Structures of three MORN repeat proteins and a re-evaluation of the proposed lipid-binding properties of MORN repeats.

Sara Sajko* (1), Irina Grishkovskaya* (1), Julius Kostan* (1), Melissa Graewert (2), Kim Setiawan (3), Linda Trübestein (1), Korbinian Niedermüller (3), Charlotte Gehin (4), Antonio Sponga (1), Martin Puchinger (1), Anne-Claude Gavin (4), Dimitri Svergun (2), Tom Leonard (1), Terry K. Smith (5), Brooke Morriswood (3), Kristina Djinovic-Carugo (1, 6)

* = equal contribution

Authors for correspondence:
Brooke Morriswood (brooke.morriswood@uni-wuerzburg.de)
Kristina Djinovic-Carugo (kristina.djinovic@univie.ac.at)

Affiliations:

[1]Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, Austria

[2]European Molecular Biology Laboratory, Hamburg Unit, c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany

[3]Department of Cell and Developmental Biology, Biocenter, University of Würzburg, 97074 Würzburg, Germany.

[4]European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany.

[5]School of Biology, BSRC, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9ST, UK

[6]Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, SI-1000 Ljubljana, Slovenia

Keywords: MORN repeat, MORN1, structure, phospholipids, *Trypanosoma,* Apicomplexa.

ABSTRACT

MORN (membrane occupation and recognition nexus) repeat proteins have a wide taxonomic distribution, being found in both prokaryotes and eukaryotes. Despite this ubiquity, they remain poorly characterised at both a structural and a functional level compared to other common repeat motifs such as leucine-rich repeats, armadillo repeats, WD40 repeats, and ankyrin repeats. In functional terms, they are often assumed to be lipid-binding modules that mediate membrane targeting, but direct evidence for this role is actually lacking. This putative activity was addressed by focusing on a protein composed solely of MORN repeats - *Trypanosoma brucei* MORN1. No evidence for binding to membranes or lipid vesicles by TbMORN1 either in vivo or in vitro could be obtained. TbMORN1 did interact with individual phospholipids, but it remains unclear if this was physiological or an artefact. High- and low-resolution structures of the MORN1 protein from *Trypanosoma brucei* and homologous proteins from the parasites *Toxoplasma gondii* and *Plasmodium falciparum* were obtained using a combination of macromolecular crystallography, small-angle X-ray scattering, and electron microscopy. The structures indicated that MORN repeats can mediate homotypic interactions, and can function as both dimerisation and oligomerisation devices.

INTRODUCTION

MORN (Membrane Occupation and Recognition Nexus) repeats were first discovered in 2000, following a screen for proteins present in the triad junctions of skeletal muscle (Takeshima et al., 2000). The junctophilins, the protein family identified in this screen, were observed to have 8 repeat motifs present in their N-terminal regions. The repeats were given the name MORN based on a proposed role in mediating plasma membrane association of the N-terminal domain of the junctophilins. The MORN repeats were initially classified as being 14 amino acids in length, with an approximate consensus sequence of YEGEWxNGKxHGYG (Takeshima, Komazaki et al., 2000). A bioinformatics analysis at the time indicated that assemblies of 8 consecutive MORN repeats were also present in a putative junctophilin orthologue in a nematode (*Caenorhabditis elegans*), a family of plant (*Arabidopsis thaliana*) lipid kinases, and a bacterial (*Cyanobacterium*) protein (Takeshima et al., 2000). Later genome-era bioinformatics has shown that MORN repeat proteins are in fact found ubiquitously, being present in both eukaryotes and prokaryotes (El-Gebali, Mistry et al., 2019).

The number of MORN repeat proteins in any given protein can vary greatly, from two to over 20, and they are found in combination with a wide range of other domains and protein repeat motifs. Most published work now favours a 23-amino acid length for a single MORN repeat, with the highly-conserved GxG motif at positions 12-14. A 14-amino acid length is still favoured by some groups, however (Habicht, Woehle et al., 2015). Notable mammalian MORN repeat proteins besides the junctophilins include ALS2/alsin, at least two radial spoke proteins (RSPH10B, RSPH1/meichroacidin), the histone methyltransferase SETD7, and retinophilin/MORN4 (Tsuchida, Nishina et al., 1998, Wilson, Jing et al., 2002, Otomo, Hadano et al., 2003, Mecklenburg, 2007).

MORN repeats are generally assumed to be lipid-binding modules, but direct evidence for this function is actually lacking. In junctophilins, there is good evidence that the N-terminal region containing the MORN repeats mediates plasma membrane targeting (Takeshima et al., 2000, Nakada, Kashihara et al., 2018, Rossi, Scarcella et al., 2019). It has not been demonstrated whether the MORN repeats or the other sequences in the N-terminal region are responsible for this however, or if this targeting is due to protein-lipid or protein-protein interactions.

Similarly, although there is good evidence that the N-terminal region (amino acids 1-452) of junctophilin-2 can directly bind lipids, it has not specifically been shown that the MORN repeats are responsible. Binding could potentially be mediated by other nearby sequences, especially the run of over 100 amino acids that occurs between repeats 6 and 7 (Bennett, Davenport et al., 2013). Work on the family of plant phosphatidylinositol(4)phosphate 5-kinases (PIPKs) that contain MORN repeats has led to suggestions that the repeats might regulate the activity of the kinase domain, bind to phospholipids, or mediate protein-protein interactions (Ma, Lou et al., 2006, Im, Davis et al., 2007, Camacho, Smertenko et al., 2009). It therefore remains unclear what role(s) this ubiquitous class of repeat motifs actually have (Mikami, Saavedra et al., 2010).

Coupled to this lack of unambiguous functional data is a lack of high-resolution structural information, exemplified by the ongoing lack of consensus as to whether a single repeat is 14

or 23 amino acids. This contrasts sharply with the considerable amount of information available on other classes of protein repeat motifs such as ankyrin repeats, leucine-rich repeats, or WD40 repeats (Andrade, Perez-Iratxeta et al., 2001). Until very recently, the structure of the SETD7 histone methyltransferase was the sole representative of the MORN repeat protein family in the protein data bank (PDB) (Jacobs, Harp et al., 2002, Wilson et al., 2002, Xiao, Jing et al., 2003). Even here, the structure of the N-terminal domain containing the repeats is incomplete, and the level of sequence similarity of the repeats to those of junctophilins and other family members makes assignment difficult. Each repeat appears to form a β-hairpin with an acidic surface, but it remains unclear if this is a general property of MORN repeats. The SETD7 structure has not been analysed in this context, with more work focusing on its catalytic methyltransferase domain. In 2019, and while this manuscript was in preparation, Li et al. published the first high-resolution structure of a canonical MORN repeat protein, specifically the MORN4/retinophilin protein in complex with its Myo3a binding partner (Li, Liu et al., 2019). This structure contains four MORN repeats. More structural analysis of MORN repeat proteins is still needed however, in particular for providing a structure-based definition of the repeat class itself.

To address this, and additionally to tackle the question of putative lipid binding, it would obviously be advantageous to utilise a protein that is composed solely of MORN repeats. In this way, the contribution of other sequences or domains could be discounted. The MORN1 protein from the early-branching eukaryote *Trypanosoma brucei* is an ideal candidate in this regard, and has the advantage of also being well-characterised at a cell biology level (Morriswood, He et al., 2009, Esson, Morriswood et al., 2012, Morriswood, Havlicek et al., 2013, Morriswood & Schmidt, 2015).

TbMORN1 consists of 15 consecutive 23-amino acid MORN repeats, with barely any intervening sequence whatsoever (Fig. 1A). In *T. brucei*, TbMORN1 is localised to an ~2 μm long cytoskeleton-associated complex (the hook complex) that is found just below the inner leaflet of the plasma membrane. The hook complex encircles the neck of a small invagination of the plasma membrane that contains the root of the cell's single flagellum (Lacomble, Vaughan et al., 2009). This invagination, termed the flagellar pocket, is the sole site of endo- and exocytosis in trypanosome cells and is thought to be analogous to the ciliary pocket that is found at the base of some mammalian primary cilia (Grunfelder, Engstler et al., 2003, Engstler, Thilo et al., 2004, Molla-Herman, Ghossoub et al., 2010).

Previous work on TbMORN1 demonstrated by fluorescence recovery after photobleaching (FRAP) that it is a stable component of the hook complex, and a list of its binding partners and near neighbours has been obtained using proximity-dependent biotin identification (BioID) (Esson et al., 2012, Morriswood et al., 2013). In functional terms, depletion of TbMORN1 by RNAi in the mammalian-infective form of the parasite resulted in a lethal phenotype (Morriswood et al., 2009). Functional analysis indicated that the protein might be involved in endocytosis, as well as regulating the flow of macromolecular cargo through the neck of the flagellar pocket (Morriswood & Schmidt, 2015).

In this study, a detailed biochemical, structural, and functional analysis of the TbMORN1 protein was carried out. A truncated form missing the first MORN repeat that was best suited

for in vitro work was found not to bind to phospholipid vesicles under any conditions, although there are indications that it might be able to associate with individual lipid molecules. In addition, high-resolution crystal structures of a truncated form of the TbMORN1 protein and its homologues from the parasites *Toxoplasma gondii* (TgMORN1) and *Plasmodium falciparum* (PfMORN1) structures enabled a first structure-based definition of the MORN repeat itself, and provided confirmation that MORN repeats can mediate homotypic interactions - a function that may unify previous observations relating to MORN repeats.

RESULTS

TbMORN1 forms tail-to-tail dimers via its C-terminus

TbMORN1 is composed of 15 consecutive 23-amino acid MORN repeats, with a 5-amino acid extension after the 6th repeat (Fig. 1A). An alignment of the repeats in TbMORN1 revealed several highly conserved glycine residues, with a rough consensus of $YxGEWx_2Gx_3GxGx_3Yx_2Gx_2$ (Fig. 1A, sequence logo). Bioinformatic analysis of TbMORN1 predicted an all-β secondary structure, with each repeat expected to form a β-hairpin (strand-loop-strand) pattern (Fig. 1A).

To determine which TbMORN1 constructs were amenable for biochemical and structural studies, limited proteolysis assays were carried out. These assays used a series of proteases with different cleavage specificities (proteinase K, trypsin, and chymotrypsin) (Fig. S1A). The resulting proteolytic fragments were analysed by mass spectrometry (Fig. S1B). Comparison of the proteolytic fragments obtained at different dilutions suggested progressive digestion occurring from the N-terminus (Fig. S1B, compare K1 and K2, T1/T2/T3). The C-terminus of the molecule (repeats 13-15) seemed to be fairly stable, with no proteolytic digestion observed at this end. In addition, the pattern of fragments generally suggested that the assignment of repeat boundaries based on bioinformatic analysis was accurate.

Consequently, a panel of different truncations were cloned according to the MORN repeat boundaries predicted by the alignment (Fig. 1A). These truncations were named according to the number of repeats they contained - for example, TbMORN1(1-15) denotes full-length protein. These truncations were expressed in bacteria and purified using a two-step protocol combining affinity purification and size exclusion chromatography (Fig. S1C).

The oligomeric state and polydispersity of the purified proteins were investigated using size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). SEC-MALS analysis of TbMORN1(1-15) elution profiles suggested the formation of either aggregates or higher-order assemblies (Fig. 1B). The yields of TbMORN1(1-15) were always very low, making this construct not generally suitable for in vitro assays. By contrast, TbMORN1(2-15) displayed a well-defined monodisperse peak in SEC-MALS, with a molecular weight corresponding to a dimer (Fig. 1B). This strongly suggested that the first MORN repeat mediated oligomerisation. Further successive truncations from the N-terminus (TbMORN1(7-

15) and TbMORN1(10-15)) also eluted as monodisperse dimers, suggesting that dimerisation of TbMORN1 is mediated by the C-terminus (Fig. S1D). Consistent with this conclusion, removal of the last MORN repeat in the construct TbMORN1(1-14) resulted in the elution of a mixture of monomers, dimers, and higher-order structures (Fig. 1B). This demonstrated that the C-terminal repeats play an important role in dimer stabilisation.

Circular dichroism (CD) measurements taken of TbMORN1(1-15), TbMORN1(2-15), and TbMORN1(7-15) indicated β-strand character in all cases, with >30% antiparallel β-strand content for each construct (Fig. S1E). Therefore, the secondary structure content of TbMORN1 was in good agreement with a priori bioinformatic predictions (Fig. 1A). Thermostability measurements of TbMORN1(1-15), TbMORN1(2-15) and TbMORN1(7-15) using CD returned values in °C of 45.6 +/- 0.1, 43.5+/-0.1, and 42.2+/-0.1 respectively. The relative similarity of these values indicated that there had been no significant destabilisation of the protein resulting from truncation, consistent with the suggested repeat motif boundaries obtained using limited proteolysis (Fig. S1B).

To map which residues were likely mediating dimerisation, cross-linking mass spectrometry (XL-MS) was used to analyse TbMORN1(1-15). Two different chemical cross-linkers were used: EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide), which has a zero-length spacer arm and forms bonds between carboxyl groups and primary amines, and BS^3 (bis(sulphosuccinimidyl)suberate), which has an 11.4 Å spacer arm and cross-links primary amines. For both chemicals, most cross-links were observed to form between repeats 13, 14, and 15, especially via repeat 14 (Table S1). These data were consistent with those obtained by SEC-MALS (Fig. 1B, Fig. S1D), suggesting that TbMORN1 molecules in solution form tail-to-tail dimers via their C-termini. It remained unclear whether the polypeptide chains in these tail-to-tail dimers were in a parallel or antiparallel orientation, however.

TbMORN1 can bind lipid side chains but not phospholipid liposomes

As a first test of the hypothesis that MORN repeats can directly interact with phospholipids, protein-lipid overlay assays were carried out using PIP strips. Purified TbMORN1(1-15) was found to interact with a number of different phosphoinositide species on the PIP strips, principally PI(3)P, PI(4)P, and PI(5)P, but also $PI(3,4)P_2$, $PI(3,5)P_2$, $PI(4,5)P_2$, and phosphatidic acid (PA) (Fig. 2A). This same trend, with a stronger interaction observed with monophosphate PIPs than diphosphate PIPs, has also been seen in protein-lipid overlay assays using purified junctophilin-2. This might be due to a lower solubility of the monophosphate PIPs, however (Bennett et al., 2013). As a positive control, the PIP strips were incubated with the pleckstrin homology (PH) domain of phospholipase C delta (PLCδ). Strong binding was observed to $PI(4,5)P_2$ only, as expected (Fig. 2B).

Of these candidates, the most intriguing was $PI(4,5)P_2$. It is a known endocytic effector, is enriched in the flagellar pocket membrane of trypanosomes, and depletion of TbMORN1 by RNAi resulted in a phenotype suggestive of an endocytosis defect (Demmel, Schmidt et al., 2014, Morriswood & Schmidt, 2015). Furthermore, a PI(4)P 5-kinase, TbPIPKA, has been

shown to have a strongly overlapping distribution with TbMORN1 in vivo (Demmel et al., 2014).

To confirm that TbMORN1 could interact with PI(4,5)P_2, fluorescence anisotropy was used as a second independent and quantitative approach. In this method, the tumbling of a fluorescently-labelled lipid in solution is reduced upon binding to a larger protein molecule, and the change in the polarisation of the emitted light can be measured. TbMORN1(2-15) showed excellent binding to BODIPY TMR-labelled PI(4,5)P_2, with a K_d of approximately 7.5 +/- 4.1 μM (Fig. 2C). To narrow down the PI(4,5)P_2 binding site(s) on TbMORN1, truncation constructs were again used. Both TbMORN1(7-15) and TbMORN1(10-15) showed strong binding to PI(4,5)P_2, with K_d values of 7.7 +/- 2.6 and 1.0 +/- 0.1 μM respectively, indicating that the interaction was occurring in the C-terminal portion of TbMORN1 (Fig. 2C).

Curiously, while the anisotropy signal for TbMORN1(7-15) was lower than that of TbMORN1(2-15), the highest anisotropy signal of all was obtained for TbMORN1(10-15) (Fig. 2C). This was surprising, as one would usually expect the largest construct containing the binding site to give the highest anisotropy signal - the larger protein means less tumbling of the fluorophore-conjugated lipid, resulting in greater polarisation of the emitted light and a higher anisotropy signal.

As an additional test, native gel electrophoresis assays were carried out. In these assays, interaction between a protein and a fluorescently-labelled lipid is detected by co-migration of the lipid with the protein. Both TbMORN1(2-15) and TbMORN1(10-15) produced a band shift of BODIPY TMR-labelled PI(4,5)P_2. In contrast to the PIP strip data, no band shift was seen for BODIPY TMR-labelled PI(4)P (Fig. S2A, B). This too was surprising as this lipid had given the strongest signal in the overlay assays (Fig. 2A). Of note, all of the assays to this point - overlay assays, fluorescence anisotropy, native gel electrophoresis - had looked at the binding of TbMORN1 to isolated lipids below their critical micelle concentration. This situation is somewhat artificial, and so for better physiological relevance the binding of TbMORN1 to phospholipid membranes was investigated using liposome pelleting assays.

The assays were carried out with liposomes of roughly 100 nm diameter, and containing 0-20% PI(4,5)P_2. Varying the amount of PI(4,5)P_2 from 0-20% did not result in any increase in the amount of TbMORN1(2-15) present in the pellet (P) fraction, while strong concentration-dependent co-sedimentation was observed with the positive control Doc2B (Fig. 2D). Quantification of the amount of protein in the pellet relative to the 0% PI(4,5)P_2 condition showed that the amount of TbMORN1(2-15) was within roughly 10% of the control condition at all times (Fig. 2E).

Additional assays were carried out using liposomes of differing diameters (100 and 400 nm) in order to vary curvature, and different cholesterol contents (0 and 40%) in order to concentrate PI(4,5)P_2 into microdomains. Neither approach produced an increase in TbMORN1(2-15) association (Fig. S3A). Therefore, TbMORN1 appeared to bind to isolated molecules of PI(4,5)P_2 but not to liposomes, in agreement with the fluorescence anisotropy assays with isolated lipids which had suggested a binding site in the C-terminal region of TbMORN1.

The TbMORN1 primary structure was next examined for putative PI(4,5)P_2-binding sites, based on similarity to known PI(4,5)P_2-binding sites in the PH, PLC-δ1, CALM-N, ENTH, FERM, PTB, and I3P-RBC domains (Franzot, Sjoblom et al., 2005). Two candidate sites were identified in the C-terminus of TbMORN1: one in repeat 13, and one in repeat 14. The location of these sites was in agreement with the fluorescence anisotropy assays, which had suggested binding to the C-terminal region of TbMORN1 (Fig. 2C).

Site-directed mutagenesis was carried out on these sites in the TbMORN1(2-15)-encoding construct either singly or in combination. These single or double mutants of TbMORN1(2-15) were then expressed, purified, and biophysically characterised. Thermostability of the purified proteins was assessed using differential scanning fluorimetry (DSF), and indicated that mutagenesis of the site in repeat 14 and the double mutant both resulted in slightly decreased stability (i.e. a lower T_m) than TbMORN1(2-15) (Table S2).

SEC-MALS analysis showed that while mutagenesis of the site in repeat 13 did not alter the dimeric state of TbMORN1(2-15), mutagenesis of the site in repeat 14 resulted in a mixture of monomers and dimers, while mutagenesis of both sites resulted in a monomeric protein (Fig. S3B). This supported the conclusion from the SEC-MALS studies that the dimerisation site also resides in the C-terminal segment (Fig. 1B, Fig. S1D). CD analyses indicated that the constructs retained an all-β secondary structure (Fig. S3C).

Despite these changes, all three TbMORN1(2-15) mutant constructs showed unimpaired binding to PI(4,5)P_2 in fluorescence anisotropy assays (Fig. S3D). Furthermore, all three mutants actually had stronger binding affinities for PI(4,5)P_2 than the wild-type TbMORN1(2-15) construct, with K_d values of ~ 1 μM. Although these results might indicate that TbMORN1(2-15) utilises a non-canonical PI(4,5)P_2 binding site, a more parsimonious explanation was that the data reflected a nonspecific interaction. Subsequent experiments favoured this latter interpretation.

All fluorescence anisotropy assays had been carried out using labelled PI(4,5)P_2 with 16-carbon aliphatic chains. When the assays were repeated using a version of PI(4,5)P_2 with 6-carbon aliphatic chains, no binding was seen (Fig. S4A). A range of lipids (PI, PI(4)P, PI(3,4)P_2, PI(3,5)P_2) with 6-carbon aliphatic chains were then tested for interaction with TbMORN1(2-15) by fluorescence anisotropy, and no binding to any was seen (Fig. S4A).

Furthermore, it was observed that TbMORN1(2-15) could also bind to PI(3,4,5)P_3 - which had not given a positive result in the overlay assays (Fig. 2A) - if the lipid had 16-carbon aliphatic chains (Fig. S4B). This interaction with 16-carbon PI(4,5)P_2 and PI(3,4,5)P_3 was seen using BODIPY-fluorescein-labelled lipids instead of BODIPY TMR-labelled ones (Fig. S4B). The binding of TbMORN1(2-15) to 16-carbon versions of either PI(4,5)P_2 or PI(3,4,5)P_3 was comparable to that seen for positive controls (Fig. S4C).

These results strongly suggested that the observed binding of TbMORN1(2-15) to PI(4,5)P_2 by both fluorescence anisotropy and native gel electrophoresis was in fact due to interaction

with the 16-carbon aliphatic chains and not with the lipid headgroup. An inability to interact with the lipid headgroup was also consistent with the observed lack of interaction of TbMORN1(2-15) with phospholipid membranes in the liposome pelleting assays (Fig. 2D,E, Fig. S3A).

In conclusion, while protein overlay assays indicated that TbMORN1(1-15) and TbMORN1(2-15) could interact with phospholipid species, subsequent fluorescence anisotropy and native gel electrophoresis assays strongly suggested that this interaction was actually with the aliphatic chains. No evidence for TbMORN1(2-15) binding to phospholipid liposomes in vitro was obtained.

TbMORN1 co-purifies with PE but does not bind to lipid vesicles in vitro

A caveat to the previous conclusion was that the purified recombinant TbMORN1 was obtained via mechanical lysis of bacterial cells in the absence of detergent (see Materials and Methods). It was therefore possible that TbMORN1 was associating with bacterial lipids that might be occupying the binding sites.

To test this, purified recombinant TbMORN1(1-15) and TbMORN1(10-15) were treated according to a de-lipidation protocol, and the resulting supernatants were analysed by mass spectrometry. Interestingly, large amounts of phosphatidylethanolamine (PE) were found to have been bound to TbMORN1(1-15) (Fig. 3A). The co-purifying PE displayed a narrow window of molecular moieties differing in the length of the aliphatic chain. No significant PE presence was detected in supernatants obtained following de-lipidation of TbMORN1(10-15) (Fig 3B). Given that TbMORN1(10-15) showed robust binding to the 16-carbon chain PI(4,5)P_2 (Fig. 2C), this suggested that PE was not occluding the binding site.

To remove co-purifying lipids, recombinant TbMORN1(1-15) was purified using hydrophobic interaction chromatography - in this regime, almost no lipids were detected in the elutions. Triton X-100 treatment was also found to efficiently remove bound lipids.

Co-purification of PE with TbMORN1(1-15) is not evidence of physiological interaction, as PE is highly abundant in bacteria and carries a net neutral charge (Cronan, 2003). As such, PE might simply have associated with the recombinant protein following lysis of the bacteria prior to purification. To investigate this, mass spectrometry analysis of whole-cell lipids from bacteria expressing recombinant TbMORN1(1-15) or TbMORN1(10-15) was carried out. Curiously, bacterial cells expressing TbMORN1(1-15) showed elevated levels of PE (Fig. 3C). This effect was not seen in bacterial cells expressing TbMORN1(10-15), which had approximately normal levels of PE (Fig. 3D).

In summary, TbMORN1(1-15) but not TbMORN1(10-15) was found to have co-purifying PE, and bacterial cells expressing TbMORN1(1-15) appeared to have elevated levels of PE. This suggested that PE might be a plausible candidate for a physiological interaction partner of TbMORN1.

To test the possibility of PE binding, and also to check whether the presence of co-purifying PE was affecting possible PI(4,5)P_2 binding, pelleting assays using sucrose-loaded vesicles (SLVs) were carried out. Sucrose loading increases the vesicle mass and enables much lower centrifugation forces to be used for pelleting, thereby reducing the risk of false positives due to protein aggregates pelleting independently of vesicle association. Three types of SLVs were used: PE-enriched SLVs, PI(4,5)P_2-enriched SLVs, and SLVs reconstituted from purified trypanosome whole-cell lipids. These were incubated with TbMORN1(2-15) purified either in the absence or presence of Triton X-100, i.e. either without or with co-purifying bacterial lipids.

As expected, in the absence of Triton X-100 treatment to remove bound lipids, purified recombinant TbMORN1(2-15) did not co-sediment with any of the liposome preparations (Fig. 4A). Even after purified recombinant TbMORN1(2-15) was treated with Triton X-100 to remove bound lipids, no association with the liposomes was seen (Fig 4B). As a positive control for pelleting, the PH domain of PLCγ was used. This showed robust and specific pelleting in presence of PI(4,5)P_2-containing liposomes, but in no other conditions (Fig4C, arrow). At this point, the candidate approach to TbMORN1 lipid binding was discontinued.

As an unbiased and high-throughput approach, TbMORN1 was next tested in a liposome microarray assay (LiMA). LiMA enables sampling of a wide range of liposome compositions and curvatures in a single experiment. Readout is via the detection of two fluorophores - one on a carrier lipid in the liposomes, and one on the protein of interest (Saliba, Vonkova et al., 2014, Saliba, Vonkova et al., 2016). Colocalised signals are regarded as positive hits. For these assays EGFP-TbMORN1(2-15) was expressed and purified. The EGFP-TbMORN1(2-15) protein showed no significant preference or affinity for liposomes across the whole range of conditions tested (Fig. 4D,E). Each lipid was tested at three different concentrations. Of note, no binding was seen to either phosphoinositide lipids or PE. In contrast, the PH domain of PLCδ, which was again used as a positive control, showed a strong and specific binding to PI(4,5)P_2 liposomes and, to a lesser extent, SM-enriched ones (Fig. 4F, note difference in y-axis scale compared to panels D, E).

In summary, purified recombinant TbMORN1(2-15) showed no binding to liposomes under any conditions assayed (Fig. 2D,E, Fig. S3A, Fig. 4A-F). TbMORN1(2-15) showed robust binding to PI(4,5)P_2 and PI(3,4,5)P_3 molecules in fluorescence anisotropy assays, but only when lipid reporters with 16-carbon aliphatic chains were used (Fig. 2C, Fig. S3D, Fig. S4A, B). TbMORN1(1-15) was found to co-purify with PE. Taken together, these data suggest that TbMORN1 requires large hydrophobic chains of the lipid for binding, which would explain the negative results in the liposome-based assays.

TbMORN1 does not associate with membranes in vivo

At this point, the only remaining positive indicators of an interaction of TbMORN1 with lipid came from the PIP strips (Fig. 2A), and the bacterial mass spectrometry data (Fig. 3A, C). The latter data showed that TbMORN1(1-15) but not TbMORN1(10-15) co-purified with PE, and

that bacteria expressing TbMORN1(1-15) had elevated PE levels. These elevated cellular PE levels might represent upregulated synthesis to compensate for something binding and sequestering the lipid (Fig. 3A, C).

Notably, all these positive pieces of evidence related to TbMORN1(1-15), which was polydisperse in vitro and formed large oligomers (Fig. 1B). The membrane-binding activity of these polydisperse oligomers was not possible to test in vitro, as the purification yields of TbMORN1(1-15) were always low. As an alternative, the possible membrane association of full-length TbMORN1 protein was examined in vivo.

For these experiments, cell lines of bloodstream form *T. brucei* were generated that inducibly expressed full-length TbMORN1(1-15) with an N-terminal Ty1 epitope tag. The presence of the ectopic gene in the genomes of the transfected cells was confirmed by PCR analysis of genomic DNA. Induction of Ty1-TbMORN1(1-15) overexpression using tetracycline (Tet) produced a strong growth defect in all three *T. brucei* clones tested (Fig. 5A). A rise in the number of so-called "BigEye" cells with grossly enlarged flagellar pockets was seen in the same time 96-hour window (Fig. 5B). Such a phenotype had previously been seen following depletion of TbMORN1 (Morriswood & Schmidt, 2015). The BigEye phenotype is thought to result from perturbations to membrane traffic, especially endocytosis (Allen, Goulding et al., 2003).

Immunoblotting with anti-TbMORN1 antibodies confirmed tight and inducible expression of the ectopic Ty1-TbMORN1 protein (Fig. 5C, left panel). The presence of the Ty1 epitope tag in the ectopic protein was confirmed by blotting with anti-Ty1 antibodies (Fig. 5C, right panel). Quantification of the immunoblots indicated only around 2-fold overexpression of Ty1-TbMORN1 relative to endogenous TbMORN1 protein (Fig. 5D).

These strong negative effects were unexpected, and could potentially be due either to the overexpression of the MORN1 protein, or to the presence of the Ty1 tag. To check the first point, cells that inducibly overexpressed untagged TbMORN1 from an ectopic locus were generated. The presence of the ectopic gene in the genomes of assayed clones was confirmed by PCR. Induction of ectopic gene expression by addition of tetracycline resulted in overexpression of TbMORN1 in immunoblots (Fig. S5A). Quantification of overexpression indicated that approximately 7 times more TbMORN1 protein was present in induced cells relative to controls (Fig. S5B). Cells overexpressing untagged TbMORN1 exhibited a very strong growth defect, stronger than that seen for overexpression of Ty1-TbMORN1 (Fig. S5C). Therefore, TbMORN1 protein levels alone were capable of producing a negative effect on growth in the absence of the Ty1 tag.

It was however not possible to obtain cells that solely expressed Ty1-TbMORN1 by endogenous replacement, despite repeated attempts. This indicated that Ty1-TbMORN1 cannot functionally compensate for the loss of the endogenous protein.

Immunofluorescence microscopy analysis of whole cells labelled with anti-TbMORN1 antibodies showed that Ty1-TbMORN1-overexpressing cells displayed a whole-cell labelling pattern, unlike controls (Fig. 5E, left panels). Although no sign of protein aggregation was

observed, this suggested that perhaps the Ty1-TbMORN1 protein might not localise correctly. Immunofluorescence microscopy analysis of detergent-extracted cytoskeletons labelled using anti-Ty1 antibodies confirmed that Ty1-TbMORN1 was able to target correctly, however; as expected, no labelling was seen in controls. (Fig. 5E, right panels).

To confirm the immunofluorescence microscopy observations, one-step biochemical fractionation using the non-ionic detergent IGEPAL was used. The detergent-soluble cytoplasmic fraction (SN) was separated from the detergent-insoluble cytoskeleton pellet (P) by centrifugation (Fig. 5F). In control cells, endogenous TbMORN1 associated almost entirely with the cytoskeletal (P) fraction (Fig. 5G). Blotting fractions from Ty1-TbMORN1-overexpressing cells with anti-TbMORN1 antibodies showed that the overexpressed Ty1-TbMORN1 was mostly extracted by the detergent (Fig. 5G, arrow 1). However, a small amount did associate with the cytoskeleton (P) fraction. This association was accompanied by a displacement of some of the endogenous protein into the cytoplasmic SN fraction (Fig. 5G, arrows 2).

Quantification of fractionation data from multiple experiments supported the qualitative analysis (Fig. 5H). Summing the signals of Ty1-TbMORN1 and TbMORN1 present in the cytoskeletal fraction in overexpressing cells indicated that the total amount of cytoskeleton-associated protein was approximately the same as in controls (Fig 5H, grey bar). This suggested that there are a finite number of Ty1-TbMORN1 molecules that can associate with the cytoskeleton. As the total amount of endogenous and ectopic TbMORN1 associated with the cytoskeleton is roughly the same in both overexpressing cells and controls, this suggested also that the dominant negative cellular effects are primarily due to the endogenous and ectopic TbMORN1 in the cytoplasmic fraction.

To determine if the cytoplasmic fraction of endogenous and ectopic TbMORN1 in the overexpressing cells was membrane-associated or cytosolic, two-step fractionations were carried out. These assays involved a first extraction with digitonin, then a second extraction with IGEPAL. Digitonin has an affinity for cholesterol and other lipids enriched in the plasma membrane, so at the right concentration it should enable the extraction of cytosol while leaving organelles relatively intact (Adam, Marr et al., 1990).

To optimise the conditions for digitonin extraction, a cell line expressing cytosolic GFP as a marker was used (Batram, Jones et al., 2014). These cells were extracted with varying concentrations of digitonin, using 1% IGEPAL as a positive control for full extraction, and fractions were separated by centrifugation (Fig. S6A). The supernatant and pellet fractions were immunoblotted with antibodies against GFP and the ER lumenal chaperone BiP (Fig. S6B), and the results quantified (Fig. S6C). At 40 µg/ml, good extraction of GFP was obtained with only minimal extraction of BiP (Fig. S6B, arrow 1). A timecourse of extraction using 40 µg/ml digitonin was then carried out (Fig. S6D). Increasing the incubation time over a range of 15-30 min did not appear to increase the amount of GFP extraction (Fig. S6E,F). As a result, a 25 min incubation time was used in the subsequent experiments.

A two-step extraction using first digitonin and then IGEPAL was then carried out (Fig. 6A). The digitonin supernatant (SN1) was enriched for cytosol, while the membrane/organelle fraction

present in the pellet (P1) was subsequently extracted using IGEPAL and partitioned into SN2. Cells expressing cytosolic GFP were spiked in alongside Ty1-TbMORN1-overexpressing cells in order to use GFP as a cytosolic marker.

Analysing equal fractions by immunoblotting showed that the Ty1-TbMORN1 and TbMORN1 proteins not associated with the cytoskeleton were predominantly cytosolic (Fig. 6B). The GFP cytosolic marker was almost wholly extracted by the digitonin, and accompanied by the majority of the cytoplasmic Ty1-TbMORN1 and TbMORN1 (Fig. 6B, arrows 1).

The Ty1-TbMORN1 and TbMORN1 that were not extracted by digitonin and therefore present in P1 were not strongly extracted by IGEPAL and barely present in the second supernatant (SN2), while the ER marker BiP was (Fig. 6B, arrows 2). Almost all the Ty1-TbMORN1 and TbMORN1 present in P1 partitioned into the second pellet, P2. Quantification of multiple independent experiments using the three separate clones produced results consistent with the exemplary blot shown (Fig. 6C,D). The presence of the overexpressed Ty1-TbMORN1 and displaced endogenous TbMORN1 in the digitonin supernatant (SN1) indicated that they are predominantly cytosolic. This indicated that TbMORN1 does not associate with membranes in vivo.

In conclusion, the extensive studies conducted on TbMORN1 here provide no evidence whatsoever that its MORN repeats are able to associate with phospholipid membranes in vivo or directly interact with phospholipid vesicles in vitro. TbMORN1 was able to bind to individual lipid molecules, notably PE. This interaction does not appear to be mediated by lipid headgroups however, and it is very hard to imagine how it would occur under physiological conditions unless TbMORN1 is a carrier and not a membrane-binding protein.

If MORN repeats do not bind membranes, then this raises the question of what they really do, and whether this other function might unify the various observations made about MORN repeat proteins to date. ALS2 has been suggested to use its MORN repeats to form an antiparallel dimer (Kunita, Otomo et al., 2004), and the evidence obtained here showed that TbMORN1 molecules also formed tail-to-tail dimers via their C-termini (Figs. 1, S1). Mammalian PI(4)P 5-kinases are also dimers, implying that the MORN repeats found at the N-termini of the family of plant PIPKs might function to mediate homotypic interactions (Rao, Misra et al., 1998). To investigate how the MORN1 dimers were being formed, high-resolution structural studies were used.

High-resolution crystal structures of three MORN repeat proteins

Crystallisation trials were initially performed with TbMORN1(2-15) and TbMORN1(7-15). Diffraction data were obtained for TbMORN1(7-15) in two different crystal forms ($P2_1$ and C2), but attempts to solve the phase problem using experimental phasing approaches (multiple isomorphous replacement, multiple anomalous scattering exploiting selenium and sulphur atom signals) and molecular replacement failed due to low reproducibility of crystals, anisotropy of diffraction data, and absence of sufficient homology of TbMORN1 to other

MORN repeat-containing proteins of known structure. At that point, the only MORN repeat-containing protein for which the crystal structure was solved was SETD7, a histone methyltransferase. SETD7 is predicted to contain up to 6 MORN repeats at its N-terminus, but they display low sequence homology to both the MORN repeats of TbMORN1 and the consensus MORN repeat sequence obtained by Pfam. This prevented successful use of molecular replacement as an approach.

As a new tactic, the MORN1 proteins from the apicomplexan parasites *Plasmodium falciparum* (PfMORN1) and *Toxoplasma gondii* (TgMORN1) were analysed. Despite their evolutionary distance, they share high (57% and 54%, respectively) sequence identity with TbMORN1 (Fig. S7A,B). CD measurements of purified recombinant protein indicated that TgMORN1, TgMORN1(7-15), PfMORN1, PfMORN1(2-15) and PfMORN1(7-15) all had an overall β-structure (Fig. S7C,D). This agrees with bioinformatic predictions and is consistent with the data obtained for TbMORN1 (Fig. S1E).

Diffracting crystals of selenomethionine-labelled PfMORN1(7-15) were obtained and its crystal structure was determined to 2.14 Å resolution using the single-wavelength anomalous dispersion (SAD) method. The structures of TgMORN1(7-15) and both $P2_1$ and C2 crystal forms of TbMORN1(7-15) were subsequently determined to 2.90, 2.35 and 2.53 Å resolution, respectively, with the PfMORN1(7-15) structure used as a search model for the molecular replacement (Fig. 7A-C). The structures of PfMORN1(7-15), TgMORN1(7-15), and both $P2_1$ and C2 forms of TbMORN1(7-15) were refined to an R_{work}/R_{free} of 23.0%/26.4%, 28.3%/32.2%, 23.2%/25.6% and 22.5%/28.2%, respectively (Table 1). TbMORN1(7-15) $P2_1$ (Fig. 7A), C2 (Fig. S8A) and TgMORN1(7-15) (Fig. 7B) crystallised with one dimer in the asymmetric unit, while PfMORN1(7-15) (Fig. 7C) crystallised with one subunit in the asymmetric unit, and the functional dimer was formed via crystallographic symmetry axis.

All MORN1(7-15) crystal structures showed subunit interaction via the C-terminal regions to form antiparallel tail-to-tail dimers, with variable inter-subunit angles and dimerisation interfaces (Fig. 7A-C, Fig. S8A). The two subunits in the TbMORN1(7-15) C2 crystal form and the TgMORN1(7-15) dimer made a rather straight assembly, while in the $P2_1$ form they displayed a bend of about 30° (Fig. S8A). The TbMORN1(7-15) and TgMORN1(7-15) dimers thus appeared as rod-shaped particles. This dimer architecture is consistent with the limited proteolysis data, which had indicated that the N-terminal regions of the molecule are more exposed (Fig. S1A,B). Interestingly, PfMORN1(7-15) displays a V-shaped dimer with an inter-subunit angle of about 45°, and incorporated a structural Zn^{2+} ion at the dimer interface (Fig. 7C). One crystal form of TgMORN1(7-15) also adopts the same V-shape seen for PfMORN1(7-15) (Fig. S9B). Here too, a Zn^{2+} ion is also found at the dimer interface.

Superimposing the TbMORN1(7-15) $P2_1$ subunit with the structures of the TgMORN1(7-15) and PfMORN1(7-15) subunits over 202 C_α atoms yielded rmsd values of 1.0 and 1.1 Å, respectively, revealing high structural similarity between the three proteins. The common structural feature of all three subunits is an elongated and twisted β-sheet. The curved shape of each constituent MORN repeat forms a groove laterally delimited by a rim. An individual MORN1(7-15) subunit is approximately 80 Å in length and displays a longitudinal groove of about 16 Å in depth (Fig. 7A).

Comparison of a single TbMORN1(7-15) subunit with known three-dimensional structures was carried out using the DALI server (Holm & Rosenstrom, 2010, Holm & Laakso, 2016). This revealed closest similarity with the G-box domain at the C-terminus of the human CPAP protein. CPAP is a centriolar protein essential for microtubule recruitment. The G-box comprises a single elongated β-sheet with all residues being solvent-exposed, and is capable of forming supramolecular assemblies [PDB entry 4LZF, (Hatzopoulos, Erat et al., 2013)]. Despite the low sequence identity (10%) between TbMORN1(7-15) and CPAP, the Z-score calculated by DALI suggested a significant similarity between the two structures (rmsd over 155 superimposed C_α atoms = 3.9 Å, Z-score = 12).

A structure-based redefinition of the MORN repeat

The crystal structures confirmed that each 23-amino acid MORN repeat is composed of a β-hairpin followed by a 6-residue loop that connects to the next MORN repeat. Each β-hairpin is composed of two 6-residue β-strands connected by a 5-residue loop. The MORN repeats from all three crystal structures could be readily superimposed, showing a high level of structural conservation (Fig. 7D). Based on this high level of conservation, a revised sequence alignment of the TbMORN1 repeats was constructed that better reflects the structural architecture of the protein (Fig. 7E, compare with Fig. 1A). The alignment of repeats 7-15 was obtained directly from structural superpositions and used to bootstrap the alignment of the upper part (repeats 1-6).

The new consensus MORN repeat sequence displays three highly-conserved features: a GxG motif at the start, a conserved glycine (G) at position 10, and a YEGEW motif at positions 13-17 (Fig. 7E). A slightly less conserved LxY motif is at positions 5-7. The GxG motif is at the beginning of the first β-strand, while the YEGEW motif comprises most of the second β-strand.

The glycine residues at position 10 are in a β-hairpin of type I, where the most commonly-observed residue at this position is a glycine (Hutchinson & Thornton, 1994). The GxG motif is strictly conserved because the first G residue adopts a main chain conformation mapping to the lower right corner of the Ramachandran plot, which is exclusively allowed for glycines. The conformation of this glycine is stabilised via a main-chain hydrogen bond with the tryptophan (W) residue of the YEGEW motif (Fig. 7F). The high conservation of the second glycine residue in the GxG motif is to accommodate the highly conserved neighbouring aromatic residues from the YEGEW and LxY motifs, as any larger side chain would create steric clashes. The tyrosine and tryptophan side chains of the YEGEW and LxY motifs provide a textbook example of aromatic stacking, filling up the groove and stabilising the tertiary structure of the TbMORN1 subunit (Fig. 7F, G). The highly conserved tyrosine of the YEGEW motif is sandwiched in a T-shaped π-stacking interaction between the highly conserved tryptophan residue from its own motif and the tryptophan residue in the next YEGEW motif (Fig. 7F). The tyrosine of the LxY motif is stabilised via hydrophobic or aromatic interactions with the leucine residues in its own and the subsequent LxY motif.

Three MORN repeats of SETD7 can be aligned with TbMORN1 repeat 7 over 22-23 aligned C_α atoms with an rmsd of 2.3, 1.5 and 1.9 Å respectively (Fig. S8B). The first glycine residue in the SETD7 MORN repeats is conserved with that in TbMORN1, while tyrosine, phenylalanine, and valine replace the tryptophan of the YEGEW motif (Fig. S8C).

TbMORN1 displays an overall negative charge

The TbMORN1(7-15) antiparallel dimer displays a two-fold symmetry perpendicular to the longitudinal axis of the assembly (Fig. 7A). Analysis of amino acid conservation derived from a sequence alignment of representative MORN repeat-containing proteins showed a well conserved stretch of residues in the groove (Fig. 8A). Due to the twofold symmetric nature of the quaternary structure assembly, surface properties are displayed on opposite sides of the elongated dimeric particle and perpetuated symmetrically along the rims (Fig. 8A-C).

Each TbMORN1 subunit displays a negative electrostatic potential. The most prominent feature of each subunit is the negatively charged groove, which contributes to the highly negative electrostatic potential of the dimer. This groove is flanked by a larger non-charged area with a central positively-charged pocket (Fig. 8B, left-hand side, 8C middle). The negatively-charged patch is formed by residues residing on loops of MORN repeats 9–13, and several of them display high sequence conservation throughout the structural alignment (Fig. 8A,B, left-hand side).

The positive charge of the pocket closer to the dimer interface is contributed by the universally conserved Lys316 residue, which is positioned on a loop connecting MORN repeats 13 and 14, and which is juxtaposed with Arg293 from the loop connecting MORN repeats 12 and 13. The Arg293 residue is additionally involved in aromatic stacking interactions with the residues Tyr278, Trp288 and Phe311, the latter forming a large surrounding non-charged area.

The rest of the subunit surface towards the N-terminus of each TbMORN1 subunit displays a non-charged and partially hydrophobic character (Fig. 8C), while the opposite side of the subunit displays a fairly uniform distribution of negative charges (Fig. 8B, right-hand side). The overall negative charge of TbMORN1(7-15) and the lack of pronounced positive patches that could serve as binding sites for negatively-charged phospholipid polar heads is in line with the negative binding data.

TbMORN1 forms an extended antiparallel dimer

The dimer interface is built from residues in MORN repeats 12-15, which connect the two subunits in an antiparallel, tail-to-tail orientation (Fig. 7A-C). The tightest overlap between the two subunits occurs at the site of MORN repeats 14 and 15. In the $P2_1$ and C2 TbMORN1(7-15) crystal forms, the dimer interfaces occupy surface areas of 747 $Å^2$ and 966 $Å^2$, respectively. Calculations of gain in solvation free energy (Δ^iG) upon dimer formation for TbMORN1(7-15) C2 and $P2_1$ performed with the PDBePISA package (Krissinel & Henrick,

2007) yielded values of -20.1 kcal/mol and -11.3 kcal/mol, respectively, with corresponding p-values of interface specificity 0.01 and 0.08. $\Delta^{i}G$ values lower than -10 to -15 kcal/mol and p-values lower than 0.5 are significant for stable protein dimers, indicating highly specific dimerisation interfaces (Krissinel & Henrick, 2007).

In both crystal forms of TbMORN1(7-15), the central core of the dimer interface is the same (Fig. 9A, Fig. S9A). Due to the twofold symmetry, the majority of the stabilising interactions are duplicated and build up an extended dimer interface. A series of hydrophobic and aromatic π-stacking interactions between residues from repeats 14 and 15, together with hydrogen bonds across the edges, stabilise the dimer (Fig. 9A, Fig. S9A).

Furthermore, the very negatively-charged C-terminal region of one TbMORN1 subunit forms an arch above the positively-charged platform contributed by Lys316 and Arg293 (Fig. 8B). Two residues from MORN repeat 14 - Lys325 (subunit A) and Asp326 (subunit B) - form a salt bridge, which further stabilises the dimer interface.

In the C2 crystal form, the dimer is additionally stabilised by two disulphide bridges formed between Cys351 at the C-terminus of repeat 15, and Cys282 from the β-hairpin loop of repeat 12 (Fig. S9A). In the $P2_1$ crystal structure, the position of the loop differs from that in the C2 dimer, and keeps the C_α atoms of Cys351 and Cys282 at a distance of 11.7 Å, preventing disulphide bond formation. Here, the side chain of Cys351 is engaged in a polar interaction with Asp303 (Fig. 9A, Fig. S9A).

A retroactive validation of the dimer interface came from the earlier $PI(4,5)P_2$-binding work. Mutagenesis of repeat 14 had unexpectedly produced a mixture of monomers and dimers, while the simultaneous mutagenesis of two candidate sites in repeats 13 and 14 had resulted in a pure monomer population (Fig. S3B). Analysis of the interaction and electrostatic maps of TbMORN1(7-15) and its mutagenised variants clearly showed that Arg293 and Lys316, residing in repeats 13 and 14 respectively, are crucial for maintaining TbMORN1 in a stable dimeric state through electrostatic interactions (Fig. 9C). Lys315 (in MORN repeat 14) does not appear to be directly involved in the dimerisation interface, but could peripherally contribute to the stabilisation of the C-terminal region of TbMORN1 through its electrostatic potential. Residues Arg292, Lys296 (both in MORN repeat 13), and Arg321, Lys325 (both in MORN repeat 14) were mapped to the outer surface of the dimer and therefore do not take part in the dimerisation interface (Fig. 9A, Fig. S9A).

The transition of $TbMORN1(2\text{-}15)^{Mut14}$ from a dimeric to a mixed monomer/dimer population is therefore a direct consequence of the single point mutation K316A, whereas the complete abrogation of dimerisation observed for the TbMORN1(2-15) double mutant can be attributed to a synergistic effect of both R293A and K316A point mutations (Fig. S3B, Fig. 9C).

The V-shaped and extended dimer forms of apicomplexan MORN1 proteins

Unlike the extended dimers of TbMORN1(7-15) and TgMORN1(7-15), the V-shaped PfMORN1(7-15) is mainly stabilised by a single Zn^{2+} ion incorporated into the core of its dimer

interface, spanning over 665 Å^2 (Fig. 7C, Fig 9B). One Cys306 and one Asp309 from each respective subunit tetrahedrally coordinate Zn^{2+} with the expected coordination distances (2.32 Å for Zn^{2+}-S and 1.94 Å for Zn^{2+}-O). In addition, residues from repeats 13-15 are involved in stabilising the dimer via a combination of hydrophobic, polar, and electrostatic interactions across the subunits.

In comparison to the TbMORN1(7-15) C2 crystal form, the extended TgMORN1(7-15) structure has an approximately 1.5-times smaller dimer interface (601 Å^2), which is contributed solely by residues from MORN repeats 13-15 (Fig. 7B, Fig. S9C). This is closer in size to the 747 Å^2 interface of the $P2_1$ crystal form of TbMORN1(7-15). The final C-terminal part of the protein seems to be flexible and does not engage in the stabilisation of the dimer. In comparison to TbMORN1(7-15), the dimer interface of TgMORN1(7-15) is not built around aromatic stacking, but instead employs hydrophobic interactions between Phe350 and neighbouring small hydrophobic residues, such as Leu335 and Leu327 (Fig. S9C).

Although TgMORN1(7-15) was predominantly found as an extended dimer (Fig. 7B), a V-shaped form similar to that of PfMORN1(7-15) was sporadically observed in the crystal lattice (Fig. S9B). The two V-shaped dimers share the same coordination sphere of a Zn^{2+} ion, which in TgMORN1(7-15) is provided by the Cys305 and Asp308 residues (Fig. S9D). The latter residue in turn interacts with the Ser310 residue from the other subunit. While the salt bridge and anion π-interactions are also conserved between the two V-shaped dimers, the aromatic stacking present in the core of the PfMORN1(7-15) dimer interface is functionally replaced in the TgMORN1(7-15) V-shaped dimer by a series of aromatic stacking interactions at its vertex. Although Asp residues 303 (TbMORN1), 308 (TgMORN1) and 309 (PfMORN1) are conserved in all the three proteins, the coordination of a Zn^{2+} ion clearly demands the presence of both cysteine and aspartate residues. Such pairs are present in PfMORN1 and TgMORN1, but not in TbMORN1, where the cysteine residue at the corresponding position is replaced by Leu301. The coordination residues map to the β-hairpin of MORN repeat 13, which in TgMORN1 and PfMORN1, but not TbMORN1, contains an insertion of a glutamate residue - Glu307 and Glu308, respectively. Taking part in an anion π-interaction with a phenylalanine residue (Phe303 and Phe304, respectively), the resulting Glu-Phe pairs effectively stabilise the TgMORN1 and PfMORN1 dimers in their V-shaped form. Moreover, these very same glutamate residues further stabilise the two V-shaped dimers by being involved in a salt bridge with lysine residues (Lys321 in TgMORN1 and Lys322 in PfMORN1). At the equivalent position in TbMORN1, Lys316 does not participate in a salt bridge, but rather points towards the negative patch at the C-terminal part of the other subunit and stabilises the extended dimer via electrostatic interactions.

To see whether it was possible to predict if a MORN repeat protein formed either extended or V-shaped dimers, or both, a comparative sequence analysis was carried out. The sequences of 15 selected MORN repeat proteins from various protist lineages were aligned with the C-terminal parts of TbMORN1, TgMORN1 and PfMORN1 encompassing repeats 12-15 (Fig. S10). The residues in the crystal structures of TgMORN1(7-15) and PfMORN1(7-15) forming the Zn^{2+}-coordination sphere and anion π-interaction were taken as a fingerprint for a V-shaped dimer. All sequences of MORN repeat proteins from kinetoplastids contain a leucine residue (Leu301 in the case of *T.* brucei) at the position of the coordinating Cys residue, and

lack the Phe-Glu either side of it forming the anion π-interaction pair - as these residues are essential for V-shape dimerisation, the kinetoplastid proteins are all thus predicted to form extended dimers only. Conversely, all protein sequences in the dataset from apicomplexans (*Toxoplasma gondii, Plasmodium falciparum, Gregarina niphandrode*s, *Babesia microti*, *Cryptosporidium parvum, Eimeria acervuline, Theileria equi*) contained these cysteine, phenylalanine, and glutamate residues. This suggests that all these apicomplexan proteins - probably MORN1 homologues - can adopt both extended and V-shaped forms. In addition, the sequences from the alveolates *Symbiodinium microadriaticum* and *Perkinsus marinus*, and the stramenopile *Aureococcus anophagefferens* also contain these three residues, implying that they too might adopt both extended and V-shaped conformations. Apicomplexans belong to the Alveolata clade, and both alveolates and stramenopiles are in the SAR supergroup (Adl, Bass et al., 2019). This suggests that the ability to adopt two conformations might have arisen within this specific clade, and possibly explains its absence from the kinetoplastid sequences, as kinetoplastids are excavates.

MORN1 proteins adopt extended conformations in solution

The crystal structures were all consistent with the earlier results of the cross-linking mass spectrometry experiments carried out on TbMORN1 in solution, which had indicated close proximity between repeats 13-15 (Table S1). To confirm the presence of V-shaped and extended dimers in solution, all three MORN1 proteins were structurally analysed using small-angle X-ray scattering (SAXS). SAXS analysis of TbMORN1(7-15) and TgMORN1(7-15) indicated an extended dimer in solution, and the crystal structures could be docked into the calculated molecular envelopes without difficulty (Fig. 10A, B, Fig. S9E). The SAXS analysis of PfMORN1(7-15) indicated an extended structure, similar to that seen for TbMORN1(7-15) and TgMORN1(7-15) (Fig. 10C). This supported the prediction that TgMORN1 and PfMORN1 are capable of adopting two different conformations. Subsequent SAXS analysis of TbMORN1(2-15) produced a molecular envelope for the almost-full-length protein, into which an extrapolated version of the model of TbMORN1(7-15) could be docked (Fig. 10D, Fig. S9E).

Rotary shadowing EM on TbMORN1(2-15) produced results consistent with the SAXS analysis, showing small kinked rods approximately 25 nm in length (Fig. 10E). The population was homogeneous, consistent with the monodispersity of this construct observed by SLS (Fig. 1B). Comparison of the maximal dimer dimension (D_{max}) for TbMORN1(2-15) obtained from SAXS and EM showed a very good agreement between the two values of 250-260 Å. Rotary shadowing EM was also carried out on the small amount of TbMORN1(1-15) that eluted from the SEC column. Full-length TbMORN1 was far more heterogeneous than TbMORN1(2-15), consistent with the polydispersity observed by SLS (Fig. 1B). In addition to single kinked rods, longer filamentous assemblies of varying length were occasionally observed (Fig. 10F). Rarely, much larger meshlike assemblies of full-length TbMORN1 could be observed (Fig. 10G), offering a tantalising clue into the higher-order assembly properties of the protein. These properties will be investigated in future work.

DISCUSSION

MORN repeat proteins are found ubiquitously in the tree of life, but the function of MORN repeats themselves has remained unclear and supported by very limited structural information to date. MORN repeats were first named almost 20 years ago, in a paper identifying the junctophilin protein family (Takeshima et al., 2000).

While there is abundant evidence that junctophilins associate with the plasma membrane, and that the MORN repeat-containing region is likely to mediate this, there is to the authors' knowledge no paper demonstrating that the junctophilin MORN repeats directly interact with lipids (Takeshima et al., 2000, Munro, Jayasinghe et al., 2016, Woo, Srikanth et al., 2016, Perni, Lavorato et al., 2017, Jayasinghe, Clowsley et al., 2018, Nakada et al., 2018, Rossi et al., 2019). Despite this, the evidence that the MORN repeat-containing region mediates plasma membrane targeting has been repeatedly cited as evidence that MORN repeats directly bind lipids. MORN repeats are thus widely assumed to be lipid-binding modules, despite there actually being no evidence for direct membrane binding. Warnings that the function of MORN repeats has not really been elucidated, and that extant data are frequently contradictory, have been largely overlooked (Mikami et al., 2010).

This study set out to provide a test of the lipid-binding hypothesis by using a protein composed exclusively of MORN repeats, TbMORN1. The data provide something of a cautionary tale.

At first, the evidence obtained using overlay assays and fluorescence anisotropy indicated an ability to bind specific phospholipids (Fig. 2A, C). This binding appears to be mediated primarily by the aliphatic chains of the phospholipids however, as no binding was ever seen in the anisotropy assays when small (6-carbon) lipids were used.

Purified recombinant TbMORN1(1-15) was at least found to co-purify with PE. Bacteria expressing TbMORN1(1-15) showed elevated levels of PE, which might additionally suggest that something was binding and sequestering the lipid, requiring them to upregulate synthesis (Fig. 3A, C). Almost no PE was found to co-purify with TbMORN1(10-15), and none was detected with the apicomplexan MORN1 proteins. This suggests that the N-terminal part of TbMORN1 is a major interaction area.

It is worth noting that neither PE nor any other lipids were observed in the crystal structure of TbMORN1(7-15). This might again be due to the fact that TbMORN1(7-15) does not contain the N-terminal part of the molecule, or possibly because the lysine methylation step used to enhance crystallisation altered its surface-exposed lysine residues, making them unavailable for electrostatic interactions with lipid molecules due to the loss of positive charge. Despite this evidence for binding to PE, no sign of direct binding by TbMORN1 to lipid vesicles in vitro (Fig. 2D,E, Fig. 4) or in vivo (Fig. 6) could be obtained.

A remaining question then is whether this ability to bind fatty acid chains greater than 6-carbon length is physiological, or an artefact. Given that TbMORN1 is a cytoskeleton-associated

protein, it is hard to imagine how it would be able to get access to the hydrophobic chains of membrane-embedded phospholipids in vivo, although it is localised directly under the cytoplasmic leaflet of the plasma membrane. The electrostatic profile of TbMORN1(7-15) is also not suggestive of membrane binding, with a strong overall negative charge profile (Fig 8B). Consequently, a physiological interaction with phospholipid membranes seems very unlikely.

It is important to note that phospholipid binding and membrane binding are two separate things. Not interacting with membranes might not preclude the TbMORN1 proteins taking some PE or other phospholipids out of membranes without stably interacting with the membranes themselves. A similar hypothesis would be that TbMORN1 functions as a carrier/transporter of lipids with aliphatic chains within a certain length, but it is difficult to envisage what the corresponding physiological role would be in the context of the flagellar pocket. For now, the simplest interpretation remains that there is no interaction of TbMORN1 with lipid membranes, and consequently no physiological interaction with membrane-embedded phospholipids either.

Due to the twofold symmetry of both extended and V-shaped MORN1 dimers – the dimeric particle displays the conserved groove on opposite sides of the dimer, therefore excluding this region for interaction with the membranes and leaving as an option the rims lining the groove (Fig. 8A-B). In the extended dimers, the two rims display a concave and a convex curvature, suggestive of membrane sculpting BAR domain proteins which interact with membranes mainly via non-specific electrostatic interactions (Salzer, Kostan et al., 2017, Carman & Dominguez, 2018). These curved surfaces nevertheless do not display a pronounced positive charge, nor the typical membrane insertion motifs characteristic for BAR domain proteins, refuting thus the hypothesis for membrane binding via non-specific electrostatic interactions. In conclusion, the quaternary structure architecture together with surface properties of the MORN1 dimers do not support membrane binding.

There are however a small number of caveats that might still allow a lipid-binding activity to be present. In vitro, the best indications for lipid binding came from the full-length protein, TbMORN1(1-15). TbMORN1(1-15) gave positive result on lipid blots (Fig. 2A), co-purified with PE (Fig. 3A), and its expression correlated with increased PE levels in bacteria (Fig. 3C). No co-purification of PE was seen with the apicomplexan MORN1 proteins despite 57% identity at an amino acid level. This lipid-binding activity therefore seems specific to TbMORN1. The elevation of PE levels seen in bacteria expressing TbMORN1(1-15) is highly unusual, and the authors do not currently have a good explanation for it that does not invoke lipid sequestration. If the PE co-purifying with recombinant TbMORN1(1-15) was just carry-over, then the lipid profile should resemble that of total bacterial cellular lipids, which is not the case. Furthermore, the bound PEs display a narrow range of aliphatic chains, suggesting specificity of binding/recognition.

Lipids in trypanosomes have, in general, much longer aliphatic chains than those in bacteria. In *T. brucei*, PE accounts for around 10-20% of total lipid, with aliphatic chains of 36:0 being the dominant isoform (Richmond, Gibellini et al., 2010). In *E. coli*, PE is the predominant zwitterionic lipid and accounts for around 80% of total lipid (Epand & Epand, 2009). As the

length of the side chains in bacterial lipids is predominantly 14 and 16 carbon atoms (Pramanik & Keasling, 1997), PE molecules with 16 carbon atoms chains are the most common. Notably, TbMORN1(1-15) overexpressed in *E. coli* co-purified with PE species of much greater length (30:0 – 36:2) (Fig. 3A), which was also reflected in increased production of these PE species in overexpressing bacteria (Fig. 3C). Taken together, these data strongly suggest the specificity of TbMORN1(1-15) towards PE species with aliphatic chains of a length characteristic for *T. brucei*.

Similarly, *in vivo*, it remains a possibility that the cytoskeleton-associated fraction is associated with the plasma membrane, potentially indirectly by binding to membrane-embedded partners. in some way. The data presented here show only that TbMORN1 and Ty1-TbMORN1, when not associated with cytoskeleton, are cytosolic rather than associated with the membrane/organelle fraction. The cause of the negative phenotypic effects resulting from overexpression of either untagged or Ty1-tagged TbMORN1 remains unclear.

As a final possibility, it might also be the case that a post-translational modification of TbMORN1 is essential for lipid binding, and is either not added in bacteria or lost during purification. Testing the activity of TbMORN1 expressed in a eukaryotic expression system or translated in vitro would be means of exploring this.

These data have clear implications for other MORN repeat proteins. Junctophilins do appear to bind lipids, but it is not clear if the MORN repeats are mediating this or just within the region/domain involved (Bennett et al., 2013). There is again evidence for the role of the N-terminal part of junctophilins mediating plasma membrane localisation, but this does not rule out an indirect association via protein-protein interactions being the primary driver (Takeshima et al., 2000). It now also appears that palmitolyation may play a significant role in junctophilin association with the plasma membrane (Jiang, Hu et al., 2019), Another recent paper on junctophilin-2 showed that upon cleavage by an endogenous protease, the N-terminal region translocates to the nucleus via a nuclear localisation signal and functions there as a transcription factor (Guo, Wang et al., 2018). The original paper on the junctophilin family noted the nuclear localisation of some truncations of junctophilin-1, but the significance of this was not appreciated at the time (Takeshima et al., 2000).

It is difficult to say whether this translocation and nuclear function is easier or harder to explain if the MORN repeats in the N-terminus are dedicated lipid-binding modules. Lipid-binding is known to cause conformational changes in junctophilin-2, and such conformational changes in reverse may well drive the protease-mediated dissociation, but again it is not clear if these changes are in any way due to the MORN repeats themselves. Additionally, it has been observed that the MORN repeats of ALS2 are not required for targeting to endosomes, although the exact domain mediating correct targeting is not agreed upon (Yamanaka, Vande Velde et al., 2003, Kunita et al., 2004).

In summary, the presence of MORN repeats in a protein should not be taken as indicative of lipid binding or lipid membrane binding without experimental evidence. Equally, evidence of binding from PIP strips alone should be interpreted with caution, given that the phospholipids are being presented to the protein in a non-physiological setting.

If MORN repeats are not lipid membrane modules by default, this raises the question of what they might actually be doing. The structural studies presented here make a case that one conserved function of MORN repeats is in homotypic interactions, and possibly also in higher-order assembly.

The three high-resolution structures described in this work are amongst the first canonical MORN repeat proteins to be detailed. Given that the structures are from representatives of two of the five eukaryotic supergroups - the excavates, and the SAR (stramenopiles, alveolates, Rhizaria) clade - this demonstrates how the fundamental structure of the MORN repeat has been conserved over evolutionary time.

All three MORN1 proteins analysed here formed tail-to-tail dimers via their C-termini with the polypeptide chains aligned in an antiparallel arrangement. The all-β structure of the proteins produces a twisting elongated structure, with a groove lined with aromatic side chains running longitudinally through it. While this manuscript was in preparation, Li et al. published a high-resolution structure of the MORN4/retinophilin protein in complex with the myosin 3a tail. Binding was mediated by this central groove, showing that it can be used for very high-affinity protein-protein interactions (Li et al., 2019).

The three high-resolution MORN1 crystal structures were consistent with the lower-resolution SAXS data obtained on the same proteins in solution (Fig. 10, Table S1). The apicomplexan MORN1 proteins appear capable of forming both extended and V-shaped conformations. The key residues mediating the dimer interface in the V-shaped form were defined, and shown to be conserved throughout the Apicomplexa. It thus seems likely that all apicomplexan MORN1 proteins can adopt these two conformations.

MORN1 proteins in Apicomplexa have been shown to be a key component of the basal complex, which undergoes a constriction event at the end of the cell division cycle (Gubbels, Vaishnava et al., 2006, Hu, 2008). It is therefore tempting to speculate that these extended and V-shaped conformations represent the pre- and post-constriction states of the MORN1 proteins in the basal complex. If so, this would constitute a remarkable rearrangement of the molecules in the dimer, which would move through about 145° (Supplementary Video 1).

The high-resolution structures also enabled a new structure-based consensus sequence for a MORN repeat to be defined (Fig. 7E). A single MORN repeat forms a β-hairpin, one of the conserved building blocks of structural biology, with a long loop attached. The structural basis for the conservation of individual residues within the MORN repeat has been defined (Fig. 7E, F). These structures strongly argue that an individual MORN repeat is longer than the 14 amino acids sometimes suggested. Despite the lower level of sequence conservation in the loop following the hairpin, a 23-amino acid length seems the most parsimonious definition from a structural perspective.

One interesting consequence of this redefinition of the repeat is that it suggests that full-length TbMORN1 begins with an incomplete repeat, a feature also noted in the MORN4/retinophilin

structure (Li et al., 2019). The (M)IYSEGE residues at the very N-terminus are predicted to form only a single β-strand rather than a complete hairpin (Fig. 7E). Li et al. suggested that this incomplete repeat could function as a capping element, but given that the first repeat in MORN1 is critical for oligomerisation, an alternative explanation is possible: the N-terminus of a second TbMORN1 molecule (itself encoding a single β-strand) could complete this hairpin through intermolecular interactions. Oligomerisation would thus be driven by a "split-MORN" mechanism where a complete hairpin is formed by the interaction of the N-termini of two proteins. Further work will be needed to test this hypothesis. The concluding data presented here suggest that TbMORN1 utilises this oligomerisation capacity to build mesh-like assemblies, which can reach considerable size in vitro (Fig. 10G). The biophysical properties of these meshworks, in particular their tensile strength, are likely to be another fruitful avenue of investigation.

MATERIALS AND METHODS:

Antibodies and other reagents

All custom antibodies have been described previously. The rabbit anti-TbMORN1 were made for a previous project (Morriswood et al., 2013). The mouse monoclonal anti-Ty1 (BB2) antibodies were a gift from Cynthia He (University of Singapore) (Bastin, Bagherzadeh et al., 1996). The mouse monoclonal anti-PFR1,2 antibodies (L13D6) were a gift from Keith Gull (University of Oxford) (Kohl, Sherwin et al., 1999). The anti-BiP antibodies were a gift from Jay Bangs (University at Buffalo) (Bangs, Uyetake et al., 1993). The rabbit anti-GFP antibodies were a gift from Graham Warren (MRC Laboratory for Molecular Cell Biology) (Pelletier, Stern et al., 2002). The following antibodies were obtained from commercial sources: anti-strep tag StrepMAB-Classic (iba), HRP-conjugated anti-mouse (ThermoFisher Scientific), anti-GST (Santa Cruz Biotechnology). Defatted BSA was purchased from Sigma-Aldrich.

Cloning and mutagenesis of expression constructs

The 1077 bp TbMORN1 open reading frame (ORF) (UniProt accession no. Q587D3; TriTrypDB database accession no. Tb927.6.4670) was amplified by PCR from genomic DNA obtained from *Trypanosoma brucei brucei* strain Lister 427 and ligated into vector pETM-13 encoding a Strep-tag at the 3' end of the insert. TbMORN1 truncations were generated using this construct as a template by ligase-independent cloning (Doyle, 2005). The sequences for TbMORN1(2-15) and TbMORN1(1-14) were additionally ligated into the pCoofy12 vector encoding a 3C protease-cleavable N-terminal Twin-Strep-tag (Scholz, Besir et al., 2013) by sequence and ligation-independent cloning (Li & Elledge, 2012). Mutagenesis constructs were generated by standard methods using the pre-existing pCoofy12_TbMORN1(2-15) construct as the template (Hemsley, Arnheim et al., 1989). All primer sequences are available upon request. For LiMA experiments, the construct encoding the EGFP-TbMORN1(2-15) was cloned in a two-step procedure by sequence and ligation-independent cloning followed by Gibson assembly using pCoofy12_TbMORN1(2-15) and the pEGFP-C1 vector (Gibson,

Young et al., 2009). The 1092 bp TgMORN1 (UniProt accession no. Q3S2E8) and 1095 bp PfMORN1 (UniProt accession no. Q8IJ93) ORFs were amplified by PCR from genomic DNA. Truncations of the TgMORN1 and PfMORN1 constructs were generated using ligase-independent cloning. The TgMORN1 constructs were additionally ligated into the pET14 vector encoding a 3C protease-cleavable N-terminal His10-tag. The PfMORN1 constructs were additionally ligated into the pCoofy32 encoding a 3C protease-cleavable N-terminal His10-tag and C-terminal OneStrep-tag.

Recombinant protein expression and purification

Rosetta 2 (DE3)pLysS bacterial cells transformed with the required expression plasmids were grown at 37 °C with shaking in the presence of the appropriate antibiotics. Large scale expression was carried out either in Luria-Broth or in auto-induction (ZY) medium (Studier, 2005), with 500 ml media being inoculated with 3-5 ml of pre-cultured cells. Cells in Luria-Broth were grown to an OD_{600} ~ 0.8 – 1.0, after which 50 μM IPTG was added to induce recombinant protein expression. The cells were then incubated at lower temperature (overnight, 20 °C). The cells were then harvested by centrifugation (5000 x g, 30 min), and either lysed immediately or stored at -80 °C. For purification, the cells were resuspended in lysis buffer (50 mM Tris-HCl pH 8.5, 200 mM NaCl, 5% (w/v) glycerol, 1 mM DTT, protease inhibitor cocktail, benzonase). The pellet emulsions were first homogenised by mixing on ice using a T 10 basic Ultra-Turrax dispersing instrument (IKA), and lysis was accomplished using a single cycle in a cell disruptor (Constant systems Ltd), with the pressure set to 1.35 kPa. Lysates were clarified by centrifugation (18,000 x g, 45 min, 4 °C), and a two-step fast protein liquid chromatography (FPLC) purification protocol using an ÄKTA Protein Purification System (GE Healthcare Life Sciences) at 8 °C was then followed to obtain the recombinant protein. The supernatants were applied to two connected Strep-Trap HP 5 ml columns packed with Strep-Tactin ligand immobilized in an agarose matrix (GE Healthcare Life Sciences) (GE Healthcare Life Sciences) previously equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.5, 200 mM NaCl, 2% (w/v) glycerol, 1 mM DTT). Flow speed was adjusted to 2.5 ml/min. When 100% step gradient of elution buffer (equilibration buffer plus 2.5 mM D-desthiobiotin) was applied, the bound proteins were eluted in a single chromatographic peak. Selected peak fractions were examined by SDS-PAGE for protein content and purity, pooled accordingly, and concentrated in Amicon Ultra centrifugal filter units (MerckMillipore, various pore sizes) according to the manufacturer's instructions. These affinity-purified protein concentrates were then applied to a previously equilibrated HiLoad 16/600 Superdex 200 pg column (GE Healthcare Life Sciences) packed with dextran covalently bound to highly cross-linked agarose, enabling separation of proteins with MW in the range of 10 – 600 kDa. Flow speed was adjusted to 1 ml/min and fractions of 1.5 ml were collected. Fractions corresponding to the targeted chromatographic peak were examined for protein content by SDS-PAGE, pooled accordingly to their purity, concentrated, and stored at -80 °C until use.

Limited proteolysis

Purified recombinant His-TbMORN1(1-15) at 1 mg/ml was separately incubated with three proteases (α-chymotrypsin, trypsin, and proteinase K) in 20 mM Tris-HCl pH 8.5, 200 mM NaCl, 2% glycerol, 0.2 mM $CaCl_2$ (15 min, RT). The proteases were used at dilutions of 1:100

- 1:2000. The reactions were stopped by the addition of SDS-Coomassie sample loading buffer for analysis by gel electrophoresis. The indicated protein bands were extracted from the gel and subjected to mass spectrometry analysis.

Size-exclusion chromatography coupled to multi-angle light scattering (SEC MALS)

The MW and oligomeric state of purified proteins were verified by size exclusion chromatography (SEC) coupled to multi-angle light scattering (MALS), using a Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences). Up to five protein samples of 100 μl were dialysed against 2x 1 L of freshly-prepared, degassed gel filtration buffer (20 mM Tris-HCl pH 8.5, 200 mM NaCl, 2% (w/v) glycerol, 0.5 mM DTT) (overnight, 4 °C). Gel filtration buffer was also used for overnight equilibration of the column and in the subsequent measurements. Protein samples were clarified by centrifugation using a TLA-55 rotor in an Optima MAX-XP table top ultracentrifuge (Beckman Coulter) (90,720 x g, 30 min, 4 °C). 100 μl of 2 – 4 mg/ml protein samples were applied to a column using the 1260 Infinity HPLC system (Agilent Technologies) coupled to a MiniDawn Treos detector (Wyatt Technologies) with a laser emitting at 690 nm. An RI-101 detector (Shodex) was used for refractive index determination and the Astra 7 software package (Wyatt Technologies) for data analysis. No correction of refractive index was necessary due to the 2% (w/v) glycerol content in the buffer.

Circular dichroism (CD)

Far-UV CD was used both for measurement of secondary structure and for validation of the thermostability of TbMORN1 constructs. To avoid the absorption of Tris and NaCl below 180 nm (Kelly, Jess et al., 2005), three protein samples of 100 μl were first dialysed against 2x 1 L of dialysis buffer (20 mM NaH_2PO_4, 20 mM Na_2HPO_4, 200 mM NaF) (overnight, 4 °C). The pH 8.0 was adjusted by mixing the mono- and dibasic sodium phosphate solutions. The dialysed proteins were clarified by centrifugation in an Optima MAX-XP tabletop ultracentrifuge (Beckman Coulter Life Sciences) (90,720 x g, 30 min, 4 °C). The concentration of protein samples was adjusted to 0.25 mg/ml. CD measurements were carried out in a quartz cuvette with an optical path length of 0.5 mm (Stana Scientific Ltd) using a Chirascan Plus spectrophotometer (Applied Photophysics) equipped with the Chirascan-plus DMS software package. The CD profiles for secondary structure calculations were obtained at RT in the range of 190 – 260 nm. Further analysis was carried out using the BeStSel server, which is specialised in the analysis of CD data from proteins rich in β-strands (Micsonai, Wien et al., 2015, Micsonai, Wien et al., 2018). Data were converted to $\Delta\varepsilon$ ($M^{-1}cm^{-1}$) and uploaded to the BeStSel online server. Melting experiments were performed in the range of 190 – 260 nm, 20 – 80 °C, with a temperature ramp of 0.8 °C/min. Data were analysed with Global 3 software package.

Chemical cross-linking coupled to mass spectrometry (XL-MS)

For chemical cross-linking with EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) or BS^3 (bis(sulfosuccinimidyl)suberate), 200-300 μl of approximately 30 μM TbMORN1 sample was dialysed twice against 1 L of EDC buffer (20 mM MES-NaOH pH 6.8, 200 mM NaCl) or BS^3 buffer (20 mM HEPES-NaOH pH 8.0, 200 mM NaCl) (overnight, 4 °C).

Following dialysis, the protein was clarified by centrifugation (90,720 x g, 30 min, 4 °C) using a TLA-55 rotor in an Optima MAX-XP tabletop ultracentrifuge (Beckman Coulter Life Sciences). EDC (ThermoFischer Scientific) was first equilibrated to RT and then a stock solution in EDC buffer was prepared. 1.3 μM TbMORN1 previously dialysed in EDC buffer was mixed with 0, 200, and 400 μM EDC (all final concentrations) in a total volume of 40 μl. After a 30 min incubation at RT, 10 μl of SDS loading buffer was added and the mixtures were further denatured by heating (95 °C, 10 min). The experiments with BS^3 were carried out identically except that BS^3 buffer and 3.4 μM TbMORN1 were used, and the incubation time was 120 min. Samples were separated by SDS-PAGE, stained with Coomassie dye, and selected bands corresponding to monomers and dimers cross-linked with 400 μM cross-linker were excised and subjected to enzymatic digestion and subsequent mass spectrometry analysis. Coomassie Brilliant Blue-stained excised bands were destained with a mixture of acetonitrile and 50 mM ammonium bicarbonate (ambic), in two consecutive steps (each 10 min, RT). The proteins were reduced using 10 mM DTT in 50 mM ambic for (30 min, 56 °C), alkylated with 50 mM iodoacetamide in 30 mM ambic in the dark (30 min, RT), and digested with trypsin (Promega, mass spectroscopy grade) (overnight, 37 °C). The reaction was stopped using 10% (v/v) formic acid and extracted peptides were desalted using C18 Stagetips (Rappsilber, Mann et al., 2007). Peptides were analysed on an UltiMate 3000 HPLC RSLCnano system coupled to a Q Exactive HF mass spectrometer, equipped with a Nanospray Flex ion source (all Thermo Fisher Scientific). Peptides were loaded onto a trap PepMap 300 C18 column of dimensions 5 mm x 300 µm i.d., packed with 5 µm particles with a pore size of 100 Å (Thermo Fisher Scientific, cat. no. 164718) and separated on an analytical C18 100 column of dimensions 500 mm x 75 µm i.d., packed with 2 µm particles with a pore size of 100 Å (Thermo Fisher Scientific, cat. no. 164942), applying a linear gradient from 2% to 40% solvent B (80% acetonitrile, 0.1% formic acid) at a flow rate of 230 nl/min over 120 min. The mass spectrometer was operated in a data-dependent mode at high resolution of both MS1 and MS2 level. Peptides with a charge of +1, +2 or of a higher than +7, were excluded from fragmentation. To identify cross-linked peptides, the spectra were searched using pLink software v1.23 (Yang, Wu et al., 2012). Q Exactive HF raw-files were pre-processed and converted to mgf-files using pParse (Yuan, Liu et al., 2012). The MaxQuant database (Tyanova, Temu et al., 2016) was used to search the spectra for the most abundant protein hits. Carbamidomethylation of cysteine and oxidation of methionine residues were set as variable modifications. Trypsin was set as an enzyme specificity, and EDC or BS^3 was set as a cross-linking chemistry. In case of EDC, aspartic and glutamic acid residues, as well as C-termini of proteins, were allowed to be linked with lysine residues. In the case of BS^3, lysine residues and N-termini of proteins were allowed to be linked with lysine residues, N-termini of proteins, as well as to serine, threonine and tyrosine residues. Search results were filtered for 1% FDR (false discovery rate) on the PSM (number of peptide-spectrum matches) level and a maximum allowed precursor mass deviation of 5 ppm. To remove low quality PSMs, an additional e-Value cutoff of < 0.001 was applied. In order to distinguish intra- from inter-molecular chemical cross-links, results from monomers and dimers were compared. A potential inter-molecular cross-link must have shown the following criteria: (1) minimally 3 peptide PSMs in dimer and (2) minimally 3-times more PSMs in dimer than in monomer.

Protein-lipid overlay assays (PIP strips)

PIP strips were purchased from Echelon Biosciences (cat. no. P-6001). PBS-T (PBS, 0.1% TWEEN-20) was used as a general buffer. Purified recombinant TbMORN1(1-15) was clarified by centrifugation (20,817 x g, 20 min, 4 °C) prior to use. The PIP strips were blocked using blocking buffer (3% (w/v) defatted BSA, PBS-T) (60 min, RT) and then incubated with 5 μg/ml of TbMORN1 in 10 ml of blocking solution (60 min, RT). After three washes with PBS-T, the membranes were overlaid with anti-strep antibodies diluted in blocking solution (60 min, RT). After a further three PBS-T washes, the membranes were overlaid with HRP-conjugated secondary antibodies (60 min, RT). The membranes were then washed three times with PBS-T and visualised by ECL (Western Blotting substrate, Thermo Fisher) using a Fusion FX imager (Vilber Lourmat). All binding and wash steps were carried out with gentle agitation of the membranes. For the positive control, the PIP strip was overlaid with GST-tagged PLC-δ1 PH domain (Echelon Biosciences) and mouse monoclonal anti-GST antibodies (Santa Cruz Biotechnology) were used.

Fluorescence anisotropy

Stocks of BODIPY TMR-labelled PI C6, PI(4)P C6, PI(3,4)P_2 C6, PI(3,5)P_2 C6, PI(4,5)P_2 C6 and PI(4,5)P_2 C16 were sonicated (5 min, RT) in a sonication bath, and in parallel with purified recombinant TbMORN1(2-15), TbMORN1(7-15), TbMORN1(10-15), were clarified by centrifugation (20,817 x g, 20 min, 4 °C). The concentrations of lipid stocks were determined with a Hitachi U-3501 UV-VIS spectrophotometer, using quartz absorbance cuvettes and an optical path length of 10 mm (Hellma Analytics). For this purpose, the maximum absorbance of BODIPY TMR dye at λ = 544 nm was measured; its extinction coefficient ε = 60,000 $cm^{-1}M^{-1}$. The total volume of each respective sample was 110 μl. The concentration of selected TbMORN1 constructs was varied from 0 to 35 μM (or more), while the concentration of added lipid was kept constant at 0.1 μM. Measurements were performed on a Perkin Elmer LS50B fluorimeter in quartz cuvettes with an optical path length of 10 x 2 mm (Hellma Analytics). To ensure a constant temperature of 20 °C in the measured sample, the measurement cell was connected to a water bath. The parameters for fluorescence anisotropy (r) measurements were: λ_{ex} = 544 nm and aperture of excitation slit = 15 nm; λ_{em} = 574 nm and aperture of emission wavelength = 20 nm; time of integration = 1 s and T = 20 °C. The grating factor (G factor), which provides grating correction for the optical system, was determined on samples with exclusively 0.1 μM lipid and kept constant during measurement of each concentration series. Triplicates of each protein concentration point were measured and afterwards averaged using Excel software. Graphs were drawn and fitted in SigmaPlot ver. 13.0. The equation used for fitting was a four parameters logistic curve where:

$$y = \min + \frac{(\max - \min)}{1+(\frac{x}{EC50})^{-Hillslope}},$$

Options were set to default; initial parameters values, as well as parameters min, max, EC50, Hillslope, were selected automatically, parameter constraints were max > min and EC50 > 0, number of iterations was 200, and tolerance was kept at $1e^{-10}$. The reduced chi-square method was used to compute parameters' standard errors. Experimental r values of respective 0.1 μM BODIPY TMR-lipid and protein-TMR BODIPY-lipid mixtures were compared with theoretical values obtained by Perrin equation.

Mass spectrometry analysis of extracted lipids

Lipid extractions from purified recombinant full length and truncated TbMORN1 were achieved by three successive vigorous extractions with ethanol (90% v/v) according to a published protocol (Fyffe, Alphey et al., 2006). The pooled extracts were dried using N_2 gas in a glass vial and re-extracted using a modified Bligh and Dyer method (Richmond et al., 2010). For whole *E. coli* lipid extracts, cells were washed with PBS and extracted following the modified Bligh and Dyer method. All extracts were dried under N_2 gas in glass vials and stored at 4 ^{o}C. Extracts were dissolved in 15 µl of choloroform:methanol (1:2) and 15 µl of acetonitrile:propan-2-ol:water (6:7:2) and analysed with an Absceix 4000 QTrap, a triple quadrupole mass spectrometer equipped with a nano-electrospray source. Samples were delivered using a Nanomate interface in direct infusion mode (~125 nl/min). Lipid extracts were analysed in both positive and negative ion modes using a capillary voltage of 1.25 kV. MS/MS scanning (daughter, precursor and neutral loss scans) were performed using nitrogen as the collision gas with collision energies between 35-90 V, allowing lipid structure assignments.

Preparation of sucrose-loaded vesicles (SLVs) and pelleting assay

To generate synthetic SLVs, lipids reconstituted in $CHCl_3$ were mixed in the following ratio: 30% DOPC; 35% DOPE; 15% DOPS; 20% cholesterol. 5 mol % $PI(4,5)P_2$ was added in place of 5 mol% DOPE in the $PI(4,5)P_2$-containing liposomes. Lipids were extracted from bloodstream form *T. brucei* according to an established protocol (Bligh & Dyer, 1959). Briefly, mid-log phase cells were harvested by centrifugation (750 x g, 10 min, RT), washed once with PBS, then resuspended in 100 µl PBS and transferred to a glass tube. 375 µl of 1:2 (v/v) $CHCl_3$:MeOH was added and the mixture was vortexed (20 s) and then incubated with continuous agitation (15 min, RT). A further 125 µl $CHCl_3$ was then added to make the mixture biphasic, and following brief vortexing 125 µl ddH_2O was added. The mixture was vortexed again and then separated by centrifugation (1000 x g, 5 min, RT). The lower organic layer was then transferred to a new glass vial, dried under a nitrogen stream, and kept at 4 °C until use. For the preparation of SLVs from trypanosomal lipids the lyophilized lipids (extract from $8x10^7$ cell equivalents) were reconstituted in 50 µl $CHCl_3$. 6 µM Rhodamine B dihexadecanoyl phosphoethanolamine (Rh-DHPE) was added to all lipid mixtures to facilitate the visualisation of the SLVs. The lipid mixtures were dried under a nitrogen stream, and the lipid films hydrated in 20 mM HEPES pH 7.4, 0.3 M sucrose. The lipid mixtures were subjected to 4 cycles of freezing in liquid nitrogen followed by thawing in a sonicating water bath at RT. The vesicles were pelleted by centrifugation (250,000 x g, 30 min, RT) and resuspended in 20 mM HEPES pH 7.4, 100 mM KCl to a total lipid concentration of 1 mM. SLVs were incubated with 1.5 µM purified TbMORN1(2-15) at a 1:1 ratio (30 min, RT). To separate soluble and SLV-bound TbMORN1(2-15), the vesicles were pelleted by centrifugation (8,700 x g, 30 min, RT), and equal volumes of supernatant and resuspended pellet were separated by SDS-PAGE and analysed by Coomassie staining.

Liposome microarray assay (LiMA)

LiMA (Saliba et al., 2014) was performed in the lab of Anne-Claude Gavin (EMBL Heidelberg, Germany) according to the standard protocol (Saliba et al., 2016). TbMORN1(2-15) tagged N-terminally with EGFP and two positive controls, PLCδ1-PH and Lactadherin-C2, both fused to superfolder GFP (sfGFP), were applied to microarrays printed with different signalling lipids. In brief, lipids of interest were combined with the carrier lipid DOPC, PEGylated PE, and PE labelled with Atto 647 dye (PE-Atto 647, 0.1 mol%). Lipid mixtures containing 2, 5, and 10 mol% of the signalling lipid were spotted onto a thin agarose layer (TAL). The agarose layers were hydrated using buffer A (20 mM Tris-HCl pH 8.5, 200 mM NaCl) and vesicles formed spontaneously. Efficiency of liposome formation was verified by fluorescence microscopy. The protein was diluted to 7 µM in buffer A and 40 µl was applied to each array. Microarrays were incubated (20 min, RT) and subsequently washed three times with 40 µl of buffer A. Chips were analysed by automated fluorescence microscopy. Positions of liposomes were determined by tracking the fluorescence of PE-Atto 647 and images were taken for 3 ms and 5 ms exposure times. In parallel, the fluorescence of EGFP was determined for 1, 5, 10, 30, 75, 100, 200 and 300 ms exposures. Images were processed using CellProfiler and CPAnalyst. Only EGFP signals that overlapped with Atto 647 signals were taken into account. Normalised binding intensity (NBI) was calculated as the ratio between EGFP and Atto 647 fluorescence, normalised by exposure time. Three microarrays were examined, carrying liposomes with the following signalling lipids; PIP-chip: DOPA, DOPE, DOPI, DOPS, DODAG, cardiolipin, BMP, DOPI(4,5)P_2, DOPG; GLP-chip: ceramide C16, ceramide(1)P C16, ceramide(1)P C18, S(1)P, S, SM, DOPI(4,5)P_2, DOPS; SL-chip: DOPI(3)P, brain PI(4)P, DOPI(5)P, DOPI(3,4)P_2, DOPI(3,5)P_2, brain PI(4,5)P_2, DOPI(3,4,5)P_3, DOPS and cholesterol. Each microarray was performed in triplicate.

Cell culture and cell line generation

Bloodstream form *T. brucei* cells were maintained in HMI-9 media supplemented with 10% foetal bovine serum (Sigma-Aldrich, St. Louis, USA) at 37 °C and 5% CO_2 in cell culture flasks with filter lids (Greiner). For overexpression studies, 427 strain "single marker" cells - which express T7 RNA polymerase and the Tetracycline repressor protein, both maintained under 2.5 µg/ml G418 selection - were used (Wirtz, Leal et al., 1999). For optimisation of digitonin extraction conditions, the GFP^{ESPro}-221^{ES}.121^{tet} cell line, which constitutively expresses GFP from the VSG expression site, was used (Batram et al., 2014). Constructs for overexpression of Ty1-tagged TbMORN1 and untagged TbMORN1 were obtained by cloning the required ORFs into the pLEW100v5-HYG plasmid; the identity of the inserts was verified by DNA sequencing followed by BLAST analysis against the TbMORN1 ORF (Tb927.6.4670). The plasmids were linearised by NotI digestion, and plasmid DNA was purified by ethanol precipitation. Linearisation was verified using agarose gel electrophoresis. Stable cell lines were generated by using 20 µg of the linearised plasmids to transfect ~$3x10^7$ "single marker" cells in transfection buffer (90 mM Na_2PO_4, 5 mM KCl, 0.15 mM $CaCl_2$, 50 mM HEPES pH 7.3) using the X-001 program of an Amaxa Nucleofector II (Lonza, Switzerland) ((Burkard, Fragoso et al., 2007, Schumann Burkard, Jutzi et al., 2011). Clones were obtained from the transfected cells by limiting dilution under 5 µg/ml hygromycin selection. Clones were verified as described in the manuscript text.

Growth curves and BigEye cell counts

22 ml cells at a defined starting concentration were divided into two flasks of 10 ml each, and overexpression of the ectopic transgene was initiated in one flask by the addition of tetracycline to a final concentration of 1 μg/ml. For overexpression of Ty1-TbMORN1, a starting concentration of $1x10^4$ cells/ml was used, and the cells were split and reseeded at this concentration after 48 h. For overexpression of untagged TbMORN1, a starting concentration of $1x10^3$ cells/ml was used with no reseeding. Tetracycline was refreshed every 24 h in both cases. Population density was measured using a Z2 Coulter Counter (Beckman Coulter, Krefeld, Germany) at the indicated timepoints. For quantification of BigEye cell incidence at the indicated timepoints, the cultures were briefly agitated to mix the cells, which were then allowed to settle for 30 min. The culture flasks were then examined directly using an inverted phase contrast microscope (Leitz Labovert) and a 10x objective lens. Three fields of view were chosen at random for each flask and the number of normal and BigEye cells manually quantified, using higher magnification where necessary. Given the low magnification used, the numbers presented are likely to be underestimates of the true incidence.

Immunoblotting

To obtain whole cell lysates, cell concentration was measured using a Z2 Coulter Counter, and a defined volume was then transferred to 15 ml Falcon tubes. The cells were pelleted by centrifugation (750 x g, 10 min, RT), resuspended in 1 ml PBS, and transferred to microfuge tubes. The cells were again pelleted (1800 x g, 2 min, RT), and the cell pellet then directly resuspended in SDS loading buffer to a final concentration of 2×10^5 cells/μl. The lysates were heated (95 °C, 10 min) before use. Lysates were separated by SDS-PAGE ($1.4x10^6$ cells/lane in a 15-well gel of 1.0 mm thickness), and the proteins then transferred to nitrocellulose membranes. The membranes were blocked in blocking buffer (10% milk, PBS, 0.3% TWEEN-20) (30 min, RT), and then incubated with the indicated primary antibodies in blocking buffer (1 h, RT). The membranes were washed three times in PBS-T (PBS, 0.3% TWEEN-20), and were then incubated with IRDye-conjugated secondary antibodies in PBS-T (1 h, RT). After a further three washes in PBS-T the membranes were briefly dried between sheets of filter paper and then imaged using an Odyssey CLx (LI-COR Biosciences, Bad Homberg, Germany). Processing and quantification was carried out using ImageStudioLite software (LI-COR Biosciences).

Immunofluorescence microscopy

Cell concentration was measured using a Z2 Coulter Counter, and 10^6 cells per coverslip were taken. The cells were transferred to 15 ml Falcon tubes, and fixed directly in media by the addition of paraformaldehyde solution to a final concentration of 4% (37 °C, 20 min). 10 ml trypanosome dilution buffer (TDB; 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM glucose) was then added, and the cells were pelleted by centrifugation (750 x g, 10 min, RT). The supernatant was removed, the cell pellet was resuspended in 500 μl TDB, and the cells were transferred to poly-L-lysine-coated coverslips in a 24-well plate. The cells were attached to the coverslips by centrifugation (750 x g, 4 min, RT), and the cells

were permeabilised using a solution of 0.25% TritonX-100 in PBS (5 min, RT). The cells were washed with PBS, blocked using a solution of 3% BSA in PBS (30 min, RT), and sequentially incubated with primary and secondary antibodies diluted in PBS (1 h, RT for each) with three PBS wash steps after each incubation. After the final wash, the coverslips were rinsed in ddH_2O, excess fluid removed by wicking, and mounted on glass slides using Fluoromount-DAPI (Southern Biotech). For analysis of detergent-extracted cytoskeletons, cells were washed using TDB and attached to poly-L-lysine coverslips as described above. The cells were detergent-extracted using extraction buffer (0.5% IGEPAL, 0.1 M PIPES-NaOH pH 6.9, 2 mM EGTA, 1 mM $MgSO_4$, 0.1 mM EDTA, cOmplete protease inhibitors [Roche]) (5 min, RT), washed three times with extraction buffer, and then fixed with ice-cold MeOH (-20 °C, 30 min). Blocking, antibody incubation steps, and mounting were as described above. All liquid handling was carried out using a P1000 micropipette, and pipetting was done as gently as possible to minimise shear forces. The coverslips were imaged using a Leica DMI6000B inverted microscope equipped with a Leica DFC365 camera and a 100x oil objective lens (NA1.4) and running Leica Application Suite X software. The same exposure times were used for acquisition of +/-Tet samples, and 40 z-slices of 0.21 μm thickness were taken per field of view. Image processing was carried out using ImageJ. Maximum intensity z-projections are shown.

Fractionation

Cell concentration was measured using a Z2 Coulter Counter, and an equal number of cells (~$2.5x10^7$ per experiment) was taken from the control (-Tet) and overexpression (+Tet) samples and transferred to 50 ml Falcon tubes. The cells were pelleted by centrifugation (750 x g, 10 min, 4 °C), and the cell pellets then resuspended in 1 ml TDB and transferred to microfuge tubes. The cells were pelleted by centrifugation (1800 x g, 2 min, 4 °C), and then resuspended in 200 μl extraction buffer (see Immunofluorescence microscopy section for composition). After a short incubation (15 min, RT, orbital mixer), a 5% (10 μl) input sample was taken, and the mixtures separated by centrifugation (3400 x g, 2 min, 4 °C) into detergent-soluble (cytoplasmic) supernatant and detergent-insoluble (cytoskeleton) pellet fractions. The supernatant was transferred to a fresh microfuge tube, its exact volume noted, and a 5% sample taken. The tube containing the pellet was centrifuged a second time (3400 x g, 2 min, 4 °C) in order to bring down material sticking to the tube wall; this second supernatant was discarded. The pellet was resuspended in 200 μl extraction buffer and a 5% sample (10 μl) taken. SDS loading buffer was added to the input, supernatant, and pellet samples to a final volume of 20 μl, and denaturation assisted by heating (95 °C, 10 min). Equal fractions were loaded onto polyacrylamide gels, separated by SDS-PAGE, and analysed by immunoblotting. In the exemplary blot shown (Fig. 5G, each sample is a 4.5% fraction, equivalent to ~10^6 cells in the Input fraction). For optimisation of extraction conditions using digitonin, essentially the same protocol was followed except that ultra-pure digitonin (Calbiochem) in TDB buffer was used and incubations were carried out at 24 °C in a heating block. For the two-step digitonin/IGEPAL fractionations (Fig. 8), cells were pelleted by centrifugation (750 x g, 10 min, RT), resuspended in 1 ml TDB, transferred to microfuge tubes, and pelleted again (750 x g, 3 min, RT). The cell pellet was resuspended in 400 μl of 40 μg/ml digitonin in TDB and extracted (25 min, 24 °C), after which a 5% input sample was taken. The mixture was then separated

by centrifugation (750 x g, 5 min, RT) and 320 µl of the cytosolic fraction (SN1) transferred to a fresh tube and a 5% sample was taken. The cell pellet was then resuspended with 1 ml TDB and the extracted cells again pelleted by centrifugation (750 x g, 5 min, RT). The extracted cells were then resuspended in 400 µl extraction buffer (see Immunofluorescence section above for composition) and incubated (24 °C, 15 min, heating block with shaker). After the incubation, a 5% sample (P1) was taken, and the extracted cells pelleted by centrifugation (3400 x g, 2 min, RT). 320 µl of the supernatant (SN2) was transferred to a fresh microfuge tube and a 5% sample was taken. The pellet was resuspended in 1 ml TDB and centrifuged again (750 x g, 5 min, RT). The pellet (P2) was then resuspended in 400 µl extraction buffer. SDS loading buffer was added to the 5% samples (I, SN1, SN2, P1, P2) to a final volume of 40 µl. Samples were analysed by immunoblotting as detailed above.

Crystallisation

Crystallisation of TbMORN1(7-15), TgMORN1(7-15) and PfMORN1(7-15) was performed at 22 °C using a sitting-drop vapour diffusion technique and micro-dispensing liquid handling robots (Phoenix RE (Art Robbins Instruments) and Mosquito (TTP labtech)). In the case of TbMORN1(7-15), crystals only appeared from reductively methylated protein, using a standard protocol (Walter, Meier et al., 2006). The best diffracting crystals of TbMORN1(7-15) were grown at a protein concentration of 3.5 mg/ml in the following conditions: 0.166 M Tris-HCl pH 8.8, 0.15 M $MgCl_2$, 0.45 M KI, 24% PEG 2000 MME, and 4% glycerol. The tetragonal crystals of TgMORN1(7-15) were obtained at a protein concentration of 10 mg/ml in the following conditions: 0.1 M Tris-HCl pH 8.2, 15% PEG 3350, 0.2 M NaCl. The diffracting crystals of both selenomethionine-containing crystals and native crystals of PfMORN1(7-15) were obtained at a protein concentration of 8 mg/ml in the conditions “B11” from the Morpheus II crystallisation screen (Molecular Dimensions): 2 mM divalents mix (0.5 mM $MnCl_3$, 0.5 mM $CoCl_2$, 0.5 mM $NiCl_2$, 0.5 mM $Zn(OAc)_2$, 0.1 M Buffer System 6, pH 8.5 (Gly-Gly, AMPD), and 50% precipitation Mix 7 (20 % PEG 8000, 40% 1,5-Pentanediol). The crystals were flash cooled in liquid nitrogen prior to data collection.

X-ray diffraction data collection and crystal structure determination

Initially, the structure of PfMORN1(7-15) was determined using the single-wavelength anomalous diffraction (SAD) method. The selenomethionine dataset was collected at the beamline ID29 (ESRF, Grenoble) at 100K at the peak of selenium using a wavelength of 0.979 Å. The data frames were processed using the XDS package (Kabsch, 2010), and converted to mtz format with the program AIMLESS (Winn, Ballard et al., 2011). The apo-PfMORN1(7-15) structure was solved using single anomalous diffraction with AUTOSOL software from the PHENIX package. The structures of TgMORN1(7-15) and TbMORN1(7-15) were then solved using the molecular replacement program PHASER (McCoy, Grosse-Kunstleve et al., 2007) with the atomic coordinates of PfMORN1(7-15) as a search model. The structures were then refined with REFMAC and Phenix Refine and rebuilt using Coot (Murshudov, Vagin et al., 1997, Emsley & Cowtan, 2004, Adams, Afonine et al., 2010). The structures were validated and corrected using the PDB_REDO server (Joosten, Long et al., 2014). The figures were produced using Pymol and Chimera software. Coordinates have been deposited in the protein

data bank (accession codes 6T4D, 6T4R, 6T68, 6T69, 6T6Q). Data collection and refinement statistics are reported in Table 3.

Small angle X-ray scattering (SAXS)

Synchrotron radiation X-ray scattering from various MORN constructs in solution were collected at different synchrotron facilities (Table S3). TgMORN1(7-15), TbMORN1(7-15) as well as PfMORN1(7-15) were collected at the EMBL P12 beamline of the storage ring PETRA III (DESY, Hamburg, Germany) (Blanchet, Spilotros et al., 2015). Images were recorded using a photon counting Pilatus-2M detector at a sample to detector distance of 3.1 m and a wavelength (λ) of 1.2 Å covering the range of momentum transfer $0.01 < s < 0.5$ Å^{-1}; with $s = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle. To obtain data from a monodisperse sample from TgMORN1(7-15) and TbMORN1(7-15), a size exclusion chromatography column directly coupled to the scattering experiment (SEC-SAXS) was employed. The parallel collection of UV and light scattering data allowed the protein to be monitored while it eluted from the column (Graewert, Franke et al., 2015). Throughout the complete chromatography process, 1 s sample exposures were recorded. As mobile phase, the various buffers were used: TbMORN1(7-15): 20mM Tris-HCl pH 8.5, 200 mM NaCl, 2% (v/v) glycerol, 1 mM DTT. TgMORN1(7-15): 20 mM Tris-HCl pH 7.5, 100 mM NaCl. 100 µl of purified sample (3.8 mg/mL TbMORN1(7-15) and 2.6 mg/ml TgMORN1(7-15)) were injected onto a Superdex 200 10/300 (GE Healthcare) column and the flow rate was set to 0.5 ml/min. SAXS data were recorded from macromolecular free fractions corresponding to the matched solvent blank. PfMORN1(7-15) was measured in batch mode from a concentration series spanning 1-8 mg/ml. 20 mM Tris-HCl pH 7.5, 100 mM NaCl buffer was measured for background subtraction. As a concentration dependent increase in size was detectable further analysis were based solely on the data collected at 1 mg/ml. In a similar manner as described above, TbMORN1(2-15) data were collected at ESRF BM29 beamline (Pernot, Round et al., 2013) in SEC-SAXS mode with the setup described in by Brennich et al. (Brennich, Round et al., 2017). SAXS data from the run were collected at a wavelength of 0.99 Å using a sample-to-detector (PILATUS 1M, DECTRIS) distance of 2.867 m. Here too, a Superdex 200 10/300 (GE Healthcare) column was used as well as 20 mM Tris-HCl pH 7.5, 100 mM NaCl as mobile phase. 100 µl of 5.8 mg/ml TbMORN1(2-15) were injected. Data reduction to produce final scattering profiles of dimeric MORN1 constructs were performed using standard methods. Briefly, 2D-to-1D radial averaging was performed using the SASFLOW pipeline (Franke, Petoukhov et al., 2017). For data collected at ESRF EDNA pipeline (Brennich, Kieffer et al., 2016) was used. CHROMIXS was used for the visualisation and reduction of the SEC-SAXS datasets (Panjkovich & Svergun, 2018). Aided by the integrated prediction algorithms in CHROMIXS the optimal frames within the elution peak and the buffer regions were selected. Single buffer frames were then subtracted from sample frames one by one, scaled and averaged to produce the final subtracted curve. The indirect inverse Fourier transform of the SAXS data and the corresponding probable real space-scattering pair distance distribution (p(r) versus r profile) of the various MORN1 constructs were calculated using GNOM (Svergun, 1992), from which the R_g and D_{max} were determined. The p(r) versus r profile were also used for *ab initio* bead modelling of selected MORN1 constructs. For this, 20 independent runs of DAMMIF (Franke & Svergun, 2009) in the case of TbMORN1(2-15) and DAMMIN (Svergun, 1999) in case of the shorter MORN1(7-15) constructs were performed. From these the most probable models

were selected by DAMAVER (Volkov & Svergun, 2003). The ab initio modelling was performed with and without symmetry constraints (p2 symmetry to reflect the dimeric state of the protein). Comparison with theoretical curves calculated from the X-tal structures described here was performed with Crysol (Svergun, Barberato et al., 1995). Due to the elongated nature of the molecules, fits were improved by increasing LM (maximum order of harmonics) to 50. The molecular mass (MM) was evaluated based on concentration independent methods according to Porod (Porod, 1951) and as implemented in the ATSAS package. Dimensionless Kratky plots were constructed according to (Receveur-Brechot & Durand, 2012)_and the reference point for globular proteins at √3, 1.1 indicated. Graphical representation was produced using Pymol Molecular Graphics System (Schrödinger, LLC.). The SAXS data (Table S3) and ab initio bead models as well as fits to the crystal structures described within this work have been deposited into the Small-Angle Scattering Biological Data Bank (SASBDB) (Valentini, Kikhney et al., 2015) under the accession codes SASDG97, SASDGA7, SASDGB7, SASDGC7.

Transmission electron microscopy (EM) with rotary shadowing

TbMORN1 and TbMORN1(2-15) were purified according to the two-step procedure detailed above. They were then diluted in spraying buffer (100 mM $NH_4CH_3CO_2$-NaOH pH 8.5, 30% (v/v) glycerol) to a final concentration of 50 – 100 μg/ml. Diluted samples were sprayed onto freshly cleaved mica chips (Christine Gröpl) and immediately transferred into a MED020 high vacuum evaporator (BAL-TEC) equipped with electron guns. While rotating, samples were coated with 0.6 nm of Platinum (BALTIC) at an angle of 4°, followed by 6 nm of Carbon (Oerlicon) at 90°. The obtained replicas were floated off the mica chips, transferred to 400 mesh Cu/Pd grids (Agar Scientific), and examined using a Morgagni 268D transmission electron microscope (FEI) operated at 80 kV. Images were acquired using an 11 megapixel Morada CCD camera (Olympus-SIS).

ACKNOWLEDGEMENTS

Financial support was from Austrian Science Fund (FWF) grants P22265-B12 FWF (to Graham Warren) P27016-B21 (to Brooke Morriswood and Kristina Djinovic-Carugo), P28135 and P30584 (both to Tom Leonard). Kristina Djinovic-Carugo's research was additionally supported by Austrian Science Fund (FWF) Projects I525 and I1593, P22276, P19060 and W1221, Federal Ministry of Economy, Family and Youth through the initiative “Laura Bassi Centres of Expertise”, funding the Centre of Optimized Structural Studies, N°253275, by the Wellcome Trust Collaborative Award (201543/Z/16), COST action BM1405 - Non-globular proteins - from sequence to structure, function and application in molecular physiopathology (NGP-NET), WWTF (Vienna Science and Technology Fund) Chemical Biology project LS17-008, and by the University of Vienna. Funding also came from Medical Research Council grant MR/Mo20118/1 (to Terry Smith). We are indebted to Graham Warren (MRC Laboratory for Molecular Cell Biology, University College London), for comments and support. Valeria Stefania carried out the SAXS measurements. The Martens lab at the Max Perutz Labs Vienna provided the Doc2B protein; Iva Lucic in the Leonard group supplied the Akt protein. Euripedes Almeida Ribeiro advised on fluorescence anisotropy measurements. Markus Engstler

(University of Würzburg) provided the bloodstream form GFP cells as well as lab space and support. Susanne Fenz (University of Würzburg) assisted with the purification of trypanosome lipids. The authors thank the staff of MX and SAXS beamlines at ESRF in Grenoble and EMBL-Hamburg for their excellent support. EM data were recorded at the EM Facility of the Vienna BioCenter Core Facilities GmbH (VBCF), a member of the Vienna BioCenter (VBC), Austria. Proteomics analyses were performed by the Mass Spectrometry Facility at Max Perutz Labs Vienna using the VBCF instrument pool.

References

Adam SA, Marr RS, Gerace L (1990) Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol* 111: 807-16

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213-21

Adl SM, Bass D, Lane CE, Lukes J, Schoch CL, Smirnov A, Agatha S, Berney C, Brown MW, Burki F, Cardenas P, Cepicka I, Chistyakova L, Del Campo J, Dunthorn M, Edvardsen B, Eglit Y, Guillou L, Hampl V, Heiss AA et al. (2019) Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *J Eukaryot Microbiol* 66: 4-119

Allen CL, Goulding D, Field MC (2003) Clathrin-mediated endocytosis is essential in Trypanosoma brucei. *Embo J* 22: 4991-5002

Andrade MA, Perez-Iratxeta C, Ponting CP (2001) Protein repeats: structures, functions, and evolution. *J Struct Biol* 134: 117-31

Bangs JD, Uyetake L, Brickman MJ, Balber AE, Boothroyd JC (1993) Molecular cloning and cellular localization of a BiP homologue in Trypanosoma brucei. Divergent ER retention signals in a lower eukaryote. *J Cell Sci* 105 (Pt 4): 1101-13

Bastin P, Bagherzadeh Z, Matthews KR, Gull K (1996) A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei. *Mol Biochem Parasitol* 77: 235-9

Batram C, Jones NG, Janzen CJ, Markert SM, Engstler M (2014) Expression site attenuation mechanistically links antigenic variation and development in Trypanosoma brucei. *Elife* 3: e02324

Bennett HJ, Davenport JB, Collins RF, Trafford AW, Pinali C, Kitmitto A (2013) Human junctophilin-2 undergoes a structural rearrangement upon binding PtdIns(3,4,5)P3 and the S101R mutation identified in hypertrophic cardiomyopathy obviates this response. *Biochem J* 456: 205-17

1588 Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D,
1589 Mark D, Zengerle R, Cipriani F, Fiedler S, Roessle M, Svergun DI (2015) Versatile sample
1590 environments and automation for biological solution X-ray scattering experiments at the P12
1591 beamline (PETRA III, DESY). *J Appl Crystallogr* 48: 431-443

1592 Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can J*
1593 *Biochem Physiol* 37: 911-7

1594 Brennich ME, Kieffer J, Bonamis G, Antolinos AD, Hutin S, Pernot P, Round A (2016) Online
1595 data analysis at the ESRF bioSAXS beamline, BM29. *Journal of Applied Crystallography* 49:
1596 203-212

1597 Brennich ME, Round AR, Hutin S (2017) Online Size-exclusion and Ion-exchange
1598 Chromatography on a SAXS Beamline. *J Vis Exp*

1599 Burkard G, Fragoso CM, Roditi I (2007) Highly efficient stable transformation of bloodstream
1600 forms of Trypanosoma brucei. *Mol Biochem Parasitol* 153: 220-3

1601 Camacho L, Smertenko AP, Perez-Gomez J, Hussey PJ, Moore I (2009) Arabidopsis Rab-E
1602 GTPases exhibit a novel interaction with a plasma-membrane phosphatidylinositol-4-
1603 phosphate 5-kinase. *J Cell Sci* 122: 4383-92

1604 Carman PJ, Dominguez R (2018) BAR domain proteins-a linkage between cellular
1605 membranes, signaling pathways, and the actin cytoskeleton. *Biophys Rev* 10: 1587-1604

1606 Cronan JE (2003) Bacterial membrane lipids: where do we stand? *Annu Rev Microbiol* 57:
1607 203-24

1608 Demmel L, Schmidt K, Lucast L, Havlicek K, Zankel A, Koestler T, Reithofer V, de Camilli P,
1609 Warren G (2014) The endocytic activity of the flagellar pocket in Trypanosoma brucei is
1610 regulated by an adjacent phosphatidylinositol phosphate kinase. *Journal of cell science* 127:
1611 2351-64

1612 Doyle SA (2005) High-throughput cloning for proteomics research. *Methods Mol Biol* 310:
1613 107-13

1614 El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson
1615 LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE,
1616 Finn RD (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47: D427-
1617 D432

1618 Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta*
1619 *Crystallogr D Biol Crystallogr* 60: 2126-32

1620 Engstler M, Thilo L, Weise F, Grunfelder CG, Schwarz H, Boshart M, Overath P (2004)
1621 Kinetics of endocytosis and recycling of the GPI-anchored variant surface glycoprotein in
1622 Trypanosoma brucei. *Journal of cell science* 117: 1105-15

1623 Epand RM, Epand RF (2009) Lipid domains in bacterial membranes and the action of
1624 antimicrobial agents. *Biochim Biophys Acta* 1788: 289-94

1625 Esson HJ, Morriswood B, Yavuz S, Vidilaseris K, Dong G, Warren G (2012) Morphology of
1626 the trypanosome bilobe, a novel cytoskeletal structure. *Eukaryot Cell* 11: 761-72

1627 Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney
1628 AG, Hajizadeh NR, Franklin JM, Jeffries CM, Svergun DI (2017) ATSAS 2.8: a
1629 comprehensive data analysis suite for small-angle scattering from macromolecular solutions.
1630 *J Appl Crystallogr* 50: 1212-1225

1631 Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in
1632 small-angle scattering. *J Appl Crystallogr* 42: 342-346

1633 Franzot G, Sjoblom B, Gautel M, Djinovic Carugo K (2005) The crystal structure of the actin
1634 binding domain from alpha-actinin in its closed conformation: structural insight into
1635 phospholipid regulation of alpha-actinin. *J Mol Biol* 348: 151-65

1636 Fyffe SA, Alphey MS, Buetow L, Smith TK, Ferguson MA, Sorensen MD, Bjorkling F, Hunter
1637 WN (2006) Reevaluation of the PPAR-beta/delta ligand binding domain model reveals why it
1638 exhibits the activated form. *Mol Cell* 21: 1-2

1639 Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO (2009)
1640 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6:
1641 343-5

1642 Graewert MA, Franke D, Jeffries CM, Blanchet CE, Ruskule D, Kuhle K, Flieger A, Schafer
1643 B, Tartsch B, Meijers R, Svergun DI (2015) Automated pipeline for purification, biophysical
1644 and x-ray analysis of biomacromolecular solutions. *Sci Rep* 5: 10734

1645 Grunfelder CG, Engstler M, Weise F, Schwarz H, Stierhof YD, Morgan GW, Field MC,
1646 Overath P (2003) Endocytosis of a glycosylphosphatidylinositol-anchored protein via
1647 clathrin-coated vesicles, sorting by default in endosomes, and exocytosis via RAB11-positive
1648 carriers. *Mol Biol Cell* 14: 2029-40

1649 Gubbels MJ, Vaishnava S, Boot N, Dubremetz JF, Striepen B (2006) A MORN-repeat
1650 protein is a dynamic component of the Toxoplasma gondii cell division apparatus. *J Cell Sci*
1651 119: 2236-45

1652 Guo A, Wang Y, Chen B, Wang Y, Yuan J, Zhang L, Hall D, Wu J, Shi Y, Zhu Q, Chen C,
1653 Thiel WH, Zhan X, Weiss RM, Zhan F, Musselman CA, Pufall M, Zhu W, Au KF, Hong J et
1654 al. (2018) E-C coupling structural protein junctophilin-2 encodes a stress-adaptive
1655 transcription regulator. *Science* 362

1656 Habicht J, Woehle C, Gould SB (2015) Tetrahymena Expresses More than a Hundred
1657 Proteins with Lipid-binding MORN Motifs that can Differ in their Subcellular Localisations. *J*
1658 *Eukaryot Microbiol* 62: 694-700

1659 Hatzopoulos GN, Erat MC, Cutts E, Rogala KB, Slater LM, Stansfeld PJ, Vakonakis I (2013)
1660 Structural analysis of the G-box domain of the microcephaly protein CPAP suggests a role in
1661 centriole architecture. *Structure* 21: 2069-77

1662 Hemsley A, Arnheim N, Toney MD, Cortopassi G, Galas DJ (1989) A simple method for site-
1663 directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res* 17: 6545-51

1664 Holm L, Laakso LM (2016) Dali server update. *Nucleic Acids Res* 44: W351-5

1665 Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res*
1666 38: W545-9

1667 Hu K (2008) Organizational changes of the daughter basal complex during the parasite
1668 replication of Toxoplasma gondii. *PLoS Pathog* 4: e10

1669 Hutchinson EG, Thornton JM (1994) A revised set of potentials for beta-turn formation in
1670 proteins. *Protein Sci* 3: 2207-16

1671 Im YJ, Davis AJ, Perera IY, Johannes E, Allen NS, Boss WF (2007) The N-terminal
1672 membrane occupation and recognition nexus domain of Arabidopsis phosphatidylinositol
1673 phosphate kinase 1 regulates enzyme activity. *J Biol Chem* 282: 5443-52

1674 Jacobs SA, Harp JM, Devarakonda S, Kim Y, Rastinejad F, Khorasanizadeh S (2002) The
1675 active site of the SET domain is constructed on a knot. *Nat Struct Biol* 9: 833-8

1676 Jayasinghe I, Clowsley AH, Lin R, Lutz T, Harrison C, Green E, Baddeley D, Di Michele L,
1677 Soeller C (2018) True Molecular Scale Visualization of Variable Clustering Properties of
1678 Ryanodine Receptors. *Cell Rep* 22: 557-567

1679 Jiang M, Hu J, White FKH, Williamson J, Klymchenko AS, Murthy A, Workman SW, Tseng
1680 GN (2019) S-Palmitoylation of junctophilin-2 is critical for its role in tethering the
1681 sarcoplasmic reticulum to the plasma membrane. *J Biol Chem* 294: 13487-13501

1682 Joosten RP, Long F, Murshudov GN, Perrakis A (2014) The PDB_REDO server for
1683 macromolecular structure model optimization. *IUCrJ* 1: 213-20

1684 Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66: 125-32

1685 Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochim*
1686 *Biophys Acta* 1751: 119-39

1687 Kohl L, Sherwin T, Gull K (1999) Assembly of the paraflagellar rod and the flagellum
1688 attachment zone complex during the Trypanosoma brucei cell cycle. *J Eukaryot Microbiol*
1689 46: 105-9

1690 Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state.
1691 *J Mol Biol* 372: 774-97

1692 Kunita R, Otomo A, Mizumura H, Suzuki K, Showguchi-Miyata J, Yanagisawa Y, Hadano S,
1693 Ikeda JE (2004) Homo-oligomerization of ALS2 through its unique carboxyl-terminal regions
1694 is essential for the ALS2-associated Rab5 guanine nucleotide exchange activity and its
1695 regulatory function on endosome trafficking. *J Biol Chem* 279: 38626-35

1696 Lacomble S, Vaughan S, Gadelha C, Morphew MK, Shaw MK, McIntosh JR, Gull K (2009)
1697 Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in
1698 trypanosomes revealed by electron microscope tomography. *Journal of cell science* 122:
1699 1081-90

1700 Li J, Liu H, Raval MH, Wan J, Yengo CM, Liu W, Zhang M (2019) Structure of the
1701 MORN4/Myo3a Tail Complex Reveals MORN Repeats as Protein Binding Modules.
1702 *Structure*

1703 Li MZ, Elledge SJ (2012) SLIC: a method for sequence- and ligation-independent cloning.
1704 *Methods Mol Biol* 852: 51-9

1705 Ma H, Lou Y, Lin WH, Xue HW (2006) MORN motifs in plant PIPKs are involved in the
1706 regulation of subcellular localization and phospholipid binding. *Cell Res* 16: 466-78

1707 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007)
1708 Phaser crystallographic software. *J Appl Crystallogr* 40: 658-674

1709 Mecklenburg KL (2007) Drosophila retinophilin contains MORN repeats and is conserved in
1710 humans. *Mol Genet Genomics* 277: 481-9

1711 Micsonai A, Wien F, Bulyaki E, Kun J, Moussong E, Lee YH, Goto Y, Refregiers M, Kardos J
1712 (2018) BeStSel: a web server for accurate protein secondary structure prediction and fold
1713 recognition from the circular dichroism spectra. *Nucleic Acids Res* 46: W315-W322

1714 Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, Kardos J (2015) Accurate
1715 secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc*
1716 *Natl Acad Sci U S A* 112: E3095-103

1717 Mikami K, Saavedra L, Sommarin M (2010) Is membrane occupation and recognition nexus
1718 domain functional in plant phosphatidylinositol phosphate kinases? *Plant Signal Behav* 5:
1719 1241-4

1720 Molla-Herman A, Ghossoub R, Blisnick T, Meunier A, Serres C, Silbermann F, Emmerson C,
1721 Romeo K, Bourdoncle P, Schmitt A, Saunier S, Spassky N, Bastin P, Benmerah A (2010)
1722 The ciliary pocket: an endocytic membrane domain at the base of primary and motile cilia. *J*
1723 *Cell Sci* 123: 1785-95

1724 Morriswood B, Havlicek K, Demmel L, Yavuz S, Sealey-Cardona M, Vidilaseris K, Anrather
1725 D, Kostan J, Djinovic-Carugo K, Roux KJ, Warren G (2013) Novel Bilobe Components in
1726 Trypanosoma brucei Identified Using Proximity-Dependent Biotinylation. *Eukaryot Cell* 12:
1727 356-67

1728 Morriswood B, He CY, Sealey-Cardona M, Yelinek J, Pypaert M, Warren G (2009) The
1729 bilobe structure of Trypanosoma brucei contains a MORN-repeat protein. *Mol Biochem*
1730 *Parasitol* 167: 95-103

1731 Morriswood B, Schmidt K (2015) A MORN-repeat protein facilitates protein entry into the
1732 flagellar pocket of Trypanosoma brucei. *Eukaryot Cell*

1733 Munro ML, Jayasinghe ID, Wang Q, Quick A, Wang W, Baddeley D, Wehrens XH, Soeller C
1734 (2016) Junctophilin-2 in the nanoscale organisation and functional signalling of ryanodine
1735 receptor clusters in cardiomyocytes. *J Cell Sci* 129: 4388-4398

1736 Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by
1737 the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240-55

1738 Nakada T, Kashihara T, Komatsu M, Kojima K, Takeshita T, Yamada M (2018) Physical
1739 interaction of junctophilin and the CaV1.1 C terminus is crucial for skeletal muscle
1740 contraction. *Proc Natl Acad Sci U S A* 115: 4507-4512

1741 Otomo A, Hadano S, Okada T, Mizumura H, Kunita R, Nishijima H, Showguchi-Miyata J,
1742 Yanagisawa Y, Kohiki E, Suga E, Yasuda M, Osuga H, Nishimoto T, Narumiya S, Ikeda JE
1743 (2003) ALS2, a novel guanine nucleotide exchange factor for the small GTPase Rab5, is
1744 implicated in endosomal dynamics. *Hum Mol Genet* 12: 1671-87

1745 Panjkovich A, Svergun DI (2018) CHROMIXS: automatic and interactive analysis of
1746 chromatography-coupled small-angle X-ray scattering data. *Bioinformatics* 34: 1944-1946

1747 Pelletier L, Stern CA, Pypaert M, Sheff D, Ngo HM, Roper N, He CY, Hu K, Toomre D,
1748 Coppens I, Roos DS, Joiner KA, Warren G (2002) Golgi biogenesis in Toxoplasma gondii.
1749 *Nature* 418: 548-52

1750 Perni S, Lavorato M, Beam KG (2017) De novo reconstitution reveals the proteins required
1751 for skeletal muscle voltage-induced Ca(2+) release. *Proc Natl Acad Sci U S A* 114: 13822-
1752 13827

1753 Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, Huet J, Kieffer J,
1754 Lentini M, Mattenet M, Morawe C, Mueller-Dieckmann C, Ohlsson S, Schmid W, Surr J,
1755 Theveneau P, Zerrad L, McSweeney S (2013) Upgraded ESRF BM29 beamline for SAXS
1756 on macromolecules in solution. *J Synchrotron Radiat* 20: 660-4

1757 Porod G (1951) Die Rontgenkleinwinkelstreuung Von Dichtgepackten Kolloiden Systemen
1758 .1. *Kolloid Z Z Polym* 124: 83-114

1759 Pramanik J, Keasling JD (1997) Stoichiometric model of Escherichia coli metabolism:
1760 incorporation of growth-rate dependent biomass composition and mechanistic energy
1761 requirements. *Biotechnol Bioeng* 56: 398-421

1762 Rao VD, Misra S, Boronenkov IV, Anderson RA, Hurley JH (1998) Structure of type IIbeta
1763 phosphatidylinositol phosphate kinase: a protein kinase fold flattened for interfacial
1764 phosphorylation. *Cell* 94: 829-39

1765 Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-
1766 fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896-
1767 906

1768 Receveur-Brechot V, Durand D (2012) How random are intrinsically disordered proteins? A
1769 small angle scattering perspective. *Curr Protein Pept Sci* 13: 55-75

1770 Richmond GS, Gibellini F, Young SA, Major L, Denton H, Lilley A, Smith TK (2010)
1771 Lipidomic analysis of bloodstream and procyclic form Trypanosoma brucei. *Parasitology*
1772 137: 1357-92

1773 Rossi D, Scarcella AM, Liguori E, Lorenzini S, Pierantozzi E, Kutchukian C, Jacquemond V,
1774 Messa M, De Camilli P, Sorrentino V (2019) Molecular determinants of homo- and
1775 heteromeric interactions of Junctophilin-1 at triads in adult skeletal muscle fibers. *Proc Natl*
1776 *Acad Sci U S A* 116: 15716-15724

1777 Saliba AE, Vonkova I, Ceschia S, Findlay GM, Maeda K, Tischer C, Deghou S, van Noort V,
1778 Bork P, Pawson T, Ellenberg J, Gavin AC (2014) A quantitative liposome microarray to
1779 systematically characterize protein-lipid interactions. *Nat Methods* 11: 47-50

1780 Saliba AE, Vonkova I, Deghou S, Ceschia S, Tischer C, Kugler KG, Bork P, Ellenberg J,
1781 Gavin AC (2016) A protocol for the systematic and quantitative measurement of protein-lipid
1782 interactions using the liposome-microarray-based assay. *Nat Protoc* 11: 1021-38

1783 Salzer U, Kostan J, Djinovic-Carugo K (2017) Deciphering the BAR code of membrane
1784 modulators. *Cell Mol Life Sci* 74: 2413-2438

1785 Scholz J, Besir H, Strasser C, Suppmann S (2013) A new method to customize protein
1786 expression vectors for fast, efficient and background free parallel cloning. *BMC Biotechnol*
1787 13: 12

1788 Schumann Burkard G, Jutzi P, Roditi I (2011) Genome-wide RNAi screens in bloodstream
1789 form trypanosomes identify drug transporters. *Mol Biochem Parasitol* 175: 91-4

1790 Studier FW (2005) Protein production by auto-induction in high density shaking cultures.
1791 *Protein Expr Purif* 41: 207-34

1792 Svergun D, Barberato C, Koch MHJ (1995) CRYSOL - A program to evaluate x-ray solution
1793 scattering of biological macromolecules from atomic coordinates. *Journal of Applied*
1794 *Crystallography* 28: 768-773

1795 Svergun DI (1992) Determination of the Regularization Parameter in Indirect-Transform
1796 Methods Using Perceptual Criteria. *Journal of Applied Crystallography* 25: 495-503

1797 Svergun DI (1999) Restoring low resolution structure of biological macromolecules from
1798 solution scattering using simulated annealing. *Biophys J* 76: 2879-86

1799 Takeshima H, Komazaki S, Nishi M, Iino M, Kangawa K (2000) Junctophilins: a novel family
1800 of junctional membrane complex proteins. *Mol Cell* 6: 11-22

1801 Tsuchida J, Nishina Y, Wakabayashi N, Nozaki M, Sakai Y, Nishimune Y (1998) Molecular
1802 cloning and characterization of meichroacidin (male meiotic metaphase chromosome-
1803 associated acidic protein). *Dev Biol* 197: 67-76

1804 Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass
1805 spectrometry-based shotgun proteomics. *Nat Protoc* 11: 2301-2319

1806 Valentini E, Kikhney AG, Previtali G, Jeffries CM, Svergun DI (2015) SASBDB, a repository
1807 for biological small-angle scattering data. *Nucleic Acids Res* 43: D357-63

1808 Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle
1809 scattering. *Journal of Applied Crystallography* 36: 860-864

1810 Walter TS, Meier C, Assenberg R, Au KF, Ren J, Verma A, Nettleship JE, Owens RJ, Stuart
1811 DI, Grimes JM (2006) Lysine methylation as a routine rescue strategy for protein
1812 crystallization. *Structure* 14: 1617-22

1813 Wilson JR, Jing C, Walker PA, Martin SR, Howell SA, Blackburn GM, Gamblin SJ, Xiao B
1814 (2002) Crystal structure and functional analysis of the histone methyltransferase SET7/9.
1815 *Cell* 111: 105-15

1816 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel
1817 EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell
1818 HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current
1819 developments. *Acta Crystallogr D Biol Crystallogr* 67: 235-42

1820 Wirtz E, Leal S, Ochatt C, Cross GA (1999) A tightly regulated inducible expression system
1821 for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. *Mol*
1822 *Biochem Parasitol* 99: 89-101

1823 Woo JS, Srikanth S, Nishi M, Ping P, Takeshima H, Gwack Y (2016) Junctophilin-4, a
1824 component of the endoplasmic reticulum-plasma membrane junctions, regulates Ca2+
1825 dynamics in T cells. *Proc Natl Acad Sci U S A* 113: 2762-7

1826 Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, Kelly G, Howell S, Taylor IA, Blackburn
1827 GM, Gamblin SJ (2003) Structure and catalytic mechanism of the human histone
1828 methyltransferase SET7/9. *Nature* 421: 652-6

1829 Yamanaka K, Vande Velde C, Eymard-Pierre E, Bertini E, Boespflug-Tanguy O, Cleveland
1830 DW (2003) Unstable mutants in the peripheral endosomal membrane component ALS2
1831 cause early-onset motor neuron disease. *Proc Natl Acad Sci U S A* 100: 16041-6

1832 Yang B, Wu YJ, Zhu M, Fan SB, Lin J, Zhang K, Li S, Chi H, Li YX, Chen HF, Luo SK, Ding
1833 YH, Wang LH, Hao Z, Xiu LY, Chen S, Ye K, He SM, Dong MQ (2012) Identification of
1834 cross-linked peptides from complex samples. *Nat Methods* 9: 904-6

1835 Yuan ZF, Liu C, Wang HP, Sun RX, Fu Y, Zhang JF, Wang LH, Chi H, Li Y, Xiu LY, Wang
1836 WP, He SM (2012) pParse: a method for accurate determination of monoisotopic peaks in
1837 high-resolution mass spectra. *Proteomics* 12: 226-35
1838

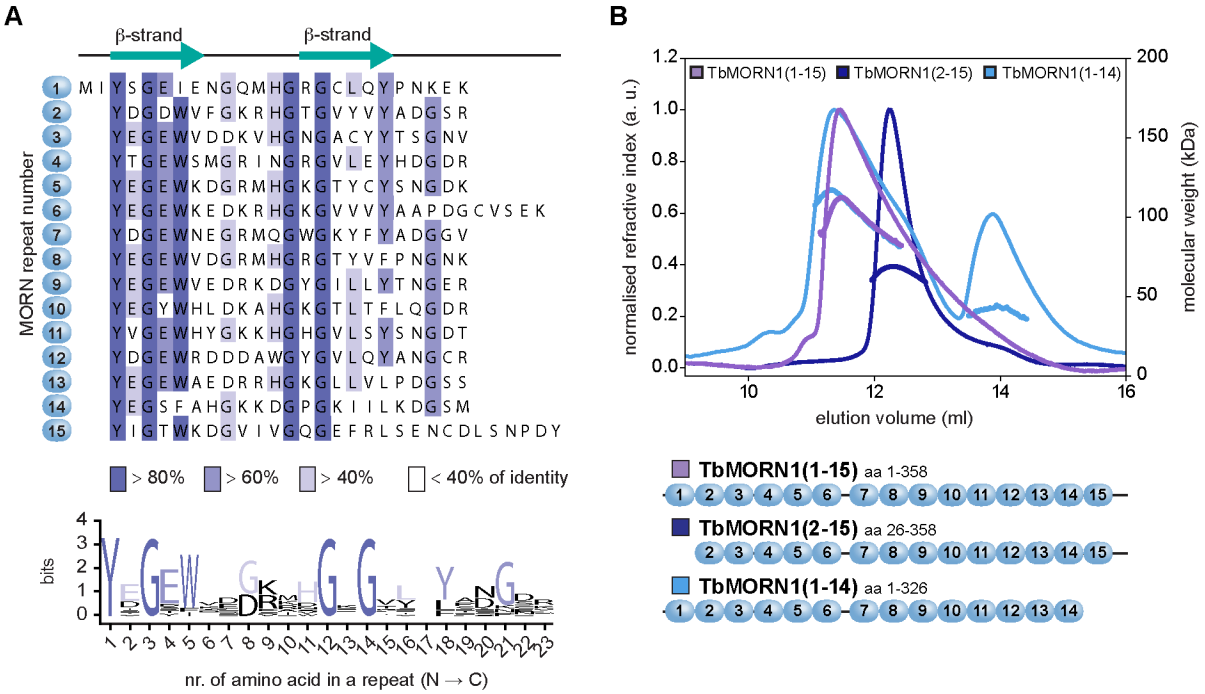


Figure 1. TbMORN1 primary structure and dimerisation. (A) Primary structure of TbMORN1, with individual MORN repeats shown in alignment, and coloured according to amino acid conservation. A schematic of the predicted secondary structure of each repeat is shown above the alignment. A consensus amino acid sequence of the individual MORN repeats from TbMORN1 based on the alignment is shown in the sequence logo. (B) TbMORN1 dimerises via its C-terminus. SEC-MALS profiles of TbMORN1(1-15), TbMORN1 (2-15), and TbMORN1 (1-14). Schematics are shown underneath. TbMORN1(1-15) tended to form high-order assemblies, whereas removal of the first MORN repeat resulted in a monodisperse dimer. Removal of the last MORN repeat in TbMORN1 (1-14) resulted in a polydisperse mixture of higher-order assemblies, dimers, and monomers.

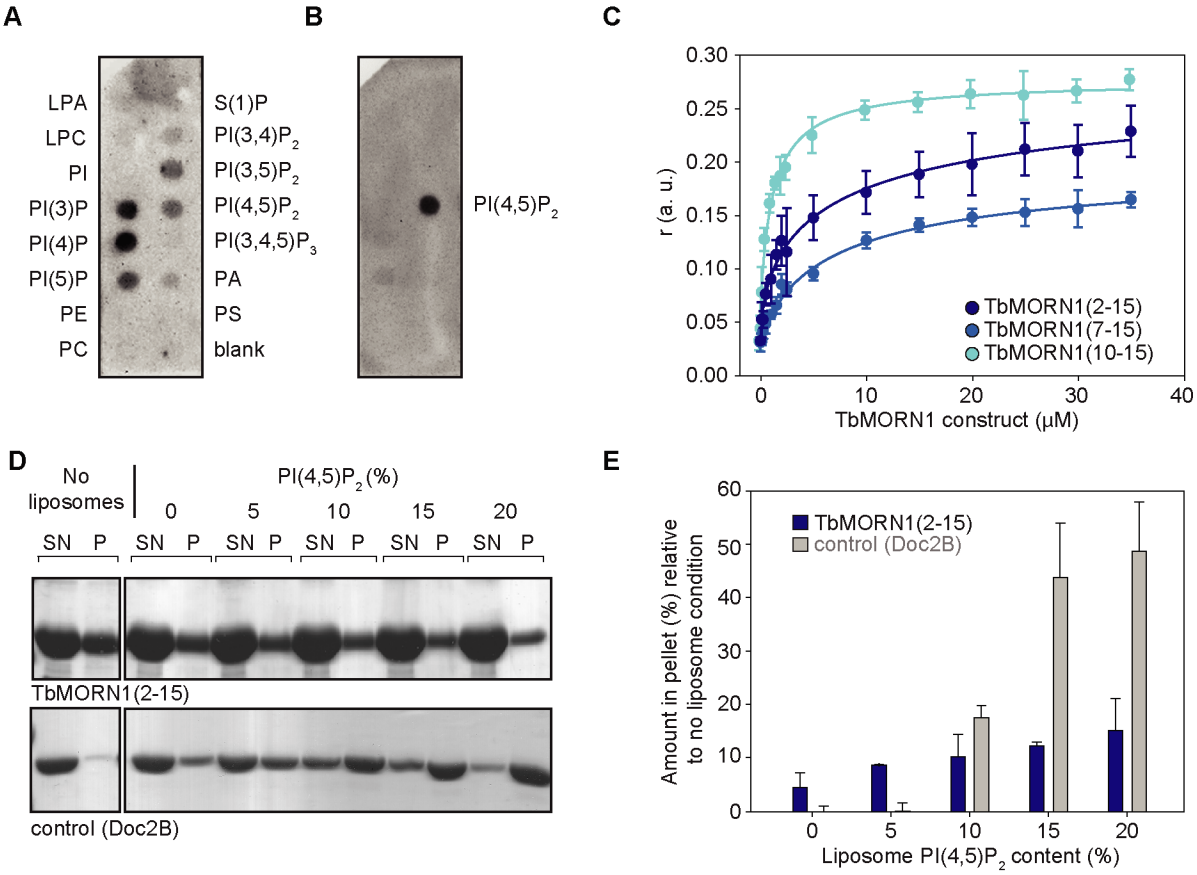


Figure 2. TbMORN1 interacts with phospholipids but not liposomes. (A) Purified recombinant TbMORN1 binds to multiple phospholipid species in protein-lipid overlay assays. PIP strips were incubated with purified recombinant TbMORN1(1-15) protein, and bound proteins were detected by immunoblotting with an anti-His tag antibody. Abbreviations: PI(n)P, phosphatidylinositol (n) phosphate; PA, phosphatidic acid; LPA, lysophosphatidic acid; LPC, lysophosphatidylcholine; PI, phosphatidylinositol; PE, phosphatidylethanolamine; PC, phosphatidylcholine; S(1)P, sphingosine-1-phosphate; PS, phosphatidylserine. Data were obtained from 3 independent experiments using 2 biological replicates; an exemplary blot is shown. (B) PIP strip overlaid with the PH domain of PLCδ, a positive control for $PI(4,5)P_2$ binding. Data were obtained from 3 independent experiments using 2 biological replicates; an exemplary blot is shown. The PIP strips presented here were exposed to the light source for the same time. (C) Fluorescence anisotropy measurements of 0.1 μM BODIPY TMR-$PI(4,5)P_2$ in the presence of TbMORN1(2-15), (7-15) and (10-15). All three truncations of TbMORN1 interacted with $PI(4,5)P_2$ with binding affinities in the low micromolar range. Data obtained from 3 independent experiments using 3 biological replicates, with 10 technical replicates for each experiment. Traces show mean values, bars are s.e.m. (D) Liposome co-sedimentation assay. POPC liposomes containing 0, 5, 10, 15 and 20% of porcine brain $PI(4,5)P_2$ were incubated with 10 μM TbMORN1(2-15). TbMORN1(2-15) was found in both pellet (P) and supernatant (SN) fractions but did not increase proportionally to $PI(4,5)P_2$ concentration. The positive control, Doc2B, bound $PI(4,5)P_2$ in a concentration-dependent manner, with a shift from SN to P fractions proportional to the increase in % of $PI(4,5)P_2$ present in the liposomes. Data were obtained from 2 independent experiments using 2 biological replicates; an exemplary blot is shown. (E) Quantification of the liposome pelleting assays. The amount of protein in the pellet fraction is presented relative to the amount present in the pellet fraction of the no liposome condition.

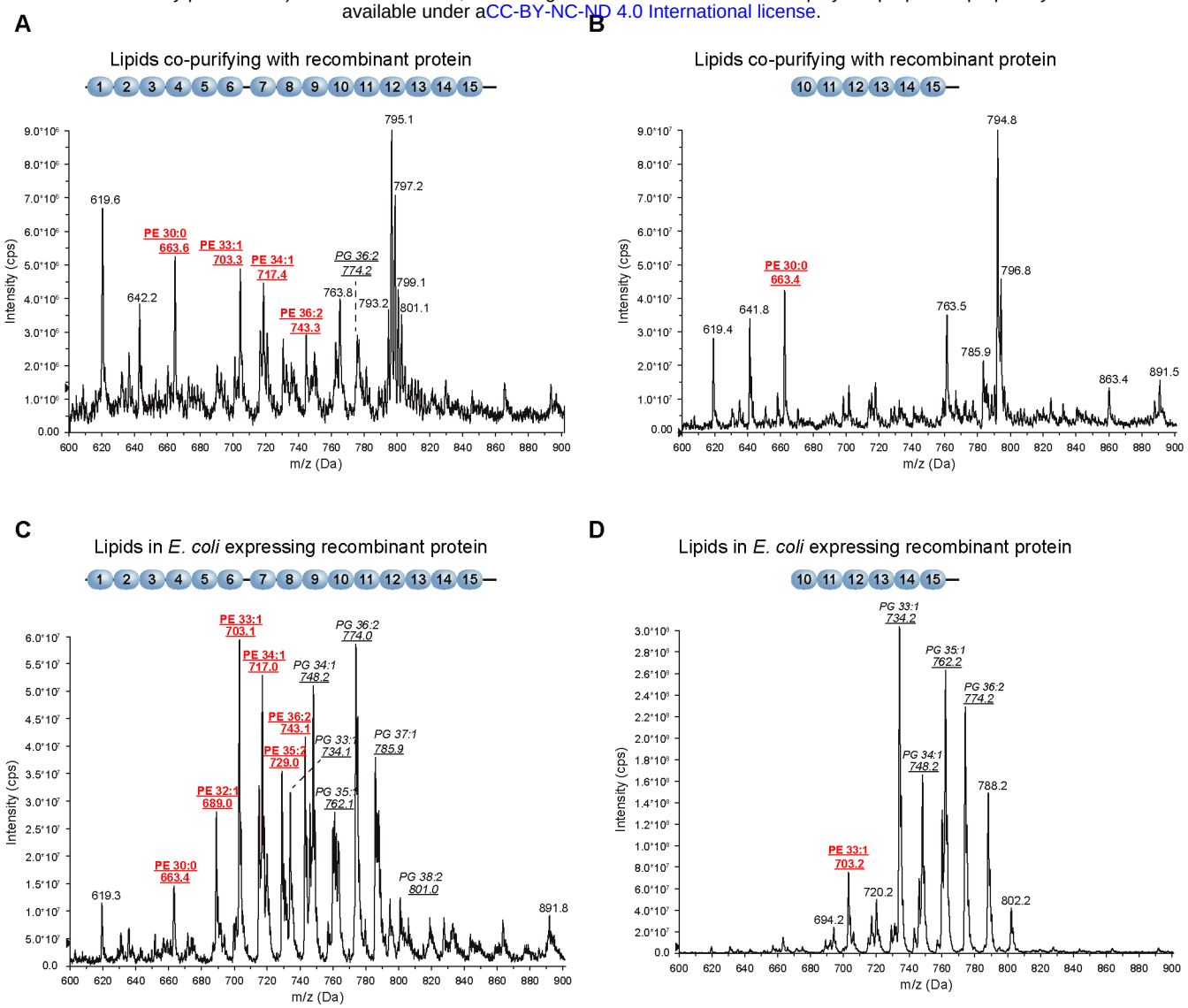


Figure 3. Recombinant TbMORN1 co-purifies with PE and increases *E. coli* PE levels. Negative ion mode survey scan (600-900 m/z) of lipid extracts from the indicated conditions. (A,B) Lipid extracts from purified recombinant TbMORN1(1-15) (A) and TbMORN1(10-15) (B). A large amount of PE co-purified with TbMORN1(1-15) but very little was associated with TbMORN1(10-15). (C,D) Lipid extracts from *E. coli* cells expressing the indicated constructs. (C) Cells expressing TbMORN1(1-15) had elevated PE levels. (D) Cells expressing TbMORN1(10-15) showed no changes to cellular lipid ratios compared to wild-type (empty vector control). In all cases, phospholipid identity was confirmed by daughter fragmentation and reported here. Schematics of the recombinant TbMORN1 constructs are shown above the traces.

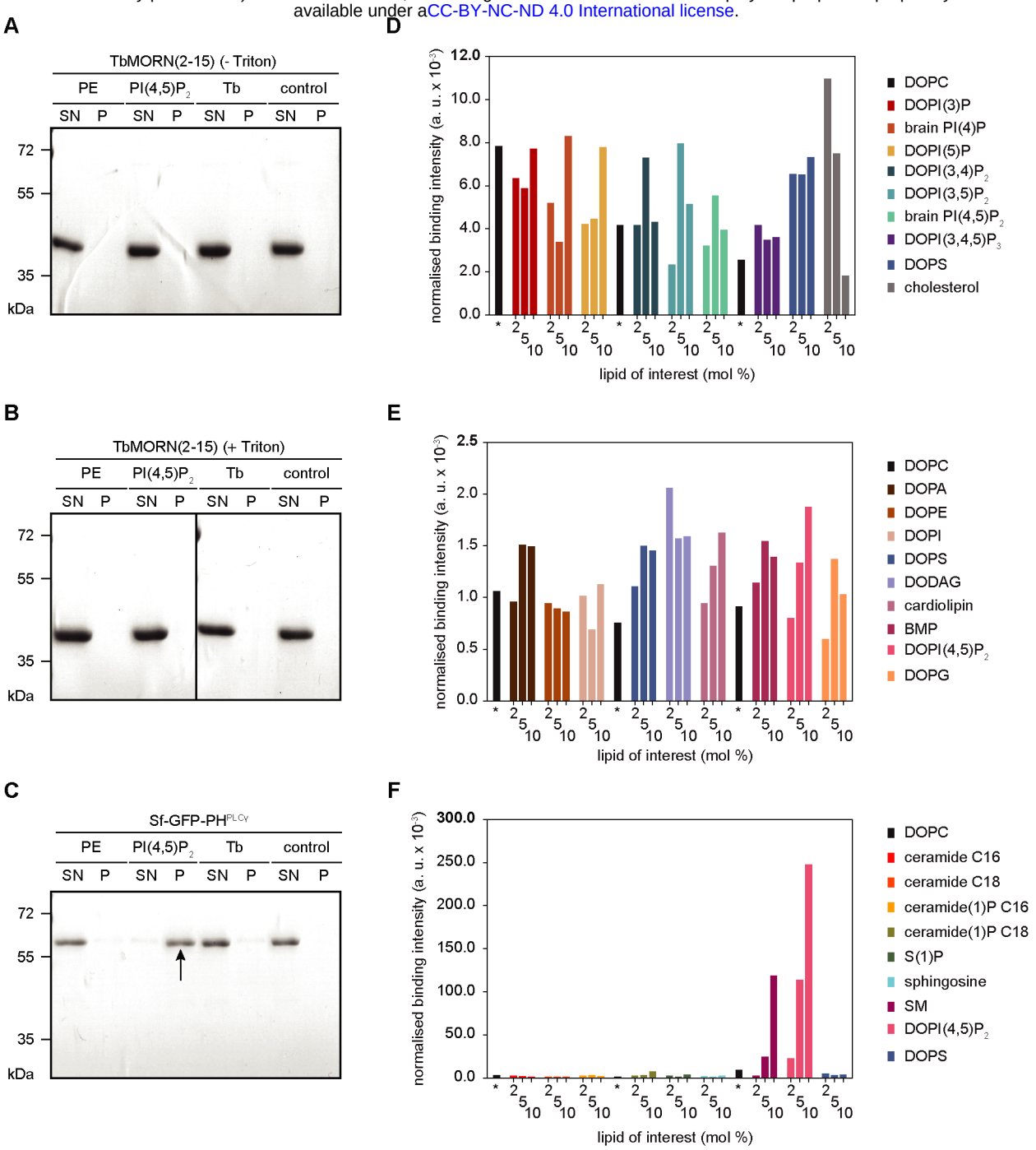


Figure 4. TbMORN1(2-15) does not bind to liposomes in vitro. (A-C) Liposome pelleting assays using sucrose-loaded vesicles (SLVs). His-TbMORN1(2-15) was purified in the absence (A) or presence (B) of Triton X-100 in the lysis buffer. The purified proteins were incubated with SLVs, which were then pelleted by centrifugation. Supernatant (SN) and pellet (P) fractions were analysed by SDS-PAGE using Coomassie staining. The SLVs were made from commercial lipids with an excess of either PE or $PI(4,5)P_2$, and also reconstituted from purified whole-cell trypanosome lipids (Tb). A no-SLV condition was included as an additional negative control. (C) The PH domains of PLCγ was used as a positive control for $PI(4,5)P_2$ binding. As expected, the PLCγ PH domain co-sedimented with $PI(4,5)P_2$-containing SLVs and was entirely present in the P fraction in this condition (arrow). The recombinant TbMORN1 proteins remained in the SN fraction in all conditions. (D-F) Liposome microarray analysis. Microchips carrying giant unilamellar vesicles (GUVs) with lipids of interest at three different concentrations (2, 5 and 10 mol %) were incubated with purified recombinant EGFP-TbMORN1(2-15). No significant binding was observed. (D) n (independent replicates) = 7; (E) n (independent replicates) = 3. (F) Microchip incubated with PLC-δ1 PH domain. A specific and concentration-dependent binding between the PLC-δ1 PH domain and $DOPI(4,5)P_2$ and SM was observed. DOPC, a carrier lipid, was used as an internal negative control of binding, as well as a marker for tracking positions of liposomes on a given microarray. n (independent replicates) = 1. Note the different scales on the y-axes of the three charts.

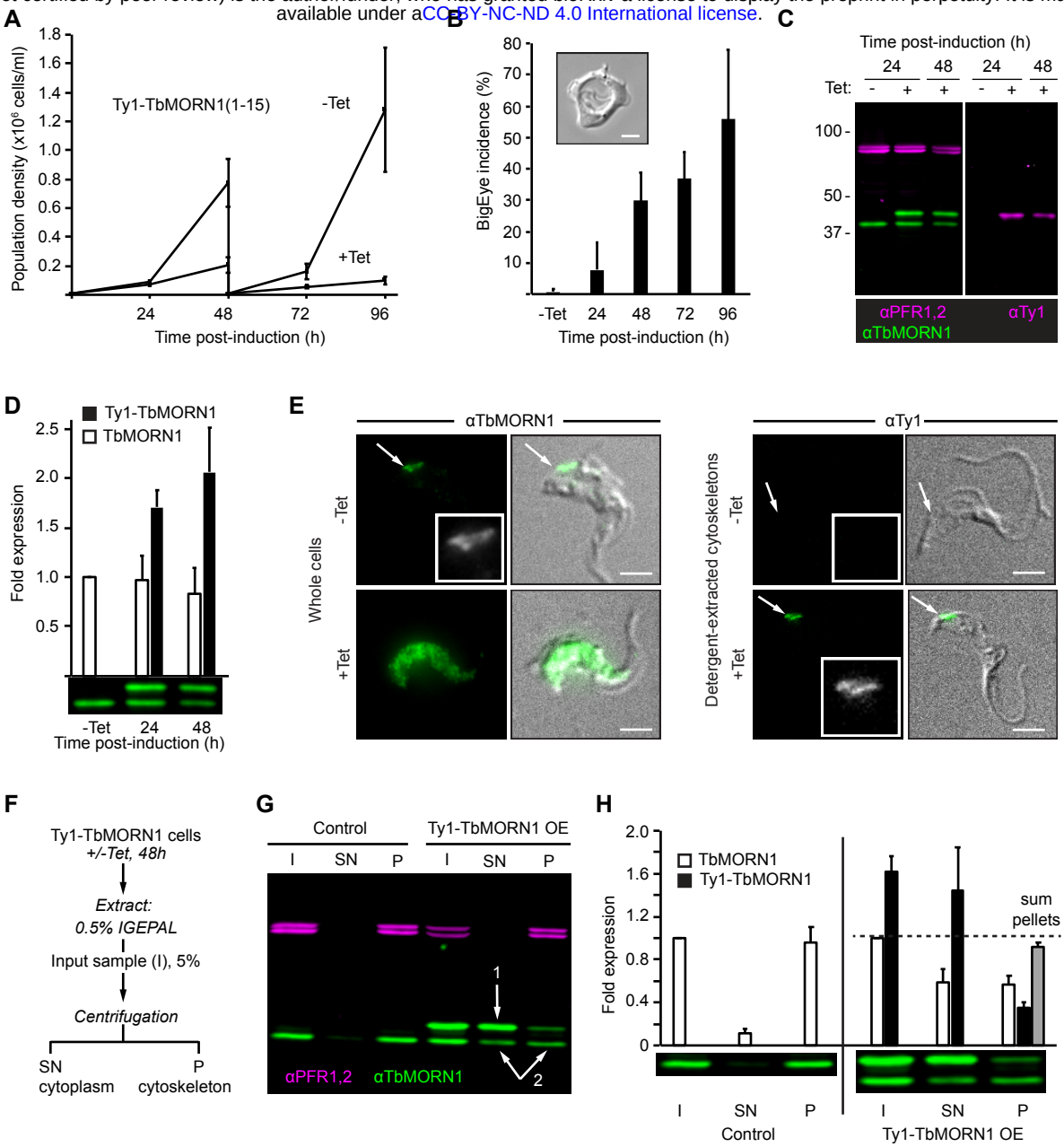


Figure 5. Overexpression of Ty1-tagged TbMORN1 causes a dominant negative phenotype. (A) Overexpression of Ty1-TbMORN1 is deleterious. Growth curves of control (-Tet) cells, and cells inducibly expressing Ty1-TbMORN1(1-15) (+Tet). Population density was measured every 24h, and the cultures split and reseeded at 48h. Data were compiled from 3 separate clones, each induced in 3 independent experiments; bars show mean +/- SD. (B) Overexpression of Ty1-TbMORN1 produces a BigEye phenotype. The incidence of BigEye cells was counted in control (-Tet) and Ty1-TbMORN1-expressing cells at the indicated timepoints. Data were compiled from 3 separate clones, each induced in 3 independent experiments; bars show mean + SD. The inset shows an example BigEye cell. Scale bar, 2 μm. (C) Tight induction of Ty1-TbMORN1 expression. Whole-cell lysates were harvested from control (-Tet) and Ty1-TbMORN1-expressing cells (+Tet) at the indicated timepoints and probed with anti-TbMORN1 and anti-Ty1 antibodies. PFR1,2 were used as a loading control. At least three independent inductions were carried out for each clone; an exemplary blot is shown. (D) Quantification of overexpression. The levels of endogenous TbMORN1 and ectopic Ty1-TbMORN1 in immunoblots were normalised relative to the PFR1,2 signal. Data were compiled using 3 separate clones, each induced in at least two independent experiments; bars show mean + SD. (E) Ty1-TbMORN1 can localise correctly to the cytoskeleton. Whole cells or detergent-extracted cytoskeletons were fixed and labelled with anti-TbMORN1 or anti-Ty1 antibodies. The fluorescence signal is shown with the transmitted light image of the cell overlaid; inset shows the fluorescence signal from the antibody in greyscale. Results confirmed for 3 separate clones, exemplary images are shown. Scale bars, 2 μm. (F) Schematic of the fractionation protocol. (G) Overexpression of Ty1-TbMORN1 displaces the endogenous protein from the cytoskeleton. Control and Ty1-TbMORN1-expressing cells were fractionated as shown in panel F and the I, SN, and P fractions were blotted. PFR1,2 was used as a marker for the cytoskeleton. Expression of Ty1-TbMORN1 was accompanied by a displacement of endogenous TbMORN1 from the insoluble (P) fraction into the soluble (SN) fraction (arrows 1,2). Equal fractions (5%) were loaded in each lane. 3 independent experiments using 3 separate clones were carried out; an exemplary blot is shown. (H) Quantification of the fractionation data. Bars show mean + SD.

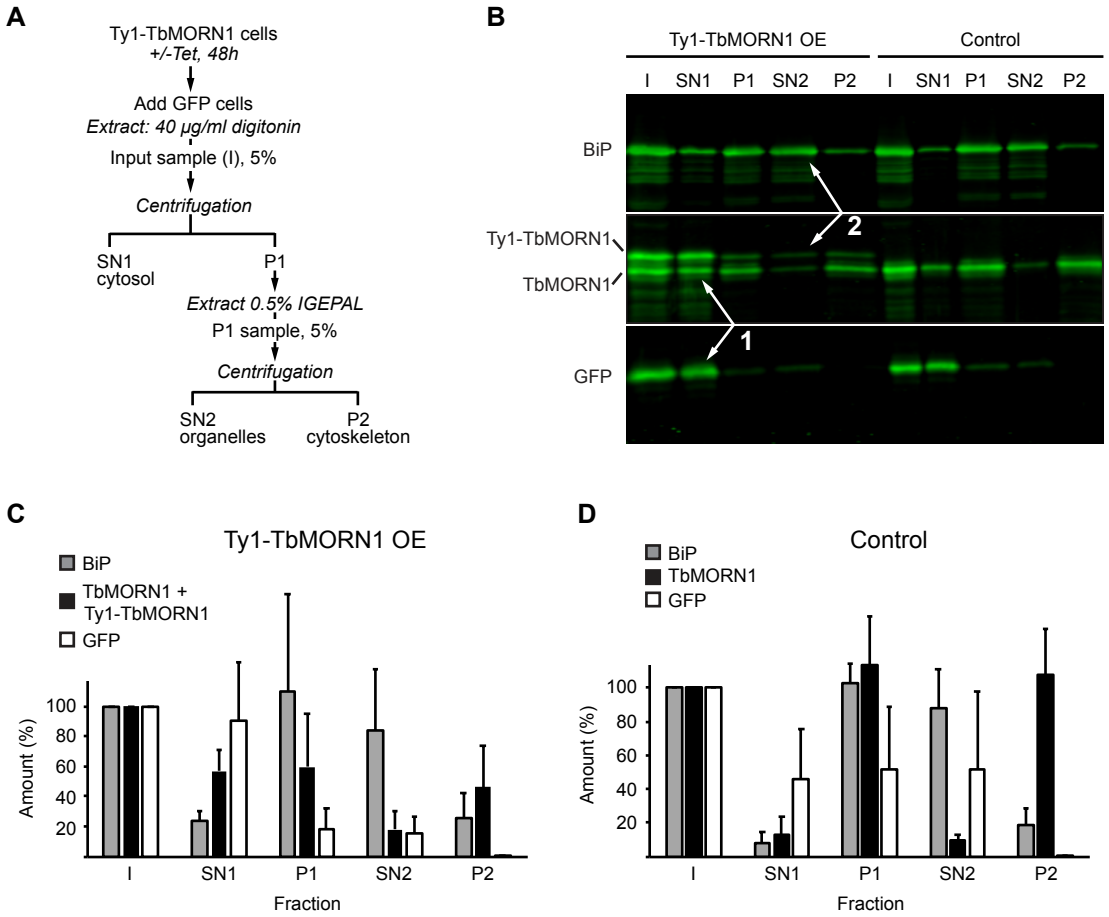


Figure 6. Overexpressed Ty1-TbMORN1 is predominantly cytosolic. (A) Schematic of the two-step fractionation scheme. (B) Immunoblots of fractions taken from control and Ty1-TbMORN1 overexpressing cells, using anti-BiP, anti-TbMORN1 and anti-GFP antibodies. Note that the membrane was cut into three strips for the immunoblot. Equal fractions (5%) were loaded in each lane. The overexpressed Ty1-TbMORN1 was predominantly extracted by digitonin and partitioned with the cytosolic GFP into the SN1 fraction (arrows 1). Very little of the remainder was subsequently extracted with non-ionic detergent into the SN2 fraction (arrows 2), with most partitioning into the cytoskeleton-associated P2 fraction. Three independent experiments were carried out using cells from three clones pooled together; an exemplary blot is shown. (C,D) Quantification of the immunoblots of the two-step fractionation. Data were compiled from three independent experiments, each using cells pooled from three separate clones. Bars show mean values + SD.

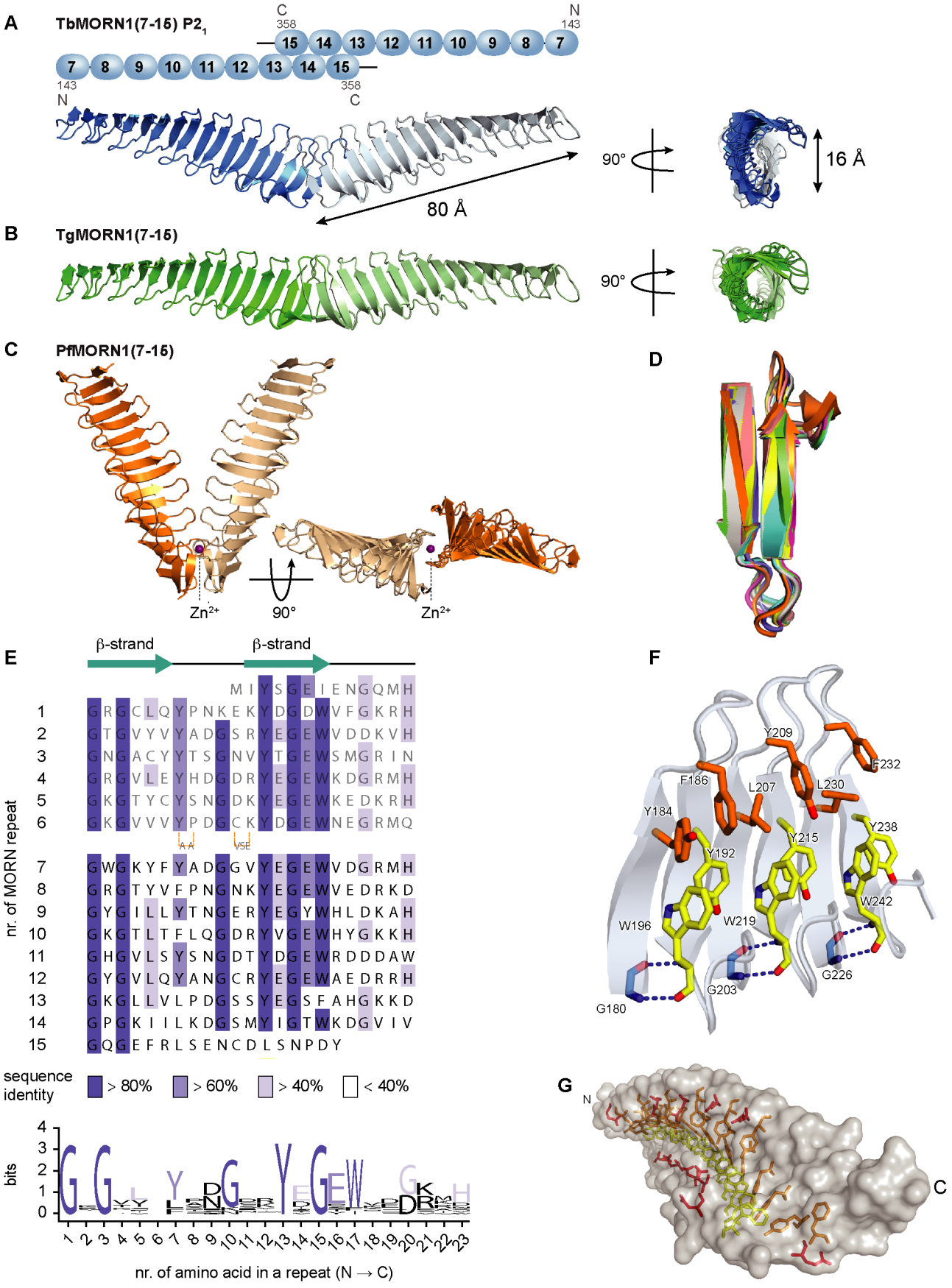


Figure 7. High-resolution structures of MORN repeat proteins and a structural redefinition of the MORN repeat. (A) Schematic depiction and crystal structure of the TbMORN1(7-15) dimer in its P21 crystal form. Amino acid numbers and N- and C-termini are indicated in the schematic. The crystal structure is shown in two orientations, with main dimensions indicated in Å. The structure contains 2 x 9 MORN repeats, and is an antiparallel homodimer with the subunits arranged in a splayed tail-to-tail configuration. The secondary structure consists of exclusively antiparallel beta-strands and peripherally positioned loops, which together form a longitudinal gutter through the middle of the protein. (B) Crystal structure of TgMORN1(7-15) shown in two orientations. The number of MORN repeats and the configuration is the same as for TbMORN1(7-15) in panel A. (C) Crystal structure of PfMORN1(7-15) shown in two orientations. The bound zinc ion is labelled. The structure contains 2x 9 MORN repeats, again in tail-to-tail configuration but with an overall V-shaped arrangement. (D) Alignment of all 9 TbMORN1(7-15) MORN repeats in the crystal structure reveals a high level of structural conservation. (E) A revised consensus MORN repeat sequence, based on the crystal structures. A new alignment of the MORN repeats in TbMORN1 is shown. Repeats 7-15 are present in the crystal structure; repeats 1-6 are inferred. Conservation of sequence identity is indicated by colour intensity. Each MORN repeat consists of a β-hairpin, built up of two 6-residue β-strands connected by a 5-residue loop. The β-hairpin is followed by a 6-residue loop that connects to the next MORN repeat. The new 23-residues long consensus MORN repeat starts with the GxG motif. (F) The tertiary structure of individual MORN repeats is stabilised by hydrogen bonds between the first G of the GxG motif and the W from the YEGEW motif. MORN repeat arrays are further stabilised by aromatic stacking between the highly conserved aromatic residues in the YEGEW and LxY motifs, and by T-shaped π-stacking interactions of the highly conserved Y of the YEGEW motif, which is sandwiched between the W residue of its own motif, and the W residue in the next YEGEW motif. (G) A single TbMORN1(7-15) subunit viewed at an oblique angle. The residues of the YEGEW and LxY motifs involved in aromatic stacking line the surface of the longitudinal gutter running through the middle of the protein.

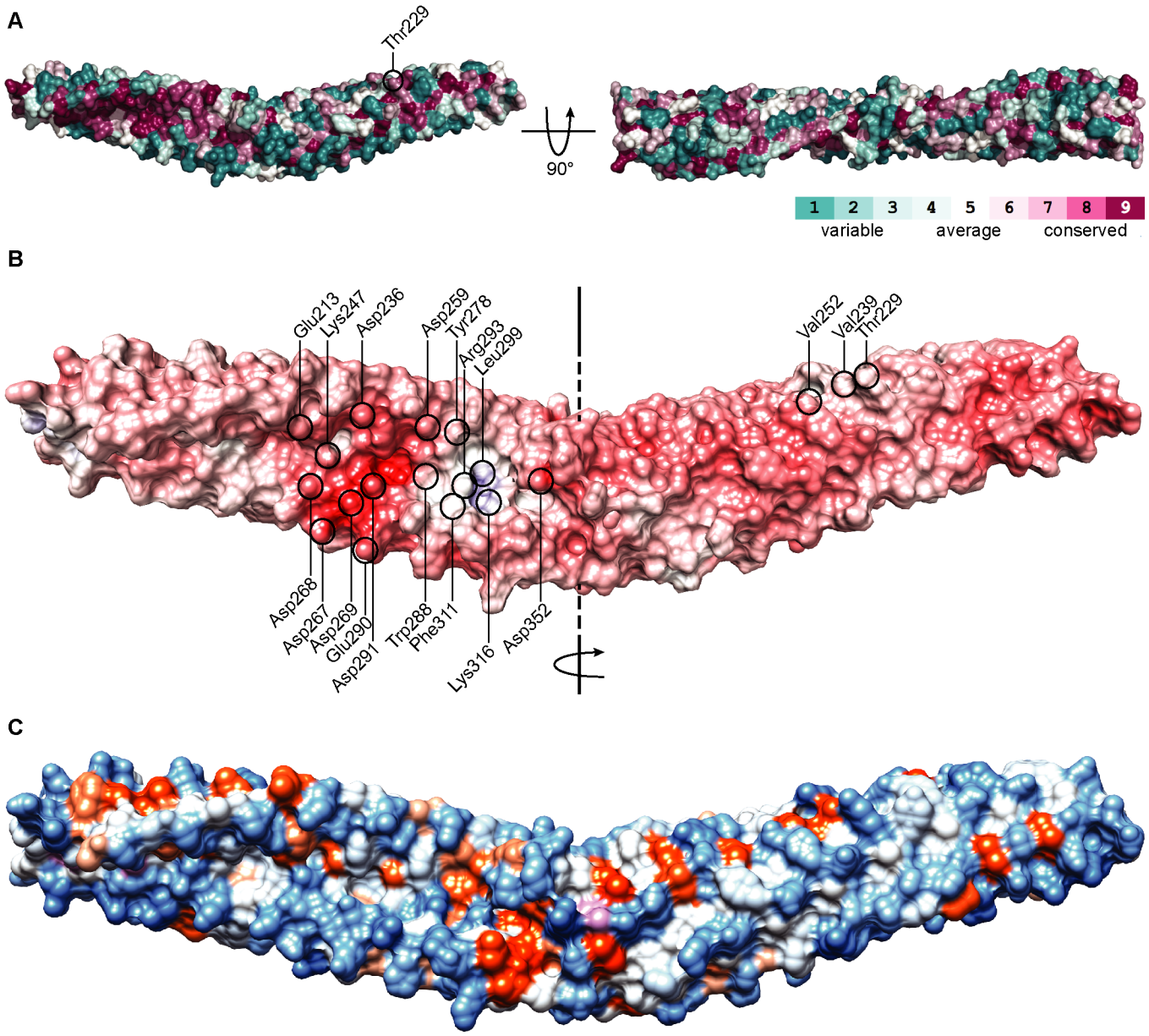


Figure 8. Conservation and properties of residues in TbMORN1(7-15). (A) Conservation map of the TbMORN1(7-15) $P2_1$ crystal structure reveals a highly conserved stretch of residues along the gutter. The structure is shown in two orientations, with residues colour-coded according to the level of conservation. (B) An electrostatic map of TbMORN1(7-15) $P2_1$. Colour scale: red = -13 kT; blue = +13 kT. Individual residues contributing to its surface electrostatics are labelled, namely those of the two negatively-charged loops building up a negative patch inside the gutter, and the residues contributing to a small positively-charged region close to the dimer interface. (C) Hydrophobic map of TbMORN1(7-15) $P2_1$. Colour scale: blue = hydrophilic; orange = hydrophobic, pink = methionine residues.

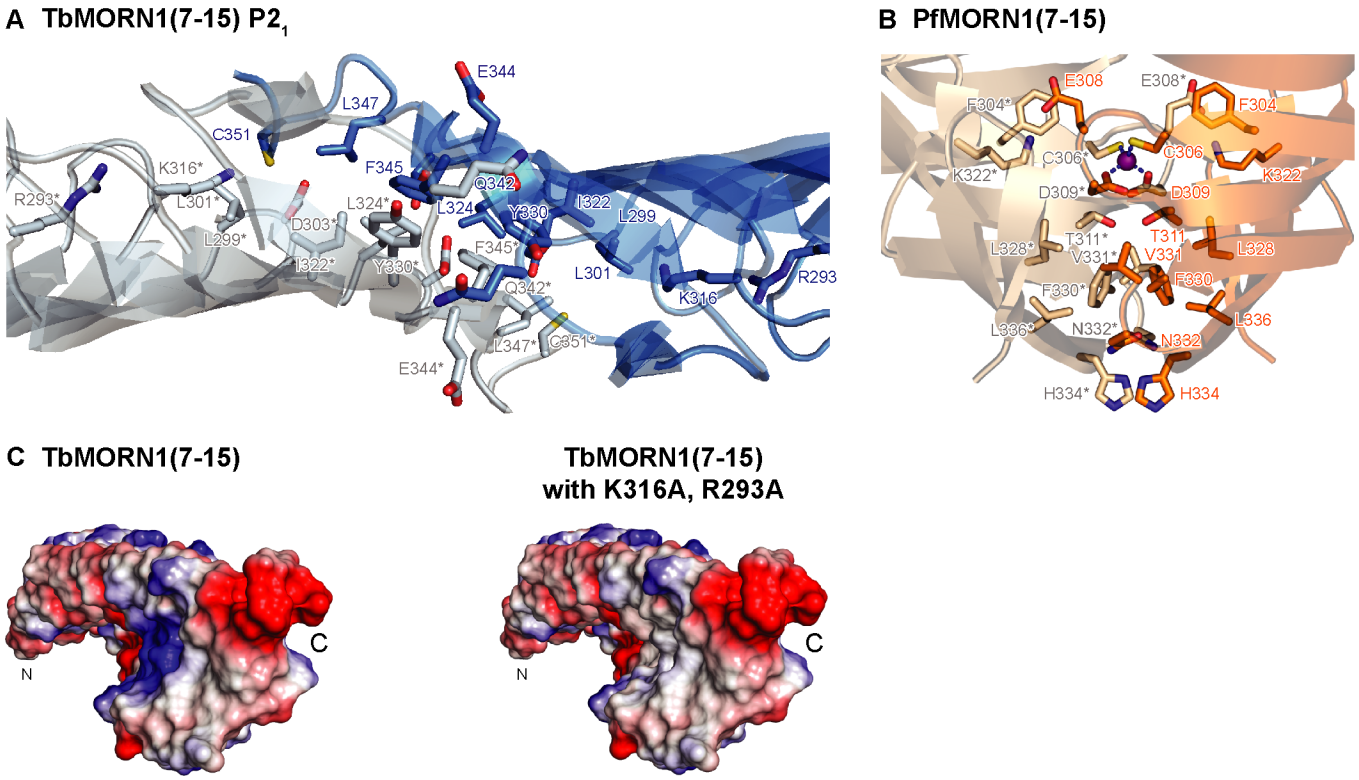


Figure 9. Dimerisation interfaces of TbMORN1(7-15) and PfMORN1(7-15). (A) The dimerisation interface of TbMORN1(7-15) $P2_1$ involves residues from MORN repeats 12-15, which stabilised the interface via aromatic π-stacking (Tyr330 and Phe345 from the respective subunits), hydrophobic interactions (Leu301, Leu347, Ile339, Ile322, Leu324), and additionally via hydrogen bonding interactions at the edges of the dimer interface. In comparison to the TbMORN1(7-15) C2 crystal structure, there are no disulphide bridges stabilising the dimerisation interface of TbMORN1(7-15) $P2_1$. (B) The dimerisation interface of the V-shaped PfMORN1(7-15) dimer is smaller and is additionally stabilised by the incorporation of a structural Zn^{2+} ion, which is tetrahedrally coordinated by Cys306 and Asp309 residues from each respective subunit. Thr311 holds the side chain of Asp309 in the appropriate orientation. The dimer interface is additionally stabilised by symmetric hydrogen bonding between the Thr311 pair, aromatic stacking between the Phe330 pair, a hydrophobic cluster formed by Leu328, Val331 and Leu336, two salt bridges between Lys322 and Glu308 from the respective subunits, anion-π interaction of a side chain of Glu308 with Phe304, and a combination of aromatic stacking (His334 pair) and polar interactions (His334, Asn332) at the vertex of the dimer. (C) An electrostatic map calculated for a single subunit of TbMORN1(7-15). The structure on the right shows the predicted effect of two point mutations, R293A from MORN repeat 13, and K316A from MORN repeat 14. The mutations are expected to result in the loss of a postively-charged patch close to the dimer interface, and consequently disrupt the dimerisation of the TbMORN1(2-15) constructs.

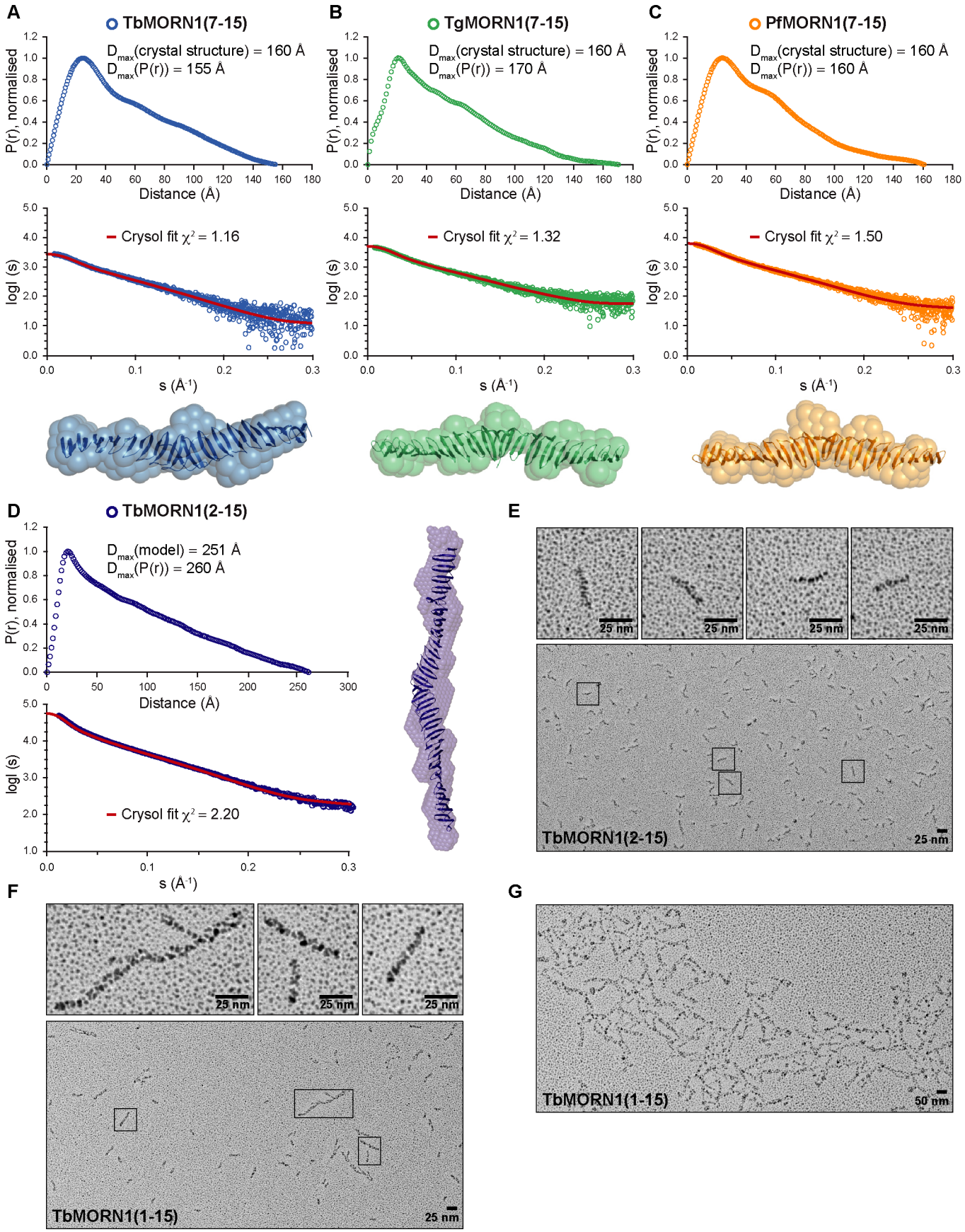
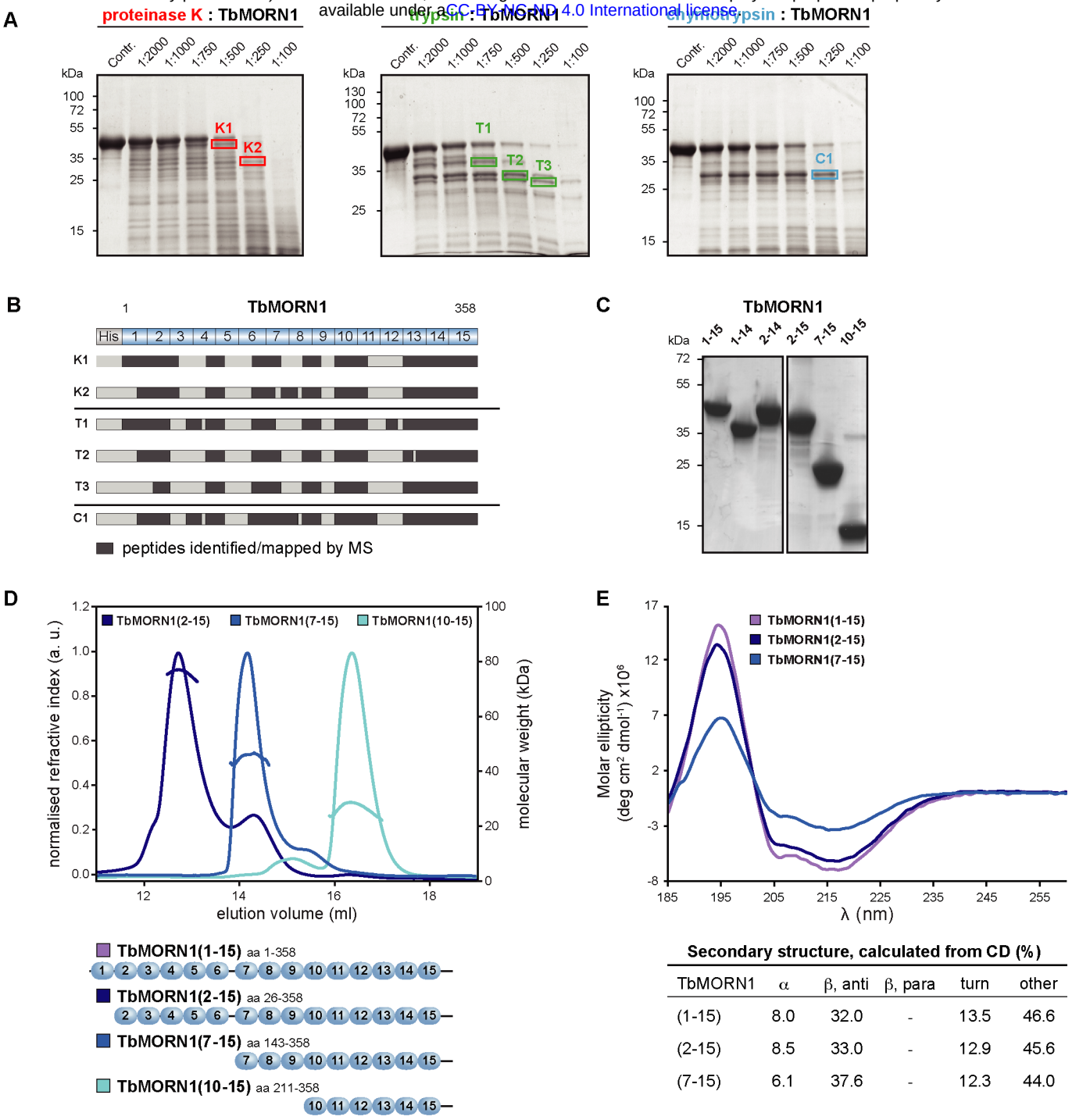


Figure 10. MORN1 proteins form extended dimers in solution. (A-D) SAXS experiments on TbMORN1(7-15) (A), TgMORN1(7-15) (B), PfMORN1(7-15) (C), and TbMORN1(2-15) (D). For each respective protein, the results include a P(r) plot with derived experimental D_{max} value compared with a D_{max} value derived from the structure, an experimental SAXS scattering data with a fit calculated by the Crysol programme, and a SAXS-based ab initio molecular envelope. In the case of TbMORN1(2-15), the theoretical D_{max} value was derived from a structural model, which was generated by spiking the TbMORN1(7-15) structure with additional structures of individual TbMORN1(7-15) subunits. (E-G) EM with rotary shadowing of TbMORN1(2-15) and full-length TbMORN1. (E) TbMORN1(2-15) forms a homogenous population of extended dimers of approximately 25 nm in length (see insets for individual examples). (F) Full-length TbMORN1 is heterogeneous and includes rare filaments of 175-200 nm in length (first inset) and individual dimers (second and third inset). (G) Large oligomers of full-length TbMORN1 assembled in a mesh-like structure. Magnification, 71,000x; scale bars, 25 nm, 50 nm as indicated. n (independent replicates) = 2, n (biological replicates) = 2.



Secondary structure, calculated from CD (%)

TbMORN1	α	β, anti	β, para	turn	other
(1-15)	8.0	32.0	-	13.5	46.6
(2-15)	8.5	33.0	-	12.9	45.6
(7-15)	6.1	37.6	-	12.3	44.0

Figure S1. Low-resolution structural analysis of TbMORN1. (A) Full-length TbMORN1 with an N-terminal HisTag was subjected to limited proteolysis using proteinase K (red), trypsin (green), and chymotrypsin (cyan) at protease:protein ratios (w/w) as indicated on each panel. Samples were resolved by SDS-PAGE and selected bands (labelled boxes) corresponding to proteolytic products were excised and analysed by mass spectrometry. Control (Contr.) corresponds to protein without protease treatment. (B) Mass spectrometry analysis of the excised proteolytic products indicated in panel A. Peptides identified and mapped by mass spectrometry are shown as dark grey boxes; a schematic of the full-length construct is shown above, with individual MORN repeats labelled. Note that the proteolytic products show progressive degradation from their N-termini, while the C-terminal part is stable. (C) Coomassie-stained SDS-PAGE gel showing purified recombinant TbMORN1 truncations. (D) SEC-MALS traces of TbMORN1 (2-15), (7-15), and (10-15). Chromatographic separation was done using a Superdex 200 Increase 10/300 GL column. The three proteins all eluted as dimers. Schematics are shown underneath. (E) Far-UV CD profiles of TbMORN1, TbMORN1(2-15) and (10-15). A positive peak at 195 nm and a negative one at 218 nm demonstrated that the constructs are all β-proteins. The secondary structure content predictions for each construct were calculated in BeStSel and are shown below the CD graph.

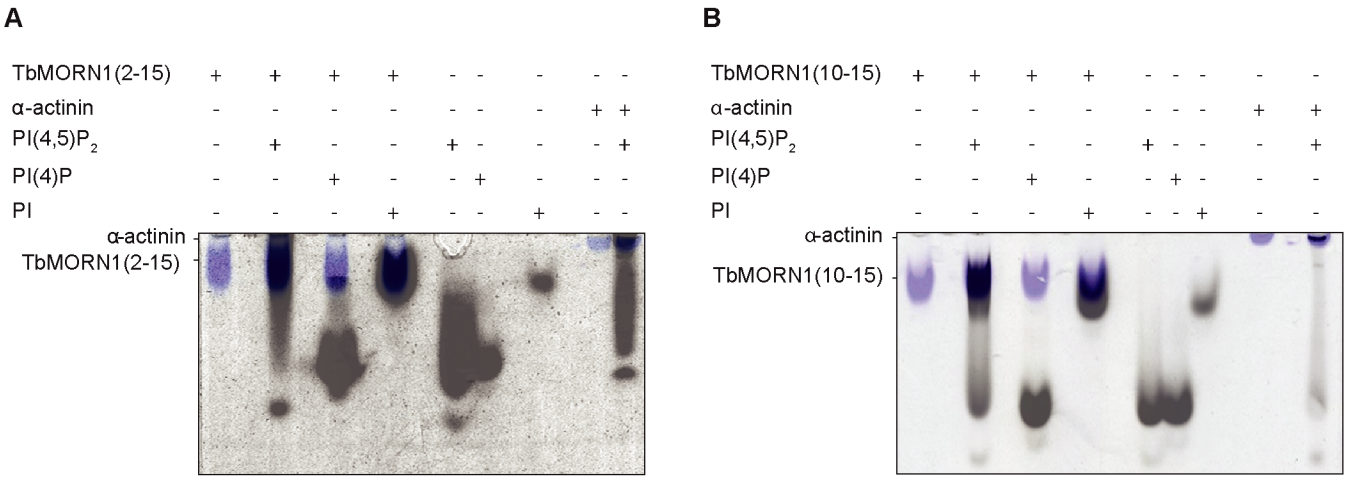


Figure S2. TbMORN1 interacts with PI(4,5)P_2 in native gel bandshift assays. (A) Native gel electrophoresis of TbMORN1(2-15) and (B) TbMORN1(10-15) in the presence and absence of PI(4,5)P_2, PI(4)P and PI, all labelled with Bodipy TMR fluorescent dye. α-actinin served as a positive control of PI(4,5)P_2 binding. Data obtained from two independent experiments, each using a different biological replicates.

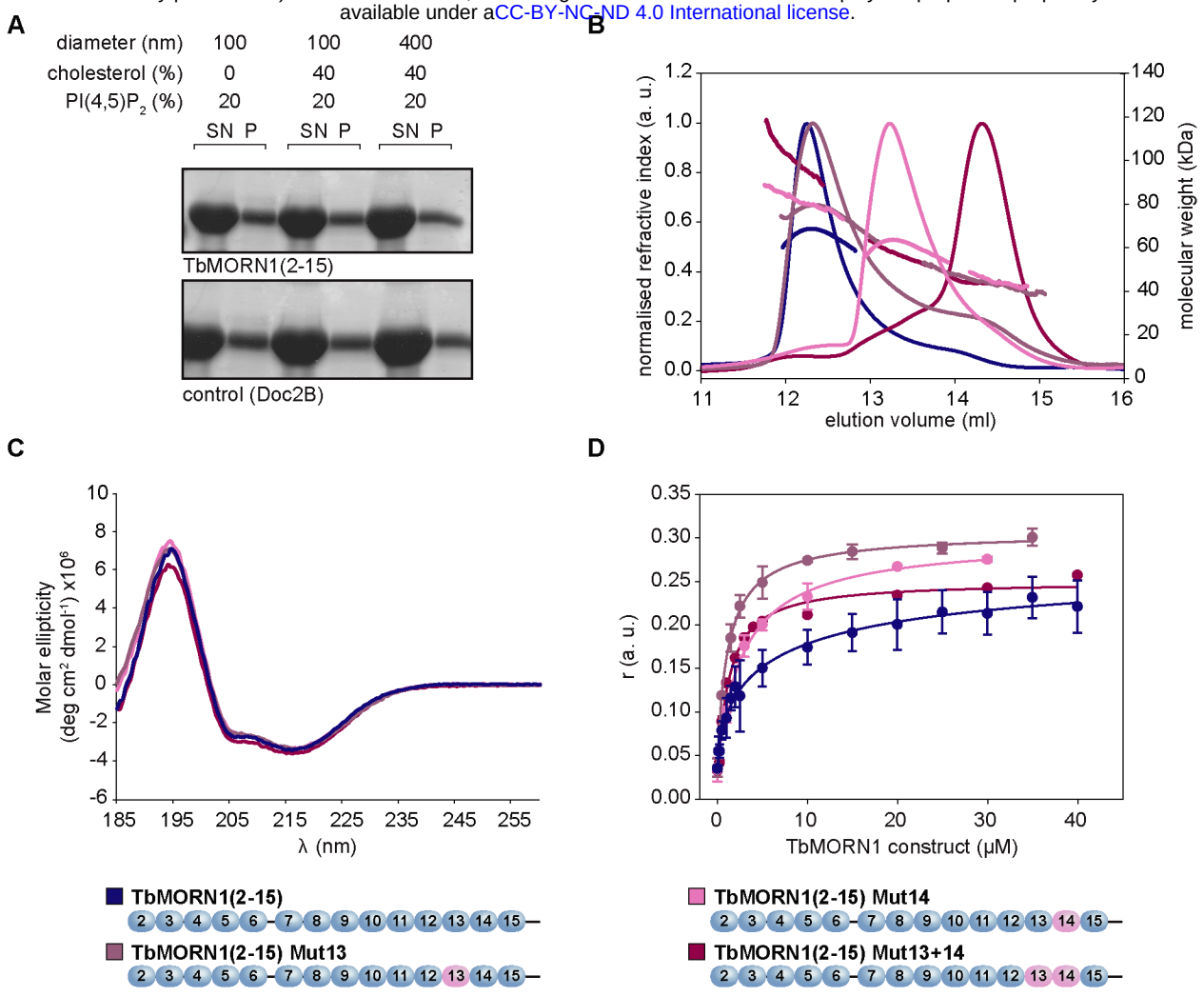


Figure S3. Mutagenesis of putative PI(4,5)P_2 binding sites in TbMORN1 has no effect on binding. (A) Liposome co-sedimentation assay performed on TbMORN1(2-15) in the presence of POPC liposomes containing 20% of porcine brain PI(4,5)P_2 and 0 or 40% cholesterol. The excess cholesterol was expected to promote local high concentrations of PI(4,5)P_2 on the surface of the liposomes. To assay for the effect of curvature, two batches of liposomes containing 20% PI(4,5)P_2 and 40% of cholesterol were tested, with the diameter of the liposomes being either 100 or 400 nm. No significant co-sedimentation of TbMORN1(2-15) and PI(4,5)P_2-containing liposomes was observed. The positive control, 10 μM Doc2B was predominantly found in the pellet (P) fractions. (B) SEC-MALS profiles of TbMORN1(2-15) and its mutagenised variants. Residues comprising the putative PI(4,5)P_2-binding sites in MORN repeats 13 and 14 were mutated to alanines. Mutagenesis of repeat 13 (Mut13) did not result in any change to the dimeric status of the protein. However, mutagenesis of repeat 14 (Mut14) resulted in a mixture of monomers and dimers being eluted, while mutagenesis of both repeats (Mut13+14) resulted in monomeric protein. (C) Far-UV CD profiles of TbMORN1(2-15) and its putative PI(4,5)P_2-binding mutants. The constructs remained β-proteins despite the site-directed mutagenesis. (D) Fluorescence anisotropy measurements of TbMORN1(2-15) and its putative PI(4,5)P_2-binding mutants, measured in the presence of 0.1 μM BODIPY TMR-PI(4,5)P_2. All constructs showed good interaction with the fluorophore-conjugated PI(4,5)P_2.

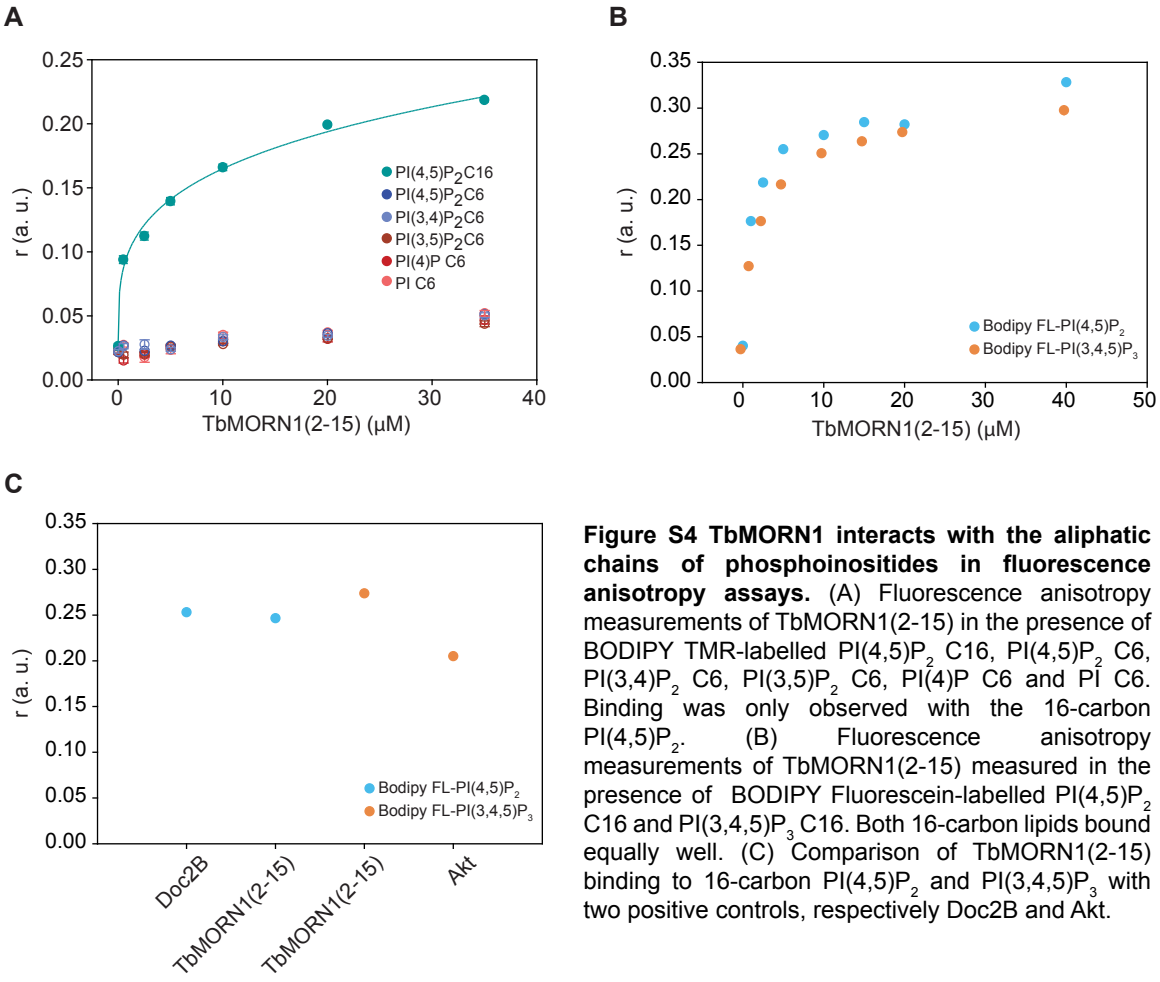


Figure S4 TbMORN1 interacts with the aliphatic chains of phosphoinositides in fluorescence anisotropy assays. (A) Fluorescence anisotropy measurements of TbMORN1(2-15) in the presence of BODIPY TMR-labelled $PI(4,5)P_2$ C16, $PI(4,5)P_2$ C6, $PI(3,4)P_2$ C6, $PI(3,5)P_2$ C6, PI(4)P C6 and PI C6. Binding was only observed with the 16-carbon $PI(4,5)P_2$. (B) Fluorescence anisotropy measurements of TbMORN1(2-15) measured in the presence of BODIPY Fluorescein-labelled $PI(4,5)P_2$ C16 and $PI(3,4,5)P_3$ C16. Both 16-carbon lipids bound equally well. (C) Comparison of TbMORN1(2-15) binding to 16-carbon $PI(4,5)P_2$ and $PI(3,4,5)P_3$ with two positive controls, respectively Doc2B and Akt.

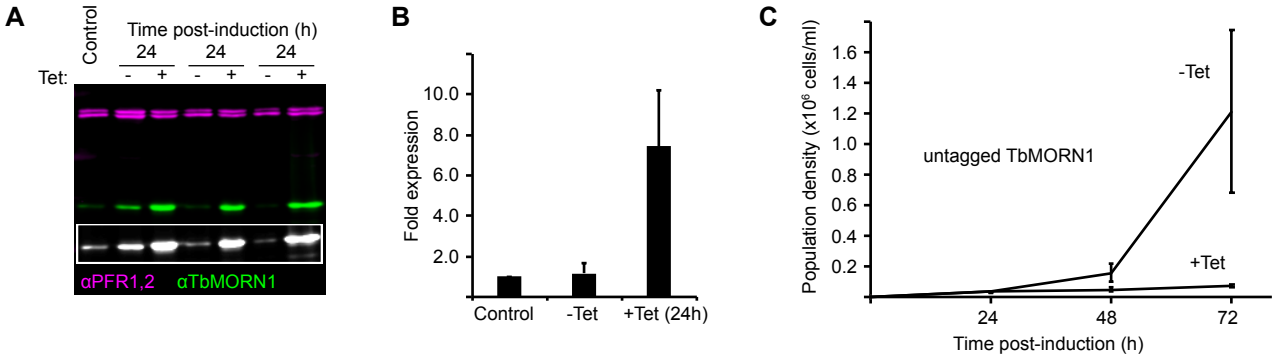


Figure S5. Overexpression of untagged TbMORN1 causes a dominant negative phenotype. (A) Inducible overexpression of untagged TbMORN1. Immunoblot of whole-cell lysates from three separate clones overexpressing untagged TbMORN1 from an ectopic locus. TbMORN1 was detected using anti-TbMORN1 antibodies; anti-PFR1,2 antibodies were used as a loading control. Inset shows a greyscale image of the TbMORN1 channel with enhanced levels so the endogenous protein is visible. Three separate clones were assayed, each in three independent experiments; an exemplary blot is shown. One of the three clones appeared to have slightly leaky expression, with TbMORN1 levels in the -Tet condition being higher than controls. (B) Quantification of overexpression. TbMORN1 levels in control, uninduced (-Tet) and induced (+Tet) were normalised relative to the loading control and expressed relative to the control cells. Approximately 7-fold overexpression was achieved relative to control cells. Data were obtained from blots using 3 separate clones, each induced in 3 independent experiments. Bars show mean + SD. (C) Overexpression of untagged TbMORN1 is deleterious. Uninduced (-Tet) and TbMORN1 overexpressing (+Tet) cells were assayed at 24 h intervals in a 3-day timecourse. Data were obtained from blots using 3 separate clones, each induced in 3 independent experiments. Mean +/- SD.

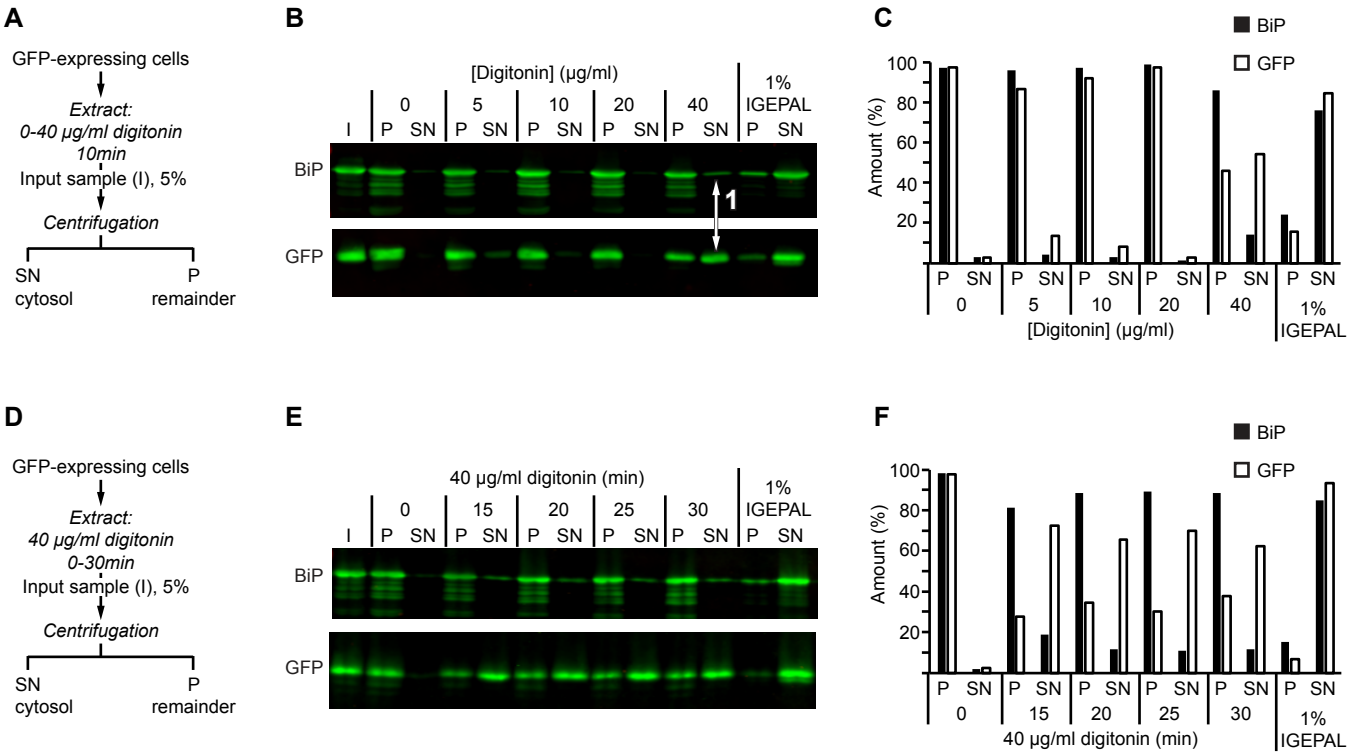
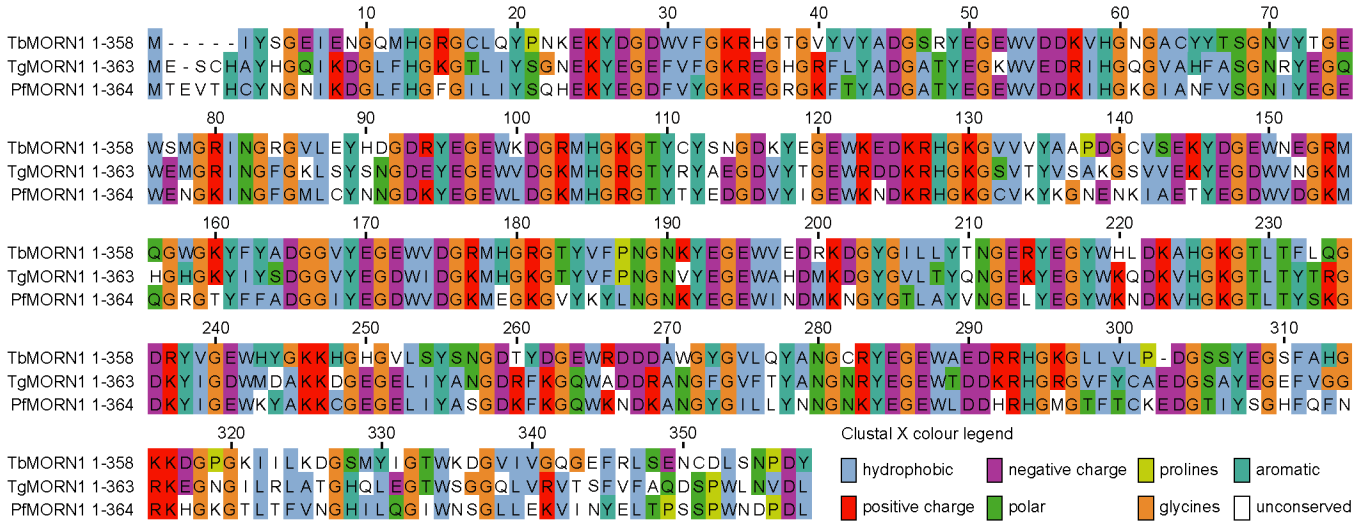


Figure S6. Optimisation of digitonin fractionation. (A) Schematic of the one-step fractionation scheme. Cell expressing cytosolic GFP were incubated in increasing concentrations of digitonin for 10 min prior to separation of fractions by centrifugation. Equal fractions (5%) were then blotted with antibodies specific for GFP and the endoplasmic reticulum chaperone BiP. (B) At 40 µg/ml digitonin, good solubilisation of GFP is achieved with only negligible solubilisation of BiP (arrow 1). Both proteins are efficiently solubilised using 1% IGEPAL as a positive control. Multiple independent experiments were carried out; an exemplary blot is shown. Note that the membrane was cut into strips prior to blotting, but the samples shown are from the same experiment. (C) Quantification of the immunoblot shown in B. (D, E, F) As per panels A-C, but with a constant 40 µg/ml digitonin concentration and considering the effect of varying incubation time (0-30 min).

A

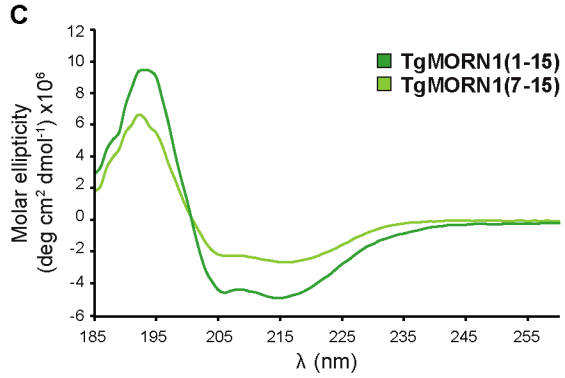


B

Comparison of MORN1 proteins and % of their sequence identity

T. brucei / P. falciparum	54
T. brucei / T. gondi	57
P. falciparum / T. gondii	65

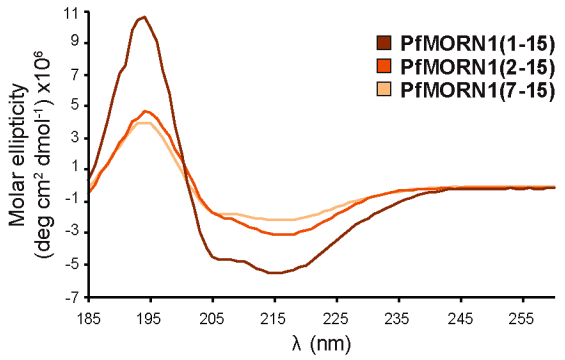
C



Secondary structure, calculated from CD (%)

TgMORN1	α	β, anti	β, para	turn	other
(1-15)	5.3	30.8	0.8	13.0	50.1
(7-15)	4.8	42.3	-	11.8	41.1

D



Secondary structure, calculated from CD (%)

PfMORN1	α	β, anti	β, para	turn	other
(1-15)	6.4	29.8	1.9	13.8	48.1
(2-15)	6.2	35.7	-	13.4	44.7
(7-15)	6.6	32.6	1.0	12.0	47.8

Figure S7. Comparison of MORN1 proteins, and secondary structure analysis of apicomplexan MORN1s. (A) Amino acid sequence alignment of MORN1 proteins from *Trypanosoma brucei, Toxoplasma gondii* and *Plasmodium falciparum*. The number of amino acids in each protein is indicated, amino acid numbers in the alignment are those for TbMORN1. The alignment is coloured according to the amino acid properties. (B) Pairwise comparison of percentage sequence identity between the three proteins. (C) Far-UV CD measurements obtained for TgMORN1(1-15) and TgMORN1(7-15). The secondary structure content predictions for each of the measured proteins were calculated in BeStSel and are shown below the CD graph. (D) As (C), but PfMORN1(1-15), (2-15) and (7-15). Like TbMORN1, TgMORN1 and PfMORN1 are also all-β proteins.

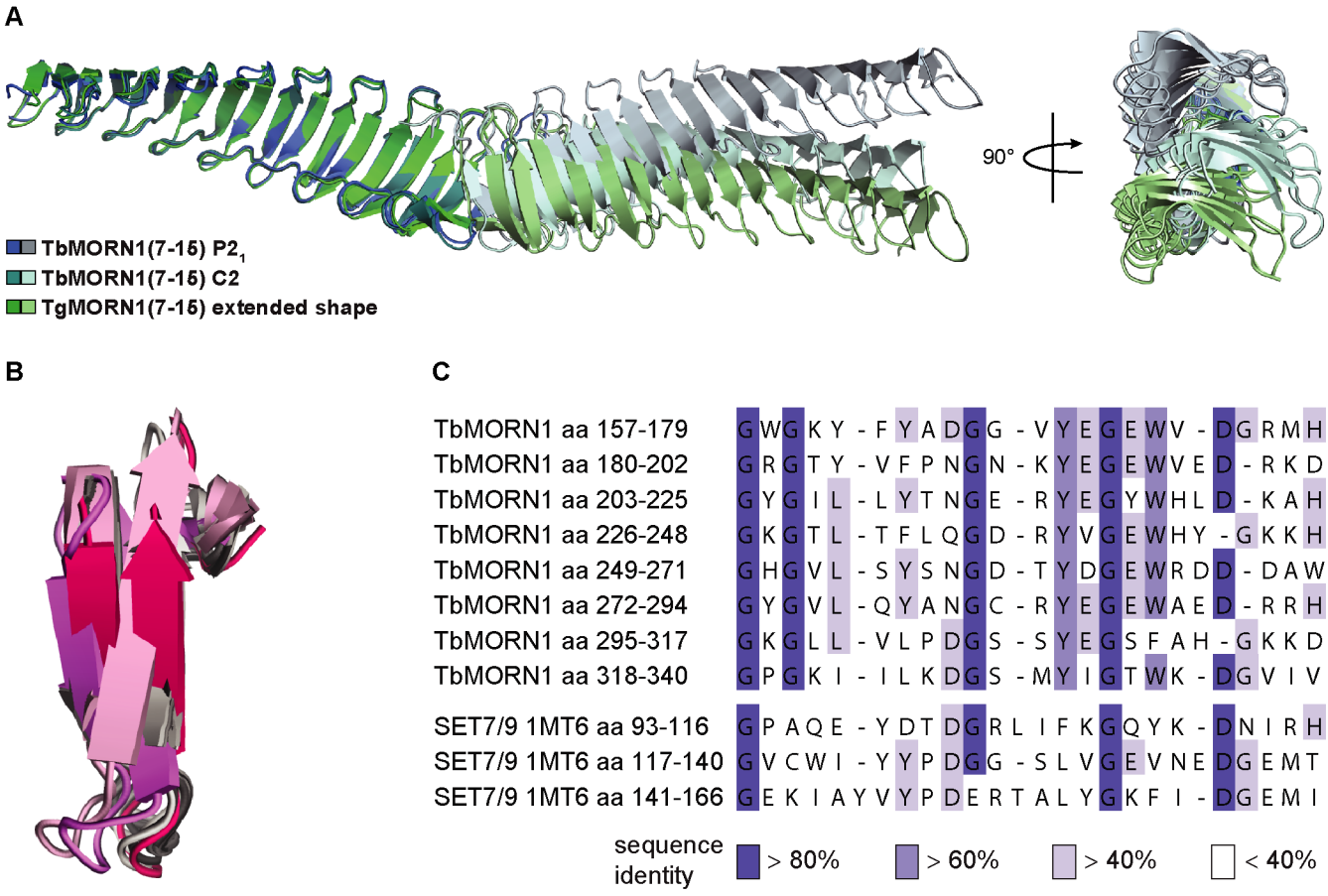
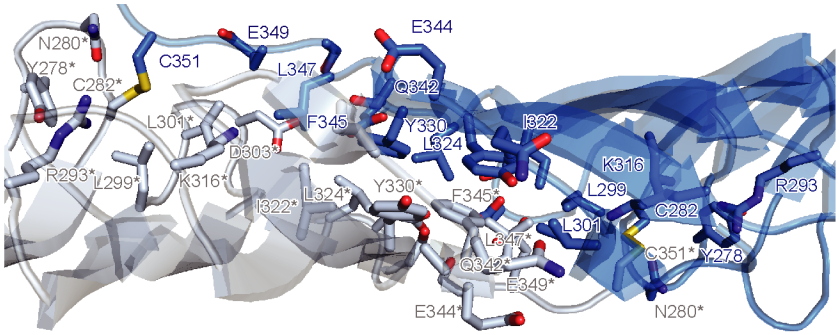
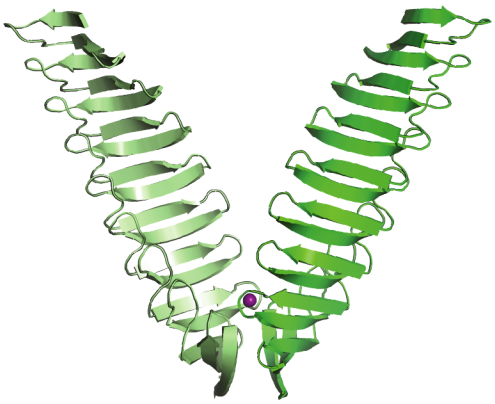


Figure S8. Comparision of MORN1 extended dimers, and with SETD7. (A) TbMORN1(7-15) P2$_1$, TbMORN1(7-15) C2, and TgMORN1(7-15) extended dimers superimposed on each other, and displayed in two orientations. In contrast to the other two proteins, the P2$_1$ crystal structure of TbMORN1(7-15) displays a bend of approximately 30° among the subunits. (B) TbMORN1(7-15) MORN repeats superimposed on three MORN repeats from SETD7 (SET7/9). Alignment of the three MORN repeats from SETD7 with MORN repeat 7 from the TbMORN1(7-15) crystal structure over 22-23 aligned C-atoms yielded rmsd values of 2.3, 1.5 and 1.9 Å respectively. (C) Sequence alignment of MORN repeats from the TbMORN1(7-15) crystal structure with three MORN repeats from SETD7. The first Gly residue is conserved in all MORN repeats of TbMORN1(7-15) and SETD7 structures.

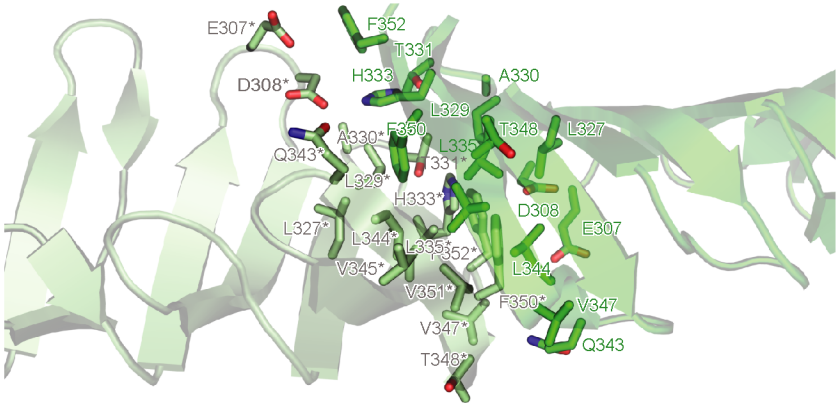
A TbMORN1(7-15) C2



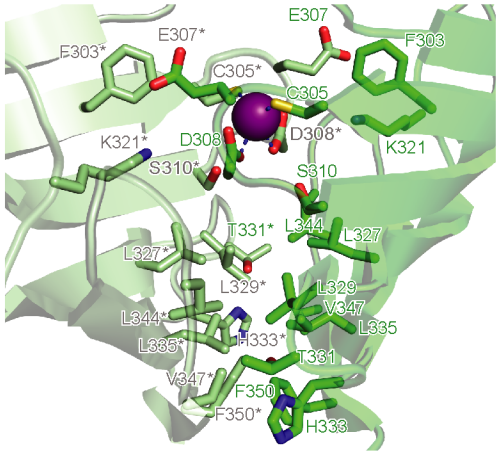
B TgMORN1(7-15)



C TgMORN1(7-15) extended dimer



D TgMORN1(7-15) V-shaped dimer



E SAXS - Kratky plots

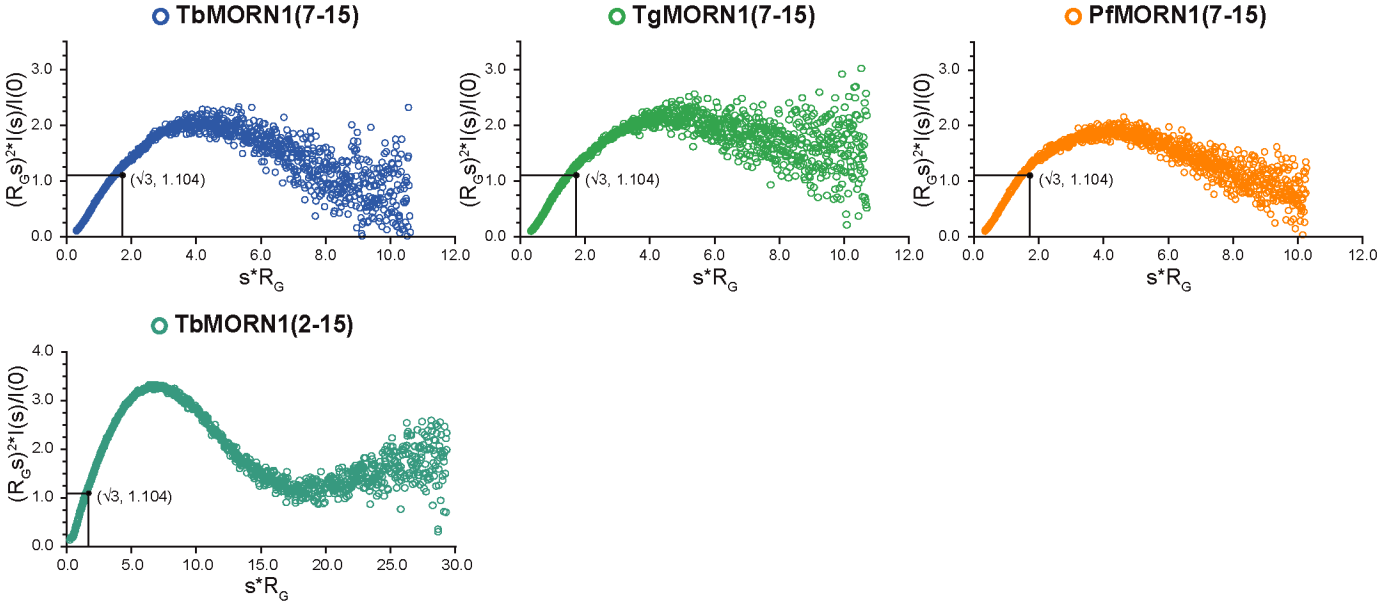


Figure S9. MORN dimer interfaces and SAXS analysis of proteins in solution. (A) Dimer interface of TbMORN1(7-15) C2 crystal form. In comparison to the $P2_1$ form, the dimer interface of C2 structure is broader, and is additionally stabilised by two disulphide bridges formed between Cys351 at the C-terminus of repeat 15 and Cys282 from the **β**-hairpin loop of repeat 12. (B) Crystal structure of the TgMORN1(7-15) V-shaped dimer, incorporating Zn2+ in its dimerisation interface. (C) Dimer interface of the TgMORN1(7-15) extended dimer, which utilises residues from MORN repeats 13-15. In contrast to TbMORN1(7-15), where the dimerisation interface is centred around aromatic stacking, a hydrophobic core plays a crucial role in the dimer interface of extended TgMORN1(7-15). Leu327, Leu329, Leu335, Leu344, Val345, Val347, Phe350 and Phe352 are part of this hydrophobic core. The dimer is stabilised by a single salt bridge formed between Asp308 of one, and His333 of the second subunit. This salt bridge is further stabilised by two hydrogen bonds between the main-chain nitrogen of Val345 and a carbonyl oxygen of Val347 of respective subunits. (D) Dimer interface of the TgMORN1(7-15) V-shaped dimer. Cys305 and Asp308 incorporate a structural Zn2+ ion, which stabilises the somewhat smaller dimerisation interface of this protein. Although its dimerisation interface is very similar to that of PfMORN1(7-15), it lacks the aromatic core of PfMORN1(7-15). The latter is replaced by a series of unique aromatic stacking interactions at the protein´s vertex, these being contributed by a pair of Phe350 residues, sandwiched between a pair of His333 residues. (E) Kratky plots derived from SAXS analysis of TbMORN1(7-15), TgMORN1(7-15), PfMORN1(7-15), and TbMORN1(2-15). The shape of the plots suggests an elongated shape of the dimers in solution.

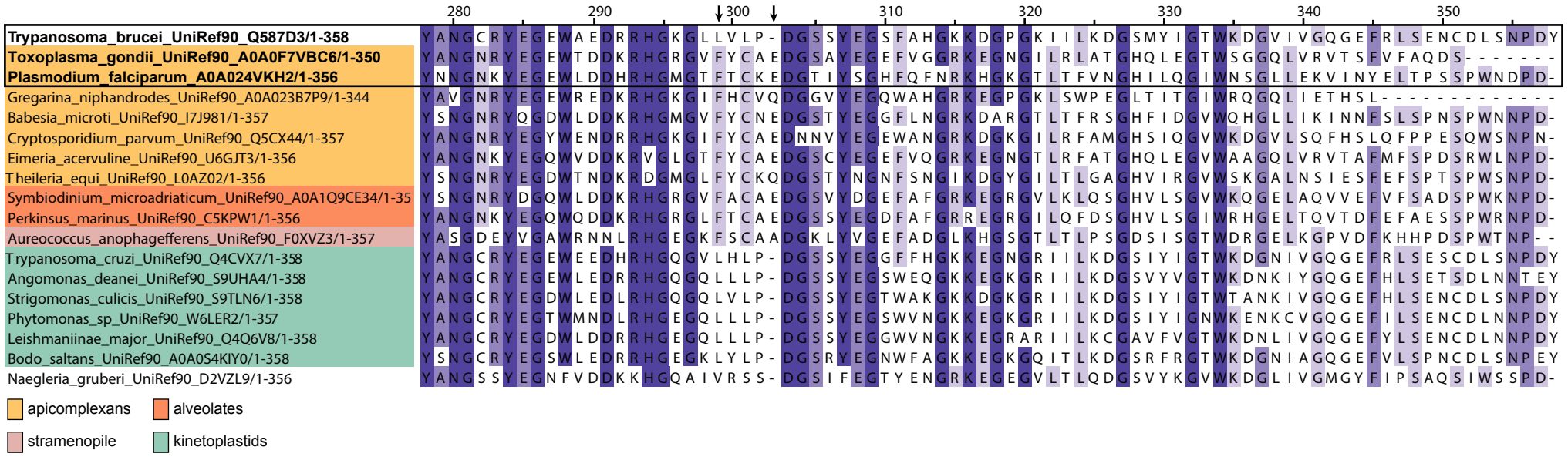


Figure S10. Conservation of residues for formation of V-shaped dimers in Apicomplexa and related clades. Amino acid sequence alignment of the C-termini of TbMORN1, TgMORN1, PfMORN1, and fifteen other MORN repeat-containing proteins from related taxa. Amino acid numbers are given according to the TbMORN1 protein, and the three proteins with experimentally-determined high-resolution structures are shown in bold within the black box. Essential for formation of a V-shaped dimer are a coordinating Cys residues and an anion-π interaction pair. In TbMORN1, the coordinating Cys residue has been substituted for Leu (Leu 301). Similarly the Phe and Glu residues of the anion-π interaction pair (indicated with black arrows) have been substituted for Leu (Leu299) and are not present (deletion after Pro302) respectively. This supports the conclusion that TbMORN1 exists only in the extended form, while the apicomplexan proteins and those from related clades are probably capable of adopting both extended and V-shaped conformations.

Table 1. X-ray data collection statistics

	PfMORN (SMet)	**PfMORN**	**TbMORN**	**TbMORN**	**TgMORN**	**TgMORN (V-shape)**
Source	ID29	ID29	ESRF ID23-1	ESRF ID23-1	ESRF ID30B	ESRF ID30B
Wavelength (Å)	0.979	0.976	0.979	1.89	0.967	1.0
Resolution (Å)	47.47-2.5 (2.59-2.5)	46.3-2.14 (2.2-2.14)	48.28-2.35 (2.43-2.35)	48.14-2.53 (2.65-2.53)	48.92-2.90 (3.08-2.90)	49.45-2.50 (2.60-2.50)
Space group	*C*222$_1$	*C*222$_1$	*P*2$_1$	*C*2	*P*4$_3$2$_1$2	*P*6$_2$22
Unit cell (Å, º)	a=57.33, b= 79.18, c= 94.42	a=57.33, b= 79.18, c= 94.42	a=69.04, b=27.63, c=114.54; β=101.83	a=192.88, b=49.74, c=41.98	a=b=53.92, c=348.85	a=b=205.86, c=40.58
Molecules (a.u.)	1	1	2	2	2	1
Unique reflections	7832(762)	12148 (974)	17460 (1247)	12419(975)	12331 (1819)	18162 (1960)
Completeness (%)	99.5(98.3)	99.6(98.1)	95.4 (71.5)	93.1(60.0)	99.1 (95.4)	99.7 (97.9)
R_{merge}[b]	0.059(0.301)	0.037 (1.283)	0.092 (0.466)	0.116 (1.365)	0.140 (2.252)	0.198 (2.116)
R_{meas}[c]	0.062(0.314)	0.042 (1.510)	0.100 (0.532)	0.146 (1.893)	0.158 (2.536)	0.204 (2.177)
CC(1/2)	0.999(0.983)	1.000 (0.467)	0.998 (0.897)	0.995 (0.303)	0.999 (0.772)	0.999 (0.426)
Multiplicity	13.0 (12.9)	4.2 (3.2)	6.1 (4.1)	6.2 (3.7)	8.4 (8.0)	17.9 (18.1)
I/sig(I)	30.3 (8.1)	17.5 (0.8)	13.4 (2.8)	9.3 (0.9)	11.8 (0.5)	14.6 (1.6)
B_{Wilson} (Å^2)	57.3	58.0	18.74	36.7	22.97	42.3
R_{work}[e]/R_{free}[f]		23.0/26.4	23.2/25.6	22.5/28.2	31.8/33.8	20.1/23.9
r.m.s.d. bonds (Å)		0.003	0.004	0.011	0.0084	0.007
r.m.s.d. angles (°)		0.6	1.24	1.612	1.417	0.811

[a] Values in parentheses are for the highest resolution shell.

[b] $$R_{merge} = \frac{\sum_{hkl}\sum_{i=1}^{N}\left|I_{i(hkl)} - \bar{I}_{(hkl)}\right|}{\sum_{hkl}\sum_{i=1}^{N} I_{i(hkl)}}$$

[c] $$R_{meas} = \frac{\sum_{hkl}\sqrt{N/(N-1)}\sum_{i=1}^{N}\left|I_{i(hkl)} - \bar{I}_{(hkl)}\right|}{\sum_{hkl}\sum_{i=1}^{N} I_{i(hkl)}}$$

where $\bar{I}_{(hkl)}$ is the mean intensity of multiple $I_{i(hkl)}$ observations of the symmetry-related reflections, N is the redundancy

[e] $$R_{work} = \frac{\sum\left||F_{obs}| - |F_{calc}|\right|}{\sum|F_{obs}|}$$

[f] R_{free} is the cross-validation R_{factor} computed for the test set of reflections (5%) which are omitted in the refinement process.

Intermolecular cross-links in TbMORN1

chemical cross-linking agent	cross-linked amino acids and respective MORN repeats	best e-value	number of PSMs in the dimer
BS^3 (spacer arm length 11.4 Å)	aa 100 x aa 123 (repeat 5 x repeat 6)	7.5×10^{-13}	4
	aa 223 x aa 246 (repeat 10 x repeat 11)	5.7×10^{-22}	7
	aa 296 x aa 321 (repeat 13 x repeat 14)	1.8×10^{-11}	56
	aa 321 x aa 321 (repeat 14 x repeat 14)	6.4×10^{-5}	1
EDC (spacer arm length 0 Å)	aa 308 x aa 321 (repeat 14 x repeat 14)	1.1×10^{-8}	5
	aa 321 x aa 326 (repeat 14 x repeat 15)	1.9×10^{-9}	3

Table S1. Intermolecular contacts in TbMORN1. A summary of intermolecular cross-links in the TbMORN1 dimer as a result of chemical cross-linking with either BS^3 or EDC followed by mass spectrometry analysis. To remove low quality peptide-spectrum matches (PSMs), an additional e-Value cutoff of < 0.001 was applied. In order to distinguish intra- from inter-molecular chemical cross-links, results from monomers and dimers were compared. Cross-links were scored as intermolecular when there were: (1) minimally 3 peptide PSMs in dimer and (2) minimally 3-times more PSMs in dimer than in monomer. Results compiled from two independent experiments.

	Thermostability
TbMORN1 construct	T_m (°C) (DSF)
2-15	44.2 ± 0.3
2-15^{Mut13}	43.8 ± 0.3
2-15^{Mut14}	41.4 ± 0.5
2-15$^{Mut13+14}$	41.7 ± 0.3

Table S2: Analysis of TbMORN1(2-15) mutagenesis constructs.
Thermostability of TbMORN1(2-15) and its mutagenised derivatives with mutations in MORN repeats 13 and 14 was measured by differential scanning fluorimetry (DSF), given here as melting temperature (T_m) values (°C).

Data collection parameters				
Radiation source	Petra III (DESY, Hamburg, Germany)			ESRF (Grenoble, France)
Beamline	EMBL P12			BM29
Detector	Pilatus 2M			Pilatus 1M
Beam geometry (mm, FWHM)	0.12 x 0.20			0.10 x 0.20
Wavelength (nm)	0.12			0.099
Sample-detector distance (m)	3.1			2.867
Momentum transfer *s* range (nm^{-1})	0.01 – 4.0			0.04 – 5.0
Exposure time (s)	1 sec (SEC-SAXS mode) 0.05 sec (Batch mode)			1 sec (SEC-SAXS mode)
Overall Parameters				
	TbMORN1 (7-15)	**TgMORN1 (7-15)**	**PfMORN1 (7-15)**	**TbMORN1 (2-15)**
Buffer	Buffer 1*	Buffer 2**	Buffer 2**	Buffer 2**
Temperature (°C)	20	20	10	20
Working concentration (mg/ml) Concentration range measured	n.a.	n.a.	1 1-8	n.a.
R_g from Guinier approximation (nm)	4.1±0.04	4.0±0.03	3.8±0.1	6.5±0.6
R_g from PDDF (nm)	4.4±0.02	4.3±0.04	4.2±0.03	7.2±0.05
D_{max} (nm)	155	17.0	16.0	260
Molecular weight, I(0) (kDa)	n.a.	n.a.	52 ±5	n.a.
Molecular weight, offline RALS (kDa)	46±5	50±5	32±5	74±5
Molecular weight from DATMOV, (kDa)	60±5	58±5	46±5	71±5
Molecular weight from sequence (dimer, kDa)	52	49	46	70
Software employed				
Primary data reduction	SASFLOW			ESRF EDNA pipeline
Data processing	PRIMUS/Chromixs			
Calculation of theoretical data	Crysol			
Ab initio modelling	DAMMIN			DAMMIF
SASBDB accession code	SASDGA7	SASDGB7	SASDGC7	SASDG97

Buffer 1* 20 mM Tris, 200 mM NaCl, 2% glycerol, 0.5 mM DTT

Buffer 2** 20 mM Tris-HCl pH 7.5, 100 mM NaCl

Table S3. SAXS data collection summary.