Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

A crystallizable form of the *Streptococcus gordonii* surface antigen SspB C-domain obtained by limited proteolysis

Nina Forsgren,[a] Richard J. Lamont[b] and Karina Persson[a]*

[a]Department of Odontology, Umeå University, Umeå SE-90 187, Sweden, and [b]Department of Oral Biology, University of Florida College of Dentistry, Gainesville, FL 32610, USA

Correspondence e-mail: karina.persson@odont.umu.se

Received 16 April 2009
Accepted 3 June 2009

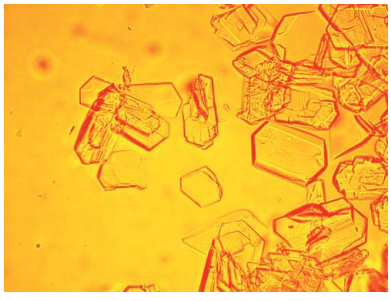
SspB is a 1500-residue adhesin expressed on the surface of the oral bacterium *Streptococcus gordonii*. Its interaction with other bacteria and host cells initiates the development of dental plaque. The full-length C-terminal domain of SspB was cloned, overexpressed in *Escherichia coli* and purified. However, the protein could not be crystallized. Limited proteolysis of the full-length C-domain identified a core fragment. The proteolysis product was cloned, expressed and purified. The protein was crystallized using the hanging-drop vapour-diffusion method. X-ray data were collected and processed to a maximum resolution of 2.1 Å with 96.4% completeness. The crystals belonged to space group $P2_1$, with one molecule in the asymmetric unit, a solvent content of 33.7% and a corresponding Matthews coefficient of 1.85 Å^3 Da^{-1}.

1. Introduction

Dental plaque is a complex mixed-species biofilm that can consist of over 700 species of bacteria (Jenkinson & Lamont, 2005). Gram-positive bacteria such as oral streptococci and *Actinomyces* spp. are primary colonizers that initiate formation of the oral biofilm. The initial bacteria provide adhesion sites for secondary colonizers such as *Fusobacterium nucleatum* and the periodontal pathogen *Porphyromonas gingivalis*.

Among the proteins involved in the interactions between streptococci and their binding partners are the antigen I/II (AgI/II) family of proteins. The AgI/II family is a family of well conserved surface adhesins that are expressed by virtually all species of oral streptococci. The precursors are between 1310 and 1653 amino acids long and contain seven discrete regions based on primary sequence. These include a signal peptide, an N-terminal region, an alanine-rich domain, a variable domain, a proline-rich domain, a C-terminal domain and a cell-wall anchoring region (Jenkinson & Demuth, 1997; Fig. 1). The AgI/II protein SspB from *Streptococcus gordonii* mediates a wide range of interactions with host proteins and other bacteria (Demuth *et al.*, 1996; Egland *et al.*, 2001; Holmes *et al.*, 1998).

Dental plaque, mainly consisting of Gram-positive and commensal bacteria, can exist in the oral cavity without causing any disease. Nevertheless, shifts in the biofilm community can increase the number of acidophiles and Gram-negative bacteria that cause common oral diseases such as caries and periodontitis (Dalwai *et al.*, 2006). *P. gingivalis* is a Gram-negative pathogen associated with adult periodontitis and is mostly found in an anaerobic environment.



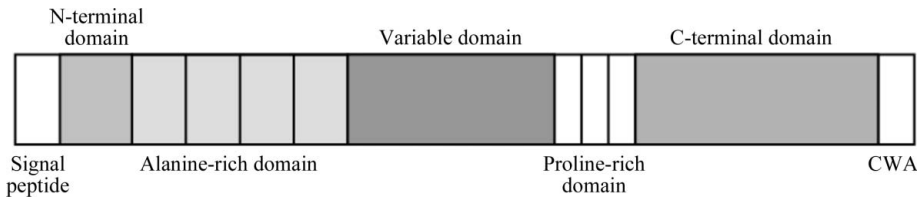


Figure 1
Domain architecture of the AgI/II proteins: a signal peptide, an N-terminal region, an alanine-rich repeat domain, a variable domain, a proline-rich repeat domain, a C-terminal domain and a cell-wall anchoring segment (CWA).

However, colonization of the oral cavity by *P. gingivalis* is likely to occur in supragingival areas where oxygen tension has been reduced by already established biofilm communities (Bradshaw *et al.*, 1998). *P. gingivalis* has been shown to adhere to the *S. gordonii* AgI/II protein SspB (Brooks *et al.*, 1997; Demuth *et al.*, 2001) *via* its minor fimbria, Mfa1. The SspB–Mfa1 interaction is essential for the development of *P. gingivalis* biofilms on a streptococcal substrate (Demuth *et al.*, 2001) and is considered to be an important target for the development of therapeutic agents to prevent *P. gingivalis* attachment. The SspB region that mediates adhesion to Mfa1 has been identified and comprises residues 1167–1193 (designated BAR for Ssp**B** **a**dherence **r**egion). The BAR motif is located within the C-terminal domain of SspB (Demuth *et al.*, 2001; Brooks *et al.*, 1997).

The three-dimensional structure of the SspB C-terminal domain will provide valuable information regarding the interacting epitopes of SspB and may also be the first step towards the design of potential agents that target *P. gingivalis* adherence. Hence, the present study targets the cloning, overexpression, purification, crystallization and preliminary X-ray diffraction studies of the C-terminal domain of *S. gordonii* SspB. The cloned fragment represents residues 1061–1413, with a molecular weight of 39.3 kDa.

2. Materials and methods

2.1. Cloning

The C-terminal domain of SspB was cloned from plasmid pEB5 (Demuth *et al.*, 1988) encoding the *sspB* gene (GenBank U40026). PCR primers were designed based on the domain arrangement suggested for the AgI/II proteins (Jenkinson & Demuth, 1997). The forward primers were 5′-TTTTTCCATGGTTCATTTCCACTATAGCAGT-3′ (full length) and 5′-TTTTTCCATGGTAACTACTCCTGGTAAA-3′ (proteolytic fragment). The reverse primer was 5′-AAAAAGGTACCTTAAGTTGTTGTACGTACTGTGTT-3′. The PCR product was digested with *Acc*65I and *Nco*I and ligated into the equivalent sites of the pET-M11 expression vector (kindly provided by G. Stier, EMBL, Germany). The final constructs encode His_6-PMSNTNIPTTENLYPEGAM-SspB912–1413 and His_6-PMSNTNIPTTENLYPEGAM-SspB1061–1413, respectively. The plasmids were transformed into *Escherichia coli* DH5α and subsequently selected on kanamycin plates. The positive clones were verified by DNA sequencing.

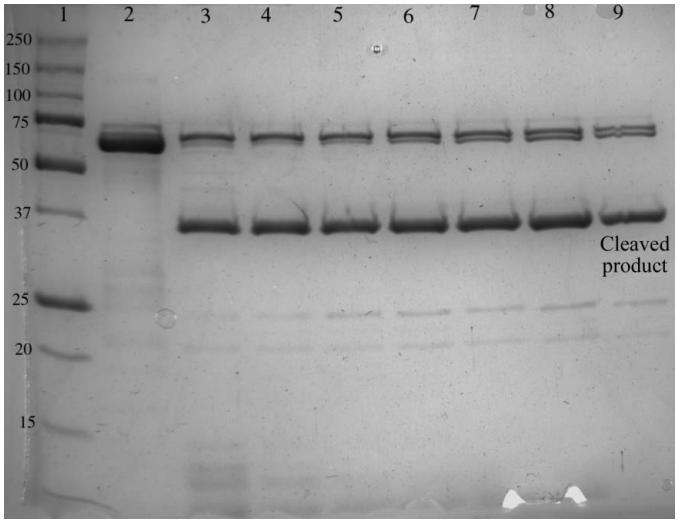


Figure 2
Limited proteolysis of the full-length SspB C-domain. Lane 1, molecular-weight markers (kDa); lane 2, untreated protein; lanes 3–9, after 2, 5, 10, 20, 30, 60 and 90 min incubation with trypsin, respectively.

2.2. Overexpression and purification

The constructs were 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 expression was induced with 0.4 m*M* IPTG, after which the cultures were grown for an additional 5 h. Cells were harvested by centrifugation at 5300*g* and the pellets were frozen at 193 K. The pellets were resuspended in 20 m*M* Tris pH 7.5, 150 m*M* NaCl and 10 m*M* 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 SspB C-domain was washed in 20 m*M* Tris pH 7.5, 150 m*M* NaCl and 20 m*M* imidazole and eluted in 20 m*M* Tris pH 7.5, 150 m*M* NaCl and 300 m*M* imidazole. The buffer was exchanged to 20 m*M* Tris pH 7.5, 150 m*M* NaCl, 0.5 m*M* EDTA and 1 m*M* DTT. The protein was further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 prep-grade column (Amersham Biosciences). The protein purity was judged by SDS–PAGE. The proteolytic and the full-length constructs were both purified following the same protocol; however, only the smaller form resulted in a single peak from gel filtration.

2.3. Limited proteolysis and N-terminal sequencing

The full-length SspB C-domain in 20 m*M* $CaCl_2$, 50 m*M* Tris–HCl pH 8.0, 1% glycerol and 2 m*M* DTT was incubated with trypsin at a ratio of 1:100 trypsin:protein at 310 K for 2.5 h. Cleavage was monitored by SDS–PAGE (Fig. 2). Western blotting was used to transfer the cleavage product to a PVDF membrane. N-terminal sequencing of the cleavage product was performed by the Protein Analysis Center (PAC) at Karolinska Institutet, Sweden.

2.4. Crystallization

The recombinant proteolytic fragment was concentrated to 10 mg ml^{-1} in 20 m*M* Tris pH 7.5, 0.5 m*M* EDTA and 1 m*M* DTT using an Amicon Ultra centrifugal filter device (Millipore). Initial crystallization trials were performed by the sitting-drop vapour-diffusion method in a 96-well MRC crystallization plate. Droplets of 1 µl protein solution were mixed with an equal volume of reservoir solution using screens from Hampton Research (Crystal Screen and Crystal Screen II) and Molecular Dimensions (PACT I and II). Crystal quality was improved by the addition of 1%(*w*/*w*) α-chymotrypsin (equivalent to 4 units ml^{-1}) to the protein prior to crystallization as described by Dong *et al.* (2007). The content of the crystals were analyzed by dissolving them and running them on a SDS–PAGE gel and by mass spectrometry.

2.5. Data collection and processing

Crystals were soaked in crystallization solution supplemented with 20% glycerol and flash-cooled in a liquid-nitrogen stream. Diffraction data were collected on a MAR 165 detector on beamline I911-2 at the MAX-lab synchrotron, Lund to a maximum resolution of 2.1 Å. A total of 180 frames of data were collected with an oscillation angle of 1°, an exposure time of 20 s and a crystal-to-detector distance of 150 mm. The diffraction data were processed with *MOSFLM* (Leslie,

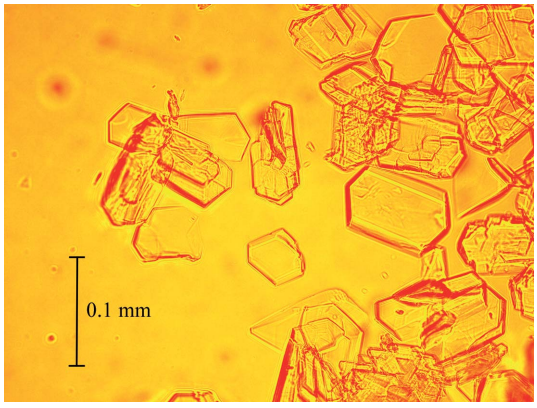


Figure 3
Crystals of the truncated SspB C-domain obtained in 0.2 *M* $MgCl_2$, 0.1 *M* Tris–HCl pH 8.5 and 30%(*w*/*v*) PEG 4000 with the addition of 1% α-chymotrypsin.

2006) and scaled with *SCALA* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

The full-length SspB C-terminal domain, comprising residues 912–1413 (theoretical molecular weight 55.6 kDa), was successfully cloned and overexpressed in *E. coli* BL21 (DE3). The protein did not purify to homogeneity during gel filtration and no crystals were obtained. Limited proteolysis was performed and the cleavage product was N-terminally sequenced. A new construct comprising residues 1061–1413 of SspB was made based on the N-terminal sequencing result. The new C-terminal protein (theoretical molecular weight 39.3 kDa) eluted as a monomer after size-exclusion chromatography. Initially, thin plate-shaped protein crystals that were stacked together were obtained, which resulted in a diffraction pattern that was difficult to process. By varying the pH, ionic strength, precipitant concentration and temperature, crystals were obtained from 0.2 *M* $MgCl_2$, 0.1 *M* Tris–HCl pH 8.5 and 30%(*w*/*v*) PEG 4000 at room temperature. The crystals were further improved by *in situ* proteolysis (Dong *et al.*, 2007) by the addition of 1%(*w*/*w*) α-chymotrypsin to the protein immediately before crystallization setup. The optimization procedures resulted in single well diffracting crystals that grew to dimensions of 0.1 × 0.08 × 0.02 mm within a week (Fig. 3). MALDI–TOF mass spectrometry of washed dissolved crystals determined the molecular weight of the crystallized protein as 36.8 kDa (the full-length construct is 39.3 kDa). X-ray diffraction data were collected to a resolution of 2.1 Å with 96.4% completeness (76% in the outer shell) and an $R_{\rm merge}$ of 5.2%. Analysis of the symmetry and systematic absences in the recorded diffraction pattern indicated that the crystal belonged to the monoclinic space group $P2_1$, with unit-cell parameters a = 35.7, b = 48.9, c = 84.1 Å, β = 100.1°. A calculated Matthews coefficient of 1.85 $Å^3\,Da^{-1}$ suggested a 33.7% solvent content with one monomer in the asymmetric unit (Matthews, 1968). The data-collection and processing statistics are given in Table 1.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the last resolution shell.

Wavelength (Å)	1.043
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 35.7, b = 48.9, c = 84.1, β = 100.1
Resolution range (Å)	19.32–2.1 (2.21–2.1)
No. of observed reflections	54592
No. of unique reflections	16254
Completeness (%)	96.4 (76.0)
Redundancy	3.4 (2.9)
$R_{\rm merge}$† (%)	5.2 (15.1)
$\langle I/\sigma(I)\rangle$	21.7 (7.3)
No. of molecules per ASU	1
Matthews coefficient ($Å^3\,Da^{-1}$)	1.85

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the average over of all observations of reflection hkl.

We have recently obtained and collected data from selenomethionine-labelled crystals of the SspB C-domain. Structure determination and refinement is under way and will be published elsewhere.

We are grateful for access to beamline I911-2 at MAX-lab, Lund and for the support of the beamline scientists. We thank Gunter Stier, EMBL, Germany for cloning vectors and Thomas Kieselbach, Umeå University for mass-spectrometric analyses. This work was supported by the Swedish Research Council, the Magn Bergvall foundation and DE12505 from the NIDCR.

References

Bradshaw, D. J., Marsh, P. D., Watson, G. K. & Allison, C. (1998). *Infect. Immun.* **66**, 4729–4732.
Brooks, W., Demuth, D. R., Gil, S. & Lamont, R. J. (1997). *Infect. Immun.* **65**, 3753–3758.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dalwai, F., Spratt, D. A. & Pratten, J. (2006). *Appl. Environ. Microbiol.* **72**, 3678–3684.
Demuth, D. R., Davis, C. A., Corner, A. M., Lamont, R. J., Leboy, P. S. & Malamud, D. (1988). *Infect. Immun.* **56**, 2484–2490.
Demuth, D. R., Duan, Y., Brooks, W., Holmes, A. R., McNab, R. & Jenkinson, H. F. (1996). *Mol. Microbiol.* **20**, 403–413.
Demuth, D. R., Irvine, D. C., Costerton, J. W., Cook, G. S. & Lamont, R. J. (2001). *Infect. Immun.* **69**, 5736–5741.
Dong, A. *et al.* (2007). *Nature Methods*, **4**, 1019–1021.
Egland, P. G., Du, L. D. & Kolenbrander, P. E. (2001). *Infect. Immun.* **69**, 7512–7516.
Holmes, A. R., Gilbert, C., Wells, J. M. & Jenkinson, H. F. (1998). *Infect. Immun.* **66**, 4633–4639.
Jenkinson, H. F. & Demuth, D. R. (1997). *Mol. Microbiol.* **23**, 183–190.
Jenkinson, H. F. & Lamont, R. J. (2005). *Trends Microbiol.* **13**, 589–595.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.