Species-independent translational leaders enable a rapid development of novel cell-free expression systems

Supplementary information

Nature Biotechnology: doi:10.1038/nbt.1556

Design and analysis of SITS sequences

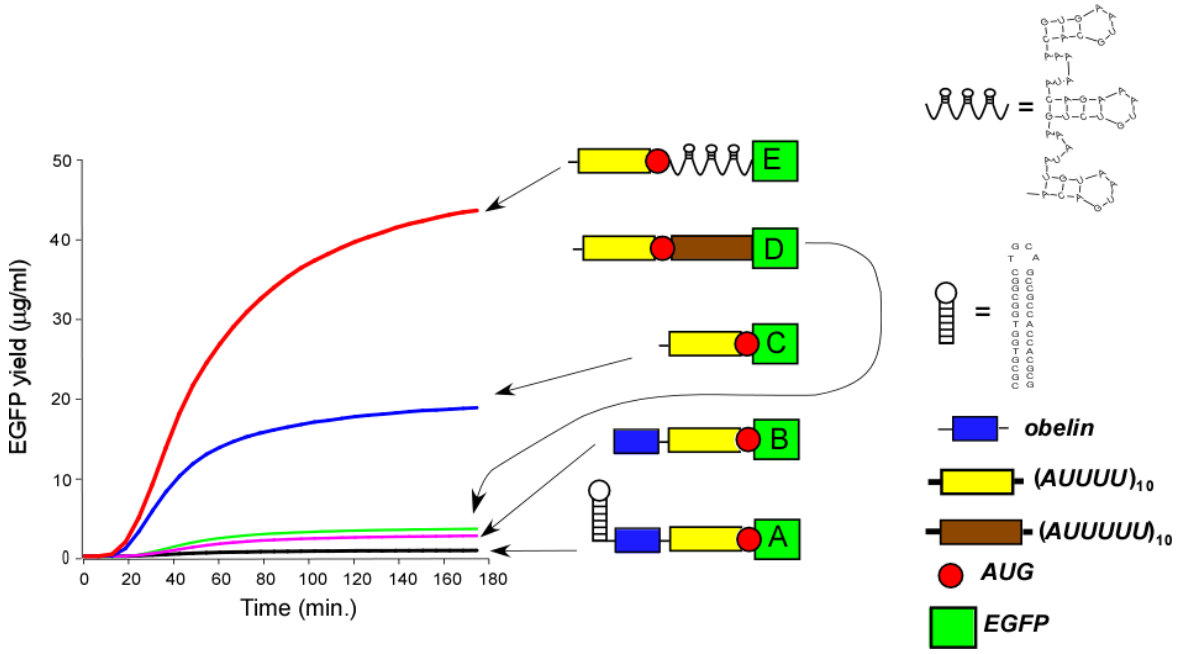

Figure S1. Translation of mRNAs coding for *egfp* gene with polyAU SITS in LTE. Equal amounts (0.6 pmol/μl) of indicated mRNAs were translated in LTE and changes in fluorescence were recorded using fluorescence plate reader. The sequence elements comprising the translated mRNAs are denoted in the inset.

Nature Biotechnology: doi:10.1038/nbt.1556

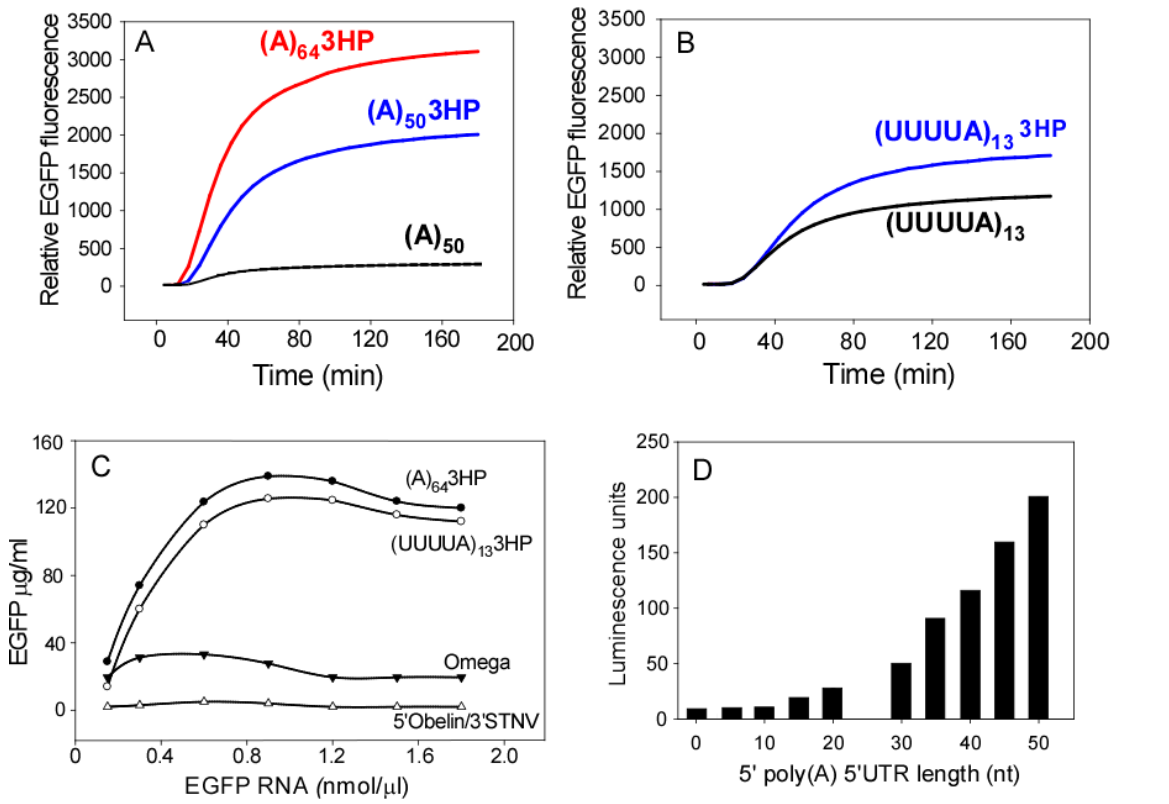

Figure S2. Analysis of the structure-activity relationship of SITS-containing mRNAs in LTE. (A) influence of the 3 hairpins downstream of AUG sequence on the efficiency of EGFP translation in LTE extract. Equal amounts (0.6pmol/μl) of indicated EGFP mRNAs were translated in LTE and changes in fluorescence were recorded (B) As in A but using $(UUUUA)_{13}$ based translational leader. (C) Dependence of the EGFP expression yield in LTE on the mRNA concentration. Identical volumes of LTE were supplemented with increasing amounts of mRNAs carrying the indicated translation leaders and the fluorescence of the reaction mixtures was measured after 2 hours. (D) Influence of the length of polyA 5′ UTR of luciferase gene on the efficiency of its translation in LTE. Equal amounts (0.4 pmol/μl) of luciferase mRNAs were translated in LTE and luminescence measured using the luciferase assay kit (Stratagene).

Nature Biotechnology: doi:10.1038/nbt.1556

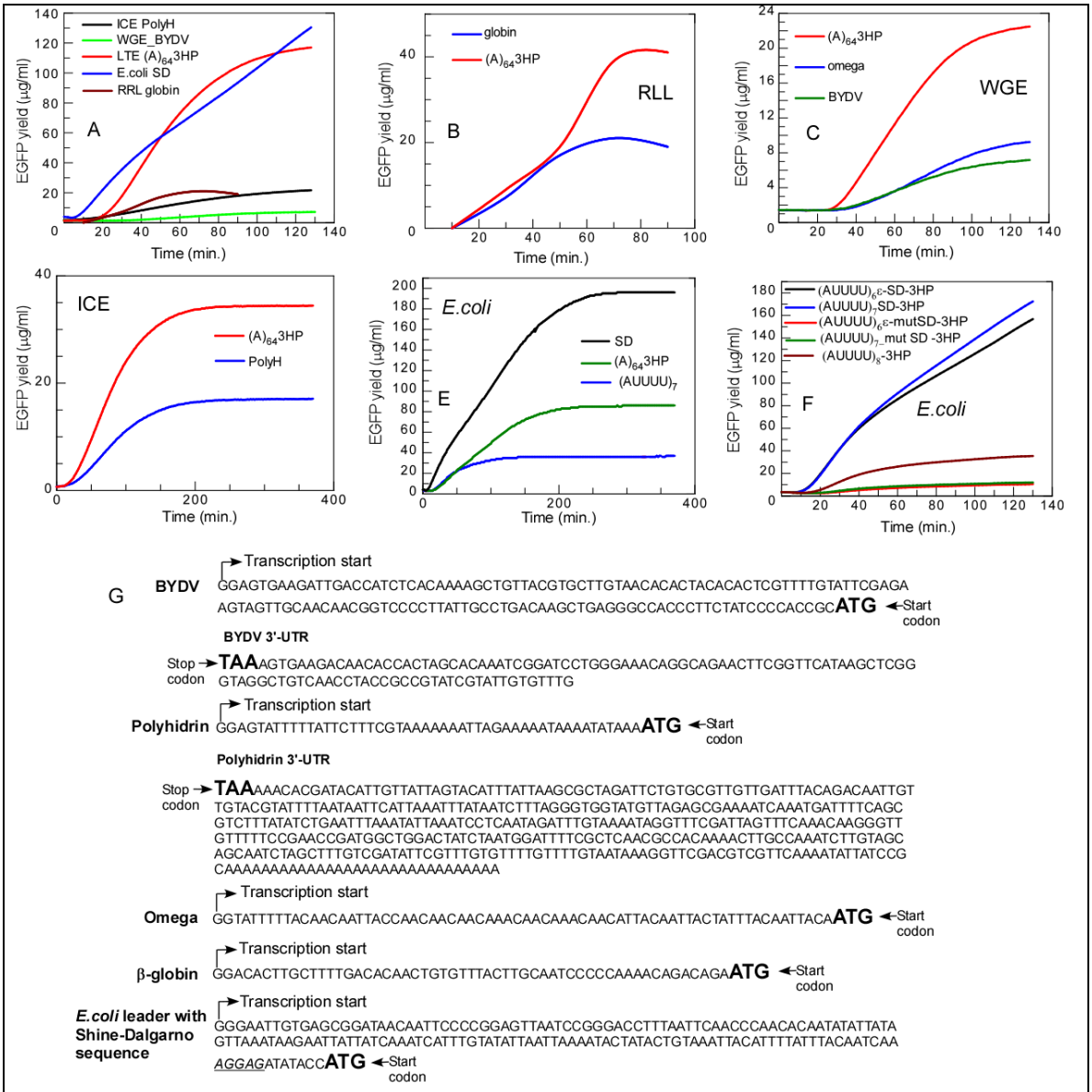

Figure S3. Analysis of translation efficiencies of EGFP encoding mRNA with different translation initiation sequences in coupled prokaryotic and eukaryotic cell-free expression systems. (A) translation of EGFP encoding PCR products in coupled LTE and commercially available coupled cell-free translation systems. The EGFP ORF was amplified by overlap-extension PCR to tag it with the indicated translational leaders at the 5′ or 3′ termini. The leaders were chosen according to the instructions of the cell-free system manufactures and are summarized in panel G. All translation reactions with the exception of RRL were incubated for 2 hours and changes in fluorescence were monitored using a fluorescent plate reader as described above. During this period all systems except *E. coli* came to complete or near saturation. The productive phase of the *E. coli* system was saturated after 4 hours resulting in total yield of approx. 200 μg/ml

Nature Biotechnology: doi:10.1038/nbt.1556

(see panel E). (B) comparison of the expression yields of Rabbit Reticulocyte Lysate (RRL) primed with EGFP PCR products bearing either a globin or $(A)_{64}$3HP leader. (C) as in B but priming Wheat Germ Extract (WGE) with EGFP templates bearing $(A)_{64}$3HP, omega or Barley yellow dwarf virus (BYDV) mRNA leader sequence. (D) as in B but priming the Insect Cell Extract (ICE) with $(A)_{64}$3HP and Polyhidrin leader (PolyH). (E) priming of *E.coli* cell-free system with the PCR products encoding for EGFP with the indicated translation leaders. (F) as in E but testing different combinations of translation initiation enhancing sequences. The latter results indicate that although SITS can substitute for the Shine-Dalgarno sequence in prokaryotes their activities are not additive. (G) translational leaders used in this study for comparative analysis of the *in vitro* translation systems. The BYDV and polyhidrin 3' UTRs are shown while in all other cases the 100 nt 3'UTR was derived from pTUB-SITS-EGFP plasmid.

Optimization of cell-disruption procedure

The ability of cell extracts to retain protein synthetic capacity is largely dependent on the approach chosen for cell disruption and the subsequent steps of the lysate fractionation [1] [2]. We assumed that the better the organelle's integrity was preserved during cell breakage the fewer of translation inhibitory factors would be released and less stress induced reactions develop. We compared several methods of *L.tarentolae* homogenization by evaluating the activity of succinate-cytochrome C reductase complex as a measure of the integrity of outer mitochondrial membrane [3,4,5].

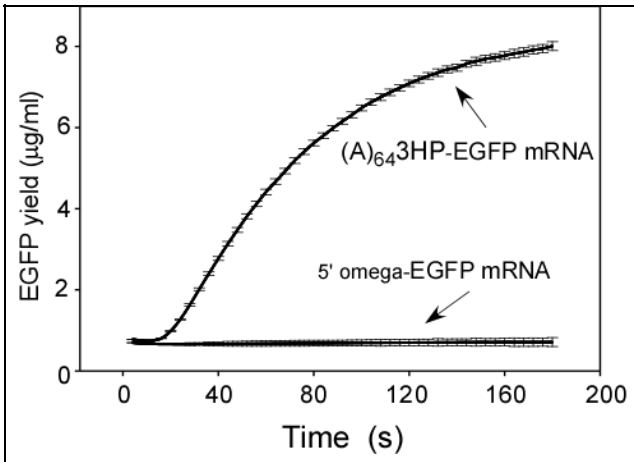

Figure S4. Translation of SITS-EGFP and omega-EGFP mRNAs in extract of *S. cerevisiae* prepared by hypotonic lysis. The lysate prepared by nitrogen cavitation produced similar results (not shown).

Nature Biotechnology: doi:10.1038/nbt.1556

Mitochondrial fraction was isolated either by Percoll gradient or by differential centrifugation in 250 mM sucrose [6] and adjusted by protein concentration. Reduction of

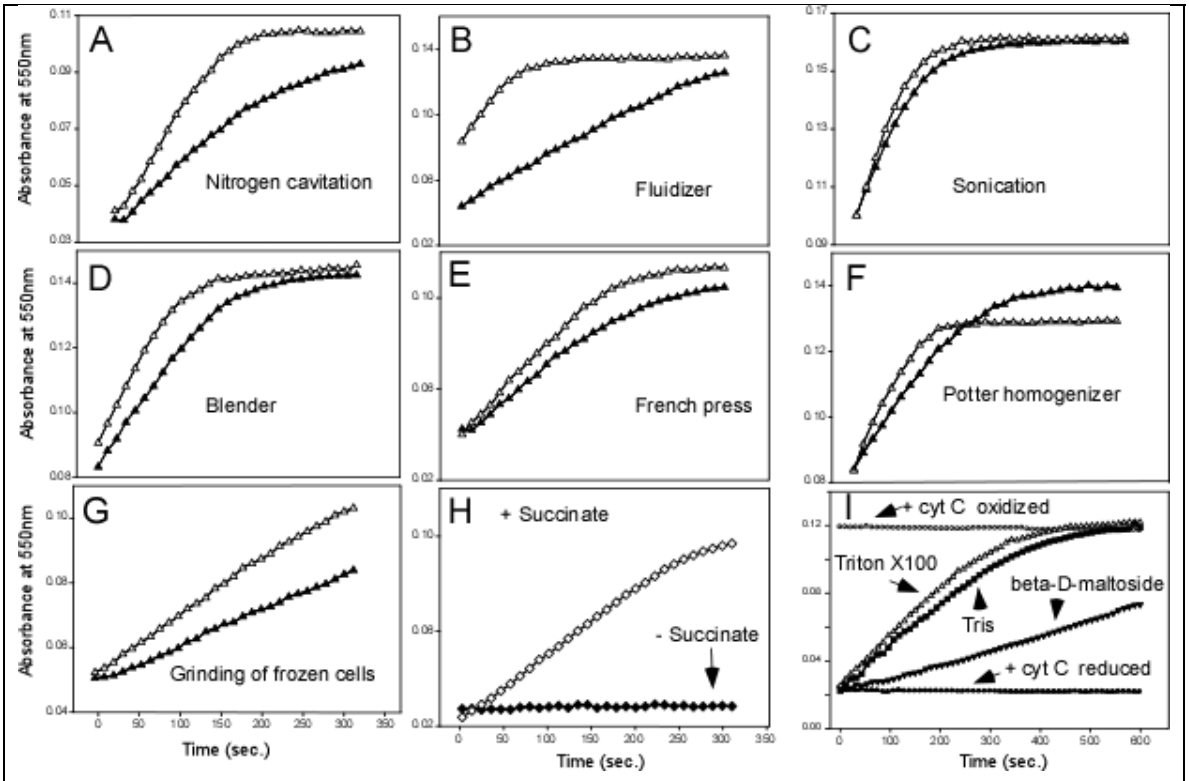

Figure S5**.** (A-G) Activity analysis of cytochrome c in mitochondria preparations isolated from *L. tarentolae* cells disrupted by different methods. The filled triangles reflect the activity of the samples. The empty triangles reflect the total activity of the preparation following TritonX-100 treatment of the fractions. (H) calibration of the assay with reduced and oxidized cytochrome c in the reaction buffer. (I) 1 mM of detergent n-dodecyl β-D-maltoside negatively influenced the enzyme's activity whereas 0.5 mM TritonX-100 had no influence on the reaction compared to the buffer control (10 mM Tris pH 7.4).

exogenously provided oxidized form of cytochrome c by mitochondria sample leads to an increase in extinction at 550 nm. The total amount of reductase activity was obtained by measuring the activity of the detergent treated samples. The ratio of activities of detergent treated and untreated samples was used as a measure of membrane integrity. Mitochondria samples obtained by nitrogen cavitation or by grinding of liquid N_2 saturated frozen cells demonstrated the highest integrity (Fig. S5) as compared to the

Nature Biotechnology: doi:10.1038/nbt.1556

samples obtained by sonication, cell-disruption in Potter homogenizer, homogenization by glass-bead blender or disruption with the French press.

Microfluidic disruption yielded mitochondria with anomalously high apparent outer membrane integrity (Fig. S5B). This is surprising considering the harshness of the procedure. To assess mitochondria integrity using an alternative method we microscopically analyzed the shape of DAPI stained kinetoplast. The kinetoplast is a large single mitochondria associated with the flagellum which appears as a single spot in intact cells. We concluded on the basis of microscopic analysis shown in figure S5 that the cell lysates obtained by nitrogen cavitation contained the highest number of intact mitochondria while fluidizing resulted in their nearly complete destruction. Thus nitrogen cavitation is believed to be the most suitable method for preparation of *L. tarentolae* lysate.

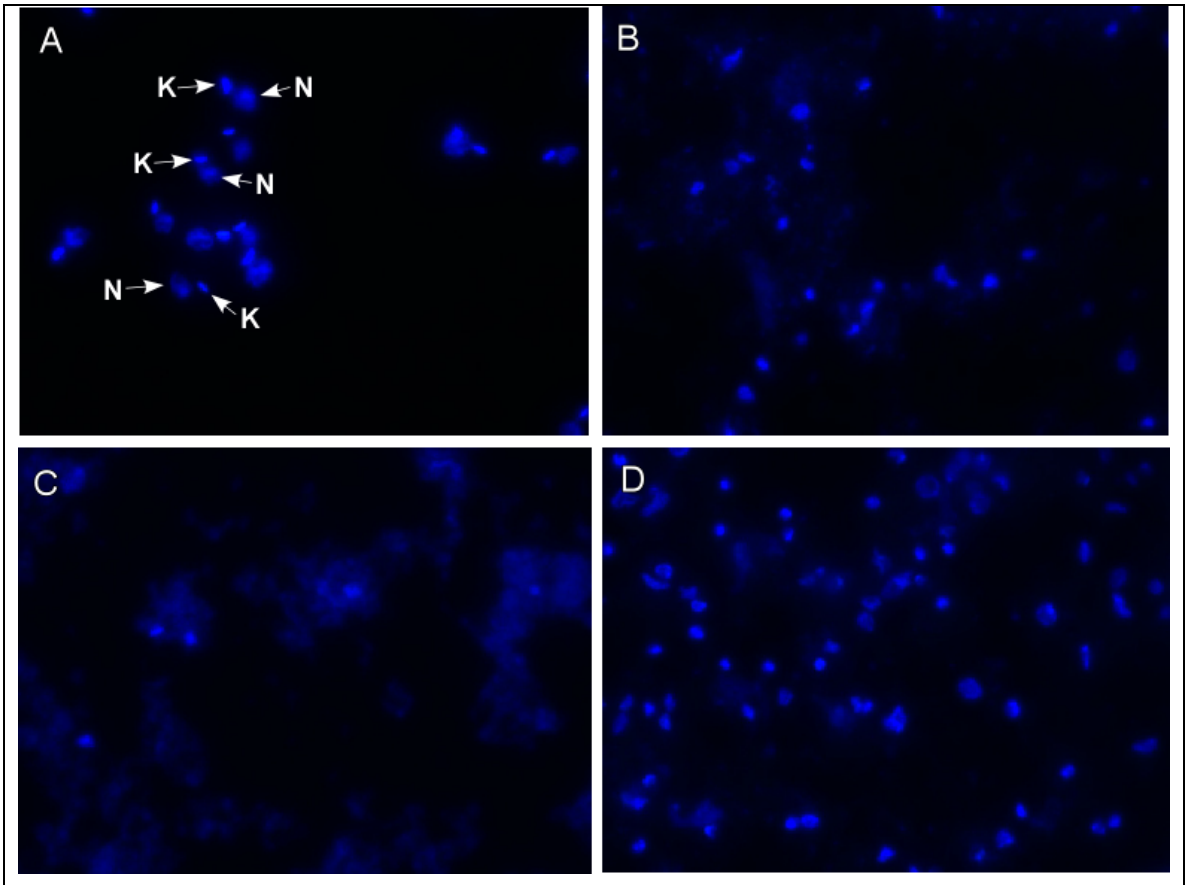

Figure S6. Fluorescence microscopy analysis of DAPI-stained kinetoplasts in *L. tarentolae* cells and mitochondria preparations obtained by different cell disruption protocols: (A) - *L. tarentolae* cells, the arrows denote nuclei -N and kinetoplasts - K; (B) - mitochondria sample obtained form cells disrupted by French press; (C) - as in B but disrupted with

Nature Biotechnology: doi:10.1038/nbt.1556

fluidizer; (D) –mitochondrial fraction after disruption by nitrogen cavitation.

Translation of endogenous and exogenous mRNAs in *L. tarentolae* extract (LTE)

Fig. S7 shows the incorporation of [C^{14}]Leu into the proteins expressed in LTE primed with endogenous mRNA. Addition of purified homologous polyA mRNA fraction led to a further increase in [C^{14}]Leu incorporation, indicating that the system could initiate translation on the exogenously added templates. To further confirm this we treated LTE with micrococcal nuclease to destroy all endogenous mRNAs. Optimization of micrococcal nuclease concentration demonstrated that incubation with 4 unit/ml for 20 min. at 20°C [2, 7] resulted in the lowest background and the highest translational activity of the extract. Addition of homologous polyA RNA to the extract treated with micrococal nuclease resulted in reappearance of translation products.

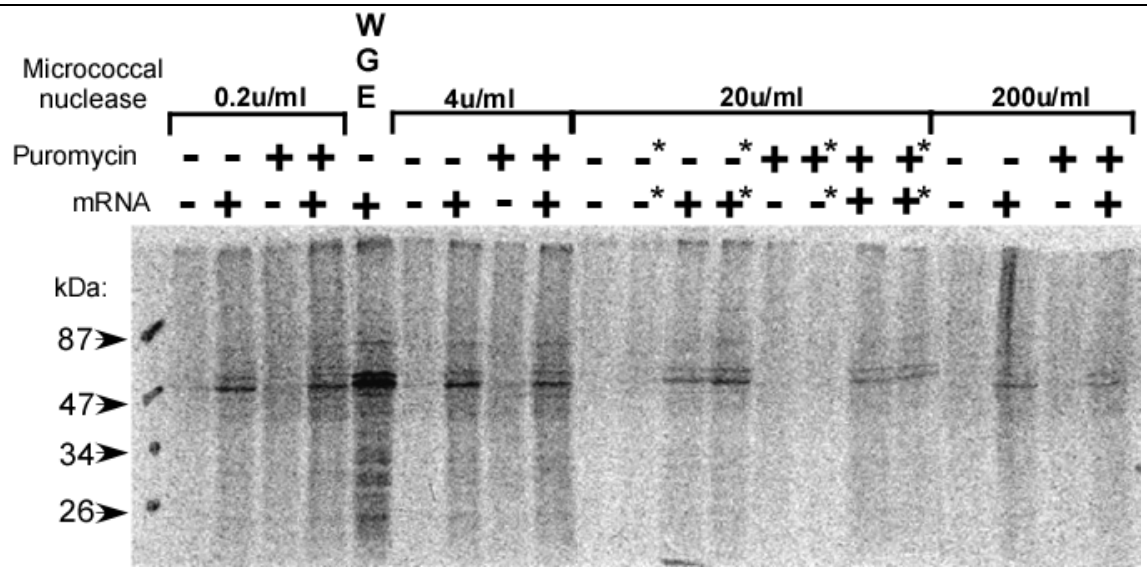

Figure S7. Translation of endogenous mRNA in LTE. SDS-PAGE analysis of the of [C^{14}] Leu supplemented LTE containing purified polyA fraction of *L. tarentoale* mRNA (60 ng/ul) and treated with puromycin or micrococal nuclease. Micrococcal nuclease was used at indicated concentrations to treat extracts prior to gel-filtration; Puromycine was added at final concentration of 1 mM to the crude cell extract prior to differential centrifugation and gel-filtration. mRNA - homologous polyA mRNA fraction. WGE wheat germ extract. * - indicates extract treated with micrococal nuclease after gel-filtration.

Nature Biotechnology: doi:10.1038/nbt.1556

Anti-SL oligonucleotide mediated inhibition of translation of SL-bearing mRNAs in LTE

Computer modeling of SL-region indicated the presence of a short imperfect hairpin structure in the middle of the sequence (Fig. S8A). Nevertheless addition of the 35 nucleotide long anti-SL oligonucleotide (CAATAAAGTACAGAAACTGATACTTATATAGCGTT) efficiently suppressed the translation of *L. tarentolae* total mRNA in LTE and WGE

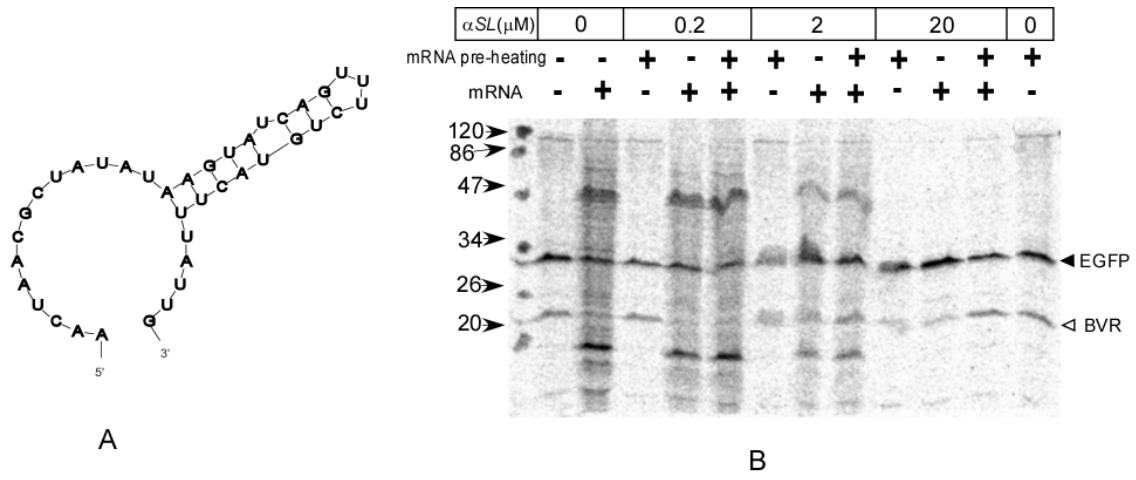

Figure S8. Suppression of *Leishmania* mRNA translation in WGE by anti-SL oligonucleotide. (A) Secondary structure of splice-leader RNA calculated with Mfold 3.2 (ΔG = -3.2 kcal/mol). **(B)** SDS-PAGE analysis of [C^{14}]Leu supplemented WGE primed simultaneously with two RNA templates: the EGFP-encoding RNA bearing obelin-derived 5'UTR and STNV-derived 3'UTRs, as well as RNA encoding for the viral capsid protein derived from brome mosaic virus (BVR). The reaction was carried out in the presence of 100 ng/μl final concentration of total *L. tarentolae* polyA RNA. Prior to translation some of the RNA samples were heated to 68°C and then cooled to 0°C while the others were kept on ice prior to translation. Positions of recombinant EGFP and BVR are indicated by the arrows.

systems at 20 μM concentration (Fig. S8B). Reduction in [C^{14}]Leu incorporation was observed independently of whether the RNA sample was pre-heated in the presence of the oligonucleotide to remove secondary structure prior to translation or not. This indicates that the heteroduplex formation is favored over intermolecular interactions. Neither the 14 nt anti-SL oligo nor 35nt oligonucleotides annealing within the EGFP

Nature Biotechnology: doi:10.1038/nbt.1556

gene repressed the translation of endogenous mRNA at the concentrations up to 20 μM. This suggests the direct mechanism of translation suppression is by interference with ribosomal interactions rather than by an RNAse H mediated process [8] . The anti-*SL* oligo treatment suppressed translation of both the endogenous and re-added mRNA indicating that it competed with pre-formed translation initiation complexes as well as inhibiting translational initiation (Fig. S8)

Translation of heterologous RNAs in LTE mediated by internal ribosome entry site (IRES) sequences of *Leishmania* viruses

We initially attempted to use IRES sequences derived from the genome of *L. guyanensis* (IRESI) and *L. major* (IRESII) viruses IRESI *L. guyanensis* LRV1-4 virus (GenBank accession number U01899) [9] and IRESII (LRV2-1 *L. major* virus (GenBank accession number U32108) [10]). However none of these elements were able to efficiently initiate translation *in vivo* in *L. tarentolae* (not shown) or *in vitro* (Fig. S9A). IRES-combinations mediated similar levels of expression in LTE and WGE (Fig. S9B). We compared these leaders to an artificial 5'UTR derived from the obelin-encoding gene of coral *Obelia longissima,* which in combination with 3'UTR derived from the gene encoding the capsid protein of satellite tobacco necrosis virus, mediates efficient translation in the WGE system [11,12].

Nature Biotechnology: doi:10.1038/nbt.1556

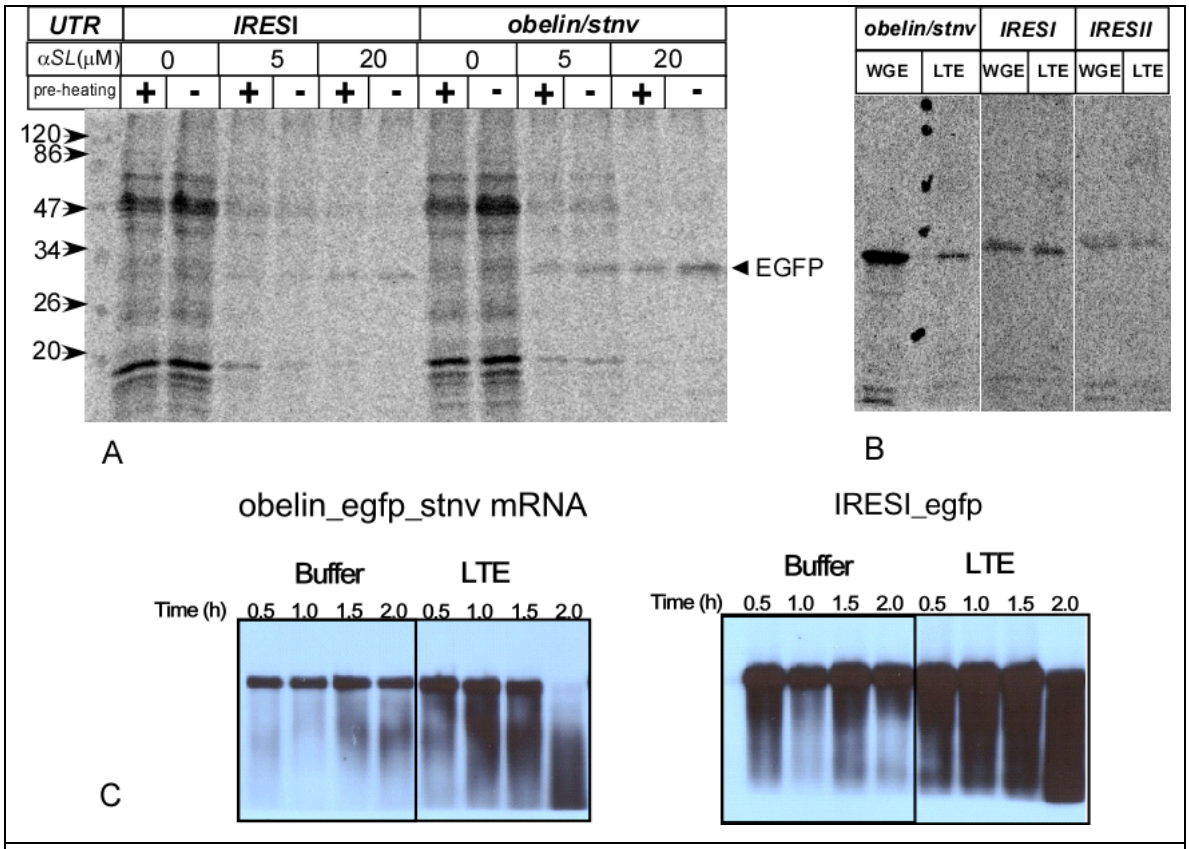

Figure S9. Analysis of translation efficiency and stability of mRNA containing IRESI and *obelin/stnv* UTR combinations in WEG and LTE. (A) The panel depicts SDS-PAGE analysis of the [C^{14}]Leu supplemented LTE primed with EGFP-encoding RNA bearing obelin-derived 5'UTR and STNV-derived 3'UTR, in the presence or absence of 100 ng/μl final concentration of total *L. tarentolae* polyA RNA. The position of EGFP is indicated by the arrow. (B) Comparison of translation efficiencies of different UTR sequences in WGE and LTE systems. mRNA flanked with the indicated translational leaders was translated in [C^{14}]Leu supplemented cell-free systems and the translation mixtures were resolved on SDS-PAGE and processed as described in Fig S9. (C) Analysis of stability of the synthetic mRNAs in LTE. Purified RNAs were added to 20μl of LTE at 50 and 0.2 pmol/μl (left panels) and 50 pmol/μl (right panels) final concentration respectively and incubated at room temperature for indicated periods of time. The samples were separated on 1% agarose gel, transferred to nitrocellulose, and subjected to Northern blot hybridization with DIG labeled EGFP probe. The hybridization products were detected with anti-DIG antibodies.

Nature Biotechnology: doi:10.1038/nbt.1556

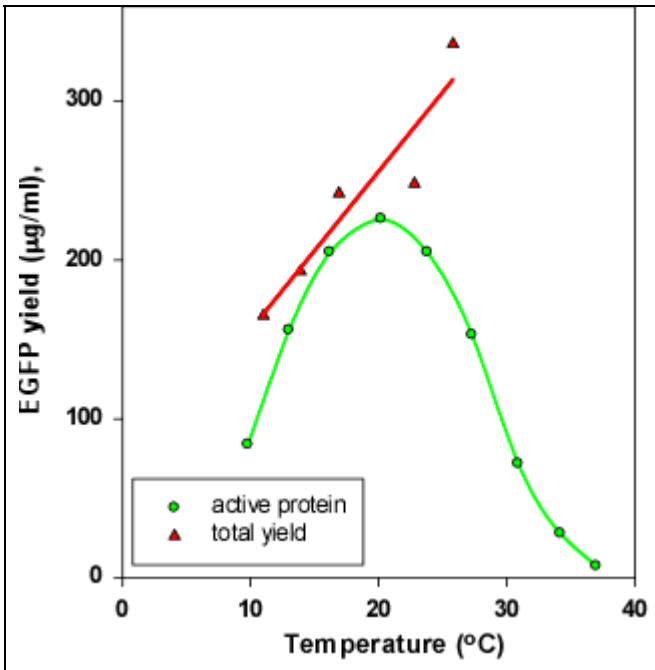

Figure S10. Analysis of EGFP expression level and fraction of active protein as a function of the incubation temperature. 12pmol of myc tagged (A) $_{64}$3HP-EGFP mRNA was incubated with 20µl of LTE at different temperatures and amounts of active EGFP measured by fluorescent scanning using purified EGFP as a standard. The total expression level was determined by Western blotting with anti-myc antibodies using purified myc-tagged meromyosin light chain as a standard.

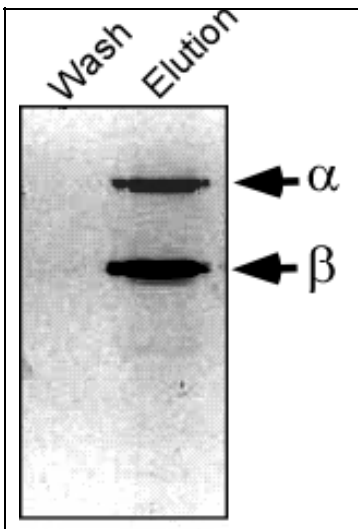

Figure S11. Co-translation in LTE of α and β subunits of Rab geranylgeranyl transferase (RabGGTase) and purification of the α/β heterodimer. The unstained SDS-PAGE gel was subjected to fluorescent scanning (the image is a superimposition of scans in two $\lambda_{ex/em}$ optimally selected for EGFP and Cherry). The arrows denote positions of RabGGTase α and β-subunits.

Nature Biotechnology: doi:10.1038/nbt.1556

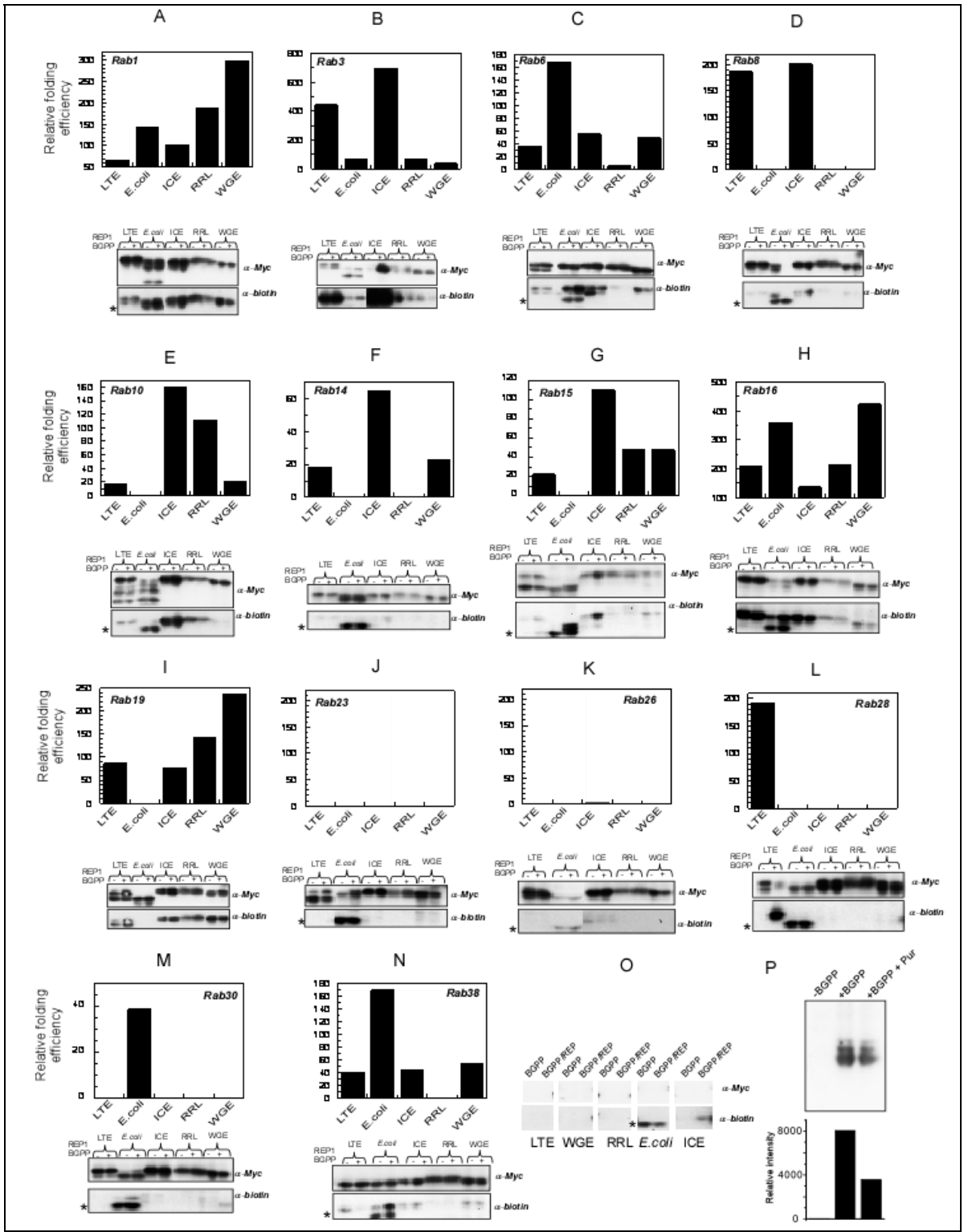

Figure S12. Analysis of folding status of human RabGTPases translated in different cell-free systems. (A-N) Fifteen human RabGTPases were N-terminally fused with myc tag and SITS sequence and translated in LTE, *E. coli,* ICE, RRL and WGE cell-free systems. The resulting translation mixtures were subjected to *in vitro* prenylation with biotingeranyl using recombinant Rab geranylgeranyl transferase (RabGGTase) and Rab

Nature Biotechnology: doi:10.1038/nbt.1556

Escort Protein 1 (REP1). The expression levels were quantified by blotting of the samples with anti-myc antibody. The folding status of the GTPases was assessed by measuring the incorporation of biotin-geranyl into Rab GTPases as determined by blotting with streptavidin-HRP and the densitometry data was plotted as shown in the top panels. Each reaction was performed in duplicate, either unsupplemented or supplemented with REP1 and BGPP in order to account for possible prenylation of recombinant Rabs with endogenous prenyltransferases in the eukaryotic lysates. (O) Prenylaiton of cell-free lysates not primed with Rab genes in order to determine unspecific prenylation and the presence of the endogenous biotinylated proteins. The star denotes the position of endogenous biotinylated *E. coli* proteins. In the case of ICE lysate, prenylation of the endogenous pool of Rab GTPase was observed indicating the presence of background translation. (P) Establishing the origin of the pool of unprenylated RabGTPases in the ICE system. The ICE was incubated with or without puromycin (pur) and subjected to prenylation with BGPP. The reduction of the signal in the presence of puromycin indicates background translation of endogenous mRNAs. The bottom panel shows quantification of the protein-conjugated biotin-geranyl signal.

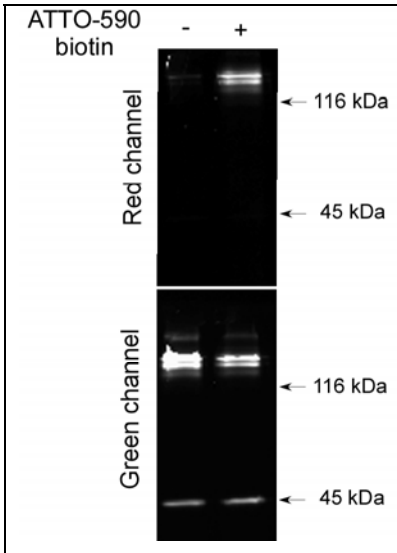

Figure S13. SDS-PAGE analysis of an LTE programmed with $(UUUUA)_7$-streptavidin-EGFP coding mRNA. 10μl of LTE reaction mixture was supplemented with ATTO-590 at 1μM final concentration. After 3 min. incubation at room temperature the samples were resolved on 15% SDS-PAGE. EGFP (green channel) and ATTO (red channel) fluorescence were detected using a Versadoc fluorescent scanner (BioRad).

Nature Biotechnology: doi:10.1038/nbt.1556

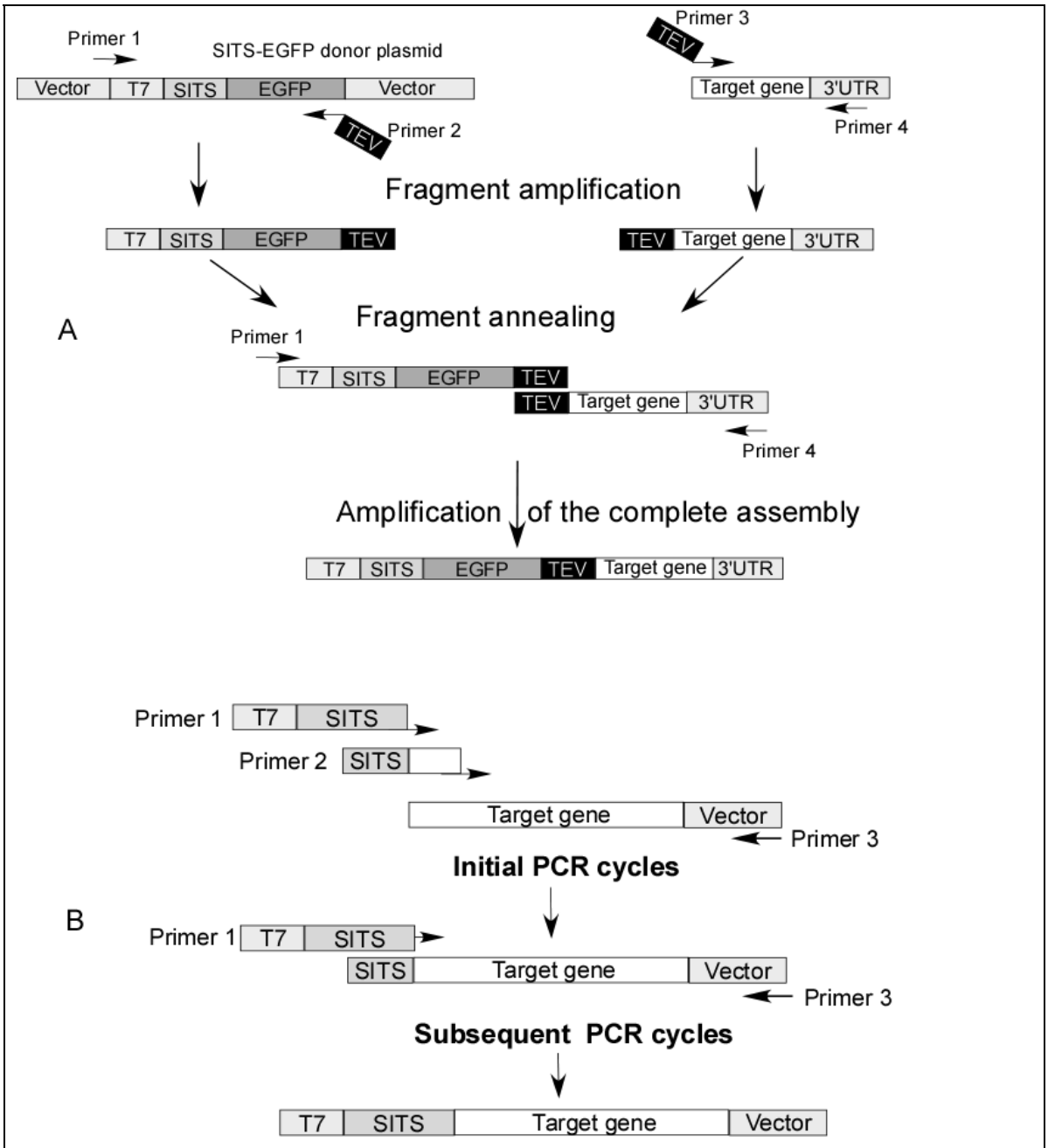

Figure S14. Schematic representation of approaches used for construction of templates. (A) Splicing Overlap Extension based PCR synthesis of the SITS-EGFP tagged ORFs. TEV denotes the recognition sequence of TEV protease. (B) Schematic representation of two primer PCR based procedure for synthesis of SITS tagged genes. Primer 2 is added at 10% of primer 1 and 3 concentrations and is consumed during the initial rounds of PCR.

Nature Biotechnology: doi:10.1038/nbt.1556

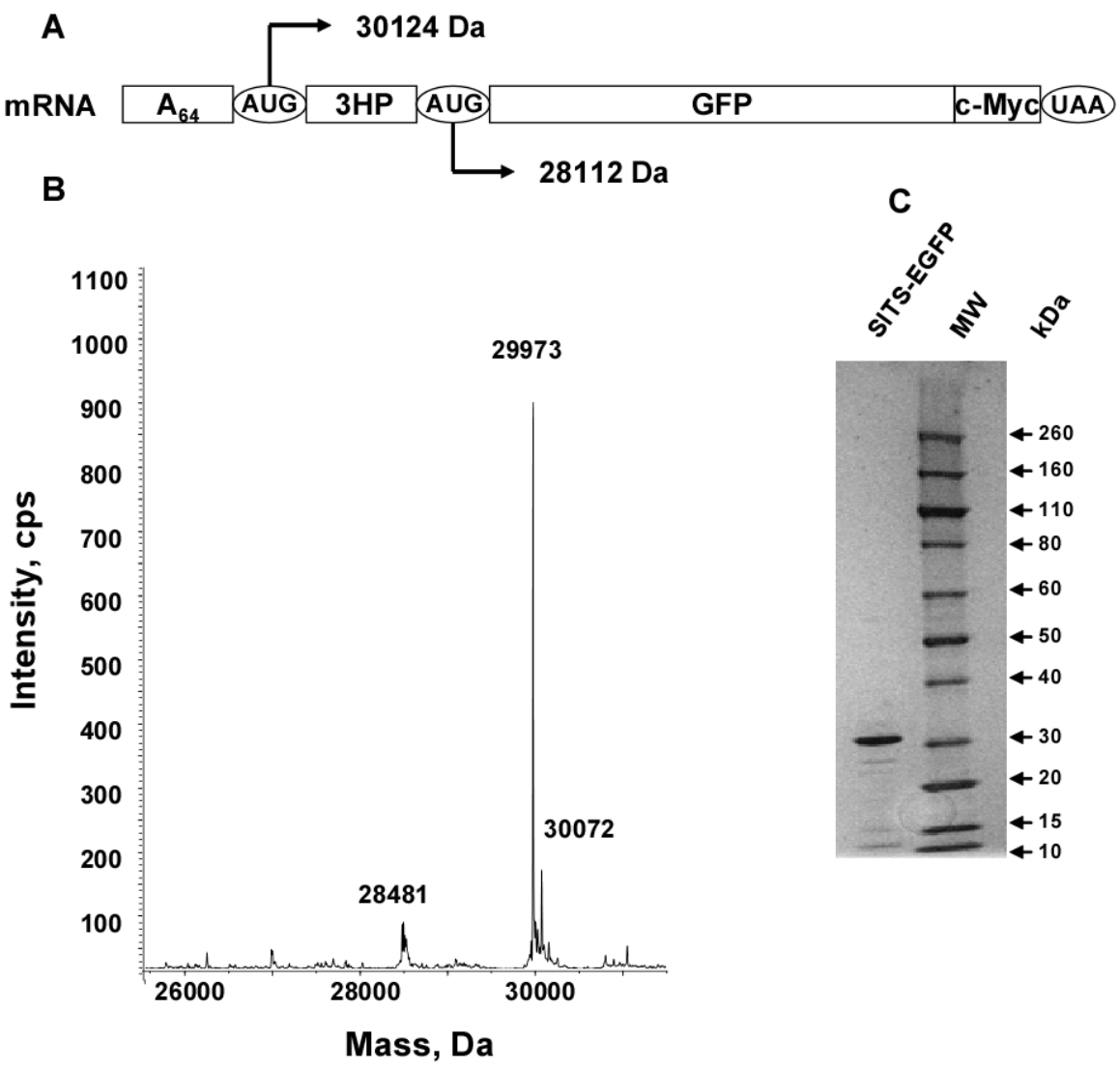

Figure S15. Mapping of the translation start in $(A)_{64}$3HP-EGFP mRNA. (A) Schematic representation of mRNA with two possible translation start sites and the calculated molecular masses of resulting polypeptides; (B) ESI-MS spectrum of purified *E. coli*-expressed SITS-EGFP. The main peak corresponds to the translation product initiated at the first AUG; (C) SDS-PAGE analysis of purified SITS-EGFP.

Nature Biotechnology: doi:10.1038/nbt.1556

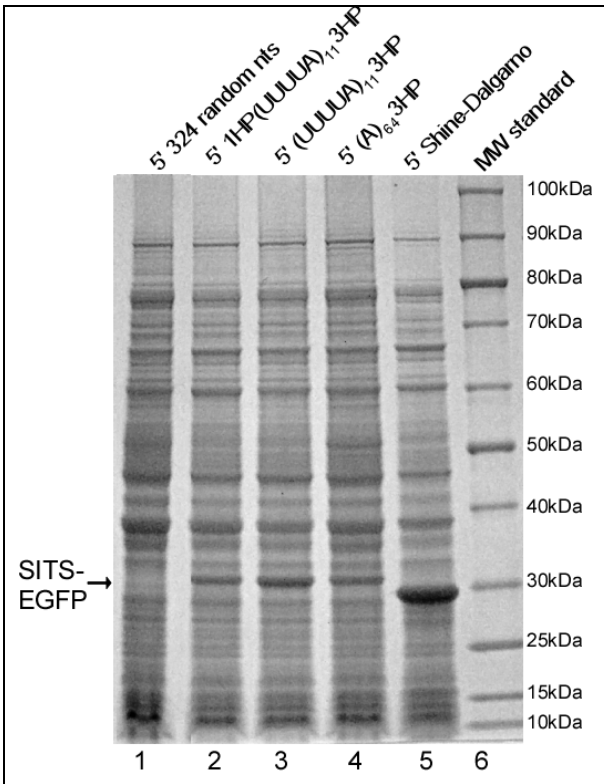

Figure S16. Full size version of the cropped gel from figure 1F. SDS-PAGE gel analysis of lysates of E.coli cells transformed with plasmids carrying indicated SITS sequences (lanes 2,3,4), Shine-Dalgarno sequence (lane 5) or randomly chosen sequence (lane 1). The 1HP of SITS in lane 2 denotes a single stable hairpin blocking the 5' end ((-61 kcal/mol; [13]) and Fig. S1). The identity of the expressed proteins was confirmed by fluorescent scanning of the gels and mass spectrometry of the purified proteins.

Nature Biotechnology: doi:10.1038/nbt.1556

Supplementary methods

PCR synthesis of SITS tagged ORFs

Overlap extension (OE) PCR and two forward primer PCR methods were employed to synthesize ORFs bearing SITS sequence on the 5′ end. In all cases the PCR reaction mixtures contained 500 nM of each primer, 0.2 mM of each dNTP and 25 unit/ml of Taq DNA polymerase (Genecraft, Germany).

Splicing Overlap Extension (SOE) PCR was performed as descried in [14] with modifications. The gene fragments for SOE PCR were generated in 20 cycles of PCR set up in a total volume of 50-100 μl containing 500 nM of each primer. The synthesized DNA products were resolved on 1% TAE agarose gel with ethidium bromide as staining agent and the DNA from respective bends was eluted with QIAQuick Gel Extraction Kit (Qiagen, Germany).

SOE PCR were set up in 20 μl volume of reaction mixture contained 5 nM of gel purified 5′ fragment and a slight excess (8 nM) of 3′ fragment. Thermal cycling included initial DNA denaturation step (2 min. at 95°C), 8 cycles of SOE (denaturation 1 min. at 95°C; annealing 3 min. at 45 or 48°C (depending on the sequence) ; 1 min. elongation for every 1000 nt of fused gene at 72°C). Subsequently 10 μl of SOE PCR mixture was mixed with 70 μl of PCR mixture containing flanking primers for amplification of the fused gene and subjected to further 20 PCR cycles to obtain the final product (Fig. S14).

Two-forward-primer-PCR was used to introduce T7 promoter and short leader sequences into templates. To achieve that an adapter oligonucleotide complementary to the gene of interest and containing 19 nucleotides of SITS sequence was used as a first forward primer. The second forward primer encoded T7 promoter and a full length leader sequence annealing to the adapter oligonucleotide. The PCR reaction mixture contained 50 nM adapter primer, 500 nM of second forward and 500 nM of the reverse primer (Fig. S14).

In vitro translation in WGE system

Nature Biotechnology: doi:10.1038/nbt.1556

In vitro translation in WGE system (Promega) was carried out at 28°C for 2 hours in 20 μl volume containing 0.12 mM of aminoacids and 600 pmol/ml of mRNA according to the instructions of the manufacturer. Translations in coupled WGE system (Promega) were carried out in 20 μl volume containing 0.05 μM of the corresponding PCR product.

In vitro translation in RLL system

In vitro translation in RRL system (Promega) was carried out at 30°C for 1.5 hours in 20 μl volume containing 600 pmol/ml of mRNA (unless stated otherwise) according to the instructions of the manufacturer. Translations in coupled RRL system (Promega) were carried out in 20 μl volume containing 0.05 μM of the corresponding PCR product.

In vitro translation in ICE system

In vitro translations in coupled ICE system (Promega) were carried out at 28°C in 20 μl volume containing 0.05 μM of the corresponding PCR product DNA.

In vitro translation in E.coli RTS system

In vitro translation in *E. coli* rapid translation system (RTS) (Roche diagnostics) was carried out at 30° for 6 hours in 20 μl volume containing 1 μg of linearized plasmid template or 50 ng/μl PCR-product according to the instructions of the manufacturer.

References

1. Carroll, R. & Lucas-Lenard, J. Preparation of a cell-free translation system with minimal loss of initiation factor eIF-2/eIF-2B activity. *Anal Biochem* **212**, 17-23 (1993).
2. Pelham, H.R. & Jackson, R.J. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem* **67**, 247-256 (1976).
3. Wojtczak, A.B., Lenartowicz, E., Rodionova, M.A. & Duszynski, J. Effect of fatty acids on pyruvate carboxylation in rat liver mitochondria. *FEBS Lett* **28**, 253-258 (1972).
4. Gentle, I.E., De Souza, D.P. & Baca, M. Direct production of proteins with N-terminal cysteine for site-specific conjugation. *Bioconjug.Chem.2004.May.-Jun.;15.(3):658.-63.* **15**, 658-663 (2004).
5. Santhamma, K.R. & Bhaduri, A. Characterization of the respiratory chain of Leishmania donovani promastigotes. *Mol Biochem Parasitol* **75**, 43-53 (1995).
6. Rovis, L. & Baekkeskov, S. Sub-cellular fractionation of Trypanosoma brucei. Isolation and characterization of plasma membranes. *Parasitology* **80**, 507-524 (1980).

Nature Biotechnology: doi:10.1038/nbt.1556

7. Palmiter, R.D. Ovalbumin messenger ribonucleic acid translation. Comparable rates of polypeptide initiation and elongation on ovalbumin and globin messenger ribonucleic acid in a rabbit reticulocyte lysate. *J Biol Chem* **248**, 2095-2106 (1973).
8. Schultz, S.J. & Champoux, J.J. RNase H activity: structure, specificity, and function in reverse transcription. *Virus Res* **134**, 86-103 (2008).
9. Stuart, K.D., Weeks, R., Guilbride, L. & Myler, P.J. Molecular organization of Leishmania RNA virus 1. *Proc Natl Acad Sci U S A* **89**, 8596-8600 (1992).
10. Scheffter, S.M., Ro, Y.T., Chung, I.K. & Patterson, J.L. The complete sequence of Leishmania RNA virus LRV2-1, a virus of an Old World parasite strain. *Virology* **212**, 84-90 (1995).
11. Berestovskaya, N.G. et al. Cotranslational formation of active photoprotein obelin in a cell-free translation system: direct ultrahigh sensitive measure of the translation course. *Anal Biochem* **268**, 72-78 (1999).
12. Shaloiko, L.A. et al. Effective non-viral leader for cap-independent translation in a eukaryotic cell-free system. *Biotechnol.Bioeng.* **88**, 730-739 (2004).
13. Kozak, M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* **9**, 5134-5142 (1989).
14. Warrens, A.N., Jones, M.D. & Lechler, R.I. Splicing by overlap extension by PCR using asymmetric amplification: an improved technique for the generation of hybrid proteins of immunological interest. *Gene* **186**, 29-35 (1997).

Nature Biotechnology: doi:10.1038/nbt.1556