The FEBS Journal

FEBS PRESS science publishing by scientists

Structural characterization of VapB46 antitoxin from *Mycobacterium tuberculosis*: insights into VapB46–DNA binding

Madhurima Roy[1], Anirban Kundu[1], Anirban Bhunia[2], Sujoy Das Gupta[3], Soumya De[4] and Amit Kumar Das[1,4]

1 Department of Biotechnology, Indian Institute of Technology Kharagpur, India
2 Department of Biophysics, Bose Institute, Kolkata, India
3 Department of Microbiology, Bose Institute, Kolkata, India
4 School of Bioscience, Indian Institute of Technology Kharagpur, India

Keywords
NMR spectroscopy; Phd/YefM; protein–DNA interaction; toxin–antitoxin; virulence-associated protein

Correspondence
A. K. Das, Department of Biotechnology & School of Bioscience, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
Tel: +91 3322 283756
E-mail: amitk@bt.iitkgp.ac.in
and
S. De, School of Bioscience, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
Tel: +91 3322 260514
E-mail: somde@iitkgp.ac.in

(Received 17 June 2018, revised 24 October 2018, accepted 17 December 2018)

doi:10.1111/febs.14737

The ability to form persister cells by *Mycobacterium tuberculosis* (Mtb) is a prime cause for the emergence of drug-resistant strains. A large number of toxin–antitoxin systems in the Mtb genome are postulated to promote bacterial persistence. The largest family of toxin–antitoxin systems encoded in the genome of Mtb is VapBC, with 47 VapBC toxin–antitoxin systems regulated by VapB antitoxins. In this study, we characterized the structure of VapB46 antitoxin and determined its interaction with its cognate DNA sequence. Using electrophoretic mobility shift assay and DNase I footprinting we showed that VapB46 binds to two sites in the upstream promoter–operator region. Using nuclear magnetic resonance (NMR)-based structural studies we found that VapB46 has a well-folded dimeric N-terminal domain, which contains a Phd/YefM motif and is involved in DNA binding. The remaining C-terminal residues are disordered but promote higher order oligomerization of VapB46. We propose a DNA-binding model in which tetrameric VapB46 binds to the two sites in its promoter–operator region, with each site bound by its dimeric N-terminal domain.

Introduction

Toxin–antitoxin (TA) systems are bacterial stress response systems which are present on both bacterial chromosomes and plasmids. TA systems participate in important functions such as plasmid stabilization [1,2], altruistic suicide [3,4], formation of persister cells [5,6], and are also involved in bacteriostasis and emergence of antibiotic-tolerant populations [7,8]. Based on the antitoxin, TA systems are classified into five different types, I–V [9–12]. Type II TA systems are the most abundant and well characterized, where both toxin and antitoxin are proteins and form a tight complex. In this TA system, the genes are arranged in a bicistronic operon, where the toxin gene is preceded by the antitoxin gene and prevents toxicity through

Abbreviations
CD, circular dichroism; dsDNA, double-stranded DNA; EMSA, electrophoretic mobility shift assay; HSQC, heteronuclear single quantum coherence; Mtb, *Mycobacterium tuberculosis*; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; SEC, size exclusion chromatography; TA, toxin–antitoxin; Vap, virulence associated protein.

translational coupling. The antitoxin also represses the transcription of the TA operon, which is further modulated by the toxin [13–15]. The majority of type II toxins target essential cellular mechanisms such as DNA replication and protein synthesis, so that slowing down these processes enables the bacteria to survive under various stress conditions [10,16–19].

Mycobacterium tuberculosis (Mtb), one of the most harmful pathogens, has 88 putative type II TA systems in its genome. Type II TA systems have eight major families: MazEF, HigBA, VapBC, ParDE, RelBE, Phd/Doc, CcdAB and HipAB [9]. Among these, the VapBC TA family is the largest [20]. VapC toxins belong to the PIN domain family of proteins [21] and exhibit RNase activity [22–25]. VapB antitoxins have an N-terminal DNA binding domain that belongs to the AbrB, helix–turn–helix, ribbon–helix–helix or PhD/YefM family [11]. Structures of VapBC TA systems provide important insight into the VapB–VapC interaction and how VapB neutralizes the toxicity of VapC [26–30].

Antitoxins alone or in complex with toxins repress the transcription of the TA operon. Therefore, the mechanism of DNA binding by the antitoxin is important for understanding the regulation of the TA systems [31]. Mycobacterial antitoxin ParD2 has been shown to regulate the ParDE2 TA operon by interacting with the promoter–operator region [32]. HigA antitoxin from the HigBA TA operon binds to the P2 promoter region and represses the expression of *higA* and *higB* genes [13]. MazE6 binds to a 99 bp sequence upstream of its start codon and autoregulates the expression of the MazEF6 TA operon [33]. Recently, the antitoxin VapB26 has been shown to bind a 19 bp palindrome-like DNA sequence in the promoter region that regulates the VapBC26 TA operon [30].

VapBC46 has been identified as an important TA locus in Mtb that is activated during cellular stress [34]. VapB46 has been shown to act as an antitoxin by its ability to alleviate the toxic activity of its cognate toxin, VapC46, when coexpressed in *Mycobacterium smegmatis* [20]. The N-terminal region of VapB46 is predicted to belong to the Phd/YefM antitoxin family, consistent with the hypothesis that it also possesses DNA binding ability [14,15]. In order to gain insight into the DNA binding mechanism of VapB46 antitoxin, we have structurally characterized VapB46 antitoxin and its interaction with the upstream promoter–operator region. Electrophoretic mobility shift assay (EMSA) and the DNase I footprinting method showed that VapB46 binds to the 148 bp upstream DNA sequence and occupies two binding sites in the promoter–operator region. Our solution NMR spectroscopy-based structural studies revealed that VapB46 has a well-folded N-terminal domain that adopts a Phd/YefM fold, while the C-terminal domain is unstructured. A chemical shift perturbation study identified residues in the N-terminal domain of VapB46 that are important for interaction with the promoter–operator DNA sequence. Thus, our findings help in understanding the DNA binding mechanism by VapB46 antitoxin and how it regulates the VapBC46 TA operon.

Results

Structural characterization of VapB46 by solution NMR and circular dichroism spectroscopy

In order to characterize the Mtb VapB46 structure at atomic resolution, we used solution NMR spectroscopy. Two-dimensional ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectra of ^{15}N-labeled VapB46 consisted of well-dispersed peaks, indicating the presence of a folded domain. Also, a few strong signals were observed in the center of the spectra (7.5–8.5 ppm on the proton scale), which are characteristic of an unfolded protein sequence (Fig. 1A). Based on sequence analysis using INTERPRO [35] and sequence alignment with Phd and YefM domains (Fig. 1D), we predicted that residues 1–54 should contain a folded domain, and hence a truncated VapB46$^{1–54}$ construct was designed (Table 1). The ^{1}H–^{15}N HSQC spectrum of VapB46$^{1–54}$ was well dispersed and sharp peaks from unfolded regions were absent (Fig. 1B). Overlay of the ^{1}H–^{15}N HSQC spectra of full-length VapB46 and VapB46$^{1–54}$ shows near-perfect overlaps of the well-dispersed peaks from the structured regions and the sharp peaks at the center missing in the VapB46$^{1–54}$ spectra (Fig. 1C). This indicates that VapB46 consists of an ~ 50-residue N-terminal folded domain and a C-terminal unstructured region, and the truncation of VapB46 did not alter the structure of its N-terminal domain.

To further confirm the domain architecture of VapB46, circular dichroism (CD) experiments were performed on both full-length VapB46 and truncated VapB46$^{1–54}$. CD spectra showed that both VapB46 and VapB46$^{1–54}$ are composed of α-helices and β-strands (Fig. 1E). Deconvolution of the CD spectra was done by SELCON3 [36] and the resulting percentage of secondary structures are listed in Table 2. In VapB46, a higher proportion of residues (57%) lack any secondary structure (helix or strand) compared to the smaller construct VapB46$^{1–54}$ (42%), which is consistent with the NMR studies predicting a disordered

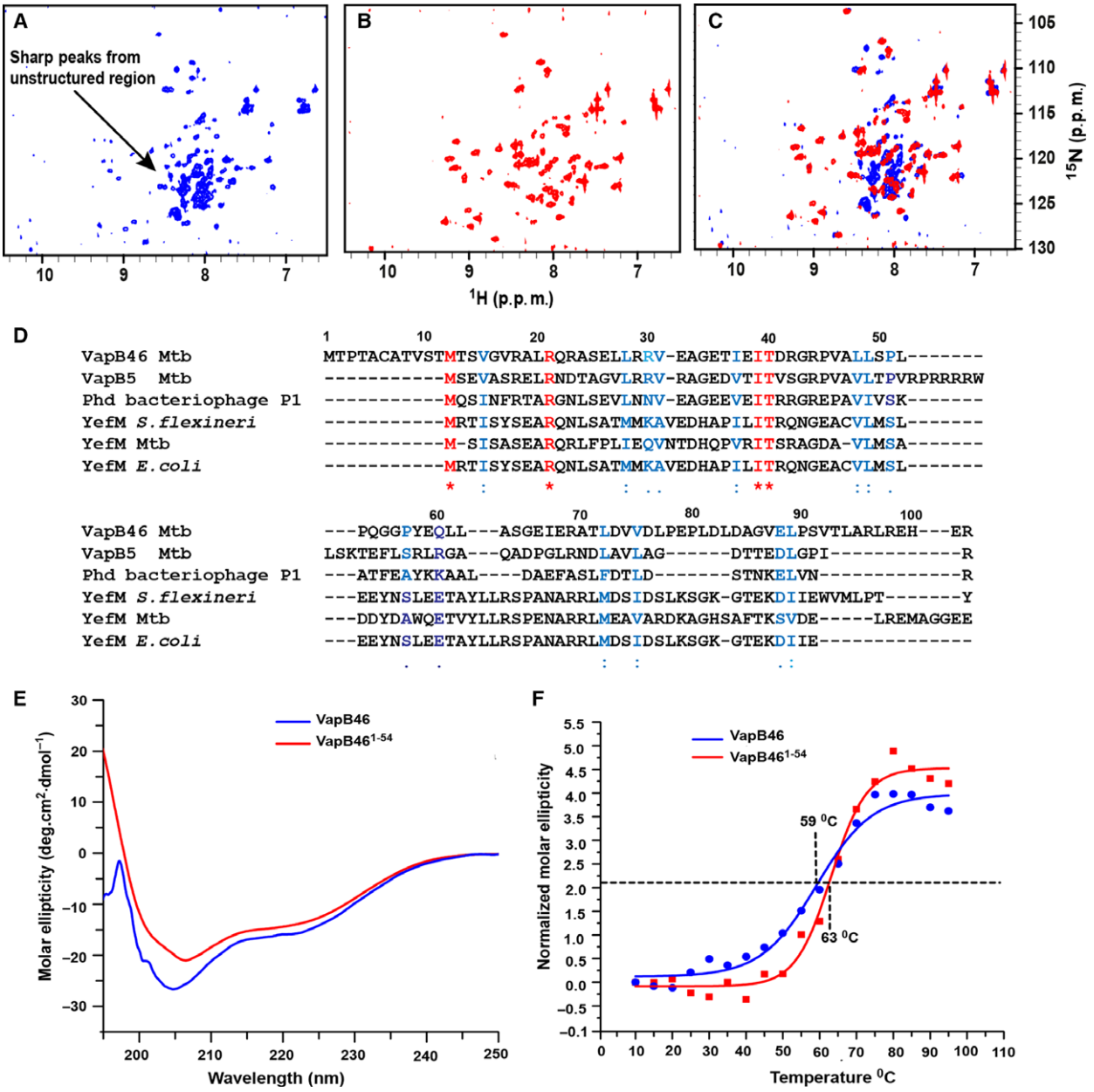

Fig. 1. Determination of folded domain in VapB46. (A) ^{1}H–^{15}N HSQC spectra of VapB46 (blue) has well dispersed peaks indicating the presence of a folded domain. Sharp peaks (black arrow) originating from the unstructured region appear at the center. (B) The sharp peaks are absent while the dispersed peaks are present in the ^{1}H–^{15}N HSQC spectra of VapB46$^{1-54}$ (red). (C) Overlay of ^{1}H–^{15}N HSQC spectra of VapB46 and VapB46$^{1-54}$ shows that the folded domain is structurally similar in both constructs. (D) Multiple sequence alignment of VapB46 with proteins containing Phd and YefM DNA binding domain. Residues shown in red are identical and conserved while similar and moderately conserved residues are shown in blue. (E) Far UV CD spectrum of VapB46 and VapB46$^{1-54}$ are superimposed. Secondary structure compositions are shown in Table 2. (F) Thermal denaturation curves of full-length VapB46 and VapB46$^{1-54}$ monitored at 222 nm from 10 to 95 °C show that the isolated N-terminal domain is marginally more stable compared to the full-length protein.

region towards the C terminus. To determine the effect of the disordered region on the stability of the protein, thermal stability studies were performed by heat denaturation of VapB46 and VapB46$^{1-54}$ from 10 °C to 95 °C and monitored at 222 nm. VapB46 has a melting temperature (T_m) of 59 °C, while VapB46$^{1-54}$ has a T_m of 63 °C (Fig. 1F). Thus, the disordered C-terminal region in VapB46 marginally lowers its stability.

Table 1. VapB46 constructs and DNA sequences

Name	Sequence	Comment
VapB46	Met1–Arg102	Full-length protein
VapB46$^{1-54}$	Met1–Gln54	Truncated protein
DNA site 1	5′-CCAGCTCAGC-3′	NMR titration
DNA site 2	5′-AACGCGCCGACTT-3′	NMR titration
Site1_EMSA_sense	5′-TCCTCCCAGCTCAGCGCCAAC-3′	EMSA studies
Site2_EMSA_sense	5′-GGAACAACGCGCCGACTTTTTCAG-3′	EMSA studies
Mut_Site1_sense[a]	5′-TCCTCCAATCATAATCGCCAAC-3′	EMSA studies
Mut_Site2_sense[a]	5′-GGAACAATATATTGACTTTTTCAG-3′	EMSA studies

[a]The mutated bases are underlined.

A ^{13}C and ^{15}N-labeled VapB46$^{1-54}$ sample was prepared to assign its backbone atoms. Triple resonance experiments HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO were used to sequentially assign the backbone amides. Out of 54 residues, 50 residues were unambiguously assigned (Fig. 2A). The ^{1}H-, ^{15}N- and ^{13}C-assigned chemical shifts of VapB46$^{1-54}$ has been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under the accession number 27580. The chemical shifts of $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{13}CO, $^{1}H^{\alpha}$, $^{1}H^{N}$ and ^{15}N were used to predict the secondary structure of VapB46$^{1-54}$ (Fig. 2B). Residues M12–Q54 contain secondary structures consisting of two α-helices, α1 and α2, and three β-strands, β1, β2 and β3. This is consistent with the Phd/YefM antitoxin fold. Residues M1–T11 were predicted to be unstructured by MICS [37] which was further confirmed by heteronuclear $\{^{1}H\}$–^{15}N steady state NOE (Fig. 2C). Small and negative hetero NOE values for residues M1–T11 are indicative of a flexible region. Positive hetero NOE values in the range of 0.65–0.90 for residues M12–Q54 indicate a well-folded rigid domain.

VapB46 forms tetramer in solution

Phd/YefM domains typically form dimers [14,15,38,39]. As the secondary structures of the N-terminal residues (1–54) predicted from NMR chemical shifts are similar to a Phd/YefM-like domain, the oligomeric state of VapB46 was investigated. Two complementary methods were used, viz. size exclusion chromatography (SEC) and chemical cross-linking experiment. In SEC, VapB46 elutes at ~ 60 mL, which corresponds to a molecular mass of 49 kDa while the calculated molecular mass of monomeric VapB46 is 12.1 kDa. This indicates that VapB46 forms a tetramer in solution (Fig. 3A,B and Table 3). On the other hand, VapB46$^{1-54}$ elutes at ~ 84 mL, which corresponds to a molecular mass of 19 kDa. The calculated molecular mass of monomeric VapB46$^{1-54}$ is 7.2 kDa. This indicates that the N-terminal domain of VapB46 forms a dimer. Chemical cross-linking shows a significant band at ~ 35 kDa for VapB46, while for VapB46$^{1-54}$ a band at ~ 14 kDa is observed (Fig. 3C,D and Table 3). Hence, the isolated N-terminal domain of VapB46 forms a stable dimer similar to a Phd/YefM domain whereas VapB46 forms a higher oligomer, most likely a tetramer.

Table 2. Summary of CD experiments on VapB46 and VapB46$^{1-54}$

	VapB46	VapB46$^{1-54}$
Helix (%)	12	45
Strand (%)	31	14
Others (%)	57	41
Melting temperature (°C)	~ 59	~ 63

It is important to note that the ^{1}H–^{15}N HSQC spectra of VapB46$^{1-54}$ have ~ 50 peaks, one for each amide N-H in the monomer. Since the amides of the same residues in the two monomeric unit have the same chemical shift, VapB46$^{1-54}$ must form a symmetric dimer, which is similar to that observed for the YefM antitoxin N-terminal domain [40]. The dimerization and tetramerization of VapB46 are also consistent with the weaker signal-to-noise obtained from these proteins in the NMR experiments. Increase in size of proteins results in slower tumbling of the molecules, which in turn results in faster relaxation of the NMR signal. The VapB46$^{1-54}$ dimer and the VapB46 tetramer are ~ 14 and ~ 49 kDa size complexes, respectively, and thus have significantly reduced signal-to-noise than that expected from the monomeric protein.

VapB46 binds at two sites in its promoter–operator region

To identify the sequence-specific binding sites of VapB46 in the 148 bp DNA sequence (Fig. 4A), present

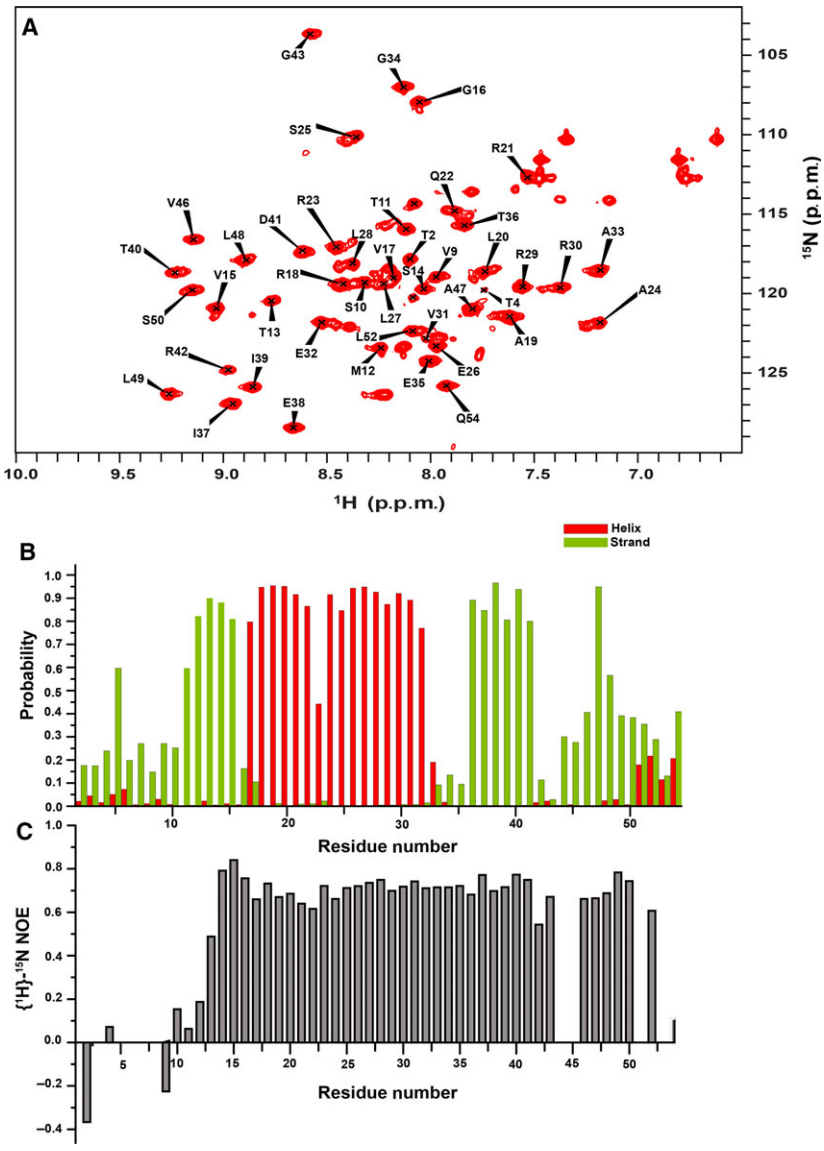

Fig. 2. Secondary structure and dynamics of VapB46$^{1–54}$. (A) Assigned ^{1}H–^{15}N HSQC spectra of ^{15}N-labeled VapB46$^{1–54}$ at 35 °C. (B) Secondary structure prediction of VapB46$^{1–54}$ from backbone chemical shifts. Helices and strands are colored red and grey, respectively. (C) {^{1}H}–^{15}N steady state NOE of VapB46$^{1–54}$ shows that residues M1–T11 are flexible while residues M12–Q54 are rigid, which is characteristic of a folded domain.

immediately upstream of the start codon of the VapB46 gene in the Mtb genome, EMSA-based DNA binding and DNase I footprinting studies were performed. Binding of VapB46 to the 148 bp DNA sequence was confirmed by concentration-dependent EMSA (Fig. 4B). To confirm that the mobility shift is due to the binding of the DNA to VapB46, unlabeled 148 bp DNA sequence was titrated in as a competitor. As the concentration of unlabeled probe increased, the band corresponding to the protein–DNA complex diminished, and at 150-fold molar excess of unlabeled DNA, the band almost disappeared (Fig. 4C). This clearly shows that mobility shifts of the labeled 148 bp DNA are due to binding with VapB46. All EMSA experiments were done in the presence of salmon sperm DNA to prevent non-specific interactions.

Specific sequences in the 148 bp DNA bound by VapB46 were identified by DNase I footprinting experiments. The DNA sequence was found to be protected by VapB46 binding at two regions, site 1 (5′-CCAGCTCAGC-3′) and site 2 (5′-AACGCGCC GACTT-3′), which are separated by a 17 bp region (Fig. 4D). This was further confirmed by EMSA. A 52 bp DNA, containing both sites separated by a 17 bp region, showed concentration-dependent binding to VapB46 (Fig. 4E), which can be competed off by unlabeled 52 bp DNA (Fig. 4F). EMSA was further performed individually using site 1 and site 2 sequences. Both sequences showed concentration-

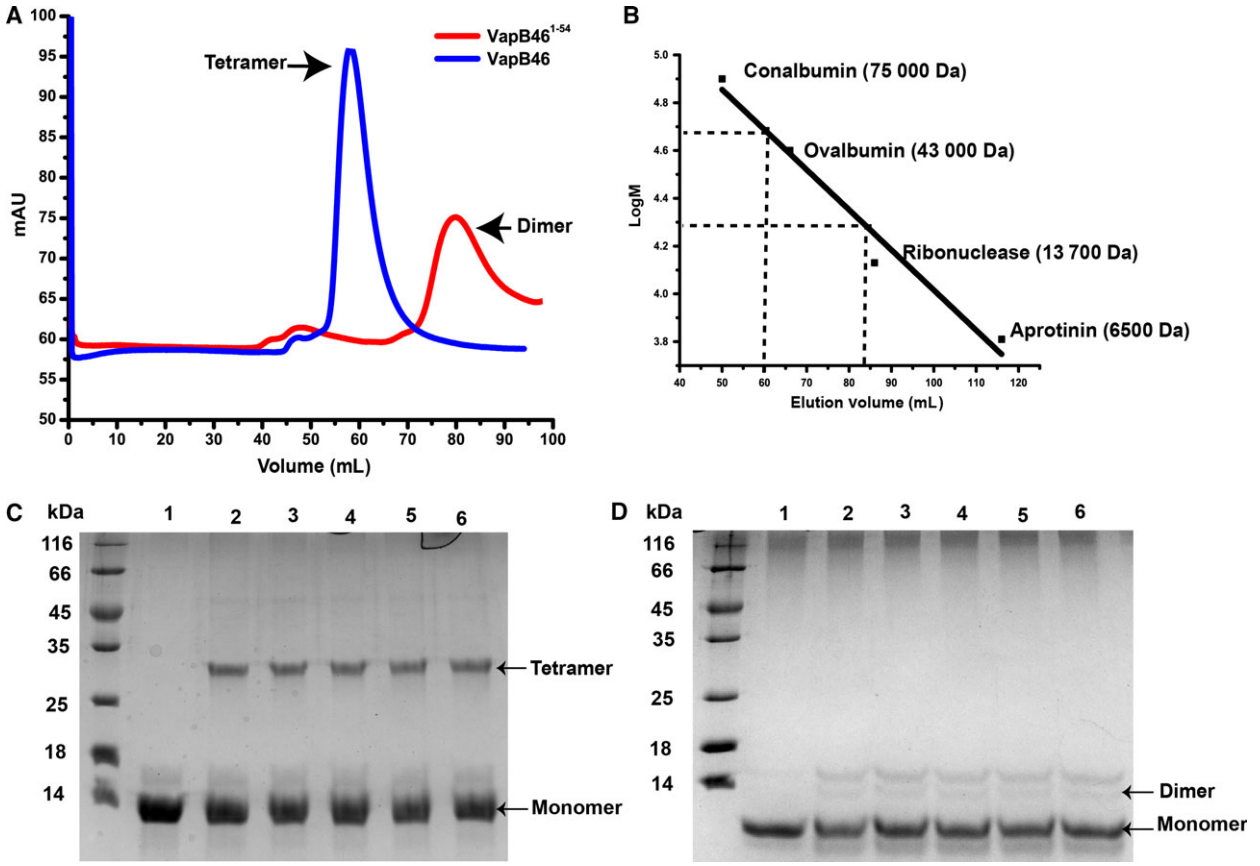

Fig. 3. Oligomerization evaluated by size exclusion chromatography and chemical cross-linking experiments. (A) Superposed SEC chromatograms of full-length VapB46 and VapB46$^{1-54}$ show the respective elution volumes. (B) Standard curve of Superdex 75 16/60 column is shown as solid line. The elution volumes of VapB46 and VapB46$^{1-54}$ projected on this line allow the determination of the corresponding molecular masses. Oligomerization of VapB46 (C) and VapB46$^{1-54}$ (D) was also evaluated by chemical cross-linking. Each protein (0.2 mg·mL^{-1}) was mixed with increasing concentration of glutaraldehyde (lane 1 to lane 6: 0%, 0.02%, 0.04%, 0.06%, 0.08% and 0.1% glutaraldehyde).

Table 3. Oligomerization of VapB46 and N-terminal domain VapB46$^{1-54}$

	Calculated molecular mass[a] (kDa)	Size exclusion chromatography (kDa)	Chemical cross-linking (kDa)	Oligomeric state
VapB46	12.1	49	~ 35[#]	Tetramer
VapB46$^{1-54}$	7.2	19	~ 14	Dimer

[a]Molecular mass of each protein is calculated including the residual fusion tag from the expression vector.
[#]Cross-linked proteins may have anomalous migration in SDS-PAGE.

dependent binding (Fig. 4G and J) that can be competed off by the respective unlabeled DNA sequence (Fig. 4H and K). Furthermore, respective unlabeled mutated sequences (Table 1) could not compete off the binding of the labeled site 1 or site 2 sequences (Fig. 4I and L). Thus, VapB46 binds to two sites in the 148 bp promoter sequence and this binding is DNA sequence specific. Although not quantified, it is clear from the EMSA experiments that VapB46 binds the 52 bp DNA tighter than the individual sites.

Identification of the DNA-binding surface of VapB46

In order to determine the DNA binding surface of VapB46, two titration experiments were performed. The ^{15}N-labeled VapB46$^{1-54}$ was titrated with site 1 or site 2 DNA fragments (Table 1) and ^{1}H–^{15}N HSQC spectra collected for each titration point. The titration experiments show that the interaction of VapB46$^{1-54}$ with both DNA sequences occurs in a fast exchange regime, which is characterized by the smooth transition of peaks from the free to the bound chemical shift

(Fig. 5A,B). In both experiments, the amide NH of residues V17, R18, L20, R21, R23, R30, V31, D41, R42 and G43 showed significant chemical shift perturbations (Fig. 5C). The chemical shift perturbation of VapB46$^{1-54}$ is very similar in both titrations, indicating that VapB46$^{1-54}$ binds site 1 and site 2 DNA sequences with the same binding interface. Chemical shift-based secondary structure prediction of VapB46$^{1-54}$ showed that residues R18, L20 and R21 are located in helix α1 while R23 is located on the loop connecting helices α1 and α2. Residues R30 and V31 are on the α2 helix. Residue D41 is located on strand β2 while R42 and G43 are located on the β-hairpin loop connecting strands β2 and β3 (Figs 5C and 7). Thus, helix α1 and the β-hairpin loop show major chemical shift perturbation.

Our NMR and SEC experiments indicate that free VapB46$^{1-54}$ exists as a symmetric dimer. Similarly, the protein in the (VapB46$^{1-54}$)$_2$–DNA complex most likely remains as a symmetric dimer, which is similar to Phd antitoxin from P1 bacteriophage bound to its operator DNA [38]. However, the DNA may introduce asymmetry in the (VapB46$^{1-54}$)$_2$–DNA interface (Fig. 7). Our NMR titration experiments with site 1 and site 2 DNA fragments show that the binding is in a fast exchange time scale. This results in an average chemical shift for the interface residues. Peak splitting for the interface residues would have been observed only if the DNA binding was in the slow time scale that typically results from very tight binding. Thus, the weak DNA binding affinity of VapB46$^{1-54}$ results in fast chemical exchange leading to average signals for interface residues that most likely masks the asymmetry of the (VapB46$^{1-54}$)$_2$–DNA interface.

The chemical shift perturbations were used to estimate the dissociation constant (K_d) of VapB46$^{1-54}$ binding to site 1 and site 2 DNA sequences (Fig. 5D, E). The dissociation constants were determined to be 12 ± 6 and 24 ± 8 μM for site 1 and site 2, respectively. The micromolar range dissociation constants are consistent with the fast exchange time scale observed in the titration experiments. The site 1 DNA sequence is bound more tightly by VapB46$^{1-54}$ than is the site 2 DNA sequence.

Chemical shift ROSETTA predicts Phd/YefM domain like structure for VapB46$^{1-54}$

Chemical shift ROSETTA (CS-ROSETTA) utilizes the backbone chemical shifts to predict torsion angles of each residue, which in turn is used to determine three-dimensional structures of the protein [41]. Since the VapB46$^{1-54}$ forms a symmetric dimer and the monomeric unit is only 7.2 kDa, the backbone chemical shifts of $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, $^{1}H^{\alpha}$, $^{1}H^{N}$ and ^{15}N were used to predict the three-dimensional structure of its monomer. The final structure of VapB46$^{1-54}$ generated by CS-ROSETTA was selected based on the convergence of the run. This structure is consistent with the secondary structure prediction from MICS and rigid domain segment determined by heteronuclear $\{^1H\}$–^{15}N steady state NOE data. The calculated VapB46$^{1-54}$ structure resembles that of the Phd/YefM DNA binding fold from residues M12 to Q54 which consists of two α-helices, α1 and α2, and three β-strands, β1, β2 and β3 (Fig. 6B). Structural alignment of VapB46$^{1-54}$ from residue M12 to Q54 with the N terminus of Phd antitoxin (PDB code: 4MZ0) [38] has a backbone rmsd of 1.7 Å and that of the N terminus of YefM antitoxin (PDB code: 2A6Q) [40,42] has a backbone rmsd of 2.0 Å (Fig. 6B,C).

To confirm the predicted structure of VapB46$^{1-54}$, ^{15}N-NOESY with 100 ms mixing time was acquired. Secondary structures can be identified by characteristic NOE patterns between H^{α} and H^{N} protons [42]. The residues in the two α-helices showed diagnostic cross peaks between neighboring amide protons ($^{1}H^{N}$), and also between α ($^{1}H^{\alpha}$) and amide protons (Fig. 6D). Similarly, the β-strands showed strong $^{1}H^{\alpha}_{i-1}$–$^{1}H^{N}_{i}$ NOEs (Fig. 6D) and more importantly, cross strand NOEs between α ($^{1}H^{\alpha}$) and amide ($^{1}H^{N}$) protons (Fig. 6E). Collectively, the backbone chemical shifts, $\{^1H\}$–^{15}N steady state NOE and NOESY cross peaks provide good confidence in the structural model of VapB46$^{1-54}$.

To better understand the DNA-binding interface, a model of the VapB46$^{1-54}$–site 1 DNA complex was built using the DNA bound structure of N-terminal Phd antitoxin (PDB code: 4MZ0) [38] as template (Fig. 7). This model is consistent with our NMR-based titration data. Residues (R18, R21, R23 and R42) that show significant chemical shift perturbation are indeed present in the DNA-binding interface (Fig. 7A). The model shows that helix α1 inserts into the major groove of the DNA with side chains of R18 and R21 making direct contact with the nucleotide bases, thus enabling a direct readout of the nucleic acid sequence. The side chain of R42, present on the β-hairpin loop connecting strands β2 and β3, inserts into the minor grove of the DNA while R23 on the helical turn interacts with the negatively charged phosphate backbone of the DNA. Residues L20, V31 on the dimeric interface of VapB46$^{1-54}$ (Fig. 7B) and R30 on the α2 helix showed significant chemical shift perturbation and are not directly involved in DNA binding. Chemical shift perturbation of negatively charged D41 most likely

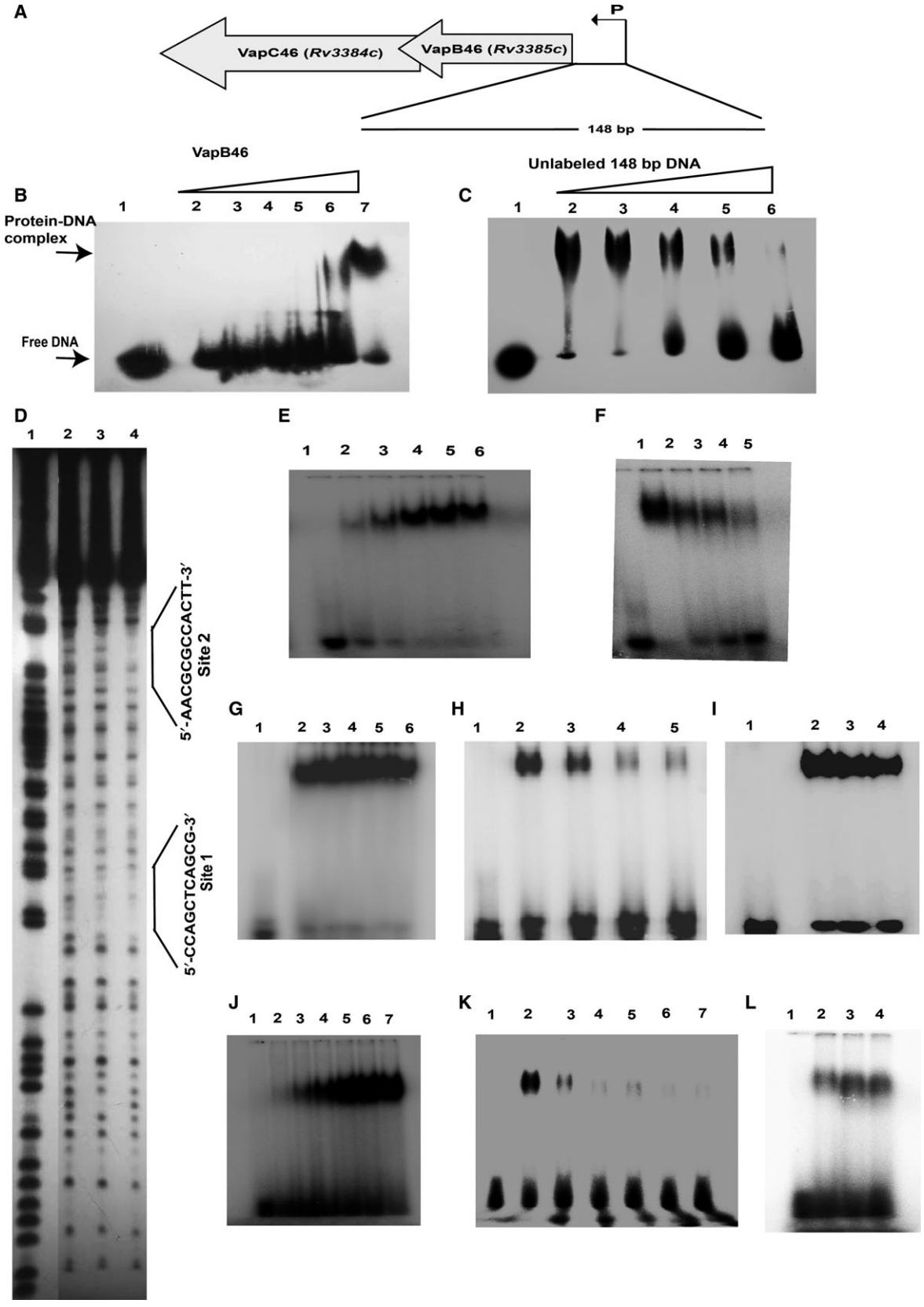
A
P
VapC46 (*Rv3384c*)
VapB46 (*Rv3385c*)
148 bp
VapB46
B
1 2 3 4 5 6 7
Protein-DNA complex
Free DNA
Unlabeled 148 bp DNA
C
1 2 3 4 5 6
D
1 2 3 4
5′-AACGCGCCACTT-3′
Site 2
5′-CCAGCTCAGCG-3′
Site 1
E
1 2 3 4 5 6
F
1 2 3 4 5
G
1 2 3 4 5 6
H
1 2 3 4 5
I
1 2 3 4
J
1 2 3 4 5 6 7
K
1 2 3 4 5 6 7
L
1 2 3 4

Fig. 4. DNA binding studies of VapB46. (A) Organization of VapBC46 toxin–antitoxin operon showing 148 bp DNA sequence upstream of the start codon of VapB46. (B) EMSA with the 148 bp double stranded DNA (dsDNA). Lane 1: no protein; lanes 2–7: 6, 12, 18, 24, 30 and 36 μM protein. (C) Competition EMSA with unlabeled 148 bp dsDNA. Lane 1: no protein; lane 2: no unlabeled 148 bp dsDNA sequence used; lanes 3–6: 20-, 60-, 100- and 150-fold molar excess unlabeled 148 bp dsDNA sequence used as competitor. In all lanes 36 μM protein was used. (D) DNase I footprinting assay. Lane 1: A+G ladder; lane 2: no protein; lanes 3 and 4: 24 and 36 μM protein. The protected sites are outlined at the right side and marked as site 1 and site 2. (E) EMSA with 52 bp DNA sequence having both site 1 and site 2 sequences. Lane 1: no protein; lanes 2–6: 10, 20, 40, 60 and 80 μM of VapB46 used. (F) Competition EMSA with unlabeled 52 bp. Lane 1: no protein; lane 2: 36 μM protein, no unlabeled DNA sequence; lanes 3–5: 10-, 20- and 50-fold molar excess unlabeled 52 bp dsDNA. (G) EMSA with site 1 dsDNA. Lane 1: no protein; lanes 2–6: 18, 24, 30, 36 and 48 μM of VapB46. (H) Competition EMSA with site 1 dsDNA. Lane 1: no protein; lane 2: 36 μM VapB46; lanes 3–5: 10-, 50- and 100-fold molar excess unlabeled site 1 dsDNA. (I) Competition EMSA using mutated site 1 dsDNA; lane 1: no protein; lanes 2–4: 50-, 100- and 150-fold molar excess unlabeled mutated site 1 dsDNA used as competitor. In all lanes 36 μM VapB46 used. (J) EMSA with site 2 dsDNA. Lane 1: no protein; lanes 2–7: 6, 12, 24, 36, 48 and 60 μM of VapB46 used. (K) Competition EMSA with unlabeled site 2 dsDNA to VapB46. Lane 1: no protein; lane 2: 36 μM VapB46; lanes 3–7: 20-, 50-, 100-, 120- and 150-fold molar excess unlabeled site 2 dsDNA used. (L) Competition EMSA using mutated site 2 dsDNA. Lane 1: no protein; lane 2: no competitor; lanes 3 and 4: 100- and 150-fold molar excess of unlabeled mutated site 2 dsDNA and 36 μM VapB46 used.

arises from the electrostatic repulsion from the negatively charged phosphate backbone.

Discussion

Toxin–antitoxin (TA) systems provide unique mechanisms for programmed cell cycle arrest or even cell death in bacteria [1,12]. Understanding the underlying mechanism of control and activation of these systems can provide novel routes for antibacterial therapies. In the deadly pathogen *Mycobacterium tuberculosis* (Mtb), VapBC constitutes the largest type II TA family with 47 members [20], and thus, is an important therapeutic target. Previous studies have mostly focused on the toxicity of Mtb VapC toxins and their interaction with cognate VapB antitoxins, but the ability of VapB antitoxins to bind the promoter–operator site remains unclear. To address the binding property of VapB with DNA, we have structurally characterized the interaction of VapB46 with its promoter–operator region to shed light on its regulatory mechanism.

The DNA binding domain of VapB antitoxins is typically found in their N-terminal region. For example, the N-terminal sequences of VapB antitoxin from *Shigella flexineri* and VapB2 from *Rickettsia felis* contain an AbrB-type DNA-binding domain [43,44]. VapB26 from *Mycobacterium tuberculosis* and FitA antitoxin from *Neisseria gonorrhoeae* have a ribbon–helix–helix DNA binding motif in their N terminus [30,45]. Structural analyses of Mtb VapB46 by CD and NMR spectroscopy reveal that the N-terminal residues 12–54 form a well-folded Phd/YefM domain, while the remaining C-terminal residues are unstructured. Using NMR titration experiments, we showed that this N-terminal domain of VapB46 is responsible for DNA binding. Thus, VapB46 also conforms to the paradigm of antitoxins with DNA binding domain in their N-terminal region.

Transcriptional autoregulation of toxin–antitoxin operons by antitoxins is mediated by binding to palindromic DNA sequence in the promoter–operator region upstream of their start codons [13,14,46–48]. We have shown that VapB46 recognizes two distinct binding sites in the promoter–operator region separated by 17 nucleotides. This is similar to antitoxins like RelB, CcdA and ParD [47,49,50]. Inspection of the two binding sites of VapB46 showed that core site 2 DNA sequence consists of a 6 bp palindrome-like sequence 5′-CGCGCC-3′, while site 1 has similar sequence 5′-aGCtCa-3′ with three out of six bases matching the core consensus site 2 DNA sequence. Compared to site 1 or site 2 alone, the binding of VapB46 is stronger to a 52 bp DNA fragment consisting of both site 1 and site 2.

Studies on various antitoxins have also shown that they mostly exist as dimers [14,30,46,49,51]. RelB from *Escherichia coli* is an exception that exists as a tetramer [47]. Using size exclusion chromatography and chemical cross-linking analysis, we found that VapB46 exists as a tetramer, while the N-terminal Phd/YefM domain ($VapB46^{1–54}$) exists as a dimer in solution. VapB46 is most likely the first VapB antitoxin that has been shown to form a tetramer. As the N-terminal VapB46 structurally resembles the N-terminal Phd and YefM antitoxins [38,40], a comparison of VapB46 sequence and structure with these antitoxins (Fig. 6A–C) suggests that the dimerization of $VapB46^{1–54}$ is mostly likely mediated by interaction of the hydrophobic patch of residues V46–S50 present on the β3 strand of both monomers with each other and also with helices α1 and α2. Moreover, the tetramerization of the full-length VapB46 is most likely mediated by its

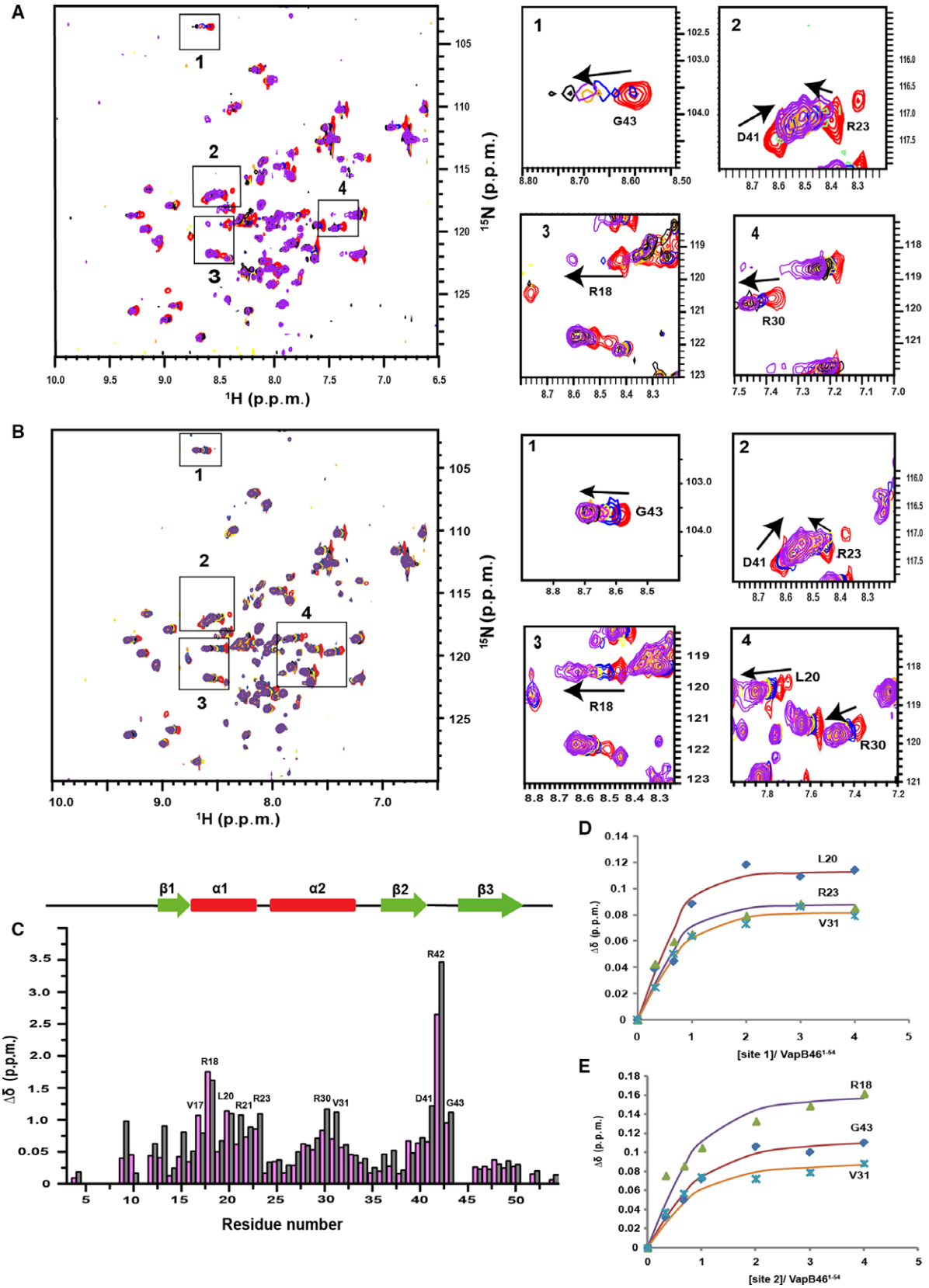
A
B
C
D
E
^{15}N (p.p.m.)
^{1}H (p.p.m.)
G43
D41
R23
R18
R30
L20
β1
α1
α2
β2
β3
Δδ (p.p.m.)
Residue number
V17
R21
V31
R42
[site 1]/ VapB46$^{1-54}$
[site 2]/ VapB46$^{1-54}$

Fig. 5. DNA titration and chemical shift mapping of DNA binding interface. The ^{15}N-labeled VapB46$^{1–54}$ was titrated with site 1 (A) and site 2 (B) dsDNA. For each titration point, ^{1}H–^{15}N HSQC spectra were collected and are shown as overlays. Expanded regions of residues with significant chemical shift changes are separately shown for both titrations. The black arrow indicates the chemical shift change from free to DNA-bound state. The [dsDNA]/[VapB46$^{1–54}$] ratios are as follows: red, 0 : 1; blue, 1 : 0.33; yellow, 1 : 0.67; magenta, 1 : 1; pink, 1 : 2; orange, 1 : 3; purple, 1 : 4. (C) Chemical shift perturbation (Δδ) between the free and the bound protein are plotted for each residue. The predicted secondary structure is shown on the top, helices as red cylinders and strands as green arrows. Residues perturbed upon titration with site 1 and site 2 are represented by magenta and grey bars. Residues with significant chemical shifts are labeled. Similar perturbation is observed for both titrations and it indicates both DNA sequences bind to the same surface on VapB46$^{1–54}$. Representative data fitting to evaluate dissociation constant (K_d) for site 1 (D) and site 2 (E). Changes in the combined amide chemical shifts (Δδ) are plotted as a function of respective DNA to VapB46$^{1–54}$ molar ratio.

C-terminal region. The helical propensity of 10 consecutive residues throughout the C-terminus of VapB46 (55–102) was determined using the program AGADIR [52]. This revealed that residues V87–E99, present towards the end, have a relatively high helical propensity (Fig. 8). Hence, it may be postulated that these residues can form a helical bundle leading to the tetrameric and higher oligomeric states of VapB46. Thus, the tetrameric VapB46 acts as a bidentate ligand that can bind the two sites in the promoter–operator region (Fig. 9). This effectively increases the affinity of VapB46 towards the promoter–operator region in

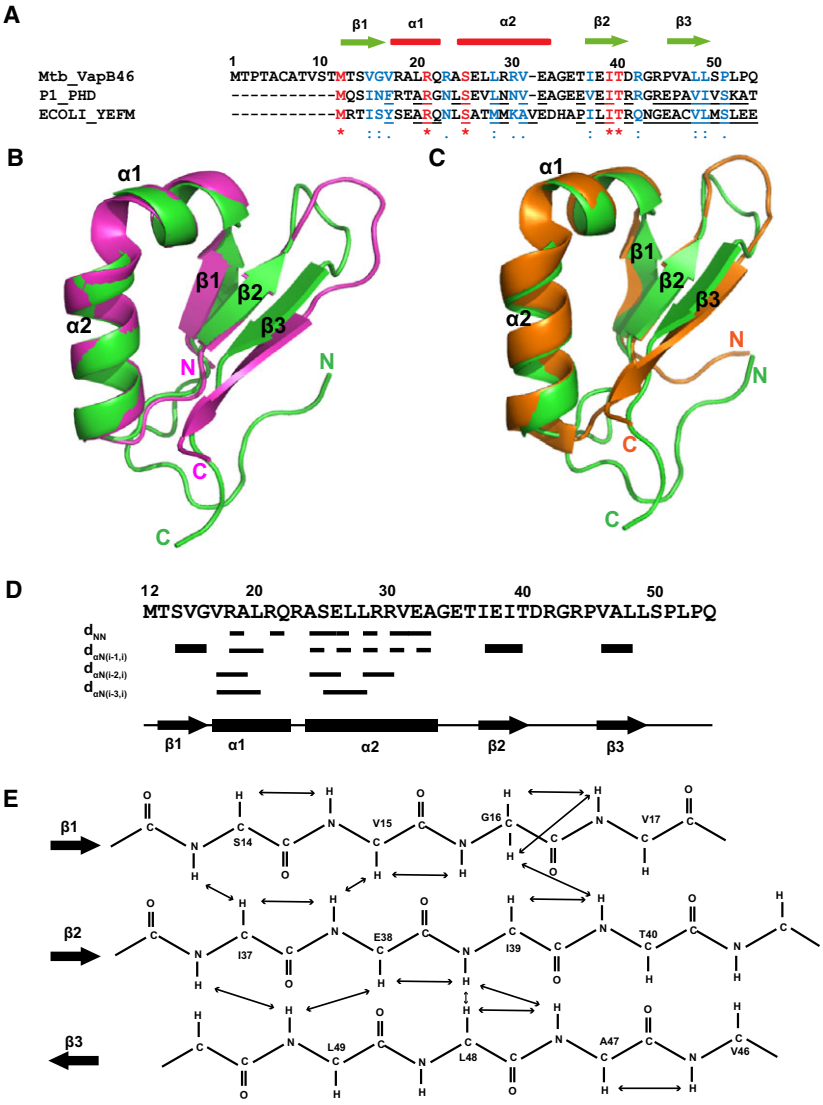

Fig. 6. Structural comparison of VapB46$^{1–54}$ with Phd and YefM antitoxins. (A) Multiple sequence alignment is shown for VapB46$^{1–54}$ with Phd antitoxin from P1 bacteriophage and YefM antitoxin from *Escherichia coli*. The interface residues of Phd and YefM dimers are underlined. Using backbone chemical shifts, a CS-ROSETTA model of VapB46$^{1–54}$ is generated. This model shows two α-helices and three β-strands forming a β-sheet. The VapB46$^{1–54}$ model is superposed with the DNA binding domain of Phd antitoxin (residues 1–43) (B) and YefM antitoxin (residues 10–53) (C) using PYMOL. VapB46$^{1–54}$ is green, Phd is magenta and YefM is orange. (D) Short range NOE patterns are shown. α-Helices and β-strands are shown as rectangles and arrows, respectively. Unambiguous NOEs between $^{1}H^{N}$ and $^{1}H^{\alpha}$ are labeled as $d_{NN(i-1,i)}$, $d_{\alpha N(i,i-1)}$, $d_{\alpha N(i,i-2)}$, $d_{\alpha N(i,i-3)}$. The relative NOE strength (weak/strong) is reflected by the bar thickness (mixing time = 100 ms). (E) The three β-strands are shown as β1, β2 and β3, and the unambiguous NOEs are indicated by arrows. Both sequential and cross strand NOEs are observed.

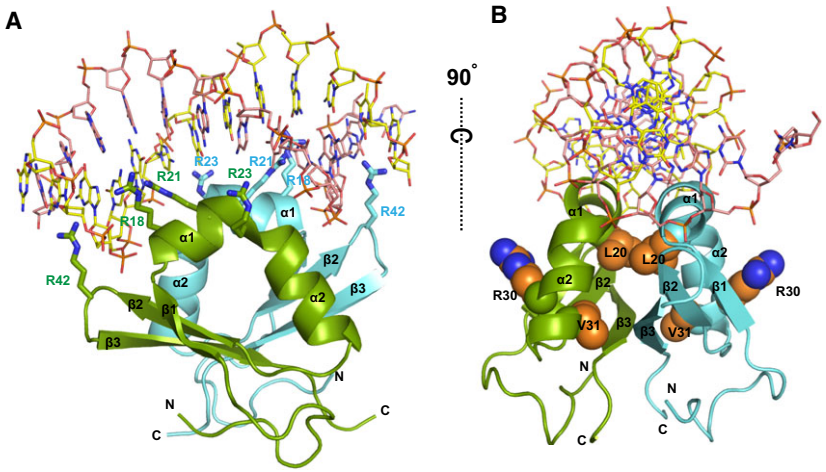

Fig. 7. Model of VapB46$^{1–54}$–DNA complex. (A) The CS-ROSETTA model of VapB46$^{1–54}$ is docked onto a dsDNA with site 1 sequence using PDB id: 4ZM0 as a template. Here VapB46$^{1–54}$ is shown as a symmetric dimer as no peak splitting is observed in the NMR titration experiments. The two monomers (green and cyan) are represented in cartoon format. This complete structure was energy minimized by YASARA. Arginine residues, which have significant chemical shift perturbation upon interaction with DNA, are shown as sticks. (B) A rotated view of the complex shows the dimer interface. Residues L20, V31 and R30, which also show significant chemical shift perturbation upon DNA titration, are shown in sphere representation. L20 and V31 are buried in the dimer interface.

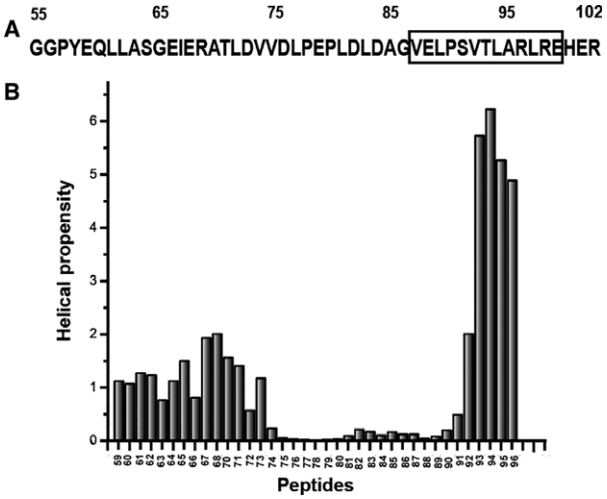

Fig. 8. Prediction of helical propensity by AGADIR. (A) The C-terminal sequence of VapB46 (VapB46$^{55–102}$) was divided into 38 peptides of 10 residues each and submitted to the AGADIR server. The residues V87–E99 (boxed) show high helical propensity. (B) The helical propensity for each 10-residue peptide is shown where the number in the *x*-axis is for the fifth residue of respective peptide.

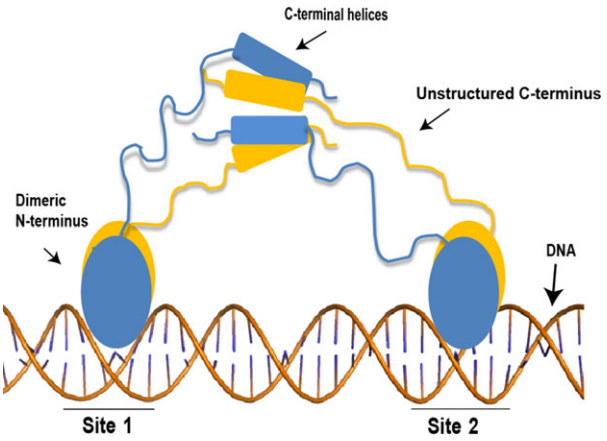

Fig. 9. Model of interaction of VapB46 with its promoter–operator region. Each site on the DNA is bound with an N-terminal dimer of VapB46. The N-terminal dimeric subunits of VapB46 are represented in blue and yellow and the two sites on DNA are shown as site 1 and site 2. The two dimers are further held together in a tetrameric form by the C-terminal helices that most likely form a helical bundle. The cylinders represent the putative helices formed at the end of the disordered C terminus of VapB46.

comparison to the dimeric N-terminal region binding to individual sites. As C-terminal regions of antitoxins typically bind their cognate toxin [30,38,40,43,45,51], the VapC46 toxin most likely binds the C terminus of VapB46. This interaction may disrupt the helical bundle formation.

The DNA-bound Phd antitoxin from P1 bacteriophage [38] shows that helix α1 from both monomers inserts into the major groove of the DNA. The structural model of VapB46$^{1–54}$, which resembles the Phd domain, is dimeric and NMR titration data show that the same interface containing helix α1 is involved in DNA binding (Figs 5C and 7A). This highlights a conserved mechanism of DNA recognition by Phd domains where the α1 helix inserts into the major groove of the

DNA and is parallel to the helical axis of the DNA. Mutation of the equivalent residues of R18 and R21 (VapB46 numbering, Fig. 6A) on the helix α1 in Phd antitoxin have been shown to completely disrupt the interaction of the protein with DNA [46]. Thus, the amino acids on helix α1 determine the sequence specificity for DNA binding (Fig. 7A). Interestingly, residues L20 and V31 on the dimeric interface as well as R30 present on the α2 helix (Fig. 5C) also showed significant chemical shift perturbation. Binding of VapB46$^{1–54}$ to DNA most likely triggers structural rearrangement in the interface, which can lead to the chemical shift perturbation of the backbone NH of these residues.

Although VapB46$^{1–54}$ structurally aligns with Phd/YefM DNA-binding domain, the identified promoter–operator sites differ from the core motif 5′-GTAC-3′ recognized by YefM and Phd antitoxins from *E. coli* and P1 bacteriophage [14,40], respectively. Interestingly, the DNA sequence recognized by the mycobacterial VapB26 also consists of GC repeats [30]. Despite structural similarities with well-known DNA binding domains of VapB antitoxins, whether all mycobacterial VapB antitoxins recognize the GC-rich palindromic-like sequence for transcription regulation of mycobacterial VapBC operon needs to be investigated.

In conclusion, we have structurally characterized the antitoxin VapB46 and identified the DNA sequence that VapB46 specifically binds via its N-terminal domain. We have also identified amino acid residues of VapB46 that interact with promoter–operator region. Thus, our findings should provide a strong basis for elucidation of the transcriptional autoregulation mechanism by VapB46 antitoxin.

Materials and methods

Bioinformatics

Sequence analysis was done using the InterPro database [35]. Multiple sequence alignment was performed using CLUSTALW [53]. Parameters such as molecular mass and amino acid composition were determined using the PROTPARAM tool [54]. Interface residues were analyzed by PISA [55].

Cloning, overexpression and purification of protein

The Gene encoding VapB46 (*Rv3385c*) was amplified from the H37Rv genome of Mtb using sequence-specific primers with BamHI and HindIII restriction sites in the forward primer (5′-CGGGATCCATGACGCCGACCGCTTGTGC-3′) and reverse primer (5′-CCCAAGCTTTCAACGCTCGTGCT CACGC-3′), respectively. The N-terminal domain of VapB46$^{1–54}$ (Table 1) was cloned using the same forward primer and a new reverse primer (5′- CCCAAGCTTTCACTG CGGCAGCGGCGAGAG- 3′) with HindIII restriction site. Genes encoding VapB46 and VapB46$^{1–54}$ were cloned into pQE30 (Qiagen, Germantown, MD, USA) vector and transformed into SG13009 and M15 *E. coli* cells, respectively. These constructs have a 6x-His affinity tag appended to their N terminus. *Escherichia coli* cells harboring VapB46 and VapB46$^{1–54}$ were cultured in 2 L and 4 L LB medium, respectively, to a D_{600} of 0.6 at 37 °C and induced with 0.5 mM IPTG for 4 h at 37 °C. For NMR studies, these cells were cultured in 1 L M9 minimal medium supplemented with 1 g·L^{-1} of ^{15}N-labeled ammonium chloride and 2 g·L^{-1} of ^{13}C-labeled glucose as the sole nitrogen and carbon source, respectively. Cells were grown to a D_{600} of ~ 0.8 at 37 °C and induced with 1 mM IPTG for 18 h at 20 °C. The 6x-His-tagged proteins were purified using Ni-NTA affinity chromatography followed by size exclusion chromatography. Sample purity was checked in 15% SDS/PAGE and mass was confirmed by matrix-assisted laser desorption/ionization time of flight. The concentration of purified protein was estimated using the Bradford assay. Yield of VapB46 from 1 L LB and 1 L M9 media was 4 mg and 1.2 mg, respectively. Also, the protein precipitated within a few days. Yield of VapB46$^{1–54}$ from both 1 L LB and 1 L M9 media was 1 mg. This protein was stable in solution for 3 weeks.

Electrophoretic mobility shift assay

A 148 bp DNA sequence, which is upstream of the gene encoding VapB46, was PCR amplified and 5′ end labeled with [γ-^{32}P]ATP (BRIT, Hyderabad, India) using T4 polynucleotide kinase. The binding reaction was set with varying VapB46 concentration, 1 μL of 0.06 μM labeled DNA, 3 μL of 10× binding buffer (100 mM Tris/HCl pH 8.0, 300 mM NaCl, 30 mM $MgCl_2$, 1 mM EDTA, 20% glycerol) and 2 μg of salmon sperm DNA in a total of 30 μL reaction volume. The mixture was incubated at 4 °C for 1 h and then separated on 5% native PAGE. For EMSA with 52 bp, site 1 and site 2 DNA sequences (22 bp each) were first oligo annealed and then labeled with [γ-^{32}P]ATP. The concentrations of labeled 52 bp, site 1 and site 2 DNA sequences used for the binding reaction were 0.18 μM, 0.03 μM and 0.06 μM, respectively. The binding reaction was set in a similar manner to that above. For the competition assay, unlabeled DNA fragment was added to the binding reaction mix and electrophoresis was carried out. Gels were vacuum dried, and the bands were visualized by autoradiography.

DNase I footprinting

Protein–DNA complex was formed by incubating 24 and 36 μM VapB46 with 2 μL of [γ-^{32}P]ATP-labeled DNA, 0.5 M $MgCl_2$, 0.1 M $CaCl_2$ in 10 μL of 10× binding buffer

(100 mM Tris/HCl pH 8.0, 300 mM NaCl, 30 mM $MgCl_2$, 1 mM EDTA, 20% glycerol) and 2 μg of salmon sperm DNA in a total of 100 μL reaction volume. DNase I (0.04 U) was added and the reaction mix was incubated for 3 min at 37 °C. It was stopped with a stop buffer (50 mM Tris/HCl pH 8.0, 2% SDS, 100 mM EDTA, 0.4 $\mu g \cdot mL^{-1}$ proteinase K). The mixture was purified by the phenol–chloroform method followed by resuspension in loading buffer (98% formamide, 10 mM EDTA, 0.025% xylene cyanol and 0.025 bromophenol blue). Samples were boiled for 10 min at 95 °C, chilled rapidly and separated by gel electrophoresis on 8% sequencing gel (PAGE) and autoradiographed.

Circular dichroism spectroscopy

Both VapB46 and VapB46$^{1–54}$ were prepared in 50 mM potassium phosphate buffer pH 6.5, 50 mM NaCl and 2 mM DTT at a concentration of 10 and 20 μM, respectively. CD spectra were obtained in Jasco J-815 (Easton, MD, USA) at a scan rate of 50 $nm \cdot min^{-1}$ with 0.1 nm data pitch, using 1 nm bandwidth for two accumulations at 25 °C. Thermal stability analysis was performed by monitoring the CD peak at 222 nm with a temperature change rate of 5 $°C \cdot min^{-1}$ in the range of 10–95 °C.

Size exclusion chromatography

SEC was performed on a Superdex 75 16/60 column (GE Healthcare, Chicago, IL, USA) using the AKTA FPLC system (GE Healthcare) at 4 °C at 1 $mL \cdot min^{-1}$ flow rate in 50 mM HEPES pH 7.2, 50 mM NaCl and 2 mM DTT. The molecular masses of the proteins were determined by comparison of the elution volume with the calibration curve plot of elution volume *vs* log*M*.

Chemical cross-linking

VapB46 and VapB46$^{1–54}$ were dialyzed in a cross-linking buffer of 50 mM HEPES pH 7.2, 50 mM NaCl and 2 mM DTT. Proteins (0.2 $mg \cdot mL^{-1}$) were mixed with an increasing concentration of glutaraldehyde (0.02–0.1%) in 20 μL reaction volume at 4 °C for 1 h. Reactions were stopped by adding SDS loading buffer, and samples were heated at 95 °C for 5 min and run on 15% SDS/PAGE.

Chemical shift assignment of VapB46$^{1–54}$

NMR samples (0.3 mM) were in 50 mM potassium phosphate buffer at pH 6.5 containing 50 mM NaCl, 2 mM DTT and 7% D_2O for spin lock. All the NMR experiments were performed with a Bruker (Billerica, MA, USA) AvanceIII 600 MHz. Data were processed using NMRPIPE and NMRDRAW [56] while spectra were analysed using SPARKY [57]. The $^1H–^{15}N$-HSQC spectrum of VapB46$^{1–54}$ was assigned by sequential assignment strategy using 3D experiments, such as HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO. Secondary structures of VapB46$^{1–54}$ were predicted based on the chemical shifts of $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^1H^{\alpha}$, $^1H^N$, ^{15}N and ^{13}CO using MICS [37]. The ^1H, ^{15}N and ^{13}C assigned chemical shifts of VapB46$^{1–54}$ have been deposited in BioMagResBank (http://www.bmrb.wisc.edu/) under the accession number 27580.

Chemical shift perturbation study

Double stranded DNA of 10 bp and 13 bp in length (Table 1) corresponding to site 1 and site 2, respectively, was titrated into ^{15}N-labeled VapB46$^{1–54}$ and the interaction was monitored by $^1H–^{15}N$ HSQC experiments. The molar ratios of site 1 dsDNA to VapB46$^{1–54}$ and site 2 dsDNA to VapB46$^{1–54}$ in both titration experiments were 0, 0.33, 0.67, 1, 2, 3 and 4. Reference $^1H–^{15}N$ HSQC spectra were recorded for VapB46$^{1–54}$ at a concentration of 0.2 and 0.17 mM, respectively, prior to titration with site 1 and site 2. The final concentration of the protein was 0.14 and 0.12 mM, respectively. The magnitude of chemical shift perturbation (Δδ) of amide protons of VapB46$^{1–54}$ upon titration with DNA was calculated using the equation:

$$\Delta\delta = [(\Delta\delta_H)^2 + (0.154\Delta\delta_N)^2]^{1/2} \quad (1)$$

where $\Delta\delta_H$ and $\Delta\delta_N$ are ^1H and ^{15}N chemical shift changes, respectively, between the free and bound states of the protein at each titration point and Δδ is the combined chemical shift change. The dissociation constant (K_d) of the protein bound to DNA sequences was calculated according to the equation:

$$\Delta\delta_{obs} = \frac{\Delta\delta_{max}\{([P]_t + [L]_t + K_d) - [([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t]^{1/2}\}}{2[P]_t} \quad (2)$$

where $\Delta\delta_{obs}$ is the observed shift from the free state, $\Delta\delta_{max}$ is the maximum shift change upon saturation, $[P]_t$ is the total concentration of protein, $[L]_t$ is the total concentration of DNA and K_d is the dissociation constant.

Model of VapB46$^{1–54}$–DNA complex

A model of the VapB46$^{1–54}$–DNA complex was built using the structure of Phd antitoxin bound to operator DNA (PDB code: 4MZ0) [38] as a template. The monomeric structure of VapB46$^{1–54}$ was modeled based on chemical shifts using CS-ROSETTA [41]. A VapB46$^{1–54}$ dimeric model was built by superimposing onto the Phd antitoxin dimer (PDB code: 4MZ0) used as the template using PYMOL [58]. The nucleotide sequence of the template DNA was mutated into VapBC46 promoter–operator site 1 using COOT version 0.8.9 [59]. The resultant model was energy minimized to

remove steric clashes using the YASARA energy minimization server [60].

Acknowledgements

We acknowledge the Central Research Facility of IIT Kharagpur for NMR data collection and the Central Research Facility of Department of Chemistry, IIT Kharagpur for CD spectroscopy facility. We also acknowledge Anirban Ghosh (Department of Biophysics, Bose Institute), for valuable discussion regarding the CD experiment and Susmitnarayan Chaudhury, Department of Chemistry, IIT Kharagpur for help in CD data interpretation. The genomic DNA of H37Rv was a kind gift from Prof. Harald G. Wiker of University of Bergen, Norway. This work was supported by SGBSI Grant IIT/SRIC/BIO/LDO/2014-15/33 of IIT Kharagpur and BT/PR12404/BRB/10/1362/2014 dated 26-08-2016 to AKD; and SERB Grant ECR/2016/000847 dated 07-03-2017 to SD.

Author contributions

MR, AKD and SD planned the work. MR performed all experiments. MR and SD analyzed the data. MR, SD and AKD wrote the manuscript. AK and SDG helped with EMSA and DNase I footprinting experiments. SD and AB helped in NMR spectroscopy and biophysical studies.

References

1 Gerdes K, Rasmussen PB & Molin S (1986) Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci USA* **83**, 3116–3120.

2 Roberts RC, Spangler C & Helinski DR (1993) Characteristics and significance of DNA binding activity of plasmid stabilization protein ParD from the broad host-range plasmid RK2. *J Biol Chem* **268**, 27109–27117.

3 Aizenman E, Engelberg-Kulka H & Glaser G (1996) An *Escherichia coli* chromosomal "addiction module" regulated by guanosine [corrected] 3′,5′-bispyrophosphate: a model for programmed bacterial cell death. *Proc Natl Acad Sci USA* **93**, 6059–6063.

4 Engelberg-Kulka H, Amitai S, Kolodkin-Gal I & Hazan R (2006) Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genet* **2**, 1518–1526.

5 Lewis K (2010) Persister Cells. *Annu Rev Microbiol* **64**, 357–372.

6 Gerdes K & Maisonneuve E (2012) Bacterial persistence and toxin-antitoxin loci. *Annu Rev Microbiol* **66**, 103–123.

7 Keren I, Shah D, Spoering A, Kaldalu N & Lewis K (2004) Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J Bacteriol* **186**, 8172–8180.

8 Singh R, Iii CEB & Boshoff HIM (2010) The three RelE homologs of *Mycobacterium tuberculosis* have individual, drug-specific effects on bacterial antibiotic tolerance. *J Bacteriol* **192**, 1279–1291.

9 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.

10 Buts L, Lah J, Dao-Thi M-H, Wyns L & Loris R (2005) Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends Biochem Sci* **30**, 672–679.

11 Gerdes K, Christensen SK & Løbner-Olesen A (2005) Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* **3**, 371–382.

12 Hayes F (2003) Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *J Bacteriol* **301**, 1496–1500.

13 Fivian-Hughes AS & Davis EO (2010) Analyzing the regulatory role of the HigA antitoxin within *Mycobacterium tuberculosis*. *J Bacteriol* **192**, 4348–4356.

14 Kędzierska B, Lian LY & Hayes F (2007) Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression. *Nucleic Acids Res* **35**, 325–339.

15 Garcia-Pino A, Balasubramanian S, Wyns L, Gazit E, De Greve H, Magnuson RD, Charlier D, van Nuland NA & Loris R (2010) Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* **142**, 101–111.

16 Christensen-Dalsgaard M & Gerdes K (2006) Two higBA loci in the Vibrio cholerae superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Mol Microbiol* **62**, 397–411.

17 Gerdes K (2000) Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J Bacteriol* **182**, 561–572.

18 Pedersen K, Christensen SK & Gerdes K (2002) Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Mol Microbiol* **45**, 501–510.

19 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.

20 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.

21 Arcus VL, McKenzie JL, Robson J & Cook GM (2011) The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. *Prot Eng Design Select* **24**, 33–40.
22 Cruz JW, Sharp JD, Hoffer ED, Maehigashi T, Vvedenskaya IO, Konkimalla A, Husson RN, Nickels BE, Dunham CM & Woychik NA (2015) Growth-regulating *Mycobacterium tuberculosis* VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat Commun* **6**, 7480.
23 Winther K, Tree JJ, Tollervey D & Gerdes K (2016) VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. *Nucleic Acids Res* **16**, 9860–9871.
24 Winther KS, Brodersen DE, Brown AK & Gerdes K (2013) VapC20 of *Mycobacterium tuberculosis* cleaves the sarcin-ricin loop of 23S rRNA. *Nat Commun* **4**, 2796.
25 Winther KS & Gerdes K (2011) Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci USA* **108**, 7403–7407.
26 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.
27 Min AB, Miallau L, Sawaya MR, Habel J, Cascio D & Eisenberg D (2012) The crystal structure of the Rv0301-Rv0300 VapBC-3 toxin-antitoxin complex from *M. tuberculosis* reveals a Mg^{2+} ion in the active site and a putative RNA-binding site. *Prot Sci* **21**, 1754–1767.
28 Das U, Pogenberg V, Subhramanyam UKT, Wilmanns M, Gourinath S & Srinivasan A (2014) Crystal structure of the VapBC-15 complex from *Mycobacterium tuberculosis* reveals a two-metal ion dependent PIN-domain ribonuclease and a variable mode of toxin-antitoxin assembly. *J Struct Biol* **188**, 249–258.
29 Lee I-G, Lee SJ, Chae S, Lee K-Y, Kim J-H & Lee B-J (2015) Structural and functional studies of the *Mycobacterium tuberculosis* VapBC30 toxin-antitoxin system: implications for the design of novel antimicrobial peptides. *Nucleic Acids Res* **43**, 7624–7637.
30 Kang S, Kim D, Lee K, Park SJ, Yoon H, Lee J, Im H & Lee B (2017) Functional details of the *Mycobacterium tuberculosis* VapBC26 toxin-antitoxin system based on a structural study: insights into unique binding and antibiotic peptides. *Nucleic Acids Res* **45**, 8564–8580.
31 Bukowski M, Rojowska A & Wladyka B (2011) Prokaryotic toxin-antitoxin systems – the role in bacterial physiology and application in molecular biology. *Acta Biochim Pol* **58**, 1–9.
32 Gupta M, Nayyar N, Chawla M, Sitaraman R, Bhatnagar R & Banerjee N (2016) The Chromosomal parDE2 toxin-antitoxin system of *Mycobacterium tuberculosis* H37Rv: genetic and functional characterization. *Front Microbiol* **7**, 886.
33 Zhao L & Zhang J (2008) Biochemical characterization of a chromosomal toxin-antitoxin system in *Mycobacterium tuberculosis*. *FEBS Lett* **582**, 710–714.
34 Gupta A, Venkataraman B, Vasudevan M & Bankar KG (2017) Co-expression network analysis of toxin-antitoxin loci in *Mycobacterium tuberculosis* reveals key modulators of cellular stress. *Sci Rep* **7**, 1–14.
35 Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L *et al*. (2009) InterPro: The integrative protein signature database. *Nucleic Acids Res* **37**, D211–D215.
36 Sreerama N & Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252–260.
37 Shen Y & Bax Ad (2012) Identification of helix capping motifs and beta-turns from NMR chemical shifts. *J Biomol NMR* **52**, 211–232.
38 Garcia-Pino A, De Gieter S, Talavera A, De Greve H, Efremov RG & Loris R (2016) An intrinsically disordered entropic switch determines allostery in Phd-Doc regulation. *Nat Chem Biol* **12**, 490–496.
39 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.
40 Kamada K & Hanaoka F (2005) Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Mol Cell* **19**, 497–509.
41 Shen Y, Bryan Philip N, Yanan H, Orban J, Baker D & Bax A (2010) De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci* **19**, 349–356.
42 Wüthrich K (1986) NMR of Proteins and Nucleic Acids, pp. 262–284. Wiley, New York, NY.
43 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.
44 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 Res* **40**, 3245–3258.
45 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.
46 Zhao X & Magnuson RD (2005) Percolation of the Phd repressor-operator interface. *J Bacteriol* **187**, 1901–1912.
47 Li G, Zhang Y, Inouye M & Ikura M (2008) Structural mechanism of transcriptional autorepression of the *Escherichia coli* RelB/RelE antitoxin/toxin module. *J Mol Biol* **380**, 107–119.
48 Khoo SK, Loll B, Ting W, Shoeman RL, Ngoo L, Yeo C & Meinhart A (2007) Molecular and structural characterization of the PezAT chromosomal toxin-antitoxin system of the human pathogen *Streptococcus pneumoniae*. *J Biol Chem* **282**, 19606–19618.
49 Madl T, Van Melderen L, Mine N, Respondek M, Oberer M, Keller W, Khatai L & Zangger K (2006) Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA. *J Mol Biol* **364**, 170–185.
50 Oberer M, Zangger 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. *Prot Sci* **16**, 1676–1688.
51 Zorzini V, Buts L, Schrank E, Sterckx YGJ, Respondek M, Engelberg-kulka H, Loris R, Zangger K & Van Nuland AJ (2015) *Escherichia coli* antitoxin MazE as transcription factor : insights into MazE-DNA binding. *Nucleic Acids Res* **43**, 1241–1256.
52 Pace NC & Scholtz JM (1988) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* **75**, 422–427.
53 Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R *et al.* (2007) ClustalW and ClustalX version 2. *Bioinformatics* **23**, 2947–2948.
54 Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (Walker JM, ed), pp. 571–607. Humana Press, New York, NY.
55 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
56 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.
57 Kneller DG & Kuntz ID (1993) UCSF Sparky an NMR display, ANNOT4TIOK and assignment tool. *J Cell Biochem* **53**, 254.
58 DeLano WL (2002) The PyMOL Molecular Graphics System. DeLANO Scientific, Palo Alto, CA, USA. Website at http://www.pymol.org.
59 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D: Biol Crystallogr* **60**, 2126–2132.
60 Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D & Karplus K (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* **77**, 114–122.