Parasitology

cambridge.org/par

Insights into the functions and RNA binding of *Trypanosoma brucei* ZC3H22, RBP9 and DRBD7

Esteban Erben*, Kevin Leiss, Bin Liu†, Diana Inchaustegui Gil, Claudia Helbig and Christine Clayton

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

Research Article

*current address: Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín (IIBIO-UNSAM)-CONICET, 25 de Mayo y Francia, San Martín, Buenos Aires, Argentina.

†current address: Hebei Viroad Biotechnology Co., Ltd., Huanghedadao 136, Shijiazhuagn 050035, Peoples Republic of China.

Cite this article: Erben E, Leiss K, Liu B, Gil DI, Helbig C, Clayton C (2021). Insights into the functions and RNA binding of *Trypanosoma brucei* ZC3H22, RBP9 and DRBD7. *Parasitology* **148**, 1186–1195. https://doi.org/10.1017/S0031182021000123

Received: 27 October 2020
Revised: 8 December 2020
Accepted: 6 January 2021
First published online: 4 February 2021

Key words:
mRNA; RNA-binding protein; *Trypanosoma brucei*

Author for correspondence:
Christine Clayton,
E-mail: cclayton@zmbh.uni-heidelberg.de

Abstract

Trypanosoma brucei is unusually reliant on mRNA-binding proteins to control mRNA fate, because its protein-coding genes lack individual promoters. We here focus on three trypanosome RNA-binding proteins. ZC3H22 is specific to Tsetse fly forms, RBP9 is preferentially expressed in bloodstream forms; and DRBD7 is constitutively expressed. Depletion of RBP9 or DRBD7 did not affect bloodstream-form trypanosome growth. ZC3H22 depletion from procyclic forms caused cell clumping, decreased expression of genes required for cell growth and proliferation, and increased expression of some epimastigote markers. Apart from decreases in mRNAs encoding enzymes of glucose metabolism, levels of most ZC3H22-bound transcripts were unaffected by ZC3H22 depletion. We compared ZC3H22, RBP9 and DRBD7 RNA binding with that of 16 other RNA-binding proteins. ZC3H22, PUF3 and ERBP1 show a preference for ribosomal protein mRNAs. RBP9 preferentially binds mRNAs that are more abundant in bloodstream forms than in procyclic forms. RBP9, ZC3H5, ZC3H30 and DRBD7 prefer mRNAs with long coding regions; UBP1-associated mRNAs have long 3′-untranslated regions; and RRM1 prefers mRNAs with long 3′or 5′-untranslated regions. We suggest that proteins that prefer long mRNAs may have relatively short or degenerate binding sites, and that preferences for A or U increase binding in untranslated regions.

Introduction

Kinetoplastids depend heavily on post-transcriptional mechanisms for control of gene expression, because transcription of nearly all their protein-coding genes is polycistronic. *Trypanosoma brucei (T. brucei)* is usually cultivated as bloodstream forms (the form that lives in mammals) and as procyclic forms (the form that lives in the midgut of Tsetse flies). In *T. brucei*, messenger RNA (mRNA) processing efficiency, translation and decay are all influenced by RNA-binding proteins (Clayton, 2019). In trypanosomes, as in other organisms (Khong and Parker, 2020), mRNAs are associated with many proteins. Some, including poly(A) binding proteins and cap-binding translation initiation complexes, are present on many different mRNAs, whereas others are specialized to promote or repress decay or translation of more specific mRNA subsets. In the past few years numerous studies have examined the binding specificities of *Trypanosoma brucei* RNA-binding proteins and their effects of their depletion or over-expression on gene expression. These have revealed some proteins that are required for normal homoeostasis, and others that are required for differentiation, or in particular life-cycle stages (Clayton, 2019).

We here studied three proteins that repress expression of a reporter protein when artificially 'tethered' to a reporter mRNA (Erben *et al.*, 2014; Lueong *et al.*, 2016). In this assay, we use trypanosomes that constitutively express a reporter mRNA that contains five copies of the bacteriophage lambda 'boxB' sequence. The protein of interest is expressed as a fusion with the lambdaN peptide, which binds to the boxBs in the reporter with high affinity. Repression of reporter expression suggests that the protein that is tethered either inhibits translation, or promotes mRNA degradation, or both. For example, RBP10, which is a repressor in the tethering assay, binds to specific mRNAs in bloodstream forms *via* its RNA Recognition Motifs (RRMs) and causes target mRNA destruction (Mugo and Clayton, 2017).

ZC3H22 (Tb927.7.2680) has two $C(x)_8C(x)_5C(x)_3H$ zinc finger RNA-binding domains. ZC3H22 is absent in bloodstream-form trypanosomes, appearing only after differentiation into the procyclic form. RNA-Seq results demonstrate the persistence of *ZC3H22* mRNA in the Tsetse fly proventriculus, and a decrease in the salivary glands. Its RNAi-mediated depletion in procyclic forms caused a strong growth defect (Domingo-Sananes *et al.*, 2015). The *ZC3H22* gene is immediately upstream of those encoding two other CCCH zinc finger proteins, ZC3H20 and ZC3H21, which activate expression of procyclic-specific proteins *via* recruitment of MKT1, part of a protein complex that stabilizes mRNAs and enhances translation (Liu *et al.*, 2020). Although the sequence of ZC3H22 is partially related to those of ZC3H20 and ZC3H21, there are sufficient differences to predict that it will not bind the same mRNAs; and ZC3H22 also lacks the MKT1 interaction motif (Liu *et al.*, 2020).

CAMBRIDGE UNIVERSITY PRESS

RBP9 (Tb927.11.12120) has a single RRM. It is of interest because its mRNA is at least ten times more abundant in *T. brucei* bloodstream forms (the form that lives in mammals) than in procyclic forms (the form that lives in the midgut of tsetse flies). Its binding to mRNA may be quite weak (Lueong *et al.*, 2016) and probably occurs *via* a single RRM. DRBD7 (Tb927.4.400) has two RRMs and is clearly bound to mRNA in bloodstream forms (Lueong *et al.*, 2016), but it is not strongly developmentally regulated at the mRNA or protein levels. RNAi targeting RBP9 or DRBD7 in bloodstream forms was reported to result in minor growth defects (Wurst *et al.*, 2009; Alsford *et al.*, 2011) and expression of RBP9 in procyclic forms was toxic (Miguel De Pablos *et al.*, 2017).

We analysed the effects of RNAi targeting ZC3H22 in procyclic forms, and RBP9 or DRBD7 in bloodstream forms, and characterized the mRNAs to which each is bound. Since the results from the RNA-binding studies were somewhat unexpected, we compared them with those for 15 other *T. brucei* proteins. This uncovered some interesting preferences related to the lengths of the coding and untranslated regions.

Materials and methods

Cell lines and growth

Plasmids and oligonucleotides are listed in Supplementary Table S1. Procyclic-form trypanosomes were routinely cultured in MEM Pros medium in which the only glucose comes from the fetal calf serum. For some experiments with the *ZC3H22* RNAi cell line, glucose was added to a final concentration of 10 mM. *ZC3H22*, *DRBD7* and *RBP9* genes were tagged within their endogenous locus to retain the endogenous 3′-UTR sequence, in the hope that native expression levels would be maintained. For this, the 5′end of each open reading frame and a region encompassing their 5′ untranslated region just upstream of the start codon were amplified by polymerase chain reaction (PCR) and subcloned into pENT6B-TAP (Kelly *et al.*, 2007) and pBS-BLAV5 (Shen *et al.*, 2001). For ZC3H22, we also generated a single knock-out line with puromycin resistance. For RNAi, a specific attB-tagged gene fragment for either *RBP9* or *DRBD7* was amplified and cloned into pGL2084 by Gateway recombination (Jones *et al.*, 2014). The resulting plasmids were digested with BamHI and HindIII and the fragment containing the stem-loop was subcloned into the pHD1146 plasmid as previously described (Bajak *et al.*, 2020*b*). For *ZC3H22* RNAi, various gene-specific fragments were cloned into p2T7-TAblue and a stem-loop construct was also generated; details are in Supplementary Table S1. RNA interference was induced by adding 100–250 ng mL^{-1} tetracycline.

The amounts of live cell material after *ZC3H22* RNAi were measured after incubation with Resazurin (Begolo *et al.*, 2018). All other methods were as previously described in Liu *et al.* (2020).

RNA binding studies

To find mRNAs associated with RNA-binding proteins, we tagged each at the N-terminus *in situ* with a tag that can be cleaved with Tobacco Etch Virus (TEV) protease (Puig *et al.*, 2001). Cleared extracts were allowed to adhere to an IgG column, then the associated tagged protein was eluted by TEV cleavage (Mugo and Clayton, 2017). Proteins in the unbound and eluate fractions were protease-digested, RNA was purified using TRIfast (VWR), and the RNA was sent for high-throughput sequencing. The accession numbers for the transcriptome RNA pull-down data are E-MTAB-9092 (DRBD7), E-MTAB-9093 (RBP9) and E-MTAB-6906 (ZC3H22). For ZC3H22 RNAi the number is E-MTAB-9705. Reads were aligned to the TREU927 and Lister427 (2018) genomes using TrypRNASeq (Leiss *et al.*, 2016), allowing each read to align once. Differential expression after RNAi was analysed using DESeqU1 (Leiss and Clayton, 2016), a user-friendly version of DeSeq2 (Love *et al.*, 2014). For this, we worked with a list of unique coding regions, in order to avoid distortions caused by repeated genes. We had previously estimated the effective gene copy numbers by comparing results when reads were aligned once, with those obtained when the same reads were allowed to align 20 times. Before analysis of the unique gene set, the reads from individual genes in the unique list were multiplied by the calculated gene copy numbers. This ensures that the large amounts of mRNA from repeated genes are assigned appropriate weights. The heat maps for RNA binding were generated from log RPM ratios, after excluding genes with low coverage, using trypclusterviewer (Mulindwa *et al.*, 2018). All details are in Supplementary Figure legends.

Results

Depletion of ZC3H22 causes trypanosome clumping and decreases mRNAs required for cell growth and division

To analyse the functions of RBP9, DRBD7 and ZC3H22, we depleted them by RNAi and examined the effects on trypanosome proliferation. To monitor the effectiveness of the RNAi, we used trypanosomes in which one gene copy had been deleted and the other had been tagged *in situ* by integration of a sequence encoding an N-terminal V5 tag or a tandem affinity purification (TAP) tag. Depletion of either DRBD7 or RBP9 in bloodstream-form trypanosomes did not have any obvious effect on trypanosome morphology or proliferation (Supplementary Fig. 1A and B). Either the two proteins are not essential – perhaps their functions can be replaced by other proteins – or else, they are essential but low levels are sufficient to maintain function. Since the depletion had no obvious effects, we did not analyse these cells further.

Another group previously reported that RNAi-mediated depletion of ZC3H22 inhibited procyclic-form trypanosome proliferation (Domingo-Sananes *et al.*, 2015). We had considerable difficulty repeating this result (see Supplementary Table S1). However, we did finally see an altered phenotype in procyclic cells that had only one *ZC3H22* allele, and RNAi driven by opposing tetracycline-inducible T7 promoters. This line also had a V5 tag integrated at the N-terminus of the remaining *ZC3H22* gene. Western blot, Northern blot and transcriptome analysis showed reduced *ZC3H22* mRNA and protein, no effect on mRNA from *ZC3H21* and a slight increase in *ZC3H20* mRNA (Supplementary Fig. S1C, D and Table S2). In the presence of tetracycline, most of the cells were present in longitudinally aligned aggregates (Fig. 1A). The same behaviour was also detected, to a lesser extent, in the absence of tetracycline. Accurate counting was impossible, but measurements using a live-cell fluorimetric assay after 3 days of RNAi induction revealed no significant differences in the total live cell content between the hemizygous cells and those containing the RNAi construct, with or without tetracycline. This was true whether we grew cells in our standard, low glucose medium or medium supplemented with glucose, as used in the previous publication. We do not know why our results differ from those previously reported, but the differences in either the starting cell line or the culture conditions might be responsible.

The very clear morphological changes in ZC3H22-depleted cells prompted us to examine them in more detail. We initially obtained single transcriptome datasets from the original *ZC3H22* hemizygous line, the RNAi line grown without

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

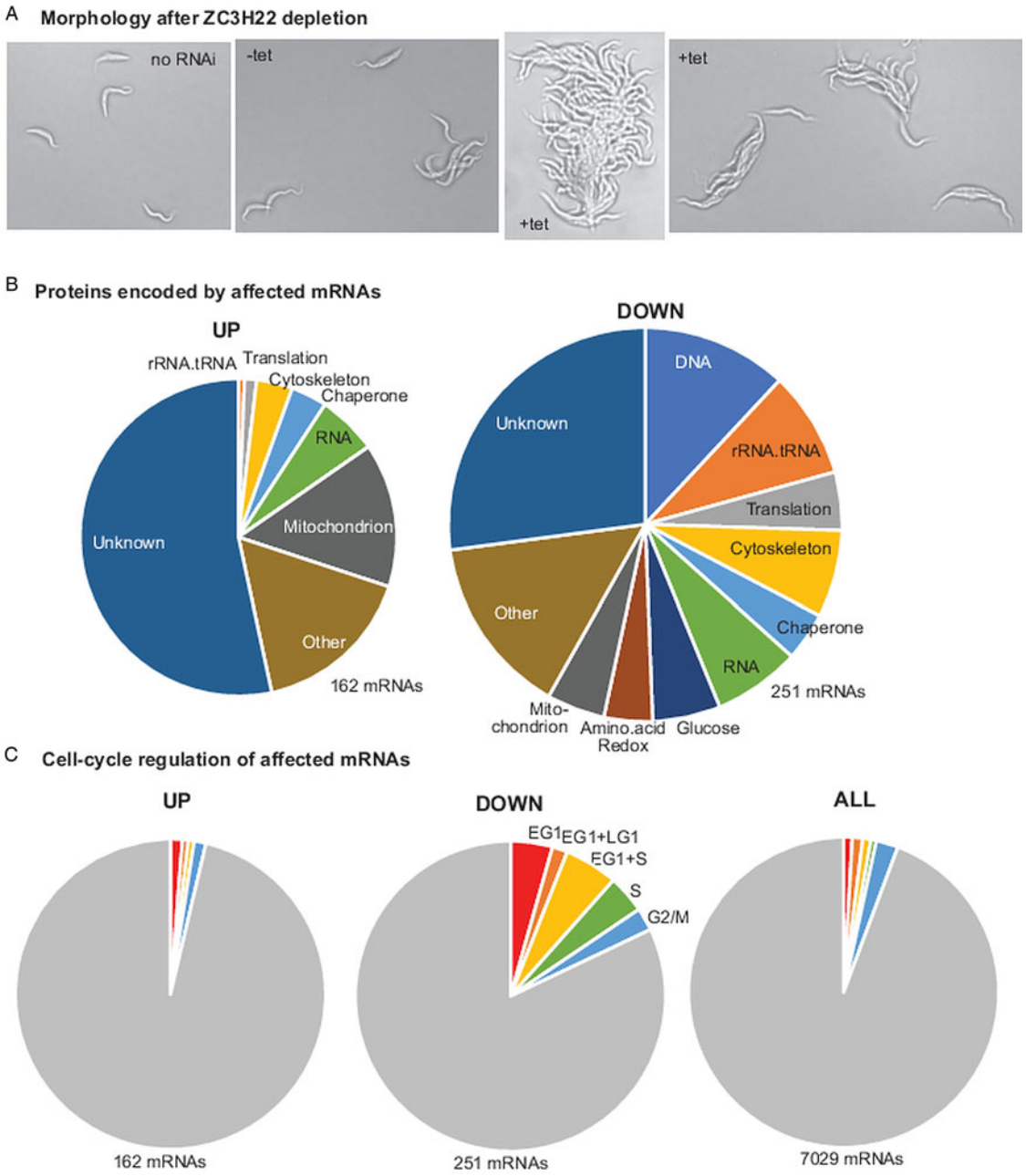

Fig. 1. Depletion of ZC3H22 decreases mRNAs implicated in cell growth and division. (A) Live-cell differential interference contrast images. One image ('no RNAi') is of trypanosomes hemizygous for ZC3H22, and with no RNAi plasmid. The other images are of cells with inducible RNAi, grown either without (−tet) or with (+tet) tetracycline. (B) Categories of protein products encoded by mRNAs that were increased or decreased by ZC3H22 depletion. For the analysis, we chose mRNAs that increased or decreased at least 1.5-fold, with an adjusted *P* value of less than 0.05. The areas of the charts are to scale. The categories of individual mRNAs are listed in Supplementary Table S1. For simplicity, several categories are grouped together in the figures. These are as follows. 'DNA' includes proteins implicated in nucleotide metabolism, DNA synthesis, chromatin, cell cycle control and nuclear-cytoplasmic transport. 'rRNA.tRNA' are enzymes implicated in tRNA and rRNA processing and modification. 'Translation' includes translation factors and amino-acyl tRNA synthetases. 'Chaperone' includes both protein-refolding chaperones and the Tric complex, which folds newly-synthesized proteins. 'RNA' includes RNA-binding proteins, helicases, enzymes of mRNA processing and of RNA synthesis. 'Glucose' includes enzymes of glucose and glycerol metabolism. Other categories are self-explanatory. (C) Cell cycle regulation of mRNAs that were increased or decreased by ZC3H22 depletion. EG1 = early G1. LG1 is late G1. The classifications were taken from Archer *et al.* (2011). Areas are not to scale. In the 'decreased' category cell cycle-regulated mRNAs were significantly decreased relative to the total transcriptome (Fisher test *P* value 2.5×10^{-12}) mainly because of dencreases in mRNAs with maximum expression in G1 and S phases (*P* value 4.3×10^{-16}).

tetracycline and the RNAi line growth with tetracycline, each grown either in our standard low-glucose medium, or the same medium supplemented with 10 mM glucose. To look for effects of glucose, we compared the data for cells without RNAi induction (the hemizygous line without RNAi and the line with RNAi but no tetracycline). This revealed no mRNAs that were significantly (Padj [adjusted *P* value]< 0.05) changed after glucose addition. We therefore treated the results with and without glucose as replicates for subsequent comparisons. (For sample details see Supplementary Table S2, sheet 9.) In comparison with the starting cell line, the *ZC3H22* mRNA level in the RNAi cell line was reduced even in the absence of tetracycline, explaining the slight aggregation phenotype. We therefore compared the transcriptomes of cells with induced *ZC3H22* RNAi with those of controls with no integrated RNAi plasmid (Supplementary Table S2). In total, 162 different mRNAs were significantly increased at least 1.5-fold after RNAi. Prominent among the encoded products were mitochondrial proteins implicated in the

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

citric acid cycle and electron transport (Fig. 1B). Interestingly, they also included the epimastigote surface protein BARP and the regulator RBP6, ectopic expression of which induces differentiation of procyclic forms to epimastigotes (Kolev et al., 2012). The downregulated mRNAs encoded numerous proteins involved in gene expression and DNA replication (Fig. 1B); correspondingly, these were strongly enriched in mRNAs that are maximally abundant during the G1 and *S* phases of the cell cycle (Fig. 1C).

ZC3H22 mRNA binding

To find out which mRNAs are bound by ZC3H22 in procyclic forms, and therefore might be directly regulated by ZC3H22, we used cells expressing tagged ZC3H22 from the endogenous locus; the other allele was deleted. Lysates were incubated with IgG beads, to allow binding by the protein A portion of the TAP tag and the unbound fraction was collected. After washing, the bound tagged protein, along with its associated RNA, was released using TEV protease, which cleaves within the tag. RNA was purified from both the unbound fraction and the eluate (bound) fractions, and sequenced for two biological replicates. To increase mapped reads, we depleted rRNAs from the unbound fraction, so enrichment of rRNAs could not be measured. In addition, the RNA was fragmented and size-selected (~300 nt) before library building, so small structural and catalytic RNAs are not reliably detected. Since the method we used has relatively low stringency (a single purification step), it might allow purification of mRNAs that are bound *via* protein interaction partners of the tagged protein, as well as those bound to the tagged protein itself.

We then looked for RNAs that were enriched in the bound, relative to the unbound fraction. In each case we considered only those mRNAs showing a minimal level of enrichment in each of two replicates; the threshold varied as described below. Only 16 mRNAs were even 1.5-fold enriched in the ZC3H22 pull-down, so to examine the data we set the threshold at 1.3-fold enrichment. Although the zinc-finger motifs and surrounding sequences of ZC3H22 resemble those of ZC3H20, there was no correlation between their binding specificities (Fig. 2A). ZC3H22 showed a significant preference (Fisher *P* value 3×10^{-24}) for mRNAs encoding ribosomal proteins (Fig. 2B, Supplementary Fig. S2B), which comprise 21 of the 74 mRNAs that were enriched at least 1.3-fold. These 74 mRNAs have longer than average half-lives (Fig. 2C). Overall, the enriched mRNAs were not significantly developmentally regulated (Fig. 2D). Nevertheless, enrichment of mRNAs of glucose and glycerol metabolism (Supplementary Fig. S2B, Fisher *P* value 1×10^{-6}) was concentrated on proteins that are implicated in procyclic-form glucose metabolism. These were enolase and several glycosomal enzymes: malate dehydrogenase, glycerol-3-phosphate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, phosphoenolpyruvate carboxykinase, glucose-6-phosphate isomerase, and hexokinase. Intriguingly, the mRNAs encoding additional proteins of energy metabolism – F1 ATPase, delta-1-pyrroline-5-carboxylate dehydrogenase and malic enzyme – were also enriched. We found no enriched linear motifs in the bound mRNAs.

The enrichment of mRNAs with ZC3H22 was poor, so the results have to be interpreted with the utmost caution. Also, any effects of ZC3H22 on translation alone would not have been detected in our experiments. Nevertheless, we next compared the 74 preferentially associated mRNAs with those that had been affected by ZC3H22 depletion. There was no overall correlation (Supplementary Fig. S2A) and none of the 163 mRNAs that increased after RNAi was in the 'ZC3H22-bound' category. Interestingly, however, 13 of the 251 mRNAs that *decreased* were indeed ZC3H22-associated. Moreover, half of these were mRNAs encode enzymes of glucose metabolism: hexokinase, NAD-linked glycerol-3-phosphate dehydrogenase, phosphoenolpyruvate carboxykinase, enolase, phosphoglycerate kinase (PGKB) and glycosomal malate dehydrogenase. Of these, all but *PGKB* were also shown to be downregulated in trypanosomes of the proventriculus compared to those of the midgut (Savage *et al*., 2016), and *PGKB* may have been missed because Savage *et al*., did not distinguish between the three PGK isoforms. The mRNAs encoding 5 other enzymes of glucose metabolism that were also decreased both after *ZC3H22* RNAi and in the proventriculus were not associated with ZC3H22. These were ATP-dependent phosphofructokinase, Aldose-1-epimerase, triosephosphate isomerase, fumarate hydratase, and glyceraldehyde 3-phosphate dehydrogenase. The remaining bound and decreased mRNAs were mostly implicated in aspects of DNA replication or translation.

Overall, the results suggested that ZC3H22 might bind to, and stabilize, mRNAs implicated in procyclic-form glucose metabolism. Loss of such mRNAs might affect the energy balance, triggering other changes associated with epimastigote differentiation. However, we had previously found that tethering of ZC3H22 to a toxic reporter mRNA in bloodstream forms suppressed its expression, giving a growth advantage to cells if the reporter mRNA was toxic (Supplementary Fig. S3A). The results of screens in which fragments of ZC3H22 were tethered had also suggested that the suppressive activity lay towards the C-terminus of the protein (Supplementary Fig. S3B). Since ZC3H22 is not normally expressed in bloodstream forms, we repeated the assay in procyclic forms. No effect at all was seen (Supplementary Fig. S3C). A version lacking the C-terminal myc tag also had no effect (not shown). It may be that ZC3H22 is differently modified in procyclic forms (Urbaniak *et al*., 2013); or the result from bloodstream forms could have been an artefact of unknown aetiology. Assuming that ZC3H22 indeed has no capacity to increase or decrease expression directly in procyclic forms, its effects might depend directly on its RNA binding alone. Under this scenario, in procyclic forms ZC3H22 may compete with a destabilizing factor for binding to mRNA targets. Loss of ZC3H22 would then enable the destabilising factor to bind and cause targeted mRNA decay.

RNA binding by RBP9 and DRBD7

To find out which mRNAs are bound by RBP9 and DRBD7 in bloodstream forms, we used the same procedure as for ZC3H22, except that the second gene copies were not deleted. Comparison of the bound and unbound mRNAs revealed 121 mRNAs that were at least 2-fold enriched in both RBP9 purifications and 605 for DRBD7. No conserved motifs were detected in the bound mRNAs. Binding of both RBP9 and DRBD7 to mRNAs positively correlated with overall mRNA length (Supplementary Fig. S4). Interestingly, the correlation was predominantly with the coding region length (Supplementary Fig. S4D and E), rather than with the 3′-UTR length (Supplementary Fig. S4H) or 5′-UTR length (see later). In all, 49 of the 121 RBP9-bound mRNAs (at least 2× in each replicate) encoded cytoskeletal proteins (Fisher *P* value 2×10^{-31}, 49 out of the 121); this probably reflects the frequency of long coding regions in this functional class (Supplementary Fig. S5). For DRBD7, no functional class, apart from GRESAG mRNAs, was significantly enriched in the bound fraction, although median binding of cytoskeletal mRNAs was again above the 75th percentile (Supplementary Fig. S5). The mRNAs encoding ribosomal proteins were not bound by either protein (Supplementary Fig. S5). Intriguingly, however, the RBP9-bound mRNAs were significantly more abundant in bloodstream forms than in

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

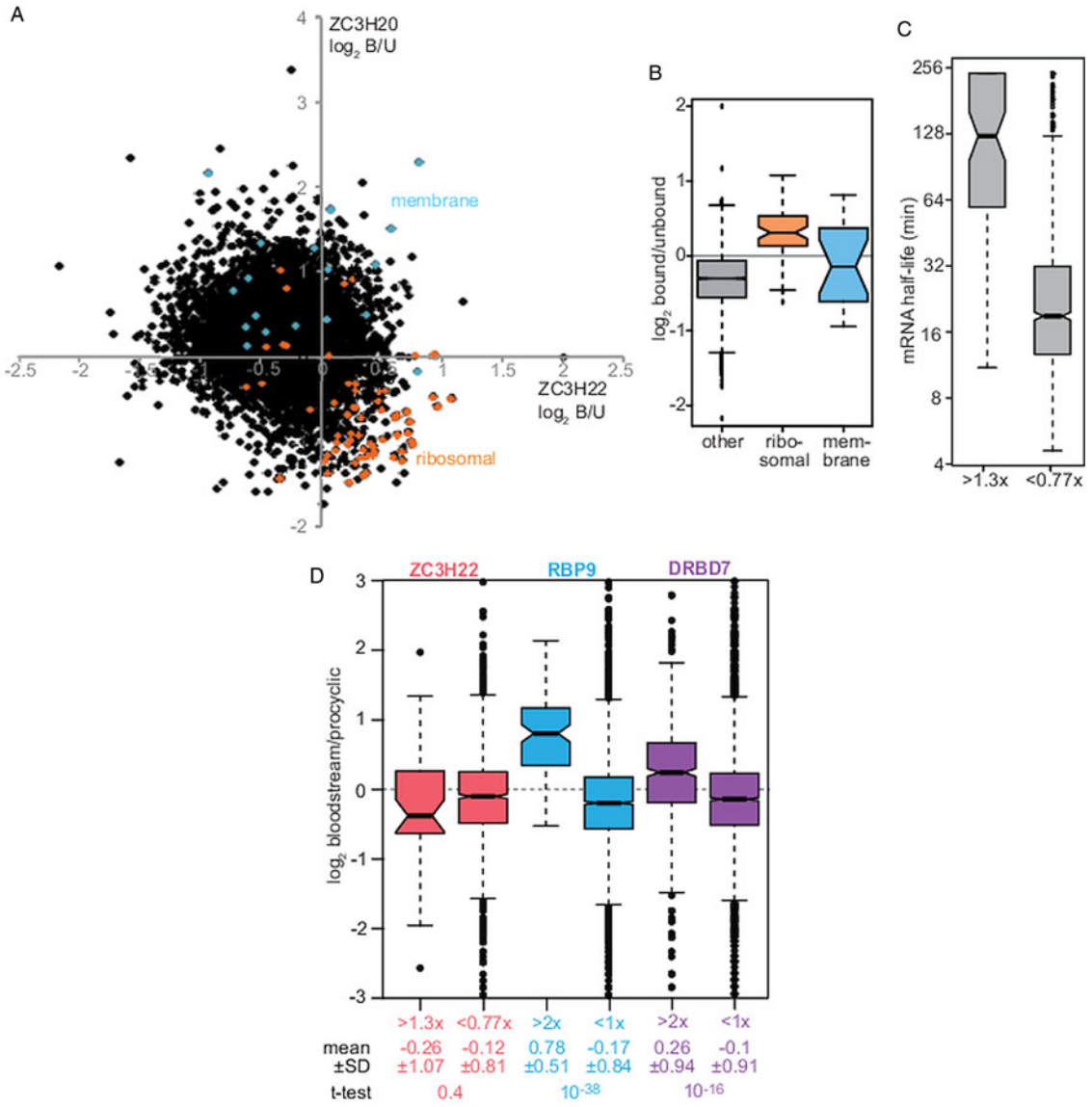

Fig. 2. Functions and regulation of mRNAs bound to ZC3H22, RBP9 and DRBD7. All analyses are of a set of unique genes, to avoid over-counting of repeated genes and gene families. Details are in Supplementary Table S2. (A) Average mRNA binding for ZC3H20 (Liu *et al*., 2020) and ZC3H22. Membrane protein mRNAs, some of which are bound by ZC3H20 and ribosomal protein mRNAs are highlighted. (B) Boxplot showing relative binding of different mRNA categories by ZC3H22. The box extends from the 25th to the 75th percentile, with the median indicated. The whiskers extend to points up to 1.5-fold the inter-quartile range, and dots are outliers. (C) Boxplot showing the half-lives (Fadda *et al*., 2014) of ZC3H22 bound (>1.3 ×) and unbound (<0.77 ×) mRNAs. (D) Boxplot showing relative abundances of bound and unbound mRNAs in the bloodstream and procyclic forms. For ZC3H22 the groups were as in (C), and for DRBD7 and RBP9, the 'bound' fraction was reproducibly at least 2-fold enriched while the 'unbound' fraction was reproducibly less than 1-fold enriched. Some more extreme outliers are not shown.

procyclic forms, with an almost 2-fold difference relative to the mRNAs in the unbound fraction (Fig. 2D). Although DRBD7-bound mRNAs were also significantly more abundant in bloodstream forms, the median difference (1.3-fold) is small and may not be biologically meaningful.

Motif searches for proteins that bind long mRNAs are problematic, because the unbound mRNAs that should act as controls are shorter than the bound ones, so are not directly comparable. We therefore tried comparing of the coding regions of mRNAs bound by RBP9 or DRBD7 with scrambled versions of the same sequences. For comparison we also tried the same for ZC3H5, which also binds mRNAs with long coding regions (see below). However, in each case this yielded variants of GGAGGA, and it is unlikely that all three proteins have the same specificity. Moreover, for the longest 120 coding regions with less than 1.5 × enrichment with RBP9, similar sequences were retrieved, suggesting that GGAGGA is simply a preferred motif in coding regions.

When RBP9 was inducibly expressed in procyclic forms, cell proliferation was inhibited and transcriptome changes occurred that suggested a loss of developmental regulation (Miguel De Pablos *et al*., 2017). There was no relationship between those reported transcriptome changes and the 122 mRNAs that were bound to RBP9 in bloodstream forms: 7 of the bound mRNAs decreased and 6 increased. Although it is possible that RBP9 RNA binding differs between bloodstream and procyclic forms, we suggest that the more likely explanation is that the transcriptome changes seen after ectopic RBP9 expression were secondary to growth inhibition.

Several RNA-binding proteins preferentially associate with long or short mRNAs

The results that we had obtained so far did not give much insight into the functions of the proteins investigated. Most of the mRNAs bound by ZC3H22 were not affected by RNAi, and it

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

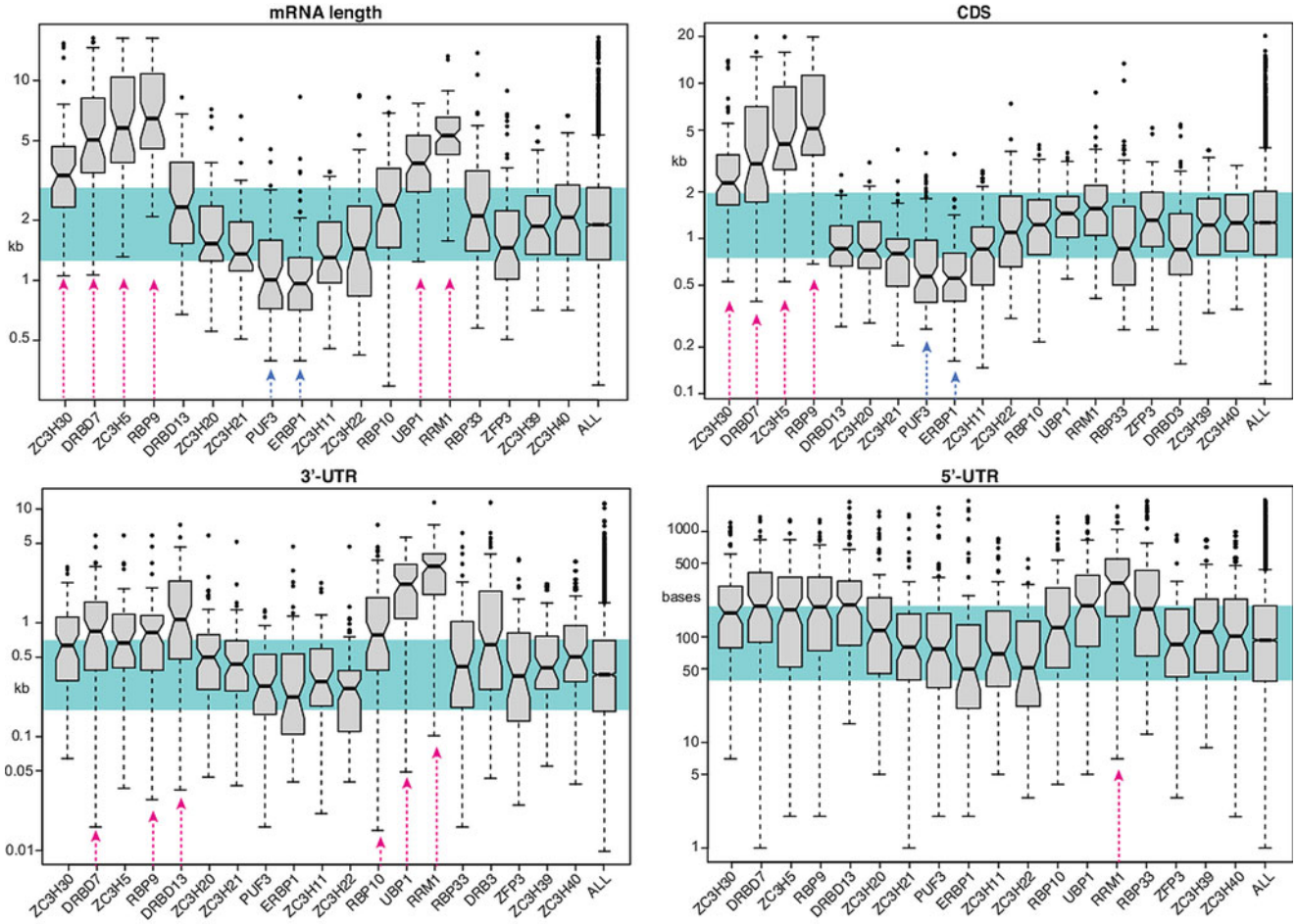

Fig. 3. Different RNA-binding proteins select mRNAs with different length characteristics. The 100 most enriched mRNAs were selected for each of the listed RNA-binding proteins, except for ZC3H22 and DRBD3 (see text). The annotated lengths of the mRNAs are shown in box plots. Some of the bound mRNAs lack annotated 5′-UTR or 3′-UTR lengths. The aqua box shows the inter-quartile range for all mRNAs. The magenta arrows indicate data for which the median is longer than the overall 75th percentile and the blue arrows, those for which the median is shorter than the overall 25th percentile.

selected ribosomal protein mRNAs; and for DRBD7 and RBP9, RNAi had no effect on growth, and they had a bias towards binding long RNAs. We had seen both types of RNA-binding preference before (Chakraborty and Clayton, 2018; Bajak *et al.*, 2020*a*, 2020*b*; Kamanyi Marucha and Clayton, 2020). Gene length has also been shown to cause technical artefacts in mammalian RNA-Seq datasets, resulting in artificial over-representation of very long and very short genes in the subsets that are significantly differentially regulated (Mandelboum *et al.*, 2019). In order to place our new results in context, and assess their specificity, we therefore decided to compare all available RNA binding results (Supplementary Table S3). These were for ERBP1 (Bajak *et al.*, 2020*b*), PUF3 (Kamanyi Marucha and Clayton, 2020), ZC3H20 and ZC3H21 (Liu *et al.*, 2020), ZC3H11 (Droll *et al.*, 2013), DRBD13 (Jha *et al.*, 2015), RBP10 (Mugo and Clayton, 2017), ZC3H5 (Bajak *et al.*, 2020*a*), ZC3H30 (Chakraborty and Clayton, 2018) and ZC3H32 (Klein *et al.*, 2017); PUF2 (Jha *et al.*, 2014), UBP1 (Jha *et al.*, 2014), RRM1 (Naguleswaran *et al.*, 2015), RBP33 (Fernandez-Moya *et al.*, 2014), ZFP3 (Walrad *et al.*, 2011), DRBD3 (Das *et al.*, 2015), ZC3H39 and ZC3H40 (Trenaman *et al.*, 2019).

First, we examined the effect of length (Fig. 3). Since the methods that had been used to find bound mRNAs varied between experiments and laboratories, as did the enrichment ratio thresholds, we chose to compare the 100 mRNAs that had been most enriched with each protein. Exceptions were for ZC3H22, where only the 79 that were at least 1.3-fold enriched were included; and DRBD3, where mRNAs with peaks in the CDS or 3′-UTR were counted. This analysis clearly showed the preferences of DRBD7, ZC3H5 and RBP9 for long coding regions (Fig. 3). RRM1, UBP1 and DRBD13 prefer long 3′-UTRs; RRM1, intriguingly, also bound mRNAs with long 5′-UTRs. ERBP1- and PUF3-bound mRNAs were significantly shorter than average: they both enrich ribosomal protein mRNAs, which are mostly relatively short (Clayton, 2019).

Quantitative comparison of RNA binding preferences

To dissect similarities and differences more quantitatively, we restricted our analysis to results that had all been obtained by the same method: affinity purification of N-terminally TAP-tagged protein and release through cleavage of the tag. Nearly all of these reproducibly pulled down their own (or 'self') coding mRNA (Table 1). This is consistent with purification of the encoding mRNA *via* the nascent polypeptide. We do not know why this was not seen with RBP9, ZC3H11 and ZC3H30. Results for mRNPs purified by other methods were more mixed (Table 1), which is expected: an N-terminal tag is particularly well placed for nascent polypeptide purification, whereas a C-terminal tag should exclude it.

We first used a customized script (Mulindwa *et al.*, 2018) to construct a heat map using all of the similar datasets. We tried various numbers of clusters; with 60, clear differentiation of RNA binding by RBP10 to 3–4 clusters was seen, and this analysis

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

Table 1. Co-purification of the encoding mRNA during purification of RNA-binding proteins

RBP	Life-cycle stage	Replicate 1	Replicate 2	Replicate 3	Purification
TAP-DRBD7	BS	115	125.3		TAP (1-step)
TAP-DRBD13	PC	4.9			TAP (2-step)
TAP-PUF3	BS	4.6	1.9		TAP (1-step)
TAP-RBP9	BS	1.7	1.8		TAP (1-step)
TAP-RBP10	BS	10.1	11.5		TAP (1-step)
TAP-ZC3H5	BS	10.5	9.2		TAP (1-step)
TAP-ZC3H11	BS	0.8			TAP (1-step)
TAP-ZC3H20	PC	48	49		TAP (1-step)
TAP-ZC3H21	PC	30.9	32.1		TAP (1-step)
TAP-ZC3H22	PC	16.5	12.5		TAP (1-step)
TAP-ZC3H30	PC	1.4	3.8		TAP (1-step)
TAP-ZC3H32	BS	12	71.9	11.6	TAP (1-step)
TAP-ERBP1	BS	0.3	0.27		TAP (1-step)
V5-PUF2	BS	2	22.5		V5-IP
ZFP3	PC	0.8			Specific antibody IP
TRRM1	PC	10			Specific antibody IP
RBP33	PC	4			Specific antibody IP
ZC3H39-GFP	BS	1.6			GFP-IP
ZC3H40-GFP	BS	1.3			GFP-IP

The Table shows the purified RNA-binding protein (RBP) in the first column, the life-cycle stage used in the next column, the enrichment of the mRNA in the pull-down in the next three columns, and the method used for the purification in the fifth column. N-terminal tags are placed before the name of the protein, and C-terminal tags after. BS, bloodstream form; PC, procyclic form; TAP, tandem affinity purification (Puig *et al*., 2001); IP, immunoprecipitation; GFP, green fluorescent protein. In some publications the individual results for replicates were not supplied, or only enriched mRNAs were listed. Results for RBP33 are the average of 3 replicates. References are ERBP1 (Bajak *et al*., 2020*b*), PUF3 (Kamanyi Marucha and Clayton, 2020), ZC3H20 and ZC3H21 (Liu *et al*., 2020), ZC3H11 (Droll *et al*., 2013), DRBD13 (Jha *et al*., 2015), RBP10 (Mugo and Clayton, 2017), ZC3H5 (Bajak *et al*., 2020*a*), ZC3H30 (Chakraborty and Clayton, 2018), ZC3H32 (Klein *et al*., 2017); PUF2 (Jha *et al*., 2014), UBP1 (Jha *et al*., 2014), TRRM1 (Naguleswaran *et al*., 2015), RBP33 (Fernandez-Moya *et al*., 2014), ZFP3 (Walrad *et al*., 2011), ZC3H39 and ZC3H40 (Trenaman *et al*., 2019).

will be discussed in detail (Fig. 4). The clusters are listed under 'A' in Supplementary Table S3, sheet 1, and details are in Supplementary Table S3, sheet 2. Using this number of clusters, most replicates for individual proteins clustered together. This was also true when the analysis was done with different numbers of clusters specified. The exceptions were ZC3H30 and DRBD7, which were reproducibly separated because one replicate co-purified much more RNA than the other one did (Fig. 4). We do not know the reason for this as it did not correlate with the degree to which the 'self' mRNA was purified. As previously noted (Klein *et al*., 2017), no mRNAs (apart from 'self') reproducibly co-purified with ZC3H32; realignment of the reads to a recently-published Lister 427 genome (Müller *et al*., 2018) also did not reveal any binding of ZC3H32 to *VSG* or other strain-specific mRNAs.

Since the binding of some mRNA-binding proteins is influenced by length, it is not surprising that classification of mRNAs according to their protein binding resulted in several clusters that included mostly very long mRNAs (Fig. 5A, Supplementary Table S3 sheet 2), either because of long coding regions or long 3′-UTRs (Fig. 5B and C). Not unexpectedly, clusters A1–4, which are preferentially bound by ZC3H5, ZC3H30, DRBD7 and RBP9, were also enriched in mRNAs encoding cytoskeletal proteins (Supplementary Table S3 sheet 2). However, the clusters also reflect differences between the four proteins: for example, cluster A7 includes mRNAs that are not particularly long but were reproducibly bound by DRBD7 (Fig. 5, Supplementary Fig. S6B and Table S4). ERBP1 (Bajak *et al*., 2020*b*) and PUF3 (Kamanyi Marucha and Clayton, 2020) had already been reported to bind to ribosomal protein mRNAs (clusters A28 and A44), many of which are very short (Fig. 5, Supplementary Fig. S6C and Table S4). Cluster A29 mRNAs were strongly excluded in the ERBP1 pull-down; these combine abnormally long 5′-UTRs with short coding regions.

RBP10 destabilizes procyclic-form-specific mRNAs, with a specific binding motif in their 3′-UTRs (Mugo and Clayton, 2017). There is no correlation between RBP10 mRNA binding and transcript length, but RBP10 does control expression of some protein kinases and RNA-binding proteins; these mRNAs tend to have long 3′-UTRs (Clayton, 2019) and are enriched in Cluster A6 (Fig. 5, Supplementary Table S4). The RBP10-bound cluster 36 (Supplementary Table S4 and Fig. S6D) was strongly enriched for mRNAs that encode components of the electron transport chain and are known to be suppressed by RBP10 in bloodstream forms (Mugo and Clayton, 2017).

Results for V5-PUF2 immunoprecipitation were not reproducible (correlation coefficient $R = 0.06$) (Jha *et al*., 2014), but one replicate (replicate 2) showed much stronger 'self' enrichment than the other (Table 1) and results for this replicate correlated very strongly with those for V5-UBP1 ($R = 0.96$); this indicates that the experiments need to be repeated. The results from the cluster analysis suggested that RNA binding by ZC3H30 should also be re-examined, since some mRNAs were very specifically bound in at least one of the two replicates (Supplementary Fig. S6E and F). The mRNAs encoding amino-acyl tRNA synthetases, which are concentrated mainly in clusters A25 and A26, are examples.

The results from UBP1-myc and V5-UBP1 pull-downs were relatively well correlated ($R = 0.63$ for log-transformed data). We previously reported that from a MEME analysis, the mRNAs

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

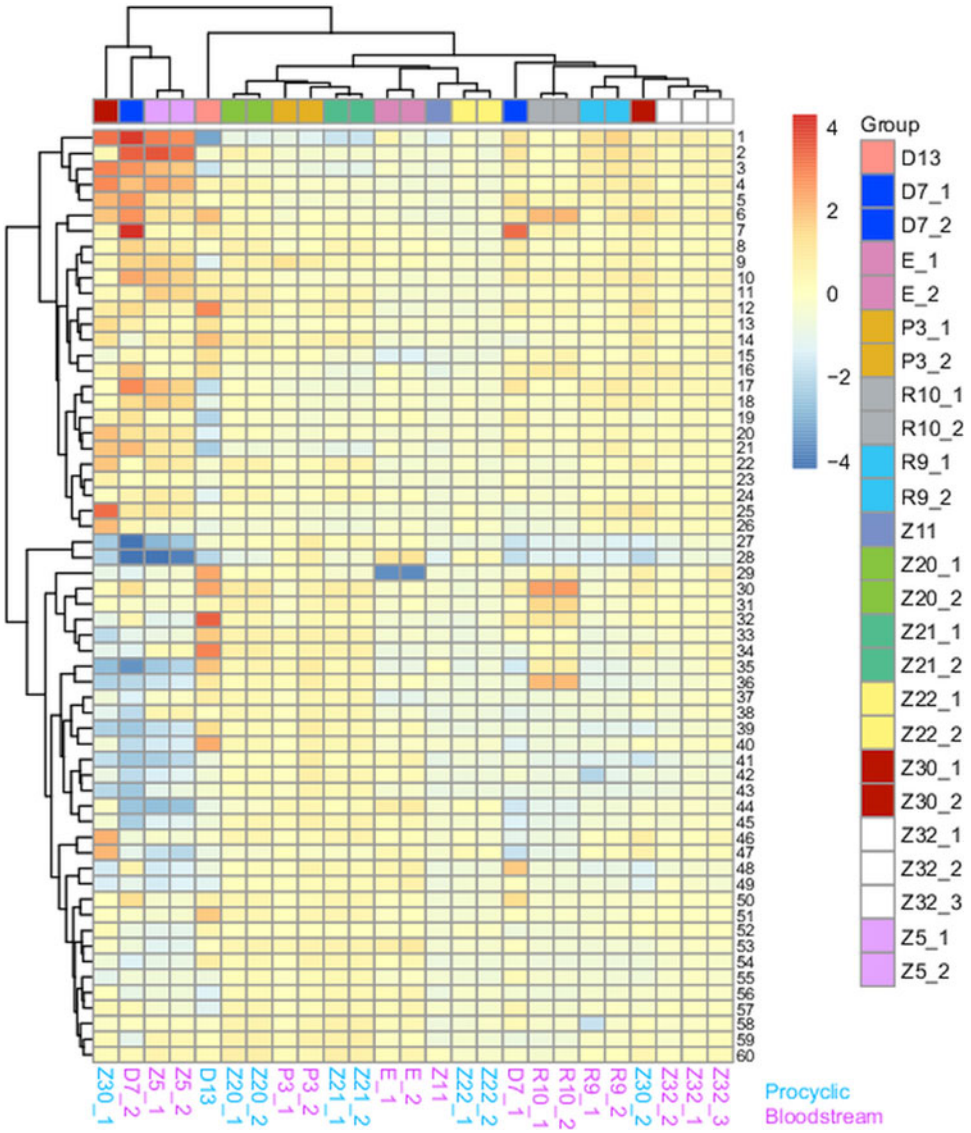

Fig. 4. Cluster analysis reveals binding similarities between different RNA-binding proteins. Binding ratios for each experiment were plotted as a heat map using ClusterViewer (Mulindwa *et al*., 2018). In this figure, 60 clusters are shown. All details are in Supplementary Table S3. The colours of the labels at the bottom indicate the life cycle stage for which RNA binding was measured.

bound to UBP1 were enriched in U-rich sequences (Jha *et al*., 2014). When, however, we compared the bound 3′-UTRs with those from a control set of unbound mRNAs, using DREME (Bailey, 2011), no enriched motif was found, presumably because U-rich sequences are generally abundant in trypanosome untranslated regions.

Clustering for a smaller group of RNA-binding proteins

The inclusion of proteins with strong binding to long mRNAs made it difficult to see lower intensity differences in a heat map. In order to detect specific binding of RBP9 and ZC3H22, we therefore compared a more limited set of RNA-binding proteins, eliminating those with strong binding or length biases. We also set the values for 'self' binding to 0, since this apparent binding is an artefact of the method. Sorting the data into 70 clusters (Supplementary Fig. S8 and 'B' in Supplementary Table S3) proved ideal to display verified specificities. For example, the mRNAs that encode chaperones, and are bound by ZC3H11, were compactly sorted into cluster B40, and the mRNAs encoding the electron transport chain and procyclic-specific membrane proteins, which are bound and repressed by RBP10, sorted into clusters B36 and B37. The overlapping specificities of ZC3H20 and ZC3H21 also became visible (clusters B58 and B61). In procyclic forms these two proteins bind to, and stabilize, some mRNAs that are repressed by RBP10 in bloodstream forms (Liu *et al*., 2020); this is easily seen in cluster B37. Unfortunately, however, this analysis did not reveal any additional specificity for RBP9 or ZC3H22.

Discussion

This paper describes attempts to analyse the functions of three mRNA-binding proteins, RBP9, DRBD7 and ZC3H22. Depletion of RBP9 and DRBD7 by RNA interference did not affect bloodstream-form trypanosome growth, and both preferentially bound long mRNAs. RBP9, which is preferentially expressed in bloodstream forms, also preferentially binds to developmentally-regulated mRNAs that are more abundant in that form, but because no RNAi effect was seen, the biological significance is unknown. It is possible that for both proteins, the protein that remained after RNAi was sufficient for function; alternatively, their functions may be redundant with those of other proteins.

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press

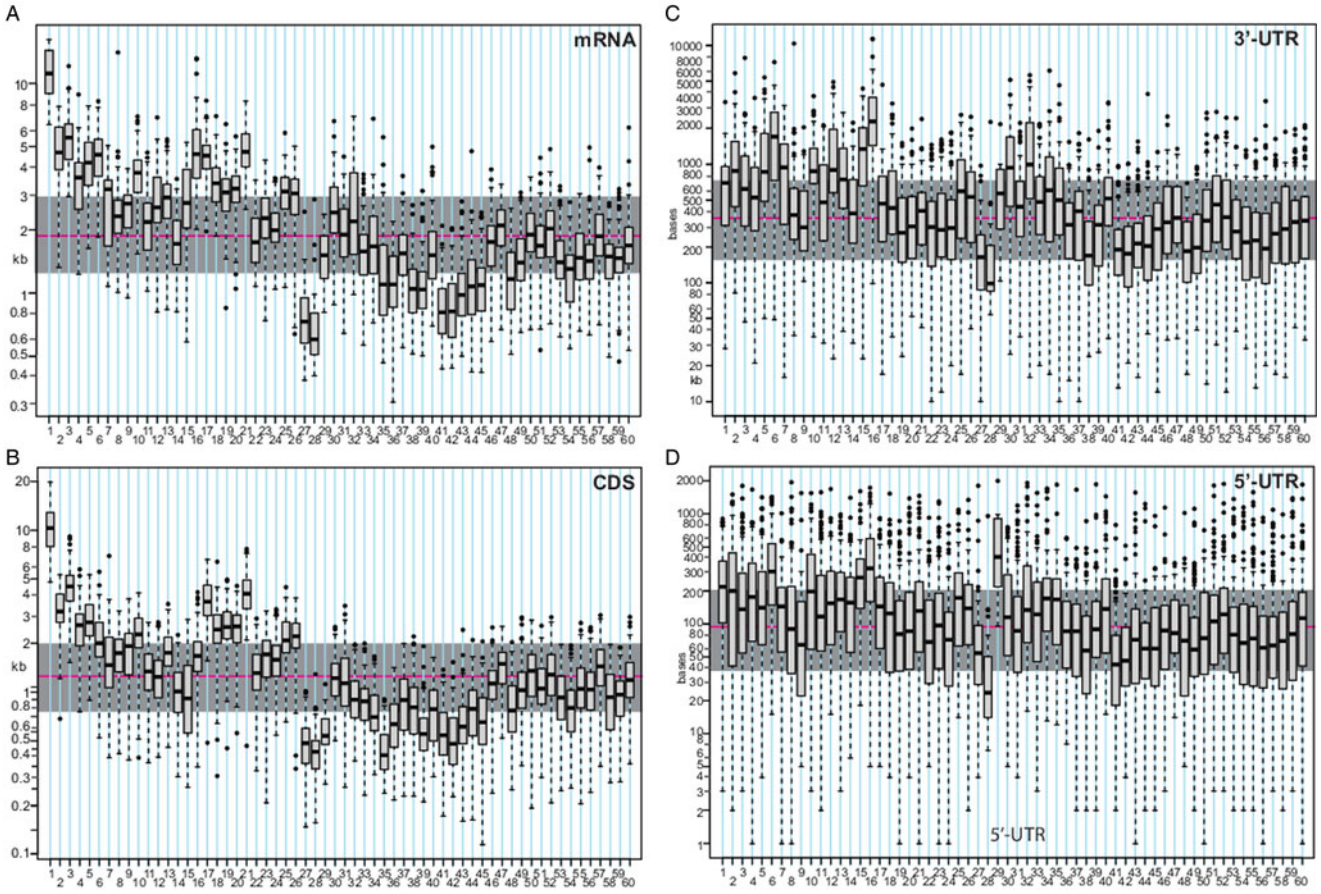

Fig. 5. Lengths of mRNAs in the different clusters are shown in Fig. 4. All details are in Supplementary Table S3 and lists of genes in each cluster are in Supplementary Table S4.

We detected only poor ZC3H22 mRNA binding, but there was some specific enrichment of mRNAs encoding ribosomal proteins and procyclic-form glucose metabolism. ZC3H22 depletion caused cells to stick together to make large longitudinally aligned clumps, with relative decreases in many mRNAs required for cell growth and division, and increases in mRNAs encoding some epimastigote-form markers. Epimastigotes also can form clusters, but the strong clumping in our cultures precluded further morphological analysis. The abundances of ribosomal protein mRNAs were not affected by the RNAi. However, there were clear decreases in several bound mRNAs encoding enzymes of procyclic energy metabolism. Since tethering of ZC3H22 to a reporter mRNA in procyclic forms had no effect on expression, we suggest that the normal role of ZC3H22 might be to antagonise a different, suppressive protein that binds to the same sequences. This would be analogous to the function of HuR and HuD, which antagonise the action of degradation-promoting AU-rich element-binding proteins in mammalian cells (Brennan and Steitz, 2001; Raineri *et al.*, 2004; Lal *et al.*, 2005). The effects of ZC3H22 depletion on other, non-bound transcripts may have been secondary.

When we compared the results for RBP9, DRBD7 and ZC3H22 with those for other proteins, some interesting patterns emerged. This was particularly the case for proteins whose binding correlated with mRNA length. Some proteins preferred mRNAs with long coding regions; and others tended to select mRNAs with long 3′-UTRs. We suggest that proteins that behave in this way bind with relatively low sequence specificity. Consequently, preferred binding sites are likely to be scattered throughout mRNAs with a frequency that is proportional to length. If the binding sites are A-rich, U-rich or repetitive, they are likely to be found more often in untranslated regions, while more G- or C-rich binding sites will be concentrated in coding regions. Consistent with this, *T. brucei* UBP1 selects mRNAs with long 3′-UTRs, and *T. cruzi* UBP1 was shown to bind *in vitro* to U-rich sequences (D'Orso and Frasch, 2001). In contrast, ZC3H5 prefers long coding regions, and a consensus sequence, (U/A)UAG(A/G), was enriched only in the coding regions and 5′-UTRs of its bound mRNAs (Bajak *et al.*, 2020*a*). These results suggest that future analyses of trypanosomatid RNA-binding protein specificity should pay close attention to the lengths of different portions of bound mRNAs, in addition to primary sequences and function.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182021000123

Acknowledgements. We thank Ute Leibfried for technical assistance.

Conflict of interest. The authors have no competing interests.

Financial support. This work was primarily supported by core funding to CC from the State of Baden-Württemberg.

Ethical standards. Not applicable.

References

Alsford S, Turner D, Obado S, Sanchez-Flores A, Glover L, Berriman M, Hertz-Fowler C and Horn D (2011) High throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research* **21**, 915–924.

Archer S, Inchaustegui D, de Queiroz R and Clayton C (2011) The cell-cycle regulated transcriptome of an early-branching eukaryote. *PloS One* **6**, e18425.

Bailey T (2011) DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics (Oxford, England)* **27**, 1653–1659.

Bajak K, Leiss K, Clayton C and Erben E (2020a) A potential role for a novel ZC3H5 complex in regulating mRNA translation in *Trypanosoma brucei*. *Journal of Biological Chemistry* **295**, 14291–14304. doi: 10.1074/jbc.RA120.014346

Bajak K, Leiss K, Clayton C and Esteban Erben E (2020b) The endoplasmic reticulum-associated mRNA-binding proteins ERBP1 and ERBP2 interact in bloodstream-form *Trypanosoma brucei*. *PeerJ* **8**, e8388.

Begolo D, Vincent I, Giordani F, Pöhner I, Witty M, Rowan T, Bengaly Z, Gillingwater K, Freund Y, Wade R, Barrett M and Clayton C (2018) The trypanocidal benzoxaborole AN7973 inhibits trypanosome mRNA processing. *PLoS Pathogens* **14**, e1007315.

Brennan CM and Steitz JA (2001) Hur and mRNA stability. *Cellular and Molecular Life Sciences* **58**, 266–277.

Chakraborty C and Clayton C (2018) Stress susceptibility in *Trypanosoma brucei* lacking the RNA-binding protein ZC3H30. *PLoS Neglected Tropical Diseases* **12**, e0006835.

Clayton C (2019) Control of gene expression in trypanosomatids: living with polycistronic transcription. *Royal Society Open Biology* **9**, 190072.

Das A, Bellofatto V, Rosenfeld J, Carrington M, Romero-Zaliz R, del Val C and Estevez AM (2015) High throughput sequencing analysis of *Trypanosoma brucei* DRBD3/PTB1-bound mRNAs. *Molecular and Biochemical Parasitology* **199**, 1–4.

Domingo-Sananes MR, Szoor B, Ferguson MA, Urbaniak MD and Matthews KR (2015) Molecular control of irreversible bistability during trypanosome developmental commitment. *Journal of Cell Biology* **211**, 455–468.

D'Orso I and Frasch ACC (2001) TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. *Journal of Biological Chemistry* **276**, 34801–34809.

Droll D, Minia I, Fadda A, Singh A, Stewart M, Queiroz R and Clayton C (2013) Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. *PLoS Pathogens* **9**, e1003286.

Erben E, Fadda A, Lueong S, Hoheisel J and Clayton C (2014) Genome-wide discovery of post-transcriptional regulators in *Trypanosoma brucei*. *PLoS Pathogens* **10**, e1004178.

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

Fernandez-Moya SM, Carrington M and Estevez AM (2014) Depletion of the RNA-binding protein RBP33 results in increased expression of silenced RNA polymerase II transcripts in *Trypanosoma brucei*. *PloS One* **9**, e107608.

Jha B, Fadda A, Merce C, Mugo E, Droll D and Clayton C (2014) Depletion of the trypanosome pumilio domain protein PUF2 causes transcriptome changes related to coding region length. *Eukaryotic Cell* **13**, 664–674.

Jha BA, Gazestani VH, Yip CW and Salavati R (2015) The DRBD13 RNA binding protein is involved in the insect-stage differentiation process of *Trypanosoma brucei*. *FEBS Letters* **589**, 1966–1974.

Jones NG, Thomas EB, Brown E, Dickens NJ, Hammarton TC and Mottram JC (2014) Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS Pathogens* **10**, e1003886.

Kamanyi Marucha K and Clayton C (2020) Roles of the Pumilio domain protein PUF3 in *Trypanosoma brucei* growth and differentiation. *Parasitology* **147**, 1171–1183.

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

Khong A and Parker R (2020) The landscape of eukaryotic mRNPs. *Rna* **26**, 229–239.

Klein C, Terrao M and Clayton C (2017) The role of the zinc finger protein ZC3H32 in bloodstream-form *Trypanosoma brucei*. *PloS One* **12**, e0177901.

Kolev NG, Ramey-Butler K, Cross GA, Ullu E and Tschudi C (2012) Developmental progression to infectivity in Trypanosoma brucei triggered by an RNA-binding protein. *Science* **338**, 1352–1353.

Lal A, Mazan-Mamczarz K, Kawai T, Yang X, Martindale JL and Gorospe M (2005) Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO Journal* **23**, 3092–3102.

Leiss K and Clayton C (2016) DESeqUI – Trypanosome RNAseq analysis made easy. *Zenodo*. doi: 10.5281/zenodo.165132.

Leiss K, Merce C, Muchunga E and Clayton C (2016) Tryprnaseq – A easy to use pipeline for *Trypanosoma* RNAseq data. *Zenodo*. http://doi.org/10.5281/zenodo.158920.

Liu B, Marucha K and Clayton C (2020) The zinc finger proteins ZC3H20 and ZC3H21 stabilise mRNAs encoding membrane proteins and mitochondrial proteins in insect-form. *Trypanosoma brucei Mol Microbiol* **113**, 430–451.

Love M, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biology* **15**, 550.

Lueong S, Merce C, Fischer B, Hoheisel J and Erben E (2016) Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Molecular Microbiology* **100**, 457–471.

Mandelboum S, Manber Z, Elroy-Stein O and Elkon R (2019) Recurrent functional misinterpretation of RNA-seq data caused by sample-specific gene length bias. *PLoS Biology* **17**, e3000481.

Miguel De Pablos L, Kelly S, Nascimento J, Sunter J and Carrington M (2017) Characterization of RBP9 and RBP10, two developmentally regulated RNA binding proteins in *Trypanosoma brucei*. *Open Biology* **7**, 160159.

Mugo E and Clayton C (2017) Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in *Trypanosoma brucei*. *PLoS Pathogens* **13**, e1006560.

Mulindwa J, Leiss K, Ibberson D, Kamanyi Marucha K, Helbig C, Melo do Nascimento L, Silvester E, Matthews K, Matovu E, Enyaru J and Clayton C (2018) Transcriptomes of *Trypanosoma brucei Rhodesiense* from sleeping sickness patients, rodents and culture: effects of strain, growth conditions and RNA preparation methods. *PLoS Neglected Tropical Diseases* **12**, e0006280.

Müller L, Cosentino R, Förstner K, Guizetti J, Wedel C, Kaplan N, Janzen C, Arampatzi P, Vogel J, Steinbiss S, Otto T, Saliba A-E, Sebra R and Siegel T (2018) Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121–125.

Naguleswaran A, Gunasekera K, Schimanski B, Heller M, Hemphill A, Ochsenreiter T and Roditi I (2015) *Trypanosoma brucei* RRM1 is a nuclear RNA-binding protein and modulator of chromatin structure. *MBio* **6**, e00114.

Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M and Seraphin B (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods (San Diego, Calif.)* **24**, 218–229.

Raineri I, Wegmueller D, Gross B, Certa U and Moroni C (2004) Roles of AUF1 isoforms, HuR and BRF1 in ARE-dependent mRNA turnover studied by RNA interference. *Nucleic Acids Research* **32**, 1279–1288.

Savage AF, Kolev NG, Franklin JB, Vigneron A, Aksoy S and Tschudi C (2016) Transcriptome profiling of *Trypanosoma brucei* Development in the tsetse fly vector *Glossina morsitans*. *PloS One* **11**, e0168877.

Shen S, Arhin GK, Ullu E and Tschudi C (2001) *In vivo* epitope tagging of *Trypanosoma brucei* genes using a one step PCR-based strategy. *Molecular and Biochemical Parasitology* **113**, 171–173.

Trenaman A, Glover L, Hutchinson S and Horn D (2019) A post-transcriptional respiratome regulon in trypanosomes. *Nucleic Acids Research* **47**, 7063–7077. doi: 10.1093/nar/gkz455

Urbaniak MD, Martin D and Ferguson MA (2013) Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *Journal of Proteome Research* **12**, 2233–2244.

Walrad P, Capewell P, Fenn K and Matthews K (2011) The post-transcriptional trans-acting regulator, TbZFP3, co-ordinates transmission-stage enriched mRNAs in *Trypanosoma brucei*. *Nucleic Acids Research* **40**, 2869–2883.

Wurst M, Robles A, Po J, Luu V, Brems S, Marentije M, Stoitsova S, Quijada L, Hoheisel J, Stewart M, Hartmann C and Clayton C (2009) An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **163**, 61–65.

https://doi.org/10.1017/S0031182021000123 Published online by Cambridge University Press