Title

Novel structure of the N-terminal helical domain of BibA, a group B streptococcus immunogenic bacterial adhesin.

Journal

Acta Crystallographica Section D: Structural Biology, 76(Pt 8)

Authors

Manne, Kartik
Chattopadhyay, Debasish
Agarwal, Vaibhav
et al.

Publication Date

2020-08-01

DOI

10.1107/S2059798320008116

Peer reviewed

STRUCTURAL BIOLOGY
ISSN 2059-7983

Novel structure of the N-terminal helical domain of BibA, a group B streptococcus immunogenic bacterial adhesin

Kartik Manne,[a] Debasish Chattopadhyay,[b] Vaibhav Agarwal,[c] Anna M. Blom,[c] Baldeep Khare,[d] Srinivas Chakravarthy,[e] Chungyu Chang,[f] Hung Ton-That[f] and Sthanam V. L. Narayana[a]*

[a]Center for Biophysical Sciences and Engineering, University of Alabama at Birmingham, Birningham, AL 35294, USA, [b]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA, [c]Department of Translational Medicine, Lund University, S-214 28 Malmö, Sweden, [d]Department of Biological Sciences, Purdue University, West Lafayette, Indiana, USA, [e]The Biophysics Collaborative Access Team (BioCAT), Department of Biological Sciences, Illinois Institute of Technology, Chicago, IL 60616, USA, and [f]Division of Oral Biology and Medicine, School of Dentistry, University of California Los Angeles, Los Angeles, California, USA. *Correspondence e-mail: narayana@uab.edu

Received 24 February 2020
Accepted 17 June 2020

Edited by Z. S. Derewenda, University of Virginia, USA

Keywords: BibA; group B streptococcus; immunogenic bacterial adhesins; three-helix-bundle-motif repeats; C4b-binding proteins.

PDB reference: BibA, 6poo

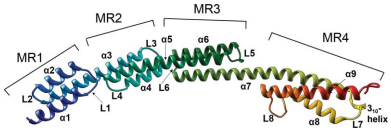
Supporting information: this article has supporting information at journals.iucr.org/d

BibA, a group B streptococcus (GBS) surface protein, has been shown to protect the pathogen from phagocytic killing by sequestering a complement inhibitor: C4b-binding protein (C4BP). Here, the X-ray crystallographic structure of a GBS BibA fragment ($BibA_{126–398}$) and a low-resolution small-angle X-ray scattering (SAXS) structure of the full-length N-terminal domain ($BibA_{34–400}$) are described. The $BibA_{126–398}$ fragment crystal structure displayed a novel and predominantly helical structure. The tertiary arrangement of helices forms four antiparallel three-helix-bundle-motif repeats, with one long helix from a bundle extending into the next. Multiple mutations on recombinant $BibA_{34–400}$ delayed the degradation of the protein, and circular dichroism spectroscopy of $BibA_{34–400}$ suggested a similar secondary-structure composition to that observed in the crystallized $BibA_{126–398}$ fragment. A model was generated for the 92 N-terminal residues ($BibA_{34–125}$) using structural similarity prediction programs, and a $BibA_{34–400}$ model was generated by combining the coordinates of $BibA_{34–126}$ and $BibA_{126–398}$. The X-ray structure of $BibA_{126–398}$ and the model of $BibA_{34–400}$ fitted well into the calculated SAXS envelope. One possible binding site for the BibA N-terminal domain was localized to the N-terminal CCP (complement-control protein) domains of the C4BP α-chain, as indicated by the decreased binding of BibA to a ΔCCP1 C4BP α-chain mutant. In summary, it is suggested that the GBS surface protein BibA, which consists of three antiparallel α-helical-bundle motifs, is unique and belongs to a new class of Gram-positive surface adhesins.

1. Introduction

Group B streptococcus (GBS) causes sepsis and meningitis in neonates (Dermer *et al.*, 2004), and is the second most common pathogen infecting infants (the first being *Escherichia coli*) during their first week of life (Ladhani *et al.*, 2019). A worldwide meta-analysis estimates that GBS affects infants at a rate of 0.3–2 per 1000 lives (with higher incidences in Africa compared with those in other regions), with early-onset infections occurring more frequently than late-onset infections (Madrid *et al.*, 2017). GBS is also responsible for infections in pregnant and peripartum women (Krohn *et al.*, 1999) and is increasingly being recognized to cause severe infections among adults with diabetes, cancer and other chronic illnesses (Farley, 2001; Sendi *et al.*, 2008; Farley *et al.*, 1993).

Human complement, which eliminates microorganisms and other antigens from blood and tissues, plays a critical role in the host defense against pathogens (Wessels *et al.*, 1995; Jarva



et al., 2003). Many enzymes, cofactors, regulatory proteins and receptors constitute complement, and they act in concert to induce inflammatory and cytocidal reactions in response to invading microbes. The activation of complement, which is essential for its biological activity, is mediated through three pathways: the classical pathway, which is activated primarily by antibody–antigen complexes, the alternative pathway, which is triggered directly by pathogens, and the lectin pathway, which is initiated by carbohydrates on the surface of microorganisms. The complement plays an important role in host defense against GBS: while the classical pathway is implicated in early bloodstream clearance of GBS (Butko *et al.*, 1999), it has been shown that antibody-mediated immunity requires C3 and not C4, thus also suggesting a role for the alternative pathway in opsonophagocytosis of GBS (Wessels *et al.*, 1995).

Santi *et al.* (2007) reported BibA to be a surface protein that was expressed in all 24 GBS strains that they analyzed, and its role in GBS adhesion was first demonstrated by the impaired ability of a *bibA*-knockout mutant strain to adhere to both cervical and lung epithelial cells. In a complement-mediated opsonophagocytic assay, 50% of mice infected with the wild-type 2603/VR strain died within three days, while only 10% of mice infected with a *bibA*-mutant 2603/VR strain died in the same period, confirming the role of BibA in GBS virulence. More importantly, when tested for the ability of GBS to replicate in human blood, the *bibA*-mutant 2603/VR strain proliferated 80% less efficiently compared with the wild type, suggesting that the survival of GBS is better in human blood when BibA is surface-expressed than when BibA is absent (Santi *et al.*, 2007).

BibA is a modular protein that is predicted to have a helix-rich N-terminal domain (34–399) followed by a proline-rich C-terminal region (400–543) with a canonical LP*X*TG cell-wall-anchoring motif LPSTG at the C-terminus (Fig. 1*a*). The sequence of BibA is highly conserved in all clinical isolates, with the N-terminal domain being identical and the proline-rich region showing 98–100% similarity (Santi, Maione *et al.*, 2009). The putative N-terminal helical domain protects GBS from phagocytic killing by attaching to the complement inhibitor C4b-binding protein (C4BP; Santi *et al.*, 2007; Pietrocola *et al.*, 2018). BibA protein was shown to be highly expressed on the GBS surface at pH 7.0 and to a lower extent at pH 5.5 (Santi, Grifantini *et al.*, 2009), and such pH-regulated BibA expression is fully reversible upon exposure to a new pH environment. This also explains the induced opsonophagocytic killing of GBS by anti-BibA antibodies at pH 7.0, while no killing activity was observed against bacteria grown at pH 5.5 (Santi, Grifantini *et al.*, 2009). Mice immunized with BibA were shown to be protected against the challenge of GBS, strengthening the argument for the beneficial effects of vaccination using BibA (Santi, Maione *et al.*, 2009; Senn *et al.*, 2011).

The complement inhibitor C4BP is a spider-shaped molecule (Fig. 1*b*). The main isoform of C4BP contains seven identical α-chains and a unique β-chain. The β-chain-containing C4BP in circulation is bound to the vitamin-K-dependent anticoagulant Protein S (PS), forming a C4BP–PS complex. A less abundant form of C4BP, which is upregulated upon inflammation, is composed exclusively of the α-chains (Fig. 1*c*). Each α-chain is made of eight CCP (complement-control protein) domains and has multiple binding sites for C4b. The binding interface of C4BP for C4b is contributed by a cluster of positively charged amino acids present in the linker between CCP1 and CCP2 (Leung *et al.*, 2006; Blom *et al.*, 1999; Blom, Zadura *et al.*, 2000; Villoutreix *et al.*, 1998; Blom, 2002). This same region of C4BP is also crucial for binding C3b (Blom *et al.*, 2003), heparin (Blom *et al.*, 1999) and several bacterial pathogens such as *Haemophilus influenzae* (Hallström *et al.*, 2007), the gingipain HRgpA from *Porphyromonas gingivalis* (Potempa *et al.*, 2008), the *Moraxella catarrhalis* UspA1/2 proteins (Nordström *et al.*, 2004) and the

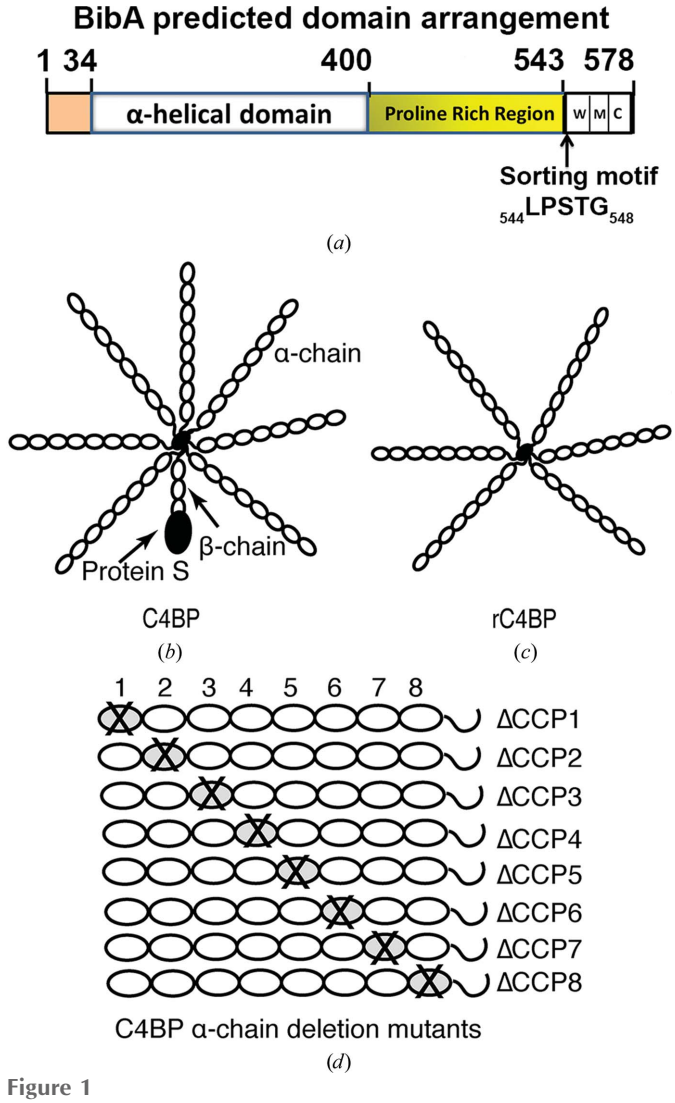


Figure 1
(*a*) Predicted domain arrangement for the GBS surface protein BibA. The C-terminal sorting motif is shown. Different C4BP variants were used: plasma-purified C4BP (*b*), rC4BP (*c*) and C4BP α-chain deletion mutants (*d*). The main isoform of C4BP contains seven identical α-chains and one unique β-chain. The β-chain-containing C4BP in circulation is bound to the vitamin-K-dependent anticoagulant Protein S (PS), forming a C4BP–PS complex. The rC4BP contains six α-chains and lacks the β-chain and the associated PS. C4BP α-chain deletion mutants lack single CCP domains (represented by gray circles with an X in each α-chain).

Table 1
Recombinant production information.

Source organism	*Streptococcus agalactiae*
Forward primer	TACTTCCAATCCAATGCACACGCGGATACTAGTTCAGG
Reverse primer	TTATCCACTTCCAATGTTATTTTTTATCTTCTTTAGTACCTCT
Forward K49N mutation primer	[Phos]GGAATATCGGCTTCAATTCCTCATAAGAACCAAGTTAATTTAGGGGCGGTTACT
Reverse K70N mutation primer	[Phos]TACTCTACTTAAAAGTATAGCAATAGCGTTGTCATTACCACGATATTTAGAAAT
Forward K112N mutation primer	[Phos]ATTTTATATCAAGGACAAATTGGTAACCAAAATAAACCAAGTGTAACTACA
Cloning vector	pMCSG7
Expression vector	pMCSG7
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced ($BibA_{34–400}$)†	*MHHHHHH*SSGVDLGTENLYFQSHADTSSGISASIPHKKQVNLGAVTLKNLISKYRGNDKAI AILLSRVNDFNRASQDTLPQLINSTEAEIRNILYQGQIGKQNKPSVTTHAKVSDQELGKQ SRRSQDIIKSLGFLSSDQKDILVKSISSSKDSQLILKFVTQATQLNNAESTKAKQMAQND VALIKNISPEVLEEYKEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQANGKKQEIAKL ENLSNDEMLRYNTAIDNVVKQYNEGKLNITAAMNALNSIKQAAQEVAQKNLQKQYAKKIE RISSKGLALSKKAKEIYEKHKSILPTPGYYADSVGTYLNRFRDKQTFGNRSVWTGQSGLD EAKKMLDEVKKLLKELQDLTRGTKEDKK

† Mutated residues are underlined and the His_6 tag is shown in italics.

Streptococcus pyogenes surface M protein (Blom, Berggård *et al.*, 2000; Buffalo *et al.*, 2016). An NMR structure of $C4BP_{CCP1–2}$ (Jenkins *et al.*, 2006) and the crystal structures of the hypervariable regions of four different M proteins (M2, M49, M22 and M28) in complex with $C4BP_{CCP1–2}$ (Buffalo *et al.*, 2016) are available and have revealed the significant contributions of residues Arg39, Arg64, Arg66 and His67 towards the binding of *S. pyogenes* M proteins to $C4BP_{CCP1–2}$. The N-terminal region of BibA has previously been implicated in its association with C4BP (Santi *et al.*, 2007).

A structural view of the association of BibA and C4BP could help in developing strategies for the prevention and treatment of GBS infections. Towards this goal, we attempted to determine the crystal structure of the N-terminal domain of BibA ($BibA_{34–400}$) in complex with $C4BP_{CCP1–2}$. However, we could only obtain crystals of a proteolytic fragment of BibA ($BibA_{126–398}$), and its structure was determined using X-ray crystallographic techniques. $BibA_{126–398}$ displays a novel helical structure that forms four antiparallel three-helix-bundle-motif repeats. Further, we performed a structural investigation into the intact $BibA_{34–400}$ protein using circular dichroism (CD) spectrometry and small-angle X-ray scattering (SAXS). Based on the above analyses, we propose BibA ($BibA_{34–400}$) to be a distorted rod-shaped molecule. We also analyzed the molecular basis of the interactions between BibA and recombinant rC4BP (α-chain) by performing surface plasmon resonance analysis and localized the binding sites for $BibA_{34–400}$ to the C4BP α-chain.

2. Materials and methods

2.1. Cloning and mutagenesis

The *bibA* gene from group B streptococcus strain 2603 (serotype V/R) was used as a template to express the recombinant forms of BibA. The gene segment corresponding to the N-terminal domain of BibA encompassing amino-acid residues 34–400, labeled $BibA_{34–400}$, was amplified by PCR with the respective primers (Table 1). The products were cloned into vector pMCSG7 (DNASU plasmid repository, USA), which harbors an N-terminal histidine tag and TEV protease cleavage site, using ligation-independent cloning (LIC). The Change-IT Multiple Mutation Site-Directed Mutagenesis Kit (USB Affymetrix, USA; Chen & Ruffner, 1998) was used to generate the mutants with their respective mutagenic primers (Table 1). The resulting constructs and mutations were confirmed by DNA sequencing.

2.2. Expression and purification of recombinant proteins

The recombinant plasmids were transformed into *E. coli* BL21(DE3) cells (Novagen, USA) for expression. An overnight seed culture was used to inoculate 1 l Luria–Bertani (LB) broth supplemented with 50 µg ml^{-1} ampicillin. The cells were grown at 37°C until the $OD_{600\ nm}$ reached ∼0.7, at which point protein expression was induced overnight by the addition of 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation at 3210*g* for 30 min at 4°C using a Beckman Allegra 6R centrifuge. The cell pellets were resuspended in lysis buffer consisting of 50 m*M* Tris–HCl pH 7.0, 200 m*M* NaCl, 5%(*v*/*v*) glycerol, 100 m*M* L-arginine and an EDTA-free protease-inhibitor cocktail tablet (Roche, USA). The bacterial cells were lysed by the freeze–thaw method and 1 mg DNase I (New England BioLabs, USA) was added to remove unwanted DNA from the cell lysate to improve the protein-extraction efficiency. The lysate was cleared by centrifugation at 48 384*g* for 1 h at 4°C using a Beckman Avanti J-25 centrifuge. The cleared lysate was passed through a 5 ml pre-packed HisTrap FF column (GE Healthcare, USA) pre-equilibrated with lysis buffer at a flow rate of 1 ml min^{-1}. The protein-bound column was washed with ten bed volumes of Ni column buffer *A* (the same as the lysis buffer) and ten bed volumes of buffer *B* [50 m*M* Tris–HCl pH 7.0, 200 m*M* NaCl, 5%(*v*/*v*) glycerol, 100 m*M* L-arginine, 250 m*M* imidazole] at a 15% concentration. Bound protein was eluted with a linear gradient of 15–100% buffer *B*. Eluted fractions with the protein of interest were pooled, treated with TEV protease to remove the N-terminal His tag and concentrated using an Amicon ultracentrifugation system to a final volume of 1 ml. This sample was further purified by

gel-filtration chromatography (HiPrep 16/60 Sephacryl S-100 HR column; GE Healthcare) in 20 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl, 10%(*v*/*v*) glycerol, 25 m*M* L-arginine, 5 m*M* EDTA, 5 m*M* β-mercaptoethanol (SEC buffer).

The substitution of $BibA_{34–400}$ with selenomethionine (SeMet) and its expression were carried out by inoculating a single colony carrying pMCSG7-$BibA_{34–400}$ into 50 ml minimal medium (Molecular Dimensions, USA) supplemented with 50 µg ml^{-1} ampicillin and incubating overnight at 37°C. The cells were harvested by centrifugation at 3210*g* for 30 min at 4°C and were washed thoroughly with distilled water before resuspending them in 1 ml water. The resuspended cell pellets were decanted into 1 l minimal medium containing L-selenomethionine, which had been prewarmed to 37°C, and were allowed to grow until the $OD_{600\ nm}$ reached ∼0.5. At this point, the medium was cooled to 30°C and protein expression was induced by the addition of 1 m*M* IPTG when the $OD_{600\ nm}$ reached ∼0.7; the culture was then grown overnight. The cells were harvested and lysed and SeMet-substituted $BibA_{34–400}$ was purified using a method similar to that used for the native protein. For both $BibA_{34–400}$ proteins (native and SeMet-substituted) the purity was checked by 12% SDS–PAGE (Laemmli, 1970), and their concentrations were determined by measuring their UV absorbance at 280 nm using an ε value of 21 890 M^{-1} cm^{-1}.

The main isoform of C4BP (570 kDa) containing α- and β-chains as well as Protein S (Fig. 1*b*) was purified from human plasma as described in Dahlbäck (1983), while recombinant wild-type C4BP composed of only α-chains (Fig. 1*c*) and rC4BP mutants lacking single CCP domains in the α-chains (Fig. 1*d*) were expressed in eukaryotic cells and purified by affinity chromatography as described previously (Blom *et al.*, 2001).

2.3. Protein crystallization and X-ray diffraction data collection

The purified $BibA_{34–400}$ protein was concentrated with an Amicon ultracentrifugation system (molecular-weight cutoff 30 kDa). The JCSG Core Suites I and II (Qiagen), The PEGs Suites I and II (Qiagen) and Crystal Screen HT (Hampton Research, USA) crystallization kits were used in initial crystallization screens. Robotic screening was performed with a Crystal Phoenix Robot (Art Robbins Instruments, USA) in a 96-well sitting-drop plate (Corning 3550) at 20°C. Optimization of the conditions to produce good-quality single crystals was performed by hand in 24-well hanging-drop plates (VDX plates; Hampton Research). The final buffer composition and the reservoir solution were as described in Table 2. X-ray diffraction data were collected at 100 K using 20%(*v*/*v*) glycerol as a cryoprotectant on the SER-CAT beamline 22-ID at the Advanced Photon Source (APS), Argonne National Laboratory, USA.

2.4. Structure determination and refinement

Data collected from SeMet $BibA_{126–400}$ crystals at the selenium edge and peak were processed using *XDS* (Kabsch, 2010) and *iMosflm* (Battye *et al.*, 2011). The crystal structure was solved by multi-wavelength anomalous dispersion (MAD) phasing to a resolution of 3.03 Å using *CRANK*2 (Skubák & Pannu, 2013), which is integrated into the *CCP*4 suite (Winn *et al.*, 2011). Model building and refinement were performed using *Phenix* (Liebschner *et al.*, 2019) and *Coot* (Emsley *et al.*, 2010). The final refined model consisting of residues Asp126–sp398 converged with R_{work} and R_{free} values of 22.7% and 26.2%, respectively, and showed an excellent Ramachandran plot with no residues in the disallowed regions.

Table 2
Crystallization information.

Method	Hanging-drop vapor diffusion
Plate type	24-well plate (VDX plate, Hampton Research)
Temperature (K)	295
Protein concentration (mg ml^{-1})	40.8
Composition of protein solution	20 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl, 10%(*v*/*v*) glycerol, 25 m*M* L-arginine, 5 m*M* EDTA, 5 m*M* β-mercaptoethanol
Composition of reservoir solution	25%(*w*/*v*) PEG 3350, 0.1 *M* bis-Tris pH 5.0, 0.2 *M* magnesium chloride
Volume and ratio of drop	2.2 µl, 1:1 (+0.2 µl crystal seed)
Volume of reservoir (ml)	1

2.5. Circular dichroism measurements

$BibA_{34–400}$ protein in SEC buffer was diluted eightfold with 10 m*M* Tris–HCl pH 7.5, and the CD spectra were measured in a quartz cell with an optical path length of 0.1 cm on a JASCO J-815 Circular Dichroism Spectrometer (JASCO, USA). The final concentration of the protein was 0.15 mg ml^{-1} and the spectra were obtained from three scans measured from 240 to 190 nm. The initial two scans were made at 20°C and 40°C, and the sample was then denatured at 80°C and cooled slowly to 20°C. The third scan was recorded from this cooled sample. The CD spectrum of the SEC buffer (blank) was measured in the same manner and was subtracted from the collected data. The input units of measured ellipticity (mdeg) were converted to $\Delta\varepsilon$ and the secondary-structure analysis was performed using the *BeStSel* web server (Micsonai *et al.*, 2015; Fig. 2). The normalized root-mean-square deviation (NRMSD) values of all the samples were 0.01 or less. The atomic coordinates from the $BibA_{126–400}$ crystal structure were used with the SP175 (soluble) reference database to generate CD spectra (Fig. 2*a*) using an empirical-based approach with the online *PDB2CD* web server (Mavridis & Janes, 2017). The $BibA_{126–400}$ CD spectra were used as input for the *BeStSel* web server for comparison with the secondary-structure analysis performed with the CD spectra of the $BibA_{34–400}$ protein (Fig. 2*b*).

2.6. Small-angle X-ray scattering (SAXS) measurements and data analysis

SAXS data sets were collected for $BibA_{34–400}$ at BioCAT (beamline 18ID at the APS) using inline size-exclusion chromatography (SEC-SAXS). This method ensures optimal sample quality by separating the sample from aggregates and other contaminants. The sample was dialyzed against 25 m*M*

Tris pH 8.0, 75 m*M* NaCl, 10%(*v*/*v*) glycerol and concentrated to 16 mg ml^{-1} before loading 700 µl onto a Superdex 200 Increase 10/300 GL column. The sample was run at 0.7 ml min^{-1} on an ÄKTA pure FPLC (GE Healthcare Life Sciences) and the eluate passed through the UV monitor flowed through the SAXS flow cell, which is modeled after a sheath-flow cell (the coflow method; Kirby *et al.*, 2016). Scattering intensity was recorded using a PILATUS3 1M detector (Dectris), and a q-range of ∼0.004 to ∼0.36 Å^{-1} was accessed by placing the detector 3.64 m from the sample. Data sets were acquired with 0.5 s exposures at 2 s intervals during elution and data were reduced using *BioXTAS RAW* 1.6.0 (Hopkins *et al.*, 2017). Buffer blanks were created by averaging regions flanking the elution peak and were subtracted from exposures selected from the elution peak to create the $I(q)$ versus q curves used in subsequent analyses. Evolving factor analysis was used to extract SAXS curves for single components (Hopkins *et al.*, 2017). Guinier fit and molecular-weight analysis was performed using *BioXTAS RAW* 1.6.0 (Hopkins *et al.*, 2017) and $P(r)$ distributions were calculated using *GNOM* (Svergun, 1992) in *ATSAS* (Franke *et al.*, 2017) through the *BioXTAS RAW* interface.

2.7. *Ab initio* and ensemble modeling

Ab initio bead models (also known as dummy-atom models) were calculated using *DAMMIF* (Franke & Svergun, 2009). The bead models were averaged by *DAMAVER* (Volkov & Svergun, 2003) to build a consensus shape and produce a final model. The crystallographic structure of BibA$_{126–398}$ was fitted into the low-resolution SAXS model using *SUPCOMB* (Kozin & Svergun, 2001) and *UCSF Chimera* (Pettersen *et al.*, 2004). An initial model for the missing N-terminal residues was generated with the help of the MR1 and MR2 motif repeats of the BibA$_{126–398}$ crystal structure, which were used as a template in the *I-TASSER* online structure-prediction server (Zhang, 2008) to generate the full-length BibA$_{34–400}$ model. The final model was idealized and refined using *Coot*. This final BibA$_{34–400}$ model was fitted into the SAXS envelope using *UCSF Chimera*. Theoretical profiles for the atomic coordinates of the full-length BibA$_{34–400}$ model were calculated using *FoXS* (Schneidman-Duhovny *et al.*, 2013, 2016), which is available through *UCSF Chimera*.

2.8. Antibodies

The polyclonal rabbit anti-human C4BP 9008 was a kind gift from Professor B. Dahlbäck (Lund University, Sweden). The 9008 antibody recognizes all of the forms of C4BP (plasma purified C4BP–Protein S complex, rC4BP and mutants) equally well (Avirutnan *et al.*, 2011). Antibodies were purified using a Protein A Sepharose column. Peroxidase-conjugated swine anti-rabbit and rabbit anti-mouse IgG were obtained

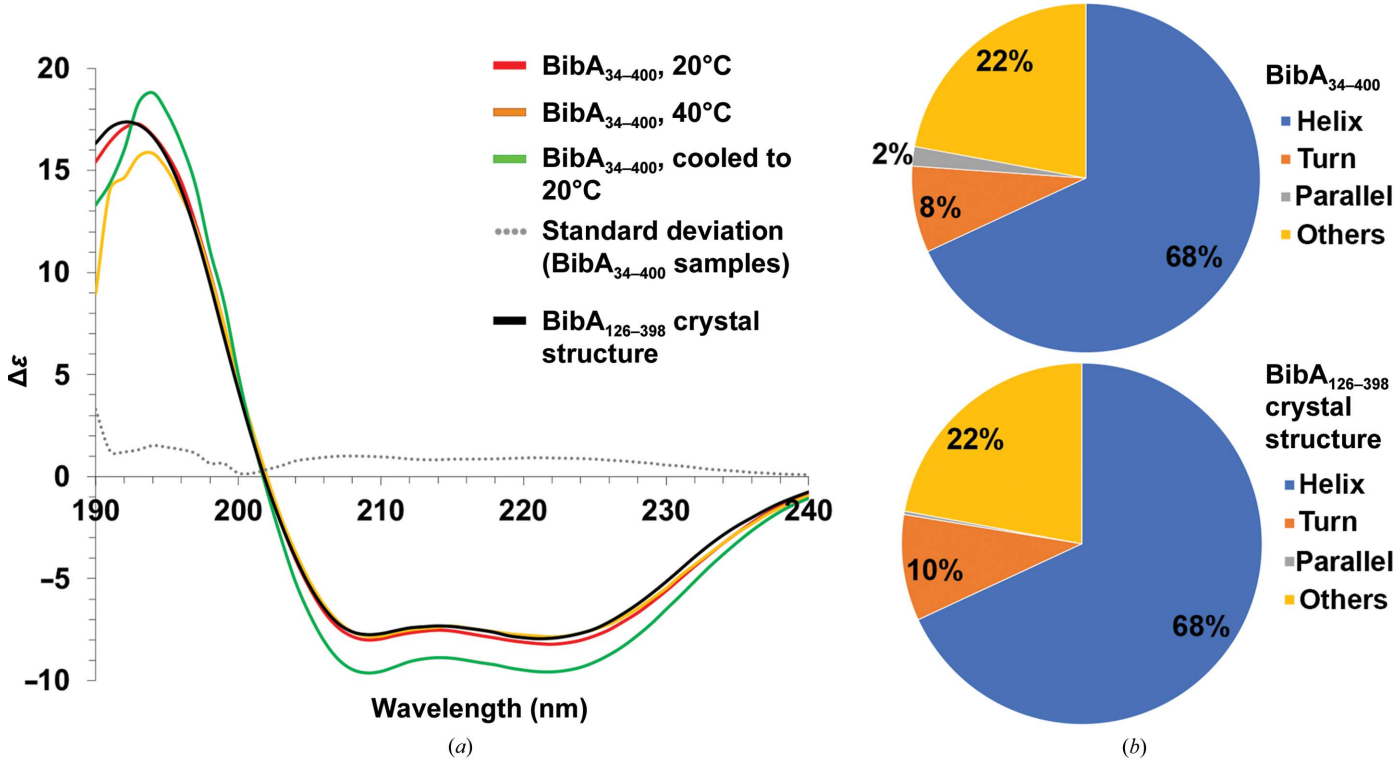


Figure 2
(*a*) Circular dichroism spectra of freshly prepared and intact BibA$_{34–400}$ sample at 20°C (red), at 40°C (orange) and denatured at 80°C followed by cooling to 20°C (green), the standard deviation error between the BibA$_{34–400}$ samples (gray dotted line) and the spectrum generated from the BibA$_{126–398}$ crystal structure (black). Two negative peaks at 208 and 222 nm typical of a α-helical secondary structure were observed for the BibA$_{34–400}$ sample. (*b*) Estimated secondary-structure content (%) of BibA$_{34–400}$ sample in solution (top) and the spectrum generated from the BibA$_{126–398}$ crystal structure (bottom).

from Dako (USA) and the bovine serum albumin (BSA) used as a negative control was purchased from AppliChem (USA).

2.9. Direct binding assays

Microtiter plates were coated with BibA (5 µg ml^{-1}) in PBS overnight at 4°C. Wells coated with BSA were used as a control. The plates were washed with 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 0.1% Tween-20 between each step. The plates were then blocked with 250 µl blocking solution [50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 0.1% Tween-20, 3% fish gelatin (Norland)] for 2 h at room temperature. After blocking, the plate was incubated with increasing concentrations of plasma-purified C4BP–PS or rC4BP for binding to BibA in binding buffer (50 m*M* HEPES pH 7.4, 150 m*M* NaCl, 2 m*M* $CaCl_2$, 50 µg ml^{-1} BSA) and was incubated for 90 min at room temperature. To assess the binding of various forms of C4BP (wild-type rC4BP and mutants lacking individual domains) to BibA, 30 µg ml^{-1} of each C4BP variant was incubated overnight at 4°C with BibA immobilized at 5 µg ml^{-1}. Bound C4BP was detected using a specific polyclonal Abs (9008), followed by a swine anti-rabbit peroxidase-conjugated Abs. The plates were developed with *o*-phenylenediamine (Dako, USA) substrate and H_2O_2, and the absorbance at 490 nm was measured (Varian Cary 50 MPR Microplate Reader).

2.10. Surface plasmon resonance analysis

To measure the kinetics of rC4BP binding to BibA, surface plasmon resonance analysis was performed using a Biacore 2000 (GE Healthcare, USA). rC4BP was diluted to 10 µg ml^{-1} in 10 m*M* sodium acetate pH 4.0 and was immobilized on the surface of a CM5 sensor chip to reach 3000 response units (RU). All experiments were performed at a continuous flow rate of 30 µl min^{-1} using Biacore buffer (150 m*M* NaCl, 10 m*M* HEPES, 2.5 m*M* $CaCl_2$, 0.002% Tween-20 pH 7.4). The analyte, $BibA_{34–400}$, was injected in a concentration gradient from 1.75 to 56 n*M* followed by two consecutive injections of 2 *M* NaCl and 100 m*M* HCl followed by 4 *M* $MgCl_2$ and 0.05% SDS for regeneration. The obtained sensorgrams were analyzed using the *Bio-evaluation* software version 3.0 (Biacore) using a 1:1 Langmuir binding model of interaction.

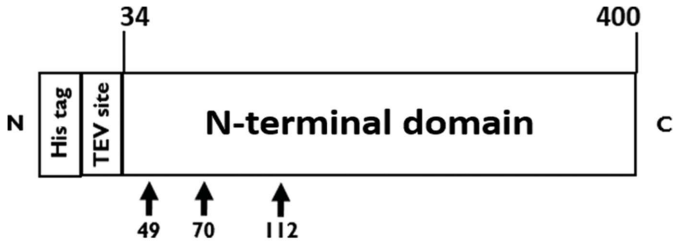
2.11. Other software

Figures were generated using *PyMOL* (version 1.8; Schrödinger; http://www.pymol.org) and *UCSF Chimera*.

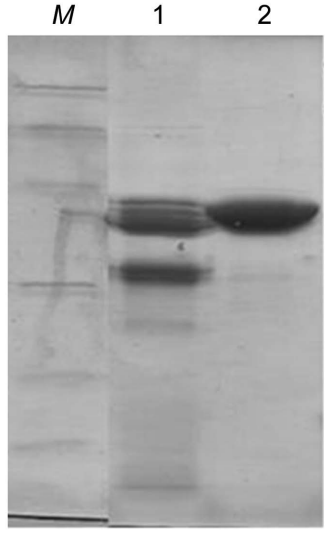
3. Results and discussion

3.1. $BibA_{126–398}$ crystal structure

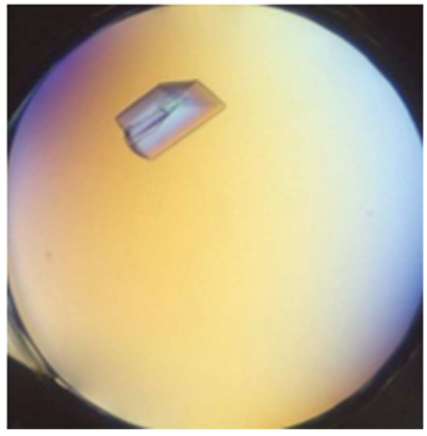
Concentrated $BibA_{34–400}$ protein (Fig. 3*a*) was subjected to initial crystallization trials, but the sample was very susceptible to proteolysis. Several attempts were made to produce an intact protein by modifying the purification protocols and using various protease inhibitors. Based on the N-terminal sequencing results of the proteolytically digested protein sample (data not shown), multiple mutations were introduced. The residues at positions 49, 70 and 112 were mutated to asparagine, delaying the breakdown of the protein (Fig. 3*b*). This mutated $BibA_{34–400}$ sample with selenomethionine substitutions was concentrated to 40.8 mg ml^{-1} and subjected to crystallization screening. Single good-quality crystals only appeared in the drops after several months (Fig. 3*c*). N-terminal sequencing was performed on the crystals and revealed the loss of 92 amino-acid residues (His34–Ser125).



(*a*)



(*b*)



(*c*)

Figure 3
(*a*) A schematic of GBS $BibA_{34–400}$ with N-terminal His tag and TEV protease cleavage site. (*b*) SDS–PAGE gel showing the breakdown of the $BibA_{34–400}$ sample. The marker, $BibA_{34–400}$ without any mutations and $BibA_{34–400}$ with K49N, K70N and K112N mutations are labeled *M*, 1 and 2, respectively. (*c*) Picture of a $BibA_{126–398}$ crystal formed after proteolysis.

Table 3
Data collection, processing and refinement.

Values in parentheses are for the highest resolution shell

Diffraction source	SER-CAT 22-ID, APS
Wavelength (Å)	0.9757
Temperature (K)	100
Space group	$P2_122_1$
a, b, c (Å)	26.42, 68.81, 199.38
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.25
Resolution range (Å)	68.81–3.03 (3.14–3.03)
Total No. of reflections	15398 (1474)
No. of unique reflections	7702 (737)
Completeness (%)	100.0 (100.0)
Multiplicity	2.0 (2.0)
$\langle I/\sigma(I)\rangle$	16.10 (3.46)
R_{merge} (%)	2.81 (18.93)
R_{meas} (%)	3.98 (26.78)
$CC_{1/2}$	0.999 (0.905)
Overall B factor from Wilson plot (Å^2)	78.73
Refinement statistics	
Resolution range (Å)	47.8–3.03 (3.14–3.03)
Completeness (%)	100.0 (100.0)
No. of reflections, working set	7699 (737)
No. of reflections, test set	462 (44)
Final R_{work} (%)	22.7 (46.1)
Final R_{free} (%)	26.2 (40.6)
No. of non-H atoms	2158
No. of protein residues	273
R.m.s. deviations	
Bond lengths (Å)	0.011
Angles (°)	1.09
Average B factors (Å^2)	
Non-H atoms	68.2
Ramachandran plot	
Most favored (%)	96.7
Allowed (%)	3.3
Outliers (%)	0
PDB entry	6poo

GBS BibA lacking the N-terminal 92 residues (that is, $BibA_{126–400}$) crystallized in space group $P2_122_1$ with one molecule in the asymmetric unit. The crystal structure of $BibA_{126–400}$ was solved by the selenomethionine multi-wavelength anomalous diffraction (MAD) method to a resolution of 3.03 Å. The experimentally phased electron-density maps of $BibA_{126–400}$ allowed the building of a 273-amino-acid single polypeptide chain (Asp126–Asp398). The Lys399 and Lys400 residues on the C-terminus could not be built, so the visible fragment of the crystal structure is referred to as $BibA_{126–398}$. Compared with other similar resolution structures deposited in the Protein Data Bank, the final refined model has favorable crystallographic refinement parameters. The data-collection, processing and refinement statistics are summarized in Table 3.

The overall structure of $BibA_{126–398}$ has an ∼150 Å-long distorted rod shape composed of nine α-helices, a 3_{10}-helix and small loops connecting the helices. Interestingly, the nine α-helices adopt a unique pattern to form four antiparallel three-helix-bundle-motif repeats labeled MR1 (Asp126–Ser182), MR2 (Thr183–Lys235), MR3 (Glu236–Leu303) and MR4 (Gln304–Asp398) (Fig. 4*a*). Topologically, the N-terminus first assembles into an 18-residue α-helix (α1; Asp126–Ser142) and a small loop (L1; Leu143–Ser147), followed by the second α-helix (α2; Ser148–Ser160), which runs antiparallel to helix α1. Loop 2 (L2; Ser161–Asp163) proceeds α2, allowing the third α-helix (α3; Ser164–Leu195) to run antiparallel to α2. Among the first three helices, α3 is more extended, comprising of 33 residues, of which residues Ser164–Ser182 are involved in forming the MR1 motif, while residues Thr183–Leu195 extend into the MR2 motif. This shared-helix

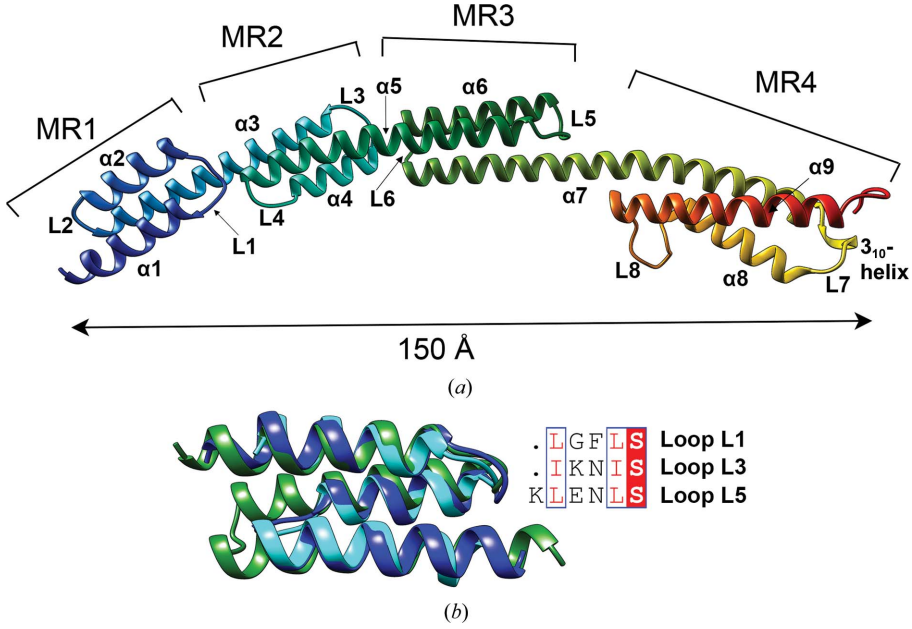


Figure 4
(*a*) Cartoon representation of the $BibA_{126–398}$ crystal structure shown as a rainbow model with color changing gradually from the N-terminus (blue) to the C-terminus (red). The length of the rod-shaped structure is depicted, and the antiparallel three-helix-bundle-motif repeats are labeled from the N-terminus to the C-terminus as MR1–MR4, respectively. (*b*) Superposition of the antiparallel three-helix-bundle-motif repeats MR1 (dark blue), MR2 (cyan) and MR3 (forest green). Similar residues such as Leu and Ile are boxed and shown in a red font and the conserved serine residue is highlighted in red.

pattern to form the motif is repeated to generate motifs MR3 and MR4, which involve the remaining α-helices (α4–α9) and loops (L3–L8). The C-terminal MR4 motif is composed of a longer stretch of residues (94) compared with MR1 (57 residues), MR2 (53 residues) and MR3 (67 residues). The α7 helix that is shared by MR3 and MR4 is the longest, with 54 residues (Ile281–Lys331). The longer MR4 motif harbors an additional 3_{10}-helix (Lys333–Ile335) between the α7 helix and loop 7 (L7), and an extended loop (L8, residues Leu357–Trp365) that is distinct from the other motifs. Helices α3, α5 and α7 bridge the three-helix-bundle motifs together to form an extended rod with a curved shape. This curvature in the protein perhaps assists target proteins such as C4BP by increasing its interaction surface (Youn *et al.*, 2017).

Although there was no significant sequence conservation between the three antiparallel three-helix-bundle-motif repeats MR1, MR2 and MR3, their structural motifs superimpose with r.m.s.d. values of 0.818 Å (MR1 and MR2) and 0.996 Å (MR1 and MR3). The loops L1 (Leu143–Ser147), L3 (Ile196–Ser200) and L5 (Lys251–Ser256) have a conserved serine residue just before their respective succeeding helices and structurally align with each other (Fig. 4*b*).

3.2. Circular dichroism spectra of $BibA_{34–400}$

Crystallization attempts using the $BibA_{34–400}$ protein always produced crystals lacking the N-terminal 92 residues, leading us to believe that this segment might be disordered. The X-ray crystal structure analysis of $BibA_{126–398}$ indicates a highly helical protein. So, to evaluate the structure of full-length $BibA_{34–400}$, the freshly purified protein was analyzed by CD spectrometry. It showed a typical spectrum for an α-helical secondary structure, with two negative peaks at 208 and 222 nm (Fig. 2*a*). Estimation of the protein secondary structure from the CD spectrum using *BeStSel* (Micsonai *et al.*, 2015) revealed a composition of about 80% α-helix, with about 20% categorized as 'others' (Fig. 2*b*). This composition is similar to that of the $BibA_{126–398}$ crystal structure. Further, to validate the secondary-structure analysis of $BibA_{34–400}$ by *BeStSel*, we generated CD spectra with the help of the *PDB2CD* web server (Mavridis & Janes, 2017) using the atomic coordinates of the $BibA_{126–398}$ crystal structure (Fig. 2*a*) and performed secondary-structure analysis with the help of the *BeStSel* web server (Fig. 2*b*). The secondary-structure content estimated from the generated CD spectrum of $BibA_{126–398}$ was similar to that from the experimental CD spectrum of $BibA_{34–400}$, and based on this secondary-structure content analysis we suggest that the N-terminal 92-residue segment (His34–Ser125) is α-helical in nature and might adopt a conformation similar to the rest of the molecule, forming antiparallel three-helix-bundle-motif repeats.

3.3. Small-angle X-ray scattering (SAXS) envelope and the $BibA_{34–400}$ model

To gain further insights into the structure of $BibA_{34–400}$, freshly purified and intact recombinant $BibA_{34–400}$ was analyzed by SEC-SAXS. The pair distance distribution function $P(r)$, calculated using *GNOM*, suggests that the protein is an elongated cylinder in shape (Kajiwara & Hiragi, 1996; Putnam *et al.*, 2007), as the asymmetric curve has a gradual slope after the peak (Fig. 5*a*). As inferred from the CD spectra and SAXS analysis, the initial model for full-length $BibA_{34–400}$ was generated using *Coot* by mimicking the MR1 and MR2 motifs for the missing N-terminal residues. After mutating the residues to the original sequence of the protein, the model from *Coot* was used as a template in *I-TASSER* to generate the full $BibA_{34–400}$ model. The final refined model predicted is ∼200 Å long and contains 367 amino acids (His34–Lys400) with six antiparallel three-helix-bundle-motif repeats. The N-terminal motif repeats were labeled MR-N1 and MR-N2 (Fig. 5*c*). Further, we have calculated theoretical profiles for the $BibA_{34–400}$ model using *FoXS* (Schneidman-Duhovny *et al.*, 2013, 2016) and compared them with the experimental scattering profile of $BibA_{34–400}$ in solution (Fig. 5*b*). The profiles fit reasonably, with a χ^2 of 2.251.86, possibly supporting the structure of the predicted $BibA_{34–400}$ model. Although the scaling of the atomic radius (c_1), which is 1.05, is within the reasonable range ($0.99 \leq c_1 \leq 1.05$), the parameter controlling the density of the water layer around the molecule (c_2) is higher than the maximum value ($-2.0 \leq c_2 \leq 4.0$), suggesting possible overfitting of the theoretical profile of the proposed $BibA_{34–400}$ model in the experimental scattering profile.

The *ab initio* envelope calculated for the full-length protein ($BibA_{34–400}$) using *DAMMIF* was an elongated and distorted rod-like shape of ∼200 Å in length. This *ab initio* envelope fits well with the experimental data, with a χ^2 of 1.084. The crystal structure of $BibA_{126–398}$ was fitted using *UCSF Chimera* and, as expected, it sits well in the *ab initio* envelope with an empty pocket on the N-terminal side of the pocket (Fig. 5*c*). By comparison, the predicted $BibA_{34–400}$ hybrid model fits very well in the SAXS-derived envelope, with 2756 of 2879 atoms fittied inside the contour, agreeing reasonably with the proposed model (Fig. 5*c*).

3.4. Binding-study results

The binding of plasma-purified C4BP (Fig. 6*a*) and rC4BP (Fig. 6*b*) to immobilized BibA was measured, and the K_d was estimated to be 28.99 n*M*, suggesting moderate to tight binding of the C4BP α-chain and immobilized $BibA_{34–400}$. However, the binding-affinity estimation for C4BP and BibA (Fig. 6*a*) is not accurate compared with that estimated for BibA and rC4BP, and we present this data only as an estimation that is useful for understanding the binding mode. Both proteins bind $BibA_{34–400}$ in a relatively similar range, indicating that the binding site is probably in the α-chain and not the β-chain. The interaction affinity between rC4BP immobilized in dextran matrix and BibA in solution analyzed by surface plasmon resonance (Fig. 6*c*), in a similar range of direct binding, results in a K_d of 1.35 n*M* ($k_a = 2.6 \times 10^5\ M^{-1}\ s^{-1}$ and $k_d = 3.5 \times 10^{-4}\ s^{-1}$).

We have localized the binding of $BibA_{34–400}$ to the C4BP α-chain, and further observed that the ΔCCP1 mutant

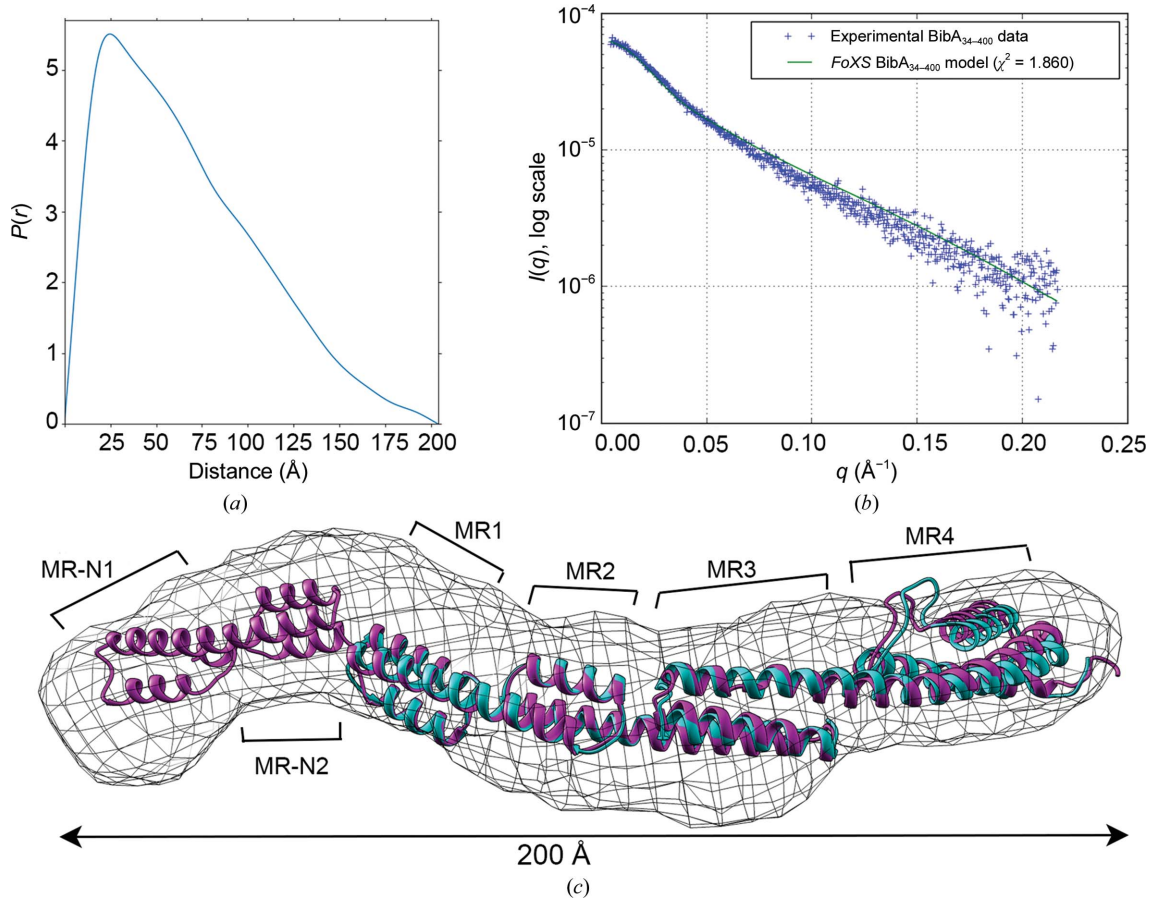


Figure 5
(*a*) Pair distance distribution [*P*(*r*)] function of intact BibA$_{34–400}$ protein. (*b*) Comparison of the experimental scattering profile (in blue) for BibA$_{34–400}$ with profiles from a theoretical model (*FoXS*; green) derived from the proposed BibA$_{34–400}$ model. (*c*) Fit of the crystal structure of BibA$_{126–398}$ (cyan) and the proposed BibA$_{34–400}$ (magenta) into the *ab initio* model of BibA$_{34–400}$ calculated with *DAMMIF*.

exhibited decreased binding (Fig. 6*d*). There appears to be a second BibA binding site in CCP7–8 indicated by the decrease in binding shown by the ΔCCP7 and ΔCCP8 mutants. The CCP7–8 domains, however, may be less accessible *in vivo* considering that they are close to the crowded C-terminal oligomerization region of C4BP, which is also in proximity to Protein S bound to the α-chain of C4BP. This hypothesis is supported by the finding that rC4BP lacking the β-chain and Protein S binds more strongly to BibA than plasma-purified protein carrying these additional components. Therefore, we focused our structural effort on the CCP1–2 region and its interactions with BibA$_{34–399}$. However, we were not able to purify the BibA$_{34–399}$–CCP1–2 complex despite trying multiple buffers and salt concentrations.

4. Conclusions

The crystal structure of BibA$_{126–398}$ was solved by the MAD phasing method. We could not determine the full-length BibA$_{34–400}$ crystal structure owing to protein degradation during purification and crystallization. Suspecting the N-terminal BibA$_{34–125}$ to be flexible or disordered, we introduced mutations into the BibA$_{34–400}$ constructs, which slowed down the degradation during purification (Fig. 3*b*). However, the single crystals obtained were of BibA$_{126–398}$, and the crystal structure of this C-terminal segment is predominantly α-helical, with four sequentially connected three-helical bundles forming a slightly curved rod-shaped molecule.

We pursued circular dichroism spectrometry and small-angle X-ray scattering analysis to gain insights into the intact BibA$_{34–400}$ protein. The CD spectra of BibA$_{34–400}$ recorded at 20 and 40°C and the CD spectrum of denatured and refolded sample recorded at 20°C suggested the structure to be 80% α-helical, similar to the content observed in the crystal structure of BibA$_{126–398}$. We generated a model for the full-length BibA$_{34–400}$ structure by combining an N-terminal segment *ab initio* model, which was built assuming it to be similarly composed of two three-helical bundle motifs (MR-N1 and MR-N2), and the C-terminal BibA$_{126–398}$ crystal structure. This hybrid model fits well in the *ab initio* SAXS envelope. However, emphasizing that the N-terminal segment *ab initio* model generated by us may be biased, we could safely suggest that the full-length BibA$_{34–400}$ is a rod-shaped molecule and is predominantly α-helical in composition.

The overall structure of the N-terminal domain of BibA is novel and provides insight into the structure of a bacterial

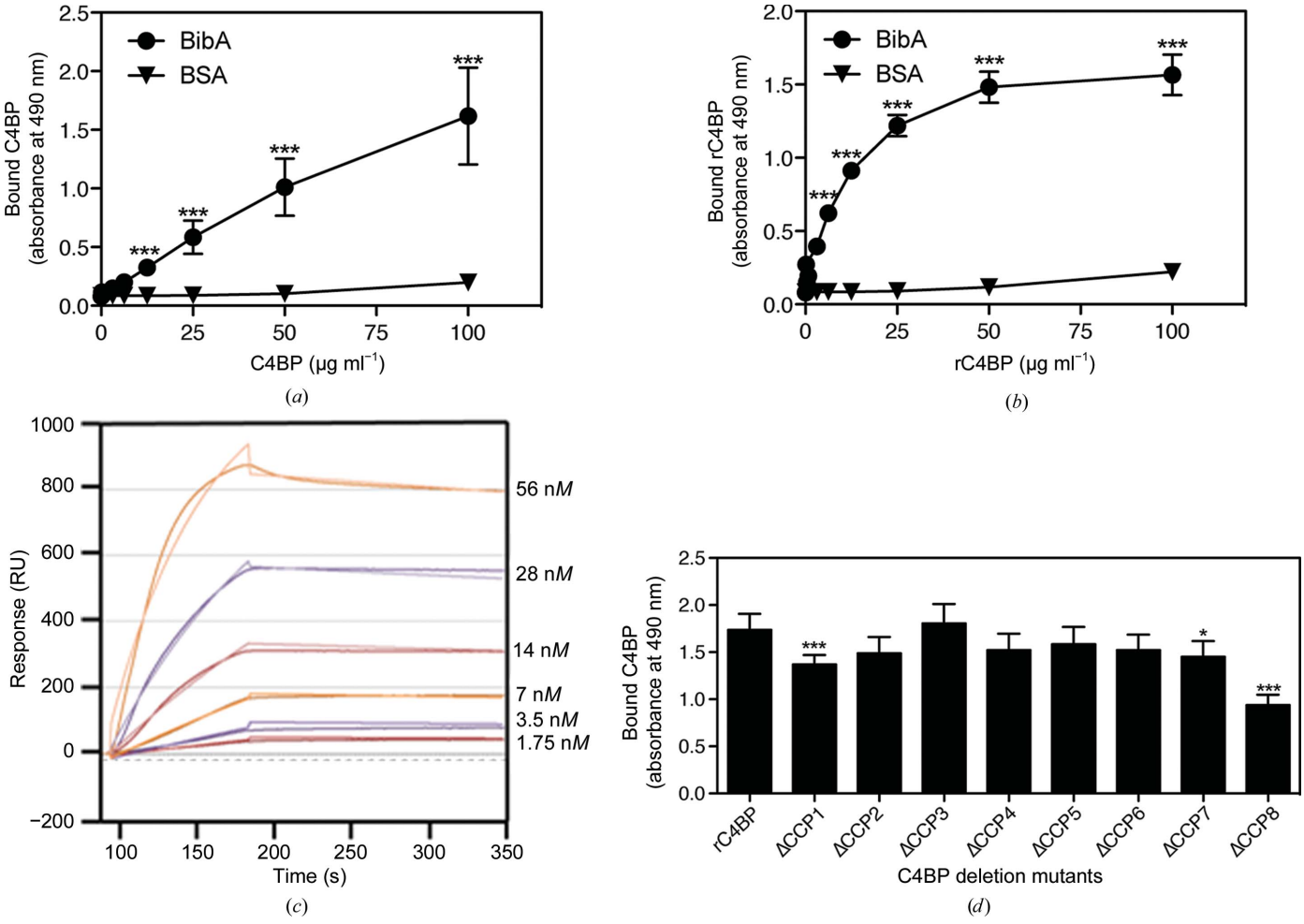


Figure 6
Microtiter plates were coated with BibA (5 μg ml^{-1}) and increasing amounts of plasma-purified C4BP (*a*) or rC4BP (*b*) were added. Binding was detected using specific polyclonal anti-C4BP antibodies. BSA was used as a negative control. The mean ± SD of three independent experiments performed in duplicate is presented. Statistical significance was calculated using a two-way Anova test. ***, $p < 0.001$. (*c*) Surface plasmon resonance analysis to measure the binding of BibA to immobilized C4BP using Biacore. (*d*) The C4BP variants were allowed to bind to BibA immobilized on a plate. Bound C4BP was detected with polyclonal Abs. The graph represents data from three independent experiments performed in duplicate ± SD. Statistical significance was calculated using a one-way Anova test. *, $p < 0.05$; ***, $p < 0.001$.

component involved in complement inhibition. The binding site of BibA to C4BP is located in the α-chain, with two potential interaction sites in CCP1 and CCP8. Currently, we have not produced a BibA–C4BP complex in solution, and we were not able to identify the precise region of BibA binding to C4BP. Moving forward, we will attempt to generate a crystal or solution structure of the BibA–C4BP complex, which may help in developing strategies for fighting GBS infections.

The amino-acid sequence of BibA is unique, and a non-redundant GenBank database *BLAST* search (Altschul *et al.*, 1997) did not reveal any significant homologies. Nevertheless, on searching for three-helix-bundle proteins similar to the $BibA_{34–400}$ model, we found a host immune response-modulating surface protein from *Staphylococcus aureus* called protein A (Youn *et al.*, 2017), which has 22% sequence identity to BibA. Protein A has antiparallel three-helix-bundle motif repeats and its secondary-structure content is 80% helical and turns, similar to BibA. However, protein A (PDB entry 5h7b) is a 237-residue polypeptide with five repeat modules, compared with the 273-residue $BibA_{126–398}$ protein, which is made of four repeat modules. Hence, the repeat modules of protein A are smaller in size and do not superimpose on the BibA structure and its structural motifs.

The GBS BibA functional homologs in group A streptococcus (GAS) are the surface-anchored virulence factor M proteins (Ghosh, 2011; McNamara *et al.*, 2008). These proteins bind to C4BP (Persson *et al.*, 2006) when recruited on the surface of GAS, and the binding has been localized to the N-terminal 50 residues of M proteins (hypervariable region, HVR; Lannergård *et al.*, 2011; Penfound *et al.*, 2010; Sandin *et al.*, 2006). The crystal structure of C4BP α-chain CCP1 and CCP2 in complex with hypervariable regions of four different M proteins (M2, M22, M28 and M49) helped to identify conserved sequence patterns within their HVRs (Buffalo *et al.*, 2016). However, we were unable to identify a similar pattern in BibA, suggesting that the interactions of C4BP with GBS surface proteins differ from those with GAS surface proteins.

Acknowledgements

This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. BioCAT is supported by grant 9 P41 GM103622 from the National Institute of General Medical Sciences of the National Institutes of Health (NIGMS). Use of the PILATUS3 1M detector was provided by grant 1S10OD018090-01 from NIGMS. We are grateful to the staff of the SER-CAT (22-ID) and BioCAT (18-ID) beamlines of the Advanced Photon Source, Illinois, Chicago for their help with data collection. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Funding information

This research was funded by the National Institute of Allergy and Infectious Diseases (grant R01-AI106808 to Sthanam V. L. Narayana) and by the National Institute of Dental and Craniofacial Research (award Nos. R01-DE025015 and R01-DE017382 to Hung Ton-That).

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Avirutnan, P., Hauhart, R. E., Somnuke, P., Blom, A. M., Diamond, M. S. & Atkinson, J. P. (2011). *J. Immunol.* **187**, 424–433.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Blom, A. M. (2002). *Biochem. Soc. Trans.* **30**, 978–982.
Blom, A. M., Berggård, K., Webb, J. H., Lindahl, G., Villoutreix, B. O. & Dahlbäck, B. (2000). *J. Immunol.* **164**, 5328–5336.
Blom, A. M., Kask, L. & Dahlbäck, B. (2001). *J. Biol. Chem.* **276**, 27136–27144.
Blom, A. M., Kask, L. & Dahlbäck, B. (2003). *Mol. Immunol.* **39**, 547–556.
Blom, A. M., Webb, J., Villoutreix, B. O. & Dahlbäck, B. (1999). *J. Biol. Chem.* **274**, 19237–19245.
Blom, A. M., Zadura, A. F., Villoutreix, B. O. & Dahlbäck, B. (2000). *Mol. Immunol.* **37**, 445–453.
Buffalo, C. Z., Bahn-Suh, A. J., Hirakis, S. P., Biswas, T., Amaro, R. E., Nizet, V. & Ghosh, P. (2016). *Nat. Microbiol.* **1**, 16155.
Butko, P., Nicholson-Weller, A. & Wessels, M. R. (1999). *J. Immunol.* **163**, 2761–2768.
Chen, Z. & Ruffner, D. E. (1998). *Nucleic Acids Res.* **26**, 1126–1127.
Dahlbäck, B. (1983). *Biochem. J.* **209**, 847–856.
Dermer, P., Lee, C., Eggert, J. & Few, B. (2004). *J. Pediatr. Nurs.* **19**, 357–363.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Farley, M. M. (2001). *Clin. Infect. Dis.* **33**, 556–561.
Farley, M. M., Harvey, C., Stull, T., Smith, J. D., Schuchat, A., Wenger, J. D. & Stephens, D. S. (1993). *N. Engl. J. Med.* **328**, 1807–1811.
Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). *J. Appl. Cryst.* **50**, 1212–1225.
Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
Ghosh, P. (2011). *Adv. Exp. Med. Biol.* **715**, 197–211.
Hallström, T., Jarva, H., Riesbeck, K. & Blom, A. M. (2007). *J. Immunol.* **178**, 6359–6366.
Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.
Jarva, H., Jokiranta, T. S., Würzner, R. & Meri, S. (2003). *Mol. Immunol.* **40**, 95–107.
Jenkins, H. T., Mark, L., Ball, G., Persson, J., Lindahl, G., Uhrin, D., Blom, A. M. & Barlow, P. N. (2006). *J. Biol. Chem.* **281**, 3690–3697.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kajiwara, K. & Hiragi, Y. (1996). *Applications of Synchrotron Radiation to Materials Analysis*, edited by H. Saisho & Y. Gohshi, pp. 353–404. Amsterdam: Elsevier.
Kirby, N., Cowieson, N., Hawley, A. M., Mudie, S. T., McGillivray, D. J., Kusel, M., Samardzic-Boban, V. & Ryan, T. M. (2016). *Acta Cryst.* D**72**, 1254–1266.
Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
Krohn, M. A., Hillier, S. L. & Baker, C. J. (1999). *J. Infect. Dis.* **179**, 1410–1415.
Ladhani, S. N., Henderson, K. L., Muller-Pebody, B., Ramsay, M. E. & Riordan, A. (2019). *Arch. Dis. Child.* **104**, 874–878.
Laemmli, U. K. (1970). *Nature*, **227**, 680–685.
Lannergård, J., Gustafsson, M. C., Waldemarsson, J., Norrby-Teglund, A., Stålhammar-Carlemalm, M. & Lindahl, G. (2011). *Cell Host Microbe*, **10**, 147–157.
Leung, E., Blom, A. M., Clemenza, L. & Isenman, D. E. (2006). *Biochemistry*, **45**, 8378–8392.
Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.
Madrid, L., Seale, A. C., Kohli-Lynch, M., Edmond, K. M., Lawn, J. E., Heath, P. T., Madhi, S. A., Baker, C. J., Bartlett, L., Cutland, C., Gravett, M. G., Ip, M., Le Doare, K., Rubens, C. E., Saha, S. K., Sobanjo-ter Meulen, A., Vekemans, J., Schrag, S., Agarwal, R., da Silva, A. R. A., Bassat, Q., Berkley, J. A., Dangor, Z., Dhaded, S., Giannoni, E., Hammoud, M., Kobayahsi, M., O'Sullivan, C., Sakata, H., Sridhar, S., Sigaúque, B., Tyrrell, G. & Paul, V. (2017). *Clin. Infect. Dis.* **65**, S160–S172.
Mavridis, L. & Janes, R. W. (2017). *Bioinformatics*, **33**, 56–63.
McNamara, C., Zinkernagel, A. S., Macheboeuf, P., Cunningham, M. W., Nizet, V. & Ghosh, P. (2008). *Science*, **319**, 1405–1408.
Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M. & Kardos, J. (2015). *Proc. Natl Acad. Sci. USA*, **112**, E3095–E3103.
Nordström, T., Blom, A. M., Forsgren, A. & Riesbeck, K. (2004). *J. Immunol.* **173**, 4598–4606.
Penfound, T. A., Ofek, I., Courtney, H. S., Hasty, D. L. & Dale, J. B. (2010). *J. Infect. Dis.* **201**, 1580–1588.
Persson, J., Beall, B., Linse, S. & Lindahl, G. (2006). *PLoS Pathog.* **2**, e47.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Pietrocola, G., Arciola, C. R., Rindi, S., Montanaro, L. & Speziale, P. (2018). *Front. Immunol.* **9**, 602.
Potempa, M., Potempa, J., Okroj, M., Popadiak, K., Eick, S., Nguyen, K. A., Riesbeck, K. & Blom, A. M. (2008). *J. Immunol.* **181**, 5537–5544.
Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). *Q. Rev. Biophys.* **40**, 191–285.
Sandin, C., Carlsson, F. & Lindahl, G. (2006). *Mol. Microbiol.* **59**, 20–30.
Santi, I., Grifantini, R., Jiang, S.-M., Brettoni, C., Grandi, G., Wessels, M. R. & Soriani, M. (2009). *J. Bacteriol.* **191**, 5387–5397.

Santi, I., Maione, D., Galeotti, C. L., Grandi, G., Telford, J. L. & Soriani, M. (2009). *J. Infect. Dis.* **200**, 564–570.
Santi, I., Scarselli, M., Mariani, M., Pezzicoli, A., Masignani, V., Taddei, A., Grandi, G., Telford, J. L. & Soriani, M. (2007). *Mol. Microbiol.* **63**, 754–767.
Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. (2013). *Biophys. J.* **105**, 962–974.
Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. (2016). *Nucleic Acids Res.* **44**, W424–W429.
Sendi, P., Johansson, L. & Norrby-Teglund, A. (2008). *Infection*, **36**, 100–111.
Senn, B. M., Visram, Z., Meinke, A. L., Neubauer, C., Gelbmann, D., Sinzinger, J., Hanner, M., Lundberg, U., Boisvert, H., Reinscheid, D., von Gabain, A. & Nagy, E. (2011). *Vaccine*, **29**, 4116–4124.
Skubák, P. & Pannu, N. S. (2013). *Nat. Commun.* **4**, 2777.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Villoutreix, B. O., Härdig, Y., Wallqvist, A., Covell, D. G., García de Frutos, P. & Dahlbäck, B. (1998). *Proteins*, **31**, 391–405.
Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
Wessels, M. R., Butko, P., Ma, M., Warren, H. B., Lage, A. L. & Carroll, M. C. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 11490–11494.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.
Youn, S.-J., Kwon, N.-Y., Lee, J. H., Kim, J. H., Choi, J., Lee, H. & Lee, J.-O. (2017). *Sci. Rep.* **7**, 2595.
Zhang, Y. (2008). *BMC Bioinformatics*, **9**, 40.