Expression of normal levels of the Pumilio domain protein PUF3 is required for optimal growth of *Trypanosoma brucei*.

Kevin Kamanyi Marucha and Christine Clayton*

Centre for Molecular Biology of Heidelberg University (ZMBH), DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany

*to whom correspondence should be addressed. e-mail cclayton@zmbh.uni-heidelberg.de

Abstract

The *Trypanosoma brucei* pumilio domain protein PUF3 is a cytosolic mRNA-binding protein that suppresses expression when tethered to a reporter mRNA. PUF3 was not essential in bloodstream-form or procyclic-form trypanosomes, and RNAi in bloodstream forms had no significant effect on the transcriptome. There was little evidence for specific binding of PUF3 to bloodstream-form mRNAs and mass spectrometry revealed no binding partners that might explain its suppressive activity. Ectopic expression of C-terminally tagged PUF3 in procyclic forms impaired viability. Reduction of PUF3 in bloodstream forms caused a slight growth defect and slightly delayed differentiation to the procyclic form, but the cells lost both defects upon prolonged cultivation. Procyclic forms without PUF3 grew somewhat slower than wild-type, but were able to transform to bloodstream forms after induced expression of the bloodstream-form RNA-binding protein RBP10. We suggest that PUF3 is implicated in fine-tuning gene expression. Since PUF3 is present throughout the Kinetoplastida, selection for its retention may occur within the arthropod vector.

Introduction

Pumilio domain proteins contain multiple "Puf" repeats, each of which can bind to a single nucleotide in RNA. Classically, there are 8 adjoined repeats, which together bind to a sequence with consensus of UGUR [1]. *Trypanosoma brucei* is a unicellular eukaryote of the Excavata supergroup. It multiplies as the bloodstream form in man and ruminants, and as the procyclic form in the midgut of tsetse flies. The genome of *T.*

brucei encodes eleven PUF proteins with different domain architectures. Only PUF2, 3, 4, and 6 have the canonical 8 pumilio repeats, and all but PUF3 (Tb927.10.310) have low complexity or coiled coil motifs. Most of the PUF proteins are cytosolic; the exceptions are PUF7, PUF8 and PUF10, which are found in the nucleolus [2]. All except PUF11 and PUF8 are conserved in Kinetoplastea, which includes not only in parasitic kinetoplastida but also free-living bodonids (S1 Fig A). Conservation between PUF3 homologues is restricted to the eight repeats, a 17-residue region immedietely C-terminal to the last repeat, and two short regions at the N-terminus (S1 Fig B).

We here describe experiments to investigate the function of *T. brucei* PUF3. PUF3 was already known to bind to mRNA in *T. brucei* bloodstream forms [3] and to repress expression when tethered (via a lambdaN peptide) to the 3'-untranslated region of a reporter mRNA [3, 4]. Results of a high-throughput RNAi screen also suggested that reduced PUF3 expression caused a significant loss of fitness in all life-cycle stages tested (bloodstream forms, and early differentiation to the procyclic form). Our results show that PUF3 is not essential in cultured *T. brucei*, but it may have a role during differentiation.

Results

PUF3 is cytosolic and rapidly represses expression

To determine the cellular location of PUF3, we used bloodstream-form trypanosomes expressing N-terminally V5-tagged PUF3 from the endogenous locus (V5-PUF3) or with inducible expression of C-terminally myc-tagged PUF3 (PUF3-myc) (S2 Fig, S1 Table). By immonofluorescence microscopy, V5-PUF3 was spread throughout the cytosol with no signal from the nucleus (Fig 1A). Cell fractionation using digitonin confirmed that PUF3-myc was not inside mitochondria or glycosomes (Fig 1B). Later, results from high-throughput N-terminal GFP tagging [2] also indicated a cytosolic location.

Cytosolic phosphoglycerol kinase (PGKB) is toxic to bloodstream-form trypanosomes. Tethering of full-length PUF3 to an mRNA encoding PGKB conferred a selective advantage [3]. Using random fragments of the *PUF3* gene [4], most selected clones encoded tethered polypeptides that commenced within the first 100 amino acids (Fig 1C). This result suggests that the suppressive sequence is N-terminal to the PUF repeats. We confirmed that tethering of lambdaN-PUF3-myc to a reporter mRNA encoding chloramphenical acetyltransferase (CAT) was repressive: CAT activity decreased within 4 hours (Fig 1D).

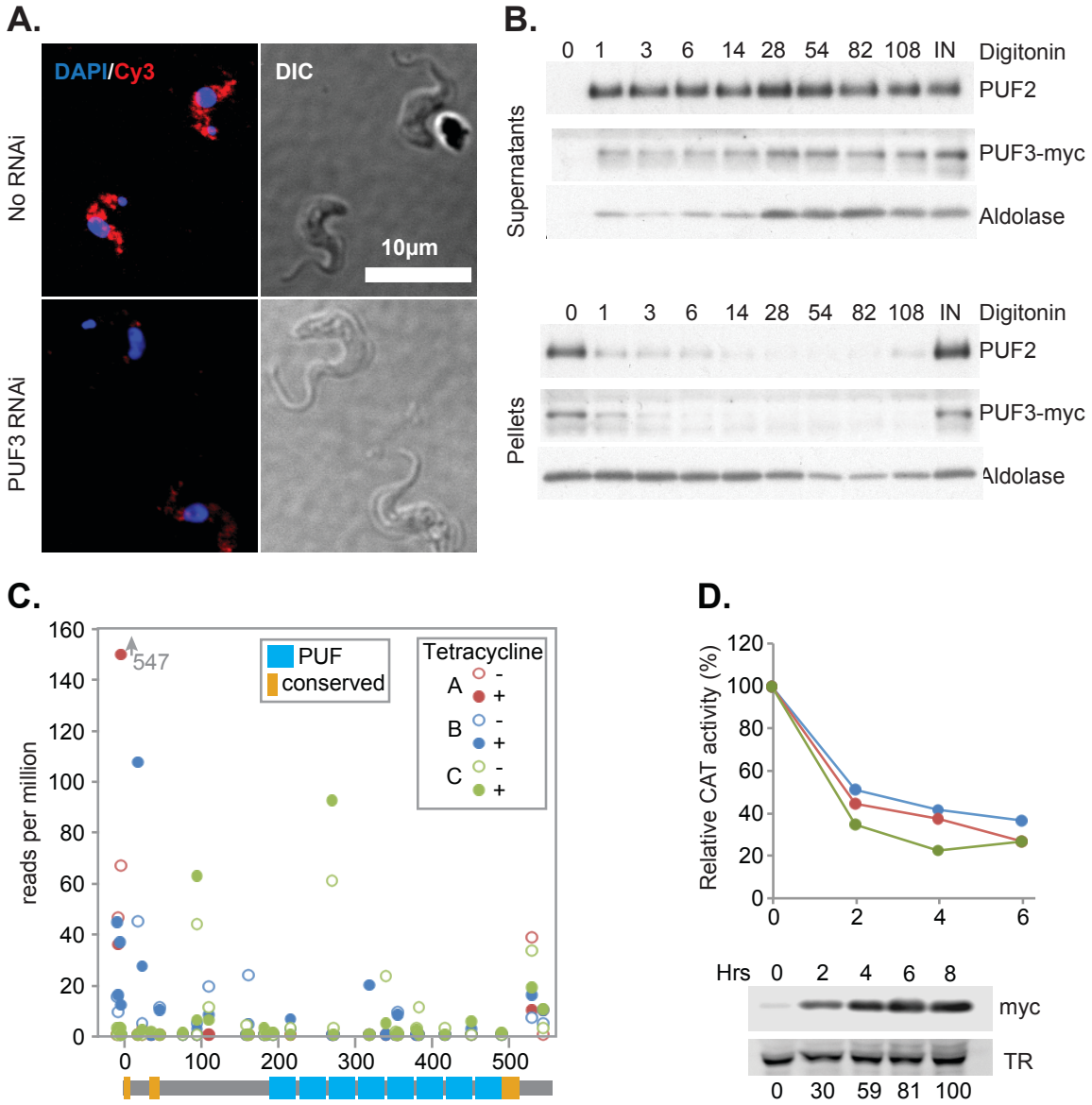

Figure 1

PUF3 location and tethering.

A) Bloodstream-form trypanosomes with in situ tagged PUF3 and inducible RNAi targeting *PUF3* were used. The upper image shows cells without tetracycline and the lower image after 24h tetracycline treatment. Stains are for V5-PUF3 (red) and DNA (blue). DIC is the differential interference contrast.

B) Cells expressing inducible PUF3-myc (24 h tetracycline) were lysed using increasing concentrations of digitonin. Organelle rupture was detected by the migration of the organelle's proteins to the supernatant. PUF2 and aldolase were used as cytosolic and glycosomal markers respectively. The amounts of digitonin (mg of digitonin/ mg of protein) are indicated above each lane.

C) Results of the tethering screen. The bloodstream-form trypanosomes used contained a tetracycline-inducible *PGKB* gene, with five copies of boxB between the open reading frame and the actin 3'-untranslated region (3'-UTR). They were transfected with a library designed for inducible expression of proteins encoded by ~1kb genomic fragments, preceded by the lambdaN peptide. Tetracycline addition resulted in growth retardation, unless the expressed lambdaN-protein repressed PGKB expression. After induction, cells were grown, then the DNA inserts were amplified and sequenced [4]. This selects for protein sequences capable of suppressing translation and/or inducing mRNA degradation. The graph shows the results for PUF3, with first codon positions on the x-axis, and the reads per million on the Y-axis. Higher values with tetracycline than without tetracycline suggest suppressive activity. The cartoon below shows domains of PUF3, to scale: PUF domains in cyan, and sequences conserved in Kinetoplastea are in orange (S1 Fig B).

D) Confirmation of tethering results for full-length lambdaN-PUF3-myc. The bloodstream-form trypanosomes used constitutively express a chloramphenicol acetyltransferase (*CAT*) mRNA with five B-boxes between the open reading frame and the actin 3'-UTR. Expression of lambdaN-PUF3-myc was induced by addition of tetracycline and CAT activity was measured. The graph shows results for 3 independent cultures.

PUF3 is not required in growing bloodstream forms.

To assess the function of PUF3 we first reduced its expression by RNAi. We used two trypanosome cell lines expressing the tet repressor. One, of the EATRO1125 strain, is capable of full differentiation from the bloodstream to the procyclic form. The EATRO1125 bloodstream forms were routinely grown in methyl cellulose-containing medium to maintain differentiation competence. When they reach a density of 2.5-3.0 x 10^6/ml, they arrest in G_0, and become "stumpy" - that is, they are shorter and fatter, and express the stumpy-form marker PAD1 and various mitochondrial proteins required for survival in tsetse. When the growth temperature is reduced from 37°C to 27°C and cis-aconitate is added, they differentiate to procyclic forms and start to grow in appropriate medium within 1-2 days. Our Lister 427 strain bloodstream forms, in contrast, can grow to densities of 6-7 x10^6/ml and never make stumpy forms or PAD1. If these cells are treated at 27°C with cis-aconitate, they are able to express some procyclic-specific mRNAs and proteins, including the procyclic surface protein EP procyclin, but they are unable to grow as procyclic forms.

Depletion of PUF3 in Lister 427 strain bloodstream-form trypanosomes had only a marginal effect on growth (S2 Fig A, B) and gene deletion was also possible (S2 Fig C, D). Induction of RNAi in pleomorphic EATRO1125 bloodstream-form trypanosomes (EATRO1125 strain) resulted in extremely mild growth retardation (Fig 2A), but once again, clones lacking both genes were readily obtained (Fig 2B, S3 Fig). Three independent clones lacking *PUF3* had the same growth rates as cells containing both, or

only one *PUF3* gene, all having doubling times between 6.9 h and 7.5 h. This suggested that the cells had been able to compensate for loss of PUF3 during clonal selection.

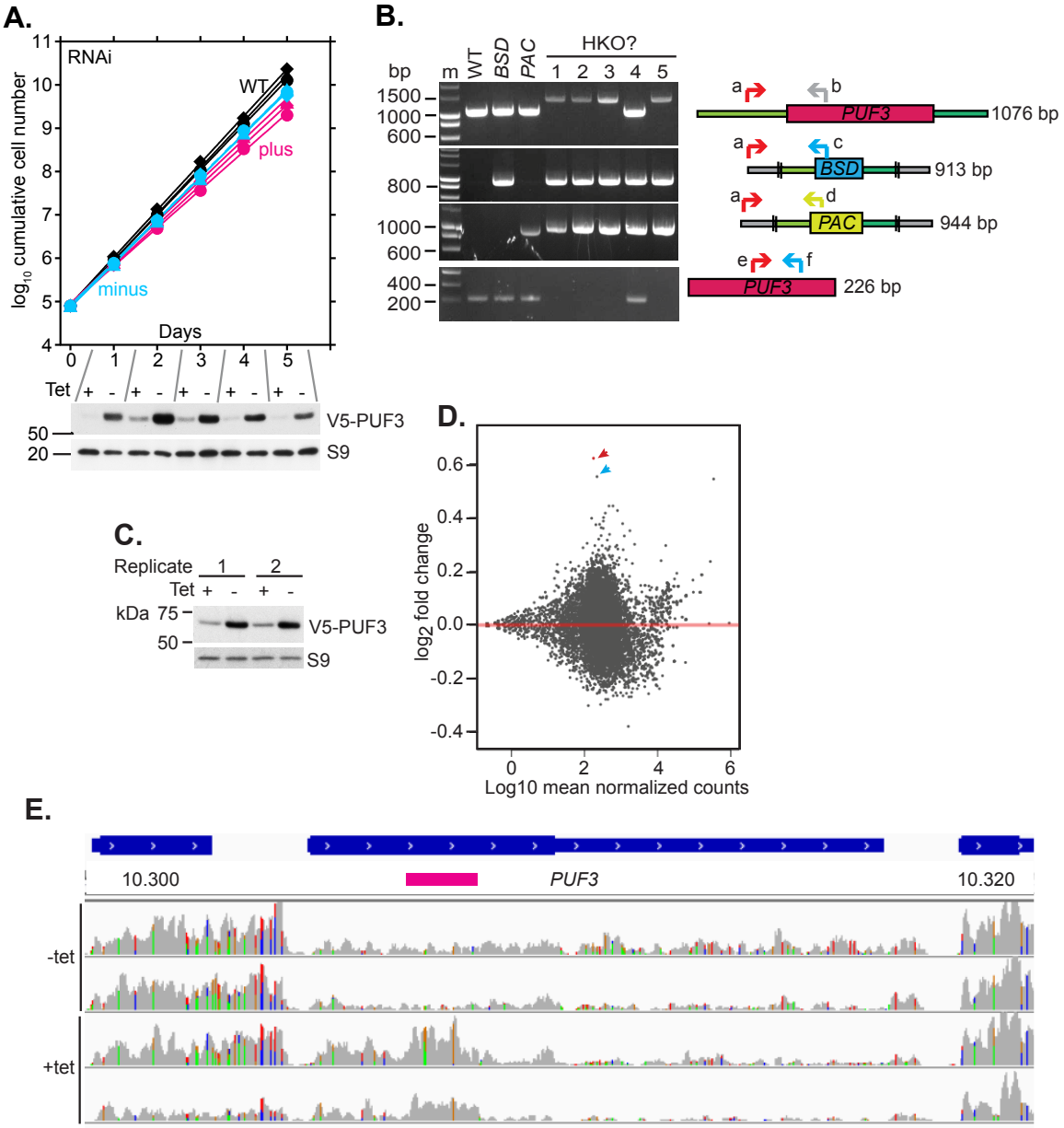

Figure 2

PUF3 is not essential for growth and survival of pleomorphic bloodstream forms.

A) Bloodstream-form EATRO1125 trypanosomes with inducible RNAi targeting PUF3 were grown with or without tetracycline and diluted as necessary. Three replicates are shown. "WT" is the parental line without RNAi. A Western blot showing depletion of V5-PUF3 from one of the replicates is shown below, with ribosomal protein S9 as loading control. The remaining replicates are shown in S3 Fig.

B) PCR test for *PUF3* knockout. Genomic DNA from wild-type, single knockouts (*BSD* and *PAC*) and candidate homozygous knockout (HKO) clones were used in a polymerase chain reaction. Double stroke lines represent the ends of the cloning fragment used for transfections. Oligonucleotides were: a) CZ6220; b) CZ3542; c) CZ5153; d) CZ5155; e) CZ6328; f) CZ6329 (S4 Table). PCR fragment sizes are on the right. Clone 4 did not have a homozygous deletion.

C) Western blot to detect PUF3 after 24h RNA interference induction, for the samples sent for RNASeq. Ribosomal protein S9 served as a loading control.

D) RNA reads that aligned to coding sequences were analysed by DEseq2 and fold changes (+/- tetracycline) represented on a scatter plot relative to the total number of normalized counts. The red arrow points to the only significantly increased mRNA; the cyan arrow is *PUF3*.

E) Distribution of reads mapping on PUF3 locus. Bam files generated from PUF3 RNAi reads were loaded in Integrative Genomics Viewer (IGV version 2.4.14) using default settings. The coverages were viewed using the same Data Range for all the bam files and screenshots taken. The cartoons in blue are the gene models of PUF3 and flanking genes while the magenta bar represents the PUF3 region targeted by RNAi.

To investigate why EATRO1125 cells with *PUF3* RNAi had slightly slower growth, we examined their trasncriptomes. RNA was made 1 day after tetracycline addition, in an attempt to ensure that compensation for PUF3 loss would not occur. Depletion of V5-PUF3 protein was confirmed (Fig 2C), but unexpectedly, the number of reads for PUF3 was about 1.5-fold increased (Fig 2D). Only a single additional mRNA (Tb927.9.10170) was significantly increased, but again only 1.5-fold (S1 Table); given the total number of genes considered this almost certainly reflects random variation. We concluded that depletion of PUF3 for 24h has no significant effect on the transcriptome.

RNAi targeting PUF3 causes an accumulation of RNA 5' to the targeted region

We were intrigued by the apparent contradiction between the effects of RNAi on PUF3 protein and mRNA, so we examined the distribution of sequence reads across the *PUF3* mRNA. Without RNAi, the sequence reads were fairly evenly distributed (Fig 2E), confirming results from other non-poly(A)-selected transcriptomes in the database. In contrast, after RNAi there was an accumulation not only over the region covered by the dsRNA, but also towards the 5'-end of the transcript. Comparison with the neighbouring gene (Tb927.10.300) shows that the read density in this 5' region was at least two-fold higher than in uninduced cells; it also clearly exceeded that seen over the region 3' to the dsRNA fragment. We have no explanation for this; perhaps, for some reason, *PUF3* mRNA remains stuck to the RISC complex, or its 5'-end is resistant to exoribonuclease digestion. The read density over the single other significantly increased mRNA (Tb927.9.10170, encoding anaphase-promoting complex subunit 11) was, in contrast, uniform in all replicates (not shown).

Interactions of TAP-tagged PUF3

To investigate mRNAs bound to PUF3, we made cells in which one *PUF3* open reading frame was fused at the 5'-end with a sequence encoding a tandem affinity purification (TAP) tag. Results described below suggest that this protein is at least partially functional. After UV cross-linking, cells were lysed, tagged PUF3 was selected, and bound and unbound RNAs were sequenced. Examination of the results suggested that the fractions in replicate 1 had been swapped at some point (S5 Fig) and for that reason, the results cannot be interpreted. Results that would be obtained on the assumption that these fractions were swapped are in S5 Fig. B-F.

The unbound RNA fraction had been subjected to rRNA depletion using RNase H and oligonucleotides complementary to rRNA [5]. To check the effect of this procedure on the transcriptome we sequenced the input RNA from the experiment, either with or without rRNA depletion. There was no significant difference between the two datasets, which validated our method for rRNA depletion (Fig 3, S2 Table, sheets 3 and 4; S4 Fig A, S5 Fig).

We also subjected TAP-PUF3 to tandem affinity purification and identified co-purifying proteins by mass spectrometry (S3 Table). The only proteins that appeared to co-purify, relative to TAP-GFP controls, were abundant proteins that had previously been identified in other purifications. The results therefore gave no hint of a mechanism by which PUF3 might inhibit mRNA translation or promote mRNA degradation.

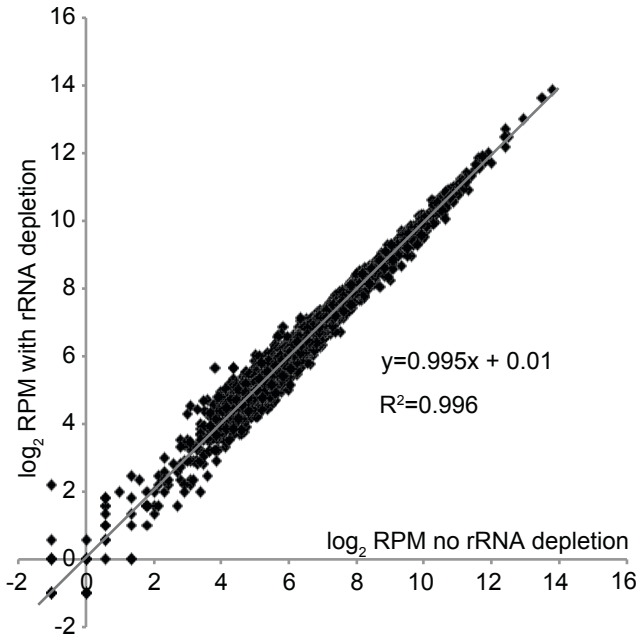

Figure 3

rRNA depletion using oligonucleotides and RNaseH does not affect the transcriptome.

Average reads per million reads for duplicate matched total RNA samples are shown, on a log2 scale. Individual comparisons are in S5 Fig.

PUF3 may assist bloodstream-to-procyclic form differentiation

We next tested whether PUF3 has a role in differentiation of bloodstream forms to procyclic forms. We first tested the effects of short-term depletion by RNAi (S3 Fig). In Lister 427 trypanosomes, depletion of PUF3 caused a delay in appearance of EP

procyclin (S6 Fig A). The defect was accentuated in Lister 427 trypanosomes lacking PUF3 (Fig 4A). To check that the defect was due to to the lack of PUF3, we inducibly re-expressed the (untagged) protein. In one cloned line, re-expression of PUF3 was tetracycline-dependent and partially restored EP procyclin expression (Fig 4B, C). A second line, with constitutive PUF3 expression, showed complete complementation (Fig 4B, C). Interestingly, C-terminally myc-tagged PUF3 could not complement the defect (Fig 4D). N-terminally myc-tagged PUF3 was extremely poorly expressed (S6 Fig B).

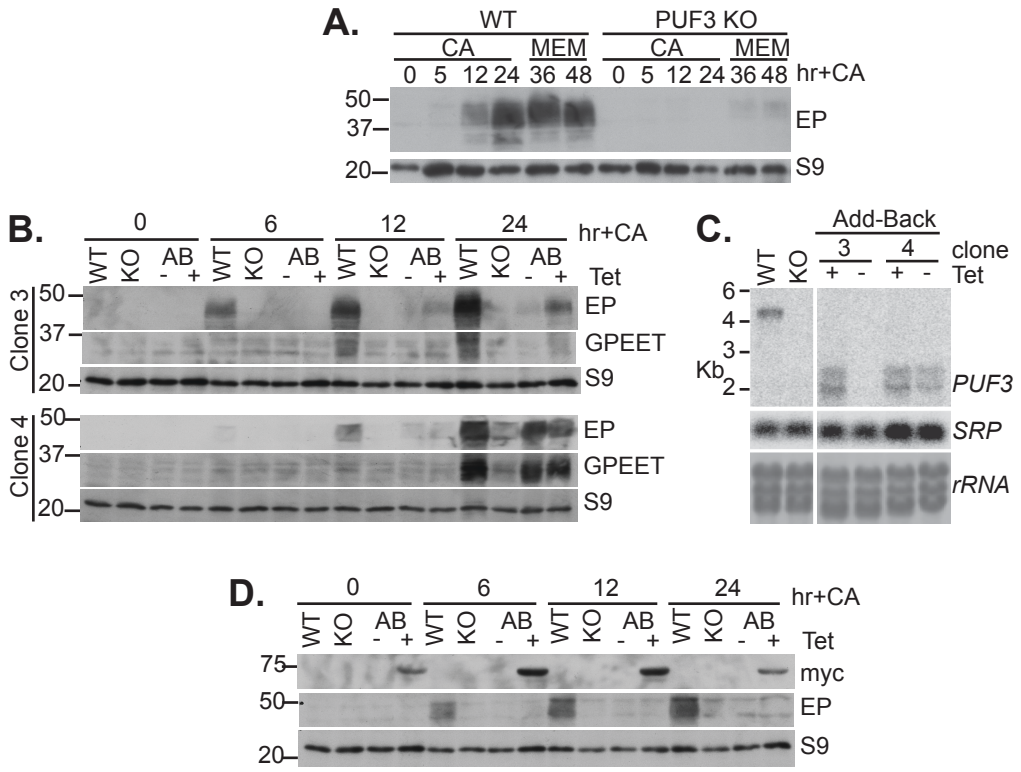

Figure 4

PUF3 is required for procyclin expression by differentiation-stimulated Lister 427 bloodstream forms.

A) Monomorphic bloodstream forms were grown to 1.8-2.0 $\times 10^6$ cells/ml, then cis-aconitate (CA) was added to the culture (final concentration 6 mM) and the temperature was reduced from 37°C to 27°C. After 24h the medium was changed to standard procyclic medium (very-low-glucose, high-proline minimal essential medium, MEM). The time in hours after CA treatment is indicated. EP procyclin, and ribosomal protein S9, were detected by Western blotting. WT: cells with wild-type complement of *PUF3* genes; KO: cells lacking *PUF3*.

B) As in (A), but cells with inducible expression of PUF3 ("add-back", AB) were also included, either with (+) or without (-) tetracycline addition (final concentration 100 ng/ml) 24h before the start of the experiment. The add-back results are for two different clones, clone 3 and clone 4. GPEET in addition to EP and S9 were detected by Western blotting.

C) Northern blot showing inducible expression of *PUF3* mRNA in the add-back clones 3 and 4. The 7SL signal recognition particle RNA (SRP) and rRNA were detected as loading controls.

D) Expression of PUF3-myc did not complement the defect in differentiation. The experiment is similar to that in (B) except that PUF3-myc was expressed.

Next, we followed differentiation in pleomorphic EATRO1125 cells with PUF3 depletion or knock-out (Fig 5, S6 Fig). These were first grown to high density in methyl cellulose-containing media to promote stumpy-form differentiation, and then switched to differentiation conditions (Fig 5A). Again, expression of EP procyclin was delayed after *PUF3* RNAi. In these cells, we could also see that *PUF3* RNAi caused reduced expression of the stumpy-form marker PAD1 (Fig 5B, S6 Fig C). However, pleomorphic EATRO1125 cells that completely lacked PUF3 expressed EP procyclin normally (Fig 5C). Preliminary results also confirmed PAD1 expression (S6 Fig D). These results all suggest that loss of PUF3 had been compensated for during selection of the knock-out clones.

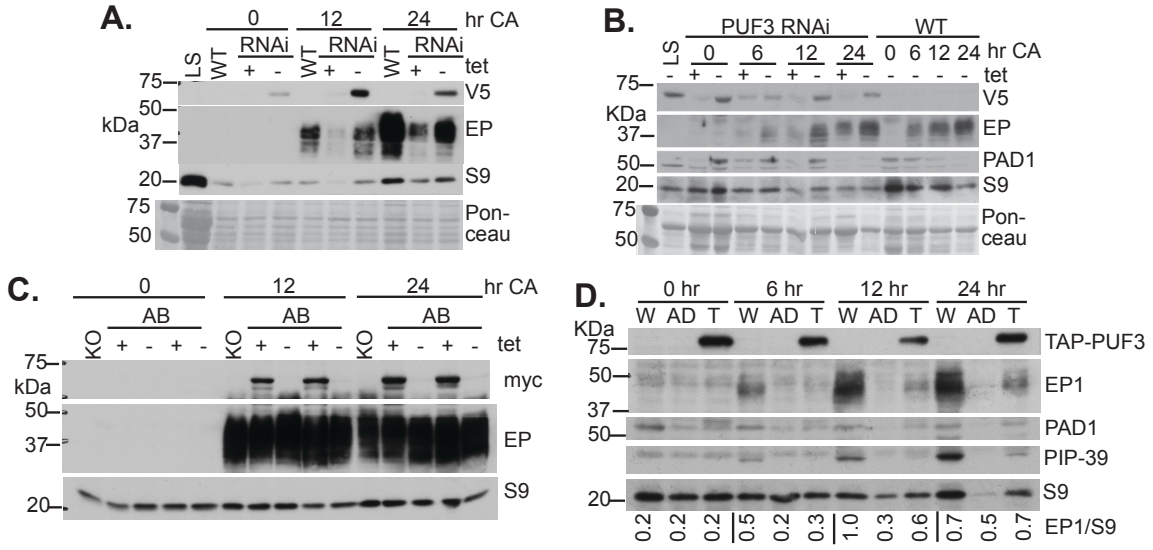

Figure 5

Differentiation of PUF3-depleted EATRO1125 bloodstream forms.

A) Expression of EP procyclin during differentiation of cells with *PUF3* RNAi. RNAi was induced in EATRO1125 bloodstream forms expressing V5-PUF3, at a density of 5 x 10^5, in methyl cellulose-containing medium. 48 hours later, cis-aconitate (CA) was added and the temperature changed to 27°C. LS: long slender cells, WT; wild-type. Ribosomal protein S9 and the Ponceau protein stain serve as controls.

B) Replica of (A) with detection of PAD1.

C) Cells adapted to the absence of PUF3 are not defective in EP expression in response to differentiation stimuli. The experiment is as in (A), but here cells completely lacking PUF3, with (+) or without (-) inducible expression of PUF3-myc, were used.

D) The experiment is as in (A) but here PUF3 single knockout cells expressing *Drosophila* adenosine deaminase-PUF3 (AD) or TAP-PUF3 from PUF3 locus were used.

Given the potential function of PUF3 in differentiation, and the potential impact of tagging (Fig 4D), we checked the differentiation capability of our trypanosomes expressing only TAP-tagged PUF3. We found that these cells showed delayed expression of both EP procyclin and the phosphatase PIP39 (Fig 5D) relative to wild-type. This could be haploid insufficiency but might also reflect an effect of the tag. Cells in which all PUF3 was N-terminally fused to *Drosophila* adenosine deaminase (to be described elsewhere) were much more defective.

Correct expression of PUF3 in procyclic forms is required for optimal growth

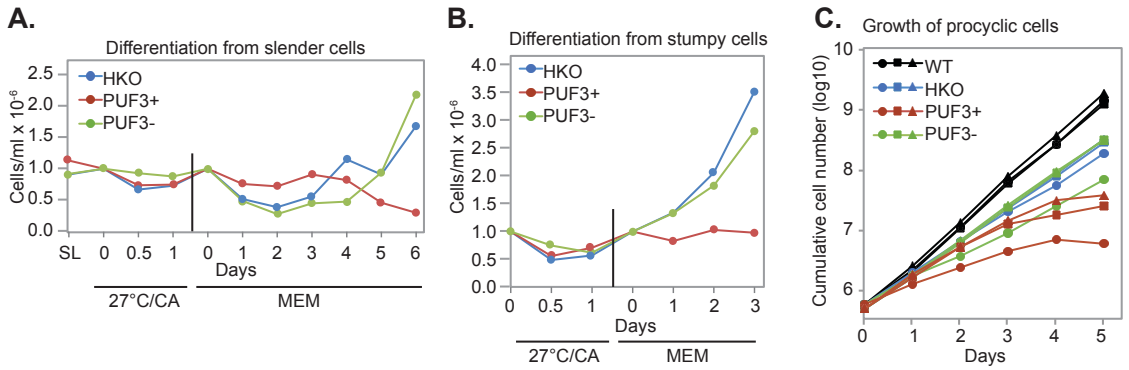

Figure 6

Correct expression of PUF3 is required for normal growth of procyclic forms

Three different EATRO1125 cell lines were used: wild-type (WT); lacking endogenous PUF3 genes (HKO), and HKO cells with inducible expression of PUF3, grown either with (PUF3+) or without (PUF3-) tetracycline.

A) Cells were grown to 10^6/ml, then cis-aconitate was added, with or without tetracycline, and the temperature was reduced to 27°C (27°C/CA). After 24h the cells were transferred to procyclic-form medium (MEM).

B) As (A) except that cells were kept at high density to induce differentiation to stumpy forms before differentiation to procyclic forms was induced.

C) Differentiated cells were grown as procyclic forms for two weeks, then tetracycline was added to induce PUF3 over-expression. Results are for three independent cultures, indicated by the different symbols.

Differentiated bloodstream cells were grown for at least four weeks as procyclics before tetracycline addition to induce PUF3 expression.

Pleomorphic bloodstream-form trypanosomes lacking PUF3 could differentiate, and grew as procyclic forms when we transferred them to appropriate medium (Fig 6A, B). Cells with inducible PUF3, but without tetracycline, grew similarly. However, if expression of PUF3 was induced during and after differentiation, growth was repressed (PUF3+ in Fig

6A, B). To examine the phenomenon in more detail, we studied the effect of expression of untagged PUF3 in well-established procyclic forms. Despite an adaptation time of several weeks, both cells lacking PUF3 and the uninduced line grew slower than wild-type (Fig 6C). However, induction of PUF3 transgene expression did not correct the defect: instead, growth was repressed (Fig 6C). This suggests that the correct level of PUF3 is required for optimal growth of procyclic forms.

Finally we investigated whether PUF3 is required for differentiation of procyclic forms to bloodstream forms. RBP10 expression in procyclic forms forces them to "jump" to metacyclic-like forms, a very small proportion of which can replicate when placed in bloodstream-form culture conditions. We therefore transfected bloodstream-form EATRO1125 cells lacking *PUF3* genes with a construct that allows inducible expression of RBP10-myc. Regulated expression was tested, then the cells were differentiated to procyclic forms. This selected against lines with insufficient suppression of RBP10 expression, and indeed only one clone survived. The cells were cultured for 2 weeks, then RBP10-myc expression was induced with tetracycline. Two days later the cells were placed in bloodstream-form culture conditions. After 2 weeks, replicating *PUF3* HKO bloodstream-form trypanosomes were readily detectable, indicting that procyclic forms lacking PUF3 are capable of RBP10-induced differentiation to the bloodstream form.

Discussion

Our results suggest that PUF3 is not essential for survival of cultured bloodstream-form or procyclic-form trypanosomes. Although loss of PUF3 initially resulted in defects in bloodstream-form growth and differentiation, these were lost after the more prolonged culture that was required to select cells completely lacking *PUF3* genes. We suggest, therefore, that the functions of PUF3 can be taken over by overproduction, or perhaps altered modification, of another protein - most likely, another member of the PUF family. This would not be unusual - the functions of Puf proteins also overlap in *Saccharomyces cerevisiae* [6]. In *T. brucei*, the most closely related proteins to PUF3 are PUF1 and PUF4 (S1 Fig, Fig 7). RNAi targeting pairs of these three mRNAs, however, had no effect on bloodstream-form trypanosome growth - although reduction in the proteins was not assessed [7]. PUF1 also suppresses expression when tethered, and was non-essential in procyclic forms, but pull-down and microarray results suggested that it binds to retroposon sequences [7]. It is probably much more abundant than PUF3 [8]. PUF5, like PUF3, is non-essential in procyclic forms [9], but PUF5 stimulated when tethered [4]. Tethering of PUF4 fragments yielded no evidence that it has suppressive function [4], but several other PUFs functioned only at full length [3]. The only other cytoplasmic PUF protein that is known to have suppressive function is PUF2, which is essential in bloodstream forms [10]; its binding specificity is unknown. PUF2 probably has an abundance similar to that of PUF3 [8] so a period of adaptation might be required if it were to compensate for the absence of PUF3.

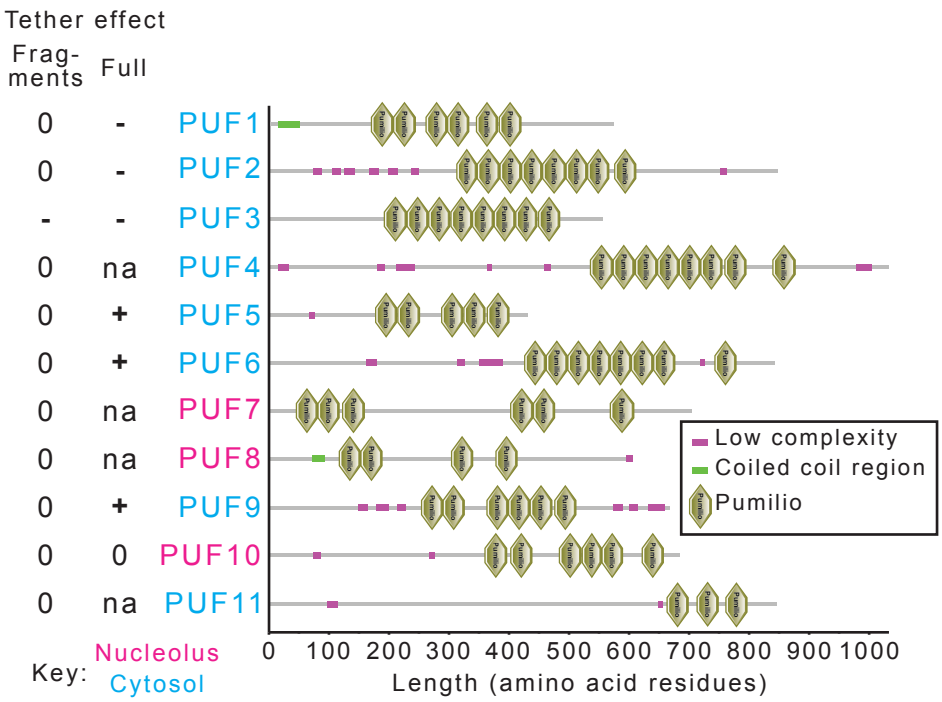

Figure 7

Domains of *T. brucei* PUF proteins.

Key to symbols and colours are on the Figure. The effects of tethering to a reporter mRNA are shown on the left. "Fragments" refers to results from a high-throughput screen using random genomic fragments [4], and "full" refers to a screen using full-length open reading frames [3]; "na" indicates that the open reading frame was not sufficiently represented. The colours of the labels refer to the location, usually for GFP fusion proteins [2].

Ectopic expression of PUF3 had no effect in bloodstream forms but was deleterious in procyclic forms. Since we have no antibody to PUF3, we cannot judge wether the absence of an effect in bloodstream forms is due to differences in PUF3 action, or to having much less expression. The effect in procyclic forms is difficult to interpret. Binding of proteins to mRNAs depends on both the binding affinities and concentrations, so over-expression of PUF3 may cause it to bind to mRNAs that have sub-optimal recognition sequences. This might then displace other PUF proteins from their target mRNAs and thus disrupt regulation. Alternatively, PUF3 may be needed in order to fine-tune expression of some mRNAs in procyclic forms, in which case excess PUF3 could cause excessive repression of genuine targets.

PUF3 can suppress expression of an mRNA when tethered, and the effect is quite rapid. Results from the previous tethering screen suggested that the region of PUF3 that is responsible for this is towards the N-terminus, with the strongest and most reproducible suppression being from fragments that started just after the initiation codon. Two short N-

terminal sequences are conserved throughout Kinetoplastea, and unique to PUF3: they are WTV[QH][QE]D[DE][TYC] at residues 3-10, and YT[KQ][SC]SQVF at residues 38-45. It is possible that one or both sequences are implicated in suppression. The mechanism, however, remains unknown since no protein partners were identified. Our affinity purifications for RNASeq and mass spectrometry were done using N-terminally TAP-tagged protein from bloodstream forms. This tagged version only partially complemented the differentiation defect in cells lacking wild-type PUF3, so it is possible that the wild-type N-terminus is required for protein-protein interactions. Disruption of RNA binding seems less likely since the puf domains are separated from the N-terminus by 200 residues - but interference in the folded protein is still possible.

We found no evidence that PUF3 depletion affected the transcriptome, although the cells grew slightly slower than normal; perhaps only translation is affected, or the effects are so minor as to escape detection. There is no evidence, from transcriptomes or proteomes, that its level is developmentally regulated. However, since PUF3 depeletion more strongly affected dfferentiation, and growth of procyclic forms, it is also possible that PUF3 is important only in these stages.

PUF3 is conserved throughout Kinetoplastea, including the free-living Bodo saltans, and in parasites restricted to arthropods. It therefore clearly does confer a selective advatage in the natural environment.

Materials and Methods

Trypanosome culture and transfection

All the cells used in this study were based on stable cell lines constitutively expressing a tetracycline repressor [11], which for simplicity are referred to as wild-type. Lister 427 bloodstream forms (monomorphic) were cultured at 37°C and 5% CO_2 in adapted HMI9 [12]: MEM including 3.24 g/L sodium bicarbonate, 23.2 mg/L bathocuprone sulphonate, 0.14 g/L hypoxanthine, 0.1 g/L sodium pyrivate, 39 mg/L thymidine, supplemented to give 1.5 mM L-cysteine and 0.2 mM β-mercaptoethanol, and with 10% fetal bovine serum (Heat inactivated at 55°C for 1 hr). EATRO 1125 bloodstream forms (pleomorphic) were cultured in HMI-9 media containing 1.1 % methyl cellulose (Sigma, M0512) to maintain their pleomorphism [13]. Densities were kept below 5x 10^6/ml for Lister 427 and below 1x10^6ml for EATRO1125.

Procyclic forms were cultured at 27°C in MEM-Pros medium [14] supplemented with 5% PenStrep (Gibco), 7.5 mg/l hemin and 10% of heat inactivated fetal bovine serum at a density between 10^5 - 4×10^6 cells/ml.

Differentiation of bloodstream slender forms to stumpy forms was done by allowing a starting density of 5-7$\times10^5$ cells/ml to peak at 3-4$\times10^6$ cells/ml, and later descend to 10^6 cells/ml. This takes 36-48 hrs. Differentiation to procyclic forms was done by adding

6mM cis-aconitate to slender or stumpy forms at 10^6 cells/ml and incubating at 27°C for 24 hrs before transferring the cells to MEM-Pros. Growth curves were performed using media without selecting drugs, with 100ng/μl of tetracycline when needed for transgene induction. Plasmids and oligonucleotides are listed in S4 Table and illustrated in S7 Fig. Transfection of linearised plasmids and drug selections were done as described in [11].

Immunofluorescence microscopy

Cells were induced for RNAi for 24 hrs and processed for immunofluorescence microscopy. Anti-V5 antibody (1:500, mouse) and secondary antibody with Cy3 fluorophore (1:700, anti-mouse) were used. DNA was stained with 4′,6-diamidino-2-phenylindole (DAPI). Cells were fixed in 4% paraformaldehyde at room temperature for 20 min, washed and smeared on slides. The slides were either stored at 4°C or analysed immediately. All the following steps were done at room temperature. The slides were placed in PBS for 5 min to rehydrate the cells then permeabilised using 0.2% Triton-X for 20 min. Cells were blocked with 0.5% gelatin for one hour and incubated with a primary antibody (dissolved in 0.5% gelatin) for one hour. The slides were washed three times with PBS and incubated in the dark with a secondary antibody containing a fluorescence dye for one hour. The slides were washed three times and incubated for 15 min with 1µg/ml DAPI solution. Three more washes were made and the slides air-dried before mounting with 90% glycerol-PBS. The cover slips were sealed to the slides using nail polish. Images were acquired using an Olympus IX81 microscope and analysed with Olympus xcellence software or image J.

RNA and DNA preparation and blotting

To obtain at least 10μg of RNA, 3×10^7 of log phase growing bloodstream-form cells (1 x 10^6/ml for Lister 427, no more than 8x 10^5/ml for EATRO1125) collected at 900×g for 10 min and RNA extracted using TriFast reagent according to the manufacturer's protocol. RNA concentrations were determined using a Nanodrop spectrophotometer and the RNA stored at -80°C or used immediately. rRNAs were depleted from whole RNA extracts before RNAseq using a cocktail of DNA oligonucleotides targeting rRNAs with RNaseH [5]. For blotting, RNA was electrophoresed on a formaldehyde agarose gel, blotted on a nylon membrane by downward capillary for at least 4 hrs followed by UV cross-linking of RNA to the membrane using Stratagene® UV cross-linker. Prior to probing, the blot was blocked using pre-hybridization buffer (5×SSC, 0.5% SDS, 5× Denhardt's solution, 100μg/ml Salmon sperm DNA). For spliced leader probing, the pre-hybridization buffer was similar to the previous one except for 6×SSC and 0.05% sodium pyrophosphate. Probes made from PCR products were prepared using PrimeIt kit (stratagene), and those from oligonucleotides using phosphonucleotide kinase (NEB) to incorporate radioactive nucleotides into the probe. The northern blot was incubated overnight with at least 2×10^6 cpm/ml of the probe at 42°C or 65°C for an oligonucleotide or a PCR-product probe respectively. The blot was washed with buffer (2 x SSC, 0.1 %

SDS) twice for 10 min per wash at room temperature and once for 15 min at 65°C then exposed to a phosphorimaging film before developing the autoradiographs with a phosphorimager (Fujifilm FLA-3000).

For Southern blotting, At least 5μg pf genomic DNA was digested using appropriate restriction enzymes and electrophoresed on an ethidium bromide-containing agarose gel. The gel was washed in hydrolysis buffer (0.25 M HCl) for 15 min, in denaturing buffer (1.5 M NaCl, 0.4M NaOH) for 15 min to denature long DNA and in neutralizing buffer (1.5 M NaCl, 0.5 M Tris-Cl pH7.5) for 15 min before blotting onto a nylon membrane and processing as described for the northern blotting.

Chloramphenicol acetyltransferase assay

2×10^7 cells were collected at 900×g for 10 min and washed three times with PBS. The pellet was re-suspended in 100µl of CAT buffer (100mM Tris-HCL pH 7.8) and freeze-thawed three times using liquid nitrogen and a 37°C heating block. The supernatants were collected by centrifugation at 10000×g for 5 min and kept in ice. The protein concentrations were determined by Bradford assay (BioRad) according to the manufacturer's protocol. For each setup, 0.5µg of protein in 50µl of CAT buffer, 10µl of radioactive butyryl CoA (^{14}C), 2µl of chloramphenicol (stock: 40mg/ml), 200µl of CAT buffer and 4ml of scintillation cocktail were mixed in a Wheaton scintillation tube HDPE (neoLab #9-0149) and the incorporation of radioactive acetyl group on chloramphenicol was measured using program 7 of Beckman LS 6000IC scintillation counter.

Tandem affinity purification and mass spectrometry

$6\text{-}7\times10^9$ cells were collected at 900×g at 4°C for 15 min in 500 ml centrifuge containers. The cell pellet was washed twice in ice cold PBS, snap frozen in liquid nitrogen and stored at -80°C or used immediately. The cells were lysed in IPP10 buffer (10mM Tris-HCL (pH 7.8), 10mM NaCl, 0.1% IGEPAL, 50µg/ml leupeptin) by passing them through a 21-gauge needle 15 – 20 times. The lysate was cleared by centrifugation at 10,000×g for 10 min and the concentration of NaCl adjusted to 150mM before loading onto a 0.8×4 cm Poly-Prep column (BioRad) containing IgG sepharose beads that had been prewashed in IPP150 buffer (10mM Tris-HCL (pH 7.8), 150mM NaCl, 0.1% IGEPAL). This was incubated at 4°C for 2 hours with gentle rocking. The beads were washed twice with 10ml IPP150 buffer and once with 10ml TEV cleavage buffer (IPP150 buffer, 1mM DTT, 0.5mM EDTA). The beads were re-suspended in 800µl TEV cleavage buffer and incubated with TEV protease at 15-20°C for 2 hours with gentle rocking. The eluate was collected and transferred to a second column containing calmodulin beads in calmodulin binding buffer (IPP150 buffer, 10 mM β- mercaptoethanol, 1 mM $MgAc_2$, 1 mM imidazole, 2 mM $CaCl_2$), and incubated for 2 hours at 4°C. The beads were washed three times with calmodulin binding buffer and the proteins were eluted using 2mM EGTA. The proteins in the eluate were concentrated using trichloroacetic acid and dissolved in

Laemmli buffer. This was electrophoresed for a short time so that the sample dye ran approximately 1.5cm. The gel was then submitted to the sequencing facility at ZMBH (Heidelberg) for mass spectrometry.

RNA pull-down and RNASeq

This procedure was done in a similar manner as the TAP purification with the exception that the cells were UV-irradiated before lysis and only one purification step was done. Approximately 3×10^9 cells were collected at 900 ×g at 4°C for 15 min, re-suspended on a petri dish in 30 ml of the supernatant on ice and UV- irradiated twice using a Stratagene UV crosslinker with an energy of 3 kilo joules. The cells were pelleted again in 4°C at 2000×g for 5 min and washed twice with ice cold PBS. Cells were lysed in IPP10 buffer consisting of KCl in place of NaCl and containing RNase inhibitor (RNasin, Invitrogen), by passing them through a 21-gauge needle 15-20 times. The lysate was cleared by centrifugation at 10,000×g for 10 min, KCl concentration adjusted to 150 mM and loaded onto a column containing prewashed sepharose beads similar as in the TAP purification. A small volume of the lysate equivalent to 7.5×10^7 cells was collected as the input fraction and treated with 0.2mg/ml proteinase K, 8mM EDTA, 0.2% SDS at 45°C for 15 min before adding TriFast-FL reagent (Peqlab). The columns were incubated at 4°C for 2 hours with gentle rocking and the unbound lysate collected. This was also subjected to proteinase K treatment as described for the input fraction. Subsequently, the beads on the column were washed twice with IPP150 buffer and once with TEV cleavage buffer, re-suspended in 500μl of TEV cleavage buffer containing TEV protease and incubated at 15-20°C for 2 hours with gentle rocking. The cleaved product (bound) was treated with proteinase K and mixed with TriFast-FL reagent. The three fractions in TriFast-FL were stored at -80°C and later used for RNA extraction according to the manufacturer's protocol. RNA concentrations were determined using a NanoDrop® or a Qubit® spectrophotometer and submitted to the Next-generation sequencing facility in Bioquant (Heidelberg) for library preparation. Sequencing was done at GeneCore EMBL (Heidelberg).

Fastq files were used as input in TrypRNAseq – an in-house pipeline for RNAseq (https://github.com/klprint/TrypRNAseq; 10.5281/zenodo.158920). Briefly, reads were taken through quality control, mapped on the *T. brucei* genome (TREU927) using Bowtie2, and the number of reads aligning to coding sequences (CDS) determined. Each sequence was allowed to align 20 times to allow for the repetitive nature of the *T. brucei* genome, but for statistical analyses, only one copy of repeated genes was considered [15]. The fold changes were compared between RNAi induced and un-induced samples using the DESeqU1 adpatation (10.5281/zenodo.165132) of DESeq2 [16].

Data availability

The sequence results are available under the following accession numbers: RNAi - E-MTAB-8062; RIP-Seq - E-MTAB-8068.

Acknowledgements

Kevin Marucha was supported by a fellowship from the Deutsches Akademisches Austauschdienst (DAAD). The work was also supported by the Deutsche Forschungsgemeinshaft, Grant CL112/24. We thank Ute Leibfried and Claudia Helbig for technical assistance, and Keith Matthews for the antibodies to PAD1 and PIP39. We thank the ZMBH mass spectrometry facility for protein analysis and David Ibberson (BioQuant, Heidelberg) for RNASeq library construction.

Literature cited

1. Lu G, Dolgner S, Tanaka Hall T. Understanding and engineering RNA sequence specificity of PUF proteins. Curr Opinion Struct Biol. 2009;19:110-5.
2. Dean S, Sunter J, Wheeler R. TrypTag.org: A trypanosome genome-wide protein localisation resource. Trends Parasitol. 2016;33:80-2.
3. Lueong S, Merce C, Fischer B, Hoheisel J, Erben E. Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. Molecular microbiology. 2016;100:457-71. PubMed Central PMCID: PMCdoi: 10.1111/mmi.13328 PMID: 26784394.
4. Erben E, Fadda A, Lueong S, Hoheisel J, Clayton C. Genome-wide discovery of post-transcriptional regulators in *Trypanosoma brucei*. PLoS pathogens. 2014;10:e1004178. doi: doi:10.1371/journal.ppat.1004178. PubMed Central PMCID: PMCdoi:10.1371/journal.ppat.1004178.
5. Minia I, Merce C, Terrao M, Clayton C. Translation regulation and RNA granule formation after heat shock of procyclic form *Trypanosoma brucei*: many heat-induced mRNAs are increased during differentiation to mammalian-infective forms. PLoS neglected tropical diseases. 2016;10:e0004982. PubMed Central PMCID: PMC10.1371/journal.pntd.0004982.
6. Fischer AD, Olivas WM. Multiple Puf proteins regulate the stability of ribosome biogenesis transcripts. RNA biology. 2018;15(9):1228-43. Epub 2018/09/27. doi: 10.1080/15476286.2018.1521211. PubMed PMID: 30251908; PubMed Central PMCID: PMCPMC6284577.
7. Luu VD, Brems S, Hoheisel J, Burchmore R, Guilbride D, Clayton C. Functional analysis of *Trypanosoma brucei* PUF1. Mol Biochem Parasitol. 2006;150:340-9. doi: DOI: 10.1016/j.molbiopara.2006.09.007. PubMed Central PMCID: PMCPMID: 17052765
8. Dejung M, Subota I, Bucerius F, Dindar G, Freiwald A, Engstler M, et al. Quantitative proteomics uncovers novel factors involved in developmental dfferentiation of *Trypanosoma brucei*. PLoS pathogens. 2016;12(2):e1005439. Epub 2016/02/26. doi: 10.1371/journal.ppat.1005439. PubMed PMID: 26910529; PubMed Central PMCID: PMCPmc4765897.

9. Jha B, Archer S, Clayton C. The trypanosome pumilio domain protein PUF5. PloS ONE. 2013;8:e77371. doi: 10.1371/journal.pone.0077371. PubMed Central PMCID: PMCPMID: 24167571.
10. Jha B, Fadda A, Merce C, Mugo E, Droll D, Clayton C. Depletion of the trypanosome pumilio domain protein PUF2 causes transcriptome changes related to coding region length. Eukaryot Cell. 2014;13:664-74. doi: doi: 10.1128/EC.00018-14. PubMed Central PMCID: PMCPMID: 24681684.
11. Alibu VP, Storm L, Haile S, Clayton C, Horn D. A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. Mol Biochem Parasitol. 2004;139:75-82.
12. Duszenko M, Mühlstädt K, Broder A. Cysteine is an essential growth factor for *Trypanosoma brucei* bloodstream forms. Mol Biochem Parasitol. 1992;50:269-74.
13. Vassella E, Kramer R, Turner CM, Wankell M, Modes C, van den Bogaard M, et al. Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of Trypanosoma brucei. Molecular microbiology. 2001;41(1):33-46. Epub 2001/07/17. PubMed PMID: 11454198.
14. Ziegelbauer K, Quinten M, Schwarz H, Pearson TW, Overath P. Synchronous differentiation of *Trypanosoma brucei* bloodstream to procyclic forms in vitro. Eur J Biochem. 1990;192(2):373-8.
15. Siegel T, Hekstra D, Wang X, Dewell S, Cross G. Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic acids research. 2010;38:4946-57.
16. Love M, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome biology. 2014;15:550. PubMed Central PMCID: PMCPMID: 25516281.
17. Mugo E, Clayton C. Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in *Trypanosoma brucei*. PLoS pathogens. 2017;13:e1006560. PubMed Central PMCID: PMCPMID: 28800584 PMCID: PMC5568443 DOI: 10.1371/journal.ppat.1006560.
18. Klein C, Terrao M, Clayton C. The role of the zinc finger protein ZC3H32 in bloodstream-form *Trypanosoma brucei*. PLoS ONE. 2017;12:e0177901. PubMed Central PMCID: PMC10.1371/journal.pone.0177901.
19. Liu B, Marucha K, Clayton C. The zinc finger proteins ZC3H20 and ZC3H21 stabilise mRNAs encoding membrane proteins and mitochondrial proteins in insect-form *Trypanosoma brucei* BIORXIV. 2019. doi: BIORXIV/2019/646547.

Supplementary Tables

S1 Table

RNASeq data: effect of RNAi targeting PUF3

S2 Table

RIPSeq data: mRNA association with TAP-PUF3 in bloodstream forms

S3 Table

Mass spectrometry data: lack of protein association with TAP-PUF3with other proteins

S4 Table

Plasmids and oligonucleotides

Supplementary Figures

S1 Figure

(A) Cladogram of PUF orthologs

(B) PUF3 protein sequences from *Trypanosoma brucei*, *T. vivax*, *T. congo*, *T. theileri, T. cruzi*, *Leishmania major*, *Blechomonas ayalai, Endotrypanum monterogeii* and *Bodo saltans* were aligned using Clustal Omega (EMBL-EBI). The shading highlights similarities. Red bars indicate the position of the PUF repeats in *T. brucei* PUF3. The apparent gap in the seventh PUF repeat for *T. vivax* is caused by a gap in the sequence assembly.

S2 Figure

PUF3 is not essential for survival of monomorphic bloodstream forms.

(**A**) Growth of bloodstream-form Lister427 trypanosomes with *in situ* N-terminally V5-tagged PUF3 and tetracycline-inducible *PUF3* RNAi. Results are shown for wild-type cells and for two independent clones grown with (+) and without (-) tetracycline.

(**B**) Western blots showing the expression of a V5-tagged PUF3 in the cell lines in (A). Antibody to aldolase was used as a loading control.

(**C**) The upper panel represents the design of a Southern blot. Coding sequences of genes are represented by rectangles and vertical double-strokes show restriction sites. Genomic DNA of cells selected for homozygous knockout (transfected with plasmid pHD2851 and pHD2852) was digested with *Eco* RI (experiment a) or *Eco* NI and *Sac* II (experiment b). The digested DNAs for wild-type (WT), single knockouts (*BSD* or *PAC*) and three candidate homozygous knockout clones were analysed by Southern blotting using a probe targeting *PUF3* (as indicated in the upper panel), with *TUT3* as a loading control. L: marker ladder.

(D) DNA from (C) was used in a polymerase chain reaction using the primers shown (Table S4). Double stroke lines represent the ends of the cloned fragments used for transfections.

S3 Figure

PUF3 is not essential for survival of pleomorphic EATRO1125 bloodstream forms.

(**A**) Western blots showing V5-PUF depletion for two additional samples from Figure 2A. Controls are ribosomal protein S9 or trypanothione reductase (TR).

(**B**) The upper panel represents the design of a Southern blot. Coding sequences of genes are represented by rectangles and vertical double-strokes show restriction sites. Genomic DNA of cells selected for homozygous knockout (transfected with plasmid pHD2851 and pHD2852) was

digested with *Eco* NI and *Sac* II (experiment a) or *Eco* RI (experiment b). The digested DNAs for wild-type (WT), single knockouts (*BSD*) and four candidate homozygous knockout clones were analysed by Southern blotting first using a probe targeting *PUF3* (as indicated in the upper panel) and second with a probe for beta-tubulin.

(**C**) Cumulative growth of the bloodstream forms, mean values. Variation (not shown) was so low that all of the values overlapped. AB: Add-back of tetracycline-inducible PUF3-myc, +tet or -tet.

S4 Figure

mRNA association with TAP-PUF3.

TAP-PUF3 was bound to IgG beads and bound and unbound RNAs were sequenced.

A) Principal component analysis for RNA-pull-down results. EL - eluate or bound fraction; FT - flow-through or unbound fraction. Results for the input RNA (total or rRNA depleted) are also included.

Eluate 1 clearly clustered with the input RNA and unbound fraction 2. We expect the unbound fraction to be very similar to total RNA, and also that both unbound fractions should be similar (S2 Table Sheet 2).

B) The ratios of bound/unbound reads per million were calculated for all unique genes. assuming that the samples from experiment 1 had been swapped. We expect the mRNA encoding the TAP-tagged protein to be enriched via the nascent polypeptide, as seen previously for the *RBP10* [17], ZC3H32 [18], ZC3H20 and ZC3H21 [19] mRNAs. With swapped samples, *PUF3* mRNA was enriched 4.6-fold in replicate 1 and 1.88-fold in replicate 2 (orange dot). There was also then a weak positive correlation between the two datasets (R= 0.35 for log-transformed results).

C) Relationship between binding and mRNA length for experiment 2. Each dot represents a single open reading frame and all values are $\log_2$-transformed. Average of bound/unbound reads per million on the y-axis; annotated mRNA length on the x-axis. mRNAs with no annotated 5'-UTR or 3'-UTR are excluded.

D) As (B), but for experiment 1, assuming that the flow-through and eluate fraction were swapped.

E) Half-lives of bound and unbound mRNAs, assuming that the flow-through and eluate fraction for experiment 1 were swapped. mRNAs were classed as "bound" if both experiments gave a ratio >2, and unbound if both ratios were less than 1. IF this analysis is correct, PUF3-associated mRNAs have shorter half-lives than unbound mRNAs, which would fit with a suppressive action of PUF3.

F) Average binding ratios for each mRNA are plotted agains the fold change after RNAi. None of the RNAi effects was statistically significant. Most mRNAs with higher bound/unbound ratios showed no RNAi effect at all. Orange dots represent mRNAs for which bound/unbound was >2 in both experiments, assuming the sample swap.

S5 Figure

rRNA depletion has no significant effect on the transcriptome

The four graphs show individual comparisons of reads per million reads for replicate RNA samples with or without rRNA depeltion using RNaseH and rRNA-complementary oligonucleotides. The effects of rRNA depletion were similar to variation between replicates and no mRNA was significantly affected (S2 Table).

S6 Figure

Differentiation of bloodstream forms lacking PUF3.

(A) RNAi targeting PUF3 in Lister 427 bloodstream forms gave a transient differentiation defect. RNAi was induced 24 h before cells attained a density of 1.8-2.0 $\times 10^6$ cells/ml. Cis-aconitate (CA) was added to the culture (final concentration 6 mM) and the temperature was reduced from 37°C to 27°C. The time in hours after CA treatment is indicated. The Western blot shows expression of EP procyclin and V5-tagged PUF3 with ribosomal protein S9 as a control.

(B) Inducible expression of N-terminally myc-tagged PUF3 in five Lister 427 bloodstream-form clones. The protein was barely detectable.

(C) Expression of PAD1 and RBP10 after *PUF3* RNAi in EATRO1125 bloodstream forms. The experiment was as in Fig 5B.

S7 Figure

Plasmids made for this work