 doi:10.20944/preprints202112.0492.v1

Secondary structure and X-ray crystallographic analysis of the Glideosome-Associated Connector (GAC) from *Toxoplasma gondii*

Amit Kumar[1], Xu Zhang[1], Oscar Vadas[2], Fisentzos A Stylianou[1], Nicolas Dos Santos Pacheco[2], Sarah L. Rouse[1], Marc L. Morgan[1], Dominique Soldati-Favre[2] & Steve Matthews[1]

[1]Department of Life Sciences, Imperial College London, South Kensington, London SW7 2AZ, UK.

[2]Department of Microbiology & Molecular Medicine, University of Geneva, 1 Rue Michel-Servet, 1211 Geneva, Switzerland.

 doi:10.20944/preprints202112.0492.v1

Abstract

A model for parasitic motility has been proposed in which parasite filamentous actin (F-actin) is attached to surface adhesins by a large component of the glideosome, known as the glideosome-associated connector protein (GAC). This large 286 kDa protein interacts at the cytoplasmic face of the plasma membrane with the phosphatidic acid-enriched inner leaflet and cytosolic tails of surface adhesins to connect them to the parasite actomyosin system. GAC is observed initially to the conoid at the apical pole and re-localised with the glideosome to the basal pole in gliding parasite. GAC presumably functions in force transmission to surface adhesins in the plasma membrane and not in force generation. Proper connection between F-actin and the adhesins is as important for motility and invasion as motor operation itself. This notion highlights the need for new structural information on GAC interactions, which has eluded the field since its discovery. We have obtained crystals that diffracted to 2.6-2.9 Å for full-length GAC from *Toxoplasma gondii* in native and selenomethionine-labelled forms. These crystals belong to space group P212121, cell dimensions are roughly a=119 Å, b=123Å, c=221Å, α=90, β=90, γ=90 with 1 molecule per asymmetric unit, suggesting a more compact conformation than previously proposed.

 doi:10.20944/preprints202112.0492.v1

1. Introduction

There are more than 5000 species of apicomplexan parasites and many are etiological agents of major diseases that are a threat to global human and animal health, particularly in low-resource settings. Most significant are malaria (Plasmodium), cryptosporidiosis (Cryptosporidium) and toxoplasmosis (Toxoplasma). The lifestyle of these obligate intracellular parasites involves crucial steps that depend on gliding motility such host cell invasion, egress from infected cell and crossing of biological barriers.

These processes are dependent upon the orchestrated release of proteins from apical secretory organelles; micronemes and rhoptries. Following initial apical organelle secretion, a moving junction is formed that participates in the active penetration of host-cells. In addition to invasion, parasites also use gliding motility to actively exit infected host cells during egress or migrate across biological surfaces, and in all cases, motility appears to be powered by the glideosome (Frenal *et al.*, 2017, Boucher & Bosch, 2015, Cowman *et al.*, 2017).

Current understanding broadly agrees upon a molecular architecture for the glideosome and explains how an actomyosin-motor drives motility in apicomplexan parasites. In this model, the TRAP/MIC family of adhesins target ligands on the surface of the host cell to mediate apical attachment (Carruthers & Tomley, 2008). At the cytoplasmic side of the plasma membrane these adhesins are connected to the parasite actin filament network, while myosin motors drag adhesins through the plane of the plasma membrane towards the parasite posterior, and consequently pull the host cell membrane around the parasite. The myosin motor is situated in the space between the inner membrane complex (IMC) and parasite plasma membrane and together with other associated proteins it spans the two membranous structures. Myosin A (MyoA), a small divergent

class XIV myosin together with and glideosome-associated proteins (GAPs) act as the motor powering gliding motility (Powell *et al.*, 2018). MyoA is situated in the space between the inner membrane complex (IMC) and parasite plasma membrane and together with the GAPs it spans the two membranous structures.

The molecular component that bridges the adhesin to F-actin is a large novel protein termed the glideosome associated connector protein (GAC), which translocates with the moving junction from the parasite apex to the basal pole during gliding motility (Jacot *et al.*, 2016). GAC is highly conserved across the entire Apicomplexa phylum and forms complexes with three binding partners (**Figure 1A**). The C-terminal region of GAC interacts with phosphatidic acid (PA) enriched membranes and its deletion results in a defective lytic-cycle phenotype. It has been suggested that GAC localisation is dependent upon PA generated by the lipid signalling cascade that regulates the apicomplexan Pleckstrin-homology domain protein (APH) to control microneme secretion (Darvill *et al.*, 2018, Bullen *et al.*, 2016). PA signalling may ensure GAC is appropriately recruited only when a productive interaction can occur with adhesins during gliding motility. GAC also binds to microneme adhesin C-terminal tails at the plasma membrane's cytoplasmic face and therefore serves as the link to the surface adhesins. Thirdly, a direct connection to the glideosome is made via an interaction between GAC and the parasite actin filaments (F-actin). It has also been shown that proper connection of actin to the adhesins (via GAC) is more important for efficient motility and invasion than motor operation itself (Whitelaw *et al.*, 2017).

We set out to solve the structure of GAC from *T. gondii*. Initial efforts to generate high-resolution diffracting crystals of TgGAC proved challenging. Eventually after by optimisation of purification conditions together with secondary structure analyses and

 doi:10.20944/preprints202112.0492.v1

extensive crystallization screening, we generated reproducible crystals that diffract to 2.67 Å, which represents a major step towards providing the first structural insight into the GAC architecture and function within the glideosome.

2. Materials and methods

2.1 Protein expression and purification

Full-length TgGAC genes with TEV cleavable N-terminal 6xHis-tag had been cloned into the pET28a vector as previous described (Jacot *et al.*, 2016). Cells were harvested and resuspended in 50 mM Tris (pH 8), 300 mM NaCl, 10 mM Imidazole and 5 mM TCEP, followed by lysis by sonication and centrifugation at 18,000 rpm for 60 minutes (Ti45 rotor, Beckmann). TgGAC was then purified by nickel chromatography followed by gel filtration using a Superdex-300 column (GE Healthcare).

2.2 TgGAC crystallization

Conditions for crystallization were initially screened by the sitting drop method of vapour diffusion at 20 °C and 4 °C using sparse matrix crystallization kits (Hampton research and Molecular Dimensions). MRC 96-well optimization plates (Molecular Dimensions) were utilised. Each drop was set with 100 nl protein solution and 100 nl reservoir solution using a Mosquito nanolitre high-throughput robot (TTP Labtech). Protein crystals were obtained in a reproducible manner from in the condition of 100 mM magnesium acetate, 100 mM sodium acetate, 6% PEG8000, and pH 5.0. These were manually optimised by screening over sodium acetate pH ranges 4.0 to 5.0 in one dimension and a PEG8000 concentration gradient of 4 – 10% in the second dimension. Crystallization was set-up at concentration of 5-60 mg ml^{-1}.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

2.3. CD Spectroscopy

CD spectrum of GAC was recorded at a concentration of 0.8 mg ml^{-1} in a variety of solution condition and temperatures. Spectra were recorded in the wavelength range of 200 – 260 nm, with scan length per point of 2 seconds. Four repeats were collected and averaged. Data was collected and processed by Chirascan CD Spectrometer and APLData Converter. 2.4, respectively.

2.4. X-ray data collection and processing

Crystals were mounted in a MicroLoop (MiTeGen), cryoprotected with 30% ethylene glycol for 5 seconds and immediately flash-cooled in liquid nitrogen. Diffraction data from a single native crystal were collected on beamline i04 of the Diamond Light Source (DLS), UK. Data were processed with CCP4, dials (Beilsten-Edmands *et al.*, 2020, Winn *et al.*, 2011, Winter *et al.*, 2018, Winter, 2010) and scaled using dials.scale (Evans, 2006) within the Xia2 package (Winter *et al.*, 2013). Multiple-wavelength anomalous diffraction (MAD) data from a single SeMet labelled crystal were collected on beamline i04 of the Diamond Light Source at the following wavelengths: peak=0.9795Å, inflection=0.9796Å and remote=0.9722Å. Data were processed initially by AutoProc (Vonrhein *et al.*, 2011). Substructure definition and initial model building were performed using AutoSHARP (Vonrhein *et al.*, 2007). This was followed by manual building in Coot (Emsley *et al.*, 2010) and further refinement using Phenix Refine (Adams *et al.*, 2010).

Data collection statistics are shown in **Table 1**. The content of the unit cell was analysed using the Matthews coefficient (Matthews, 1968). Molecular replacement (MR)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

attempts were carried out using computationally derived structures using the following servers: RaptorX, Alphafold and iTasser (Senior *et al.*, 2020, Källberg *et al.*, 2012, Yang *et al.*, 2015)

3. Results and discussion

While the identity and function of crucial genetic components of the parasite life cycle and infectivity are known, a detailed mechanistic understanding of parasite motility and invasion remains limited. Despite GAC's essential role in efficient motility and invasion, no high-resolution experimental structural information is available. A small-angle X-ray scattering (SAXS) study presented TgGAC as a ~27nm club-shaped molecule that stretches across the space between the parasitic plasma membrane and F-actin (Jacot *et al.*, 2016) (**Figure 1A**). However, this model is inconsistent with our current understanding of the glideosome, as GAC would be unable to fit lengthways across this space together with the other essential components. To fully understand how GAC carries out its role, new experimental structural insight is required. We therefore isolated and purified the full-length TgGAC (**Figure 1B**). TgGAC contains 75 cysteines residues that are predicted not to participate in disulphide bonds. After assessment of structure and stability of GAC with CD spectroscopy (**Figures 1C & 1D**), we found that maintaining strict reducing conditions throughout purification was a crucial step in maintaining GACs full secondary structure (62% α-helix, 17% β-sheet, 11% turn). Under these experimental conditions, crystals reliably grew to 100-200 μm^3 in size over the course of 5-7 days (**Figure 2A**). Furthermore, to ensure the highest resolution it was necessary to acquire X-ray diffraction images immediately after harvesting and freezing. Frozen crystals stored for any longer than a few days experienced significant deterioration in high resolution diffraction.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

Native diffraction data were collected to 2.92 Å (**Figure 2B**) and indexed in space group P212121. Analysis of the crystal content indicated cell dimensions are a=119 Å, b=123Å, c=221Å, α=90, β=90, γ=90 with 1 molecule per asymmetric unit with a solvent content of 56%. This suggests that GAC adopts a more compact and distinct conformation from that suggested by the earlier SAXS analysis and will likely explain the architecture of within the confined. Data-collection and processing statistics are listed in **Table 1**.

Molecular replacement attempts were made using idealized structures based on various prediction algorithms as search models: RaptorX, Alphafold and iTasser. No solutions were found. We subsequently prepared selenomethionine-substituted and heavy-atom derivatives to provide accurate phases using anomalous dispersion techniques. A single crystal of selenomethionine-labelled TgGAC was successfully obtained in the same conditions as native protein. The crystals of native and SeMet-derivatized SD protein diffracted to 2.92 Å and 2.67 Å resolution, respectively. Matthews coefficient (VM = 2.70 $Å^3$ Da^{-1}) and solvent-content (VS = 56.8%) calculations indicated that one molecule was present in the asymmetric unit. The multiple-wavelength anomalous dispersion method was used to determine the initial phases of the SeMet-substituted SD protein. The autoSHARP processing pipeline obtained 73 Se sites. Structure solution is currently under way and the phasing statistics will be reported elsewhere.

Acknowledgements

This work was supported by a BBSRC and Leverhulme Trust awards to SJM (RPG_2018_107). The crystallisation facility at Imperial College London is supported

by the Biotechnology and Biological Sciences Research Council (BB/D524840/1) and Wellcome Trust (202926/Z/16/Z). We are grateful to staff at Diamond Light Source beamline I04 for their help with data collection.

References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-221.

Beilsten-Edmands, J., Winter, G., Gildea, R., Parkhurst, J., Waterman, D. & Evans, G. (2020). *Acta Crystallogr D Struct Biol* **76**, 385-399.

Boucher, L. E. & Bosch, J. (2015). *J Struct Biol* **190**, 93-114.

Bullen, H. E., Jia, Y., Yamaryo-Botte, Y., Bisio, H., Zhang, O., Jemelin, N. K., Marq, J. B., Carruthers, V., Botte, C. Y. & Soldati-Favre, D. (2016). *Cell Host Microbe* **19**, 349-360.

Carruthers, V. B. & Tomley, F. M. (2008). *Subcell Biochem* **47**, 33-45.

Cowman, A. F., Tonkin, C. J., Tham, W. H. & Duraisingh, M. T. (2017). *Cell host & microbe* **22**, 232-245.

Darvill, N., Dubois, D. J., Rouse, S. L., Hammoudi, P. M., Blake, T., Benjamin, S., Liu, B., Soldati-Favre, D. & Matthews, S. (2018). *Structure (London, England : 1993)* **26**, 1059-1071.e1056.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501.

Evans, P. (2006). *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72-82.

Frenal, K., Dubremetz, J. F., Lebrun, M. & Soldati-Favre, D. (2017). *Nat Rev Microbiol* **15**, 645-660.

Jacot, D., Tosetti, N., Pires, I., Stock, J., Graindorge, A., Hung, Y. F., Han, H., Tewari, R., Kursula, I. & Soldati-Favre, D. (2016). *Cell Host Microbe* **20**, 731-743.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H. & Xu, J. (2012). *Nat Protoc* **7**, 1511-1522.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491-497.

Powell, C. J., Ramaswamy, R., Kelsen, A., Hamelin, D. J., Warshaw, D. M., Bosch, J., Burke, J. E., Ward, G. E. & Boulanger, M. J. (2018). *Proceedings of the National Academy of Sciences of the United States of America* **115**, E10548-e10555.

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K. & Hassabis, D. (2020). *Nature* **577**, 706-710.

Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol Biol* **364**, 215-230.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011). *Acta Crystallogr D Biol Crystallogr* **67**, 293-302.

Whitelaw, J. A., Latorre-Barragan, F., Gras, S., Pall, G. S., Leung, J. M., Heaslip, A., Egarter, S., Andenmatten, N., Nelson, S. R., Warshaw, D. M., Ward, G. E. & Meissner, M. (2017). *BMC Biology* **15**, 1.

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., 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 Crystallogr D Biol Crystallogr* **67**, 235-242.

Winter, G. (2010). *J. Appl. Crystallogr.* **43**, 186-190.

Winter, G., Lobley, C. M. & Prince, S. M. (2013). *Acta Crystallogr D Biol Crystallogr* **69**, 1260-1273.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Crystallogr D Struct Biol* **74**, 85-97.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. & Zhang, Y. (2015). *Nat Methods* **12**, 7-8.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

Table 1

Values in parentheses are for the highest resolution shell.

Crystal	Native	SeMet
Space group	P212121	P212121
Cell dimensions (Å)	*a*=120.63, *b*=123.96, *c*=221.86	*a*=119.08, *b*=123.60, *c*=221.51
Angles (°)	α=90.00, β=90.00, γ=90.00	α=90.00, β=90.00, γ=90.00
Resolution (Å)	82.66-2.92 (2.97-2.92)*	110.75-2.67 (2.67-2.72)*
Wavelength (Å)	0.97950	0.97950
Total reflections	2899820 (140928)	904519 (44214)
Unique observations	72969 (3564)	92922 (4581)
Completeness (%)	100 (99-100)	100 (99.7 - 100)
Multiplicity	39.7 (39.5)	9.7 (9.7)
R_{pim}†	0.045 (1.486)	0.039 (0.861)
$\langle I \rangle / \sigma I$	13.1 (88 -0.5)	14 (47.2 - 0.8)
Molecules per asymmetric unit‡	1	1
Solvent content (%)	56.8	56.8

*Values in parentheses correspond to the highest resolution shell

†$R_{merge} = \Sigma_{hkl} \Sigma_i \mid I_i(hkl) - \langle I(hkl) \rangle \mid / \Sigma_{hkl} \Sigma_i I_i(hkl)$ where $\langle I(hkl) \rangle$ is the mean intensity of the observations $I_i(hkl)$ of reflection *hkl*. ‡Most probable value.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

Figures and legends

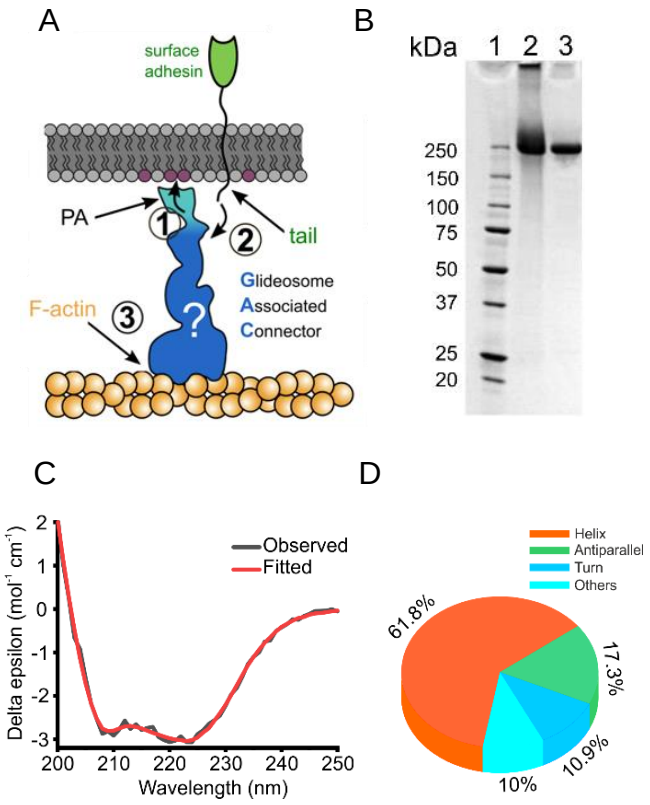

Figure 1. TgGAC interactions, purification, and characterisation. (A) Schematic representation of GAC interactions with 1) the plasma membrane, 2) surface adhesins and 3) F-actin (B) SDS–PAGE of purified full-length TgGAC used for the crystallization trials. Lane 1: molecular-weight markers (kDa), lane 2: Ni-NTA eluted fraction and lane 3: eluted fraction from an S-300 SEC column. (C) CD spectrum of purified TgGAC at 298 K. Black spectrum showed the observed spectrum in 25 mM Tris.HCl, pH 8.0. Red spectrum indicated the fitted line for secondary structure analysis. (D) Secondary structure analysis based upon the CD spectrum as indicated in (C). Orange indicated the helical contents, green as antiparallel, turn as blue and cyan as other contents in the GAC.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2021 doi:10.20944/preprints202112.0492.v1

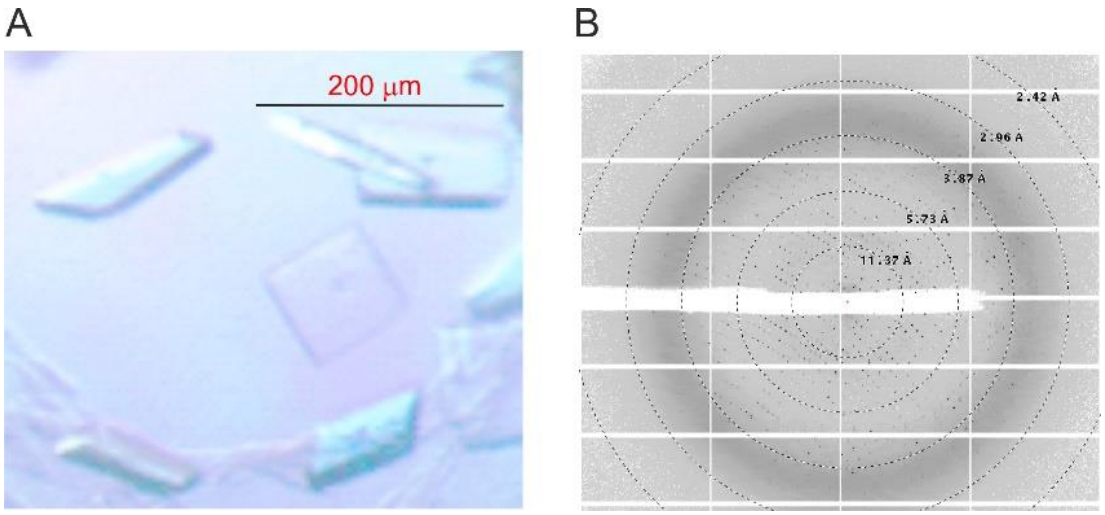

Figure 2. TgGAC full length crystallography. (A) Representative native crystals of TgGAC. Scale bar represents 200 μm. (B) Diffraction image from a TgGAC crystal. Resolution rings are annotated.