The RNA-binding protein RBP10 controls a regulatory cascade that defines bloodstream-form trypanosome identity.

Elisha Mugo and Christine Clayton
DKFZ-ZMBH Alliance, Zentrum für Molekulare Biologie der Universität Heidelberg, University of Heidelberg. Im Neuenheimer Feld 282, D69120 Heidelberg, Germany.

Abstract

Gene expression control in the pathogen *Trypanosoma brucei* relies almost exclusively on post-transcriptional mechanisms, so RNA binding proteins must assume the burden that is usually borne by transcription factors. *T. brucei* multiply in the blood of mammals as bloodstream forms, and in the midgut of Tsetse flies as procyclic forms. We show here that a single RNA-binding protein, RBP10, defines the bloodstream-form trypanosome differentiation state. Depletion of RBP10 from bloodstream-form trypanosomes gives cells that can grow only as procyclic forms; conversely, expression of RBP10 in procyclic forms converts them to bloodstream forms. RBP10 binds to procyclic-specific mRNAs containing an UAUUUUUU motif, targeting them for translation repression and destruction. Products of RBP10 target mRNAs include not only the major procyclic surface protein and enzymes of energy metabolism, but also protein kinases and stage-specific RNA-binding proteins: consequently, alterations in RBP10 trigger a regulatory cascade.

Introduction

In multicellular organisms, cell differentiation is usually unidirectional, with stabilisation of differentiated states by epigenetic mechanisms, and reversal only in when healthy gene regulation is disrupted. In contrast, differentiation in unicellular organisms can be reversible, or involve cycles in which the direction of each transition is fixed. In each case differentiation is initiated by environmental stimuli, which activate signal transduction cascades; these in turn result in post-translational modifications that modify transcription factor activity. The consequent changes in mRNA synthesis, and hence protein synthesis, trigger a cascade of downstream regulatory events, affecting additional signalling molecules, transcription factors, and post-transcriptional modulators. Differentiation states are then stabilised by modification of DNA and chromatin.

Kinetoplastids are unicellular flagellates, and include many pathogens of mammals and plants, including the leishmanias and trypanosomes that cause human disease. Their life cycles involve a series of unidirectional transitions, which are driven by environmental changes as they move between mammalian and invertebrate hosts. The kinetoplastid studied in this paper, *Trypanosoma brucei*, causes human sleeping sickness in Africa and diseases of ruminants throughout the tropics. *T. brucei* multiply as bloodstream form trypomastigotes in the blood and tissue fluids of mammals, escaping immunity by antigenic variation of Variant Surface Glycoprotein (VSG) [1].

43 High cell density triggers growth arrest and differentiation to stumpy forms [2], which,
44 upon uptake by a Tsetse fly, convert to proliferative procyclic forms in the midgut.
45 Procyclic forms express EP and GPEET procyclin surface proteins [3], and have
46 numerous metabolic adaptations including increased mitochondrial metabolism. In
47 the following 2 weeks, trypanosomes migrate towards the salivary glands [4].
48 Epimastigotes, expressing the surface protein BARP [5], are found attached to the
49 proventriculus and salivary glands; here, a subset of parasites undergoes meiosis
50 and mating [6,7]. Finally, growth-arrested, VSG-expressing metacyclic forms are
51 seen swimming freely in the salivary glands; these are now ready for infection of a
52 new mammal. Trypanosome developmental transitions are marked by many changes
53 in mRNA [8-12] and protein [13-16] abundance.

54 Despite extensive changes in gene expression through the *T. brucei* life cycle, and
55 the existence of stable differentiation states, the normal paradigm for regulation does
56 not apply. This is because - as in other Kinetoplastid protists - transcription by *T.*
57 *brucei* RNA polymerase II is polycistronic, with no regulation at the level of individual
58 mRNAs. Monocistronic mRNAs are excised by 5' *trans* splicing of a 39nt capped
59 leader sequence, and by 3' polyadenylation [17,18]. Regulated is thus exclusively at
60 the levels of mRNA processing, degradation, and translation [10-12,19-22]. The
61 sequences required for regulation of translation and mRNA decay often lie in the 3'-
62 untranslated regions (3'-UTRs) of the mRNAs. Many groups of trypanosome mRNAs
63 are coordinately regulated [23] and in a few cases, co-regulated genes have been
64 shown to share regulatory elements that are bound by specific RNA-binding proteins
65 [24,25]. RNA-binding proteins thus assume the burden that is carried by transcription
66 factors in other eukaryotes, by regulating distinct mRNA subsets [19,26]. Although an
67 RNAi machinery exists, it is not essential for cell proliferation or differentiation [27],
68 and microRNAs have not been found [28].

69 The very abundant *VSG* and procyclin mRNAs are made, exceptionally, by RNA
70 polymerase I [29], and their transcription is developmentally regulated by epigenetic
71 mechanisms [1]. Procyclin is not subject to antigenic variation, and the presence of
72 even a small amount on the bloodstream-form surface would enable the
73 development of host immunity. Suppression of procyclin expression is therefore
74 essential for trypanosome pathogenicity. However, procyclin gene transcription is
75 only 10-fold less active in bloodstream forms than in procyclic forms [30]. This means
76 that additional suppressive mechanisms are required: the mRNAs are both extremely
77 rapidly degraded and very poorly translated [31-35], and trafficking of procyclin
78 proteins to the surface is suppressed [36]. A 26nt motif within the 3'-UTR of procyclin

79 mRNAs is required for their degradation and translational suppression [33,34].
80 However, until now, the protein(s) responsible for preventing procyclin expression
81 were unknown.

82 RBP10 is a cytosolic protein with a single RNA recognition motif. Proteome analyses
83 [16] and Western blot results [37] indicate that RBP10 is expressed in proliferating
84 bloodstream forms, but not in stumpy forms or procyclics; *RBP10* mRNA also has
85 low abundance in procyclic forms and in salivary gland trypanosomes [38]. Depletion
86 of RBP10 by RNA interference (RNAi) in the bloodstream form is lethal, and forced
87 RBP10 expression in procyclics inhibits growth [37]. Microarray analyses on cells
88 with RBP10 RNAi or forced expression suggested a function for RBP10 either in
89 stabilizing bloodstream-form specific mRNAs, or destabilizing procyclic-specific
90 mRNAs, but it was unclear which effects were due to growth inhibition and the direct
91 RNA targets of RBP10 were unknown.

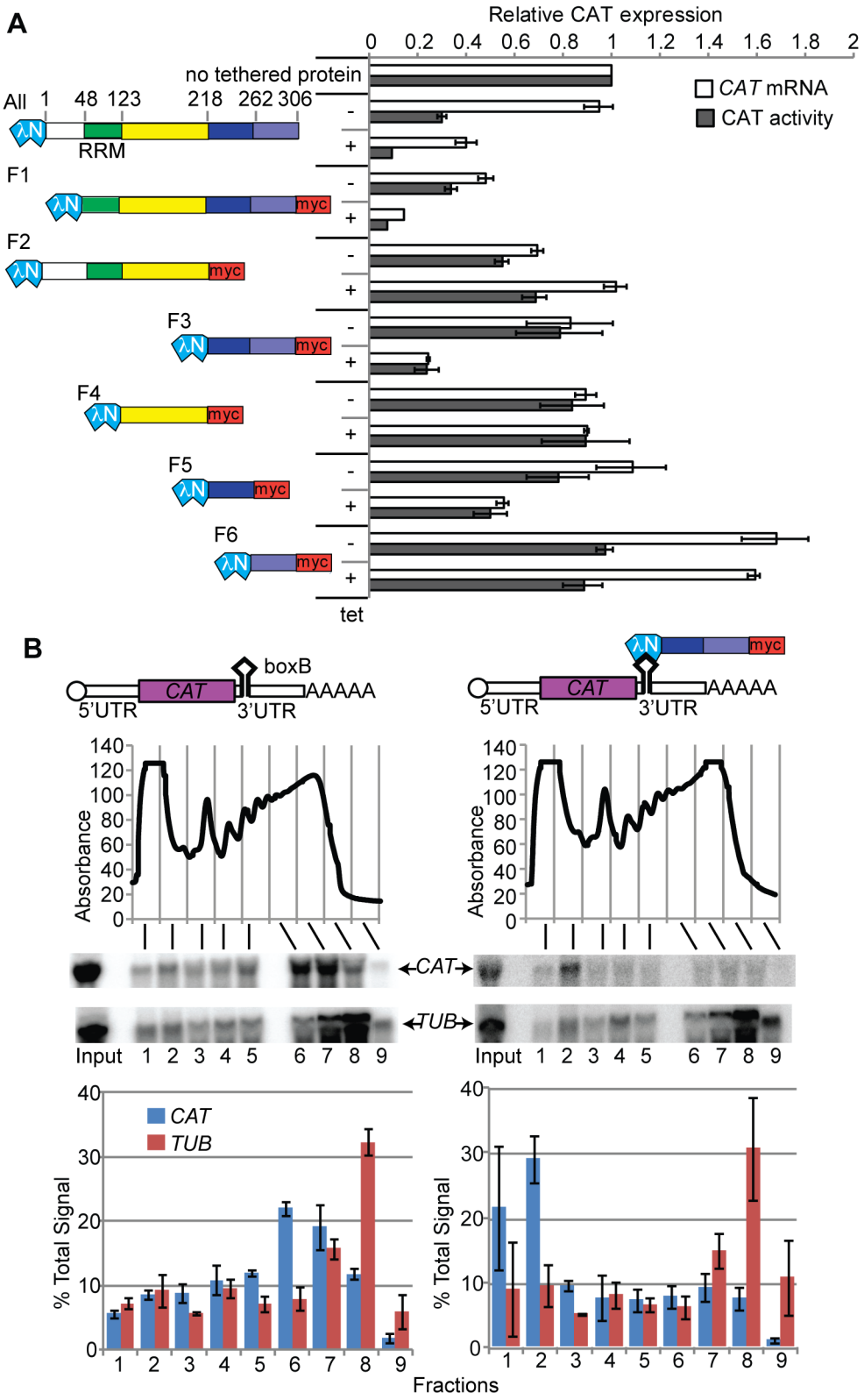
92 In this paper we set out to identify the direct mRNA targets of RBP10 and to
93 determine its role in mRNA control. We discovered that the mRNAs that are bound to
94 RBP10 in bloodstream forms are preferentially targeted for translation inhibition and
95 destruction, and that cells without RBP10 can survive only by differentiating to the
96 procyclic form. Conversely, cells expressing RBP10 can survive only as bloodstream
97 forms. The presence of RBP10 thus defines bloodstream-form trypanosome identity.

98 Results

99 Attachment of the RBP10 C-terminal domain to a reporter mRNA causes
100 translation inhibition and mRNA degradation

101 Inducible expression of lambdaN-peptide-tagged RBP10 in cells containing a *CAT-*
102 *boxB* reporter will result in attachment of the RBP10 to the reporter RNA via the
103 lambdaN-boxB interaction. We previously showed that after 24h lambdaN-RBP10
104 induction, there was a decrease in boxB-tagged chloramphenicol acetyltransferase
105 (*CAT*) reporter expression [37,39,40]. New deletion analyses (Fig 1) revealed that the
106 C-terminal 88 amino acids of RBP10 was sufficient for the repressive effect (Fig 1 A,
107 fragment F3). To investigate the mechanism in more detail we examined whether the
108 *CAT-boxB* mRNA comigrated with polysomes in sucrose gradients. In the absence of
109 tethered protein, the *CAT-boxB* reporter migrated in the dense sucrose fraction with
110 large polysomes (Fig 1 B). With lambdaN-RBP10-F3 expression, *CAT-boxB* mRNA
111 was reduced in abundance and the remainder migrated towards the top of the
112 gradient. The control mRNA (encoding alpha-tubulin) was unaffected (Fig 1 B). The

113 RBP10 C-terminus thus prevents translation initiation and promotes mRNA
114 destruction. Interestingly, although a homologue of RBP10 is present in all
115 sequenced Kinetoplastids, only the RRM domain is conserved: this C-terminal
116 sequence is specific to *Trypanosoma* (S1 Fig).



117

118 **Fig 1**

119 Tethering of the C-terminal portion of RBP10 to a reporter mRNA inhibits its
120 translation and causes its destruction.

121 A. Various fragments of RBP10 were inducibly expressed with an N-terminal
122 lambdaN peptide and a C-terminal myc tag, in a cell line expressing a *CAT* reporter
123 mRNA with 5 boxB sequences in the 3'-UTR. Cartoons showing the fragments are on
124 the left. The numbers above the full-length protein are amino acid residues. The
125 effects on the amounts *CAT* mRNA and CAT protein were measured by Northern blot
126 and enzyme assay, respectively. Levels are expressed as mean ± standard deviation
127 for at least 3 replicates, relative to a line with no lambdaN peptide protein (top bars)

128 B. Bloodstream-form trypanosomes expressing the reporter in (A), and with
129 tetracycline-inducible expression of lamdaN-RBP10-F3, were used. The reporter with
130 or without tethered protein is shown at the top. Lysates from cells grown without (left)
131 or with (right) tetracycline were subjected to sucrose gradient centrifugation. The
132 panel below the cartoons show the absorption profiles at 254 nm as the fractions
133 were collected. The migration of *CAT* and beta-tubulin (*TUB*) mRNAs on the gradient
134 was detected by Northern blotting; representative blots are shown. The graphs show
135 Northern signals, expressed as the percentage of the total signal, with arithmetic
136 mean and standard deviation for three independent biological replicates.

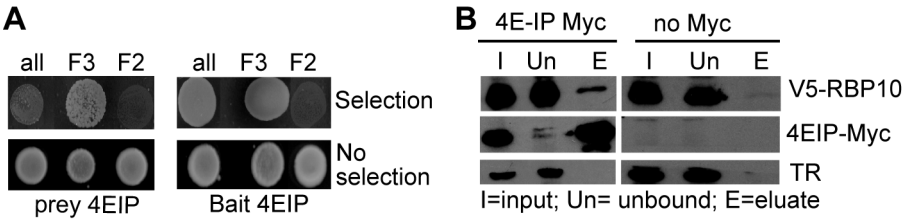
137

138 **RBP10 can interact with an eIF4E interacting protein**

139 We investigated the mechanism of RBP10-mediated repression by looking for
140 interaction partners. We made a bloodstream-form cell line containing a single copy
141 of RBP10 bearing a cleavable N-terminal affinity purification tag (S2 Fig A). After
142 tandem affinity purification the most consistent co-purifying proteins were
143 components of the proteasome (S1 Table). While this could be a mechanism for
144 repression, it is also possible that some of the affinity-tagged RBP10 was targeted to
145 the proteasome because it was incorrectly folded. The only interaction partner linked
146 to mRNA degradation was CAF40, a component of the NOT deadenylase complex,
147 but since no other components of the complex were present the significance of the
148 interaction was unclear. Furthermore, V5-CAF40 pull-down failed to identify any
149 RBP10 peptides (data not shown). To find out whether other proteins - including
150 those of unknown function - might be important we investigated direct interactions of
151 RBP10 by 2-hybrid screening. Nearly 80 potential partners were seen (S2 Table), of

152 which 40 had not been seen in other screens. This suggested that RBP10 is prone to
153 non-specific interactions. Potential partners included RBP10 itself, and four other
154 RNA binding proteins. Interestingly, 8 of the 33 specifically interacting proteins had
155 previously been show to suppress expression in the tethering assay [39,40]. There
156 was however no overlap between the two-hybrid results and those from affinity
157 purification.

158 A single fragment of a trypanosome eIF4E-interacting protein (4EIP) was scored as a
159 possible RBP10 partner in the two-hybrid screen. Leishmania 4EIP interacts with
160 eIF4E1 [41], which is one of the six trypanosomatid homologues of the cap-binding
161 translation initiation factor eIF4E. In *T. brucei*, both 4EIP and eIF4E1 are extremely
162 strong suppressors of reporter expression in the tethering assay [39,40]. We
163 confirmed the two-hybrid interaction between full length RBP10 and 4EIP (Fig 2 A);
164 and found that the interaction was mediated by the C-terminal 88 amino acids (F3)
165 that are active in the tethering assay. This would be consistent with RBP10 acting via
166 4EIP recruitment. To test the interaction *in vivo*, we used trypanosomes expressing
167 RBP10 that was N-terminally V5-tagged at the endogenous locus, then expressed C-
168 terminally myc-tagged 4EIP from a strong promoter. Immunoprecipitation of myc-
169 tagged 4EIP indeed pulled down a very small proportion of V5-RBP10 (Fig 2 B). It is
170 therefore possible that RBP10 acts by recruiting 4EIP. However, depletion of 4EIP
171 does not reproducibly inhibit bloodstream-form trypanosome growth ([42] and our
172 preliminary results), which argues against this hypothesis.

173 

174 **Fig 2**

175 Interaction of RBP10 with 4E-IP

176 A. Two-hybrid assay showing yeast growth. Upper panels show selection against
177 non-interacting pairs.

178 B. Immunoprecipitation of 4E-IP-Myc pulls down V5-tagged RBP10. The number of
179 cell equivalents loaded for the eluate was 33x that of input and unbound fractions.
180 The inverse precipitation did not confirm the interaction, but this could be because of
181 an imbalance of protein abundances in the cells. (The abundance of 4E-IP is not
182 known).

183

184 **RBP10 represses procyclic-form specific mRNAs**

185 To analyse the role of RBP10 in gene expression in bloodstream forms, we depleted
186 RBP10 by RNAi for 15h (S2 Fig B), before parasite proliferation and translation (S3
187 Fig) were affected. High-throughput cDNA sequencing (RNA-Seq) revealed nearly
188 200 mRNAs that changed significantly in total and/or polysomal RNA fractions (S3
189 Table, sheets 3-7). 36% of the mRNAs that increased after RNAi were preferentially
190 expressed in procyclic forms, while 44% of those that decreased were preferentially
191 expressed in bloodstream forms (Table 1 and Fig 3 A). Increases in polysomal RNA
192 after RNAi were very slightly stronger than for total RNA. *RBP10* was the only mRNA
193 with strongly decreased polysomal loading after RNAi, whereas about 60 mRNAs
194 moved into polysomes (S3 Table sheet 3). 19 of these 60 mRNAs have a higher
195 ribosome density in procyclic forms than in bloodstream forms [11], suggesting that
196 their translation is inhibited in bloodstream forms. An example is the *COXVI* mRNA
197 (S3 Table sheet 3, S4 Fig A,B), which is known to be developmentally regulated
198 mainly at the level of translation [43]. Only 4 mRNAs had lower polysome occupancy
199 after RNAi (S1 Table sheet 3). Our results almost certainly under-estimate the extent
200 of translation changes, since we measured only whether an mRNA was in the
201 polysomal fraction or not. For example, if an mRNA had a few ribosomes before
202 RNAi, and more ribosomes after RNAi, this would not have been detected.

203 **Fig 3**

204 mRNAs bound to and/or regulated by RBP10.

205 Results for mRNA levels are displayed as proportional Venn diagrams made using
206 http://jura.wi.mit.edu/bioc/tools/venn.php and
207 http://www.eulerdiagrams.org/eulerAPE/

208 BS = Bloodstream form, PC = procyclic form. Categories were: Bound:
209 Supplementary Table S3 sheet 4. PC>BS and BS>PC: at least 2x significantly
210 regulated [10]. BS rbp10 down or up: affected by RNAi, Supplementary Table S1
211 sheets 6 & 7; PC+RBP10 down or up: affected by RBP10 expression in procyclics,
212 Supplementary Table S2 sheets 6 & 7. Fischer exact test values, calculated for the
213 dataset in the centre, are in blue italics.

214 A) Relationship between developmental regulation and the effects of RBP10
215 depletion in bloodstream forms

216 B) Relationship between developmental regulation and the effects of RBP10
217 expression in procyclic forms

218 C) Expression of RBP10 in procyclic forms related to the effects *rbp10* RNAi in
219 bloodstream forms.

220 D) Developmental regulation of the mRNAs bound by RBP10.

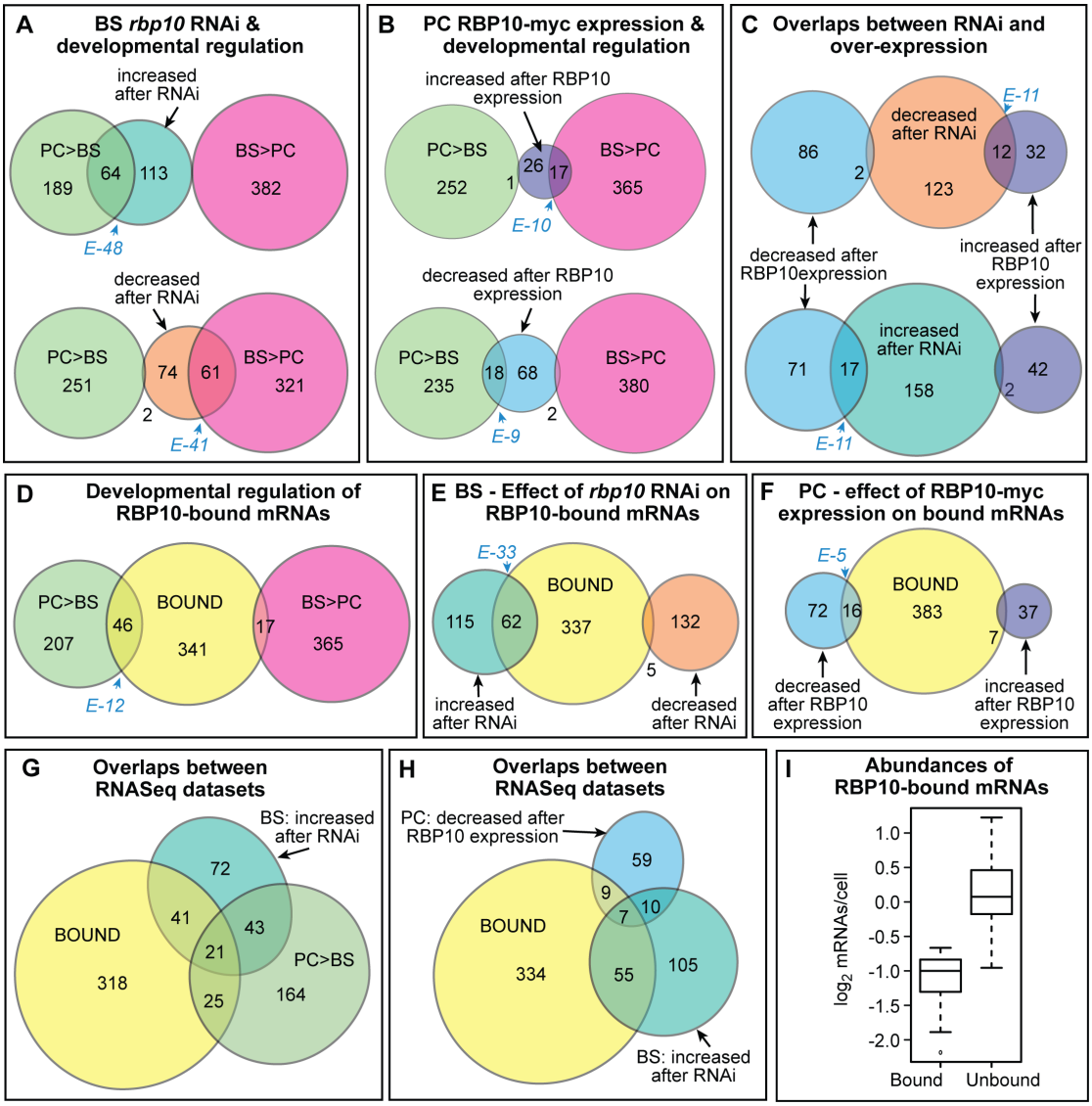
221 E) Effect of bloodstream-form *rbp10* RNAi on the mRNAs bound by RBP10.

222 D) Effect of procyclic-form RBP10 expression on the mRNAs bound by RBP10.

223 G) Overlaps between bound mRNAs, procyclic-specific mRNAs, and the mRNAs that
224 increased after *rbp10* RNAi in bloodstream forms

225 H) Overlaps between bound mRNAs, the mRNAs that increased after *rbp10* RNAi in
226 bloodstream forms, and the mRNAs that decreased after RBP10 expression in
227 procyclic forms

228 I) Abundances of RBP10-bound mRNAs compared to all other mRNAs.



Next we analysed the effect of expressing C-terminally myc-tagged RBP10 (RBP10-myc) in procyclic forms. After only 6h induction, (S2 Fig C, S4 Fig) we saw decreases in procyclic-form-specific mRNAs, and increases in bloodstream-form specific mRNAs (Fig 3 B and S4 Table). A subset of mRNAs showed opposite reactions in the bloodstream RNAi and procyclic over-expression experiments, consistent with being immediate RBP10 targets (Fig 3 C).

We had previously attempted to identify RBP10-bound mRNAs, but the methods that we used had insufficient sensitivity [37]. Since a subsequent study showed that RBP10 is indeed bound to mRNA [40], we made another attempt. This time, we used our bloodstream-form cell line in which all RBP10 bears a cleavable tag (S2 Fig A), and compared bound and unbound mRNAs by RNASeq. 400 mRNAs were identified (S5 Fig A and S5 Table, sheet 3). 12% of them are more abundant in procyclic forms

243 than in bloodstream forms (Fig 3 D). Consistent with the tethering results, 15% of the
244 RBP10 targets increased after *rbp10* RNAi in bloodstream forms, but only 1%
245 decreased (Fig 3 E). Only 16 target mRNAs decreased after RBP10 expression in
246 procyclic forms (Fig 3 F); this low number might be due to the very short duration of
247 RBP10-myc expression. The RBP10-bound mRNAs that are more abundant in
248 procyclic forms and increased after *rbp10* RNAi (Fig 3 G) are listed in Table 1: six of
249 them are also decreased after RBP10 expression in procyclics (Fig 3 H). Scatter
250 plots showing the effects of RNAi on bound RNA abundance and polysomal loading
251 are shown in S5 Fig C and D.

252

253

254 **Table 1**

255 21 mRNAs that are bound to RBP10, are more abundant (RNA PC>BS) or better
256 translated in procyclic forms than in bloodstream forms according to ribosome
257 profiling (Riboprof PC>BS), and are increased after *rbp10* RNAi in bloodstream
258 forms. For details see S5 Table sheet 8. The latter includes many additional mRNAs
259 that fulfilled fewer conditions.

id	Annotation	RNA PC> BS	RiboProf PC>BS	de-creased by RBP10 in PC
Metabolism				
9.5900	glutamate dehydrogenase (GDH)	yes	yes	no
10.2350	pyruvate dehydrogenase complex	yes	yes	no
10.4280	Complex III subunit	yes	yes	no
1.4100	cytochrome oxidase subunit IV (COXIV)	yes	yes	no
11.15550	NADH-cytochrome b5 reductase,	yes	yes	no
4.4990	ubiquinol-cytochrome c reductase	yes	yes	no
7.210	proline dehydrogenase	yes	yes	no
9.4310	tricarboxylate carrier	yes	yes	no
9.7470	purine nucleoside transporter NT10	yes	yes	yes
11.6280	pyruvate phosphate dikinase (PPDK)	yes	yes	no
11.8990	cation transporter, putative	yes	yes	no
Surface				
10.10250	Procyclins #	yes	yes	yes
7.6850	trans-sialidase (TbTS)	yes	yes	yes
8.7340	trans-sialidase	no	yes	yes
4.3500	Amastin-like protein	yes	yes	no
RNA binding				
7.2660	ZC3H20	yes	no	no
7.2680	ZC3H22	yes	yes	yes
Unknown function				
9.13200	hypothetical protein	yes	yes	yes
11.16300	hypothetical protein	yes	yes	no
7.5550	hypothetical protein	yes	yes	no
10.3300	DUF4206 (cysteine-rich)	yes	yes	no
10.2360	hypothetical protein	yes	no	no

260

261

262 Overall, the RBP10-bound mRNAs are significantly less abundant than unbound
263 mRNAs (Fig 3 I). At the proteome level [16], the products of RBP10-bound mRNAs
264 are enriched in proteins that increase within 2-24h of the initiation of stumpy form
265 differentiation, or early during conversion to procyclic forms. There is no enrichment
266 for proteins that appear later in differentiation, or only in mature procyclics, and
267 unregulated proteins or those that are decreased in procyclic forms are under-
268 represented (S5 Table, sheet 6).

269 **RBP10-bound mRNAs share a UAUUUUUU motif**

270 A comparison of the 3'-untranslated region (3'-UTR) sequences of the most reliably
271 RBP10-bound mRNAs with those of unbound mRNAs (S5 Table, sheet 5 and S1-S3
272 texts) revealed that the motif $UA(U)_6$ is highly enriched in the 3'-UTRs of RBP10
273 targets (Fig 4 A). It is present 1-9 times in 225 out of the 255 most strongly bound 3'-
274 UTRs (Fig 4 B, S1 text). Transcriptome-wide, the number of $UA(U)_6$ motifs was
275 strongly associated with RBP10 binding (Fig 4 C). Nevertheless, many mRNAs that
276 contain $UA(U)_6$ were not pulled down with RBP10. Although there are some
277 exceptions [44,45], RRM domains usually bind to single-stranded RNA [46-48]; the
278 unbound motifs may have inappropriate conformations, or they might instead be
279 bound by competing proteins.

280 **Fig 4**

281 RBP10 targets a UAUUUUUU motif

282 A) The 3'-UTRs of 188 mRNAs that were at least 3x enriched in both RBP10-bound
283 samples (Supplementary Table S5 sheet 5) were compared with those of mRNAs
284 that did not bind (<0.7x enrichment in both experiments) using DREME. Only 3'-
285 UTRs annotated in tritrypDB were used (S1 text, S2 text, S3 text). The best-scoring
286 motif found is shown.

287 B) Numbers of $UA(U)_6$ motifs in 255 manually annotated bound 3'-UTRs (S1 text, S5
288 Table sheet 5). One mRNA each had 7, 8 and 9 motifs. The numbers of motifs
289 shown here are sometimes higher than with automatically annotated 3'-UTRs from
290 the genome database, because some of the automatic 3'-UTRs are truncated.

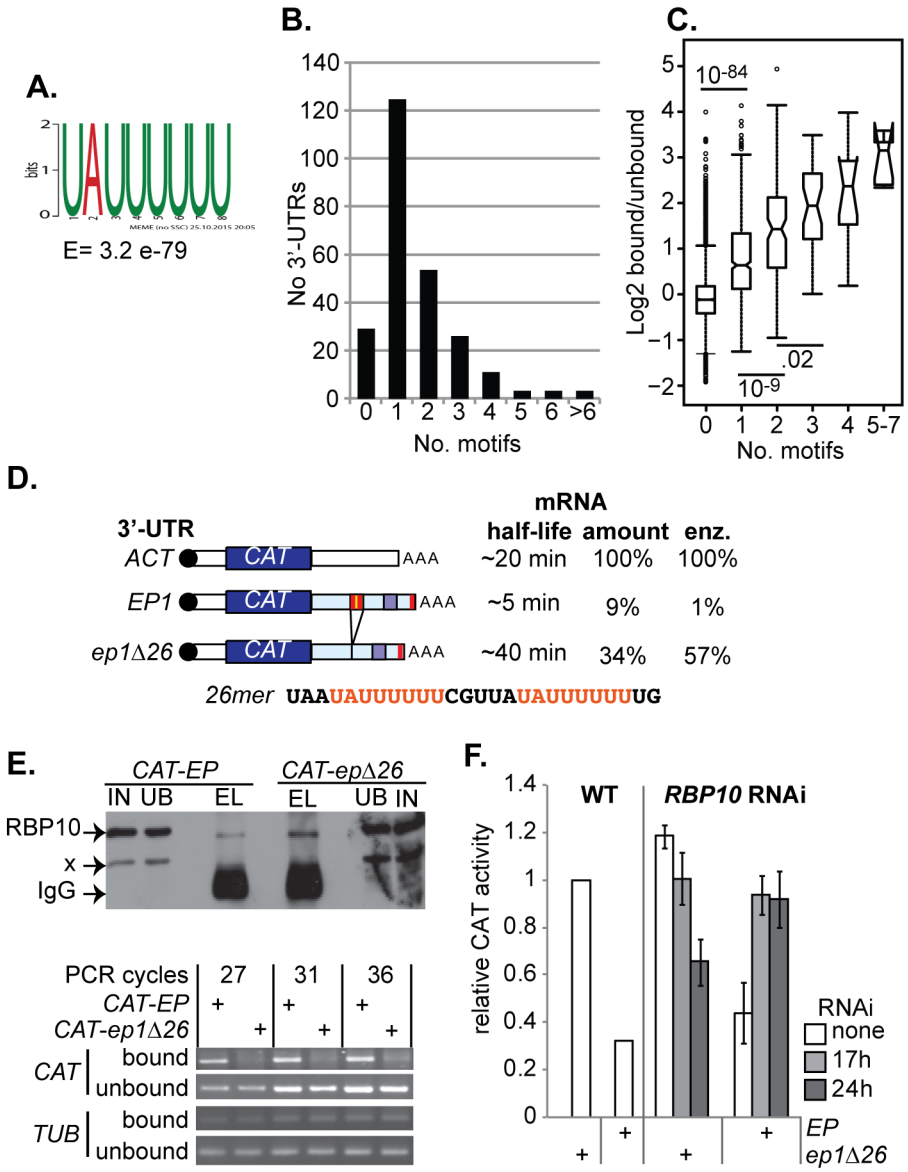
291 C) Effect of the $UA(U)_6$ motif on RBP10 binding: median ± 25th percentiles, with 95%
292 confidence limits and outliers. Database 3'-UTRs were analyzed. Since these are
293 often truncated, and are missing for 30% of genes, the numbers of motifs are under-
294 estimated. Significant differences (Student t-test) are shown, For the mRNAs that

295 were enriched according to DeSeq (S5 Table, sheet 3), 284/353 annotated UTRs
296 contained UA$(U)_6$.

297 D) Reporters used in (E) and (F) [34]. The red bars are UAUUUUUU elements, which
298 are highlighted in the 26mer sequence.

299 E) RBP10 binding to the *EP* 3'-UTR requires the 26mer. Cells were UV-irradiated,
300 and TAP-RBP10 was pulled down with IgG beads then released with TEV protease
301 (upper panel, Western blot). RNAs were detected by reverse transcription and PCR
302 using gene -specific primers (lower panel).

303 F) *rbp10* RNAi specifically affects expression of a reporter containing the 26mer. CAT
304 activity was measured 24h after RNAi induction.



305

306

307 UA(U)$_6$ was previously identified in several mRNAs encoding procyclic-specific
308 proteins [43]. The 26mer 3'-UTR element that is required for *EP1* procyclin mRNA
309 instability [33,34] contains two UA(U)$_6$ motifs (Fig 4 B) and these are single-stranded
310 *in vivo* [49], as is required for RRM domain binding. Deletion of the region containing
311 UA(U)$_6$ from the COXV 3'-UTR increased protein expression without much effect on
312 the mRNA level [43], consistent with translation control. The mRNAs encoding
313 phosphoglycerate kinase (PGK) constitute a third example. In bloodstream-form
314 trypanosomes, most PGK is encoded by the gene *PGKC* and is targeted to
315 microbodies called glycosomes (S6 Fig A); in procyclic forms, a different gene,
316 *PGKB*, is expressed and the enzyme is in the cytosol [50,51] (S6 Fig C). The *PGKB*
317 *and PGKC* coding regions are identical except for a C-terminal targeting sequence in
318 *PGKC*, so they were not distinguished in our initial RNASeq analysis. However, the
319 3'-UTRs are completely different, and determine developmental regulation [52]. The
320 UA(U)$_6$ motif is present twice within a 3-UTR segment that is required for suppression
321 of PGKB expression in bloodstream forms [53] (S1 text). Using our RNASeq pull-
322 down data, we compared the abundances of several 15nt sequences specific to the
323 *PGKB* and *PGKC* mRNAs and found that that *PGKB* mRNA is almost exclusively in
324 the bound fraction, whereas most *PGKC* mRNA is unbound. A similar analysis gave
325 RBP10 binding ratios of 2.8:1 for *EP1* and 1.3:1 for *GPEET* mRNAs. The *GPEET*
326 mRNA 3'-UTR has UA(U)$_5$ repeats, but no UA(U)$_6$, and it is not yet clear whether it is
327 bound by RBP10.

328 To test whether the 26mer was required for binding of the *EP1* 3'-UTR to RBP10, we
329 used cell lines that expressed chloramphenicol acetyltransferase (CAT) reporter
330 mRNAs with either the intact wild-type *EP1* 3'-UTR, or a version without the 26mer
331 (*ep1Δ26*) (Fig 4 D) [34]. Co-immunoprecipitation of the *CAT* mRNA with RBP10
332 depended on the presence of the 26mer (Fig 4 E). Depletion of RBP10 by RNAi also
333 increased expression from the *CAT-EP* reporter but not from *CAT-ep1Δ26* (Fig 4 F).
334 These results demonstrated both that RBP10 binds to the *EP* 3'-UTR via the 26mer
335 sequence, and that the 26mer is necessary for RBP10-mediated regulation of an
336 mRNA containing the *EP1* 3'-UTR.

337 **Expression of RBP10 determines bloodstream form identity**

338 To test the role of RBP10 in differentiation, we first examined conversion from the
339 bloodstream form to the procyclic form. Our monomorphic bloodstream forms with
340 *rbp10* RNAi, or treated with the differentiation-inducer *cis* aconitate, are unable to
341 multiply under procyclic culture conditions. We therefore used differentiation-

342 competent trypanosomes (EATRO 1125, Antat1.1 strain). As expected, growth of
343 these cells at high density in methyl cellulose [3] for three days resulted in expression
344 of the a -form marker PAD1 [54]. RBP10 was also reduced (Fig 5 A, lanes 2 &3),
345 consistent with its absence in the stumpy-form proteome [54]. The cells were now
346 treated with 6mM cis-aconitate in bloodstream-form medium at 27°C for 17h. Upon
347 transfer to procyclic medium without cis-aconitate, the cells started to grow as
348 procyclics within 24h (Fig 5B).

349

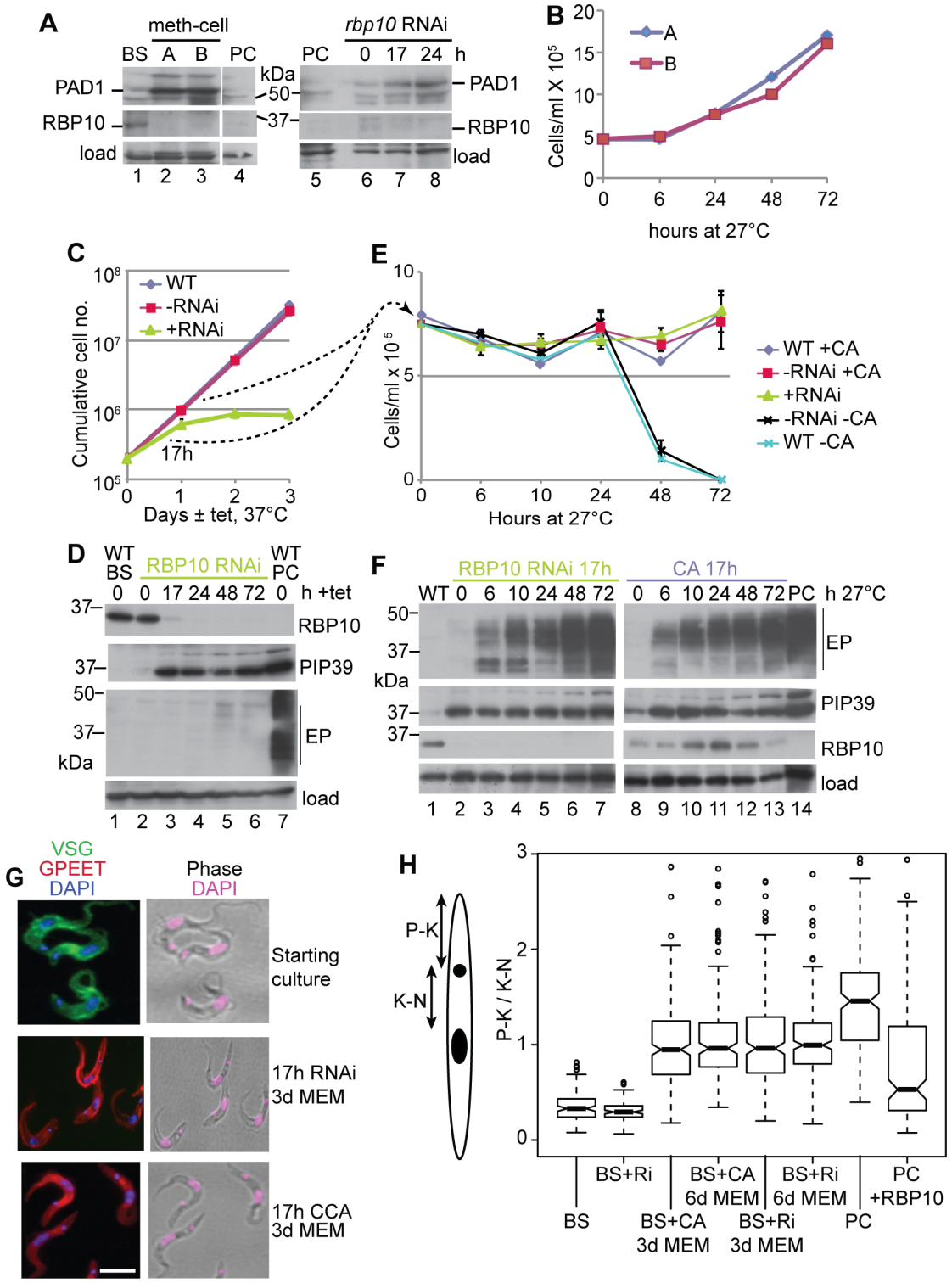


Fig 5

RBP10 depletion primes bloodstream forms for differentiation to procyclic forms

A) Expression of PAD1 and RBP10 in trypanosomes incubated at maximal density in methyl cellulose-containing medium for 3 days (left panel), or after *rbp10* RNAi (right

355 panel). The RBP10 signal is weak, probably because the samples were not
356 denatured in order to allow PAD1 detection.

357 B) Growth of trypanosomes incubated at maximal density in methyl cellulose-
358 containing medium for 3 days, followed by cis-aconitate treatment for 17h at 27°C.
359 Time 0 is the time of transfer to procyclic conditions.

360 C) Cumulative growth curve of bloodstream-form (BS) trypanosomes with and
361 without RNAi. WT = wild-type (*tet* repressor only); -RNAi: RNAi cell line with no
362 tetracyline: +RNAi: RNAi cell line with tetracycline.

363 D) Expression of RBP10, PIP39, and EP procyclin were measured by Western
364 blotting up to 3d after tetracycline induction of RNAi.

365 E) 17h after RNAi induction at 37°C in bloodstream form medium (+RNAi), or
366 incubation at high density in procylic medium with 6mM cis-aconitate at 27°C (+CA),
367 trypanosomes were placed in procyclic medium (MEM-pros) at 27°C with neither
368 tetracycline nor cis-aconitate. The cell densities of the cultures were monitored.
369 Negative controls were either wild-type or uninduced RNAi lines, without cis-
370 aconitate. Cell densities are shown; they started to increase at 72h.

371 F) Protein expression after transfer to procyclic medium at 27°C after 17h pre-
372 treatment as in (E).

373 G) Immunofluorescence: Cell morphology and expression of GPEET procyclin after
374 3d cultivation in procyclic medium. DNA is stained with DAPI.

375 H) The cartoon shows the way in which the position of the kinetoplast, relative to the
376 nucleus (P-K/K-N), was measured. The box plot shows the median with 25th
377 percentiles and 95% confidence limits. Some outliers have been deleted. The full
378 results, together with the numbers of trypanosomes measured, are in S7 Fig A.

379 BS - normal bloodstream forms; PC- normal procyclic forms; CA - 17h cis-aconitate
380 pretreatment; RNAi or Ri; 17h *rbp10* induction; +RBP10 - induced expression or
381 RBP10 for 2 days. There were no significant differences between the 3-day and 6-
382 day differentiating populations. All other pairs were significantly different (Student t-
383 test, $p<0.05$).

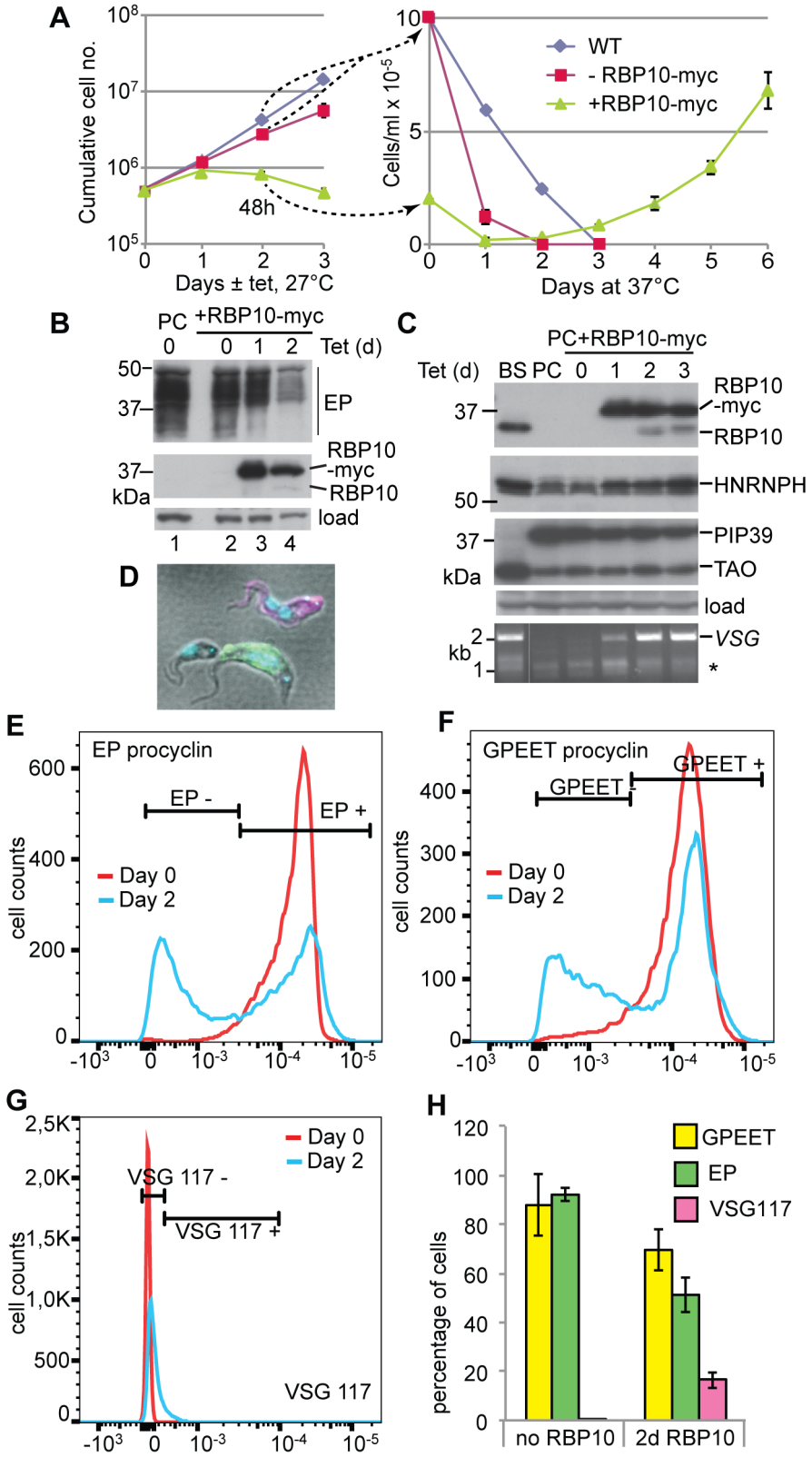
384 Induction of *rbp10* RNAi in differentiation-competent cells in log phase resulted in
385 growth inhibition (Fig 5 C), weak PAD1 expression (Fig 5 A, lanes 6-8) and, within
386 17h, expression of the differentiation-regulating phosphatase PIP39 [55] (Fig 5 D).
387 After 17h RNAi, transfer of the cells to procyclic-form medium at 27°C (without cis-
388 aconitate) allowed them to survive and eventually start to grow (Fig 5 E), with

389 expression of EP procyclin (Fig 5 F, lanes 1-7) and GPEET procyclin (Fig 5 G)
390 proteins. As expected, cells without RNAi induction died (Fig 5 E). The RNAi
391 experiment had to be done without high density pre-incubation in methyl cellulose
392 since RNAi can only be induced in growing trypanosomes. As a control, we therefore
393 grew differentiation-competent cells to maximum density in liquid medium, and
394 treated them with cis-aconitate for 17h at 27°C [56] before transfer to procyclic
395 medium without cis-aconitate, also at 27°C. In these cells, weak RBP10 expression
396 persisted for 2 days (Fig 5 F, lanes 8-13) but the growth and procyclin expression
397 kinetics were similar to those seen after *rbp10* RNAi (Fig 5 E, F, G).

398 The rather slow resumption of growth after both *rbp10* RNAi and 17h cis-aconitate
399 treatment, and the persistence of RBP10 in the cis-aconitate experiment, suggested
400 that the population might be a mixture of differentiation-competent and incompetent
401 cells. Indeed, both differentiating cultures were mixtures of normal-looking,
402 proliferating, and dying or abnormal forms, and even after 3 days in procyclic culture,
403 about 9% of the cells in both cultures retained bloodstream-form morphology and
404 PGK localization (S6 Fig D). In bloodstream forms, the kinetoplast is at the posterior
405 end of the cell (Fig 5 G), which means that the distance between the posterior and
406 the kinetoplast (P-K) is much less than the distance from the kinetoplast to the
407 nucleus (K-N) (Fig 5 H and S7 Fig A, B; "BS"). Procyclic forms are longer and the
408 kinetoplast is nearer to the nucleus (Fig 5 G), giving a much higher P-K/K-N ratio (Fig
409 5 H and S7 Fig A, B; "PC"). The differentiating cultures showed very heterogeneous
410 P-K/K-N ratios, with medians that were intermediate between those of bloodstream
411 and procyclic forms (Fig 5 H, S7 Fig A-C). As expected, cells with low ratios typical of
412 bloodstream forms were preferentially GPEET negative (S7 Fig D,E).

413 We next inducibly expressed RBP10-myc in differentiation-competent procyclic
414 forms. For the results shown, we made bloodstream forms with inducible RBP10-
415 myc, differentiated them into procyclic forms, grew these for more than 3 months,
416 and then induced RBP10 expression. However similar results were obtained by
417 differentiating bloodstream forms into procyclic forms, and then making three new
418 cell lines with the inducible RBP10-myc plasmid. After RNAi induction, cell growth
419 was inhibited, as expected [37] (Fig 6 A). After 48h, expression of procyclin was
420 strongly reduced, alternative oxidase increased, and untagged RBP10 was detected
421 (Fig 6 B, C). At this point the cell population was very heterogeneous. Some cells
422 had normal procyclic morphology and expressed EP and GPEET procyclin; some
423 were clearly abnormal; while another subset had clear bloodstream-form morphology
424 (Fig 6D). This was reflected in the P-K/K-N ratios, which ranged from procyclic to

425 bloodstream-form patterns (Fig 5 G, S7 Fig F). Flow cytometry analysis revealed that
426 about 50% of cells had decreased, or absent, procyclin expression (Fig 6 E, F). Cells
427 with bloodstream-form morphology were usually (but not always) procyclin negative
428 (S7 Fig G). Notably, *VSG* mRNA was present (Fig 6 C). We do not know which VSGs
429 are expressed in the bloodstream-form-morphology cells, but a subset of them
430 stained with a polyclonal antibody against VSG117 (Fig 6 D, G, H). PGK staining in
431 the bloodstream-form-like cells seemed reduced and somewhat more punctate than
432 in procyclic forms, but had not yet assumed a purely glycosomal pattern (S6 Fig E).
433 Remarkably, the RBP10-induced cells could now survive only if they were transferred
434 to bloodstream-form growth conditions (Figure 6A). After 6 days in these conditions,
435 all surviving cells had bloodstream-form morphology and 60% of them cross-reacted
436 with the anti-VSG117 antibody.


A
Cumulative cell no.
Days ± tet, 27°C
48h
Cells/ml x 10-5
Days at 37°C
WT
- RBP10-myc
+RBP10-myc
B
PC
+RBP10-myc
Tet (d)
EP
RBP10 -myc
RBP10
load
kDa
C
PC+RBP10-myc
Tet (d)
BS PC
RBP10 -myc
RBP10
HNRNPH
PIP39
TAO
load
VSG
kb
D
E
EP procyclin
EP -
EP +
Day 0
Day 2
cell counts
F
GPEET procyclin
GPEET
GPEET +
G
VSG 117 -
VSG 117 +
VSG 117
H
GPEET
EP
VSG117
percentage of cells
no RBP10
2d RBP10

437

438 **Fig 6**

439 RBP10 expression in procyclic forms causes differentiation to bloodstream forms

440 A) Cell counts in a typical experiment. Expression of RBP10-myc was induced using
441 tetracycline at 27°C in procyclic-form medium. The left-hand panel is a cumulative
442 growth curve. After 48h cells were transferred to bloodstream-form medium at 37°C;
443 in the right-hand panel cell densities are shown.

444 B) Expression of EP procyclin and RBP10 after induction of RBP10-myc expression.
445 Cells were cultured in procyclic medium at 27°C. Expression of EP Procyclin protein
446 decreased after 48h of RBP10-myc induction, while untagged RBP10 becomes
447 detectable.

448 C) As in (B) except that culture continued for 4 days. Proteins that are expressed
449 more in bloodstream forms than in procyclic forms were detected by Western
450 blotting. TAO is trypanosome alternative oxidase; HNRNPH is an RNA-binding
451 protein. *VSG* mRNA was detected by RT-PCR using a spliced leader primer and a
452 primer that hybridizes to a conserved region within the *VSG* 3'-UTR.

453 D) Trypanosomes were taken after 48h RBP10-myc induction. They were stained for
454 EP procyclin and phosphorylated GPEET procyclin (green), VSG117 (magenta) and
455 DNA (cyan). The stain is overlaid with differential interference contrast (grey). The
456 panel shows a typical green procyclic form, a bloodstream form with two kinetoplasts,
457 and a third cell which resembles a bloodstream form (terminal kinetoplast) but was
458 not stained by anti-VSG117. This cell may express a VSG that does not react with
459 the anti-VSG117 antibodies.

460 E) Quantitation of surface EP procyclin by flow cytometry, comparing cells with (day
461 2) and without (day 0) induced RBP10-myc expression.

462 F) Quantitation of surface phospho-GPEET procyclin by flow cytometry, comparing
463 cells with (day 2) and without (day 0) induced RBP10-myc expression.

464 G) Quantitation of staining with anti-VSG117 by flow cytometry, comparing cells with
465 (day 2) and without (day 0) induced RBP10-myc expression.

466 H) Measurement of surface protein expression by FACSscan. Values for the different
467 windows are mean and standard deviation from 4 replicates. Similar results were
468 obtained by manual counting of stained smears.

469

470 The conversion of procyclic forms to bloodstream forms was highly reproducible. It
471 was not due to persistence of a few bloodstream forms in the procyclic cultures,
472 since survival always depended on induced RBP10 expression (Fig 6 A). RBP10
473 induction for two days was required; this was also the time needed to detect native
474 RBP10 expression by Western blotting. Moreover, our preliminary results indicate
475 that transient expression of RBP10 from an episome is sufficient to generate
476 bloodstream forms. In contrast, attempts to obtain bloodstream forms by RBP10
477 expression in Lister 427 cells that have been in procyclic form culture for more than
478 two decades were not successful.

479 Interestingly, epimastigote-like cells were never seen after induction of RBP10-myc
480 expression in our differentiation-competent cells (24h or 48h), and the epimastigote-
481 specific surface protein BARP was not detected by Western blotting or
482 immunofluorescence. Although we may have missed transient formation of
483 epimastigotes, this is unlikely since the behaviour of the trypanosomes was clearly
484 not synchronous. Moreover, the presence of dividing cells with bloodstream-form
485 morphology, suggested that RBP10 induction might be causing a direct conversion of
486 procyclic forms to bloodstream forms, "jumping" past the non-dividing metacyclic
487 form. We therefore compared the mRNA changes after 6h RBP10-myc induction with
488 those found in salivary gland parasites (a mixture of epimastigotes and metacyclics)
489 (S4 Table, sheets 5 and 6). The transcriptome comparisons and morphological
490 analysis suggested that after 2 days of RBP10 expression the cell population might
491 include both metacyclic and bloodstream forms. Interestingly, however, the mRNA
492 encoding the meiosis protein MND1, which is specific to a pre-metacyclic
493 intermediate form, is bound by RBP10 and *decreases* upon RBP10 induction.
494 Evidence so far thus suggests that RBP10 induction causes a developmental switch
495 that by-passes the epimastigote stage.

496 Discussion

497 Our results indicate that the presence of RBP10 defines the identity of a
498 trypanosome as a replication-competent bloodstream form.

499 RBP10 binds to procyclic-specific mRNAs with a $UA(U)_6$ motif in the 3'-UTR,
500 suppressing translation and causing mRNA destruction. However, many RNAs that
501 are bound by RBP10 did not change in either abundance, or in the percentage in
502 polysomes, after RBP10 RNAi or forced expression. As noted above, our assay did
503 not measure ribosome density, so it is likely that many changes in translation went

504 undetected. However, our results are completely consistent with work on other RNA-
505 binding proteins: usually, only some of the bound mRNAs change after the protein is
506 removed (e.g. [57]). For most mRNAs, many different proteins can bind along the
507 same 3'-UTR - a sequence of 300nt is likely to bind at least twenty proteins - and it is
508 their combinatorial effects that determine mRNA behaviour.

509 RBP10 target mRNAs include several that have been studied in considerable detail,
510 including those encoding cytochrome complexes and the procyclins. However, the
511 direct effects of RBP10 are by no means sufficient to explain its far-reaching
512 influence, since many mRNAs that change in translation or abundance after RBP10
513 depletion are not bound by RBP10. This suggests that RBP10 depletion may initiate
514 a cascade of events. Our results show how this could occur. Notably, the direct
515 RBP10 targets include the mRNAs encoding three potential RNA-binding proteins,
516 ZC3H20, ZC3H21, and ZC3H22, These proteins are normally expressed only in
517 procyclic forms [15,16,58]. ZC3H20 binds to, and stabilises, at least two procyclic-
518 specific mRNAs [58], of which one, *MCP12*, is not bound by RBP10 but was
519 increased after *rbp10* RNAi. ZC3H22 suppresses expression in the tethering assay
520 [39], and is required for proliferation of procyclic forms [15,42]. Suppression of
521 ZC3H21 and ZC3H22 expression is clearly very important to bloodstream-form
522 trypanosome survival: their mRNA 3-UTRs contain, respectively, no fewer than 5 and
523 7 RBP10 binding motifs.

524 In addition to controlling ZC3H20-22, RBP10 binds to and represses mRNAs
525 encoding two procyclic-specific protein kinases (Tb927.8.6490, Tb927.11.15010) and
526 one procyclic specific protein phosphatase (Tb927.10.8050). Also bound by RBP10,
527 but with no detected effect of RNAi, are the transcripts encoding and protein kinases
528 RDK1 and NRKA. NRKA [15] is up-regulated after the onset of stumpy form
529 differentiation; it is possible that RBP10 indeed regulates its translation but we did not
530 detect this because of the insensitivity of our assay. RDK1 is a differentiation
531 repressor [59] and is more abundant in bloodstream forms, making regulation by
532 RBP10 unlikely. Similar to RBP10, depletion of *RDK1* primes bloodstream forms to
533 differentiate to procyclic forms. More than 20 mRNAs show similar changes after
534 RNAi targeting either *RDK1* or *RBP10*, and 16 of these are both bound and regulated
535 by RBP10. Surprisingly, the *RBP10* transcript was not significantly affected after
536 *RDK1* RNAi; however, the protein level and phosphorylation status of RBP10 were
537 not determined.

538 The changes in RNA regulators, protein kinases and phosphatases, and perhaps
539 also the changes in metabolism that are directly induced by RBP10 loss (e.g. [60])

540 will trigger downstream effects on signaling pathways. Such secondary effects of
541 *rbp10* RNAi must include the increase in PIP39 (Tb927.9.6090), a key regulator of
542 differentiation [55], and decreases in mRNAs encoding the RNA-binding proteins
543 PUF11, RBP9, ZC3H31, ZC3H46, ZC3H48, DRBD5 and HNRNPH (S1 Table, sheet
544 7). These in turn will trigger further changes in gene expression.

545 So far, the only condition known to allow differentiation of *T. brucei* procyclic forms to
546 bloodstream forms *in vitro* is induced expression of the putative RNA-binding protein
547 RBP6. This mimics the natural differentiation pathway, with formation of
548 epimastigotes within 24h and appearance of VSG-expressing metacyclic forms 4-5
549 days later [61]. *RBP6* mRNA has low abundance in both procyclic an bloodstream
550 forms, and is maximal in salivary gland parasites [38]. RBP6 action is therefore
551 transient, although it might possibly define the epimastigote form; its mechanism of
552 action is unknown. Depletion of the RNA-binding protein DRBD18 from procyclic
553 forms caused a partial switch to a bloodstream-form or epimastigote expression
554 pattern; the mRNAs that increased included those encoding RBP6 and RBP10, but
555 the ability to differentiate further was not tested [62]. Heat shock can also cause
556 small increases in epimastigote-specific transcripts [63]. The rapid conversion of
557 procyclic forms to bloodstream forms by RBP10 expression, in contrast, might be
558 regarded as a form of *trans*-differentiation since the intermediate epimastigote stage
559 was not detected. In animal cells, *trans*-differentiation can be achieved through
560 changes in expression of transcription factors, either by direct manipulation or by
561 changing expression of microRNAs that control transcription factor expression [64-
562 66]. However, these changes have so far always been unidirectional. *Trans*
563 differentiation that works in both directions according to the presence or absence of a
564 single protein has, to our knowledge, not previously been observed.

565 It has long been known that stumpy forms cannot revert to long-slender bloodstream
566 forms, and recent results indicate that the differentiation of stumpy forms to
567 procyclics is an irreversible bi-stable switch [15]. We suggest that expression of
568 RBP10 governs a bi-stable switch during the transition from bloodstream forms to
569 stumpy forms, and from procyclic forms to bloodstream forms. The bi-stable
570 character can be explained by the fact that RBP10 directly suppresses procyclic-
571 specific post-transcriptional regulators. While RBP10 is present, the bloodstream-
572 form state is self-reinforcing; loss of RBP10 enables procyclic regulators to be
573 expressed, while expression of RBP10 in procyclic forms tips the balance towards a
574 bloodstream-form expression pattern. In conclusion, our results demonstrate how
575 differentiation states can be maintained by a post-transcriptional cascade.

576 ## Materials and methods

577 ### Parasite culture and plasmid constructs

578 *T. brucei* Lister 427 [67] and pleomorphic AnTat 1.1 (EATRO 1125) [56] cells
579 expressing the Tet-repressor were used. Lister 427 were used for all experiments
580 except those shown in Figs 5 and 6. Bloodstream form parasites were cultured in
581 HMIi-9 medium supplemented with 10% fetal bovine serum at 37°C with 5% CO2.
582 The procyclic forms were grown at 27°C in MEM-Pros medium supplemented with
583 10% heat inactivated fetal bovine serum. Stable cell lines were generated and
584 maintained as described in [56,67] except that selection of the pleomorphic cells was
585 done with 8µg/ml hygromycin and 2µg/ml blasticidin.
586 The tetracycline inducible constructs for *RBP10* RNAi and over expression were
587 described in [37]. All new constructs and oligonucleotides are in S6 Table. A
588 bloodstream-form cell line expressing all RBP10 with a tandem affinity purification tag
589 at the N-terminus was generated by inserting the tag at the 5' end of one
590 endogenous *RBP10* open reading frame (pHD2506) and deletion of the second copy
591 (pHD2061). Plasmids for the tethering assays are in S6 Table. The tethering
592 constructs were separately transfected in a cell line constitutively expressing the *CAT*
593 reporter with 5 copies of *boxB* preceding the *ACT* 3' UTR. Expression of the fusion
594 protein was induced for 24h using tetracycline (100ng/ml) and the CAT assays
595 carried out as described in [25].

596 ### Cross-linking and RNA immunoprecipitation

597 3x10^{9} bloodstream form cells expressing TAP-RBP10 were irradiated using UV (400
598 mj/cm2), washed in cold PBS and the cell pellet snap frozen in liquid nitrogen. The
599 RNA immunoprecipitation was done as described in [24]. Briefly, the extracts were
600 incubated with the beads, and the unbound fraction was collected. After washing,
601 bound RBP10 was eluted using TEV protease. After digestion with 50µg proteinase K
602 at 42°C for 15 minutes to reduce cross-linked protein, RNA was isolated from both
603 the bound and unbound fractions using Trifast reagent (Peqlab, GMBH). To assess
604 the quality of the purified RNA, an aliquot of the sample was analysed by Northern
605 blotting and the blot hybridized with a splice leader probe. Total RNA from the
606 unbound fraction was depleted of ribosomal RNA (rRNA) using RNAse H as
607 described in [68] except that a cocktail of 50-base DNA oligos complementary to
608 trypanosome rRNAs was used [63]. The recovered RNA from both bound and
609 unbound samples was then analysed by RNA-Seq.

610 **Tandem affinity purification (TAP) and genome-wide yeast two-hybrid screen**

611 Approximately $2x10^{10}$ bloodstream form cells expressing in-situ TAP-RBP10 were
612 harvested and used for TAP as previously described [69]. Three biological replicates
613 were done, with or without RNAse A treatment; triplicate results from a cell line with
614 inducible GFP-TAP served as the background control. The eluate was separated on
615 a 10% SDS-polyacrylamide gel for only 2 cm and the proteins visualized by colloidal
616 Coomassie staining. The gel area was excised and analysed by mass spectroscopy.
617 For the yeast two-hybrid screen, the RBP10 open reading frame was cloned into the
618 pGBKT7 plasmid (pHD2361). The construct was used as a bait to screen a library
619 made up of *T. brucei* random genomic DNA fragments. The screening, high
620 throughput sequencing and analysis were done as previously described [70].

621 **Polysome fractionation**

622 Approximately $3\text{-}5x10^{8}$ cells were collected by centrifugation (850 x g, 10 minutes,
623 21°C), and then treated in serum free media with 100µg/ml cycloheximide for 7
624 minutes at room temperature. The pellet was washed in 1 ml ice cold PBS, lysed in
625 350µl lysis buffer (20mM Tris pH7.5, 20mM KCl, 2mM MgCl2, 1mM DTT, 1000u
626 RNasin (Promega), 10µg/ml leupeptin, 100µg/ml cycloheximide, 1× complete
627 protease Inhibitor without EDTA (Roche), 0.2% (vol/vol) IGEPAL) by passing 15-30
628 times through a 21G needle, followed by centrifugation (15000 x g, 10 minutes, 4°C)
629 to clear the lysate. KCl was adjusted to 120 mM and the clarified lysate loaded on top
630 of a 4 ml continuous linear 15-50% sucrose (w/v) gradient in polysome buffer (20mM
631 Tris pH7.5, 120mM KCl, 2mM MgCl2, 1mM DTT, 10µg/ml leupeptin, 100µg/ml
632 cycloheximide). After 2 hours of ultracentrifugation (40000 rpm, 4°C; Beckman SW60
633 rotor), 400µl fractions were collected using Teledyne Isco Foxy Jr. gradient
634 fractionator system and RNA was isolated using Trifast reagent (Peqlab, GMBH). In
635 the case of RBP10 tethering, lambdaN-RBP10 was induced (24h) and the distribution
636 of the *CAT* reporter and *alpha-tubulin* mRNAs in the collected fractions detected by
637 Northern blotting as described in [71].
638 For RNASeq, bloodstream form cells plus or minus *RBP10* RNAi for 15h and, the
639 procyclic form cells with or without RBP10-myc over-expression for 6h were used. In
640 this case 250µl fractions were collected from each gradient and RNA was isolated
641 after pooling the fractions into two groups; i) lighter fractions including monosomes, ii)
642 denser fractions with at least two ribosomes. The amount of mRNA in the pooled
643 samples was assessed by Northern blotting using splice leader RNA as probe. Also,
644 total RNA was prepared from the input samples (~10% of total cell lysate) to quantify

645 changes in the steady state mRNAs levels. All samples were depleted of rRNAs (as
646 above) prior to analysis by RNA-seq.

647 **High throughput RNA sequencing and bioinformatic analysis**

648 RNA-seq was done at the CellNetworks Deep Sequencing Core Facility at the
649 University of Heidelberg. For library preparation, NEBNext Ultra RNA Library Prep Kit
650 for Illumina (New England BioLabs Inc.) was used. The libraries were multiplexed (6
651 samples per lane) and sequenced with a HiSeq 2000 system, generating 50 bp
652 single-end sequencing reads.
653 The quality of the raw sequencing data was checked using FastQC
654 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and the sequencing
655 primers removed using Cutadapt [72]. The data was aligned to the *T. brucei* TREU
656 927 reference genome using Bowtie [73], then sorted and indexed using SAMtools
657 [74]. Reads aligning to open reading frames of the TREU 927 genome were counted
658 using custom python scripts. Analysis for differentially expressed genes was done in
659 R using the DESeq package [75] with an adjusted p-value cut-off of 0.05.
660 Comparative analysis was limited to a list of unique genes modified from [9]. Gene
661 annotations are manually updated versions of those in TritrypDB
662 (http://tritrypdb.org/tritrypdb/), and categories were assigned manually. Statistical
663 analysis was done in R. The 3' UTR motif enrichment search was done using
664 DREME [76]; annotated 3' UTRs sequences were downloaded from tritrypDB and we
665 considered only the mRNAs with 3' UTR >20 nt. Manual 3'-UTR annotation was done
666 using the RNASeq reads and poly(A) site data [9,11,77] in tritrypDB.

667 **Analysis of 3'-UTR reporters**

668 Bloodstream form cells expressing the *CAT* reporter with either full length *EP1* 3'
669 UTR (pHD1610) or a mutant version (pHD1611) lacking the 26mer (*EP1Δ26*)
670 instability element [34] were used for RNA immunoprecipitation . For the pull down,
671 $4x10^8$ cells (without the stem-loop) were irradiated using UV (400 mj/cm^2), washed in
672 cold PBS and the cell pellet lysed in 350µl lysis buffer (10mM Tris pH 7.5, 10mM
673 NaCl, 1000u RNasin (Promega), 1× complete protease Inhibitor without EDTA
674 (Roche), 0.1% IGEPAL) by passing 15 times through a 21G needle. The lysate was
675 cleared by centrifugation at 15000g for 10 minutes at 4^oC, NaCl was adjusted to
676 150mM followed by incubation with 50µl anti-RBP10 [37] coupled agarose beads for
677 2 hours at 4^oC. After washing the beads 5 times with IPP150 buffer (10mM Tris pH
678 7.5, 150mM NaCl, 0.1% IGEPAL), the beads were treated with 20µg proteinase K at
679 42^oC for 15 minutes and RNA was isolated from both bound and unbound fractions

680 using Trifast reagent (peqlab, GMBH). Equal amounts of the recovered RNA (eluate
681 and flow through fractions) were reverse-transcribed using RevertAid First Strand
682 cDNA Synthesis Kit (Thermal Scientific) according to the manufacturer's instructions.
683 2µl of the cDNA was used as template in a 50µl PCR reaction to detect the *CAT* and
684 *alpha-tubulin* genes. PCR was done using Q5 DNA polymerase and buffer (NEB)
685 with 0.5µm of the following primers, for *CAT* (CZ5725; CZ689) and *alpha-tubulin*
686 (CZ5725; CZ6168); the forward primer is the same for both since it anneals to the
687 splice leader. Aliquots (10µl) were removed after 27, 31 and 36 cycles and analysed
688 by agarose gel electrophoresis.

689 To determine if the regulation of the *EP* mRNA by RBP10 depends on the 26mer
690 instability element, a stem-loop construct (pHD1984) targeting *RBP10* was
691 transfected into the two cell lines [37]. *RBP10* RNAi was induced using 100ng/ml
692 tetracycline for 17h or 24h and CAT activity was measured.

693 **RBP10 and trypanosome differentiation**

694 For bloodstream-procyclic form conversion, pleomorphic AnTat 1.1 bloodstream form
695 cells with stem-loop RNAI (pHD1984) targeting *RBP10* were used. 17h after RNAi
696 induction, the cells were pelleted, resuspended (~$8x10^5$ cells/ml) in procyclic form
697 (MEM-pros) medium and incubated at 27°C. As positive controls, wild type or
698 uninduced cells ($2x10^6$ cells/ml) were treated with 6mM cis-aconitate (Sigma) at
699 27°C; after 17h the cells were transferred into procyclic form media (~$8x10^5$ cells/ml)
700 and maintained at 27°C. Samples were taken at different times for Western blotting
701 (~$5x10^6$) and for morphological analysis.

702 To convert procyclic to bloodstream forms, AniTat 1.1 bloodstream form cells with an
703 inducible RBP10-myc construct (pHD2098) were differentiated to procyclic forms
704 using cis-aconitate as described above. The cells were cultured in presence of
705 hygromycin (8µg/ml) and phleomycin (0.2µg/ml) for more than 3 months to generate
706 well-established procyclic forms. RBP10-myc was induced using 100ng/ml
707 tetracycline. Marker proteins were detected by Western blot. VSG transcripts were
708 detected by semi-quantitative RT-PCR as previously described [61] using primers
709 CZ6308/CZ6309. To obtain growing bloodstream forms, RBP10-myc was induced for
710 48 hours, the cells were pelleted, resuspended ($2x10^5$ cells/ml) in HMI-9 medium and
711 then incubated at 37°C with 5% CO_2. The cell density was monitored for 6 days or
712 more; wild type or uninduced cells served as control.

713 **Morphological analysis**

714 To stain surface proteins, dried smears were fixed with 100% methanol at -20°C for
715 15 minutes, rehydrated in 1x PBS for 15 minutes, blocked for 20 minutes with 20%
716 FCS, then labeled with mouse anti-EP (1:500; Cedarlane, Canada), mouse anti-
717 phospho-GPEET (1:500; Cedarlane, Canada) and rabbit anti-VSG-117 (1:500, from
718 either G. Cross or P. Overath). Anti-BARP staining was done as described in [5],
719 using a procyclic cell line with the *BARP* open reading frame in the procyclin locus as
720 a positive control. For PGK staining, 2x10^{6} cells were collected by centrifugation, re-
721 suspended, fixed in 4% paraformadehyde (in 1x PBS) for 18 minutes, sedimented
722 again for 2 minutes, re-suspended in PBS and allowed to settle on poly-L-lysine
723 coated slides for 30 minutes. Before staining, slides were blocked with 20% fetal calf
724 serum (1x PBS) for 20 minutes. Cells were permeabilised with 0.2% (v/v) Triton-X
725 100 (in 1xPBS) for 15 minutes at room temperature, washed twice then incubated for
726 one hour with rabbit anti PGK antibody (1:1500, [78]). The cells were washed 3 times
727 before being stained with fluor-conjugated secondary antibody (1:500; mouse-Cy3 or
728 rabbit-Alexa-488; Molecular probes, Eugene). Cellular DNA was stained with
729 100ng/ml DAPI in 1x PBS for 15 minutes. Images were taken using Olympus Cell-R
730 microscope. Random fields were photographed (by E.M.) and analysed (by C.C.)
731 using Fiji [79]. We attempted blind analysis but this was not possible because the
732 differences between samples were too obvious.

733 **Flow cytometry**

734 Approximately 5x10^{6} cells were fixed with 2% formaldehyde/0.05% glutaraldehyde at
735 4°C for at least 1 hour. The cells were pelleted, washed twice with PBS then
736 incubated with 200µl (2% BSA in PBS) mouse anti-EP (Cedarlane, Canada; 1:500),
737 anti phosph-GPEET (Cedarlane, Canada; 1:500), or rabbit anti-VSG-117 (1:500) for
738 one hour on ice. After washing twice, the cells were stained with the secondary
739 antibody (1:500; mouse-Cy5 or rabbit-Alexa-488; Molecular probes, Eugene) for one
740 hour. Cells stained only with the secondary antibody and the unstained cells served
741 as the negative controls. Flow cytometry was performed with BD FACSCanto II flow
742 cytometer and the data analysed using FlowJo software (TreeStar Inc.).

743 **Western blotting**

744 Western blotting was done as previously described [71]. Antibodies were: anti-
745 RBP10 (rat, 1:2000, [37]), anti-HNRNPH (rabbit,1:5000, [20]), anti-TAO (rabbit,1:100,
746 [78]), anti-EP (mouse,1:2000, Cedarlane, Canada), anti-phospho-GPEET
747 (mouse,1:2000, Cedarlane, Canada), anti-BARP (rabbit,1:2500, [5]), and anti-PIP39

748 (rabbit,1:1000, [55]) and anti-PAD1 (rabbit,1:1000, [54] with electrophoresis under
749 non-denaturing conditions) .

750 **Data availability.**

751 All sequence datasets are available at Array Express with the accession number E-
752 MTAB-4564.

753 **Ethical approval**

754 The research did not involve animals or humans and did not require ethical approval.

755 Acknowledgements

756 For antibodies we thank Keith Matthews (Edinburgh University, Scotland; PIP39 and
757 PAD1), Isabel Roditi (University of Bern, Switzerland, BARP); Shula Michaeli (Bar-
758 Ilan University, Israel, anti-HNRNPH), Paul MIchels (anti-PGK, Edinburgh University,
759 Scotland) and Minu Chauduri (Meharry Medical College, Nashville, Tennessee, USA,
760 AOX). The anti-VSG117 antibody was either from Peter Overath (Max-Planck Institut,
761 Tübingen, Germany) or George Cross (Rockefeller University, USA). We also thank
762 Keith, Isabel and Shula for useful discussions. All RNA-Seq libraries were prepared
763 by David Ibberson at Bioquant, Heidelberg, and sequenced at EMBL. We thank
764 Claudia Helbig and Ute Leibfried for technical assistance and Kevin Leiss for
765 counting the numbers of elements in all 3'-UTRs.

766 References

767 1. Horn D (2014) Antigenic variation in African trypanosomes. Mol Biochem Parasitol
768 195: 123-129.
769 2. Vassella E, Reuner B, Yutzy B, Boshart M (1997) Differentiation of African
770 trypanosomes is controlled by a density sensing mechanism which signals
771 cell cycle arrest via the cAMP pathway. J Cell Sci 110: 2661-2671.
772 3. Vassella E, Acosta-Serrano A, Studer E, Lee SH, Englund PT, et al. (2001)
773 Multiple procyclin isoforms are expressed differentially during the
774 development of insect forms of Trypanosoma brucei. J Mol Biol 312: 597-607.
775 4. Rotureau B, Van Den Abbeele J (2013) Through the dark continent: African
776 trypanosome development in the tsetse fly. Front Cell Infect Microbiol 3: 53.
777 5. Urwyler S, Studer E, Renggli CK, Roditi I (2007) A family of stage-specific alanine-
778 rich proteins on the surface of epimastigote forms of Trypanosoma brucei.
779 Mol Microbiol 63: 218-228.
780 6. Peacock L, Bailey M, Carrington M, Gibson W (2013) Meiosis and Haploid
781 Gametes in the Pathogen Trypanosoma brucei. Curr Biol.
782 7. Peacock L, Ferris V, Sharma R, Sunter J, Bailey M, et al. (2011) Identification of
783 the meiotic life cycle stage of *Trypanosoma brucei* in the tsetse fly. Proc Natl
784 Acad Sci U S A 108: 3671-3676.

785 8. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, et al. (2010) Spliced
786 leader trapping reveals widespread alternative splicing patterns in the highly
787 dynamic transcriptome of *Trypanosoma brucei*. PLoS Pathog 6: e1001037.
788 9. Siegel T, Hekstra D, Wang X, Dewell S, Cross G (2010) Genome-wide analysis of
789 mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and
790 identification of splicing and polyadenylation sites. Nucleic Acids Res 38:
791 4946-4957.
792 10. Fadda A, Ryten M, Droll D, Rojas F, Färber V, et al. (2014) Transcriptome-wide
793 analysis of mRNA decay reveals complex degradation kinetics and suggests
794 a role for co-transcriptional degradation in determining mRNA levels. Mol
795 Microbiol 94: 307-326.
796 11. Jensen BC, Ramasamy G, Vasconcelos EJ, Ingolia NT, Myler PJ, et al. (2014)
797 Extensive stage-regulation of translation revealed by ribosome profiling of
798 *Trypanosoma brucei*. BMC Genomics 15: 911.
799 12. Vasquez JJ, Hon CC, Vanselow JT, Schlosser A, Siegel TN (2014) Comparative
800 ribosome profiling reveals extensive translational complexity in different
801 *Trypanosoma brucei* life cycle stages. Nucleic Acids Res 42: 3623-3637.
802 13. Gunasekera K, Wuthrich D, Braga-Lagache S, Heller M, Ochsenreiter T (2012)
803 Proteome remodelling during development from blood to insect-form
804 Trypanosoma brucei quantified by SILAC and mass spectrometry. BMC
805 Genomics 13: 556.
806 14. Urbaniak MD, Guther ML, Ferguson MA (2012) Comparative SILAC proteomic
807 analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages.
808 PLoS One 7: e36619.
809 15. Domingo-Sananes MR, Szoor B, Ferguson MA, Urbaniak MD, Matthews KR
810 (2015) Molecular control of irreversible bistability during trypanosome
811 developmental commitment. J Cell Biol 211: 455-468.
812 16. Dejung M, Subota I, Bucerius F, Dindar G, Freiwald A, et al. (2016) Quantitative
813 proteomics uncovers novel factors involved in developmental dfferentiation of
814 *Trypanosoma brucei*. PLoS Pathog 12: e1005439.
815 17. Michaeli S (2011) Trans-splicing in trypanosomes: machinery and its impact on
816 the parasite transcriptome. Future Microbiol 6: 459-474.
817 18. Clayton C, Michaeli S (2011) 3' processing in protists. Wiley Interdiscip Rev RNA
818 2: 247-255.
819 19. Clayton C (2013) The regulation of trypanosome gene expression by RNA-
820 binding proteins. PLoS Pathog 9: e1003680.
821 20. Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, et al. (2013) The hnRNP F/H
822 homologue of *Trypanosoma brucei* is differentially expressed in the two life
823 cycle stages of the parasite and regulates splicing and mRNA stability.
824 Nucleic Acids Res 41: 6577-6594.
825 21. Gupta SK, Chikne V, Eliaz D, Tkacz ID, Naboishchikov I, et al. (2014) Two
826 splicing factors carrying serine-arginine motifs, TSR1 and TSR1IP, regulate
827 splicing, mRNA stability, and rRNA processing in Trypanosoma brucei. RNA
828 Biol 11.
829 22. Antwi E, Haanstra J, Ramasamy G, Jensen B, Droll D, et al. (2016) Integrative
830 analysis of the *Trypanosoma brucei* gene expression cascade predicts
831 differential regulation of mRNA processing and unusual control of ribosomal
832 protein expression. BMC Genomics 17: 306.
833 23. Queiroz R, Benz C, Fellenberg K, V S, Hoheisel J, et al. (2009) Transcriptome
834 analysis of differentiating trypanosomes reveals the existence of multiple
835 post-transcriptional regulons. BMC Genomics 10: 495.
836 24. Archer SK, van Luu D, de Queiroz R, Brems S, Clayton CE (2009) *Trypanosoma*
837 *brucei* PUF9 regulates mRNAs for proteins involved in replicative processes
838 over the cell cycle. PLoS Pathog 5: e1000565.

839 25. Droll D, Minia I, Fadda A, Singh A, Stewart M, et al. (2013) Post-transcriptional
840 regulation of the trypanosome heat shock response by a zinc finger protein.
841 PLoS Pathog 9: e1003286.
842 26. Kolev NG, Ullu E, Tschudi C (2014) The emerging role of RNA-binding proteins in
843 the life cycle of *Trypanosoma brucei*. Cell Microbiol 16: 482-489.
844 27. Janzen CJ, van Deursen, F., Shi H, Cross GAM, Matthews KR, et al. (2006)
845 Expression site silencing and life-cycle progression appear normal in
846 Argonaute1-deficient *Trypanosoma brucei* Mol Biochem Parasitol 14: 102-
847 107.
848 28. Tschudi C, Shi H, Franklin J, Ullu E (2012) Small interfering RNA-producing loci
849 in the ancient parasitic eukaryote *Trypanosoma brucei*. BMC Genomics 13:
850 427.
851 29. Gunzl A, Bruderer T, Laufer G, Schimanski B, Tu LC, et al. (2003) RNA
852 polymerase I transcribes procyclin genes and variant surface glycoprotein
853 gene expression sites in *Trypanosoma brucei.* Eukaryot Cell 2: 542-551.
854 30. Biebinger S, Rettenmaier S, Flaspohler J, Hartmann C, Peña-Diaz J, et al. (1996)
855 The PARP promoter of *Trypanosoma brucei* is developmentally regulated in a
856 chromosomal context. Nucleic Acids Res 24: 1202-1211.
857 31. Furger A, Schürch N, Kurath U, Roditi I (1997) Elements in the 3' untranslated
858 region of procyclin mRNA regulate expression in insect forms of
859 *Trypanosoma brucei* by modulating RNA stability and translation. Mol Cell
860 Biol 17: 4372-4380.
861 32. Hehl A, Vassella E, Braun R, Roditi I (1994) A conserved stem-loop structure in
862 the 3' untranslated region of procyclin mRNAs regulates expression in
863 *Trypanosoma brucei*. Proc Natl Acad Sci USA 91: 370-374.
864 33. Schürch N, Furger A, Kurath U, Roditi I (1997) Contribution of the procyclin 3'
865 untranslated region and coding region to the regulation of expression in
866 bloodstream forms of *Trypanosoma brucei*. Mol Biochem Parasit 89: 109-121.
867 34. Hotz H-R, Hartmann C, Huober K, Hug M, Clayton CE (1997) Mechanisms of
868 developmental regulation in *Trypanosoma brucei*: A polypyrimidine tract in the
869 3'-untranslated region of a trypanosome surface protein mRNA affects RNA
870 abundance and translation. Nucleic Acids Res 25: 3017-3025.
871 35. Hotz H-R, Biebinger S, Flaspohler J, Clayton CE (1998) *PARP* gene expression:
872 regulation at many levels. Mol Biochem Parasitol 91: 131-143.
873 36. Engstler M, Boshart M (2004) Cold shock and regulation of surface protein
874 trafficking convey sensitization to inducers of stage differentiation in
875 *Trypanosoma brucei*. Genes Dev 18: 2798-2811.
876 37. Wurst M, Selinger B, Jha B, Klein C, Queiroz R, et al. (2012) Expression of the
877 RNA Recognition Motif protein RBP10 promotes a bloodstream-form
878 transcript pattern in *Trypanosoma brucei*. Mol Microbiol 83: 1048-1063.
879 38. Telleria EL, Benoit JB, Zhao X, Savage AF, Regmi S, et al. (2014) Insights into
880 the trypanosome-host interactions revealed through transcriptomic analysis of
881 parasitized tsetse fly salivary glands. PLoS Negl Trop Dis 8: e2649.
882 39. Erben E, Fadda A, Lueong S, Hoheisel J, Clayton C (2014) Genome-wide
883 discovery of post-transcriptional regulators in *Trypanosoma brucei*. PLoS
884 Pathogens 10: e1004178.
885 40. Lueong S, Merce C, Fischer B, Hoheisel J, Erben E (2016) Gene expression
886 regulatory networks in *Trypanosoma brucei*: insights into the role of the
887 mRNA-binding proteome. Mol MIcrobiol 100: 457-471.
888 41. Zinoviev A, Leger M, Wagner G, Shapira M (2011) A novel 4E-interacting protein
889 in *Leishmania* is involved in stage-specific translation pathways. Nucleic
890 Acids Res 39: 8404-8415.
891 42. Alsford S, Turner D, Obado S, Sanchez-Flores A, Glover L, et al. (2011) High
892 throughput phenotyping using parallel sequencing of RNA interference targets
893 in the African trypanosome. Genome Res 21: 915-924.

894 43. Mayho M, Fenn K, Craddy P, Crosthwaite S, Matthews K (2006) Post-
895 transcriptional control of nuclear-encoded cytochrome oxidase subunits in
896 *Trypanosoma brucei*: evidence for genome-wide conservation of life-cycle
897 stage-specific regulatory elements. Nucleic Acids Res 34: 5312-5324.
898 44. Arieti F, Gabus C, Tambalo M, Huet T, Round A, et al. (2014) The crystal
899 structure of the Split End protein SHARP adds a new layer of complexity to
900 proteins containing RNA recognition motifs. Nucleic Acids Res 42: 6742-
901 6752.
902 45. Morgan CE, Meagher JL, Levengood JD, Delproposto J, Rollins C, et al. (2015)
903 The First Crystal Structure of the UP1 Domain of hnRNP A1 Bound to RNA
904 Reveals a New Look for an Old RNA Binding Protein. J Mol Biol 427: 3241-
905 3257.
906 46. Wang X, Tanaka Hall T (2001) Structural basis for recognition of AU-rich element
907 RNA by the HuD protein. Nat Struct Biol 8: 141-145.
908 47. Deo R, Bonanno J, Sonenberg N, Burley S (1999) Recognition of polyadenylate
909 RNA by the poly(A)-binding protein. Cell 98: 835-845.
910 48. Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, et al. (1999) Structural basis
911 for recognition of the tra mRNA precursor by the Sex-lethal protein. Nature
912 398: 579-585.
913 49. Drozdz M, Clayton CE (1999) Structure of a regulatory 3'-untranslated region
914 from *Trypanosoma brucei*. RNA 5: 1632-1644.
915 50. Misset O, Opperdoes FR (1987) The phosphoglycerate kinases from
916 *Trypanosoma brucei*. Eur J Biochem 162: 493-500.
917 51. Gibson WC, Swinkels BW, Borst P (1988) Post-transcriptional control of the
918 differential expression of phosphoglycerate kinase genes in *Trypanosoma*
919 *brucei*. J Mol Biol 201: 315-325.
920 52. Blattner J, Clayton CE (1995) The 3'-untranslated regions from the *Trypanosoma*
921 *brucei* phosphoglycerate kinase genes mediate developmental regulation.
922 Gene 162: 153-156.
923 53. Quijada L, Hartmann C, Guerra-Giraldez C, Drozdz M, Irmer H, et al. (2002)
924 Expression of the human RNA-binding protein HuR in *Trypanosoma brucei*
925 induces differentiation-related changes in the abundance of developmentally-
926 regulated mRNAs. Nucleic Acids Res 30: 1-11.
927 54. Dean S, Marchetti R, Kirk K, Matthews K (2009) A surface transporter family
928 conveys the trypanosome differentiation signal. Nature 459: 213-217.
929 55. Szoor B, Ruberto I, Burchmore R, Matthews KR (2010) A novel phosphatase
930 cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal
931 signaling pathway. Genes Dev 24: 1306-1316.
932 56. Benz C, Mulindwa J, Ouna B, Clayton C (2011) The *Trypanosoma brucei* zinc
933 finger protein ZC3H18 is involved in differentiation. Mol Biochem Parasitol
934 177: 148-151.
935 57. Kershaw CJ, Costello JL, Talavera D, Rowe W, Castelli LM, et al. (2015)
936 Integrated multi-omics analyses reveal the pleiotropic nature of the control of
937 gene expression by Puf3p. Sci Rep 5: 15518.
938 58. Ling A, Trotter J, Hendriks E (2011) A zinc finger protein, TbZC3H20, stabilises
939 two developmentally regulated mRNAs in trypanosomes. J Biol Chem 286:
940 20152-20162.
941 59. Jones NG, Thomas EB, Brown E, Dickens NJ, Hammarton TC, et al. (2014)
942 Regulators of Trypanosoma brucei cell cycle progression and differentiation
943 identified using a kinome-wide RNAi screen. PLoS Pathog 10: e1003886.
944 60. Vassella E, Probst M, Schneider A, Studer E, Renggli C, et al. (2004) Expression
945 of a major surface protein of *Trypanosoma brucei* insect forms is controlled by
946 the activity of mitochondrial enzymes. Mol Biol Cell 15: 3986-3993.

947 61. Kolev NG, Ramey-Butler K, Cross GA, Ullu E, Tschudi C (2012) Developmental
948 progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding
949 protein. Science 338: 1352-1353.
950 62. Lott K, Mukhopadhyay S, Li J, Wang J, Yao J, et al. (2015) Arginine methylation
951 of DRBD18 differentially impacts its opposing effects on the trypanosome
952 transcriptome. Nucleic Acids Res 43: 5501-5523.
953 63. Minia I, Merce C, Terrao M, Clayton C (2016) Many mRNAs that increase after
954 heat shock of procyclic-form *Trypanosoma brucei* are also induced during
955 differentiation to mammalian-infective forms. submitted.
956 64. Wang H, Li X, Gao S, Sun X, Fang H (2015) Transdifferentiation via transcription
957 factors or microRNAs: Current status and perspective. Differentiation 90: 69-
958 76.
959 65. van der Meulen T, Huising MO (2015) Role of transcription factors in the
960 transdifferentiation of pancreatic islet cells. J Mol Endocrinol 54: R103-117.
961 66. Gao T, McKenna B, Li C, Reichert M, Nguyen J, et al. (2014) Pdx1 maintains
962 beta cell identity and function by repressing an alpha cell program. Cell Metab
963 19: 259-271.
964 67. Alibu VP, Storm L, Haile S, Clayton C, Horn D (2004) A doubly inducible system
965 for RNA interference and rapid RNAi plasmid construction in *Trypanosoma*
966 *brucei*. Mol Biochem Parasitol 139: 75-82.
967 68. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, et al. (2013)
968 Comparative analysis of RNA sequencing methods for degraded or low-input
969 samples. Nat Methods 10: 623-629.
970 69. Estévez A, Kempf T, Clayton CE (2001) The exosome of *Trypanosoma brucei*.
971 EMBO J 20: 3831-3839.
972 70. Singh A, Minia I, Droll D, Fadda A, Clayton C, et al. (2014) Trypanosome MKT1
973 and the RNA-binding protein ZC3H11: interactions and potential roles in post-
974 transcriptional regulatory networks. Nucleic Acids Res 42: 4652-4668.
975 71. Minia I, Clayton C (2016) Regulating a post-transcriptional regulator: protein
976 phosphorylation, degradation and translational blockage in control of the
977 trypanosome stress-response RNA-binding protein ZC3H11. PLoS
978 Pathogens 12: e1005514.
979 72. Martin M (2011) Cutadapt removes adapter sequences from high-throughput
980 sequencing reads. EMBnetjournal 17: 10-12.
981 73. Langmead B, Trapnell C, Pop M, Salzberg S (2009) Ultrafast and memory-
982 efficient alignment of short DNA sequences to the human genome. Genome
983 Biol 10: R25.
984 74. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence
985 Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.
986 75. Anders S, Huber W (2010) Differential expression analysis for sequence count
987 data. Genome Biol 11: R106.
988 76. Bailey T (2011) DREME: Motif discovery in transcription factor ChIP-seq data.
989 Bioinformatics 27: 1653-1659.
990 77. Kolev N, Franklin J, Carmi S, Shi H, Michaeli S, et al. (2010) The transcriptome of
991 the human pathogen *Trypanosoma brucei* at single-nucleotide resolution.
992 PLoS Pathog 6: e1001090.
993 78. Zomer AW, Allert S, Chevalier N, Callens M, Opperdoes FR, et al. (1998)
994 Purification and characterisation of the phosphoglycerate kinase isoenzymes
995 of *Trypanosoma brucei* expressed in *Escherichia coli*
996 . Biochim Biophys Acta 1386: 179-188.
997 79. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, et al. (2012) Fiji:
998 an open-source platform for biological-image analysis. Nat Methods 9: 676-
999 682.

1000 80. Fritz M, Vanselow J, Sauer N, Lamer S, Goos C, et al. (2015) Novel insights into
1001 RNP granules by employing the trypanosome's microtubule skeleton as a
1002 molecular sieve. Nucl Acids Res 43: 8013-8032.
1003 81. Mony BM, Macgregor P, Ivens A, Rojas F, Cowton A, et al. (2013) Genome-wide
1004 dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*.
1005 Nature 505: 681-685.

1006

1007 ## Supporting information

1008 **S1 Text**

1009 3'-untranslated regions of mRNAs that were at least 3x enriched in both RBP10 pull-
1010 downs.

1011 **S2 Text**

1012 3'-untranslated regions of mRNAs that were less than 0.7x enriched in both RBP10
1013 pull-downs.

1014 **S3 Text**

1015 3'-untranslated regions of mRNAs that were at least 3x enriched on average in
1016 RBP10 pull-downs, according to DESeq (Padj <8 E-5), but one replicate was less
1017 than 3x enriched.

1018 **S1 Table**

1019 Proteins that co-purified with TAP-tagged RBP10.

1020 Endogenously TAP-tagged RBP10 was purified three times in the presence or
1021 absence of RNase. The controls were three independent purifications of GFP-TAP

1022 Sheet 1: Proteins enriched in the purifications. Peptide counts and ratios are shown.
1023 The highest peptide count from the GFP-TAP replicates was used as a baseline. The
1024 proteins shown in this table were present as at least 2 peptides, present in at least
1025 two similar purifications, and were on average at least 2-fold enriched relative to
1026 GFP-TAP. Functions of interest are highlighted in different colours. "Average PGKB
1027 =tet/-tet" are tethering screen results [39]; a value above 1.5 suggests that the
1028 protein can suppress expression. "in starvation granules?" is 1 if positive and 0 if
1029 negative [80]. The number of peptides found by purification of MKT1-TAP are also
1030 indicated [70].

1031 Sheet 2: All raw data.

1032 **S2 Table**

1033 Results of a yeast 20-hybrid screen with a random fragment library.

1034 The library [70] was screened with pGAD-RBP10 and results evaluated by
1035 sequencing.

1036 Sheet 1: Raw data for pGAD-RBP10. All reads that started from position -36 onwards
1037 (relative to the ATG) are included.

1038 Sheet 2: Raw data for pGAD-T7 (unselected background, filtered for at least 4
1039 reads)

1040 Sheet 3: Results for pGAD-RBP10, filtered to exclude loci with less than 10 reads,
1041 and open reading frames with less than 2 independent fragments. Numerous
1042 cysteine protease fragments have been removed.

1043 Sheet 4: Possible interaction partners. The ratio to the maximum background count is
1044 at least 4. Proteins that also scored as interacting with MKT1 or CFB1 [70] (indicating
1045 less specificity) are listed separately. The "regulator" column shows results from the
1046 tethering screen [39].

1047 **S3 Table**

1048 Effect of RBP10 depletion on gene expression in bloodstream forms.

1049 RNAi was induced for 15h and mRNA was collected for RNASeq from total cells, the
1050 free and monosomal fractions, and the polysomal fractions. Data are for two
1051 independent experiments which are labelled A and B.

1052 Throughout the Tables, the colour code for ratios is as follows: Red is >2, orange is
1053 1.5-2.0, green is 0.667 - 0.5 and blue is <0.5. Shading of columns does not follow a
1054 particular colour code but is designed to make the tables easier to navigate. Many
1055 numbers have been rounded to 2 or 3 points beyond the decimal in order to
1056 decrease the sizes of the files. Ribosome profiling results (Riboprof) are from [11],
1057 and developmental regulation data for total RNA are from [10].

1058 Sheet 1: Raw read counts for all unique genes, chosen according to [9] with some
1059 later additions.

1060 Sheet 2: Columns E - P Results were converted to reads per million reads (RPM).
1061 For each experiment, we then corrected the RPM for the monosomes and polysomes
1062 according to the fraction of the total mRNA that was in each fraction (sheet 7 and S3
1063 Fig, numbers not included in the Table). We then calculated the percentage of each
1064 mRNA that was on polysomes (columns Q-T). Column U contains the average of

1065 ratio between RNAi and control, calculated for the total input RNA - (K+L) / (E+F).
1066 Column V is the same ratio for the fraction of RNA not on polysomes, and column W
1067 is the ratio for the polysomal fractions (S+T) / (Q+R). To find the lowest possible
1068 estimate of the ratio in W, we took the lowest value for the proportion in polysomes
1069 after *rbp10* RNAi, and divided by the highest value from wild-type. From this column
1070 (X) we could find mRNAs that reliably move *into* polysomes after RNAi. Column Y is
1071 similar but shows the *highest* estimate for the ratio; this can be used to find mRNAs
1072 that reliably move *away* from polysomes after RNAi. Since there were only 3 genes in
1073 this category (column Y) they were not considered further.

1074 Sheet 3: mRNAs that move onto polysomes after *rbp10* RNAi. Same as sheet 2,
1075 except that only genes for which column W is more than 1.5, and column X is more
1076 than 1.2, are included. (This means that the average increase is at least 1.5x and the
1077 minimum is 1.2x.) Columns Y and Z are ribosome profiling results from [11], giving
1078 the ribosome density ratios for bloodstream forms (BS) divided by procyclic forms
1079 (PC). Slender bloodstream forms from blood are slBS and cultured bloodstream
1080 forms are cBS. Column AB is results from S5 Table, sheet 2 (binding of the RNA to
1081 RBP10, ratio of bound to unbound) and column AC is the adjusted P-values for the
1082 results in column AB.

1083 Sheet 4: DeSeq analysis of the data in Sheet 1.

1084 Sheet 5: Genes showing at least a 1.5-fold increase either in the amount of total
1085 RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-
1086 value threshold 0.05). Column AA is the RNA-binding ratio from S5 Table, sheet 2,

1087 Sheet 6: As sheet 5, except that these are genes showing at least a 1.5-fold
1088 decrease either in the amount of total RNA, or the amount of RNA in polysomes, or
1089 both, according to DeSeq (adjusted P-value threshold 0.05).

1090 Sheet 7: Quantitation of spliced leader signals for the free/monosome and polysome
1091 fractions used for RNA-Seq

1092 Sheet 8: Statistical analyses to compare the proteomes of differentiating
1093 trypanosomes [16] with the effects of RBP10 RNAi on RNA levels.

1094 **S4 Table**

1095 Effect of RBP10 expression on gene expression in procyclic forms.

1096 Expression of RBP10 was induced with tetracycline addition for 6 hours, and RNA
1097 samples prepared from duplicates as in Supplementary Table S3. Controls are the
1098 same cells without tetracycline.

1099 Sheet 1: Columns D-O are raw reads. Columns P-S are corresponding RPM values
1100 for input samples.

1101 Sheet 2: Proportions in free/monosome and polysome fractions were calculated as
1102 described for Supplementary Table S3, sheet 2. Columns Q-X: Effects on both
1103 polysomes and free/monosome fractions were also calculated in a similar way.

1104 Sheet 3: As sheet 2, but only genes showing regulated translation according to the
1105 following criteria: (a) movement to the free/monosome fraction from the polysomes.
1106 This is the top section (rows 2-320). The smallest value for the proportion in
1107 free/monosomes with RBP10 expression was divided by the largest value calculated
1108 value for the proportion in free/monosomes without RBP10 induction. If this number
1109 was greater than 1.5, the gene was included. In addition, we divided the largest
1110 number for the proportion in polysomes plus RBP10 expression by the smallest value
1111 for the proportion in polysomes without RBP10 induction. If this number was less
1112 than 0.667 the gene was also included. Genes that satisfied both criteria are in bold.
1113 Additional columns are as in Table S3. Rows 322-326 are the four RNAs showing
1114 movement onto polysomes.

1115 Sheet 4: DeSeq analysis of the results in Sheet 1.

1116 Sheet 5: Genes showing at least a 1.5-fold increase either in the amount of total
1117 RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-
1118 value threshold 0.05).

1119 Sheet 6: As sheet 5, except that these are genes showing at least a 1.5-fold
1120 decrease either in the amount of total RNA, or the amount of RNA in polysomes, or
1121 both, according to DeSeq (adjusted P-value threshold 0.05). Columns V-Z are results
1122 for salivary gland trypanosomes, calculated from the raw data in [38]; statistical
1123 analysis was not possible due to the absence of replicas.

1124 Sheet 7: Quantitation of spliced leader signals for the free/monosome and polysome
1125 fractions used for RNASeq.

1126 **S5 Table**

1127 RNAs bound to tagged RBP10.

1128 Results for two independent experiments, A and B, are shown.

1129 Sheet 1: Raw read counts for all samples, calculation of reads per million reads
1130 (RPM), and ratios of bound/unbound for each experiment.

1131 Sheet 2: DeSeq analysis of the raw data in sheet 1

1132 Sheet 3: Set of mRNAs that were at least 3-fold bound according to DeSeq. This is a
1133 slightly more relaxed stringency than that used in sheet 5. The largest P-value was
1134 7E-5. H-J: Results for developmental regulation [10]; K 0- calculated from [38]; M-T:
1135 results from Supplementary Tables S3 and S4; U, W, Z: [13]; V: [14]; X, Y, AA: [15];
1136 AB: [16]; AC: [81]; AD: [42].

1137 Sheet 4: Statistics used to create the Venn diagrams in Figure 2.

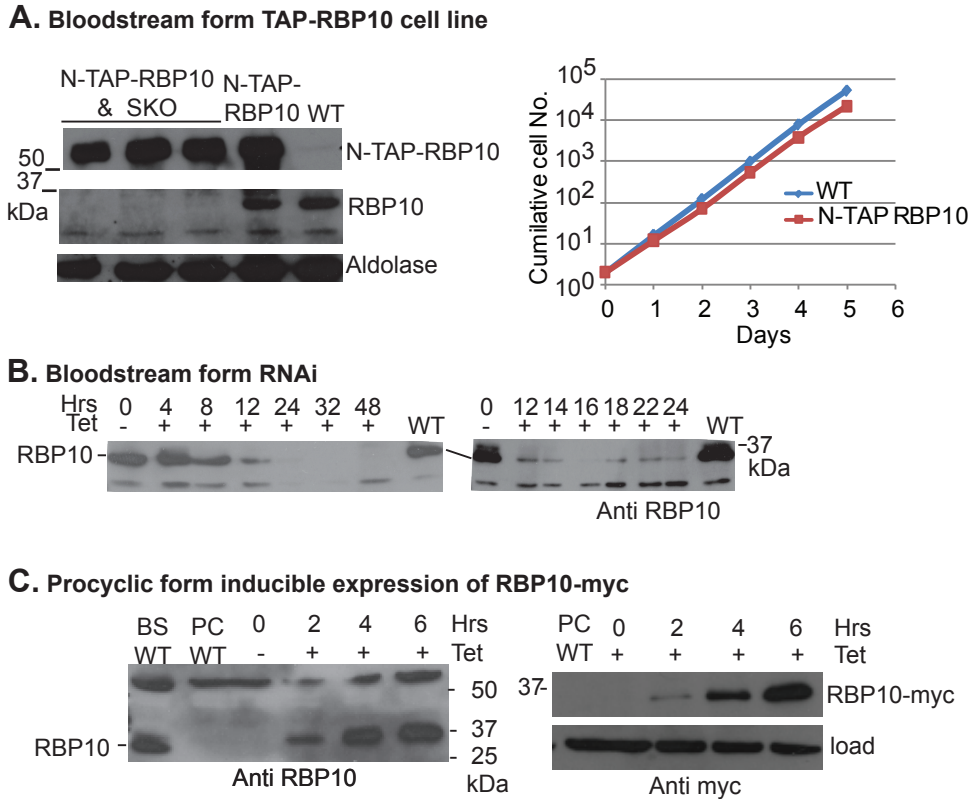
1138 Sheet 5: Set of bound RNAs used for the motif search. Filters were (a) at least 10
1139 reads in all samples and (b) a bound/unbound ratio of >3 in both experiments. The 3'-
1140 UTR lengths are those annotated in TritrypDB, except where noted. Only annotated
1141 3'-UTRs were used for the initial motif search.

1142 **S6 Table**

1143 Plasmids and oligonucleotides

1144 **S1 Fig**

1145 Alignment of RBP10 from different kinetoplastids
1146



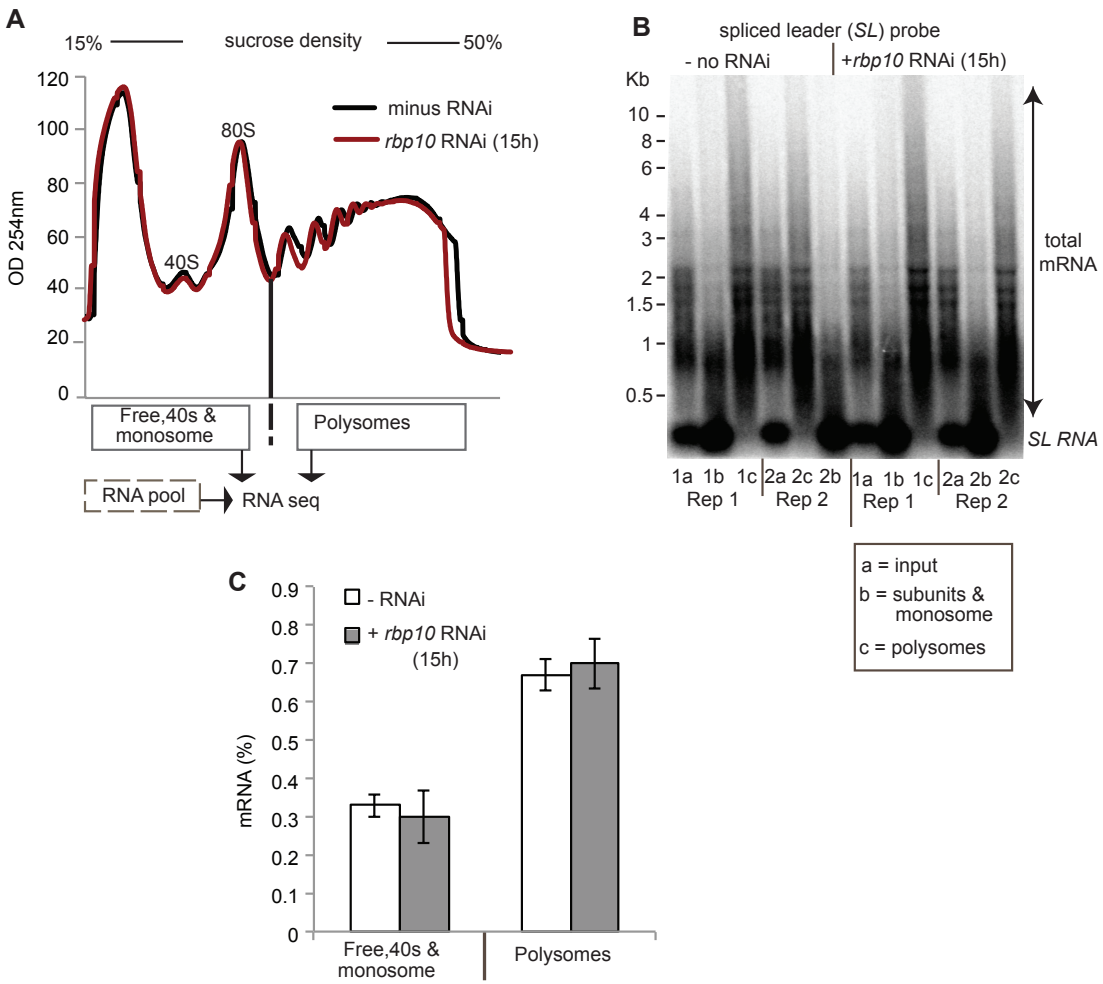
S2 Fig

Cell line characterisation and RBP10 expression

A) Characterisation of bloodstream-form trypanosomes with a sequence encoding an N-terminal tandem affinity purification tag integrated in frame with the *RBP10* coding region. The left hand panel is a western blot showing the presence or absence of TAP-RBP10 and native RBP10, with genotypes above the lanes. The right-hand panel is a cumulative growth plot for the cell line used in the interactome experiments.

B) RNAi cell line: Western blots showing the time course of RBP10 decrease after tetracycline addition

C) Western blots showing the time course of expression of RBP10-myc in Lister 427 procyclic forms after tetracycline addition.



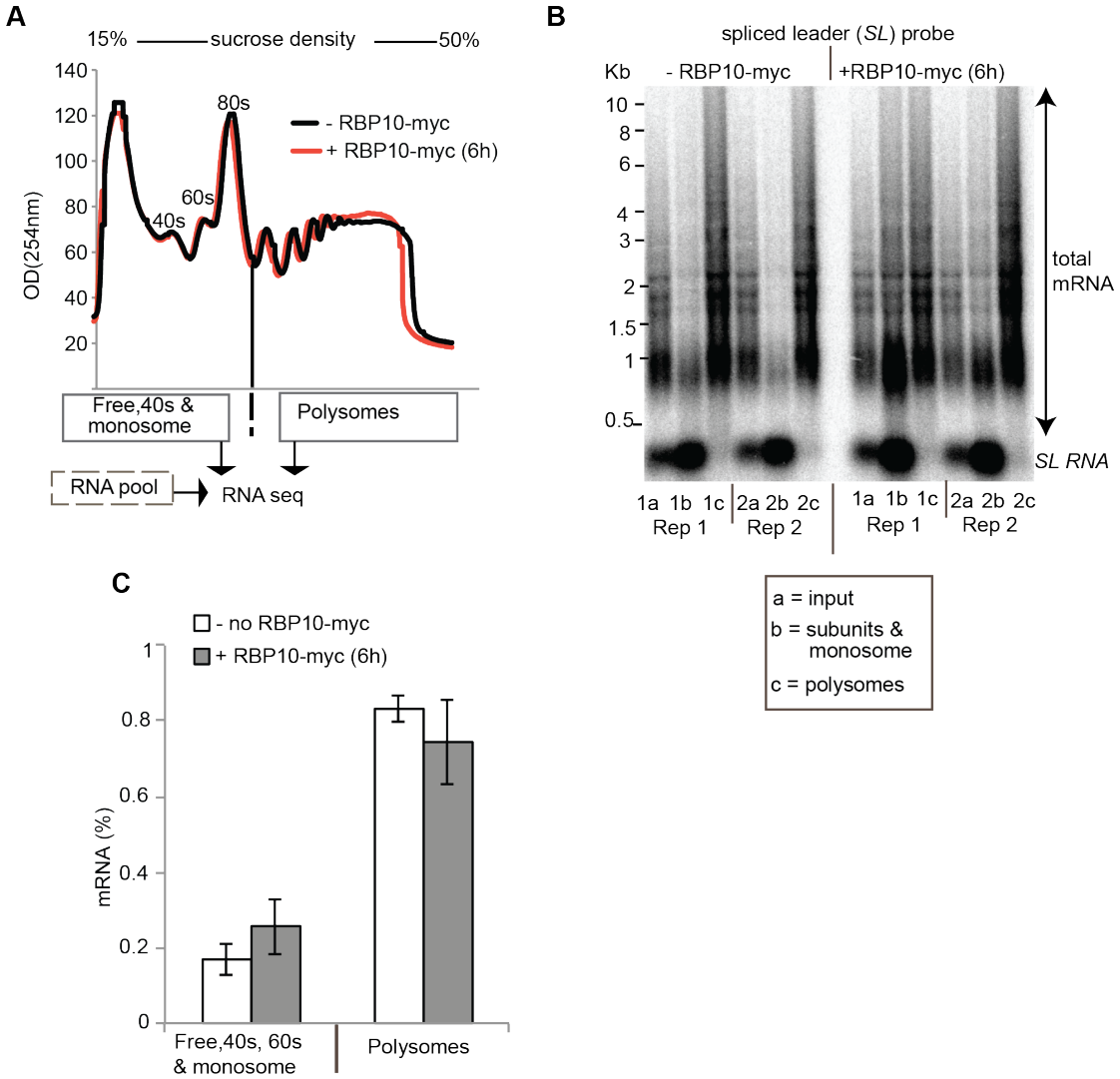
S3 Fig

RNAi targeting *RBP10* in bloodstream forms: samples used for RNASeq

A. Typical Sucrose gradient profiles of samples

B. Total mRNA in the sucrose gradient fractions, detected using the spliced leader probe

C. Average quantitation of the spliced leader signal



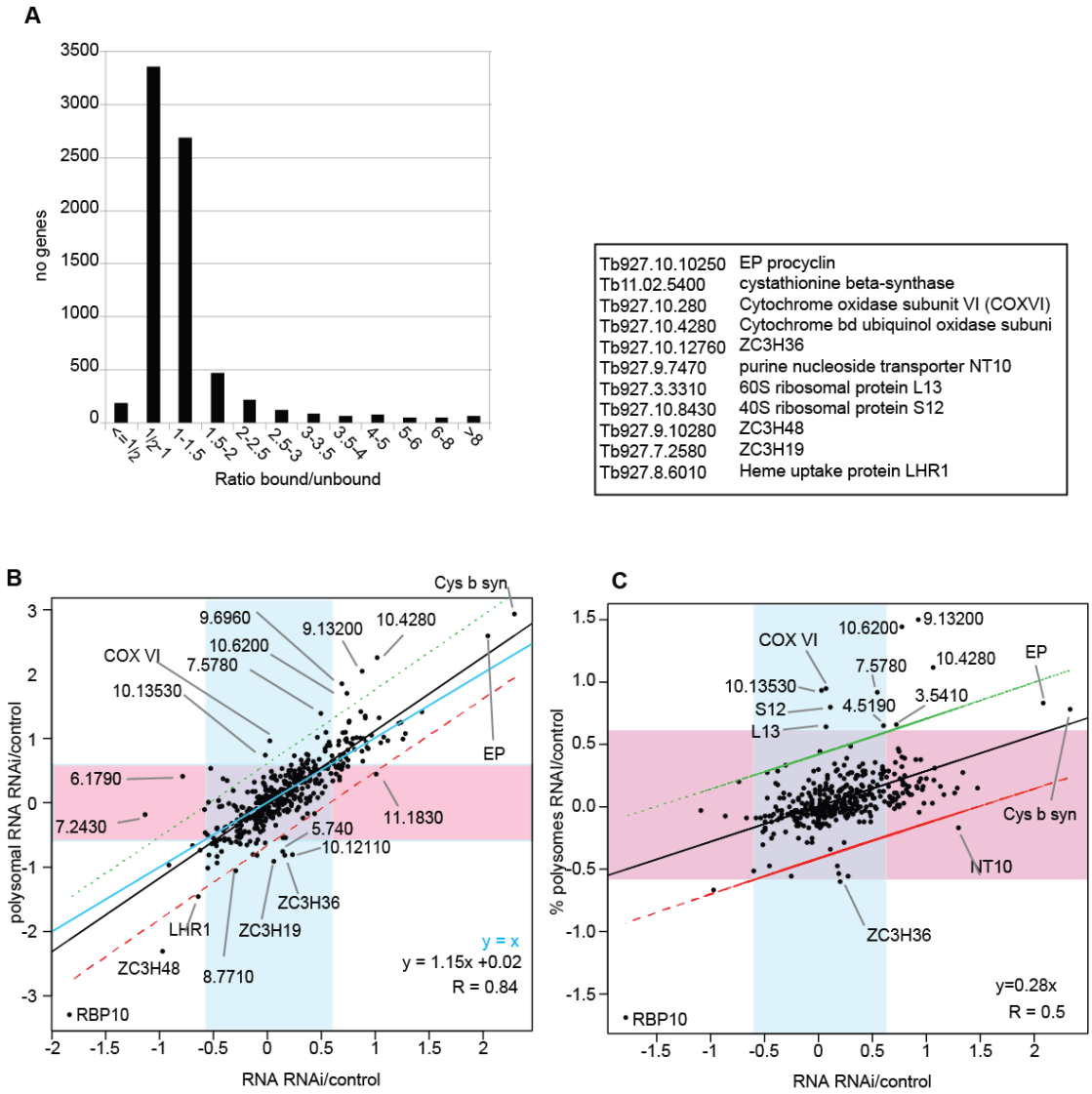
S4 Fig

Expression of RBP10 in procyclic forms: samples used for RNASeq

A. Typical Sucrose gradient profiles of samples

B. Total mRNA in the sucrose gradient fractions, detected using the spliced leader probe,

C. Average quantitation of the spliced leader signal



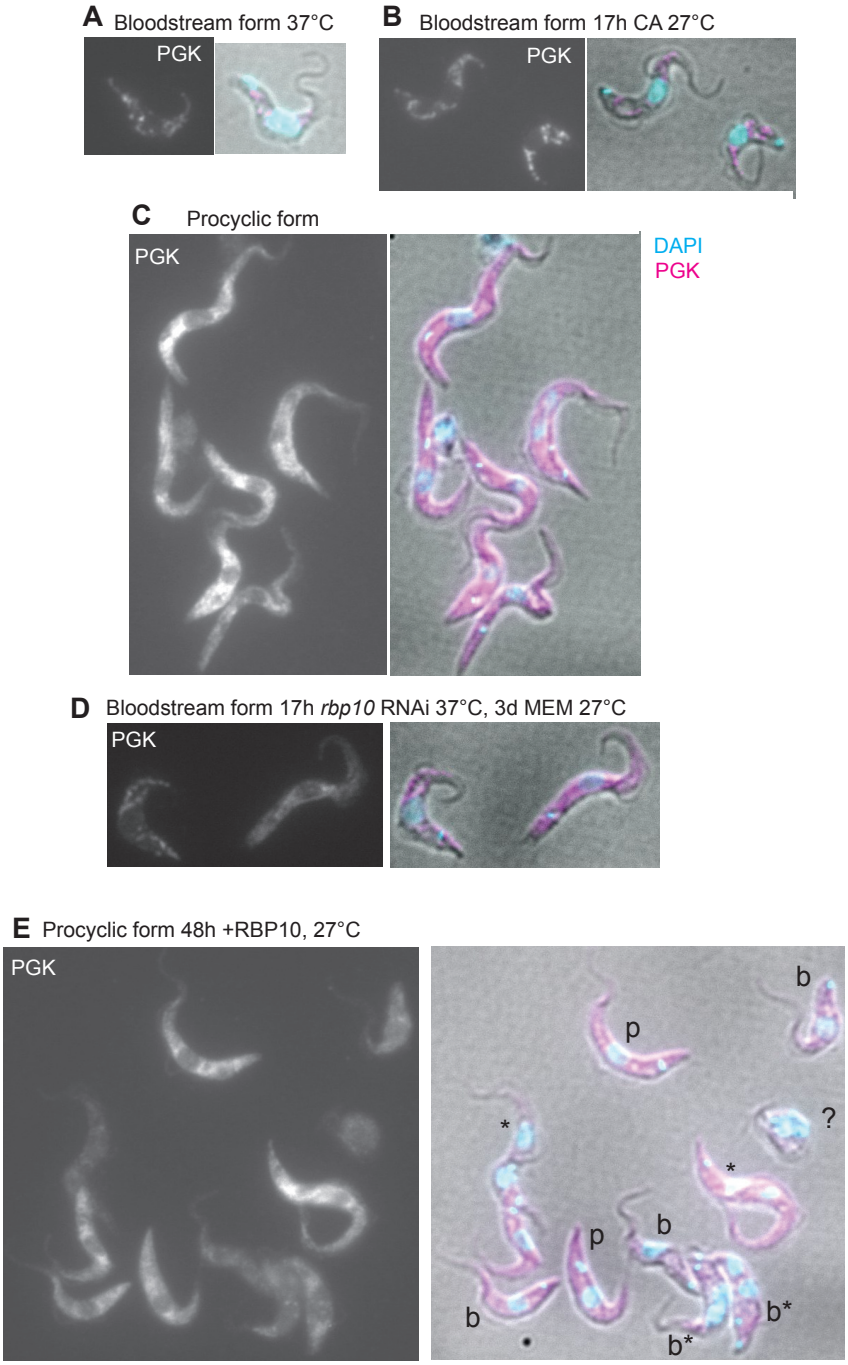
S5 Fig

Effects of RNAi and analysis of RBP10-bound RNAs.

A. Analysis of binding of individual mRNAs to RBP10. The classes are ratios of bound/unbound; the number of different open reading frames in each class is on the y axis.

B. Scatter plot comparing the effect of *rbp10* RNAi on total RNA (x axis) with the effect on polysomal RNA (y axis) for mRNAs that were at least 3x enriched according to DeSeq; all P-values were less than 8E-5. The black line is the regression line and the red and green lines show the 95% confidence limits for the data. The blue shadow encloses total mRNAs that were less than 1.5x affected, and the pink shadow encloses polysomal RNAs that were less than 1.5x affected. The cyan line is perfect correlation. The box beneath the graph lists relevant TritrypDB accession numbers. The gene numbers on the plot are also accession numbers, with "Tb927." removed.

1193 C. As (B), but here the y axis shows the effect of RNAi on the percentage of the
1194 mRNA in polysomes.



1195

1196 **S6 Fig**

1197 Relocation of phosphoglycerate kinase during differentiation.

1198 Cells were fixed with fomaldehyde, pernmeabilized with triton x-100, and stained
1199 using a polyclonal antibody to phosphglycerate kinase (PGK). The grey-scale panels
1200 show PGK alone and the differential interference contrast panels show DNA in cyan
1201 and PGK in magenta.

1202 A. Bloodstream forms

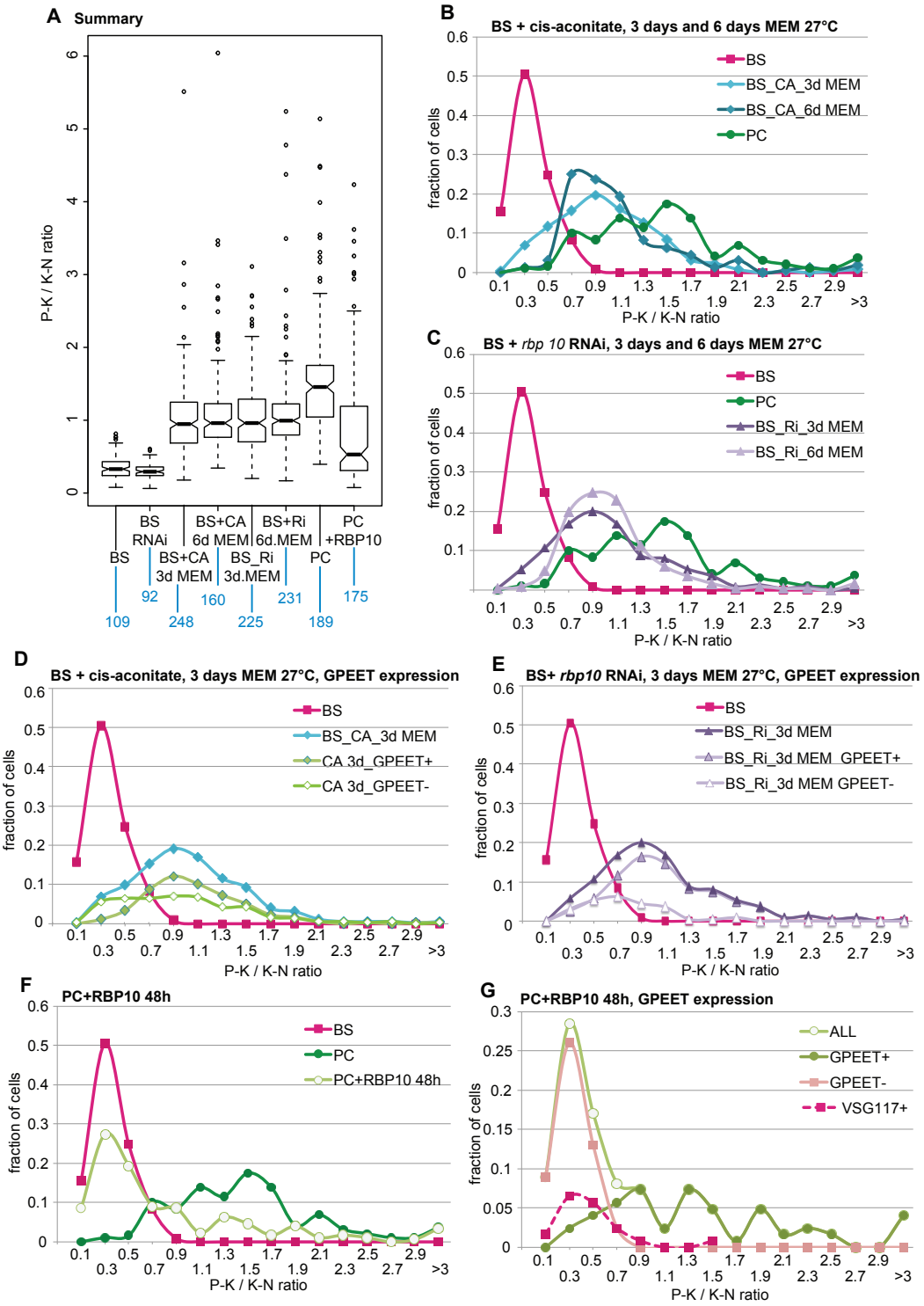
1203 B. Bloodstream forms after incubation with cis aconitate for 17h at 27°C.

1204 C. Procyclic forms.

1205 D. Bloodstream forms with 17h *rbp10* RNAi followed by culture for 3 days under
1206 procyclic-form culture conditions. The selection shows on procyclic-like trypanosome
1207 (left) and one which still has bloodstream-form morphology (right).

1208 E. Procyclic forms after 2 days induction of expression of RBP10-myc. Cells with
1209 terminal kinetoplasts are labelled "b" and cells with more procyclic morphology are
1210 labelled "p". Cells that appear to be dividing are indicated with asterisks. A very
1211 abnormal cell is indicated with "?". Cells if unclear status are not labelled.

1212
1213



S7 Fig

Morphological analysis of differentiation.

Cells were treated as indicated, fixed, and stained for DNA (cyan), EP-procyclin, phopho-GPEET procyclin, or VSG. (magenta). The posterior-kinetoplast and kinetoplast-nucleus distances were measured in Image J and the ratios calculated.

1220 A. Summary of all results. This is the same as Fig 5H, but all of the outliers are
1221 included, and the numbers of cells analysed are below each lane. BS - normal
1222 bloodstream forms; PC- normal procyclic forms; CA - 17h cis-aconitate pretreatment;
1223 RNAi or Ri; 17h *rbp10* induction; +RBP10 - induced expression or RBP10 for 2 days.

1224 B. Cell distributions after cis-aconitate-induced differentiation. The x axis shows the
1225 P-K/K-N ration, and y-axis shows the percentage of cells with that ratio. Controls are
1226 bloodstream forms (magenta) and established procyclic forms (green). The other
1227 lines are for 3 and 6 days after transfer to procyclic medium (MEM).

1228 C. As B, but with transfer after 17h *rbp10* RNAi.

1229 D. Phospho-GPEET expression 3 days after cis-aconitate-stimulated differentiation.
1230 The cells indicated with the cyan line (BS_CA_3d MEM) were subdivided into
1231 GPEET positive and GPEET negative.

1232 E. Phospho-GPEET expression 3 days after RNAi-stimulated differentiation; other
1233 details as in D.

1234 F. Cell distributions in procyclic forms after induced expression of RBP10.

1235 G. As (F), but also showing staining with anti-phospho-GPEET and anti-VSG117.
1236 Cells wit no GPEET have bloodstream-form kinetoplast positions, and a subset of
1237 them cross-reacts with anti-VSG117 antibodies. The remainder may express VSGs
1238 that do not cross react with the anti-117 antibody.
1239