Heliyon 9 (2023) e16661

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Evolutionary analysis and structure modelling of the Rcs-repressor IgaA unveil a functional role of two cytoplasmic small β-barrel (SBB) domains

Leticia Rodríguez [a], Marcos Peñalver [a,b], Patricia Casino [c,d,e], Francisco García-del Portillo [a,*]

[a] *Laboratory of Intracellular Bacterial Pathogens, National Center for Biotechnology-Consejo Superior de Investigaciones Científicas (CNB-CSIC), Madrid, Spain*
[b] *Departamento de Biología Molecular, Universidad Autónoma de Madrid (UAM), Madrid, Spain*
[c] *Departamento de Bioquímica y Biología Molecular, Universitat de València, Burjassot, Spain*
[d] *Instituto Universitario de Biotecnología y Biomedicina BIOTECMED, Universitat de València, Burjassot, Spain*
[e] *CIBER de Enfermedades Raras (CIBERER-ISCIII), Madrid, Spain*

ARTICLE INFO

Keywords:
Salmonella
IgaA
Evolution
Enterobacterales
SBB domain
Rcs system

ABSTRACT

The Rcs sensor system, comprising the RcsB/RcsC/RcsD and RcsF proteins, is used by bacteria of the order *Enterobacterales* to withstand envelope damage. In non-stress conditions, Rcs is repressed by IgaA, a membrane protein with three cytoplasmic regions (cyt-1, cyt-2 and cyt-3). How the Rcs-IgaA axis evolved within *Enterobacterales* has not been yet explored. Here, we report phylogenetic data supporting co-evolution of IgaA with RcsC/RcsD. Functional exchange assays showed that IgaA from *Shigella* and *Dickeya*, but not from *Yersinia* or the endosymbionts *Photorhabdus* and *Sodalis*, repress the Rcs system of *Salmonella*. IgaA from *Dickeya*, however, repress only partially the Rcs system despite being produced at high levels in the complementation assay. The modelled structures of these IgaA variants uncovered one periplasmic and two cytoplasmic conserved β-rich architectures forming partially closed small β-barrel (SBB) domains. Conserved residues map in a connector linking cytoplasmic SSB-1 and SBB-2 domains (E180-R265); a region of cyt-1 facing cyt-2 (R188-E194-D309 and T191-H326); and between cyt-2 and cyt-3 (H293-E328-R686). These structures validated early *in vivo* studies in *Salmonella* that assigned a role in function to R188, T191 and G262, and in addition revealed a previously unnoticed "hybrid" SBB-2 domain to which cyt-1 and cyt-2 contribute. IgaA variants not functional or partially functional in *Salmonella* lack H192-P249 and R255-D313 interactions. Among these variants, only IgaA from *Dickeya* conserves the helix α6 in SSB-1 that is present in IgaA from *Salmonella* and *Shigella*. RcsF and RcsD, which interact directly with IgaA, failed to show structural features linked to specific IgaA variants. Altogether, our data provide new insights into IgaA by mapping residues selected differently during evolution and involved in function. Our data also infer contrasting lifestyles of *Enterobacterales* bacteria as source of variability in the IgaA-RcsD/IgaA-RcsF interactions.

* Corresponding author.
E-mail address: fgportillo@cnb.csic.es (F. García-del Portillo).

https://doi.org/10.1016/j.heliyon.2023.e16661
Received 28 October 2022; Received in revised form 17 May 2023; Accepted 23 May 2023
Available online 26 May 2023

1. Introduction

The Rcs system plays a key sensory role in bacteria of the order *Enterobacterales* for combating insults to their envelope [1,2]. Upon signal perception, the Rcs system is activated by a phosphorelay cascade involving the inner membrane proteins RcsC/RcsD, which transduce the signal to the cytoplasmic transcriptional regulator RcsB [1]. Studies performed mostly in *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) and *Escherichia coli* have provided compelling evidence for a tight control of the Rcs system under non-stress conditions. This control impedes unnecessary activation, which has been shown to be detrimental for viability [3–7].

An important element in the control of the Rcs system is the negative regulator IgaA, an essential inner membrane protein identified in *S.* Typhimurium as a factor that reduces the growth rate of this pathogen following the entry into eukaryotic cells [3,8]. Previously, an IgaA ortholog of *Proteus mirabilis* named UmoB, was shown to control flagellar synthesis and swarming [9], functions controlled by the Rcs system. IgaA has five transmembrane domains with three defined cytoplasmic regions and a large periplasmic loop. The mechanism by which IgaA represses Rcs signaling is under intense investigation. An interaction between the outer membrane protein RcsF and the periplasmic region of IgaA was proposed to relieve its repression over RcsC/RcsD [10]. More recent studies have reported interactions between outer membrane proteins and RcsF, modulating the RcsF-IgaA interaction [11,12], and, between RcsD and IgaA [6,7], this latter involving periplasmic and cytoplasmic domains [7].

Besides *E. coli* and *S.* Typhimurium, IgaA orthologs have been studied in other *Enterobacterales* genera, mainly in relation to virulence given the control that this protein exerts over the Rcs system, required for expression of flagella and capsules among other virulence-related factors [1,2]. Ortholog members of the IgaA family characterized to date include UmoB of *Proteus miriabilis* [9] and GumB of *Serratia marcenses* [13]. In contrast to *S.* Typhimurium and *E. coli* [3,5–7], the lack of the IgaA orthologs UmoB and GumB does not compromise viability in *P. mirabilis* or *S. marcenses* [9,13]. These findings indicate that, although conserving its master role as repressor of the Rcs system, IgaA may have evolved differently within the *Enterobacterales* order. This variability probably relies on factors related to the lifestyle or the environmental niche that is occupied by the bacterium, resulting in different signals perceived by the IgaA-Rcs axis. Consistent with this hypothesis, studies performed in *S. marcenses* demonstrated that a null *gumB* mutation could be complemented by expressing IgaA of *S.* Typhimurium or *E. coli* whereas expression of a predicted IgaA ortholog from *Klebsiella pneumoniae*, KumO, restored wild type colony morphology but not pigmentation [13]. Differences in the capacity of RcsC from pathogenic *E. coli*, *S. enterica* or *Yersinia pestis* to complement an *rcsC* mutation in *E. coli* K-12 were also reported [14].

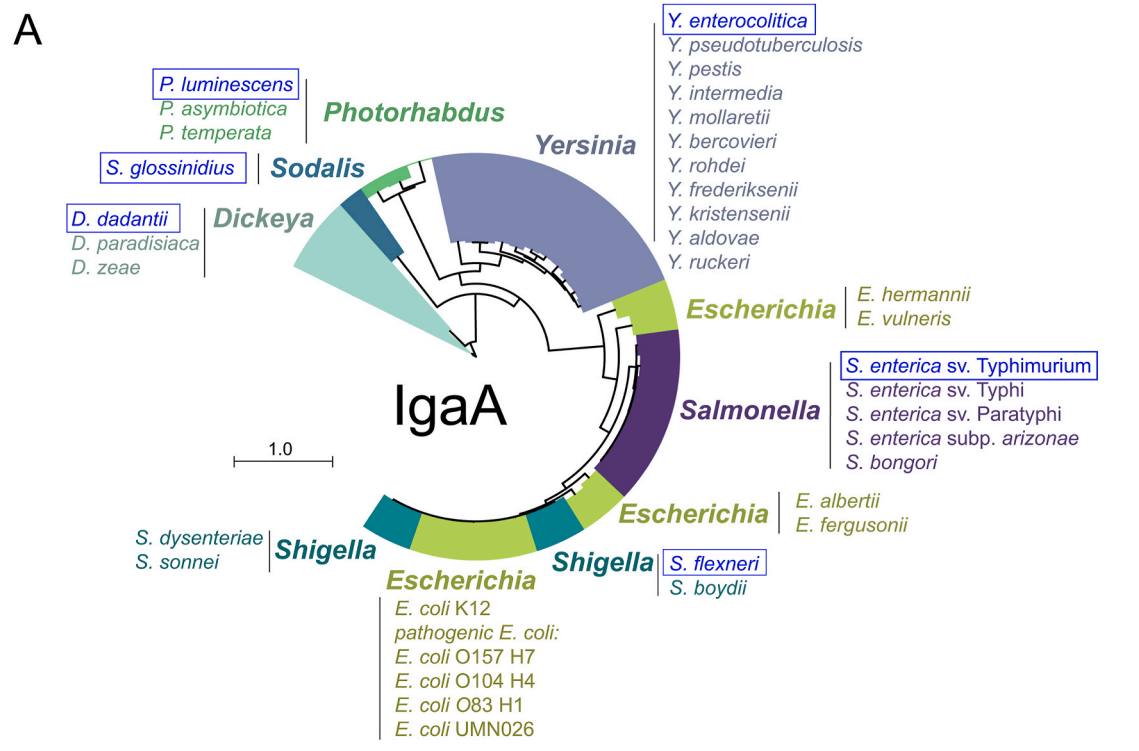
In this study, we performed a phylogenetic analysis of IgaA and Rcs proteins in *Enterobacterales* genera with complete genome information available in databases. This was followed by expression of IgaA from several *Enterobacterales* genera in *S.* Typhimurium to interrogate, via a lethality assay, for their capacity to replace endogenous IgaA in its role as Rcs repressor. The complementation, obtained for only a few of the heterologous IgaA expressed, support divergence in the evolution of IgaA within the *Enterobacterales* order. Modelling and structural analyses of the IgaA tested allowed us to map conserved and varied positions relevant for Rcs repression and to assign a functional role to a previously unnoticed partially-closed small β-barrel (SBB) domain [15] related to OB (oligonucleotide/oligosaccharide binding motif) fold domain [16]. One of these SBB domains is formed by residues of two cytoplasmic regions of IgaA, cyt-1 and cyt-2.

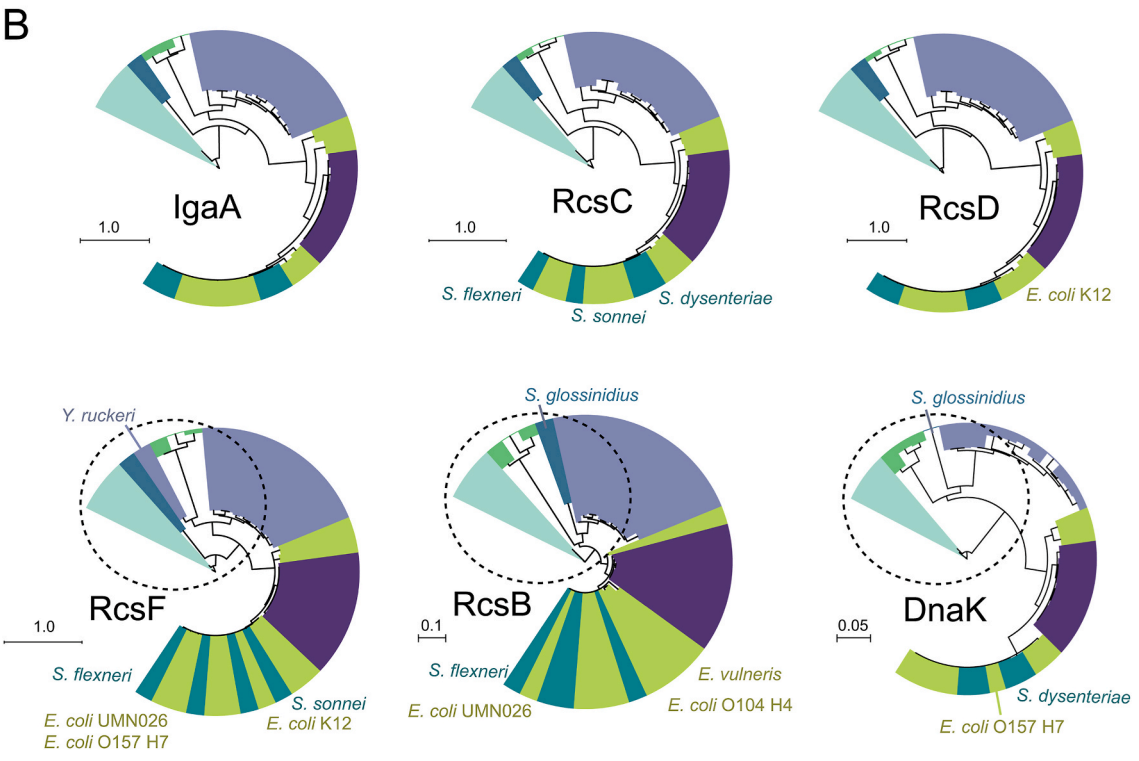
2. Results

2.1. Phylogenetic analysis of IgaA and Rcs proteins within the Enterobacterales order

Although IgaA orthologs have been characterized in some *Enterobacterales* genera such as *Salmonella, Escherichia, Proteus* and *Serratia*, to our knowledge no systematic analysis of the distribution and evolution of the membrane components of the Rcs system and its repressor IgaA has been reported to date. Using as source the "Bacterial and Viral Bioinformatics Resource Center" (BV-BRC) (https://www.bv-brc.org/), we identified IgaA, RcsC, RcsD and RcsF orthologs by BLASTp in most of the *Enterobacterales* genera, except for the obligate endosymbionts *Buchnera aphidicola* and *Wigglesworthia glossinidia,* which lack the entire set of proteins (Tables S1-S2). Interestingly, other bacteria show loss or truncation of specific components. Thus, all isolates of *Y. pestis* with complete genome sequenced encode a truncated RcsD as well as some *S. enterica* subp. *arizonae* isolates (Table S1). Truncation of RcsD in *Y. pestis* was already inferred in an early study [17]. Other interesting cases of bacteria with complete genome sequenced are *Budvicia aquatica*, which lacks RcsB, RcsC and RcsD but encodes full-length versions of IgaA and RcsF; *Pleisomonas shigelloides*, which lacks IgaA, RcsB, and RcsD, but encodes full-length RcsF and a truncated RcsC; and the insect symbiont *Arsnophonus nasoniae*, which has a genome encoding a truncated RcsF (Table S1).

We next performed a phylogenetic analysis of the IgaA/Rcs proteins encoded by 114 complete genomes deposited in the BV-BRC database covering all known *Enterobacterales* genera (Tables S1-S2). This analysis revealed more similarity in the phylogenetic trees obtained for RcsC and RcsD in comparison to those of IgaA, RcsF and RcsB (Figure S1). This finding suggests that the interaction in the phosphorelay known for RcsC and RcsD could have been the origin of the IgaA/Rcs regulatory axis. With this information, we re-analyzed the phylogenetic data of IgaA and Rcs proteins with 40 species from a reduced number of *Enterobacterales* genera representing different lifestyles and occupying distinct ecological niches: *Shigella, Dickeya*, *Yersinia*, *Photorhabdus* and *Sodalis* (Table S3). These genera are classified in different families of the *Enterobacterales* order: *Shigella* (*Enterobacteriaceae*); *Dickeya* (*Pectobacteriaceae*); *Yersinia* (*Yersiniaceae*), *Photorhadbus* (*Morganellaceae*), and *Sodalis* (*Bruguierivoracaceae*) [18]. Importantly, the phylogenetic tree for IgaA of these representative genera, including species of the genus *Escherichia* (to which *Shigella* and *Salmonella* are close phylogenetically), showed a marked overlap with those of RcsC and RcsD (Fig. 1A and B). This feature was however not observed for RcsF, RcsB or an unrelated protein as the chaperone DnaK (Fig. 1B). Taken together, these phylogenetic analyses indicated that in bacteria of





(caption on next page)

Fig. 1. Phylogeny of IgaA and proteins of the Rcs system in selected genera of the *Enterobacterales* order. (A) Phylogenetic tree obtained for IgaA after analyzing a total of 38 representative species within families with different lifestyles belonging to the *Enterobacterales* order: *Salmonella, Escherichia* and *Shigella* (family *Enterobacteriaceae*); *Dickeya* (*Pectobacteriaceae*); *Yersinia* (*Yersiniaceae*), *Photorhadbus* (*Morganellaceae*); and *Sodalis* (*Bruguierivoracaceae*). In blue boxes are highlighted the species for which IgaA variants were cloned and expressed in the functional assays; (B) Phylogenetic trees obtained from representative species for IgaA, RcsC, RcsD, RcsB, RcsF and DnaK, the latter an unrelated protein chosen as control. Protein trees for each set were built using the BV-BRC Gen Tree tool with no-end trimming or gappy sequences removal and using RAxML algorithm with LG evolutionary model. Trees were visualized using iTOL version 6.5.7. Bars indicate phylogenetic distance.

the *Enterobacterales* order, IgaA could have integrated into the Rcs regulatory network primarily by interacting with RcsC/RcsD.

2.2. Expression of heterologous IgaA proteins in S. Typhimurium to repress the Rcs system

Based on the phylogenetic data (Fig. 1A and B), we cloned orthologs of representative species with different lifestyles to test their capacity to replace endogenous IgaA and, therefore, to repress the *S.* Typhimurium Rcs system. To this aim, *igaA* was cloned from the following bacterial genera and species: the plant pathogen *D. dadantii* (formerly *Erwinia chrysanthemi*), the animal pathogens *S. flexneri* and *Y. enterocolitica,* and, the endosymbionts *S. glossinidius* and *P. luminescens.* The rationale for testing function of IgaA from *S. glossinidius* (tsetse fly -*Glossina* spp.- endosymbiont) and *P. luminescens* (nematode endosymbiont) was the intriguing loss of IgaA together with the entire set of Rcs proteins in obligate endosymbionts like *B. aphidicola* or *W. glossinidia* (see Table S1). These observations suggested that IgaA function might become dispensable once bacteria adapt to the invariant and stable intracellular niche of host cells. The IgaA variants were engineered with a Myc-tag in the C-terminus to monitor their production and were expressed from pBAD24, a plasmid bearing a promoter that is repressed with glucose and inducible with arabinose [19]. Immunoassays with anti-Myc antibody confirmed the expression of these heterologous IgaA variants expressed in *S.* Typhimurium (Fig. 2A). The lack of mutations in the respective sequences of these heterologous IgaA variants was confirmed by sequencing.

Next, we took advantage of the essentiality of IgaA in *S.* Typhimurium to test whether the ectopically expressed heterologous IgaA could replace endogenous IgaA. To that purpose, we used P22 bacteriophage to transduce a null Δ*igaA2*::KXX allele [3]. This allele was transduced to *S.* Typhimurium strains bearing the respective IgaA variant-expressing plasmid in the absence/presence of the inducer (0.2% arabinose) (Fig. 2B). An additional positive control consisting of pBAD24 expressing the *S*. Typhimurium IgaA, was included.

Transductants bearing the Δ*igaA2*::KXX allele following induction with arabinose were obtained in *S.* Typhimurium strains expressing IgaA of *S. flexneri* and *D. dadantii* (Fig. 2C). This result showed that IgaA of *Y. enterocolitica*, *P. luminescens*, and *S. glossinidius*, which represent the *Yersiniaceae*, *Morganellaceae* and *Bruguierivoracaceae* families respectively, cannot repress the Rcs system when produced in *S.* Typhimurium, which belongs to the *Enterobacteriaceae* family. Interestingly, even in the presence of 2% glucose, some transductants bearing the Δ*igaA2*::KXX were obtained when producing ectopically IgaA of *S.* Typhimurium. Most of these transductants exhibited a mucoid phenotype (Fig. 2D), denoting overactivation of the Rcs system [3]. These mucoid transductants lacking IgaA were reported to arise spontaneously in *S.* Typhimurium by compensatory mutations that decrease levels of RscC or RcsD [5]. The reason for obtaining these mucoid clones in glucose-containing plates only in the strain bearing the plasmid with *igaA* from *S.* Typhimurium is at present unknown. Minute amounts of highly active endogenous IgaA produced from the plasmid even in the absence of inducer could be sufficient to confer the survival phenotype. Interestingly, we observed that transductants expressing IgaA from *D. dadantii* were also mucoid (Fig. 2D). In this case, these mucoid transductants were however obtained exclusively in 0.2% arabinose-containing plates, i.e. in the presence of high levels of *D. dadantii* IgaA (Fig. 2A). These findings indicated that IgaA from *D. dadantii* represses but only partially the *Salmonella* Rcs system.

The data obtained following the expression in *S.* Typhimurium of heterologous IgaA, in which complementation was detected only in defined cases, supported variability in IgaA within the *Enterobacterales* order. We reasoned this variability could rely on structural differences that limit interactions with the endogenous Rcs proteins from *S.* Typhimurium and/or the recognition of yet unknown signaling molecules.

2.3. Modelled structure of IgaA from Enterobacterales unravel conserved positions among families of the order and the presence of three small β-barrel (SBB) domains

Based on the phenotypic data obtained with the different IgaA expressed in *S.* Typhimurium (Fig. 2), we searched for regions or residues that could vary in the non-complementing versus complementing IgaA variants by a sequence-based structural approach (Fig. 3). To that purpose, we used the publicly available tool AlphaFold (https://alphafold.ebi.ac.uk/) [20,21] to predict the structure of the different IgaA expressed in *S.* Typhimurium. As previously proposed [4], the overall structure predicted by AlphaFold (with high and confident accuracy per residue, pLDDT >90–70%, for the majority of the sequence) comprises five transmembrane helices (TM), three defined cytoplasmic regions (cyt-1, cyt-2, cyt-3) and a periplasmic region (Fig. 4A and B). A first unexpected finding resulting from modelling of IgaA was the presence of three β-rich architectures forming partially-closed small β-barrel (SBB) domains [15], which are related to OB (oligonucleotide/oligosaccharide binding motif) fold domains [16]. The two SSB domains in cytoplasmic regions were termed SBB-1 and SBB-2 while the periplasmic domain was named SBB-3 (Fig. 4A and B). SBB-1, SBB-2 and SBB-3 share similarity to the fold present in a domain of the IgaA ortholog from *E. coli,* YrfF (residues 36–154, PDB: 4UZM) [22], showing root-mean-square deviation (RMSD) values of 1.5 Å for 101 residues in SBB-1; 2.1 Å for 73 residues for SBB-2; and, 1.6 Å for 105 residues for SBB-3 (Figure S2). Each of the three SBB domains present in IgaA are formed by two β-sheets (β-sheet I and β-sheet II) where the first β-strand is connected to an α-helix and the second β-strand is curved being shared by both β-sheets (Fig. 4C). In SSB-3,

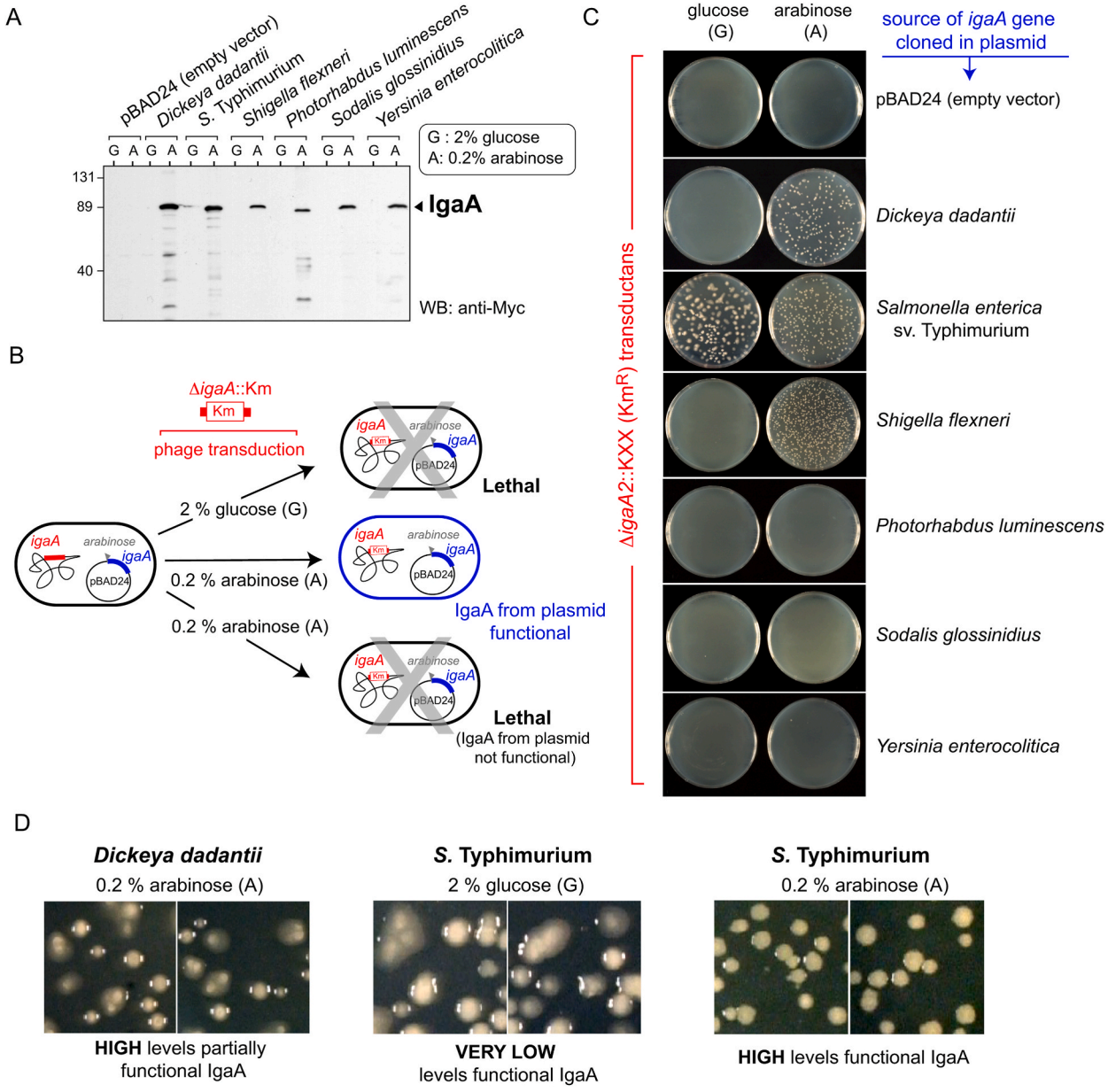


Fig. 2. Functional assay of IgaA from several *Enterobacterales* genera expressed in *S.* Typhimurium. (A) *igaA* wild type alleles from the indicated *Enterobacterales* genera and species were cloned in pBAD24 plasmid [19], a vector bearing an arabinose-induced promoter to drive expression of the insert. These *igaA* alleles were cloned as tagged versions with a Myc epitope located at the 3′ end. Shown is a Western assay of total extracts of the indicated *S.* Typhimurium strains grown either in the presence of glucose, a repressor of the pBAD promoter, or arabinose as inducer. Numbers in the left side refer to the molecular weight markers (in kDa); (B) Scheme of the experimental design used to test function of IgaA from distinct *Enterobacterales* genera in *S.* Typhimurium based on suppression of lethality due to the absence of endogenous IgaA; (C) Result of the phage transduction of an Δ*igaA2*::KXX allele in the presence of glucose or arabinose. Transductants were obtained only when expressing IgaA from *S.* Typhimurium and from the genera *Shigella* and *Dickeya* (*Enterobacteriaceae* and *Pectobacteriaceae* families). The transduction assay was repeated in two independent biological replicates with identical results; (D) Phenotype of *S.* Typhimurium transductants expressing IgaA from *Dickeya* or *Salmonella*. The colonies appear mucoid when expressing IgaA of *Dickeya*, which denotes partial complementation and lack of a completely repressed Rcs system [4]. In the case of the IgaA from *Salmonella*, the production in the presence of glucose of minute amounts of this functional IgaA may confer survival with a partially activated Rcs system and, as a result, the mucoid phenotype.

another β-strand (β19) is however curved and shared by the two β-sheets contributing to the β-barrel shape (Fig. 4C). Remarkably, SSB-1 is located entirely in the first cytoplasmic region cyt-1 whereas SSB-2 is "hybrid". Thus, SSB-2 is built by residues of region cyt-1 that comprise the β7 strand (sequence 183-ELLNIRQ-189 in IgaA of *S.* Typhimurium) that is connected to α7 to form the "SBB-connector" as well as most of region cyt-2 (Fig. 4C).

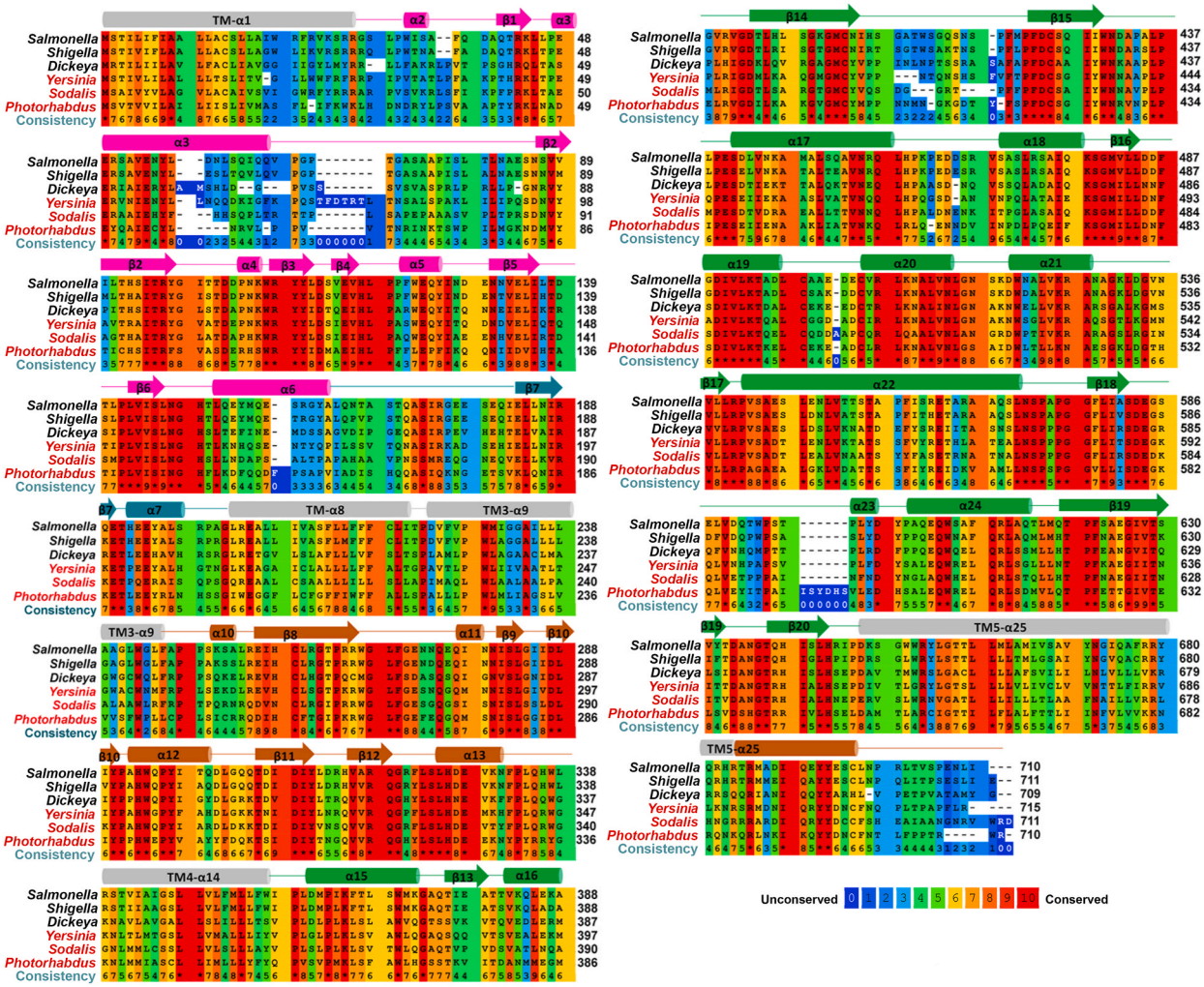


Fig. 3. Alignment of IgaA sequences from the selected *Enterobacterales* genera that were expressed in *S.* Typhimurium. IgaA sequences from the *Salmonella, Shigella, Dickeya, Yersinia*, *Sodalis* and *Photorhadbus* genera were aligned with the PRALINE server. Conservation of the residues is shown according to the color key. The secondary structure is shown on the sequence where the structural elements are colored according to Fig. 4B (helix α in TMs in gray, SSB-1 domain in magenta, SBB connector in cyan, SBB-2 in orange and the periplasmic region having SBB-3 in green).

We then modelled IgaA variants from the diverse *Enterobacterales* genera used in our functional study. Sequence identity with *Salmonella* is high for *Shigella* (84.1%) but moderate with the rest of the genera tested (49.3% with *Dickeya*, 52.8% with *Yersinia*, 39.3% with *Photorhabdus*, and 47.7% with *Sodalis*). All these IgaA variants conserve the three SBB domains (Figure S3). However, we observed sequence and length variability in several regions between the complementing and non-complementing IgaA variants (Fig. 3). One was the loop between α3 and β2 in the SBB-1 domain, which is rather long in *Yersinia* and shorter in *Photorhabdus* (Figs. 3 and 5A-F), existing also sequence divergence between the region spanning from the C-terminus of α3 and the N-terminus of the loop (Fig. 3). Another variable region was the α6 helix in cyt-1, which is almost absent in *Yersinia* and *Sodalis* whereas it is five residues shorter in *Photorhabdus*, showing sequence variability together with the N-terminal of the SBB connector. Interestingly, the loop connecting α6 with β7 in the SBB connector showed very low confidence (pLDDT <50%) in the modelled structure, indicative of sequence and structural variability in that region (Figs. 3 and 5D-F). Other regions showing sequence and length variability in the non-complementing IgaA variants include: i) the loop between β14-β15 in SBB-3; ii) the C-terminal end of TM-α1 connected to the N-terminal of β1 in SBB-1; and iii) the C-terminal end in cyt-3 displaying a high degree of sequence variability (Fig. 3).

2.4. The IgaA variants tested functionally in S. Typhimurium differ in some interacting sites involving the cyt-1/cyt-2 interface of the SBB-2 domain

We further exploited AlphaFold to search in IgaA for distances ≤4 Å between residues, inferring in this manner interactions putatively relevant for domain stability. Among these interactions, we mapped E180-R265, which involves moderately conserved

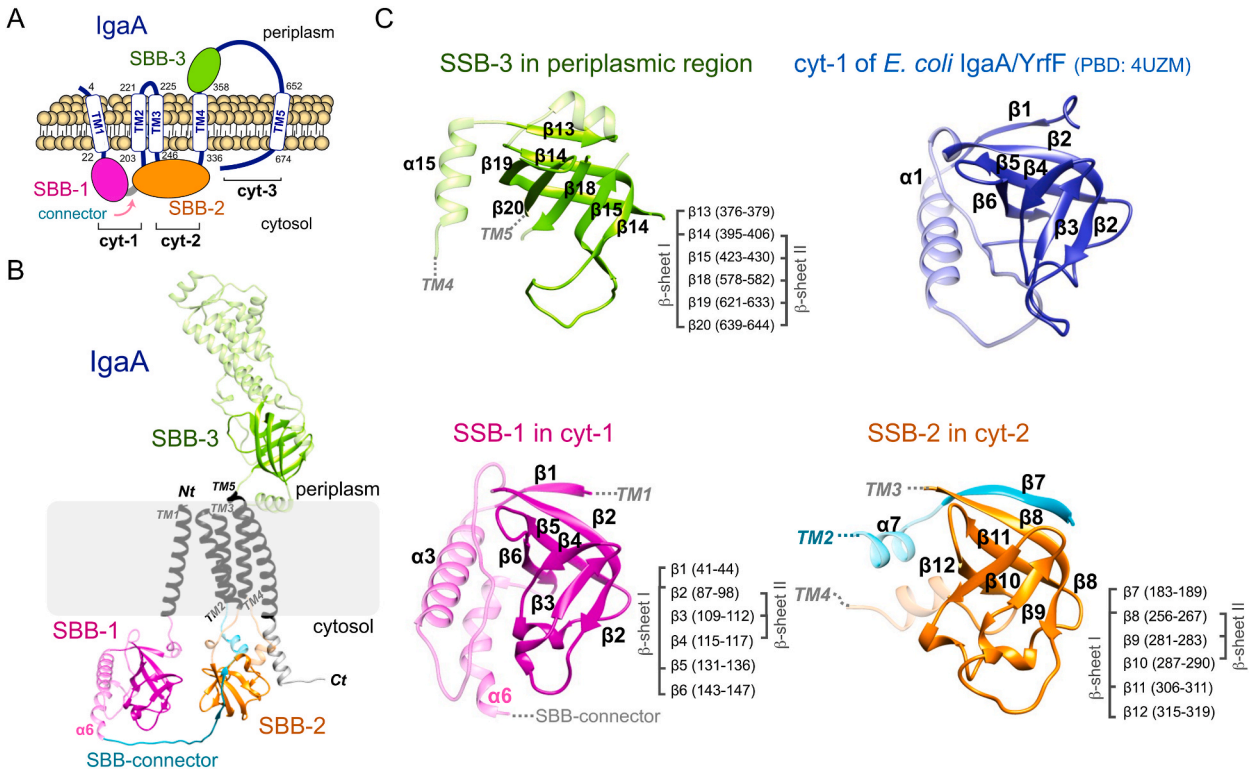


Fig. 4. Modelling of IgaA structure reveals the presence of three SBB domains. (A) Cartoon showing the new domain configuration predicted for IgaA with two cytoplasmic SBB fold domains, with the named SBB-2 composed by residues of regions cyt1 and cyt 2. The C-terminus of region cyt3 is shown close to SBB-2 since at least residue R686 is predicted to interact and stabilize this domain (see Fig. 6). The drawing also shows the presence of a connector between SBB-1 and SBB-2. Numbers indicate the positions of residues at the start and end of each of the five transmembrane domains; (B) IgaA structure predicted by AlphaFold in which the five transmembrane regions (*TM*) are shown in dark gray, SSB-1 domain in magenta, SBB connector in cyan, SBB-2 in orange and the periplasmic SBB-3 in green. *Nt, Ct*: N- and C-termini, respectively; (C) The three SBB domains of *S.* Typhimurium IgaA predicted by AlphaFold and compared to the domain previously reported from a fragment of *E. coli* IgaA (YrfF) [22]. For each SBB domain the residues involved in formation of β-sheet I and β-sheet II, are indicated.

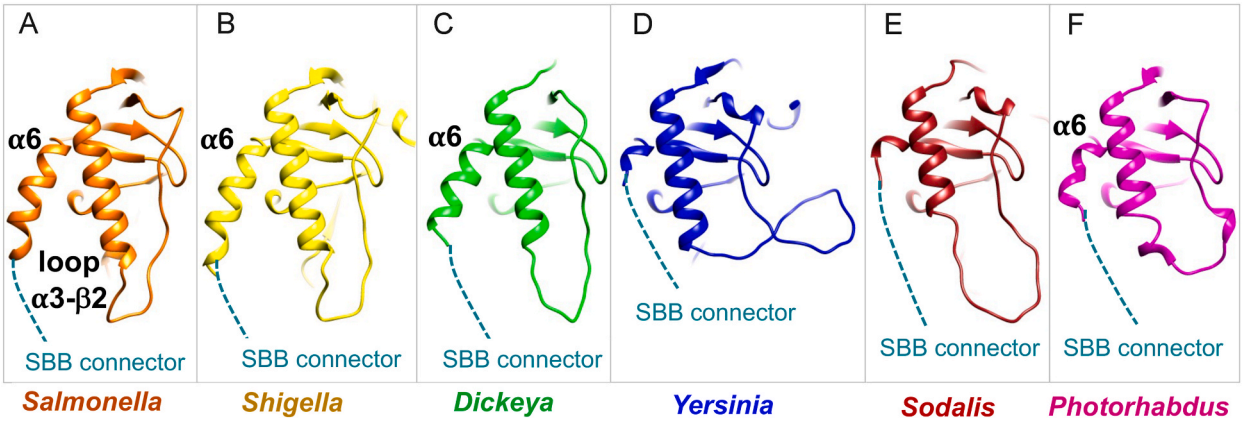


Fig. 5. Structural features of the SBB-1 domain show divergence in the non-complementing IgaA variants. Modelled structure of the SBB-1 domain in cartoon representation showing differences in the length of loop α3-β2 and α6 in the *Enterobacterales* tested in this study: (A) *Salmonella;* (B) *Shigella;* (C) *Dickeya*; (D) *Yersinia*; (E) *Sodalis*; and (F) *Photorhadbus*. The connection of α6 to the SBB connector is shown as a dashed line in cyan.

residues, one present in a stretch acting as SBB-1/SBB-2 connector (E180) to face another SBB-2 residue (R265) (Fig. 6A). Additional interactions involving conserved residues seem to ensure folding of the SBB-2 domain. These include: i) the triad E194-R188-D309, which facilitates proximity between strand β7 of region cyt-1 and SBB-2 residues that are part of the cytoplasmic region cyt-2

(Fig. 6B); and ii) the triad H293-E328-R686, which implicates two residues of cytoplasmic region cyt-2 forming part of SBB-2 and a residue of cytoplasmic region cyt-3 (R686) (Fig. 6B). Interestingly, an R188H mutation, probably disrupting the E194-R188-D309 interaction, was the first spontaneous mutation reported to affect IgaA function *in vivo* [8]. The modelled structure obtained by AlphaFold explains another mutation reported to affect IgaA function *in vivo*, G262R [4], which impairs proper folding of the β8 strand. The putative stabilizing role of R686 (Fig. 6B) would implicate for the first time a functional role for the cyt-3 region. Also, Y694, a conserved residue in all genera, has several hydrophobic interactions with SBB-2. Importantly, these predicted interactions sites are conserved in all IgaA that were expressed in *S.* Typhimurium (Figs. 3 and 6C). This evidence supports the idea of a "hybrid" SBB-2 domain in IgaA formed by residues of cyt-1 and cyt-2 regions, stabilized by residues of cyt-3, which is absolutely required to repress the Rcs system.

Next, we sought to get insights into the structural variations that could explain the lack of complementation exhibited by some of the IgaA expressed in *S.* Typhimurium (Fig. 2). We searched in the AlphaFold structural predictions for residues that were not conserved when comparing those IgaA functional in *S.* Typhimurium (versions of *S. flexneri* and *D. dadantii*, this latter only partially) versus those non-complementing (*Y. enterocolitica*, *P. luminescences* and *S. glossidinius*). We reasoned these divergent residues could be involved in interactions required for proper folding of IgaA in *S.* Typhimurium but probably not existing in IgaA of other *Enterobacterales* genera. H192-P249 was found as a putative interacting site diverging among the two groups of IgaA proteins, complementing versus non-complementing (Fig. 7A–F). In *D. dadantii,* this putative interaction could be established between leucine and proline residues (L191-P248) (Fig. 7C). Interestingly, H192 is in α7, within the SBB-connector, contributing to the "hybrid" SBB-2 domain formed by residues of cyt-1 and cyt-2 regions. H192 is also next to T191, a conserved residue previously described to be important for IgaA function in *S.* Typhimurium [4] and whose role can now be explained by its interaction with the conserved H326 (Fig. 7A). Thus, sequence diversity in the β7-α7 region within the SBB-connector contributing to the "hybrid" SBB2 domain seems to have an important impact in IgaA function not only providing anchoring to the SSB-2 domain but also a structural support. Other two sites differing between these two sets of IgaA proteins were R255-D313 and D287-R314, interaction sites that involve again the SBB-2 domain creating an electrostatic surface that varies largely in the non-complementing IgaA variants from *S. glossidinius* and *P. luminescences* but to a lesser extent in those from *D. dadantii* and *Y. enterocolitica* (Figure S4). One of these interactions, R255-D313 in *S.* Typhimurium, could have been lost in IgaA from *Dickeya dadantii*, which has a threonine residue replacing the charged aspartate residue (R254-T312) (Figure S4). Collectively, these findings reinforce the relevance that the "hybrid" SBB-2 domain has for IgaA function.

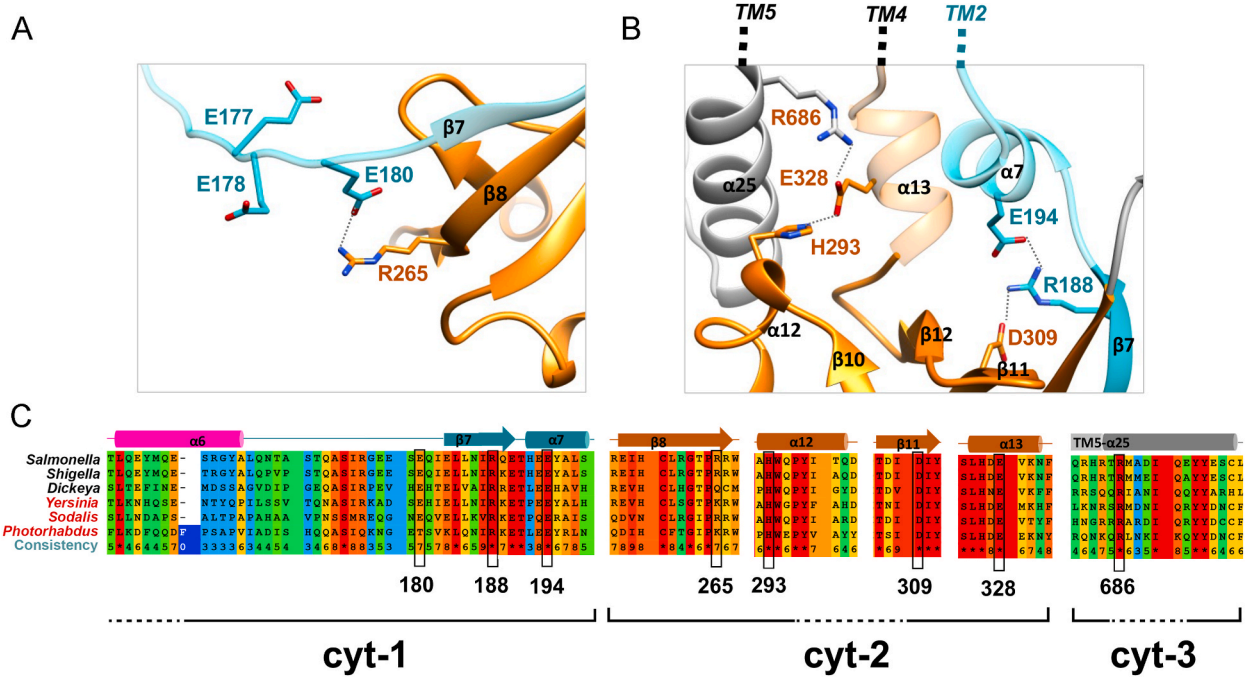


Fig. 6. Conserved interactions in IgaA of diverse *Enterobacterales* genera connecting residues of the cyt-1, cyt-2 and cyt-3 regions. Modelled structure in cartoon representation of *S.* Typhimurium IgaA showing salt bridge interactions conserved in the *Enterobacterales* tested. (A) Residue E180 in the SBB-connector (cyan) belonging to cyt-1 interact with R265 located in the β8 of the SBB-2 domain in cyt-2 (orange); (B) Residues in the β7 (R188) and α7 (E194) within the SBB-connector of cyt-1 (cyan) interact with D309 of the SBB-2 (orange). Residues E328 and H293 in cyt-2 (orange) interact with R686 in cyt3 (gray); (C) Sequence alignment, as shown in Fig. 3, of the structural features shown in panels A and B. Dashed lines between residues indicate interacting distances in the modelled structures.

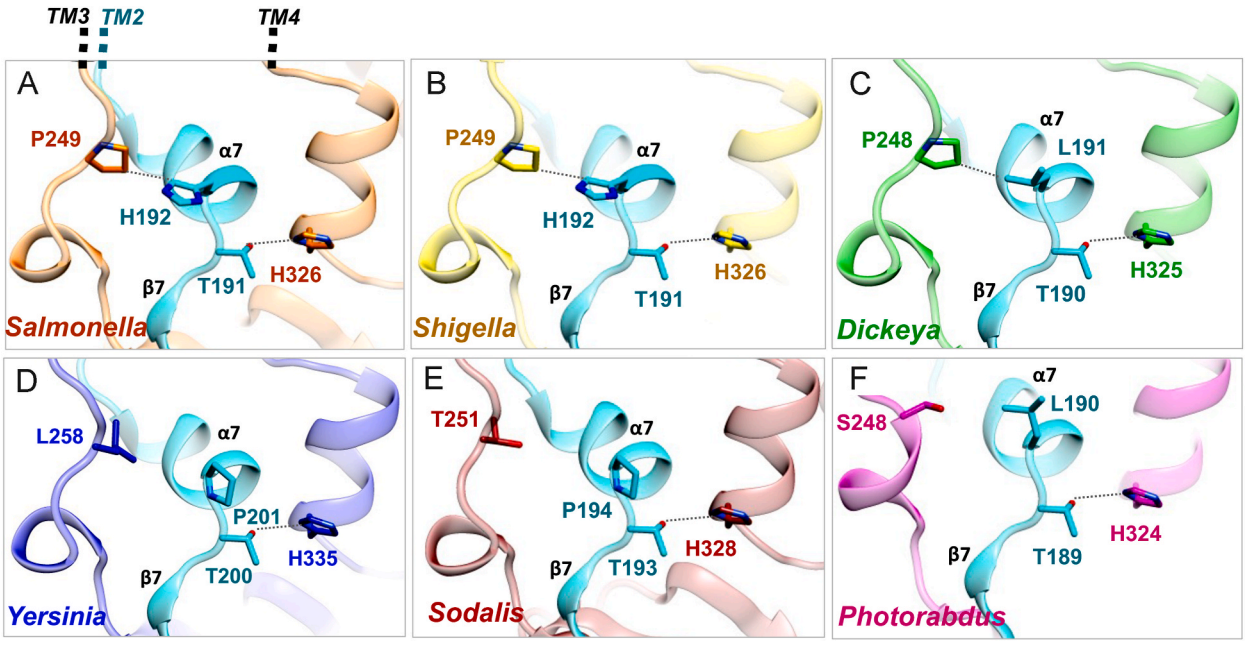


Fig. 7. The SBB-connector provides specific anchoring to the SSB-2 domain as well as a structural support. Modelled structures by cartoon representation of the SBB-2 "hybrid" domain in the *Enterobacterales* tested in this study: (A) *Salmonella*; (B) *Shigella*; (C) *Dickeya*; (D) *Yersinia*; (E) *Sodalis*; and (F) *Photorhabdus*. Residues in the β7-α7 structural feature of the SBB-connector from cyt1 (in cyan) show divergence and conservation to contact the SBB-2. Dashed lines between residues indicate interacting distances in the modelled structures.

2.5. *Structural analyses in RcsF and RcsD of Enterobacterales, two proteins interacting directly with IgaA*

Various studies have shown that IgaA interacts directly with outer membrane lipoprotein RcsF [10,12] and the inner membrane protein RcsD [6,7]. These interactions represent the activated and resting states of the Rcs phosphorelay, respectively. Based on these findings, we reasoned that the IgaA variants that do not replace endogenous IgaA in *S.* Typhimurium could not be functional in that heterologous host due to marked structural differences in either RcsF and/or RcsD impeding these interactions. In this reasoning, we hypothesized that RcsF and RcsD could have co-evolved with IgaA showing sequence and structural variability among *Enterobacterales*. To test this, we aligned RcsF and RcsD sequences of the *Enterobacterales* genera included in our analysis and obtained their modelled structures, as we did for IgaA. All RcsF structures showed a disordered N-terminus (~50 residues) linked to a C-terminal domain modelled with high confidence for all bacterial genera (AlphaFold confidence score pLDDT >90%) given that the experimental structure of *E. coli* RcsF (PDB: 2Y1B) is available. This RcsF region shows high structural similarity (Figure S5A). Noteworthy, the unstructured N-terminus shows high sequence variability among species (Figure S5B), albeit RcsF of *Salmonella* is very similar to that of *Shigella* (88.5% similarity, overall sequence 90.3%). Interestingly, RcsF of *Salmonella* is more similar to *Yersinia* RcsF (51.9%, all sequence 67.9%) than to the *Dickeya* ortholog (43.8%, all sequence 54.5%) despite IgaA of *Yersinia* being not functional in *S.* Typhimurium and, vice versa, the IgaA of *Dickeya* partially functional (see Fig. 2). At the C-terminus of RcsF, there is low sequence variability, which can be mapped in the loop connecting β3-β4 (Figure S5B).

For RcsD, aligned sequences were also compared with modelled structures (obtained with high and confident accuracy per residue, pLDDT >90–70%, for most of the sequence) (Figures S6 and S7). Sequence comparison indicated that *Salmonella* shares high overall identity with *Shigella* (84.6%) and similar identity values with *Dickeya* (43.7%), *Yersinia* (47.1%) and *Sodalis* (42.8%), being more distant from *Photorhabdus* (35.4%). Sequence diversity is reflected in the length of the N-terminus that connects the first transmembrane segment (TM-α1), which is longer for *Salmonella* and *Shigella* (Figure S6). The periplasmic region of RcsD, which interacts with IgaA *in vivo* [7] contains two variable regions, one comprising the loop α2-β1 and another more variable in the loop between β10-β11 (Figure S6). However, the periplasmic region shows structural similarity in all RcsD variants, with a fold that can be ascribed to a double-PDC domain [23] as denoted by a search through the DALI server, which shows structural flexibility at the sequence variable loop β10-β11 (Figure S7A). We also used for this analysis the Foldseek server [24], which predicted structural similarities of the RcsD periplasmic domain with double-Cache chemoreceptor ligand-binding domains. This finding is consistent with the proposal of PDC domains belonging to the Cache superfamily [25]. In the cytoplasmic region, RcsD shows structural similarity among bacterial genera at the domain level: the PAS-like and His-phosphotransfer domains (Hpt) (Figure S7B) as well as at the HATPase-like and ABL domains (Figure S7C). However, each modelled RcsD shows some differences in the 3-D spatial orientation of the PAS-like and Hpt domains due to the flexibility of the loops that connect those domains (Figure S7C). In relation to the sequence, there is variability at the connection between the transmembrane helix (TM-α7) and the PAS-like domain, including the α8, which locate close to the membrane being able to interact with cytoplasmic regions of IgaA, as suggested by *in vivo* data [7]. The connection of TM-α7 with α8 is

five residues longer in *Salmonella* and *Shigella,* but not in *Dickeya*, which could contribute to explain the partial function exhibited by IgaA from *Dickeya* when expressed in *S.* Typhimurium. Another interesting feature of RcsD is helix α11, which shows in the models two different spatial orientations respect to α12. The similar α11-α12 orientation predicted in *Salmonella* and *Shigella* differs from that predicted for the rest of RcsD proteins analyzed (Figure S7C). Curiously, the region comprising the C-terminal end of α11 and the N-terminus of α12 shows high sequence variability, being eight residues shorter for *Salmonella* and *Shigella* (Figure S6). Although this region may locate far away to directly interact with IgaA, it might contribute to downstream signaling. Overall, the modelled structures of RcsF and RcsD did not show obvious structural changes among the bacterial genera examined.

3. Discussion

In this study, we examined from an evolutionary perspective the repression that the inner membrane protein IgaA exerts over the Rcs system. Our strategy involving expression of different IgaA from genera representing distinct families of the *Enterobacterales* order was insightful to uncover regions critical for function and to establish the degree of conservation or divergence of this protein within this order. The functional data obtained together with the *in silico* analyses of the structures predicted by AlphaFold all point to a previously unnoticed presence of an SBB-2 domain, related to the OB fold domain [16]. SBB-2 is formed by most of cytoplasmic region cyt-2, a β-strand (β7) supplied by the C-terminal residues of cyt-1 creating a clear cyt-1/cyt-2 interface that is stabilized by defined residues of the cyt-3 region like R686. A previous study reported the presence of an OB fold domain in a fragment of *E. coli* IgaA (formerly YrfF) encompassing residues 37–154, all positioned in region cyt-1 [22], and which corresponds to what we termed SSB-1 in IgaA of *S.* Typhimurium. This consistency provides reliability to the AlphaFold prediction extended to full-length IgaA showing the presence of two additional SBB domains.

SSB domains differ from β-barrels of integral membrane proteins in having much less β-sheets, in the order of five or six, and of shorter length [15]. SSB domains are also of small size, generally encompassing $\leq$100 residues. Other striking features of the SSB domains are the orthogonal package of the β-sheets and the presence of short invariant α-helices connecting such β-sheets [15]. Remarkably, a distant OB fold domain was previously assigned to region cyt-1 of *E. coli* IgaA [22] whereas our structural analysis reveal that this domain should be more strictly considered as SSB (Figure S2). Other studies involving periplasmic proteins have predicted OB-fold domains, as it is the case of YdeI and YgiW of *S.* Typhimurium, two proteins that interact with outer membrane porins and play a role in antimicrobial resistance, stress and virulence [26,27]. Intriguingly, YdeI is regulated positively by the Rcs system and important for *Salmonella* persistent infections [28]. Considering our findings in IgaA, the presence of OB fold domains in these proteins should be revisited to dissect whether SSB domains, besides its wide presence in proteins related to DNA, RNA and protein metabolism, could also be common in proteins involved in stabilizing envelope architecture. To our knowledge, IgaA represents the first example of protein possessing SBB domains facing at both the outer and inner leaflets of the plasma membrane, opening the possibility of SBB domains being also involved in signal transduction.

In a recent study, Wall et al. provided compelling evidence in *E. coli* for IgaA interacting with RcsD at both their periplasmic and cytoplasmic regions [7]. These authors generated IgaA constructs deficient in either cyt-1 or cyt-2 or the periplasmic region, with all of them exhibiting loss-of-function. Of major interest, the cyt-1 deletion encompassed residues 36–181 [7], therefore making possible the presence an intact hybrid SBB-2 domain folded with the contribution of the β7 strand (see Fig. 4). In the context of that study and our analyses, it can be hypothesized that the three SBB domains act in concert to transduce signal from the periplasm to the cytoplasm, being the basis of the mechanism by which IgaA could repress the Rcs system. The essential role played by the cytoplasmic regions of IgaA, as pointed by Wall et al. [7] and further inferred from our study, also agrees with the data of Hussein et al. [6]. Despite restoring viability in *E. coli* transiently depleted of IgaA but expressing IgaA cyt-1 and cyt-2 fragments, these authors were unable to complement a Δ*igaA*::Km null allele with such individual fragments [6].

The most accepted model of communication between IgaA and the Rcs system is a control at the periplasmic side exerted by an IgaA-RcsF interaction [10,12], which is relieved upon stress, an alteration that is supposed to be translated to cytoplasmic regions of IgaA. This model, however, contrasts at some extent with our functional studies with IgaA from different *Enterobacterales* families and the structural *in silico* analyses presented here. Thus, in the primary sequences of the IgaA tested functionally and, basically with the only exception of the region linking β14-β15 sheets, we did not find other motifs in the periplasmic region diverging between the complementing IgaA variants and those non-complementing from *Y. enterocolitica*, *S. glossidinius* and *P. luminescences* (Fig. 3). By contrast, more variable sequences and structures were mapped in the cyt-1, cyt-2 and cyt-3, regions, including the α3-β2 junction, α6 helix, the SBB-connector α6-β7, P249, R255, D287, D313, R314, and the C-terminus (Figs. 3, 5–7, and S4). Although we are uncertain for the reason of detecting more divergence in the cytoplasmic versus the periplasmic regions, the structural requirements in the periplasm might be less strict compared to cyt-1 and cyt-2. Supporting this idea is the capacity to repress the Rcs system shown by a recombinant IgaA in which its periplasmic domain was replaced for the equivalent region of MalF [6]. Moreover, among the spontaneous mutations selected in *S.* Typhimurium for partial loss of function in IgaA, those mapping in cyt-1 and cyt-2 like R188H, T191P and G262R have a more profound phenotypic effect regarding the loss of repression over the Rcs system and attenuation of virulence [4]. It is remarkable that R188H, the original mutation selected in *S.* Typhimurium for a phenotype of increased growth inside eukaryotic cells that was used to coin this protein [8], involves a residue mapping in the β7 strand that links cyt-1 to cyt-2 to conform the hybrid SSB-2 domain (Figs. 4 and 6).

Overall, our work represents a novel approach that has exploited evolutionary data to unveil the functional relevance of the cytoplasmic regions of IgaA that is sustained by *in silico* analyses. Moreover, our approach complements recent work by other groups [6,7] and provides a detail map of variable and conserved sites as well as structures directly related to function. These differences may have established structural constrains in the interaction of IgaA with RcsF or RcsD for defined *Enterobacterales* genera, which might not

occur when the protein is expressed in *S.* Typhimurium. Unfortunately, the modelled structures of RcsF and RcsD did not show obvious differences. Instead, there are sequence constrains among the *Enterobacterales* genera tested that do not apply to *Shigella,* phylogenetically closer to *Salmonella* than to the other genera analyzed in this work. The loss of a few structural interactions of those outlined in this work may be probably behind the partial function exhibited by IgaA of *Dickeya* when expressed in *S.* Typhimurium. Future experimental structures for the protein complexes will provide information about the events of protein-protein recognition as well as the associated conformational changes. Moreover, it cannot be ruled out that IgaA from *Y. enterocolitica*, *P. luminescences* and *S. glossidinius* may not function in *S.* Typhimurium due to the absence in this organism of a putative signaling molecule to which these bacteria respond in the natural environments that they inhabit. These signaling molecules could be certainly a source of variability determining the interaction of IgaA with RcsD and/or RcsF.

4. Materials and methods

4.1. Bacterial strains and growth conditions

All *S. enterica* serovar Typhimurium strains used in this study derive from the virulent strain SL1344 [29]. *E. coli* strains used for cloning are derivate of strain DH10B [30]. The bacterial strains and plasmids used in this study are listed in Table S4 of the Supplementary Material. Bacteria were cultured in Luria-Bertani (LB) broth at 37 °C in shaking (150 rpm) conditions. To repress/induce expression of *igaA* gene cloned in plasmid, bacteria were grown in LB containing 2% glucose or 0.2% arabinose, respectively, during 2 h. The sugars were added to the medium at mid exponential phase (OD_{600}~0.5). When required, the medium was supplemented with ampicillin (50 μg/mL), or kanamycin (30 μg/mL).

4.2. Cloning of igaA from different Enterobacterales genera

The *igaA* gene from selected *Enterobacterales* genera were cloned using genomic DNA from these sources: *S.* Typhimurium strain SL1344 [29]; *S. flexneri* strain M90T [31]; *D. dadantii* strain 3937 [32]; *S. glossinidius* strain *morsitans* [33]; *Y. enterocolitica* strain 8081 [34]; and, *P. luminescens* strain TTO1 [35]. *igaA* was PCR-amplified with the oligonucleotides listed in Table S5 of Supplementary Material using TaKaRa DNA polymerase. Oligonucleotides were designed with NheI/HindIII sites to allow cloning of PCR-products in pGEM-T cloning vector (Promega Corporation, Madison, WI) and the sequence of the Myc epitope (5′- CAG ATC CTC TTC TGA GAT GAG TTT TTG TTG -3′) in one of the primers (Table S5). From the pGEM-T plasmid, inserts were transferred to the expression vector pBAD24 [19] for inducing expression with 0.2% arabinose. *igaA* from *S.* Typhimurium, *D. dadantii*, *S. glossinidius* and *Y. enterocolitica* were cloned in pBAD24 using NheI/HindIII enzymes. *igaA* of *S. flexneri* and *P. luminescens* were cloned in pBAD24 using XbaI/EcoRI sites to insert a SpeI/EcoRI product obtained from pGEM-T. In all cloning procedures, *E. coli* DH10B [30] was used a host. The series of pBAD24 plasmids with the different *igaA* variants were finally used to transform *S.* Typhimurium strain SL1344. The complementation assay with the entire list of *igaA* variants was repeated in two independent biological replicates with similar results.

4.3. Phage transduction and viability assays

IgaA is an essential protein in *S.* Typhimurium in a Rcs^+ genetic background [3]. To monitor the capacity of IgaA from other *Enterobacterales* genera to replace the endogenous protein, a P22H105/1 *int 201* transducing phage lysate obtained from *S.* Typhimurium strain SV4441 (*igaA2*::KXX *rcsC*::MudA), was used (gift from Prof. J. Casadesús, University of Sevilla, Spain). This lysate was employed to transduce the null *igaA2*::KXX allele and select for transductants resistant to kanamycin, only obtained when the exogenous IgaA produced from plasmid complemented endogenous IgaA. The appearance of transductants on plates containing 30 μg/mL kanamycin provided a direct proof for the activity of exogenous IgaA in *S.* Typhimurium.

4.4. Immunoassays for detection of IgaA-Myc recombinant proteins

IgaA-Myc of different *Enterobacterales* genera produced in *S.* Typhimurium were detected with mouse monoclonal anti-Myc-tag antibody (clone #9B11, Cell Signaling Technology, Danvers, MA) at 1:1000 dilution and secondary goat anti-mouse peroxidase at 1:2000 dilution. The amount of sample loaded per well corresponded to ~5×10^7 bacteria. Preparation of protein extracts, electrophoresis and immunoassay conditions were as described [4].

4.5. Phylogenetic analysis of IgaA and Rcs proteins

A total of 114 genomes belonging to *Enterobacterales* order bacteria were retrieved from "The Bacterial and Viral Bioinformatics Resource Center (BV-BRC)" (https://www.bv-brc.org/) based on completeness, total contig number and the consideration of being "reference" or "representative" genomes. These genomes were interrogated using the BLASTp tool using as queries the following protein sequences from *S. enterica* serovar Typhimurium strain SL1344 (in brackets Uniprot accession number): IgaA (E1WIS2), RscB (A0A0H3NNS0), RcsC (A0A0H3NF64), RscD (A0A0H3NDQ0), RcsF (A0A0H3N9F1) and DnaK (A0A0H3NCG3). The assignment for presence/absence considered BLASTp scores and the BV-BRC annotations. For those entries lacking functional annotation, a search of orthologs was performed using eggNOG 5.0 sequence search web service [36] in order to confirm that they belonged to the predicted orthologous group.

To compose the phylogenetic trees, two sets of amino acids sequences of IgaA, RcsB, RcsC, RcsD, RcsF and DnaK were built for each protein; one containing all the sequences predicted to be homologous found in the 114 proteomes and, the second with only with those belonging to some specific genera: *Salmonella, Shigella, Escherichia, Dickeya, Sodalis*, *Yersinia* and *Photorhabdus*. Protein trees for each set were built using the BV-BRC GenTree tool with no end trimming or gappy sequences removal and using the "Randomized Axelerated Maximum Likelihood" (RAxML) algorithm with the LG general amino acid replacement matrix [37]. Trees were visualized using iTOL version 6.5.7 [38].

4.6. Structural analyses on models obtained by AlphaFold

The structure of *S.* Typhimurium IgaA was obtained from the AlphaFold database [21], code AF-E1WIS2-F1. The modelled structures of the diverse IgaA tested in this study, from *Shigella*, *Dickeya*, *Sodalis*, *Yersinia* and, *Photorhadbus* were obtained loading its sequence in the ColabFold [39], an open-source software to predict protein structure using AlphaFold2. All modelled structures for IgaA showed high and confident accuracy per residue modelled (pLDDT >90%–70%) for most of the sequence. The modelled structures of RcsD were obtained using also the open-source software ColabFold to predict protein structure using AlphaFold2 and showed high and confident accuracy per residue modelled (pLDDT >90%–70%). The modelled structures of RcsF using AlphaFold2 were downloaded from the Uniprot database (A0A0H3N9F1 for *Salmonella*, A0A4P7TT85 for *Shigella*, E0SDI5 for *Dickeya*, Q2NRN2 for *Sodalis*, A1JP50 for *Yersinia* and Q7N8M5 for *Photorhabdus*) and showed high confident accuracy per residue (pLDDT >90%). Structural superposition was produced with Superpose and Gesamt, programs supported in the CCP4 Suite [40]. Multiple sequence alignment was performed using PRALINE [41]. The structure of the SBB domains found in modelled IgaA were compared with other proteins in the PDB using the DALI server [42]. The cartoon representation of the structures was produced using UCSF Chimera [43].

Author contribution statement

Leticia Rodriguez: Conceived and designed experiments. Performed experiments. Analyzed and interpreted the data. Revised the draft article. Final approval de the submitted manuscript.

Marcos Peñalver: Conceived and designed experiments. Performed experiments. Analyzed and interpreted the data. Revised the draft article. Final approval de the submitted manuscript.

Patricia Casino: Conceived and designed experiments. Performed experiments. Analyzed and interpreted the data. Revised the draft article. Final approval de the submitted manuscript.

Francisco García-del Portillo: Conceived and designed experiments. Analyzed and interpreted the data. Wrote the paper.

Data availability statement

Data will be made available on request.

Funding

This work was supported by grant PID2020-112971GB-I00/10.13039/501100011033 (to F.G.dP.) and PID2019-110630GB-I00 (to P·C.) from the Spanish Ministry of Science and Innovation.

Additional information

Supplementary content related to this article has been published online at [URL].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank L.A. Fernández (Centro Nacional de Biotecnología-CSIC, Madrid, Spain) for *E. coli* strain DH10B; H. Niki (National Institute of Genetics, Shizuoka, Japan) for pBAD24 plasmid; J. Casadesús (University of Seville, Spain) for *S.* Typhimurium strain SV4441; D. Clarke (University College Cork, Ireland) for *P. luminescens* strain TTO1; and Henar González (CNB-CSIC) for technical support.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.heliyon.2023.e16661.

References

[1] E. Wall, N. Majdalani, S. Gottesman, The complex Rcs regulatory cascade, Annu. Rev. Microbiol. 72 (2018) 111–139, https://doi.org/10.1146/annurev-micro-090817-062640.

[2] J. Meng, G. Young, J. Chen, The Rcs system in Enterobacteriaceae: envelope stress responses and virulence regulation, Front. Microbiol. 12 (2021), 627104, https://doi.org/10.3389/fmicb.2021.627104.

[3] D.A. Cano, G. Domínguez-Bernal, A. Tierrez, F. Garcia-Del Portillo, J. Casadesús, Regulation of capsule synthesis and cell motility in Salmonella enterica by the essential gene igaA, Genetics 162 (2002) 1513–1523, https://doi.org/10.1093/genetics/162.4.1513.

[4] G. Domínguez-Bernal, M.G. Pucciarelli, F. Ramos-Morales, M. García-Quintanilla, D.A. Cano, J. Casadesús, F. García-del Portillo, Repression of the RcsC-YojN-RcsB phosphorelay by the IgaA protein is a requisite for Salmonella virulence, Mol. Microbiol. 53 (2004) 1437–1449, https://doi.org/10.1111/j.1365-2958.2004.04213.x.

[5] J.F. Mariscotti, F. García-Del Portillo, Instability of the Salmonella RcsCDB signalling system in the absence of the attenuator IgaA, Microbiology (Read.) 154 (2008) 1372–1383, https://doi.org/10.1099/mic.0.2007/015891-0.

[6] N.A. Hussein, S.-H. Cho, G. Laloux, R. Siam, J.-F. Collet, Distinct domains of Escherichia coli IgaA connect envelope stress sensing and down-regulation of the Rcs phosphorelay across subcellular compartments, PLoS Genet. 14 (2018), e1007398, https://doi.org/10.1371/journal.pgen.1007398.

[7] E.A. Wall, N. Majdalani, S. Gottesman, IgaA negatively regulates the Rcs phosphorelay via contact with the RcsD phosphotransfer protein, PLoS Genet. 16 (2020), e1008610, https://doi.org/10.1371/journal.pgen.1008610.

[8] D.A. Cano, M. Martínez-Moya, M.G. Pucciarelli, E.A. Groisman, J. Casadesús, F. García-Del Portillo, Salmonella enterica serovar Typhimurium response involved in attenuation of pathogen intracellular proliferation, Infect. Immun. 69 (2001) 6463–6474, https://doi.org/10.1128/IAI.69.10.6463-6474.2001.

[9] A. Dufour, R.B. Furness, C. Hughes, Novel genes that upregulate the Proteus mirabilis flhDC master operon controlling flagellar biogenesis and swarming, Mol. Microbiol. 29 (1998) 741–751, https://doi.org/10.1046/j.1365-2958.1998.00967.x.

[10] S.-H. Cho, J. Szewczyk, C. Pesavento, M. Zietek, M. Banzhaf, P. Roszczenko, A. Asmar, G. Laloux, A.-K. Hov, P. Leverrier, C. Van der Henst, D. Vertommen, A. Typas, J.-F. Collet, Detecting envelope stress by monitoring β-barrel assembly, Cell 159 (2014) 1652–1664, https://doi.org/10.1016/j.cell.2014.11.045.

[11] K. Dekoninck, J. Létoquart, C. Laguri, P. Demange, R. Bevernaegie, J.-P. Simorre, O. Dehu, B.I. Iorga, B. Elias, S.-H. Cho, J.-F. Collet, Defining the function of OmpA in the Rcs stress response, Elife 9 (2020), e60861, https://doi.org/10.7554/eLife.60861.

[12] S.R. Lach, S. Kumar, S. Kim, W. Im, A. Konovalova, Conformational rearrangements in the sensory RcsF/OMP complex mediate signal transduction across the bacterial cell envelope, PLoS Genet. 19 (2023), e1010601, https://doi.org/10.1371/journal.pgen.1010601.

[13] N.A. Stella, K.M. Brothers, J.D. Callaghan, A.M. Passerini, C. Sigindere, P.J. Hill, X. Liu, D.J. Wozniak, R.M.Q. Shanks, An IgaA/UmoB family protein from Serratia marcescens regulates motility, capsular polysaccharide biosynthesis, and secondary metabolite production, Appl. Environ. Microbiol. 84 (2018), e02575-17, https://doi.org/10.1128/AEM.02575-17.

[14] Y.-H. Huang, L. Ferrières, D.J. Clarke, Comparative functional analysis of the RcsC sensor kinase from different Enterobacteriaceae, FEMS Microbiol. Lett. 293 (2009) 248–254, https://doi.org/10.1111/j.1574-6968.2009.01543.x.

[15] P. Youkharibache, S. Veretnik, Q. Li, K.A. Stanek, C. Mura, P.E. Bourne, The small β-barrel domain: a survey-based structural analysis, Structure 27 (2019) 6–26, https://doi.org/10.1016/j.str.2018.09.012.

[16] A.G. Murzin, OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences, EMBO J. 12 (1993) 861–867, https://doi.org/10.1002/j.1460-2075.1993.tb05726.x.

[17] S.J. Hinchliffe, S.L. Howard, Y.H. Huang, D.J. Clarke, B.W. Wren, The importance of the Rcs phosphorelay in the survival and pathogenesis of the enteropathogenic yersiniae, Microbiology (Read.) 154 (2008) 1117–1131, https://doi.org/10.1099/mic.0.2007/012534-0.

[18] M. Adeolu, S. Alnajar, S. Naushad, R.S. Gupta, Genome-based phylogeny and taxonomy of the "Enterobacteriales": proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov, Int. J. Syst. Evol. Microbiol. 66 (2016) 5575–5599, https://doi.org/10.1099/ijsem.0.001485.

[19] L.M. Guzman, D. Belin, M.J. Carson, J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, J. Bacteriol. 177 (1995) 4121–4130, https://doi.org/10.1128/jb.177.14.4121-4130.1995.

[20] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.

[21] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res. 50 (2022) D439–D444, https://doi.org/10.1093/nar/gkab1061.

[22] M.D. Allen, M. Christie, P. Jones, B.T. Porebski, B. Roome, S.M.V. Freund, A.M. Buckle, M. Bycroft, D. Christ, Solution structure of a soluble fragment derived from a membrane protein by shotgun proteolysis, Protein Eng. Des. Sel. 28 (2015) 445–450, https://doi.org/10.1093/protein/gzv021.

[23] Z. Zhang, W.A. Hendrickson, Structural characterization of the predominant family of histidine kinase sensor domains, J. Mol. Biol. 400 (2010) 335–353, https://doi.org/10.1016/j.jmb.2010.04.049.

[24] M. van Kempen, S. Kim, C. Tumescheit, M. Mirdita, C.L.M. Gilchrist, J. Söding, M. Steinegger, Fast and Accurate Protein Structure Search with Foldseek, BioRxiv., 2022.

[25] A.A. Upadhyay, A.D. Fleetwood, O. Adebali, R.D. Finn, I.B. Zhulin, Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes, PLoS Comput. Biol. 12 (2016), e1004862, https://doi.org/10.1371/journal.pcbi.1004862.

[26] M.C. Pilonieta, K.D. Erickson, R.K. Ernst, C.S. Detweiler, A protein important for antimicrobial peptide resistance, YdeI/OmdA, is in the periplasm and interacts with OmpD/NmpC, J. Bacteriol. 191 (2009) 7243–7252, https://doi.org/10.1128/JB.00688-09.

[27] A. Arunima, S.K. Swain, S.D. Patra, S. Das, N.K. Mohakud, N. Misra, M. Suar, Role of OB-fold protein YdeI in stress response and virulence of Salmonella enterica serovar enteritidis, J. Bacteriol. 203 (2020), e00237-20, https://doi.org/10.1128/JB.00237-20.

[28] K.D. Erickson, C.S. Detweiler, The Rcs phosphorelay system is specific to enteric pathogens/commensals and activates ydeI, a gene important for persistent Salmonella infection of mice, Mol. Microbiol. 62 (2006) 883–894, https://doi.org/10.1111/j.1365-2958.2006.05420.x.

[29] S.K. Hoiseth, B.A. Stocker, Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines, Nature 291 (1981) 238–239, https://doi.org/10.1038/291238a0.

[30] S.G. Grant, J. Jessee, F.R. Bloom, D. Hanahan, Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 4645–4649, https://doi.org/10.1073/pnas.87.12.4645.

[31] N.T. Onodera, J. Ryu, T. Durbic, C. Nislow, J.M. Archibald, J.R. Rohde, Genome sequence of Shigella flexneri serotype 5a strain M90T Sm, J. Bacteriol. 194 (2012) 3022, https://doi.org/10.1128/JB.00393-12.

[32] J.D. Glasner, C.-H. Yang, S. Reverchon, N. Hugouvieux-Cotte-Pattat, G. Condemine, J.-P. Bohin, F. Van Gijsegem, S. Yang, T. Franza, D. Expert, G. Plunkett, M. J. San Francisco, A.O. Charkowski, B. Py, K. Bell, L. Rauscher, P. Rodriguez-Palenzuela, A. Toussaint, M.C. Holeva, S.Y. He, V. Douet, M. Boccara, C. Blanco, I. Toth, B.D. Anderson, B.S. Biehl, B. Mau, S.M. Flynn, F. Barras, M. Lindeberg, P.R.J. Birch, S. Tsuyumu, X. Shi, M. Hibbing, M.-N. Yap, M. Carpentier, E. Dassa, M. Umehara, J.F. Kim, M. Rusch, P. Soni, G.F. Mayhew, D.E. Fouts, S.R. Gill, F.R. Blattner, N.T. Keen, N.T. Perna, Genome sequence of the plant-pathogenic bacterium Dickeya dadantii 3937, J. Bacteriol. 193 (2011) 2076–2077, https://doi.org/10.1128/JB.01513-10.

[33] H. Toh, B.L. Weiss, S.A.H. Perkin, A. Yamashita, K. Oshima, M. Hattori, S. Aksoy, Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of Sodalis glossinidius in the tsetse host, Genome Res. 16 (2006) 149–156, https://doi.org/10.1101/gr.4106106.

[34] N.R. Thomson, S. Howard, B.W. Wren, M.T.G. Holden, L. Crossman, G.L. Challis, C. Churcher, K. Mungall, K. Brooks, T. Chillingworth, T. Feltwell, Z. Abdellah, H. Hauser, K. Jagels, M. Maddison, S. Moule, M. Sanders, S. Whitehead, M.A. Quail, G. Dougan, J. Parkhill, M.B. Prentice, The complete genome sequence and comparative genome analysis of the high pathogenicity Yersinia enterocolitica strain 8081, PLoS Genet. 2 (2006), e206, https://doi.org/10.1371/journal.pgen.0020206.

[35] E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles, E. Dassa, R. Derose, S. Derzelle, G. Freyssinet, S. Gaudriault, C. Médigue, A. Lanois, K. Powell, P. Siguier, R. Vincent, V. Wingate, M. Zouine, P. Glaser, N. Boemare, A. Danchin, F. Kunst, The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens, Nat. Biotechnol. 21 (2003) 1307–1313, https://doi.org/10.1038/nbt886.

[36] J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S.K. Forslund, H. Cook, D.R. Mende, I. Letunic, T. Rattei, L.J. Jensen, C. von Mering, P. Bork, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses, Nucleic Acids Res. 47 (2019) D309–D314, https://doi.org/10.1093/nar/gky1085.

[37] S.Q. Le, O. Gascuel, An improved general amino acid replacement matrix, Mol. Biol. Evol. 25 (2008) 1307–1320, https://doi.org/10.1093/molbev/msn067.

[38] I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, Nucleic Acids Res. 49 (2021) W293–W296, https://doi.org/10.1093/nar/gkab301.

[39] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all, Nat. Methods 19 (2022) 679–682, https://doi.org/10.1038/s41592-022-01488-1.

[40] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G. N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242, https://doi.org/10.1107/S0907444910045749.

[41] P. Bawono, J. Heringa, PRALINE: a versatile multiple sequence alignment toolkit, Methods Mol. Biol. 1079 (2014) 245–262, https://doi.org/10.1007/978-1-62703-646-7_16.

[42] L. Holm, Dali server: structural unification of protein families, Nucleic Acids Res. (2022), https://doi.org/10.1093/nar/gkac387 gkac387.

[43] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera–a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612, https://doi.org/10.1002/jcc.20084.