

UNIVERSIDADE ESTADUAL DE CAMPINAS  
SISTEMA DE BIBLIOTECAS DA UNICAMP  
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELLECTUAL DA UNICAMP

**Versão do arquivo anexado / Version of attached file:**

Versão do Editor / Published Version

**Mais informações no site da editora / Further information on publisher's website:**

<https://www.mdpi.com/2076-2607/10/10/1945>

**DOI: 10.3390/microorganisms10101945**

**Direitos autorais / Publisher's copyright statement:**

©2022 by MDPI. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>



## Review

# EbfC/YbaB: A Widely Distributed Nucleoid-Associated Protein in Prokaryotes

Tamires Fernanda Vilas Boas Cordeiro \* , Marco Túlio Pardini Gontijo , Genesey Perez Jorge and Marcelo Brocchi \*

Department of Genetics, Evolution, Microbiology and Immunology, Institute of Biology, University of Campinas (UNICAMP), Rua Monteiro Lobato 255, Campinas 13083-862, Brazil

\* Correspondence: tamiresfvbc@gmail.com (T.F.V.B.C.); mbrocchi@unicamp.br (M.B.)

**Abstract:** Genomic compaction is an essential characteristic of living organisms. Nucleoid-associated proteins (NAPs) are a group of small proteins that play crucial roles in chromosome architecture and affect DNA replication, transcription, and recombination by imposing topological alterations in genomic DNA, thereby modulating global gene expression. EbfC/YbaB was first described as a DNA-binding protein of *Borrelia burgdorferi* that regulates the expression of surface lipoproteins with roles in virulence. Further studies indicated that this protein binds specifically and non-specifically to DNA and colocalises with nucleoids in this bacterium. The data showed that this protein binds to DNA as a homodimer, although it can form other organised structures. Crystallography analysis indicated that the protein possesses domains responsible for protein–protein interactions and forms a “tweezer” structure probably involved in DNA binding. Moreover, sequence analysis revealed conserved motifs that may be associated with dimerisation. Structural analysis also showed that the tridimensional structure of EbfC/YbaB is highly conserved within the bacterial domain. The DNA-binding activity was observed in different bacterial species, suggesting that this protein can protect DNA during stress conditions. These findings indicate that EbfC/YbaB is a broadly distributed NAP. Here, we present a review of the existing data on this NAP.



**Citation:** Cordeiro, T.F.V.B.; Gontijo, M.T.P.; Jorge, G.P.; Brocchi, M. EbfC/YbaB: A Widely Distributed Nucleoid-Associated Protein in Prokaryotes. *Microorganisms* **2022**, *10*, 1945. <https://doi.org/10.3390/microorganisms10101945>

Academic Editor: Ricardo Amils

Received: 17 August 2022

Accepted: 20 September 2022

Published: 30 September 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** nucleoid-associated protein; histone-like protein; DNA-binding; chromosome organisation; transcription; binding domains

## 1. Introduction

The genome of a living organism must be compacted in an organised process to fit into the cell. Therefore, genomic compaction is an essential characteristic for all organisms. In prokaryotes, the chromosome is compacted by a combination of factors such as molecular crowding, DNA supercoiling, DNA transactions, and association with different DNA-binding proteins [1]. Nucleoid-associated proteins (NAPs), a group of low molecular weight proteins, are among the primary proteins that play an essential role in chromosome architecture [2]. NAPs can bind to DNA as monomers, dimers, or other organised structures [2]. H-NS, IHF, HU, Fis, and Dps are well-characterised NAPs of *Escherichia coli* and other species [3].

NAPs show variations in expression level and the DNA sequences they bind to, and some exhibit preferences for sequences and/or DNA topological features located along the bacterial genome. For instance, in *Salmonella enterica*, Fis is characteristically expressed during exponential growth [4,5], whereas IHF is expressed during the transition from the exponential to the stationary phase [6]. In addition, some NAPs possess bridging activities that stabilise DNA hairpins commonly observed during the exponential growth phase, whereas others wrap or bend DNA [2]. For instance, H-NS can form H-NS–DNA filaments or DNA–H-NS–DNA bridges [1]. Furthermore, the DNA-binding properties of NAPs can be modulated by factors such as protein–protein or ion interactions, which may affect the type of interaction of NAPs with DNA.

In addition to their topological effects on DNA, NAPs modulate many biological functions [2,7,8]. For example, H-NS plays an important architectural role in genome structuring and is also a transcriptional silencer for many genes. This protein binds AT-rich sequences and inhibits the expression of genes acquired by lateral gene transfer (LGT). This activity of H-NS is considered a determinant of genomic stabilisation of exogenous DNA acquired by LGT [9–11]. The inhibitory effect of H-NS can be disrupted by counter-silencing factors, thereby releasing many genes from the inhibitory effects of H-NS [12]. In some pathogens, this is critical for expressing virulence genes [13]. Notably, some NAPs act as classical transcription factors; they can be inhibitors or transcriptional activators. For instance, NAPs may block promoter–RNA polymerase (RNAPol) binding and/or RNAPol progression in transcription bubbles by bending the DNA, thereby negatively affecting transcription.

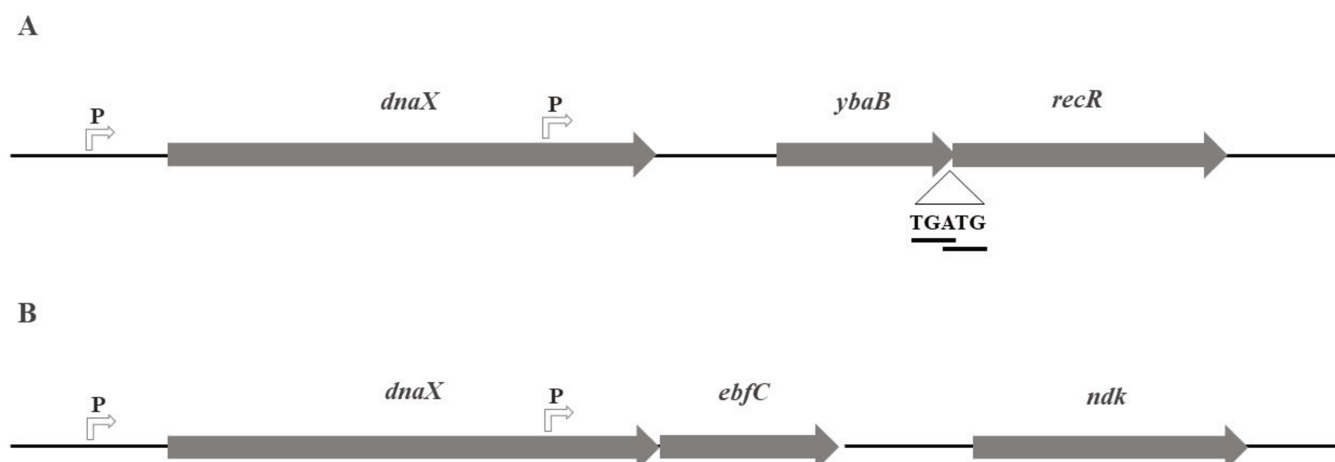
In contrast, NAPs can allow contact between enhancer and promoter sequences by bending the DNA, thereby promoting transcription. Direct contact between NAPs and RNAPol has also been reported, where the former guides the latter in the transcription process [7]. However, there have been instances where differentiation between transcription factors and NAPs was not clear [14]. NAPs are also involved in other DNA transactions such as initiation of DNA replication, recombination, site-directed recombination, transposition, DNA protection, and supercoiling [1–3,15]. Interestingly, a link exists between some NAPs, such as Fis, and DNA supercoiling [8]. All these events affect gene expression globally [3,8,16], and many studies with model organisms demonstrated that NAPs have essential functions in regulating different physiological and pathogenic traits. Here, we have cited a few of them [2,6,17–26].

EbfC (*erp-binding factor, chromosomal*), also known as YbaB, is a less-studied and widely distributed NAP among prokaryotes [27]. EbfC was demonstrated to have significant roles in the pathogenesis of *Borrelia burgdorferi* sensu lato (referred here as *B. burgdorferi*), which causes Lyme disease [28–32]. The protein structures of EbfC/YbaB of *Haemophilus influenzae* [33], *E. coli* (<http://www.rcsb.org/structure/1PUG>, accessed on 10 June 2022), and other bacterial species have been determined. Structural analyses did not reveal a characteristic DNA-binding domain; however, experimental studies have demonstrated that these proteins bind to DNA with distinct sequence preferences [30,34]. YbaB has been suggested to have a role in DNA repair in *E. coli*, but this does not seem to be the case for *Streptomyces* [35]. Preliminary results from our laboratory indicated that YbaB of *Salmonella enterica* is not essential for pathogenesis in a mouse model of systemic infection (unpublished results).

Despite its broad distribution, the biological role of EbfC/YbaB in many prokaryotic groups is not yet understood. Therefore, considering the critical role of NAPs in prokaryotes, this review aims to describe what is known regarding EbfC/YbaB.

## 2. Operon Structure and Expression

In *E. coli*, *ybaB* is located between the structural sequences of *dnaX* and *htpG* upstream of the *recR* gene. In this configuration, *ybaB* and *recR* are probably transcribed as a transcriptional unit, forming an operon [36,37]. Sequence analyses indicated the presence of an internal promoter in *dnaX*, which was shown to be a functional promoter when fused to a gene reporter [36]. Another characteristic of the *ybaB-recR* operon in *E. coli* and other bacteria is that the open reading frames of the two genes overlap, with the last “A” of the stop codon of *ybaB* corresponding to the first “A” of the start codon of *recR*. This indicates that the translation of YbaB and RecR is coupled [36,37]. Indeed, sequence analysis revealed a putative ribosome binding site (RBS) upstream of the start codon of *ybaB*, and no putative RBS was found close to the ATG of *recR*. Analyses of codon usage in *E. coli* indicated that *recR* has rarer codons than *ybaB*, which could reflect the differences in translation efficiency [36]. The operon structure of *ybaB-recR* in *E. coli* is shown in Figure 1A.



**Figure 1.** Operon structure of *dnaX-ybaB-recR* in *Escherichia coli* and that of *dnaX-ebfC* in *Borrelia burgdorferi*. (A) Operon structure of *dnaX-ybaB-recR* observed in *E. coli*. The initiation codon of *ybaB* and the termination codon of *recR* show overlap. (B) *dnaX-ebfC* operon of *B. burgdorferi*. The two promoters, one upstream of *dnaX* and the other located in the *dnaX* structural gene, are shown. Arrows indicate the promoter (P) and direction of transcription. The schematic structure of the operons is not to scale.

The operon structure of *ebfC* in *B. burgdorferi*, a bacterium in which the biological function of *ebfC* is best characterised, has also been described. In this bacterium, the operon *dnaX-ebfC* is regulated by a promoter located upstream of *dnaX* [30]. In addition, *ebfC* can also be expressed by a promoter located in the *dnaX* structural sequence, as determined by RT-PCR and GFP-transcriptional fusions [30,38]. In addition, DnaA, one of the main proteins of chromosomal replication, has been suggested to be involved in the regulation of expression of the *dnaX-ebfC* operon in *B. burgdorferi*, as discussed by Stevenson et al. [38]. The operon structure of *dnaX-ebfC* in *B. burgdorferi* is shown in Figure 1B.

RecR is a component of the RecFOR recombination system found in different bacterial species and is involved in DNA recombination and repair [39]. *dnaX* encodes subunits of DNA polymerase (DNAPol) [40]. The organisation of *dnaX*, *ybaB*, and *recR* in the chromosome of *E. coli* has also been observed in other bacteria. The colocalisation of *ybaB* with *dnaX* and *recR* suggests a possible role for *ybaB* in DNA recombination, repair, and/or replication. The involvement of NAPs in the modulation of DNA transactions corroborates this suggestion. However, synteny among *dnaX-ebfC/ybaB-recR* has not been observed in many other bacterial species, which contradicts this hypothesis.

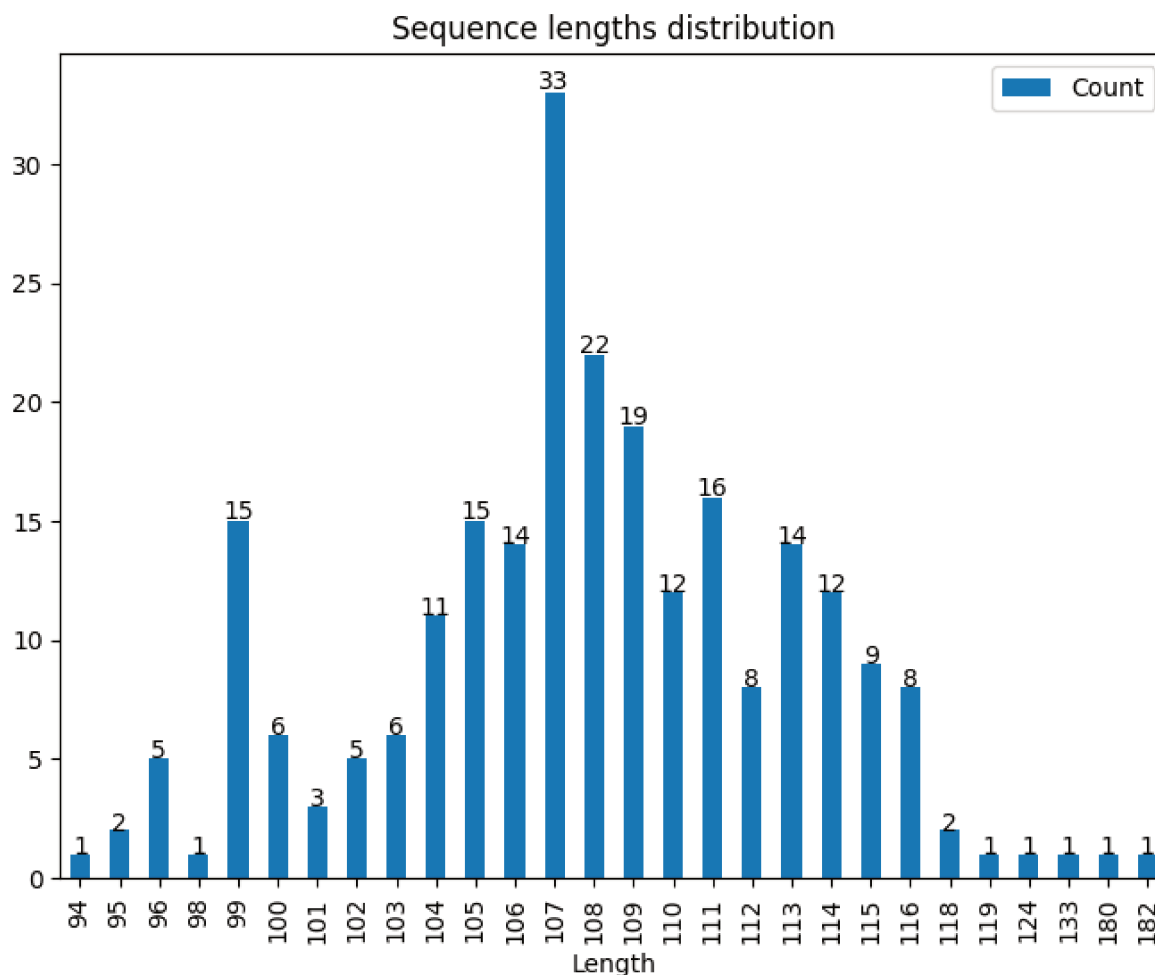
Using a two-plasmid system to identify promoters recognised by RNA polymerase containing the sigma E factor, *ybaB* belonged to the sigma E ( $\sigma$ E) regulon in *E. coli* [41]. The  $\sigma$ E regulon is crucial for responding to stress conditions, which could perturb cell envelope integrity [42]. Interestingly, it was demonstrated that YbaB potentiates the expression of heterologous membrane proteins in *E. coli* by an unknown mechanism [43]. This observation reinforces the possible involvement of YbaB in the envelope stress response; however, further studies are necessary to investigate their role. The biological functions assigned to *ebfC/ybaB* are discussed further in this paper.

### 3. EbfC/YbaB Orthologs Are Widely Distributed in Prokaryotes

In a data bank composed of 790 non-redundant, phylogenetically distinct bacterial genomes, including 704 genomes of bacteria and 86 of archaea, the most frequently found NAPs were HU, IHF, and EbfC/YbaB [27]. While this text is being written, based on the InterPro protein families and domains database [44], 585 putative reviewed sequences belong to the YbaB/EbfC family (IPR004401), and putative YbaB/EbfC was found in approximately 370 bacterial species (<https://www.ebi.ac.uk/interpro/entry/InterPro/IPR004401/protein/UniProt/#table>, accessed on 10 June 2022). The majority (99.19%) of

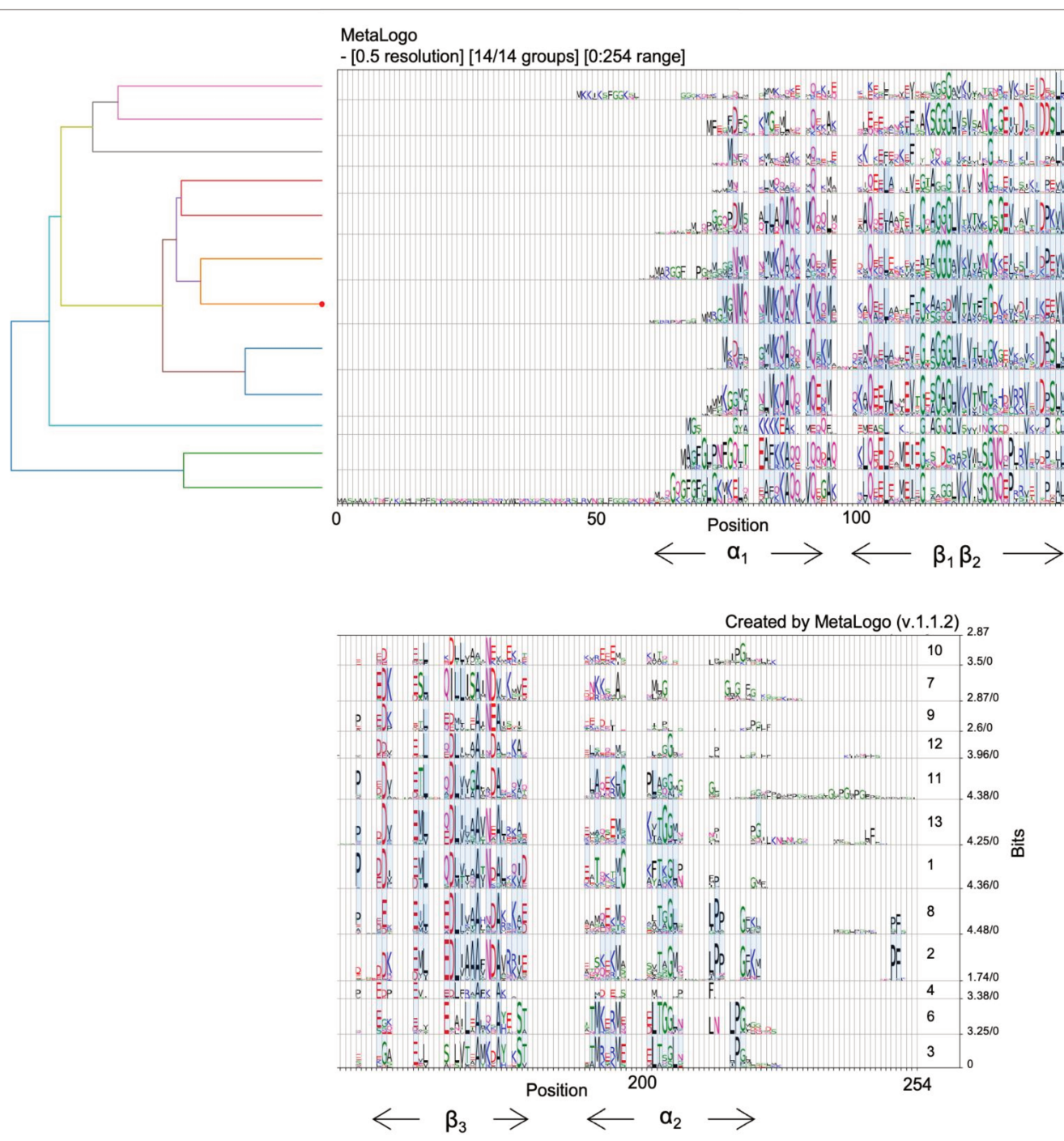


putative YbaB/EbfC proteins contain 94–133 amino acids. The largest sequences comprising 180 and 182 amino acids also belonged to the YbaB/EbfC family (Figure 2). However, they are found in the genome of *Arabidopsis thaliana*, outside the bacterial domain. This might indicate that YbaB/EbfC orthologs found in bacteria might have been present in the last common ancestor of *Eukarya*, *Archaea*, and *Bacteria* or were acquired by LGT.



**Figure 2.** Sequence length distribution results from MetaLogo [45]. The sequences were grouped with a sequence identity cut-off of 90% using CD-HIT Suite [46], yielding 248 representative proteins. Check Supplementary Table S1 to view the accession numbers of the sequences.

Figure 3 shows the multi-group sequence logo of all 248 representative orthologs identified in the bacterial domain. The sequences were divided into 12 groups according to their evolutionary relationships, with a resolution of 0.5. Two other groups were also predicted but are not present in the cladogram due to the low number of representatives. The light blue bands highlight conserved motifs shared by adjacent groups. All regions of *ybaB* seem to be relatively conserved across bacteria (details in Section 4). However, the most conserved patterns were found to be in the  $\beta 1/\beta 2/\beta 3$  and  $\alpha 2$  regions (also highlighted in Supplementary Figure S2B). The  $\alpha 1$  region was also conserved in some groups, but this region appears to be less conserved than the other two. Interestingly, the  $\alpha 2$ -helix is believed to be the region directly interacting with DNA, at least in *Caulobacter crescentus* [47]. Notably, the  $\beta 1/\beta 2/\beta 3$  region is composed of amino acids with either hydrophobic side chains or small side chains. These properties may facilitate the formation of homodimers.



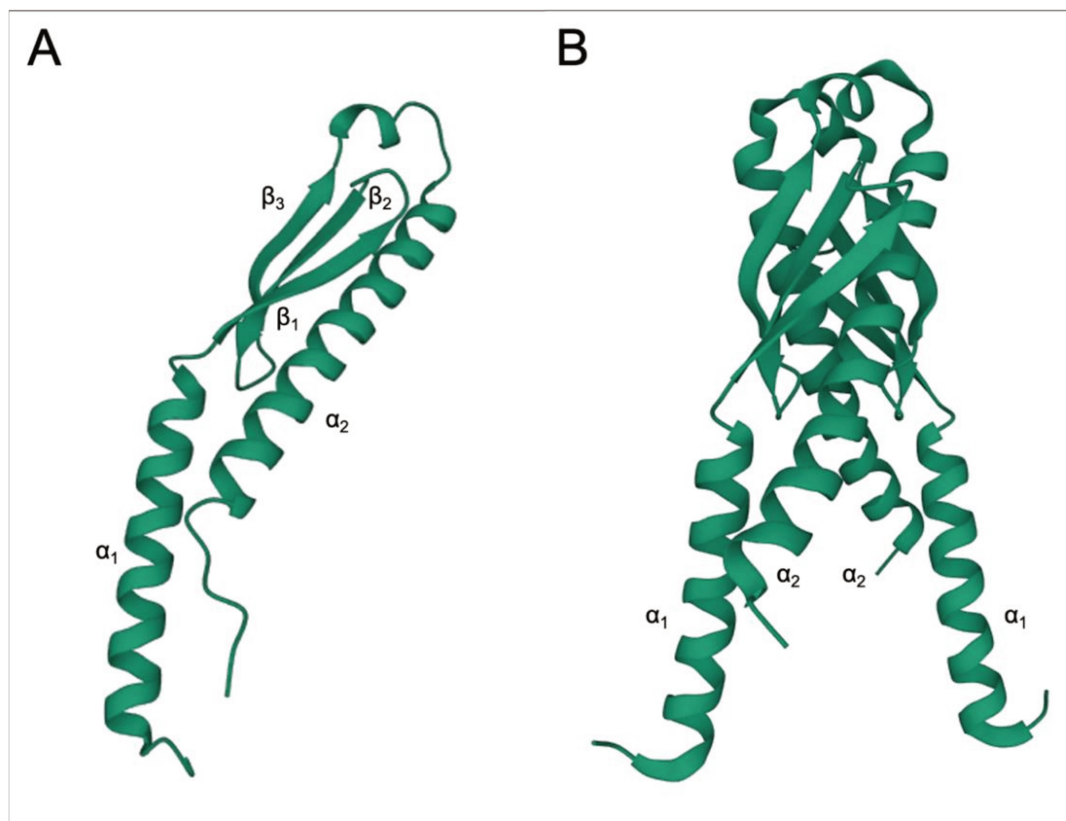
**Figure 3.** Sequence logos for all the reviewed putative proteins belonging to the YbaB/Ebfc family (IPR004401). The putative sequences are deposited in the InterPro protein families and domains database [44]. The database contains 585 reviewed sequences. Sequence logos were constructed using MetaLogo [45]. The tree on the left indicates the relationships among the groups. The red dot on the tree shows the group containing the target sequence. Light blue coloured strips connect conserved positions among groups. Check Supplementary Figure S1 for a complete view of Figure 3.

Supplementary Figure S2A shows the sequence counts for each group. Groups 1, 2, 8, 11, and 13 accounted for more than 70% of the total sequences. Figure S2B shows the entropy values for each position in each group. The higher the entropy (yellow), the less convergent the position. Figure S2C shows boxplots of the entropies for each group. Group 4 showed the most conserved sequence pattern with the lowest median entropies.

In contrast, groups 9 and 13 showed conserved sequences. Figure S2D shows the pairwise clustering results for these groups. The data indicated four distinct clusters, one composed of homologous groups 3 and 6 (correlation coefficient > 0.75) and the other comprising groups 7, 9, and 10 (correlation coefficient > 0.50). Homologous group 4 differed the most from the others with no indicated cluster. The other homologous groups comprised the fourth cluster with a correlation coefficient greater than 0.75. However, the single Gaussian distribution pairwise distance of the sequences (Figure S2E) indicated that all homologous groups were homogeneous and must be evaluated as a whole.

#### 4. EbfC/YbaB Structure

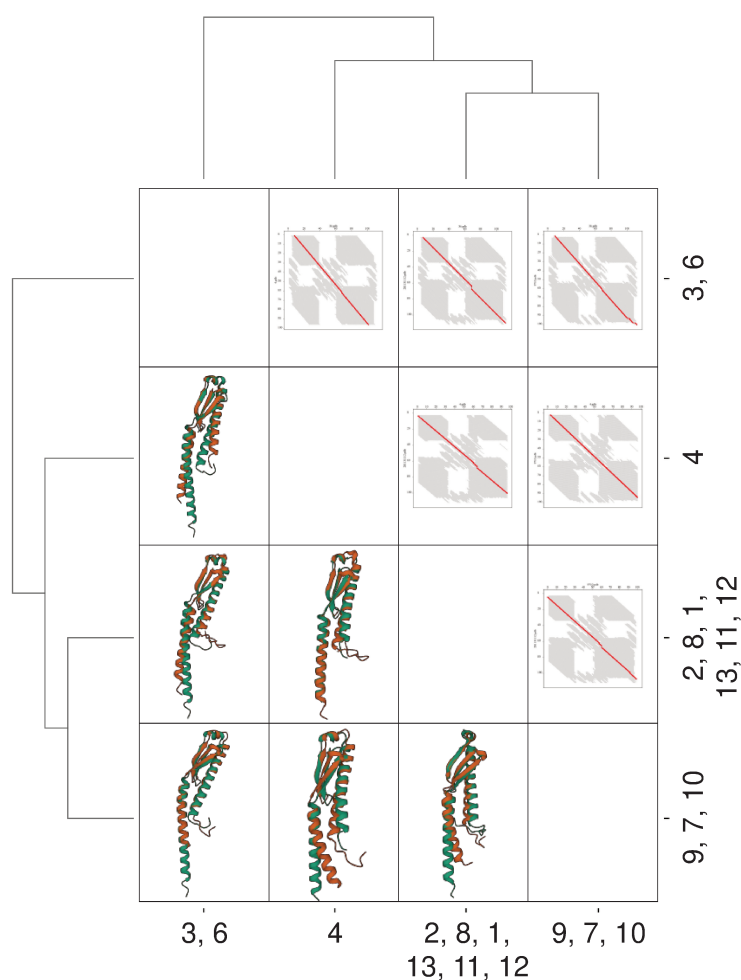
So far, the investigations have revealed that YbaB forms homodimers composed of  $\alpha$ -helices and  $\beta$ -sheets with two  $\alpha$ -helices protruding from a globular region, forming a tweezer-like structure [33,34]. To date, the crystal structure of YbaB has been determined for five organisms: *H. influenzae* (PDB accession 1j8b) [33], *E. coli* (PDB accession 1pug), *Clostridium thermocellum* (PDB accession 1ybx), *Helicobacter pylori* (PDB accession 3f42), and *Mycobacterium tuberculosis* (PDB accession 5yrx) [48]. The canonical structure of EbfC/YbaB consists of an  $\alpha + \beta$  structure with the topology N- $\alpha_1/\beta_1/\beta_2/\beta_3/\alpha_2$ -C (Figure 4A). The crystals revealed a homodimer structured like a “pair of tweezers”, where the  $\beta$ -sheets interact, and the  $\alpha$ -helices are extended like arms and form the putative DNA-binding domain (Figure 4B) [28,33,48]. The EbfC/YbaB protein can also form other organised structures, such as tetramers and octamers, in solution, but their properties are not apparent [29,49]. The space between the arm structures of EbfC is 15–22 Å, which would be a suitable space to fit double-stranded DNA [28].



**Figure 4.** Canonical structure of EbfC/YbaB: (A) ribbon diagram of YbaB monomer highlighting the  $\alpha + \beta$  structure with the topology N- $\alpha_1/\beta_1/\beta_2/\beta_3/\alpha_2$ -C; (B) ribbon diagram of the dimer, which is the most observed functional structure [33]. Structures were predicted using AlphaFold [50] and visualised using Mol\*Viewer [51].

This structure has no characteristic DNA-binding domain; however, experimental evidence showed that the EbfC (YbaB) of *B. burgdorferi*, *H. influenzae*, and *E. coli* binds to DNA [30,34]. Additionally, protein structure analyses indicated the presence of negatively charged protein-binding domains that can mimic the surface of DNA, thus permitting speculation that YbaB can compete with the DNA-binding activity of other proteins [33]. This topic is discussed in the next section.

Figure 5 shows the conservation of the tridimensional structure of YbaB/EbfC across the bacterial domain. All putative proteins belonging to the YbaB/EbfC family (IPR004401) had a conserved structure (N- $\alpha$ 1/ $\beta$ 1/ $\beta$ 2/ $\beta$ 3/ $\alpha$ 2-C). Some differences were observed in the length of the  $\alpha$ 1-helix and the length of the additional amino acid chains at the N-terminus from  $\alpha$ 1 and/or the C-terminus from  $\alpha$ 2. The graph of FATCAT chaining result in Figure 5, upper right corner, revealed the significant alignment of the structures as shown by the almost straight red diagonal line, despite the small changes. Each axis represents the position of specific amino acids in the respective structure, and the red line represents their positional match compared to the pairwise structure. A straight  $-45^\circ$  indicates a perfect pairwise alignment.



**Figure 5.** Pairwise alignment of monomeric structures of putative proteins belonging to the YbaB/EbfC family (IPR004401). Accession: Q3ANM6 (Group 3, 6), Q5L6F1 (Group 4), Q3KFA7 (Group 2, 8, 1, 13, 11, 12), and P75502 (Group 9, 7, 10). Structures were predicted using AlphaFold [50]; pairwise structural alignment was created using FATCAT [52]; and superimposed structures were visualised using Mol\*Viewer [51]. The graph of the FATCAT chaining result is represented in the upper right corner, and the superimposed structures are shown in the lower left corner.



## 5. Ability to Bind to DNA

Thus far, all the data indicate that EbfC/YbaB could be grouped into a family of DNA-binding proteins. The members of this family are encoded by a broad range of bacteria [27,30,34,35,47,49,53–55]. Interestingly, a chloroplast-localised protein encoded by a gene (*lta1*) that controls the tiller angle and gravity response in rice and has homologs in other plants presents a conserved YbaB DNA-binding domain [56]. However, further studies are required to characterise better eukaryotic proteins with DNA-binding domains similar to EbfC/YbaB.

The DNA binding ability of EbfC/YbaB has been demonstrated in *B. burgdorferi*, *H. influenzae*, *E. coli*, *Deinococcus radiodurans*, and *C. crescentus* [28,30,34,47,55]. In *B. burgdorferi*, EbfC has been shown to bind to DNA as a homodimer, but other organised structures can also be formed [28,29]. The DNA-binding properties of *B. burgdorferi* EbfC were demonstrated using an electrophoretic mobility shift assay (EMSA). Experimental data indicated that this protein binds to DNA through its  $\alpha$ -helical domains. As discussed above, the two  $\alpha$ -domains form a protruding tweezer, which is hypothesised to form a DNA-binding domain. Indeed, EMSA analyses showed that EbfC protein variants carrying mutations in nine amino acids located in the  $\alpha$ -helical regions (K16, D20, K23, N77, D78, K82, K84, E85, and K88) were unable to bind to the promoter region of *erp* genes, even at a higher concentration of mutant proteins than that of the wild-type. Protein variants were constructed by mutating wild-type amino acids with alanine [28]. These results demonstrated that the  $\alpha$ -domains are responsible for the DNA-binding properties of EbfC. In particular, residue N77, which was predicted to be located adjacent to the  $\beta$ -sheet region of EbfC, was the only residue among the studied amino acids that appeared to play a role in dimerisation [28]. Interestingly, some of the residues reported to be necessary for binding EbfC to DNA (K16, D20, K23, K82, and K88) in *B. burgdorferi* [28] are not conserved in *H. influenzae*. Although mutations at these residues lead to a loss of the DNA-binding capacity of the *B. burgdorferi* EbfC protein, the mutants still formed higher-order structures such as tetramers and octamers in solution, except for N77. The location of the N77 residue near the  $\beta$ -sheet region probably affects oligomerisation. These results indicated that EbfC multimerisation is independent of its ability to bind DNA [28].

The data obtained for the EbfC protein of *B. burgdorferi* indicated that both the N- and C-termini are involved in the formation of the  $\alpha$ -domains and participate in DNA-binding activity [29]. To determine whether this is true for EbfC/YbaB of another bacterial species, a truncated version of the YbaB protein from *C. crescentus* lacking the putative C-terminal DNA-binding domain was constructed and tested for its binding activity to the operator sequences of *erpAB* operon. The truncated YbaB protein could not form stable protein–DNA complexes, indicating that this domain is essential for DNA-binding activity. Its removal leads to the loss of double-stranded DNA-binding activity in *C. crescentus* [47].

Available evidence suggests that EbfC/YbaB binds to DNA in sequence-specific and non-sequence-specific manners [28,30,33,34]. To date, sequence-specific binding has been described for the EbfC protein of *B. burgdorferi* [28]. EMSA, dissociation constant determination for EbfC–DNA interactions, and chromatin immunoprecipitation (ChIp) followed by sequence analysis, demonstrated that EbfC binds to the 4 bp palindromic sequence 5'-GTnAC-3' with high affinity, which is spread throughout the genome of *B. burgdorferi* [28–30]. In addition, fluorescence microscopy analyses of *B. burgdorferi* cells containing a fusion GFP-tagged EbfC protein indicated that EbfC colocalises with bacterial nucleoids in multiple centres [30]. Further experiments demonstrated that EbfC could bind to partial or identical consensus sequences [28,29]. Although EbfC binds to the 5'-GTnAC-3' sequence with high affinity, it can also bind to other DNA fragments lacking the palindromic sequence with lower affinity [28]. Although the YbaB orthologs of *H. influenzae* and *E. coli* exhibited preferences for specific DNA sequences, they did not show high-affinity binding for the palindromic sequence 5'-GtnAC-3', as observed for *B. burgdorferi* [34]. Therefore, the preferred DNA sequences bound by YbaB in *H. influenzae* and *E. coli* are unknown, but they are probably different from those of *B. burgdorferi* [34]. Similar results were found

for YbaB orthologous proteins from *D. radiodurans* and *C. crescentus* [47,55]. Both proteins were found to bind to the DNA. These proteins seem to have similar characteristics to the members of the YbaB/EbfC family, exhibiting the same tweezer-like conformation, the presence of DNA-binding domains in the N- and C-termini, and probably binding preferentially as a homodimer. The results obtained with *C. crescentus* suggested that the binding of the YbaB protein to DNA is not sequence-dependent; however, further studies are needed to clarify this question.

## 6. EbfC/YbaB Has Important Biological Functions

The functions of EbfC/YbaB have been described for several bacterial species (Table 1). For example, in *B. burgdorferi*, EbfC regulates the expression of pathogenic traits [29]. This bacterium contains multiple plasmids and independent DNA replication elements [57]. Cp32s, one of these elements, is circular and presents characteristics of a bacteriophage genome [57]. Among other traits, Cp32s contain genes that encode a polymorphic family of surface lipoproteins, the Erp proteins, involved in pathogenesis by promoting binding to plasminogen, laminin, and complement factor H, thus conferring adhesion to host tissues and inhibiting complement activation [38]. EbfC was found to regulate the transcription of *erp* operons by binding to operator sequences in the promoter region of *erp* genes. Both EbfC and BpaB (*borrelial ParB*) bind to the regulatory region of the *erp* operons. BpaB is a repressor that inhibits the transcription of *erp*. This protein binds cooperatively to the *erp* operator, thereby blocking the binding of RNAPol to the promoter. This repression is enhanced by the binding of BpuR (*borrelial PUR* domain), another DNA-binding protein [38,58]. EbfC functions as an anti-repressor of *erp* operons and competes with BpaB to bind to the regulatory region, thus stimulating the transcription of *erp* genes [31,38]. *ebfC* exhibited a higher transcription level during the exponential growth phase than during the stationary phase, where transcripts were not detected. The expression of *bpaB* and *bpuR* was higher in unfed ticks in the insect's midgut; however, *ebfC* expression was higher during the transmission of the bacteria from the feeding tick to the host, which explains the activation of *erp* transcription [30,31,38]. Therefore, the transcription of *erp* operons is inhibited by BpaB in the non-fed tick vector, but it is highly induced during the infection of mammals [30,31,38,59]. This explains why this protein is called *erp*-binding factor, chromosomal (EbfC) [28–31]. Interestingly, *bpaB* is a gene located in extra-chromosomal elements, whereas *bpuR* and *ebfC* are chromosomal genes [38].

As discussed previously, Stevenson et al. [38] recently suggested the involvement of DnaA in *dnaX-ebfC* expression. DnaA is necessary for DNA replication and functions as a regulatory protein [60]. However, because *dnaX* encodes DNA polymerase III subunits, these findings suggest a link between DNA replication and *ebfC* expression in *B. burgdorferi* [38].

Attempts to construct mutants of *B. burgdorferi* by *ebfC* deletion have been unsuccessful, suggesting that this is an essential protein in this species [28,31]. However, this is not the case for other species such as *E. coli*, *Streptomyces lividans*, and *Deinococcus radiodurans* [35,55,61].

As an alternative method for analysing the effects of EbfC on global gene expression, the EbfC protein was overexpressed in *B. burgdorferi*, and the transcriptome was analysed using RNA-seq [30]. Upon comparison, the abundance of *ebfC* transcripts was 28-fold in the bacteria overexpressing EbfC than in the uninduced control. This overexpression did not measurably affect the growth rate; however, transcriptome analyses indicated that EbfC affects gene expression globally in *B. burgdorferi*. In the EbfC-overexpressing strain, the expression levels of 52 genes were affected positively or negatively, representing approximately 4.5% of the gene content of the bacterial strain studied [30]. The genomic distribution of genes affected by EbfC overexpression was biased toward small replicons, an exciting observation for a NAP. However, the genomic characteristics of *B. burgdorferi* may explain this observation. The authors specified that smaller replicons carry most infection-related genes, whereas the chromosome is rich in housekeeping genes [30]. The functions of most of the genes affected by EbfC overexpression were unknown (29%), followed



by metabolic processes (23%); lipoproteins (21%); DNA recombination, replication, and repair (9%); chemotaxis (8%); and others (10%) [30]. For instance, genes induced by EbfC overexpression included the genes encoding outer membrane proteins, such as *dbpB*, involved in adherence to host tissues, enzymes such as alanine racemase, and the gamma subunit of exodeoxyribonuclease V. The genes repressed by EbfC included *flhB*, which encodes a flagellar assembly/export protein, and *bbb07*, an outer membrane protein involved in tick colonisation [30].

**Table 1.** Functions of EbfC/YbaB in gene regulation and in other cellular processes as described in the literature.

Biological Roles of EbfC/YbaB	Bacterial Species	Ref.
<b>Regulatory functions and DNA Repair/Protection</b>		
Component of <i>sigma E</i> regulon (envelop stress response)	<i>Escherichia coli</i>	[41]
Regulation of pathogenic factors (EbfC as anti-repressor, stimulating the transcription of the <i>erp</i> operon)	<i>Borrelia burgdorferi</i>	[29]
Global regulator of gene expression	<i>Borrelia burgdorferi</i>	[30]
UV survival	<i>Deinococcus radiodurans</i> <i>Escherichia coli</i>	[55,61]
DNA protection from enzymatic action	<i>Deinococcus radiodurans</i> <i>Caulobacter crescentus</i>	[47,55]
DNA protection under iron-limiting conditions	<i>Paenibacillus riograndensis</i>	[62]
<b>Cellular processes involving EbfC/YbaB</b>		
Higher expression of several heterologous membrane proteins	<i>Escherichia coli</i>	[43]
EbfC/YbaB target of ClpYQ protease	<i>Escherichia coli</i>	[63]

In some bacterial species, *ebfC/ybaB* has been suggested to play a role in DNA repair by forming an operon with *recR*. Indeed, results indicate that YbaB plays a role in DNA repair because a null-mutant *E. coli* exhibited higher sensitivity to radiation than the wild-type strain [61]. However, this is not a general characteristic. *Streptomyces* is a genus of bacteria frequently used to study DNA repair because genetic instability is a common characteristic in this group. For example, a study to characterise the *recR* gene of *Streptomyces* demonstrated that mutants for *recR* or *orf1recR* of *S. lividans* presented an increased susceptibility to DNA-damaging agents [35]. Both mutant strains were complemented with the *recR* gene alone. Considering that *orf1* is orthologous to *ebfC/ybaB*, *orf1* does not seem essential for the repair function in *S. lividans* [35]. Interestingly, the same mutant strains were complemented by *orf107recR* from *Bacillus subtilis*; however, attempts to complete them with *B. subtilis recR* alone or with the *ybaBrecR* region of *E. coli* failed [35]. *orf107* is also orthologous to the *ebfC/ybaB* gene. In addition, *dr0199*, a gene orthologous to *ebfC* in *Deinococcus radiodurans*, has been suggested to play a role in DNA repair [55]. However, this activity may be due to the binding and protection of DNA, which is discussed further ahead in the text. During the construction of *ybaB* mutants, the occurrence of polar effects on *recR* should be considered. This is particularly important in species in which *ybaB* and *recR* translations are coupled, such as *E. coli*. Based on all these results, further studies are necessary to characterise better the role of EbfC/YbaB protein in DNA repair.

Studies have also indicated that EbfC/YbaB may participate in the stress response [41,55,62]. As discussed above, *ybaB* appears to belong to the  $\sigma^E$  regulon of *E. coli*, as demonstrated by a two-plasmid system [41]. However, no data indicate its partic-

ipation in the *S. enterica*  $\sigma^E$  regulon [64,65], a phylogenetically related species belonging to the *Enterobacteriaceae* family. The  $\sigma^E$  regulon is activated by a cascade of events in response to misfolded and/or mis-translocated outer membrane proteins or LPS in Gram-negative bacteria as a consequence of exposure to different stress conditions [66]. In *E. coli*,  $\sigma^E$  is an essential gene; however, compensatory mutations can generate variants that can survive.

In contrast, this gene is not essential in *S. enterica*, and the mutants are viable [66]. Interestingly, when co-expressed with recombinant prokaryotic or eukaryotic membrane proteins, YbaB can enhance the expression and accumulation of these proteins in *E. coli*, in some cases by a factor of 10-fold [43]. Therefore, YbaB appears to be a general “enhancer” of membrane protein production in *E. coli* [43]. These results reinforce the possible link between YbaB and the cell envelope; however, further analyses are needed to describe better their functions related to the envelope stress response.

Studies with different bacteria have suggested a possible role for EbfC/YbaB in DNA protection. *D. radiodurans* is a model organism for studying DNA repair based on its high capacity to repair DNA damage caused by mutagenic agents. This bacterium contains a gene (*dr0199*) orthologous to EbfC. A strain carrying a deletion mutation in *dr0199* exhibited high sensitivity to hydrogen peroxide and UV and gamma radiation, which was complemented by the reintroduction of the wild-type gene [55]. The authors suggested that the protective effect on DNA can be mediated by direct DNA binding; however, this protein seemed to modulate the expression of different genes in *D. radiodurans*. In *Paenibacillus riograndensis*, a plant growth-promoting bacterium, the *ebfC/ybaB* ortholog (0116.0006 0.0231) was highly expressed during growth in an iron-limiting medium when compared to growth under sufficient iron conditions [62]. The mechanism involved in activating 0116.0006 0.0231 expression was not determined; however, the authors suggested a function of DNA protection for this EbfC/YbaB orthologous gene [62]. Furthermore, recent characterisation of YbaB in *C. crescentus* demonstrated that this protein binds to DNA in a non-specific manner, can compact DNA, and protects it against enzymatic degradation [47]. The role of DNA-binding proteins, including some NAPs, in protecting DNA has recently been reviewed [67].

Protein turnover is an essential characteristic of any organism and is necessary for modulating peptide and protein levels, including misfolding and regulatory proteins. In *E. coli*, ClpYQ is an ATP-dependent protease, and YbaB/EbfC was demonstrated to be one of the targets of this protease [63]. These data indicated that the regulation of YbaB levels in cells is essential for *E. coli*, reinforcing the need for further studies to better characterise the roles of this NAP in prokaryotes.

## 7. Conclusions

EbfC/YbaB is a widely distributed family of prokaryote proteins capable of binding to DNA, as demonstrated in different studies. Furthermore, the structure of these proteins is conserved, exhibiting a general topology of N- $\alpha 1/\beta 1/\beta 2/\beta 3/\alpha 2$ -C, with regions involved in dimerisation and DNA-binding. The DNA-binding region, structured as a pair of tweezers, probably represents a DNA-binding domain that is not well-characterised yet. Despite its wide distribution, the roles of EbfC/YbaB have been described in a few bacterial species. In *B. burgdorferi*, EbfC binds specifically to DNA sequences with high affinity. However, this protein can also bind to DNA sequences in a lower affinity non-specific manner. Data from the literature suggest that EbfC/YbaB orthologs may be involved in the bacterial envelope stress response, DNA protection, and repair. However, further studies are required to clarify the role of EbfC/YbaB orthologs in these processes. EbfC of *B. burgdorferi* is undoubtedly the best-characterised EbfC/YbaB ortholog in terms of DNA binding, regulation of pathogenic traits, and as a global regulator. Taken together, these data suggest that EbfC/YbaB is a NAP, but its biological function has not yet been characterised in most prokaryotic species. Therefore, further studies are necessary to better describe the biological roles of this NAP in prokaryotes.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms10101945/s1>, **Figure S1:** Sequence logos for all the reviewed putative proteins belonging to the YbaB/EbfC family (IPR004401). The putative sequences are deposited in the InterPro protein families and domains database [1]. The database contains 585 reviewed sequences. Sequence logos were constructed using MetaLogo [2]. The tree on the left indicates the relationships among the groups. The red dot on the tree indicates the group containing the target sequence. Light blue coloured strips connect conserved positions among groups; **Figure S2:** Results obtained from MetaLogo [2]. (a) Sequence counts of each group; (b) Entropy heatmap of each group; (c) Boxplot of entropies of positions in all groups; (d) Clustering result of sequence logo groups revealing the pairwise relationships; (e) Distribution of pairwise distances of the nodes in the phylogenetic tree; **Table S1:** Homologous group, accession to InterPro database, strain ID, sequence length, and sequence match for YbaB/EbfC.

**Author Contributions:** Conceptualisation, T.F.V.B.C. and M.B.; writing—original draft preparation, T.F.V.B.C., M.T.P.G., G.P.J. and M.B.; writing—review and editing, M.B.; supervision, M.B.; project administration, M.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the São Paulo Research Foundation (FAPESP), grant numbers 2021/00465-0 and 2017/10051-2. T.F.V.B.C. was supported by a fellowship from the National Council for Scientific and Technological Development (CNPq), grant number 142190/2017-9. M.T.P.G. was supported by fellowships from CNPq, grant number 130553/2020-4, and FAPESP, grant number 2020/01535-9. G.P.J. was supported by the Administrative Department of Science, Technology, and Innovation (ColCiencias), grant number 2016/772, and M.B. is a scientific fellow of CNPq, grant number 309380/2019-7.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Dame, R.T.; Rashid, F.-Z.M.; Grainger, D.C. Chromosome Organization in Bacteria: Mechanistic Insights into Genome Structure and Function. *Nat. Rev. Genet.* **2020**, *21*, 227–242. [\[CrossRef\]](#)
2. Dillon, S.C.; Dorman, C.J. Bacterial Nucleoid-Associated Proteins, Nucleoid Structure and Gene Expression. *Nat. Rev. Microbiol.* **2010**, *8*, 185–195. [\[CrossRef\]](#)
3. Amemiya, H.M.; Schroeder, J.; Freddolino, P.L. Nucleoid-Associated Proteins Shape Chromatin Structure and Transcriptional Regulation across the Bacterial Kingdom. *Transcription* **2021**, *12*, 182–218. [\[CrossRef\]](#)
4. Ó Cróinín, T.; Carroll, R.K.; Kelly, A.; Dorman, C.J. Roles for DNA Supercoiling and the Fis Protein in Modulating Expression of Virulence Genes during Intracellular Growth of *Salmonella enterica* Serovar Typhimurium. *Mol. Microbiol.* **2006**, *62*, 869–882. [\[CrossRef\]](#)
5. Kelly, A.; Goldberg, M.D.; Carroll, R.K.; Danino, V.; Hinton, J.C.D.; Dorman, C.J. A Global Role for Fis in the Transcriptional Control of Metabolism and Type III Secretion in *Salmonella enterica* Serovar Typhimurium. *Microbiology* **2004**, *150*, 2037–2053. [\[CrossRef\]](#)
6. Mangan, M.W.; Lucchini, S.; Danino, V.; Cróinín, T.Ó.; Hinton, J.C.D.; Dorman, C.J. The Integration Host Factor (IHF) Integrates Stationary-Phase and Virulence Gene Expression in *Salmonella enterica* Serovar Typhimurium. *Mol. Microbiol.* **2006**, *59*, 1831–1847. [\[CrossRef\]](#)
7. Dorman, C.J. Function of Nucleoid-Associated Proteins in Chromosome Structuring and Transcriptional Regulation. *J. Mol. Microbiol. Biotechnol.* **2014**, *24*, 316–331. [\[CrossRef\]](#)
8. Dorman, C.J. Co-Operative Roles for DNA Supercoiling and Nucleoid-Associated Proteins in the Regulation of Bacterial Transcription. *Biochem. Soc. Trans.* **2013**, *41*, 542–547. [\[CrossRef\]](#)
9. Navarre, W.W.; Porwollik, S.; Wang, Y.; McClelland, M.; Rosen, H.; Libby, S.J.; Fang, F.C. Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in *Salmonella*. *Science* **2006**, *313*, 236–238. [\[CrossRef\]](#)
10. Lucchini, S.; Rowley, G.; Goldberg, M.D.; Hurd, D.; Harrison, M.; Hinton, J.C.D. H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria. *PLoS Pathog.* **2006**, *2*, e81. [\[CrossRef\]](#)
11. Dillon, S.C.; Cameron, A.D.S.; Hokamp, K.; Lucchini, S.; Hinton, J.C.D.; Dorman, C.J. Genome-Wide Analysis of the H-NS and Sfh Regulatory Networks in *Salmonella* Typhimurium Identifies a Plasmid-Encoded Transcription Silencing Mechanism. *Mol. Microbiol.* **2010**, *76*, 1250–1265. [\[CrossRef\]](#)
12. Will, W.R.; Navarre, W.W.; Fang, F.C. Integrated Circuits: How Transcriptional Silencing and Counter-Silencing Facilitate Bacterial Evolution. *Curr. Opin. Microbiol.* **2015**, *23*, 8–13. [\[CrossRef\]](#)
13. Dorman, M.J.; Dorman, C.J. Regulatory Hierarchies Controlling Virulence Gene Expression in *Shigella flexneri* and *Vibrio cholerae*. *Front. Microbiol.* **2018**, *9*, 2686. [\[CrossRef\]](#)

14. Dorman, C.J.; Schumacher, M.A.; Bush, M.J.; Brennan, R.G.; Buttner, M.J. When Is a Transcription Factor a NAP? *Curr. Opin. Microbiol.* **2020**, *55*, 26–33. [\[CrossRef\]](#)
15. Flåtten, I.; Skarstad, K. The Fis Protein Has a Stimulating Role in Initiation of Replication in *Escherichia coli* In Vivo. *PLoS ONE* **2013**, *8*, e83562. [\[CrossRef\]](#)
16. Dorman, C.J.; Colgan, A.; Dorman, M.J. Bacterial Pathogen Gene Regulation: A DNA-Structure-Centred View of a Protein-Dominated Domain. *Clin. Sci.* **2016**, *130*, 1165–1177. [\[CrossRef\]](#)
17. Leite, B.; Werle, C.H.; do Carmo, C.P.; Nóbrega, D.B.; Milanez, G.P.; Culler, H.F.; Sircili, M.P.; Alvarez-Martinez, C.E.; Brocchi, M. Integration Host Factor Is Important for Biofilm Formation by *Salmonella enterica* Enteritidis. *Pathog. Dis.* **2017**, *75*, ftx074. [\[CrossRef\]](#)
18. Fitzgerald, S.; Dillon, S.C.; Chao, T.-C.; Wiencko, H.L.; Hokamp, K.; Cameron, A.D.S.; Dorman, C.J. Re-Engineering Cellular Physiology by Rewiring High-Level Global Regulatory Genes. *Sci. Rep.* **2015**, *5*, 17653. [\[CrossRef\]](#)
19. Ali, B.M.; Amit, R.; Braslavsky, I.; Oppenheim, A.B.; Gileadi, O.; Stavans, J. Compaction of Single DNA Molecules Induced by Binding of Integration Host Factor (IHF). *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10658–10663. [\[CrossRef\]](#)
20. Duprey, A.; Reverchon, S.; Nasser, W. Bacterial Virulence and Fis: Adapting Regulatory Networks to the Host Environment. *Trends Microbiol.* **2014**, *22*, 92–99. [\[CrossRef\]](#)
21. Cameron, A.D.S.; Stoebel, D.M.; Dorman, C.J. DNA Supercoiling Is Differentially Regulated by Environmental Factors and FIS in *Escherichia coli* and *Salmonella enterica*. *Mol. Microbiol.* **2011**, *80*, 85–101. [\[CrossRef\]](#)
22. Mangan, M.W.; Lucchini, S.; Ó Cróinín, T.; Fitzgerald, S.; Hinton, J.C.D.; Dorman, C.J. Nucleoid-Associated Protein HU Controls Three Regulons That Coordinate Virulence, Response to Stress and General Physiology in *Salmonella enterica* Serovar Typhimurium. *Microbiology* **2011**, *157*, 1075–1087. [\[CrossRef\]](#)
23. Lucchini, S.; McDermott, P.; Thompson, A.; Hinton, J.C.D. The H-NS-like Protein StpA Represses the RpoS (Sigma 38) Regulon during Exponential Growth of *Salmonella* Typhimurium. *Mol. Microbiol.* **2009**, *74*, 1169–1186. [\[CrossRef\]](#)
24. Cho, B.-K.; Knight, E.M.; Barrett, C.L.; Palsson, B.Ø. Genome-Wide Analysis of Fis Binding in *Escherichia coli* Indicates a Causative Role for A-/AT-Tracts. *Genome Res.* **2008**, *18*, 900–910. [\[CrossRef\]](#)
25. Boubrik, F.; Bonnefoy, E.; Rouvière-Yaniv, J. HU and IHF: Similarities and Differences. In *Escherichia coli*, the Lack of HU Is Not Compensated for by IHF. *Res. Microbiol.* **1991**, *142*, 239–247. [\[CrossRef\]](#)
26. Castaing, B.; Zelwer, C.; Laval, J.; Boiteux, S. HU Protein of *Escherichia coli* Binds Specifically to DNA That Contains Single-Strand Breaks or Gaps. *J. Biol. Chem.* **1995**, *270*, 10291–10296. [\[CrossRef\]](#)
27. Perez-Rueda, E.; Ibarra, J.A. Distribution of Putative Xenogeneic Silencers in Prokaryote Genomes. *Comput. Biol. Chem.* **2015**, *58*, 167–172. [\[CrossRef\]](#)
28. Riley, S.P.; Bykowski, T.; Cooley, A.E.; Burns, L.H.; Babb, K.; Brissette, C.A.; Bowman, A.; Rotondi, M.; Miller, M.C.; DeMoll, E.; et al. *Borrelia burgdorferi* EbfC Defines a Newly-Identified, Widespread Family of Bacterial DNA-Binding Proteins. *Nucleic Acids Res.* **2009**, *37*, 1973–1983. [\[CrossRef\]](#)
29. Babb, K.; Bykowski, T.; Riley, S.P.; Miller, M.C.; Demoll, E.; Stevenson, B. *Borrelia burgdorferi* EbfC, a Novel, Chromosomally Encoded Protein, Binds Specific DNA Sequences Adjacent to *erp* Loci on the Spirochete's Resident Cp32 Prophages. *J. Bacteriol.* **2006**, *188*, 4331–4339. [\[CrossRef\]](#)
30. Jutras, B.L.; Bowman, A.; Brissette, C.A.; Adams, C.A.; Verma, A.; Chenail, A.M.; Stevenson, B. EbfC (YbaB) Is a New Type of Bacterial Nucleoid-Associated Protein and a Global Regulator of Gene Expression in the Lyme Disease Spirochete. *J. Bacteriol.* **2012**, *194*, 3395–3406. [\[CrossRef\]](#)
31. Jutras, B.L.; Verma, A.; Adams, C.A.; Brissette, C.A.; Burns, L.H.; Whetstone, C.R.; Bowman, A.; Chenail, A.M.; Zückert, W.R.; Stevenson, B. BpaB and EbfC DNA-Binding Proteins Regulate Production of the Lyme Disease Spirochete's Infection-Associated Erp Surface Proteins. *J. Bacteriol.* **2012**, *194*, 778–786. [\[CrossRef\]](#)
32. Steere, A.C.; Strle, F.; Wormser, G.P.; Hu, L.T.; Branda, J.A.; Hovius, J.W.R.; Li, X.; Mead, P.S. Lyme Borreliosis. *Nat. Rev. Dis. Primers* **2016**, *2*, 16090. [\[CrossRef\]](#)
33. Lim, K.; Tempczyk, A.; Parsons, J.F.; Bonander, N.; Toedt, J.; Kelman, Z.; Howard, A.; Eisenstein, E.; Herzberg, O. Crystal Structure of YbaB from *Haemophilus influenzae* (HI0442), a Protein of Unknown Function Coexpressed with the Recombinational DNA Repair Protein RecR. *Proteins Struct. Funct. Bioinform.* **2003**, *50*, 375–379. [\[CrossRef\]](#)
34. Cooley, A.E.; Riley, S.P.; Kral, K.; Miller, M.C.; DeMoll, E.; Fried, M.G.; Stevenson, B. DNA-Binding by *Haemophilus influenzae* and *Escherichia coli* YbaB, Members of a Widely-Distributed Bacterial Protein Family. *BMC Microbiol.* **2009**, *9*, 137. [\[CrossRef\]](#)
35. Peláez, A.I.; Ribas-Aparicio, R.M.; Gómez, A.; Rodicio, M.R. Structural and Functional Characterization of the *recR* Gene of *Streptomyces*. *Mol. Genet. Genom.* **2001**, *265*, 663–672. [\[CrossRef\]](#)
36. Yeung, T.; Mullin, D.A.; Chen, K.S.; Craig, E.A.; Bardwell, J.C.; Walker, J.R. Sequence and Expression of the *Escherichia coli* *recR* Locus. *J. Bacteriol.* **1990**, *172*, 6042–6047. [\[CrossRef\]](#)
37. Mahdi, A.A.; Lloyd, R.G. The *recR* Locus of *Escherichia coli* K-12: Molecular Cloning, DNA Sequencing and Identification of the Gene Product. *Nucleic Acids Res.* **1989**, *17*, 6781–6794. [\[CrossRef\]](#)
38. Stevenson, B.; Krusenstjerna, A.C.; Castro-Padovani, T.N.; Savage, C.R.; Jutras, B.L.; Saylor, T.C. The Consistent Tick-Vertebrate Infectious Cycle of the Lyme Disease Spirochete Enables *Borrelia burgdorferi* to Control Protein Expression by Monitoring Its Physiological Status. *J. Bacteriol.* **2022**, *204*, e0060621. [\[CrossRef\]](#)



39. Michel, B.; Leach, D. Homologous Recombination—Enzymes and Pathways. *EcoSal Plus* **2012**, *5*, 21. [\[CrossRef\]](#)
40. Oakley, A.J. A Structural View of Bacterial DNA Replication. *Protein Sci.* **2019**, *28*, 990–1004. [\[CrossRef\]](#)
41. Rezuchova, B.; Miticka, H.; Homerova, D.; Roberts, M.; Kormanec, J. New Members of the *Escherichia coli*  $\sigma^E$  Regulon Identified by a Two-Plasmid System. *FEMS Microbiol. Lett.* **2003**, *225*, 1–7. [\[CrossRef\]](#)
42. Saha, S.; Lach, S.R.; Konovalova, A. Homeostasis of the Gram-Negative Cell Envelope. *Curr. Opin. Microbiol.* **2021**, *61*, 99–106. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Skretas, G.; Georgiou, G. Simple Genetic Selection Protocol for Isolation of Overexpressed Genes That Enhance Accumulation of Membrane-Integrated Human G Protein-Coupled Receptors in *Escherichia coli*. *Appl. Environ. Microbiol.* **2010**, *76*, 5852–5859. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Blum, M.; Chang, H.-Y.; Chuguransky, S.; Grego, T.; Kandasamy, S.; Mitchell, A.; Nuka, G.; Paysan-Lafosse, T.; Qureshi, M.; Raj, S.; et al. The InterPro Protein Families and Domains Database: 20 Years On. *Nucleic Acids Res.* **2021**, *49*, D344–D354. [\[CrossRef\]](#)
45. Chen, Y.; He, Z.; Men, Y.; Dong, G.; Hu, S.; Ying, X. MetaLogo: A Heterogeneity-Aware Sequence Logo Generator and Aligner. *Brief. Bioinform.* **2022**, *23*, bbab591. [\[CrossRef\]](#)
46. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152. [\[CrossRef\]](#)
47. Pal, P.; Modi, M.; Ravichandran, S.; Yennamalli, R.M.; Priyadarshini, R. DNA-Binding Properties of YbaB, a Putative Nucleoid-Associated Protein from *Caulobacter crescentus*. *Front. Microbiol.* **2021**, *12*, 733344. [\[CrossRef\]](#)
48. Gopalan, A.; Deka, G.; Prabhavathi, M.; Savithri, H.S.; Murthy, M.R.N.; Raja, A. Structural and Biophysical Characterization of Rv3716c, a Hypothetical Protein from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* **2018**, *495*, 982–987. [\[CrossRef\]](#)
49. Tian, S.; Chen, H.; Sun, T.; Wang, H.; Zhang, X.; Liu, Y.; Xia, J.; Guo, C.; Lin, D. Expression, Purification and Characterization of Esx-1 Secretion-Associated Protein EspL from *Mycobacterium tuberculosis*. *Protein Expr. Purif.* **2016**, *128*, 42–51. [\[CrossRef\]](#)
50. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [\[CrossRef\]](#)
51. Sehna, D.; Bittrich, S.; Deshpande, M.; Svobodová, R.; Berka, K.; Bazgier, V.; Velankar, S.; Burley, S.K.; Koča, J.; Rose, A.S. Mol\* Viewer: Modern Web App for 3D Visualization and Analysis of Large Biomolecular Structures. *Nucleic Acids Res.* **2021**, *49*, W431–W437. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Li, Z.; Jaroszewski, L.; Iyer, M.; Sedova, M.; Godzik, A. FATCAT 2.0: Towards a Better Understanding of the Structural Diversity of Proteins. *Nucleic Acids Res.* **2020**, *48*, W60–W64. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Carnielli, C.M.; Artier, J.; Franco de Oliveira, J.C.; Novo-Mansur, M.T.M. Interaction Network and Mass Spectrometry Data of *Xanthomonas citri* Subsp. *citri* Surface Proteins from Differential Proteomic Analysis of Infectious and Non-Infectious Cells. *Data Brief* **2016**, *8*, 1400–1411. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Turab Naqvi, A.A.; Rahman, S.; Rubi, Zeya, F.; Kumar, K.; Choudhary, H.; Jamal, M.S.; Kim, J.; Hassan, M.I. Genome Analysis of *Chlamydia trachomatis* for Functional Characterization of Hypothetical Proteins to Discover Novel Drug Targets. *Int. J. Biol. Macromol.* **2017**, *96*, 234–240. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Wang, H.; Wang, F.; Hua, X.; Ma, T.; Chen, J.; Xu, X.; Wang, L.; Tian, B.; Hua, Y. Genetic and Biochemical Characteristics of the Histone-like Protein DR0199 in *Deinococcus radiodurans*. *Microbiology* **2012**, *158*, 936–943. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Pan, X.; Li, Y.; Zhang, H.; Liu, W.; Dong, Z.; Liu, L.; Liu, S.; Sheng, X.; Min, J.; Huang, R.; et al. The Chloroplast-Localized Protein LTA1 Regulates Tiller Angle and Yield of Rice. *Crop J.* **2022**, *10*, 952–961. [\[CrossRef\]](#)
57. Casjens, S.R.; Di, L.; Akther, S.; Mongodin, E.F.; Luft, B.J.; Schutzer, C.M.; Qiu, W.-G. Primordial Origin and Diversification of Plasmids in Lyme Disease Agent Bacteria. *BMC Genom.* **2018**, *19*, 218. [\[CrossRef\]](#)
58. Jutras, B.L.; Chenail, A.M.; Carroll, D.W.; Miller, M.C.; Zhu, H.; Bowman, A.; Stevenson, B. Bpur, the Lyme Disease Spirochete's PUR Domain Protein: Identification as a Transcriptional Modulator and Characterization of Nucleic Acid Interactions. *J. Biol. Chem.* **2013**, *288*, 26220–26234. [\[CrossRef\]](#)
59. Medrano, M.S.; Ding, Y.; Wang, X.-G.; Lu, P.; Coburn, J.; Hu, L.T. Regulators of Expression of the Oligopeptide Permease A Proteins of *Borrelia burgdorferi*. *J. Bacteriol.* **2007**, *189*, 2653–2659. [\[CrossRef\]](#)
60. Menikpurage, I.P.; Woo, K.; Mera, P.E. Transcriptional Activity of the Bacterial Replication Initiator DnaA. *Front. Microbiol.* **2021**, *12*, 662317. [\[CrossRef\]](#)
61. Sargentini, N.J.; Gularte, N.P.; Hudman, D.A. Screen for Genes Involved in Radiation Survival of *Escherichia coli* and Construction of a Reference Database. *Mutat. Res./Fundam. Mol. Mech. Mutagenes.* **2016**, *793–794*, 1–14. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Sperb, E.R.; Tadra-Sfeir, M.Z.; Sperotto, R.A.; Fernandes, G.D.C.; Pedrosa, F.D.O.; de Souza, E.M.; Passaglia, L.M.P. Iron Deficiency Resistance Mechanisms Enlightened by Gene Expression Analysis in *Paenibacillus riograndensis* SBR5. *Res. Microbiol.* **2016**, *167*, 501–509. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Tsai, C.-H.; Ho, Y.-H.; Sung, T.-C.; Wu, W.-F.; Chen, C.-S. *Escherichia coli* Proteome Microarrays Identified the Substrates of ClpYQ Protease. *Mol. Cell. Proteom.* **2017**, *16*, 113–120. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Skovierova, H.; Rowley, G.; Rezuchova, B.; Homerova, D.; Lewis, C.; Roberts, M.; Kormanec, J. Identification of the SigmaE Regulon of *Salmonella enterica* Serovar Typhimurium. *Microbiology* **2006**, *152*, 1347–1359. [\[CrossRef\]](#) [\[PubMed\]](#)

- 
65. Li, J.; Overall, C.C.; Johnson, R.C.; Jones, M.B.; McDermott, J.E.; Heffron, F.; Adkins, J.N.; Cambronne, E.D. ChIP-Seq Analysis of the  $\sigma^E$  Regulon of *Salmonella enterica* Serovar Typhimurium Reveals New Genes Implicated in Heat Shock and Oxidative Stress Response. *PLoS ONE* **2015**, *10*, e0138466. [[CrossRef](#)]
  66. Hews, C.L.; Cho, T.; Rowley, G.; Raivio, T.L. Maintaining Integrity Under Stress: Envelope Stress Response Regulation of Pathogenesis in Gram-Negative Bacteria. *Front. Cell. Infect. Microbiol.* **2019**, *9*, 313. [[CrossRef](#)]
  67. Molan, K.; Žgur Bertok, D. Small Prokaryotic DNA-Binding Proteins Protect Genome Integrity throughout the Life Cycle. *IJMS* **2022**, *23*, 4008. [[CrossRef](#)]