Nuclear periphery granules of trypanosomes

A characterization of composition and function

DISSERTATION

FOR A DOCTORAL DEGREE

AT THE JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG

SUBMITTED BY

CARINA GOOS

BORN IN NÜRNBERG

WÜRZBURG, 2021

Submitted on: ……

Office stamp

Members of the Promotionskomitee:

Chairperson:

Primary Supervisor: Dr. Susanne Kramer

Second Supervisor: Prof. Dr. Thomas Dandekar

Date of Public Defense: ……

Date of Receipt of Certificates: ……

Für meine Eltern

Table of Contents

Summary

The nuclear envelope serves as important mRNA surveillance system. In yeast and humans, several control mechanisms act in parallel to prevent nuclear export of unprocessed mRNAs. However, trypanosomes lack homologues to most of the proteins involved. In addition, gene expression in trypanosomes relies almost completely on post-transcriptional regulation as they transcribe mRNAs as long polycistrons, which are subsequently processed into individual mRNA molecules by trans-splicing. As trans-splicing is not error-free, unspliced mRNAs may be recognized and prevented from reaching the cytoplasm by a yet unknown mechanism.

When trans-splicing is inhibited in trypanosomes, the formation of a novel RNA granule type at the cytoplasmic periphery of the nucleus, so called nuclear periphery granules (NPGs) was previously observed. To identify potential regulators of nuclear export control, changes in protein localization which occur when trans-splicing is inhibited, were globally analyzed during this work. For this, trypanosome nuclei were purified under conditions maintaining NPG attachment to the nucleus, in the absence and presence of trans-splicing. Mass spectrometry analyses identified 128 proteins which are specifically enriched in nuclear preparations of cells inhibited for trans-splicing. Amongst them are proteins, which change their localization to the nucleus or to the nuclear pores as well as many proteins that move into NPGs. Some of these proteins are promising candidates for nuclear export control proteins, as the changes in localization (to the nucleus or nuclear pores) were specific to the accumulation of unspliced mRNAs. The NPG proteome almost exclusively contains proteins involved in mRNA metabolism, mostly unique to trypanosomes, notably major translation initiation factors were absent. These data indicate that NPGs are RNP complexes which have started or completed nuclear export, but not yet entered translation. As a byproduct of these proteomic studies, a high-quality dataset of the yet unknown *T. brucei* nuclear proteome is provided, closing an important gap in knowledge to study trypanosome biology, in particular nuclear related processes.

NPGs were characterized in more detail by microscopy. The granules are cytoplasmic and present in at least two different trypanosome life cycle stages. There are at least two distinct granule subsets, with differences in protein composition. A closer analysis of NPGs by electron microscopy revealed that the granules are electron dense structures, which are connected to nuclear pores by string-like structures.

In order to approach the function of NPGs, on the one hand, the hypothesis that NPGs might be related to perinuclear germ granules of adult gonads of *C. elegans* was tested: we found no relation between the two granule types. On the other hand, initial single molecule mRNA FISH experiments

performed in trypanosomes showed no accumulation of unspliced transcripts in NPGs, arguing against an involvement of the granules in mRNA quality control.

The mechanism of nuclear export control of mRNAs is still poorly understood in any organism. During this work several new potential regulators were identified, an important step towards a deeper knowledge of this important pathway. Overall, the extensive dataset of changes in protein localizations at inhibition of trans-splicing paves the way towards a better understanding of mRNA processing and its control mechanisms, not only in trypanosomes, but also in eukaryotes in general.

Zusammenfassung

Die Kernhülle um den Zellkern dient als wichtiges mRNA-Überwachungssystem in Eukaryoten. Bei Hefen und Menschen wirken dabei beispielsweise mehrere Kontrollmechanismen parallel, um den Export von unverarbeiteten mRNAs aus dem Kern heraus zu verhindern. Trypanosomen fehlen jedoch Homologe zu den Meisten hierbei beteiligten Proteinen. Außerdem basiert Genexpression in Trypanosomen fast ausschließlich auf posttranskriptioneller Kontrolle, da die Parasiten mRNAs als lange Polycistrons transkribieren, die anschließend durch Transspleißen zu einzelnen mRNA-Molekülen verarbeitet werden. Da der Prozess des Transspleißen nicht fehlerfrei zu sein scheint, gibt es möglicherweise einen noch unbekannten Mechanismus, der nicht gespleißte mRNAs erkennt und diese daran hindert, das Zytoplasma zu erreichen.

Unter Bedingungen, in denen Transspleißen in Trypanosomen blockiert ist, konnte eine neue Art von RNA-Granula an der zytoplasmatischen Peripherie des Zellkerns beobachtet werden, sogenannte Nuclear Periphery Granules (NPGs). Um potentielle Regulatoren einer Kontrolle des mRNA-Exports zu identifizieren, wurde während dieser Arbeit umfassend analysiert, inwieweit sich die Lokalisation von Proteinen ändert, wenn Transspleißen gehemmt wird. Zu diesem Zweck wurden Zellkerne von Trypanosomen unter Bedingungen aufgereinigt, bei denen die Bindung der NPGs an den Zellkern in Abwesenheit und Gegenwart von Transspleißen erhalten blieb. Mit Hilfe von massen-spektrometrischen Analysen konnten 128 Proteine identifiziert werden, die spezifisch in den Kernpräparaten von Zellen angereichert sind, in denen Transspleißen blockiert ist. Darunter befinden sich Proteine, die ihre Lokalisation in den Kern hinein oder zu den Kernporen hin verändern, sowie viele Proteine, die sich in NPGs bewegen. Einige dieser Proteine stellen vielversprechende Kandidaten für eine potenzielle Rolle in der Kernexportkontrolle dar, da die Veränderungen in der Lokalisation (zum Kern oder zu den Kernporen) spezifisch für die Akkumulation von nicht gespleißten mRNAs waren. Das hier erarbeitete NPG-Proteom enthält fast ausschließlich Proteine, die am mRNA-Metabolismus beteiligt sind und nur in Trypanosomen vorkommen. Insbesondere fehlen im NPG-Proteom wichtige Translationsinitiationsfaktoren. Die Daten zeigen, dass NPGs RNP-Komplexe sind, die den Export aus dem Zellkern bereits begonnen oder abgeschlossen haben, aber noch nicht mit dem Translationsprozess begonnen haben. Als Nebenprodukt dieser proteomischen Analyse, kann ein qualitativ hochwertiger Datensatz des noch unbekannten Kernproteoms von *T. brucei* bereitgestellt werden. Damit wird eine wichtige Wissenslücke bei der Forschung zur Trypanosomenbiologie, insbesondere zu nuklearen Prozessen geschlossen.

Im Rahmen dieser Arbeit wurden außerdem NPGs anhand mikroskopischer Untersuchungen detaillierter charakterisiert. Die Granula sind zytoplasmatisch und liegen in mindestens zwei

verschiedenen Lebenszyklusstadien von Trypanosomen vor. Es gibt mindestens zwei Untergruppen der Granula mit Unterschieden in der Proteinzusammensetzung. Eine genauere elektronenmikroskopische Analyse ergab, dass es sich bei NPGs um elektronendichte Strukturen handelt, die durch fadenartige Strukturen mit den Kernporen verbunden sind.

Um mehr über die potenzielle Funktion von NPGs herauszufinden, wurde zum einen die Hypothese untersucht, dass NPGs mit perinuklearen Keimgranula adulter Gonaden in *C. elegans* verwandt sind: Es konnte keine Beziehung zwischen den beiden Granulatypen gefunden werden. Zum anderen zeigten erste Experimente mittels Einzelmolekül-mRNA-FISH in Trypanosomen keine Akkumulation von ungespleißten Transkripten in NPGs, was gegen eine Beteiligung an mRNA-Qualitätskontrollmechanismen spricht.

Es wurde bisher in keinem Organismus aufgeklärt, welche exakten Mechanismen der nuklearen Exportkontrolle von mRNAs zu Grunde liegen. Während dieser Arbeit wurden mehrere neue potenzielle Regulatoren dieses Prozesses identifiziert - ein wichtiger Schritt in Richtung Aufklärung dieses wichtigen Transportweges. Der hier zur Verfügung gestellte Datensatz enthält umfangreiche Informationen zu Änderungen der Proteinlokalisationen bei Hemmung von Transspleißen und ebnet dadurch den Weg für ein besseres Verständnis von mRNA-Verarbeitung und den notwendigen Kontrollmechanismen, nicht nur in Trypanosomen, sondern auch in Eukaryoten im Allgemeinen.

1 Introduction

1.1 *Trypanosoma brucei* – neglected pathogen, health and economic risk factor and model organism

Trypanosoma brucei are single cell, protozoan parasites of the order *Kinetoplastida,* which share the common feature of a kinetoplast, that consists of the mitochondrial DNA of the cell. The parasitic flagellate has a digenic lifecycle and shuttles between an insect vector and its mammalian host. *T. brucei* is transmitted by the bite of an infected tsetse fly (*Glossina spp.)* taking its blood meal. The parasite has a major impact on human health in two ways. Not only does it cause human African trypanosomiasis (HAT), commonly known as Sleeping Sickness[1], it also causes the cattle disease Nagana in animals, resulting in major economic losses and food shortages. This neglected tropical disease mainly affects areas with bad infrastructure, high poverty and limited access to medical care or drugs, e.g. areas of sub-Saharan Africa. If not treated, African Sleeping Sickness is always fatal.

Trypanosomes undergo a complex life cycle as they alternate between their mammalian host and their insect vector. This shuttling behavior requires several complex changes and differentiation steps, since trypanosomes are exposed to drastic changes in environment including changes in temperature, pH or availability of nutrients. To cope with these alterations, the parasite frequently changes gene expression, morphology and metabolism. Of the several different developmental stages of trypanosomes, the procyclic form (PCF), which colonizes the midgut of the tsetse fly, and the bloodstream form (BSF), which naturally occurs in the mammalian host, are the ones mostly studied in the lab because these stages can be cultivated in vitro.

Trypanosomes branched off the eukaryotic lineage very early during evolution and in some cases evolved unique and unusual biological mechanisms. For example, they possess the ability to survive extracellularly in the bloodstream of vertebrates, with constant exposure to the immune system. The cells are evading the immune system by frequently changing their surface protein variants (antigenic variation)[2]. Also, the organism completely relies on polycistronic transcription, which functions without classical promoters but instead is initiated from regions defined by histone variants[3]. All polycistronic transcripts are subsequently processed by trans-splicing. But why are trypanosomes a suitable model for studying mRNA metabolism? 1) Genetic manipulations like RNAi and gene knock-out, as well as forwards genetics via RNAi library screens, are easy to perform and well established. In *T. brucei*[4] RNAi was discovered almost at the same time as in *Caenorhabditis elegans*[5] and it opened new possibilities for reverse genetics. Other molecular tools such as recombination-based gene deletion and *in situ* tagging[6] as well as inducible ectopic gene expression[7] are available[8]. More recently,

the CRISPR/Cas 9 system was established[9,10]. 2) The parasite is almost completely dependent on post-transcriptional regulation of gene expression, the proposed main functions of ribonucleoprotein particles (RNP) granules (see chapter 1.2). Many translational control features of higher eukaryotes, such as µRNAs and IRES are likely absent[11]. 3) The genome of *T. brucei*[12] and many other kinetoplastid species[13–15] has been fully sequenced. Also, RNA sequencing data for total mRNA abundance and mRNA changes during life cycle, as well as mapping data of untranslated regions (UTRs) are available. 4) Two of the proliferating *T. brucei* life-cycle stages, the procyclic form (PCF) and the bloodstream form (BSF) can be easily cultivated *in vitro* and therefore provide an experimentally accessible system to study gene expression, post transcriptional regulation or mRNA maturation processes. 5) Orthologs of many genes are conserved from trypanosomes to humans. Thus, members of this group are a good model for studying not only the origin of parasitism[16] but also the conservation and divergence of genes and various biological processes within eukaryotes.

1.2 Introducing RNP granules

No eukaryotic mRNA molecule exists isolated: it is accompanied by a whole set of proteins throughout its entire life. Under certain circumstances mRNAs and their accompanying proteins aggregate to very large, microscopically visible structures, so called RNP granules[17]. RNP granules play an essential role in the posttranscriptional regulation of messenger RNA (mRNA) fate. There are several different types of RNP granules, of which the following are best studied:

P-bodies (processing bodies) are highly dynamic structures, which are present in the cytoplasm of all eukaryotic cells. RNA of P-bodies is in equilibrium with translating ribosomes and their size correlates with the levels of cytoplasmic, non-polysomal mRNAs. They are smallest when most mRNAs are in translation or entrapped in polysomes but since mRNAs can move between these two compartments[18,19], they increase with higher levels of non-polysomal RNAs[20–23]. P-bodies dissociate when ribosomes are stalled by cycloheximide[20,24] whereas puromycin causes polysome disassembly and increases P-bodies[21,22]. P-bodies contain enzymes involved in 5'- 3' mRNA decay, for example the decapping complex Dcp1/2 and the 5'-3' exoribonuclease XNR1[20], but their exact function remains unclear, in particular, if they actually function in mRNA degradation[25,26].

Stress granules form in response to a variety of cellular stresses and are believed to temporarily store and protect mRNAs during a respective stress condition, thus modulating the cell's stress response[27]. The granules contain stalled translation initiation complexes; some subtypes also contain small ribosomal subunits. Their formation is as dynamic as P-body formation and both granule types

overlap in their protein and RNA content and can interact with each other. While in yeast, P-bodies and stress granules partially overlap[28–30], they seem to exist next to each other in mammalian cells[21,31].

Other granule types are restricted to specific developmental stages. **Germ granules** for example occur in the germ line of animals and are known to control gene expression during early development, when transcription is inactive, and cells rely entirely on maternal mRNAs[32]. Germ granules regulate translational repression by storing and protecting maternal mRNAs until these are recruited to ribosomes for translation. In adult gonads, germ granules localize to the cytoplasmic site of nuclear pores (perinuclear germ granules) and, in *C. elegans*, it was shown that these perinuclear granules not only interact with nuclear pores, but also contain newly transcribed mRNAs[33].

However, the exact functions of most RNP granules remain elusive, as do their compositions. Only recently, it was possible to, for the first time, to successfully determine the proteome and RNA content of trypanosome starvation stress granules[34]. Later, stress granules were purified from yeast and human[35].

1.3 RNP granules in trypanosomes

Some types of RNP granules are conserved throughout the eukaryotic kingdom, from man to trypanosomes, for example P-bodies and stress granules. However, there are also granules that appear to be specific to a certain group of organisms, perhaps as an adaptation to their individual features and needs[11,36].

One such unique feature in the group of Kinetoplastids is the full dependency on mRNA trans-splicing. Kinetoplastids, including Trypanosomes co-transcribe their entire genomes as long polycistronic pre-mRNAs of tens to hundreds of functionally unrelated genes[12,37,38]. Promoters are largely absent and transcription start and stop sites are defined by histone variants[3]. Trans-splicing transforms these polycistronic transcripts into individual mRNA molecules: a capped 39 nt mini-exon of the spliced leader RNA (SL RNA) is added to the 5’ end of each individual mRNA, coupled to polyadenylation of the upstream gene[39]. With two exceptions, trypanosome genes have no introns; cis-splicing is present, but affects only these two genes[12,13,40,41]. As a consequence, gene expression in Trypanosomes almost entirely relies on posttranscriptional regulation[42,43]. With this almost complete lack of transcriptional control it seems to be a logical consequence to possess a large repertoire of RNP granules[44]. At least six different RNP granule types are present in trypanosomes:

P-body- like granules[45], nutrient starvation stress granules[46], heat shock stress granules[47], the posterior pole granule [47], tRNA half granules[48] and **nuclear periphery granules (NPGs)**.

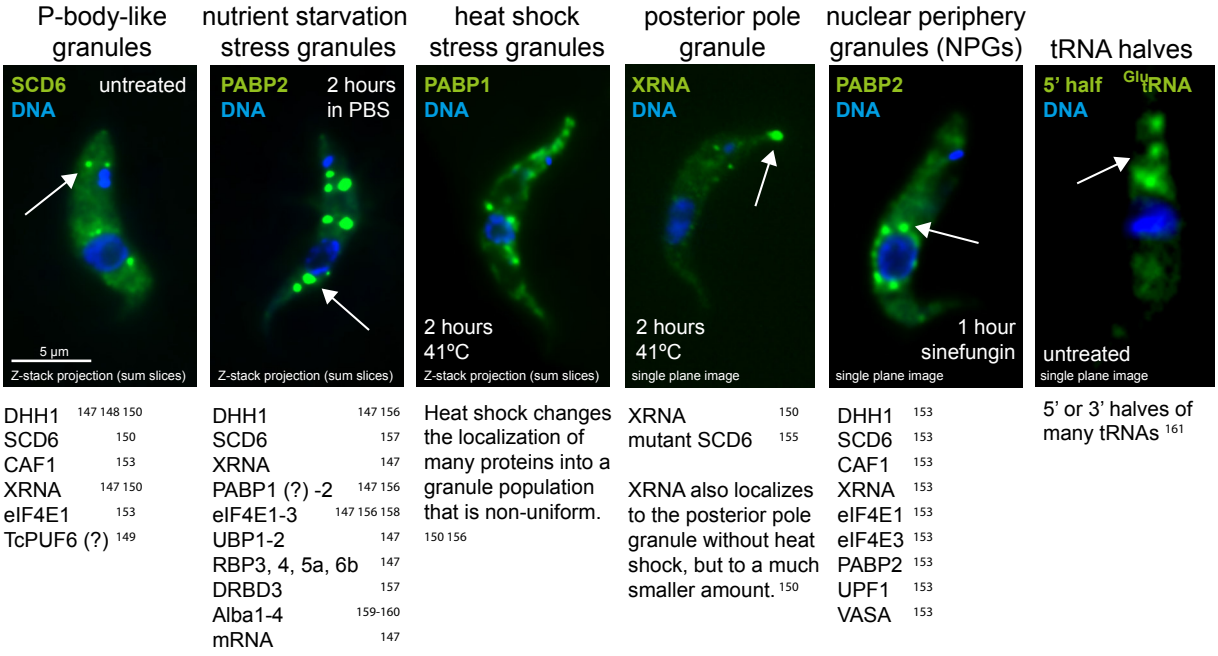


Figure 1.1: RNP granules of trypanosomes[11]

Fluorescence microscopy images of the different types of trypanosome RNP granules from *T. cruzi* (tRNA half-granules) and *T. brucei* (all others) are shown. Proteins/RNA that were unequivocally identified in the granules in at least one of the Tritryps are indicated below the images. Arrows point to the specific granules. All references for the localization to granules are indicated for each and can be found in Kramer 2014. The figure and caption are taken and modified from Kramer 2014[11].

NPGs resemble perinuclear germ granules of adult gonads[49]. They form, when trans-splicing is inhibited by sinefungin, an S-adenosylmethionine analogue, which inhibits the methylation of the SL RNA[50]. Unlike P-bodies, NPGs are dependent on active transcription, but do not dissociate with cycloheximide[27,49]. 20 proteins have been previously tested for NPG localization by the expression of eYFP fusion proteins. All proteins involved in the early, nuclear mRNA processing steps (splicing, capping and nuclear export) were absent from NPGs. Nine cytoplasmic RNA-binding proteins localized to NPGs upon sinefungin treatment, namely the translation initiation factors eIF4E1 and eIF4E3, the deadenylase CAF1, the 5´-3´ exonuclease XRNA, the NMD helicase UPF1, the RNA helicase VASA/DBP1, Poly(A) binding protein 2 (PABP2), the DEAD box RNA helicase DHH1 and the Lsm domain protein SCD6[34,49].

Even if the function of NPGs remains unknown, the fact that these granules depend on active transcription but not on translation suggests that they may operate as a cytoplasmic mRNA quality control compartment.

1.4 Controlling mRNA quality - maturation process and nuclear export

Eukaryotic messenger RNAs (mRNAs) require several nuclear maturation steps prior to nuclear export. A complex, coordinated control system ensures that only fully processed mature mRNAs reach the cytoplasm[51,52]. This prevents incorrectly or incompletely processed mRNAs from disturbing or interfering with cellular functions[53]. mRNA transport is mediated by transport factors that guide the mRNAs through nuclear pores, which are highly specialized transport channels consisting of 500 to 1000 copies of different proteins - the nucleoporins (NUPs) - embedded in the nuclear envelope. The nuclear envelope thus acts as main control barrier and ensures that only mature mRNAs reach the cytoplasm.

Within the nucleus, during transcription and processing, each mRNA molecule associates with a range of proteins, which either target the molecule to degradation or facilitate it's export[54]. The so-called TRAMP complex marks faulty mRNAs for degradation by the nuclear RNA exosome[55]. In contrast, mRNAs are marked as export competent by a set of shuttling adaptor proteins, often serine arginine (SR) rich proteins. These in turn recruit the nuclear export receptor complex Mex67/Mtr2, which mediates transport across the nuclear pore via interactions with distinct nucleoporins. The messenger ribonucleoproteins (mRNPs) to be exported, are further controlled at the entrance site of the nuclear pore - the nuclear basket. The multiprotein complex TREX (Transcription/Export) is essential for the complete mRNA maturation process, since it couples mRNA transcription, recruitment of RNA binding proteins and nuclear export[56].

Even though several factors that play a role in mRNA quality control are known, the underlying mechanism still remains largely unknown. In particular, it remains unclear whether an mRNP is actively selected for nuclear export if it is intact, or actively retained if it is faulty, or, whether both models are involved[57]. Furthermore, it remains largely unclear how faulty mRNAs are distinguished from intact ones, although post-translational modifications of SR proteins may play a role[58].

Trypanosomes mostly regulate gene expression through post-transcriptional mechanisms, particularly mRNA stability. They have a TRAMP-like complex[59], a nuclear RNA exosome[60], which degrades unspliced RNA[61] and a functional homologue to the Mex67/Mtr2 receptor complex[62,63]. Surprisingly, in trypanosomes the export factors NMD3 and XPO1, which are usually involved in export of ribosomal RNA, regulate both rRNA and mRNA export, in addition or in parallel to the Mex67/Mtr2 pathway[64]. Whether trypanosomes have a TREX complex remains controversial[65]. Moreover, the nuclear pore structure of trypanosomes is highly unusual in that the typical asymmetry of nuclear pore components found in opisthokonts is mostly missing[66,67].

Eukaryotic nuclear export control in general as well as underlying processes in trypanosomes are poorly understood. It was previously observed through fractionation methods, that any control system in trypanosomes is not completely tight and unspliced (e.g. polycistronic) mRNAs can reach the cytoplasm[49,68]. Since trypanosomes have almost no introns, this may be the reason why they can afford a certain leakage of unspliced mRNAs. It might be tolerable to the parasite, with the worst damage being a misregulation in gene expression, but no production of faulty proteins. However, at least some control system must be present, since the concentration of unspliced *tubulin* mRNA was found to be 8-fold higher in the nucleus than in the cytoplasm[49] and polycistronic transcripts accumulate upon RNA interference (RNAi) depletion of the nuclear RNA exosome[61].

1.5 Goals of the thesis

Trypanosomes are used as a model organism during this work, because they provide a unique system, in which trans-splicing can be globally inhibited by drug treatment, enabling the analysis of changes on protein and mRNA localizations. By massively flooding the system with very long, polycistronic mRNA transcripts, the mRNA export pathway of trypanosomes is interrupted. As a result, RNP complexes at the cytoplasmic site of the nucleus, in between nuclear export and translation, form - nuclear periphery granules (NPGs).

The aim of this work was to extend and deepen the current knowledge on NPGs and to get insight into the pathway of nuclear export control of unspliced transcripts.

This work includes the following key experimental approaches:

1. Trypanosome nuclei of untreated and sinefungin-treated cells under conditions that maintain the attachment of NPGs to the nucleus were purified and analyzed by mass spectrometry. The localization of all proteins enriched in the nuclei of sinefungin-treated cells was determined. Several proteins were further characterized.
2. Various fluorescence studies and ultrastructural analysis by electron microscopy were performed to further characterize NPGs.
3. Single molecule mRNA FISH, specific to unspliced mRNAs, was planned to give insight into a possible function of NPGs.

2 Materials

2.1 Organisms and cell lines

Bacterial strains

E. coli Top 10

Genotype: mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rspL(SmR), endA1, nupG

Trypanosomes

Wild type cell lines

Trypanosoma brucei brucei Lister 427 (PCF)

Transgenic cell lines

Trypanosoma brucei brucei MITat 118 (BSF), 2845 expression (see Sup. Table 7.2) (provided by Susanne Kramer)

Trypanosoma brucei brucei Lister 427 pSPR2.1 (Sunter et al., 2012)[69]

All further transgenic trypanosomes, which were used or newly generated during this work, were produced using standard procedures[70]. The respective cell lines are described in detail, at the point of mentioning.

C. elegans

WEH172

Genotype: unc119(ed3)III, wurIs66[cgh-1::gfp, unc-119(+)] (provided by Ann Wehman, Rudolf-Virchow-Center, Würzburg, Germany)

JJ2102

Genotype: zuls242 [nmy2::PGL-1-GFP line O05-1] (provided by Sebastian Markert, Biocenter, Würzburg, Germany)

2.2 Antibodies

Table 2.1: Overview of used antibodies.

Antibody	Organism	Origin	Application/Dilution
Primary Antibodies			
anti-DHH1	Rabbit	Mark Carrington (Department of Biochemistry, Cambridge, UK)	IF 1:2000 WB 1:10000 EM 1:100
anti-SCD6	Rabbit	Susanne Kramer (Biocenter, Würzburg, Germany)	IF 1:500 WB 1:20000 EM 1:500
anti-BB2	Mouse	Keith Gull (Sir William Dunn School of Pathology, Oxford, UK)	IF 1:300 WB 1:1000 EM 1:200
anti-PFR	Mouse	Keith Gull (Sir William Dunn School of Pathology, Oxford, UK)	WB 1:50
anti-BIP	Rabbit	Jay Banks (University at Buffalo, NY, USA)	WB 1:2000
anti-PABP2	Rabbit	Osvaldo de Melo Neto (Instituto Osvaldo Cruz, Recife, Brazil)	WB 1:1000
anti-H3	Rabbit	Christian Janzen (Biocenter, Würzburg, Germany)	WB 1:50000
anti-MAB414	Mouse	Abcam, Cambridge, UK	IF 1:1000/ 1:5000
Secondary Antibodies			
anti-mouse IgG Alexa594	Goat	Thermo Fisher Scientific, Cat. No. A11005	IF 1:500
anti-rabbit IgG Alexa594	Goat	Thermo Fisher Scientific, Cat. No. A11012	IF 1:500
anti-rabbit IgG CF350	Goat	Sigma-Aldrich, Cat. No. SAB4600015	IF 1:500
anti-rabbit IgG IRDye800CW	Goat	LI-COR Bioscience, Cat. No. 926-32211	WB 1:10000
anti-mouse IgG IRDye680LT	Goat	LI-COR Bioscience, Cat. No. 926-68020	WB 1:20000
anti-rabbit IgG 12nm colloidal gold	Goat	Dianova, Cat. No. 111-205-144	EM 1:10

2.3 Probes and dyes

Table 2.2: Overview of probes.

Probe	Description	Fluorophore	Origin	Sequence
AF4	beta to alpha tubulin IR	type 4	Affymetrix	ATCCTCCACCTTTATGGGTCCCATTGTTTGCCT CTTCCGCTGTGTGGAGTGCGCCTACACGCACT TCTCACTTCGTAAGTGGTGGTGGCGTAAGTA TTCCCTAATGTTGACTCTATATTCTCCTCTCCT CACCCCCTCGCGGTGCTGATTTCTGACAG
SKA48	ME antisense	Cy3	Sigma-Aldrich	CAATATAGTACAGAAACTGTTCTAATAATAGC GTT

Table 2.3: Overview of fluorescent dyes.

Dye	Origin	Concentration	Target
DAPI	AppliChem, Darmstadt	5 mg/ml	DNA

2.4 Oligonucleotides

All oligonucleotides used during this work were ordered from Sigma-Aldrich with a synthesis scale of 0.025 µmol and purification method desalinated. Sup. Table 7.1 contains a detailed list of all primers used during this work for: validation of the nuclear proteome (Chapter 4.2.3.3), validation of the NPG proteome (Chapter 4.2.7), expressing *C. elegans* P granule proteins in trypanosomes (Chapter 4.3.3.1), RNAi of RNST proteins (Chapter 4.3.5) and other primers used for sequencing or during the cloning process.

2.5 Plasmids

All plasmids used or generated during this work are listed in Sup. Table 7.2. Details about the cloning strategies that were used to generate plasmids are described in Chapter 3.4.

2.6 Chemicals

Chemicals were obtained from Carl Roth (Karlsruhe, DE), AppliChem (Darmstadt, DE) or Sigma-Aldrich (St. Louis, USA) if not stated otherwise in the text.

2.7 Antibiotics

Table 2.4: Overview of antibiotics used in trypanosome cell culture.

Antibiotic	Stock solution (mg/ml)	Working concentration (µg/ml)	Dilution
G418	10	15	1:666
Hygromycin	10	25	1:400
Phleomycin	10	2.5	1:4000
Puromycin	1	1	1:1000
Blasticidin	10	10	1:1000
Tetracycline	10	1	1:10000

2.8 Cultivation media

LB medium for *E. coli*

For 1 l medium: 10 g tryptone, 5 g yeast extract, 10 g NaCl, pH 7.4 (autoclaved). 100 µg/ml ampicillin was added for selection of transformed bacteria. For agar plates 1.5% agar was added to LB medium.

HMI-9 medium for BSF *T. brucei* (according to Hirumi and Hirumi)[71]

For 1 l medium: 17.66 g Iscove's modified Dulbecco's Medium (IMDM), 3.024 g $NaHCO_3$, 136 mg hypoxanthine, 28.2 mg bathocuproin sulfonate, 14 µl ß-mercaptoethanol, 39 mg thymidine, 100,000 U penicillin, 100 mg streptomycin, 182 mg cysteine, 10% (v/v) heat inactivated fetal calf serum (Sigma Aldrich); pH 7.5, sterile filtered, stored at -20°C.

SDM-79 medium for PCF *T. brucei*

For 10 l medium: SDM79-powder dissolved in 9 l ddH_2O, 20 g $NaHCO_3$, 18.642 g dextrose, 5.134 g L-glutamine, 6.15 g L-proline, 1 g sodium pyruvate, 4.068 g L-threonine, 105 mg sodium acetate (anhydrous), 224.2 mg L-glutamic acid, adjust pH with 10 M NaOH (pH 7.3). The medium was supplemented with 10% (v/v) heat inactivated fetal calf serum (Sigma Aldrich), 3 ml autoclaved heme (2.5 mg/ml in 0.05 M NaOH) and 10 ml penicillin-streptomycin-solution (penicillin 10 kU/ml, streptomycin 10 mg/ml).

NGM lite plates for *C. elegans*

For 500 ml medium: 1 g NaCl, 2 g tryptone, 1.5 g KH_2PO_4, 0.25 g K_2HPO_4, 10 g agar, 490 ml ddH_2O (autoclaved). 0.5 ml 5 mg/ml cholesterol in EtOH and 2.5 ml nystatin solution were added under sterile conditions, the NGM lite solution was poured into 60 mm petri dishes and the surface flamed with a Bunsen burner to remove bubbles and contaminations. After the agar had solidified it was seeded with *E. coli* strain OP50 liquid culture (4 drops out of a 10 ml pipette).

2.9 Buffers and solutions

Table 2.5: Buffers and solutions for working with *E. coli*.

Buffer/Solution	Composition
TSS	LB medium with 10% PEG 3350, 5% DMSO and 50 mM $MgCl_2$, sterile filtrated (pH 6.5)

Table 2.6: Buffers and solutions for working with *T. brucei.*

Buffer/Solution	Composition
Transfection buffer	90 mM NaH_2PO_4, 5 mM KCl, 0.15 mM $CaCl_2$, 50 mM HEPES (pH 7.3)
1x PBS	10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl (pH 7.4)
SDM-79 w.s.h.	SDM-79 medium without supplementation of fetal calf serum, heme and penicillin-streptomycin solution, used for washing
2x freezing mix for PCF *T. brucei*	SDM-79 with 20% (v/v) glycerol (sterile filtered, stored at 4°C)
2x freezing mix for BSF *T. brucei*	HMI-9 with 20% (v/v) glycerol (sterile filtered, stored at 4°C)
8% or 4% Paraformaldehyde	8 g or 4 g PFA in 100 ml PBS sterile, dissolved by heating for ≈ 10 min. 50µl 1 M NaOH can be added if pH is not 7.0

Table 2.7: Buffers and solutions for working with *C. elegans.*

Buffer/Solution	Composition
M9 medium	64 g Na_2HPO_4 $7H_2O$, 15 g KH_2PO_4, 2.5 g NaCl, 5.0 g NH_4Cl in 1 l ddH_2O (sterile)
Agarose	2% agarose in ddH_20

Table 2.8: Buffers and solutions for DNA work.

Buffer/Solution	Composition
Self-made buffer P1	50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 100 µg/ml RNase A
Self-made buffer P2	200 mM NaOH, 1% SDS
Self-made buffer P3	3 M potassium-acetate pH 5.5
10x DNA loading dye	0.1 M EDTA (pH 8.0), 0.05% (w/v) bromphenol blue, 0.05% (w/v) xylene cyanol, 40% (w/v) sucrose
DNA size marker	GeneRuler DNA Ladder Mix (Thermo Fisher Scientific)
Elution buffer	10 mM Tris-Cl (pH 8.5)
TAE buffer	40 mM TRIS, 40 mM concentrated acetic acid, 1 mM EDTA (pH 8.0)
Extraction buffer (EB buffer)	10 mM TRIS (pH 8.0), 10 mM NaCl, 10 mM EDTA (pH 8.0), 0.5% SDS
Ethidium bromide bath	3.3 mg/ml ethidium bromide (Applichem, A1152,0025) in ddH_2O

Table 2.9: Buffers and solutions for RNA work.

Buffer/Solution	Composition
Denhardt's solution (50x)	1% (w/v) Ficoll 400, 1% (w/v) polyvinylpyrolidone, 1% (w/v) BSA
SSC (20x)	3 M NaCl, 0.3 M trisodium citrate, pH 7.0
Hybridization solution	2% BSA, 5 x Denhardt's solution, 4x SSC, 5% dextran sulphate, 35% deionized formamide, 0.5 mg/ml tRNA yeast, 40 U/ml Ribolock, fill up with RNase free water
Self-made Affymetrix wash buffer	0.1x SSC, 0.1% SDS in ddH_2O

Table 2.10: Buffers and solutions for protein work.

Buffer/Solution	Composition
Coomassie stain	0.5 g Coomassie Brilliant Blue R 250, 50 ml acetic acid, 100 ml isopropanol, 850 ml H_20
Destain solution	50 ml acetic acid, 100 ml isopropanol, 850 ml H_20
Separating gel buffer	1.5 M Tris-Cl (pH 8.8), store at 4°C
Stacking gel buffer	0.5 M Tris-Cl (pH 6.8), store at 4°C
1 x SDS running buffer	20 mM Tris base, 192 mM Glycin, 0.1% (w/v) SDS

Buffer/Solution	Composition
1 x SDS sample buffer	2% (w/v) SDS, 80 mM Tris-Cl (pH 6.8), 10% (v/v) glycerol, 0.004% (w/v) bromphenol blue, 5% (v/v) beta-mercaptoethanol (add freshly); (boiling samples at 100°C)
1x NuPAGE LDS sample puffer (Thermo Fisher Scientific)	Complement with 100 mM DTT; (boiling samples at 70°C)
Separating gel mix	Exemplary for 15% gel: 12 ml acrylamide/bisacrylamide (37.5:1), 7.5 ml separating gel buffer, 10.5 ml ddH_2O, 100 µl 10% (w/v) APS, 20 µl TEMED
Stacking gel mix	1.95 ml acrylamide/bisacrylamide (37.5:1), 3.75 ml stacking gel buffer, 9.15 ml ddH_2O, 75 µl 10% (w/v) APS, 15 µl TEMED
Protein size marker	PageRulerTM Prestained Protein Ladder (Thermo Fisher Scientific)
Blotting buffer	20% (v/v) methanol in 1 x SDS Gel Running buffer
1x PBS-T	1x PBS with 0.1% Tween20

Table 2.11: Buffers and solutions for purification of trypanosome nuclei and nuclear envelopes.

Buffer/Solution	Composition
Protease inhibitor solution	Complete Protease Inhibitor cocktail tablets EDTA free (1187358001 Roche), 25x stock solution in ddH_2O.
PVP solution	8% (w/v) polyvinylpyrrolidone (PVP-40, Sigma-Aldrich), 11.5 mM KH_2PO_4, 8.5 mM K_2HPO_4, 750 µM $MgCl_2$, adjust to pH 6.53 with concentrated H_3PO_4 (~15 µl for 1 l), sterile filtrated, store at 4°C
Lysis buffer (1 volume = 20 ml)	0.05% Triton X-100, 5 mM DTT, 1x protease inhibitor solution (from 25x stock) in PVP solution, prepare fresh
Sucrose/PVP solutions	2.01 M: 183.3 g sucrose (D (+)-sucrose pure, Apllichem, A1125), add PVP solution to final weight 338 g (RI 1.4370); 2.10 M: 193 g sucrose, add PVP solution to final weight 340 g (RI 1.4420); 2.30 M: 216 g sucrose, add PVP solution to final weight 340 g (RI 1.4540). To dissolve the sucrose, solutions were heated to 37°C. Before measuring the refractive index (RI) with a refractometer, solutions were allowed to cool down to RT. The RI was adjusted by stepwise adding PVP while stirring the solution. The permitted error for the RI was 0.0003. Solutions were sterile filtrated and stored at -20°C.
Underlay buffer (1 volume = 10 ml)	5 mM DTT, 1x protease inhibitor solution (from 25x stock) in 0.3 M sucrose/PVP (dilute 2.1 M sucrose/PVP solution 1:7 in PVP solution), prepare fresh

Buffer/Solution	Composition
Resuspension buffer (1 volume = 8 ml)	5 mM DTT, 1x protease inhibitor solution (from 25x stock) in 2.1 M sucrose/PVP, prepare fresh
BT/Mg buffer	0.01 M bis-Tris-Cl (pH 6.50), 0.1 mM $MgCl_2$
Shearing buffer	1 mM DTT, 1 mg/ml heparin, 20 µg/ml DNaseI, 2 µg/ml RNase A, protease inhibitor in BT/Mg buffer
Sucrose solutions	2.1 M sucrose (D (+)-sucrose pure, Apllichem, A1125) in 20% Accudenz (Gentaur, AN7050) in BT/mg buffer
	2.5 M sucrose in BT/Mg buffer (RI: 1.4533)
	2.2.5 M sucrose in BT/Mg buffer (by diluting stock solution)
	1.50 M sucrose in BT/Mg buffer (by diluting stock solution)

Table 2.12: Buffers and solutions for microscopy.

Buffer/Solution	Composition
LR White Resin	Medium Grade Acrylic Resin, London Resin Company Ltd.
Epon	
Ponceau solution	0.1% PonceauS in 1% acidic acid
Elution solution	0.2 M glycine, 1 mM EDTA (pH 2.5)
Neutralization solution	1 M Tris-Cl (pH 8.0)

2.10 Devices and equipment

Table 2.13: Overview of used devices and equipment.

Device/Equipment	Supplier	Central office
Cell Culture		
Amaxa NucleofectorTM 2b Device	Lonza	Basel, CH
Electroporation cuvettes	BTX	Holliston, USA
CO_2 Incubator	Binder	Tuttlingen, DE
Steriflip filter	Merck	Darmstadt, DE
Vented cell culture flasks	Greiner Bio-One	Frickenhausen, DE
24-well suspension plates	Sarstedt	Nümbrecht, DE
Neubauer chamber	Marienfeld	Lauda-Königshofen, DE
Cryotubes (2 ml)	Sarstedt	Nümbrecht, DE

Device/Equipment	Supplier	Central office
Centrifuges and rotors		
Sigma 6-16 K, rotors: 11650, 12169, 12170	SIGMA Laborzentrifugen	Osterode, DE
Hettich Mikro 200	Hettich	Bäch, CH
Hermle Z 383 K	HERMLE Labortechnik	Wehingen, DE
Hermle Z 216 MK, rotor: 220.87	HERMLE Labortechnik	Wehingen, DE
Sorvall R6 plus, rotor: HB-6	Thermo Fisher Scientific	Waltham, USA
Beckmann L7, rotors: AH629, Ty50Ti	Beckman Coulter	Brea, USA
Shaker		
Shaker Certomat-S	B. Braun	Melsungen, DE
Shaker KL 2	Edmund Bühler	Hechingen, DE
Shaker KS-15	Edmund Bühler	Hechingen, DE
Microscopes and cameras		
iMIC	FEI-TILL Photonics	Gräfelfing, DE
PCO Sensicam qe	PCO	Kehlheim, DE
Nikon Eclipse TS100	Nikon	Tokio, JP
DMI6000 wide-field microscope	Leica	Wetzlar, DE
DNA, RNA and protein analysis		
Mini-PROTEAN Tetra Cell	Bio-RAD	Hercules, USA
Mini-Trans Blot Cell	Bio-RAD	Hercules, USA
Tecan Infinite M200 Plate reader	Tecan Group	Männedorf, Schweiz
Odyssey Infrared Imaging System	LI-COR Bioscience	Lincoln, USA
Immobilon-FL membrane	Merck Millipore	Burlington, USA
Bachofer UV-table, 302	Vilber Lourmat	Eberhardzell, DE
GeneRuler^{TM} DNA Ladder mix	Thermo Fisher Scientific	Waltham, USA
Enzymes		
Restriction enzymes	Thermo Fisher Scientific	Waltham, USA
Phusion High Fidelity DNA Polymerase	Thermo Fisher Scientific	Waltham, USA
Expand High Fidelity Polymerase	Roche Diagnostics	Rotkreuz, CH
Proteinase K	Serva	Heidelberg, DE
Klenow	Thermo Fisher Scientific	Waltham, USA

Device/Equipment	Supplier	Central office
T4 Ligase	Thermo Fisher Scientific	Waltham, USA
Fast AP	Thermo Fisher Scientific	Waltham, USA
DNase	Thermo Fisher Scientific	Waltham, USA
RNase A	Sigma-Aldrich	St. Louis, USA
Kits		
DNase I Set, RNase-free	QIAGEN	Venlo, NL
QIAGEN Plasmid Midi Kit	QIAGEN	Venlo, NL
QIAEXII Gel Extraction Kit	QIAGEN	Venlo, NL
RNeasy Mini Kit	QIAGEN	Venlo, NL
miRNeasy Mini Kit	QIAGEN	Venlo, NL
RevertAid First Strand cDNA Synthesis Kit	Thermo Fisher Scientific	Waltham, USA
CloneJET PCR Cloning Kit	Thermo Fisher Scientific	Waltham, USA
QuantiGene View RNA ISH Cell Assay Kit	Affymetrix	Santa Clara, USA
NucleoSpin Gel and PCR clean-up Kit	Macherey-Nagel	Düren, DE
Electron microscopy		
EM HPM100	Leica Microsystems	Wetzlar, DE
EM AFS2 freeze substitution system	Leica Microsystems	Wetzlar, DE
JEM-2100	JEOL	München, DE
TemCam F416 4k x 4k camera	Tietz Video and Imaging Processing Systems	Gauting, DE
Histo Jumbo Diamond Knife	Diatome AG	Nidau, CH
Others		
POLYTRON® homogenizer PT 1200E, PT-DA 12/2 EC-E123	Kinematica	Luzern, CH
Bioruptor Plus	Diagenode	Seraing, BE
PAP pen	Kisker	Steinfurth, DE

2.11 Software

Table 2.14: Overview of used software.

Software/Version	Supplier	Central office
4 Peaks	Mekentosj	Dordrecht, NL
Adobe Illustrator CS6	Adobe Systems	San José, USA
Adobe Photoshop CS6	Adobe Systems	San José, USA
CLC Main Workbench 6	CLC bio	Aarhus, DK
Fiji	rsbweb.nih.gov/ij/	
Huygens Essential	Scientific Volume Imaging	Hilversum, NL
iControl	Tecan Group	Männedorf, CH
ImageJ 64	rsbweb.nih.gov/ij/	
Image Studio Lite	LI-COR Biosystems	Licoln, USA
LA Acquisition	FEI-TILL Photonics	Gräfelfing, DE
Microsoft Office 2011	Microsoft	Redmond, USA
Papers2	Mekentosj	Dordrecht, NL
Mendeley Version 1.19.4	Mendeley Ltd.	London, UK

3 Methods

3.1 Working with *E. coli*

Generation of chemically competent *E. coli*

For long term storage: 200 ml LB medium was inoculated with 5 ml bacteria suspension of an overnight culture and incubated at 37°C on a shaker (150 rpm) until the OD600 reached 0.5-0.6. Bacteria were harvested (900x g, 4°C, 10 min), transferred to ice and resuspended in 5 ml TSS. In the cold-room, the bacteria suspension was immediately aliquoted a 100 µl in precooled Eppendorf tubes and the aliquots were shock-frozen in liquid nitrogen and stored at -80°C until usage. Competence of bacteria was tested by transfecting 1 pg, 10 pg, and 50 pg of the pUC vector, aiming a competence of at least $3x10^6$ (transfection of 10 pg pUC results in at least 30 colonies on plate).

For instant transformation: 10 ml LB medium in a 50 ml falcon was inoculated with a bacterial colony from an agar plate and incubated on a shaker (37°C, 150 rpm) until the OD600 reached 0.3-0.5. The bacterial culture was placed on ice for 10 min, harvested (3000 rpm, 10 min, 4°C), resuspended in 1/3 original volume TFB and kept on ice for 15 min. Then cells were centrifuged (3000 rpm, 10 min, 4°C) and resuspended in 0.8 ml TFB. 28 µl DND was added and after 10 min on ice another 28 µl DND was added. After 10 min on ice, cells were ready for transformation.

Transformation of chemical competent *E. coli*

Chemically competent *E. coli* Top10 were thawed on ice. They were either added to a plasmid (10 ng for re-transformation) or to 10 µl of a ligation mixture and incubated on ice for at least 30 min. Bacteria were heat shocked (42°C, 60 sec), spread on agar plates containing ampicillin (100 µg/ml) and incubated over night at 37°C. For difficult transfections (e.g. ligations of large plasmids) 1ml LB medium without ampicillin was added prior to plating and bacteria were incubated on a shaker (1 h, 37°C, 150 rpm). Bacteria were then harvested (3000 rpm, 10 min, 4°C), resuspended in 100µl LB medium containing ampicillin and proceeded further as described above.

Liquid bacteria culture

LB medium with ampicillin was inoculated with one bacterial colony from an agar plate and incubated overnight on a shaker (37°C, 150 rpm). For the preparation of plasmid DNA, 1.5 ml (Miniprep) or 25 ml (Midiprep) LB medium was incubated in an Eppendorf tube or in a conical flask, respectively.

3.2 Working with *T. brucei*

Cultivation of PCF

T. brucei PCF strains were cultured as suspension culture in SDM-79 medium at 27°C in a humid environment with 5% CO_2. The parasites were kept at densities between $5x10^5$ to $1x10^7$ cells/ml to ensure their logarithmic growth. The density of trypanosomes was determined with a Neubauer chamber. The parasites were harvested via centrifugation at 1400x g for 10 min at 4°C if not stated otherwise. Washing steps were also conducted in the same manner.

Cultivation of BSF

T. brucei BSF strains were cultured as suspension culture in HMI-9 medium at 37°C in a humid environment with 5% CO_2.The parasites were kept at densities below $1x10^6$ cells/ml to ensure their logarithmic growth. The density of trypanosomes was determined by counting with a Neubauer chamber. The parasites were harvested via centrifugation at 1400x g for 10 min at 4°C if not stated otherwise. Washing steps were also conducted in the same manner.

Freezing and thawing of PCF trypanosomes

For long-term storage 1 ml of PCF trypanosome culture at about $1x10^7$ cells/ml was mixed with 1 ml 2x freezing mix in a cryotube by inverting the tube twice. Then cells were frozen at -80°C immediately. For long term storage cryo-stabilates should be transferred to -150°C. Stabilates were thawed quickly in a 37°C water bath and 1 ml was transferred to 9 ml of fresh SDM-79. If required, antibiotic selection was added after ≈ 24 h.

Freezing and thawing of BSF trypanosomes

For long-term storage, approximately $4x10^6$ cells were harvested, resuspended in 1 ml HMI-9 medium and mixed with 1 ml 2x freezing mix. Cells were frozen at -80°C immediately. Stabilates were thawed quickly in a 37°C water bath and washed once with 10 ml of HMI-9. After centrifugation the supernatant was discarded, and cells were taken up in 10 ml fresh pre-warmed HMI-9 medium. After cultivation for 45-60 min the cell density was determined, and cells were diluted. If required, antibiotic selection was added after ≈ 24 h.

Transfection

Stable transgenic PCF trypanosomes were generated by electroporation. For one transfection $1x10^7$ cells were harvested at room temperature, resuspended in 400 µl transfection buffer and transferred into an electroporation cuvette containing the DNA (either 10 µg of linearized plasmid or

the PCR product of one PCR reaction in 10 µl sterile ddH_2O). Cells were electroporated using the *Amaxa Nucleofector II (Program X-001)*, immediately transferred to 25 ml fresh SDM-79 containing 20% FCS and cultivated overnight. The required selection was added after approximately 14 h and serial dilutions (1:10 and 1:100) with SDM-79 containing 20% FCS were plated onto 24-well plates, with each well containing 1 ml. After about 10-15 days, when wells started to get yellow-cloudy, cells were transferred into 10 ml culture flasks and cultivated with selection for at least another four days. For clonal cell lines, only plates that had cells in less than eight wells were chosen.

Treatments

Sinefungin: Cells were treated with 2 µg/ml sinefungin for 50 min. For the large-scale nuclear purification experiments 0.2 µg/ml were used.

PBS – Starvation: Cells were centrifuged for 3 min in a picofuge and resuspended in 1 ml PBS. From that point on the timer was set to 2 h. Cells were centrifuged again, resuspended in 1 ml fresh PBS and incubated at 27°C for exactly 2 h in total.

Heat-shock: Cells were incubated in a 1.5 ml Eppendorf tube at 41°C for exactly 2 h in a heating block.

Actinomycin D: Cells were treated with 10 µg/ml actinomycin D for 1 h, using a 1000x stock solution (10 mg/µl actinomycin D in ethanol).

Fixation

For microscopic analysis 1 ml of trypanosome culture with an approximate density of $5x10^6$ cells/ml was centrifuged (3 min, 1400 g) and resuspended in 190 µl SDM-79 w.s.h. 300 µl 4% paraformaldehyde (freshly thawed) was added for fixation. Samples were then stored at 4°C in the dark overnight. Prior to microscopic analysis cells were washed twice with 1 ml PBS and finally resuspended in 200 µl PBS.

DAPI staining

For staining of DNA in fixed trypanosomes, a 5 mg/ml stock solution (in ddH_2O) was diluted 1:10,000 and one volume was added to one volume of fixed trypanosome cells directly before microscopy.

3.3 Working with *C. elegans*

Cultivation of *C. elegans*

Worms were cultivated on NGM lite agar plates seeded with *E. coli* strain OP50 and incubated at 20°C or 15°C. Examination and transfer of worms were performed under a binocular. Worms were transferred to new plates regularly to maintain the culture by either picking (worm pick consisting of a platinum wire molten onto a Pasteur pipette) single worms, or "chunking" (sterilized spatula) a part of the old plate onto a new one. For expansion, worms were incubated at 20°C, for storage they were kept at 15°C.

DNA bombardment

Integrated transgenic lines were created by microparticle bombardment of unc-119(ed3) according to established protocols[72] in the lab of Ann Wehman (Rudolf-Virchow-Centre, Würzburg, Germany).

Treatments

To inhibit splicing, worms were treated with spliceostatin A stock solution 200 µg/ml in methanol, further diluted with M9) at either 2.5 µg/ml or 5 µg/ml. As controls, worms were incubated with either M9 medium only or with 2.5% methanol in M9. Young adult worms were freshly picked from a plate and transferred to a 2 ml Eppendorf tube containing 40 µl of the respective solution. Worms were incubated for 5 h on a shaker to enable oxygen exchange and were provided with bacteria to prevent starvation.

Preparation for microscopic analysis

Microscopic slides with small agarose pads (2% agarose in ddH_2O) were prepared and worms were picked into a drop of either 25 mM sodium azide to immobilize the worms or into M9 cultivation medium. To investigate *C. elegans* embryos or gonad arms, the adult worm was cut with a razor blade.

Isolation of RNA from *C. elegans*

Ten *C. elegans* worms were picked from a fresh NGM plate and washed twice with M9 medium (1500 rpm, 4°C) in a 1.5 ml RNase free Eppendorf tube. From now on all steps were conducted under a fume cupboard. To lyse the cells 700µl QIAzol lysis reagent (Qiagen) were added to the pellet. Then the sample was vortexed (30 sec) followed by 5 min incubation time at RT. Next, 140 µl chloroform were added. The sample was mixed by vigorous shaking (15 sec) and incubated for 3 min at RT. After centrifugation (12,000x g, 15 min, 4°C) the upper, aqueous phase was transferred to a new Eppendorf

tube; 525 µl 100% ethanol (1.5-fold volume) was added and the sample mixed by pipetting. From this point on work was continued on the bench. The sample was transferred to a spin column (RNeasy Mini Kit, Qiagen), centrifuged (10,000 rpm, 15 sec, RT) and the flow through discarded. Then the column was washed with 350 µl RWT buffer (miRNeasy Mini Kit, Qiagen) (10,000 rpm, 15 sec, RT). In a next step the DNA present in the sample was digested on the column for 15 min with a mixture of 10 µl DNase I and 70 µl RDD buffer (RNase-Free DNase Set, Qiagen). After DNA digestion 350 µl RW1 buffer (RNeasy Mini Kit, Qiagen) were added to the column, the sample was centrifuged (10,000 rpm, 15 sec, RT) and the flow through was discarded. The procedure was repeated twice with 500 µl RPE buffer (RNeasy Mini Kit, Qiagen), with the last centrifugation step being extended to 2 min to dry the column. The spin column was placed in a new 1.5 ml collection tube and 30 µl RNase free water was added to elute the RNA (10,000 rpm, 1 min, RT). RNA concentration and purity were determined by measuring the absorption of the solution at 260 nm and 280 nm using the Tecan Infinite M200 Plate reader. RNA was stored at -80C°.

cDNA synthesis

For first strand cDNA synthesis the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) was used, following the manufactures' instructions. For the protocol, 0.1-5 µg total *C. elegans* RNA (isolation described above) was used as template and Oligo(dT)18 as primer. GAPDH RNA was used as a control. cDNA was stored at -20°C. 2 µl of cDNA were used for subsequent standard PCR protocols.

3.4 Cloning strategies

Below, three general cloning strategies employed for the generation of new plasmids are described. All plasmids used or newly generated during this work are summarized in more detail in Sup. Table 7.2.

3.4.1 Endogenous tagging

'Traditional' tagging after Kelly et al. 2007[6]

Some proteins surveyed for the validation of the NPG proteome (compare chapter 4.2.7, details see Sup. Table 7.7 A, for details on plasmids see Sup. Table 7.2) were C-terminally tagged with eYFP-4xTy1. Also, NUP93 was either C-terminally tagged with mChFP-4xTy1 (chapter 4.2.7) or eYFP-4xTy1 (chapter 4.1.4 and 4.3.2).

All plasmids were designed according to[6], which allowed the directional cloning of fragments derived from blunt/BamHI or BglII compatible fragments. DNA fragments were generated by PCR amplification with primers specifically designed for each gene of interest, covering the complete ORF excluding the stop codon. Primers were designed to add HindIII and BamHI compatible ends[6]. In some cases, the sequence of the gene of interest contained one or more BamHI restriction sites and the reverse primer was designed to add a BglII site instead: Forward primer: 5'-**AAGCTT**CCGCCACC**ATG** followed by the next 15 bases of the ORF; the HindIII site and the initiation codon are shown in bold. Reverse primer: 5'-TGATCA**GGATCC**AGAACC followed by reverse complement of last 18 bases of the ORF, excluding the stop codon; the BamHI site is shown in bold. Alternative reverse primer: 5'-TGATCA**AGATCT**AGAACC followed by reverse complement of last 18 bases of the ORF; the BglII site is shown in bold. The PCR products were subsequently cloned into a pJET 1.2 cloning vector for sequencing. Afterwards, fragments were excised from the cloning vector with a suitable blunt cutter (cutting within the gene of interest) and BamHI or BglII (depending on the target gene derived restriction enzyme fragment), so that DNA fragments with a length between 300 bp and 1500 bp emerged. Those fragments were then cloned into the suitable expression vector. SK141[34] was used as mother plasmid for tagging with eYFP, SK197 (produced during this work) was used for mChFP tagging. Both expression vectors were opened for insertion of fragments by cutting with SwaI/BamHI. The correct size of all fusion proteins was confirmed by western blots (data not shown). Expression vectors were linearized with a suitable single-cutter restriction enzyme that cut within the gene of interest's sequence (trypanosomes require only 50 nt of identity for targeted homologous recombination) for subsequent transfection and genomic integration in *T. brucei*.

PCR-tagging after Dean et al.2015[73]

Most proteins of the NPG proteome (chapter 4.2.7, details see Sup. Table 7.7 A) and all proteins of the nuclear proteome (chapter 4.2.3.4), as well as proteins analyzed under Mex67 RNAi conditions (chapter 4.3.5) were C-terminally tagged with eYFP using the method by Dean et al. 2015.

With this system, PCR products of a long-primer PCR (amplicons) can be directly transfected into trypanosomes. Primers were designed as described in Dean et al. 2015 and consisted (1) of the last 80 nt of the target ORF excluding the stop codon followed by the first 18 nt of the 5' GS linker sequence (5' to 3' forward primer) or (2) of the first 80 nt of the 3' UTR of the target gene in reverse complement followed by the 20 nt 3' pPOTv4 primer binding sequence in reverse complement (5' to 3' reverse primer). Primers were used in a hot start PCR with pPOTv4[73] (SK342) as template for eYFP tagging (protocol see Table 3.2) and subsequently transfected into *T. brucei*.

3.4.2 Inducible expression

Experiments for inducible expression of *C. elegans* proteins (chapter 4.3.3.1) were performed in trypanosome Lister 427 pSPR2.1 cells expressing the tetracycline repressor TetR[69].

DNA fragments were generated by PCR with primers specifically designed for each gene of interest (CGH-1, PGL-2, OMA-2 and POS-1) covering the complete ORF excluding the stop codon. Primers were designed as described above to add HindIII and BamHI or BglII compatible ends[6]. The generated fragments were subsequently cloned into a pJET 1.2 cloning vector. Fragments were then excised with the respective restriction enzymes (HindIII/BamHI or BglII) and cloned into the expression vector p3888[74], which was modified from p3383[69] for C-terminal eYFP tagging. Expression vectors were linearized with NotI for subsequent transfection into *T. brucei*.

3.4.3 RNAi

Experiments for RNAi against RNST proteins (chapter 4.3.5 and 4.3.6) were performed in Lister 427 pSPR2.1 cells expressing the tetracycline repressor TetR[69]. DNA fragments were generated by PCR with primers specifically designed for each gene of interest (Tb927.10.7580, ZFP1, ZFP2, RBP12, ZC3H29, Tb927.11.6600 and DRBD3) covering the complete ORF excluding the stop codon. Primers were designed as described above to add HindIII and BamHI or BglII compatible ends[6]. The generated fragments were subsequently cloned into a pJET 1.2 cloning vector for sequencing. Afterwards, fragments were excised from the cloning vector with a suitable blunt cutter (cutting within the gene of interest) and BamHI or BglII (depending on the target gene derived restriction enzyme fragment), so that DNA fragments with a length between 250 bp and 1000 bp emerged. These fragments were then cloned into the mother plasmid p3666[69], which was designed to express a RNA stem loop upon tetracycline induction. The vector contains a spacer sequence, which is flanked by two stem multiple cloning sites, that allow directional cloning of two fragments derived from blunt/BamHI and BclI/blunt compatible fragments[6], so that they have opposite orientations and produce a stem loop upon expression. The same target gene derived restriction enzyme fragment was used in two successive rounds of sub-cloning to produce the final stem loop RNAi plasmid. The expression vector was first opened with SwaI/BclI and after inserting the fragment again opened with PmeI/BamHI for a second insertion. For use of the BclI restriction enzyme, which is sensitive to Dam-methylation, plasmids were isolated from Dam negative *E. coli*. Expression vectors were linearized with NotI for subsequent transfection into *T. brucei*.

3.5 Working with DNA

Isopropanol precipitation

One volume of isopropanol was added to the DNA solution, as well as 3 M sodium acetate (1/10 of the total volume). Subsequently, the mixture was centrifuged (14,000x g, 20 min, 4°C), the supernatant discarded, and the pellet washed with 500 µl 70% ethanol (14,000x g, 10 min, 4°C). The pellet was air dried and dissolved in the required volume of ddH_2O. If the DNA was used for transfection of *T. brucei* all steps following the ethanol washing were conducted under sterile conditions and the DNA was solved in 10 µl sterile ddH_2O per transfection.

Isolation of plasmid DNA from *E. coli*

Small amounts of plasmid DNA were extracted from 1.5 ml of an overnight bacteria culture using an alkaline lysis protocol similar to the QIAGEN Plasmid Mini Kit, but with self-made buffers. Bacteria were pelleted (1 min, maximum speed, RT) and resuspended in 250 µl buffer P1. First 250 µl buffer P2 and then 250 µl buffer P3 were added and the sample was mixed by inverting 5 times, respectively. After centrifugation (10 min, maximum speed, RT), the supernatant was transferred to a new Eppendorf tube and mixed with 700 µl isopropanol; the DNA was precipitated as described above. The low purity of this extraction method was sufficient for analytical digestions and subsequent sequencing (GATC Biotech).

For high-quality plasmid DNA (e.g. for transfections) a 25 ml overnight culture and the QIAGEN Plasmid Midi Kit were used following the manufactures' instructions. DNA concentration and purity were determined by measuring the absorption of the solution at 260 nm using the Tecan Infinite M200 Plate reader. The plasmid DNA was stored at -20°C.

Isolation of genomic DNA from *T. brucei*

For isolation of genomic DNA about $8x10^7$ trypanosomes were harvested, transferred into a 2 ml Eppendorf tube and once washed once with 1 ml of SDM w.s.h. The pellet was stored at -80°C until the extraction was continued. For the following phenol/chloroform extraction 500 µl EB buffer containing 100 µg Proteinase K was added to the cell pellet without resuspending it and incubated at least 2 h at 37°C. Then 1 ml phenol was added, and the sample mixed by vigorous shaking. After separating the aqueous phase from the phenolic phase by centrifugation (21,000x g, 3 min, 4°C) the aqueous phase was transferred into a new 2 ml Eppendorf tube. The procedure was first repeated with 1 ml of a 1:1 phenol chloroform mixture and then with 1 ml chloroform. The aqueous phase was transferred to a new Eppendorf tube and finally the genomic DNA was isolated via isopropanol

precipitation using 1/25 volume 5 M NaCl instead of sodium acetate and resuspended in 100 µl ddH_20. Genomic DNA was stored at 4°C.

Agarose gel electrophoresis and gel extraction of DNA fragments

DNA fragments or PCR products were separated by agarose gel electrophoresis (0.8% agarose (peqGOLD Universal Agarose, Peqlab) in TAE buffer for fragments up to 500 bp and up to 2% agarose in TAE buffer for fragments below 500 bp). For subsequent gel extraction of DNA fragments Ultra-Pure Agarose (Invitrogen) was used. Samples were mixed with 10x loading dye prior to loading onto the gel and the GeneRuler DNA Ladder Mix (Thermo Fisher Scientific) was used as size reference. Agarose gels were stained in ethidium bromide (3.3 mg/ml in ddH_20) (Applichem, A1152,0025), DNA was visualized by UV-light (312 – 350 nm) and recorded using a gel documentation system. If DNA fragments were excised for cloning, gels were visualized with a low radiant intensity during cutting to minimize DNA damage. Then, DNA was purified from the gel according to the manufactures' protocol using the QIAquick Gel Extraction Kit (Qiagen).

Restriction enzyme digestion and modification of DNA

Restriction enzyme digestions were used to linearize plasmids for transfection, to analyze the correct integration of an insert into a vector or to generate suitable overhangs for the following ligation of insert and backbone. Digestions with two different enzymes were either conducted as double digestion, with both enzymes at once or sequential digestion, with an isopropanol precipitation in between, both following the manufactures' instructions. In order to prevent re-ligation of blunt ended vectors and those opened with one enzyme, the phosphate groups were removed with the thermo-sensitive alkaline phosphatase Fast AP (Thermo Fisher Scientific). If 5' or 3' restriction site overhangs were not compatible for subsequent cloning, they were either filled up or removed by Klenow fragment (Thermo Fisher Scientific).

Ligations of plasmid fragments

To ligate DNA fragments into a suitable vector, a total amount of 50 ng was used at a molar ration of 3:1 (insert:backbone). Ligations were performed with the T4-ligase (Thermo Fisher Scientific) at room temperature overnight. The ligation mixture was directly used for *E. coli* transformation.

Polymerase chain reaction (PCR)

DNA fragments were amplified from genomic or plasmid DNA, separated by agarose gel electrophoresis and excised from the gel for further use. Either fragments were used for ligation into

the pJET1.2 cloning vector using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific) or were directly transfected into trypanosomes[73].

Table 3.1: Standard PCR protocol (for ligation into pJET1.2).

Pipetting scheme	
DMSO	2.5 µl
5x HF buffer (Thermo Fisher Scientific)	10 µl
dNTPs (10 mM)	1 µl
forward primer (100 µM)	1 µl
reverse primer (100 µM)	1 µl
plasmid DNA or	10–100 ng
genomic DNA	1 µl
Phusion High Fidelity DNA Polymerase (2 U/µl)	1 µl
ddH_2O	add to 50 µl

PCR program			
initial denaturation	94°C	3 min	
denaturation	94°C	10 sec	30 x
annealing	65°C	30 sec	
elongation	72°C	30 sec per 1 kbp	
final elongation	72°C	10 min	

Table 3.2: Long-primer PCR protocol (for direct transfection)[73].

Pipetting scheme	
DMSO	1 µl
10x buffer nr. 2 (Roche)	5 µl
dNTPs (10 mM)	1 µl
forward primer (100 µM)	0.5 µl
reverse primer (100 µM)	0.5 µl
SK342 (25 ng/µl)	1 µl
ddH_2O	40.5 µl
Once mixture has reached 94°C add Phusion High-Fidelity DNA Polymerase	0.5 µl

PCR program			
initial denaturation	94°C	5 min	
denaturation	94°C	15 sec	30 x
annealing	65°C	30 sec	
elongation	72°C	2 min	
final elongation	72°C	7 min	

3.6 Working with RNA

mRNA FISH

$1x10^8$ PCF trypanosomes were harvested, washed once with SDM79 w.s.h., resuspended in 1 ml SDM79 w.s.h. and fixed in 4% paraformaldehyde for 30 min with gentle rotation at RT. After the addition of 12 ml PBS, cells were pelleted (1400x g, 5 min) and resuspended in 1 ml PBS. A PAP pen was used to draw up to four hydrophobic circles on baked Superfrost® Plus slides (Thermo Fisher Scientific). Cells (about 100 µl cell suspension per circle) were allowed to settle down for 20 min on the slide, controlled by phase contrast microscopy. Slides were washed 10 min in 25 mM NH_4Cl in a coplin jar. All further incubation steps were performed in a sealed humid chamber. Samples were blocked and permeabilized with 0.5% saponin/2% BSA for 1 h and washed quickly with PBS before pre-hybridization with hybridization solution (2 h, 37°C). Slides were incubated with fluorescently labelled oligonucleotide-probes diluted 1:100 in hybridization solution (overnight, 37°C, in the dark). Slides were washed successively in 4x SSC/ 35% formamide, 4x SSC, 2x SSC and PBS; each washing step was done for 5 min. To combine mRNA FISH with immunofluorescence, a standard IF protocol was included here, starting with the blocking step. Finally, DNA was stained with DAPI (1 µg/ml in PBS) for 1 min. Slides were washed once in PBS before mounting the cover slips with Fluor Save (Calbiochem, Merck Millipore). Slides were stored in the dark at 4°C.

Single cell mRNA FISH – Affymetrix

To visualize single mRNA molecules in trypanosomes a modified protocol of the QuantiGene ViewRNA ISH Cell Assay Kit (Affymetrix, Santa Clara) was used[34]. Ideally, a 1100 nt sequence of the gene of interest was used by Affymetrix to generate specific probe sets for the experiments. In some cases, shorter sequences had to be used for example to generate a probe set to the tubulin intergenic region, which is only 158 nt long. In general, the probe consisted of a mixture of up to 20 oligonucleotide pairs each covering a 40 nt fragment of the target mRNA thereby leading to a highly

amplified signal, as a 1100 nt mRNA fragment could be hybridized with up to 20 oligonucleotide pairs in parallel. For the assay, about $1x10^8$ cells (previously treated or untreated e.g. with sinefungin) were harvested, washed once with SDM79 w.s.h., resuspended in 1 ml PBS and fixed in 4% paraformaldehyde for 10 min with gentle rotation. After the addition of 13 ml PBS, cells were pelleted again and resuspended in 1 ml PBS. A PAP pen was used to draw up to four hydrophobic circles on baked Superfrost® Plus slides (Thermo Fisher Scientific). Cells (about 100 µl cell suspension per circle) were allowed to settle down for 20 min. Since the PAP pen gets destroyed during the assay and the probe spreads between different circles, only one probe set combination can be used on one slide. Next, slides were washed once in PBS for 5 min in a coplin jar. Unless stated otherwise, all washing steps were performed in this manner. The residual PBS was removed by a tissue and 50 µl detergent solution QC was added to each sample (5 min, RT). After two washing steps with PBS, 100 µl of protease solution (1:6000 in PBS) was added to each circle for exactly 10 min at RT. For this step the room temperature should be close to 25°C. Since the protease activity can strongly vary, a serial dilution should be tested when performing the assay for the first time. For that reason, only samples from the same experiment can be compared with each other. Slides were washed three times with PBS. For target hybridization the probe set was diluted 1:100 in Probe Set Diluent and pre-warmed to 40°C. Then 100 µl were applied to each circle and slides were incubated for 3 h at 40°C in a sealed humid chamber. Afterwards slides were washed three times with self-made wash buffer (0.1x SSC, 0.1% SDS in ddH_2O) and once with 20 ml Affymetrix wash buffer (plastic bowl instead of coplin jar to allow the smaller volume (reduction of costs), gentle shaking). The signal was then amplified by first applying the PreAmplifier Mix solution, second the Amplifier Mix solution and finally the Label Probe Mix solution. At each step 100 µl of the solution was applied to each circle and slides were incubated for 30 min at 40°C in a sealed humid chamber. After each incubation step, the slides were washed three times with self-made wash buffer and once with 20 ml Affymetrix wash buffer (as above). Finally, DNA was stained with DAPI (1 µg/ml in PBS) for 2 min at RT. Slides were washed twice with PBS. All residual PBS was removed by a tissue before mounting the cover slips with Prolong® Gold Antifade (Thermo Fisher Scientific). After drying over night at room temperature, slides were stored at 4°C in the dark until image acquisition.

For some experiments Affymetrix single cell mRNA FISH was combined with an immunofluorescence. In this case, slides were washed twice with PBS after the last washing step of the FISH experiment (before DAPI staining), blocked with 2% BSA in PBS for 60 min at RT, incubated with the primary antibody (diluted in 2% BSA in PBS) for 60 min at RT, washed four times in PBS, incubated with secondary antibody (diluted in 2% BSA in PBS) and washed four times with PBS. All

steps were performed protected from light. DAPI staining was omitted, if blue secondary antibodies were used. Note that the 3 h target hybridization with the Affymetrix probes was conducted over night.

3.7 Working with proteins

Whole protein lysates from *T. brucei*

Trypanosomes were harvested, washed twice with SDM-79 w.s.h. and resuspended in 1x SDS sample buffer to a concentration of 1x or 5x 10^5 cells/µl. The sample was boiled 5 min at 100°C and stored until analysis at -20°C or -80°C for long-term storage.

Protein precipitation and sample preparation (NPG purification)

Samples from the NPG purification were methanol/chloroform precipitated and sonicated before analyzing them on an SDS gel or western blot. To precipitate proteins, 200 µl of a sample was stepwise mixed and vortexed vigorously with 800 µl methanol, 150 µl chloroform and 450 µl ddH_2O. After centrifugation (max. speed, 5 min) the upper, aqueous phase was discarded, another 650 µl methanol was added and the sample was mixed by inverting three times. Centrifugation was repeated, and the supernatant discarded. The pellet, containing the proteins, was then air-dried and resuspended in 100 µl 1x SDS sample buffer or 1x NuPAGE LDS (Thermo Fisher Scientific) sample puffer and boiled at 99°C (SDS buffer) or 70°C (NuPAGE LDS buffer) for 5 to 10 min. Subsequently, samples were sonicated with a Bioruptor Plus (Diagenode) sonication device at settings: high, 10 cycles, 30 sec ON/30 sec OFF.

Discontinuous SDS polyacrylamide gel electrophoresis

SDS polyacrylamide gel electrophoresis (SDS PAGE) was performed according to standard protocols. 1 x SDS running buffer and a voltage of 100 V was used to separate proteins according to their molecular weight in a one-dimensional gel using the Mini-PROTEAN system (Bio-Rad). Protein gels containing 12.5% acrylamid/bisacrylamid were used to resolve proteins with a size between 20-150 kDa. Proteins with a size between 5-70 kDa were separated on a 15% or 16% SDS gel. For analysis of whole cell lysates 1x10^6 cells were applied per lane. If samples of the NPG purification were analyzed, the amount of loaded cell equivalents was adapted to the type of the sample; details are described with the respective experiments. The PageRulerTM Prestained Protein Ladder (Thermo Fisher Scientific) was used as size marker of the proteins. Subsequently, gels were either stained with Coomassie Blue or specific proteins were further analyzed by western blotting.

Staining of proteins in SDS gels

Total proteins separated via SDS PAGE were stained with Coomassie Blue for 1 h and incubated in de-staining solution over night until no background staining was visible. Results were documented with the Odyssey Infrared Imaging System (LI-COR Bioscience).

Western Blot

After SDS PAGE proteins were transferred to an Immobilon-FL membrane (Merck Millipore) (1 h, 350 mA) via wet/tank blotting in 1x blotting buffer using the Mini-Trans Blot system (Bio-Rad). Western blots were performed according to standard protocols, using the manufactures' instructions. All components were soaked in blotting buffer prior to assembling the system and the membrane was activated with 100% methanol. Afterwards proteins were detected as described below.

Detection of proteins on membrane

After western blotting, the membrane was blocked in 3% (w/v) milk in PBS for 1 h at RT or overnight at 4°C. Next, the membrane was incubated with the primary antibody diluted in 1% (w/v) milk in PBS-T for 1 h at RT and washed 3x 10 min in 1x PBS-T. Afterwards the membrane was incubated with the secondary antibody diluted in 0.02% SDS/ 1% milk in PBS-T for 1 h at RT in the dark. The membrane was washed 3x 10 min in 1x PBS-T and dried between Whatman papers. Antibodies and dilution instructions for western blots are listed above (see chapter 2.2 and 2.9). Proteins were detected and quantified by the Odyssey Infrared Imaging System (LI-COR Biosciences). Background method for quantification: the average of a three-pixel width line at the top and bottom of each band was subtracted from each pixel.

Antibody affinity purification

For immuno-gold labelling of EM sections primary antibodies were affinity purified. A standard SDS PAGE with 1 to 10 µg of the recombinant protein was conducted and subsequently blotted onto an Immobilon-FL membrane (Merck Millipore) by standard procedure. The membrane was stained with Ponceau solution (5 min) and quickly washed with ddH_2O. A membrane piece containing the protein was cut out, transferred to an Eppendorf tube and incubated with 5% (w/v) milk in PBS (1 h, RT, rotating). Then the membrane was washed four times with PBS, followed by incubation with 0.5 ml antiserum (3 h, RT, rotating). Again, the membrane was washed four times with PBS and finally the antibody was eluted with 1 ml elution solution for exactly 10 min. The eluate was transferred into a new Eppendorf tube containing 200 µl neutralization solution, as well as 200 µg/ml BSA for stabilization. Finally, the antibody was dialyzed against PBS for 24 h, with three changes of the PBS.

The antibody was stored at 4°C containing 0.03% NaN3. The membranes were washed twice in PBS, dried and stored at -20°C for reuse.

3.8 Purification of trypanosome nuclei and nuclear envelops

Nuclei

The purification protocol was based on the purification of trypanosome nuclei described 2001 by Field and Rout[75] and 2008 by DeGrasse et al.[76]. For each purification approximately 10^{10} PCF cells at about $6x10^6$ cells/ml were grown in 2 l SDM79 medium in sealed conical glass flasks (5 l volume) with gentle shaking. For this, it was essential to use cells of a densely grown culture (at least $2x10^7$ cells/ml) to set up the 2 l culture at a density of $1.5x10^6$ cells/ml 24 h before starting the purification. Cells were either treated with SF (200 ng/ml) or left untreated. NPG formation was monitored microscopically (the cells expressed a fluorescent granule marker protein). Due to the large volume, cells were then pelleted stepwise. For this, 2 l of cell suspension were first pelleted in four 800 ml centrifuge buckets (1700x g, 10 min, 27°C, swing out rotor 11650, Sigma 6-16K), the supernatant discarded until about 100 ml volume was left, cells were resuspended and transferred to about ten 50 ml falcons. Cells were pelleted again (1400x g, 10 min, 27°C, Hermle Z 383 K) and finally pooled in two 50 ml falcons where they were washed twice with 50 ml SDM79 w.s.h. (1400x g, 10 min, 27°C, Hermle Z 383 K). From this point on, all further steps were conducted in a cold room. Cells were resuspended on ice in 20 ml lysis buffer and disrupted with a POLYTRON® homogenizer (PT 1200E, PT-DA 12/2 EC-E123, Kinematica AG, Switzerland) for at least 5 min at 2/3 of its maximum speed. Cell lysis was monitored by brightfield and fluorescence microscopy. Subsequently, the whole cell lysate was underlaid with 10 ml underlay buffer in a 30 ml COREX (No. 8445) glass tube (with a long syringe) and centrifuged (10,500x g, 20 min, 4°C, rotor HB-6 in a Sorvall R6 plus centrifuge). The supernatant, containing mainly soluble material from the cytosol, was decanted and stored at -80°C. The pellet, containing all crude cell material, was immediately resuspended in 8 ml resuspension buffer followed by further homogenization with the POLYTRON® (5 min, 2/3 of maximum speed). After complete cell lysis (monitored by brightfield and fluorescence microscopy), the material was loaded on a three-step sucrose gradient (2.01 M/ 2.1 M/ 2.3 M sucrose in PVP, 8 ml respectively) in a Sorvall AH629 rotor tube (PA, thin-wall, 38.5ml, No 253050). The gradient was filled to the rim with 2.1 M sucrose/PVP. After ultracentrifugation (25,000 rpm, 3.5 h, 4°C, Beckmann L7 centrifuge), the gradient was harvested from the top (with cut off blue pipette tips) in 2 ml fractions. The ring-shaped pellet at the bottom of the tube was resuspended in 2 ml 2.3 M sucrose/PVP. To monitor the purification, samples from each step (total cell lysate,

supernatant, resuspended pellet, and fractions of the gradient) were taken and stored at -80°C for further analysis.

Nuclear envelopes

The separation of nuclear membranes from purified nuclei of cells treated with SF was done exactly as described in DeGrasse et al. 2009[77]. The purified nuclei from cells treated with SF were supplemented with PVP, vortexed and centrifuged in a sealed Ty50Ti Rotor tube (PA, thin-wall, 11.5ml) (40,000 rpm, 1 h, 4°C). The pellet was resuspended in 3 ml shearing buffer and nuclear envelopes were sheared by vigorous vortexing. After a 5 min incubation at RT, 10 ml 2.1 M sucrose in 20% Accudenz (Gentaur, AN7050) in BT/Mg was added and the sample was mixed by vortexing. Then the mixture was transferred to a Sorvall AH629 rotor tube (PA, thin-wall, 38.5ml, No 253050) and overlaid with first 12 ml 2.25 M sucrose and second 10 ml 1.50 M sucrose. After centrifugation (25,000 rpm, 4 h, 4°C), the gradient was harvested from the top and all fractions were analyzed microscopically.

3.9 Mass spectrometry analysis (NPG proteome)

All mass spectrometry analyses were performed by Mario Dejung and Falk Butter at the Institute of Molecular Biology (IMB), Mainz, Germany. The samples were in-gel digested and MS measurement performed as previously described[78] with the following adaptations: the measurement time per sample was extended to 240 min. The triplicates were analyzed with MaxQuant version 1.5.0.25[79] with standard settings except LFQ quantitation and match between runs was activated. The trypanosome protein database TREU927 version 8.0 (11,567 entries) was downloaded from https://tritrypdb.org[80]. Further analysis was conducted in the Perseus environment[81] with filtering for proteins only identified by site, reverse entries, potential contaminants and quantitation values in at least 2 of the 3 replicates. Prior to imputation of missing LFQ values with a normal distribution (width 0.3, downshift 1.8), the LFQ values were log2 transformed. Significant enriched proteins were determined using a Welch t-test with 250 randomizations at FDR=0.05 and s0=0.5. The data was exported (Sup. Table 7.7 C). The volcano plot (Figure 4.11) was generated with the R ggplot2 package.

3.10 Electron microscopy

High pressure freezing (HPF)

For high pressure freezing, at least $2x10^7$ cells (treated with sinefungin or untreated) were pelleted (750x g, 3 min, RT), washed once with 50% FCS/SDM-79, and transferred into a freezing chamber

(specimen carriers type A, 100 µm, covered with specimen type B, 0 µm, Leica Microsystems). Protocols for high-pressure freezing and freeze substitution were adapted from Weimer 2006[82] and performed as described in Stigloher et al. 2011[83]. The samples were cryo-immobilized with an EM HPM100 (Leica Microsystems) at >20,000 K/s freezing speed and >2100 bar pressure and immediately stored in liquid nitrogen in an EM AFS2 freeze substitution system (Leica Microsystems) which afterwards enabled the gradually raising of the temperature for freeze substitution.

Freeze substitution and embedding in Epon

Samples were incubated in 0.1% tannic acid and 0.5% glutaraldehyde in anhydrous acetone at -90°C for 96 h, washed four times for 1 h with anhydrous acetone at -90°C, and fixed in 2% OsO_4 in anhydrous acetone at -90°C for 28 h. Then the temperature was gradually raised to -20°C within 14 h, samples were incubated at -20°C for 16 h and again the temperature was gradually raised to 4°C within 4 h. Afterwards samples were immediately washed with anhydrous acetone at 4°C four times at 0.5 h intervals, followed by gradually increasing the temperature to 20°C within 1 h. Subsequently, samples were transferred for embedding into a series of freshly prepared solutions with increasing concentrations of Epon (50% Epon in acetone for 3 h at room temperature, 90% Epon in acetone overnight at 4°C, followed by 2 times 100% Epon at room temperature for 2 h). Epon infiltrated samples were polymerized for 72 h at 60°C.

Freeze substitution and embedding in LR-White

Samples were incubated in 0.1% $KMnO_4$ in anhydrous acetone at −90°C for 65, the solution was exchanged once with fresh 0.1% $KMnO_4$ in anhydrous acetone and temperature was ramped for 11 h to −45°C. At −45°C samples were washed four times within 3 h with anhydrous acetone, one time with 1/3 ethanol in acetone for 30 min, one time 2/3 ethanol in acetone for 30 min and finally two times with 100% ethanol for 30 min each. Afterwards the temperature was gradually raised to 4°C within 16 h and the samples were washed twice with 100% ethanol for 30 min each. Subsequently, samples were transferred for embedding into a series of freshly prepared solutions with increasing concentrations of LR-White (50% LR-White in ethanol for 3 h, at 4°C, fresh 50% LR-White in ethanol over night at 4°C, followed by three times 100% LR-White for 1 h, 3 h and overnight at 4°C). LR-White infiltrated samples were polymerized for 72 h at 52°C. If LR-White contained additional accelerator, samples were cured for at least 48 h at 4°C in UV light and then further 48 h at RT in daylight.

Serial sections

60 nm serial sections were cut with a histo Jumbo Diamond Knife (Diatome AG) and placed on either pioloform coated slotted copper grids (no subsequent immuno-gold labelling) or 100 mesh nickel grids (subsequent immuno-gold labelling) depending on the subsequent analysis.

Immuno-gold labelling

For immuno-gold labelling sections on nickel grids were rehydrated with PBS (15 min) and blocked with 0.1% BSA/ 0.1% Tween20 (25 min), then stained with an affinity purified anti *T. brucei* SCD6 antibody (2 h, RT, 1:500 in PBS) followed by four washing steps with 1% BSA/0.1% Tween20 (each 20 min). Then sections were incubated with a secondary antibody, 12 nm gold conjugated goat anti rabbit IgG (Jackson Immuno research (111-205-144)) (2 h, RT, 1:10 in 0.1% BSA/ 0.1% Tween20). Finally, sections were washed twice in 1% BSA/ 0.1% Tween20, twice in PBS, once in 1.25% glutaraldehyde in PBS and three times in ddH_2O (each 15 min).

Staining and contrasting

For staining and contrasting, Epon embedded sections were incubated in 2% aqueous uranyl acetate for 10 min followed by incubation in Reynolds lead citrate for 5 min. LR White embedded sections incubated in 2% aqueous uranyl acetate for 5 min followed by incubation in Reynolds lead citrate for 1.5 min.

Image Acquisition

A 200 kV JEM-2100 (JEOL) transmission electron microscope equipped with a TemCam F416 4k x 4k camera (Tietz Video and Imaging Processing Systems) was used for imaging.

3.11 Immunofluorescence

10 ml trypanosome cell culture at about $5x10^6$ cells/ml were harvested, washed once with SDM-79 w.s.h. and resuspended in 1 ml SDM-79 w.s.h. Cells were fixed in 4% paraformaldehyde with mild shaking for 30 min at RT. Afterwards cells were washed in 12 ml PBS (1400x g, 5 min), resuspended in 1 ml PBS. A PAP pen was used to draw circles on Superfrost® Plus slides (Thermo Fisher Scientific). Cells (about 100 µl cell suspension per circle) were spread on the slides and allowed to settle down for 20 min. This was controlled by phase contrast microscopy. Slides were washed 10 min in 25 mM NH_4Cl in a coplin jar. All further incubation steps were performed in a humid chamber. Samples were blocked and permeabilized with 0.5% saponin/2% BSA (100 µl per circle) for 1 h and washed quickly with PBS

before incubating 1 h with the primary antibody (diluted in 2% BSA in PBS) on the slide (100 µl per circle). Slides were washed four times in PBS for 5 min in a coplin jar, the secondary antibody (diluted in 2% BSA in PBS) was applied to the slide (100 µl per circle) and incubated for 1 h protected from light. After four washing steps in PBS, each 5 min (coplin jar), cells were stained with DAPI (1 µg/ml in PBS) for 1 min and washed once for 5 min in PBS. Cover slips were mounted with Fluor Save (Calbiochem, Merck Millipore) and stored in the dark at 4°C.

3.12 Fluorescence microscopy and image processing

Z-stack images (20-100 stacks, depending on protein expression level and signal intensity with 100 nm distance) were taken with a custom build iMic microscope (FEI-TILL Photonics) equipped with a CCD camera (Sensicam, pixel size 6.45 µm, PCO) using the Live acquisition software (TILL Photonics). eYFP was monitored with the FRET-CFP/YFP-B-000 filter, mCherry, anti-mouse/ anti-rabbit IgG Alexa594, and CAL type 1 Affymetrix probes with the ET-mCherry-Texas-Red filter, GFP and type 4 Affymetrix probes with the ET-GFP filter and DNA and anti-rabbit IgG Sigma CF350 with the DAPI filter (Chroma Technology CORP). A 100x (NA 1.4) oil objective (Olympus) was used for image acquisition.

Images were deconvolved using the Huygens Essential software (Scientific Volume Imaging B. V.) (maximum-likelihood estimation; 20 iterations) and afterwards processed by the ImageJ software (National Institutes of Health). Images are presented as either z-stack projections (method summed slices) or single plane images, depending on the illustrative purpose.

4 Results

4.1 Nuclear periphery granules (NPGs) and their properties

4.1.1 Tools for the analysis of NPGs

In the following, some basic NPG features and different tools for their analysis will be introduced. NPGs form at the nuclear periphery upon treatment with sinefungin (SF), which inhibits the process of trans-splicing[49,50]. During this work different marker proteins, tagged with either eYFP or mChFP, were used to microscopically analyze the formation, location and content of NPGs. Three previously identified NPG proteins served as reference for NPG formation: the DEAD box RNA helicase DHH1, the poly(A) binding protein PABP2 and the Lsm domain protein SCD6 (Figure 4.1 A-C)[49]. To determine the exact location of NPGs relative to the nucleus, nuclear periphery or cytoplasm, the nucleus was either detected by DAPI staining or with the aid of a cell line expressing NUP93-eYFP (p3597). Being part of the inner ring of the nuclear pore complex[66,77], NUP93 defines the border between nucleus and cytoplasm more accurately than DAPI stain. NUP93-eYFP localizes into distinct spots (nuclear pores). NPGs do not colocalize with NUP93 but form a ring around NUP93, indicating cytoplasmic localization of NPGs (Figure 4.1 A). Note that in trypanosomes DAPI also stains the kinetoplast, which contains a network of circular mitochondrial DNA. Depending on the cell cycle phase, one or two kinetoplasts are present in addition to the nucleus.

A helpful tool for the analysis of NPGs is the overexpression of the NPG component SCD6. Tetracycline induction causes a 3-fold overexpression of SCD6 and consequently a larger number of NPGs upon SF treatment in comparison to native expression of SCD6[74] (Figure 4.1 C-D) (plasmid 3924[74]). The SCD6 overexpression cell line was in particular useful when the detection of the protein became limiting, for example in immunocytochemistry. In comparison, the expression of SCD6-eYFP from its endogenous locus at native expression levels resulted in a smaller fraction of SCD6-eYFP localizing in granules (Figure 4.1 C). More details on the overexpression cell line were previously published[74]. Details about all plasmids and cloning methods can be found in chapter 2.5.

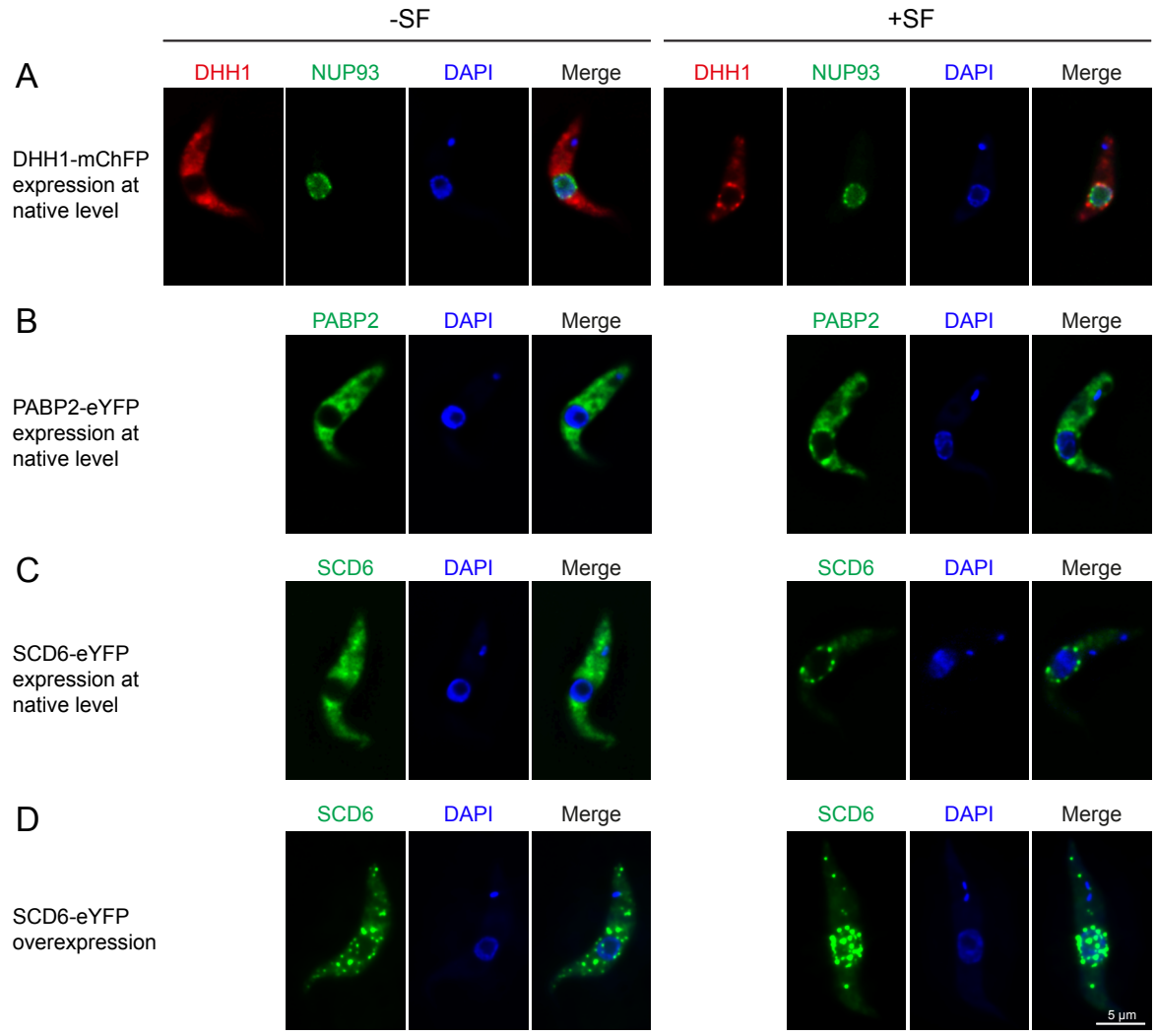


Figure 4.1: NPG localization, marker proteins and tools.

Various NPG marker proteins are displayed in untreated (-SF) and sinefungin-treated (+SF) cells. DNA was stained with DAPI (blue). Images are shown as single plane images of a deconvolved z-stack. **(A)** NPG marker protein mChFP-DHH1 (red) co-expressed with the nuclear pore marker protein NUP93-eYFP (green). **(B)** NPG marker protein PABP2-eYFP at native expression level. **(C)** NPG marker protein SCD6-eYFP at native expression level. **(D)** Overexpression of NPG marker protein SCD6-eYFP.

4.1.2 Different NPG components localize to different NPG subpopulations with different kinetics

As described above, in this work SCD6, DHH1 and PABP2 were employed as NPG marker proteins. Whereas SCD6 and DHH1 localized to NPGs with similar kinetics upon treatment with SF (Fig 1A in [49]) other known NPG components like eIF4E3 appeared to localize to NPGs at much later time-points (Fig S4 in [49]). To better understand the localization of PABP2 and DHH1 to NPGs, they were analyzed over a time-course of SF treatment. A cell line co-expressing both NPG marker proteins, PABP2-eYFP (p3295) and mChFP-DHH1 (p2845), enabled the direct comparison of both proteins in one cell. Figure 4.2 shows fluorescent images of representative cells over a time course of treatment with SF. In untreated cells, both proteins localize to the cytoplasm, with a small fraction of DHH1 (but not PABP2) found in P-bodies. Upon induction with SF, DHH1 is already detectable in NPGs within 10 min. Within 30 min, DHH1 has reached its maximal NPG localization. In contrast, PABP2 is first detectable in NPGs after 40 min and reaches its maximal NPG localization after 60 min. It should thus be considered that individual components move to NPGs with different kinetics.

Based on these findings, the time for SF treatment was set to 50 min for all further experiments. The actual time points for all experiments are 60 min, due to roughly another 10 min elapsing while cells are collected and fixed.

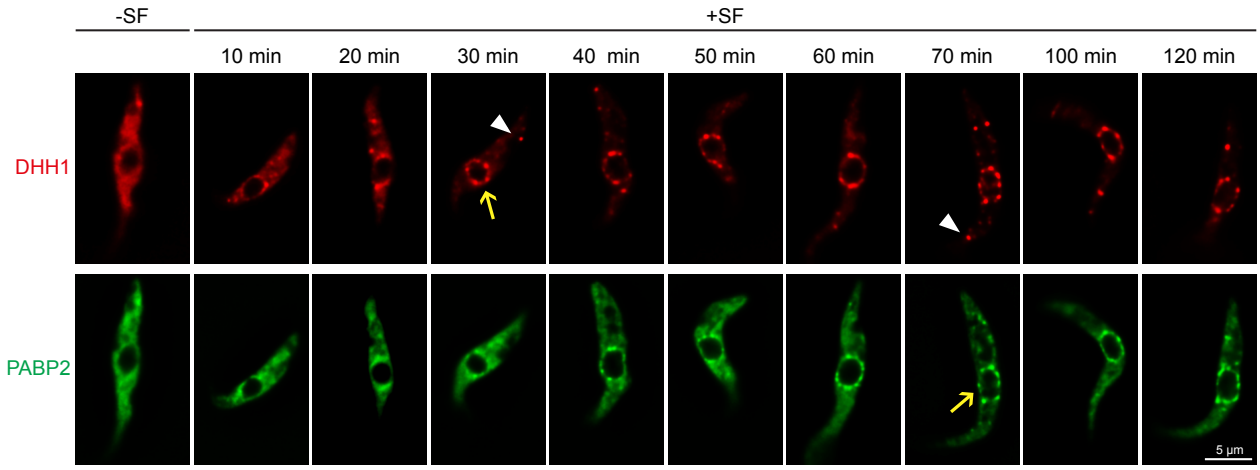


Figure 4.2: Different NPG components localize to NPGs with different kinetics.

Time course of re-localization of two NPG marker proteins upon sinefungin treatment. Two NPG granule marker proteins were co-expressed from the endogenous locus: DHH1-mChFP (red) and PABP2-eYFP (green). Images in both channels were taken as a time course 10-120 min after the initial sinefungin treatment. Examples of P-bodies are marked with a white triangle, examples of NPGs are labelled with yellow arrows. Untreated cells (-SF) served as control. Images are shown as single plane images of a deconvolved z-stack.

4.1.3 NPGs are present in different trypanosome life cycle stages

Previous investigations on NPGs were only conducted in PCF trypanosomes[84]. The question arose if NPGs were unique to this life cycle stage. To address this question mChFP-DHH1 was endogenously expressed in PCF and BSF trypanosomes, using the same construct in both cases (p2845). In both life cycle stages mChFP-DHH1 re-localizes from the cytoplasm to NPGs after SF treatment (Figure 4.3). Hence NPGs are not unique to the PCF life cycle stage of trypanosomes.

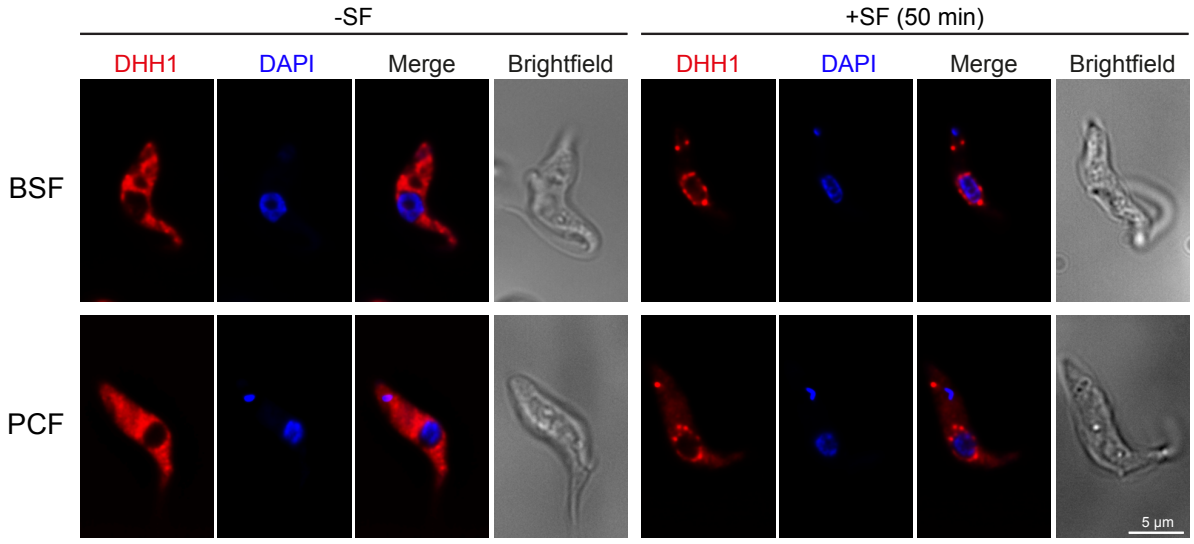


Figure 4.3: NPGs are present in PCF and BSF trypanosomes.

NPGs form in response to sinefungin treatment in BSF as well as in PCF. The NPG granule marker DHH1 was expressed as mChFP fusion protein in BSF and PCF trypanosomes (red). Representative images of untreated (-SF) and sinefungin-treated (+SF) cells are shown. DNA was stained with DAPI (blue). Images are shown as single plane images of a deconvolved z-stack.

4.1.4 NPGs are associated with nuclear pores – an ultrastructural analysis

NPGs in trypanosomes appear similar to perinuclear germ granules (P granules) of adult gonads in *C. elegans*[85,86] due to similarities in localization, sensitivity to actinomycin D, insensitivity to cycloheximide and the presence of the germ granule marker VASA/DBP1[49]. In *C. elegans* different studies also revealed that P granules are associated with clusters of nuclear pores and extend the nuclear pore complex[87,88]. It was shown that perinuclear P granules and their associated nuclear pore complexes are the major sites of mRNA export[33]. Thus, perinuclear P granules might have a function in the regulation or control of nuclear export or might play a role in the determination of mRNA fate. To investigate whether a similar connection between nuclear pores and NPGs also exists in trypanosomes, ultrastructural analyses of NPGs using electron micrographs were conducted.

Pilot experiments - Testing cell lines, sample preparation and antibodies

Trypanosomes expressing PABP2-eYFP-4xTy1 from the endogenous locus (SK148) (details on plasmid, see chapter 2.5 were high pressure frozen to best preserve the fine structure of the cell. In comparison to chemical fixation, rapid freezing by high pressure is faster and thus shows better results in simultaneous stabilization of all cellular components. Especially for studies on membranes (e.g. nuclear envelope), which appear wrinkled after chemical fixation but remain smooth using cryo-fixation, the benefits of this method show. Subsequent embedding at low temperature has furthermore proven to be advantageous for immunocytochemistry studies, since it contributes to preserving and accurately localizing antigens at high resolution).

In a first step, antibodies against all three NPG marker proteins, anti-DHH1, anti-SCD6 and BB2 (to detect PABP2-4xTy1) were tested on ultrathin electron micrograph sections of either wild-type or SK148 PCF trypanosomes. Primary antibodies were detected using a gold conjugated anti-rabbit or anti-mouse antibody, which were validated by a no primary antibody control beforehand (on four conducted images on average only one gold grain was visible; data not shown).

Anti-SCD6 showed the highest amount of labelled protein (highest amount of gold grains per cell). Therefore, this antibody was used for all further immuno-labelling studies. Images of cells stained with anti-DHH1 and anti-Ty1 BB2 (detecting PABP2) are displayed in Sup. Figure 7.1.

For all further experiments either trypanosomes with native SCD6 expression levels (SK148) or trypanosomes carrying a tetracycline-based SCD6-eYFP overexpression system (3925/3924/PSPR2.1)[69,74] were used. Compared to native expression, the overexpression of SCD6 provides a higher amount of detectable antigen per cell for a subsequent immuno-labelling study in general and in particular, after SF treatment, more protein re-localizes to NPGs. The eYFP tag allowed visualization of NPG formation during sample preparation.

Ultrastructural analysis of NPGs

To gain first insight into the ultrastructure of NPGs, untreated and SF treated trypanosomes with native SCD6 expression and overexpressed SCD6 were analyzed. Samples were embedded in LR-white and SCD6 was labelled with immuno-gold as described above. Whereas in untreated cells (left panel Figure 4.4) the majority of SCD6 is localized within the nucleus, the cytoplasm and large aggregates herein (most likely P-bodies) in SF treated cells SCD6 is present at distinct spots at the nuclear periphery (right panel, white circles Figure 4.4). The granule arrangement around the nucleus and their cytoplasmic position resembles fluorescence images of NPGs. Furthermore, in cells with induced SCD6 overexpression, the protein seems to be so densely packed in NPGs that granules itself are visible as

electron dense structures; this was not the case in cells with native SCD6 expression. Moreover, due to the smaller amount of expressed protein, the number of gold grains labelling the NPGs is lower.

With this experiment the cytoplasmic localization of NPGs to the nuclear periphery was confirmed on an ultrastructure level. Furthermore, it could be demonstrated that with a high amount of SCD6 present in the cell, NPGs are visible as electron dense structures. However, the method of LR-white embedding used here, did not allow to visualize nuclear pores and thus to address the question whether NPGs interact with pores.

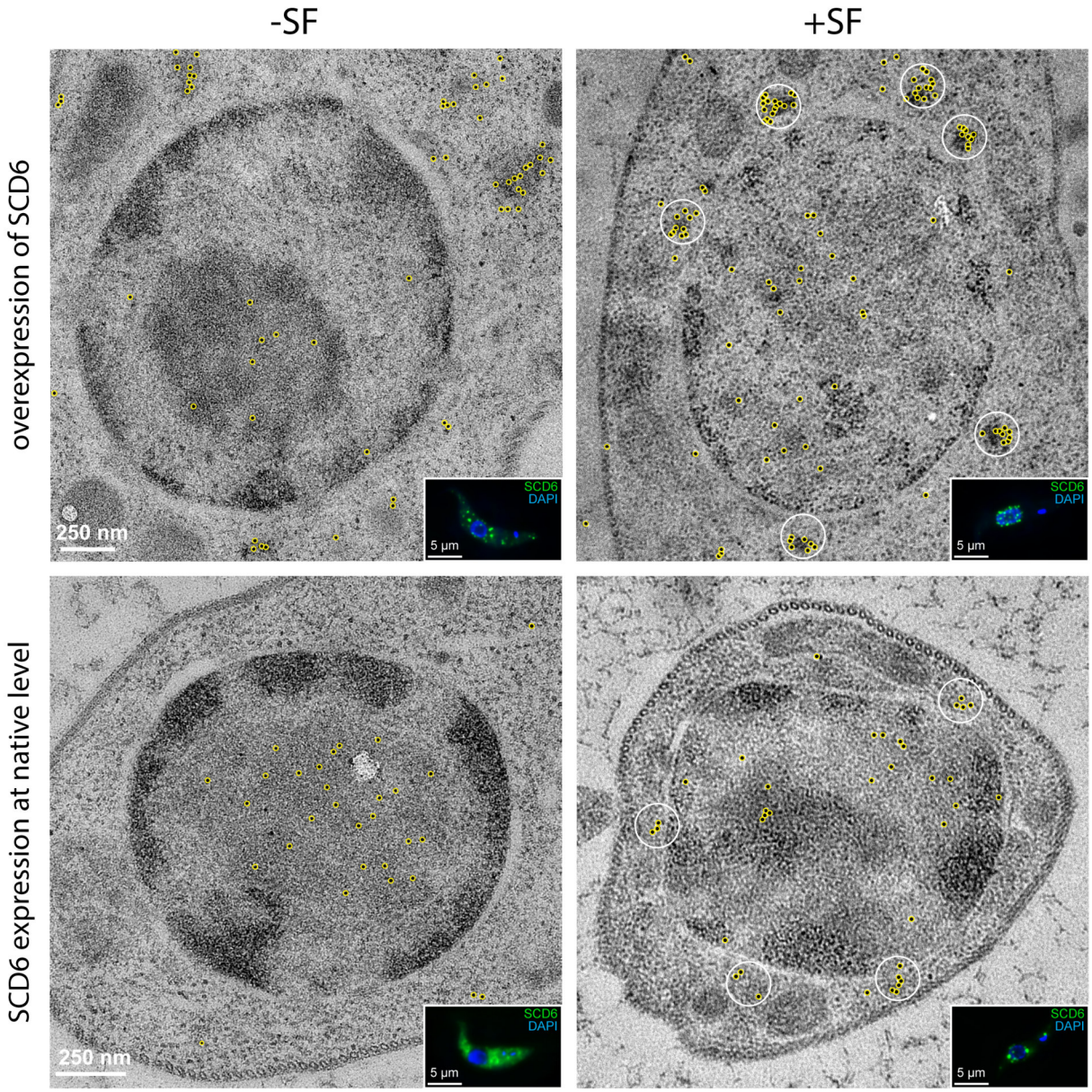


Figure 4.4: Ultrastructural analysis of NPGs.

Immuno-gold labelling of the NPG marker protein SCD6 on EM sections. Representative images of a cell line with inducible overexpression of the NPG granule marker SCD6 **(A)** and a cell line with native SCD6 expression **(B)** are shown for one untreated (-SF) and one sinefungin-treated (+SF) cell respectively. Cells were high pressure frozen, embedded in LR-White,

cut into 60 nm thin sections and contrasted with aqueous uranyl acetate and Reynolds lead citrate. SCD6 was labelled with anti-SCD6 and 12 nm conjugated gold anti-rabbit. For better visualization gold grains are marked with yellow circles, NPGs are circled in white. A fluorescence image of one representative cell expressing SCD6-eYFP (green) is displayed as deconvolved z-stack projection. DNA was stained with DAPI (blue).

Connecting NPGs and nuclear pores (on an ultrastructural level)

To investigate whether there is a direct connection between NPGs and nuclear pores, co-labelling of both structures was aimed. First, the generation of a cell line endogenously expressing NUP93-eYFP-4xTy1 (SK362) (for details on plasmid, see chapter 2.5) enabled the detection of NUP93 with an anti-Ty1 BB2 antibody. However, on ultrathin sections, nuclear pores could not be sufficiently labelled, probably due to the small amount of protein (antigen) present per pore (data not shown). Next, tests with a commercial anti-nuclear pore complex antibody MAB414 (Abcam, Cambridge), which specifically labels a conserved family of nuclear pore complex proteins in different organisms (e.g. human, mouse, yeast, *C. elegans,* drosophila) were performed on wild-type PCF trypanosomes. However, the antibody showed unspecific labelling throughout the whole trypanosome cell (data not shown). Note that at a later time point a study was published, in which the MAB414 antibody indeed was successfully used to label nuclear pores in PCF *T. brucei* cells[89]. Still, since co-labelling of NPGs and nuclear pores could not be accomplished by antibody staining at the time, the strategy was changed at this point.

If the nuclear pore structure is preserved, pores can be identified through two indicators: (1) The nuclear envelope, consisting of an inner and outer membrane is interrupted at the position of a nuclear pore. (2) Inside the nucleus, the densely packed heterochromatin shows a gap, forming a passage towards the pores. Taking advantage of these facts, a new strategy for detecting nuclear pores arose. While the so far used LR-white with its hydrophilic nature has proven to be the best embedding agent for subsequent immunocytochemistry studies, Epon is the better choice for morphological studies since it preserves membranes (e.g. nuclear envelope) and other cellular structures much better. Hence, the ultrastructural analysis of NPGs and nuclear pores was repeated with Epon embedded cells, with the aim to preserve the structure of nuclear pores. Trypanosomes carrying the tetracycline-based SCD6-eYFP overexpression system (3925/3924/PSPR2.1)[74] were used, because of the previously found visibility of NPGs as electron dense structures in this system. Figure 4.5 A shows cross sections of nuclei from untreated and SF treated cells. Nuclear pores are clearly visible in both samples (white arrowheads). In SF treated cells NPGs formed around the nucleus. Here some NPGs (labelled) are clearly connected to nuclear pores, whereas others show no association with pores. The estimated diameter of an NPG (about 200 nm) exceeds the diameter of one nuclear pore complex (about 100 nm)[75]. Thus, for some NPGs the attached nuclear pore may simply be not present in the same

section of the cryo-cut material. Figure 4.5 B shows details from cells imaged at higher resolution for better visualization of the pores and the connection to NPGs. More examples of NPG/nuclear pore connections are displayed in Figure 4.6. Most interestingly, in some images a strand of electron dense material is pervading from inside of the nucleus through the pore into an NPG (indicated with *).

Comparing all available images (12 for untreated and 28 for SF treated cells), on average, one granule is located close to the nuclear periphery per image of untreated cells, whereas in SF treated cells five granules are present. Also, in untreated cells, none of the visible granules (n= 12) could be correlated with a nuclear pore, whereas in SF treated cells about 38% of NPGs (n= 145) were connected to nuclear pores (on average 2 granules per cross section). Interestingly, in SF treated cells in 46% of the cases a visible nuclear pore (n= 97) was not associated with an NPG.

This experiment finally proved the existence of a direct connection between NPGs and nuclear pores in trypanosomes.

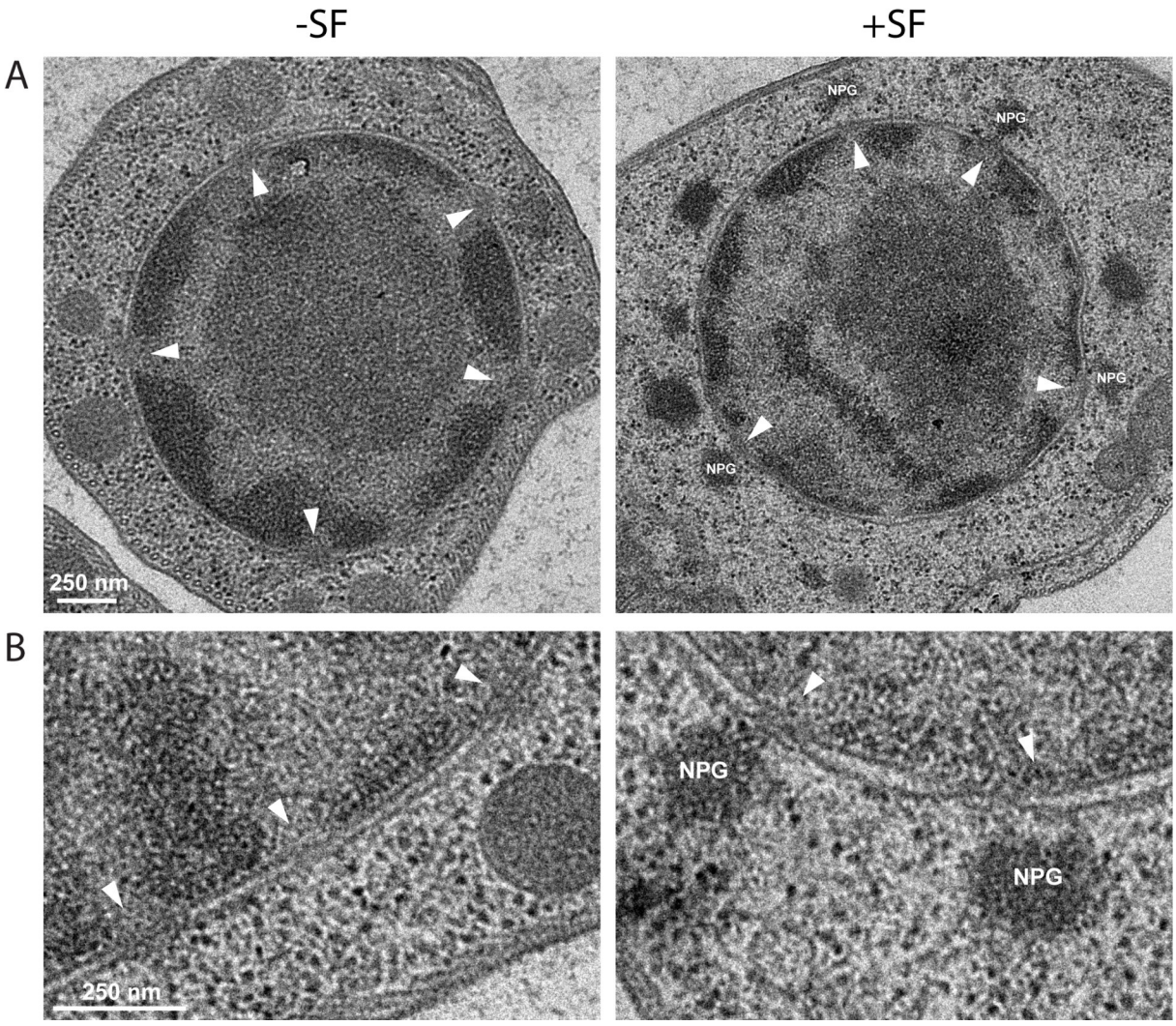


Figure 4.5: NPGs are associated with nuclear pores.

Specific correlation of NPGs with nuclear pores. **(A)** Representative images of a cell line with induced overexpression of SCD6 are displayed for untreated (-SF) and sinefungin treated (+SF) cells. **(B)** shows more detailed images acquired with higher magnification. Cells were high pressure frozen, embedded in Epon, cut into 60 nm thin sections and contrasted with aqueous uranyl acetate and Reynolds lead citrate. Arrowheads point to nuclear pores, NPGs are labelled in white.

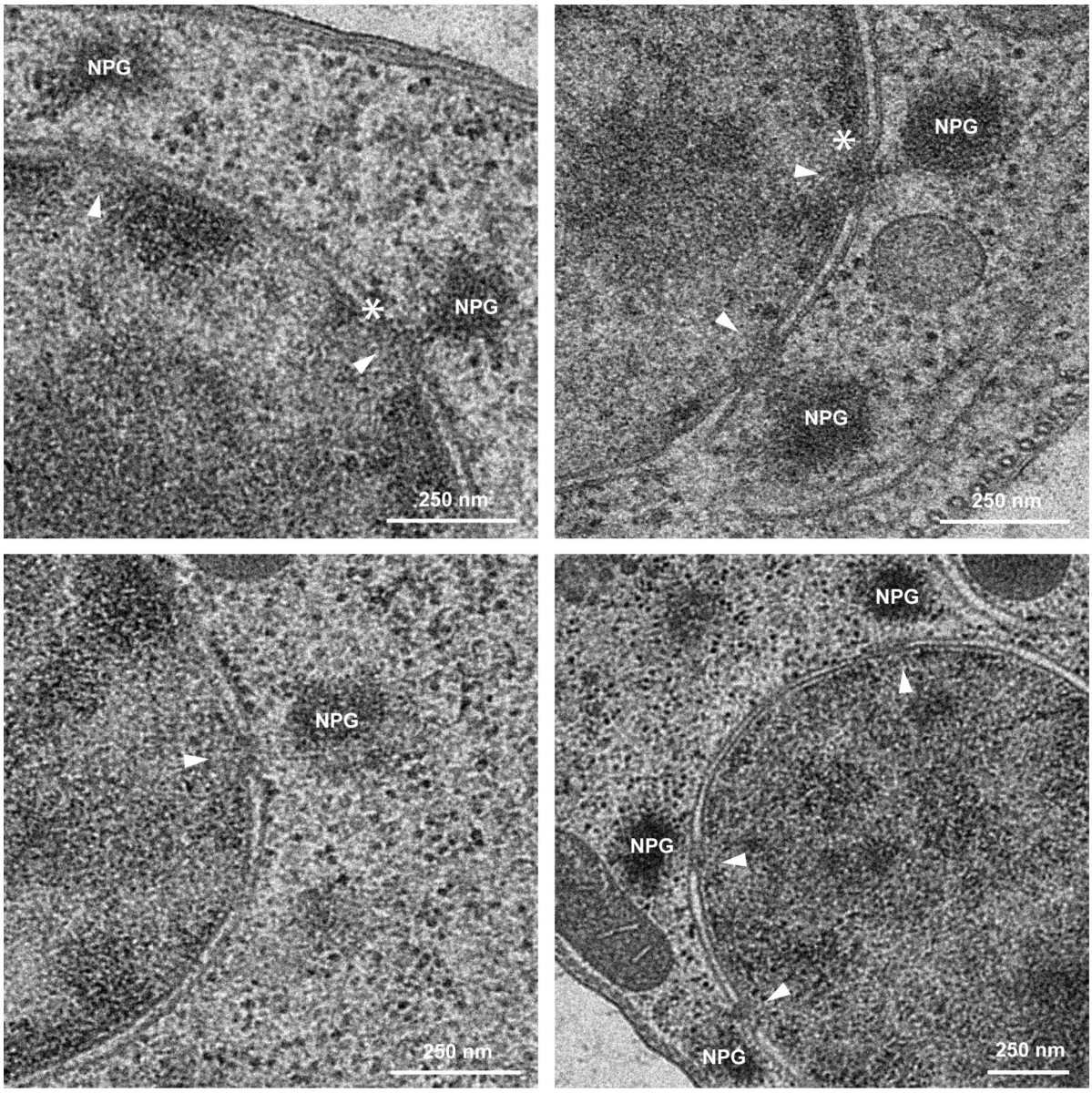


Figure 4.6: NPGs are associated with nuclear pores.

Additional examples for correlation of NPGs with nuclear pores. Images of sinefungin treated cells with induced overexpression of SCD6 are shown (compare Figure 4.5). Arrowheads point to nuclear pores, NPGs are labelled in white and stars indicate spots where it seems as if a protein strand is pervading through a pore.

4.1.5 Conclusion

In summary, the fluorescence and electron microscopy experiments shown above revealed several novel features of NPGs: NPGs are cytoplasmic and present in different trypanosome life cycle stages (PCF and BSF). The granule population is not homogenous but composed of distinct granule subsets, with differences in protein composition. Most interestingly, granules are clearly connected to nuclear pores and it is tempting to speculate about an involvement in nuclear export.

4.2 The NPG proteome

4.2.1 Ideas, modifications and trial attempts

Initial idea

Trypanosome NPGs form at the nuclear periphery upon treatment with sinefungin, which causes the inhibition of trans-splicing. My electron microscopy data (chapter 4.1.4) now clearly show that NPGs are associated with nuclear pores at the cytoplasmic site of the nucleus. This might indicate a physical connection between NPGs and the nucleus or the nuclear membrane, which could be exploited for the development of a granule purification protocol. In 2001 Rout and Field[75] published a protocol for isolation of subnuclear compartments, where in a first step trypanosome nuclei were purified and in a second step the nuclear membranes were isolated from these nuclei. I set out to reproduce this protocol using PCF trypanosomes with the attempt to modify it for the purification of NPGs. The aim was to determine the proteome of NPGs by mass spectrometry analysis.

Purification of trypanosome nuclei: establishment of the method

While establishing the purification of trypanosome nuclei as a first step, different problems appeared. These were solved step by step and resulted in a protocol for nuclei purification that slightly differed to the previously published method[75,76] (described in chapter 4.2.2). The major problems that were solved during the establishment of the method were:

- Cultivation of cells in conical flasks:

Due to the high number of cells needed per purification, growing cells in cell culture flasks was inefficient. Thus, cultivation of trypanosomes was performed in sterile conical glass flasks (5 l volume). These flasks could be filled with a maximum of 2 l and cells grew when the flask was gently shaking. However, different test runs showed that it was essential to inoculate the 2 l culture with cells from a culture grown to a very high cell density (at least $2x10^7$ cells/ml), 24 h before starting the purification. Otherwise cells were severely impaired in growth (PDT ≈24 h instead of 8 h).

- Finding the right number of cells for the purification:

The original protocol suggested $2x10^{10}$ cells for each gradient and noted not to exceed this cell number, because the purity of the nuclei would drop with increasing cell numbers. Tests showed that a lower number of cells separated even better on the gradient and resulted in purer fractions. Hence, $1x10^{10}$ cells were proceeded per gradient.

- Efficient cell lysis:

Since the original protocol provided no exact time and intensity for a mechanical cell lysis with the POLYTRON® homogenizer, this had to be tested, too. Usually a complete cell lysis was achieved after 5 min at $^2/_3$ of the maximum POLYTRON® speed. Still, this had to be monitored microscopically for every experiment. Complete cell lysis and a homogenous sample turned out to be essential for the efficiency of the subsequent separation on a sucrose gradient.

- Recovery of nuclei from sucrose gradient:

In contrast to the previously published protocol, that collected nuclei from an interface section of the sucrose gradient (between 2.1 M and 2.3 M sucrose), in my hand nuclei were best recovered from a ring-shaped pellet at the bottom of the gradient. A recent purification of nuclei from the related parasite *Trypanosoma cruzi* that used the same method also reported such a pelleting of the nuclei instead of interface localization[90]. For the stepwise collection of all fractions, in the end a truncated blue pipette tip proved most suitable.

NPGs do not survive further fractionation of the nuclei into nuclear envelopes

After the successful purification of nuclei (with NPGs still attached after sinefungin treatment, more details see below 3.2.4), it was planned to separate nuclear membranes with attached NPGs from the rest of the nuclei. Buffers, materials and devices were applied as described in DeGrasse et al. 2009[77]. The use of a cell line expressing the nuclear pore protein NUP93-eYFP (3597) and the NPG marker mChFP-DHH1 (2845) enabled the detection of both, the nuclear envelope and NPGs throughout the process. The separation of nuclear membranes from purified nuclei of cells treated with SF was done exactly as described in DeGrasse et al. 2009[77] (for details see chapter 3.8). Nuclear envelopes were visible in brightfield and fluorescence microscopy, indicating a successful purification (data not shown). However, NPGs were not detectable. Further tests revealed that NPGs resolved as soon as shearing buffer was added to the nuclei. A closer look at the buffer components (DTT, heparin, DNaseI, RNaseA, protease inhibitor) indicated that DTT and protease inhibitors are unlikely to have caused granule dissociation, since both were also components of the lysis buffer used during the first purification step. Heparin is often used as an affinity ligand and captures positively charged structures. Its role in dissolving NPGs remains unclear. During tests with different buffer compositions, the strongest dissociation of NPGs was seen when the buffers contained DNaseI or RNaseA. Both enzymes were, however, essential components for the successful shearing of nuclear envelopes, and could not be omitted from the buffer. Thus, nuclear envelopes with intact NPGs could not be isolated from trypanosome nuclei. The fact that DNase and RNase cause granule dissociation is interesting by itself and indicates that granules may require both intact DNA and intact RNA.

Salt treatment of purified nuclei was not advantageous

Since a further fractionation of nuclei could not be performed without the loss of NPGs, enhancement of the purity of the original, nuclei containing pellet fraction was aimed. Classical purification approaches often use salts and detergents to dissolve cellular structures. Since it was previously found that NPGs are sensitive to detergent[49], the testing of detergents in this case seemed unreasonable. Neither sodium chloride (NaCl) nor potassium chloride (KCl) nor calcium chloride ($CaCl_2$) at concentrations ranging from 150 mM to 300 mM dissolved flagellar or other cell remnants, instead NPGs disappeared, and nuclei seemed to burst (data not shown). Therefore, the protocol for the purification of nuclei was used as described above, without any further purification steps.

4.2.2 Purification of *T. brucei* nuclei

The following improved method (Figure 4.7 A) is a result of incorporating the above described modifications to the previously published protocol[75,76]. The protocol was designed to purify PCF nuclei under conditions that maintain the attachment of NPG to the nuclei and allow subsequent mass spectrometry analysis.

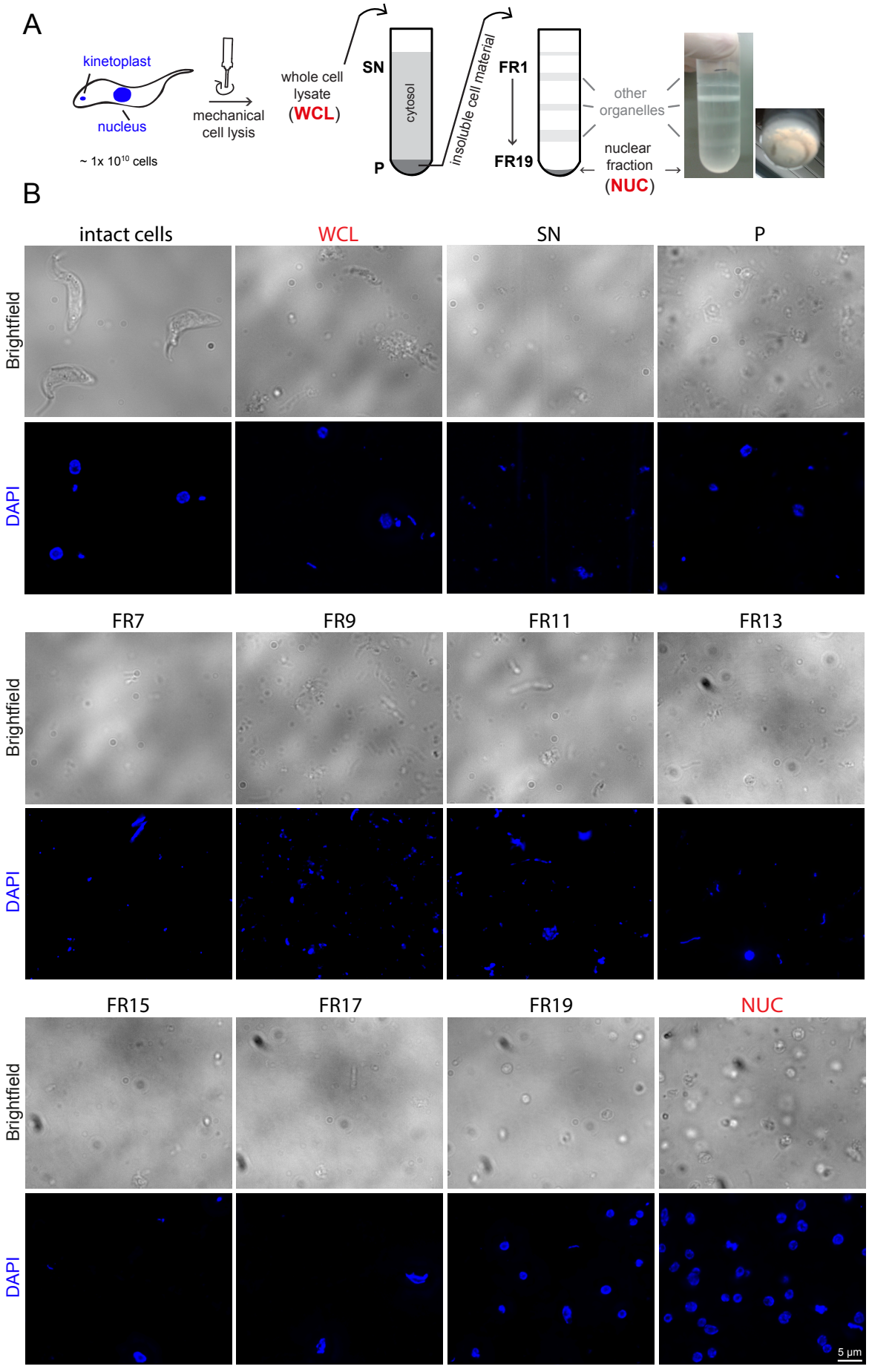
Purification protocol

For each purification approximately 10^{10} PCF cells at about $6x10^6$ cells/ml were grown in 2 l SDM79 medium. Cells were either treated with SF or left untreated. NPG formation was monitored microscopically. Cells were pelleted stepwise and washed with SDM79 without serum and heme. From this point on, all further steps were conducted in a cold room. Cells were resuspended on ice in lysis buffer and disrupted by a POLYTRON® homogenizer ("whole cell lysate (WCL)" Figure 4.7 A and B). Cell lysis was monitored by brightfield and fluorescence microscopy. Subsequently, the whole cell lysate was underlaid with underlay buffer and centrifuged: the supernatant contained mainly soluble material from the cytosol ("supernatant (SN)" Figure 4.7 A and B), the pellet contained mostly insoluble cell material. The pellet was immediately resuspended in resuspension buffer followed by further homogenization with the POLYTRON® ("pellet (P)" Figure 4.7 A and B). After complete cell lysis (monitored by brightfield and fluorescence microscopy), the material was loaded on a three-step sucrose gradient and after ultracentrifugation the gradient was harvested from the top (with truncated blue pipette tips) in 2 ml fractions ("FR1-19" Figure 4.7 A and B). The ring-shaped pellet at the bottom of the tube was resuspended in 2 ml 2.3 M sucrose/PVP ("nuclei (NUC)" Figure 4.7 A and B). From every step during the purification process, samples were collected and stored at -80°C for subsequent analysis. A photo of the gradient after ultracentrifugation, with different milky layers and the ring-

shaped pellet shows the successful separation of the cell lysate (Figure 4.7 A) (for details of the protocol see chapter 3.8).

Microscopic monitoring of the procedure

Each sample collected during the procedure was stained with DAPI to detect nuclei and kinetoplasts and analyzed by brightfield and fluorescence microscopy. Figure 4.7 B shows representative images of fractions WCL, SN, P, FR7-FR19 (every second fraction) and NUC. As a control, an image of intact trypanosome cells is shown. After the first round of mechanical cell lysis almost all cells were disrupted. The sample (WCL) consisted of cell fragments of various sizes and nuclei were mostly still connected to other cell fragments. After the sucrose cushion fractionation, most crude and insoluble cell material (nuclei and kinetoplasts visible with DAPI staining; flagella and cytoskeleton visible in brightfield was present in the pellet fraction (P), while the supernatant fraction (SN) contained mostly soluble cytosolic material (only few undefined structures visible with DAPI staining and in brightfield) (Figure 4.7 B, upper panel). After the second round of mechanical cell lysis, nuclei of fraction P were separated from the remaining cell material by further separation on the sucrose gradient. Fractions FR1 to FR7 of the sucrose gradient showed no obvious difference in composition. Since only few structures were visible (Figure 4.7 B, FR7), they presumably contained mainly soluble material or very small particles. Further down the gradient every fraction appeared different. FR7 or FR17 for example, which were collected within one sucrose molarity, contained only few structures compared to FR9 or FR15, which were collected directly at the interface between two different sucrose molarities. In general (1) during the course of the gradient, flagella, kinetoplasts, cytoskeletal elements, free DNA and remnants of ruptured nuclei were captured; (2) the higher the sucrose molarity and the further down the gradient, the more nuclei were present in the fraction; (3) finally, fraction NUC (resuspended ring-shaped pellet) contained the highest concentration of nuclei and the least visible contaminants such as flagella or kinetoplasts. Hence, all subsequent nuclear analyses were conducted with the NUC fraction. Notably this differed from the previously published protocol[75,76]. Here, nuclei were collected directly from the interface between 2.1 M and 2.3 M sucrose and no ring-shaped pellet was reported; the reason for these differences remains unknown. To prove the purity and quality of the procedure, fractions were not only microscopically analyzed, but also by SDS-Page (Coomassie stain), western blot and mass spectrometry (see chapters 4.2.3 and 4.2.4).



(caption on next page)

Figure 4.7: Establishing a protocol for purification of NPGs.

(A) Schematics of the procedure. About 10^{10} trypanosome cells were mechanically lysed with a POLYTRON mixer (whole cell lysate, WCL). Afterwards the cell lysate was separated with a sucrose cushion into an insoluble cell fraction (pellet, P) and a soluble fraction (supernatant, SN). The insoluble pellet fraction was resuspended and further separated on a three-step discontinuous sucrose gradient by ultracentrifugation. The gradient was then collected stepwise (FR1-FR19). At the interfaces of the sucrose layers various organelles and cell fragments accumulated and became visible as milky white interphase layers. At the bottom of the gradient, a ring-shaped pellet fraction formed, which was highly enriched in nuclei (NUC). A representative picture of a sucrose gradient after ultracentrifugation is shown on the right. **(B)** Composition and content of fractions of each purification step. Samples of each fraction were stained with DAPI and analyzed microscopically. Fractions FR1 to FR7 showed no difference in composition of organelles and cell fragments. Fractions FR7 to FR19 are represented by images of every second fraction. After the whole cell lysate (WCL) was separated, the supernatant (SN) contained mostly soluble material from the cytosol, whereas the pellet (P) contained crude cell material like flagella, kinetoplasts and nuclei. This crude cell material was then further separated, leaving flagella and kinetoplasts at upper fractions (FR1- FR17) and showing an enrichment of nuclei in lower fractions (FR18-FR19 +NUC). In the NUC fraction, isolated nuclei are highly enriched and clearly visible as ovoids together with only few other structures, such as remnants of flagella (brightfield image). Nuclei are intact (native shape, nucleolus is visible by absence of DAPI staining) and only few kinetoplasts are visible (DAPI image). Samples were not fixed to the slide and moved during imaging, thus the two channels do not completely overlap. The DAPI image is shown as deconvolved z-stack projection, the brightfield image is a single plane image.

4.2.3 The nuclear proteome of *T. brucei* (Goos et al. 2017)[91] – A tool for monitoring the quality and purity of the purification process

During the development of the purification procedure, Prof. Christian Janzen became aware of this work and suggested to collaborate on a long-intended project: obtaining the complete nuclear proteome of *T. brucei*. We decided to pursue this idea by using mass spectrometry analysis of purified nuclei and, for comparison, of whole cell lysates. With this, two aims could be achieved at the same time: 1) Obtain the yet unknown *T. brucei* nuclear proteome 2) Confirm purity and quality of the purification procedure, which is essential for the final aim to determine the NPG proteome from the purified nuclei. Our nuclear proteome of PCF trypanosomes was published in 2017[91]. With this, an important gap in the knowledge of trypanosome biology, in particular nuclear-related processes, was closed. In the following paragraphs, results from the publication are explained and summarized.

4.2.3.1 Fraction NUC is highly enriched in nuclei with only little contaminants

Samples of fractions NUC and WCL of wild-type trypanosomes were stained with DAPI and analyzed microscopically. As expected, whole cell lysates contained whole cell remnants with both nuclei and kinetoplasts (Figure 4.8 A, left panel), whereas the nuclear fraction only contained a large

number of nuclei and little visible contaminants (Figure 4.8 A, right panel). Protein samples of the two fractions were also analyzed on a Coomassie-stained SDS gel. Histones were highly enriched in the nuclear fraction in comparison to whole cell lysates, while the amount of total proteins decreased. Additionally, histone H3 was detected by western blot with an anti-H3 antibody (Figure 4.8 B). This way, the successful enrichment of nuclei was verified. Furthermore, all collected fractions were tested for two non-nuclear organelle marker proteins. The ER protein TbBiP[92] and the paraflagellar rod protein PFR[93] were detected by western blot (Figure 4.8 C). Only selected fractions are shown. PFR as part of the flagellum was mainly present in the pellet P (containing crude cell material), and in the top fraction of the gradient (not shown here). Consistent with the light microscopic analysis that showed the absence of flagellar remnants, fraction NUC contained only minor amounts of PFR, if any. BiP was present throughout all samples in the top fractions of the gradient, with a minor reduction in the bottom fractions. The residual BiP in fraction NUC is very likely caused by perinuclear ER that cannot be fully separated from the nucleus by the mechanical lysis. BiP is present in the perinuclear space between inner and outer nuclear envelope membrane, witch directly merges into the ER[94]. The same phenomenon was previously reported by Field and Rout 2001[75]. With the high purity of nuclei in fraction NUC being confirmed, samples were prepared for mass spectrometry analysis and sent to the laboratory of Falk Butter, Mainz, Germany.

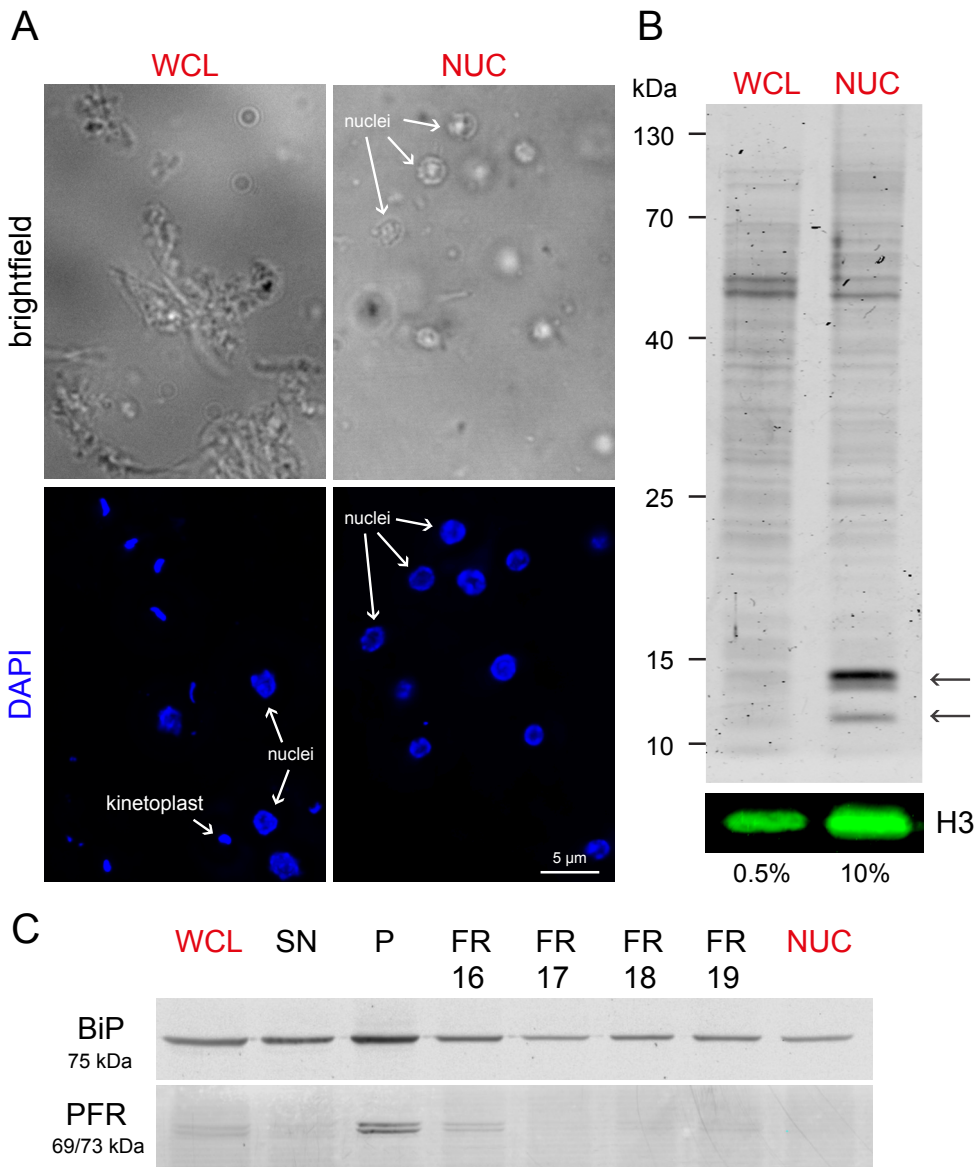


Figure 4.8: Purification of *T. brucei* nuclei/ Purity and quality of purification.

(A) Direct comparison of WCL and NUC fraction. Samples of WCL and the NUC fraction were stained with DAPI and analyzed microscopically. The WCL sample contains remnants of whole cells, including both nuclei and kinetoplasts. After purification, the NUC fraction contains highly enriched, intact nuclei with only few cellular remnants. Samples were not fixed to the slide and moved during imaging, thus the two channels do not completely overlap. The DAPI image is shown as deconvolved z-stack projection, the brightfield image is a single plane image. **(B)** Enrichment of histones in NUC fraction. A representative Coomassie-stained gel loaded with 0.5% of the WCL fraction and 10% of the NUC fraction shows the enrichment of histones in the NUC fraction (arrows point to the bands corresponding to histones). In addition, histone H3 was detected by western blot (lower panel). **(C)** Reduction of non-nuclear proteins. The reduction in non-nuclear proteins during the purification was controlled by western blot detecting the ER marker BiP and the flagellar component PFR. In the NUC fraction no PFR and only a small amount of BiP was detected. In contrast, the pellet fraction P contained high amounts of both proteins. Since the outer nuclear membrane is part of the ER, BiP in this case likely corresponds to the membrane of the enriched nuclei. The results of the western blot correspond to the microscopic analysis in Figure 4.7.

4.2.3.2 Obtaining the nuclear proteome: a list of 764 candidates

Protein samples of four biological replicates of fractions WCL and NUC were analyzed by label free quantitative (LFQ) mass spectrometry with the help of Falk Butter and Mario Dejung (IMB, Mainz, Germany). 3447 protein groups were detected in at least two of the samples, which correspond to more than 1/3 of all proteins encoded by the *T. brucei* genome[12] Sup. Table 7.3 A. Nuclear enrichment scores (NES) for each protein group were calculated as the ratio of LFQ intensities of fraction NUC divided by the LFQ intensities of fraction WCL. NES (transformed by log2) ranged from +7.7 to -9.2. A NES of 0.7 was chosen as threshold because it corresponded to a local minimum in the NES histogram[91]. Proteins with an NES below 0.7 or a p-value above 0.05 were removed, resulting in a list of 760 protein groups with potential nuclear localization. With these very stringent cut off criteria, even some histones were excluded, because the high abundance of a protein reduced the difference between the NUC LFQ score and the WCL LFQ score. Therefore, all proteins with an NES above 0.7, that were also amongst the 20% of most abundant proteins, were added to the list, independent of their p-value. As a result, four more proteins extended the list: additional to Tb927.7.4180 and Tb927.11.2510, two of the histones were included. Thus, the final list of nuclear protein candidates contained 764 protein groups (Sup. Table 7.3 B).

4.2.3.3 Validation of nuclear proteome (done in collaboration with Susanne Kramer)

Less than 2% false positives

With the aim to evaluate the previously chosen thresholds and to produce a reliable list of nuclear proteins, with only few false positives, the number of non-nuclear proteins within the nuclear proteome was estimated. For this, the proteome was compared with six already experimentally characterized, non-nuclear structures/pathways: the lipid metabolism pathway[95], the flagellome[96], the mitochondrial proteome[97] proteins that associate with the cilium transition zone[98], the glycosome[99] and the cell surface[100] (Figure 4.9 A).

96 proteins are known to be involved in *T. brucei* lipid metabolism[95]. This list of proteins was prepared by homologies to yeast enzymes and/or experimental characterizations. 7 of 96 proteins were present in our nuclear proteome (Sup. Table 7.4 A). This included three fatty acid elongases (ELO1-3) that were found during an initial quality control by Gene Ontology (GO) enrichment analysis (data not shown see Goos et al. 2017[91]). In yeast, fatty acid elongases are known to localize to the perinuclear region of the ER membrane[101]. Evidence exists that this localization might be conserved

for *T. brucei* ELOs[102]. Many of the 96 lipid metabolism proteins that were absent from the nuclear proteome also localize to the ER, for example enzymes involved in glycosylphosphatidylinositol (GPI) biosynthesis. Apparently, the contamination with the ELO proteins was not due to a general contamination with ER proteins but caused by a co-purification of the nucleus-adjacent ER membrane (previously explained for BiP remnants in NUC fraction).

331 proteins were identified by mass spectrometry in purified flagella of *T. brucei*[96]. Of these, 16 were present within the nuclear proteome (Sup. Table 7.4 B). However, the number of true potential flagellar proteins was reduced from 16 to three because ten of the proteins had known nuclear localization, demonstrated by the expression of GFP-fusion proteins[77,103] or by specific antibody staining[104] and another three proteins were clear homologues to proteins with nuclear localization in other organisms (GLE2, Kre33 and ERB1). The remaining three proteins with a potential flagellar localization were Tb927.5.940, Tb927.8.2290 and Tb927.3.5010.

From mitochondria enriched fractions, about 1000 proteins were determined within the mitochondrial proteome of *T. brucei*[97]. For the comparison with the nuclear proteome, only the 401 proteins that were assigned to mitochondria with high confidence in Panigrahi et al. 2009[97] were used. Only four of these proteins were present in the nuclear proteome (Sup. Table 7.4 C). They are likely true mitochondrial proteins as they are described by mitochondrial GO-terms. Additionally, two of them were experimentally characterized, one is an RNA editing component and one is found in the small subunit of the mitochondrial ribosome[105,106].

68 proteins were recently characterized as the proteome of the cilium transition zone[98]. They were identified as non-nuclear structures by successful localization of eYFP fusion proteins. They localized to the cilium transition zone, the basal body, the pro-basal body, the flagellar pocket collar, the Inv-like compartment (a region distally adjacent to the transition zone), a longitudinal structure close to where the flagellum emerges from the flagellar pocket, the flagellum, the Golgi and combinations of these localizations. None of the 68 proteins were present in the nuclear proteome (Sup. Table 7.4 D).

129 proteins were identified as the trypanosome glycosome proteome by a combination of glycosome purification (via an epitope tag) and SILAC labelling[99]. Based on this data set, the nuclear proteome is contaminated with up to three glycosomal proteins (Tb927.5.2590, Tb927.8.920, Tb927.9.15260) (Sup. Table 7.4 E).

295 proteins, corresponding to 198 unique protein groups, were assigned as cell surface proteome of PCF trypanosomes by mass spectrometry analysis of biotinylated surface proteins[100]. Nine of 295 proteins are present in our nuclear proteome (Sup. Table 7.4 F). Of these, six have strong experimental

evidence for nuclear localization[107–109]. The remaining three proteins (Tb927.1.120, Tb927.2.1330, Tb927.2.470) are likely to be non-nuclear.

In summary, out of 1162 unique proteins from the six control groups of proteins with likely non-nuclear localization only 20 proteins were present in the nuclear proteome, corresponding to 1.7%. Thus, the nuclear proteome is likely to contain less than 2% false positives.

About 80% true positives

For a statement on the completeness of the nuclear proteome, it was compared with four experimentally characterized nuclear structures: the nuclear pores[66,77], the exosome[60,110,111], kinetochores[108,112] and the spliceosome[113] (Figure 4.9 B).

In two studies, 27 structural components of the nuclear pore complex were identified in *T. brucei*[66,77]. This was confirmed by localization of GFP fusions to distinct punctate structures at the nuclear rim. For two proteins the tagging attempt had failed (TbNup59, TbNup62). It turned out that the nuclear proteome contains 25 of the 27 proteins; only TbNup75 and TbNup65 are missing (Sup. Table 7.5 A).

So far, 11 proteins are identified for the *T. brucei* exosome[60,110,111]. Whether exosome localization is entirely or only partially nuclear has been debated in the past. Recently it was suggested that the majority or all of the exosome is nuclear, because all reported functions of the *T. brucei* exosome are nuclear and an eYFP-Rrp6 fusion (an essential exosome component) localized mainly to the nucleus[61,110,114,115]. Ten of the 11 exosome proteins were present in our nuclear proteome; only RRP41B was absent due to a p-value slightly higher than 0.05 (Sup. Table 7.5 B).

20 kinetoplast kinetochore proteins (KKT1-20), seven kinetoplast kinetochore interacting proteins (KKIP1-7)) and seven further nuclear proteins were identified as components of the trypanosome kinetochore[108,112]. Their nuclear localization was confirmed by eYFP tagging. The nuclear proteome contained 27 of these 34 proteins. KKTs 1, 5, 10, 15 and 16 and KKIPs 5 and 7 were absent (Sup. Table 7.5 C).

As one of the best-characterized trypanosome structures, the spliceosome with 59 known proteins was identified by a combination of bioinformatics and tandem tag affinity purification with four different bait proteins[113] (and references within). 44 of the known 59 spliceosomal proteins were present in the nuclear proteome (Sup. Table 7.5 D). Most of the 15 missing spliceosomal proteins were small Lsm and Sm proteins.

In summary, 106 of 131 control proteins with known nuclear localization are present in the nuclear proteome, corresponding to 80.9%. Thus, the completeness of the nuclear proteome was estimated to about 80%.

Size exclusion

With an average molecular weight of 37.4 kDa the 25 nuclear proteins of the known nuclear complexes that were absent from the nuclear proteome were relatively small. In comparison, the average molecular weight of proteins in the nuclear proteome data set was 66 kDa. The result of an unpaired, two-tailed students t-TEST ($p=0.01$) proved this difference to be significant (Figure 4.9 C). This could be caused by a preferential loss of smaller proteins during the purification procedure. Smaller proteins are more likely to leak out of the nuclei and thus result in fewer unique peptides detectable by the mass spectrometer.

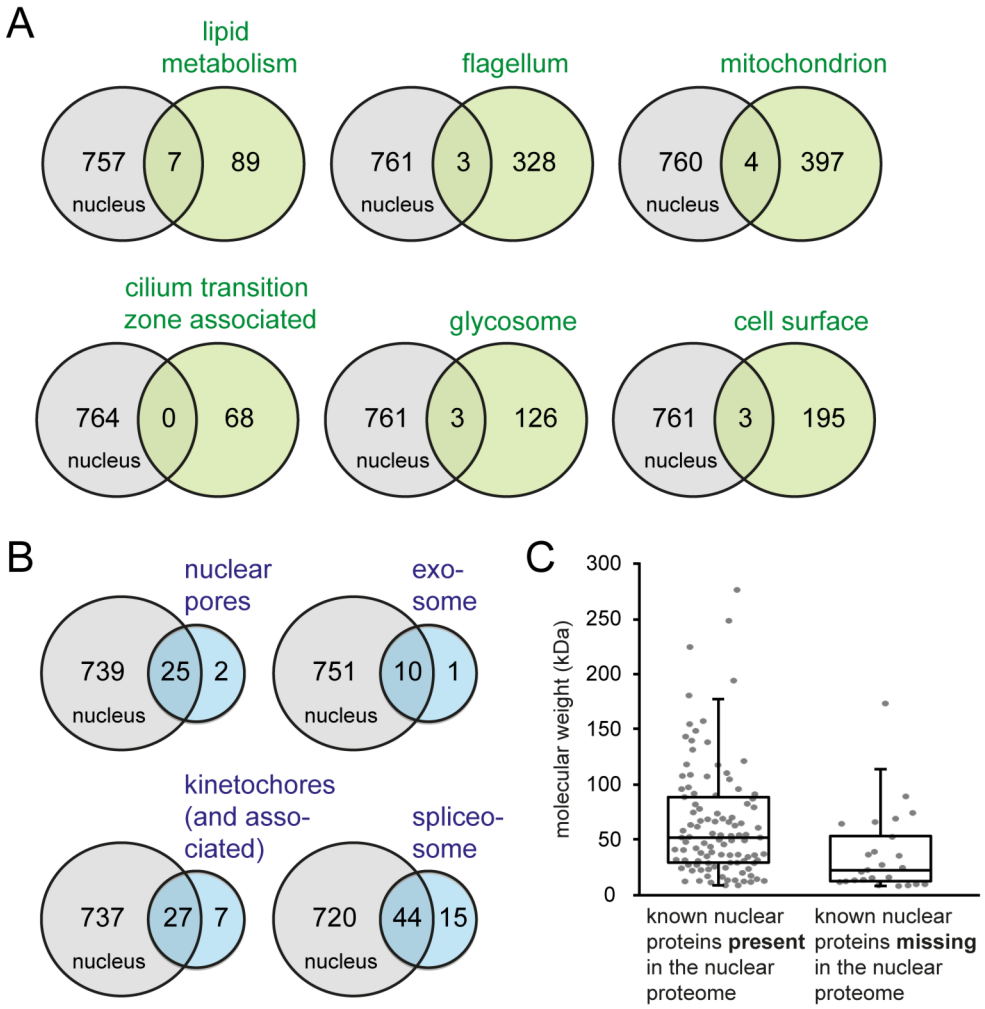


Figure 4.9: Validation of nuclear proteome.

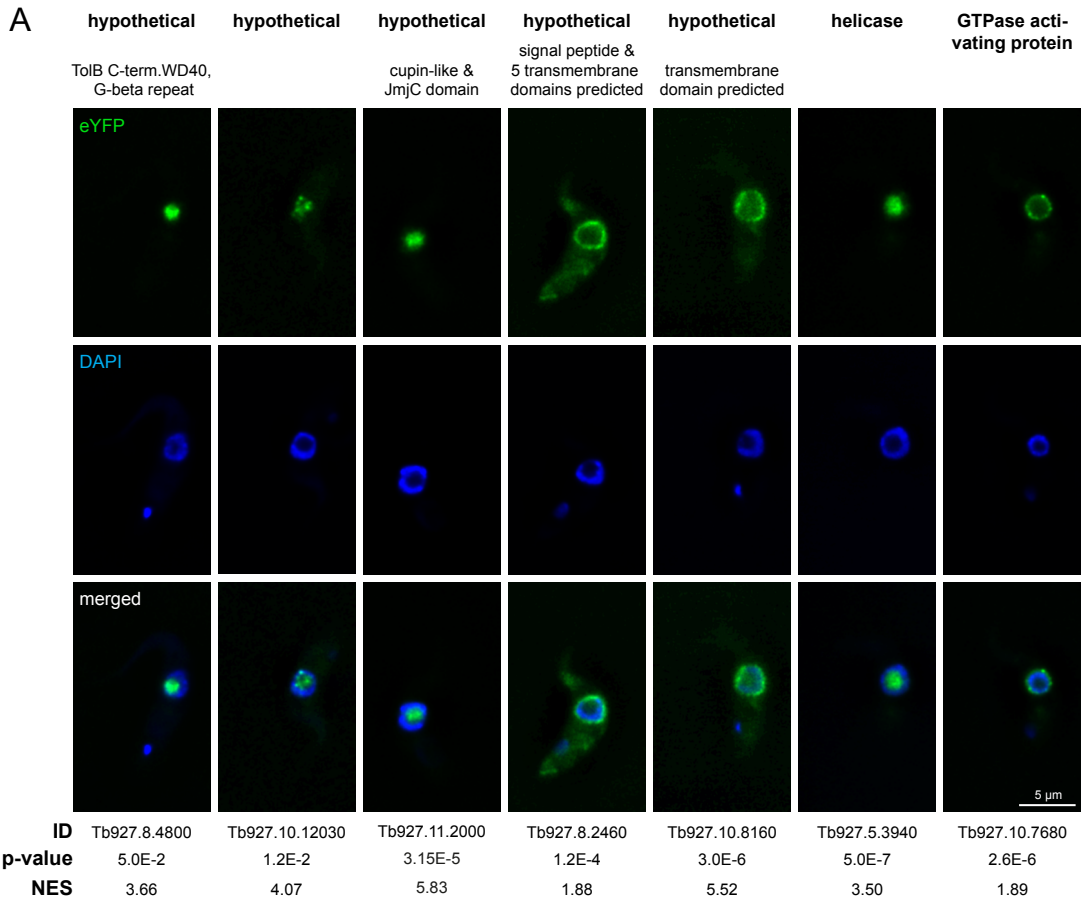
Comparison of the nuclear proteome with known nuclear and non-nuclear structures. **(A)** The content of the nuclear proteome was compared with proteins of non-nuclear structures such as proteins involved in lipid metabolism, the flagellum, the mitochondrion, the cilium transition zone, the glycosome and the cell surface. The number of proteins present in both proteomes is displayed in the overlap of the circles. **(B)** The content of the nuclear proteome was compared with proteins of known nuclear structures: the nuclear pores, the exosome, the kinetochores and the spliceosome. The number of proteins present in both proteomes is displayed in the overlap of the circles. **(C)** The molecular weight of proteins from known nuclear structures (characterized in B) is shown, for proteins that are present in the nuclear proteome (left) and proteins that are absent from the nuclear proteome (right). Figure taken from Goos et al. 2017 [91].

4.2.3.4 Localization of nuclear proteome candidates

To experimentally confirm the nuclear localization of the proteins and to investigate whether the proteome data set could be used to localize previously uncharacterized proteins, ten randomly chosen nuclear proteome candidates were successfully expressed as eYFP fusions from their endogenous locus[73] (Sup. Table 7.6). The candidates comprised seven previously uncharacterized proteins and three proteins with expected or proven nuclear localization. Their NES ranged from 6.0 to 1.8.

Amongst the three proteins with previously indicated or published nuclear localization were one nuclear pore protein (TbNup48) and two hypothetical proteins (Figure 4.10 A). TbNup48 (Tb927.11.4540) was previously identified as outer ring protein of the nuclear pore complex (NPC) by proteomic analyses of NPC-containing fractions. Its localization to the NPC was confirmed by GFP tagging[77]. One hypothetical protein (Tb927.9.1410) was already described as component of the trypanosome kinetochore (KKIP1-interacting protein) with a confirmed nuclear localization by eYFP tagging (nuclear rim)[108]. The other hypothetical protein (Tb927.2.5810) is annotated as putative "holiday-junction resolvase-like of SPT6/SH2 domain containing protein", which is crucial for chromosome segregation during mitosis and localizes to the nucleoplasm (http://tryptag.org)[107]. It localized to the nucleoplasm. For all three proteins, the previously known nuclear localizations could be demonstrated during this work: TbNup48 localizes to nuclear pores; Tb927.9.1410 localizes to the nuclear rim; Tb927.2.5810 localizes to the nucleoplasm (Figure 4.10 A). These proteins served as proof of principle.

Amongst the seven uncharacterized proteins were five hypothetical proteins (including one with a p-value slightly above the threshold of 0.05), one helicase and one GTPase activating protein (Figure 4.10 B). Three hypothetical proteins (Tb927.10.12030, Tb927.8.4800, Tb927.11.2000) and the helicase (Tb927.5.3940) mainly localized to the nucleolus (visualized by the absence of DAPI staining). The GTPase activating protein (Tb927.10.7680) localized to a spot-like structure at the nuclear rim, highly reminiscent of nuclear pores. The two remaining hypothetical proteins (Tb927.10.8160, Tb927.8.2460) also localized to the nuclear rim, but the fluorescence was distributed more homogenously around the rim rather than accumulating in spots. Since both proteins have predicted trans-membrane domains, this indicates a localization to the nuclear membrane (localization to the nucleus-adjacent ER membrane cannot be excluded due to the limits of light microscopy) (Figure 4.10 B).



(caption on next page)

Figure 4.10: Localization as evidence for novel nuclear proteins.

Validation of nuclear localization by expressing eYFP fusion proteins. Protein candidates were expressed as eYFP fusion proteins from their endogenous loci. Representative images (single plane images of deconvolved z-stacks) are depicted. For each protein the amount of enrichment is indicated by the nuclear enrichment score. (NES), the p-values indicate the significance. DNA was stained with DAPI. For Details see Sup. Table 7.6. **(A)** Three proteins of the nuclear proteome with previously indicated nuclear localization **(B)** Seven proteins of the nuclear proteome with previously unknown localization.

Taken together, we provide a high-quality nuclear proteome of procyclic *T. brucei*. A comparison with proteomes of several experimentally characterized nuclear and non-nuclear structures and pathways shows that the proteome contains about 80% of all nuclear proteins and less than 2% false positives. Furthermore, for ten proteins (seven of these with previously unknown localization) a nuclear localization was additionally confirmed by expressing eYFP fusions.

4.2.4 Purification of *T. brucei* nuclei with associated NPGs

The data above show successful purification of nuclei. The method was now used to determine the proteome of NPGs by comparing nuclei purified with and without sinefungin incubation, resulting in cells with and without NPGs. For initial test experiments, a cell line expressing the NPG granule marker eYFP-DHH1 from the endogenous locus (p2829) was used to microscopically follow NPGs throughout the procedure (Figure 4.11 A-C). Later, a cell line co-expressing mChFP-DHH1 (p2845) and PABP2-eYFP-4xTy1 (SK148) (for details on plasmids see chapter 2.5) from the endogenous loci was used, to additionally enable the detection of PABP2 by western blot (Figure 4.11 D and E). Subsequently, nuclei were purified as described above (Figure 4.11 A). Applying the POLYTRON® mixer for mechanical cell lysis presumably constitutes the critical step for NPG survival. Therefore, nuclei of disrupted cells expressing the NPG marker e-YFP-DHH1 were stained with DAPI and analyzed for NPG attachment directly after cell lysis. I found that NPGs were still present at the nuclei from sinefungin treated cells and, as expected, no granules were observed at the nuclei of untreated cells (Figure 4.11 B). After the complete purification process, a sample of fraction NUC of SF treated cells was stained with DAPI and microscopically analyzed (Figure 4.11 C). NPGs, visualized by eYFP-DHH1, still remained attached to the nuclei (Figure 4.11 C, right panel).

Moreover, quantitative western blotting showed that all three NPG marker proteins (PABP2, SCD6 and DHH1) were between 2- and >5-fold enriched in nuclei of sinefungin treated cells in comparison to nuclei of untreated cells (Figure 4.11 D). Protein samples of the two NUC fractions were also analyzed on a Coomassie-stained SDS gel. Here, no obvious differences on the Coomassie-stain pattern between untreated and SF treated NUC fractions were visible (Figure 4.11 E): the identification of proteins that are enriched in SF treated nuclei requires a more sensitive approach.

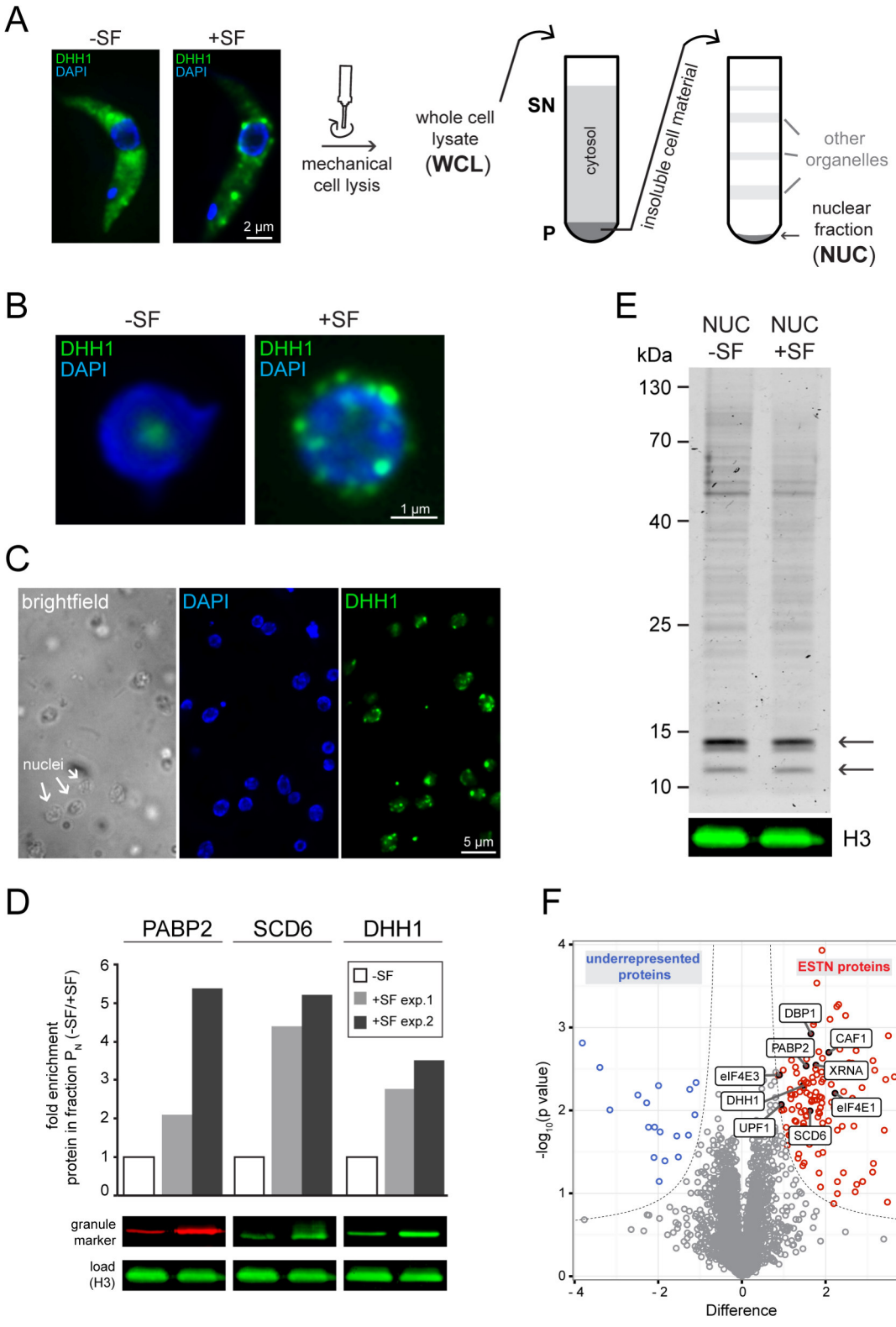


Figure 4.11: Purification of nuclei with attached NPGs.

(A) Trypanosome cells were incubated in the presence and absence of sinefungin (-/+SF) resulting in cells with and without NPGs. Subsequently, the purification was conducted for both conditions simultaneously. Images of cells expressing the NPG marker eYFP-DHH1 (green) are shown as single plane images of deconvolved z-stacks. DNA was stained with DAPI. **(B)** Nuclei with and without NPGs after mechanical cell lysis. In purified nuclei from cells expressing the NPG marker eYFP-DHH1 (green),

the presence of NPGs is specific to sinefungin treated cells. The granules survive the mechanical cell lysis by the POLYTRON mixer. Representative images are shown as deconvolved z-stack projections. DNA was stained with DAPI. **(C)** Microscopic analysis of NUC fraction of sinefungin-treated cells. Isolated nuclei are clearly visible as ovoids and only few other structures are present, such as remnants of flagella (brightfield image). Nuclei are intact (native shape, nucleolus is visible by absence of DAPI staining) and only few kinetoplasts (DAPI image) are present (Results are identical to purification of untreated cells (compare Figure 4.8)). NPGs are still attached to purified nuclei of cells expressing the NPG marker eYFP-DHH1 (green). Samples were not fixed to the slide and moved during imaging, thus the different channels do not completely overlap. Fluorescence images are shown as deconvolved z-stack projections, the phase image is a single plane image. DNA was stained with DAPI. **(D)** Three NPG marker proteins are enriched in fraction NUC of sinefungin-treated cells. Fraction NUC of control cells (-SF) and NPG-induced cells (+SF) was quantified by western blot for the NPG marker proteins PABP2, SCD6 and DHH1. For each protein, the fold-enrichment upon sinefungin treatment from two independent experiments is plotted. The lower panels show a representative western blot, with histone H3 (H3) serving as loading control. **(E)** Fractions NUC of untreated and SF treated cells. A representative Coomassie-stained gel loaded with NUC fractions of untreated (-SF) and SF treated cells (+SF) shows no obvious differences in protein pattern. Bands corresponding to highly enriched histones in fraction NUC are marked with arrows. Additionally, histone H3 was detected by western blot (lower panel). A comparison of fractions NUC and WCL of the purification procedure is shown separately in Figure 4.8 (nuclear proteome). **(F)** Volcano plot of the mass spectrometry data. Proteins enriched in sinefungin treated nuclei (ESTN) are displayed in red, proteins that were underrepresented in nuclei of sinefungin treated cells in blue and proteins with no enrichment in grey. All nine previously known NPG proteins are labelled.

4.2.5 128 proteins are enriched in sinefungin-treated nuclei (ESTN)

Protein samples of fraction NUC (enriched in nuclei) of untreated and SF treated cells from three independent biological replicates were analyzed by quantitative label-free mass spectrometry. 3195 protein groups were detected, which corresponded to about 1/3 of all proteins encoded by the *T. brucei* genome[12] (Sup. Table 7.7 C). By subtractive proteomics, 131 proteins were identified as significantly (FDR<0.05) enriched in the nuclear fraction of SF treated cells in comparison to untreated cells. Results are shown as volcano plot in Figure 4.11 F. The 131 ESTN proteins (enriched in sinefungin-treated nuclei), including all nine previously known NPG proteins (names are indicated) are labelled in red. The fact that all previously known NPG proteins were detected with this method indicates a successful co-purification of NPGs together with nuclei[49]. Furthermore 18 proteins were underrepresented in fraction NUC of SF treated cells (blue) (Sup. Table 7.7 B). During this work these underrepresented proteins were not further analyzed. The number of 131 ESTN proteins was reduced to 128 by removing two retrotransposon hot spot proteins (Tb11.v5.0670, Tb927.2.830) and one protein that seemed to be a sequence orphan with partial overlap to PAN2 (Tb11.v5.0553) (Sup. Table 7.7 A).

4.2.6 Bioinformatic characterization of the 128 ESTN proteins

4.2.6.1 ESTN proteins are highly enriched in mRNA metabolism proteins, but lack most translation factors

Inhibition of trans-splicing most likely affects proteins involved in mRNA metabolism. Therefore, all 128 ESTN proteins were examined for a potential function in mRNA processing pathways. We classified proteins as 'involved in mRNA metabolism', when they either had a known function in mRNA metabolism (shown experimentally or by being a homologue to a protein with a known function in mRNA metabolism) or possessed RNA binding domains. Proteins were also classified as 'involved', when two independent lines of literature evidence[34,39,62,109,116–123] supported a function in mRNA metabolism, for example identification in the screens for mRNA binding and posttranscriptional regulation described in[12,13,40,41,109] together with confirmed localization to stress granules[34] (or this work, data not shown). Proteins were classified as 'probably involved' in mRNA metabolism, if evidence was based on only one such evidence. The vast majority of the ESTN proteins were either involved (93/128) or probably involved (13/128) in mRNA metabolism (Figure 4.12 A). Analyzing only the subgroup of validated NPG proteins with a confirmed sinefungin-induced change in localization (see chapter 4.2.7 below) (55/128) essentially had the same outcome (Sup. Figure 7.2 A). Details of the analysis and the classification of every protein can be found in Sup. Table 7.7 A.

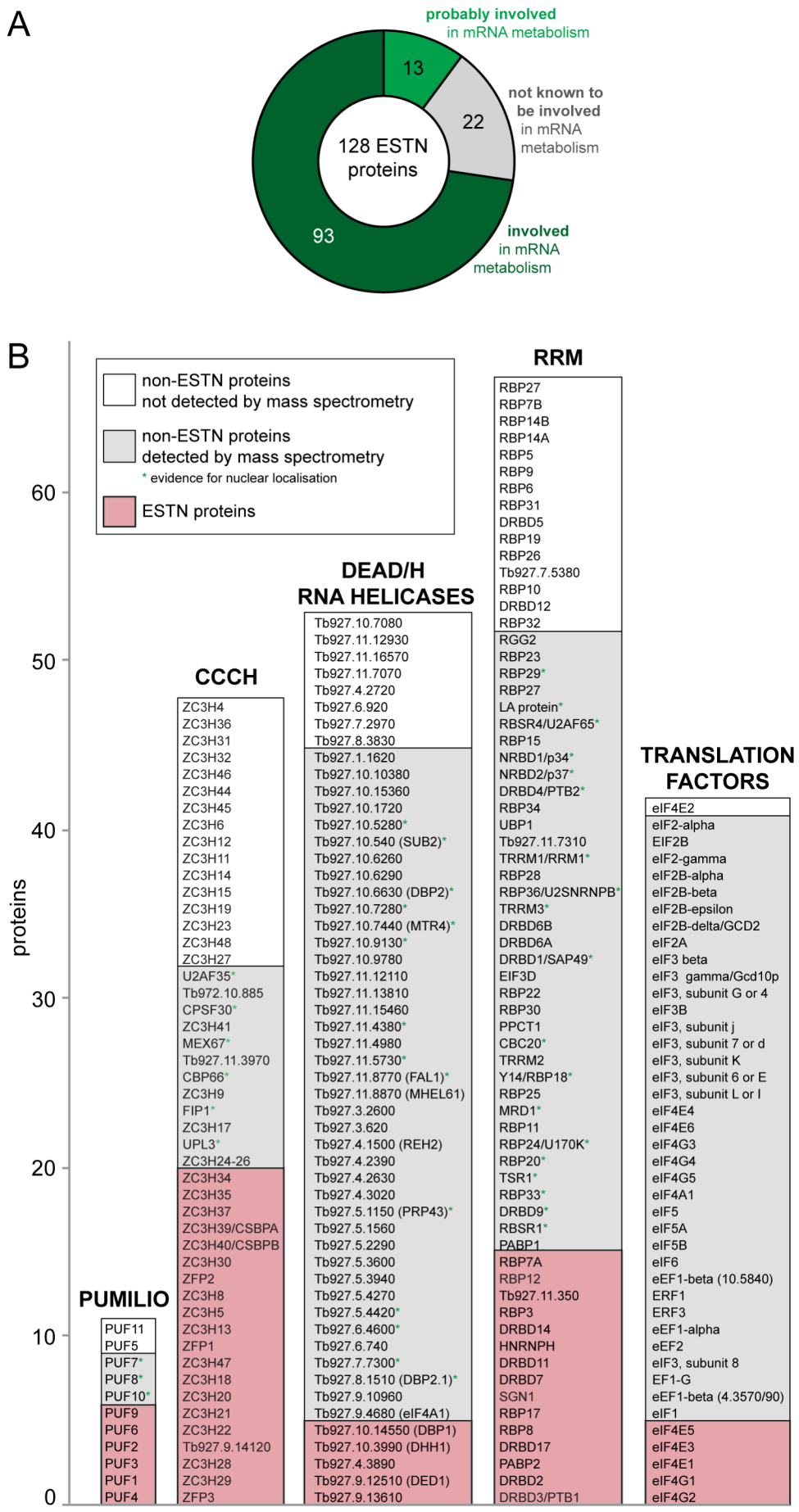
Next, ESTN proteins involved in mRNA metabolism were examined in more detail. To analyze whether a particular group of mRNA metabolism proteins was over- or underrepresented in the group of ESTN proteins, five groups of trypanosome mRNA metabolism proteins were analyzed for their presence of ESTN proteins: Pumilio domain proteins, CCCH type zinc finger, RRM-domain proteins, DEAD/H RNA helicases and translation factors (groups are not necessarily exclusive). Each protein within the five groups was compared with the mass spectrometry data sets of purified nuclei. They were characterized as (1) not detected by mass spectrometry (white), (2) not enriched in SF treated nuclei but within the mass spectrometry data (grey) or (3) within the 128 ESTN proteins (red) (Figure 4.12 B). The list of trypanosome mRNA metabolism proteins was originally taken from Fritz et al. 2015[34], where these were compared with a mass spectrometry data set of proteins enriched in starvation stress granules (Figure 4.12 C).

The group of ESTN proteins (red background Figure 4.12 B) contained over half of the Pumilio domain proteins, many CCCH type zinc finger proteins and some of the RRM domain proteins. Interestingly, only a very small fraction of the translation factors and DEAD/H RNA helicases were ESTN proteins. While the small content in DEAD/H RNA helicases can be explained by the fact that many of

these proteins have either nuclear localization or very low expression levels, this does not apply to the translation factors, which are cytoplasmic and have very high expression levels. The comparison of ESTN proteins (in this work) and stress granule proteins (Fritz et al. 2015) in their content of mRNA metabolism proteins (Figure 4.12 C) revealed interesting differences between the two granule types. These differences were most obvious when looking at the translation factors: in contrast to stress granules, NPGs seemed to lack canonical translation factors. A closer look revealed, that all translation factors that act downstream of the eIF4F complex, including elongation and release factors, are absent from the group of ESTN proteins. The three isoforms of the eIF4F complex that are likely involved in translation of bulk mRNA in PCF trypanosomes, eIF4A1, eIF4E4 and eIF4G3, are also absent from NPGs. The same is true for poly(A) binding protein 1 (PABP1), which is part of the same complex[84,124,125]. Only five isoforms of subunits of the eIF4F complex, eIF4G1, eIF4G2, eIF4E1, eIF4E3 and eIF4E5 are present in NPGs (Figure 4.12 B). Neither of these proteins is likely to be involved in canonical cap dependent translation initiation of bulk mRNAs based on the following reasons: (1) Depletion of eIF4G1, G2, E5 causes only a very minor reduction in translation[122,126] (2) *Leishmania* eIF4E1 interacts with the 4E-interacting protein, but not with any of the six *Leishmania* eIF4Gs[127,128] (3) eIF4E3 binds mainly to m7GTP, but not to the type 4 cap that is present on trypanosome mRNAs[84,126]. The significant absence of canonical translation factors indicates, that NPGs are RNP complexes that have not yet entered translation.

A large number of mRNA metabolism proteins were detected by mass spectrometry but were not enriched in nuclei of SF treated cells, consequently non-ESTN proteins (grey background Figure 4.12 B).). Amongst these were the vast majority of translation factors, a large fraction of RRM domain proteins as well as DEAD/H RNA helicases and some Pumilio and CCCH proteins. Some of these non-ESTN proteins localize to the nucleus in both SF and untreated cells. Indeed, for many of the non-ESTN Pumilio, CCCH, RRM domain proteins and DEAD/H RNA helicases evidence for a nuclear localization (42/92 proteins; green * Figure 4.12 B) exists. Translation factors in contrast are mostly cytoplasmic and were probably detected due to their high abundance (see above) or because they shuttle between nucleus and cytoplasm.

Some of the proteins containing RNA binding domains and some DEAD/H RNA helicases as well as one translation initiation factor (eIF4E2) were not detected by mass spectrometry (white background Figure 4.12 B). Either this is due to a very low expression level, or these proteins are neither nuclear nor ESTN proteins.


A
probably involved in mRNA metabolism
13
not known to be involved in mRNA metabolism
22
128 ESTN proteins
93
involved in mRNA metabolism
B
non-ESTN proteins not detected by mass spectrometry
non-ESTN proteins detected by mass spectrometry
* evidence for nuclear localisation
ESTN proteins
proteins
0
10
20
30
40
50
60
PUMILIO
PUF11
PUF5
PUF7*
PUF8*
PUF10*
PUF9
PUF6
PUF2
PUF3
PUF1
PUF4
CCCH
ZC3H4
ZC3H36
ZC3H31
ZC3H32
ZC3H46
ZC3H44
ZC3H45
ZC3H6
ZC3H12
ZC3H11
ZC3H14
ZC3H15
ZC3H19
ZC3H23
ZC3H48
ZC3H27
U2AF35*
Tb972.10.885
CPSF30*
ZC3H41
MEX67*
Tb927.11.3970
CBP66*
ZC3H9
FIP1*
ZC3H17
UPL3*
ZC3H24-26
ZC3H34
ZC3H35
ZC3H37
ZC3H39/CSBPA
ZC3H40/CSBPB
ZC3H30
ZFP2
ZC3H8
ZC3H5
ZC3H13
ZFP1
ZC3H47
ZC3H18
ZC3H20
ZC3H21
ZC3H22
Tb927.9.14120
ZC3H28
ZC3H29
ZFP3
DEAD/H RNA HELICASES
Tb927.10.7080
Tb927.11.12930
Tb927.11.16570
Tb927.11.7070
Tb927.4.2720
Tb927.6.920
Tb927.7.2970
Tb927.8.3830
Tb927.1.1620
Tb927.10.10380
Tb927.10.15360
Tb927.10.1720
Tb927.10.5280*
Tb927.10.540 (SUB2)*
Tb927.10.6260
Tb927.10.6290
Tb927.10.6630 (DBP2)*
Tb927.10.7280*
Tb927.10.7440 (MTR4)*
Tb927.10.9130*
Tb927.10.9780
Tb927.11.12110
Tb927.11.13810
Tb927.11.15460
Tb927.11.4380*
Tb927.11.4980
Tb927.11.5730*
Tb927.11.8770 (FAL1)*
Tb927.11.8870 (MHEL61)
Tb927.3.2600
Tb927.3.620
Tb927.4.1500 (REH2)
Tb927.4.2390
Tb927.4.2630
Tb927.4.3020
Tb927.5.1150 (PRP43)*
Tb927.5.1560
Tb927.5.2290
Tb927.5.3600
Tb927.5.3940
Tb927.5.4270
Tb927.5.4420*
Tb927.6.4600*
Tb927.6.740
Tb927.7.7300*
Tb927.8.1510 (DBP2.1)*
Tb927.9.10960
Tb927.9.4680 (eIF4A1)
Tb927.10.14550 (DBP1)
Tb927.10.3990 (DHH1)
Tb927.4.3890
Tb927.9.12510 (DED1)
Tb927.9.13610
RRM
RBP27
RBP7B
RBP14B
RBP14A
RBP5
RBP9
RBP6
RBP31
DRBD5
RBP19
RBP26
Tb927.7.5380
RBP10
DRBD12
RBP32
RGG2
RBP23
RBP29*
RBP27
LA protein*
RBSR4/U2AF65*
RBP15
NRBD1/p34*
NRBD2/p37*
DRBD4/PTB2*
RBP34
UBP1
Tb927.11.7310
TRRM1/RRM1*
RBP28
RBP36/U2SNRNPB*
TRRM3*
DRBD6B
DRBD6A
DRBD1/SAP49*
EIF3D
RBP22
RBP30
PPCT1
CBC20*
TRRM2
Y14/RBP18*
RBP25
MRD1*
RBP11
RBP24/U170K*
RBP20*
TSR1*
RBP33*
DRBD9*
RBSR1*
PABP1
RBP7A
RBP12
Tb927.11.350
RBP3
DRBD14
HNRNPH
DRBD11
DRBD7
SGN1
RBP17
RBP8
DRBD17
PABP2
DRBD2
DRBD3/PTB1
TRANSLATION FACTORS
eIF4E2
eIF2-alpha
EIF2B
eIF2-gamma
eIF2B-alpha
eIF2B-beta
eIF2B-epsilon
eIF2B-delta/GCD2
eIF2A
eIF3 beta
eIF3 gamma/Gcd10p
eIF3, subunit G or 4
eIF3B
eIF3, subunit j
eIF3, subunit 7 or d
eIF3, subunit K
eIF3, subunit 6 or E
eIF3, subunit L or I
eIF4E4
eIF4E6
eIF4G3
eIF4G4
eIF4G5
eIF4A1
eIF5
eIF5A
eIF5B
eIF6
eEF1-beta (10.5840)
ERF1
ERF3
eEF1-alpha
eEF2
eIF3, subunit 8
EF1-G
eEF1-beta (4.3570/90)
eIF1
eIF4E5
eIF4E3
eIF4E1
eIF4G1
eIF4G2

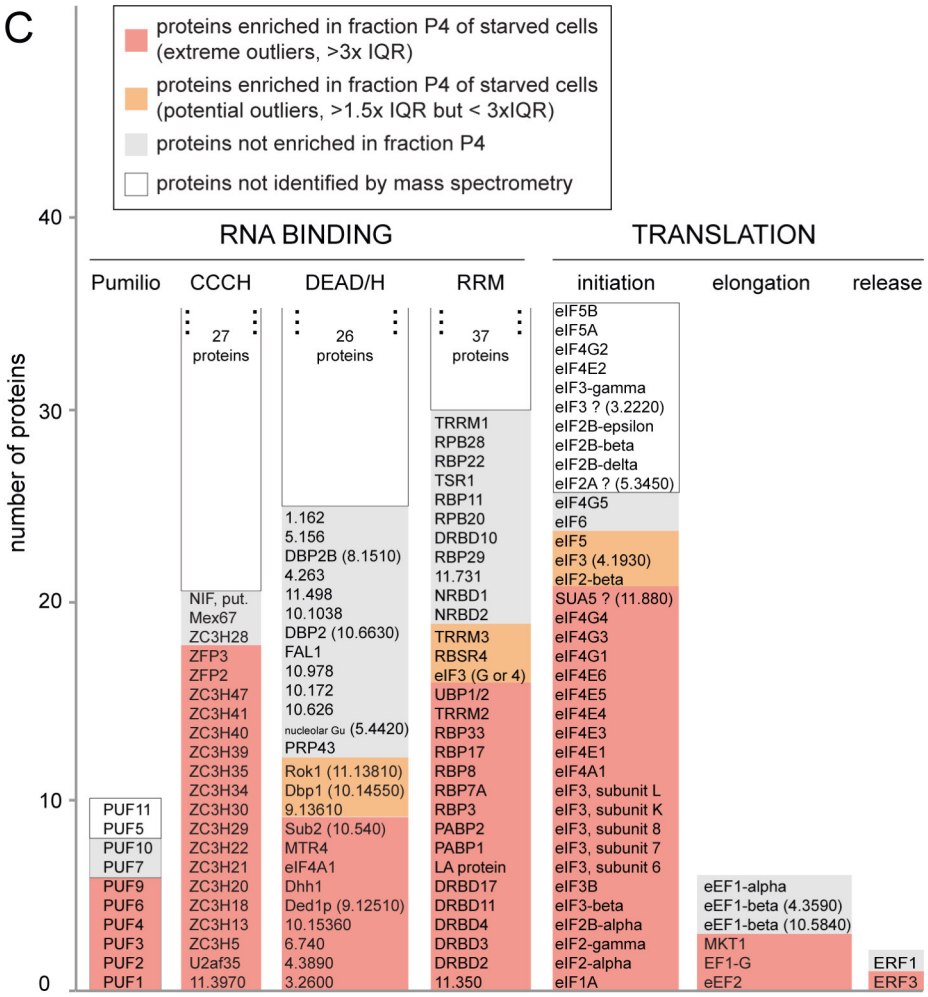


Figure 4.12: ESTN proteins are highly enriched in mRNA metabolism proteins but lack most translation factors.

(A) The majority of ESTN proteins are involved in mRNA metabolism. 128 ESTN proteins were classified as involved (green), probably involved (light green) and not known to be involved (grey) in mRNA metabolism. Details of the analysis are depicted in Sup. Table 7.7 A. **(B)** Mass spectrometry dataset of nuclear preparations in comparison with control proteins. All proteins with RNA binding features (Pumilio, CCCH, DEAD/H RNA helicases, RRM) and translation factors (initiation, elongation, release) that are present in the *T. brucei* genome were compared with the mass spectrometry dataset of the nuclear preparations (treated with SF and untreated). Proteins are shown with orange background if they were among the ESTN proteins, with grey background if they were not among the ESTN proteins but detected by mass spectrometry and with white background if they were not detected by mass spectrometry. Non-ESTN Proteins that were detected by the mass spectrometry were marked with a green *, if there was evidence for nuclear localization. For hypothetical proteins the unique gene identifier is shown. **(C)** Mass spectrometry dataset of granule preparations (starvation stress granules) in comparison with control proteins. The same dataset of control proteins used in (A) was previously compared with a mass spectrometry data set of granule preparations (starved cells and non-starved cells). For details see Fritz et al. 2015[34]. Proteins are shown with red (extreme outliers) or orange (potential outliers) background if they were among the starvation granule-enriched proteins, with grey background if they were not among the starvation granule-enriched proteins but detected by mass spectrometry and with white background if they were not detected by the mass spectrometry. Non-detected proteins are not listed but shown as total numbers. For hypothetical proteins the unique gene identifier is displayed, excluding 'Tb927.' (Figure taken and modified from Fritz et al. 2015[34])

4.2.6.2 ESTN proteins are enriched in features that contribute to granule localization

Previous studies revealed a correlation between glutamate (Q)- and asparagine (N)- enriched regions of proteins and protein localization to RNP granules. It was suggested, that in proteins of P-bodies, these enriched regions have self-propagating prion-like features and function as modulators of specific protein-protein interactions, thus contributing to granule aggregation[129,130].

Since the data set of 128 ESTN proteins arose from a purification method with the goal to purify granule proteins (NPGs), it could be hypothesized that ESTN proteins may have similar features. The content in Q- and N-rich regions for all 128 ESTN proteins and a randomly chosen group of control proteins was calculated based on the method described in Reijns et al. 2009[129] and Michelitsch and Weissman 2000[130]. Briefly, amino acid sequences of all proteins were divided into 60-mers. A 60-mer was defined as Q- or N-rich region, if it contained at least seven glutamate or six asparagine residues. There was a significant enrichment in Q-rich regions in the group of ESTN proteins in comparison to the control group (result of unpaired, two-tailed students t-TEST<$1*10^{-4}$) (Figure 4.13 A). There was also an enrichment in N-rich regions, but with no statistical significance (result of unpaired, two-tailed students t-TEST<0.07). The screen was repeated with the validated group of NPG proteins (55/128) (validation by localization study, see below in chapter 4.2.7), displaying essentially the same outcome (Sup. Figure 7.2 B). The results pointed to an involvement of ESTN proteins in RNP granule aggregation, further evidence for a successful NPG purification.

Recently, a group of British scientists founded the TrypTag project, with the aim to determine the localization of every trypanosomal protein within the cell. The still on-going project fluorescently tags every protein gene in the genome and provides images on their website http://tryptag.org[107]. The complete workflow of the procedure is available as well. It seems that due to the workflow, cells are in a stressed state at the time of image acquisition (high cell density, no dilution step between growth of clones in 24-well plates and imaging of cells). Thus, potential granule proteins may react by re-localizing to cytoplasmic stress granules. A rough screen of all 128 ESTN proteins that were already analyzed by this project in July 2017 (72/128) confirmed that the majority of ESTN proteins (52/72) indeed localized to cytoplasmic granules (Figure 4.13 B) (Sup. Table 7.7 A). To summarize, the majority of ESTN proteins seem to be potential RNA granule proteins, based on their enrichment in Q-rich regions and protein localization to stress granules.

4.2.6.3 ESTN proteins are mostly unique to kinetoplastids

The 128 ESTN proteins were screened by BLAST for potential homologues in four eukaryotic model organisms: *S. cerevisiae* (taxid4932), *C. elegans* (taxid6239), *M. musculus* (taxid10090) and *P. falciparum* (taxid5833). The first hit of the BLAST analysis was used for a reciprocal BLAST against the *T. brucei* genome. If the initial protein was again the first hit and had an e-value of < $1x10^{-5}$, the protein was classified as having a homologue in the respective organism (details about the analysis, see Sup. Table 7.7 D). The majority (82/128) had no homologue in any of the organisms and are thus likely unique to kinetoplastids. The remaining 46 proteins had putative homologues in either one, two, three or four of the screened model organisms. 14 proteins had putative homologues in all four organisms and thus seem to be conserved among eukaryotes (Figure 4.13 C). The analysis was repeated with the group of validated NPG proteins (55/128) (validation by localization study see below), with essentially the same outcome (Sup. Figure 7.2 C).

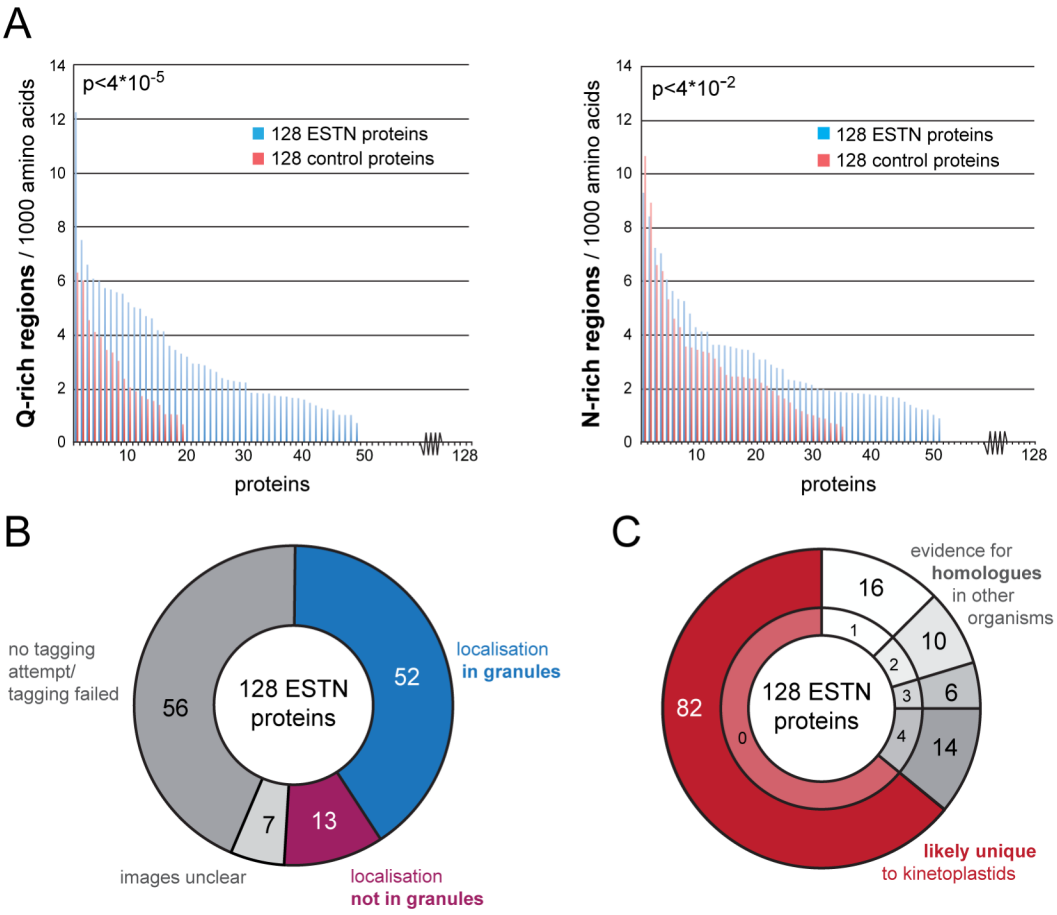


Figure 4.13: Characterization of ESTN proteins.

(A) Enrichment of Q-rich and N-rich regions. The number of Q-rich regions and N-rich regions per 1000 amino acids is shown for the 128 ESTN proteins (blue) in comparison to a randomly chosen group of control proteins (red), respectively. One bar represents one protein. **(B)** Comparison with TrypTag.org database. All ESTN proteins were screened for localization to

granules in the TrypTag.org database. Cells imaged for this database are probably under stress due to the high-throughput work flow of the project. Thus, potential granule proteins react by re-localizing to cytoplasmic stress granules. Proteins with unequivocal eYFP signal were classified as localized in granules (blue) or not localized in granules (pink). Some proteins had a low eYFP signal or the localization was unclear (light grey). For half of the proteins no images were available due to failed or unattempted tagging (dark grey). **(C)** The majority of ESTN proteins are likely unique to kinetoplastids. 128 ESTN proteins were screened for putative homologues in *S. cerevisiae, C. elegans, M. musculus* and *P. falciparum* by BLAST (default parameters). The first hit of the BLAST analysis was then used for a reciprocal BLAST against the *T. brucei* genome. If the initial protein was again the first hit and had an e-value of < $1x10^{-5}$, the protein was classified as having a homologue in the respective organism. Using these definitions, the majority of proteins were unique to kinetoplastids (*T. brucei*) (red). The remaining proteins had putative homologues in either one, two, three or four of the screened organisms (grey). Details of the analysis are shown in Sup. Table 7.7 D.

To summarize, the majority of ESTN proteins are involved in mRNA metabolism and unique to kinetoplastids. ESTN proteins contain a fraction of all known trypanosomal cytoplasmic RNA binding proteins (Pumilio, CCCH and RRM domain,) and RNA helicases, but lack canonical translation factors. The proteins are enriched in regions that are known to contribute to granule localization. A screen of fluorescence images from TrypTag.org confirmed this.

4.2.7 Localization of ESTN proteins

ESTN stands for 'enriched in sinefungin treated nuclei' and ESTN proteins should therefore partially or fully change their localization upon sinefungin treatment from the cytoplasm to either the nucleus or to nucleus-attached NPGs. To investigate the localization of the ESTN proteins in the absence and presence of sinefungin, proteins were expressed as eYFP fusion proteins.

Ten of the 128 ESTN proteins have previously been analyzed for their localization (grey font in boxes Figure 4.14 A)[49]. SF treatment caused eight of them to localize to NPGs, one (DRBD3) to localize to the nucleus and one (VASA/DBP1) to localize to both the nucleus and NPGs. The remaining 118 proteins were expressed as eYFP fusions from their endogenous loci in a cell line also expressing mChFP-DHH1 (p2845). The fusion proteins were made either according to Kelly et al. 2007[6] (17/118 proteins, using SK141, which adds a double tag out of eYFP and 4xTy1) or to Dean et al. 2015[73] (101/118 proteins, using pPOTv4) (details for each protein, see Sup. Table 7.7 A, details on plasmids, see Sup. Table 7.2). In the first case, the correct size of these fusion proteins was controlled by a western blot probed for the Ty1 epitope tag by the BB2 antibody (data not shown). For 15 proteins two transfection attempts failed and at this point further attempts were abandoned. For 43 proteins, the eYFP signal was too low for microscopic analysis. The remaining 60 proteins were analyzed for their localization in the presence and absence of SF (black font in boxes Figure 4.14 A). By doing so, there were four different conditions, which could be observed during the validation procedure: proteins re-localized

to NPGs (green) or to the nucleus (blue), but also to nuclear pores (orange) or did not change their localization (red) (Figure 4.14 A) (further details, see Sup. Table 7.7 A).

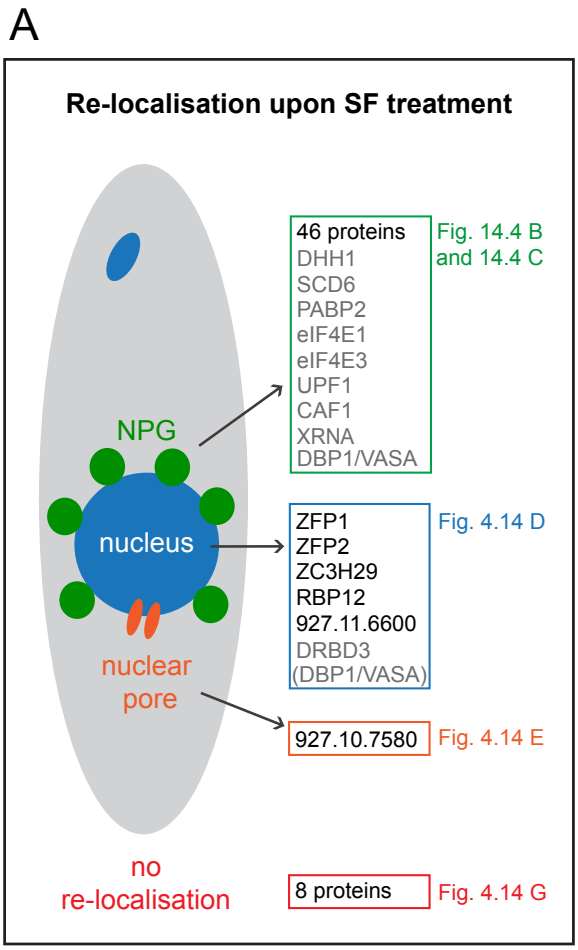
46 of the 60 tagged proteins showed an unequivocal change in localization to NPGs upon sinefungin treatment (Figure 4.14 B and C): the proteins co-localized with the NPG marker protein mChFP-DHH1. In the absence of SF, all proteins were mainly cytoplasmic (complete images see Sup. Figure 7.3). These findings are consistent with previous data, increasing the number of known NPG proteins up to 55.

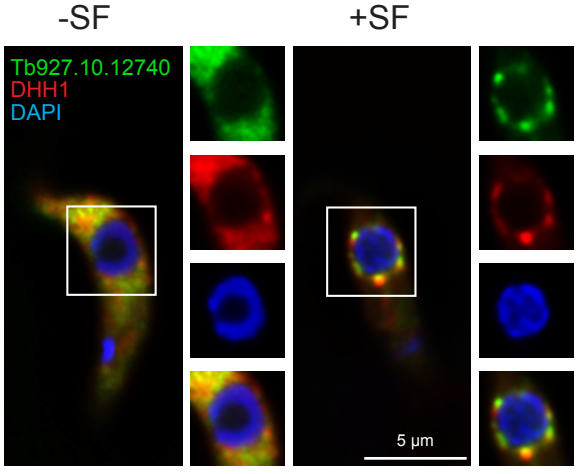
Five of the 60 proteins re-localized from the cytoplasm to the nucleus upon treatment with SF, either completely (ZFP1, ZFP2 and RBP12) or partially (ZC3H29 and Tb927.11.6600) (Figure 4.14 D). These are likely proteins that shuttle between the nucleus and the cytoplasm. Treatment with actinomycin D, which inhibits transcription, did not cause a change in protein localization. This resembles previous results for DRBD3 and VASA/DBP1[74]. The results suggest that the re-localization to the nucleus is not caused by the absence of fully processed mRNAs but is a consequence of the accumulated incompletely processed mRNA, caused by SF treatment.

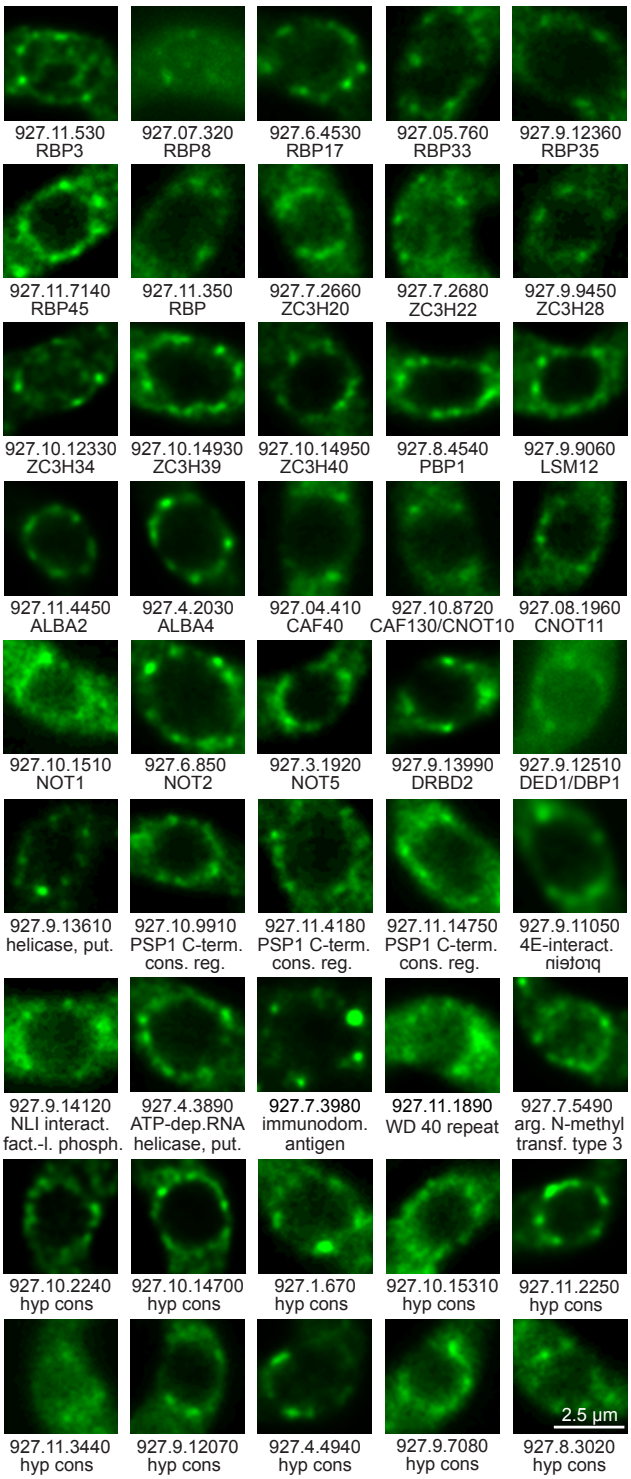
One protein (Tb927.10.7580) changed its localizations from the nucleoplasm to a structure that resembled nuclear pores (Figure 4.14 E). To confirm this assumption, the protein was expressed in a cell line also expressing the nuclear pore marker protein NUP93-mChFP (SK432) (for details on plasmid, see chapter 2.5) (Figure 4.14 F). Since the re-localization from nucleoplasm to nuclear pores does not change the total protein amount in purified nuclei (which were used for the mass spectrometry analysis), it should not be detected as enriched in SF treated nuclei. However, assuming that a fraction of nucleoplasmic proteins gets lost during the preparation of nuclei, this loss could potentially be smaller in the presence of SF, because part of the protein is ‘anchored’ to the nuclear pores.

For eight proteins, SF caused no change in localization (Figure 4.14 G). Proteins were either nuclear (UMSBP1, UMSBP2, Tb927.9.8520, Tb927.10.12980, SGN1), mitochondrial (COXX, Tb927.11.9450) or localized in an undefined cytoplasmic structure (Tb927.10.200). They most likely represent false positives. However, in case of the proteins with nuclear localization, in theory a similar mechanism as described for the protein with localization to nuclear pores could have resulted in the detected enrichment. In the presence of SF these proteins could be attached to mRNA within the nucleus.

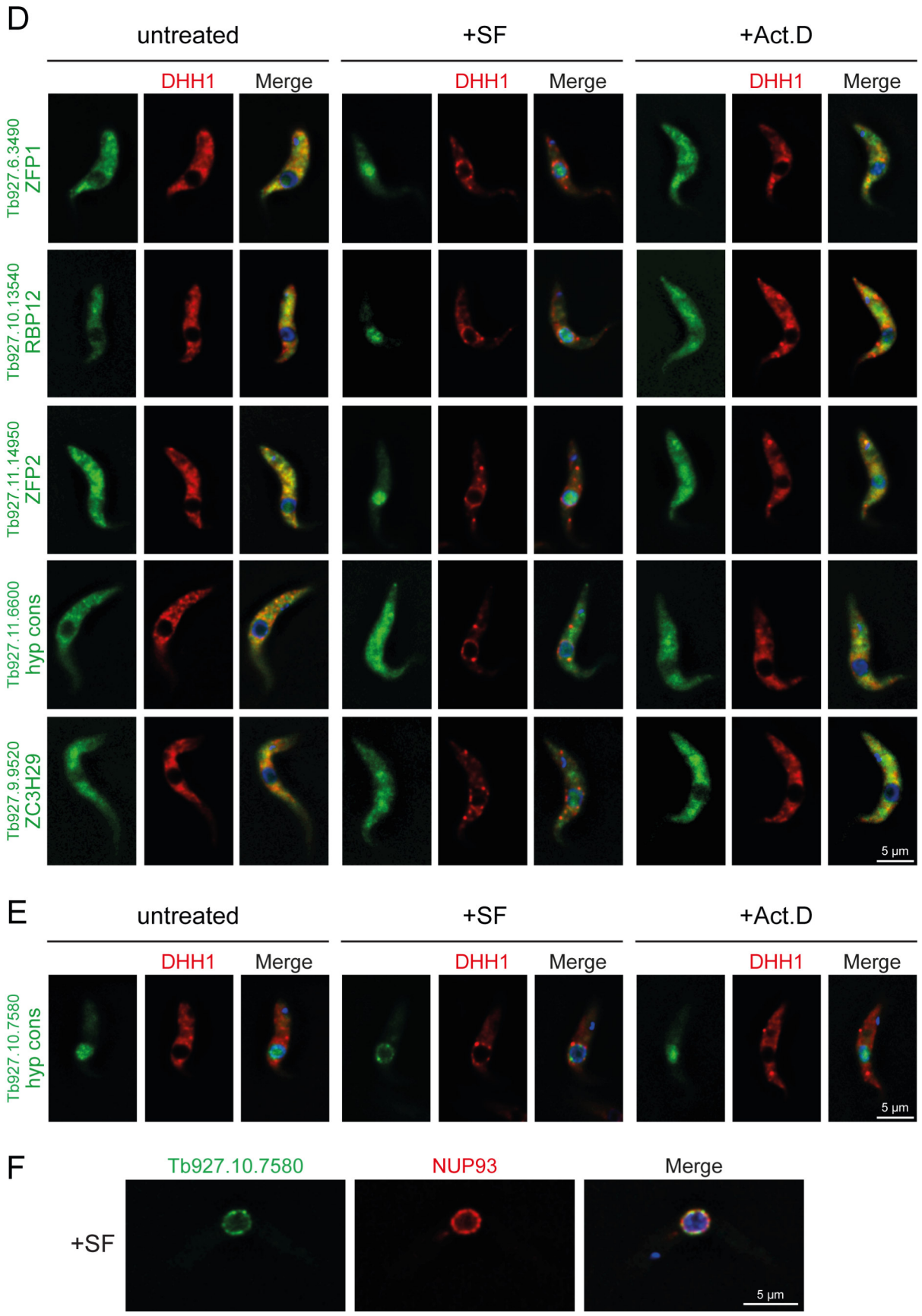
A pie chart of the tagging statistics of all 118 ESTN proteins is shown in Sup. Figure 7.4. For 13% of the tagged proteins the transfections failed and for another 36% the detected eYFP signal was too low for a clear statement. The tagging project http://tryptag.org[107] (personal communication with Samuel Dean, University of Oxford, UK) reports similar numbers of failure.



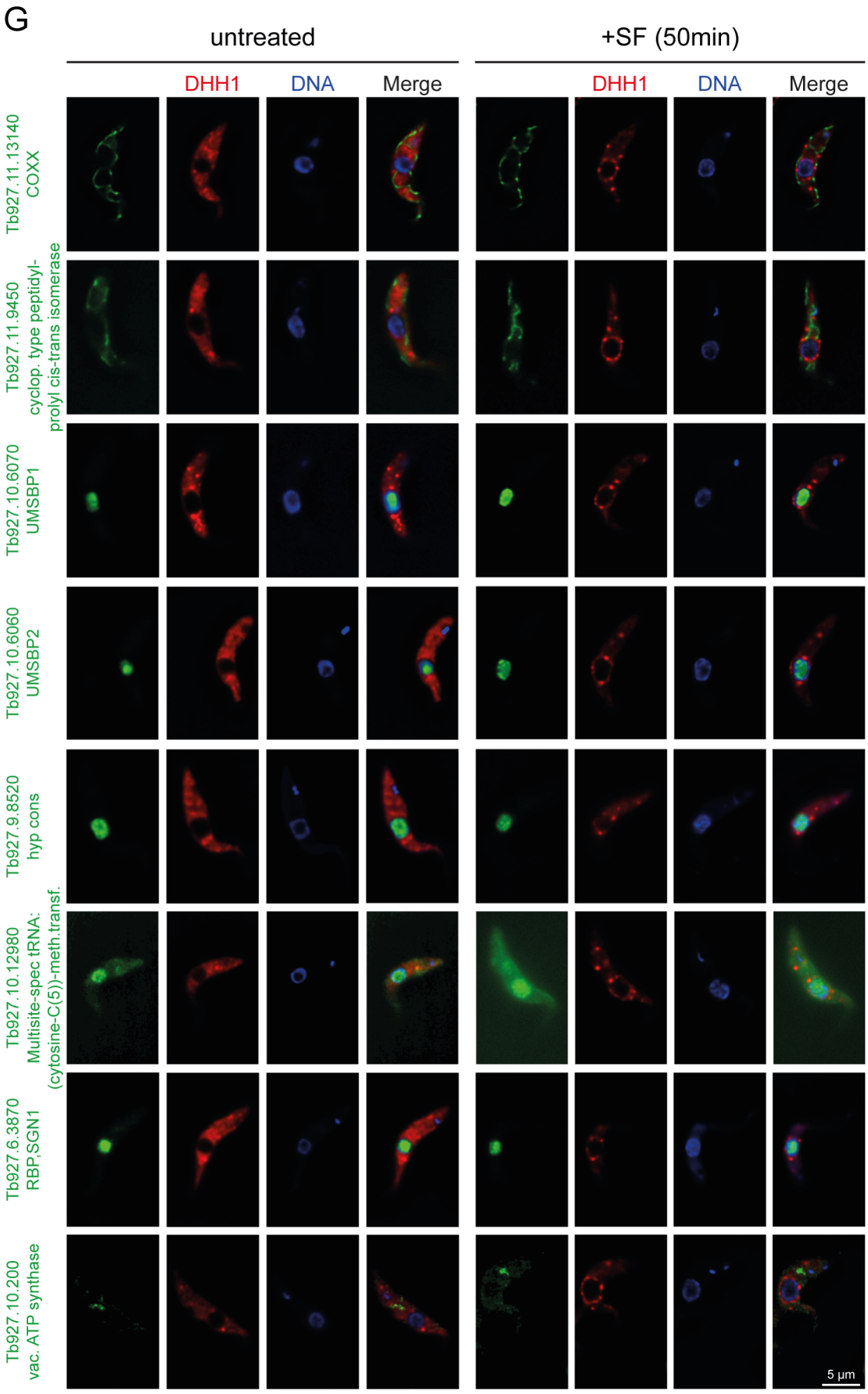




(next part of figure and caption on following pages)



(next part of figure and caption on next page)



(caption on next page)

Figure 4.14: Localization of ESTN proteins.

All ESTN proteins were expressed as eYFP fusion proteins from the endogenous locus (green) in a cell line co-expressing mChFP-DHH1 (red) as NPG granule marker. DNA was stained with DAPI (blue). Images are either single plane images of a deconvolved z-stack or z-stack projections (details in Sup. Table 7.7 A). **(A)** Summary of all validated ESTN proteins and their change in localization upon sinefungin treatment. Previously identified and localized NPG proteins are written in grey[49], proteins localized during this study are written in black with reference to the respective figure. Upon SF treatment, ESTN proteins either changed their localization to NPGs (green box), the nucleus (blue box) or to nuclear pores (orange box) or did not change their localization (red box). **(B and C)** ESTN proteins that re-localize to NPGs in response to sinefungin treatment. For one representative protein a complete cell treated and untreated with sinefungin is shown with the nuclear region as inset. **(B)** Images of complete cells (treated and untreated with sinefungin) including co-localization with mChFP-DHH1 and DAPI staining, are shown in Sup. Figure 7.3. Names and gene identifiers are shown, excluding 'Tb.' **(C)** For the remaining 45 proteins sections only of the nucleus and the surrounding NPGs are shown. **(D)** ESTN proteins that re-localize to the nucleus in response to sinefungin treatment. Representative images of one untreated, one sinefungin-treated and one actinomycin D-treated cell are shown for each protein. **(E)** ESTN protein Tb927.10.7580 re-localizes to the nuclear pore in response to sinefungin treatment. Representative images of one untreated, one sinefungin-treated and one actinomycin D-treated cell are shown. **(F)** ESTN protein Tb927.10.7580 expressed as eYFP fusion protein from the endogenous locus (green), co-expressed with the nuclear pore marker protein NUP93-DHH1. **(G)** ESTN proteins that do not change their localization in response to sinefungin treatment. Representative images of one untreated and one sinefungin-treated cell are shown.

4.2.8 Conclusion

In summary, a high-quality purification method for trypanosome nuclei under conditions that maintain NPG attachment to the nucleus was established. Mass spectrometry analysis revealed 128 proteins that were significantly enriched in nuclei with attached NPGs (in comparison to nuclei without NPGs) – the NPG proteome. The majority of these proteins are involved in mRNA metabolism, unique to kinetoplastids and contain features that contribute to granule formation. NPGs consist almost entirely of RNA binding proteins, but in contrast to stress granules, lack canonical translation factors, which indicates that NPGs are RNP complexes having not yet entered translation. The localization of 70 ESTN proteins ±SF was either known from previous published studies (10) or determined by expressing eYFP fusions (60). 55 proteins moved from the cytoplasm to NPGs when trans-splicing was inhibited, therefore representing true NPG proteins. Furthermore, seven proteins moved from the cytoplasm to the nucleus and one from the nucleoplasm to nuclear pores at inhibition of trans-splicing. For these eight proteins, the changes in localization were specific to the accumulation of unspliced mRNAs and did not occur in the absence of newly transcribed mRNAs, making them strong candidates for nuclear export control proteins.

4.3 On the way to NPG function

4.3.1 Where do NPG protein components originate from?

Where are the protein components necessary for NPG formation recruited from? In other words, do NPG proteins originate from the nucleus or the cytoplasm? The answer to this long outstanding question about NPGs could give a first hint about their function. Various previous findings added up to a first idea how to answer this question.

(1) It was previously shown that NPGs depend on transcription but not translation[49]. Upon treatment with actinomycin D, which inhibits transcription in trypanosomes, NPGs did not form. *Vice versa*, once NPGs formed, they dissociated when transcription was inhibited. Hence NPGs are depended on newly synthesized mRNAs. In contrast, cycloheximide, which traps ribosomes in polysomes, preventing their disassembly and causing a decrease in P-bodies had no effect on NPG formation and did not dissolve them. Taken together, these data suggest that NPGs contain newly transcribed mRNAs (in the process of leaving the nucleus) and not mRNAs that have been previously in translation (originating from the cytoplasm).

(2) Another important evidence for NPGs representing a compartment originating from the nucleus and not the cytoplasm is provided by the NPG proteome, obtained in this work (see chapter4.2.5 to 4.2.7). The list of 128 ESTN proteins, of which the majority localizes to NPGs, does not contain canonical translation factors, indicating, that NPGs are RNP complexes that have started or completed nuclear export, but not yet entered translation.

(3) Short after the first report about NPGs, a detailed characterization of the two poly(A) binding proteins PABP1 and PABP2 revealed that both proteins localize to different sets of RNP granules in response to different stress conditions and to inhibition of trans-splicing. PABP2 co-localized with the P-body marker DHH1 into RNP granules with similarity to P-bodies, such as nuclear periphery granules and nutrient starvation stress granules, whereas PABP1 localized to heat shock induced stress granules only[84]. Most interestingly, when SF treatment was combined with heat shock, PABP2 accumulated in the nucleus, whereas PABP1 remained in the cytoplasm[84]. This combined treatment of cells most likely severely inhibits nuclear export of mRNAs because i) heat shock has been shown to inhibit nuclear export in yeast[131–133] and ii) unspliced mRNAs which accumulate after sinefungin treatment are likely impaired in nuclear export. This raises the possibility that PABP2 as a component of NPGs may bind to mRNA in the nucleus. PABP1, which is not a component of NPGs, may bind to its mRNA targets later in the cytoplasm[49] (and this work).

All these findings (1-3) let to the hypothesis that protein components of NPGs might originate from the nucleus rather than from the cytoplasm and bind their mRNA targets in the nucleus. With a subset of the newly identified NPG proteins from this work it was therefore tested whether they re-localize to the nucleus upon heat shock and SF treatment. If such re-localization to the nucleus would be a hallmark for all NPG proteins, and absent from all non-NPG proteins this would be very strong evidence for NPGs originating from nuclear components.

Six randomly chosen proteins from the list of 55 validated NPG proteins were expressed as eYFP fusions from their endogenous locus[73] in a cell line also expressing mChFP-DHH1 (p2845). Proteins were analyzed for their subcellular localization in untreated, SF treated, heat shocked or heat shocked + SF treated cells (Figure 4.15). As shown above (Figure 4.14 B), all six proteins showed cytoplasmic localization in untreated cells and moved to NPGs in the presence of SF. Upon heat shock, only one protein of the six, namely DRBD2 localized to cytoplasmic stress granules. DRBD2 was the only protein with unequivocal nuclear localization after combined heat shock and SF treatment. Two further proteins (NOT2, PRMT7) had partial localization to the nucleus and three remained in the cytoplasm (Tb927.10.14700, Tb927.4.4940, ZC3H39). The experiments were discontinued after these results, because they could neither provide proof nor disproof for proteins being of nuclear origin. The reason for this is that the effect that heat shock and sinefungin and the combination of the two drugs have on nuclear export is unknown and there is no simple way to investigate these most likely complex effects. Furthermore, the fact that two proteins (Tb927.4.4940, ZC3H39) could still localize to NPGs after a combined treatment with heat shock and SF is interesting, but for the moment remains unresolved.

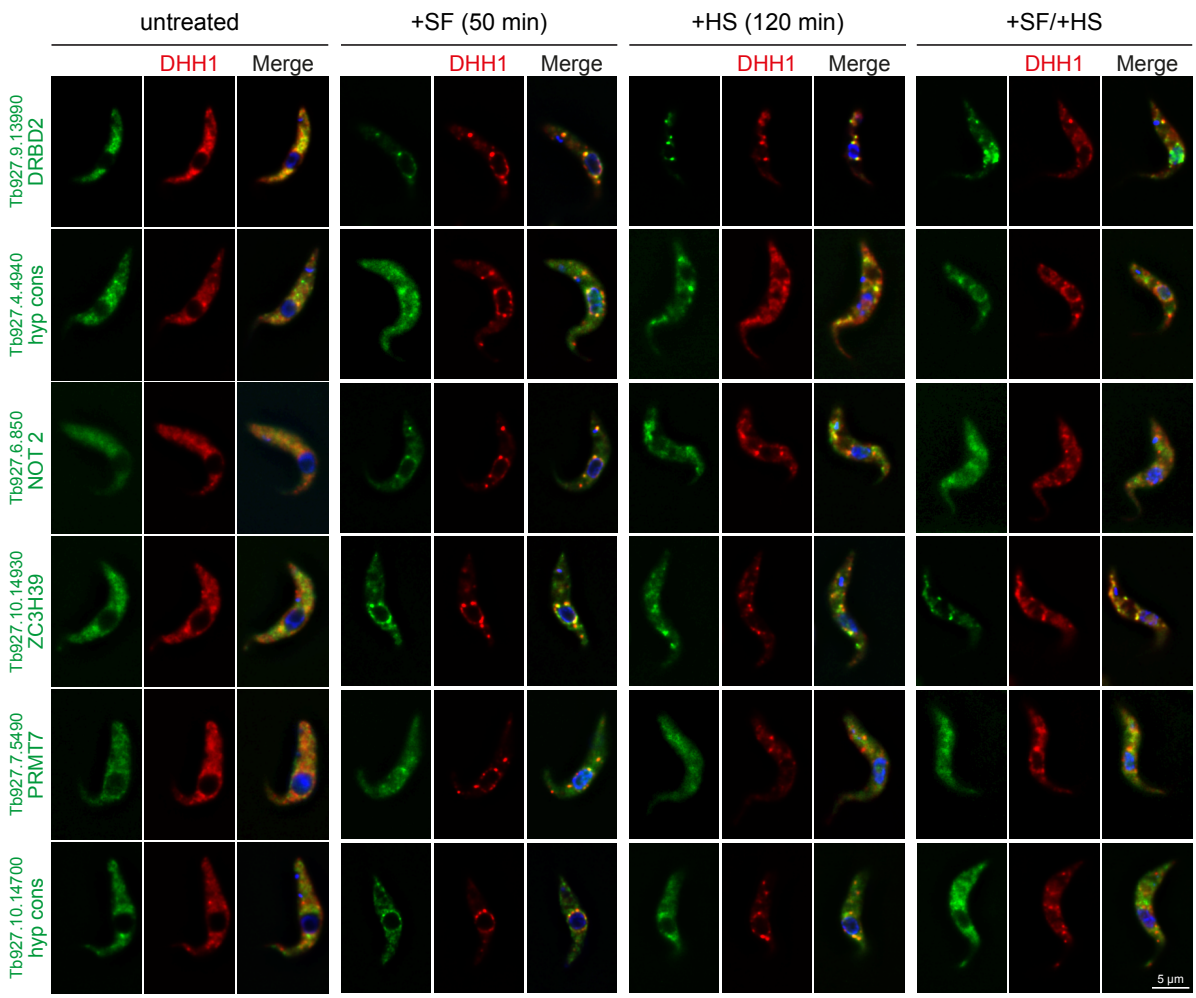


Figure 4.15: Recruitment of NPG proteins.

ESTN proteins were expressed as eYFP fusion proteins from the endogenous locus (green) in a cell line co-expressing mChFP-DHH1 (red) as NPG granule marker. Representative images of one untreated, one sinefungin-treated, one heat-shocked and one heat-shocked plus sinefungin-treated cell are shown. DNA was stained with DAPI (blue). Images are shown as single plane images of a deconvolved z-stack.

4.3.2 Are NPGs involved in quality control of unspliced mRNA?

Based on similarities to perinuclear germ granules in *C. elegans*, it was previously suggested that NPGs function as a quality control compartment, preventing entry of unspliced mRNA into translation[49]. If this hypothesis is correct, unspliced mRNAs should accumulate in NPGs.

In search of a reliable and specific method to visualize single mRNA molecules, our group recently adapted the *QuantiGene ViewRNA ISH Cell Assay Kit* (Affymetrix, Santa Clara) for usage in trypanosomes[34] (for more details see chapter 3.6). With this method, up to 20 pairs of adjacent antisense oligos are hybridized to a target mRNA. Afterwards the signal gets amplified by branched DNA technology[134,135], which allows the detection and quantification of single mRNA molecules with great specificity and low background. To visualize unspliced mRNAs, a single molecule mRNA FISH probe antisense to the intergenic region in between α -tubulin and β -tubulin was designed (AF4, beta to alpha tubulin IR). This intergenic region is only 158 nt long and can therefore be probed with only three oligo pairs instead of 20. The signal proved still sufficient for detection. Initial tests to implement Affymetrix FISH on ultrathin EM sections were not further pursued due to technical problems (the total amount of detected unspliced tubulin precursor mRNA molecules did not change between untreated and SF treated cells).

Untreated and SF treated wild-type trypanosome cells were probed for beta to alpha tubulin IR by Affymetrix FISH, stained with DAPI and analyzed microscopically (Figure 4.16 A). The amount of unspliced tubulin precursor mRNA molecules per cell was determined. As expected, there was a significant increase in tubulin precursor mRNAs upon SF treatment from in average 1.2±1.0 molecules per cell to 8.1±3.2 ($N>200$, result of unpaired, two-tailed students t-TEST$<10^{-106}$) (Figure 4.16 B). With DAPI as nuclear marker, there was no obvious accumulation of mRNA molecules at the nuclear periphery of SF treated cells (data not shown). Moreover, about half of all mRNA molecules were unequivocally cytoplasmic (e.g. distant to the nucleus), independent on whether cells were untreated or treated with SF.

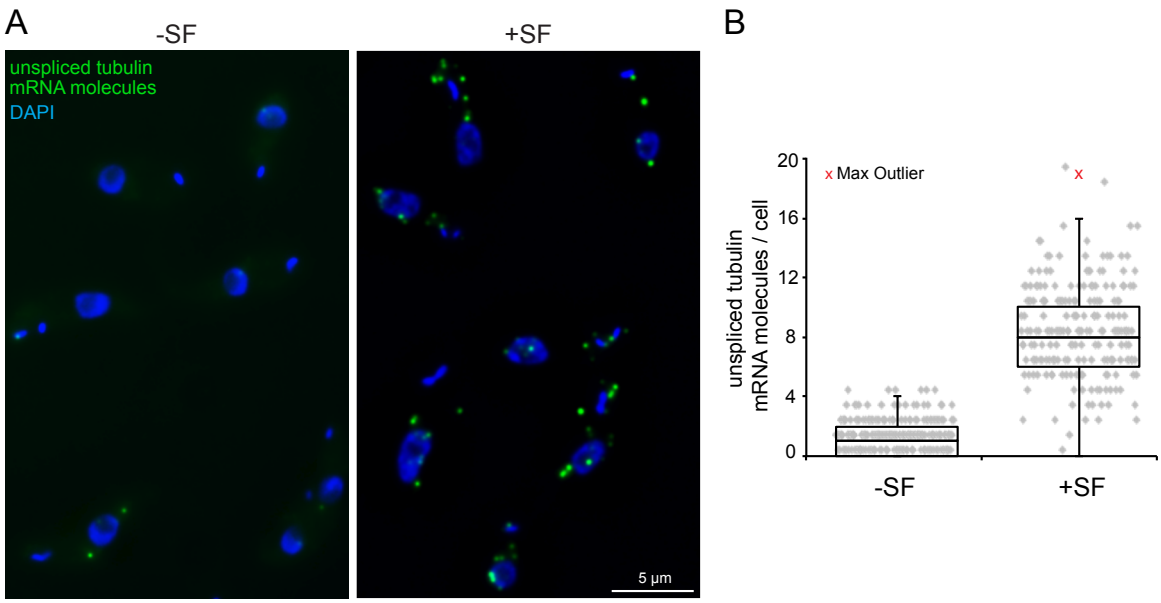


Figure 4.16: Quantification of unspliced mRNAs.

(A) Unspliced tubulin mRNA molecules in untreated and sinefungin-treated cells. Unspliced tubulin mRNA molecules are detected by single molecule mRNA FISH using probes antisense to the intergenic region between β-tubulin and α-tubulin (green spots). Nuclei were stained with DAPI (blue). Z-stack projection images are shown of untreated cells (-SF) and sinefungin-treated cells (+SF). **(B)** The number of unspliced tubulin mRNA molecules per cell for untreated (-SF) and sinefungin-treated (+SF) cells was quantified. Results are shown as box plot (waist is median; box is interquartile range (IQR); whiskers are ±1.5 IQR; only the largest outliers are shown; n >200).

To more precisely examine a possible co-localization of unspliced mRNAs and NPGs, Affymetrix FISH was combined with antibody staining of the NPG marker protein DHH1 (anti-DHH1). Additionally, the use of a cell line expression NUP93-eYFP-4xTy1 from the endogenous locus (SK362) enabled the staining of nuclear pores with the anti-Ty1 BB2 antibody. This way a clear border between nucleus and cytoplasm could be defined and signals of single mRNA molecules at this border could be more accurately assigned to either the nucleus or the cytoplasm. SF treated cells were probed for beta to alpha tubulin IR by Affymetrix FISH and stained with antibodies against DHH1 and NUP93-Ty. On average, 4.9 unspliced tubulin mRNA molecules were detected in each cell (N=114). The majority of all mRNA molecules (54%) were located in the cytoplasm. Of these, 39% were not in proximity to the nucleus, 11% were adjacent to the nucleus with no overlap to NPGs and only 4% showed overlap with NPGs (Figure 4.17 A and B).

By single molecule mRNA FISH, it was revealed that, in agreement with earlier observations[49,68], about half of all unspliced mRNAs is able to reach the cytoplasm, indicating that any control system is not be perfect. However, these cytoplasmic, unspliced mRNAs do not accumulate in NPGs, arguing against a function of the majority of granules in mRNA quality control.

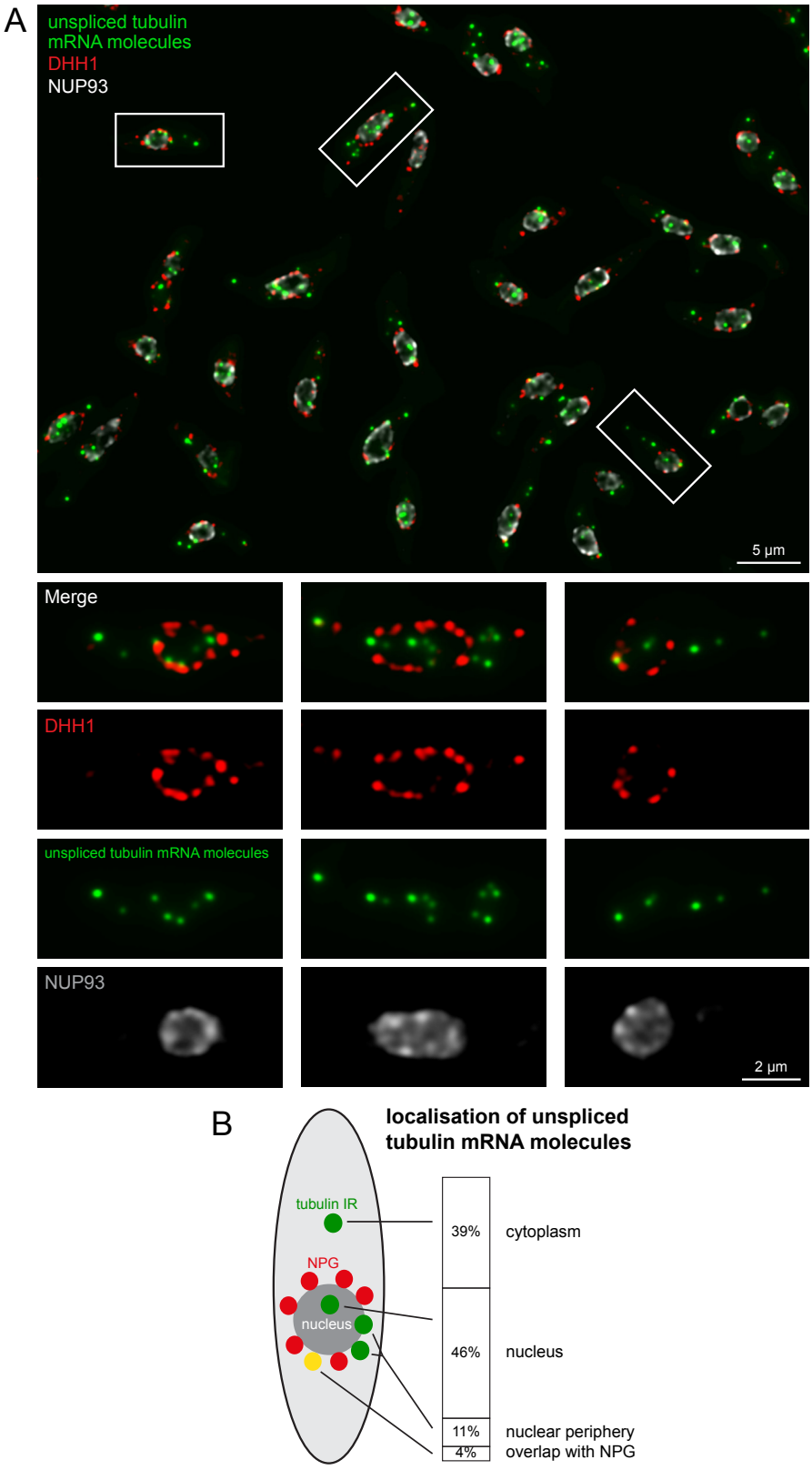


Figure 4.17: Subcellular localization of unspliced mRNAs.

(A) Localization of unspliced tubulin mRNA molecules in sinefungin-treated cells. Unspliced tubulin mRNA molecules are detected as above (green fluorescence). In addition, the NPG marker protein DHH1 is detected by anti-DHH1 (shown in red). The nuclear pore marker protein NUP93 is detected by anti-Ty1 BB2 (shown in white). Insets of three selected cells are shown as merged images as well as individual channels. **(B)** The intracellular localization of unspliced tubulin mRNA molecules to the cytoplasm, nucleus, nuclear periphery or NPGs was determined from 114 sinefungin-treated cells.

4.3.3 Are NPGs related to perinuclear germ granules (P granules) of *C. elegans*?

Due to similarities in localization, sensitivity to actinomycin D, insensitivity to cycloheximide and the presence of the germ granule marker VASA[49], NPGs were previously suggested to be related to perinuclear germ granules (P granules) in *C. elegans*[86]. Actual proof of such a relation, combined with new findings about NPGs from this work, may help to understand the function of NPGs.

4.3.3.1 Some *C. elegans* P granule proteins localize to NPGs in trypanosomes

The question of a potential relationship between NPGs and P granules was first addressed by expressing known P granule proteins in trypanosomes and test, weather these localize to NPGs. Based on a list published by Strome et al. in 2005[136], four *C. elegans* P granule proteins were chosen for expression in *T. brucei*: CGH-1, PGL-2, OMA-2 and POS-1. These proteins were selected according to the following criteria: they are expressed in adult worms, have an appropriate sequence length for the applied cloning strategy, as well as suitable restriction sides. cDNA was synthesized from RNA extracted from *C. elegans* worms kindly provided by Sebastian Markert (University of Würzburg, Germany). Subsequently, DNA fragments were amplified with primers, specifically designed for each gene (Sup. Table 7.1) and cloned as described in chapter 3.4.2. Utilizing the system for inducible expression of transgenes in trypanosomes[69], the generated constructs (Sup. Table 7.2) were transfected into a PCF cell line co-expressing TetR (PSPR2.1) and mChFP-DHH1 (p3925). This allowed the tetracycline-inducible, transgenic expression of the four *C. elegans* P granule proteins as eYFP fusion proteins in trypanosomes. All *C. elegans* proteins were analyzed for their subcellular localization in untreated, SF treated, and PBS treated trypanosome cells (Figure 4.18). The marker protein mChFP-DHH1 served as reference for granule formation. The *C. elegans* proteins POS-1 (no homology to NPG proteins) and CGH-1 (homology to NPG protein DHH1) unequivocally localized to NPGs upon treatment with SF and to starvation stress granules upon PBS treatment. PGL-2 and OMA-2 (both no homology to NPG proteins) did not change their localization with those treatments. The localization of a unique *C. elegans* P granule protein to NPGs might indicate a relation between these two granule types. However, one could argue if *C. elegans* P granule proteins contain the same classical RNA binding domains as *T. brucei* granule proteins, localization to NPGs (or stress granules) may be mediated solely through this RNA binding feature. Indeed, both *C. elegans* CGH-1 and its *T. brucei* homologue DHH1 have a DEAD/H RNA helicase motive and POS-1 has two CCCH type zinc finger domains[136], a classical RNA binding motif present in many trypanosome proteins[119]. *Vice versa*, the absence of a *C. elegans*

protein from NPGs does not exclude that the granules are related, as we cannot test whether the transgenic protein is functional (properly folded etc.). Thus, data from experiments based on transgenic expression of P granule or NPG proteins in trypanosomes or *C. elegans* have only limited values. The technically more challenging *vice versa* experiments, the expression of trypanosome NPG proteins in *C. elegans*, were therefore not performed.

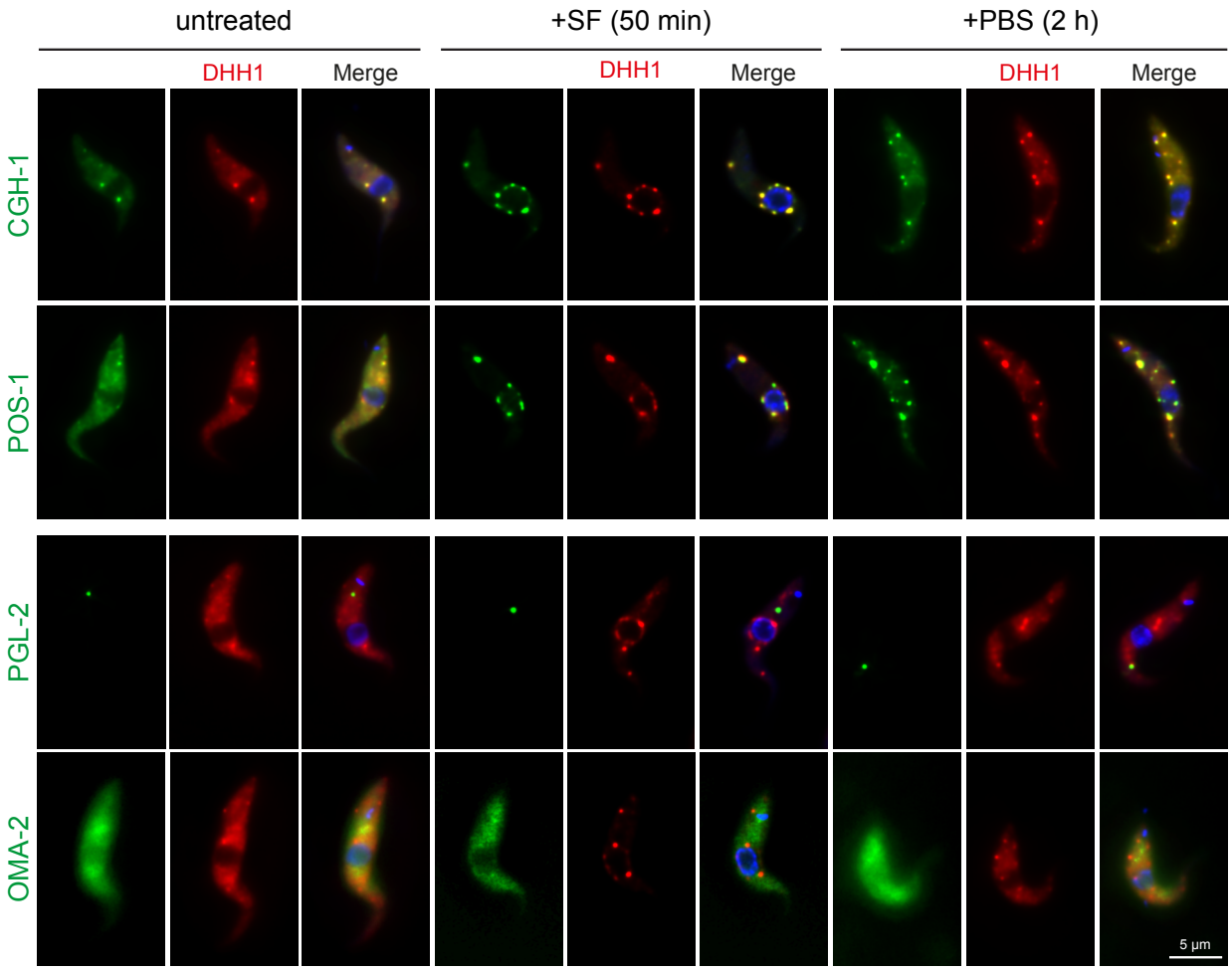


Figure 4.18: Localization of *C. elegans* P granule proteins in *T. brucei*.

C. elegans P granule proteins were inducible expressed as e-YFP fusions in trypanosomes (green) in a cell line co-expressing mChFP-DHH1 (red) as granule marker. Proteins were analyzed for their subcellular localization in untreated, SF treated, and PBS treated cells. DNA was stained with DAPI (blue). Images are either single plane images (+SF) of a deconvolved z-stack or z-stack projections (untreated, +PBS).

4.3.3.2 No homologies between NPG and P granule protein content

Another approach to determine a possible relationship between NPGs and P granules was to analyze similarities in their protein content. Previous homology studies between those two granule types were based on only nine known NPG proteins[49]. Now, with a comprehensive list of 55 NPG proteins available, a search for homologues will allow a much more precise statement. In theory, this analysis could also open the possibility to identify new P granule protein candidates in *C. elegans*. Such proteins would have to (1) be unequivocally validated as NPG proteins (eYFP tagging), (2) possess a *C. elegans* orthologue (BLAST analysis), (3) be unknown as P granule proteins and (4) ideally be expressed within the germline of *C. elegans*. The selected proteins could then be tagged fluorescently and expressed in *C. elegans* to analyze their localization.

For this, the 55 validated NPG candidates were searched for homologous proteins in *C. elegans* (compare also chapter 4.2.6.3). The majority of validated NPG proteins had no homologues in *C. elegans*; only 13/55 had homologues in *C. elegans* (Sup. Table 7.8 A). Of these, only four were reported to be in P granules, namely LAF-1 (a putative DEAD-box RNA helicase related to VASA/DBP1), CCF-1 (homologue to TbCAF1, part of the CCR4/NOT complex and a P-body component in *C. elegans*), CAR-1 (homologue to TbSCD6, *C. elegans* P-body component) and CGH-1 (homologue to TbDHH1, putative DEAD-box RNA helicase, stress granule and P-body component in *C. elegans*) (Figure 4.19 green)[137–141].

Two further proteins are very likely P granule proteins. First, IFE-3 (homologue to TbeIF4E1) is one of five *C. elegans* homologues of the mRNA cap-binding protein eIF4E. The proven P granule localization of IFE-1, another eIF4E homologue with similar properties, indicates a potential P granule function[142]. Second, ATX-2 (homologue to PABP1-binding protein PBP1 in *T. brucei*) interacts with the two P granule core proteins PAB-1 and CGH-1[143] and could thus also be a P granule component (Figure 4.19 orange). However, all six proteins are rather general RNA metabolism proteins with multiple functions and are not specific for perinuclear germ granules.

For the remaining seven proteins, no indications for a P granules localization is published so far (Figure 4.19 red). All of these proteins fulfilled at least the first three (some of them even all four) characteristics that were defined above for identification of potentially new P granule protein candidates in *C. elegans*. They may be worth looking into in more detail, regarding general *C. elegans* and P granule research. Still, in total only six of the 55 NPG proteins had a homologue in *C. elegans* with proven or possible P granule localization.

Vice versa, 31 characterized P granule proteins of *C. elegans*[136,144] were screened by BLAST for potential homologues in *T. brucei* (Sup. Table 7.8 B), with essentially the same outcome. Only 7/31 P granule proteins had a homologue in *T. brucei* with proven localization to NPGs. Overall, the data do not support a compositional relationship between NPGs and germ granules.

		T. brucei		*C. elegans*		
		ID	Description	ID	Description	
evidence for p-granule localisation	YES	Tb927.10.14550	VASA/DBP1	NP_001254858.1	LAF-1	* Hubert & Anderson (2009)
		Tb927.6.600	CAF1	NP_499553.1	CCF-1	* Gallo et al (2008)
		Tb927.11.550	SCD6	CAA22317.1	CAR-1	* Boag et al. (2005), Audhya et al. (2005)
		Tb927.10.3990	DHH1	NP_498646.1	CGH-1	* Navarro et al. (2001)
	Possible*	Tb927.11.2260	eIF4E1	NP_741502.1	IFE-3	* Amiri et al. (2001)
		Tb927.8.4540	PABP1-binding protein PBP1	NP_001021230.1	ATX-2	* Boag et al. (2008)
	NO	Tb927.10.1510	NOT1	AGT18675.1	NTL-1b/ LET-711	
		Tb927.3.1920	NOT5	NP_001076652.1	NTL-3	
		Tb927.4.410	CAF40	NP_498048.2	NTL-9	
		Tb927.7.5490	arginine N-methyltransferase, type III (PRMT7)	3WST_A	Prmt7	
		Tb927.5.2140	UPF1	NP_490829.1	SMG-2	
		Tb927.9.13990	DRBD2	NP_493023.1	R09B3.3	
		Tb927.7.4900	XRNA	NP_496945.3	XRN-1	

Figure 4.19: Comparing protein composition between NPGs and P granules.

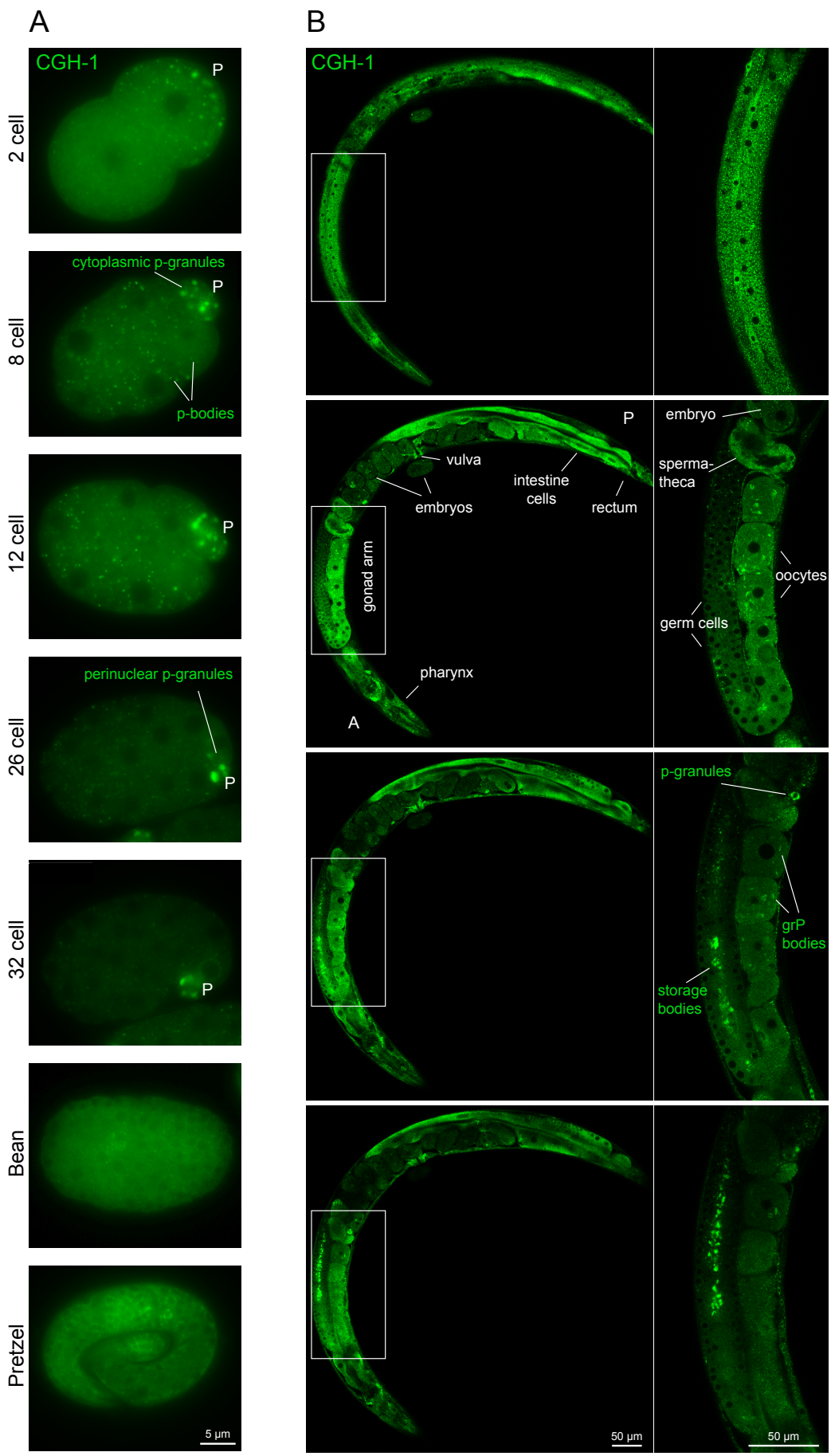
A BLAST screen revealed homologies between 13 validated trypanosomal NPG proteins and proteins in *C. elegans* (taxid6239). The identified proteins were classified for a localization to *C. elegans* P granules (yes, possible, no) by literature and experimental evidence (indicated by *). IDs displayed for *C. elegans* proteins are the IDs provided by the BLAST search. For details on the BLAST analysis and further literature evidence see Sup. Table 7.7.

4.3.3.3 Inhibition of splicing does not cause the formation of perinuclear granules in *C. elegans*

To investigate the possibility that other organisms may have functional counterparts to NPGs (even though these may not be compositionally related, see above), one essential question needed to be answered: Does inhibition of splicing cause the formation of granules at the nuclear periphery also in other organisms? Previously, this was already tested in HeLa cells[49]. There, the RNA helicase p54 (the human orthologue to DHH1), served as granule marker protein[145]. Splicing was inhibited by the addition of spliceostatin A, which blocks splicing in HeLa cells by binding to the splicing factor SF3B[146]. In a second independent approach, splicing was inhibited by a morpholino oligonucleotide antisense to the U2 snRNA. In both cases, the successful inhibition of splicing did not cause the formation of NPGs[49]. Since in trypanosomes all mRNAs (with only a few exceptions) are trans-spliced, whereas in humans only cis-splicing occurs, it was necessary to exclude that granule formation depended on the

inhibition of trans-splicing rather than cis-splicing. Therefore, the effects of splicing inhibition were analyzed in *C. elegans*, where about 70% of all transcripts are trans-spliced to one of two spliced leader RNAs[147]. It was described by Shiimori et al. 2012[148] that the *C. elegans* spliceosome can also be inhibited with spliceostatin A.

P granules are specific to the germ line of *C. elegans* and granule composition changes, when cells develop from germ cells to oocytes and embryos. CGH-1 (the *C. elegans* DHH1 homologue) is also a well-known *C. elegans* granule protein. While CGH-1 is a main component of P-bodies, stress granules and germ line storage bodies[149] it is also present in P granules of the early to late embryo. Within the embryo's P-cell, CGH-1 is first present in large cytoplasmic granules (1-cell to 12-cell stage) and at later developmental stages relocates around the nuclear periphery[139,141]. CGH-1 is not only expressed within the germline, but also in somatic cells of the adult animal, which makes it the ideal marker protein to analyze a potential formation of perinuclear granules in somatic cells. Worms expressing CGH-1 fused to GFP (WEH172) were generated in cooperation with Ann Wehman (Rudolf-Virchow-Centre, Würzburg, Germany), prepared and microscopically analyzed as described in chapter 3.3. Different embryonic developmental stages (Figure 4.20 A), and young adult worms (Figure 4.20 B) were analyzed for CGH-1 expression.



(caption on next page)

Figure 4.20: Localization of CGH-1 in *C. elegans*.

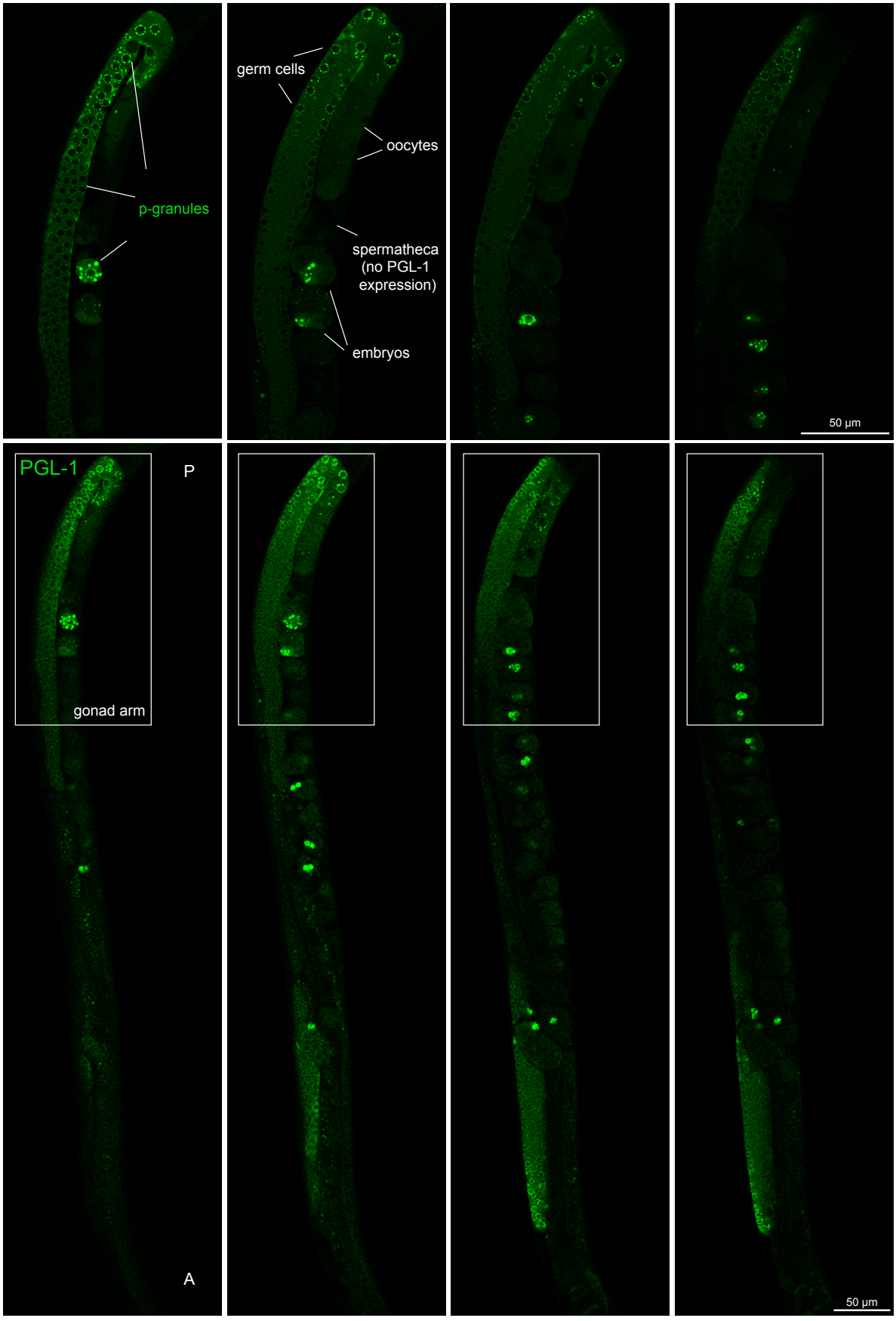
Expression of the P-body factor CGH-1 in WEH172 embryos **(A)** and young adults **(B)**. Treatment with spliceostatin A caused no change in the subcellular localization of CGH-1, thus only images of untreated worms are displayed for illustration. **(A)** Embryos at different developmental stages. CGH-1 is highly expressed within the P-cell (germ cell blastomere) (indicated by "P"), where in early stages of cell division P granules are cytoplasmic and become perinuclear at about the 12-cell stage. CGH-1 is also expressed within all other cell types, where it forms cytoplasmic P-bodies. Single plane images are shown. **(B)** Image sections of an adult worm. CGH-1 is expressed in all cell types and forms P-bodies (somatic cells), so called 'storage bodies' (gonad arm), grP bodies (arrested oocytes) or P granules (embryo). Four single plane images at different focal layers of a z-stack, originally consisting of 52 stacks are shown. Insets (right panel) of one gonad arm conducted at higher magnification are displayed. Note that bacteria (serving as food) inside the worm's intestine display auto-fluorescence, hence intestine cells appear as bright green. The anterior end of the worm is marked with (A), the posterior end with (P).

Young adult worms were treated with spliceostatin A to inhibit splicing and compared with untreated animals. To control the inhibition of splicing, RNA from both conditions was isolated and cDNA was analyzed by PCR as described in Shiimori et al. 2012[148]. However, primers designed to cover the tra-2 sequence from exon 21 to exon 23 as well as a primer pair covering exon 7 (the largest exon within the sequence) failed and no clear results were obtained due to technical problems (for details on primer see Sup. Table 7.1).

Since embryos form an eggshell after fertilization, which shields them from their surroundings, they lose the ability of drug uptake via skin and are thus most likely unaffected by any drug treatment. Figure 4.21 B only shows images of untreated animals, since in four independent experiments, no obvious localization of the germ granule marker proteins CGH-1 to perinuclear granules in gonads or somatic cells of adult worms (pharynx, intestine, vulva) was observed. CGH-1 remained cytoplasmic or in granules distant to the nucleus upon treatment with spliceostatin A. In some cell types, P-bodies or stress dependent storage bodies were enlarged (data not shown).

In a second experiment, animals expressing PGL-1 fused to GFP (JJ2101) were analyzed. PGL-1 is another granule marker protein, which is highly specific for P granules and is only expressed within the germline of *C. elegans*[150,151]. Again, upon treatment with spliceostatin A no obvious changes in localization were observed.

To summarize, neither bioinformatical nor experimental approaches could identify any relationship between NPGs of *T. brucei* and germ granules of *C. elegans*. Taken together, the data indicate that NPGs may be a unique phenomenon in kinetoplastids.



(caption on next page)

Figure 4.21: Localization of PGL-1 in *C. elegans*.

Expression of the P granule core component PGL-1 in JJ2101 young adult worms. Treatment with spliceostatin A caused no change in PGL-1 localization, thus only images of untreated worms are displayed. PGL-1 is solely expressed within the worm's germ line (germ cells, oocytes, embryos) where it forms P granules. P granules are perinuclear for the majority of the life cycle and only dissociate in the oocyte. They remain cytoplasmic until around the 12- cell embryo stage. Four single plane images at different focal layers of a z-stack, originally consisting of 68 stacks are shown. Insets (upper panel) of one gonad arm conducted at higher magnification are displayed. The anterior end of the worm is marked with (A), the posterior end with (P).

4.3.4 Proteins re-localizing to nucleus/nuclear pores upon sinefungin treatment (RNST) may function in nuclear export control

Within the group of ESTN proteins that changed their localization upon treatment with SF, six proteins moved from the cytoplasm to the nucleus and one protein moved from the nucleoplasm to nuclear pores. These proteins were named RNST proteins (proteins re-localizing to nucleus/nuclear pores upon sinefungin treatment) (compare Figure 4.14 A). When transcription was inhibited with actinomycin D, this re-localization did not occur (Figure 4.14 D), which means that the phenomenon is solely based on the accumulation of incompletely processed mRNAs and not on the absence of fully processed mRNAs.

This re-localization of proteins to the nucleus upon SF treatment can potentially be caused in two ways: either, a protein is actively transported to the nucleus or, it is somehow prevented from leaving the nucleus. At least in the latter case, the protein would then need to be shuttling between the cytoplasm and the nucleus. In trypanosomes such a shuttling behavior was already discovered for one of the RNST proteins, namely DRBD3[152], but also for many non-RNST proteins, for example PABP2[84] or UBP1[153]. Shuttling proteins are not only thought to play an important role in regulating gene expression, but they are also strong candidates for a function in nuclear export control. Since only little is known about nuclear export control in trypanosomes, finding new key players is important.

4.3.5 RNAi depletion of Mex67 phenocopies the effect of sinefungin treatment on RNST proteins: re-localization to the nucleus

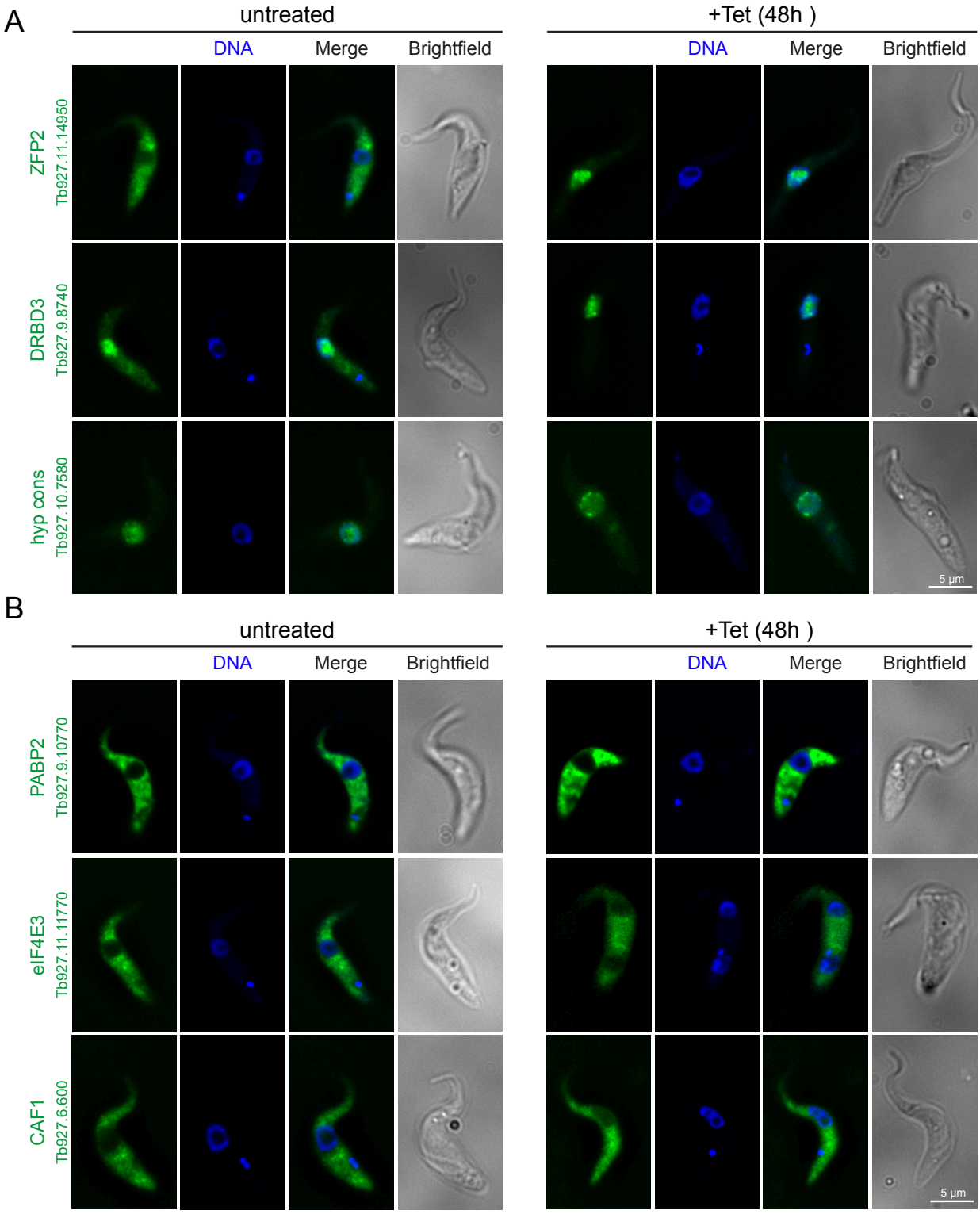
To investigate whether re-localization of RNST proteins to the nucleus upon SF treatment is caused by active transport or nuclear retention, a method to block nuclear export was needed. If this block in nuclear export would then results in an accumulation of the RNST protein inside the nucleus, this would be a clear indication that (1) the protein shuttles and (2) is prevented from leaving the nucleus rather than being actively transported into the nucleus.

One of the known key players involved in nuclear export of mRNAs in trypanosomes is the putative homologue to Mex67 in yeast. TbMex67 (Tb927.11.2370) localizes to the nucleoplasm and to spots that resemble nuclear pores[119]. The down-regulation of TbMex67 inhibits cell growth and causes an accumulation of polyadenylated mRNA within the nucleus[63]. Thus, an essential role of Mex67 for cell growth and nuclear mRNA export was suggested[62]. Recently, a *T. cruzi* RNA helicase named Hel45 (TcCLB.506587.40) (actually annotated as FAL1 at TriTrypDB[80]), whose shuttling properties depend on a nuclear export route involving the Mex67 receptor, was discovered[154]. This conclusion was based on the fact that RNAi against Mex67 caused the *T. brucei* homologue (Tb927.11.8770) of Hel45/Fal1 to re-localize from the cytoplasm to the nucleus together with polyadenylated mRNA.

So, I asked: does RNAi against Mex67 specifically cause RNST to re-localize to the nucleus and thus reproduces the effect of SF treatment? This would suggest that RNST proteins have indeed shuttling properties and get trapped inside the nucleus either due to a blocked nuclear export (Mex67 RNAi) or due to the accumulation of unspliced precursor mRNAs (SF treatment) while they are bound to their target mRNA. If other ESTN proteins (control proteins) still can leave the nucleus, this would indicate that RNST have either a function in nuclear export control by preventing defective mRNA from nuclear export, or they are involved in the same export pathway as Mex67/ are co-factors of Mex67 depended export.

To address this idea, six RNST proteins (ZFP1, ZFP2, RBP12, ZC3H29, DRBD3 and Tb927.11.6600) and six control proteins (PABP2, DHH1, eIF4E3, CAF1, NOT1 and UPB2) were expressed as eYFP fusions[73] in an inducible Mex67 RNAi cell line (3896/PSPR2.1) (for RNST proteins and UBP2 method by Dean et al. 2015 was used, for other control proteins SK252, SK85, p3935, p4062 and p3842 were used. Details see chapter 3.4.3 and Sup. Table 7.2). The localization of each protein was microscopically analyzed in uninduced cells and cells after 48h of tetracycline induced Mex67 RNAi. Up to the point of writing this thesis, convincing fluorescent images of 2/6 RNST proteins and 3/6 control proteins were obtained during a first screen. Supporting the previously proposed hypothesis, indeed both RNST proteins (ZFP2 and DRBD3) re-localized from the cytoplasm to the nucleus upon Mex67 RNAi (Figure 4.22 A) indicating a shuttling behavior, whereas none of the three control proteins (PABP2, eIF4E3 and CAF1) changed their localization (Figure 4.22 B). Interestingly, PABP2 has a known shuttling behavior[155], yet does not re-localize to the nucleus when nuclear export of mRNA is blocked. Furthermore, the seventh RNST protein that changed localization from the nucleoplasm to nuclear pores (Tb927.10.7580) upon SF treatment was analyzed (compare chapter 4.2.7). Similar to the RNST proteins ZFP2 and DRBD3, RNAi against Mex67 phenocopied the effect of SF treatment (Figure 4.22 A).

Hence, this behavior is specific to the RNST proteins and not a general feature of all shuttling proteins. It is tempting to speculate that RNST proteins are novel regulators involved in nuclear export control of unspliced transcripts. For an unambiguous statement, multiple clones of each cell line need to be analyzed, as well as cell lines of the remaining ESTN and control proteins.



(caption on next page)

Figure 4.22: Effect of Mex67 RNAi on RNST protein localization.

RNST proteins **(A)** and control proteins **(B)** were expressed as eYFP fusions from the endogenous locus (green) in a Mex67 RNAi cell line. The subcellular localization of proteins was determined in uninduced cells and in cells induced for Mex67 RNAi (+Tet 48h). DNA was stained with DAPI (blue). Fluorescence images are shown as single plane images of a deconvolved z-stack. Brightfield images are single plane images.

4.3.6 Effects of RNAi depletion of RNST proteins on cell growth

It was previously reported that RNAi of the nuclear export factor Mex67 resulted in a severe growth defect and caused the accumulation of polyadenylated mRNA within the nucleus[62,63]. So, to investigate a possible role of the RNST proteins (Tb927.10.7580, ZFP1, ZFP2, RBP12, ZC3H29, Tb927.11.6600 and DRBD3) in nuclear export control, cells down-regulated for the respective protein were (1) examined for a growth defect and (2) analyzed by RNA-FISH. Proteins were targeted with a tetracycline inducible RNAi system, using the expression vector p3666[69] (details see chapter 3.4.3, used plasmids see Sup. Table 7.2). Constructs were transfected into a PCF cell line expressing the tetracycline repressor TetR (PSPR2.1) (growth curve shown in Figure 4.23 A); the growth phenotype of Mex67 RNAi cells served as a reference for all further experiments (Figure 4.23 A). On average, resistant clonal cell lines were obtained 12 days after the addition of selection and were subsequently analyzed for their growth rate under uninduced and RNAi induced conditions.

For 5/7 RNST proteins, a down-regulation by RNAi caused no obvious growth defect, only 2/7 proteins (RBP12, DRBD3) showed a minor reduction in growth after 48 of RNAi (Figure 4.23 B). A severe growth defect similar to the one seen after Mex67 RNAi (p3896) starting as early as 24 h (Figure 4.23 A) was not observed for any of the RNST proteins. However, a subpopulation of the cells with down-regulated RBP12 had morphological changes that were very similar to the morphological changes reported in cells with down-regulated Mex67[63]. Cells had enlarged nuclei and in general a more short and roundish shape, accompanied by a reduced motility (starting at 48h after tetracycline induction) (data not shown).

A nuclear retention of polyadenylated mRNAs as seen in cells induced for Mex67 RNAi [62,63], would point to a role of RNST proteins in nuclear export control of mRNA. In a first screen, cells with down-regulated RBP12, DRBD3 or Tb927.11.6600 were analyzed by RNA-FISH. Cells with down-regulated Mex67 served as a positive control. Total mRNA was labelled with a fluorescent probe antisense to the mini-exon (ME) (SKA48), which is a sequence present at the 5'end of all mature mRNAs in trypanosomes. Cells were analyzed 0 h, 24 h, 48 h and 72 h after induction of RNAi. The previously described shift of mRNA from the cytoplasm (0 h) to the nucleus (48 h, 72 h) could only be seen for Mex67 RNAi. For all other proteins, a down-regulation caused no such effect (data not shown). An

essential role of RNST proteins in cell metabolism and especially nuclear export control could not be demonstrated with this approach. However, due to time-constraints, these experiments were only performed once, and the RNAi experiments have not yet been controlled for successful reduction in RNA/ protein levels. The results should therefore be considered preliminary.

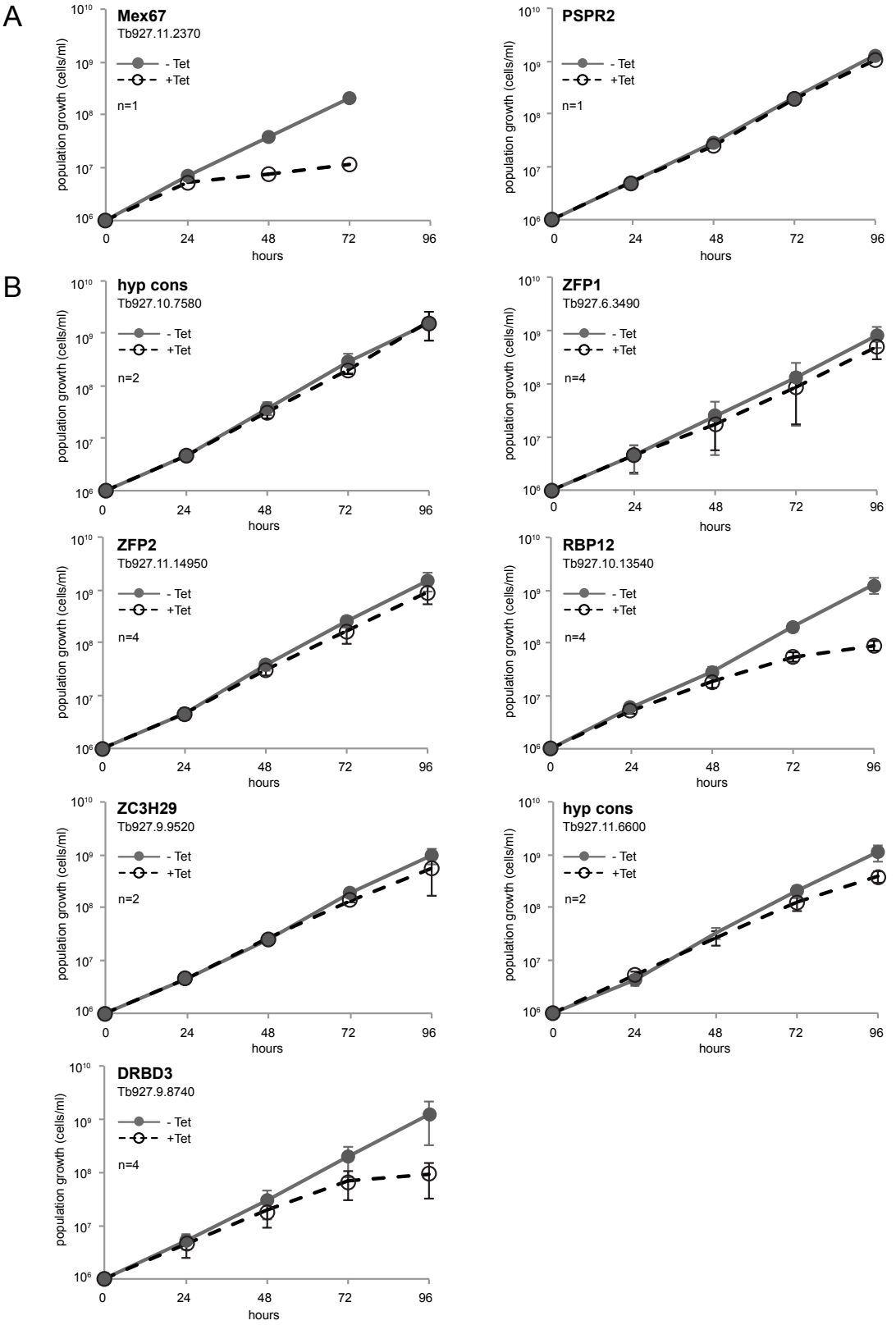


Figure 4.23: Effect of RNAi against RNST proteins on cell growth.

Growth curves of cells induced for RNAi (+Tet) and uninduced cells (-Tet). Cell numbers were determined every 24 h with a Neubauer chamber. **(A)** Growth curves of Mex67 and the parental cell line (PSPR2) served as control. Cell numbers of one clone are displayed (n=1). **(B)** Growth curves of RNST proteins. The mean, cumulative cell numbers of two or four independent clones (n=2; n=4) are displayed (±SD). In some cases, error bars are not visible due to small standard deviations.

5 Discussion

5.1 Discovering the NPG proteome allows first insights into a function of NPGs

Until quite recently, RNP granules were thought to possess liquid-like physical properties[156]. This implicated instability would make a purification of these structures almost impossible. Nevertheless, our group accomplished to purify stress granules from trypanosomes[34] while at the same time stress granules were purified from yeast and mammals[35]. In trypanosomes, the inhibition of trans-splicing by SF causes the formation of a special RNP granule type outside the nuclear periphery: nuclear periphery granules (NPGs)[49]. The recent breakthrough in purifying granules encouraged us to also develop a purification method for NPGs. This would not only enable us to draw conclusions about the granule's function, but also provide new insight into general RNP granule biology and specifically into global changes of mRNA metabolism proteins upon inhibition of trans-splicing.

The tight bond between NPGs and nuclei in the absence of detergent allowed the co-purification of both compartments. Furthermore, the ability to specifically induce NPGs (+SF) provided the option of a plus granule and minus granule state. Hence, the established purification method not only allowed resolving the NPG proteome (comparing purified nuclei of untreated and SF treated cells), but also to obtain the long wanted nuclear proteome (comparing whole cell lysates and purified nuclei of untreated cells) of *T. brucei*[91].

The proteome of the *T. brucei* nucleus was comprised of 764 proteins and proved to be about 80% complete, while containing less than 2% non-nuclear proteins. Moreover, six proteins with previously unknown localization were identified as nuclear proteins by expressing eYFP fusion proteins[91]. The data will be of use to unveil nuclear processes in trypanosomes and may complement the current efforts of TrypTag.org.

The comprehensive nuclear proteome data were a first indicator for the purity and high quality of the purified nuclei. Likewise, the quality of the NPG proteome was high: all nine previously known NPG proteins were identified. The NPG proteome was comprised of 128 ESTN proteins (enriched in sinefungin-treated nuclei), which (1) were predominantly involved in mRNA metabolism, (2) mostly unique to kinetoplastids, (3) contained many of all known trypanosomal cytoplasmic RNA binding proteins (Pumilio, CCCH and RRM domain,) and RNA helicases, but lacked canonical translation factors and (4) were enriched in regions that are known to contribute to granule localization (Q- and N- rich regions[129]. Of the 128 ESTN proteins, 70 proteins could be successfully tested for their subcellular

localization by fusing them to eYFP. 46/70 unequivocally localized to NPGs upon treatment with SF, thus increasing the number of known NPG proteins from previously 9 proteins[49] to 55.

Certainly, the most interesting outcome of the NPG proteome was the absence of canonical translation factors, since this gives a first hint on NPG function. A closer look revealed, that all translation factors that act downstream of the eIF4F complex, including elongation and release factors, were absent from the group of ESTN proteins. By this, NPGs differ in their protein content from stress granules[34]. Together with previous findings that showed the absence of proteins with functions in early nuclear processing steps[49], these data indicate that NPGs are RNP complexes that have already started or completed nuclear export but not yet entered translation.

5.2 From where are protein components recruited to NPGs?

It still remains unclear where NPG protein components derive from; do they emerge from the cytoplasm or the nucleus? Based on three different findings, the current theory suggests that NPG proteins rather derive from the nucleus than the cytoplasm. (1) NPG formation requires active transcription, but NPG formation is neither prevented nor are the granules dissolved by any interference with translation[49]. (2) Canonical translation factors are absent from NPGs, which indicates that NPGs are RNP complexes that have started or completed nuclear export, but not yet entered translation (see above). (3) Upon combined heat shock (which presumably blocks nuclear export) and SF treatment, PABP2 accumulates in the nucleus, whereas PABP1 remains in the cytoplasm[84]. This last finding raises the possibility that PABP2 as a component of NPGs may bind to mRNA in the nucleus. PABP1, which is not a component of NPGs, may bind to its mRNA targets later in the cytoplasm. In general, this could mean that NPG proteins bind to their target mRNA inside the nucleus und are then exported as RNP complexes. This would also mean that these proteins display a shuttling behavior, which allows them to return to the nucleus after they have been exported to the cytoplasm.

With the NPG proteome at hand, the idea arose to investigate, whether all NPG proteins share this common shuttling feature and bind to their mRNA targets inside the nucleus, afterwards being exported to NPGs. Cells expressing eYFP fusions of randomly chosen NPG proteins were treated with a combination of heat shock, which was shown to inhibit nuclear export in yeast[133,157], and SF, which causes the accumulation of unspliced mRNAs that are most likely impaired in nuclear export. This treatment captured only one of six NPG proteins, namely DRBD2 inside the nucleus. Since the effect that a combination of heat shock and SF has on nuclear export is not completely understood and there is no simple way to investigate these most likely complex effects, no clear results could be drawn from this experiment. At this point, the fact that 2/6 proteins (Tb927.4.4940, ZC3H39) could still localize to

NPGs after a combined treatment remains unresolved. But could it be a hint that at least some protein components are recruited to NPGs from the cytoplasm? At this point, the experiments were discontinued, because the results could neither proof nor disproof the theory that NPG proteins have a nuclear origin.

The general idea of the described experiment above was to block nuclear export by combining the effect of heat shock with the effect of SF (flooding the nucleus with unspliced mRNA). So, might there be other ways to block or interfere with nuclear export in trypanosomes and thereby identify the origin of NPG protein components?

In eukaryotic cells, protein traffic between the nucleus and the cytoplasm is restricted to nuclear pore complexes (NPC), which perforate the nuclear envelope. Small molecules with less than 40 kDa can freely diffuse through NPCs[158] but macromolecules depend on active transport across this border. This directed trafficking, in and out of the nucleus, is regulated by various signal-mediated pathways. For instance, proteins entering the nucleus from the cytoplasm are equipped with a nuclear localization signal (NLS). Classical NLS are comprised of clusters of basic amino acids distributed in monopartite or bipartite stretches[159]. Briefly, this sequence gets recognized by a receptor, which then mediates the transport of the protein/receptor complex through the NPC. Inside the nucleus, the receptor dissociates and returns to the cytoplasm. The NLS as well as the corresponding receptors display heterogeneity throughout the eukaryotic kingdom[160].

If NPG protein components indeed shuttle between the cytoplasm and the nucleus and subsequently bind to their mRNA targets inside the nucleus as suggested above, this would require all protein components being endowed with an NLS. If on the contrary, NPG proteins lack an NLS this would indicate that NPG proteins originates from the cytoplasm. This hypothesis could be investigated by bioinformatically analyzing the NPG proteome candidates for predicted NLS. Furthermore, by experimentally removing the NLS and thus inhibiting a protein's nuclear export ability, or *vice versa* by adding an NLS to a protein, as previously attempted for PABP2 (NPG component) and PABP1 (non-NPG component)[84], the ability of single proteins to localize to NPGs could be investigated. However, the identification of NLS sequences in trypanosome proteins proved too difficult for these approaches: In trypanosomes, only few classical NLS are known. Only one monopartite NLS has been reported: the sequence RGHKRSRE of the La protein is both necessary and sufficient to mediate nuclear localization[160]. All other published nuclear localization requirements of *T. brucei* or *T. cruzi* proteins are more complex: H2B histone and ESAG8 have bipartite NLS[160,161] and *T. cruzi* UBP1 and p14 proteins even require part of their RRM domains for nuclear localization[153,162]. Thus, the complexity of NLS signals in *T. brucei* resembles that in higher eukaryotes, where NLS can be a mixture of monopartite and bipartite basic residues and a cell might need several distinct nuclear import signals to regulate

import of distinct classes of substrates separately[162,163]. A bioinformatic analysis for NLS sequences in all NPG proteins is therefore not possible.

Other ways to block or interfere with nuclear export have been previously tested in trypanosomes, but none proved suitable for the investigation of NPG formation.

XPO1/CRM1 is a major component of the nuclear export machinery and responsible for the transport of U-rich snRNAs and other various protein cargos in metazoans[164]. By specifically and irreversibly binding to this karyopherin, the cytotoxin Leptomycin B can cause the inhibition of nuclear export[165]. In trypanosomes, Leptomycin B was previously used to inhibit nuclear export in two contradictory studies, which analyzed the export of spliced leader (SL) RNA[166,167]. But only one of these studies could confirm that *T. brucei* is indeed sensitive to LMB treatment, implying an involvement of XPO1 in nuclear export; cells displayed growth and morphological phenotypes, including enlarged nuclei and multiple kinetoplasts upon treatment with Leptomycin B[166].

Mex67 was identified as a major mRNA nuclear export factor in trypanosomes. RNAi against Mex67 was shown to inhibit cell growth and cause the accumulation of polyadenylated mRNA within the nucleus[62]. This way of blocking nuclear export messes with the cell's metabolism and morphology in such a way, that NPG formation is prohibited per se and can no longer be analyzed.

Since to this day no simple method exists to manipulate nuclear export in trypanosomes and thus to resolve whether protein components are recruited to NPGs from the nucleus or the cytoplasm, this question remains unanswered. NPGs may be comprised of a mixture of both, cytoplasmic and nuclear proteins; some proteins bind to their mRNA targets within the nucleus and are then transported as mRNA/protein complexes to the cytoplasm, where subsequently other proteins bind.

5.3 NPGs appear not to be related to germ granules, in particular P granules of *C. elegans*

The sexual reproduction of every organism depends on germ cells, because these are the only cells with the ability to transfer genetic material from one generation to the next. A feature of these highly specialized cell types are so called germ granules. This subtype of RNP granules contains various mRNA and protein molecules and is presumably involved in regulating the translation of mRNA molecules inside the germ cells during development[168]. Germ granules are for example thought to store maternal mRNAs until these are recruited to ribosomes for translation, by minimizing the exposure of the mRNA to the cytoplasmic environment[32]. Many of the germ granule components are conserved between distantly related species but are especially well studied in *C. elegans*.

In *C. elegans* germ granules are called 'P granules', because directly after the first cell division, the embryo develops a polarity and germ granules are only present at the posterior end of the zygote ('p' being short for the p lineage of cells that forms at the posterior pole). P granules are perinuclear for the majority of the life cycle and only dissociate from the nucleus in the oocyte stage, where they remain cytoplasmic up to the P3 embryo cell stage[169]. P granules that hold a tight connection to nuclear pores, for example P granules in adult gonads (perinuclear germ granules) extend the nuclear pore complex[87,88], and are the major sites of mRNA export[33]. Thus, it was speculated, that perinuclear P granules might regulate nuclear export and/or play a role in the determination of mRNA fate.

Trypanosome NPGs have some obvious similarities to perinuclear P granules of *C. elegans*. The two granule types share common features in localization, sensitivity to actinomycin D, insensitivity to cycloheximide and the presence of the germ granule marker VASA[49]. Still, trypanosomes do not sexually reproduce, thus a connection between germ granules and NPGs might seem farfetched. However, germ granules share a significant feature with the unicellular parasite: their strong or complete reliance on posttranscriptional gene regulation. In germ cells, transcription is inhibited during late stage oocyte development and early embryogenesis, and the storage or release of maternal mRNAs from germ granules provides the only source for new protein production[85,144]. Based on these findings it was hypothesized that NPGs could be related to perinuclear P granules of *C. elegans* and by extending the nuclear pore complex might have a function as quality control compartment during nuclear export of mRNA.

As part of my work, this hypothesis and the possible relation between P granules and NPGs was tested by the following four approaches: (1) Transgenic expression of *C. elegans* P granule protein components in trypanosomes and investigation of their ability to form NPGs. (2) Comparison of the protein contents of P granules and NPGs. (3) Investigation whether the formation of perinuclear granules in *C. elegans* can be induced in the same way as NPGs: by inhibition of trans-splicing. (4) Analysis of NPGs for a potential connection to nuclear pores on an ultrastructural level.

(1) Of four different *C. elegans* P granule proteins (POS-1, CGH-1 PGL-2, OMA-2) expressed in trypanosomes, two proteins (POS-1, CGH-1) unequivocally localized to NPGs upon treatment with SF and to starvation stress granules upon PBS treatment. With CGH-1 as direct homologue to TbDHH1 but POS-1 showing no homology to any NPG protein, it is tempting to speculate that the localization of a unique *C. elegans* P granule protein to NPGs could indicate a relation between those two granule types. However, if *C. elegans* P granule proteins contain the same classical RNA binding domains as *T. brucei* granule proteins, a localization to NPGs (or stress granules) may be mediated merely through this RNA binding feature. And in fact, CGH-1 (one DEAD/H RNA helicase domain) and POS-1 (two CCCH type zinc finger motifs) both have classical RNA binding motifs that are conserved throughout

eukaryotes. *Vice versa*, the absence of a *C. elegans* protein from NPGs does not exclude a relation between the granules, as it remains unclear whether the transgenic protein is functional. Hence, data from the transgenic expression of P granule proteins in trypanosomes have only a limited value.

(2) With only nine known NPG proteins[49], a comparison of the protein content of NPGs and P granules and an investigation about potential homologies, was not meaningful. Now, with a comprehensive NPG proteome and 55 validated NPG proteins at hand, any compositional relationship between the two granule types could be tested. By applying BLAST analyses, we found no similarities in protein contents between P granules and NPGs: the majority of validated NPG proteins had no homologues in *C. elegans* (42/55); *vice versa*, the majority of known P granules proteins had no homologues in *T. brucei* (24/31). In fact, of the 13 NPG proteins with homologues in *C. elegans*, only four have been reported to be in P granules, namely LAF-1 (related to TbVASA/DBP1), CCF-1 (homologue to TbCAF1), CAR-1 (homologue to TbSCD6) and CGH-1 (homologue to TbDHH1)[137–141]. Of these, the DEAD-box RNA helicase LAF-1 may not even be the true orthologue to TbVASA/DBP1: in *vice versa* BLAST analyses the *T. brucei* helicase has almost identical similarities to a whole range of DEAD box RNA helicases in various organisms.

Importantly, all four proteins that are present in both NPGs and P granules are rather general RNA metabolism proteins, and neither is specific for perinuclear germ granules or NPGs. Thus, our data do not support any compositional relationship between NPGs and germ granules.

(3) This still left the possibility that other organisms or even *C. elegans* may have functional counterparts to NPGs that are inducible in the same way as NPGs: by inhibition of splicing. Previously, this was already addressed in HeLa cells, where inhibition of splicing by spliceostatin A[146] as well as by a morpholino oligonucleotide antisense to the U2 snRNA resulted in no accumulation of p54 (the human homologue to DHH1, which served as granule marker) in perinuclear granules[49]. In contrast to humans, in trypanosomes all mRNAs (with only a few exceptions) are trans-spliced. To rule out the possibility that granule formation depended on the inhibition of trans-splicing rather than cis-splicing, this time, the effects were analyzed in *C. elegans*, where about 70% of all transcripts are trans-spliced to one of two spliced leader RNAs[147]. Like in HeLa cells, the *C. elegans* spliceosome was inhibited with spliceostatin A[148]. In four independent experiments, no altered localization of the general granule marker protein CGH-1 and the specific P granule marker proteins PGL-1 to perinuclear granules in either gonads or somatic cells of young adult worms could be observed. Hence, no similar underlying mechanism of NPG formation in *T. brucei* and perinuclear P granule formation in *C. elegans* was found. However, the success of the inhibition of splicing was not shown: this important control needs to be done for final conclusions.

(4) Different studies revealed that P granules are associated with clusters of nuclear pores and extend the nuclear pore complex in *C. elegans*[87,88]. Perinuclear P granules with a connection to nuclear pore complexes most likely represent the major site of mRNA export[33] and thus were previously suggested to function as nuclear export control compartment. NPGs in trypanosomes were investigated for a similar connection to nuclear pores by electronmicroscopy. Indeed, in sinefungin treated trypanosomes, NPGs were visible as electron dense structures at the nuclear periphery and unequivocally connected to nuclear pores (compare chapter 4.1.4 and 5.4). But does this structural similarity already allow conclusions about a similar granule function, namely nuclear export control? For *C. elegans* it was shown that mRNA passes from the nucleus through P granules into the cytoplasm and that this movement is supposedly slower than diffusion[33], suggesting an active process as cause of the slowdown. In trypanosomes so far, there is no evidence for such an active process regarding NPGs. Performing life experiments similar to those in *C. elegans* could bring clarity.

Even though NPGs of *T. brucei* and germ granules of *C. elegans* share the common structural feature of a tight association with nuclear pores, neither bioinformatical nor experimental approaches could identify any further relationship between the two granule types. Taken together, the data indicate that NPGs may be a unique phenomenon in kinetoplastids.

5.4 Does the connection between NPGs and nuclear pores implicate an involvement of NPGs in nuclear export control?

Electronmicroscopy studies conducted during this work not only for the first time revealed the ultrastructure of NPGs but also confirmed a previously suggested association of NPGs with nuclear pores. In sinefungin treated trypanosome cells over-expressing SCD6, NPGs were visible as electron dense structures, comparable to the previously resolved ultrastructure of P-bodies in human cells[170,171], at the nuclear periphery. It was possible to identify nuclear pores by two indicators: (1) an interrupted nuclear membrane at the position of a nuclear pore and (2) heterochromatin, which is usually densely packed inside the nucleus, forming a passage towards the pores. Thus, visible NPGs could be clearly connected with nuclear pores.

In cells induced for NPG formation, not all visible NPGs appeared connected to nuclear pores; only 38% (n=55) were unequivocally located close to a nuclear pore. *Vice versa*, in almost half of the cases a visible nuclear pore was not associated with a granule; 46% of visible nuclear pores had no attached NPG (n=45). The former lack of correlation could originate from the sectional plane that samples were cut at. The estimated diameter of an NPG (about 200 nm) exceeds the diameter of one nuclear pore complex (about 100 nm)[75]. Thus, for some NPGs the attached nuclear pore may simply be out of focus.

On the contrary, to explain why not all nuclear pores have an associated NPG is not that simple and might be an even more interesting outcome from a functional standpoint. Images of PCF nuclei (+SF) conducted during this work revealed that on average 3-4 nuclear pores were clearly recognizable per nuclear section (n= 28) depending on the sectional plane. Previously, the total number of nuclear pores per trypanosome nucleus was estimated to 200-300; the approximate nuclear diameter averages 1.5 µm[75]. For a more sophisticated statement about the exact number of nuclear pores, and the amount that is connected to NPGs, 3D analysis of images conducted by scanning electron microscopy tomography could be a feasible solution. Alternatively, NPGs could be counted on fluorescence images; here resolution limits need to be considered and make the approach difficult. Since this work was mainly focused on investigating a potential connection of NPGs with nuclear pores and not so much on analyzing the situation of the complete nucleus, the data at this point only allow speculations. The 46% of nuclear pores without a connected NPG could point towards the existence of different subtypes of nuclear pores, each possibly holding different functions. With this hypothesis, not all pores might be responsible for the transport of the same cargo. Nuclear pore complexes could even differ in a way that some are preferably involved in import, others in export, or some transport mRNA and some proteins. The mechanisms underlying nuclear import/export in eukaryotes are still only understood to some extend and especially in trypanosomes only very little is known. One conceivable explanation could be that nuclear pores with NPGs have a function in the export of proteins or mRNA and pores without NPGs are responsible for import. Another possibility is that not all pores are active at the same time and again only at active sites NPGs can be observed. So, could these differences already be the first hint for a possible function of NPGs? At least the latter could be tested by looking at the kinetics of NPGs over time.

Based only on the connection between NPGs and nuclear pores, there is no adequate answer to this question. For example, the nature of the connection between NPGs and the nucleus itself remains completely unclear. The fact that NPGs survive the harsh physical forces that are applied to them during the mechanical purification process, points towards a firm bond between the two structures. On the opposite, addition of detergent dissolves NPGs as witnessed during the approach of purifying nuclear envelops. Furthermore, electronmicroscopy images revealed, that in some cases a strand of electron dense material is pervading from inside of the nucleus through the pore into an NPG. It is tempting to speculate that this strand represents an RNA strand and its associated proteins, anchoring the granule to the nucleus. In fact, evidence for a co-transcriptional export of mRNA in trypanosomes was found recently[172], meaning the 5’ end of an mRNA transcript can already pass the pore, while the 3’ end is still in transcription. This would fit with the idea of NPGs being anchored to the interior of the nucleus via mRNA.

5.5 Do trypanosomes possess a control system to prevent the nuclear export of incorrect mRNAs?

If co-transcriptional export (as described above) allows the 5' end of a premature mRNA to leave the nucleus while the rest of the mRNA is still processed, what assures that the exported mRNA has been finally processed? In general, what feature does an mRNA need to be able to exit the nuclear pore? And do trypanosomes possess mechanisms that control the quality of mRNA before it leaves the nucleus? The hypothesis that NPGs could be involved in this control of mRNA quality and nuclear export was first based on the tight connection between NPGs and nuclear pores. But this connection alone, cannot serve to prove the hypothesis. The following chapter will discuss evidence and theories that are either in favor of or argue against an existing nuclear export control in trypanosomes and will clarify the involvement of NPGs within this process.

Arguments against an existing control

Fact is, that in trypanosomes a fraction of unspliced mRNAs is able to leave the nucleus and does enter the cytoplasm[49,68]. Thus, any existing control system is not perfect. But do trypanosomes need a tight nuclear export control? Since their gene expression uniquely relies on post-transcriptional control mechanisms and thus high levels of mRNA are constantly produced, wouldn't it be efficient and consistent to also not restrict nuclear export? More important, trypanosomes have virtually no introns. This may be the reason why they can afford a certain leakage of unspliced mRNAs to the cytoplasm. A leakage may be tolerable for the parasite, with the worst damage being a misregulation in gene expression, but no production of faulty proteins.

Furthermore, trypanosomes have either no clear homologues of nuclear export control key players that were previously identified in yeast[57,173–175] (Pml1/Pml3) or, if they have a potential homologue, it remains unclear whether the protein is a true functional homologue (Mlp)[77].

There is evidence that, even though the trypanosomatid lineage diverged very early during the development of eukaryotes, the majority of NPC architecture is conserved throughout the eukaryotic kingdom and was already established in the last common eukaryotic ancestor[77]. Still, the trypanosome NPC structurally differs from other eukaryotic NPCs. In contrast to opisthokonts, the trypanosome NPC displays almost complete nucleocytoplasmic symmetry; with exception of the nuclear basket. This absence of a clear nucleocytoplasmic asymmetry in the trypanosome NPC is remarkable, especially since NPC asymmetry is crucial for driving opisthokont mRNA export[176–178], and might reflect the basic mechanism of nucleocytoplasmic transport[66]. Possibly, this symmetry could also point towards a difference, or even lack of nucleocytoplasmic transport control mechanisms in trypanosomes.

Most importantly, during this work, the previously reported leakage of unspliced mRNA[179,180] was quantified by single molecule mRNA FISH, which revealed that indeed about half of all unspliced tubulin precursor mRNA molecules are cytoplasmic. If NPGs would play a role in this system, for example by restraining mRNA from passing through into the free cytoplasm, one would expect that unspliced transcripts accumulate in the granules. However, no accumulation of unspliced mRNAs in NPGs was observed, arguing against such a control function. If not caused by difficulties of probe access, in theory, the invisibility of these transcripts in NPGs could be due to their fast degradation within the granules. However, the NPG proteome contains only one mRNA decay enzyme, XRNA, and this has been previously shown not to be involved in the degradation of unspliced transcripts: RNAi depletion of XRNA caused no accumulation of mRNA precursors[61]. Still, the data do not per se exclude the presence of a control system, because the stability of tubulin precursor mRNAs within the different cellular compartments remains unknown. If there is active degradation of unspliced transcripts within the nucleus but not in the cytoplasm, the majority of unspliced mRNAs might still be degraded before reaching the cytoplasm. In fact, it was previously shown that unspliced mRNAs get degraded by the nuclear exosome, but not by the cytoplasmic 5'-3' exoribonuclease XRNA or the NMD protein UPF1[61].

To summarize the facts and circumstances that argue against an existing nuclear export control in trypanosomes: (1) unspliced mRNA can leave the nucleus (2) mRNA gets co-transcriptionally exported (3) key factors of nuclear export control are missing (4) trypanosome mRNAs mostly lack introns and thus may not need a tight quality control (5) the nuclear pore is built symmetrically. Most importantly, it became apparent that NPGs are most likely not involved in mRNA quality and nuclear export control.

Proposed model for role of NPGs

Alternatively, NPGs could be the result of a transport jam at the nuclear pore caused by flooding the system with very long, unspliced transcripts (SF treatment). If mRNA export in trypanosomes occurs co-transcriptional, the 5' end of the mRNA transcript would have already passed the pore, while the 3' end is still in transcription. Even assuming some kind of nuclear export control exists (see below), for example by proteins that specifically recognize and bind to unspliced mRNAs, it is possible that nuclear export of an incorrect/incomplete mRNA is started. In the case of inhibited trans-splicing, transcript size, transcription time and therefore the time any mRNA resides at the cytoplasmic site of the pore will increase. With the 5' end of the polycistronic mRNAs, together with any accompanying proteins being stuck in the pore, microscopically visible granules (NPGs) will appear at the cytoplasmic site of the pore. The remaining part of the mRNA would be anchored inside the nucleus awaiting processing.

The mechanism of nuclear retention of unspliced mRNAs is still poorly understood in any organism[57] and the study of mRNA export pathways in general is challenging, due to the fast kinetics. Uniquely in trypanosomes we can massively increase the amount of unprocessed mRNA by drug-dependent inhibition of trans-splicing. Interrupting the mRNA export pathway by flooding the system with very long, polycistronic mRNA transcripts appears to freeze nuclear export and provides a novel tool for the study of this important pathway.

Arguments for an existing control

But do all the facts above, and the insight that NPGs are probably not actively involved in the control of nuclear export, mean that trypanosomes do not possess a control system at all? Even though they show major differences in nuclear mRNA metabolism and lack most of the control mechanisms and key players typically involved in nuclear export control on other eukaryotes, at least some steps during an mRNA's life in trypanosomes underlies a certain control. The main question is how and whether trypanosomes regulate the export of polycistronic mRNA precursors and how they prevent their accumulation in the cytoplasm.

In simplified terms: In eukaryotes unspliced mRNAs are retained within the nucleus with the help of the nuclear pore complex (NPC)[57]. Within the nucleus, faulty mRNAs get actively degraded by the nuclear exosome. In case this system fails, intron containing mRNAs are recognized and degraded by the cytoplasmic nonsense mediated decay (NMD) system, part of this system is not completely cytoplasmic but still associated with the nucleus[181].

In trypanosomes, the main mRNA processing steps appear to be conserved. However, with the exception of Mex67-Mtr2, all complexes and proteins involved in regulating mRNA export in opisthokonts are either absent (TREX, TREX-2, RES, 1049 DBP5, probably Mlp1-2) or have no reported functions in mRNA export (non-classical CTD, SR proteins except perhaps RRM1, EJC, TRAMP complex)[182]. Still, there is some evidence for an existing control to prevent the accumulation of polycistronic mRNAs. It was for example shown that the half-life of tubulin dicistrons is significantly shorter than the half-life of mature tubulin mRNA[49], indicating the presence of an active mechanism for the removal of unspliced mRNAs.

Trypanosomes have a conserved exosome that is essential[110]. They possess orthologues to all RNA exosome subunits[110,111] and most importantly, evidence points, amongst others, to an essential function in removal of unspliced mRNAs[61]. Whether the exosome specifically targets these RNAs, or, whether these RNAs are degraded because they have an extended exposure time to the exosome is still not known. It is possible that mRNA precursors get degraded co-transcriptionally, because mRNA processing and degradation act as opponents, and longer mRNAs are more likely degraded than short

mRNAs simply because processing time and thus exosomal exposure is longer[183]. Is has also been shown in yeast that the exosome constantly competes with the spliceosome for intron containing mRNAs[184]. This idea would be in agreement with the current model in opisthokonts[185,186]. Or the exosome rather unspecifically degrades RNA, which is supported by the finding that several short-lived RNA species are stabilized, when nuclear export is inhibited in various ways[64] and that developmentally regulated mRNAs are enriched in nuclear fractions in the related parasite Trypanosoma cruzi[187].

It is possible that the levels of defective mRNA in general might simply be regulated by kinetics, based on the 'survival of the fittest'. Defective mRNAs could possibly be slowed down in comparison to correct mRNAs, regarding transport or control processes. For example, unspliced mRNA could get stuck inside the nuclear pore during translocation, with their 5' end already reaching into cytoplasm, thus slowing down the export process. By this capture, the mRNA might be more susceptible to the major cytoplasmic degradation pathway by the 5'-3' exoribonuclease XRNA (the trypanosome orthologue to Xrn1). The second of the two main decay routes, degradation by the cytoplasmic 3'-5' exosome, would play no or only a minor role in this scenario[61,115,188]. This co-transcriptional processing, which is supported by the preferential cytoplasmic degradation, might be a sufficient way of keeping the leakage of unprocessed mRNAs under control. This would also explain why NPGs are visible as electron dense structures when the system is flooded with polycistronic mRNA precursors. They could physically block the pores, perhaps because some remain attached to the transcription site and fail to exit the nucleus completely.

Alternatively, a yet unknown quality control checkpoint could act at the nuclear basket that recognizes an mRNA as not fully processed, slowing or preventing export. In this work several additional candidate proteins for a nuclear export control of mRNAs were identified based on their re-localization to the nucleus or nuclear pores upon inhibition of trans-splicing. Whether either of these five proteins acts as such a checkpoint in RNA export control remains to be investigated.

RNST proteins and their potential role in nuclear export control

One major outcome of this work was the identification of further candidate proteins for nuclear export control, in addition to the previously known protein DRBD3. Within the group of ESTN proteins, that were initially analyzed for localization to NPGs, some proteins fully or partially re-localized from the cytoplasm to the nucleus (ZFP1, ZFP2, RBP12, ZC3H29 and Tb927.11.6600), and in once case to nuclear pores (Tb927.10.7580), when trans-splicing was inhibited (SF); they were named 'RNST proteins' (proteins that re-localized to nucleus/nuclear pores upon sinefungin treatment). At transcriptional inhibition (Actinomycin D), RNST proteins did not change their localization. Therefore, the results indicate that the re-localization is not caused by the absence of fully processed mRNAs but

is a consequence of the accumulated incompletely processed, polycistronic mRNA, caused by SF treatment.

The re-localization of RNST proteins to the nucleus can potentially be caused in two ways: either, a protein is actively recruited to the nucleus upon treatment with SF or, the protein shuttles between the cytoplasm and the nucleus and is somehow prevented from leaving the nucleus in the presence of SF. In the latter case, such proteins could bind to unspliced transcripts within the nucleus and, perhaps, mark them as export-incompetent or as target for exosomal degradation. In trypanosomes such a shuttling behavior has already been discovered for one of the RNST proteins, namely DRBD3[152], but also for many non-RNST proteins, for example PABP2[84] or UBP1[153].

One identified RNST protein re-localized to nuclear pores upon inhibition of trans-splicing. This was unexpected since a re-localization from the nucleoplasm to nuclear pores should not change the total protein amount within the purified nuclei. However, assuming that a fraction of nucleoplasmic proteins gets lost during the preparation of nuclei, this loss could potentially be smaller in the presence of SF, because part of the protein is 'anchored' to the nuclear pores. This makes Tb927.10.7580, and other proteins displaying the same behavior, very strong candidates for a role in nuclear export control. Such proteins could bind to unspliced mRNAs and serve as a retaining signal during the process of nuclear export. If this signal protein is still bound to an mRNA during nuclear export it could get trapped within the pore and mediate the exosomal degradation of the mRNA.

The further analysis of these two groups of RNST proteins may help to answer the question how mRNA export is controlled in trypanosomes.

5.6 Identification of new candidates for nuclear export control in trypanosomes

One of the already identified key players involved in nuclear export of mRNAs in trypanosomes is the putative homologue to Mex67 in yeast. TbMex67 localizes to the nucleoplasm and to spots that resemble nuclear pores[119]. An essential role of Mex67 for cell growth and nuclear mRNA export was suggested, because Mex67 down-regulation inhibits cell growth and causes an accumulation of polyadenylated mRNA within the nucleus[62]. This block of nuclear export caused by Mex67 RNAi now provided a tool to investigate the following issues: (1) Is the observed re-localization of RNST proteins to the nucleus upon SF treatment caused by nuclear retention or by active transport? (2) Do proteins get trapped by the blocked nuclear export of mRNA (Mex67 RNAi) or by to the accumulation of unspliced precursor mRNAs within the nucleus (SF treatment)? (3) Is any nuclear re-localization upon Mex67 RNAi specific to RNST proteins or do random ESTN proteins (control proteins) display the same

behavior? Ultimately this could point towards a function of RNST proteins in nuclear export control, either by specifically preventing defective mRNA from leaving the nucleus, or by being involved in the same export pathway as Mex67/ being co-factors of Mex67 depended export.

During this work it was shown that Mex67 RNAi phenocopied the effect of SF treatment on three RNST proteins (ZFP2, DRBD3, Tb927.10.7580), whereas none of the control proteins (PABP2, eIF4E3 and CAF1) changed their localization. The observed effect was specific to RNST proteins and not a general feature of all shuttling proteins, since PABP2 as one of the control proteins has a known shuttling behavior[155] and did not re-localize to the nucleus under Mex67 RNAi conditions. Due to time limitations this analysis could not be completed for all proteins (RNST and control proteins).

RNAi depletion of RNST proteins resulted in no or only minor changes in cell growth and morphology. Also, the RNA-FISH analysis of some RNST depleted cells revealed no changes in intracellular distribution of polyadenylated mRNAs. Both effects were previously reported for the down regulation of the identified nuclear export factor Mex67; RNAi resulted in a severe growth/morphology defect and caused the accumulation of polyadenylated mRNA within the nucleus[62,63]. Again, due to time limitations this analysis could not be completed and can only serve as a first indication.

To summarize, the initial results support the proposed hypothesis of RNST proteins being novel candidates of nuclear export control, in which they either serve to mark of premature or defective mRNA for degradation or as a retaining signal that anchors these mRNAs within the nucleus/ nuclear pores. With this, a first step into the better understanding of the quality control and export mechanism in trypanosomes was achieved.

5.7 Final conclusion - The bigger picture

Kinetoplastids are responsible for several neglected tropical diseases, including African Sleeping Sickness, Chagas disease or Leishmaniasis, that thread millions of lives world-wide. One unique feature of these parasites is their full reliance on polycistronic mRNA transcription and subsequent processing by trans-splicing. Trypanosomes were used as a model organism during this work, because they provide a unique system, in which trans-splicing could be inhibited globally by drug treatment, enabling the analysis of changes on protein and mRNA level. By massively flooding the system with very long, polycistronic mRNA transcripts, we have interrupted the mRNA export pathway of trypanosomes. As a result, we observed RNP complexes at the cytoplasmic site of the nucleus (NPGs), in between nuclear export and translation, as well as several proteins that changed their localization to the nucleus or to nuclear pores.

The study of mRNA export pathways in any organism is challenging, due to its fast kinetics. In mammalian cells, an mRNA is transported through the NPC by three steps: docking (80 ms), transport (5–20 ms) and release (80 ms), totaling 180±10 ms[189]. Even though there is extensive knowledge on the physical structure and composition of the nuclear pore complex in eukaryotes, also particularly in trypanosomes, the transport processes and the dynamics of mRNA export at nuclear pores remain unknown. The global and massive increase in polycistronic mRNAs that we achieve in trypanosomes appears to freeze nuclear export and provides a novel tool for the study of this important pathway and the regulators involved in it. We propose a model, in which an mRNA is co-transcriptionally exported, meaning the 5' end of an mRNA transcript can already pass the pore, while the 3' end is still in transcription, by this "anchoring" the mRNA and its associated proteins inside the nucleus. In that state, the mRNP complex awaits to be further processed or cleared for nuclear export. Normally, the amount of protein/mRNA within this state is small, since most RNPs are transported across the pores very fast. As soon as this process is paused, because the majority of mRNA is defective due to inhibition of trans-splicing (unspliced polycistrons), microscopically visible granules (NPGs) appear at the cytoplasmic site of the pore.

The inhibition of trans-splicing not only resulted in the formation of NPGs, but also revealed proteins that changed their localization to the nucleus or nuclear pores. These proteins were identified as novel, potential nuclear export factors or factors involved in quality control of mRNA. The mechanism of how unspliced or defective mRNA is retained inside the nucleus is still poorly understood in any organism[57]. It also remains unknown, which factors exactly need to bind to an mRNA for enabling it to leave the nucleus. Identifying new potential regulators during this work is a first step towards a better understanding of this important pathway. Overall, the global dataset of changes in protein localizations at inhibition of trans-splicing paves the way towards a better understanding of mRNA processing and its control mechanisms, not only in trypanosomes, but also in eukaryotes in general.

With this work, I provide two high-quality proteome datasets: the nuclear proteome of *T. brucei* as well as the proteome of NPGs (a subtype of RNP granules). Both are essential complements to the already existing databases for trypanosome research, TriTrypDB[80] and TrypTag.org[107].

6 References

1. Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. The Lancet *375*, 148–159.

2. Horn, D., and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Current Opinion in Microbiology *13*, 700–705.

3. Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A.M. (2009). Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. Genes and Development *23*, 1063–1076.

4. Ngô, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America *95*, 14687–14692.

5. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. Nature *391*, 806–811.

6. Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B., and Carrington, M. (2007). Functional genomics in Trypanosoma brucei: A collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. Molecular and Biochemical Parasitology *154*, 103–109.

7. Wirtz, E., Leal, S., Ochatt, C., and Cross, G.A.M. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Molecular and Biochemical Parasitology *99*, 89–101.

8. Akiyoshi, B., and Gull, K. (2013). Evolutionary cell biology of chromosome segregation: Insights from trypanosomes. Open Biology *3*.

9. Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J., and Gluenz Sir William, E. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastidsdoi:10.1098/rsos.170095.

10. Rico, E., Jeacock, L., Kovářová, J., and Horn, D. Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes OPENdoi:10.1038/s41598-018-26303-w.

11. Kramer, S. (2014). RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. Wiley interdisciplinary reviews. RNA *5*, 263–84.

12. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., El-Sayed, N.M., et al. (2005). The genome of the African trypanosome Trypanosoma brucei. Science *309*, 416–422.

13. Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Myler, P.J., et al. (2005). The genome of the kinetoplastid parasite, Leishmania major. Science *309*, 436–442.

14. El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Andersson, B., et al. (2005). The genome sequence of Trypanosoma cruzi, etiologic agent of chagas disease. Science *309*.

15. Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S.,

Bringaud, F., Baldauf, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R.O., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., and Berriman, M. (2007). Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nature Genetics *39*, 839–847.

16. Lukeš, J., Skalický, T., Týč, J., Votýpka, J., and Yurchenko, V. (2014). Evolution of parasitism in kinetoplastid flagellates. Molecular and Biochemical Parasitology *195*, 115–122.

17. Anderson, P., and Kedersha, N. (2009). RNA granules: Post-transcriptional and epigenetic modulators of gene expression. Nature Reviews Molecular Cell Biology *10*, 430–436.

18. Brengues, M., Teixeira, D., and Parker, R. (2005). Cell biology: Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science *310*, 486–489.

19. Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Relief of microRNA-Mediated Translational Repression in Human Cells Subjected to Stress. Cell *125*, 1111–1124.

20. Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science *300*, 805–808.

21. Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fitzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. Journal of Cell Biology *169*, 871–884.

22. Anderson, P., and Kedersha, N. (2006). RNA granules. Journal of Cell Biology *172*, 803–808.

23. Teixeira, D., and Parker, R. (2007). Analysis of P-body assembly in Saccharomyces cerevisiae. Molecular Biology of the Cell *18*, 2274–2287.

24. Cougot, N., Molza, A.E., Giudice, E., Cavalier, A., Thomas, D., and Gillet, R. (2013). Structural organization of the polysomes adjacent to mammalian processing bodies (P-bodies). RNA Biology *10*, 314–320.

25. Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. Molecular and Cellular Biology *27*, 3970–3981.

26. Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225–229.

27. Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. Trends in Cell Biology *26*, 668–679.

28. Brengues, M., and Parker, R. (2007). Accumulation of polyadenylated mRNA, Pab1p, eIF4E, and eIF4G with P-bodies in Saccharomyces cerevisiae. Molecular Biology of the Cell *18*, 2592–2602.

29. Hoyle, N.P., Castelli, L.M., Campbell, S.G., Holmes, L.E.A., and Ashe, M.P. (2007). Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. Journal of Cell Biology *179*, 65–74.

30. Buchan, J.R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in Saccharomyces cerevisiae. Journal of Cell Biology *183*, 441–455.

31. Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. Journal of Cell Science *118*, 981–992.

32. Sengupta, M.S., and Boag, P.R. (2012). Germ granules and the control of mRNA translation. IUBMB Life *64*, 586–594.

33. Sheth, U., Pitt, J., Dennis, S., and Priess, J.R. (2010). Perinuclear P granules are the principal sites of mRNA export in adult C. elegans germ cells. Development *137*, 1305–1314.

34. Fritz, M., Vanselow, J., Sauer, N., Lamer, S., Goos, C., Siegel, T.N., Subota, I., Schlosser, A., Carrington, M., and Kramer, S. (2015). Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. Nucleic Acids Research *43*, 8013–8032.

35. Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

36. Alves, L.R. (2016). RNA-binding proteins related to stress response and differentiation in protozoa. World Journal of Biological Chemistry *7*, 78.

37. El-Sayed, N.M.A., Ghedin, E., Song, J., MacLeod, A., Bringaud, F., Larkin, C., Wanless, D., Peterson, J., Hou, L., Taylor, S., Tweedie, A., Biteau, N., Khalak, H.G., Lin, X., Mason, T., Hannick, L., Caler, E., Blandin, G., Bartholomeu, D., Simpson, A.J., Kaul, S., Zhao, H., Pai, G., Van Aken, S., Utterback, T., Haas, B., Koo, H.L., Umayam, L., Suh, B., Gerrard, C., Leech, V., Qi, R., Zhou, S., Schwartz, D., Feldblyum, T., Salzberg, S., Tait, A., Turner, C.M.R., Ullu, E., White, O., Melville, S., Adams, M.D., Fraser, C.M., and Donelson, J.E. (2003). The sequence and analysis of Trypanosoma brucei chromosome II. Nucleic Acids Research *31*, 4856–4863.

38. Hall, N., Berriman, M., Lennard, N.J., Harris, B.R., Hertz-Fowler, C., Bart-Delabesse, E.N., Gerrard, C.S., Atkin, R.J., Barron, A.J., Bowman, S., Bray-Allen, S.P., Bringaud, F., Clark, L.N., Corton, C.H., Cronin, A., Davies, R., Doggett, J., Fraser, A., Grüter, E., Hall, S., Harper, A.D., Kay, M.P., Leech, V., Mayes, R., Price, C., Quail, M.A., Rabbinowitsch, E., Reitter, C., Rutherford, K., Sasse, J., Sharp, S., Shownkeen, R., MacLeod, A., Taylor, S., Tweedie, A., Turner, C.M.R., Tait, A., Gull, K., Barrell, B., and Melville, S.E. (2003). The DNA sequence of chromosome I of an African trypanosome: Gene content, chromosome organisation, recombination and polymorphism. Nucleic Acids Research *31*, 4864–4873.

39. Michaeli, S. (2011). Trans-splicing in trypanosomes: Machinery and its impact on the parasite transcriptome. Future Microbiology *6*, 459–474.

40. Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H.O., Bishop, J.R., Falcone, F.H., Gavrilescu, C., Montgomery, J.L., Santori, M.I., Stern, L.S., Wang, Z., Ullu, E., and Tschudi, C. (2000). A new twist in trypanosome RNA metabolism: Cis-splicing of pre-mRNA. RNA *6*, 163–169.

41. Siegel, T.N., Hekstra, D.R., Wang, X., Dewell, S., and Cross, G.A.M. (2010). Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites. Nucleic Acids Research *38*, 4946–4957.

42. Clayton, C., and Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Molecular and Biochemical Parasitology *156*, 93–101.

43. Haile, S., and Papadopoulou, B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. Current Opinion in Microbiology *10*, 569–577.

44. Cassola, A. (2011). RNA Granules Living a Post-transcriptional Life: the Trypanosomes' Case. Current chemical biology *5*, 108–117.

45. Holetz, F.B., Correa, A., Ávila, A.R., Nakamura, C.V., Krieger, M.A., and Goldenberg, S. (2007). Evidence of P-body-like structures in Trypanosoma cruzi. Biochemical and Biophysical Research Communications *356*, 1062–1067.

46. Cassola, A., De Gaudenzi, J.G., and Frasch, A.C. (2007). Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. Molecular Microbiology *65*, 655–670.

47. Kramer, S., Queiroz, R., Ellis, L., Webb, H., Hoheisel, J.D., Clayton, C., and Carrington, M. (2008). Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2α phosphorylation at Thr169. Journal of Cell Science *121*, 3002–3014.

48. Garcia-Silva, M.R., Frugier, M., Tosar, J.P., Correa-Dominguez, A., Ronalte-Alves, L., Parodi-Talice, A., Rovira, C., Robello, C., Goldenberg, S., and Cayota, A. (2010). A population of tRNA-derived small RNAs is actively produced in Trypanosoma cruzi and recruited to specific cytoplasmic granules. Molecular and Biochemical Parasitology *171*, 64–73.

49. Kramer, S., Marnef, A., Standart, N., and Carrington, M. (2012). Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate. Journal of Cell Science *125*, 2896–2909.

50. McNally, K.P., and Agabian, N. (1992). Trypanosoma brucei spliced-leader RNA methylations are required for trans splicing in vivo. Molecular and Cellular Biology *12*, 4844–4851.

51. Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA Processing Reaches Back toTranscription and Ahead to Translation. Cell *136*, 688–700.

52. Kelly, S.M., and Corbett, A.H. (2009). Messenger RNA Export from the Nucleus: A Series of Molecular Wardrobe Changes. Traffic *10*, 1199–1208.

53. Kallehauge, T.B., Robert, M.C., Bertrand, E., and Jensen, T.H. (2012). Nuclear Retention Prevents Premature Cytoplasmic Appearance of mRNA. Molecular Cell *48*, 145–152.

54. Soheilypour, M., and Mofrad, M.R.K. (2018). Quality control of mRNAs at the entry of the nuclear pore: Cooperation in a complex molecular system. Nucleus *9*, 202–211.

55. Schmidt, K., Butler, J.S., and Scott, J. (2013). Nuclear RNA Surveillance: Role of TRAMP in Controlling Exosome Specificity. Wiley Interdiscip Rev RNA *4*, 217–231.

56. Carmody, S.R., and Wente, S.R. (2009). mRNA nuclear export at a glance. Journal of Cell Science *122*, 1933–1937.

57. Bonnet, A., and Palancade, B. (2015). Intron or no intron: A matter for nuclear pore complexes. Nucleus *6*, 455–461.

58. Huang, Y., Yario, T.A., and Steitz, J.A. (2004). *A molecular link between SR protein dephosphorylation and mRNA export*at <www.pnas.orgcgidoi10.1073pnas.0403533101>.

59. Etheridge, R.D., Clemens, D.M., Gershon, P.D., and Aphasizhev, R. (2009). Identification and characterization of nuclear non-canonical poly(A) polymerases from Trypanosoma brucei. Molecular and Biochemical Parasitology *164*, 66–73.

60. Clayton, C., and Estevez, A. (2010). The exosomes of trypanosomes and other protists. Advances in experimental medicine and biology *702*, 39–49.

61. Kramer, S., Piper, S., Estevez, A., and Carrington, M. (2016). Polycistronic trypanosome mRNAs are a target for the exosome. Molecular and Biochemical Parasitology *205*, 1–5.

62. Schwede, A., Manful, T., Jha, B.A., Helbig, C., Bercovich, N., Stewart, M., and Clayton, C. (2009). The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. Nucleic Acids Research *37*, 5511–5528.

63. 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. Molecular Microbiology *88*, 728–739.

64. Bühlmann, M., Walrad, P., Rico, E., Ivens, A., Capewell, P., Naguleswaran, A., Roditi, I., and Matthews, K.R. (2015). NMD3 regulates both mRNA and rRNA nuclear export in African trypanosomes via an XPOI-linked pathway. Nucleic Acids Research *43*, 4491–4504.

65. Serpeloni, M., Moraes, C.B., MunizJoã, J.R.C., Motta, M.C.M., Ramos, A.S.P., Kessler, R.L., Inoue, A.H., daRocha, W.D., Yamada-Ogatta, S.F., Fragoso, S.P., Goldenberg, S., Freitas-Junior, L.H., and ÁvilaAndré, A.R. (2011). An essential nuclear protein in trypanosomes is a component of mrna transcription/export pathway. PLoS ONE *6*.

66. Obado, S.O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N.E., Chait, B.T., Field, M.C., and Rout, M.P. (2016). Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. PLoS Biology *14*, 1–30.

67. Obado, S.O., Field, M.C., and Rout, M.P. (2017). Comparative interactomics provides evidence for functional specialization of the nuclear pore complex. Nucleus *8*, 340–352.

68. Jäger, A. V., De Gaudenzi, J.G., Cassola, A., D'Orso, I., and Frasch, A.C. (2007). mRNA maturation by two-step trans-splicing/polyadenylation processing in trypanosomes. Proceedings of the National Academy of Sciences of the United States of America *104*, 2035–2042.

69. Sunter, J., Wickstead, B., Gull, K., and Carrington, M. (2012). A new generation of T7 RNA polymerase-independent inducible expression plasmids for Trypanosoma brucei. PloS one *7*.

70. McCulloch, R., Vassella, E., Burton, P., Boshart, M., and Barry, J.D. (2004). Transformation of monomorphic and pleomorphic Trypanosoma brucei. Methods in molecular biology (Clifton, N.J.)doi:10.1385/1-59259-761-0:053.

71. Hirumi, H., and Hirumi, K. (1989). Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. Journal of Parasitology *75*, 985–989.

72. Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217–1226.

73. Dean, S., Sunter, J., Wheeler, R.J., Hodkinson, I., Gluenz, E., and Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. Open biology *5*, 140197.

74. Krüger, T., Hofweber, M., and Kramer, S. (2013). SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains. Molecular Biology of the Cell *24*, 2098–2111.

75. Rout, M.P., and Field, M.C. (2001). Isolation and characterization of subnuclear compartments from Trypanosoma brucei. Identification of a major repetitive nuclear lamina component. Journal of Biological Chemistry *276*, 38261–38271.

76. Degrasse, J.A., Chait, B.T., Field, M.C., and Rout, M.P. (2008). High-yield isolation and subcellular proteomic characterization of nuclear and subnuclear structures from trypanosomes. Methods in Molecular Biology *463*, 77–92.

77. DeGrasse, J.A., Dubois, K.N., Devos, D., Siegel, T.N., Sali, A., Field, M.C., Rout, M.P., and Chait, B.T. (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Molecular and Cellular Proteomics *8*, 2119–2130.

78. Bluhm, A., Casas-Vila, N., Scheibe, M., and Butter, F. (2016). Reader interactome of epigenetic histone marks in birds. Proteomics *16*, 427–436.

79. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology *26*, 1367–1372.

80. Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., Gardner, M.J., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Logan, F.J., Miller, J.A., Mitra, S., Myler, P.J., Nayak, V., Pennington, C., Phan, I., Pinney, D.F., Ramasamy, G., Rogers, M.B., Roos, D.S., Ross, C., Sivam, D., Smith, D.F., Srinivasamoorthy, G., Stoeckert, C.J., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., and Wang, H. (2009). TriTrypDB: A functional genomic resource for the Trypanosomatidae. Nucleic Acids Research *38*.

81. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods *13*, 731–740.

82. Weimer, R.M. (2006). Preservation of C. elegans tissue via high-pressure freezing and freeze-substitution for ultrastructural analysis and immunocytochemistry. Methods in molecular biology (Clifton, N.J.) *351*, 203–21.

83. Stigloher, C., Zhan, H., Zhen, M., Richmond, J., and Bessereau, J.L. (2011). The presynaptic dense projection of the Caenorhabiditis elegans cholinergic neuromuscular junction localizes synaptic vesicles at the active zone through SYD-2/liprin and UNC-10/RIM-dependent interactions. Journal of Neuroscience *31*, 4388–4396.

84. Kramer, S., Bannerman-Chukualim, B., Ellis, L., Boulden, E.A., Kelly, S., Field, M.C., and Carrington, M. (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*.

85. Updike, D., and Strome, S. (2010). P granule assembly and function in Caenorhabditis elegans germ cells. Journal of Andrology *31*, 53–60.

86. Updike, D.L., and Strome, S. (2009). A genomewide RNAi screen for genes that affect the stability, distribution and function of P granules in Caenorhabditis elegans. Genetics *183*, 1397–1419.

87. Pitt, J.N., Schisa, J.A., and Priess, J.R. (2000). P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. Developmental Biology *219*, 315–333.

88. Updike, D.L., Hachey, S.J., Kreher, J., and Strome, S. (2011). P granules extend the nuclear pore complex environment in the C. elegans germ line. Journal of Cell Biology *192*, 939–948.

89. Farine, L., Niemann, M., Schneider, A., and Bütikofer, P. (2015). Phosphatidylethanolamine and phosphatidylcholine biosynthesis by the Kennedy pathway occurs at different sites in Trypanosoma brucei. Scientific Reports *5*, 1–11.

90. De Castro Moreira Dos Santos, A., Kalume, D.E., Camargo, R., Gómez-Mendoza, D.P., Correa, J.R., Charneau, S., De Sousa, M.V., De Lima, B.D., and Ricart, C.A.O. (2015). Unveiling the trypanosoma Cruzi nuclear proteome. PLoS ONE *10*, 1–12.

91. Goos, C., Dejung, M., Janzen, C.J., Butter, F., and Kramer, S. (2017). The nuclear proteome of Trypanosoma brucei. PLoS ONE *12*.

92. Bangs, J.D., Uyetake, L., Brickman, M.J., Balber, A.E., and Boothroyd, J.C. (1993). Molecular cloning and cellular localization of a BiP homologue in trypanosoma brucei. Divergent ER retention signals in a lower eukaryote. Journal of Cell Science *105*, 1101–1113.

93. Schlaeppi, K., Deflorin, J., and Seebeck, T. (1989). The major component of the paraflagellar rod of Trypanosoma brucei is a helical protein that is encoded by two identical, tandemly linked genes. Journal of Cell Biology *109*, 1695–1709.

94. Field, H., Ali, B.R.S., Sherwin, T., Gull, K., Croft, S.L., and Field, M.C. (1999). TbRab2p, a marker for the endoplasmic reticulum of Trypanosoma brucei, localises to the ERGIC in mammalian cells. Journal of Cell Science *112*, 147–156.

95. Smith, T.K., and Bütikofer, P. (2010). Lipid metabolism in Trypanosoma brucei. Molecular and Biochemical Parasitology *172*, 66–79.

96. Broadhead, R., Dawe, H.R., Farr, H., Griffiths, S., Hart, S.R., Portman, N., Shaw, M.K., Ginger, M.L., Gaskell, S.J., McKean, P.G., and Gull, K. (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. Nature *440*, 224–227.

97. Panigrahi, A.K., Ogata, Y., Zíková, A., Anupama, A., Dalley, R.A., Acestor, N., Myler, P.J., and Stuart, K.D. (2009). A comprehensive analysis of trypanosoma brucei mitochondrial proteome. Proteomics *9*, 434–450.

98. Dean, S., Moreira-Leite, F., Varga, V., and Gull, K. (2016). Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes. Proceedings of the National Academy of Sciences of the United States of America *113*, E5135–E5143.

99. Güther, M.L.S., Urbaniak, M.D., Tavendale, A., Prescott, A., and Ferguson, M.A.J. (2014). High-confidence glycosome proteome for procyclic form trypanosoma brucei by epitope-tag organelle enrichment and SILAC proteomics. Journal of Proteome Research *13*, 2796–2806.

100. Shimogawa, M.M., Saada, E.A., Vashisht, A.A., Barshop, W.D., Wohlschlegel, J.A., and Hill, K.L. (2015). Cell surface proteomics provides insight into stage-specific remodeling of the host-parasite interface in trypanosoma brucei. Molecular and Cellular Proteomics *14*, 1977–1988.

101. Kohlwein, S.D., Eder, S., Oh, C.-S., Martin, C.E., Gable, K., Bacikova, D., and Dunn, T. (2001). Tsc13p Is Required for Fatty Acid Elongation and Localizes to a Novel Structure at the Nuclear-Vacuolar Interface inSaccharomyces cerevisiae. Molecular and Cellular Biology *21*, 109–125.

102. Lee, S.H., Stephens, J.L., Paul, K.S., and Englund, P.T. (2006). Fatty Acid Synthesis by Elongases in Trypanosomes. Cell *126*, 691–699.

103. DuBois, K.N., Alsford, S., Holden, J.M., Buisson, J., Swiderski, M., Bart, J.-M., Ratushny, A. V., Wan, Y., Bastin, P., Barry, J.D., Navarro, M., Horn, D., Aitchison, J.D., Rout, M.P., and Field, M.C. (2012). NUP-1 Is a Large Coiled-Coil Nucleoskeletal Protein in Trypanosomes with Lamin-Like Functions. In: Misteli, T (ed.). PLoS Biology *10*, e1001287.

104. Bessat, M., and Ersfeld, K. (2009). Functional characterization of cohesin SMC3 and separase and their roles in the segregation of large and minichromosomes in *Trypanosoma brucei*. Molecular Microbiology *71*, 1371–1385.

105. Tarun, S.Z., Schnaufer, A., Ernst, N.L., Proff, R., Deng, J., Hol, W., and Stuart, K. (2008). KREPA6 is an RNA-binding protein essential for editosome integrity and survival of Trypanosoma brucei. RNA *14*, 347–358.

106. Zíková, A., Panigrahi, A.K., Dalley, R.A., Acestor, N., Anupama, A., Ogata, Y., Myler, P.J., and Stuart, K. (2008). Trypanosoma brucei mitochondrial ribosomes. Molecular and Cellular Proteomics *7*, 1286–1296.

107. Dean, S., Sunter, J.D., and Wheeler, R.J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. Trends in Parasitology *33*, 80–82.

108. D'Archivio, S., and Wickstead, B. (2017). Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. Journal of Cell Biology *216*, 379–391.

109. 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. Molecular Microbiology *100*, 457–471.

110. Estévez, A.M., Kempf, T., and Clayton, C. (2001). The exosome of Trypanosoma brucei. EMBO Journal *20*, 3831–3839.

111. Estévez, A.M., Lehner, B., Sanderson, C.M., Ruppert, T., and Clayton, C. (2003). The roles of intersubunit interactions in exosome stability. Journal of Biological Chemistry *278*, 34943–34951.

112. Akiyoshi, B., and Gull, K. (2014). Discovery of unconventional kinetochores in kinetoplastids. Cell *156*, 1247–1258.

113. Günzl, A. (2010). The pre-mRNA splicing machinery of trypanosomes: Complex or simplified? Eukaryotic Cell *9*, 1159–1170.

114. Cristodero, M., and Clayton, C.E. (2007). Trypanosome MTR4 is involved in rRNA processing. Nucleic Acids Research *35*, 7023–7030.

115. Fadda, A., Färber, V., Droll, D., and Clayton, C. (2013). The roles of 3'-exoribonucleases and the exosome in trypanosome mRNA degradation. RNA *19*, 937–947.

116. Lesénéchal, M., Becquart, L., Lacoux, X., Ladavière, L., Baida, R.C.P., Paranhos-Baccalà, G., and Da Silveira, J.F. (2005). Mapping of B-cell epitopes in a Trypanosoma cruzi immunodominant antigen expressed in natural infections. Clinical and Diagnostic Laboratory Immunology *12*, 329–333.

117. Gupta, S.K., Carmi, S., Ben-Asher, H.W., Tkacz, I.D., Naboishchikov, I., and Michaeli, S. (2013). Basal splicing factors regulate the stability of mature mRNAs in trypanosomes. Journal of Biological Chemistry *288*, 4991–5006.

118. Erben, E.D., Fadda, A., Lueong, S., Hoheisel, J.D., and Clayton, C. (2014). A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in Trypanosoma brucei. PLoS Pathogens *10*.

119. 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*.

120. Schwede, A., Ellis, L., Luther, J., Carrington, M., Stoecklin, G., and Clayton, C. (2008). A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. Nucleic Acids Research *36*, 3374–3388.

121. Singh, A., Minia, I., Droll, D., Fadda, A., Clayton, C., and Erben, E. (2014). Trypanosome MKT1 and the RNA-binding protein ZC3H11: Interactions and potential roles in post-transcriptional regulatory networks. Nucleic Acids Research *42*, 4652–4668.

122. Freire, E.R., Vashisht, A.A., Malvezzi, A.M., Zuberek, J., Langousis, G., Saada, E.A., Nascimento, J.D.F., Stepinski, J., Darzynkiewicz, E., Hill, K., De Melo Neto, O.P., Wohlschlegel, J.A., Sturm, N.R., and Campbell, D.A. (2014). eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in Trypanosoma brucei. Rna *20*, 1272–1286.

123. Bercovich, N., Levin, M.J., Clayton, C., and Vazquez, M.P. (2009). Identification of core components of the exon junction complex in trypanosomes. Molecular and Biochemical Parasitology *166*, 190–193.

124. de Melo Neto, O.P., da Costa Lima, T.D.C., Xavier, C.C., Nascimento, L.M., Romão, T.P., Assis, L.A., Pereira, M.M.C., Reis, C.R.S., and Papadopoulou, B. (2015). The unique Leishmania EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. RNA Biology *12*, 1209–1221.

125. Moura, D.M.N., Reis, C.R.S., Xavier, C.C., Costa Da Lima, T.D., Lima, R.P., Carrington, M., and Melo De Neto, O.P. (2015). Two related trypanosomatid eiF4G homologues have functional differences compatible with distinct roles during translation initiation. RNA Biology *12*, 305–319.

126. Yoffe, Y., Zuberek, J., Lerer, A., Lewdorowicz, M., Stepinski, J., Altmann, M., Darzynkiewicz, E., and Shapira, M. (2006). Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania. Eukaryotic Cell *5*, 1969–1979.

127. da Costa Lima, T.D., Moura, D.M.N., Reis, C.R.S., Vasconcelos, J.R.C., Ellis, L., Carrington, M., Figueiredo, R.C.B.Q., and de Melo Neto, O.P. (2010). Functional characterization of three Leishmania poly(A) binding protein homologues with distinct binding properties to RNA and protein partners. Eukaryotic Cell *9*, 1484–1494.

128. Zinoviev, A., Léger, M., Wagner, G., and Shapira, M. (2011). A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways. Nucleic Acids Research *39*, 8404–8415.

129. Reijns, M.A.M., Alexander, R.D., Spiller, M.P., and Beggs, J.D. (2009). UKPMC Funders Group Author Manuscript UKPMC Funders Group Author Manuscript A role for Q / N-rich aggregation-prone regions in P-body localization. Cell *121*, 2463–2472.

130. Michelitsch, M.D., and Weissman, J.S. (2000). A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions. Proceedings of the National Academy of Sciences of the United States of America *97*, 11910–11915.

131. Saavedra, C., Tuug, K.S., Amberg, D.C., Hopper, A.K., and Cole, C.N. (1996). Regulation of mRNA export in response to stress in Saccharomyces cerevisiae. Genes and Development *10*, 1608–1620.

132. Tani, T., Derby, R.J., Hiraoka, Y., and Spector, D.L. (1996). *Errata Nucleolar Accumulation of Poly (A)+ RNA in Heat-shocked Yeast Cells: Implication of Nucleolar Involvement in mRNA Transport*.

133. Liu, Y., Liang, S., and Tartakoff, A.M. (1996). Heat shock disassembles the nucleolus and inhibits nuclear protein import and poly(A)+ RNA export. EMBO Journal *15*, 6750–6757.

134. Collins, M.L., Irvine, B., Tyner, D., Fine, E., Zayati, C., Chang, C.A., Horn, T., Ahle, D., Detmer, J., Shen, L.P., Kolberg, J., Bushnell, S., Urdea, M.S., and Ho, D.D. (1997). A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. Nucleic Acids Research *25*, 2979–2984.

135. Kern, D., Collins, M., Fultz, T., Detmer, J., Hamren, S., Peterkin, J.J., Sheridan, P., Urdea, M., White, R., Yeghiazarian, T., and Todd, J. (1996). An enhanced-sensitivity branched-DNA assay for quantification of human immunodeficiency virus type 1 RNA in plasma. Journal of Clinical Microbiology *34*, 3196–3202.

136. Strome, S. (2005). Specification of the germ line. WormBook : the online review of C. elegans biology, 1–10doi:10.1895/wormbook.1.9.1.

137. Hubert, A., and Anderson, P. (2009). The C. elegans sex determination gene laf-1 encodes a putative DEAD-box RNA helicase. Developmental Biology *330*, 358–367.

138. Gallo, C.M., Munro, E., Rasoloson, D., Merritt, C., and Seydoux, G. (2008). Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos. Developmental Biology *323*, 76–87.

139. Boag, P.R., Nakamura, A., and Blackwell, T.K. (2005). A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in C. elegans. Development *132*, 4975–4986.

140. Audhya, A., Hyndman, F., McLeod, I.X., Maddox, A.S., Yates, J.R., Desai, A., and Oegema, K. (2005). A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in Caenorhabditis elegans. Journal of Cell Biology *171*, 267–279.

141. Navarro, R.E., Shim, E.Y., Kohara, Y., Singson, A., and Blackwell, T.K. (2001). cgh-1, a conserved predicted RNA helicase required for gametogenesis and protection from physiological germline apoptosis in C. elegans. Development *128*, 3221–3232.

142. Amiri, A., Keiper, B.D., Kawasaki, I., Fan, Y., Kohara, Y., Rhoads, R.E., and Strome, S. (2001). An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in C. elegans. Development *128*, 3899–3912.

143. Boag, P.R., Atalay, A., Robida, S., Reinke, V., and Blackwell, T.K. (2008). Protection of specific maternal messenger RNAs by the P body protein CGH-1 (Dhh1/RCK) during Caenorhabditis elegans oogenesis. Journal of Cell Biology *182*, 543–557.

144. Schisa, J.A. (2012). *New Insights into the Regulation of RNP Granule Assembly in Oocytes*. International Review of Cell and Molecular Biology *295*, Elsevier Inc., 233–289pp.

145. Minshall, N., Kress, M., Weil, D., and Standart, N. (2009). Role of p54 RNA helicase activity and its c-terminal domain in translational repression, p-body localization and assembly. Molecular Biology of the Cell *20*, 2464–2472.

146. Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., Tani, T., Horinouchi, S., and Yoshida, M. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nature Chemical Biology *3*, 576–583.

147. Zorio, D.A.R., Cheng, N.N., Blumenthal, T., and Spieth, J. (1994). Operons as a common form of chromosomal organization in C. elegans. Nature *372*, 270–272.

148. Shiimori, M., Inoue, K., and Sakamoto, H. (2013). A Specific Set of Exon Junction Complex Subunits Is Required for the Nuclear Retention of Unspliced RNAs in Caenorhabditis elegans. Molecular and Cellular Biology *33*, 444–456.

149. Rajyaguru, P., and Parker, R. (2009). CGH-1 and the control of maternal mRNAs. Trends in Cell Biology *19*, 24–28.

150. Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. Cell *94*, 635–645.

151. Noble, S.L., Allen, B.L., Goh, L.K., Nordick, K., and Evans, T.C. (2008). Maternal mRNAs are regulated by diverse P body-related mRNP granules during early Caenorhabditis elegans development. Journal of Cell Biology *182*, 559–572.

152. Fernández-Moya, S.M., García-Pérez, A., Kramer, S., Carrington, M., and Estévez, A.M. (2012). Alterations in DRBD3 Ribonucleoprotein Complexes in Response to Stress in Trypanosoma brucei. PLoS ONE *7*, 1–10.

153. Cassola, A., and Frasch, A.C. (2009). An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein. Journal of Biological Chemistry *284*, 35015–35028.

154. Inoue, A.H., Serpeloni, M., Hiraiwa, P.M., Yamada-Ogatta, S.F., Muniz, J.R.C., Motta, M.C.M., Vidal, N.M., Goldenberg, S., and Ávila, A.R. (2014). Identification of a novel nucleocytoplasmic shuttling RNA helicase of trypanosomes. PLoS ONE *9*.

155. Ishigaki, Y., Li, X., Serin, G., and Maquat, L.E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. Cell *106*, 607–617.

156. Courchaine, E.M., Lu, A., and Neugebauer, K.M. (2016). Droplet organelles? The EMBO Journal *35*, 1603–1612.

157. Saavedra, C.A., Hammell, C.M., Heath, C. V., and Cole, C.N. (1997). Yeast heat shock mRNAs are exported through a distinct pathway defined by Rip1p. Genes and Development *11*, 2845–2856.

158. Keminer, O., and Peters, R. (1999). Permeability of single nuclear pores. Biophysical Journal *77*, 217–228.

159. Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: Definition, function, and interaction with importin α. Journal of Biological Chemistry *282*, 5101–5105.

160. Marchetti, M.A., Tschudi, C., Kwon, H., Wolin, S.L., and Ullu, E. (2000). Import of proteins into the trypanosome nucleus and their distribution at karyokinesis. Journal of Cell Science *113*, 899–906.

161. Hoek, M., Engstler, M., and Cross, G.A.M. (2000). Expression-site-associated gene 8 (ESAG8) of Trypanosoma brucei is apparently essential and accumulates in the nucleolus. Journal of Cell Science *113*, 3959–3968.

162. Westergaard, G.G., Bercovich, N., Reinert, M.D., and Vazquez, M.P. (2010). Analysis of a nuclear localization signal in the p14 splicing factor in Trypanosoma cruzi. International Journal for Parasitology *40*, 1029–1035.

163. Görlich, D. (1997). Nuclear protein import. Current Opinion in Cell Biology *9*, 412–419.

164. Kuersten, S., Ohno, M., and Mattaj, I.W. (2001). Nucleocytoplasmic transport: Ran, beta and beyond. Trends in Cell Biology *11*, 497–503.

165. Kudo, N., Matsumor, i N., Taoka, H., Fujiwara, D., Schreiner, E., Wolff, B., Yoshida, M., and Horinouchi, S. (1999). Leptomycin B inactivates CRM1 exportin 1 by covalent modification at a cysteine residue in the. Proc Natl Acad Sci USA *96*, 9112–9117.

166. Zeiner, G.M., Sturm, N.R., and Campbell, D.A. (2003). Exportin 1 mediates nuclear export of the kinetoplastid spliced leader RNA. Eukaryotic Cell *2*, 222–230.

167. Biton, M., Mandelboim, M., Arvatz, G., and Michaeli, S. (2006). RNAi interference of XPO1 and Sm genes and their effect on the spliced leader RNA in Trypanosoma brucei. Molecular and Biochemical Parasitology *150*, 132–143.

168. Seydoux, G., and Braun, R.E. (2006). Pathway to Totipotency: Lessons from Germ Cells. Cell *127*, 891–904.

169. Voronina, E. (2013). The diverse functions of germline P-granules in Caenorhabditis elegans. Molecular Reproduction and Development *80*, 624–631.

170. Souquere, S., Mollet, S., Kress, M., Dautry, F., Pierron, G., and Weil, D. (2009). Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. Journal of Cell Science *122*, 3619–3626.

171. Cougot, N., Cavalier, A., Thomas, D., and Gillet, R. (2012). The dual organization of P-bodies revealed by immunoelectron microscopy and electron tomography. Journal of Molecular Biology *420*, 17–28.

172. Kramer, S. (2017). Simultaneous detection of mRNA transcription and decay intermediates by dual colour single mRNA FISH on subcellular resolution. Nucleic Acids Research *45*, 1–12.

173. Dziembowski, A., Ventura, A.P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprévote, O., and Séraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. EMBO Journal *23*, 4847–4856.

174. Palancade, B., Zuccolo, M., Loeillet, S., Nicolas, A., and Doye, V. (2005). Pml39, a novel protein of the nuclear periphery required for nuclear retention of improper messenger ribonucleoparticles. Molecular Biology of the Cell *16*, 5258–5268.

175. Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear Retention of Unspliced mRNAs in Yeast Is Mediated by Perinuclear Mlp1. Cell *116*, 63–73.

176. Schmitt, C., Von Kobbe, C., Bachi, A., Panté, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M., Séraphin, B., Carmo-Fonseca, M., and Izaurralde, E. (1999). Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO Journal *18*, 4332–4347.

177. Hurwitz, M.E., Strambio-De-Castillia, C., and Blobel, G. (1998). Two yeast nuclear pore complex proteins involved in mRNA export form a cytoplasmically oriented subcomplex. Proceedings of the National Academy of Sciences of the United States of America *95*, 11241–11245.

178. Folkmann, A.W., Noble, K.N., Cole, C.N., and Wente, S.R. (2011). Dbp5, Gle1-IP6 and Nup159: A working model for mRNP export. Nucleus *2*, 540–548.

179. Legrain, P., and Rosbash, M. (1989). Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. Cell *57*, 573–583.

180. Yost, H.J., and Lindquist, S. (1988). Translation of unspliced transcripts after heat shock. Science *242*, 1544–1548.

181. Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.H., Grierson, D.S., Bertrand, E., Tazi, J., and Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. Journal of Cell Biology *178*, 1145–1160.

182. Kramer, S. (2021). Nuclear mRNA maturation and mRNA export control: from trypanosomes to opisthokonts. Parasitology, 1-97. doi:10.1017/S0031182021000068 (in press)

183. Fadda, A., Ryten, M., Droll, D., Rojas, F., Färber, V., Haanstra, J.R., Merce, C., Bakker, B.M., Matthews, K., and Clayton, C. (2014). Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. Molecular Microbiology *94*, 307–326.

184. Gudipati, R.K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L.M., Jacquier, A., and Libri, D. (2012). Extensive Degradation of RNA Precursors by the Exosome in Wild-Type Cells. Molecular Cell *48*, 409–421.

185. Schmid, M., and Jensen, T.H. (2018). Controlling nuclear RNA levels. Nature Reviews Genetics *19*, 518–529.

186. Tudek, A., Schmid, M., and Jensen, T.H. (2019). Escaping nuclear decay: the significance of mRNA export for gene expression. Current Genetics *65*, 473–476.

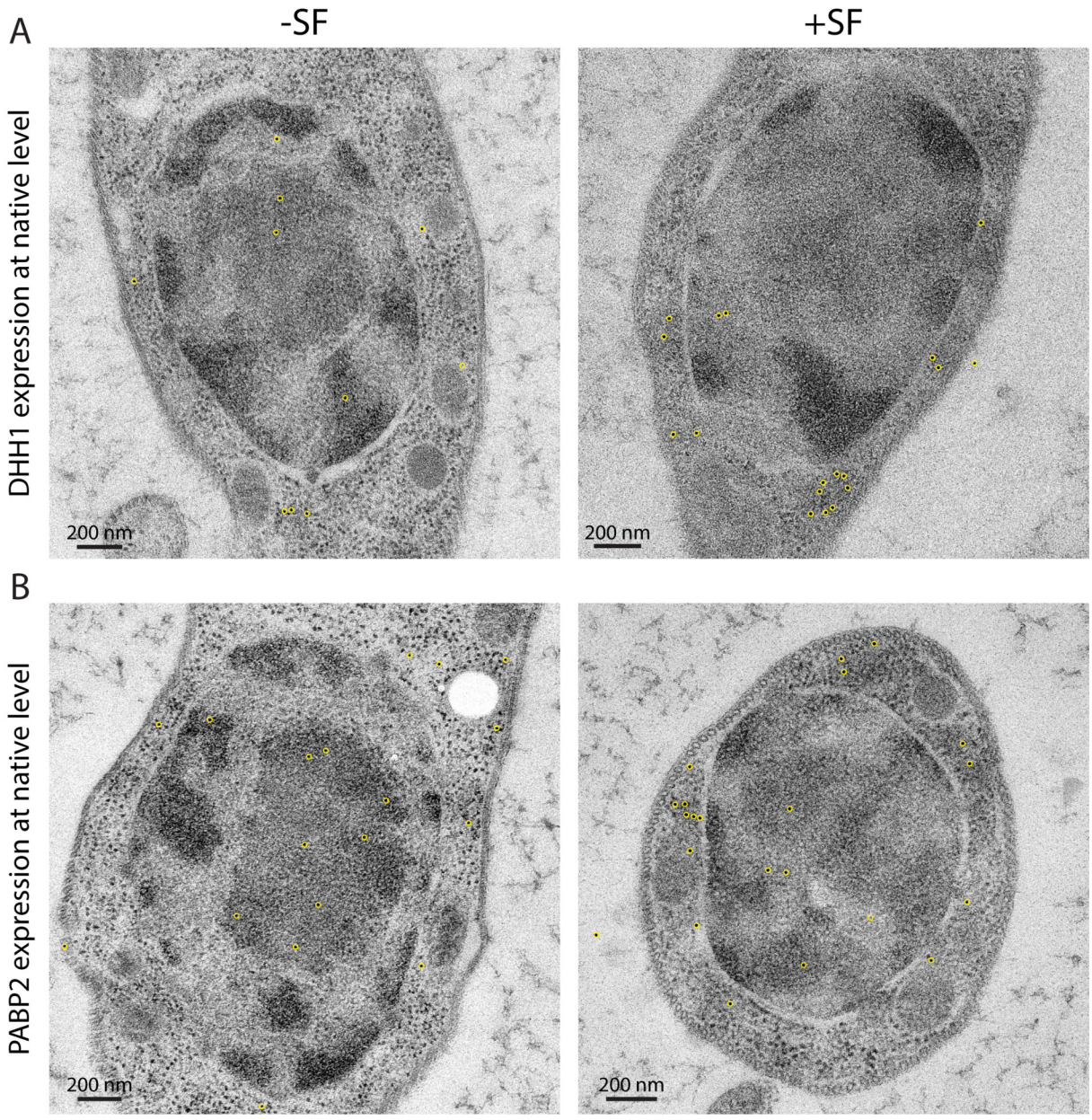
187. Pastro, L., Smircich, P., Di Paolo, A., Becco, L., Duhagon, M.A., Sotelo-Silveira, J., and Garat, B. (2017). Nuclear Compartmentalization Contributes to Stage-Specific Gene Expression Control in Trypanosoma cruzi. Frontiers in Cell and Developmental Biology *5*, 8.

188. Manful, T., Fadda, A., and Clayton, C. (2011). The role of the 5'-3' exoribonuclease XRNA in transcriptome-wide mRNA degradation. RNA *17*, 2039–2047.

189. Grünwald, D., and Singer, R.H. (2010). In vivo imaging of labelled endogenous B-actin mRNA during nucleocytoplasmic transport. Nature *467*, 604–607.

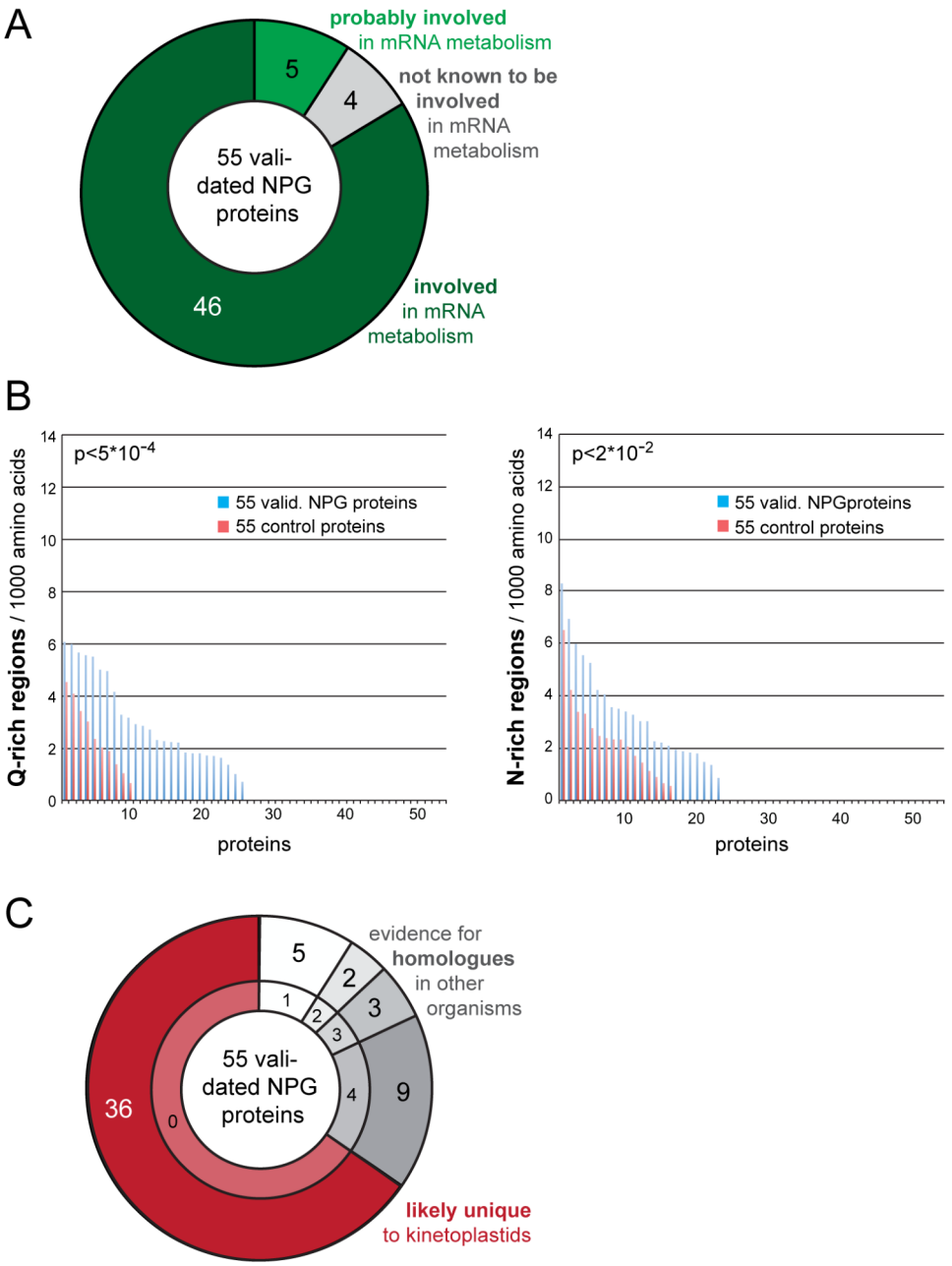
7 Supplementary Data

7.1 Supplementary Figures



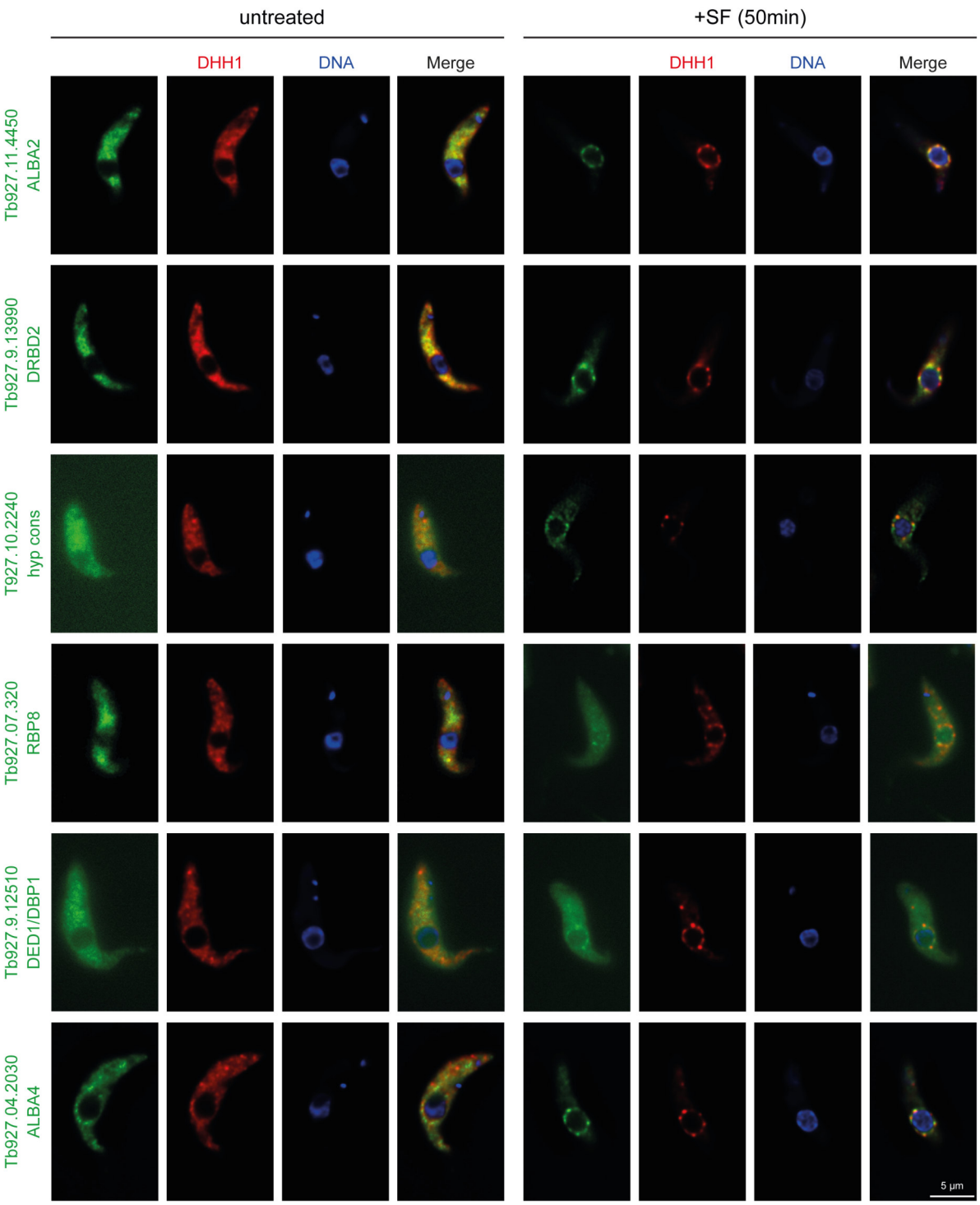
Sup. Figure 7.1: Immuno-gold labelling of DHH1 and PABP2 on EM sections.

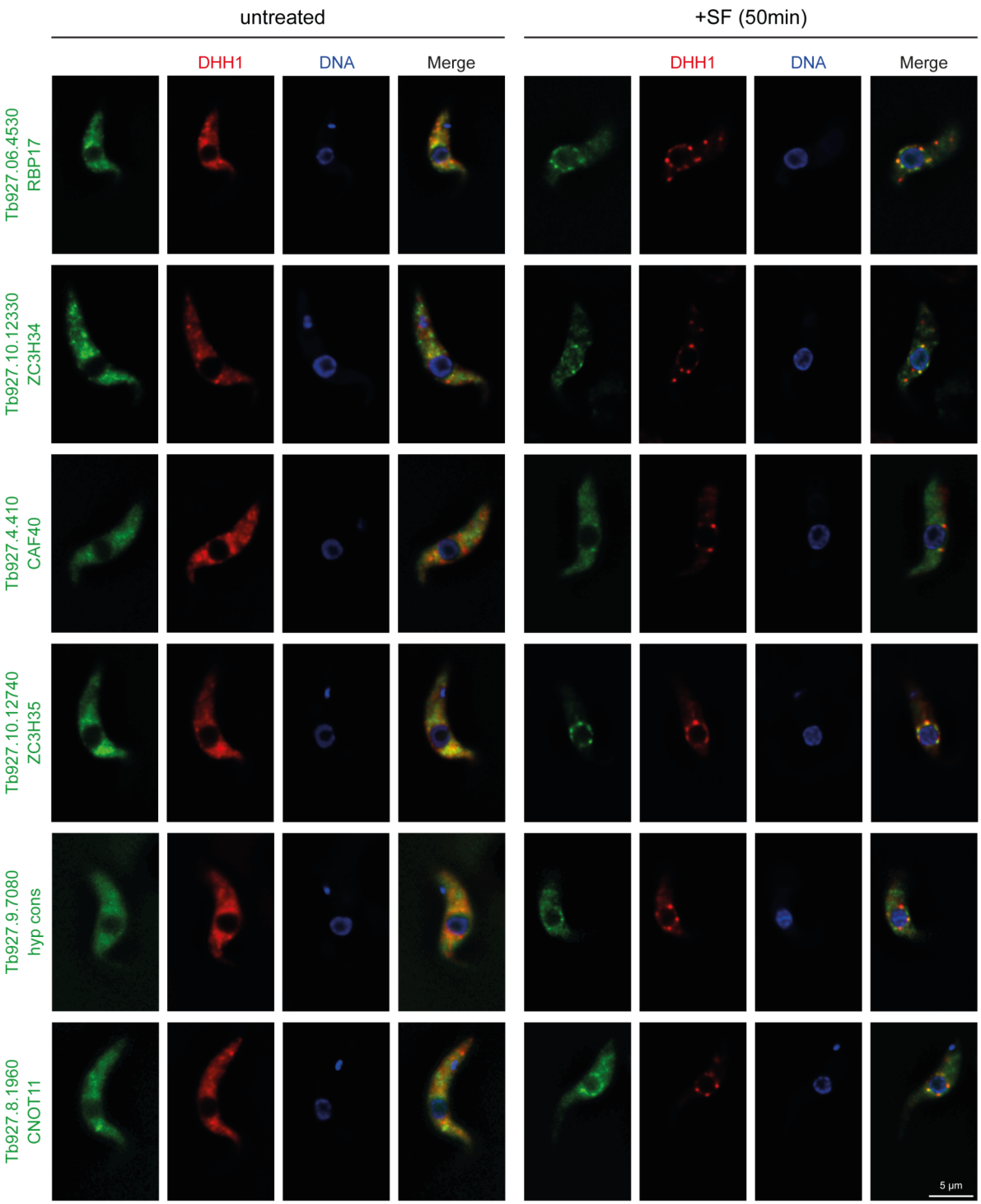
Representative images of a cell line with native expression of the NPG granule marker DHH1 **(A)** and a cell line expressing PABP2-4xTy-eYFP **(B)** are shown for one untreated (-SF) and one sinefungin-treated (+SF) cell respectively. Cells were high pressure frozen, embedded in LR-White, cut into 60 nm thin sections and contrasted with aqueous uranyl acetate and Reynolds lead citrate. DHH1 was labelled with anti-DHH1 and 12 nm conjugated gold anti-rabbit, PABP2 was labelled with anti-Ty1 BB2 and 12 nm conjugated gold anti-mouse. For better visualization gold grains are marked with yellow circle.

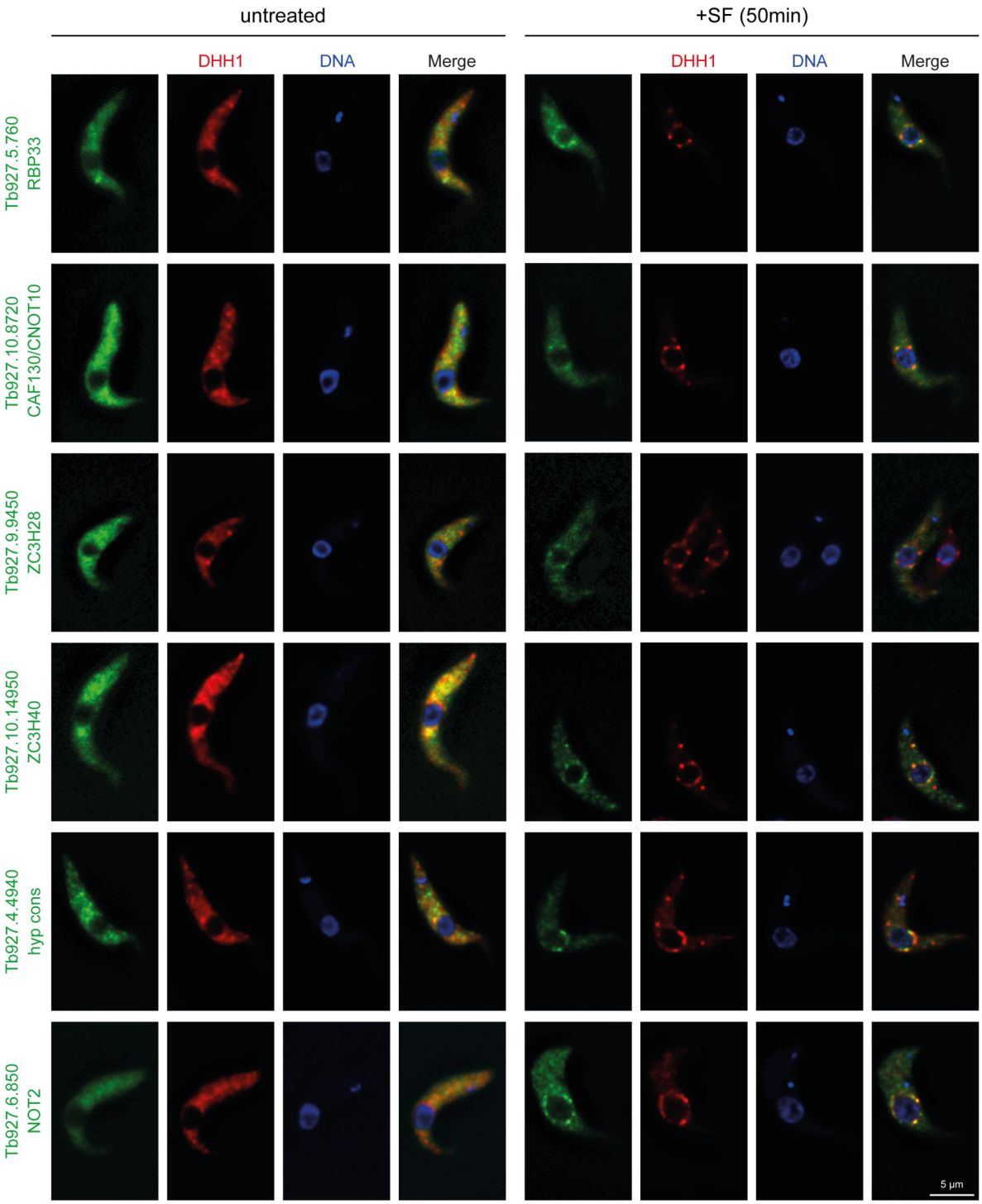


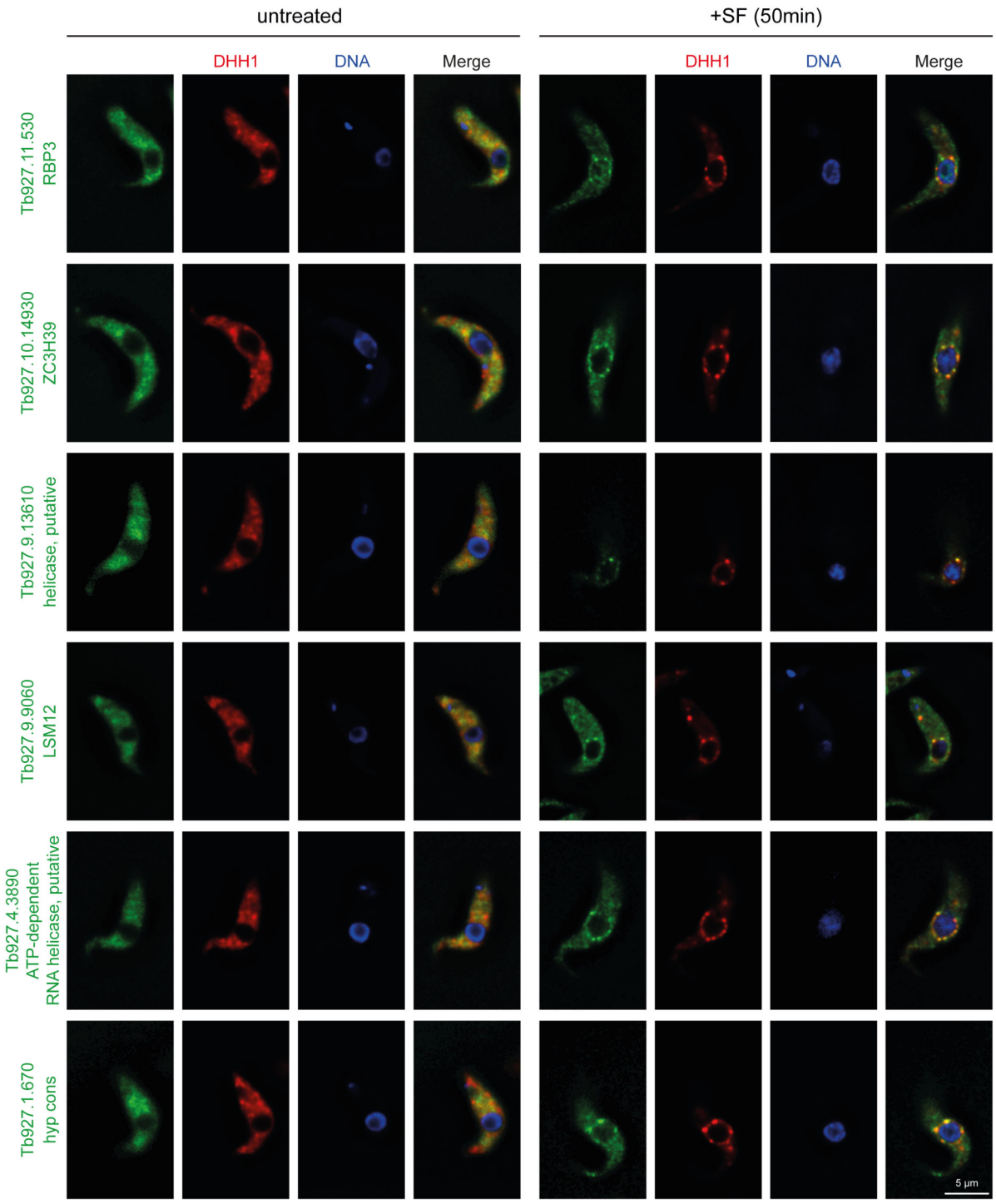
Sup. Figure 7.2: Characterization of validated NPG proteins

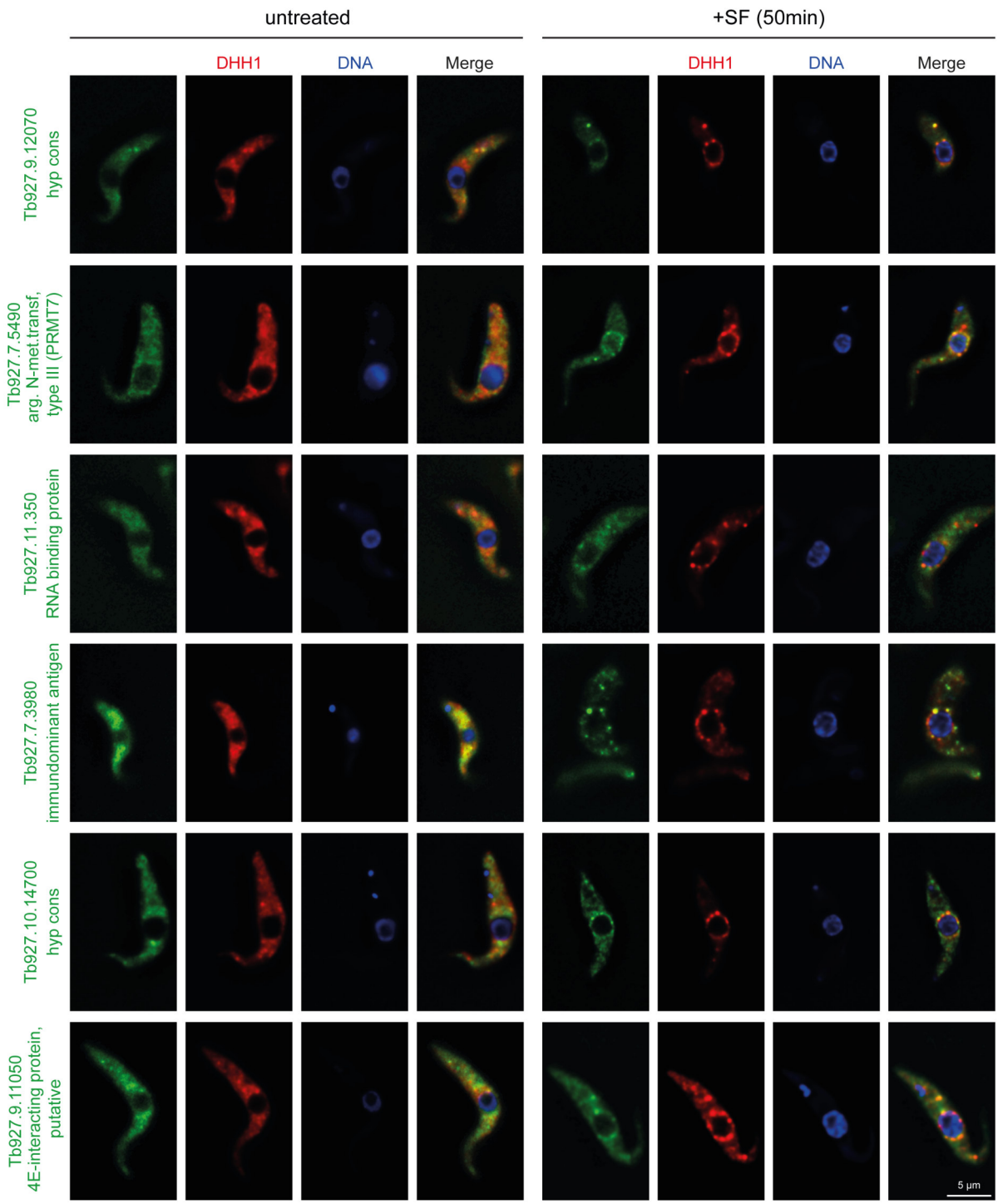
(A) The majority of NPG proteins are involved in mRNA metabolism. 55 NPG proteins were classified as involved (green), probably involved (light green) and not known to be involved (grey) in mRNA metabolism. Details of the analysis are depicted in Sup. Table 7.7 A. **(B)** Enrichment of Q-rich and N-rich regions. The number of Q-rich regions and N-rich regions per 1000 amino acids is shown for the 55 validated NPG proteins (blue) in comparison to a randomly chosen group of control proteins (red), respectively. One bar represents one protein. Proteins with no enriched regions are lacking a bar. **(C)** The majority of NPG proteins are likely unique to kinetoplastids. 55 NPG proteins were screened for putative homologues in *S. cerevisiae, C. elegans, M. musculus* and *P. falciparum* by BLAST (default parameters). The first hit of the BLAST analysis was then used for a reciprocal BLAST against the *T. brucei* genome. If the initial protein was again the first hit and had an e-value of $< 1x10^{-5}$, the protein was classified as having a homologue in the respective organism. Using these definitions, the majority of proteins were unique to kinetoplastids (*T. brucei*) (red). The remaining proteins had putative homologues in either one, two, three or four of the screened organisms (grey). Details of the analysis are shown in Sup. Table 7.7 D.

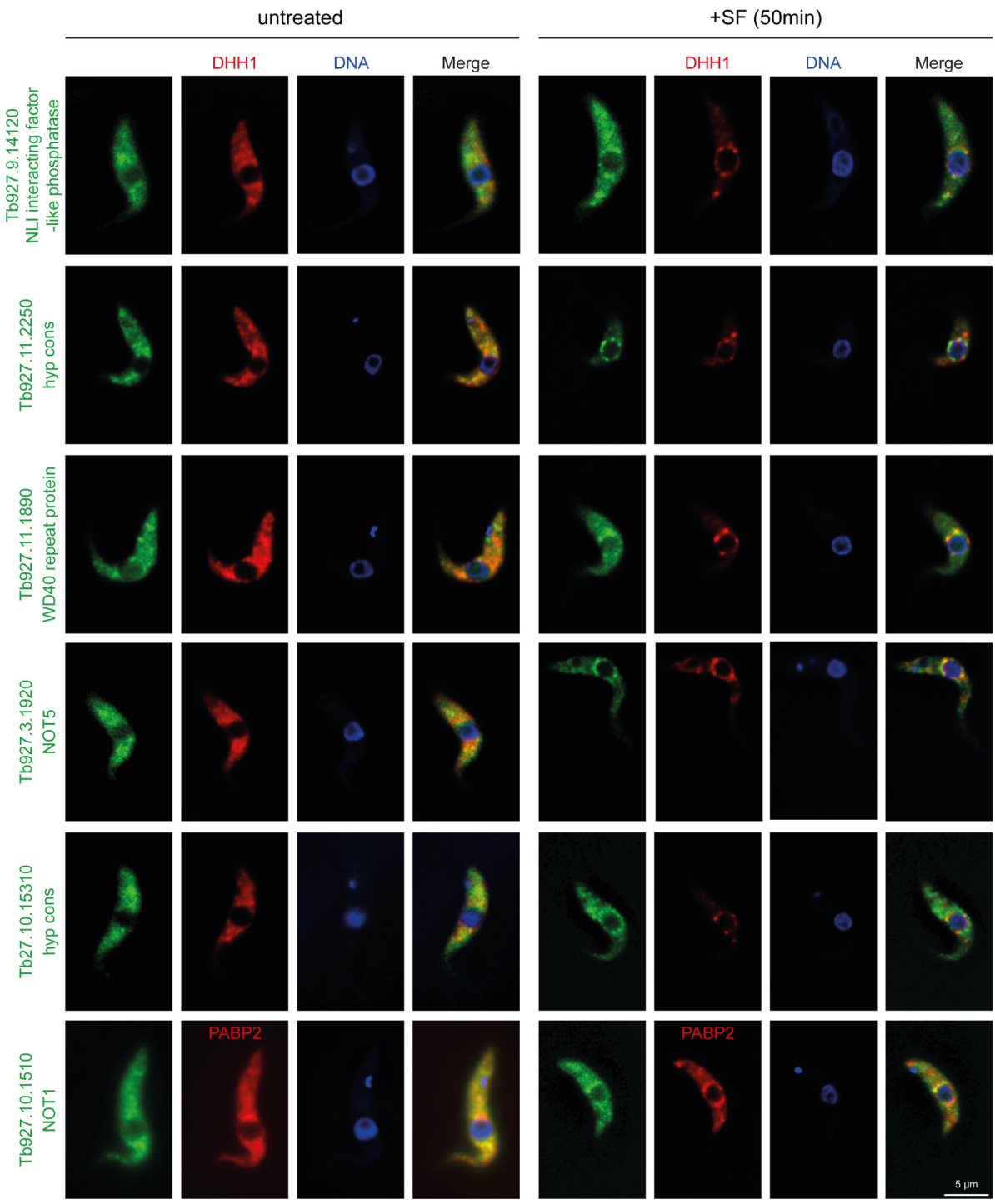

untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.11.4450
ALBA2
Tb927.9.13990
DRBD2
T927.10.2240
hyp cons
Tb927.07.320
RBP8
Tb927.9.12510
DED1/DBP1
Tb927.04.2030
ALBA4
5 µm

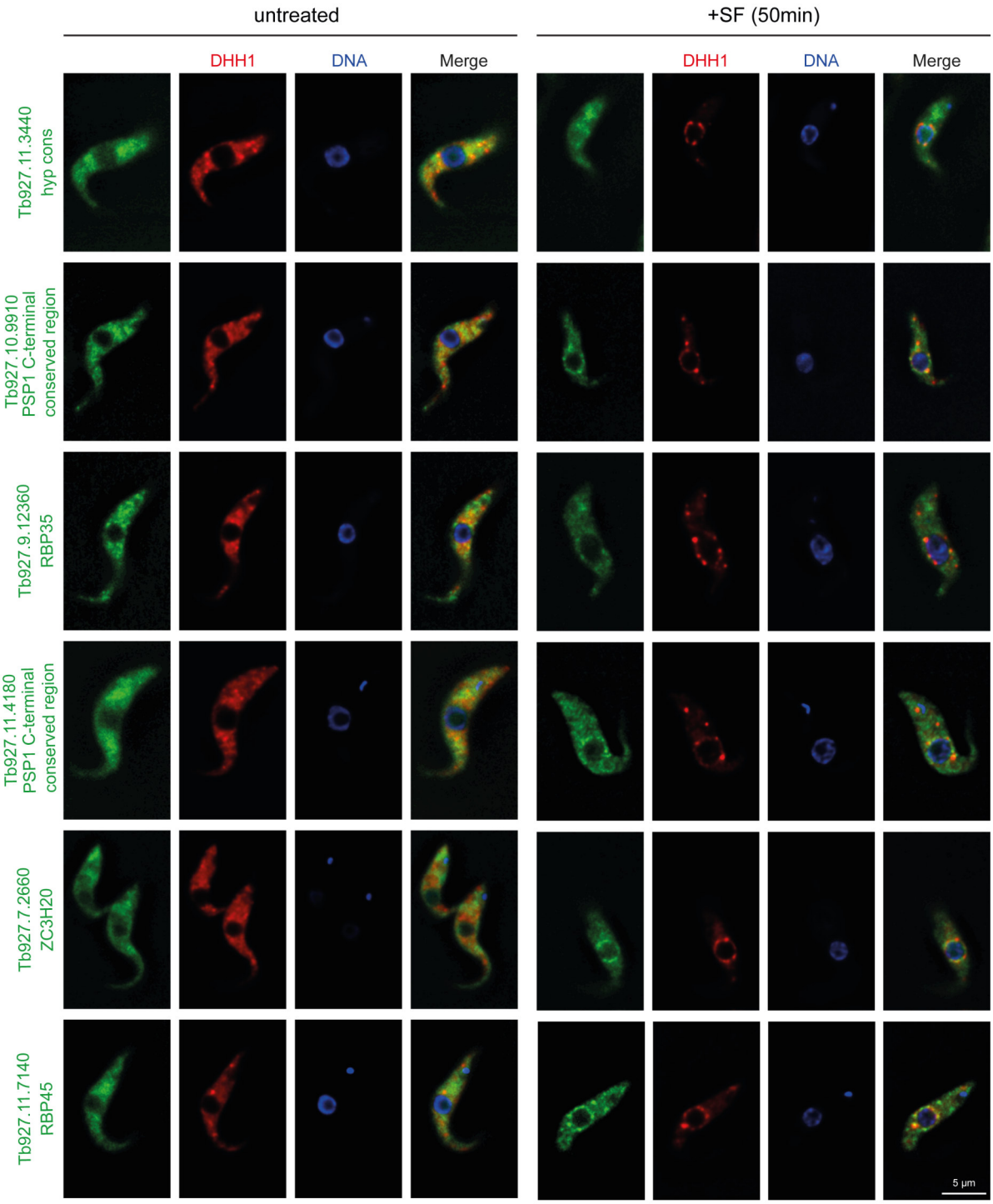

untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.06.4530
RBP17
Tb927.10.12330
ZC3H34
Tb927.4.410
CAF40
Tb927.10.12740
ZC3H35
Tb927.9.7080
hyp cons
Tb927.8.1960
CNOT11
5 µm

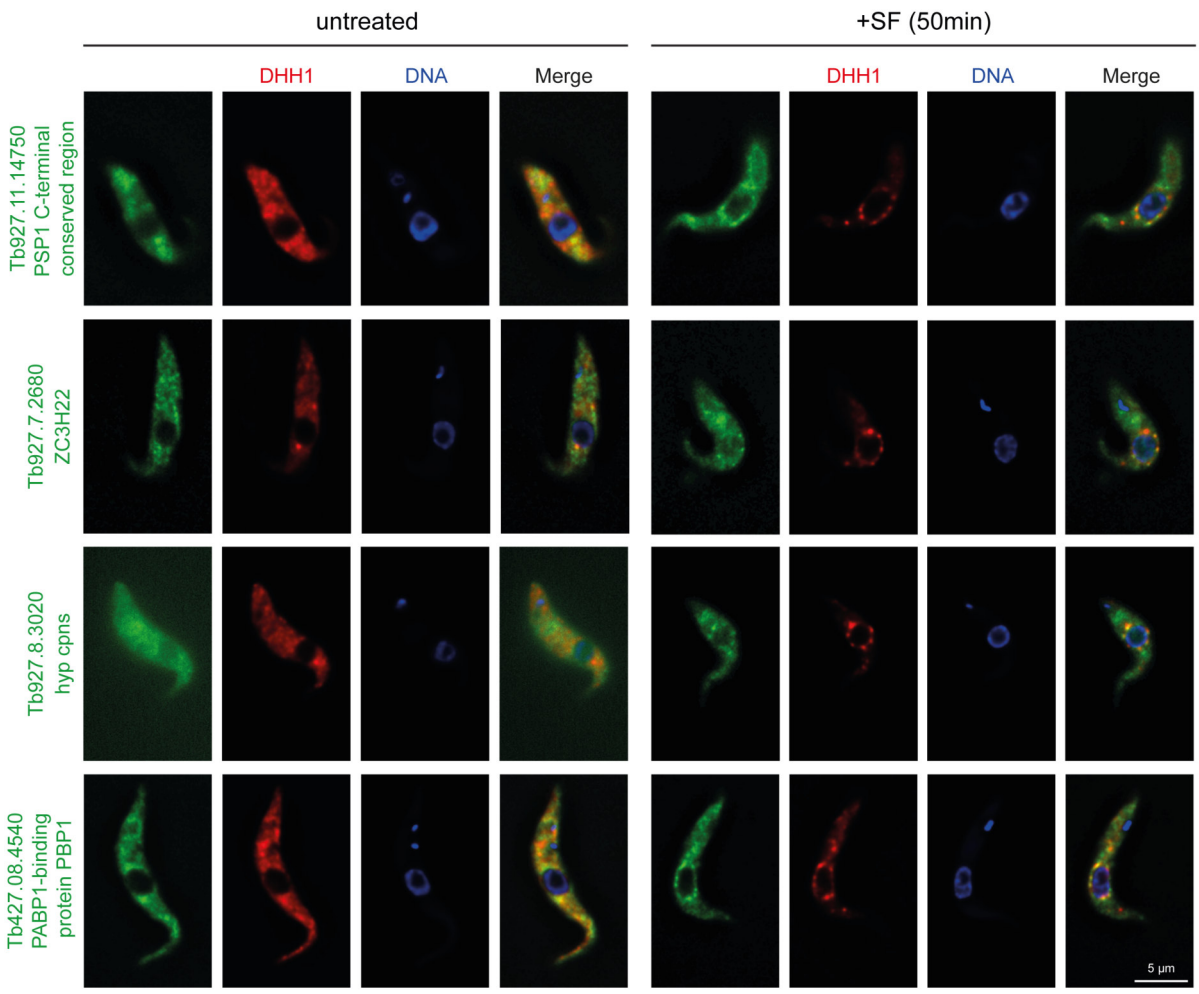

untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.5.760
RBP33
Tb927.10.8720
CAF130/CNOT10
Tb927.9.9450
ZC3H28
Tb927.10.14950
ZC3H40
Tb927.4.4940
hyp cons
Tb927.6.850
NOT2
5 µm


untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.11.530
RBP3
Tb927.10.14930
ZC3H39
Tb927.9.13610
helicase, putative
Tb927.9.9060
LSM12
Tb927.4.3890
ATP-dependent
RNA helicase, putative
Tb927.1.670
hyp cons
5 µm


untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.9.12070
hyp cons
Tb927.7.5490
arg. N-met.transf,
type III (PRMT7)
Tb927.11.350
RNA binding protein
Tb927.7.3980
immundominant antigen
Tb927.10.14700
hyp cons
Tb927.9.11050
4E-interacting protein,
putative
5 µm

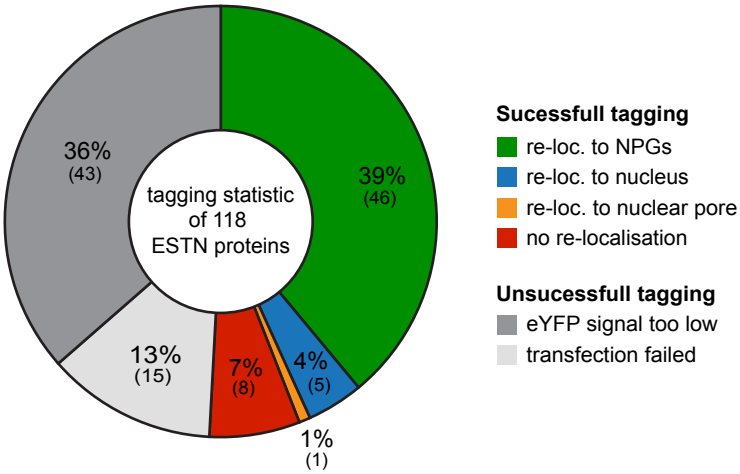

untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.9.14120
NLI interacting factor
-like phosphatase
Tb927.11.2250
hyp cons
Tb927.11.1890
WD40 repeat protein
Tb927.3.1920
NOT5
Tb27.10.15310
hyp cons
Tb927.10.1510
NOT1
PABP2
PABP2
5 µm


untreated
+SF (50min)
DHH1
DNA
Merge
DHH1
DNA
Merge
Tb927.11.3440
hyp cons
Tb927.10.9910
PSP1 C-terminal
conserved region
Tb927.9.12360
RBP35
Tb927.11.4180
PSP1 C-terminal
conserved region
Tb927.7.2660
ZC3H20
Tb927.11.7140
RBP45
5 µm



Sup. Figure 7.3: 46 ESTN proteins, which localize to NPGs upon sinefungin treatment.

All ESTN proteins were expressed as eYFP fusion proteins from the endogenous locus (green) in a cell line co-expressing mChFP-DHH1 (red) as NPG granule marker. DNA was stained with DAPI (blue). Images are either single plane images of a deconvolved z-stack or z-stack projections (further details in Sup. Table 7.7 A). Representative images of complete cells (treated and untreated with sinefungin) are shown.



Sup. Figure 7.4: Tagging statistics of all 118 ESTN proteins.

7.2 Supplementary Tables

All supplementary tables are attached to this work in digital form as CD.

Sup. Table 7.1: Oligonucleotides

All oligonucleotides used during this work for: validation of nuclear proteome (Chapter 4.2.3.3), validation of NPG proteome (Chapter 4.2.7), expressing *C. elegans* P granule proteins in trypanosomes (Chapter 4.3.3.1), RNAi of RNST proteins (Chapter 4.3.5) and other primers used for sequencing or during the cloning process.

Sup. Table 7.2: Plasmids

All plasmids used or generated during this work are listed. Details about the cloning strategies that were used to generate plasmids are described in Chapter 3.4

Sup. Table 7.3: Obtaining the nuclear proteome[91]

(A) List of all proteins that were detected by mass spectrometry. **(B)** List of the nuclear proteome. Further legend included in Excel file.

Sup. Table 7.4: Estimation of false positives[91]

Lists of non-nuclear proteins, including overlaps with the nuclear proteome **(A)** Lipid metabolism **(B)** Flagellum **(C)** Mitochondrion **(D)** Cilium transition zone and associated **(E)** Glycosome **(F)** Cell surface.

Sup. Table 7.5: Estimation of missing proteins[91]

Lists of nuclear proteins, including overlap with the nuclear proteome **(A)** Nuclear pores **(B)** Exosome **(C)** Kinetochores **(D)** Spliceosome.

Sup. Table 7.6: Localization of nuclear proteome candidates

List and additional information of the ten chosen nuclear proteome candidates to experimentally confirm their nuclear localization.

Sup. Table 7.7: The NPG proteome

(A) 128 ESTN proteins (NPG proteome) with details on analysis and classification. **(B)** List of underrepresented proteins **(C)** List of all proteins that were detected by mass spectrometry. **(D)** BLAST screen of 128 ESTN proteins for potential homologues in four eukaryotic model organisms. Further legends included in Excel file.

Sup. Table 7.8: BLAST Analysis - No homologies between NPGs and P granules

(A) 55 validated NPG candidates were searched for homologous proteins in *C. elegans*, including overlap with P granule protein content. **(B)** 31 characterized P granule proteins of *C. elegans* were screened for potential homologues in *T. brucei*, including overlap with localization to NPG.

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Prefix	Factor
p	pico-	10^{-12}
n	nano-	10^{-9}
µ	micro-	10^{-6}
m	milli-	10^{-3}
c	centi-	10^{-2}
k	kilo-	10^{3}

°C	degrees Celsius
bp	base pair
BSA	bovine serum albumin
BSF	bloodstream form
cDNA	coding deoxyribonucleic acid
Da	Dalton
DAPI	4’,6-diamidino-2-phenylindole
ddH_2O	double-distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
EDTA	ethylenediaminetetraacetic acid
e.g.	exempli gratia (Latin “for example”)
EM	electron microscopy
ESTN	enriched in sinefungin-treated nuclei
et al.	et alteri; and others
eYFP	enhanced yellow fluorescent protein
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
FISH	fluorescence in situ hybridization
fw	forward
g	gram

GFP	green fluorescent protein
h	hours
HS	heat shock
i.e.	id est (Latin “that is”)
IF	immunofluorescence
l	liter
LFQ	label-free quantification
M	molarity
m	meter
mChFP	mCherry fluorescent protein
min	minutes
mRNA	messenger ribonucleic acid
NES	nuclear enrichment score/ nuclear export sequence
NLS	nuclear localization signal
NPC	nuclear pore complex
NPG	nuclear periphery granule
nt	nucleotide
PBS	phosphate-buffered saline
PCF	procyclic form
PCR	polymerase chain reaction
PDT	population doubling time
PFA	paraformaldehyde
RI	refractive index
RNA	ribonucleic acid
RNAi	RNA-interference
RNase	ribonuclease
RNP	ribonucleoprotein
RNST	re-localizing to nucleus/nuclear pores upon sinefungin treatment
rpm	revolutions per minute
RT	room temperature
rv	reverse
SDM-79 w.s.h.	SDM-79 without serum and hemin
sec	seconds

SF	sinefungin
Tet	tetracycline
V	volt
VSG	variant surface protein
v/v	volume/volume %
WB	western blot
wt	wildtype
w/v	weight/volume %
x g	fold gravitational acceleration (g = 9.81 m/s^2)

Acknowledgement

An dieser Stelle möchte ich mich bei allen Begleitern während meiner Doktorarbeit bedanken.

Als allererstes möchte ich mich von Herzen bei **Dr. Susanne Kramer** bedanken. Es hat mir große Freude bereitet meine Doktorarbeit in deiner Arbeitsgruppe zu schreiben und dich als deine erste Doktorandin beim Aufbau einer Arbeitsgruppe zu unterstützen. Vielen Dank für die vielen bereichernden, fachlichen wie auch freundschaftlichen Gespräche und deine intensive und engagierte Betreuung. Deine gewissenhafte Art sich mit wissenschaftlichen Themen auseinanderzusetzen hat mich grundlegend geprägt und ich habe die stets vertrauensvolle Zusammenarbeit sehr genossen.

Prof. Dr. Thomas Dandekar danke ich für die bereitwillige Übernahme des Zweitgutachtens.

Ein ganz besonderer Dank geht an **Prof. Dr. Markus Engstler**, der als Leiter des Lehrstuhls immer die AG Kramer und damit auch mich nach allen Kräften gefördert hat. Du hast ein Umfeld zur Verfügung gestellt, in dem ich mich als Wissenschaftlerin entwickeln konnte. Vielen Dank für die stets hilfsbereite und konstruktive Zusammenarbeit und deine uneingeschränkte Unterstützung.

Dr. Falk Butter und **Dr. Mario Dejung** danke ich für alle massenspektrometrischen Analysen.

Dankeschön an **Prof. Dr. Christian Stigloher** und das gesamte Team der Elektronenmikroskopie, sowie an **Dr. Ann Wehman** für die mir zur Verfügung gestellte Zeit, Expertise und Hilfe.

Bei **allen Freunden und Kollegen des Lehrstuhls** möchte ich mich von Herzen für ihre Hilfsbereitschaft, viele kurzweilige und lustige Momente und die tolle gemeinsame Zeit bei der Arbeit sowie nach Feierabend bedanken. Danke **Marie, Irene** und **Erick**, dass ihr so viel mehr als Kollegen wart.

Janne, Susanne, Brooke, Ines, Tim, Manfred und Nicola danke ich für viele wertvolle Diskussionen, hilfreiche Anregungen und Rat bei jeglichen technischen und wissenschaftlichen Problemen.

Ein großes Dankeschön geht an alle guten Seelen des Lehrstuhls, insbesondere an **Uli**, **Lidia** und **Patrick**, weil ihr einfach jeden Tag den Laden am Laufen haltet. **Elisabeth** und **Beate**, bei euch bedanke ich mich ganz besonders für eure tatkräftige Unterstützung in allen Laborangelegenheiten.

Danke auch an alle anderen **aktuellen und ehemaligen Teammitglieder** für eine unvergessliche Zeit.

Mein ganz besonderer Dank geht an meine **Familie und Freunde** für ihre bedingungslose und uneingeschränkte Unterstützung. Danke, dass ihr mir immer euer Vertrauen schenk und mich in jeder Lebenslage begleitet und unterstützt.

Affidavit

I hereby confirm that my thesis entitled ‘‘Nuclear periphery granules of trypanosomes - A characterization of composition and function’’ is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not been submitted as part of another examination process neither in identical nor in similar form.

_____ _____

Place, date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Kern-assoziierte RNA Granula in Trypanosomen - Eine Untersuchung von Struktur und Funktion“ eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

_____ _____

Ort, Datum Unterschrift

Erklärungen nach §4 Abs. 3 Satz 3, 5, 8 der Promotionsordnung der Fakultät für Biologie