Chad A. Brautigam,^{a*} Ranjit K. Deka^b and Michael V. Norgard^b^aDepartment of Biophysics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA, and^bDepartment of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USACorrespondence e-mail:
chad.brautigam@utsouthwestern.edu

Received 5 February 2013

Accepted 4 March 2013

Purification, crystallization and preliminary X-ray analysis of TP0435 (Tp17) from the syphilis spirochete *Treponema pallidum*

Syphilis, caused by the bacterial spirochete *Treponema pallidum*, remains a prominent sexually transmitted infection worldwide. Despite sequencing of the genome of this obligate human pathogen 15 years ago, the functions of a large number of the gene products of *T. pallidum* are still unknown, particularly with respect to those of the organism's periplasmic lipoproteins. To better understand their functions, a structural biology approach has been pursued. To this end, the soluble portion of the *T. pallidum* TP0435 lipoprotein (also known as Tp17) was cloned, hyper-expressed in *Escherichia coli* and purified to apparent homogeneity. The protein crystals obtained from this preparation diffracted to 2.4 Å resolution and had the symmetry of space group R3. In the hexagonal setting, the unit-cell parameters were $a = b = 85.7$, $c = 85.4$ Å.

1. Introduction

Syphilis remains a prominent sexually transmitted infection worldwide (Gerbase *et al.*, 1998). It is caused by the bacterial spirochete *Treponema pallidum*, which is an obligate human pathogen. Although penicillin treatment is effective in eradicating the organism, the lack of a vaccine and the persistence of syphilis as a global public health problem (Chao *et al.*, 2006) motivates continued research on the disease and its etiologic agent.

The genome of *T. pallidum* has been sequenced (Fraser *et al.*, 1998). Perhaps the most surprising outcome of this effort was the fact that this organism has a modest protein-encoding capacity (~1100 genes). Indeed, this bacterium appears to lack genes for many key metabolic processes, including fatty-acid synthesis and the *de novo* syntheses of amino acids and many essential vitamins. *T. pallidum* is therefore dependent on its human host for most of its nutrients and it must import them from rather unaccommodating environments. Because essential nutrients must traverse the organism's dual membrane system, understanding the membrane biology of *T. pallidum* is likely to hold the key to elucidating salient aspects of the ability of the spirochete to sustain itself and thrive within its human host. Of note, the outer membrane of *T. pallidum* contains a paucity of proteins (Radolf, 1995), but there are a large number of periplasmic lipoproteins that appear to be lipid-anchored to the periplasmic leaflets of the cytoplasmic and outer membranes (Cox *et al.*, 2010). Their positions as intermediaries between the extracellular and intracellular environments place the lipoproteins in a strategic position to markedly influence the parasitic strategy of *T. pallidum*. However, the functions of the vast majority of these lipoproteins remain unknown.

Because *T. pallidum* cannot be continuously cultivated *in vitro* (Norris, 1993), it is not genetically manipulatable. Thus, researchers do not have access to a key tool for the elucidation of protein functions. In response, we have adopted a 'structure-to-function' approach for studying *T. pallidum* membrane lipoproteins. In this strategy, we computationally identify putative lipoproteins, then clone and hyper-express their soluble (non-acylated) domains in *Escherichia coli*. After purification, we crystallize them and determine their crystal structures. Often, their structures have revealed important clues to their functions, which then have prompted hypotheses that are subsequently tested with various biophysical, biochemical and bioinformatic methods (Deka *et al.*, 2006, 2007, 2012;

Machius *et al.*, 2007; Brautigam, Deka, Ouyang *et al.*, 2012; Brautigam, Deka, Schuck *et al.*, 2012).

The crystallization target described in this report is the product of the gene *tp0435*. The protein, TP0435 (also known as Tp17 in the literature), has some sequence homology (20% identity) to NlpE. In *E. coli*, this latter protein is anchored to the outer membrane and can activate the Cpx cell envelope stress response pathway (Snyder *et al.*, 1995; Otto & Silhavy, 2002). In *T. pallidum*, however, it is likely that TP0435 does not have this function, inasmuch as the genome of this organism apparently does not encode other members of the Cpx pathway; also, TP0435 is likely localized to the periplasm (*i.e.* not to the outer surface of the organism). As a starting point for the structural characterization of TP0435, we describe here its cloning, hyper-expression, purification and crystallization. The resulting crystals diffracted X-rays to a minimum Bragg spacing ($d_{\min}$) of 2.4 Å and exhibited the symmetry of space group R3.

2. Materials and methods

2.1. Cloning

To produce a non-lipidated recombinant derivative of TP0435 in *E. coli*, the DNA fragment encoding amino-acid residues 9–132 (cloned without the post-translationally modified N-terminal Cys plus seven other residues; the numbering reflects the assignment of this Cys as residue 1 of the processed protein) of TP0435 was PCR-amplified from *T. pallidum* genomic DNA using primer pairs encoding the predicted 5'- and 3'-termini. The PCR primers were 5'-ata**GGTCTCAAGGTGGGAAGGCCAAAGCGGAAAAG**-3' and 5'-ggaaggtctc**TCTAGATCACTATTTCTTTGTTTTTTGAGCAC**-3'. In the preceding sequences, the region of each primer complementary to the *tp0435* sequence is underlined. The forward primer contained both an ata overhang (lowercase) and a *BsaI* site (bold); the reverse primer contained a ggaaggtctc overhang (lowercase) and an *XbaI* site (bold). The PCR product was subcloned into a *BsaI/XbaI*-digested pE-SUMOpro-based bacterial expression vector (LifeSensors) in-frame with an N-terminal His-SUMO tag. The construct was confirmed by DNA sequencing.

2.2. Hyper-expression in *E. coli* and purification

The recombinant protein was hyper-expressed in *E. coli* BL21(DE3) cells using an overnight express auto-induction system 2 (Novagen). Cells derived from 1 l of culture were harvested by centrifugation and lysed at room temperature with gentle rocking for 20 min using 50 ml of B-PER II (Thermo Scientific). The resulting suspension was centrifuged at 25 000g for 15 min to remove cell

Figure 1
Crystals of TP0435. Shown are typical crystals of TP0435. Crystals can reach sizes of 500 µm (in all three dimensions).

debris. The recombinant TP0435 (rTP0435) was isolated from the supernatant by affinity chromatography using Ni-NTA agarose (Qiagen). The His-SUMO tag was removed by the SUMO-specific protease 1 (LifeSensors) and the digestion mixture was further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 preparative-grade column (GE Healthcare) equilibrated with buffer A (20 mM HEPES, 0.1 M NaCl, pH 7.5, 2 mM *n*-octyl β-D-glucoside). Peak fractions were analyzed by SDS-PAGE. Fractions containing purified rTP0435 were pooled and stored at 277 K in buffer A.

Protein concentrations were determined spectrophotometrically using Beer's law. The extinction coefficient of $11\,585\,M^{-1}\,cm^{-1}$ at 280 nm was calculated using the *ProtParam* utility of ExPASy (<http://us.expasy.org>). The yield of purified proteins tended to be about 10 mg per litre of bacterial culture.

2.3. Crystallization

TP0435 crystals were obtained using the hanging-drop vapor-diffusion method. Crystals of rTP0435 appeared after 3 weeks at 293 K from solutions containing 2 µl protein ($\sim 11\,mg\,ml^{-1}$ in buffer A) and 2 µl of precipitant solution consisting of 40%(v/v) PEG 400, 0.2 M lithium sulfate, 0.1 M Tris pH 8.5. The crystals had an irregular, multifaceted appearance (Fig. 1). Upon transfer to cryoprotectant [60%(v/v) PEG 400, 0.2 M lithium sulfate, 0.1 M Tris pH 8.5], the crystals sometimes cracked. A shard broken away from one of these cracked crystals was mounted in a nylon loop, flash-cooled in liquid nitrogen and stored under liquid nitrogen until ready for data collection.

2.4. Data collection and processing

X-ray diffraction data (Fig. 2) were collected on beamline 19-ID at the Structural Biology Center of the Advanced Photon Source at Argonne National Laboratory. The temperature of data collection was 100 K and the incident radiation had a wavelength of 0.97912 Å;

Figure 2
Diffraction data from crystals of TP0435. The diffraction limit of the data set, 2.4 Å, is shown as a black circle on the image.

Table 1

Data-collection statistics for TP0435.

Space group	<i>R</i> 3
Unit-cell parameters	
<i>a</i> (Å)	85.7
<i>b</i> (Å)	85.7
<i>c</i> (Å)	85.4
α (°)	90.0
β (°)	90.0
γ (°)	120.0
Wavelength (Å)	0.97912
Temperature (K)	100
Resolution range (Å)	37.8–2.40 (2.44–2.40)
Unique reflections	9033 (419)
Multiplicity	5.4 (4.2)
Data completeness (%)	99.2 (89.1)
$R_{\text{merge}}^{\dagger}$ (%)	4.8 (54.9)
$\langle I/\sigma(I) \rangle$	40.1 (1.75)
Wilson <i>B</i> value (Å ²)	68.1

$$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl).$$

data were collected from the single shard-like crystal. The data were integrated, merged and scaled using *HKL-2000* (Otwinowski & Minor, 1997). The *CTRUNCATE* procedure implemented in *CCP4* (Winn *et al.*, 2011) was used to place the reflections on an absolute scale and give all reflections non-negative values (French & Wilson, 1978). A search of a self-Patterson map to detect the presence of translational pseudosymmetry was performed in *phenix.xtriage* in *PHENIX* (Adams *et al.*, 2010) and the self-rotation function was examined using *MOLREP* from *CCP4* with data between 9 and 4 Å.

3. Results and discussion

The cloning and purification strategy employed for TP0435 produced soluble non-lipidated pure recombinant protein (rTP0435). The protein could be crystallized from a buffered basic (pH 8.5) solution containing PEG 400 and lithium sulfate (Fig. 1). Although these crystals have irregular shapes, we observed no signs of nonmerohedral twinning in the X-ray diffraction patterns resulting from the irradiation of a shard broken away from one of them (Fig. 2).

The diffraction patterns (see, for example, Fig. 2) obtained from this shard-like crystal displayed the symmetry of space group *R*3 (Table 1). A complete data set was collected; the d_{min} spacing was 2.4 Å. Fitting the high-resolution portion of a Wilson plot resulted in an estimated Wilson *B* factor of 68.1 Å². Several tests for twinning, including examining several acentric moments of *E* (Stanley, 1972; Yeates, 1997), the *L* test (Padilla & Yeates, 2003) and the cumulative intensity distribution did not reveal any evidence of pathologies in the data set. Given the unit-cell dimensions and the molecular mass of the protein construct (14 261 Da), we calculated Matthews coefficients (Matthews, 1968) of 4.23 and 2.12 Å³ Da^{−1} for one and two monomers per asymmetric unit, respectively. Although the latter value (corresponding to a solvent content of 42%) is more probable based on entries in the PDB (Kantardjieff & Rupp, 2003), it is possible that only one monomer is present in the asymmetric unit, resulting in a solvent content of approximately 71%. No prominent

peaks are present in a self-Patterson map, but a self-rotation map betrayed the presence of a peak with 65% of the strength of the origin peak at $\theta = 90.0$, $\varphi = -86.8$, $\chi = 180.0^\circ$, indicating the likely presence of a twofold rotation axis in the asymmetric unit. Both SAD (single-wavelength anomalous diffraction) and MAD (multi-wavelength anomalous diffraction) using a selenomethionyl derivative of the protein are being pursued as phasing strategies. Additionally, initial molecular-replacement trials with the low-homology models deposited as PDB entries 3lhn (Joint Center for Structural Genomics, unpublished work) and 2z4h (Hirano *et al.*, 2007) have not yielded a clear solution.

We thank Dr Diana Tomchick for helpful discussions. This research was supported by an NIH grant (AI056305) to MVN. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne LLC for the US Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Brautigam, C. A., Deka, R. K., Ouyang, Z., Machius, M., Knutsen, G., Tomchick, D. R. & Norgard, M. V. (2012). *J. Bacteriol.* **194**, 6771–6781.
 Brautigam, C. A., Deka, R. K., Schuck, P., Tomchick, D. R. & Norgard, M. V. (2012). *J. Mol. Biol.* **420**, 70–86.
 Chao, J. R., Khurana, R. N., Fawzi, A. A., Reddy, H. S. & Rao, N. A. (2006). *Ophthalmology*, **113**, 2074–2079.
 Cox, D. L., Luthra, A., Dunham-Ems, S., Desrosiers, D. C., Salazar, J. C., Caimano, M. J. & Radolf, J. D. (2010). *Infect. Immun.* **78**, 5178–5194.
 Deka, R. K., Brautigam, C. A., Goldberg, M., Schuck, P., Tomchick, D. R. & Norgard, M. V. (2012). *J. Mol. Biol.* **416**, 678–696.
 Deka, R. K., Brautigam, C. A., Tomson, F. L., Lumpkins, S. B., Tomchick, D. R., Machius, M. & Norgard, M. V. (2007). *J. Biol. Chem.* **282**, 5944–5958.
 Deka, R. K., Brautigam, C. A., Yang, X. F., Blevins, J. S., Machius, M., Tomchick, D. R. & Norgard, M. V. (2006). *J. Biol. Chem.* **281**, 8072–8081.
 Fraser, C. M. *et al.* (1998). *Science*, **281**, 375–388.
 French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
 Gerbase, A. C., Rowley, J. T., Heymann, D. H., Berkley, S. F. & Piot, P. (1998). *Sex. Transm. Infect.* **74**, S12–S16.
 Hirano, Y., Hossain, M. M., Takeda, K., Tokuda, H. & Miki, K. (2007). *Structure*, **15**, 963–976.
 Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
 Machius, M., Brautigam, C. A., Tomchick, D. R., Ward, P., Otwinowski, Z., Blevins, J. S., Deka, R. K. & Norgard, M. V. (2007). *J. Mol. Biol.* **373**, 681–694.
 Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
 Norris, S. J. (1993). *Microbiol. Rev.* **57**, 750–779.
 Otto, K. & Silhavy, T. J. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 2287–2292.
 Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
 Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* **D59**, 1124–1130.
 Radolf, J. D. (1995). *Mol. Microbiol.* **16**, 1067–1073.
 Snyder, W. B., Davis, L. J., Danese, P. N., Cosma, C. L. & Silhavy, T. J. (1995). *J. Bacteriol.* **177**, 4216–4223.
 Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.
 Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
 Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.