Nucleic Acids Research, 2015 **1**
doi: 10.1093/nar/gku1352

Escherichia coli antitoxin MazE as transcription factor: insights into MazE-DNA binding

Valentina Zorzini[1,2], Lieven Buts[1,2], Evelyne Schrank[3], Yann G.J. Sterckx[1,2], Michal Respondek[3], Hanna Engelberg-Kulka[4], Remy Loris[1,2], Klaus Zangger[3] and Nico A.J. van Nuland[1,2,*]

[1]Molecular Recognition Unit, Structural Biology Research Center, VIB, Pleinlaan 2, 1050 Brussels, Belgium, [2]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium, [3]Institute of Chemistry/Organic and Bioorganic Chemistry, University of Graz, Heinrichstrasse 28, A-8010 Graz, Austria and [4]Department of Microbiology and Molecular Genetics, Institute for Medical Research Israel-Canada (IMRIC), The Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel

Received July 15, 2014; Revised December 04, 2014; Accepted December 15, 2014

ABSTRACT

Toxin-antitoxin (TA) modules are pairs of genes essential for bacterial regulation upon environmental stresses. The *mazEF* module encodes the MazF toxin and its cognate MazE antitoxin. The highly dynamic MazE possesses an N-terminal DNA binding domain through which it can negatively regulate its own promoter. Despite being one of the first TA systems studied, transcriptional regulation of *Escherichia coli mazEF* remains poorly understood. This paper presents the solution structure of C-terminal truncated *E. coli* MazE and a MazE-DNA model with a DNA palindrome sequence ∼10 bp upstream of the *mazEF* promoter. The work has led to a transcription regulator-DNA model, which has remained elusive thus far in the *E. coli* toxin–antitoxin family. Multiple complementary techniques including NMR, SAXS and ITC show that the long intrinsically disordered C-termini in MazE, required for MazF neutralization, does not affect the interactions between the antitoxin and its operator. Rather, the MazE C-terminus plays an important role in the MazF binding, which was found to increase the MazE affinity for the palindromic single site operator.

INTRODUCTION

Toxin–antitoxin (TA) systems are ubiquitous on bacterial chromosomes and bacterial plasmids. Depending on the nature of the antitoxin and the mechanism by which it neutralizes the toxin, TA modules can be categorized into five distinct types (1–8). Type II TA systems, where both toxin and antitoxin are proteins, are the most common. Their expression is regulated at the level of transcription through the antitoxin, which acts as a repressor, the activity of which is modulated by the toxin (9–12). Multiple roles have been suggested for TA modules ranging from plasmid stabilization (13–16) to altruistic suicide (17,18). Recent reports indicate that TA modules are associated with generation of nondividing but viable persister cells (19–21). Modulation of the persister state requires entangled molecular mechanisms that link protein activity to transcription regulation via the intrinsically disordered nature of the antitoxin C-terminal domain (22).

The *mazEF* operon was the first TA system found on the *Escherichia coli* chromosome (17). It is related to the *kis/kid* module on plasmid R1 (1,23), and it is homologous to the *E. coli* chromosomal TA module *chpBIK* (24). The *mazEF* operons encode a long-lived ribonuclease MazF, which cleaves mRNAs at specific sites (25).

In addition, MazF targets the 16S rRNA within the *E. coli* 30S ribosomal subunit at the decoding center, thereby removing 43 nucleotides from the 3′-terminus. The resulting truncated ribosomes preferentially translate the subset of leaderless mRNAs (26). The activity of the MazF toxin is neutralized by the short-lived antitoxin MazE, which is degraded by the ClpPA serine protease (17).

MazE proteins consist of two domains. The N-terminal domain has a DNA binding function and adopts a swapped-hairpin β-strand motif, typically of the AbrB/MazE/MraZ superfamily (27,28). This fold is common among bacterial transcription regulators and is found e.g. in the transition state regulator Abh (29) and the transcription regulator SpoVT (30,31). X-Ray diffraction studies have shown that the C-terminal domain of MazE is intrinsically disordered and upon binding to MazF adopts a unique and mostly extended conformation (27,32). The MazE binding to MazF stabilizes and protects its own

*To whom correspondence should be addressed. Tel: +32 2 629 3553; Fax: +32 2 629 1963; Email: nvnuland@vub.ac.be

vulnerable C-terminus from specific protease's cleavage, by which the TA system would be induced, leading to antitoxin degradation and toxin activation.

The *E. coli mazEF* operon contains a 47-bp operator region that contains three binding sites for the MazE dimer (*'cab'*) (27,33,34). These sites are termed *'c'*, *'a'* and *'b'* when moving downstream toward the *mazE* start codon. Of these, only site *'a'* contains the perfect palindrome 5′-ATATAT-3′, a hallmark of TA operator sequences (35). The details of the interaction between *E. coli* MazE and its operator DNA are not known yet and two distinct models were proposed based on the crystal structures of a *Ec*MazEF complex (32) and a complex between *Ec*MazE and a dromedary heavy chain antibody fragment (27,34). DNA binding by MazE/MazF complexes is thought to be primarily due to the antitoxin, with the toxin serving to enhance binding affinity (33).

In this work we studied in detail the structural and (thermo-)dynamic features of *E. coli* MazE binding to the *'a'* site DNA. We show that the N-terminal domain of the *E. coli* MazE (*Ec*MazE$^{1–50}$) is solely involved in DNA binding, excluding any participation of the disordered C-terminus in DNA interaction. We confirm that the functional role of the intrinsically disordered region is purely related to toxin neutralization, which is essential for transcription regulation. Moreover, *Ec*MazF cooperatively increases the *Ec*MazE-DNA affinity on the single palindrome suggesting how the toxin/antitoxin molar ratio can control the self-regulation of the TA locus transcription.

MATERIALS AND METHODS

Expression and purification of full-length *Ec*MazE

The pQE30-*mazE* plasmid containing a 21 residue N-terminal tag, six of which are histidines, was transformed into *E. coli* MC4100ΔmazEF-lacI^q (relA$^+$) and cells were grown in 1 l of M9 minimal medium at 310 K supplemented with ^{15}N-labeled NH_4Cl and/or ^{13}C-labeled glucose, respectively, and 100 μg/ml ampicillin. At an OD_{600} of 0.8 protein expression was induced with 1 mM IPTG and after additional 5 h of incubation the cells were harvested by centrifugation (30 min/4.000 g/277 K) and resuspended in 50 mM K_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0. After sonication the suspension was again centrifuged (60 min/15.000 g/277 K) and the supernatant filtered through a 0.45 μm sterile filter prior to loading it onto a Ni-CAM column (Sigma Aldrich, St. Louis, MO, USA), pre-equilibrated in 50 mM K_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0. The protein was eluted from the column with 50 mM K_2HPO_4, 300 mM NaCl, 500 mM imidazole, pH 8.0 with a 10–500 mM imidazole gradient as a single peak at an approximate concentration of 100-mM imidazole. The fractions containing the protein were combined and dialyzed first against 2 l of distilled water followed by 1 l of 20 mM KH_2PO_4, 100 mM NaCl, pH 6.5 as nuclear magnetic resonance (NMR) buffer. In a final step the concentrated protein was heated up to 358–363 K for 2 min and slowly cooled down to room temperature. This heat treatment makes *Ec*MazE more stable for long-term storage at room temperature, since it denatures contaminating proteins, such as proteases, while *Ec*MazE can be refolded upon heat denaturation like other bacterial antitoxins (36,37).

Expression and purification of truncated *Ec*MazE$^{1–50}$

In parallel, the pQE30-*mazE*-truncated plasmid containing *Ec*MazE$^{1–50}$ with 18 residues N-terminus tag, six of which are histidines, was transformed into *E. coli* BL21 (DE3) competent cells. Cells were grown in ^{13}C, ^{15}N-enriched minimal medium (SPECTRA 9, purchased from Cambridge Isotope Laboratories). Expression of ^{13}C, ^{15}N-labeled *Ec*MazE$^{1–50}$ was induced with 1 mM IPTG at an OD_{600nm} of 0.6 and the culture was incubated overnight at 310 K, 120 rpm. The cells were harvested by centrifugation (30 min/4.000 g/277 K) and resuspended in lysis buffer, 20 mM Tris-HCl, 150 mM NaCl, 10 mM imidazole, pH 7.0, 0.1 mg/ml p-aminoethylbenzenesulfonyl fluoride (AEBSF) and 1 μg/ml leupeptin. After breaking the cells, passing them twice though a french press (1000–1200 bar, 12000 psi), the suspension was again centrifuged (20 min/15.000 g/277 K) and the supernatant filtered through a 0.45 μm sterile filter prior to loading it onto a 5 ml Ni-NTA resin (Qiagen) pre-equilibrated with 20 mM Tris-HCl, 150 mM NaCl, 10 mM imidazole, pH 7.0. The proteins were eluted using 20 mM Tris-HCl, 150 mM NaCl, 1 M imidazole, pH 7.0. The *Ec*MazE$^{1–50}$-containing peak started to elute at 330 mM of imidazole concentration. To obtain highly pure samples, *Ec*MazE$^{1–50}$ was consecutively loaded on a high-resolution Superdex 75PG 16/60 in 20 mM Tris-HCl, 150 mM NaCl, pH 7.0. The ^{13}C, ^{15}N-labeled *Ec*MazE$^{1–50}$ was further dialyzed against 50 mM Na phosphate pH 6.5, 50 mM NaCl, as suitable NMR buffer, adding a proteases inhibitor cocktail (10 mM ethylenediaminetetraacetic acid, 50 μg/ml AEBSF, 100 μg/ml leupeptin).

NMR spectroscopy on full-length *Ec*MazE

All full-length *Ec*MazE spectra for the assignment were acquired at 298 K on a Varian Unity INOVA 600-MHz NMR spectrometer. DNA binding experiments and ^{15}N relaxation data were obtained on a Bruker Avance III 700 MHz NMR spectrometer, equipped with a cryogenically cooled 5 mm TCI probe. For the NMR experiments *Ec*MazE was dissolved in 90% aqueous buffer (50 mM KPi pH 6.5, 50 mM NaCl) and 10% D_2O, except for the 3D HCCH-TOCSY (100% D_2O). Data were processed using NMRPipe (38) and analyzed in NMRView (39). ^{1}H, ^{15}N and ^{13}C resonances were assigned using 2D homonuclear and standard triple resonance experiment (40). {^{1}H}-^{15}N heteronuclear Nuclear Overhauser Enhancements (NOEs) of full-length *Ec*MazE were measured at 700 MHz and 298 K, and determined from the ratio of peak intensities (Ion/Ioff) with and without the saturation of the amide protons for 3 s. Average heteronuclear NOE values and their errors were obtained from a duplicate set of experiments.

NMR spectroscopy on truncated *Ec*MazE$^{1–50}$

^{13}C, ^{15}N-labeled-truncated *Ec*MazE$^{1–50}$ was prepared at 1.0 mM in 50 mM Na phosphate pH 6.5, 50 mM NaCl, 10% D_2O. All NMR spectra used for the assignment were recorded at 298 K using a Varian 600 MHz NMR Direct-Drive System. A 2D NOESY with a 100 ms mixing time was recorded on a Varian 800 MHz NMR Direct-Drive System,

equipped with a salt tolerant triple-resonance PFG-Z cold probe, on the same sample. All NMR data were processed using NMRPipe (38) and analyzed by CCPNMR (41).

Semi-automatic assignment of the protein backbone was performed using CCPNMR software (41). The ^{1}H,^{15}N frequencies of the ^{15}N-HSQC spectrum were used to correlate each peak with its $^{13}C\alpha$ and $^{13}C\beta$ and the ones of the preceding amino acid (by using HNCACB and CBCA(CO)NH spectra) and with the preceding ^{13}CO (via the HNCO spectrum). $^{1}H\alpha$ and $^{1}H\beta$ were assigned using the HBHA(CO)NH spectrum.

Assignments were extended to the side chain signals using correlations within the C(CO)NH and HCCH-TOCSY for the aliphatic side-chains. Aromatic ^{1}H and ^{13}C frequencies of the single Trp residue were assigned from the ^{13}C-HSQC and ^{13}C-NOESY-HSQC spectra. Side-chain $^{15}N^{1}H_2$ frequencies of glutamines and asparagines and $^{15}N\epsilon^{1}H\epsilon$ of arginines were assigned from HNCACB, CBCA(CO)NH and 3D ^{15}N-NOESY-HSQC spectra. All ^{1}H, ^{13}C and ^{15}N resonances were verified from 3D ^{15}N- and ^{13}C-NOESY-HSQC spectra (with 100 ms mixing times).

$\{^{1}H\}$-^{15}N heteronuclear NOEs of *Ec*MazE$^{1–50}$ were measured at 600 MHz and 298 K, and determined from the ratio of peak intensities (Ion/Ioff) with and without the saturation of the amide protons for 3 s. Average heteronuclear NOE values and their errors were obtained from a duplicate set of experiments.

*Ec*MazE$^{1–50}$ NMR structure calculations

Truncated *Ec*MazE$^{1–50}$ NMR solution structure calculations were performed using CYANA version 2.1 (42,43). Sixty-one inter-monomeric NOEs were identified based on the *Ec*MazE X-ray structure (PDB entry 1MVF). These manually assigned NOEs were used together with non-assigned NOEs and dihedral restraints from Talos+ (44) as input for CYANA (42,43)). Non-assigned NOEs were assigned using the automated NOE assignment procedure of CYANA. A standard protocol was used with seven cycles of combined automated NOE assignment and structure calculation of 100 conformers in each cycle. From the three NOESY data sets, 946 NOEs were unambiguously assigned, including 166 inter-monomeric NOEs (Table 1). These unambiguously assigned restraints were used for a final structure refinement in explicit solvent using the RECOORD protocol (45), which runs under CNS (46). The twenty lowest-energy structures were used for final analysis.

Isothermal titration calorimetry to study MazE-DNA binding

Isothermal Titration Calorimetry (ITC) experiments were performed on a MicroCal iTC200 system (GE Healthcare). Investigation of the *Ec*MazE binding to its own palindrome promoter sequence was carried out using the *'a'* site proposed previously by Marianovky *et al.* (33); (forward: 5′-TTGATATATACTGT-3′; reverse: 3′-ACAGTATATATCAA-5′). Besides the DNA '*a*' site as main target in this work, we performed ITC experiments on other biologically relevant sites of the *mazEF* operon, the full three sites *'cab'* (forward 5′-CTCGTATCTACAATGTAGATTGATATATACTGTA TCTACATATGATAGCGT-3′), and the two other single sites *'c'* (forward 5′-GTATCTACAATGTAGATTG-3′) and *'b'* (forward 5′-ATATACTGTATCTACATAT-3′), all purchased from Sigma Aldrich. A control experiment was done with DNA fragment *'X'* (forward: 5′-GATTTTTGATTTT-3′; reverse: 3′-AAAATCAAAAAC-5′), purchased from VBC Biotech (Vienna, Austria), and treated as all the other samples. The double-stranded DNA fragment solutions were generated by dissolving equimolar amounts of single-strands oligonucleotide in water, heated up the solution to 368 K at 275 K/min and then slowly cooled down to 298 K to allow annealing. To exactly match buffer composition, the double strand DNA fragments and both *Ec*MazEs, full-length and truncated, were dialyzed overnight against 2 l of 50 mM phosphate buffer at pH 6.5 and 50 mM NaCl. Prior to titration, the samples were filtrated with 0.22 μm filters and degassed for 10 min at temperature corresponding to the titration temperature. A 14 μM solution of DNA fragments was titrated with full-length *Ec*MazE and *Ec*MazE$^{1–50}$ (both solutions at 280 μM). The *Ec*MazEs-DNA *'a'* titrations were measured at three different temperatures: 292, 298, 305 K, while the others only at 305 K. Additionally, heats of dilution, determined by titrating the proteins into solution buffer, were subtracted from the raw titration data before analysis. Data analysis was performed with MicroCal Origin software accompanying the ITC instrument. The binding affinity (K_D) and change in enthalpy associated with the binding event (ΔH) were calculated after fitting each data set by least-squares procedures assuming an n identical and independent site-binding model. The change in heat capacity of binding (ΔC_p) was determined from the slope of the linear dependence of ΔH with the temperature.

Electrophoretic mobility shift assay

Binding of *Ec*MazE to the DNA '*a*' was followed by mobility shift electrophoresis (electrophoretic mobility shift assay, EMSA). Prior to hybridization, the DNA fragments were 5′-end labeled with [γ-32P]-ATP by T4 polynucleotide kinase (New England Biolabs). The double strand DNA *'a'* fragment was purchased from Sigma Aldrich. Labeled probes were incubated with purified proteins (MazE variable concentrations from 7.5–100 μM for following MazE to DNA binding; MazE fixed concentration of 1 μM for following the effect of MazF on the MazE-to-DNA binding using variable concentration of MazF from 0.5–10 μM in 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 50 μg/ml bovine serum albumin (BSA). Reactions were incubated for 20 min at 310 K. DNA bound complexes were separated by native polyacrylamide gel electrophoresis in 6% acrylamide gels with 0.5 X TBE for 3 h at 8 V cm^{-1}. The separation was followed by phosphorimaging.

NMR chemical shift mapping of MazE-DNA binding

Investigation of the *Ec*MazE binding to its own palindrome promoter sequence was carried out by chemical shift mapping using the *'a'* site. A 1.0 mM DNA stock solution was prepared by dissolving the two single-strands oligonucleotide in 50 mM KH_2PO_4 pH 6.5, 50 mM NaCl buffer

Table 1. Structural statistics over the 20 lowest-energy water-refined NMR structures of *Ec*MazE$^{1–50}$

	*Ec*MazE$^{1–50}$
Distance restraints total[a]	1892 (946 A, 946 B)
Short range ($i - j = 0$)	492 (246 A, 246 B)
Medium range ($1 \leq \|i - j\| \leq 4$)	920 (460 A, 460 B)
Long range ($\|i - j\| \geq 5$)	480 (240 A, 240 B)
Inter-monomer	332 (166 A->B, 166 B->A)
Dihedral restraints	166 (83 A, 83 B)
Phi angles	80 (40 A, 40 B)
Psi angles	86 (43 A, 43 B)
CNS energies (kcal/mol)	
E_{total}	-4967.6 ± 103.2
E_{vdw}	-526.2 ± 30.9
E_{elec}	-5554.1 ± 127.8
Restraint statistics	
NOE violations >0.5 Å	0
Dihedral violations >5°	0
RMSD from average[b] (Å)	Residues 3–47
Backbone N, CA, C′	0.50 ± 0.16
Heavy atoms	0.85 ± 0.15
Ramachandran plot	Residues 2–49
Most favored regions (%)	84.1
Additional allowed regions (%)	15.2
Generously allowed regions (%)	0.7
Disallowed regions (%)	0.0

Values are reported for the *Ec*MazE$^{1–50}$ homodimer consisting of monomers with chain ID: A and B.
[a]Statistics for residues from −6 to 50. Flexible N-terminal His-tag and the C-terminal residue were omitted from the RMSD analysis and Ramachandran statistics obtained from PROCHECK analysis.
[b]Values with their corresponding standard deviations are reported for the *Ec*MazE$^{1–50}$ homodimer.

and annealing over night. For the chemical shift mapping a full-length *Ec*MazE reference ^{1}H-^{15}N-HSQC NMR spectrum at a protein concentration of 0.3 mM was recorded prior to a six-step titration series with the corresponding DNA sequence at a concentration range from 0 to 0.3 mM. The final concentration of the protein in the last point of the titration was 0.21 mM. A control experiment was done using DNA fragment *'X'*. The NMR titration experiments were performed under the same conditions and concentrations used for the DNA *'a'* chemical shift mapping.

In order to investigate the DNA binding to truncated *Ec*MazE$^{1–50}$, a slightly different palindrome sequence was selected, differing from the *'a'* fragment used for the full-length *Ec*MazE by only 3 bp at the 5′ and 3′ extremities (forward: 5′-CGTGATATATACTGC-3′; reverse: 3′-GCAGTATATATCACG-5′, purchased from Sigma Aldrich). A 1.4 mM DNA stock solution was prepared by dissolving equimolar amounts of single-strands oligonucleotide in water, heated up the solution to 368 K at 275 K/min and then slowly cooled down to 298 K to allow annealing. Prior to titration, the double strand DNA *'a'* fragment was dialyzed against the same NMR buffer, 50 mM Na phosphate pH 6.5, 50 mM NaCl. For the chemical shift mapping a *Ec*MazE$^{1–50}$ reference ^{1}H-^{15}N-HSQC NMR spectrum at a protein concentration of 0.4 mM was recorded prior to the titration. A titration series was done in six steps using the DNA stock solution leading to a concentration range between 0 and 0.4 mM DNA. A ^{1}H-^{15}N-HSQC NMR spectrum was recorded at each step in order to follow the chemical shift perturbations upon DNA binding. The final concentration of the protein in the last point of the titration was 0.33 mM.

The magnitude of the chemical shift perturbation (Δδ) was calculated by $\Delta\delta = [(\Delta\delta_H)^2+(\Delta\delta_N/6.51)^2]^{1/2}$ where Δδ is the difference between the bound and free form combined chemical shifts.

Structure calculations of the *Ec*MazE:DNA complex

The structural model of the complex between *Ec*MazE$^{1–50}$ and DNA was obtained using the HADDOCK software (47). Ambiguous Interaction Restraints (AIRs) for *Ec*MazE$^{1–50}$ were obtained from the chemical shift perturbation data. All the atoms showing a higher difference than the corresponding mean were investigated in terms of solvent accessibility with NACCESS software (48) and location in the *Ec*MazE$^{1–50}$ structure. The active residues used in the docking were 7–12, 16, 18–20 in both *Ec*MazE$^{1–50}$ monomers. Histidine 3 was included in the docking as passive residue. Active and passive nucleotides in both DNA strands were defined as nucleotides 5–11, and 3–4 and 12–13, respectively. Overall, a total of 34 AIRs were defined between *Ec*MazE$^{1–50}$ and DNA with upper distances fixed at 2.0 Å.

Docking was started from the whole ensemble of 20 lowest-energy *Ec*MazE$^{1–50}$ free structures. For the DNA *'a'* fragment we used the X-ray structure of the VapBC$_2$–DNA complex (PDB entry 3ZVK) as model (49). We mutated the VapBC$_2$ DNA using UCSF Chimera (50) to obtain the structure of our 15 bp DNA *'a'* fragment. During docking the DNA was kept rigid, while the protein was kept semi-flexible. The final step of the structure refinement was done in explicit water. The seven structures with the lowest interaction energies and lowest AIR violations were selected for further analysis.

Small angle X-ray scattering

Complexes of full-length *Ec*MazE DNA and *Ec*MazE$^{1–50}$-DNA from the same pQE30-*mazE* construct (18 N-terminal His-tag) were analyzed by SAXS; data were collected at Swing, Soleil Synchrotron (Paris, France) at 7–8 mg/ml concentration. The complexes were beforehand dialyzed against 20 mM Tris-HCl pH 7.0, 150 mM NaCl. All samples were subsequently centrifuged at 10 000 rpm for 2 min at 277 K and loaded on Shodex packed HPLC column coupled to the beam capillary. For each data set, 250 frames were collected, averaged and background-subtracted. The initial data process was realized using the program PRIMUS (51,52) for scaling and merging. A Guinier analysis was performed at very low scattering angle and used to estimate the radius of gyration (R_g) of the particle. The indirect Fourier transform package GNOM (51) was used to compute the distance distribution $p(r)$ function from the scattering curve and calculate the maximum dimension of the particle (D_{max}). To accurately determine the molecular mass, assess model-data agreement and verify that we did not over-fitted our data, metrics like Q_R, χ^2 free and R_{sas} were calculated based on Rambo and Tainer (53). To define the minimal set of MazE$^{1–50}$-DNA NMR structures that can explain the SAXS data, the minimal ensemble algorithm (Minimal Ensemble Search, MES) was used (54). This algorithm searches for the minimal ensemble set of conformations from the pool of all given conformations, systematically evaluating combinations of five models or less. The full-length *Ec*MazE in complex with DNA was built using our best NMR *Ec*MazE$^{1–50}$ structure in complex with DNA and modeled the missing C-terminal disordered tails using MODELLER (AllosMod FOXS) (55,56). A minimal ensemble was defined for this protein complex as well. The comparison between the theoretical scattering curves of both protein–DNA complexes with the experimental data, which was expressed in χ^2 goodness of fit, was done using FoXS (54).

Paramagnetic relaxation enhancement

In order to validate the structural model of the *Ec*MazE$^{1–50}$–DNA complex we monitored intensity changes in ^{1}H-^{15}N-HSQC spectra of full-length *Ec*MazE upon the binding of double stranded, paramagnetically tagged DNA corresponding to the *'a'* region in the *mazEF* operon. A paramagnetic iodoacetamido-proxyl tag was attached to DNA oligos containing a PTO modification on either the 5′- or 3′-end. 0.5 mM double-stranded DNA stock solutions were prepared by dissolving the modified and unmodified oligonucleotide strands, respectively, in 200 mM trishydroxymethyl-aminomethane, pH 8.0 buffer and left the DNA annealed over-night. Subsequently, the spin-label 3-(2-iodoacetamido)-proxyl (Sigma-Aldrich, St. Louis, MO, USA) was added to a final concentration of 20 mM to the modified dsDNA and the solution was stirred for 48 h at room temperature in the dark. Subsequently, the free spin-label was extracted from the solution by phase separation after addition of 500 μl CH_2Cl_2. This solvent extraction step was repeated twice prior to dialyzing the sample twice against 1 l of distilled water for a complete removal of the organic solvent. Finally, the dialyzed sample was lyophilized. For each titration, with spin-labeled DNA, a *Ec*MazE reference ^{1}H-^{15}N-HSQC spectrum at 0.2 mM final protein concentration was recorded prior to dissolving the lyophilized DNA in the 300 μl ^{15}N-labeled *Ec*MazE sample. After recording the *Ec*MazE spectrum with paramagnetically labeled DNA and adding 10 mM sodium dithionite to reduce the proxyl group to its diamagnetic form, another ^{1}H-^{15}N-HSQC was acquired. Comparison of peak intensities in the diamagnetic and paramagnetic form yields signal reductions due to paramagnetic relaxation enhancements (PREs). High mobility of the tag due to its length and position at the flexible DNA ends prevented a quantitative analysis of the PRE data.

RESULTS

NMR solution structure and dynamics of *Ec*MazE

In order to reveal insights on the *Ec*MazE–DNA interaction, we firstly conducted a structural and dynamic comparison between the 'wild-type' full-length *Ec*MazE antitoxin and its truncated version *Ec*MazE$^{1–50}$. The *Ec*MazE C-terminal intrinsically disordered tail was strategically truncated in order to distinctly characterize the N-terminal DNA binding domain and its direct interaction with the TA promoter fragment. In contrast to the strong overlap in the crowded full-length *Ec*MazE HSQC spectrum (Figure 1), *Ec*MazE$^{1–50}$ shows a nicely dispersed HSQC, which guaranteed a straightforward peak assignment, and consequently an accurate structural determination. The ^{1}H-^{15}N peaks for 10G and 11N (including those of the side-chain NH_2) are not detectable in the full-length *Ec*MazE HSQC, while they show up as weaker signals in the truncated *Ec*MazE$^{1–50}$ HSQC. Interestingly, overlay of the two monomers in the X-ray structure (PDB entry 1MVF) shows different conformation of the 10G-11N loop, indicating conformational exchange explaining the weakening/disappearing of the NMR signals. Moreover, in one of the monomers, there is no density for the N11 side chain in the X-ray structure. Difference in chemical shifts between *Ec*MazE$^{1–50}$ and full-length *Ec*MazE is mainly evident for residue E50, which corresponds to the C-terminus in *Ec*MazE$^{1–50}$ (see Supplementary Figure S1).

The ^{1}H, ^{15}N and ^{13}C assigned resonances of *Ec*MazE$^{1–50}$ and of full-length *Ec*MazE have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession number 25086 and 25093, respectively.

The NMR structure of the truncated version *Ec*MazE$^{1–50}$ was obtained from the combined use of distance and dihedral restraints. Figure 1C shows the ensemble of the 20 lowest energy conformations. NMR structural statistics are summarized in Table 1. The structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2MRN.

The *Ec*MazE$^{1–50}$ structure possesses a typical swapped hairpin β-strand motif consisting of two N-terminal β-strands, followed by an α-helix and two C-terminal β-strands. The N-terminal and C-terminal strands form two 4-stranded β-sheets in the homodimer. Our *Ec*MazE$^{1–50}$ solution structure resembles closely the crystal structure of full-length MazE in complex with a nanobody (27), which

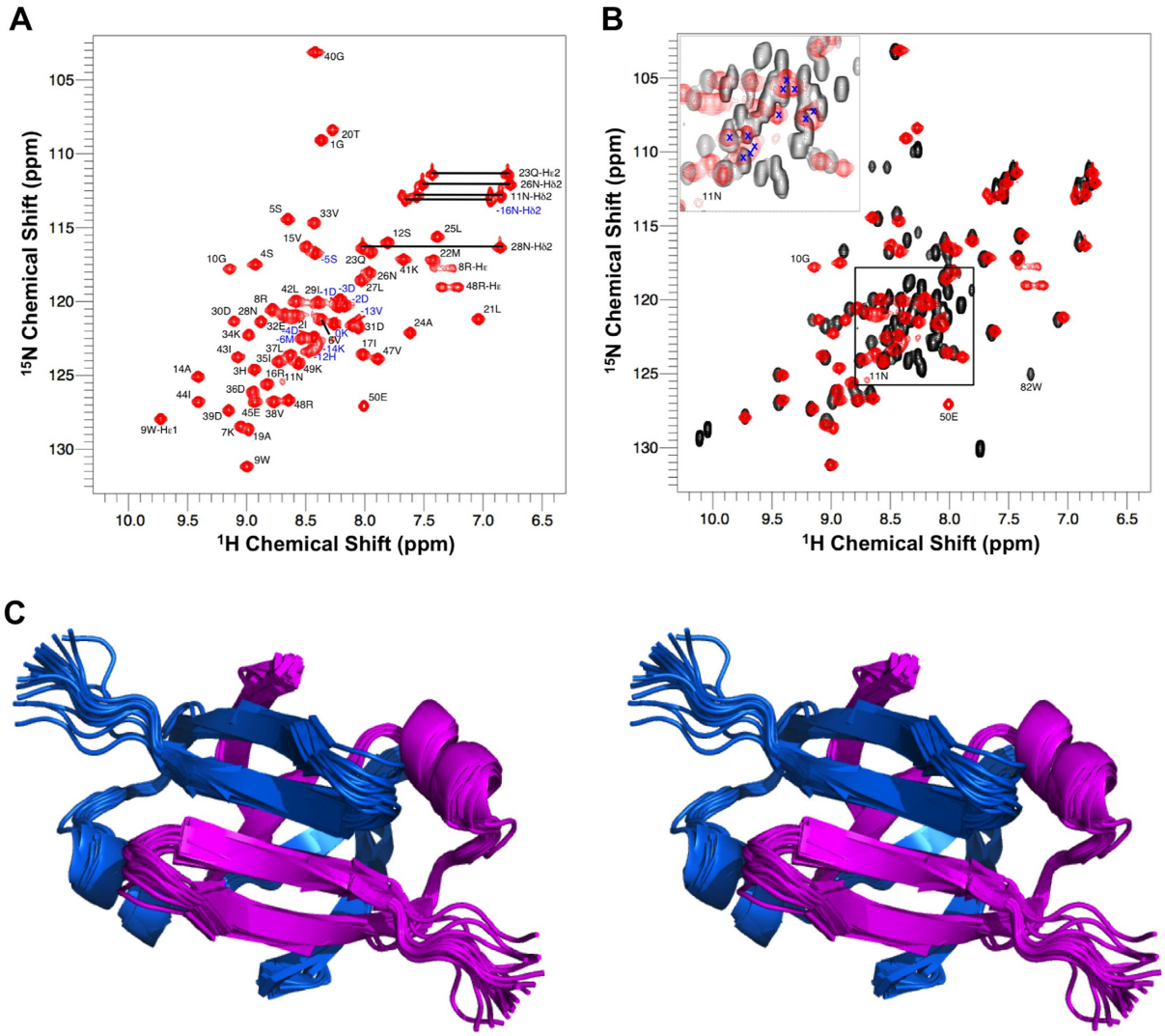

Figure 1. NMR characterization of *Ec*MazE$^{1–50}$ versus full-length *Ec*MazE. (**A**) Assigned 600 MHz ^{1}H-^{15}N-HSQC spectrum of ^{13}C-^{15}N-labeled *Ec*MazE$^{1–50}$ at 298K. *Ec*MazE$^{1–50}$ peaks of backbone ^{1}H, ^{15}N pairs are numbered with their corresponding position in the amino-acid sequence. Peaks labeled in blue belong to the *Ec*MazE$^{1–50}$ His-tag. (**B**) In red the 600 MHz ^{1}H-^{15}N-HSQC spectrum of *Ec*MazE$^{1–50}$ and in black the 600 MHz ^{1}H-^{15}N-HSQC spectrum of full-length *Ec*MazE. The inset contains peaks of the N-terminal His-tag and all peaks of the C-terminus of full-length *Ec*MazE (aa 50–81), except C-terminal Trp82. The same region contains the peaks of the *Ec*MazE$^{1–50}$ His-tag, indicated with a blue X. Peaks belonging to G10 and N11 visible in the *Ec*MazE$^{1–50}$ HSQC and not detectable in the full-length *Ec*MazE HSQC are indicated. (**C**) Stereo cartoon representation of the 20 lowest energy *Ec*MazE$^{1–50}$ NMR structures; the two monomers are colored in sky-blue and magenta. The highly flexible N-terminal (His-tag) was removed from the NMR ensemble for clarity. Figure created in PyMol.

shows electron density for residues 4 to 47 only, thus missing the disordered C-terminal tail. The backbone rmsd for residues 4–47 of both monomers in the dimer is 0.864 Å between the closest-to-average NMR structure and the X-ray structure, and is 1.07 ± 0.21 Å between all NMR structures and the X-ray structure, using Profit (http://www.bioinf.org.uk/profit/).

Information about the dynamical behavior of both full-length and truncated *Ec*MazEs in their free form was obtained by measuring {^{1}H}-^{15}N steady state NOEs (Supplementary Figure S2). Small and negative {^{1}H}-^{15}N NOEs are indicative of higher flexibility and they are observed mainly for the N-terminal residues (His-tag) and in the disordered C-terminal domain, in agreement with the lack of density in the crystal for the C-terminal tails of full-length *Ec*MazE. Moreover, lack of chemical shift dispersion and high intensity peaks of the C-terminal domain in full-length *Ec*MazE (Figure 1B, inset) confirms the high flexibility of this region.

Additionally, the predicted secondary structure elements from ^{1}Ha, ^{13}Ca, ^{13}C' and ^{13}Cb chemical shifts, presented as Chemical Shift Index (CSI) patterns in Supplementary Figure S3, are the same for the N-terminal domain of *Ec*MazE in truncated and full-length *Ec*MazE, and correspond well to the secondary structure elements present in the solution structure. The CSI patterns for the C-terminal domain in the full-length *Ec*MazE as well as the low-dispersed and high intensity NMR signals and the small and/or negative {^{1}H}-^{15}N NOEs indicate random structure for this region

and point out that the disordered C-terminal tail does not affect the N-terminal domain structure and dynamics.

Isothermal titration calorimetry on *Ec*MazE binding to its operator DNA fragments

DNA binding is supposed to be limited to the N-terminal domain of *Ec*MazE and specific for the palindrome DNA sequence *'a'*. We performed ITC experiments using both full-length and truncated *Ec*MazE and titrate them with the selected DNA *'a'* fragment (Figure 2B). These ITC experiments show an equal behavior of these two proteins upon binding the oligonucleotide. The binding of *Ec*MazE$^{1–50}$ to the other biological relevant operator fragments *'b'*, *'c'* and *'cab'* was also measured (Figure 2A and C). Supplementary Figure S4 shows all the ITC data for the systems studied. Table 2 reports the thermodynamic parameters for every ITC measured. The dissociation constants (K_D) vary from ∼0.5–5 μM for the different oligonucleotide fragments tested. The enthalpy of DNA *'a'* binding is rather constant over a temperature range between 292 and 305 K, which leads to a ΔCp close to zero for both *Ec*MazE and *Ec*MazE$^{1–50}$ (Supplementary Figure S4C and Supplementary Table S1). This indicates very little if any structuring of *Ec*MazE upon DNA binding, neither of its DNA binding domain nor of its disordered tails and agrees with the disordered tails not being involved in the process. This is likely happening as well for all the other fragments, since the thermodynamic values are closely similar. ITC experiments using random DNA segment *'X'* under the same conditions (Supplementary Figure S4G) show very low affinity binding to *Ec*MazE, indicating that the antitoxin binding is specific to its own operator.

MazE shows higher affinity for the single palindrome operator when titrated with MazF

To elucidate the role of the C-terminus extended domain of antitoxin *Ec*MazE, we performed EMSAs using full-length *Ec*MazE and DNA *'a'* and consecutively titrating the *Ec*MazE–DNA *'a'* complex with increasing amounts of toxin *Ec*MazF. First, a gel shift analysis using fixed concentrations of DNA *'a'* and variable amounts of *Ec*MazE was carried out to probe the antitoxin binding to the palindrome single site. A clear shift corresponding to complex formation is observed, accompanied with diminishing amounts of free DNA (Figure 3A), showing that *Ec*MazE binds specifically to the operator fragment *'a'*.

However, addition of a variable amount of *Ec*MazF to a lower concentration of *Ec*MazE (1 μM), which is not sufficient to cause a shift of the DNA band by itself, results in an increase in affinity for *Ec*MazE to DNA (Figure 3B). At very high *Ec*MazF:*Ec*MazE ratios, this effect is abolished and coincides with a reduced shift of the band corresponding to the complex. Thus, *Ec*MazF enhances the binding of *Ec*MazE to their DNA operator fragment *'a'* though its interaction with the C-terminal region of *Ec*MazE.

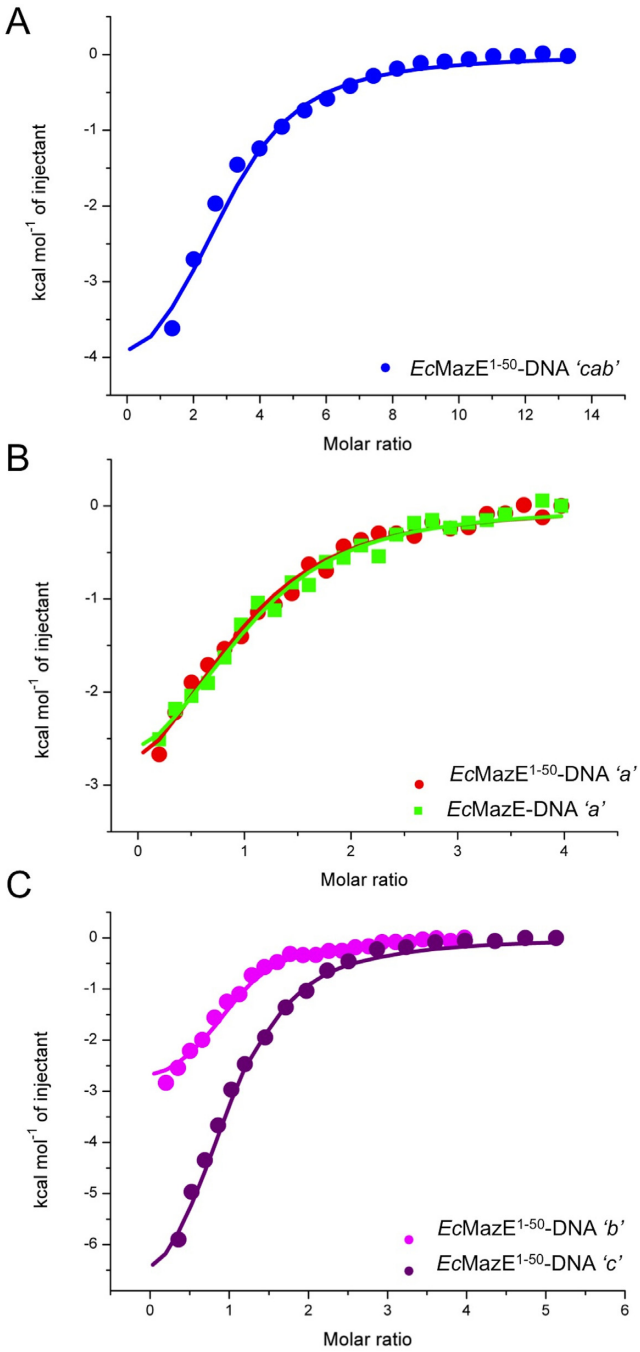

Figure 2. Isothermal titration calorimetry on *Ec*MazE-DNA binding. (**A**) ITC titration curve for *Ec*MazE$^{1–50}$ binding to DNA *'cab'* at 305 K. (**B**) ITC titration curves for *Ec*MazE$^{1–50}$ and full-length *Ec*MazE binding to DNA *'a'* at 305 K in red (circles) and in green (squares), respectively. (**C**) ITC titration curve for *Ec*MazE$^{1–50}$-DNA *'c'* in purple and *Ec*MazE$^{1–50}$-DNA *'b'* in magenta measured at 305 K. The solid lines in panels (A), (B) and (C) correspond to the best fit using a *n* equal to 1 binding site model. The thermodynamic parameters for the *Ec*MazE-DNA binding are reported in Table 2.

Structural model of the *Ec*MazE–DNA complex from NMR and SAXS

Since no *Ec*MazE-DNA structure is available, we aimed at obtaining an *Ec*MazE-DNA structure using a combination

Table 2. Thermodynamic parameters for *Ec*MazE-DNA binding from ITC

	*Ec*MazE$^{1–50}$-DNA *'cab'*	*Ec*MazE-DNA *'a'*	*Ec*MazE$^{1–50}$-DNA *'a'*	*Ec*MazE$^{1–50}$-DNA *'b'*	*Ec*MazE$^{1–50}$-DNA *'c'*
Thermodynamic parameters					
K_D (μM)	0.6 ± 0.5	5.1 ± 0.2	5.6 ± 0.3	2.1 ± 0.6	0.6 ± 0.2
ΔH (kcal/mol)	−5.1 ± 0.3	−3.5 ± 0.3	−3.8 ± 0.4	−3.1 ± 0.1	−9.1 ± 0.5
TΔS (kcal/mol)	3.4 ± 0.2	3.9 ± 0.2	3.5 ± 0.1	4.3 ± 0.3	1.0 ± 0.6
n	3	1	1	1	1

The error indicated corresponds to standard deviation. Number of binding site, n, was fixed at the value indicated.

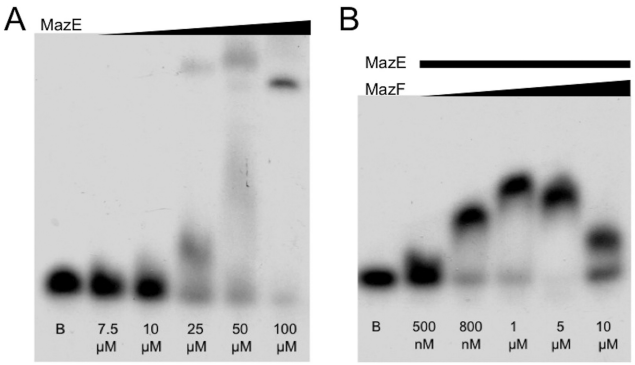

Figure 3. *Ec*MazEF toxin–antitoxin system shows higher affinity binding than the antitoxin alone for its single site operator fragment. (**A**) Concentration-dependent binding of full-length *Ec*MazE to the single site *'a'* operator fragment. As the *Ec*MazE concentration increases the DNA shifts from the free state to a *Ec*MazE–DNA complex. Blank sample (DNA without protein) is indicated by a B. (**B**) Binding of *Ec*MazE to the single site *'a'* operator fragment enhanced by MazF. All samples include equal concentration of antitoxin *Ec*MazE (1 μM), which is not sufficient to cause a shift of the DNA band. As the *Ec*MazF concentration increases and approaches the 1:1 ratio with the antitoxin, however, a clear mobility shift is observed. At higher ratios, the shift disappears again. B: blank sample (DNA without protein).

of techniques. For this, we choose a 15 bp operator fragment containing the ATATAT palindrome sequence labeled as *'a'* site (33). Using double labeled full-length *Ec*MazE and truncated *Ec*MazE$^{1–50}$, we carried out NMR chemical shift mapping experiments, titrating both protein solutions with the *'a'* site DNA (Figure 4). The NMR titration led to close-to-maximum chemical shift changes already at stoichiometric ratios. Substantial chemical shift perturbations and/or signal disappearance in the *Ec*MazE$^{1–50}$ titration are detected for residues K7, R8, W9, G10, N11, S12, A14, V15, R16, I17, A19, T20, Q23, L27, N28 and I29 (Figure 5A). These chemical shift perturbations induced by the DNA *'a'* binding on *Ec*MazE$^{1–50}$ were mapped on the *Ec*MazE structure, as shown in Figure 5C and D.

We performed the NMR titrations also on full-length *Ec*MazE. The chemical shift perturbations upon DNA binding are very similar between the *Ec*MazE$^{1–50}$ and full-length *Ec*MazE (Figure 5A and B, respectively), underlining that both proteins bind specifically to the oligonucleotide fragment used and that the C-terminal region in *Ec*MazE is not involved in direct binding to DNA. Additionally, an NMR titration with the random DNA sequence *'X'* was performed as negative control (Supplementary Figure S5). The DNA-binding-induced *Ec*MazE shifts in this NMR titration experiment are much smaller and for some residues different than the ones upon binding DNA *'a'*, indicating that *Ec*MazE indeed binds DNA *'a'* specifically with much higher affinity.

The ^{1}H and ^{15}N assigned resonances of *Ec*MazE$^{1–50}$ and of full-length *Ec*MazE in complex with DNA *'a'* have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession numbers 25092 and 25094, respectively.

Based on the chemical shift perturbations shown in Figure 5, we used this information to drive the docking of the DNA *'a'* fragment onto the *Ec*MazE$^{1–50}$. The chemical shift perturbations are in agreement with the X-ray structure of the *Rickettsia felis* VapBC$_2$–DNA complex (PDB entry 3ZVK), which adopts the same fold as the *Ec*MazE N-terminal domain (Supplementary Figure S6). Besides the similar fold of the homologous proteins, the central region of the DNA involved in binding VapB$_2$ (ATATATACT) is identical to that in the DNA *'a'* fragment we used, which we have demonstrated to bind specifically to the antitoxin *Ec*MazE.

Figure 6 shows the structural models of the *Ec*MazE$^{1–50}$–DNA complex resulting from the structure calculation procedure using HADDOCK. The structural statistics are summarized in Table 3. The structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2MRU.

A closer view into this structural model reveals that the homodimer *Ec*MazE$^{1–50}$ binds into the major groove of double-stranded DNA *'a'*, involving key residues W9, N11, R16 for the main interactions with the oligonucleotide. The complex shows a large concave surface for protein interaction in the center of the oligonucleotide fragment, resulting from widening of the major groove. In this model, the R16 side-chain and the A19 backbone amide make specific hydrogen bonds with the nucleotide bases at positions 6 and 4 or 5, respectively. The formation of an H-bond of the A19 amide proton is in agreement with the downfield shift observed for this proton upon DNA binding (Figure 4B). Additional electrostatic interactions between the positive-charged residues K7 and R8 with the DNA backbone stabilize the *Ec*MazE$^{1–50}$–DNA complex. Residues W9, N11 and R16 are three key residues in *Ec*MazE$^{1–50}$–DNA complex which correspond to N9, Q11 and R16 in VapB$_2$ for DNA binding (49). The side-chains of these homologous residues show striking similar structural conformations (Supplementary Figure S6B), indicating the common key-role for these three residues in DNA recognition.

Table 3. Structural statistics over the seven Haddock structures of *Ec*MazE$^{1–50}$-DNA *'a'*

Ambiguous interaction restraints		34	
CNS interaction energies (kcal/mol)			
E_{total}		− 208.1 ± 25.8	
E_{vdw}		− 70.8 ± 8.6	
E_{elec}		− 501.9 ± 17.5	
Restraint statistics			
AIR violations >0.3 Å		1.57 ± 0.79	
Buried surface area[a] (Å^2)	*Ec*MazE$^{1–50}$		DNA
	953.2 ± 44.8		898.5 ± 48.6
RMSD from average[b] (Å)		Residues 2–49	
Backbone N, CA, C′		0.52 ± 0.12	
Heavy atoms		0.87 ± 0.10	
Ramachandran plot[b]		Residues 2–49	
Most favored regions (%)		79.6 ± 3.9	
Additional allowed regions (%)		19.3 ± 4.1	
Generously allowed regions (%)		1.0 ± 1.3	
Disallowed regions (%)		0.0	

[a]BSAs with their corresponding standard deviations calculated between the *Ec*MazE$^{1–50}$ dimer and the double-stranded DNA using PDBePISA (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver).
[b]Values with their corresponding standard deviations are reported for the *Ec*MazE$^{1–50}$ homodimer. Ramachandran statistics obtained from PROCHECK analysis.

In order to compare the structural and dynamical characteristics between truncated *Ec*MazE$^{1–50}$-DNA with full-length *Ec*MazE–DNA complexes, we employed SAXS on both *Ec*MazE-DNA *'a'* complexes. SAXS is particularly suitable for studying less structured systems, and especially complementary with NMR (37,57,58). No other techniques so far have reported the behavior of *Ec*MazE with its long disordered tails in solution. Complementary to validating our NMR *Ec*MazE$^{1–50}$-DNA structure, we aimed to investigate the structural dynamics of the DNA-binding domain and the extended toxin-neutralizing domain in the full-length protein. The SAXS data collection, structural parameters and model statistics derived from the Guinier analysis of both full-length *Ec*MazE-DNA and truncated *Ec*MazE$^{1–50}$–DNA complexes are given in Table 4. The estimated molecular masses determined by Guinier $I(0)$ analysis, SAXSMoW (59) and Q_R (53) agree well with the one predicted from the corresponding sequences. The overall size of both systems was examined by monitoring the R_g and D_{max} values. As expected, a comparison of these parameters for *Ec*MazE$^{1–50}$-DNA and full-length *Ec*MazE-DNA, respectively, indicates that there is a substantial difference in terms of size between both scattering particles (R_g 25 Å versus 30 Å, and D_{max} 71 Å versus 91 Å; see also Figure 7C). Furthermore, a comparison between the normalized Kratky plots of both *Ec*MazE–DNA complexes (Figure 7D) reveals that the *Ec*MazE$^{1–50}$-DNA shows diminished internal flexibility compared to full-length *Ec*MazE-DNA. This is in perfect agreement with our NMR relaxation data (Supplementary Figure S2), and can be explained by the absence of the long disordered C-terminal tail in the truncated *Ec*MazE$^{1–50}$. Such plots show a maximum value of ∼1.3 at a $q{\cdot}Rg$ value of around 2.2 for the *Ec*MazE$^{1–50}$–DNA *'a'* complex and around 2.6 for the full-length *Ec*MazE–DNA *'a'* complex. None of the two normalized Kratky return to zero, indicating the presence of highly flexible regions in the scattering particle mainly due to the His-tag in both complexes and increased flexibility is even more present in the full-length *Ec*MazE–DNA *'a'* complex due to the extended disordered C-terminal tails, missing in the *Ec*MazE$^{1–50}$–DNA *'a'* complex. Because of the significant degree of flexibility present in both systems, it is unlike that a single conformer can account for the experimental SAXS data. We therefore determined the minimal ensemble of structures (MES) sufficient to describe the SAXS data and the peculiar dynamics of the complex. In the case of *Ec*MazE$^{1–50}$-DNA *'a'*, the MES turned out to be as little as two models ($\chi^2 = 0.64$), whereas a minimal ensemble of three structures is needed for full-length *Ec*MazE-DNA *'a'* ($\chi^2 = 0.96$) (Figure 7A and B). The major source of variability required for a good agreement with the SAXS data is likely attributed to the flexible C-terminus more than at the N-terminal His-tag.

To verify that our model-data were accurately determined, we calculated the χ^2 free and R_{sas} for both complexes (53). From the high quality of these values (see Table 4), we can conclude that our analysis was not overfitted and the models are in good agreement with the experimental SAXS data.

To confirm the correctness of our *Ec*MazE-DNA structural models, paramagnetic spin labels were introduced into the *'a'* oligomer. The PRE was measured by monitoring peak intensities in 2D ^{1}H-^{15}N-HSQC spectra upon the reduction of the paramagnetic iodo-acetamido-proxyl labels to its diamagnetic form. Paramagnetic probes like this nitroxide spin-label influence the relaxation behavior of nearby signals (distance <∼10 Å). Transverse relaxation (T_2) enhancement leads to broader signals, which in turn lowers their intensity. Due to the relatively short-range effect only NH signals close to the paramagnetic DNA tag are expected to show an effect. While most signals of *Ec*MazE are only slightly or not affected by the paramagnetic probe, the signals of H3, but especially of A19, T20 and M22, change their intensities significantly, in agreement

Table 4. SAXS data collection and scattering-derived parameters for *Ec*MazE–DNA complexes

	*Ec*MazE$^{1–50}$-'*a*'	*Ec*MazE-'*a*'
Data collection parameters		
Beam line	SWING	SWING
HPLC column	KW402.5–4F	KW402.5–4F
Wavelength (Å)	1.03	1.03
q range	0.012–0.299	0.012–0.331
Injected concentrations	90 µl at 7 mg/ml	90 µl at 8 mg/ml
Temperature (K)	283	283
Structural parameters		
$I(0)$ (from Guinier)	$0.52 \pm 0.30 \cdot 10^{-2}$	$0.24 \cdot 10^{-3} \pm 0.38 \cdot 10^{-6}$
R_g (Å) (from Guinier)	25.51 ± 0.19	30.26 ± 0.07
$I(0)$ (from $p(r)$)	$0.51 \pm 0.10 \cdot 10^{-2}$	$0.24 \cdot 10^{-3} \pm 0.21 \cdot 10^{-6}$
R_g (Å) (from $p(r)$)	24.78 ± 0.06	30.96 ± 0.03
Dmax (Å)	71	91
Molecular mass determination		
SAXSMoW (kDa)[a]	29.0	32.3
Q_R (kDa)[a]	28.1	30.2
Theoretical MM from sequence (kDa)	24.3	31.7
Model statistics		
χ^2	0.64	0.96
χ^2_{free}	0.92	1.15
R_{SAS} (%)	0.9927	0.0126

[a]For the q range reported above in the table.

with the position and orientation of the DNA relative to the *Ec*MazE$^{1–50}$ dimer in the complex (Supplementary Figure S7).

DISCUSSION

The *E. coli mazEF* operon was the first TA module that was identified on a chromosome (33) and remains one of the best characterized TA modules in terms of biochemical and physiological functions of the toxin. *E. coli mazEF* is autoregulated with MazE being the primary transcription factor and MazF modulating its activity. The *mazEF* operator consists of three consecutive independent MazE binding sites that differ in affinity up to one order of magnitude. Our binding data thus improve the model of three non-interacting, quasi-equivalent binding sites published earlier (34). Three *Ec*MazE dimers bind at the promoter *'cab'* sequence with apparent binding constants in the micromolar range and favorable enthalpic components dominating the Gibbs free energy. The interaction is specific and the presence of the intrinsically disordered *Ec*MazF-neutralizing tail does not significantly influences the affinity of the protein for either a single site or the complete operator (34).

We determined an accurate structural model of the *Ec*MazE–DNA complex using a combination of NMR and SAXS. This structure is in agreement with previous mutagenesis data and confirms that the C-terminal tail of *Ec*MazE remains disordered and is not directly involved in DNA-binding upon interaction between the N-terminal *Ec*MazE domain with DNA. *Ec*MazE binds into the major groove of double-stranded DNA *'a'*, involving side-chains of residues W9, N11, R16 for the main interactions with the oligonucleotide. Indeed, the R16A mutant is essentially inactive (27). In addition we could identify further electrostatic interactions that likely participate in stabilizing the *Ec*MazE–DNA complex, in particular between the positive-charged residues K7 and R8 and the DNA backbone. K7 and R8 were previously pointed as the primary DNA anchors for the MazE/MazF heterocomplex (32) while the R8A mutant shows reduced binding to the operator (27). Moreover, superposition of our *Ec*MazE–DNA complex on the complex between *Ec*MazE and a dromedary heavy chain antibody fragment indicates no structural clash (Supplementary Figure S8B) and thus confirms the correctness of our structural model as the presence of the heavy chain antibody fragment was shown to have no effect on the DNA-binding properties of MazE (27).

Our structural model resembles strongly the structure of *Rickettsia felis* VapB$_2$ (*Rf*VapB$_2$) in complex with its operator. While cataloged as a 'VapB' due to its association with a VapC toxin, this antitoxin contains an AbrB-type DNA binding domain similar to *Ec*MazE. Interestingly, *Rf*VapB$_2$ recognizes the same palindrome as *Ec*MazE (5′-ATATAT-3′) using identical interactions with the N-terminal β-strand and hairpin. Interactions differ nevertheless at the periphery of the combining site, where the structures of both proteins diverge. There alternative contacts are seen, such as the backbone NH of A19 in *Ec*MazE mimicking the interaction of the side chain of K19 from *R. felis* VapB with a DNA backbone phosphate.

The presence of *Ec*MazF influences operator recognition by *Ec*MazE (33). MazF proteins structurally resemble CcdB proteins and have a similar binding site for their cognate antitoxin. The antitoxins and toxins from both *ccdAB* and *mazEF* modules form chains of alternating toxin and antitoxin dimers (60,61). The F-plasmid *ccdAB* operon also contains an operator with multiple sites for the antitoxin, and the enhanced affinity of CcdA for its opera-

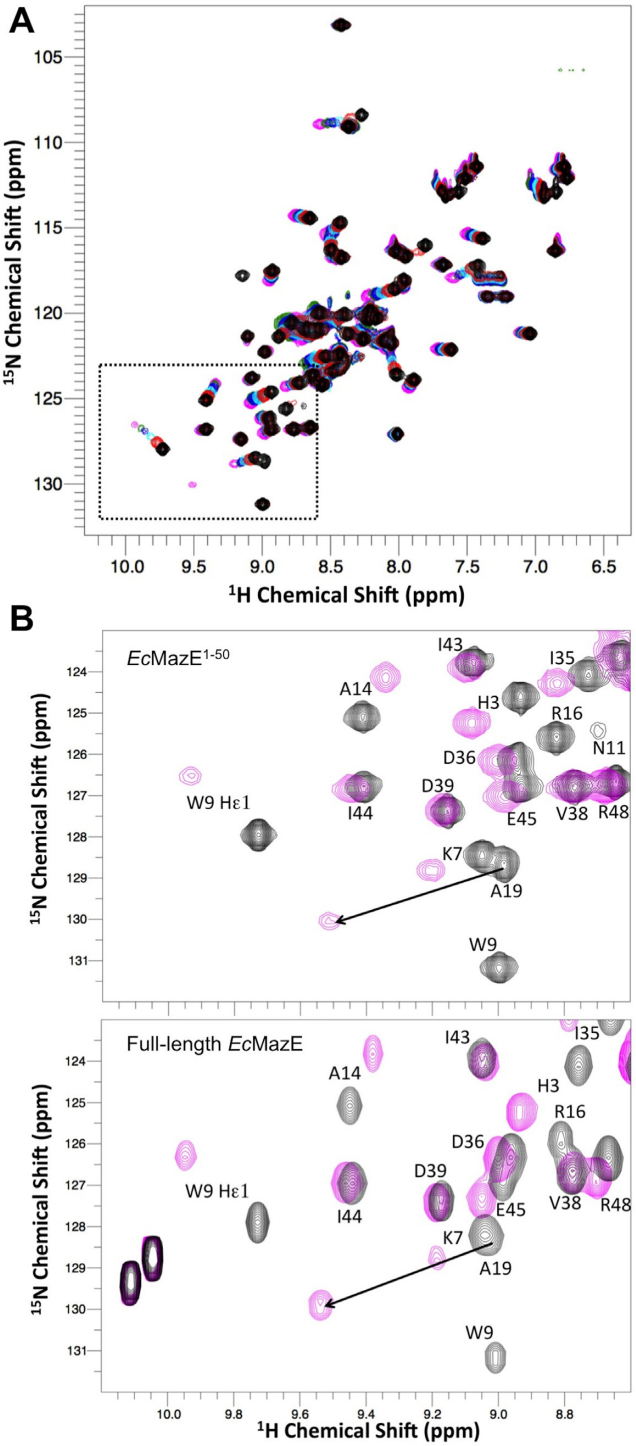

Figure 4. Binding of DNA *'a'* to *Ec*MazE monitored by NMR. (**A**) ^{1}H-^{15}N-HSQC spectra recorded during titrations of *Ec*MazE$^{1-50}$ with DNA *'a'*. All spectra are plotted at the same contour level and are colored from black (free form) to magenta (last titration point). (**B**) Selected region of ^{1}H-^{15}N-HSQC spectra recorded during titrations of *Ec*MazE$^{1-50}$ with DNA *'a'*. Only NMR spectra of the *Ec*MazE$^{1-50}$ free form (black) and of the last titration point (magenta) are shown. For clarification, the black arrow indicates the direction of the chemical shift changes for Ala 19. (**C**) The same selected region of the ^{1}H-^{15}N-HSQC spectra of free full-length *Ec*MazE (black) and in complex with DNA *'a'* (magenta).

tor in the presence of CcdB is believed to stem from an avidity effect (10). Native mass spectrometry data on the Kis/Kid module (36), a homolog of *E. coli* MazEF, support a similar model of regulation within the *mazEF* modules. We observed, however, that MazF can increase the affinity of MazE even for a single operator site where no avidity effects are present. This is unexpected, but can be explained by either direct interactions between *Ec*MazF and DNA or by thermodynamic stabilization of the N-terminal domain of *Ec*MazE through interaction with MazF. Such stabilization-induced affinity enhancement was previously observed in the *phd/doc* module (9). Superposition of our *Ec*MazE–DNA complex on the crystal structure of the *Ec*MazE–*Ec*MazF complex (32) (Supplementary Figure S8C) indicates additional protein–DNA interaction via the flanking basic regions of the *Ec*MazF homodimer. This favors a model where the enhancement in DNA binding by *Ec*MazF is caused by co-operative binding of the antitoxin and toxin to the DNA instead of an allosteric effect.

At very high *Ec*MazF to *Ec*MazE ratios, the affinity of MazE for the *'a'* operator site diminishes again. This resembles the conditional co-operativity phenomenon previously observed for the *ccdAB*, *phd/doc* and *relBE* modules (9–12). The phenomenon however occurs only at *Ec*MazF to *Ec*MazE ratios that are never attained *in vivo*, and its physiological relevance is thus uncertain. We are currently also unable to provide a satisfactory mechanistic explanation.

The antitoxin *Ec*MazE belongs to a large family of transcription regulators called the AbrB family, for which structural data on DNA recognition are relatively underrepresented compared to other major families of DNA binding domains. In order to gain more insight into the DNA recognition and specificity in binding by AbrB-like domains, we performed a comparative study using all structurally homologous proteins in the PDB. AbrB-like transcriptional factors show a swapped hairpin β-strand motif. Conservation of the βαβ core as the main structural unit supports a common evolutionary origin between this AbrB-like fold and the double-psi β barrels (62). A combined sequence and secondary structure alignment of representative proteins of this superfamily show two main families: one formed by AbrB itself and its closest relatives, and the other by the bacterial *Ec*MazE-like domain (Figure 8). The two families mainly differ in the positions of residues crucial for DNA binding: while family I is characterized by the presence of key residues in turn LP1 at positions 9 (W/N/S), 11 (Q/N/R) and the additional arginine at position 16, family II lacks this loop extension and DNA binding is driven by four key arginines at positions 8, 15, 23 and 24 (61). Figure 8B highlights the similar binding characteristics within family I, represented by *Ec*MazE and *Rf*VapB$_2$, and the different binding surface within family II, represented by *Bs*AbrB. Despite their similarities, *Ec*MazE and *Rf*VapB$_2$ differ in the position of additional positively charged residues important for protein–DNA stabilization, N-terminal of the LP1 turn in the case of *Ec*MazE and C-terminal in the case of *Rf*VapB$_2$.

Based on the binding characteristics of both families we included and accordingly categorized some uncharacterized bacterial proteins, *Pectobacterium atrosepticum* ChpR suppressor of growth inhibitor (*Pa*ChpR) and *Gloeobacter vi-*

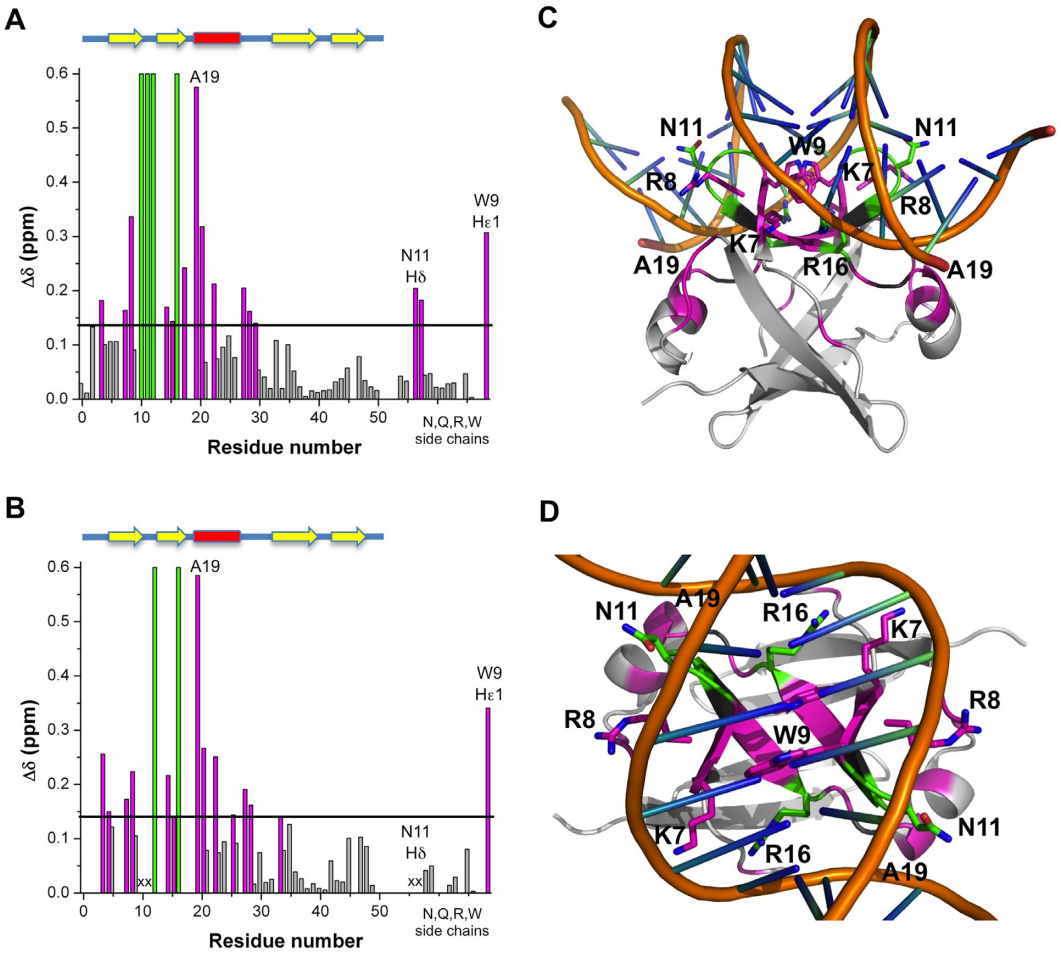

Figure 5. Chemical shift mapping of the DNA *'a'* binding sites on *Ec*MazE$^{1–50}$ and full-length *Ec*MazE. (**A**) Residue-specific DNA *'a'*-induced chemical shift changes *Ec*MazE$^{1–50}$. The chemical shift perturbations $\Delta\delta < \Delta\delta_{av} + \frac{1}{2}$ SD are colored gray, $\Delta\delta > \Delta\delta_{av} + \frac{1}{2}$ SD colored magenta, in green the residues of which their peak disappear upon addition of the DNA. The black line represents the chemical shift perturbations $\Delta\delta_{av} + \frac{1}{2}$ SD. Secondary structure elements within the *Ec*MazE$^{1–50}$ structure are indicated by yellow arrows (β-strands) and red bars (α-helices). (**B**) Residue-specific DNA *'a'*-induced chemical shift changes full-length *Ec*MazE. Color coding as in (A). Residues not visible in the HSQC spectra of free full-length *Ec*MazE are labeled as x. (**C, D**) Chemical shift mapping on the representative free NMR structure of *Ec*MazE$^{1–50}$ superimposed on VapB within the VapBC$_2$–DNA complex (showing only the DNA within the complex, PDB entry 3ZVK) as in Supplementary Figure S6. Color coding as in (A). Figures prepared using PyMol.

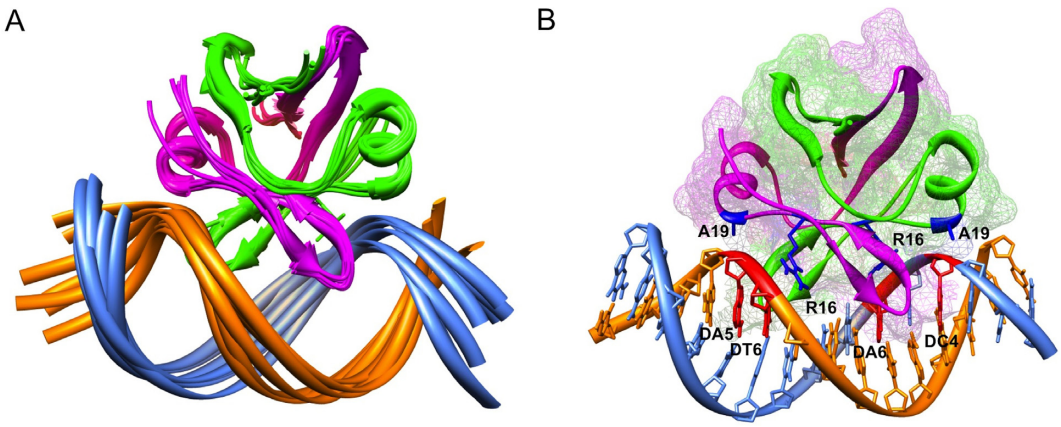

Figure 6. Structural model of the *Ec*MazE$^{1–50}$–DNA *'a'* complex. (**A**) Cartoon representation of the ensemble of the seven HADDOCK structures with the lowest interaction energies and lowest AIR violations. The two *Ec*MazE$^{1–50}$ monomers are colored green and magenta, the two DNA strands in orange and sky-blue. (**B**) Details of the *Ec*MazE$^{1–50}$–DNA complex showing the lowest-interaction energy structure of the ensemble. Color coding as in (A). The *Ec*MazE$^{1–50}$ dimer is also shown in mesh surface. Residues and nucleotides involved in H-bonding common in the ensemble are shown in blue (*Ec*MazE$^{1–50}$) and red (DNA) sticks, respectively. Figures prepared using Chimera.

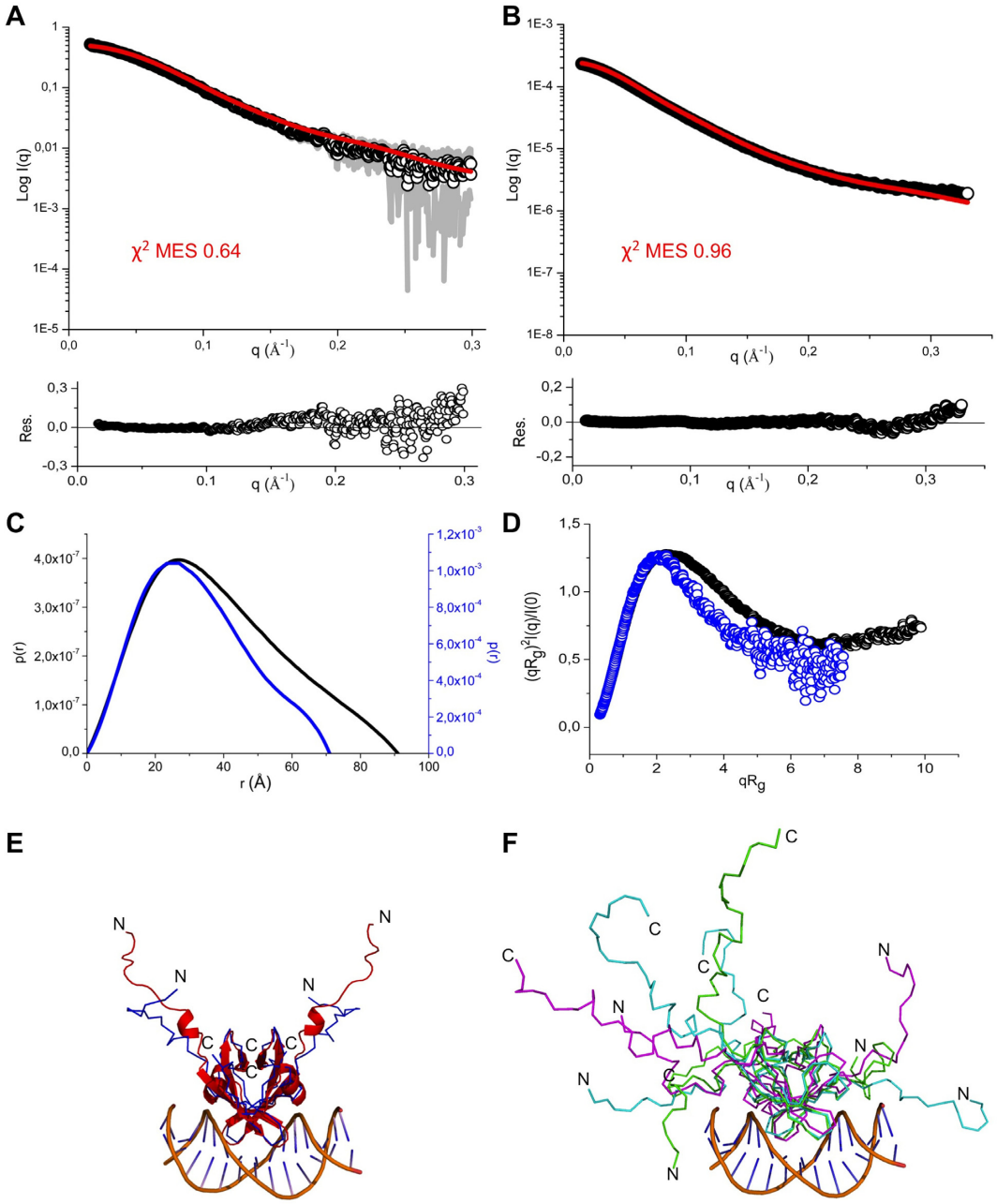

Figure 7. Small angle X-ray scattering of *Ec*MazE-DNA *'a'*. (**A, B**) The experimental SAXS curve for *Ec*MazE$^{1–50}$–DNA *'a'* and full-length *Ec*MazE–DNA *'a'* complexes are shown in open dots, while the error margins are shown in gray. The fit of the minimal search ensemble (MES) of two structures for *Ec*MazE$^{1–50}$-DNA *'a'* is reported in red (A), while for full-length *Ec*MazE-DNA *'a'*, the fit of the MES is with three structures (red line in (B)). The residual fitting is reported below for both systems. (**C**) Overlay of two *p(r)* function of *Ec*MazE$^{1–50}$-DNA *'a'* (blue) and full-length *Ec*MazE-DNA *'a'* (black). (**D**) Overlay of two normalized Kratky plots corresponding to the *Ec*MazE$^{1–50}$-DNA *'a'* (blue) and full-length *Ec*MazE-DNA *'a'* (black) shown as open dots. (**E, F**) A cartoon representation of the minimal set of two NMR structures of the *Ec*MazE$^{1–50}$–DNA *'a'* complex and three structures of the full-length *Ec*MazE–DNA *'a'* complex. The N-terminal His-tag and the extended flexible C-terminal tails are indicated by N and C. Panels (E) and (F) were created using PyMol.

olaceus cell growth regulatory protein (*Gv*), and archaea structures, such as the *Pyrococcus horikoshii* S018 putative uncharacterized protein (PhS018) and thermoacidophilic *Sulfolobus solfataricus* Sso7c4 (SSOL). Both *Pa*ChpR and *Gv* possess the key and additional residues for DNA binding as in *Ec*MazE and are thus predicted to have the same DNA-recognition site as the other members of this family.

PhS018 represents an archaeal intermediate between a double-psi (six β-strands) and a swapped-hairpin β-barrel consisting of four β-strands with a βαβ core. Despite the fact that PhS018 differs from the AbrB-like fold by the addition of one β-strand, it has been considered to be the dimeric ancestor of swapped-hairpin dimer (62). Alignment

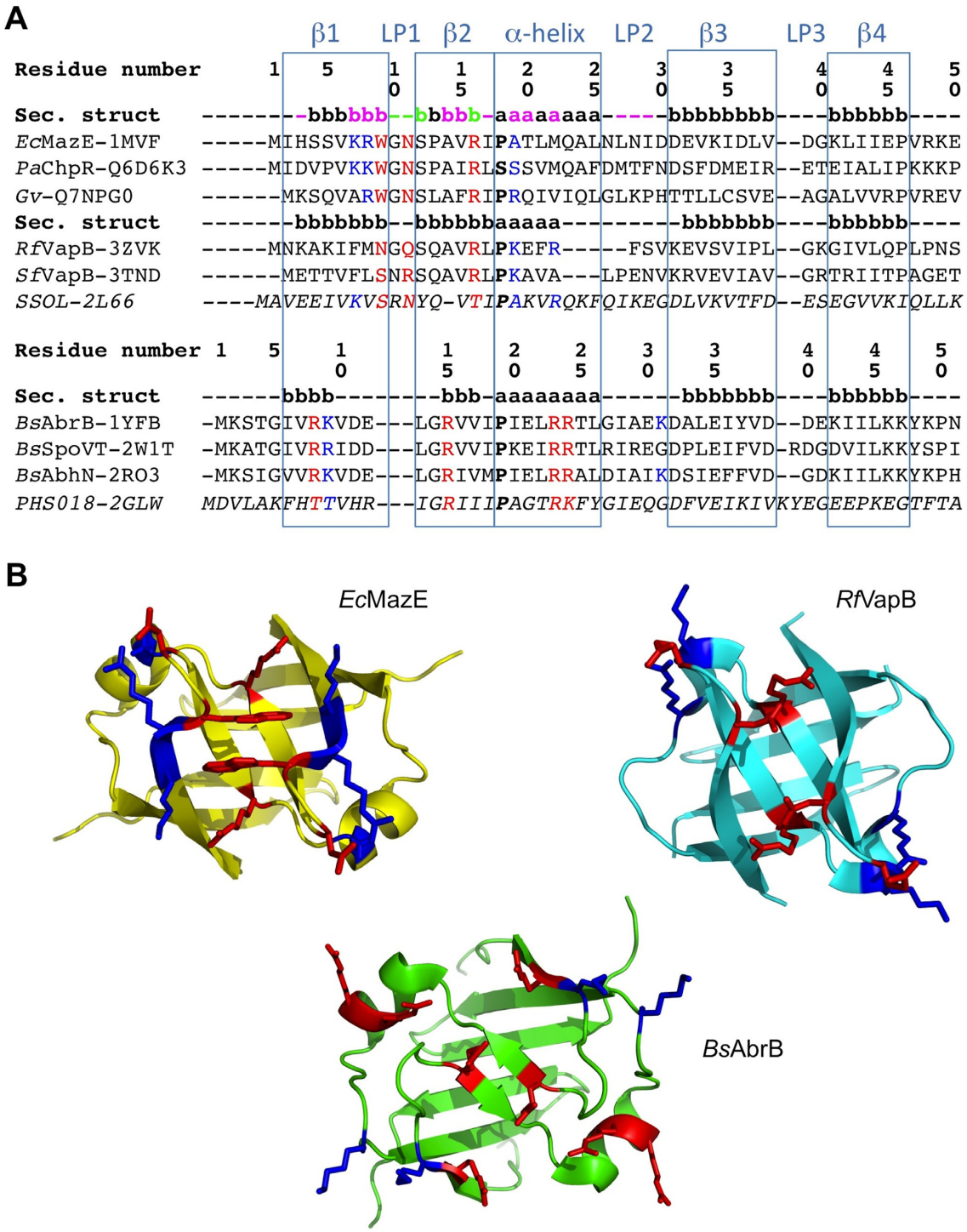

Figure 8. Structure-based sequence alignment of AbrB-like domain superfamily members. (**A**) Sequence alignment of the superfamily divided into two main families; the first one contains two subgroups. The consensus secondary structure within the superfamily is highlighted in light blue squares. Secondary structure elements within each family and sub-group are also given, representing the first one in each. The one belonging to *Ec*MazE is color-coded by the CSP given in Figure 5 ($\Delta\delta > \Delta\delta_{av} + \frac{1}{2}$ SD colored magenta, in green the residues of which their peaks disappear upon addition of the DNA). Residue numbering for the two families corresponds to that of *Ec*MazE and *Bs*AbrB, respectively. Key residues for DNA interaction are colored red, additional residues stabilizing the protein–DNA interaction in blue. The sequences from top to bottom are (corresponding PDB or Uniprot entries are given between parenthesis): *Ec*MazE: *Escherichia coli* MazE antitoxin (1MVF); *Pa*ChpR: *Pectobacterium atrosepticum* ChpR suppressor of growth inhibitor (Q6D6K3); Gv: *Gloeobacter violaceus* cell growth regulatory protein (Q7NPG0); *Rf*VapB: *Rickettsia felis* VapB antitoxin (3ZVK); *Sf*VapB: *Shigella flexneri* VapB antitoxin (3TND); *SSOL: Sulfolobus solfataricus* transcription regulator (2L66); *Bs*AbrB: *Bacillus subtilis* AbrB transition state regulator (1YFB); *Bs*SpoVT: *Bacillus subtilis* SpoVT stage V sporulation protein T (2W1T); *Bs*AbhN: *Bacillus subtilis* AbhN putative transition state regulator (2RO3); *PHS018: Pyrococcus horikoshii* S018 putative uncharacterized protein (2GLW). (**B**) Representative structures of the two AbrB-like domain families. Structures were superimposed using Pymol and thus are in the same orientation. Residues important for DNA binding are given in sticks and colored as defined in (A).

predicts this protein to recognize DNA in a similar way as members belonging to family II.

Differently, the archaeal DNA-binding homodimer Sso7c4 possesses residues that are highly conserved in the archaeal homologs (R11, N12, R22). Based on our alignment (Figure 8A), Sso7c4 aligns well with the *Ec*MazE family, including key residues at positions 9, 11 and 16 important for the DNA binding. Interestingly it contains two residues necessary for DNA stabilization, K8 like in the case of *Ec*MazE and R22 like in *Rf*VapB$_2$, thus representing a characteristic of both subgroups in family I.

DATABASE DEPOSITION

The ^{1}H, ^{15}N and ^{13}C assigned resonances of *EcMazE*$^{1–50}$ have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession number 25086. The structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2MRN.

The ^{1}H and ^{15}N assigned resonances of *Ec*MazE$^{1–50}$ in complex with DNA *'a'* have been deposited in the BioMagResBank under accession number 25092. The structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2MRU.

The ^{1}H, ^{15}N and ^{13}C assigned resonances of full-length *EcMazE* have been deposited in the BioMagResBank under accession number 25093. The ^{1}H and ^{15}N assigned resonances of full-length *Ec*MazE in complex with DNA *'a'* have been deposited in the BioMagResBank under accession number 25094.

ACCESSION NUMBERS

25086, 25092, 25093, 25094, 2MRN and 2MRU.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO); Vlaams Instituut voor Biotechnologie (VIB); Onderzoeksraad of the Vrije Universiteit Brussel (VUB); Hercules Foundation; Biostruct-X; FWO [to V.Z.]; Austrian Science Foundation (FWF) [19902, 25880 to K.Z.]; DOCfFORTE Fellowship, Austrian Academy of Sciences [to E.S.]; 'European Community-Access to Research Infrastructure Action of the Improving Human Potential Programme' EU-NMR. Funding for open access charge: Internal VIB grant.
Conflict of interest statement. None declared.

REFERENCES

1. Bravo,A., de Torrontegui,G. and Díaz,R. (1987) Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. *Mol. Gen. Genet.*, **210**, 101–110.
2. Brown,J.M. and Shaw,K.J. (2003) A novel family of *Escherichia coli* toxin-antitoxin gene pairs. *J. Bacteriol.*, **185**, 6600–6608.
3. Pandey,D.P. and Gerdes,K. (2005) Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.*, **33**, 966–976.
4. Gerdes,K. and Wagner,E.G.H. (2007) RNA antitoxins. *Curr. Opin. Microbiol.*, **10**, 117–124
5. Buts,L., Lah,J., Dao-Thi,M., Wyns,L. and Loris,R. (2005) Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends Biochem. Sci.*, **30**, 672–679.
6. Fineran,P.C., Blower,T.R., Foulds,I.J., Humphreys,D.P., Lilley,K.S. and Salmond,G.P.C. (2009) The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 894–899.
7. Wang,X., Lord,D.M., Cheng,H.Y., Osbourne,D.O., Hong,S.H., Sanchez-Torres,V., Quiroga,C., Zheng,K., Herrmann,T., Peti,W. *et al.* (2012) A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat. Chem. Biol.*, **8**, 855–861.
8. Wang,X., Lord,D.M., Hong,S.H., Peti,W., Benedik,M.J., Page,R. and Wood,T.K. (2013) Type II toxin/antitoxin MqsR/MqsA controls type V toxin/antitoxin GhoT/GhoS. *Environ. Microbiol.*, **15**, 1734–1744.
9. Garcia-Pino,A., Balasubramanian,S., Wyns,L., Gazit,E., De Greve,H., Magnuson,R.D., Charlier,D., van Nuland,N.A. and Loris,R. (2010) Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell*, **142**, 101–111.
10. De Jonge,N., Garcia-Pino,A., Buts,L., Haesaerts,S., Charlier,D., Zangger,K., Wyns,L., De Greve,H. and Loris,R. (2009) Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell*, **35**, 154–163.
11. Overgaard,M., Borch,J., Jørgensen,M.G. and Gerdes,K. (2008) Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Mol. Microbiol.*, **69**, 841–857.
12. Afif,H., Allali,N., Couturier,M. and Van Melderen,L. (2001) The ratio between CcdA and CcdB modulates the transcriptional repression of the *ccd* poison-antidote system. *Mol. Microbiol.*, **41**, 73–82.
13. Gerdes,K., Rasmussen,P.B. and Molin,S. (1986b) Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 3116–3120.
14. Roberts,R.C., Ström,A.R. and Helinski,D.R. (1994) The parDE operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. *J. Mol. Biol.*, **237**, 35–51.
15. Christensen,S.K., Mikkelsen,M., Pedersen,K. and Gerdes,K. (2001) RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 14328–14333.
16. Albrethsen,J., Agner,J., Piersma,S.R., Højrup,P., Pham,T.V., Weldingh,K., Jimenez,C.R., Andersen,P. and Rosenkrands,I. (2013) Proteomic profiling of *Mycobacterium tuberculosis* identifies nutrient-starvation-responsive toxin-antitoxin systems. *Mol. Cell. Proteomics.*, **12**, 1180–1191.
17. Aizenman,E., Engelberg-Kulka,H. and Glaser,G. (1996) An *Escherichia coli* chromosomal 'addiction module' regulated by guanosine 3′, 5′-bispyrophosphate: a model for programmed bacterial cell death. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 6059–6063.
18. Engelberg-Kulka,H., Amitai,S., Kolodkin-Gal,I. and Hazan,R. (2006) Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genet.*, **2**, 1528–1526.
19. Lewis,K. (2010) Persister cells. *Annu. Rev. Microbiol.*, **64**, 357–372.
20. Gerdes,K. and Maisonneuve,E. (2012) Bacterial persistence and toxin-antitoxin loci. *Annu. Rev. Microbiol.*, **66**, 103–123.
21. Tripathi,A., Dewan,P.C., Siddique,S.A. and Varadarajan,R. (2014) MazF-induced growth inhibition and persister generation in *Escherichia coli*. *J. Biol. Chem.*, **289**, 4191–205.
22. Loris,R. and Garcia-Pino,A. (2014) Disorder- and dynamics-based regulatory mechanisms in toxin-antitoxin modules. *Chem. Rev.*, **114**, 6933–6947.
23. Tsuchimoto,S., Ohtsubo,H. and Ohtsubo,E. (1988) Two genes, *pemK* and *pemI*, responsible for stable maintenance of resistance plasmid R100. *J. Bacteriol.*, **170**, 1461–1466.
24. Masuda,Y., Miyakawa,K., Nishimura,Y. and Ohtsubo,E. (1993). *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. *J. Bacteriol.*, **175**, 6850–6856.
25. Zhang,J., Zhang,Y., Hoeflich,K.P., Ikura,M., Qing,G. and Inouye,M. (2003) MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Mol. Cell*, **12**, 913–923.

26. Vesper,O., Amitai,S., Belitsky,M., Byrgazov,K., Kaberdina,A.C., Engelberg-Kulka,H. and Moll,I. (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell*, **147**, 147–157.
27. Loris,R., Marianovsky,I., Lah,J., Laeremans,T., Engelberg-Kulka,H., Glaser,G., Muyldermans,S. and Wyns,L. (2003) Crystal structure of the intrinsically flexible addiction antidote MazE. *J. Biol. Chem.*, **278**, 28252–28257.
28. Coles,M., Djuranovic,S., Söding,J., Frickey,T., Koretke,K., Truffault,V., Martin,J. and Lupas,A.N. (2005) AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels. *Structure*, **13**, 919–928.
29. Chumsakul,O., Takahashi,H., Oshima,T., Hishimoto,T., Kanaya,S., Ogasawara,N. and Ishikawa,S. (2011) Genome-wide binding profiles of the *Bacillus subtilis* transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation. *Nucleic Acids Res.*, **39**, 414–428.
30. Bagyan,I., Hobot,J. and Cutting,S. (1996) A compartmentalized regulator of developmental gene expression in *Bacillus subtilis*. *J. Bacteriol.*, **178**, 4500–4507.
31. Asen,I., Djuranovic,S., Lupas,A.N. and Zeth,K. (2009) Crystal structure of SpoVT, the final modulator of gene expression during spore development in *Bacillus subtilis*. *J. Mol. Biol.*, **386**, 962–975.
32. Kamada,K., Hanaoka,F. and Burley,S.K. (2003) Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Mol. Cell*, **11**, 875–884.
33. Marianovsky,I., Aizenman,E., Engelberg-Kulka,H. and Glaser,G. (2001) The regulation of the *Escherichia coli mazEF* promoter involves an unusual alternating palindrome. *J. Biol. Chem.*, **276**, 5975–5984.
34. Lah,J., Marianovsky,I., Glaser,G., Engelberg-Kulka,H., Kinne,J., Wyns,L. and Loris,R. (2003) Recognition of the intrinsically flexible addiction antidote MazE by a dromedary single domain antibody fragment. Structure, thermodynamics of binding, stability, and influence on interactions with DNA. *J. Biol. Chem.*, **278**, 14101–14111.
35. Bailey,S.E.S. and Hayes,F. (2009) Influence of operator site geometry on transcriptional control by the YefM-YoeB toxin-antitoxin complex. *J. Bacteriol.*, **191**, 762–772.
36. Monti,M.C., Hernandez-Arriaga,A.M., Kamphuis,M.B., Lopez-Villarejo,J., Heck,A.J.R., Boelens,R., Diaz-Orejas,R. and van den Heuvel,R.H.H (2007) Interactions of Kid–Kis toxin–antitoxin complexes with the *parD* operator-promoter region of plasmid R1 are piloted by the Kis antitoxin and tuned by the stoichiometry of Kid–Kis oligomers. *Nucleic Acids Res.*, **35**, 1737–1749.
37. Oberer,M., Zangger,K., Gruber,K. and Keller,W. (2007) The solution structure of ParD, the antidote of the ParDE toxin antitoxin module, provides the structural basis for DNA and toxin binding. *Protein Sci.*, **16**, 1676–1688.
38. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
39. Johnson,B.A. and Blevins,R.A. (1994) NMR View: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
40. Cavanagh,J., Fairbrother,W.J., Palmer,A.G., Rance,M. and Skelton,N.J. (2007) *Protein NMR Spectroscopy*, 2nd edn, Academic Press, San Diego, CA.
41. Vranken,W.F., Boucher,W., Stevens,T.J., Fogh,R.H., Pajon,A., Llinas,M., Ulrich,E.L., Markley,J.L., Ionides,J. and Laue,E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, **59**, 687–696.
42. Guntert,P., Mumenthaler,C. and Wuthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–298.
43. Herrmann,T., Guntert,P. and Wuthrich,K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR*, **24**, 171–189.
44. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
45. Nederveen,A.J., Doreleijers,J.F., Vranken,W., Miller,Z., Spronk,C.A., Nabuurs,S.B., Guntert,P., Livny,M., Markley,J.L., Nilges,M. *et al.* (2005) RECOORD: a recalculated coordinate database of 500 +proteins from the PDB using restraints from the BioMagResBank. *Proteins*, **59**, 662–672.
46. Brünger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
47. Dominguez,C., Boelens,R. and Bonvin,A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, **125**, 1731–1737.
48. Hubbard,S. and Thornton,J. (1993) Naccess—Computer Program. Department of Biochemistry and Molecular Biology, University College London.
49. Maté,M.J., Vincentelli,R., Foos,N., Raoult,D., Cambillau,C. and Ortiz-Lombardía,M. (2011) Crystal structure of the DNA-bound VapBC2 antitoxin/toxin pair from *Rickettsia felis*. *Nucleic Acids Res.*, **40**, 3245–3258.
50. 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.
51. Konarev,P.V., Volkov,V.V., Sokolova,A.V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst*, **36**, 1277–1282.
52. Konarev,P.V., Petoukhov,M.V., Volkov,V.V. and Svergun,D.I. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Cryst.*, **39**, 277–286.
53. Rambo,R.P. and Tainer,J.A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481.
54. Pelikan,M., Hura,G.L. and Hammel,M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.*, **28**, 174–189.
55. Schneidman-Duhovny,D., Hammel,M. and Sali,A. (2010) FoXs: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.*, **38**, W540–W544.
56. Weinkam,P., Pons,J. and Sali,A. (2012) Structure-based model of allostery predicts coupling between distant sites. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 4875–4880.
57. Sterckx,Y.G., Volkov,A.N., Vranken,W.F., Kragelj,J., Jensen,M.R., Buts,L., Garcia-Pino,A., Jové,T., Van Melderen,L., Blackledge,M. *et al.* (2014) Small-angle X-Ray scattering- and nuclear magnetic resonance-derived conformational ensemble of the highly flexible antitoxin PaaA2. *Structure*, **22**, 854–865.
58. Zorzini,V., Buts,L., Sleutel,M., Garcia-Pino,A., Talavera,A., Haesaerts,S., Greve,H.D., Cheung,A., van Nuland,N.A. and Loris,R. (2014) Structural and biophysical characterization of Staphylococcus aureus SaMazF shows conservation of functional dynamics. *Nucleic Acids Res.*, **42**, 6709–6725.
59. Fischer,H., de Oliveira Neto,M., Napolitano,H.B., Craievich,A.F. and Polikarpov,I. (2010) The molecular weight of proteins in solution can be determined from a single SAXS measurement on a relative scale. *J. Appl. Cryst.*, **43**, 101–109.
60. Dao-Thi,M.H., Charlier,D., Loris,R., Maes,D., Messens,J., Wyns,L. and Backmann,J. (2002) Intricate interactions within the *ccd* plasmid addiction system. *J. Biol. Chem.*, **277**, 3733–3742.
61. Kamphuis,M.B., Monti,M.C., van den Heuvel,R.H., Santos-Sierra,S., Folkers,G.E., Lemonnier,M., Díaz-Orejas,R., Heck,A.J. and Boelens,R. (2007) Interactions between the toxin Kid of the bacterial *parD* system and the antitoxins Kis and MazE. *Proteins*, **67**, 219–231.
62. Coles,M., Hulko,M., Djuranovic,S., Truffault,V., Koretke,K., Martin,J. and Lupas,A.N. (2006) Common evolutionary origin of swapped-hairpin and double-psi beta barrels. *Structure*, **14**, 1489–1498.