Published in final edited form as:
Structure. 2014 November 4; 22(11): 1650–1656. doi:10.1016/j.str.2014.08.018.

Structure and DNA-binding traits of the Transition State Regulator AbrB

Andrew L. Olson[1], **Ashley T. Tucker**[1], **Benjamin G. Bobay**[1], **Erik J. Soderblom**[2], **M. Arthur Moseley**[2], **Richele J. Thompson**[1], and **John Cavanagh**[1,*]

[1]Department of Molecular and Structural Biochemistry, North Carolina State University, 128 Polk Hall, Raleigh, NC, 27659, USA

[2]Duke Proteomics Core Facility, Institute for Genome Sciences and Policy, School of Medicine, B02 Levine Sciences Research Center, 450 Research Drive, Duke University, Durham, NC, 27710, USA

Summary

The AbrB protein from *Bacillus subtilis* is a DNA-binding global regulator controlling the onset of a vast array of protective functions under stressful conditions. Such functions include biofilm formation, antibiotic production, competence development, extracellular enzyme production, motility and sporulation. AbrB orthologues are known in a variety of prokaryotic organisms, most notably in all infectious strains of *Clostridia, Listeria* and *Bacilli.* Despite its central role in bacterial response and defense, its structure has been elusive, due to its highly dynamic character. Orienting its N- and C-terminal domains with respect to one another has been especially problematic. Here we have generated the first structure of full-length, tetrameric AbrB using NMR, chemical crosslinking and mass spectrometry. We note that AbrB possesses a strip of positive electrostatic potential encompassing its DNA binding region and that its C-terminal domain aids in DNA binding.

Keywords

AbrB structure; NMR; chemical crosslinking/MS; DNA-binding

*Corresponding author. Address: Department of Molecular and Structural Biochemistry, NC State University, Raleigh, NC 27695, USA. Fax: 919-515-2047. jcavana@ncsu.edu.

Accession Numbers
The chemical shift assignments have been deposited in the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu) under ID code 18939. The atomic coordinates of AbrBC have been deposited to the Protein Data Bank (http://www.rcsb.org) under ID code 2MJG.

Supplemental Information
Supplementary Table S1 and Figures S1–S3.

Introduction

Bacteria respond to environmental fluctuations by expressing transcription factors known as transition state regulators (TSRs). These proteins regulate countless processes, helping the bacteria survive (Sonenshein et al., 2002). Currently, the most extensively characterized TSR is AbrB from *Bacillus subtilis*. Many pathogenic organisms make use of AbrB-like TSRs to control 'survival' gene expression. These include clinically important species such as *Listeria*, *Clostridia* and *Bacilli*. Together these organisms are responsible for approximately 1 million infections annually, with up to 14,000 resulting in death (Scallan et al., 2011; Centers for Disease Control and Prevention, 2012; Battone, 2010). Of the strategies regulated by TSRs for survival, one of the most problematic, medically speaking, is their ability to control biofilm formation. Biofilms are communities of bacteria encased within a self-produced matrix of extracellular polymeric substance (EPS) (Donlan et al., 2002). Within the biofilm, bacteria are especially resilient compared to their planktonic counterparts (Stewart et al., 2001). It is estimated that 65% of all nosocomial infections directly involve biofilms (Costerton et al., 1999). Indeed, there has been specific clinical interest in the ability of *B. subtilis* biofilms to protect pathogenic bacteria (e.g., *S. aureus*) from microbicides in hospitals (Bridier et al., 2012). As AbrB is central in coordinating this response, it is vital to fully understand its mechanisms of action.

From a genetic and biochemical standpoint, AbrB directly regulates the transcription of more than 100 genes and influences hundreds more indirectly (Bobay et al., 2004; Chumsakul et al., 2011; Xu and Strauch, 1996). Interestingly, although AbrB regulates numerous genes in *B. subtilis*, it does not recognize a clear consensus DNA sequence. It has been suggested that AbrB's ability to recognize and bind an array of promoter regions may be due, in part, to its flexibility (Strauch, 1995; Bobay et al., 2004). Despite AbrB's critical role in survival and protection, the structure of the functional, full-length tetramer is not known. At present, only the structure of the N-terminal domain (AbrBN) has been elucidated (Bobay et al., 2004; Bobay et al., 2005; Coles et al., 2005). Attempts to study the full structure, via both NMR and X-ray crystallography, have been unsuccessful. This is likely due to its dynamic nature – ironically the very trait that facilitates its broad influence. (Bobay et al., 2005; Coles et al., 2005; Sullivan et al., 2008).

Full-length AbrB exists as a homotetramer with 4 identical 94-residue subunits (Cavanagh et al., 2002). Each monomer is composed of 2 separate domains: an N-terminal DNA binding domain (AbrBN – a dimer) and a C-terminal multimerization domain (AbrBC – a dimer). While both domains affect DNA binding, AbrBN's influence is much greater since it is the only domain that is currently known to directly contact DNA (Xu and Strauch, 2001; Bobay et al., 2005). To this point, the structure of AbrBC and the orientations of AbrBN and AbrBC within the full-length structure have remained notably elusive. Here, we characterized the NMR structure of AbrBC and subsequently oriented AbrBN (PDB: 1Z0R) and AbrBC with respect to one another using distance restraints from chemical crosslinking and LC/MS/MS analysis. This enabled us to produce the first structural model of full-length AbrB. AbrB is found to exhibit surface electrostatic characteristics that likely influence its DNA binding. This facet was explored computationally by docking AbrB to a known DNA target. Finally, a comparative dynamics investigation of AbrBC from *B. subtilis* and *B.*

anthracis revealed that a sequential variation leads to significant differences in flexibility in this multimerization domain. This may play a role in species-specific regulation.

Results

NMR Structure of AbrBC from *Bacillus subtilis*

When isolated, HPLC shows AbrBC (residues 52–94) exists as a dimer (Figure S1A) and its HSQC overlays with the HSQC of AbrB from *B. subtilis* (AbrB-BS) (Figure S1B). The NMR structure of AbrBC was solved using a combination of restraints derived from NOEs, dihedral angles, HNHA couplings and residual dipolar ($^1D_{HN}$) couplings (RDCs) (Table S1). Chemical shift assignments for AbrBC were obtained for 100% of backbone H^N and N nuclei as well as 93% of all backbone/sidechain nuclei (residues 55–94). The resulting ensemble of the 10 lowest AMBER energy conformers had excellent agreement with experimental data, well-defined folding and low backbone and heavy atom RMSDs (residues 55–91) of 0.251Å and 0.673Å (Figure 1A), respectively, as well as minimal distance/torsion/RDC violations with little deviation to idealized geometry.

AbrBC consists of a domain swapped-folded homodimer with structured regions at residues 65–67 (β1), 72–74(β2) and 77–91(α3). A ^{12}C-^{13}C edited/filtered NOESY showed a core consisting of extensive interactions between two sets of anti-parallel βsheets that comprise the dimer interface. This interface is comprised of strand β1 from each monomer interacting with strand β2′ of its dimer pair (Figure 1B). The α3 helix trails the β2 strand of each monomer, running anti-parallel to the dimer α3′ helix. This planar helical structure is set underneath the two sets of β-sheets. The first 13 residues of each monomeric subunit are relatively unstructured since the majority of this region links AbrBN to AbrBC, providing flexibility in binding DNA. The linker (LR) for monomeric subunit A is placed next to β2′ strand and α3′ helix of subunit B. This folding is supported by multiple NOEs found in both ^{15}N and ^{13}C NOESY data, notably between the sidechain and backbone hydrogens of residues C54 to L74 (confirmed as an intermolecular interaction by the ^{12}C-^{13}C edited/filtered NOESY).

Dimerization of AbrBC is driven by the domain swapped-fold of the β-sheets, stemming primarily from where loop residues 68–71, after the β1 strand of each subunit, cross each other. The crossing of these residues was somewhat unexpected, because hits from the Dali Database with Z-scores greater than 2.0 aligned solely to the α-helices of AbrBC, making this folding pattern unique (Holm et al., 2010). Structural orientation was confirmed with RDC refinement in both structure calculation and water refinement. The resulting correlations between calculated (D_{calc}) and experimental (D_{obs}) residual NH dipolar couplings show good agreement with a corresponding Q-factor of 0.17. Electrostatic surface profiles are shown in Figure 1C.

Chemical Crosslinking and Mass Spectrometry Analysis

Attempts to orient AbrBN and AbrBC by NOEs and RDCs proved unsuccessful due to the dynamic nature of the protein and long distances between the domains. We were finally able to orient the domains using their individual NMR structures and by chemical crosslinking,

using the reagents DSG, BS3 and previously determined crosslinks using the CID cleavable crosslinker SuDP (Olson et al., 2013). DSG, BS3, and SuDP (7.7Å, 11.4Å, and 11.2Å reagent length and maximal lysyl crosslinking distances of 20.4Å, 24.1Å, and 23.9Å, respectively) contain bifunctional amine reactive groups, crosslinking lysine amines together (Liu et al., 2012; Soderblom and Goshe, 2006; Soderblom et al., 2007). Employing varying crosslinker lengths allows for more accuracy in the modeling of AbrB (Sinz, 2006).

Successful crosslinking was analyzed by SDS–PAGE (Figure S2A). Crosslinked AbrB (Figure S2A: lanes B and C) resulted in two bands compared to un-crosslinked AbrB (lane A). While un-crosslinked AbrB has two bands consisting of monomer and dimer states (10.1 and 22 kDa, respectively), crosslinked AbrB has two bands, one for each dimer and tetramer state (~22 and ~38 kDa, respectively). Increasing amounts of crosslinking reagent did not shift the crosslinking equilibrium to the full tetramer. Following in-gel digestion of the 38 kDa band and LC/MS/MS data acquisition, all spectra were initially searched against *B. subtilis* taxonomy in the SwissProt database and dynamic mass modifications corresponding to hydrolyzed DSG and BS3 (Figure S2B–D). The crosslinking reagent modified the majority of the lysine residues within AbrB (dead-end hydrolyzed) as expected, indicating suitable reaction conditions. To identify through-space crosslinks, these spectra were then submitted to the MassMatrix (Xu and Freitas, 2009) search engine with the AbrB sequence appended to an *E. coli* proteome background.

Orientation of AbrB Domains From Observed Crosslinks

Of the 11 lysine residues in AbrB, 6 belong to AbrBN and 4 to AbrBC with the remaining lysine in the flexible linking region (K49). While the lysines of AbrBN are dispersed throughout the protein, 3 of the 4 lysines in AbrBC are in the domain-swapped β-sheet region (K66, K71 and K76). A list of inter-domain crosslinks resulting from treatment of AbrB with DSG, BS3 and SuDP are in Table 1. Among the 3 reagents there were 27 observed species with: (i) 5 crosslinks in a single AbrBC homodimer (combined intra- and inter-AbrBC domain), (ii) 9 crosslinks to a single AbrBN homodimer (combined intra- and inter-AbrBN domain) and (iii) 13 inter-domain crosslinks (AbrBN-AbrBC). Of the 13 inter-domain crosslinks, 7 were used to arrange the AbrB domains with distances spanning from 11.3 to 24.0Å. The remaining 6 crosslinks were located in the highly flexible linker and could not be reliably used.

The 7 valid crosslinking restraints were used in PyMOL to position the AbrBC domain relative to the AbrBN domain (based on the maximum allowed distances – PDB: 1Z0R) (PyMOL Molecular Graphics System, Schrödinger, LLC). Due to AbrB's tetrameric nature, there are 8 possible crosslinks for each pair of crosslinked lysyl groups. However, half can be eliminated since AbrBN has two faces, with one exceeding the maximal crosslinking distance. AbrBC domain C1 can have crosslinks to residues in both N1 and N4 AbrBN domains, but not to N2 and N3 AbrBN domains. The remaining 4 crosslinks were used to orient AbrBC and AbrBN. The crosslinks used between domains are shown on the fully assembled tetrameric AbrB structure (Figure 2A and 2B).

Between the 2 crosslinking reagents used here and the previous crosslinks from SuDP, there are 7 inter-domain crosslinks usable for orientation. All inter-domain crosslinks drive

orientation to make AbrBC's domain-swapped region face into and split the set of AbrBN domains (Figure 2B) medially. This orientation is supported by 4 inter-domain crosslinks to K71 (K2, K9, K31 and K46) from AbrBN, ultimately positioning the helical core outwards from the structure. In this arrangement both AbrBN domains are located so that DNA binding regions are coextended up.

Crosslinks from any of the three reagents place AbrBC in this particular orientation. However a combination of crosslinkers of *different* lengths helps derive distance restraints for modeling. The 4Å difference between DSG and BS3/SuDP provides the distance gap for this purpose. The K9-K76 crosslink is seen with both BS3 and SuDP but is absent with DSG, providing a restraint of >20.4Å between the Cαs of K9 and K76 (Figure 2B – orange lines). Since each domain was positioned independently, linking regions for each monomer (connecting AbrBN to AbrBC through residues 48–52) were added and minimized in explicit solvent using AMBER (Case et al., 2012) (Figure 2B). The alignment of the domains forms a disc-like quaternary structure 63Å and 55Å in diameter across AbrBN domains and AbrBC domains respectively. The AbrBC domain swapped-fold region is oriented so that it fits aptly into the space formed by adjacent AbrBN dimers. The electrostatic plot of AbrB's DNA-binding surface shown in Figure 2C (see later). tCONCOORD conformational space analysis (Seeliger and De Groot, 2009) reveals that the linking regions of AbrB allow significant domain movement of AbrBN relative to AbrBC and vice versa (Figure 2D and Figure S3A). The wide array of space sampled by AbrBN allows alternate conformations in binding many tertiary structures of DNA while the AbrBC domains keep the active tetrameric form intact.

AbrBC mobility in AbrB homologs

Structural studies on the full length AbrB protein from *B. anthracis* have proven unsuccessful thus far. While the AbrBN domains from *B. anthracis* are identical to *B. subtilis* both sequentially and structurally, the AbrBC domains are quite different (based on HSQC differences shown in Figure S2C). AbrBC from *B. subtilis* (AbrBC-BS) has well dispersed chemical shifts, while AbrBC from *B. anthracis* (AbrBC-BA) has ^{1}H chemical shifts clustered around ~7.9–8.2 ppm. Chemical shift studies and secondary structure prediction of AbrB-BA suggested that AbrBC-BA likely contains two helices, possibly forming a bundle of 4 helices and the multimerization domain (Olson et al., 2012). This would differ from the AbrB-BS structure elucidated here, that contains a mix of β-sheets and an α-helix. Interestingly, while the resonance intensities of all peaks (AbrBN and AbrBC) in the AbrB-BS spectra were similar, there was an average ~7-fold difference in intensity between AbrBC-BA (stronger) and AbrBN-BA resonances.

To investigate the nature of this disparity, the backbone flexibility of both AbrB homologs were investigated using ^{1}H-^{15}N heteronuclear NOE experiments (Figure S1D). Exchange broadening occurs in two loop regions of AbrBN (residues 29–32 and 40–44) of both homologs and the linking region (48–54) of AbrBN-BS. The AbrBN domains of both homologs are overall structurally rigid with nearly identical average ^{1}H-^{15}N NOE values over residues 9–50 (0.720 and 0.728 for AbrB-BS and AbrB-BA respectively). Conversely, the AbrBC domains of both homologs are very different (residues 55–89). AbrBC-BS has a

slightly lower average ^{1}H-^{15}N NOE value of 0.567, suggesting slightly more flexibility than AbrBN-BS, which is not surprising considering the main role of AbrBC is to provide quaternary structure in multimerization. Unexpectedly, AbrBC-BA has a resoundingly lower average ^{1}H-^{15}N NOE value of 0.201, which indicates significant mobility.

Discussion

Since its discovery almost 25 years ago (Strauch et al., 1989), AbrB has been extensively studied, reflecting its importance in gene transcription. With its pivotal role in a host of other bacterial defenses, including biofilm formation, AbrB's mechanism of action is crucial to understanding bacterial resistance, yet its specific mode of gene recognition has been elusive. While it has been suggested that gene identification by AbrB depends, to some degree, on DNA tertiary structure, a comprehensive structural study of AbrB itself has been noticeably absent.

In this study, structural/biophysical methods were used to characterize the full structure of the TSR AbrB from *B. subtilis*. Initially, the AbrBC domain of AbrB-BS was determined to be an independent dimer and its NMR structure solved. The orientation of N- and C-terminal domains in the full-length AbrB-BS tetramer was resolved using long-range distance restraints identified from a variety of chemical crosslinkers and mass spectrometry. Computational docking was then used with the AbrB-BS structure and a target DNA sequence to evaluate binding proclivities. Differences in the flexibility of AbrBC between AbrB homologs were evaluated through NMR dynamics analysis.

Of the three crosslinking reagents used to assemble AbrB, SuDP was the most useful in obtaining inter-domain restraints. Five additional inter-domain crosslinks were observed with the SuDP reagent compared to BS3 and DSG, because SuDP possesses heightened flexibility, hydrophobicity, and reactivity, enabling further crosslink modification (Soderblom and Goshe, 2006). However, the critical distance restraint necessary (K9-K76) could only be derived using information from all three reagents.

The disc-like structure of AbrB and the extended AbrBN-to-AbrBC linker allows AbrB to bind many tertiary DNA elements. The overall conformation of AbrB exhibits an extraordinary electrostatic arrangement. The positive surface on the interior of AbrBC (domain-swapped region) contributes to a "lane" of positive electrostatic potential encompassing the DNA binding regions of AbrBN. The negatively charged exterior of AbrBC (α-helical core) could prevent DNA from binding in any other fashion.

To better understand the interaction between AbrB and DNA, the structural model of the complex between AbrB and its cognate promoter, *abrB8,* was created using HADDOCK molecular docking (Dominguez et al., 2003), utilizing previous mutational and docking studies (Strauch, 1995; Xu and Strauch, 1996; Sullivan et al., 2008). Docking resulted in 5 unique clusters (6Å Cα RMSD with a minimum of 4 structures/cluster) with the majority of structures in cluster 1 (141 out of 200 solutions) – see Figure 3A. The interface, as suggested in previous studies (Cavanagh et al., 2002; Sullivan et al., 2008), extends across both AbrBN domains, utilizing the positively charged residues necessary for DNA binding (R8, R15, R23

and R24). The AbrB tetramer is positioned into consecutive major grooves of the DNA with an overall buried surface area of 3956Å^2. In accordance with previous studies, the AbrBN loop regions (containing positively charged residues) extend into the major grooves and toward the phosphate backbone of DNA, while the helices expand towards the neighboring minor grooves. While the tertiary structure of DNA-bound and unbound AbrB is primarily the same (Cα RMSD difference of 2.55Å), the AbrBC domains move 4.3Å closer together (Figure S3C). In some cases, when the whole protein was allowed to be fully flexible, the AbrBC domains occasionally came into close proximity to one another (15Å between K71 Cα atoms - grey in Figure 3B). This is 9Å closer than the docked complex in Figure 3A. This docking result corroborates previous pseudo cell-size SEC-MS experiments, which indicated that AbrB decreases in hydrodynamic volume upon DNA binding (Cavanagh et al., 2002). This is consistent with AbrBC domains compressing around the target DNA (Figure 3A–B). It is no surprise that AbrB utilizes positively charged residues to facilitate DNA binding; however, the electrostatic surface potential in AbrB forms a “lane” of positive charge across both AbrBN dimers (Figures 2C and 3C–D). Although AbrBC is primarily a multimerization domain, its electrostatic characteristics likely aid DNA binding. The electrostatic surface of AbrBC reveals the two halves have opposing charges - the α-helical core of the protein is negatively charged, while the domain-swapped β-sheets are positively charged (Figure 1C). The electrostatic surface potential of the docked AbrB/DNA complex of Figure 3A is shown in Figures 3C–D.

AbrBN is known to be AbrB’s primary DNA binding domain. Its DNA recognition and specificity determinants have been proposed to lie solely within its N-terminal amino acid sequence. (Xu and Strauch, 2001). Currently, the precise role of AbrBC is still unclear, with no conclusive data indicating its direct interaction with DNA. Previous studies have shown that AbrBC mutations at C54, L67 and Q81-termination codon decreased DNA binding likely due to altered multimeric interactions (Xu et. al. 1996). AbrBC mutations at N64 and L67 are also known to affect the multimerization state and hence DNA binding (Strauch – unpublished data). AbrBC mutations at Q55, K71, E80, Q81, E85 and E90 have also been proposed to affect DNA-binding (Neubauer et al., 2014). Since no controls showing wild-type AbrBC binding to DNA are available at this time, it is not possible to fully discern whether DNA-binding is impeded due to direct binding or by changes in charge packing/ multimerization of AbrBC. AbrBC mutations that alter the spacing of adjacent sets of AbrBN dimers or disrupt AbrB’s multimeric structure will certainly indirectly influence AbrBN’s DNA-binding proclivities. Based on our structural model, Q55 and K71 would be the residues most likely to affect DNA-binding directly, whereas E80, Q81, E85 and E90 are more distal and could influence binding indirectly. Recently, it was discovered that S84 (reported as S86) becomes phosphorylated during transition and stationary phases (Soufi et al., 2010). S84 is located on the outward face of AbrB on the helical segment of AbrBC (Figure S3D), openly available to be phosphorylated (as opposed to the inward face unable to be bound by a serine kinase), further validating the orientation of the individual AbrB domains.

AbrB proteins from both *B. subtilis* and *B. anthracis* share an identical amino acid sequence from residues 1–62 but are significantly different in the remaining 32 residues. Both proteins

were found to interchangeably regulate the same cellular environment, indicating AbrBN domains are mainly responsible for the recognition of target sequences. While AbrBC does not contribute directly to DNA binding specificity, it aids in electrostatic binding of DNA, adding to its established role in multimerization. But why is there such a divergence in the flexibility of the AbrBC domain? One possibility is that the highly dynamic AbrBC domain in *B. anthracis* has a species-specific function in binding unique DNA tertiary structure(s) not found in *B. subtilis*.

In summary, by orienting the individually solved domains using crosslinking and mass spectrometry techniques, we were able to provide the first full-length tetrameric structural model of the TSR AbrB from *B. subtilis*. Computational docking supported previous suggestions on global conformational changes following DNA binding. Finally, NMR dynamic analyses provided insight into differences between AbrB from *B. subtilis* and its ortholog in *B. anthracis*.

Materials and Methods

Protein Expression and Purification

AbrB from *B. subtilis* and *B. anthracis* was expressed and purified as described previously (Bobay et al., 2005; Olson et al., 2012). AbrBC was cloned into pET-28a containing a thrombin cleavable N-terminal histidine tag. AbrBC was expressed and purified using same protocol as AbrB from *B. subtilis*.

NMR Spectroscopy and Structure Calculation

Samples for NMR experiments were dialyzed into buffer (10 mM KH_2PO_4, 15 mM KCl, 1 mM DTT, 1 mM EDTA and 0.02% NaN_3, in 10% D_2O and 100% D_2O for associated experiments) at pH 7.0 for both AbrBC and AbrB-BA and pH 7.9 for AbrB-BS. Standard NMR protocols were used to determine backbone, sidechain, and aromatic chemical shifts using 700 MHz Bruker Avance and 600 MHz Varian Inova spectrometers. $^1D_{NH}$ RDCs were measured with 1.0 mM uniformly labeled ^{15}N samples of AbrBC using IPAP-HSQC experiments on a 6.0 mm to 4.2 mm radially compressed 7% polyacrylamide gel and analyzed with REDCAT (Valafar and Prestegard, 2004). Distance restraints were obtained from ^{15}N-NOESY, ^{13}C aliphatic NOESY and ^{13}C aromatic NOESY at 75 ms and 120 ms mixing times. Structure calculations were performed with RDC restraints used in water refinement protocols. ^1H-^{15}N NOE spectra were acquired with and without saturation using a recycle delay of 8 ms. NMR spectra were processed using NMRPipe (Delaglio et al., 1995) and analyzed using NMRView (Johnson and Blevins, 1994) and Sparky (Goddard and Kneller, 2006).

Nano-Flow Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC-MS/MS) Analysis and Crosslinked Peptide Identification

Following SDS-PAGE separation, the molecular weight region corresponding to AbrB-BS (~38 kDa) was excised and subjected to an in-gel reduction, iodoacetamide alkylation, and trypsin digestion as previously described (Wilm, et al., 1996). Extracted peptides were then analyzed by nano-flow liquid chromatography on a Waters NanoACQUITY UPLC coupled

to Waters Synapt QToF high-resolution mass spectrometer as previously described (Tucker, et al., 2014). Raw LC-MS/MS data files were subjected to Mascot searches against a SwissProt (taxonomy *B. subtilis*) database (4290 forward sequences, updated Dec 2012) appended with the reverse sequence of all of the forward entries. Search tolerances were 20 ppm for precursor ions and 0.04 Da for product ions using trypsin specificity with up to two missed cleavages. Carbamidomethylation (+57.0214 Da on C) was set as a fixed modification, whereas oxidation (+15.9949 Da on M) and hydrolyzed DSG (+114.031694 Da on K) or hydrolyzed BS3 (+156.0786 Da on K) were considered as variable modifications. All searched spectra were imported into Scaffold (Proteome Software) and identification confidences were set to a <1% false discovery rate (Keller, et al., 2002; Nesvizhskii, et al., 2003). To identify crosslinked species, Mascot distiller generated MGF files were submitted to MassMatrix (v 2.4.2, Feb 2012) searches against a forward/reverse SwissProt database (taxonomy *B. subtilis*) appended with the N-terminal thrombin cleavage site modified sequence of AbrB (Xu and Freitas, 2007). Search mass tolerances and modifications were as described for Mascot searches, with the “advanced search” option enabled to allow for inter- or intrapeptide crosslinking of DSG (+96.0211 Da) or BS3 (+138.0681 Da). Specificity of the crosslinker was initially confined to lysine-lysine residues, with a secondary search towards lysine-glycine residues to allow for the mapping of crosslinked sites to the protein N-terminal primary amine. Trypsin rules were set to not allow cleavage at a crosslinked modified residue and only one crosslink per peptide pair was allowed. A peptide match within MassMatrix was only considered if peptide scoring thresholds were above that required for a matching probability less than p-value <0.05.

AbrB and DNA Docking with HADDOCK

Default HADDOCK parameters were used throughout the docking procedure (Dominguez et al., 2003). Active residues for AbrB and DNA were determined from previous studies of residues important for the ability of AbrB to bind DNA. Passive AbrB residues were defined as residues following and proceeding actives ones. A total of 2000 structures were generated for the first iteration (ridged docking), 200 for the second iteration (semi-flexible docking), and the 200 lowest energy structures were water refined. The Cα (AbrB) and backbone phosphate (DNA) RMSD values of the complexes were calculated using ProFit. A cluster analysis was performed on the final docking solutions using a minimum cluster size of 4. The RMSD cut-off for clustering was manually determined to be 6.5Å (lower than the default 7.5Å). The RMSD matrix was calculated over the backbone atoms of the interface residues. The lowest energy structure from molecular docking within the highest populated cluster was further analyzed by PSVS (Bhattacharya et al., 2007) to confirm stereochemical quality of the protein structure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This paper is dedicated to the memory of Mark Strauch. This work was supported by NIH grant RO1 GM55769 (JC) and the V Foundation for Cancer Research (JC).

References

Battone EJ. *Bacillus cereus*, a volatile human pathogen. Clin Microbiol Rev. 2010; 23:382–398. [PubMed: 20375358]

Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins: Structure, Function, and Bioinformatics. 2007; 66:778–795.

Bobay BG, Benson L, Naylor S, Feeney B, Clark AC, Goshe MB, Strauch MA, Thompson R, Cavanagh J. Evaluation of the DNA Binding Tendencies of the Transition State Regulator AbrB. Biochemistry. 2004; 43:16106–16118. [PubMed: 15610005]

Bobay BG, Andreeva A, Mueller GA, Cavanagh J, Murzin AG. Revised structure of the AbrB N-terminal domain unifies a diverse superfamily of putative DNA-binding proteins. FEBS Lett. 2005; 579:5669–5674. [PubMed: 16223496]

Bridier A, Sanchez-Vizuete MP, Le Coq D, Aymerich S, Meylheuc T, Maillard J, Thomas V, Dubois-Brissonnet F, Briandet R. Biofilms of a *Bacillus subtilis* hospital isolate protect Staphylococcus aureus from biocide action. PLoS One. 2012; 7:e44506. [PubMed: 22973457]

Case, D.; Darden, T.; Cheatham, T., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K. AMBER. Vol. 12. University of California; San Francisco: 2012.

Cavanagh J, Thompson R, Bobay B, Benson LM, Naylor S. Stoichiometries of Protein Protein/DNA Binding and Conformational Changes for the Transition-State Regulator AbrB Measured by Pseudo Cell-Size Exclusion Chromatography-Mass Spectrometry. Biochemistry. 2002; 41:7859–7865. [PubMed: 12069574]

Centers for Disease Control and Prevention. HAI. 2012. Retrieved. From: http://www.cdc.gov/HAI/

Chumsakul O, Takahashi H, Oshima T, Hishimoto T, Kanaya S, Ogasawara N, Ishikawa S. Genome-wide binding profiles of the *Bacillus subtilis* transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation. Nucleic Acids Research. 2011; 39:414–428. [PubMed: 20817675]

Coles M, Djuranovic S, Söding J, Frickey T, Koretke K, Truffault V, Martin J, Lupas AN. AbrB-like Transcription Factors Assume a Swapped Hairpin Fold that Is Evolutionarily Related to Double-Psi β Barrels. Structure. 2005; 13:919–928. [PubMed: 15939023]

Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: a common cause of persistent infections. Science. 1999; 284:1318–22. [PubMed: 10334980]

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomolecular NMR. 1995; 6:277–293.

Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc. 2003; 125:1731–1737. [PubMed: 12580598]

Donlan RM, Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. Clin Microbiol Rev. 2002; 15:167–93. [PubMed: 11932229]

Goddard, T.; Kneller, D. Sparky—NMR assignment and integration software. University of California; 2006.

Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Research. 2010; 38:W545–W549. [PubMed: 20457744]

Johnson BA, Blevins RA. NMR View: A computer program for the visualization and analysis of NMR data. J Biomolecular NMR. 1994; 4:603–614.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002; 74:5383–5392. [PubMed: 12403597]

Liu F, Wu C, Sweedler JV, Goshe MB. An enhanced protein crosslink identification strategy using CID-cleavable chemical crosslinkers and LC/MSn analysis. Proteomics. 2012; 12:401–405. [PubMed: 22213719]

Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003; 75:4646–4658. [PubMed: 14632076]

Neubauer S, Dolgova O, Präg G, Borriss R, Makarewicz O. Substitutional analysis of the C-terminal domain of AbrB revealed its essential role in DNA-binding activity. PLoS One. 2014; 9:e97254. [PubMed: 24832089]

Olson A, Bobay B, Melander C, Cavanagh J. 1H, 13C, and 15N resonance assignments and secondary structure prediction of the full-length transition state regulator AbrB from *Bacillus anthracis*. Biomolecular NMR Assignments. 2012; 6:95–98. [PubMed: 21845362]

Olson AL, Liu F, Tucker AT, Goshe MB, Cavanagh J. Chemical crosslinking and LC/MS analysis to determine protein domain orientation: Application to AbrB. Biochem Biophys Res Commun. 2013; 431:253–257. [PubMed: 23313475]

Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM. Foodborne illness acquired in the United States--major pathogens. Emerg Infect Dis. 2011; 17:7–15. [PubMed: 21192848]

Seeliger D, De Groot BL. tCONCOORD-GUI: Visually supported conformational sampling of bioactive molecules. J Computational Chemistry. 2009; 30:1160–1166.

Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom Rev. 2006; 25:663–682. [PubMed: 16477643]

Soderblom EJ, Bobay BG, Cavanagh J, Goshe MB. Tandem mass spectrometry acquisition approaches to enhance identification of protein-protein interactions using low-energy collision-induced dissociative chemical crosslinking reagents. Rapid Communications in Mass Spectrometry. 2007; 21:3395–3408. [PubMed: 17902198]

Soderblom EJ, Goshe MB. Collision-Induced Dissociative Chemical Cross-Linking Reagents and Methodology: Applications to Protein Structural Characterization Using Tandem Mass Spectrometry Analysis. Anal Chem. 2006; 78:8059–8068. [PubMed: 17134140]

Sonenshein, AL.; Hoch, JA.; Losick, R. *Bacillus subtilis* and Its Closest Relatives: From Genes to Cells. Washington, D.C: ASM Press; 2002.

Soufi B, Kumar C, Gnad F, Mann M, Mijakovic I, Macek B. Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) Applied to Quantitative Proteomics of *Bacillus subtilis*. J Proteome Res. 2010; 9:3638–3646. [PubMed: 20509597]

Stewart PS, Costerton JW. Antibiotic resistance of bacteria in biofilms. Lancet. 2001; 358:135–138. [PubMed: 11463434]

Strauch MA. In vitro binding affinity of the *Bacillus subtilis* AbrB protein to six different DNA target regions. J Bacteriology. 1995; 177:4532–4536.

Strauch M, Spiegelman G, Perego M, Johnson W, Burbulys D, Hoch J. The transition state transcription regulator AbrB of *Bacillus subtilis* is a DNA binding protein. EMBO J. 1989; 8:1615–1621. [PubMed: 2504584]

Sullivan DM, Bobay BG, Kojetin DJ, Thompson RJ, Rance M, Strauch MA, Cavanagh J. Insights into the Nature of DNA Binding of AbrB-like Transcription Factors. Structure. 2008; 16:1702–1713. [PubMed: 19000822]

Tucker AT, Bobay BG, Banse AV, Olson AL, Soderblome EJ, Moseley MA, Thompson RJ, Varney KM, Losick R, Cavanagh J. A DNA Mimic: The Structure and Mechanism of Action for the Anti-Repressor Protein AbbA. J Mol Bio. 2014; 9:1911–1924. [PubMed: 24534728]

Valafar H, Prestegard JH. REDCAT: a residual dipolar coupling analysis tool. J Magn Reson. 2004; 167:228–241. [PubMed: 15040978]

Wilm M, Schevchenko A, Houthaeve T, Breitt S, Mann M. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature. 1996; 379:466–469. [PubMed: 8559255]

Xu H, Freitas MA. A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics. 2007; 8:133. [PubMed: 17448237]

Xu H, Freitas MA. MassMatrix: A database search program for rapid characterization of proteins and peptides from tandem mass spectrometry data. Proteomics. 2009; 9:1548–1555. [PubMed: 19235167]

Xu K, Strauch MA. In vitro selection of optimal AbrB-binding sites: comparison to known in vivo sites indicates flexibility in AbrB binding and recognition of three-dimensional DNA structures. Mol Microbiol. 1996; 19:145–158. [PubMed: 8821944]

Xu K, Strauch MA. DNA-binding activity of amino-terminal domains of the *Bacillus subtilis* AbrB protein. J Bacteriol. 2001; 18:4094–4098. [PubMed: 11395475]

Xu K, Clark D, Strauch MA. Analysis of abrB mutations, mutant proteins, and why abrB does not utilize a perfect consensus in the −35 region of its sigma A promoter. J Biol Chem. 1996; 271:2621–2626. [PubMed: 8576231]

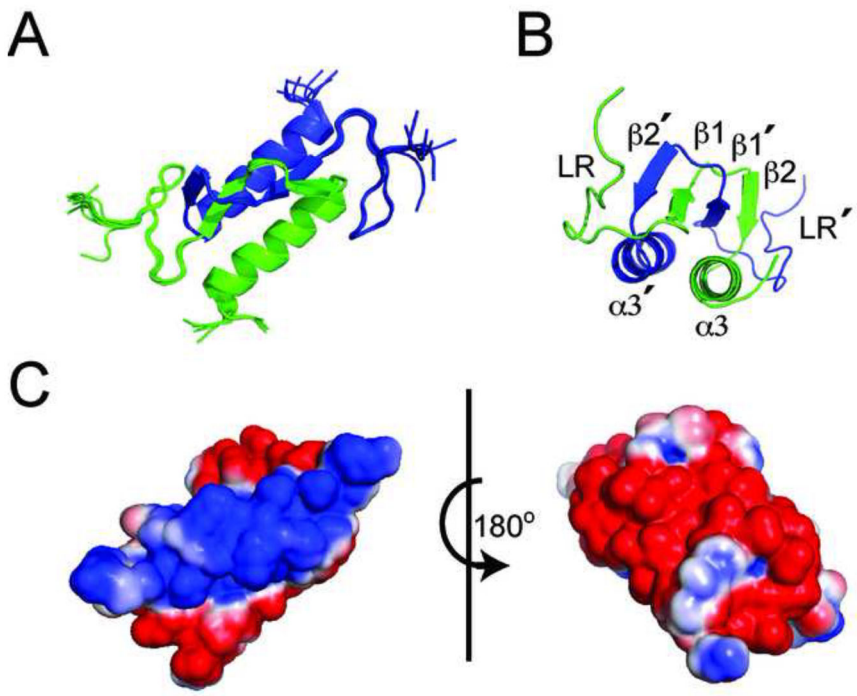

Figure 1.

The NMR structure of AbrBC. (A) The 10 lowest energy structures from water refinement. Individual monomers are green and blue. (B) Alternate view of AbrBC with relevant structural features labeled. (C) Electrostatic surface potential of AbrBC (same view as Figure 1A, rotated 180°); blue is positive charge and red is negative.

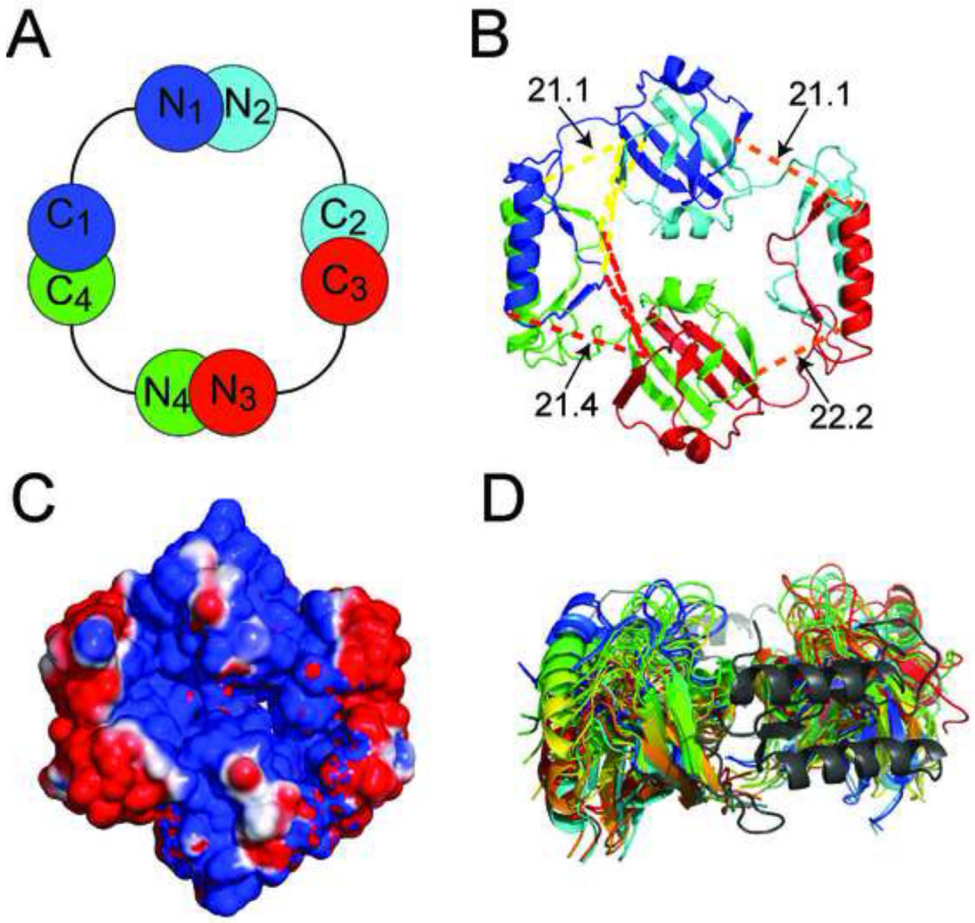

Figure 2.

The fully assembled structure of AbrB. (A) Schematic of the AbrB-BS tetramer with alternating colors corresponding to individual AbrB monomers. (B) Ribbon structure with crosslinks formed from AbrBC lysines to AbrBN lysines (in Å); yellow and red lines represent crosslinks to the neighboring AbrBN domains; orange lines represent the crosslink from K9-K76. N- and C- domains were visually docked according to the cross linking/MS data and then energy minimized via AMBER resulting in distances shown. (C) Electrostatic surface potential of AbrB (blue = positive, red = negative). (D) The conformational space sampling of AbrBN (red to cyan) with fixed positional restraints on AbrBC (grey).

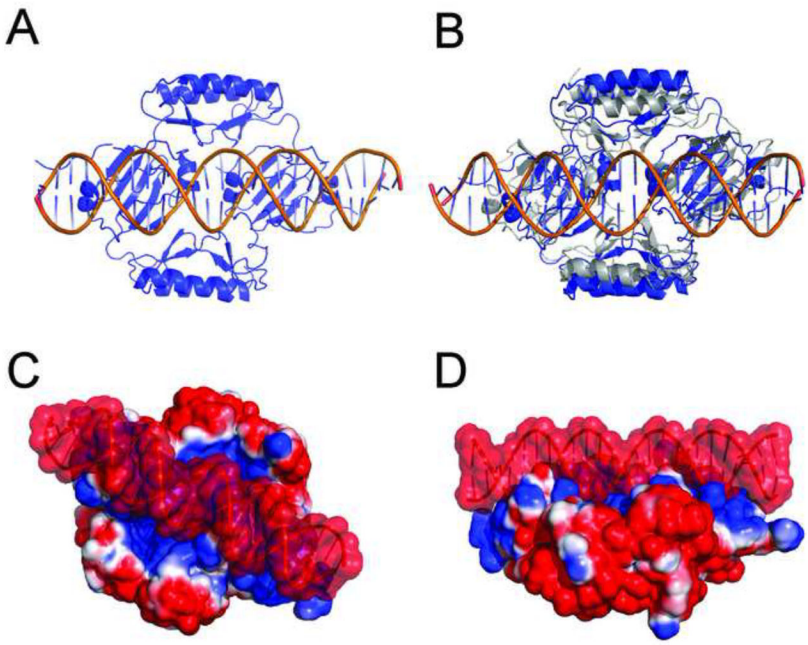

Figure 3.

AbrB upon binding DNA. (A) Model of the docked complex of AbrB and target promoter *abrB8* using HADDOCK. (B) Overlay of Figure 2A and a docked complex (blue) with all residues allowed to be fully flexible (grey). (C) and (D) alternate views of electrostatic surface potential of AbrB bound to *abrB8* DNA.

Table 1

Inter-domain crosslinks of AbrB by DSG, BS3, and SuDP.

Peptide 1	Peptide 2	Crosslink	DSG Crosslinker (20.4 Å)	BS3 Crosslinker (24.1 Å)	SuDP Crosslinker (23.9 Å)
GSHMK* STGIVR	LAGGK* LVLSK	K2-K71			X
K* VDELGR	YK* PNMTCQVTGEVSDDNLK	K9-K49	X	X	
K* VDELGR	YKPNMTCQVTGEVSDDNLK* LAGGK	K9-K66			X
K* VDELGR	LAGGK* LVLSK	K9-K71	X	X	X
K* VDELGR	LVLSK* EGAEQIISEIQNQLQNLK	K9-K76		X	X
K* VDELGR	EGAEQIISEIQNQLQNLK*	K9-K94	X	X	
TLGIAEK* DALEIYVDDEK	LAGGK* LVLSK	K31-K71			X
TLGIAEK* DALEIYVDDEK	LVLSK* EGAEQIISEIQNQLQNLK	K31-K76			X
IILK* K	YKPNMTCQVTGEVSDDNLK* LAGGK	K46-K66			X
IILK* K	LAGGK* LVLSK	K46-K71			X
IILK* K	EGAEQIISEIQNQLQNLK*	K46-K94	X	X	
YK* PNMTCQVTGEVSDDNLK	LAGGK* LVLSK	K49-K71	X	X	
YK* PNMTCQVTGEVSDDNLK	LVLSK* EGAEQIISEIQNQLQNLK	K49-K76			X

* Lysine involved in crosslinking.