An SOS Regulon under Control of a Noncanonical LexA-Binding Motif in the *Betaproteobacteria*

Neus Sanchez-Alberola,^{a,b} Susana Campoy,^a David Emerson,^c Jordi Barbé,^a Ivan Erill^b

Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain^a; Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland, USA^b; Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA^c

ABSTRACT

The SOS response is a transcriptional regulatory network governed by the LexA repressor that activates in response to DNA damage. In the *Betaproteobacteria*, LexA is known to target a palindromic sequence with the consensus sequence CTGT-N₈-ACAG. We report the characterization of a LexA regulon in the iron-oxidizing betaproteobacterium *Sideroxydans lithotrophicus*. *In silico* and *in vitro* analyses show that LexA targets six genes by recognizing a binding motif with the consensus sequence GAACGaaCGTTC, which is strongly reminiscent of the *Bacillus subtilis* LexA-binding motif. We confirm that the closely related *Gallionella capsiferiformans* shares the same LexA-binding motif, and *in silico* analyses indicate that this motif is also conserved in the *Nitrosomonadales* and the *Methylophilales*. Phylogenetic analysis of LexA and the alpha subunit of DNA polymerase III (DnaE) reveal that the organisms harboring this noncanonical LexA form a compact taxonomic cluster within the *Betaproteobacteria*. However, their *lexA* gene is unrelated to the standard *Betaproteobacteria* *lexA*, and there is evidence of its spread through lateral gene transfer. In contrast to other reported cases of noncanonical LexA-binding motifs, the regulon of *S. lithotrophicus* is comparable in size and function to that of many other *Betaproteobacteria*, suggesting that a convergent SOS regulon has reevolved under the control of a new LexA protein. Analysis of the DNA-binding domain of *S. lithotrophicus* LexA reveals little sequence similarity with that of other LexA proteins targeting similar binding motifs, suggesting that network structure may limit site evolution or that structural constraints make the *B. subtilis*-type motif an optimal interface for multiple LexA sequences.

IMPORTANCE

Understanding the evolution of transcriptional systems enables us to address important questions in microbiology, such as the emergence and transfer potential of different regulatory systems to regulate virulence or mediate responses to stress. The results reported here constitute the first characterization of a noncanonical LexA protein regulating a standard SOS regulon. This is significant because it illustrates how a complex transcriptional program can be put under the control of a novel transcriptional regulator. Our results also reveal a substantial degree of plasticity in the LexA recognition domain, raising intriguing questions about the space of protein-DNA interfaces and the specific evolutionary constraints faced by these elements.

In order to survive, bacteria react to transient environmental insults through the activation of coordinated transcriptional programs, collectively known as stress responses. Among these, the SOS response to DNA damage is one of the most widely conserved and well characterized (1–3). In *Escherichia coli*, where it was originally described, the SOS response controls the expression of almost 40 genes involved in DNA repair, translesion synthesis, and cell division inhibition (2, 4). The SOS response is coordinated by the transcriptional repressor LexA, which binds to specific sites upstream of regulated operons and blocks transcription initiation (1, 2). DNA damage is sensed by the RecA protein, which forms active nucleoprotein filaments upon binding single-stranded DNA fragments formed at stalled replication forks (5). Activated RecA promotes self-cleavage of the LexA repressor, leading to the induction of the SOS response, which includes the *lexA* and *recA* genes (2, 4, 6).

Most major bacterial clades possess a functional *lexA* gene whose product controls an SOS regulon of variable size, with notable exceptions, including the *Bacteroidetes*, the *Epsilonproteobacteria* or the *Streptococcaceae* (3, 7, 8). In contrast to other broadly conserved transcriptional regulators, the DNA-binding motif of LexA has undergone several drastic changes in the course of evolution (3). For instance, LexA targets a 16-bp palindromic binding

motif with a consensus sequence (CTGT-N₈-ACAG) in *E. coli*, a 12-bp palindrome in *Bacillus subtilis* (GAAC-N₄-GTTC), and a 15-bp direct repeat in *Caulobacter vibrioides* (GTTC-N₇-GTTC) (9–11). These motifs are known to be monophyletic in their respective major clades (*Gamma*proteobacteria, *Firmicutes*, and *Alphaproteobacteria*) (12–14), but the description of three distinct LexA-binding motifs in the *Deltaproteobacteria* bears testament to the substantial recognition plasticity of the LexA repressor over different evolutionary scales. In spite of these changes to the LexA-binding motif, the SOS response maintains a consistent regulatory

Received 21 January 2015 Accepted 9 May 2015

Accepted manuscript posted online 18 May 2015

Citation Sanchez-Alberola N, Campoy S, Emerson D, Barbé J, Erill I. 2015. An SOS regulon under control of a noncanonical LexA-binding motif in the *Betaproteobacteria*. *J Bacteriol* 197:2622–2630. doi:10.1128/JB.00035-15.

Editor: R. L. Gourse

Address correspondence to Ivan Erill, erill@umbc.edu.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/JB.00035-15>.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JB.00035-15

function across bacterial groups, providing a unique perspective into the coevolution of transcriptional regulators and their regulatory networks (3, 4, 9, 11).

In the *Betaproteobacteria*, *in vivo* and *in silico* studies have established that the SOS response is regulated by a LexA protein targeting a binding motif closely related to that observed in *E. coli* and other *Gammaproteobacteria* (12, 15, 16). Comparative genomic and transcriptomic data, however, indicate that the SOS regulon of the *Betaproteobacteria* displays some fundamental differences from that reported for many *Gammaproteobacteria*. For instance, the *Betaproteobacteria* LexA protein does not regulate the *recN* gene or the recombination repair operon *ruvAB*. Beyond the core SOS genes shared with the *Gammaproteobacteria* (*lexA*, *recA*, and *uvrA*), the *Betaproteobacteria* SOS response regulates a gene coding for a predicted photolyase (*splB*) and a mutagenesis cassette (*imuAB-dnaE2*) also regulated by LexA in many bacterial groups (15, 17). *Sideroxydans lithotrophicus* and *Gallionella capsiferriformans* are chemolithoautotrophic *Betaproteobacteria* capable of using iron as their main electron source (18). The complete genome sequences of these two representatives of the *Gallionellales* order reveal the presence of a *lexA* gene with low sequence identity to its well-characterized homologs in other *Betaproteobacteria* orders (19). Here, we combine *in silico* and *in vitro* techniques to characterize the LexA-binding motif and SOS regulon of these two bacterial species. Our results confirm that the *lexA* gene of the *Gallionellales* and other related bacteria is not closely related to the canonical *Betaproteobacteria* *lexA* and show that its product targets a binding motif resembling that recognized by *B. subtilis* LexA. Furthermore, the composition of the SOS regulon in *S. lithotrophicus* suggests that the SOS regulatory network under the control of this novel *lexA* arose through a process of convergent evolution.

MATERIALS AND METHODS

Genome, motif, and sequence data. Available complete genome sequences for the species under study were downloaded from the NCBI RefSeq database in GenBank (.gbk) format (20). Experimentally validated transcription factor-binding motifs for individual species and clade aggregates were downloaded from the CollecTF database (21). Orthologs for *S. lithotrophicus* and *G. capsiferriformans* LexA on complete genome sequences were identified with PSI-BLAST and validated as best reciprocal BLAST hits using BLASTP with E values lower than 10^{-20} (14, 22). LexA and DnaE orthologs from complete genome sequences were compiled based on their annotation in RefSeq and BLAST searches and were then validated through reciprocal BLAST. LexA and DnaE orthologs were compiled to provide broad coverage of selected phyla and classes, while maintaining a relatively low total number of taxa for phylogenetic inference. For comparative genomics, orthologs of putatively regulated genes were determined as best reciprocal BLAST hits of their encoded protein sequences. Upstream regions (−250, +50) of *lexA* orthologs were obtained from the NCBI GenBank database using custom Python scripts. Circular genome maps for *S. lithotrophicus* and *G. capsiferriformans* were generated using DNAPlotter (23) with default parameters. Sequence identity from pairwise alignments and molecular weights for LexA proteins were computed using the BioWord add-on for Microsoft Word (24).

Site search, motif discovery, and comparative genomics. Whole-genome sequences were scanned for inferred and experimentally validated motifs using xFITOM with the sequence information content (R_i) scoring method and a motif-normalized threshold as reported previously (25–27). Putative LexA-binding sites were associated with annotated genes if they were located within bp −250 to +50 from the predicted translation

start point for the gene and their score was less than one standard deviation below the average score of the original collection (13, 15). For the comparative genomics analysis of the *Gallionellales* taxonomic cluster, genes were defined as LexA-regulated if candidate sites were located in at least two species of the cluster. The LexA-binding motif of the *Gallionellales* cluster was identified using MEME on the upstream (−250, +50) regions of identified *lexA* gene orthologs, searching for palindromic 10- to 30-bp motifs with *N* instances per sequence and with otherwise default parameters (28). Putative −35 and −10 promoter elements were identified with the phiSITE PromoterHunter web service and the PePPER toolbox with default parameters (29, 30).

Multiple sequence alignments and phylogenetic inference. Multiple sequence alignments of DNA and protein sequences were performed using CLUSTAL W (31). For LexA and DnaE sequence alignments, the information of the *E. coli* LexA and DnaE crystal structures (32, 33) subsumed into the UniProt records P0A7C2 and P10443 were used to define penalty masks in profile alignment mode. Editing and graphical representation of alignments were performed with BioEdit (34). Sequence logos were created using aligned DNA or protein sequences with the WebLogo service (35). For phylogenetic inference, multiple sequence alignments of LexA and DnaE protein sequences were generated with CLUSTAL W using variable (5, 10, and 25) gap opening penalties for the multiple-alignment stage (17). These alignments were integrated with a local LALIGN alignment for optimization using T-COFFEE (36), and the resulting alignment was processed with Gblocks with the half-gaps setting and otherwise default parameters to select conserved positions for reliable phylogenetic inference (37). As the closest relatives to *Gammaproteobacteria* and *Betaproteobacteria*, *Alphaproteobacteria* species were included as the outgroup for DnaE tree inference (38). *Alphaproteobacteria* LexA protein sequences are known to present a 33-amino-acid insertion that disrupts the multiple-sequence alignment (39). To enhance alignment quality and phylogenetic resolution, Gram-positive LexA protein sequences were used as the outgroup for LexA tree inference. Bayesian phylogenetic inference on T-COFFEE-optimized alignments was performed with MrBayes (40). Twelve Metropolis-Coupled Markov Chain Monte Carlo runs with four independent chains were carried out for 500,000 generations, and the resulting consensus tree was plotted with TreeView (41).

EMSAs. Electrophoresis mobility shift assays (EMSAs) were performed to validate computationally predicted LexA-binding sites in the promoter region of putative *G. capsiferriformans* and *S. lithotrophicus* SOS genes. *G. capsiferriformans* ES-2 DNA was extracted from phosphate buffered (50 mM) saline (pH 8.0)-washed pellets containing cells and iron oxides using a MoBio PowerSoil device (MoBio, Carlsbad, CA) DNA isolation kit (42). The *G. capsiferriformans* ES-2 *lexA* gene was amplified using suitable primers (see Table S1 in the supplemental material) and cloned into a pET15b vector (see Table S2 in the supplemental material). Overexpression and purification of the *G. capsiferriformans* ES-2 and *E. coli* and *B. subtilis* LexA proteins was performed as described previously for other LexA proteins (26, 43, 44). DNA probes were constructed using two complementary 61-bp synthetic oligonucleotides centered on the LexA-binding sites of interest (see Table S1 in the supplemental material). EMSAs were performed as described previously (15), using 40 nM *G. capsiferriformans* ES-2 LexA, 200 nM *E. coli* LexA, or 40 nM *B. subtilis* LexA protein and 20 ng of each digoxigenin-marked DNA probe in the binding mixture. For EMSA competitive assays, 200-fold of either specific or nonspecific unlabeled DNA was added to the binding mixture. In all cases, samples were loaded onto 6% nondenaturing Tris-glycine polyacrylamide gels. Digoxigenin-labeled DNA-protein complexes were detected using the manufacturer's protocol (Roche Applied Science, Indianapolis, IN). Information for all transcription factor-binding sites reported in the present study is available on the CollecTF database (<http://www.collectf.org>).

FIG 1 Sequence alignment of the region immediately upstream of the *S. lithotrophicus* (Slit_2300) and *G. capsiferriformans* (Galf_0635) *lexA* genes. Predicted -35 and -10 promoter elements are boxed. The predicted translational start site is underlined. The conserved palindromic sequences corresponding to putative LexA-binding sites are shaded in gray.

RESULTS AND DISCUSSION

In silico analysis of the *lexA* promoter of *S. lithotrophicus* and *G. capsiferriformans*. The complete genome sequences of both *S. lithotrophicus* and *G. capsiferriformans* contain a gene (Slit_2300 and Galf_0635, respectively) with a predicted product annotated as LexA repressor. The two predicted *lexA* genes are in opposite ends of the genome (see Fig. S1 in the supplemental material) and, following the overall pattern reported for these two genomes, the genomic surroundings of these *lexA* genes reveal no evidence of synteny (19). Multiple-sequence alignment of the amino acid sequence of these predicted LexA proteins shows that they are homologs, sharing 78% sequence identity with each other, but only 24% sequence identity versus the LexA protein of *E. coli* (coded by b4043) or that of the more closely related *Burkholderia thailandensis* (coded by BTH_I2481). The latter figure is comparable to the sequence identity (28%) that *S. lithotrophicus* and *G. capsiferriformans* LexA proteins share with their ortholog (coded by BSU17850) in the more distantly related *B. subtilis* (see Fig. S2 in the supplemental material). An *in silico* analysis of the *S. lithotrophicus* and *G. capsiferriformans* genome sequences using the LexA-binding motif targeted by *E. coli* and *B. thailandensis* LexA did not yield any significant hits in the promoter region of known SOS genes, including *lexA*. Sequence alignment of the intergenic region upstream of these two *lexA* genes revealed the presence of two conserved imperfect palindromes in a tandem configuration and overlapping the region encompassing the predicted -35 and -10 promoter elements (Fig. 1). Except for one reported case (45), the *lexA* gene is typically regulated by its functional product through binding at specific target sites on its upstream sequence, and this property has been effectively leveraged to infer the binding motif of LexA and other transcriptional repressors through comparative

genomics (14, 46). Tandem configurations resembling the conserved palindromes shown in Fig. 1 have been reported previously for functional LexA-binding sites upstream of *lexA* and other SOS regulated genes (9, 12, 14, 47, 48). Furthermore, the identified palindromes display a consensus sequence (GAACG-N₂-CGTTC) that is strongly reminiscent of the *B. subtilis* LexA-binding motif. Functional LexA-binding motifs bearing strong similarity to the *B. subtilis* one have been reported in *Firmicutes*, *Actinobacteria*, *Cyanobacteria*, and *chloroflexi* and for a LexA paralog in *Gammaproteobacteria* (26, 48, 49–54). Taken together, these data strongly suggested that the identified tandem palindromes were functional LexA-binding sites.

Characterization of the SOS regulon in the *Gallionellales*. To study the composition of the SOS regulatory network under the control of the *Gallionellales* LexA protein, we performed an *in silico* genome-wide search for putative LexA-binding sites using the four identified conserved palindromes as the putative LexA-binding motif. The results shown in Table 1, reveal both similarities and differences in the composition of the putative *S. lithotrophicus* and *G. capsiferriformans* SOS response. The inferred LexA regulons comprise five shared genes coding for essential SOS proteins: the LexA repressor, the recombination proteins RecA and RecX, and the *umuDC*-encoded DNA polymerase V. In *G. capsiferriformans*, the predicted LexA regulon encompasses two copies of DNA polymerase V, since its *lexA* gene (Galf_0635) forms a putative operon with *umuDC* (Galf_0634–Galf_0633). These results hence point toward a core LexA regulon devoted to translation synthesis and are in broad agreement with previous reports on the composition of the SOS response in multiple bacterial clades (3). In addition to these core genes, the LexA protein of *S. lithotrophicus* appears to regulate three additional genes: a *splB*

TABLE 1 Putative LexA-binding sites identified through a genome-wide *in silico* search^a

Score	Site	Strand	Distance (bp)		Position	Locus	Product
			d_{TLS}	d_{-10}			
19.00	AAGAACGTTTCGTTCTG	+	−53	−16	671728	Galf_0635	LexA
17.41	GAGAACGTTTCGTTCTG	+	−50	−16	2589520	Galf_2402	UmuDC
17.41	CAGCACGTTTCGTTCTC	−	−28	+9	671719	Galf_0635	LexA
17.00	AAGAACGATCGTTCTG	+	−72	−37 ^b	554879	Galf_0508	RecA
18.77	AAGAACGTTTCGTTCTG	+	−38	+8	819364	Slit_0831	UmuDC
18.77	CAGAACGTTTCGTTCTG	−	−60	−16	2456160	Slit_2431	Fic dom.
18.77	AAGAACGTTTCGTTCTG	−	−29	+8	2329266	Slit_2300	LexA
18.77	AAGAACGTTTCGTTCTG	−	−29	+3	819355	Slit_0831	UmuDC
18.77	CAGAACGTTTCGTTCTG	−	−57	+11	723063	Slit_0718	SplB
18.77	CAGAACGTTTCGTTCTG	−	−60	−16	2456160	Slit_2431	(hyp)
18.18	CAGAACGTTTCGTTCTC	+	−60	−37	2133987	Slit_2105	RecA
17.18	CAGAACGCTCGTTCTT	−	−76	−16	2329290	Slit_2300	LexA

^a Site scores are reported in bits. Positive and negative symbols denote the strand on which sites were detected with a maximal score with respect to the leading strand in the annotated genome. The d_{TLS} and d_{-10} columns indicate the distance from the 5' end of the site to the 5' end of the predicted translation start site (TLS) and the -10 promoter element, respectively. The absolute genome position of the site is also given, as well as the NCBI locus tag for the gene with the closest TLS.

^b For Galf_0508, two alternative -10 elements were predicted (see Fig. S5 in the supplemental material). The distances of the putative LexA-binding site to these elements are, respectively, -37 and $+12$ bp.

FIG 2 EMSAs were performed with 61-bp fragments of *G. capsiferiformans* and *S. lithotrophicus* DNA centered on putative LexA-binding sites predicted *in silico*. The numbers between brackets denote the distance of the targeted site to the predicted TLS, as reported in Table 1. Locus tags for putative SOS genes are reported. Lanes E and B designate assays with *E. coli* and *B. subtilis* LexA proteins, respectively. For all other lanes, *G. capsiferiformans* LexA protein was used. The “+” and “–” symbols denote, respectively, the presence or absence of LexA protein. The molecular masses of these LexA proteins are as follows: *E. coli* (22.35 kDa), *B. subtilis* (22.85 kDa), and *G. capsiferiformans* (21.97 kDa).

homolog encoding a predicted photolyase (Slit_0718), a hypothetical protein (Slit_2431), and a gene encoding a predicted filamentation protein containing a Fic domain (Slit_2431).

The *in silico* search results presented in Table 1 are congruent with LexA defining a canonical SOS response by recognizing *B. subtilis*-type LexA-binding sites in the promoter region of SOS genes. To validate that this putative LexA-binding motif was involved in the recognition of specific promoters by LexA, we performed EMSAs with purified LexA protein from *G. capsiferiformans*, *E. coli*, and *B. subtilis* on the promoter region of putative SOS genes from *G. capsiferiformans* and *S. lithotrophicus*. The EMSA results shown in Fig. 2 reveal that *G. capsiferiformans* LexA is able to bind targets identified *in silico* on both organisms, indicating that the LexA-binding motif of the *Gallionellales* is a 14-bp palindrome (AGAACG-N₂-CGTTCT) that closely resembles the *B. subtilis* motif. Titration experiments with competing DNA show the binding of *G. capsiferiformans* LexA binds is specific to *Gallionellales* promoters (see Fig. S3 in the supplemental material). Further evidence for the association between the *Gallionellales* and Gram-positive LexA-binding motifs is provided by the fact that purified *B. subtilis* LexA binds canonical SOS targets (*lexA* and *recA* promoters) of both organisms. In contrast, *E. coli* LexA, which targets a LexA-binding motif very similar to that reported in the *Betaproteobacteria* *B. thailandensis*, *Bordetella pertussis*, and *Ralstonia solanacearum* (12, 16, 55), does not bind canonical target promoters in the *Gallionellales*. In spite of the overall similarity of their binding motifs, *B. subtilis* and *G. capsiferiformans* LexA proteins show distinct binding specificities (see Fig. S4 in the supplemental material). As suggested by its strong conservation in the identified *G. capsiferiformans* and *S. lithotrophicus* LexA-binding sites (Table 1), the spacer and terminal positions of the AGAACG-N₂-CGTTCT motif play an important role in site recognition by *G. capsiferiformans* LexA. Together with the *in silico* analyses, the binding assays reported in Fig. 2, Fig. S3, and Fig. S4 demonstrate that LexA recognizes a noncanonical binding motif in the *Gallionellales*, and indicate that these noncanonical sites define a small but prototypical SOS network in *S. lithotrophicus* and *G. capsiferiformans*. In this regard, it is worth noting that the proximal end of all of the LexA-binding sites reported in Table 1 is at the most 6 bp away from the predicted core promoter sequence and that in 78% of the analyzed sequences there is one LexA-binding site located within the predicted core promoter sequence (see Fig. S5 in the supplemental material). The location of predicted LexA-binding sites hence not only supports the notion that these are

functional but also indicates that in all of these instances LexA most likely operates as a direct repressor, inhibiting transcriptional initiation by blocking access to the RNA polymerase binding site (10, 56, 57, 58).

Phylogenetic and comparative genomics analysis of the *Gallionellales* LexA. The identification of a LexA-binding protein targeting a noncanonical binding motif within the *Betaproteobacteria* can be most readily explained by two alternative evolutionary scenarios: lateral gene transfer (LGT) or duplication, followed by divergence and deletion of the canonical *Betaproteobacteria* *lexA* gene. Previous research has postulated the former as the likely mechanism for the presence of an *Alphaproteobacteria*-like LexA-binding motif in the *Acidobacteria* and diverging motifs in the *Leptospira* genus and the latter as the originating process for the existence of an additional LexA protein targeting a *Cyanobacteria*-like motif in several *Gammaproteobacteria* (39, 51, 59). To elucidate the likely origin of the *Gallionellales* LexA protein, we performed phylogenetic inference on the amino acid sequence of the housekeeping α subunit of DNA polymerase III (DnaE) and the LexA protein (17). In agreement with previous work, the DnaE phylogenetic tree (Fig. 3) establishes the *Gallionellales* as a distinct clade within the *Betaproteobacteria* (19, 60). Furthermore, it confirms that this order is most closely related with the *Nitrosomonadales* and the *Methylophilales* and that this cluster delineates an early branching point in the *Betaproteobacteria* radiation (19, 60). The LexA tree also clusters the *Gallionellales* with the *Nitrosomonadales* and the *Methylophilales* but places them in a highly divergent group far removed from the main *Betaproteobacteria* clade. In addition to members of these orders, the cluster encompassing the *Gallionellales* LexA also includes LexA proteins from several unrelated *Burkholderiales* and from a *Deltaproteobacteria* (*Syntrophus aciditrophicus*). Phylogenetic analysis of a nonhousekeeping SOS protein, UmuC, is consistent with the results obtained for DnaE (see Fig. S6 in the supplemental material), indicating that the origin of the *Gallionellales* *lexA* is not linked with a broader rearrangement of SOS-related components.

The robust clustering of the *Gallionellales* with both related and unrelated bacteria in the LexA tree suggests that these LexA sequences share a origin and that, in accordance, their LexA protein might also target the same binding motif. Relying on the assumption that the *lexA* gene is regulated by its product (14), we used the MEME motif discovery algorithm to identify overrepresented motifs in the promoter region of *lexA* homologs from species clustering with the *Gallionellales* LexA. The most significant result

FIG 3 Consensus tree of DnaE (top) and LexA (bottom) protein sequences. Branch support values are provided as Bayesian posterior probabilities. LexA-binding motifs for different clades are illustrated using sequence logos made with inferred binding sites (*Gallionellales* cluster) or aggregate experimental LexA-binding data downloaded from the CollecTF database. Different clades are shaded in alternating shades of gray. The cluster containing the *Gallionellales* DnaE and LexA proteins is boxed and lightly shaded. The scale bar on each tree denotes expected substitutions per residue. To enhance alignment quality, the DnaE tree was rooted using two *Alphaproteobacteria* LexA sequences as outgroup and the LexA tree was rooted with four Gram-positive sequences as outgroup. A DnaE tree rooted with Gram-positive sequences is provided in Fig. S7 in the supplemental material for a direct comparison. The protein accession numbers and full species names for all represented protein sequences are available in Table S3 in the supplemental material.

returned by MEME is a LexA-binding motif encompassing the one experimentally validated here for the *Gallionellales* (AGAAC-N₄-GTTCT) (Fig. 3). This motif was identified in the promoter region of *lexA* in the *Gallionellales*, *Nitrosomonadales*, *Methylobacter* *versatilis*, *Burkholderia phytofirmans*, and *S. aciditrophicus*. In addition, a related motif (TGTAC-N₄-GTACA) was identified in the *Methylophilales* *Methylobacillus flagellatus* and *Methylovorus glucosotrophus* (Fig. 3). To further verify that the identified motifs constituted the LexA-binding motif of this heterogeneous cluster, we used these motifs to perform a comparative genomics analysis of the predicted LexA regulon in the cluster of species encompassing the *Gallionellales* LexA. The comparative genomics results (Table 2) reveal that the shared LexA regulon in this cluster is limited to *lexA*, with only the two *Gallionellales* species showing evidence of *recA* regulation. In spite of this, the inferred LexA regulons of several species present significant similarities, such as the regulation of *umuDC* operons and of the *splB* gene encoding a predicted photolyase, providing further support for the *in vivo* functionality of the LexA-binding sites experimentally validated in this work.

Given the short and variable nature of the LexA protein sequence, phylogenetic inference methods cannot fully resolve the origin of the *Gallionellales* LexA. Nonetheless, the combined results of Fig. 3 and Table 2 do provide conclusive evidence of several facts regarding the evolutionary history of this protein. The presence in *B. phytofirmans*, *Alicyclophilus denitrificans*, and *Ralstonia pickettii* of another LexA protein clustering with the canonical *Betaproteobacteria* LexA clearly establishes that the *Gallionellales* LexA does not derive directly from the canonical *Betaproteobacteria* LexA. Instead, the presence of two *lexA* copies in these genomes is suggestive of an early duplication scenario at the branching point between *Gammaproteobacteria* and *Betaproteobacteria*, followed by selective deletion of the canonical *lexA* gene (17). On the other hand, the presence of a *Gallionellales*-type LexA in a single *Deltaproteobacteria* indicates that this *lexA* has been involved in at least one lateral gene transfer event. Furthermore, the results also establish that the LexA proteins clustering with the *Gallionellales* LexA share a ancestor and target variations of a LexA-binding motif similar to that reported in *B. subtilis*. The small size of the putative SOS regulon in these *Gallionellales* cluster species, and specifically the loss of *recA* regulation by LexA, has been described before in association with rapidly diverging *lexA* genes, *lexA* duplications and putative LGT (26, 39, 59, 61–64). In this setting, the description in *S. lithotrophicus* of a more elaborate LexA regulon, including a signature gene of the canonical *Betaproteobacteria* SOS response (*splB*) points to a process of convergent evolution toward the regulation of this gene. This notion is reinforced by the consistent regulation of *umuDC* operons in *Gallionellales* cluster species, since their phylogeny is congruent with a stable genetic composition of SOS genes in the *Betaproteobacteria* (see Fig. S6 in the supplemental material). The hypothesis of convergent evolution toward the regulation of a prototypical *Betaproteobacteria* SOS response applies not only to scenarios involving LGT, where novel binding sites for the transferred LexA must necessarily be evolved on target genes, but also to scenarios involving gene duplication, followed by deletion of the canonical *lexA*. In the latter, the dramatic changes observed in the sequence of the noncanonical LexA protein and its binding motif are suggestive of relaxed or absent selection. This is difficult to reconcile with preservation of regulon composition, implying that sites for the non-

TABLE 2 Results of the comparative genomics analysis of the *Gallionellales* cluster^a

Gene	Site	Strand	<i>d</i> _{TLS} (bp)
<i>lexA</i>			
Bphyt_3718	CGAATGAACGTTCTG	+	−4
Galf_0635 ^b	AGAACGAACGTTCT	−	−66
	AGAACGAACGTTCT	+	−13
Mfla_1781	TGTACATTCGTACA	−	−42
M301_2214	CGAACGAACGTTCTG	+	−69
	CGAACGAACGTTCTG	−	−65
Msip34_2108	TGTACGTTTCGTACA	−	−30
Nmul_A2596 ^c	AGAACAATGTTCT	−	−54
	AGAGCAACTGTTCT	+	−32
Slit_2300	AGAACGAACGTTCT	+	−14
	AGAACGAGCGTTCT	+	−38
SYN_00359	AGAACAATCGTTCT	+	−13
<i>recA</i>			
Galf_0508	AGAACGATCGTTCT	+	−71
Slit_2105	AGAACGAACGTTCT	−	−89
<i>umuDC</i>			
Galf_2402	AGAACGAACGTTCT	−	−63
Msip34_1483	TGTACGTTTCGTACA	+	−15
Slit_0831	AGAACGAACGTTCT	+	−14
	AGAACGAACGTTCT	−	−67
<i>splB</i>			
Mfla_1759	TGTACGTTTCGTACA	+	−8
Slit_0718	AGAACGAACGTTCT	+	−57

^a Putative LexA-binding sites resulting from a genome-wide search are reported only if present upstream of at least two orthologs for a given gene. Positive and negative symbols denote the strand on which sites were detected with a maximal score with respect to the leading strand in the annotated genome. The *d*_{TLS} column indicates the distance from the 5′ end of the site to the 5′ end of the predicted translation start site (TLS). The NCBI locus tag for the gene with the closest TLS is also given.

^b Galf_0635 forms a putative operon with Galf_0634 and Galf_0633 (*umuDC*).

^c Nmul_A2596 forms a putative operon with Nmul_A2597, Nmul_A2598, and Nmul_A2599 (an *imuAB-dnaE2* cassette).

canonical *lexA* on target genes such as *splB* must have evolved anew after the deletion of the canonical *lexA* gene (65).

Comparative analysis of the LexA-binding domain. LexA proteins recognizing *B. subtilis*-type motifs have been reported in the *Actinobacteria*, the *Cyanobacteria*, the chloroflexi, and some *Gammaproteobacteria* (26, 48, 49–54). Even though phylogenetic inference cannot fully resolve the origin of the *Gallionellales* LexA protein, the fact that it targets a *B. subtilis*-like binding motif suggested that a detailed analysis of its DNA-binding domain might shed light onto its evolutionary origin. The crystal structure of *E. coli* LexA bound to DNA, as well as docked models of *B. subtilis* LexA and numerous biochemical and genetic studies, has clearly established that the α3 helix of the N-terminal helix-turn-helix motif is responsible for the majority of the specific contacts with the symmetric conserved elements of LexA-binding motifs in these two species (33, 66–69). Leveraging this structural information, we performed a multiple-sequence alignment of amino acid sequences for LexA proteins targeting similar half-sites and analyzed the alignment segments corresponding to the α3 helix.

The multiple-sequence alignment results summarized in Fig. 4 show that the α3 helix of the *Gallionellales* LexA is not directly related to the α3 helix of any other LexA targeting a conserved

FIG 4 Sequence logo of aligned LexA $\alpha 3$ helix sequences and experimentally validated LexA-binding sites in different clades. Each $\alpha 3$ helix logo was generated using 8 to 10 aligned representative sequences from each clade (see Fig. S4 in the supplemental material). Clade aggregate LexA-binding data were obtained from the CollecTF database. The $\alpha 3$ helix and binding motif of *Betaproteobacteria* and *Xanthomonadaceae* (1) LexA are shown for comparison. The LexA protein targets a GAAC/GTAC half-site in all of the other depicted clades. Accession numbers and species names for all proteins used in the multiple-sequence alignment are available in Table S4 in the supplemental material.

GAAC half-site. Rather, the *Gallionellales* cluster LexA $\alpha 3$ helix displays a mosaic of features from other $\alpha 3$ helices, such as the conserved $S_1K_2S_3$ pattern seen in the *Alphaproteobacteria* or the $R_{10}L_{11}$ motif observed in the *Firmicutes* and the *Cyanobacteria*. Most importantly, the *Gallionellales* cluster LexA data illustrate that a similar binding motif (GAAC- N_4 -GTTC) can be targeted by LexA proteins with seemingly unrelated $\alpha 3$ helix sequences. Although the $\alpha 3$ helices of *Actinobacteria*, *Firmicutes*, and chloroflexi display substantial sequence similarity, the highly divergent $\alpha 3$ helix sequences of the *Xanthomonadaceae*/*Pseudomonadaceae*, the *Cyanobacteria*, and the *Gallionellales* cluster target a structurally similar motif that has been shown to be bound by several non-cognate LexA proteins (50, 51). Recent work on the PhoQ-PhoP two-component signal transduction system has shown that there is a remarkable degree of sequence plasticity in the PhoQ-PhoP interface and that evolution has explored only a small fraction of the space of the PhoP-specific PhoQ motifs, due to both physiological and biochemical constraints on available mutational pathways (70–72). The results presented here suggest that the same is true for the LexA protein-DNA interface, which displays a varied palette of $\alpha 3$ helix and LexA-binding motif pairs. Importantly, however, the available data indicate that the coevolution of the LexA recognition domain and its binding motif may face more restrictions than those displayed by two-component systems (70, 71, 73). In particular, the abundance of highly divergent $\alpha 3$ helix sequences targeting a similar, *B. subtilis*-type LexA-binding motif yields two nonexclusive scenarios. On the one hand, it suggests that the need for coordinated changes in multiple binding sites

might limit the evolution of the DNA component of the regulatory system. On the other hand, in the same way that the PhoQ-PhoP interface sequence distribution is highly nonuniform (72), the *B. subtilis*-type LexA-binding motif may provide a broadly accessible interface to the canonical LexA structure, enabling its interaction with a large number of $\alpha 3$ helix sequences and therefore substantially concentrating the evolutionary exploration of the $\alpha 3$ helix sequence space.

ACKNOWLEDGMENTS

This study was supported by UMBC Special Research Assistantship/Initiative Support (SRAIS) and U.S. National Science Foundation (MCB-1158056) awards to I.E. and by Spanish Ministry of Science and Innovation (BFU2011-23478) and Generalitat de Catalunya (2014SGR572) awards to J.B. N.S.-A. was funded by a Fundació Cellex fellowship.

REFERENCES

- Little JW, Mount DW. 1982. The SOS regulatory system of *Escherichia coli*. Cell 29:11–22. [http://dx.doi.org/10.1016/0092-8674\(82\)90085-X](http://dx.doi.org/10.1016/0092-8674(82)90085-X).
- Walker GC, Smith BT, Sutton MD. 2000. The SOS response to DNA damage, p 131–144. In Storz G, Hengge-Aronis R (ed), Bacterial stress responses. American Society for Microbiology, Washington, DC.
- Erill I, Campoy S, Barbe J. 2007. Aeons of distress: an evolutionary perspective on the bacterial SOS response. FEMS Microbiol Rev 31:637–656. <http://dx.doi.org/10.1111/j.1574-6976.2007.00082.x>.
- Fernandez De Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R. 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. Mol Microbiol 35:1560–1572.
- Sassanfar M, Roberts JW. 1990. Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. J Mol Biol 212:79–96.

6. Little JW. 1991. Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* 73:411–421. [http://dx.doi.org/10.1016/0300-9084\(91\)90108-D](http://dx.doi.org/10.1016/0300-9084(91)90108-D).
7. Black CG, Fyfe JAM, Davies JK. 1998. Absence of an SOS-like system in *Neisseria gonorrhoeae*. *Gene* 208:61–66. [http://dx.doi.org/10.1016/S0378-1119\(97\)00653-7](http://dx.doi.org/10.1016/S0378-1119(97)00653-7).
8. Boutry C, Delplace B, Clippe A, Fontaine L, Hols P. 2013. SOS response activation and competence development are antagonistic mechanisms in *Streptococcus thermophilus*. *J Bacteriol* 195:696–707. <http://dx.doi.org/10.1128/JB.01605-12>.
9. Au N, Kuester-Schoeck E, Mandava V, Bothwell LE, Canny SP, Chachu K, Colavito SA, Fuller SN, Groban ES, Hensley LA, O'Brien TC, Shah A, Tierney JT, Tomm LL, O'Gara TM, Goranov AI, Grossman AD, Lovett CM. 2005. Genetic composition of the *Bacillus subtilis* SOS system. *J Bacteriol* 187:7655–7666. <http://dx.doi.org/10.1128/JB.187.22.7655-7666.2005>.
10. Wade JT, Reppas NB, Church GM, Struhl K. 2005. Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. *Genes Dev* 19:2619–2630. <http://dx.doi.org/10.1101/gad.1355605>.
11. Paes da Rocha R. 2008. Characterization of the SOS regulon of *Caulobacter crescentus*. *J Bacteriol* 190:1209–1218. <http://dx.doi.org/10.1128/JB.01419-07>.
12. Erill I, Escibano M, Campoy S, Barbe J. 2003. In silico analysis reveals substantial variability in the gene contents of the gamma proteobacteria LexA-regulon. *Bioinformatics* 19:2225–2236. <http://dx.doi.org/10.1093/bioinformatics/btg303>.
13. Erill I, Jara M, Salvador N, Escibano M, Campoy S, Barbe J. 2004. Differences in LexA regulon structure among *Proteobacteria* through in vivo assisted comparative genomics. *Nucleic Acids Res* 32:6617–6626. <http://dx.doi.org/10.1093/nar/gkh996>.
14. Cornish JP, Matthews F, Thomas JR, Erill I. 2012. Inference of self-regulated transcriptional networks by comparative genomics. *Evol Bioinform Online* 8:449–461. <http://dx.doi.org/10.4137/EBO.S9205>.
15. Sanchez-Alberola N, Campoy S, Barbe J, Erill I. 2012. Analysis of the SOS response of *Vibrio* and other bacteria with multiple chromosomes. *BMC Genomics* 13:58. <http://dx.doi.org/10.1186/1471-2164-13-58>.
16. Ulrich RL, Deshazer D, Kenny TA, Ulrich MP, Moravusova A, Opperman T, Bavari S, Bowlin TL, Moir DT, Panchal RG. 2013. Characterization of the *Burkholderia thailandensis* SOS response by using whole-transcriptome shotgun sequencing. *Appl Environ Microbiol* 79:5830–5843. <http://dx.doi.org/10.1128/AEM.00538-13>.
17. Erill I, Campoy S, Mazon G, Barbe J. 2006. Dispersal and regulation of an adaptive mutagenesis cassette in the bacteria domain. *Nucleic Acids Res* 34:66–77. <http://dx.doi.org/10.1093/nar/gkj412>.
18. Emerson D, Fleming EJ, McBeth JM. 2010. Iron-oxidizing bacteria: an environmental and genomic perspective. *Annu Rev Microbiol* 64:561–583. <http://dx.doi.org/10.1146/annurev.micro.112408.134208>.
19. Emerson D, Field E, Chertkov O, Davenport K, Goodwin L, Munk C, Nolan M, Woyke T. 2013. Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. *Evol Genomic Microbiol* 4:254. <http://dx.doi.org/10.3389/fmicb.2013.00254>.
20. Pruitt KD, Tatusova T, Maglott DR. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 33:D501–D504.
21. Kiliç S, White ER, Sagitova DM, Cornish JP, Erill I. 2014. CollecTF: a database of experimentally validated transcription factor-binding sites in bacteria. *Nucleic Acids Res* 42:D156–160. <http://dx.doi.org/10.1093/nar/gkt1123>.
22. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <http://dx.doi.org/10.1093/nar/25.17.3389>.
23. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. 2009. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 25:119–120. <http://dx.doi.org/10.1093/bioinformatics/btn578>.
24. Anzaldi LJ, Muñoz-Fernández D, Erill I. 2012. BioWord: a sequence manipulation suite for Microsoft Word. *BMC Bioinformatics* 13:124. <http://dx.doi.org/10.1186/1471-2105-13-124>.
25. Schneider TD. 1997. Information content of individual genetic sequences. *J Theor Biol* 189:427–441. <http://dx.doi.org/10.1006/jtbi.1997.0540>.
26. Abella M, Campoy S, Erill I, Rojo F, Barbe J. 2007. Cohabitation of two different lexA regulons in *Pseudomonas putida*. *J Bacteriol* 189:8855–8862. <http://dx.doi.org/10.1128/JB.01213-07>.
27. Bhargava N, Erill I. 2010. xFITOM: a generic GUI tool to search for transcription factor binding sites. *Bioinformatics* 26:49–50. <http://dx.doi.org/10.1093/bioinformatics/btq5049>.
28. Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34:W369–W373. <http://dx.doi.org/10.1093/nar/gkl198>.
29. Klucar L, Stano M, Hajduk M. 2010. phiSITE: database of gene regulation in bacteriophages. *Nucleic Acids Res* 38:D366–D370. <http://dx.doi.org/10.1093/nar/gkp911>.
30. De Jong A, Pietersma H, Cordes M, Kuipers OP, Kok J. 2012. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC Genomics* 13:299. <http://dx.doi.org/10.1186/1471-2164-13-299>.
31. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680. <http://dx.doi.org/10.1093/nar/22.22.4673>.
32. Lamers MH, Georgescu RE, Lee S-G, O'Donnell M, Kuriyan J. 2006. Crystal structure of the catalytic alpha subunit of *Escherichia coli* replicative DNA polymerase III. *Cell* 126:881–892. <http://dx.doi.org/10.1016/j.cell.2006.07.028>.
33. Zhang APP, Pigli YZ, Rice PA. 2010. Structure of the LexA-DNA complex and implications for SOS box measurement. *Nature* 466:883–886. <http://dx.doi.org/10.1038/nature09200>.
34. Hall TT. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95.
35. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <http://dx.doi.org/10.1101/gr.849004>.
36. Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302:205–217. <http://dx.doi.org/10.1006/jmbi.2000.4042>.
37. Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552. <http://dx.doi.org/10.1093/oxfordjournals.molbev.a026334>.
38. Gupta RS. 2000. The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiol Rev* 24:367–402. <http://dx.doi.org/10.1111/j.1574-6976.2000.tb00547.x>.
39. Mazon G, Campoy S, Erill I, Barbe J. 2006. Identification of the *Acidobacterium capsulatum* LexA box reveals a lateral acquisition of the *Alpha-proteobacteria* LexA gene. *Microbiology* 152:1109–1118. <http://dx.doi.org/10.1099/mic.0.28376-0>.
40. Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574. <http://dx.doi.org/10.1093/bioinformatics/btg180>.
41. Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12:357–358.
42. Emerson D, Rentz JA, Lilburn TG, Davis RE, Aldrich H, Chan C, Moyer CL. 2007. A novel lineage of proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. *PLoS One* 2:e667. <http://dx.doi.org/10.1371/journal.pone.0000667>.
43. Cambay G, Sanchez-Alberola N, Campoy S, Guerin E, Da Re S, Gonzalez-Zorn B, Ploy MC, Barbe J, Mazel D, Erill I. 2011. Prevalence of SOS-mediated control of integron integrase expression as an adaptive trait of chromosomal and mobile integrons. *Mob DNA 2:6*. <http://dx.doi.org/10.1186/1759-8753-2-6>.
44. Cornish JP, Sanchez-Alberola N, O'Neill PK, O'Keefe R, Gheba J, Erill I. 2014. Characterization of the SOS meta-regulon in the human gut microbiome. *Bioinformatics* 30:1193–1197. <http://dx.doi.org/10.1093/bioinformatics/btt753>.
45. Cune J, Cullen P, Mazon G, Campoy S, Adler B, Barbe J. 2005. The *Leptospira interrogans* LexA gene is not autoregulated. *J Bacteriol* 187:5841–5845. <http://dx.doi.org/10.1128/JB.187.16.5841-5845.2005>.
46. Sahota G, Stormo GD. 2010. Novel sequence-based method for identifying transcription factor binding sites in prokaryotic genomes. *Bioinformatics* 26:2672–2677. <http://dx.doi.org/10.1093/bioinformatics/btq501>.
47. Tapias A, Fernandez S, Alonso JC, Barbe J. 2002. *Rhodobacter sphaeroides* LexA has dual activity: optimizing and repressing *recA* gene tran-

- scription. *Nucleic Acids Res* 30:1539–1546. <http://dx.doi.org/10.1093/nar/30.7.1539>.
48. Smollett KL, Smith KM, Kahramanoglu C, Arnvig KB, Buxton RS, Davis EO. 2012. Global analysis of the regulon of the transcriptional repressor LexA, a key component of SOS response in *Mycobacterium tuberculosis*. *J Biol Chem* 287:22004–22014. <http://dx.doi.org/10.1074/jbc.M112.357715>.
 49. Fernandez de Henestrosa AR, Cune J, Erill I, Magnuson JK, Barbe J. 2002. A green nonsulfur bacterium, *Dehalococcoides ethenogenes*, with the LexA binding sequence found in gram-positive organisms. *J Bacteriol* 184: 6073–6080. <http://dx.doi.org/10.1128/JB.184.21.6073-6080.2002>.
 50. Mazon G, Lucena JM, Campoy S, Fernandez de Henestrosa AR, Candau P, Barbe J. 2004. LexA-binding sequences in Gram-positive and cyanobacteria are closely related. *Mol Genet Genomics* 271:40–49. <http://dx.doi.org/10.1007/s00438-003-0952-x>.
 51. Abella M, Erill I, Jara M, Mazon G, Campoy S, Barbe J. 2004. Widespread distribution of a *lexA*-regulated DNA damage-inducible multiple gene cassette in the *Proteobacteria* phylum. *Mol Microbiol* 54:212–222. <http://dx.doi.org/10.1111/j.1365-2958.2004.04260.x>.
 52. Jochmann N, Kurze A-K, Czaja LF, Brinkrolf K, Brune I, Huser AT, Hansmeier N, Puhler A, Borovok I, Tauch A. 2009. Genetic makeup of the *Corynebacterium glutamicum* LexA regulon deduced from comparative transcriptomics and in vitro DNA band shift assays. *Microbiology* 155:1459–1477. <http://dx.doi.org/10.1099/mic.0.025841-0>.
 53. Van der Veen S, van Schalkwijk S, Molenaar D, de Vos WM, Abee T, Wells-Bennik MHJ. 2010. The SOS response of *Listeria monocytogenes* is involved in stress resistance and mutagenesis. *Microbiology* 156:374–384. <http://dx.doi.org/10.1099/mic.0.035196-0>.
 54. Walter BM, Rupnik M, Hodnik V, Anderluh G, Dupuy B, Paulič N, Žgur-Bertok D, Butala M. 2014. The LexA regulated genes of the *Clostridium difficile*. *BMC Microbiol* 14:88. <http://dx.doi.org/10.1186/1471-2180-14-88>.
 55. Favre D, Viret JF. 1990. Nucleotide sequence of the *recA* gene of *Bordetella pertussis*. *Nucleic Acids Res* 18:4243. <http://dx.doi.org/10.1093/nar/18.14.4243>.
 56. Brent R, Ptashne M. 1981. Mechanism of action of the *lexA* gene product. *Proc Natl Acad Sci U S A* 78:4204–4208. <http://dx.doi.org/10.1073/pnas.78.7.4204>.
 57. Bertrand-Burggraf E, Hurstel S, Daune M, Schnarr M. 1987. Promoter properties and negative regulation of the *uvrA* gene by the LexA repressor and its amino-terminal DNA binding domain. *J Mol Biol* 193:293–302. [http://dx.doi.org/10.1016/0022-2836\(87\)90220-8](http://dx.doi.org/10.1016/0022-2836(87)90220-8).
 58. Quinones M, Kimsey HH, Ross W, Gourse RL, Waldor MK. 2006. LexA represses CTX ϕ transcription by blocking access of the alpha C-terminal domain of RNA polymerase to promoter DNA. *J Biol Chem* 281:39407–39412. <http://dx.doi.org/10.1074/jbc.M609694200>.
 59. Fonseca LS, da Silva JB, Milanez JS, Monteiro-Vitorello CB, Momo L, de Moraes ZM, Vasconcellos SA, Marques MV, Ho PL, da Costa RMA. 2013. *Leptospira interrogans* serovar Copenhageni harbors two *lexA* genes involved in SOS response. *PLoS One* 8:e76419. <http://dx.doi.org/10.1371/journal.pone.0076419>.
 60. Hedrich S, Schlömann M, Johnson DB. 2011. The iron-oxidizing proteobacteria. *Microbiology* 157:1551–1564. <http://dx.doi.org/10.1099/mic.0.045344-0>.
 61. Jara M, Nunez C, Campoy S, Fernandez de Henestrosa AR, Lovley DR, Barbe J. 2003. *Geobacter sulfurreducens* has two autoregulated *lexA* genes whose products do not bind the *recA* promoter: differing responses of *lexA* and *recA* to DNA damage. *J Bacteriol* 185:2493–2502. <http://dx.doi.org/10.1128/JB.185.8.2493-2502.2003>.
 62. Fernandez de Henestrosa AR, Cune J, Mazon G, Dubbels BL, Bazylnski DA, Barbe J. 2003. Characterization of a new LexA binding motif in the marine magnetotactic bacterium strain MC-1. *J Bacteriol* 185:4471–4482. <http://dx.doi.org/10.1128/JB.185.15.4471-4482.2003>.
 63. Campoy S, Salvador N, Cortes P, Erill I, Barbe J. 2005. Expression of canonical SOS genes is not under LexA repression in *Bdellovibrio bacteriovorus*. *J Bacteriol* 187:5367–5375. <http://dx.doi.org/10.1128/JB.187.15.5367-5375.2005>.
 64. Mazon G, Campoy S, Fernandez de Henestrosa AR, Barbe J. 2006. Insights into the LexA regulon of *Thermotogales*. *Antonie Van Leeuwenhoek* 90:123–137. <http://dx.doi.org/10.1007/s10482-006-9066-x>.
 65. Pougach K, Voet A, Kondrashov FA, Voordeckers K, Christiaens JF, Baying B, Benes V, Sakai R, Aerts J, Zhu B, Van Dijck P, Verstrepen KJ. 2014. Duplication of a promiscuous transcription factor drives the emergence of a new regulatory network. *Nat Commun* 5:4868. <http://dx.doi.org/10.1038/ncomms5868>.
 66. Oertel-Buchheit P, Lamerichs RM, Schnarr M, Granger-Schnarr M. 1990. Genetic analysis of the LexA repressor: isolation and characterization of LexA(Def) mutant proteins. *Mol Gen Genet* 223:40–48. <http://dx.doi.org/10.1007/BF00315795>.
 67. Otteleben G, Messori L, Rüterjans H, Kaptein R, Granger-Schnarr M, Schnarr M. 1991. 1H-NMR investigation of the interaction of the amino-terminal domain of the LexA repressor with a synthetic half-operator. *J Biomol Struct Dyn* 9:447–461. <http://dx.doi.org/10.1080/07391102.1991.10507928>.
 68. Thliveris AT, Mount DW. 1992. Genetic identification of the DNA-binding domain of *Escherichia coli* LexA protein. *Proc Natl Acad Sci U S A* 89:4500–4504. <http://dx.doi.org/10.1073/pnas.89.10.4500>.
 69. Groban ES, Johnson MB, Banky P, Burnett PG, Calderon GL, Dwyer EC, Fuller SN, Gebre B, King LM, Sheren IN, Von Mutius LD, O'Gara TM, Lovett CM. 2005. Binding of the *Bacillus subtilis* LexA protein to the SOS operator. *Nucleic Acids Res* 33:6287–6295. <http://dx.doi.org/10.1093/nar/gki939>.
 70. Podgornaia AI, Casino P, Marina A, Laub MT. 2013. Structural basis of a rationally rewired protein-protein interface critical to bacterial signaling. *Structure* 21:1636–1647. <http://dx.doi.org/10.1016/j.str.2013.07.005>.
 71. Rowland MA, Deeds EJ. 2014. Crosstalk and the evolution of specificity in two-component signaling. *Proc Natl Acad Sci U S A* 111:5550–5555. <http://dx.doi.org/10.1073/pnas.1317178111>.
 72. Podgornaia AI, Laub MT. Pervasive degeneracy and epistasis in a protein-protein interface. *Science* 347:673–677. <http://dx.doi.org/10.1126/science.1257360>.
 73. Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT. 2008. Rewiring the specificity of two-component signal transduction systems. *Cell* 133:1043–1054. <http://dx.doi.org/10.1016/j.cell.2008.04.040>.