

Supplementary Fig. S1 - Characterization of anti-FXN polyclonal antibodies and recombinant FXN proteins used in this study

Rec-FXN⁵⁶⁻²¹⁰ was used to produce two polyclonal antisera in rabbits (PAC 2517 and PAC 2518). The two antibodies were validated via western blotting analysis of total extracts from lymphoblastoid cells from two FRDA carriers and their affected children, and from normal controls. The indicated samples were analyzed by western blotting with PAC 2517 (approximately 35 µg of total IgG in 50 ml) as described in the legend of Fig. 1B. Afterwards, the membrane was stripped per manufacturer's protocol (acetonitrile for 10 min, followed by incubation for 30 min at 50 °C in 62.5 mM Tris-HCl buffer, pH 6.7, in the presence of 2% SDS and 100 mM β -mercaptoethanol) and cut in two halves (designated Blot 1 and Blot 2), which were probed with anti-FXN monoclonal antibody (MitoSciences, Inc.) and PAC 2518, respectively (approximately 140 µg of total IgG in 40 ml and 35 µg in 50 ml, respectively). All blots shown in this figure were developed by incubation with ECL Plus Western Blotting Detection Reagents (GE Healthcare) for 15 min. All three antibodies recognized FNX isoforms although with decreasing sensitivities (Monoclonal<PAC 2518<PAC 2517). Each antibody detected a different set of cross-reacting bands in the high-molecular-weight region of the blot (denoted by the bracket). Both PAC 2517 and PAC 2518 detected non-specific bands in the region of the membrane between FXN⁴²⁻²¹⁰ and FXN⁸¹⁻²¹⁰ (denoted by asterisks), which unlike FXN isoforms were present in equivalent levels in controls, carriers and patients.







Supplementary Fig. S2 - Human frataxin protein profiles in different control cells and tissues

(a) Total cell extracts were prepared from the indicated cell lines and tissues from non-FRDA controls as well as a yeast frataxin knock-out strain expressing the wild type human frataxin precursor ($vfh1\Delta[FRDA^{1-210}]$) as described in Experimental Procedures. Fifty µg total protein was loaded per lane (5 µg for yeast) and analyzed by western blotting with PAC 2517 anti-FXN antibody as described in Supplementary Fig. S1. The asterisk denotes an abundant protein present in heart, probably myoglobin, which co-migrates with FXN⁵⁶⁻²¹⁰. The double asterisk denotes a non-specific cross-reacting band as described in Supplementary Fig. S1. (b) The FXN protein profile was analyzed by western blotting with PAC 2517 in lymphoblastoid cell lines (Coriell Cell Repository) from 5 healthy individuals: GM07521 (19 years old female); GM14907 (28 years old male); GM14406 (41 years old female); GM05398 (44 years old male); GM03798 (10 years old male). The double asterisk denotes a non-specific cross-reacting band as described above. (c) Duplicates of the samples from Patient #1 analyzed in Fig. 2A were analyzed by western blotting with anti-FXN monoclonal antibody. Afterwards, this blot was stripped and reprobed with PAC 2518. The initial blot with PAC 2517 (the same blot shown in Fig. 2A) is shown for reference. The triple asterisk denotes a band in cerebellum and spinal cord, which does not co-migrate with any of the known FXN isoforms; this band was detected with both PAC 2517 and PAC 2518 as well as the monoclonal antibody and its significance is undetermined. The single asterisk denotes an abundant protein present in heart, probably myoglobin, which migrates close to FXN⁵⁶⁻²¹⁰. The white asterisk denotes a band that was detected with PAC 2517 but not with PAC 2518 or the monoclonal antibody. Immunodetection with monoclonal antibody required ~4 fold higher amounts of total purified IgG compared to polyclonal PAC

2517 or PAC 2518 (~140 μ g vs. ~35 μ g). (d) The FXN protein profile was analyzed once again by western blotting with PAC 2518 in a non-FRDA heart sample (the same sample analyzed in Supplementary Fig. S2a) side by side with heart samples from FRDA Patients #1 and #2. The single and triple asterisks are as described in (c) above. Native FXN42-210

1

42

MWTLGRRAVAGLLASPSPAQAQTLTRVPRPAELAPLCGRRGLR<u>TDIDATCTPR</u>R **81**

ASSNQR<mark>GLNQIWNVK</mark>K<mark>QSVYLMNLR</mark>K<mark>SGTLGHPGSLDETTYER/LAEETLDSLAE FFEDLADKPYTFEDYDVSFGSGVLTVK/LGGDLGTYVINKQTPNKQIWLSSPSSGP KRYDWTGKNWVYSHDGVSLHELLAAELTKALK<mark>TKLDLSSLAYSGKDA</mark></mark>

Rec-FXN42-210

42

AR<mark>TDIDATCTPR</mark>R

81

ASSNQR<mark>GLNQIWNVK</mark>K<mark>QSVYLMNLR</mark>K<mark>SGTLGHPGSLDETTYER/LAEETLDSLAE</mark> FFEDLADKPYTFEDYDVSFGSGVLTVK/LGGDLGTYVINK/QTPNKQIWLSSPSSG PKRYDWTGKNWVYSHDGVSLHELLAAELTKALKTKLDLSSLAYSGKDA

Native FXN81-210

1 42 MWTLGRRAVAGLLASPSPAQAQTLTRVPRPAELAPLCGRRGLRTDIDATCTPRR 81

ASSNQRGLNQIWNVKKQSVYLMNLRK<mark>SGTLGHPGSLDETTYER/LAEETLDSLAE</mark> FFEDLADKPYTFEDYDVSFGSGVLTVK/LGGDLGTYVINKQTPNKQIWLSSPSSGP KRYDWTGKNWVYSHDGVSLHELLAAELTKALKTKLDLSSLAYSGKDA

Rec-FXN81-210

81

SGTLGHPGSLDETTYER/LAEETLDSLAE FFEDLADKPYTFEDYDVSFGSGVLTVK/LGGDLGTYVINK/QTPNKQIWLSSPSSG PKRYDWTGKNWVYSHDGVSLHELLAAELTKALKTKLDLSSLAYSGKDA

Supplementary Fig. S3 - Peptide mass finger-printing analysis of native FXN⁴²⁻²¹⁰ and FXN⁸¹⁻²¹⁰

Native FXN proteins were isolated by immunoprecipitation and analyzed by in-gel trypsin digestion followed by nanoLC-MS/MS with hybrid orbitrap/linear ion trap mass spectrometry and peptide mass finger-printing, as described in detail in Supplementary Methods. The corresponding recombinant proteins were analyzed in parallel. The tryptic peptides identified for each protein are highlighted in yellow; contiguous peptides are separated by slashes. Note how the corresponding native (unknown N-terminus) and recombinant (known N-terminus) proteins yielded overlapping peptide profiles except for greater coverage in the C-terminal regions of the recombinant proteins. Three N-terminal peptides that conclusively distinguish FXN⁴²⁻²¹⁰ from FXN⁸¹⁻²¹⁰ are underlined. In the case of native FXN⁴²⁻²¹⁰, the close proximity of arginine 40 to arginine 43 does not allow to exclude that an additional very short peptide (GLR) was present upstream of peptide TDIDATCTPR. Together, these data allow mapping the N-terminus of native FXN⁴²⁻²¹⁰ between residues 41-43.



Supplementary Fig. S4 - Recombinant proteins used in this study

Proteins were expressed and purified as described in Experimental Procedures and Supplementary Methods. (a) One µg each of the indicated proteins was analyzed by 15% SDS-PAGE (Tris-HCl, Criterion precast gel, Bio-Rad) and protein staining with SYPRO Orange. (b) Five µg of purified oligomeric rec-FXN⁴²⁻²¹⁰ were analyzed by 15% SDS-PAGE (Tris-HCl, Criterion precast gel, Bio-Rad) and protein staining with Coomassie blue. Several weak bands were present in addition to the predominant FXN⁴²⁻²¹⁰ band. Indicated bands were excised from the gel, mixed together, digested with trypsin, and analyzed by nano-liquid chromatography/tandem mass spectrometry at the Mayo Proteomics Research Center. Peptide mass fingerprinting was used to match MS/MS spectra against the proteins in the Swiss-Prot database. This analysis revealed that the oligomeric rec-FXN⁴²⁻²¹⁰ preparation contained *E. coli* proteins not implicated in iron metabolism or iron-sulfur cluster synthesis. (c) Ten µg of purified human HisNFS1-ISD11 complex were analyzed by 18% SDS-PAGE (Tris-HCl, Criterion precast gel, Bio-Rad) and protein staining with Coomassie blue. Several weak bands were present in addition to the predominant NFS1 and ISD11 bands. Indicated bands were excised from the gel, and subjected to MS/MS analysis as described above. This analysis revealed the presence of the E. coli NifU-like protein (POACD6; 19 peptides) and other E. coli proteins not implicated in iron metabolism or iron-sulfur cluster synthesis.





Supplementary Fig. S5 - Interactions of native FXN isoforms in yeast mitochondrial lysate Yeast cells lacking endogenous Yfh1 and expressing FXN^{1-210} (yfh1 Δ [FRDA¹⁻²¹⁰]) were grown for ~20 h at 30 °C in rich medium with galactose as the carbon source, after which mitochondria were isolated as described in Supplementary Methods. The soluble mitochondrial fraction (~5 mg total protein) was analyzed on a Superdex 75 column. Fractions comprising the entire molecular mass fractionation range of the column were analyzed by western blotting with (a) anti-Nfs1, (b) anti-FXN or (c) anti-Isu1 antibody [(1) and this study]. The high- and lowmolecular-weight fractions (HMW and LMW boxes) were pooled; one half of each sample was incubated overnight with anti-FXN antibodies, and the second half with anti-Nfs1 antibodies (previously immobilized on Protein A Magnetic beads from New England BioLabs by crosslinking with dimethyl pimelidate dihydrochloride) in 10 mM HEPES-KOH, pH 7.3, 100 mM NaCl, 0.01% Triton X-100, with protease inhibitors. The beads were washed 3 times with 1 ml of buffer, and bound proteins were eluted with 80 µl of 100 mM Glycine, pH 2.5. Co-eluted proteins were analyzed by western blotting with (d) anti-Nfs1, (e) anti-FXN or (f) anti-Isu1 antibody. The band denoted by an asterisk is the heavy IgG chain. No cross reactivity was detected between the anti-Nfs1 antibody and FXN isoforms in the low molecular weight fractions (LMW). Higher levels of FXN are detected in individual HMW fractions (panel b) relative to the HMW fraction pool (panel e; *Input* and *Not bound*) because ~4 times less total protein was loaded in the latter analysis.

		Leu42 Ser56	
		↓ 64 75	
Human	MUTI GRRAVAGI ASP. SP ADADTI TRVPRP	- RGL RT DI DATC TPRRASSNORGI - NDI WNVKKOSVYL - MN R. K	
Chimp		PSL PTG I DATC	
Macaque	WITEGRAVAGUASP.SP		
Bovine		POLPLOTAKAP APSOSSI SI POL NO TI DVKKOSVOW, IN I P T	
Mouse		RGI HV TVNAGATRHAHI NI HYIO LI NIKKOSVCV. VHI.RN	
Rat		PRE HV TANADA	
Opossum			
Dufforfich		TRV. DADVW. BRIMHI SI H PR. R	
Zobrafish		A H H T H H H H H H H H H H H H H H H H	
Zebialish	51500K5		
Arabidanaia		. NAVRVSSY	
Arabidopsis			
Mosquito (Aedes)			
Mosquito (Anopheles)	IMNIV	D T ADEVE NON VORUS OF A	
Fly (D. persimilis)	MFTRRVVA	-PIRRL	
Fly (D. melanogaster)	······································	IGRWSKPOAHA.SOVILPSIPATA.AVAIQCEEFTANR.RLFS	
Irematode (S. mansoni)		·····MRYSTI-FQHLRRPRTCVHSFFLKSL-DL······VYI-TH····VAKIR	
Nematode (C. elegans)			
Nematode (C. briggsae)			
Nematode (C. remanei)		FS	
Yeast (S. cerevisiae)	······································	RYMIAAAGGE-RARFCPAVTNKK-NHTVNTFQKRFVE-SSTD	
Yeast (S. pombe)		······································	
0			
S			
S	er81		
Se	er81 87 <mark>50</mark> <mark>TLGHP</mark> GSLDETT <mark>YERLAEETLD SLAEFFEDL AD</mark> KPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KOIWL-SSPS	······ <mark>sgpkrydwtg·······knwvyshdgvslhellaaeltkalkt·kldlsslaysgkda······</mark>	
Su Human Chimp	er81 87 50TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSF6S6VLTVKLGGDLGTYVINKQTPN.KQIWL.SSPS 30TLGHPDETYERLAEETLDFLAEFFEDLADKPYTFEDYDVSF6S6VLTVKLGGDLGTYVINKQTPN.KQIWL.SSPS	······SGPKRYDWTG·······KNWVYSHDGVSLHELLAAELTKALKT-KLDLSSLAYSGKDA······ sopkrydwto·······knwvyshdovslhellaaeltkalkt-kldlsslaysokda······	
Su Human Chimp Macaque	er81 SGTLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN.KAIWL.SSPS SGTLGHPOSLDETTYERLAEETLDFLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN.KAIWL.SSPS SGTLGHPGSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN.KAIWL.SSPS	SGPKRYDWTG······· KNWVYSHDGVSLHELLAAELTKALKT·KLDLSSLAYSGKDA······ SGPKRYDWTG······· KNWVYSHDGVSLHELLAAELTKALKT·KLDLSSLAYSGKDA······ SGPKRYDWTG······· KNWVYSHDGVSLHELLGAELTKALKT·KLDLSSLAYSGKDA······	
Su Human Chimp Macaque Bovine	er81 SGTLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN-KAIWL-SSPS SGTLGHPOSLDDTTYERLAEETLDFLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN-KAIWL-SSPS SGTLGHPGSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN-KAIWL-SSPS AGTLGDAGTLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKATPN-KAIWL-SSPS	SGPKRYDWTG KNWVYSHDG VSLHELLAAEL TKALKT - KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDG VSLHELLAAEL TKALKT - KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDG VSLHELLG AEL TKALKT - KLDLSSLAYSGKDA SGPKRYDWTG RNWVYSHDG VSLHELLATEL TGALKT - KLDLSALAYSGKDTCCPA	
Su Human Chimp Macaque Bovine Mouse	er81 S6TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWL.SSPS S0TLGHPGSLDETTYERLAEETLDFLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWL.SSPS S6TLGHPGSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWL.SSPS A6TLGHPSSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWL.SSPS LGTLDNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWL.SSPS	SGPKRYDWTG · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · · SOPKRYDWTG · · · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · · SOPKRYDWTG · · · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · · SOPKRYDWTG · · · · · RNWVYSHDGVSLHELLATELTGALKT · KLDLSALAYSGKDT · · CCPA	4Q C
Su Human Chimp Macaque Bovine Mouse Rat	er81 56TLGHPGSLDETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 50TLGHPGSLDETTYERLAEETLD FLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGHPGSLDDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 66TLGDAGTLDDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 16TLGDASLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGDASLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGDASLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGNPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSPS 56	SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT.KLDLSSLAYSGKDA SOPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT.KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT.KLDLSSLAYSGKDA SGPKRYDWTG RNWVYSHDGVSLHELLATELTGALKT.KLDLSALAYSGKDTCCPA SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT.KLDLSSLAYSGKGT KGPKRYDWTG KNWVYSHDGVSLHELLARELTEALNT.KLDLSSLAYSGKGT	4Q C
Si Human Chimp Macaque Bovine Mouse Rat Opossum	er81 87 56TLGHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDLGTYVINKOTPN.KQIWL.SSPS 50TLGHPGSL DETTYERLAEETLD FLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDLGTYVINKOTPN.KQIWL.SSPS 64TLGHPGSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDLGTYVINKOTPN.KQIWL.SSPS 65TLGHPSSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFG SGVL TVKLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG SGVL TI KLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG SGVL TI KLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG SGVL TI KLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG SGVL TI KLGGDLGTYVINKOTPN.KQIWL.SSPS 56TLGNFNSL DETTYEKLAEETLD SLAFFEDLGD KPFTS KDYDVSLGSGVLTI KLGGDLGTYVINKOTPN.KQIWL.SSPT	SGPKRYDWTG · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · SOPKRYDWTG · · · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · SOPKRYDWTG · · · · · · KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA · · · · · SOPKRYDWTG · · · · · PNWVYSHDGVSLHELLATELTGALKT · KLDLSSLAYSGKOT · · · · CCPA · · SOPKRYDWTG · · · · · KNWVYSHDGVSLHELLARELTKALNT · KLDLSSLAYSGKOT · · · · · · · · · · · · KNWVYSHDGVSLHELLARELTKALNT · KLDLSSLAYSGKOT · · · · · · · · · · · · · · · · · · ·	4Q C
Si Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish	er81 87 56 ···· TLGHP ··GSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 56 ···· TLGHP ··GSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 66 ···· TLGHP ··GSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 66 ···· TLGHP ··SSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 66 ···· TLDNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 56 ···· TLGNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFGSGVLTVKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 56 ···· TLGNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFGSGVLTIKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 56 ···· TLGNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFGJGVLTIKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 56 ···· TLGNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFGJGVLTIKLGGDLGTYVINKOTPN·KQIWL·SSPS ·· 57 ···· TLGNP ··SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFGJGVLTIKLGGDLGTYVINKOTPN·KQIWL·SSPT ·· 58 ···· TLGDK ·· NSL DETTYEKLAEETLD SLADFFEDLOD KPFTS KDYDVSLGSGVLTIKLGGDLGTYVINKATPN·KQIWL·SSPT ··	SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA SOPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT · KLDLSSLAYSGKDA SGPKRYDWTG RNWVYSHDGVSLHELLATELTQALKT · KLDLSALAYSGKDTCPA SGPKRYDWTG KNWVYSHDGVSLHELLATELTKALNT · KLDLSSLAYSGKGT XGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT · KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT · KLDLSSLAYSGKOT SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT · KLDLSSLAYSGKD SGPKRYDWTG KNWVYNDGVSLHELLARELTFALNT · KLDLSSLAYSGKD SGPKRYDWTG KNWVYNDGVSLHELLARELTFALNT · KLDLSSLAYSGKD	AQ C
Si Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish	er81 SG TLGHP - GSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGHP - GSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGHP - GSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - LG TLGHP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGSGVLTVKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGJGVLTIKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGJGVLTIKLGGDLGTYVINKOTPN-KQIWL-SSPS - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGJGVLTIKLGGDLGTYVINKOTPN-KQIWL-SSP5 - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGJGVLTIKLGGDLGTYVINKOTPN-KQIWL-SSP5 - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKOTPN-KQIWL-SSP5 - SC KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGADYDVVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL-SSP5 - - E KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL-SSP5 - - E KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL-SSP5 - - E KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL SSP5 - - E KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL SSP5 - - E KAHHL- REISEAAYEKLVDETLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL SSP5 - - E KAHHL- REISEAAYEKLVGTLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL SSP5 - - E KAHHL- REISEAAYEKLVGTLDALAEYFEDLTDAAFTGLDYDVFSNGVLTVKVGSDHGTYVINKOTPN-RQIWL SSP5 - - E KAHHL- REISEAAF	SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT - KLDLSSLAYSGKDA SOPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT - KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLGAELTKALKT - KLDLSSLAYSGKDA SGPKRYDWTG RNWVYSHDGVSLHELLATELTGALKT - KLDLSALAYSGKDTCPA SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT - KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT - KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELT EALNT - KLDLSSLAYSGKDT SGPKRYDWTG KNWVYSHDGVSLHELLARELT EALNT - KLDLSSLAYSGKDT SGPKRYDWTG KNWVYSHDGVSLHELLARELT EALNT - KLDLSSLAYSGKDT SGPKRYDWTG KNWVYSHDGVSLHELLEMEF SQTLKT - QLDLSSLVYSGKDT SGPKRYDWTG ERWVYTHDGVSLHQLLSQEF SA IF SR - DIDLLDLPCS	AQ C
Si Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice	er81 S6 TLGRP - GSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - S6 TLOH P DETYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - S6 TLGDH GSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - AC TLGDA GTLD DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - C6 TLGDN SSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - S6 TLGDN SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN - KQIWL - SSP S - S6 TLGDN SSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG SGVL TI KLGGDL GTYVINKOTPN - KQIWL - SSP T - S6 TLGDN SSL DETAYERLAEETLD SLAFFEDL COKPFTS KOYDVSLG SGVL TI KLGGDL GTYVINKOTPN - KQIWL - SSP T - . S APVQI SSL DETAYERLAEETLD SLAFFEDL TDAAFTG ADYDVVF SV SSL PAXXXXXXXXXXPRLSPHARAL L SLPP C - E KAHHL RE I SEAEYERLAEETLD ALADYFEDL TDENFTG LDYDVVF SN GVL TVKNG SDH GTYVINKOTPN - RQIWL - SSP T - ARRLR SIL P EDEYHKLADET IHDLLEKLEEYGD SL - QMDGFD I DYGN QVL TLRLG - EL GTYVNKQAPN - RQIWL - SSP T - ARRLR SIL P EDEYHKLADET IHDLEKLEEYGD SL - QMDGFD I DYGN QVL TLRLG - EL GTYVNKQAPN - RQIWL - SSP T - 	SGPKRYDWTG	AQ C
Si Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis	er81 SGTLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGHPGSLDETTYERLAEETLDFLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWLSSP5 AGTLGHPGSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWLSSP5 AGTLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWLSSP5 SGTLGNPSSLDETAYERLAEETLDSLAFFEDLGAKPFTSKDYDVSFSSVSLPXXXXXXXXXLPRLSPHARALLSEPC SAPVQI.SSLDETAYERLAEETLDALAEFFEDLGDKPFTSLDYVFSNGVLTVKGSDHGTYVINKOTPN.KQIWLSSP5 CKAHHLREISEAEYERLAEETLDALAEFFEDLGDKPFTGLDYDVFSSSVLPXXXXXXXXXXXXXXXXXXXXXXXXXXXX ARRLR.SILPEDEYHKLAEETLDALAEFFEDLGDAFFEGLSONVFSNGVLTKKGSDHGTYVINKOTPN.KQIWLSSP5 ARRLR.SILPEDEFHKLAAETIDALEKEEYEGDSL.QMDFFDIDYGNQVLTKKGSDHGTYVINKOTPN.RQIWLSSP5 ARRLR.SILPEDEFHKLAAETINHLEKEEYEGDSL.QMDFFDIDYGNQVLTKKGSLGTYVINKOTPN.RQIWW.SSPV	SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLGAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLATELTGALKT KLDLSALAYSGKDT COPA SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELTEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYVHDGVSLHELLARELTEALNT KLDLSSLAYSGK SGPKRYDWTG KNWVYVHDGVSLHELLARELTEALNT KLDLSSLAYSGKDT SGPKRYDWTG KNWVYVHDGVSLHELLARELTEALNT KLDLSSLAYSGKDT SGPKRYDWTG KNWVYVHDGVSLHELLARELTEALNT KLDLSSLAYSGKDT SGPKRYDWTG FRWVYTHDGVSLHALLSGEFSAIFSR DIDLLDLPCS SGPKRYDWTG FRWVYTHDGVSLHALLSGEFSAIFSR DIDLSHLIHS SGPKRYDWTG ANLYRTEAKLHKLLEEELENLCGE PIGLS	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes)	er81 56TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWU-SSPS 50TLGHPGSLDETTYERLAEETLDFLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGHPGSLDDTTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTEDYDVSFGSGVLTVKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGDGVLTIKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGNPSSLDETAYERLAEETLDSLAEFFEDLGNKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKOTPN.KQIWU-SSPS 56TLGOKNSLDETTYEKLAEETLDSLAFFEDLGNKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKOTPN.KQIWU-SSPT 5APVQI.SELSEAAYEKLVDETLDALAEYFEDLTDAAFTGADYDVFSVSLPAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SOPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLATEL TGALKT KLDLSSLAYSGKDT CCPA SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKDT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKDT SGPKRYDWTG KNWVYSHDGVSLHELLEMEF SGTLKT GLDLSSLVYSGKDT SGPKRYDWTG FRWVTHDGVSLHALLSGEF SGTF NIDLSSLVYSGKDT SGPKRYDWTG FRWVTHDGVSLHALLSGEF SGTF NIDLSHL HS SGPSRFDWDAP TNCW IYRTG ANLVELLEKEI GELCGT FVELS SGPSRFDWDRD ANAWIYRFTEAKLHKLLE EEL EN LCGG FIGLS.	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Anopheles)	er81 56TLGHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 66TLGHPGSL DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 66TLGDAGTL DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 66TLGDAGTL DTTYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG SGVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGDASSL DETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG SGVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGDKNSL DETTYEKLAEETLD SLAEFFEDL GD KPFTS KOYDVSLG SGVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGDKNSL DETTYEKLAEETLD SLAEFFEDL GD KPFTS KOYDVSLG SGVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGDKNSL DETTYEKLAEETLD SLAPFFEDL GD KPFTS KOYDVSLG SGVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGDKNSL DETTYEKLAEETLD ALAEFFEDL TDAAFTG ADYDVVF SV SSL PAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	SGPKRYDWTG KNWVYSHDG VSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDG VSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDG VSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDG VSLHELLAREL TKALKT KLDLSALAYSGKDT CCPA SGPKRYDWTG KNWVYSHDG VSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDG VSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDG VSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDG VSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDG VSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDG VSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWYTHDG VSLHELLSKELS SGT SA I SR DI DLLDLPCS SGPKRYDWTG KNYTHD AVPLHSLLSKELS I I SR DI DLLDLPCS SGPSRFDWDAP T NCW I YRRTG ANLVELLE KEI GEL CGT F VELS SGPSRFDWD RD ANAWI YRRTE AKLHKLLE EEL EN LGGE F VIJSLPHSKRLD SGPKRYD F I PDKSKVNE GF WLYRHOG VSLHELLARE I SA I VGR ELEFF ALPHSGRPQEPASDGR	AQC
Su Human Chimp Macaque Bovine Bovine Mouse Rat Opossum Pufferfish Zebrafish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis)	er81 56TLOHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSF 6 S 6VL TVKL GGDL 6TYVINKOTPN.KQI WL.SSP 5 50TLOHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSF 6 S 6VL TVKL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLOHPSSL DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSF 6 S 6VL TVKL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSF 6 S 6VL TVKL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSF 6 S 6VL TVKL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSF 6 S 6VL TI KL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSF 6D 6VL TI KL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD ALAEFFEDL AD KPYTL KDYDVSF 6D 6VL TI KL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONPSSL DETAYERLAEETLD ALAEFFEDL AD KPYTL KDYDVSF 6D 6VL TI KL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONFSSL DETAYERLAEETLD ALAEFFEDL AD KPYTL KDYDVSF 6D 6VL TI KL GGDL 6TYVINKOTPN.KQI WL.SSP 5 56TLONKNSL DETTYEKLAEETLD ALAEFFEDL TDAAFT6 ADYDVVF SV SSL PAXXXXXX XXXX FRLSPPHARAL L SLPP C 57APVQ 1SEL SEAAYEKL VDETLD ALAEFFEDL TDAAFT6 ADYDVVF SV SSL PAXXXXXX XXXX FRLSPPHARAL L SLPP C 58APVQ 1SEL SEAAYEKL VDETLD ALAEFFEDL TDAAFT6 LDYDVVF SV SSL PAXXXXXX XXXX FRLSPPHARAL L SLPP C 59ARRLR.SILP EDYHKLABET INDLLEKLEEYOD SL. QMDGFD I DYGN QVL TVKVG SDH GTYVINKQAPN.RQI WL.SSP V 70ARRLR.SILP EDYHKLABET INDLLEKLEYOD SL. QMDGFD I DYGN QVL TVKVG SDH GTYVINKQAPN.RQI WL.SSP V 71DF IAVSL ID SITF FAVCSDTLE SLODYFEQ I VE FAAFL KSADI TYGD GVL TVNF GEPY GTYVINKQAPN.RQI WL.SSP V 72DF IAVSL ID SITF FAVCSDTLE SLODYFEQ I VE FAAFL KSADI TYGD GVL TVNF GEPY GTYVINKQAPN.RQI SSP V 75QIETESSLD TATYERVOSETLDGLCDYFEEL TE NATDL IGTDVAYG GVL TVNLGKSH GTYVINROTPN.KQI V.SST	SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALKT KLDLSALAYSGKDT CCPA SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKDT SGPKRYDWTG SCH KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SCH KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SCH KNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SCH KNWVYNHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SCH KNWVYNHDGVSLHELLSGEF SATF SR DTDLLDLPGS SGPKRYDWTG SCH KNWVYNHDGVSLHELLSKELST SGF SCH ST SGPKRYDWTG SCH SCH SCH ST SGPSRFDWDAP SCH	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Anopheles) Fly (D. persimillis) Fly (D. melanogaster)	er81 56TLGHPGSLDETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 50TLGHPGSLDETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 50TLGHPGSLDDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 64TLGHPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 55TLGHPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGNPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGNPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGNPSSLDETAYERLAEETLD SLAEFFEDL AD KPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSPS 56TLGNPSSLDETAYERLAEETLD ALAEFFEDL AD KPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 56TLGNPSSLDETAYERLAEETLD ALAEFFEDL AD KPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 56TLGNK.NSLDETTYEKLAEETLD ALAEFFEDL AD KPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 57APVQ1.SELSEAAYEKLVDETLD ALAEFFEDL TD AAFTG ADYDVVFSVSLGSGVLTIKLGGDLGTYVINKGTPN.RGIWL.SSP5 58APVQ1.SELSEAAYEKLVDETLD ALAEFFEDL TD AAFTG ADYDVVFSVSSLPAXXXXXXXXXXX FRLSPPHARALLSLPPC 59ARRLR.SILPEDYHKLADETIHDLLEKLEEYGDSL-GMDGFD IDYGNAVLTKVGSDHGTYVINKGTPN.RGIWL.SSP5 79ARRLR.SILPEDYHKLADETIHDLLEKLEEYGDSL-GMDGFD IDYGNAVLTRLG.ELGTYVINKGAPN.RGIWL.SSP5 75OFTAV.SLIDSITFEAVCSDTLESLCDYFEGL TEAAFLKKADD TYGD GVLTVNFGEPYGTYVINKGAPN.RGIWL.SSP5 75OFTAV.SLIDSITFEAVCSDTLESLCDYFEGL TE KANFDVINGD VING FFFFGL VVINKGAPN.RGIWL.SSP5 75QIETESSLDTATYERVCSSTLGGLCDYFEELTENATDL IGTDVAYGOVLTVNLGGNGTYVINKGTPN.KGIWL.SSP5 75QIETESSLDTATYERVCSSTLDALCDYFEELTENATDL IGTDVAYGOVLTVNLGGNGTYVINKGTPN.KGIWL.SSP5 75QIETESSLDTATYERVCSSTLDALCDYFEELTENATDL IGTDVAYSOVLTVNLGGNGTYVINKGTPN.KGIWL.SSP5 75QIETESSLDTATYERVCSSTLDALCDYFEELTENATDL IGTDVAYSOVLTVNLGGNGGVLTVNLGGNGTYVINKGTPN.KGIWL.SSP5 75QIETESSLDTATYERVCS	SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALKT KLDLSALAYSGKDT CCPA SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG SKNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SKNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SKNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG SKNWVYSHDGVSLHELLAREL TEALNT KLDLSSLAYSGK SGPKRYDWTG SKNWVYSHDGVSLHELLEMEF SQTLKT SKNWSGKDT SGPKRYDWTG SKNWYSHDGVSLHELLEMEF SQTLKT SKNWSGK SGPKRYDWTG SKNWSYNDGVSLHELLEKEL SFI SR SFI DLLDLPCS SGPKRYDWTG SKNWSYNDGVSLHELLEKEL SGEL SGT SKNWS SGPSFDWD AP ST NCW YRTGANLVELLEKE I GEL SGT SKNWS SGPKRYDF I PDKSKVNE GFWLYKHDGVSLHELLQKE I GVI VQK SVDFLSPHSKRLD SGPKRYDF I PDKSKVNE GFWLYKHDGVSLHELLQKE I GVI VQK SUDFLSPHSKRLD SGPKRYDF VGTP SKAGKW YRHSGQSLHELLQEI PN I VKSQTVDFMRLPHSS SOPKRYDF VGTV SAGKW YRHSGQSLHELLQEI PN I LKSQSVDFLRLPYCS S	AQ C
Su Human Chimp Macaque Bovine Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis) Fly (D. melanogaster) Trematode (S. mansoni)	er81 SG TLG HP GSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN . KGI WL . SSP S SG TLG HP GSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN . KGI WL . SSP S SG TLG HP GSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN . KGI WL . SSP S SG TLG DA STLD DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DA STLD DTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DA SSLD ETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG SGVL TVKLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DK NSLD ETTYEKLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG D GVL TI KLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DK NSLD ETTYEKLAEETLD ALAEFFEDL AD KPYTL KDYDVSFG D GVL TI KLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DK NSLD ETTYEKLAEETLD ALAEFFEDL AD KPYTL KDYDVSFG D GVL TI KLGGDL GTYVINKGTPN KGI WL SSP S SG TLG DK NSLD ETTYEKLAEETLD ALAEFFEDL TD AAFTG AD YDVVF SV SSLPAXXXXXX XXXXLFRLSPHARALL SLPP C E KAHHL REI SEAEYERLAEETLD ALAEFFEDL TD ENFTG LDYDVVF SN SVL TVKVG SD HGTYVINKGTPN KGI WL SSP T ARRLRR SILP EDEYHKLADETI HDLLEKLE EVGD SL GMDGFDI DYGN GVL TVKVG SD HGTYVINKGTPN RGI WL SSP VRL PA SVDYS SVLGEEFFKLANFT INHLLEKIE DYGD NV. GI DGFDI DYGN GVL TVK GSD HGTYVINKGTPN RGI WL SSP V PN DF I AV SL I D SITFEAVCSDTLE SLCDYFEGL VEEAAFL K.SADI TYGD GVL TVK FGEPY GTYVINKGTPN RGI WL SSP T SSMNPNDFI AT SLI D SITFEAVCSDTLE SLCDYFEGL VEEAAFL K SDI TYGD GVL TVN FGEPY GTYVINRGPN RGI WL SSP T SSMNPNDFI AT SLI D SITFEAVCSDTLE SLCDYFEGL VEEAAFL K.SADI TYGD GVL TVN FGEPY GTYVINRGPN RGI WL SSP T SSMNPNDFI AT SLI D SITFEAVCSDTLE SLCDYFEGL TENATEL I GTDVAYGD GVL TVN FGEPY GTYVINRGPN RGI WL SSP T SS QI E TE SSLD TATYERVCSDTLDALGDYFEL TENASEL GOTDVAYS O	SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLAAELTKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWVYSHDGVSLHELLATELTGALKT KLDLSALAYSGKDT CCPA SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT KNWVYSHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYSHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYVHDGVSLHELLARELTKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWVYVHDGVSLHELLEMEFSGTLKT QLDLSSLVYSGKDT SGPKRYDWTG KNWVYVHDGVSLHELLEMEFSGTLKT VLDLSSLAYSGK SGPKRYDWTG KNWVYVHDGVSLHELLEMEFSGTLKT NIDLSHLINS SGPKRYDWTG KNWVYHDGVSLHELLEKEIGELGGT PVELS SGPKRYDWTG KNWVYHDGVSLHELLEKEIGELGGT VGLS SGPKRYDWTG KNW SHTGANLVELLEKEIGELGGT PVELS SGPSRFDWD AP T NCW YRRTGANLVELLEKEIGEL NG GE PIGLS SGPKRYDFI PDKSKVNE GFWLYKHDGVSLHELLGKEIGVI VGK EVDFLSLPHSKRLD SGPKRYDFI PDKSKVNE GFWLYKHDGVSLHELLGREISAIVGR ELEFFALPHSGRPGEPASDGR SGPKRYDFI PDKSKVNE GFWLYRHSGGSLHELLGREI SAIVGG PVEFGT.	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis) Fly (D. melanogaster) Trematode (S. mansoni) Nematode (C. elegans)	er81 SG TLGHP - GSLDETTYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGHP - GSLDETTYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGHP - GSLDDTTYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - AG TLGHP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTF EDYDVSFGSGVLTVKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNP - SSLDETAYERLAEETLDSLAEFFEDLADKPYTLKDYDVSFGSGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNP - SSLDETAYERLAEETLDALAEFFEDLADKPYTLKDYDVSFGGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNK - NSLDETTYEKLAEETLDALAEFFEDLADKPYTLKDYDVSFGGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNK - NSLDETTYEKLAEETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNK - NSLDETTYEKLAEETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG TLGNK - NSLDETTYEKLAEETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG APVQI - SELSEAAYEKLVDETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN-KQIWL SSP5 - SG APVQI - SELSEAAYEKLVDETLDALAEFFEDLGDKPFTSKDYDVSLGSGVLTKKGSDHGTYVINKGTPN-KQIWL SSP5 - SG APVQI - SELSEAAYEKLVDETLDALAEFFEDLGDKFEQLTGAAFTGADVSVSSSVGTVKKGSDHGTYVINKGTPN-KQIWL SSP5 - SS APVQI - SELSEAAYEKLVDETLDALAEFFEDLGTGDVSQU TVSNGVTVKSSSVGTVVINKGAPN-PGIWL SSP5 - SS APVQI - SELSEAAYEKLVDETLDALAEFFEDLGUTFEGLVEGAAFTGADVSVSSST SS APVQI - SELSEAYEKLYDETLGLCDYFEGLVEEAAFTGADVSSSSVGTVVINKGSPN-KQIWL SSP5 - SS QIETE- SSLDTATYERVSSETLSLCDYFEGLVEEAAFTGADVSVSSSS SS QIETE- SSLDTATYERVSSETLGLCDYFEELTENATDLIGTDVAYSDGVLTVNFGGPNGTVVINKGPN-KQIWL SSP5 - SS QIETE- SSLDTATYERVSSETLGLCDYFEELTENATDLIGTDVAYSDGVLTVNSSSVGTVVINKGPN-KQIWL SSP5 - SS QIETE- SSLDTATYERVSSTTTELLASSEQFFEGLTENASELGTVVI	SGPKRYDWTG KNWYSHDG VSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWYSHDG VSLHELLAAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWYSHDG VSLHELLGAEL TKALKT KLDLSSLAYSGKDA SGPKRYDWTG KNWYSHDG VSLHELLATEL TGALKT KLDLSSLAYSGKDT CCPA SGPKRYDWTG KNWYSHDG VSLHELLAREL TKALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWYSHDG VSLHELLAREL TEALNT KLDLSSLAYSGKGT SGPKRYDWTG KNWYSHDG VSLHELLEMEF SGTLKT OLDLSSLAYSGKDT SGPKRYDWSG SERWYTHDG VSLHELLEMEF SGTLKT OLDLSSLAYSGKDT SGPKRYDWTG SERWYTHDG VSLHELLEMEF SGTLKT NIDLSHLINS SGPKRYDWTG SERWYTHDG VSLHELLEMEF SGTLKT SGTLKT SGTL SGPKRYDFG SERWYTHDG SLHELLGGE SGT SGTLST SGPKRYDFF SGTLYRHGGYTLHELGKEI GELCGT SELFF ALPHSGRPGEPASDGR SGPKRYDF I PDKRTINE GYMLYRHGGVTLHELLGRE I SATVGG SVDFLRHPSSS SGPKRYD Y GF SAGKWTYRHGGSLHELLGGE I GTLKSG SVDFLRHPSSS SGPKRYD Y FS MRLWTYRHGGSLHSLLGGE I SATVGG PVEFGT SGPKRYD I PS MRLWTYRHGGSLHSLLGE I SATVGG PVEFGT	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis) Fly (D. melanogaster) Trematode (S. mansoni) Nematode (C. briggsae)	Er81 SG TLGHP GSLDETTYERLAEETLD SLAEFFEDL ADKPYTF EDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGHP GSLDETTYERLAEETLD SLAEFFEDL ADKPYTF EDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGHP GSLDETTYERLAEETLD SLAEFFEDL ADKPYTF EDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 AG TLGHP SSLDETAYERLAEETLD SLAEFFEDL ADKPYTF EDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGHP SSLDETAYERLAEETLD SLAEFFEDL ADKPYTF EDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD SLAEFFEDL ADKPYTLEDYDVSFGSGVLTVKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL ADKPYTL KDYDVSFGSGVLTIKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL ADKPYTL KDYDVSFGSGVLTIKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL ADKPYTL KDYDVSFGSGVLTIKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL ADKPYTL KDYDVSFSGVSTIKLGGDLGTVVINKOTPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL ADKPYTL KDYDVSFSGVLTIKLGGDLGTVVINKATPN.KQIWL.SSP5 SG TLGNP SSLDETAYERLAEETLD ALAEFFEDL GDKPFTS KDYDVSFSVSLPAXXXXXXXXLPRLSPHARALLSEPPC S APVQ1 SSLDETYEKLAEETLD ALAEFFEDL TO AAFTGA ADYDVFSVS SNGVLTVKVGSDHGTVVINKATPN.KQIWL.SSP5 SG TLGNP SILP EDEYHKLAEETLD ALAEFFEDL TO AAFTGA ADYDVFSNGVLTVKVGSDHGTVVINKATPN.KQIWL.SSP5 SG APVQ1 SELSEAAFEKLDETLD ALAEFFEDL TO AAFTGA ADYDVFSNGVLTVKGSDHGTVVINKATPN.KQIWL.SSP5 SG APVQ1 SELSEAAFEKLDETLD ALAEFFEDL TO AAFTGA ADYTVFSNGVLTVKGSDHGTVVINKATPN.KQIWL.SSP5 SS APVQ1 SELSEAFFEDL ALAEFFEDL TO AAFTGA ADYTVFSNGVLTVKGSDHGTVVINKATPN.KQIWL.SSP5 SS APVQ1 SELSEAFFEDL ALAEFFEDL TO AAFTGA ADYTON ADYTYSNG SOVLTVKGSDHGTVVINKATPN.KQIWL.SSP5 SS QIETEE. SILD SATFYEQUE ADYFOUTINELEKEEYGO SL. QMDFDUN ADYTON ADYTNNKATPN.KQIWL.SSP5 SS QIETEE. SSLD	SGPKRYDWTG- KNWVYSHDGVSLHELLAAELTKALKT-KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLAAELTKALKT-KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLAAELTKALKT-KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLATELTGALKT-KLDLSSLAYSGKDT- CCPA SGPKRYDWTG- KNWVYSHDGVSLHELLARELTKALNT-KLDLSSLAYSGKGT- SGPKRYDWTG- KNWVYSHDGVSLHELLARELTKALNT-KLDLSSLAYSGKGT- SGPKRYDWTG- KNWVYSHDGVSLHELLARELTFALNT-KLDLSSLAYSGK0T- SGPKRYDWTG- KNWVYSHDGVSLHELLARELTKALNT-KLDLSSLAYSGK- SGPKRYDWTG- KNWVYSHDGVSLHELLARELTEALNT-KLDLSSLAYSGK- SGPKRYDWTG- KNWVYSHDGVSLHELLARELTEALNT-KLDLSSLAYSGK- SGPKRYDWTG- FRWVYTHDGVSLHELLEMEFSATFSR-DIDLLDLPCS- SGPKRYDWTG- FRWVYTHDGVSLHELLEMEFSATFSR-DIDLLDLPCS- SGPSRDWDAP- TNCWTYRRTGANLVELLEKETGELCGT-PVELS- SGPSRDWDAP- ANAWTYRRTEAKLHKLLEEELENLCGE-PTGLS- SGPKRYDFTPDKSKVNEGFWLYKHDGVSLHELLGKETGVTVGK-EUFFSLPHSKRLD- SGPKRYDFTPDKSKVNEGFWLYKHDGVTLHELLGRETSATGR-ELEFFALPHSQRPQEPASDGR SGPKRYDFTPVGTP-KAGKWTYRHGGSLHELLGETPGTLKSGSVDFLRLPYCS- SOPKRYDTFVGTV-AAGRWTYNGGGSLHELLGETSATVGG-PVELS- SGPKRYDTFVGS- MRLWTNGGGSLHSLLSETSATVGG-PVERV- SGPKRYDTFSKRV SGPKRYDTFS- MRLWTYKHDGKSLHSLLSETSATVGG-PVERV	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis) Fly (D. persimillis) Fly (D. melanogaster) Trematode (S. mansoni) Nematode (C. elegans) Nematode (C. remanei)	er81 SG TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 AG TLGHPGSLDETTYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 AG TLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTFEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGHPSSLDETAYERLAEETLDSLAEFFEDLADKPYTEEDYDVSFGSGVLTVKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNPSSLDETAYERLAEETLDSLAEFFEDLADKPYTLEDYDVSFGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNPSSLDETAYERLAEETLDSLAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNPSSLDETAYERLAEETLDSLAEFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNKNSLDETTYEKLAEETLDSLAPFFEDLGDKPFTSKDYDVSLGSGVLTIKLGGDLGTYVINKGTPN.KGIWL.SSP5 SG TLGNKNSLDETTYEKLAEETLDSLAPFFEDLGDKPFTSKDYDVSSGVVTYKVGSDHGTVVINKGTPN.KGIWL.SSP5 SG APVGI.SELSEAAYEKLVDETLDALAEYFEDLGDKPFTGLDYVFSNSULTKKGGDLGTVVINKGTPN.RGIWL.SSP5 SG APVGI.SELSEAYEKLAEETLDALAEYFEDLTDAAFTGGDVVFSNSULTKKGSDHGTVVINKGTPN.RGIWL.SSP5 SVDYS.SVLGEEFHKLANFTIHHLEKLEEYGSSL.GMGFDIDYGNVUTUKGSDHGTVVINKGPN.RGIWL.SSP5 SVDYS.SVLGEEFHKLANFTIHHLEKLEDYGONV.GIDGFDIDYGNSVLTKNGSDHGTVVINKGPN.RGIWL.SSP5 SS GIETE.SLDTATYERVCSDTLESLCDYFEGLVEEAAFLKSADITYGDGVLTVNFGEPYGTVINKGPN.RGIWL.SSP5 SS GIETE.SLDTATYERVCSDTLESLCDYFEGLTENATDLIGTDVAYSDGVLTVNLGKSHGTVVINKGPN.KGIWL.SSP5 SS GIETE.SLDTATYERVSSTLELLADSFEQLPERFALNKEYDVENGVLTVNLGGHGTVVINKGPN.KGIWL.SSP5 SS GIETE.SLDAATYERVSSTLELLADSFEQLPERFALNKEYDVENGVLTVNLGGHGTVVINKGPN.KGIWL.SSP5 SS GIETE.SLDAATYERVSSTLELLADSFEGLPERFALNKEYDVENGVLTVNLGGHGTVVINKGSPN.KGIWL.SSP5 SS GIETE.SLDAATYERVSSTLELLADSFEGLPERFALNKEYDVENGVLTVNLGGVLTVNSKSVGTVVINKGSPN.KGIWL.SSP5.	SGPKRYDWTG- KNWVYSHDG VSLHELLAAEL TKALKT - KLDLSSLAYSG KDA- SOPKRYDWTG- KNWVYSHDG VSLHELLAAEL TKALKT - KLDLSSLAYSG KDA- SGPKRYDWTG- KNWVYSHDG VSLHELLAAEL TKALKT - KLDLSSLAYSG KDA- SGPKRYDWTG- KNWVYSHDG VSLHELLATEL TQALKT - KLDLSSLAYSG KDA- SGPKRYDWTG- KNWVYSHDG VSLHELLAREL TKALNT - KLDLSSLAYSG KDT - CCPA SGPKRYDWTG- KNWVYSHDG VSLHELLAREL TKALNT - KLDLSSLAYSG KDT - CCPA SGPKRYDWTG- KNWVYSHDG VSLHELLAREL TFALNT - KLDLSSLAYSG KDT - SGPKRYDWTG- KNWVYSHDG VSLHELLAREL TFALNT - KLDLSSLAYSG KDT - SGPKRYDWTG- FRWVYTHDG VSLHELLAREL TFALNT - KLDLSSLAYSG KDT - SGPKRYDWTG- FRWVYTHDG VSLHELLSG FSA IFSR - DIDLLDLPCS - SGPKRYDWTG- FRWVYTHDG VSLHALLSG FSA IFSR - DIDLLDLPCS - SGPSRFDWD AP - T NOW IYRRTG ANLVELLE KE IG EL GGT - PVELS - SGPSRFDWD AP - T NOW IYRRTG ANLVELLE KE IG EL GGT - PVELS - SGPKRYDF I PDKSKVNE GF WLYKHDG VSLHELLARE I SA I VGR - ELEFF ALPHSQRPQEPASDGR SGPKRYDF I PDKSKVNE GF WLYKHDG VSLHELLQ KE IG VI VQK - EVDFL SLPHSKRLD - SGPKRYDF I PDKSKVNE GF WLYKHDG VSLHELLQ KE IG VI VQK - EVFF ALPHSQRPQEPASDGR SGPKRYDF VGTV - KAGKWI YRHSG SLHQLLQ E I PN I VKSQ TVDFMRLPHCS - SGPKRYDF VGTV - KAGKWI YRHSG SLHQLLQ E I PN I VKSQ SVDFLRLPVCS - SGPKRYDF VGTV - AAGRWI YKHGG SLHELLQ ZE IS AI VGG - PVEFSRHV - SGPKRYD I PS - MRLWI YKHOG KSLHSLLS EE IS AI VGG - PVEFSRHV - SGPKRYD L AEE - O GWYYSHDG EKLD SLINREF RK I LAD DR I DFSRHV - SGPKRYD LAEE - O SWYSHDG EKLD ELLNREF RK I LAD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LAD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LAD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD R I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - O SWWYSHDG EKLD ELLNREF RK I LGD DR I DFSRHV - SGPKRYD LAEE - SGWYSHDG EKLD ELLN	AQ C
Su Human Chimp Macaque Bovine Mouse Rat Opossum Pufferfish Zebrafish Rice Arabidopsis Mosquito (Aedes) Mosquito (Aedes) Mosquito (Anopheles) Fly (D. persimillis) Fly (D. persimillis) Fly (D. melanogaster) Trematode (S. mansoni) Nematode (C. elegans) Nematode (C. remanei) Nematode (C. remanei)	er81 56TLGHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SOVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGHPGSL DETTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SOVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGHPGSL DDTTYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SOVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGHPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SOVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGHPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTF EDYDVSFG SOVL TVKLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL EDYDVSFG DOVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG DOVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGNPSSL DETAYERLAEETLD SLAEFFEDL AD KPYTL KDYDVSFG DOVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 56TLGNK.NSL DETTYEKLAEETLD SLADFFEDL OD KPTS KDYDVSLG SOVL TI KLGGDL GTYVINKOTPN.KQI WL.SSP5 57APVQI.SEL SEAAYEKLVDETLD ALAEFFEDL TD AAFTG AD YDVYF SVSL PAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	SGPKRYDWTG- KNWVYSHDGVSLHELLAAEL TKALKT - KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLAAEL TKALKT - KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLAAEL TKALKT - KLDLSSLAYSGKDA- SGPKRYDWTG- KNWVYSHDGVSLHELLATEL TQALKT - KLDLSSLAYSGKDT - CCPA SGPKRYDWTG- KNWVYSHDGVSLHELLATEL TQALKT - KLDLSSLAYSGKDT - CCPA SGPKRYDWTG- KNWVYSHDGVSLHELLAREL TKALNT - KLDLSSLAYSGKGT SGPKRYDWTG- KNWVYSHDGVSLHELLAREL TEALNT - KLDLSSLAYSGKGT SGPKRYDWTG- KNWVYSHDGVSLHELLAREL TEALNT - KLDLSSLAYSGKGT SGPKRYDWTG- KNWVYHDGVSLHELLEMEF SGTLKT - QLDLSSLVYSGKDT SGPKRYDWSG- ERWVYTHDGVSLHELLSGEF SA IF SR - DIDLLDLPCS SGPSRFDWDAP - TNOW IYRRTGANLVELLSKELS I IF KT - NIDLSHL IHS SGPSRFDWDAP - TNOW IYRRTGANLVELLEKEI GELCGT - PVELS SGPSRFDWDAP - ANAWIYRRTEAKLHKLLE EEL EN LGGE - P1QLS SGPSRFDWDAP - ANAWIYRRTEAKLHKLLE EEL EN LGGE - P1QLS SGPKRYDF I PDKSKVNE GF WLYKHDGVSLHELLQKE I GV I VQK - EVDFL SLPHSKRLD SGPKRYDF I PDKSKVNE GF WLYKHDGVSLHELLQKE I GV I VQK - EVDFL SLPHSKRLD SGPKRYDF I PDKSKVNE GF WLYKHDGVSLHELLQKE I GV I VQK - EVDFL SLPHSKRLD SGPKRYDF I PDKSKVNE GF WLYKHDGVSLHELLQKE I GV I VQK - EVDFL SLPHSKRLD SGPKRYDF I PDKSKVNE GF WLYKHDGVSLHELLQKE I SA I VGR - ELEFF ALPHSQRPQEPASDGR SGPKRYDF VGTV - AAGRWIYKHSGGSLHELLQGE I PG I LKSQ SVDFL RLPYGS SGPKRYD Y I PS MRLWIYKHDGK SLH SLLS EE I SA I VGG - PVEFGT - SGPKRYD Y I PS MRLWIYKHDGGSLHELLQGE I PG I LKSQ SVDFL RLPYGS SGPKRYD Y I PS MRLWIYKHDGE SLLNREF RK I LADD RI DFSRHV SGPKRYD LEEE - GKWYYSHDGENLD LLNREF RK I LGD DR I DFSRHV SGPKRYD LAEE - O SWKYSHDGEKLD ELLNREF RK I LGD DR I DFSRHV- SGPKRYD LAEE - O RWYYSLRMG TKLTD IL TEEVEKAI SK SQ	AQ C

Supplementary Fig. S6 - N-terminal sequence conservation among mammalian frataxin proteins

The sequence alignment of frataxin homologues was created with ProbAlign (2), and sequences were taken from a TreeFam phylogenetic tree of genes with fully sequenced genomes (3). The alignment reveals conservation of residues 81-210 across a broad range of eukaryotes (shown in red) and conservation of preceding residues only among mammals (shown in blue). Color density indicates the level of conservation of a residue. *Leu42*, *Ser56*, and *Ser81* indicate the N-termini of FXN⁴²⁻²¹⁰, FXN⁵⁶⁻²¹⁰, and FXN⁸¹⁻²¹⁰, respectively; *64-87* highlights a putative oligomerization domain that may enable mammalian frataxins to achieve iron-independent oligomerization.

REFERENCES CITED IN SUPPLEMENTARY FIGURE LEGENDS

- 1. Li, H., Gakh, O., Smith, D. Y. IV, and Isaya, G. (2009) J. Biol. Chem. 284, 21971-21980
- 2. Roshan, U., and Livesay, D. R. (2006) *Bioinformatics* 22, 2715-2721
- Ruan, J., Li, H., Chen, Z., Coghlan, A., Coin, L. J., Guo, Y., Heriche, J. K., Hu, Y., Kristiansen, K., Li, R., Liu, T., Moses, A., Qin, J., Vang, S., Vilella, A. J., Ureta-Vidal, A., Bolund, L., Wang, J., and Durbin, R. (2008) *Nucleic Acids Res.* 36, D735-740