Journal of Structural Biology xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Structural Biology

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

Structure Report

An arm-swapped dimer of the *Streptococcus pyogenes* pilin specific assembly factor SipA

Paul G. Young *, Hae Joo Kang [1], Edward N. Baker

School of Biological Sciences, University of Auckland, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 26 April 2013
Received in revised form 28 May 2013
Accepted 30 May 2013
Available online xxxx

Keywords:
Streptococcus pyogenes
SipA
LepA
Pilus assembly
Signal peptidase

ABSTRACT

Streptococcus pyogenes (group A streptococcus [GAS]) is a major human pathogen. Attachment of GAS to host cells depends in large part on pili. These assemblies are built from multiple covalently linked subunits of a backbone protein (FctA), which forms the shaft of the pilus, and two minor pilin proteins, FctB anchoring the pilus to the cell wall and Cpa functioning as the adhesin at the tip. Polymerisation of the pilin subunits is mediated by a specific sortase, which catalyzes the formation of peptide bonds linking successive subunits. An additional gene, SipA, is also essential for GAS pilus polymerisation, but its function remains undefined. Here we report the crystal structure of a truncated SipA protein from GAS, determined at 1.67 Å resolution. The structure reveals that SipA has the same core fold as the *Escherichia coli* type-I signal peptidase (SPase-I), but has a much smaller non-catalytic domain. The truncated protein, which lacks 9 N-terminal residues, forms an arm-swapped dimer in which the C-terminal β-strand of each monomer crosses over to interact with an N-terminal strand from the other monomer. In addition, there is no peptide binding cleft and significant differences in the putative membrane association region.

1. Introduction

The gram-positive pathogen *Streptococcus pyogenes* (Group A Streptococcus [GAS]) is a common human pathogen causing a spectrum of diseases ranging from mild localized infections such as strep throat, through to serious illnesses such as scarlet fever, rheumatic fever, and pneumonia (Cunningham, 2000). Recently it has been discovered that GAS produces pili on its surface (Mora et al., 2005). These pili are instrumental in mediating attachment of GAS to host cells and subsequent disease development (Abbot et al., 2007; Manetti et al., 2007). Pili on GAS are composed of multiple covalently-linked subunits of a major backbone pilin (FctA), and two minor pilin proteins, FctB and Cpa (Mora et al., 2005). The major pilin forms the polymeric backbone of the pilus (Kang et al., 2007; Mora et al., 2005), whereas the minor pilin Cpa forms the adhesin at the tip of the pilus (Quigley et al., 2009; Smith et al., 2010) and the basal pilin FctB covalently links the pilus to peptidoglycan of the cell wall (Hendrickx et al., 2011; Linke et al., 2010; Smith et al., 2010). Polymerisation of the pilin subunits is mediated by a specific sortase (SrtC), which catalyzes the formation of an isopeptide bond that joins one subunit to the next (Hendrickx et al., 2011; Kang et al., 2007; Mora et al., 2005). A second gene (*sipA*) has also shown to be essential for pilus polymerisation (Nakata et al., 2009; Zahner and Scott, 2008). The SipA protein (also known as LepA) has significant sequence similarity with type-I signal peptidases, which are membrane-bound serine proteases that cleave the N-terminal signal sequence from secreted proteins (Dalbey et al., 1997; Paetzel et al., 1998). These enzymes have a characteristic serine–lysine catalytic dyad, in which the serine acts as the nucleophile while the amino group of lysine provides the general base that deprotonates the serine hydroxyl group (Paetzel et al., 1998). Most bacteria typically have only one active signal peptidase, which is essential for growth and survival (Dalbey and Wickner, 1985; Inada et al., 1989). However, some gram-positive bacteria have several signal peptidases that appear to have overlapping sequence specificities. In addition, genomic sequencing has identified a growing number of gram-positive signal peptidase-like proteins that are predicted to have the same protein architecture but lack an identifiable catalytic dyad.

Here we report the first structure from this family of non-active signal peptidases. This truncated SipA protein, which lacks 9 N-terminal residues that were removed to enhance solubility, shows 23% sequence identity with the *Escherichia coli* type-I signal peptidase (SPase-I). The structure confirms that SipA shares the core SPase-I fold, but reveals significant differences in the catalytic domain, loss of the peptide binding cleft present in *E. coli* SPase-I, and concomitant formation of an arm-swapped dimer.

* Corresponding author. Address: Structural Biology, School of Biological Sciences, University of Auckland, 3a Symonds St., Auckland 1010, New Zealand.
E-mail address: p.young@auckland.ac.nz (P.G. Young).
[1] Present address: Membrane Protein Crystallography Group, Division of Molecular Biosciences, Imperial College London, London SW7 2AZ, UK.

1047-8477/$ - see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.jsb.2013.05.021

2. Cloning, expression and purification of SipA

The *sipA* gene comprising the entire extracellular region of the protein ($SipA_{36–173}$) and a truncated version lacking 9 membrane-proximal residues 36–44 ($SipA_{45–173}$) were PCR-amplified from *S. pyogenes* strain 90/306S genomic DNA. For simplicity $SipA_{36–173}$ and the truncated $SipA_{45–173}$ will be referred to as $SipA_{WT}$ and $SipA_{\Delta 9}$, respectively. $SipA_{WT}$ and $SipA_{\Delta 9}$ were amplified using the gene specific primers SPY0127 F1 5′- AAA **GGCGCC** CAG TAT GTT TTT GGT GTT ATG ATT A -3′ ($SipA_{WT}$) or SPY0127 F2 5′- AAA **GGCGCC** AAC ACT AAT GAT ATG AGT CCT GCT TTA AG -3′ ($SipA_{\Delta 9}$) and SPY0127 R1 5′- AAA **GAATTC** TTA AAT TCC TCT CAC TCT TAA TAG AGT TGA G -3′ (5′ *KasI* and 3′ *EcoRI* restriction nuclease recognition sites are shown in bold). The amplified fragments were cloned into the vector pProEXHTa (Invitrogen).

For recombinant SipA protein expression, *E. coli* BL21 (λDE3) pRIL cells were transformed with SipA plasmids and grown in Luria–Bertani (LB) media supplemented with the required antibiotics at 310 K until OD_{600} reached 0.6. The cultures were induced with 0.2 mM IPTG at 301 K or 291 K for 16 h and the cells harvested by centrifugation. Cell pellets were resuspended buffer A (50 mM Tris.Cl pH 8.0, 500 mM NaCl, 2% (v/v) glycerol, 5 mM imidazole) containing Complete Protease Inhibitor Cocktail Mini Tablets EDTA-free (Roche), snap frozen, and stored at 253 K.

The recombinant proteins were purified from frozen cells, which were thawed in buffer A with Complete Protease Inhibitor Cocktail Mini Tablets EDTA-free (Roche) and 2 μg/ml DNase I, and then lysed using a cell disruptor (Constant Cell Disruption Systems) at 18 kpsi. Insoluble matter was sedimented by centrifugation and the soluble phase was loaded onto a HiTrap Chelating 5 ml column (GE Healthcare) for purification by IMAC. Bound protein was washed with buffer B (buffer A + 20 mM imidazole) and eluted in a gradient with buffer C (50 mM Tris.Cl pH 8.0, 150 mM NaCl, 2% (v/v) glycerol, 500 mM imidazole). The His_6 affinity tag was cleaved from $SipA_{\Delta 9}$ recombinant protein with a 1:50 ratio of rTEV-His_6 and concurrently dialyzed against buffer D (25 mM Tris.Cl pH 7.5, 150 mM NaCl, 0.05 mM EDTA) at 277 K for 16 h. $SipA_{\Delta 9}$ was separated from the rTEV-His_6 protease and uncleaved protein by IMAC. The unbound protein containing $SipA_{\Delta 9}$ was concentrated using a 3 kDa MWCO protein concentrator (VivaScience) and purified by size exclusion chromatography on a Superdex S75 10/300 column (GE healthcare) in crystallization buffer (10 mM Tris.Cl pH 8.0, 100 mM NaCl). $SipA_{\Delta 9}$ eluted in a single peak that corresponds to a dimer of approximately 30 kDa and was 99% pure as indicated by SDS–PAGE. Dynamic light scattering (DLS) data confirmed the protein was monodisperse with a radius of gyration that equates to a molecular weight of 30 kDa, in agreement with the size exclusion chromatography. In contrast, the $SipA_{WT}$ construct produced aggregated protein that could not be cleaved from the His_6-tag and predominately eluted in the void volume with size exclusion chromatography.

3. Crystallization

Vapour diffusion crystallization trials were carried out at 291 K using a Cartesian nanolitre dispensing robot (Genomic systems) and a locally compiled crystallization screen (Moreland et al., 2005). Initial $SipA_{\Delta 9}$ crystals were grown in 0.1 μl format and subsequently optimised in a hanging-drop vapour diffusion format. The crystals used for X-ray data collection grew by mixing 1 μl protein solution (10 mg/ml in 10 mM Tris.Cl pH 8.0, 100 mM NaCl) with 1 μl precipitant (15% ethanol, 0.1 M Tris.Cl pH 8.0) at 291 K.

4. Data collection and structure determination

Crystals of $SipA_{\Delta 9}$ were transferred to cryoprotectant (100 mM Tris.Cl pH 8, 50 mM NaCl, 15% ethanol, 20% ethylene glycol) prior to flash-freezing in liquid nitrogen. For phase determination experiments crystals were soaked in cryoprotectant supplemented with 500 mM NaI for 3 min before freezing.

X-ray diffraction data were collected in-house (Micromax-007HF, Rigaku; MAR345DTB, MAR Research) at 110 K. All datasets were integrated using XDS (Kabsch, 1993), reindexed using POINTLESS (Evans, 2006) and scaled using SCALA (Evans, 2006). The

Table 1
Data collection and refinement statistics.

Data collection	Native	NaI
Wavelength (Å)	1.54179	1.54179
Resolution range (Å)*	19.31–1.67 (1.76–1.67)	26.3–1.86 (1.96–1.86)
Space group	$P4_12_12$	$P4_12_12$
Unit cell axial lengths (Å)	$a = b = 53.28$, $c = 131.75$	$a = b = 53.79$, $c = 131.27$
Angles (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Total No. of observations*	808,555 (95,927)	402,042 (47,171)
Unique reflections*	22,923 (3178)	16,684 (2049)
Redundancy*	35.3 (30.2)	24.1 (23.0)
Completeness (%)*	99.4 (96.9)	97.7 (85.3)
Mean $I/\sigma(I)$*	60.8 (13.3)	35.1 (6.5)
R_{merge} (%)*,†	4.3 (26.2)	8.0 (57.9)
Anomalous completeness*		97.6 (84.2)
Anomalous multiplicity*		13.4 (12.5)
DelAnom correlation between half-sets		0.83
Mid-slope of anom normal probability		0.61
No. of heavy atoms		20
FOM (phaser EP acentric/centric)		0.43/0.14
Refinement		
Resolution range (Å)	19.3–1.67	
R_{work}/R_{free} (%)	18.5/21.4	
No. atoms per AU	1061	
Average B-factors (Å^2)	20.2	
RMS deviations		
Bond lengths	0.010	
Bond angles	1.35	
Ramachandran plot		
Residues in most favored regions, allowed, disallowed (%)	99.1, 0.9, 0	

† $R_{merge} = \Sigma_{hkl} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle | / \Sigma_{hkl} \Sigma_i I_i(hkl)$.
* Values in parentheses are for the outermost resolution shell.

Please cite this article in press as: Young, P.G., et al. An arm-swapped dimer of the *Streptococcus pyogenes* pilin specific assembly factor SipA. J. Struct. Biol. (2013), http://dx.doi.org/10.1016/j.jsb.2013.05.021