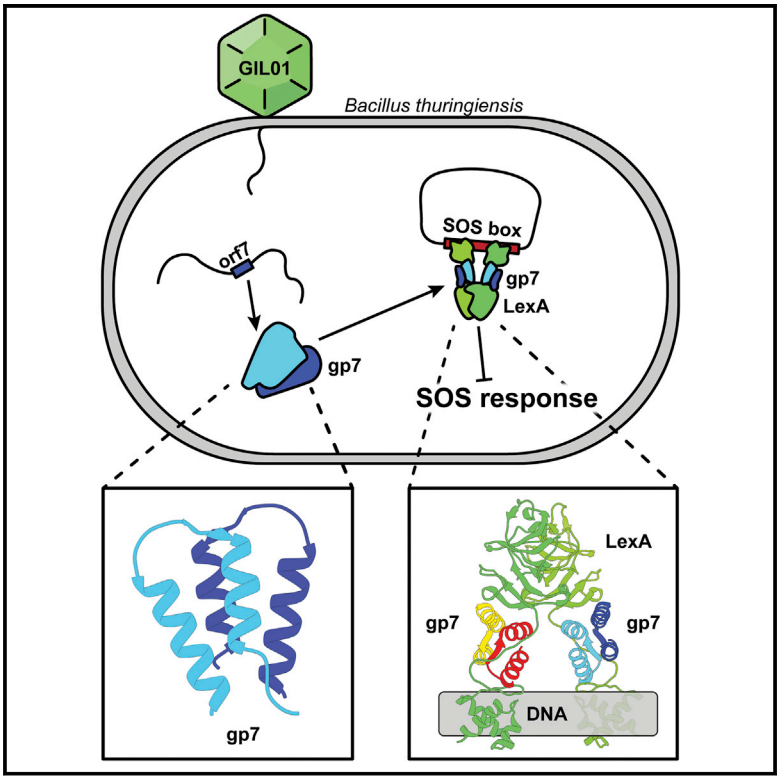
Article

Structure

Structural Insights into Bacteriophage GIL01 gp7 Inhibition of Host LexA Repressor

Graphical Abstract



Highlights

- Crystal structure of GIL01 gp7 has been solved
- A hybrid approach provides a model for gp7 scaffolding of LexA
- gp7 is seen to interact with phylogenetically distinct *Staphylococcus aureus* LexA
- Structural evidence of a phage factor associating with LexA to modulate the SOS response

Authors

Nathanael A. Caveney, Anja Pavlin, Guillermo Caballero, ..., Nadine Fornelos, Matej Butala, Natalie C.J. Strynadka

Correspondence

matej.butala@bf.uni-lj.si (M.B.), ncjs@mail.ubc.ca (N.C.J.S.)

In Brief

Developing a comprehensive understanding of the SOS response is of importance, as the response is proposed to play a key role in the development of antibacterial resistance. Here, Caveney et al. structurally characterize the bacteriophage protein gp7, which is seen to modulate the SOS response during host infection. Using a variety of techniques, a model for gp7 association and modulation of host LexA and SOS response is deduced.

Caveney et al., 2019, Structure *27*, 1094–1102
July 2, 2019
https://doi.org/10.1016/j.str.2019.03.019

CellPress

CellPress

Structure
Article

Structural Insights into Bacteriophage GIL01 gp7 Inhibition of Host LexA Repressor

Nathanael A. Caveney,[1] Anja Pavlin,[2] Guillermo Caballero,[1] Miha Bahun,[2] Vesna Hodnik,[2] Liza de Castro,[1] Nadine Fornelos,[3] Matej Butala,[2,*] and Natalie C.J. Strynadka[1,4,*]

[1]Department of Biochemistry and Molecular Biology and the Centre for Blood Research, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

[2]Department of Biology, Biotechnical Faculty, University of Ljubljana, Ljubljana 1000, Slovenia

[3]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[4]Lead Contact

*Correspondence: matej.butala@bf.uni-lj.si (M.B.), ncjs@mail.ubc.ca (N.C.J.S.)

https://doi.org/10.1016/j.str.2019.03.019

SUMMARY

Bacteria identify and respond to DNA damage using the SOS response. LexA, a central repressor in the response, has been implicated in the regulation of lysogeny in various temperate bacteriophages. During infection of *Bacillus thuringiensis* with GIL01 bacteriophage, LexA represses the SOS response and the phage lytic cycle by binding DNA, an interaction further stabilized upon binding of a viral protein, gp7. Here we report the crystallographic structure of phage-borne gp7 at 1.7-Å resolution, and characterize the 4:2 stoichiometry and potential interaction with LexA using surface plasmon resonance, static light scattering, and small-angle X-ray scattering. These data suggest that gp7 stabilizes LexA binding to operator DNA via coordination of the N- and C-terminal domains of LexA. Furthermore, we have found that gp7 can interact with LexA from *Staphylococcus aureus*, a significant human pathogen. Our results provide structural evidence as to how phage factors can directly associate with LexA to modulate the SOS response.

INTRODUCTION

The SOS response is a process by which bacteria can identify and respond to DNA damage. Typical to this system are two components, RecA recombinase and LexA repressor. Under regular growth conditions, LexA binds SOS-box sequences upstream of DNA repair genes to suppress the expression of these genes. RecA binds to and polymerizes on single-stranded DNA, which is a product of DNA damage. As a complex, it activates the autocleavage of LexA thereby lifting repression of SOS genes under DNA-damage conditions (Butala et al., 2009).

Due to the implications in the repair of damaged DNA, LexA and the SOS response are increasingly associated with the propagation of resistance phenotypes in response to antibacterial agents. It has been shown that LexA represses biochemical pathways that facilitate mutation and resistance (Blázquez et al., 2012). Also, LexA is associated with a variety of other functions unrelated to rescue from a DNA-damage phenotype, such as a role in the movement and regulation of mobile genetic elements (Fornelos et al., 2016) and virulence factors (Kelley, 2006). It is of note that recently LexA was shown to directly control the excision of the staphylococcal cassette chromosome conferring methicillin resistance, potentially resulting in horizontal gene transfer of these cassettes and the spread of antibiotic resistance (Liu et al., 2017). LexA has also been implicated in the regulation of lysogeny in various temperate bacteriophages that lack their own stress-inducible repressor (Fornelos et al., 2011, 2015; Úbeda et al., 2007).

Tectiviruses are bacteriophages that share structural similarities with several archaeal and eukaryotic viruses (Krupovic and Koonin, 2015). Tectiviruses are broadly distributed, having been identified in Gram-negative and Gram-positive hosts (Dariba et al., 2011; Jalasvuori and Koskinen, 2018). The tectiviral bacteriophage GIL01 is known to infect *Bacillus thuringiensis*, an insect pathogen of use in agriculture (Jouzani et al., 2017) and mosquito control (Pruszynski et al., 2017). *B. thuringiensis* belongs to the *Bacillus cereus sensu lato* group, which includes a variety of closely related *Bacillus* species (e.g., *Bacillus anthracis*), which are relevant to human health (Helgason et al., 2000). In *B. thuringiensis*, GIL01 phage is seen to produce a stable lysogenic state upon infection, switching to a lytic state when its host undergoes DNA damage (Verheust, 2003). This link between the transition to lytic state and DNA damage is characteristic of many temperate phages, such as the well-characterized λ phage (Gandon, 2016).

GIL01 has been shown to have a 15-kb linear double-stranded DNA genome with two major functional domains (Fornelos et al., 2011). The first domain contains the replication and regulatory genes of the genome, with transcriptional initiation at promoters *P1* and *P2*. The second domain contains the structural and lytic genes, with initiation at *P3* (Fornelos et al., 2011). Flanked by *P1* and *P2* is the canonical LexA-binding site, *dinBox1*, and an additional non-canonical site, *dinBox1b* (Fornelos et al., 2011, 2015). Two additional canonical sites, *dinBox2* and *dinBox3*, flank *P3* (Fornelos et al., 2011). LexA is seen to bind and repress expression through binding of *dinBox1/1b*, with LexA affinity for this region increased through interaction with a viral accessory protein, gp7 (Fornelos et al., 2015). gp7 interacts with both DNA bound and free LexA, with the latter interaction having an equilibrium

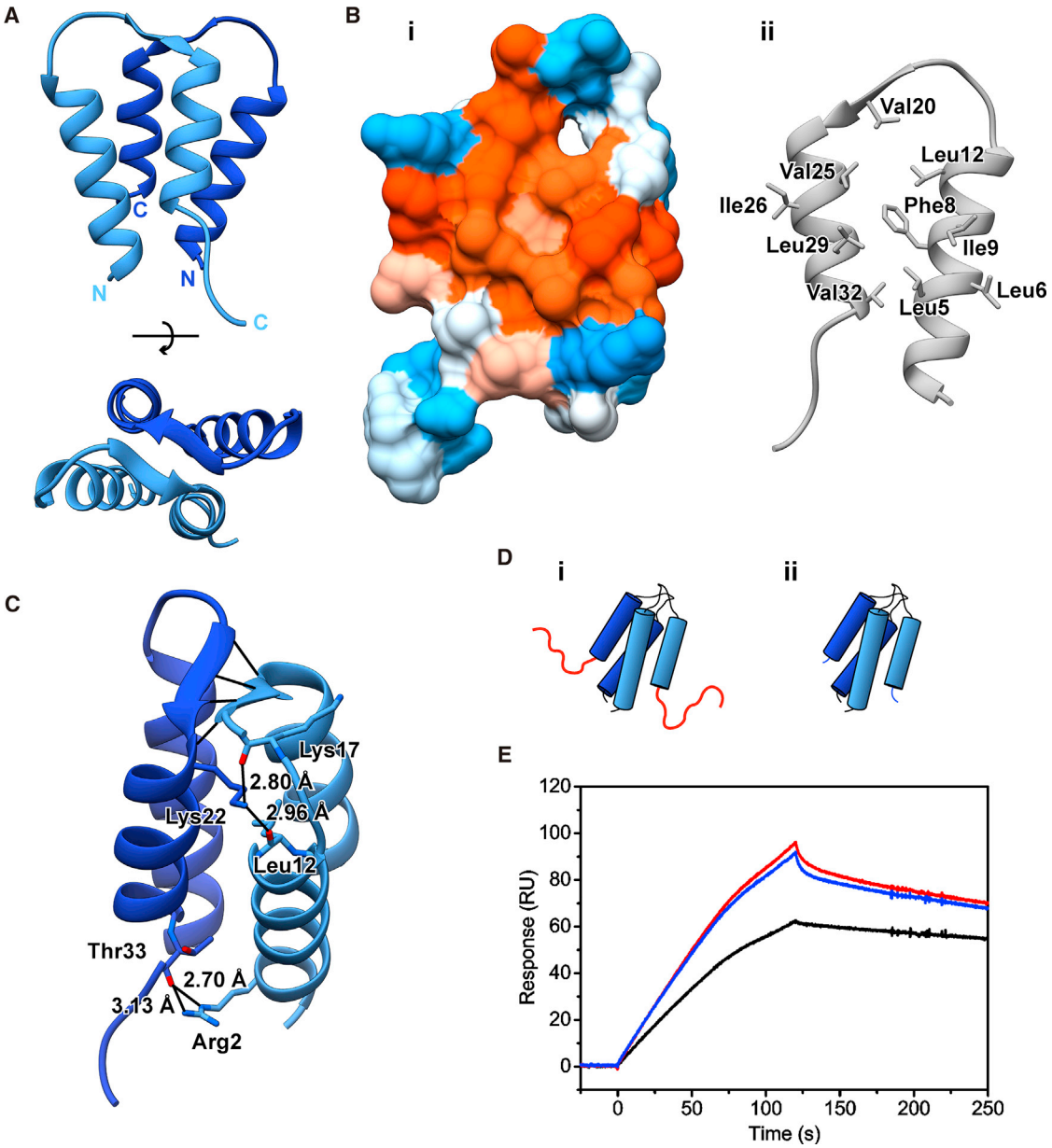


Figure 1. Crystal Structure of GIL01 gp7

(A–E) gp7 (A) is seen to form a dimeric 4-helix bundle in the crystal structure, with 19.5° helical packing between monomeric subunits and 27° packing within each monomer. (B) Hydrophobicity of gp7 dimer interface. A hydrophobicity surface representation is shown with hydrophobicity assigned based on the Kyte-Doolittle scale (Kyte and Doolittle, 1982), with residues ranging from hydrophilic (blue) to hydrophobic (orange) (i). The key hydrophobic residues contributing to the dimerization interface are shown on a monomer of gp7 (ii). (C) Dimerization of β-sheet and key hydrogen bonds of arginine 2 and lysine 22 across the dimer interface. The residues of the unstructured gp7 C-terminal region are not essential for the interaction with LexA. (E) The SPR sensorgram of an experiment investigating the interaction of either His_6-LexA (10 nM) alone (black line) or preincubated with either 100 nM wild-type His_6-gp7 (Di) (red line) or the derivative His_6-gp7Δ45–50 (Dii) (blue line) with the chip-immobilized *dinBox*1/1b DNA fragment. Proteins were injected for 120 s at a flow rate of 50 μL/min.

dissociation constant in the nanomolar range (Fornelos et al., 2015). gp7 is seen to be crucial for proper establishment of a lysogenic state *in vivo*, with mutations to gp7 often resulting in a lytic phenotype (Fornelos et al., 2011, 2015).

In this work, we report the structure of phage-borne gp7 and describe its interaction with host LexA. The crystal structure of a gp7 dimer was solved to 1.7-Å resolution and shows that a 4-helix bundle is formed at the dimeric interface. The C-terminal 13 residues were not clearly resolved in the crystal structure and are likely unstructured elements. Removal of the last six C-terminal residues of gp7 did not prevent its interaction with LexA. We propose a model of LexA-gp7 interaction in which gp7 acts as a small scaffold to orient the N-terminal and C-terminal domains of LexA such that the binding affinity to DNA is increased. The

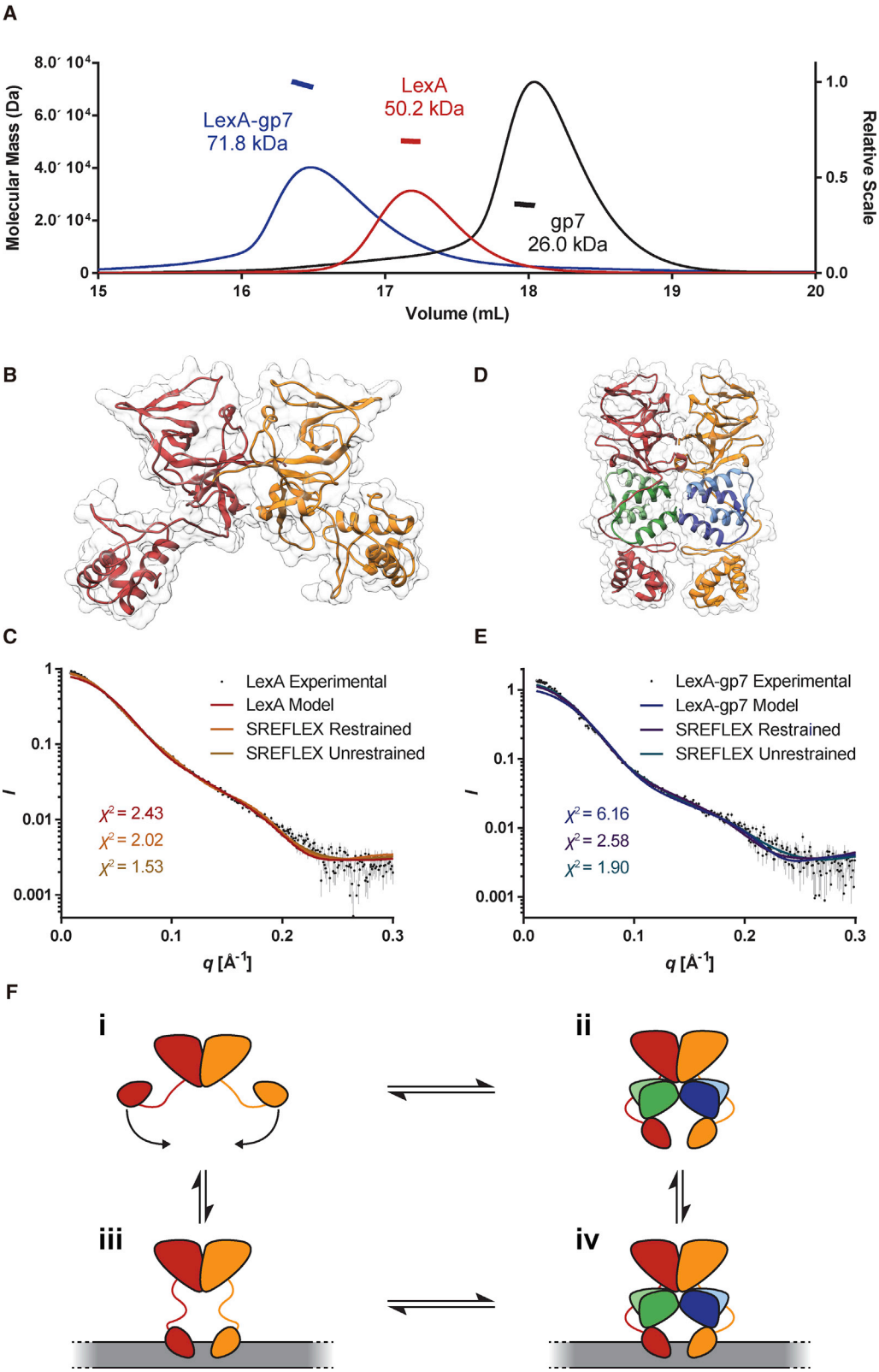


Figure 2. SEC-MALS, SAXS Analysis, and Modeling of LexA and LexA-gp7 Complex

For SEC-MALS, each protein or protein complex was run at a concentration between 2 and 6 mg/mL, and the horizontal line corresponds to the calculated molecular weight. LexA-gp7 heterohexamer is blue, LexA dimer red, and gp7 tetramer black (A). A model of LexA generated with I-TASSER (Yang et al., 2014),

(legend continued on next page)

stoichiometry of the LexA-gp7 interaction was determined as a heterohexameric complex of four gp7 units interacting with a LexA dimer. We assessed the ability of gp7 to interact with LexA proteins from four phylogenetically distinct bacteria and found that gp7 also interacts with LexA from the human pathogen, *Staphylococcus aureus*. Our results provide structural evidence of how a factor can directly associate with LexA to modulate the SOS response and, in turn, affect potential LexA-mediated antimicrobial resistance effects.

RESULTS AND DISCUSSION

LexA-gp7 Interaction Forms a Heterohexamer with 2:4 Stoichiometry

We have shown previously by surface plasmon resonance (SPR) that LexA and gp7 form a complex, but the stoichiometry of this interaction remained unknown (Fornelos et al., 2015). To determine this fundamental property of the interaction, we collected size-exclusion chromatography-multiangle light scattering (SEC-MALS) data on *B. thuringiensis* LexA, GIL01 gp7, and the LexA-gp7 complex. Analysis of the SEC-MALS results revealed that the 6-kDa gp7 is present in solution as a tetramer of 26 kDa (Figure 2). Similarly, the 26-kDa LexA was found as a dimer of 50 kDa and, when applied together, a monodisperse mixture of gp7 and LexA appeared as a 72-kDa heterohexameric complex composed of a tetrameric gp7 and dimeric LexA. The LexA repressor protein is well known to form a dimer in the literature (Luo et al., 2001), and *Bacillus* LexA is no exception. The formation of a tetrameric gp7 is of interest, as this observation raises the question as to whether the oligomerized gp7 tetramer binds LexA in a symmetric or asymmetric fashion or whether the tetramer dissociates to a pair of dimers with each binding a monomer of LexA.

GIL01 gp7 Dimer Forms a 4-Helix Bundle in the Crystal Structure

To further investigate the LexA-gp7 interaction, attempts were made to elucidate the crystal structures of GIL01 gp7, *B. thuringiensis* LexA, and the corresponding LexA-gp7 complex. A crystal of GIL01 gp7 was obtained with $P2_12_12_1$ symmetry, dimensions of $a = 39.69$ Å, $b = 44.70$ Å, $c = 46.30$ Å, and a diffraction resolution of 1.7 Å. The crystal was phased on a single-wavelength anomalous dispersion dataset of the same crystal, using iodine incorporated from the mother liquor. The resulting maps showed well-resolved electron density for the N-terminal 37 residues of gp7, which formed a dimer in the crystal structure. After refinement, the R_{work} for the structure was 18.4% and R_{free} was 21.8% (Table S1). The C-terminal 13 residues were seen to extend into the solvent channel of the crystal, but they could not be fully resolved due to local disorder in the electron density at the peripheral end. The majority of the structured residues form a tight α-helical bundle (Figure 1A) that resembles the dimerization domains of rabies virus phosphoprotein (PDB: 3L32 [Ivanov et al., 2010]) and human centromere protein B (PDB: 1UFI [Tawaramoto et al., 2003]) based on the results of a structural homology comparison with the server, Dali (Holm and Rosenström, 2010). This bundle conforms to the typical 4-helical bundle motif with 19.5° helical packing across the dimer and 27° packing among the monomer helices, as determined by UCSF Chimera Structure Measurements tool (Pettersen et al., 2004). Evidence suggests that the dimer seen in the crystal structure is indeed a physiologically relevant form of the protein. The interfacial surface area of the interaction is ~940 Å^2, which is approximately one-third of the 3,281 Å^2 total surface area of the structured portion of a gp7 monomer (surface area predicted by PISA [Krissinel and Henrick, 2007]). Furthermore, the interaction interface is largely hydrophobic at its core due to the packing of ten hydrophobic residues from each monomer (Figure 1B). Around the periphery of the interaction interface, there is also a number of stabilizing hydrogen-bond interactions. Residues 19 through 22 of one monomer form a β sheet with residues 22 through 19 of the adjacent monomer. In addition, the guanidinium group of Arg2 forms hydrogen bonds to the backbone carbonyl of Thr33 in the adjacent monomer, while Lys22 forms hydrogen bonds with the backbone carbonyls of Leu12 and Lys17 (Figure 1C). No evidence of a symmetry-derived tetramer could be found when looking beyond the asymmetric unit. Rather, gp7 dimers form the crystal lattice in a repetitive side-to-side fashion, with the interface between asymmetric units unlikely to be the tetramerization interface given the low interfacial surface area (553 Å^2, as predicted by PISA [Krissinel and Henrick, 2007]).

The C-Terminal Region of gp7 Is Unstructured

Because the 13 C-terminal residues of gp7 could not be fit reliably due to local disorder in this region of the electron density maps, secondary structure prediction and circular dichroism (CD) spectroscopy were used to assess the type of secondary structure therein. Secondary structure prediction using Jpred 4 (Drozdetskiy et al., 2015) provides little consensus on the region in question, with hidden Markov modeling predicting no structure for all 13 residues, while position-specific scoring matrices predict helical character for the first 7 of 13 residues (Figure S1). This results in overall very low (0–2 out of 10) Jpred confidence scores for the seven residues, with higher confidence that the terminal six are unstructured. The experimental CD spectrum is illustrated in Figures S1B and S1C, together with the predicted spectra of

SASREF (Petoukhov and Svergun, 2005), and ModLoop (Fiser and Sali, 2003) is shown in (B). The model is seen to show the dimeric form of LexA as predicted by SEC-MALS. The N-terminal domains are seen to be splayed outward from a DNA-binding conformation on the flexible linker regions between N- and C-terminal domains. The computed SAXS data of the model, in red, is fit to the experimental LexA data (40 and 5 mg/mL merged), in black (C). A model of the LexA-gp7 heterohexameric complex generated with the solved gp7 dimer structure, I-TASSER (Yang et al., 2014), ClusPro (Kozakov et al., 2017), SAXS guided modeling, and ModLoop (Fiser and Sali, 2003) is shown in (D). The model shows the structured elements of gp7 (blue-green) forming a scaffold between the N- and C-terminal domains of LexA (red-orange). The N-terminal domains of LexA are oriented in a similar fashion to what would be expected for DNA binding. The computed SAXS data of the model, in blue, is fit to the experimental LexA-gp7 data (12 and 0.75 mg/mL merged), in black (E). To account for flexibility, two additional curves generated with SREFLEX are fit to the experimental data for both models (C and E). χ^2 values are reported for all curves. (F) A schematic representation of the potential binding cycle between LexA (i), LexA-gp7 complex (ii), DNA-LexA-gp7 complex (iv), and DNA-LexA complex (iii). LexA is shown in red-orange, gp7 in blue-green, and DNA in gray.

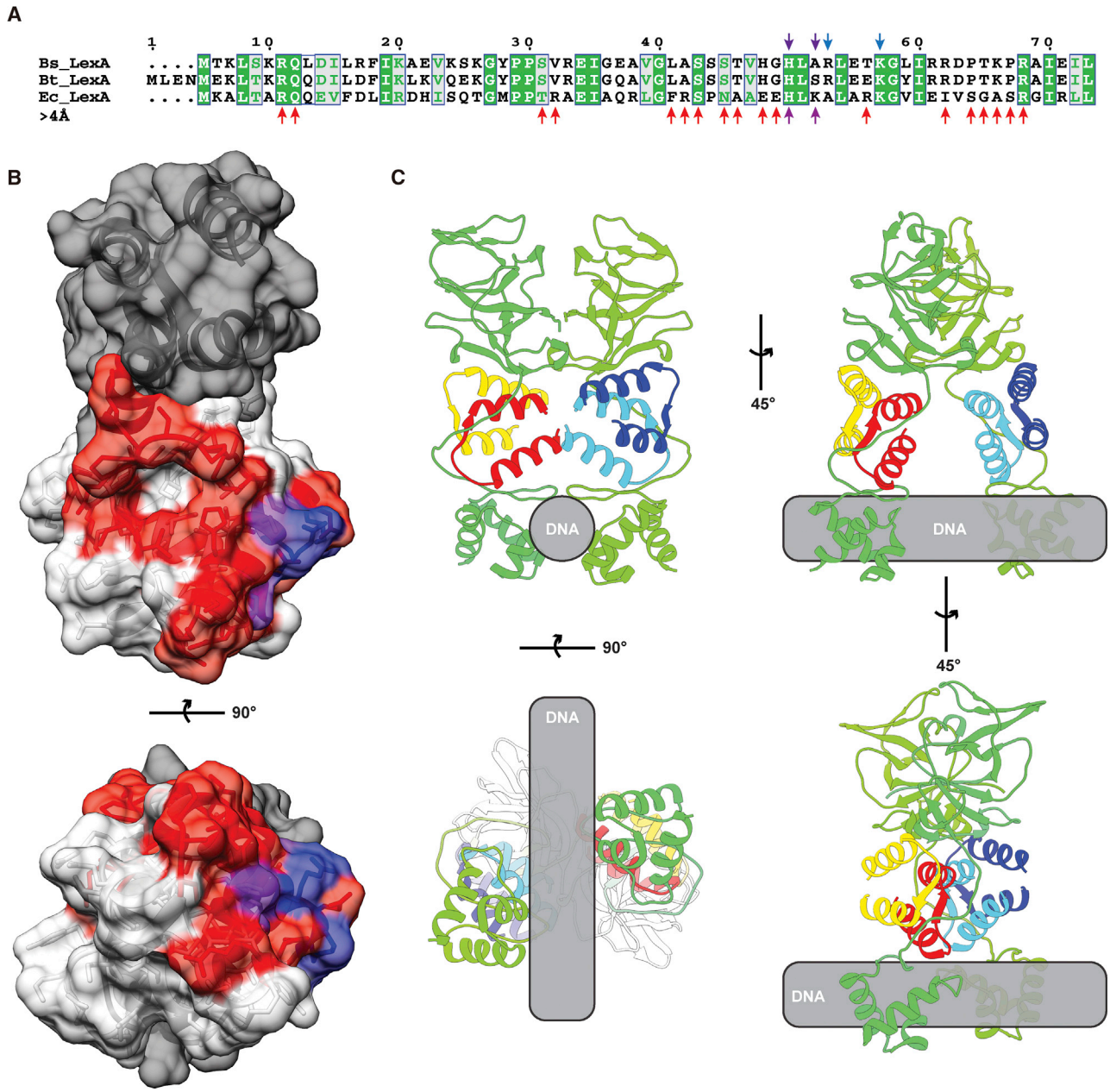


Figure 3. Surface-Exposed DNA-Binding Residues and Proposed DNA-Binding Mode

A sequence alignment of the N-terminal domains of *B. subtilis*, *B. thuringiensis*, and *E. coli* LexA is shown (A). Residues within 4 Å of DNA in *E. coli* LexA-DNA complex (PDB: 3JSO) are marked by red arrows. Residues that have been shown to mediate LexA-DNA interaction in *B. subtilis* are marked by blue arrows. Residues that fall under both aforementioned categories are marked by purple arrows. These surface-exposed potential DNA-binding residues of *B. thuringiensis* LexA are mapped (red, blue, and purple) onto the N-terminal domain (light gray) (B). GIL01 gp7 (dark gray) is not seen to occlude these residues in the proposed model. Proposed DNA-binding mode of the LexA-gp7 complex model is shown in (C), with LexA in green, gp7 in blue, red, and yellow, and DNA in gray.

hypothetical gp7 structures modeled with a completely α-helical C-terminal tail in one case and a completely unstructured C terminus in the other. Comparison of these spectra reveal the collected gp7 spectrum correlates with the calculated unstructured spectrum to a far greater extent, with a normalized root-mean-square deviation (NRMSD) of 0.238 for the spectra between 195 and 240 nm and an NRMSD of 0.029 between 210 and 240 nm, a region known to contain α-helical character (Holzwarth and Doty, 1965). The spectra compared with the α-helical calculated spectrum had NRMSD values of 0.487 and 0.121 for the same regions. NRMSD values between replicate experimental collections were 0.044 and 0.029 for the full and truncated spectra, respectively. Despite the lack of structured elements, it is known that the C terminus of gp7 is important in the development of a

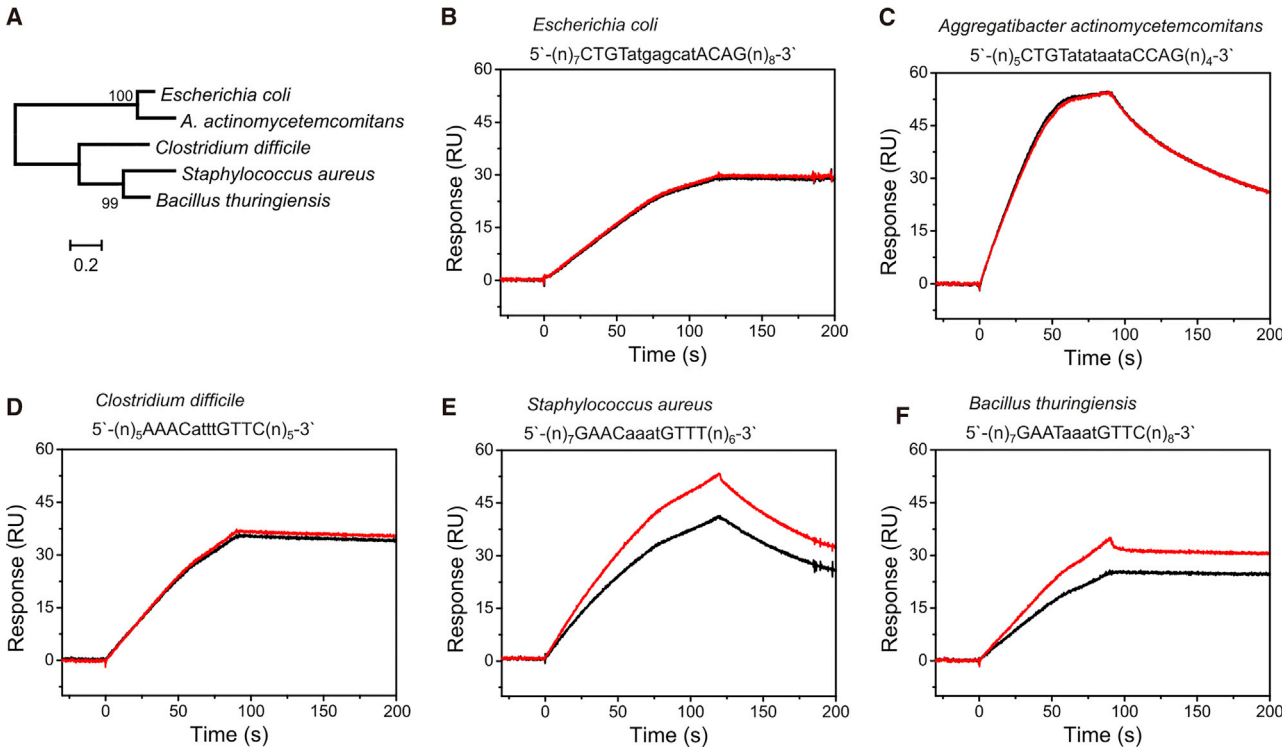


Figure 4. The GIL01 Bacteriophage gp7 Protein Interacts with *Staphylococcus aureus* LexA Bound to Target DNA

(A) Phylogenetic tree of five LexA protein sequences analyzed using the maximum-likelihood principle (PhyML). Significant bootstrapping values (%) are shown on the nodes. The scale bar corresponds to 0.2 change per amino acid.

(B–F) SPR sensorgrams showing the interaction of His_6-LexA repressors (10 nM) either alone (black) or combined with His_6-gp7 (100 nM, red), with a chip-immobilized target DNA fragment (~50 response units [RU]) for 90–120 s at a flow rate of 50 μL/min. Source organisms of the LexA repressor and the *lexA* operators of *A. actinomycetemcomitans* (C), *C. difficile* (D), *S. aureus* (E), or the *recA* operators of E. coli (B) or *B. thuringiensis* (F) are presented above the sensorgrams. Three independent measurements were performed and representative sensorgrams are shown.

lysogenic state, as various truncation mutants to this region result in a lytic phenotype (Fornelos et al., 2011). It may be that this region is typical of other intrinsically disordered regions that adopt a structured state only upon binding to cognate interaction partners. To explore the significance of the gp7 C-terminal region, we prepared a gp7 derivative lacking residues 45–50, and the influence of this deletion on complex formation of LexA and gp7 was assessed by SPR. This analysis revealed that at least six residues of the gp7 C terminus are not essential for the formation of the complex (Figure 1E).

gp7 Acts As a Scaffold to Orient LexA N- and C-Terminal Domains for DNA Binding

Given the potential for understanding and preventing the spread of antimicrobial resistance, the molecular mechanisms by which gp7 acts to modulate *Bacillus* LexA are of interest to unravel at the atomic level. To develop a model of the LexA-gp7 interaction, we collected bioSAXS data on the *B. thuringiensis* LexA and LexA-gp7 complex. Atomistic and *ab initio* models of LexA and the LexA-gp7 complex were constructed that fit the experimental scattering data of each sample (Figures 2, S2, and S3). The atomistic model of LexA fits with the current understanding of the flexible nature of the linker region between the N- and C-terminal domains. The two DNA-binding N-terminal domains are seen to be folded upward toward the C-terminal dimerization and catalytic domains, similar to what we have previously observed in the first crystal structure of *Escherichia coli* LexA (Luo et al., 2001). It is likely that upon the presence of DNA, the *B. thuringiensis* LexA N-terminal domains would rotate downward on this linker region into a DNA-binding orientation similar to that seen by Zhang et al. (2010) in their *E. coli* LexA-DNA complex structures. The atomistic model of LexA-gp7 complex suggests that gp7 is interacting with both the N- and C-terminal domains of *B. thuringiensis* LexA, although the most ordered regions of the interaction appear to be largely C-terminal. In the bioSAXS-guided model, the N-terminal domains are reoriented in comparison with the aforementioned LexA dimer model. The unstructured C-terminal residues of gp7 could not be reliably modeled and are excluded in this analysis. In the model, the N-terminal domains of LexA are in closer proximity to one another and in an orientation similar to what one would expect for DNA binding. Although no structures of a Gram-positive LexA or operator have yet been determined, it is predicted that the architecture of *Bacillus* SOS "boxes" and dinBox1/1b of GIL01 will differ from the canonical *E. coli* SOS box observed in the only prior LexA/operator crystal structures (Erill et al., 2004). Based on SOS-box sequence comparisons, there is expected to be a significant rotational variation between *E. coli* and *B. thuringiensis*

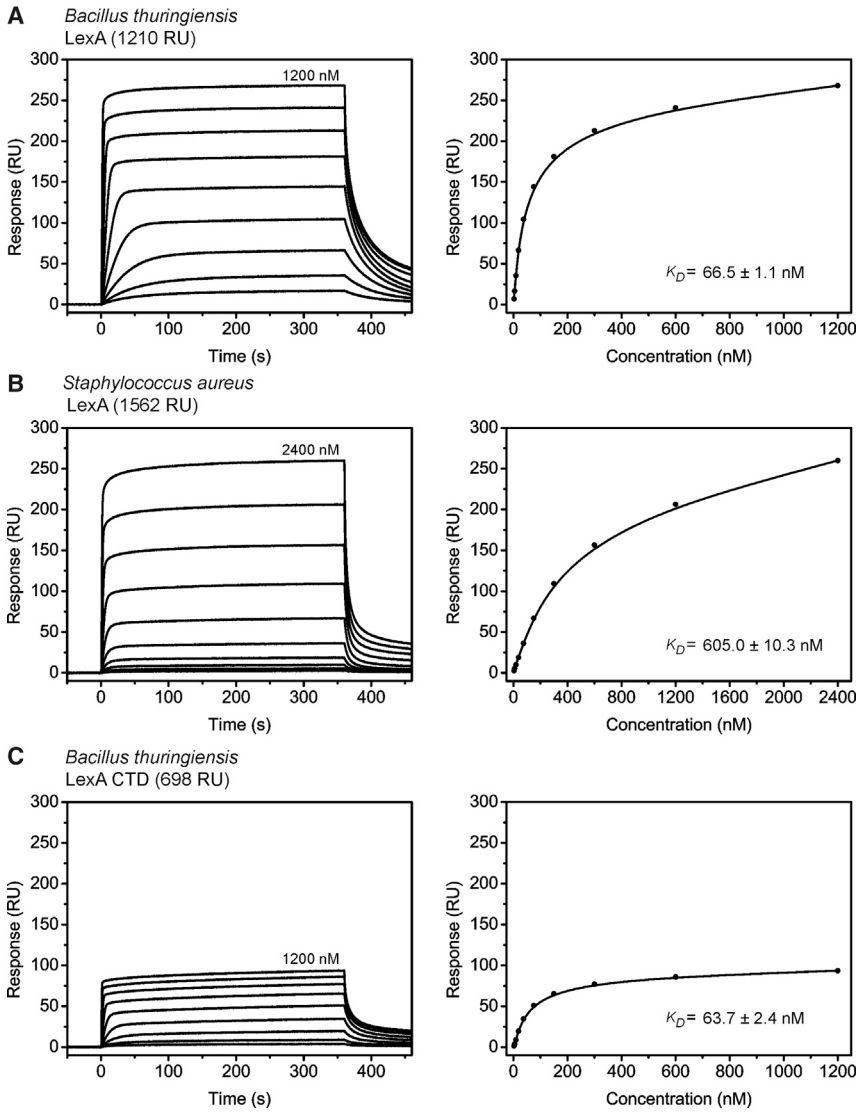


Figure 5. gp7 Interacts with *B. thuringiensis* LexA C-Terminal Domain and with Free *S. aureus* LexA Repressor

SPR analysis of the His_6-gp7 interaction with the immobilized (A) full-length *B. thuringiensis* His_6-LexA, (B) its His_6-CTD, or (C) *S. aureus* His_6-LexA. The amount of each sensor surface-immobilized LexA protein is indicated in response units (RU). His_6-gp7 was injected across immobilized His_6-LexA proteins in 2-fold serial dilutions ranging from 2,400 or 1,200 to 2.35 nM for 360 s at a rate of 30 μL/min and dissociation was followed for 120 s. Sensorgrams are shown along with the apparent equilibrium dissociation constant (K_D) determined from the response curves as a function of the His_6-gp7 concentration injected across His_6-LexA. K_D values are the mean ± standard deviation of three titrations of each protein.

DNA due to a 4-nt length difference between the two SOS sites. Our proposed model places the two N-terminal domains in an orientation that would be more representative of a proposed GIL01 dinBox and Gram-positive SOS-box binding. We note that in the model, gp7 binds directly at the interface between both C- and N-terminal domains but does not occlude predicted surface-exposed DNA-binding residues of the N-terminal domain (Groban et al., 2005; Zhang et al., 2010) (Figure 3). It is possible that the gp7 binding and subsequent reorientation of these N-terminal domains primes LexA for DNA binding, reducing the entropic penalty of DNA-LexA binding and resulting in the increased affinity for DNA seen previously (Fornelos et al., 2015).

GIL01 gp7 Interacts with LexA from *S. aureus*

Considering that the gp7 interaction with *B. thuringiensis* LexA prevents LexA cleavage (Fornelos et al., 2015), it was of interest to investigate whether gp7 can form a complex with LexA proteins from unrelated bacteria. Four species of medical relevance were selected for investigation: the Gram-positive pathogens *Clostridium difficile* and *S. aureus* as well as the Gram-negative *E. coli* and periodontopathogen *Aggregatibacter actinomycetemcomitans*. The sequence alignment of the different LexA proteins showed the highest sequence identity between *B. thuringiensis* and *S. aureus* LexA repressors (65.1% ID, Figures 4A and S4). The respective LexA proteins were expressed and purified and the potential interactions between gp7 and the different LexA proteins bound to their cognate DNAs were assessed using SPR. In accordance with the highest level of LexA sequence conservation, gp7 interacted with *S. aureus* LexA (Figures 4B–4F). Despite the conserved DNA-binding motif across Gram-positive bacteria (CGAAC-N_4-GTTYC) (Erill et al., 2004), this result shows that this conservation of binding is primarily dependent on LexA sequence conservation, as opposed to SOS site conservation. The interaction was confirmed to be dose dependent by using increasing concentrations of gp7 in presence of DNA-free LexA (Figures 5A and 5B), further showing the stability of the interaction over time. We determined that gp7 forms a moderate affinity complex with *S. aureus* LexA, exhibiting an apparent equilibrium dissociation constant (K_D) of 605 nM. In contrast, the gp7 interaction with LexA homologs from *C. difficile* and *A. actinomycetemcomitans* is negligible (Figure S5). Interaction with *E. coli* LexA exhibits a K_D in the micromolar range (Figure S5). To narrow down the sequence requirements of a LexA-gp7 interaction, we expressed and purified the C-terminal domain of *B. thuringiensis* LexA and analyzed its interaction with gp7. gp7 interacted with LexA CTD in similar fashion as it did with full-length LexA (Figure 5C).

Conclusion

GIL01 gp7 plays a crucial role in the modulation of the SOS response through its interaction with LexA. In the absence of a functional gp7, phage GIL01 is no longer able to establish

lysogeny (Fornelos et al., 2011). Our results provide a structural basis for the ability of gp7 to interact with LexA, to act as a scaffold to orient domain architecture of LexA, and to increase the overall affinity of LexA-DNA binding. Study of the modulation of LexA and the SOS response through cofactor binding is of importance as the spread of antibiotic-resistant bacteria increases. Modulation of the SOS response not only provides an attractive target for coupled therapeutics but can prevent the development and spread of further resistance phenotypes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Cloning and Protein Expression
  - Protein Purification
  - Surface Plasmon Resonance Assays
  - Phylogenetic Analysis
  - Crystallography and Determination of the gp7 Structure
  - Circular Dichroism Spectroscopy
  - SEC-MALS Analysis of Quaternary Structure
  - bioSAXS Analysis of Quaternary Structure and bioSAXS Guided Complex Modelling
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2019.03.019.

ACKNOWLEDGMENTS

We thank Fred Rossell for advice on CD spectroscopy, Cene Gostinčar for advice on phylogenetic analysis, and Michael Krisinger for advice on SPR. Research described in this paper was performed using beamline 08ID-1 at the Canadian Light Source, which is supported by the Canada Foundation for Innovation (CFI), Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research (CIHR). This work was funded by operating grants to N.C.J.S. from CIHR and the Howard Hughes International Senior Scholar program and to M. Butala and A.P. from the Slovenian Research Agency (J1-8150 and P1-0207) and infrastructure funding from the CFI and British Columbia Knowledge Development Fund. N.A.C. holds a Center for Blood Research Graduate Award and N.C.J.S. is a Tier I Canada Research Chair in Antibiotic Discovery.

AUTHOR CONTRIBUTIONS

Conceptualization, N.A.C., N.F., M. Butala, and N.C.J.S.; Methodology, N.A.C., A.P., V.H., N.F., M. Butala, and N.C.J.S.; Investigation, N.A.C., A.P., G.C., M. Bahun, V.H., and M. Butala; Resources, L.d.C.; Writing – Review & Editing, N.A.C., N.F., M. Butala, and N.C.J.S.; Supervision, N.F., M. Butala, and N.C.J.S.; Funding Acquisition, M. Butala and N.C.J.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 14, 2018
Revised: March 4, 2019
Accepted: March 25, 2019
Published: May 2, 2019

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Blázquez, J., Couce, A., Rodríguez-Beltrán, J., and Rodríguez-Rojas, A. (2012). Antimicrobials as promoters of genetic variation. Curr. Opin. Microbiol. *15*, 561–569.

Butala, M., Zgur-Bertok, D., and Busby, S.J.W. (2009). The bacterial LexA transcriptional repressor. Cell. Mol. Life Sci. *66*, 82–93.

Butala, M., Klose, D., Hodnik, V., Rems, A., Podlesek, Z., Klare, J.P., Anderluh, G., Busby, S.J.W., Steinhoff, H.J., and Žgur-Bertok, D. (2011). Interconversion between bound and free conformations of LexA orchestrates the bacterial SOS response. Nucleic Acids Res. *39*, 6546–6557.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics *27*, 1164–1165.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res. *43*, W389–W394.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Erill, I., Jara, M., Salvador, N., Escribano, M., Campoy, S., and Barbé, J. (2004). Differences in LexA regulon structure among Proteobacteria through in vivo assisted comparative genomics. Nucleic Acids Res. *32*, 6617–6626.

Fiser, A., and Sali, A. (2003). ModLoop: automated modeling of loops in protein structures. Bioinformatics *19*, 2500–2501.

Fodje, M., Grochulski, P., Janzen, K., Labiuk, S., Gorin, J., and Berg, R. (2014). 08B1-1: an automated beamline for macromolecular crystallography experiments at the Canadian Light Source. J. Synchrotron Radiat. *21*, 633–637.

Fornelos, N., Bamford, J.K.H., and Mahillon, J. (2011). Phage-borne factors and host LexA regulate the lytic switch in phage GIL01. J. Bacteriol. *193*, 6008–6019.

Fornelos, N., Butala, M., Hodnik, V., Anderluh, G., Bamford, J.K., and Salas, M. (2015). Bacteriophage GIL01 gp7 interacts with host LexA repressor to enhance DNA binding and inhibit RecA-mediated auto-cleavage. Nucleic Acids Res. *43*, 7315–7329.

Fornelos, N., Browning, D.F., and Butala, M. (2016). The use and abuse of LexA by mobile genetic elements. Trends Microbiol. *24*, 391–401.

Gandon, S. (2016). Why be temperate: lessons from bacteriophage λ. Trends Microbiol. *24*, 356–365.

Groban, E.S., Johnson, M.B., Banky, P., Burnett, P.G.G., Calderon, G.L., Dwyer, E.C., Fuller, S.N., Gebre, B., King, L.M., Sheren, I.N., et al. (2005). Binding of the *Bacillus subtilis* LexA protein to the SOS operator. Nucleic Acids Res. *33*, 6287–6295.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Helgason, E., Økstad, O.A., Dominique, A., Johansen, H.A., Fouet, A., Hegna, I., Kolstø, A., Økstad, O.L.E.A., Caugant, D.A., Mock, L.E., et al. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*—one species on the basis of genetic evidence one species on the basis of genetic evidence. Appl. Environ. Microbiol. *66*, 2627–2630.

Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38*, 545–549.

Holzwarth, G., and Doty, P. (1965). The ultraviolet circular dichroism of polypeptides. J. Am. Chem. Soc. *87*, 218–228.

Ivanov, I., Crepin, T., Jamin, M., and Ruigrok, R.W.H. (2010). Structure of the dimerization domain of the rabies virus phosphoprotein. J. Virol. *84*, 3707–3710.

Jalasvuori, M., and Koskinen, K. (2018). Extending the hosts of Tectiviridae into four additional genera of Gram-positive bacteria and more diverse *Bacillus* species. Virology *518*, 136–142.

Jouzani, G.S., Valijanian, E., and Sharafi, R. (2017). *Bacillus thuringiensis*: a successful insecticide with new environmental features and tidings. Appl. Microbiol. Biotechnol. *101*, 2691–2711.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

Kelley, W.L. (2006). Lex marks the spot: the virulent side of SOS and a closer look at the LexA regulon. Mol. Microbiol. *62*, 1228–1238.

Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., Beglov, D., and Vajda, S. (2017). The ClusPro web server for protein-protein docking. Nat. Protoc. *12*, 255–278.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Krupovic, M., and Koonin, E.V. (2015). Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. Nat. Rev. Microbiol. *13*, 105–115.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. *157*, 105–132.

Liu, P., Wu, Z., Xue, H., and Zhao, X. (2017). Antibiotics trigger initiation of SCCmectransfer by inducing SOS responses. Nucleic Acids Res. *45*, 3944–3952.

Luo, Y., Pfuetzner, R.A., Mosimann, S., Paetzel, M., Frey, E.A., Cherney, M., Kim, B., Little, J.W., and Strynadka, N.C.J. (2001). Crystal structure of LexA: a conformational switch for regulation of self-cleavage. Cell *106*, 585–594.

Mao, D., Wachter, E., and Wallace, B.A. (1982). Folding of the mitochondrial proton adenosinetriphosphatase proteolipid channel in phospholipid vesicles. Biochemistry *21*, 4960–4968.

Mavridis, L., and Janes, R.W. (2017). PDB2CD: a web-based application for the generation of circular dichroism spectra from protein atomic coordinates. Bioinformatics *33*, 56–63.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Obradović, D., Gašperšič, R., Caserman, S., Leonardi, A., Jamnik, M., Podlesek, Z., Seme, K., Anderluh, G., Križaj, I., Maček, P., et al. (2016). A cytolethal distending toxin variant from aggregatibacter actinomycetemcomitans with an aberrant CdtB that lacks the conserved catalytic histidine 160. PLoS One *11*, 1–16.

Panjkovich, A., and Svergun, D.I. (2016). Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. Phys. Chem. Chem. Phys. *18*, 5707–5719.

Petoukhov, M.V., and Svergun, D.I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys. J. *89*, 1237–1250.

Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. *45*, 342–350.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Pruszynski, C.A., Hribar, L.J., Mickle, R., and Leal, A.L. (2017). A large scale biorational approach using *Bacillus thuringiensis israeliensis* (strain AM65-52) for managing *Aedes aegypti* populations to prevent Dengue, Chikungunya and Zika transmission. PLoS One *12*, e0170079.

Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. *76*, 2879–2886.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. *28*, 768–773.

Tawaramoto, M.S., Park, S.Y., Tanaka, Y., Nureki, O., Kurumizaka, H., and Yokoyama, S. (2003). Crystal structure of the human centromere protein B (CENP-B) dimerization domain at 1.65-A resolution. J. Biol. Chem. *278*, 51454–51461.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2007). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D Biol. Crystallogr. *64*, 61–69.

Úbeda, C., Maiques, E., Tormo, M., Campoy, S., Lasa, Í., Barbé, J., Novick, R.P., and Penadés, J.R. (2007). SaPI operon I is required for SaPI packaging and is controlled by LexA. Mol. Microbiol. *65*, 41–50.

Verheust, C. (2003). pGIL01, a linear tectiviral plasmid prophage originating from *Bacillus thuringiensis* serovar *israelensis*. Microbiology *149*, 2083–2092.

Walter, B.M., Rupnik, M., Hodnik, V., Anderluh, G., Dupuy, B., Paulič, N., Žgur-Bertok, D., and Butala, M. (2014). The LexA regulated genes of the *Clostridium difficile*. BMC Microbiol. *14*, 88.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. Nat. Methods *12*, 7–8.

Zhang, A.P.P., Pigli, Y.Z., and Rice, P.A. (2010). Structure of the LexA-DNA complex and implications for SOS box measurement. Nature *466*, 883–886.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli BL21 (DE3) pLysE	NEB	C2527I
Chemicals, Peptides, and Recombinant Proteins		
Bacillus thuringiensis LexA	This study	N/A
GIL01 gp7	This study	N/A
Aggregatibacter actinomycetemcomitans LexA	This study	N/A
Staphylococcus aureus LexA	This study	N/A
Bacillus thuringiensis LexA CTD	This study	N/A
Escherichia coli LexA	This study	N/A
Clostridium difficile LexA	This study	N/A
HisTrap HP Sepharose cartridge	GE Lifesciences	17524801, 29051021
cOmplete protease inhibitor	Roche	4693116001
HIS-Select HF nickel affinity gel	Merck	H0537
Deposited Data		
gp7 atomic coordinates	This study	6N7O
LexA and LexA-gp7 SAXS data	This study	SASDEB8 and SASDEC8
LexA-DNA complex	Zhang et al., 2010	3JSO
Oligonucleotides		
Oligonucleotides used in this study are presented in Table S3		
Recombinant DNA		
pET-28a *B. thuringiensis lexA*	This study	N/A
pET-28a GIL01 *ORF7*	This study	N/A
pET-8c *A. actinomycetemcomitans lexA*	This study	N/A
pET-8c *S. aureus lexA*	This study	N/A
pET-8c GIL01 *ORF7Δ45-50*	This study	N/A
pET-8c *B. thuringiensis lexA CTD* (*G110-H223*)	This study	N/A
pQE-30 GIL01 *ORF7*	Fornelos et al., 2015	N/A
pQE-30 *B. thuringiensis lexA*	Fornelos et al., 2015	N/A
pET-8c *B. thuringiensis lexA*	Fornelos et al., 2015	N/A
pET-8c *C. difficile lexA*	Walter et al., 2014	N/A
pET-8c *E. coli lexA*	Butala et al., 2011	N/A
Software and Algorithms		
Biacore T100 Evaluation Software	GE Healthcare	N/A
PhyML	Guindon et al., 2010	N/A
ProtTest 3.2.1 3.4	Darriba et al., 2011	N/A
Autoprocess	Fodje et al., 2014	N/A
XDS	Kabsch, 2010	N/A
Phaser	McCoy et al., 2007	N/A
AutoBuild	Terwilliger et al., 2007	N/A
Phenix	Adams et al., 2010	N/A
Coot	Emsley and Cowtan, 2004	N/A
PDB2CD	Mavridis and Janes, 2017	N/A
ModLoop	Fiser and Sali, 2003	N/A
Astra6	Wyatt Technology	N/A
SAXSLab	Rigaku	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ATSAS	Petoukhov et al., 2012	N/A
I-TASSER	Yang et al., 2014	N/A
SASREF	Petoukhov and Svergun, 2005	N/A
ClusPro	Kozakov et al., 2017	N/A
CRYSOL	Svergun et al., 1995	N/A
DAMMIN	Svergun, 1999	N/A
SREFLEX	Panjkovich and Svergun, 2016	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, NCJS (ncjs@mail.ubc.ca).

METHOD DETAILS

The GIL01 host genome sequence used as the reference in this study is that of the *B. cereus* strain G9842 (GenBank: CP001186). We note however that GIL01 originates from a strain that was identified as *B. thuringiensis* based on presence of the plasmid that encodes insecticidal toxins (Verheust, 2003). *B. thuringiensis* and *B. cereus* are otherwise extremely close on a genomic level, with LexA sequences being 100% conserved amongst the two.

Cloning and Protein Expression

Full-length *B. thuringiensis lexA* (coordinates 3,624,082 through 3,624,714 in strain G9842) and GIL01 ORF7 (coordinates 4,564 through 4,716; GenBank: AJ536073) were cloned between restriction sites *Bam*HI and *Hin*dIII into the expression vector pET28a in a fusion with a thrombin-cleavable N-terminal hexahistidine tag. Expression constructs were transformed into *E. coli* BL21 (DE3) (NEB). Cells were cultured in ZYP-5052 autoinduction media for four hours at 37°C followed by overnight protein expression at 25°C.

The full-length *lexA* genes of *Aggregatibacter actinomycetemcomitans* (coordinates 658956 through 659576; GenBank: CP003099.1), *Staphylococcus aureus* and the carboxy-terminal domain of *B. thuringiensis lexA*, containing residues from the glycine (G110) of the cleavage site to the terminal histidine (H223) (UniProt ID: B7ISN2_BAC2), were PCR amplified from the genomic DNAs of the serotype b *A. actinomycetemcomitans* strain (Obradović et al., 2016), *Staphylococcus aureus* subsp. *aureus* Rosenbach ATCC 29213 (coordinates 8738 through 9358; GenBank: LHUS02000114.1) and from *B. thuringiensis* (Fornelos et al., 2015), respectively. To prepare a gp7 derivative lacking residues glutamate (E45) to the terminal residue aspartate (D50), primers gp7Δ45-50_F and gp7Δ45-50_R were used to PCR amplify the truncated gp7 gene from *B. thuringiensis*. PCR products were cloned between the *Bam*HI and *Mlu*I restriction sites of expression vector pET8c in a fusion with a thrombin-cleavable N-terminal hexahistidine tag (Novagen). The generated plasmid constructs and the plasmid to overexpress the *C. difficile* 630 LexA (Walter et al., 2014) and the *E. coli* K12 LexA (Butala et al., 2011) were transformed into *E. coli* BL21 (DE3) pLysE (NEB). Strains were grown in 500 mL L-broth to an OD_{600} of 0.6 at 37°C when 0.8 mM IPTG was added and the cultures were further grown at 20°C for 4 h. Cells were pelleted and stored at -80°C until required.

Protein Purification

For purification of both *B. thuringiensis* LexA and GIL01 gp7 used in structural experiments, cell pellets with pET28a constructs were resuspended in Buffer A (20 mM Hepes, pH 8.0, 300 mM NaCl, 10% glycerol) and lysed by processing twice with a homogenizer (15 kPa; Avestin). Cellular debris was pelleted by centrifugation at 125,000 x g for 1 hour. The resultant supernatant was loaded onto a 1/5 mL Ni^{2+}-saturated HisTrap HP Sepharose cartridge (GE Lifesciences), washed with 75 mM imidazole in Buffer A, and the protein was eluted with 300 mM imidazole in Buffer A. 1 U of thrombin was added per mg of protein to remove the N-terminal His-tag overnight at 4°C. Samples were purified further by size exclusion chromatography (SEC) with a Superdex 200 10/300 GL column (GE Lifesciences) equilibrated in Buffer A. Fractions containing pure LexA or gp7 were pooled and concentrated to 15 to 100 mg/mL. Protein was frozen rapidly in liquid nitrogen and stored at -80°C until required.

For use in SPR analysis, GIL01 gp7 and *B. thuringiensis* LexA were expressed from the pQE-30 vector as N-terminal hexahistidine fusions as described in (Fornelos et al., 2015). His_6-gp7Δ45-50, His_6-LexA homologs or the *B. thuringiensis* His_6-CTD of LexA were expressed from pET8c as fusions to a thrombin-cleavable N-terminal hexahistidine tag. Tags were not cleaved for experiments in SPR. Cell pellets with overexpressed recombinant proteins were resuspended in 50 mM NaH_2PO_4 and 0.3 M NaCl, pH 8.0 (Buffer B), supplemented with lysozyme (0.5 mg/mL), DNAse (10 μg/mL), RNAse (20 μg/mL), a cOmplete protease inhibitor cocktail tablet (Roche) and sonicated on ice three times for 30 s at a 40% amplitude (Vibracell, Sonics). Recombinant proteins were purified

from bacterial cytoplasm by HIS-Select HF nickel affinity gel (Merck), washed and eluted with Buffer B containing 20 mM or 300 mM imidazole, respectively. Proteins were stored at -80°C in Buffer C containing 20 mM NaH_2PO_4 (pH 7.4), 140 mM NaCl.

Surface Plasmon Resonance Assays

The SPR measurements were performed at the Infrastructural Centre for Analysis of Molecular Interactions at the Department of Biology, University of Ljubljana on a Biacore T100 (GE Healthcare) at 25°C. To prepare double-stranded DNA fragments encompassing single or tandem *lexA* operator site(s), complementary primers (Sigma-Aldrich) marked as in the Table S3 were mixed in a 1:1.5 (_u:_d) mole-to-mole ratio in Buffer C and annealed using a temperature gradient from 94°C to 4°C. Approximately, 50 resonance units (RU) of the annealed 22-42 bp DNA fragments carrying a 15-nucleotide overhang was hybridized to the complementary S1 primer that is immobilized on the flow cell 2 of the sensor chip SA. DNA fragments were injected at a flow rate of 2 μL/min. The interaction of repressor with the target DNA was assayed by injecting 10 nM His_6-LexA, either alone or in combination with 100 nM His_6-gp7 at 50 μL/min over the chip-immobilized DNA. Proteins were diluted in running buffer composed of 25 mM Tris-HCl (pH 7.4), 140 mM NaCl, 5 mM EDTA, 2 mM DTT, 0.1 mg/mL BSA, 0.005% surfactant P20. To assay the interaction of gp7 with LexA repressor homologs or with the *B. thuringiensis* LexA CTD, we immobilized the repressor variants on carboxymethyl dextran-coated gold surface of the series S sensor chip CM5 (GE Healthcare). The carboxymethyl groups were first activated with 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride and 0.1 M N-hydroxysuccinimide to which the His_6-LexA proteins were afterwards covalently coupled *via* the free amino groups. Any uncoupled protein was removed by a short pulse of 50 mM NaOH and any uncoupled active carboxymethyl groups on the chip were blocked with ethanolamine. His_6-gp7 protein was serially diluted in running buffer composed of 20 mM HEPES (pH 7.4), 140 mM NaCl, 5 mM EDTA, 0.1 mg/mL BSA, 0.005% P20. To assay the titration reproducibility, each His_6-gp7 concentration tested (ranging from 2.35 to 2400 nM) was injected three times for 360 s at a flow rate of 30 μL/min at 25°C. The dissociation phase was followed for 120 s. The experiments were performed in duplicate. Regeneration of the sensor surface was performed with 1.6 mM NaOH solution for 6 s at a flow rate of 30 μL/min. Sensorgrams were doubly referenced for the untreated surface flow-cell 1 response and the buffer response. The data were analysed with the Biacore T100 Evaluation Software and equilibrium dissociation constants (K_D) were determined by fitting the data to the steady state affinity model. The average K_Ds and standard deviations were determined from three titrations of each protein.

Phylogenetic Analysis

Sequences of *B. thuringiensis* (UniProt ID: A0A2B4EEI3), *S. aureus* Mu50 (UniProt ID: P65819), *E. coli* K12 (UniProt ID: P0A7C2), *C. difficile* 630 (UniProt ID: Q187P1) or *A. actinomycetemcomitans* serotype b strain (UniProt ID: A0A1C7BWL4) LexA protein sequences were aligned with MAFFT 7.215 (Katoh and Standley, 2013). The best relatedness model (LG) was estimated with ProtTest 3.2.1 3.4 (Darriba et al., 2011). The aligned amino acid LexA sequences were used to construct a phylogenetic tree with the PhyML software based on the maximum-likelihood principle (Guindon et al., 2010) with aLRT implementation for the calculation of branch supports as Chi2 based support. The ProtTest estimate of the alpha parameter (3.273) of the gamma distribution of substitution rate categories and the estimated proportion of invariable sites (0.112) were used in the analysis.

Crystallography and Determination of the gp7 Structure

GIL01 gp7 protein was crystallized at 20°C by sitting drop vapour diffusion using 0.2 μL protein solution (49 mg/mL purified protein in Buffer A) and 1 μL of mother liquor (0.2 M sodium iodide, 0.1 M Bis Tris propane, pH 8.5, 20% (w/v) PEG 3.35k). X-ray diffraction data was collected on Canadian Light Source beamline 08B1-1 using a single crystal flash-frozen in liquid nitrogen. For phasing, a single wave single-wavelength anomalous dispersion (SAD) experiment was carried out using the Iodine of the mother liquor at wavelength 1.771 Å. All datasets were processed with Autoprocess (Fodje et al., 2014) running XDS (Kabsch, 2010). For the SAD dataset, Phaser (McCoy et al., 2007) was used for phasing, model building was performed by AutoBuild (Terwilliger et al., 2007) and refined using Phenix (Adams et al., 2010) and Coot (Emsley and Cowtan, 2004). For the native dataset, the structure was solved by molecular replacement using Phaser (McCoy et al., 2007). The structure was refined using Phenix (Adams et al., 2010) and Coot (Emsley and Cowtan, 2004).

Circular Dichroism Spectroscopy

The circular dichroism (CD) spectrum of GIL01 gp7 was collected with a Jasco J-810 circular dichroism spectrometer (Jasco Incorporated). The spectrum of a 20 μM protein solution (20 mM Tris, pH 8.0, 300 mM NaF) was collected at 25°C with a 0.1 cm path length quartz cuvette. Three scans were collected and averaged over the range of 195 nm to 240 nm using a spectral bandwidth of 1 nm and a response time of 2 seconds. PDB2CD (Mavridis and Janes, 2017) was used to generate simulated CD spectra of gp7 with both unstructured and α-helical C-termini for comparison. Models for the simulated spectra were generated using Coot (Emsley and Cowtan, 2004) and ModLoop (Fiser and Sali, 2003). NRMSD values were calculated as described previously (Mao et al., 1982).

SEC-MALS Analysis of Quaternary Structure

Purified protein was applied to a Superose 6 10/300 GL column (GE Lifesciences), equilibrated in Buffer A, using an Agilent 1100 series HPLC (Agilent Technologies) equipped with an in-line coupled Dawn Heleos II 18-angle MALS light scattering detector and an Optilab T-rEX differential refractometer monitor (Wyatt Technology). Bovine serum albumin (Sigma-Aldrich) was used to normalize the light scattering detectors and data was analysed with the Astra 6 software package (Wyatt Technology).

bioSAXS Analysis of Quaternary Structure and bioSAXS Guided Complex Modelling

Biological small angle X-ray scattering (bioSAXS) data were collected with in-house X-rays (1.54 Å) and a Rigaku bioSAX-2000 (Rigaku Corporation). Scattering profiles of purified *B. thuringiensis* LexA were collected at 40, 20, 10, 5, 2.5 and 1.25 mg/mL and gp7-LexA complex at 12, 6, 3, 1.5, 0.75 and 0.375 mg/mL. 12 consecutive frames of 5 minutes in length were collected for each profile, corrected by subtracting the background scattering of SEC buffer, and normalized in concentration. Processing up to this point was carried out using SAXSLab (Rigaku Corporation). Further processing was performed with components of the ATSAS software package (Petoukhov et al., 2012). Scattering curves were merged to create final scattering curves for each of the samples. For the LexA sample, a homology model of *B. thuringiensis* LexA generated with I-TASSER (Yang et al., 2014) was used in conjunction with SASREF (Petoukhov and Svergun, 2005) to place the N- and C-terminal domains. ModLoop (Fiser and Sali, 2003) was used to model the linker region between the two domains. For the LexA-gp7 complex sample the I-TASSER model of LexA and the solved gp7 dimer structure were used in ClusPro (Kozakov et al., 2017) bioSAXS guided modelling to place the structured domains, while the linker region of LexA was once again modelled with ModLoop (Fiser and Sali, 2003). Theoretical scattering curves of the developed models were computed using CRYSOL (Svergun et al., 1995) and SREFLEX (Panjkovich and Svergun, 2016) and compared to the collected scattering curves. DAMMIN (Svergun, 1999) was used to generate *ab initio* bead models for both LexA and LexA-gp7 data. These bead models are represented by a surface generated at 25 Å by UCSF Chimera (Pettersen et al., 2004). Full SAXS sample details, data collection parameters, software, structure parameters, and modelling statistics can be found in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

SPR analysis of K_Ds was performed from three titrations of each protein and reported as an average and standard deviation. CD spectroscopy was performed in triplicate and averaged. NRMSD values between computed and collected CD spectra were calculated as described previously (Mao et al., 1982). SAXS data was collected in twelve consecutive frames per sample at six varying concentrations. High and low q data was merged in ATSAS, as outlined in Figure S2.

DATA AND SOFTWARE AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structure has been deposited with the Protein Data bank under accession number PDB: 6N7O. SAXS data has been deposited with the Small Angle Scattering Biological Data Bank under the accession numbers SASBDB: SASDEB8 and SASBDB: SASDEC8.