



## Chemical crosslinking and LC/MS analysis to determine protein domain orientation: Application to AbrB

Andrew L. Olson<sup>1</sup>, Fan Liu<sup>1</sup>, Ashley T. Tucker, Michael B. Goshe, John Cavanagh<sup>\*</sup>

Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695, USA

### ARTICLE INFO

#### Article history:

Received 26 December 2012

Available online 9 January 2013

#### Keywords:

NMR  
Solution structure  
Chemical crosslinking  
Domain orientation  
AbrB  
Gas-phase cleavable

### ABSTRACT

To fully understand the modes of action of multi-protein complexes, it is essential to determine their overall global architecture and the specific relationships between domains and subunits. The transcription factor AbrB is a functional homotetramer consisting of two domains per monomer. Obtaining the high-resolution structure of tetrameric AbrB has been extremely challenging due to the independent character of these domains. To facilitate the structure determination process, we solved the NMR structures of both domains independently and utilized gas-phase cleavable chemical crosslinking and LC/MS<sup>n</sup> analysis to correctly position the domains within the full tetrameric AbrB protein structure.

© 2013 Elsevier Inc. All rights reserved.

### 1. Introduction

To fully understand the modes of action of multi-protein complexes, it is essential to determine their overall global architecture and the specific relationships between domains and subunits. Certainly, X-ray crystallographic studies can provide such information, but only when suitable crystals are available. In many cases, independently moving domains may make crystallization difficult [1]. Using NMR to provide such information is not always practical. Often, protein domains are separated by long distances, making NOE driven NMR investigations problematic. Both classical NOE experiments and isotope-filtered/edited NOE experiments are limited to rather short distance restraints (5–6 Å) [2]. Techniques such as paramagnetic relaxation enhancement (PRE) and Förster resonance energy transfer (FRET) have been developed to provide long-range distance information. However, these methods often require difficult and deleterious modifications to be made to the protein to obtain the distance restraints [3,4]. An appealing alternative approach comes in the form of chemical crosslinking coupled with mass spectrometry. Chemical crosslinking can be performed under physiological conditions with modest amounts of crosslinking

reagents that do not alter protein structure [5,6], which makes it a viable approach to complement high-resolution structures and offers insight into spatial orientation [7–9].

Chemical crosslinking covalently attaches the side chains of specific residues (in most cases lysines) throughout the protein and defines maximal inter-residue distances that can be used as constraints in modeling. A particularly useful crosslinking agent is the collision-induced dissociative crosslinker disuccinimidyl-succinamyl-aspartyl-proline (SuDP), which incorporates a gas-phase labile aspartyl-protol amide bond in the linker region [7]. SuDP crosslinks lysine residues with a maximum inter-residue distance of 23.9 Å, the maximal distance of C $\alpha$ –C $\alpha$  between the two crosslinked lysines. During LC/MS<sup>n</sup> analysis, the D–P bond of the crosslinked peptide (M<sub>1</sub>-SuDP-M<sub>2</sub>) is preferentially cleaved in MS<sup>2</sup> and allows the identification of each released peptide (M<sub>1</sub>-SuD and P-M<sub>2</sub>) by MS<sup>3</sup> sequencing [7–9]. In this study, we show the value of SuDP chemical crosslinking to correctly orient the independent N- and C-terminal domains of the transition state regulator protein AbrB. In this case, NMR NOE distance restraints were unable to provide any useful information, and the extra constraints provided by the crosslinking enabled the complete structure of the protein to be solved.

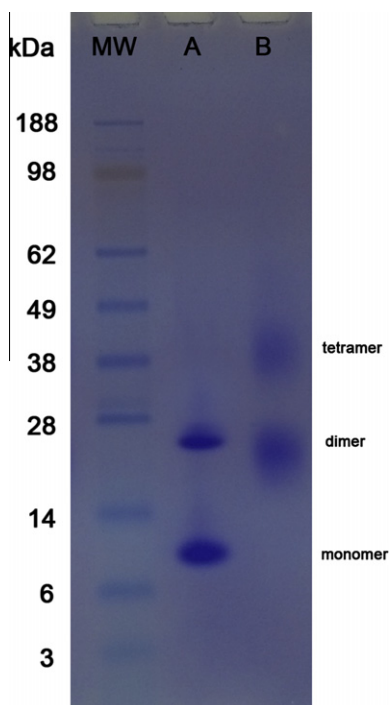
AbrB is a transcription factor that has homologues in a variety of organisms [10]. Its primary role is to regulate processes needed for both growth and survival during times of bacterial decision making. AbrB consists of 94 residues and is a functional homotetramer, consisting of two domains per monomer [11,12]. The DNA-binding N-terminal domain (AbrBN) has a looped-hinge helix fold and forms a homodimer with another AbrBN monomer [13–15,18]. The C-terminal domain (AbrBC) forms a homodimer with another

**Abbreviations:** SuDP, disuccinimidyl-succinamyl-aspartyl-proline; PRE, paramagnetic relaxation enhancement; FRET, Förster resonance energy transfer; NOE, nuclear overhauser effect; NMR, nuclear magnetic resonance; LC/MS<sup>n</sup>, multi-stage liquid chromatography–tandem mass spectrometry.

<sup>\*</sup> Corresponding author. Fax: +1 919 515 2047.

**E-mail addresses:** [alolson@ncsu.edu](mailto:alolson@ncsu.edu) (A.L. Olson), [fliu4@ncsu.edu](mailto:fliu4@ncsu.edu) (F. Liu), [attucker@ncsu.edu](mailto:attucker@ncsu.edu) (A.T. Tucker), [michael\\_goshe@ncsu.edu](mailto:michael_goshe@ncsu.edu) (M.B. Goshe), [john\\_cavanagh@ncsu.edu](mailto:john_cavanagh@ncsu.edu) (J. Cavanagh).

<sup>1</sup> These authors contributed equally to this work.



**Fig. 1.** SDS–PAGE analysis of Lane (A) un-crosslinked full length AbrB, and Lane (B) crosslinked AbrB with SuDP reagent at 1:100 M ratio.

AbrBC monomer, but is not involved directly in binding DNA. Obtaining the high-resolution structure of tetrameric AbrB has been extremely challenging due to the independent character of these domains. To facilitate the structure determination process, we solved the NMR structures of both AbrBN and AbrBC independently (pdb code 1ZOR and unpublished data, respectively) [18]. However, because of their disconnected nature, we have been unable to orient each domain with respect to one another using NOE approaches. As described below, SuDP chemical crosslinking and subsequent LC/MS<sup>n</sup> analysis were used to correctly position the domains within the full tetrameric AbrB protein structure.

## 2. Materials and methods

### 2.1. Materials

The SuDP reagent was synthesized as previously described [7–9,17]. Sequencing-grade modified trypsin was purchased from

Promega ([www.promega.com](http://www.promega.com)). Acetonitrile (HPLC grade) and formic acid (ACS reagent grade) were from Aldrich ([www.sigmaaldrich.com](http://www.sigmaaldrich.com)). Water was distilled and purified using a High-Q 103S water purification system ([www.high-q.com](http://www.high-q.com)). All other reagents and chemicals were purchased from Sigma–Aldrich–Fluka ([www.sigmaaldrich.com](http://www.sigmaaldrich.com)) unless otherwise stated.

### 2.2. Protein expression and purification

Full length *Bacillus subtilis* AbrB (residues 1–94) was expressed and purified as described previously [16]. Briefly, full length AbrB (residues 1–94) from *B. subtilis* was cloned into expression vector pET-28a (Novagen) with a Thrombin cleavable N-terminal histidine tag and transformed into BL21(DE3) cells (Genesee) for expression in LB media at 37 °C. At OD<sub>600</sub> of ~0.7 the temperature was reduced to 30 °C and expression induced with 1 mM IPTG. The cells were harvested by centrifugation 4 h postinduction at 7000g for 15 min. Cell pellets were suspended in lysis buffer (10 mM Tris–HCl, 300 mM KCl, 1 mM DTT, 5 mM imidazole, 0.02% sodium azide at pH 7.9) and sonicated with resulting cell lysate clarified by centrifugation at 15,000g and the resulting supernatant was passed over Ni–NTA agarose resin (Qiagen). The N-terminal histidine tag was removed from the purified protein by incubation with Thrombin and loaded onto a Q-Sepharose ion exchange column (GE Healthcare). Samples for crosslinking were dialyzed into working buffer (10 mM KH<sub>2</sub>PO<sub>4</sub>, 15 mM KCl, 1 mM DTT, 1 mM EDTA, 0.02% NaN<sub>3</sub>) at pH 7.5 and concentrated to 0.5 mg/mL.

### 2.3. Crosslinking and proteolytic digestion of AbrB

AbrB was crosslinked and digested as previously described [7]. Briefly, protein samples were prepared in PBS buffer and crosslinked with SuDP using a final protein-to-crosslinker ratio of 1:100. The crosslinked proteins were digested with trypsin (protein-to-trypsin ratio of 1:50, w/w) overnight at 37 °C.

### 2.4. Liquid chromatography–tandem mass spectrometry analysis

Crosslinked peptide samples were separated using an Easy-nLC system (<http://www.thermoscientific.com>) coupled online to an Orbitrap Elite mass spectrometer (<http://www.thermoscientific.com>). Reversed-phase separation of the peptides from the protein digest was accomplished using a 75 μm i.d. × 25 cm column packed in-house with 3 μm 200 Å Magic C18AQ stationary phase (Michrom Bioresources Inc., <http://www.michrom.com>) coupled to a Acclaim PepMap100 μm i.d. × 2 cm trap column (<http://www.thermoscientific.com>). The mobile phases consisted of (A) 0.1% formic acid in 2% acetonitrile, and (B) 0.1% formic acid in

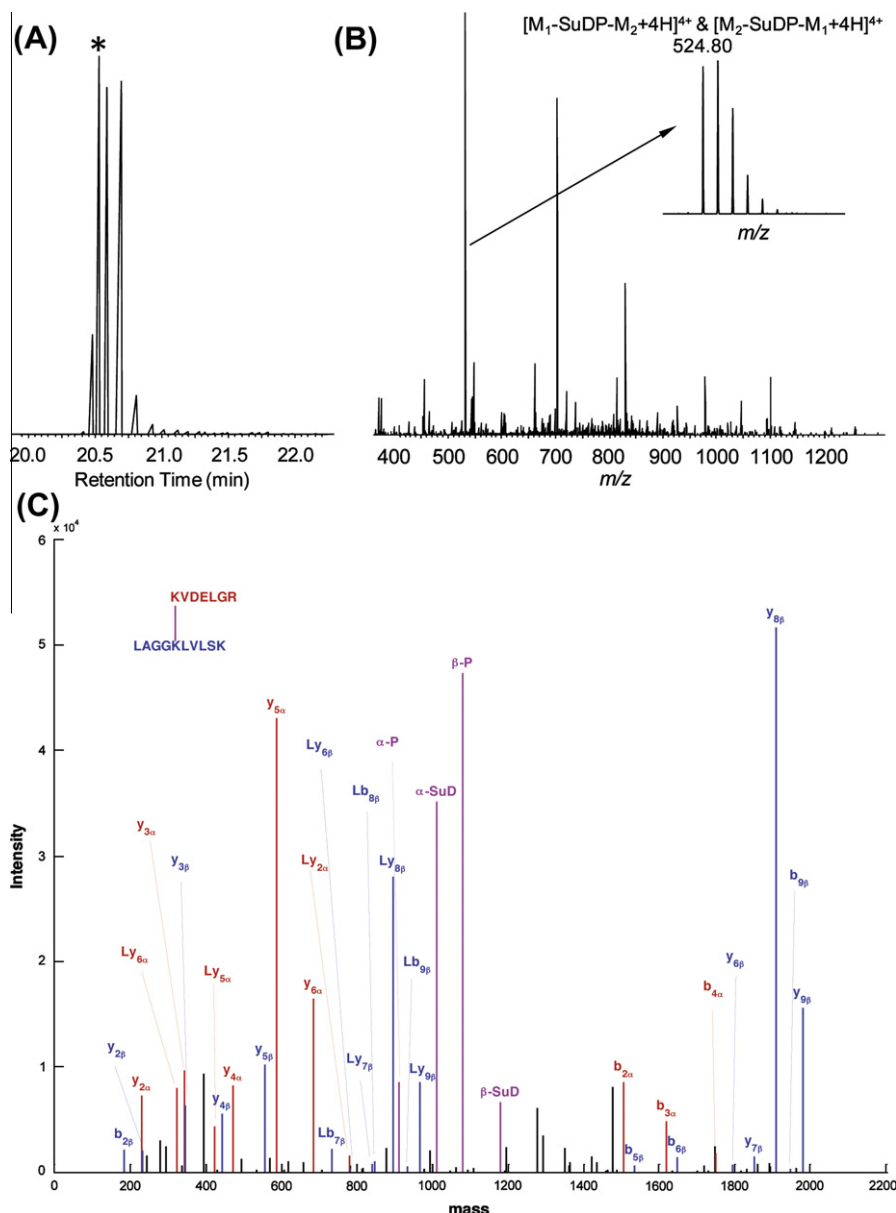
**Table 1**  
Interdomain crosslinks of AbrB by SuDP labeling.

Peptide 1	Mass calculated (Da)	Mass measured (Da)	Peptide 2	Mass Calculated (Da)	Mass Measured (Da)	Crosslink (Cα–Cα)
ILLK*K	613.4526	613.4543	LAGGK*LVLSK	984.6336	984.6361	K46–K71
K*VDELGR	815.4501	815.4518	LAGGK*LVLSK	984.6331	983.6276	K9–K71
GSHMK*STGIVR	1171.6131	815.4518	LAGGK*LVLSK	984.6331	984.6334	K2–K71
TLGIAEK*DALEIYVDDEK	2021.0204	2020.0180	LAGGK*LVLSK	984.6331	984.6338	K31–K71
YK*PNM#TC^QVTGEVSDDNLKLAGGK	2640.2522	2640.2711	LAGGK*LVLSK	984.6331	984.6341	K49–K71
YK*PNMTC^QVTGEVSDDNLKLAGGK	2624.2574	2624.2712	LVLSK*EGAEQIIEIQNLQNLK	2594.4278	2594.4534	K49–K76
TLGIAEK*DALEIYVDDEK	2021.0204	2021.0312	LVLSK*EGAEQIIEIQNLQNLK	2594.4278	2594.4534	K31–K76
ILLK*K	613.4526	613.4580	YKPNMTC^QVTGEVSDDNLK*LAGGK	2624.2574	2624.2593	K46–K66
K*VDELGR	815.4501	815.4513	YKPNMTC^QVTGEVSDDNLK*LAGGK	2624.2573	2623.2565	K9–K66
K*VDELGR	815.4501	815.4514	LVLSK*EGAEQIIEIQNLQNLK	2594.4278	2594.4491	K9–K76

\*Lysine involved in crosslinking.

^Carboxamidomethylation of Cys.

#Oxidation of Met.



**Fig. 2.** LC/MS<sup>n</sup> analysis of an AbrB inter-peptide crosslink KVDELGR-LAGGKLVLISK. (A) The extracted chromatograph of the crosslink precursor  $m/z$  524.80. The each peak maximum corresponds to the MS scan with the intervening time between them corresponding to the tandem MS scans. (B) The MS spectrum containing the crosslink precursor. The time point of the selected MS spectrum is indicated by the asterisk in panel (A). In this example, the crosslinked precursor is the most abundant ion in the spectrum, which is enlarged to illustrate its isotopic distribution. It is a +4 charged ion with a monoisotopic  $m/z$  value of 524.80, where  $M_1$  is KVDELGR (Peptide 1) and  $M_2$  is LAGGKLVLISK (Peptide 2) as listed in Table 1. (C) The MS<sup>2</sup> product ion spectrum of crosslink precursor. The product ion spectrum is reconstructed based on the deconvoluted neutral monoisotopic masses generated by CXLinkS using the raw data and is manually inspected to verify the crosslink identification. The product ions from peptide alpha KVDELGR ( $M_1$ ), peptide beta LAGGKLVLISK and the crosslinker fragmentation are shown in red, blue and magenta, respectively. The letter L indicates the remaining portion of the cleaved crosslinker with respect to each peptide fragment ion. Each individual linked peptide (magenta) is subjected to LC/MS<sup>3</sup>CID fragmentation for sequence identification (MS<sup>3</sup> data not shown). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acetonitrile. A total of 100 nmol of peptide digest was loaded onto the reversed-phase trap column for each injection and then separated using a linear gradient of 1% B/min from 5% B to 40% B at a flow rate of 300 nl/min.

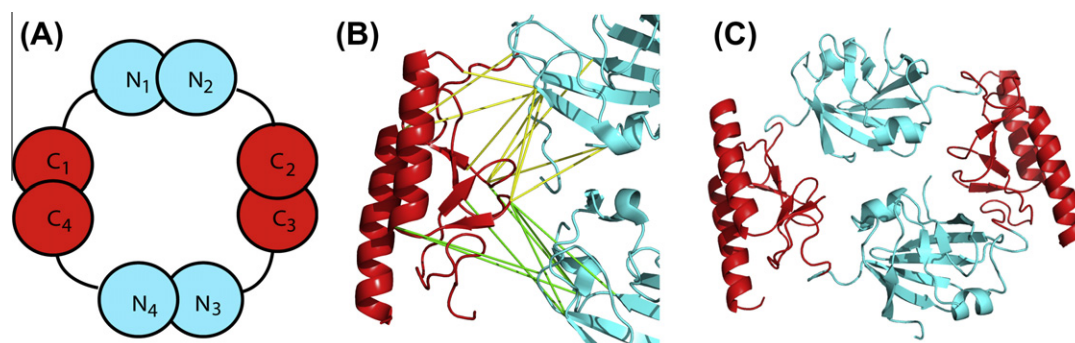
### 2.5. Two-step LC/MS<sup>n</sup> strategy and data analysis

LC/MS<sup>n</sup> data-dependent acquisition and data analysis were performed as previously described [9]. Briefly, an LC/MS<sup>2</sup> experiment was acquired in the orbitrap under low-energy collision induced dissociation (CID) to favor the preferential cleavage of our gas-phase cleavable crosslinker SuDP. Subsequently, an LC/MS<sup>3</sup>

analysis targeting each of the released peptides from the inter-peptide crosslinks was performed based on the inclusion list generated from the LC/MS<sup>2</sup> data via our in-house developed algorithm CXLinkS. The LC/MS<sup>3</sup> data were analyzed as described previously [9] with the database containing the AbrB sequence.

### 3. Results and discussion

Successful AbrB crosslinking with 1:100 M ratio of SuDP reagent were analyzed by SDS-PAGE and shown in Fig. 1. Crosslinking (Lane B) resulted in two crosslinked bands compared to un-crosslinked AbrB (Lane A). While un-crosslinked AbrB has two bands



**Fig. 3.** (A) AbrB tetramer with corresponding monomers and domains labeled. (B) Cartoon representation of crosslinks formed by C-terminal lysine residues (red) to lysine residues of N-terminus (cyan) of AbrB by SuDP reagent where yellow and green lines represent crosslinks to neighboring N-termini. (C) Assembled AbrB structure based on crosslinks detected in Table 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

consisting of monomer and dimer states (10.1 and 22 kDa, respectively), crosslinked AbrB is comprised of two bands at dimer and tetramer states (~22 and ~40 kDa, respectively). The smearing in the dimer and tetramer bands of crosslinked AbrB (Lane B) suggest highly crosslinked species with differences in molecular weight, protein shape, and net charge from the attachment of different numbers of crosslinkers.

A list of inter-domain crosslinks is presented in Table 1 and an example of an inter-peptide crosslink is provided in Fig. 2. In total, twenty-two crosslinked species were observed, including (i) nine crosslinks in a single AbrBN homodimer (combined intra- and inter-domain crosslinks), (ii) three crosslinks in a single AbrBC homodimer (combined intra- and inter-domain crosslinks), and (iii) ten inter-domain crosslinks spanning from AbrBN to AbrBC. Of the ten AbrBN–AbrBC crosslinks seven were used in organizing the domains with measured distances of 10 Å–22.1 Å (Table 1). The remaining three crosslinks (K2–K71, K49–K71, K49–K76) involve peptides of the highly unstructured region of AbrBN (K2) and the especially flexible domain linker (K49). Since those portions of the protein are very dynamic, these crosslinks could not be employed to reliably define distances.

Usable restraints were used in PyMOL to model AbrB by manually positioning the AbrBC NMR structure relative to AbrBN based on the maximum distance constraint of 23.9 Å using SuDP reagent (PyMOL Molecular Graphics System, Schrödinger, LLC). For each pair of crosslinked lysyl groups, there are eight possible crosslinks due to AbrB's tetrameric nature. However, four of the eight crosslinks can be eliminated since there are two faces to the N-terminal domain, with the one, more distant face, exceeding the crosslinking distance. This is illustrated schematically in Fig. 3(A). As shown, the C-terminal domain C1 can have crosslinks to residues in both N1 and N4 N-terminal domains, but crosslinks to N2 and N3 N-terminal domains would be disallowed. The remaining four allowed crosslinks are used to orient the AbrBC (red) to AbrBN (cyan) to satisfy the maximal distance constraints. The suite of crosslinks used between domains is shown in Fig. 3(B) and the subsequent fully assembled tetrameric AbrB structure is shown in Fig. 3(C).

In summary, we have shown that chemical crosslinking in conjunction with mass spectrometry can be used to accurately elucidate domain organization in a protein where the individual, independent domain structures are known, but cannot easily be connected by other means (e.g. X-ray crystallography and NMR). The crosslinking data facilitated spatial placement of the N- and C-terminal domains of AbrB based on the maximal distance constraints using SuDP crosslinking reagent. The long-range distance constraints provided by this method successfully enabled us to obtain a model of full tetrameric AbrB.

## Acknowledgments

This work was supported by grant NIH GM055769 (JC). We would like to thank the DOD DMRDP program (W81XWH-11-2-0115) for their support. The DMRDP program is administered by the Department of Army; The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick, MD 21702-5014 is the awarding and administering office. The content of this manuscript does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. Part of this work was supported by a Major Research Instrumentation grant from the National Science Foundation (DBI-1126244). We thank the research agencies of North Carolina State University and the North Carolina Agricultural Research Service for continued support of our bioanalytical mass spectrometry research.

## References

- [1] C.V. Robinson, A. Sali, W. Baumeister, The molecular sociology of the cell, *Nature* 450 (2007) 973–982.
- [2] Y. Yang, T.A. Ramelot, R.M. McCarrick, S. Ni, E.A. Feldmann, J.R. Cort, H. Wang, C. Ciccosanti, M. Jiang, H. Janjua, T.B. Acton, R. Xiao, J.K. Everett, G.T. Montelione, M.A. Kennedy, Combining NMR and EPR Methods for homodimer protein structure determination, *J. Am. Chem. Soc.* 132 (2010) 11910–11913.
- [3] T. Ha, Single-molecule fluorescence resonance energy transfer, *Methods* 25 (2001) 78–86.
- [4] W.L. Hubbell, D.S. Cafiso, C. Altenbach, Identifying conformational changes with site-directed spin labeling, *Nat. Struct. Mol. Biol.* 7 (2000) 735–739.
- [5] O. Rinner, J. Seebacher, T. Walzthoeni, L. Mueller, M. Beck, A. Schmidt, M. Mueller, R. Aebersold, Identification of cross-linked peptides from large sequence databases, *Nat. Methods* 5 (2008) 315–318.
- [6] A. Sinz, Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions, *Mass Spectrom. Rev.* 25 (2006) 663–682.
- [7] E.J. Soderblom, B.G. Bobay, J. Cavanagh, M.B. Goshe, Tandem mass spectrometry acquisition approaches to enhance identification of protein–protein interactions using low-energy collision-induced dissociative chemical crosslinking reagents, *Rapid Commun. Mass Spectrom.* 21 (2007) 3395–3408.
- [8] F. Liu, M.B. Goshe, Combinatorial electrostatic collision-induced dissociative chemical cross-linking reagents for probing protein surface topology, *Anal. Chem.* 82 (2010) 6215–6223.
- [9] F. Liu, C. Wu, J.V. Sweedler, M.B. Goshe, An enhanced protein crosslink identification strategy using CID-cleavable chemical crosslinkers and LC/MS<sup>n</sup> analysis, *Proteomics* 12 (2012) 401–405.
- [10] M.A. Strauch, P. Ballar, A.J. Rowshan, K.L. Zoller, The DNA-binding specificity of the *Bacillus anthracis* AbrB protein, *Microbiology* 151 (2005) 1751–1759.
- [11] L.M. Benson, J.L. Vaughn, M.A. Strauch, B.G. Bobay, R. Thompson, S. Naylor, J. Cavanagh, Macromolecular assembly of the transition state regulator AbrB in its unbound and complexed states probed by microelectrospray ionization mass spectrometry, *Anal. Biochem.* 306 (2002) 222–227.
- [12] Z.E.V. Phillips, M.A. Strauch, Role of Cys54 in AbrB multimerization and DNA-binding activity, *FEMS Microbiol. Lett.* 203 (2001) 207–210.
- [13] B.G. Bobay, L. Benson, S. Naylor, B. Feeney, A.C. Clark, M.B. Goshe, M.A. Strauch, R. Thompson, J. Cavanagh, Evaluation of the DNA binding tendencies of the transition state regulator AbrB, *Biochemistry* 43 (2004) 16106–16118.

- [14] B.G. Bobay, A. Andreeva, G.A. Mueller, J. Cavanagh, A.G. Murzin, Revised structure of the AbrB N-terminal domain unifies a diverse superfamily of putative DNA-binding proteins, *FEBS Lett.* 579 (2005) 5669–5674.
- [15] M. Coles, S. Djuranovic, J. Söding, T. Frickey, K. Koretke, V. Truffault, J. Martin, A.N. Lupas, AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi  $\beta$  barrels, *Structure* 13 (2005) 919–928.
- [16] A.L. Olson, B. Bobay, C. Melander, J. Cavanagh,  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignments and secondary structure prediction of the full-length transition state regulator AbrB from *Bacillus anthracis*, *Biomol. NMR Assign.* 6 (2012) 95–98.
- [17] E.J. Soderblom, M.B. Goshe, Collision-induced dissociative chemical cross-linking reagents and methodology: applications to protein structural characterization using tandem mass spectrometry analysis, *Anal. Chem.* 16 (2006) 8059–8068.
- [18] D.M. Sullivan, B.G. Bobay, D.J. Kojetin, R.J. Thompson, M. Rance, M.A. Strauch, J. Cavanagh, Insights into the nature of DNA binding of AbrB-like transcription factors, *Structure* 16 (2008) 1702–1713.