Elsevier Editorial System(tm) for Protist
Manuscript Draft

Manuscript Number: PROTIS-D-16-00074R3

Title: Variation in basal body localisation and targeting of trypanosome RP2 and FOR20 proteins

Article Type: Original Paper

Keywords: basal body; ciliogenesis; FOP; protein targeting; Trypanosoma brucei; YL1/2

Corresponding Author: Professor Michael L. Ginger, PhD

Corresponding Author's Institution: University of Huddersfield

First Author: Jane Harmer

Order of Authors: Jane Harmer; Xin Qi; Gabriella Toniolo; Aysha Patel; Hannah Shaw; Fiona E Benson; Michael L. Ginger, PhD; Paul G McKean

Abstract: TOF-LisH-PLL motifs defines FOP family proteins; some members are involved in flagellum assembly. The critical role of FOP family protein FOR20 is poorly understood. Here, we report relative localisations of the four FOP family proteins in parasitic Trypanosoma brucei: TbRP2, TbOFD1 and TbFOP/FOP1-like are mature basal body proteins whereas TbFOR20 is present on pro- and mature basal bodies - on the latter it localises distal to TbRP2. We discuss how the data, together with published work for another protist Giardia intestinalis, informs on likely FOR20 function. Moreover, our localisation study provides convincing evidence that the antigen recognised by monoclonal antibody YL1/2 at trypanosome mature basal bodies is FOP family protein TbRP2, not tyrosinated α-tubulin as widely stated in the literature. Curiously, FOR20 proteins from T. brucei and closely related African trypanosomes possess short, negatively-charged N-terminal extensions absent from FOR20 in other trypanosomatids and other eukaryotes. The extension is necessary for protein targeting, but insufficient to re-direct TbRP2 to probasal bodies. Yet, FOR20 from the American trypanosome T. cruzi, which lacks any extension, localises to pro- and mature basal bodies when expressed in T. brucei. This identifies unexpected variation in FOR20 architecture that is presently unique to one clade of trypanosomatids.

Variation in basal body localisation and targeting of trypanosome RP2 and FOR20 proteins

Jane Harmer[a], Xin Qi[a], Gabriella Toniolo[a], Aysha Patel[a], Hannah Shaw[a], Fiona E. Benson[a], Michael L. Ginger[b,1], and Paul G. McKean[a,1]

[a]Faculty of Health and Medicine, Division of Biomedical and Life Sciences, Lancaster University, Lancaster LA1 4YQ, UK.

[b]Department of Biological Sciences, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK.

[1]Co-senior and co-corresponding authors: M.Ginger@hud.ac.uk; p.mckean@lancaster.ac.uk

TOF-LisH-PLL motifs defines FOP family proteins; some members are involved in flagellum assembly. The critical role of FOP family protein FOR20 is poorly understood. Here, we report relative localisations of the four FOP family proteins in parasitic *Trypanosoma brucei*: *Tb*RP2, *Tb*OFD1 and *Tb*FOP/FOP1-like are mature basal body proteins whereas *Tb*FOR20 is present on pro- and mature basal bodies – on the latter it localises distal to *Tb*RP2. We discuss how the data, together with published work for another protist *Giardia intestinalis*, informs on likely FOR20 function. Moreover, our localisation study provides convincing evidence that the antigen recognised by monoclonal antibody YL1/2 at trypanosome mature basal bodies is FOP family protein *Tb*RP2, not tyrosinated α-tubulin as widely stated in the literature. Curiously, FOR20 proteins from *T. brucei* and closely related African trypanosomes possess short, negatively-charged N-terminal extensions absent from FOR20 in other trypanosomatids and other eukaryotes. The extension is necessary for protein targeting, but insufficient to re-direct *Tb*RP2 to probasal bodies. Yet, FOR20 from the American trypanosome *T. cruzi*, which lacks any extension, localises to pro- and mature basal bodies when expressed in *T. brucei*. This identifies unexpected variation in FOR20 architecture that is presently unique to one clade of trypanosomatids.

Keywords: basal body; ciliogenesis; FOP; protein targeting; *Trypanosoma brucei*; YL1/2

Running title: FOP family protein targeting in trypanosomatids

Introduction

In eukaryotic cells, cilia (or flagella) are often central to cell swimming, cell feeding, reproduction, and sensory perception. Length (cilia tend to be thought of as shorter), number (large numbers of cilia tend to be arrayed across cell surfaces whereas examples of flagellate protists with more than two flagella are fewer than taxa possessing one flagellum or two flagella), and principal mode of motion (an oar-like ciliary waveform versus more whip-like flagellar beating) are the obvious determinants commonly used for distinguishing cilia from flagella. Yet, these terms refer essentially to variants of the same organelle that have as their defining structure a microtubule axoneme (Moran et al. 2014).

In the classic '9+2' configuration, axonemes are composed of nine outer-doublet microtubules surrounding two singlet microtubules. Irrespective of the number of outer-doublet microtubules present (structures with as few as three outer-doublets have been described (Prensier et al. 2008)), axoneme elongation occurs from a barrel-shaped microtubule organising centre (MTOC) or centriole, which in the context of flagellum assembly is better described as a basal body. *De novo* basal body biogenesis is known, but in many organisms a probasal body, or pro-centriole, comprised of triplet microtubules, rather than doublets, is physically associated with a mature basal body (Fritz-Laylin et al. 2016). Thus, in trypanosomatids, the flagellate parasitic protists featured in this work, probasal body-to-basal body maturation occurs when doublet microtubules extend from A- and B-tubules of the triplets of the probasal body, thereby forming the transition zone of the mature basal body (Vaughan and Gull 2015). The transition zone is capped at the distal end by a basal plate from which the axoneme proper subsequently extends by a process that, as in many other flagellate cells, is dependent upon an intraflagellar transport (IFT) system (Absalon et al. 2008, Davidge et al. 2006). Coincident with trypanosome probasal body maturation, biogenesis of two new probasal bodies and their association to either the newly matured basal body or the basal body matured in a previous cell cycle, also occurs (Vaughan and Gull 2015).

Estimates derived from the proteomics of organelles isolated from diverse taxa suggest 200-300 different proteins are likely to be *bona fide* components of mature basal bodies (and their associated appendages) *e.g.* (Keller et al. 2005, Kilburn et al. 2007). Such estimates provide an interesting contrast with bioinformatics-based comparisons that indicate structural conservation and complexity of centriole/basal body symmetry across the breadth of eukaryotic evolution may be dependent upon only a handful of conserved proteins (Carvalho-Santos et al. 2011, Hodges et al. 2010). One basal body/centriole protein conserved in evolutionarily diverse flagellate eukaryotes is FOR20.

The only recognisable architectural features of the small FOP-related protein of 20 kDa (or FOR20) are also shared with other FOP family proteins (*e.g.* FOP, TONNEAU1 and OFD1), namely N-terminally localised TOF, LisH and 'PLL' motifs (Sedjai et al. 2010). The protein was initially described as present at the distal end of the basal body in the ciliate *Tetrahymena thermophila* (where it is known as Bbc20 (Kilburn et al. 2007)) and subsequently as a component of the granular pericentriolar satellites that surround the centrosome (a centriole-bearing MTOC) in animal cells (Fritz-Laylin and Cande 2010, Sedjai et al. 2010). FOR20 is required for assembly of the non-motile primary cilium that extends as a sensory antenna from the surface of many animal cell types (Sedjai et al. 2010). In such cells, primary cilium formation occurs following centrosome relocation from a normally central intracellular position to the cortical cytoskeleton and the maturation of the mother centriole to a basal body (Dawe et al. 2007) although how FOR20 contributes to primary cilium assembly is not certain. Experimental analysis by gene-specific RNA interference (RNAi) in the ciliate *Paramecium* indicates its FOR20 is a stable component of the ciliate basal body, rather than subject to turn over, and is required for basal body docking at the plasma membrane and/or transition zone maturation (Aubusson-Fleury et al. 2012). Similar to Bbc20 in *Tetrahymena* and consistent with a proposed role in membrane docking, *Paramecium* FOR20 also localises to the distal end of both older (ciliated) and their associated non-ciliated, younger basal bodies (Aubusson-Fleury et al. 2012).

Generally speaking, FOP family proteins are found in flagellate eukaryotes, but not in organisms that lack a capacity to build a flagellum – acentriolar plants are the exception to this rule. In each of the taxa examined thus far, the FOP protein family is small in number (Azimzadeh et al. 2008, Hodges et al. 2010) and among flagellates the other family members are also centriolar and required for flagellum assembly (André et al. 2014, Aubusson-Fleury et al. 2012, Sedjai et al. 2010, Singla et al. 2010). In acentriolar plant cells another FOP-related protein, TONNEAU1, is found; it is required for organisation of cortical

microtubule arrays and interacts with the classic MTOC protein centrin (Azimzadeh et al. 2008, Spinner et al. 2010). In kinetoplastid protists (which include the parasitic trypanosomatids), four proteins comprise the FOP family: three conserved family members, FOP/FOP1-like, FOR20 and OFD1, plus a lineage-specific protein known as *Tb*RP2 in the African trypanosome *Trypanosoma brucei* (André et al. 2014). In this lineage-specific protein the TOF-LisH-PLL motif sequence lies, apparently uniquely, upstream of a tubulin cofactor C domain – although there is indication of a conserved requirement for a tubulin cofactor C domain-containing protein *per se* in flagellum assembly. Thus, *Tb*RP2 is found at mature basal bodies and in *T. brucei* is required for assembly of a full-length flagellum and an intact axoneme (André et al. 2014, Stephan et al. 2007). Intriguingly, we reported recombinant *Tb*RP2 is recognised by monoclonal antibody YL1/2 (André et al. 2014). YL1/2 is classically used to detect tyrosinated α-tubulin in eukaryotic cells, including by many in the trypanosomatid community. There is no doubt that YL1/2 recognises *T. brucei* tyrosinated α-tubulin, but likely additional recognition of *Tb*RP2 calls into question whether there is a mature basal body pool of tyrosinated α-tubulin specifically recognised by YL1/2.

Peculiarities of the kinetoplastid FOP family are not restricted to the presence of a lineage-specific family member: the only candidate orthologue of FOP, which in animal cells is required for ciliogenesis (Lee and Stearns 2013), is a protein we term FOP/FOP1-like and is required in *T. brucei* for assembly of an essential (lineage-specific) extra-axonemal structure, the paraflagellar rod, but has no discernible effect on axoneme assembly (André et al. submitted). Among African trypanosome species *FOR20* is predicted to encode a protein possessing a predicted short N-terminal extension not seen in either other kinetoplastids or other eukaryotes. Here, we demonstrate the N-terminal extension predicted for *T. brucei* FOR20 is real, rather than an artefact of gene annotation, and essential but not sufficient for protein localisation to both pro- and mature basal bodies throughout the trypanosome cell cycle. Interestingly, FOR20 from the distantly related American trypanosome *T. cruzi*, which lacks an N-terminal extension, is also targeted to pro- and mature basal bodies when expressed in *T. brucei*. Collectively, data presented here provide new insights into the functional diversification of the conserved, yet small FOP protein family, the spatial organisation of trypanosome basal body biogenesis, and provide further indication that *Tb*RP2 is the basal body antigen specifically recognised by YL1/2.

Results

*Tb*FOR20 is present at pro- and mature basal bodies

The *T. brucei* FOR20 orthologue is encoded by Tb927.11.3090. Expressed from an endogenous chromosomal locus as an N-terminal fusion with YFP, *Tb*FOR20 is targeted to both pro- and mature basal bodies, with YFP fluorescence lying distal to the indirect immunofluorescence signal observed using monoclonal antibody BBA4 (Fig. 1A). BBA4 recognises an unknown trypanosome antigen from the proximal end of pro- and mature basal bodies (Woodward et al. 1995); thus YFP::*Tb*FOR20 localises to, or towards, the distal end of probasal bodies and is retained on mature basal bodies. In trypanosomatids, pro- and mature basal bodies are physically attached at their proximal ends to the mitochondrial genome (or kinetoplast) by a complex filament network, the tri-partite attachment complex (TAC), that traverses outer and inner mitochondrial membranes (Ogbadoyi et al. 2003); thus, the fluorescence from 6-diamidino-2-phenylindole (DAPI)-stained kinetoplasts lies proximal the BBA4-immunolabelled basal bodies.

The relative position of YFP::*Tb*FOR20 on mature basal bodies (Fig. 1B) was determined by dual fluorescence experiments with monoclonal antibody YL1/2, which recognises the C-terminal ‘D-D-F’ and ‘E-E-Y’ epitopes of mature basal body localised

*Tb*RP2 and/or tyrosinated α-tubulin, respectively (André et al. 2014, Wehland et al. 1984). Here we performed immunofluorescence analysis of detergent extracted cytoskeletons (Fig. 1B) and isolated flagella (Fig. 2). On cytoskeletal images there was ambiguity as to whether YL1/2 lay distal, proximal or coincident with YFP::*Tb*FOR20. Detergent and NaCl extraction of *T. brucei* cells yields axonemes plus associated paraflagellar rod, transition zone microtubules, and basal bodies. Analysis of flagella isolated by detergent and then NaCl extraction revealed spatial separation of YL1/2 immunofluorescence and YFP::*Tb*FOR20 signals, with the former lying proximal to YFP::*Tb*FOR20 (Fig. 2A). In our earlier work (André et al. 2014) we proposed that at mature trypanosome basal bodies *Tb*RP2 is an antigen recognised by YL1/2. Given the spatial resolution of YL1/2 immunofluorescence and YFP::*Tb*FOR20 signals we were able to further question the identity of the trypanosome antigen(s) recognised by YL1/2. Indirect immunofluorescence of isolated flagella using a specific anti-*Tb*RP2 antibody (André et al. 2014) revealed that on mature basal bodies *Tb*RP2 is also proximal to YFP::*Tb*FOR20 (Fig. 2D), consistent with the idea that YL1/2 recognises *Tb*RP2 rather than tyrosinated α-tubulin. The final piece of experimental evidence that YL1/2 indeed specifically recognises *Tb*RP2 as its basal body antigen is the observation that YL1/2 detects by immunoblot not merely recombinant *Tb*RP2 (as we reported previously, André et al. 2014,) but also native *Tb*RP2 from trypanosome extracts (Fig. 3). We return to this point in the discussion.

Analysis of cells from early stages of the cell division cycle, when maturation of the probasal body assembled in the previous cell cycle occurs and elongation of the new flagellum begins, indicated duplication of YFP::*Tb*FOR20 fluorescence signals was coincident with new probasal body biogenesis and kinetoplast replication (Fig. 1C). Localisation of YFP::*Tb*FOR20 to both pro- and mature basal bodies was then retained throughout the remainder of the cell cycle (data not shown). In terms of length and motif architecture full length *Tb*FOR20 closely resembles a C-terminal truncated variant of *Tb*RP2, *Tb*RP2$^{\Delta 134\text{-}463}$::myc, which localises only to mature basal bodies in *T. brucei* (André et al. 2014). To confirm pro- and mature basal body localisations for YFP::*Tb*FOR20 was not an artefact of the N-terminal YFP tag, we also expressed in *T. brucei Tb*FOR20 with a C-terminal myc-epitope (*Tb*FOR20::myc). Again, tagged protein was targeted to both pro- and mature basal bodies (Fig. 1D). Moreover, on isolated flagella *Tb*FOR20::myc, as with YFP::*Tb*FOR20, lies distal to *Tb*RP2 (Fig. 2E). Collectively, our localisation data indicate neither the nature nor the position of the tag influence protein localisation.

An N-terminal extension particular to FOR20 from African trypanosome species

We looked for differences between *Tb*FOR20, *Tb*RP2, and the other two trypanosome proteins with coupled TOF-LisH motifs (*Tb*OFD1, *Tb*FOP/FOP1-like, which are both mature basal body proteins; Fig. 2B-C) that could offer insight into why only *Tb*FOR20 is additionally targeted to probasal bodies, we noted (i) the distance from start methionine to the TOF motif is approximately twenty amino acids longer in the protein encoded by the Tb*FOR20* gene model and (ii) this short N-terminal extension, enriched for negatively charged amino acids, appeared to be unique to mammal-infective African trypanosome species. Thus, the start methionine for FOR20 orthologues from a taxonomically diverse range of flagellate eukaryotes, a free-living trypanosomatid relative, *Bodo saltans*, and for other trypanosomatids, including the American trypanosome species *T. cruzi* and *T. rangeli*, plus their phylogenetically close relation *T. grayi*[1] (Kelly et al. 2014), and different *Leishmania* species all lie immediately upstream of the TOF motif (Fig. 4). We checked whether in the gene models predicted for FOR20 from *T. cruzi* and *L. major* (where, along with *T. brucei*, genome assembly and annotation is possibly the most exhaustive) the start codon could

[1] Although *T. grayi* is a parasite of African crocodiles.

have been called incorrectly, and that the open reading frame therefore extended upstream to an alternative start codon. We found no evidence to suggest that this was the case. Moreover, a near absence of introns in trypanosomatid genomes means examples of *cis*-splicing are extremely rare. Thus, we believe the start codons predicted for FOR20 orthologues from *T. cruzi* and *Leishmania*, and by inference other trypanosomatids, are correct.

In *Tb*FOR20, we also noted a methionine immediately upstream of the TOF motif that could conceivably provide the start codon for a shorter protein, although this downstream methionine is not present in the syntenic orthologues from two other African trypanosome species, *T. congolense* and *T. vivax*, for which nuclear genome sequences are available. Nonetheless, to confirm the authenticity of the predicted start codon for *Tb*FOR20 we compared the localisation of C-terminally myc-tagged *Tb*FOR20 variant translated from the downstream methionine (*Tb*FOR20$^{\Delta 1\text{-}21}$::myc) with the localisations of full length *Tb*FOR20::myc and YFP::*Tb*FOR20 (in which YFP was fused in frame with the upstream methionine). Despite expression from a strong, doxycycline-inducible procyclin transcription promoter, *Tb*FOR20$^{\Delta 1\text{-}21}$::myc did not localise to pro- or mature basal bodies, but instead accumulated throughout the cell body (Fig. 5A). Thus, the short N-terminal extension unique to FOR20 from mammal-infective African trypanosome species is essential for basal body targeting, and moreover its role in basal body targeting is not masked by fusion of *Tb*FOR20 at its N-terminus to YFP.

To query the possible structural conformation adopted by the N-terminal extension, we submitted *Tb*FOR20 to analysis at the Phyre^2 web portal (Kelley et al. 2015). The α-helical conformations of the TOF and LisH motifs were accurately predicted and an equally confident prediction of disorder was noted for the short trypanosome FOR20 N-terminal extension (Fig. 6). Limited site directed mutagenesis within TOF and LisH motifs of *Tb*FOR20 revealed that both motifs were required for efficient basal body targeting. Three site-directed mutants were examined (E37A, F54A, Y88A; Fig. 7): for the E37A mutation overall protein expression levels were lower than for the other two site-directed mutants (Supp. Fig. 1B) and basal body targeting was completely abrogated (Fig. 7A). For F54A and Y88A mutations, where steady state accumulation of protein was much higher than for the E37A mutation, some protein was present at the basal/probasal body region (Fig. 7B-C) in detergent-extracted cytoskeletons although the localisation did not replicate the basal body/probasal body localisation seen for both YFP::*Tb*FOR20 and *Tb*FOR20::myc (Fig. 1) in that the typical two-dot basal/probasal body *Tb*FOR20 signals were not seen. Instead a single focus or sometimes multiple foci of *Tb*FOR20 fluorescence was observed; where a single TbFOR20 signal was seen this was often offset from the mature basal body signal of *Tb*RP2. The images shown in Fig. 7 were captured using identical acquisition parameters. Since F54A and Y88A mutant proteins readily accumulated in the cytoplasm of intact cells the representative localisation data (Fig.7B-C) is also suggestive of inefficient protein targeting to pro- and mature basal bodies.

*Tb*FOR20 N-terminal extension is insufficient to re-target *Tb*RP2$^{\Delta 134\text{-}463}$::myc

To investigate further the possible role of the extension in *Tb*FOR20 and the orthologues from African trypanosome species closely related to *T. brucei*, we determined whether the addition of the N-terminal 21 amino acids from *Tb*FOR20 to amino acids 2-133 of *Tb*RP2 would be sufficient to direct the protein chimera *Tb*FOR20$^{\Delta 22\text{-}151}$::RP2$^{\Delta 1/\Delta 134\text{-}463}$::myc to both pro- and mature basal bodies in *T. brucei* – since *Tb*RP2$^{\Delta 134\text{-}463}$::myc closely resembles *Tb*FOR20 in length and motif architecture and incorporates efficiently into mature basal bodies (André et al. 2014). Despite efficient expression of protein, localisation to mature basal bodies only, was observed throughout the cell division cycle (*i.e.* at the start of the cell

cycle when cells possess one flagellum, one probasal body, one nucleus and one kinetoplast (1K1N); following kinetoplast replication but pre-mitosis when cells possess two flagella, two probasal bodies and one nucleus (2K1N); and post-mitosis (2K2N)) (Fig. 5B).

Heterologously expressed *T. cruzi* FOR20 (with no extension) localises to pro- and mature basal bodies in *T. brucei*

In a final set of targeting experiments we asked where in a *T. brucei* cell would a trypanosomatid FOR20 in which an N-terminal extension is not normally present be found when expressed heterologously. For this analysis we amplified by PCR the FOR20 open-reading frame from *T. cruzi* (Sylvio X10 strain) and cloned the resulting amplicon into HindIII-XhoI-digested expression plasmid pDEX377$_{TbRP2::myc}$ such that *Tc*FOR20 would be expressed with a C-terminal myc-epitope tag in *T. brucei*. Dual immunofluorescence studies with either BBA4 (Fig. 5C) or affinity-purified anti-*Tb*RP2 antibodies (Fig. 5D) indicated that despite the absence of the N-terminal extension seen in *Tb*FOR20, *Tc*FOR20 is readily incorporated into pro- and mature basal bodies. The images shown in Fig. 5C-D are of detergent-extracted cytoskeletons, thereby confirming the stable incorporation of *Tc*FOR20 into *T. brucei* basal bodies.

Discussion

In a wider evolutionary context trypanosomatids and their free-living kinetoplastid relatives belong to the eukaryotic super group Excavata. Widely accepted as evolutionarily divergent unicellular eukaryotes (Akiyoshi and Gull 2013, Hampl et al. 2009, Katz and Grant 2015, Rogozin et al. 2009), cell shape and morphogenesis in these organisms is defined by a microtubule-based cytoskeleton. Extensive cytoskeletal remodelling and major alterations in cell morphology are observed during generally complex trypanosomatid life cycles (Sharma et al. 2008, Peacock et al. 2014) although changes in expression of just a single regulatory protein can control the developmental pattern of morphogenesis undertaken by *T. brucei* in its tsetse fly vector (Kolev et al. 2012). Gene duplication events and variations in the abundance of some individual cytoskeletal proteins also provide stage-specific regulation or life stage-relevant influences on trypanosome morphology (Hayes et al. 2014, Olego-Fernandez et al. 2009, Portman and Gull 2014, Sunter et al. 2015, Vedrenne et al. 2002). Differences between trypanosomatids with respect to domain or motif architectures in a few, lineage-specific cytoskeletal proteins have also been reported (*e.g.* Vaughan et al. 2008). Identification of a curious N-terminal extension in FOR20 orthologues from mammal-infective African trypanosomes not evident in any other eukaryote and enriched in acidic amino acids, adds a further molecular variation to our understanding of the sculpture and evolution of trypanosomatid cytoskeletons. Yet, trypanosomatids also provide tractable models for studying cytoskeletal processes conserved widely in eukaryotic evolution. Flagellum assembly, structure and function provide prime examples of where trypanosomes provide a useful model system for interrogating conserved gene function (Broadhead et al. 2006, Vincensini et al. 2011).

Probing FOR20 function and localisation

FOR20 localisation and function was previously studied in the ciliate *Paramecium tetraurelia* (Bengueddach et al. 2017, Aubusson-Fleury et al. 2012). Differences in basal body assembly and maturation processes between *T. brucei* and *P. tetraurelia* are

potentially informative in questioning FOR20 function. In both *Paramecium* and trypanosomatids, a non-ciliated basal body is physically associated with a mature basal body from which an axoneme is extended. In ciliates, the basal body transition zone is present and docked at the plasma membrane, irrespective of whether the basal body is ciliated or not (Aubusson-Fleury et al. 2012, Tassin et al. 2015). FOR20 is present at or immediately above the terminal plate (the most proximal boundary) of the ciliate transition zone of both ciliated and non-ciliated *Paramecium* basal bodies (Aubusson-Fleury et al. 2012). The distal region of the transition zone in non-ciliated basal bodies is subject to re-modelling coincident with the start of axoneme elongation (Aubusson-Fleury et al. 2012, Tassin et al. 2015). In contrast, *T. brucei* probasal bodies form orthogonal to their associated mature basal body. New probasal bodies form early in the cell cycle, coincident with maturation of the probasal body formed in the previous cell cycle, and re-orientate from orthogonal to parallel with the associated mature basal body prior to mitosis (~0.4 of a cell cycle later). Even when reoriented there is clearly a region of cytosol between the distal end of the probasal body and the flagellar pocket membrane (*e.g.* Höög et al. 2014, Lacomble et al. 2009). Following reorientation, the transition zone forms in the next cell cycle, as doublet microtubules extend from the triplet microtubule barrel (Vaughan and Gull 2015). Transition zone extension is accompanied by transitional fibre formation, membrane docking and continues with elongation of the axoneme proper (Lacomble et al. 2010, Woodward et al. 1995). In our view, early recruitment of *Tb*FOR20 to the distal end of orthogonal trypanosome probasal bodies and the distance between the re-orientated probasal bodies and the plasma membrane argue against a direct role for little-understood FOR20 in basal body-membrane docking, or at least that a role for FOR20 in docking is not conserved in all flagellate eukaryotes. We also note there is no evidence of motifs or domains associated with protein-lipid/membrane interaction seen in FOR20 proteins.

Additional to studying TbFOR20 localisation, we also subjected *Tb*FOR20 to gene-specific RNAi. Despite ~90% reduction in YFP::*Tb*FOR20 detectable by immunoblot (Supp. Fig. 1E), no morphology phenotype was discernible over the course of six successive cell cycles - normally ample enough time for phenotypes to develop following RNAi against essential flagellum or basal body components. Failure to observe any RNAi phenotype may indicate that the small amount of FOR20 still produced is sufficient for normal cell function, a critical importance for *Tb*FOR20 only in other life cycle stages – our experiments were all performed with cultured procyclic trypomastigotes, which normally replicate within a tsetse fly midgut – or functional redundancy. We note that a previous genome-wide survey of gene function also failed to elicit a *Tb*FOR20 RNAi phenotype in either procyclic or pathogenic bloodstream form *T. brucei* (Alsford et al. 2011). Further clues to the role(s) of FOR20 in flagellum assembly or function, however, perhaps comes from the localisation of FOR20 in *Giardia intestinalis*, another excavate protist. Here, hemagglutinin-tagged *Giardia* FOR20 localises along the length of paraflagellar dense rods associated with the cytoplasmic regions of axonemes that form anterior, posterior-lateral and immotile caudal flagella (Lauwaet et al. 2011). Using transmission electron microscopy, Hoeng et al. (2008) demonstrated the cytoplasmic regions of *Giardia* axonemes, which run for a considerable distance through the cytoplasm before exit from the cell body, do not correspond to elongated transition zones, but instead exhibit conventional '9+2' architecture. Thus, localisation of FOR20 orthologues in both *T. brucei* and *G. intestinalis* are not consistent with conservation of a direct role for FOR20 in basal body-membrane docking. Given its small size and paucity of recognisable domain architecture (other than TOF-LisH motifs, which assume α-helical secondary structure and mediate protein oligomerisation (Mikolajka et al. 2006, Sedjai et al. 2010)), it appears unlikely that FOR20 exhibits any intrinsic enzymatic activity. FOR20 also stands alone among FOP family proteins in that the protein lacks extensive amino acid sequence and domain architecture downstream of N-terminal TOF-LisH-PLL motifs. Conceivably, FOR20 acts as an adapter or linking protein for scaffolding and/or stabilisation of other proteins or complexes for flagellum assembly and/or basal body

docking, albeit that protein function appears non-essential in procyclic *T. brucei* under standard culture conditions.

On the antigen specificity of YL1/2 at the *T. brucei* basal body

Our ongoing interests in characterising the functions of trypanosome FOP family proteins has seen us question the identity of the basal body antigen recognised by the monoclonal antibody YL1/2. For over 30 years, YL1/2 has been used as a marker for tyrosinated α-tubulin in eukaryotic cells (Kilmartin et al. 1982). In trypanosomatids, YL1/2 has been used widely in studies defining critical events in cell morphogenesis and division (*e.g.* Sherwin et al. 1987, Wheeler et al. 2013). With cell shape so heavily dependent on a sub-pellicular microtubule corset remodelled and inherited in a semi-conservative manner (Sherwin and Gull 1989), YL1/2 is perfect for marking growth of new microtubules in which the C-terminal epitope of α-tubulin is recognisable prior to detyrosination. Yet, the prominent mature basal body signal observed in YL1/2 immunofluorescence of trypanosomatids presents a conundrum in that the centriole barrel is composed of mature microtubules and other anti-tubulin antibodies do not yield an immunofluorescence focus comparable to the YL1/2 basal body signal. It was proposed that this YL1/2 basal body signal represents a dynamic pool of tyrosinated α-tubulin awaiting IFT-mediated transport to the flagellar distal tip (Stephan et al. 2007). In that regard, we note that the axoneme of new, elongating flagella in trypanosomes is also readily detected using YL1/2 (Sherwin et al. 1987). With our demonstration that YL1/2 readily detects by immunoblot both recombinant TbRP2 (André et al. 2014) and native *Tb*RP2 (this study), the notion that basal body localised YL1/2 detects tyrosinated α-tubulin must be called into question. Note we do not call into question the value of using YL1/2 to detect tyrosinated α-tubulin at other cellular locations in trypanosomatids. The caution that we encourage echoes caveats offered by Wehland and colleagues when they mapped the epitope specificity of YL1/2 in their seminal study (Wehland et al. 1984).

In questioning where along the length of pro- and mature basal bodies *Tb*FOR20 is found, we have mapped the localisation of *Tb*FOR20, *Tb*OFD1 and *Tb*FOP/FOP1-like proteins relative to basal body antigen recognised by YL1/2, which we believe is the fourth trypanosome FOP family protein *Tb*RP2. Although one considers the transitional fibres where *Tb*RP2 is found to be at the junction of the triplet to doublet microtubule basal body transition and attach to the base of the flagellar pocket membrane contributing to a physical ciliary gate (Garcia-Gonzalo FR and Reiter JF 2017), our immunofluorescence analysis of isolated flagella indicates *Tb*RP2 proximal to *Tb*FOR20 with no evidence of *Tb*FOR20 re-localisation concomitant with transition zone elongation during pro- to basal body maturation. Returning to the original immunogold localisation of YL1/2 in *T. brucei* (Stephan et al. 2007) it is clear that there is extensive gold labelling at the level of the triplet microtubule barrel, consistent with the immunolocalisation data presented here. Collectively, relative localisations of the complete trypanosome FOP protein family can be summarised by the model shown in Fig. 8. The model highlights subtle, yet realistically functionally significant differences in the basal body localisations of trypanosome FOP family proteins. Future studies are likely to provide insight into whether these localisation differences are more broadly conserved and how they might relate to functional specialisation.

Subtleties in TOF-LisH motif-dependent targeting of centriolar proteins

The prediction of a short N-terminal extension for FOR20 orthologues from mammal-infective African *Trypanosoma* species was intriguing, not least because it is among *Leishmania* species where the average length of orthologous coding sequences is predicted

to be larger by ~15-20% than in other trypanosomatids (El-Sayed et al. 2005). Following initial annotations of the *T. brucei* nuclear genome, RNAseq analyses saw identification of an essential 'small proteome' and re-annotation of many trypanosome protein-coding genes, revealing alternative sites for *trans*-splicing (the mechanism by which the 5'-end of all protein-coding transcripts are capped) and the use of alternative start codons (Ericson et al. 2014, Kolev et al. 2010, Nilsson et al. 2010, Siegel et al. 2010). Thus, it was necessary to confirm the presence of a candidate N-terminal extension in *Tb*FOR20 and exclude the use of a potential downstream start methionine. Our analyses revealed that the extension is essential for localisation of *Tb*FOR20 at pro- and mature basal bodies, but its function in ensuring protein targeting is not masked by the presence of a large (~28 kDa), globular N-terminal YFP tag.

The TOF-LisH motif combination is a conserved, if seldom used, feature of centriolar proteins. In mammalian cells inappropriate use of TOF-LisH targeting can have pathological consequences (Lelièvre et al. 2008). Thus, there is wider interest in understanding how TOF-LisH motif combinations mediate centriolar localisation. The cryptic nature of how the *Tb*FOR20 N-terminal extension contributes to pro- and mature basal body targeting, combined with the observation that *Tc*FOR20, which naturally lacks any N-terminal extension, readily localises to *T. brucei* pro- and mature basal bodies indicate there remains much to learn regarding how TOF-LisH motif combinations, and additional motif elements, contribute to the targeting of different proteins to distinct basal body sites. Successive C-terminal deletions of *Tb*RP2 all target myc-epitope tagged proteins to mature trypanosome basal bodies providing the TOF-LisH motifs remain intact (André et al. 2014). This suggests a principal functional role of these motifs is to confer protein targeting. RNAi phenotypes of *T. brucei* FOP family proteins show no discernible overlap, and proteomic screening has not identified any other FOP family member as an interacting or near neighbour protein to *Tb*RP2[2]. These observations emphasise a likely importance of molecule-specific interactions in underpinning the use of this efficient targeting determinant to eukaryotic basal bodies.

Methods

Cell culture: Procyclic *T. brucei* (cell line 927smox; Poon et al. 2012; Lister strain 427) were cultured in SDM-79 medium containing 10% v/v heat-inactivated foetal bovine serum, as described previously (Brun and Schönenberger, 1979). Constitutive expression of YFP- and myc epitope-tagged proteins occurred in a 427 genetic background. *Tb*FOR20::myc proteins (full length and without the N-terminal extension) were also expressed in a 927smox background. 927smox is genetically modified to express a tetracycline-repressor protein and T7 RNA polymerase, meaning these cells are amenable to inducible gene expression or RNAi. The *Tb*RP2 RNAi mutant (on a 927smox background) was generated as described previously (André et al. 2014). Tb*FOR20* RNAi was also induced on a 927smox background. *Tb*FOR20$^{\Delta 22\text{-}151}$::RP2$^{\Delta 1/\Delta 134\text{-}463}$::myc (sub-cloned as described below) and *Tc*FOR20::myc were constitutively expressed in a 427 background. Cells were transfected and stable transformants selected using blasticidin (10 μg ml^{-1}) or hygromycin B (50 μg ml^{-1}) according to standard methods (McCulloch et al. 2004).

Plasmid constructs: For constitutive expression of YFP-tagged *Tb*FOR20 and *Tb*OFD1 (André et al. 2016) from endogenous chromosomal loci amplicons corresponding to XbaI-XhoI digested partial open reading frames and XhoI-BamHI-digested upstream intergenic regions were sub-cloned in three-way ligations into XbaI-BamHI-digested pEnT6B-Y (Kelly et al. 2007). For expression of myc epitope-tagged *Tb*FOR20 variants and *Tc*FOR20, open reading frames minus the stop codon were PCR amplified using forward and reverse priming oligonucleotides synthesised with HindIII and XhoI restriction sites at the 5' end, respectively. HindIII-XhoI-digested PCR amplicons were then ligated with HindIII-XhoI-digested pDEX377$_{TbRP2::myc}$ (André et al. 2014). To prepare a chimeric gene for expression of the first 21 amino acids of *Tb*FOR20 fused to amino acids 2-133 of *Tb*RP2 a two-step

[2] Qi X et al. in preparation.

PCR strategy was used. Using gDNA template, forward primer 5'-ggcgagagtcctttaacgcacccgaggcgactacaacctaccaagcgaagg-3' and reverse primer 5'tta*ggatcc*gctattggcacccgccgcgcccggtg-3' (BamHI site italicised) was used to amplify coding sequence for amino acids 2-133 of *Tb*RP2 preceded in frame by amino acids 11-21 of *Tb*FOR20. The resulting PCR amplicon was purified and used as template for a second PCR using the same reverse primer as PCR 1, but with a new forward primer, 5'-cgcaagcttatggaggaaagggaggagggagaggtgcggcgagagtcctttaacgc (HindIII site underlined), to create coding sequence with the amino acids 1-10 from *Tb*FOR20 also added. The resulting purified PCR amplicon was digested with BamHI and HindIII and ligated with HindIII-XhoI-digested pDEX377$_{TbRP2::myc}$, thereby creating pDEX377$_{TbFOR20\Delta22\text{-}151::RP2\Delta1/\Delta134\text{-}463::myc.}$ For Tb*FOR20* RNAi the open reading frame was sub-cloned into BamHI-HindIII-digested p2T7$_{177}$ (Wickstead et al. 2002). NotI-digested constructs with genes encoding C-terminal myc-tagged proteins or Tb*FOR20* RNAi insert were transfected as described above. All plasmids were sequenced using ABI prism sequencing technology (Source Bioscience).

Microscopy and immunoblotting: For fluorescence microscopy, cells were settled onto glass coverslips and either fixed directly with paraformaldehyde (3.7 % w/v in PBS) or extracted for 0.5 min with PEME containing 1% v/v NP-40 prior to fixation. For preparation of isolated flagella, exponentially growing cells were harvested and cytoskeletons extracted on ice for 10 minutes in PEME 1% v/v NP-40. Cytoskeleton pellets were harvested and flagella extracted twice on ice for 10 minutes in PEME 1M NaCl; flagella were collected by centrifugation after each extraction (13000 *g*, 30 min, 4°C), before settling onto glass coverslips. Fixed preparations were decorated for indirect immunofluorescence with the monoclonal antibodies BBA4 or YL1/2 as described previously (Woodward et al. 1995 and Sherwin et al. 1987, respectively) and decoration with the anti-myc monoclonal antibody was performed following the instructions of the supplier (Myc, Abcam). Cells were imaged at 60x magnification using either an Applied Precision DeltaVision Microscope and a Roper Scientific Photometrics Cool SNAP HQ camera or for the images shown in Fig. 2D-E a LSM880 Laser Scanning Confocal microscope (Zeiss). Expected sizes of all tagged proteins expressed in the study were confirmed by immunoblotting (Supp. Fig. 1); 10% acrylamide gels were used for SDS-PAGE prior to protein transfer onto Hybond P membranes (GE Healthcare).

Acknowledgements

This work was in part supported by grants BBG0210581 andBBF0109311 awarded to MLG and PGM by the UK Biological and Biotechnological Sciences Research Council. We thank Kevin Tyler (UEA, UK) for the kind gift of *T. cruzi* genomic DNA.

References

Absalon S, Blisnick T, Kohl L, Toutirais G, Dore G, Julkowska D, Tavenet A, Bastin P (2008) Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes. Mol Biol Cell 19:929-944.

Akiyoshi B, Gull K (2013) Evolutionary cell biology of chromosome segregation: insights from trypanosomes. Open Biol 3:130023.

Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, Hertz-Fowler C, Horn D (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res 21:915-924.

André J, Harrison S, Towers K, Qi X, Vaughan S, McKean PG, Ginger ML (2013) The tubulin cofactor C family member TBCCD1 orchestrates cytoskeletal filament formation. J Cell Sci 126:5350-5356.

André J, Kerry L, Qi X, Hawkins E, Drizyte K, Ginger ML, McKean PG (2014) An alternative model for the role of RP2 protein in flagellum assembly in the African trypanosome. J Biol Chem 289:464-475.

Aubusson-Fleury A, Lemullois M, de Loubresse NG, Laligne C, Cohen J, Rosnet O, Jerka-Dziadosz M, Beisson J, Koll F (2012) The conserved centrosomal protein FOR20 is required for assembly of the transition zone and basal body docking at the cell surface. J Cell Sci 125:4395-4404.

Azimzadeh J, Nacry P, Christodoulidou A, Drevensek S, Camilleri C, Amiour N, Parcy F, Pastuglia M, Bouchez D (2008) *Arabidopsis* TONNEAU1 proteins are essential for preprophase band formation and interact with centrin. Plant Cell 20:2146-2159.

Bengueddach H, Lemullois M, Aubusson-Fleury A, Koll F (2017) Basal body positioning and anchoring in the multiciliated cell *Paramecium tetraurelia*: roles of OFD1 and VFL3. Cilia 6:6.

Broadhead R, et al. (2006) Flagellar motility is required for the viability of the bloodstream trypanosome. Nature 440:224-227.

Brun R, Schönenberger M (1979) Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. Acta Trop **36**: 289-292.

Carvalho-Santos Z, Azimzadeh J, Pereira-Leal JB, Bettencourt-Dias M (2011) Evolution: Tracing the origins of centrioles, cilia, and flagella. J Cell Biol 194:165-175.

Davidge JA, Chambers E, Dickinson HA, Towers K, Ginger ML, McKean PG, Gull K (2006) Trypanosome IFT mutants provide insight into the motor location for mobility of the flagella connector and flagellar membrane formation. J Cell Sci 119:3935-3943.

Dawe HR, Farr H, Gull K (2007) Centriole/basal body morphogenesis and migration during ciliogenesis in animal cells. J Cell Sci 120:7-15.

El-Sayed NM, et al. (2005) Comparative genomics of trypanosomatid parasitic protozoa. Science 309:404-409.

Ericson M, Janes MA, Butter F, Mann M, Ullu E, Tschudi C (2014) On the extent and role of the small proteome in the parasitic eukaryote *Trypanosoma brucei*. BMC Biol 12:14.

Fritz-Laylin LK, Cande WZ (2010) Ancestral centriole and flagella proteins identified by analysis of *Naegleria* differentiation. J Cell Sci 123:4024-4031.

Fritz-Laylin LK, Levy YY, Levitan E, Chen S, Cande WZ, Lai EY, Fulton C (2016) Rapid centriole assembly in *Naegleria* reveals conserved roles for both de novo and mentored assembly. Cytoskeleton (Hoboken) 73:109-116.

Garcia-Gonzalo FR, Reiter JF (2017) Open Sesame: how transition fibres and the transition zone control ciliary composition. Cold Spring Harb Perspect Biol 9:a028134.

Gleunz E, Povelones ML, Englund PT, Gull K (2011) The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. Mol Cell Biol 31:1012-1021.

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AG, Roger AJ (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci U S A 106:3859-3864.

Hayes P, Varga V, Olego-Fernandez S, Sunter J, Ginger ML, Gull K (2014) Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosome morphology. J Cell Biol 206:377-384.

Hodges ME, Scheumann N, Wickstead B, Langdale JA, Gull K (2010) Reconstructing the evolutionary history of the centriole from protein components. J Cell Sci 123:1407-1413.

Hoeng JC, Dawson SC, House SA, Sagolla MS, Pham JK, Mancuso JJ, Lowe J, Cande WZ (2008) High-resolution crystal structure and in vivo function of a kinesin-2 homologue in *Giardia intestinalis*. Mol Biol Cell 19:3124-3137.

Höög JL, Lacomble S, O'Toole ET, Hoenger A, McIntosh JR, Gull K (2014) Modes of flagellar assembly in *Chlamydomonas reinhardtii* and *Trypanosoma brucei*. ELife 3:e01479.

Katz LA, Grant JR (2015) Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. Syst Biol 64:406-415.

Keller LC, Romijn EP, Zamora I, Yates JR, 3rd, Marshall WF (2005) Proteomic analysis of isolated *Chlamydomonas* centrioles reveals orthologs of ciliary-disease genes. Curr Biol 15:1090-1098.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre^2 webportal for protein modelling, prediction and analysis. Nat Protoc 10:845-858.

Kelly S, Ivens A, Manna PT, Gibson W, Field MC (2014) A draft genome for the African crocodilian trypanosome *Trypanosoma grayi*. Sci Data 1:140024.

Kelly S, Reed J, Kramer S, Ellis L, Webb H, Sunter J, Salje J, Marinsek N, Gull K, Wickstead B, Carrington M (2007) Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged protein from endogenous and ectopic gene loci. Mol Biochem Parasitol 154:103-109.

Kilburn CL, Pearson CG, Romijn EP, Meehl JB, Giddings TH, Jr., Culver BP, Yates JR, 3rd, Winey M (2007) New *Tetrahymena* basal body protein components identify basal body domain structure. J Cell Biol 178:905-912.

Kilmartin JV, Wright B, Milstein C (1982) Rat monoclonal antitubulin antibodies derived by using a new nonsecreting rat cell line. J Cell Biol 93:576-582.

Kolev NG, Ramey-Butler K, Cross GA, Ullu E, Tschudi C (2012) Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. Science 338:1352-1353.

Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, Tschudi C (2010) The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. PLoS Pathog 6:e1001090.

Lacomble S, Vaughan S, Gadelha C, Morphew MK, Shaw MK, McIntosh JR, Gull K (2010) Basal body movements orchestrate membrane organelle division and cell morphogenesis in *Trypanosoma brucei*. J Cell Sci 123:2884-2891.

Lacomble S, Vaughan S, Gadelha C, Morphew MK, Shaw MK, McIntosh JR, Gull K (2009) Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in trypanosomes revealed by electron microscope tomography. J Cell Sci 122:1081-1090.

Lauwaet T, Smith AJ, Reiner DS, Romijn EP, Wong CC, Davids BJ, Shah SA, Yates JR, 3rd, Gillin FD (2011) Mining the *Giardia* genome and proteome for conserved and unique basal body proteins. Int J Parasitol 41:1079-1092.

Lee JY, Stearns T (2013) FOP is a centriolar satellite protein involved in ciliogenesis. PLoS One 8:e58589.

Lelièvre H, Chevrier V, Tassin AM, Birnbaum D (2008) Myeloproliferative disorder FOP-FGFR1 fusion kinase recruits phosphoinositide-3 kinase and phospholipase Cγ at the centrosome. Mol Cancer 7:30.

Mikolajka A, Yan X, Popowicz GM, Smialowski P, Nigg EA, Holak TA (2006) Structure of the N-terminal domain of the FOP (FGFR1OP) protein and implications for its dimerization and centrosomal localization. J Mol Biol 359:863-875.

McCulloch R, Vassella E, Burton P, Boshart M, Barry JD (2004) Transformation of monomorphic and pleomorphic *Trypanosoma brucei*. Methods Mol Biol **262**: 53-86.

Moran J, McKean PG, Ginger ML (2014) Eukaryotic flagella: variations in form, function, and composition during evolution. Bioscience 64:1103-1114.

Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi I, Ochsenreiter T (2010) Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. PLoS Pathog 6:e1001037.

Ogbadoyi EO, Robinson DR, Gull K (2003) A high-order trans-membrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes. Mol Biol Cell 14:1769-1779.

Olego-Fernandez S, Vaughan S, Shaw MK, Gull K, Ginger ML (2009) Cell morphogenesis of *Trypanosoma brucei* requires the paralogous, differentially expressed calpain-related proteins CAP5.5 and CAP5.5V. Protist 160:576-590.

Peacock L, Bailey M, Carrington M, Gibson W (2014) Meiosis and haploid gametes in the pathogen *Trypanosoma brucei*. Curr Biol 24:181-186.

Poon SK, Peacock L, Gibson W, Gull K, Kelly S (2012) A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. Open Biol 2:110037.

Portman N, Gull K (2014) Identification of paralogous life-cycle stage specific cytoskeletal proteins in the parasite *Trypanosoma brucei*. PLoS One 9:e106777.

Prensier G, Dubremetz JF, Schrevel J (2008) The unique adaptation of the life cycle of the coelomic gregarine Diplauxis hatti to its host *Perinereis cultrifera* (Annelida, Polychaeta): an experimental and ultrastructural study. J Eukaryot Microbiol 55:541-553.

Rogozin IB, Basu MK, Csuros M, Koonin EV (2009) Analysis of rare genomic changes does not support the unikont-bikont phylogeny and suggests cyanobacterial symbiosis as the point of primary radiation of eukaryotes. Genome Biol Evol 1:99-113.

Sedjai F, et al. (2010) Control of ciliogenesis by FOR20, a novel centrosome and pericentriolar satellite protein. J Cell Sci 123:2391-2401.

Sharma R, Peacock L, Gluenz E, Gull K, Gibson W, Carrington M (2008) Asymmetric cell division as a route to reduction in cell length and change in cell morphology in trypanosomes. Protist 159:137-151.

Sherwin T, Schneider A, Sasse R, Seebeck T, Gull K (1987) Distinct localisation and cell cycle dependence of COOH terminally tyrosinolated α-tubulin in the microtubules of *Trypanosoma brucei brucei*. J Cell Biol 104:439-446.

Sherwin T, Gull K (1989) Visualization of detyrosination along single microtubules reveals novel mechanisms of assembly during cytoskeletal duplication in trypanosomes. Cell 57:211-221.

Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GA (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Res 38:4946-4957.

Singla V, Romaguera-Ros M, Garcia-Verdugo JM, Reiter JF (2010) Ofd1, a human disease gene, regulates the length and distal structure of centrioles. Dev Cell 18:410-424.

Spinner L, Pastuglia M, Belcram K, Pegoraro M, Goussot M, Bouchez D, Schaefer DG (2010) The function of TONNEAU1 in moss reveals ancient mechanisms of division plane specification and cell elongation in land plants. Development 137:2733-2742.

Stephan A, Vaughan S, Shaw MK, Gull K, McKean PG (2007) An essential quality control mechanism at the eukaryotic basal body prior to intraflagellar transport. Traffic 8:1323-1330.

Sunter JD, Benz C, André J, Whipple S, McKean PG, Gull K, Ginger ML, Lukes J (2015) Modulation of flagellum attachment zone protein FLAM3 and regulation of the cell shape in *Trypanosoma brucei* life cycle transitions. J Cell Sci 128:3117-3130.

Tassin AM, Lemullois M, Aubusson-Fleury A (2015) *Paramecium tetraurelia* basal body structure. Cilia 5:6.

Vaughan S, Gull K (2015) Basal body structure and cell cycle-dependent biogenesis in *Trypanosoma brucei*. Cilia 5:5.

Vaughan S, Kohl L, Ngai I, Wheeler RJ, Gull K (2008) A repetitive protein essential for the flagellum attachment zone filament structure and function in *Trypanosoma brucei*. Protist 159:127-136.

Vedrenne C, Giroud C, Robinson DR, Besteiro S, Bosc C, Bringaud F, Baltz T (2002) Two related subpellicular cytoskeleton-associated proteins in *Trypanosoma brucei* stabilize microtubules. Mol Biol Cell 13:1058-1070.

Vincensini L, Blisnick T, Bastin P (2011) 1001 model organisms to study cilia and flagella. Biol Cell 103:109-130.

Wehland J, Schroder HC, Weber K (1984) Amino acid sequence requirements in the epitope recognized by the alpha-tubulin-specific rat monoclonal antibody YL 1/2. EMBO J 3:1295-1300.

Wheeler RJ, Scheumann N, Wickstead B, Gull K, Vaughan S (2013) Cytokinesis in *Trypanosoma brucei* differs between bloodstream and tsetse trypomastigote forms: implications for microtubule-based morphogenesis and mutant analysis. Mol Microbiol 90:1339-1355.

Wickstead B, Ersfeld K, Gull K (2002) Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. Mol Biochem Parasitol 125:211-216.

Woodward R, Carden MJ, Gull K (1995) Immunological characterization of cytoskeletal proteins associated with the basal body, axoneme and flagellum attachment zone of *Trypanosoma brucei*. Parasitology 111:77-85.

Figure Legends

Figure 1. Localisation of *Tb*FOR20. (**A**) Localisation of YFP::*Tb*FOR20 at pro- and mature basal bodies in procyclic *T. brucei*; the main panel images show detection in detergent-extracted cytoskeletons of YFP::FOR20 distal to the antigen detected by monoclonal antibody BBA4. The inset shows YFP::FOR20 localises only to basal bodies in intact cells. 6-Diamidino-2-phenylindole (DAPI) was used to detect nuclear DNA (N) and the mitochondrial genome (or kinetoplast, K) to which pro- and mature basal bodies are physically attached. (**B**). YFP::FOR20 localisation in detergent-extracted cytoskeletons in comparison with the mature basal body antigen(s) detected by monoclonal antibody YL1/2. (**C**) Duplication of YFP::FOR20 signals coincides with maturation of the existing probasal body and the biogenesis of new probasal bodies during kinetoplast replication (as denoted by the ‘domed’ kinetoplast; Gluenz et al. 2011). (**D**) Localisation of *Tb*FOR20::myc at pro-

and mature basal bodies in detergent-extracted procyclic *T. brucei*. Formaldehyde fixed cytoskeletons and whole cells (not shown) were decorated for indirect immunofluorescence using anti-myc monoclonal primary antibody. Scale bars in all main panels indicate 5 µm and in the inset of (A) 1 µm.

Figure 2. Localisation of *T. brucei* FOP family proteins on isolated flagella. Spatial resolution of indirect immunofluorescence signals for (**A**) YL1/2 from *Tb*YFP::FOR20, (**B**) polyclonal anti-RP2 from *Tb*YFP::OFD1, (**C**) anti-RP2 from *Tb*YFP::FOP, (**D**) BBA4 and anti-RP2 from *Tb*YFP::FOR20, (**E**) BBA4, anti-RP2, and *Tb*FOR20::Myc. Main panel scale bars, 5 µm.

Figure 3. Immunoblot detection of *Tb*RP2 by monoclonal antibody YL1/2. (**A**) Detection of *Tb*RP2 depletion in procyclic *Tb*RP2 RNAi mutants by immunoblot. The left-hand immunoblot shows detection of both *Tb*RP2 (predicted mass 50.6 kDa) and tyrosinated α-tubulin (predicted mass 49.8 kDa) from whole cell extracts (2 x 10^6 cell equivalents loaded per lane) by YL1/2; proteins were separated by SDS-PAGE using a 10 % polyacrylamide gel. The right-hand immunoblot shows the depletion of *Tb*RP2 from *Tb*RP2 RNAi mutants using polyclonal, affinity-purified antibodies raised against recombinant *Tb*RP2. Cells were induced for RNAi for 48 h before preparation of cell lysates (RNAi+ lanes). (**B**) Immunoblot detection of recombinant *Tb*RP2 by YL1/2; the amount of recombinant protein loaded per lane is indicated; reproduced from André et al. (2014) under the terms of a Creative Common Attribution 3.0 unported licence.

Figure 4. An N-terminal extension unique to FOR20 proteins from *T. brucei* and closely related African trypanosome species. (**A**) Evolutionary relationships between trypanosomatid species. (**B**) Clustal Omega alignment of amino acid sequences for FOR20 orthologues from kinetoplastid protists and other, evolutionary diverse flagellates. Positions of amino acid identity (*) and conservation (:) are indicated; negatively charged amino acids within the N-terminal extensions of FOR20 from *T. vivax*, *T. congolense*, and *T. brucei* are italicised. 'African' trypanosome species are in red; trypanosomatid species more closely related phylogenetically to *T. cruzi* than to the *T. brucei* clade are in purple; other kinetoplastids are in blue. Taxonomic abbreviations: Bs, *Bodo saltans*; Cr, *Chlamydomonas reinhardtii*; Dr, *Danio rerio*; Gg, *Gallus gallus*; Gl, *Giardia lamblia*; Hs, *Homo sapiens*; Lm, *Leishmania major*; Ng, *Naegleria gruberi*; Nv, *Nematostella vectensis*; Pt, *Paramecium tetraurelia*; Tb, *Trypanosoma brucei*; Tc, *T. cruzi*; Tcg, *T. congolense*; Tg, *T. grayi*; Tr, *T. rangeli*; Tt, *Tetrahymena thermophilia*; Tv, *Trypanosoma vivax*; Tva, *Trichomonas vaginalis*; Xl, *Xenopus laevis*.

Figure 5. Targeting of trypanosomatid FOR20 to pro- and mature basal bodies in procyclic *T. brucei*. (**A**) Inducible expression of *Tb*FOR20$^{\Delta 1\text{-}21}$::myc: main panel, a formaldehyde-fixed cell decorated for indirect immunofluorescence using anti-myc monoclonal primary antibody; inset, absence of *Tb*FOR20$^{\Delta 1\text{-}21}$::myc from detergent-extracted cytoskeletons. (**B**) Mature basal body localisation only for constitutively expressed *Tb*FOR20$^{\Delta 22\text{-}151}$::RP2$^{\Delta 1/\Delta 134\text{-}463}$::myc; panels and merged insets correspond to kinetoplast-basal body regions from 1K1N, 2K1N and 2K2N cells where the mature basal body is decorated by a YFP fusion of the *T. brucei* orthologue of FOP family protein OFD1. (**C-D**) Localisation of *Tc*FOR20::myc at both pro- and mature basal bodies; cells were decorated with anti-myc plus either BBA4 (**C**) or anti-

*Tb*RP2 antibodies (**D**). K, kinetoplast; m, mature basal body; p, probasal body. Scale bars correspond to 5 µm (**A-D**) and 1 µm (inset **B** and **C-D**).

Figure 6. Phyre[2] prediction of *Tb*FOR20 secondary structure. Relative predictions of confidence for disorder and α-helical regions are indicated. TOF and LisH motifs are known to adopt α-helical conformations: the *Tb*FOR20 TOF motif spans amino acids 30-60 and the LisH motif amino acids 78-104.

Figure 7. Site-directed mutagenesis of *Tb*FOR20 TOF and LisH motifs. (**A**) E37A mutation abrogates protein targeting to the basal/probasal bodies. Main panels show punctate accumulation of protein in whole cells; inset shows the absence of protein from basal body region in detergent extracted cytoskeletons. (**B-C**) Basal body localisation of F54A (B) and Y88A (C) site-directed mutants.

Figure 8. Organisation of FOP family proteins at mature trypanosome basal bodies. Positions of FOP family proteins are shown relative to the proximal end antigen recognised by monoclonal antibody BBA4 and TAC component TBCCD1 (an additional tubulin cofactor C domain-containing protein (André et al. 2013)).

Supplementary Figure 1. Immunoblot analysis of YFP- and epitope-tagged proteins analysed in this study. Whole cell lysates corresponding to 2 x10^6 cell equivalents were loaded in all lanes for all panels. (**A**) Constitutive expression of YFP::*Tb*FOR20 (expected molecular mass ~45 kDa). (**B**) Doxycycline-inducible expression of *Tb*FOR20::myc E37A, F54A, and Y88A site-directed mutants; protein levels were determined relative to *Tb*AKF, an adenylate kinase isoform detected using protein-specific polyclonal antisera. (**C**) Doxycycline-inducible expression of *Tb*FOR20::myc (~20 kDa) and *Tb*FOR20$^{\Delta 1\text{-}21}$::myc (~18 kDa); -, no doxycycline added; +, doxycycline-dependent induction of gene expression for 24 h. (**D**) Constitutive expression of *Tc*FOR20::myc (~18 kDa; lane 2) and *Tb*FOR20$^{\Delta 22\text{-}151}$::RP2$^{\Delta 1/\Delta 134\text{-}463}$::myc (~21 kDa; lane 3); *Tb*FOR20::myc (~20 kDa; lane 1) was immunoblotted for comparison. (**E**) Time-dependent depletion of *Tb*YFP::FOR20 expression in *Tb*FOR20 RNAi mutants. RNAi was induced in cells modified at an endogenous chromosomal location for expression of *Tb*YFP::FOR20; levels of the YFP-fusion protein were determined at the time points indicated, relative to β-tubulin detected by KMX-1 monoclonal antibody. Other primary monoclonal antibodies used to probe immunoblots were BB2 (panels A and E, recognising the Ty-epitope engineered into the YFP-fusion) or AbCam anti-myc 9E10 (panels B-D).

Figure 1 re-revised

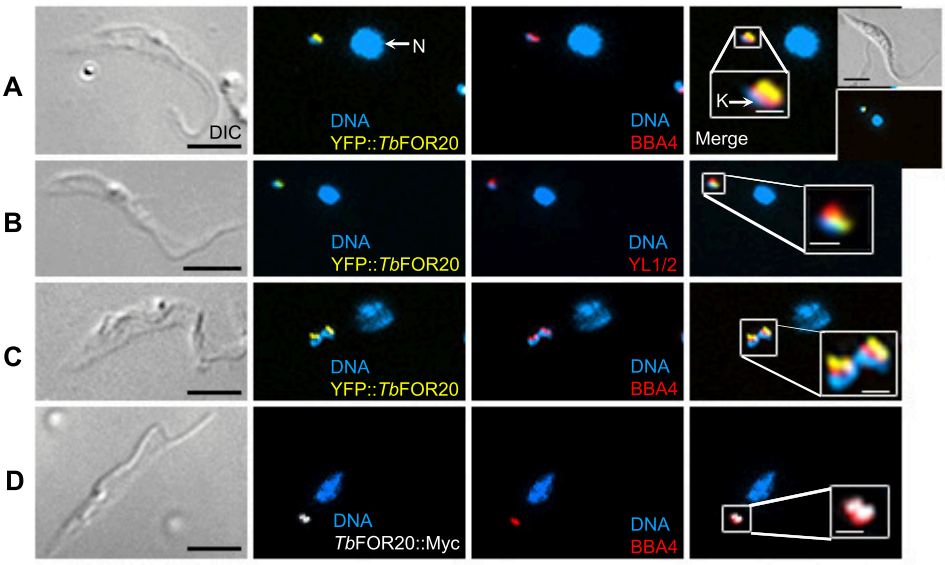


Fig. 1

Figure 2 re-revised

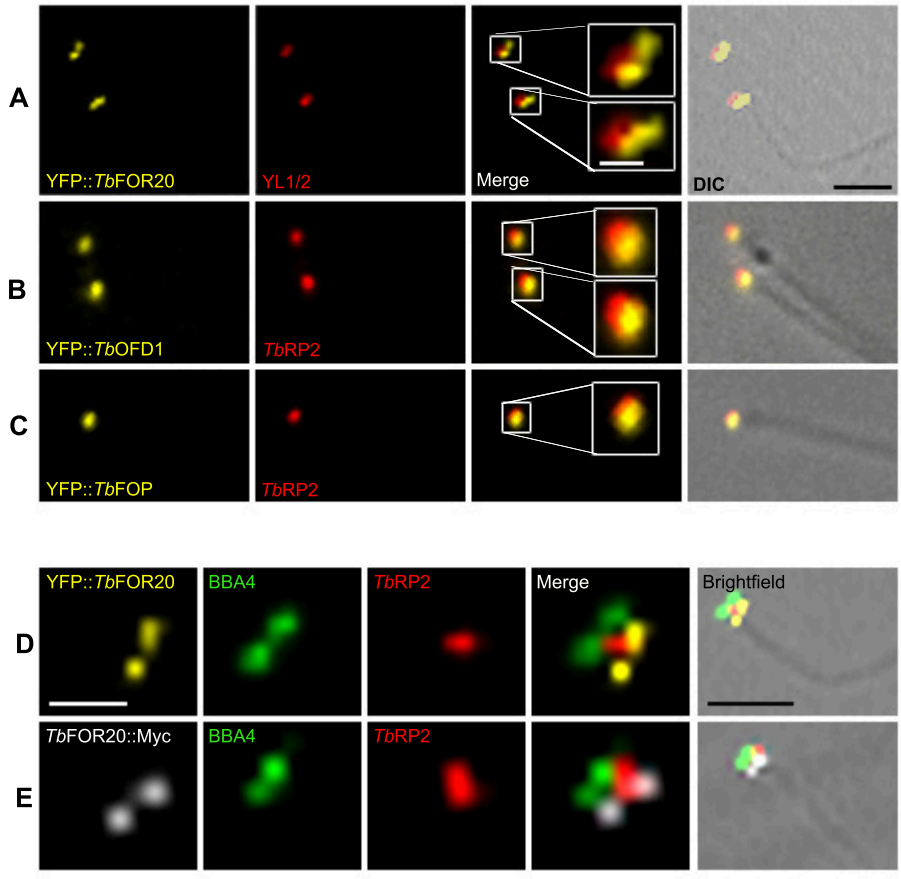
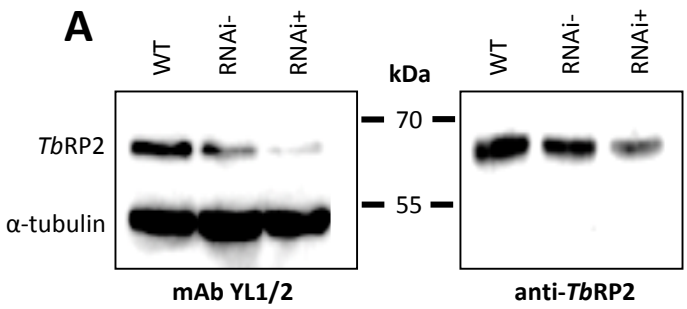


Fig. 2

Figure 3



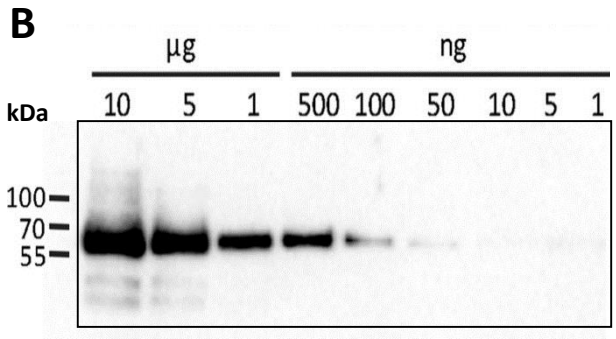
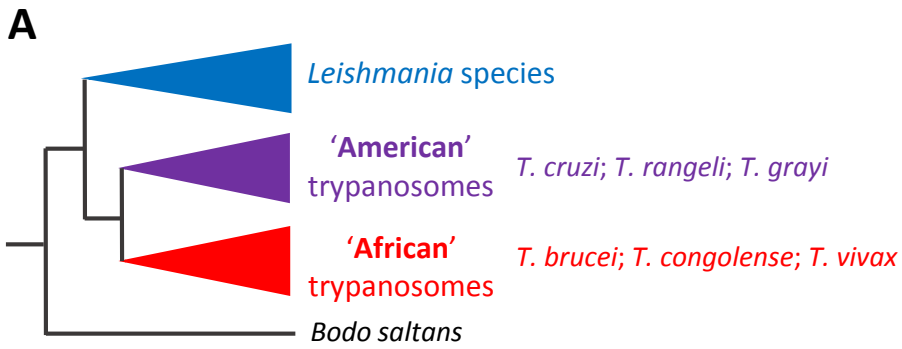


Fig. 3

Figure 4



B

```
Xl      ----------------------MATVGDLKAVVKDTLEKRGVMGQLKARVRAEVFEALDD     38
Gg      ----------------------MATIAELKAVLKDTLEKRGALRQIKARIRAEVFNALDD     38
Hs      ----------------------MATVAELKAVLKDTLEKKGVLGHLKARIRAEVFNALDD     38
Dr      ----------------------MATITELKSALRETLEARGVLGQLKARVRAEVFSALDD     38
Nv      ---------------------MATSSEKLKEVLKETLDNRGILGQIRARIRAEVFSALDD     39
Cr      ----------------------MASVEDLKDALRENLDRSGKLRQLKAQLRADVYNALHN     38
Tt      ----------------------MASVNELKDVLKETLEEKGVLTKIRAKIRAEIFNTLND     38
Pt      ----------------------MTTINEMKDALKETLESRGVLSQLRARIRAEIFNALNE     38
Ng      ----------------------MSLEEIKDILKDNLEKRGVLNKIRANLRAEIFKTFEE     37
Gl      --------------------------MDEIKSIILQNMKQRGELAVLQAKVRESVLRALEG     35
Tva     ---------------------MSVSLGDLNDAVVASLRETGKLGQITAQIRAEIYRILTE     39
Tb      MEEREEGEVRRESFNAPE-ATTMSKHGSLKDAMRQVLETKGVIDHVKAELRAAIFHSLQE     59
Tcg     --MEEEDGIPYKSFSAPASVAGTTKHGSLKDAMREVLETKGVIDHVKAELRAAIFHSLQE     58
Tv      --MDEPGE------HPVAPATGVPRQEALKDAMREIMETKGVINHVKAELRAAIFHSLQE     52
Tc      --------------------MAMARQESLKEAMREVLETKGVMDHVKAELRAAIFHALQD     40
Tr      ----------------------MTKQESLKEAMREVLETKGVMDHVKAELRAAIFHALQD     38
Tg      ----------------------MSNQESLKGAMREVLEAKGVIDHVKAELRAAIFQSLQD     38
Lm      ----------------------MPNQESLKAAMRDSLEANGTISRIKAELRAAIFERLSD     38
Bs      --------------------MATNDAAALKAAMRETLENNGQLDNIKAQLRAVVFQALDA     40
                                    ::  :   :   * :  : *.:*  :   :
```

Fig. 4

Figure 5

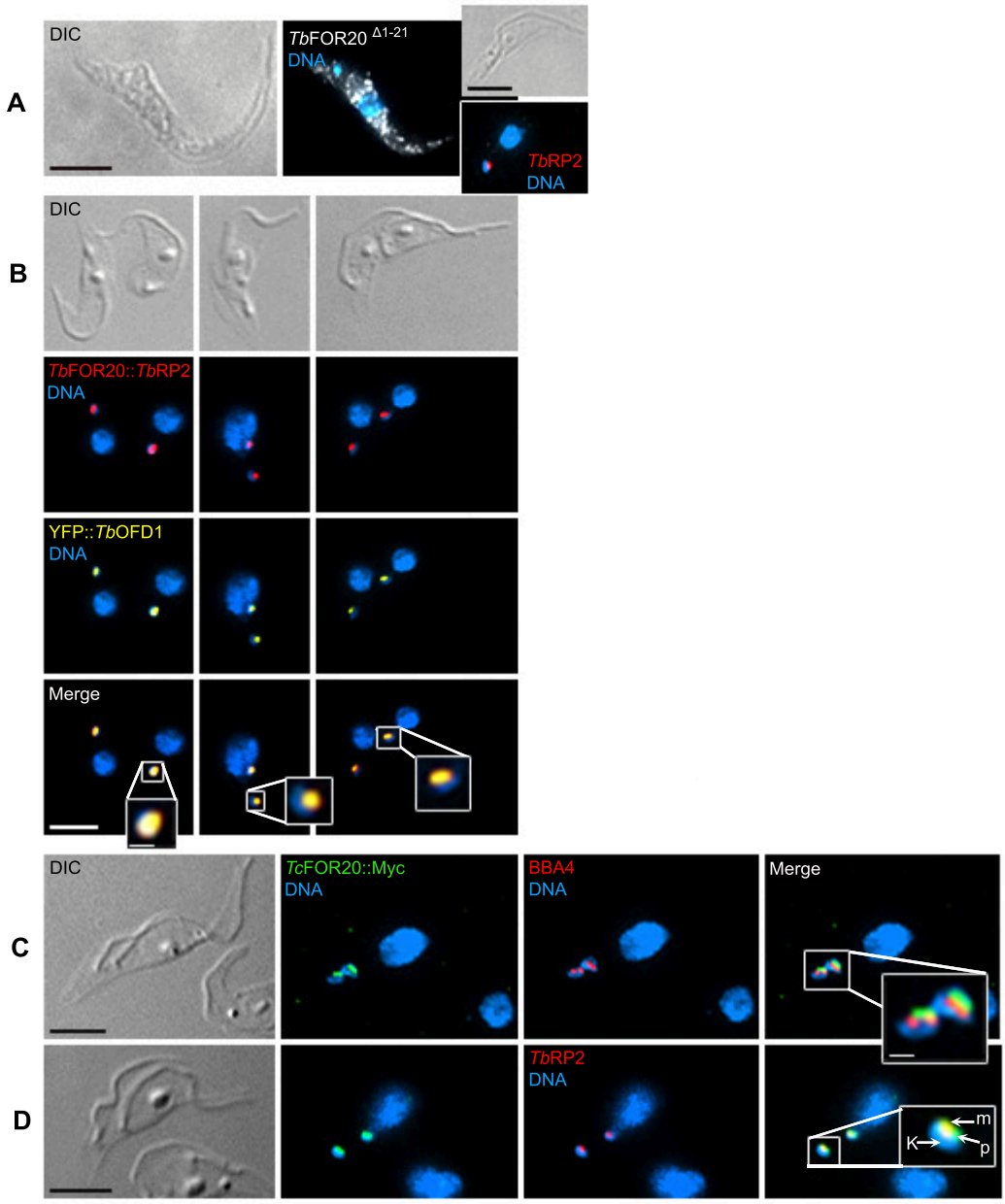


Fig. 5

Figure 6

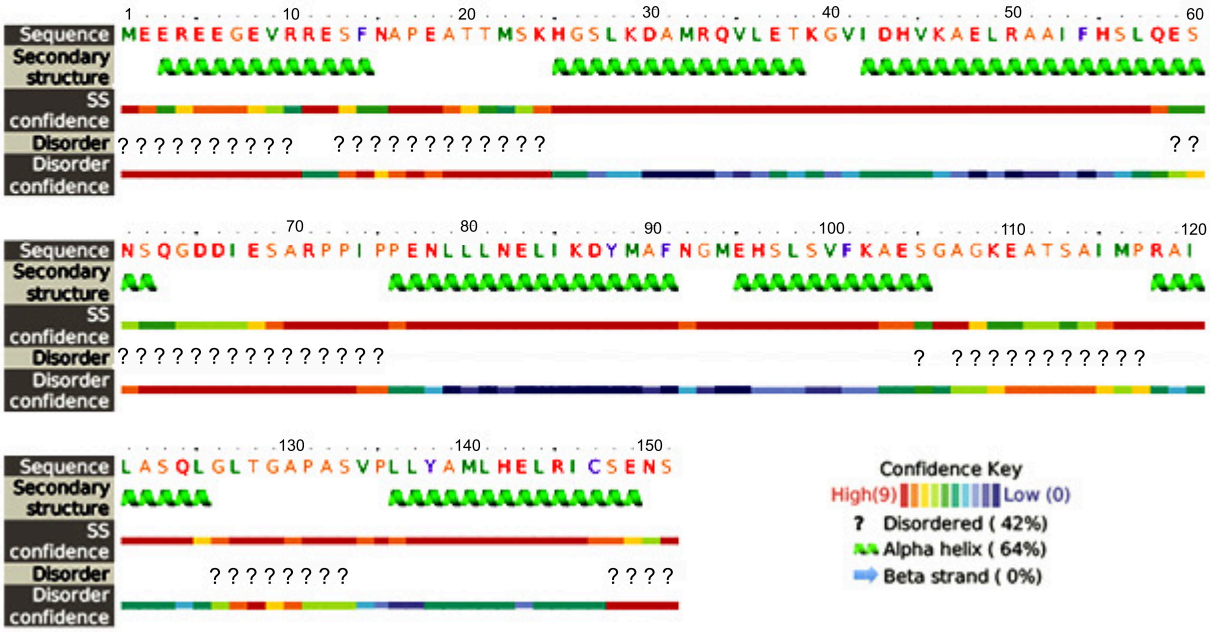


Fig. 6

Figure 7

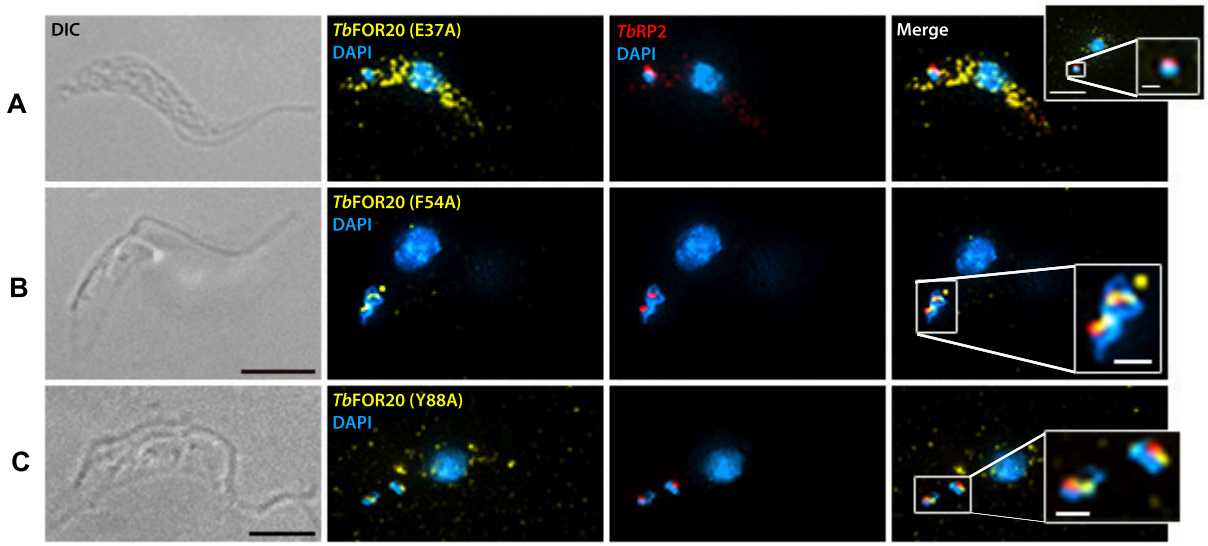


Fig. 7

Figure 8 as requested

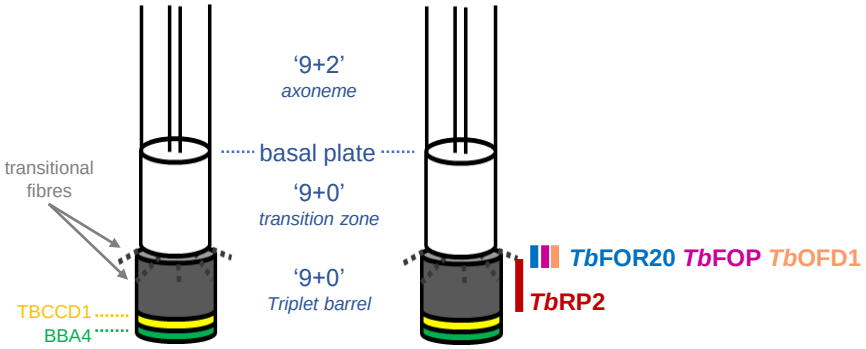


Fig. 8

Supplement Figure 1

Click here to download Supplement Material: for20_Supp_Fig_1.pdf