



# Impact of inherent biases built into proteomic techniques: Proximity labeling and affinity capture compared

Received for publication, August 16, 2022, and in revised form, October 26, 2022. Published, Papers in Press, November 19, 2022.  
<https://doi.org/10.1016/j.jbc.2022.102726>

Claudia Maria do Nascimento Moreira<sup>1,2</sup>, Cristina D. Kelemen<sup>3</sup>, Samson O. Obado<sup>4</sup>, Farnaz Zahedifard<sup>3</sup>, Ning Zhang<sup>5</sup>, Fabiola B. Holetz<sup>2</sup>, Laura Gauglitz<sup>1</sup>, Bruno Dallagiovanna<sup>2</sup>, Mark C. Field<sup>5,6</sup>, Susanne Kramer<sup>1,\*</sup>, and Martin Zoltner<sup>3,\*</sup>

From the <sup>1</sup>Department of Cell and Developmental Biology, Biocenter, University of Würzburg, Würzburg, Germany; <sup>2</sup>Carlos Chagas Institute (ICC), FIOCRUZ/PR, Curitiba, Brazil; <sup>3</sup>Department of Parasitology, Faculty of Science, Charles University in Prague, Biocev, Vestec, Czech Republic; <sup>4</sup>The Rockefeller University, Laboratory of Cellular and Structural Biology, New York, New York, USA; <sup>5</sup>School of Life Sciences, University of Dundee, Dundee, United Kingdom; <sup>6</sup>Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic

Edited by Dennis Voelker

The characterization of protein–protein interactions (PPIs) is of high value for understanding protein function. Two strategies are popular for identification of PPIs direct from the cellular environment: affinity capture (pulldown) isolates the protein of interest with an immobilized matrix that specifically captures the target and potential partners, whereas in BioID, genetic fusion of biotin ligase facilitates proximity biotinylation, and labeled proteins are isolated with streptavidin. Whilst both methods provide valuable insights, they can reveal distinct PPIs, but the basis for these differences is less obvious. Here, we compare both methods using four different trypanosome proteins as baits: poly(A)-binding proteins PABP1 and PABP2, mRNA export receptor MEX67, and the nucleoporin NUP158. With BioID, we found that the population of candidate interacting proteins decreases with more confined bait protein localization, but the candidate population is less variable with affinity capture. BioID returned more likely false positives, in particular for proteins with less confined localization, and identified low molecular weight proteins less efficiently. Surprisingly, BioID for MEX67 identified exclusively proteins lining the inner channel of the nuclear pore complex (NPC), consistent with the function of MEX67, whereas the entire NPC was isolated by pulldown. Similarly, for NUP158, BioID returned surprisingly few PPIs within NPC outer rings that were by contrast detected with pulldown but instead returned a larger cohort of nuclear proteins. These rather significant differences highlight a clear issue with reliance on a single method to identify PPIs and suggest that BioID and affinity capture are complementary rather than alternative approaches.

Most proteins function as part of multisubunit complexes, and identification of protein–protein interactions (PPIs) is valuable for understanding function. Identification of PPIs can uncover a wide range of interactions, which include direct,

indirect, static, or dynamic binding. Furthermore, proteins can moonlight and engage in multiple different specific complexes, whereas complex composition can change in a temporal manner and/or a spatial manner.

Presently, there are two common methods used to identify PPIs in a cellular context: affinity capture (colloquially pulldown) (1) and proximity labeling (2). While evidence indicates that both methods provide valuable insights, there are considerable differences between them, both in technical requirements and consequently the interactome revealed. Each method has its adherents and while it would be fallacious to view one or the other approach as superior, it is unclear what choice of method implies in terms of the types of PPIs identified and hence critical assessment of a given dataset.

Affinity capture requires cell rupture with the lysate or extract then exposed to a solid phase with a specific affinity to the protein of interest or bait. The solid phase is commonly coupled to an antibody against the bait, or alternatively, the bait is genetically fused to a tag and purified using a solid phase with affinity to the tag. The bait along with copurified interaction partners is then eluted, and this eluate is analyzed. A major disadvantage here is that interactions are determined from the cell extract and not the true cellular environment, and thus, some interactions may be lost, whereas spurious and nonphysiological interactions may occur as a result of membrane breakage and decompartmentalization of protein complexes. Screening of extraction conditions including detergent solubilization, pH, ionic strength, and other parameters is therefore necessary. A variation on this approach is cryomilling, which avoids the use of detergents during lysis, but still disrupts cellular organization (3, 4). In general, affinity capture requires a lot of optimization for each individual bait, and results tend to be more variable between experiments and/or labs because of minor differences in cell harvesting, lysis, and buffer conditions.

In proximity labeling, some of the pitfalls of affinity isolation are avoided by capturing interactions *in vivo*. Here, the bait is fused to an enzyme that converts a substrate into a reactive radical that is covalently linked to nearby proteins. Modified

\* For correspondence: Martin Zoltner, [zoltnerm@natur.uni.cz](mailto:zoltnerm@natur.uni.cz); Susanne Kramer, [susanne.kramer@uni-wuerzburg.de](mailto:susanne.kramer@uni-wuerzburg.de).

## Proximity labeling and affinity capture compared

proteins are purified frequently under extremely stringent conditions. The most commonly used method is BioID and variants. The bait is coupled to BirA\*, a mutant form of biotin ligase from *Escherichia coli*. Wildtype BirA converts biotin to biotinol-5'-AMP, which is retained by BirA until it is transferred to acetyl-CoA carboxylase. A mutant version, BirA\*, is modified to release biotinol-5'-AMP, causing biotinylation of lysine residues of proteins nearby (5–7). Additional BioID variants have been developed, one of which is TurboID, which exhibits greater biotinylation efficiency (8). TurboID biotinylates within minutes after addition of exogenous biotin, in contrast to other BirA\* variants (8), but also has some biotinylation activity in the absence of exogenous biotin, leading to an increased labeling radius (9).

Given the distinct underlying principles of these two methods, we have carried out a direct comparison between them, employing four different proteins from trypanosome mRNA metabolism that encompass a variation of location and positional constraints. For affinity capture, we used data previously published by us (10, 11) and repeated the experiment for one target, nucleoporin NUP158, for optimal comparability. Cell lysis was performed by cryomilling, where rapid freezing and mechanical cell disruption at 77 K preserves PPIs and has been successful for isolating nuclear pore complexes (NPCs) and many other complexes from many organisms (3, 12–14). All proteins were expressed as GFP/YFP fusions from the endogenous locus and captured with an anti-GFP single-chain nanobody (10, 11). For BioID, we expressed the same proteins from their endogenous loci fused to the biotin ligase TurboID and used steady-state biotinylation, facilitated by biotin in the culture medium.

We found surprisingly little concordance between PPIs identified from affinity capture and BioID. With their less-confined cytoplasmic localizations, BioID is error prone for poly(A)-binding proteins (PABPs), with many false identifications. In contrast, for the identification of MEX67 interactors, BioID was more adequate and exclusively identified FG-repeat nucleoporins, NPC components that line its inner surface, rather than the entire NPC, which was isolated by the affinity method. Likewise, NUP158 affinity capture identified most part of the NPC cellular structure including distant proteins that are indirectly but stably associated with the bait, which are outside the BioID labeling radius. Whilst affinity capture delivers a snapshot of stable PPIs at the time of lysis, proximity labeling records a history of protein interactions occurring during the duration of labeling, which can bias against detection of a stable “core” complex in favor of dynamic associations. Altogether, our data indicate that the utility of each method is context dependent and hence should be viewed as complementary rather than as alternatives.

## Results and discussion

### Establishing TurboID in trypanosomes

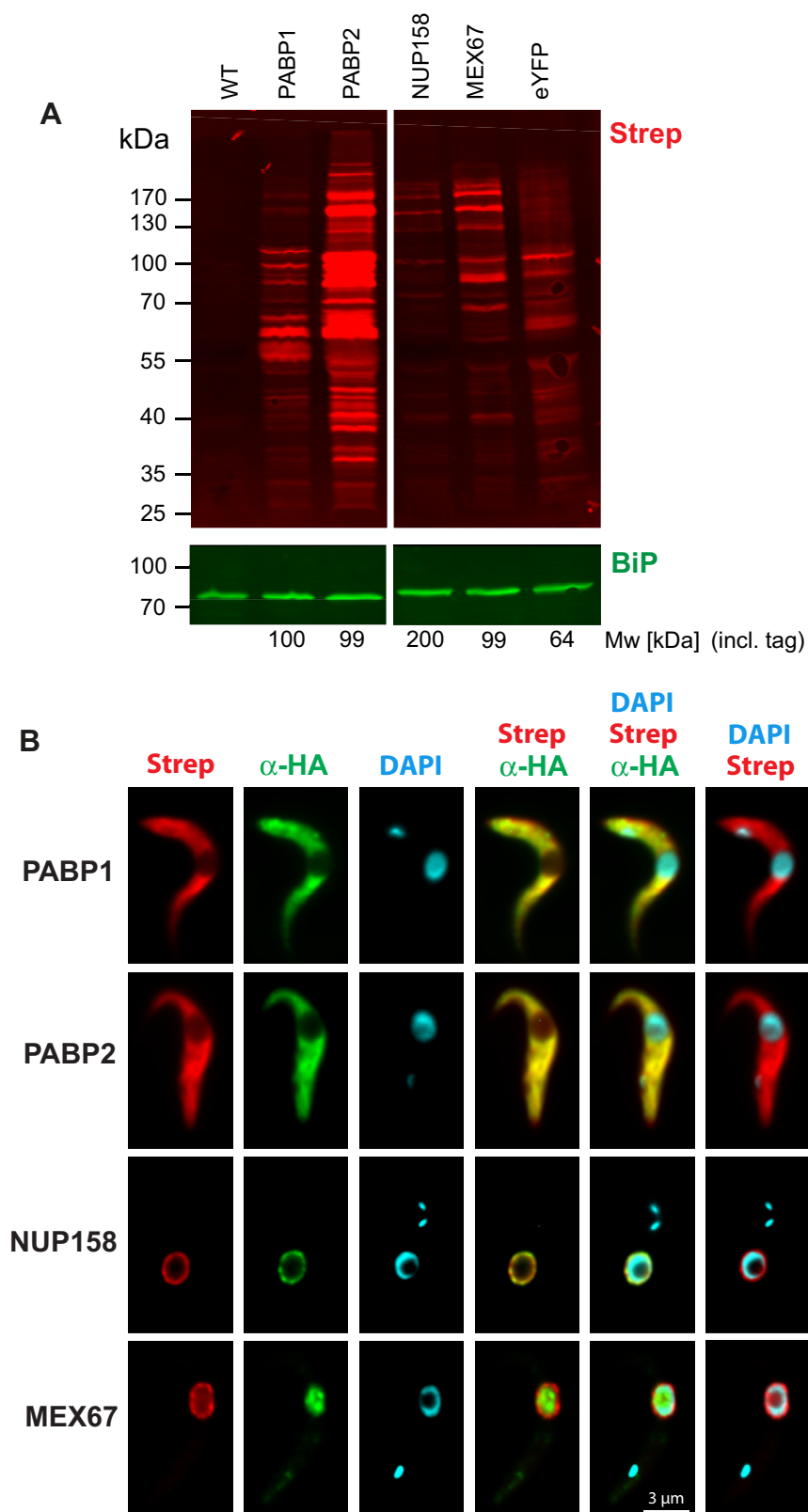
We selected four trypanosome proteins as bait to establish BioID with TurboID biotin ligase fusion. PABP1 and PABP2 are cytoplasmic and by light microscopy appear unconfined

within the cytoplasm under standard culture conditions (10, 15). The nuclear export receptor MEX67 shuttles between the nucleus and the cytoplasm with predominant localization at NPCs (16, 17). The NPC protein NUP158 (ortholog to Nup145 in *Saccharomyces cerevisiae* and Nup98/96 in *Homo sapiens*) is localized to the outer rings of the NPC (3). We selected these bait proteins for the following reasons: (i) we anticipate differential levels of nonspecific background biotinylation dependent on how rigidly a protein is confined, and including a range of protein localizations (not confined, semi-confined, and confined) allows this to be considered and (ii) prior work means we have validated cell lines and in some cases mass spectrometry (MS) data (10, 11).

All four proteins were expressed from their endogenous loci fused C-terminally to TurboID and hemagglutinin (HA), replicating the tagging strategy used for cryomill affinity capture, except that in that case, fusion was to enhanced YFP (eYFP). We used two control cell lines: unmodified parental cells and cells expressing eYFP fused to TurboID and HA (using an inducible expression system). All experiments were done in procyclic form cells, the *Trypanosoma brucei* life-cycle stage in the insect host. Generated lines were subjected to Western blotting to detect biotinylated proteins by a streptavidin probe (Fig. 1A). Almost no biotinylated proteins were detected in parental cells, whereas the tagged cell lines, including the eYFP control, had many biotinylated proteins. There were major differences in the number of biotinylated proteins obtained for each bait and also in the identity of the detected proteins, indicating specificity. Notably, PABP2-TurboID resulted in a significantly higher number of biotinylated proteins than PABP1-TurboID in both life cycle stages, consistent with previous findings that PABP2 interacts with a significantly bigger cohort of proteins compared with PABP1 (10). In addition, we confirmed correct localization of the TurboID bait proteins by anti-HA immunofluorescence and, in parallel, traced biotinylation *via* a fluorescent streptavidin probe (Fig. 1B). For the cytoplasmic PABPs, we observed cytoplasmic localization of the two probes and for NUP158 colocalization at distinct structures of the nuclear envelope. MEX67 TurboID yielded a similar streptavidin labeling pattern as NUP158 with some additional weak signal from the nucleus, but immunofluorescence signals were restricted to nucleoplasm and nucleolus. The absence of immunofluorescence from the nuclear envelope can be explained by the phase-separated environment created by the FG-type nucleoporins, which presumably prevents antibody binding in this region of the pore but not streptavidin binding. Altogether, the observed labeling pattern indicates high spatial labeling selectivity of our TurboID approach. Hence, we purified biotinylated proteins by streptavidin affinity for each cell line in triplicate and analyzed by LC–MS/MS.

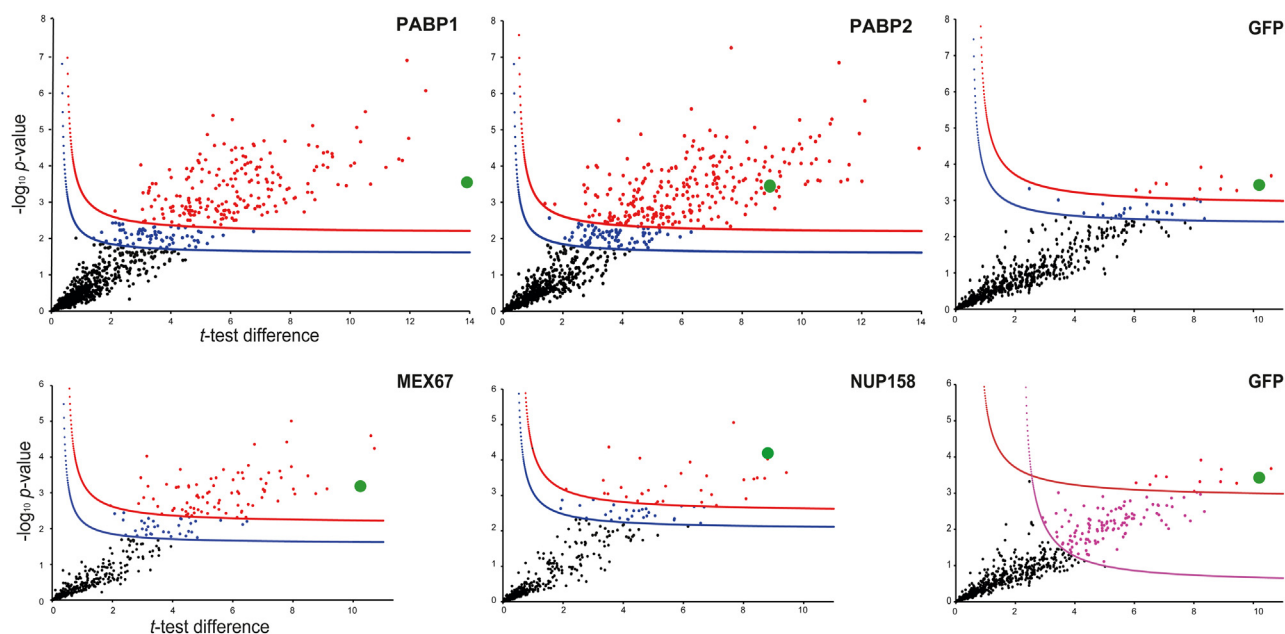
### Defining confidence intervals

Proteins enriched in the BioID samples were grouped into confidence intervals (significant classes; SigA, SigB, and SigC) defined by cutoff curves in volcano plots (Fig. 2) and based on



**Figure 1. Biotinylation in the different BioID cell lines.** *A*, Western blot probed with fluorescently labeled streptavidin to detect biotinylated proteins in the control cell lines (eYFP = TurboID fused to eYFP) or cell lines expressing the respective bait proteins (PABP1, PABP2, MEX67, and NUP158) fused to TurboID. The expected molecular weight of all bait proteins fused to TurboID is indicated. *B*, localization of bait proteins and biotin labeling. Cells were probed with anti-HA antibody and IRDye 800CW streptavidin to detect biotinylated proteins, and then DAPI stained analyzed by fluorescence microscopy. DAPI, 4',6-diamidino-2-phenylindole; eYFP, enhanced YFP; HA, hemagglutinin; PABP, poly(A)-binding protein.

## Proximity labeling and affinity capture compared



**Figure 2. Hawaii plot for the statistical analysis of BioID experiments.** Hawaii plot (multiple volcano plots) of LFQ results of the BioID experiments for PABP1, PABP2, MEX67, NUP158, and the GFP-control. All samples were prepared in triplicate. To generate the volcano plots, the  $-\log_{10} p$  value was plotted versus the  $t$  test difference (difference between means), comparing each respective bait experiment to the *wt* control. Potential interactors were classified according to their position in the plot, applying cutoff curves for “significant class A” (SigA; drawn in red; FDR = 0.01,  $s_0 = 0.1$ ) and “significant class B” (SigB; drawn in blue; FDR = 0.05,  $s_0 = 0.1$ ), respectively, and, for the GFP control “significant class C” (SigC; drawn in pink; FDR = 0.05,  $s_0 = 2$ ). Bait proteins are indicated by a green dot. FDR, false discovery rate; LFQ, label-free quantification; PABP, poly(A)-binding protein.

statistical analysis in Perseus (Max Planck Institute of Biochemistry, Martinsried) (18) (for details, see the [Experimental procedures](#) section). All BioID samples were compared with parental cells to identify proteins naturally biotinylated as well as proteins that nonspecifically bind to the affinity matrix. A second control, cells expressing a GFP-TurboID fusion, served to identify proteins biotinylated in a nonspecific manner.

For the abundant cytoplasmic proteins PABP1 and PABP2, we chose strict parameters to define significance intervals for both the bait proteins and the GFP control (SigA: false discovery rate [FDR] = 0.01;  $s_0 = 0.1$ ; SigB: FDR = 0.05;  $s_0 = 0.1$ ). This defined 35 proteins as GFP positive, 13 in SigA and 24 in SigB (Table S1, A and B, Figs. 2 and 3A), which were removed from the list of PABP1 and PABP2 candidate interactors. The usage of less strict parameters for the GFP-control (SigC: FDR = 0.05;  $s_0 = 2$ ; as used for NUP158 or MEX67 mentioned later) would have resulted in the subtraction of *bona fide* PABP1/2 interactors, as for example, PAB1-binding protein PBP1 and even the bait PABP1 itself. Among the GFP-positive proteins were many involved in translation or associated with cytoskeleton or membranes (19–24), as well as large proteins (26% > 100 kDa; whole genome: 10% > 100 kDa). Only two of this cohort were unique to the GFP BioID experiment, likely reflecting nonspecific interactions, whereas the remaining proteins were also identified in other BioID experiments (usually in more than one; nine in all five experiments).

We also applied SigA and SigB cutoffs to the less abundant proteins with a confined localization at the NPC, NUP158, and MEX67. However, we chose less strict parameters to define the GFP control (sigA: FDR = 0.01;  $s_0 = 0.1$ , SigC: FDR = 0.05;  $s_0 =$

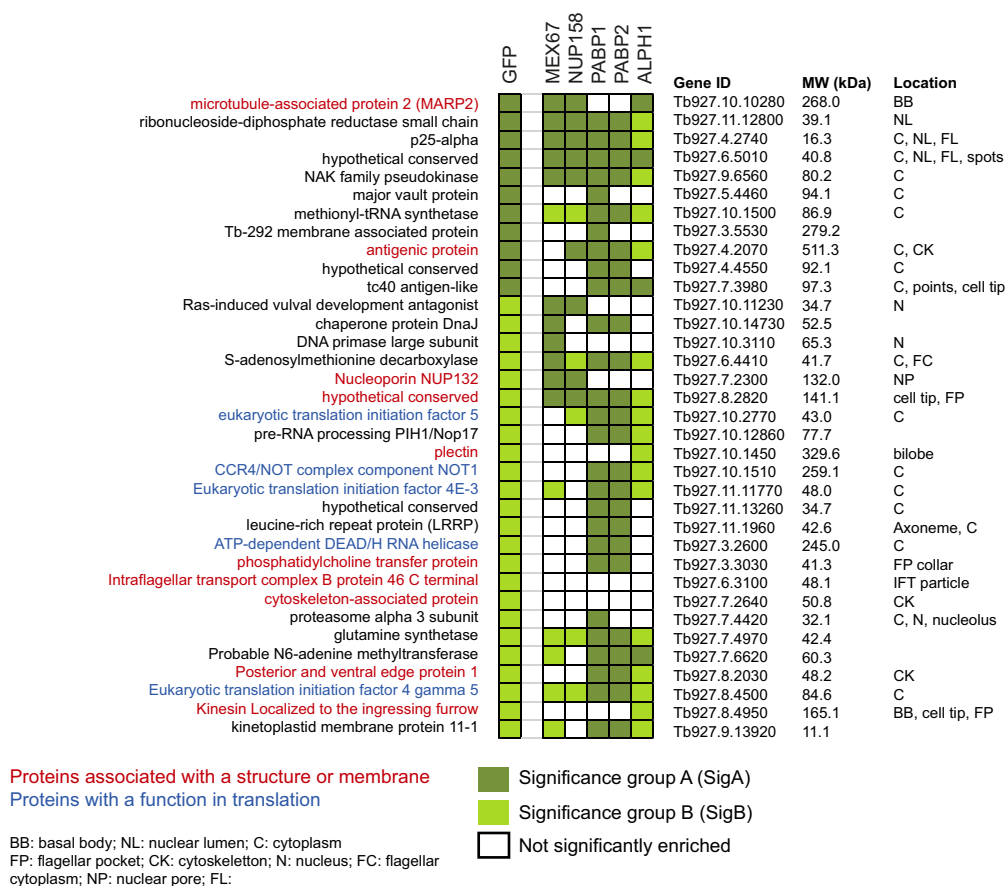
2) (Table S1, A and C, Figs. 2, and 3B). With these parameters, 139 proteins were defined as GFP positive (13 of these in SigA). About 50% of these 139 proteins were uniquely enriched in the GFP control. The others were removed from the list of NUP158- or MEX67-enriched proteins, except for proteins that were in SigA for the bait proteins and in SigC for GFP (listed in Fig. 3B). The latter cohort contained an enrichment in proteins with nuclear localization and even two proteins with localization to the NPC, including the *bona fide* *T. brucei* NPC component NUP132 (11, 25). In contrast, very few nuclear-localized proteins were among confirmed GFP-positive proteins (Fig. 3B).

In conclusion, we determined the appropriate filtering parameters to define the GFP control data for each individual BioID bait to exclude obvious false positives, while including known interaction partners. Notably, being GFP positive does not necessarily exclude the protein from being a true interacting protein; MARP2 (TbBBP268) was identified as GFP positive here but was identified by BioID with basal body protein TbPAC11 and TbBBP46 as baits. MARP2 localization to the basal body indicates that this is likely a true interactor (21). Likewise, some of the translation initiation factors and the RNA helicase within the cohort of GFP positives may be true PABP interactors, highlighting the need for careful consideration of the dataset.

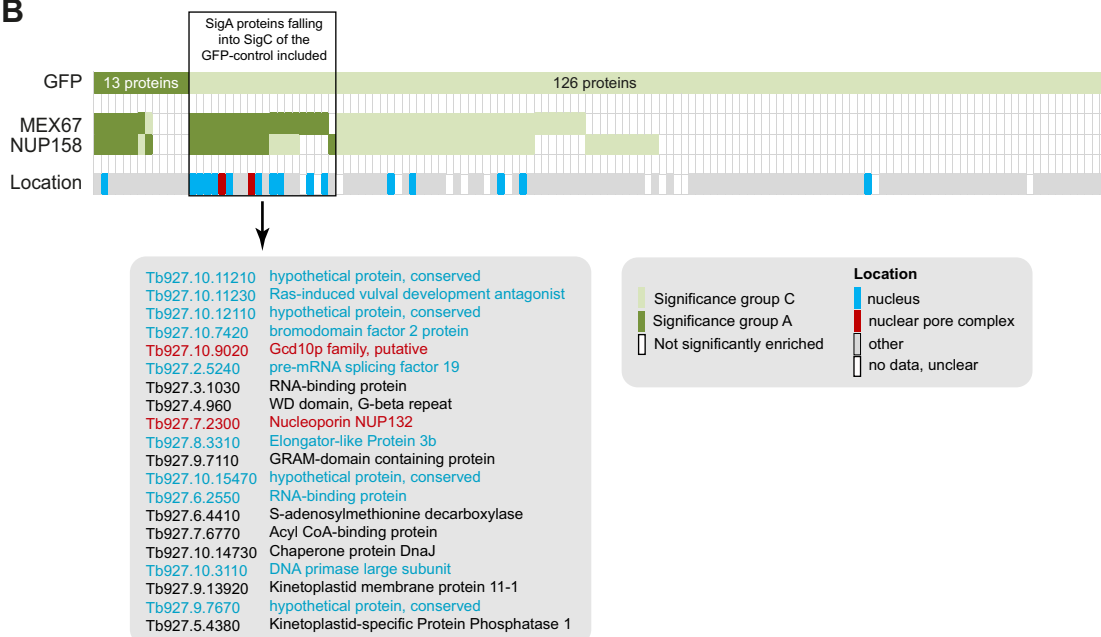
### Cytoplasmic PABPs

*T. brucei* has two cytoplasmic PABPs. Both PABP1 and PABP2 have cytoplasmic localizations under physiological conditions, but PABP2 localizes to stress granules when cells are starved (26, 27) and can also localize to the nucleus under

A



B

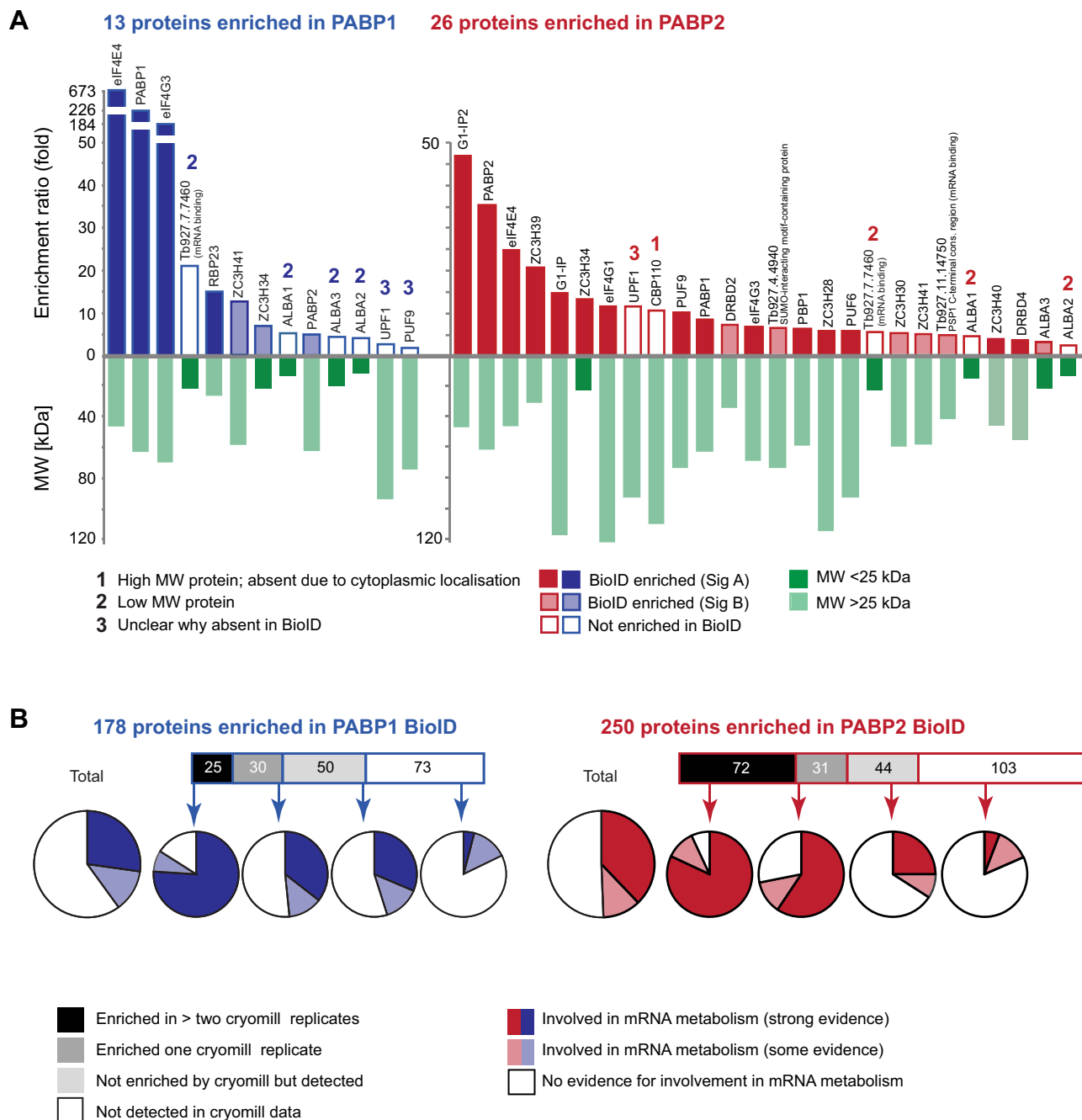


**Figure 3. Proteins biotinylated by the GFP-TurboID control.** A cell line expressing GFP-TurboID was used to detect background biotinylation. We have used different thresholds to define a protein as GFP positive for the different bait proteins (SigA, SigB, and SigC; details in the text). For further comparison, we have added respective TurboID data for the trypanosome mRNA decapping enzyme ALPH1 (63), which localizes to the cytoplasm and posterior pole granules. A, to control the BioID experiment of PABP1 and PABP2, proteins biotinylated by GFP-TurboID were defined using strict parameters (35 proteins, falling into SigA and SigB); only the proteins shown here were considered GFP positive. All GFP-positive proteins were removed from the list of proteins identified in the BioID with PABPs as baits, independent of the significance group. B, to control the BioID experiment of MEX67 and NUP158, we used less strict parameters to define the control (139 proteins, falling into SigA and SigC, were considered GFP positive). However, proteins that were identified with very high confidence with MEX67 and NUP158 BioID were included, if the GFP control was in SigC. Protein localization is mainly taken from TrypTag (20). PABP, poly(A)-binding protein; SigA, significant class A; SigB, significant class B; SigC, significant class C.

## Proximity labeling and affinity capture compared

certain conditions (26, 28). We have previously determined an interactome for PABP1 and PABP2 by cryomill affinity capture using GFP single-chain antibodies, from cells expressing C-terminal fusions of the PABPs to eYFP from endogenous loci (10). Data from four experiments (two with high and two with low salt conditions) were analyzed to yield a high confidence list of PABP PPIs that included only proteins that were at least twofold enriched in all four experiments, had no more than

one zero-detection in the bait sample replicates and were no obvious contaminants. This resulted in a cohort of 13 proteins coimmunoprecipitated with PABP1, of which two proteins (eIF4E4 and eIF4G3) dominated with extremely high enrichment ratios of >150. PABP2 coimmunoprecipitated 26 proteins (Fig. 4A). All proteins coimmunoprecipitated with either PABP1 or PABP2 have either a known function in *T. brucei* mRNA metabolism, are known mRNA binders (Tb927.7.7460



**Figure 4. Poly(A)-binding proteins (PABPs): comparing BioID and pulldown.** A, all proteins significantly enriched in all four replicates of the PABP1 and PABP2 pulldown experiment are shown ranked according to their enrichment ratio (10). Proteins that are also identified in the respective BioID experiment and matching significance group A or B criteria are shown with black or gray filling, respectively. The molecular weight of all proteins is shown below, as we noticed a certain bias for loss of small proteins in BioID. Possible reasons of why a protein is absent from the BioID data are indicated. B, all proteins identified with the PABP BioID experiments in significance group A are listed, split into four groups, dependent on whether these proteins were found enriched in more than one of the pulldown experiments, enriched in only one of the pulldown experiments, detected in the pulldown experiments but not enriched or not detected. For each group of proteins, the fraction of proteins involved in mRNA metabolism is shown as a pie diagram (Table S2, B and C).

and Tb927.11.14750 (29)) or have a localization reflecting RNA granules (Tb927.4.4940 (20)), suggesting a cohort of genuine PPIs. In contrast, the number of proteins identified with BioID was larger and included 178 (PABP1) and 250 (PABP2) proteins in SigA and 263 (PABP1) and 330 (PABP2) proteins in SigB (Table S2).

We first asked which of the affinity-purified proteins were also identified with BioID (Fig. 4A). Of 13 proteins identified in the PABP1 pull-down, four were in SigA of the BioID experiment, including the bait and the two highly enriched proteins eIF4E4 and eIF4G3. A further three proteins were in SigB. Six proteins were not enriched in the BioID experiment. Four of these are low molecular weight proteins (Fig. 4A) and may be missed because of their lower number of detectable unmodified peptides, which we recognized as a systematic problem of BioID (discussed in Conclusion section). The remaining two proteins, PUF9 and UPF1, were the least enriched proteins in the PABP1 pull-down and hence may suggest a less stable or physically more distant and indirect association with PABP1. Of the 26 proteins identified in the PABP2 pull-down, 15 are also in BioID SigA, including the bait and six proteins with highest enrichment ratios in the pull-down. A further six proteins were in the SigB cohort. Of the remaining five, three are low molecular weight and may not be detected as suggested previously. The fourth is a large nuclear protein (CBP110) that possibly remains in the cytoskeleton fraction removed prior to incubation with the streptavidin beads. The fifth protein is, again, UPF1.

We next asked how many proteins identified by BioID were also present in the pull-down datasets (Fig. 4B). Of the 178 proteins identified in the PABP1 BioID (SigA), 25 were present in at least two replicates of the pull-down, 30 were enriched in one replicate, 50 were not enriched but detected, and 73 were not detected. Of the 250 proteins from the PABP2 BioID (SigA), we detected 72 in at least two of the replicates of the pull-down, a further 31 in one replicate, 44 not enriched but detected, and 103 proteins undetected. PABP-interacting proteins are expected to be mainly involved in mRNA metabolism. To evaluate these data further, we manually assessed evidence for involvement in mRNA metabolism for each protein, based on conserved domains, known homologs, and published data. About 40 and 49% of PPIs identified by BioID in confidence group SigA for PABP1 and PABP2, respectively, had evidence for a role in mRNA metabolism. As expected, the number of proteins with a role in mRNA metabolism was highest (>75%) among the BioID proteins that were also identified by affinity capture and lowest (<25%) among the BioID proteins undetected by affinity isolation.

While there is overlap between BioID and pull-down PPI datasets, BioID identifies a significantly larger number of potential interactors, partly explained by the complexity of PABP interactions and the dynamic nature of ribonucleoprotein complexes in which the PABP function. Consistent with this, a recent study (30) applied BioID to elucidate the PPIs of two *T. brucei* RNA-binding proteins, RBP9 and RBP10, uncovering a similar number (>200) of high-confidence interactors.

BioID appears to have a false-positive rate among the high-confidence SigA cohort of about 50%. This is likely because of the less-confined (cytoplasmic) localization of PABP1 and PABP2 and perhaps in addition by their function: PABPs are mobile and interact with many complexes that themselves may have peripheral proteins not obviously involved in mRNA metabolism. Moreover, PABPs are involved in translation and thus in close proximity to many nascent proteins that they potentially biotinylate despite the absence of direct interaction. Note, however, that ribosomal proteins were not found biotinylated, excluding random biotinylation of the entire polysomal complex.

### MEX67

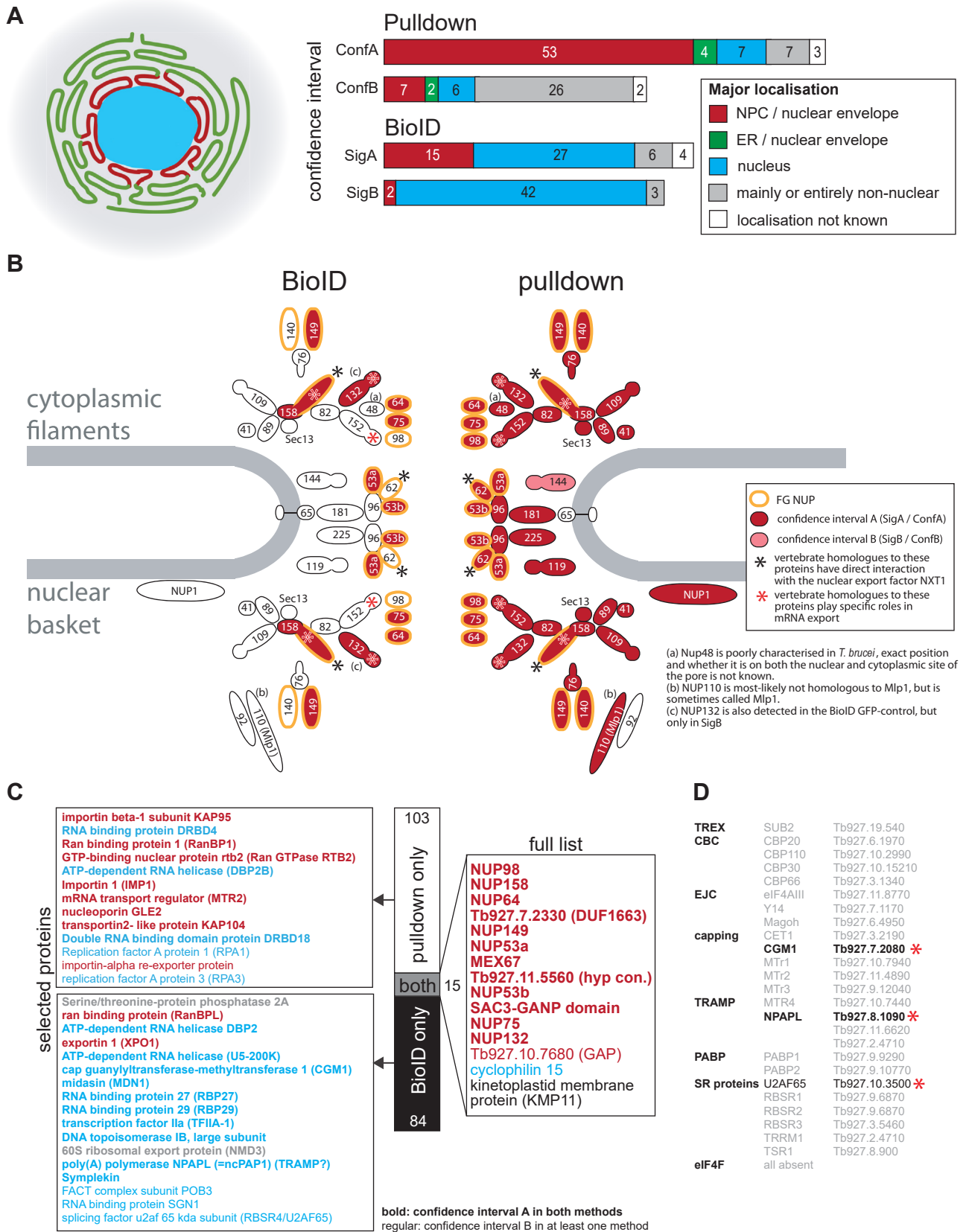
MEX67 forms a heterodimer with Mtr2, and this complex is the major mRNA export complex and is conserved across most eukaryotes (31, 32). MEX67 binds its mRNA targets both directly or indirectly *via* adaptor proteins and mediates mRNA export to the cytoplasm by interacting with FG-type nucleoporins. While mRNA export in trypanosomes differs in several aspects from the pathway described in opisthokonts (33), TbMEX67 (Tb927.11.2370) has been well studied in the past and appears to have a conserved function in mRNA export (17, 33–35). Two datasets from pull-downs are available for *T. brucei* MEX67. A classical immunoprecipitation with a C-terminal PTP (Protein C- tobacco etch virus protease cleavage site - Protein A) tag identified Mtr2 and importin 1, a transport receptor for nuclear import (35), whereas a cryomill affinity capture using a C-terminally GFP-tagged MEX67 as bait identified many nucleoporins (11). We reanalyzed the latter affinity capture data (11) for three different extraction conditions to generate a list of enriched proteins (Table S4B). We have then performed BioID with MEX67 to compare the PPIs identified by these different methods.

In contrast to the PABP1 and PABP2 datasets, the number of MEX67 PPIs identified with BioID (99) was similar to the number of PPIs identified by pull-down (118 proteins) (Fig. 5A). We suggest that this is likely a consequence of the more confined localization of the MEX67 protein, which results in less bystander labeling.

MEX67 BioID identified only nine proteins with a non-nuclear localization, spread over both SigA and SigB cohorts, and the majority of proteins were localized to the nucleoplasm (69 proteins) or the NPC (17 proteins) (Fig. 5A). SigA contained 15 core nucleoporins, whereas SigB showed enrichment of proteins localized to the nucleoplasm. Proteins with nuclear localization only identified by BioID could be proteins that interact with nuclear proteins only transiently, preventing their detection in the pull-down experiment.

Pull-down-identified PPIs differ from the corresponding BioID PPIs in containing a larger number of nucleoporins and a smaller number of nucleoplasmic proteins, and, in particular, for confidence group B, a higher number of non-nuclear proteins. Also specific to the pull-down PPIs were proteins localized to the endoplasmic reticulum (ER) and nuclear envelope: these likely coprecipitated with the NPCs but are not

# Proximity labeling and affinity capture compared





close contact with MEX67 and hence not sampled with BioID (Fig. 5A). However, it should be considered that any transmembrane, vesicular, or intraluminal (perinuclear cisterna) proteins might also be inaccessible to the BioID labeling but not to direct physical connection-mediated affinity capture.

While MEX67 coprecipitated essentially the entire NPC, MEX67 BioID selectively identified six of nine FG-repeat proteins exposed to central channel (Fig. 5B). The only non-FG-repeat NUP identified by BioID is NUP132, which is also enriched with the GFP-bait control (SigB) and may thus have higher exposure to random biotinylation. Thus, for MEX67, BioID captured only proteins directly interacting with MEX67, whereas the pulldown identified the entire NPC, likely driven by high-affinity interactions between individual nucleoporins.

Similar to PABPs, there was little concordance between the PPIs identified by the two different methods, and only 15 proteins of 202 were common (Fig. 5C and Table S3B). Additional to MEX67, there were 12 proteins with localization to the NPC or nuclear envelope, namely (i) eight nucleoporins (compare Fig. 5B), (ii) Tb927.6.890, unique to kinetoplastida, but possessing a SAC3/GANP/THP3 domain at its N terminus; Sac3 from yeast is a well-known interactor of MEX67-Mtr2 that localizes to the cytoplasmic side of the NPC (36), (iii) Tb927.7.2330 and Tb927.11.5560, two proteins of unknown function, and finally (iv) Tb927.10.7680, a GTPase-activating protein (RabGAP-TBC domain, TBC-RootA (37)); TBC-RootA is likely involved in the GTP-dependent mRNA transport *via* the small GTPase Ran (11). Ran, Ran-binding protein, and MEX67b (38), all high-confidence interactors in the pulldown, were also enriched in the BioID PPI dataset but below a significance cutoff. Two further proteins common to both BioID and pulldown are cyclophilin (localization to the nucleolus, unknown function) and KMP11 (a protein of the basal body and flagellum attachment zone) (20, 39), both likely contaminants. Surprisingly, the major MEX67 interactor, Mtr2, was only identified by pulldown, likely because of the low molecular weight of Mtr2 (15.8 kDa).

About 103 proteins were unique to the pulldown, 61 of these in confidence group A. Most (40 proteins) of these localized to the NPC or nuclear envelope and included most nucleoporins (compare Fig. 5B), (putative) transport proteins including Ran-binding protein 1, importin 1, importin beta, Ran RTB2, and many hypothetical proteins. Four proteins localized to the nuclear envelope and ER. Seven had nuclear localizations, including the RBPs DRBD4, DBP2B, and DRBD18, whereas 10 had non-nuclear or unknown localizations.

In contrast, of 84 proteins unique to the BioID dataset, 68 had a nuclear localization, four with NPC and only 12 had non-nuclear or unknown localization. NPC-associated proteins unique to the BioID were exportin-1 (XPO1), NMD3, and Ran-binding protein RanBPL. Interestingly, NMD3 is implicated in nuclear export of mRNA of procyclin-associated genes, a process shown to be disrupted by silencing of NMD3,

XPO1, or MEX67 (40). This finding extended the prototypic function of NMD3 and XPO1 in rRNA export and is suggestive of overlapping and interdependent nuclear export routes for mRNA and rRNA, for which there is also evidence in yeast (41). While the absence of NMD3 and XPO1 in the pulldown argues against a stable interaction with MEX67, their high enrichment in BioID datasets suggests transient interactions with MEX67.

Among the nuclear-localized proteins were many RBPs and proteins with functions in nuclear mRNA processing, including spliceosomal factors (U5-200K, RBSR4/U2AF65, splicing factor 3B, U4/U6 small nuclear ribonucleoprotein PRP3), an exosome subunit RRP6, the noncoding poly(A) polymerase NPAPL, the cap guanylyltransferase-methyltransferase 1 (CGM1), and proteins involved in transcription. However, the majority of proteins known to be involved in nuclear mRNA metabolism (33) was not identified with either method (Fig. 5D).

To summarize, both BioID and pulldown detected the expected interactions of MEX67 with FG-type nucleoporins. In fact, these are likely the major interactions; as in yeast, MEX67 can be fused to such a nucleoporin and still fulfill its essential functions (42). Apart from these FG-type nucleoporins, there was little coincidence between the PPIs identified by the two methods. While the pulldown mainly coimmunoprecipitated the entire NPC, BioID identified nuclear proteins involved in mRNA metabolism and many RBPs, potentially reflecting transient interactions between MEX67 and its mRNA substrates that are missed in the pulldown.

Altogether, BioID is a valuable tool for interrogation of MEX67 function as it clearly discriminates components of the NPC engaging in direct interaction, as the subset NUPs exposed the inner pore channel. On the other hand, the pulldown detects many indirect interactions because of high stability of NPC subunit interactions. The latter provides equally meaningful information, as MEX67 indeed, is considered to exhibit some of the characteristics of a mobile nucleoporin, because of its significant presence at the NPC (42). Furthermore, BioID appears to identify transient interactions, as indicated by the enrichment of proteins with a function in mRNA binding and nuclear localization, albeit requiring validation.

### The nucleoporin NUP158

NUP158 is an FG-NUP/alpha-solenoid and ortholog of outer ring nucleoporins ScNup145 and HsNup98-96. Cryo-mill affinity capture was repeated as described (11) but subjecting the entire elution to triplicate LC-MS/MS analysis. We detected coprecipitated 23 nucleoporins, eight falling into SigA and five into SigB (Figs. 6B, S1 and Table S3). Among these are NUP109, Sec13, NUP41, NUP82, NUP89, NUP132, and NUP152; together with NUP158, these eight proteins

Only three of these proteins were identified in the BioID experiment (*red asterisk*) in significance group A (*black, bold*) or B (*black, regular*). Full details on the MEX67 BioID experiment are listed in Table S3B. A corresponding color-code scheme is shown on the left. CBC, cap binding complex; EJC, exon junction complex; NPC, nuclear pore complex; PABP, poly(A)-binding protein; TRAMP, Trf4-Air2-Mtr4, polyadenylation; TREX, couples transcription and export.

# Proximity labeling and affinity capture compared

A

## BioID

### Nuclear pore complex, envelope

Tb927.3.3180 NUP98  
 Tb927.4.4310 NUP64  
 Tb927.11.980 NUP158  
 Tb927.7.2330 Protein of unknown function(DUF1663)  
 Tb927.11.13080 hypothetical protein, conserved  
 Tb927.10.8170 NUP144  
 Tb927.2.4230 NUP-1  
 \* Tb927.10.9020 Gcd10p family  
 \* Tb927.7.2300 NUP132  
 Tb927.11.11080 NUP149  
 Tb927.10.8650 RanBPL  
 Tb927.10.7680 GTPase activating protein  
 Tb927.11.15560 NUP53a  
 Tb927.11.14340 XPO1  
 Tb927.11.15990 NUP109  
 Tb927.8.5770 ELP3a

### Nucleus

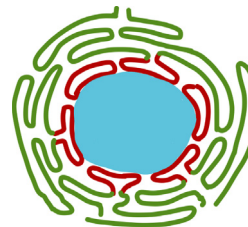
\* Tb927.10.11210 hypothetical protein, conserved  
 \* Tb927.10.11230 Ras-induced vulval development antagonist  
 \* Tb927.10.12110 hypothetical protein, conserved  
 \* Tb927.10.7420 bromodomain factor 2 protein  
 \* Tb927.2.5240 pre-mRNA splicing factor 19  
 \* Tb927.8.3310 Elongator-like Protein 3b  
 Tb927.10.6630 ATP-dependent RNA helicase, DBP2A  
 Tb927.5.2290 ATP-dependent RNA helicase, U5-200K  
 Tb927.10.15330 CYTH domain containing protein  
 Tb927.1.3270 hypothetical protein, conserved  
 Tb927.6.3870 RNA-binding protein, SGN1  
 Tb927.10.14390 FACT complex subunit, POB3  
 Tb927.7.2080 cap guanylyltransferase-methyltransferase 1  
 Tb927.1.3450 hypothetical protein

### Localisation includes ER

Tb927.11.13230 VAMP-associated protein, VAP (mainly FAZ, FP)  
 Tb927.5.4500 ras-like small GTPase, SAR1  
 Tb927.9.9550 EnigmRBP2, ERBP2  
 Tb927.10.13210 dolichyl-P-Man:GDP-ManGlcNAc2-PP-dolichyl  
 beta-1,4-mannosyltransferase  
 (function indicates ER localisation)

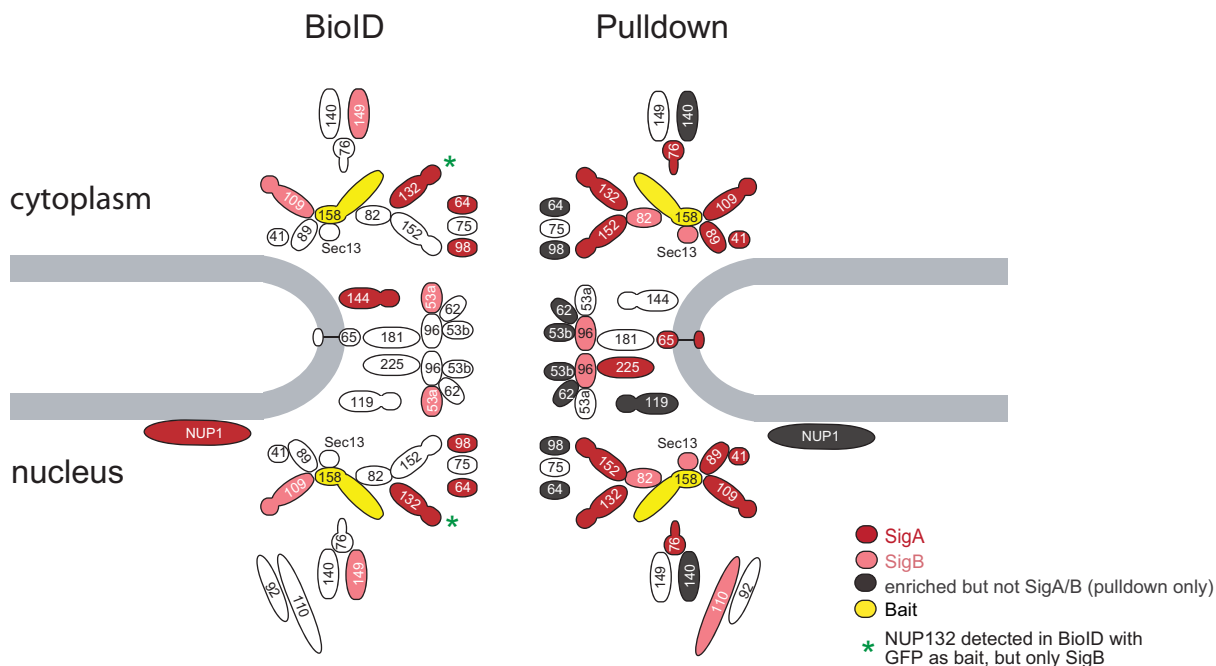
### Other localisations

Tb927.3.1240 Serine/threonine-protein phosphatase 2A  
 \* Tb927.3.1030 RNA-binding protein  
 \* Tb927.4.960 WD domain, G-beta repeat  
 \* Tb927.9.7110 GRAM domain containing protein  
 \* Tb927.5.4380 Kinetoplast-specific Protein Phosphatase 1  
 Tb927.10.3100 glycerol-3-phosphate acyltransferase, putative  
 Tb927.7.4060 cysteine peptidase, Clan CA, family C2  
 Tb927.11.9040 conserved protein  
 Tb927.7.1680 Uncharacterised conserved protein (DUF2305)  
 Tb927.5.800 casein kinase I, isoform 2  
 Tb927.2.2360 Ankyrin repeats/C3HC4 zinc finger (RING finger)  
 Tb927.5.940 NADH-dependent fumarate reductase  
 Tb927.10.6070 universal minicircle sequence binding protein 1  
 Tb927.11.810 hypothetical protein, conserved  
 Tb927.11.11725 ribosomal protein S14p/S29e  
 Tb927.3.4120 hypothetical protein, conserved  
 Tb927.5.2320 Mak10 subunit, NatC N(alpha)-terminal acetyltransferase  
 Tb927.9.6470 hypothetical protein, conserved  
 Tb927.11.3230 60S ribosomal protein L44



\* Also found in GFP BioID  
 significance group B  
**Significance Group A (SigA)**  
 Significance Group B (SigB)

B



**Figure 6. BioID with NUP158.** A, all proteins identified in the NUP158 BioID are listed, grouped according to their intracellular localizations based on TrypTag (20). B, schematic representation of the trypanosome NPC (11). Proteins identified in the NUP158 BioID or pull-down experiments are filled in red or light red for significance intervals A or B, respectively. The bait protein NUP158 is filled in yellow. Proteins enriched in the pull-down falling outside the SigA/B cutoff are filled in gray. NPC, nuclear pore complex. SigA, significant class A; SigB, significant class B.

form the outer ring complex (11). Further significant interactors are the nuclear basket protein NUP110, NUP76, NUP48 and inner ring proteins NUP65, NUP96, and NUP225. Many additional NPC-associated proteins such as NUP140, the inner ring proteins NUP62, NUP53b, and

NUP119, the FG NUPs NUP64 and NUP98, and the lamina protein NUP-1 were also enriched but fall outside the SigA/B cutoff. About 18 additional proteins were identified as significant interactors (Table S3), of which five localize to the NPC and further five to the nucleus (20).

NUP158 BioID identified 53 proteins (Fig. 6A and Table S3): 10 proteins in SigA (including five nucleoporins), a further 12 in SigA that also appear in the GFP-control SigC detections, and 31 proteins in SigB. In comparison to the PABP and MEX67 BioID experiments, this is the shortest list of biotinylated proteins, likely reflecting the highly confined localization of NUP158. Of these 53 proteins, 13 localized to the NPC, 16 to the nucleus, and five to the ER/nuclear envelope (20).

Concordance between BioID and pulldown PPIs was once again poor (Fig. 6B). Many nucleoporins identified as interaction partners in the pulldown were not identified in the BioID (NUP140, NUP76, NUP41, NUP89, SEC13, NUP82, NUP152, and NUP110) and, vice versa, others were unique to the BioID (NUP144, NUP149, and NUP53a) or outside the pulldown SigA/B cutoff (NUP98 and NUP64). The only nucleoporins in common were FG-NUPs NUP109 and NUP13, and there was just one additional interactor Tb927.11.13080 shared between pulldown and BioID datasets (Table S3). This latter is a protein of unknown function but localizes to the NPC (20)). Hence, both methods identify meaningful, but significantly different, PPIs for NUP158. The pulldown identifies adjacent outer ring proteins that are partly absent in BioID. A potential reason for the absence could be that NUP158, located within the rigid structure forming the outer rings of the NPC, has a limited labeling radius because of immobility, steric constraints, or even the local biochemical environment. Therefore, although pulldown is superior identification of PPIs for a strictly localized nucleoporin, additional valuable data can be obtained with BioID.

## Conclusions

Collectively, we found surprisingly little concordance between proteins identified by BioID and by affinity capture. This strongly suggests that these methods should be viewed as providing a distinct PPI for a given bait rather than offer equivalents or alternatives. A similar comparative interactomics study to the one here has been performed for several chromatin-associated protein complexes in human cell lines (43, 44) with broadly similar outcome; BioID generally produced larger interactomes, and the concordance between pulldown and BioID was poor, *albeit* that both techniques identified meaningful PPIs. Our dataset however analyzed a cohort of bait proteins with considerably greater variation of location and positional constraints to provide additional insight.

For soluble and unconstrained cytoplasmic proteins with a high potential for nonspecific interactions, BioID has a significant level of bystander labeling. By contrast, BioID with MEX67, a more restricted bait, identified exclusively FG-NUPs of the inner pore channel, whereas pulldown coprecipitated the entire NPC. Thus, the optimal method depends on the localization of the protein of interest. Ideally, both methods are used in parallel, and recently, a hybrid tag was introduced combining both biotin ligase and an epitope tag for pulldown (45). Several variants of BioID aiming to overcome some of the pitfalls of BioID are available and include an inducible system

with target-specific biotinylation only occurring when the biotin ligase is attached to the bait using a dimerizing agent (46) and a split BioID, where biotinylation can only occur once two proteins carrying partial BioID tags interact (conditional biotinylation) (47).

A major weakness of BioID beyond bystander identifications is that many low-molecular weight proteins are unidentified. Most strikingly, here was the absence of Mtr2 from the MEX67 BioID data. In general, small proteins are harder to detect by MS than large proteins as there are simply less peptides to detect, but this applies to both pulldown and BioID. However, first, BioID relies on surface-exposed lysine residues, and probability of occurrence depends on protein size amongst other parameters. Second, in pulldown analysis, peptides of the entire protein are potentially available for MS, as proteins are fully, or near fully, eluted from the beads, but in BioID, peptides are eluted from beads by trypsin digestion, and any biotinylated peptides remain attached to the beads (Table S4). For low-molecular weight proteins, the likelihood that lysine residues are inaccessible is lower than for larger proteins, as small proteins are less likely to “bury” lysine residues in their core. This can explain the systematic absence of expected interactors with a low molecular weight from BioID but not from the pulldown. One potential way to increase elution of biotinylated peptides would be to add excess biotin before the proteolytic digest to saturate all biotin-binding sites and hence restrict rebinding of biotinylated peptides to the beads after their release by the protease. There is also an alternative approach for purification: trypsination of cell lysates prior to purification followed by purification of biotinylated peptides using a biotin-antibody rather than streptavidin has been shown to result in a higher recovery rate of biotinylated peptides (48, 49). However, this approach comes at the cost of increased background from nonspecific binding, excluding the use of stringent washes as facilitated by the strong biotin–streptavidin interaction.

Altogether, this study offers guidance for choosing the most appropriate method for protein complex characterization, dependent on localization and positional constraints. Furthermore, it highlights potential pitfalls concerning experimental design and data interpretation. Overall, BioID and affinity capture are complementary with each approach elucidating a unique subset of PPIs for a given target protein. Hence, combination of both methods can be leveraged for more complete interactome mapping.

## Experimental procedures

### Cell culture and genetic manipulation

*T. brucei* Lister 427 procyclic cells were cultured in SDM-79 (50). The generation of transgenic trypanosomes was done using standard methods (51). All fusion proteins were expressed from their endogenous locus as described (52), modified to result in fusing TurboID and one HA tag to the C terminus of the protein. eYFP was expressed fused to TurboID-HA using an inducible expression system based on tetracycline (52).

## Proximity labeling and affinity capture compared

### Western blot

Western blotting was performed using standard methods. Detection of biotinylated proteins was done with Streptavidin-IRDye680LT (LI-COR) (1:10,000 dilution). Binding immunoglobulin protein (BiP) was detected using anti-BiP (1:200,000 dilution) (kind gift of James D. Bangs, University at Buffalo).

### Immunofluorescence

About  $1 \times 10^7$  cells at  $5 \times 10^6$  cells/ml were washed once in 1 ml SDM79 without hemin and serum and resuspended in 500  $\mu$ l PBS. For fixation, 500  $\mu$ l of 8% paraformaldehyde was added for 20 min while rotating. About 7 ml PBS with 20 mM glycine were added, cells were pelleted, resuspended in 150  $\mu$ l PBS, and spread on polylysine-coated slides (in circles drawn with a hydrophobic pen). After 15 min, cells had settled to the slide, surplus PBS was removed, and cells were permeabilized with 0.5% Triton X-100 in PBS. Slides were rinsed in PBS, and cells were then blocked in 3% bovine serum albumin (BSA) in PBS for 30 min, followed by 60 min incubation with rabbit-mAb-anti-HA C29F4 (1:500 dilution; Cell Signaling Technology) and with Streptavidin-Cy3 (1:200 dilution; Jackson Laboratories) in PBS/3% BSA. Slides were washed in PBS (three times for 5 min) and incubated with anti-rabbit Alexa Fluor Plus 488 (1:500 dilution) in PBS/3% BSA for 1 h; the last 10 min were done in the presence of 4',6-diamidino-2-phenylindole (DAPI) (0.1  $\mu$ g/ml). Slides were washed 3  $\times$  5 min in PBS and embedded into ProLong Diamond Antifade Mountant (Thermo Fisher Scientific). Images were recorded as Z-stacks (100 images with 100 nm distance) on a custom build TILLPhotonics iMic microscope equipped with a Sencam camera (PCO AG, 6.45 m/pixel) and 100 $\times$  oil immersion (numerical aperture = 1.4) objectives (Olympus). Filter sets were (i) excitation (ex): 320 to 380 nm, dichroic (dc): 400 to 430 nm, and emission (em): 438 to 486 nm (DAPI), (ii) ex: 430 to 474 nm, dc: 585 nm, and em: 489 to 531 nm (Alexa Fluor Plus 488) and (iii) ex: 540 to 580 nm, dc: 585 nm, and em: 592 to 664 nm (Cy3). For each image, the exposure times were 50 ms for DAPI and 500 ms for all other fluorophores. Images were deconvolved using Huygens Essential software (Scientific Volume Imaging BV) and are either presented as a single plane or as a Z-stack projection (sum slices), as indicated.

### Affinity enrichment of biotinylated proteins and on-bead tryptic digests

Cells were maintained at a density of  $1 \times 10^6$  to  $10^7$  cells per ml and harvested at a cell density of  $5 \times 10^6$  cells per ml. The eYFP control was induced by addition of tetracycline at a concentration of 1  $\mu$ g/ml 24 h prior to harvesting. No extra biotin was added for induction of labeling, as we found the biotin concentration in the SDM79 medium (827 nM) to be sufficient for high-level biotinylation.

About  $5 \times 10^8$  cells were used in each replicate. Cells were harvested at 1400g, washed once with serum-free medium, and pellets were rapidly frozen in liquid nitrogen and stored at  $-80$   $^{\circ}$ C. For isolation of biotinylated proteins, each cell pellet was resuspended in 1 ml lysis buffer (0.5%

octylphenoxypolyethoxyethanol, 0.1 M Pipes-NaOH (pH 6.9), 2 mM ethylene glycol-bis( $\beta$ -aminoethyl ether)-*N,N,N',N'*-tetraacetic acid, 1 mM MgSO<sub>4</sub>, 0.1 mM EDTA, complete EDTA-free protease inhibitor cocktail [Roche]) and incubated for 15 min at room temperature (RT) in an orbital mixer. Soluble and nonsoluble fractions were separated by centrifugation (14,000g, 5 min, 4  $^{\circ}$ C) and the soluble fraction was incubated with 100  $\mu$ l streptavidin-linked Dynabeads (MyOne Streptavidin C1; Thermo Fisher Scientific) for 1 h at 4  $^{\circ}$ C under gentle mixing. Beads were washed twice in 1 ml buffer 1 (2% [w/v] SDS in water) once in 1 ml buffer 2 (0.1% [w/v] deoxycholate, 1% Triton X-100, 1 mM EDTA, 50 mM Hepes [pH 7.5], and 500 mM NaCl), once in 1 ml buffer 3 (250 mM LiCl, 0.5% octylphenoxypolyethoxyethanol, 0.5% [w/v] deoxycholate, 1 mM EDTA, 10 mM Tris-HCl [pH 8.1]), and once in 1 ml buffer 4 (50 mM Tris-HCl [pH 7.4], 50 mM NaCl); each washing step was 8 min at RT under orbital shaking. Beads were then prepared for tryptic digestion by washing three times in 500  $\mu$ l ice-cold 50 mM NH<sub>4</sub>HCO<sub>3</sub>, resuspension in 40  $\mu$ l of the same buffer supplemented with 10 mM dithiothreitol, and incubation in a thermomixer at RT for 1 h. Iodoacetamide was added to a concentration of 20 mM, followed by incubation in the dark at RT for 30 min. Finally, 5  $\mu$ g/ml proteomics-grade trypsin (SOLu-Trypsin; Sigma-Aldrich) was added to the beads. The digest was done overnight at 30  $^{\circ}$ C in a thermomixer (1000 rpm). After removal of the first eluate, beads were resuspended in 50  $\mu$ l 50 mM NH<sub>4</sub>HCO<sub>3</sub> supplemented with 10 mM dithiothreitol and 5  $\mu$ g/ml MS-grade trypsin and incubated in a thermomixer at 37  $^{\circ}$ C for 1 h. The eluate was combined with the first eluate, and both were lyophilized in a Speed-vac (Christ alpha 2–4). LoBind tubes (Eppendorf) were used throughout.

### MS and proteomics analysis

BioID-eluted peptides were resuspended in 50 mM NH<sub>4</sub>HCO<sub>3</sub> and passed over C<sub>18</sub> stage tip columns as described (53) and then analyzed by LC-MS/MS on an Ultimate3000 nano rapid separation LC system (Dionex) coupled to an LTQ Q-exactive mass spectrometer (Thermo Fisher Scientific). Resulting spectra were processed using the intensity-based label-free quantification (LFQ) in MaxQuant, version 1.6.16 (Max Planck Institute of Biochemistry, Martinsried) (54, 55). Minimum peptide length was set at six amino acids allowing a maximum of two missed cleavages, and FDRs of 0.01 were calculated at the levels of peptides, proteins, and modification sites based on the number of hits against the reversed sequence database. Data from TurboID were in addition searched for lysine biotinylated peptides (+226.078 Da). LFQ data were analyzed using Perseus (18). For statistical analysis, LFQ values were log<sub>2</sub> transformed, and missing values were imputed from a normal distribution of intensities around the detection limit of the mass spectrometer. These values were subjected to a Student's *t* test comparing an untagged control (*wt* cells) triplicate sample group to the TurboID-tagged protein triplicate sample groups, including the GFP control.  $-\log_{10}$  *t* test *p* values were plotted versus *t* test

difference to generate multiple volcano plots (Hawaii plot, Fig. 2). Potential interactors were classified according to their position in the Hawaii plot, applying cutoff curves for “significant class A” (SigA; FDR = 0.01,  $s_0 = 0.1$ ), “significant class B” (SigB; FDR = 0.05,  $s_0 = 0.1$ ), and “significant class C” (SigC; FDR = 0.05,  $s_0 = 2$ ), respectively. The cutoff is based on the FDR and the artificial factor  $s_0$ , controlling the relative importance of the  $t$  test  $p$  value and difference between means (at  $s_0 = 0$  only the  $p$  value matters, whereas at nonzero  $s_0$ , the difference of means contributes).

### Cryomill affinity capture

For affinity capture of PABP1, PABP2, and MEX67, we used data from our previous studies (10, 11). MEX67 data were reanalyzed for three different extraction conditions (higher stringency buffer 1: 20 mM Hepes, pH 7.4, 250 mM NaCl, 0.5% Triton, plus protease inhibitors, and two low stringency stabilizing buffers: buffer 2A—20 mM Hepes, pH 7.4, 250 mM sodium citrate, 0.5% Tween-20, plus protease inhibitors; buffer 2B—20 mM Hepes, pH 7.4, 20 mM NaCl, 50 mM sodium citrate, 0.5% Tween-20, plus protease inhibitors) and compared with a negative control, omitting the GFP tag. Potential interactors were assigned to two confidence intervals termed confidence A (confidence group A) for proteins enriched greater than twofold under all three conditions and confidence B for proteins enriched greater than twofold in at least one condition. The NUP158 cryomill affinity enrichment purification analysis was repeated essentially as described (11) but subjecting the entire elution to triplicate LC-MS/MS analysis. In brief, NUP158 endogenously tagged with eYFP at the C terminus was extracted in buffer 3 (20 mM Hepes [pH 7.4], 250 mM NaCl, 0.5% CHAPS, complete EDTA-free protease inhibitor cocktail), captured on magnetic anti-GFP nanobody beads (GFP-Trap Magnetic Agarose; Chromotek) and washed four times with buffer 3. Proteins were eluted by on-bead tryptic digest and analyzed by LC-MS/MS on an Ultimate3000 nano rapid separation LC system (Dionex) coupled to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Data were analyzed as described (56) and detailed previously.

### Data availability

All proteomics data have been deposited at the ProteomeXchange Consortium *via* the PRIDE partner repository (57) with the dataset identifier PXD031245. The proteomics data for the PABP1 and PABP2 pulldown have the dataset identifier PXD008839.

**Supporting information**—This article contains supporting information(19–25, 58–62).

**Acknowledgments**—We thank Silke Braune for excellent technical assistance and the TrypTag consortium, which is supported by the Wellcome Trust (grant no.: 108445/Z/15/Z) for freely providing data, which enabled part of this work. We are grateful to the

FingerPrints proteomics facility at the University of Dundee, which is supported by the “Wellcome Trust Technology Platform” award (grant no.: 097945/B/11/Z), to Mark Carrington (University of Cambridge, UK) for plasmids, and to Jay Bangs (University at Buffalo, United States) for provision of the BiP antibody.

**Author contributions**—S. K. and M. Z. conceptualization; C. M. d. N. M., C. D. K., S. O. O., F. Z., N. Z., F. B. H., L. G., and B. D. investigation; S. K., M. Z., M. C. F., and C. M. d. N. M. writing—original draft; S. K. and M. Z. supervision; C. D. K., S. O. O., and F. Z. writing—review & editing.

**Funding and additional information**—The work was financed by a bilateral GACR/DFG grant (project IDs.: 21-19503J and KR4017/9-1; to M. Z. and S. K., respectively), the bilateral DAAD/CAPES grant (project ID: 57390761; to F. B. H. and S. K.), the Czech Ministry of Education (project ID: OPVVV/0000759; to M. Z.), the Wellcome Trust (grant no.: 204697/Z/16/Z; to M. C. F.), and the Medical Research Council (grant no.: MR/P009018/1; to M. C. F.). The MEX67 work was funded by Natural Science Foundation 1818129 (to S. O. O.), National Institutes of Health (NIH) R01 AI140429 (to S. O. O. and Michael P. Rout [to M. P. R.]), NIH R01 GM112108 (to M. P. R.), NIH P41 GM103314 (to Brain T. Chait [B. T. C.]), NIH U54 GM103511, and NIH P41 GM109824 (to M. P. R. and B. T. C.); and Wellcome Trust grant 082813/Z/07/Z (to M. C. F.). We thank M. P. R. and B. T. C. for sharing the MEX67 pulldown data. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: BiP, binding immunoglobulin protein; BSA, bovine serum albumin; DAPI, 4',6-diamidino-2-phenylindole; dc, dichroic; em, emission; ER, endoplasmic reticulum; ex, excitation; eYFP, enhanced YFP; FDR, false discovery rate; HA, hemagglutinin; LFQ, label-free quantification; MS, mass spectrometry; NPC, nuclear pore complex; PABP, poly(A)-binding protein; PPI, protein–protein interaction; RBP, RNA-binding protein; RT, room temperature; SigA, significant class A; SigB, significant class B; SigC, significant class C.

### References

- Dunham, W. H., Mullin, M., and Gingras, A. (2012) Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics* **12**, 1576–1590
- Bosch, J. A., Chen, C., and Perrimon, N. (2021) Proximity-dependent labeling methods for proteomic profiling in living cells: an update. *Wiley Inter. Rev. Dev. Biol.* **10**, e392
- Obado, S. O., Field, M. C., Chait, B. T., and Rout, M. P. (2016) High-efficiency isolation of nuclear envelope protein complexes from trypanosomes. *Met. Mol. Biol. (Clifton, NJ)* **1411**, 67–80
- LaCava, J., Fernandez-Martinez, J., Hakhverdyan, Z., and Rout, M. P. (2016) Optimized affinity capture of yeast protein complexes. *Cold Spring Harb. Protoc.* <https://doi.org/10.1101/pdb.prot087932>
- Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810
- Chapman-Smith, A., and Cronan, J. E. (1999) *In vivo* enzymatic protein biotinylation. *Biomol. Eng.* **16**, 119–125

## Proximity labeling and affinity capture compared

- Choi-Rhee, E., Schulman, H., and Cronan, J. E. (2004) Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci.* **13**, 3043–3050
- Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinikina, T., Carr, S. A., *et al.* (2018) Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887
- May, D. G., Scott, K. L., Campos, A. R., and Roux, K. J. (2020) Comparative application of BioID and TurboID for protein-proximity biotinylation. *Cells* **9**, 1070
- Zoltner, M., Krienitz, N., Field, M. C., and Kramer, S. (2018) Comparative proteomics of the two *T. brucei* PABPs suggests that PABP2 controls bulk mRNA. *PLoS Negl. Trop. Dis.* **12**, e0006679
- Obado, S. O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N. E., Chait, B. T., *et al.* (2016) Interactome mapping reveals the evolutionary history of the nuclear pore complex. *PLoS Biol.* **14**, e1002365
- Boehm, C. M., Obado, S., Gadelha, C., Kaupisch, A., Manna, P. T., Gould, G. W., *et al.* (2017) The trypanosome exocyst: a conserved structure revealing a new role in endocytosis. *PLoS Pathog.* **13**, e1006063
- Luo, Y., Jacobs, E. Y., Greco, T. M., Mohammed, K. D., Tong, T., Keegan, S., *et al.* (2016) HIV–host interactome revealed directly from infected cells. *Nat. Microbiol.* **1**, 16068
- Heider, M. R., Gu, M., Duffy, C. M., Mirza, A. M., Marcotte, L. L., Walls, A. C., *et al.* (2016) Subunit connectivity, assembly determinants and architecture of the yeast exocyst complex. *Nat. Struct. Mol. Biol.* **23**, 59–66
- Kramer, S., Queiroz, R., Ellis, L., Webb, H., Hoheisel, J. D., Clayton, C. E., *et al.* (2008) Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169. *J. Cell Sci.* **121**, 3002–3014
- Stewart, M. (2019) Polyadenylation and nuclear export of mRNAs. *J. Biol. Chem.* **294**, 2977–2987
- Kramer, S., Kimblin, N. C., and Carrington, M. (2010) Genome-wide in silico screen for CCCH-type zinc finger proteins of *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. *BMC Genomics* **11**, 283
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., *et al.* (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Met.* **13**, 731–740
- Hu, H., Zhou, Q., and Li, Z. (2015) SAS-4 protein in *trypanosoma brucei* controls life cycle transitions by modulating the length of the flagellum attachment zone filament. *J. Biol. Chem.* **290**, 30453–30463
- Dean, S., Sunter, J. D., and Wheeler, R. J. (2017) TrypTag.org: a trypanosome genome-wide protein localisation resource. *Trends Parasitol.* **33**, 80–82
- Dang, H. Q., Zhou, Q., Rowlett, V. W., Hu, H., Lee, K. J., Margolin, W., *et al.* (2017) Proximity interactions among basal body components in *trypanosoma brucei* identify novel regulators of basal body biogenesis and inheritance. *mBio*. <https://doi.org/10.1128/mbio.02120-16>
- Fritz, M., Vanselow, J., Sauer, N., Lamer, S., Goos, C., Siegel, T. N., *et al.* (2015) Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. *Nucl. Acids Res.* **43**, 8013–8032
- Goos, C., Dejung, M., Janzen, C. J., Butter, F., and Kramer, S. (2017) The nuclear proteome of *Trypanosoma brucei*. *PLoS One* **12**, e0181884
- Morriswood, B., Havlicek, K., Demmel, L., Yavuz, S., Sealey-Cardona, M., Vidilaseris, K., *et al.* (2013) Novel bilobe components in *trypanosoma brucei* identified using proximity-dependent biotinylation. *Eukaryot. Cell* **12**, 356–367
- DeGrasse, J. A., DuBois, K. N., Devos, D., Siegel, T. N., Sali, A., Field, M. C., *et al.* (2009) Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Mol. Cell Proteomics* **8**, 2119–2130
- Kramer, S., Bannerman-Chukualim, B., Ellis, L., Boulden, E. A., Kelly, S., Field, M. C., *et al.* (2013) Differential localization of the two *T. brucei* poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. *PLoS One* **8**, e54004
- Cassola, A., Gaudenzi, J. G. D., and Frasch, A. C. (2007) Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Mol. Microbiol.* **65**, 655–670
- Lima, T. D. da C., Moura, D. M. N., Reis, C. R. S., Vasconcelos, J. R. C., Ellis, L., Carrington, M., *et al.* (2010) Functional characterization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners. *Eukaryot. Cell* **9**, 1484–1494
- Lueong, S., Merce, C., Fischer, B., Hoheisel, J. D., and Erben, E. D. (2016) Gene expression regulatory networks in *trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol. Microbiol.* **100**, 457–471
- Pablos, L. M. D., Kelly, S., Nascimento, J. de F., Sunter, J., and Carrington, M. (2017) Characterization of RBP9 and RBP10, two developmentally regulated RNA-binding proteins in *Trypanosoma brucei*. *Open Biol.* **7**, 160159
- Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Lührmann, R., *et al.* (1997) Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *EMBO J.* **16**, 3256–3271
- Katahira, J., Strässer, K., Podtelejnikov, A., Mann, M., Jung, J. U., and Hurt, E. (1999) The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* **18**, 2593–2609
- Kramer, S. (2021) Nuclear mRNA maturation and mRNA export control: from trypanosomes to opisthokonts. *Parasitology* **148**, 1196–1218
- Schwede, A., Manful, T., Jha, B. A., Helbig, C., Bercovich, N., Stewart, M., *et al.* (2009) The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. *Nucl. Acids Res.* **37**, 5511–5528
- Dostalova, A., Käser, S., Cristodero, M., and Schimanski, B. (2013) The nuclear mRNA export receptor Mex67-Mtr2 of *Trypanosoma brucei* contains a unique and essential zinc finger motif. *Mol. Microbiol.* **88**, 728–739
- Lei, E. P., Stern, C. A., Fahrenkrog, B., Krebber, H., Moy, T. I., Aebi, U., *et al.* (2003) Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. *Mol. Biol. Cell* **14**, 836–847
- Gabernet-Castello, C., O'Reilly, A. J., Dacks, J. B., and Field, M. C. (2013) Evolution of tre-2/bub2/Cdc16 (TBC) rab GTPase-activating proteins. *Mol. Biol. Cell* **24**, 1574–1583
- [preprint] Obado, S. O., Stein, M., Hegedúsová, E., Wenzhu Zhang, W., Hutchinson, S., Brillantes, M., *et al.* (2022) Mex67 paralogs mediate division of labor in trypanosome RNA processing and export. *bioRxiv*. <https://doi.org/10.1101/2022.06.27.497849>
- Zhou, Q., Hu, H., He, C. Y., and Li, Z. (2015) Assembly and maintenance of the flagellum attachment zone filament in *Trypanosoma brucei*. *J. Cell Sci.* **128**, 2361–2372
- Buhlmann, M., Walrad, P., Rico, E., Ivens, A., Capewell, P., Naguleswaran, A., *et al.* (2015) NMD3 regulates both mRNA and rRNA nuclear export in African trypanosomes via an XPO1-linked pathway. *Nucl. Acids Res.* <https://doi.org/10.1093/nar/gkv330>
- Yao, W., Lutzmann, M., and Hurt, E. (2008) A versatile interaction platform on the Mex67–Mtr2 receptor creates an overlap between mRNA and ribosome export. *EMBO J.* **27**, 6–16
- Derrer, C. P., Mancini, R., Vallotton, P., Huet, S., Weis, K., and Dultz, E. (2019) The RNA export factor Mex67 functions as a mobile nucleoporin. *J. Cell Biol.* **218**, 3967–3976
- Lambert, J.-P., Tucholska, M., Go, C., Knight, J. D. R., and Gingras, A.-C. (2015) Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J. Proteomics* **118**, 81–94
- Hesketh, G. G., Youn, J. Y., Samavarchi-Tehrani, P., Raught, B., and Gingras, A. C. (2017) Parallel exploration of interaction space by BioID and affinity purification coupled to mass spectrometry. *Met. Mol. Biol.* **1550**, 115–136
- Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R. G., Öhman, T., *et al.* (2018) An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.* **9**, 1188
- Chojnowski, A., Sobota, R. M., Ong, P. F., Xie, W., Wong, X., Dreesen, O., *et al.* (2018) 2C-BioID: an advanced two component BioID system for precision mapping of protein interactomes. *Science* **10**, 40–52
- Schopp, I. M., Ramirez, C. C. A., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., *et al.* (2017) Split-BioID a conditional proteomics approach to

- monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* **8**, 1–14
48. Udeshi, N. D., Pedram, K., Svinkina, T., Fereshetian, S., Myers, S. A., Aygun, O., *et al.* (2017) Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Met.* **14**, 1167–1170
  49. Kim, D. I., Cutler, J. A., Na, C. H., Reckel, S., Renuse, S., Madugundu, A. K., *et al.* (2018) BioSITe: a method for direct detection and quantitation of site-specific biotinylation. *J. Proteome Res.* **17**, 759–769
  50. Brun, R., and Schönenberger. (1979) Cultivation and *in vitro* cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. *Acta Tropica* **36**, 289–292
  51. McCulloch, R., Vassella, E., Burton, P., Boshart, M., and Barry, J. D. (2004) Transformation of monomorphic and pleomorphic *Trypanosoma brucei*. *Met. Mol. Biol. (Clifton, NJ)* **262**, 53–86
  52. Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., *et al.* (2007) Functional genomics in *trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol. Biochem. Parasitol.* **154**, 103–109
  53. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906
  54. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
  55. Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13**, 2513–2526
  56. Zoltner, M., Pino, R. C. del, and Field, M. C. (2020) Trypanosomatids, methods and protocols. *Met. Mol. Biol.* **2116**, 645–653
  57. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., *et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucl. Acids Res.* **47**, D442–D450
  58. Lacomble, S., Vaughan, S., Deghelt, M., Moreira-Leite, F. F., and Gull, K. (2012) A *trypanosoma brucei* protein required for maintenance of the flagellum attachment zone and flagellar pocket ER domains. *Ann. Anatomy* **163**, 602–615
  59. DuBois, K. N., Alsford, S., Holden, J. M., *et al.* (2012) NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol* **10**, e1001287
  60. Yavuz, S., and Warren, G. (2017) A role for Sar1 and ARF1 GTPases during Golgi biogenesis in the protozoan parasite *Trypanosoma brucei*. *Mol. Biol. Cell* **28**, 1782–1791
  61. Yorimitsu, T., Sato, K., and Takeuchi, M. (2014) Molecular mechanisms of Sar/Arf GTPases in vesicular trafficking in yeast and plants. *Front. Plant Sci.* **5**, 411
  62. Milman, N., Motyka, S. A., Englund, P. T., Robinson, D., and Shlomai, J. (2007) Mitochondrial origin-binding protein UMSBP mediates DNA replication and segregation in trypanosomes. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19250–19255
  63. Kramer, S. (2017) The ApaH-like phosphatase TbALPH1 is the major mRNA decapping enzyme of trypanosomes. *PLoS Pathog.* **13**, e1006456