5 mM $MgCl_2$, 0.5 mM β-mercaptoethanol (βME)). Each 100 ml of resuspended cells were incubated for 30 min at 4°C (while stirring) with 200 μl protease inhibitor cocktail stock (which contained 100 μM aprotinin, 1 mM leupeptin, 1 mM pepstatin A) (Roche) and 40 U of a solution of 100 mg/ml DNase I (Roche). The resuspended cells were disrupted, first by sonication (4× at 80% amplitude) and then twice by microfluidizer. Cell debris were removed by centrifugation at 18 600 g at 4°C for 60 min. The supernatant was then loaded onto a cobalt NTA column (TALON Superflow histidine-tagged protein purification resin). The column was generated by pouring the resin into a Bio-Rad glass column and equilibrating before protein addition with Buffer A (with 900 mM NaCl) overnight. It was noted after several YejK protein purifications that variable amounts of DNA co-purified with YejK. Hence, a high salt concentration was used for purification as well as including low concentrations of urea (1.5 M) in the initial wash step that would perturb oligomers but not lead to unfolding. Hence, after protein application, the column was washed twice (500 ml each), during the first wash 1.5 M urea was added and during the next wash, 20 mM imidazole was included with Buffer A and no urea. Finally, the proteins were collected in Buffer A in an increased gradient of imidazole. Fractions were analyzed by SDS-PAGE and those containing the protein were combined. The fractions were concentrated using Sigma-Millipore concentrators (Amicon) and then subjected to a second purification step using size exclusion chromatography (SEC) on a SUPERDEX™ 200 pg Hiload™ 26/600 column and an AKTA prime plus. The buffer used for this step was 25 mM Tris–HCl pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 5 mM $MgCl_2$ and 0.5 mM βME. The His-tag was removed by thrombin digestion overnight at 37°C using a thrombin cleavage capture kit (EMD Millipore). The cleaved His-tag was removed by loading the cleavage reaction onto a Ni-NTA column and collecting the flow through. The tag-free protein was concentrated using centricons (Amicon) with a 30 kDa molecular weight (MW) cutoff.

Generation and purification of YejK(R173A-R176A-K177A), YejK(R276A-R280A-K284A), YejK(L245M-L313M-L326M), YejK(1–290) (YejK C-trunc) and YejK(Δ134–147) (YejK N-trunc)

The DNA-binding YejK mutant proteins, YejK(R173A-R176A-K177A) and YejK(R276A-R280A-K284A), were produced from the WT artificial *yejK* gene from Genscript that had been mutated via QuikChange. These mutants were made to test the importance of pore lining residues on DNA binding. YejK(1–290) (also called YejK C-trunc), YejK(Δ134–147) (termed YejK N-trunc) and YejK(L245M–L313M–L326M) were generated as artificial genes, codon optimized for *E. coli* expression, from Genscript and cloned into pET15b using NdeI and BamHI restriction sites. All the resulting expressed proteins harbored His-tags and were expressed and purified as per the WT protein.

Crystallization and structure determination of *E. coli* YejK

For crystallization, the His-tag-free *E. coli* YejK protein was concentrated to 40 mg/ml. Wizard screens I to IV were used for screening at room temperature (rt) by the hanging drop vapor diffusion method. Crystals were produced by mixing the protein 1:1 with a solution consisting of 0.1 M 2-[4-(2-hydroxyethyl)piperazine-1-yl]ethanesulfonic acid (HEPES) pH 7.5, 1.8 M Na/K phosphate. The same crystals were produced with the His-tagged protein. Crystals were cryopre-

served by dipping them in a solution consisting of the crystallization reagent supplemented with 20% (v/v) glycerol for ∼5 s before plunging into liquid nitrogen. X-ray intensity data were collected at ALS beamline 5.0.2 and processed with XDS (42). The data were noticeably anisotropic but diffraction was observed to ∼3.4 Å in the best direction. An *E. coli* YejK Alphafold model was used in attempts to solve the structure by molecular replacement (MR) but was not successful. We note that the Alphafold model ultimately harbored overall similarity to the final subunit structure (i.e. monomer), but there are several differences in loop regions and secondary structural orientations within each domain that likely prevented MR from being successful. The data are also relatively low resolution, likely impacting the success of MR. Because of these issues it was deemed critical to obtain unbiased phases. Hence, selenomethionine substituted YejK protein was produced (as per the WT) for use in multiple wavelength anomalous diffraction (MAD) experiments. Crystals were successfully produced of the selenomethionine substituted protein under the same conditions as the native protein and MAD data were collected at ALS beamline 5.0.2 and processed with XDS (42). The resulting electron density map was not of sufficient quality to construct a complete model, in particular, of the C-domain. Hence, a YejK mutant was obtained in which several leucines in the C-domain were substituted with methionines, YejK(L245M–L313M–L326M), to use in MAD. Crystals of YejK(L245M–L313M–L326M) grew under the same conditions as WT and MAD data were collected at ALS beamline 5.0.2 (Table 1). Phenix Autosol (43) was used to determine selenium sites, perform phasing and carry out density modification. The addition of the extra methionines enhanced the phasing power of the MAD data, leading to an improved electron density map. The inclusion of the extra methionines in the C-domain was also critical in model construction as the C-domain electron density was poor in the initial MAD electron density maps, impeding construction of this domain. The addition of extra C-domain methionines provided markers to permit tracing this domain with the proper amino acid register (44). After construction, the model was subjected to refinement in Phenix. And following multiple rounds of fitting and refitting in Coot and refinement in Phenix (43,44) the structure converged to R_{work}/R_{free} values of 26.9%/29.0% to 3.56 Å resolution (Table 2).

Table 1. Crystallographic data collection statistics: MAD data collection for *E. coli* Semet YejK(L245M-L313M-L326M)

	Wavelength 1	Wavelength 2	Wavelength 3
Data collection			
Space group	$P4_322$	$P4_322$	$P4_322$
Wavelength (λ)	0.97965	0.95690	0.97976
Cell dimensions			
a, *b*, *c* (Å)	84.4, 84.4, 113.0	84.4, 84.4, 113.0	84.4, 84.4, 112.9
α, β, γ (°)	90.0, 90.0, 90.0	60.7, 87.2, 68.4	90.0, 90.0, 90.0
Wavelength (λ)			
Resolution (Å)	67.7–3.60 (3.82–3.60)*	59.4–3.60 (3.82–3.60)	84.4–3.60 (3.82–3.60)
R_{sym}	0.072 (0.871)	0.071 (0.870)	0.070 (0.881)
R_{pim}	0.033 (0.544)	0.035 (0.599)	0.045 (0.488)
$I/\sigma I$	12.5 (1.4)	11.8 (1.2)	12.7 (1.9)
Completeness (%)	99.9 (99.8)	99.2 (99.3)	99.7 (99.8)
Redundancy	5.6 (5.7)	5.2 (5.3)	5.6 (5.7)
CC(1/2)	0.999 (0.665)	0.999 (0.506)	0.999 (0.667)

Table 2. Crystallographic data collection and refinement statistics: *E. coli* YejK

	***E. coli* YejK**
Data collection	
Space group	$P4_322$
Pdb code	9BE2
Cell dimensions	
a, *b*, *c* (Å)	85.1, 85.1, 113.8
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	42.6–3.56 (4.05–3.56)*
R_{sym}	0.081 (1.869)
R_{pim}	0.031 (0.698)
$I/\sigma I$	18.1 (2.1)
Completeness (%)	88.2 (56.0)
Redundancy	7.4 (7.5)
CC(1/2)	0.999 (0.886)
Refinement	
Resolution (Å)	36.1–3.56
No. reflections	4762 (478)
R_{work} / R_{free} (%)	26.9/29.0
No. atoms	
Protein	2648
Water	0
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.38
Ramachandran analyses	
Favored (%)	90.1
Disallowed (%)	0.0
Molprobity score	2.30

*Values in parentheses are for highest-resolution shell.

Fluorescence polarization (FP) binding experiments

To measure DNA binding to *E. coli* YejK, the following fluoresceinated DNA targets were used (top strand); AT-rich: 5′-ATTATTAATTAATTAATAAT-3′, GC-rich: 5′-CGCGCCGGCCGGCCGGCGCG-3′, mixed: 5′-TTGCGACAGAAACGGTCGTA-3′. In each experiment, increasing concentrations of YejK were titrated into the sample cell containing 1 nM of the DNA in a buffer of 25 mM Tris–HCl pH 7.5, 75 mM NaCl, 2.5% (v/v) glycerol. DNA binding analyses of mutant YejK, both DNA binding and N-trunc and C-trunc mutants, were performed as per WT with the AT-rich DNA site. Samples were excited at 490 nM and fluorescence emission was measured at 530 nM. The resultant data were plotted using KaleidaGraph and the curves fit with single site binding. The FP equation is: $P = P_f + (P_b - P_f)\,[M] / (K_d + [M])$, where P is the measured polarization, P_f is the polarization of free DNA, P_b is the polarization of the completely bound DNA, K_d is the equilibrium dissociation constant, and M is the concentration of protein. P_b and K_d were determined by nonlinear least-squares analysis. Three technical repeats were performed for each curve and the standard errors from the three affinities were determined.

Glutaraldehyde crosslinking

Glutaraldehyde crosslinking experiments were carried out on YejK proteins at 0.4 mg/ml in a buffer composed of 25 mM HEPES pH 7.5, 150 mM NaCl. For the reactions, glutaraldehyde was added to a final concentration of 0.1% and the reaction was run for 15 min. SDS running buffer was then added

1:1 and the samples immediately boiled for 2 min prior to running on SDS gels.

Size exclusion chromatography (SEC) experiments

SEC studies were performed using a SUPERDEX™ 200 pg Hiload™ 26/600 column and an AKTA prime plus. The buffer used for the runs was 25 mM HEPES pH 7.5, 150 mM NaCl, 5% (v/v) glycerol and 0.5 mM βME. Fractions were concentrated using Sigma-Millipore concentrators (Amicon) prior to column application. Samples were loaded using a 1 ml (final volume) syringe. The SEC studies were carried out on WT YejK and YejK C-trunc and N-trunc mutants. The elution volumes of each sample were compared with a series of protein standards to determine the molecular weights. The standards used for calculation of the standard curve were cytochrome c (12.4 kDa), carbonic anhydrase (29.0 kDa), bovine serum albumin (66.0 kDa), alcohol dehydrogenase (150.0 kDa), β-amylase (200 kDa) and Blue Dextran (2.0 mDa). The calibration curve was plotted using the gel-phase distribution coefficient (Kav) versus logarithm of the molecular weight (Log Mw). Kav = (Ve – Vo) / (Vc – Vo), where Ve = elution volume for the standard, Vo = column void volume = elution volume for Blue Dextran, Vc = total column volume.

Circular dichroism (CD) spectrometry

For CD, WT YejK, YejK C-trunc and YejK N-trunc were each diluted 50-fold to a final concentration of 0.25 mg/ml with CD Buffer (50 mM NaH_2PO_4, 200 mM NaF, 5% (v/v) glycerol, 1 mM βME, pH 7.5). Readings were done for 3 replicates for each corresponding sample, including the control (Buffer CD), using a Circular Dichroism Spectrometer from AVIV Biomedical Inc. Data were collected using the AVIV 435 CD software.

Mass photometry (MP) experiments

MP experiments were done using a Refeyn OneMP instrument (Oxford, UK). The experiments were performed using microscope coverslips, cleaned by sequential sonication in isopropanol (HPLC grade, and Milli-Q H_2O (15 min each)), followed by drying with a clean, pressurized air stream. Clean coverslips were assembled into the flow chamber, and silicone gaskets were positioned on the glass surface for sample loading to hold the sample drops with 4 × 4 wells prior to measurements. Contrast-to-mass calibration was achieved by measuring the contrast of proteins in the native marker protein standard mixture (NativeMark Protein Standard, Thermo Fisher Scientific) to generate a standard calibration curve at 100-fold dilution; 1048, 480, 146 and 66 kDa Molecular masses were used to fit the calibration curve with an R^2 value of 0.998 and a Max mass error of 11.3% in the Refeyn DiscoverMP v2023 R2 software. Calibration was applied to each sample measurement to calculate the molecular mass of each histogram distribution during analysis.

For these experiments, WT YejK was buffer exchanged in 25 mM Tris–HCl pH 7.5, 150 mM NaCl, and diluted in the working buffer of 20 mM Tris–HCl pH 7.5, 150 mM NaCl to a final concentration of 100 nM prior to sample analysis with 2-fold dilution on buffer droplet to a final concentration of 50 nM. To find focus, 10 μl of fresh buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl) adjusted to rt was pipetted into a well and the focal position was identified and locked using the autofocus function of the instrument. For data acquisition, 10 μl of diluted protein was added to the well and mixed, and movies of 60 s duration with 6000 frames (10 frame rate/Hz) were recorded using Refeyn AcquireMP 2023 R1 software using normal measurement mode with regular image acquisition settings. Three individual measurements/replicates were performed. All mass photometry movies of each measurement were processed and analyzed by Refeyn DiscoverMP v2023 R2 software, and Gaussian curves were fit to each histogram distribution, and the mass (kDa), sigma (kDa) and counts were determined.

Results

E. coli YejK binds DNA non-sequence specifically

To perform structure/function studies on *E. coli* YejK, we expressed and purified the protein to homogeneity (Materials and Methods). Previous EMSA studies revealed that YejK bound DNA sites with mixed sequences that were a minimum of 12 base pairs (bp) and displayed increased binding affinity to longer DNA sites (33). These assays revealed that *E. coli* YejK bound 18 bp, 38 bp and 200 bp DNA sites with K_ds of 5.9 μM, 224 nM and 70.1 nM, respectively (33). Many NAPs have shown a preference for interacting with AT-rich DNA sites (12). The DNA sites used in EMSA analyses with YejK (33) had equal CG/AT content. Hence, it was unclear if YejK displays any sequence selectivity for AT-rich or GC-rich DNA. To address this question, we obtained three 20-mer dsDNA sites for use in fluorescence polarization (FP) binding experiments, one with a random/mixed arrangement of nucleobases (top strand: 5′-TTGCGACAGAAACGGTCGTA-3′), one that was AT-rich (5′-ATTATTAATTAATTAATAAT-3′) and one that was GC-rich (5′- CGCGCCGGCCGGCCG-GCGCG -3′). YejK bound each of these sites with essentially the same affinity; the YejK binding affinities for the AT-rich 20-mer, GC-rich 20-mer and mixed 20-mer were 6.8 ± 0.5, 5.8 ± 0.4 and ± 6.7 0.9 μM, respectively (Figure 1A).

Crystal structure of *E. coli* YejK

How YejK proteins might function as DNA binding NAPs has been unknown as these proteins show no sequence homology to any other characterized DNA-binding protein and lack any discernable DNA-binding motif. Thus, to gain insight into YejK function we obtained crystals of the full-length *E. coli* YejK protein. To obtain phases, we used multiple wavelength anomalous diffraction (MAD) to solve the structure with crystals of Selenomethionine (SeMet)-substituted YejK (Materials and Methods; Figure 1B–D; Table 1). The crystallographic asymmetric unit (ASU) contain one YejK subunit (Figure 1B; Supplementary Figure S1A-B) and was refined to a final resolution of 3.56 Å (Table 2).

YejK proteins are conserved across the Gram-negative clade (Figure 2A). For the *E. coli* YejK, the homology is highest with the corresponding proteins in the Enterobacteria *Klebseilla*, *Shigella* and *Salmonella* with sequence identities of 92%, 97% and 96%, respectively. The identity remains high in the Enterobacterial *Yersinia* YejK proteins (75–78%). The YejK proteins from γ-proteobacteria *Shewanella, Vibrio, Haemophilus* and *Pseudomonas* show less homology to the *E. coli* protein (45%, 50%, 53% and 35% identities, respectively) but have similar enough sequences that allow them to be identified as

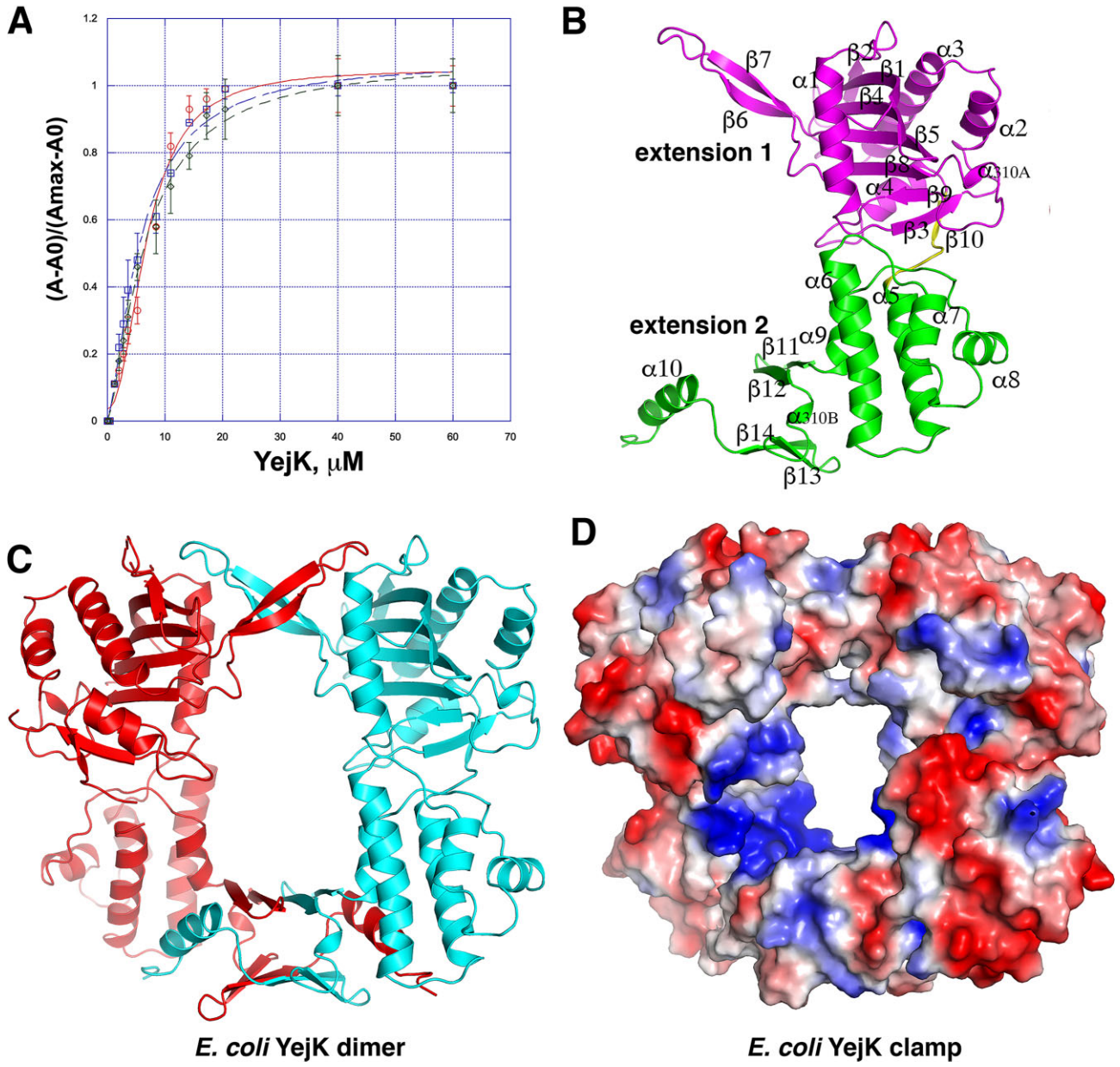


Figure 1. Structure of *E. coli* YejK and non-sequence specific DNA binding by YejK. (**A**) FP isotherms for WT YejK binding to fluoresceinated AT-rich 20 bp DNA (top strand, 5′-ATTATTAATTAATTAATAAT-3′) (red open circles), GC-rich 20 bp DNA (5′-CGCGCCGGCCGGCCGGCGCG-3′) (blue open squares), mixed 20 bp DNA (5′-TTGCGACAGAAACGGTCGTA-3′) (green open diamonds). The *x* and *y* axes are concentration of YejK in μM and normalized D millipolarization units (mP) $((A - A_0)/(A_{max} - A_0))$, respectively. *A* is change in mP reading, A_0 is the initial mP value before addition and A_{max} is the maximal mP reading upon binding saturation. Normalization was done to account for slightly different A_{max} values obtained for the different DNA sites. Data points represent mean values ± SD with the error bars centered at the mean. The error in K_d was determined as the SD between the calculated K_ds for three technical replicate runs. (**B**) Subunit structure of YejK, with the N-domain colored magenta and the C-domain colored green. Secondary structural elements are labeled as are the two extensions that emanate from each domain, and which mediate dimerization. (**C**) Structure of the YejK dimer. One subunit is colored cyan and shown in the same orientation as Figure 1B. the other subunit is colored red. (**D**) Electrostatic surface potential of the YejK dimer. Blue regions represent electropositive regions and red, electronegative regions. Figures 1B–D, 2B–C, 3A–B, 5A and 6A were generated in PyMOL (61).

YejK proteins (Figure 2A). Identical residues among all YejK proteins largely map to the hydrophobic core of in the *E. coli* YejK structure (Figure 2A–C). In addition, several basic residues are conserved as either lysine or arginine (see below). Interestingly, in the less well conserved YejK homologs, there are insertions in the sequences between β3 and α2 and between β1 and β2. Notably, the structure shows that these regions are in surface exposed loops, which would allow the inclusion of extra residues without impacting the overall fold (Figure 2A, B).

Each YejK subunit consists of two domains, an N-domain and C-domain; the N-domain (residues 1–187) is composed of a six stranded anti-parallel β-sheet, surrounded by four α-helices with topology β1-β2-α1-β3-α2-α3-β4-β5-β6-β7-β8-α_{310}-β9-α4-β10. A linker composed of residues 188 to 195 connects the N-domain to the C-domain. The C-domain

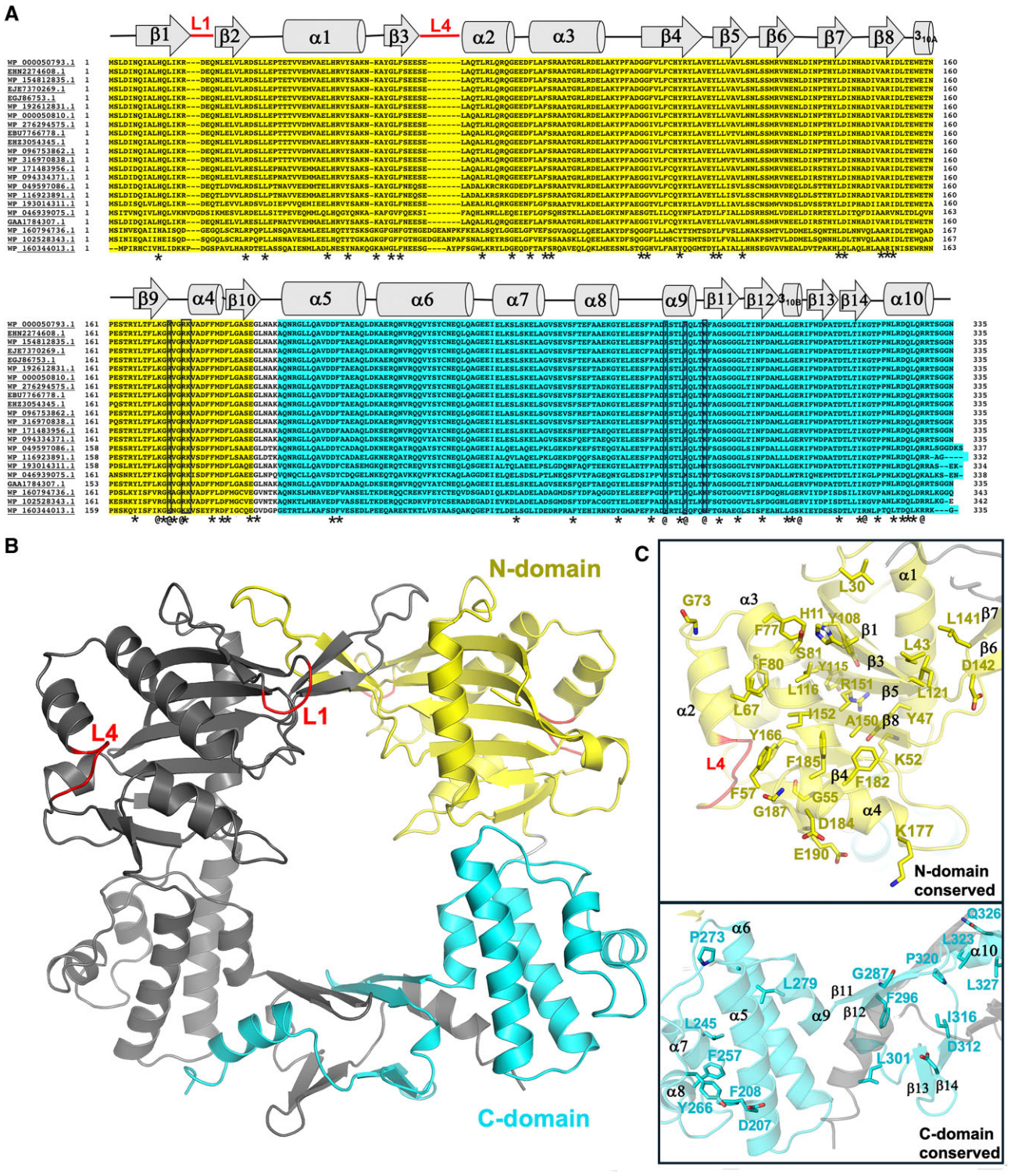


Figure 2. Multiple sequence alignment of YejK homologs. (**A**) Multiple sequence alignment of YejK proteins with the N-domain colored yellow and the C-domain, cyan. Secondary structural elements from the structure are indicated over the sequence and completely conserved residues are indicated by an asterisk under the alignment and residues conserved as K/R are indicated by @. Boxed are residues that were mutated to assess their roles in DNA binding. Two loop regions where insertions are present in some homologs are indicated by red, where L1 and L4 indicate Loop 1 and Loop 4. The YejK proteins from the following organisms are included (code; organism): WP_000050793.1; *Enterobacteriaceae*/100% identical to *E. coli*, EHN2274608.1; *Shigella sonnei*, WP_15412835.1; Enterobacteriaceae, EJE7370269.1; *Shigella dysenteriae*, EGJ86753.1; *Shigella flexneri* 2747–71, WP_192612831.1; *Citrobacter* sp. R56, EBU7766778.1; *Salmonella enterica* subsp. enterica serovar Rovaniemi, EHZ3054345.1; *Salmonella enterica* subsp. enterica, WP_096753862.1; *Citrobacter koseri*, WP_316970383.1; *Salmonella enterica*, WP_171483956.1; *Klebsiella pneumoniae*, WP_094334371.1; *Klebsiella quasipneumoniae*, WP_049597086.1; Yersinia, WP_116923891.1; *Yersinia pestis*, WP_193014311.1; Gammaproteobacteria, WP_046939075.1; *Haemophilus haemolyticus*, GAA1784307.1; *Nocardioides hankookensis*, WP_160794736.1; *Shewanella insulae*, WP_102528343.1; Shewanella, WP_160344013.1;*Pseudomonas xionganensis*. (**B**) YejK dimer with the N-domain colored yellow and the C-domain, cyan as in Figure 2A. Also colored red are L1 and L4, highlighting that these variable loops are surface exposed, hence inclusion of residues in these regions would be accommodated in these loops. (**C**) Close up of conserved residues mapped onto the N-domain and C-domain. The residues are shown as sticks and labeled.

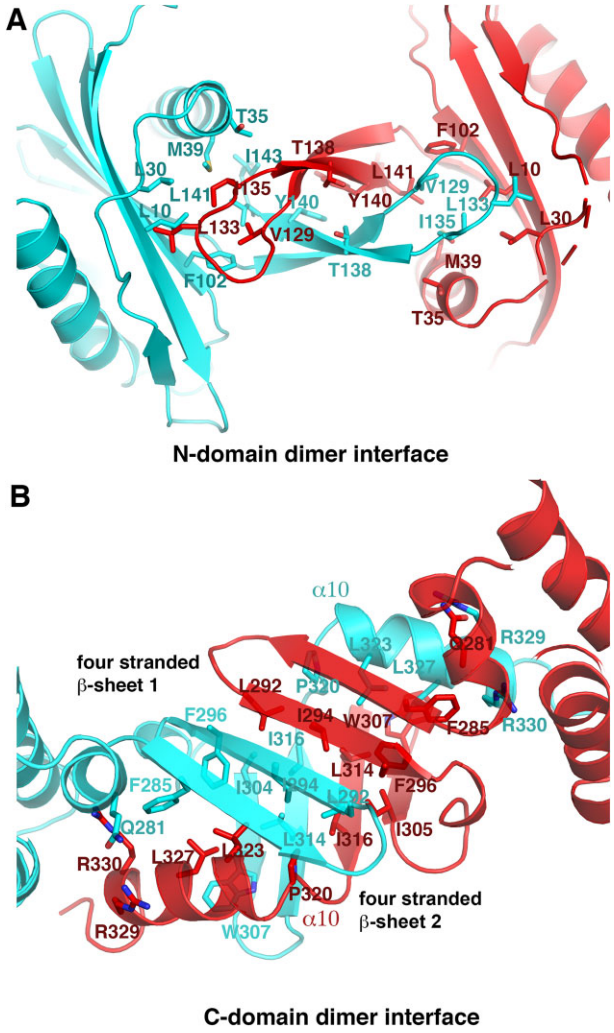


Figure 3. Two dimer interfaces in YejK. (**A**) N-domain dimerization interface. Shown as sticks and labeled are residues that are involved in dimer formation. (**B**) C-domain dimer interface. Residues involved in the YejK C-domain dimer interface are shown as sticks and labeled. Also indicated are the two stacked four stranded β-sheets that form a tight dimer interface in the C-domain, abutted by helices on each end.

(residues 196–335) has the topology α5-α6-α7-α8-α9-β11-β12-α_{310B}-β13-β14-α10 (Figure 1B; Figure 2A). A notable feature of both domains is the presence of a region that extends far from the main domain. In the C-domain, this extension consists of β11-β12 and β13-β14-α10, which projects out from the C-domain core by 48 Å. Similarly, the N-domain contains an elongated structure, a β-hairpin, comprised of β6-β7, that extends ∼40 Å from the core (Figure 1B). Strikingly, these extensions from the N- and C-domains combine with two-fold related extensions in an adjacent subunit in the crystal structure to create a ring-like dimer (Figure 1B, C; Supplementary Figure S1A-B).

In the YejK dimer, the curved β6-β7 hairpin of the N-domain interacts with the β6′-β7′ hairpin from the other subunit to form a β barrel like structure. Hydrophobic residues line the inside of this dimeric structure (Figure 2A). Tyr140 makes key contributions to the interface through stacking interactions (Figures 3A and 2C). Multiple sequence alignments show that this residue is strongly conserved as a tyrosine in most YejK proteins except *Pseudomonas* homologs, where the residue is a histidine, which could similarly mediate stacking interactions. Other hydrophobic resides in the interface are also highly conserved. For example, Leu133 is conserved and Ile135 is either a valine, leucine or isoleucine in YejK homologs.

In the C-domain, the two extended β hairpins combine to create two intertwined, four-stranded antiparallel β-sheets that pack tightly together. The C-terminal helices, α10 from each subunit, pack at the sides of the two β-sheets, further bolstering the intimate C-domain dimerization interface (Figures 1C and 3B). This dimer interface is highly hydrophobic with residues Phe285, Gly287, Leu292, Ile294, Phe296, Ile304, Trp307, Leu314, Ile316, Pro320, Leu323 and Leu327 from the interlaced β-sheets (Figures 2C and 3B). Of these residues, Gly287, Phe296, Ile316, Pro320, Leu323 and Leu327 are highly conserved among YejK homologs while the remaining dimer interface residues have conservative substitutions that would mediate the same interface. For example, Phe285 is either a phenylalanine or tyrosine, Leu292 is either a leucine or valine, Ile294 is isoleucine, valine or leucine and Leu314 is found to be either a leucine or isoleucine.

YejK forms dimers in solution

Our crystal structure revealed a YejK dimer. However, to investigate the oligomeric state of YejK in solution we employed several biochemical assays. First, size exclusion chromatography (SEC) experiments was performed at a YejK concentration of ∼50 μM. In these experiments YejK eluted at a position consistent with a dimer, with a calculated molecular weight (MW) of 61 kDa compared to the expected dimer MW of 74 kDa (Figure 4A, B). Because these experiments were performed at relatively high protein concentrations, we also carried out mass photometry (MP). These analyses, which were done at a YejK concentration of 50 nM, revealed that ∼90% of the protein was present as a dimeric species (calculated MW of 73 ± 1 kDa), indicating that YejK forms dimers at low protein concentrations (Figure 4C). Finally, we also employed glutaraldehyde crosslinking experiments (at a YejK concentration of 10 μM), which revealed that YejK was crosslinked to a dimer (Supplementary Figure S2). Collectively, these analyses support that YejK is a dimer in solution.

Structure similarity searches reveal YejK harbors a newly described fold

The YejK structure appears to possess a unique fold. However, to determine if there are other structures with homology to YejK we performed structural homology searches with the program SSM/pdbfold (45,46). These searches failed to identify structures with significant similarity to the overall fold of YejK. Hence, we next interrogated the Protein Data Bank for structures with similarity to just the YejK N- and C-domains (45,46). These searches revealed two structures with weak homology to the YejK N-domain. The first structure, *E. coli* RdgC (pdb code: 2OWL), superimposed on 97 Cα atoms of the N-domain of YejK with an rmsd of 3.2 Å (Supplementary Figure S3). Interestingly, RdgC was identified along with YejK as an *E. coli* NAP isolated from DNase I treated, spermidine isolated nucleoids (34). RdgC also forms a ring-like structure and functions in recombination and fork repair in *E. coli* (47,48). Though both YejK and RdgC form clamps, the ring structures in each are distinct. Indeed, the similarity between

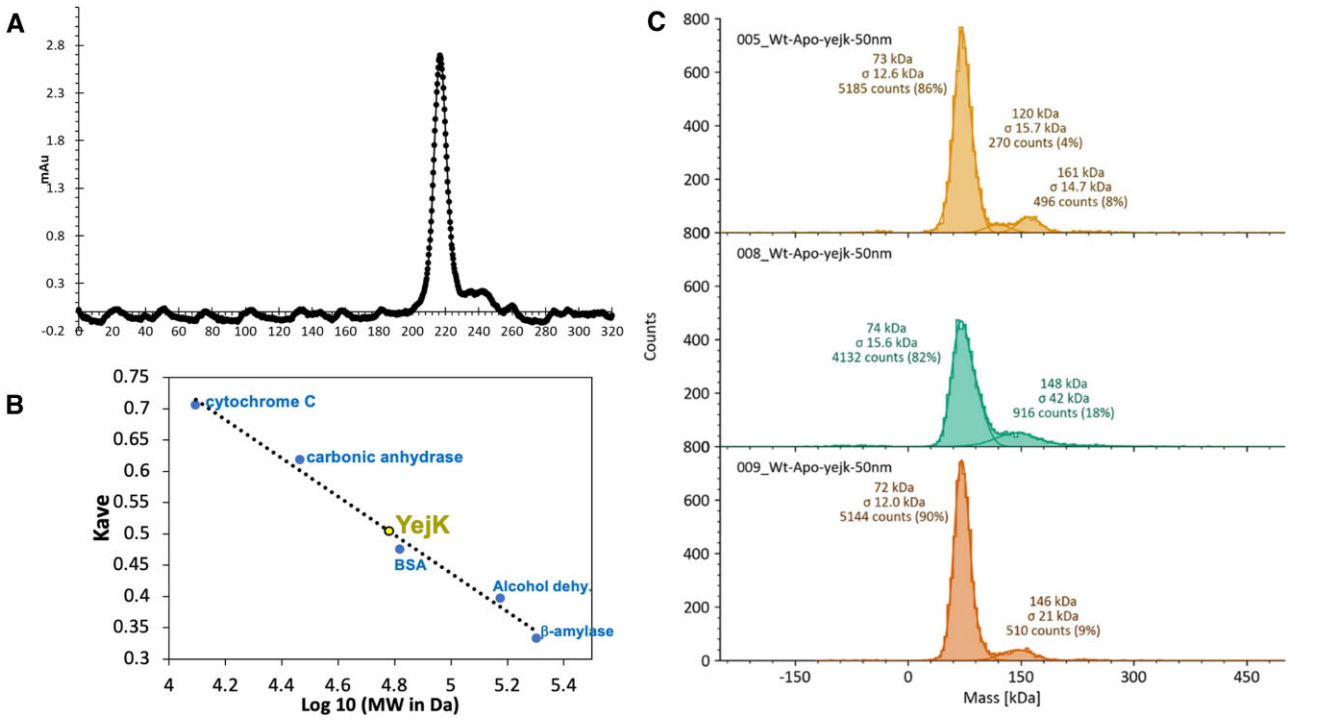


Figure 4. WT *E. coli* YejK forms dimers in solution. (**A**) WT YejK SEC profile from run on the SUPERDEX^{TM} 200 pg Hiload^{TM} 26/600 column and an AKTA prime plus. (**B**) SEC analyses used to calculate the MW of YejK from the SEC run in (A). The x and y axes are Log MW and elution parameter (Kav), respectively. Elution parameter Kav calculated by Kav = (elution volume for the standard – void volume)/(column volume – void volume). WT YejK eluted (yellow circle) at a calculated molecular weight (MW) of 61 kDa, consistent with a dimer. The standards used for calculation of the standard curve are shown (light_blue circles) and were cytochrome c (12.4 kDa), carbonic anhydrase (29.0 kDa), bovine serum albumin (BSA) (66.0 kDa), alcohol dehydrogenase (150.0 kDa), β-amylase (200 kDa) and Blue Dextran (2.0 mDa). (**C**) Triplicate analyses of WT *E. coli* YejK by mass photometry (MP) at a concentration of 50 nM. In the replicates, most of YejK was present as a dimer (MWs of 73, 74 and 72 kDa). These peaks were present at 86%, 82% and 90%. A small fraction was present as a tetrameric species or aggregates.

the YejK and RdgC structures is restricted to the N-domain β strand region of YejK (Supplementary Figure S3).

The second structure showing weak homology with the YejK N-domain was the Cenp-N protein from the yeast inner kinetochore CCAN-centromeric nucleosome complex (pdb code: 6QLE) (49) (Supplementary Figure S4). The Cenp-N in this complex superimposes onto the YejK N-domain with an rmsd of 3.4 Å for 127 similar Cα atoms. In this CCAN complex, Cenp-N was found to bind the unwrapped DNA termini that emanate from the Cenp-A containing nucleosome and in the superimposition with YejK, this DNA is positioned within the central region of the YejK dimer (see below). Finally, there were no structures that we could identify that showed even weak homology to the YejK C-domain in these searches.

Mutagenesis supports DNA-binding within the YejK clamp

Our data support that YejK forms a dimer. Strikingly, the central pore created by dimerization is unobstructed with a diameter of 30 Å, which is large enough to accommodate a dsDNA molecule (Figures 1C and 2B). Moreover, despite the protein's low overall pK_a of 4.4, the pore is lined with basic residues, thus generating an electropositive cavity (Figure 1D). By contrast, the outside surface of the pore is mostly electronegative. Also, as noted, in overlays, the YejK pore overlaps with a region in CenpN that is involved in DNA binding in the latter protein. Hence, we hypothesized that YejK might bind DNA within its central pore, thus acting as a DNA binding clamp.

To first assess whether dsDNA would fit within the pore, we performed docking experiments with dsDNA. The resultant model showed that, indeed, a dsDNA of ∼12 bps can be well fit within the YejK pore (Supplementary Figure S5). In fact, the model revealed little clash between YejK and the dsDNA, which could be alleviated by slight movements or shifts in either the DNA or protein (Figure 5A). Despite extensive co-crystallization attempts with dsDNA substrates of different lengths (from 12 to 38 bp in length) and nucleotide sequences, we were unable to obtain a YejK crystal form that contained well-ordered DNA. Thus, we reasoned that if the DNA binds within the positively charged YejK cavity, substitutions of conserved, basic residues with alanines would impair DNA binding in FP DNA binding assays.

Specifically, to ascertain if the N-domain or C-domain of the clamp, or both of these regions, are involved in DNA-binding, we made substitutions of basic residues that cluster within the pore, YejK(R173A-R176A-K177A) and YejK(R276A-R280A-K284A) (Figure 5A). These residues were also chosen as they showed strong conservation among YejK homologs as either lysine or arginine residues (Figure 2A). Both mutant proteins appeared to be folded; glutaraldehyde crosslinking analyses showed both proteins formed dimers under conditions where the WT protein dimerizes (Supplementary Figure S2). We then performed DNA binding assays with the 20 bp AT-rich site, which WT YejK bound with a K_d 6.8 ± 0.5 μM. In these assays, the YejK(R173A-R176A-K177A) showed weak, nonsaturable binding and the YejK(R276A-R280A-K284A) mutant bound with a K_d of 40 ± 5.6 μM (Figure

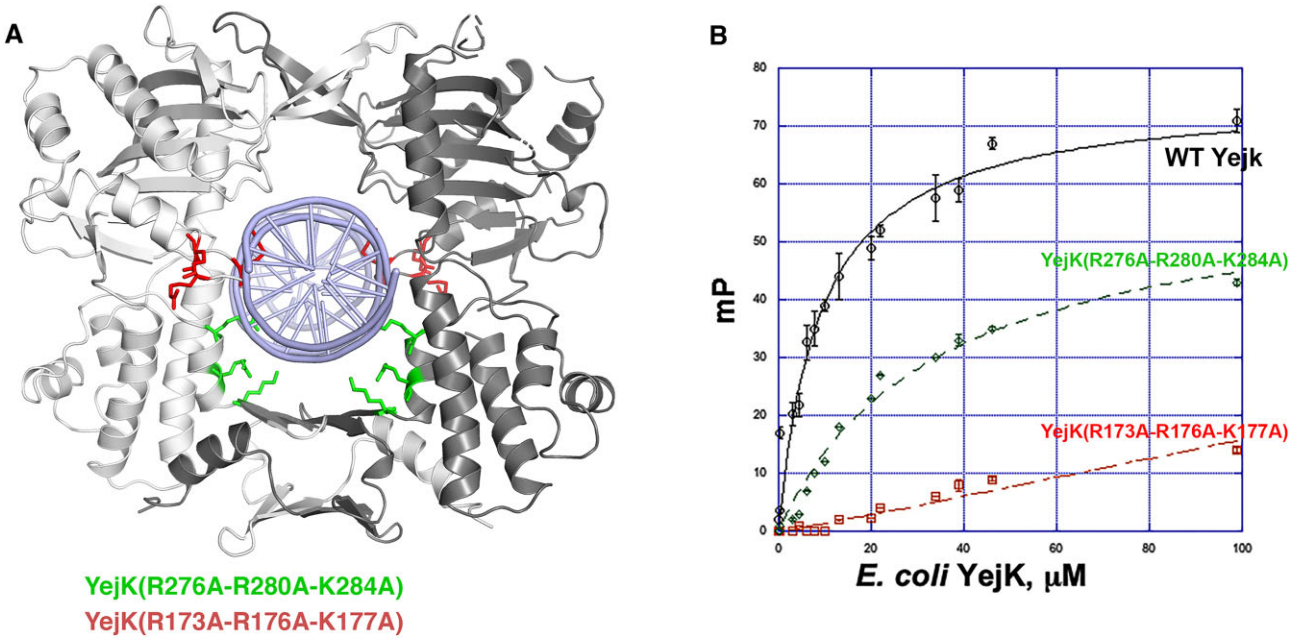


Figure 5. The YejK central pore mediates DNA binding. (**A**) Mapping of conserved basic (K/R) residues within the YejK central pore that were mutated to assess effects on DNA binding. Two mutant proteins were generated, one in which basic residues within the N-domain were substituted, YejK(R173A-R176A-K177K), and one in which basic residues in the C-domain were substituted, YejK(R276A-R280A-K284A). (**B**) FP binding analyses comparing binding of the WT YejK (black open circles) with the two mutants. The WT bound robustly while the mutants displayed no binding (YejK(R173A-R176A-K177K)) (red open boxes) or significantly reduced binding (YejK(R276A-R280A-K284A)) (green +).

5B). Thus, these data support that YejK can bind DNA within its central pore and hence functions as a DNA binding clamp and indicate that both the N- and C-domains provide contacts to the DNA.

YejK ring loading

Our crystal structure captured YejK in a dimeric, closed ring conformation. But DNA-binding clamps that encircle DNA must be cracked open to allow the DNA to load. Our SEC, MP and crosslinking data indicate that YejK is dimeric. However, the fully closed YejK pore is created from two separate dimerization regions, one between the two N-domains and the other between the C-domains. If one of the dimerization interfaces is stable, holding the dimer intact and the other interface is weak and can adopt an open state, DNA could enter into the pore prior to clamp closure. Such a scenario would provide a mechanism for YejK clamp loading. However, if both interfaces form stable dimers, a clamp loader or other means of clamp opening would be required to enable DNA docking.

To address the question of YejK clamp loading, we generated truncation mutants in which each of the dimerization arms were removed (Figure 6A). For the C-domain truncation mutant (C-trunc), residues 291–335, which form the C-terminal dimerization extension, were truncated from the construct. Importantly, this truncation did not remove residues shown to be involved in DNA binding, Arg276, Arg280 and Lys284. For the N-domain mutant (N-trunc), residues 135–146, which form the β6-β7 arm, were removed. The latter removal preserved enough residues in the arm to form a turn between β6 and β7 and also retained the N-domain residues that interact with DNA. The N-trunc and C-trunc proteins were purified to homogeneity and circular dichroism (CD) experiments showed that both were properly folded (Supplementary Figure S6).

To next assess if the mutants formed dimers, we performed SEC analyses. These studies showed that, like the WT, the N-trunc (at 2 mg/ml) eluted as a dimer with a calculated MW of 70.8 kDa, compared to the theoretical MW of 70 kDa. Thus, the N-trunc protein retains the dimeric state. By contrast, the C-trunc mutant protein (at 2 mg/ml) eluted as a monomer-dimer mixture, but was primarily monomeric with the main peak eluting at a calculated MW of 36 kDa (compared to a theoretical MW of 33 kDa for the monomer) and a smaller peak corresponding to a MW of 60 kDa (compared to the theoretical MW of 66 kDa) for the C-trunc dimer (Figure 6B). These analyses indicated that the N-domain can adopt an open form that could allow DNA entry, while the C-domain forms a stable dimer, which could serve as a foundation for initial binding. These data are consistent with what is known about physiological dimers, which typically have buried surface areas (BSA) of 1000 Å^2 per subunit, as the N-domain has a relatively small BSA of 570 Å^2 per subunit, while the C-domain interface is extensive, with 1578 Å^2 BSA per subunit.

To analyze if either of the truncation mutants can bind DNA, we next performed FP binding assays with our 20 bp AT-rich DNA site. While the C-trunc mutant showed essentially no binding to the DNA, the N-trunc mutant showed weak interaction, suggesting that the DNA could dock into the partially formed pore composed of an open N-domain stabilized by the dimerized C-domain (Figure 6C). However, the lack of interaction of the C-trunc mutant with DNA and weak binding observed for the N-trunc indicate that stable, saturable DNA binding to YejK requires the protein adopt a closed clamp conformation.

Discussion

The genetic code for all organisms is contained on a large DNA substrate the size of which exceeds the cell in which it is retained. Eukaryotes utilize histones to organize their genomes. Unlike eukaryotes, bacteria lack histones. Instead, studies have revealed that small, basic proteins called

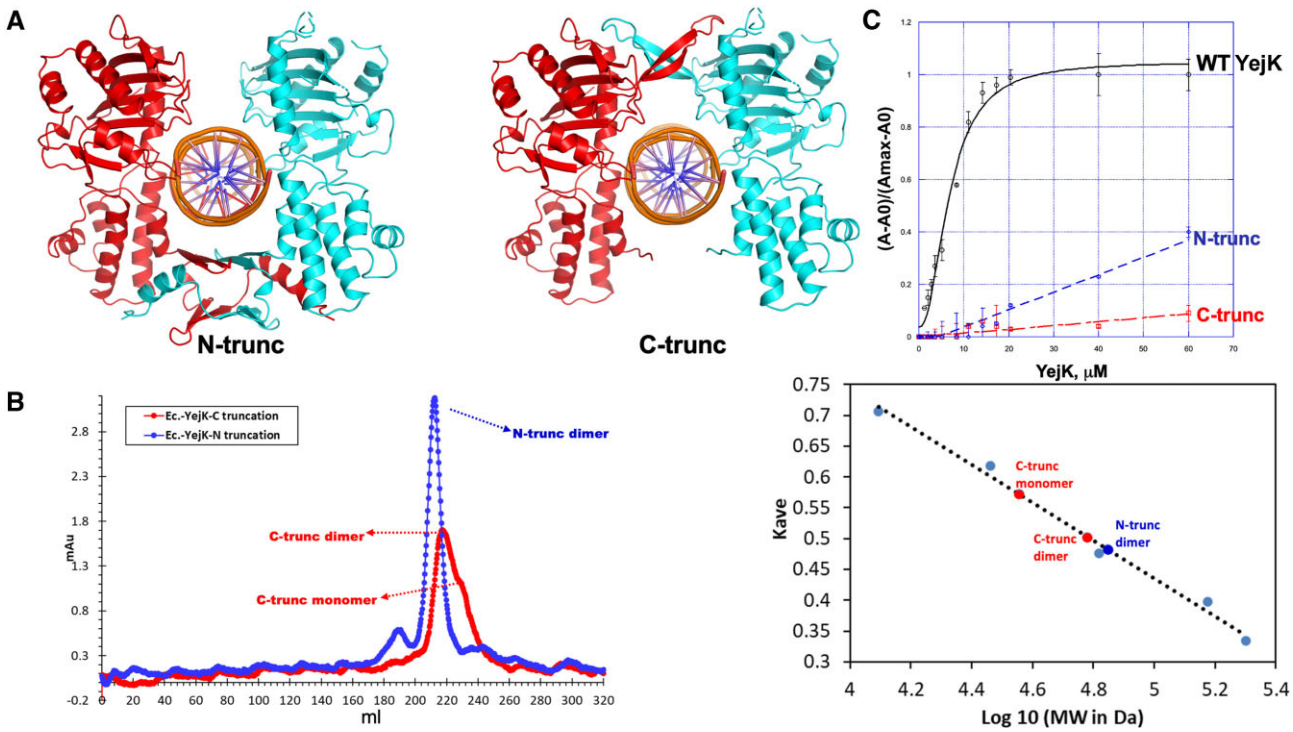


Figure 6. YejK contains an open interface to allow DNA docking into the pore. (**A**) Models of the two YejK truncation mutants made to test oligomer state and DNA binding. The N-trunc and C-trunc mutants removed the N-domain extensions and C-domain extensions, leaving only one dimer interface intact in each mutant. (**B**) SEC analyses of YejK trunc mutants. Left is the SEC profiles. Right is the plot used to calculate the MW of the mutants. The x and y axes are Log MW and elution parameter (Kav), respectively. The standards used for calculation of the standard curve were the same as Figure 4B. cytochrome c (12.4 kDa), carbonic anhydrase (29.0 kDa), bovine serum albumin (66.0 kDa), alcohol dehydrogenase (150.0 kDa), β-amylase (200 kDa) and Blue Dextran (2.0 mDa). The calibration curve was plotted using the gel-phase distribution coefficient (Kav) versus logarithm of the molecular weight (Log Mw). Kav = (Ve - Vo) / (Vc - Vo), where Ve = elution volume for the standard, Vo = column void volume = elution volume for Blue Dextran, Vc = total column volume. One peak is present for the N-trunc corresponding to a MW of 70 kDa (compared to the theoretical MW of 70.8 kDa). There are two overlapping peaks for the C-trunc with the major peak aligned at a calculated MW of 36 kDa and minor peak at 60 kDa. The computed MW of the C-trunc monomer is 33 kDa and the dimer is 66 kDa. Hence, most of the protein is present as a monomer. (**C**) FP binding assay examining the binding of WT, C-trunc and N-trunc YejK binding to the AT-rich 20mer (labled). The C-trunc mutant showed essentially no binding while the N-trunc showed very weak binding.

nucleoid-associated proteins (NAPs) are involved in bacterial DNA organization and dynamics. Interestingly, data indicate that bacteria encode multiple NAPs with surprisingly diverse structures (1–34). And while the *E. coli* NAPs are arguably the best understood, gaps in our knowledge remain as to how many of these proteins interact with and organize DNA at the molecular level. Arguably, one of the least understood *E. coli* NAP is the YejK protein. YejK was first identified after DNase I treatment of *E. coli* nucleoids and subsequent work in the Marians lab revealed that YejK is present at high abundance and is localized throughout the *E. coli* nucleoid, consistent with its role as a NAP (33,34). Yet how YejK binds DNA has been unknown.

Here, we carried out structure/function analyses on the *E. coli* YejK protein and show that it harbors a novel fold, consisting of two domains. Each domain contains an elongated extension that combines with the same extension from another YejK subunit to form an asymmetric dimer with two distinct dimer interfaces. Despite the fact that YejK is, overall, highly acidic, the formation of the dimer creates a toroidal structure with an internal surface that is electropositive. The basic nature of the pore and its diameter (30 Å), which would accommodate dsDNA, suggested that YejK may function as a DNA binding clamp. Consistent with this hypothesis, we showed that mutation of basic regions lining the center of the toroid significantly impaired DNA binding by YejK. Future crosslinking experiments will be explored to assess residues interacting with DNA. To date, numerous DNA binding clamps have been structurally characterized, including *D. radiodurans* RecR (50), *E. coli* RdgC (47,48) and *C. crescentus* GapR (32,51) as well as processivity factors, β subunit of bacterial DNA polymerase III (52), gp45 of T4 DNA polymerase (53), the eukaryotic 9–1–1 clamps and the proliferating cell nuclear antigen (PCNA) proteins (54–58). YejK is similar to these clamps in harboring a central positive region. However, its pore structure is comprised of β-strands, loops and helical regions, which differs from most characterized clamps, which have pores composed of α helices (Supplementary Figure S7). It also distinct from the RdgC clamp, which is generated by β-strands and loops. These data indicate that DNA-binding clamps can be formed from multiple structural elements (Supplementary Figure S7). They also underscore that DNA-binding clamps can be generated from diverse oligomeric states, including dimers, trimers and tetramers.

Consonant with YejK functioning as a DNA binding clamp, we showed that double stranded B-DNA can be docked into the YejK pore with almost no clash. This modeling though showed that the DNA would be tightly encased within the 30 Å YejK clamp, which would likely prevent YejK from freely

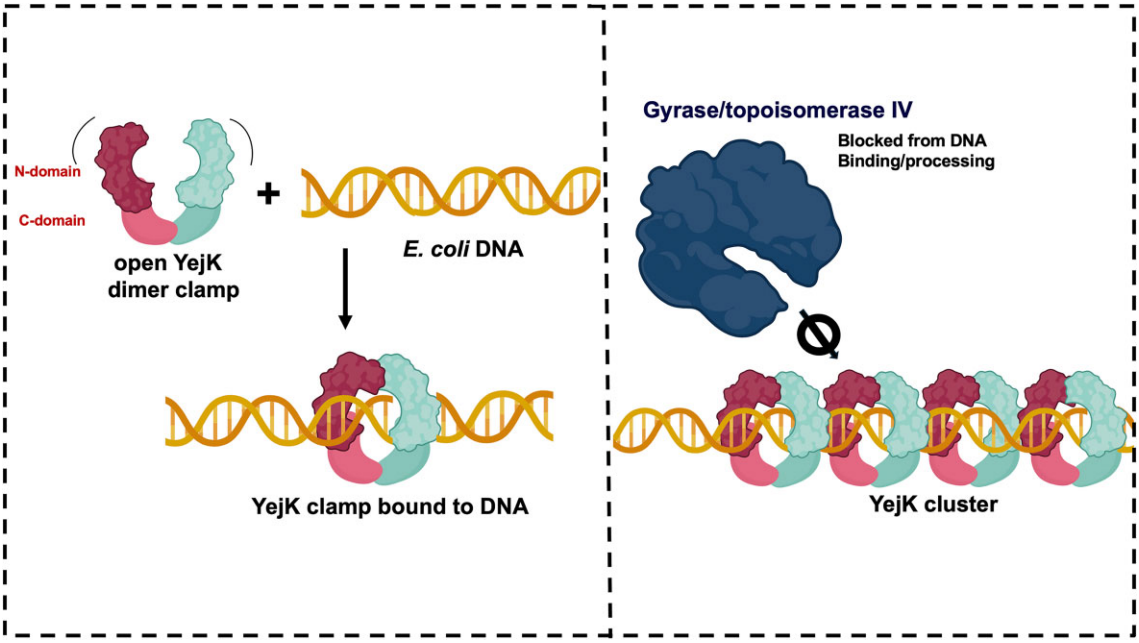


Figure 7. Model for YejK binding to *E. coli* chromosomal DNA and impact on topoisomerase activity. Left, cartoon depicting YejK as a dimer with two subunits, one red and the other blue. In the apo state, the N-domain region is open (not dimerized) allowing DNA to gain entry to the central pore. Upon DNA binding YejK encases the DNA and forms a clamp on the DNA. Right, multiple YejK dimer clamps binding the DNA may form clusters that impede topoisomerase binding and/or processive catalysis. This figure was created with BioRender.com.

sliding along DNA. This is distinct from the DNA processivity factors, which function as moveable clamps. These clamps, which have 35 Å pores, are thought to maintain a topological interaction with DNA, but not directly interact with it (53–58). Nonetheless, a key question regarding all DNA-binding clamps is how they are loaded onto DNA substrates that lack a free end. In fact, many of these proteins require AAA+ clamp loaders that use ATP to open and close them around their DNA target (59,60). The *C. crescentus* GapR is one clamp that does not appear to require a loader, as it adopts an open structure that closes upon DNA binding (32). We propose that YejK uses a somewhat similar mechanism whereby the YejK N-domain can form an opening to allow DNA entry, while the C-domain forms a stable dimer to act as the base (Figure 7). In this way, YejK can function as a self-assembling clamp.

Notably, YejK has also been shown to bind both topoisomerase IV and gyrase and effect their activities in complex manners (33). Specifically, in the presence of YejK, topoisomerase IV switches from a processive to a distributive mode and YejK binding to gyrase inhibits it from relaxing negatively supercoiled DNA (33). The structure suggests a possible mechanism to explain these effects as binding by the YejK clamp could function as a roadblock to interaction or processive movement of the topoisomerases. An interesting feature of YejK is that its affinity for DNA increases with increasing DNA length (33). This could be explained if multiple YejK dimers dock onto the DNA, possibly via cooperative interactions. The formation of such YejK clusters on the DNA could serve as an even more effective roadblock to topoisomerase activity (Figure 7). These effects on topoisomerase activity could also have large scale consequences on chromosome topology. However, future studies will be needed to determine the role of YejK in modulating topoisomerase/gyrase activity and possible impacts on DNA conformation. In conclusion, our studies have uncovered a novel *E. coli* NAP that forms a self-loading, DNA-binding clamp and underscores that much remains to be learned, even in *E. coli*, regarding NAP structure and function.

Data availability

The atomic coordinates and structure factors for the *E. coli* YejK structure have been deposited in the Protein Data Bank (https://www.rcsb.org) and are publicly available as of the data of publication. The accession code is: 9BE2. Any additional information required to reanalyze the data reported in the paper is available from the corresponding author upon request.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We acknowledge beamline 5.0.2 for X-ray diffraction data collection. The ALS (Berkeley, CA) is a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the US Department of Energy under Contract DE-AC02-05CH11231, Office of Basic Energy Sciences. Beamline 5.0.2 of the ALS, a US Department of Energy Office of Science User Facility under Contract DE-AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the NIH, National Institute of General Medical Sciences, Grant P30 GM124169-01. Mass photometry was performed at the UNC Macromolecular Crystallography core lab supported by the National Cancer Institute of the National Institutes of Health under award number P30CA016086.

Funding

Nanaline H. Duke Endowed Chair and National Institutes of Health grants [R35GM130290 to M.A.S.]. Funding for open access charge: NIH [R35GM130290].

Conflict of interest statement

Professor Maria A. Schumacher holds the position of Executive Editor for Nucleic Acids Research and has not peer reviewed or made any editorial decisions for this paper.

References

1. Luijsterburg,M.S., White,M.F., van Driel,R. and Dame,R.T. (2008) The major architects of chromatin. Architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 393–418.
2. Dillon,S.C. and Dorman,C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.*, **8**, 185–195.
3. Takeda,T., Yun,C.S., Shintani,M., Yamane,H. and Nojiri,H. (2011) Distribution of genes encoding nucleoid-associated protein homologs in plasmids. *Int. J. Evol. Biol.*, **2011**, 685015,
4. Dorman,C.J. (2013) Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. *Biochem. Soc. Trans.*, **41**, 542–547.
5. Grainger,D.C. (2016) Structure and function of bacterial H-NS protein. *Biochem. Sco. Trans.*, **44**, 1561–1569.
6. Qin,L., Erkelens,A.M., Ben Bdira,F. and Dame,R.T. (2019) The architects of bacterial DNA bridges: a structurally and functionally conserved family of proteins. *Open Biol*, **9**, 190223.
7. Dorman,C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat. Rev. Microbiol.*, **2**, 391–400.
8. Doetsch,M., Gstrein,T., Schroeder,R. and Furtig,B. (2010) Mechanisms of StpA-mediated RNA remodeling. *RNA Biol*, **7**, 735–743.
9. Grimwade,J.E. and Leonard,A.C. (2021) Blocking, bending, and binding: regulation of initiation of chromosome replication during the *Escherichia coli* cell cycle by transcriptional modulators that interact with origin DNA. *Front. Microbiol.*,**12**, 732270.
10. Browning,D.F., Grainger,D.C. and Busby,S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.*, **13**, 773–780.
11. Stojkova,P., Spidlova,P. and Stulik,J. (2019) Nucleoid-associated protein HU: a lilliputian in gene regulation of bacterial virulence. *Front. Cell Infect. Microbiol.*, **9**, 159.
12. Azam,T.A. and Ishihama,A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.*, **274**, 33105–33113.
13. Azam,T.A., Hiraga,S. and Ishihama,A. (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells*, **5**, 613–626.
14. Yokoyama,H. and Fujii,S. (2014) Structures and metal binding properties of *Helicobacter pylori* neutrophil-activating protein with a di-nuclear ferroxidase center. *Biomolecules*, **4**, 600–615.
15. Amemiya,H.M., Schroeder,J. and Freddolino,P.L. (2021) Nucleoid-associated proteins shape chromatin structure and transcription regulation across the bacterial kingdom. *Transcription*, **12**, 182–218.
16. Joyeux,M. (2021) Impact of self-association on the architectural properties of bacterial nucleoid proteins. *Biophys. J.*, **120**, 370–378.
17. Melekhov,V.V., Shvyreva,U.S., Timchenko,A.A., Tutukina,M.N., Preobrazhenskaya,E.V., Burkova,D.V., Artiukhov,V.G., Ozoline,O.N. and Antipov,S.S. (2015) Modes of *Escherichia coli* dps interaction with DNA as revealed by atomic force microscopy. *PLoS One*, **10**, e0126504.
18. Preito,A.I., Kahramanoglou,C., Ali,R.M., Fraser,G.M., Seshasayee,A.S. and Luscombe,N.M. (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli*. *Nucleic Acids Res.*, **40**, 3524–3537.
19. Macvanin,M. and Adhya,S. (2012) Architectural organization in *E. coli* nucleoid. *Biochim. Biophys. Acta*, **1819**, 830–835.
20. Keane,O.M. and Dorman,C.J. (2003) The *gyr* genes of *Salmonellaenterica* serovar *Typhimurium* are repressed by the factor for inversion stimulation, Fis. *Mol. Genet. Genomics*, **270**, 56–65.
21. Weinstein-Fischer,D. and Altuvia,S. (2007) Differential regulation of *Escherichia coli* topoisomerase I by Fis. *Mol. Microbiol.*, **63**, 1131–1144.
22. Menzel,R. and Gellert,M. (1983) Regulation of the genes for *E. coli* DNA gyrase. Homeostatic control of DNA supercoiling. *Cell*, **34**, 105–113.
23. McLeod,S.M., Aiyar,S.E., Gourse,R.L. and Johnson,R.C. (2002) The C-terminal domains of the RNA polymerase α subunits. Contact site with Fis and localization during co-activation with CRP at the *Escherichia coli* proP P2 promoter. *J. Mol. Biol.*, **316**, 517–529.
24. Grainger,D.C., Goldberg,M.D., Lee,D.J. and Busby,S.J. (2008) Selective repression by Fis and H-NS at the *Escherichia coli* dps promoter. *Mol. Microbiol.*, **68**, 1366–1377.
25. Kahramanoglou,C., Seshasayee,A.S.N., Prieto,A.I., Ibberson,D., Schmidt,S., Zimmermann,J., Benes,V., Fraser,G.M. and Luscombe,N.M. (2011) Direct and indirect effect of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res.*, **39**, 2073–2091.
26. Holowka,J. and Zakrzewska-Czerwinska,J. (2020) Nucleoid associated proteins: the small organizers that help to cope with stress. *Front. Microbiol.*, **11**, 005990.
27. Odermatt,N.T., Sala,C., Benjak,A. and Cole,S.T. (2018) Essential nucleoid associated protein mIHF (Rv1388) controls virulence and housekeeping genes in *Mycobacterium tuberculosis*. *Sci. Rep.*, **8**, 14214.
28. Gordon,B.R.G., Li,Y., Wang,L., Sintsova,A., van Bakel,H., Tian,S., Navarre,W.W., Xia,B. and Liu,J. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 5154–5159.
29. Riccardi,E., van Mastbergen,E.C., Navarre,W.W. and Vreede,J. (2019) Predicting the mechanism and rate of H-NS binding to AT-rich DNA. *PLoS Compt. Biol.*, **15**, e1006845,
30. Navarre,W.E.W., Porwollik,S., Wang,Y., McClelland,M., Rosen,H., Libby,S.J. and Fang,F.C. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. *Science*, **313**, 236–238.
31. Bouffartigues,E., Buckle,M., Badaut,C., Travers,A. and Rimsky,S. (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat. Struct. Mol. Biol.*, **14**, 441–448.
32. Guo,M.S., Haakonsen,D.L., Zeng,W., Schumacher,M.A. and Laub,M.T. (2018) A bacterial chromosome structuring proteins binds overtwisted DNA to stimulate type II topoisomerases and enable DNA replication. *Cell*, **175**, 583–597.
33. Lee,C. and Marians,K.J. (2013) Characterization of the nucleoid-associated protein *YejK*. *J. Biol. Chem.*, **288**, 31503–31516.
34. Murphy,L.D., Rosner,J.L., Zimmerman,S.B. and Esposito,D. (1999) Identification of two new proteins in spermidine nucleoids isolated from *Escherichia coli*. *J. Bacteriol.*, **181**, 3842–3844.
35. Billane,K., Harrison,E., Cameron,D. and Brockhurst,M.A. (2022) Why do plasmids manipulate the expression of bacterial phenotypes?. *Philos. Trans. R Soc. Load. B Biol. Sci.*, **377**, 20200461.

36. Beard,S., Moya-Beltrán,A., Siva-García,D., Valenzuela,C., Pérez-Acle,T., Loyola,A. and Quatrini,R. (2023) Pangenome-level analysis of nucleoid-associated proteins in the *Acidithiobacillia* class: insights into the functional roles in Mobile genetic elements biology. *Front. Microbiol.*, **14**, 1271138.
37. Flores-Rios,R., Quantrini,R. and Loyola,A. (2019) Endogenous and foreign nucleoid-associated proteins of bacteria: occurrence, interactions and effects on mobile genetic elements and host biology. *Comp. Struct. Biotech.*, **17**, 746–756.
38. Dorman,C.J. (2013) Genome architecture and global gene regulation in bacteria: making progress towards a unified model. *Nat. Rev. Microbiol.*, **11**, 349–355.
39. Doyle,M., Fookes,M., Ivens,A., Mangan,M.W. and Dorman,C.J. (2007) An H-NS-like stealth protein aides horizontal DNA transmission in bacteria. *Science*, **315**, 251–252.
40. Dorman,C.J. (2014) H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid*, **75**, 1–11.
41. Navarre,W.W. (2016) The impact of gene silencing on horizontal gene transfer and bacterial evolution. *Adv. Microb. Physiol.*, **69**, 157–186.
42. Kabsch,W. (2010) XDS. *Acta Cryst.*, **D66**, 125–132.
43. Liebschner,D., Afonine,P.V., Baker,M.L., Bunkoczi,G., Chen,V.B., Croll,T.I., Hintze,B., Hung,L.W., Jain,S., McCoy,A.J., *et al.* (2019) Macromolecular structure determination using X rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.*, **75**, 861–877.
44. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Cryst.*, **D66**, 486–501.
45. Krissinel,E. and Hendrick,K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst*, **D60**, 2256–2268.
46. Krissinel,E. and Hendrick,K. (2004) Common subgraph isomorphism detection by backtracking search. *Software: Pract. Exp.*, **34**, 591–607
47. Ha,J.Y., Kim,H.K., Kim,D.J., Kim,K.H., Oh,S.J., Lee,H.H., Yoon,H.J., Song,H.K. and Suh,S.W. (2007) The recombination-associated protein RdgC adopts a novel toroidal architecture for DNA binding. *Nucleic Acids Res.*, **35**, 2671–2681.
48. Briggs,G.S., McEwan,P.A., Yu,J., Moore,T., Emsley,J. and Lloyd,R.G. (2007) Ring structure of the *Escherichia coli* DNA-binding protein RdgC associated with recombination and replication fork repair. *J. Biol. Chem.*, **282**, 12353–12357.
49. Yan,K., Yang,J., Zhang,Z., McLaughlin,S.H., Chang,L., Fasci,D., Ehrenhofer-Murray,A.E., Heck,A.J.R. and Barford,D. (2019) Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. *Nature*, **574**, 278–282.
50. Lee,B.I., Kim,K.H., Park,S.J., Eom,S.H., Song,H.K. and Suh,S.W. (2004) Ring-shaped architecture of RecR: implications for its role in homologous recombinational DNA repair. *EMBO J.*, **23**, 2029–2038.
51. Guo,M.S., Kawamura,R., Littlehale,M.L., Marko,J.F. and Laub,M.T. (2021) High-resolution, genome-wide mapping of positive supercoiling in chromosomes. *eLife*, **10**, e67236.
52. Kong,X.P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.
53. Moarefi,I., Jeruzalmi,D., Turner,J., O'Donnell,M. and Kuriyan,J. (2000) Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. *J. Mol. Biol.*, **296**, 1215–1223.
54. Dieckman,L.M., Freudenthal,B.T. and Washington,M.T. (2012) PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell. Biochem.*, **62**, 281–299.
55. Gonzalez-Magana,A. and Blanco,F.J. (2020) Human PCNA structure, function and interactions. *Biomolecules*, **10**, 570.
56. Xu,M., Bai,L., Gong,Y., Xie,X., Hang,H. and Jiang,T. (2009) Structural and functional implications of the human rad9-hus1-rad1 cell cycle checkpoint complex. *J. Biol. Chem.*, **284**, 20457–20461.
57. Doré,A.S., Kilkenny,M.L., Rzechorzek,N.J. and Pearl,L.H. (2009) Crystal structure of the rad9-rad1-hus1 DNA damage checkpoint complex—Implications for clamp loading and regulation. *Mol. Cell*, **34**, 735–745.
58. De March,M., Merino,N., Barrera-Vilarmau,S., Crehuet,R., Onesti,S., Blanco,F.J. and De Blasio,A. (2017) Structural basis of human PCNA sliding on DNA. *Nat. Commun.*, **8**, 13935.
59. Yao,N.Y. and O'Donnell,M. (2012) The RFC clamp loader: structure and function. *Subcell. Biochem.*, **62**, 259–279.
60. McNally,R., Bowman,G.D., Goedken,E.R., O'Donnell,M. and Kuriyan,J. (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct. Biol.*, **10**, 3.
61. Schrödinger (2023) *The PyMOL Molecular Graphics System, Version 2.55*. Schrödinger, LLC.

Received: April 22, 2024. **Revised:** May 2, 2024. **Editorial Decision:** May 3, 2024. **Accepted:** May 15, 2024