

Nucleotide excision repair in *Trypanosoma brucei*: specialization of transcription-coupled repair due to multigenic transcription.

Carlos R. Machado ^{1#*}, João P. Vieira-da-Rocha ^{1#}, Isabela Mendes ^{1,3}, Matheus A. Rajão ², Lucio Marcello ³, Mainá Bitar ¹, Marcela G. Drummond ⁵, Priscila Grynberg ¹, Denise A. A. Oliveira ⁵, Catarina Marques ³, Ben Van Houten ⁴, Richard McCulloch ^{3*}

¹ Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Caixa Postal 486, Belo Horizonte 30161-970, MG, Brazil.

² Coordenação de Pesquisa, Instituto Nacional de Câncer, Rua André Cavalcanti, 37 Fátima 20231-050 - Rio de Janeiro, RJ, Brazil.

³ The Wellcome Trust Centre for Molecular Parasitology, College of Medical, Veterinary and Life Sciences, Institute of Infection, Immunity and Inflammation, University of Glasgow, Sir Graeme Davies Building, 120 University Place, Glasgow, G12 8TA, UK.

⁴ Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine and The University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, PA 15213, USA.

⁵ Laboratório de Genética Animal, Escola de Veterinária, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Caixa Postal 486, Belo Horizonte 30161-970, MG, Brazil.

#These authors contributed equally to this work.

* Corresponding authors: richard.mcculloch@glasgow.ac.uk; FAX: +44(0)141 330 8269 and crmachad@icb.ufmg.br; FAX: (5531) 3409-2984

Supporting Figure legends

Supporting figure 1: *T. brucei* XPC. The upper diagram shows predicted conserved domains in TbXPC when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a sequence alignment of TbXPC with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 2: *T. brucei* RAD23. The upper diagram shows predicted conserved domains in TbRAD23 when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a sequence alignment of TbRAD23 with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 3: *T. brucei* DDB. A sequence alignment of TbDDB (only the larger DDB subunit could be identified) with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 4: *T. brucei* CSB. The upper diagram shows predicted conserved domains in TbCSB when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a sequence alignment of TbCSB with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 5: *T. brucei* XPD. The upper diagram shows predicted conserved domains in TbXPD when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a sequence alignment of TbXPD with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 6: *T. brucei* XPB homologues. The upper diagrams shows predicted conserved domains in TbXPB and TbXPBz when the polypeptides are used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a neighbor-joining tree of the two *T. brucei* XPB homologues with related proteins from other eukaryotes and from an archaea.

Supporting figure 7: *T. brucei* XPG. The upper diagram shows predicted conserved domains in TbXPG when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The middle diagram shows a summary of the comparative sizes and predicted conserved domains in XPG from *T. brucei* and from *H. sapiens*. The lower diagram shows a sequence alignment of TbXPG with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 8: *T. brucei* XPF. The upper diagram shows predicted conserved domains in TbXPF when the polypeptide is used to interrogate the proteins in the NCBI database by BLASTp analysis. The lower diagram shows a sequence alignment of TbXPF with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 9: *T. brucei* ERCC1. A sequence alignment is shown of TbERCC1 with related proteins from other eukaryotes; residues identical or conserved in >50% of sequences are shaded black or grey, respectively.

Supporting figure 10: TbCSB depletion increases UV-resistance for a short period post-irradiation. Survival curve of TbCSB depleted cell line and its uninduced control, 4 hours after treatment with 1500 J/m² of UVC. Percent survival was calculated as the relative percentage of cells present in the treated groups compared with the untreated control. Values are means of three experiments, and vertical lines denote standard deviation.

Table S1: Primers used in RNAi construct production and mRNA level quantification.

Primer	Sequence
XPBrnai_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTAGCCACTGTTAAGAAGCGGA
XPBrnai_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCCACTCGAGCAAGGAAAC
XPBzrna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCGACGCTTGGTTTTACAC
XPBzrna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGCCTACTAACCTTTGCAGCG
XPCrna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTTGTGGGTAACCGCTAA
XPCrna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGCGCCACAGAAAAGGTGTAT
XPDrna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGACGGAGCTCAAGAAAGGT
XPDrna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTACCTTTAGCGCCTTCCTAGC
XPERnai_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGCACCAGGGTAGTTTTGC
XPERnai_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCTAGAGTCGGCCAAGCAT
XPFrna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAATGGGAGCGATGGACTAA
XPFrna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCACTAAACGAACGGCATT
ERCC1rna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGTGTTGTGCTTCTTCTCCG
ERCC1rna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATAAATGCAGCGAAGTGC
XPGrna_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCTGGTGAAAGTGTGGAGG
XPGrna_R	GGGGACCACTTTGTACAAGAAAGCTGGGTAACCTCCCGAGTTCGTTTTT
CSBrnai_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCGGAGAAATACGAACGTG
CSBrnai_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTGTTACTGCGCATCGAAAA
XPCrnaq_F	GAAGTCGCTGCAGCCAAGA
XPCrnaq_R	GCGTCAGCGAAGTCATCAAG
XPEFrnaq_F	GAATTGATTTCCGTGACGTTTCT
XPERrnaq_R	CGGCGACCGCATTTC
XPBzrnaq_F	TAACGTTAACATTGCGTTGAAGTCT
XPBzrnaq_R	GCGTCAACAGCAGCGATCT
XPBrnaq_F	AGTTGGTGGAGGGCAATCC
XPBrnaq_R	AGCAAGCGGTCCACAACAG
XPDrnaq_F	GACTGACCAGCGTTGTGTGACTT

XPDrnaq_R	CGGGCGGAGATCAAAGC
XPGrnaq_F	CCAGCGCCGTCAACCA
XPGrnaq_R	AAGTCGTCGTTGAGCAGAATCA
XPFrnaq_F	TCCCGACTGGACCCCTTATC
XPFrnaq_R	GGGTTGACCTCCGGAACAG
ERCC1rnaq_F	CGGAACTGCCGGGTTTC
ERCC1rnaq_R	CGTTAAGTACGGCGTGAAGTCTT
CSBrnaq_F	GAGCACTGCCGGTTGCTT
CSBrnaq_R	TGCCCTCGTCAAGAATAACA
HIPB1rnaq_F	ACGCTGCAAATTATGCTTTACG
HIPB1rnaq_R	GGCACGTCTATCTCCAAAGAAAA
MANB2rnaq_F	CGGCTCGCTGTGGATCTT
MANB2rnaq_R	CCCAGTGGGCGTGTCAA
TCPB3rnaq_F	GACCTCGCGGCTCACTTC
TCPB3rnaq_R	CTTACGCAAACGGCGAATG
HIP2B4rnaq_F	GCACAGGCGGGAAACG
HIP2B4rnaq_R	ATCCCCATGTCGGCAGAA
ACTA1rnaq_F	CGGCTCGTGGGCTTTAAAG
ACTA1rnaq_R	ACGCCGCCGTAAGTTCTG
DNATA3rnaq_F	CGTGAGGAGTTGCCACAGAA
DNATA3rnaq_R	TACGGCACCCACTCCTGAA
HIP3A4rnaq_F	GTTGAGCGAATGTGCATGCT
HIP3A4rnaq_R	CGCGCCGAAGTGTAAGC
HIP4A5rnaq_F	CGTTGAAGACCGCGCAGTA
HIP4A5rnaq_R	GAGGCGCAGTTCAAGCATCT
GPI8rnaq_F	TCTGAACCCGCGCACTTC
GPI8rnaq_R	CCACTCACGGACTGCGTTT

All primer names contain the name of target gene (XPE refers to DDB). “rna” label is designed for primers used in production of RNAi constructs, while “rnaq” label is for all primers utilized in mRNA quantification analysis. “_F” and “_R” mean primer forward and reverse, respectively.

T. brucei RAD23

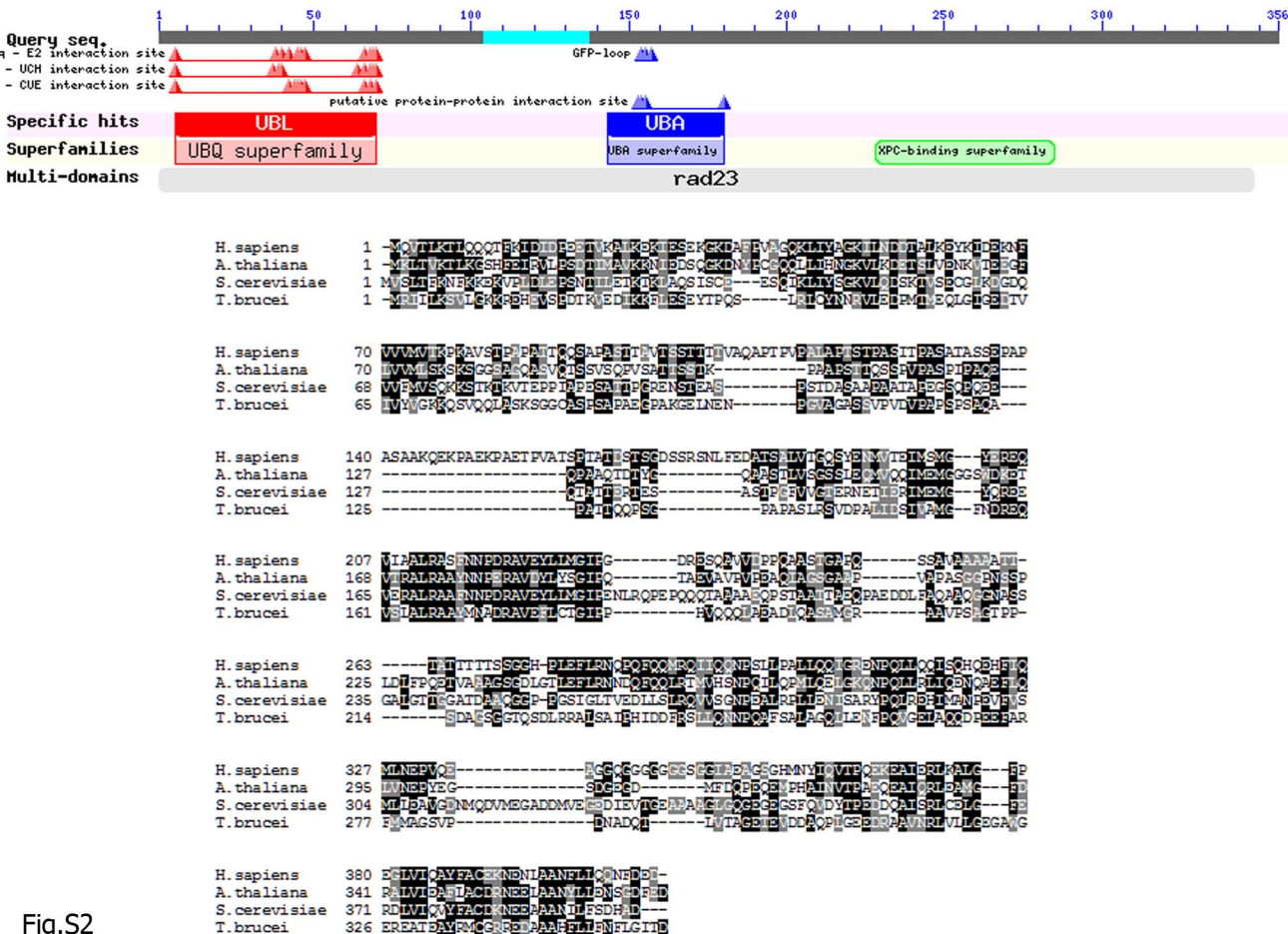
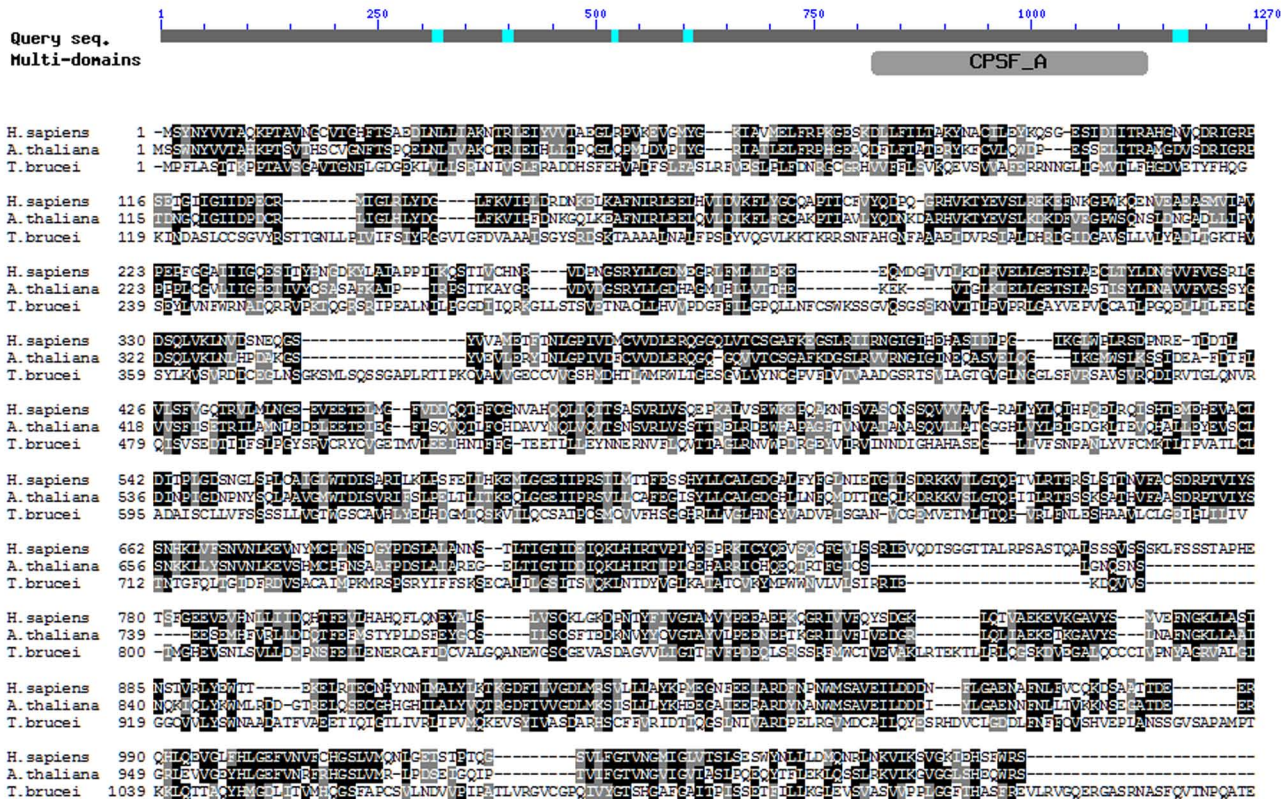
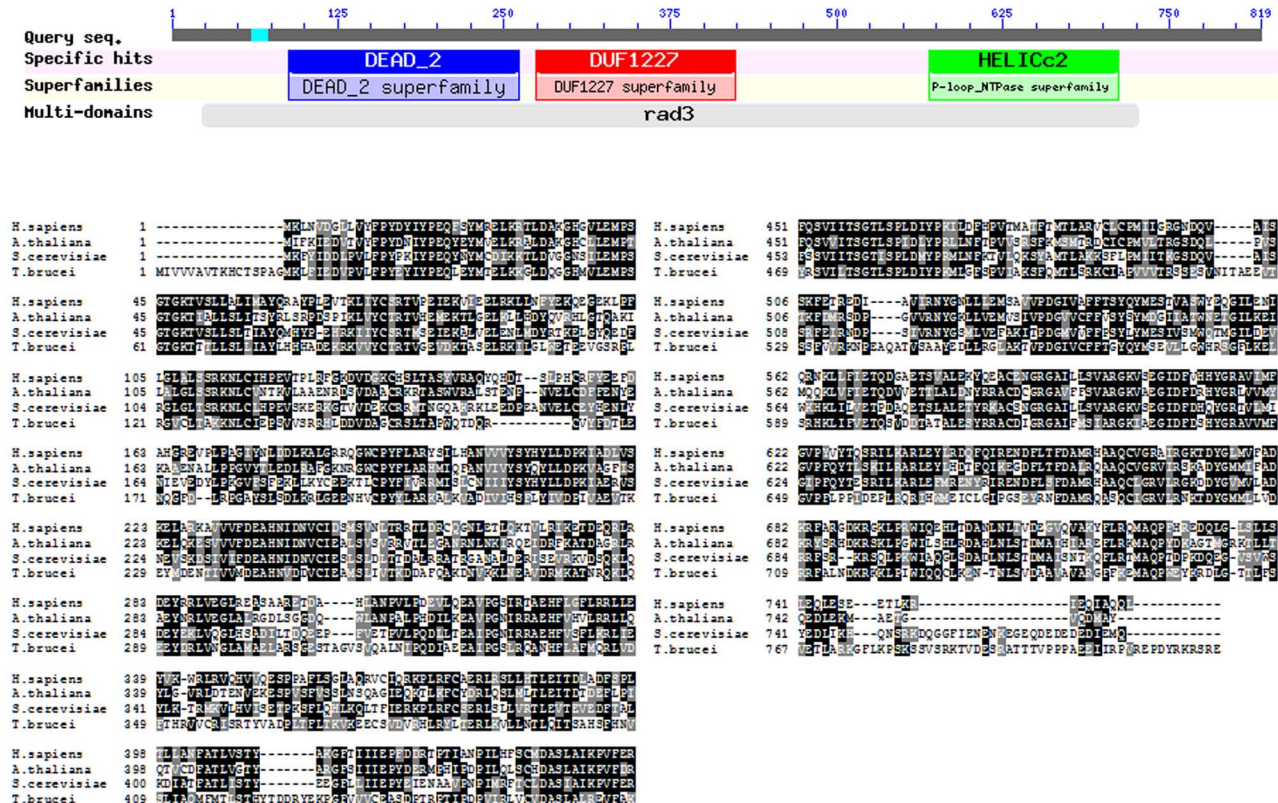


Fig.S2

T. brucei DDB





T. brucei XPB



T. brucei XPBz

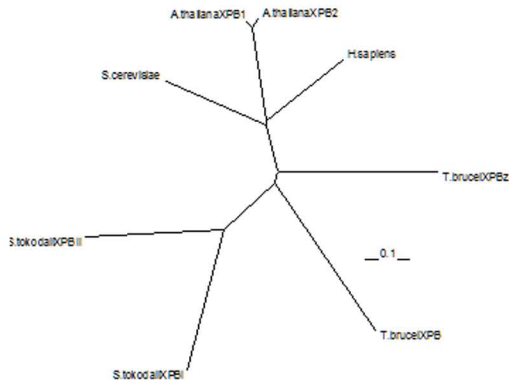


Fig.S6

T. brucei XPG

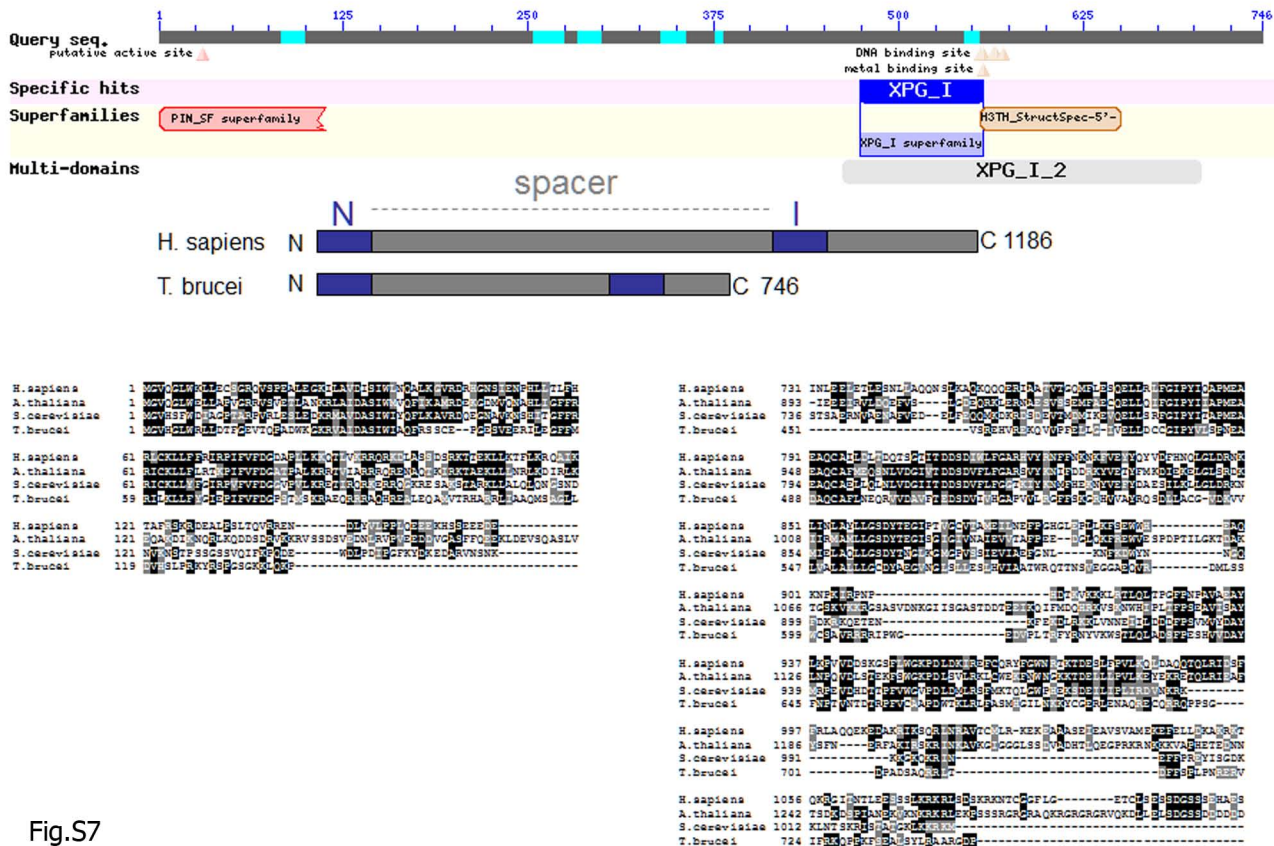


Fig.S7

T. brucei XPF

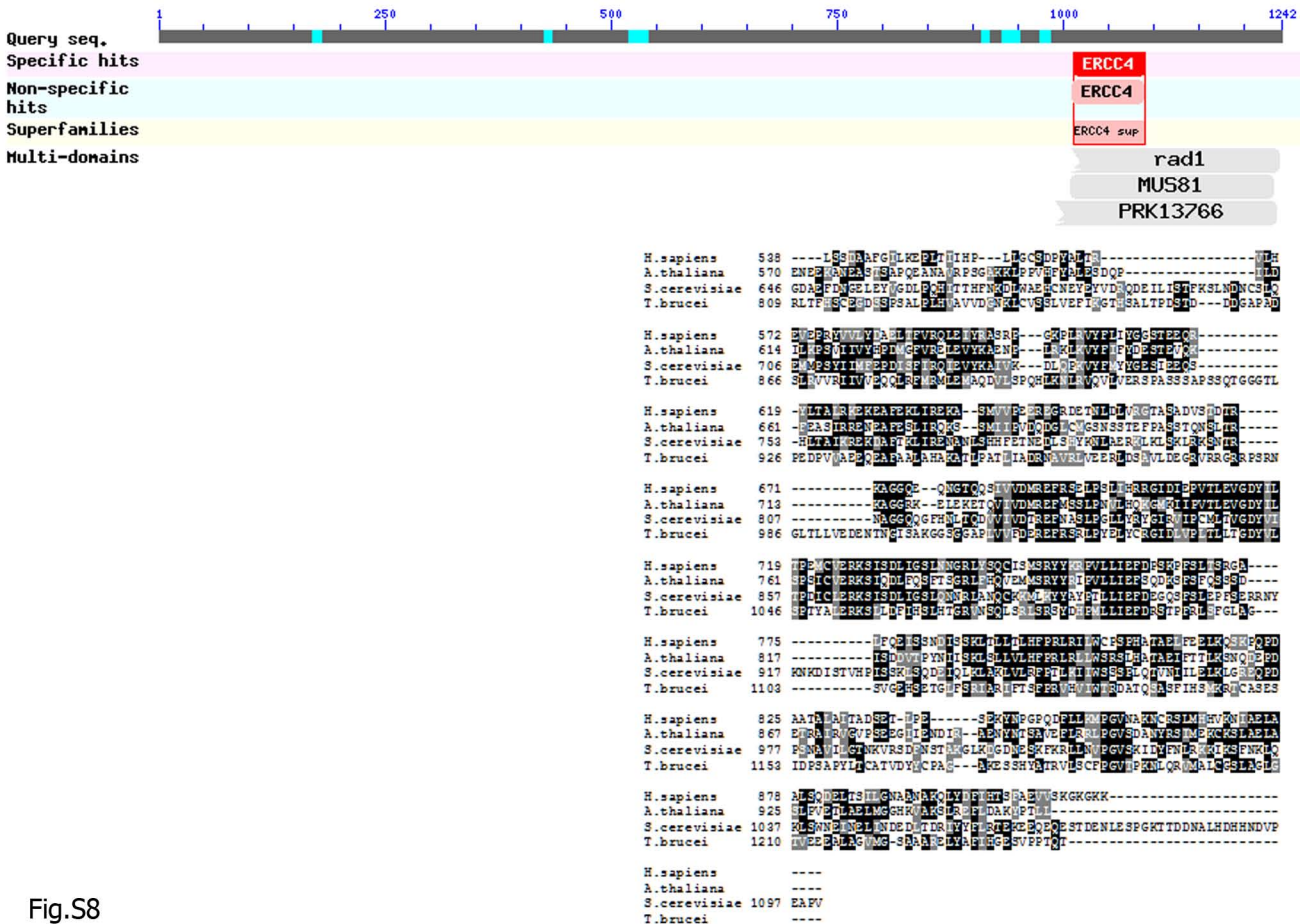


Fig.S8

T. brucei ERCC1

H.sapiens	1	----MDPEKDKESVYEQPSGEPARKKFFVIFID-----EDEVPPGVAKPLFRSTQSL
A.thaliana	1	MANEDDDGEGKSRSLHQIARKPKTQIVIGVPSYQEVLESSQTKSTPPSLEKPSQSFSQAF
S.cerevisiae	1	-----MNNIDETSEESILAGVA-----KLRKEKSGADTTGSQSL
T.brucei	1	-----MPPLPRGVVVKVVGALRGDNIIRIMQRHR-----YVIEEADLPIYDFLCGGTCV
H.sapiens	47	PTVDTIS-AGAAPQTYAEYALISQPIEGAGATCPTGSEPLAGETPNQALKPGAK-----
A.thaliana	61	A FV KSSD VY SP PP SSAAASSSQPSGASQVPHSSSQTHQTDGASSSSTPVATG SV PSNTT
S.cerevisiae	35	EIDASKLQ Q Q E PQTSRRINSNQVINA F NQ Q KPEEWTD S KATDDY N RKR P FRS-----T
T.brucei	49	V F VD D A D DL S DAARRTKV S QQLS V LKTHTGASWRCV V LL R V V SEEV R PDIL-----
H.sapiens	98	--S N SI I V S FR O GNP V LK F V R N V PE E FG-----D V I P DY V L G Q S T C A L F L S L R Y H N L H P
A.thaliana	121	Q N R N A I L V S H R O K G N P IL K H I R N V K W V FS-----D I I P DY V L G Q N S C A L Y L S L R Y H L L H P
S.cerevisiae	88	R P E K T V L V N T Q K EN P IL N H L K S T N R V VS S TGIN M I Y D Y L V R G R S V L F L T L T Y H K L Y V
T.brucei	101	--E W L N L H C S V E Q G C G V L F W T D E E C A A Y -----L E G L S D S N V A T A D Y C V G V R D S T P M
H.sapiens	151	D Y L H G R L O S L G K N E A L R V L L V Q V D V K D P Q A L K E L A K M C T L A D C T L I L A W S P E E A G R Y L E
A.thaliana	176	D Y L Y F R I R E L Q K N E K L S V V L C H V D V E D I V K P L L E V T K T A L L H D C T L L C A W S M T E C A R Y L E
S.cerevisiae	148	D Y L S R R M O P L S R N E N N -I L L F I V D D N N S E D T L N D I T K L C M F N G F T L L L A F N F E Q A A K Y T E
T.brucei	153	Q L I D A L T Q T P Q L M T R N D V V R A V N S F G S V A G L L I T A T ----AE Q L T E L P G F A Q K K A G R L H A
H.sapiens	211	T Y K A Y E Q K P A D L L M E K L E Q D F V S R V T E C L T T V K S V N K T D S Q T I L L T T F G S L E Q L I A A S R E D
A.thaliana	236	T L K V Y E N K P A D L L Q Q M D T D Y L S R L N H S L T S I R H V N K S D V V T L G S T F G S L A H I I D A S M E D
S.cerevisiae	207	Y L N L -----
T.brucei	209	V L N A P F N T S R C L V A V L Q R D-Q D E S N D E S S E R R P A C E T M K Q A L R C I Y D R E D E D V Q E G G Q -
H.sapiens	271	L A L C P G L G P Q K A R R L F D V L H E P L K V P -----
A.thaliana	296	L A R C P G L G E R K V R L Y D T F H E P K R A T S S Y P S V V E P P I P E A P V E K D V N S E E P V E E D E D F V
S.cerevisiae		-----
T.brucei		-----
H.sapiens		-----
A.thaliana	356	E D S R K R K K K E P E P E K T V K T A L S A V F A R Y S D R L S K K K E K Q K E K D T T T A S D A E T H Q N
S.cerevisiae		-----
T.brucei		-----

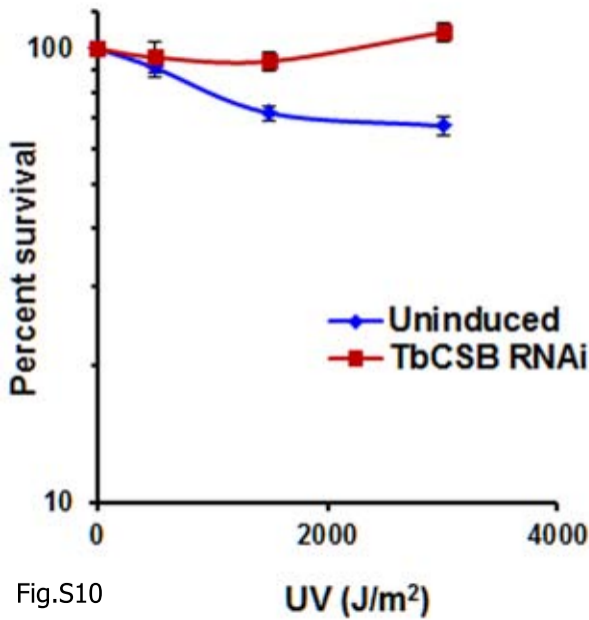


Fig.S10