



(43) International Publication Date
18 September 2014 (18.09.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/144583 A2

- (51) **International Patent Classification:** Not classified
- (21) **International Application Number:**
PCT/US2014/029056
- (22) **International Filing Date:**
14 March 2014 (14.03.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/792,290 15 March 2013 (15.03.2013) US
- (71) **Applicant:** **NORTHWESTERN UNIVERSITY**
[US/US]; 633 Clark Street, Evanston, IL 60208 (US).
- (72) **Inventors:** **JEWETT, Michael, C.**; 2760 Central Park,
Evanston, IL 60201 (US). **HODGMAN, Charles, E.**; 1930
Ridge Avenue, Apt. C211, Evanston, IL 60201 (US).
GAN, Rui; 800 Hinman Avenue, Apt. 203, Evanston, IL
60202 (US).
- (74) **Agent:** **CELANDER, Daniel, W.**; Klintworth & Rozen-
blat IP LLC, 850 West Jackson Boulevard, Suite 525,
Chicago, IL 60607 (US).
- (81) **Designated States** (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM,
ZW.

- (84) **Designated States** (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

- *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** METHODS FOR CELL- FREE PROTEIN SYNTHESIS

(57) **Abstract:** Cell-free protein synthesis systems and methods of using the same for producing in vitro protein materials in high yield are disclosed. The cell-free protein synthesis platform includes (a) a *Saccharomyces cerevisiae* cellular extract prepared from mid-exponential to late-exponential batch cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀ or fed-batch cultures harvested in mid-exponential to late-exponential phase; (b) a reaction buffer; and (c) a translation template or (c') a transcription template from which a translation template can be prepared *in situ* with an RNA polymerase. A method of performing high-throughput protein synthesis *in vitro* is also provided that utilizes a combined transcription/translation reaction with the cell-free protein synthesis platform from *Saccharomyces cerevisiae*, an RNA polymerase and a transcription template prepared from a source DNA using an amplification procedure.



WO 2014/144583 A2

METHODS FOR CELL-FREE PROTEIN SYNTHESIS

CROSS-REFERENCE TO RELATED APPLICATIONS

[001] This application claims benefit of priority to U.S. provisional application No. 61/792,290, filed on March 15, 2013, which is incorporated by reference in its entirety.

SEQUENCE LISTING

[002] The instant application contains a Sequence Listing, which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on March 14, 2014, is named NWN01-035-PCT_ST25.txt, and is 212,664 bytes in size.

FIELD OF THE INVENTION

[003] This invention pertains to cell-free protein synthesis systems and methods of using the same for producing *in vitro* protein materials in high yield.

BACKGROUND OF THE INVENTION

[004] Cell-free protein synthesis (CFPS) platforms have emerged as a powerful technology for protein expression. Prominent applications include the production of pharmaceutical proteins and vaccines (Goerke, A. R. *et al.* "Development of cell-free protein synthesis platforms for disulfide bonded proteins," *Biotechnol. Bioeng.* **99**, 351-367 (2008); Kanter, G. *et al.* "Cell-free production of scFv fusion proteins: An efficient approach for personalized lymphoma vaccines," *Blood* **109**, 3393-3399, (2007); Stech, M. *et al.* "Production of functional antibody fragments in a vesicle-based eukaryotic cell-free translation system," *J. Biotechnol.* **164**, 220-231 (2012); Yang, J. *et al.* "Rapid expression of vaccine proteins for B-cell lymphoma in a cell-free system," *Biotechnol. Bioeng.* **89**, 503-511 (2005); Yin, G. *et al.* "Aglycosylated antibodies and antibody fragments produced in a scalable *in vitro* transcription-translation system," *MAbs* **4**, 217-225 (2012); Zawada, J. F. *et al.* "Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines," *Biotechnol. Bioeng.* **108**, 1570-1578 (2011)). Such systems enable expression *in vitro* of proteins that are difficult to produce *in vivo*, as well as high-throughput production of protein libraries for protein

evolution, functional genomics, and structural studies (Madin, K. *et al.* “A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: Plants apparently contain a suicide system directed at ribosomes,” *Proc. Natl. Acad. Sci. U.S.A.* **97**, 559-564 (2000); Takai, K *et al.* “Practical cell-free protein synthesis system using purified wheat embryos,” *Nat. Protoc.* **5**, 227-238 (2010)). Prokaryotic *Escherichia coli* extract based cell-free systems have developed rapidly (for a review, see Carlson, E. D. *et al.* “Cell-free protein synthesis: Applications come of age,” *Biotechnol. Adv.* **30**, 1185-1194, (2012)). Yet an integrated eukaryotic platform with similar productivity, scalability, protein folding capability, and cost effectiveness has lagged behind.

[005] The major eukaryotic CFPS platforms previously developed include systems made from wheat germ extract (WGE) (Goshima, N. *et al.* “Human protein factory for converting the transcriptome into an in vitro-expressed proteome,” *Nat. Methods* **5**, 1011-1017 (2008); Hoffmann, M. *et al.* in *Biotechnol Annu Rev* Vol. 10, 1-30 (Elsevier, 2004); Takai *et al.* (2010)), rabbit reticulocyte lysate (RRL) (Jackson, R. J. *et al.* in *Methods Enzymol* Vol. Vol. 96 (eds. Becca Fleischer, Sidney Fleischer) Ch. 4, 50-74 (Academic Press, 1983)); insect cell extract (ICE) (Ezure, T *et al.* “A cell-free protein synthesis system from insect cells,” *Methods Mol. Biol.* **607**, 31-42 (2010); Kubick, S *et al.* in *Current Topics in Membranes*, Vol. 63 (ed. Larry DeLucas) 25-49 (Academic Press, 2009); Tarui, H. *et al.* “Establishment and characterization of cell-free translation/glycosylation in insect cell (*Spodoptera frugiperda* 21) extract prepared with high pressure treatment,” *Appl. Microbiol. Biotechnol.* **55**, 446-453 (2001)); *Leishmania tarentolae* extract (Kovtun, O. *et al.* “Towards the construction of expressed proteomes using a *Leishmania tarentolae* based cell-free expression system,” *PLoS One* **5**, e14388 (2010); Mureev, S. *et al.* “Species-independent translational leaders facilitate cell-free expression,” *Nat. Biotechnol.* **27**, 747-752 (2009)); and HeLa and hybridoma cell extract (Mikami, S. *et al.* in *Cell-Free Protein Production* Vol. 607 *Methods in Molecular Biology* (eds. Yaeta Endo, Kazuyuki Takai, & Takuya Ueda) Ch. 5, 43-52 (Humana Press, 2010)).

[006] Compared to the *E. coli* system, these methods have advantages for producing some types of complex proteins and can achieve post-translational modifications not found in bacteria (Chang, H.-C. *et al.* “De novo folding of GFP fusion proteins: High efficiency in eukaryotes but not in bacteria,” *J. Mol. Biol.* **353**, 397-409 (2005)). Insect cell-extract systems, for example, have demonstrated acetylation and N-myristoylation (Suzuki, T. *et al.*

“N-terminal protein modifications in an insect cell-free protein synthesis system and their identification by mass spectrometry,” *Proteomics* **6**, 4486-4495 (2006)); isoprenylation (Suzuki, T. *et al.* “Protein prenylation in an insect cell-free protein synthesis system and identification of products by mass spectrometry,” *Proteomics* **7**, 1942-1950 (2007)); ubiquitination (Suzuki, T. *et al.* “Preparation of ubiquitin-conjugated proteins using an insect cell-free protein synthesis system,” *J. Biotechnol.* **145**, 73-78 (2010)), core glycosylation (Merk, H. *et al.* “Cell-free synthesis of functional and endotoxin-free antibody Fab fragments by translocation into microsomes,” *Biotechniques* **53**, 153-160 (2012); Tarui *et al.* (2001)); disulfide bond formation in single chain antibody fragments (Stech *et al.* (2012)); and significant advances in expression and modification of membrane bound proteins (Kubick *et al.* (2009)). However, eukaryotic cell-free platforms often have limited batch protein yields (Carlson *et al.* (2012)), or depend on costly and inefficient continuous exchange reactions that do not scale commercially (Zawada *et al.* (2011)). Furthermore, eukaryotic CFPS systems are generally limited by laborious and expensive extract preparation methods. For example, WGE, which is the most common eukaryotic system, requires lengthy preparation steps that include grinding, sieving, extensive washing, and eye selection of the embryo to ensure the embryo is in the proper stage of development (Takai *et al.* (2010)). An additional challenge of this approach is that approximately 5 mL of active extract is produced from 5 to 6 kg of starting material after 4 to 5 days of processing (*Id.*) In contrast, *E. coli* can be processed quickly and under precise growth conditions to develop a highly active and robust CFPS platform, where 60 g of cells (wet weight) can be converted to 120 mL of extract in only 4-6 hours of preparation (Liu, D.V. *et al.* “Streamlining *Escherichia coli* S30 extract preparation for economical cell-free protein synthesis,” *Biotechnol Prog* **21**, 460-465 (2005)). The above limitations motivate the need for a new eukaryotic CFPS platform that is robust, easy to prepare, highly active, and amenable to economical scale-up.

[007] *S. cerevisiae*, like *E. coli*, is a microbe that can be grown quickly under precise conditions in either a bioreactor or shake flasks. Furthermore, *S. cerevisiae* as a eukaryotic organism is suited to fold eukaryotic proteins and has previously shown some ability for post-translational modifications *in vitro*, such as glycosylation (Rothblatt, J. A. *et al.* “Secretion in yeast: Reconstitution of the translocation and glycosylation of alpha-factor and invertase in a homologous cell-free system,” *Cell* **44**, 619-628 (1986)). Because it is a model organism for molecular study, *S. cerevisiae* is well understood at the biochemical level, has a wealth of documented “omics” that can prove useful when trying to characterize a cell-free

system, and genetic tools are readily available for facile changes to the host strain (Nielsen, J. *et al.* “Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*,” *FEMS Yeast Res.* **8**, 122-131 (2008)). *S. cerevisiae* is also an important bio-manufacturing production platform and accounted for 18.5% of all FDA and EMA licensed recombinant protein biopharmaceuticals as of January 2009 (Ferrer-Miralles, N., *et al.* “Microbial factories for recombinant pharmaceuticals,” *Microb. Cell. Fact.* **8**, 17 (2009)).

[008] Despite these attractive features, yeast based CFPS systems have not been extensively developed as a protein synthesis platform since their origin in the 1970s and early 1980s (Gasior, E. *et al.* “The analysis of intermediary reactions involved in protein synthesis, in a cell-free extract of *Saccharomyces cerevisiae* that translates natural messenger ribonucleic acid,” *J. Biol. Chem.* **254**, 3970-3976 (1979); Gasior, E. *et al.* “The preparation and characterization of a cell-free system from *Saccharomyces cerevisiae* that translates natural messenger ribonucleic acid,” *J. Biol. Chem.* **254**, 3965-3969 (1979)). Instead, the majority of research involving yeast cell-free translation systems has focused on investigating translation from a fundamental perspective, such as elucidating cap-dependent translation (Iizuka, N. *et al.* “Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccharomyces cerevisiae*,” *Mol. Cell. Biol.* **14**, 7322-7330 (1994); Iizuka, N. & Sarnow, P. “Translation-competent extracts from *Saccharomyces cerevisiae*: Effects of L-A RNA, 5' cap, and 3' poly(A) tail on translational efficiency of mRNAs,” *Methods* **11**, 353-360 (1997)) and characterizing translation initiation factors (Algire, M. A. *et al.* “Development and characterization of a reconstituted yeast translation initiation system,” *RNA* **8**, 382-397 (2002); Hinnebusch, A. G., *et al.* “Mechanism of translation initiation in the yeast *Saccharomyces cerevisiae*,” pp. 225–268 in *Translational Control in Biology and Medicine*, (eds. M. B. Mathews, N. Sonenberg and J. W. B. Hershey) (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2007); Kurata, S. *et al.* “Ribosome recycling step in yeast cytoplasmic protein synthesis is catalyzed by eEF3 and ATP,” *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10854-10859 (2010); Saini, P. *et al.* “Hypusine-containing protein eIF5A promotes translation elongation,” *Nature* **459**, 118-121 (2009); Thompson, S. R. *et al.* “Internal initiation in *Saccharomyces cerevisiae* mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element,” *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12972-12977 (2001)). Despite this focus, some recent work has shown the potential to use yeast CFPS for making proteins of interest, such as virus-like particles (Wang, X. *et al.* “An optimized yeast cell-free system: Sufficient for translation of human

papillomavirus 58 L1 mRNA and assembly of virus-like particles,” *J. Biosci. Bioeng.* **106**, 8-15 (2008); Wang, X. *et al.* “Translational comparison of HPV58 long and short L1 mRNAs in yeast (*Saccharomyces cerevisiae*) cell-free system,” *J. Biosci. Bioeng.* **110**, 58-65 (2010)) and additional viral proteins (Pogany, J. *et al.* “Authentic replication and recombination of tomato bushy stunt virus RNA in a cell-free extract from yeast,” *J. Virol.* **82**, 5967-5980 (2008)).

BRIEF SUMMARY OF THE INVENTION

[009] In first aspect, a cell-free protein synthesis platform for preparing protein from a translation template is disclosed. The cell-free protein synthesis platform includes (a) a *Saccharomyces cerevisiae* cellular extract, (b) a reaction buffer, and (c) the translation template. The *Saccharomyces cerevisiae* cellular extract is prepared from mid-exponential to late-exponential batch cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀ or fed-batch cultures harvested in mid-exponential to late-exponential phase.

[010] In a second aspect, a cell-free protein synthesis platform for preparing protein from a transcription template is disclosed. The cell-free protein synthesis platform includes (a) a *Saccharomyces cerevisiae* cellular extract, (b) a reaction buffer, (c) an RNA polymerase, and (d) the transcription template. The *Saccharomyces cerevisiae* cellular extract is prepared from mid-exponential to late-exponential cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀. The RNA polymerase is capable of transcribing the transcription template to form a translation template and the *Saccharomyces cerevisiae* cellular extract can sustain protein synthesis through a combined transcription/translation reaction.

[011] In a third aspect, a method of performing high-throughput protein synthesis *in vitro* is disclosed. The method includes several steps. The first step is providing a source nucleic acid. The second step is preparing a transcription template from the source nucleic acid. The third step is synthesizing protein *in vitro* using a cell-free protein synthesis platform utilizing the transcription template. The cell-free protein synthesis platform comprises (i) a *Saccharomyces cerevisiae* cellular extract, (ii) a reaction buffer, and (iii) an RNA polymerase. The *Saccharomyces cerevisiae* cellular extract is prepared from mid-exponential to late-exponential cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀. The RNA polymerase is capable of transcribing from the transcription template to form the translation

template, and the *Saccharomyces cerevisiae* cellular extract can sustain protein synthesis through a combined transcription/translation reaction.

BRIEF DESCRIPTION OF THE FIGURES

[012] **FIG. 1A** depicts a schematic of an exemplary prior art method for yeast extract preparation and CFPS system.

[013] **FIG. 1B** depicts a schematic of the disclosed innovative method for yeast extract preparation and CFPS system.

[014] **FIG. 1C** illustrates active luciferase yield from cell-free translation only reactions, wherein the two extract preparation methods of **FIG. 1A** and **1B** are compared in the presence or absence of Micrococcal Nuclease (MNase) pre-treatment. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments.

[015] **FIG. 1D** illustrates that MNase pre-treatment degrades ribosomal RNA (rRNA), as shown by the disappearance of the original rRNA bands. Lane 1: 200 bp DNA step ladder (Promega, Madison, WI), lane 2: Crude extract with no MNase pre-treatment, lanes 3-6: Crude extract after MNase pre-treatment for 0, 10, 20, and 30 min, respectively.

[016] **FIG. 2** depicts exemplary expression constructs (SEQ ID NOS:1-24) to evaluate the effects of 5'-UTR and 3'-UTR elements in the DNA transcription template.

[017] **FIG. 3A** depicts three 5'-UTRs from tobacco mosaic virus (Ω), tobacco etch virus (TEV), and tobamovirus (Tbm) were tested for the ability to enable translation initiation in combined yeast CFPS.

[018] **FIG. 3B** depicts results of protein expression from templates that include Kozak sequences in yeast CFPS reactions. Yeast Tx/Tl CFPS reactions primed with PCR products containing the Ω sequence combined with variations of the Kozak sequence are compared over CFPS reaction lifetime. At the appropriate time point, 2 μ L from each reaction were sampled and assayed for fluorescence intensity. Values show means with error bars representing high and low values of 2 independent experiments. Values are displayed relative to the final yield for a CFPS reaction without inclusion of any Kozak sequence.

[019] **FIG. 3C** depicts different lengths of Poly(A) tails, 25nt, 50nt, 90nt, and 170nt were tested for the ability to enable translation initiation in combined yeast CFPS.

[020] **FIG. 3D** depicts the effect of 5' Ω sequence and 3' poly(A) tail in cap-independent translation initiation of yeast CFPS demonstrates that both the leader sequence and poly(A) tail are required for efficient translation with Ω . The structures of the expression templates are shown in **FIG. 2**.

[021] **FIG. 3E** The effect of various 3'-UTRs in combined yeast CFPS. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments. Luciferase data are presented in RLUs, or relative light units.

[022] **FIG. 4A** depicts a time course for combined transcription and translation activated in yeast CFPS from the Ω leader sequence. Active luciferase synthesis is shown over the course of a standard batch reaction.

[023] **FIG. 4B** depicts the physicochemical environment of the CFPS reaction was optimized by altering temperature. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments. Luciferase data are presented in RLUs, or relative light units.

[024] **FIG. 4C** depicts the physicochemical environment of the CFPS reaction was optimized by altering DNA template amount. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments. Luciferase data are presented in RLUs, or relative light units.

[025] **FIG. 4D** depicts the physicochemical environment of the CFPS reaction was optimized by altering DTT concentration. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments. Luciferase data are presented in RLUs, or relative light units.

[026] **FIG. 5** depicts a schematic of a robust and rapid two-step method for assembly of linear DNA templates for high-throughput protein expression in yeast CFPS. In a first PCR assay (PCR 1; designated "(1)"), protein-coding sequences can be amplified directly from plasmid, genome, cDNA or other genetic material using gene-specific primers with universal tail sequence. In a second PCR assay (PCR 2; designated "(2)"), the T7 promoter, Ω

sequence, and poly(A)50 are added to protein-coding region by overlapping PCR to the universal tail sequences. The assembled linear PCR product can be used directly as template in combined Tx/Tl yeast CFPS.

[027] **FIG. 6** demonstrates linear expression templates activating high-level yeast combined transcription and translation.

[028] **FIG. 6A** illustrates a comparison of the efficiency in CFPS of PCR product amplified directly from plasmid, assembled linear DNA template, and plasmid. The concentration of template is 5.3 nM for each sample. These data show that linear templates outperform plasmid DNA. Values show means with error bars representing standard deviations (s.d.) of at least 4 independent experiments. Luciferase data are presented in RLUs, or relative light units.

[029] **FIG. 6B** illustrates expression of CFPS of luciferase (Luc), green fluorescence protein (GFP), and chloramphenicol acetyl transferase (CAT) using assembled, linear DNA templates. The [³⁵S]-methionine-labeled products were analyzed in SDS-PAGE and autoradiography. Lane 1 and 2: soluble and insoluble fraction of Luc; Lane 3 and 4: soluble and insoluble fraction of GFP; Lane 5 and 6: soluble and insoluble fraction of CAT; Lane 7 and 8: soluble and insoluble fraction of negative control reaction without template. Notably, nearly the entire product for each protein is soluble.

[030] **FIG. 7** depicts a time course assay of active SuperFolder GFP (sfGFP) in yeast CFPS batch reactions.

[031] **FIG. 8** illustrates a comparison of the disclosed yeast CFPS method to an exemplary prior art method, as measured by active protein synthesis yield ($\mu\text{g mL}^{-1}$; left axis) and relative product yield (μg protein synthesized per \$ reagent cost; right axis). Substrate cost includes all substrates used to treat the crude extract, make the genetic template, and assemble the CFPS reaction.

[032] **FIG. 9A** depicts the effects of typical growth curve for yeast culture on YPAD media, pH 5.5 at 30° C.

[033] **FIG. 9B** depicts active luciferase yield in combined transcription and translation reactions from extracts made with cells harvested at 3, 6, 9, 12, 15, and 18 OD₆₀₀. Extracts

were grouped together according to their protein synthesis activity. The groupings were 3, 6-12, and 15-18 OD₆₀₀ with most active extract at mid-exponential phase harvest. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments.

[034] **FIG. 10** depicts the effect of adding inorganic phosphate to the growth media affects CFPS activity. Cells were grown in shake flasks in YPAD media, pH 5.5 supplemented with 0, 50, 100, and 150 mM potassium phosphate (pH 5.5). Crude extract was subsequently made from the variable cell cultures and compared to each other for protein synthesis activity. Maximum protein synthesis occurred with media supplemented with 50 mM potassium phosphate.

[035] **FIG. 11A** depicts the physicochemical environment of the CFPS reaction as a function of magnesium concentration.

[036] **FIG. 11B** depicts the physicochemical environment of the CFPS reaction as a function of temperature.

[037] **FIG. 11C** depicts the physicochemical environment of the CFPS reaction as a function of anions used with potassium and magnesium salts (chloride, acetate, and glutamate as shown).

[038] **FIG. 11D** depicts the physicochemical environment of the CFPS reaction as a function of polyamine concentration (putrescine and spermidine).

[039] **FIG. 11E** depicts the effect of adding RNase Inhibitor to the CFPS reaction. RNase Inhibitor (Qiagen, Valencia, CA) and RNase Inhibitor "Storage Buffer" (2 mM KH₂PO₄, 8.0 mM Na₂HPO₄, 3.0 mM KCl, 150 mM NaCl, pH 7.4, and 50% glycerol) were each added to the cell-free reaction in equal volumes and compared for active luciferase yield.

[040] **FIG. 11F** depicts the physicochemical environment of the CFPS reaction as a function of glycerol concentration. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments.

[041] **FIG. 12A** depicts commercially available RNAP (New England Biolabs, Ipswich, MA) from bacteriophage SP6 and T3 were compared to the established T7 RNAP for activity

in yeast CFPS reactions. Values are displayed relative to T7 RNAP at the given concentration. Values represent means with error bars displaying high and low values of 2 independent experiments.

[042] **FIG. 12B** depicts an exemplary comparison of T7 RNAP that is commercially available to T7 RNAP prepared in house. Values are displayed relative to T7 RNAP prepared in house. Values represent means with error bars displaying high and low values of 2 independent experiments.

[043] **FIG. 13A** depicts synthesis of active luciferase over the course of a batch reaction. Fifteen μ L batch reactions were prepared in different tubes for each time point and sampled for active luciferase yield.

[044] **FIG. 13B** illustrates an exemplary experimental design schematic of “pre-incubation” experiments.

[045] **FIG. 13C** depicts active luciferase yield from extracts pre-incubated for the specified time. Values show means with error bars representing standard deviations (s.d.) of at least 3 independent experiments.

DETAILED DESCRIPTION OF THE INVENTION

[046] A novel cell-free protein synthesis (CFPS) platform from yeast extract is disclosed. The platform provides a 250-fold increase in protein synthesis yield and 2000-fold reduction in protein synthesis cost over extant prior art methods. The new platform displays robust protein synthesis from combined transcription-translation systems having the capability of efficiently utilizing linear transcription templates as input substrates. These improvements have direct implications for high-throughput protein expression, industrial synthesis of pharmaceutical or biotechnological relevant proteins, bench-top laboratory protein expression using an *in vitro* protein expression kit, protein expression for crystallography and proteomics.

Definitions

[047] To aid in understanding the invention, several terms are defined below.

[048] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of skill in the art. Although any methods and materials similar to or equivalent to those described herein can be used in the practice or testing of the claims, the exemplary methods and materials are described herein.

[049] Moreover, reference to an element by the indefinite article “a” or “an” does not exclude the possibility that more than one element is present, unless the context clearly requires that there be one and only one element. The indefinite article “a” or “an” thus usually means “at least one.”

[050] The term “about” means within a statistically meaningful range of a value or values such as a stated concentration, length, molecular weight, pH, time frame, temperature, pressure or volume. Such a value or range can be within an order of magnitude, typically within 20%, more typically within 10%, and even more typically within 5% of a given value or range. The allowable variation encompassed by “about” will depend upon the particular system under study.

[051] The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to,”) unless otherwise noted.

[052] Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, and includes the endpoint boundaries defining the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein.

[053] The terms “nucleic acid” and “oligonucleotide,” as used herein, refer to polydeoxyribonucleotides (containing 2-deoxy-D-ribose), polyribonucleotides (containing D-ribose), and to any other type of polynucleotide that is an N glycoside of a purine or pyrimidine base. There is no intended distinction in length between the terms “nucleic acid”, “oligonucleotide” and “polynucleotide”, and these terms will be used interchangeably. These terms refer only to the primary structure of the molecule. Thus, these terms include double- and single-stranded DNA, as well as double- and single-stranded RNA. For use in the present invention, an oligonucleotide also can comprise nucleotide analogs in which the base, sugar or phosphate backbone is modified as well as non-purine or non-pyrimidine nucleotide analogs.

[054] Oligonucleotides can be prepared by any suitable method, including direct chemical synthesis by a method such as the phosphotriester method of Narang *et al.*, 1979, *Meth. Enzymol.* **68**:90-99; the phosphodiester method of Brown *et al.*, 1979, *Meth. Enzymol.* **68**:109-151; the diethylphosphoramidite method of Beaucage *et al.*, 1981, *Tetrahedron Letters* **22**:1859-1862; and the solid support method of U.S. Pat. No. 4,458,066, each incorporated herein by reference. A review of synthesis methods of conjugates of oligonucleotides and modified nucleotides is provided in Goodchild, 1990, *Bioconjugate Chemistry* **1**(3): 165-187, incorporated herein by reference.

[055] The term “primer,” as used herein, refers to an oligonucleotide capable of acting as a point of initiation of DNA synthesis under suitable conditions. Such conditions include those in which synthesis of a primer extension product complementary to a nucleic acid strand is induced in the presence of four different nucleoside triphosphates and an agent for extension (for example, a DNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature.

[056] A primer is preferably a single-stranded DNA. The appropriate length of a primer depends on the intended use of the primer but typically ranges from about 6 to about 225 nucleotides, including intermediate ranges, such as from 15 to 35 nucleotides, from 18 to 75 nucleotides and from 25 to 150 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template nucleic acid, but must be sufficiently complementary to hybridize with the template. The design of suitable primers for the amplification of a given target sequence is well known in the art and described in the literature cited herein.

[057] Primers can incorporate additional features which allow for the detection or immobilization of the primer but do not alter the basic property of the primer, that of acting as a point of initiation of DNA synthesis. For example, primers may contain an additional nucleic acid sequence at the 5' end which does not hybridize to the target nucleic acid, but which facilitates cloning or detection of the amplified product, or which enables transcription of RNA (for example, by inclusion of a promoter) or translation of protein (for example, by inclusion of a 5'-UTR, such as an Internal Ribosome Entry Site (IRES) or a 3'-UTR element, such as a poly(A)_n sequence, where n is in the range from about 20 to about 200). The region

of the primer that is sufficiently complementary to the template to hybridize is referred to herein as the hybridizing region.

[058] The term “promoter” refers to a cis-acting DNA sequence that directs RNA polymerase and other trans-acting transcription factors to initiate RNA transcription from the DNA template that includes the cis-acting DNA sequence.

[059] The terms “target”, “target sequence”, “target region”, and “target nucleic acid,” as used herein, are synonymous and refer to a region or sequence of a nucleic acid which is to be amplified, sequenced or detected.

[060] The term “hybridization,” as used herein, refers to the formation of a duplex structure by two single-stranded nucleic acids due to complementary base pairing. Hybridization can occur between fully complementary nucleic acid strands or between “substantially complementary” nucleic acid strands that contain minor regions of mismatch. Conditions under which hybridization of fully complementary nucleic acid strands is strongly preferred are referred to as “stringent hybridization conditions” or “sequence-specific hybridization conditions”. Stable duplexes of substantially complementary sequences can be achieved under less stringent hybridization conditions; the degree of mismatch tolerated can be controlled by suitable adjustment of the hybridization conditions. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length and base pair composition of the oligonucleotides, ionic strength, and incidence of mismatched base pairs, following the guidance provided by the art (see, e.g., Sambrook et al., 1989, *Molecular Cloning—A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; Wetmur, 1991, *Critical Review in Biochem. and Mol. Biol.* 26(3/4):227-259; and Owczarzy et al., 2008, *Biochemistry*, 47: 5336-5353, which are incorporated herein by reference).

[061] The term “amplification reaction” refers to any chemical reaction, including an enzymatic reaction, which results in increased copies of a template nucleic acid sequence or results in transcription of a template nucleic acid. Amplification reactions include reverse transcription, the polymerase chain reaction (PCR), including Real Time PCR (see U.S. Pat. Nos. 4,683,195 and 4,683,202; *PCR Protocols: A Guide to Methods and Applications* (Innis et al., eds, 1990)), and the ligase chain reaction (LCR) (see Barany et al., U.S. Pat. No. 5,494,810). Exemplary “amplification reactions conditions” or “amplification conditions”

typically comprise either two or three step cycles. Two-step cycles have a high temperature denaturation step followed by a hybridization/elongation (or ligation) step. Three step cycles comprise a denaturation step followed by a hybridization step followed by a separate elongation step.

[062] As used herein, a “polymerase” refers to an enzyme that catalyzes the polymerization of nucleotides. “DNA polymerase” catalyzes the polymerization of deoxyribonucleotides. Known DNA polymerases include, for example, *Pyrococcus furiosus* (Pfu) DNA polymerase, *E. coli* DNA polymerase I, T7 DNA polymerase and *Thermus aquaticus* (Taq) DNA polymerase, among others. “RNA polymerase” catalyzes the polymerization of ribonucleotides. The foregoing examples of DNA polymerases are also known as DNA-dependent DNA polymerases. RNA-dependent DNA polymerases also fall within the scope of DNA polymerases. Reverse transcriptase, which includes viral polymerases encoded by retroviruses, is an example of an RNA-dependent DNA polymerase. Known examples of RNA polymerase (“RNAP”) include, for example, T3 RNA polymerase, T7 RNA polymerase, SP6 RNA polymerase and *E. coli* RNA polymerase, among others. The foregoing examples of RNA polymerases are also known as DNA-dependent RNA polymerase. The polymerase activity of any of the above enzymes can be determined by means well known in the art.

[063] As used herein, a primer is “specific,” for a target sequence if, when used in an amplification reaction under sufficiently stringent conditions, the primer hybridizes primarily to the target nucleic acid. Typically, a primer is specific for a target sequence if the primer-target duplex stability is greater than the stability of a duplex formed between the primer and any other sequence found in the sample. One of skill in the art will recognize that various factors, such as salt conditions as well as base composition of the primer and the location of the mismatches, will affect the specificity of the primer, and that routine experimental confirmation of the primer specificity will be needed in many cases. Hybridization conditions can be chosen under which the primer can form stable duplexes only with a target sequence. Thus, the use of target-specific primers under suitably stringent amplification conditions enables the selective amplification of those target sequences that contain the target primer binding sites.

[064] As used herein, “expression template” refers to a nucleic acid that serves as substrate for transcribing at least one RNA that can be translated into a polypeptide or protein.

Expression templates include nucleic acids composed of DNA or RNA. Suitable sources of DNA for use a nucleic acid for an expression template include genomic DNA, cDNA and RNA that can be converted into cDNA. Genomic DNA, cDNA and RNA can be from any biological source, such as a tissue sample, a biopsy, a swab, sputum, a blood sample, a fecal sample, a urine sample, a scraping, among others. The genomic DNA, cDNA and RNA can be from host cell or virus origins and from any species, including extant and extinct organisms. As used herein, “expression template” and “transcription template” have the same meaning and are used interchangeably.

[065] As used herein, “translation template” refers to an RNA product of transcription from an expression template that can be used by ribosomes to synthesize polypeptide or protein.

[066] As used herein, the term “cap” (or “5’-cap”) refers to a chemical modification of the 5’-terminus of a translation template. A cap for eukaryotic translation templates can include a guanine nucleotide connected to the mRNA via a 5’ to 5’ triphosphate linkage (“5’, 5’-GpppG” or “G(5')ppp(5')G”). The N-7 position guanine cap can be methylated (“m⁷GpppG” or “m⁷G(5')ppp(5')G”). Translation templates that include cap can be designated by 5’,5’-GpppG-, G(5')ppp(5')G-, m⁷G(5')ppp(5')G- or m⁷GpppG-translation templates.

[067] As used herein, “cap-dependent,” as the term modifies “translation” or “translation template,” refers to the requirement of the translation template to include a 5’-cap for efficient protein synthesis from that translation template.

[068] As used herein, “cap-independent,” as the term modifies “translation” or “translation template,” refers to the lack of a requirement that the translation template include a 5’-cap for efficient protein synthesis from that translation template.

[069] The term “reaction mixture,” as used herein, refers to a solution containing reagents necessary to carry out a given reaction. An “amplification reaction mixture”, which refers to a solution containing reagents necessary to carry out an amplification reaction, typically contains oligonucleotide primers and a DNA polymerase in a suitable buffer. A “PCR reaction mixture” typically contains oligonucleotide primers, a DNA polymerase (most typically a thermostable DNA polymerase), dNTPs, and a divalent metal cation in a suitable buffer. A “CFPS reaction mixture” typically contains a crude or partially-purified yeast extract, an RNA translation template, and a suitable reaction buffer for promoting cell-free

protein synthesis from the RNA translation template. In some aspects, the CFPS reaction mixture can include exogenous RNA translation template. In other aspects, the CFPS reaction mixture can include a DNA expression template encoding an open reading frame operably linked to a promoter element for a DNA-dependent RNA polymerase. In these other aspects, the CFPS reaction mixture can also include a DNA-dependent RNA polymerase to direct transcription of an RNA translation template encoding the open reading frame. In these other aspects, additional NTP's and divalent cation cofactor can be included in the CFPS reaction mixture. A reaction mixture is referred to as complete if it contains all reagents necessary to enable the reaction, and incomplete if it contains only a subset of the necessary reagents. It will be understood by one of ordinary skill in the art that reaction components are routinely stored as separate solutions, each containing a subset of the total components, for reasons of convenience, storage stability, or to allow for application-dependent adjustment of the component concentrations, and that reaction components are combined prior to the reaction to create a complete reaction mixture. Furthermore, it will be understood by one of ordinary skill in the art that reaction components are packaged separately for commercialization and that useful commercial kits may contain any subset of the reaction components of the invention.

Preparation of active yeast cellular extract using scalable techniques

[070] Because CFPS exploits an ensemble of catalytic proteins prepared from the crude lysate of cells, the cell extract (whose composition is sensitive to growth media, lysis method, and processing conditions) is the most critical component of extract-based CFPS reactions.

[071] Yeast extracts for CFPS platforms disclosed herein can be prepared in a variety of ways. **FIG. 1A** depicts a scheme for making yeast extracts using a prior art procedure based upon the method of Sarnow and coworkers (Iizuka *et al.* (1994); Iizuka & Sarnow (1997)). **FIG. 1B** depicts the innovative method disclosed herein for producing extracts. The disclosed method can include three steps: (1) expanding a yeast cell culture in a bioreactor; (2) performing mechanical lysis of the cells by high-pressure homogenization; (3) performing a buffer exchange to generate the resultant extracts for the CFPS platform. Tangential flow filtration can be used to generate the resultant extract, where CFPS platforms are prepared on a large-scale process in industry. In most cases, however, dialysis is preferred in part for ease of use where CFPS platforms are prepared on a smaller-scale process in the laboratory.

[072] The composition of the cellular machinery at the time of harvest can directly affect the CFPS potential of the crude extract. Historically, yeast cells used for cell-free translation experiments have been harvested in early exponential phase (for examples, 1.5 OD₆₀₀ or 3-5 OD₆₀₀). Recovery of yeast cells during growth at mid-exponential to late-exponential phase (for example, a range from about 6 OD₆₀₀ to about 18 OD₆₀₀) can provide surprising benefits for translation using yeast extract-based CFPS platforms. For example, source cells for the yeast extracts disclosed herein can be obtained from mid-exponential to late-exponential batch cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀ or fed-batch cultures harvested in mid-exponential to late-exponential phase. Since the cells are rapidly dividing in this phase, they have a highly active translation machinery. Moreover, from a scaling standpoint, the ability to harvest at a later optical density can allow for larger cell mass recovery per fermentation, thereby leading to a larger volume of total crude extract prepared per fermentation for improved overall system economics. Typically, 1 L of cell culture yields about 6 g of wet cell mass when harvested at 12 OD₆₀₀ compared to ~1.5 g of wet cell mass when harvest at 3 OD₆₀₀. Subsequently, 1 g of wet cell mass leads to ~2 mL of crude extract.

[073] Yeast culturing techniques and culture media are well known in the art. Exemplary yeast culture media include YPD media (yeast extract (10 g/l), bacto-peptone (20 g/l; Difco) and dextrose (20 g/l), adjusted to pH5.5) and YPAD media (yeast extract (10 g/l), bacto-peptone (20 g/l; Difco), dextrose (20 g/l) and adenine hemisulfate (30 mg/l), adjusted to pH5.5). For *Saccharomyces cerevisiae* cellular extracts prepared from the mid-exponential to late-exponential cultures having a range of about 6 OD₆₀₀ to about 18 OD₆₀₀, the yeast cells were cultured in YPAD media. Other yeast culture media, including variations of YPD and YPAD, as well as synthetic dextrose, which is composed of 6.7 g L⁻¹ Yeast Nitrogen Base (YNB) (Sigma-Aldrich, St. Louis, MO), 20 g L⁻¹ glucose and 50 mM potassium phosphate buffer, pH 5.5, and its variations, can be used to culture the source *Saccharomyces cerevisiae* cells for the preparation of the crude yeast extracts for the CFPS systems, platforms and reactions disclosed herein.

[074] Furthermore, a step of adding inorganic phosphate to the growth media can increase protein synthesis capability for extracts generated. Typically, cells can be grown in media containing any source of inorganic phosphate, such as potassium phosphate, sodium

phosphate, magnesium phosphate, calcium phosphate, among others, including mixed metal phosphates (for example, sodium potassium phosphate). Preferred concentrations of inorganic phosphate range from about 15 mM to about 250 mM, including about 50 mM, about 75 mM, about 100 mM, about 125 mM and about 150 mM, among other concentrations within this range. Without the claimed subject matter being bound to any particular theory, the addition of phosphate to the growth media can reduce phosphatase expression in the cells during growth that can stabilize nucleoside triphosphate (NTP) concentrations during the cell-free reaction using cellular extracts prepared from such cells.

CFPS reaction conditions for translation-only reactions with yeast S30 and S60 extracts

[075] The ionic composition and temperature can have a profound effect on the efficiency and robustness of many protein-nucleic acid interactions and the proper function of protein biosynthesis. Altering the physicochemical environment of the CFPS reaction to better mimic the cytoplasm can improve protein synthesis activity. The following parameters can be considered alone or in combination with one or more other components to improve robust CFPS reaction platforms based upon crude yeast cellular extracts (for examples, yeast S30 and S60 extracts).

[076] The temperature may be any temperature suitable for CFPS. Temperature may be in the general range from about 10° C to about 40° C, including intermediate specific ranges within this general range, include from about 15° C to about 35° C, from about 15° C to about 30° C, from about 15° C to about 25° C. In certain aspects, the reaction temperature can be about 15° C, about 16° C, about 17° C, about 18° C, about 19° C, about 20° C, about 21° C, about 22° C, about 23° C, about 24° C, about 25° C. Preferably, the reaction temperature can be about 21° C.

[077] The CFPS reaction can include any organic anion suitable for CFPS. In certain aspects, the organic anions can be glutamate, acetate, among others. In certain aspects, the concentration for the organic anions is independently in the general range from about 0 mM to about 200 mM, including intermediate specific values within this general range, such as about 0 mM, about 10 mM, about 20 mM, about 30 mM, about 40 mM, about 50 mM, about 60 mM, about 70 mM, about 80 mM, about 90 mM, about 100 mM, about 110 mM, about 120 mM, about 130 mM, about 140 mM, about 150 mM, about 160 mM, about 170 mM, about 180 mM, about 190 mM and about 200 mM, among others.

[078] The CFPS reaction can also include any halide anion suitable for CFPS. In certain aspects the halide anion can be chloride, bromide, iodide, among others. A preferred halide anion is chloride. Generally, the concentration of halide anions, if present in the reaction, is within the general range from about 0 mM to about 200 mM, including intermediate specific values within this general range, such as those disclosed for organic anions generally herein.

[079] The CFPS reaction may also include any organic cation suitable for CFPS. In certain aspects, the organic cation can be a polyamine, such as spermidine or putrescine, among others. Preferably polyamines are present in the CFPS reaction. In certain aspects, the concentration of organic cations in the reaction can be in the general about 0 mM to about 3 mM, about 0.5 mM to about 2.5 mM, about 1 mM to about 2 mM. In certain aspects, more than one organic cation can be present.

[080] The CFPS reaction can include any inorganic cation suitable for CFPS. For example, suitable inorganic cations can include monovalent cations, such as sodium, potassium, lithium, among others; and divalent cations, such as magnesium, calcium, manganese, among others. In certain aspects, the inorganic cation is magnesium. In such aspects, the magnesium concentration can be within the general range from about 1 mM to about 50 mM, including intermediate specific values within this general range, such as about 1 mM, about 2 mM, about 3 mM, about 5 mM, about 6 mM, about 7 mM, about 8mM, about 9mM, about 10mM, among others. In preferred aspects, the concentration of inorganic cations can be within the specific range from about 4 mM to about 9 mM and more preferably, within the range from about 5 mM to about 7 mM.

[081] The CFPS reaction includes NTPs. In certain aspects, the reaction use ATP, GTP, CTP, and UTP. In certain aspects, the concentration of individual NTPs is within the range from about 0.1 mM to about 2 mM.

[082] The CFPS reaction can also include any alcohol suitable for CFPS. In certain aspects, the alcohol may be a polyol, and more specifically glycerol. In certain aspects the alcohol is between the general range from about 0% (v/v) to about 25 % (v/v), including specific intermediate values of about 5% (v/v), about 10% (v/v) and about 15% (v/v), and about 20 % (v/v), among others.

[083] The CFPS reaction preferably includes glutamate salts, NTPs, spermidine, putrescine, glycerol and magnesium.

Activating combined transcription and translation reactions in the CFPS platform.

[084] The present disclosure provides a novel transcription/translation system to circumvent the disadvantages associated with prior art eukaryotic CFPS platforms that rely exclusively upon exogenous RNA translation templates generated in separate costly and inefficient *in vitro* transcription reactions. According to one aspect, the present invention seeks to activate combined transcription and translation (Tx/Tl) in a one-pot reaction. The advantages of the combined transcription/translation (Tx/Tl) system in the CFPS platform include the following observations: (i) eliminates an extraneous processing step (*in vitro* transcription); (ii) removes the dependence of the reaction on the costly and potentially inhibitory m⁷G(5')ppp(5')G RNA cap structure analog; (iii) eliminates inconsistency issues with the capping reaction, a known problem for eukaryotic CFPS reactions (Takai *et al.* (2010)); and (iv) improves overall yields ~2-fold over cell-free translation only reactions when using a linear DNA template.

[085] Consequently, a combined Tx/Tl system is preferred for the disclosed CFPS reactions. This leads to considerations of template requirements for carrying out combined transcription and translation.

(1) Translation template considerations

(a) Optimization of 5'-UTR elements

[086] The CFPS strategy can leverage the use of certain translational elements in the untranslated region 5' (5'-UTR) of the open reading frame to be translated. In particular, a preferred translational element to include in the 5'-UTR include Internal Ribosome Entry Site (IRES) elements or cap-independent translation enhancer sequences to initiate translation. Such sequence elements can circumvent the need to utilize 5'-capped mRNA templates for efficient protein translation in the CFPS platforms.

[087] To evaluate cap-independent translation initiation, 15 µL batch cell-free translation-only reactions were carried out at 24° C for 1 h using different expression template constructs (see, for example, **FIG. 2**). These reactions were charged with 0.3 pmol purified *in vitro* transcribed luciferase mRNA having a 90-mer poly(A) tail. Initially, several yeast native internal ribosome entry site (IRES) sequences were evaluated, including the 5'-UTR of TFIID, HAP270, and YAP1 genes (Iizuka *et al.* (1994)), as well as the 5'-UTR of

the gene TIF4631 that is the yeast homolog of the mammalian translation initiation factor eIF4G (the mRNA is called p150; *see* Verge, V. *et al.*, “Localization of a promoter in the putative internal ribosome entry site of the *Saccharomyces cerevisiae* TIF4631 gene,” *RNA* 10, 277-286 (2004); Edwards, S. R., and Wandless, T. J., “Dicistronic regulation of fluorescent proteins in the budding yeast *Saccharomyces cerevisiae*,” *Yeast* 27, 229-236 (2010)).

[088] As a control, non-capped mRNAs harboring these different cap-independent translation leader sequences placed upstream of luciferase gene were compared to capped mRNA. Compared to capped luciferase mRNA, 5'-UTRs of HAP4 and TFIID showed low activities, while YAP1 and p150 did not direct translation (**Table 1**).

[089] Certain non-native, viral cap-independent sequences were next considered. The Ω leader sequence (also referred to herein as “ Ω leader,” “ Ω sequence,” “ Ω ” or “ Ω cap-independent translation enhancer”) showed surprisingly high activity among all tested cap-independent translation sequences, outperforming the capped mRNA by almost 2-fold (**Table 1**). As the next best leader sequence, the polyhedrin 5'-UTR sequence was ~17% as efficient in initiating translation as capped mRNA (**Table 1**). The species-independent translational mRNA sequence A64pA90 showed a low efficiency of translation initiation. Finally, the intergenic region (IGR) IRES from cricket paralysis virus (CrPV), which initiates translation in yeast cells without initiation factors, was examined. Unfortunately, when compared to capped mRNA and mRNA harboring the Ω sequence, the CrPV IRES showed little activity in disclosed CFPS platform assay (**Table 1**). Thus, certain non-coding IRES elements function better than others in the disclosed CFPS platform for initiating combined Tx/Tl, wherein the Ω sequence proved superior to the other tested IRES elements or cap-independent translation enhancer sequences.

[090] **Table 1. The efficiency of cap-independent and IRES mediated yeast cell-free translation as compared to capped mRNA.**

Template ¹	Efficiency ²
CappA90 (SEQ ID NO: 1)	100 ± 20%
Ω pA90 (SEQ ID NO: 9)	187 ± 13%
HedrinpA90 (SEQ ID NO: 8)	17 ± 4%
A64pA90 (SEQ ID NO: 7)	1.5 ± 0.2%
IGRpA90 (SEQ ID NO: 6)	<1 ± 0.02%
TFIIDpA90 (SEQ ID NO: 3)	<1 ± 0.07%

Template ¹	Efficiency ²
YAP1pA90 (SEQ ID NO: 5)	<1 ± 0.03%
p150pA90 (SEQ ID NO: 4)	<1 ± 0.08%
N5UpA90 (SEQ ID NO: 22)	2±0.3%

¹All templates include a luciferase coding sequence having a poly(A) tail of 90 nucleotides (pA90) (see **FIG. 2**). Cap – capped message. All other abbreviations described in the text.

²Efficiencies are normalized relative to the capped message (CappA90).

[091] The 5'-UTRs from tobacco etch virus (TEV) and Crucifer-infecting tobamovirus (Tbm) were also evaluated. The TEV 5'-UTR showed ~5% lower activity than that of the Ω sequence; the activity of Tbm 5'-UTR is half of the Ω sequence (**FIG. 3A**).

[092] Additionally, introduction of Kozak sequence elements in the 5'-UTR can lead to improved translation of expression templates. As shown in **FIG. 3B**, about a 40% improvement can be realized in sfGFP synthesis yields when different forms of the *S. cerevisiae* Kozak sequence are included in the 5'-UTR of expression templates.

(b) Optimization of 3'-UTR elements

[093] Furthermore, the sequences found in the untranslated region 3' (3'-UTR) to the open reading frame to be translated can also affect translation. In particular, sequences that include poly(A)_n tail can interact with Poly(A)-Binding Protein (PABP) to enhance protein synthesis and can promote enhance stability. The 3'-terminal poly(A)_n sequence can include different lengths of adenosine residues, where n can range from about 20 to about 200. Different lengths of poly(A) tail were evaluated in the 3'-UTR of luciferase RNA transcripts for its ability to support efficient protein synthesis, wherein the 3'-terminal poly(A)_n sequence had n of 25 nt, 50 nt, 90 nt and 170 nt. The length of poly(A) tail was optimized; 50 nt and 170 nt showed similar activities, while those of 90 nt and 25 nt showed 1.5~2-fold decrease in activity by comparison (**Fig. 3C**).

[094] Notably, the poly(A) tail is essential for yeast CFPS. Without the poly(A) tail, luciferase synthesis is decreased to 8.3% of the complete template (**Fig. 3D**). After deletion of Ω , the protein yield decreased to 1.7%, with the protein yield decreased to 0.1% when missing both Ω and poly(A)90 tail (**Fig. 3D**).

[095] Finally, the contribution of various 3'-UTRs (for examples, SEQ ID NOs: 97, 98, 104, 106, 107 and 110) in combined Tx/Tl was investigated. As compared to the no 3'-UTR control, the protein yields of various 3'-UTRs changed slightly (86%-136%) (**Fig. 3E**). Notably, the overall system performance decreases in these experiments because the templates shown in **FIG. 3E** do not have poly(A) tails. These results suggest that 3'-UTRs in yeast CFPS have less functional importance than in other CFPS systems, such as the wheat germ CFPS.

(c) Optimization of physiological solutes in combined transcription-translation reactions with yeast S60 extract

[096] The combined Tx/Tl reactions from plasmid vectors equipped with the Ω leader sequence were specifically optimized using a yeast S60 extract as the CFPS platform. Specifically, a series of optimization experiments were conducted to explore the effect of temperature, DTT concentration, DNA template concentration, magnesium concentration, and nucleotide concentrations on batch Tx/Tl reactions. Excepting temperature, these variables were selected because they were newly required for the combined Tx/Tl system, as opposed to the translation only reactions described above. Notably, these variables are also interdependent, as has been observed before in the development of crude extract based CFPS systems. Here, trends for the aforementioned optimizations are disclosed with only a single variable deviating from the finalized solute concentrations as reported in Examples.

[097] **FIG. 4A** shows active luciferase yield throughout the duration of a batch combined Tx/Tl reaction, monitored by samples taken at 0, 0.5, 1, 1.5, and 2 hours. The synthesis of luciferase shows slight lag during the first 0.5 h and then progresses linearly from 0.5-1.5h, with reaction termination occurring by 2 hours. The lag in the first 30 minutes is possibly associated to a delay resulting from combining transcription and translation, and has been previously observed. The temperature optimum was observed to be approximately 24° C (**FIG. 4B**). The combined Tx/Tl system was insensitive to template concentration above 3.2 nM, reaching saturation by ~5.3 nM (**FIG. 4C**). The system was also insensitive to DTT concentration over a range of 2-7 mM (**FIG. 4D**), suggesting that the S60 extract has a comparatively low potential of oxidation. This is important because the T7 RNA polymerase used to drive transcription in these particular examples requires a reducing environment for maximal activity.

[098] Unique insights about the interdependence of the magnesium and nucleotide concentrations were obtained through the following optimization experiments. The four nucleotide triphosphates play a role in both transcription and translation, yet CTP and UTP were not present in initial, translation-only reactions. Since ATP is maintained through the creatine phosphate secondary energy system, the concentration of the other 3 types of nucleoside triphosphates, GTP, UTP, and CTP (abbreviated “GUC”) were adjusted. Importantly, this required a change in magnesium concentration. It is well known that nucleotide concentration has a strong buffering/chelation effect on magnesium concentration and that optimal magnesium concentration is necessary for highly active CFPS. Thus, increasing the total nucleotide pool from 1.7 mM total (ATP, GTP only) to 7.5 mM total (ATP, GTP, CTP, UTP) when shifting from cell-free translation-only reactions (as described above) to combined Tx/Tl reactions, also required higher concentrations of magnesium (an increase of 1 mM to 6 mM). The optimal concentration of GUC with different magnesium concentrations was also investigated (**Table 2**). The maximum protein synthesis yield occurred when using 12 mM magnesium and 3.5 mM GUC.

[099] **Table 2. Optimization of [Mg]-[GUC]**

GUC (mM/each)	Percent yield (%) ¹				
	Mg (mM)				
	6	8	10	12	14
2	<i>100 ± 13.5</i>	120 ± 0.5	105 ± 4.4	81 ± 10.3	55 ± 0.1
3	48 ± 5.8	121 ± 5.8	126 ± 8.2	100 ± 1.1	82 ± 4.4
3.5	3 ± 0.0	93 ± 2.1	119 ± 15.8	128 ± 3.1	102 ± 4.5
4	0 ± 0.0	49 ± 6.3	92 ± 31.2	115 ± 22.4	110 ± 4.5

¹All reactions were performed using standard combined cell-free transcription-translation conditions except magnesium glutamate (Mg) and GTP/UTP/CTP (GUC) concentrations were varied. The luminescence value of standard reaction (6 mM Mg, 2 mM/each GUC) was counted as 100 (Bold and italic value). The luminescence values of all other samples were represented as the ratio to standard reaction. All values are the average of 3 individual reactions with standard deviations shown. The concentrations of Mg do not include Mg from S60 extract (see Examples). The values (Bold font) of highlighted the comparatively high-yield samples among different concentrations of Mg and GUC. The expression construct used in these experiments corresponds to SEQ ID NO: 9.

[0100] Notably, the highest yielding samples occurred along the diagonal of **Table 2**, indicating that unbalanced concentrations of magnesium and GUC significantly reduced the protein yield. With an eye towards ultimately developing a cost-effective CFPS system, the ~25% increase in yield with 150% additional nucleotide was insufficient motivation to keep the higher nucleotide concentrations, given the cost increase. As compared to cell-free translation alone, the newly designed combined Tx/Tl system improved overall protein synthesis yields more than 2-fold (up to 7 $\mu\text{g mL}^{-1}$). More importantly, it eliminated inconsistency issues with the capping reaction and further removed the dependence of the reaction on the costly and potentially inhibitory m⁷GpppG RNA cap structure analog.

(2) Transcription template considerations

[0101] For high-level and high-throughput expression of protein libraries using CFPS platforms, the use of linear transcription templates as the source of DNA is preferred. This advantage allows CFPS reactions to be primed with DNA transcription templates that can be prepared by an amplification reaction (for example, polymerase chain reaction (PCR)) without performing laborious cloning steps and without the need to use a circular DNA transcription template. A two-step overlap PCR method has been developed here that can be used to prime the yeast based CFPS reactions in a high-throughput fashion.

[0102] Referring to **FIG. 5**, in the first PCR reaction (PCR 1), the genes of interest are amplified using gene-specific primers complementary to the gene of interest with overhang regions for the second PCR reaction (PCR 2). The primers for PCR 1 can be comprised of short oligonucleotides (for example, about 30 nt to about 40 nt in length) that include a sequence having a length typically from about 15 nt to about 20 nt that can hybridize to the gene of interest and having about 15 nt to about 20 nt of non-gene-specific, unique sequences for secondary primer hybridization during PCR 2 (**FIG. 5**; for example, SEQ ID NOS: 25 and 26 for the luciferase gene (SEQ ID NO: 24)). PCR 2 can then be conducted with secondary primers that include a sequence complementary to the non-gene-specific, unique sequences found in the PCR 1 primers and also a sequence for a polymerase promoter element and Ω IRES element on the forward primer and a poly(A)₅₀ tail complement on the reverse primer (**FIG. 5**; for example, SEQ ID NOS: 27 and 28).

[0103] Preferred polymerases for use in the combined transcription/translation CFPS platform disclosed herein can be any polymerase that supports *in vitro* transcription in the

yeast CFPS platform extract and reaction. Examples of suitable polymerases include *E. coli* RNA Polymerase, T3 RNA Polymerase, T7 RNA Polymerase and SP6 RNA Polymerase, among others. Phage RNA polymerases, such as T3, T7 and SP6 RNA Polymerases, are generally preferred for use in the yeast CFPS platform reaction disclosed herein, owing to the small size of the promoter element for these polymerases (for example, 17-20 nucleotides in length). Such polymerases are particular preferred when linear DNA transcription templates are generated by DNA amplification methods, because the polymerase promoter sequence must be included in the primer of PCR2 to generate the transcription unit having the promoter 5' to the gene encoding the open reading frame to be transcribed into RNA and subsequently translated as protein. In one aspect, T7 RNA Polymerase-directed synthesis of RNA in the combined transcription/translation system of the CFPS platform is preferred. Following construction according to this aspect, the entire linear template can be amplified using a universal T7 primer as one of the primers. Similar approaches can be used with linear templates that contain a promoter sequence specific for a different polymerase.

[0104] After construction of linear templates, CFPS reactions were conducted using three different DNA templates: (i) PCR product amplified directly from plasmid; (ii) assembled linear DNA template produced by the two-step overlap PCR procedure described above; and (iii) circular plasmid. All three DNA templates can successfully synthesize luciferase, but surprisingly, the linear DNA templates performed ~40-60% better than the plasmid (**FIG. 6A**). To demonstrate utility of the approach, two other proteins, GFP and CAT, were also expressed using DNA templates assembled by the disclosed two-step PCR method. By comparison to commercial standard proteins, the yields of active protein are estimated as ~7 to 12.5 µg/ml (**Table 3**).

[0105] Table 3. Yield of active proteins from yeast CFPS programmed with linear DNA templates.¹

Protein	Yield ($\mu\text{g mL}^{-1}$ active protein)
Luciferase (SEQ ID NO: 30)	7 ± 2
GFP (SEQ ID NO: 33)	12.5 ± 2.5
CAT (SEQ ID NO: 36)	10 ± 1

¹Luciferase (SEQ ID NO: 30) was prepared from coding sequence corresponding to SEQ ID NO: 24 and expression construct corresponding to SEQ ID NO: 29; green fluorescence protein (GFP; SEQ ID NO: 33) was prepared from coding sequence corresponding to SEQ ID NO: 31 and expression construct corresponding to SEQ ID NO: 32; and chloramphenicol acetyltransferase (CAT; SEQ ID NO: 36) was prepared from coding sequence corresponding to SEQ ID NO: 34 and expression construct corresponding to SEQ ID NO: 35.

[0106] The solubility of each protein is also demonstrated in [³⁵S]-methionine autoradiography (**FIG. 6B**). Notably, more than 95% of the total protein produced was soluble in all cases. Overall, high-throughput combined Tx/Tl method enables one having ordinary skill in the art to proceed from DNA sequence to protein in less than six hours.

[0107] Due to the open architecture platform of both PCR and CFPS reactions, this process can be easily automated for high-throughput protein expression. For example, a reaction module containing programmed reagent additions and thermally-controlled reaction vessel incubation protocols can be used to generate linear transcription templates via amplification using PCR1 and PCR2 in a single-pot reaction. Once the linear transcription templates are prepared by amplification, the PCR mixtures can be processed to recover the linear transcription templates for use in the CFPS reaction in a separate module. Optionally, the PCR mixtures can be used directly in the CFPS reaction in a separate module without performing the processing step to recover the linear transcription templates. This latter option may be preferred in cases where the PCR mixture contribution to the CFPS reaction mixture is sufficiently small to not interfere with combined transcription/translation in the CFPS reaction.

Ability to express a variety of proteins

[0108] The developed cell-free protein synthesis platform has also demonstrated utility to express a variety of different proteins, including reporter proteins such as firefly luciferase, SuperFolder green fluorescent protein, and chloramphenicol acetyltransferase, as well as therapeutically relevant proteins such as single-chain antibody variable fragments (scFvs) and virus-like particles (VLPs) (**Table 4**). A major advantage of the disclosed yeast CFPS platform compared to prior art platforms is the expression of proteins in soluble and active form. Nearly 85-100% of complex proteins expressed (scFv and luciferase) are in their soluble form, compared to *E. coli* based CFPS where only ~15 % of these same proteins synthesized are soluble. Additionally, the reaction is capable of synthesizing proteins for up to 6 h (**FIG. 7**), and this is the longest yeast CFPS batch synthesis disclosed.

[0109] **Table 4. Exemplary protein synthesis yields with the yeast CFPS system**

Protein ¹	Molecular Weight (kDa)	Yield(s) (μg/mL)
Firefly Luciferase (SEQ ID NO: 30)	61	12.4 ± 1.00 active protein
Chloramphenicol Acetyltransferase (SEQ ID NO: 36)	25	10.00 ± 1.00 active protein
Superfolder Green Fluorescent Protein (SEQ ID NO: 39)	27	42.99 ± 0.38 active protein
ANX scFv (SEQ ID NO: 42)	27	4.97 ± 0.05 total protein 4.96 ± 0.13 soluble protein
BOT scFv (SEQ ID NO: 45)	28	5.42 ± 0.03 total protein 4.83 ± 0.07 soluble protein
MS2 scFv (SEQ ID NO: 48)	28	2.04 ± 0.04 total protein 1.94 ± 0.06 soluble protein
MS2-A scFv (SEQ ID NO: 51)	28	1.41 ± 0.00 total protein 1.25 ± 0.01 soluble protein
2E2 scFv (SEQ ID NO: 54)	26	1.03 ± 0.01 total protein 0.91 ± 0.0. soluble protein

Protein ¹	Molecular Weight (kDa)	Yield(s) (µg/mL)
2E2-3d scFv (SEQ ID NO: 57)	26	1.60 ± 0.02 total protein 1.50 ± 0.02 soluble protein
Human Papillomavirus 16 L1 (SEQ ID NO: 60)	51	3.26 ± 0.04 total protein 2.99 ± 0.03 soluble protein

¹Corresponding source nucleic acid sequences encoding the open reading frame (ORF) information (SEQ ID NOs: 24, 34, 37, 40, 43, 46, 49, 52, 55 and 58, respectively) were used for preparing linear templates for the expression constructs (SEQ ID NOs: 29, 35, 38, 41, 44, 47, 50, 53, 56 and 59, respectively) to express the identified proteins in the yeast CFPS system.

[0110] The present yeast crude extract preparation method and the physicochemical environment of the *in vitro* system provide substantial advantages for protein synthesis. Overall, one can increase protein synthesis yield ~250-fold to $7.69 \pm 0.53 \mu\text{g mL}^{-1}$ and increased relative product yield ~2000-fold to 0.39 µg protein synthesized per \$ reagent cost using the disclosed yeast CFPS platform relative to the prior art yeast extract CFPS method (**FIG. 8**).

[0111] The improvements disclosed herein have implications for using yeast CFPS as a model to study translation. Referring to **Table 5**, both active protein synthesis of luciferase (SEQ ID NO: 30) from a linear luciferase expression template (SEQ ID NO: 29) using (i) extract derived from the prior art “benchtop” protocol (*e.g.*, glass beads lysis) for combined Tx/Tl reactions and (ii) extract derived from optimized extract preparation protocol (“Innovation” in **Table 5**) for translation only reactions was improved.

[0112] Table 5: Relative comparison of CFPS methods

Extract Preparation Method	MNase	Combined¹/ Separated² Tx/Tl	CFPS Reaction Conditions	Active Luciferase Yield (µg/mL)	Fold Increase Relative to Prior Art MNase Treated Tx/Tl Reaction³
Prior Art	Yes	Separated	Prior Art	0.03 +/- 0.00	1
Prior Art	No	Combined	Innovation	1.47 +/- 0.24	47
Innovation	No	Separated	Innovation	2.41 +/- 0.03	77
Innovation	No	Combined	Innovation	7.69 +/- 0.53	245

¹Combined Tx/Tl means that transcription and translation occur together in a one-pot reaction.

²Separated Tx/Tl means that transcription of mRNA and capping was performed in a separate reaction prior to cell-free translation.

³Fold increase is normalized relative to the active luciferase yield obtained for extracts prepared by the prior art method using CFPS reactions conditions of the prior art (set at a value of 1).

[0113] For example, by altering the reaction conditions and taking advantage of the developed Tx/Tl method, a 47-fold increase in active protein synthesis over prior art methods was realized. Translation only reactions primed with extract generated with the optimized protocol developed in this disclosure show a similar increase with a 77-fold improvement over prior art methods. Furthermore, in both cases the yields crested the 1 µg mL⁻¹ threshold and demonstrate a significant improvement in utility of the cell-free system.

EXAMPLES

Example 1. Strains and reagents

[0114] Yeast strains MBS and S288c were used. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. DNA polymerase, T4 polynucleotide kinase, T4 DNA ligase, and restriction endonucleases were purchased from New England Biolabs (Ipswich, MA). T7 polymerase was prepared in lab (following the protocol developed by Swartz, J.R. *et al.*, "Cell-free protein synthesis with prokaryotic combined

transcription-translation,” *Methods in molecular biology* (Clifton, N.J.) 267, 169-182 (2004)). Plasmids were extracted using Omega Kits (Omega Bio-Tek, Norcross, GA). All DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa).

Example 2. Nucleic acid manipulations

[0115] The schematic structures of exemplary expression templates described below are listed in part in **FIG. 2**. All nucleotide and amino acid sequences and their corresponding SEQ ID NOs are compiled in the **SEQUENCE TABLE** at the end of the **Examples** section. The luciferase-coding region was amplified from pK7LUC plasmid (SEQ ID NO: 61) (Jewett, M.C. *et al.*, “In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation,” *Mol. Syst. Biol.* 2013, 9, 678) using primers pET23LucA-f (SEQ ID NO: 62) and pET23LucA-r (SEQ ID NO: 63), and inserted into pET23c plasmid (SEQ ID NO: 64) with NdeI and XhoI sites to construct plasmid pET23LucA (SEQ ID NO: 65). Poly(A) tails with 25nt and 50nt were introduced into pET23LucA plasmid to replace 90nt poly(A) using the primer pairs: PolyA-f (SEQ ID NO: 66)/PolyA25-r (SEQ ID NO: 67) and PolyA-f (SEQ ID NO: 66)/PolyA50-r (SEQ ID NO: 68). We were unable to produce a plasmid with a correctly identified 170nt poly(A) due to the poor efficiency of PCR and sequencing for a 170nt poly(A) region. However, we introduced a 170nt poly(A) into linear DNA template by PCR using the primer PolyA-f (SEQ ID NO: 66) with the primer PolyA170-r (SEQ ID NO: 69) purchased from IDT with PAGE purification. Three yeast native internal ribosome entry site sequences, the 5'-UTR of TFIID, HAP270, and YAP1 genes (Iizuka *et al.* (1994)), were amplified from yeast genomic DNA using primer pairs TF5UTR-f (SEQ ID NO: 70)/TF5UTR-r (SEQ ID NO: 71), HAP270-f (SEQ ID NO: 72)/HAP270-r (SEQ ID NO: 73), and YAP1-f (SEQ ID NO: 74)/YAP1-r (SEQ ID NO: 75), respectively, and inserted into plasmid pET23LucA (SEQ ID NO: 65) between the T7 promoter and luciferase with NdeI and XbaI. These three plasmids are identified as pET23TFIIDLucA (SEQ ID NO: 76), pET23HAP270LucA (SEQ ID NO: 77), and pET23YAP1LucA (SEQ ID NO: 78). The 5'-UTR of p150 gene (SEQ ID NO: 79; Zhou, W., Edelman, G. M., and Mauro, V. P., “Transcript leader regions of two *Saccharomyces cerevisiae* mRNAs contain internal ribosome entry sites that function in living cells,” *Proc. Natl. Acad. Sci. U.S.A.* 98, 1531-1536 (2001)) was amplified from yeast genomic DNA using primers P150-f (SEQ ID NO: 80) and P150-r (SEQ ID NO: 81) flanking with XbaI and BamHI sites, and inserted into pET23LucA

plasmid (SEQ ID NO: 65) where NdeI had been replaced with BamHI, since the insert fragment contains an NdeI site.

[0116] The Ω sequence (65nt) from TMV (SEQ ID NO: 82) was introduced into pET23LucA upstream of luciferase with primers Sf-f (SEQ ID NO: 83) and Omega-r (SEQ ID NO: 84). The 5'-UTR of polyhedrin gene (44nt) (SEQ ID NO: 85) was introduced into pET23LucA (SEQ ID NO: 65) with primers Sf-f (SEQ ID NO: 83) and Polyhedrin-r (SEQ ID NO: 86). A 5'-end poly(A)₆₄ sequence (SEQ ID NO: 87) was introduced into pET23LucA (SEQ ID NO: 65) with primers Sf-f (SEQ ID NO: 83) and PolyA₆₄-r (SEQ ID NO: 88). A 5'-UTR fragments (143 nt) (SEQ ID NO: 89) from tobacco etch virus (TEV) genome (Accession number: NC_001555) was cloned into pET23LucA (SEQ ID NO: 65) upstream of the luciferase gene by oligo TEV-r (SEQ ID NO: 90); another plant viral 5'-UTR fragment (65 nt) (SEQ ID NO: 91) from Crucifer tobamovirus (CfTbm) genome (Accession number: NC_003355.1) was inserted into pET23LucA (SEQ ID NO: 65) upstream of luciferase gene using oligo CfTbm-r (SEQ ID NO: 92). An IRES sequence of the cricket paralysis virus (CrPV) intergenic region (IGR) (SEQ ID NO: 93) was amplified from the plasmid pSall-IGR (SEQ ID NO: 94) (Deniz, N. *et al.*, "Translation initiation factors are not required for Dicistroviridae IRES function *in vivo*," *RNA* 15, 932-946 (2009)) using primers IGR-f (SEQ ID NO: 95) and IGR-r (SEQ ID NO: 96).

[0117] Two fragments (SEQ ID NOs: 97 and 98) were cloned from the 3'-UTR of yeast *FBA1* gene (Accession number: NM_001179626) that encodes fructose 1, 6-bisphosphate aldolase. The shorter fragment containing 662 nt from the first nucleotide after the stop codon (SEQ ID NO: 97) was amplified by primers FBA3UTR-f (SEQ ID NO: 99) and FBA3UTR1-r (SEQ ID NO: 100). The longer fragment containing 1465 nt from the first nucleotide after the stop codon (SEQ ID NO: 98) was amplified by primers FBA3UTR-f (SEQ ID NO: 99) and FBA3UTR2-r (SEQ ID NO: 101). The two 3'-UTR fragments were then placed after the stop codon of luciferase with XhoI and SacI sites. Two DNA fragments of TMV genome were synthesized (GenScript, Piscataway, NJ). The first fragment, TMV1, ranges from 4920 to 5711 of genome (792 nt in length) (SEQ ID NO: 102) containing the sequence between two open-reading frame TMVgp1 and TMVgp6; the second fragment, TMV2 ranges from 6192 to 6395 genome (204 nt in length) (SEQ ID NO: 103). Three fragments, TMV13U200 (SEQ ID NO: 104), TMV13U400 (SEQ ID NO: 105), and TMV13U700 (SEQ ID NO: 106), were amplified from TMV1 with the length of 200nt,

400nt, and 700nt respectively using primer pairs: TMV13U-f (SEQ ID NO: 107) and TMV13U200-r (SEQ ID NO: 108); TMV13U-f (SEQ ID NO: 107) and TMV13U400-r (SEQ ID NO: 109); and TMV13U-f (SEQ ID NO: 107) and TMV13U700-r (SEQ ID NO: 110), respectively. One fragment was amplified from TMV2 with the length of 204 nt (SEQ ID NO: 111) by using primer pairs TMV23U-f (SEQ ID NO: 112) and TMV23U-r (SEQ ID NO: 113). All four fragments amplified from TMV1 and TMV2 were placed after the stop codon of the luciferase-coding frame with XhoI and SacI sites as 3'-UTRs.

[0118] To assemble linear expression templates of luciferase, green fluorescence protein (GFP), and chloramphenicol acetyl transferase, the coding region of the 3 enzymes were amplified with primer pairs QEluc-f (SEQ ID NO: 114) and QEluc-r (SEQ ID NO: 115); QEGFP-f (SEQ ID NO: 116) and QEGFP-r (SEQ ID NO: 117); and QECAT-f (SEQ ID NO: 118) and QECAT-r (SEQ ID NO: 119), respectively. Therefore, T7 promoter (SEQ ID NO: 120), Ω sequence (SEQ ID NO: 121), and poly(A)50 tail (SEQ ID NO: 122) were overlapped to the coding region by primer pairs QET7Ome-f (SEQ ID NO: 123) and PolyA50-r (SEQ ID NO: 68), respectively.

[0119] Plasmid pET23c-GFP-cyc3 (SEQ ID NO: 124) was kindly provided by Dr. Markus Pech at the Max Planck Institute for Molecular Genetics (Iskakova M.B. *et al.*, "Troubleshooting coupled in vitro transcription-translation system derived from Escherichia coli cells: Synthesis of high-yield fully active proteins," *Nucleic Acids Res.* 34(19):e135 (2006)). The gene encoding for firefly luciferase (60,755 Da) (SEQ ID NO: 24) in place of GFP-cyc3 was inserted into the plasmid backbone as the reporter proteins. The luciferase gene was inserted between the T7 promoter (SEQ ID NO: 120) and T7 terminator (SEQ ID NO: 125) sequences using NdeI and XhoI restriction digest enzymes using the luciferase sense primer (SEQ ID NO: 126) and luciferase anti-sense primer (SEQ ID NO: 127). Encoded in the anti-sense primer was the addition of a 90-mer poly(A) tail to 3' end of the coding sequence. Plasmids were harbored in *E. coli* DH5 α competent cells and purified using E.Z.N.A. Omega Plasmid DNA Maxi Kits (Omega Bio-Tek, Norcross, GA). For combined transcription/translation (Tx/Tl) reactions, the Ω cap-independent translation enhancer from tobacco mosaic virus was inserted in the 5'-UTR to remove the dependence of the 5'-cap for translation initiation to construct the pET23 Ω LucA plasmid (SEQ ID NO: 128). For CFPS reactions, the coding sequence containing the Ω cap-independent translation enhancer, protein of interest, and polyA tail, was PCR amplified with backbone sense primer (SEQ ID

NO: 129) and anti-sense primer (SEQ ID NO: 130) using Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). The PCR product was purified using the Qiagen PCR Purification Kit (Qiagen, Valencia, CA).

Example 3. In vitro transcription

[0120] *In vitro* transcription and mRNA capping was performed with the Ambion mMessage mMachine® Kit (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. The capped mRNA was purified following a phenol-chloroform extraction and desalted using a Micro Bio-Spin® 6 chromatography column (Biorad, Hercules, CA). Non-capped RNA was prepared according to Mureev *et al.* (2009). Poly(A)_n-containing mRNA was isolated using Dynabeads Oligo(dT)₂₅ magnetic beads (Life Technologies, Grand Island, NY).

Example 4. Growth conditions

[0121] The cultures of yeast cells were grown to 3, 6, 9, 12, 15 and 18 OD₆₀₀ and used to prepare individual batches of crude extract originating from each of these fermentations. Thereafter, combined transcription and translation in a 15 µL batch reaction with these different extracts were carried out for a period of 2 h. The most active extracts were obtained from yeast culture harvested at mid-exponential phase: $1.34 \pm 0.25 \mu\text{g mL}^{-1}$ for OD₆₀₀ 6-12 compared to $0.32 \pm 0.05 \mu\text{g mL}^{-1}$ for OD₆₀₀ of 3 and $0.49 \pm 0.08 \mu\text{g mL}^{-1}$ for OD₆₀₀ 15-18 (**FIG. 9A, B**). Furthermore, the addition of inorganic phosphate to the growth media resulted in a 21% increase in protein synthesis capability for extracts generated from cells grown with 50 mM potassium phosphate (**FIG. 10**).

Example 5. Preparation of yeast S30 extract

[0122] For extract preparation, *S. cerevisiae* strain MBS (Thompson, S.R. *et al.*, "Internal initiation in *Saccharomyces cerevisiae* mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element," *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12972-12977 (2001)) was grown in either the BIOSTAT Cplus 10 L bioreactor (Sartorius Stedim Biotech S.A., Aubagne Cedex, France), or 1 L of culture in 2.5 L Tunair (Sigma-Aldrich, St. Louis, MO) shake flasks in YPAD media, pH 5.5 supplemented with 50 mM potassium phosphate to 12 OD₆₀₀. When the OD₆₀₀ reached 12, the yeast culture was cooled quickly to between 4-8° C by either harvesting through a stainless steel coil immersed in an ice bath (when grown in the

10 L bioreactor) or by adding 1 L of ice per L of yeast culture (when grown in the shake flasks). To pellet the cells, the cell suspension was centrifuged for 10 min at 3,000×g and 4° C. The cell pellet was washed with 60 mL of Mannitol Buffer A (30 mM HEPES, pH 7.4 w/5M KOH, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM dithiothreitol, 8.5% (w/v) mannitol) per liter of starting culture followed by centrifugation for 5 min at 3,000×g and 4° C. This step was repeated 3 times with 20 mL of Mannitol Buffer A, with the final wash centrifuged at 4,000×g for 5 min. Potassium and magnesium glutamate were used in place of potassium and magnesium acetate as denoted in the text. Extra buffer was removed by placing the centrifugation bottle upside down and tapping against a paper towel. The cell pellet was weighed, flash-frozen on liquid N₂, and stored at -80°C. Alternatively, the extract preparation can be continued with cellular lysis.

[0123] Cell lysis was performed using one of two methods, either with 0.5 mm glass beads (Sigma-Aldrich, St. Louis, MO) or high-pressure homogenization. For glass beads lysis, a prior art method was used (Iizuka *et al.* (1994)). Briefly, 5-6 g of wet cell mass was combined with 1.5 mL of cold Lysis Buffer A (Mannitol Buffer A + 0.5 mM PMSF) per 1 gram of wet cell mass in a 50 mL falcon tube and the suspension was thawed on ice. Note, PMSF was first dissolved in 100% ethanol and was added fresh before each use. In the cold room, the cells were lysed in capped 50 mL falcon tubes by five 1-min cycles of hand shaking (2 Hz) over a 50 cm hand path, with cooling on ice water for 1 min between cycles.

[0124] For high-pressure homogenization lysis, a minimum of 3 g of wet yeast cell pellet was lysed at one time. Again, 1.5 mL of cold Lysis Buffer A per 1 g of wet cell mass was added to the cell pellet and the suspension was thawed on ice. Immediately after cell thawing was completed, the cells were lysed by passing through an EmulsiFlex-C5 Homogenizer (Avestin, Ottawa, ON, Canada) at 30,000 psi and a flow rate of approximately 1-3 mL per minute. The sample was collected through a cooling coil immediately upon exit that was submerged in ice water.

[0125] After cell disruption, the lysate was centrifuged at 4°C and 25,000×g for 5 min. Immediately, the supernatant was transferred with a pipette into a clean Nalgene spherical bottom high-speed centrifuge bottle for the second centrifugation at 4° C 25,000×g for 5 min. The aqueous fraction was carefully removed by avoiding crude cell debris at the bottom.

[0126] The lysate subsequently underwent buffer exchange through either dialysis or fast protein liquid chromatography (FPLC). For dialysis, the extract was dialyzed against four exchanges of 200-volumes of Buffer A/PMSF (Lysis Buffer A without the addition of mannitol) for 30 min each at 4°C using Slide-A-Lyzer Dialysis Cassettes (2,000 Da MWCO; Thermo Fisher Scientific, Waltham, MA). The dialyzed extract was centrifuged at 12,000×g at 4°C for 20 min to remove any degraded proteins.

[0127] For FPLC, the extract was loaded onto a Sephadex G-25 Superfine (GE Healthcare Biosciences, Pittsburgh, PA) column at 25% of the bed volume using the BioLogic DuoFlow FPLC (Bio-Rad, Hercules, CA). The extract was exchanged against Buffer A/PMSF with a flow rate of 0.65 mL per min at 4°C. Fractions were collected in 0.5 mL volumes. All fractions with an A₂₆₀ reading greater than 0.45 after 200-fold dilution were pooled together.

[0128] After buffer exchange, the extract was immediately aliquoted into 50, 100 and 200 µL samples as desired. The aliquots were rapidly frozen in liquid nitrogen and stored at -80°C for long-term storage. No decrease in activity was seen after several months of storage at -80°C or after up to 4 freeze-thaw cycles.

Example 6. Micrococcal Nuclease pre-treatment

[0129] For Micrococcal Nuclease (MNase) pre-treatment, 1 µL of 50 mM CaCl₂ and 0.72 µL of 25 U µL⁻¹ Micrococcal Nuclease from *Staphylococcus aureus* (Sigma-Aldrich, St. Louis, MO; final concentration of 0.5 mM CaCl₂ and 0.18 U µL⁻¹ MNase) were added to 100 µL of crude extract on ice. The solution was mixed by pipetting up and down and the reaction was incubated at room temperature for 5 min. To quench the reaction, 1 µL of 250 mM EGTA was added to the reaction to a final concentration of 2 mM. The reaction was promptly mixed by pipetting up and down and the treated crude extract was placed back on ice and used for downstream CFPS reactions.

Example 7. Preparation of yeast S60 extract

[0130] Colonies of yeast strain were cultivated in rich media (2% peptone, 1% yeast extract, 2% glucose), shaking at 250 rpm at 30 °C overnight to saturation. The seeding culture was used to inoculate 1 L of fresh rich media with 1:1000 in 2.5 L Tunair (Sigma-Aldrich, St. Louis, MO), shaking at 250 rpm at 30°C. Cells were harvested at

mid-logarithmic phase (OD_{600} 10 - 12) by centrifugation at 3000g for 10 min. Cell pellets were resuspended and washed three times in Buffer A (20 mM HEPES-KOH pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate). The wet pellet was weighed and suspended by vortex in lysis buffer (20 mM HEPES-KOH pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM DTT, 0.5 mM PMSF) with 1 mL buffer per gram of wet cell weight. Cells were lysed using an Avestin EmulsiFlex-C5 High Pressure Homogenizer (Avestin, Ottawa, ON, Canada) one time under 30,000 psig. The lysate was centrifuged at 4°C and 30,000g for 30 min, the supernatant was removed, placed in a clean spherical bottom high-speed centrifuge bottle and clarified again. Supernatant was desalted using dialysis tubing (Spectra/Por 3 MWCO 3500, Spectrum Labs, Rancho Dominguez, CA) against four exchanges of 50-volumes of lysis buffer (20 mM HEPES-KOH pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM DTT, 0.5 mM PMSF) for 30 min each at 4°C. After dialysis, extract was centrifuged at 60,000g for 20 min at 4°C. Final extract was distributed into 100 μ L aliquots in 1.5-mL Eppendorf tubes, frozen in liquid nitrogen and stored at -80 °C. The protein concentration was determined using Quick StartTM Bradford Protein Assay (Bio-Rad Laboratories, Hercules, CA).

Example 8. Cell-Free Protein Synthesis using crude yeast extracts

[0131] CFPS reactions were carried out in 1.5 mL Eppendorf tubes at 21°C in a temperature-controlled water bath in 15 μ L reactions. The cell-free reaction mixture was assembled on ice from stock solutions to the following working concentrations for translation only reactions: 22 mM HEPES-KOH pH 7.4, 120 mM potassium glutamate, 2 mM magnesium glutamate, 0.75 mM adenosine triphosphate (ATP), 0.1 mM guanosine triphosphate (GTP), 0.04 mM of each of 20 amino acids, 25 mM creatine phosphate, 1.7 mM DTT, 1 mM putrescine, 0.5 mM spermidine, 0.27 mg mL⁻¹ creatine phosphokinase (from rabbit muscle; Sigma-Aldrich, St. Louis, MO), 26.7 U mL⁻¹ RNase Inhibitor (Qiagen, Valencia, CA), 600 ng *in vitro* transcribed mRNA, and 50% (v/v) yeast extract. For combined transcription and translation reactions the working concentrations varied slightly to: 22 mM HEPES-KOH pH 7.4, 120 mM potassium glutamate (unless otherwise noted), 5 mM magnesium glutamate (unless otherwise noted), 1.5 mM of each ATP, GTP, CTP and UTP, 0.08 mM of each of 20 amino acids, 25 mM creatine phosphate, 1.7 mM DTT, 1 mM putrescine, 0.5 mM spermidine, 0.27 mg mL⁻¹ creatine phosphokinase (from rabbit muscle, Sigma-Aldrich, St. Louis, MO), 26.7 U mL⁻¹ RNase Inhibitor (Qiagen, Valencia, CA), 250

ng Ω LucA PCR amplified DNA, 0.027 mg mL⁻¹ T7 RNA Polymerase (made in house following the protocol developed by Swartz *et al.* (2004) and 50% (v/v) yeast extract. The final concentration of yeast extract proteins was 25.7±1.0 mg mL⁻¹, as determined by Bradford Assay using commercially available assay reagents (Bio-Rad, Hercules, CA) compared to a bovine serum albumin protein standard. All other reagents were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. The amount of active firefly luciferase produced was determined by adding 12 μ L of CFPS sample to 30 μ L of ONE-Glo Luciferase Assay System (Promega, Madison, WI) in a white 96-well plate. The total luminescence was read every two minutes over a 20-minute interval using a BioTek (Winooski, VT) Synergy 2 plate reader. The maximum amount of relative light units (RLUs) was recorded for each cell-free reaction. RLUs were then compared to a linear standard curve of recombinant luciferase (Promega, Madison, WI) added directly to the ONE-Glo reaction mixture.

[0132] For the pre-incubation experiments, all soluble components of the cell-free reaction were initially assembled on ice except the crude extract, T7 polymerase (T7 Pol), and creatine phosphokinase (CK). For each individual reaction, the T7 Pol and CK were added to the cell-free reaction immediately before the extract had finished “pre-incubating”. After the pre-incubated extract was added to the reaction mixture, the CFPS reaction proceeded for an additional 2 h. In order to assay all of the samples simultaneously, the cell-free reactions were quenched by fast freezing in liquid nitrogen, stored at -20°C and thawed simultaneously before being assayed for active luciferase yield.

Example 9. Yeast cell-free translation only reactions using mRNA template using the yeast S60 extract

[0133] Yeast cell-free translation was prepared as described by Sarnow with some modifications (Iizuka & Sarnow (1997)). CFPS reactions were primed with 20 nM mRNA in 15 μ L reactions. The cell-free reaction mixture was assembled on ice from stock solutions to the following working concentrations: 25 mM HEPES-KOH pH 7.4, 120 mM potassium glutamate, 1 mM magnesium glutamate, 1.5 mM adenosine triphosphate (ATP), 0.2 mM guanosine triphosphate (GTP), 0.1 mM of each of 20 amino acids, 25 mM creatine phosphate, 1 mM DTT, 0.27 mg/mL creatine phosphokinase (C3755-1KU, Sigma), 200 U/mL RNase Inhibitor (Qiagen), and 50% (v/v) yeast S60 extract.

Example 10. Combined transcription/translation (Tx/Tl) cell-free protein synthesis using the yeast S60 extract

[0134] Combined cell-free transcription-translation reactions were carried out in 1.5-mL Eppendorf tubes in 15 μ L reactions. The reaction was primed with 3.2 nM PCR product. The cell-free reaction mixture was prepared on ice from stock solutions to the following working concentrations for translation only reactions: 25 mM HEPES-KOH pH 7.4, 120 mM potassium glutamate, 6 mM magnesium glutamate, 1.5 mM ATP, 2 mM of each GTP, CTP and UTP, 0.1 mM of each of 20 amino acids, 25 mM creatine phosphate, 2 mM DTT, 0.27 mg/mL creatine phosphokinase (C3755-1KU, Sigma), 200 U/mL RNase Inhibitor (Qiagen), 27 μ g/mL T7 RNA Polymerase, and 50% (v/v) yeast S60 extract. All combined cell-free transcription-translation reactions were performed using above conditions unless specified otherwise. For the analysis of [^{35} S]-methionine-labeled protein products, combined Tx/Tl cell-free protein synthesis was performed as described above except that [^{35} S]-methionine was supplemented with the final concentration of 0.58 μ M. The protein products were resolved by NuPAGE® Novex® 4-12% Bis-Tris Gels (Invitrogen, Grand Island, NY).

Example 11. Luciferase activity assay

[0135] The amount of active firefly luciferase was determined by ONE-GLO™ Luciferase Assay System (Promega), in a white 96-well plate. Five μ L of CFPS sample was added to 30 μ L of Luciferase Assay Buffer. Luminescence was read every 2 min over a 20 min period using a BioTek Synergy 2 plate reader (Winooski, VT).

Example 12. Chloramphenicol acetyl transferase assay

[0136] Active chloramphenicol acetyl transferase (Catherine, C. *et al.*, “Cell-free platforms for flexible expression and screening of enzymes,” *Biotechnol Adv* 31, 797-803 (2013)) was measured as previously described (Jewett, M.C. *et al.*, “An integrated cell-free metabolic platform for protein production and synthetic biology,” *Mol Syst Biol* 2008, 4 (2008)).

Example 13. Autoradiography

[0137] Autoradiography was used to determine the size of protein synthesized. 1.7 μ L of ^{35} S- Methionine (~18 μ Cu) (PerkinElmer, San Jose, CA) was added to each 15 μ L CFPS

reaction. Following 3 h incubation, the CFPS reaction was loaded onto a NuPAGE 4-12% Bis-Tris Gel (Life Technologies, Grand Island, NY) following the manufacturer's instructions. The NuPAGE gels were stained with SimplyBlue SafeStain (Life Technologies, Grand Island, NY) and all proteins present in the cell-free reaction were visualized using the Gel Doc XR+ (Bio-Rad, Hercules, CA). The gels were dried and exposed overnight on a Storage Phosphor Screen (GE Healthcare Biosciences, Pittsburgh, PA) and imaged with the Storm 860 Phosphoimager (GE Healthcare Biosciences, Pittsburgh, PA). This image was digitally compared to the SimplyBlue stained image that included a protein standard ladder to determine the length of synthesized proteins.

Example 14. Comparison of CFPS platforms

[0138] **FIG. 1A** depicts a scheme for making yeast extracts using a prior art procedure based upon the method of Sarnow and coworkers (Iizuka *et al.* (1994); Iizuka, N. & Sarnow, P. (1997)). **FIG. 1B** depicts the method disclosed herein for preparing yeast extracts. The crude extracts prepared by these two methods were evaluated by assessing the total active luciferase yield from batch cell-free translation reactions with capped *in vitro* transcribed luciferase RNA prepared according to Example 3. Translation of luciferase was carried out in a 15 μ L batch reaction for 2 hours at 21°C. Strikingly, extract from innovation method disclosed herein synthesized nearly an order or magnitude more active luciferase (380.9 ± 2.2 ng mL⁻¹) relative to the prior art method (53.7 ± 0.7 ng mL⁻¹) (**FIG. 1C**).

Example 15. Removal of non-essential processing steps

[0139] The merits of extraneous processing steps, specifically Micrococcal Nuclease (MNase) treatment and uncoupled *in vitro* transcription, were evaluated with the system. Both of these steps include costly reagents and were not obviously beneficial to the overall extract preparation design and CFPS reaction. MNase was originally introduced to the extract preparation protocol as a means of digesting endogenous mRNA and mitigating unwanted competition with the gene of interest, as it preferentially digests single stranded nucleic acids. When cell-free translation was performed using the disclosed extract preparation method with the removal of MNase pre-treatment disclosed CFPS reaction had an increase in active luciferase yield from 380.9 ± 2.2 ng mL⁻¹ to 681.2 ± 10.2 ng mL⁻¹ (**FIG. 1C**).

[0140] Without the claimed subject matter being bound to any particular theory, the MNase decreased activity in the extract by non-productively degrading other RNA species

that were beneficial for CFPS, specifically ribosomal RNA. To explore this hypothesis, RNA samples were prepared from extracts generated with and without MNase pre-treatment. The results suggest that in addition to digesting endogenous mRNA as expected, MNase also digests ribosomal RNA as an undesirable consequence (**FIG. 1D**). Consequently, MNase pre-treatment can be removed from for the CFPS reactions disclosed herein.

Example 16. CFPS reaction condition variables

[0141] Because the combined transcription and translation reactions required an increase in NTP concentrations from 0.75 mM ATP, 0.1 mM GTP, 0 mM CTP, and 0 mM UTP to 1.5 mM of all NTPs, we initially sought to identify the magnesium concentration resulting in the largest luciferase yields (**FIG. 11A**). This is because preservation of the ionic composition, which includes free magnesium, is essential for many protein-nucleic acid interactions and the proper function of protein biosynthesis. We found the highest luciferase yields for 2-hour combined Tx/Tl reactions when magnesium concentration was increased from 2 mM, the concentration used for translation only (Iizuka *et al.* (1994)), to 7 mM. Next, the cell-free reaction temperature was optimized, thereby revealing that active luciferase yield was optimal at 21°C (**FIG. 11B**).

[0142] The physicochemical environment of the CFPS reaction was optimized to better mimic the cytoplasm and improve protein synthesis activity. As an initial step, changes in the ionic composition were targeted, first seeking to use glutamate as the primary anion instead of acetate. Glutamate, which is the most predominant anion used in the cell, is also the preferred anionic species used *in vitro* because of its dispersed electron charge density compared to acetate or chloride. Substituting glutamate salts for acetate salts, improved active luciferase yield more than 2-fold from $1.35 \pm 0.11 \mu\text{g mL}^{-1}$ to $3.18 \pm 0.25 \mu\text{g mL}^{-1}$ (**FIG. 11C**). The inclusion of polyamines, specifically spermidine and putrescine, which act to modify the function of and stabilize DNA, RNA, and tRNA was investigated. Polyamines can improve bacterial CFPS. In the yeast CFPS system disclosed herein, the addition of polyamines to the CFPS reaction proved beneficial (**FIG. 11D**). The optimal concentrations of putrescine and spermidine for luciferase expression were 1 mM and 0.5 mM, respectively. In order to account for the increase in positively charged small molecules in the cell-free reaction, magnesium concentration was decreased from 7 to 5 mM (**FIG. 11E**). These combined changes improved active luciferase yield to $3.89 \pm 0.17 \mu\text{g mL}^{-1}$ (**FIG. 11E**).

Example 17. Effect of added RNase inhibitor and glycerol on CFPS platform performance.

[0143] A technical design criteria for the disclosed system is that it maximizes protein synthesis yield, while minimizing reaction cost. Therefore, the necessity of adding RNase Inhibitor (Qiagen, Valencia, CA) to the cell-free reaction was investigated. RNase Inhibitor is expensive and may no longer be a productive component during combined Tx/Tl reactions (*that is*, mRNA could be continually synthesized with sufficient nucleotides present). Interestingly, the addition of the storage buffer of the RNase Inhibitor (2 mM KH_2PO_4 , 8.0 mM Na_2HPO_4 , 3.0 mM KCl, 150 mM NaCl, pH 7.4, and 50% glycerol) to the CFPS reaction had the same effect of adding the RNase Inhibitor itself in its storage buffer to the CFPS reaction (**FIG. 11E**). Furthermore, the removal of RNase Inhibitor had no obvious effect on overall yield. This result, unexpectedly, led to the discovery that the addition of glycerol (50% of the RNase Inhibitor storage buffer) improved active luciferase yields to $7.69 \pm 0.53 \mu\text{g mL}^{-1}$ (**FIG. 11F**). Overall, the disclosed CFPS system, which includes glutamate salts, NTPs, spermidine, putrescine, glycerol, and magnesium, resulted in a 6-fold improvement of CFPS yield systems lacking these components.

Example 18. Comparison of T3, T7 and SP6 RNA Polymerases in combined transcription/translation yeast CFPS platform systems.

[0144] The activity of three commonly used and commercially available phage RNA polymerases including T7, T3, and SP6 RNAP (New England Biolabs, Ipswich, MA) were compared in the combined Tx/Tl yeast CFPS platform. Motivation for this work stems from a desire to use the yeast CFPS platform with other polymerases. In addition, this example showcases the ease at which different templates can be investigated by simply varying the second round forward primer using our two-step PCR method. In this example, the three variable RNA Polymerase promoter regions were attached to the sfGFP gene used as the reporter protein (**Table 6**).

[0145] Table 6. Round 2 PCR Forward Primers with Variable RNAP Promoter Sequences.

Primer Name	5' → 3' Sequence ¹
T7-Ω- <i>f</i> (SEQ ID NO: 131)	ccgcgaaatTAATACGACTCACTATAGGGAGA <i>tatTTTTacaacaattaccaacaacaac</i>
SP6-Ω- <i>f</i> (SEQ ID NO: 132)	ccgcgaaatATTTAGGTGACACTATAGAAGAG <i>tatTTTTacaacaattaccaacaacaac</i>
T3-Ω- <i>f</i> (SEQ ID NO: 133)	ccgcgaaatAATTAACCCTCACTAAAGGGAA <i>tatTTTTacaacaattaccaacaacaac</i>

¹Lowercase standard: 5' end GC clamp; Uppercase standard: RNAP Promoter sequence; lowercase italics standard: Ω sequence overlap.

[0146] As evident in **FIG. 12A**, both T7 and T3 RNA Polymerases are active in yeast CFPS reactions with comparable yield and SP6 RNA Polymerase is active, although with reduced productivity. In this experiment, each polymerase was normalized for the amount of Units added (1 U is defined as amount of enzyme that will incorporate 1 nmol ATP in a 50 μL reaction in 1 h at 37°C). The storage buffer for each polymerase was identical (50 mM Tris-HCl, 100 mM NaCl, 20 mM β-mercaptoethanol, 1 mM EDTA, 50% Glycerol, and 0.1% Triton[®] X-100, pH 7.9) and differences in volume added were accounted for by supplying the reaction with additional storage buffer.

[0147] Because equal Units of each RNA Polymerase were supplied to each reaction, one would expect equivalent amounts of mRNA to be synthesized. To confirm activity of each polymerase on the generated PCR templates, separate *in vitro* transcription reactions confirmed amplification is possible under the appropriate conditions. Therefore, one possible explanation for the result in **FIG. 12A** is that only T7 and T3 RNA Polymerases are active in the given reaction conditions for yeast CFPS. Alternatively, the initial nucleotides transcribed by SP6 polymerase could negatively affect the Ω sequence function. Ultimately, this result confirms that T3 RNA Polymerase is an option for use in yeast CFPS reactions, but SP6 RNA Polymerase should be avoided in this particularly optimized platform. Furthermore, the commercially available T7 RNA Polymerase gives similar reaction yields to T7 RNA Polymerase prepared in house using the method Swartz *et al.* (2004) confirming that either source would be an appropriate option for general use (**FIG. 12B**).

Example 19. Effect of Kozak sequence elements on yeast CFPS platform.

[0148] The inclusion of the Kozak sequence, which is a consensus sequence found in the 5'-UTR of mRNA, can assist in eukaryotic translation initiation (Kozak, M., "An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs," *Nucleic Acids Res.* 15, 8125-8148 (1987)). The effects of including the Kozak sequence in the DNA template was investigated. Using the developed 2-step PCR method and only varying the initial forward primer, alternate forms of the Kozak sequence were inserted directly downstream of the Ω sequence. The specific primers are illustrated in **Table 7**, where the lower case font indicates gene-specific sequence hybridization, italics font indicates the *S. cerevisiae* consensus Kozak sequence; the bold font indicated the T7 RNA Polymerase promoter sequence; the underlined font indicates the Tobacco Mosaic Virus Ω Sequence (contains 5' overlap between PCR1- and PCR2-specific primers; and the double-underlined font indicates the 3' overlap between PCR1- and PCR2-specific primers for the superfolder GFP ("sfGFP") ORF (SEQ ID NO: 37) (**Table 7**).

[0149] **Table 7. Primer Design for PCR Template Production**

Primer	Sequence (5' → 3')
P1.1- Ω -Kozak(<i>S. cer partial</i>)-sfGFP-f (SEQ ID NO.: 134)	<u>ACAAACAACATTACAATTACTATTTACAATTA</u> <i>AAAAAatgagcaaaggtgaagaactgt</i>
P1.2-sfGFP-r (SEQ ID NO.: 135)	<u>AGCAGCCGGATCTCAGT</u> <i>ttatttttcgaactgcggatgg</i>
P2.1-T7- Ω -f (SEQ ID NO.: 136)	<u>CCGCGAAATTAATACGACTCACTATAGGGAGT</u> <u>ATTTTACAACAATTACCAACAACAACAAACA</u> <u>CAACAACATTACAATTACTATTTACAATTA</u>
P2.2-PolyA50-r (SEQ ID NO.: 137)	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTGTT <u>AGCAGCCGGATCTCAGT</u>

[0150] In the first PCR stage, sense and anti-sense primers were designed to anneal to the sfGFP sequence and add overlapping regions to each end of the PCR product. The 5'-end of the first stage PCR product contained a flanking sequence encoding the canonical yeast Kozak sequence (AAAAAA) and partially encoding the Ω sequence from tobacco mosaic virus. The 3'-end of the first stage PCR product contained an additional 17 nt flanking

sequence. The second stage PCR sense and anti-sense primers were designed to anneal to the 5'- and 3'- flanking sequence, respectively. The 5'-region of the full-length PCR product encoded the full Ω sequence (65 nt) as well as the T7 Promoter sequence. The 3'-end encoded a 50 nt poly(T) sequence to extend the mRNA with a poly(A)₅₀ tail.

[0151] Three Kozak sequences were evaluated in sfGFP expression constructs including: (i) the consensus eukaryotic Kozak sequence, and the *S. cerevisiae* specific sequence (ii) with and (iii) without adjusting the second translated codon after the canonical AUG (**Table 8**).

[0152] **Table 8. Kozak Sequences Investigated in Yeast CFPS Reactions**

Name ¹	5' → 3' Sequence ²
No Kozak (SEQ ID NO: 138)	<i>cat</i> ATG
Consensus Full (SEQ ID NO: 139)	cCC ACC ATG G
<i>S. cerevisiae</i> Partial (SEQ ID NO: 140)	aAa AaA ATG
<i>S. cerevisiae</i> Full (SEQ ID NO: 141)	aAa AaA ATG TCt

¹Expression constructs (5'-UTR containing an Ω sequence and the sfGFP open reading frame) containing these sequences are indicated in parentheses.

²Sequences shown coincide to sequence region candidate Kozak sequence. Lowercase italics font: Restriction digest scar (NdeI); lowercase standard font: most common base found at that position; uppercase standard font: highly conserved base found at that position; bold ATG: start codon.

[0153] **FIG. 3B** shows 40% improvement in sfGFP synthesis yields when either form of the *S. cerevisiae* Kozak sequence is included. Furthermore, the difference in synthesis yields is dominated by the rate of synthesis (as opposed to duration effects). This result would be representative of the fact that translation initiation is indeed catalytically limiting and enhancing the rate of initiation improves the overall protein synthesis yield in a finite reaction. Additionally, there was no significant difference between the *S. cerevisiae* partial and full Kozak sequences, suggesting the second translated codon present in sfGFP (AGC) is sufficient for enhanced expression. Furthermore, there was no significant difference between

the consensus eukaryotic Kozak sequence and no Kozak sequence, thus supporting the notion that the yeast translation apparatus is evolutionally distinct from higher eukaryotes.

Example 20. Extending the reaction lifetime

[0154] **FIG. 13A** shows active luciferase yield throughout the duration of the CFPS batch reaction. The final yield of luciferase after a 120-minute incubation was $7.69 \pm 0.53 \mu\text{g mL}^{-1}$. This duration of synthesis is very robust, being the longest demonstrated for a yeast CFPS batch reaction. There are several potential reasons the cell-free reaction may stop synthesizing protein after 2 h. This list includes but is not limited to: DNA/mRNA degradation, substrate limitations, toxic molecule accumulation, and loss of crude extract activity. Activity loss of crude extract proteins would be particularly detrimental. To test the validity of this concern, a set of experiments were designed that were directed toward unveiling if degradation of factors in the extract alone could be responsible for reaction termination. The extract was pre-incubated at 21°C for 0, 15, 30, 60, 90, 120, 150, and 180 minutes and then the pre-incubated extract was used for 2-hour batch CFPS assays (**FIG. 13B, C**). Strikingly, no change in final CFPS luciferase yield was found after up to three hours of pre-incubation. These data suggest that catalyst activity is not responsible for reaction cessation.

[0155] SEQUENCE TABLE

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
CappA90 (SEQ ID NO: 1)	gagaccacaacggttccctctagaaataatttgtttaactttaagaaggagatatacatatggaagacgccaaaa acataaagaaaggcccgccgcccattctatccgctagaggatggaaccgctggagagcaatgcataaggctat gaagagatacgccctgggttcctggaacaattgctttacagatgcacatcgcagggtgaacatcacgtacgcgga atacttcgaaatgtccgttcggttgccagaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgt tcagtgaaaactctctcaattcttatgccgggtgtgggcgcgttattatcggagttgcagttgcgccgcgaac gacatttataatgaacgtgaattgctcaacagtatgaacattcgcagcctaccgtagtggttttccaaaaagg gttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggattctaaacggatta ccagggtttcagtcgatgtacacgttcgtcacatctcatctacccgggtttaatgaatacgaatttgtaccaga gtcctttgatcgtgacaaaacaattgcaactgataatgaactcctctggatctactgggttacctaagggtgtgccc ttccgatagaactgcctgcgtcagattctcgcagatccagagatcctatttttgcaatcaaatcattccggatactg cgattttaagtgtgttcattccatcacgggttttgaatgtttactacactcggatatttgatgtggatttcgagtcg tctaatgtatagattgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgtgtagtacc aacctatttcattcttcgcaaaagcactctgattgacaaatacatttatcaatttacagaaattgctctgggg ggcaccctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccatcttcagggtacgacaaagat gggctcactgagactacatcagctattctgattacaccggaggggatgataaacggggcgcgggtcgtgtaaagt tgttccatttttgaagcgaagggtgtggtatcggataccgggaaaacgctgggcgttaatacagagaggcgaatta tgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgacaaggatg

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	gatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctgaagcttta attaaatacaaaggataccaggtggccccgctgaattggagtcgatattgttacaacaccccaacatcttcgac gcgggctggcaggtctcccgacgatgacgccggtgaactcccgccgctgtgtttggagcacggaaa gacgatgacggaaaaagagatcgtggattacgtcggcagtcagtaacaaccgccaaaaagttgctggagg agttgtgtttgtggacgaagtaccgaaaggtctaccggaaaactcgacgcaagaaaaatcagagagatcctcat aaaggccaagaaggcggaaggtcgaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
HAPpA90 (SEQ ID NO: 2)	gagaccacaacggtttccctctagataaaccccagttttatatcgtatatgctatctacaggtccactttacactta aatataaaaatactactataaaggaaccagaaaaataaaaaagggtcattatttttggcagatcattatcaaac gcataggaagagaaaaaacagttttttttccacacatatatttggctcctagtagatcaaaagacatttta atgggtgtgtgtttttacacattttctagtacaaaaaaacaaaaaagacatatggaagacgccaaaa acataaagaaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgcataaggctat gaagagatacgcctggtcttggaacaattgctttacagatgcacatcgcaggtgaacatcacgtacgcgga atacttcgaaatgtccgttcggtggcagaagctatgaacgatatgggctgaatacaaatcacagaatcgtcgt tgcagtgaaaactctctcaattcttatgccggtgtgggcgcgttattatcggagttgcagttgcgccgcgaac gacatttataatgaacgtgaattgctcaacagatgaacattcgcagcctaccgtagtgttttccaaaaagg gttgcaaaaaatttgaacgtgcaaaaaaattaccaataccagaaaattattatcatgattctaaaaaggatta ccagggtttcagtcgatgtacacgttcgtcacatctcatctcccggttttaataacagattttgtaccaga gtcccttgatcgtgacaaaacaattgactgataatgaactcctctggatctactgggttacctaagggtgtgccc ttccgatagaactgctgcgtcagattctgcgatgccagagatcctattttggcaatcaaatcattccggatactg cgattttaagtgtgttcattccatcacgggttttgaatgtttactacactcggaattttgatgtggttcgagtcg tctaatgtatagattgaagaagagctgttttacgatccctcaggattacaaaattcaagtgcgtgtgtagtacc aacctatttctattcttcccaaaagcactctgattgacaaatcagattatctaattacacgaaattgcttctggg gcgcacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccatctccaggatcagacaaggat gggctcactgagactacatcagctattctgattacaccgagggggatgataaacggggcgcggtcgtgtaagt tgttccatttttgaagcgaaggtgtggtatcggataccgggaaaacgctgggcgttaatacagagaggcgaatta tgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgacaaggatg gatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctgaagcttta attaaatacaaaggataccaggtggccccgctgaattggagtcgatattgttacaacaccccaacatcttcgac gcgggctggcaggtctcccgacgatgacgccggtgaactcccgccgctgtgttttggagcacggaaa gacgatgacggaaaaagagatcgtggattacgtcggcagtcagtaacaaccgcgaaaaagttgctggagg agttgtgtttgtggacgaagtaccgaaaggtctaccggaaaactcgacgcaagaaaaatcagagagatcctcat aaaggccaagaaggcggaaggtcgaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
TFIIDpA90 (SEQ ID NO: 3)	gagaccacaacggtttccctctagatcgtgcggccgcgaattcgggacgtgaaaattacagtagttactgttttt ttggactataagatcgggggaaagataacacataagaaataaaacgactactagttgactgctctgcggaaga agcaaggaaagtaaggctgcattttttttcttctagtccaacataaaacaggtgatcaagagaactttttaag agctcgtcgcggtccatattggaagacgccaaaaacataaagaaaggcccgccgcttctatccgtagag gatggaaccgctggagagcaactgcataaggctatgaagagatacgcctggttcttggaacaattgctttaca gatgcacatcgcaggtgaacatcacgtacgcgggaatacttgaaatgtccgttcggttggaagaagctatgaaa cgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactcttcaattcttatgccggtgtggg cgcgttattatcggagttgcagttgcgccgcgaacgacatttataatgaacgtgaattgtcacaacagatgaaca ttctgcagcctaccgtagtgtttgttccaaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaata atccagaaaattattatcatgattctaaaacggattaccagggtttcagtcgatgtacacgttcgtcacatctcat ctacctcccggttttaataacagattttgtaccagagtccttgatcgtgacaaaacaattgactgataatgaact cctctggatctactgggttacctaaagggtgtggccctccgatagaactgcctgcgtcagattctgcagtcag agatcctatttttggcaatcaaatcctccgatactgcgattttaagtgtgttcattccatcacgggtttggaatgt tactacactcggaattttgatattgtggttcgagtcgtctaatgtatagattgaagaagagctgtttttacgatccc ttacaggattacaaaattcaagtgcgttctagtaccaacctattttcattcttcccaaaagcactctgattgaaa atacgattatctaatcacgaaattgcttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgca

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	aaacgcttccatctccaggatagacagaaggatagggtcactgagactacatcagctattctgattacacccg aggggggatgataaacgggagcggtcggttaaggtgttccatttttgaagcgaagggttggtatcgataaccg ggaaaacgctgggcgttaatcagagaggcggaattatgtgtcagaggacctatgattatgtccgggtatgtaaca atccgggaagcgaccaacgcttgattgacaaggatggatggctacattctggagacatagcttactgggacgaa gacgaacacttctcatagttgaccgcttgaagtctttaaataacaaaggataaccaggtggcccccgtgaatt ggagtcgatattgttacaacaccccaacatcttcgacgcggcggtggcaggtctcccagcgatgacgccgggtg aactccccgccgctgtgtgttggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgcc agtcaagtaacaaccgcaaaaagttgcgcggagaggtgtgtgttgacgaagtaccgaaaggtcttaccgg aaaactcgacgcaagaaaaatcagagagatctcataaaggccaagaggcggaaggtcgaattgtaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaa
P150pA90 (SEQ ID NO: 4)	gagaccacaacggtttccctctagacccagttcgatcctgggcgaaatcatttttgaattacattaataaggct ttttcaatatctctggaacaacgtttgttctacttactaataagccttgaacatcatgatggcagact tccatcgtagtgggatgatcatatgatgggcgtatcctcgcgactcgataacgacgtgagaacgattttttt ttcttttaccggtattttgtgcgtcctttcaattatagcttttttttttttttctgtactgtttcactgacaaaag ttttttcaagaaaaatttcgatccgcgttctgtgtgcaacggatggatggtagatggaattcaatatgttgctt gaaattttaccaatcttgatattgtgataatttacttaattatgattcttcttccctcaattcttaagcttctacttta ctccttctgtcataaataagcaaggtaagaggacaactgtaattacattacataagatccataggaagacg caaaaacataaagaaaggcccgccgcatctatccgtagaggatggaaccgctggagagcaactgcata aggctatgaagagatacgccttggttcttggaacaattgctttacagatgcacatcgcaggtgaacatcacgta cgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatggcgtgaatacaaatcacagaat cgtcgtatgcagtgaaaactcttcaattctttatgccggtgttgggcgcgtattatcggagttgcagttgcgcc cgcgaacgacatttataatgaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtgtttgttccaa aaagggttgcaaaaaatttgaaactgcaaaaaaattaccaataatccagaaaattattatcatgattctaaaa cggattaccagggttccagtcgatgtacacgttcgtcacatctcatctaccccggttttaataacacgattttgt accagagtccttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaagggtg tgcccttccgcatagaactgctgcgtcagattctgcagatccagagatcctattttggcaatcaaatctccg gatactgcgattttaagtggtgttcattccatcacggttttgaatgttactacactcggaatttgatatgtggatttc gagtcgtcttaagttagatttgaagaagagctgttttacgatccctcaggattacaaaattcaagtgcgttgct agtaccaacctattttcattcttcgcaaaaagcactctgattgacaaatcagatttatcaatttacgaaattgctt ctgggggcgcaccttttcgaaagaagtcggggaagcgggtgcaaaacgcttccatctccaggatagacaga aggatattgggtcactgagactacatcagctattctgattacacccgaggggggatgataaacgggagcggtcg gtaaggtgttccatttttgaagcgaagggtgtggatctggataccgggaaaacgctgggcgttaatcagagagg cgaattatgtgcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgac aaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttga agctttaataaatacaaaaggatacagggtggcccccgtgaattggagtcgatattgttacaacaccccaacat cttcgacgcggcggtggcaggttcccagcgatgacgccggtgaacttcccgccgctgtgtgtttggagca cggaaagacgatgacggaaaagagatcgtgattacgtcggcagtaagtaacaaccgcaaaaagttgcg cggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccgaaactcgacgcaagaaaaatcagagag atcctcataaaggccaagaaggcggaaggtcgaattgtaaaaaataaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
YAP1pA90 (SEQ ID NO: 5)	gagaccacaacggtttccctctagatagtaaccagccctagctgtttggttgatttgacctaggttactcttttctttt ctgggtgcgggtaacaatttgggccccgaaagcgccgcttctgtcatgggaaccggaacccctccgatgaag agtaggagggttgcaactgatgatgcgtaaggtcttaagagatacatttcttaatagtcttccgtttaccgatta agcacagtacctttacgttatataggttggtgttagcttttttctgagccctggttgactgtgcatgaacac gagccatttttagtttgaagggaagtttttgcacccaaaacgttaagaaggaaaagttgtttctaaacccat atggaagacgcaaaaacataaagaaaggcccgccgcatctatccgctagaggatggaaccgctggagag caactgcataaggctatgaagagatacgccttggttcttgaacaattgctttacagatgcacatcgcaggtga acatcacgtacgcggaatacttgaaatgtccgttcggttggcagaagctatgaacgatatggctgaatacaa atcacagaatcgtcgtatgcagtgaaaactcttcaattctttatgccggtgttgggcgcgttattatcggagttgc

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	agttgcgcccgcgaacgacattataatgaacgtgaattgctcaacagtatgaacatttcgagcctaccgtagtg ttgtttccaaaaaggggttgcaaaaaatttgaacgtgcacaaaaaattaccaataatccagaaaattattatcatg gattctaaaacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgttttaaatgaat acgattttgtaccagagtcctttgatcgtgacaaaaaattgcactgataatgaactcctctggatctactgggttac ctaagggtgtggcccttcgcatagaactgcctgcgtcagattctcgcagatccagagatcctattttggcaatcaa atcattccggatactgcgattttaagtgtgttcattccatcacggtttgggaatgtttactacactcggatatttgata tgtggattcagtcgtcttaattgtatagattgaagaagagctgtttacgatccctcaggattacaaaattcaaa gtgcgttgtagtaccacacctatttcattcttcgcaaaagcactctgattgacaaatcagttatctaatttacac gaaattgcttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgcacaaacgctccatctccagg gatacacaaggatatgggtcactgagactacatcagctattctgattacacccgagggggatgataaacggg gcgcggtcggtaaaagtgttcatttttgaagcgaaggtgtggatctggataccgggaaacgctgggcgftaa tcagagaggcgaattatgtgtcagaggacctatgattatgtccggtatgtaaacaatccggaagcgaccaacgc cttgattgacaaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagtt gaccgctgaagtccttaataatacaaaaggataccaggtggcccccgctgaattggagtcgatattgttacaac accccaacatcttcgacgcgggcgtggcaggtctcccgacgatgacgccggtgaactcccggccgctgtgtt gtgttgagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgcagtcagtaacaaccgccaa aaaagttgcgcggaggagttgtgttgggacgaagtaccgaaaggtcttaccggaaactcgacgcaagaaa aatcagagagatcctcataaaggccaagaaggcggaaagtccaattgtaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
IGRpA90 (SEQ ID NO: 6)	agaccacaacggtttccctctagagcaaaaatgtgatcttgccttgaataacaattttgagaggtaataaattacaa gtagtgctattttgtatttaggttagctatttagctttacgttcaggatgcctagtgccagccccacaatatccagg aagccctctctgcggttttcagattaggtagtcgaaaaacctaagaaatttacctgctacattcaagattcatatgg aagacgccaaaaacataaagaaaggcccgccgcttctatccgctagaggatggaaccgctggagagcaac tgcataaaggctatgaagagatacgcctgttcctggaacaattgctttacagatgcacatatcgaggtgaacat cacgtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaacgatatgggtgaatacaaatca cagaatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgttgggcgcgttatttatcgagttgcagt gcgccgcgaacgacattataatgaacgtgaattgtcaacagtatgaacatttcgagcctaccgtagtggttgt ttccaaaaaggggttgcaaaaaatttgaacgtgcacaaaaaattaccaataatccagaaaattattatcatggatt ctaaaacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgttttaataatgaacg attttgtaccagagtcctttgatcgtgacaaaaaattgcactgataatgaactcctctggatctactgggttacctaa ggggtgtggcccttcgcatagaactgcctgcgtcagattctcgcagatccagagatcctattttggcaatcaaatca ttccggatactgcgattttaagtgtgttcattccatcacggtttgggaatgtttactacactcggatatttgatgtg gatttcgagtcgtcttaattgtatagattgaagaagagctgttttacgatccctcaggattacaaaattcaaaagtc gttgctagtaccaacctatttcattcttcgcaaaagcactctgattgacaaatacatttatctaatttacacgaaa ttgcttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgcacaaacgcttccatcttcagggtac gacaaggatatgggtcactgagactacatcagctattctgattacacccgagggggatgataaacggggcgc ggtcggtaaaagtgttcatttttgaagcgaaggtgtggatctggataccgggaaacgctgggcgftaatcag agaggcgaattatgtgtcagaggacctatgattatgtccggtatgtaaacaatccggaagcgaccaacgccttg attgacaaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgacc gcttgaaagtcttaataataatacaaaaggataccaggtggcccccgctgaattggagtcgatattgttacaacacc caacatcttcgacgcgggcgtggcaggtctcccgacgatgacgccggtgaacttcccgccgctgtgtgtttt ggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgcagtcagtaacaaccgccaaaaa agttgcgcggaggagttgtgttgggacgaagtaccgaaaggtcttaccggaaactcgacgcaagaaaaatc agagagatcctcataaaggccaagaaggcggaaagtccaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
A64pA90 (SEQ ID NO: 7)	gagaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacatat ggaaagacgccaaaaacataaagaaaggcccgccgcttctatccgctagaggatggaaccgctggagagc aactgcataaaggctatgaagagatacgcctgttctggaacaattgctttacagatgcacatatcgaggtgaa catcacgtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaacgatatgggtgatacaaaa tcacagaatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgttgggcgcgttatttatcgaggtgca

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	gttgcgccgcgaacgacattataatgaacgtgaattgctcaacagatgaacatttcgagcctaccgtagtgtt tgtttcaaaaagggttgcaaaaattttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgg attctaaaacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacccgggtttaatgaata cgattttgtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacc taagggtgtggcccttcgcataagaactgcctgcgtcagattctcgcagatccagagatcctattttggcaatcaaa tcattccggatactgcgattttaagtgtgtccattccatcacgggtttggaatgtttactacactcggatatttgat gtggatttcgagtcgtctaatgtatagattgaagaagagctgttttacgatcccttcaggattacaaaattcaag tgcgttgctagtaccaaccctatttcatttcgcacaaagcactctgattgacaaatcagttatctaatttacacg aaattgcttctgggggcgcaccttttcgaaagaagtcggggaagcgggtgcaaacgctccatctccaggga tacgacaaggatatgggtcactgagactacatcagctattctgattacaccggaggggatgataaacgggc gcggtcgtaaatgtgtccattttgaagcgaaggtgtggtatcgataccgggaaacgctgggctgtaatc agagaggcgaattatgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgcct tgattgacaaggatggatggctacattctggagacatagcttactgggacgaagcgaacactcttcatagttga ccgcttgaagtcttaataatacaaaaggataccaggtggcccccgtgaattggagtcgatattgttacaacac cccaacatcttcgacgcgggcgtggcaggtctcccgacgatgacgccggtgaactcccgcgccggttgtgtt ttggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgccagtcagtaacaaccgccaaa aagttgcgcggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccggaaactcgacgcaagaaaaat cagagagatcctcataaaggccaagaaggcggaagtcctcaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
HedrinpA90 (SEQ ID NO: 8)	gagtattttattcttctgtaaaaaattagaaaaataaaatataaacatatggaagacgccaaaaacataaaagaaa ggccccggcgccattctatccgctagaggatggaaccgctggagagcaactgcataaggctatgaagagatacgc ccctgggtcctggaacaattgctttacagatgcacatcgcaggtgaacatcacgtacgcggaatacttcgaaat gtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaa ctctctcaattctttatgccggtgttgggcgcgttattatcggagttgcagttgcgccgcgaacgacattataat gaacgtgaattgctcaacagatgaacatttcgagcctaccgtagtgtttgtttcaaaaagggttgcaaaaat tttgacgtgcaaaaaaattaccaataatccagaaaattattatcatggattctaaaacggattaccagggatttca gtcgtatgtacacgttcgtcacatctcatctacccgggtttaatgaatacagttttgtaccagagtcctttgatcgtg acaaaacaattgcactgataatgaactcctctggatctactgggttacctaagggtgtggcccttcgcatagaac tgctgcgtcagattctcgcagcagagatcctattttggcaatcaaatcatcgggatactgcgattttaaagtgtt gttccattccatcacgggtttggaatgtttactacactcggatatttgatatgtggatttcgagtcgtctaatgtataga tttgaagaagagctgttttacgatcccttcaggattacaaaattcaaaagtcggttgctagtaccaaccctatttcatt cttcgcaaaaagcactctgattgacaaatcagttatctaatttacacgaaattgcttctgggggcgcacctcttc gaaagaagtcggggaagcgggttgcaaacgcttccatctccaggatgacgaaaggatggggtcactgag actacatcagctattctgattacaccggaggggatgataaacgggcgcggtcggtaagttgttccatttttga agcgaaggtgtggtatcgtgataccgggaaacgctgggcgttaatacagagaggcgaattatgtgtcagagga cctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgacaaggatggatggctacatt ctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttgaagtcttaataatacaaaa ggataccaggtggcccccgtgaattggagtcgatattgttacaacaccccaacatcttcgacggggcgtggc aggtcttcccgacgatgacgccggtgaactcccgcgccggttgttttggagcacggaaagacgatgacgg aaaaagagatcgtggattacgtcgcagtcagtaacaaccgccaaaaagttgcgcggaggagttgtgtttgtg gacgaagtaccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagagagatcctcataaaggccaag aaggcggaaggtcctcaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ΩpA90 (SEQ ID NO: 9)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaaagaaaggccggcgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgcctggttctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgttgggcgcgttattatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgagcctaccgtagtgtttgttcc aaaaagggttgcaaaaattttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggattctaa

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	aaccgattaccagggtttcagtcgatgtacacgttcgtcacatctcatctaccccggtttaatgaatacgtttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaagg tgtggcccttcgcatagaactgcctgcgtcagattctcgcagatccagagatcctattttggcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacggtttggaaatgtttactacactcggatattgatattgtggatt tcgagtcgtcttaattgatagatttgaagaagagctgttttacgatcccttcaggattacaaaattcaaaagtgcgttg ctagtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacgtttatctaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccatctccagggtacacgac aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtc ggtaaagtgttccatttttgaagcgaaggtgtggatctggataccgggaaaacgctgggcgttaatacagagag gcgaattatgtgcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacactcttcatagttgaccgcttg aagtcTTAATAAATACAAAGGATACCAGGTGGCCCCGCTGAATTGGAGTCGATATTGTACAACACCCCAACA TCTTCGACGCGGGCGTGGCAGGTCTCCCGACGATGACGCCGGTGAACCTCCCGCCGCTGTGTGTTGGAGC ACGGAAAGACGATGACGGAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCCAAAAGTTGC GCGGAGGAGTGTGTTGTGGACGAAGTACC GAAAGGTCTACCGAAAACGACGCAAGAAAAATCAGAGAGA GATCCTATAAAGGCCAAGAAGGGCGGAAAGTCAAATTGTAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
TEVpA90 (SEQ ID NO: 10)	gaaataacaaatctcaacacaacatatacaaaaacaaacgaatctcaagcaatcaagcattctacttattgcagc aatttaaatcatttctttaaaagcaaaagcaatttctgaaaatttaccatttacgaacgataagcaatggaagacgc caaaaacataaaagaaaggcccgccgcatctatccgctagaggatggaaccgctggagagcaactgcataa ggctatgaagagatacggcctggttctggaacaattgcttttacagatgcacatacagaggtgaacatcacgtac gcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggtgtaatacaatcacagaatc gtcgtatgcagtgaaaactcttcaattctttatgcccgtgttggcgcttatttatcgaggtgcagttgcgccc gcgaacgacatttataatgaacgtgaattgtcgaacagatgaacatttcgcagcctaccgtagtgtttgttccaaa aaggggttgcaaaaattttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggattctaaac ggattaccagggtttcagtcgatgtacacgttcgtcacatctcatctaccccggttttaatagaatacgttttga ccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaagggtgt ggcccttcgcatagaactgcctgcgtcagattctcgcagatccagagatcctattttggcaatcaaatcattccgg atactgcgattttaagtgtgttcattccatcacggtttggaaatgtttactacactcggatattgatattgtgatttcg agtcgtcttaattgatagatttgaagaagagctgttttacgatcccttcaggattacaaaattcaaaagtgcgttgcta gtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacgtttatctaatttacacgaaattgcttc tgggggcgcacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccatctccagggtacacgacaa ggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtcgg taaagtgttccatttttgaagcgaaggtgtggatctggataccgggaaaacgctgggcgttaatacagagaggc gaattatgtgcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgaca aggatggatggctacattctggagacatagcttactgggacgaagacgaacactcttcatagttgaccgcttgaa gtctttaattaaatacaaaaggataaccaggtggccccgctgaattggagtcgattgttacaacaccccaacatct tcgacgcggcggtggcaggtcttccgacgatgacggcggtgaacttcccgccgctgtgttttggagcac ggaaagacgatgacggaaaagagatcgtggattacgtcggcagtcgaagtaacaacggccaaaaggtgcgc ggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccggaaaactcgcgcaagaaaaatcagagaga tcctcataaaggccaagaaggcggaaggtccaaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
TbmpA90 (SEQ ID NO: 11)	gatttaattattgcaacaacaacaacaattacaataataacaacaaaatacaacaacaacaacatggaagac gccaaaacataaagaaaggcccgccgcatctatccgctagaggatggaaccgctggagagcaactgcatt aaggctatgaagagatacggcctggttctggaacaattgcttttacagatgcacatacagaggtgaacatcacgt acgcggaatacttcgaaatgtccgttcgggtggcagaagctatgaacgatatgggtgaatacaaatcacagaa tcgtcgtatgcagtgaaaactcttcaattctttatgccgggtgttggcgcttatttatcgaggtgcagttgcgcc cgcgaacgacatttataatgaacgtgaattgtcgaacagatgaacatttcgcagcctaccgtagtgtttgttccaa aaagggttgcaaaaattttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgattctaaaa cggattaccagggtttcagtcgatgtacacgttcgtcacatctcatctaccccggttttaatagaatacgtttgt

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	accagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaagggtg tgcccttccgcatagaactgcctgcgtcagattctcgcatgccagagatcctattttggcaatcaaatcattccg gatactgcgattttaagtgtgttcattccatcacggtttgaatgtttactacactcggatatttgatatgtggatttc gagtcgtcttaattgatagattgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgtgct agtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacagttatctaatttacgaaattgctt ctgggggagcaccctttcgaagaagtcggggaagcgggtgcaaaacgcttccatcttccagggatacgcaca aggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccggcgcggtcgc gtaaagtgttccatttttgaagcgaaggtgtggatctggataccgggaaaacgctggcggttaatacagagagg cgaattatgttcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgac aaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttga agtctttaataatacaaaaggataccagggtggccccgctgaattggagtcgatattgttacaacaccccaacat cttcgacgcggcggtggcaggtcttccgacgatgacgcgggtgaactcccgcgcggtgtgttttggagca cggaaagacgatgacggaaaaagagatcgtggattacgtcgcagtcgaagtaacaaccgcaaaaagtgtgcg cggaggagttgtgtttgtggacgaagtaccgaaaggcttaccggaaaactcgcgcaagaaaaatcagagag atcctcataaaggccaagaaggcggaagtccaaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ΩpA25 (SEQ ID NO: 12)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgcctggttctcgaacaattgctttacagatgcacatcagaggtaacatcac gtacgcggaatacttcgaaatgtccgttcggtggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccggtgttggcgcggtatttatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgtcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggtattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgttttaataacacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttccgcatagaactgcctgcgtcagattctcgcagatgccagagatcctatttttggcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacggtttgaatgtttactacactcggatatttgatatgtggatt tcgagtcgtcttaattgatagattgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgttg ctagtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacagttatctaatttacgaaattgc ttctgggggagcaccctttcgaagaagtcggggaagcgggtgcaaaacgcttccatcttccagggatacgcac aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccggcgcggtc ggtaaaagtgttccatttttgaagcgaaggtgtggatctggataccgggaaaacgctggcggttaatacagagag gcaattatgttcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttg aagtctttaataatacaaaaggataccagggtggccccgctgaattggagtcgatattgttacaacaccccaaca tcttcgacgcggcggtggcaggtcttccgacgatgacgcgggtgaactcccgcgcggtgtgttttggagc acggaaagacgatgacggaaaaagagatcgtggattacgtcgcagtcgaagtaacaaccgcgaaaaagtgtc gcggaggagttgtgtttgtggacgaagtaccgaaaggcttaccggaaaactcgcgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaagtccaaattgtaaggatccgtacgagctcatgcgaattcctcga gcaccaccaccaccactgagatccggctgtaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ΩpA50 (SEQ ID NO: 13)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgcctggttctcgaacaattgctttacagatgcacatcagaggtaacatcac gtacgcggaatacttcgaaatgtccgttcggtggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccggtgttggcgcggtatttatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgtcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggtattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgttttaataacacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttccgcatagaactgcctgcgtcagattctcgcagatgccagagatcctatttttggcaatcaaatcattcc

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	ggatactgcgattttaaagtgtgttcattccatcacggttttgaatgttactacactcggatatttgatatgtggatt tcgagtcgtcttaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaattcaaatgcgttg ctagtaccaaccctattttcattcttcgcaaaagcactctgattgacaaatacgaattatcaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccattctccagggatacgcac aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtc ggtaaaagtgttccatttttgaagcgaaggtgtggtatctggataccgggaaaacgctgggcgttaacagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccggagcgcaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacactcttcatagttgaccgctg aagtccttaattaatacaaaaggataaccaggtggccccgctgaattggagtcgataattgtacaacacccaaca tcttcgacgcgggcgtggcaggtcttccgacgatgacgcgggtgaacttcccgccgctgtgttttggagc acggaaagacgatgacggaaaagagatcgtggattacgtcggcagtcagtaacaaccgcgaaaaagtgc gcggaggagttgtgttggacgaagtaccgaaaggtcttaccgaaaactcagcgaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccttaaggatccgtacgagctcatcgaaattcctcga gcaccaccaccaccactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa
ΩpA170 (SEQ ID NO: 14)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactattacaattacatattggaag acgcaaaaacataaaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccttggctcctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggtggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccgggtgtggcgcggtatttatcgaggtgcagttgcg cccgcgaaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatctgattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacccgggtttaaataacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggccctccgcatagaactgcctgcgtcagattctcgcagccagagatcctatttttgcaatcaaatcattcc ggatactgcgattttaaagtgtgttcattccatcacggttttgaatgttactacactcggatatttgatatgtggatt tcgagtcgtcttaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaattcaaatgcgttg ctagtaccaaccctattttcattcttcgcaaaagcactctgattgacaaatacgaattatcaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccattctccagggatacgcac aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtc ggtaaaagtgttccatttttgaagcgaaggtgtggtatctggataccgggaaaacgctgggcgttaacagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccggagcgcaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacactcttcatagttgaccgctg aagtccttaattaatacaaaaggataaccaggtggccccgctgaattggagtcgataattgtacaacacccaaca tcttcgacgcgggcgtggcaggtcttccgacgatgacgcgggtgaacttcccgccgctgtgttttggagc acggaaagacgatgacggaaaagagatcgtggattacgtcggcagtcagtaacaaccgcgaaaaagtgc gcggaggagttgtgttggacgaagtaccgaaaggtcttaccgaaaactcagcgaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccttaaggatccgtacgagctcatcgaaattcctcga gcaccaccaccaccactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ΩFBAL (SEQ ID NO: 15)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactattacaattacatattggaag acgcaaaaacataaaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccttggctcctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggtggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccgggtgtggcgcggtatttatcgaggtgcagttgcg cccgcgaaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatctgattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacccgggtttaaataacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	tgtggcccttccgcatagaactgcctgcgtcagattctcgcagccagagatcctatttttggaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacggtttggaatgtttactacactcggatatttgatgtggatt tcgagtcgtctaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgtg ctagtaccaacctattttcattcttcgcaaaagcactctgattgacaaatacgaattatcatttacacgaaattgc ttctgggggagcaccctttcgaaagaagtcggggaagcgggtgcaaaacgcttccatctccaggagatcacgac aaggatatgggctcactgagactacatcagctattctgattacaccgagggggatgataaaccggcgcggtc ggtaaaagttgtccatttttgaaagcgaaggtgtggtatctggataccgggaaaacgctgggcgttaacagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctg aagctttaataatacaaaaggataaccaggtggcccccgctgaattggagtcgataattgtacaacaccccaaca tcttcgacgcgggcggtggcaggtcttccgacgatgacgcccgtgaactccccgccgctgtgtgtttggagc acggaaagacgatgacggaagaagagatcgtggattacgtcgcagtcgaagtaacaaccgcgaaaaagttgc gcggaggagttgtgtgtgtggacgaagtaccgaaaggtctaccgaaaactcgacgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtcgaattgtaaggatccgtacgagctcagattgaatctgtt tagaaataatggaatattttttattttattttattttggtcggctctttcttgaaggtcaatgacaaaatgata tgaaggaaataatgatttctaaattttacaacgaagatattttacaaaagcctagctcatctttgtcatgcactatt ttactcacgcttgaaattaacggccagtcactgcggagtcatttcaaatgcacctaactcagctatcgttttgata gtcattttggagtcgcgattgtcttctgttattacaactgttttaattttatttcattctggaacttccgagttctgt aaagtcttcatagtagcttactttatcctccaacatatttaacttcatgcaatttcggctctaaattttccacatcatca agttcaacatcatcttttaactgaattattctctagctcttccaaccaagcctcattgctccttgattactggtgaaa agtatacactttgcccgaatccaggtcaaaacttctcgaagaattcaccaattctcgacatcatagtaaa ttgtttttgtctcccatcacaatttaataacatgatggattcttatgaaagcgtgggtaatggacgtgactctactt cgctttttccctactccttttagtacggaagacaatgctaataaataagagggtataataatattataatcggcaa aaaagattaacgccaaagcgttaattatcagaaagcaaacgtcgtaccaatcctgaatgcttcccaattgtatatt aagagtcacacagcaacatattctgttattaaattaattattattgattttgatattgtataaaaaaccaaataatgta taaaaaaagtgaataaaaaataacaaagtatggagaaatattagaagctatacgttaaaaccagaacgtgcac aattttttaatctgccaatggaaaaacggaaatatacggaaaagaagtgaaagtaatgtagaaaggcaaaa aaggaaaagaacaatttaaaatatcttaagattatattagaacaacaccaatgttcatttccttagaataat ccgaatgaaatgaccaactactgtttgtaaactgaggaagaagaatattttctccgaaaactgtcataccg tagctgtctgtctttattgttttgacttatttttcaaaaatcaccgtgtttgtgagttttatagttgtgtataaa ttgtcacttactgaatttttttcacagaacatacaagcaaaaggcggtccgttgatgaacggatcttatatgcc tttcaagtgtcactgcgtcgaatgccttctcacagctttacacttgaaaggtttattttatcgtagttgtggggtc aatg
ΩFBAS (SEQ ID NO: 16)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatattggaag acgcaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacggcctggtcctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatactcgaatgtccgttcggttgccagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccgggtgtggcgcggtatttatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtgtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgggttaataatgacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttccgcatagaactgcctgcgtcagattctcgcagccagagatcctatttttggaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacggtttggaatgtttactacactcggatatttgatgtggatt tcgagtcgtctaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgtg ctagtaccaacctattttcattcttcgcaaaagcactctgattgacaaatacgaattatcatttacacgaaattgc ttctgggggagcaccctttcgaaagaagtcggggaagcgggtgcaaaacgcttccatcttccaggagatcacgac aaggatatgggctcactgagactacatcagctattctgattacaccgagggggatgataaaccggcgcggtc ggtaaaagttgtccatttttgaaagcgaaggtgtggtatctggataccgggaaaacgctgggcgttaacagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctg

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	aagctttaattaatacaaaaggataaccaggtggcccccgtgaattggagtcgatattgttacaacaccccaaca tcttcgacgcggcggtggcaggtcttcccgacgatgacgccgggtgaacttcccgccgccgtgtgttttggagc acggaaagacgatgacggaaaaagagatcgtggattacgtcgccagtgcaagtaacaaccgcgaaaaagtgc gcggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccaaattgtaaggatccgtacgagctcgagtattgaatctgtt tagaaataatggaatattattttatttatttatttatttggctggctctttctctgaagggtcaatgacaaaatgata tgaaggaaataatgatttctaaaattttacaacgtaagatattttacaaaagcctagctcatctttgtcatgcactatt ttactcacgcttgaaattaacggccagtcactcgaggagtcattcaaaagtcacctaactgatctatcgttttgata gtcatctttggagttcgcgattgtcttctgttattcacaactgttttaattttatttcattctggaactcttcgagttcttgt aaagctttcatagtagcttactttatcctccaacatatttaacttcatgcaatttcggctctaaattttccacatcatca agttcaacatcatcttttaactgaatttattcttagctctccaaccaagcctcattgctccttgattactggtgaaa agtatacactttgcgcgcaatccaggtcaaaactttcctgcaagaattcaccaatttctgcacatcatagtaaa ttgttttgtctcccatcacaatttaataacatctgatggattcttatgaagcgctgggtaattggacgtgtca
ΩTMV13U200 (SEQ ID NO: 17)	gagtatttttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatattggaag acgcaaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccttggctcctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccggtgttggcgcggtattatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtggtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaaattaccaataatccagaaaattattatcatgattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgggtttaaataacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttcccatagaactgcctgcgtcagattctcgcagatgccagagatcctatttttggcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacgggtttggaatgtttactacactcggatatttgatatgtggatt tcgagtcgtctaatgtatagatttgaagaagagctgtttttacgatccctcaggattacaaaattcaagtgcgttg ctagtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacatttatcatttacacgaaattgc ttctgggggcgcaccttttcgaaagaagtcggggaagcggttgcaaaacgcttccatcttcaggggatacgc aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccggcgcggtc ggtaaaattgttccatttttgaagcgaaggttggatctggataccgggaaaacgctgggcgttaacagagag gcgaattatgtgcagaggacatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctg aagctttaattaatacaaaaggataaccaggtggcccccgtgaattggagtcgatattgttacaacaccccaaca tcttcgacgcggcggtggcaggtcttcccgacgatgacgccgggtgaacttcccgccgccgtgtgttttggagc acggaaagacgatgacggaaaaagagatcgtggattacgtcgccagtgcaagtaacaaccgcgaaaaagtgc gcggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccggaaaactcgacgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccaaattgtaaggatccgtacgagctcaggaaaagtgaata tcaatgagtttatcaccctgacaaaatggagaagatcttaccgtcagtggtttaccctgtaaagagtggtatgtgt ccaaagtgataaaataatggttcagagaatgagtcattgtcagagggtgaaccttctaaaggagttgaagcttattg atagtgatacgtctgtttgacgggttgg
ΩTMV13U400 (SEQ ID NO: 18)	gagtatttttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatattggaag acgcaaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccttggctcctggaacaattgctttacagatgcacatcagaggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccggtgttggcgcggtattatcggagttgcagttgcg cccgcgaacgacatttataatgaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtggtttgttcc aaaaaggggttgcaaaaaatttgaacgtgcaaaaaaaattaccaataatccagaaaattattatcatgattctaa aacggattaccagggatttcagtcgatgtacacgttcgtcacatctcatctacctcccgggtttaaataacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttcccatagaactgcctgcgtcagattctcgcagatgccagagatcctatttttggcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacgggtttggaatgtttactacactcggatatttgatatgtggatt

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	tcgagtcgtcttaatgtatagatttgaagaagagctgttttacgatcccttcaggattacaaaattcaagtgcgttg ctagtaccacccctattttcattcttcgcaaaagcactctgattgacaaatacatttatcaatttacacgaaattgc ttctgggggcgccacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccattctccagggatacgc aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtc ggtaaaagttgtccatttttgaagcgaaggttggtgatctggataccgggaaaacgctgggcgttaatcagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccgggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttg aagtccttaataatacaaaaggataaccaggtggcccccgctgaattggagtcgatattgtacaacaccccaaca tcttcgacgcggcggtggcaggtcttcccgacgatgacgcgggtgaacttcccgccgctgtgtgtttggagc acggaaagacgatgacggaaaagagatcgtggattacgtcggcagtcagtaacaaccgcgaaaaagttgc gcggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccgaaaactcagcgaagaaaaatcagaga gatcctcataaaggccaagaaggcggaagtcgaaattgtaaggatccgtacgagctcaggaaaagtgaata tcaatgagtttatcgacctgacaaaaatggagaagatcttaccgtcgtatgtttaccctgtaaagagtggtatgtgt ccaaagtgataaaataatggttcagagaatgagtcattgtcagagggtgaaccttctaaaggagtaagcttattg atagtggatacgtctgttagccgggttggtcgtcacgggcgaatggaaactgcctgacaattgcagaggaggtgt gagcgtgtgtctgtggacaaaaggatggaaagagccgacgaggccactctcggatcttactacacagcagct gcaaaagaaaagatttcagttcaaggctgtcccaattatgtataaccaccaggacgcgatgaaaaacgtctgg caagtt
ΩTMV13U700 (SEQ ID NO: 19)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaaagaaaggccgggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacggcctgttctggaacaattgcttttacagatgcacatatcgaggtaacatcac gtacgcggaatacttcgaaatgtccgttgggtggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattctttatgccgggtgtgggcgcgttattatcgagggtgcagttgcg ccgcaacgacatttataatgaacgtgaattgtcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgattctaa aacggattaccagggtttcagtcgatgtacacgttcgtcacatctcatctacctccgggtttaatgaatacgtttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggccctccgcatagaactgctgcgtcagttctcgcagagatcctattttggcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacgggtttggaatgtttactacactcggatattgatattggatt tcgagtcgtcttaatgtatagatttgaagaagagctgttttacgatcccttcaggattacaaaattcaagtgcgttg ctagtaccacccctattttcattcttcgcaaaagcactctgattgacaaatacatttatcaatttacacgaaattgc ttctgggggcgccacctcttcgaaagaagtcggggaagcgggtgcaaaacgcttccattctccagggatacgc aaggatatgggctcactgagactacatcagctattctgattacacccgagggggatgataaaccgggcgcggtc ggtaaaagttgtccatttttgaagcgaaggttggtgatctggataccgggaaaacgctgggcgttaatcagagag gcgaattatgtgtcagaggacctatgattatgtccgggtatgtaacaatccgggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgcttg aagtccttaataatacaaaaggataaccaggtggcccccgctgaattggagtcgatattgtacaacaccccaaca tcttcgacgcggcggtggcaggtcttcccgacgatgacgcgggtgaacttcccgccgctgtgtgtttggagc acggaaagacgatgacggaaaagagatcgtggattacgtcggcagtcagtaacaaccgcgaaaaagttgc gcggaggagttgtgtttgtggacgaagtaccgaaaggtcttaccgaaaactcagcgaagaaaaatcagaga gatcctcataaaggccaagaaggcggaagtcgaaattgtaaggatccgtacgagctcaggaaaagtgaata tcaatgagtttatcgacctgacaaaaatggagaagatcttaccgtcgtatgtttaccctgtaaagagtggtatgtgt ccaaagtgataaaataatggttcagagaatgagtcattgtcagagggtgaaccttctaaaggagtaagcttattg atagtggatacgtctgttagccgggttggtcgtcacgggcgaatggaaactgcctgacaattgcagaggaggtgt gagcgtgtgtctgtggacaaaaggatggaaagagccgacgaggccactctcggatcttactacacagcagct gcaaaagaaaagatttcagttcaaggctgtcccaattatgtataaccaccaggacgcgatgaaaaacgtctgg caagttttagtttaattagaaatgtaaaagatgcagcgggttctgtccgcttctctggagttgtgtcgggtgtatt gtttatagaaataataaaattaggtttgagagagaagattacaacgtgagagacggaggggcccatggaactt acagaagaagtcgttgatgagttcatggaagatgtccctatgtcaatcaggcttgcaaaagttcgtatcggaccg gaaaaagagtgatgtccgtaaaagggaataatgtagtagtgaccgggtcagtgccgaacaagaactatagaaa t

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
ΩTMV23U (SEQ ID NO: 20)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccctggttcctggaacaattgctttacagatgcacatatcgaggtaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgttgggcgcgtatttatcggagtgcagttgcg ccgcgaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggtattctaa aacggattaccagggttcagtcgatgtacacgttcgtcacatctcatctacctcccgggtttaatgaatacgtttt gtaccagagtcctttgatcgtgacaaaacaattgactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttcgcatagaactgcctgcgtcagatttcgcagatccagagatcctattttggcaatcaaatcattcc ggatactgcgattttaagtgttgcattccatcacgggtttggaatgtttactacactcggatatttgatatgtgatt tcgagtcgtctaatgtatagattgaagaagagctgttttacgatccctcaggattacaaaattcaagtgcgttg ctagtaccaacctatttcttctcgcctcgaagcactctgattgacaaatcagatttatcaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcgttgcaaaacgttccatcttccagggtacacgac aaggatatgggctcactgagactacatcagctattctgattacaccgagggggatgataaaccgggcgcggtc ggtaaaagttgtccatttttgaagcgaaggttggatctggataccgggaaaacgctgggcgttaatacagagag gcgaattatgtgcagaggacctatgattatgtccggttatgtaacaatccggagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgttg aagtcfttaataatacaaaaggataaccagggtggccccgcgtgaattggagtcgatattgtacaacaccccaaca tcttcgacgcggcggtggcaggtcttccgacgatgacggcggtgaacttcccgccgctgtgttttgagc acggaaagacgatgacggaaaagagatcgtggattacgtcgcagtcgaagtaacaaccgcgaaaaagtgc gcggaggagttgtgttgggacgaagtaccgaaaggtcttaccgaaaactcgcgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccaaattgtaaggatccgtacgagctcggtatgtaagatgc ataataataacggattgtgtccgtaatcacacgtggtgcgtacgataacgcatagtgttttccctcacttagatc gaagggtgtgtcttgatgcgcgggtcaaatgtatatggtcatatacatccgcaggcacgtaataaagcgag gggttcgaatcccccggtacccccggtaggggccca
ΩN3U (SEQ ID NO: 21)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatatggaag acgcaaaaacataaagaaggcccggtgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgccctggttcctggaacaattgctttacagatgcacatatcgaggtaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgttgggcgcgtatttatcggagtgcagttgcg ccgcgaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatggtattctaa aacggattaccagggttcagtcgatgtacacgttcgtcacatctcatctacctcccgggtttaatgaatacgtttt gtaccagagtcctttgatcgtgacaaaacaattgactgataatgaactcctctggatctactgggttacctaaggg tgtggcccttcgcatagaactgcctgcgtcagatttcgcagatccagagatcctattttggcaatcaaatcattcc ggatactgcgattttaagtgttgcattccatcacgggtttggaatgtttactacactcggatatttgatatgtgatt tcgagtcgtctaatgtatagattgaagaagagctgttttacgatccctcaggattacaaaattcaagtgcgttg ctagtaccaacctatttcttctcgcctcgaagcactctgattgacaaatcagatttatcaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcgttgcaaaacgttccatcttccagggtacacgac aaggatatgggctcactgagactacatcagctattctgattacaccgagggggatgataaaccgggcgcggtc ggtaaaagttgtccatttttgaagcgaaggttggatctggataccgggaaaacgctgggcgttaatacagagag gcgaattatgtgcagaggacctatgattatgtccggttatgtaacaatccggagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgttg aagtcfttaataatacaaaaggataaccagggtggccccgcgtgaattggagtcgatattgtacaacaccccaaca tcttcgacgcggcggtggcaggtcttccgacgatgacggcggtgaacttcccgccgctgtgttttgagc acggaaagacgatgacggaaaagagatcgtggattacgtcgcagtcgaagtaacaaccgcgaaaaagtgc gcggaggagttgtgttgggacgaagtaccgaaaggtcttaccgaaaactcgcgcaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtccaaattgtaa

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
N5UpA90 (SEQ ID NO: 22)	gagaccacaacggttccctctagaataattttgttaactttaagaaggagatatacatatggaagacgccaaa acataaagaaaggcccgccattctatccgctagaggatggaaccgctggagagcaactgcataaggctat gaagagatacgcctggttcttggaacaattgctttacagatgcacatatcgagggtgaacatcacgtacgcgga atacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgt tgcagtgaaaactcttcaattcttatgccgggttgggcgcgttattatcgagggttcagttgcgccgcgaac gacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgttttccaaaaagg gttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgattctaaaacggatta ccagggtttcagtcgatgtacacgttcgtcacatctcatctacctcccgggttttaataacacgattttgaccaga gtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaagggtgtggccc ttccgcatagaactgcctgcgtcagattctgcgatgccagagatcctatttttgcaatcaaatcattccggatactg cgattttaagtgtgttcattccatcacgggtttggaatgtttactacactcggatattgatattgtggttcgagtcg tcttaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgtgctagtacc aacctattttcattcttcgcaaaaagcactctgattgacaaatacatttatctaatttacacgaaattgctctgggg gcgcacctcttcgaaagaagtcggggaagcggttgcaaaacgctccattctccagggtacgacaaggat gggctcactgagactacatcagctattctgattacaccgagggggatgataaaccggcgcggtcggtaaaagt tgttccatttttgaagcgaagggttggtatctggataccgggaaaacgctgggcgttaatcagagaggcgaatta tgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattgacaaggatg gatggctacattctggagacatagcttactgggacgaagacgaacacttctcatagttgaccgctgaagtcttta attaatacaaaaggataccaggtggccccgctgaattggagtcgatattgtacaacacccaacatcttcgac gcgggcgtggcaggtctcccgacgatgacggcggtgaactcccgcggctgtgtttggagcacggaaa gacgatgacggaaaagagatcgtggattacgtcgccagtcgaagtaacaaccgcaaaaagttgcgcggagg agttgtgtttgtggacgaagtaccgaaaaggttaccgaaaactcagcgaagaaaaatcagagagatcctcat aaaggccaagaaggcggaaggtcgaattgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ΩNpA (SEQ ID NO: 23)	gagtattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaattacatattggaag acgcaaaaacataaagaaaggcccgccattctatccgctagaggatggaaccgctggagagcaactgc ataaggctatgaagagatacgcctggttcttggaacaattgctttacagatgcacatatcgagggtgaacatcac gtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacag aatcgtcgtatgcagtgaaaactcttcaattcttatgccgggttgggcgcgttattatcgagggttcagttgcg cccgcgaacgacatttataatgaacgtgaattgctcaacagatgaacatttcgcagcctaccgtagtgtttgttcc aaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaataatccagaaaattattatcatgattctaa aacggattaccagggtttcagtcgatgtacacgttcgtcacatctcatctacctcccgggttttaataacgatttt gtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaactcctctggatctactgggttacctaaggg tgtggccctccgcatagaactgcctgcgtcagattctgcagccagagatcctatttttgcaatcaaatcattcc ggatactgcgattttaagtgtgttcattccatcacgggtttggaatgtttactacactcggatattgatattgtggatt tcgagtcgtctaatgtatagatttgaagaagagctgttttacgatccctcaggattacaaaattcaaatgctgttg ctagtaccaaccctattttcattcttcgcaaaaagcactctgattgacaaatacatttatctaatttacacgaaattgc ttctgggggcgcacctcttcgaaagaagtcggggaagcggttgcaaaacgctccattctccagggtacgac aaggatattgggctcactgagactacatcagctattctgattacaccgagggggatgataaaccggcgcggtc ggtaaaagttgtccatttttgaagcgaagggttggtatctggataccgggaaaacgctgggcgttaatcagagag gcgaattatgtgtcagaggacctatgattatgtccggttatgtaacaatccggaagcgaccaacgccttgattga caaggatggatggctacattctggagacatagcttactgggacgaagacgaacactcttcattagttgaccgcttg aagtctttaattaatacaaaaggataccaggtggccccgctgaattggagtcgatattgtacaacacccaaca tcttcgacgcgggcgtggcaggtctcccgacgatgacggcggtgaactcccgcggctgtgtttttggagc acggaaagacgatgacggaaaagagatcgtggattacgtcgccagtcgaagtaacaaccgcaaaaagttgc gcggaggagttgtgtttgtggacgaagtaccgaaaaggttaccgaaaactcagcgaagaaaaatcagaga gatcctcataaaggccaagaaggcggaaggtcgaattgtaa
Luciferase ORF (SEQ ID NO: 24)	atggaagacgcaaaaacataaagaaaggcccggaaccgctggagagcaactgcataaggctatgaagaga tacgcctggttcttggaacaattgctttacagatgcacatatcgagggtgaacatcacgtacgcggaatacttcg aaatgtccgttcggttggcagaagctatgaaacgatatgggctgaatacaaatcacagaatcgtcgtatgcagtg

[illegible]

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	gggcggaagtccaattgtaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaa
Luciferase (Protein) (SEQ ID NO: 30)	MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIE VNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSNSLQFFMPVLGA LFIGVAVAPANDIYNERELLSNMNISQPTVVVFVSKKGLQKILNVQKKLP IIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALI MNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVPFHH GFGMFTTLGYLICGFRVVLMYRFEELFLRSLQDYKIQSALLVPTLFSF FAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLT ETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELC VRGPMIMSGYVNNPEATNALIDKDWLHSGDIAYWDEDEHFFIVDRL KSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDDAGELPAVVV LEHGKTMTEKEIVDYVASQVTTAKKLRGGVVVFDEVKGLTGKLDAR KIREILIKAKKGGKSKL
GFP (ORF) (SEQ ID NO: 31)	atggagaaaaaatcactactcctccacccaccaccgccgtattccagagggactagatatcttgcgcagcct agtggcaatactagtcttagtgccctaatacgaaggcctcaaaaggcccccaaaagccttcacagaacctagtcct gtcactccctcaacaactaagtctttaaaaatgcgccagcgcagcatccatggtgagcaagggcgaggagct gttcacgggggtgtgcccactctggtcagctggacggcgacgtaaacggccacaagttcagcgtgcggg cgaggcgagggcgatgccaccaacggcaagctgacctgaagttcatctgcaccaccggcaagctgcccgt gccctggcccaccctctgtgaccacctgacctacggcgtgcagtgcttcagcgcgtaccccgaccacatgaag cagcacgacttctcaagtccgccatgccgaaggctacgtccaggagcgcaccatcaccttaaggacgacg gcacctacaagacccgcgcgaggtgaagttcaggggcgacacctggtgaaccgcatcgagctgaagggc atcgactcaaggaggacggcaacatctggggcacaagctggagtacaacttcaacagccacaacgtctatat caccggcgacaagcagaagaacggcatcaaggccaacttcaagatccgccacaacgtggaggacggcagc gtgcagctcggcgaccactaccagcagaacacccccatcggcgacggccccgtgctgctcccgacaacca ctacctgagcaccagtcgaagctgagcaaaagaccccaacgagaagcgcgacacatggtcctgctggagttc gtgaccgccggggatcactcaggcatggacgagctgtacaaggagctcgcatgggtcaccaccatcat catcattaa
GFP Expression Construct (Table 3) (SEQ ID NO: 32)	acgctgcccagatctcgatccgcgaattaatacactactataggagatattttacaacaattaccaacaa caacaacaacaacaacattacaattactattfacaattacatatgagcaaaaggtgaagaactgttaccggcgt gtgccattctgtggaactggatggcgtgtgaacggtcacaaatcagcgtgctggtgaaggtgaaggcg atgccacgattggcaaaactgacgctgaaatttatctgcaccaccggcaaaactgcccgtgctggccgacgctg gtgaccacctgacctatggcgttcagtggttttagtcgctatccggatcacatgaaacgtcacgattctttaaact gcaatgccggaaggctatgtgcaggaacgtacgattagctttaaagatgatggcaaatataaaacgcgcgccgt tgtgaaatttgaaggcgataccctggtgaaccgattgaactgaaaggcacggattttaaagaagatgcaatat cctgggccataaaactggaatacaacttaataagccataatgtttatattacggcgataaaacagaaaaatggcatc aaagcgaattttaccgttcgccataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaatacc ccgattggtgatggtccggtgctgctgcccggataatcattatctgagcacgcagaccgttctgtctaaagatccga acgaaaaaggcacgcgggaccacatggtctgcacgaatatgtgaatgcggcaggtattacgtggagccatcc gcagttcgaaaaataaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaa
GFP (Protein) (SEQ ID NO: 33)	MEKKITTPSTPPPPYSRGTRYLAQPSGNTSSSALMQGQKAPQKPSQNLV PVTPSTTKSFKNAPAPGSMVSKGEELFTGVVPILVELDGDVNGHKFSV RGEGEDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDH MKQHDFFKSAMPEGYVQERTITFKDDGTYKTRAEVKFEGLTLVNRIEL KGIDFKEDGNILGHKLEYNFNHNVYITADKQKNGIKANFKIRHNVED GSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMLV LEFVTAAGITHGMDELYKELGMGHHHHHH

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
CAT (ORF) (SEQ ID NO: 34)	atggaaaaaaaaatcaccggctacaccaccgttgacatctctcagtgaccgttaaagaacactttgaagcgttc cagctctgtcgcgcagtgtagcttacaaccagaccgttcagctagacatcaccgcgttcctgaaaaccgttaaaaaa aacaacacaaaattctaccggcggttcattcacatcctggcgcgtctgatgaacgcgcacccggaatttcgtatg gcatgaaagacggggaactgggtatctgggactctgtcaccgtgctacaccgtttccacgaacagaccgaa accttcttctctgtggtctgaataccacgacgacttccgtcagttcctgcacatctactctcaggacgttgctgc tacggtgaaaacctggcgacttcccgaagggttcatcgaacatgttcttcgtttctgcgaaccctgggtttct ttcactcttgcacctgaacgtggcgaaacatggacaacttctgcgccggtttcactatgggtaaatactacac ccagggtgacaaagtctgatgccgtggcgatccagggtcaccacgcggttgcgacggtttccacgttggtcg tatgctgaacgaactccagcagattgacgacgaatggcaggggtggtgcgtaa
CAT Expression Construct (Table 3) (SEQ ID NO: 35)	acgctgcccagatctcgcacccgcgaattaatacgaactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacataggaaaaaaaaaatcaccggctacaccaccgt tgacatctctcagtgaccgttaaagaacactttgaagcgttcagctctgtcgcgcagtgtagcttacaaccagac cgttcagctagacatcaccgcgttctctgaaaaccgttaaaaaaacaacaaaattctaccggcggttcattcac atcctggcgcgtctgatgaacgcgcacccggaatttcgtatggcgatgaaagacggggaactgggtatctggga ctctgttcaccgtgctacaccgtttccacgaacagaccgaaaccttctctctgtggtctgaataccacgacg actcctcagttcctgcacatctactctcaggacgttgctgctacggtgaaaacctggcgacttcccgaagg ttcatcgaacatgttcttcgtttcgcgaaccctgggttttccactcttgcacctgaacgtggcgaaacatg gacaacttctgcgccggtttcactatgggtaaatactacaccagggtgacaaagtctgatgccgtggcgaa tccagggtcaccacgcggttgcgacggtttccacgttggtcgtatgctgaacgaactccagcagattgacgacg aatggcaggggtggtgcgtaaactgagatccggctgtaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaa
CAT (Protein) (SEQ ID NO: 36)	MEKKITGYTTVDISQWHRKEHFEAFQSV AQCTYNQTVQLDITAF LKTV KKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVF HEQTETFSSLWSEYHDDFRQLHIYSQDVACYGENLAYFPKGF IENMF FVSANPWVSFTSFDLNVANMDNFFAPVFTMGKYYTQGDKVL MPLAIQ VHHA VCDGFHVGRMLNELQQYCDEWQGGA
Superfolder GFP (ORF) (SEQ ID NO: 37)	atgagcaaaagtgagaactgtttaccggcggttgccgattctggtggaactggatggcgatgtgaacgggtca caaattcagcgtgctggtgaagggtgaaggcgatgccacgattggcaaaactgacgctgaaattatctgcacca ccggcaaaactgccgggtgcccgtggcgacgctggtgaccaccctgacctatggcggttcagtggtttagtcgctatc cggatcacatgaaacgtcacgatttcttaaatctgcaatgccggaaggctatgtgcaggaacgtacgattagcttt aaagatgatggcaatataaaacgcgcgccgttgtaaatgtgaaggcgataccctgggtgaaccgcattgaact gaaaggcacggattttaaagaagatggcaatatcctgggccataaactggaatacaacttaataagccataatggt tatattacggcggataaacagaaaaatggcatcaaagcgaattttaccgttcgcataacgttgaagatggcagt gtgcagctggcagatcattatcagcagaataccccgattggtgatggtccggtgctgctccggataatcattatc tgagcacgcagaccgttctgtctaaagatccgaacgaaaaaggcacgcgggaccacatggttctgcacgaata tgtgaatgcggcaggtattacgtggagccatccgcagttcgaaaaataa
Superfolder GFP Expression Construct (Table 4) (SEQ ID NO: 38)	acgctgcccagatctcgcacccgcgaattaatacgaactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattaaaaaaatgagcaaaagggtgaagaactgtttaccgg cgttggtgccgattctggtggaactggatggcgatgtgaacggtcacaaattcagcgtgctggtgaagggtgaag gcatgcccagattggcaaaactgacgctgaaattatctgcaccaccggcaaaactgccgggtgccgtggcgac gctggtgaccaccctgacctatggcggttcagtggtttagtcgctatccggatcacatgaaacgtcacgatttcttaa atctgcaatgccggaaggctatgtgcaggaaacgtacgattagctttaaagatgatggcaatataaaacgcgcg ccgttgtaaatgtgaaggcgataccctgggtgaaccgattgaactgaaaggcacggattttaaagaagatggca atatacctgggccataaactggaatacaacttaataagccataatgttatattacggcggataaacagaaaaatggc atcaaagcgaattttaccgttcgcataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaata ccccgattggtgatggtccggtgctgctgccgataatcattatctgagcacgcagaccgttctgtctaaagatcc gaacgaaaaaggcacgcgggaccacatggttctgcacgaatatgtgaatgcggcgaggtattacgtggagccat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	ccgcagttcgaaaaataaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa
Superfolder GFP (protein) (SEQ ID NO: 39)	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATIGKLTCLKFICT TGKLPVPWPTLVTTLTYGVCFSRYPDHMKRHDFFKSAMPEGYVQER TISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYN FNHNHYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDG PVLLPDNHYLSTQTVLSKDPNEKGTRDHMVLHEYVNAAGITWSHPQF EK
ANX scFv (ORF) (SEQ ID NO: 40)	atggctcaggtacagttacaacaatcaggattagaactcgtaaaaccaggtgctagtgttaaatctctcgaaaa caagtgggtatacttttacagaatatacagatgcattgggtgtaaaacaatctcatggtgaaagttagaattgatcgga ggcatcaacccaacaatggcggcactcttataatcaaaatttaaggcaagcaatccttacagtcgacaaat cttcatccaccgcctatctcgaattacgtagttaacatcagaagattcagccgtttattattgctgcacgtgacgac gttatccagcttggttgcttattgggggcaaggtaccacagttacagtatcctctggcggcggcggatctggcgg cggcggctccggtggaggtggttctactgacatccaactgactcaatctcctcatcccttctgcatcacttggcg aacgcgtctcgtacacgtgctgcgcactcgaagacatcgatcaaatftaaattggctcaacaaaaacctgatgg cacgatcaaacgccttattatgccactcctctctcgatagcggcgtcccgaaacgttttctggtctcgtagcgg ttcagactattcattgacaatcagctcactcgaaagcgaagactttgtagattactattgccttcagtatgccagcag cccactacatttggcggaggtacaaaattggaattaaacgcgcggccgcatggagccatccgcagttcgag aataaa
ANX scFv Expression Construct (Table 4) (SEQ ID NO: 41)	acgctgcccagatctcgatcccgcaaaataatcagactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacataggctcaggtacagttacaacaatcaggatta gaactcgtaaaaccaggtgctagtgttaaaatctctcgaaaaacaagtggttatactttacagaatatacagtgca ttgggtgtaaaacaatctcatggtgaaagttagaattgagtcggagggcatcaacccaacaatggcggcactcttat aatcaaaaatttaaggcaagcaatccttacagtcgacaaatctcatccaccgcctatctcgaattacgtagtta acatcagaagattcagccgtttattattgctgcacgtgacgacgttatccagcttggttgcttattgggggcaaggt accacagttacagtatcctctggcggcggcggatctggcggcggcggctccggtggaggtggttctactgacat ccaactgactcaatctcctcatcccttctgcatcacttggcgaacgcgtctcgtacacgtgctgcgcacatcaag acatcgatcaaaatttaattggcttcaacaaaaacctgatggcacgatcaaacgccttattatgccacctcctctc tcgatagcggcgtcccgaaacgttttctggtctcgtagcgggtcagactattcattgacaatcagtcactcgaa agcgaagactttgtagattactattgccttcagtatgccagcagccacctacatttggcggaggtacaaaattgg aaattaaacgcgcggccgcatggagccatccgcagttcgagaataaactgagatccggctgtaacaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ANX scFv (Protein) (SEQ ID NO: 42)	MAQVQLQSQGLELVKPGASVKISCKTSGYTFTEYTMHWVKQSHGKSL EWIGGINPNNGGTSYNQKFKGKAILTVDKSSSTAYLELRSLTSEDSAVY YCARDDRYPAWFAYWGQGTITVTVSSGGGGSGGGGSGGGGSTDIQLT QSPSSLSASLGERVSITCRASQDIGSNLNLWLQKPDGTIKRLIYATSSLD SGVPKRFSGRSGSDYSLTISSLESEDFVDYYCLQYASSPPTFGGGTKLE IKRAAAWSHPQFEK
BOT scFv (ORF) (SEQ ID NO: 43)	atggctcaagttcagttacaagaatctggtggcggttagttcaaccaggtggctctctcgtctctcatgtccgc atcggggttcaccttttctgaccattatgtactgggtccgtcaagcggcgcaaggactgaatgggtagca acaatctctgatggtggtcttatacctattactctgactcagtcgaaggtcgtttacaactctcgtgataactcaa aaaatactctctatttacaatgaacagcttacgtgccgaagatactgcaattattattgtcccggtatcgttatgac gacgctatggattattggggccaagtcacttttagtaacagttcatccgggtggcggcggcggcgggtgg ctctggcgggtgggaagtacagaaattgtttaactcagagtcggcgacattatcactctccccggcggaactg gtacaatctcctgtcgtgcctctgaaagcgtagattcatacggacactcctttatgcatgggtatcaacaaaaacc gggacaagcaccacgtctcttaattatcgtgcatcaacttagaacctgggattccagccggttcagtggtctctg gatcaggtaccgattttacattaaccatctctagtttgaaccagaagacttcgagttattattgccagcaaggaa

[illegible]

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	YYCARQGDFGDWYFDVWGAGTTVTVSSGGGSGGGGSGGGGSTDV VLTQTPLSLPVSLGDQASISCRSSQSLVHSNGNTYLHWY LQKPGQSPKL LIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCQQSTHVP WTFGGGTKLEIKAAAWSHPPQFEK
MS2-A scFv (ORF) (SEQ ID NO: 49)	atggctgaagtaaattagtagaaagtgggtgggtttggtaaacctgggtgatctctaaactctcgtgcaaagc aagcggctttactttttctcttatgctatgcatgggtccgcaaacctccgaaaaacgcttagaatgggtagcaac aatttcaacaggcggaggctatacatatttccagattctgttaaaggcgctttacaatttccgcgataatgcga aaaatgctttatatttacaatgaaatccttacgttcagaaagacacagctacgtattattgtgctcgtcaaggcgactt tggtgattgggtacttcgatgtatggggcgaggcagacagttacagtatcttcaggcggcggttctggtgg cgggtggctccgggtgggtgggaagcacggatgtgtactgacccaaactccctatttaccagtcctattaggc gatcaagcaaccatttcgtcgtcttctcaatctctgttactcctaagcgaatactacttaccattggtatcttcaa aaaccaggccaatctcctaactccttattataaagttcaaatcgttttcaggcgtccagatcgttttccggctc cggcagtggcaccgattttacctaataaatttctcgtgtagaagctgaagacttaggtgtatattttgcttcaatca acttacgttccctggacttttgggtgggtacgaaattagaataaagcggccgcatggagccatccgcagttcg aaaaataa
MS2-A scFv Expression Construct (Table 4) (SEQ ID NO: 50)	acgctgcccagatctcgcacccgcgaaattaatacgaactactataggagatttttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacatatggctgaagtaaattagtagaaagtgggtgt ggtttgtaaaacctgggtgatctcttaaaactctcgtgcaaagcaagcggctttacttttcttcttatgctatgcatg ggctccgtcaaacctcccgaataacgcttagaatgggtagcaacaatttcaacaggcggaggtatataatttccc agattctgttaaaggcgctttacaatttcccgataatgcgaaaaatgctttatatttacaatgaaatccttacgt tcagaagacacagctacgtattattgtgctcgtcaaggcgactttgggtgattgggtacttcgatgtatggggcgag gcacgacagttacagtatcttcaggcggcggttctggtggcggtggctccgggtgggtgggaagcacggga tgtgtactgacccaaactccctatttaccagtcctattaggcgatcaagcaaccatttcgtcgtcttctcaat ctctgttactcctaagcgaatactacttaccattggtatcttcaaaaaccaggccaatctcctaactccttattata aagtttcaaatcgttttcaggcgtccagatcgttttccggctccggcagtggcaccgattttacctaataaatttct cgtgtagaagctgaagacttaggtgtatattttgcttcaatcaacttacgttccctggacttttgggtgggtacga aattagaataaagcggccgcatggagccatccgcagttcgaaaaataaactgagatccggctgctaacaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
MS2-A scFv (Protein) (SEQ ID NO: 51)	MAEVKLVESGGGLVPGGSLKLSCKASGFTFSSYAMSWVRQTPEKRL EWVATISTGGGYTYFPDSVKGRFTISRDNALYLMKSLRSEDAT YYCARQGDFGDWYFDVWGAGTTVTVSSGGGSGGGGSGGGGSTDV VLTQTPLSLPVSLGDQATISCRSSQSLVHSNGNTYLHWY LQKPGQSPK LLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCFQSTYV PWTFGGGTKLEIKAAAWSHPPQFEK
2E2 scFv (ORF) (SEQ ID NO: 52)	atggctgaagtgcagttggtgaatcaggtgggggttagtacagccgggtggtagtttacgtttgcatgtgcgg catcaggttttttttagtagtgattggtgaattgggtacgtcaagcaccgggaaaaggattagaatgggtggc gaatattaatcaagatggttcagaaaaatattatgttgattcagttaaaggctgtttacaatcagccgtgacaacgc acaaaatagcttatacttacaatgaacagtttacggcgagaagacacagcagtagtatattgtgcaaaggaatta gggccgtgggggcaaggacattatgtacgggtgagtagcgggggagggggcagcggcggtggtggttcgg gagggggaggttcacacagcagtagttatcaggaaagcgactcacgacatctccgggggggacgggtta ttctcattgccgcagcagtagcaggaacgattacgacttctaactatgcaattgggtccagaaaaaacgaatca tgtgtttacgggttaattggggcaacgagcattcgcgcggcgagtgccgggtacgttttagcgggtttctattg gtgaaaggcagcattactattacaggagcgcaaacgaagatgatgctatgtattttgcgcgttatggtataa cacacactatgttttggaggtggcacgaagggtacagttatggggcaagcggccgcatggagccatccgcagtt tcgagaaataa
2E2 scFv Expression Construct (Table 4)	acgctgcccagatctcgcacccgcgaaattaatacgaactactataggagatttttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacatatggctgaagtgacgttgggtgaatcaggtggg

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
(SEQ ID NO: 53)	ggtttagtacagccgggtgtagtttacgtttgtcatgtcgccgcatcaggtttatttttagtagtgattggatgaattg ggtacgtcaagcaccgggaaaaggattagaatgggtggcgaatattaatcaagatggttcagaaaaatattatgt ggattcagttaaaggctgtttacaatcagccgtgacaacgcacaaaatagcttatacttacaatgaacagtttac gggcagaagacacagcagtatattattgtgcaaaggaattagggccgtgggggcaaggacattagtgacggt gagtagcgggggagggggcagcggcgggtggttcgggagggggaggttcgacacaggcagtagttattc aggaaagcgactcacgacatctccgggggggacgggtattctcactgccgcagcagtagcaggaacgattac gacttctaactatgcaaattgggtccagaaaaaacgaatcatgtgtttacgggtttaattggggcaacgagcattc gcgcgccgggagtgccggtagcttttagcgggtttcttattggtggaaaggcagcattaactattacaggagcgc aaaccgaagatgatgctatgtattttgcgcgttatggtataacacacactatgttttggaggtggcacgaaggta cagtattggggcaagcggccgcatggagccatccgcagttcgagaaataaactgagatccggctgctaacaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
2E2 scFv (Protein) (SEQ ID NO: 54)	MAEVQLVESGGGLVQPGGSLRLSCAASGFISSDWMNWRQAPGKGL EWVANINQDGSEKYYVDSVKGRFTISRDNQNSLYLQMNSLRAEDTA VYYCAKELGPWGQGLTVTVSSGGGGSGGGSGGGGSTQAVVIQESAL TTSPGGTVILTCRSSTGTITTSNYANWVQKKPNHVFTGLIGATSIRAPG VPVRFSGFLIGGKAALTITGAQTEDDAMYFCALWYNTHYVFGGGTKV TVLGQAAAWSHPPQFEK
2E2-3d scFv (ORF) (SEQ ID NO: 55)	atggctgaagtgcagttggtgaatcaggtgggggttagtacagccgggtgtagtttacgtttgtcatgtaagg catcaggtttatttttagtagtgattggatgaattggtccgtcaagcaccgggaaaaggattagaatgggtggcg aatattaatcaagatggttcagaaaaatattatgtggattcagttaaaggctgtttacaatcagccgtgacaacga caaaataccttatacttacaatgaacagtttacgggcagaagacacaggagtagtattattgtgcaaaggaattag ggcctgtggggcaaggacattagtgacggtagtagcgggggagggggcagcggcggtggtggttcggg agggggaggttcgacacaggcagtagttactcaggaaagcgactcacgacatctccgggggggacgggttac tctcactgccgcagcagtagcaggaacgattacgacttctaactatgcaaattgggtccagaaaaaacgaatcat gtgtttacgggtttaattggggcaacgagcattcgcgcgccgggagtgccggtacgttttagcgggtctcttattg gtggaaaggcagcattaactattacaggagcgcaaaccgaagatgatgctatgtattttgcgcgttatggtataa cacacactatgttttggaggtggcacgaagggtacagtattggggcaagcggccgcatggagccatccgcagt tcgaaaaataa
2E2-3d scFv Expression Construct (Table 4) (SEQ ID NO: 56)	acgctgcccgagatctcgatccgcgaaattaatacgaactactatagggagtagttttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacataggtcgaagtgtagttggtgaatcaggtggg ggttttagtacagccgggtgtagtttacgtttgtcatgtaaggcatcaggtttatttttagtagtgattggatgaattg gtccgtcaagcaccgggaaaaggattagaatgggtggcgaatattaatcaagatggttcagaaaaatattatgt ggattcagttaaaggctgtttacaatcagccgtgacaacgcacaaaataccttatacttacaatgaacagtttac gggcagaagacacaggagtagtattattgtgcaaaggaattagggccgtgggggcaaggacattagtgacggt gagtagcgggggagggggcagcggcgggtggtggttcgggagggggaggttcgacacaggcagtagttact caggaaagcgactcacgacatctccgggggggacgggtactctcactgccgcagcagtagcaggaacgatta cgacttctaactatgcaaattgggtccagaaaaaacgaatcatgtgtttacgggtttaattggggcaacgagcatt cgcgcgccgggagtgccggtacgttttagcgggtctcttattggtggaaaggcagcattaactattacaggagc gcaaaccgaagatgatgctatgtattttgcgcgttatggtataacacacactatgttttggaggtggcacgaagg ttacagtattggggcaagcggccgcatggagccatccgcagttcgaaaaataaactgagatccggctgctaaca aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
2E2-3d scFv (Protein) (SEQ ID NO: 57)	MAEVQLVESGGGLVQPGGSLRLSCKASGFISSDWMNWRQAPGKGL EWVANINQDGSEKYYVDSVKGRFTISRDNQNTLYLQMNSLRAEDTG VYYCAKELGPWGQGLTVTVSSGGGGSGGGSGGGGSTQAVVTQESA LTTSPGGTVTLTCRSSTGTITTSNYANWVQKKPNHVFTGLIGATSIRAP GVPVRFSGSLIGGKAALTITGAQTEDDAMYFCALWYNTHYVFGGGTK VTVLGQAAAWSHPPQFEK

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
HPV16L1 scFv (ORF) (SEQ ID NO: 58)	atgagcctgtggctgccagcgaggccaccgtgtacctgccccctgtcccgtgagcaaggtggtgagcacc gacgagtacgtggcaggaccaactctactaccacgccggcaccagcaggtgtgcccgtggccacccc ctacttccccatcaagaagcccaacaacaagaatcctgtgtcccgaaggtgagcggtctgcagtacaggggtg ttcaggatccacctgcccgaacccaagaagttcggcttccccgacaccagcttctacaaccccgacaccagag gctggtgtgggcctgctgtgggcgtggaggtgggcagggggccagccccctgggctgtggcatcagcgggccac cccctgctgaacaagctggacgacaccgagaacgccagcgcttacgccgcaacgccggcggtggacaaca gggagtgcatcagcatggactacaagcagaccagctgtgctgacgggtgcaagcccccatcggcgagc actggggcaagggcagccccctgcaccaacgtggcctgaaccccgcgactgccccccctggagctgac aacaccgtgatccaggacggcgacatgtgtggacaccggcttcggcgccatggacttaccacctgcaggcc aacaagagcgaggtgccccctggacatctgcaccagcatctgcaagtaccccgactacatcaagtgtgtgagcg agccctacggcgacagcctgttcttctacctgaggagggagcagatgttctgtgaggcacctgttcaacaggggcc ggcgccgtggcgagaaagtgtcccgcacacctgtacatcaagggcagcggcagcaccgccaacctggcca gcagcaactacttccccacccccagcggcagcatgtgtgaccagcgacgccagatcttcaacaagccctactg gctgcagagggcccaggggccacaacaacggcatctgctggggcaaccagctgttctgtgacctgtgtggacac caccaggagcacaacatgagcctgtgcgcccatcagcaccagcgagaccacctacaagaacaccaactt caaggagtacctgaggcacggcgaggagtacgacctgcagttcatcttcagctgtgcaagatcacctgacc ggcgacgtgatgacctacatccacagcatgaacagcaccatcctggaggactggaaacttcggcctgcagcccc ccccggcgccacctggaggacacctacaggttcgtgaccagccagggcatcgctgccagaagcacacc ccccccgccccaaaggaggacccccgaagaagtacaccttctggaggtgaacctgaaggagaagtgcagc ggcgacctggaccagttccccctgggcagggaagttcctgtgtagggccggcctgaaggccaagcccaagttc accttgggcaagagggaagggccacccccaccaccagcagcaccagcaccaccgccaaggaagaagagg aagctgtga
HPV16L1 scFv Expression Construct (Table 4) (SEQ ID NO: 59)	acgtgtcccagagatctcgatcccgcaaaattaatacactactatagggagtattttacaacaattaccaaca caacaacaacaacaacattacaattactatttacaattacatagagcctgtggctgcccagcgaggccaccgt gtacctgccccctgtcccgtgagcaaggtgtgtgagcaccgacgagtagctggccaggaccaactctacta ccacggcgccaccagcaggtgtgtgcccgtgggccaccctacttccccatcaagaagcccaacaacaaca gatcctgtgtcccgaaggtgagcggtctgcagtacaggtgttccaggatccacctgcccgaacccaagaagttc gggttccccgacaccagcttctacaaccccgacaccagaggtgtgtgtggcctgctgtggcggtggaggtg ggcagggggccagccccctgggcgtgggcatcagcgccacccccctgtgaacaagctggacgacaccgaga acgccagcgcttacggcgccaacggcggtggacaacaggagtgatcagcatggactacaagcagacc cagctgtgctgacgggtgcaagcccccatcgcgagcactggggcaagggcagccccctgcaccaacgt ggcctgtgaacccccggcgactgccccccctggagctgatcaaacctgtatccaggacggcgacatggtgg acaccggcttcggcgccatggacttcaccacctgcaggccaacaagagcgaggtgccccctggacatctgca ccagcatctgcaagtaccccgactacatcaagatggtgagcgagccctacggcgacagcctgttcttctacctg aggagggagcagatgttctgtgaggcacctgttcaacaggcgccggcgctgggcgagaacgtgcccacga cctgtacatcaaggcagcgggcagcaccgccaacctggccagcagaactacttccccacccccagcgga gcatgtgtgaccagcgacggccagatcttcaacaagccctactggctgcagagggcccaaggccacaacaac ggcatctgctggggcaaccagctgttctgtgaccgtgtgtggacaccaccaggagcaccacatgagcctgtgc ggcgccatcagcaccagcgagaccacctacaagaacaccaactcaaggagtacctgaggcacggcgagga gtacgacctgcagttcatcttcagctgtgcaagatcacctgaccggcgacgtgatgacctacatccacagcat gaacagcaccatcctggaggactggaacttcggcctgcagccccccccggcgggcaccctggaggacacct acagggttcgtgaccagccaggccatcgctgcagaagcacaccccccccgccccaaaggaggacccccctg aagaagtacaccttctgggaggtgaacctgaaggagaagttcagcgccgacctggaccagttccccctgggca ggaggttcctgctgaggccggcctgaaggccaagcccaagttcacctgggcaagagggaagggcacccccc accaccagcagcaccagcaccaccgccaagagggaagaagagggaagctgtgaactgagatccggctgtaac aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
HPV16L1 scFv (Protein) (SEQ ID NO: 60)	MSLWLPSEATVYLPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHP YFPIKKPNNKILVPKVSGLQYRVFRIHLPDPNKFPGPDTSFYNPDTQR LVWACVGVGVGRGQPLGVGISGHPLLNLDDTENASAYAANAGVDN RECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	TVIQDGDMDVDTGFGAMDFTTLQANKSEVPLDICTSICKYPPDYIKMVSE PYGDSLFFYLRRQMFVRHLFNRAVGENVPDDL YIKGSGSTANLAS SNYFPTPSGSMVTSDAQIFNKPYWLQRAQGHNNGICWGNQLFVTVVD TTRSTNMSLCAAISTSETTYKNTNFKEYLRHGEEYDLQFIFQLCKITLA DVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTRYFVTSQAIACQKHTP PAPKEDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFT LGKRKATPTTSSTSTTAKRKRKL
pK7LUC (SEQ ID NO: 61)	tcgacgggatcggtccactgagcgtcagaccccgtagaaaagatcaaggatcctcttgagatcctttttctgcgc gtaatctgctgcttgcaaacaaaaaacaccgctaccagcgggtggttggtgcccggatcaagagctaccaact cttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtcctctagtgtagccgtagtaggcca ccacttcaagaactctgtagcacgcctacatacctcgctctgctaactcgtgtaccagtggtgctgccagtggc gataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctgggctgaacgg ggggttcgtgcacacagcccagcttgagcgaacgacctaccgaactgagatacctacagcgtgagcattg agaaagcggcacgcttcccgaaggagagaaaggcggacaggtatccggtaagcggcagggtcggaaacagga gagcgcacgagggagcttccagggggaaacgcttggtatcttatagtcctgtcgggttcgccacctgtgactt gagcgtcgattttgtgatgtcgtcagggggggcggagcctatggaaaaacggcgaacggcgcttttacg gttctcggccttttctggccttttctcacatgttcttctgctggtatccctgattctgtggataaccgtattaccg ctttgagtgagctgataccgctcgccgcagcggacgaccgagcgcagcgagtcagtgagcggaggaagcgg aagaagctcgacgccaatcgcgaacccgctctccccgcgcgttgccgattcattaatgcagctggcacga cagggttcccgactggaaagcgggcagtgagcgaacgcaattaatgtgagttagtcactcattaggcacccc aggctttacactttatgcttcggctcgtatgtgtgtggaattgtgagcggataacaatttcacacaggaacagct atgacctgattacgaattcagatctcgatcccgcaaatataacgactcactataggagaccacaacgggttc cctctagaaataattttgtaactttaagaaggagatatacatatggaagacgccaaaaacataaagaaaggccc ggcgccattctatccgctagaggatggaaccgctggagagcaactgcataaggctatgaagagatacgccctg gttcttgaacaattgctttacagatgcacatatcgagtggaacatcacgtacgcggaatacttcgaaatgtccgt tcggttggcagaagctatgaacgatatgggctgaatacaatcacagaatcgtcgtatgcagtgaaaactctctt caattctttatgccggtgttgggcgcgttatttatcgaggttcgagttgcgccgcgaacgacatttataatgaacgt gaattgtcaacagtatgaacatttcgcagcctaccgtagtgtttgtttcaaaaaagggttgcaaaaaattttgaac gtgcaaaaaaattaccaataatccagaaaattattatcatggtattctaaaacggattaccagggttcagtcgat gtacacgttcgtcacatctcatctacctcccggttttaataatgaatacgtttgtaccagagtcctttgatcgtgacaaa acaattgcactgataatgaactcctctggatctactgggttacctaaagggtgtggtgcccttcgcatagaactgcctg cgtcagattctcgcatgccagagatcctattttggcaatcaatcattccggatactgcgattttaaagtggtgtccat tccatcacgggttttggaatgtttactacactcggatatttgatatgtggatttcgagtcgtcttaattgtatagatttgag aagagctgttttacgatccctcaggattacaaaattcaaagtcggttgctagtaccaaccctattttcattcttcgcc aaaagcactctgattgacaatacagatttatcaatttacacgaaattgcttctgggggcgcacctcttcgaaaga agtcggggaagcgggttgcaaaacgcttccatctccaggggatacgacaaggatatgggctcactgagactacat cagctattctgattacacccgagggggatgataaacggggcgggtcggttaaagttgttcatttttgaagcgaa ggttgtggtatcgataccgggaaacgctgggcgttaatcagagaggcgaaattatgttcagaggacctatga ttatgtccgggttatgtaaacatccggaaagcgaccaacgccttgattgacaaggatggatggctacattctggaga catagcttactgggacgaagacgaacacttctcatagttgaccgctfgaagctttaaataaatacaaaaggatacc agggtggccccgctgaattggagtcgatattgttacaacaccccaacatctcgacgcgggcgtggcaggtcttc ccgacgatgacgccggtgaactccccgcgcggtgtgttttgaggcacggaaagacgatgacggaaaaaga gatcgtggattacgtcgccagtcagtaacaaccgccaaaaagttgcgcggaggagttgtgttggacgaag taccgaaaggtcttaccgaaaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaaggggcg gaaagtcgaattgtaagtcgaccggctgctaacaagccgaaagggaagctgagttggctgctgccaccgct gagcaataactagcataacccctggggcctctaaacgggtcttgagggggtttttgctgaaaggaggaactatat ccggataacctcgagctgcagggcatgcaagcttggcactggccgtcgtttacaacgtcgtgactgggaaaac cctggcggttaccgaacttaatgccttgcagcacatcccccttcgccagctggcgtaatagcgaagaggccccg caccgatcgccctcccaacagttgcgcagcctgaatggcgaatcgatttatcaacaagccgccgtcccgtc aagtcagcgtaatgtctgccagtgttacaaccaattaaccaattctgattagaaaaactcatcgagcatcaaatga aactgcaattttatcatatcaggattatcaatfaccatattttgaaaaagccgtttctgtaatgaaggagaaaactcac

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	cgaggcagttccataggatggcaagatcctggtatcggctgcgattccgactcgtccaacatcaatacaacctat taatttcccctcgtaaaaaataagggttatcaagtgagaaatcaccatgagtgacgactgaatccgggtgagaatggc aaaagcttatgcatttcttccagacttggtcaacaggccagccattacgctcgtcatcaaaatcactcgcataacc aaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgtgttaaaaaggacaattacaaca ggaatcgaatgcaaccggcgaggaacactgccagcgcatcaacaatatttcacctgaatcaggatattctct aatacctggaatgctgtttccggggatcgcagtggtgagtaaccatgcatacaggagtagcgataaaatgc ttgatggtcgaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgct acctttgccatgtttcagaacaactctggcgatcgggcttccatacaatcgatagattgtcgacactgattgcc cgacattatcgcgagccatttataccataataatcagcatccatgttggaatttaatcgggcttcgagcaagac gtttccggttgaaatggctcataacaccccttgattactgttatgtaagcagacagtttattgttcatgatgatata ttttatctgtgcaatgtaacatcagagattttgagacacaacgtggcttgttgataaatcgaacttttgctgagttg aaggatcagatcacgcatttcccgacaacgcagaccgttccgtggcaaaagcaaaagtcaaaatcaccaactg gccacctacaacaaagctctcatcaaccgtggctccctcactttctggctggatgatggggcgattcaggcctg gtatgagtcagcaacacctcttcacgaggcagacctc
pET23LucA-f (SEQ ID NO: 62)	ggtggtcatatggaagacgccaaaaacat
pET23LucA-r (SEQ ID NO: 63)	ggtggctcgcagttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttac aatttggactttccgc
pET23c (SEQ ID NO: 64)	gaatgacttggttgagtactcaccagtcacagaaaagcatcttaccggatggcatgacagtaagagaattatgcag tgctgccataaccatgagtgataaactcgggccaacttactctgacaacgatcggaggaccgaaggagctaa ccgctttttgcacaacatgggggatcatgtaactcgcctgacgttggaaccggagctgaatgaagccatacc aaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaactattactggcgaacta cttactctagcttccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctc gcccctccggctggctgtttattgctgataaatctggagccggtagcgtgggtctcgcggatcattgcagcac tggggccagatggtgaagccctccgtagctgtattatctacacgacggggagtcaggcaactatggatgaacga aatagacagatcgtgagatagggtcctcactgattaaagcattggttaactgtcagaccaagttactcatataact tagattgatttaaaacttcatttttaattaaaaggatctaggtgaagatccttttgataatctcatgacaaaaatccct aacgtgagtttctgtccactgagcgtcagaccccgtagaaaagatcaaaaggatctcttgagatcctttttctgc gcgtaatctgctgcttgcaacaaaaaaaccaccgctaccagcgggtgtttgttggccggatcaagagctacca ctcttttccgaaggtaactggcttcagcagagcgcagataccaataactgtcctctagtgtagccgtagttaggc caccactcaagaactctgtagcaccgcctacatacctcgtctgctaactctgttaccagtggctgctgccagt gcgataagtcgtgttaccgggttgactcaagacgatagtaccggataaggcgcagcggctgggctgaac gggggggttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatactacagcgtgagct atgagaaaagcggcagcttcccgaagggaagggcggacaggtatccggtaagcggcagggtcggaaaca ggagagcgcacgaggagcttcagggggaaacgcctggtatctttatagtcctgtcgggttccgccacctctg actgagcgtcgtttttgtgatgctcgtcagggggcggagcctatggaaaacgccagcaacgcggcctttt acggttcctggccttttctggccttttctcacatgttttctcgtgtatccctgattctgtgataaccgtattac cgctttgagtgagctgataccgtcggcgagccgaacgaccgagcgcagcagtcagtgagcgaggaag cggaagagcgcctgatcggtattttctccttacgcatctgtgcggtatttcacaccgcatatagggtgactctca gtacaatctgctctgatccgcatagttaagccagtatacctcgcgtatcgtactgtgactgggtcatggctgcg ccccgacacccgccaacaccgctgacgcgcctgacgggctgtctgtctccggcatccgcttacagacaag ctgtgaccgtctccggagctgcatgtgcagaggtttaccgtcatcaccgaaacgcgcgaggcagctgcgg taaagctcatcagcgtggtcgtgaagcgattcacagatgtctgctgttcatccgcgtccagctcgttgagtttct cagaagcgtaatgtctggtctgtataaagcgggccatgttaaggcggttttttctgtttgtgactgatgcctc cgtgtaagggggatttctgtcatgggggtaatgataccgatgaacgagagaggatgctcacgatacgggtta ctgatgatgaacatcccgttactggaacgttgtgagggtaaacaactggcggatggtatggcggggacca gagaaaaatcactcagggtcaatgccagcgcttcgttaatacagatgtagggtgtccacagggtagccagcagc atcctgcgatgcagatccggaacataatggtgcaggggcgtgacttccgcgtttccagactttacgaaacacgg aaaccgaagaccattcatgttgtgtcaggtcgcagacgtttgcagcagcagtcgcttcacgttcgctcgcgta tcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctaaccgacaggagcagcagc

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	atgcgacccgtggccaggaccaacgctgcccagatctcgatccgcgaaattaatacgaactactatagg gagaccacaacgggttccctctagaaataatttgttaactttaagaaggagatatacatatggctagcatgactgg tgacagcaaatgggtcgcggatccgaattcgagctccgtcgacaagcttgcggccgactcgagcaccacc accaccaccactgagatccggctgctaacaagcccgaagggaagctgagttggctgctgccaccgctgagc aataactagcataacccttggggcctctaaccgggtcttgagggggttttgcgaaaggaggaactatatccgg attggcgaatgggacgcgccctgtagcggcgcaataagcgcgggcggtgtgtgtgttacgcgcagcgtgacc gtacacttgccagcgccctagcgcccgtccttctgcttcttcccttcttctcggcacgttcgccggcttccc cgtcaagctctaaatcgggggctcccttaggggtccgatttagtgccttacggcacctcgaccccaaaaacttg attagggtgatgttcacgtagtgggcatcgccctgatagacggttttcgcccttgacgttgagtgccacgttc ttaatagtggactctgttccaaactggaacaactcaaccctatctcggtctattctttgattataagggattttg ccgatttcggcctattggttaaaaatgagctgatttaaaaaatttaacgcgaatttaacaaaatattaacgtta caatttcaggtggcacttttcggggaatgtgcgcggaaccctatttgtttatttttaataacattcaaatatgat ccgctcatgagacaataaccctgataaatgctcaataatattgaaaaaggagagatgatttcaacatttccg tgtcgccctatttcccttttgcggcatttgccttctgttttgcacccagaaacgctggtgaaagtaaaagatg ctgaagatcagttgggtgcacgagtggtgtacatgaactggatctcaacagcggtgaagatccttgagagtttgc ccccgaagaacgtttccaatgatgagcacttttaagttctgctatgtggcgcggtattatcccgattgacgccg ggcaagagcaactcggtcgccgcatacactatttca
pET23LucA (SEQ ID NO: 65)	gccagcgccctagcgcccgtccttctgcttcttcccttcttctcggcacgttcgccggcttccccgtcaagct ctaaatcgggggtccctttaggggtccgatttagtgccttacggcacctcgaccccaaaaacttgattagggtg atggttcacgtagtgggcatcgccctgatagacggttttcgcccttgacgttgagtgccacgttcttaatagt gactcttgttccaaactggaacaactcaaccctatctcggtctattctttgattataagggtttgccgatttcg gcctattggttaaaaaatgagctgatttaaaaaatttaacgcgaatttaacaaaatattaacgtttacaatttcag gtggcacttttcggggaatgtgcgcggaaccctatttgtttatttttaataacattcaaatatgtatccgctcatg agacaataaccctgataaatgctcaataatattgaaaaaggagagatgatttcaacatttccgtgtcgccctt attccctttttgcggcatttgccttctgttttgcacccagaaacgctggtgaaagtaaaagatgctgaagatc agttgggtgcacgagtggtgtacatgaactggatctcaacagcggtgaagatccttgagagtttgcggcgaag aacgtttccaatgatgagcacttttaagttctgctatgtggcgcggtattatccgtattgacggcggaagag caactcggtcgccgcatacactattctcagaatgacttggtgagtgactaccagtcacagaaaagcatcttacgg atggcatgacagtaagagaattatgcagtgtgccataacctgagtgataacactgcggccaacttactctgac aacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgttg ggaaccggagctgaatgaagccataccaacgacgagcgtgacaccacgatccgtcagcaatggcaacaa cgttgcgcaaaactattaaactggcgaaactacttctgacttcccggcaacaattaatagactggatggaggcgg ataaagttgcaggaccacttctgcgctcgcccttccggctggctggttattgctgataaatctggagccggtga gcgtgggtctcgcggtatcattgcagcactggggccagatggtgaagccctcccgtatcgtatctacacgac ggggagtcaggcaactatggatgaacgaatagacagatcgctgagataggtgcctcactgattaagcattgggt aactgtcagaccaagttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaaga tccttttgataatctcatgacaaaatcccttaacgtgagtttcttccactgagcgtcagaccccgtagaaaaga tcaaaggatcttctgagatcccttttctgcgctaactctgctgcttgcacacaaaaaaaccaccgctaccagcg gtggtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgagataccaa atactgtcctctagtgtagcgtgatttagccaccactcaagaactctgtagcaccgctacatacctcgtctg ctaactcgttaccagtggtgctgctgccagtgggcgataagctgtgtcttaccgggttgactcaagacgatgttac cggataaggcgacgcggtcgggctgaacggggggttcgtgcacacagcccagcttgagcgaacgacctac accgaactgagatacctacagcgtgagctatgagaagcgccacgctcccgaaggagaaaggcggacag gtatccggtgaagcggcagggtcggaaacagagagcgcacgagggaagcctccagggggaaacgctgtgatac ttatagtcctgtcgggttccgacactctgactgagcgtcgtttttgtgatgtcgtcagggggggcggagcctat ggaaaaacgccagcaacgcggccttttacggttctggtccttttgcgtccttttgcacatgttcttctcgtt atccccgtattctgtgataaccgtattaccgctttgagtgagctgataccgtcggcgagccgaacgaccga gcgcagcagtgatgagcaggaagcggagagcgcctgatgcggtattttctccttacgcactgtgctgta tttcaaccgcataatggtgactctcagtaaatctgctctgatccgcatagttaagccagtatacactccgcta tcgctacgtgactgggtcatggtgcgccccgacacccgcaacacccgctgacgcgccctgacgggctgtc tgctcccgcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttcaccgtcatc

[illegible]

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
(SEQ ID NO: 72)	
HAP270-r (SEQ ID NO: 73)	tctagagggaaaccgtgtgtgt
YAP1-f (SEQ ID NO: 74)	gggtggttctagatagtaaccagccctagctgtt
YAP1-r (SEQ ID NO: 75)	gggtggtcatatgggtttaagaacaacttttccttc
pET23TFIIDLucA (SEQ ID NO: 76)	ttccgatttagtgctttacggcacctcgacccccaaaaaacttgattaggggtgatgggtcacgtagtgggccatcgc cctgatagacggtttttcgccctttgacgttgagtcacgttcttaatagtgactctgttccaaactggaacaac actcaaccctatctcggctctattctttgattataagggattttgccgatttcggcctatttggttaaaaaatgagctgat ttaacaaaaatttaacgcgaattttaacaaaatatttaacgtttacaatttcaggtggcacttttcggggaaatgtgcgc ggaacccctatttgtttatttttctaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttca ataatattgaaaaaggaagagtagtattcaacatttccgtgtcgccttattcccttttgcggcattttgccttc ctgttttgcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacat cgaactggatctcaacagcggtaagatccttgagagtttcgccccgaagaacgtttccaatgatgagcactttta aagttctgctatgtggcgcggtattatccgattgacggcgggaagagcaactggtcgccgcatacactatc tcagaatgacttggtgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatg cagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggag ctaaccgctttttgcacaacatgggggatcatgtaactcgcttgatcgttgggaaccggagctgaatgaagcca taccaaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttcgcaaacattataactggcga actacttactctagcttccggcaacaattaatagactggatggaggcggataaagtgcaggaccacttctgcg ctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggatcattgca gactggggccagatggaagccctcccgatcgtatgtatctacacgacggggagtcaggcaactatggatga acgaaatagacagatcgtgagataggtgcctcactgattaagcattggttaactgacagcaagtttactcatat atactttagattgatttaaaacttcatttttaattaaaaggatctaggtgaagatccttttgataatctcatgacaaaa tcccttaacgtgagtttctgctcactgagcgtcagaccccgtagaaaagatcaaaaggatctcttgagatccttttt tctgcgcgaatctgctgcttgcacaacaaaaaaaccaccgctaccagcgggtggtttgttgcggatcaagagct accaactcttttccgaaggtaactggcttcagcagagcgcagataccaataactgtccttctagtgtagccgtagt taggccaccactcaagaactctgtagcaccgcctacatacctcgtctgtctaatcctgttaccagtggtgctgc cagtggtgcgataagtcgtgtcttaccgggtggactcaagacgatagttaccggataaggcgcagcggtcgggct gaacggggggtctgtgcacacagccagcttggagcgaacgacctacaccgaactgagatacctacagctg agctatgagaaagcggcacgcttccggaagggaagggcggacaggtatccggtaagcggcagggtcggga acaggagagcgcacgagggagcttcagggggaacgcctggtatctttatagctctgtcgggttcgccacct ctgacttgagcgtcgtttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcct tttacgggtcctggccttttctgctgacctttgtcacatgttcttctgcgttatccctgattctgttgataaccgat taccgctttgagtgagctgataaccgtcgcgcagccgaacgaccgagcgcagcagtcagtgagcagga agcgggaagagcgcctgatgcggtattttctcttacgcacatctgtgcggtatttcacaccgcatataggtgactct cagtacaatctgctctgatgccgatagttaagccagtatcactccgctatcgtacgtgactgggtcatggctg cgccccgaccccgcaacacccgctgacgcgcctgacgggctgtctgctccggcatccgcttacagaca agctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgtcatcaccgaaacgcgcgaggcagctgc ggtaaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagttt tccagaagcgttaatgtctggcttctgataaagcgggccatgtaaggcggttttctctgttggtcactgatgcc tccgtgtaaggggatttctgttcatgggggtaatgataccgatgaaacgagagaggtgctcacgatacgggt actgatgatgaacatgcccgttactggaacgttgtaggggtaaacaactggcggatggatgcggcgggacc agagaaaaatcactcagggtcaatgccagcgttcgttaatacagatgtaggtgtccacagggtagccagcag catcctgcgatgcagatccggaacataatgggtgcagggcgtgacttccgcgtttccagactttacgaaacacg gaaaccgaagaccattcatgtgtgtcaggtgcagacgttttgacgagcagtcgcttcacgttcgctcgcgt atcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgat catgcgcacccgtggccaggaccaacgctgcccagatctcgtatccgcgaaattatacgaactcactatag ggagaccacaacggttccctctagatcgtatgcggcgcgaattcgggacgtgaaattacagtagttactgttt tttggactataagatcgggggaaagataacacataagaaataaaacgactactagttagactgctctgcggaag aagcaagggaagtaaggctgcattttattttcttctagttcaacataaacaggtgtatcaagagaactttttaa

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	gagctcgtcgacggatccatattggaagacgccccaaacataaagaaaggcccgccattctatccgctaga ggatggaaccgctggagagcaactgcataaggctatgaagagatacgccctggctcctggaacaattgcttttac agatgcacatatcgaggtgaacatcacgtacgcggaatacttcgaaatgtccgttcggttggcagaagctatgaa acgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactctctcaattctttatgccgggtgttg gcgcgttattatcggagttgcagttgcgcccgcgaacgacattataatgaacgtgaattgctcaacagtatgaa catttcgcagcctaccgtagtgtttgttccaaaaagggttgcaaaaattttgaacgtgcaaaaaaattaccaa taatccagaaaattattatcatgattctaaaacggattaccagggttcagtcgatgtacacgttcgtcacatctc atctaccccccgttttaataacgattttgtaccagagtcctttgatcgtgacaaaacaattgcactgataatgaa ctcctctggatctactgggttacctaagggtgtggccctccgcatagaactgcctgcgtcagattctcgcatgcc agagatcctattttggcaatcaaatcattccggatactgcgattttaagtgtgttcattccatcacggttttggaat gtttactacactcggatatttgatgtggatttcgagtcgtcttaatgtatagattgaagaagagctgttttacgac ccttcaggattacaaaattcaagtgcgttgctagtaccaaccctattttcattcttcgcaaaaagcactctgattgac aaatacagttatctaatttacagaaattgcttctgggggcgcacccctttcgaagaagtcgggggaagcgggtg caaaacgcttccatcttccagggtacgacaaggatattgggtcactgagactacatcagctattctgattacacc cgaggggggatgataaaccgggcgcggtcggtaaaagtgttccattttgaaagcgaaggtgtgtgatctggatac cgggaaaacgctgggcgttaacagagagggcgaattatgtgtcagaggacctatgattatgtccggttatgtaa caatccggaagcgaccaacgccttgattgacaaggatggatggctacattctggagacatagcttactgggacg aagacgaacacttctcatagttgaccgctgaagtctttaataatacaaaaggataccagggtggccccgcgtga attggagtcgatattgtacaacaccccaacatcttcgacgcggcggtggcaggtcttccgacgatgacgccg gtgaactcccgccgctgtgtgttttgagcacggaagacgatgacggaaaaagagatcgtggattacgtc gccagtcaagtaacaaccgcaaaaagtgcgcggaggagtgtgttttggtgacgaagtaccgaaaggctttac cggaaaactcgacgcaagaaaaatcagagagatcctataaaggccaagaaggcggaaagtccaaattgta aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaactcgagcaccaccaccaccactgagatccggctgctaacaaggccgaa aggaagctgagttgggtgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttg aggggtttttgctgaaaggaggaactatatccggattggcgaatgggacgcgccctgtagcggcgcaataagc gcggcggtgtgtgtgttacgcgcagcgtgaccgctacacttgcagcgccctagcggcgctccttcgctt cttcccttcttctcgcacgttcgccggttccccgtcaagctctaaatcgggggtccctttagggtg
pET23HAP270LucA (SEQ ID NO: 77)	tggcgaatgggacgcgccctgtagcggcgcaatgaagcgcggcggtgtgtgtgttacgctgcagcgtgaccgc tacacttgccagcgccttagcggccgctcctttcgtttcttcccttctctcgcacgttcgccggcttccccgt caagctctaaatcgggggtccctttagggttcgatttagtgccttacggcacctcgacccccaaaaaactgatta gggtgatgggtcacgtagtgggcatcgcctgtatagacggttttccgctttgacgttgagtcacgttcttta atagtggaactctgttccaaactgaacaacactcaaccctatctcgtctattcttttgattataagggttttgccg atttcggcctattgttaaaaaatgagctgatttaaaaaaattaacgcgaattttaaaaaatattaacgtttacaat ttcaggtggcacttttcggggaatgtgcgcggaaccctattgttttttaataacattcaaatatgtatccgc tcatgagacaataaccctgataaatgctcaataatattgaaaaaggaagatgatgattcaacatttccgtgc gccctattccctttttgcggcattttgccttctgttttgcaccagaaacgctggtgaaagttaaagatgctg aagatcagttgggtgcacgagtggttacatgaactggatctcaacagcggttaagatccttgagagtttccg ccgaagaacgtttccaatgatgagcacttttaagttctgctatgtggcgcggtattatccggtattgacgggg caagagcaactcggcgcgcacatactattctcagaatgacttggtgagtagtaccagtcacagaaaagcat cttacggatggcatgacagtaagagaattatgcagtctgccataaccatgagtataacactgcggccaactta cttctgacaacgatcggaggaccgaaggagtaaccgctttttgcacaacatgggggatcatgtaactgcctt gatcgttgggaaccggagctgaatgaagccataccaacgacgagcgtgacaccacgatccctgcagcaatg gaacaacgttgcgcaactattaactggcgaactacttacttagcttccggcaacaattaatagactggatgg aggcgataaagttgcaggaccactctgcgctcggccctccggctggctgtgtttattgctgataaatctggagc cgggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctccgctatcgtatgtatcta cacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgattaag cattgtaactgtcagaccaagtttactcatatatacttttagattgatttaaaacttatttttaattaaaggatctagg tgaagatccttttgataatctcatgacaaaatcccttaacgtgagtttcttccactgagcgtcagacccgtag aaaagatcaaaggatcttctgagatccttttttgcgcgtaactctgctgcttgcacacaaaaaaccaccgctac cagcgggtggttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	accaaatactgtccttctagttagccgtagtaggccaccacttcaagaactctgtagcaccgcctacatacctc gctctgctaatacctgttaccagtggtgctgctgccagtgggcagataagtcgtgttaccgggttgactcaagacgat agttaccggataaggcgcagcggctgggctgaacggggggttcgtgcacacagcccagcttgagcgaaacg acctacaccgaactgagatacctacagcgtgagctatgagaaagcggcagctcccgaaggaggagaaaggcg gacaggtatccggtaagcggcaggggtcggaaacaggagagcgcacgagggagcttcagggggaaacgcct ggatcttctatagtcctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggcgga gcctatggaaaaacgccagcaacgcggcctttttacgggtcctggccttttctggtccttttctcacatgttcttc tgcgttatccctgattctgtggataaccgtattaccgcctttgagttagctgataccgctcggcgagccgaacg accgagcgcagcagtcagtgagcaggaagcggagagcgcctgatgcgggtttttctcttacgcattctgt gcgggtattcacaccgcatatatgtgacactctcagtacaatctgctctgatccgcatagttaagccagtatac tccgctatcgtacgtgactgggtcatggctgcgccccgacaccgccaacaccgctgacgcgcctgacgg gcttctgtcctccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttcac cgtcatcaccgaaacgcgcgaggcagctgcggtaagctcatcagcgtggctgtgaagcgattcacagatgtc tgcctgttcacccgcgtccagctcgttgagtttccagaagcgttaagtgtctggtctgataaagcgggcatgtt aaggggcggtttttctgtttgtcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgataccgat gaaacgagagaggatgctcacgatacgggttactgatgatgaacatgccgggttactggaacgtgtgagggtga aacaactggcggtatggatgcggcgaggaccagagaaaaatcactcagggtcaatgccagcgcttcgttaatac agatgtaggtgtccacagggtagccagcagcatcctgcgatgcagatccggaacataatgtgtcagggcgct gactccgcgtttccagactttacgaaacacggaaaccgaagaccattcatgtttgtcaggtcgcagacgtttt gcagcagcagtcgttcacgttcgtcgcgtatcgggtgattcattctgtaaccagtaaggcaaccccgccagcc tagccgggtcctcaacgacaggagcagcatatgcgcacccgtggcaggacccaacgctgccgagatctc gatcccgcaaatatacgaactcactataggagaccacaacggttccctctagataaaccagtttatatc gtatatgctatctacaggtccactttacacttaataataaaaaatactataaagggaaccagaaaaataaaaaa gggtcattattttttagcagatcattatcaaacgcatagggaagagaaaaaacacagttttttttccacacata ttattggctcctagtagcatcaagagcattttaatgggtgtgctgatttgtttacacattttctagtaaaaaaaa aacaataaaagacatatggaagacgccaaaaacataaagaaaggcccggtgcattctatccgctagaggatg gaaccgctggagagcaactgcataaggctatgaagagatacgccttggtcctggaaacattgttttacagatg cacatacaggggtgaacatcacgtacgcgggaatactcgaatgtccgttcggttggcagaagctatgaaacgat atgggctgaatacaaatcacagaatcgtcgtatgcagtgaactctctcaattcttatgccggtgttggcgcg ttattatcggagttgcagttgcggccgcgaacgacattataatgaacgtgaattgctcaacagatgaacatttcg cagcctaccgtagtgtttttccaaaaagggttgcaaaaaatttgaacgtcaaaaaaattaccaataatcca gaaaattattatcatggattctaaaacggattaccagggtttcagtcgatgtacacgttcgtcacatctcatctacct cccggtttaatgaatacattttgtaccagagtcctttgatcgtgacaaaacattgcactgataatgaactcctctg gatctactgggttacctaagggtgtggccctccgcatagaactgcctgcgtcagattctcgtatgccagagatcc tatttttggcaatcaaatcattccggatactgcgattttaagtgtgttccattccatcacgggtttggaatgtttactaca ctcggatattgatattgtgatttcgagtcgtcttaattgtatagatttgaagaagagctgtttttacgatcccttcagga ttacaaaattcaagtgcgtttagtagtaccacccattttcattcttcgcaaaagcactctgattgacaaatagat ttatctaatttacacgaaattgcttctggggcgacacctttcgaagaagtcggggaagcggttgcaaacgct tccatcttcagggtacgacaaggatattgggtcactgagactacatcagctattctgattacaccgagggggg atgataaacggggcggtcggtaaaagtgttcatttttgaagcgaaggttggatctggataccgggaaaa cgctgggcgttaatcagagagcgaattatgttcagagagacattatgctcgggttatgtaaacatccgga agcgaccaacgccttgattgacaaggatggatggctacattctggagacatagcttactgggacgaagacgaac acttctcatagttgaccgcttgaagctttaaataatacaaaaggataccaggtggcccccgctgaattggagtcg atattgttacaacaccccaacatcttcgacgcggcggtggcaggtctccgacgatgacgggtgaacttcc gccgccgttgttttggagcacggaaagacgatgacggaaaaagagatcgtggattacgtcgccagtcgaagt aacaaccgcgaaaaagtgcgcggaggagttgtgttgggacgaagtaccgaaaggtcttaccggaaaaactc gacgcaaaaaatcagagagatctcataaaggccaagaaggcggaaggtccaaattgtaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaactcagcaccaccaccaccactgagatccggctgctaacaagccgaaagggaagctgagttg gctgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttgaggggttttctaaa aggagaaactaacctat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
pET23YAP1LucA (SEQ ID NO: 78)	ttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggatggttcacgtagtgggccatcgc cctgatagacgggttttcgccctttgacgttgagtcacgttcttaatagtggactctgttccaaactggaacaac actcaaccctatctcgggtctattctttgatttataagggttttgcgatttcggcctattgggttaaaaaatgagctgat ttaacaaaaatttaacgcgaattttaacaaaatattacgtttacaatttcaggtggcacttttcggggaaatgtgcgc ggaacccctatttgtttattttctaaatacattcaaatatgtatccgctcatgagacaataaacctgataaatgtctca ataattattgaaaaagggaagagatgaggtattcaacattccgtgtcggccttattccctttttgcggcattttgccttc ctgttttctcaccacagaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttcatc cgaactggatctcaacagcggtaagatccttgagagttttcggccgaagaacgtttccaatgatgagcactttta aagttctgtatgtggcgcggtattatccgtattgacggcgggaagagcaactcggtcggcgatacactattc tcagaatgacttgggtgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatg cagtgtcgcataaccatgagtgataacactcggccaacttacttctgacaacgatcggaggaccgaaggag ctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagcca taccaaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttcgcaaaactattactggcga actacttactctagcttccggcaacaattaatagactggatggaggcggataaagtgcaggaccacttctgcg ctggcccttccggctggctgtttattgctgataaatctggagccgggtgagcgtgggtctcgcggatcattgca gcaactggggccagatgtaagccctcccgatcgtatgtatctacacgacggggagtcaggcaactatggatga acgaaatagacagatcgtgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatat atactttagattgatttaaaacttcatttttaattaaaaggatctaggtgaagatccttttgataatctcatgacaaaa tcccttaacgtgagtttctgctcactgagcgtcagaccccgtagaaaagatcaaaaggatcttctgagatccttttt tctgcgcgtaatctgctgttgcacaacaaaaaacaccgctaccagcgggtgttggtttgcggatcaagagct accaactcttttccgaaggtaactggcttcagcagagcgcagataccaatactgtccttctagttagccgtagt tagggccaccactcaagaactctgtagcaccgctacatacctcgtctgtctaatcctgttaccagtggctgctgc cagtggcgataagtcgtgttaccgggtggactcaagacgatagttaccggataaggcgcagcggtcgggct gaacgggggggtcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatactacagcgtg agctatgagaaagcgccacgcttcccgaagggaagaaaggcggacaggtatccggtaagcggcagggtcggga acaggagagcgcacgagggagcttccagggggaacgcctggtatctttatagctctgcgggttcgccacct ctgacttgagcgtcgattttgtgatgctcgtcagggggggcggagcctatggaaaaacgccgaacgcggcct ttttacggctcgtggccttttctggtccttttctcacatgttcttctcgtttatccctgattctgttgataaccgtat taccgctttgagtgagctgataccgctcgcgcagccgaacgaccgagcgcagcagtcagtgagcagga agcgggaagagcgcctgatcgggtattttctccttacgcacatctgtgcgggtatttcacaccgatatatggtgactct cagtacaatctgctctgatccgcatagttaagccagtatacactccgctatcgtactgactgggtcatggctg cgccccgacaccgccaacaccgctgacgcgcctgacgggctgtctgctccggcatccgcttacagaca agctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgtcatcaccgaaacgcgcgaggcagctgc ggtaaaagctcatcagcgtgtgtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagtttc tccagaagcgtaattgtctggttctgataaagcgggcatgttaaggcgggttttctcgttggctactgatgcc tccgtgtaagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgtcacgatacgggtt actgatgatgaacatgcccgggttactggaacgttgtgagggtaaacaactggcggatggatggcgggggacc agagaaaaatcactcagggtcaatgccagcgccttctgtaatacagatgtagggttccacagggttagccagcag catcctgcgatgcagatccggaacataatgggtcagggcgtgacttccgcgttccagactttacgaaacacg gaaaccgaagaccattcatgttgtgtcaggtcgcagacgttttgcagcagcagtcgttcaagttcgtcgcgt atcgggtgattcattctgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgat catgcgcacccgtggccaggacccaacgctgcccagatctcgtatcccgcaaatatacgaactcactatag ggagaccacaacgggttccctctagatagtaaccagccctagctgtttgggtgatttgacctagggttactctttttt tctgggtgcgggtaacaatttgggccccgcaaagcgccgtcttgcgtatgggaaccgaaacccctccgatgaag agtaggaggggtggcaactgatgatgcgtaaggcttaagagatacattgtcttaatagtcttccgttaccgatta agcacagtacctttacgttatatataaggattggtgttagcttttttctgagccctggttgactgtgcatgaacac gagccatttttagttgttaagggaagtttttgcaccccaaaacgtttaaagaaggaaaagttgttcttaaacccat atggaaagacgcaaaaacataaaagaaaggccggcgccattctatccgctagaggatggaaccgctggagag caactgcataaggctatgaagagatacgcctggttctggaacaattgctttacagatgcacatagcaggtga acatcacgtacgcggaataacttcgaaatgtccgttcgggtggcagaagctatgaacgatatgggtgtaatacaa atcacagaatcgtcgtatgcagtgaaaactctcttcaattcttatgccggtgttggcgcggtattttatcgaggttc agttgcggccgcgaacgacatttataatgaacgtgaattgctcaacagtatgaacatttcgcagcctaccgtagtg

[illegible]

[illegible]

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	accgaactgagatacctacagcgtgagctatgagaagcgcacgctcccgaaggagaaaggcggacag gtatccggtaagcggcagggcgcgaacaggagagcgcacgagggagcttcagggggaaacgcctggatc ttatagtcctgtcgggttcgccacctctgacttgagcgtcgtatgttgatgctcgtcagggggcgagcctat ggaaaaacgccagcaacgcggccttttacggctcctggccttttgctggccttttgctcacatgttcttctcgtt atccccgtattctgtggataaacgtattaccgctttgagtgagctgataccgctcggcgcagccgaacgaccga gcgcagcagtgagtgagcgcgaggaagcgaagagcgcctgatgcgggtatgttctccttacgcatctgtgcggta ttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagtaagccagtatacactccgcta tcgctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgcctgacgggctgtc tgctcccgcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgcagagggtttaccgctc accgaaacgcgcgagcagctgcggtaaagctcatcagcgtggctgtgaagcgattcacagatgtctgcctgtt catccgcgtccagctcgttgagtttccagaagcgttaagtctggtctgataaagcgggcatgttaaggc ggtttttctgttggctactgatgcctccgtgtaagggggatttctgtcatgggggtaataaccgatgaacg agagaggatgctcacgatacgggttactgatgatgaacatgcccggtactggaacgttgagggttaacaac tggcggtatggatgcggcgggaccagagaaaaatcactcaggggtcaatgccagcgttcgttaatacagatgta ggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatgtgacgggcgctgactcc gcgttccagactttacgaaacacggaaccgaagaccattcatgttgtgtcaggtcgcagacgtttgcagca gcagtcgcttcacgttcgctcgcgtatcggtgattcattctgtaaccagtaaggcaaccccgccagcctagccg ggtcctcaacgacaggagcacgatcatgcgcaccggtggccaggaccaacgctgcccagatctcgtatccc gcgaaattaatacgaactactatagggagaccacaacggttcccttagagcaaaaatgtgattctgttgaat acaattttgagagggttaataaattacagtagtgctattttgtatttaggttagctatttagctttacgttcaggatgc ctagtggcagccccacaatatccaggaaagccctctctgcgggtttcagattaggtagtcgaaaaacctaagaaat ttacctgtacatttcaagattcatatggaagacgcaaaaacataaagaaaggccggccattctatccgcta gaggatggaaccgctggagagcaactgcataaggctatgaagagatacgccttggttcctggaacaattgctttt acagatgcacatatcagggtgaacatcacgtacgcggaatactcgaatgtccgttcggttggcagaagctatg aaacgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactctctcaattctttatgccggtgtg ggcgcggtatttatcgaggtgcagttgcgccgcgaacgacattataatgaacgtgaattgtcaacagtatga acatttcgagcctaccgtagtgtttgttccaaaaagggttgcaaaaaatttgaacgtgcaaaaaaattacca ataatccagaaaattattatcatggattctaaaacggattaccagggttccagtcgatgtacacgtcgtcacatct catctacctcccggttttaataacgattttgtaccagagtcctttgatcgtgacaaaaaattgcactgataatga actcctctggatctactgggttacctaagggtgtggccttcgcatagaactgcctgcgtcagattctcgcgtgc cagagatcctattttggcaatcaaatcattccggatactgcgattttaagtgtgttccattccatcacggtttggaa tgtttactacactcggatatttgatgtggttgcagtcgtctaatgtatagatttgaagaagagctgttttacgat cccttcaggattacaaaattcaaaagtgcgttgtagtaccacccctattttcattcttcgcaaaaagcactctgattga caaatcagatttatctaatttacacgaaattgcttctgggggcgcacctcttcgaaagaagtcggggaagcgggtt gcaaaacgcttccatcttcaggagatacgaaggtatgggctcactgagactacatcagctattctgattacac ccgagggggatgataaaccggcgcggtcggtaagttgttccattttgaagcgaagggttgatctggata ccgggaaaacgctggcggttaatcagagaggcgaattatgtgtcagaggacctatgattatgtccggttatgtaa acaatccggaagcgaccaacgccttgattgacaaggatggatggctacattctggagacatagcttactgggac gaagacgaacacttctcatagttgaccgctgaagtccttaattaatacaaaaggataccagggtggccccgcgtg aattggagtcgatattgtacaacaccccaacatcttcgacgcggcggtggcaggcttcccacgatgacgcc ggtgaactcccgccgccgttgttttggagcacggaaagacgatgacggaaaaagagatcgtggattacgt cgccagtcagtaacaaccgccaaaaagtgcgcggaggaggtgtgtttgtggacgaagtaccgaaaggctcta ccggaaaaactgcagcgaagaaaaatcagagagatcctcataaaggccaaagaaggcggaagtcgaaattgt aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaactcgagcaccaccaccaccactgagatccggctgtaacaaagcccgga aaggaagctgagttggctgctgccaccgctgagcaataactagcataacccttggggccttaaacgggtctt gaggggtttttgctgaaaggaggaaactatccggattggcgaatgggacgcgccctgtagcggcgcatgaag cgcggcggtgtgtgtgttacgcgcagcgtgaccgctacatt
Primer IGR-f (SEQ ID NO: 95)	gggtggttctagagcaaaaatgtgatcttgccttga
Primer IGR-r	gggtggtcatatgaatcttgaaatgtagcaggtaaat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
(SEQ ID NO: 96)	
3'-UTR of yeast <i>FBA1</i> gene (Accession number: NM_001179626) (short fragment) (SEQ ID NO: 97)	gagtattgaatctgtttagaataatggaatattttttattttattttatattattggcggccttttcttgaaggtc aatgacaaaatgatatgaaggaaataatgatttctaaaattttacaacgtaagatattttacaaaagcctagctcatc ttttgcatgcactattttactcacgcttgaaattaacggccagtcactgcggagtcatttcaaaagtcacctaactg atctatcggtttgatagctcattttggagttcgcgattgtctctgtattcacaactgttttaattttattcattctggaa ctcttcgagttctttgtaaagctttcatagtagcttactttatcctccaacatatttaactcatgtcaatttcggctctta aattttccacatcatcaagtcaacatcatcttttaacttgaattattctctagctctccaaccaagcctcattgtcctt tgatttactggtgaaaagtatacactttgcgcgaatccagggtcaaaacttctcgaagaattcaccatttctc gacatcatagtacaattgtttgtctcccatcacatattaatactgatggattcttatgaagcgtgggtaattgg acgtgtcactc
3'-UTR of yeast <i>FBA1</i> gene (Accession number: NM_001179626) (long fragment) (SEQ ID NO: 98)	gagtattgaatctgtttagaataatggaatattttttattttattttatattattggcggccttttcttgaaggtc aatgacaaaatgatatgaaggaaataatgatttctaaaattttacaacgtaagatattttacaaaagcctagctcatc ttttgcatgcactattttactcacgcttgaaattaacggccagtcactgcggagtcatttcaaaagtcacctaactg atctatcggtttgatagctcattttggagttcgcgattgtctctgtattcacaactgttttaattttattcattctggaa ctcttcgagttctttgtaaagctttcatagtagcttactttatcctccaacatatttaactcatgtcaatttcggctctta aattttccacatcatcaagtcaacatcatcttttaacttgaattattctctagctctccaaccaagcctcattgtcctt tgatttactggtgaaaagtatacactttgcgcgaatccagggtcaaaacttctcgaagaattcaccatttctc gacatcatagtacaattgtttgtctcccatcacatattaatactgatggattcttatgaagcgtgggtaattgg acgtgtcactctacttcgctttttccctactccttttagtacggaagacaatgctaataaataagagggtataataa tattattaatcgcaaaaaagattaaacgccaagcggttaattatcagaagcaaacgctgaccaatccttgaatg cttcccaattgtatattaagagtcacacagcaacatattcttgttattaataattattattgattttgatattgtataaa aaaaccaaatatgtataaaaaagtgataaaaaataccaagtatggagaatatattagaagctatagcgttaaaa accagaacgtgcacaatttttaacttgccaaatgaaaaaacggaaatatacggaaaagaagtggaagtaata gttagaaaggcaaaaaaggaaagaacaatttaaaatatcttaagattatattagaacaacaccaatgttcattt cattccttagaatatatacgaatgaaatgaccaactactgttttgtaaactgaggaagaaagaatattttctccg aaaacttgatatacctagcttgcttctgtttattgtctttgacctattttttcaaaaatcaccgtgcttttgtgagttt ttagatgttgataaattgtcacttctactgaattttttctcacagaacatacaagcaaaaggcggtccgttgatg aacggatcttatgcttttcaagtgtcactgcgtctgaatgccttctcacagctttacacttgaaggtttattttt atcgtagttgttgggtcaatg
FBA3UTR-f (SEQ ID NO: 99)	gggtggtgagctcgagtattgaatctgtttagaataatgg
FBA3UTR1-r (SEQ ID NO: 100)	gggtggtctcgagtgcacgtccattaccagc
FBA3UTR2-r (SEQ ID NO: 101)	gggtggtctcgagcattgacccaacaactacg
TMV1, ranges from 4920 to 5711 of genome (792 nt in length) (SEQ ID NO: 102)	aggaaaagtgaatatcaatgagttatcgacctgacaaaaatggagaagatcttaccgtcgatgtttaccctgta aagagtgttatgtgtccaaagttgataaaataatggttcattgagaatgagtcattgtcagaggtgaaccttctaaa ggaggttaagcttattgatagtgatagctgtgttagccggtttggtcgtcacgggcgaatggaacttgctgacaa ttgcagaggaggtgtgagcgtgtgtctgttgacaaaaggatggaagagccgacgagggcactctcggatct tactacacagcagctgcaaaagaaagatttcagttcaaggctgttcccaattatgtcataaccaccaggacgcg atgaaaaacgtctggcaagttttagtttaattatagaatgtaaaatgtcagcgggtttctgtcgttctcttgaggt ttgtgtcgggtgtgtattgtttatagaataatataaaattaggtttgagagagaagattacaaacgtgagagacgga gggcccattggaacttacagaagaagtcgttgatgagttcatggaagatgtccctatgtcaatcaggcttgcaaaag tttcgatctcggaccggaaaaaagagtgatgtccgtaaaagggaataatagtagtagtaccgggtcagtgccgaa caagaactatagaatgttaaggattttggaggaatgagtttaaaaaagaataatttaacgatgatgattcggagg ctactgtcggcgaatcggttcgttttaaat
TMV2 ranges from 6192 to 6395 genome (204 nt in	ggtagtcaagatgcataataaataacggattgtgtccgtaatcacacgtggtgcgtacgataacgcatagtgtttt ccctccacttagatcgaagggtgtgtcttgatcgcgcgggtcaaatgtatatggttcataatccgcaggca cgtataaaagcgaggggttcgaatccccgttaccctggtaggggcccc

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
length) (SEQ ID NO: 103)	
TMV13U200 (SEQ ID NO: 104)	ggtggtctcgagccaaaccggctaaacaga
TMV13U400 (SEQ ID NO: 105)	ggtggtctcgagaacttgccagacgttttcat
TMV13U700 (SEQ ID NO: 106)	ggtggtctcgagatttctatagtcttgttcggca
TMV13U-f (SEQ ID NO: 107)	ggtggtgagctcaggaaaagtgaatatcaatgagttatc
TMV13U200-r (SEQ ID NO: 108)	ggtggtctcgagccaaaccggctaaacaga
TMV13U400-r (SEQ ID NO: 109)	ggtggtctcgagaacttgccagacgttttcat
TMV13U700-r (SEQ ID NO: 110)	ggtggtctcgagatttctatagtcttgttcggca
TMV2 with the length of 204 nt (SEQ ID NO: 111)	ggtagtcaagatgcataataaataacggattgtgtccgtaatcacacgtggtgcgtacgataacgcatagtgtttt ccctccacttagatcgaagggtgtgtcttggatcgcgcgggtcaaatgtatatggtcatatacatccgcaggca cgtataaaagcgaggggttcgaatccccccgtacccccggtaggggcccc
TMV23U-f (SEQ ID NO: 112)	ggtggtgagctcggtagtcaagatgcataata
TMV23U-r (SEQ ID NO: 113)	ggtggtctcgagtgggcccctaccggggg
QEluc-f (SEQ ID NO: 114)	ttactatttacaattacatatggaagacgccaaaaac
QEluc-r (SEQ ID NO: 115)	agcagccggatctcagtttacaatttggactttccgc
QEGFP-f (SEQ ID NO: 116)	ttactatttacaattacatatgagcaaagggtgaagaac
QEGFP-r (SEQ ID NO: 117)	agcagccggatctcagtttatttttgaactgcgga
QECAT-f (SEQ ID NO: 118)	ttactatttacaattacatatggaaaaaaaaatcaccgg
QECAT-r (SEQ ID NO: 119)	agcagccggatctcagtttacgcaccaccctgcc
T7 promoter (SEQ ID NO: 120)	taatacgactcactataggag
Ω sequence (SEQ ID NO: 121)	tattttacaacaattaccaacaacaacaacaacaacattacaattactatttacaatta
poly(A)50 tail (SEQ ID NO: 122)	aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
QET7Ome-f (SEQ ID NO: 123)	acgtgccccgagatctcgatccgcgaaattaatacgactcactataggagatttttacaacaattaccaaca caacaacaacaacaacattacaattactatttacaattacat
pET23c-GFP-cyc3 (SEQ ID NO: 124)	tggcgaatgggacgcgccctgtagcggcgcatgaagcgcggcggtgtggtggttacgcgcagcgtgaccgc tacacttgccagcgcctagcggccgctcttctgctttcttcccttcttctcgccacgttcgccggcttccccgt caagctctaatacgggggtcccttaggggtccgatttagtgcttacggcacctcgacccccaaaaacttgatta gggtgatggttcacgtagtgggccatcgccctgatagacgggttttcgcccttgacgttgagtcacgttctta atagtggaactctgttccaaactggaacaacactcaaccctatctcggctattcttttgattataagggaatttgcg atttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgttacaat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	ttcaggtggcacttttcggggaaatgtgcgcggaaccctatttgttttttctaatacattcaaatatgtatccgc tcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttcctgtgc gcccttattccctttttcgggcattttgcttcctgttttgcaccagaaacgctggtgaaagtaaaagatgctg aagatcagttgggtgcacgagtggttacatcgaactggatctcaacagcggtaagatccttgagagtttccgc ccgaagaacgtttccaatgatgagcacttttaaagtctgctatgtggcgcggtattatccgtattgacgccggg caagagcaactcggcgcgcgcatacactattctcagaatgacttgggtgagtactaccagtcacagaaaagcat ctacggatggcatgacagtaagagaattatgcagtctgcataaccatgagtataacactgcggccaactta ctctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaactgggggatcatgtaactgcctt gatcgttgggaaccggagctgaatgaagccatacacaacgacgagcgtgacaccagatgctgcagcaatg gcaacaacgttgcgcaactattaactggcgaaactactacttagcttcccggaacaattaatagactggatgg aggcggataaagttgcaggaccactctgcgctcggccttccggctggtgttattgctgataaatctggagc cgggtgagcgtgggtctcgggtatcattgcagcactggggccagatggttaagccctcccgatcgtagtattcta cacgacggggagtcaggcaactatggatgaacgaatagacagatcgtgagataggtgcctcactgattaag cattggttaactgtcagaccaagtttactcatatatacttttagattgatttaaaactcatttttaatttaaaagatctagg tgaagatccttttgataatctcatgacaaaatcccttaacgtgagtttctggtccactgagcgtcagaccccgtag aaaagatcaaaagatcttctgagatcctttttctgcgcgtaactctgctgctgcaacaaaaaaaccaccgctac cagcgggtggtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagat accaaatactgtccttctagttagccgtatgttagccaccactcaagaactctgtagcaccgcctacatactc gctctgtaactcctgttaccagtggctgctgccagtggcgataagctgtgtcttaccgggttgactcaagacgat agttaccggataaggcgcagcggctgggctgaacggggggtcgtgcacacagcccagcttgagcgaacg acctacaccgaactgagatacctacagcgtgagctatgagaaagcggcagcctccgaaggagaaaggcg gacaggtatccggttaagcggcagggcgggaacaggagagcgcacgagggagcttcagggggaaacgcct ggtatctttatagctcgtcgggtttcggccactctgacttgagcgtcgattttgtgatgctcgtcagggggcgga gcctatggaaaaacgccagcaacgcggccttttacggttctggccttttctggtccttttctcacatgttcttcc tgcgttatccccgtattctgttgataaccgtattaccgctttgagtgagctgataccgctcggcgagccgaacg accgagcgcagcagtcagtgagcaggaagcggagagcgcctgatgcgggtatttttctccttacgcatctgt gcgggtatttcacaccgcatatagtgctcactctcagtaacaatctgctctgatccgcatagtaagccagtatac tccgctatcgtactgactgggtcatggtcgcgccccgacaccgccaacaccgctgacgcgcctgacgg gcttgtctgctccggcatccgcttacagacaagctgtgaccgtctcgggagctgcatgtgtcagaggtttcac cgtcatcaccgaaacgcgcgagggagctgcggtaagctcatcagcgtggtcgtgaagcgattcacagatgtc tgcctgttcatccgctccagctcgttgagtttctcagaagcgttaatgtctggctctgataaagcgggcatgtt aagggcggttttctgtttggctactgatgcctcgtgtaaggggatttctgttcatgggggtaatgataccgat gaaacgagagaggatgctacgatacgggttactgatgatgaacatgccggttactggaacgttgtaggggta aacaactggcggtatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgcttcgtaatac agatgtaggtgtccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcaggcgct gacttccgctttcagactttacgaaacacggaaaccgaagaccattcatgttgtgctcaggtcgcagacgttt gcagcagcagtcgcttcacgttcgctcgcgtatcgggtgattcattctgtaaccagtaaggcaacccgccagcc tagccgggtcctcaacgacaggagcagcatatgcgcaccgtggccaggaccaacgctgccgagatctc gatcccgcgaaattaatacgaactactataggagaccacaacggttccctctagaaataattttgttaacttta gaaggagatatacatatggctagcatgactagcaaaaggagaagaacttttactggagttgtcccaattctgttg aattagatgggtgatgtaaatgggcacaaattttctgtagtggaaggggtgaaggtgatgtacatacggaaagct taccctaaatttattgactactggaaaactacgtgtccatggccaacacttgcactactttcttattggtgtca atgtctttcccggtatccggatcatatgaacggcatgacttttcaagagtgccatgccgaagggttatgtacagg aacgcactatatcttcaaagatgacgggaactacaagacgcgtgctgaagtcaagttgaaaggtgataccctgt taatcgtatcgagttaaaaggattgttttaagaagatggaacattctcggacacaaactcagatcaactata actcacacaatgtatacatcacggcagacaaaagaatggaatcaaagctaacttcaaaattcggcacaac attgaagatggatccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtcctttacc agacaaccattacgtgcacacaatctgcccttaccgaagatcccaacgaaaaagagagaccacatggtccttc ttgagttgtaacagctgctgggattacacatggcatggatgaactatacaaacgggagatccggcgggcgcc gcaactgagcaccaccaccaccactgagatccggctgctaacaagccggaaggaagctgagttggct gctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttgaggggtttttgctgaaa ggaggaactatatccggat

[illegible]

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	gcgaaattaatacgaactactatagggagatattttacaacaattaccaacaacaacaacaacaacattac aattactatttacaattacatatggaagacgccaaaaacataaagaaaggccggcgccattctatccgctagag gatggaaccgctggagagcaactgcataaggctatgaagagatacgcctggttcttggaacaattgcttttaca gatgcacatatcgagggtgaacatcacgtacgcgggaatacttcgaaatgtccgttcggttggcagaagctatgaaa cgatatgggctgaatacaaatcacagaatcgtcgtatgcagtgaaaactctctcaattcttattgccgggtgtggg cgcgttatttatcgagggtgcagttgcgccgcgaacgacatttataatgaacgtgaattggtcaacagtatgaaca tttcgcagcctaccgtagtgtttgttccaaaaagggttgcaaaaaatttgaacgtgcaaaaaaattaccaata atccagaaaattattatcatggattctaaaacggattaccagggttcagtcgatgtacacgttcgtcacatctcat ctacctcccggttttaataatgaatacgtttgtaccagagtccttgcgtgacaaaaaattgcactgataatgaact cctctggatctactgggttacctaagggtgtggccctccgcatagaactgcctgcgtcagattctgcagtcgcag agatcctattttggcaatcaaatattccggatactgcgattttaagtgtgttccattccatcacggttttggaatgtt tactacactcggatatttgatatgtggatttcgagtcgtcttaatgtatagattgaagaagagctgttttacgatccc ttcaggattacaaaattcaaaagtcgttgctagtaccaacccattttcattcttcgcaaaaagcactctgattgacaa atacgtttatctaatttacacgaaattgcttctgggggcgcacctcttcgaaagaagtcggggaagcgggtgca aaacgcttccatcttcagggtacgacaaggatattgggtcactgagactacatcagctattctgattacaccg aggggggatgataaaccgggcgcggtcggtaaaagtgttccatttttgaaagcgaagggtgtggatctggataccg ggaaaacgctgggcgttaatcagagaggcgaattatgtgtcagaggacctatgattatgtccggttatgtaaca atccgggaagcgaccaacgccttgattgacaaggatggatggctacattctggagacatagcttactgggacgaa gacgaacacttctcatagttgaccgcttgaagtccttaattaaatacaaaaggataaccaggtggccccgcgtgaatt ggagtcgatattgttacaacaccccaacatcttcgacgcgggcgtggcaggtctcccgacgatgacgccggtg aactcccgcgcgctgtgttttgagacacggaaagacgatgacggaaaaagagatcgtggattacgtcgc agtcaagtaacaaccgcaaaaaggtgcgcggaggaggtgtgtgttgagcgaagtaccgaaaggtcttaccgg aaaactcgacgcaagaaaaatcagagagatcctcataaaggccaagaaggcggaagtcgcaaaattgtaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaactcgagcaccaccaccaccactgagatccggctgtaacaaaagcccgaagg aagctgagttggctgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttgagg ggtttttgtcgaaggagggaactatatccgattggcgaatgggacgcgcctgtagcggcgcatgaagcgcg gcgggtgtggtggttacgcgcagcgtgaccgctacatt
backbone sense primer (SEQ ID NO: 129)	gtgattcattctgctaaccag
backbone anti-sense primer (SEQ ID NO: 130)	ccccaagggttatgctagt
T7-Ω-f (SEQ ID NO: 131)	ccgcgaaattaatacgaactactatagggagatattttacaacaattaccaacaacaac
SP6-Ω-f (SEQ ID NO: 132)	ccgcgaaatatttaggtgacactatagaagagtattttacaacaattaccaacaacaac
T3-Ω-f (SEQ ID NO: 133)	ccgcgaaataattaacccctactaaagggaatattttacaacaattaccaacaacaac
P1.1-Ω-Kozak(No Kozak)-sfGFP-f (SEQ ID NO: 134)	acaaacaacattacaattactatttacaattacatatgagcaaaaggtgaagaactgt
P1.1-Ω-Kozak(Cons ensus Full)-sfGFP-f (SEQ ID NO: 135)	acaaacaacattacaattactatttacaattaccaccatggagcaaaaggtgaagaactgt
P1.1-Ω-Kozak(<i>S. cer</i> <i>partial</i>)-sfGFP-f (SEQ ID NO: 136)	acaaacaacattacaattactatttacaattaaaaaaatgagcaaaaggtgaagaactgt
P1.1-Ω-Kozak(<i>S. cer</i> <i>full</i>)-sfGFP-f	acaaacaacattacaattactatttacaattaaaaaaatgtctagcaaaaggtgaagaactgt

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
(SEQ ID NO: 137)	
Ω-(No Kozak)-sfGFP (SEQ ID NO: 138)	acgctgcccagagatctcgatccgcgaaattaatacgactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattacatatgtctagcaaaagggtgaagaactgtttaccggc gttggtgccgattctggtggaactggatggcgtgtgaacggtcacaaattcagcgtgcgtggtgaagggtgaagg cgatgccacgattggcaaaactgacgctgaaatttatctgcaccaccggcaaaactgccggtgccgtggccgacg ctggtgaccaccctgacctatggcgttcagtggttttagtcgctatccggatcacatgaacgtcacgatttctttaa tctgcaatgccggaaggctatgtgcaggaacgtacgattagctttaagatgatggcaaatataaaacgcgcgc cgttgtaaatgtgaaggcgataccctggtgaaccgattgaactgaaaggcacggattttaagaagatggcaa tatcctgggccataaactggaatacaactttaatagccataatgtttatattacggcggataaacagaaaaatggca tcaaagcgaattttaccgttcgccataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaatac cccgattggtgatggtccggtgctgctgccggataatcattatctgagcacgcagaccgttctgtctaaagatccg aacgaaaaaggcacgcgggaccacatggttctgcacgaatatgtgaatgcggcaggtattacgtggagccatc cgagttcgaaaaataaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaa
Ω-Kozak (Consensus)-sfGFP (SEQ ID NO: 139)	acgctgcccagagatctcgatccgcgaaattaatacgactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattaccaccatggagcaaaagggtgaagaactgtttaccg gcgttggtccgattctggtggaactggatggcgtgtgaacggtcacaaattcagcgtgcgtggtgaagggtgaa ggcgtatgccacgattggcaaaactgacgctgaaatttatctgcaccaccggcaaaactgccggtgccgtggccga cgttggtgaccaccctgacctatggcgttcagtggttttagtcgctatccggatcacatgaacgtcacgatttctta aatctgcaatgccggaaggctatgtgcaggaacgtacgattagctttaagatgatggcaaatataaaacgcgc gccgttgtaaatgtgaaggcgataccctggtgaaccgattgaactgaaaggcacggattttaagaagatggc aatatcctgggccataaactggaatacaactttaatagccataatgtttatattacggcggataaacagaaaaatgg catcaaagcgaattttaccgttcgccataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaat accccattggtgatggtccggtgctgctgccggataatcattatctgagcacgcagaccgttctgtctaaagatc cgaacgaaaaaggcacgcgggaccacatggttctgcacgaatatgtgaatgcggcaggtattacgtggagcca tccgcagttcgaaaaataaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaa
Ω-Kozak (<i>S. cer</i> partial)-sfGFP (SEQ ID NO: 140)	acgctgcccagagatctcgatccgcgaaattaatacgactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattaaaaaaatgagcaaaagggtgaagaactgtttaccgg cgttggtgccgattctggtggaactggatggcgtgtgaacggtcacaaattcagcgtgcgtggtgaagggtgaag gcgatgccacgattggcaaaactgacgctgaaatttatctgcaccaccggcaaaactgccggtgccgtggccgac gctggtgaccaccctgacctatggcgttcagtggttttagtcgctatccggatcacatgaacgtcacgatttcttaa atctgcaatgccggaaggctatgtgcaggaacgtacgattagctttaagatgatggcaaatataaaacgcgcgc ccgttgtaaatgtgaaggcgataccctggtgaaccgattgaactgaaaggcacggattttaagaagatggca atatcctgggccataaactggaatacaactttaatagccataatgtttatattacggcggataaacagaaaaatggc atcaaagcgaattttaccgttcgccataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaata ccccgattggtgatggtccggtgctgctgccggataatcattatctgagcacgcagaccgttctgtctaaagatcc gaacgaaaaaggcacgcgggaccacatggttctgcacgaatatgtgaatgcggcaggtattacgtggagccat ccgcagttcgaaaaataaactgagatccggctgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaa
Ω-Kozak(<i>S. cer</i> full)-sfGFP (SEQ ID NO: 141)	acgctgcccagagatctcgatccgcgaaattaatacgactactatagggagtattttacaacaattaccaacaa caacaacaacaacaacattacaattactatttacaattaaaaaaatgtctagcaaaagggtgaagaactgtttacc ggcgttggtccgattctggtggaactggatggcgtgtgaacggtcacaaattcagcgtgcgtggtgaagggtga aggcgatgccacgattggcaaaactgacgctgaaatttatctgcaccaccggcaaaactgccggtgccgtggccg acgctggtgaccaccctgacctatggcgttcagtggttttagtcgctatccggatcacatgaacgtcacgatttctt aaatctgcaatgccggaaggctatgtgcaggaacgtacgattagctttaagatgatggcaaatataaaacgcgc gccgttgtaaatgtgaaggcgataccctggtgaaccgattgaactgaaaggcacggattttaagaagatggc aatatcctgggccataaactggaatacaactttaatagccataatgtttatattacggcggataaacagaaaaatgg catcaaagcgaattttaccgttcgccataacgttgaagatggcagtggtgcagctggcagatcattatcagcagaat

Name (SEQ ID NO:)	Sequence (5' → 3' if nucleotides; N → C if amino acids)
	accccgattggtgatggtccggtgctgctgccggataatcattatctgagcacgcagaccgttctgtctaaagatc cgaacgaaaaaggcacgcgggaccacatggttctgcacgaatatgtgaatgcggcaggtattacgtggagcca tccgcagttcgaaaaataaactgagatccgggtgctaacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaa

Additional Acknowledgements

[0156] All references, including publications, patent applications, and patents, cited herein are hereby incorporated by reference to the same extent as if each reference were individually and specifically indicated to be incorporated by reference and were set forth in its entirety herein.

[0157] All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0158] Preferred aspects of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those preferred aspects may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect a person having ordinary skill in the art to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

CLAIMS

What is claimed is:

1. A cell-free protein synthesis platform for preparing protein from a translation template, comprising:
 - (a) a *Saccharomyces cerevisiae* cellular extract prepared from mid-exponential to late-exponential batch cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀ or fed-batch cultures harvested in mid-exponential to late-exponential phase;
 - (b) a reaction buffer; and
 - (c) the translation template.
2. The cell-free protein synthesis platform of claim 1, wherein the *Saccharomyces cerevisiae* cellular extract comprises a crude extract.
3. The cell-free protein synthesis platform of claim 1, wherein the *Saccharomyces cerevisiae* cellular extract comprises an S30 extract.
4. The cell-free protein synthesis platform of claim 1, wherein the *Saccharomyces cerevisiae* cellular extract comprises an S60 extract.
5. The cell-free protein synthesis platform of any of claims 1-4, wherein the reaction buffer comprises NTPs, spermidine, putrescine, a glutamate salt, a magnesium salt and glycerol.
6. The cell-free protein synthesis platform of any of claims 1-4, wherein the reaction buffer comprises glycerol.
7. The cell-free protein synthesis platform of any of claims 1-4, wherein the reaction buffer comprises at least one component selected from the group consisting of NTPs, a polyamine, an organic anion, a divalent cation, an alcohol and combinations thereof.
8. The cell-free protein synthesis platform of claim 7, wherein the polyamine is selected from spermidine and putrescine, the organic anion is selected from glutamate and acetate, the divalent cation is selected from magnesium, calcium and manganese; and the alcohol comprises glycerol.
9. The cell-free protein synthesis platform of any of claims 1-8, wherein the *Saccharomyces cerevisiae* cellular extract is not pre-treated with a micrococcal nuclease.
10. The cell-free protein synthesis platform of any of claims 1-9, wherein the translation template comprises at least one RNA.
11. The cell-free protein synthesis platform of claim 10, wherein the at least one RNA comprises an open reading frame, a 5'-UTR and a 3'-UTR.

12. The cell-free protein synthesis platform of claim 11, wherein the 5'-UTR comprises a cap-independent translation enhancing element.
13. The cell-free protein synthesis platform of claim 12, wherein the cap-independent translation enhancing element is selected from a TMV Ω sequence, a TEV 5'-UTR element, and a Tbm 5'-UTR element.
14. The cell-free protein synthesis platform of claims 12 or 13, wherein the 5'-UTR further comprises a yeast Kozak sequence or variant thereof.
15. The cell-free protein synthesis platform of any of claims 11-14, wherein the 3'-UTR comprises a poly(A)_n 3'-terminus, where n ranges from about 20 nucleotides to about 200 nucleotides in length.
16. The cell-free protein synthesis platform of claim 15, wherein n is selected from 25 nucleotides, 50 nucleotides, 90 nucleotides, and 170 nucleotides.
17. A cell-free protein synthesis platform for preparing protein from a transcription template, comprising:
 - (a) a *Saccharomyces cerevisiae* cellular extract prepared from mid-exponential to late-exponential cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀;
 - (b) a reaction buffer;
 - (c) an RNA polymerase; and
 - (d) the transcription template,wherein the RNA polymerase is capable of transcribing the transcription template to form a translation template and the *Saccharomyces cerevisiae* cellular extract can sustain protein synthesis through a combined transcription/translation reaction.
18. The cell-free protein synthesis platform of claim 17, wherein the *Saccharomyces cerevisiae* cellular extract comprises an S30 extract.
19. The cell-free protein synthesis platform of claim 17, wherein the *Saccharomyces cerevisiae* cellular extract comprises an S60 extract.
20. The cell-free protein synthesis platform of any of claims 17-19, wherein the reaction buffer comprises NTPs, spermidine, putrescine, a glutamate salt, a magnesium salt and glycerol.
21. The cell-free protein synthesis platform of any of claims 17-19, wherein the reaction buffer comprises glycerol.
22. The cell-free protein synthesis platform of any of claims 17-19, wherein the reaction buffer comprises at least one component selected from the group consisting of NTPs, a polyamine, an organic anion, a divalent cation, an alcohol and combinations thereof.

23. The cell-free protein synthesis platform of claim 22, wherein the polyamine is selected from spermidine and putrescine, the organic anion is selected from glutamate and acetate, the divalent cation is selected from magnesium, calcium and manganese; and the alcohol comprises glycerol.
24. The cell-free protein synthesis platform of any of claims 17-23, wherein the *Saccharomyces cerevisiae* cellular extract is not pre-treated with a micrococcal nuclease.
25. The cell-free protein synthesis platform of any of claims 17-24, wherein the RNA polymerase is selected from SP6 RNA Polymerase, T3 RNA Polymerase and T7 RNA polymerase.
26. The cell-free protein synthesis platform of any of claims 17-24, wherein the RNA polymerase is selected from T3 RNA Polymerase and T7 RNA polymerase.
27. The cell-free protein synthesis platform of any of claims 17-24, wherein the RNA polymerase comprises T7 RNA polymerase.
28. The cell-free protein synthesis platform of any of claims 17-28, wherein the transcription template comprises at least one DNA.
29. The cell-free protein synthesis platform of claim 28, wherein the at least one DNA comprises a linear DNA or a circular DNA.
30. The cell-free protein synthesis platform of claim 28, wherein the at least one DNA encodes an open reading frame, a 5'-UTR and a 3'-UTR operably linked to a promoter specific for the RNA polymerase.
31. The cell-free protein synthesis platform of claims 29 or 30, wherein the at least one DNA comprises a linear DNA prepared from an amplification reaction.
32. The cell-free protein synthesis platform of claim 31, wherein the amplification reaction comprises a polymerase chain reaction.
33. The cell-free protein synthesis platform of claim 30, wherein the 5'-UTR comprises a cap-independent translation enhancing element.
34. The cell-free protein synthesis platform of claim 33, wherein the cap-independent translation enhancing element is selected from a TMV Ω sequence, a TEV 5'-UTR element, and a Tbm 5'-UTR element.
35. The cell-free protein synthesis platform of claims 33 or 34, wherein the 5'-UTR further comprises a yeast Kozak sequence or variant thereof.
36. The cell-free protein synthesis platform of claim 30, wherein the 3'-UTR comprises a

poly(A)_n 3'-terminus, where n ranges from about 20 nucleotides to about 200 nucleotides in length.

37. The cell-free protein synthesis platform of claim 36, wherein n is selected from 25 nucleotides, 50 nucleotides, 90 nucleotides, and 170 nucleotides.

38. A method of performing high-throughput protein synthesis *in vitro*, comprising:

- (a) providing a source nucleic acid;
- (b) preparing a transcription template from the source nucleic acid; and
- (c) synthesizing protein *in vitro* using a cell-free protein synthesis platform utilizing

the transcription template, wherein the cell-free protein synthesis platform comprises:

- (i) a *Saccharomyces cerevisiae* cellular extract prepared from mid-exponential to late-exponential cultures in the range from about 6 OD₆₀₀ to about 18 OD₆₀₀;
- (ii) a reaction buffer; and
- (iii) an RNA polymerase,

wherein the RNA polymerase is capable of transcribing from the transcription template to form the translation template, and the *Saccharomyces cerevisiae* cellular extract can sustain protein synthesis through a combined transcription/translation reaction.

39. The method of claim 38, wherein the source nucleic acid comprises genomic DNA, cDNA, RNA or a combination thereof.

40. The method of claims 38 or 39, wherein the preparing a transcription template from the source nucleic acid comprises amplifying the source nucleic acid with a first primer and second primer in the presence of a DNA polymerase, wherein the first and second primers comprise gene-specific sequences capable of hybridizing the gene encoding an open reading frame in the source nucleic acid.

41. The method of any of claims 38-40, wherein the transcription template comprises a linear DNA encoding an open reading frame, a 5'-UTR and a 3'-UTR operably linked to a promoter specific for the RNA polymerase.

42. The method of any of claims 38-41, wherein the RNA polymerase is selected from SP6 RNA Polymerase, T3 RNA Polymerase and T7 RNA polymerase.

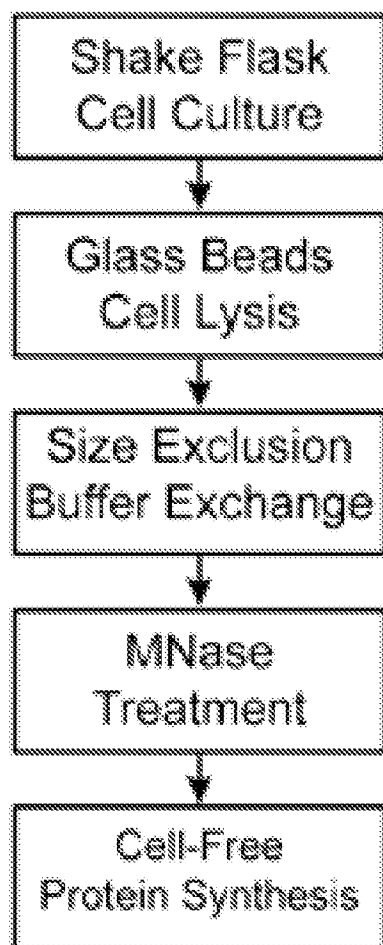
43. The method of any of claims 38-41, wherein the RNA polymerase is selected from T3 RNA Polymerase and T7 RNA polymerase.

44. The method of any of claims 38-41, wherein the RNA polymerase comprises T7 RNA polymerase.

45. The method of any of claims 31-44, wherein the 5'-UTR comprises a cap-independent translation enhancing element.

46. The method of claim 45, wherein the cap-independent translation enhancing element is selected from a TMV Ω sequence, a TEV 5'-UTR element, and a Tbm 5'-UTR element.
47. The method of claims 45 or 46, wherein the 5'-UTR further comprises a yeast Kozak sequence or variant thereof.
48. The method of any of claims 41-47, wherein the 3'-UTR comprises a poly(A)_n 3'-terminus, where n ranges from about 20 nucleotides to about 200 nucleotides in length.
49. The method of claim 48, wherein n is selected from 25 nucleotides, 50 nucleotides, 90 nucleotides, and 170 nucleotides.

A. PRIOR ART



B. INNOVATION

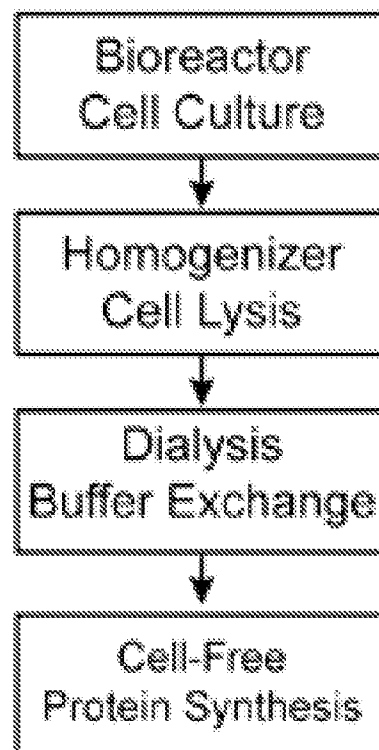


FIG. 1(A-B)

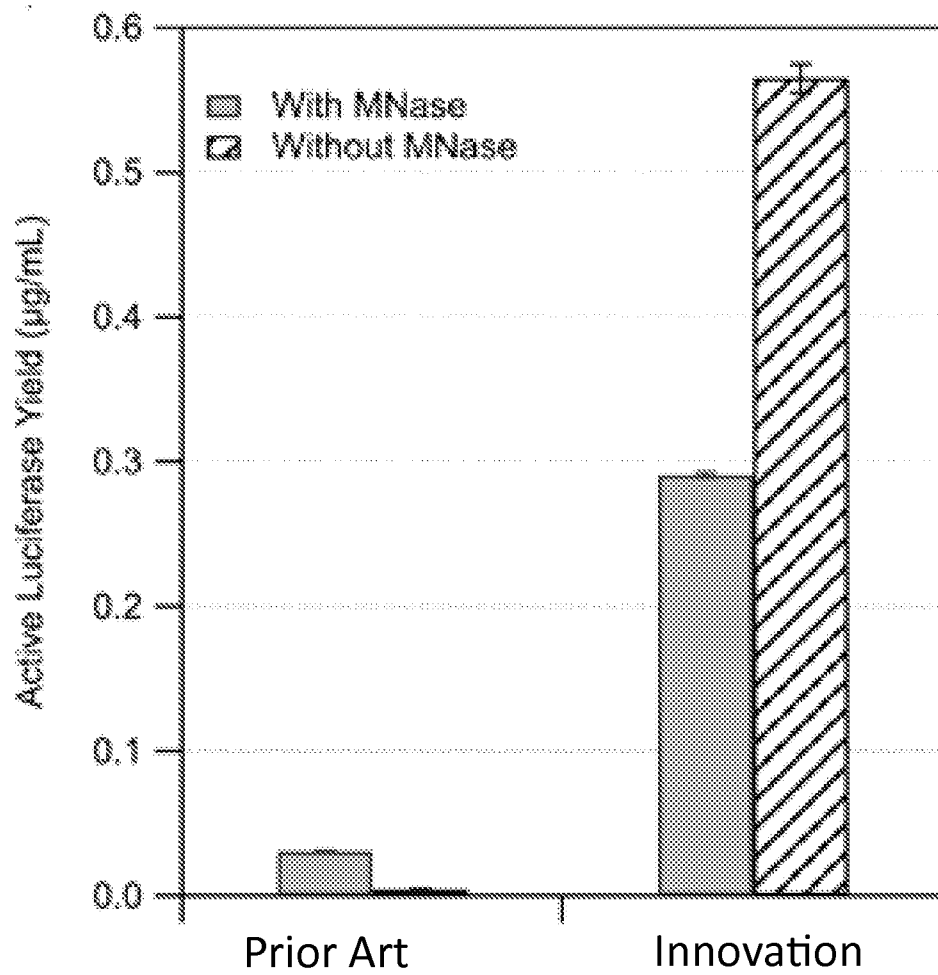


FIG. 1C

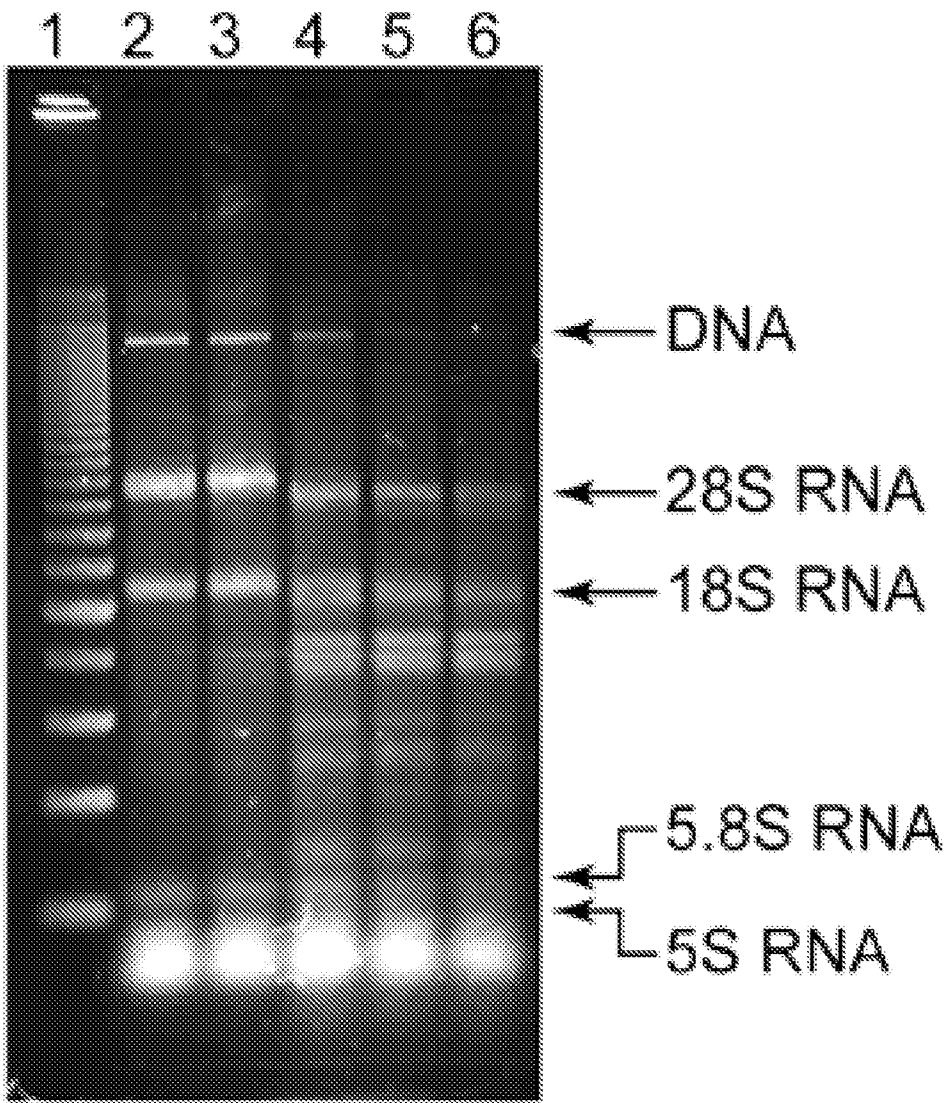


FIG. 1D

Construct	← 5' UTR →	← ORF →	← 3' UTR →	← PolyA →	SEQ ID NO:
CappA90	Cap analog	luciferase	None	90	1
HAPpA90	HAP270	luciferase	None	90	2
TFIIDpA90	TFIID	luciferase	None	90	3
p150pA90	p150	luciferase	None	90	4
YAP1pA90	YAP1	luciferase	None	90	5
IGRpA90	IGR	luciferase	None	90	6
A64pA90	PolyA64	luciferase	None	90	7
HedrinpA90	Polyhedrin	luciferase	None	90	8
ΩpA90	Ω	luciferase	None	90	9
TEVpA90	TEV	luciferase	None	90	10
TbmpA90	CFTbm	luciferase	None	90	11
ΩpA25	Ω	luciferase	None	25	12
ΩpA50	Ω	luciferase	None	50	13
ΩpA170	Ω	luciferase	None	170	14
ΩFBAL	Ω	luciferase	FBA1 1465nt	0	15
ΩFBAS	Ω	luciferase	FBA1 659nt	0	16
ΩTMV13U200	Ω	luciferase	TMV13U 200nt	0	17
ΩTMV13U400	Ω	luciferase	TMV13U 400nt	0	18
ΩTMV13U700	Ω	luciferase	TMV13U 700nt	0	19
ΩTMV23U	Ω	luciferase	TMV23U	0	20
ΩN3U	Ω	luciferase	None	0	21
N5UpA90	None	luciferase	None	90	22
ΩpA	Ω	luciferase	None	0	23
Luc	None	luciferase	None	0	24

FIG. 2

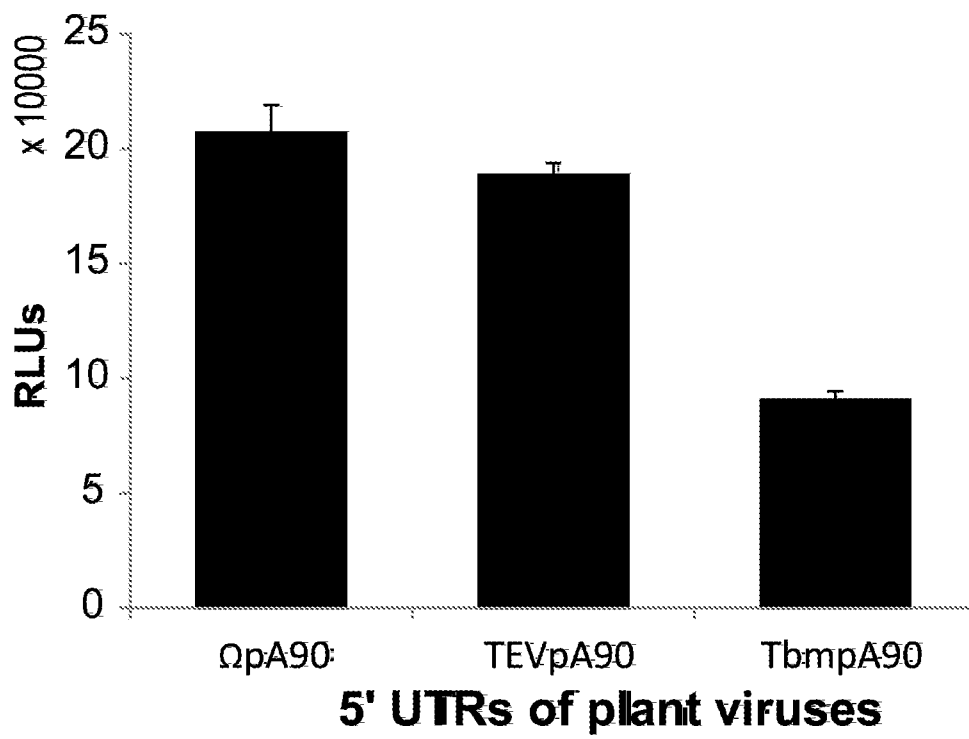


FIG. 3A

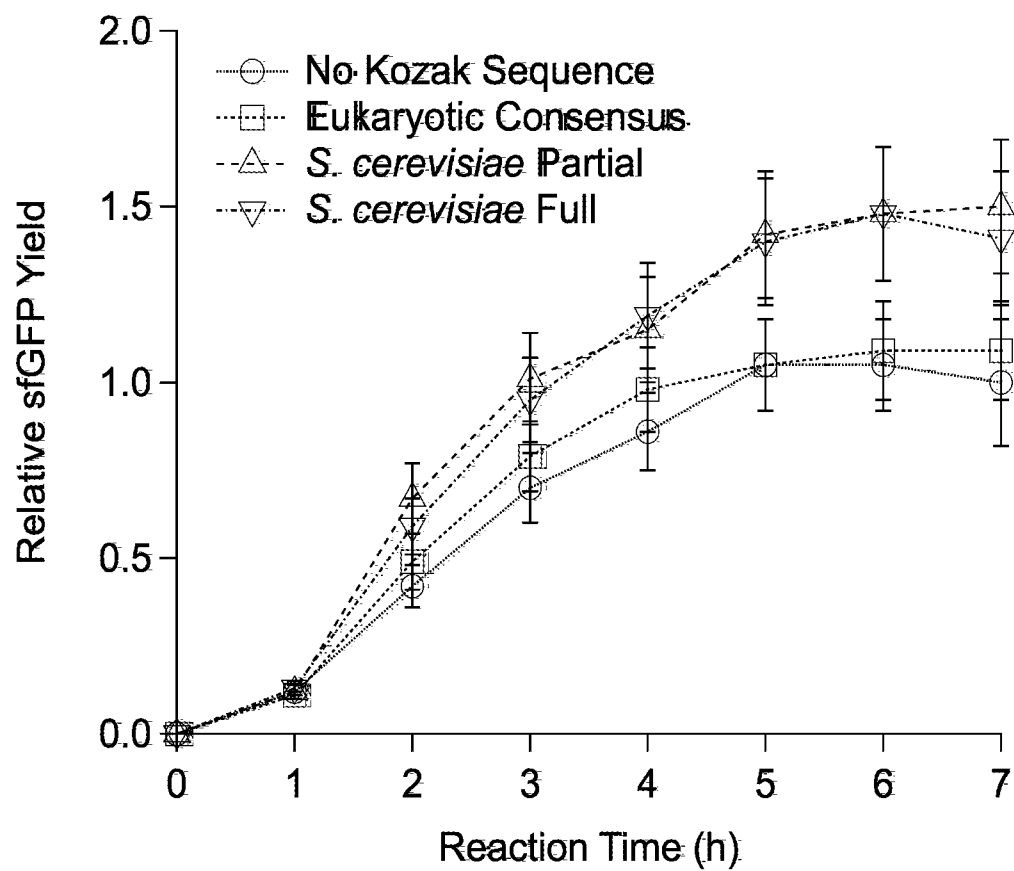


FIG. 3B

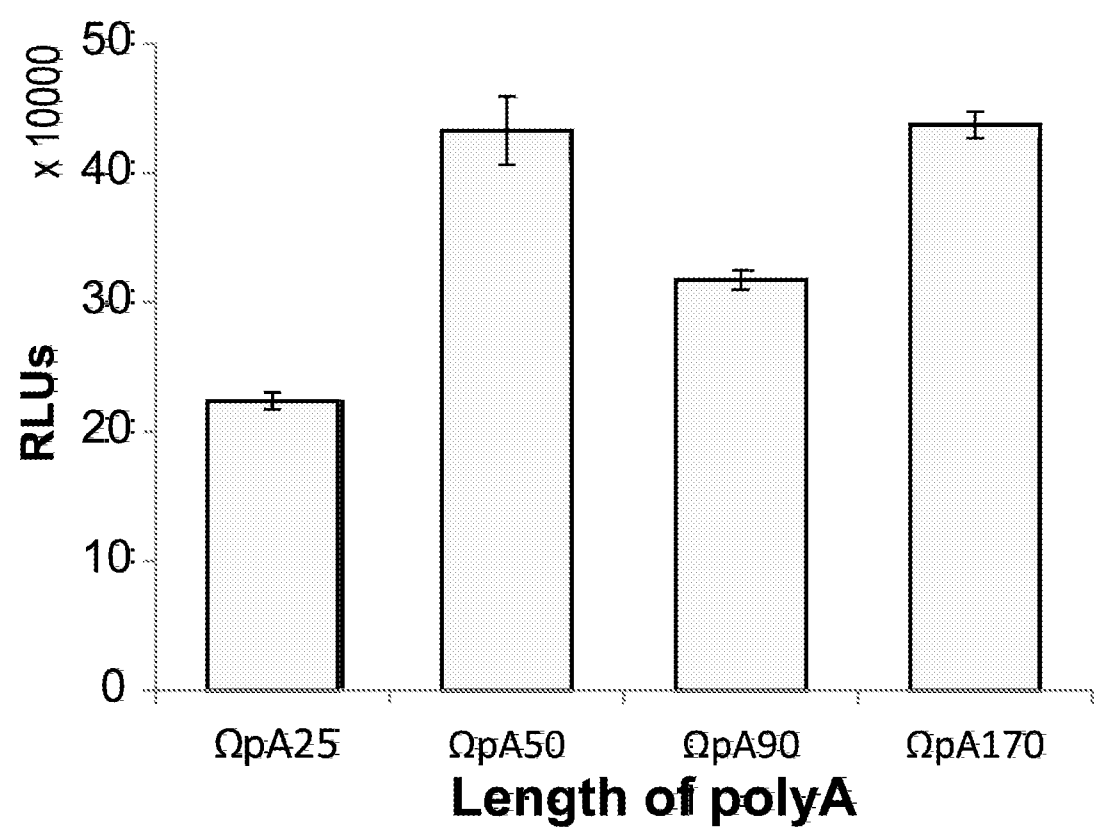


FIG. 3C

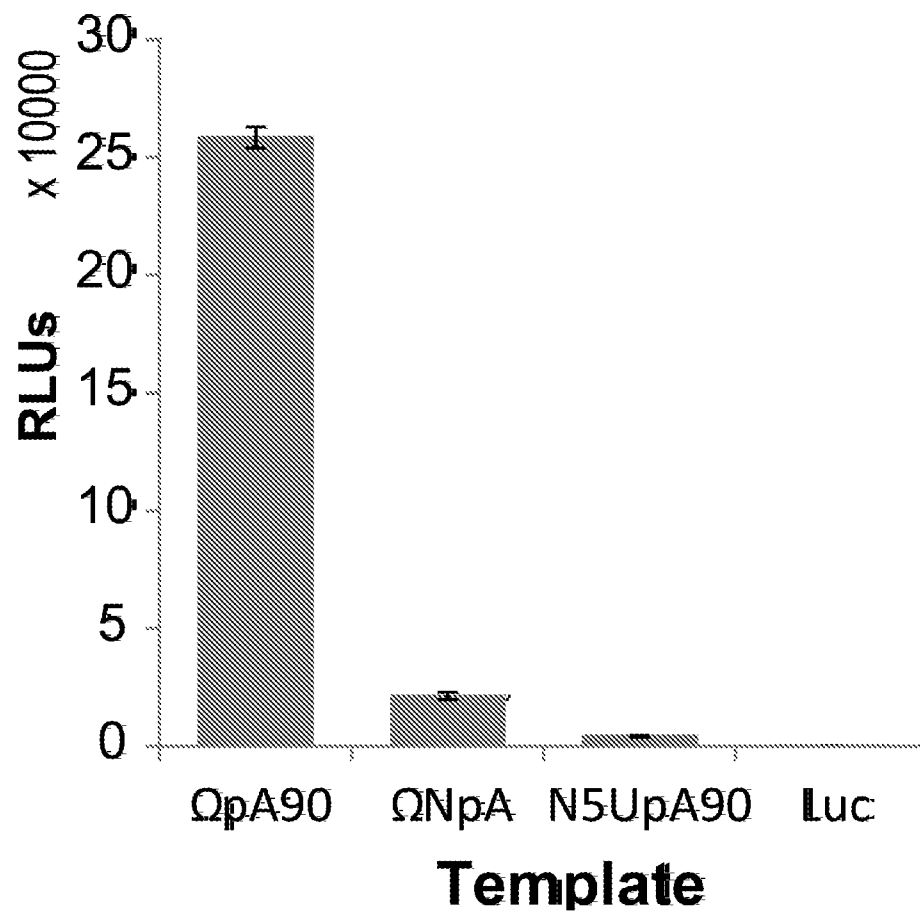


FIG. 3D

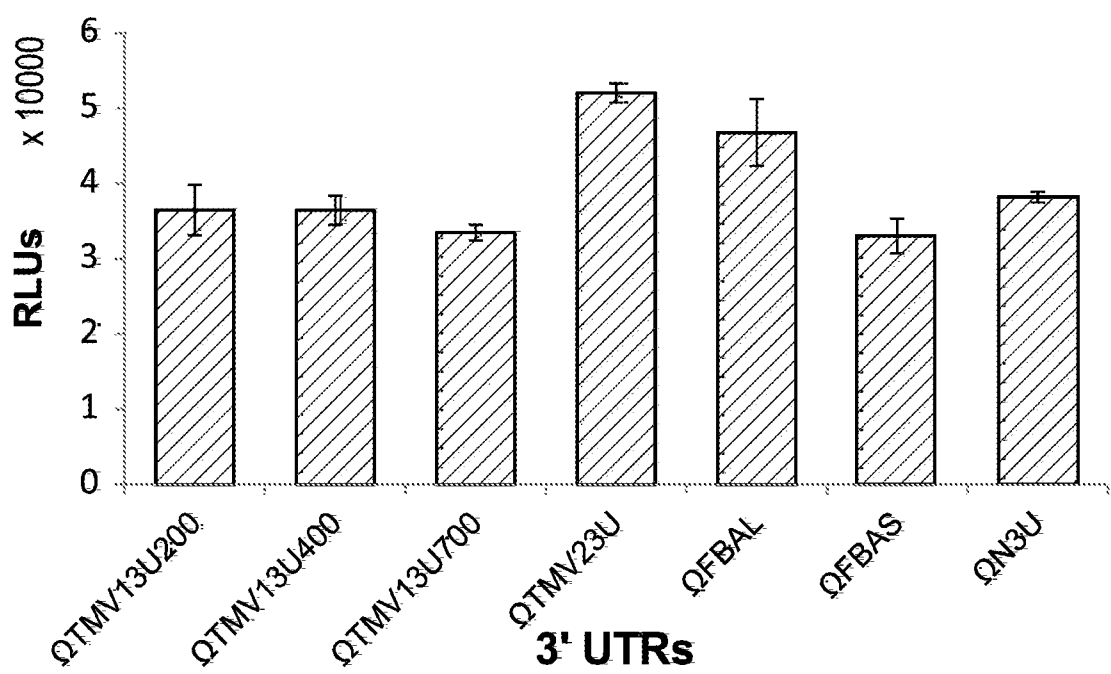


FIG. 3E

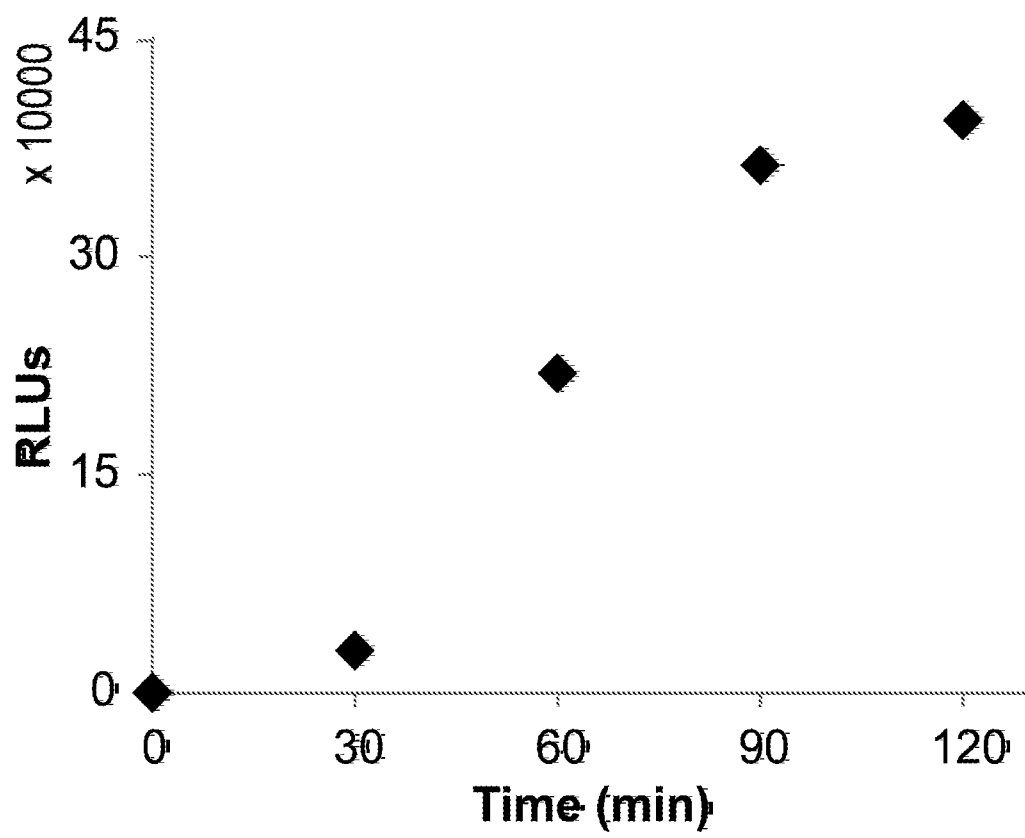


FIG. 4A

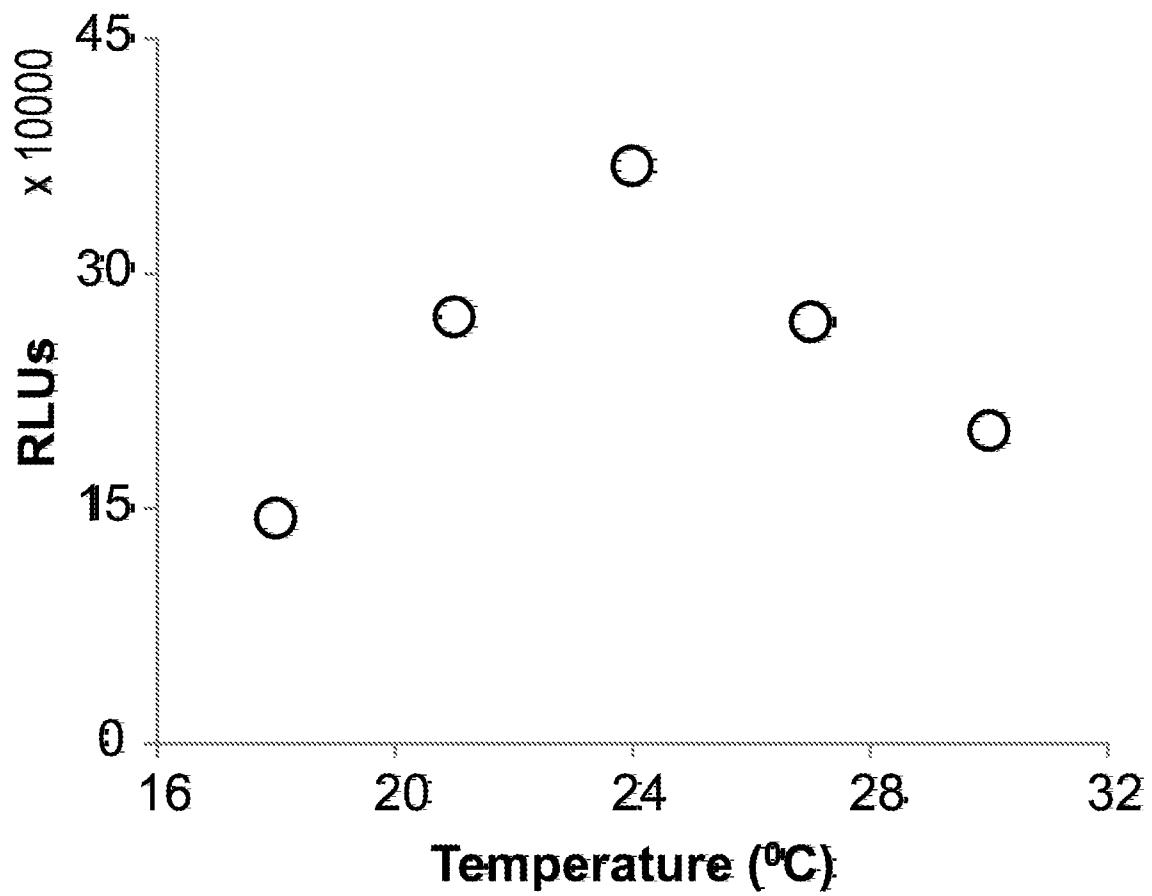


FIG. 4B

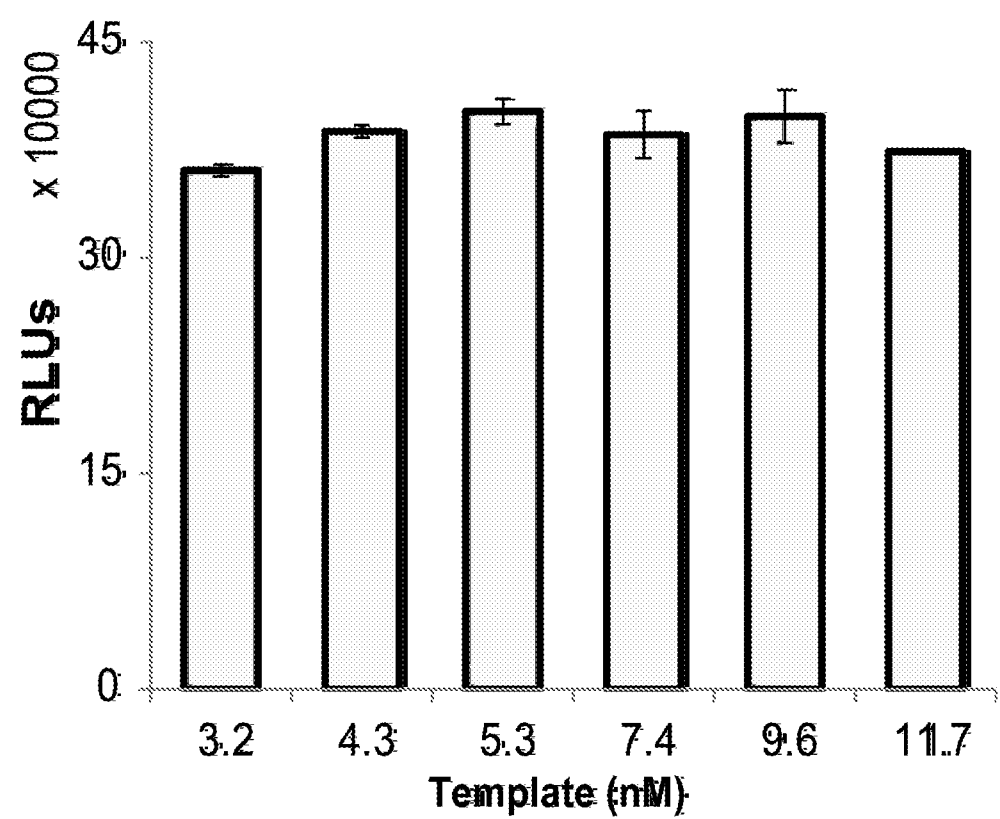


FIG. 4C

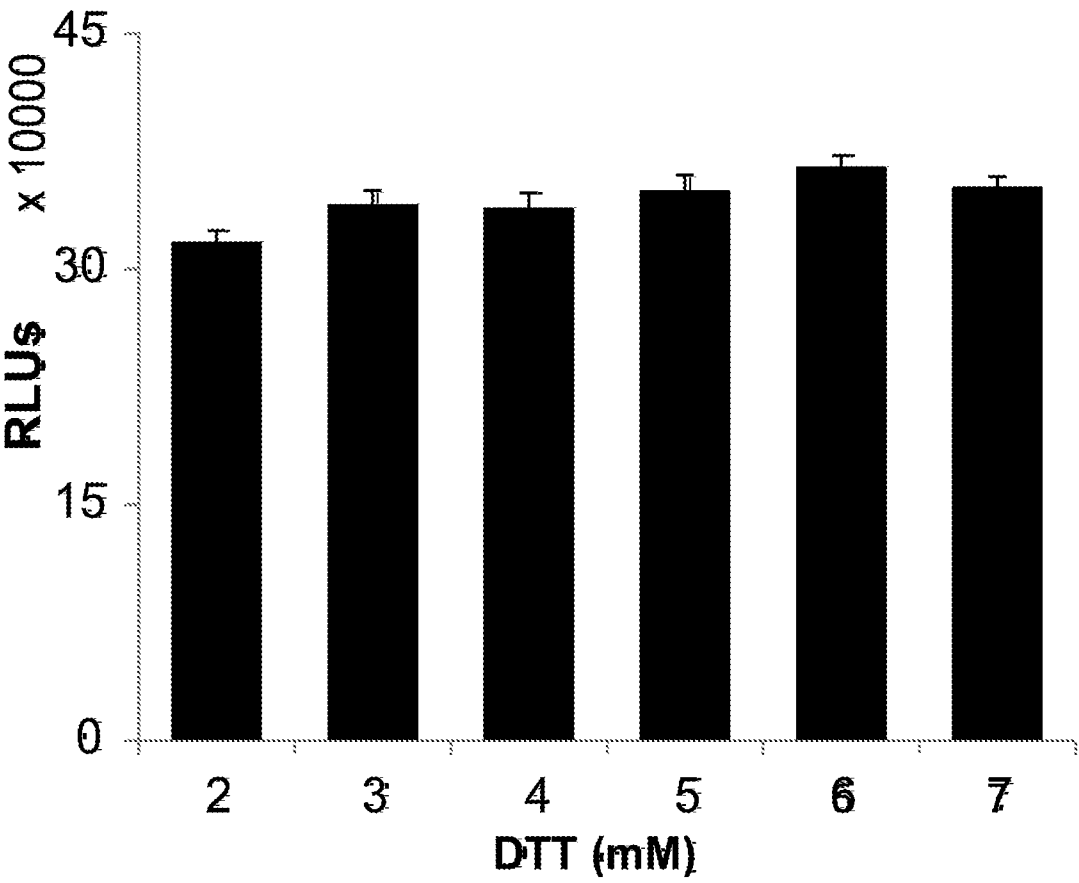


FIG. 4D

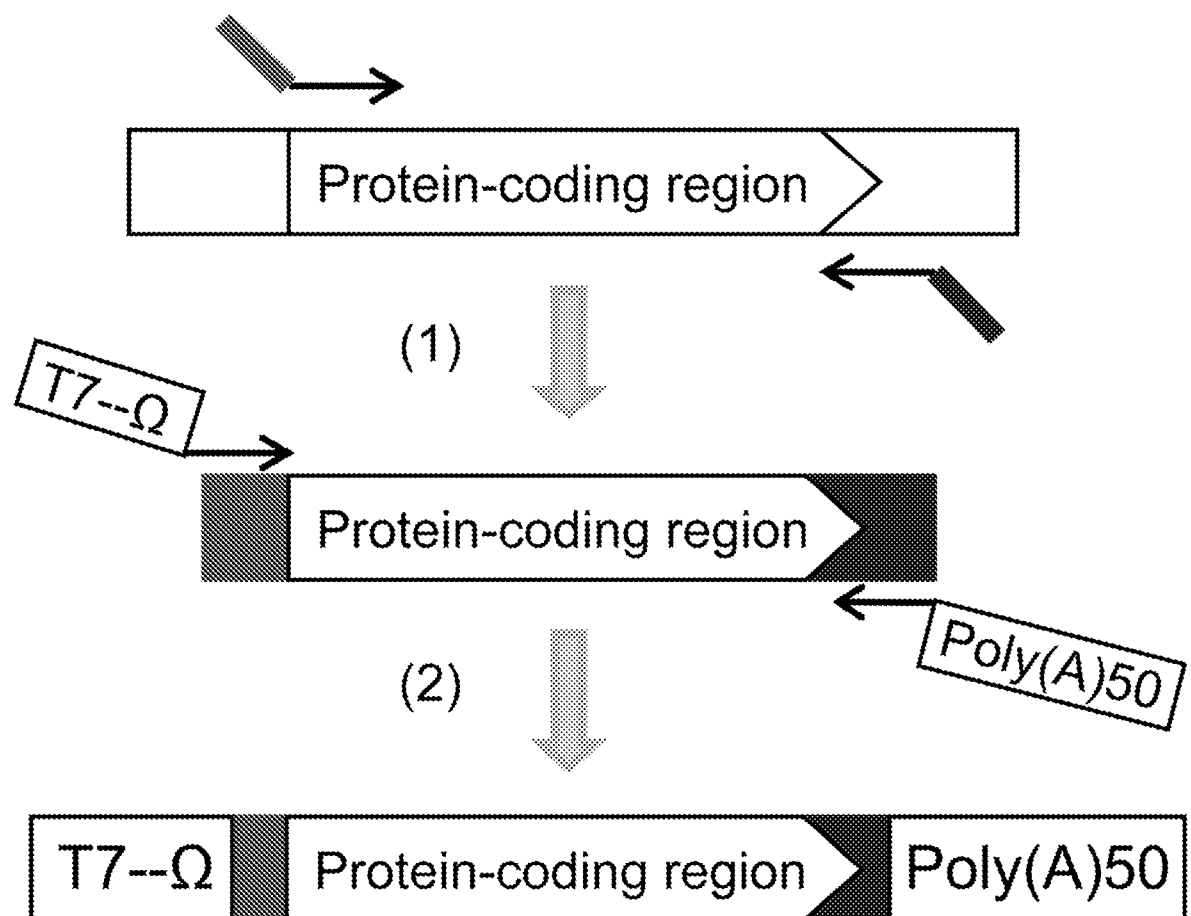


FIG. 5

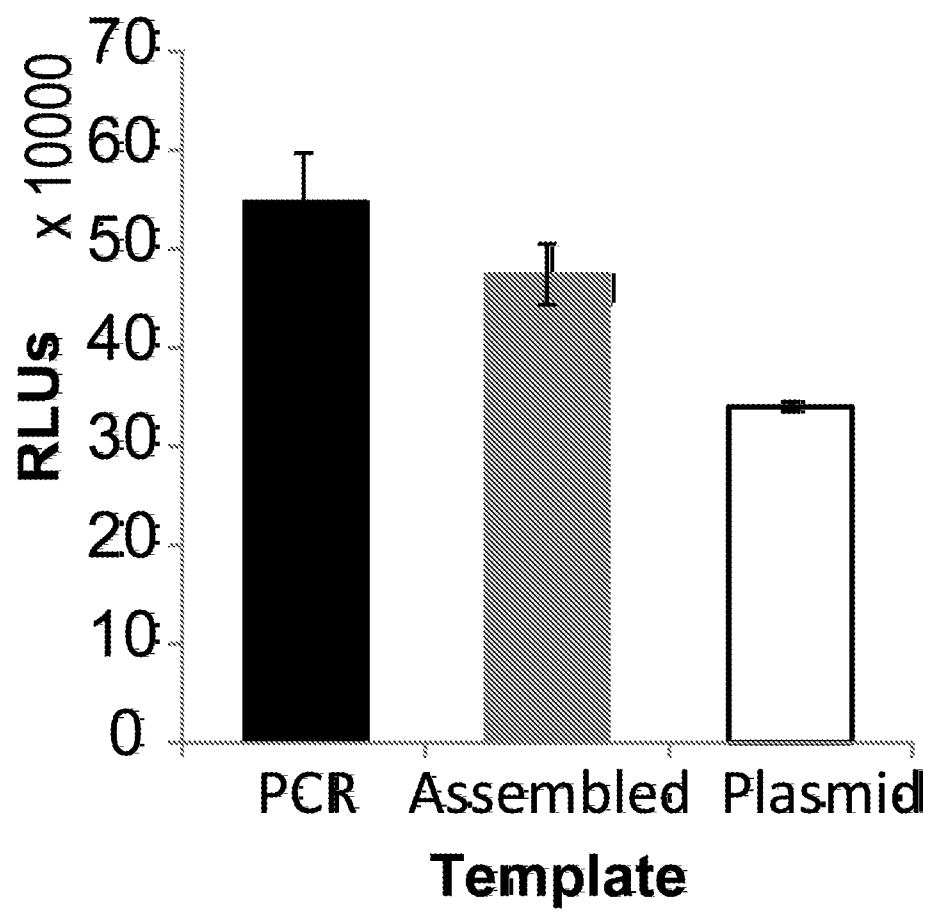


FIG. 6A

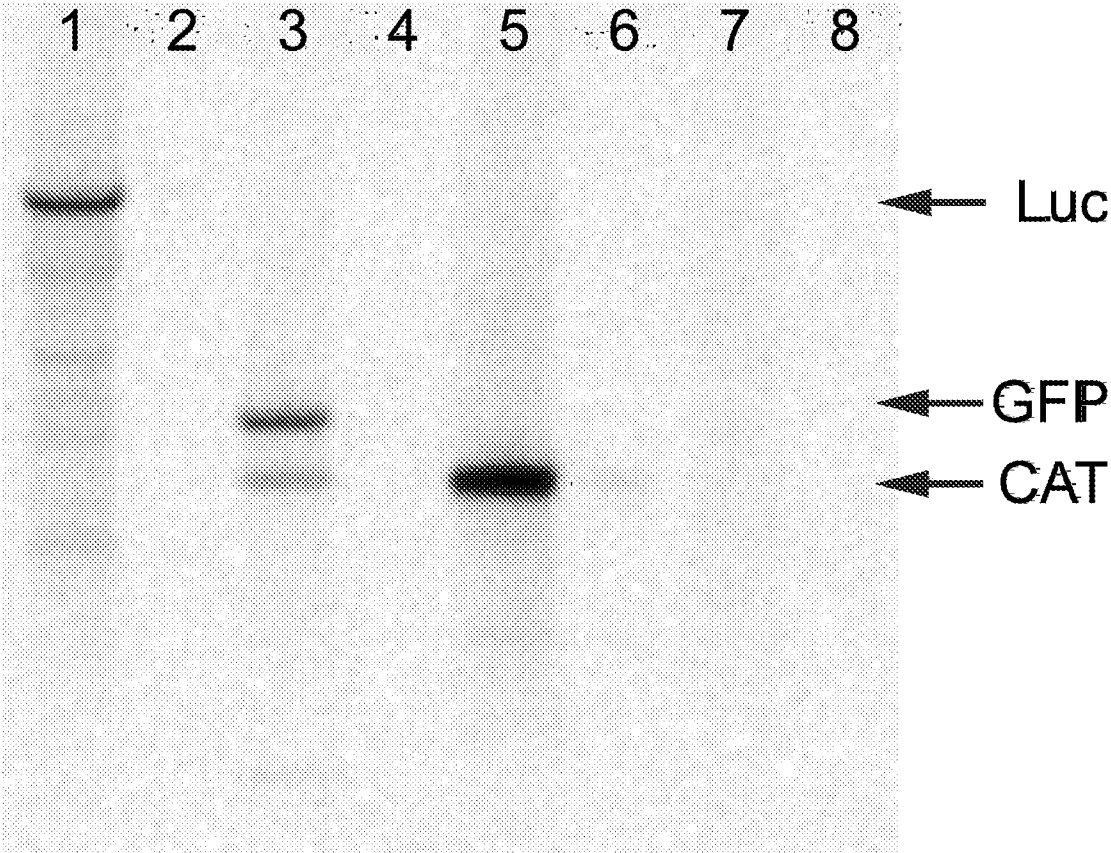


FIG. 6B

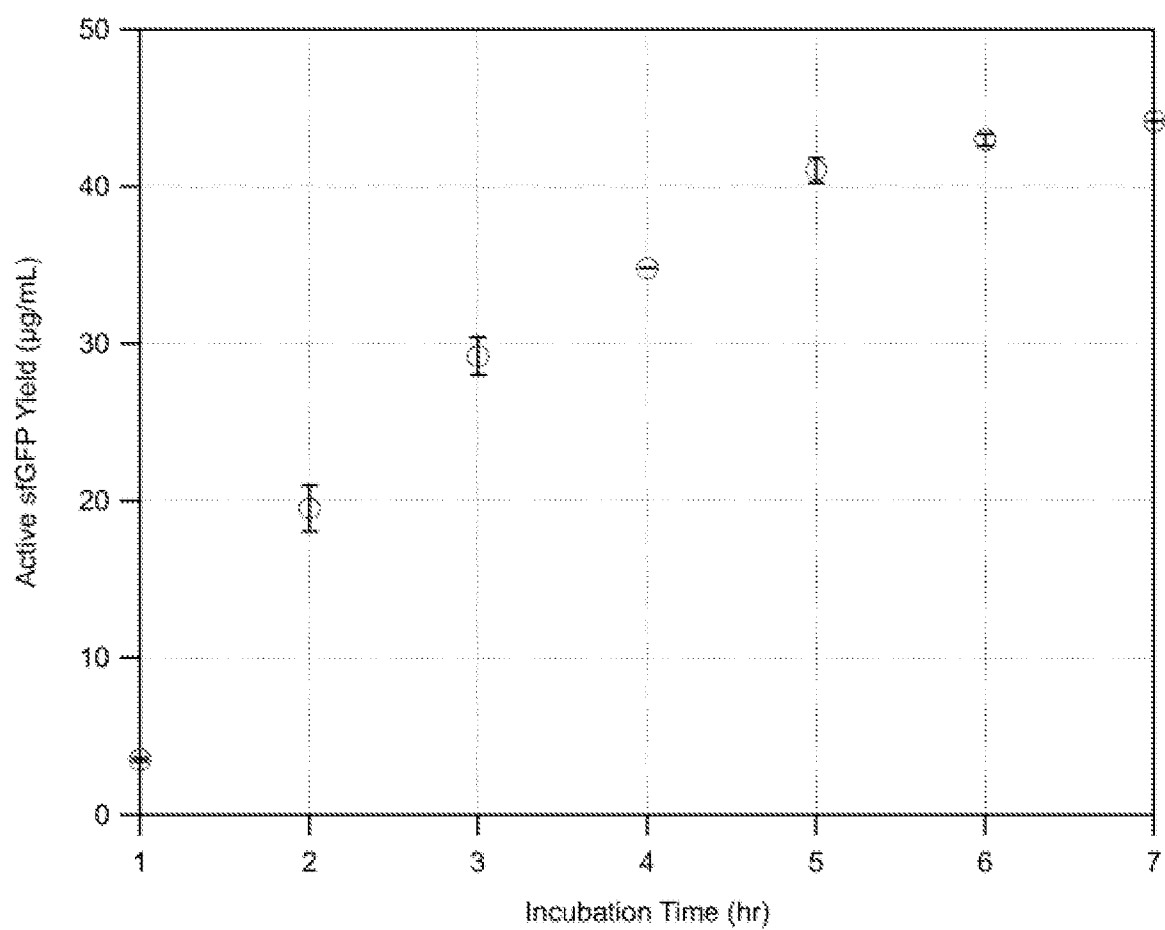


FIG. 7

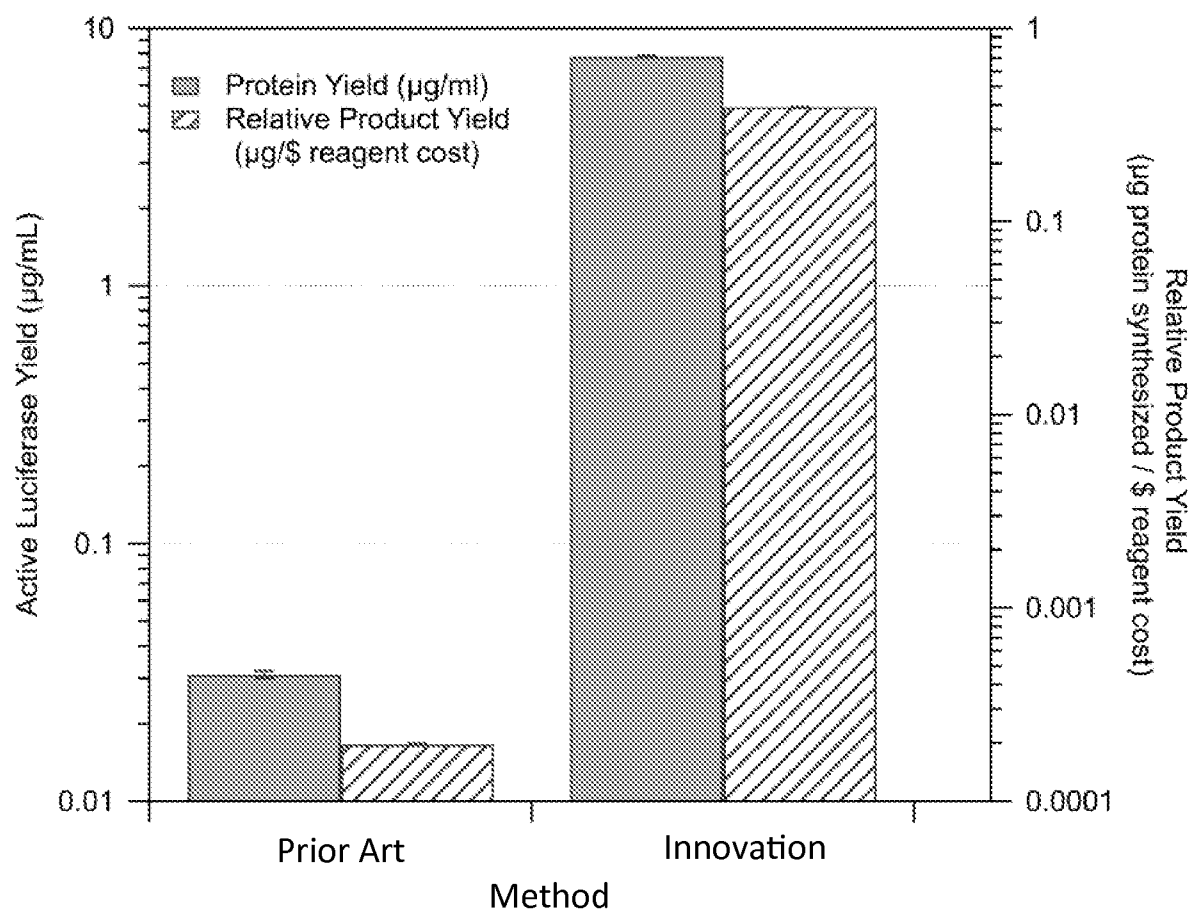


FIG. 8

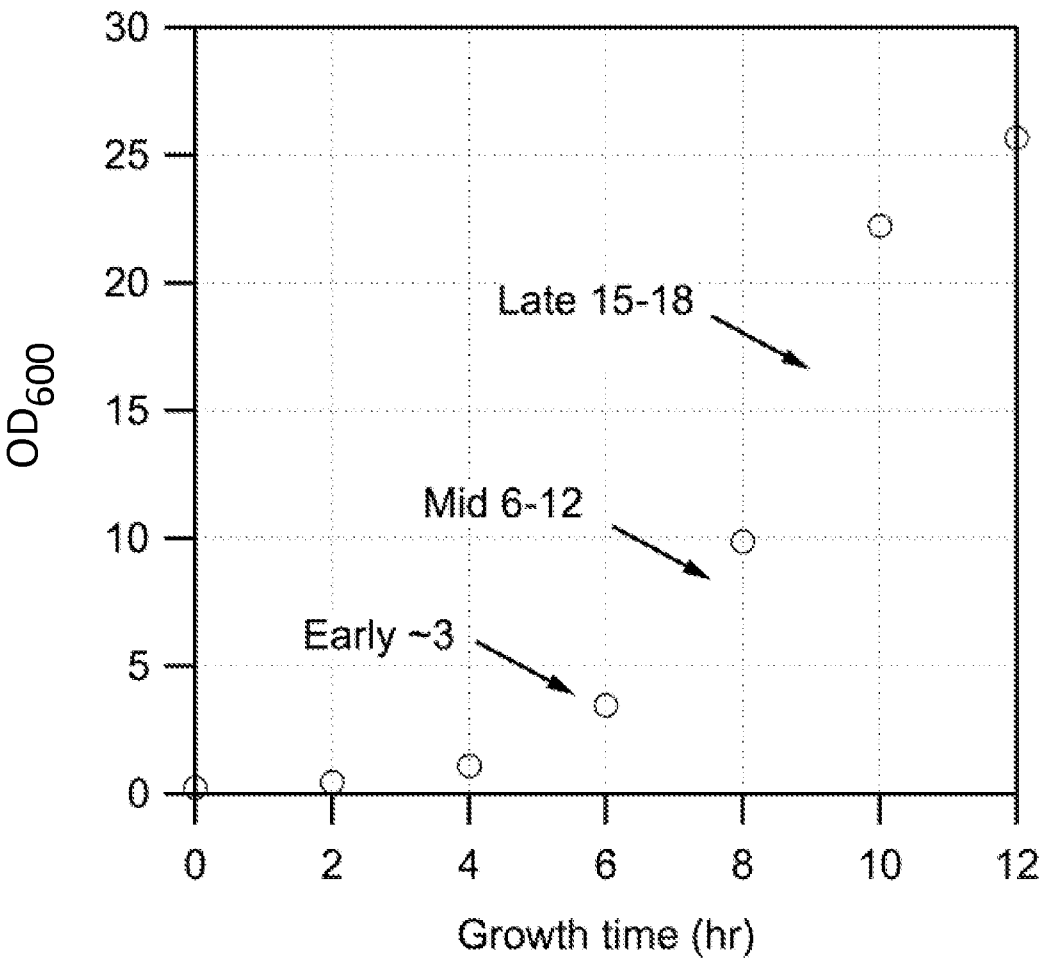


FIG. 9A

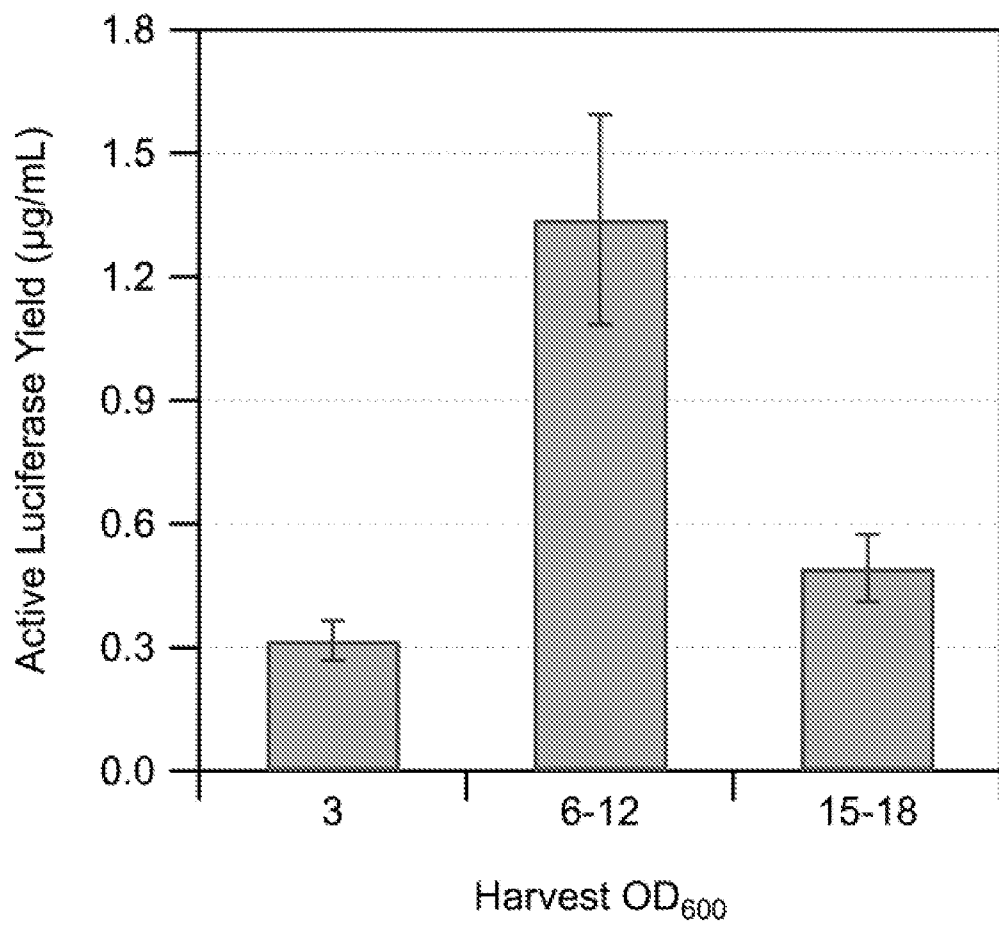


FIG. 9B

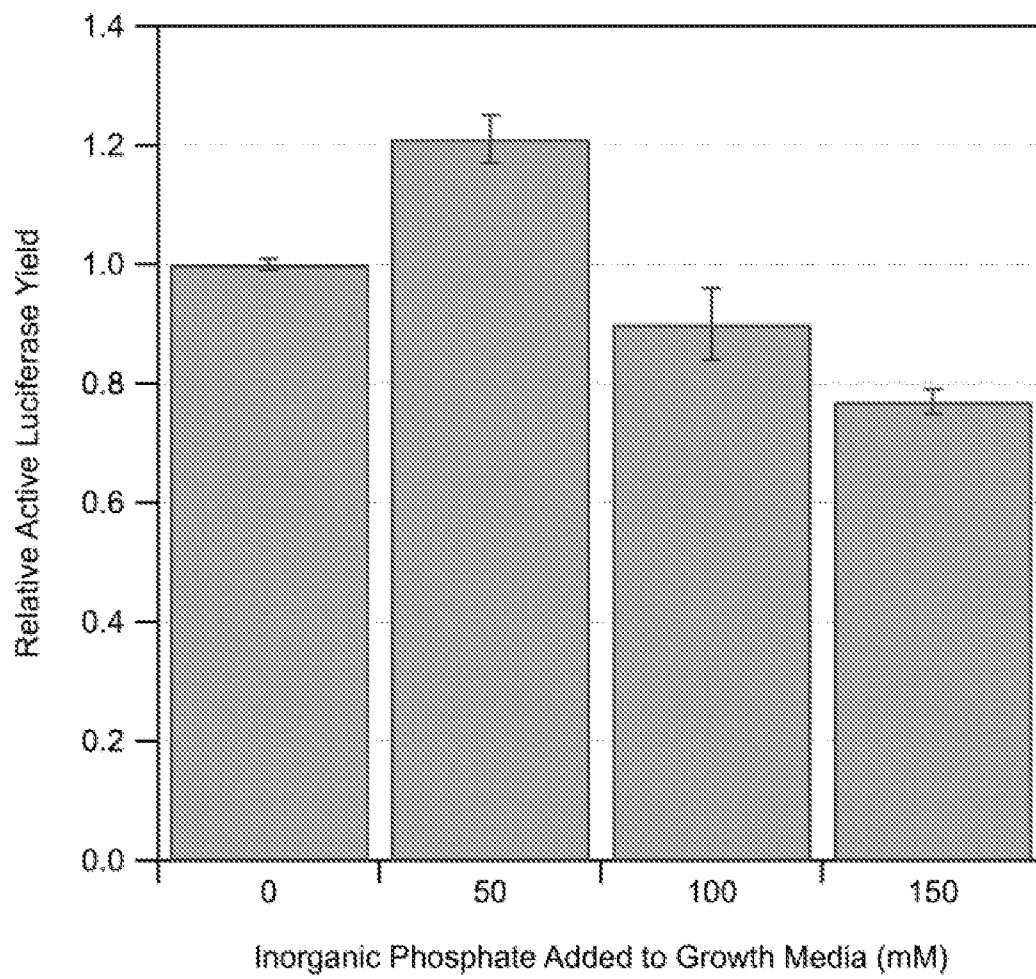


FIG. 10

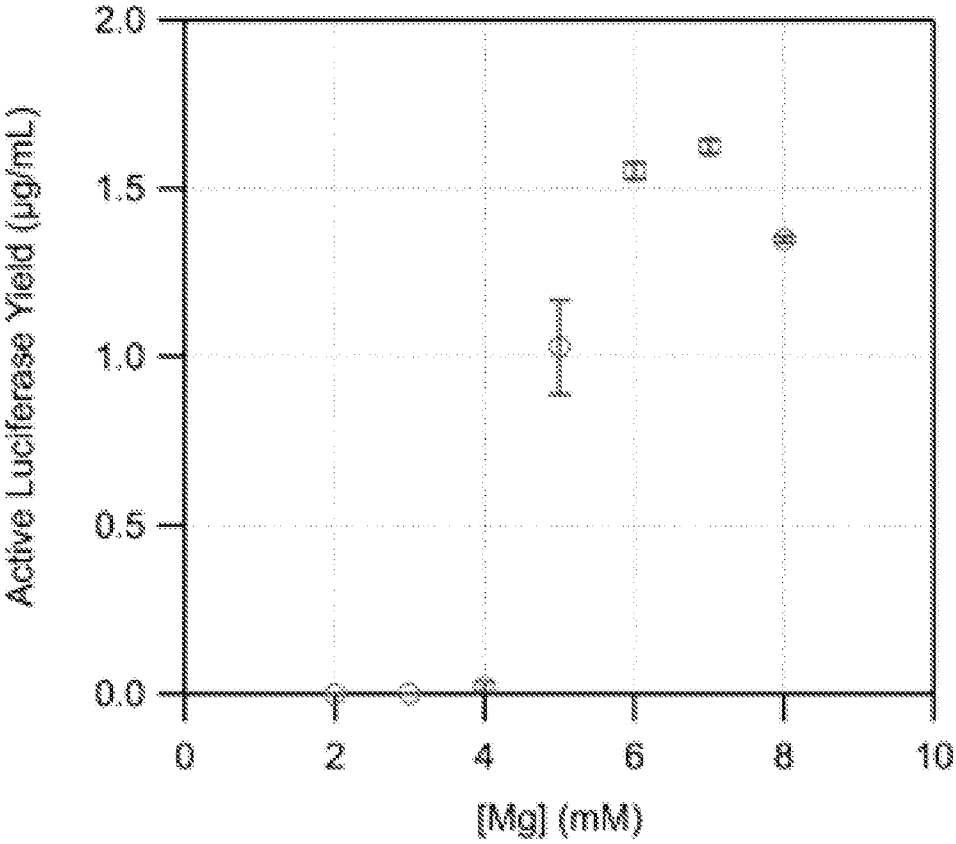


FIG. 11A

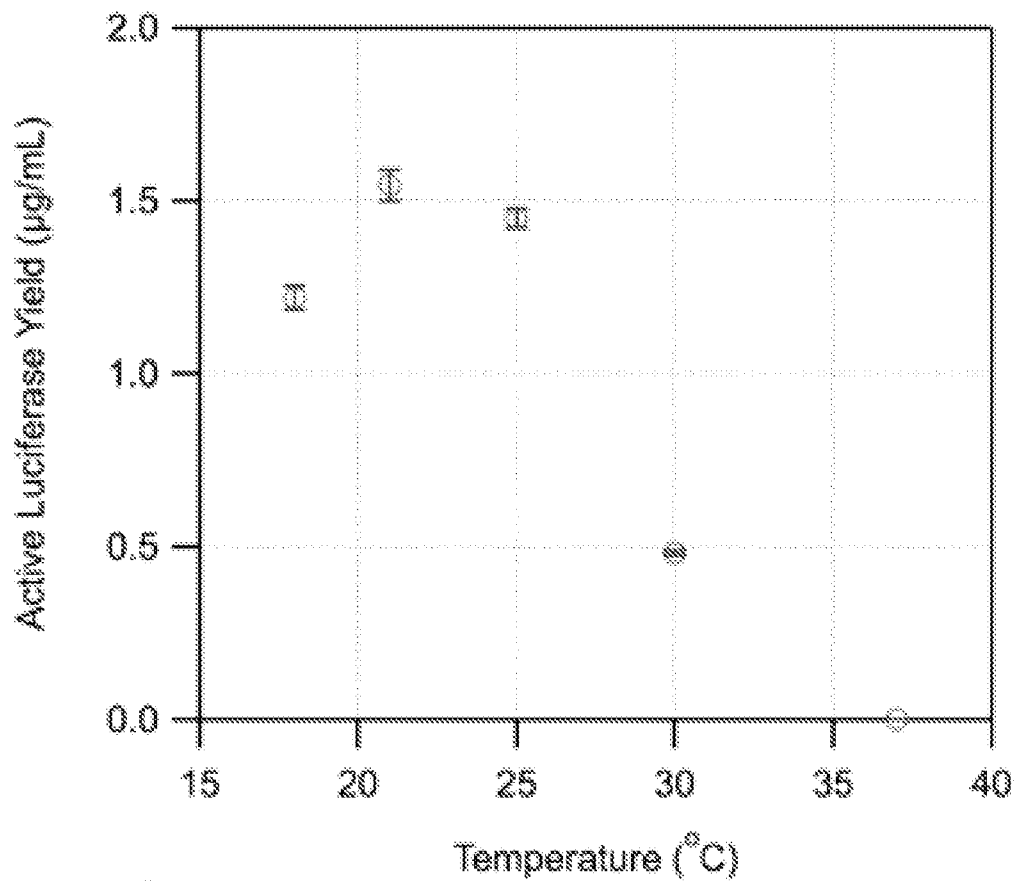


FIG. 11B

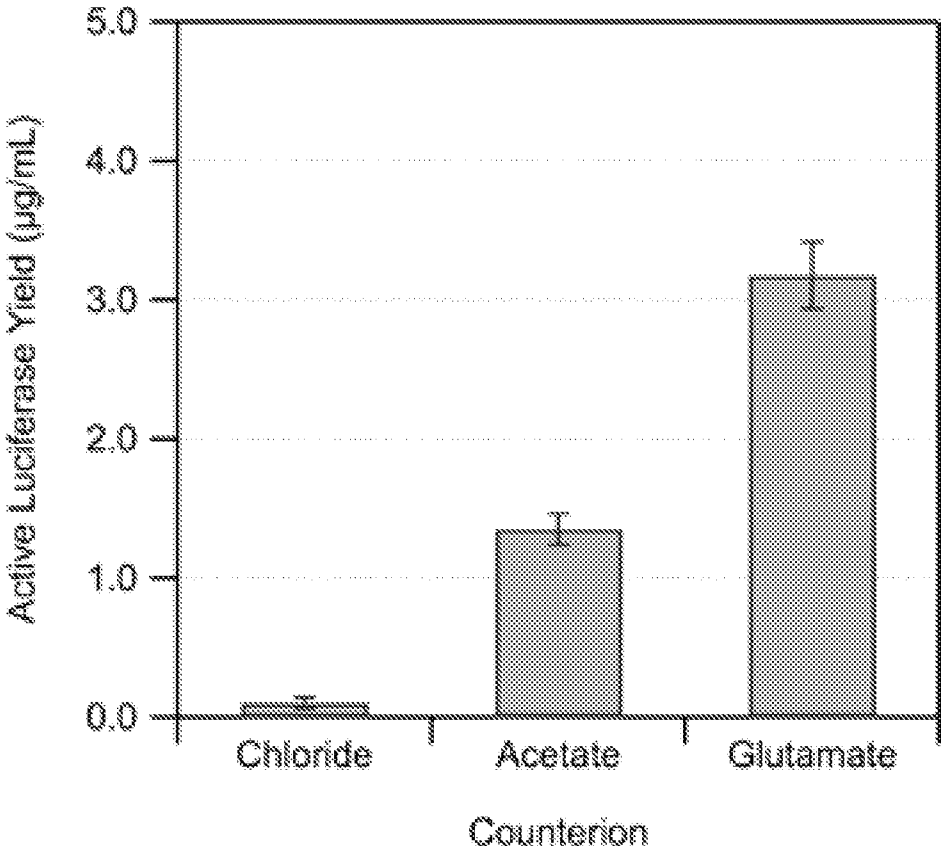


FIG. 11C

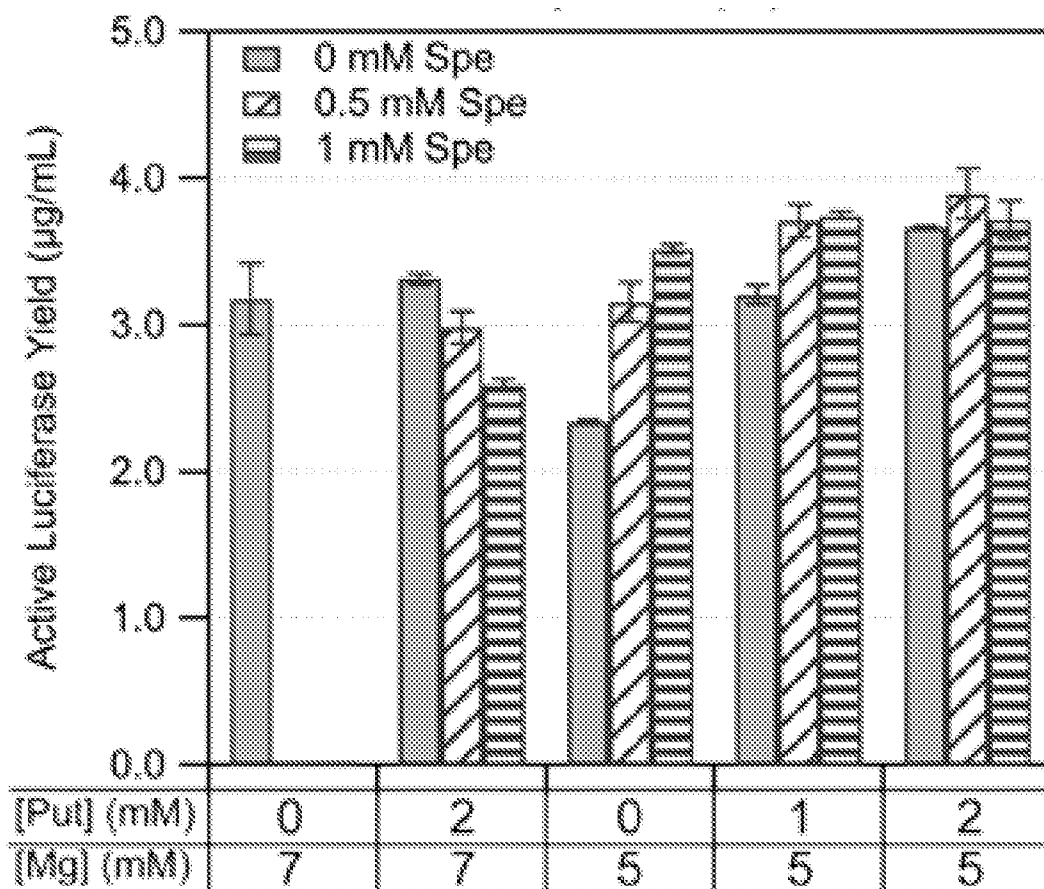


FIG. 11D

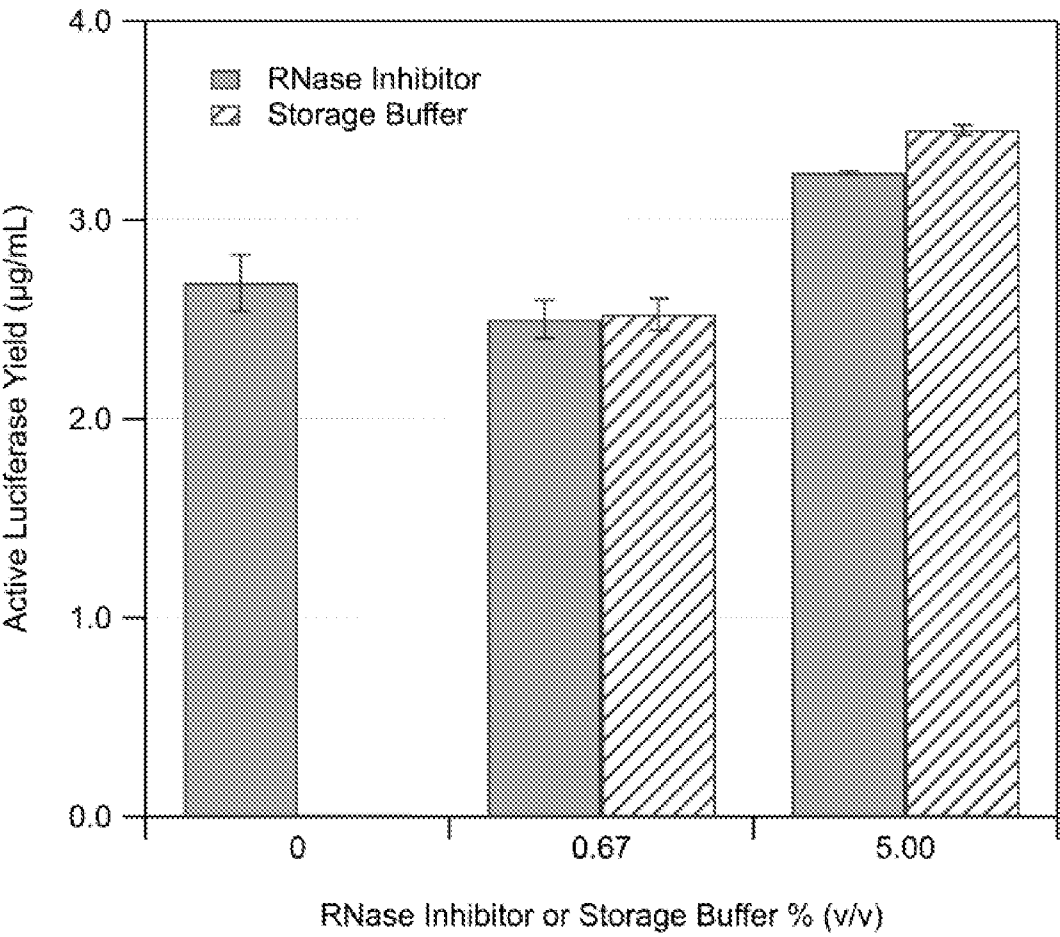


FIG. 11E

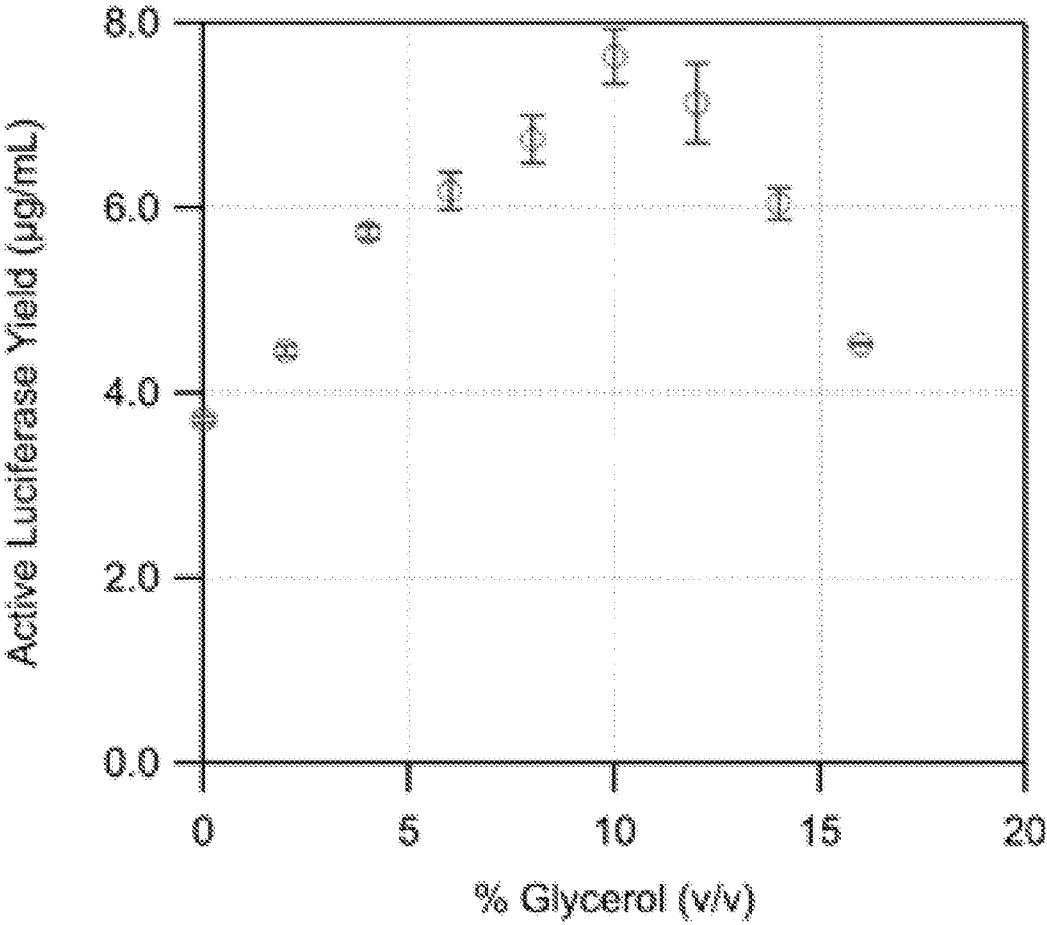


FIG. 11F

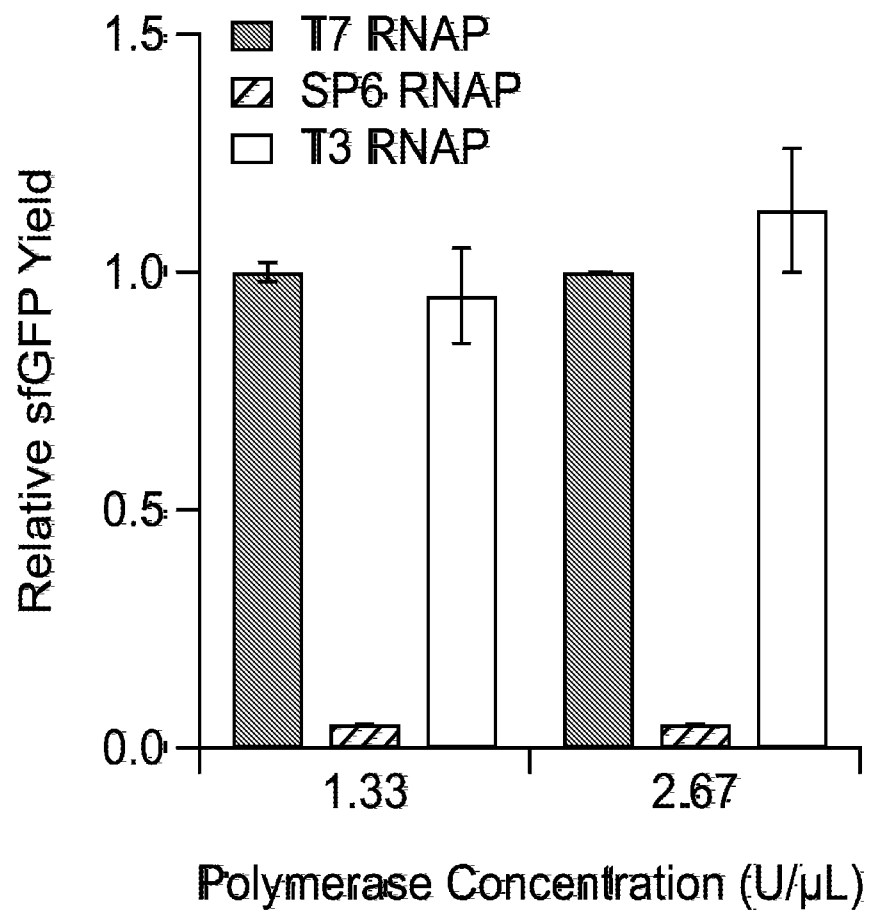


FIG. 12A

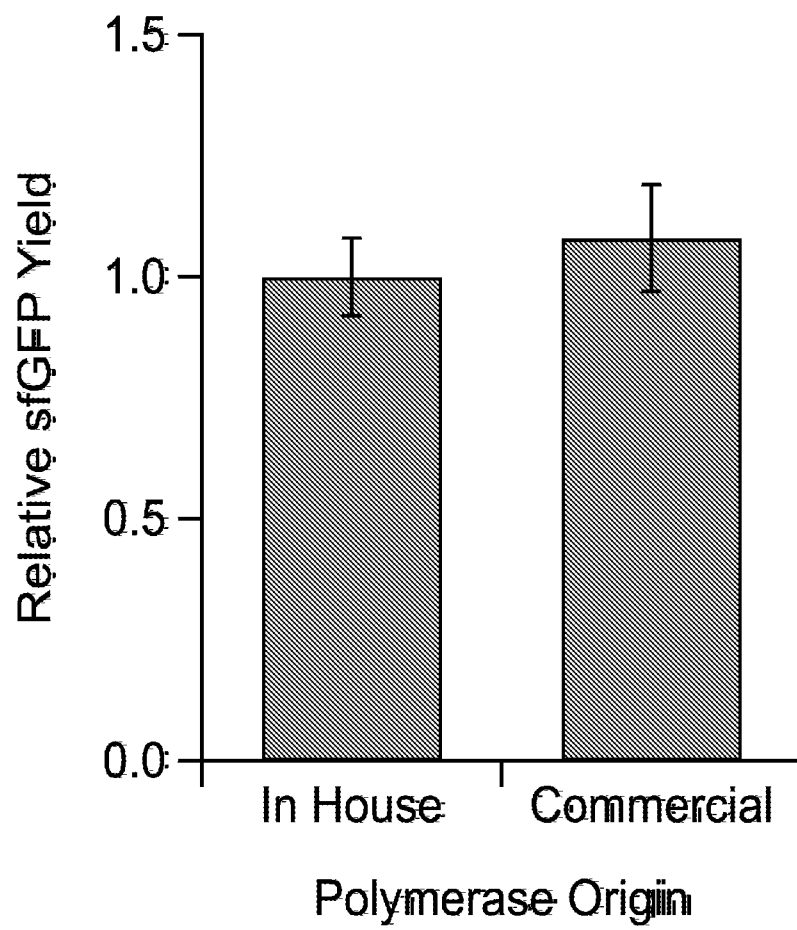


FIG. 12B

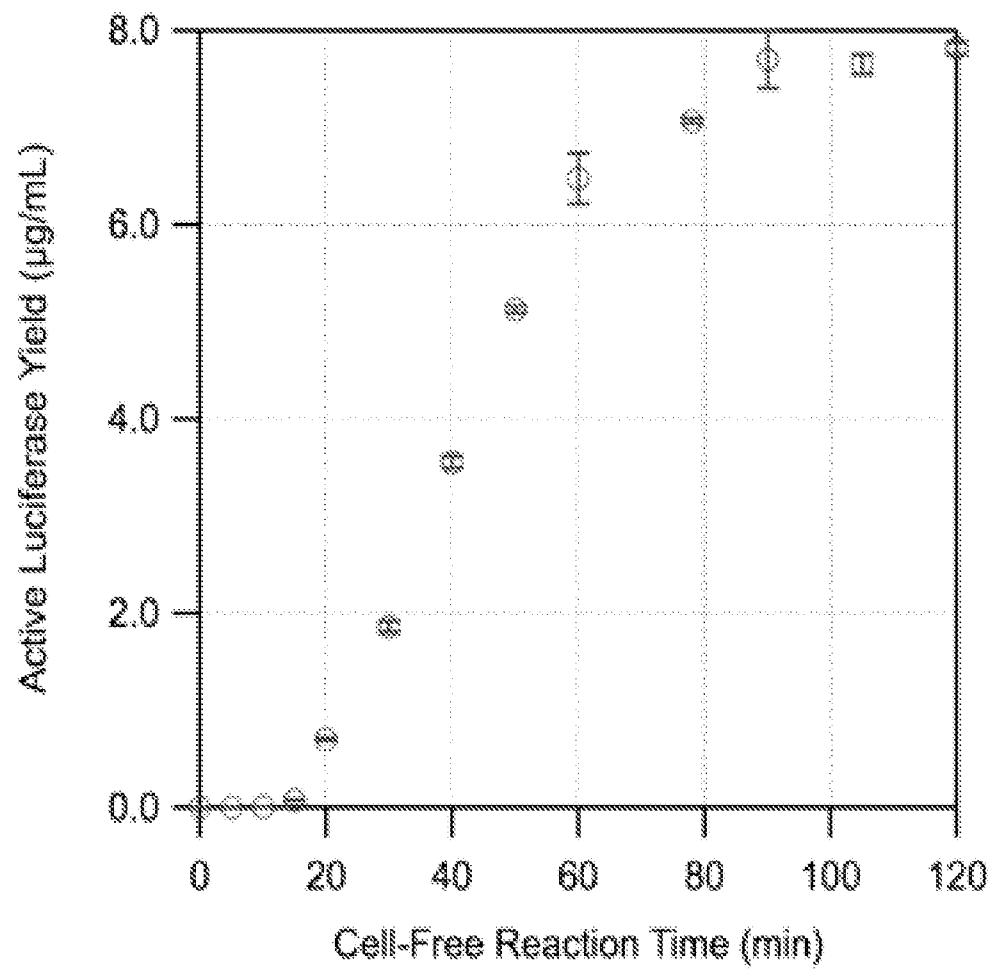


FIG. 13A

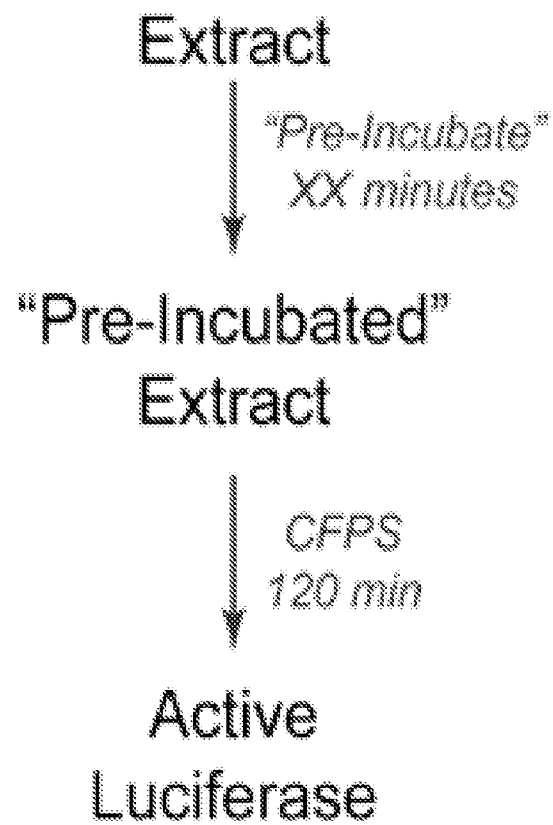


FIG. 13B

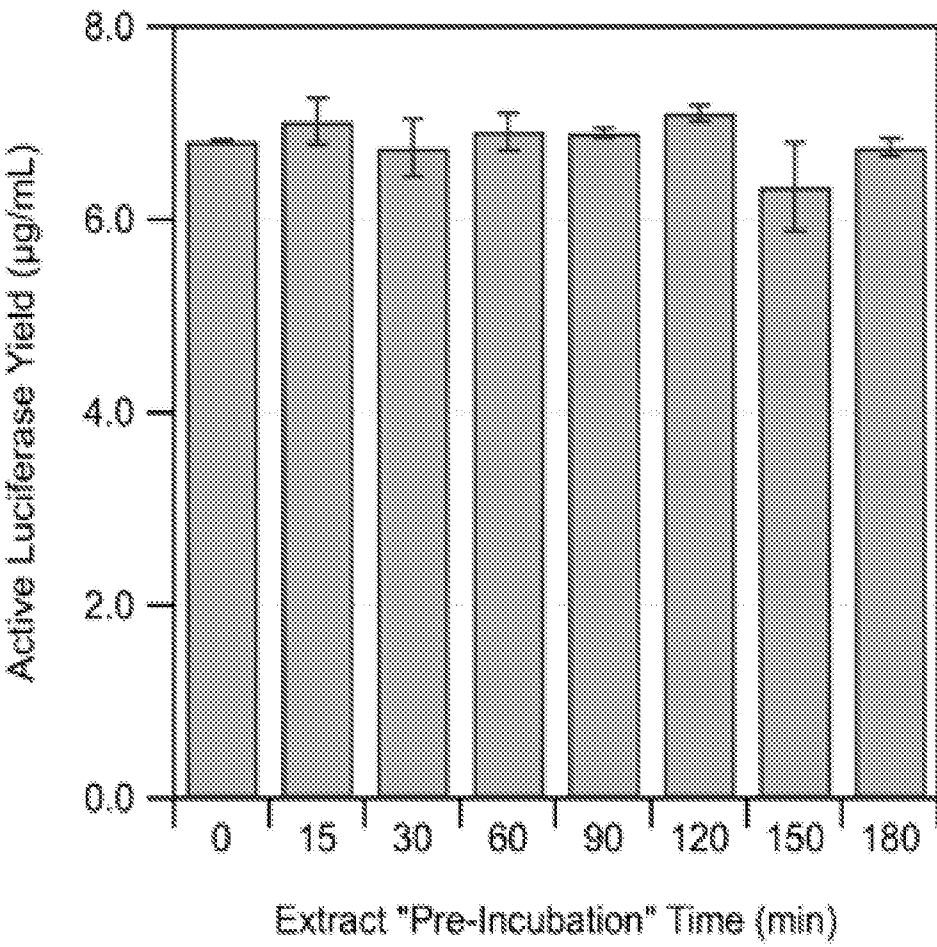


FIG. 13C