



(51) International Patent Classification:

C12N 15/09 (2006.01) C12N 15/33 (2006.01)
C12N 15/87 (2006.01) C07K 14/195 (2006.01)
C07K 14/005 (2006.01) C12N 15/31 (2006.01)

(21) International Application Number:

PCT/US2012/049625

(22) International Filing Date:

3 August 2012 (03.08.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/514,825 3 August 2011 (03.08.2011) US

(71) Applicant (for all designated States except US):

PROSETTA ANTIVIRAL, INC. [US/US]; 670 5th Street, San Francisco, CA 94107 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LINGAPPA, Vishwanath, R. [US/US]; 626 Pacheco St., San Francisco, CA 94116 (US).

(74) Agent: MANN, Jeffry, S.; Morgan, Lewis & Bockius LLP, One Market, Spear Street Tower, San Francisco, CA 94105 (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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: CELL FREE TRANSLATION SYSTEM FOR COMPOUND SCREENING AND RELATED USES

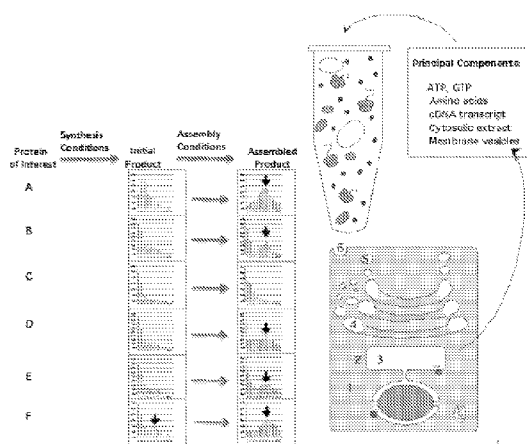


Figure 14

(57) Abstract: The invention provides a cell-free system comprising not more than about 5% wheat germ extract for expressing proteins such as viral proteins and proteins required for viral capsid assembly, and proteins that assemble into multiprotein complexes in a manner analogous to viral capsids, are provided. Further provided are methods for expressing proteins such as viral proteins, proteins required for capsid assembly, and proteins that assemble into multiprotein complexes in a manner analogous to viral capsids using a cell-free system comprising not more than about 5% wheat germ extract. Further provided are methods to assay for compounds that modulate viral protein, viral capsid assembly, and assembly of proteins into multiprotein complexes whose disruption can ameliorate bacterial, parasitic, metabolic, oncologic, immunologic, or CNS disease in a cell-free system comprising not more than about 5% wheat germ extract.

**CELL FREE TRANSLATION SYSTEM FOR COMPOUND SCREENING AND
RELATED USES**

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Application No.
5 61/514,825, filed August 3, 2011, the content of which is expressly incorporated herein by
reference in it entirely for all purposes.

**STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY
SPONSORED RESEARCH AND DEVELOPMENT**

[0002] Not applicable

10 **FIELD OF THE INVENTION**

[0003] The present invention relates to a wheat germ extract for use in compositions for
cell-free translation, methods of expressing proteins using a composition including wheat
germ extract, methods of assembling viral capsids, and methods of detecting compounds that
modulate protein-protein interactions including but not limited to the assembly of viral
15 capsids.

BACKGROUND OF THE INVENTION

[0004] Cell-free translation systems utilize cellular extracts to express proteins of interest.
In this paradigm, an *in vitro* transcribed or isolated RNA template is added to the cellular
extract along with additional components, i.e. amino acids, to allow protein synthesis.

20 [0005] Wheat germ extract is commonly used for cell-free translation reactions, and was
initially described by various investigators (e.g., Roberts, B. E. and Paterson, B. M. (1973,
PNAS 70, p. 2330)). Wheat germ extract is desirable because it supports translation of
prokaryotic, eukaryotic, and viral RNAs. Wheat germ extract has been further shown to be
useful in cell-free systems designed to assemble viral capsids (Lingappa et al. *J Cell Bio* 125:
25 99-111 (1994); Lingappa et al. *J Cell Bio* 136:567-581 (1997); Singh et al. *Virology* 279:257-
270 (2001); Zimmerman et al. *Nature* 415:88-92 (2002); Lingappa and Thielen, *Methods Mol
Biol.* 485:185-95 (2009)). Viral capsids are the protein shell of a virus that protects the viral
genome, and its assembly is a process, catalyzed by host factors, that can be targeted to
develop anti-viral drugs.

DB2/ 22365473.4

[0006] Proper wheat germ extract concentration is required for efficient protein expression. Optimal wheat germ extract concentration for robust protein expression has generally been shown to be around 40 to 50% of the total translation mix (Erikson and Blobel, *Methods Enzymology* 96:38-50 (1983)). Reports have specified that wheat germ extract is optimally
5 used at 20% of the final reaction volume for proper capsid assembly (Lingappa and Thielen, *Methods Mol Biol.* 485:185-95 (2009)).

[0007] It is an important aspect of antiviral drug development processes utilizing cell-free expression systems to show that antiviral modulator analogs are effective in cell-free systems in similar relative proportion to their activity against viruses in cell culture. Currently used
10 cell-free systems do not show useful correlation between the ability of therapeutic agents that disrupt capsid assembly *in vitro* and the ability of these therapeutic agents to prevent virus replication. There is therefore a need to develop a cell-free system having levels and ratios of components that achieve maximum sensitivity to drugs in the cell-free system and then to correlate that efficacy against live viruses in cell culture.

15 BRIEF SUMMARY OF THE INVENTION

[0008] Quite surprisingly, it has now been discovered that cell-free systems in which the concentration of wheat germ extract is maintained at or below about 5%, provide significantly better correlation between drug sensitivity and infectious virus drug sensitivity at 5% wheat germ extract with little apparent correlation at 10% or 20%. The art has to date
20 failed to recognize a connection between wheat germ extract concentration and drug sensitivity, recommending the use of high concentrations of wheat germ extract to maximize protein synthesis. Thus, the literature's teaching that wheat germ extract should be used at higher concentrations (e.g., 20% or higher) does not result in a cell-free system providing optimum sensitivity for drug screening, as observed at 5% wheat germ or lower, as provided
25 by the present invention.

[0009] The present invention generally relates to a cell-free system for expressing proteins using not more than about 5% wheat germ extract. Accordingly, the claimed subject matter provides compositions and methods useful for expressing proteins and for assaying modulators of viral proteins and capsid assembly.

[0010] In an exemplary embodiment, a cell-free system for expressing a protein of interest is provided. The system comprises not more than about 5% wheat germ extract and further comprises other components necessary for expression of the protein of interest. In an exemplary embodiment, the protein of interest is a viral protein. In various embodiments, the viral protein is a viral capsid protein. In a selected embodiment, the protein of interest is a capsid interacting protein. In a selected embodiment, the protein of interest is a host protein which undergoes assembly in a manner analogous to capsid proteins, i.e. most likely due to catalysis of formation of specific multi-protein complexes by other proteins in the cytoplasm.

[0011] In an exemplary embodiment, the protein is a non-viral protein that forms a multiprotein complex. In selected embodiments, the multiprotein complex participates in diseases comprising central nervous system disorders, metabolic disorders, oncologic disorders, or immunologic disorders. In an exemplary embodiment, the protein of interest is a bacterial or parasitic protein.

[0012] The viral protein can be from any virus or family of viruses. In an exemplary embodiment, the viral protein is from a virus which is a member of a viral family selected from the group consisting of *Flaviviridae*, *Togaviridae*, *Bunyaviridae*, *Arenaviridae*, *Filoviridae*, *Poxviridae*, *Orthomyxoviridae*, *Rhabdoviridae*, *Herpesviridae*, *Coronaviridae*, *Paramyxoviridae*, *Hepadnaviridae*, *Bornaviridae*, *Picornaviridae*, *Retroviridae*, *Reoviridae*, *Papillomaviridae*, *Adenoviridae*, *Astroviridae*, *Polyomaviridae*.

[0013] In addition to the not more than about 5% wheat germ extract, an exemplary composition of the invention further includes components necessary or useful to ensure that the expression system expresses the desired protein in the desired amount and/or form. In an exemplary embodiment, the further components of the composition include one or more of a buffer, an amino acid, and a nucleic acid transcript. In various embodiments the composition further comprises one or more of a detectable moiety, ATP, GTP, creatine phosphate, a labeled amino acid, myristoyl CoA lithium salt, an RNase inhibitor, creatine kinase, and a tRNA. In an exemplary embodiment, the labeled amino acid comprises [35S] methionine. In an exemplary embodiment, the nucleic acid transcript is derived from an *in vitro* transcription reaction. In an exemplary embodiment, the nucleic acid transcript encodes a viral protein, e.g., a viral capsid protein. In an exemplary embodiment, the nucleic acid transcript encodes a viral capsid interacting protein. In an embodiment, the buffer comprises a member selected

from potassium acetate, spermine, and dithiothreitol or other reducing agents and a combination thereof.

[0014] In another embodiment, a method of expressing a protein of interest using a cell-free system as described herein is provided. In an exemplary embodiment, the protein of interest is a viral protein, e.g., a viral capsid protein. In an exemplary embodiment, the protein of interest is a viral capsid interacting protein.

[0015] In another embodiment, the invention provides a method of assembling a viral capsid using a cell-free system of the invention.

[0016] In an exemplary embodiment, the invention provides a method assembling non-viral proteins in a multiprotein complex. In selected embodiments, the methods assemble a multiprotein complex that participates in diseases comprising central nervous system disorders, metabolic disorders, oncologic disorders, or immunologic disorders. In an exemplary embodiment, the methods assemble a multiprotein complex of bacterial or parasitic proteins.

[0017] In various embodiments, there is provided a method of testing whether a compound modulates a cellular function. An exemplary method comprises introducing the compound to a cell-free system of the invention, and determining whether the cellular function is modulated.

[0018] The invention also provides a method of testing whether a compound modulates viral capsid assembly due to an effect on host proteins or on the viral capsid proteins themselves. The method includes introducing the compound to a cell-free system of the invention and determining whether the viral capsid (or other protein) assembly is modulated.

[0019] In another embodiment, a kit for cell-free expression of a protein is provided. An exemplary kit comprises a cell-free mixture comprising not more than about 5% wheat germ extract, components necessary for expression of the protein, and instructions sufficient for use of the kit in a cell-free expression experiment.

[0020] In another embodiment, a kit for the assembly of viral capsid is provided. The kit comprises a cell-free mixture comprising not more than about 5% wheat germ extract,

components necessary for expression of proteins required for viral capsid assembly and instructions sufficient for use of the kit in a cell-free expression experiment.

[0021] In another embodiment, a kit for determining whether a compound modulates expression of a protein is provided. The kit comprises a cell-free mixture comprising not more than about 5% wheat germ extract, components necessary for expression of the protein, and instructions sufficient for use of the cell-free mixture to determine whether the compound modulates expression of the protein.

[0022] In another embodiment, a kit for determining whether a compound modulates viral capsid assembly is provided. The kit comprises a cell-free mixture comprising not more than about 5% wheat germ extract, components necessary for expression of proteins required for viral capsid assembly, and instructions sufficient for determining whether the compound modulates viral capsid assembly in the cell-free mixture.

[0023] In another embodiment, a kit for determining whether a compound modulates multiprotein assembly is provided. The kit comprises a cell-free mixture comprising not more than about 5% wheat germ extract, components necessary for expression of non-viral proteins, and instructions sufficient for determining whether the compound modulates multiprotein assembly.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] **Figure 1** illustrates quantitation of protein synthesis at a composition of 5%, 10%, or 20% wheat germ extract at 26 °C by western blotting 2 µL of translation extract in triplicates, and plotted as arbitrary density units upon quantitation of band intensity using an alkaline phosphatase conjugated secondary antibody-based colorimetric method. Cell-free translation was performed in the WG cell-free protein synthesizing system, programmed with mRNA encoding Hepatitis C virus core polypeptide, at a composition of 5%, 10% or 20% wheat germ extract, in the absence (DMSO control), or presence of compounds whose efficacy had been assessed against infectious virus in cell culture and found to represent a gradation of potencies from weak to strong. A quantitation of protein synthesis at each percentage of wheat germ by western blotting of 2µL of translation extract (triplicates), plotted as arbitrary density units upon quantitation of band intensity using an alkaline phosphatase conjugated secondary antibody-based colorimetric method. Conclusion: the

level of protein synthesis is roughly the same at each percentage wheat germ, but slightly lower at 5% WG than either 10% or 20%.

[0025] **Figure 2** illustrates the sensitivity and correlation between compound potency against Hepatitis C (HCV) synthesized at a composition of 5%, 10%, or 20% wheat germ extract in the absence (DMSO control) or presence of compounds. (A) Compound efficacy ranges from extremely weak (compound 138-85) to the most potent (compound 138-136). Conclusion: Compound potency against infectious virus in cell culture ranged from weak ($EC_{50} > 50\mu M$) to strong ($EC_{50} = 100nM$). (B) Correlation of drug potency on plate screen at each wheat germ percentage to drug potency against infectious virus. Conclusion: Drug sensitivity displays greatest correlation to infectious virus drug sensitivity at 5% wheat germ extract with little apparent correlation at 10 or 20%. Thus, the literature's teaching that wheat germ extract should be used at 20% or higher misses the point of optimum sensitivity for drug screening which is observed at 5% wheat germ or lower, as provided by the present invention.

[0026] **Figure 3** illustrates the sensitivity and correlation between compound potency against influenza (FLUV) synthesized at a composition of 5%, 10%, or 20% wheat germ extract in the absence (DMSO control) or presence of compounds. (A) Compound efficacy ranges from extremely weak (compound 138-165) to the most potent (compound 150-120). (B) Correlation of drug potency using plate screen at each wheat germ percentage to drug potency against FLUV. As can be seen for FLUV, as for HCV in the previous figure, the sensitivity and correlation between compound potency against infectious virus and compound potency on the plate screen is greatest at 5% WG and is lost at 10% WG concentration, and reappears at 20% to some extent but not in full correlation with potency against infectious virus. Thus, the points made for HCV in Figure 1, apply to significant extent also to FLUV.

[0027] **Figure 4** illustrates the sensitivity and correlation between compound potency against Venezuelan Equine Encephalitis (VEEV) synthesized at a composition of 5%, 10%, or 20% wheat germ extract in the absence (DMSO control) or presence of compounds. (A) Compound efficacy ranges from extremely weak (compound 158-80) to the most potent (compound 150-133). (B) Correlation of drug potency using plate screen at each wheat germ percentage to drug potency against VEEV. As can be seen for VEEV, as for HCV in the previous figure, the sensitivity and correlation between compound potency against infectious virus and compound potency on the plate screen is greatest at 5% WG and is dampened at

10% and largely lost at 20% WG concentrations. Thus, the points made for HCV in Figure 1, apply also to VEEV.

[0028] **Figure 5** illustrates the sensitivity and correlation between compound potency against Human Immunodeficiency Virus (HIV) synthesized at a composition of 5%, 10%, or 20% wheat germ extract in the absence (DMSO control) or presence of compounds. (A) Compound efficacy ranges from extremely weak (compound 6027) to the most potent (compound 6051). As can be seen for HIV, as for the previous three viral families represented in Figures 1-4, the sensitivity of the plate screen is greatest at 5% WG, at which concentration the greatest correlation to potency against infectious virus is seen. This sensitivity and correlation is dampened or lost at higher WG concentrations. Thus, in four of four cases, as demonstrated in figures 1-4, the cell-free system displays greatest drug sensitivity and best correlation to potency against infectious virus, at low WG concentrations.

[0029] **Figure 6** illustrates the EC_{50} observed in the HCV capsid protein-programmed cell-free system. (A) In order to establish the fidelity of the cell-free translation-based anti-capsid assembly drug screen, the EC_{50} observed in the HCV capsid protein-programmed cell-free system was plotted versus the EC_{50} against infectious HCV in cell culture. If activity in the cell-free system correlates with activity against infectious virus, the data points from structure-activity relationship optimization should project into the right upper quadrant, as observed. Furthermore, analogs that are more potent on other viral families are indicated in blue, off of the high correlation line. (B) Zoom of the most potent analogs from part (A) reveals more clearly the projection of analogs into the right upper quadrant.

[0030] **Figure 7** illustrates the EC_{50} observed in the HIV capsid protein-programmed cell-free system. Correlation of efficacy of analogs against HIV capsid assembly in the cell-free system (x-axis) versus against infectious virus in cell culture (y-axis) as was shown for HCV in figure 6. The parent compound is shown in red, with analogs depicted in purple. Structure-activity relationships are indicated by dotted lines connecting the related compounds that comprise a given sub pharmacophore.

[0031] **Figure 8** illustrates the EC_{50} observed in the FLUV capsid protein-programmed cell-free system. Correlation of efficacy of analogs against FLUV capsid assembly in the cell-free system (x-axis) versus against infectious virus in cell culture (y-axis) as was shown for HCV in figure 6 and HIV in figure 7. The parent compound is shown in red. Structure-

activity relationship (SAR) is indicated by dotted lines connecting the related compounds that comprise a given sub-pharmacophore. As for HIV and FLUV, SAR project towards the right upper quadrant of the graph, as expected for good cell-free to authentic virus anti-viral compound correlation.

5 **[0032]** **Figure 9** illustrates the uniformity of 5 % wheat germ preparations among source and preparation conditions. Samples of 13 wheat germ extract preps were diluted 50 μ L to 1 mL and 20 μ L was taken and mixed with 20 μ L of SDS gel sample buffer with 10% β ME and heated at 100°C/5 min. 10 μ L was loaded on SDS PAGE. The gel was fixed in 10% acetic acid and 50% methanol when the dye front reached the bottom and then stained with
10 0.1% Coomassie brilliant blue and destained to completion.

[0033] **Figure 10** illustrates the quantitation of scanned lanes from Figure 9 plotted as arbitrary integrated density units. This analysis quantifies the compositional similarity of different wheat germ preps and reveals the extracts to be highly reproducible in protein content as measured by Coomassie stainable protein.

15 **[0034]** **Figure 11** illustrates the A260 absorbance spectrum of 5% wheat germ preps. The absorbance spectrum scan reveals a major A260 peak in the 5% wheat germ extract. This peak is used to define the range of absorbance values associated with each of the 13 5% wheat germ extract preps of Figure 9.

[0035] **Figure 12** illustrates the A280 absorbance spectrum of 13 5% wheat germ preps of
20 Figure 9. The absorbance spectrum scan reveals no major A280 peaks in the 5% wheat germ extracts, due largely to the enormity of the A260 absorbance spectrum.

[0036] **Figure 13** illustrates the A260/A280 absorbance spectrum ratio. Because of the uniformity of A260 nm and A280 nm values from various wheat germ extract preparations, irrespective of source and a range of buffer conditions, the A260/A280 absorbance ratio was
25 calculated as another measure of preparation uniformity.

[0037] **Figure 14** illustrates cell-free systems for protein biogenesis applied to drug discovery for CNS disease. The biogenesis of six proteins (A-F) implicated in various CNS disorders were explored using the 5% wheat germ cell-free protein-synthesizing systems. The sucrose gradient profiles of initial and final products (to the right in blue), synthesis and
30 assembly were dissociated in all cases. In all except C, conditions were established to

successfully reconstitute formation of high molecular weight structures. For several putative CNS disorders, small molecule inhibitors were identified and validated in cells.

[0038] **Figure 15** illustrates sucrose gradient analysis of newly synthesized carbonic anhydrase (panels above) and alpha globin (panels below), two proteins whose mature oligomeric state does not involve large high molecular weight structures of >20S. Left panel is of gradient profile quantified from SDS PAGE and autoradiography after ³⁵S labeled methionine synthesis at 26°C/1h, followed by treatment with puromycin to terminate protein synthesis and release chains. Right panel is of gradient profile quantified from SDS PAGE and autoradiography after subsequent incubation at 34°C/2h. Failure of conversion into high molecular weight structures of these newly synthesized proteins whose mature forms are limited to low molecular weight oligomeric forms suggests that assembly as observed for viral capsid proteins or known oligomeric non-viral proteins is not a form of non-specific aggregation, but rather represents substrate-selective assembly. Thus, this figure suggests that the present invention is selective for a subset of biologically relevant and medically important proteins, but not to the many proteins that do not form large multiprotein complexes upon cell-free translation and incubation under assembly conditions.

DETAILED DESCRIPTION OF THE INVENTION

Introduction

[0039] In various embodiments, the invention provides a cell-free system comprising not more than about 5% wheat germ extract and methods for expressing viral proteins using these systems. The cell-free systems of the invention are useful for assembling viral capsids. Also provided are methods for testing whether a compound modulates cellular function using a cell-free system of the invention.

[0040] Cell-free systems are generally useful for capsid assembly in that a variety of manipulations are possible. For example, reactions within the system can be separated into a protein synthesis phase and an assembly phase. This slows down the naturally occurring capsid assembly phase and allows manipulations to be performed to determine optimal modulation of viral protein synthesis and capsid assembly. One advantage of slowing down capsid assembly in a cell-free system is this allows for a correlation between the efficacy of known modulators of live viruses in cell culture and the efficacy of the same modulators in

the cell-free system. Another advantage is that transient targets (e.g. host proteins that work catalytically to promote capsid formation) can be identified.

[0041] The present invention surprisingly demonstrates that those concentrations of wheat germ extraction, which are significantly lower than those long accepted in the art as standard and necessary concentration are very efficacious. It has long been accepted in the art that wheat germ extract should be used at concentrations above 20% in cell-free expression systems (See e.g., Erikson and Blobel, *Methods Enz* 96:38-50 (1983); and Lingappa and Thielen, *Methods Mol Biol.* 485:185-95 (2009)). According to the compositions and methods of the present invention, capsid assembly is as robust at 5% wheat germ extract as it is at 20%. However, it is demonstrated herein that little or no correlation exists between antiviral efficacy against live viruses in cell culture and the ability to inhibit capsid assembly in cell-free systems using 20% wheat germ extract. What one wishes to achieve is the optimal balance between the level of synthesis and the amount of factors. Importantly, it is shown herein that efficacy of antiviral compounds in cell-free systems is most closely correlated to virus sensitivity in cell culture using wheat germ extract concentrations of 5%. At greater than 5%, drug sensitivity is diminished.

Definitions

[0042] Unless otherwise noted, the technical terms used herein are according to conventional usage as understood by persons skilled in the art. Definitions of common terms in molecular biology may be found in standard texts (e.g. Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19854287-9); Kendrew et al. (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd, 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8)).

[0043] The phrase “cell-free system,” as used herein, refers to any type of system capable of synthesizing proteins *in vitro* in the absence of viable cells. An exemplary system is a cell-free protein synthesis system derived from wheat germ extract.

[0044] The term “expressing” and “expression,” as used herein, refer to the production of a protein, peptide, or nucleotide sequence, and include transcription into an RNA product, post-

transcriptional modification and/or translation into a protein product or polypeptide from a DNA encoding that product, as well as possible post-translational modifications.

[0045] The terms “polypeptide” or “peptide” or “protein” are used interchangeably herein, to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer. Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization (*see, e.g., Alberts et al., Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980)). “Primary structure” refers to the amino acid sequence of a particular peptide. “Secondary structure” refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains, e.g., enzymatic domains, extracellular domains, transmembrane domains, pore domains, and cytoplasmic tail domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 15 to 350 amino acids long. Exemplary domains include domains with enzymatic or other functional activity. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. “Tertiary structure” refers to the complete three dimensional structure of a polypeptide monomer. “Quaternary structure” refers to the three dimensional structure formed by the noncovalent association of independent tertiary units.

[0046] A “protein of interest,” “desired polypeptide,” “desired protein,” or “target protein,” as used herein, are interchangeable and refer to a whole protein molecule, including but not limited to, viral capsid proteins, or a portion thereof, i.e., cytoplasmic domain or other domain of a protein.

[0047] The term “virus” or “viral,” as used herein, refers to minute infectious agents, which, with certain exceptions, are not observable by light microscopy, lack independent metabolism, and are able to replicate only within a living host cell. Some exceptions include, but are not limited to, the cell-free translation system described herein. These individual particles (i.e., for example, virions) typically comprise nucleic acid and a protein shell or coat; some virions also have a lipid containing membrane. The term virus encompasses all types of viruses, including animal, plant, phage, and other viruses.

[0048] The term “viral capsid” or “capsid,” as used herein, refers to the protein coat that surrounds the viral nucleic acid. Viral capsids have interior surfaces and exterior surfaces. The interior surface of a viral capsid is the surface that is normally exposed to the viral nucleic acid. The exterior surface of a viral capsid is the surface that is generally exposed to the environment. The phrase “viral capsid assembly” refers to the process of arranging viral capsid proteins in a manner sufficient to generate a viral capsid.

[0049] The term “capsid interacting protein,” as used herein, refers to protein that interacts with a viral capsid either during or after its assembly. The capsid interacting protein may be endogenous to a virus, may be exogenously added, or present with the cell-free extract.

Capsid interacting proteins can include, but are not limited to, capsid chaperones and proteins that have catalytic actions favoring capsid formation.

[0050] The term “components,” as used herein, refers to constituents of the cell-free system necessary to incorporate a naturally occurring or non-natural amino acid into a growing polypeptide chain. For example, components can include, but are not limited to, buffers, amino acids, nucleic acid transcripts, ATP, GTP, creatine phosphate, labeled amino acids, myristoyl CoA lithium salts, RNase inhibitors, creatine kinases, and tRNAs. Components are described in greater detail herein.

[0051] The phrase “detectable moiety” or “conjugate” refers to any atom, molecule or a portion thereof, the presence, absence or level of which is directly or indirectly monitorable.

A variety of detectable moieties are well known to those skilled in the art, and can be any material detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Such detectable labels can include, but are not limited to, magnetic beads, fluorescent dyes, radiolabels, enzymes, and colorimetric labels such as colloidal gold or colored glass or plastic beads.

[0052] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code. Amino acids can be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, can be referred to by their commonly accepted single-letter codes. Amino acid substitutions, deletions or additions to individual or

a small percentage of amino acids in the encoded sequence is a conservatively modified variant, where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g., Creighton, Proteins (1984)*).

[0053] The term “nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions can be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide. A particular nucleotide sequence also implicitly encompasses “splice variants,” which as the name suggests, are products of alternative splicing of a gene. After transcription, an initial nucleic acid transcript can be spliced such that different (alternate) nucleic acid splice products encode different polypeptides. Mechanisms for the production of splice variants vary, but include alternate splicing of exons. Alternate polypeptides derived from the same nucleic acid by read-through transcription are also encompassed by this definition. Any products of a splicing reaction, including recombinant forms of the splice products, are included in this definition.

[0054] The phrase “*in vitro* transcription reaction,” as used herein, refers to a transcription reaction that takes place in a cell-free environment using largely purified components, for example, purified DNA template and purified DNA-dependent RNA polymerase.

[0055] The term “modulates” or “modulated,” as used herein, means to interact with a target either directly or indirectly so as to alter the activity of the target, including, for example, to inhibit the activity of the target, or to limit or reduce the activity of the target. Accordingly, the phrase “modulates a cellular function” means to alter the function of a way, which can include, but is not limited to, inhibition of protein synthesis or inhibition of protein assembly into molecular structures such as viral capsids.

[0056] The term “test compound” or “compound” or “drug candidate” or “modulator” or grammatical equivalents, as used herein, describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc., to be tested for the capacity to directly or indirectly modulation tumor cell proliferation. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity. Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis. Compounds can be inhibitors, activators, or modulators of, for example, nucleic acid and polypeptide sequences, and are used to refer to activating, inhibitory, or modulating molecules identified using *in vitro* and *in vivo* assays of the nucleic acid and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate activity or expression of, for example, nucleic acids or polypeptides derived from the cell-free system described herein, e.g., antagonists. Activators are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate, for example, nucleic acids or polypeptides derived from the cell-free system described herein, e.g., agonists. Inhibitors, activators, or modulators also include genetically modified versions of the proteins derived from the cell-free system, e.g., versions with altered activity, as well as naturally occurring

and synthetic ligands, substrates, antagonists, agonists, antibodies, peptides, cyclic peptides, nucleic acids, antisense molecules, ribozymes, or small chemical molecules, for example.

[0057] The phrase “small organic molecule” refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of from about 50 to about 2500 daltons, preferably less than about 2000 daltons, preferably between about 100 and about 1000 daltons, more preferably between about 200 and about 500 daltons.

[0058] The term “biopharmaceutical” or “biopharmaceutical compound,” as used herein, refers to any compound such as a protein, peptide, polypeptide, antibody, or the like, which can be expressed endogenously in a biological system under genetic control and which confers biological activity toward pharmaceutical or therapeutic use. The biopharmaceutical or biopharmaceutical compound can be constitutively or inducibly expressed. The biological system can be an *in vivo* biological system and/or an *in vitro* biological system.

Cell-free systems

[0059] Cell-free extracts of wheat germ provide cytosolic factors critical for translation, and support the *in vitro* translation of a wide variety of mRNAs into protein. The translation mechanism is sufficiently conserved so that a cell-free system can translate both prokaryotic and eukaryotic mRNAs with high efficiency. The optimal concentration of wheat germ extract can be determined for each RNA sequence to be translated.

[0060] In an exemplary embodiment, the invention provides a cell-free system (composition) for expressing a protein of interest, including viral proteins, comprising 5% wheat germ extract. In various embodiments, the invention provides a cell-free system for viral capsid assembly comprising 5% wheat germ extract. However, one skilled in the art would recognize that wheat germ extract amounts of less than about 5%, such as about 4%, about 3%, about 2%, or about 1%, can be used to express viral proteins. The composition of the invention can also include higher amounts of wheat germ extract, such as about 6%, about 7%, about 8%, about 9%, about 10%, etc., such that the total concentration of wheat germ extract is less than about 20% vol/vol of the composition.

[0061] The method for producing the wheat germ extract used in the present invention is not particularly limited, as long as the wheat germ extract enables the cell-free protein synthesis when template nucleic acid, amino acids, and an energy source and the like are

supplied. In an exemplary process, an extract obtained by separating wheat germ from albumen in a wheat seed, followed by extraction and purification from the wheat germ, is used as a wheat germ extract. One skilled in the art recognizes that wheat germ extract can be prepared from a wheat seed (See, e.g., Erikson and Blobel, *Methods Enz* 96:38-50 (1983);
5 *PNAS* 97:559-564 (2000)). One skilled in the art also recognizes that wheat germ can be commercially purchased (e.g. Promega, American Biosciences, etc).

[0062] Cell-free translation systems using wheat germ extract can utilize a wide variety of components in addition to the wheat germ extract for efficient translation. Described herein are basic components useful for efficient translation of viral proteins using wheat germ
10 extract. However, one skilled in the art recognizes that additional components might be useful for translation and/or capsid assembly depending on the desired properties of the proteins and/or capsids produced the desired production conditions and other variables. The choice of the proper components for a cell-free system of the invention is well within the capabilities of those of skill in the art. The use of wheat germ extract for capsid assembly is
15 also known in the art, and is described in more detail, for example, in Lingappa and Thielen, *Methods Mol Biol.* 485:185-95 (2009), and US Pat. No. 7,638,269.

[0063] In an exemplary embodiment, the composition of the invention further includes a buffer. Cell-free translation systems generally require an appropriate compensating buffer. Potassium and magnesium concentrations of the wheat germ translation system can have
20 dramatic effects on the efficiency of translation, and the compensating buffer is used to adjust the ion concentration of the total translation reaction to an optimum that can be determined for each mRNA being translated. Buffers can include further components for efficient protein expression, which include, but are not limited to, potassium acetate, amines (e.g., spermine), and sulfur compounds (e.g., dithiothreitol).

[0064] The composition of the invention also optionally includes a nucleic acid encoding a protein or a portion thereof. In an exemplary embodiment, the translation mixture contains transcript nucleic acid or a fragment thereof that encodes one or more viral protein. In this
25 embodiment, the cell-free translation system involves two linked reactions: *in vitro* transcription and cell-free translation. RNA can be obtained by any method known in the art including, but not limited to, isolating mRNA or by making *in vitro* RNA transcripts from
30 DNA cloned into a vector containing an RNA polymerase promoter. RNA molecules can also be generated in the same reaction vessel used for the translation reaction. In an

exemplary embodiment, the RNA is generated *in situ* by addition of, for example, SP6 polymerase to the reaction mixture along with the viral protein coding region or cDNA.

[0065] In an exemplary embodiment, a sample containing a virus of interest or a bodily fluid of an individual infected with a virus of interest, or infected cells from an individual, is used a source of viral nucleic acid encoding the protein for the virus. This can then be engineered behind an appropriate promoter (e.g. for SP6 polymerase), amplified by PCR and purified for transcription-linked translation. The fluid may be any bodily fluid including, without limitation, blood, serum, plasma, lymphatic fluid, urine, sputum, cerebrospinal fluid, and the like.

[0066] The endogenous mRNA present in the wheat germ extract can compete with the exogenous RNA for ribosomes and factors required for translation. It is therefore optionally advantageous to reduce the concentration of endogenous RNA by treating the prepared extract with nuclease. Such nucleases are well known in the art, and can include, but are not limited to, micrococcal nuclease from *Staphylococcus aureus*.

[0067] Methods known in the art are used to maintain energy levels sufficient to maintain protein synthesis. In an exemplary embodiment, additional nucleotide energy sources are added during the reaction. In various embodiments, energy is maintained by addition of an energy source such as creatine phosphate/creatine phosphokinase. In an exemplary embodiment, ATP and GTP concentrations present in a standard translation mixture known in the art are sufficient to support both protein synthesis and capsid formation.

[0068] In an exemplary embodiment, the virus has a myristolated intermediary. For such viruses, one can add sufficient myristoyl coenzyme A (MCoA) with or without acceptable salts to the system to enable capsid assembly. The concentration required may vary according to the particular experimental conditions, and can therefore be determined empirically.

[0069] In various embodiments, the cell-free translation systems of the invention comprise further components including, but not limited to, RNase inhibitors, ribonuclease inhibitors, protease inhibitors, microsomal membranes, and tRNAs, either alone or in combination.

Viral protein expression and capsid assembly

[0070] The present invention also provides a method of expressing a protein using a composition of the invention. In an exemplary embodiment, the cell-free system is used to mimic capsid biogenesis and assembly. In the cell-free system, viral capsid transcripts are translated in the presence of wheat germ extract that contains soluble factors necessary for capsid protein translation and subsequent capsid assembly (Lingappa et al. *J.Cell Biol* 136:567-581 (1997)). Both cytosolic and membrane proteins present in the wheat germ extract may be involved in capsid assembly and/or viral replication. Integral membrane proteins can include transmembrane proteins. In those embodiments utilizing viruses requiring membrane proteins for capsid assembly, appropriate membranes can be added to the cell-free translation mixture. It is further possible to supplement the cell-free translation mixture with other exogenous proteins, such as chaperone proteins that can for example, facilitate the assembly of capsid intermediates. Assembly of capsids in the cell-free system minimally requires expression of only the particular viral protein(s) that are involved in capsid assembly. Once expressed, polypeptides proceed to assemble into capsids that are catalyzed by host factors.

[0071] After incubation for a time sufficient to produce capsids, products of the cell-free reaction can be analyzed to determine sedimentation value, buoyant density, and electron microscopy appearance. Together these form a sensitive set of measurements for integrity of capsid formation.

[0072] Synthesized viral proteins can be detected in any manner known in the art. In an exemplary embodiment, radiolabeled capsid polypeptides using labeled amino acids are used. In one embodiment using this approach, ³⁵S methionine is added to the translation mixture. Following *in vitro* expression, velocity sedimentation gradients can generate fractions that are aliquoted into loading buffers and run on a standard SDS-PAGE gel. The gel can be exposed to film that generates autoradiographs showing the amount of ³⁵S labeled viral protein in different fractions of the velocity sedimentation gradients. Alternatively, as is known in the art, a phosphoimager can be used to visualize radiolabeled viral proteins.

[0073] In various embodiments, antibodies, e.g., commercially available antibodies, are used to detect successful protein expression or capsid assembly. Alternatively, antibodies can be specifically raised against proteins of interest, as is well known in the art.

[0074] Viral protein expression and capsid assembly in the cell-free system of the invention is useful for viruses from any family including, without limitation, *Flaviviridae*, *Togaviridae*, *Bunyaviridae*, *Arenaviridae*, *Filoviridae*, *Poxviridae*, *Orthomyxoviridae*, *Rhabdoviridae*, *Herpesviridae*, *Coronaviridae*, *Paramyxoviridae*, *Hepadnaviridae*,
5 *Bornaviridae*, *Picornaviridae*, *Retroviridae*, *Reoviridae*, *Papillomaviridae*, *Adenoviridae*, *Astroviridae*, *Polyomaviridae*.

Non-viral proteins

[0075] The present invention also provides a method of expressing non-viral proteins using a composition of the invention. In an exemplary embodiment, the cell-free system is used to
10 mimic a pathway of assembly into a multi-protein complex for which the viral capsid serves as an analogous model. Figure 14 illustrates the use of 5% wheat-germ extract to dissect various CNS-related proteins, including the DISC1 gene into a synthesis phase and an assembly phase. Because host proteins involved in viral capsid assembly must exist in the host for other purposes, the non-viral proteins that are the substrates for those endogenous
15 assembly pathways should be equally effectively usable for drug screening by the present invention. Thus, the present invention should be applicable not only to CNS disease, but also to proteins involved in other disorders including but not limited to metabolic, oncologic, and immunologic diseases. The present invention can be further used to mimic a bacterial or parasitic protein that forms a multiprotein complex that when disrupted, ameliorates bacterial
20 or parasitic disease. What is required for applicability of the present invention is that the newly synthesized proteins in question share the ability to use other proteins in the extract to assist or facilitate their assembly into distinct multiprotein complexes.

Assays for modulators

[0076] In some embodiments, modulation of proteins, including viral proteins, can be
25 assessed by determining the effect of a compound on expression, folding, and assembly of the protein in the cell-free system of the invention (e.g., with about 5% wheat germ extract). In other embodiments, modulation of viral capsid assembly can be assessed by determining the effect of a compound on capsid assembly using a cell-free system of the invention (e.g., with about 5% wheat germ extract). In some embodiments, modulation of capsid interacting
30 proteins, including but not limited to capsid assembly chaperone proteins, can be assessed by determining the effect of a compound on expression of the protein in the cell-free system of

the invention (e.g., using 5% wheat germ extract). Modulation can further include, but is not limited to, modulation of infection, replication, receptor binding, cell entry, particle formation, and the like.

[0077] An advantage of using the cell-free system of the present invention is that the process of capsid formation is slowed down, thus allowing for the targeting of capsid assembly processes for modulators of capsid assembly. An advantage of using cell-free systems of the invention, such as those having about 5% wheat germ extract, is an increased sensitivity for detecting compounds that otherwise would not be detected at higher wheat germ concentrations.

[0078] Measurement of modulation of a viral protein and/or viral capsid assembly can be performed using a variety of assays, *in vitro*, *in vivo*, and *ex vivo*. The assays described herein can use a full length viral protein, a variant, a mutant or a fragment thereof. A suitable physical, chemical or phenotypic change that affects activity, e.g., enzymatic activity, cell surface marker expression, viral replication and proliferation can be used to assess the influence of a test compound on the proteins expressed. The assay can also make use of one or more drug designed to block or alter protein activity, capsid assembly, or the associations of chaperones with viral proteins. The assay can also identify modulators of viral capsid assembly intermediates. Moreover, genomic nucleic acid can also be encapsidated into capsids, which can be used to design drugs that interfere with encapsidation and with the design of assay systems that examine the mechanism of action of drugs that inhibit encapsidation.

[0079] The binding assay can be either solid state or soluble. The protein or viral capsid can be bound to a solid support, either covalently or non-covalently. Often, the *in vitro* assays of the invention are substrate or ligand binding or affinity assays, either non-competitive or competitive. Other *in vitro* assays include measuring changes in spectroscopic (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein.

[0080] A high throughput binding assay can be performed in which the viral protein is contacted with a potential modulator and incubated for a suitable amount of time. The potential modulator can be bound to a solid support, and the protein added. Alternatively, the protein is bound to a solid support. A wide variety of modulators can be used, including, but

not limited to, small organic molecule, or a biopharmaceutical or biological entity, such as a protein, e.g., an antibody or peptide, a sugar, a nucleic acid, e.g., an antisense oligonucleotide or a ribozyme or siRNA, or a lipid. A wide variety of assays can be used to identify viral protein-modulator binding, including labeled protein-protein binding assays, electrophoretic mobility shifts, immunoassays, enzymatic assays, and the like. In some cases, the binding of the candidate modulator is determined through the use of competitive binding assays, where interference with binding of a known ligand or substrate is measured in the presence of a potential modulator. Either the modulator, the known ligand, or substrate is bound first; then the competitor is added. After the protein is washed, interference with binding, either of the potential modulator or of the known ligand or substrate, is determined. Often, either the potential modulator or the known ligand or substrate is labeled.

[0081] In high throughput assays, either soluble or solid state, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 350 (e.g., 384) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay many plates per day; assay screens for up to about 6,000, 20,000, 50,000, or more than 100,000 different compounds are possible using integrated systems.

[0082] High throughput screening methods involve providing a combinatorial small organic molecule or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds) can be used. Such “combinatorial chemical libraries” or “ligand libraries” are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

[0083] A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e., the number of

amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

[0084] Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton et al., *Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication No. WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091), benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs et al., *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara et al., *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann et al., *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen et al., *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho et al., *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell et al., *J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (see Ausubel, Berger and Sambrook, *supra*), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn et al., *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang et al., *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (see, e.g., benzodiazepines, Baum *C&EN*, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

Mammalian extracts

[0085] The concepts illustrated by the present invention are equally applicable to non-wheat germ extracts capable of supporting both protein synthesis and protein assembly. Thus, the lowest concentration of mammalian protein extracts capable of de novo synthesis of proteins is likely to be the most effective system for multiprotein assembly, by analogy to the present invention. One skilled in the art would be capable of using the information provided here for wheat germ extracts and applying it to development of extracts from mammalian cells.

Kits

[0086] In an exemplary embodiment, the invention provides kits for the cell-free expression of a protein. In various embodiments, the kit is of use to assemble a viral capsid. Exemplary kits of the invention include a cell-free mixture comprising not more than about 5% wheat germ extract and each of the components necessary for expression of the protein. The wheat germ extract and components included in the kit can be the wheat germ extract and components as described herein. The invention further provides kits for determining whether a compound modulates expression of a protein. In some embodiments, additional components can be included to modify expression of a viral protein or assembly of a capsid for suited purposes. Kits typically include instructions for optimization of protein expression and/or capsid assembly. Kits can optionally include instructions to practice a high-throughput method of assaying for an effect on activity of the proteins and/or capsid assembly of the invention, one or more containers or compartments (*e.g.*, to hold the probe, labels, or the like), a control modulator of the activity, and/or a robotic armature for mixing kit components or the like.

[0087] The following examples are intended to illustrate exemplary embodiments of the invention and do not limit the scope of the invention.

EXAMPLE 1

[0088] Viral capsid assembly was examined in cell-free protein synthesizing systems having a composition of 5%, 10%, or 20% wheat germ extract in the presence of compounds (or DMSO) control whose efficacy had been assessed against infectious viruses in cell culture.

Results

[0089] Protein synthesis was first quantified at a composition of 5%, 10%, or 20% wheat germ extract by blotting 2 μ L of translation extract (triplicates), and plotted as arbitrary density units upon quantitation of band intensity using an alkaline phosphatase conjugated secondary antibody-based colorimetric method using the Hepatitis C virus core polypeptide. Protein synthesis was roughly the same at each percentage of wheat germ extract (**Figure 1**).

Addition, protein synthesis was uniform regardless of wheat germ extract source or preparation conditions (**Figures 9-13**), and selective for a subset of biologically relevant and medically important proteins that form large multiprotein complexes (**Figure 15**).

[0090] Compounds that were shown to be potent against Hepatitis C, Influenza,

- 5 Venezuelan Equine Encephalitis, and Human Immunodeficiency viruses in cell culture were then correlated for potency against respective viral capsid proteins using the cell-free protein synthesis system at different wheat germ extract concentrations.

Hepatitis C (HCV)

- [0091] Optimal wheat germ extract concentration was determined for compounds shown to
10 be efficacious against Hepatitis C. Compound potency against infectious virus in cell culture ranged from weak (138-85; $EC_{50} > 50\mu M$) to strong (138-136; $EC_{50} = 100nM$) (**Figure 2A and 6**). Drug sensitivity displayed a greatest correlation to infectious virus drug sensitivity at 5% wheat germ extract with little apparent correlation at 10 or 20% (**Figure 2B**).

Influenza (FLUV)

- 15 [0092] Optimal wheat germ extract concentration was determined for compounds shown to be efficacious against Influenza. Compound potency against infectious virus in cell culture ranged from weak (138-65; $EC_{50} 1-20\mu M$) to strong (150-120; $EC_{50} < 0.16-4nM$) (**Figure 3A and Figure 8**). Drug sensitivity displayed a greatest correlation to infectious virus drug sensitivity at 5% wheat germ extract. The correlation is lost at 10% wheat germ extract, and
20 reappears to some extent at 20%, but not in full correlation with the potency against the infectious virus cell culture assay. (**Figure 3B**).

Venezuelan Equine Encephalitis Virus (VEEV)

- [0093] Optimal wheat germ extract concentration was determined for compounds shown to be efficacious against Venezuelan Equine Encephalitis Virus. Compound potency against
25 infectious virus in cell culture ranged from weak (158-80; $EC_{50} > 20\mu M$) to strong (150-133; $EC_{50} < 0.4\mu M$) (**Figure 4A**). Drug sensitivity displayed a greatest correlation to infectious virus drug sensitivity at 5% wheat germ extract and is greatly reduced at 10% wheat germ extract and mostly lost at 20% wheat germ extract. (**Figure 4B**).

Human Immunodeficiency Virus (HIV)

[0094] Optimal wheat germ extract concentration was determined for compounds shown to be efficacious against Human Immunodeficiency Virus. Compound potency against infectious virus in cell culture is shown in **Figure 7**. Drug sensitivity displayed a greatest correlation to infectious virus drug sensitivity at 5% wheat germ extract and is greatly reduced or lost at higher wheat germ extract concentrations (**Figure 5**).

Methods

[0095] Cloned cDNAs encoding the protein of interest were engineered behind an SP6 bacteriophage promoter and mammalian 5' untranslated coding region, and PCR products were generated, gel purified, phenol/chloroform extracted, ethanol precipitated, and dissolved in sterile water and adjusted to 1 mg/mL. The products were used to generate transcripts.

[0096] A transcription reaction is set up containing the following components at the indicated final concentration: Tris pH 7.9 45mM, MgAc 6mM, Spermidine 2mM, dithiothreitol 10mM, and 480μM each ATP, GTP, UTP, CTP, 500 units/mL SP6 polymerase, 25 units/mL of Rnasin, and 20 μg/mL of PCR product were incubated at 40°C for 2 1/2 hours, aliquoted, and frozen in liquid N2 and stored at -80°C.

[0097] The translation reaction is carried out in a 20μL volume in a 384 well plate format, with 1/5 of the volume comprising the transcription reaction. The balance of the reaction including the following components at the indicated concentration in the final translation reaction: ATP and GTP 1mM each, creatine phosphate 4mM, each of the 20 amino acids used for protein synthesis at 40μM, Rnasin 4 μg/mL, creatine kinase 50 μg/mL, wheat germ extract with the A260/280 absorbance values indicated in **Figures 11-13**, comprising approximately 5% or less of the translation volume.

Discussion

[0098] Families of viruses known to cause human disease are shown in Table 1, and are amenable to the approach that a putative capsid has been assembled and a pathway comprising distinctive assembly intermediates identified for every viral family studied. For 13 of these families corroboration against infectious virus in cell culture has been sought for compounds active in the cell-free screen – and multiple active pharmacophores identified in

each and every case. Each of virus families has assembled capsids in the cell-free capsid assembly system of the invention.

[0099] The results show that previous notions suggesting that wheat germ extract should be used at 20% or higher in cell-free capsid assembly systems are unfounded. This is because the greatest correlation between compound potency against viruses in cell culture and capsid assembly in cell-free systems are observed when wheat germ extract concentrations are at 5% as opposed to 10 or 20%. As a result, previous attempts to discover viral modulators using standard wheat germ extracts likely overlooked promising antiviral drug candidates.

[00100] **Table 1.** Families of viruses known to cause human disease.

Viral family	Virus	Putative Capsid Assembled	Putative Pathway Identified	Screen Established	Infectious virus validated hits	SAR improvement
Flaviviridae	HCV	+	+	+	+	+
	DENV	+	+	+	+	+
	WNV	+	+		+	
	YFV				+	
Togaviridae	VEEV	+	+	+	+	+
	WEEV	+			+	+
	EEEV	+			+	+
	CHIKV	+			+	+
Bunyaviridae	BVPU	+	+	+	+	+
	HNTV	+	+	+	+	
	PTV				+	
Arenaviridae	LASV	+	+	+	+	+
Filoviridae	EBOV	+	+	+	+	+
	MARV	+	+	+	+	+
Parvoviridae	MPXV	+	+	+	+	+
	RFXV	+			+	+
Orthomyxoviridae	FLUV	+	+	+	+	+
Rhabdoviridae	RAEV	+	+	+	+	+
Herpesviridae	CMV	+	+	(+)	+	+
Coronaviridae	SARS-CoV	+	+	(+)	+	
Paramyxoviridae	NIPV	+	+	(+)		
	RSV	<i>in progress</i>	<i>in progress</i>		+	
Hepadnaviridae	HBV	+	+	(+)		
Bornaviridae	BONV	+	+	(+)		
Picornaviridae	RHIV	+	+	(+)	+	
Reoviridae	HV1	+	+	+	+	+
	HV2	+	+		+	
	SV				+	
Reoviridae	ROTV	+	+	(+)		
Papillomaviridae	HPV	+	+	(+)		
Adenoviridae	ADNV	+	+	(+)	+	
Asenoviridae	ASTV	+	<i>in progress</i>	<i>in progress</i>		
Polyomaviridae	BKV	+	<i>in progress</i>			

[00101] It is understood that the example and embodiment described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of

this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

WE CLAIM:

- 1 1. A cell-free system for expressing a protein of interest, wherein the cell-free system
2 comprises not more than about 5% wheat germ extract and further comprises
3 components necessary for expression of the protein of interest
- 1 2. The cell-free system of claim 1, wherein the protein of interest is a viral protein.
- 1 3. The cell-free system of claim 2, wherein the viral protein is a viral capsid protein.
- 1 4. The cell-free system of claim 1, wherein the protein of interest is a capsid interacting
2 protein.
- 1 5. The cell free system of claim 2, wherein the viral protein is expressed by a virus
2 which is a member of a viral family selected from the group consisting of
3 *Flaviviridae, Togaviridae, Bunyaviridae, Arenaviridae, Filoviridae, Poxviridae,*
4 *Orthomyxoviridae, Rhabdoviridae, Herpesviridae, Coronaviridae, Paramyxoviridae,*
5 *Hepadnaviridae, Bornaviridae, Picornaviridae, Retroviridae, Reoviridae,*
6 *Papillomaviridae, Adenoviridae, Astroviridae, Polyomaviridae, Anelloviridae,*
7 *Parvoviridae, Caliciviridae.*
- 1 6. The cell-free system of claim 1, wherein the components comprise a member
2 selected from a buffer, an amino acid, a nucleic acid transcript and a combination
3 thereof.
- 1 7. The cell-free system of claim 6, wherein the components further comprise a member
2 selected from a detectable moiety, ATP, GTP, creatine phosphate, a labeled amino
3 acid, myristoyl CoA lithium salt, an RNase inhibitor, creatine kinase, a tRNA and a
4 combination thereof.
- 1 8. The cell-free system of claim 7, wherein the labeled amino acid is [35S] methionine.
- 1 9. The cell-free system of claim 6, wherein the nucleic acid transcript is assembled in
2 an *in vitro* transcription reaction.
- 1 10. The cell-free system of claim 6, wherein the nucleic acid transcript encodes a viral
2 protein.

- 1 11. The cell-free system of claim 10, wherein the nucleic acid transcript encodes a viral
2 capsid protein.
- 1 12. The cell-free system of claim 6, wherein the nucleic acid transcript encodes a viral
2 capsid interacting protein.
- 1 13. The cell-free system of claim 6, wherein the buffer comprises a member selected
2 from the group consisting of potassium acetate, spermine, dithiothreitol and a
3 combination thereof.
- 1 14. A cell-free system for assembling a viral capsid, wherein the cell-free system
2 comprises not more than about 5% wheat germ extract and components necessary for
3 expression of at least one viral proteins comprising the viral capsid.
- 1 15. The cell-free system of claim 14, wherein the viral capsid is from a viral family
2 selected from *Flaviviridae*, *Togaviridae*, *Bunyaviridae*, *Arenaviridae*, *Filoviridae*,
3 *Poxviridae*, *Orthomyxoviridae*, *Rhabdoviridae*, *Herpesviridae*, *Coronaviridae*,
4 *Paramyxoviridae*, *Hepadnaviridae*, *Bornaviridae*, *Picornaviridae*, *Retroviridae*,
5 *Reoviridae*, *Papillomaviridae*, *Adenoviridae*, *Astroviridae*, *Polyomaviridae*,
6 *Anelloviridae*, *Parvoviridae*, *Caliciviridae*.
- 1 16. The cell-free system of claim 14, wherein the components comprise a buffer, an
2 amino acid, and a nucleic acid transcript encoding the viral protein.
- 1 17. The cell-free system of claim 16, wherein the components further comprise a
2 detectable moiety, ATP, GTP, creatine phosphate, a labeled amino acid, myristoyl
3 CoA lithium salt, an RNase inhibitor, creatine kinase, and a tRNA.
- 1 18. The cell-free system of claim 17, wherein the labeled amino acid is [35S]
2 methionine.
- 1 19. The cell-free system of claim 16, wherein the nucleic acid transcript is assembled in
2 an *in vitro* transcription reaction.
- 1 20. The cell-free system of claim 16, wherein the nucleic acid transcripts encode
2 proteins that comprise the viral capsid.

- 1 21. The cell-free system of claim 16, wherein the buffer comprises a member selected
2 from the group consisting of potassium acetate, spermine, and dithiothreitol.
- 1 22. A method of expressing a protein of interest, wherein the protein of interest is
2 expressed using the cell-free system of claim 1.
- 1 23. The method of claim 22, wherein the protein of interest is a viral protein.
- 1 24. The method of claim 23, wherein the viral protein is a viral capsid protein.
- 1 25. The method of claim 22, wherein the protein of interest is a viral capsid interacting
2 protein.
- 1 26. A method of assembling a viral capsid using the cell-free system of claim 14.
- 1 27. A method of testing whether a compound modulates a cellular function, the method
2 comprising
- 3 (a) introducing the compound to the cell-free system of claim 1; and
- 4 (b) determining whether the cellular function is modulated.
- 1 28. The method of claim 27, wherein the compound is selected from the group
2 consisting of a small organic molecule and a biopharmaceutical.
- 1 29. A method of testing whether a compound modulates viral capsid assembly, the
2 method comprising
- 3 (a) introducing the compound to the cell-free system of claim 14; and
- 4 (b) determining whether the viral capsid assembly is modulated.
- 1 30. The method of claim 29, wherein the compound is selected from the group
2 consisting of a small organic molecule and a biopharmaceutical.
- 1 31. A kit for cell-free expression of a protein comprising:
- 2 (a) a cell-free mixture comprising not more than about 5% wheat germ extract;
- 3 (b) components necessary for expression of the protein; and

4 (c) instructions sufficient for use of the kit.

1 32. A kit for the assembly of a viral capsid comprising:

2 (a) a cell-free mixture comprising not more than about 5% wheat germ extract;

3 (b) components necessary for expression of proteins required for viral capsid
4 assembly; and

5 (c) instructions sufficient for use of the kit.

1 33. A kit for determining whether a compound modulates expression of a protein
2 comprising:

3 (a) a cell-free mixture comprising not more than about 5% wheat germ extract;

4 (b) components necessary for expression of the protein; and

5 (c) instructions sufficient for determining whether the compound modulates
6 expression of the protein.

1 34. A kit for determining whether a compound modulates viral capsid assembly
2 comprising:

3 (a) a cell-free mixture comprising not more than about 5% wheat germ extract;

4 (b) components necessary for expression of proteins required for viral capsid
5 assembly; and

6 (c) instructions sufficient for determining whether the compound modulates viral
7 capsid assembly.

1 35. A kit for determining whether a compound modulates multiprotein assembly
2 comprising:

3 (a) a cell-free mixture comprising not more than about 5% cellular extract;

4 (b) components necessary for expression of non-viral proteins; and

5 (c) instructions sufficient for determining whether the compound modulates
6 multiprotein assembly.

1 36. The cell-free system of claim 1, wherein the protein of interest is a non-viral protein
2 that forms a multiprotein complex.

1 37. The cell-free system of claim 36, wherein the multiprotein complex participates in
2 diseases comprising central nervous system disorders, metabolic disorders, oncologic
3 disorders, or immunologic disorders.

1 38. The cell-free system of claim 36, wherein the protein of interest is a bacterial or
2 parasitic protein.

1 39. The method of claim 26, wherein the protein of interest is a non-viral protein that
2 forms a multiprotein complex.

1 40. The method of claim 39, wherein the multiprotein complex participates in diseases
2 comprising central nervous system disorders, metabolic disorders, oncologic
3 disorders, or immunologic disorders.

1 41. The method of claim 39, wherein the protein of interest is a bacterial or parasitic
2 protein.

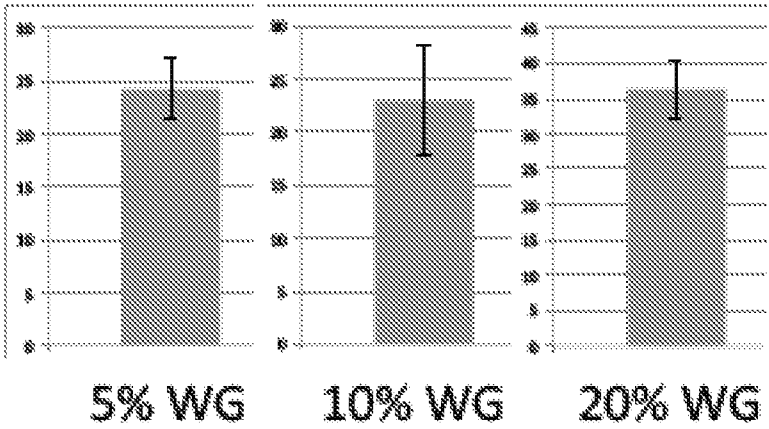


Figure 1

2/19

	EC50	EC99
138-85	>50uM	>50uM
150-114	0.25-2.5uM	2.5-7.5uM
105-123	0.25-1uM	1-2.5uM
138-136	0.1uM	0.25-1uM

Figure 2A

3/19

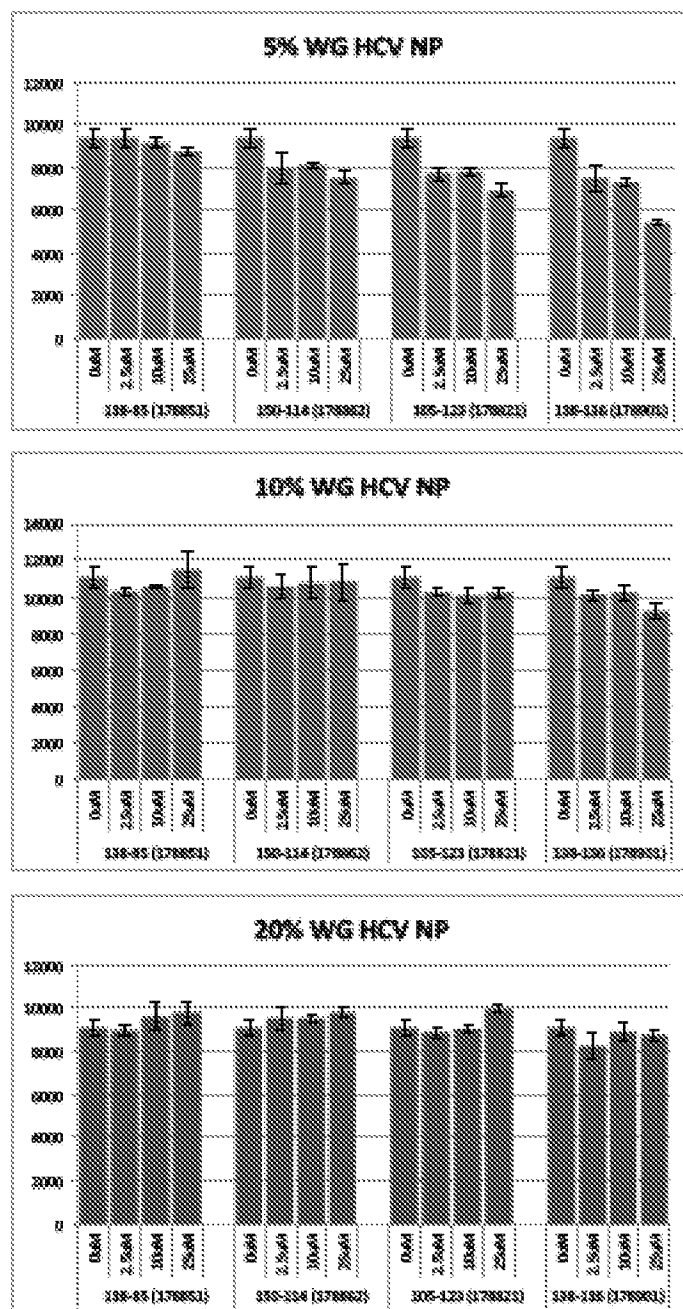


Figure 2B

4/19

A

	EC50	EC99
138-165	1-20uM	1-20uM
150-114	<0.5uM	<0.5uM
162-131	<0.1uM	>0.1uM
120-130A	<0.01uM	<0.14uM
150-120	<0.16-4nM	20-100nM

Figure 3A

5/19

B

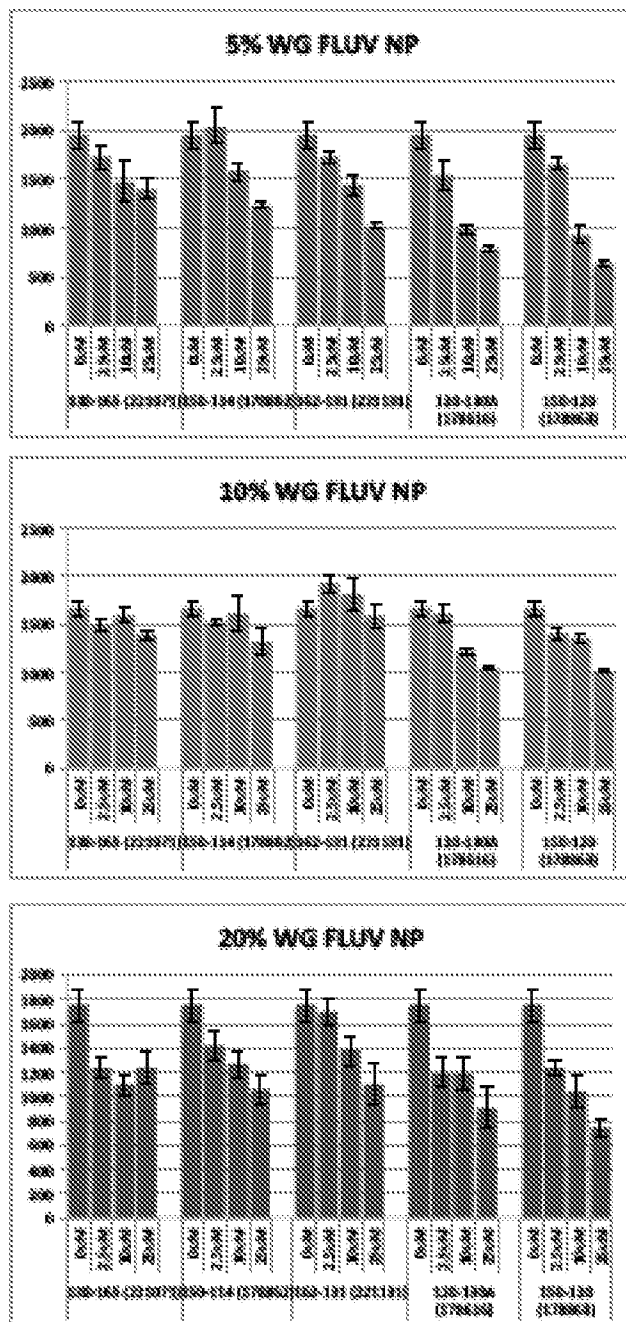


Figure 3B

6/19

A	EC50	EC99
158-80	>20uM	>20uM
120-69	1-7.5uM	7.5-20uM
138-159A	<1uM	5-10uM
162-31	0.2-0.75uM	>1uM
150-133	<0.04uM	0.2-0.75uM

Figure 4A

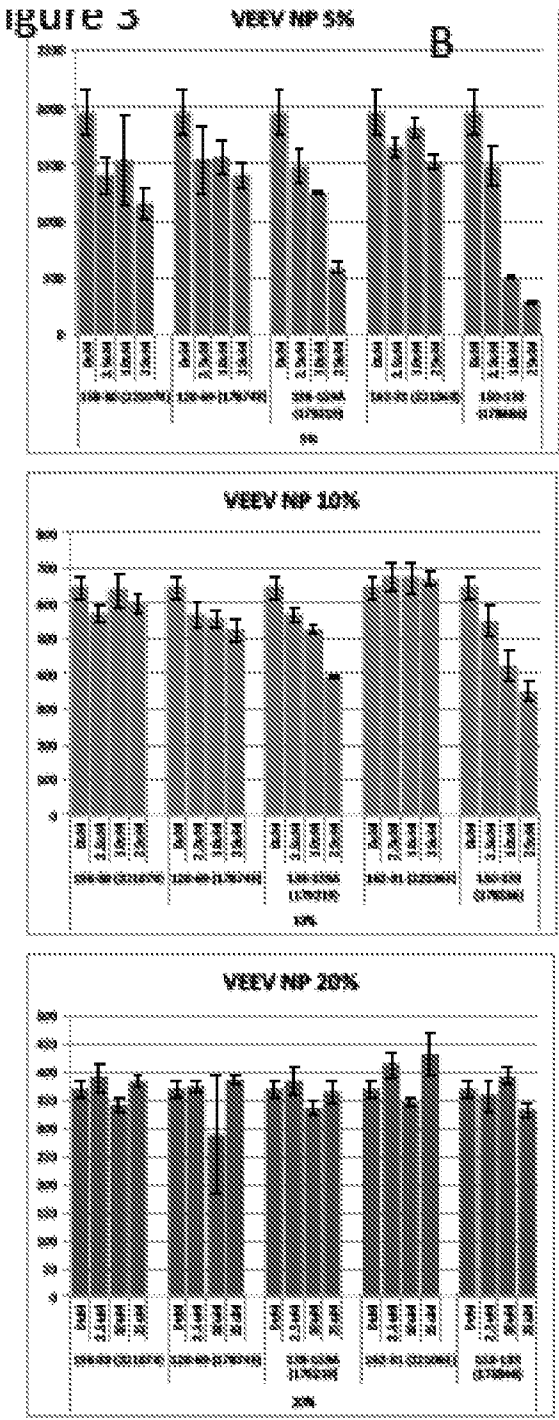


Figure 4B

8/19

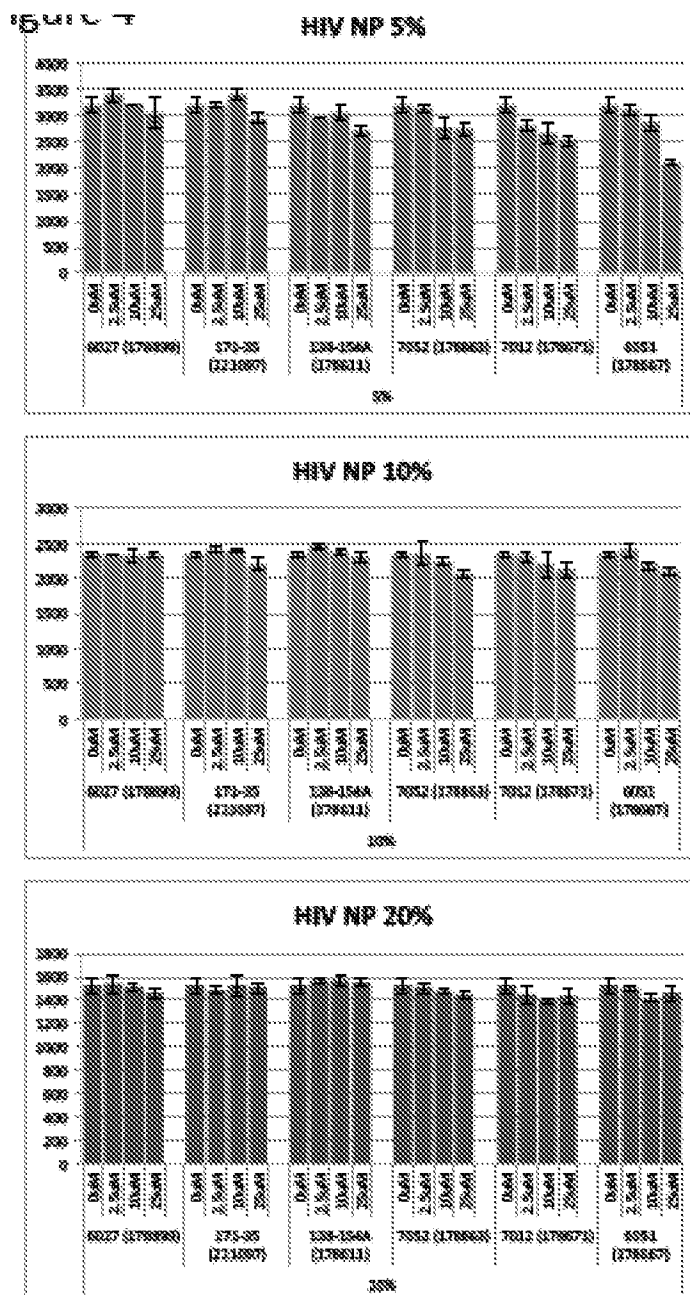


Figure 5

9/19

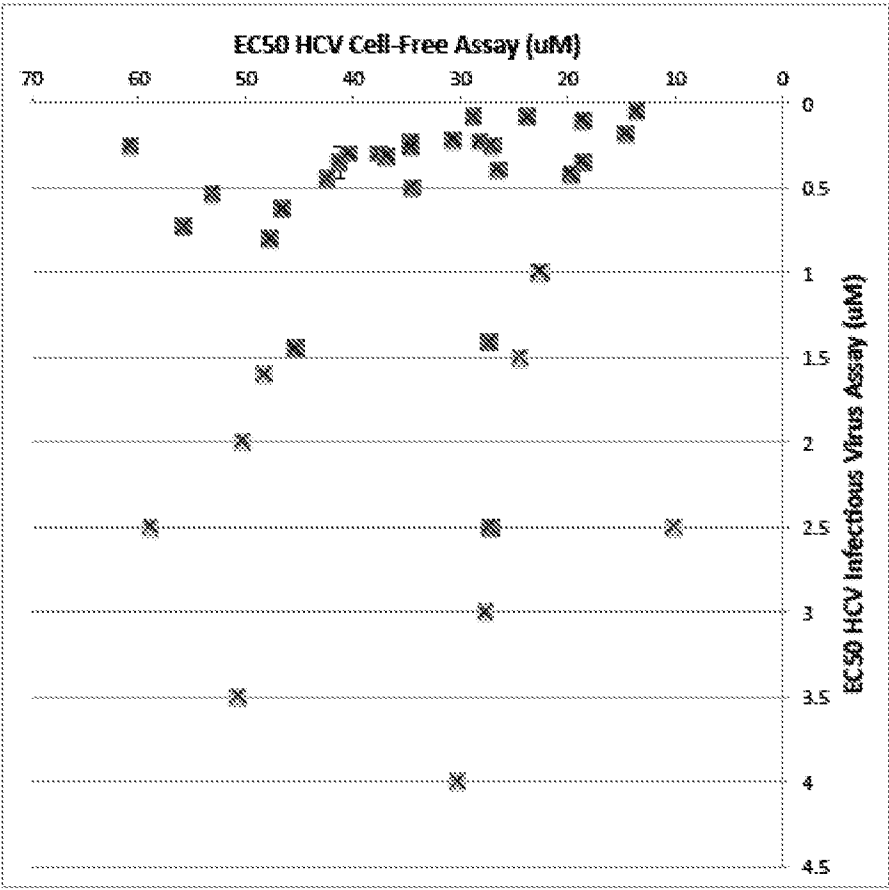


Figure 6A

10/19

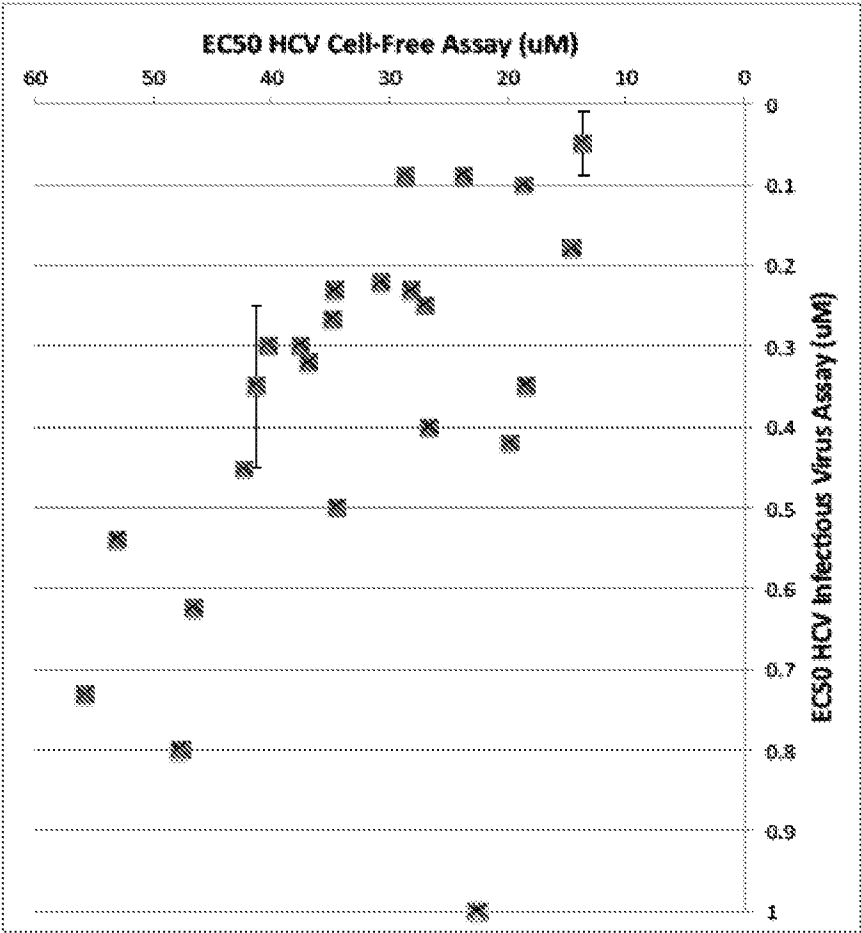


Figure 6B

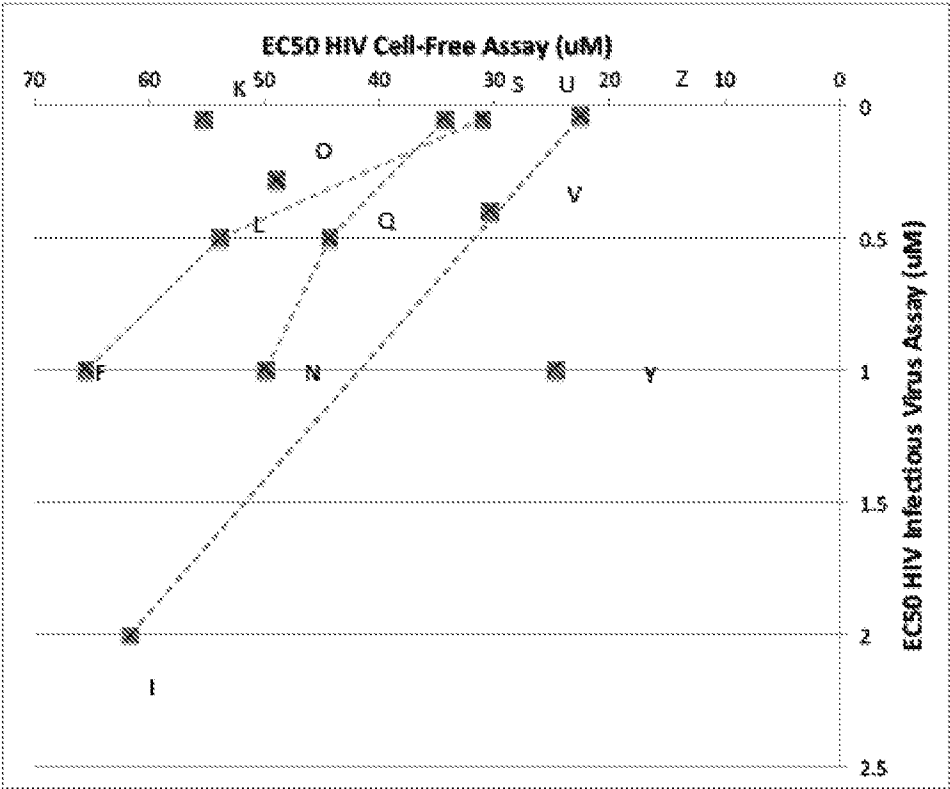


Figure 7

12/19

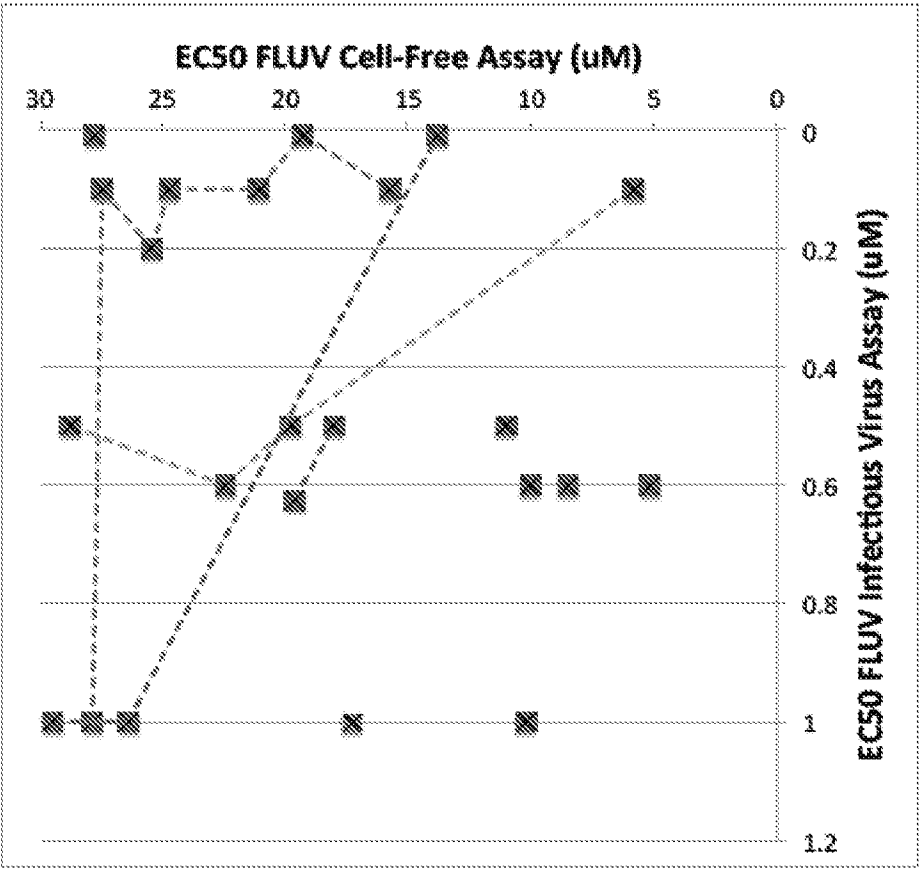


Figure 8

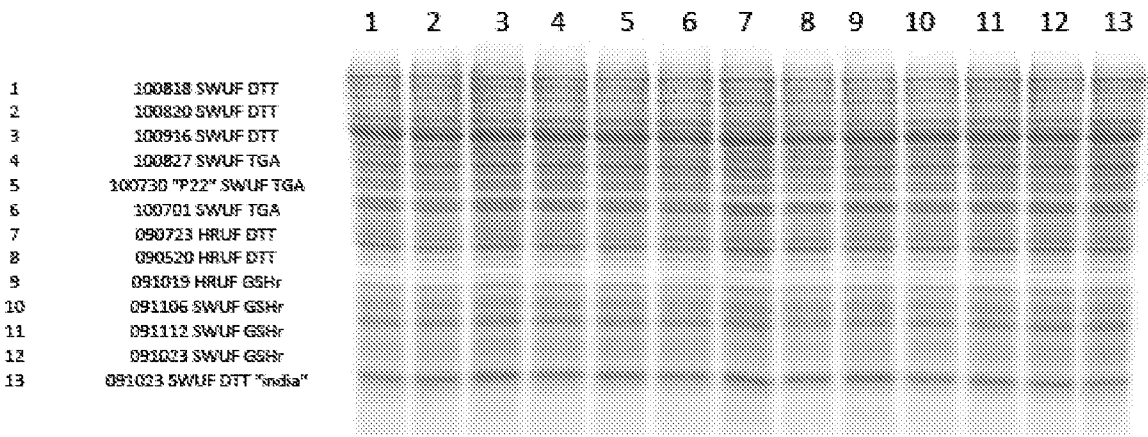


Figure 9

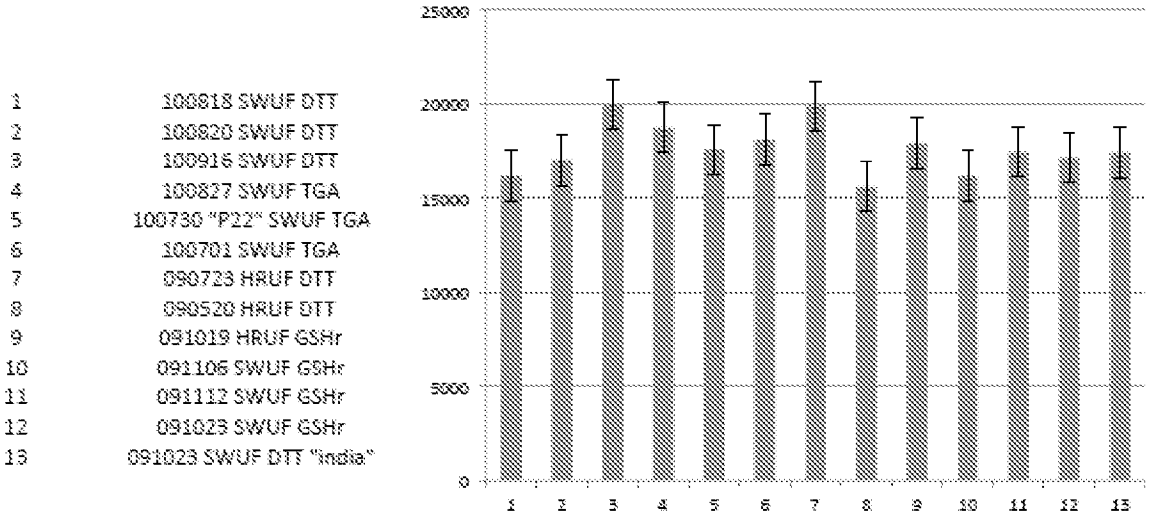


Figure 10

15/19

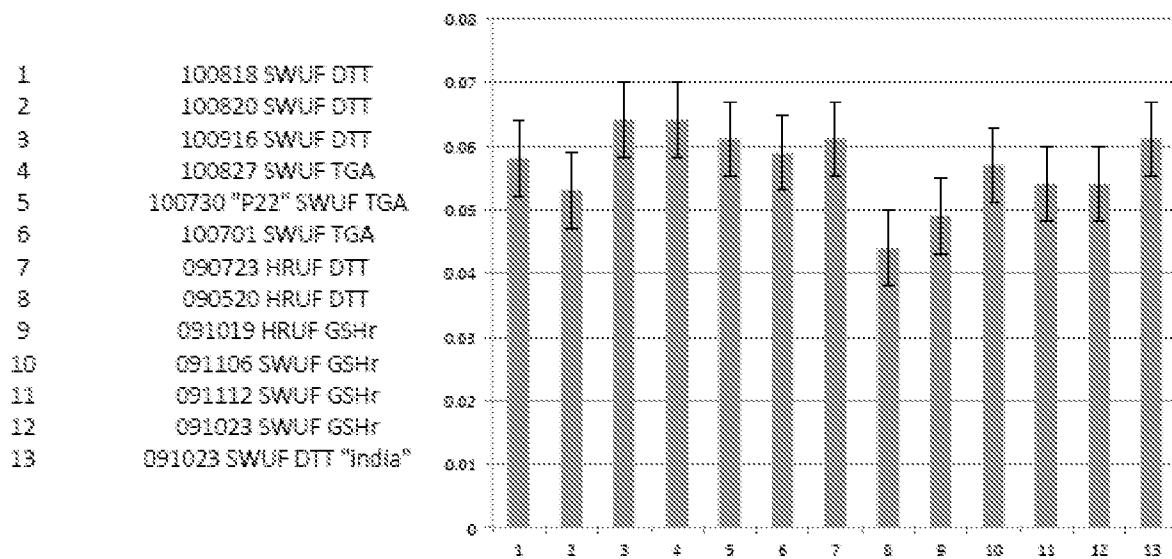


Figure 11

16/19

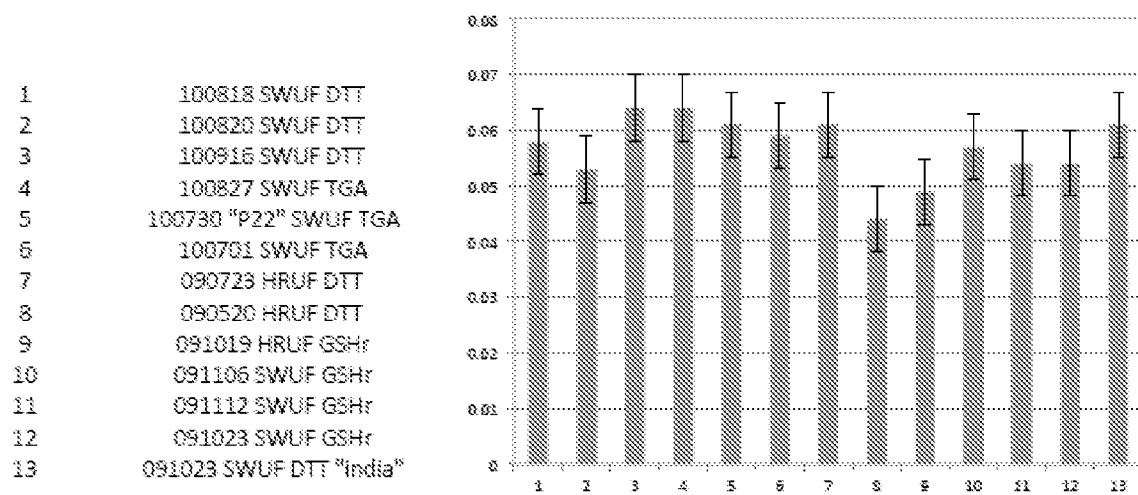


Figure 12

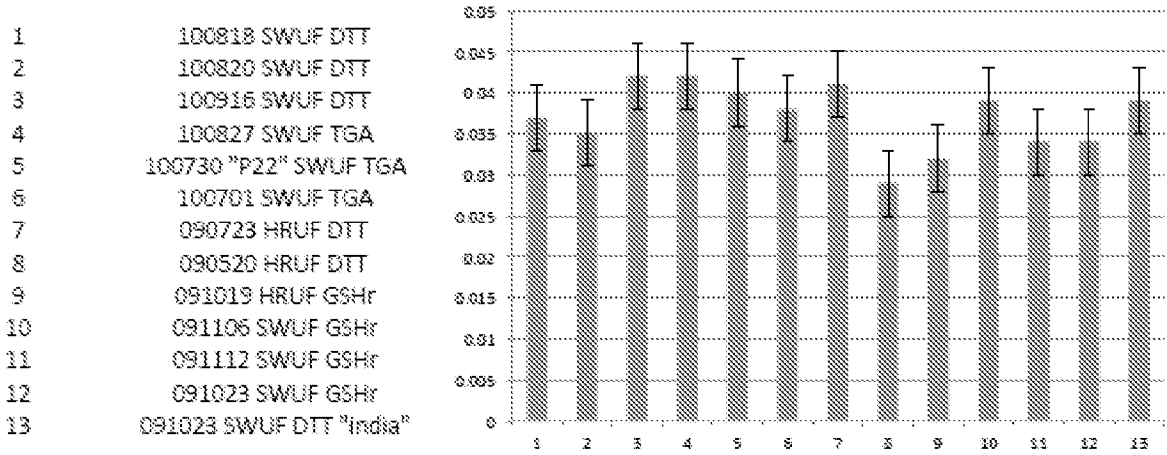


Figure 13

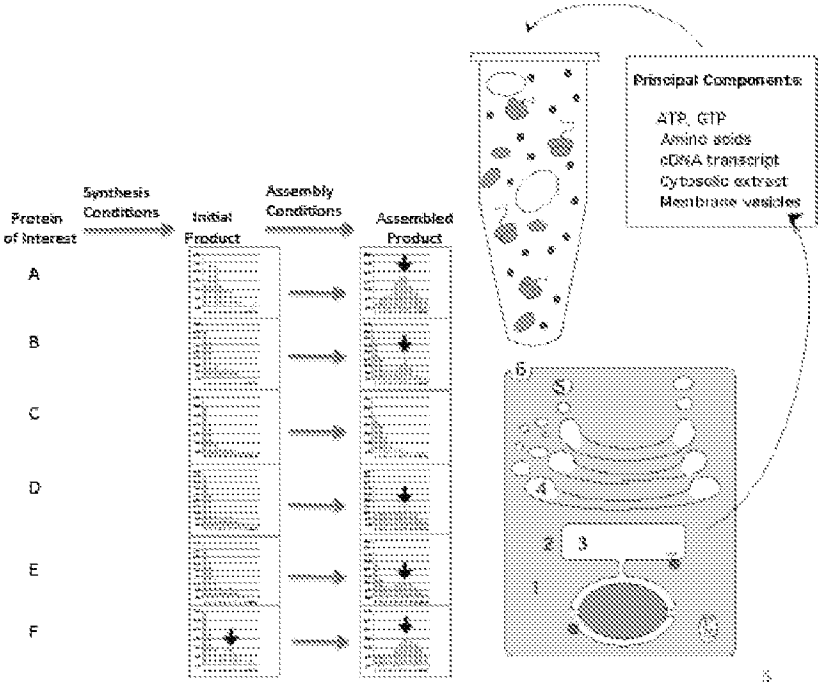


Figure 14

19/19

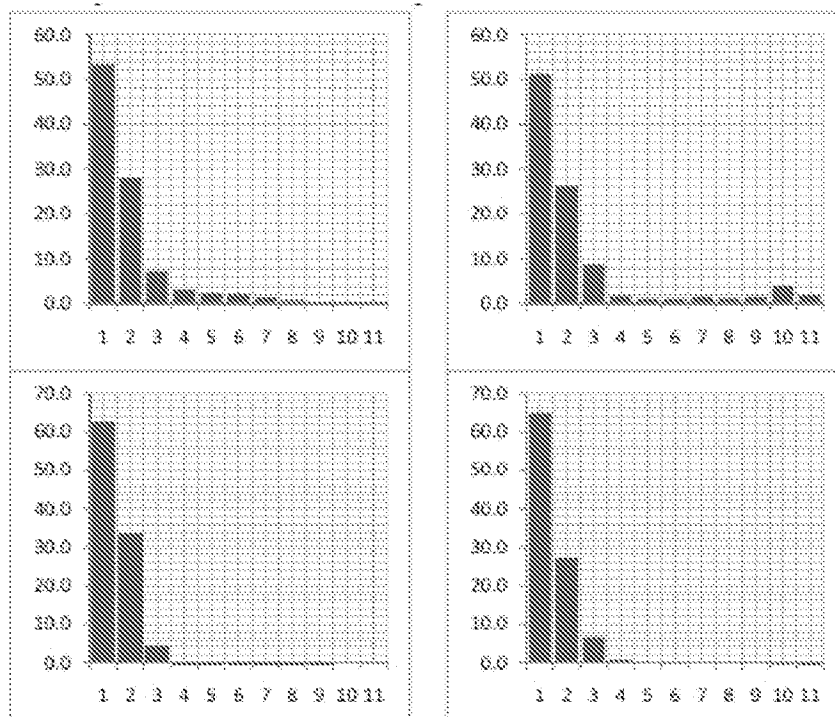


Figure 15