

# The crystal structure of the Rv0301-Rv0300 VapBC-3 toxin–antitoxin complex from *M. tuberculosis* reveals a Mg<sup>2+</sup> ion in the active site and a putative RNA-binding site

Andrew B. Min,<sup>1</sup> Linda Miallau,<sup>1†</sup> Michael R. Sawaya,<sup>1</sup> Jeff Habel,<sup>2‡</sup> Duilio Cascio,<sup>1</sup> and David Eisenberg<sup>1\*</sup>

<sup>1</sup>Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, UCLA-DOE Institute of Genomics and Proteomics, UCLA, Los Angeles, California 90095-1570

<sup>2</sup>E. O. Lawrence Berkeley National Laboratory, University of California at Berkeley, CA 94720

Received 10 July 2012; Revised 11 September 2012; Accepted 12 September 2012

DOI: 10.1002/pro.2161

Published online 25 September 2012 proteinscience.org

**Abstract:** VapBC pairs account for 45 out of 88 identified toxin-antitoxin (TA) pairs in the *Mycobacterium tuberculosis* (Mtb) H37Rv genome. A working model suggests that under times of stress, antitoxin molecules are degraded, releasing the toxins to slow the metabolism of the cell, which in the case of VapC toxins is via their RNase activity. Otherwise the TA pairs remain bound to their promoters, autoinhibiting transcription. The crystal structure of Rv0301-Rv0300, an Mtb VapBC TA complex determined at 1.49 Å resolution, suggests a mechanism for these three functions: RNase activity, its inhibition by antitoxin, and its ability to bind promoter DNA. The Rv0301 toxin consists of a core of five parallel beta strands flanked by alpha helices. Three proximal aspartates coordinate a Mg<sup>2+</sup> ion forming the putative RNase active site. The Rv0300 antitoxin monomer is extended in structure, consisting of an N-terminal beta strand followed by four helices. The last two helices wrap around the toxin and terminate near the putative RNase active site, but with different conformations. In one conformation, the C-terminal arginine interferes with Mg<sup>2+</sup> ion coordination, suggesting a mechanism by which the antitoxin can inhibit toxin activity. At the N-terminus of the antitoxin, two pairs of Ribbon-Helix-Helix (RHH) motifs are related by crystallographic twofold symmetry. The resulting hetero-octameric complex is similar to the FitAB system, but the two RHH motifs are about 30 Å closer together in the Rv0301-Rv0300 complex, suggesting either a different span of the DNA recognition sequence or a conformational change.

**Keywords:** tuberculosis; toxin-antitoxin; toxin; antitoxin; vapBC; vapB; vapC; protein complex; Rv0301; Rv0300; PIN-domain

---

**Abbreviations:** Mtb, *Mycobacterium tuberculosis*; RHH, ribbon-helix-helix; RMSD, root mean square deviation; SASA, solvent accessible surface area; sc, shape complementarity; TA, toxin-antitoxin.

Additional Supporting Information may be found in the online version of this article.

<sup>†</sup>Current address: Uniformed Services University of the Health Sciences, Bethesda, MD 20814.

<sup>‡</sup>Current address: Protein Technologies, Bio-Rad Laboratories, Hercules, CA 94547.

Grant sponsor: National Science Foundation; Grant number: MCB-0958111; Grant sponsor: National Institutes of Health; Grant number: A1068135-04; Grant sponsor: U.S. Department of Energy; Grant number: DE-FC02-02ER63421.

<sup>\*</sup>Correspondence to: David Eisenberg, Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, UCLA-DOE Institute of Genomics and Proteomics, Box 951570, UCLA, Los Angeles CA 90095-1570. E-mail: david@mbi.ucla.edu

**Table I.** Statistics of X-Ray Data Collection and Atomic Refinement

	SeMet	Native
Space group	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å)	84.7, 84.7, 154.0	85.6, 85.6, 155.6
Resolution range (Å)	74.19–1.85 (1.89–1.85)	60.52–1.49 (1.53–1.49)
<i>R</i> <sub>merge</sub> (%)	7.3 (38.6)	4.9 (46.1)
Total number of reflections	45,018 (3283)	89,761 (6920)
Redundancy	21.2 (21.6)	8.7 (8.2)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
<i>I</i> / <i>σ</i> ( <i>I</i> )	37.3 (7.1)	36.9 (4.0)
Number of residues/asymmetric unit	413	418
Number of protein atoms	3237	3331
Number of ligand atoms	24	40
Number of water atoms	258	435
Matthews' coefficient (Å <sup>3</sup> /Da)	2.69	2.78
<i>R</i> (%)	17.4 (20.6)	15.6 (21.3)
<i>R</i> <sub>free</sub> (%)	19.0 (25.4)	17.4 (27.0)
Test set size (%), selection	5.0, random	5.0, random
RMSD bond lengths (Å)	0.034	0.015
RMSD bond angles (°)	2.052	1.486
Ramachandran angles		
Most favored (%)	97.0	97.3
Additionally allowed (%)	3.0	2.7
Generously allowed (%)	0.0	0.0
Disallowed (%)	0.0	0.0
Errat (44) overall quality factor (%)	99.7	97.0
Verify3D (45) residues with score >0.2 (%)	84.4	91.4
Average <i>B</i> factor for protein atoms (Å <sup>2</sup> )	9.1	20.6
Average <i>B</i> factor for ligand atoms (Å <sup>2</sup> )	24.5	24.5
Average <i>B</i> factor for water atoms (Å <sup>2</sup> )	13.9	22.8

Numbers in parentheses refer to the outer shell of data.

## Introduction

Toxin-antitoxin (TA) loci are found in many prokaryotes and archaea. TA are typically organized as bicistronic operons that encode a stable toxin and a labile antitoxin.<sup>1</sup> The toxin protein inhibits growth, or kills the cell at high enough doses, and its cognate antitoxin neutralizes the toxin by forming a tight complex. Under certain conditions, such as nutritional stress, the antitoxin is degraded and leaves an active toxin.<sup>2</sup>

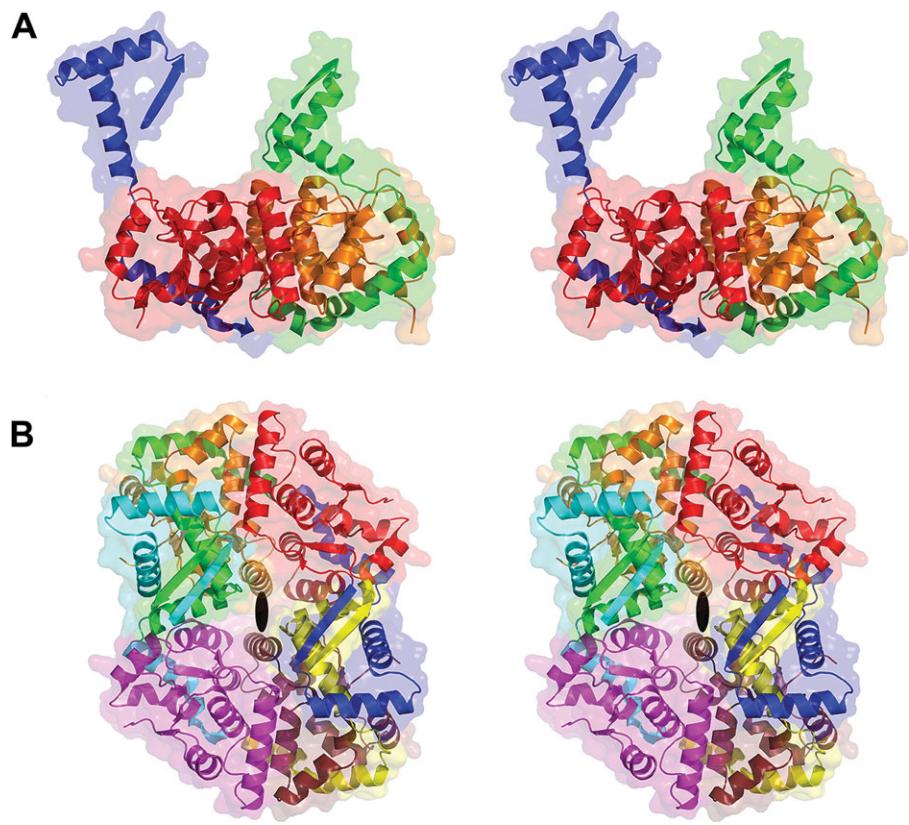
The first toxin structure to be determined was the *E. coli* F plasmid CcdB toxin.<sup>3</sup> The structures of many toxins, antitoxins, and TA complexes have been determined since, including *E. coli* plasmid R1 Kid toxin,<sup>4</sup> *E. coli* MazE-MazF TA complex,<sup>5</sup> broad-host-range low-copy-number plasmid pSM19035 from *Streptococcus pyogenes* ε<sub>2</sub>ζ<sub>2</sub> TA complex,<sup>6</sup> *E. coli* YoeB toxin and YefM2-YoeB TA complex,<sup>7</sup> archaeal *Pyrococcus horikoshii* aRelE-aRelB TA complex,<sup>8</sup> *Neisseria gonorrhoeae* FitAB TA complex,<sup>9</sup> solution structure of the broad host range low-copy-number plasmid pRK2/RP4 ParD antitoxin,<sup>10</sup> *Mycobacterium tuberculosis* (Mtb) YefM antitoxin,<sup>11</sup> and Mtb VapBC-5 TA complex.<sup>12</sup>

The genome of Mtb H37Rv strain encodes at least 88 TA loci covering five families of toxins.<sup>13</sup> The mechanisms have been elucidated for at least one member from four out of the five families. A few members of the RelE, HigB, and MazF family members have been shown to be mRNA-specific endonucleases, and have been termed interferases.<sup>14–18</sup>

ParE toxins have been shown in *E. coli* to block DNA replication by inhibiting DNA gyrase activity.<sup>19</sup>

The VapBC family is by far the largest TA family by numbers, with 45 identified in Mtb H37Rv,<sup>13</sup> but its function remains unclear. The VapC toxins are members of the PIN-domain family which are characterized by an α/β/α sandwich topology and four conserved acidic residues.<sup>20</sup> Computational analysis has suggested that VapC toxins are ribonucleases,<sup>20</sup> and there is increasing biochemical data that support the idea. For example, VapC-1 from *Haemophilus influenzae*,<sup>21</sup> VapC-5 from Mtb,<sup>12</sup> VapC-1 and VapC-2 from *Rickettsia felis*, and VapC-1 from *Rickettsia bellii*<sup>22</sup> have been reported to exhibit ribonuclease activity.

In spite of the knowledge gained regarding the biochemical function of VapC toxins, their physiological roles are largely unknown. The fitAB operon of the sexually transmitted pathogen *Neisseria gonorrhoeae*<sup>23</sup> was discovered through genetic screens, unlike the majority of vapBC loci which were discovered through bioinformatics. FitAB deletion mutants grow faster than wild-type inside epithelial cells, and subsequent structures have shown that the FitA antitoxin consists of an N-terminal ribbon-helix-helix (RHH) motif, and a C-terminal FitB toxin-binding interface.<sup>24</sup> The RHH motif is a conserved three-dimensional structural motif that binds to DNA in sequence-specific manner. FitA binds to its own promoter and the FitAB complex binds the same DNA



**Figure 1.** Stereo views of the Rv0301-Rv0300 complex. A: The asymmetric unit consists of two Rv0301 toxin molecules, represented as red and orange, and two Rv0300 antitoxin molecules, represented as blue and green. Notice how the green antitoxin (closed conformation) wraps around the orange toxin (Mg-unbound) more tightly than the blue antitoxin (open conformation) does the red toxin (Mg-bound). B: Two asymmetric units related by a 2-fold symmetry form a closed heter-octamer. The symmetry related pairs are red and magenta (Mg-bound toxins), orange and ruby (Mg-unbound toxins), blue and cyan (open antitoxins), and green and yellow (closed antitoxins).

sequence with a much higher affinity than the FitA alone. This autoregulatory nature is common to many TA loci including vapBC loci. The FitAB complex was described as a tetramer of FitAB heterodimers. No ribonuclease activity was detected for FitB but it was suggested that it may target a specific sequence.<sup>9,24</sup>

Mtb is known to lie dormant inside macrophages, and the biological roles of at least some of its numerous vapBC loci may be to slow cell growth for presumed survival advantages, similar to the roles proposed for FitAB. FitAB has been hypothesized to allow *N. gonorrhoeae* to hide inside host cells to evade immune responses and to allow some hosts to remain asymptomatic carriers.<sup>9,23,24</sup> Mtb follows parallel strategies. The structure of the FitAB TA complex bound to DNA revealed an octamer.<sup>9</sup> Recently, two more VapBC complexes have been described as forming octamers: VapBC from *Shigella flexneri*<sup>25</sup> and VapBC2 from *Rickettsia felis*.<sup>26</sup>

Our 1.49 Å crystal structure of the VapBC complex Rv0301-Rv0300 from Mtb is the highest resolution TA complex to date, and reveals a bound Mg<sup>2+</sup> ion at the active site. Furthermore, the detailed structure of the VapC toxin suggests distinct RNA-binding grooves and antitoxin-binding grooves.

Our interest in Mtb TA pairs is the possibility that they are key players in tuberculosis persistence, which is thought to be an essential mechanism by which Mtb acquires drug resistance.

## Results

### **Overall structure of the Rv0301-Rv0300 VapBC-3 toxin-antitoxin complex**

We determined the structure of the Rv0301-Rv0300 complex by single anomalous wavelength dispersion, using selenomethionine substitutions. Six selenium atoms were located in the asymmetric unit, which turned out to be three per TA pair out of a total of six methionine residues. Electron density was visible for T2 to R137 of the chain A toxin, T2 to P139 of the chain B toxin (out of a wild type sequence of 141 amino acids), and all but the start methionines of the antitoxins chain C and chain D (out of a wild type sequence of 73 amino acids). Additionally, five glycerol molecules, one imidazole molecule, one Mg<sup>2+</sup> ion, one β-mercaptoethanol molecule, and 426 water molecules were modeled. The native structure was refined at 1.49 Å resolution to a *R*<sub>work</sub> of 0.16 and an *R*<sub>free</sub> of 0.17. Statistics for data collection and refinement are

shown in Table I. The asymmetric unit consists of two Rv0301 toxin molecules and two Rv0300 antitoxin molecules [Fig. 1(A)]. The two toxin molecules are related by a pseudo twofold symmetry with small differences in conformations, in which chain A coordinates a Mg<sup>2+</sup>-ion (Mg-bound toxin) and chain B does not (Mg-unbound toxin). The two antitoxin molecules, however, have significantly different conformations with chain C (closed antitoxin) wrapped around the Mg-unbound toxin and chain D (open antitoxin) wrapped around the Mg-bound toxin. Two asymmetric units related by a twofold symmetry form a closed hetero-octamer [Fig. 1(B)].

#### **The Rv0301 toxin has a PIN-domain fold and coordinates a Mg<sup>2+</sup> ion**

The Rv0301 toxin molecule forms a tightly packed globular protein with the characteristics of a canonical PIN-domain protein, such as an  $\alpha/\beta/\alpha$  sandwich topology and four conserved acidic residues. The conserved residues are D9, E43, D99, and D117. The sandwich consists of five parallel beta strands flanked by four N-terminal alpha helices and three C-terminal alpha helices [Fig. 2(A)]. Despite low sequence identity, the PIN-domain fold is structurally well conserved [Table II and Fig. 2(B)]. In one of the subunits (chain A), the last two of these residues (D99 and D117) cluster together with another residue, D119, and three water molecules to coordinate an atom [Fig. 3(A,B)]. The 2.2–2.3 Å separations and the octahedral geometry indicate that the atom is a Mg<sup>2+</sup> ion, which is why we term the subunit Mg-bound toxin. No such coordination is observed in the other toxin subunit (chain B), which we term the Mg-unbound toxin. The two toxin subunits are related by a pseudo twofold symmetry: structural alignment of the Mg-bound and Mg-unbound toxins result in a C<sub>x</sub> RMSD of 0.6 Å. The major differences are observed around the peptide bond linking G94 to P95, which is trans in the Mg-bound form and cis in the Mg-unbound form, and the loop formed by the residues R89 to S96 which shifts positions. Another major difference is in the ion binding site. The 2.3 Å distances between the Mg<sup>2+</sup> ion and D99, D117, and D119 in the Mg-bound form would be increased to estimated distances of 2.5 Å, 2.5 Å, and 3.6 Å respectively in the Mg-unbound form [Fig. 3(C)].

Interestingly, of the three Mg<sup>2+</sup>-coordinating aspartates, only D99 and D117 are conserved in PIN-domain sequences, and only D99 is conserved structurally [Fig. 3(D)]. The only other VapC structure solved with a metal ion bound at the active site is PAE0151 (PDBid 2FE1), which coordinates a Mn<sup>2+</sup>-ion. In the PAE0151 structure, only two residues, D100 and D118 (corresponding respectively to D99 and D117 of Rv0301), participate in Mn<sup>2+</sup>-coordination, with D100 making indirect contact via two water molecules and D118 making direct contact.<sup>29</sup>

#### **The Rv0300 antitoxin has an extended structure consisting of an N-terminal RHH domain and a C-terminal toxin-binding domain joined by a flexible hinge loop**

The Rv0300 antitoxin molecule has an extended structure consisting of an N-terminal β-strand followed by four α-helices, which are linked by loops and a short 3<sub>10</sub>-helix [Fig. 4(A)]. The N-terminal β1-α1-α2 is identified as a RHH domain by Pfam,<sup>30</sup> a structural motif common in transcription factors that allow sequence-specific DNA-binding. The long loop between α2 and α3 (Q41-T48) forms a hinge, and part of the hinge along with the C-terminal α3-linker-α4 comprises the interface with the toxin. Structural alignments by Dahl<sup>31</sup> indicate that the closest homolog to Rv0300 is the orthologous VapB antitoxin FitA from *Neisseria gonorrhoeae*. Figure 4(B) shows a structure-based sequence alignment of the Rv0300 antitoxin (chain C) to FitA (PDBid 2H1O, chain F). Rv0300 and FitA share a sequence identity of 17% and align across all residues to a C<sub>x</sub> RMSD of 5.5 Å. This large deviation arises mostly from the hinge and the toxin-binding domain: aligning just the first 44 residues results in a C<sub>x</sub> RMSD of 0.49 Å.

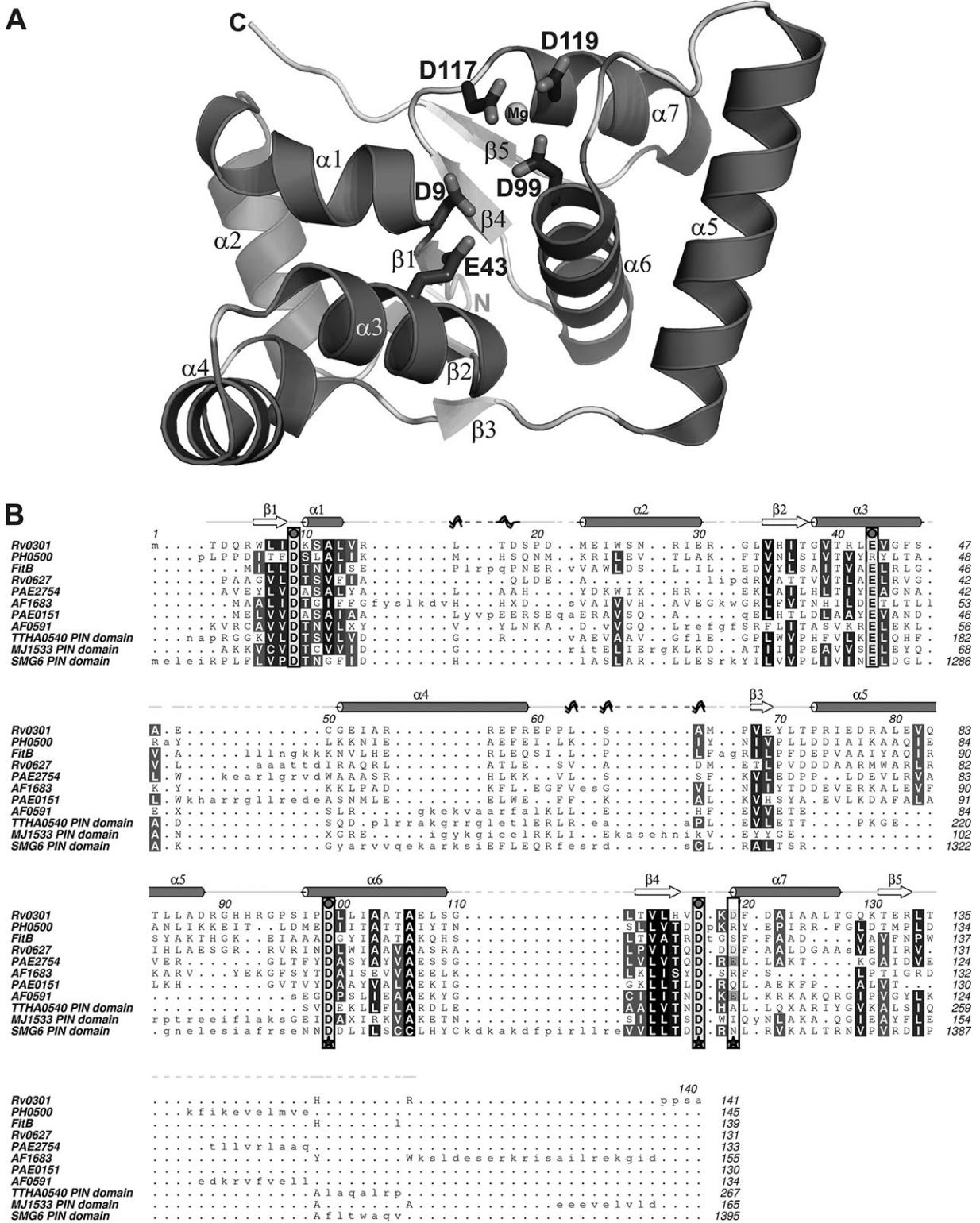
Two antitoxin molecules interact through their N-terminal strands, forming a dimer to complete the RHH motif. The β-strands from each antitoxin form an antiparallel sheet [Fig. 4(C)]. Nearly 40% of the residues participate in the antitoxin-antitoxin dimerization and over 1300 Å<sup>2</sup> of each antitoxin solvent accessible surface area (SASA) is buried in the interface. The antitoxin dimer is further stabilized with salt bridges formed by (chain D-chain C) D3-R7 (β1), R7-D3 (β1) and D39-R35 (α2).

In contrast to the limited conformational differences between the toxin molecules, the antitoxins have significantly different conformations due to their differing hinges [Fig. 4(D,E)]. In the open TA interface, the hinge twists outwards with respect to its cognate toxin, shifting it slightly out of position. In particular, residues Q41, T42, and A43 do not interact with the toxin at all, whereas in the closed interface, these residues have respectively 9%, 76%, and 48% of their accessible surface areas buried in the toxin interface.

The antitoxins have one other segment of flexibility spanned by residues 68–73 [Fig. 4(F)]. Residues 68–73 of the two antitoxins in the asymmetric unit clash sterically, and in our crystal structure this segment was modeled as alternately being present in one chain and disordered in the other chain. The occupancy of residues 68–73 was modeled at 0.6 and 0.4 respectively for the closed chain and the open chain.

#### **The Rv0301-Rv0300 complex forms an octamer**

Two asymmetric units related by a twofold symmetry complete a closed hetero-octamer, in which the bulk of the interactions between the two asymmetric units are formed within the two antitoxin-antitoxin dimers.



**Figure 2.** Structure of the Rv0301 VapC toxin. A: A cartoon diagram of Rv0301 toxin showing the  $\alpha/\beta/\alpha$  topology. D9, D99, and D117 are conserved in PIN-domain proteins. Notice the three aspartates – D99, D117, and D119 – that coordinate a  $Mg^{2+}$ -ion represented as a sphere. B: Structure-based sequence alignment of Rv0301 and close homologues. Secondary structure elements are based on the Rv0301 structure. The conserved residues are coded in shades of grey with darker shades representing more highly conserved residues. The columns corresponding to the four highly conserved acidic residues of PIN-domain proteins are boxed with a circle on top, and the residues involved in  $Mg^{2+}$ -coordination are marked with a star on the bottom. Notice that D9 and E43 are highly conserved but do not interact with  $Mg^{2+}$ , while D119 is poorly conserved but does interact with  $Mg^{2+}$ . The alignment diagram was generated with Aline.<sup>27</sup>

**Table II.** Comparison of Rv0301 to Structural Homologues Using DALI

Name	PDB ID (Chain)	Organism	Z	RMSD (Å)	lali	nres	%id	Description
Rv0301	3H87 (A)	<i>Mycobacterium tuberculosis</i>	28.9	0	136	136	100	Hypothetical protein (VapC toxin)
PH0500	1V96 (B)	<i>Pyrococcus horikoshii</i>	15.6	2.9	130	144	18	Hypothetical protein
FitB	2H1C (A)	<i>Neisseria gonorrhoeae</i>	15	2.3	123	139	17	Trafficking protein B (VapC toxin)
Rv0627	3DBO (B)	<i>Mycobacterium tuberculosis</i>	13.7	2.6	115	126	22	Hypothetical protein (VapC toxin)
PAE2754	1V8P (I)	<i>Pyrobaculum aerophilum</i>	12.2	2.8	114	132	25	Hypothetical protein
AF1683	1W8I (A)	<i>Archaeoglobus fulgidus</i>	11.3	3.1	121	155	21	Hypothetical protein
PAE0151	2FE1 (A)	<i>Pyrobaculum aerophilum</i>	10.7	2.8	112	130	17	Conserved hypothetical protein (VapC toxin)
AF0591	1O4W (A)	<i>Archaeoglobus fulgidus</i>	9.3	3.1	97	125	16	PIN domain protein
TTHA0540	3IX7 (B)	<i>Thermus thermophilus</i>	9.1	3.3	100	130	17	PIN domain of uncharacterized protein
MJ1533	3I8O (A)	<i>Methanocaldococcus jannaschii</i>	8.5	3.1	96	140	14	PIN domain of KH domain-containing protein
SMG6	2HWW (B)	<i>Homo sapiens</i>	7.9	2.7	97	157	10	PIN domain of telomerase-binding protein EST1A

lali gives the number of aligned positions; nres gives the total number of residues in matched structure.

The shape complementarity score (sc) for the anti-toxin-antitoxin interface is 0.64.<sup>32</sup> The sc scores and the extent of buried surface areas in the interfaces indicate that all of the protein-protein interaction interfaces are tight and extensive. Dynamic light scattering (data not shown) and small-angle X-ray scattering data (Supporting Information) indicate that the complex is monodisperse in solution and confirms that it exists as a stable octamer.

The toxin homodimers interact primarily by docking  $\alpha$ 5 between  $\alpha$ 3 and  $\alpha$ 4 giving a pseudo two-fold symmetry [Fig. 5(A)]. Thirty-eight residues from each toxin make contributions to the interface, burying  $\sim$ 1400 Å<sup>2</sup> of the total SASA of each toxin with a sc score of 0.69. The dimer is further stabilized by five salt bridges, formed by (chain B-chain A) E49-R93, R55-D88, D88-R55, R93-E49, and the symmetry-breaking H92-E49. All known structures of prokaryotic PIN domain proteins form dimers.

The asymmetric unit can best be described as a dimer of TA heterodimers, held together by the toxin-toxin interactions. In the closed TA interaction, 1490 Å<sup>2</sup> of the toxin and 1620 Å<sup>2</sup> of the antitoxin SASA are buried in the interface. Out of 141 toxin residues 51 and out of 73 antitoxin residues 35 contribute to the interface with a sc score of 0.73. In the open TA, 1090 Å<sup>2</sup> of the toxin and 1180 Å<sup>2</sup> of the antitoxin are buried. Thirty-six toxin residues and 23 antitoxin residues contribute to the interface with a sc score of 0.70. A toxin E57 to antitoxin R54 salt bridge is present in both closed and open interfaces, but a toxin D33 to antitoxin R56 salt bridge is seen only in the

closed interface, and a toxin E30 to antitoxin R44 salt bridge is seen only in the open interface.

### **The C-terminus of the antitoxin binds two toxins together**

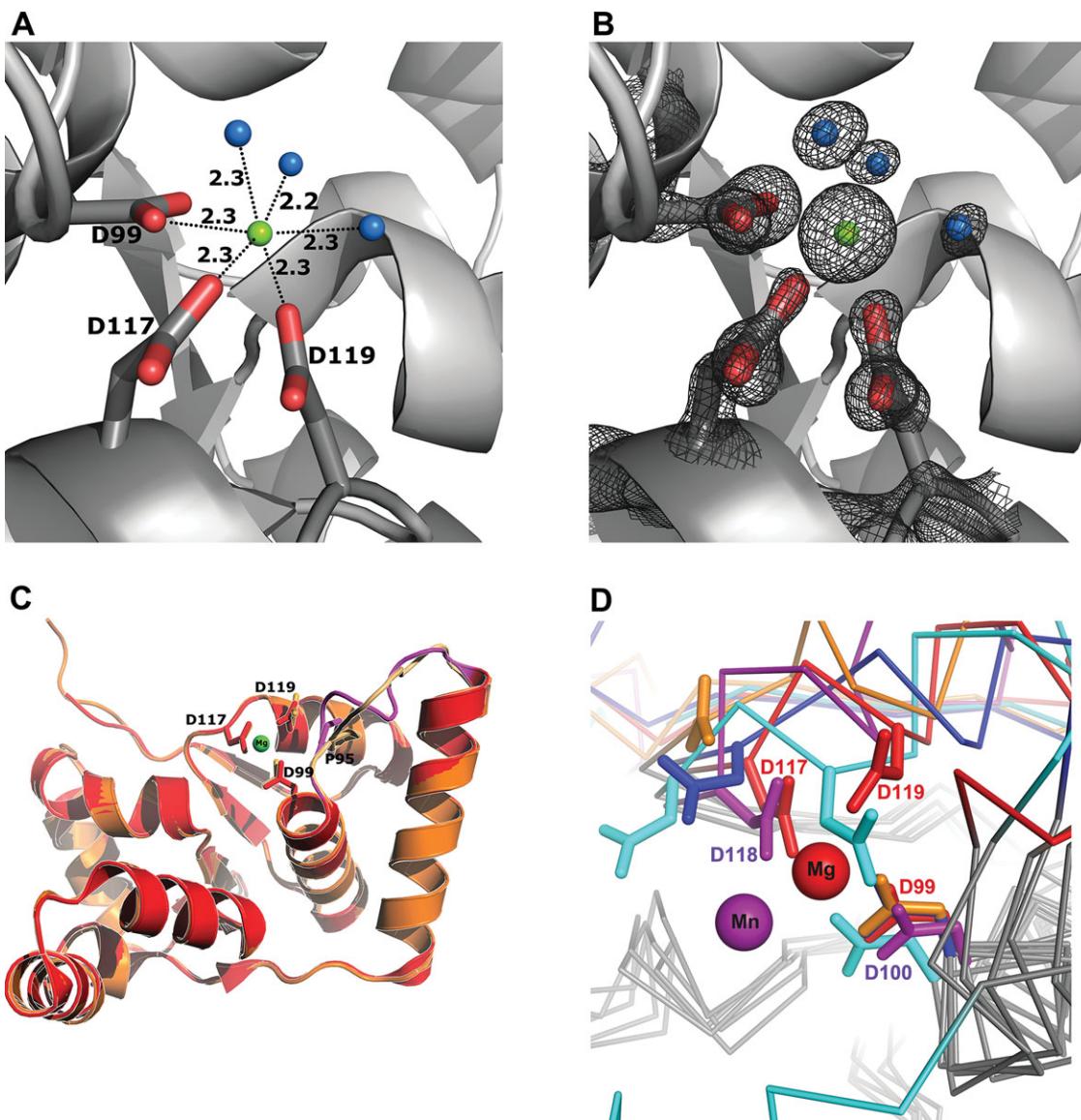
The side chain atoms of the antitoxin W72 and R73, which are the last two residues, interact with the toxin homodimer in a specific manner [Fig. 5(B,C)]. The indole moiety of the antitoxin W72 is buried inside a crevice formed by the toxin-toxin interface [Fig. 5(B)]. The crevice is located near the toxin homodimer pseudo twofold symmetry rotational axis,  $\sim$ 10 Å from each putative active site pocket. The guanidinium moiety of the antitoxin R73 sits in the pocket making close contact with all three Mg<sup>2+</sup>-coordinating aspartates [Fig. 5(C)].

The toxin homodimer contains two distinct channels with possible biological function, as discussed below. Each toxin molecule has two channels, which line up to form the longer channels in the homodimer. One channel, which has a slightly negative electrostatic potential, is a deep continuous groove occupied by the antitoxins. The other channel, which has mostly positive electrostatic potential, is a shallower groove that lies at an angle to the antitoxin groove [Fig. 5(D,E)].

### **Discussion**

#### **Rv0301 toxin is a Mg<sup>2+</sup>-dependent ribonuclease with a single metal ion at the active site**

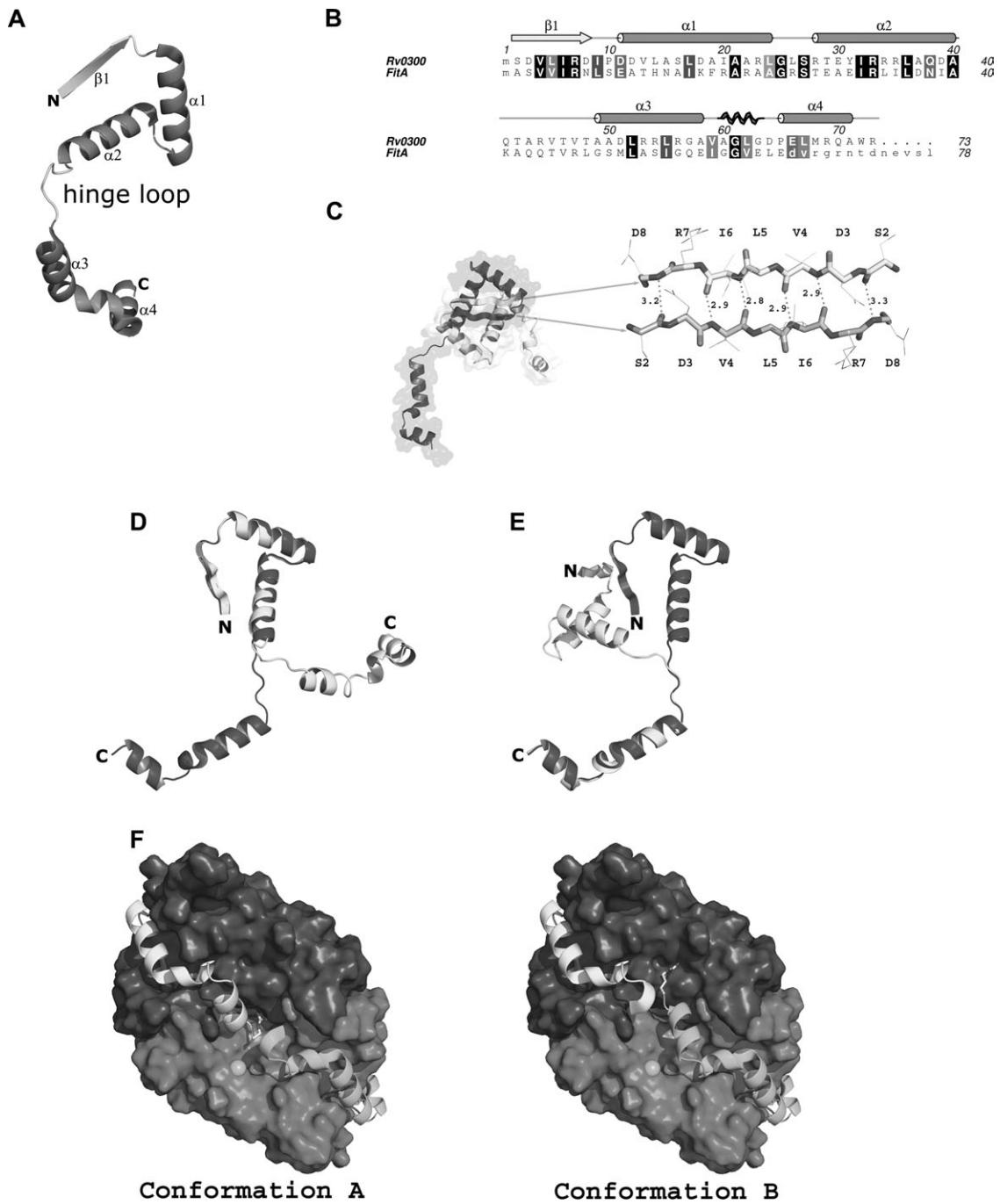
There is a growing consensus that VapC toxins are ribonucleases that cleave single-stranded RNA.



**Figure 3.** The Rv0301 toxin coordinates a  $Mg^{2+}$ -ion. A: Three aspartates (D99, D117, and D119) and three water molecules coordinate a  $Mg^{2+}$ -ion. The  $Mg^{2+}$ -ion is represented as a green sphere and the water molecules are represented as blue spheres. Distances are given in angstroms. Notice the octahedral geometry and the 2.2–2.3 Å separations. B: A  $2F_o - F_c$  composite omit map of (A) with contours at  $2\sigma$  (omit map calculated with PHENIX<sup>28</sup>). C: Mg-unbound toxin, in orange, aligned to Mg-bound toxin, in red. Among the three  $Mg^{2+}$ -coordinating aspartates, D119 is most out of position in the Mg-unbound toxin. The largest deviations in the two structures occur at the loop (R89 to S96) preceding D99, colored magenta for the Mg-bound and light orange for the Mg-unbound. In the Mg-unbound, there is a cis bond between G94 and P95. D: Ribbon representation of a structural alignment of five homologues of Rv0301; Rv0301 (3H87 chain A) is red, PH0500 (1V96 chain B) is orange, FitB (2H1C chain A) is blue, Rv0627 (3DBO chain B) is cyan, and PAE0151 (2FE1 chain A) is purple. Structurally conserved regions are shown in grey. The red sphere is the  $Mg^{2+}$ -ion coordinated by Rv0301 and the purple sphere is the  $Mn^{2+}$ -ion coordinated by PAE0151. D99, D117, and D119 are the three aspartates of Rv0301 that coordinate the  $Mg^{2+}$ -ion, and D100 and D118, which align to D99 and D117 respectively, are the two aspartates of PAE0151 that coordinate the  $Mn^{2+}$ -ion. D99 is structurally conserved among PIN-domain proteins.

Indeed, Rv0301 has been reported to exhibit moderate ribonuclease activity *in vitro*,<sup>13</sup> cleaving single-stranded viral MS2 RNA. However, the catalytic mechanism is not well understood beyond the fact that it depends on divalent metal ions. Specifically, there is an ongoing debate whether the catalytic activity of VapC toxins depend on one or two metal ions. Ours is the first structure of a VapC toxin that

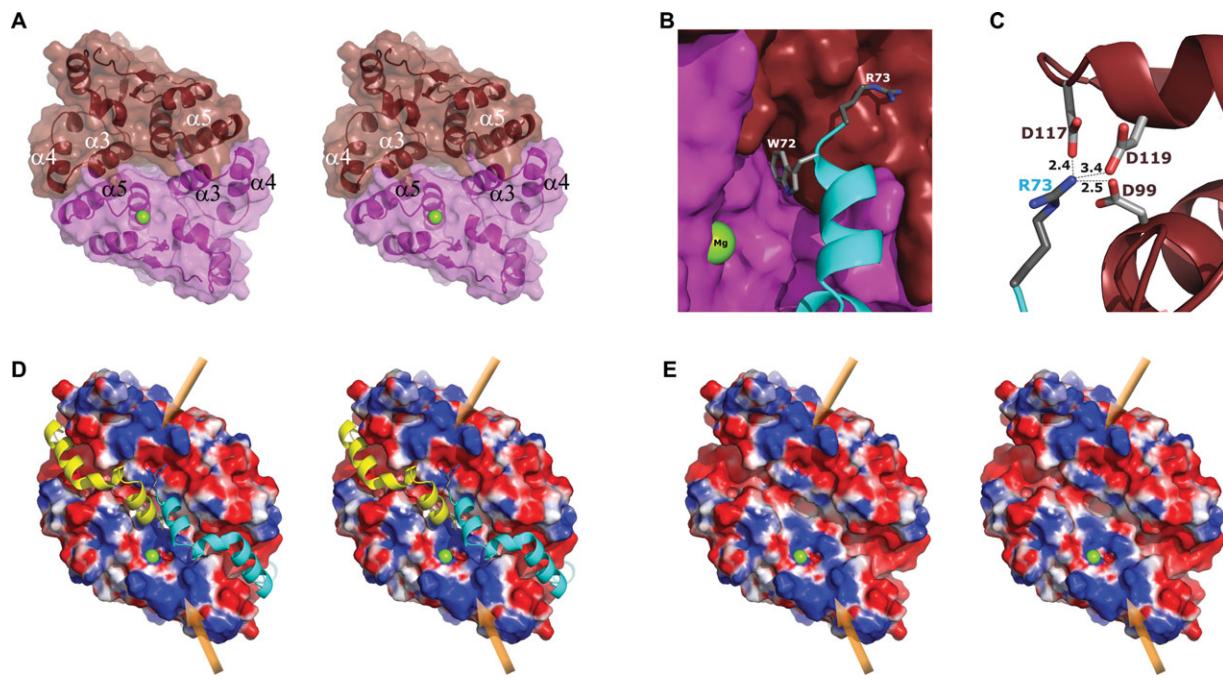
shows a coordinated  $Mg^{2+}$  ion, and our structure of Rv0301 suggests that the mechanism itself relies on a single divalent metal ion. However, the structure does not rule out the possibility that Rv0301 uses a second  $Mg^{2+}$  ion to stabilize the enzyme-substrate complex as has been suggested for other VapC toxins based on structural similarities to FEN-1 flap nucleases.<sup>12,33</sup> The side chain carboxyl carbons of D9 and



**Figure 4.** Structure of the Rv0300 VapB antitoxin. A: Cartoon diagram of the Rv0300 antitoxin shows its extended structure. B: Structure-based sequence alignment of Rv0300 and FitA. C: The Rv0300 antitoxin dimer.  $\beta$ 1- $\alpha$ 1- $\alpha$ 2 make up the antitoxin-antitoxin homodimerization interface. Two beta strands, consisting of residues S2 to D8, form an anti-parallel  $\beta$ -sheet, completing the RHH-motif. D: N-terminal superposition of the open antitoxin (light) to the closed antitoxin (dark). Residues S2 to Q41 aligned to a RMSD of 0.56 Å. E: C-terminal superposition of the two antitoxins. Residues T46-R73 aligned to a RMSD of 0.31 Å. F: Alternate conformations of the antitoxin C-terminal  $\alpha$ 4 helices. In conformation A, the light antitoxin is fully present but residues L67 to R73 of the dark antitoxin are disordered. In conformation B, the dark antitoxin is fully present but residues L67 to R73 of the light antitoxin are disordered. The Rv0301 toxin dimer is represented as a surface. The  $\alpha$ 4 helices of the two antitoxins cannot both occupy the toxin binding sites simultaneously due to steric clashes.

E43, the two conserved acidic residues of PIN domain proteins that have not been observed to bind  $Mg^{2+}$  ions, are  $\sim 7$  Å apart, and could possibly coordinate another  $Mg^{2+}$  ion with the putative RNA

substrate in place. Considering that D119, one of the  $Mg^{2+}$ -coordinating residues, is not conserved in VapC toxins, it is possible that there are different subfamilies of VapC toxins.



**Figure 5.** Proposed mechanism of Rv0301 toxin inhibition by Rv0300 antitoxin and RNA binding site. A: Stereo view of toxin dimer. The  $\alpha_5$  helices insert between the  $\alpha_3$  and  $\alpha_4$  helices. B: Close-up of last two residues of antitoxin in relation to the toxin dimer. W72 lodges in a crevice formed by the dimer interface. C: R73 of the antitoxin contacts all three  $Mg^{2+}$ -coordinating aspartates. D: Stereo view of toxin dimer surface electrostatics potential with antitoxins superimposed. The electrostatics calculations take into account the  $Mg^{2+}$  ion. E: Stereo view of toxin dimer surface electrostatics potential with orange arrows indicating our proposed RNA binding site. Compare the negative potential (red) of the deep groove occupied by the antitoxin molecules in (D) to the positive potential (blue) of the shallower groove of our proposed RNA binding site.

#### The functional unit of Rv0301 toxin is a homodimer

The functional unit of the Rv0301 toxin is most likely the homodimer. About  $1400\text{ \AA}^2$  is buried in the interface, and the sc score of 0.69 is also high, comparable to 0.64–0.68 sc values typical of antibody/antigen interactions.<sup>34</sup> The two facts taken together suggest that the dimer is stable *in vivo*. Furthermore, the dimer creates a crevice into which W72 of the Rv0300 antitoxin docks, and it forms a continuous groove that covers the putative active sites containing the  $Mg^{2+}$ -binding pockets [Fig. 5(E)]. Assuming two  $Mg^{2+}$  ions in the dimer, the two metals would be  $\sim 20\text{ \AA}$  apart, so it is likely that the two active sites would work independently.

#### The Rv0301 toxin dimer has a putative RNA-binding groove distinct to the antitoxin-binding groove

Our proposed RNA-binding site on the Rv0301 toxin dimer lies at an angle to the antitoxin-binding groove. The Rv0300 antitoxins occupy long grooves in the toxin dimer, but electrostatics analysis shows that these grooves have a net negative potential, which makes it unlikely that RNA substrates would bind there. However, there is a shallow groove at an angle to the antitoxin-groove which shows positive potential and could accommodate RNA [Fig. 5(D,E)].

The electrostatic surface shows a pocket of negative potential in the toxin lacking a  $Mg^{2+}$  ion but the presence of the divalent cation would result in positive potential, as can be seen in the corresponding pocket with the  $Mg^{2+}$ . If both toxin molecules were to coordinate single  $Mg^{2+}$  ions, our proposed RNA binding site would have a continuous positive potential, and it would place an RNA substrate in contact with both  $Mg^{2+}$  ions.

#### The hinge segment of the Rv0300 antitoxin is extended and flexible but the N-terminal DNA-binding RHH domain and the C-terminal toxin-binding domain are well defined

The crystal structure shows that the antitoxin has two flexible segments. The first is a long loop (Q41 to T48) connecting  $\alpha_2$  and  $\alpha_3$  that forms a hinge. Changes in the hinge configuration lead to the differences between the closed and open conformations. Figure 4(D,E) illustrate the conformational switches by aligning the N-termini (S2-Q41) and C-termini (T46-R73) respectively of the closed and open antitoxin chains. The RMSD of the N-termini and C-termini are  $0.56\text{ \AA}$  and  $0.31\text{ \AA}$  respectively, which reveals that the DNA-binding region and the toxin-binding regions have well-defined structures.

Rv0300 homodimers are formed by extensive interactions of the N-terminal strand-helix-helix, including an antiparallel sheet formed by the

$\beta$ -strands, which is a signature of the DNA-binding RHH motif. Rv0300 also has an R7, which is a conserved residue of the RHH motif. Several VapB proteins have been shown to be auto-inhibitors that bind near their promoter regions.<sup>24,29,33,35</sup> It is likely that this is true of Rv0300 as well, with the N-terminal RHH motif of the dimer constituting the DNA-binding region, and possibly enabling the complex to regulate other genes.

Our structure is consistent with the proposal that a single antitoxin inhibits both toxin molecules of the dimer. The two flexible loops of the C-terminal loop-helix-loop-helix of the antitoxin Rv0300 form the toxin-binding interface. Conformational changes in the first loop result in the differences between the open and closed TA interfaces. In the open TA interface, the first loop twists slightly outwards with respect to its cognate toxin, shifting it slightly out of position. In particular, the first three residues of the hinge (Q41, T42, and A43) seem to determine whether the antitoxin adopts the closed or open conformation. The second loop either brings the C-terminal helix of Rv0300 in contact with the toxin or shifts it away. In our structure, there is an overlap in the last five residues of the antitoxins, and only one can be in contact with the toxins at a time. While the possibility that the overlap is an artifact resulting from crystal dehydration (due to the microbatch method of crystallization under oil) cannot be discounted, the position of the C-terminal helix is consistent with the structure of DNA-bound VapBC2 complex from *R. felis* and the authors' proposal that one antitoxin inhibits both toxin molecules of the toxin dimer.<sup>26</sup>

#### **The Rv0301-Rv0300 complex octamer is homologous to the FitAB octamer**

The overall octamer topology of the Rv0301-Rv0300 complex is similar to the *N. gonorrhoeae* FitAB,<sup>9</sup> *S. flexneri* VapBC,<sup>30</sup> and *R. felis* VapBC2<sup>26</sup> complexes. The structures of FitAB and *R. felis* VapBC2 were solved bound to DNA. Our Rv0301-Rv0300 octamer is closed and compact like the *S. flexneri* and *R. felis* octamers, compared to the open FitAB octamer. However, the Rv0300 and FitA antitoxins both have RHH DNA-binding domains, while the *S. flexneri* VapB and *R. felis* VapB2 have related AbrB-like DNA-binding domains which are characterized by swapped  $\beta$ -hairpins. The authors of the *R. felis* VapBC2 structure have stated that the complex forms a (VapB2)<sub>2</sub>(BapC2)<sub>4</sub> hexamer when unbound to DNA based on size exclusion chromatography and multiangle light scattering.<sup>26</sup> However, structures of FitAB, *S. flexneri* VapBC, and the Rv0301-Rv0300 complex have been observed only as octamers bound or unbound to DNA, so these most likely exist as stable octamers *in vivo*.

DALI<sup>31</sup> results show that Rv0301 (3H87 chain A) has a RMSD of 2.3 Å aligning 123 out of 139 residues to FitB (2H1C chain A), and that Rv0300 (3H87 chain C) has a RMSD of 5.5 Å aligning 63 out of 64 residues to FitA (2H1O chain F). FitA binds with higher affinity to its target DNA when complexed with FitB, and the structure of FitAB complexed to DNA reveals that the octamer stabilizes the two resulting RHH motifs in position to bind the inverted repeats in the promoter region.<sup>9,24</sup> The structural homology to FitAB suggests that the biological role of the Rv0301-Rv0300 octamer is similar, although the distance separating the two anti-parallel  $\beta$ -sheets is ~60 Å in the FitAB octamer and only ~30 Å in the Rv0301-Rv0300 octamer. The FitAB octamer binds an inverted repeat with 12 base pairs separating the repeats. The 30 Å shorter distance in Rv0301-Rv0300 suggests that either the RHH ribbons insert into repeats separated only by one or two base pairs or that there is a conformational change when bound to DNA that opens up the Rv0301-Rv0300 octamer. The structure of the *R. felis* VapBC2 bound to DNA suggests the former possibility. In addition to enhanced DNA-binding affinity, the tightly closed topology of the octamer may serve to protect the antitoxins from protease degradation.

#### **Two factors contribute to the inhibition of Rv0301 toxin by Rv0300 antitoxin**

The Rv0300 antitoxin inhibits the Rv0301 toxin by creating steric hindrance for potential RNA targets and by competing away the necessary Mg<sup>2+</sup> ions. In the discussion above of the Rv0301 toxin, we show that electrostatic calculations make it unlikely that putative RNA substrates bind in the groove where the Rv0300 antitoxin binds. The theoretical pI of Rv0300, as calculated by ProtParam<sup>36</sup> is 10.4, which is consistent with the electrostatics results. However, the presence of Rv0300 would hinder access by RNA molecules to our proposed RNA-binding groove. Furthermore, the C-terminal helix of the antitoxin partially blocks the proposed RNA-binding groove, and may completely block access to the active sites. In addition, when comparing the Mg-bound toxin molecule to the Mg-unbound toxin molecule, there is a trans to cis transition in the G109-P110 peptide bond, which shifts D117 and D119 out of position from coordinating a Mg<sup>2+</sup> ion. Finally, our structure indicates that the last two residues of the antitoxin, W72 and R73, are critical for toxin inhibition [Fig. 5(B,C)]. W72 is often a purine mimic and could compete directly with an adenine or a guanosine. The guanidinium moiety of R73, the last residue of Rv0300, fits snugly in the Mg<sup>2+</sup>-binding pocket. Thus, tight binding of the antitoxin would remove the bound Mg<sup>2+</sup> ion. The role of C-terminal arginines may be a common mechanism of VapB antitoxins as the arginines of *N. gonorrhoeae* FitAB<sup>9</sup> and

Mtb VapBC-5<sup>12</sup> have also been observed to interact with putative active sites.

### Possible biological role of Rv0301-Rv0300

We propose that Rv0301 targets certain RNAs preferentially to regulate cell growth rates and plays a role in Mtb persistence. Ramage *et al.* have shown that expression of Rv0301 in *Mycobacterium smegmatis* inhibits translation which leads to retarded cell growth. They have further shown that Rv0301 cleaves viral MS2 RNA *in vitro*, although with less activity compared to the other toxins they tested.<sup>13</sup> It is reasonable to conclude that the Rv0301 toxin is a ribonuclease that slows cell growth by inhibiting translation via mRNA degradation. The *N. gonorrhoeae* FitAB system may provide clues by analogy. FitB was first discovered through genetic screens to slow intracellular trafficking. Mattison *et al.* have proposed that the FitAB complex is bound to its promoter in an extracellular environment, but is released when the cell enters epithelial cells, which allows transcription and accumulation of FitAB, and ultimately results in free FitB. The FitB slows growth so that *N. gonorrhoeae* cells remain inside the host cells.<sup>9</sup> A similar function may be provided by the Rv0301-Rv0300 pair in Mtb that helps regulate the cell's transition to and from a persistent state depending on whether it resides inside macrophages. Mattison *et al.* were unable to detect any ribonuclease activity by FitB in *in vitro* studies. They argued that FitB may target a specific sequence. This is a plausible hypothesis that could explain the relatively weak activity of Rv0301, and provide an explanation for why the Mtb genome carries so many VapBC loci. The position of the antitoxin W72 provides a shred of circumstantial evidence that putative target RNA molecules must have a purine at a specific location.

Since only the C-terminal helix of the Rv0300 antitoxin directly occludes our proposed RNA-binding site, an Rv0300 mutant with the last helix truncated should bind the Rv0301 toxin with high affinity and reduced inhibitory effect, and a peptide that mimics the last helix could effectively inhibit toxin activity. Either strategy could potentially aid in the design of new drugs; the former to compete with the antitoxin in the hopes that the toxin would kill its cell, and the latter to inhibit toxin function in the hopes that it would hinder the cell from entering a persistent state.

### Conclusions

The Rv0301 VapC toxin forms a homodimer which functions as a ribonuclease. Each toxin molecule coordinates a single Mg<sup>2+</sup> ion in its active site, and a shallow groove, which connects the two active sites, forming the probable RNA-binding site. The Rv0300 VapB antitoxin does not bind toxin in the same site as RNA. The C-terminal helix of the antitoxin obstructs the RNA-binding groove and blocks access to the

active sites. Additionally, the C-terminal arginines of the antitoxins remove Mg<sup>2+</sup> ions from the active sites.

### Accession Number

The atomic coordinates and structure factors of the *Mycobacterium tuberculosis* Rv0301-Rv0300 complex have been deposited in the Protein Data Bank with accession code 3H87.

### Materials and Methods

#### Cloning of the Rv0301-Rv0300 complex

Rv0301 and Rv0300 genes were PCR-amplified from *M. tuberculosis* H37Rv genomic DNA using Sure-Pol DNA polymerase (Denville Scientific) and the following primers: Rv0301Fwd (5'-GACGACGACAA-GATGGTGACTGACCAGCGCTGGCTGATCGACAAGTCGG), Rv0301Rev (5'-CGCGGGCGGCCGTATTAAAGCGGAAGGCGGGCGATCGTGAGCCG), Rv0300Fwd (5'-GCGGGCCCGGCCTTGATGAGTGATGTACTGATTCGGGACATCCCCGACGACGTG), and Rv0300Rev (5'-GAGGAGAAGCCCGTATTACCTCACGCCTGACGCATAAGCTCGGGATCG).

The PCR products were treated with T4 polymerase and dATP according to the protocol provided with the pET-46 Ek/LIC Vector Kit (Novagen). Following T4 treatment, the products were incubated with the LIC DUET minimal adaptor and Ek/LIC vector according to the LIC DUET Adaptor kit protocol (Novagen). The primers were designed so that Rv0301 would be in the upstream open reading frame and Rv0300 in the downstream open reading frame of the final vector. The annealed vector was transformed into *E. coli* NovaBlue GigaSingles Competent Cells (Novagen), which were then plated on LB-agar supplemented with 100 µg/mL of ampicillin. The LIC DUET adaptor introduced an N-terminal MQAGPAL sequence to the Rv0300 gene, and this sequence was removed with the QuikChange mutagenesis kit (Agilent Technologies) using the following primers: Rv0300FixFwd (5'-aacttaagaaggagatataccatgagtgtatgtactcg), Rv0300FixRev (5'-ccgaatcgtacatcaatcatggtatatccctttaagg). The final construct had two tandem open reading frames, in which the upstream encoded Rv0301 with an N-terminal MAHHHHHVDDDK affinity tag and enterokinase site, and the downstream encoded Rv0300 unmodified. All vectors were purified using the Qiagen MiniPrep Kit (Qiagen), and the sequences were confirmed by DNA sequencing (UCLA Web-Seq; <http://genoseq.ucla.edu/webseq/>).

#### Protein expression and purification

The Rv0301-Rv0300 expression plasmid was transformed into *E. coli* BL21 Gold (DE3) (Agilent Technologies), and a colony was picked to grow overnight cultures at 37°C in LB broth supplemented with 100 µg/mL ampicillin. The following day, 6 L of fresh media

was inoculated with 10 mL of overnight culture per liter and grown at 37°C until an OD<sub>600nm</sub> of 0.6 was reached. At this point protein expression was induced with IPTG to a final concentration of 1.0 mM and the cultures were grown for an additional 16 hours at 25°C. The cells were harvested by centrifugation and the cell pellet was resuspended in 60 mL lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl, 25 mM imidazole pH 8.0, 10 mM β-mercaptoethanol). Protease inhibitor cocktail (Sigma-Aldrich), DNase I (NEB), and RNaseA (Sigma-Aldrich) were added. Cells were lysed by sonication and the lysate was then centrifuged at 29,800g for 30 minutes at 4°C. The supernatant was then passed through a 0.22 μm filter and loaded on to a 5 mL HisTrap Column (GE Healthcare) pre-equilibrated with lysis buffer. The column was then washed with 50 mL of wash buffer (50 mM Tris pH 8.0, 500 mM NaCl, 50 mM imidazole pH 8.0, 10 mM β-mercaptoethanol), and eluted in a step gradient. The complex (as determined by SDS PAGE) coeluted in three peaks with the largest peak eluting at 150 mM imidazole. Following concentration of this fraction, TCEP (1 M stock solution in 50 mM Tris; NaOH was added to a final pH of 7.0) was added to a final concentration of 10 mM. The final concentration of the protein solution used for crystallization was ~40 mg/mL (in 50 mM Tris, 500 mM NaCl, 150 mM imidazole, 10 mM TCEP, pH 8.0).

The selenomethionine (SeMet) derivative of the Rv0301-Rv0300 complex was prepared by growing an overnight culture of *E. coli* BL21 Gold (DE3) (Agilent Technologies) in M9 minimal media supplemented with 100 μg/mL ampicillin. The overnight culture was used to inoculate 6 L of fresh M9 media supplemented with 100 μg/mL ampicillin. The cells were grown at 37°C with shaking. Upon reaching an OD<sub>600nm</sub> of 0.5, 50 mg leucine, 50 mg isoleucine, 50 mg valine, 100 mg phenylalanine, 100 mg lysine, 100 mg threonine, and 60 mg selenomethionine were added per liter. After 15 minutes, the cultures were induced with IPTG to a final concentration of 1.0 mM and grown at 25°C for 18 hours. Purification of the SeMet derivative followed the same protocol as for the native complex, and eluted at 120 mM imidazole. Following addition of TCEP to 10 mM, the final concentration of the SeMet derivative protein solution was ~30 mg/mL (in 50 mM Tris, 500 mM NaCl, 150 mM imidazole, 10 mM TCEP, pH 8.0).

#### **Crystallization of the Rv0301-Rv0300 complex**

The initial crystallization screens were carried out with the complex containing the MQAGPAL artifact. Multiple attempts resulted in poorly diffracting crystals with low reproducibility. Cleaving the affinity tag followed by additional purification steps including size exclusion did little to improve the crystals. New crystallization trials using the complex with MQAGPAL removed immediately resulted in well diffracting

crystals. The best native crystals were formed under the microbatch method as follows: 60 μL of protein solution was mixed with 30 μL of precipitant solution (200 mM potassium acetate, pH 7.0; 7.5% PEG-3350). The mixed solution was then filtered through a 0.1 μm filter. The filtered protein mix was aliquoted in 3 μL drops under 50 μL of paraffin oil. Large single crystals formed in most of the wells and the largest crystals grew to ~500 μm in 60 hours.

SeMet crystals were formed as follows: 60 μL of protein solution was mixed with 30 μL of precipitant solution (200 mM potassium acetate, pH 6.5; 4.5% PEG-3350). The mixed solution was filtered through a 0.1 μm filter and 3 μL drops were aliquoted under 50 μL of paraffin oil. After 48 hours, most drops contained one or two crystals with the largest crystals reaching ~300 μm in size.

For cryoprotection, both native and SeMet crystals were soaked sequentially in precipitant solution with 5%, 15%, and 25% glycerol for 1 minute each. The crystals were mounted and flash frozen in liquid nitrogen. Data were collected on the 24-ID-C beamline at the Advanced Photon Source of Argonne National Lab. Data from SeMet crystals were used to determine P4<sub>1</sub>2<sub>1</sub>2 symmetry. For the native crystals, in order to maximize high resolution data, data was collected at 10% transmission for 104 images and 1% transmission for 90 images.

#### **Rv0301-Rv0300 structure determination**

The data for both native and SeMet crystals were processed with Denzo and Scalepack.<sup>37</sup> Data for phasing was obtained by single-wavelength anomalous dispersion (SAD) at 0.9795 Å. The program HKL2MAP (SHELX C, D, and E)<sup>38</sup> was used to calculate phases to 1.85 Å by first determining a substructure of selenium atoms and its correct enantiomorph. DM<sup>39</sup> was used to improve the experimental phases through density modification and Arp/wARP<sup>40</sup> was used for automated model building. Refmac5<sup>41</sup> and Coot<sup>42</sup> were used for iterative rounds of model refinement and manual extensions. For higher resolution data, the native dataset was used with the phases determined from the SeMet dataset. The native crystals had slightly larger unit cell dimensions, so a rigid body fit was used for the first round of refinement. Refmac5, Coot, and Phenix<sup>28</sup> were used for additional rounds of refinement, and the final rounds of refinement used TLS parameters<sup>43</sup> determined through the TLSmd server (<http://skuld.bmsc.washington.edu/~tlsmd/>). The final structure was validated with PROCHECK,<sup>44</sup> ERRAT,<sup>45</sup> and Verify3D<sup>46</sup> through the NIH-MBI SAVS server (<http://nihserver.mbi.ucla.edu/SAVS/>). PISA<sup>47</sup> and SC<sup>32</sup> were used to analyze subunit interfaces, EZProt<sup>48</sup> was used to calculate electrostatic surface potentials, and DALI<sup>31</sup> was used for structural comparisons. All graphic representations of molecular structures were

prepared with PyMol (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC).

### Acknowledgments

The authors thank Janet Chiang for the initial construct, and Dr. Mark Arbing and Dr. Arthur Laganowsky for many helpful discussions. They also thank the UCLA-DOE X-ray Crystallography Core Facility, which is supported by DOE Grant DE-FC02-02ER63421; M. Capel, K. Rajashankar, N. Sukumar, J. Schuermann, I. Kourinov and F. Murphy at NECAT beamlines 24-ID at APS, which are supported by grants from the National Center for Research Resources (5P41RR015301-10), and the National Institute of General Medical Sciences (8 P41 GM103403-10) from the National Institutes of Health. Use of the APS is supported by DOE under Contract DE-AC02-06CH11357.

### References

1. Gerdes K (2000) Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J Bacteriol* 182:561–572.
2. Pandey DP, Gerdes K (2005) Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* 33:966–976.
3. Loris R, Dao-Thi MH, Bahassi EM, Van Melderen L, Poortmans F, Liddington R, Couturier M, Wyns M (1999) Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *J Mol Biol* 285:1667–1677.
4. Hargreaves D, Santos-Sierra S, Giraldo R, Sabariego-Jareño R, de la Cueva-Méndez G, Boelens R, Díaz-Orejas R, Rafferty JB (2002) Structural and functional analysis of the kid toxin protein from *E. coli* plasmid R1. *Structure* 10:1425–1433.
5. Kamada K, Hanaoka F, Burley SK (2003) Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Mol Cell* 11:875–884.
6. Meinhart A, Alonso JC, Sträter N, Saenger W (2003) Crystal structure of the plasmid maintenance system epsilon/zeta: functional mechanism of toxin zeta and inactivation by epsilon 2 zeta 2 complex formation. *Proc Natl Acad Sci U S A* 100:1661–1666.
7. Kamada K, Hanaoka F (2005) Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Mol Cell* 19:497–509.
8. Takagi H, Kakuta Y, Okada T, Yao M, Tanaka I, Kimura M (2005) Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nat Struct Mol Biol* 12:327–331.
9. Mattison K, Wilbur JS, So M, Brennan RG (2006) Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA reveals a tetramer of toxin-antitoxin heterodimers containing pin domains and ribbon-helix-helix motifs. *J Biol Chem* 281:37942–37951.
10. Oberer M, Zanger K, Gruber K, 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.
11. Kumar P, Issac B, Dodson EJ, Turkenburg JP, Mande SC (2008) Crystal structure of *Mycobacterium tuberculosis* YefM antitoxin reveals that it is not an intrinsically unstructured protein. *J Mol Biol* 383:482–493.
12. Miallau L, Faller M, Chiang J, Arbing M, Guo F, Cascio D, Eisenberg D (2009) Structure and proposed activity of a member of the VapBC family of toxin-antitoxin systems. VapBC-5 from *Mycobacterium tuberculosis*. *J Biol Chem* 284:276–283.
13. Ramage HR, Connolly LE, Cox JS (2009) Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet* 5: e1000767.
14. Zhang Y, Zhu L, Zhang J, Inouye M (2005) Characterization of ChpBK, an mRNA interferase from *Escherichia coli*. *J Biol Chem* 280:26080–26088.
15. Zhu L, Zhang Y, Teh JS, Zhang J, Connell N, Rubin H, Inouye M (2006) Characterization of mRNA interferases from *Mycobacterium tuberculosis*. *J Biol Chem* 281:18638–18643.
16. Overgaard M, Borch J, Jørgensen MG, Gerdes K (2008) Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Mol Microbiol* 69: 841–857.
17. Christensen-Dalsgaard M, Gerdes K (2008) Translation affects YoeB and MazF messenger RNA interferase activities by different mechanisms. *Nucleic Acids Res* 36:6472–6481.
18. Christensen-Dalsgaard M, Jørgensen MG, Gerdes K (2010) Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol* 75:333–348.
19. Jiang Y, Pogliano J, Helinski DR, Konieczny I (2002) ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Mol Microbiol* 44:971–979.
20. Arcus VL, Bäckbro K, Roos A, Daniel EL, Baker EN (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* 279:16471–16478.
21. Daines DA, Wu MH, Yuan SY (2007) VapC-1 of non-typeable *Haemophilus influenzae* is a ribonuclease. *J Bacteriol* 189:5041–5048.
22. Audoly G, Vincentelli R, Edouard S, Georgiades K, Mediannikov O, Gimenez G, Socolovschi C, Mège JL, Cambillau C, Raoult D (2011) Effect of rickettsial toxin VapC on its eukaryotic host. *PLoS One* 6:e26528.
23. Hopper S, Wilbur JS, Vasquez BL, Larson J, Clary S, Mehr IJ, Seifert HS, So M (2000) Isolation of *Neisseria gonorrhoeae* mutants that show enhanced trafficking across polarized T84 epithelial monolayers. *Infect Immun* 68:896–905.
24. Wilbur JS, Chivers PT, Mattison K, Potter L, Brennan RG, So M (2005) *Neisseria gonorrhoeae* FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif. *Biochemistry* 44:12515–12524.
25. Dienemann C, Bøggild A, Winther KS, Gerdes K, Brodersen DE (2011) Crystal structure of the VapBC toxin-antitoxin complex from *Shigella flexneri* reveals a hetero-octameric DNA-binding assembly. *J Mol Biol* 414:713–722.
26. Maté MJ, Vincentelli R, Foos N, Raoult D, Cambillau C, Ortiz-Lombardía M (2012) Crystal structure of the DNA-bound VapBC2 antitoxin/toxin pair from *Rickettsia felis*. *Nucleic Acids* 40:3245–3258.
27. Bond CS, Schüttelkopf AW (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Cryst D* 65:510–512.
28. Adams PD, Afonine PV, Bunkóczki G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS,

- Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* 66:213–221.
29. Bunker RD, McKenzie JL, Baker EN, Arcus VL (2008) Crystal structure of PAE0151 from *Pyrobaculum aerophilum*, a PIN-domain (VapC) protein from a toxin-antitoxin operon. *Proteins* 72:510–518.
  30. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A (2010) The Pfam protein families database. *Nucleic Acids Res* 38:D211–222.
  31. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38:W545–549.
  32. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234:946–950.
  33. Dienemann C, Bøggild A, Winther KS, Gerdes K, Brodersen DE (2011) Crystal structure of the VapBC toxin-antitoxin complex from *Shigella flexneri* reveals a heterooctameric DNA-binding assembly. *J Mol Biol* 414:713–722.
  34. Epa VC, Colman PM (2001) Shape and electrostatic complementarity at viral antigen-antibody complexes. *Curr Top Microbiol Immunol* 260:45–53.
  35. Robson J, McKenzie JL, Cursons R, Cook GM, Arcus VL (2009) The vapBC operon from *Mycobacterium smegmatis* is an autoregulated toxin-antitoxin module that controls growth via inhibition of translation. *J Mol Biol* 390:353–367.
  36. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 112:531–552.
  37. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
  38. Pape T, Schneider TR (2008) Hkl2map: a graphical user interface for phasing with shelx program. *J Appl Cryst* 37:843–844.
  39. Cowtan K (1994) Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31:34–38.
  40. Perrakis A, Sixma TK, Wilson KS, Lamzin VS (1997) wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. *Acta Cryst D* 53:448–455.
  41. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Cryst D* 60:2184–2195.
  42. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst D* 66:486–501.
  43. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Cryst* 39:109–111.
  44. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK—a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291.
  45. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519.
  46. Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85.
  47. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
  48. Petit FK, Bowie JU (2007) Ezprot class library for analysis of protein sequence and structure. <http://www.doe-mbi.ucla.edu/local/software/ezprot>.