nature biotechnology

Species-independent translational leaders facilitate cell-free expression

Sergei Mureev[1], Oleksiy Kovtun[1], Uyen T T Nguyen[2] & Kirill Alexandrov[1,2]

Cell-free protein synthesis enables the rapid production and engineering of recombinant proteins. Existing cell-free systems differ substantially from each other with respect to efficiency, scalability and the ability to produce functional eukaryotic proteins. Here we describe species-independent translational sequences (SITS) that mediate efficient cell-free protein synthesis in multiple prokaryotic and eukaryotic systems, presumably through bypassing the early translation initiation factors. We use these leaders in combination with targeted suppression of the endogenous *Leishmania tarentolae* mRNAs to create a cell-free system based on this protozoan. The system can be directly programmed with unpurified PCR products, enabling rapid generation of large protein libraries and protein variants. *L. tarentolae* extract can produce up to 300 μg/ml of recombinant protein in 2 h. We further demonstrate that protein-protein and protein–small molecule interactions can be quantitatively analyzed directly in the translation mixtures using fluorescent (cross-) correlation spectroscopy.

Cell-free protein expression systems have been invaluable in studies of protein translation mechanisms, protein engineering and interactions, *in vitro* evolution and structural research[1,2]. The best characterized and most widely used cell-free translation systems are based on *Escherichia coli*, one of which (PURE) exclusively comprises recombinant components[3]. The main shortcomings of *E. coli*–based cell-free systems are the lack of chaperone-based mechanisms needed to fold complex eukaryotic proteins and the absence of post-translational modifications specific to eukaryotes. Eukaryotic cell-free systems based on wheat germ extract (WGE), rabbit reticulocyte lysate (RRL) and insect cell extract (ICE) overcome some of these problems[4–6]. Their ability to couple transcription and translation reactions on PCR-generated templates in multiplexed format greatly simplifies and accelerates protein production. This methodology has been used for protein engineering and production of protein arrays for many applications including drug screening and diagnostics (for review, see ref. 7). However, these eukaryotic cell-free systems do not scale up or down easily, are laborious and expensive to prepare, and are derived from source organisms that are difficult to manipulate genetically. Large batch-to-batch variation is also inevitable because of the long life cycle of the host organisms subject to many environmental influences. An alternative system based on a rapidly growing, fermentable organism that is amenable to technically straightforward genetic modifications is therefore highly desirable.

Three main obstacles must be overcome to create an efficient cell-free expression system. First, the system has to be programmed with exogenous mRNA that can efficiently engage the translational machinery. Second, endogenous mRNAs must be degraded or otherwise prevented from being translated. Finally, practical protocols must be developed for preparing cell-free extracts that retain a functional translational apparatus. The biggest hurdle of cell-free protein synthesis is reprogramming the system with exogenous mRNA. Although the problem is readily solved in an *E. coli*–based system by using natural or synthetic ribosome binding sites, it presents a substantial challenge in eukaryotic systems where assembly of the translational complex requires a 5′-capped mRNA and numerous translational factors[8,9]. As cell-free expression systems typically use RNA that is synthesized *in vitro* and is therefore uncapped, they fail to engage the cap-binding complex and early translation factors that are required to form the 43S preinitiation complex. The problem is partially overcome by enzymatic *in vitro* capping or by using cap-independent untranslated regions (UTRs), which mimic the cap structure and promote ribosomal complex assembly on the mRNA[10,11]. However, context dependence and species specificity of the latter sequences often complicate their generic application.

Undesired background translation is often prevented by digesting the endogenous mRNAs with micrococcal nuclease, which is more active toward mRNA than toward the ribonucleoprotein translational complexes. However, the inevitable partial degradation of tRNAs and rRNAs by such treatment reduces the translational activity of the lysate[12], making alternative strategies highly desirable.

We describe the design of species-independent, universal translation-initiation leaders that engage ribosomes directly and thereby bypass the cap-dependent pathway. We used these leaders to develop a eukaryotic cell-free system based on extracts of the protozoan *L. tarentolae*.

RESULTS

Development of a universal translation-initiating sequence

Ribosomes comprise an evolutionarily conserved catalytic core of translation machinery. This core has been extended in eukaryotes by a large array of translation initiation factors. RNA secondary

[1]Institute for Molecular Bioscience and Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, Queensland, Australia. [2]Max-Planck-Institute for Molecular Physiology, Dortmund, Germany. Correspondence should be addressed to K.A. (k.alexandrov@uq.edu.au).

Received 5 June; accepted 6 July; published online 2 August 2009; doi:10.1038/nbt.1556

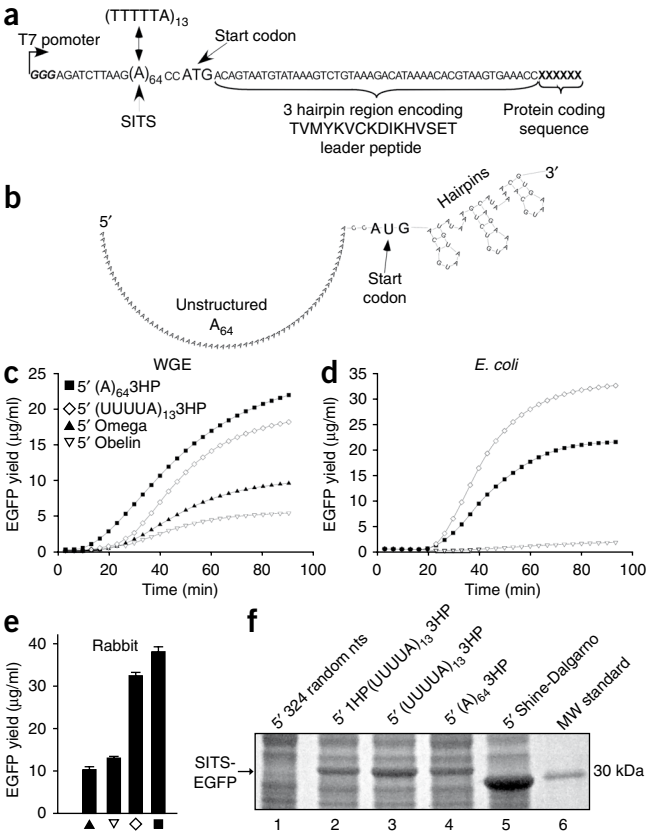

Figure 1 Design of unstructured species-independent translational sequences (SITS) and analysis of their ability to initiate translation in cell-free expression systems. (**a**) Structure of SITS with functional elements indicated by arrows. (**b**) Predicted secondary structure of $(A)_{64}$ 3HP-EGFP mRNA calculated with Mfold ($\Delta G = -3.2$ kcal/mol). Only the first 112 nucleotides are displayed. (**c**) Translation of EGFP-coding mRNAs carrying different 5′ translational leaders in WGE. The reactions were performed in microtiter plates and EGFP fluorescence was detected at 510 nm using a plate reader. (**d**) As in **c**, but using *E. coli* cell-free translation system. (**e**) As in **c**, but using RRL. Owing to high hemoglobin content of the lysate, EGFP fluorescence could not be analyzed directly. Therefore protein expression levels were measured by resolving the unboiled samples on SDS-PAGE gels followed by fluorescence scanning of the unstained gels. Bars are labeled with the same symbols as in **c**. (**f**) Coomassie-stained SDS-PAGE gel analysis of lysates of *E. coli* cells transformed with plasmids carrying indicated SITS sequences (lanes 2,3,4), the Shine-Dalgarno sequence (lane 5) or a randomly chosen sequence (lane 1). The 1HP of SITS in lane 2 denotes a single stable hairpin blocking the 5′ end (**Supplementary Fig. 1**). The identity of the expressed proteins was confirmed by fluorescent scanning of the gels and mass spectrometry of the purified proteins. The larger size of EGFP in lanes 2–4 reflects the inclusion of the SITS sequence into the final product as confirmed by mass spectrometry (**Supplementary Figs. 15** and **16**).

structure is believed to present a major obstacle to the assembly of the translation-initiation complex, and one of the major functions of the eukaryotic cap-binding complex is activation of eIF4A helicase, which untangles the 5′ terminus of mRNA[13]. This enables presentation of the mRNA to the 43S preinitiation complex, which comprises the ribosomal 40S subunit and initiation factors eIF3, eIF2, tRNAi, eIF1, eIF1A and eIF5 (refs. 14,15). Secondary structures within the 5′ UTR substantially reduce the rate of 48S preinitiation complex assembly and thus translation efficiency[16]. Several kinetic and biochemical studies support this idea by demonstrating rapid assembly of the preinitiation complex on templates devoid of secondary structure. Interestingly, the late poxvirus genes contain poly(A)-enriched sequences as their 5′ UTRs[17]. Translation of such mRNAs was proposed to be cap independent and was demonstrated to function efficiently in a plant-based cell-free system[18].

The presumed ability of unstructured translation leaders to bypass the early initiation steps and drive formation of the initiation complex may have several practical advantages. First, because they directly interact with the ribosomes, which are more conserved than the translation initiation factors, it is likely that such leaders display less species-dependence then sequences specifically binding initiation factors. Second, such a mechanism would be largely unaffected by the fused open reading frames (ORFs) because unstructured sequences have a low propensity for long-range base pairing. To test these ideas experimentally, we synthesized a series of mRNA molecules coding for *egfp* prefaced with poly(A) or mixed poly(U)/(A) sequences. These templates were assessed for their ability to mediate synthesis of EGFP in WGE. **Figure 1a** shows the most successful design, which includes a 64-residue poly(A) or $(TTTTTA)_{13}$ stretch followed by the start ATG and another 50-residue AT-rich sequence prefacing the *egfp* ORF. The latter sequence forms three hairpin (3HP) elements comprising three very short stems and thermodynamically preferred tetraloops (**Fig. 1a**,**b**).

This arrangement was designed to promote assembly of the preinitiation complex and to enable one-dimensional sliding of the small ribosomal subunit in search of the start codon. The 3HP structure allows melting of the region downstream of AUG upon advance of the initiation complex and increases the residence time of the preinitiation complex in the vicinity of the start codon. This is expected to facilitate AUG recognition by the small ribosomal subunit and subsequent recruitment of the large ribosomal subunit. This model is supported by the observation that the hairpin structure substantially enhanced the translation efficiency of the poly(A) leaders (**Supplementary Figs. 1** and **2**).

The mRNAs containing the unstructured leaders were translated efficiently in WGE with expression levels of ~30 μg/ml (**Fig. 1a**,**c**)—yields exceeding those achieved by using either an "omega" leader (one of the most efficient translation enhancers for WGE[19]) or a combination of the 5′ UTR derived from the obelin-encoding gene of *Obelia longissima* with the satellite tobacco necrosis virus–derived 3′ UTR (**Fig. 1c**)[11,20]. Remarkably, the same mRNA was efficiently translated in rabbit reticulocyte lysate and in *E. coli* extract (**Fig. 1d**,**e** and **Supplementary Fig. 3**). We speculated that unstructured leader sequences could potentially compete with endogenous mRNA and mediate mRNA translation *in vivo*. To test this directly, we transformed the *E. coli* BL21(DE3) RIL strain with plasmids bearing a T7 promoter and encoding EGFP prefaced either with the Shine-Dalgarno sequence or with $(A)_{64}$ 3HP or $(UUUUA)_{13}$ 3HP sequences. As can be seen in **Figure 1f**, $(UUUUA)_{13}$3HP and $(A)_{64}$3HP-containing plasmids mediated expression of EGFP in *E. coli*, but less efficiently than the Shine-Dalgarno sequence (**Fig. 1f**).

The observed species-independence is strong evidence of a cap-independent mechanism of translational complex assembly and implies that in principle, this new leader could initiate translation in any organism. We used two *Saccharomyces cerevisiae* extracts, one prepared by hypotonic lysis and one by nitrogen cavitation, to demonstrate translation of heterologous synthetic mRNA using this system (**Supplementary Fig. 4**). Further systematic analysis of translation efficiency of $(A)_{64}$3HP-EGFP constructs in coupled RRL, WGE and ICE systems demonstrated that this sequence mediated higher expression yields in all systems than species-specific leaders did (**Supplementary Fig. 3**). The developed leaders are SITS.

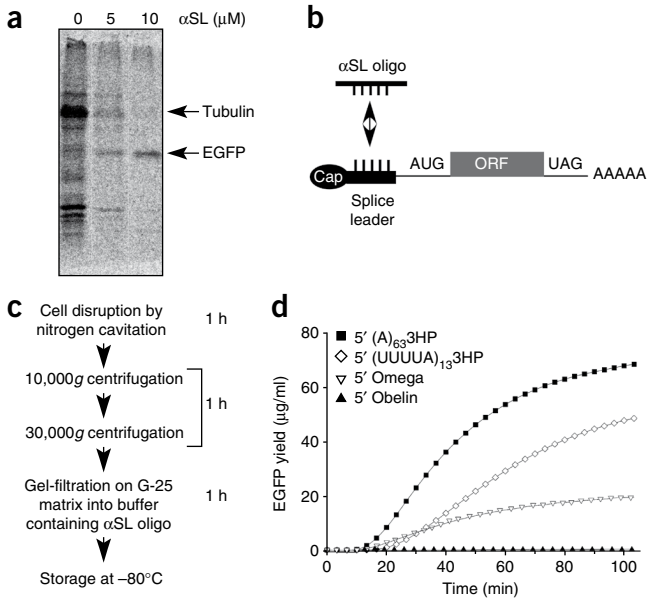

Figure 2 Preparation of cell-free translational lysate based on *L. tarentolae*. (**a**) Autoradiograph of SDS-PAGE gel loaded with the [^{14}C]leucine-supplemented LTE. Equal amounts of mRNA coding for the obelin-*egfp* gene were added to all reactions together with the indicated concentrations of α-splice leader oligonucleotide. The arrows indicate the positions of the dominant endogenous tubulin band and the recombinant EGFP. The identity of the latter was confirmed by western blot analysis. (**b**) Principle scheme of anti-splice leader (α-splice leader) oligonucleotide-mediated suppression of *L. tarentolae* mRNA translation in a cell-free system. (**c**) Flow chart of *L. tarentolae* cell-free lysate preparation. αSL, α-splice leader. (**d**) Translation of EGFP-coding mRNAs carrying different 5′ translational leaders in LTE. The reactions were performed and analyzed as in **Figure 1c**.

An *L. tarentolae*–based cell-free translation system

The availability of SITS expands the range of organisms that can be used for obtaining lysates for cell-free protein production. This enables development of better and cheaper cell-free systems based on rapidly growing eukaryotes that are easy to manipulate genetically but have never been considered for this purpose. We chose cell-free lysates from Trypanosomatidae, a large group of primitive and predominantly parasitic protozoa, to test this idea. Compared to other eukaryotes gene expression in Trypanosomatidae is a collection of exceptions and innovations that feature RNA editing, virtually no transcriptional control, polycystronic transcription and mRNA trans-splicing[21]. The mRNA splicing generates an identical 39-nucleotide splice-leader sequence on the 5′-end of all mature mRNAs, thus providing a universal and potentially targetable signature.

L. tarentolae has been used extensively as a Trypanosomatidae model. It grows rapidly in a range of media, is nonpathogenic to mammals and is easily manipulated genetically[22–24]. We tested several disruption procedures to obtain a translation-competent extract from cultured *L. tarentolae*. Nitrogen cavitation was identified as the best in terms of integrity of intracellular compartments and translation efficiency of the extracts (**Supplementary Figs. 5** and **6**). Radioautographs of SDS-PAGE gels used to resolve lysate incubated with [^{14}C]leucine revealed numerous radioactively labeled proteins, indicating active protein synthesis of endogenous mRNA (**Fig. 2a** and **Supplementary Fig. 7**). Further analysis also confirmed the system's ability to initiate translation of exogenously added mRNAs (**Supplementary Figs. 7** and **8**). Although translation of endogenous templates could be suppressed by predigestion with nucleases (**Supplementary Fig. 7**), we sought a less invasive and more robust suppression method, and therefore, we decided to exploit the presence of an identical 5′ splice leader in all protein-coding mRNAs of *L. tarentolae* (**Fig. 2b**). Treatment of *L. tarentolae* extract (LTE) with an anti-splice-leader oligonucleotide was expected to form high-affinity DNA:RNA heteroduplex clamp on the 5′ termini of endogenous translationally competent mRNAs that would selectively interfere with assembly of the translation initiation complex.

Addition of a 35-nucleotide anti-splice-leader oligonucleotide to LTE and WGE suppressed the translation of total *L. tarentolae* mRNA in a dose-dependent manner. Importantly, mRNA encoding EGFP, flanked with an obelin 5′ UTR, displayed increased translation under these conditions (**Fig. 2a** and **Supplementary Figs. 8** and **9**). The same effect could be achieved by adding the anti-splice-leader oligonucleotide at the last stage of lysate preparation, thus substantially simplifying the procedure (**Fig. 2c**). The entire process of extract preparation takes <4 h and yields ~2.5 ml of translation-competent lysate per liter of *L. tarentolae* culture. We tested the lysate for its ability to translate a set of EGFP-coding mRNAs carrying SITS, obelin or omega 5′ UTRs. Consistent with the observations on other eukaryotic cell-free systems, $(A)_{64}$3HP or $(UUUUA)_{13}$3HP translational leaders mediated the highest levels of protein synthesis in LTE (**Fig. 2d**).

A single-step, PCR-compatible, LTE-based *in vitro* translation protocol

The LTE cell-free system was then adapted for high-throughput automated applications, with the ultimate aim of becoming a one-tube template amplification and protein synthesis procedure. An LTE lysate was formulated, which contained recombinant T7 RNA polymerase, appropriate nucleotides and a buffer system capable of supporting both transcription and translation steps in a coupled fashion. When primed with SITS-EGFP-coding plasmid, the coupled system was ~30% less efficient than the mRNA-primed system. Unpurified PCR mixtures added to the LTE lysate were only 10% less efficient than plasmid DNA (**Fig. 3a**). This modified system is now well suited for rapid protein engineering and production in a high-throughput format.

We next compared the efficiency of the coupled LTE system with commercially available coupled *E. coli*, RRL, ICE and WGE systems by first generating a set of linear DNA templates encoding EGFP and carrying the optimal translational leaders. Translation using the corresponding systems demonstrated that, in the chosen configuration, LTE has several fold higher yields than commercial eukaryotic systems but was ~40% less efficient than an *E. coli*–based system (**Supplementary Fig. 3**).

Fluorescence readout of the translation assay allowed further optimization of the system. Temperature and buffer conditions for the LTE system were optimized using 384-well plate-based fluorescent assays. Using the optimal configuration, we could produce ~300 μg of protein/ml of lysate in 90 min using the optimal reaction temperature of 20 °C for RNA-driven systems and 26 °C for DNA-driven systems (**Supplementary Fig. 10**).

The ability of the LTE system to translate targets of diverse origins was next assessed by PCR-tagging a set of arbitrarily chosen genes combined with the EGFP-coding sequence. In all experiments, the fluorescent products could be detected directly in the gel (**Fig. 3b**). Furthermore, four independent RNAs were co-translated using this system to yield the four protein products in comparable amounts (**Fig. 3c**). Using a combination of LTE synthesis and affinity purification, we demonstrated the *in vitro* expression and assembly of the heterodimer Rab geranylgeranyl transferase (**Supplementary Fig. 11**).

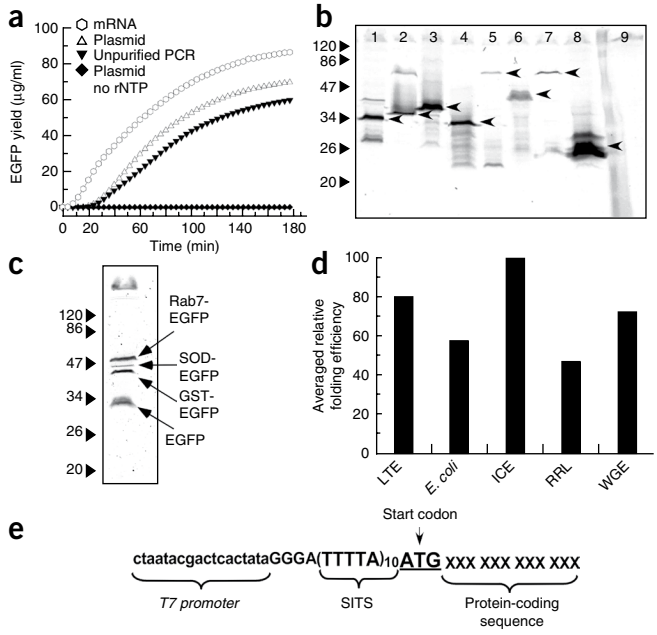

Figure 3 Development of coupled LTE system. (**a**) Comparison of expression levels in the LTE system primed with mRNA, plasmid DNA and PCR products coding for $(A)_{64}$3HP-EGFP. (**b**) Examples of genes that were PCR-tagged with *egfp* and translated in the LTE. The products were detected by fluorescence of unboiled samples in SDS-PAGE gels. Arrowheads indicate, by lane number: (1) GST-EGFP, (2) SOD-EGFP, (3) Rab7-EGFP, (4) EGFP-GST, (5) EGFP-Rab GGTase-α subunit, (6) EGFP-Rab GGTase-β subunit, (7) parvovirus surface protein VP1-EGFP fusion, (8) EGFP and (9) no mRNA. Smearing of the bands is due to formation of the higher order structures and partial unfolding of the unboiled EGFP-tagged proteins in the gel. (**c**) As in **b**, but with co-translation of the four *egfp*-tagged mRNAs. (**d**) Averaged relative folding efficiency of cell-free systems tested on the set of 15 human RabGTPases. The individual data set is shown in **Supplementary Figure 12** and is described in Online Methods. High apparent folding capacity of the ICE system may to an extent reflect the background translation of the endogenous Rab GTPases (**Supplementary Fig. 12p**). (**e**) Sequence of the PCR primer-encodable SITS.

To further ascertain that the LTE system could produce eukaryotic proteins in the folded state and compare it to other cell-free systems, we used SITS-tagged PCR products to translate 15 human Rab GTPases in coupled LTE, *E. coli*, ICE, RRL and WGE systems[25] (**Supplementary Fig. 12** and **Supplementary Methods**). The proteins were N-terminally fused to a Myc tag, which allowed their absolute quantification in translation mixtures. To assess the folding status of recombinant GTPases, we took advantage of their post-translational prenylation mediated by Rab geranygeranyl transferase. This modification can occur only when the GTPase domain of RabGTPase is properly folded and forms an extensive protein:protein interface with an accessory protein termed Rab escort protein (REP)[26]. To quantify the prenylation reaction, we took advantage of a biotinylated analog of RabGGTase's substrate (geranylgeranylpyrophosphate), biotin-geranyl pyrophosphate (BGPP). Incorporation of BGPP into protein can be readily detected by western blot analysis[27].

Comparative analysis of *in vitro* prenylation efficiency of the synthesized RabGTPases demonstrated that LTE and ICE systems were the most efficient for folding mammalian RabGTPases (**Fig. 3d** and **Supplementary Fig. 12**). Surprisingly, the mammalian RRL method performed relatively poorly in the test, even though possible interference of RRL components with the Rab prenylation machinery was ruled out by the appropriate controls. We observed that folding patterns of individual proteins significantly deviated from the statistical average. This indicates that similar to *in vivo* expression systems, screening of the protein target against cell-free systems should enable identification of the optimal expression system. The developed SITS enables such screening using a single-gene construct.

To simplify the procedure of template preparation for high-throughput applications, we sought to develop a smaller translational leader that could be encoded into PCR primers to enable translation of PCR products (**Fig. 3e**). To this end, we reduced the length of the UUUUA repeats to ten units and deleted the hairpin sequence. The modified design was based on the observation that a UUUUA-based SITS is less dependent on the post-AUG hairpin sequences than on the poly(A) sequence. The simplified leaders mediated protein synthesis from PCR-amplified templates with an efficiency ~65% of that observed for the full-length SITS (**Supplementary Fig. 2a**,**b**).

Analysis of LTE-translated proteins by fluorescence correlation spectroscopy

One of the main advantages of cell-free protein production is its speed and amenability to miniaturization and multiplexing. Analysis of protein:protein and protein:small molecule interactions requires the isolation—and in many cases—labeling of the recombinant proteins. Although pull-down–based interaction analysis can in principle be used for high-throughput applications, it provides no quantitative information about the interaction and is not easily expanded to analyze protein:small-molecule interactions. These limitations can be overcome by monitoring interactions of the molecules directly in solution using spectroscopic methods. Fluorescent correlation and cross-correlation spectroscopy (FCS and FCCS) have emerged as promising techniques for the analysis of molecular interactions in complex mixtures such as cell lysates or even in living cells[28].

These single molecule–based methods provide information on diffusion rates, binding affinities and interaction kinetics for fluorescent molecules in solution[29]. The availability of genetically encoded fluorophores enables application of these methods to monitor protein interactions directly in cell-free expression systems, thus compressing the entire cycle of recombinant protein production, derivatization and functional analysis into a single experiment. We initially sought to test whether the developed cell-free system is compatible with FCS measurements. Coupled *in vitro* translation of the SITS-*egfp* gene was undertaken in the chamber of a confocal microscope configured in the FCS mode. Fluorescent molecules were detected within 5 min and the optimal measuring density was reached by 15 min. The average autocorrelation curve for EGFP was compared with purified EGFP in solution. The curves demonstrated low noise but also exhibited a slightly anomalous behavior, possibly because of molecular crowding in the LTE (**Fig. 4a**,**b**). Incubations longer than 20 min further increased the signal and led to a reduction in autocorrelation due to the simultaneous presence of multiple EGFP molecules in the focal volume. Translation of EGFP fused with the β-subunit of Rab geranylgeranyl transferase yielded a protein with a lower diffusion coefficient, reflecting its larger size (**Fig. 4b**). This result demonstrated that single molecule analysis could be performed directly in the cell-free translation mixture.

Analysis of LTE-synthesized protein interactions by FCCS

Having shown that diffusion rates of LTE-synthesized proteins can be analyzed directly in the translation mixture, we next asked whether the same direct analysis is possible for protein:small molecule and protein:protein interactions. The well-characterized interaction

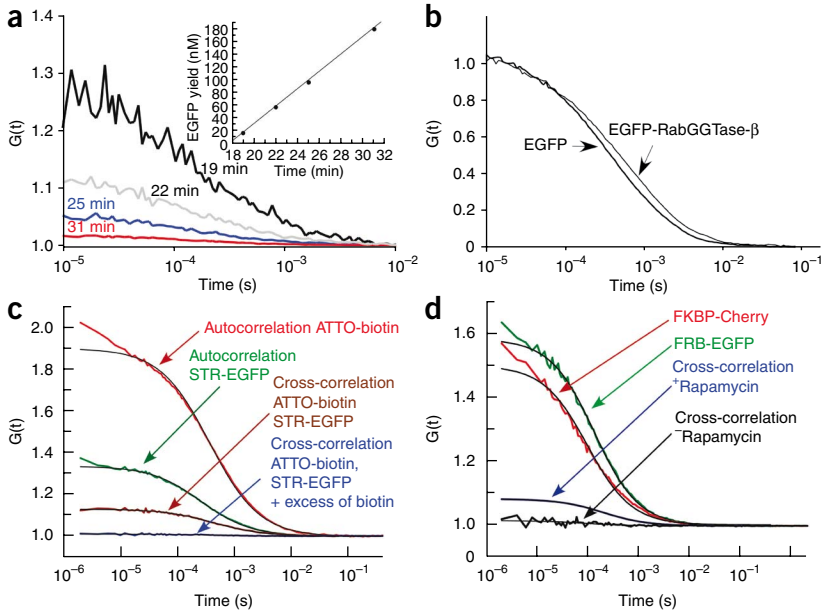

Figure 4 Fluorescence correlation (FCS) and cross-correlation (FCCS) analysis of fluorescent proteins synthesized using the LTE system. (**a**) FCS analysis of EGFP molecules synthesized in the coupled LTE system primed with the SITS-EGFP-coding plasmid. The data were collected at the indicated times and fitted to the autocorrelation function. The inset shows the changes in EGFP concentration during the measurement time. (**b**) Fit of the diffusion rates of unpurified EGFP and EGFP-RabGGTase β-subunit expressed using the LTE system. (**c**) FCS analysis of streptavidin-EGFP fusion protein expressed in LTE and FCCS analysis with the ATTO 590-biotin conjugate. The control samples (blue curve) were supplemented with excess unlabeled biotin. The individual curves were fitted to the autocorrelation function (black line). (**d**) Interaction analysis of the LTE-expressed FRB-EGFP and FKBP-Cherry in the presence (blue curve) and absence (black curve) of rapamycin.

between streptavidin and biotin was chosen as the first test example. A fusion protein of streptavidin and EGFP was synthesized in LTE, and its ability to interact with the fluorescent ATTO 590-biotin was analyzed by FCCS (**Supplementary Methods**). This analysis successfully demonstrated cross-correlation between channels, indicating complex formation between the streptavidin-EGFP and ATTO 590-biotin (**Fig. 4c**). The observed interaction was specific as addition of excess unlabeled biotin abrogated the signal owing to competition for binding to streptavidin. Independent proof that the LTE-synthesized streptavidin was correctly folded was obtained by gel co-migration experiments that confirmed its tetramerization and interaction with ATTO 590-biotin (**Supplementary Fig. 13**).

Finally, we examined whether more complex systems synthesized in LTE could be analyzed by FCCS. We chose to analyze FKBP and FRB proteins, which form a high-affinity complex in the presence of the small-molecule rapamycin[30]. EGFP-FRB was translated in the LTE cell-free system in the presence of recombinant FKBP-Cherry that was purified by nickel-nitrilotriacetic acid (Ni-NTA) chromatography from a separate translation reaction (**Supplementary Fig. 14**). Although co-translation of both proteins in one setup is also possible, the slow maturation of Cherry compared with EGFP complicates the direct interaction analysis. FCCS analysis of the translation mixture revealed cross-correlation of the fluorescent signals that could be fitted to a two-component binding model with a K_d value of 176 nM. The interaction was specific because no signal was present in the absence of rapamycin (**Fig. 4d**). The obtained K_d value was higher than the K_d value of 32 nM obtained with purified recombinant proteins (data not shown) and the previously reported affinity estimates[30]. The apparent discrepancy of the obtained values prompted us to repeat the measurement of the purified recombinant complex in the context of the LTE lysate. This yielded a K_d value of 90 nM, indicating that the observed discrepancy reflected interaction(s) of the complex components with the LTE extract and that the method faithfully determined the affinity changes.

DISCUSSION

Our study offers solutions to four main bottlenecks of cell-free protein production and analysis: initiating translation, suppressing translation of endogenous mRNAs, rapid preparation of eukaryotic translation lysates and interaction analysis of the synthesized proteins. SITS promote assembly of the active ribosome:mRNA complexes in both pro- and eukaryotic cell-free systems, apparently without the requirement for translation initiation factors. This is evidenced by the observation that SITS-containing mRNA can initiate translation on *E. coli* ribosomes both *in vivo* and *in vitro*. The SITS-mediated initiation of translation is efficient in all cell-free systems tested and more efficient in the eukaryotic systems than the established translation-initiation leaders. Availability of such universal translation leaders enables rapid development of cell-free expression systems based on organisms previously not considered suitable as sources of *in vitro* translation lysates.

We exploited SITS to develop a cell-free system using a *L. tarentolae*-based *in vitro* translation system. Amenability to large-scale and inexpensive fermentation of *L. tarentolae* enables the routine production of ~50 ml of cell-free extract in a standard laboratory setup. The development of methods for *L. tarentolae* genetic manipulation enabled further improvement of the host strains, including the overexpression of chaperone systems, kinases and other enzymes to mediate post-translational modifications. The potential for such modifications represent a major advantage of *L. tarentolae*–based systems over other eukaryotic systems based on complex multicellular organisms that are less amenable to genetic modification.

The ability of the developed cell-free system to co-express multiple proteins also prompted us to seek a generic solution for protein:protein and protein:small molecule interaction analysis. We demonstrated the use of FCCS to quantitatively and directly analyze interactions between the *in vitro*–translated proteins and small molecules in the translation mixture. Development of a single-tube, coupled PCR-to-protein synthesis protocol allows high-throughput protein engineering and binding studies in 384-well or higher-density formats. A very appealing application of such a technology is the high-throughput engineering of multi-protein complexes, which has until now not been possible due to exorbitant numbers of truncation variants and their combinations.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This work was supported in part by a Deutsche Forschungsgemeinschaft grant AL 484/5-4 and Heisenberg fellowship to K.A. U.T.T.N was supported by the predoctoral fellowship of Fonds der chemischen Industrie. We thank P. Lommerse for help with FCCS measurements and to V. Rajagopalan, M. Terbeck and A. Sander for excellent technical assistance. We are grateful to M. Gruen, R. Breitling, F. Seebeck and A. Goodall for critical reading of the manuscript and stimulating discussions and K. Browning (University of Texas at Austin) for a gift of reagents. We are very grateful to C. Clayton, S. Beverley, L. Shaloiko, I. Granowsky and P. Dittrich for support in the early stages of the project. We are indebted to R.S. Goody for long-term support.

AUTHOR CONTRIBUTIONS

S.M. designed and performed research, O.K. designed and performed research, U.T.T.N designed and performed research and K.A. designed research and wrote the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Nirenberg, M.W. & Matthaei, J.H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. USA* **47**, 1588–1602 (1961).
2. Katzen, F., Chang, G. & Kudlicki, W. The past, present and future of cell-free protein synthesis. *Trends Biotechnol.* **23**, 150–156 (2005).
3. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).
4. Sawasaki, T., Morishita, R., Gouda, M.D. & Endo, Y. Methods for high-throughput materialization of genetic information based on wheat germ cell-free expression system. *Methods Mol. Biol.* **375**, 95–106 (2007).
5. McCallum, C.D., Do, H., Johnson, A.E. & Frydman, J. The interaction of the chaperonin tailless complex polypeptide 1 (TCP1) ring complex (TRiC) with ribosome-bound nascent chains examined using photo-cross-linking. *J. Cell Biol.* **149**, 591–602 (2000).
6. Ezure, T. *et al.* Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol. Prog.* **22**, 1570–1577 (2006).
7. He, M. & Wang, M.W. Arraying proteins by cell-free synthesis. *Biomol. Eng.* **24**, 375–380 (2007).
8. Pestova, T.V. *et al.* Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. USA* **98**, 7029–7036 (2001).
9. Merrick, W.C. Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332**, 1–11 (2004).
10. Lomakin, I.B., Hellen, C.U. & Pestova, T.V. Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Mol. Cell. Biol.* **20**, 6019–6029 (2000).
11. 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).
12. Pelham, H.R. & Jackson, R.J. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* **67**, 247–256 (1976).
13. Korneeva, N.L., First, E.A., Benoit, C.A. & Rhoads, R.E. Interaction between the NH2-terminal domain of eIF4A and the central domain of eIF4G modulates RNA-stimulated ATPase activity. *J. Biol. Chem.* **280**, 1872–1881 (2005).
14. Kozak, M. The scanning model for translation: an update. *J. Cell Biol.* **108**, 229–241 (1989).
15. Kozak, M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867–19870 (1991).
16. Pestova, T.V. & Kolupaeva, V.G. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev.* **16**, 2906–2922 (2002).
17. Ahn, B.Y. & Moss, B. Capped poly(A) leaders of variable lengths at the 5′ ends of vaccinia virus late mRNAs. *J. Virol.* **63**, 226–232 (1989).
18. Bablanian, R., Goswami, S.K., Esteban, M., Banerjee, A.K. & Merrick, W.C. Mechanism of selective translation of vaccinia virus mRNAs: differential role of poly(A) and initiation factors in the translation of viral and cellular mRNAs. *J. Virol.* **65**, 4449–4460 (1991).
19. Kamura, N., Sawasaki, T., Kasahara, Y., Takai, K. & Endo, Y. Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. *Bioorg. Med. Chem. Lett.* **15**, 5402–5406 (2005).
20. 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).
21. Clayton, C. & Shapira, M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* **156**, 93–101 (2007).
22. Breitling, R. *et al.* Non-pathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr. Purif.* **25**, 209–218 (2002).
23. Niculae, A. *et al.* Isotopic labeling of recombinant proteins expressed in the protozoan host *Leishmania tarentolae*. *Protein Expr. Purif.* **48**, 167–172 (2006).
24. Fritsche, C., Sitz, M., Wolf, M. & Pohl, H.D. Development of a defined medium for heterologous expression in *Leishmania tarentolae*. *J. Basic Microbiol.* **48**, 488–495 (2008).
25. Leung, K.F., Baron, R. & Seabra, M.C. Thematic review series: lipid posttranslational modifications. geranylgeranylation of Rab GTPases. *J. Lipid Res.* **47**, 467–475 (2006).
26. Rak, A. *et al.* Structure of the Rab7:REP-1 complex: insights into the mechanism of Rab prenylation and choroideremia disease. *Cell* **117**, 749–760 (2004).
27. Nguyen, U.T. *et al.* Exploiting the substrate tolerance of farnesyltransferase for site-selective protein derivatization. *ChemBioChem* **8**, 408–423 (2007).
28. Bacia, K. & Schwille, P. Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. *Nat. Protoc.* **2**, 2842–2856 (2007).
29. Oyama, R. *et al.* Protein-protein interaction analysis by C-terminally specific fluorescence labeling and fluorescence cross-correlation spectroscopy. *Nucleic Acids Res.* **34**, e102 (2006).
30. Banaszynski, L.A., Liu, C.W. & Wandless, T.J. Characterization of the FKBP. rapamycin.FRB ternary complex. *J. Am. Chem. Soc.* **127**, 4715–4721 (2005).

ONLINE METHODS

Preparation of LTE. *L. tarentolae* Parrot laboratory strain was obtained from Jena Bioscience GmbH and was cultivated as agitated suspension in LEXSY BHI medium (Jena Bioscience) at 26 °C. Cells were grown to a final OD_{600} of 1.9. Cells were pelleted for 10 min at 2.500*g* in Beckman Avanti TM centrifuge using J-20 XP Rotor. The pellet was washed once by with sucrose extraction buffer (SEB) containing 250 mM sucrose, 45 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 3 mM magnesium acetate. The cell suspension at 5×10^9 cells/ml was transferred into cavitation chamber of cell disruption bomb (Parr Instruments) and incubated 30 min on ice under 70 bar nitrogen pressure. After the instant release of the pressure the suspension of the broken cells was subjected to two rounds of differential centrifugation. First centrifugation was carried out at 10,000*g* for 15 min and the supernatant was transferred to a fresh tube and centrifuged at 30,000*g* for 15 min To obtain post-microsomal supernatant (30S extract) the upper ¾ portion of the supernatant was loaded onto the NAP-25 gel-filtration columns (GE Healthcare) pre-equilibrated with SEB buffer lacking sucrose (EB). Protein concentration in fractions was determined with Coomassie Protein Assay kit (Pierce), the fractions with $OD_{595} \geq 0.7$ were pooled together, supplemented with anti–splice leader oligonucleotide to final concentration of 20 μM and snap frozen in liquid nitrogen. The extracts were kept at −80 °C for 6 months without the detectable loss of activity.

***In vitro* transcription and purification of mRNA.** *In vitro* transcription reactions were carried out in buffer containing 40 mM HEPES pH 7.9, 18 mM magnesium acetate, 2 mM spermidine, 5 mM of each ribonucleotide, 40 mM DTT, 10 unit/μl of T7 RNA polymerase (Fermentas), 2.5 unit/ml of yeast inorganic pyrophosphatase (NEB), 0.3 unit/μl of RNase Inhibitor (Clontech) and 50–70 ng/μl of template DNA. The mixture was incubated for 90 min at 37 °C, diluted threefold with RNase-free water, and the RNA was precipitated by addition of 0.5 volume of 8 M LiCl followed by 30 min incubation on ice and subsequent centrifugation. The pellet was dissolved in 300 μl of water and precipitated again by addition of 2.5 volumes of absolute ethanol in the presence of 0.3 M NH_4OAc at pH 5.2. The pellet was rinsed twice with 70% ethanol and dissolved in an appropriate volume of distilled water. RNA concentration and purity was measured by UV-spectroscopy using Nanodrop (Thermo). The integrity of RNA was verified by denaturing agarose gel.

Preparation of translation extract from *S. cerevisiae* and translation of $(A)_{64}$3HP-EGFP coding mRNA. Yeast cells were grown in 1 liter of yeast peptone dextrose medium to a density of 5×10^7 cells/ml, pelleted by centrifugation at 2,500*g* for 10 min and washed with 0.5 liter of PBS. 15 g of cell pellet was resuspended in 30 ml of buffer containing Tris pH 9.4, and 10 mM DTT and incubated for 15 min at 30 °C with 100 r.p.m. agitation. The suspension was then centrifuged and resuspended in 75 ml of buffer containing 20 mM KP_i, pH 7.4, 1.2 M sorbitol and 30 mg zymolyase and incubated for 45 min with shaking as above. The resulting spheroplasts were washed in 500 ml of buffer containing 100 mM potassium acetate, 2 mM magnesium acetate, 30 mM HEPES pH 7.6 and 1.2 M sorbitol. Half of the pellet was washed once in cavitation buffer (8.5% mannitol, 30 mM HEPES pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM DTT) and resuspended in 14 ml of the same buffer to the density of 5×10^9 cells/ml, incubated on ice in nitrogen cavitation reservoir (Parr Systems) at 70 bar pressure. After the 25 min of incubation the pressure was released. Another half of the pellet was resuspended in 14 ml buffer containing 20 mM HEPES-KOH pH 7.6 and 2 mM DTT and subjected to hypotonic lysis on ice for 15 min. Both lysates were cleared by centrifugation and rebuffered as described for LTE preparation. The gel-filtration columns were equilibrated with cavitation buffer without mannitol. The final extracts were treated with micrococcal nuclease at a concentration of 10 units/ml for 20 min at 20 °C and snap frozen in liquid nitrogen.

***In vitro* translation in LTE system. Translation of mRNA.** Translation reactions in LTE were carried out at 20–26 °C using 600 pmol/ml of mRNA. To simplify translation protocol all additives were formulated as 5× feeding master mix and added to the translation reaction mix before translation reaction. The final *in vitro* translation reaction contained 50% of LTE in 40 mM HEPES, pH 7.6, 50 mM potassium acetate, 0.25 mM spermidine, 2 mM DTT, 2.5 mM magnesium acetate, 1% wt/wt PEG 3000, protease cocktail (Roche), 133 μM of each of 20 amino acids as well as the energy regeneration system composed of 1.2 mM ATP, 0.12 mM GTP and 40 unit/ml skeletal muscle creatine kinase and 40 mM of creatine phosphate (EMD Biosciences). The translation reaction was allowed to proceed for 2 h at 20 °C. Typically a 10 μl aliquot was withdrawn, mixed with 2× SDS loading buffer and resolved on the 15% SDS-PAGE gel. In case of proteins tagged with fluorescent domain the samples were not boiled before electrophoresis, and the fluorescent products were visualized by EGFP fluorescence using VersaDoc gel documentation system (BioRad). Fluorimetric kinetic assays of EGFP expression were performed in 384-well plates in 20 μl reaction volume at 20 °C during 2 h with one measurement every 2 min. The measurements were carried out using the Fluoroscan Ascent FL fluorescent plate reader (Labsystems). The excitation/emission wavelengths were set to 485/510 nm for EGFP.

To prepare a coupled transcription-translation system the LTE extract was additionally supplemented with 2 unit/μl T7 polymerase and rNTP mixture to a final concentration of 1 mM each. The transcription-translation reactions were programmed either by SITS containing plasmids, purified PCR products at 0.05 μM or 10 μl of 20% wt/wt of a crude PCR reaction.

***In vivo* expression, purification and ESI-MS analysis of EGFP expressed in *E. coli* using SITS leader.** *E. coli* BL 21(DE3) strain was transformed with pTUB-NEO vector containing EGFP-coding ORF with $(A)_{64}$3HP leader under control of T7 promoter. The transformed cells were grown in 500 ml of LB media at 37 °C to OD_{600} = 0.6 and the protein synthesis was induced with 0.5 mM of IPTG. The cells were grown for an additional 2 h and harvested by centrifugation for 15 min at 8,000*g*. Harvested bacterial cells were resuspended in 3 ml of ice-cold buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA and 0.1 mg/ml of DNase I) and disrupted by sonication. The obtained crude cell extract was cleared by centrifugation for 15 min at 30,000*g* followed by ethanol extraction of EGFP according to the previously described protocol[32]. Next 0.4 ml of EGFP-containing ethanol fraction was subjected to gel-filtration on a Superdex 75 10/300 column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 8.0, 50 mM NaCl buffer. Fractions of 0.2 ml were collected and analyzed by SDS-PAGE followed by Coomassie Blue staining. Selected EGFP-containing fractions were pooled, and total protein from 100 μl aliquot was precipitated with acetone. Dried acetone pellet was dissolved in 5% formic acid and subjected to ESI-MS analysis on Q-STAR Elite mass-spectrometer (Applied Biosystems) fitted with a C-18 reversed-phase column driven by 1100 series nano-HPLC system (Agilent). Nano-HPLC reversed-phase chromatography was performed under 0–100% gradient of solvent B (90% acetonitrile in water, 0.1% formic acid) in solvent A (0.1% formic acid in water) with a flow rate of 4 μl/min.

Translation of RabGTPases in *E. coli*, WGE and LTE extracts. Rab GTPases 1, 3A, 6A, 8, 10, 14, 15, 16, 19, 23, 26, 26, 28, 30 and 38 were PCR amplified using the splicing overlap extension (SOE) PCR procedure, but the SITS sequence–Myc assembly was amplified and the Myc sequence used as a fusion junction. The procedure resulted in linear DNA fragments bearing sequences encoding for T7 promoter, $(A)_{64}$2HP, myc and Rab ORF. The oligonucleotides annealing 150 nt downstream from the Rab stop-codon triplet were used as the reverse primers. Purification of PCR products was carried out as described above. This resulted in insertion of the sequence: AGATCTTAAGAA AAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCATGACAGTAATGTA TGTCTGTAAAGACAAAACCGAGCAGAAGCTGATCTCGGAGGAGGACCTG between T7 promoter and the first codon of Rab ORFs. The respective PCR products were added to the final concentration of 40 nM to 20 μl of LTE, *E. coli* RTS (Roche), TnT-ICE (Promega), TnT-RRL (Promega), TnT-WGE (Promega) and coupled cell-free systems. The translation reactions were carried out at 26 °C for LTE and at 28 °C for *E. coli*, TnT-ICE, TnT-RRL and TnT-WGE. The total product yield was estimated by western blot analysis with the mouse anti-Myc antibodies and secondary antibodies conjugated to HRP using Myc-tagged meromyosin light chain as standard. The antibody staining was developed using ECL chemiluminescent detection system (GE Healthcare) and luminescence was imaged using Lumi-Imager FM (Boehringer Mannheim). The densitometry was performed with Image J software package.

***In vitro* prenylation of the translated RabGTPases.** *In vitro* prenylation of the translated RabGTPases was carried out to assess the folding status of the

recombinant GTPases. To exclude incorporation of endogenous isoprenoids present in eukaryotic cell–free lysates by endogenous protein prenyltransferases, we supplemented duplicates of the translation reactions with BGPP/REP1 mixture during the translation reaction. After 25 min incubation the reactions were supplemented with RabGGTase and incubated for a further 2.5 h. The reactions that were not supplemented with prenylation machinery have been subjected for prenylation upon completion of cell-free synthesis and incubated for 3 h at 25 °C[27]. The reactions were stopped by addition of an equal volume of the SDS-sample buffer and 2 µl aliquots were used for western blot analysis as described above. To detect incorporation of biotin-geranyl into RabGTPases the membranes were probed with streptavidin-HRP and the staining was visualized as described above. In control experiments the unprimed cell-free lysates were subjected to prenylation and analyzed by western blot analysis as described above. The density of the bands on the scanned luminescent images was quantified using program Image J. The densities of the incorporated biotin label were adjusted using loading correction factors derived from protein expression level quantification obtained with anti-Myc antibody. The data were further corrected for incorporation of biotin-geranyl into endogenous RabGTPases and plotted using Grafit 5.0 (Erithacus Software). To obtain the averaged prenylation efficiency the data for individual Rabs were scaled to 100% scale and the average for each cell-free system was calculated.

***In vitro* co-translation and subunit interaction analysis of Rab geranylgeranyl transferase (RabGGTase).** Genes coding for EGFP-GGTase β and 6His-mCherry-GGTase α-subunits with $(A)_{27}$ leader and T7 promoter were synthesized using SOE PCR and used as templates for *in vitro* transcription. Both mRNAs were co-translated in 30 µl of LTE cell-free system. As a negative control mRNA coding for EGFP-GGTase β subunit was translated under the same conditions. The translation reaction was carried out as described above but the concentration of each mRNA was set at 0.3 pmol/µl and β-mercaptoethanol in the translation reaction mix was substituted for DTT. The formation of α/β heterodimer was analyzed by affinity chromatography using the 6His tag on the N terminus of the α-subunit. For that, the translation reactions were mixed with 100 µl of binding buffer (300 mM NaCl, 50 mM HEPES, pH 7.6, 0.01% Triton X-100) and 30 µl of Ni-NTA resin. The slurry was incubated for 10 min at 22 °C, loaded into a spin-column and centrifuged at 500*g* for 1 min. The resin was further washed twice with 250 µl of washing buffer, and the bound proteins were eluted with 30 µl of binding buffer supplemented with 300 mM of imidazole. 15 µl aliquots of second wash and the eluate were resolved unboiled on 15% SDS-PAGE. The polypeptides were visualized by multi-wavelength scanning for mCherry and EGFP fluorescence using the VersaDoc fluorescent scanner (BioRad).

Expression and purification of recombinant FRB-EGFP and FKBP-Cherry fusion proteins. ORFs of FRB and FKBP were PCR amplified using the synthetic genes for FRB and FKBP as templates. The PCR fragments were cloned in-frame into mCherry or EGFP containing pOPINE vectors yielding plasmids 6His-EGFP-FRB and 6His-mCherry-FKBP[33]. *E. coli* transformation, cultivation, protein synthesis induction and harvesting were done as described elsewhere[33].

Harvested cells were resuspended in buffer A (40 mM HEPES KOH, pH 7.6, 300 mM NaCl with 1 mM PMSF, 2 mM β-mercaptoethanol and protease inhibitor cocktail; Roche) before lysis and disrupted using fluidizer. The crude extract was cleared by centrifugation for 1.5 h at 50,000*g*. The supernatant was loaded on Hi-Trap Ni-NTA chelating column (GE Healthcare), and the bound proteins were eluted with a linear gradient of 10–150 mM imidazole in buffer A. Fractions containing the protein of interest were pooled and dialyzed against 25 mM HEPES 7.6, 40 mM NaCl and 5 mM β-mercaptoethanol. Dialyzed protein solution was cleared by centrifugation and concentrated to 10 ml with Centriprep-30 (Amicon) concentrator. The solution was cleared again by centrifugation and subjected to further purification by size exclusion chromatography on Superdex-200 column. The protein eluting at the expected position, was concentrated and snap-frozen in liquid nitrogen.

FCS and FCCS analysis. All single-molecule measurements were carried out on the inverted LSM 510 confocal microscope with ConfoCor2 module (Zeiss) operated by Zeiss software. EGFP fluorescence was excited by 488 nm line of an argon laser beam; ATTO 590, by 594 nm HeNe laser; and mCherry, by 561 nm diode laser. The excitation beams were focused in 20 µl of the analyzed sample placed in the well of an 8-well chamber slide (Nunc) using water immersion Apochromat X40/1.2NA objective. Emitted light passed through an adjustable pinhole and was separated on a 565 nm dichroic mirror; 505–540 and 615–680 nm segments of the spectra were then selected with bypass filters for EGFP (green channel) and mCherry or ATTO 590 (red channel) fluorescence detection, respectively.

The power of exciting beams was set up to achieve maximum count per molecule (CPM) rate without visible photo-bleaching, and typically, CPM was kept in a range between 3 and 6 kHz. Data acquisition time was 100 s and included 10 runs of 10 s each. The confocal volumes were estimated by measuring rhodamine green (green channel) and Alexa 546 (red channel), whose diffusion coefficients are known to be 230 and 260 $\mu m^2 s^{-1}$ respectively, in the buffer containing 10 mM TrisHCl, pH 7.6, and 0.1 mM EDTA. The overlap of green/red channel confocal volumes was assumed to be 50%.

The collected data was analyzed using Zeiss software package. The theoretical basics of autocorrelation and cross-correlation analysis are described everywhere[34]. Experimental autocorrelation functions (ACF) were fitted in three-dimensional diffusion model[35] for a two-components model. The dissociation constants (K_d) for binary interacting systems were calculated as described previously[29]. FKBP:FRB interaction was treated as a simple binary interaction owing to the high affinity of rapamycin:protein interactions and the excess of antibiotic over the proteins.

FCCS interaction analysis of streptavidin-EGFP and ATTO 590-biotin. PCR, T7 transcription and cell-free translation procedures were carried out as described above. The gene, encoding T7 promoter, $(UUUUA)_7$ leader sequence, streptavidin-EGFP fusion protein and 94-nt long 3′ UTR was synthesized by SOE PCR and used for *in vitro* transcription. The mRNA obtained was purified and used to program the LTE as described above. After 2 h of incubation at 20 °C, the translation mixture was clarified by centrifugation at 18,000*g* for 10 min and diluted with buffer (100 mM NaCl, 40 mM HEPES KOH, pH 7.6) to adjust the concentration of the synthesized protein to ~70 nM. An aliquot of reaction mixture was analyzed on 15% SDS-PAGE gel in the presence or absence of labeled biotin to verify the ability of the synthesized streptavidin-EGFP fusion protein to tetramerization and biotin binding. ATTO 590-biotin was added directly to the translation mixture to a final concentration of 40 nM and the sample was incubated for 3 min at 20 °C and used for FCCS measurements. To confirm the specificity of the interaction between streptavidin-EGFP and ATTO 590-biotin, the experiment was repeated in the presence of 1 µM of unlabeled biotin.

FCCS interaction analysis of FRB and FKBP synthesized in cell-free system. To analyze the interaction of FRB and FKBP, we used SOE PCR to generate four variants of each gene fused either to mCherry or EGFP and containing either C-terminal 6his tag or no tag. The PCR fragments were designed to contain the T7 promoter and 5′ $(UUUUA)_7$ leader sequence. Respective mRNAs were prepared and translated at 20 °C for 3 h as described above. To evaluate the quality of the synthesized proteins the translation mixtures were resolved on 15% SDS-PAGE gel and the fluorescent products were detected by fluorescence imaging as described above. Concentrations of the fluorescently labeled FRB and FKBP in translation mixtures or in elution fraction from Ni-NTA agarose resin (in the case of hexahistidine-tagged purified proteins) was estimated from fluorescence autocorrelation spectroscopy data. The protein complex formation was analyzed using 50–80 nM concentration of each partner proteins and 2 mM of rapamycin (Invitrogen). For dilution the buffer containing 40 mM HEPES, pH 7.6, and 100 mM NaCl was used.

31. 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).
32. Yakhnin, A.V., Vinokurov, L.M., Surin, A.K. & Alakhov, Y.B. Green fluorescent protein purification by organic extraction. *Protein Expr. Purif.* **14**, 382–386 (1998).
33. Berrow, N.S. *et al.* A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* **35**, e45 (2007).
34. Keese, M. *et al.* Quantitative imaging of apoptosis commitment in colorectal tumor cells. *Differentiation* **75**, 809–818 (2007).
35. Rudolf, R. & Ulo, M. Diffusion of single molecules through a Gaussian laser beam in *Proc. SPIE*, **1921** (ed. Jouko, E.K.-T.) 239–248 (1993).

doi:10.1038/nbt.1556

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

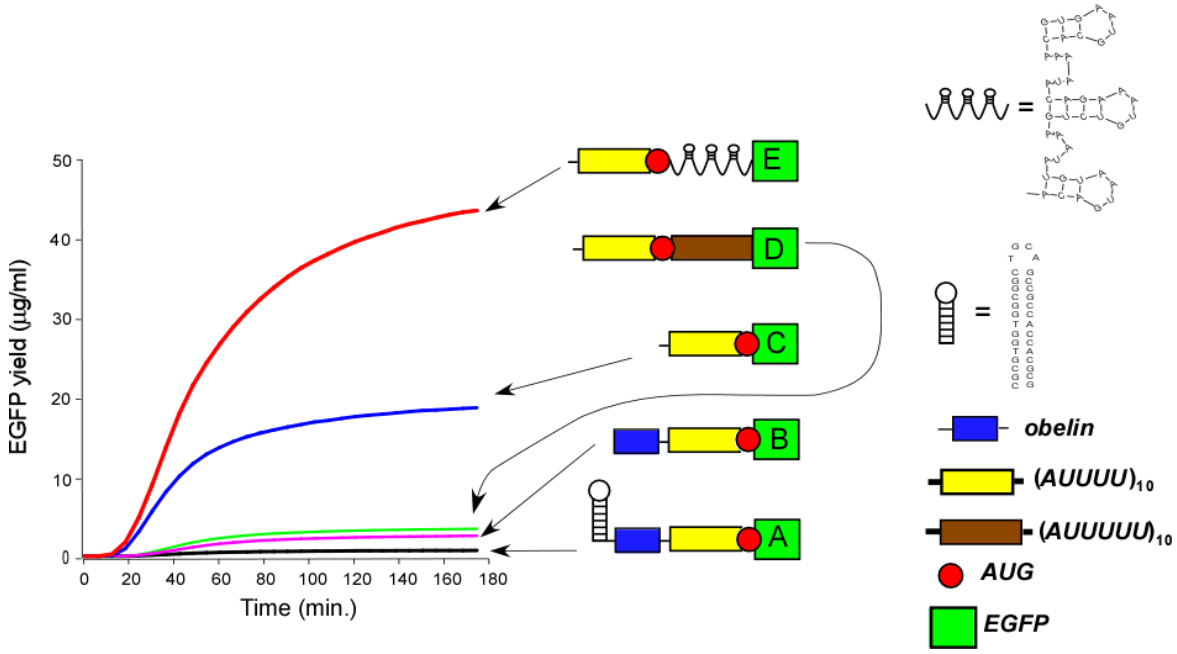

Figure S1. Translation of mRNAs coding for *egfp* gene with polyAU SITS in LTE. Equal amounts (0.6 pmol/μl) of indicated mRNAs were translated in LTE and changes in fluorescence were recorded using fluorescence plate reader. The sequence elements comprising the translated mRNAs are denoted in the inset.

Nature Biotechnology: doi:10.1038/nbt.1556

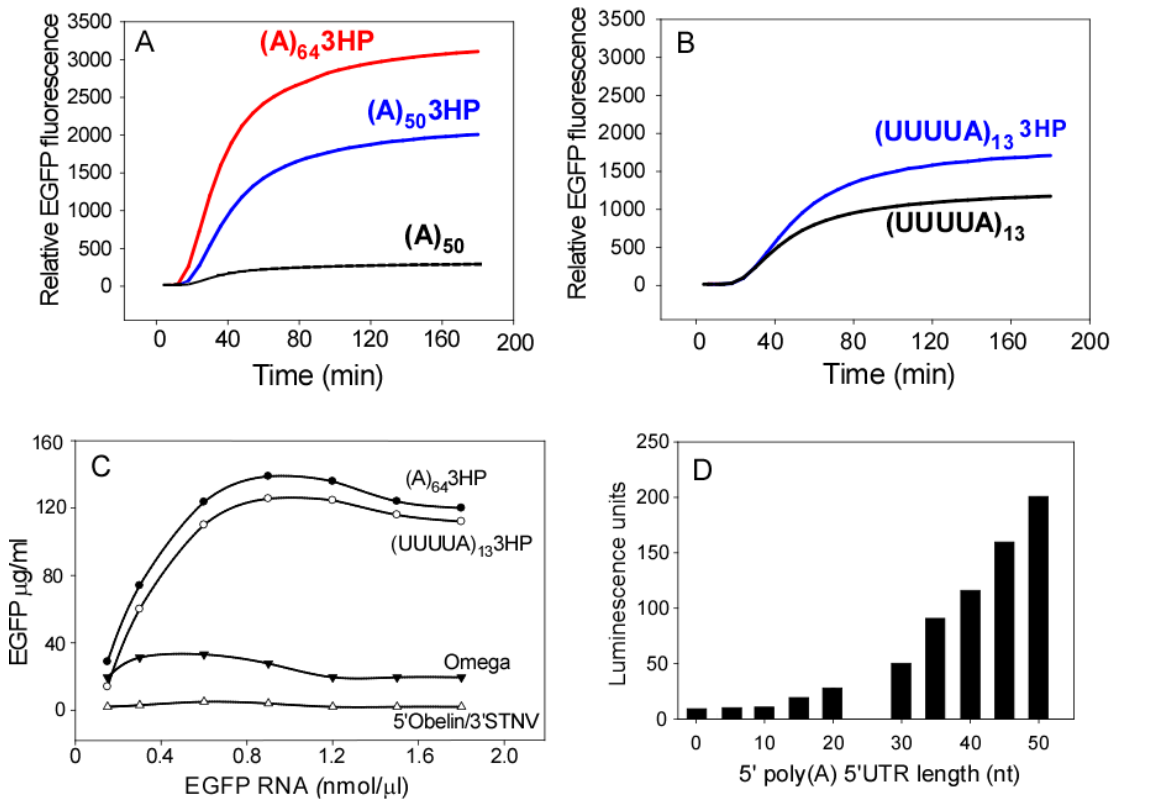

Figure S2. Analysis of the structure-activity relationship of SITS-containing mRNAs in LTE. (A) influence of the 3 hairpins downstream of AUG sequence on the efficiency of EGFP translation in LTE extract. Equal amounts (0.6pmol/μl) of indicated EGFP mRNAs were translated in LTE and changes in fluorescence 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

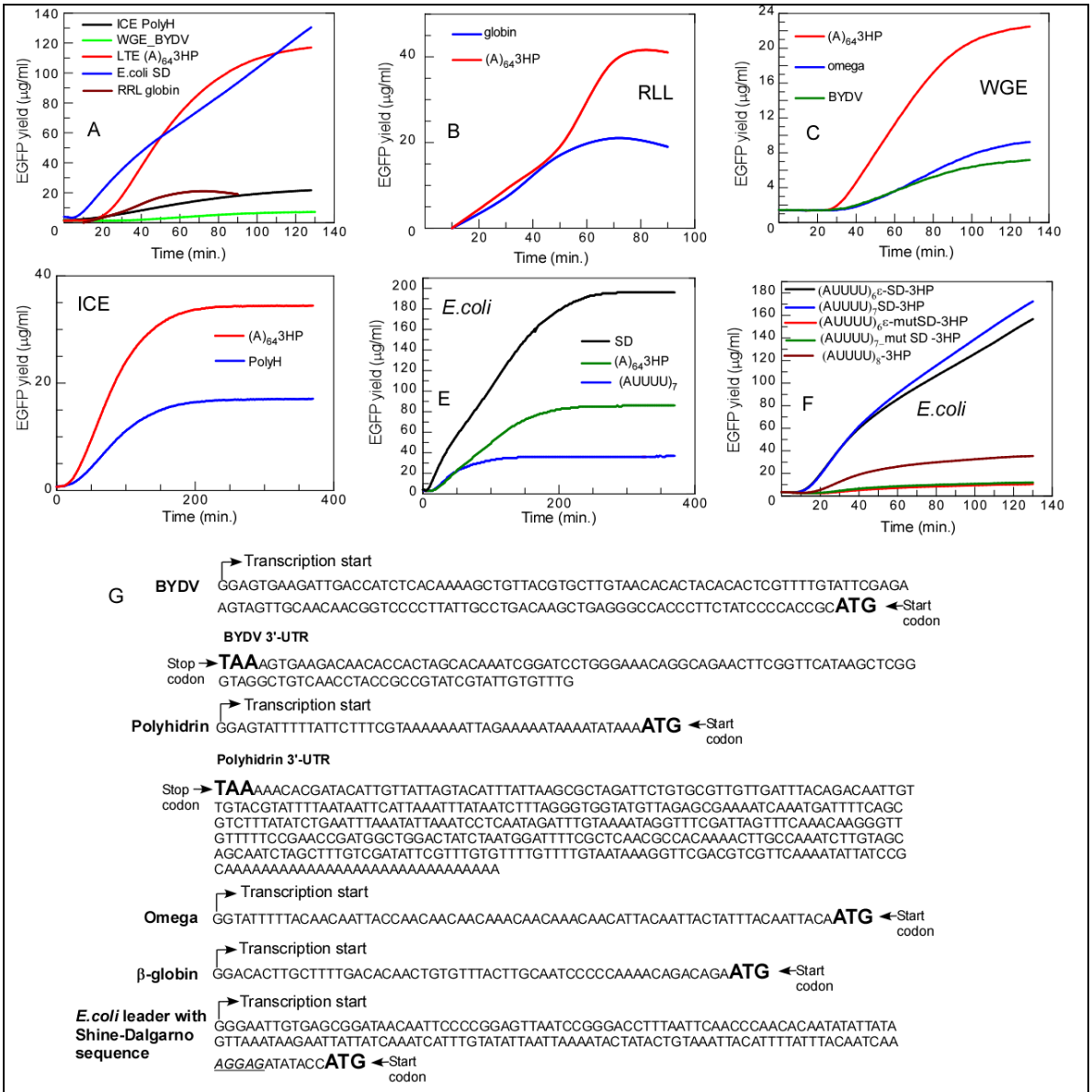

Figure S3. Analysis of translation efficiencies of EGFP encoding mRNA with different translation initiation sequences in coupled prokaryotic and eukaryotic cell-free expression systems. (A) translation of EGFP encoding PCR products in coupled LTE and commercially available coupled cell-free translation 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].

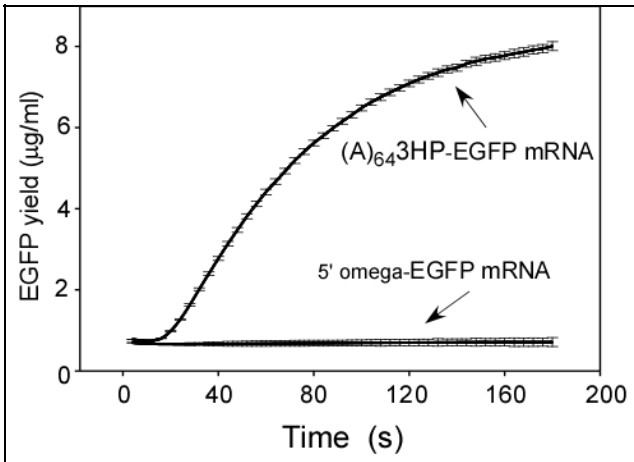

Figure S4. Translation of SITS-EGFP and omega-EGFP mRNAs in extract of *S. cerevisiae* prepared by hypotonic lysis. The lysate prepared by nitrogen cavitation produced similar results (not shown).

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

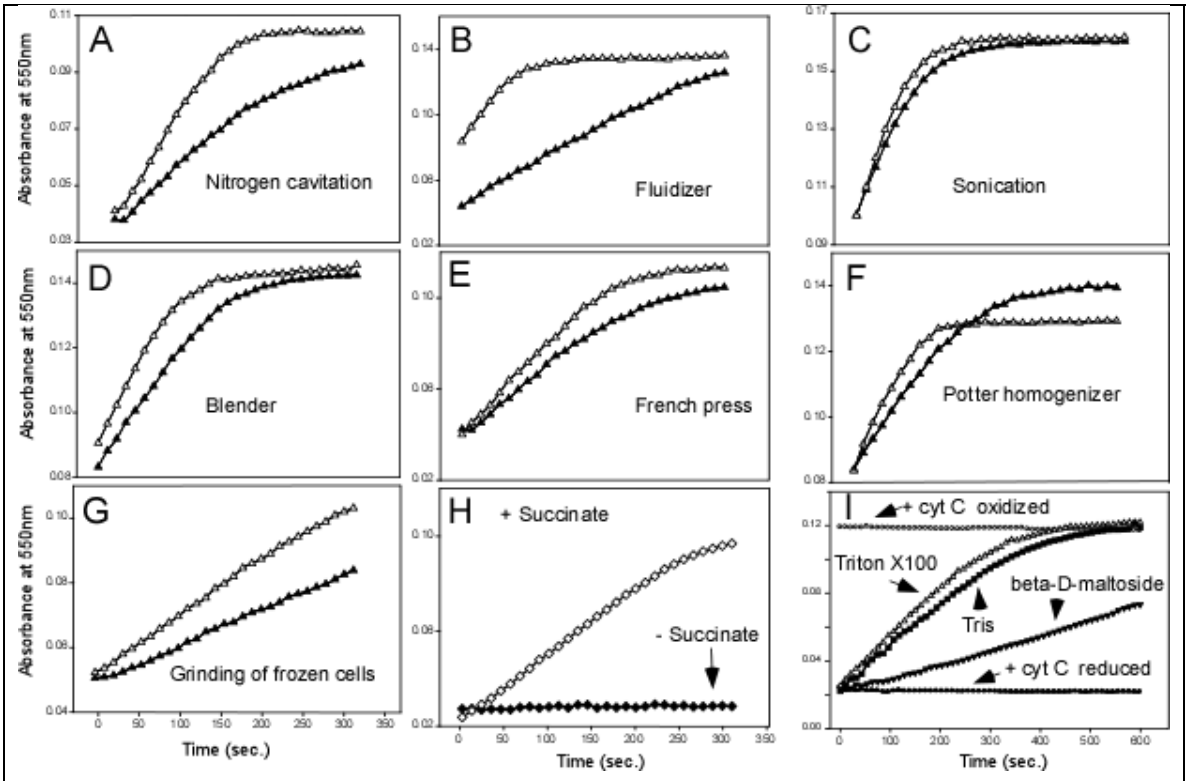

Figure S5**.** (A-G) Activity analysis of cytochrome c in mitochondria preparations isolated from *L. tarentolae* cells disrupted by different methods. The filled triangles reflect the activity of the samples. The empty triangles reflect the total activity of the preparation following TritonX-100 treatment 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.

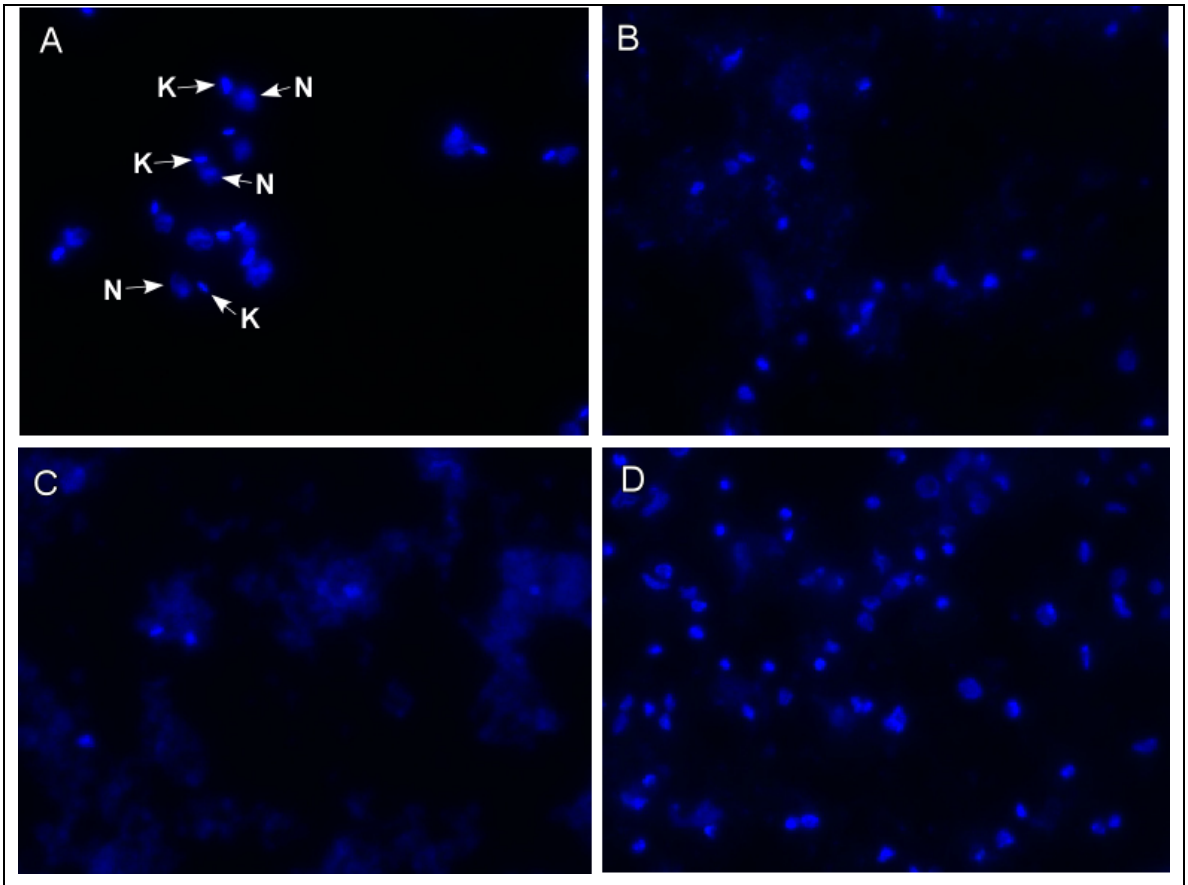

Figure S6. Fluorescence microscopy analysis of DAPI-stained kinetoplasts in *L. tarentolae* cells and mitochondria preparations obtained by different cell disruption protocols: (A) - *L. tarentolae* cells, the arrows denote nuclei -N and kinetoplasts - K; (B) - mitochondria sample obtained form cells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

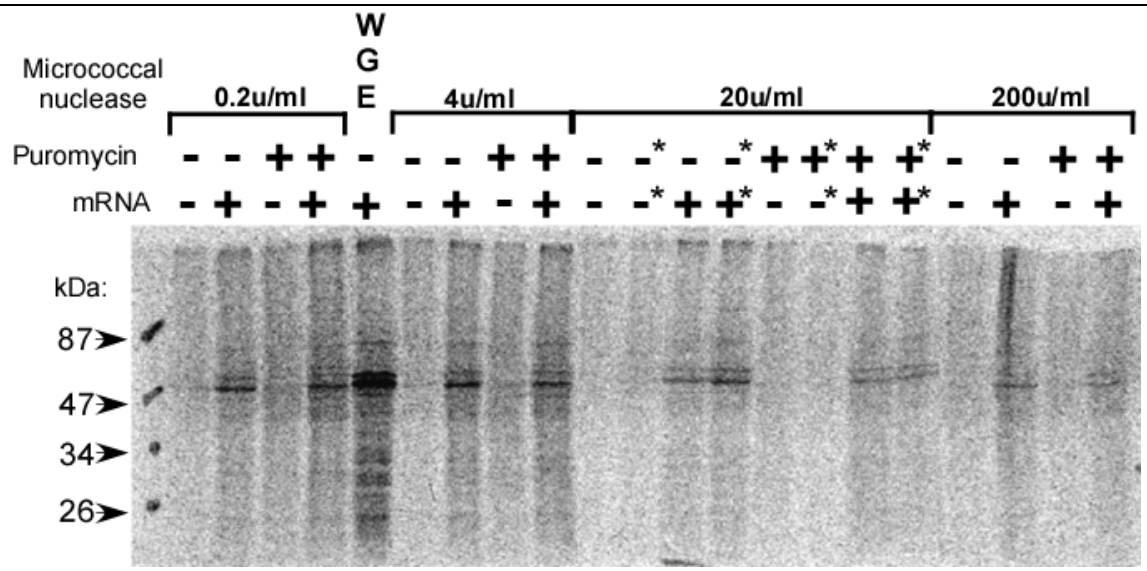

Figure S7. Translation of endogenous mRNA in LTE. SDS-PAGE analysis of the of [C^{14}] Leu supplemented LTE containing purified polyA fraction of *L. tarentoale* mRNA (60 ng/ul) and treated with puromycin or micrococal nuclease. Micrococcal nuclease was used at indicated concentrations to treat extracts 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

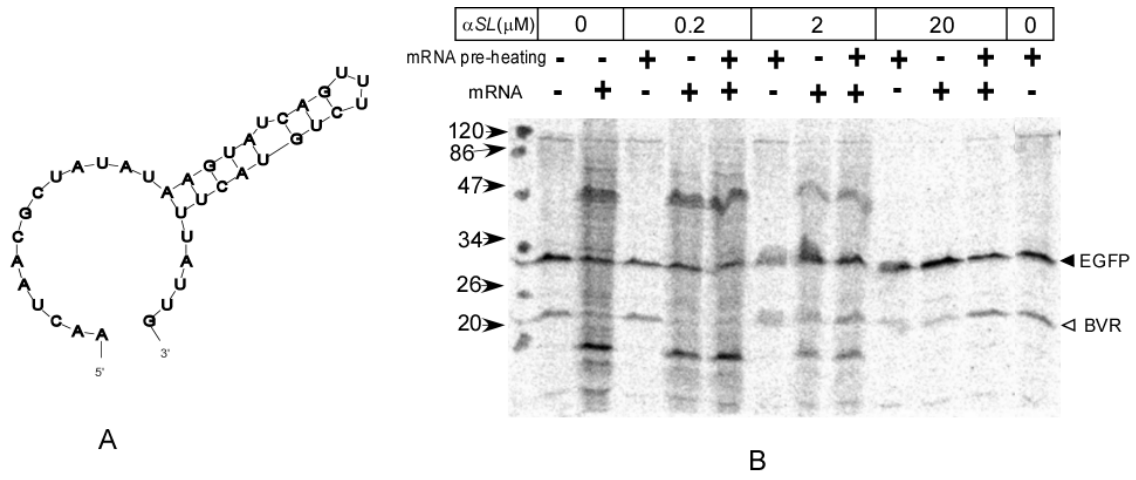

Figure S8. Suppression of *Leishmania* mRNA translation in WGE by anti-SL oligonucleotide. (A) Secondary structure of splice-leader RNA calculated with Mfold 3.2 (ΔG = -3.2 kcal/mol). **(B)** SDS-PAGE analysis of [C^{14}]Leu supplemented WGE primed simultaneously with two RNA templates: the EGFP-encoding 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

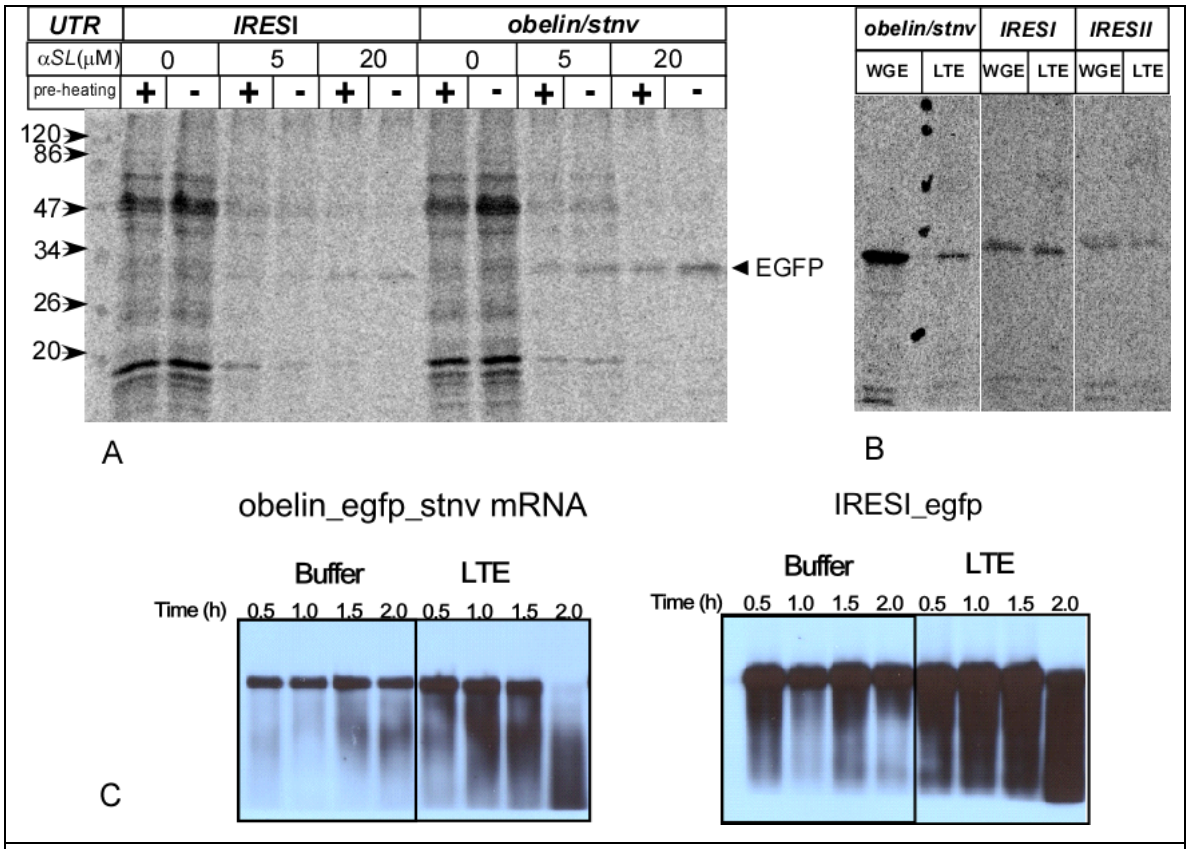

Figure S9. Analysis of translation efficiency and stability of mRNA containing IRESI and *obelin/stnv* UTR combinations in WEG and LTE. (A) The panel depicts SDS-PAGE analysis of the [C^{14}]Leu supplemented LTE primed with EGFP-encoding RNA bearing obelin-derived 5'UTR and STNV-derived 3'UTR, in the 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

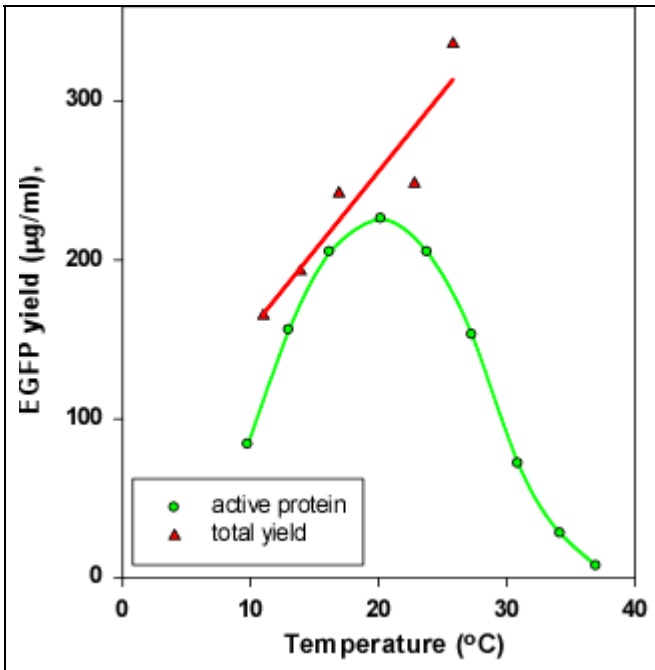

Figure S10. Analysis of EGFP expression level and fraction of active protein as a function of the incubation temperature. 12pmol of myc tagged (A) $_{64}$3HP-EGFP mRNA was incubated with 20µl of LTE at different temperatures and amounts of active EGFP measured by fluorescent scanning using purified EGFP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

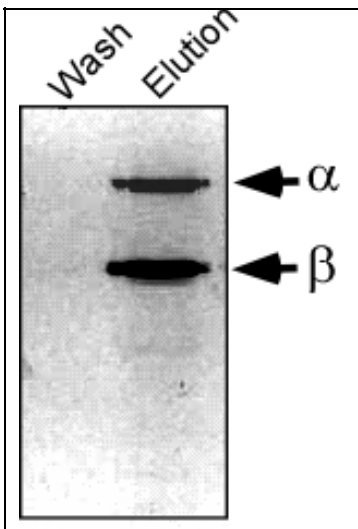

Figure S11. Co-translation in LTE of α and β subunits of Rab geranylgeranyl transferase (RabGGTase) and purification of the α/β heterodimer. The unstained SDS-PAGE gel was subjected to fluorescent scanning (the image is a superimposition of scans in two $\lambda_{ex/em}$ optimally selected for EGFP and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

Nature Biotechnology: doi:10.1038/nbt.1556

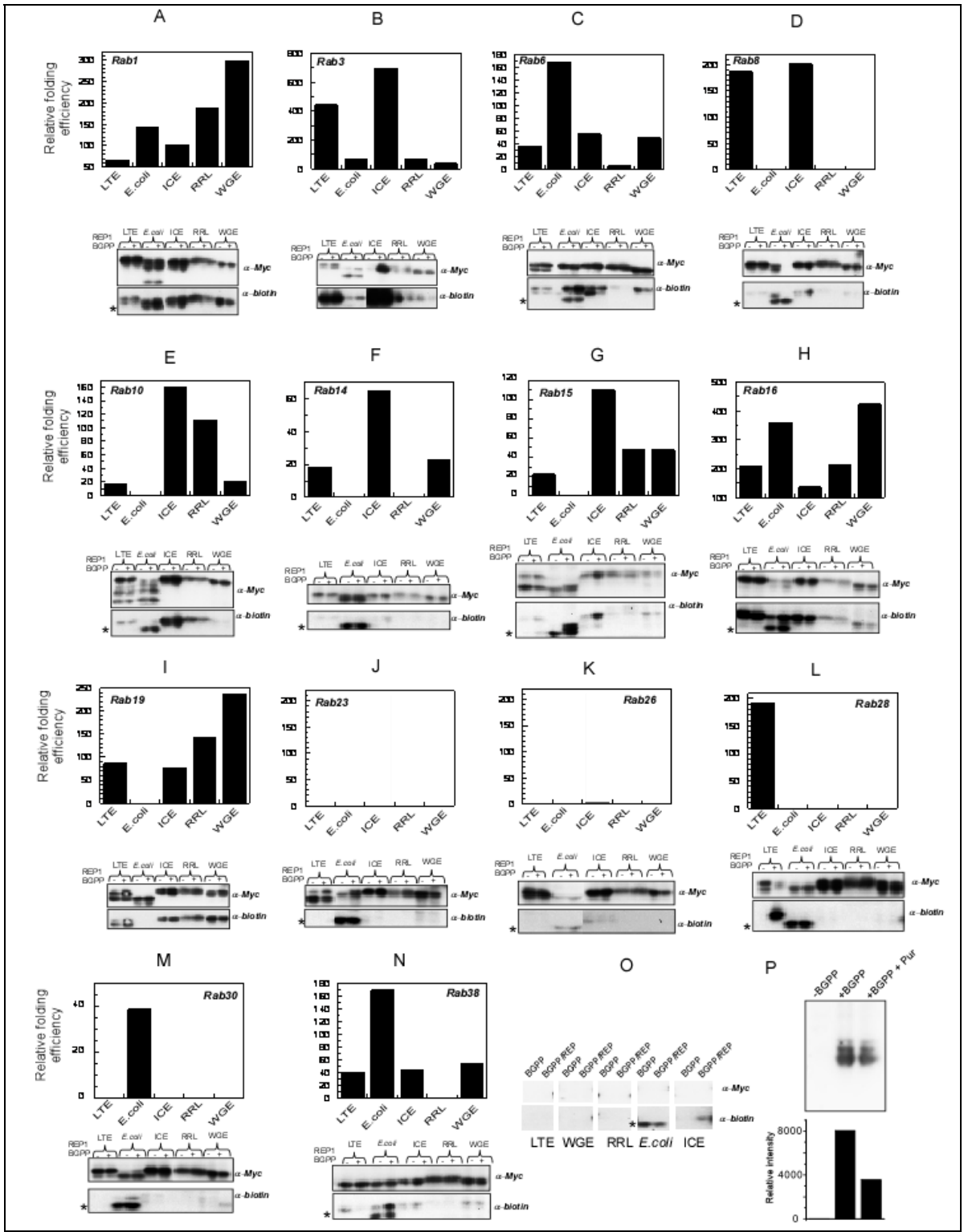

Figure S12. Analysis of folding status of human RabGTPases translated in different cell-free systems. (A-N) Fifteen human RabGTPases were N-terminally fused with myc tag and SITS sequence and translated in LTE, *E. coli,* ICE, RRL and WGE cell-free systems. The resulting translation mixtures were subjected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

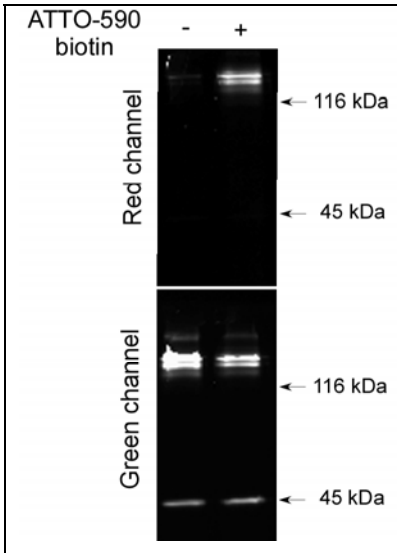

Figure S13. SDS-PAGE analysis of an LTE programmed with $(UUUUA)_7$-streptavidin-EGFP coding mRNA. 10μl of LTE reaction mixture was supplemented with ATTO-590 at 1μM final concentration. After 3 min. incubation at room temperature the samples were resolved on 15% SDS-PAGE. EGFP (green channel) and ATTO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

Nature Biotechnology: doi:10.1038/nbt.1556

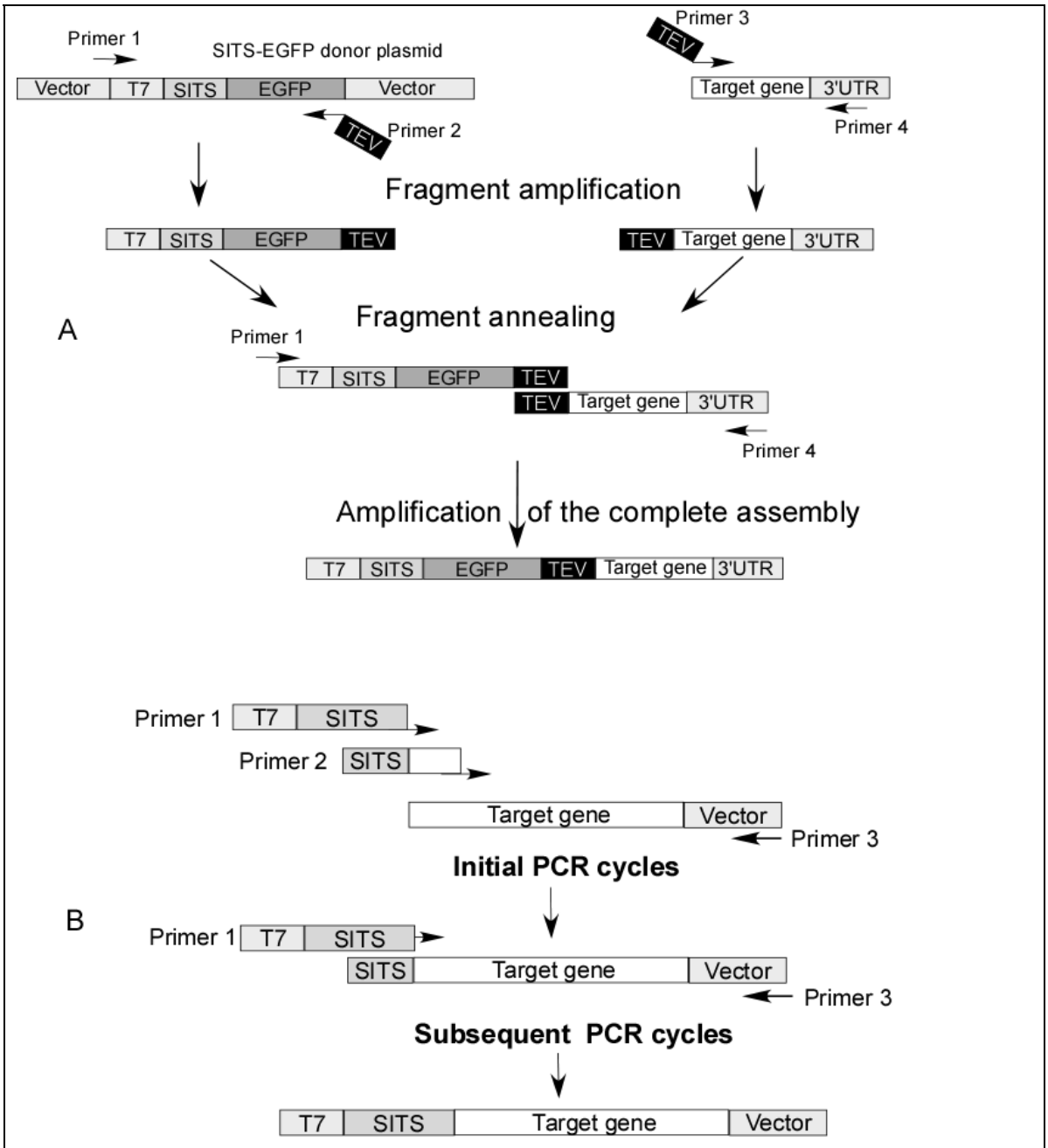

Figure S14. Schematic representation of approaches used for construction of templates. (A) Splicing Overlap Extension based PCR synthesis of the SITS-EGFP tagged ORFs. TEV denotes the recognition sequence of TEV protease. (B) Schematic representation of two primer PCR based procedure for synthesis of 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

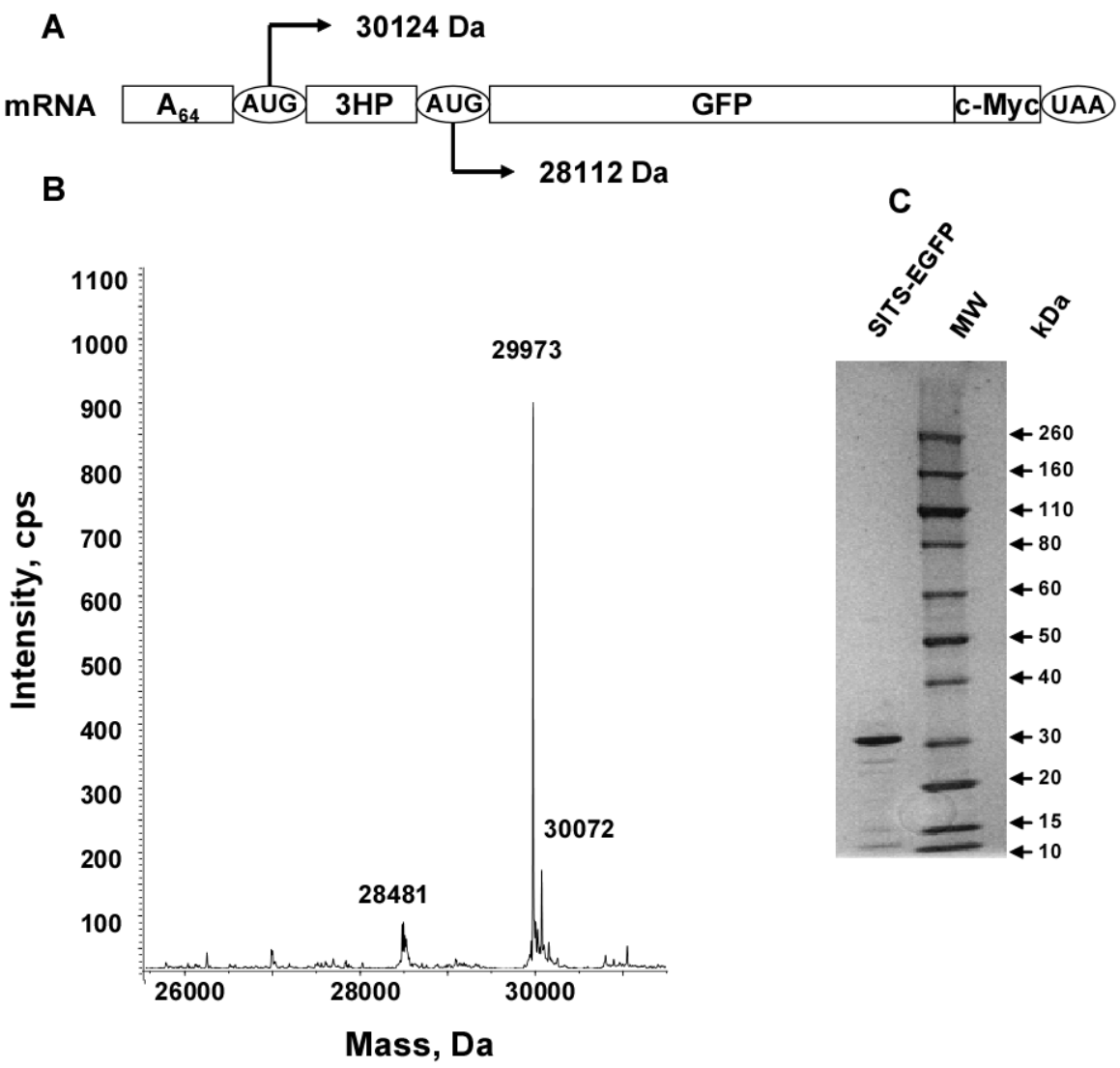

Figure S15. Mapping of the translation start in $(A)_{64}$3HP-EGFP mRNA. (A) Schematic representation of mRNA with two possible translation start sites and the calculated molecular masses of resulting polypeptides; (B) ESI-MS spectrum of purified *E. coli*-expressed SITS-EGFP. The main peak corresponds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

Nature Biotechnology: doi:10.1038/nbt.1556

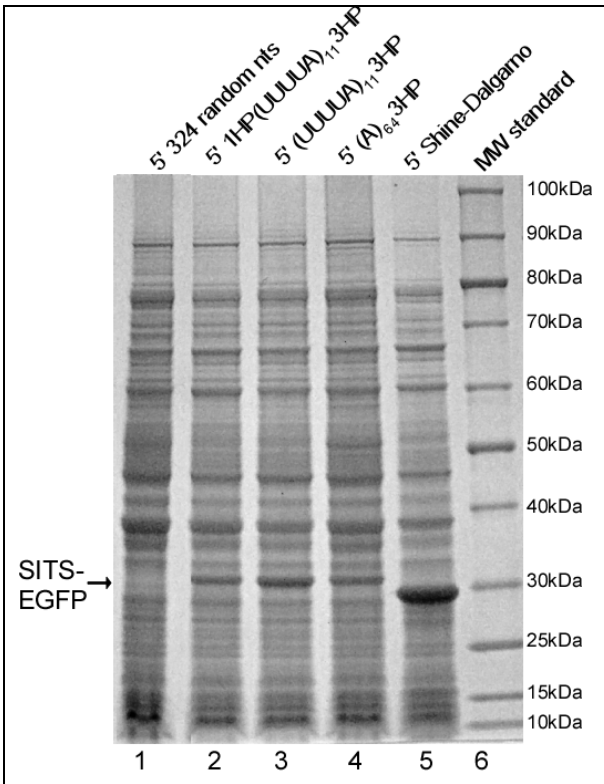

Figure S16. Full size version of the cropped gel from figure 1F. SDS-PAGE gel analysis of lysates of E.coli cells transformed with plasmids carrying indicated SITS sequences (lanes 2,3,4), Shine-Dalgarno sequence (lane 5) or randomly chosen sequence (lane 1). The 1HP of SITS in lane 2 denotes a single 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