



Supplementary Materials to

Import of *Entamoeba histolytica* mitosomal ATP sulfurylase relies on internal targeting sequences

Herbert J. Santos^{1,2,3}, Yoko Chiba^{2,3,4}, Takashi Makiuchi^{2,5}, Saki Arakawa^{3*}, Yoshitaka Murakami^{3*}, Kentaro Tomii⁶, Kenichiro Imai⁷, and Tomoyoshi Nozaki^{1,2,3**}

Plasmid name	Orientation	Sequence (5' - 3')
EhAS-HA	sense	ACAAACACATTAACAGATCTATGAGCATTCAAGAAAACTTAAAC
	antisense	TCATATGGATACATAGATCTTTTCATGGCATCACCAGTAGC
EhAS(DvA)-HA	sense	ATGTCAAAGTTAGTTCCAG
	antisense	TTCACCTTCTGATAAACTTTAAC
DvAS(EhA)-HA	sense	TTATCAGAAGGTGAATTTCCAATT
	antisense	AACTAACTTTGACATGTTTTCTTGAATGCTCATAGA
$EhAS(D_{2}B)$ UA	sense	TTTCCAACAAAGTTTGCAG
<i>LIIA3(D0</i> D)-11A	antisense	TCTACCTATAATCATACGATTAAT
DvAS(EhB)-HA	sense	ATGATTATAGGTAGAGATCATGCT
	antisense	AAACTTTGTTGGAAATTCTCCTTCTGAAA
EhAS(DvC)-HA	sense	GGAGTTGGAGATTTCTAT
	antisense	TCATATGGATACATTACTGAACCTGATGCA
DvAS(EhC)-HA	sense	TCAGGTTCAGTAATGTATCCATATGATGTTCCAGA
	antisense	GAAATCTCCAACTCCAGCAT
EhAS(Dv1-37)-HA		(phos)TAAAAGACAACAAACAAGTCCTTTTCCTCCATGAGCTGGAACTAAC
	sense	IIIGACAIAGAICIGIIAAIGIGIIIGICAGIICA (phon)TTTTCCTCCATCACCTCCAACTAACTTTCAACATCTCTTAATCTCCTT
	antisense	ТС
EhAS(Dv16-37)-HA		(phos) TTAAAAAAAGCTGCTGGTCTTAAACAAATAGAGATTTCATCACGAG
	sense	AATTTGG
		(phos)TTCAGCTTCTATCAGCTCCTTCTAAAAGACATTTAATGAGTTTTCC
	antisense	ICCAIG
EhAS(Dv42-47)-HA	sense	(phos)GACTTAATTATGATGGGAATTGGAGG
	antisense	(phos)TCCTTTTGCTCGTGATGAAATATCAATC
EhAS(Dv64-73)-HA	sense	(Phos)GTTTGTGAAAAAATGACACTTGCTGATGGAACA
	antisense	(phos)AGATTTCCAATACGCTTTTTTCATAAATCCATTTAATGG
EhAS(Dv125-139)-HA	sense	(phos)TGAGTTAGTATTTAAAGGAGAATCAGACCATCCAGG
	antisense	(phos)CATTCCCATTTCTTGTCTGCTTCCTTAGTCATTTCATAAACTTGG
EhAS(Dv165-174)-HA	sense	(phos)GAACAGTTAAAGTTCTTTCAGAAGGAGAATTTC
	antisense	(phos)CTGCAATGTTAAATTCTTTTTGTTCCATAACTTTTTTAAC

Table S1. List of primer sets used in this study.





E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	1	69 69 69 69 69 100 64
E.histolytica E.nuttalli E.dispar E.mosokkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	70 DWF SVCKNFTLADG TFWPIPITINS SEEEAKKLKRGOKVALKYNKDINDISG TIEIDQVYENTKKDKEEGKDIFTTTD SDHPGVKKVM 70 DWF SVCKNFTLADG TFWPIPITINS SEEEAKKLKRGOKVALKYNKDINDISG TIEIDQVYENTKKDKEEGKDIFTTD SDHPGVKKVM 70 DWF SVCKNFTLADG TFWPIPITINS SEEEAKKLKRGOKVALKYNKDINDISG TIEIDQVYENTKKDKEEGKDIFTTD SDHPGVKKVM 70 DWF SVCKNFTLADG TFWPIPITINS SEEEAKKLKRGOKVALKYNKDINDISG TIENDQVYENTKKDKEEGKDIFTTD DHPGVKKVM 70 DWF SVCKNTLADG TFWPIPITINS SEEEAKKLKRGOKVALKYKKDNDISG TIENDQVYEL KKDKEMGKDIFTTD DHPGVKKVM 70 DWF SVCKNTLADG TFWPIPITINS SEEDAKKLKRGOKVALKYKKDNDISG TIENDQVYEL KKDKEMGKDIFTTD DHPGVKKVM 70 DWF SVCKNTLADG TFWPIPITISS SEEDAKKLKRGOKVALKYKKDNDISG TIENDQVYEL KKDKEMGKDIFTTD PEHPGVKKVM 70 DWF SVCKNTLADG TFWPIPITISS SEEDARKLRRGOKVALSYKKDQ0PISGTIAVDEIYEMTKKDKENGKDIFTTD KDFPGVKKVM 70 DVFSVCMNTLADG TFWPIPITISSSEEDARKLRRGOKVALSYKKDQ0PISGTIAVDEIYEMTKKKARGOKALSYKENGORVALSYKKDQ0 SKKDKRGKAND 70 DVFSVCMTLADG TFWPIPITISSSEEDARKLRRGOKVALSYKKDQPISGTIAVDEIYEMTKKKARGORVALSYKENGORVAL	158 158 158 158 158 154 183 161
E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	159 EQ. KPFNVAGKVTTLSEGEFFI KYKGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KAFNVAGKVTTLSEGEFFI KYKGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KAFNVAGKVTTLSEGEFFI KYKGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KPFNVSGKVTTLSEGEFFI KYKGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KPFNVSGKVTTLSEGEFFT KYKGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KPFNVSGKVTLSEGEFFT KYAGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 EQ. KPFNVSGKVTLSEGEFFT KYAGIYMTPEESRLNFAKKGWKTI AALOLRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 159 KAAOVYVGSEGEGI OLIFHVYOFVELROTPKEIRASMAAKSWSRM/AACIRNPMHRSHEFLAKIAV. EVCDGVFIHSLVGNLKPGDIPAEVRVKCID 154 NQ. RRFNVGSSKVVLSEGEFFT KYAGIYMRPEESRKFFAERGWRTI AALOLRNPMHRSHEFLAKIAV. EVCDGVVIHSLVGSLKPGDIPAEVRVKCID	254 254 254 254 254 251 279 257
E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	255 ALV D KYF V KKN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G O T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Y N A D P K K R L L T Q PM K I D M T F 255 ALV D KYF V KKN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G C T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Y N A D P K K R L L T Q PM K I D M T F 255 ALV D KYF V KKN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G C T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Y N A D P K K R L L T Q PM K I D M T F 256 ALV D KYF V K KN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G O T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Y N A D P K K R L L T Q PM K I D M T F 256 ALV D KYF V K KN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G O T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Y N A D P K K R L L T Q PM K I D M T F 256 ALV D KYF V K KN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G C T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Q N X D F K K R L U T Q PM K I D M T F 256 ALV D KYF V K KN VL Q G G Y PL D M R Y G G P RE ALL HA T F R Q N Y G C T HM I I G R D HA G Y G D YY G PF D A Q K I F D K I P Q N X	344 344 344 344 344 338 369 351
E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	345 YCH K CDOMASLET CPHTKKDRVIVSOT MVRKMLSEG K TLPDHFOR AESLKIL ADYYQH DKSKKVTIKLQKFA TODAM	423 423 423 423 423 427 443 427
E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulgaris	428 <mark>KKNROLL</mark> SS ELOFSRTDRDLNMRRNOYVA SENTKSOGVAVLTANSPYEEGRRTDRDMVSRHOGFFENYNSTPLEVCEERDRKGLYAKARROLNTDFTG <mark>V</mark> N	527
E.histolytica E.nuttalli E.dispar E.moshkovskii E.invadens A.castellanii M.balamuthi D.vulcaris	528 DPYEPPKSPENTINDASRVSWTNAWOTILIAALEEEGYLAOWK	569

Figure S1

Multiple sequence alignment of ATP sulfurylase of various *Entamoeba* species namely, *Entamoeba* histolytica, *E. nuttalli, E. dispar, E. moshkovskii,* and *E. invadens,* as well as that of *Acanthamoeba* castellanii, Mastigamoeba balamuthi, and Desulfovibrio vulgaris. The figure was created using Clustal Omega with the default parameters [1]. The alignment was displayed using Jalview [2]. The hydrophobic, positively charged, negatively charged, hydrophilic, glycine and proline residues are colored green, blue, red, light blue, yellow and orange respectively. Red box indicate the N-terminal targeting sequence of *MbAS* predicted by MitoFates [3]. Black boxes indicate the position of the two internal targeting sequence in *EhAS*.







(a)





Figure S2

Structural features of *Eh*AS internal targeting sequences. (a) Surface charge of *Eh*AS. Positively (blue) and negatively (red) charged surfaces of both sides, including two ITS segments: 133-147:KKDKEMECKDIFTTT (left panel) and 179-203:IKYKGIYMTPEESRLNFAKKGWKTI (right panel), of *Eh*AS are depicted. The surface, other than two identified regions, of AS is transparent. (b) The locations and structures of two ITS segments in *Eh*AS. Two ITS segments: 133-147:KKDKEMECKDIFTTT (left panel) and 179-203:IKYKGIYMTPEESRLNFAKKGWKTI (right panel) and their surface lysine and arginine residues are shown in magenta. Three blocks are indicated by red, blue, and green, respectively.

Reference:

- Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011, 7, doi:10.1038/msb.2011.75.
- 2. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191, doi:10.1093/bioinformatics/btp033.
- 3. Fukasawa, Y.; Tsuji, J.; Fu, S.-C.; Tomii, K.; Horton, P.; Imai, K. MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites. *Mol. Cell. Proteomics* **2015**, *14*, 1113–1126, doi:10.1074/mcp.M114.043083.