FEBS Open Bio 4 (2014) 283–289

FEBS openbio

journal homepage: www.elsevier.com/locate/febsopenbio

Structure of the C-terminal domain of AspA (antigen I/II-family) protein from *Streptococcus pyogenes*

Michael Hall [a], Åsa Nylander [b], Howard F. Jenkinson [c], Karina Persson [a,*]

[a] Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden
[b] Department of Odontology, Division of Oral Microbiology, Umeå University, SE-901 87 Umeå, Sweden
[c] School of Oral and Dental Sciences, University of Bristol, Bristol BS1 2LY, UK

ARTICLE INFO

Article history:
Received 15 January 2014
Revised 27 February 2014
Accepted 28 February 2014

Keywords:
Antigen I/II
Isopeptide bond
Adhesin
X-ray crystallography

ABSTRACT

The pathogenic bacteria *Streptococcus pyogenes* can cause an array of diseases in humans, including moderate infections such as pharyngitis (strep throat) as well as life threatening conditions such as necrotizing fasciitis and puerperal fever. The antigen I/II family proteins are cell wall anchored adhesin proteins found on the surfaces of most oral streptococci and are involved in host colonization and biofilm formation. In the present study we have determined the crystal structure of the C_{2-3}-domain of the antigen I/II type protein AspA from *S. pyogenes* M type 28. The structure was solved to 1.8 Å resolution and shows that the C_{2-3}-domain is comprised of two structurally similar DEv-IgG motifs, designated C_2 and C_3, both containing a stabilizing covalent isopeptide bond. Furthermore a metal binding site is identified, containing a bound calcium ion. Despite relatively low sequence identity, interestingly, the overall structure shares high similarity to the C_{2-3}-domains of antigen I/II proteins from *Streptococcus gordonii* and *Streptococcus mutans*, although certain parts of the structure exhibit distinct features. In summary this work constitutes the first step in the full structure determination of the AspA protein from *S. pyogenes*.

1. Introduction

The Gram-positive bacterium *Streptococcus pyogenes* (or Group A Streptococcus; GAS) is a causative agent of many human diseases, including pharyngitis (strep throat), impetigo, scarlet fever, cellulitis, and also potentially life threatening invasive disease such as necrotizing fasciitis. In order to colonize and proliferate on host tissues *S. pyogenes* must first adhere to the host cell surface, where it has the potential to form biofilms in the oral cavity, nasopharynx and on skin and wounds. Several mechanisms for biofilm formation by different strains and serotypes of *S. pyogenes* have been proposed. The first involves the formation of complexes between cell surface lipoteichoic acid (LTA) and members of the M protein family, which together control the hydrophobic properties of the cell surface, thereby mediating adhesion and biofilm formation in M1 serotype *S. pyogenes* [1]. A second mechanism involves the use of pili structures on the bacterial cell surface [2,3]. A third alternative is that biofilm formation may be mediated by so called antigen I/II (AgI/II) type proteins. The AgI/II proteins are a family of cell surface adhesins which previously have been shown to play many roles in host colonization in the oral viridans group of streptococci. The proteins of the AgI/II family are large, comprising 1310–1653 amino acids, which have a distinct domain organization [4]. The primary sequences of AgI/II proteins begin with a signal peptide, approximately 30–40 amino acid residues in length, followed in sequence by a small N-terminal domain, an alanine rich repeat region (A), a variable domain (V), a proline rich repeat region (P), a C-terminal region, and finally a cell wall anchor segment containing an LPxTG-motif. The LPxTG sequence is recognized by housekeeping sortases, which mediate cell wall anchoring and presentation of the AgI/II protein on the bacterial surface. The crystal structures of several domains of the AgI/II proteins from the oral streptococci *Streptococcus mutans* and *Streptococcus gordonii* revealed that the AgI/II proteins fold into highly elongated structures [5–9], with the V-domain being presented at the tip of the protein, furthest away from the cell wall [4]. The elongated structure is achieved by the interaction of the A-domain α-helical repeat with the P-domain polyproline II (PPII) helix, which intertwine, creating a supercoiled fibrillar stalk structure, projecting the V-domain >50 nm out from the cell surface [4,6]. The N- and C-domains are the domains that are closest to the cell wall, and form the base of the protein.

In the genome of *S. pyogenes* serotype M28, a single AgI/II type protein is encoded by the *aspA* gene. The gene product, AspA, is comprised of 1352 amino acids and is expected to exhibit the typical AgI/II protein domain architecture (Fig. 1A). Similarly as ob-

* Corresponding author. Tel.: +46 907865926.
E-mail address: karina.persson@chem.umu.se (K. Persson).

http://dx.doi.org/10.1016/j.fob.2014.02.012
2211-5463/

served for SpaP from *S. mutans* and SspA and SspB from *S. gordonii*, the AspA protein contains three repeat regions in the A-domain, three repeat regions in the P-domain, and also the C-domain can be divided into three subdomains. Although the general structural characteristics are predicted to be similar, the overall sequence identity between AspA and SpaP is only 27% (and 28% between AspA and SspA/B). A sequence alignment of the C_{2-3} regions of AspA, SpaP and SspB is shown in Fig. 1B. Especially the AspA V-domain, which shares only 8% sequence identity with the V-domain of SpaP, appears to be functionally distinct from those of SpaP, SspA and SspB. Interestingly, the V-domain shares high sequence similarity with the V-domain of the AgI/II protein BspD from *Streptococcus agalactiae*, and they both contain conserved histidine–aspartic acid clusters, typically associated with divalent metal binding sites [4]. In agreement with this observation it was recently shown that the V-domain of AspA could indeed bind divalent metal ions, specifically Zn^{2+} ions [10]. However, the functional significance of the Zn^{2+} binding remains unknown. In the same study it was also concluded that AspA can bind immobilized salivary agglutinin gp-340 and that deletion of the *aspA* gene abolished the abilities of two different M28 strains of *S. pyogenes* to form biofilms on saliva coated surfaces [10].

In order to increase our understanding of the mechanisms and function of the AspA protein in *S. pyogenes* host colonization and biofilm formation we have undertaken to determine the crystal structure of the AspA polypeptide. In this study we present the structure of the C_2 and C_3 components of the C-domain, determined by X-ray crystallography to 1.8 Å resolution.

2. Materials and methods

2.1. Cloning

The C-terminal domain of AspA was cloned from plasmid pET-46 Ek/LIC-NAVPC [10] encoding the *S. pyogenes aspA* gene (GeneID: 3574034). PCR primers were designed based on the crystal structure of the *S. gordonii* SspB C_{2-3}-domain. Forward primer was 5′-TTTTTCCATGGATAATCTGATTCAGCCAACT-3′ and reverse 5′-AAAAAGGTACCTTACGTATGAGTGGTGACTTT-3′. The PCR product was digested with Acc65I and NcoI and ligated into the equivalent sites of the pET-M11 expression vector. The final construct encodes MKHHHHHHPM–AspA-C_{2-3}. The plasmids were transformed into *Escherichia coli* DH5α and subsequently selected on kanamycin plates. The positive clones were verified by DNA sequencing.

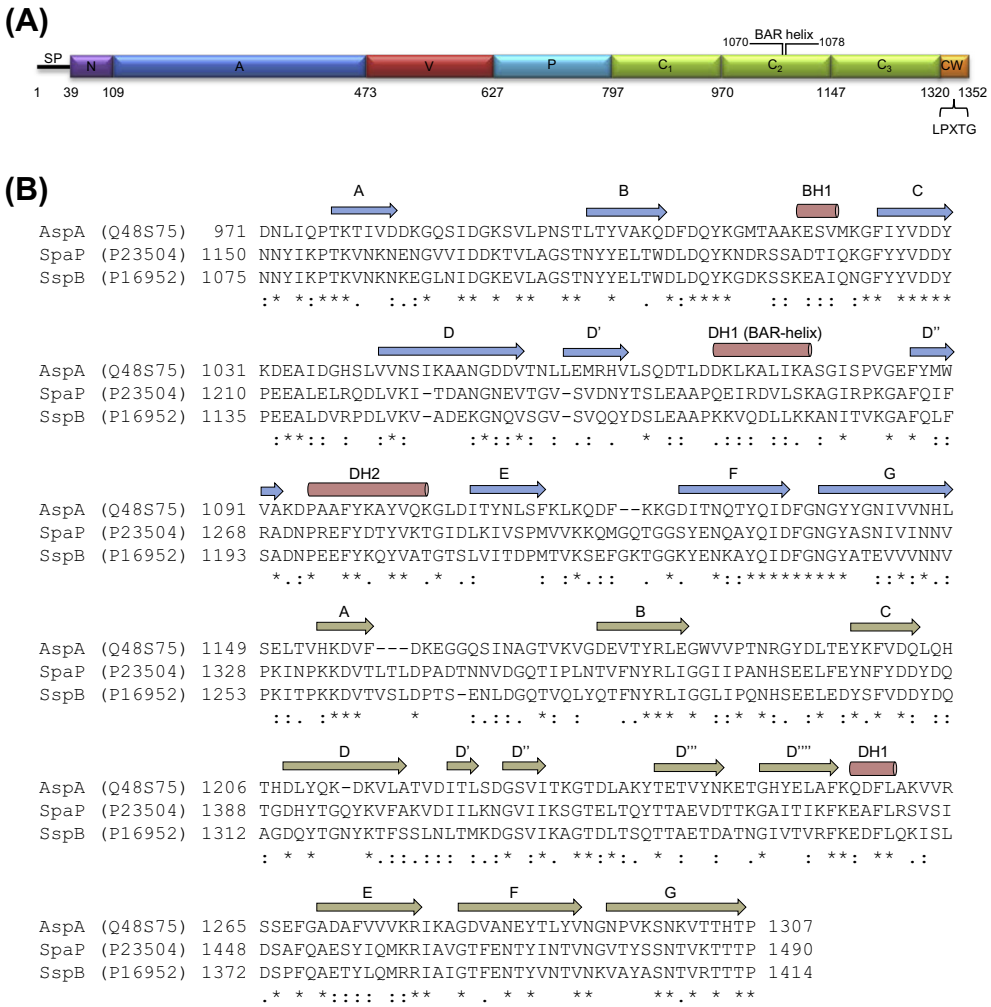

Fig. 1. Domain organization of *S. pyogenes* AspA and sequence alignment of the C_{2-3} domains of AspA, SpaP and SspB. (A) The primary sequence of AspA begins with a signal peptide (SP), followed in sequence by a small N-terminal domain (N), an alanine rich repeat region (A), a variable domain (V), a proline rich repeat region (P), a C-terminal region (C), and finally a cell wall anchor segment (CW) containing an LPxTG-motif for sortase mediated cell wall attachment. (B) Sequence alignment of the AspA (Uniprot id. Q48S75), SpaP (P23404) and SspB (P16952) C_{2-3} domains. Secondary structure elements from the structure of AspA-C_{2-3} are marked above the alignment.

2.2. Overexpression and purification

The AspA-C_{2-3} construct was overexpressed in *E. coli* BL21 (DE3) at 310 K in Luria Broth supplemented with 50 μg mL^{-1} kanamycin. When the cultures reached an OD_{600} of 0.6, the temperature was lowered to 303 K and the expression was induced with 0.4 mM IPTG after which the cultures were grown for an additional 4 h. Cells were harvested by centrifugation at 5300*g* and the pellets were frozen at 193 K. Cell pellets were resuspended in 50 mM NaH_2PO_4 pH 7.5, 300 mM NaCl and 10 mM imidazole supplemented with EDTA-free protease inhibitor cocktail (Roche). The suspension was lysed on ice by sonication and cellular debris was removed by centrifugation at 39,000*g* for 60 min. The supernatant was loaded onto a column packed with Ni–NTA agarose (Qiagen). The protein was washed in 50 mM NaH_2PO_4 pH 7.5, 300 mM NaCl and 20 mM imidazole and eluted with 50 mM NaH_2PO_4 pH 7.5, 300 mM NaCl and 300 mM imidazole. The buffer was exchanged to 30 mM Tris–HCl pH 8.0, 200 mM NaCl, 1 mM EDTA and 0.5 mM TCEP. The protein was further purified by size-exclusion chromatography using a HiLoad™ 16/60 Superdex™ 200 prep-grade column (GE Healthcare). The protein purity was judged by SDS–PAGE to be >95% and the protein was finally concentrated to 85 mg mL^{-1} in 20 mM TrisHCl pH 8.0 using an Amicon Ultra centrifugal filter device (Millipore).

2.3. Crystallization, data collection and processing

Initial crystallization trials were performed by the sitting-drop vapor-diffusion method in 96-well MRC-crystallization plates (Molecular Dimensions) using a Mosquito (TTP Labtech) pipetting robot and standard crystal screening kits (Hampton Research and Molecular Dimensions). Crystals were obtained in condition 44 of the MIDAS screen (20% glycerol ethoxylate, 10% tetrahydrofuran and 0.1 M Tris–HCl pH 8.0). This condition was subsequently optimized using the hanging-drop method by mixing 1 μL protein, at a concentration of 85 mg mL^{-1} and 2 μL reservoir solution. Crystals suitable for data collection were obtained after 3 weeks. Diffraction data from a single crystal were collected on a Bruker Proteum X8 X-ray diffraction system, equipped with a MICROSTAR-H rotating anode generator and a Platinum 135 detector. The diffraction images were integrated with SAINT and scaled and merged using SADABS, components of the Proteum software (Bruker).

2.4. Structure determination and refinement

The structure was solved using the BALBES molecular replacement pipeline [11] and automated model building was performed using phenix.autobuild [12]. The model was completed by iterative rounds of manual model building using COOT [13] and refinement using phenix.refine [14]. Translation-libration-screw (TLS) refinement was used in the last rounds of refinement, treating each domain (amino acids 970–1146 for the C_2 domain and 1147–1306 for the C_3 domain) as separate TLS groups. Evolutionary conservation analysis was performed using ConSurf [15] and figures were prepared using CCP4MG [16] and PyMol [17].

3. Results and discussion

3.1. Crystallization and structure determination

The full length AspA protein consists of 1352 amino acid residues, with the C-terminal domain constituting 511 residues. A construct representing the C_2 and C_3 components of the C-terminal domain (residues 971–1307) was expressed, purified and crystallized by the hanging drop vapor diffusion method. The crystals diffracted to a maximum resolution of 1.8 Å and belonged to space group $P2_12_12$, with unit cell parameters a = 99.0, b = 105.0, c = 74.5 Å. Relevant processing, refinement and model quality statistics are presented in Table 1. The asymmetric unit contained two AspA-C_{2-3} molecules (A and B). The structure was solved by molecular replacement and subsequently automated model building was performed. The model was completed by iterative manual model building and refinement, finally yielding a model with an R_{work} of 18.2% and R_{free} of 23.1%. The final model is well ordered with an average atomic displacement factor of 22.8 $Å^2$. The structure consists of residues 971–1306 in chain A and residues 972–1306 in chain B, using numbering based on the full length AspA protein. Additionally 984 water molecules are included in the model. No electron density was observed for the flexible poly-histidine tag and linker region and they were hence not included in the model.

3.2. Overall structure

The overall topology of the AspA-C_{2-3} structure is presented in Fig. 2, and is comprised of two distinct domains, referred to as the C_2- and C_3-domains (residues 971–1149 and 1150–1306 respectively). Together these two domains form an elongated structure, approximately 95 Å long and 35 Å wide. Each domain adopts the DEv-IgG fold [18], which similarly to the classical IgG folds is comprised of two major antiparallel β-sheets, designated ABED and CFG. The ABED sheet is formed by the A, B, E and D strands while the CFG sheet is correspondingly formed by strands C, F and G. The main variation from the classical IgG folds, including additional helices and strands, is found between the D and E strands. For the C_2-domain, there are two additional strands on the CFG sheet, designated D′ and D″ as well as two α-helices, DH1 and DH2, located respectively on the loop region between strands D′ and

Table 1
Data collection, refinement and model statistics.

Wavelength (Å)	1.5418
Resolution range (Å)	31.67–1.80 (1.86–1.80)
Space group	$P2_12_12$
Unit cell (Å, °)	a = 99.0, b = 105.0 c = 74.4, α, β, γ = 90
Total reflections	779400 (26976)
Unique reflections	70996 (6553)
Multiplicity	10.8 (4.1)
Completeness (%)	97.6 (82.7)
Mean $I/\sigma(I)$	24.6 (3.2)
Wilson B-factor ($Å^2$)	15.5
R-merge[a] (%)	6.2 (41.1)
R-work[b] (%)	18.2 (31.1)
R-free[b] (%)	23.1 (35.8)
Number of atoms	6365
Macromolecules	5379
Ligands	2
Water	984
Protein residues	672
RMS bonds (Å)	0.007
RMS angles (°)	1.04
Ramachandran favored (%)	99
Ramachandran outliers (%)	0
Average B-factor ($Å^2$)	22.8
Macromolecules ($Å^2$)	21.0
Metal ions ($Å^2$)	13.1
Solvent ($Å^2$)	32.6
PDB code	4OFQ

Values in parentheses indicate statistics for the highest resolution shell.

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th observation of reflection *hkl and* $\langle I(hkl)\rangle$ is the average over of all observations of reflection *hkl.*

[b] $R_{work} = \Sigma||F_{obs}| - |F_{calc}|| / \Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. R_{free} is equivalent to R_{work} but is calculated using a 5% randomly selected set of reflections which is excluded from refinement.

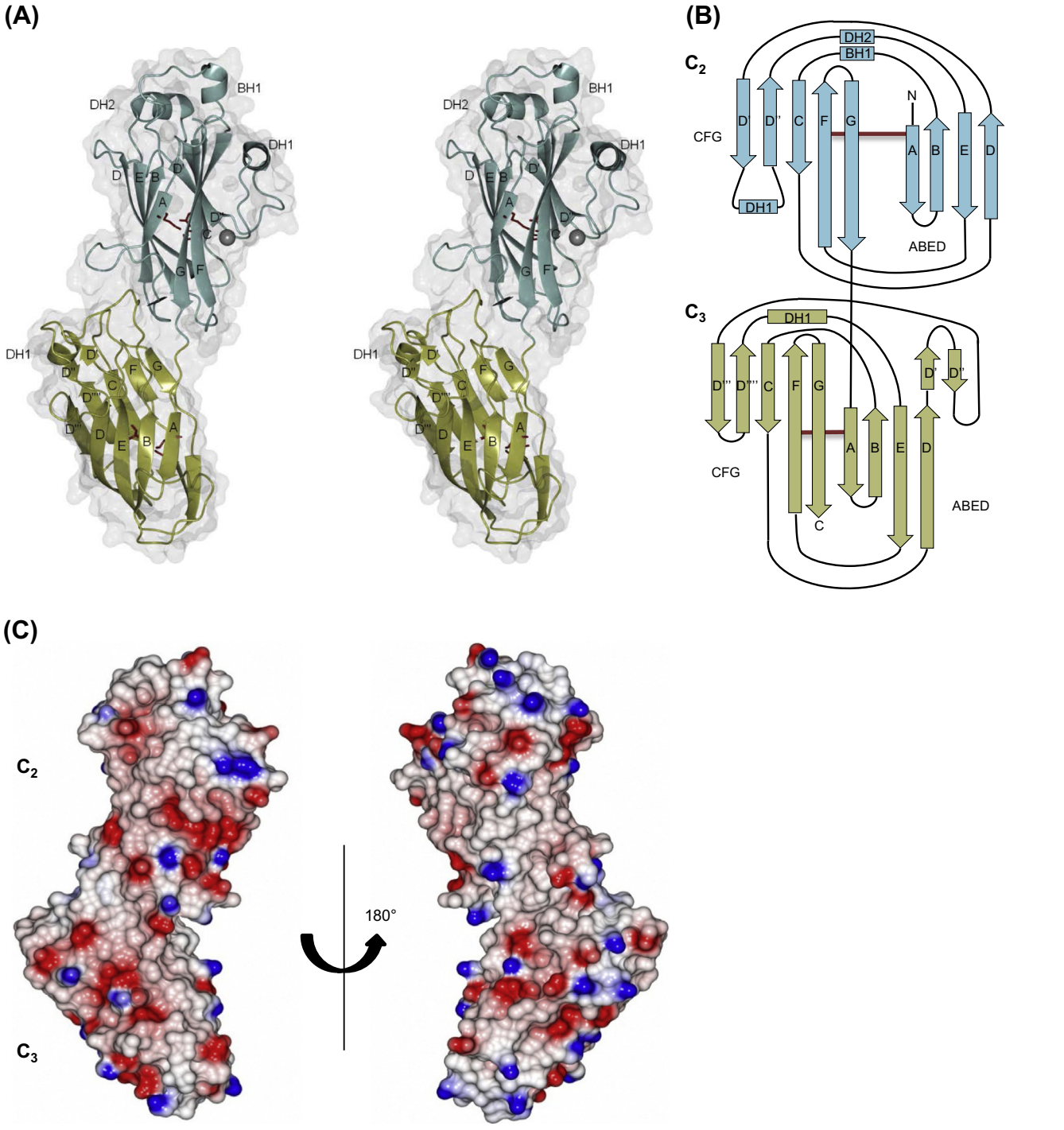

Fig. 2. Overall topology of AspA-C_{2-3}. (A) Ribbon diagram of AspA-C_{2-3} presented in stereo. The C_2 domain (residues 971–1150) is depicted in blue while the C_3 domain (1151–1306) is depicted in yellow. A bound calcium ion is shown as a grey sphere. (B) Topology diagram of AspA-C_{2-3} where α-helices are represented as rectangles, β-strands as arrows and loops as lines. The isopeptide bonds between K987 and N1128 and between K1155 and N1286 are marked as red lines. (C) Electrostatic surface potential rendering of AspA-C_{2-3} colored from −0.5 V (red) to 0.5 V (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

D″ and between strands D″ and E. In addition an α-helix (BH1) is located in the loop region between strands B and C. The central DEv-IgG motif of the C_3-domain is in a similar manner formed from the four stranded ABED and three stranded CFG main β-sheets. Two additional strands (D‴ and D⁗) extend the CFG sheet into a five stranded sheet. As for the C_2-domain, sheets ABED and CFG are interconnected by several cross-connecting loops and one α-helix (DH1) between strand D⁗ of the CFG sheet and strand E of the ABED sheet. The C_2- and C_3-domain are connected by a linker extending from strand G in the C_2-domain to strand A in the C_3-domain. Additionally, in the interface region between the two domains, the side chains of D982 and N996 in the C_2-domain are involved in hydrogen bonding with the side chains of R1264 and N1295 in the C_3 domain. Main chain hydrogen bonding can also be observed between S992 in C_2 and N1189/G1191 in C_3, furthermore stabilizing the interaction between the domains.

Finally the main chain atoms of the interface region exhibit low temperature factors, in the range similar to those for main chain atoms of the central DEv-IgG folds, suggesting that the domains are fixed in position and that the structure is rigid. The C_2 domain contains one bound metal ion, modeled as Ca^{2+}, and both the C_2- and C_3-domains are stabilized by conserved isopeptide bonds, which connect the β-sheets of the central DEv-IgG motifs. An electrostatic surface potential rendering shows that positively and negatively charged residues are heterogeneously distributed across the protein surface (Fig. 2C).

3.3. Isopeptide bonds and metal binding sites

An isopeptide bond can be observed between K978 (on strand A) and N1128 (strand F) in domain C_2, forming a covalent link between the S1 and S2 β-sheets of the β-sandwich motif. The interaction is completed by hydrogen bonding between the C═O and NH isopeptide groups and the side chain of D1028 (Fig. 3A). Similarly K1155 (strand A) and N1286 (strand F) in domain C_3 are connected by an isopeptide bond, stabilized by the side chain of D1201. The isopeptide bonds are surrounded by aromatic and hydrophobic residues. This may result in an increased pK_a of the aspartic acid and decreased pK_a of the lysine residue, facilitating nucleophilic attack on the Asn CG carbon by the deprotonated lysine amino group, finally yielding a covalent bond and the release of ammonium [19]. This type of self-generated stabilizing isopeptide bond has now been identified in a number of different pilin/adhesin proteins from Gram-positive bacteria [5,7,19–22]. Probably the extra structural stabilization inferred by the isopeptide bonds is essential for resistance against physical and chemical stress. Oral bacteria, for example, are constantly exposed to strong physical shear forces from salivary flow as well as tongue movement, and must remain firmly attached to host surfaces in order to persist and colonize.

The previously resolved crystal structures of the C-terminal domains of SspB and SpaP, AgI/II type proteins from *S. gordonii* and *S. mutans*, respectively, revealed tightly bound Ca^{2+} ions in both the C_2- and C_3-domains [5,7,8]. Also in the AspA-C_{2-3} structure Ca^{2+} binding can be observed (Fig. 3B), but only in the C_2-domain, where a single Ca^{2+} ion is coordinated by D1029 (OD1) and Y1030 (O) from the strands C–D loop, V1084 (O) and E1086 (O) from the loop connecting helix DH1 with strand D″, and one water molecule. The position of the Ca^{2+} ion correlates well with that observed in the SspB-C_{2-3} and SpaP-C_{2-3} structures, and the binding residues on the strands C–D loop are highly conserved.

3.4. Comparative structural and evolutionary conservation analyses

While *S. mutans* SpaP-C_{2-3} and *S. gordonii* SspB-C_{2-3} share 64% sequence identity (for 332 aligned amino acids), Asp-C_{2-3} only shares 36% (326 aligned amino acids) and 34% (317 aligned amino acids) sequence identity with SpaP-C_{2-3} and SspB-C_{2-3}, respectively. Despite the low overall sequence identity the structure of AspA-C_{2-3} is very similar to that of *S. mutans* SpaP-C_{2-3} (r.m.s.d. 1.84 Å) and *S. gordonii* SspB-C_{2-3} (r.m.s.d 1.9 Å) (Fig. 4A). The main region of difference is found in and in proximity to the α-helix (DH1) lying perpendicular to the C_2-domain central DEv-IgG motif. In *S. gordonii* SspB this region has been described as a recognition handle for the short fimbria Mfa1 from the periodontal pathogen *Porphyromonas gingivalis* and is referred to as BAR (for SspB Adherence Region) [23]. Within BAR, two structural motifs, corresponding to the DH1-helix and the loop region following it, have been identified as being important for *P. gingivalis* binding [5,24]. Interestingly, although the structures of SpaP-C_{2-3} and SspB-C_{2-3} are highly similar, also in secondary structure in BAR [7], *P. gingivalis* does not bind to *S. mutans* SpaP [23]. It rather appears that the binding is dependent on the specific composition of BAR, perhaps especially on the surface charge distribution around the α-helix. The AspA-C_{2-3} BAR helix also has a variable amino acid sequence, (DDKLKALIKAS, as compared to KKVQDLLKK in SspB and QEIRDVLSK in SpaP) which gives BAR a different surface charge distribution pattern compared to both SspB and SpaP (Fig. 4B). Also in the AspA-C_{2-3} structure, residues D1066 and T1067 produce a distinct bulge in the short loop preceding the BAR helix, which protrudes from the protein surface. The interactions in which BAR of AspA may play a role and how these structural features may be involved remains unknown and requires further study. Of additional note is that all three AgI/II type proteins for which the structure of the C_{2-3} domain has been determined so far (AspA, SspB, SpaP) have bound Ca^{2+} ions located after the short loop following the DH1 α-helix, which are thought to stabilize the position of the helix. In order to better understand the differences and variability among the C-terminal domains of the AgI/II type proteins, the AspA-C_{2-3} structure was subjected to evolutionary conservation analysis using ConSurf [16,25]. For the ConSurf analysis 51 homologous sequences were collected using a BLAST search and subsequently used for

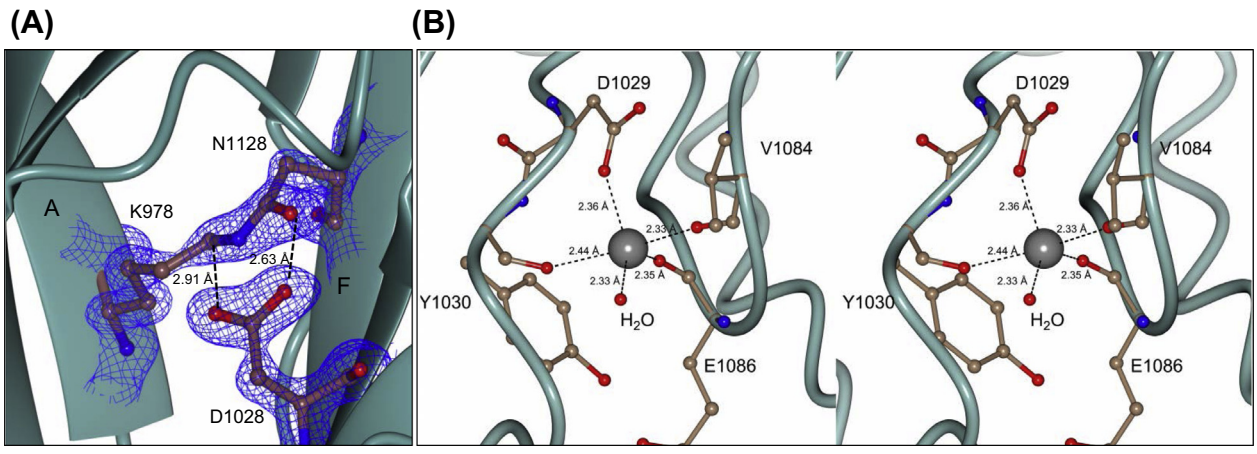

Fig. 3. Isopeptide bonds and metal binding site. (A) The isopeptide bond between K978 on strand A and N1128 on strand F in the C_2 domain is represented as a stick model in a 2Fo-Fc map, contoured at 0.90 e/Å^3. It is stabilized by hydrogen bonding with the side chain of D1028 (dashed lines). (B) In the C_2 domain a metal ion, modeled as a Ca^{2+} ion, is coordinated by D1029 and Y1030 on the strands C-D loop, V1084 and E1086 from the loop connecting helix DH1 with strand D″, and one water molecule. The figure is presented in stereo.

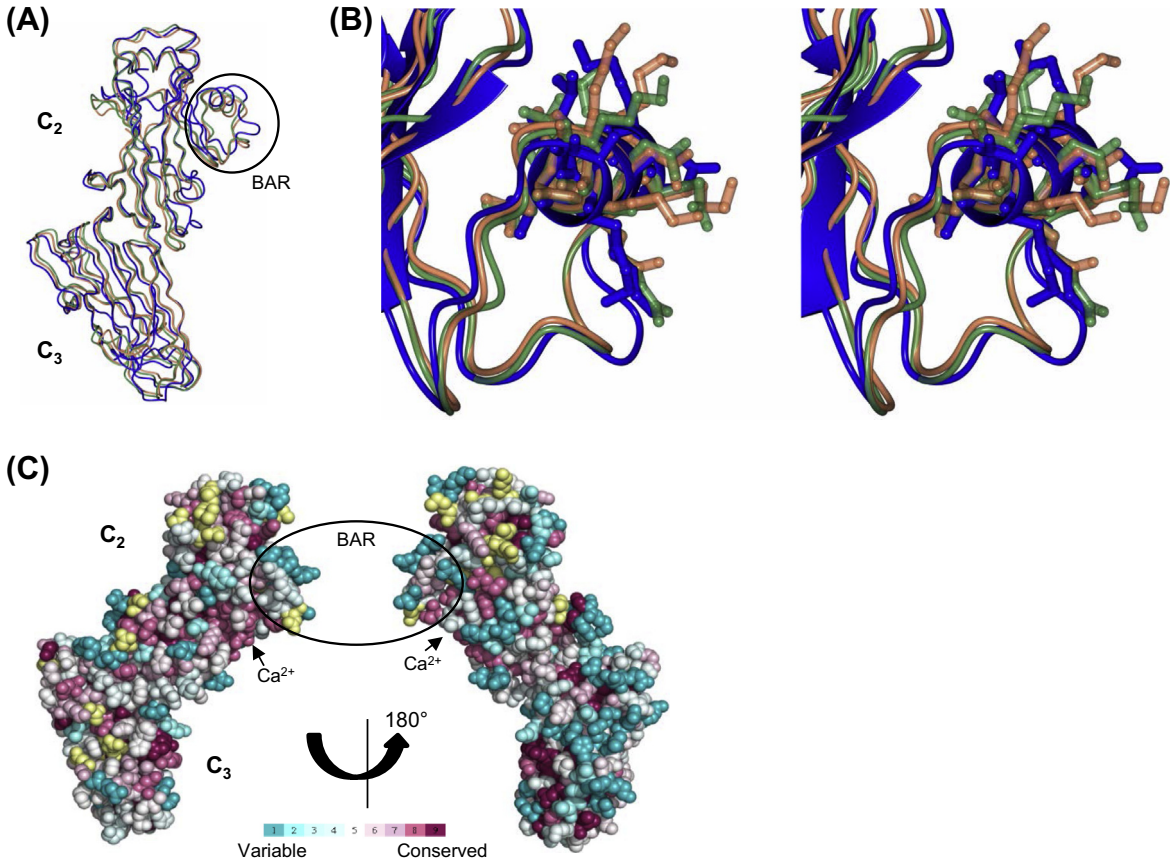

Fig. 4. Comparative structural and evolutionary conservation analyses of AspA-C_{2-3}. (A) Superposition of the AspA-C_{2-3} structure (blue) with the structures of SspB-C_{2-3} from *S. gordonii* (orange, PDB: 2WOY) and SpaP-C_{2-3} from *S. mutans* (green, PDB: 3OPU). The BAR (SspB Adherence Region), recognized by *P. gingivalis* short fimbrial protein Mfa1, is highlighted by a black circle and is shown in more detail as a stereo image in (B) together with a sequence alignment of the BAR helix. (C) Space filling model representation of evolutionary conservation analyses for AspA-C_{2-3} performed using ConSurf [25]. The level of conservation of individual amino acids is indicated from variable (turquoise, 1) to highly conserved (maroon, 9) according to the color coding bar. Positions for which the level of conservation was assigned with low confidence are marked with yellow color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

multiple sequence alignment and evolutionary conservation analysis. The level of conservation of each amino acid, on a scale from 1–9 (where 1 = variable and 9 = very highly conserved), was mapped onto the structural model (Fig. 4C). In good agreement with the observations discussed above, the conservation analysis shows that BAR, and the BAR helix in particular, is highly variable, while the closely located Ca^{2+} binding pocket in contrast is well conserved. Generally the β-sheet residues with side chains pointing to the interior of the central DEv-IgG motifs in each domain, constituting the hydrophobic core, are well conserved, while those with surface exposed side chains are variable. This is especially evident for one side of the C_3 domain (Fig. 4C, right).

4. Conclusions

In conclusion we have here presented the 1.8 Å crystal structure of the C_2 and C_3 domains of the AgI/II type adhesin protein AspA from *S. pyogenes*. This is the first part of the AspA protein to be structurally characterized and this work, together with successful structure determination of the other domains, especially of the variable (V) domain, will give valuable insights into the molecular mechanisms underlying adhesion and infection by *S. pyogenes*.

Acknowledgements

This work was supported by the Swedish Research Council Grant number 2011-4186, Umeå Centre for Microbial Research and the National Research School in Odontology. We thank Gunter Stier, EMBL, Germany for cloning vectors.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fob.2014.02.012.

References

[1] Courtney, H.S., Ofek, I., Penfound, T., Nizet, V., Pence, M.A., Kreikemeyer, B., Podbielski, A., Hasty, D.L. and Dale, J.B. (2009) Relationship between expression of the family of M proteins and lipoteichoic acid to hydrophobicity and biofilm formation in *Streptococcus pyogenes*. PLoS One 4, e4166.
[2] Nakata, M., Koeller, T., Moritz, K., Ribardo, D., Jonas, L., McIver, K.S., Sumitomo, T., Terao, Y., Kawabata, S., Podbielski, A. and Kreikemeyer, B. (2009) Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. Infect. Immun. 77, 32–44.
[3] Manetti, A.G.O., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., Gambellini, G., Bensi, G., Mora, M., Edwards, A.M., Musser, J.M., Graviss, E.A., Telford, J.L., Grandi, G. and Margarit, I. (2007) Streptococcus pyogenes pili promote pharyngeal cell adhesion and biofilm formation. Mol. Microbiol. 64, 968–983.
[4] Brady, L.J., Maddocks, S.E., Larson, M.R., Forsgren, N., Persson, K., Deivanayagam, C.C. and Jenkinson, H.F. (2010) The changing faces of streptococcus antigen I/II polypeptide family adhesins. Mol. Microbiol. 77, 276–286.
[5] Forsgren, N., Lamont, R.J. and Persson, K. (2010) Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. J. Mol. Biol. 397, 740–751.
[6] Larson, M.R., Rajashankar, K.R., Patel, M.H., Robinette, R.A., Crowley, P.J., Michalek, S., Brady, L.J. and Deivanayagam, C. (2010) Elongated fibrillar

structure of a streptococcal adhesin assembled by the high-affinity association of alpha- and PPII-helices. Proc. Natl. Acad. Sci. USA 107, 5983–5988.
[7] Nylander, A., Forsgren, N. and Persson, K. (2011) Structure of the C-terminal domain of the surface antigen SpaP from the caries pathogen *Streptococcus mutans*. Acta Crystallogr. F 67, 23–26.
[8] Larson, M.R., Rajashankar, K.R., Crowley, P.J., Kelly, C., Mitchell, T.J., Brady, L.J. and Deivanayagam, C. (2011) Crystal structure of the C-terminal region of *Streptococcus mutans* antigen I/II and characterization of salivary agglutinin adherence domains. J. Biol. Chem. 286, 21657–21666.
[9] Troffer-Charlier, N., Ogier, J., Moras, D. and Cavarelli, J. (2002) Crystal structure of the V-region of *Streptococcus mutans* antigen I/II at 2.4 angstrom resolution suggests a sugar preformed binding site. J. Mol. Biol. 318, 179–188.
[10] Maddocks, S.E., Wright, C.J., Nobbs, A.H., Brittan, J.L., Franklin, L., Stromberg, N., Kadioglu, A., Jepson, M.A. and Jenkinson, H.F. (2011) Streptococcus pyogenes antigen I/II-family polypeptide AspA shows differential ligand-binding properties and mediates biofilm formation. Mol. Microbiol. 81, 1034–1049.
[11] Long, F., Vagin, A.A., Young, P. and Murshudov, G.N. (2008) BALBES: a molecular-replacement pipeline. Acta Crystallogr. D 64, 125–132.
[12] Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J. and Adams, P.D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D 64, 61–69.
[13] Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60, 2126–2132.
[14] Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D 68, 352–367.
[15] Ashkenazy, H., Erez, E., Martz, E., Pupko, T. and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 38, W529–W533.
[16] Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A. and Noble, M. (2004) Developments in the CCP4 molecular-graphics project. Acta Crystallogr. D 60, 2288–2294.
[17] L.L.C. DeLano, The PyMol Molecular Graphics System, Version 0.99rc6 (2006).
[18] Deivanayagam, C.C., Wann, E.R., Chen, W., Carson, M., Rajashankar, K.R., Höök, M. and Narayana, S.V. (2002) Embo J. 21, 6660–6672.
[19] Kang, H.J., Coulibaly, F., Clow, F., Proft, T. and Baker, E.N. (2007) Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science 318, 1625–1628.
[20] Kang, H.J., Paterson, N.G., Gaspar, A.H., Ton-That, H. and Baker, E.N. (2009) The Corynebacterium diphtheriae shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. Proc. Natl. Acad. Sci. USA 106, 16967–16971.
[21] Persson, K., Esberg, A., Claesson, R. and Stromberg, N. (2012) The pilin protein FimP from *Actinomyces* oris: crystal structure and sequence analyses. PloS One 7, e48364.
[22] Budzik, J.M., Marraffini, L.A., Souda, P., Whitelegge, J.P., Faull, K.F. and Schneewind, O. (2008) Amide bonds assemble pili on the surface of bacilli. Proc. Natl. Acad. Sci. USA 105, 10215–10220.
[23] Brooks, W., Demuth, D.R., Gil, S. and Lamont, R.J. (1997) Identification of a *Streptococcus gordonii* SspB domain that mediates adhesion to *Porphyromonas gingivalis*. Infect. Immun. 65, 3753–3758.
[24] Daep, C.A., James, D.M., Lamont, R.J. and Demuth, D.R. (2006) Structural characterization of peptide-mediated inhibition of *Porphyromonas gingivalis* biofilm formation. Infect. Immun. 74, 5756–5762.
[25] Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. and Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 33, W299–302.