

Supplementary Figures

Parasitology

Analyses of the essential C82 subunit uncovered some differences in RNA polymerase III transcription between *Trypanosoma brucei* and *Leishmania major*

Andrés Cano-Santiago, Luis E. Florencio-Martínez, Daniel E. Vélez-Ramírez, Adrián J. Romero-Chaveste, Rebeca G. Manning-Cela, Tomás Nepomuceno-Mejía and Santiago Martínez-Calvillo

Fig. S1. C82 sequence comparisons among eukaryotes. (A) Percent identities among amino acid sequences of C82. (B) Sequence alignment of the N-terminal region of C82. The trypanosomatid-specific loop is shown in bold letters. The sequences analyzed were: *T. brucei* (Tb, Tb927.2.2990), *L. major* (Lm, LmjF.27.2600), *Schizosaccharomyces pombe* (Sp, SPAPB1E7.03), *Saccharomyces cerevisiae* (Sc, NP_015516), *Homo sapiens* (Hs, NP_006459.3), *Mus musculus* (Mm, NP_083201), *Rattus norvegicus* (Rn, XP_008759532), *Arabidopsis thaliana* (At, NP_566914), *Caenorhabditis elegans* (Ce, NP_491807), *Plasmodium falciparum* (Pf, ETW47020.1), *Toxoplasma gondii* (Tg, XP_002366133), and *Naegleria fowleri* (Nf, KAF0973179).

Fig. S2. Sequence comparison of C82 in trypanosomatids. (A) C82 protein sequence alignment. (B) Percent identities among amino acid sequences of C82. The species analyzed were: *Leishmania braziliensis* (Lbr, LbrM.27.2720), *Leishmania tarentolae* (Lta, LtaP27.2650), *Leishmania mexicana* (Lmx, LmxM.27.2600), *Leishmania major* (Lmj, LmjF.27.2600), *Leishmania donovani* (Ld, LdCL_270033100), *Crithidia fasciculata* (Cfa, CFAC1_210027500), *Leptomonas seymouri* (Lse, Lsey_0405_0050), *Trypanosoma cruzi* (Tcr, TcCLB.509127.40), *Trypanosoma rangeli* (Tra, TRSC58_03156), *Trypanosoma brucei* (Tbr, Tb927.2.2990), and *Trypanosoma vivax* (Tvi, TvY486_0200650).

Fig. S3. Two isoforms of C82 are present in the *Leishmania* species. (A) Sequence comparison between C82 (LmjF.27.2600) and C82-short (LmjF.02.0680) in *L. major*. The peptide specific to C82-short is indicated with the black bar at the boundaries of the eWH3 and eWH4 domains. (B) Sequence alignment of C82-short. (C) Percent identities among amino acid sequences of C82-short. The species analyzed were: *L. major* (Lmj, LmjF.02.0680), *L. aethiopica* (Lae, LAEL147_000034900), *L. arabica* (Lar, LARLEM1108_020011500), *L. braziliensis* (Lbr, LbrM.02.0660), *L. donovani* (Ldo, LdBPK_020650.1), *L. gerbilli* (Lgr, LGELEM452_020011500), *L. infantum* (Lin, LINF_020012100), *L. mexicana* (Lmx, LmxM.02.0680), *L. tropica* (LtrLTRL590_020012900), and *L. turanica* (Ltu, LTULEM423_020011100).

Fig. S1

A

	Tb 563 aa	Lm 604 aa	Sp 591 aa	Sc 654 aa	Hs 534 aa	Mm 533 aa	Rn 533 aa	At 523 aa	Ce 544 aa	Pf 915 aa	Tg 823 aa	Nf 611 aa
Tb	100	32.78	13.97	12.25	13.64	14.25	14.53	16.32	12.43	8.80	18.45	16.83
Lm		100	13.62	13.32	16.05	16.09	16.09	15.34	16.62	11.22	21.04	17.59
Sp			100	25.78	22.57	22.57	22.36	17.73	20.08	15.26	19.85	22.20
Sc				100	19.25	18.45	18.87	16.05	16.19	15.74	15.50	18.54
Hs					100	94.00	94.00	21.14	26.04	15.25	18.34	24.53
Mm						100	97.37	20.55	26.29	15.29	18.16	24.17
Rn							100	20.76	26.29	15.29	18.38	24.79
At								100	19.44	18.28	17.56	24.15
Ce									100	16.30	14.44	21.41
Pf										100	18.03	17.41
Tg											100	19.53
Nf												100

B

Trypanosomatid-specific loop

```

Tb MS-RRAEAL-----GALNPQQQQQIIHDEHPRMIGVQNELLETSSVSEQLSPLAGFLVRTLNRCP-LTLKALAEAVRQEPFGNPLRSSKGGGASVSGGSVN--EAVKQ 99
Lm MP-FFCVGD-----AATCPSEVCEELAAKPITAHQLIISTVTHQLGFLAGICRTLIQSGP-MSLKDISDAVHRDEEVRAVAAAAANRASATGEKTTAVPPTHQ 98
Sp -----MSQYAVELCEILVEEFFGDCCSAVASALLRHGR-LTIPMLQKR----- 42
Sc M-DELLGEALSAENQTGESTVESEKLVTPEDVMTISSLEQRTLNPDLFLYKELVKAHLGERAASVIGMLVALGR-LSVRELVEK----- 82
Hs -----MTQAEIKLCSLLQEHFGEIVEKIGVHLIRTGS-QPLRVIAHD----- 42
Mm -----MTQAEIKLCSLLQEHFGEIVEKIGVHLIRTGS-QPLRVIAHD----- 42
Rn -----MTQAEIKLCSLLQEHFGEIVEKIGVHLIRTGS-QPLRVIAHD----- 42
At MSMEFGIVYAIHIITVQFQSVVSKVCECLLRKGP-LSSRDISRLAESDI----- 49
Ce MGGGKYETQMCVIIIEDIFGKIVAKVMETILKESCQLSTLFYKLG--GQI----- 48
Pf MIKNELDYLYKYIICDIFGSCCSDIIVEIILLYGNKLSIYELINL--SSY----- 46
Tg MFRYEVQFACLLLEDAFQPVVSTCARLLLLHGS-LTILELRDLFLAVPSASGVSPPLSRCSRNAAFA--SGCGRPSTHA----- 77
Nf MN--KRTLALEILTQDFGDVIGKVGGLMQFGP-QTMAEL-----IQA----- 40
  
```

```

Tb VSSPIPL---PVTAPSDIVTDAAIKEIMTRLLLRVQVH----- 135
Lm VGSSFVHNLGDIASVNDVIHEVAVKELVTRLVHRLIHA----- 137
Sp -----TSLPGPKIRQALVSLMQHHMVLYVTVIENVR-EVT---- 76
Sc -----IDGMDVDSVKTTLVSLTQLRCVKYLQETAI----- 115
Hs -----TGTSLDQVKKALCVLVQHNLVSYQV-HKRGVVEYEAQCS 80
Mm -----TKASLDQVKKALCVLIHNNLVLYHV-HKRGVVEYEAQCS 80
Rn -----TKASLDQVKKALCVLIHNNLVLYHV-HKRGVVEYEAQCS 80
At -----NHNKVKDILYLLIQHNCVQAFS-----IEPP 75
Ce -----SNTLIRKSVSTLINFGVTF-----HLDTN 73
Pf -----EFNVVRNVLCLLIHNIVDVNIYYSKSIETCNHLSPE 82
Tg -----GQDVYPLVRNALLVLLQHNLLVTPIDGQGAPSLRAAPQ 116
Nf -----TSMQFPHIRNALLVLIQHNMCIIYDKNIYQL-KLE---- 74
  
```

Fig. S2

A

eWH1 **Trypanosomatid-specific loop**

```

Lbr  MPLFRVSDAA---TRPSEVCELDVAKPIAAHYQLIIISTITDQLGPLAGAVCRALTQSGPMSLKDISDAVHRDEEVRVAAAAANRASATGETTTAVPATHQLGSSFHVLDGDIASVNDVI 117
Lta  MPLFRVGDGA---TRPSEVCELDAAKPIIAHHQLIISTIVTHQLGPLAGAICTRLIQSGPMSLKEISDAVHRDEELRAVAAAAANRASAAGETTAVPPTHQVGSNFVHVLGDIAVSNDVI 117
Lmx  MPLFCVGDVT---TRPSEVCELDAAKPIIAHHQLIISTIVTHQLGPLAGAICTRLIQSGPMSLKDISDAVHRDEEVRVAAAAANRASATGETTAVPPTHQVGSNFVHVLGDICTVNDVI 117
Lmj  MPFFCVGDA---TCPSEVCELDAAKPIIAHHQLIISTIVTHQLGPLAGAICTRLIQSGPMSLKDISDAVHRDEEVRVAAAAANRASATGETTAVPPTHQVGSNFVHVLGDIAVSNDVI 117
Ldo  MPFFCVGDA---TCPSEVCELDAAKPIIAHHQLIISTIVTHQLGPLAGAICTRLIQSGPMSLKDISDAVHRDEEVRVAAAAANRASATGETTAVPPTHQVGSNFVHVLGDIAVSNDVI 117
Cfa  MPLFCVGSKA---TQPSEVCELDAAKPIIAHHQLIISTIVTHQLGPLAGAAVCRALTTHAGPMTLSDISDAVHRDEEVRVAAAAANRASATGETTTAVPATHVGTNLHVLDVGLSAAKELI 117
Lse  MPLFCVGA---TQPSEVCELDAAKPIIAHYELLVSTVAHQGLGPLAGAVCRALTHTGPMTLSDISEAVRDEEDVRVAAAAASHASATGETTTAVPPTHRIIGSNLHVLDLGLSANELEI 117
Tcr -----MTRRHGDHHTHARILTEHQELLESSISEQLGPLAGFVLRNRRGALTKRGLADAVKREPFNPLHTG-----DATHKIGCPLAVPV---TA-GDMA 89
Tra -----MARVWHDSHHDHARMLTEHAELLESSVSEQLGPLAGFVLRNRRGALTKRGLADAVKREPFNPLHTG-----DAARHVGCP LAVSA---VA-GDVG 90
Tbr MSRRABALGALNPQQQQIIHDEHPRMIGVQNELLETSVSEQLSPLAGFLVTRLNRCGPLTKLALAEAVRQEPFNGPLRNSKGGGASVSGS---VNEAVKQVSSPIPLPV---TAPSDIV 115
Tvi MVQRVSDAASRNQRQSQSVVTAHNELLESIARQLGPLAEFLVTRLNRCGPLTKLELADTTKRSEFNNPLRSGNDGET-----DATHKTRSPILPLV---TVGADVA 107
      . . . . . : : : : : * : * * : : : . . . . . : : . . : : .
  
```

eWH2

```

Lbr  HEVAVKELVTRLVVHRLVHTDPVTHLYEMRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVVRVLEDGWLPVPVPLTAAAPDAATPA 237
Lta  HEVAVKELVTRLVHRLIHADPITHLYELRHGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVVRMVENGWLVPLVPSAAGAAATV 236
Lmx  HEVAVKELVTRLVVHRLIHADPATHLYEIRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVVRMEDGWLVVPVNSPTAAPGAAATA 237
Lmj  HEVAVKELVTRLVVHRLIHADPATHLYEIRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVVRMEDGWLVPLVPLTAAAGATA 237
Ldo  HEVAVKELVTRLVVHRLIHADTATHLYEIRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVVRMEDGWLVPLVNSPTAAPGATTA 237
Cfa  HEVAVKEVVTRLVHRLVYADPARYLGMRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVRMAEDGWLVPLVNSPTAAPGATTA 229
Lse  HEVAVKEVVTRLVHRLVYADPARYLGMRYGSAILLRVLPPLLLHCARQYGEAAKCVLLVAVQLGVVPPRAAVQVALSRTPSITRDAIEYAVRMAEDGWLVPLVNSPTAAPGATTA 229
Tcr TDAAIKEMARLVLRHVLVADPSTENYVSLGYLELRLTFLPLVMQFFQRRYGETGVTLLMVFYQLAVVPRDAAVRLAMERRPALSEA-LRCHAEMERLGFTEFVSAAGSAAEAAA 204
Tra TDAIKELMARLVLRHVVVYDAERHVVGYLGYGLLRTFLPLVMQFFQRRYGETGVTLLMVFYQLAVVPRDAAVRLAMERRPALSEA-LRCHAEMERLGFTEFVSAAGSAAEAAA 205
Tbr TDAAIKEMTRLLHRLVQHPDSSSYVSLGYGLLRTFLPLVMQFFQRRYGETGVTLLMVFYQLAVVPRDAAVRLAMERRPALSEA-LRCHAEMERLGFTEFVSAAGSAAEAAA 234
Tvi ADAAIKEVVTRLHRLVYVDAESQYVSLGYGLLRTFLPLVMQFFQRRYGETGVTLLMVFYQLAVVPRDAAVRLAMERRPALSEA-LRCHAEMERLGFTEFVSAAGSAAEAAA 224
      . . . . . : : * : * * * : : . . . . . : : : : : * : * * : : : . . . . . : : . . : : .
  
```

Insertion loop **eWH3**

```

Lbr  ---LRADGASSHLGGSIDLWSATRHPCVLEQAALHLCLFSDAIE-QVAERYADG-----RLALTLVRALRRRTPANSEYTEAVASFQELAAELPPTA-----GVRDRSGEP---RVG 338
Lta  ---TRAADVASSHPDNDILWNAATTPYCVLEEAALHLYFNDAIE-QVAERYADG-----RLALTLVRTLRRAAPNSGYTEAVASFQELAAELPPTA-----GVRDRSGEP---RMG 338
Lmx  NAARLADLADPSHPGNDLWNAATTPYCVLEEAALHLYFNDAIE-QVAERYADG-----RLALTLVRVLRRAAPNSGYTEAVASFQELAAELPPTA-----GVRDRSGEP---LMG 341
Lmj  NAARLADLADPSHPGNSDWNATTPYCVLEEAALHLYFNDAIE-QVAERYADG-----RLALTLVRVLRRAAPNSGYTEAVASFQELAAELPPTA-----GVRDRSGEP---RMG 341
Ldo  NAARLADLADPSHPGNDLWNAATTPYCVLEEAALHLYFNDAIE-QVAERYADG-----RLALTLVRALRRRTPANSEYTEAVASFQELAAELPPTA-----GVRDRSGEP---RMG 341
Cfa  -----AAVKESADHVDAATAPIRYILESALFSLMDAIE-QVAERYADG-----GLALTLVRALRRRAAPNSRYTESVASFQELAAELPHTASVGGFGDWDGKGP---VAL 330
Lse  -----PTS-SCDVAIFPSATTPYCVLEESALFFLMDALE-QVAERYADG-----GLALTLVRALRRRAAPNSQYDGAASQELAAELPPTA-----GVRDRSGEP---VQL 324
Tcr -----TT-THSTGLSPDPCRLVIGNILAEMLHDAIQQLLAERFADG-----GVAAVMEAFATLGRTRLVERSH-----GFPFVHPRTSASPLRLVL 288
Tra -----T-ATPTATLSPDPCRLVIGNILAEMLHDAIQQLLAERFADG-----GVAAVMEAFATLGRGRLLEERSP-----MFPFVHPRASVSLPLRLV 288
Tbr NAGSGTRVEGPGE-TVAAADALQELCRLLVIGNILAEMLRDAIHRHLLAERFADGSGSGGGVSAVAIEAFATSA-RCRRFRSERVS-----GFPAAPRPSASMPRLAL 336
Tvi EAPSRQST---TT-ANLALPLSDPCRLVIGNILAEMLRDAIHRHLLAERFADGSGNS--SSVAVI VDCFTKAT-RTRHLEKRLD-----GFPFMPRVPVSLRLAL 321
      . . . . . : : * : * * * : : . . . . . : : : : : * : * * : : : . . . . . : : . . : : .
  
```

```

Lbr  TDSVSSGSAEAVRCLERLQCPVIFAL-GCNAAPMTPAAGSS-----PNFTMPSTRALLVKPQSESADYALDDVTAVQLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 446
Lta  TDNLPSSSVEAQRCLRQCPVIF-Y-SCDAAPMTSATGVS-----PSLTMPTSTRALLVKPHSOSNVYSMHVAAVHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 445
Lmx  TDNVPSSEVAVQRCLRQCPVIF-Y-SCDAAPMTSVAGIS-----PSFTMPSTRALLVKPHSOSNFYALDHTAVHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 448
Lmj  TDNVPSSEVAVQRCLRQCPVIF-Y-SCDAAPMTSAAGIS-----PSFTMPSTRALLVKPHSOSNFYALDHTAVHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 448
Ldo  TDNVPSSEVAVQRCLRQCPVIF-Y-SCDAAPMTSVAGIS-----PSFTMPSTRALLVKPHSOSNFYALDHTAVHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 448
Cfa  PET-TQSAEAQRCLRQCPVIF-L-SCDAASATATPLSPGATASSSSAAFPPLATRALLVQPPNQRGFYALDDVTATHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 446
Lse  PEN-TRSAAEAQRCLRQCPVIF-L-SCDGNASTAAAAPH---ASASSSTLPVALPATRALLVQPHKQHGFAVHLLQEAVCERVVYSGVGLVGRIMKLLQHHPFLEE 436
Tcr LREVP-MAEGVVAVSRMCRPDGDSLLSNDGTAT-VA-----LEGPYRFRYDAAIEAMQRDCERLVARHGVGLVGRIMKLLQHHPHMD 372
Tra LREVP-LAESAVMTVSRMCRPDGDSLLSNDGTAT-VA-----LEGPYRFRYDAAIDAMRDCERLVARHGVGLVGRIMKLLQHHPHMD 372
Tbr LREVP-HPEPTVTEVLRNRCRAEADGINSMSGGGAGVDA-----LEGPYRFRYDAAIEAMQRDCERLVARHGVGLVGRIMKLLQHHPHMD 421
Tvi LCEVP-HDEQAVLETLSRMCSDGDFLVCSDGASVSGS-----MDGSFSFYHSAVEAMRDCERLVARHGVGLVGRIMKLLQHHPHMD 406
      . . . . . : : * : * * : : . . . . . : : : : : * : * * : : : . . . . . : : . . : : .
  
```

eWH4 **Cleft loop** **C. coil**

```

Lbr  RTLAEQSVATYVKREVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---CV-----RSSPAASGGIKSSA 554
Lta  RTLAEQSVATYVKREVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVVPKPA---RV-----GSTPAGGFFNSV 553
Lmx  RTLAEQSVATYVKREVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 556
Lmj  RTLAEQSVATYVKREVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 556
Ldo  RTLAEQSVATYVKREVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 556
Cfa  RTLAEQSVATYVKTRDVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 562
Lse  RTLAEQSVATYKTRDVLHQMFKDGLLQEQVPRTSALVERSAKASVYLWGLSWSATLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 539
Tcr RMLAEEAATLPTREVLRHMDRQVNRGVPKNSVFLWGCSLEKDLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 492
Tra RMLAEEAATLPTREVLRHMDRQVNRGVPKNSVFLWGCSLEKDLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 480
Tbr RMLAEEAATLPTREVLRHMDRQVNRGVPKNSVFLWGCSLEKDLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 518
Tvi RMLAEEAATLPTREVLRHMDRQVNRGVPKNSVFLWGCSLEKDLLPVRERLAKLTIIAWVKLREAOQQA SAVAKPAA---RV-----RPTPAAGGINSNA 503
      * * * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * : * : * * * :
  
```

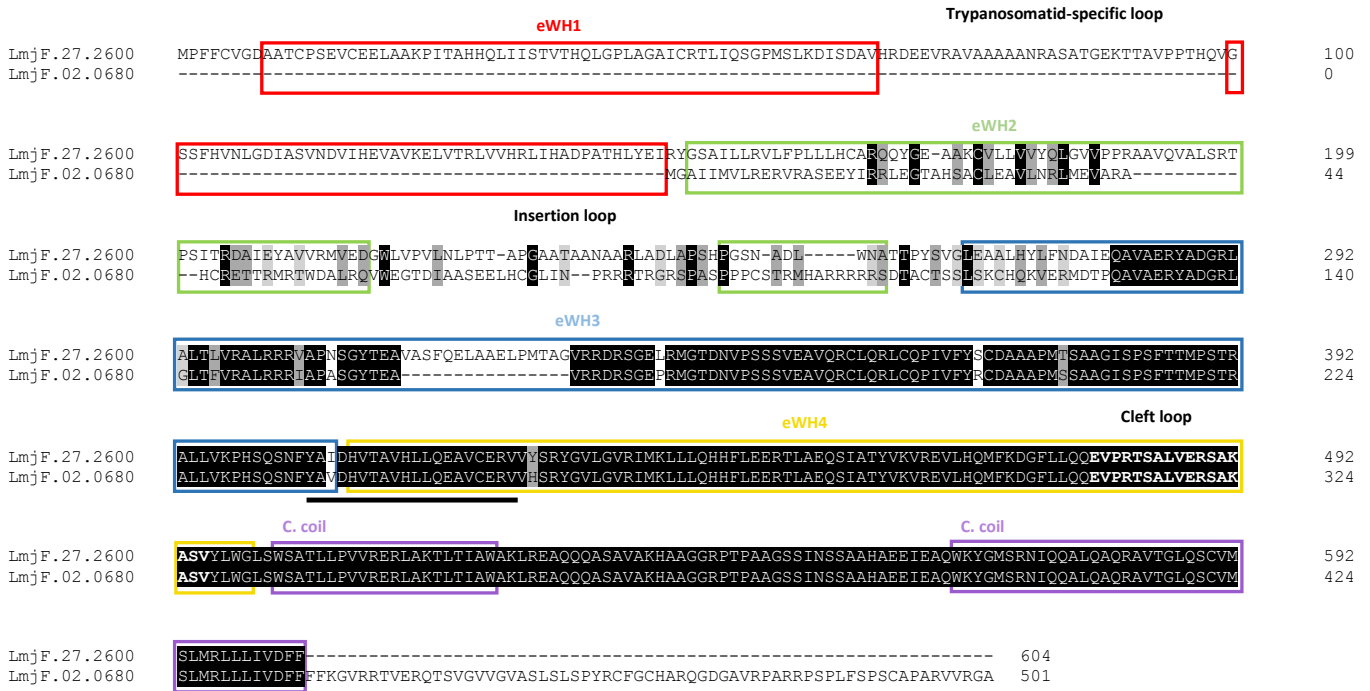
C. coil

```

Lbr  AHAEVEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 602
Lta  AHTEVEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 601
Lmx  AHAEIEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 604
Lmj  AHAEIEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 604
Ldo  AHAEIEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 604
Cfa  SVMDEAQNKSGMSRNIQQALQARTITGLQSCVMSLMRLLIVDF 610
Lse  QAADAEAMRKYSLTRMQALRVQRTITGLQSCVMSLMRLLIVDF 587
Tcr AAGLSLTAGIGTGEVIAETLRQRAVALESSVALMGTMLVDFY 540
Tra VATSEAAAGVRRPTEAATLPRTRQRAVALESSVALMGTMLVDFY 528
Tbr VASATATIN---GDVHTPGGSYESKHVALESSVALMGTMLVDFY 563
Tvi VAAHSSVSG---VNSQGAKNVDGNYVPIALLESSVALMGTMLVDFY 548
      . . . . . : : * : * * : : . . . . . : : : : : * : * * : : : . . . . . : : . . : : .
  
```


Fig. S3

A



B

	eWH2	Insertion loop	
Lmj	MGAIIMVLRERVASEEYIRRLEGTAHSACLEAVLNRLMEVARAHCRETTMRMTWDALRQ	WEGTDIAASEELHCGLINPRRRTRGRSPAS	99
Lae	-----	-----	0
Lar	-----	-----	52
Lbr	-----	-----MNR-----	3
Ldo	-----	-----	0
Lgr	-----	-----	0
Lin	-----	-----	0
Lmx	-----	-----	0
Ltr	MGPLIMVLRREGVRVPEYIRGLEGPAHSACLEAVLNRLMEVARARCRETKPMRTWDALRRV	WEGMDIAASEELHCGLINPRRRTRGRSSAS	100
Ltu	MDPIIMVLRERVVSEEYIRRLEGPAHSACLEAVLNRLMEVARARCRETTMRMTWDALRQ	WEGMDIAASEELHCGLINPQRRTGRSSAS	100

	eWH3	
Lmj	HARRRRSDTACTSSLKCHQKVERMDTPQAAVAERYADGRL-----GLTFVRLRRRIAP-ASGYTEAVRRDRSGEPRMGTDNVPSSSVEAV	185
Lae	-----MHRPRPS-----AVQTDGF-----PLTLVRLRRCIAP-SSGYTEAVRRDRSGELRMGTDNVPSSSVEAV	60
Lar	HARRRRADTACTSSLKCHQKVERMDTPQAAAERYADGRL-----ALTLVRLRRCIAP-DSGYTEAVRRDRSGEPRGTDNVPSSSAEAV	138
Lbr	-----SADQHVAAKYPDKRRVRTSPHPSQGREIKTHMRRRLRGRGREGGDDTGQVADGECGGEPRVGTDSVSSGSAEAV	77
Ldo	-----	14
Lgr	-----	0
Lin	-----	14
Lmx	-----	14
Ltr	HARRRRRAGTACTSSLKCHRKVERMDAPQAAAERYADGRL-----ALTLVRLRRRVAP-NSGYTEAVRRDRSGELRMGTDNVPSSSVVAV	186
Ltu	HARRRRRADTACTSSLKCHQKVERMDTPQAAAERYADGRL-----ALTLVRLRRRVAP-DSGYTEAVRRDRSGEPRMGTDNVPSSSVEAV	186

	eWH4	
Lmj	QRCLQRLCQPIV-FYRCDAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVRIMKLLQHHFLEERTLA	284
Lae	QRCLQRLCQPIV-FYSCDAAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVRIMKLLQHHFLEERTLA	159
Lar	QRCLQRLCQPIV-FYSCDAAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVCIMKLLQHHFLEERTLA	237
Lbr	RRCLERLCQPVAFAGCNAAAPMTPAAGSSPNFTTTPSTRALFVKPQSESDLYALDDVTAVQLLQEAVCERVVHSRYGVLGVRIMRLLQHHFLEERTLA	177
Ldo	QRCLQRLCQPIV-FYSCDAAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVRIMKLLQHHFLEERTLA	113
Lgr	-----	16
Lin	QRCLQRLCQPIV-FYSCDAAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVRIMKLLQHHFLEERTLA	113
Lmx	---QPLCQTVV-FYSCDAAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVCIMKLLQHHFLEERTLA	109
Ltr	QRCLQRLCQPIV-FV-----	200
Ltu	QRCLQRLCQPIV-FYSCDPAAPMSSAAGISPSFTTTPSTRALLVKPHSQSNFYAIDHVTAVHLLQEAVCERVVHSRYGVLGVRIMKLLQHHFLGERTLA	285

	Cleft loop	C. coil	
Lmj	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQQASAVAKHAAGRPTPAAGSSI		384
Lae	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQQASAVAKPAAGVRPTPAAGSSI		259
Lar	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQQASAVAKPAAGVRPTPAAGSSI		337
Lbr	EQSVATYVKAREVLHRMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPAVRERLAKTLTIWAKLREAQQASAVAKPAACVRSSPAAGGGI		277
Ldo	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQQH-----		192
Lgr	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKALTIWAKLREAQQASAVAKPAAGVRPTPAAGSSI		116
Lin	EQSIATYVKVREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQQH-----		192
Lmx	EQSVATYVKVREVLHQMFKDGFLQQEVPRTSALVERPAKASVYLWGGG-----		158
Ltr	-----		200
Ltu	EQSIATYVKAREVLHQMFKDGFLQQEVPRTSALVERSAKASVYLWGLSWSATLLPVVRERLAKTLTIWAKLREAQLQASAVAKPAAGVRPTPAAGSSI		385

	C. coil	
Lmj	NSSAAHAEIEAQWKYGMSRNIQQALQAQRAVTLGQSCVMSLMRLLLVDFEFPKGVRRVTRVQTSVGVVGVASLSLSPYRCFGCHARQDGAVRPARRP	484
Lae	NSSAAHAEIEAQWKYGMSRNIQQALQAQRAVTLGQSCVMSLMRLLLVDFE-----	311
Lar	NSSAAHAEIEAQWKYGMSRNIQQALQAQRAVTLGQSCVMSLMRLLLVDFE-----	389
Lbr	KSSAAHAEIEAQWKSMSRSIQALQQTQRTITGLQSCVMSLMRLLLVDFE-----	329
Ldo	-----	192
Lgr	NSSAAHAEIEAQWKYGMSRNIQQALQAQRAVTLGQSCVMSLMRLLLVDFE-----	168
Lin	-----	192
Lmx	-----	158
Ltr	-----	200
Ltu	NSSAAHAEIEAQWKYGMSRNIQQALQAQRAVTLGQSCVMSLMRLLLVDFE-----	437

Lmj	SPLFSPSCAPARVVRGA	501
Lae	-----	311
Lar	-----	389
Lbr	-----	329
Ldo	-----	192
Lgr	-----	168
Lin	-----	192
Lmx	-----	158
Ltr	-----	200
Ltu	-----	437

