Depletion of the Trypanosome Pumilio Domain Protein PUF2 or of Some Other Essential Proteins Causes Transcriptome Changes Related to Coding Region Length

Bhaskar Anand Jha,* Abeer Fadda, Clementine Merce, Elisha Mugo, Dorothea Droll, Christine Clayton
Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance, Heidelberg, Germany

Pumilio domain RNA-binding proteins are known mainly as posttranscriptional repressors of gene expression that reduce mRNA translation and stability. *Trypanosoma brucei* has 11 PUF proteins. We show here that PUF2 is in the cytosol, with roughly the same number of molecules per cell as there are mRNAs. Although PUF2 exhibits a low level of *in vivo* RNA binding, it is not associated with polysomes. PUF2 also decreased reporter mRNA levels in a tethering assay, consistent with a repressive role. Depletion of PUF2 inhibited growth of bloodstream-form trypanosomes, causing selective loss of mRNAs with long open reading frames and increases in mRNAs with shorter open reading frames. Reexamination of published RNASeq data revealed the same trend in cells depleted of some other proteins. We speculate that these length effects could be caused by inhibition of the elongation phase of transcription or by an influence of translation status or polysomal conformation on mRNA decay.

Pumilio domain proteins are involved in various aspects of RNA metabolism and are found throughout eukaryotic evolution (1–3). They usually contain approximately 8 Puf repeats, each of which binds a single nucleotide with a consensus target sequence of UGUA (4). Pumilio proteins were first discovered as regulators of differentiation in *Drosophila melanogaster*, and subsequently they have been found to function in differentiation of other animal species (3, 5). Most pumilio proteins studied so far have been found to be cytosolic, but there are also conserved pumilio proteins that are located in the nucleolus and are involved in rRNA processing (1, 6).

The yeast *Saccharomyces cerevisiae* has 6 cytosolic pumilio domain proteins, none of which is essential: even cells lacking Pufs1, 2, 3, 4, and 5p are still viable, although they have extensive changes in gene expression (7). The yeast cytosolic Puf proteins assist in mRNA localization, suppress translation, and promote mRNA decay. Puf1p and Puf5p bind to adjacent sites in the 3’-untranslated region (3’-UTR) of the HO mRNA, repressing its expression and causing deadenylation (8); they seem to act via interaction with the eukaryotic initiation factor 4E (eIF4E)-binding proteins Caf20p and Eap1p (9). Puf1p acts in conjunction with Puf3p (10), which is involved in the localization of mRNAs that encode mitochondrial proteins to the mitochondrial membrane (11). Meanwhile, Puf6p is involved in translational repression and localization of the ASH1 mRNA (12). Pumilio domain proteins in malaria parasites effect stage-specific translational repression (13, 14). In contrast, some pumilio proteins have also been shown to increase expression: for example, in *Caenorhabditis elegans* olfactory sensory neurons, Fas-binding factor 1 (FBF-1) promotes translation of egl-4 mRNA by binding to its 3’-UTR (15).

African trypanosomes are excellent models for the study of posttranscriptional regulation: transcription is polycistronic, with no apparent control of RNA polymerase II initiation. Individual mRNAs are produced by trans splicing and polyadenylation (16). Two life cycle stages are routinely cultured in the laboratory: the bloodstream form, which is similar to the forms that multiply in mammalian blood and tissue fluids, and the procyclic form, which is similar to the parasites in the tsetse fly midgut. Different mRNA levels—and also regulation during differentiation—are achieved by modulation of mRNA decay (17) and trans-splicing (18–20). Trypanosome mRNA decay pathways are similar to those of other eukaryotes: mRNAs are deadenylated by a CAF1-NOT complex (21), then the mRNA body is attacked from the 3’ end by the exosome (22–24) and in the 5’-3’ direction by XRNA (25). Depletion of the deadenylase CAF1 decreases deadenylation and therefore inhibits degradation of most mRNAs (22, 26); depletion of the core exosome subunit RRP45 decreases 3’-5’ degradation by the exosome (22, 23). Meanwhile, depletion of XRNA stabilizes mRNAs that are normally very unstable (17). Numerous trypanosome RNA-binding proteins have been implicated in the control of mRNA fate (27–29).

*Trypanosoma brucei* has 11 pumilio domain proteins. Three of them, PUF7, PUF8, and PUF10, are in the nucleolus and have been implicated in aspects of rRNA metabolism (30), although PUF10 also has functions in expression of mRNAs that are synthesized by RNA polymerase I (31). Neither PUF1 (32) nor PUF5 (33) is essential in procyclic forms, and they may also be dispensable in bloodstream forms; their roles are not known. PUF9 binds to, and stabilizes, a few mRNAs in the S phase of the cell cycle (34). We describe here studies of PUF2.

**MATERIALS AND METHODS**

**Trypanosome culture.** Culturing and transfection of trypanosomes (Lister 427 strain expressing the tet repressor from phD1313) [1313 try-
panosomes) were performed as described in reference 35. For the expression of tetracycline-inducible genes, 100 ng ml−1 of tetracycline (Tet) (36) was used. For monitoring the growth, procyclic cells were grown in minimal essential medium-ProS medium at 27°C and were diluted to a density of 5 × 10^5 cells ml−1. Bloodstream-form trypanosomes were grown in HMI-9 medium at 37°C with 5% CO2 and diluted to 1 × 10^6 cells ml−1. To monitor growth after RNA interference (RNAi), trypanosome cultures were initially diluted and grown without selecting antibiotics, then diluted to 2 × 10^6 cells ml−1 for bloodstream forms and 5 × 10^6 cells ml−1 for procyclic forms before tetracycline addition. Thereafter, cells were similarly diluted daily as required. Cells with MKT1 RNAi have been described elsewhere (37).

**Plasmid constructs.** Plasmids and oligonucleotides are listed in Tables S1A and S1B, respectively, in the supplemental material.

To express V5-tagged PUF2 from the PUF2 locus, 400-bp long fragments from the 5'UTR and the open reading frame (ORF) of PUF2 were amplified by using primers CZ3539 (SacII) and CZ3540 (XbaI) and CZ3537 (Xhol) and CZ3538 (ApaI), respectively. The PCR products were ligated into the plasmid Bla/V5 (38). The resulting plasmid, pH2D180, was digested with SacII and ApaI preceding transfection into bloodstream-form and procyclic-form cells. For in situ tagging using eYFP (enhanced yellow fluorescent protein), a 708-bp fragment from the C terminus of PUF2 was amplified using the primers CZ4024 and CZ4025. The resulting PCR fragment was ligated into the KpnI and BamHI sites of p3043 (39) [kind gift from Mark Carrington]. The resulting plasmid, pH2D230, was linearized with EcoNI before use in transfecting bloodstream-form trypanosomes.

For tandem affinity purification (TAP) tagging, the PUF2 ORF was amplified using the primers CZ3942 (HindIII) and CZ3943 (HpaI), digested with the corresponding restriction enzymes, and ligated into the linearized pH918 (24) to generate pH2D226. After linearizing with NotI, pH2D226 was transfected into 1313 bloodstream-form trypanosomes.

RNA interference against PUF2 was done by making a stem-loop RNAi plasmid as described in reference 36. A 432-bp region (bp 399 to 831 of the ORF) of the PUF2 gene was selected for RNAi (the sequence was verified for its specificity by using the RNAi program [http://trypanofan.path.cam.ac.uk/software/RNAi.html]). The stem-loop PUF2 RNAi vector was referred to as pH2D2183, and after digestion with NotI the vector was used to transfect 1313 bloodstream- and procyclic-form cells.

For tethering, the PUF2 ORF was PCR amplified using the primers CZ3913 (ApaI) and CZ4025 (BamHI) and ligated into vector pH2D202 (40) to give pH2D204. The plasmid was cotransfected into induced expression of the fusion protein, AN-PUF2-myc, into cell lines constitutively expressing a chloramphenicol acetyltransferase (CAT) reporter with an actin 3'-UTR (pHD1991) and box B actin 3'-UTR (pHD2277). CAT assays were done as described in reference 41.

**Antibody production and affinity purification.** The first 942 bp of the PUF2 ORF from the start codon (excluding the PUF domain) were cloned in the pQTEV-based (Addgene, Cambridge, MA) vector pH1D746 (10X His tags), to give pH2D2181. Escherichia coli BL21 cells were used to express the recombinant protein. The PUF2 fragment was purified from the soluble fraction under native conditions by using a Ni-nitrilotriacetic acid superflow apparatus (Qiagen, Germany). Antibody against the purified PUF2 protein fragment was raised in a rabbit (Charles River Laboratories, Kisslegg, Germany). Antibodies were affinity purified using His-tagged PUF2 bound to nitrocellulose membrane (42). To check the sensitivity and specificity, 0.1 to 5.0 ng of the recombinant protein was loaded on a 12% SDS-PAGE, and the blot was probed with antibody diluted to 1:10,000. The antibody was able to detect a minimum of 0.2 ng of protein by Western blotting with little cross-reactivity with other trypanosome proteins, and cleavage of the His tag did not affect the signal (data not shown).

**Indirect immunofluorescent staining and Western blotting.** For immunofluorescent staining, the steps of slide preparation and immunofluorescent staining were done as described in reference 43, with incubation at room temperature with gentle shaking unless otherwise specified. The Z-stacks of the cells were obtained and the images were deconvoluted (Wiener algorithm) using the Olympus Cell-R microscope.

Antibodies used for immunofluorescence and/or Western blotting were the following: anti-PUF2 (this study), anti-myc (Santa Cruz Biotechnology), anti-aldolase (44), anti-peroxiredoxin (kind gift from Luis Krauth-Siegel), anti-SCD6 (kind gift from Mark Carrington), and anti-V5 (AbD SeroTec).

**Northern blotting.** The 1313 bloodstream-form trypanosomes in the logarithmic growth phase (0.8 × 10^9 to 1.2 × 10^9 cells ml−1) were harvested. RNA was prepared using Trizast reagent (Peqlab, Germany). Ten micrograms of total RNA was run on a denaturing formaldehyde agarose gel and then blotted on a nylon membrane (Hybond-N; Amersham, GE Healthcare). Radioactive DNA probes (Prime-IT RmT random primer labeling kit; Stratagene) were used for hybridization, and the signals were detected by using a phosphorimager (FLA 7000; Fujifilm).

**Cell fractionation and polysome preparation.** Approximately 5 × 10^6 1313 bloodstream-form cells (density, ~1 × 10^6 cells ml−1) were fractionated as described in reference 45. Polysomes were purified from the same number of cells according to the method described in reference 46. The polysomal samples were trichloroacetic acid precipitated and resuspended in 4X Laemmli sample buffer. Blots were incubated with antibodies against PUF2, trypanothione reductase (TR; kind gift from L. Krauth-Siegel’s lab), and small ribosomal subunit protein S9.

**TAP and commounprecipitation.** A total of 1 × 10^8 bloodstream-form cells expressing the C-terminally TAP-tagged PUF2 (TAP-PUF2) were harvested after 24 h of induction with 100 ng ml−1 tetracycline. Tandem affinity purification was done as described previously (47). The proteins were separated by SD-SPACE, and the entire lane from the TAP-PUF2 purification was examined by mass spectrometry. Commounprecipitation was done as described in reference 48. Proteins were detected by Western blotting using antibodies against V5 tag (1:2,000; AbD seroTec), the Myc tag (1:2,000; Santa Cruz Biotechnology), and PUF2 (1:2,000; present study).

**Cross-linking and immuno precipitation.** PUF2 cross-linking to RNA followed by immunoprecipitation (CLIP) was performed. The method used was a combination of the CLIP (49) and the PAR-CLIP (photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation) (50, 51) methods with minor changes and without in vivo labeling with 4-thio-uracil (52). Approximately 2 × 10^6 bloodstream-form trypanosomes expressing V5-PUF2 were used for the pulldown assay. An equal number of 1313 bloodstream-form cells without any tagged protein expression was taken as a negative control. Around 2 × 10^6 bloodstream-form cells were harvested at a density of approximately 1.0 × 10^6 cells ml−1. All cultures were cross-linked two times with 0.4 J cm−2 254-nm UV light in a Stratalinker apparatus with mixing in between. Immunoprecipitation was done using anti-V5 antibody-coupled agarose beads (Bethyl Laboratories, TX). RNase T1 (10 U ml−1 of the original bead volume; Fermentas, Germany) used to digest the RNA-protein complexes bound to the beads. The remaining RNAs were dephosphorylated, followed by 5'-end labeling with 32P. A 4 to 12% Novex NuPAGE bis-Tris gel (Novex, Germany) was used to separate the RNA-protein complexes. Aliquots of around 5% of the lysate and unbound and bound fractions were used for Western blotting to check the efficiency of the immunoprecipitation.

**Transcriptome analysis.** For the RNAi experiment, trypanosomes were grown for 1 day in the presence of tetracycline, and poly(A)+ mRNA was sheared and prepared for sequencing according to the standard Illumina protocol. One PUF2 RNAi batch was sequenced on the GAII and the second with HiSeq. Data analysis, including motif searches and enrichment calculations, were performed as described previously (22, 54). All data were log transformed before the correlation coefficient calculations.
FIG 1 Expression and localization of PUF2. (A) Amounts of PUF2 in cells. A 1:10,000 dilution of the antibody was used to probe 5 × 10⁶ bloodstream-form wild-type cells, along with cells overexpressing TAP-tagged PUF2 and PUF2 RNAi cells. Different amounts of purified PUF2 fragment were used: 5.0 to 200 ng was loaded to check the sensitivity and specificity of the antibody. The antibody was also used to estimate the expression level of the protein (in terms of number of molecules per cell). (B) Intracellular location of V5-PUF2. Bloodstream-form cells expressing V5-PUF2 were used. Anti-V5 antibody was used to detect the localization of V5-PUF2 protein, while peroxiredoxin was used as a control for cytoplasmic localization. Cells without a V5-tagged protein were included as negative control.

in order to reduce the influence of outliers and correct for a nonnormal distribution of the data.

RESULTS
PUF2 is a cytoplasmic protein that is essential for normal growth of bloodstream-form trypanosomes. PUF2 expression was examined in two ways. First, a sequence encoding a V5 tag was integrated in frame to the 5′ end of the open reading frame. Later, we expressed the N-terminal 314 amino acids (without the PUF domains) as a His-tagged fusion in E. coli and raised an antibody. Figure 1A (lane 1) shows that the antibody recognized a protein from bloodstream-form trypanosomes that migrated slightly faster than the 100-kDa marker. In cells in which one PUF2 gene was N-terminally tagged with a sequence encoding the V5 epitope, the total level of PUF2 remained similar (Fig. 1A, lane 2). (The tag is too short to affect mobility.) Cells bearing a plasmid carrying a version with a tandem affinity tag showed the appearance of an additional, slower-migrating band after tetracycline addition (Fig. 1A, lanes 3 to 6). The specificity of the antibody was demonstrated by the absence of a signal after RNAi targeting of PUF2 (Fig. 1A, compare lanes 7 and 8). To obtain an approximate estimate of the abundance of PUF2 in bloodstream-form cells, we loaded different amounts of cell extract and the recombinant protein onto gels and compared the signals by Western blotting. In several experiments, 5 ng (1.5 × 10⁳ molecules) of the recombinant protein fragment gave a signal of similar intensity to that from 5 × 10⁶ cells (an example is shown in Fig. 1A), which means that there are probably 1 × 10⁴ to 2 × 10⁵ molecules of PUF2 per bloodstream-form cell or 0.5 to 1 molecules of PUF2 per mRNA (55). Unfortunately, we were unable to obtain any soluble PUF2 for in vitro studies.

An immunofluorescence assay using the bloodstream-form cells with constitutive expression of V5-PUF2 showed that the tagged protein was localized in discrete foci in the cytoplasm (Fig. 1B). The cytoplasmic localization was confirmed in an immunofluorescence using the affinity-purified PUF2 antibody (see Fig. S1 in the supplemental material) and PUF2 tagged at the C terminus with eYFP (see Fig. S2A and B in the supplemental material). This was consistent with a role in mRNA metabolism. The pattern for PUF2-YFP was clearly distinct from that seen with a soluble protein, peroxiredoxin (see Fig. S2C), but neither V5-PUF2 nor PUF2-YFP formed granules after heat shock and neither colocalized perceptibly with the stress granule marker SCD6 (see Fig. S2D). The levels of PUF2 in bloodstream forms were not affected by 1-h incubations at 16°C or 41°C or by treatment with the protease inhibitor MG132 (10 μg/ml) (data not shown).

To examine the role of PUF2, we decreased its expression by RNAi. PUF2 knock-down lead to a growth defect in the bloodstream form of the parasite after 2 days of induction. (Fig. 2; see also Fig. S3 in the supplemental material). No effect on procyclic
growth was seen. These results are consistent with those previously seen in a high-throughput RNAi screen (56).

To find out whether PUF2 interacts with other proteins, we subjected a tagged version to tandem affinity purification. In an initial experiment, we examined only those bands that were absent in the control (TAP-only) purification. In the second experiment, all copurifying proteins were analyzed. To find out whether proteins were specifically associated with PUF2, we compared the new results with those for other purifications (unpublished results) (57). Numerous proteins copurified specifically with PUF2, but none of them was found in both purifications (see Table S2 in the supplemental material). Of those found in the second experiment, which was likely to yield more comprehensive results, several were associated with RNA metabolism. They included several other proteins with RNA-binding domains: ZC3H41, RBP20, RBP33, ALB4, and another protein with two weak RNA recognition motifs. It was tempting to speculate that the Nudix hydrolase encoded by Tb927.11.9810 could be the decapping enzyme, but this sequence matches *S. cerevisiae* Ysa1P (E-values of 7.5e-19 and 3.3e-18 for BLASTP and reciprocal BLASTP, respectively). Results from mutagenesis experiments indicate that Ysa1P is involved in ADP ribose and NAD<sup>−</sup> metabolism (58). Its closest human homologue is Nudt5, which lacks decapping activity (59).

**Some PUF2 is bound to RNAs in vivo.** In order to verify the RNA-binding ability of PUF2, a modified in vivo CLIP analysis was performed. The cells were irradiated at 254 nm, which leads to covalent bond formation between RNA and bound proteins. After immunoprecipitation, all unbound RNA except that protected by the protein was digested. Remaining RNAs were end labeled with <sup>32</sup>P, and RNA-protein complexes were detected after denaturing gel electrophoresis. As a negative control, we took extracts from cells with no V5-tagged proteins and immunoprecipitated the extracts by using anti-V5 antibody-coated beads. A strong background smear was seen (Fig. 3A). As a positive control, we used extracts from cells expressing V5-UBP1 (53, 60). The control Western blot showed that most of the V5-UBP1 was attached to the beads (Fig. 3B), and the autoradiogram revealed a clear band of <sup>32</sup>P-labeled material running at about 40 kDa that was absent from the negative control (Fig. 3A). Using cells expressing V5-PUF2, the anti-V5 immunoprecipitation was apparently not quite as efficient as for V5-UBP1 (Fig. 3B), but there was nevertheless a clear signal at the expected position of 100 kDa that was absent in both the negative and positive controls. Thus, PUF2 is clearly able to bind mRNAs.

To determine the nature of the bound mRNAs, we first tried to amplify RNA from the CLIP experiment, but we were unable to obtain sufficient products. Therefore, we repeated the immunoprecipitation of V5-tagged PUF2 and sent the RNA for sequencing without prior RNase treatment or gel electrophoresis. Cells expressing Myc-tagged and V5-tagged UBP1 served as controls, since we expected UBP1 to bind to different sets of mRNAs. After UV cross-linking and immunoprecipitation with appropriate antibody-coupled beads, the bound RNAs were purified and sequenced. For each cell line, RNA in the unbound fraction served as the control. Results are shown in Table S3 of the supplemental material. Although RNAs were specifically selected by the procedures, there was almost no difference between the mRNAs selected by V5-PUF2 and those selected by V5-UBP1. Either these two proteins have similar specificities or the binding is independent of V5-tagged protein. Surveys of the selected 3′-UTRs using MEME suggested enrichment of the sequence AUAAU (E-value, 6.3e-1,220), and Motifsampler revealed UUGUUGGU (consensus score, 1.76; information content, 1.68).

**PUF2 is mostly not associated with polysomes and can decrease levels of bound mRNA.** To find out whether PUF2 was associated with polysomes in bloodstream-form cells, we subjected extracts to sucrose gradient centrifugation (Fig. 4A) and examined protein distributions. Ribosomal protein S9 (from the small ribosomal subunit [positive control]) was enriched in the polysomal fractions (Fig. 4B and C). Although PUF2 signal was also observed in the polysomal fractions, the distribution was similar to that of the negative control, TR (Fig. 4B and C). Therefore, only a small proportion, if any, of the PUF2 pool is associated with polysomes.
Eukaryotic Cell

We next assayed the ability of PUF2 to increase or decrease the abundance and translation of associated mRNA. To do this, we used a bloodstream-form cell line that expresses a CAT mRNA sequence and therefore "tethers" any protein to which it is fused to the myc tag on its C terminus (Fig. 5C). We later reexamined the cell lines and found that expression of the AN-PUF2-myc protein was now easily detected in the absence of tetracycline (Fig. 5D). We therefore compared CAT expression with that for lines carrying no inducible AN plasmid. Again, the presence of AN-PUF2-myc had no effect in the absence of boxB (Fig. 5E); with boxB, the AN-PUF2-myc caused a decrease in CAT activity to 30%. Previous results showed that tethering of other, neutral proteins has no significant effect on reporter expression (37, 54).

Although it is clearly possible that the artificial tethering, or the presence of the N- and C-terminal tags, could affect PUF2 function, the tethering result suggests that PUF2 may decrease the abundance of mRNAs to which it is bound. The decrease in CAT protein was similar to the decrease in the RNA, suggesting that the CAT protein was reduced because there was less CAT RNA. We cannot, however, rule out the possibility that PUF2 first inhibits translation; in this case, the loss of binding of the mRNA by ribosomes or translation factors could lead to destruction of the mRNA as a secondary effect. Both mechanisms would be consistent with the fact that PUF2 is mainly not associated with polysomes. If PUF2 binding leads to RNA destruction, the bound RNA will not be translated anymore, and if PUF2 inhibits translation initiation, the RNA will not be associated with polysomes.

**PUF2 depletion results in specific increases and decreases in mRNAs.** To find out whether PUF2 expression influences steady-state mRNA levels, we examined the transcriptomes of PUF2-depleted cells (see Table S4 in the supplemental material). The experiment was done in duplicate. We found 65 open reading frames that were more than 2-fold increased in both experiments, and 107 that were reproducibly decreased. The proteins encoded by the increased mRNAs (Fig. 6A) were preferentially located in the mitochondrion (28%, P value 4.2e−5); moreover, five of the proteins that were of unknown function and location had strongly predicted mitochondrial targeting signals. Meanwhile, the down-regulated genes (Fig. 6B) were enriched in those encoding cytoskeletal proteins, including several that are known to be in the flagellum (19%; P = 1e−4).

In RNaseq experiments, the amounts of RNAs are relative. That means that by default, a large increase in some mRNAs leads to apparent decreases in others, and vice versa. We therefore used Northern blotting, with stable RNA standards, to examine the decreases in mRNAs coding for paralflagellar rod proteins 1 and 2 (see Fig. S4A and B in the supplemental material), two intraflagellar transport proteins (see Fig. S4C and D), CCCH zinc finger protein ZC3H41 (see Fig. S4E), RAD50 (see Fig. S4F), the RNA recognition motif protein RBP20 (see Fig. S4G), and cyclin 6 (see Fig. S4H). In most cases, the results were confirmed, although the effects were modest. An increase in the zinc finger protein ZC3H20 mRNA was observed (see Fig. S4I), but two other RNA-binding protein mRNAs that supposedly increased were not much affected (see Fig. S4J and K). The results showed that mRNAs both increase and decrease after PUF2 RNAi.

**The protein coding region or mRNA length affects the response to RNAi.** While examining the results, we noticed that the proteins encoded by the mRNAs that increased were all rather large, while those that decreased seemed small. For example, mRNAs encoding dynein heavy chains were increased, while those encoding the intermediate and light chains were unaffected. We therefore plotted the effect of PUF2 RNAi against the coding se-
sequence length. Remarkably, there was a strong negative correlation: short open reading frame mRNAs increased, and long ones decreased (Fig. 7A; see also Table S4 in the supplemental material). The apparent differential effects of PUF2 RNAi on mRNAs for mitochondrial and cytoskeletal proteins could be artifactual, since the former have shorter protein coding regions than the latter (see Fig. S5A). In an attempt to find mRNAs whose regulation was not just related to open reading frame length, we calculated a correction factor (see Fig. S5B and C and Table S3, sheet 3, in the supplemental material). A total of 75 mRNAs were reproducibly at least 2-fold increased after the length correction, of which 18% encoded mitochondrial proteins. For the 177 reproducibly decreased mRNAs, 8 (4.5%) were cytoskeletal and 9 (5%) were mitochondrial. The expression differences are small, however, and could reflect stochastic variations in the length effect.

We wondered whether protein coding region length would always influence expression in cells experiencing growth inhibition, so we compared our results with RNASeq or microarray data from other cell lines with RNAi that targeted expression of essential proteins (see Table S4, sheet 4, in the supplemental material).

FIG 5 Tethering of PUF2 leads to a decrease in reporter the mRNA level. (A) Schematic representation of the RNA tethering strategy. The left panel shows that the fusion protein binds to the reporter mRNA through the interaction between AN (in the fusion protein) and the stem-loop structure (boxB sequence, in the reporter). The reporter has five stem-loop structures, but for simplicity only one is shown in the figure. When the boxB is not present in the mRNA (3'-UTR), the fusion protein is unable to bind. (B, upper two panels) Expression levels of the AN-PUF2-myc fusion proteins, measured by Western blotting with an anti-myc antibody. Tetracycline was included in the cultures for 24 h where indicated. Aldolase was used as a loading control. (Lower two panels) A Northern blot showing expression of CAT mRNA in the same experiment, with methylene blue-stained rRNA as a loading control. (C) Expression of CAT activity in cells with or without expression of AN-PUF2-myc, after tetracycline induction for 1 day. Results are from three different experiments, with standard deviations shown. A typical AN-PUF2-myc expression result for this experiment is shown in panel B. (D) Expression of AN-PUF2-myc fusion proteins, in the same cell lines as used for panel B (the details are described above in the panel B legend), but after storage in liquid nitrogen. (E) Expression of CAT activity and mRNA in cells with expression of AN-PUF2-myc, with tetracycline induction for 1 day. A typical AN-PUF2-myc expression result for this experiment is shown in panel D. Results are from three different experiments, and standard deviations are shown.
MKT1 is partially associated with polysomes and may be required for translation (37). We were startled to find that there was a correlation between the effects of RNAi targeting PUF2 and those for MKT1 (Fig. 7B) and that the MKT1 RNAi effect also showed an influence of coding region length (see Fig. S6A in the supplemental material). Intriguingly, RNAi targeting CTR9, which might be involved in transcription elongation (57), showed an even stronger effect (Fig. 7C). There was no correlation with 3′-UTR length (R² value always under 0.12). ZC3H11 is required mainly for stabilization of chaperone transcripts in procyclic forms, but RNAi causes rapid death of bloodstream forms (54), while RBP10 levels correlate with expression of a subset of blood-stream-form-specific mRNAs (52). The effects of depleting ZC3H11 showed only a minor length influence (Fig. 7D), while for RBP10 there was none at all (see Fig. S6B). Neither stumpy forms (63) (see Fig. S6C) nor bloodstream forms grown to high density (63, 64) (data not shown) showed preferential loss of long coding sequence transcripts. Thus, the length effect is not simply a result of growth inhibition.

We also examined data obtained after depletion of mRNA degradation enzymes. After depletion of the CAF1 deadenylase or the core exosome subunit RRP45 (22), the steady-state levels of short open reading frame mRNAs were preferentially increased in the same way as for PUF2 (Fig. 7E; see also Fig. S6C in the supplemental material). In clear contrast, after depletion of the 5′-3′ exonuclease XRNA (17), longer-coding-region mRNAs were preferentially retained (Fig. 7E). We determined whether these effects depended on the coding region or the overall mRNA length by repeating the analysis using the annotated “processed mRNA” length. The correlation coefficients are compared in Fig. 7, and scatter plots are provided in Fig. S7 in the supplemental material. The results were similar, which was not really surprising since the mRNA length is strongly influenced by the coding region length. It therefore is not possible to distinguish between the two possibilities. There was no correlation with 3′-UTR length (the lengths used are provided in Table S4 in the supplemental material).

Finally, in an attempt to find PUF2-specific targets that were not related to the 3′-5′ decay pathway, we searched for mRNAs that were increased after PUF2 depletion but not after depletion of PAN2, CAF1, or RRP45. Only 41 genes qualified, of which 39% encoded mitochondrial proteins. However, many of these also showed increased mRNA after RNAi targeting MKT1, RBP10, or CTR9 (see Table S4, sheet 5, in the supplemental material). It is therefore difficult to tell which, if any, of the PUF2 RNAi effects on the transcriptome are direct.

**DISCUSSION**

**Role of PUF2 in bloodstream-form trypanosomes.** Trypanosome PUF2 is required for normal growth of bloodstream forms, can bind to RNA, is not associated with polysomes, and can suppress the level of both mRNA and protein in a tethering assay. The number of PUF2 molecules per cell, 10 to 20,000, is of the same order of magnitude as the number of mRNA molecules (55). The results from tethering experiments suggested that PUF2 may target a subset of mRNAs for degradation.

Identification of PUF2 targets through analysis of RNA that coprecipitated with V5-PUF2 yielded results that were difficult to interpret, since the bound mRNAs looked very similar to those selected when we used V5-tagged UBP1. Motif searches in the 3′-UTRs of putative target mRNAs did not reveal any canonical pumilio-binding domains (UGUA) but did select some sequences with U alternating with A or G. The significance of this is dubious, since the specificity of the precipitation was highly questionable. Both the inability to find mRNAs that are specifically associated with PUF2 and the lack of polysome association of PUF2 are consistent with a role for PUF2 in stimulating decay of its targets, since its binding will lead to destruction of the RNA target.

An alternative method to find PUF2 targets was to examine the transcriptome after PUF2 RNAi, in the expectation that the targets would be increased. We initially thought that PUF2 depletion caused specific increases in mRNAs encoding cytoskeletal proteins and decreases in mRNAs encoding mitochondrial proteins. Upon closer scrutiny, though, we observed a remarkable correlation between the effects of the RNAi and the coding sequence length. Long-coding-region mRNAs were depleted, while mRNAs with short coding regions increased in abundance. Some of the losses and increases were confirmed by Northern blotting. Intriguingly, the same phenomenon was seen in cells in which deadenylation pathways (CAF1, CNOT10, PAN2), or the exosome (RRP45), were depleted, and also after RNAi targeting both MKT1 and CTR9. MKT1 and CTR9 do not have the same function as PUF2—indeed, MKT1 increased mRNA levels in a tethering assay (37), the opposite effect to that of PUF2. One possibility is that the length-related effects of PUF2 depletion are indirect, through reductions in mRNAs encoding degradation machinery components. Indeed, PUF2 RNAi reduced the NOT1 mRNA by 70%, but...
it is difficult to distinguish cause and effect in this case, since the NOT1 coding region is 7 kb long.

One possibility is that trypanosome PUF2 depletion triggers a more general pathway that results in length-dependent changes in mRNA abundance. However, at present, this conclusion is very speculative and the mechanism is unknown.

The influence of coding region length on RNAi responses. The effects of RNAi targeting of 3’ degradation pathways, PUF2, CTR9, and MKT1 on mRNA abundance were all inversely correlated with coding region or mRNA length. Since the mRNA length mostly correlates with the coding region length, we do not know which is the determining factor. We initially suspected that the length effect was an artifact from RNASeq, since similar phenomena have not been observed among trypanosome microarray data. However, several different control data sets were used, over a period of more than 2 years, using both GAII and HiSeq technolo-
gies, and the most recent PUF2 experimental samples (with and without tetracycline) were sequenced together in the same run. Also, depletion of XRNA, which inhibits 5'-3' mRNA degradation, had the opposite effect (an increase in the longer mRNAs), and the effects of ZC3H11 depletion did not correlate with coding sequence length. In all experiments, the RNA was fragmented before library building. It is therefore unclear how a technical artifact of this sort could arise.

In wild-type trypanosomes, there is no relationship between coding region length and RNA decay rate (see Fig. S5E in the supplemental material), and it is difficult to envisage a mechanism by which either polyadenylation or trans-splicing could be affected by the mRNA length. The strongest length-dependent effect, as judged by both correlation coefficients and the slope of the regression line, was from depletion of CTR9, which likely functions as a transcription elongation factor (57); the next strongest were for the exosome subunit RRP45 and the deadenylase CAF1. (We appreciate that variations in the effectiveness of RNAi could influence this.) Therefore, one possibility is that the length dependence is linked to transcription elongation. Under this scenario, a slowing in elongation would lead to a decrease in the abundances of longer mRNAs, and the effects of depleting other proteins would be caused by some sort of feedback inhibition that decreased the transcription rate. It is already known that RNAi against CAF1 does not result in an overall increase in mRNA abundance (22), even though it almost abolishes RNA decay. This suggests some sort of cross talk between factors required for mRNA decay and transcription, as has already been documented for yeast (65–68). Defects in the CAF1-NOT complex might inhibit mRNA turnover sufficiently to reduce the supply of nucleotides; alternatively, more specific feedback mechanisms could exist. XRNA depletion has different consequences, perhaps because it affects only a small subset of low-abundance, unstable mRNAs (17).

An alternative explanation for the length effects is that there are CDS-length-dependent differences in translation initiation or elongation. It has long been known that mRNAs can circularize in vivo or in ovo, or in translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. Nucleic Acids Res. 38:8039–8050. http://dx.doi.org/10.1093/nar/gkq686.

Ulbricht RJ, Olivas WM. 2008. Puf1p acts in combination with other yeast Puf proteins to control mRNA stability. RNA 14:246–262. http://dx.doi.org/10.1080/09739794.2010.1216847.

11. Nop9 is an RNA binding protein present in pre-40S ribosomes and required for 18S rRNA synthesis in yeast. RNA 13:2165–2174. http://dx.doi.org/10.1261/rna.747607.

12. Olivas W, Rarker G. 2000. The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. EMBO J. 19:6602–6611. http://dx.doi.org/10.1093/emboj/19.23.6602.

13. Hook BA, Goldstrohm AC, Seay DJ, Wickens M. 2007. Two yeast PUF proteins negatively regulate a single mRNA. J. Biol. Chem. 282:15430–15438. http://dx.doi.org/10.1074/jbc.m612153200.

14. Miller MA, Olivas WM. 2011. Roles of Puf proteins in mRNA degradation and translation. Wiley Interdiscip. Rev. RNA 2:471–492. http://dx.doi.org/10.1002/wrna.69.

15. Cridge AG, Castelli LM, Smirnova JB, Selley JN, Rowe W, Hubbard SJ, McCarthy JE, Ashe MP, Grant CM, Pavitt GD. 2011. Identifying elf4E-binding protein translationally-controlled transcripts reveals links to human PUMILIO proteins. Structure 19:104–112. http://dx.doi.org/10.1016/j.str.2010.12.019.

16. Mandel-Gutfreund Y, Michaeli S. 2008. Puf1p acts in combination with other yeast Puf proteins to control mRNA stability. RNA 14:246–262. http://dx.doi.org/10.1080/09739794.2010.1216847.

17. Miller MA, Olivas WM. 2011. Roles of Puf proteins in mRNA degradation and translation. Wiley Interdiscip. Rev. RNA 2:471–492. http://dx.doi.org/10.1002/wrna.69.

18. Ulbricht RJ, Olivas WM. 2008. Puf1p acts in combination with other yeast Puf proteins to control mRNA stability. RNA 14:246–262. http://dx.doi.org/10.1080/09739794.2010.1216847.
and has been published in a scientific journal.
57. Ouna B, Nyambega B, Manful T, Helbig C, Males M, Fadda A, Clayton C. 2012. Depletion of trypanosome CTR9 leads to gene expression defects. PLoS One 7(4):e34256. http://dx.doi.org/10.1371/journal.pone.0034256.

58. Tong L, Lee S, Denu JM. 2009. Hydrolase regulates NAD+ metabolites and modulates cellular redox. J. Biol. Chem. 284:11256–11266. http://dx.doi.org/10.1074/jbc.M809790200.

59. Song M-G, Li Y, Kiledjian M. 2010. Multiple mRNA decapping enzymes in mammalian cells. Mol. Cell 40:423–432. http://dx.doi.org/10.1016/j.molcel.2010.10.010.

60. Hartmann C, Clayton C. 2008. Regulation of a transmembrane protein gene family by the small RNA-binding proteins TbUBP1 and TbUBP2. Mol. Biochem. Parasitol. 157:112–115. http://dx.doi.org/10.1016/j.molbiopara.2007.09.009.

61. Keryer-Bibens C, Barreau C, Osborne HB. 2008. Tethering of proteins to RNAs by bacteriophage proteins. Biol. Cell 100:125–138. http://dx.doi.org/10.1042/BC20070067.

62. De Gregorio E, Preiss T, Hentze MW. 1999. Translation driven by an eIF4G core domain in vivo. EMBO J. 18:4865–4874. http://dx.doi.org/10.1093/emboj/18.17.4865.

63. Jensen B, Sivam D, Kifer C, Myler P, Parsons M. 2009. Widespread variation in transcript abundance within and across developmental stages of Trypanosoma brucei. BMC Genomics 10:482. http://dx.doi.org/10.1186/1471-2164-10-482.

64. Queiroz R, Benz C, Fellenberg K, Hoheisel JD, Clayton C. 2009. Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. BMC Genomics 10:495. http://dx.doi.org/10.1186/1471-2164-10-495.

65. Haimovich G, Medina DA, Causse SZ, Garber M, Millan-Zambrano G, Barkai O, Chavez S, Perez-Otín JE, Darzacq X, Choder M. 2013. Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153:1000–1011. http://dx.doi.org/10.1016/j.cell.2013.05.012.

66. Goler-Baron V, Selitrennik M, Barkai O, Haimovich G, Lotan R, Choder M. 2008. Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. Genes Dev. 22:2022–2027. http://dx.doi.org/10.1101/gad.473608.

67. Harel-Sharvit L, Eldad N, Haimovich G, Barkai O, Duck L, Choder M. 2010. RNA polymerase II subunits link transcription and mRNA decay to translation. Cell 143:552–563. http://dx.doi.org/10.1016/j.cell.2010.10.033.

68. Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC. 2011. The multifunctional Ccr4-Not complex directly promotes transcription elongation. Genes Dev. 25:581–593. http://dx.doi.org/10.1101/gad.2020911.

69. Preiss T, Hentze MW. 1999. From factors to mechanisms: translation and translational control in eukaryotes. Curr. Opin. Genet. Dev. 9:515–521. http://dx.doi.org/10.1016/S0959-437X(99)00005-2.

70. Christensen AK, Kahn LE, Bourne CM. 1987. Circular polysomes predominate on the rough endoplasmic reticulum of somatotropes and mammotropes in the rat anterior pituitary. Am. J. Anat. 178:1–10.

71. Kopeina GS, Afonina ZA, Gromova KV, Shirokov VA, Vasiliev VD, Spirin AS. 2008. Step-wise formation of eukaryotic double-row polyribosomes and circular translation of polysomal mRNA. Nucleic Acids Res. 36:2476–2488. http://dx.doi.org/10.1093/nar/gkm1177.

72. Beilharz TH, Preiss T. 2007. Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. RNA 13:982–997. http://dx.doi.org/10.1261/rna.031898.111.

73. Park EH, Zhang F, Warringer J, Sunnerhagen P, Hinnebusch AG. 2011. Depletion of eIF4G from yeast cells narrows the range of translational efficiencies genome-wide. BMC Genomics 12:68. http://dx.doi.org/10.1186/1471-2164-12-68.