

Figure S1: Schematic drawn to scale of the T7-stl vector for conditional dsRNA expression in *T. brucei*. pT7-stl is a derivative of pLEW100 (Wirtz *et al*, 1999) with two gene cassettes arranged head-to head. To the left, a T7 promoter (black arrow) drives the expression of the bleomycin resistance gene (ble^r, green arrow) which is bounded by the actin A gene flanks providing *trans* splicing and polyadenylation signals (act 5' and act 3', minor gray boxes). The downstream ribosomal sequence (rib. spacer, blue) facilitates targeting of the construct into the transcriptionally silent spacer region of an *RRNA* locus. To the right, a T7 promoter under the control of two tetracycline operators (2x Tet op., purple) drives the expression of a stem-loop RNA whose coding region (red) can be introduced into pT7-stl analogously to the established stem-loop cloning strategy (Shi *et al*, 2000). The stem-loop RNA coding region is followed by two T7 transcription terminators (T7 trm.) and an aldolase gene 3' flank (ald 3').

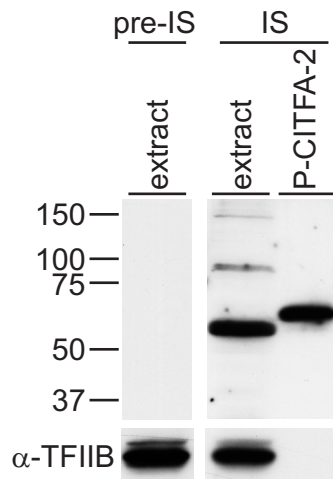


Figure S2: Immunoblots of procyclic extract were probed with anti-CITFA-2 pre-immune serum (pre-IS) and immune serum (IS). In the latter, purified CITFA-2 (P-CITFA-2) was co-analyzed. Detection of TFIIB served as a loading control.

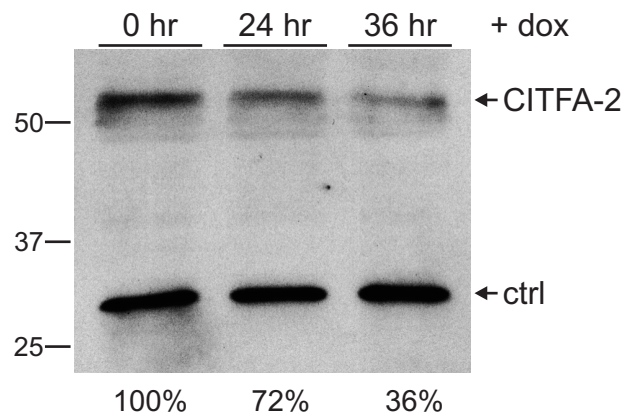


Figure S3: Immunoblot analysis of *CITFA-2* silencing in procyclic trypanosomes. Whole lysates of 10^6 cells harvested before and 24 and 36 hr after induction of *CITFA-2* dsRNA synthesis were separated on a 12% SDS/polyacrylamide gel, immunoblotted and detected with anti-CITFA-2 serum. The non-specific protein band of 30 kDa served as a loading control. The relative CITFA-2 signal strength normalized with the control signal is given below each lane. In contrast to bloodstream RNAi cells, the pZJM-transfected procyclics grew poorly and the silencing of CITFA-2 was less effective.

Tb	MPEVGTQVYWHDFEDAPPPWKNEEELAKMIEISSSIEGPI SDARH-----	45	
Tv	MPDVTTOVYWHDI PGAQAPWODEOEYARMWVI SSKIKGLVSDRSH-----	45	
Tc	MPEFGTOVYWHNI PACPPPWKDEEERARTWRI SORIVCAISDORH-----	45	
Lm	MPIKETOVYWHDL PTPRPPWRNVEELKOCVLESKRIAAQMARQWRSVCGSVPALQSSIP	60	
Li	MPAKETQVYWHDL PAPRPPWRNAEELKOCVLESKRIAAQMARQWRSVCGAVPALQSSIP	60	
Tb	MI VOPREVIYI SMRTKLRNRQOPCNR----YWSIVLEPHFGPI I LGNMVMREIKVIEMVTRE	101	
Tv	LILOPLRKEVLOMRARIRRKPPPSDR----FWGIVLEPYLSPLILGLNMIMKELOVPGVTYE	101	
Tc	LIIEPPRAEILAKRILPGRFAPKRVHR----YWGIVCEPHLSPLHGLNMVMSELOVRPIATE	101	
Lm	RALPSRAEQVSRAERLRDRATRHIERPVLFRRVVPEPHATPLOVNLIMEHLRORPVLTA	120	
Li	RALPSRAEQVSRAERLRDRATRHIERPVLFRRVVPEPHATPLOVNLIMEHLRORPVLTA	120	
Tb	EAAEAVNGVVEEFNAMSOLYSATDGAND-----CLG-----TALSGKS-----	139	
Tv	EAAAIVSVVEEYNI LTOQFGTOOGSQW-----SEYEANVTQRSTVSRRNA-----	147	
Tc	EAKEAVSVVCELYNAMTOQQ--OKSTQS-----GILSQGGVKTEILTOOSN-----	145	
Lm	EAAEVAERVTLHFNMLTSSQHRREGGASRPRCNRGACPRAGCVATESVLSOVHQDDA	180	
Li	EAAEVAERVTLHFNMLTSSQHRREGGASRPRCNRGACPRGGGTATESVLSQVQEQDDT	180	
Tb	-----AQAMDVKVAPFLEVVRKLCPEPRNLDVVEWNRDELRRRGRI	179	
Tv	-----TPAMDVKEAPLLEVVKILCADRKLIVTPLSKEELRRKGR	187	
Tc	-----TAAMDIREAPLLOVVKILCASRGLVSPFSKAELRRSKGR	185	
Lm	RLGLTSSQATIGTSGGTAGSLTMALOSLLLDITVAEFCAARGLVFRRLSRADLRGMRHRV	240	
Li	RLGLTSSQATIG---GTAGSLSMALQDSLLLDITVAEFCAARELVFRRLSRADLRDMHRV	237	
Tb	AFTDLLYHKRVVMI VDTSKPHFLICLVTHKKRLT-----DKR- EV	219	
Tv	SFSDLVSHKRVVIAFDRSKPHFAFALGVVKKKAE-----AVG-QV	227	
Tc	GVADLLPRRRVIMAFDRSKPKFVFATAVPAKKKKD-----AEKSTV	226	
Lm	SSVEVVPYRRVVVLADRHDPIFAVLFVGSROVTPPEPFVVSALQLDGKTPESIARYRKAFA	300	
Li	SSVEAVPYRRVVVLADRHDPIFAVLFVGSROVTPPEPFVVSALQLDGKTPESIARYRKTFA	297	
Tb	YCMOLLIAPTASVVPWRRITKYALVSPRAREFVITITTKRDSVK-----NSNLSIDVIGAA	274	
Tv	RCMOLVAPTANPIPWRRIVKVGMERLSANDCVVVVVORQDD-----CSQKGLGLOS	277	
Tc	CYIOLVAPNSNPVWROLKREIIPYLPNECVVAVORHHHDEEYDDDDEKNSGDGPYL	286	
Lm	PVTOVFCMAPAPVWRRITITNISVSTLRLDEFVVAVIELDEFAEE--TADADRNGDOVSTHR	358	
Li	PVTQVFCMAPAPVWRRITITNISVSTLRLDEFVVAVIERDEDEE--TAGADKNGDOVLTNR	355	
Tb	PEAKSASOKDLEDGEDGDYDDEEAEDEYDNDDDDDDEDDGDEDENSNDGDNRPKRARVA	334	
Tv	DHENEDDDVQIND-----GEGGEPIPRIPHKR--PREEGNERDGESGRSVAMSESS	326	
Tc	SDANVRSEVLLGGNESDKYDLGEEGDAOPROHEEGETPOKRHKCEYEDKIDFSDMSEVS	346	
Lm	SCAKRGRTGVTGATALPRKMORLPVDILGGDEEEEGEMGLVSDAGEDGDAEAAILNPA	418	
Li	SCAKRGRTAVTSAAARPRKMORQLVDILGGDEEEEGEMGLVSDAGEDGDAEAAILNPA	415	
Tb	DSVDNSETDLSLNNEDYPFLDEIDAYERMRTGSAPIFSAEDRVLD-----	379	
Tv	ESVSEGGDEERNY-----DKIR---FCRNPIFTEKDYIQD-----	358	
Tc	ESIGSTTASCRSD-----EAKRKKFVGREPIFTAEEKIQN-----	381	
Lm	ATPTVASSSATER-----VPLERRWLGSAPLFSGTTQLAETAPGEVAVEMVDPTR	468	
Li	ATPTVVSSSATER-----VPLERRWLGSAPLFSGTTQLAETAPGEVAVELAGPTR	465	
Tb	---DPLFNTPVATYRRASRATREGRDWVYRFTHONDRNKLFAITN	421	100%/100%
Tv	---DPLFRTPVAILRKNKRVAGEVHDWIORLVRNKERNKLFAINI	400	38.1%/55.3%
Tc	---DPLFETPVAVLRKLRKTRGVHDWIORLINPVARKRLYAINY	423	33.4%/50.1%
Lm	GAYDPLINLPRSEG-AARRRGSLORAWYRRLIPKGSDDFMARMHTYQADE	517	22.9%/37.0%
Li	GAYDPLINLPRSEG-AARRRGSLORAWYRRLIPKGSDDFMARMHTYQANE	514	23.4%/35.1%

Figure S4. Sequence alignment of trypanosomatid CITFA-2 orthologues. ClustalW alignment of CITFA-2 sequences from *T. brucei* (Tb, GeneDB accession number Tb09.211.3440), *Trypanosoma vivax* (Tv, tviv913g02.p1k_2), *Trypanosoma cruzi* (Tc, Tc00.1047053510741.100), *Leishmania major* (Lm, LmjF35.3150), and *Leishmania infantum* (Li, LinJ35.3250). Identities and similarities are shaded in black and gray, respectively. Only positions with a minimum of three identical or conserved residues are shaded. Identity / similarity values specified at the end of each sequence were determined by pair-wise comparison with the *T. brucei* sequence using the EMBOSS program (<http://www.ebi.ac.uk/emboss/align/>) at default settings. *T. cruzi* has a nearly identical CITFA-2 paralogue (Tc00.1047053510661.150) which was not included in this analysis.

Tb	MSN-----EVESSAAEPPAPDEATNVAG----HRVILINNILCRTEETSD	40	
Tv	MNDRDANNVEVVIIVDDASTTSSPTKQASNKRGSAVNVHFLILINGILCRTEAPD	53	
Tc	MMALDDMVAPFREEFTDRSCRDERNRHRFFFLINGILCRTEHNSD	43	
Lm			
Li	MSRQPSRSSSPGTALPSEATVTTASQLPSYPSASRSCQKVDEPGAFLLLNGFVCRTADAS	60	
Tb	GSFRPSDVVFGLD-PPPRGTATLGSAAEDNCNHLVSGEFDPTAYVDSISKDCSIFNVFGV	99	
Tv	ASFRPSDVVFGLD-PPPRGTTTLGSASVDNQN--SIYQFHS----ESVSRDCSIFNFTFGV	106	
Tc	GSFRPADVVFGRD-PPPRGTAMLGEASRIFDDSAVAPSFAD-----VDRDCSIFNFTFGV	96	
Lm	MPSANAHFSLSAGSAEETR----RSIDSDCPLFNCFGV	34	
Li	GRFRPVEVFSGVYREAATAPPTMPSANAHFSLSAGSAEETR----RSIDSDCPLFNCFGV	116	
Tb	MSCGVLVDSESRGFIDFTDDMIFRPPLPPPVE-----GNDTHNEGSRD	142	
Tv	LSGGIIVDPQTLRFMDFTDDMI FOPPTFGDDA-----ENCWHYDNL SY	149	
Tc	ESGGVLIDPNTRCFVDFVDDIVVTPTLCKE EAN-----GNVPHYKSLEY	140	
Lm	LSGGVIIDPIDGSMFEFTDAAVYTPSTVAAQVDITGNDTQDDLARCTLQOANEHAFRTGAT	94	
Li	LSGGVIIDPIDGSMFEFTDAAVYTPSTVAAQVDITGNDTQDDLARCTLQOANEQAFRTGEA	176	
Tb	GQIRADGVVSLVSSIFPFNGLGTV-----DTPLPWP-RSLPRLRSORISYRLAAKT	192	
Tv	WKIRANAMASLMSTIHPMNAPGTE-----ESPLPWP-RSMPRLRCSRVANSLAROS	199	
Tc	AKIRGEGVMCLCSSIEPLNVACKR-----DTPLSWP-RSM-RLTCSAEARFLALRN	189	
Lm	HLIRPECPILILCTSELPLPPRSDDRPLNRFKCRSSPLSAAGSTGPTVTSARHAAQMGMEN	154	
Li	HLIRPECPILILCTSELPLPPRSSDPPLKRFKCCSSPFSAAAGAGPTVTSARHAAQMGMEN	236	
Tb	VHRMLSRCVGNVKCNK-----AVLVVLESYRIHRKHLHSPQLT----FLPNEIPFR	238	
Tv	VDRMGARLAVNVVFDPE-----TKLGLVEEFRHHLKLTFFPQIT----FLPGEIPYR	245	
Tc	VDRILRRCLTEVSANR-----KMFRLVEGYRCHIKSKRPPOS----FMTQNIPTI	235	
Lm	VEKWLNLAPHPGHRPQSPRRQAWTTVWADAEEFRLSMRAHSDVRASEASLYLPPRIPTFP	214	
Li	VEKWLKRLAPHPGHRPQSPRRQAWTTVWADAEEFRLSMRAHSDVRASEASLYLPPRIPTFP	296	
Tb	FOWHGMTFR--VKDGGAPKRFDANGNRTHEYRSPIGRINHMDGKTG-SAYSSFHAFATG	294	
Tv	FOHGCRTLRSWAENAGHCNNTLAVVEIADSNGRRPVDSEARAAPVRIITSTSPFSEFHITFG	305	
Tc	FOYGGRPFL-----LQDDSDKEKENRS-----FPTPGPLFSEFHAFATG	273	
Lm	LRFRGQPISLIEGANESPAASQOQQPDREANSATAVPPDGVRCAAASFTCYTGVPRFAT	274	
Li	LRFRGQPISLMEEKTNESPTASQOQQPDREANSATAVPPDGVRCAAASFTCYTGVPRFAS	356	
Tb	----LPKMAFRRYRESARRSRHLENLADGYSRRLERMRIQNIIVHPIREEMHIDENIEVV	350	
Tv	----LPKMIFRRYREAIKRDLSGLVKKSLSYAQRIEMRVONVHFPIRKNLQISDRITAFG	361	
Tc	----LPKMVRMKTGEVDTYN-----HGVEFGLRLEKRLQSMKPLRKNLFIGRRVLRD	322	
Lm	ATDTAPRFDSAEFAKGTAVARLRKRLRPYCETSSLGRRRSSSDYDEADREGFLIPFTDST	334	
Li	TTDTAPRFDLAEFAKGTAVARLRKRLRPYCETASLGRRRASDDYDEEDKERFLIPFTDST	416	
Tb	FRRRSGVHLWSAFORYFTILG-----SRCTSCNTPNCNAWYLDNDDKVDYPSALR	401	
Tv	KRRKGYKRHWNIHRYFTLLR-----AKPCLTCGNPNCNVWHFGADDDVDYVEKVA	412	
Tc	KSMTRMRKYWYTFWRYFKIFR-----VYPAPSPNADGRSEWHLADDRHVNYEKAVL	373	
Lm	PRHAAFALFLLYFNHLGVLCNGVVRTGETAASSCLDGLDVGSTTWGLCPKYLYDVRAEVE	394	
Li	PRHAAFALFLLYFNHLGVLCNGVVRTGETAASSCLDGLDVGSTTWGLCPKYLYDVRAEVE	476	
Tb	KLEEDLKYRR-----TFNLSSSILADSWDWLLAKERGPITERIRKTIKYAR	447	
Tv	AEEEEAKKIHR-----SDKVNSAVVSDAWDWLFLNEKEELOQRIGKIVRAIK	458	
Tc	EFEERARKYG-----NSEWDRITAVVDAWDWLVHERHSFLRRLAKVERCLD	419	
Lm	RKRRVTRELREHPERTFAGGTRWSTTARNRDAAAATAALLISSRTALLERITELNRHIH	454	
Li	RKRRVTRELREHPERTFAGGTRWSTTARNRDAAAATAALLISSRTALLERITELNRHIH	536	
Tb	QIDAENKREVINTDKDLR	465	100%/100%
Tv	AIKAQRKK--LITSEC	471	42.2%/59.0%
Tc	KIQL	423	35.3%/48.2%
Lm	KLHDEKQRRRHADDQENQ	472	20.0%/32.0%
Li	KLHDEKQRRRHADDQENQ	554	23.5%/37.1%

Figure S5. ClustalW sequence alignment of trypanosomatid CITFA-1 orthologues corresponding to Figure S4. GeneDB accession numbers of the sequences are Tb11.47.0010 (Tb), tviv163a09.p1k_0 (Tv), Tc00.1047053508505.30 (Tc), LmjF27.0830 (Lm), and LinJ31.0920 (Li). The *T. cruzi* paralogue (Tc00.1047053503893.160) is nearly identical to the TcCITFA-1 analyzed.

Tb	MSHRISSSTNNPFSERDFN----LLKSGALEKSRSVVSRREQ-VWVNDIVDVGEGVKANDVE	55
Tv	MSRRMSSSYRDTYSERDYEERRHRLRSGAIVTSRSRITGTEV-ODEKIVNVGEGVKANIYE	59
Tc	MSYWATSEGRALSAGHTSRLSWCMRSGAMATSRSAMTOTELELRDRIVDVGEGVIANFEE	59
Lm	MPSEADLDRLRCR-----LHOSGATVSSRSVASRADANFDLEVVNVGDGRRATPVE	51
Li	MPSEADLDRLRSR-----LHOSGATVSSRSVASRADANFDLEVVNVGDGRRATPVE	51
Tb	ELMARLEOHGLRRVLAE-----NEGGTGALHAADGTHGTENCDSITIEDIDEDDSC	105
Tv	OLLARLEOHGLRRCFVEDADEANLEAREESCD---DATIVDDDGRLSLISGWTYGNPCTS	116
Tc	OLLARLEOHGIRRCILHD--DKETIYEDGTNNGDKNLHHSISLSDFCSMASSLRPSVKORME	117
Lm	HLLARLEOHGLRRRLVAEDDAAAYAVDAAARADTEENADDDGDEMYASEEGNTEDPRAVS	111
Li	HLLARLEOHGLRRRLVAEDDAAAYAVDVAARADTEENADDDSDETHASGDGNAEDPRAVS	111
Tb	-----EKRAPYYYYMCVPCELSLTEISTKPATMRMLODAHYHFSSAEH	147
Tv	-----ESRSPYYYYMCVPCEMKLTATCTTAATLRMLODAHYHFSSAKH	158
Tc	-----DAREPYYYYVCPCELOLTETSTSAATLAMLQDVHYHCSSAEH	159
Lm	SCSSGDGAFSSPAPSKAPEETREAYYYLCVPCELRLTRISRRPWRRALEDVHFHFSSAAH	171
Li	SCSSGDGAFSSPAPSOAPEETREAYYYLCVPCELRLTRISRRPWRRALEDVHFHFSSAAH	171
Tb	RCIASWVGVDIEKTLAVTSRLEVGG-YIRIFVNGIPFLISARPGGGGMFYPLPHEITVEA	206
Tv	RCEASWVGQDIDRITLNNSAEVDVNG-YMRIYVNGIPMLLPRRPGGGDMFYPLPHEEFDY	217
Tc	RRIASWVGEPDIDRITLONSAKIDPTG-YAWIHVNGIPMLLPRRPGGGDMFFPLPHEAHD-	217
Lm	RATASWMADDDDIDETLHSTPLVTPPTHYYSRIYVNGVPTLLSRRPGGGDMFYPLPHEQDLV	231
Li	RATASWMADDDDIDETLHSTPLITPTNYYSRIYVNGVPTLLSRRPGGGDMFYPLPHEQDLV	231
Tb	REOSGRNAEP--SKGTAV-----RKEITWYRPIOS	233
Tv	RDRATSVGRDGHLLHTIMR-----RKEYIYWRPIOS	246
Tc	QERIAPSGMDGLLFATQQ-----HTEVWNRPLRS	246
Lm	APPSAAAARDGNRVDCLHGSRTVSNSSSEASGAPAGRFFPPSTLRGAQGGQPVLWHRALPS	291
Li	APYSAAAARDGERVDCWHGSQTVSNSSGASGALAGRFFPSSTLRGAQGGAPVLWHRALPS	291
Tb	VFTRCKOILYTAASR--QKLRGESSPFRAORRNGIVLL---HIPMKTYERESFIRGPSVP	288
Tv	IFTGTKOLVYSADSR--VRLRTKKTPLRFORRNGVVIL---HVPMDTYAKESFVDGPTVP	301
Tc	IYTGTKOFLYSVDNN--RRICKKKWTKLKRNRNFVLO---HIPLDVYTTIGTFVGDATVP	301
Lm	LCTRTVQVVRRTRSSSLTDLPADRRRYRVARTRSRVMVGVGDHCSLDEYRQOSWMLPHKVP	351
Li	LYTRTVQVVRRTARSSCLADLPADRRRYRVARRRSRVMVGVGDHCSLDEYRQOSWMLPHKVP	351
Tb	---ALFEVQRERVSNNKRR-HKENG-----TRTCYVMREKP	320
Tv	---ALFOVDRKRMKPKVGKN-ETSDG-----YRVAYVASKKP	333
Tc	---ALFEITTOORLPKKAQYGAETT-----YRMKYVVS DKP	334
Lm	LSLAVYRLRRRERVPKPLLARPESAGPQGQAVPLRFSESAVRAAQAADVAHRTVYVYTPYHP	411
Li	LWLAVYRLRRRERVPKPLLARPESAGPHGQAVPLRCSESAVRAAQAADVAHRTVYVYTPYHP	411
Tb	IAAWED-----EDDSTDAISK-VITVKEEGVYRVLVILCSDDVRRSMRQAEHEHEPE----	370
Tv	MMVWNE-----KDARPTALNR-ITVSDGDIYRVALIHA-DVAKALLSGEP-VEEPD----	381
Tc	CMVWN-----DAMEGLSS-FVVRNNGIYRVSLICAGDAAKWASQGPPPAGSTQ----	382
Lm	IAEGSRDYHRQGDAFATAAVPVPTTVFEDECYRVALVSAAEQARREGLSHPEASRKVSHSL	471
Li	IAEGSRDYQRQGDAFATAAVPVPTTVFEDECYRVALVSAAEHARHEGLSHPVTSRPLSRSL	471
Tb	--TRVSGEVTHP--LITVEALQMTGGTIRRES-----NOSDGKSLSL-YDSS	410
Tv	--HGVIADNEPATITLLEATRRIGAISSVVG-----FGSRSGSPSDSRSDSP	426
Tc	--DAHTGCEGHISQMLTVEALIQLRRASTVGRN-----FHVLSTSELDA-SDSL	427
Lm	FPADVASEEIIHRPVLTLELLMOHTQSTOSWRPAFGLNTALDSSFVPSAPPSPSRSSI	531
Li	CPADVASEEINRPVLTLELLMOHTQSTAQSWRPAVGSSTALNSSFVPTAPPSPSCSSI	531
Tb	SWSRKT	416 100%/100%
Tv	SDSRSSSFIIIDSVADSSNNSIQLDDI	453 43.4%/59.8%
Tc	IHSEVTSYVSS	438 38.4%/51.8%
Lm	SHLTRSSSSGSTSSSTSSEHSSTSRSSSTHTRKRVKLDGCL	571 29.5%/40.2%
Li	SHLTRSSSSGSTSSSPSFEHSSTSRSLTHTRKRVKRDGCL	571 28.8%/39.8%

Figure S6. ClustalW sequence alignment of trypanosomatid CITFA-3 orthologues. GeneDB accession numbers of the sequences are Tb11.47.0008 (Tb), tviv163a09.p1k_2 (Tv), Tc00.1047053508505.10 (Tc), LmjF27.0850 (Lm), and LinJ27.0440 (Li). The *T. cruzi* gene Tc00.1047053503893.181 is at the end of a contig and predicted to encode the identical N-terminal 128 amino acid sequence of the TcCITFA-3 analyzed suggesting the existence of a paralogous gene.

Tb	MSTLASALPLLA	TKNVL	CGVTG	STIOFF	CDLTRDY	CP	ST	TKKSV	IASS	CGNR	P	GT	58
Tv	MPIPTTAVPL	LLMSK	NVM	IDVS	EETILL	VYCDI	HONSG	QSSTGRS	II	IATS	G	G	58
Tc	MESSTLANA	IPFC	GKNI	LVALE	ESNLY	IYCDLS	ODLGR	TSSGKN	I	LIAT	T	GG	60
Lm	MSE	RC	DM	K	AK	N	L	R	M	V	V	O	55
Li	MSE	RC	DM	K	AK	N	L	R	M	V	V	O	55
Tb	GAHIVLN	VFFAK	E	KPOL	DED	TLAP	L	RTR--	E	V	E	G	115
Tv	GSYMCLN	L	FC	H	S	F	SS	R	L	D	E	A	115
Tc	NAFLGLN	VFC	N	IE	G	PKL	T	D	E	T	I	S	116
Lm	GAFLGLN	I	F	T	K	S	L	D	K	R	D	L	115
Li	GAFLGLN	V	F	I	K	S	L	D	K	R	D	L	115
Tb	KKV	G	K	R	G	T	V	L	A	S	R	C	174
Tv	RKADI	T	G	A	I	L	L	A	S	S	G	N	174
Tc	K	L	A	T	T	C	N	S	V	V	L	A	175
Lm	R	L	G	A	S	G	K	S	M	L	L	A	175
Li	R	L	G	A	S	G	K	S	M	L	L	A	175
Tb	D	I	G	N	G	F	I	M	N	V	E	S	222
Tv	E	I	G	E	G	F	S	L	T	V	E	S	222
Tc	E	L	N	E	G	F	V	V	H	V	E	S	223
Lm	S	M	E	G	G	F	V	S	Y	T	P	T	235
Li	S	M	E	G	G	F	V	S	Y	T	P	T	235
Tb	R	C	G	V	K	R	N	A	A	Y	E	N	270
Tv	K	R	S	G	T	R	K	V	R	T	N	---	267
Tc	R	F	V	D	K	L	R	S	E	E	---	---	263
Lm	K	T	A	M	R	A	T	E	D	E	K	Y	295
Li	K	T	A	M	R	A	T	E	D	E	K	Y	295
Tb	-----	NTE	---	R	I	R	L	E	V	R	F	D	317
Tv	-----	NQE	---	V	I	V	D	I	R	F	D	P	314
Tc	-----	N	N	S	R	G	T	V	D	F	Y	L	312
Lm	E	A	A	V	S	A	G	S	O	T	C	A	355
Li	E	A	A	V	S	A	G	S	O	T	C	A	355
Tb	H	R	S	P	V	S	L	A	S	D	T	V	377
Tv	O	R	T	P	E	S	I	L	T	S	A	E	374
Tc	H	K	P	A	P	H	L	S	P	A	E	I	372
Lm	G	R	P	A	P	S	L	S	E	A	E	I	415
Li	G	R	S	A	P	P	L	S	E	A	K	I	415
Tb	A	I	I	S	H	F	H	R	R	A	F		388
Tv	A	V	I	K	Y	L	K	E	R	G	Y		385
Tc	A	V	K	A	Y	M	G	S	L	K	L	S	472
Lm	A	V	V	A	F	V	O	A	A	A	T		426
Li	A	V	V	A	F	V	O	A	A	A	T		426

Figure S7. ClustalW sequence alignment of trypanosomatid CITFA-4 orthologues. GeneDB accession numbers of the sequences are Tb11.01.0240 (Tb), tviv140b09.q1k_4 (Tv), Tc00.1047053508819.20 (Tc), LmjF28.0430 (Lm), and LinJ28.0320 (Li). The *T. cruzi* paralogue (Tc00.1047053509595.40) is nearly identical to the TcCITFA-4 analyzed.

Tb5a	1	MCLSKIFSPVDOSLKVLVEVM--TRLAGOEVVKGTRK--RGRSKECGTDEEFAHFVIRINGRV	58
Tb5b	1	MCLSKIFSPVDOSLKVLVEVM--TRLAGOEVVKGTRK--RGRSKECGTDEEFAHFVIRINGRV	58
Tco-1	1	MSLGRKRARSDDAYNVEVSGHPTRATYDSGVOOVHNPGRGRSTSPGVNGE--FFLIRVNGKS	59
Tco-2	1	MSLGRKRARSDDAYNVEVSGHPTRATYDSGVOOVHNPGRGRSTSPGVNGE--FFLIRVNGKS	59
Tc	47	FHTNEESMPLHSSTREEFDRLGHVITTPSALPHKMDVETVGLTDKNEDGSFFIHLNERP	107
Lm	58	HCYGDRPVLLDCTGALLORRQRORROELQSRASOLVAARADLLKTKNLWSMRRIRARR	118
Li	58	HCYGDRPVLLDCTGALLORRQRORROELQSRASOLVAARADLLRTKNLWSMRRIRARR	118
Tb5a		TOVPSAAPSINLROLLGOAG-----	78
Tb5b		TOVPSAAPSINLROLLGOAG-----	78
Tco-1		WRVDAAAPFINVTOLLDTFC-----	79
Tco-2		WRVDAAAPFINVTOLLDTFC-----	79
Tc		SRVAVTTPYINPROLFGHAG-----	127
Lm		QNKELSVIPVSAVGDAGNDADRDNEDSTARAPSGAAEGGDSTSOLLSELSRQRKAITHEMK	178
Li		QNKELSAVSAVGDAGGDAGRGSGDSTARAPSGAAEGGDSTSHLLSELSRQRKAITHEMK	178
Tb5a		-----VLLODITDEETP-----IILLPCPNTEGNVTVHPNHNYRLVKFVGEHMCLAG---	124
Tb5b		-----VLLODITDEETP-----IILLPCPNTEGNVTVHPNHNYRLVKFVGEHVCPEAG---	124
Tco-1		-----LILLQDVTVPSCP-----IVLLPCPRTGNVTVTPNREYKLVKFI GEHLCLAG---	125
Tco-2		-----LILLQDVTVPSCP-----IVLLPCPRTGNVTVTPNREYKLVKFI GEHLCLAG---	125
Tc		-----LILLQDVTVPRHP-----IILLPCPHTGNVTVHPRRRYRLMKFI GEHMCLAG---	173
Lm		EIRRELLTILLOEKRKQAAATRDGCVLMLPSPETGEVQLFPGHYRTVCFVGEDELREVLPSD	238
Li		EIRRELLTILLOEKRKQAAATRDGCVLMLPSPETGEVQLFPGHYRTVCFVGEDELREVLPSD	238
Tb5a		-----AMSHANNNSKTDHVAIAAKDGEA--E-KKSPLRSDIWAMAELW	164
Tb5b		-----DFFPAN--DHEIDAITFTAEGNKL--H-KKRVFOSKKLADSTLG	163
Tco-1		-----DISRNGAHGPESAE TNNGKSSNE--T-GKKPSHLDIWEMTELY	165
Tco-2		-----DLAEVNDVNIITIKETKE-----QOOSRFDLWAKTSLG	157
Tc		-----DFYPVLDEEYOQRONSSE-----AFRTRQRESWAMSELM	208
Lm		ARDTDVATAHHWQRRTPPLRKLASEGVAREVSEKDEGGAGGEHTSAKVGERPLWAMAELL	298
Li		A---DVATAHHWQHRTPPPLRKLASEGVARELSGNDEGGAEGEHTSAKMAERPLWAMAELL	295
Tb5a		AVIAFMOEHRAS-----CVODFLSLPVERMATMYSRGLVNLRRCCFSGVESFSSDEEVE	220
Tb5b		SLMVFLQSNKIS-----YARDFIMTKPEEVEIYK GKATAAMMRHCTKGVLEFSSDEEVE	219
Tco-1		AFVCYLKDHGGS-----CLQDFLSLP-----VDKIOFTIIPKG	197
Tco-2		GVLVYLKDKHGS-----TPRDFLSLHHSEAERYNVRGTADITQRCIKMAESTKHI GEVE	211
Tc		GMMKFLQSNPON-----SLDDEFWALPPEEAARMHSRGNVALIERSEGIQSGNDDEMESV	262
Lm		GCVLVLRQCRRNPHYRPRRGATFDDDNIECOEGRGRVDFDKIAASTISYDMVDQYLALPA	358
Li		GVVLVLRQCRRNPHYRSRRGATFDDDSIECOEGRGRVDFDKIAASTISYDMVDQYLALPA	255
Tb5a		ESEVCEYEGKED EED	232 100%/100%
Tb5b		ESEVCEYEGKED EED	231 72.0%/80.2%
Tco-1		TEISK	197 32.1%/45.3%
Tco-2		TEISK	215 39.0%/53.4%
Tc		SDDELGGASSSSSSSNGDS	280 27.2%/35.3%
Lm		CEMAMYREWAI DKFKLLPP	377 11.5%/18.9%
Li		CEMAMYREWAI DKFKLLPPQ	375 13.9%/21.4%

Figure S8. ClustalW sequence alignment of trypanosomatid CITFA-5 homologues. Since the available *T. vivax* CITFA-5 sequences were incomplete, CITFA-5 sequences of *Trypanosoma congolense* (Tco), another African trypanosome, were analyzed instead. GeneDB accession numbers of the sequences are Tb927.8.4030 and Tb927.8.4080 (Tb5a), Tb927.8.4130 (Tb5b), congo400e01.p1k_3 (Tco-1), congo240a09.q1k_19 (Tco-2), Tc00.1047053503897.140 (Tc), LmjF10.0650 (Lm), LinJ10.0990 (Li). In contrast to the CITFA-5 paralogues of African trypanosomes, the *T. cruzi* paralogue (Tc00.1047053509561.50) was nearly identical to its TcCITFA-5 counterpart and therefore not analyzed. Putative N-terminal extensions of the *T. cruzi* and *Leishmania* proteins were not included in this alignment.

```

Tb          MAVQP-GKVLAEQFSROGIQYEESS-IDV 26
Tv  MIQIHFRDQFWRNIF---YMSISLILLIEEMRP-GNSLAVQFCROGMRYEES-SDT 52
Tc  MNIVLVVVFICHLCSVNF---VYLVLVLFILKLEMAS-GKVLADQFRROGIRYEEA-SGI 54
Lm  MQELAQQFEQOGRCSADAPQRLWCSECTIYREDIAKFNDPLSEREGGVGVKAEGETGT 60
Li  MQELAQQFEQOGRCSADAPQRLWCSECTIYREDIAKSNDRHSELKGRVGVKAEGETGT 60

Tb  LWCSLCTKADSLLEIISRA-----TVEDVLEHCRLROHLLIMEKET 66
Tv  LWCSLCTTAOKLEIITAP-----TKENVLEHCCKOKRHLLCMEKEM 92
Tc  LWCSLCTNAKGLEIIEVE-----SKDSVLEHCALKRHLLIMEKOL 94
Lm  LDEAARADSDSGSSPSTPGPTRARHTRTPSPSLCTLEATKKAVLAHCATRTHLYLYECHV 120
Li  LDEAARADSDFGSSPSTPGQTRARHTSTPSPSLCTLEATKKAALLAHCATRTHLYLYEYHV 120

Tb  VK--EHYCPVEINGRSELLDHNCVYPSIMFGKGRMLLDDTLG--CVMTEDEVCGGVKLLIPR 122
Tv  LK--EYYCPVEINGRFRMLLDHNCIYPVVMFGRGRMLLDDTLG--CTMAVDVCGGVKLLIPR 148
Tc  AM--DQYCPVEINGCYMLLDHHCVYPAVMFGRGRMLLDDTLG--CLIPEDVCGGVKLLIPR 150
Lm  QNGLOHWCVELEHGYRMLLSHHCTIYPSRMFGDGRILLMDTSSGGMLLGVDCGGVKLWPO 180
Li  QNGLOHWCVELEHGYRMLLSHHCTIYPSRMFGDGRILLMDTSSGGMLLGVDCGGVKLWPO 180

Tb  HKYTVVVELVLP-----PAGTPIPK-----141
Tv  HKYTVVVELLIS-----PSTAPVPT-----167
Tc  HKYTVVVELLIP-----HSTVPLPH-----169
Lm  HOYTAVELLIPSAISGVREVLVPPVSMGAGEESQORASEAGSGKQORASSPPLPRNLPGSR 240
Li  HOYTAVELLIPSAISGVREVLVPPVSMGAEESQORVSEAESGKQORASSPPLPRNLPGSR 240

Tb  -----DPSYVVE 148
Tv  -----EPEYVVE 174
Tc  -----KPDYFVE 176
Lm  VIRIHEASEIPAENVAGSYNSALNGSRRKGGSKAMRSLDRGPAGTSGGRASSSIVPRYFDE 300
Li  VVRIHEASEIPAENVASSYNSALNGSRRKGGSKAMRSLDRGPAGRSGGRVSSSIVPRYFDE 300

Tb  RP-VRSRRCHEK-FDSWLOVENSVEGESALMKRRRIAEFSQINM 190 100%/100%
Tv  RP-IRSRKYITG-VKLLKQHNKSVDEGESALMKRRRVSEFQINM 216 57.4%/70.8%
Tc  RE-VRSRRCRLG-MDICLAYNKSVEGETAVMKRRRTVEGRVNMQGI 221 52.5%/66.1%
Lm  FHRIRSORVMEGSLKTVLEATNASTIGEARI RRRRRYAVKHNKTIETKGGKTAAGDEQ 358 22.3%/33.5%
Li  FHRIRSORVMEGSLKTVLEATNASTIGEARI RRRRRYAVKHNKTIETKGGKTAGSDEQ 358 22.1%/33.1%

```

Figure S9. ClustalW sequence alignment of trypanosomatid CITFA-6 orthologues. GeneDB accession numbers of the sequences are Tb927.5.970 (Tb), tviv1195b03.q1k_2 (Tv), Tc00.1047053508303.44 (Tc), LmjF35.1270 (Lm), LinJ35.1360 (Li). The *T. cruzi* paralogue (Tc00.104705351179.114) is nearly identical to the TcCITFA-6 analyzed.

Table S1: List of peptides identified by LC/MS/MS

CITFA-1: Tb11.47.0010 (19% coverage)

YILINNILCR
EVINTDKDLR
SGVHLWSAFQR
HRLLENLADGYSR
TGSAYSSFHAFHTGLPK
GTATLGSAEDNCNHLSVGEFDPTAYVDSISK

CITFA-2: Tb09.211.3440 (30% coverage)

KLCEPR
HMLVQPR
EVYISMR
YALVSPR
ELKVEMVTR
VAPFLEVVR
EFVIITITKR
IAFTDLLYHKR
MLELSSSIEGPISDAR
EVYCMQLLAPTASPVPWR
TGSAPIFSAEDRVLDDPLFNTVPVATYR

CITFA-3: Tb11.47.0008 (30% coverage)

TCYVMR
DFNLLK
QILYTAASR
LRGESSPFR
ANDVEELMAR
ISSTNNPFSE
LVLLCSDDVRR
GPSVPALFEVQR
MLQDAHYHFSSAEHR
EQVWDNIVDVGEGVK
VSGEVTHPLTVEALQMIGGTTR

CITFA-4: Tb11.01.0240 (37% coverage)

FKEVYR
GTVLATSR
LAIISHFHR
YTVTPGVNTER
STLASALPLLATK
ETKPQLDEDTLAPLR
SPVSLASDTVVDVAVR
EVLGGFSPEELAYLR
VGTANAEEDDMKEEVR
GCRPLGNTGIYCSFNCLR
TLLLSNSSVFAKPSLTAR

CITFA-5a: Tb927.8.4030 / Tb927.8.4080 (33% coverage); * peptides specific for this paralogue

MATYYSR*
LAGQEVVK
ECGTDEEAHFVIR
VTQVPSAAPSLNLR
ASCVQDFLSLPVER*
SKECGTDEEAHFVIR
FVGEHMCLAGAMSHANNNSK*

CITFA-5b: Tb927.8.4130 (22% coverage); * peptides specific for this paralogue

LAGQEVVK
ECGTDEEAHFVIR
DFLMLKPEEVEIYK*
VTQVPSAAPSLNLR
SKECGTDEEAHFVIR

CITFA-6: Tb927.5.970 (80% coverage)

IAFSQLNM
VLAEQFSR
QHLLLMEK

ADSLEIISR
RIAFSQLNM
ATVEDVLEHCR
EHYCPVEINGR
ETVKEHYCPVEINGR
SFLLDHNCVYPSTMFGK
YTVVEIVIPPAGTPIPK
QGIQYEESTDVLWCSLCTK
MLLDDTLGCVMTEDEVCGGVK
FDSWIQVHNSSVEGESALMK

DYNLL1: Tb11.50.0007 / Tb11.0845 (18% coverage); ** subtilisin digest

NADMPEDMQSDAVEVA**

Initial purification and analysis of the promoter-binding complex (detailed protocol)

Thirty liters of procyclic cells were harvested, washed with 200 ml of medium without FBS or hemin, and resuspended in lysis buffer (20 mM PIPES pH 7.5, 15 mM NaCl, 60 mM KCl, 0.5 mM EGTA, 4 mM EDTA, 0.5 mM DTT, and protease inhibitors — 0.5 mM PMSF, 1 μ M leupeptin, 2 μ g/ml aprotinin, 1 μ M pepstatin A, 0.1 mM TLCK) at a density of 4×10^9 cells/ml. Cells were disrupted by two cycles of freezing at -20°C and thawing. Nuclei were pelleted by centrifugation for 20 min at 15,000 g and 4°C , resuspended in nuclei buffer (20 mM HEPES pH 7.6, 420 mM KCl, 1.5 mM MgCl_2 , 0.2 mM EDTA, 25% glycerol, and protease inhibitors) at a concentration equivalent to 2×10^{10} cells/ml, and homogenized with 20 strokes of a type B pestle in a Dounce homogenizer. The homogenate was centrifuged for 30 min at 25,000 g and 4°C . The soluble nuclear extract was subjected to chromatographic purification by a protocol that was devised by empirical testing. All steps were carried out at 4°C and purification was monitored by EMSA. A 25–50% saturation ammonium sulfate precipitate was extensively dialyzed against buffer A1 (20 mM HEPES pH 7.6, 20% glycerol, 0.2 mM EDTA, 2mM DTT, and protease inhibitors) and passed through a column of S-Sepharose cation exchanger (Amersham) equilibrated with the buffer A1. The flow-through was adjusted to 2.5 mM MgCl_2 and 0.2 M NaCl and applied to a HiTrap Heparin column (Amersham) pre-equilibrated with buffer A1 containing 2.5 mM MgCl_2 and 0.2 M NaCl. After eluting with 0.8 M NaCl, precipitating with 50% ammonium sulfate, and dialyzing extensively against buffer A3 (10 mM HEPES pH 7.6, 10% glycerol, 60 mM NaCl, 0,2 mM EDTA, 2 mM DTT), the sample was adjusted to 40 mM DTT, 0.1 mM spermine, 1 mM spermidine and 0.05% Tween20, and loaded onto a Streptavidin Sepharose High Performance column (Amersham) pre-equilibrated with PBS and loaded with an 82-bp *VSG* ES promoter, PCR-amplified from plasmid ESP3 using one biotinylated and one normal primer. After elution with buffer A3 containing 1 M NaCl, and extensive dialysis against buffer B1 (bis-Tris-propane pH 7.0, 10% glycerol, 100 mM NaCl and 2 mM DTT), the extract was loaded to a Resource Q column (Amersham) pre-equilibrated with buffer B1. After a stepwise increase in NaCl concentration to 350 mM, the complex was eluted with a linear gradient 350-500 mM NaCl. The active fractions were loaded onto 4 ml 15–45% (v/v) linear glycerol gradients in buffer A4 (10 mM HEPES, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT) and centrifuged at 50,000 rev/min in a Beckman SW 60 rotor for 16 hr at 7°C . Fractions of 250 μ l were collected from the bottom.

The purified protein complex was reduced (20 mM dithiothreitol) and alkylated (100 mM iodoacetamide) in Laemmli loading buffer, resolved on a 10% SDS-PAGE Tris-glycine gel and silver-stained. Each visible band was excised, destained, dehydrated and treated with 100 ng of modified TPCK-trypsin for 8 hr at 37°C . Peptides were extracted by sonication in an aqueous mixture of 2.5% formic acid and 0.1% TFA for 5 min at 50°C , then transferred to a ZipTip (Millipore) containing 200 μ g Poros 20R2 beads (Applied Biosystems). After multiple washes with 0.1% aqueous TFA, peptides were eluted with 20 μ l 40% acetonitrile, 0.1% TFA, followed by 20 μ l 80% acetonitrile, 0.1% TFA, and these eluates were combined prior to lyophilization. The resulting peptide mixtures were resuspended in 0.01% (v/v) TFA in water/methanol/acetic acid (945:50:5 v/v/v) and subjected to LC-ESI-MS/MS analysis using an HPLC system (Smart System, Amersham) connected directly to an electrospray ion trap mass spectrometer (LCQ DECA, Thermo Fisher Scientific). Peptide mixtures were resolved on a reverse-phase column (Magic C_{18} , 50 mm x 0.2 mm, 100 \AA ; 5 μ m; Michrom Bioresources) with a fast-rising methanol gradient (7 min) at a flow rate of 40 μ l/min pre-column split to 2.8 μ l/min. MS fragmentation analysis was performed on the most intense ion observed at 2-second intervals. Protein searches were performed with SonarMSMS (Genomic Solutions) using locally collated versions of the available *T. brucei* genome data in 2001, with a re-analysis in May 2005.