UNIVERSITY of York

This is a repository copy of *Divergent cytochrome c maturation system in kinetoplastid protists*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/172747/

Version: Accepted Version

Article:
Belbelazi, Asma, Neish, Rachel, Carr, Martin et al. (2 more authors) (Accepted: 2021) Divergent cytochrome c maturation system in kinetoplastid protists. MBio. ISSN 2150-7511 (In Press)

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Divergent cytochrome *c* maturation system in kinetoplastid protists

Asma Belbelazi[a], Rachel Neish[b], Martin Carr[a], Jeremy C. Mottram[b]#, and Michael L. Ginger[a]#

[a]School of Applied Sciences, University of Huddersfield, Huddersfield, UK

[b]York Biomedical Research Institute and Department of Biology, University of York, York, UK

Running head: kinetoplastid cytochrome *c* maturation

#Address correspondence to Michael L. Ginger, m.ginger@hud.ac.uk or Jeremy C. Mottram, jeremy.mottram@york.ac.uk

Asma Belbelazi and Rachel Neish contributed equally to this work. Author order was determined on the basis of alphabetical order

ABSTRACT In eukaryotes, heme attachment through two thioether bonds to mitochondrial cytochromes *c* and c_1 is catalysed by either multi-subunit cytochrome *c* maturation system I or holocytochrome *c* synthetase (HCCS). The former was inherited from the α-proteobacterial progenitor of mitochondria; the latter is a eukaryotic innovation for which prokaryotic ancestry is not evident. HCCS provides one of few exemplars of *de novo* protein innovation in eukaryotes, but structure-function insight of HCCS is limited. Uniquely,

euglenozoan protists, which include medically relevant kinetoplastids *Trypanosoma* and *Leishmania* parasites, attach heme to mitochondrial *c*-type cytochromes by a single thioether linkage. Yet the mechanism is unknown as genes encoding proteins with detectable similarity to any involved in cytochrome *c* maturation in other taxa are absent. Here, a bioinformatics search for proteins conserved in all hemoprotein-containing kinetoplastids identified kinetoplastid cytochrome *c* synthetase (KCCS), which we reveal as essential, mitochondrial, and catalyses heme attachment to trypanosome cytochrome *c*. KCCS has no sequence identity to other proteins, apart from slight resemblance within four short motifs suggesting relatedness to HCCS. Thus, KCCS provides a novel resource for studying eukaryotic cytochrome *c* maturation, possibly with wider relevance since mutations in human HCCS leads to disease. Moreover, many examples of mitochondrial biochemistry are different in euglenozoans as compared to many other eukaryotes; identification of KCCS thus, provides another exemplar of extreme, unusual mitochondrial biochemistry in an evolutionarily divergent group of protists.

IMPORTANCE Cytochromes *c* are essential proteins for respiratory and photosynthetic electron transfer. They are post-translationally modified by covalent attachment of a heme cofactor. Kinetoplastids include important tropical disease-causing parasites; many aspects of their biology differ from other organisms, including their mammalian or plant hosts. Uniquely, kinetoplastids produce cytochromes *c* with a type of heme attachment not seen elsewhere in nature and were the only cytochrome *c*-bearing taxa without evidence of

protein machinery to attach heme to the apo-cytochrome. Using bioinformatics, biochemistry and molecular genetics we report how kinetoplastids make their cytochromes *c*. Unexpectedly, they use a highly diverged version of an enzyme used for heme-protein attachment in many eukaryotes. Mutations in the human enzyme lead to genetic disease. Identification of kinetoplastid cytochrome *c* synthetase, thus, solves an evolutionary unknown, provides a possible target for anti-parasite drug development, and an unanticipated resource for studying the mechanistic basis of a human genetic disease.

KEYWORDS cytochrome *c*, *Leishmania*, mitochondrial metabolism, post-translational modification (PTM), protist, *Trypanosoma brucei*

Trypanosomatid parasites of the genera *Trypanosoma* and *Leishmania* are responsible for a variety of serious neglected tropical diseases and belong to the class of flagellate protists called Kinetoplastea. Numerous aspects of kinetoplastid mitochondrial biology, including genome organisation (**1**), RNA editing, protein/tRNA import (**2**) and cristae formation (**3**) are highly divergent or unique compared with many eukaryotes.

A fundamental aspect of mitochondrial function is covalent attachment of heme to mitochondrial cytochromes *c* and c_1 within the mitochondrial intermembrane space. For this post-translational modification, thioether bonds form between heme vinyl groups and cysteine sulphydryl groups of a CxxCH heme-binding motif within the apo-cytochrome. The stereochemistry of this heme attachment is conserved across evolution: the 2-vinyl group of

63 heme attaches to the first cysteine, the 4-vinyl group to the second, and histidine provides
64 an axial ligand to the heme iron. In most eukaryotes, holocytochrome *c* synthetase (HCCS),
65 associated with the outer leaflet of the mitochondrial inner membrane, catalyses heme
66 attachment to mitochondrial cytochromes *c* (**4**). In the proto-mitochondrion, however, the
67 multi-subunit, integral membrane cytochrome *c* maturation System I provided an ancestral
68 pathway for *c*-type cytochrome biogenesis. It is retained, partially mitochondrially encoded,
69 in many plants and a few protists. Instances of eukaryotes containing both maturation
70 systems are extremely rare: orphan taxon and predatory flagellate *Ancoracysta twista*
71 reportedly contains HCCS and System I, but is now extinct in the laboratory (**5**) and survey
72 of the 1000 plant transcriptome resource (**6**) suggests club mosses *Phylloglossum*
73 *drummondii* and *Huperzia squarrosa* possess HCCS plus nuclear-encoded System I
74 fragments or a mitochondrial CcmF pseudogene, respectively[1].

75 Kinetoplastid protists are the only eukaryotes where mitochondrial cytochromes are
76 present but evidence of a cytochrome *c* maturation system is absent (**4**, **7**, **8**). Moreover,
77 kinetoplastids and other euglenozoans (*e.g. Euglena gracilis*), are unique in that heme is
78 bound through only a single thioether linkage in mitochondrial cytochromes *c*: In
79 Euglenozoa AAQ**C**H and FAP**C**H are the conserved heme-binding motifs in cytochromes *c*
80 and c_1, respectively (the residue at the proximal heme-binding cysteine in normal *c*-type
81 cytochromes is underlined; the heme-binding cysteine conserved in all cytochromes *c* is in
82 bold). Why euglenozoans possess mitochondrial cytochromes *c* with heme bound by a

[1] AB, MC, MLG, unpublished observations

single thioether bond is a mystery of almost fifty years standing. No noticeable difference in the physicochemical properties of euglenozoan cytochromes *c* is known (**9**). Yet, the activities of *Euglena* cytochrome *c* reductase and oxidase vary dependent upon the source of the cytochrome *c* used plus there are fitness costs in trypanosomes engineered to express only CxxCH heme-binding cytochrome *c* (**10**). This leaves it possible single cysteine linkage affects electron transport through the mitochondrial respiratory chain (**11**). The strict conservation of phenylalanine and proline within heme-binding motifs of kinetoplastid and *Euglena* cytochrome c_1 is another puzzle and potentially unique. Perhaps, the proline introduces a local bend in the polypeptide that allows accommodation of the phenylalanine side-chain with the proteins tertiary structure (**11**).

To resolve how kinetoplastids mature their unique mitochondrial *c*-type cytochromes we sorted candidate mitochondrial proteins to identify those conserved in all kinetoplastids, except for plant-pathogenic *Phytomonas*. In *Phytomonas*, adaptation to carbohydrate-rich plant latex correlates with secondary loss of mitochondrial cytochromes and other hemoproteins (**12**); thus, we reasoned that in *Phytomonas* a cytochrome *c* maturation system would also be lost. Candidate mitochondrial proteins were then screened for motifs similar to any present in proteins belonging to the four biogenesis systems known to catalyse heme attachment to a cysteine sulphydryl (**Text S1**). We identified a single hypothetical protein, highly conserved across the Kinetoplastea (**Fig. S1**), but absent from *Phytomonas* (encoded by Tb927.3.3890 in *T. brucei*; LmxM.08_29.1300 in *Leishmania mexicana*) that exhibited co-linearity, but very limited sequence similarity, to four HCCS motifs required for thioether bond formation (**Fig. 1A**). Similarity between this kinetoplastid

105 protein and human HCCS was too limited to be detected by PSI-BLAST, but the histidine
106 (His154) essential in HCCS for heme attachment to apo-cytochrome *c* (**13**) was present at an
107 analogous position in candidate kinetoplastid cytochrome *c* synthetase (KCCS).

108 To assess KCCS candidature, we co-expressed recombinant Tb927.3.3890 and *T.*
109 *brucei* cytochrome *c* in *Escherichia coli*. We reported previously *T. brucei* apo-cytochrome *c*
110 is neither subject to spontaneous maturation in the *E. coli* cytoplasm nor a substrate for the
111 endogenous periplasmic *E. coli* cytochrome *c* maturation system (which is expressed
112 minimally under the aerobic conditions we used to cultivate our *E. coli* cultures) (**7**). Here,
113 recombinant expression of *T. brucei* cytochrome *c* bearing an N-terminal hexa-histidine tag
114 also resulted in no detectable holo-cytochrome *c* formation. Co-expression of Tb927.3.3890
115 and His_6-tagged *T. brucei* cytochrome *c*, however, resulted in heme attachment to the latter,
116 as shown by SDS-PAGE of purified protein and staining for covalently bound heme (**Fig.**
117 **1B**). Pyridine hemochrome spectra of purified recombinant trypanosome holo-cytochrome *c*
118 confirmed heme attachment via a single thioether bond (**Fig. 1C**): in the spectra shown, the
119 pyridine haemochrome α-band maximum of the recombinant cytochrome was 553 nm, and
120 clearly red-shifted in comparison with the corresponding 550 nm α-band maximum of
121 cytochromes *c*, which bind heme via two thioether bonds (equine cytochrome *c* in **Fig. 1C**).
122 Thus, Tb927.3.3890 is appropriately referred to as Tb*KCCS*.

123 For molecular genetics analyses of KCCS, we used *L. mexicana* engineered for
124 tractable CRISPR-Cas9 genome editing (denoted as T7 in **Fig. 2B** and T7Cas9 in **Fig. 2C-E**)
125 (**14**). *Leishmania* promastigotes have no capacity for anaerobic growth; thus, mitochondrial

126 cytochromes are essential (**15**). Tagged with mNeonGreen and expressed from an
127 endogenous chromosomal locus, *Lm*KCCS showed mitochondrial localisation (**Fig. 2A**),
128 consistent with a role in mitochondrial cytochrome *c* maturation. Mitochondrial localisation
129 was observed irrespective of whether the mNeonGreen tag was N- or C-terminal providing
130 evidence for an internal hydrophilic mitochondrial import signal, as described in yeast HCCS
131 (**16**). We were unable to generate Lm*KCCS* null mutants and were only able to delete both
132 chromosomal copies of Lm*KCCS* without further genome rearrangements following
133 episomal expression of GFP-tagged Lm*KCCS* or Tb*KCCS* (**Fig. 2B-E**). In the absence of
134 pNUS-derived episomes, CRISPR-Cas9-mediated disruption of both Lm*KCCS* alleles in
135 diploid *L. mexicana* resulted in genome duplication, as well as the site-specific integration of
136 drug-resistance cassettes, as revealed by propidium iodide (PI) staining and flow cytometry
137 of methanol-fixed logarithmic phase parasites (**Fig. 2C**). Thus, CRISPR-Cas9 genome editing
138 indicated Lm*KCCS* is an essential gene.

139 Our comparative genomics approach identified KCCS and provides a mechanism by
140 which kinetoplastids mature their mitochondrial cytochromes *c* with heme attached via a
141 single thioether bond. Other than the missing thioether bond, x-ray structures of
142 trypanosomatid and yeast holo-cytochromes *c* are very similar, although the former (with
143 either AxxCH or an engineered CxxCH heme-binding motif) is a very poor substrate for
144 yeast HCCS (**17**), presumably because HCCS requires interaction with amino acids upstream
145 of the heme-binding motif (**18**, **19**).

Inability to detect proteins homologous to KCCS outside of the Kinetoplastea suggests KCCS is more likely to be an extreme or highly divergent HCCS, rather than derived through convergent evolution. Yet, irrespective of its origin, KCCS provides another example of 'extreme biology' within a group of protists already well-known for pushing the boundaries (**20**).

Further characterisation should indicate whether the amino acid differences and insertions observed in *Perkinsela* KCCS reflect this taxon's basal position in kinetoplastid phylogenies or the particular, unusual niche occupied by this obligate symbiont of *Paramoeba* (**21**). There is also wider significance to our observations: HCCS mutation in humans causes MLS (**22**, **23**), but insight into HCCS catalysis, and thus the mechanistic consequence(s) of HCCS mutation, are in their infancy (**13**, **24**-**26**). Our earlier work replacing 'AxxCH' cytochrome *c* in *T. brucei* with a 'CxxCH' variant and characterising the resultant mis-matured cytochrome (**10**), indicates KCCS is a highly diverged HCCS that cannot act on the first cysteine of a conventional CxxCH heme-binding motif. Indels, deletions, and substitutions within motifs I and II of KCCS, including substitution of the Glu159 MLS mutation (**22**), may be particularly informative for future study of HCCS catalysis since these are the motifs believed to mediate heme binding and release (**13**, **26**). Intriguingly, Arg217, the other residue mutated in MLS patients is also not conserved in KCCS.

Methods are available as supplemental material in Text S1.

ACKNOWLEDGEMENTS

AB was supported by a School of Applied Sciences PhD studentship. JCM is supported by the Wellcome Trust (200807/Z/16/Z).

REFERENCES

1. Aphasizheva I, Alfonzo J, Carnes J, Cestari I, Cruz-Reyes J, Göringer HU, Hajduk S, Lukeš J, Madison-Antenucci S, Maslov DA, McDermott SM, Ochsenreiter T, Read LK, Salavati R, Schnaufer A, Schneider A, Simpson L, Stuart K, Yurchenko V, Zhou ZH, Zíková A, Zhang L, Zimmer S, Aphasizhev R. 2020. Lexis and grammar of mitochondrial RNA processing in trypanosomes Trends Parasitol 36:337-355.
2. von Känel C, Muñoz-Gómez SA, Oeljeklaus S, Wenger C, Warscheid B, Wideman JG, Harsman A, Schneider A. 2020. Homologue replacement in the import motor of the mitochondrial inner membrane of trypanosomes. Elife 9:e52560.
3. Kaurov I, Vancová M, Schimanski B, Cadena LR, Heller J, Bílý T, Potěšil D, Eichenberger C, Bruce H, Oeljeklaus S, Warscheid B, Zdráhal Z, Schneider A, Lukeš J, Hashimi H. 2018. The diverged trypanosome MICOS complex as a hub for mitochondrial cristae shaping and protein import. Curr Biol 28:3393-3407.e5.
4. Babbitt SE, Sutherland MC, san Francisco B, Mendez DL, Kranz RG. 2015. Mitochondrial cytochrome *c* biogenesis: no longer an enigma. Trends Biochem Sci 40:446-455.

187 5. Janouškovec J, Tikhonenkov DV, Burki F, Howe AT, Rohwer FL, Mylnikov AP, Keeling
188 PJ. 2017 A new lineage of eukaryotes illuminates early mitochondrial genome
189 reduction. Curr Biol 27:3717-3724.

190 6. One Thousand Plant Transcriptomes Initiative. 2019 One thousand plant
191 transcriptomes and the phylogenomics of green plants. Nature 574:679-685.

192 7. Allen JWA, Ginger ML, Ferguson S.J. 2004. Maturation of the unusual single-cysteine
193 (XXXCH) mitochondrial *c*-type cytochromes found in trypanosomatids must occur
194 through a novel biogenesis pathway. Biochem J 383:537-542.

195 8. Allen JW, Jackson AP, Rigden DJ, Willis AC, Ferguson SJ, Ginger ML. 2008. Order
196 within a mosaic distribution of mitochondrial *c*-type cytochrome biogenesis systems.
197 FEBS J 275:2385-2402.

198 9. Pettigrew GW, Aviram I, Schejter A. 1975. Physicochemical properties of two atypical
199 cytochromes *c*, *Crithidia* cytochrome *c*-557 and *Euglena* cytochrome *c*-558. Biochem J
200 149:155-167.

201 10. Ginger ML, Sam KA, Allen JWA. 2012. Probing why trypanosomes assemble atypical
202 cytochrome *c* with an AxxCH haem-binding motif instead of CxxCH. Biochem J
203 448:253-260.

204 11. Priest JW, Hajduk SL. 1992 Cytochrome *c* reductase purified from *Crithidia fasciculata*
205 contains an atypical cytochrome c_1. J Biol Chem 267:20188-20195.

12. Kořený L, Sobotka R, Kovářová J, Gnipová A, Flegontov P, Horváth A, Oborník M, Ayala FJ, Lukeš J. 2012. Aerobic kinetoplastid flagellate *Phytomonas* does not require heme for viability. Proc Natl Acad Sci USA 109:3808-3813.

13. san Francisco B, Bretsynder EC, Kranz RG. 2013. Human mitochondrial cytochrome *c* synthase's heme binding, maturation determinants, and complex formation with cytochrome *c*. Proc Natl Acad Sci USA 110:E788-E797.

14. Beneke T, Madden R, Makin L, Valli J, Sunter J, Gluenz E. 2017. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. R Soc Open Sci 4:170095.

15. van Hellemond JJ, van der Meer P, Tielens AG. 1997. *Leishmania infantum* promastigotes have a poor capacity for anaerobic functioning and depend mainly on respiration for their energy generation. Parasitology 114:351-360.

16. Diekert K, Kispal G, Guiard B, Lill R. 1999. An internal targeting signal directing proteins into the mitochondrial intermembrane space. Proc Natl Acad Sci USA 96:11752-11757.

17. Fülöp V, Sam KA, Ferguson SJ, Ginger ML, Allen JWA. 2009. Structure of a trypanosomatid mitochondrial cytochrome *c* with heme attached via only one thioether bond and implications for the substrate recognition requirements of heme lyase. FEBS J 276:2822-2832.

18. Zhang Y, Stevens JM, Ferguson SJ. 2014. Substrate recognition of holocytochrome *c* synthase: N-terminal region and CXXCH motif of mitochondrial cytochrome *c*. FEBS Lett 588:3367-3374.

19. Babbitt SE, Hsu J, Kranz RG. 2016. Molecular basis behind inability of mitochondrial holocytochrome *c* synthase to mature bacterial cytochromes: defining a critical role for cytochrome *c* α helix-1. J Biol Chem 291:17523-17534.

20. Butenko A, Hammond M, Field MC, Ginger ML, Yurchenko V, Lukeš J. 2020. Reductionist pathways for parasitism in euglenozoans? Expanded datasets provide new insights. Trends Parasitol doi:10.1016/j.pt.2020.10.001.

21. Tanifuji G, Cenci U, Moog D, Dean S, Nakayama T, David V, Fiala I, Curtis BA, Sibbald SJ, Onodera NT, Colp M, Flegontov P, Johnson-MacKinnon J, McPhee M, Inagaki Y, Hashimoto T, Kelly S, Gull K, Lukeš J, Archibald JM. 2017. Genome sequencing reveals metabolic and cellular interdependence in an amoeba-kinetoplastid symbiosis. Sci. Rep. 7:11688.

22. Wimplinger I, Shaw GM, Kutsche K. 2007. HCCS loss-of-function missense mutationin a female with bilateral microphthalmia and sclerocornea: a novel gene for severe ocular malformations. Mol Vis 13:1475-1482.

23. Wimplinger I, Morleo M, Rosenberger G, Iaconis D, Orth U, Meinecke P, Lerer I, Ballabio A, Gal A, Franco B, Kutsche K. 2006. Mutations of the mitochondrial holocytochrome *c*-type synthase in X-linked dominant microphthalmia with linear skin defects syndrome. Am J Hum Genet 79:878-899.

24. Babbitt SE, san Francisco B, Bretsnyder EC, Kranz RG. 2014. Conserved residues of the human mitochondrial cytochrome *c* synthase mediate interactions with heme. Biochemistry 53:5261-5271.

25. Mendez DL, Babbitt SE, King JD, D'Alessandro J, Watson MB, Blankenship RE, Mirica LM, Kranz RG. 2017. Engineered holocytochrome *c* synthases that biosynthesize new cytochromes *c*. Proc Natl Acad Sci USA 114:2235-2240.

26. Babbitt SE, Hsu J, Mendez DL, Kranz RG. 2017. Biosynthesis of single thioether *c*-type cytochromes provides insight into mechanisms intrinsic to holocytochrome *c* synthase (HCCS). Biochemistry 46:3337-3346.

27. Hirota S, Hattori Y, Nagao S, Taketa M, Komori H, Kamikubo H, Wang Z, Takahashi I, Negi S, Sugiura Y, Kataoka M, Higuchi Y. 2010. Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix. Proc Natl Acad Sci USA 107:12854-12859.

FIGURE LEGENDS

FIG 1 (A) KCCS is a novel protein exhibiting very little sequence similarity to HCCS. MAFFT was used to align *T. brucei* (Tb) KCCS with HCCS from diverse taxa. HCCS-defining motifs I-IV are boxed; His154 is denoted by a red asterisk; mutations seen in MLS patients are indicated by blue asterisks; grey asterisks indicate residues analysed by site-directed mutagenesis of *Hs*HCCS (**10**, **21**); circles denote where site-directed mutation decreased HCCS activity. Cr, *Chlamydomonas reinhardtii* (XP_001697002.1); Hs, *Homo sapiens* (NP_001116080.1); Sc, *Saccharomyces cerevisiae* (NP_009361.1). NCBI reference sequences are provided in the parentheses. (B-C) *Tb*KCCS-catalysed maturation of trypanosome cytochrome *c* (*Tb*CYTC). (B) Soluble fractions from *E. coli* induced for expression (described

269 in Text S1) of either His_6-tagged *Tb*CYTC or His_6-tagged *Tb*CYTC plus *Tb*KCCS from
270 pCDFDuet-1 (Novagen) were purified by Ni^{2+}-affinity chromatography. Flow-through (FT)
271 and elution fractions (E6, E7) from each culture subject to acetone precipitation and
272 analysed by SDS-PAGE under non-reducing conditions. Duplicate 12% gels were stained
273 with either Instant Blue (to confirm protein loading) or 3,3',5,5'-tetramethylbenzidine (TMB,
274 to detect covalent attachment of heme to protein). 50 ng of equine holocytochrome *c* was
275 loaded in the control lane of both gels. Dimerization of cytochrome *c*, evident in all lanes is
276 explained by 'domain-swapping' of the C-terminal α-helix (**27**, Text S1). (C) Pyridine
277 hemochrome spectra for recombinant trypanosome (dashed line) or equine heart
278 cytochrome *c* (solid line) were recorded at 25°C following disodium dithionite addition.
279 Concentration of cytochromes analysed was 13 μM; spectra were normalised by Soret band
280 intensity; the inset expands the 500-600 nm region of the two spectra, indicating the
281 diagnostic α-band maximum at 553 nm for cytochrome *c* with heme bound by a single
282 thioether bond, red-shifted relative to the 550 nm α-band maximum for cytochrome *c* with
283 a CxxCH heme-binding motif. Instant blue- and TMB-stained gels of the purified
284 cytochrome preparation used for spectroscopy are shown in **Fig. S2**.

285

286 **FIG 2** (A) Mitochondrial localisation of *Lm*KCCS::mNeonGreen in live, CyGEL^{TM}-immobilized
287 *L. mexicana*. (B-E) CRISPR-Cas9 genome editing of *L. mexicana* reveals *Lm*KCCS is an
288 essential gene. (B) Homologous recombination of drug-resistance cassettes into Lm*KCCS*
289 loci with amplicons from PCR-mapping indicated. P1, amplicon within the Lm*KCCS* coding

sequence (or CDS); P2, downstream and within Lm*KCCS*; P3, downstream of Lm*KCCS* and within *PUR* or *BSD* resistance cassettes. Diagnostic PCR from gDNA templates extracted from *Leishmania* populations (X and Y) after transfection with sgDNA and template donor DNA for CRISPR-Cas9 gene editing. (C) Analysis of DNA content in propidium iodide-stained cell populations T7Cas9, X and Y by flow cytometry. (D) Episomal expression of Lm*KCCS*::*GFP* (EC-L) with PCR amplicons for P1 and P4 or (E) episomal expression of Tb*KCCS*::*GFP* (EC-T) and PCR amplicons for P5 and P6 indicated. Also shown, PCR mapping of T7Cas9 parental *L. mexicana* and FKO clones LX and LY (D) or TX and TY (E) together with immunoblot analysis of *Lm*KCCS::GFP or *Tb*KCCS::GFP expression in wild-type *L. mexicana*, episome-transfected *L. mexicana*, and FKO clones. For the loading control on the immunoblots, expression of oligopeptidase B (anti-OPB) was detected.

Guide to the Supplemental Material

Text S1 The complete description of methods covers (i) bioinformatics; (ii) biochemical validation of KCCS candidature (including details of plasmid construction, recombinant protein expression and purification, analysis of cytochrome *c* maturation by SDS-PAGE, 3,3',5,5'-tetramethylbenzidine-staining and uv/vis spectroscopy); (iii) explanation of the cytochrome *c* dimerization seen in Fig. 1B; (iv) genetic manipulation of *Leishmania mexicana*; (v) analysis of CRISPR-Cas9 *L. mexicana* mutants by flow cytometry; and (vi) analysis of KCCS localisation by fluorescence microscopy.

Supplemental Figure legends

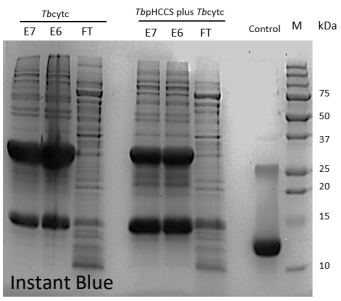
FIG S1 KCCS is a conserved kinetoplastid protein. Ad, *Angomonas deanei* (EPY32355.1); Bs, *Bodo saltans* (CUF09763.1); Lm, *Leishmania mexicana* (XP_003872427.1); Pk_sp, *Perkinsela* (KNH04224.1); Tb, *Trypanosoma brucei* (XP_843981.1); Tbrr, *Trypanoplasma borreli* (). Genbank accession numbers for the sequences used in the alignment are provided in the parentheses.

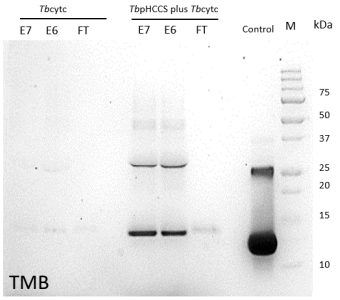
FIG S2 Purity of and covalent heme attachment to *Tb*^{His}CYTC purified for uv/vis spectroscopy. *Tb*^{His}CYTC was purified from 8 l of *E. coli* induced for recombinant expression of *Tb*^{His}CYTC and *Tb*KCCS as described in Text S1 and concentrated to 0.5 ml using a Vivaspin-20 centrifugal concentrator with a m.w. cut-off of 3 kDa. 1/250th of the purified protein was taken without acetone precipitation for analysis by SDS-PAGE under non-reducing conditions. Duplicate 12% polyacrylamide gels were stained with either Instant Blue (to confirm purity) or 3,3',5,5'-tetramethylbenzidine (to detect covalent attachment of heme to protein). Equine holocytochrome *c* was loaded as indicated.

A

```
Hs                MGLSPSAPAVAVQASNASASPPSGCPMHEGKMKGCPVNTEPSGPTCEKKTY   51
Sc              MGWFWADQKTTGKDIGGAAVSSMSGCPV----------MHESSS--------   34
Cr      MGNQQSASAPPPATSAAPCAEAAAAAGAEPPSSCPVNPKY----------KNPAV----Y   46
Tb              MWVRTFLRLCGCKSPNAAAITS-GS----------SWMTA--------AAW   32

Hs      SVPAHQERAYE----YVECPIRGTAAENKENLDPSNLM-PPPNQ---TPAPDQPFALSTV  103
Sc      -------SSPP----SSECPVMQ---GDNDRINPLNNMP-ELAA---SKQPGQKMDLPVD   76
Cr      NVYGQRINDPN----SQAKPSPLASITGADVLDPKNNMPLEPNQ---LPCPGQRKPLSTE   99
Tb      ASLGSEFSSVSESKFLQQVPDGF-----LTSRATTDMMPAEQLLLSMVEENEERYKGVDV   87

Hs      REESSIPRA-DSEKKWVYPSEQMFWNAMLKKGWKWKDEDISQKDMYNIIRIH--NQNNEQ  160
Sc      RTISSIPKSPDSNEFWEYPSPQQMYNAMVRKGKIGGSGEVAEDAVESMVQVH--NFLNEG  134
Cr      RVASNIPKG-GTESTWLFPSPQMVFNALKRKG---KGDDVTEDDMDGFIAAH--NSMNEA  153
Tb      RD------------------PSSMAVYEG-ERPRWMTLGGQVRA--VSEFVSSHLCHHISLP  128

Hs      AWKEILK----------WEA-LHAAECPCGPS---LIRFGGKAKEYSPRARIRSWMG---  203
Sc      CWQEVLE----------WEKP-HTDESHVQPK---LLKFMGKPGVLSPRARWMHLCGLLF  178
Cr      TWQRVAQ----------WEM-LHRGECDT-PT---LLRFQGKPHDLSPLAWVRHMLG---  195
Tb      AWKELFDLQYAEMDLTYWLYVLHVHMVSRRATSVPIEKFNRRREVLEE----------IL  176

Hs      -----YELPFDRHDWIINRCG----------TEVRYVIDYYDGGEVNKDYQFTILDVRPA  246
Sc      PSHFSQELPFDRHDWIVLRGERKAEQQPPTFKEVRYVLDFYGGPDDENGMPTFHVDVRPA  238
Cr      -----GPAPFDRHDWVIDRCG----------KEVRYIIDFYFFDDKAGTPQAFEIVARPA  238
Tb      LTM------FD--SWAATSEDV--MGRPPLNKIRFYIKDMYYVTAVNFEEALL--HDGPG  224

Hs      LDSLS---------------------------------AVWDRMKVAWWRWTS         268
Sc      LDSLD---------------------------------NAKDRMTRFLDRMISGPSSSSSA 268
Cr      VDSVE---------------------------------AALDRVKMNIYLKFAEWGLPCPI 268
Tb      ADLMLLGFLMKFCPLPRPEDVPLYTYYSLVHYIRFHTALLDRIPDESIAKGNFNFLSPTD  286

Hs
Sc      P                                                       269
Cr      TGQAGAVAQAAAAAGGQQAASGSS                                292
Tb      PRIFEQYSEVTLDQVIRSWTVEASEEEVKCHAAP                      320
```

B





C

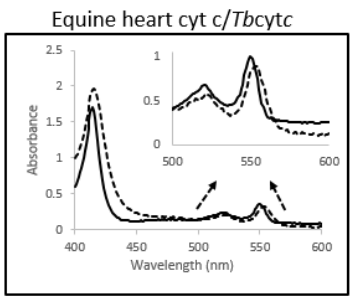


FIG 1

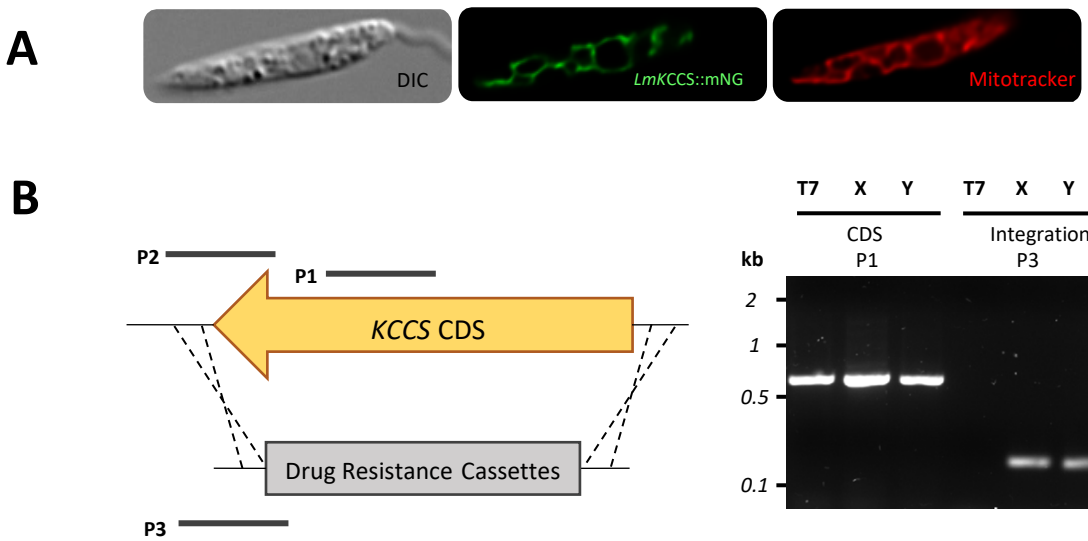

A
DIC
LmKCCS::mNG
Mitotracker
B
P2
P1
KCCS CDS
Drug Resistance Cassettes
P3
T7 X Y
CDS
P1
T7 X Y
Integration
P3
kb
2
1
0.5
0.1
C
T7Cas9
X
Y
Counts
R3 R4
PE-Cy5 Lin
D
EC-L
P4
P1
T7Cas9
FKO LX
FKO LY
kb
P1 P2 P3 P4
2
1
0.5
0.1
FKO
kDa
WT EC- L L X LY
75
50
anti-GFP
100
75
anti-OPB
E
EC-T
P6
P5
T7Cas9
FKO TX
FKO TY
kb
P1 P2 P3 P5 P6
2
1
0.5
0.1
FKO
kDa
EC- T TX TY
75
50
anti-GFP
100
75
anti-OPB

FIG 2

Extreme divergence of the kinetoplastid cytochrome *c* maturation system

Asma Belbelazi, Rachel Neish, Martin Carr, Jeremy C. Mottram, and Michael L. Ginger

TEXT S1 Materials and Methods

Bioinformatics. KCCS was identified from a bioinformatics sift of proteins present in *T. brucei* mitochondrial proteomes (**1**, **2**). 295 proteins identified as without orthologues in *Phytomonas* were then analysed manually for presence of any possible motifs or similarity to motifs present in any protein previously characterised as involved in any of the four biogenesis systems known to catalyse thioether bond formation between heme and a cysteine sulphydryl (**3**-**6**). Disorder predictions were made using IUPredA and MFDp2 (**7**, **8**). Multiple sequence alignment (MAFFT) was used to align peptide sequences (**9**).

Biochemical validation of KCCS candidature. For expression of recombinant cytoplasmic *Tb*HCCS and/or *Tb*CYTC, CDSs were sub-cloned into pCDFDuet-1 (Novagen): Tb*CYTC* was sub-cloned into HindIII-EcoRI-digested multiple cloning site 1 (or MCS1), *Tb*HCCS was sub-cloned into XhoI-NdeI-digested multiple cloning site 2 (or MCS2). Forward and reverse primer combinations for PCR amplication of *Tb*HCCS or *Tb*CYTC, CDSs, respectively, from genomic DNA templates were tt*gaattc*gcatgccaccaaaggagcgtgc and gc*aagctt*ttagtcctttaatgtctcgagg or cac*catatg*tgggtgaggacattcctgc and aag*ctcgag*tcacggtgccgcatggcattttac. Restriction sites introduced into the primers are italicised. Recombinant protein expression of *Tb*HCCS, His_6-TbCYTC, or *Tb*HCCS and His_6-*Tb*CYTC was induced in *E. coli* Rosetta (Novagen) by addition of 1mM IPTG; induced cultures

were allowed to grow for 24 h at 18°C at 100 rpm under aerated conditions. *E. coli* cultures were grown in Luria broth (Melford L24400-500.0) without addition of exogenous heme. Following induction of recombinant protein expression, bacterial cells were collected by centrifugation and re-suspended in 20 ml lysis buffer (Tris-HCl (20 mM pH 8.0); NaCl (500 mM); Triton X-100 (0.02% v/v); imidazole (20 mM); glycerol (10% v/v)) per l culture. Protease inhibitor PMSF (100 mM) was immediately added (10 µl per ml of resuspended cells) and the resuspension left shaking (50 rpm, room temperature, 30 min). After this incubation, lysing cells were subject to further disruption by ultra-sonication using a burst frequency of 5 sec on/15 sec off for 10 min at an amplitude of 85%. Following sonication, the suspension was centrifuged at 15 000 x *g* (30 min; 4°C) and the supernatant stored at -20°C prior to protein purification. Apo- and holocytochromes *c* were purified by Ni^{2+}-affinity chromatography using Amintra Ni-NTA resin (1 ml per l culture harvested) under native conditions using wash (Tris-HCl (20 mM pH 8.0); NaCl (300 mM); Triton X-100 (0.02% v/v); imidazole (20 mM); glycerol (10% v/v)) and elution (Tris-HCl (20 mM pH 8.0); NaCl (300 mM); Triton X-100 (0.02% v/v); imidazole (500 mM); glycerol (10% v/v)) buffers. Purified cytochromes were concentrated using a Vivaspin-20 centrifugal concentrator with a m.w. cut-off of 3 kDa Prior to SDS-PAGE proteins were typically subject to acetone precipitation; SDS-PAGE was carried out under non-reducing conditions; and prior to gel-loading samples were heated to 95°C for 5 min in loading buffer containing SDS (2%); glycerol (10%); bromophenol blue (0.01%); and Tris-HCl (100 mM pH 6.5). Heme-staining of SDS-PAGE gels using 3,3',5,5'-tetramethylbenzidine (TMB) was carried out as described previously (**10**). Pyridine hemochrome spectra were acquired using a Cary 4000 uv/vis spectrophotometer

following the protocol laid out by Barr and Guo (**11**). In our experiments purified recombinant *Tb*CYTC or horse heart cytochrome *c* were soluble in 50 mM Tris-HCl (pH 7.5). To measure oxidized heme spectra, protein solutions were mixed 1:1 with a solution of NaOH (0.2 M), pyridine (40% v/v) and $K_3Fe(CN)_6$ (500 µM) in a final volume of 1 ml; to subsequently measure reduced heme spectra, 10 µl of $Na_2O_4S_2$ (0.5 M) in 0.5 M NaOH was added to protein/pyridine/ $K_3Fe(CN)_6$ solutions. Difference (reduced minus oxidized) spectra are shown in Fig. 1C.

Explanation of cytochrome *c* dimers. From Fig. 1B: oligomerization of mitochondrial cytochrome *c* has been known for almost 60 years (**12**) and is evident in cytochrome *c* preparations from the 1940s (**13**). Dimerization results from displacement of the C-terminal α-helix from monomeric cytochrome *c* and replaced by the corresponding (also displaced) helix from another cytochrome *c* molecule (**14**). This readily explains the dimerization evident in all lanes where cytochrome *c* is present in the Instant Blue-stained gel from Fig. 1B albeit that there is a large proportion of dimer relative to monomer purified from *E. coli* expressing either *Tb*cytc or *Tb*cytc plus *Tb*KCCS. Where *Tb*KCCS is expressed simultaneously with its apo-cytochrome substrate, comparison of TMB- and Instant Blue-stained gels indicates only a proportion of trypanosome cytochrome is present in the holo-state (note the intensity of the TMB stain for equine cytochrome *c* relative to the amount of protein loaded). This we ascribe to the absence of additional, exogenous heme from the LB medium used to culture our *E. coli* and/or low expression of soluble *Tb*KCCS relative to apo-*Tb*cytc from the pCDFDuet-1 expression plasmid. Potentially, an extremely faint detection of heme is detected in lane E6 for *E. coli* expressing Tb*CYTC* but not Tb*KCCS*; we suggest if TMB-

staining is present it possibly reflects coordination of heme iron by the N-terminal hexa-histidine tag in the recombinant trypanosome cytochrome *c*.

Cell culture and transfection. *L. mexicana* (M379-T7Cas9) promastigotes were grown in HOMEM medium (Gibco) supplemented with 10 % heat inactivated fetal calf serum (Gibco) with 1 % Penicillin/Streptomycin (Sigma-Aldrich) at 25ºC. Transfections were carried out using the 4D Nucleofactor™ (Lonza).

Gene deletion studies of *Lm*KCCS were carried out using a CRISPR-Cas9 toolkit in *L. mexicana* promastigotes genetically modified for constitutive expression of Cas9 nuclease and T7 RNA polymerase (**15**). This was done by providing resistance cassettes using primers with homology sites of 30 nucleotides upstream (F: ACGTCGATTCGCACGACGTCCACAAGGAGAgtataatgcagacctgctgc) and downstream (R: CTTGGCCAGCGCTGCAGAAAGGGAAAGCGGccaatttgagagacctgtgc) of the break site and were amplified from pPLOT plasmids. Guide DNA primers were used as a template for the sgRNA induced break site (5' gaaattaatacgactcactataggGGCGGTAATTGTGGCGGCAG gttttagagctagaaatagc and 3' gaaattaatacgactcactataggAGCGGTACCACACGAGCGCG gttttagagctagaaatagc).

pNUS-GFPcN (**16**) was used for episomal expression of *Lm*KCCS or *Tb*HCCS, C-terminally tagged with GFP. For cloning *Lm*KCCS, the fragment was amplified using forward (cacttgtcaagcgaattccatatgATGGCGGGGGCGGCGTGG) and reverse primer (gctcatggtaccagatctcatatgCGAGTGAGGCGGTGCCGCTTC). For cloning *Tb*HCCS, the fragment was amplified using forward (cacttgtcaagcgaattccatatgATGTGGGTGAGGACATTCCTG) and

reverse primer (gctcatggtaccagatctcatatgCGGTGCCGCATGGCATTT). The PCR products were cloned into pNUS-GFPcN using Gibson assembly (NEB) as per manufacturer's instructions. Following confirmation of episomal *Lm*KCCS::GFP (EC-L in **Fig. 2D**) or *Tb*HCCS::GFP (EC-T in **Fig. 2E**) expression, CRISPR-Cas9 was again used for gene deletion of Lm*KCCS*. PCR was used as a diagnostic tool to confirm facilitated knock out (FKO) using the following primer sets P1, (forward primer) TGGGCAGCGAGTTCAAGAAT and (reverse primer) AAGCGGGCAGAACTTCATCA (amplicon size 0.62 kb); P2, TGTCATTGTGACAGTGC and AAGCGGGCAGAACTTCATCA (1.027 kb); P3, TGTCATTGTGACAGTGC and GCAGCAGGTCTGCATTATAC (0.178 kb); P4, TGGGCAGCGAGTTCAAGAAT and GCATCACCTTCACCCTCTCC (1.025 kb); P5, CGGGTGCAAATCACCCAATG and CCTCCGAGTGTCATCCATCG (0.294 kb); P6, CGGGTGCAAATCACCCAATG and GCATCACCTTCACCCTCTCC (1.088 kb). As the protein loading control for immunoblot analysis of *Lm*KCCS::GFP or *Tb*KCCS::GFP expression, polyclonal antibodies detecting oligopeptidase B (anti-OPB) (**17**) were used.

Flow cytometry. Logarithmic cells (~$5x10^6$-10^7 cells ml^{-1}) were washed in PBS and resuspend in 70 % methanol for 30 minutes. Cell were pelleted and washed with PBS containing 10 $\mu g ml^{-1}$ propidium iodide and 10 $\mu g ml^{-1}$ RNAase A. Fluorescence was measured using the PE-Cy5-Lin channel on Cyan and analyses carried using Summit V4 software (Beckman Coulter). Gating included all singlet cells.

KCCS localisation. Live imaging of promastigotes expressing *Lm*KCCS::mNeonGreen was carried out using CyGEL^{TM} (Biostatus) to immobilize cells. Samples were imaged

immediately using a Zeiss AxioObserver microscope with 488 and 405 nm lasers. Images were processed using Zen Black (Zeiss) and Microvolution™ deconvolution software.

REFERENCES

1. Acestor N, Panigrahi AK, Ogata Y, Anupama A, Stuart KD. 2009. Protein composition of *Trypanosoma brucei* mitochondrial membranes. Proteomics 9:5497-5508.
2. Niemann M, Wiese S, Mani J, Chanfon A, Jackson C, Meisinger C, Warscheid B, Schneider A. 2013. Mitochondrial outer membrane proteome of *Trypanosoma brucei* reveals novel factors required to maintain mitochondrial morphology. Mol Cell Proteomics 12:515-528.
3. Babbitt SE, Sutherland MC, san Francisco B, Mendez DL, Kranz RG. 2015. Mitochondrial cytochrome *c* biogenesis: no longer an enigma. Trends Biochem Sci 40:446-455.
4. Verissimo AF, Daldal F. 2014. Cytochrome *c* biogenesis system I: an intricate process catalysed by a maturase supercomplex? Biochim Biophys Acta 1837:989-998.
5. Simon J, Hederstedt L. 2011. Composition and function of cytochrome *c* biogenesis system II. FEBS J 278:4179-4188.
6. Saint-Marcoux D, Wollman FA, de Vitry C. 2009. Biogenesis of cytochrome b_6 in photosynthetic membranes. J Cell Biol 185:1195-1207.

7. Mészáros B, Erdős G, Dosztányi Z. 2018. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acid Res 46:W329-W337.

8. Mizianty MJ, Uversky V, Kurgan L. 2014. Prediction of intrinsic disorder in proteins using MFD2p. Methods Mol Biol 1137:147-162.

9. Katoh K, Standley DM. 2013. MAFFT multiple alignment software version 7: improvements in performance and sustainability. Mol Biol Evol 30:772-780.

10. Thomas PE, Ryan D, Levin W. 1976. An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. Anal Biochem 75:168-176.

11. Barr I, Guo F. 2015. Pyridine hemochromagen assay for determining the concentration of heme in purified solutions. Bio Protoc 5:e1594.

12. Margoliash E, Lustgarten J. 1962. Interconversion of horse heart cytochrome *c* monomer and polymers. J Biol Chem 237:3397-3405.

13. Paul K-G. 1948. The stability of cytochrome *c* at extreme pH values. Acta Chem Scand 2:430-439.

14. Hirota S, Hattori Y, Nagao S, Taketa M, Komori H, Kamikubo H, Wang Z, Takahashi I, Negi S, Sugiura Y, Kataoka M, Higuchi Y. 2010. Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix. Proc Natl Acad Sci USA 107:12854-12859.

15. Beneke T, Madden R, Makin L, Valli J, Sunter J, Gluenz E. 2017. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. R Soc Open Sci 4:170095.

16. Tetaud E, Lecuix I, Sheldrake T, Baltz T, Fairlamb AH. 2002. A new expression vector for *Crithidia fasciculata* and *Leishmania*. Mol Biochem Parasitol 120:195-204.

17. Munday JC, McLuskey K, Brown E, Coombs GH, Mottram JC. 2011. Oligopeptidase B deficient mutants of *Leishmania major*. Mol Biochem Parasitol 175:49-57.

Extreme divergence of the kinetoplastid cytochrome *c* maturation system

Asma Belbelazi, Rachel Neish, Martin Carr, Jeremy C. Mottram, and Michael L. Ginger

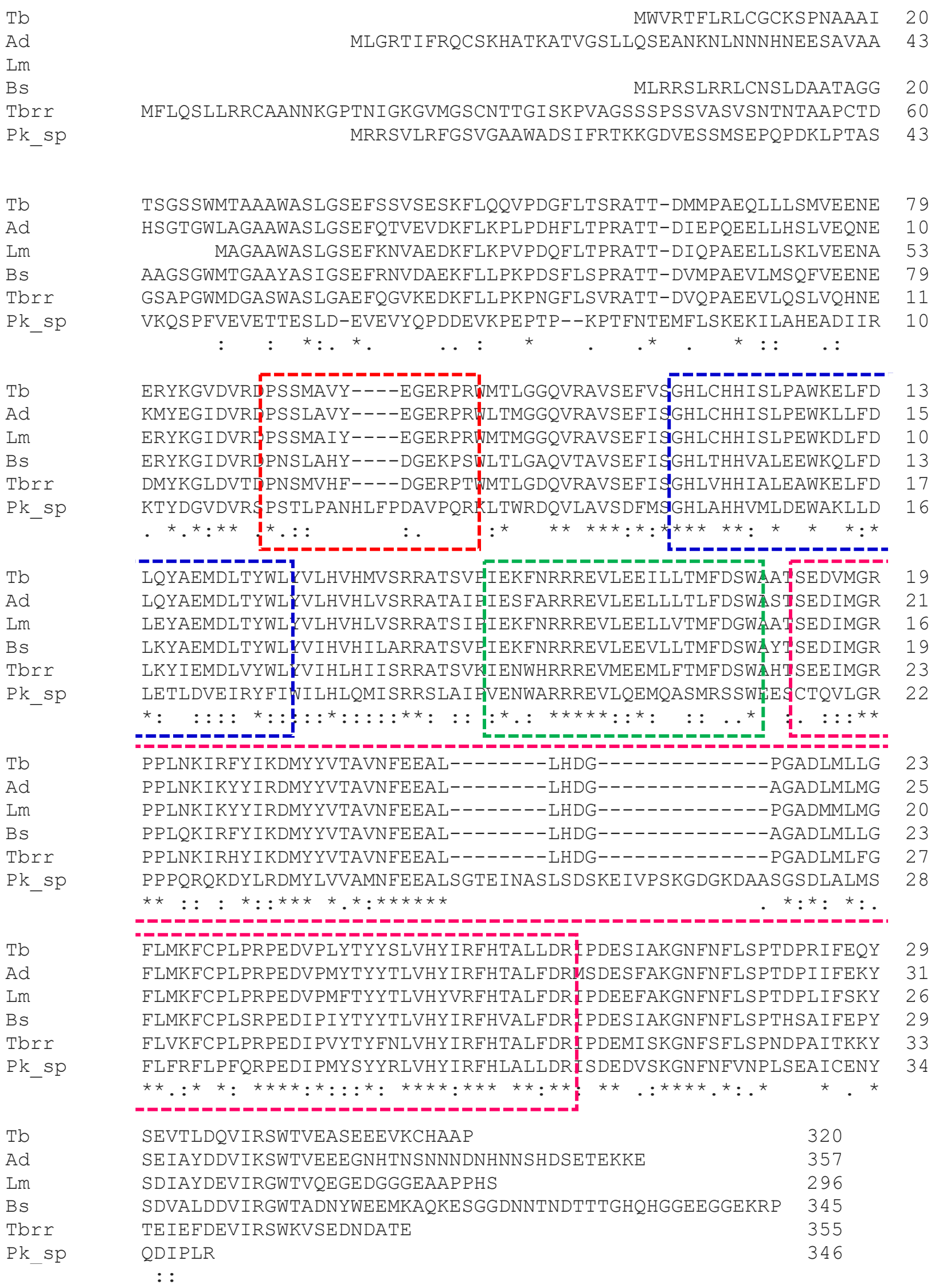
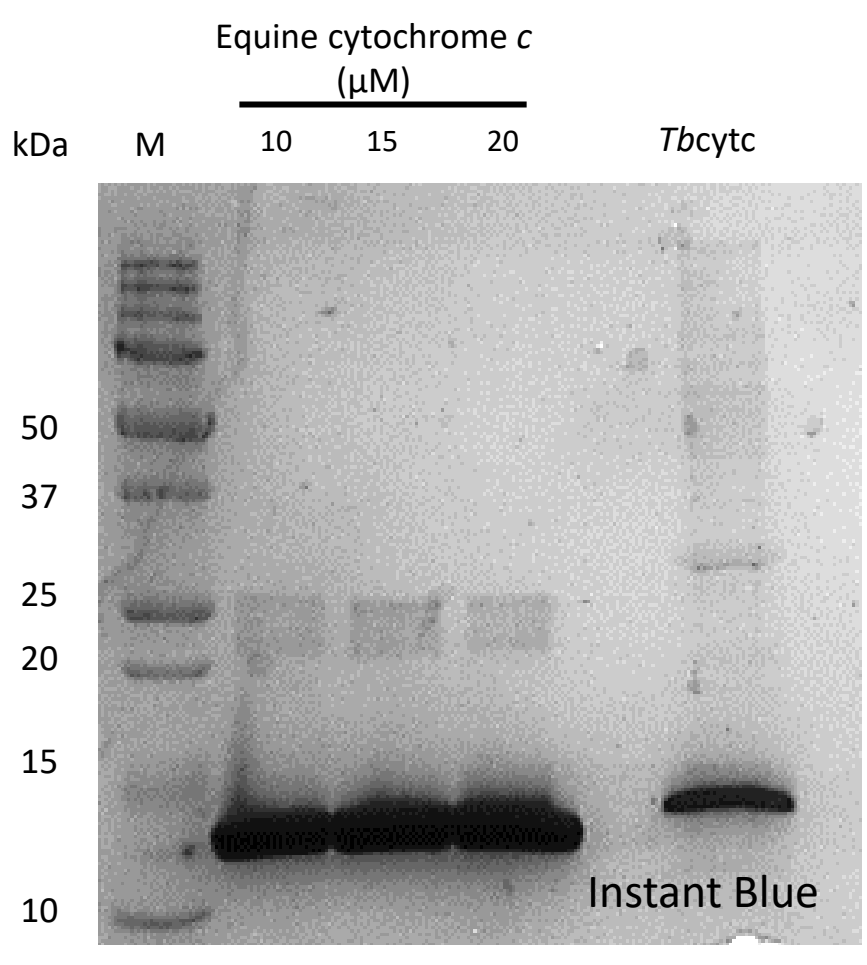


FIG S1 KCCS is a conserved kinetoplastid protein. Ad, *Angomonas deanei* (EPY32355.1); Bs, *Bodo saltans* (CUF09763.1); Lm, *Leishmania mexicana* (XP_003872427.1); Pk_sp, *Perkinsela* (KNH04224.1); Tb, *Trypanosoma brucei* (XP_843981.1); Tbrr, *Trypanoplasma borreli* (). Genbank accession numbers for the sequences used in the alignment are provided in the parentheses.



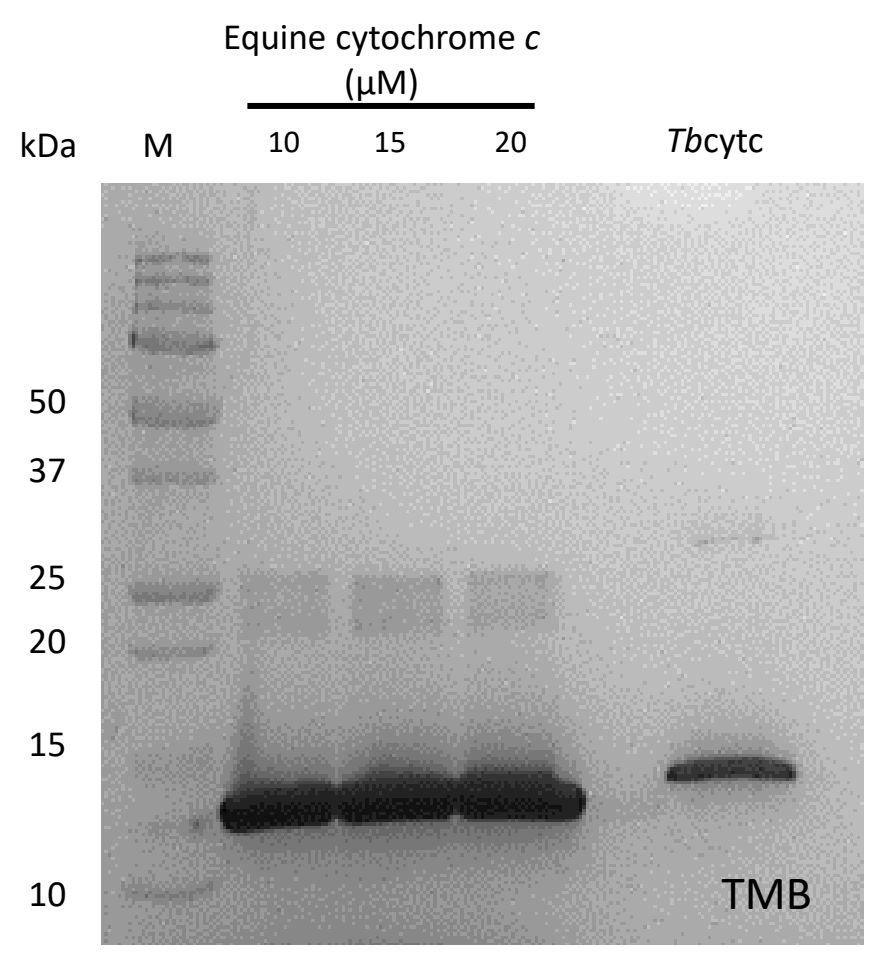


FIG S2 Purity of and covalent heme attachment to *Tb*^{His}CYTC purified for uv/vis spectroscopy. *Tb*^{His}CYTC was purified from 8 l of *E. coli* induced for recombinant expression of *Tb*^{His}CYTC and *Tb*KCCS as described in Text S1 and concentrated to 0.5 ml using a Vivaspin-20 centrifugal concentrator with a m.w. cut-off of 3 kDa. 1/250th of the purified protein was taken without acetone precipitation for analysis by SDS-PAGE under non-reducing conditions. Duplicate 12% polyacrylamide gels were stained with either Instant Blue (to confirm purity) or 3,3′,5,5′-tetramethylbenzidine (to detect covalent attachment of heme to protein). Equine holocytochrome *c* was loaded as indicated.