

Lineage-specific proteins essential for endocytosis in trypanosomes

Paul T. Manna¹, Samson O. Obado², Cordula Boehm¹, Catarina Gadelha³,
Andrej Sali⁴, Brian T. Chait², Michael P. Rout² and Mark C. Field^{1*}

¹School of Life Sciences, University of Dundee, Dundee, Scotland, DD1 5EH, UK, ²The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA, ³School of Life Sciences, University of Nottingham, Nottingham NG2 7UH and ⁴California Institute for Quantitative Biosciences, University of California, San Francisco, CA 94158, USA.

Corresponding author: Telephone: +44 (0)751-550-7880, email:
mfield@mac.com

Key words: *Trypanosoma*, trafficking, clathrin, cytoskeleton, flagellum, evolution, morphology

Abstract

Clathrin-mediated endocytosis (CME) is the most ancient endocytic mechanism known, and in many lineages the sole mechanism for internalization. Significantly, in mammalian cells CME is responsible for the vast bulk of endocytic flux and has likely undergone multiple adaptations to accommodate specific requirements by individual species. In African trypanosomes we previously demonstrated that CME is AP-2 independent, that orthologues to many of the animal and fungal CME protein cohort are absent, and that a novel, trypanosome-restricted protein cohort interacts with clathrin and drives CME. Here we used a novel cryo-milling affinity isolation strategy to preserve transient low affinity interactions, giving the most comprehensive trypanosome clathrin interactome to date. We identified the trypanosome AP-1 complex, TbEpsinR, several endosomal SNAREs plus orthologues of SMAP and the AP-2 associated kinase AAK1. Novel lineage-specific proteins were identified, which we designate TbCAP80 and TbCAP141. Their depletion produced extensive defects in endocytosis and endomembrane system organisation, revealing a novel molecular pathway subtending an early-branching and highly divergent form of CME, which is conserved and likely functionally important across the kinetoplastid parasites.

Introduction

Eukaryotes originated from a lineage within the Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota (TACK) archaeal clade (Guy and Ettema, 2011; Raymann et al., 2015). Subsequent elaboration of the eukaryotic cell during eukaryogenesis gave rise to the nucleus, endomembrane system and acquisition of the mitochondrion (Martin et al., 2015). Following emergence of a true eukaryotic cell, the lineage rapidly diversified into multiple kingdoms or supergroups, represented for example by plants, animals, fungi, amoeba and many protist lineages. The ~1.5 billion year period since this radiation is vast, and while core metabolic and gene expression pathways are frequently well conserved, many cellular features have experienced extensive specialisations, in part as a response to adaptive evolutionary forces.

One aspect of this diversity that has received considerable attention is the endomembrane system, on account of this feature representing one of the more unique, yet flexible, aspects of eukaryotic cells. The primitive endomembrane system in the earliest eukaryotes gave rise to all of the endogenously derived internal compartments present in modern lineages (Schlacht et al., 2014). As a consequence of this extensive evolutionary history, compartments have experienced significant divergence, resulting in vastly different morphologies and functions reflected in diversification of the Golgi complex, lysosomes and endosomes. Further, organelles unique to specific lineages such as the rhoptries of Apicomplexa and glycosomes of kinetoplastids are now known and are derived from endolysosomal ancestors and peroxisomes respectively (Szöör et al., 2014; Tomavo, 2014). Residing within this compartmental variation are also apparently ancient and well conserved molecular aspects of trafficking systems, for example the exocytic apparatus comprising the exocyst and endocytic pathways mediated by clathrin (Field et al., 2007; Koumandou et al., 2007)

One of the more diverse eukaryotic phyla are the Kinetoplastida, a group of flagellate protists, many of which are parasites. The African trypanosome *Trypanosoma brucei* lives and multiplies in the bloodstream and lymphatic system of its mammalian host, yet persists in the face of challenge from both the innate and adaptive immune systems (Brun et al., 2010). Immune evasion relies on multiple adaptations of the parasite cell surface composition and

membrane trafficking dynamics (Manna et al., 2014). The canonical defence of *T. brucei* is the expression of a dense coat of a single variant surface glycoprotein (VSG) which, through antigenic variation (the switching of expression between VSG genes), periodically creates an antigenically distinct cell surface (Schwede et al., 2015). As an additional defense, the VSG coat is constantly endocytosed and recycled, removing surface-bound immune effectors (Engstler et al., 2007; Field and Carrington, 2009). Perturbations of endocytic recycling severely limit parasite survival and infectivity *in vitro* and *in vivo* (Allen et al., 2003; Engstler et al., 2007; Natesan et al., 2011).

Common to all trypanosomatids is the polarisation of endocytosis and exocytosis to the flagellar pocket, a specialised plasma membrane invagination surrounding the base of the flagellum (Field and Carrington, 2009). This organelle is the gateway to and from the cell surface and essentially controls the host-parasite interface, but significantly also contributes towards interactions with multiple therapeutics (Alford et al., 2013; Zoltner et al., 2015). Less elaborate, but similar structures, are present at the base of the cilium in other lineages, e.g. the proposed ciliary pocket in mammals and a ciliary partitioning system in *Tetrahymena* (Benmerah, 2013; Ounjai et al., 2013). It is unknown if all of these flagellum-associated structures share common components and functionalities, and the potential that the flagellar pocket contains parasite-specific features limits the utility of comparative approaches that facilitated characterising much of the trypanosome endomembrane system (Field and Carrington, 2009). Recently the surface-exposed proteome of the trypanosome cell surface, the flagellar pocket and endomembrane system have been described (Gadelha et al., 2015; Shimogawa et al., 2015). While the precise definition and discrimination between these interconnected membrane systems is unclear, what has emerged is a remarkable level of novelty within the trypanosome surface membrane and endomembrane proteomes, with very few surface proteins demonstrating significant conservation across eukaryotes (Gadelha et al., 2015; Jackson et al., 2015). This extreme divergence potentially implies novel mechanisms of protein trafficking.

To directly explore the diversity of proteins associated with flagellar pocket function, we previously described affinity isolation of clathrin heavy chain (TbCHC)-complexes using polyclonal antibodies, which identified several

trypanosomatid-specific proteins, together mediating clathrin budding from the flagellar pocket membrane (Adung'a et al., 2013). However, due to the weak and transient nature of many clathrin-partner interactions, the analysis was most probably under-sampled, and many interacting proteins eluded isolation. Here we applied a cryomilling method that better preserves protein-protein interactions, whilst physically disrupting the robust sub-pellicular microtubule corset of the trypanosome cell (Obado et al., 2016). We report an endocytic protein cohort encompassing the majority of the expected early endocytic proteins, together with several novel kinetoplastid-specific gene products. Two of this latter group, TbCAP80 and TbCAP141, control not only clathrin-mediated endocytosis (CME), but also the architecture and organisation of the broader endomembrane system. We suggest that these lineage-specific proteins are central to the coordination of membrane transport at the flagellar pocket, part of a growing list of proteins that mediate the control of function in this crucial organelle.

Methods and materials

Cell culture and transfection: Procyclic form (PCF) Lister 427 *Trypanosoma brucei brucei*, bloodstream form (BSF) Lister 427 and single marker bloodstream (SMB) form parasites were cultured as previously described (Brun and Schönenberger, 1979; Hirumi and Hirumi, 1989; Wirtz et al., 1999). For generation and maintenance of lines harbouring selectable markers, antibiotics were used at the following concentrations: G418 1 µg/ml, puromycin 0.2 µg/ml, hygromycin B 5 µg/ml. COS-7 cells were cultured in DMEM with 10% fetal bovine serum, 100 U/ml penicillin, 100 U/ml streptomycin and 2 mM L-glutamine. For transfection of bloodstream stage parasite 3-4 x 10⁷ cells were transfected with 10 µg of DNA using an AMAXA nucleofection system and the human T-cell nucleofection kit (Lonza). Monoclonal populations were obtained by limiting dilution. Procyclic stage parasite were transfected by electroporation with a Bio-Rad Gene Pulser (1.5 kV, 25 µF). COS-7 cells were transfected with Fugene HD according to manufacturer's instructions.

RNAi and endogenous-locus tagging: Procedures were carried out as described previously (Alibu et al., 2005; Manna et al., 2015; Oberholzer et al.,

2006). Endogenous locus tagging cassettes were generated by PCR using the pMOTag system and gene specific primer pairs as follows; TbCHC F: CCACCCGCAACCGGGCTACGGTGGTGTGCCCGGTCAGGGATATGCTGG AGGGATGGGAAACCCTAACATGATGCCATACGGTACCGGGCCCCCCT CGAG, TbCHC R: ATTTCTTCCCCTCCACCTACTCACCTTTTCTCCTCCCATCTCCCTTCC CTGTGTTTCTTTTGTCTTTTGGGCTGGCGGCCGCTCTAGAACTAGTGGA T, TbCAP141 F: GAGTTCGCTCATCCTCCTGGTCGTGGTGTATGTGCTGAATGAGGGTCGT TATACACCTT TTGCCGCCCGTTTTCCAGTAGGTACCGGGCCCCCCTCGAG, TbCAP141 R: CACCAAACCAGCCTTGATATTAATTCCTTCACTTCTCTATACGGGTTCTT CATATCACTTCCACGAATGCAAGTGGCGGCCGCTCTAGAACTAGTGAT, TbCAP80 F: CTTTGGGTCCGATCACCTCAGTGTGAGCAAAGATAAGAGGGAGAGTGGG AATCACACGC TTACTTTCAACTTTGGTAGTGGTACCGGGCCCCCCTCGAG, TbCAP80 R: AGCACATTGTTGCAGGCGTTAGAACCCTTTTATCTCTTTCCCTTGTGT GTTTTCACTATTCTTGAAGATACCTGGCGGCCGCTCTAGAACTAGTGAT , TbSMAP F: TACTCCGAGTAATCAAGGTCCTCCGCACGTATATAGTGCTT GGGCCCATCCGGTTCCTCAAATGCTTTTCTCCTCAGGGTACCGGGCC CCCCCTCGAG, TbSMAP R: ATTTCAAGTAACCTTTTTTTTGTAAATTTTGCAACTTCAACTCTTTCTTCGCTAA TGTAGGGCCTTCACCGCGAGTGTGGCGGCCGCTCTAGAACTAGTGAT, TbGLP1 F: TACATCTCCAACATCCCAACCGCCGGTGCGGCGGTGAACTCGGCCGGT ACGAAAGGCTCCGTCATTGAGGTGGAGGATGGTACCGGGCCCCCCTC GAG, TbGLP1 R: TTTTCTCGTTCCGGAAGGGTCCATGTTCTTCAAACACCCACTGCTACCT GGTACTCTCGTCAACCGTTCAATGGCGGCCGCTCTAGAACTAGTGAT, TbAAK F: GGAAACGACCTATTCCAAGGCCACAGCAGCAACAGCCACAGCAACCAG AGAAGGACCCCTTCGCCAGTCTCTTCAAGGGTACCGGGCCCCCCTCG

AG, TbAAK R:
GACCCCCCTTTTTTTTTGGGGGGGGGGTGGCGTATGATGCTGTTCTGT
GCCAGTATTGCAGTCATGTAAAACATGGCGGCCGCTCTAGAACTAGTGG
AT. Correct integration of the tagging cassette was verified by western blot.
Gene specific RNAi fragments were selected with RNAit (Redmond et al., 2014)
and amplified from genomic DNA using the following primer pairs; TbAAK F:
TTCTGCTTCTCGCAGACTGA, TbAAK R: AAGAGGTCATCCGTTGTTGG,
TbCAP141 F: GCAGTTGGAGGAGCTACTGG TbCAP141 R:
TTTCTCTTTTCGAAGTGCGGT, TbCAP80 F: CATGCGGAAAGAAAACCAAT,
TbCAP80 R: GCTCTTGTCTGTGGAGCC, TbSMAP F:
CGAGGACGCAGAAAAGAAC, TbSMAP R: TGGGCAAGTACTAACCTCGG.
PCR fragments were cloned into p2T7^{TA}blue and constructs were verified by
sanger sequencing. Multiple clonal RNAi lines were generated for each
construct in the corresponding genomic locus tagged SMB background and
assayed for robust and reproducible protein loss following tetracycline (1 µg/ml)
addition by western blot prior to phenotypic characterization.

Cryoimmunoisolation and proteomics: Approximately 1×10^{10} Mid-log phase BSF or 5×10^{10} PCF parasites were harvested by centrifugation and washed with chilled PBS supplemented with 5.5 mM glucose. Cells were re-suspended in chilled PBS + glucose with protease inhibitors and 10 mM dithiothriitol to a density of 5×10^9 cells/ml and flash frozen in liquid nitrogen to preserve protein-protein interactions as close to native state as possible. Cells were cryomilled into a fine grindate in a planetary ball mill (Retsch). For a detailed protocol refer to (Obado et al., 2015) or the National Center for Dynamic Interactome Research website (www.ncdir.org/public-resources/protocols/) Aliquots of cryo-lysate powder were re-suspended in immuno-precipitation buffer (100 mM sodium phosphate, pH 7.4, 250 mM Na-citrate, 0.5% Tween 20 and protease inhibitors without EDTA (Roche). Subsequent affinity capture was performed as previously described using either magnetic beads coupled to anti-GFP antibody or uncoupled beads to serve as a control for non-specific interaction with the beads (Obado et al., 2016). Affinity purified proteins were eluted with 2%SDS/40 mM Tris pH 8.0, reduced in 50 mM DTT and alkylated with 100 mM iodoacetamide prior to downstream analysis (SDS-PAGE followed by protein identification using mass spectrometry – electrospray ionization (ESI) or

MALDI-TOF). Eluates were fractionated on pre-cast Novex 4-12% Bis Tris gels (Life Technology), stained using colloidal coomassie (GelCode Blue – Thermo) and analyzed by Mass Spectrometry as previously described (DeGrasse et al., 2009).

Immunofluorescence localisation and endocytosis assays: Parasites were cultured, harvested and prepared as described previously (Manna et al., 2015). Anti-HA (3F10, Roche) was used at 1:1000 and polyclonal rabbit anti TbCHC (Morgan et al., 2001) was used at 1:2500. For localisation of GFP tagged TbCAP80 and TbCAP141 truncations in bloodstream form trypanosomes the high-level protein expression required to image native GFP fluorescence led to mislocalisation as compared to endogenously expressed tagged proteins. Therefore, in these experiments expression levels were restricted to near endogenous levels and the fluorescence signal was boosted by staining with polyclonal rabbit anti-GFP (Abcam) followed by Alexfluor-488 conjugated secondary antibody (Invitrogen). Wide-field fluorescence images were acquired using a Nikon Eclipse E600 microscope with a Hamamatsu ORCA CCD camera and MetaMorph software (Universal Imaging, Marlow, UK). For structured illumination microscopy (SIM) images were acquired under a 60X, 1.42NA objective using a DeltaVision OMX V3 system (Applied Precision, Preston, UK) and deconvolved using softWoRx 5.0 software (Applied Precision). Endocytic activity was assessed by fluorescent ligand uptake over a 45 minute period as described previously (Gabernet-Castello et al., 2009; Manna et al., 2015). Quantitation of fluorescent intensities was carried out in Fiji software (National Institutes of Health, USA, <http://fiji.sc/Fiji>). Image panels were prepared in Photoshop (Adobe).

Homology searches, domain predictions and phylogenetic analyses: For the identification of homologous sequences and reconstruction of protein phylogenies, relevant excavate sequences were first identified. To identify related proteins from the kinetoplastids, BLAST searches were carried out at TriTrypDB (<http://tritrypdb.org>) against *Trypanosoma vivax*, *Trypanosoma cruzi*, *Trypanosoma brucei*, *Trypanosoma grayi*, *Trypanosoma congolense*, *Leishmania seymouri*, *Leishmania tarentolae*, *Leishmania mexicana*, *Leishmania major*, *Leishmania donovani*, *Leishmania infantum*, *Crithidia fasciculata* and *Endotrypanum monterogeii*. *Bodo saltans* was searched at

GeneDB (<http://www.genedb.org>), *Naegleria gruberi* was searched at the Joint Genome Institute (<http://genome.jgi.doe.gov>). Any identified excavate sequences were then used in an attempt to identify homologues from other eukaryotic super-groups. Initial protein-BLAST and protein-PSI-BLAST searches were carried out at NCBI (<http://www.ncbi.nlm.nih.gov>) against the non-redundant protein sequence database or species-specific protein databases where appropriate. Where no clear homologues could be identified outside of the excavates, multiple protein alignments of excavate sequences were generated and used as queries against the UniProt KB database via the Jackhammer iterative search algorithm (<http://www.ebi.ac.uk/Tools/hmmer/search/jackhammer>). Finally all identified sequences were verified by reciprocal BLAST and/or jackhammer searches and manual inspection of alignments. To identify conserved domains, retrieved sequences were run through HHPRED (<http://toolkit.tuebingen.mpg.de/hhpred>), hmmscan (<http://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>) and InterProScan (<http://www.ebi.ac.uk/interpro/search/sequence-search>). For phylogenetic reconstruction, sequences were aligned using MergeAlign (Collingridge and Kelly, 2012) and edited for gaps and low conservation. The correct evolutionary model was assessed using ProtTest (http://darwin.uvigo.es/software/protest2_server.html) and phylogenetic trees were constructed using Bayesian (MrBayes, <https://www.phylo.org>) and maximum likelihood (PhyML, <http://www.atgc-montpellier.fr/phyml/>) approaches. MrBayes analyses were run using a mixed model for at least 1×10^6 generations and until convergence, assessed by standard deviation of splits frequencies <0.05 .

TbCAP141 and TbCAP80 truncations: For mammalian expression full length and truncated forms of TbCAP141 and TbCAP80 were amplified from bloodstream form Lister 427 *Trypanosoma brucei brucei* genomic DNA and inserted into the pEGFP-N2 (Clontech) vector for C-terminal GFP tagging and mammalian expression. TbCAP141 full length was amplified using primers Forward: GAGCTCATGCCGCTATACAATATCACTATTCA and reverse: GAATTCTACTGGAAAACGGGCGGC. The N-terminal folded domain (residues 1 - 440) was amplified using the same forward primer and the reverse

primer: GAATTCGGGACCGCCGGAGGCTGCT. The extreme N-terminal putative beta-rich region (residues 1 - 165) was cloned using the same forward primer and the reverse primer: GAATTCGAGGCGAACTGTTTCCACTG. These PCR fragments were inserted between SacI and EcoRI sites of pEGFPN-N2. TbCAP80 fragments were generated using the common forward primer: AGATCTATGCCCCGAGGCGCATATT and reverse primers 1-421: AAGCTTGAACGGGCCCATTTGAGCAC and 1-175: AAGCTTGAATCGGTAGCCATCCGCTT. These fragments were cloned between BglII and HindIII sites. For doxycycline inducible expression in bloodstream form trypanosomes at near endogenous levels the corresponding constructs were cloned into a T7 RNA polymerase-independent inducible expression plasmid as previously described (Sunter et al., 2012).

Thin-section transmission electron microscopy: Samples were prepared as previously described (Gadelha et al., 2009). Briefly, cells were fixed isothermally in culture with 2.5% (v/v) glutaraldehyde at 37°C prior to harvesting by centrifugation. Following fixation, samples were post-fixed in 1% (w/v) osmium tetroxide in PBS for 30 min at room temperature and en-bloc stained with 1% (w/v) aqueous uranyl acetate. Following dehydration through acetone series, samples were embedded in epoxy resin. Ultrathin (70 nm) sections were post-stained with 2% uranyl acetate and lead citrate. Samples were imaged on a Tecnai G2 transmission electron microscope (FEI, USA).

Results

Cryomilling and mass spectrometry identifies novel clathrin-interacting proteins: A previous analysis of clathrin-interacting proteins in trypanosomes yielded a series of lineage-specific proteins we termed clathrin-associated proteins or CAPs. Several CAPs were characterised in some detail and demonstrated to mediate clathrin-budding events at the flagellar pocket membrane (Adung'a et al., 2013). However, many expected factors, including the adaptin AP-1 complex, were not retrieved despite being encoded in the trypanosome genome (Berriman et al., 2005). This suggests that additional proteins are likely present, but were not captured by our earlier procedures and which used quite stringent conditions. Seeking a more complete view of the clathrin interactome, we harvested, flash froze and cryogenically lysed

parasites by cryomilling them under liquid nitrogen to better preserve protein-protein interactions (Obado et al., 2016). Insect or bloodstream form cells expressing clathrin heavy chain with a single allele tagged with a C-terminal GFP (CHC::GFP) were subjected to cryomilling. CHC::GFP was affinity captured under a range of differing salt and detergent conditions using polyclonal anti-GFP llama antibodies coupled to magnetic beads (Fridy et al., 2014).

Coomassie-stained SDS-PAGE of immunoprecipitates using optimised conditions and 50 mg of cell powder ($\sim 5 \times 10^8$ cell equivalents) suggested candidate interacting proteins present at substoichiometric levels, as expected (Figure 1A). Mock Immunoprecipitation performed with beads to which no antibody had been coupled served as the non-specific binding control. Western blotting and MALDI-TOF MS verified the approach, showing enrichment of known or predicted clathrin interactors (Adung'a et al., 2013; Borner et al., 2006; Borner et al., 2012; Borner et al., 2014; Pearse, 1976) (Figure 1A). Larger scale immuno-precipitations using 300 mg of starting material for each condition were analysed by LC-ESI-MS and all identified proteins, from all conditions, are listed in supplementary Table S1. The top ranked proteins, based upon a probabilistic $\log(e)$ score for correct protein identification combined with label-free quantitation (exponentially modified protein abundance index (emPAI) (Ishihama *et al.*, 2005)), from each condition were compared and those proteins enriched in the anti-GFP isolates versus the negative control selected (Table 1).

BLAST searches using the *T. brucei* 927 genome sequence database as queries were used to identify orthologues for conserved proteins, which were verified by reciprocal BLAST. As well as many known or predicted clathrin-interacting proteins, including subunits of the adaptor protein (AP)-1 complex and TbEpsinR (Allen et al., 2007; Gabernet-Castello et al., 2009; Tazeh et al., 2009), multiple partially characterised soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins (Murungi et al., 2014; Venkatesh et al., 2017) and proteins with no apparent homology to known human proteins were identified (Table 1).

Predicted domain architectures revealed several uncharacterised soluble cytosolic proteins, transmembrane-domain containing SNARE proteins

and the previously characterised trypanosomal Golgi-lysosomal glycoprotein 1 (TbGLP-1) (Figure 1B) (Lingnau et al., 1999). In addition, TbEpsinR, Tb927.9.6560, Tb927.8.2820, Tb927.8.3840, Tb927.11.9180 and Tb927.10.13520 contain short, linear clathrin-binding motifs within predicted disordered regions (red dots in Figure 1B), a common feature amongst many known clathrin-interacting proteins (Kirchhausen et al., 2014).

In light of these analyses, five candidate proteins were selected for further investigation. As we were predominantly interested in potential cytoplasmic, soluble, flagellar pocket associated proteins we chose only a single integral membrane protein, TbGLP-1. The remaining four candidate proteins are predicted cytosolic with folded N-terminal domains and disordered C-termini containing predicted clathrin interaction motifs. Tb927.9.6560 has an N-terminal serine/threonine kinase domain and belongs to the same family of kinases as mammalian cyclin G-associated kinase (GAK) and AP-2 associated kinase (AAK). These kinases are involved in clathrin-mediated trafficking processes through phosphorylation of the medium chains of the AP-1 and AP-2 adaptin complexes (Conner and Schmid, 2002; Ricotta et al., 2002; Umeda et al., 2000). This protein is hereafter referred to as *T. brucei* AAK-like (TbAAKL). Tb927.11.9180 (TbSMAP) has an N-terminal ArfGAP domain and belongs to the SMAP family of ArfGAPs implicated in clathrin-dependent vesicle transport (Tanabe et al., 2005; Schlacht et al., 2013). Of particular interest were two uncharacterised proteins with no apparent homology by BLAST to non-trypanosomatid proteins and no identifiable domains, Tb927.8.2820 and Tb927.8.3840 (TbCAP140 and TbCAP80 respectively).

Localisation of putative clathrin-interacting proteins: For immuno-localisation of candidate proteins in bloodstream form trypanosomes, the lifecycle stage with the highest endocytic activity and the clearest dependence upon endocytic function for viability, PCR-generated DNA fragments encoding triple haemagglutinin tags were introduced into the endogenous genes *via* homologous recombination. C-terminal tagging was verified by Western blotting and immunofluorescence. All five candidate proteins co-localised with clathrin, although to varying degrees. This variable colocalisation is to be expected given the diversity and functional segregation of inter-organelle membrane trafficking events utilising the clathrin coat (Fig. 2A). An antibody to clathrin

heavy chain widely labels the posterior region of the trypanosome cell, specifically between the two DAPI foci, the larger one representing the nuclear DNA and a second smaller one representing the mitochondrial DNA, the kinetoplast. This clathrin-rich region of the cell contains the single Golgi apparatus and almost the entirety of the endocytic compartments and, unsurprisingly, TbGLP-1 was observed here as multiple punctae. The distribution of TbSMAP closely resembled that of TbGLP-1 and both proteins colocalised extensively with the clathrin heavy chain. The remaining three proteins were striking in their apparently restricted co-distribution with clathrin. All three, TbAAKL, TbCAP80 and TbCAP141 co-localised with a distinct subset of clathrin punctae located close to the kinetoplast and therefore representing clathrin coated structures at or near to the flagellar pocket membrane. Nascent clathrin coated pits and vesicles are generally between 100 – 200 nm diameter with several pits forming simultaneously at the flagellar pocket. These clathrin structures are not readily resolved by conventional light microscopy, making it challenging to discern true associations with clathrin coated membrane structures versus close apposition. We therefore employed structured illumination microscopy (SIM), which demonstrated a high level of co-localisation between all three candidate proteins and clathrin-positive structures (Fig. 2B, supplemental movies M1-M3). This is in contrast to a known exocytic protein TbSec15, shown by SIM to form punctae interspersed with clathrin-positive structures at the flagellar pocket, but with essentially no co-localisation (Boehm et al., 2017).

Sequence and phylogenetic analysis of T. brucei clathrin-associated proteins:

The evolutionary distribution of candidate genes was investigated to assess their potential relevance to the broader trypanosomatid lineage. The phylogenetic distribution of the SMAP family of ARFGAPs has recently been characterised in detail and demonstrated pan-eukaryotic (Schlacht et al., 2013). Hence, we focused on the remaining newly identified clathrin interactors, TbAAKL, TbCAP141 and TbCAP80. Orthologues of TbAAKL are found across the trypanosomatids and also in the sister group of free living *Bodo saltans*, suggesting that the protein is conserved and an ancient aspect of the kinetoplastid membrane trafficking system (Fig. 3 A). However, in trypanosomes that lack the adaptin complex 2 (AP-2) and also the closely

related AP-2-retaining *T. grayi*, TbAAKL mutations within the conserved D-(X)⁴-N motif of the VIb subdomain are predicted to render the kinase inactive (Hanks and Hunter, 1995) (Fig. 3 B). This suggests a possible relaxation of the requirement for AP-2 directed kinase activity and, perhaps most significantly, predating the loss of the adaptor complex, and which may suggest this as an enabling mutation for subsequent loss of AP-2.

As TbCAP80 and TbCAP141 were completely uncharacterised, we first examined their predicted secondary structures. Unexpectedly, these were strikingly similar despite no apparent sequence similarity, consisting of an extreme N-terminal, predominantly β -region, followed by a relatively large predominantly α -helical region containing putative coiled-coils (Fig. 3C). Downstream of the folded N-terminal domain, both TbCAP80 and TbCAP141 are predicted to contain a large disordered region. Domain prediction tools (HHpred, HMMER, InterproScan) returned no high-confidence predictions, although there is very weak support for the extreme N-terminal β -region representing a degenerate C2 domain. Furthermore, clathrin-binding motifs are located within the disordered regions of both TbCAP80 and TbCAP141. Both are found as single copy genes across the trypanosomatids and also in the sister group Bodonidae, but neither protein has identifiable orthologues outside of the Discicristata (Fig. 3D). Significantly, this overall predicted architecture for TbCAP80 and TbCAP141 is reminiscent of several characterised monomeric clathrin adaptors in which the N-terminal region is a membrane-binding domain while the C-terminus appears flexible and is predicted disordered; examples include the ENTH/ANTH domain epsin/epsinR/CALM family proteins, suggesting a general architectural conservation consistent with clathrin-binding (Kirchhausen et al., 2014).

For functional assessment of CAP141 and CAP80 domains, full length and truncated forms were cloned into vectors for expression in bloodstream stage trypanosomes and mammalian cells (Fig. 4A) (Sunter et al., 2012). In bloodstream form trypanosomes CAP141 and CAP80 full length GFP fusions localised to punctae at, or near, the flagellar pocket as was seen for the endogenously tagged proteins. However, this flagellar pocket localisation was apparently readily saturable, likely reflecting a limited number of available binding sites, with higher expression levels leading to mis-localisation when

compared to endogenously expressed fusion proteins (Fig. 4B and supplemental Fig. S1). Removal of the predicted disordered domains, containing the putative clathrin binding motifs, from either protein had no apparent effect upon the flagellar pocket localization, whereas expression of only the extreme N-terminal β -rich region as a GFP fusion led to a loss of flagellar pocket targeting (Fig. 4B).

In a further attempt to determine if the predicted architectures of TbCAP80 and TbCAP141 functionally correspond to an N-terminal membrane-addressing domain and potential recognition of specific lipid species at the membrane, we expressed TbCAP80 and TbCAP141 as GFP-fusion proteins in COS7 cells (Fig. 4A). This approach was chosen as expression in the mammalian cell background presumably isolates each protein from trypanosome interaction partners that may influence localization as well as represents a better characterized membrane system, where the localisations of specific phosphoinositide-species are well documented. Full length TbCAP141 had a largely perinuclear and reticular distribution (Fig. 4C) *albeit* with some clear plasma membrane association. The internal localisation for Tb141 may be a result of overexpression as seen in the homologous system. The trypanosome expression studies suggest that the large disordered region of this protein is likely dispensable for membrane localisation. Therefore a truncated version of TbCAP141, encoding only the N-terminal predicted folded domain (residues 1 - 440) fused to GFP, was expressed. This construct showed a much more diffuse cytosolic signal whilst also retaining plasma membrane association (Fig. 4C). Finally, we expressed the extreme N-terminus of TbCAP141, corresponding to the potential C2-like fold (residues 1 – 165), as a GFP fusion. This construct failed to localise to the flagellar pocket in bloodstream form trypanosomes but showed the most pronounced plasma membrane localization, suggesting that the N-terminal β -rich domain contains a membrane localization signal (Fig. 4C). We note however that in the absence of a full understanding of the lipid environment at the trypanosome flagellar pocket these data are merely supportive of this conclusion.

In contrast to TbCAP141, TbCAP80 constructs encoding both full-length protein and the extreme N-terminal β -rich domain fused to GFP were largely cytosolic and lacked clear membrane association (Fig. 4C). This suggests that

the targeting information required for localisation of TbCAP141 to the plasma membrane, or the flagellar pocket membrane in the native trypanosome cell, is conserved across eukaryotes, whereas TbCAP80 likely relies upon a trypanosome specific factor as an accessory determinant.

TbSMAP, TbCAP80 and TbCAP141 are essential for endocytosis: Identification via CHC-GFP immuno-precipitation and co-localisation with clathrin structures suggest endocytic functions for TbAAKL, TbSMAP, TbCAP80 and TbCAP141. Many characterised endocytic proteins in *T. brucei* are essential for viability in bloodstream form trypanosomes, as is endocytosis itself, suggesting that if TbAAKL, TbSMAP, TbCAP80 and TbCAP141 are indeed major players in endocytic trafficking, they also should have a significant impact on proliferation and viability.

To address these questions, we generated inducible RNAi cell lines in endogenous-locus tagged bloodstream form backgrounds for each candidate protein. As for the localisation studies we have focused on bloodstream form parasites as this life cycle stage has the highest endocytic activity as well as the clearest link between endocytic activity and viability. Following RNAi induction, protein depletion was assayed by Western blotting and cell proliferation was recorded (Figure 5A). TbSMAP depletion led to a proliferative inhibition after approximately 24 hours. Depletion of either TbCAP80 or TbCAP141 produced strong proliferative inhibition from as little as 12 hours post induction, which was consistent with previous whole genome studies (Alsford et al., 2011). By contrast, despite strong protein depletion, silencing of TbAAKL had no obvious effect upon parasite proliferation. Further, the proliferative defects following TbSMAP, TbCAP80 or TbCAP141 depletion were accompanied by morphological abnormalities (Figure 5B). In particular cells became round and this was frequently accompanied by flagellar pocket swelling, seen as a large phase-light vacuolar structure resulting from failure of clathrin-mediated endocytosis at the flagellar pocket, but also potentially arising from an impact on many trafficking events (Adung'a et al., 2013; Allen et al., 2003; Spitznagel et al., 2010). These morphological effects were absent in TbAAKL depleted cells, but more severe and faster in onset following TbCAP80 and TbCAP141 depletion compared to TbSMAP.

As the enlarged flagellar pocket phenotype of TbSMAP, TbCAP80 or TbCAP141 depleted cells suggested endocytic inhibition, we assayed this directly by following the uptake of fluorescent-labelled probes. Transferrin is selectively acquired *via* the transferrin receptor (Steverding, 2000) whereas the lectin concanavalin A is a non-selective binder of mannose-containing cell surface glycoproteins, predominantly VSG in the trypanosome (Allen et al., 2003). When compared to non-induced control cells, the uptake of both ligands over a 45 minute period was strongly inhibited, following depletion of TbSMAP, TbCAP80 and TbCAP141 (Fig. 5C-D). Consistent with this absence of a proliferative or morphological phenotype, TbAAKL depletion had no apparent impact upon endocytic cargo uptake, indicating that TbAAKL is not required for maintaining endocytic transport.

TbCAP141 and TbCAP80 are required for multiple aspects of flagellar pocket function: The effects of depleting several endocytic proteins are well characterised at the ultra-structural level, and typically lead to enlargement of the flagellar pocket membrane area and luminal volume, together with additional changes associated with delivery or removal of membrane carriers to or from the cell surface (Adung'a et al., 2013; Allen et al., 2003; García-Salcedo et al., 2004; Manna et al., 2015).

We therefore prepared ultrathin sections of bloodstream form cells depleted of TbSMAP, TbCAP80 and TbCAP141 by fast isothermal fixation for transmission electron microscopy. Control cells show the polarised distribution of the *T. brucei* endocytic system, with the posterior flagellar pocket often associated with clathrin-coated pits and vesicles (Figure 6A). Phase-light vacuoles, observed by conventional microscopy in TbSMAP, TbCAP80 and TbCAP141 depleted cells (Figure 5), were confirmed to represent swelling of the flagellar pocket (Figure 6A). Whilst TbSMAP depleted cells displayed normal endosome and Golgi complex profiles with clear tubular sorting endosomes present (Fig 6 B and C), extensive vacuolisation of the cytoplasm was seen in cells depleted of either TbCAP80 or TbCAP141 (Fig. 6B). These swollen vacuolar structures were prominent in the vicinity of the flagellar pocket and also associated with the *trans* face of the Golgi stack (Fig. 6C) indicating more widespread endomembrane system defects. Finally, we examined the morphology of clathrin-coated structures associated with the flagellar pocket.

Control cells displayed many clear clathrin structures, ranging from early, shallow pits to deeply invaginated and pre-scission necked pits (Fig 6D). No such structures could be identified in TbSMAP depleted cells, whilst TbCAP80 and TbCAP141 depleted cells formed extended shallow pits. This suggests that the endocytic defects in these cells arise from a failure of curvature generation and coated pit progression. Depletion of the characterised clathrin adaptor proteins TbCALM and TbEpsinR produces a similar phenotype (Manna et al., 2015), supporting common and non-redundant functions for these proteins. Although of uncertain significance, we also note that in TbCAP80 and TbCAP141 depleted cells occasionally (2 of 28 thin sections, 1 of 20 thin sections, respectively), enlarged flagellar pockets were associated with subpellicular microtubules, a non-canonical feature of this specialized membrane domain, likely due to loss of differentiation from the plasma membrane (Fig S2). Furthermore, depletion of either TbCAP80 or TbCAP141 also led to the appearance of cytoplasmic axonemes, often associated with a paraflagellar rod (3 of 28, 5 of 20 thin sections, respectively) (Fig S2). Although this observation could suggest a loss of cell polarity, and has been observed in knockdowns of functionally related proteins (Price *et al.*, 2007), this phenotype has also been observed as a consequence of ablation of proteins with functions apparently unrelated to those studied here (Morgan et al, 2005, Baines and Gull, 2008; Morriswood and Schmidt, 2015; Sheader et al., 2005). In summary, the most representative ultrastructural defects observed suggest crucial roles for TbCAP80 and TbCAP141 in both endocytosis and Golgi complex and flagellar pocket function and are notable by their extensive scope, impacting a considerable proportion of the endomembrane system.

Discussion

Trypanosomes likely diverged over one billion years ago and many examples of extreme specialisations within membrane trafficking, gene expression and nuclear organisation in these organisms are known (e.g. Daniels et al., 2010; DuBois et al., 2012; Field and Carrington, 2009; Obado et al., 2016B). Comparative genomics demonstrates that trypanosome endocytosis is divergent (Adung'a et al., 2013; Field et al., 2007; Manna et al., 2013; Manna et al., 2015), based both on the absence of orthologues of multiple

endocytic genes of animals and fungi, and recent identification of several lineage-specific proteins associated with clathrin. We have further investigated the clathrin interactome as our earlier study failed to identify many expected and conserved gene products. By cryomilling and improved affinity purification we indeed identified many of these anticipated interaction partners (Figure 7) and more lineage-specific proteins, expanding the evidence that novel mechanisms are operating.

Specifically, we recovered all subunits of AP-1, many SNAREs expected to be associated with the endosomal/post-Golgi trafficking systems, trypanosome orthologues of AAK and SMAP, as well as Hsc70 and TbEpsinR, these latter two common with our earlier work (Adung'a et al., 2013). We also recovered Vps45, the endosomal Sec1/SM protein which in higher eukaryotes interacts with multiple SNAREs and may well have acted as a bridge between clathrin and the SNARE proteins recovered. Vps45 also associates with Rabenosyn5 in mammalian cells, providing an additional connection to early endosomes. The SNAREs in particular are likely associated with endosomal and recycling trafficking based on recent phylogenetics and immunoisolation investigations (Venkatesh et al., 2017) We also recovered TbGLP-1, an abundant trypanosome-specific transmembrane domain protein associated with the Golgi complex of trypanosomes (Lingnau et al., 1999). The identification of several transmembrane domain proteins, together with AP-1 and Vps45 gives us confidence that we have now sampled to a considerable depth the trypanosome endosomal machinery, as does the observation that a very similar cohort of CME proteins are present in *T. cruzi* (Kalb et al., 2016). It is of interest that we detect AP-1 and not any other adaptin complex, and likely that AP-1 is involved in pathways delivering multiple proteins to the lysosome, but not directly in endocytosis from the plasma membrane (Allen et al., 2007; Morgan et al., 2001; Zoltner et al., 2015). The trypanosome orthologue of Atg27 is also of interest; this protein lacks a mammalian homologue but has been characterised in yeast where it cycles between the Golgi and peripheral membrane structures, and controls trafficking of another autophagy protein Atg9 (Backues et al., 2015; Yen et al., 2007).

The current study and recent publications from our laboratory and others allows us to assemble a model for a protein interaction network underlying

flagellar pocket endocytic function (Figure 7) (Adung'a et al., 2013; Demmel et al., 2014; Gabernet-Castello et al., 2009; Manna et al., 2015). We suggest that the local generation of a phosphoinositide signal, likely PI(4,5)P₂, is an essential step in clathrin-coated pit formation, and this signal is sensed by the phosphoinositide-binding clathrin adaptors TbEpsinR and TbCALM and possibly TbCAP80 and TbCAP141 (Demmel et al., 2014; Manna and Field, 2016; Manna et al., 2015). *T. brucei* homologues of Hsc70 and GAK/AAK family kinase are also associated with clathrin structures but their roles remain unclear due to apparently minor phenotypes following RNAi depletion.

TbCAP80 and TbCAP141 have essential roles in CME; that these apparently abundant proteins were not identified in our earlier study highlights the improved capacity for identifying likely weak and transient interactions inherent to cryogenic cell disruption based approaches. Depletion of either TbCAP80 or TbCAP141 gives strikingly similar phenotypes, suggesting related yet non-redundant functions. This is further supported by similarity in predicted secondary structure. Whilst unrelated at the sequence level to canonical monomeric clathrin adaptors, such as epsin or CALM/AP180, TbCAP80 and TbCAP141 do share a striking similarity of domain organisation, with folded N-terminal domains and disordered C-terminal regions, harbouring predicted clathrin binding motifs. This is suggestive of a degree of mechanistic similarity and we have shown that the N-terminal folded domain of TbCAP141 carries a membrane targeting signal which may conceivably be functionally analogous to the ENTH and ANTH domains of epsin and CALM. Further support for this proposition is found in the similarity of specific defects seen in clathrin coated pit morphology following depletion of TbCAP80, TbCAP141, TbCALM or TbEpsinR. It appears that ablation of any of these proteins leads to endocytic inhibition through a block of endocytic pit maturation, causing an accumulation of broad, flat clathrin-coated structures at the flagellar pocket. Interestingly both whole genome RNAi and proteomics studies indicate that TbCAP80 and TbCAP141 are essential in the procyclic stage as well, but are expressed at somewhat lower levels (~0.36-fold and ~0.25-fold respectively, relative to bloodstream stage), suggestive of roles in endocytic transport in the insect form as well (Alsford et al., 2011, Urbaniak et al., 2012).

That depletion of any of these proteins, suggested to independently drive clathrin recruitment, causes malformed pits, rather than a failure of clathrin assembly *per se*, may be explained by an emerging model of membrane vesicle formation. In addition to proposed membrane curvature driving domains, such as amphipathic helices, the steric interactions between coat proteins themselves are suggested to generate positive curvature through molecular crowding (Busch et al., 2015; Copic et al., 2012; Miller et al., 2015; Stachowiak et al., 2012). The concentration of luminal cargoes packaged into forming pits generates an opposing force that must be overcome during coat assembly. *T. brucei*, with its dense VSG coat, would therefore be predicted to be particularly sensitive to perturbations of this curvature-driving force provided by molecular crowding of coat proteins. We suggest therefore that depletion of any one of the major components of the endocytic coat in *T. brucei*, while likely not sufficient to prevent clathrin assembly, leads to a failure of pit closure and subsequent endocytic inhibition.

Heavily studied in yeast and metazoan systems, attention has only relatively recently turned to the degree of lineage-specific adaptation in the molecular mechanisms of CME. In contrast with the adaptin-related TPLATE/TSET complex recently characterised in plants and amoebozoia (Gadeyne et al., 2014; Hirst et al., 2014), TbCAP80 and TbCAP141 appear to be true lineage-specific innovations. This adds to the remarkably large number of lineage-specific clathrin-associated proteins with roles in trypanosome endocytosis (Adung'a et al., 2013) and is further evidence for significant remodelling of endosomal systems to enable specific adaptation, in this instance facilitating parasitism. While less extensive than the evolution of rhoptries and dense granules as unique organelles intimately associated with cell invasion in Apicomplexa (Tomavo, 2014), it is clear that the endosomal system of trypanosomes plays vital roles in the pathogenesis of these organisms.

In summary, this study reveals the molecular pathways subtending an early-branching and highly divergent form of clathrin-mediated endocytosis, arising via the acquisition of a largely novel set of gene products including TbCAP80 and TbCAP141 which are conserved and likely functionally important across the kinetoplastid parasites.

Acknowledgements

This work was supported by the Wellcome Trust, Medical Research Council (grant G0900255 to CG and MCF) and the NIH (P41 GM109824 and R21 AI096069 to MPR). We also thank Divya Venketesh for annotation of trypanosome SNAREs.

Author contributions

MCF, PTM conceived the study; PTM, CB, SOO, CG performed the research and PTM, CB, SOO, CG, MPR and BTC analysed the data. PTM, SOO, CG, MPR and MCF wrote and edited the manuscript.

References

- Adung'a, V. O., Gadelha, C. and Field, M. C.** (2013). Proteomic analysis of clathrin interactions in trypanosomes reveals dynamic evolution of endocytosis. *Traffic* **14**, 440–57.
- Alibu, V. P., Storm, L., Haile, S., Clayton, C. and Horn, D.** (2005). A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **139**, 75–82.
- Allen, C. L., Goulding, D. and Field, M. C.** (2003). Clathrin-mediated endocytosis is essential in *Trypanosoma brucei*. *EMBO J.* **22**, 4991–5002.
- Allen, C. L., Liao, D., Chung, W.-L. and Field, M. C.** (2007). Dileucine signal-dependent and AP-1-independent targeting of a lysosomal glycoprotein in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **156**, 175–90.
- 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**, 915-24.
- Alsford, S., Field, M. C. and Horn, D.** (2013). Receptor-mediated endocytosis for drug delivery in African trypanosomes: fulfilling Paul Ehrlich's vision of chemotherapy. *Trends Parasitol.* **29**, 207–12.
- Backues, S. K., Orban, D. P., Bernard, A., Singh, K., Cao, Y. and Klionsky, D. J.** (2015). Atg23 and Atg27 act at the early stages of Atg9 trafficking in *S. cerevisiae*. *Traffic* **16**, 172–90.
- Baines, A. and Gull, K.** (2008). WCB is a C2 domain protein defining the plasma membrane - sub-pellicular microtubule corset of kinetoplastid parasites. *Protist* **159**, 115–25.
- Benmerah, A.** (2013). The ciliary pocket. *Curr. Opin. Cell Biol.* **25**, 78–84.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renaud, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., et al.** (2005). The Genome of the African Trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422.
- Boehm, C., Obado, S., Manna, P.T., Gadelha, C., O'Reilly, A.J., Munson, M., Chait, B.T., Rout, M.P., and Field, M.C.,** (2016). The trypanosome exocyst: A conserved structure revealing a new role in endocytosis. *PLoS Pathogens* in press

- Borner, G. H. H., Harbour, M., Hester, S., Lilley, K. S. and Robinson, M. S.** (2006). Comparative proteomics of clathrin-coated vesicles. *J. Cell Biol.* **175**, 571–8.
- Borner, G. H. H., Antrobus, R., Hirst, J., Bhumbra, G. S., Kozik, P., Jackson, L. P., Sahlender, D. A. and Robinson, M. S.** (2012). Multivariate proteomic profiling identifies novel accessory proteins of coated vesicles. *J. Cell Biol.* **197**, 141–60.
- Borner, G. H. H., Hein, M. Y., Hirst, J., Edgar, J. R., Mann, M. and Robinson, M. S.** (2014). Fractionation profiling: a fast and versatile approach for mapping vesicle proteomes and protein-protein interactions. *Mol. Biol. Cell* **25**, 3178–94.
- Brun, R. and Schönenberger** (1979). Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. *Acta Trop.* **36**, 289–92.
- Brun, R., Blum, J., Chappuis, F. and Burri, C.** (2010). Human African trypanosomiasis. *Lancet* **375**, 148–59.
- Busch, D. J., Houser, J. R., Hayden, C. C., Sherman, M. B., Lafer, E. M. and Stachowiak, J. C.** (2015). Intrinsically disordered proteins drive membrane curvature. *Nat. Commun.* **6**, 7875.
- Collingridge, P. W. and Kelly, S.** (2012). MergeAlign: improving multiple sequence alignment performance by dynamic reconstruction of consensus multiple sequence alignments. *BMC Bioinformatics* **13**, 117.
- Conner, S. D. and Schmid, S. L.** (2002). Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J. Cell Biol.* **156**, 921–9.
- Copic, A., Latham, C. F., Horlbeck, M. A., D’Arcangelo, J. G. and Miller, E. A.** (2012). ER cargo properties specify a requirement for COPII coat rigidity mediated by Sec13p. *Science* **335**, 1359–62.
- Daniels, J.-P., Gull, K. and Wickstead, B.** (2010). Cell biology of the trypanosome genome. *Microbiol. Mol. Biol. Rev.* **74**, 552–69.
- Demmel, L., Schmidt, K., Lucast, L., Havlicek, K., Zankel, A., Koestler, T., Reithofer, V., de Camilli, P. and Warren, G.** (2014). The endocytic activity of the flagellar pocket in *Trypanosoma brucei* is regulated by an adjacent phosphatidylinositol phosphate kinase. *J. Cell Sci.* **127**, 2351–64.

- DuBois, K. N., Alsford, S., Holden, J. M., Buisson, J., Swiderski, M., Bart, J.-M., Ratushny, A. V., Wan, Y., Bastin, P., Barry, J. D., et al.** (2012). NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol.* **10**, e1001287.
- Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. and Overath, P.** (2007). Hydrodynamic Flow-Mediated Protein Sorting on the Cell Surface of Trypanosomes. *Cell* **131**, 505–515.
- Field, M. C. and Carrington, M.** (2009). The trypanosome flagellar pocket. *Nat Rev Micro* **7**, 775–786.
- Field, M. C., Gabernet-Castello, C. and Dacks, J. B.** (2007). Reconstructing the Evolution of the Endocytic System: Insights from Genomics and Molecular Cell Biology. *Eukaryotic Membranes and Cytoskeleton*. pp. 84–96. Springer New York.
- Fridy, P. C., Li, Y., Keegan, S., Thompson, M. K., Nudelman, I., Scheid, J. F., Oeffinger, M., Nussenzweig, M. C., Fenyö, D., Chait, B. T., et al.** (2014). A robust pipeline for rapid production of versatile nanobody repertoires. *Nat. Methods* **11**, 1253–60.
- Gabernet-Castello, C., Dacks, J. B. and Field, M. C.** (2009). The single ENTH-domain protein of trypanosomes; endocytic functions and evolutionary relationship with epsin. *Traffic* **10**, 894–911.
- Gadelha, C., Rothery, S., Morphew, M., McIntosh, J. R., Severs, N. J. and Gull, K.** (2009). Membrane domains and flagellar pocket boundaries are influenced by the cytoskeleton in African trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17425–30.
- Gadelha, C., Zhang, W., Chamberlain, J. W., Chait, B. T., Wickstead, B. and Field, M. C.** (2015). Architecture of a host-parasite interface: complex targeting mechanisms revealed through proteomics. *Mol. Cell. Proteomics* **14**, 1911–26.
- Gadeyne, A., Sánchez-Rodríguez, C., Vanneste, S., Di Rubbo, S., Zauber, H., Vanneste, K., Van Leene, J., De Winne, N., Eeckhout, D., Persiau, G., et al.** (2014). The TPLATE adaptor complex drives clathrin-mediated endocytosis in plants. *Cell* **156**, 691–704.

- García-Salcedo, J. A., Pérez-Morga, D., Gijón, P., Dilbeck, V., Pays, E. and Nolan, D. P.** (2004). A differential role for actin during the life cycle of *Trypanosoma brucei*. *EMBO J.* **23**, 780–9.
- Guy, L. and Ettema, T. J. G.** (2011). The archaeal “TACK” superphylum and the origin of eukaryotes. *Trends Microbiol.* **19**, 580–7.
- Hanks, S. K. and Hunter, T.** (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–96.
- Hirst, J., Schlacht, A., Norcott, J. P., Traynor, D., Bloomfield, G., Antrobus, R., Kay, R. R., Dacks, J. B. and Robinson, M. S.** (2014). Characterization of TSET, an ancient and widespread membrane trafficking complex. *Elife* **3**, e02866.
- Hirumi, H. and Hirumi, K.** (1989). Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J. Parasitol.* **75**, 985–9.
- Jackson, A. P., Otto, T. D., Aslett, M., Armstrong, S. D., Bringaud, F., Schlacht, A., Hartley, C., Sanders, M., Wastling, J. M., Dacks, J. B., et al.** (2015). Kinetoplastid Phylogenomics Reveals the Evolutionary Innovations Associated with the Origins of Parasitism. *Curr. Biol.* **26**, 161–172.
- Kalb, L. C., Frederico, Y.C.A., Boehm, C., Moreira, C., Soares, M.J., and Field, M.C.**, (2016). Conservation and divergence within the clathrin interactome of *Trypanosoma cruzi*. *Nature Sci. Repts.* **6**, 31212
- Kirchhausen, T., Owen, D. and Harrison, S. C.** (2014). Molecular structure, function, and dynamics of clathrin-mediated membrane traffic. *Cold Spring Harb. Perspect. Biol.* **6**, a016725.
- Koumandou, V. L., Dacks, J. B., Coulson, R. M. R. and Field, M. C.** (2007). Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evol. Biol.* **7**, 29.
- Lingnau, A., Zufferey, R., Lingnau, M. and Russell, D. G.** (1999). Characterization of tGLP-1, a Golgi and lysosome-associated, transmembrane glycoprotein of African trypanosomes. *J. Cell Sci.* **112 Pt 18**, 3061–70.

- Manna, P. T. and Field, M. C.** (2016). Phosphoinositides, kinases and adaptors coordinating endocytosis in *Trypanosoma brucei*. *Commun. Integr. Biol.* **8**, e1082691.
- Manna, P. T., Kelly, S. and Field, M. C.** (2013). Adaptin evolution in kinetoplastids and emergence of the variant surface glycoprotein coat in African trypanosomatids. *Mol. Phylogenet. Evol.* **67**, 123–8.
- Manna, P. T., Boehm, C., Leung, K. F., Natesan, S. K. and Field, M. C.** (2014). Life and times: synthesis, trafficking, and evolution of VSG. *Trends Parasitol.* **30**, 251–8.
- Manna, P. T., Gadelha, C., Puttick, A. E. and Field, M. C.** (2015). ENTH and ANTH domain proteins participate in AP2-independent clathrin-mediated endocytosis. *J. Cell Sci.* **128**, 2130–42.
- Martin, W. F., Garg, S. and Zimorski, V.** (2015). Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140330.
- Miller, S. E., Mathiasen, S., Bright, N. A., Pierre, F., Kelly, B. T., Kladt, N., Schauss, A., Merrifield, C. J., Stamou, D., Höning, S., et al.** (2015). CALM Regulates Clathrin-Coated Vesicle Size and Maturation by Directly Sensing and Driving Membrane Curvature. *Dev. Cell* **33**, 163–75.
- Morgan, G. W., Allen, C. L., Jeffries, T. R., Hollinshead, M. and Field, M. C.** (2001). Developmental and morphological regulation of clathrin-mediated endocytosis in *Trypanosoma brucei*. *J. Cell Sci.* **114**, 2605–15.
- Morgan, G. W., Denny, P. W., Vaughan, S., Goulding, D., Jeffries, T. R., Smith, D. F., Gull, K. and Field, M. C.** (2005). An evolutionarily conserved coiled-coil protein implicated in polycystic kidney disease is involved in basal body duplication and flagellar biogenesis in *Trypanosoma brucei*. *Mol. Cell. Biol.* **25**, 3774–83.
- Morriswood, B. and Schmidt, K.** (2015). A MORN Repeat Protein Facilitates Protein Entry into the Flagellar Pocket of *Trypanosoma brucei*. *Eukaryot. Cell* **14**, 1081–93.
- Murungi, E., Barlow, L. D., Venkatesh, D., Adung'a, V. O., Dacks, J. B., Field, M. C. and Christoffels, A.** (2014). A comparative analysis of trypanosomatid SNARE proteins. *Parasitol. Int.* **63**, 341–8.

- Natesan, S. K. A., Black, A., Matthews, K. R., Mottram, J. C. and Field, M. C.** (2011). Trypanosoma brucei brucei: endocytic recycling is important for mouse infectivity. *Exp. Parasitol.* **127**, 777–83.
- Obado, S. O., Field, M. C., Chait, B. T. and Rout, M. P.** (2016A). High-Efficiency Isolation of Nuclear Envelope Protein Complexes from Trypanosomes. In *Methods in molecular biology (Clifton, N.J.)*, pp. 67–80.
- Obado, S. O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N. E., Chait, B. T., Field, M. C. and Rout, M. P.** (2016B). Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. *PLoS Biol.* **14**, e1002365.
- Oberholzer, M., Morand, S., Kunz, S. and Seebeck, T.** (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of Trypanosoma brucei genes. *Mol. Biochem. Parasitol.* **145**, 117–20.
- Ounjai, P., Kim, K. D., Liu, H., Dong, M., Tauscher, A. N., Witkowska, H. E. and Downing, K. H.** (2013). Architectural insights into a ciliary partition. *Curr. Biol.* **23**, 339–44.
- Pearse, B. M.** (1976). Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 1255–9.
- Price, H. P., Stark, M., Smith, B. and Smith, D. F.** (2007). TbARF1 influences lysosomal function but not endocytosis in procyclic stage Trypanosoma brucei. *Mol. Biochem. Parasitol.* **155**, 123–127.
- Raymann, K., Brochier-Armanet, C. and Gribaldo, S.** (2015). The two-domain tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6670–5.
- Ricotta, D., Conner, S. D., Schmid, S. L., von Figura, K. and Honing, S.** (2002). Phosphorylation of the AP2 mu subunit by AAK1 mediates high affinity binding to membrane protein sorting signals. *J. Cell Biol.* **156**, 791–5.
- Schlacht, A., Mowbrey, K., Elias, M., Kahn, R. A. and Dacks, J. B.** (2013). Ancient complexity, opisthokont plasticity, and discovery of the 11th subfamily of Arf GAP proteins. *Traffic* **14**, 636–49.

- Schlacht, A., Herman, E. K., Klute, M. J., Field, M. C. and Dacks, J. B.** (2014). Missing pieces of an ancient puzzle: evolution of the eukaryotic membrane-trafficking system. *Cold Spring Harb. Perspect. Biol.* **6**, a016048.
- Schwede, A., Macleod, O. J. S., MacGregor, P. and Carrington, M.** (2015). How Does the VSG Coat of Bloodstream Form African Trypanosomes Interact with External Proteins? *PLoS Pathog.* **11**, e1005259.
- Sheader, K., Vaughan, S., Minchin, J., Hughes, K., Gull, K. and Rudenko, G.** (2005). Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8716–21.
- Shimogawa, M. M., Saada, E. A., Vashisht, A. A., Barshop, W. D., Wohlschlegel, J. A. and Hill, K. L.** (2015). Cell surface proteomics provides insight into stage-specific remodeling of the host-parasite interface in *Trypanosoma brucei*. *Mol. Cell. Proteomics* **14**, 1977–88.
- Spitznagel, D., O'Rourke, J. F., Leddy, N., Hanrahan, O. and Nolan, D. P.** (2010). Identification and characterization of an unusual class I myosin involved in vesicle traffic in *Trypanosoma brucei*. *PLoS One* **5**, e12282.
- Stachowiak, J. C., Schmid, E. M., Ryan, C. J., Ann, H. S., Sasaki, D. Y., Sherman, M. B., Geissler, P. L., Fletcher, D. A. and Hayden, C. C.** (2012). Membrane bending by protein-protein crowding. *Nat. Cell Biol.* **14**, 944–9.
- Steverding, D.** (2000). The transferrin receptor of *Trypanosoma brucei*. *Parasitol. Int.* **48**, 191–198.
- Sunter, J., Wickstead, B., Gull, K., Carrington, M., Wirtz, E., Clayton, C., Wirtz, E., Leal, S., Ochatt, C., Cross, G., et al.** (2012). A New Generation of T7 RNA Polymerase-Independent Inducible Expression Plasmids for *Trypanosoma brucei*. *PLoS One* **7**, e35167.
- Szöör, B., Haanstra, J. R., Gualdrón-López, M. and Michels, P. A. M.** (2014). Evolution, dynamics and specialized functions of glycosomes in metabolism and development of trypanosomatids. *Curr. Opin. Microbiol.* **22**, 79–87.

- Tazeh, N. N., Silverman, J. S., Schwartz, K. J., Sevova, E. S., Sutterwala, S. S. and Bangs, J. D.** (2009). Role of AP-1 in developmentally regulated lysosomal trafficking in *Trypanosoma brucei*. *Eukaryot. Cell* **8**, 1352–61.
- Tomavo, S.** (2014). Evolutionary repurposing of endosomal systems for apical organelle biogenesis in *Toxoplasma gondii*. *Int. J. Parasitol.* **44**, 133–8.
- Umeda, A., Meyerholz, A. and Ungewickell, E.** (2000). Identification of the universal cofactor (auxilin 2) in clathrin coat dissociation. *Eur. J. Cell Biol.* **79**, 336–42.
- Urbaniak MD, Guther ML, Ferguson MA.** (2012). Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. *PLoS One.* *7*, e36619.
- Venkatesh, D., Boehm, C., Barlow, L.D., Nankissor, N.N., O'Reilly, A.J., Kelly, S., Dacks, J.B., and Field, M.C.** (2017). Evolution of the endomembrane systems of trypanosomatids: Conservation and specialisation. *J. Cell Sci.* in press
- Wirtz, E., Leal, S., Ochatt, C. and Cross, G. M.** (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101.
- Yen, W.-L., Legakis, J. E., Nair, U. and Klionsky, D. J.** (2007). Atg27 is required for autophagy-dependent cycling of Atg9. *Mol. Biol. Cell* **18**, 581–93.
- Zoltner, M., Leung, K. F., Alsford, S., Horn, D. and Field, M. C.** (2015). Modulation of the Surface Proteome through Multiple Ubiquitylation Pathways in African Trypanosomes. *PLoS Pathog.* **11**, e1005236.

Figures

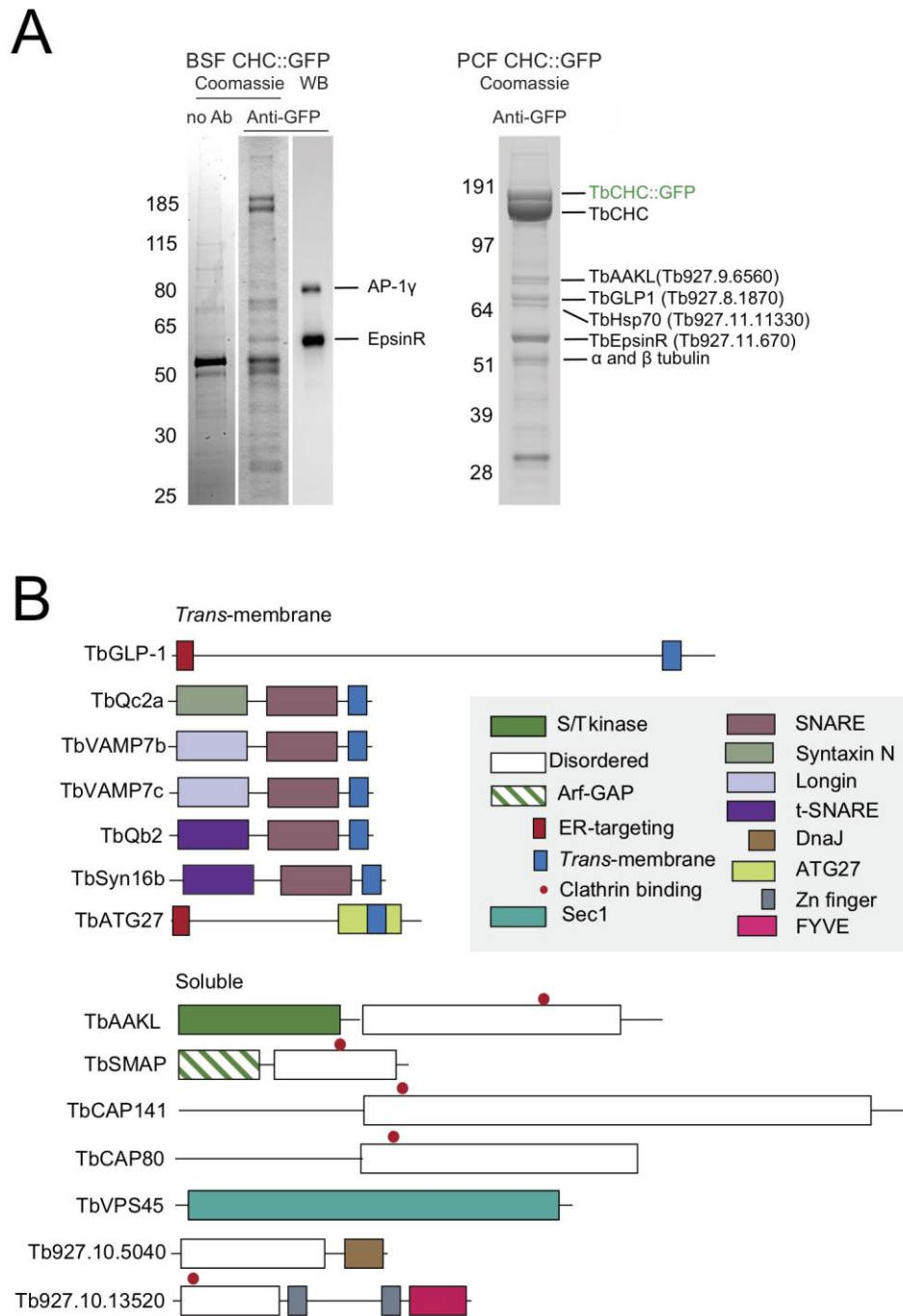


Figure 1. Identification of *T. brucei* clathrin associated proteins. Clathrin and associated proteins were immuno-precipitated from cryomilled powder generated from GFP-tagged clathrin heavy chain (CHC::GFP) bloodstream or procyclic stage cells. (A) Coomassie stained SDS-PAGE shows bands specific to anti-GFP affinity isolates versus controls, confirmed as known CHC

associated proteins by western blot and MALDI-TOF MS. (BSF = bloodstream form parasites, PCF = procyclic form parasites) (B) Domain organisation of selected high-confidence clathrin-associated proteins identified by LC-MS² (see Table 1 and Table S1). Domain predictions were generated via HHPRED, HMMER and InterProScan.

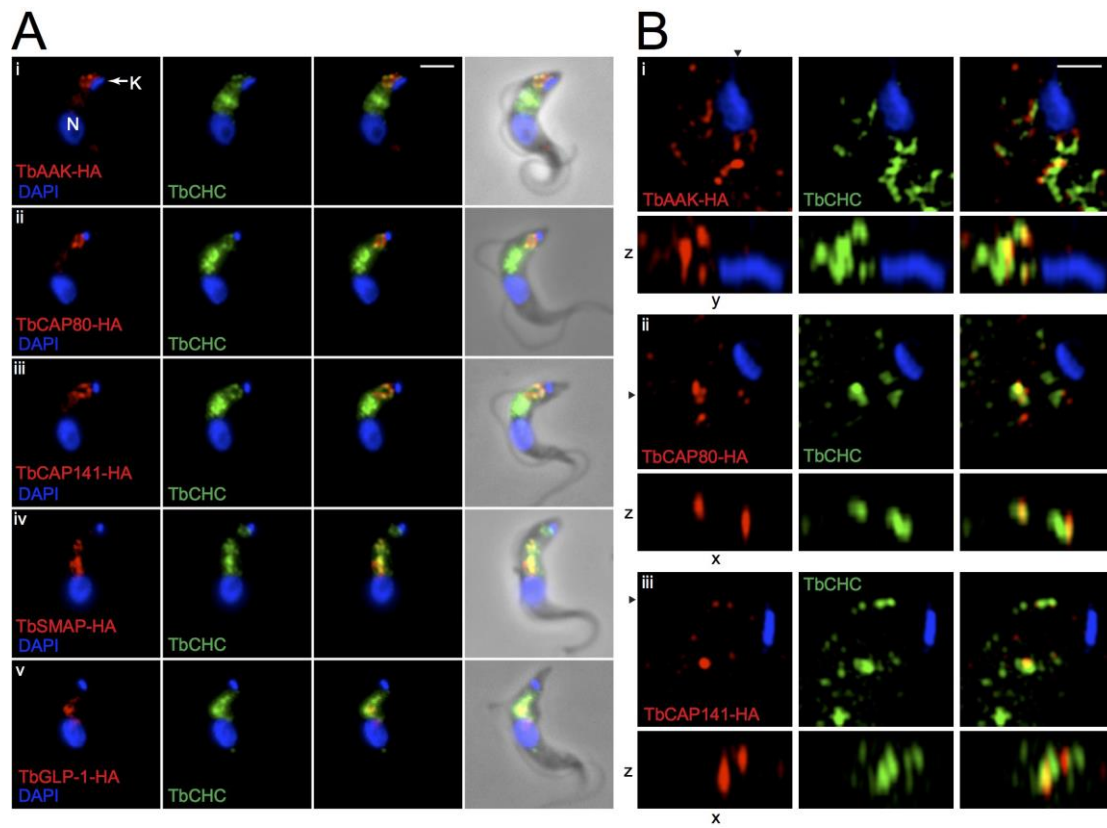


Figure 2. Localisation of selected *T. brucei* clathrin-associated proteins. C-terminal triple HA tags were added to selected clathrin-associated proteins at endogenous genomic loci *via* homologous recombination in bloodstream stage parasites. (A) Immunofluorescence co-localisation showed correlation between clathrin-associated proteins (red, left panels) and clathrin heavy chain (green, central panels). Note that TbAAKL (i), TbCAP80 (ii) and TbCAP141 (iii) showed preferential co-localisation with clathrin punctae in the vicinity of the flagellar pocket (blue DAPI, N nucleus, K kinetoplast, indicated on top panel only). Scale bar, 2 μ m. (B) Super-resolution structured illumination microscopy shows close association of clathrin-associated proteins (red, left panels, anti-HA antibody) with the flagellar pocket and clathrin-positive structures (green, centre panels, anti-TbCHC antibody) close to the kinetoplast. Boxes below the main panels show z-plane resections along the axis denoted by the black arrowhead for each image. Scale bar, 500nm.

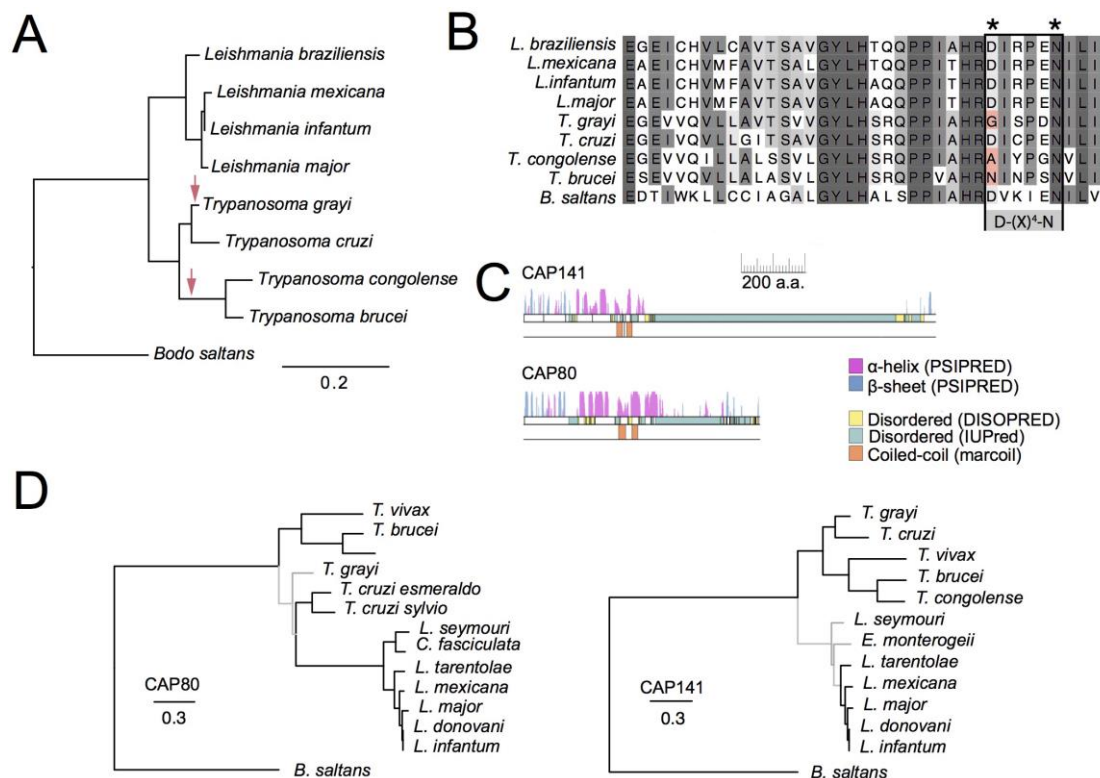


Figure 3. Sequence and phylogenetic analysis of newly identified clathrin-associated proteins. (A) Phylogenetic tree of the TbAAKL family across the kinetoplastids. Red arrows mark potential occurrences of mutations predicted to render the kinase inactive. (B) Sequence alignment showing putative inactivating mutations in kinetoplastid TbAAKL family kinases. (C) Secondary structure predictions for TbCAP141 and TbCAP80. Note similarities with predicted extreme N-terminal β -rich regions followed by α -rich regions containing putative coiled coils and large disordered domains at the C-termini. (D) Phylogenetic analyses of TbCAP141 and TbCAP80 families showing distribution across the kinetoplastids. Trees shown are best scoring ML trees (PhyML). Black branches are supported >80, >80, >0.99 (RaxML, PhyML, MrBayes), grey branches are below this level of support.

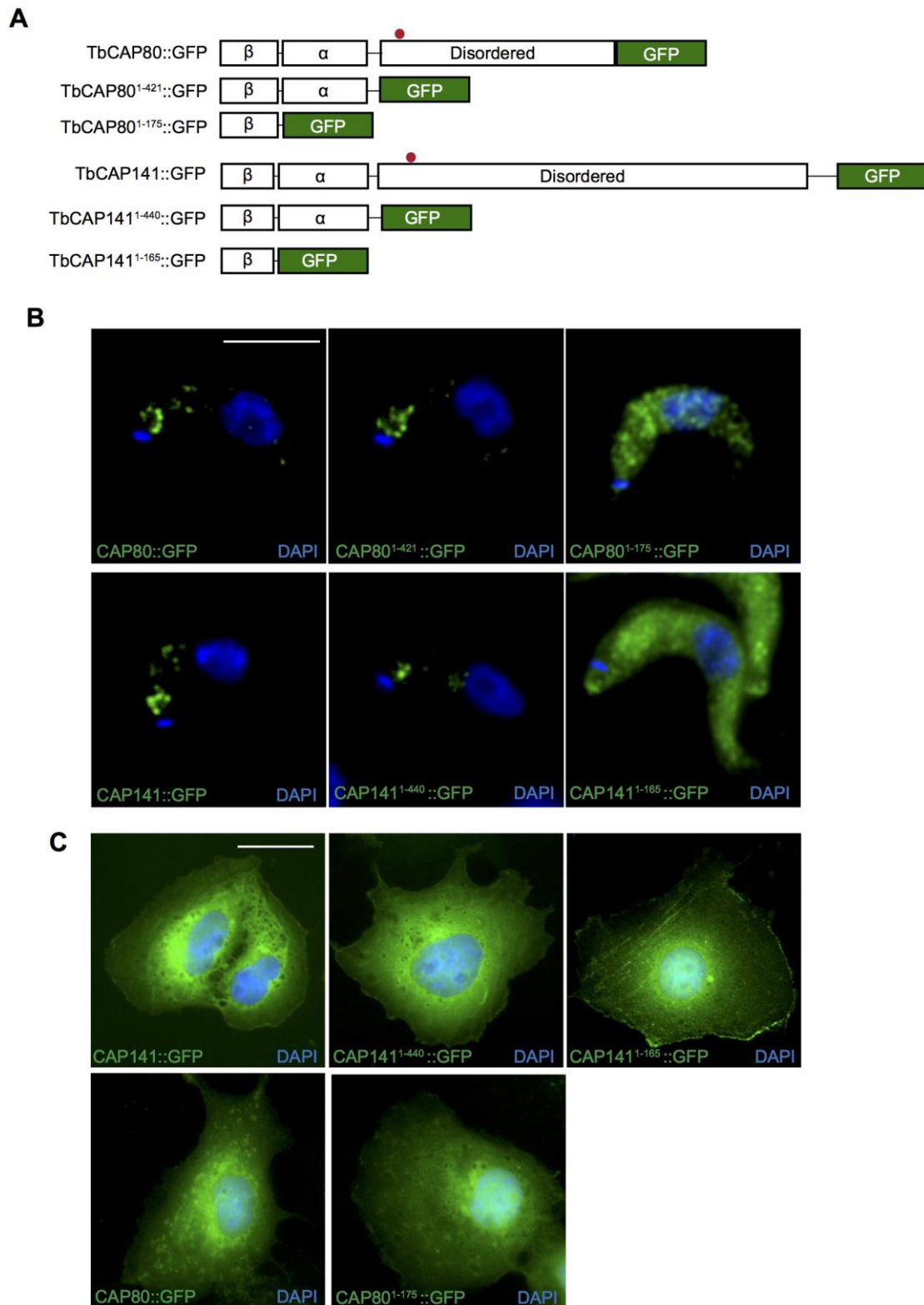


Figure 4. **Heterologous expression of TbCAP-GFP fusions.** (A) C-terminal GFP fusions were constructed of full length TbCAP141 and TbCAP80,

truncated proteins bearing the entire predicted N-terminal globular domain and the predicted β -rich extreme N-terminus alone. (B) Expression of truncation constructs in bloodstream form parasites under the control of tetracycline. Low level expression necessitated amplification of native GFP fluorescence by staining with anti-GFP followed by fluorescent secondary antibodies. Scale bar equals 5 μm . (C) Expression of TbCAP truncation constructs in COS-7 cells. Scale bar equals 20 μm .

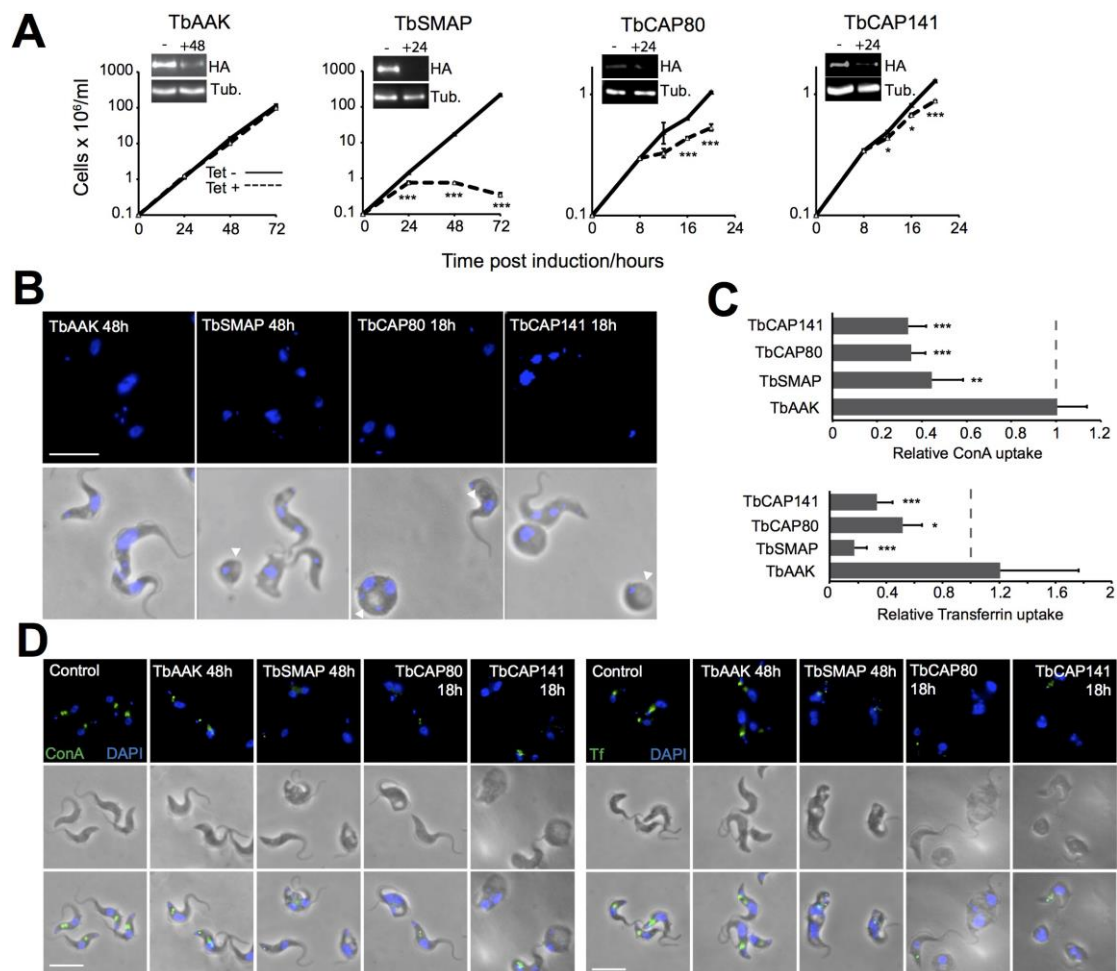


Figure 5. Phenotypic consequences of clathrin-associated protein depletion. Tetracycline inducible RNAi cell lines for selected clathrin-associated proteins were generated in HA-tagged bloodstream form cells. (A) Effects of tetracycline addition on protein levels, assessed by Western blot using an anti-HA antibody, and cell proliferation expressed as mean \pm s.e.m. from three independent repeats (* denotes t-test $p < 0.05$, ** denotes $p < 0.005$, *** denotes $p < 0.001$) (B) Morphology of RNAi induced cells at the indicated time-points (blue; DAPI, white arrowheads denote phase-light vacuolar structures reminiscent of swollen flagellar pockets). Scale bar, 5 μ m. (C, D) Effects of clathrin-associated protein depletion on uptake of FITC-transferrin or FITC-ConA. (C) Quantitation of FITC-transferrin or FITC-ConA uptake in RNAi induced cells versus uninduced controls. Data are mean \pm standard deviation of at least 50 cells per condition from two independent experiments normalized

to non-induced controls (* denotes t-test $p < 0.05$, ** denotes $p < 0.005$, *** denotes $p < 0.001$). (D) Representative images showing FITC-transferrin or FITC-ConA (green) accumulation following 45-minute pulse. Scale bar, 5 μm , blue is DAPI stain.

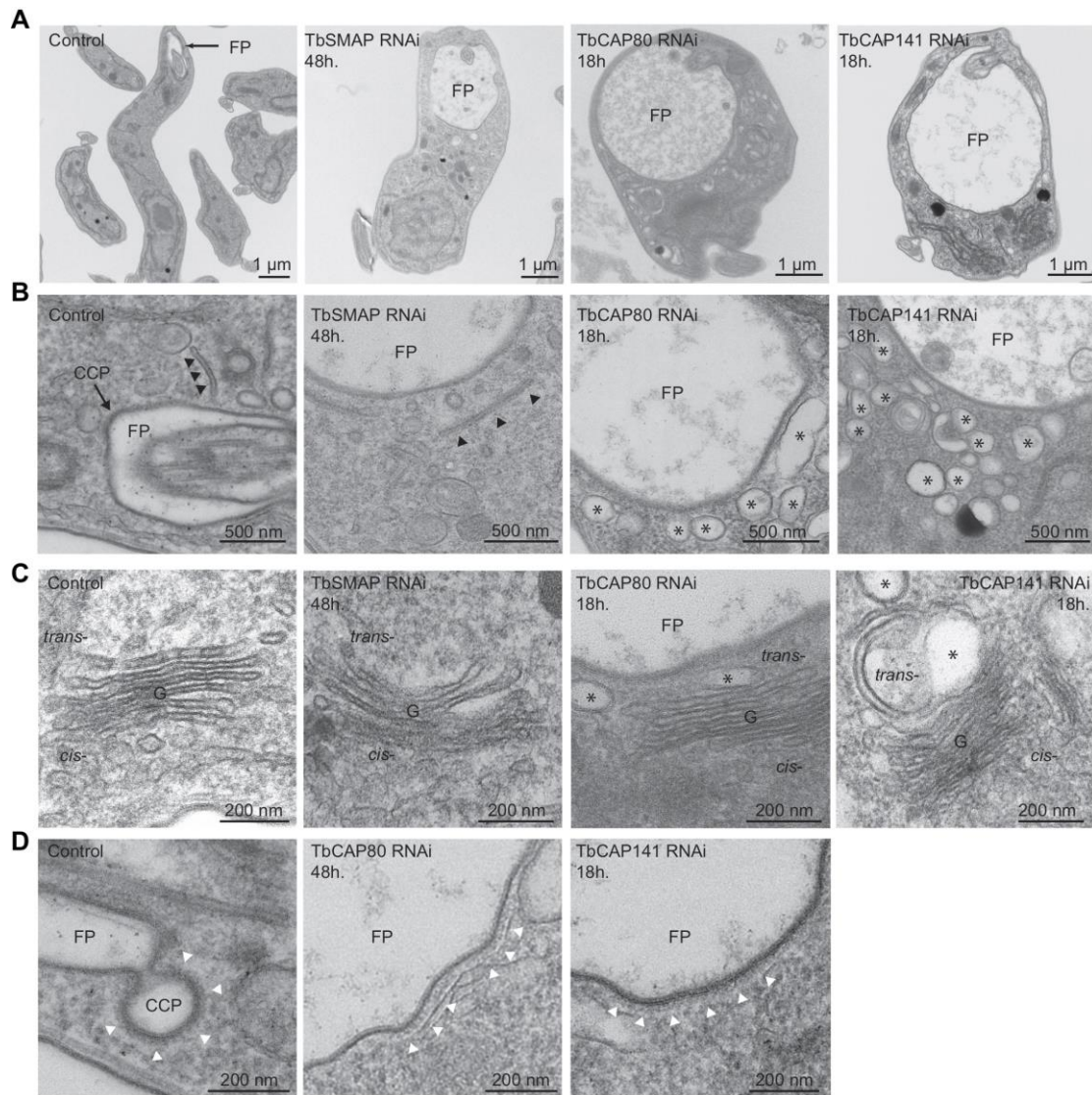


Figure 6. Effects of clathrin-associated protein depletion on endomembrane system ultra-structure. Representative transmission electron micrographs of ultra-thin resin sections of bloodstream form parasites after RNAi-mediated depletion of clathrin-associated proteins. (A) Gross ultrastructural defects, in particular swelling of the flagellar pocket (FP) and cytoplasmic vacuolization. (B) Tubular endosomes (black arrowheads) are apparent in the vicinity of the flagellar pocket of control and TbSMAP depleted cells. A clathrin-coated pit profile is seen at the flagellar pocket membrane of the control section (black arrow). Extensive vacuolization (asterisks) is apparent in both TbCAP80 and TbCAP141 depleted cells. (C) Golgi profiles;

control cell shows numerous ER-Golgi transport intermediates. TbSMAP depletion has little apparent effect upon Golgi ultrastructure, whereas depletion of TbCAP80 or TbCAP141 causes vacuolization and swelling (asterisks) of the *trans*-cisternae. (D) Clathrin coated profiles (white arrowheads) of control and TbCAP80 or TbCAP141 depleted cells. Aberrant, large flat coated profiles are seen in TbCAP80 or TbCAP141 depleted cells.

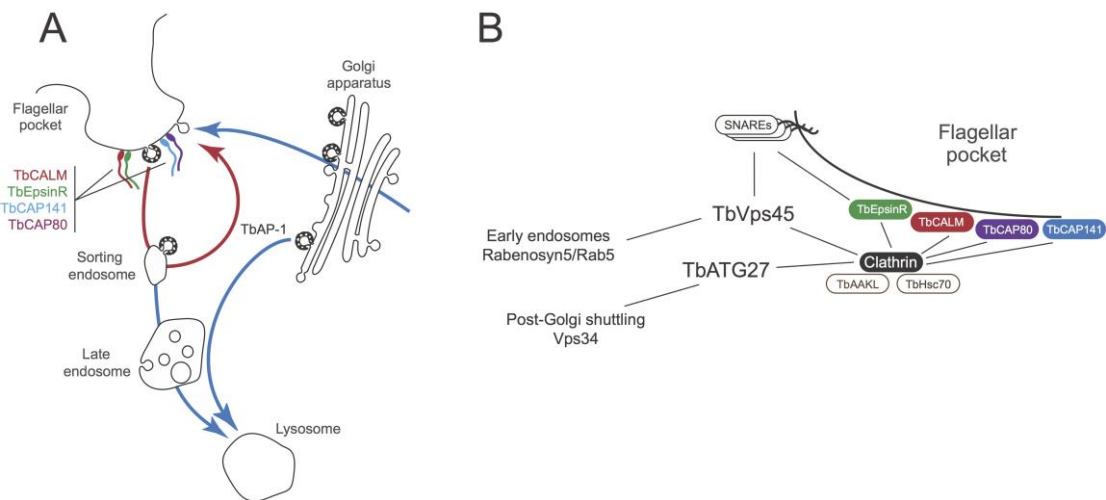


Figure 7. Interactome subtending clathrin-mediated endocytosis in trypanosomes. Whilst clathrin is distributed widely throughout the trypanosome endomembrane system, flagellar pocket specific factors act in the earliest stages of endocytosis. TbCALM and TbEpsinR are likely recruited via localised lipid accumulation. TbCAP141 and TbCAP80 are also enriched in flagellar pocket associated clathrin structures. We propose that the large disordered regions of these proteins likely contribute to the generation of local membrane curvature through molecular crowding. Additionally, a conserved core of clathrin associated accessory proteins, TbVps45, TbAAKL, TbHsc70, likely act to coordinate clathrin coat dynamics and regulate the sorting of the conserved SNARE machinery.

Supplementary material

Lineage-specific proteins essential for endocytosis in trypanosomes

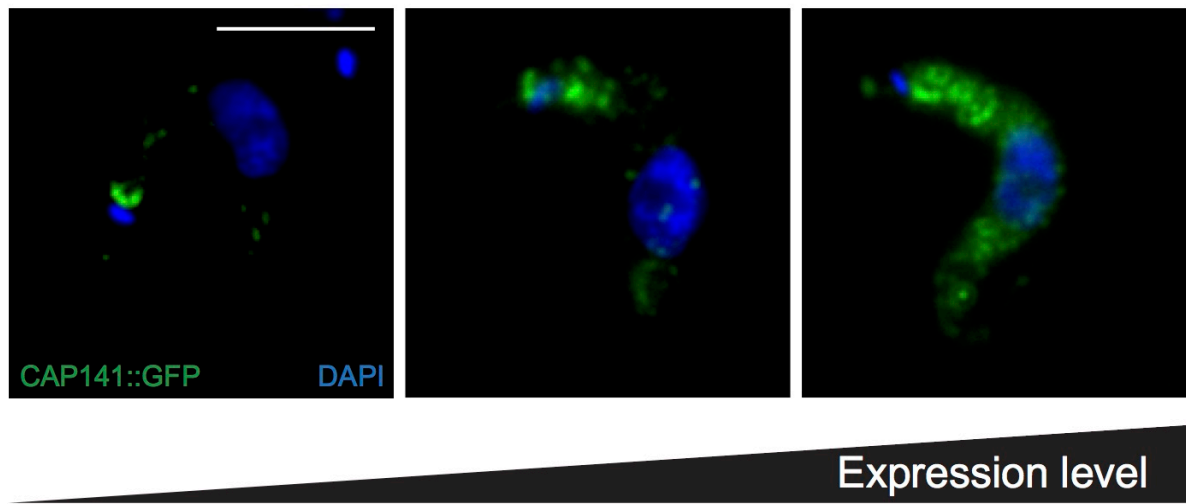
Paul T. Manna¹, Cordula Boehm¹, Samson O. Obado², Catarina Gadelha³, Andrej Sali⁴, Brian T. Chait², Michael P. Rout² and Mark C. Field^{1*}

¹Wellcome Trust Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, Scotland, DD1 5EH, UK, ²The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA, ³School of Life Sciences, University of Nottingham, Nottingham NG2 7UH and ⁴California Institute for Quantitative Biosciences, University of California, San Francisco, CA 94158, USA.

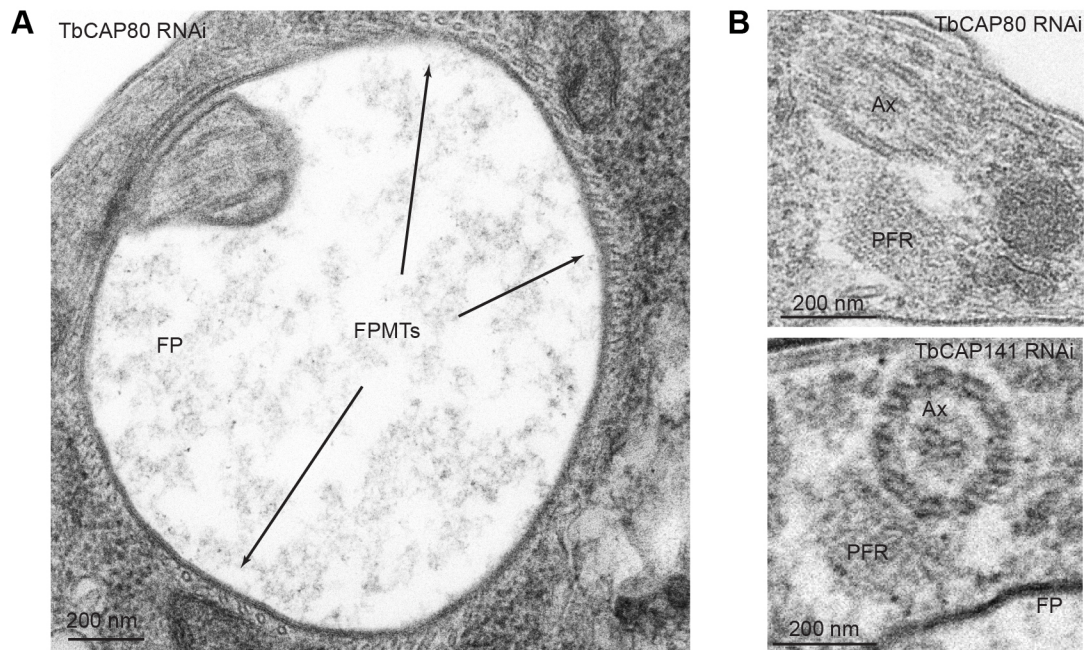
Supplemental table ST1. Peptide coverage, Log(e) confidence scores, and label-free quantitation for proteins identified in all affinity isolation experiments.

[Click here to Download Table S1](#)

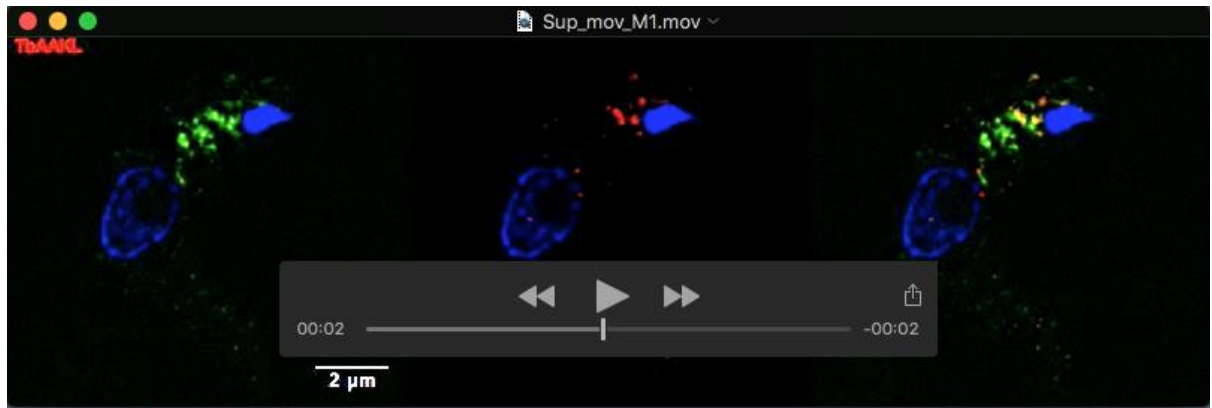
Supplementary figures



Supplemental Figure S1. Saturability of flagellar pocket localisation and mis-targeting of ectopically expressed CAP141 due to overexpression.



Supplemental Figure S2. (A) Microtubule invasion of the flagellar pocket region following TbCAP80 depletion. (B) Intracellular axoneme and paraflagellar rod material following TbCAP80 or TbCAP141 depletion. FP: flagellar pocket, FPMT: flagellar pocket microtubules, Ax: axoneme, PFR: paraflagellar rod.



Supplemental movie M1: 3D SIM reconstruction revealing co-localisation of clathrin heavy chain (green) and TbAAKL. The antibodies used were in house rabbit polyclonal anti-CHC followed by Alexa-488 anti rabbit and rat anti-HA (3f10 clone, Roche) followed by Alexa-568 anti-rat. DNA is shown in blue. Images were acquired on a Delta Vision OMX V3. Scale bar is 2 μ m.



Supplemental movie M2: 3D SIM reconstruction revealing co-localisation of clathrin heavy chain (green) and TbCAP80. The antibodies used were in house rabbit polyclonal anti-CHC followed by Alexa-488 anti rabbit and rat anti-HA (3f10 clone, Roche) followed by Alexa-568 anti-rat. DNA is shown in blue. Images were acquired on a Delta Vision OMX V3. Scale bar is 2μm.



Supplemental movie M3: 3D SIM reconstruction revealing co-localisation of clathrin heavy chain (green) and TbCAP141. The antibodies used were in house rabbit polyclonal anti-CHC followed by Alexa-488 anti rabbit and rat anti-HA (3f10 clone, Roche) followed by Alexa-568 anti-rat. DNA is shown in blue. Images were acquired on a Delta Vision OMX V3. Scale bar is 2μm.