Title: CELL-FREE METHODS FOR DETECTING PROTEIN-LIGAND INTERCTIONS

(57) Abstract: Provided are rapid and sensitive cell-free assay methods for detecting and/or measuring specific bimolecular or higher order interactions via reassembly of a split monomeric reporter protein, and methods of detecting or identifying modulators of such interactions by the effect on the signal provided by the reassembled split reporter protein. This methodology is adaptable to protein-protein, protein-peptide, protein-nucleic acid, protein-methylated or nonmethylated nucleic acid and other small or large molecule ligands and binding proteins.
CELL-FREE METHODS FOR DETECTING PROTEIN-LIGAND INTERACTIONS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U. S. Provisional Application No. 61/001,370, filed November 1, 2007; U. S. Provisional Application No. 61/072,581, filed April 1, 2008; and U. S. Provisional Application No. 61/072,616, filed April 1, 2008, all of which are incorporated by reference herein to the extent there is no inconsistency with the present disclosure.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] This invention was made with government support under R21 CA122630, R01 AI068414 and GM077403 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0003] The present invention relates to the field of molecular biology, especially as related to methods for sensitive assays for assessing protein-protein, protein-ligand or protein-nucleic acid interactions, and antagonists and/or agonists of such interactions. Specifically this invention relates to split monomeric reporter protein systems including, but not limited to, split luciferase, β-lactamase or fluorescent protein reporter systems, and excluding beta-galactosidase, where a functional protein results when the portions interact, with the result that there is a detectable signal produced in the assay. In particular, the split reporter is expressed in a cell-free system.

[0004] Protein-protein (1) and protein-nucleic acid (2) interactions are central to cellular function and are also emerging targets for pharmacological intervention
when implicated in a particular disease pathway. Thus numerous in vitro and in vivo methods have been developed to target (3-7) and study these biomolecular interactions. Widely utilized in vitro methods for interrogating protein-protein and protein-DNA interactions and their antagonists include variations of enzyme linked immunosorbent assays (ELISAs), surface plasmon resonance (SPR), fluorescence resonance energy transfer (FRET) and fluorescence polarization (FP), which either require the use of antibodies or purified proteins and in some cases require chemical derivatization. On the other hand, powerful in vivo methods such as the yeast two-hybrid (8) assays have the advantage of speed by eliminating the need for protein purification but can be subject to false positives and negatives due to the multifactorial nature of signal generation (9). In between these two extremes lies the protein fragment based methods, where a specific biomolecular interaction drives the reassembly of a previously split reporter protein (10) (Figure 1).

[0005] Whereas there are various methods employing split reporter proteins, the present inventors are not aware of any methods in which there is cell-free expression of one or both of the split monomeric reporter proteins and subsequent assay of the expressed, assembled reporter in such an assay. Examples of methods employing living cells or transgenic organisms are provided in US Patent Publications 2005/0144661, 2004/0235064; 2007/0161067; 2006/0224331; and US Patents 6,897,017; 6,872,871; 7,166,424; 7.160,691; 6,828,099; 6,428,951; 6,929,916; 7,062,219; and 7,176,287. See also Kim et al. (130); Porter et al. (23); Porter et al. (58); Paulmurugan et al. (131).

[0006] There is a need in the art for assays of molecular interactions which are fast, require relatively little culture and handling of samples, and are sensitive, accurate and precise.

SUMMARY OF THE INVENTION

[0007] The present invention provides methods for rapid and sensitive assays for detecting protein-protein, protein-nucleic acid, protein-small molecule or other protein-ligand interactions, and antagonists and/or agonists of such an interaction using split monomeric protein reporter systems including, but not limited to those generating enzymatic activity, bioluminescence, chemiluminescence, fluorescence or
absorbance, for example using luciferase, β-lactamase or a fluorescent protein reporter system, but excluding beta-galactosidase in a cell-free assay system. The two portions of the reporter protein come together in a cell-free assay and their association is mediated by an interaction of an attached protein and its specific binding ligand, which can be an antibody or other protein, a specific nucleic acid sequence or a methylated or nonmethylated nucleic acid molecule, a single- or double-stranded RNA molecule, a small molecule, hormone or growth factor, among others. Protein-ligand and protein-small molecule interactions can be assessed when at least one portion of the reporter protein is covalently or noncovalently linked to either a ligand or to an antagonist or agonist of a bimolecular interaction and the second, complementing portion of the reporter protein is expressed in a cell-free translation system. Interaction of the two binding partners, with either their ligands or each other, brings the two portions of the split reporter protein into sufficiently close proximity that the two portions reassemble into a functional protein with, for example, detectable enzymatic or other activity. Antagonists or agonists of such interactions can be assessed by detecting the displacement of one binding partner, and the resulting decrease in reporter signal or by detecting enhanced interaction via increased reporter signal, respectively. Within the present methods, at least one portion of the reporter protein is synthesized in an in vitro translation assay, and it may be synthesized after in vitro transcription of the mRNA encoding that protein.

[0008] The fusion protein supplying the ligand binding portion associated with a split reporter can be one which interacts specifically with another protein. In the context of the assay, the ligand binding portion can be a protein, modified protein (e.g. phosphorylated, glycosylated), enzyme, hormone, antibody (Ab), single chain Ab, antigen-binding fragment of an Ab (e.g., Fab) or other protein. The Ab, single chain Ab, antigen-binding fragment of an Ab can be recombinant or derived from a natural source including, without limitation, e.g. camel, chicken, rabbit, mouse, rat, monkey, sheep, and goat. The ligand can be a small molecule, peptide, protein, single-stranded or double-stranded DNA or RNA molecule, or methylated or nonmethylated DNA molecule. Where the ligand binding domain and ligand are both proteins, the protein can be, without limitation, p53 and HDM2; Bcl and Bak; FKBP and FRAP; BAD and BCLXL; p38α MAPK and MAPK-activated protein kinase 2; cMyc and Max; HIF1α and p300; Fos and Jun; PIN1 and Jun; and PKA and PKI, or
an antigen and its cognate Fab fragment or antigen-binding fragment of a single chain Ab. Where the ligand is a small molecule or peptide, the ligand can be an agonist or antagonist of the ligand binding protein.

[0009] Where the ligand is a DNA, the protein segment binding the DNA can be a zinc finger, a helix-turn-helix protein, a leucine zipper protein, a helix-loop-helix protein, a transcriptional activation factor or a negative regulatory protein or other protein involved in transcription or DNA recognition. The protein segment binding the DNA is advantageously fused to a split reporter protein. Advantageously the DNA-binding segment and the split reporter are encoded as a fusion protein coding sequence.

[0010] Where the ligand is an RNA molecule, the RNA-binding segment can be a pumilio domain, a KH domain, RRM domain, Argonaute, MS2 coat protein, eukaryotic initiation factor 4a, or other proteins or protein-RNA complexes involved in translation or RNA recognition. The protein segment binding the RNA is advantageously bound or fused to a split reporter protein component. Advantageously the RNA-binding segment and the split reporter are encoded as a fusion protein coding sequence.

[0011] It is understood that there can be a linker region between the ligand binding portion and the reporter fragment portion of one or both components of the split reporter system, especially if necessary to avoid steric hindrance of the bound ligand with respect to the reassembly of the split reporter proteins.

[0012] A fluorescent protein can be a naturally occurring or engineered or enhanced green, blue, yellow, red or other fluorescent protein. A green fluorescent protein or variant can be one derived in sequence from or modified from Aequoria, or Discosoma. Luciferase can be one derived in sequence from or modified from firefly (i.e. Photinus pyralis), Renilla or Gaussia. Beta-lactamases are known to the art, as are its chromogenic or fluorogenic or luminescent substrates, for example, nitrocefin or CCF2FA.

[0013] The cell-free translation machinery can be a mammalian, plant, fungal or bacterial translation system. A cell-free translation system, often a crude cell extract, contains all the macromolecular components (70S or 80S ribosomes, tRNAs,
aminoacyl tRNA synthetases, initiation, translocation and termination factors, etc) necessary for translation of exogenous RNA (mRNA). For efficient translation, the cell-free translation system is supplemented with amino acids, energy sources (e.g., ATP, GTP) energy regenerating systems (such as creatine phosphate and creatine phosphokinase for eukaryotic systems or phosphoenol pyruvate and pyruvate kinase for *Escherichia coli* lysate) and other cofactors including magnesium and potassium cations. The mammalian system can be rabbit reticulocyte lysate, HeLa cell extract, among others, and the plant cell extract can be from wheat, wheat germ, corn, pea, tobacco or other plant. The fungal cell-free extract can be from a fungus such as *Aspergillus nidulans* or *Neurospora crassa*, among others, and it can be from a yeast such as *Saccharomyces cerevisiae*, *Pichia pastoris* or *Candida albicans*. A bacterial cell-free extract can be prepared from *Escherichia coli*, among others. Besides cell-free extracts, the translation machinery can be in the form of purified components, as known to the art. Many of the foregoing systems are commercially available. An advantageous use of purified translation machinery is with the addition of unnatural amino acids used in translation, for example, amino acid analogues. In such systems tRNAs charged with natural and/or unnatural amino acids, as desired. A coupled transcription-translation system is one in which DNA serves as the template for the synthesis of RNA, which is not isolated or purified but is directly translated into protein in the assay system. Such systems generally employ a bacteriophage RNA polymerase and promoter (especially T7, T3 or SP6) used to drive expression of the split reporter protein in *vitro*. There are commercially available products of various types and the art knows the appropriate vector and sequence modifications for the system in which the split reporter(s) are produced.

[0014] With respect to nucleic acids, there can be sensitive and quantitative measurement of particular sequences of DNA or RNA, thus enabling the assessment of disease markers, for example, identification of up-regulated genes associated with diseased cells (including but not limited to cancer cells or cells with metabolic abnormalities), and deletions and recombination events in the genome or nucleic acids associated with genome or expression products of a pathogen, thereby permitting prediction of particular diseases. In addition, particular polymorphisms can be detected, as relevant to personalized medicine or predicting or diagnosing a disease. Finally, specific nucleic acid molecules characteristic of a particular
pathogen can be detected, thereby permitting confirmation of the presence of the pathogen in a biological, environmental, commercial, pharmaceutical, food (e.g. vegetables, fruit, dairy product, meat, poultry, fish for human or animal consumption) or water sample.

[0015] With respect to proteins, there can be sensitive and quantitative measurement of particular proteins, thus enabling the assessment of disease markers, for example, identification of up-regulated, down-regulated, mutated, or post-translationally modified proteins associated with particular diseases and therapies as relevant to personalized medicine or predicting or diagnosing a disease or following the progress of therapies. Finally, specific proteins characteristic of a particular pathogen can be detected, thereby permitting confirmation of the presence of the pathogen in a biological, environmental, commercial, pharmaceutical, food (e.g. vegetables, fruit, dairy product, meat, poultry, fish for human or animal consumption) or water sample. In addition, the ligand can be an amyloidogenic protein comprising beta-amyloid(1-40, 1-41, 1-42, 1-43), prion protein, alpha-synuclein, tau, immunoglobulin, islet amyloid polypeptide or huntington protein, and there can be achieved a diagnosis or prognosis of Alzheimers disease or a prion disease or contamination of a sample with a prior, especially one associated with human or animal disease.

[0016] The ligand can be contained within a sample, biological or otherwise as set forth above or it can be a recombinant or synthetic molecule.

[0017] In the case where a nucleic acid molecule, binding or interacting region of a protein, small molecule or other ligand is tethered to a segment of a split reporter protein, the tether can be from 0-50 amino acids in length. In other words, the binding or interacting portion is fused to the split reporter portion via a sequence of 0-50 amino acids.

[0018] The present invention further embodies a method for identifying an antagonist or an agonist of protein-protein interaction using a cell-free system comprising coupled or uncoupled transcription and translation machinery; containing RNA or DNA encoding a first fragment of a reporter operably linked and in frame to RNA or DNA encoding a first interacting protein, wherein the cell-free system
expresses at least one of the split reporter-binding segment proteins and a mixture containing both (complementing) fragments is in contact with the cognate ligand in the presence and absence of a test composition. The signal of the detectable reporter is measured in the presence and absence of the test composition, and an antagonist is identified when the signal in the presence of the test compositions is lower than in its absence and an agonist of binding is identified when the signal in the presence of the test composition is greater than in its absence. This methodology is fully adaptable to other ligand-ligand binding protein interactions (protein-small molecule, protein-nucleic acid, and the like).

[0019] In any of the methods provided herein it is advantageous that the transcription product encoding one or both of the binding region-split reporter portions comprises a stabilizing element such as a 5' stem-loop, including but not limited to a 5' stem-loop derived in sequence from bacteriophage T7 and advantageously a 3' stem-loop such as that derived in sequence from bacteriophage T3 (see, e.g., 134, 135); and also advantageously a Kozak sequence 5' to the translation start site (see, e.g., 133) and a polyadenylated 3' end (see, e.g., 136) when the cell-free translation system is a eukaryotic system or a Shine-Delgarno sequence when the cell-free translation system is a bacterial translation system.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1. Cell-free detection utilizing split-proteins. (a) Cartoon representation of a split protein-system with zinc-fingers tethered to the split-proteins in the presence of target dsDNA oligonucleotide. Different split-protein reporters tethered to sequence specific zinc-fingers in the presence and absence of target dsDNA; (b) split-Venus (a GFP variant); (c) split-β-lactamase; (d) split-firefly luciferase as described by Luker (27) et. al.; (e) split-firefly luciferase as described by Paulmurugan (28) et. al.; (f) split-Gaussia luciferase as described by Remy et al. (20).

[0021] Figure 2. Cell-free detection of a wide variety of biomolecular interactions utilizing split-luciferase starting from mRNA. Detection of (a) protein-protein interaction between PKI-NFluc and CFluc-PKA; (b) rapamycin induced interaction between FRB-NFluc and CFluc-FKBP; (c) protein-DNA interaction between PBSII-
NFluc, CFluc-Zif268, and a target dsDNA oligonucleotide; (d) methylation dependent protein-DNA interaction between MBD2-NFluc, CFluc-Zif268, and a target methylated CpG dsDNA oligonucleotide; and (e) protein-RNA interaction between Pum2-NFluc, CFluc-Pum1, and a target RNA oligonucleotide.

[0022] Figure 3. Cell-free detection of biomolecular interactions with split-luciferase starting from DNA, cartooned in (a). Utilizing transcription and translation coupled unpurified cell-free lysate system: (b) protein-DNA interaction between PBSII-NFluc, CFluc-Zif268, and 5nM target dsDNA oligonucleotides and (c) rapamycin (5 nM) induced interaction between FRB-NFluc and CFluc-FKBP. Detection of biomolecular interactions utilizing purified DNA in the PURESYSTEM classic II system consisting of completely purified transcriptional and translational components (d) protein-DNA interaction between PBSII-NFluc, CFluc-Zif268, and 5nM target dsDNA oligonucleotide and (e) rapamycin (5 nM) induced interaction between FRB-NFluc and CFluc-FKBP.

[0023] Figure 4. Interrogation of protein-nucleic acid interactions utilizing split-luciferase cell-free assay. (a) Cartoon showing dissociation of dsDNA dependent firefly luciferase ternary complex by the addition of a competitor hairpin DNA containing one of the two zinc-finger binding sites. (b) Dissociation of the reassembled PBSII-NFluc, CFluc-Zif268, dsDNA ternary complex by the addition of Zif268 hairpin DNA targets containing TGG (wild type), TAG, TTG, and TCG triplet basepairs. (c) and (d) Previously reported relative affinities (49) of target oligonucleotides with Zif268 with IC50 values derived from the cell-free firefly luciferase reassembly assay and their correlation. (e) Dissociation of the Pum2-NFluc, CFluc-Pum1, RNA ternary complex by the addition of an RNA target containing a Pum1 binding site.

[0024] Figure 5. Interrogation of small molecule and peptide modulators of protein-protein interactions utilizing the split-firefly cell-free assay. (a) Concentration dependent association of FRB-NFluc and CFluc-FKBP mediated by rapamycin (inset). (b) Concentration dependent dissociation of PKI-NFluc/CFluc-PKA complex by PKI peptide. (c) Concentration dependent dissociation of the reassembled p300-NFluc/CFluc-Hif1α complex by chetomin (inset).
[0025] Figure 6. Protein-protein and protein-DNA interactions in artificial cells interrogated by split-luciferase dependent bioluminescence. (a) White light microscope image of a water-in-oil emulsion containing a PKI-NFLuc/CFLuc-PKA translation in wheat germ extracts. Scale bar is equal to 75 μm. (b) Protein-protein association (PKI-NFLuc/CFLuc-PKA) dependent split-luciferase reassembly and bioluminescence within water-in-oil emulsion. (c) Protein-DNA interaction dependent firefly luciferase reassembly within water-in-oil emulsion.

[0026] Figure 7. A split-luciferase sandwich assay for the detection of extracellular proteins. (a) A general schematic of the designed system is shown. Specific recognition elements are used to reassemble luciferase in the presence of a target extracellular protein leading to the generation of light. (b) A schematic of the VEGF assay is shown. Flt-1 (red and blue) is attached to both the N- and C-terminal halves of luciferase and is used to directly detect the VEGF homodimer. (c) Luminescence from reassembled luciferase in the presence and absence of VEGF. In the presence of 50 nM VEGF dimer a >15-fold increase in luminescence is observed.

[0027] Figure 8. An antibody enabled split-luciferase assay for gp120 detection. (a) A schematic of the solution phase detection system for gp120 is shown. (b) The specificity of the solution phase gp120 detection system is shown. Assays were performed on the indicated wild-type or mutant gp120s; D368R and I420R mutations are known to decrease CD4 and 17b binding respectively.9,10 The inset shows the luminescence signal generated from the assay when either DTT or PDI are included during translation. (c) The specificity of the gp120 detection system, as determined by luciferase reassembly, across a panel of gp120s from the indicated clades is shown. The observed luminescence highlights the ability to rapidly categorize HIV-1 clades using this assay.

[0028] Figure 9. A split-luciferase sandwich assay for the direct detection of HER2 on human cells. (a) An overlay of HER2 (tan and light blue) with the bound luciferase fusion proteins is shown. (b) A HER2 sandwich assay performed on purified HER2 expressed from Lec1 cells. The inset shows a western blot analysis of the purified HER2 protein, lane 1 molecular weight standards and lane 2 purified HER2 protein. (c) A HER2 sandwich assay performed on human breast cancer cells;
SK-BR-3 or MCF7 cells were added after translation and luminescence was monitored after 30 min (1 x 10^4 cells during luminescence assay). The inset shows the expression levels of HER2 in the indicated cell line (15).

[0029] Figure 10. Panel a shows the ribbon models of HTB1 and TJ10, displaying the site of the 8 mutations in TJ10 relative to the patent HTB1 protein. Panel b shows that class 1 molecules are inhibited by TJ10 (blue, central bars) but not HTB1 (red bars, right bars). Panel c shows that Class II aggregation of various protein interactions are inhibited by both TJ10 (blue, central) and HTB1 (red, rightmost). Panel d shows that Class III aggregation interactions are inhibited by neither both TJ10 (blue, center bars) nor HTB1 (red, rightmost).

[0030] Figure 11. Panel a summarizes the association reaction and the luciferase reaction catalyzed by the reassembled luciferase fragments when amyloid aggregation mediates the generation of a functional luciferase from its split fragments. Panels b-d show time course results for the detection of early aggregation intermediates of Aβ40, prion fragment 106-126, IAPP and Aβ42 using TJ10 fusion constructions blue or HTB1 fusion constructs (red). Notably, there is little signal generated for the Aβ40 with the HTB1 construct, while the signal is greater for HTB1 than TJ10 with the the prion fragment. In Panels c and d only the TJ10 fusion proteins were used (solid lines, blue). Fluorescence is marked by the dotted lines.

[0031] Figure 12. Summary of the progressive association, assembly stages and conformational changes for Class I, Class II snf VClass III proteins.

[0032] Figure 13. Ternary complex formation driven by Jun-staurosporine. Small molecule inhibitors induce loss of activity.

[0033] Figure 14. ALU values for four kinases assayed with three known small-molecule kinase inhibitors. All inhibitor samples contain 125 nM Jun-staurosporine (jun-st.) in addition to 50 μM inhibitor.

[0034] Figure 15. Ternary complex formation facilitated by Tri-Staur. Addition of a kinase active site binding small molecule induces complex dissociation and loss of luciferase activity.
[0035] Figure 16. (a) Tri-Staur mediated luciferase activity. (b) Dissociation of the Tri-Staur, DHFR-NFluc, and CFluc-PKA ternary complex by the addition of free staurosporine.

[0036] Figure 17. ssRNA detection strategies using various polynucleotide binding domains. (A) Cartoon representations of pumilio, argonaute, and zinc finger binding domains attached to split-luciferase. The presence of target ssRNA results in luciferase reassembly and a luminescent signal. (B) NFluc-Pum2 and CFluc-Pum1 detect 10 nM cognate RNA. (C) 10 nM RNA annealed to the Pum2 guide (tan) allows for Ago and Pum1 binding. (D) NFluc-Pum2 and CFluc-Ago detect 10 nM RNA with an annealed Pum1 guide. (E) Employing both Ago constructs, NFluc-Ago and CFluc-Ago, along with guides for the Pum1 and Pum2 binding sites results in modest signal over background in the presence of target (10 nM).

[0037] Figure 18. Zinc finger-mediated single-stranded nucleic acid detection. (A) The designed hairpin-guides target a cognate single-stranded nucleic acid, allowing zinc finger mediated split-luciferase reassembly. (B) Hairpin-guides were designed to target 1 nM ssRNA and ssDNA, resulting in a 4.5-fold relative signal over background for each. A single G to T mutation in the ssDNA target knocked signal down to background levels. (C) Hairpin-guides were designed to detect a 295 nucleotide VEGF transcript (1 nM), resulting in a relative signal of 39-fold. (D) Hairpin-guides were designed to bind to a 201 nucleotide HER2 transcript (1 nM), resulting in a 4.3-fold relative signal.

[0038] Figure 19. Cartoons showing NFluc-Argonaute and CFluc-Pumilio, NFluc-Pumilio and CFluc-Argonature and CFluc-E2C and NFluc-Aart associations.

[0039] Figure 20. Split luciferase detection of a target nucleic acid (chemically modified or sequence specific). Inactive luciferase fragments are tethered to target binding domains which form a ternary complex upon nucleic acid target binding results in split-luciferase reassembly and luminescence.

[0040] Figure 21. Sequence specific reassembly of split-firefly luciferase. a) Cartoon representation of the detection of a 36-mer dsDNA sequence utilizing two 6-finger zinc fingers. b) Luciferase reassembly utilizing the 3-finger zinc finger Zif268 and the 6-finger zinc finger Aart in the presence or absence of the dsDNA target.
Zif268-1-Aart. c) Luciferase reassembly utilizing the 6-finger zinc fingers E2C and Aart in the presence of decreasing amounts of the target dsDNA E2C-1-Aart. (inset) Linear fit from 100 to 2 pM E2C-1-Aart.

[0041] Figure 22. Interrogation of 6-finger zinc finger, Aart, specificity. a) X-ray crystal structure of finger 3 of Aart in complex with target DNA 5'-AAA-3'. Specific hydrogen bond contacts are shown between residue N92 and the DNA position of interrogation. b) Cartoon showing approach for the interrogation of Aart specificity, where the addition of a competitor hairpin DNA containing a zinc finger binding site results in the disruption of dsDNA-firefly luciferase ternary complex formation and a loss in signal. c) Competition experiments in the presence of increasing concentrations of the competitor Aart hairpin DNA containing the consensus (con) and designed (des) recognition sequences and the 3rd finger 5'-AAA-3' to 5'-ATA-3' mutations of both the consensus (Con ATA) and designed (Des ATA) recognition sequences. d) Tabulated IC~50 values of Aart for the respective hairpin DNA targets.

[0042] Figure 23. Detection of DNA Methylation and Dimeric Transcription Factor Binding. a) C5-Cytosine methylation by DNA methyltransferase; b) Cartoon showing the site-specific determination of dsDNA methylation utilizing MBD2 and a sequence-specific zinc finger. c) Profile of different split-protein reporters tethered to MBD2 and Zif268 in the presence of a dsDNA target as a function of increasing number of base pairs between the mCpG and Zif268 sites. d) Luminescence of split-firefly luciferase tethered to MBD2 and the 6-finger zinc finger E2C in the presence of decreasing concentrations of the methylated dsDNA target mCpG-2-E2C. (inset) Linear fit from 100 pM to 10 pM mCpG-2-E2C. e) Cartoon showing the detection of dsDNA by split-luciferase utilizing MBD2 and the dimeric transcription factor λ-Cro. f) Luminescence in the presence of methylated (m) and non-methylated (u) target mCpG-2-λ-Cro containing the dimeric λ-Cro recognition site (10 nM).

[0043] Figure 24. Determination of total methylation at two sites. a) Cartoon showing the detection of di-methylated dsDNA utilizing fragmented luciferase tethered to MBD2. Addition of a di-methylated dsDNA target results in ternary complex formation and luminescence. b) Luminescence above background in the presence of dsDNA targets containing either 6 or 21 base pairs between mCpG
sites. c) Split-luciferase signal generation in the presence of decreasing amounts of the methylated target mCpG-6-mCpG. (inset) Linear fit from 20 to 0.2 nM mCpG-6-mCpG.

[0044] Figure 25. Direct detection of poly(ADP-ribose). a) Cartoon showing the detection of poly(ADP-ribose) utilizing split-luciferase tethered to the poly(ADP-ribose) binding zinc finger APLF. b) Ternary complex formation of APLF-NF\textsubscript{luc}/CF\textsubscript{luc}-APLF and PBSII-NF\textsubscript{luc}/CF\textsubscript{luc}-Zif268 in the presence of either 10 nM poly(ADP-ribose)(PAR) or 10 nM of the dsDNA target Zif-0-PBS. c) Luminescence as a function of PAR concentration. (inset) Linear fit from 2 to 0.125 nM PAR.

DETAILED DESCRIPTION OF THE INVENTION

[0045] Whereas there are various methods employing split reporter proteins, the present inventors are not aware of any methods in which there is cell-free expression of one or both of the split reporter proteins and subsequent assay of the expressed, assembled reporter in such an assay wherein there is detection of protein-protein, protein-nucleic acid or protein-ligand interactions or of agonists or antagonists of such interactions. Important advantages are that there is no requirement for introducing nucleic acid molecules encoding the segments of the reporter protein into living cells and subsequently obtaining gene expression therein and there is no need to purify the split reporter proteins prior to the assay. The methods of the present invention are readily adapted for high throughput (HTP) assays. As specifically shown herein, the measurement of bioluminescence in a split luciferase system is especially useful in an interaction monitoring assay.

[0046] Split-protein reporters have emerged as a powerful methodology for imaging biomolecular interactions which are of much interest as targets for chemical intervention. Here we describe a systematic evaluation of split-proteins, specifically the green fluorescent protein, beta-lactamase, and several luciferases, for their ability to function as reporters in completely cell-free systems to allow for the extremely rapid and sensitive determination of a wide range of biomolecular interactions without the requirement for laborious transfection, cell-culture, or protein purification (12-48 hours). We demonstrate that the cell-free split-luciferase system in particular is amenable for directly interrogating protein-protein, protein-DNA, and
protein-RNA interactions in homogenous assays with very high sensitivity (22 – 1800 fold) starting from the corresponding mRNA or DNA. Importantly, we show that the cell-free system allows for the rapid (2 hours) identification of target site specificity for protein-nucleic acid interactions and in evaluating antagonists of protein-protein and protein-peptide and protein-small molecule complexes circumventing protein purification bottle necks. Moreover, we show that the cell-free split-protein system is adaptable for analysis of both protein-protein and protein-nucleic acid interactions in artificial cell systems comprising water-in-oil emulsions. Thus this study provides a general and enabling methodology for the rapid interrogation of a wide variety of biomolecular interactions and their antagonists without the limitations imposed by current in vivo and in vitro approaches.

[0047] The present methods can readily be adapted for use in the sensitive, qualitative or quantitative determination of RNA and DNA, for example, specific disease markers such as those which are up-regulated in cancer or for specific sequences which are not expressed due to a genetic defect or a disease, for identifying deletions and recombination events in the genome which are associated with genetic defect or a genotype of interest, detecting single nucleotide polymorphisms associated with disease and detection of a pathogenic organism based on known DNA or RNA sequences or proteins.

[0048] The reconstitution of a functional protein from split-peptide fragments was first demonstrated for ribonuclease in 1959. Since then "split-protein reassembly" or "protein complementation" has been applied to the in vivo detection of a wide variety of protein-protein interactions utilizing numerous split-protein hosts including ubiquitin (12), beta-galactosidase(13), dihydrofolate reductase (14), beta-lactamase (15), GFP (16), GFP-variants (17,18), firefly luciferase (19) and Gaussia luciferase (20). Recently, we and others have also described methods for detecting nucleic acids and their chemical modification by the reassembly of ternary complexes of split-GFP and split-β-lactamase tethered to nucleic acid binding proteins (21-24). Thus split-protein systems or "protein complementation assays" (PCAs) can directly image most biomolecular interactions. Though the current methods are useful, all of these split-protein methods have certain limitations for interrogating protein-protein and protein-nucleic acid interactions and their inhibitors in a rapid and high-
throughput fashion. For example, current *in vitro* methods require extensive protein purification and also rely on proper folding of recombinant proteins, while *in vivo* methods require lengthy transfection and propagation of cellular cultures prior to analysis, both approaches being time intensive (16,24-26). Such methods are also prone to problems arising from proteolysis of intracellularly expressed proteins and peptides as well as a lack of control over interfering co-expressed cellular factors, as is also the case with yeast n-hybrid methods.

[0049] To provide a rapid and general method that circumvents many of the limitations discussed above, we determined that fragmented reporter proteins fused to functional (ligand-binding) proteins can be rapidly generated directly from mRNA utilizing cell-free translation systems and immediately interrogated for biomolecular interaction-dependent signal generation. The use of split reporter proteins in cell-free translation expression systems takes advantage of fast protein synthesis rates, from 60 to 90 minutes, and easy adaptation to homogeneous assays and high throughput analyses; these systems also avoid immobilization, washing and/or purification protocols. This cell-free approach provides a general platform for rapidly detecting protein-protein, protein-small molecule, protein-DNA, protein-methylated DNA, and protein-RNA interactions starting from mRNA or directly from DNA corresponding to the desired interaction pair in less than two hours. Moreover, we demonstrate how this approach aids in determining specificity of protein-nucleic acid interactions as well as in determining small molecule antagonists and/or agonists of protein-protein interactions.

**Evaluation of Split Protein Reporters for the Cell-Free Interrogation of Biomolecular Interactions.**

[0050] Initially we evaluated the ability of our previously reported split-GFP (22) and split-β-lactamase (23,25) systems appended to specific zinc-fingers to reassemble in the presence of target DNA utilizing *in vitro* transcribed mRNA in a purified wheat germ extract translation system (Figure 1b and 1c). Signal from the DNA-dependent reassembled GFP (22) in the early experiments was too low to observe over background using standard fluorescence measurements, while DNA dependent β-lactamase activity (25) yielded measurable but low signal-to-background ratios. Thus, we turned to *in vivo* split-luciferase systems which have
the significant advantage of negligible background from the translation system due to
the generation of a bioluminescent signal (Figure 1a). We first examined the
fragmented firefly luciferase (Fluc) reported by Luker et al (27) which when
 appended to our zinc fingers showed significant signal over background
 luminescence upon addition of target DNA (Figure 1d). This constitutes the first
demonstration of the bioluminescent read-out of a specific nucleic acid sequence,
and this split-luciferase system was chosen for further studies in cell-free systems.
The split-Gaussia luciferase (20) and alternatively split-firefly luciferase
complementation systems (27,28) were also tethered to our zinc-finger proteins and
displayed surprisingly strong signal over background bioluminescence (Figure 1e
and 1f) in a DNA-dependent fashion in cell-free assays.

[0051] To test the general applicability of the cell-free split-luciferase approach to
monitoring protein-protein and protein-nucleic acid interactions, we utilized seven
well characterized and widely studied biomolecular interactions (Figure 2) including
(a) the catalytic subunit of cAMP-dependent protein kinase (PKA) with its inhibitor
PKI (PKA/PKI) (29-31); (b) the rapamycin-dependent interaction between human
FK506-binding protein 12 (FKBP) and the FKBP12-rapamycin binding (FRB) domain
of human mTOR (FKBP/FRB) (32-34). Akin to the yeast three hybrid systems (35),
we also investigated the ternary association of (c) two sequence specific zinc-fingers
(36) with a target DNA (Zif268/PBSII); (d) a zinc-finger and methyl CpG-binding
domain with a target CpG-methylated DNA (Zif268/MBD2) (37-40); and finally (e)
two RNA-specific pumilio domains41 with a target RNA (Pum1/Pum2). Additionally,
the widely utilized coiled-coil domains of the transcription factors Fos and Jun
(Fos/Jun) (42-44) and the interaction between hypoxia inducible factor-1α (HIF-1α)
and the CH1 domain of the transcriptional coactivator p300 (HIF-1α/p300) (45,46)
were also tested.

[0052] The overall sensitivity (signal/background) of these systems (Figure 2,
panels a–e) was excellent and varied from 22 to 1800-fold, and the total assay time
from translation to analysis was less than two hours.

[0053] We also investigated whether it was possible to directly couple
transcription and translation in a cell-free lysate system, which could eliminate the
need for the separate in vitro transcription step as was used in these experiments
(Figure 3). These experiments were successful for both DNA-protein and protein-small molecule dependent interactions (Figure 3, panels a and b). In addition to using purified lysates, we further interrogated whether we could detect the above interactions using a system composed entirely of purified translational components (47) (Figure 3, panels c and e, the so-called “PURE System”), to demonstrate that reconstituted transcription and translation machinery is sufficient for detecting biomolecular interactions after split reporter protein expression. This set of experiments clearly demonstrated that a cell-free split-luciferase assay format allows the rapid, sensitive, and direct detection of protein-protein, protein-small molecule, protein-DNA, protein-methylated DNA, and protein-RNA interactions starting from either mRNA or directly from DNA corresponding to the desired interaction pair. Having established that our methods provide robust signals for a wide variety of biomolecular interactions, we also investigated whether this system is amenable to reporting upon inhibitors of protein-nucleic acid and protein-protein interactions.

[0054] In order to detect antagonists of protein-protein or protein nucleic acid interaction, we first established the thermodynamic reversibility of the ternary complex consisting of reassembled firefly luciferase fragments tethered to two zinc fingers (PBSII and Zif268) and the cognate target DNA (Figure 4a). Towards this goal, translations using mRNA encoding PBSII-NFluc and CFluc-Zif268 were initiated in the presence of the target oligonucleotide (Zif268-0-PBSII). Post DNA-dependent firefly luciferase reassembly, a hairpin DNA (hpDNA-Zif268) which is a competitor for only Zif268 binding was added at increasing concentrations and the system was allowed to equilibrate for 30 min. A concentration-dependent decrease in luminescence was observed, clearly demonstrating that the formation of the ternary complex of the firefly luciferase was reversible and it could be inhibited by addition of the dominant-negative (hpDNA-Zif268) oligonucleotide (Figure 4b, TGG containing hpDNA). The general applicability of the cell-free system for probing protein-nucleic acid inhibition was further demonstrated with translations containing mRNA encoding Pum2-NFluc and CFluc-Pum1 in the presence of target RNA. As earlier, a concentration-dependent decrease in luminescence was observed only upon the addition of increasing amounts of a competitor half-site RNA target that is known to selectively bind one of the pumilio domains (41) (Figure 4e). We recognized that our split-luciferase based cell-free system can be readily utilized to
analyze the relative target site specificity of nucleic acid binding proteins through competition binding experiments.

[0055] We studied protein-DNA target site specificity using the present methods. A number of methods have been developed to interrogate the relative affinity of DNA-binding proteins for their target site, including traditional EMSAs and DNA microarrays (48,49). Though powerful, these techniques require the use of purified components, specialized equipment, and/or radioactive materials. Having established that ternary zinc finger-DNA complexes can be disrupted by a competitor oligonucleotide added *in trans*, we next correlated the known binding affinities of Zif268 for single nucleotide changes in its binding site to IC$_{50}$ values obtained from our cell-free firefly luciferase approach in a 96-well format. Separate translation reactions of PBSII-NFluc and CFluc-Zif268 mRNA in the presence of the dsDNA target oligonucleotide Zif268-0-PBSII were initiated. Duplicate experiments containing ternary complexes were allowed to assemble for 90 minutes followed by the addition of increasing concentrations of competitor hpDNA, containing one of four different Zif268 binding sites having either A, T, C, or G at the central position. In each case a competitor hpDNA-concentration dependent decrease in luminescence was observed (Figure 4b) within 30 min. IC$_{50}$ values for each competitor hpDNA (Figure 4c) were shown to correlate extremely well ($R^2 = 0.996$) (Figure 4d) with previously reported relative affinities of these target sites (49). These results validate the use of the split-luciferase cell-free system for the determination of relative binding affinities of nucleic acid-binding proteins for their cognate target sites, and more generally, in studying inhibitors of protein-nucleic acid interactions. Thus, this cell-free system provides an advantageous alternative to current methods for interrogating protein-nucleic acid binding as they can be performed in a simple, rapid, high-throughput and homogeneous format without having to purify or refold the protein of interest and without having to transform and manipulate living cells.

[0056] Having demonstrated the ability to measure antagonists of protein-nucleic acid interactions, we next sought to interrogate the ability of cell-free firefly luciferase reassembly to assess antagonists and agonists (for example, small molecules) of protein-protein interactions. As a first test of small molecule modulation of split-
luciferase activity, we chose the well characterized rapamycin-dependent interaction between the human FK506-binding protein 12 (FKBP) and the FKBP12-rapamycin binding (FRB) domain of human mTOR (residues 2024-2113) (27,28,32). This system has been a standard test for several split-protein reporter systems. A rapamycin concentration dependent increase in luminescence was observed from the cell-free translations of the split reporters (Figure 5, panel a) (27).

[0057] As our first test for determining antagonism of protein-protein interactions, we chose the well characterized interaction between the catalytic subunit of cAMP-dependent protein kinase (PKA) with its inhibitor PKI (residues 5-24) (29). Initial experiments had demonstrated that the fusion proteins PKI-NF\text{Luc} and CFLuc-PKA could be translated in vitro from mRNA and that their association could be monitored via luminescence (Figure 2, panel a). Given that reassembly of fragmented firefly luciferase is dependent on PKA/PKI complex formation, the inhibition of this interaction was interrogated by the addition of increasing concentrations of a PKI peptide (44), with the result of an observed IC_{50} value of 11 nM (Figure 5, panel b). To provide further evidence of the general applicability of this cell-free format, we interrogated the inhibition of the interaction of hypoxia inducible factor-1α (HIF-1α) and the CH1 domain of the transcriptional coactivator p300 (50). HIF-α is an emerging anti-cancer target. Initial experiments demonstrated that the fusion proteins p300-NF\text{Luc} and CFLuc-HIF-1α could be translated in vitro from mRNA and their association could be monitored via luminescence. The small molecule chetomin has been identified as a first-in class inhibitor of the interaction between HIF-1α and p300 (46). To evaluate if our method can aid in the identification of small-molecule inhibitors, mRNA encoding p300-NF\text{Luc} and CFLuc-HIF-1α fusion proteins were translated followed by the addition of increasing concentrations of the chetomin post reassembly. Luminescence measurements following chetomin incubation revealed a concentration-dependent decrease in signal, with an IC_{50} value of 290 nM (Figure 5, panel c). Importantly, control experiments with excess chetomin or PKI-peptide showed no effect on signal generation in the irrelevant zinc finger/DNA cell-free assay, which verified that the loss in signal was dependent on the disruption of specific protein-protein interactions rather than on off-target effects such as inhibition of luciferase activity. Thus, these three systems demonstrate that both peptide and small molecule modulators of protein-protein interactions can be
rapidly evaluated in the cell-free split-luciferase system without the need for transfection and cell-culture (current yeast n-hybrids and PCA methods); protein purification, selective fluorophore labeling (FP); or immobilization on solid surfaces (SPR and ELISA).

[0058] Elegant experiments by Tawfik and Griffiths have recently established that water-in-oil emulsions of in vitro translation reactions can function as "artificial cells" or "man-made" cells, and they have been used for linking genotype with phenotype in protein evolution experiments (51). To demonstrate that our cell-free split luciferase assay is compatible with in vitro compartmentalization, translations containing mRNA encoding PKI-NFluc and CFuc-PKA; or PBSII-NFluc, CFuc-Zif268 and target DNA; were performed in wheat germ extract encapsulated in water-in-oil emulsions (Figure 6a) and subjected to luminescence analysis (Figure 6b and 6c). These experiments clearly demonstrate that the cell-free split-firefly luciferase system can be adapted in screens for protein-protein and protein-nucleic acid interactions in artificial cell based methodologies and may also find applications in studying proteins incorporating unnatural amino acids where significantly greater control over translational machinery components is often desirable (52-55).

[0059] Previous cell based and in vitro strategies, though powerful, depend on cumbersome steps which can include transfection, cell culture, purification, washing steps, and/or covalent modification, yielding overall experimental times in excess of 12-48 hours starting from appropriate clones. By contrast this present methods provide a general platform for interrogating biomolecular interaction in homogeneous assays based on cell-free split-protein systems, and results are available within two hours. This cell-free assay can employ a variety of split protein reporters to provide fluorescent (β-lactamase) or bioluminescent (luciferase) signal outputs. One drawback of the previous methods, also found in ELISA and in vivo approaches, is that only relative affinities and IC50 values can be determined, as compared to methods such as FP and SPR. However, the speed and ease of implementation of the present methods, which does not require cell culture, protein purification, or chemical derivatization, can be used to rapidly address biological and chemical questions with appropriate controls, as we have demonstrated with either dominant negative inhibition or known small molecule ligand.
[0060] We have demonstrated the ability to detect a wide variety of protein-ligand interactions, including the well studied heterodimerization of the leucine zippers Fos and Jun, the interaction between the protein kinase PKA and its inhibitor PKI, and the small molecule-dependent interaction between FKBP and FRB. Additionally, we provide the first example for a rapid method for interrogating the interaction between HIF-1α and p300, an emerging protein-protein target implicated in cancer progression. Furthermore, we detail the first examples of sensitive split-luciferase-mediated detection of a wide range of protein-nucleic acid interactions, including zinc finger domains with specific dsDNA, a methyl CpG-binding domain with specific methylated DNA, and RNA binding pumilio domains with target RNA. We have also shown that this methodology can be used to interrogate the relative binding affinities of nucleic-acid binding proteins for their target sites and the evaluation of small molecule, peptide and nucleic acid modulators of protein-protein interactions. In addition to using purified lysates and wheat germ extracts cell-free translation systems, we have demonstrated the detection of protein-protein and protein-nucleic acid using a system composed entirely of purified components that minimize non-specific interactions from cellular components and allow control over the translational machinery, for example in applications including, but not limited to, unnatural amino acid incorporation (52-55). Finally, we have demonstrated that the split-protein reporters are functional in water-in-oil emulsions providing artificial cell systems for studying protein-protein and protein-nucleic acid interactions that can potentially be utilized in screening methodologies (51).

[0061] This cell-free format can be used in a wide-variety of applications that include, without limitation, screening of DNA or RNA target sites for nucleic acid-binding proteins and the determination of target site preference. Importantly, this approach can also be used for screening small molecules, nucleic acids, peptides or proteins for inhibition of specific protein-protein or protein-nucleic acid or protein-small molecule interactions (1,2) Moreover, since the split-protein approaches have been widely utilized in a cellular context, the initial hits from the rapid cell-free system can be rapidly applied in a cellular context (10,19). This rapid, sensitive, and homogeneous assay system can be widely utilized for interrogating user-defined natural and unnatural biomolecular interactions and for evaluating agonists and antagonists of these interactions.
[0062] The invention may be further understood by the following non-limiting examples and information provided in the present Specification.

General Materials and Methods

[0063] All materials were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. ZnCl₂ was obtained from EM Sciences (Pt. Washington, PA). Restriction enzymes were obtained from NEB (Ipswich, MA) and in vitro translational products from Promega (Madison, WI). Oligonucleotide primers and targets were from IDT (Coralville, IA).

[0064] **Plasmid construction and mRNA production.** The fusion protein constructs used in this study are shown in Tables 6-28. Fragments coding for reporter protein fragments (GFP, Beta-Lactamase, and three luciferases) were generated by PCR with appropriate primers and subsequently cloned into either the pETDuet-1 vector (Novagen, Madison, WA) or the pMAL-c2x vector (NEB) using standard techniques and verified by sequencing. Fragments encoding the nucleic acid-binding proteins or associating proteins used in this study were generated by PCR starting from the specific plasmids. The fusion protein constructs were generated using standard cloning techniques and verified by sequencing. The mRNA necessary for cell-free assays was generated as follows: PCR fragments corresponding to the desired fusion constructs were generated using a forward primer containing a T7 RNA polymerase promoter and Kozak sequence and a reverse primer containing a 3’ hairpin loop. The purified PCR products were subsequently used as template for in vitro transcription using the RiboMAX Large Scale RNA Production System-T7 (Promega) following the manufacturer's protocols.

[0065] **Target DNA preparation.** All nucleic acid targets were obtained from IDT. A dsDNA target containing a zero base pair separation between the Zif268 and PBSII zinc finger sites (Z0P) was annealed as previously described. Hairpin DNA targets were annealed in 1x BamHI buffer by heating at 95°C for 7 minutes followed immediately by cooling on ice.

[0066] **Reassembly of the GFP variant Venus.** Duplicate 150 μL translations were carried out in Wheat Germ Plus extracts (Promega) according to the manufacturer's protocol using 4 pmol of each mRNA encoding for NVenus(residues
1-157)-Zif268 and PBSII-CVenus(residues 158-238), 10 μM ZnCl₂, 0.5 μL of Rnasin™ Plus (Promega), and either 50 nM Z0P target DNA or no DNA. Translations were incubated at 25 °C for 2 hours (no fluorescence was observed) or alternatively interrogated for fluorescence followed by a 20 hour incubation at room temperature. Fluorescence spectra were acquired by exiting at 515 nm and monitoring emission at 528 nm.

[0067] **Reassembly of split β-lactamase-zinc finger fusions.** Four duplicate 25 μL translations were carried out in wheat germ plus extracts (Promega) according to the manufacturer’s protocol using 0.5 pmol of each mRNA encoding for NβLac(residues 26-196)-Zif268 and PBSII-CβLac(residues 198-290), 10 μM ZnCl₂, 0.5 μL of Rnasin™ Plus (Promega), and either 20 nM Z0P target DNA or no DNA. Translations were incubated at 25 °C for 2 hours and assayed by adding 25 μL of translation to 75 μL of PBS containing a final concentration of 10 μM Fluorocillin Green soluble β-Lactamase substrate (Invitrogen, Carlsbad, CA). The final concentration of DNA in the assay was 5 nM. The rate of Fluorocillin Green hydrolysis was determined by exciting at 495 nm and monitoring emission at 525 nm with a 515 nm emission cutoff using a SPECTRAMAX™ Gemini plate reader (Molecular Devices, Sunnyvale, CA). Emission was read every 30 seconds for 10 minutes.

[0068] **Reassembly of split firefly luciferase.** Duplicate 25 μL translations were carried out in Flexi-Rabbit Reticulocyte Lysates (Promega) according to the manufacturer’s protocol using 2 pmol of each mRNA encoding for either PBSII-NFluc (residues 2-416) and CFLuc-Zif268 (residues 398-550) or NFluc(residues 2-398)-Zif268 and PBSII-CFluc(residues 394-550), 10 μM ZnCl₂, 0.5 μL of Rnasin™ Plus (Promega), and either 25 nM Z0P target DNA or no DNA. Translations were incubated at 30 °C for 90 minutes and assayed by adding 20 μL of translation to 80 μL of STEADY-GLO™ Luciferase Assay System (Promega). The final concentration of DNA in the assay was 5 nM. Light emission was monitored 1 minute after STEADY-GLO™ substrate addition using a Turner TD-20e luminometer (Turner Designs, Inc, Sunnyvale, CA) with a 3 second delay and a 10 second integration time.
[0069] **Initial cell-free assays.** Duplicate 25 μL translations were carried out in Flexi-Rabbit Reticulocyte Lysates (Promega) according to the manufacturer’s protocol using 2 pmol of each mRNA encoding the fusion proteins being analyzed, and 0.5 μL of RNasin™ Plus (Promega). For translations containing zinc finger proteins 10 μM ZnCl₂ was also added to the translation mixture. Translations were incubated at 30 °C for 90 min and assayed by adding 20 μL of translation mix to 80 μL of STEADY-GLO™ Luciferase Assay System (Promega). In the case of nucleic acid-binding proteins target oligonucleotides were either present or absent during translation. For the rapamycin induced interaction between FRB and FKBP either 5 nM rapamycin or control, DMSO, was added after translation followed by a 30 minute incubation at room temperature. Light emission was monitored 1 minute after STEADY-GLO™ substrate addition using a Turner TD-20e luminometer with a 3 second delay and a 10 second integration time.

[0070] **Reassembly of split firefly luciferase in a coupled transcription/translation system.** Coupled transcription/translation reactions were carried out in TNT T7 Coupled Rabbit Reticulocyte Lysates (Promega) according to the manufacturer’s protocol. Coupled reactions using split firefly luciferase-zinc finger fusions contained 0.5 pmols of each DNA encoding PBSII-NFLuc and CFluc-Zif268, 10 μM ZnCl₂, 1 μL of RNasin™ Plus™ (Promega), and either 100 nM Z0P target DNA or no DNA in a total of 25 μL. Coupled reactions using split firefly luciferase-FKBP and FRBP fusions contained 0.5 pmol of DNA encoding FRB-NFLuc and CFluc-FKBP, and 1 μL of RNasin™ Plus (Promega) in a total of 25 μL. Solutions were incubated at 30 °C for 90 minutes. Reactions were diluted at a 1:4 ratio into PBS containing 1% BSA (1% BSA and either 25 nM rapamycin or DMSO in the case of FRB/FKBP) and equilibrated at room temperature for 30 minutes. Samples were assayed for luciferase activity by mixing 20 μL of lysate with 80 μL of STEADY-GLO™ Luciferase Assay System (Promega). Luminescence readings were taken on a Turner TD20e luminometer using a 3 second delay and 10 second integrations, the average of replicate experiments is shown. The final concentration of Z0P or rapamycin in the assay was 5 nM.

[0071] **Reassembly of split firefly luciferase in a purified transcription/translation system.** Coupled transcription/translation reactions were
carried out using the PURESYSTEM classic II system (Post Genome Inst. Co. Ltd., Tokyo, JP; available from NEB) according to the manufacturer’s protocol. Coupled reactions using split firefly luciferase-zinc finger fusions contained 0.5 pmols of each DNA encoding PBSII-NFluc and CF1uc-Zif268, 10 μM ZnCl₂, 1 μL of RNasin™ Plus (Promega), and either 100 nM Z0P target DNA or no DNA in a total of 25 μL. Coupled reactions using split firefly luciferase-FKBP and FRBP fusions contained 0.5 pmol of DNA encoding FRB-NFluc and CF1uc-FKBP, and 1 μL of RNasin™ Plus (Promega) in a total of 25 μL. Solutions were incubated at 37 °C for 60 minutes, followed by the addition of either 25 nM rapamycin or DMSO in the case of FKBP/FRB. Samples were assayed for luciferase activity by mixing 20 μL of lysate with 80 μL of STEADY-GLO™ Luciferase Assay System (Promega). Luminescence readings were taken on a Turner TD20e luminometer using a 3 second delay and 10 second integrations, the average of replicate experiments is shown. The final concentration of Z0P or rapamycin in the assay was 5 nM.

[0072] **Competition assay to identify protein-DNA target site specificity.**
Duplicate 25 μL translations were carried out in Rabbit Reticulocyte Lysates (Promega) according to the manufacturer’s protocol using 0.05 pmols of mRNA encoding PBSII-NFluc and CF1uc-Zif268, 10 μM ZnCl₂, and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 750 pM Z0P dsDNA target. Following translation and firefly luciferase reassembly, increasing concentrations of each Zif268 hairpin DNA being tested were added followed by a 30 minute incubation at room temperature. Light emission was monitored 1 minute after STEADY-GLO™ addition using a Wallac 1420 VICTOR 3™ V luminometer with a 1 second integration time.

[0073] **Dissociation of the reassembled Pum2-NFluc, CF1uc-Pum1, RNA ternary complex.** Duplicate 25 μL translation reactions were carried out in Rabbit Reticulocyte Lysates (Promega) according to the manufacturer’s protocol using 0.1 pmols of mRNA encoding Pum2-NFluc and CF1uc-Pum1 and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 2.5 nM RNA oligonucleotide target. Following translation and firefly luciferase reassembly, increasing concentrations of a competitor RNA oligonucleotide were added followed by a 30 minute incubation at room temperature. Light emission was
monitored 1 minute after STEADY-GLO™ addition using a Turner TD-20e luminometer with a 3 second delay and a 10 second integration time.

[0074] Detection of small-molecule and peptide modulators of protein-protein interactions. Duplicate 25 μL translation reactions were carried out in Rabbit Reticulocyte Lysates (Promega) according to the manufacturer's protocol using 2 pmols of mRNA encoding either FRB-NFluc and CFluc-FKBp, PKI-NFluc and CFluc-PKA, or p300-NFluc and CFluc-HIF-1α and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30°C in. For analysis of p300/HIF-1α interaction 10 μM ZnCl₂ was added to the translation reaction. Following translation lysates were diluted 1:4 with PBS containing 1% BSA followed by the addition of increasing concentrations of either rapamycin in DMSO, PKI, or chetomin in DMSO followed by a 30 minute incubation at room temperature. Light emission was obtained by adding 20 μL of the translation solution to 80 μL of STEADY-GLO™ Luciferase Assay System (Promega). Light emission was monitored 1 minute after STEADY-GLO™ addition using a Wallac 1420 VICTOR 3™ V luminometer (PerkinElmer, Waltham, MA) with a 1 second integration time.

[0075] Reassembly of split firefly luciferase in water-in-oil emulsions. Water-in-oil emulsions were prepared in 2 mL round bottom cryogenic vials by adding 50 μL of aqueous phase, over 2 minutes, into 950 μL of mineral oil containing 4.5% Span 80 and 0.5% Tween 80 while stirring at 1,150 rpm using a 2 x 9 mm stir bar. Stirring was continued for one minute after the complete addition of the aqueous phase. Translations were prepared on ice using Wheat Germ Plus extracts (Promega) according the manufacturer's protocol using 4 pmol of each mRNA encoding either PKI-NFluc, CFluc-PKA, or both and 0.5 μL of RNasin™ Plus (Promega). For DNA dependent reassembly, 4 pmol of each mRNA encoding for CFluc-Zif268 and P52II-NFluc, 10 μM ZnCl₂, 0.5 μL of RNasin™ Plus (Promega), and either 25 nM Z0P target DNA or no DNA in a total of 50 μL. Emulsions were prepared using the ice-cold translation as the aqueous phase. Emulsions were incubated at 25 °C for 2 hours and assayed by adding 20 μL of emulsion to 80 μL of STEADY-GLO™ (Promega). Luminescence readings were taken on a Turner TD20e luminometer using a 3 second delay and a 10 second integration time.
Antibody Enabled Cell-Free Split-Luciferase Detection Systems

[0076] The present methods provide a robust solution phase split-luciferase assay that can, directly and sensitively, detect a protein or other molecule of interest, including but not limited to clinically relevant extracellular growth factors, such as VEGF; distinguish HIV-1 clades based on gp120-antibody specificities; and, record the abundance of cell-surface markers, such as HER2, without chemical derivatization, microscopy, or FACS.

[0077] High-resolution methods for imaging extracellular proteins often rely on laborious transfection and/or chemical derivatization for selective labeling. (56,57) In addition analysis is generally performed using expensive microscopy or Fluorescence-Activated Cell Sorting (FACS) instrumentation employing complex deconvolution algorithms. Alternatively, the simple Enzyme-Linked Immunosorbent Assay (ELISA) can be used to detect almost any analyte, but it requires that either the antigen or antibody be captured on a solid support prior to detection followed by vigorous washing and subsequent detection by an enzyme-secondary-antibody conjugate. This limits the utility of the ELISA for the direct detection of analytes in complex heterogeneous mixtures or biological samples such as blood and lysates. Development of a one-step solution phase sandwich assay in which the activity of an attached split-luciferase reporter would be dependent on the recognition of an extracellular protein or other molecule of interest is needed (Fig. 7a). Such methodology would allow for the direct detection of any protein in complex environments without the need for immobilization, chemical derivatization or microscopy/FACS analysis.

[0078] Herein, we provide a cell-free split-luciferase assay in which the luminescence of fragmented luciferase (see also 27) fused to interacting proteins, provides a direct measure of heterodimeric protein-protein interactions (58). This rapid method takes advantage of the in situ production of signaling proteins from mRNA in in vitro translation reaction and eliminates the need for laborious cell culture or protein purification steps (10). This cell-free split-luciferase methodology provides a general solution for the rapid and direct detection of an important and previously unaddressed class of clinically relevant proteins that include growth
factors, cell- and viral-surface receptors, and it is applicable to other proteins, carbohydrates, glycoproteins and other antigens or epitopes of interest.

[0079] We first confirmed that a dimeric receptor fragment could be used to detect its extracellular ligand. We chose as a model extracellular ligand Vascular Endothelial Growth Factor (VEGF), which is implicated in tumor angiogenesis and which binds its extracellular receptor Flt-1 in a 2:1 stoichiometry (59). With this in mind, we attached the N- and C-terminal halves of luciferase (residues 2-416 and 398-550, respectively) to separate Flt-1 domain 2 fragments (Fig. 7b), with the expectation that a statistical distribution of Flt-1-luciferase halves would still lead to ~50% of split-luciferase complementation. Expression of the split luciferase-Flt-1 fusion proteins in rabbit reticulocyte lysates leads to an increase in luminescence of >15-fold only in the presence of 50 nM VEGF dimer, clearly demonstrating the ability of dimeric receptor fragments to bind their ligands and mediate split reporter reassembly in this system (Fig. 7c). Thus, this approach conceptually allows for an expedient and general method for targeting a wide-variety of dimeric growth factors and their receptors through ternary complexation. It is also understood that present methods are adaptable to virtually all epitopes or antigens of interest, especially biomolecules.

VEGF-Flt-1 Sandwich Assay

[0080] Flexi-Rabbit Reticulocyte Lysate, RNasin™, Steady-Glo™ Luciferase Assay System and the T7 Ribomax Transcription Kit were purchased from Promega. G50 ProbeQuant™ (Pharmacia Biotech AB Corporation, Uppsala, Sweden) columns were obtained from GE Healthcare. XL1-Blue™ E. coli cells were purchased from Stratagene (La Jolla, CA). Ni-NTA agarose resin was purchased from Qiagen (Valencia, CA). All other reagents were obtained through Research Products International (Mt. Prospect, IL).

[0081] A pQE30-VEGF expression plasmid was transformed into XL1-Blue™ E. coli by electroporation according to the manufacturer’s instructions. An overnight culture of these cells was used to inoculate a 1 liter culture of 2xYT media supplemented with 100 µg/mL ampicillin at an initial OD600 of 0.05. Protein expression was induced at an OD600 of 0.8 with 1 mM IPTG. Protein expression was allowed to proceed overnight at 37 °C. Cells were pelleted by centrifugation and
resuspended in lysis buffer (Tris-HCl at pH = 8 containing 8 M Urea). Resuspended cells were lysed by sonication. The lysate was cleared by centrifugation at 18,000 rcf for 30 min. His-tagged VEGF was purified under denaturing conditions using Ni-NTA resin using the manufacturer’s instructions. Imidazole wash fractions were collected, pooled, and stored at -20 °C until use. Collected fractions were thawed on ice, concentrated and FPLC purified using a preparative Hi-Load 16/60 Superdex™ 75 (Pharmacia Biotech AB Corporation) column equilibrated with denaturing buffer (Tris-HCl at pH = 8 containing 6 M Urea). Full length monomeric VEGF was isolated, pooled, and stored at -20 °C until required for refolding.

[0082] The pooled fractions containing full-length monomeric VEGF were diluted to 50 μg/mL with buffer containing 6M Urea, 0.1 M Na₂HPO₄, 10 mM Tris-HCl at pH = 8.5, 1 mM EDTA, and 20 mM DTT. This solution was incubated for 3 hrs at room temperature to facilitate reduction. Reduced monomeric VEGF was then dialyzed against 100 mM Tris-HCl at pH = 8.5, 5 mM cysteine, 1 mM cystine, 0.5 M Urea, and 2 mM EDTA overnight at room temperature.

[0083] To separate dimeric VEGF from monomeric and multimeric species the refolded VEGF was concentrated and purified by FPLC using a Superdex™ 75 column equilibrated with PBS. Fractions containing refolded dimeric VEGF were collected, pooled, concentrated, and reapplied to the Superdex™ 75 column. Refolded VEGF was characterized by SDS-PAGE under reducing and non-reducing conditions to visualize the monomeric versus dimeric form. Protein concentrations were calculated based on UV absorbance.

[0084] To produce Flt-1 Luciferase Fusion mRNA, open reading frames encoding domain 2 of the Flt-1 receptor were cloned into bacterial vectors containing either the N- or C-terminal portions of firefly luciferase, residues 2-416 and 398-550 respectively (27), separated by a flexible amino acid linker. These plasmid sequences were confirmed by the sequencing. These constructs were PCR amplified using a 5’ primer encoding a T7 promoter and Kozak sequence and 3’ primer containing a stem loop. mRNA was generated using the T7 Ribomax Transcription Kit and purified using a G50 ProbeQuant column. Concentrations of each mRNA were determined from UV absorbance.
To perform the VEGF-Flt-1 sandwich assay, translations using Flexi-Rabbit Reticulocyte Lysate were carried out according to the manufactures procedure using 2 pmols of each mRNA encoding for the Flt-1 fusions, 0.5 \( \mu \)L RNaSin\textsuperscript{TM}, 70 mM KCl, 200 \( \mu \)M of each amino acid, 66% Lysate, and either 500 nM VEGF monomer or an equivalent volume of PBS in a 25 \( \mu \)L reaction. Reactions were incubated at 30 \( ^\circ \)C for 90 min after which luminescence was monitored on a Turner TD20e luminometer by mixing 20 \( \mu \)L of translation with 80 \( \mu \)L of Steady-Glo\textsuperscript{TM} Luciferase Assay System giving a final concentration of 100 nm VEGF monomer. Luminescence was monitored 1 min after mixing with a 10 sec integration. Reactions were performed in duplicate and averaged.

We envisioned that the cell-free split-luciferase sandwich assay could rapidly and sensitively detect and categorize HIV-1 clades based on antibody specificities. Accordingly, we turned our attention towards the interaction of CD4 with the gp120 glycoprotein from HIV-1, which leads to infection of susceptible T-lymphocytes by HIV-1. The crystal structure of the complex between CD4, gp120, and the Fab portion of a neutralizing antibody 17b (60), served as a model for the development of our gp120 sandwich assay (Fig. 8a). We fused domains 1 and 2 (D1D2, residues 1-182) of CD4, which have been shown to bind to gp120 with a \( K_d \) of \( \sim 3 \) nM (61), to the N-terminal half of luciferase. As our second recognition element we fused the C-terminal half of luciferase to the 17b single-chain antibody (scFv), which binds a CD4-induced epitope of gp120 (Fig. 8a) (60). Initial experiments showed a negligible increase in luminescence in the presence of 20 nM \( B_{al} \) gp120. Importantly, the elimination of DTT and addition of protein disulfide isomerase (PDI) allowed for luminescence and the first functional demonstration of antibody mediated targeting in the split-luciferase system (Fig 8b Inset). Having established conditions for favorable protein folding, we sought to verify the specificity of our gp120 assay. Accordingly, we first investigated luciferase reassembly in the presence of different \( B_{al} \) gp120s containing single amino acid mutations, D368R and I420R, known to reduce CD4 (62) or 17b (63) binding respectively. Indeed these mutant gp120s considerably reduce luminescence relative to the wild-type, confirming that both functional CD4 and 17b binding are required for luciferase reassembly (Fig. 8b). We also interrogated split-luciferase activity as a function of gp120 concentration; our assay system can reporting on the presence of as little as 12 ng/mL of \( B_{al} \) gp120, a
sensitivity comparable to commercially available gp120 ELISAs. This is likely a function of antibody/D1D2 affinities. The titration experiment also indicated that ~5 nM of active complex (folding capable split-halves) is translated under our current cell-free conditions, which is sufficient for most ELISA-like applications.

[0087] To use our sandwich assay for the rapid characterization of HIV-1 clades, we investigated gp120s from isolates CN54 and 96ZM651, both of which are clade C viruses. Maximal luciferase signal was observed only in the presence of B<sub>al</sub> gp120 (clade B) while a slight increase in luminescence was observed for CN54 gp120 and no detectable signal was generated for 96ZM651 gp120 (Fig. 8c). This highlights the potential utility of this rapid and inexpensive approach for rapidly distinguishing HIV-1 clades and sub-types using known antibody specificities.

gp120 Sandwich Assay

[0088] Flexi-Rabbit Reticulocyte Lysate, RNasin<sup>TM</sup>, Steady-Glo<sup>TM</sup> Luciferase Assay System, and the T7 Ribomax Transcription Kit were purchased from Promega. G50 ProbeQuant<sup>TM</sup> columns were obtained from GE Healthcare. B<sub>al</sub>, CN54, and 96ZM651 gp120s were obtained from the NIH AIDS Reference and Reagent Program, catalog numbers 4961, 7749, and 10080 respectively. Wild-type B<sub>al</sub> gp120 and the B<sub>al</sub> gp120 D368R and I420R mutants used in Figure 8, panel b (excluding the inset) were a generous gift of R. Wyatt. PDI was purchased from Sigma (St. Louis, MO). All other reagents were obtained through Research Products International (Mt. Prospect, IL).

[0089] mRNA Encoding for the split-luciferase fusions was prepared as follows. Open reading frames encoding for residues 1-182 of CD4 and the V<sub>H</sub> and V<sub>L</sub> regions of 17b separated by a (GGGGS)<sub>3</sub> linker (SEQ ID NO:91) were cloned into vectors containing the N- and C-terminal portions of luciferase respectively. These plasmid sequences were confirmed by DNA sequencing. These constructs were PCR amplified using a 5’ primer encoding a T7 promoter and Kozak sequence and 3’ primer containing a stem loop. mRNA was generated using the T7 Ribomax Transcription Kit and purified using a G50 ProbeQuant<sup>TM</sup> column. Concentrations of each mRNA were determined from UV absorbance.
[0090] To perform the initial gp120 sandwich assay, translations using Flexi-Rabbit Reticulocyte Lysate were carried out according to the manufactures procedure using 2 pmoles of each of the mRNAs encoding the CD4-NFLuc and CFluc-17b fusions, 0.5 μL RNasin™, 70 mM KCl, 200 μM of each amino acid, 66% Lysate, and either 100 nM B\textsubscript{al} gp120 or an equivalent volume of PBS in a 25 μL reaction. Reactions were incubated at 30 °C for 90 min after which luminescence was monitored on a Turner TD20e luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System giving a final concentration of 20 nm B\textsubscript{al} gp120. Reactions were performed in duplicate and averaged; significantly greater signal was generated with 20 nM gp120 than in its absence.

[0091] We sought to determine the effect of dithiothreitol (DTT) and PDI on the gp120 sandwich assay as this has previously been shown to increase scFv folding efficiency in cell-free translation systems (137). Translations using the Flexi-Rabbit Reticulocyte Lysate were carried out according to the manufactures procedure using 2 pmoles of each of the mRNAs encoding the CD4-NFLuc and CFluc-17b fusions, 70 mM KCl, 200 μM of each amino acid, 66% Lysate, 0.5 μL RNasin™ (160 μM DTT, from the RNasin™ storage buffer, during translation) were indicated, 90 μg/mL PDI were indicated, and either 100 nM B\textsubscript{al} gp120 or an equivalent volume of PBS in a 25 μL reaction. Reactions were incubated at 30 °C for 90 min after which luminescence was monitored on a Turner TD20e luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System giving a final concentration of 20 nm B\textsubscript{al} gp120. Luminescence was monitored 1 min after mixing with a 10 sec integration. Reactions were performed in duplicate and averaged.

**gp120 Titration**

[0092] Translations using Flexi-Rabbit Reticulocyte Lysate were carried out according to the manufacturer’s procedure using 2 pmoles of each of the mRNAs encoding the CD4-NFLuc and CFluc-17b fusions, 90 μg/mL PDI, 70 mM KCl, 200 μM of each amino acid, 66% Lysate, and decreasing concentrations of B\textsubscript{al} gp120 or an equivalent volume of PBS in a 25 μL reaction. Reactions were incubated at 30 °C for 90 min after which luminescence was monitored on a Turner TD20e luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System. Luminescence was monitored 1 min after mixing with a 10 sec integration.
Reactions were performed in duplicate, background subtracted (using samples containing no gp120), and averaged.

[0093] Luminescence from reassembled luciferase was monitored as a function of the concentration of gp120. Initially luciferase fusions are in excess however as the concentration of gp120 increases a maximum is reached were the concentration of luciferase fusions capable of forming a functional complex is equivalent to that of gp120. As the gp120 concentration is increased, further luminescence decreases due to localization of the luciferase fusions to different gp120s, rather than allowing reassembly of complementary fragments.

[0094] Translations and luciferase detection were carried out as described above except that 25 nM of the indicated gp120 was added during translation, giving a final concentration of 5 nM gp120.

[0095] Having identified suitable expression conditions for using scFvs in our split-luciferase system, we next established that the assay can be utilized to determine relative abundance of cell surface proteins, as specifically exemplified on human cells. For proof of concept, we chose the extracellular domain (ECD, residues 1-631) of HER2, which is over-expressed in ~30% of human breast cancers and is directly correlated with poor clinical outcomes; therefore, an expedient method to assess the relative amount of HER2 on the surface of human breast cancer cells would be of considerable utility. Two antibodies HERCEPTIN™ and Omnitarg (Genentech, South San Francisco, CA) bind distinct epitopes of the human HER2 ECD. Overlaying the crystal structures of these bound antibodies indicated that they are likely capable of binding HER2 simultaneously (64, 65). Moreover, the reported binding constants for a scFv version of HERCEPTIN™ and the Fab portion of Omnitarg for the HER2 ECD are 150 pM (66) and 8.5 nM (67), respectively, well within the detection limits of the present methods. Thus, we constructed mRNAs in which the scFv of Omnitarg was fused to the N-terminal portion of luciferase and the C-terminal portion of luciferase was fused to the scFv of HERCEPTIN™ (Fig. 9, panel a). As an initial test of the HER2 sandwich assay, the HER2 ECD was expressed, purified, and added at varying concentrations to the two tethered scFv fusions translated in rabbit reticulocyte lysates. A concentration dependent increase in luminescence in the presence of the HER2 ECD was observed (Fig. 9b),
indicating that this new antibody enabled sandwich assay was indeed capable of reporting on the presence of HER2 at sub-nanomolar levels in a complex mixture. Preliminary experiments indicate that these cell-free translations can be stored at least 7 days at -80 °C prior to the addition of HER2, potentially allowing for the long term storage of reagents and the detection of extracellular proteins within 30 min at the point of care.

[0096] Finally, we determined that the present assay could directly detect different relative expression levels of HER2 on the surface of human breast cancer cells. We chose the SK-BR-3 and MCF7 cell lines which have been shown to produce \( \sim 1 \times 10^5 \) and \( \sim 2.5 \times 10^4 \) copies of HER2 per cell, respectively (68). Translations were conducted as above, after which cells were added and the reactions were gently shaken at room temperature for 30 min. In the presence of SK-BR-3 cells (\( 1 \times 10^4 \) cells, theoretically \( \sim 170 \) pM HER2) an 18-fold increase in luminescence was observed with respect to the MCF7 cells (\( 1 \times 10^4 \) cells, theoretically \( \sim 4 \) pM HER2) indicating that our sandwich assay is capable of directly reporting on the relative amount of HER2 expression on the surface of human breast cancer cell lines (Fig. 9c). Additional experiments indicate that as few as 2,600 SK-BR-3 cells can be detected using this assay format. Importantly this split-luciferase sandwich assay allowed us to specifically label and visualize the HER2 cancer-specific antigen on the surface of human breast cancer cells without the need for FACS analysis (68).

**HER2 Sandwich Assay**

[0097] Flexi-Rabbit Reticulocyte Lysate, Steady-Glo\textsuperscript{TM} Luciferase Assay System, and the T7 Ribomax Transcription Kit were purchased from Promega. G50 ProbeQuant\textsuperscript{TM} columns were obtained from GE Healthcare. SK-BR-3 cells were obtained from the American Type Culture Collection (Manasas, VA, ATCC) (HTB-30). MCF7 cells were a generous gift of the B. Olenyuk laboratory. Cell culture media and reagents were purchased from HyClone (Logan, UT). Plasmids encoding the \( V_H \) and \( V_L \) regions of both Herceptin (138) and Omnitarg (67) separated by a \((GGGS)_3\) linker (SEQ ID NO:91) were purchased from Bio Basic (Markham, Ontario, CA). G418 and methotrexate were purchased from Research Products International. PDI, the Ala-Gln dipeptide, and Trypan Blue were purchased from
Sigma. Ni-NTA agarose resin was purchased from Qiagen. All other reagents were obtained through Research Products International.

[0098] To produce mRNA Encoding for the split-luciferase fusions, open reading frames encoding for the \( V_H \) and \( V_L \) regions of Omnitarg and Herceptin separated by a \((GGGGS)_3\) linker (SEQ ID NO:91) were cloned into vectors containing the N- and C-terminal portions of luciferase respectively. These plasmid sequences were confirmed by DNA sequencing. These constructs were PCR amplified using a 5' primer encoding a T7 promoter and Kozak sequence and 3' primer containing a stem loop. mRNA was generated using the T7 Ribomax Transcription Kit and purified using a G50 ProbeQuant™ column. Concentrations of each mRNA were determined based on UV absorbance.

[0099] Expression, purification, and western blot analysis of the HER2 ECD were carried out as follows. Lec1 cells stably expressing a human growth hormone-histidine tagged-HER2 ECD protein (139) were grown in \( \alpha \)MEM (without nucleotides or L-Gln) 95% and FBS 5% supplemented with 100 nM methotrexate, 0.5 mg/mL G418, 584 mg/L Ala-Gln, 100 units/mL penicillin, and 100 \( \mu \)g/mL streptomycin. Cell cultures were allowed to grow for three days, after which protein was purified from 50 mLs of culture media using Ni-NTA affinity chromatography. Protein was eluted with 10 mM Tris-HCl at pH = 7.5, 50 mM NaCl, and 500 mM Imidazole. This solution was used directly for the experiments described below.

[0100] Western blot analyses were performed using a rabbit anti-His-tag polyclonal primary antibody (QED Biosciences, San Diego, CA, 18814) and an IR dye conjugated anti-rabbit secondary goat antibody (Li-Cor Biosciences, Lincoln, NE, IgG IRDye 800CW, 926-32211). A Li-Cor Biosciences Odyssey scanner was used for imaging. HER2 ECD concentration was estimated from SDS-PAGE analysis.

[0101] The HER2 sandwich assay using purified HER2 ECD was carried out as follows. Translations using Flexi-Rabbit Reticulocyte Lysate were carried out according to the manufacturer’s procedure using 2 pmols of each of the mRNAs encoding the Omnitarg-NFluc and CFluc-Herceptin fusions, 70 mM KCl, 200 \( \mu \)M of each amino acid, 66% Lysate, and 90 \( \mu \)g/mL PDI in a 25 \( \mu \)L reaction. Reactions were incubated at 30 °C for 90 min after which purified HER2 ECD or an equivalent
volume of storage buffer (10 mM Tris-HCl at pH = 7.5, 50 mM NaCl, and 500 mM Imidazole) was added to the translation. These solutions were allowed to equilibrate at room temperature for 30 min. Luminescence was monitored on a Turner TD20e luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System. Luminescence was monitored 1 min after mixing with a 10 sec integration. Reactions were performed in duplicate, background subtracted (using samples containing no HER2 ECD), and averaged. HER2 ECD concentrations after rapid dilution are shown in Fig. 9, panel b.

[00102] The HER2 sandwich assay was carried out after storage at -80 °C as follows. Two 25 μL translations were carried out as described above. Reactions were incubated at 30 °C for 90 min, flash frozen, and stored at -80 °C for 7 days. Solutions were thawed and purified HER2 ECD or an equivalent volume of storage buffer (10 mM Tris-HCl at pH = 7.5, 50 mM NaCl, and 500 mM Imidazole) was added to the vials. These solutions were allowed to equilibrate at room temperature for 30 min. Luminescence was monitored on a Turner 20/20° luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System. Luminescence was monitored 1 min after mixing with a 10 sec integration.

[00103] Cell-free translations of the HER2 sandwich assay were flash frozen and stored at -80 °C for 7 days; after which purified HER2 ECD was added and luminescence was monitored. HER2-dependent luciferase activity is still observed after storage for 7 days at -80 °C.

[00104] The HER2 sandwich assay using human breast cancer cells was carried out as follows. SK-BR-3 and MCF7 cells were grown in RPMI 1640 90% and FBS 10% supplemented with 100 units/mL penicillin and 100 μg/mL streptomycin. Cells were detached using PBS containing 25 mM EDTA, washed, and resuspended in PBS; after which they were counted by Trypan Blue exclusion.

[00105] Luminescence as observed from the HER2 sandwich assay performed on 2,600 cells from the indicated human breast cancer cell lines. The relative luminescence was 1 for the SK-BR-3 cells, while the same number of MCF7 cells resulted in less a relative luminescence of less than 0.05.
Translations using Flexi-Rabbit Reticulocyte Lysate were carried out as above. Reactions were incubated at 30 °C for 90 min after which cells or an equivalent volume of PBS was added. These solutions were allowed to equilibrate at room temperature for 30 min with gentle shaking. Luminescence was monitored on a Turner 20/20° or TD20e luminometer by mixing 20 μL of translation with 80 μL of Steady-Glo™ Luciferase Assay System. Luminescence was monitored 1 min after mixing with a 10 sec integration. Reactions were performed in duplicate, background subtracted (using samples containing no cells), and averaged. The number of cells in the luminescence assay is reported.

Herein we have described a split-luciferase sandwich ELISA-like assay for the rapid analysis of proteins and receptors in complex mixtures. (70,71,68). The simplicity of generating reagents, low cost of instrumentation, sensitive bioluminescent read-out, and most importantly, the generality of scFv mediated targeting allow this method to rapidly detect virtually any target protein in complex heterogeneous systems, and thus, have utility in point of care diagnostics.

Amyloid β-peptide and related assemblies have been studied with new probes and split-protein reporters. The amyloid hypothesis is that specific proteins and peptides misfold and aggregate in a nucleation dependent manner to form fibrils with a characteristic cross-β pattern. Though much insight has been gained regarding the final fibrillar state of amyloidogenic peptides and proteins from both NMR and X-ray crystallography, far less is known regarding the multistep process involving the transition of monomers to metastable oligomers and their further assembly into mature fibrils (Fig. 11, panel a and Fig. 12). We have focused our attention on the amyloid β-peptide (Ab) implicated in Alzheimer's disease, perhaps the most studied among over 30 known protein misfolding disorders, which include Parkinson's, dialysis related amyloidosis (β2-microglobulin), Huntington's disease, and prion diseases.

Ligands identified to block the Aβ aggregation process include certain aromatic small molecules, antibodies, chaperones and synthetic peptides and peptidomimetics derived from the amyloidogenic core of the parent protein. Synthetic peptides used to block fibrillization have been further modified by incorporation of proline residues, α-α disubstitution and N-methylation of amino
acids and addition of charged residues at the N and C termini to increase their potency by either disrupting H-bonding or sterically hinder the self assembly process.

[00110] Inhibitors of aggregation have emerged as structural and mechanistic probes to explore two issues that provide useful insight into the aggregation mechanism of amyloid proteins. Recent studies have shown these ligands to bind and stabilize transient intermediates that can be useful in therapeutic and/or preventative strategies. For example, affibody Z_Aβ3 binds to Aβ40 with nanomolar affinity and was shown by 2D-NMR studies to fold and stabilize Aβ in a beta-hairpin structure, while transthyretin, a 55 kD homotetramer present in the cerebral fluid, prevented Aβ40 fibrillization by suppressing the growth of pre-existing aggregates. Chaperones (Hsp 70/40 and Hsp90), in a similar study, appeared to bind early spheroid like intermediates of Aβ42 to prevent its further self-assembly. Recently, Glabe and coworkers detected a dodecameric intermediate of Aβ isolated from transgenic mice using the polyclonal antibody A11. Secondly, the common pathological features shared by fibrils of different proteins and the detection of common soluble oligomers by the polyclonal antibody A11 (generated against micellar Aβ by Glabe and co-workers) have led to hypothesis that different proteins follow common fibril formation pathways. However, photo-crosslinking of Aβ40 and Aβ42 by Bitan et al entrapped distinct early intermediates for these two proteins.

[00111] Drawing in part from these strategies, we designed (by directed evolution) a beta-sheet mini protein (TJ10), which inhibits the aggregation of Aβ40. The β-sheet scaffold chosen for this purpose was a 56-residue hyperthermophilic IgG binding protein redesigned by Malakauskas and Mayo, termed HTB1. Eight positions on adjacent strands of HTB1 were randomized to yield a β-sheet presenting phage display library. After five rounds of panning, only two specific HTB1 library members were preferentially selected. Of these, TJ10, which contains a large number of aromatic residues (2 Trp and 2 Tyr, postulated to be important motifs in amyloid inhibitors), was selected and found to effectively inhibit Aβ40 aggregation (Fig. 11).

[00112] The interaction of TJ10 was evaluated by designing constructs of TJ10 fused to fragmented halves of firefly luciferase. The binding of TJ10 to Aβ40 led to the reassembly of luciferase enzyme which was monitored by an increase in luminescence. This assay provides a rapid, sensitive and non-invasive method of
studying the interaction between a ligand and its target amyloid protein under physiological conditions.

[00113] Because of conflicting reports, we evaluated whether different amyloid proteins assemble into common early oligomers that proceed to form fibrils or whether different early intermediates are formed by different proteins which during the aggregation process undergo structural transition to form intermediates with generic features that over a period of time mature into characteristic long fibrils. To address this question, we studied the influence of TJ10 and its parent scaffold protein, HTB1, on the aggregation properties of several unrelated amyloid proteins. TJ10 and HTB1 share a common beta-sheet epitope but vary in 8 amino acid residues displayed on the surface of two adjacent beta-strands. This subtle difference in their structure resulted in differences in their interaction with different amyloidogenic proteins. Without wishing to be bound by any particular theory, it is believed that the results obtained with TJ10 and HTB1 indicate that the early steps in fibrillization pathway are distinct processes for different proteins.

Effect of TJ10 and HTB1 on inhibition of Aβ40

[00114] We have recently described a beta-sheet mini protein (TJ10) (Fig. 10, panel a) which was selected by phage display screening and was found to effectively inhibit aggregation of Aβ40 for over 2 weeks. This experiment was repeated with a fresh Aβ40 sample and monitored by ThT fluorescence for 24h. Along with TJ10, the effect of the parent scaffold, HTB1 (Fig. 10, panel b), on the aggregation kinetics of Aβ40 was also evaluated. Surprisingly, while TJ10 showed effective inhibition in a 1:1 molar ratio, HTB1 did not influence the aggregation of Aβ40 under the same conditions. As controls, TJ10 and HTB1 solutions (50 μM each) were also incubated under aggregating condition of Aβ40 but did not show any ThT positive aggregates over a period of 3 days. Even at sub-stoichiometric concentrations, TJ10 appear to inhibit Aβ40 effectively. After 18h, while 80% of 50 μM Aβ40 aggregated under shaking conditions, in presence of 5 μM TJ10 (TJ10:Aβ40 ratio of 1:10), only 38 % Aβ40 had aggregated (Figs. 10-11).

[00115] The interaction of TJ10 with Aβ40 was further interrogated by CD and photo-induced crosslinking (PICUP) experiments. The CD spectrum of the mixture of
Aβ40 with TJ10 (after incubation at 37°C for only 3h) was different from the simple arithmetic sum of the spectra of Aβ40 and TJ10 alone, suggesting an early interaction between the two proteins.

[00116] Photo-induced crosslinking of Aβ40, TJ10 and their mixture (1:1) was carried out after 0.5 and 10 h of incubation at 37°C and 250 rpm. Crosslinking of Aβ40 at 0 h resulted in formation of dimers and trimers as reported before, however after 5 and 10 h of incubation, crosslinking resulted in formation of aggregates of Aβ40 of molecular weight > 200 kD that could not enter the pores of the acrylamide gel and got stained in wells above the separating gel. TJ10, due to the presence of Tyr and Trp residues on its surface (which are more susceptible to free radical formation) underwent random association upon crosslinking and showed a ladder of bands on SDS/PAGE. Crosslinking of Aβ40/TJ10 mixture after incubation at 0, 5 and 10h showed a pattern of bands similar to that of crosslinked TJ10, however it is noteworthy that the bands representing crosslinked aggregates of Aβ40 were not observed for the mixture, indicating that TJ10 prevented the self-assembly of Aβ40 into large prefibrillar species.

[00117] The interaction of TJ10 with Aβ40 was examined using cell-free split luciferase based assay. While CD and crosslinking experiments indicated binding of TJ10 to early intermediates of Aβ40, more direct evidence of this interaction was demonstrated using reassembly of split firefly luciferase in the cell-free assay. In extending this method to the interaction of TJ10 with Aβ40, fusion constructs of TJ10 with each half of fragmented firefly luciferase were prepared (TJ10-NFluc and CFLuc-TJ10) and transcribed into respective mRNAs as described herein. Aβ40 solution (50 μM) was incubated under its aggregation conditions and 5 μL aliquots were taken out at different time periods and added to the translation mix along with the mRNAs to initiate the translation of the proteins (NFluc-TJ10 and CFLuc-TJ10). As the proteins were synthesized over 90 minutes, the binding of TJ10 to Aβ40 oligomers (at 2,4,6,8 h time points) or to Aβ40 fibrils (at 18h timepoint) was monitored by measuring the luminescence resulting from the activity of reassembled luciferase (Fig. 10). As a control, at each time point, reassembly of the fragmented luciferase in absence of Aβ40 was also measured. A sharp increase in luminescence from 2 fold (over control, at 2 and 4h timepoints) to 7 fold (over control) at 6h timepoint shows
preferential binding of TJ10 to Aβ40 oligomers (Fig. 10, panel b). The subsequent decrease in signal at 8 and 18h shows that TJ10 did not bind substantially to either late intermediates (at 9h some ThT positive aggregates are formed, Fig. 12) or fibrillar Aβ40 respectively. To ensure that the decrease in signal at 8 and 18h timepoints were due to reduced binding of TJ10 to prefibrillar and fibrillar species of Aβ40 and not due to deleterious effect of Aβ40 on the translation system, a second control assay was simultaneously carried out. Constructs of two heterodimerizing leucine zippers (acidic and basic zippers) with each half of fragmented firefly luciferase were also prepared (RR-NFluc and CFluc-EE), transcribed and translated in a similar fashion. The two leucine zippers, once formed, dimerized spontaneously, resulting in reassembly of luciferase and a high luminescence signal. Aβ40 aliquots (5 μL) at different time points of aggregation were added to this translation mix and the luminescence monitored after 90 minutes. Similar signal for samples with and without Aβ40 proved that Aβ40 intermediates or fibrils did not affect the translation machinery or the dimerization of the leucine zippers. Hence, the changes in luminescence at various timepoints can be correlated to the interaction of TJ10 with Aβ40. We would also like to mention here that in the present assay, the final concentration of Aβ40, at the time of detection, is 2.5 μM. The detection of Aβ40 oligomers at such low concentrations indicate low nM binding affinity between Aβ40 and TJ10 while other data collected suggest low μM binding between the two proteins. Hence, the correct stoicheometry of binding of TJ10 with Aβ40 could not be concluded.

[00118] The interaction of TJ10 and HTB1 with other amyloid proteins was also studied using the cell-free split reporter system. The effect of TJ10 and HTB1 on the aggregation kinetics of Aβ40 showed that although they share a common beta-sheet template (Fig. 10, panel a), the presence of aromatic residues on the surface of this beta-sheet template of TJ10 was necessary to disrupt the further association of Aβ40 into fibrils (Fig. 12). Based on this observation, the effect of TJ10 and HTB1 on the rate of aggregation of other amyloid proteins (Aβ42, IAPP, prion fragment 106-126, insulin and lysozyme) was also monitored (by ThT fluorescence) (Fig. 11). Three distinct trends were observed based on which these proteins were classified into three classes: class I consisted of Aβ40, IAPP and insulin. For these proteins, TJ10 effectively inhibited their aggregation while HTB1 had no influence on their rate of
fibrillization, indicating that aromatic-aromatic and aromatic-hydrophobic interactions
between the surface residues of TJ10 and early oligomers of these proteins probably
played a significant role in preventing their further self-assembly. Rayleigh and co-
workers in a recent publication showed that substituting the aromatic residues in
IAPP sequence with leucine significantly delayed its aggregation kinetics, thus
concluding that aromatic interactions did influence fibril formation of IAPP. Our data
indicate that this may also be true for Aβ40 and insulin. Class II consisted of prion
fragment 106-126 and Aβ42, which were inhibited equally by both TJ10 and HTB1. It
is likely that the beta-sheet epitope common to TJ10 and HTB1 was also involved in
stabilizing the early intermediates of prion fragment and Aβ42. This result is
interesting because prion protein deposits are also found in AD senile plaques. A
recent review discusses the similarities in post-translational modifications and metal
binding domains of prion protein and Aβ (7). AChE which promotes Aβ aggregation,
has also been shown to trigger fibrillization of prion 106-126 (8). Hence, the present
methods can be useful for simultaneous targeting of Aβ42 and prion protein.

[00119] Finally, Class III consists of lysozyme, which aggregates in 3 days under
the certain conditions without being affected by either TJ10 or HTB1. Previous
studies on human lysozyme and its variants have shown that aggregation of
lysozyme proceeds by cooperative unfolding of the beta-domain of the enzyme,
followed by self association of this species to form beta-sheet rich fibrils. This
transient intermediate does not appear to interact with either TJ10 or HTB1 and
aggregates in 3 days at 37 °C with shaking (Fig. 10, panel d).

[00120] Much recent work has been carried out to identify conditions for isolation
and characterization of amyloid protein aggregation intermediates that are now
considered the true toxic species. These intermediates have been examined using
techniques including CD, FRET, fluorescence polarization and 2D NMR, analytical
ultracentrifugation, pulse-labeling hydrogen/deuterium exchange coupled with mass
spectrometry, immunological detection, TEM and atomic force microscopy. Most of
these techniques are elaborate, some requiring chemical derivatization, while others
analyze the amyloid samples under non-physiological conditions and most utilize
expensive, specialized instruments for measurements.
In the present work, we attempted to define the stage of the aggregation reaction at which TJ10 binds Aβ40, by fusing TJ10 to fragmented halves of a reporter protein, firefly luciferase and monitoring its reassembly mediated by aggregated Aβ40 via an increase in luminescence. It was evident that TJ10 did not interact with low molecular weight Aβ40, late intermediates or fibrils of Aβ40 (Fig. 10, panel b). Instead, TJ10 showed higher affinity for oligomers formed after 6h of incubating Aβ40 under aggregation condition (Fig. 11, panel b), a time period when no significant ThT fluorescence was detected in Fig. 10, panel b.

Reporter proteins, most commonly GFP, have been used in *in vivo* studies to directly monitor the aggregation of Aβ, polyglutamine and prion. Cell-based luciferase systems have been used to study the effect of presenilin proteins on Notch signaling (Am 1) and in screening for chemical compounds that inhibit APP processing (Am2, Am3). GFP fused to Aβ42 in *E. coli* has also been reported for screening for small molecule inhibitors of Aβ42 aggregation (Am4). However, to our knowledge, this is the first report utilizing reassembly of split reporter protein for detecting soluble oligomers of amyloid proteins.

It is advantageous that the transcription product encoding one or both of the binding region-split reporter portion comprises a stabilizing element such as a 5' stem-loop, including but not limited to a 5' stem-loop derived in sequence from bacteriophage T7 and advantageously a 3' stem-loop such as that derived in sequence from bacteriophage T3 (see, e.g., 134, 135); and also advantageously a Kozak sequence 5' to the translation start site (see, e.g., 133) and a polyadenylated 3' end (see, e.g., 136) when the cell-free translation system is a eukaryotic system or a Shine–Delgarno sequence when the cell-free translation system is a bacterial translation system. Exemplary Shine Delgarno sequences include UAAGGAGGUGA (SEQ ID NO:3), AGGAG or variants as well known in the art. Examples of Kozak sequences are GCCACCAGG (SEQ ID NO:4), CCACCCTAGG and variants thereof, also as well known in the art. Useful examples of T7 and T3 promoter sequences are TAATACGACTCACTATA (SEQ ID NO:5) and AATTAACCCCTCACTAAA (SEQ ID NO:7), respectively. *Escherichia coli* transcription initiation signals are widely known in the art; typically they are identified by TTGACA, followed by 15-19nucleotides, and TATAAT 5' to the start site. The
SP6 promoter signal is ATTAGGTGACACTATA (SEQ ID NO:8) or a functional variant thereof. RNAPII polymerases can be used with the appropriate choice of promoter and in the appropriate eukaryotic system, for example, a couple transcription-translation system.

[00124] When RNA or DNA is expressed as proteins in the assay, the first and second fragments of the reporter associate to give a detectable signal (above background) when the two fragments associate in a manner which is mediated by the interacting protein portions fused to those reporter fragments, either directly or via a ligand or other molecule which binds to each of the interacting proteins. That is, the two reporter fragments are brought in sufficiently close proximity to allow their reassociation to form a functional protein that provides, directly or indirectly, a detectable signal.

[00125] It is understood that the direct interaction of the interacting proteins to form a reassociated, functional reporter can be disrupted by an antagonist of the interaction, which can be a small or a large molecule. In other choices of interacting protein fragments (ligand binding) there can be association of the reporter fragments mediated by an additional molecule which brings the two interacting proteins together. Again, an antagonist of the interaction of the molecule which binds to both interaction proteins can be identified by a decrease in signal resulting from decreased reassociation of the two fragments of the split reporter. Similarly, agonists can be identified where the signal is greater in the presence of the agonist than in its absence. It is understood that compositions tested for antagonist or agonist activity can be pure or relatively pure compounds, or libraries of compounds.

[00126] The present method is relatively rapid (requires about 90 minutes), does not require specialized skills or expensive instrumentation and is sensitive enough to detect low concentrations of transient oligomers formed during the lag phase of the aggregation process that are not detected in the ThT fluorescence assay. This assay can also be extended to create fusion constructs of other inhibitors with amyloid proteins.

[00127] TJ10 and its parent protein HTB1 were also used to delineate similarities and differences in the aggregation pathway of unrelated amyloid proteins. Despite
much work in this area, there is still no clear understanding of the mechanism of aggregation and the related cytotoxicity. The detection of spherical oligomers and pore-like annular assemblies for α-synuclein and polyglutamine, that had previously been observed for Aβ and increased membrane permeability by soluble oligomers of these proteins, implied a common aggregation pathway. This theory was further advanced by detecting common soluble amyloid oligomers using polyclonal antibodies generated against micellar Aβ by Glabe and co-workers. However, about the same time, crosslinking experiments on Aβ40 and Aβ42 by Bitan et al showed that these two proteins oligomerized through distinct pathways. Also, chaperones Hsp70 and Hsp40 were shown to reduce the density of spherical and annular assemblies of polyglutamine by increasing the density of fibrils, and these chaperones prevented fibrillization of Aβ42 by stabilizing spheroid like intermediates. Pruisner and coworkers have reported a preamyloid state for prion protein which they propose forms a steady-state trimeric complex that can be stacked to form fibrils (Am5).

[00128] In view of inconsistent reports in the art, we evaluated the aggregation pathway of different amyloidogenic proteins based on their interaction with HTB1 and its aromatic rich variant, TJ10. Three trends were observed: Class I proteins, including of Aβ40, IAPP and insulin, were inhibited from fibrillization by TJ10 but were not influenced by the presence of HTB1 (Fig. 10, panels b and c). TJ10, selected to inhibit Aβ40 aggregation, also prevented IAPP aggregation for 18h, since, both these proteins share more than 70% sequence similarity, and a recent study showed the suppression of Aβ aggregation by an IAPP mimic block. It is probable that the mode of inhibition of IAPP by TJ10 is similar to its inhibition of Aβ40. To test this possibility, the split-luciferase assay was carried out with IAPP samples at different time periods. A 5 fold increase in luminescence (over control) was observed within 15 and 45 min of incubating IAPP under its aggregation conditions, which decreased to 2.5 fold in 1.5h and further to 1.5 fold (over control) at 4.5 and 6h. IAPP is highly prone to aggregation and shows significant ThT fluorescence at 4 and 6h (Fig 10b). Hence, the split-luciferase assay with IAPP showed that as in the case of Aβ40, TJ10 interacts with early intermediates of IAPP. These results indicate that the early intermediates of Aβ40 and IAPP may share some features that are recognized by TJ10. Class II includes Aβ42 and prion
fragment 106-126, which were inhibited equally by TJ10 and HTB1; and Class III includes lysozyme which did not interact with either TJ10 or HTB1. It is noteworthy that the interaction of TJ10 and HTB1 with Aβ40 is different from their interaction with Aβ42, thus our data seem to correlate with the crosslinking experiments by Bitan et al.

[00129] Comparing these results with results obtained with polyclonal antibody isolated by Glabe and coworkers and also with monoclonal antibodies, WO1 and WO2 isolated by ONualin and Wetzel, it is believed that the initial misfolding and association into low molecular weight assemblies may be distinct processes for different proteins. Hence, the early intermediates of Class I proteins (l₁) are different from those of Class II proteins (l₂) and Class III protein, lysozyme (l₃) (Fig. 12). However, as aggregation proceeds, these initial intermediates self associate and undergo conformational transition to form generic late intermediates such as protofibrils (l₄) which further mature into long fibrils that have been detected by monoclonal antibodies WO1 and WO2. The polyclonal antibody isolated by Glabe and coworkers is a heterogenous mixture of antibodies that bind a broad spectrum of soluble oligomers ranging from octamers and dodecamers to large spheroids and protofibrils of Aβ. The split-luciferase assay showed TJ10 to interact only with early intermediates of Aβ40 and IAPP.

Cell-Free Split Luciferase Enabled Assays for Small Molecule Inhibitors of Kinases

[00130] Protein fusion constructs of Fos-NLuc(residues 2-416) and CLuc(residues 398-550)-Kinase form a ternary complex and generate active luciferase upon addition of the peptide-inhibitor conjugate Jun-staurosporine (Fig. 13). Complex formation is driven by specific interactions between staurosporine and kinase active-site as well as Fos and Jun. RNA encoding each fusion construct (0.5 pmol Fos-NLuc and 0.2 pmol CLuc-Kinase) was added to duplicate 25 µl rabbit reticulocyte lysate translation reactions and incubated at 30°C for 1.5 hrs. Water (negative control) or Jun-staurosporine (positive control) was subsequently added to the lysate reaction to a final concentration of 125 nM and incubated for 1 hr at room temperature in the dark to equilibrate. Luminescence measurements were taken by adding 80µl STEADY-GLO™ Luciferase Assay Reagent to 20 µl of translation mix
followed by one minute incubation at room temperature. Luminescence was measured using a single tube luminometer with a 10 second integration (Fig. 16, panel a).

[00131] Complex formation is reversible and the addition of small molecules competitive with staurosporine for kinase active-site binding can be added into the mixture to induce dissociation of the complex, resulting in a commensurate loss in luciferase activity (Fig. 13). This allows one to screen a kinase against a library of potential or known kinase inhibitors. By adding individual inhibitors to separate lysate reactions also containing Jun-staurosporine, one can directly compare the loss in activity to the positive control which contains only Jun-staurosporine and no inhibitor.

[00132] Because of the conserved nature of protein kinase active sites and staurosporine’s ability to bind to many kinases, this system can additionally be applied to any kinase exhibiting sufficiently high affinity for staurosporine. Alternatively, staurosporine can be replaced in the peptide conjugate with an inhibitor which has high affinity for kinases that staurosporine does not bind. A diverse library of kinase active domains fused to the C-terminal fragment of luciferase can then be screened against multiple inhibitors that act competitively with staurosporine.

[00133] A panel of three inhibitors was assayed against four CLuc-fusions containing unique kinase domains (Fig. 15). Translations were prepared as above, in duplicate. Separate reactions were prepared for CLuc-PKA, CLuc-PDGFRB, CLuc-DK2, and CLuc-Fyn to contain 0.2 pmol of each RNA with 0.5 pmol Fos-NLuc RNA per 25 μl reaction. Each mixture was incubated at 30°C for 1.5 hrs. From stock lysate mixtures, several reactions were prepared. For the negative control, 1 μl water was mixed with 24 μl lysate, followed by adding 1 μl dimethylsulfoxide (DMSO) to 24 μl of this mixture. For all inhibitor reactions and the positive control, jun-staurosporine was added to lysate translations to a final concentration of 125 nM. This mixture was then divided into aliquots containing DMSO or inhibitor dissolved in DMSO to create mixtures containing inhibitor at a final concentration of 50 μM and a positive control (DMSO only). The inhibitors tested were Sunitinib (LC Laboratories, Woburn, MA), PP1 (A.G. Scientific, Inc., San Diego, CA), and Roscovitine (LC
Laboratories). Upon mixing all reactions were incubated for an hour at room temperature in the dark. Luminescence readings were taken as described above.

[00134] Similarly, a designed small molecule can be used in place of the peptide-inhibitor conjugate described above to facilitate complex formation and signal generation. The N-terminal (residues 2-416) portion of firefly luciferase is tethered to dihydrofolate reductase (DHFR) while the C-terminal (residues 398-550) portion is tethered to the protein kinase PKA. The designed small molecule consists of the small molecule trimethoprim, which is a known inhibitor of the E. coli DHFR conjugated through a tetraethylene glycol linker to the broad spectrum kinase inhibitor staurosporine creating the molecule Tri-Staur (Fig. 14). Simultaneous binding of trimethoprim by DHFR and Staurosporine to the kinase (PKA) active site results in ternary complex formation and reassembly of active firefly luciferase. Addition of staurosporine competitive small molecules results in the dissociation of the kinase-staurosporine complex and subsequent loss of luciferase activity (Fig. 15-16).

[00135] Duplicate translations were initiated by the addition of in vitro transcribed RNA encoding each fusion construct (0.5 pmol DHFR-NLuc and 0.2 pmol CLuc-PKA) to 25 µl rabbit reticulocyte lysate and incubated at 30°C for 1.5 hrs. DMSO (negative control) or Tri-Staur (positive control) was subsequently added to the lysate reaction to a final concentration of 125 nM and incubated for 0.5 hr at room temperature in the dark to equilibrate. Luminescence measurements were taken by adding 80µl STEADY-GLO™ Luciferase Assay Reagent to 20 µl of translation mix followed by a one minute incubation at room temperature. Luminescence was measured using a single tube luminometer with a 10 second integration (Fig. 16).

[00136] To evaluate the ability of Tri-Staur to identify kinase active site binding small-molecules duplicate translation were initiated as described above and used to assay the kinase inhibitor, staurosporine. From a stock lysate mixture, several reactions were prepared. For the negative control, 2 µl of DMSO was mixed with 23 µl of lysate. For all inhibitor reactions and the positive control, Tri-Staur was added to all translations to a final concentration of 125 nM. 24 µl of this mixture was then added to 1 µl of staurosporine dissolved in DMSO (final concentrations of 1.25 nM – 12.5 µM) or DMSO only. Upon mixing all reactions were incubated for 0.5 hours at
room temperature in the dark. Luminescence readings were taken as stated above. Luminescence measurements revealed a concentration dependent decrease in luciferase activity upon the addition of free staurosporine (Fig. 16b), demonstrating the use of this methodology for the identification of kinase active site binding small molecules.

[00137] We set out to develop a general approach for the sensitive, sequence-specific targeting of ssRNA by building on our cell-free split-firefly luciferase (Fluc) system (Figure 17) (19,27,58). We have recently shown that the use of native and designed pumilio domains7 attached to split-Fluc (Fig. 17, panel B) results in ssRNA target-dependent protein reassembly and luminescence.6c Though useful, this approach is inherently limited by the need to design pumilio domains with tailored specificity for each new ssRNA target. Thus, as a first step we amended our earlier design by replacing each pumilio domain with a ssRNA binding protein, the PAZ domain of argonaute 2 (residues 219-363) (76,77). Argonaute (Ago), which is central to the dicer complex formed in RNA interference (RNAi), binds with high affinity to the 2-nucleotide, 3' overhangs of short dsRNA. We postulated that by replacing the sequence-specific pumilio domains with the Ago PAZ domain, we would be able to direct our split-proteins toward any user-defined sequence of RNA by providing short, complementary guide oligonucleotides similar to those used in RNAi. Thus, we attached an Ago domain to each half of split-Fluc to generate NFluc-Ago and CFluc-Ago. We initially tested this approach by using NFluc-Pum2 and CFluc-Ago to detect a specific target RNA. The Pum1 ssRNA guide (Table 1), designed for directing CFluc-Ago to its binding site, was pre-hybridized to the target. Following incubation of the split-Fluc constructs with 10 nM of the guide-target complex, a significant signal over background was observed (Fig. 17, panel C). To our knowledge this is the first example of Ago being employed as a detection domain, providing a general module in protein-based nucleic acid detection. In a complimentary set of experiments, NFluc-Ago and CFluc-Pum1 were successfully used to detect 10 nM of target (Fig. 17, panels B-E). Finally, we attempted to provide a general solution to ssRNA detection by hybridizing both the Pum2 and Pum1 guides to the target (Fig. 17, panel E). However, upon incubation with NFluc-Ago and CFluc-Ago, less than a two-fold signal increase was observed in the presence of the ternary guide-target complex. This may be attributable to the affinity of Ago for its
target as well as to the statistical distribution of guides in the presence of excess RNA present in our translational systems. Thus, these experiments suggest that Ago may be employed successfully in conjunction with other available sequence-specific RNA binding domains, and its use allows one to access a larger detectable sequence space than pumilio domains alone.

**Cloning of NFluc-Ago and CFuc-Ago and Argonaute Refolding Experiments**

Restriction enzymes, dNTPs, Antarctic phosphatase, and T4 DNA ligase were purchased from New England Biolabs (NEB, Ipswich, MA). Pfu Ultra polymerase was obtained from Stratagene. All DNA and RNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). RNasin® Plus RNase Inhibitor, T7 RiboMAX™ Large Scale RNA Production kit, Flexi® Rabbit Reticulocyte Lysate, and Steady-Glo® Luciferase Assay System were acquired from Promega.

<table>
<thead>
<tr>
<th><strong>Table 1. Cloning Primers</strong></th>
<th>DNA and RNA oligonucleotides used in Ago cloning and RNA detection.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFluc-Ago Primers (5’→3’)</td>
<td>GGATACCCGGGCGACAGGCCGTCATGAG (SEQ ID NO:46)</td>
</tr>
<tr>
<td>Ago Xmal FWD</td>
<td>CCCCTATTCGACGGCCTAGACCTGCTCT (SEQ ID NO:47)</td>
</tr>
<tr>
<td>CFuc-Ago Primers (5’→3’)</td>
<td>GGATACCGGATCCGGGCGACAGGCCGTCATGAG (SEQ ID NO:48)</td>
</tr>
<tr>
<td>Ago Xhol REV</td>
<td>CCGCTATACCCGGGCGACAGGCCGTCATGAG (SEQ ID NO:49)</td>
</tr>
</tbody>
</table>

**RNA Target**

RNA target 5’-CAUGUGUAGUAUAAGUCUUUGAUAAUGCAGGC (SEQ ID NO:50)

**RNA guides**

Pum1 guide 5’-CUAAUAACACCAUAAUG (SEQ ID NO:51)
Pum2 guide 5’-GCGCCCUAAUCAUU (SEQ ID NO:52)

**in vitro transcription primers**

NFluc-Ago FWD 5’-CGAGCTTAATAGCTCAGCTATAGGGAAACAGACACCCATGCGGCTTCTCTCAGGAAATGATGCTGAATTTAGGATCTGGA (SEQ ID NO:53)
NFluc-Ago REV 5’-CCGCCACACCACCTTCTCTCAGGAAACAGACACCCATGCGGCTTCTCTCAGGAAATGATGCTGAATTTAGGATCTGGA (SEQ ID NO:53)
CFuc-Ago FWD 5’-GCGAGCTAATAGCTCAGCTATAGGGAAACAGACACCCATGCGGCTTCTCTCAGGAAATGATGCTGAATTTAGGATCTGGA (SEQ ID NO:53)
CFuc-Ago REV 5’-CCGCCACACCACCTTCTCTCAGGAAACAGACACCCATGCGGCTTCTCTCAGGAAATGATGCTGAATTTAGGATCTGGA (SEQ ID NO:53)

The restriction sites in the primers are shown in bold. The Pum1 and Pum2 binding sites in the RNA target are shown in bold and blues, respectively. The regions complementary to the RNA target are shown in bold for the RNA guides. For the in vitro transcription primers, T7 promoters are shown in italicized, KOZAK sequences in bold, and 3’ stem-loops underlined.

The RNA binding PAZ domain of *Homo sapiens* argonaute-2 (Ago) was PCR amplified from pIRESneo-FLAG/HA Ago2 corrected (Addgene plasmid 10822; Ref 9b of main text), which encoded residues 1-856 of hsAgo-2, using primers
indicated in Table 1 under the heading "Cloning Primers". Only the RNA binding domain (residues 219-363) was amplified, since adjacent domains have endonuclease activity.\(^1\) Plasmids containing NFluc-PBSI\(\text{I}\) and CF\(\text{I}\)uc-Zif268 were digested at Xmal/Xho\(\text{l}\) and BamH\(\text{I}\)/A\(\text{g}\)el, respectively. The digested plasmids were phosphatased, then ligated to the Ago inserts generated by PCR. Sequences were confirmed by dideoxynucleotide sequencing. All methods concerning cloning of NFluc-Pum2 and CF\(\text{I}\)uc-Pum1 have been described elsewhere (58).

[00140] We also amplified the same sequence, with 20 additional residues on the N- and C-terminus to simulate flexible linkers. Although no data are shown herein, these constructs functioned essentially the same as the PAZ domain-only constructs in all cases tested.

[00141] Annealing of guides to RNA target was carried out as follows. Guides (Table 1, RNA Guides), containing regions of complementarity to the pumilio target followed by two 3'-rU's to facilitate Ago binding, were annealed to the pumilio target using the following procedure: heating of target (Table 1, RNA Target) and guide in NEBuffer 4 (20 mM Tris-acetate, pH 7.9, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol) supplemented with RNasin\textsuperscript{TM} to 90°C for 1 minute, followed by cooling to 37°C over one hour. Samples were stored at -80°C.

[00142] Genes encoding NFluc-Ago and CF\(\text{I}\)uc-Ago were PCR amplified using \textit{in vitro} transcription primers (Table 1, \textit{in vitro} transcription primers) containing a T7 promoter and a KOZAK sequence in the forward primer and a stem-loop sequence in the reverse primer. These primers were designed so that the complementary regions had melting temperatures greater than or equal to 70°C. A typical PCR amplification included an initial heat denaturation of 95°C for 5 min, followed by 40 cycles of heating to 95°C, cooling at a rate of 6°C/min to an annealing temperature of 53°C. Elongation at 72°C for 6 minutes completed the cycle. The PCR products were then used as templates for \textit{in vitro} transcription using a T7 Ribomax RNA production kit according to the manufacturer's suggestions. Generally, 3 µg of amplified DNA template was incubated at 37°C for 3 hours in the presence of 1x T7 transcription buffer, 7.5 mM rNTPs, and T7 enzyme mix. The mRNA generated was purified over illustra ProbeQuant\textsuperscript{TM} G-50 Micro Columns (GE Healthcare) and analyzed by agarose gel electrophoresis. The mRNA was then introduced into the Flexi Rabbit
Reticulocyte Lysate System for in vitro translation to yield the protein constructs. 25 
µL reactions were set up in duplicate according to the manufacturer's instructions. A 
typical reaction was performed at 30°C for 1.5 hours and consisted of the following 
components: 0.5µL amino acid mix, 70 mM KCl, 1 mM DTT, 0.8 U/µL RNasin™, 0.1-
2 pmol each mRNA transcript, and Nuclease-free H₂O (NEB) to final volume. 
Specific mRNA and target conditions used in each of the experiment types (NFluc-
Pum2/CFIuc-Pum1, NFluc-Pum2/CFIuc-Ago, NFluc-Ago/CFIuc-Pum1, and NFluc-
Ago/CFIuc-Ago) are given in Table 2.

<table>
<thead>
<tr>
<th>mRNA 1</th>
<th>mRNA 2</th>
<th>Final target concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 pmol NFluc-Pum2</td>
<td>2 pmol CFIuc-Pum1</td>
<td>10 nM target</td>
</tr>
<tr>
<td>0.1 pmol NFluc-Pum2</td>
<td>0.1 pmol CFIuc-Ago</td>
<td>10 nM Pum1 guide + 10 nM target</td>
</tr>
<tr>
<td>1 pmol NFluc-Ago</td>
<td>1 pmol CFIuc-Pum1</td>
<td>10 nM Pum2 guide + 10 nM target</td>
</tr>
<tr>
<td>2 pmol NFluc-Ago</td>
<td>2 pmol CFIuc-Ago</td>
<td>10 nM Pum1 guide + 10 nM Pum2 guide + 10 nM target</td>
</tr>
</tbody>
</table>

[00143] In the case of NFluc-Pum2 and CFIuc-Pum1, the ssRNA target was added 
at the beginning of the translation reaction. In all other cases the target with 
annealed guide(s) or a buffer blank (NEBuffer4) was added after completion of 
translation, and binding was allowed to occur for 1 hour at 4°C. The presence of 
target with annealed guide(s) should provide a binding site for the domains attached 
to each of the luciferase halves, resulting in formation of a functional enzyme. 
Activity was monitored as a luminescent signal produced upon addition of Steady-
Glo™ Luciferase Assay system. 20 µL of each translation reaction equilibrated with 
target (or buffer) was added to 80 µL of Steady-Glo reagent and allowed to 
equilibrat at room temperature for 1 minute. Luminescence readings were acquired 
using a Turner TD-20e Luminometer with a 10 second integration time. Two 
duplicate translation reactions were vperformed simultaneously. Luminescence 
readings were averaged and normalized to one. The data summarized are results for 
two sets of duplicate translation reactions performed on separate days. The results 
from each day were averaged and normalized to one, then averaged together. The 
standard deviation presented for the sample in the presence of target is that of the 
set of translations with the greatest variation.

[00144] Additional control experiments were used to solidify the experimental 
results: The following results demonstrate selectivity of the Ago detection domain. 2
pmol of NFLuc-Pum2 and CFluc-Ago mRNA were translated as described above, except the targets were present during translation. Luminescence readings demonstrated that both target and guide must be present for Fluc reassembly to occur. The signal generated with the combination of Pum1 guide (3'-UGUACCACAUAUAUC, SEQ ID NO:57) and Target RNA (5'-CAUGGUGUAUAUAGUCUUUUUGUAUAUCG, SEQ ID NO:58) was approximately threefold that obtained with either Target RNA or no Target RNA or Guide and ½ site Target RNA (5'-CCGAAAUUGUAUAUAUCG, SEQ ID NO:59). Additionally the Pum1 guide annealed to a ½-site target site, in which no Pum2 binding site exists, resulted in only background luminescence (compare bars 1, 3, and 4). Signal is lower than that in certain earlier experiments since this experiment was performed using non-optimized mRNA amounts (2 pmols as compared to 0.1 pmol each. mRNA).

[00145] We created a general sequence-specific ssRNA detection strategy without the need for programming nucleic acid binding proteins for each new ssRNA target. Learning from the Ago-guide strategy, we envisioned that attachment of high affinity (Kd ~ low pM), sequence-specific zinc fingers (ZF s) (22,25,78,79) to our split-proteins could serve as a motif for displaying any user-defined ssDNA guide when appended to a ZF hairpin (hp) binding site, providing a hp-guide (Fig. 18, panel a). These hp-guides would allow ZF-mediated split-Fluc reassembly only in the presence of a targeted ssRNA or ssDNA sequence. Towards this goal, we designed DNA hps as binding sites for two well-characterized and high affinity six-finger ZFs (E2C and Aart) (80,81,82) and to these we attached guide sequences of ssDNA complementary to the ssRNA target (Table 3). Addition of our designed ZF-modified split-Fluc constructs to 1 nM target (ssRNA or ssDNA) hybridized to complementary hp-guides resulted in luciferase reassembly and a signal of 4.5-fold as compared to background (Fig. 18). To directly interrogate sequence selectivity, a single G to T mutation was introduced in the ssDNA, which reduced signal to background levels.

[00146] To probe the generality of our ssRNA detection approach, we selected two disease relevant targets, vascular endothelial growth factor (VEGF) and human epidermal growth factor receptor 2 (HER2) mRNA1(83,84). By judicious exchange of the guide portion of the existing hp-guides to designed ssDNA oligonucleotides
complementary to two adjacent 19 nt sequences in the 295 nt VEGF mRNA transcript (Table 3), we observed a 39-fold signal in the presence of 1 nM target (Fig. 18, panel C) and as low as 1 pM (100 attomoles) VEGF mRNA was detectable. Finally, hp-guides were designed (Table 3) for detecting two adjacent 19 nt sequences present in a 201 nt HER2 mRNA sequence, where a 4.3-fold signal over the presence of hp-guides alone was achieved (Fig. 18D), further confirming the general applicability of this nucleic acid detection system.

[00147] In conclusion, we have developed a general ssRNA and ssDNA detection methodology utilizing split-protein reassembly, which allows for distinguishing single base substitutions and detecting attomoles of a user-defined target. Thus this new bioluminescence based methodology complements existing ssRNA and ssDNA detection methods (70,71,75) and future studies demonstrate that this methodology is applicable to in vivo imaging.

Rapid Interrogation of Transcription Factor Binding, DNA Methylation, and poly(ADP-ribosyl)ation

[00148] The sequencing of the human genome revealed that transcription factors comprise the largest single group of proteins, which is perhaps not surprising since the ability to accurately recognize, bind, regulate specific DNA sequences is central to the regulation of almost all cellular processes. DNA activity is not only regulated by protein factors but also by the reversible chemical modification of both DNA and associated proteins, impacting cellular biology at multiple levels. Accordingly, the study of native DNA, its associated proteins, and their chemical regulation is of profound importance to the study of biology. Thus, there is a need for new, generally applicable methodologies for detecting and studying DNA and its associated proteins, which in turn aid in the development of designed transcription factors (85-88), and allow for the development of small molecules that modulate transcription (89-92) DNA methylation (93,94), and poly(ADP-ribose) metabolism (95,96).

[00149] General approaches for the direct detection of native dsDNA include the use of triplex forming oligonucleotides (TFOs) as well as sequence specific polyamides. Though powerful, these techniques also have associated limitations. TFOs bind the major groove of dsDNA through Hoogsteen or reverse Hoogsteen
base pairing and are thus restricted to detecting sequences comprised of purines on one strand (97). Additionally, TFOs composed of polypyrimidines require cytosine protonation in order to bind, imposing a pH regime outside of physiological conditions. On the other hand, sequence specific polyamides, which are comprised of designed N-methylpyrrole and N-methylimidazole heterocycles, are capable of recognizing all four base pairs with affinities that rival naturally occurring DNA-binding proteins (98,99). Polyamides have proven to be extremely useful for the direct in vitro detection of dsDNA when modified with environmentally sensitive fluorophores, however this approach has yet to be tested against dsDNA sequences beyond 9 bp (100,101). Furthermore the aforementioned methods are currently not capable of recognizing and reporting on chemical modifications to dsDNA such as cytosine methylation. This DNA methylation dependent epigenetic regulation is mediated by the action of methyltransferases at CpG dinucleotide sites and is currently of particular interest due to its link to transcriptional repression and cancer (102). Though distributed throughout the genome, CpG methylation is primarily excluded from promoter-associated CG-rich regions of sequence known as CpG islands. The aberrant hypermethylation of these promoters, particularly those associated with tumor-suppressor genes, has been shown to occur in a sequence specific and tumor-type specific manner, leading to the elucidation of gene hypermethylation profiles for a number of cancer types (103). Additionally, overall genome wide hypomethylation has also been associated with tumor cells (104,105), making the determination of DNA methylation a potential biomarker for cellular states, particularly tumorigenesis. Moreover, tools for the measurement of methylation and demethylation can also be potentially utilized for measuring the activities of associated DNA modifying enzymes.

[00150] In addition to modifications to dsDNA, modifications to proteins associated with DNA are also of much interest, for example, proteins associated with DNA may be modified by poly(ADP-ribosylation) (PAR), which usually occurs in response to DNA damage (106). Poly(ADP-ribosylation) is catalyzed by poly(ADP-ribose) polymerases (PARPs), while deribosylation is catalyzed by polyADP-glycohydrolysases (PARGs) (107). This modification is implicated in transcriptional regulation (108,109), apoptosis (110,111), and tumorigenesis (112,113). The ability to directly measure the presence and extent of protein poly(ADP-ribosylation) using
designed reagents has yet to be realized and could provide a potent marker for DNA damage as well as a method for interrogating the activity of enzymes such as PARP and PARG. As described herein, we have demonstrated the application of the split-protein reassembly approach to interrogate the aforementioned targets.

[00151] Generally, split-protein reassembly or protein complementation utilizes a protein reporter dissected into two inactive fragments, each of which when appended to a member of an interacting protein/peptide pair results in reassembly of the dissected protein reporter whose activity can be measured. This approach beginning with ubiquitin has been utilized for the in vivo detection of a wide variety of biomolecular interactions utilizing monomeric split-reporters such as beta-lactamase, green fluorescent protein, and luciferase (12-19,27). Recently, designed split-protein reassembly methods have also been applied towards the study of protein-DNA interactions (22,23,25,58), as well as the determination of DNA methylation and protein phosphorylation (23,27,37). Additionally, Varshavsky has proposed how DNA sequence enabled split-protein reassembly may be utilized as an enabling therapeutic strategy (115), while Barbas and coworkers have elegantly utilized this method for directed methylation of a dsDNA target (116). In each of these cases two nucleic acid-binding domains direct the formation of a ternary complex in the presence of a target dsDNA providing a sensitive “turn on” sensor (Fig. 20), which is conceptually similar to a yeast-three hybrid(117) or chemical dimerizer approach (118,119). Herein we demonstrate the utility of the fragmented firefly luciferase system in constructing extremely rapid and sensitive reporters capable of the direct detection of a) specific 36-basepair DNA sequences; b) sequence specific as well as overall DNA methylation; c) dimeric DNA-binding transcription factors; and d) the direct determination of poly(ADP-ribosyl)ation.

Direct and sensitive detection of 36-mer dsDNA sequences.

[00152] We have previously shown that 3-finger zinc finger proteins can be utilized to directly detect an 18 bp dsDNA molecule using split- GFP, beta-lactamase, and luciferase while polyamide based approaches have been shown to detect dsDNA molecules shorter than 9 bp. We then tested if it is possible to detect shorter and longer dsDNA molecules and examine whether sequence specific 6-finger zinc fingers (ZFs) offer a means for greater affinity and selectivity than their 3-fingered
counterparts and whether it is possible to recognize and directly detect 36 bp of dsDNA. Beyond the diagnostic capabilities of such a reagent, the ability to design both specific as well as high affinity ZFs is of particular relevance for ZF mediated gene therapy approaches.

[00153] As an initial test we chose the well characterized 6-finger ZF Aart, a designed finger which binds the 18 base pair sequence 5'-ATGTAGGGAAAGCAGGG-3' (SEQ ID NO:1) with a reported $K_d$ of 50 pM (81,82). To create a platform capable of dsDNA dependent luciferase reassembly (our most sensitive reporter), the N-terminal fragment of luciferase was fused to Aart creating Aart-NFLuc(residues 2-416) while the C-terminal fragment of luciferase was fused to the 3-finger ZF Zif268 creating CFLuc(residues 398-550)-Zif268. Cell-free translations were initiated by adding in vitro transcribed mRNA corresponding to the fragmented luciferase-ZF fusions in either the presence or absence of the dsDNA target Zif268-2-Aart. A 15-fold increase in luminescence was observed in the presence of 100 pM (10 fmols) Zif268-1-Aart, demonstrating the ability of a 6-finger ZF to facilitate split-luciferase reassembly in a dsDNA templated fashion (Fig. 21, panel b). Having demonstrated the ability of Aart to direct the reassembly of fragmented luciferase, we next designed a DNA sensor in which both fragments of luciferase were attached to 6-finger ZFs, where we chose another designed 6-finger ZF, E2C, which was designed to bind the 18 base pair sequence 5'-GGGGCAGGCCGAGCAGTG-3' (SEQ ID NO:2) with a $K_d$ of 500 pM(53). To recognize 36 bp of dsDNA, the 3-finger ZF Zif268 was replaced by E2C, creating the fusion CFLuc-E2C (Fig. 21, panel a). Initial experiments demonstrated a 70-fold increase in luminescence over background in the presence of 100 pM (10 fmols) of the dsDNA target E2C-1-Aart). Importantly, minimal signal was generated in the presence of 100-fold excess genomic herring sperm DNA, suggesting low non-specific binding for these 6-finger ZFs. Furthermore, titrations with decreasing amounts of E2C-1-Aart in the presence of the two 6-finger proteins attached to split-luciferase (Fig. 21, panel c) show that as low as 2 pM (200 amols) of target dsDNA is clearly detectable above background (2.7-fold). In comparison, an analogous system utilizing the 3-finger ZFs Zif268 and PBSII allowed for the detection of 10 pM (1 fmol) of target dsDNA with a 1.3-fold signal above background. These results demonstrated the possibility of improving the detection limit associated with the use
of 6-finger ZFs, presumably due to their higher affinity as compared to 3-finger ZFs. Next we wanted to directly address whether the split-luciferase approach provides a method for directly interrogating the ZF specificity of these 18bp binders.

Selectivity of the designed 6-finger zinc finger Aart.

[00154] Interestingly Aart, which was designed to bind the aforementioned A-rich DNA sequence, was found, through cyclic amplification and selection of targets (CAST) assays, to prefer an alternate and more G-rich consensus sequence 5'-ATGTAGGGAAGCCGG-3' (SEQ ID NO:1) (120). Of particular interest is finger 3, which shows a very strong preference for the DNA triplet 5'-AAA-3' (underlined) in both the consensus and the designed sequences. A recently available co-crystal structure suggests that the origin of the observed specificity likely arises from the specific hydrogen bond contacts provided by residues Q89 and N92 (Figure 21, panel a) (82). To functionally interrogate the specificity of Aart for this triplet in the context of the consensus and designed targets, cell-free translations of the split-proteins, Aart-NFluc and CFluc-E2C, were initiated in the presence of the target (consensus) oligonucleotide E2C-1-Aart, and a series of competitor hairpin DNAs (hpDNAs) (Fig. 22, panel b). The hpDNAs contained either the designed or consensus Aart recognition sequences or the designed or consensus sequence containing the mutated 3rd triplet 5'-ATA-3', which should lack the possibility of the N92 hydrogen bond to adenine. In all cases a competitor hairpin concentration dependent decrease in luminescence was observed, facilitating the determination of IC50 values for each competitor hpDNA (Fig. 22, panels c and d). The observed 6-fold difference in IC50 between the designed and consensus hpDNAs confirms the previously reported preference of Aart for the consensus target. More interestingly, competitor hpDNAs with single A to T bp substitutions in the consensus and designed binding sites reduced the signal intensity by 60- and 20-fold when compared to the parent hpDNA targets, speaking to the surprising specificity of these 6-finger proteins. Thus this approach allows for a rapid method for evaluating ZF specificity, an application of great interest to investigators engaged in designing ZFs for use in gene therapy. Having established that 6-finger proteins are amenable for targeting dsDNA in the context of split-protein assays, we turned to demonstrating their utility in specifically targeting sites of DNA methylation. The 18bp ZF targeting
domains offer the potential for the unique recognition of specific sites of methylation in genomic DNA, which was not possible with our previous designs incorporating three-finger ZFs.

**Interrogation of dsDNA methylation**

[00155] As previously discussed, cytosine methylation at CpG dinucleotides is dependent on the transfer of a methyl group from S-adenosylmethionine to the C5-position of cytosines within these CpG dinucleotides (Fig. 23, panel a), a mechanism regulated by DNA methyltransferases (121). Previously, split-protein reassembly systems for GFP and β-lactamase have been used to directly detect site-specific determination of dsDNA methylation utilizing a mCpG targeting domain, MBD2, attached to one half and a three-finger ZF attached to the other half (Fig. 23, panel b) (37, 58). These studies had demonstrated a geometric/length dependence on GFP and β-lactamase reassembly, thus before evaluating our new 6-finger targeting domains, we evaluated the effect of both distance and geometry on split-luciferase reassembly.

**Table 4.** Protein fusions, nucleic acid binding domains, and nucleic acid targets used.

<table>
<thead>
<tr>
<th>Reassembly Pairs</th>
<th>Nucleic Acid Binding Domains</th>
<th>Nucleic Acid Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aart-NFuc</td>
<td>Aar1</td>
<td>GCGTACGGTGGCGAGATGAGGGGAAAAGCCCGGTACC (SEQ ID NO:72)</td>
</tr>
<tr>
<td>CFuc-Zi268</td>
<td>Zi268</td>
<td></td>
</tr>
<tr>
<td>Aart-NFuc</td>
<td>Aart</td>
<td>GCGTACGGGCGAGCCGGGACGTGGTGGGAAAAGCCCGGTACC (SEQ ID NO:73)</td>
</tr>
<tr>
<td>CFuc-E2C</td>
<td>E2C</td>
<td></td>
</tr>
<tr>
<td>MBD2-NFuc</td>
<td>MBD2 (147-215)</td>
<td>GGCGTAAGCGCCGCACGCCTACC (SEQ ID NO:74)</td>
</tr>
<tr>
<td>CFuc-Zi268</td>
<td>Zi268</td>
<td></td>
</tr>
<tr>
<td>MBD2-NFuc</td>
<td>MBD2 (147-215)</td>
<td>GGCGTAAGCGCCGCACGCCTACC (SEQ ID NO:75)</td>
</tr>
<tr>
<td>CFuc-E2C</td>
<td>E2C</td>
<td></td>
</tr>
<tr>
<td>MBD2-NFuc</td>
<td>MBD2 (147-215)</td>
<td>GGCGTAAGCGCCGCACGCCTACC (SEQ ID NO:76)</td>
</tr>
<tr>
<td>CFuc-1.Cro</td>
<td>1.Cro (1-88)</td>
<td>GGCTAGCGACTATACCGGGGTTGATACAGCC (SEQ ID NO:77)</td>
</tr>
<tr>
<td>MBD2-NFuc</td>
<td>MBD2 (147-215)</td>
<td>GGCTAGCGACTATACCGGGGTTGATACAGCC (SEQ ID NO:78)</td>
</tr>
<tr>
<td>CFuc-MBD2</td>
<td>MBD2 (147-215)</td>
<td>GGCTAGCGACTATACCGGGGTTGATACAGCC (SEQ ID NO:79)</td>
</tr>
<tr>
<td>APLF-NFuc</td>
<td>APLF (376-441)</td>
<td>poly(ADP-ribose)</td>
</tr>
<tr>
<td>CFuc-APLF</td>
<td>APLF (376-441)</td>
<td></td>
</tr>
<tr>
<td>PBSII-NFuc</td>
<td>PBSII</td>
<td>GCGTACGGTGGCGAGATGAGGGGAAAAGCCCGGTACC (SEQ ID NO:80)</td>
</tr>
<tr>
<td>CFuc-Zi268</td>
<td>ZI268</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Methylated dsDNA targets and Aart competitive hairpins used

<table>
<thead>
<tr>
<th>Methylated dsDNA Targets</th>
<th>Aart Competitive Hairpins</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCpG-0-Zi268 GCGTA_CCGGCCACGCCACCG (SEQ ID NO:81)</td>
<td>Consensus GCATGTAAGGGAAAAAGCCCGGCGTGCTCCTCGGGCCTTTTACATGC (SEQ ID NO:87)</td>
</tr>
<tr>
<td>mCpG-2-Zi268 GCGTA_CGTGGCCACGCCACCG (SEQ ID NO:82)</td>
<td>Consensus ATA GCCATGTAAGGAATAGCCCCGGGCTCTCCTCGGGCCTTTATCCTACATG (SEQ ID NO:88)</td>
</tr>
<tr>
<td>mCpG-3-Zi268 GCGTA_CGTAAGGACCCAGCCACCCGG (SEQ ID NO:83)</td>
<td>Designed GCATGTAAGGAATAGCCCCGGGCTCTCCTCGGGCCTTTTACATGC (SEQ ID NO:89)</td>
</tr>
<tr>
<td>mCpG-4-Zi268 GCGTA_CGTAAGGACCCAGCCACCCGG (SEQ ID NO:84)</td>
<td>Designed ATA GCCATGTAAGGAATAGCCCCGGGCTCTCCTCGGGCCTTTTACATGC (SEQ ID NO:90)</td>
</tr>
</tbody>
</table>

[00156] Aart and E2C mediated firefly luciferase reassembly. Duplicate 25 μL reactions were carried out in Flexi-Rabbit Reticulocyte according to manufacturer’s protocol using 0.2 pmols of Aart-NFluc(residue 2-416) and 0.1 pmols of CFLuc(398-550)-E2C, 10 μM ZnCl₂, 0.5 μL RNasin Plus (Promega), and 1.25 μL of either 10 nM E2C-1-Aart target or water. Translations were incubated for 90 minutes at 30°C and assayed by adding 80 μL of Steady-Glo™ Luciferase Assay System (Promega) to 20 μL of translated lysate. Light emission was monitored 1 minute after substrate addition using a Turner TD-20e luminometer with a 3 sec. delay and a 10 sec. integration time.

[00157] Experiments were carried out to confirm that reassembly of the split luciferase was sequence specific. There was essentially no signal generated in response to the inclusion of sheared herring sperm DNA rather than target, and supplementation of the assay containing target DNA did not result in increase reporter reassembly.
[00158] **PBSII and Zif268 mediated firefly luciferase reassembly.** Duplicate 25 µL reactions were carried out in Flexi-Rabbit Reticulocyte according to manufacturer’s protocol using 0.2 pmols of PBSII-NFluc(residue 2-416) and 0.2 pmols of CFluc(398-550)-Zif268, 10 µM ZnCl₂, 0.5 µL RNasin™ Plus (Promega), and 1.25 µL of decreasing concentrations of Zif268-0-PBSII target dsDNA (100 nM – 1 nM) or water. Translations were incubated for 90 minutes at 30°C and assayed by adding 80 µL of Steady-Glo™ Luciferase Assay System (Promega) to 20 µL of translated lysate. Light emission was monitored 1 minute after substrate addition using a Turner TD-20e luminometer with a 3 sec. delay and a 10 sec. integration time.

[00159] **Aart and E2C mediated firefly luciferase reassembly in the presence of sheared Herring Sperm (HS) DNA.** Duplicate 25 µL reactions were carried out in Flexi-Rabbit Reticulocyte according to manufacturer’s protocol using 0.2 pmols of Aart-NFluc(residue 2-416) and 0.4 pmols of CFluc(398-550)-E2C, 10 µM ZnCl₂, 0.5 µL RNasin™ Plus (Promega), and either 1.25 µL of 30 nM E2C-1-Aart (Target), Target plus 1.25 µL of 91.65 ng/µL HS-DNA, 1.25 of 91.65 ng/µL HS-DNA, or water. Translations were incubated for 90 minutes at 30°C and assayed by adding 80 µL of Steady-Glo™ Luciferase Assay System (Promega) to 20 µL of translated lysate. Light emission was monitored 1 minute after substrate addition using a Turner TD-20e luminometer with a 3 sec. delay and a 10 sec. integration time.

[00160] **Methylation mediated luciferase reassembly.** For initial MBD2 and E2C mediated luciferase reassembly duplicate 25 µL reactions were carried out in Flexi-Rabbit Reticulocyte according to manufacturer’s protocol using 0.2 pmols of MBD2-NFluc(residue 2-416) and of CFluc(398-550)-E2C, 10 µM ZnCl₂, 0.5 µL RNasin™ Plus (Promega), and 1.25 µL of either 1 µM mCpG-2-E2C target or water. For MBD2 and Zif268 duplicate 25 µL reactions were carried out in Flexi-Rabbit Reticulocyte according to manufacturer’s protocol using 0.2 pmols of MBD2-NFluc(residue 2-416) and of CFluc(398-550)-Zif268, 10 µM ZnCl₂, 0.5 µL RNasin™ Plus (Promega), and 1.25 µL of decreasing concentrations of mCpG-2-Zif268 target (100 nM – 1 nM) or water. Translations were incubated for 90 minutes at 30°C and
assayed by adding 80 µL of Steady-Glo\textsuperscript{TM} Luciferase Assay System (Promega) to 20 µL of translated lysate. Light emission was monitored 1 minute after substrate addition using a Turner TD-20e luminometer with a 3 sec. delay and a 10 sec. integration time.

[00161] Translations with MBD2-NFluc(2-416) (SEQ ID NO: 32-33) and CFLuc(398-550)-Zif268 (SEQ ID NO: 26-27) in the presence of a series of methylated dsDNA targets incorporating increasing distances between the mCpG and Zif268 binding sites was tested and showed a profile similar to split-lactamase but distinct from split-GFP (Fig. 23, panel c). Essentially no activity is observed when the DNA binding domains are directly adjacent to each other (0 bp separation), a likely result of the inability of MBD2 and Zif268 to simultaneously bind their DNA targets. Beyond 1 bp the pattern, likely recapitulate the helical nature of dsDNA, as also observed when fragmented beta-lactamase or GFP are used as signaling domains in place of luciferase (Fig. 23, panel c, center and righthand bars in each trio of bars, respectively). The manner in which these 3 fragmented systems are assembled provides some insight with respect to the differences in signal generation observed. Both split-GFP and split-beta-lactamase were rationally designed, such that the point of dissection and new-protein attachment sites lie between loops. GFP is perhaps the most sensitive to geometrical constraints as the secondary structural elements adjacent to the dissected loop are beta-strands whereas the region adjacent to the dissection site in beta-lactamase is less structured (15,16). In contrast the firefly luciferase fragments, selected from a library, have the interacting proteins fused to the native wild type termini (27) that are separated by ~40 Å as seen in the crystal structure. This inherent separation between protein attachment sites likely allows split-firefly luciferase to tolerate targets at larger separation distance compared to GFP and beta-lactamase though surprisingly still showing highest signal at a 1 bp separation distance. These initial results suggest that the direct detection of sites of specific promoter methylation will likely be primarily dictated by the specificity of the sequence specific dsDNA binding domain, with the site of adjacent methylation lying anywhere between 2-10 base pairs away. Thus, this lack of stringency may be a boon, allowing for greater latitude in choosing sites for targeting given that the design/selection of zinc fingers, though very powerful, is still not capable of providing ZFs capable of targeting any dsDNA sequences with high selectivity. Noting that a 6-
finger ZF targeting domain (18 bp) ensures unique targeting within the genome we next tested if new domains could function in the site-specific determination of dsDNA methylation.

[00162] Given the enhancement in dsDNA detection gained by the use of 6-finger ZFs discussed above, we next utilized the 6-finger ZF E2C for the site-specific determination of dsDNA methylation. Initial experiments demonstrated the methylated dsDNA dependent reassembly of fragmented firefly luciferase-MBD2/E2C fusions (not shown). To determine the minimal amount of methylated dsDNA needed to generate an observable signal over background, cell-free translations were initiated using 0.2 pmols of mRNA corresponding to MBD2-NFluc (SEQ ID NO: 32-33) and CFluc-E2C (SEQ ID NO: 28-29) in the presence of decreasing amounts of the methylated dsDNA target mCpG-2-E2C (100pM-10pM) (Fig. 23, panel d). These experiments clearly show that 5 pM (500 amols) of mCpG-2-E2C is detectable above background and that activity scales linearly across the range of concentrations tested. When compared to the analogous platform utilizing the 3-finger ZF Zif268 that is capable of detecting 10 fmols of methylated dsDNA (not shown), these results demonstrate a 20-fold improvement in the minimal amount of dsDNA needed to generate signal above background. Importantly, this demonstrates that the use of a single 6-finger ZF capable of specifically recognizing 18 contiguous base pairs potentially allows for the detection of an unique promoter within the human genome, a feature not available using a single 3-finger ZF.

**DNA-Mediated RNA Detection Experiments**

**Note:** The cloning and initial characterization of the 6-fingers, E2C and Aart with NFluc and CFluc, are in the accompanying manuscript that focuses upon a different topic.

<table>
<thead>
<tr>
<th>Table 3.</th>
<th>in vitro transcription primers and DNA oligonucleotides used in the zinc finger-mediated RNA detection method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VEGF FWD</strong></td>
<td>5'GCAGCTATTACGACTCATACTAGGCATACGAAAGTTGAGGATCGTCTCATGCAG (SEQ ID NO:60)</td>
</tr>
<tr>
<td><strong>VEGF REV</strong></td>
<td>5'CTTCTTTGCTGTGGCATTCAGATTTGTTGCTGTAGGAACG (SEQ ID NO:61)</td>
</tr>
<tr>
<td><strong>HER2 FWD</strong></td>
<td>5'GCAGCTATATTACGACTCATACTAGGGCTAGAGCACCCAGATCGTCGCGG (SEQ ID NO:62)</td>
</tr>
<tr>
<td><strong>HER2 REV</strong></td>
<td>5'GTGCTTGGGCCCCTGCAG (SEQ ID NO:63)</td>
</tr>
<tr>
<td><strong>Hairpin-guides</strong></td>
<td></td>
</tr>
<tr>
<td>Pum1-E2C</td>
<td>5'GAGGGGGGCGGAGCCGGAGTGGGTTCCCGGAGGGGCTCGCGGCCCCTCAAAAACATAT ATACCGC (SEQ ID NO:64)</td>
</tr>
<tr>
<td>Pum2-Aart</td>
<td>5'GCCGCTATATCAAAAAACTCCGGTGTCCCACGTCGATGTAGGGGAAAG</td>
</tr>
</tbody>
</table>
Hairpin-guides are colored according to the following regions: sequences forming zinc finger binding site hairpins are bold, the hairpin loop region is underlined, and the sequence complementary to the target is italicized. Numbering in the names of the VEGF and HER2 hairpin-guides is based on the first base of the transcribed mRNA target being considered as position one. The T7 promoter present in the in vitro transcription primers is blue. The single base mutation in the ssDNA G to T target is underlined.

**Generation of target mRNA**

[00163] VEGF dsDNA was PCR amplified from an existing plasmid, pQE30-VEGF, which contained nucleotides 109-403 of VEGF cDNA, isoform 165. This 295 nucleotide region was amplified using the primers indicated in Table 3 under the heading in vitro transcription primers. Note the exclusion of a KOZAK sequence and start codon in the FWD primer and a stem-loop structure in the REV primer, as the VEGF mRNA is not meant to be subsequently translated. in vitro transcription was carried out according to the manufacturer's instructions (using 2-3 μg dsDNA template, as limited by PCR yields), and purity was assessed with agarose gel electrophoresis. To select a VEGF mRNA region to target, the following considerations were made: 1) High melting temperature (thermodynamic stability) between guide and target, 2) no tetr glycine motif, 3) no secondary structure in guide region, 4) no (or minimal) complementarity to other sites in the target, 5) presence of a suitable adjacent binding site (currently we have only employed 1-4 nucleotide separations, and the use of more distal sites has not been investigated), 6) site accessibility based on secondary structure prediction (mfold, internet address frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) of VEGF transcript. Based on these considerations, the following regions were chosen: 60-78 and 81-99 of the 295 nucleotide VEGF transcript.
[00164] HER2 mRNA was generated essentially as described above. HER2 dsDNA (nucleotides 480-681) was PCR amplified from an existing plasmid, pSGHVO-HER2, which contained nucleotides 1-1983 of the HER2 extracellular domain, using primers indicated in Table 3. Due to the presence of contaminating PCR products, a gel extraction was performed using a QIAquick PCR purification kit (Qiagen), resulting in isolation of a pure product, as visualized by agarose gel electrophoresis. Considerations regarding targetable areas were as described above, although difficulty was encountered in finding guides without secondary structure. The following regions were chosen: 100-118 and 122-140 of the 201 nucleotide HER2 transcript.

[00165] Annealing of ternary complex was achieved as follows. DNA hairpins (Table S3, Hairpin-guides) were pre-formed in NEBuffer Sall (10 mM Tris-HCl, pH 7.9, 150 mM NaCl, 10 mM MgCl₂, 1 mM dithiothreitol) by an annealing procedure consisting of heating to 95°C for 7 minutes, cooling at a rate of 1°C/min to 56°C for 5 minutes, followed by cooling at a rate of 1°C/min to 25°C for 10 minutes, followed by storage at -20°C. Targets (RNA, DNA, or DNA(G to T) targets; VEGF mRNA; HER2 mRNA; or Sall buffer blank) were heat denatured in NEBuffer Sal1 supplemented with 0.8 U/μL RNASin™ at 90°C for 7 minutes, followed by cooling at a rate of 6°C/min to 37°C for 10 minutes. Heat denaturation was followed directly by dilution of the target (or buffer blank) into the corresponding set of pre-formed hairpins, which was then held at 37°C for 3 hours. We have since reduced the holding time to 30 minutes, with no adverse effects. The annealed ternary complex stock was stored and held at all times at 4°C.

[00166] Protein reassembly was accomplished according to the following protocol. NFluc-Aart was generated by PCR amplification of Aart from an existing plasmid, followed by ligation into a plasmid containing NFluc. CFluc-E2C was generated by ligating an E2C dsDNA insert into a plasmid containing CFluc. Details regarding these cloning procedures are available (58).

[00167] Translation experiments occurred as described above with the addition of 10 μM ZnCl₂. Optimal conditions required 0.2 pmol of each mRNA, NFluc-Aart and CFluc-E2C, and annealed ternary targets or a “background” were added after completion of the translation reaction. Background consisted of buffer only and of the
hairpin-guides without target RNA. The presence of non-cognate nucleic acids (i.e. hairpin-guides) consistently resulted in a reduction of background as compared to buffer alone. This may be attributable to prevention of non-specific interactions between the split-Fluc constructs. Luminescence readings were taken following 30 minutes of incubation with the ternary target or background at 4°C. 80 μL of Steady-Glo™ reagent was added to 20 μL of each reaction (final target concentration = 1 nM) and allowed to incubate at room temperature for 1 minute prior to acquiring luminescence readings on a Turner TD-20e Luminometer with a 10 second integration time, except in the case of HER2 readings, which were collected on a Turner Biosystems 20/20th luminometer. Two duplicate translation reactions performed simultaneously. Luminescence readings were averaged and normalized to one.

**VEGF detection**

[00168] The lower limits of VEGF mRNA detection were determined according to protocols indicated above, using target concentrations of 10 nM, 1 nM, 100 pM, 10 pM, 1 pM, or buffer. Background in this case is buffer only. The buffer luminescence was subtracted from each reading in the presence of decreasing concentrations of target mRNA, followed by normalization of the signal to one. The presence of 10 pM (1 fmol) VEGF mRNA definitively produces signal over background, while the 1 pM (100 amol) sample is very close to background levels.

**Direct detection of DNA binding by dimeric DNA binding domains.**

[00169] Having validated the use of monomeric DNA binding domains (DBDs), including ZFs and MBD, we next attempted the development of a platform utilizing dimeric sequence specific DBDs. This would not only serve to augment the available DBDs, but also potentially allow for detecting any dsDNA sequence that can be targeted by a known natural DBD, whether monomeric or dimeric. As an initial test we interrogated DNA binding by the bacteriophage Lambda Cro repressor protein (λ-Cro). This prototypical helix-turn-helix DBD recognizes a 17 base pair dsDNA sequence (two copies of a 7-mer palindromic half-site separated by 3 unrecognized base pairs) through the formation of an obligate DNA binding homodimer, in which each monomer binds to the half-site 5'-TATCACC-3'. To test
the use of this dimeric DNA binder, the mCpG dinucleotide recognized by MBD2 was placed upstream of the 17 base pair λ-Cro recognition site separated by a 2 base pair spacer, generating the dsDNA target mCpG-2-λ.Cro. Molecular modeling suggested that attaching CFLuc to the N-terminus of λ-Cro and attaching MBD2 to the N-terminus of NFluc would result in a viable protein-DNA complex positioned for reassembly of fragmented firefly luciferase (Fig. 23, panel e). It is noted that studies with split-GFP utilizing similar targeting domains were unsuccessful in generating fluorescence.

[00170] To assess the ability of split-luciferase to report on λ-Cro binding its cognate dsDNA, translations with mRNA corresponding to MBD2-NFluc and CFLuc-λ.Cro in the presence of the methylated (m), non-methylated dsDNA target CpG-2-λ.Cro, or no DNA (Fig. 23, panel f) were tested. A reproducible 22-fold increase in luminescence signal was observed in the presence of 10 nM mCpG-2-λ.Cro compared to the non-methylated target. These results demonstrate for the first time that dimeric DBD domains can be utilized in split-luciferase reassembly to potentially allow for targeting sites of methylation not easily amenable to ZFs. Additionally, this approach provides the possibility for interrogating the interaction between dimeric transcription factors and their target DNA sequences.

Determination of global DNA methylation.

[00171] These approaches towards the determination of site-specific methylation status are useful as diagnostic and research tools, however they do not provide a means to determine global changes in genomic methylation, which have been observed in both disease progression (102) and aging (122). Furthermore, a global methylation sensor would provide a convenient means for interrogating the activity of methyltransferases and demethylating enzymes, as well as allow for a new method for determining the activity of small molecules that perturb their activity. Thus, we asked whether a sensor containing MBD2 to both the N-terminal and C-terminal fragments of firefly luciferase creating the fusions MBD2-NFluc and CFLuc-MBD2 would allow for the detection of any dsDNA target containing multiple methylation sites (Fig. 24, panel a) with the caveat that we would statistically expect to see 50% of the total possible signal. As our first test, cell-free translations of MBD2-NFluc and
CFluc-MBD2 were carried out in the presence of dimethylated dsDNA targets containing either a short separation distance of 6 bp or a long separation distance of 21 bp (our maximum calculated distance for split-luciferase assembly is ~140 Å) between methylated CpG sites. Both the 6 and 21 bp dimethylated targets allowed for reproducible increase in luminescence as compared to no DNA (Fig. 24, panel b). Titrations were carried out to determine the minimal amount of methylated dsDNA needed to generate an observable signal for the 6 bp site, (Fig. 24, panel c). These experiments show that 0.2 nM (20 fmols) mCpG-6-mCpG is detectable above background. Thus this a conceptually new approach, that is utilizing the same targeting domain, for detecting DNA modification and may find utility in detecting global levels of methylation as well as the activity of associated enzymes and their inhibitors.

**Direct detection of poly(ADP-ribose) using a poly(ADP-ribose)-binding zinc finger.**

[00172] Building on our success in using a single domain to detect multiple sites of modification, we next used this strategy for interrogating poly(ADP-ribose)ylation. This ubiquitous post-translational modification has been linked to carcinogenesis and is considered a possible marker for cancer detection (123). We chose the poly(ADP-ribose) (PAR) binding domain from aprataxin PNK-like factor (APLF) for the detection of PAR. APLF is a protein involved in the cellular response to DNA damage (124-126) and contains two putative Cys₂His₂ ZF domains capable of binding PAR with high affinity (127). Our designed sensor incorporated the zinc finger domain of APLF (residues 376-441) to both the N-terminal and C-terminal fragments of firefly luciferase to create APLF-NFluc and CFluc-APLF. We envisioned that the simultaneous binding of each APLF-luciferase fusion to poly(ADP-ribose) would result in firefly luciferase reassembly (Fig. 25, panel a). Translations with mRNA for APLF-NFluc and CFluc-APLF in presence of poly(ADP-ribose), non-cognate Zif268-0-PBSII dsDNA, and no added target. Additional complementary controls were also carried out where mRNA corresponding to the ZF PBSII-NFluc and CFluc-Zif268 were carried out with added poly(ADP-ribose) or Zif268-0-PBSII (Fig. 25, panel b). A 25-fold increase in luminescence was observed for APLF-NFluc/CFluc-APLF pair in the presence of 10 nM PAR, while no
luminescent signal was observed in the presence of dsDNA, Zif268-0-PBSII. Conversely, translations containing the dsDNA binding ZFs PBSII-NFLuc/CFLuc-Zif268 did not generate signal in the presence of PAR. Additionally, cell-free translations in the presence of decreasing amounts of poly(ADP-ribose) showed that 0.125 pM (12.5 fmols) of PAR is detectable above background (Fig. 25, panel c). Thus, this represents the first split-luciferase sensor for sensitive detection of PAR. This method as well as variations with other split-protein sensors may provide a also provide a means for PAR detection within a cell and may also be a valuable addition to the tool-kit for interrogating of the activity and chemical perturbation of the PAR associated proteins, PARP and PARG.

[00173] Considerable effort has been focused on the development of new and enabling technologies to elucidate the function of biological macromolecules (59,128, 129). We build upon these efforts with the development of a fragmented-firefly luciferase tool-kit that allows for the rapid, sensitive, and direct interrogation of specific dsDNA sequences, site-specific and overall DNA methylation, monomeric and dimeric DNA binding domains, and the presence of poly(ADP-ribose). Specifically, tandem 6-finger ZF based sensors will potentially allow for the rapid and direct detection of attomole quantities of a specific 36-mer dsDNA sequence while the use of 6-finger ZF in conjunction with a methylation specific binding domain may allow for the sensitive characterization of unique sites of dsDNA methylation. Furthermore, the incorporation of λ.Cro provides the first example of the use of dimeric dsDNA binding domains in split-protein systems, which not only increases the repertoire of available targeting domains but may also provide a means for directly probing DNA binding of this important class of transcription factors. Finally, we provide two conceptually new approaches utilizing split-proteins that allow for the direct and sensitive detecting global levels of dsDNA methylation and poly(ADP-ribosyl)ation. Thus, these new methods provide a valuable tool-kit to examine DNA and DNA-associated chemical modifications and perhaps more importantly, provide access to rapid assays for related enzymes and their small molecule perturbants.

[00174] Plasmid construction and mRNA production. Fusion protein constructs used in this study are shown in Tables 6-28. DNA coding for firefly luciferase fragments were generated by PCR and cloned into the pETDuet-1 vector
(Novagen, Madison, WI) using standard cloning techniques. Fragments encoding the nucleic acid-binding proteins were generated by PCR. Fusion protein constructs were generated using standard cloning techniques.

[00175] mRNA for cell-free assays was generated as follows: PCR fragments corresponding to the desired fusion constructs were generated using a forward primer containing a T7 RNA polymerase promoter and a Kozak sequence and a reverse primer containing a 3' hairpin loop. The purified PCR fragments were used as the template for in vitro transcription using the Ribomic Large Scale RNA Production System-T7 (Promega) following the manufacturer's protocols.

[00176] Target DNA preparation. All nucleic acid targets (Supplementary Table 2 – please provide) were obtained from IDT. All dsDNA targets were annealed as previously described (23,58). Hairpin DNA targets were annealed in 1x BamHI buffer by heating at 95°C for 7 minutes followed immediately by cooling on ice.

[00177] Reassembly of fragmented firefly luciferase-zinc finger fusion. Duplicate 25 μL translation reactions were carried out in Flexi-Rabbit Reticulocyte Lysate (Promega) according to the manufacturer's protocol using 0.2 pmols of Aart-NFluc (residues 2-416) and CFLuc(residues 398-550)-Zif268 mRNA, 10 μM ZnCl$_2$, 0.5 μL of RNasin Plus (Promega), and either 1.25 μL 10 nM Zif268-2-Aart target dsDNA or no dsDNA. For the case where two 6-finger zinc fingers were used, duplicate translations were carried out in Flexi-Rabbit Reticulocyte Lysate using 0.2 pmols of Aart-NFluc (residues 2-416) and 0.1 pmols of CFLuc(residues 398-550)-E2C mRNA, 10 μM ZnCl$_2$, 0.5 μL of RNasin$^{TM}$ Plus (Promega), and the concentrations of E2C-1-Aart target dsDNA indicated. Translations were incubated at 30 °C for 90 minutes and assayed by adding 80 μL of Steady-Glo$^{TM}$ Luciferase Assay System (Promega) to 20 μL of translated lysate. Light emission was monitored 1 minute after STEADY-GLO$^{TM}$ addition using a Turner TD-20e luminometer with a 3 second delay and a 10 second integration time.

[00178] Interrogation of Aart specificity. Duplicate 25 μL translations were carried out in Rabbit Reticulocyte Lysate according to the manufacturer's protocol using 0.2 pmols of Aart-NFluc(residues 2-416) and 0.4 pmols of CFLuc(residues 398-550)-E2C mRNA, 10 μM ZnCl$_2$, and 0.5 μL of RNasin Plus (Promega) and allowed to
incubate for 90 minutes at 30 °C in the presence of 300 pM Aart-1-E2C dsDNA target and increasing concentrations of hairpin DNA. Samples were assayed for luciferase activity as described above.

[00179] The effects of distance and proximity of fragmented luciferase reassembly were explored as discussed below. Duplicate 25 μL translations were carried out in rabbit reticulocyte lysates (Promega) according to the manufacturer’s protocol using 2 pmols of mRNA corresponding to MBD2-NFluc(residues 2-416) and CFLuc(residues 398-550)-Zif268, 10 μM ZnCl₂, and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 1.25 μL of 500 nM methylated target dsDNA containing either 0, 1, 2, 3, 6, or 10 basepairs between the mCpG and Zif268 binding sites. Samples were assayed for luciferase activity as described above.

[00180] Reassembly of fragmented luciferase utilizing MBD2 and E2C was performed as follows. Duplicate 25 μL translations were carried out in rabbit reticulocyte lysate according to the manufacturer’s protocol using 0.1 pmols of mRNA corresponding to MBD2-NFluc(residues 2-416) and CFLuc(residues 398-550)-E2C, 10 μM ZnCl₂, and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 1.25 μL of decreasing concentrations of the dsDNA target mCpG-2-E2C. Samples were assayed for luciferase activity as described above.

[00181] Reassembly of fragmented luciferase utilizing MBD2 and λ-Cro was examined using the following protocol. Duplicate 25 μL translations were carried out in rabbit reticulocyte lysates according to the manufacturer’s protocol using 2 pmols of MBD2-NFluc(residues 2-416) and CFLuc(residues 398-550)-λ-Cro mRNA, and 0.5 μL of RNasin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 1.25 μL of 1 μM of the dsDNA target mCpG-2-λ-Cro(m), mCpG-2-λ-Cro(u), or water. Samples were assayed for luciferase activity as described above.

[00182] Reassembly of fragmented luciferase utilizing two MBD2s was carried out as follows. Duplicate 25 μL translations were carried out in rabbit reticulocyte lysates according to the manufacturer’s protocol using 2 pmols of MBD2-NFluc (residues 2-416) and CFLuc (residues 398-550)-MBD2 mRNA, and 0.5 μL of RNasin™ Plus
(Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 1.25 µL of 1 µM of the dsDNA targets mCpG-6-mCpG, mCpG-21-mCpG, or water. Samples were assayed for luciferase activity as described above.

[00183] Reassembly of fragmented luciferase utilizing two poly(ADP-ribose) binding zinc finger domains was examined as follows. Duplicate 25 µL translations were carried out in rabbit reticulocyte lysates according to the manufacturer’s protocol using 1.3 pmols of APLF-NFluc(residues 2-416) and CFluc(residues 398-550)-APLF, or 2 pmols of PBSII-NFluc(residues 2-416) (SEQ ID NO: 24-25) and CFluc(residues 398-550)-Zif268 mRNA (SEQ ID NO: 26-27), 10 µM ZnCl₂ and 0.5 µL of RNASin™ Plus (Promega) and allowed to incubate for 90 minutes at 30 °C in the presence of 1.25 µL of 1 µM poly(ADP-ribose) (BioMol International, Plymouth Meeting, PA), 1 µM Zif268-0-PBSII, or water. Samples were assayed for luciferase activity as described above.

[00184] Regulation at the level of DNA is controlled by protein binding factors as well as chemical modifications such as DNA methylation, histone acetylation, and protein poly(ADP-ribose)lation. Alterations to “normal” DNA and its associated factors, which include mutations and chromosomal translocations, aberrant methylation, and deregulated poly(ADP-ribose)lation, can provide diagnostic signatures for the status of a cell and are often disease markers. Thus, there is much interest in developing new reagents for the direct detection of dsDNA, interrogation of transcription factor/DNA binding, monitoring DNA-methylation, and measuring poly(ADP-ribose)lation. Towards this goal we have developed a simple and general cell-free split-luciferase system that can be decorated with user-defined protein targeting modules allowing for the detection of unique dsDNA sequences at the attomolar level, site-specific and overall DNA methylation levels, dimeric transcription factor DNA binding events, and the direct measurement of poly(ADP-ribose). These new reagents are not only useful for chemical diagnostics, but they also provide a tool-kit for discovering specific small molecules that can perturb DNA regulation.

[00185] We have systematically designed a class of split-firefly luciferase sensors for ssRNA detection, each utilizing different detection domains. The first class comprises sequence-specific pumilio domains, which although successful in
detecting ssRNA, are limited by the necessity of designing new domains for each ssRNA target of interest. Thus, our second class of sensors incorporated the RNA binding domain of argonaute, which specifically recognizes 2-nucleotide, 3’ overhangs of dsRNA. By introducing short guide sequences of user-defined ssRNA, we successfully detected cognate ssRNA target when using argonaute in conjunction with pemilio domains. However, the use of two argonaute domains, which would allow one to recognize any ssRNA of interest did not provide sufficient signal. Building on these designs our third and most general design provides a convenient method of both ssRNA and ssDNA detection through the use of ssDNA guides tethered to high affinity zinc finger DNA binding hairpin guides. When a target contains adjacent sequences complementary to the guides, the hairpins are brought into proximity, allowing for zinc finger binding and luciferase reassembly. This general approach has been used to sensitively (~100 amol) and specifically detect physiologically relevant targets, including VEGF and HER2 mRNA.

[00186] In the following tables, bold type represents residues derived from luciferase, and italic type represents sequences derived from an attached protein.

[00187] Sequences of split reporter fusions useful in the detection of gp120 and her2 with protein and natobody conjugated to luciferase halves are shown in Tables 6-9.

Table 6. CFLuc-17b (see also SEQ ID NOs:8-9).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATGATTGTCCGATTGTTATGAAAAATTCCGCCAGCGAGACATCGCTTGGATTGAACAGAGATGGAA</td>
<td>CFLuc-17b (see also SEQ ID NOs:8-9)</td>
</tr>
</tbody>
</table>

...
Table 7. CD4-NF Luc (see also SEQ ID NOs:10-11).

<table>
<thead>
<tr>
<th>Seq ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>atgggcagacagccatcaccatcatccacacacccacacccacagctggaagtggctgggcaaa</td>
<td>CD4-NF Luc sequence</td>
</tr>
<tr>
<td>MG S S H H H S D P K V V L G K</td>
<td></td>
</tr>
<tr>
<td>aaggggatacaagtggaactctacgcagccctccagggagaagagcaactacaactacctccac</td>
<td></td>
</tr>
<tr>
<td>K G D T V E L T C T A S Q K K S I Q P H</td>
<td></td>
</tr>
<tr>
<td>tggaaaaactccacaccagataagattctgggaactcagggctctctcttaactaaaggt</td>
<td></td>
</tr>
<tr>
<td>W K N S N Q I K I L G N Q G S F L T K G</td>
<td></td>
</tr>
<tr>
<td>ccacatcaagctgcatagctcggctgtacccaagagagctccttttggaaccaggaacttc</td>
<td></td>
</tr>
<tr>
<td>F S K L N D R A D S R R S L W D Q G N F</td>
<td></td>
</tr>
<tr>
<td>cccctgatcatacataggtattattagagagactagctacatactcagttgagggag</td>
<td></td>
</tr>
<tr>
<td>P L I K N L K I E D S D T Y T G E</td>
<td></td>
</tr>
<tr>
<td>gaccaggaagggaggtcaattctgtcagattgctcgcctggacacaccacaccac</td>
<td></td>
</tr>
<tr>
<td>D R K E D L V F G L T A N S D T H</td>
<td></td>
</tr>
<tr>
<td>ctgcctcagccaggaagctcccctgccatttgagagggcctccctgttagatgacccctca</td>
<td></td>
</tr>
<tr>
<td>L L C Q G Q T L T L E S P G S S S S</td>
<td></td>
</tr>
<tr>
<td>gttcataatggagttcagcagctcgatagctcggctgtacccaagagagctccttttggaaccaggaacttc</td>
<td></td>
</tr>
<tr>
<td>V Q C R S P R G K N I Q G G K T L S V S</td>
<td></td>
</tr>
<tr>
<td>cagctggaagctccagagattagtggcactctggatggtacacttcagacaccccaggaagag</td>
<td></td>
</tr>
<tr>
<td>Q L E D L O S G T W T C T V L Q N Q K K</td>
<td></td>
</tr>
<tr>
<td>tgtagctcaaatagacactctcggctgtacctctccagagggcctccacaccctgggaggt</td>
<td></td>
</tr>
<tr>
<td>V E F K I D I V V L A F Q K A S T G G G</td>
<td></td>
</tr>
<tr>
<td>gccggttcagcagcgggtgggttctgtgtgggtgtacggaacagcaccagggaaaaaaaaatctaaag</td>
<td></td>
</tr>
<tr>
<td>K G P A F P Y P L E D G T A G E Q L H K</td>
<td></td>
</tr>
<tr>
<td>gcttagaagagataggccctgcttgggaaaactcttcattttacagagtgcacatagag</td>
<td></td>
</tr>
<tr>
<td>A M K R Y A L V P G T I A F T D A H I E</td>
<td></td>
</tr>
<tr>
<td>gtgaacatatcagtcgacacagaccacagtgtcctgggtcagggagcagctagaaa</td>
<td></td>
</tr>
<tr>
<td>V N I T Y A E Y F E M S V R L A E A M K</td>
<td></td>
</tr>
<tr>
<td>cgataacatatcagtcgacacagaccacagtgtcctgggtcagggagcagctagaaa</td>
<td></td>
</tr>
<tr>
<td>R Y G L N T N H R I V V C S E N S L Q P</td>
<td></td>
</tr>
<tr>
<td>tttagcgctttgaggggcgtttatattttacagagtgcagggagcagctagaaa</td>
<td></td>
</tr>
<tr>
<td>F M P V L G A L F I G V A V A P A N D I</td>
<td></td>
</tr>
</tbody>
</table>
Table 9. 2C4-NFluc (see also SEQ ID NOs:14-15)

atgggccagcagcataacctatcatctacacacgccagctggagaaggtccagctgtt
M G S S H S D P L E V Q L V L
agggcggccagcctgttgagcgcagctggagaaggtccagctgtt
D Y W C Q C T L V T V S S R L L G S
agggcggaggtggcttcgctggagaaggtccagctgtt
D Y W C Q C T L V T V S S R L L G S
agggcggaggtggcttcgctggagaaggtccagctgtt
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
V S I G V A W Y Q K P G K A P K L L
agagctccgacacaggttcagactaacaccagttcgacacacagttcgacaca
Y S A S Y R T G V P S R F S G S S G
agagctccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
dgagaacagtccgacacaggttcagactaacaccagttcgacacacagttcgacaca
Table 10. PKI-NFluc(2-416) (see also SEQ ID NOs:16-17).

```
atggaggtactagctagttactgtacatttatttagcgaggttgtcgaacaagagaaga
M G T Y A D F I A S G R T G R
aggaatgccattcatatggtgagcagcgcctgtcaggtgggtggtctgtttgtgggtgggt
R N A I H D G G A G G A A G S G G G
gg台账gaagcgcaccaacacataaagaagcagccgcgcaccttcacctccctcagggat
G T E D A N I M K G P A P P Y P L E D
gpaaccgcctgcagcgacatcgctataaggtcagagtacgcgctctgtctcctcgcgaca
G T A E Q G L H K A M K R Y A L V P G T
atgctttcagatgtcagatatgtcgagttacatgataacacagctagcgcgataacacatc
I A P T D A H I E V N I T Y A E Y P E M
 tecgtcctgtgccgcaagctatgtaaacagatctggtctaataaaacacagcactgct
S Y R L A E A N K R Y G L T N H R I V
 gtatcgtactgccgccgcctcgacctatatatatagtaacagctatgctcagcagcactgct
V C S S E N L Q P F M P V L G A L F I G
gtgcagttggtgccgccgccggaaccacctacattatatagtaacagctatgctcagcag
V A V A P A N I Y N E R E L L N S M N
atcccgacocctacgcctgtttgtgtgttcccaaaagggggtgtcgaacaagatgtg
```
Table 11. CFuc(398-550)-PKA (see also SEQ ID NOs:18-19).

atgatgtccggttattgttaaaacaactcgcagacgcaacactcgcctctg
M M S G Y V N N P E A T N A L
attgccagatagtgtactgctacatcctggagcacatgctactgggacagagcagacaac
I D K D G W L H S G D I A Y W D E D E H
ttctctatattgaccgcctgtgaagcttactaatataacaaagggatatcaggtggcccc
F F I V D R L K S L I K Y K G Y Q V A P
gcttgaattggaatcagatattgtttacaaaccaccccaaccttgccacgccggtggtggacgt
A E L S I L L Q H P N I F D A G V A G
ttccccagctagactgcgccgtgtaaccttccccgccccggtgttttt tggtggacgaggaag
L P D D E A G E L P A A V V V L E H G K
cagatgcagaaaaaagatctgtgattacgtggcactgccccgttttaacacccactgggaaag
T M T E R I V D Y V A S Q V T T A K K
ttgccgaggaagggttggttgggagagtttcctaccggaaacactccgac
L R G G V V F V D E V P K G L T G K L D
gcaagaaaaatacaggagatctcctataaaagccagaaaggggccagagatcctaatgcccc
A R K I R E I I L I K A K K G K S K L G
ttcgcagggcgtttccagcaggtgggttgctggagccgggctgccccggaagccgcc
L Q C G S G G G G G G G G G G P N A A
gcggccagagagggcagcgcagagcagagctgaagaggtttctagccaaagcaggaag
A A K K G S E Q S E V K E F L A K A K E
gattttcctaaataagggacacccctttctagataacccggccagttggtgccagtctgttgatt
d F L K K W E T P S Q N T A Q L D Q F D
agaatacgacccctttgccaccgcgtctctttggtggctgtctggtgtgagccagcagagag
R I K T L G T G S F G R V M L V K H K E
agtggagacacctacgcagatagatcttagaacaacgcaagaggtgttaggtgaaggctaagcag
S G N H Y A M K I L D K Q K V V K L K Q
atcgacgacacctcgtgaatgagaacgcagcccagcgcgtcaacctccccggtctctggttc
I E H T L N E K R I L Q A V N P F F L V
Table 12. p300-NF Luc(2-411) and also see SEQ ID NOS:20-21.

| ATGGGCGAGCGCAGCAGCATAACCCAGCCGATAAAGCAARGTGAATCCAG | MGSGAHTADPEKRKLIQ |
| CACGACGTGTCAGGTGTCAGGACGTGACGCAATAATGCAACGCAGGTAAACAGGAATGCGC |
| QQLVLLLHAKHCQREREQANG | GAAAGTGGCAGCTGACTAATCTGCCAACAGTTGAAGAGTTCGACATGAGT |
| EVRLQCNLPHCRTHKKNVNLHM | ACCATATGACAGCAGCTGATAAAGCTGCCAGTTGGTACATTGGTACGAGG |
| THCQGKSCQVAHACASSQRQI | ATTACCACTGGAATACCTGACCCGCACATGATGGCTCAGGGTGGTAAACAC |
| ISWHKNCNCTRHDPCVLPLKNC | GCAGGCTGATAAAGCCTGGTACGAGGAATAAGCCGAGTGAGGAAACAGATCAG |
| TEGAKNIRKGPAPFYPLEDPD | ACCGCTGGAGAAGCAACTGCAATAAGCTGATAGAAGGATACGACCTGTGTTCTGGTAACAATC |
| TAGEQLHKAMKRYALVPGTI | GCTTTTACAGTACACATATCCAGGTTGAACACATCAGTACGGGATAACTCTGGAATGTC |
| AFTDHAEVNAEBFEMS | GTGCTGCTGACCCACATATGAAAGATATGGCTGAAACATACACAGACATCTGGTA |
| VRLEAAEMKRYGLNTNHRIVV | TCGTGAATAATCTTCTCTTCATTCATTTTGCTCCGCTTGGTCAATTATTTACAGTT |
| CSNSELQFPMPVGLALFIGV | GACATTGTGCCCCGCGGAACAGGAAATTTGTAACGATGGT |
| AVAPANDYINERELLNSMNI | TGGAGAGCAACTCTGAGGATTATCTGAGGAGAGTGGCTAAACAGGAGGTT |
| SQPTVFSKKGKLQKILNVO | AAAATTACACCTAATCCAGAAATATTTATGATCTTAACACGATTACCAAGAG |
| KKLPIQIXIIIIMDSKTQDG | TTGGTACGTAGTACCAAGCTTCATCTCATCTCCCTCCGTTTATTGATAATACAGT |
| FQSMYTFVTSHLPFGNEYD | TTGGTACGAGCTTCCTTGTACGTCGACAAAAATGCACTGATAATCTCCTGGT |

K E F T E F
Table 13. CFuc(398-550)-Hif-1α (see also SEQ ID NOs: 22-23).

```plaintext
atgatgacgggttgattatgtaaacaatccggaagcgcaccacgcccttgattgcaacag
MMSVYNPNPEATNALIDK
```
Table 14. PBSII-NFluc(2-416) (see also SEQ ID NOs:24-25).

```
atgggacagcagccatcactcactctcaccacacacgagatcgaatcggag
MGSSHHHHSSQQPPSSESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESESE
Table 15. CFuLC(398-550)-Zif268 (see also SEQ ID NOs: 26-27).

atgatgtcgcggttatgtaaanacacctgccgaagcaccacaccgctggattgacaag
m m s g y v n n p e a t n a l i d k
atgagatgcacattcctgggacagcataattactgtgggacagacagcaccacaccctggatatg
 D G W L H S G D I A Y W D E D E H F F I
agctgcggcttgcagacttttaaataaataacaggatatcggggtgccccgctgtaattg
V D R L S L I K Y K G Y Q V A P A E L
acatgcgatattgtagttacacacccccacacctttcgacgcggcggctggaggtttgcttcccagc
E S I L L Q H P N I F D A G V A G L P D
agtggcgcgcgctgactccccgcgcgcgctttgtgatttgagacccccagagagagatgagacg
D D A G E P A A V V V L E H G K T M T
ggatcagatctgaggattacgcggcagctcaagtacccaaaaccgcggaaaaatgttcgacag
E K E I V D Y V A S Q V T T A K K L R G
agtagagatctctcataaagcgcagaagaggcggagagttccacatctgacatggcgcagggc
I R E I L K A K K G K S K L G L Q G
agtagagccgcagcgggtttcctgggacagacagcaccacaccctttcgatcggagtacg
G S G G G S G G G S G S P G E R P Y A C
cggaggttgctgtgagtcgcggctttccagctgccgcagctgacattgtagttacacacccccacac
P V E S C R D R F S R S D E L T R H I R
atccacacaacggcggagcctccacatctgtgacgtagcttcagctgcagctgcagctgcagctgcagc
I H T G Q K P F Q C R I C M R N F S R S
gaccacactccaccacccacatccgacccccacccaaacgaggaaaaagcccccttttgcttgctgagacatc
D H L T P I R H T T G E K P F A C D I
tggagaaaaattttgccagagcgtagatcggcggagcaccacaccctttcgatcggagtacg
C G R K F A R S D E R K R H T K I H L R
caagagacctcgcggctttgatgtaaaagaaacggtgctggagaattttgagacccccagacatagcg
Q K D L E S G K E T A A A A K F E R Q H M
gacctgcgtactacgcagcttttaaaaacctggggtcttgctgctgccccgctgagacatc
D S S T S A A -

Table 16. CFuLC-E2C (see also SEQ ID NOs: 28-29)

atgatgtcgcggttatgtaaanacacctgccgaagcaccacaccgctggattgacaag
m m s g y v n n p e a t n a l i d k g
tggatcacattcctgggacagcataattactgtgggacagacagcaccacaccctggatatg
W L H S G D I A Y W D E D E H F F I V D
cggaggttgctgtgagtcgcggctttccagctgccgcagctgacattgtagttacacacccccacac
R L K S L I K Y K G Y Q V A P A E L S
atccacacaacggcggagcctccacatctgtgacgtagcttcagctgcagctgcagctgcagc
I L Q H P N I F D A G V A G L P D D D
gccctgtaaatcctccccgcgcgtttggttttggagacccccagagagacagatcggggaagaag
A G E L P A A A V V V L E H G K T M T E K
agatcgtggtagatcgcgtccagctcaagtaaaccacccgctgaaatgttcgctgagagattg
E I V D Y V A S Q V T T A K K L R G G V
tgtttttgagacagttacccaaacactttcagctgcagctgcagctgcagc
V F V D E P V K G L T G K L D A R K I R
agatcctcaaaagcgcagacagagcggagaggtcctcgacagggcgttca
E I L K A K K G K S K L G L Q G G S
gccctggtggggtttcctgggagagttacgctgggtgagacccggagagacccctattgtgttgctggaag
tgctgatgtcctcctagtgaagagaggattctctggtggaccacacgcttcaccacacaggg
C G K S F S R K D S L V R H Q R T H T G
gaaaaacgtagtataaacctcccagagctgacgcaaatattttatcagtctcgggggtacattagg
E K P Y K P C E C G K S F S Q S G D L R
cgtccataaccgcaggctaccatcagtgggagagacccccataaatgcctcagaatgtagtggcaagttct
R H Q R T H T G E K P Y K P C E C G K S
Table 17. Aart-NFLuc (see also SEQ ID NOs:30-31).

atggccgacgccatcaccatcatcaccacacgccaggtactcccaggggaagccctat
MGSSSHSHHHHHSQDPGKPY
gcttgccggaatgtgtgtatgaatctctacgacgcagcatcaccttgaggccacaaacagctgt
ACPECGKSFSQRSDHLAEHQR
acccacacgggtgaaaaaccgtataaatgcccagatgcccgaatttttaattgctaaag
THTGEKYKCPCECGKSFSDK
aaatgatcgcaccgctaccagcaactcatcactcaatctggcgagaacccatatcaaatgtcagaa
KDLTRHQRHTGEKYKCPCE
ptgaccaggattttcagccgagcgcacaaaaattctgcgccgccccaccaccagtactcaccacccg
CGKFSQSRANLRAHRTHHTG
agaaactttatggtctccggaatgtgtatagcctctctctcatggttcggccacacgcgc
EKYKCPCECGKSFQSLAHRL
ccccacacgcttacaccacacggtgaaaacccgtataaatgcccagatgcggcaaatct
AHQRTHTGEKYKCPCECGKS
ttacgccgccagataactgctgacaccacatcaactcatatctggcgagaacccatatc
FSDREDNLHTHRHTGEKY
aatagtgccagaatgtggaagctttttctccccgcccaggtactctgaagcgtgcacacaccagt
KCPCECGKSFSSRDALNVHR
acccacacggcacaaaaaaactgcaccggtgggttggccgcttgacgctggggttcct
THGKKSTSGGGGSGCGGS
agtgggggtgtacggagacccgcacaaaaactgagccggtttacatctcatcatcctctctcctcctctc
tGCGGTTDAKNIKFGAPPYP
ctagggatgtggacacaccggaagagcaaccgtataagctgctagagagatatacgccttattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Table 18. MBD2-NFluc(2-416) (see also SEQ ID NOs:32-33)

atgggccagcagccatcactcactcatcaacagccaggaattcgaagaacgggaaaa
MGSSHHHHHHHSSQDPNSSEGK
ccgcatgagtgcggcgcgtcgccgggtggtgaagaaagaagaacctgtagttctaa
RMDCPAFLPPGWKKEEVRKS
ggccctagcgccggcacaagagcagcgttttagcttagccccggcacaagccccatttt
GLSAGKSDVYYFSPSGKKFR
agaacacgaggacttggtcgccttgaggcaacagcgtgatccaagcagccgttttatttt
SKPQLARYLGNVDLSSFDF
ctagccgccacaaattgacgtaggtgagggtcgccttgaccctgtgctcggccggtg
RTGMHTGGGGGGGSGGGG
aggtacagagacgccaaacataaagacaagcagggccggcacttgctattctatagagt
GTEDAKNIKKGPPAFYPLED
ggaaacgcctggagaccaactcataaggctataggaagataagccggctttctcggacaac
GTAGEQLHKAMKRYALVPGT
attcgctttacagatgcacatatcgagggtaaacatacagctacgcgtggaatattt
dahtdahievnyayefem
tccgtcggttggcagacgttagaagcagatgggtgataaaatcagaatacgtc
SVRLAEMKRYGLDNTHRIV
gtatgacgtgaaacactcctcttcatatttcctcgggttttgccggtttttattctcga
VCSENSLQQFMPVLGALFIG
gttgcagtgcggcggccgaaacgagccatttataataggaacgtgaaattgctcaacagctagac
VAVAPANDIYNELLNSMN
attcgccagctaccctagtgttgotttttccaaaaaggggtttgcaaaaaattttgaacgtg
ISQPTVVVFVSGBKILNV
caaaaaaaattaccataaatcgcagaaaaatttatcatggattttaaacaggaattaccag
[00190] Tables 19-24 provide split reporter fusion protein sequences useful in assays for the detection of small-molecule inhibitors of kinases.

Table 19. Cfluc-PKA_(see also SEQ ID NOs:34-35).

atgatgcctgcgtatgatgtaacactcggacaaccggctttagctgaaagtggaaa
HMSGYVNNPEATNALIDKDG

tggctactattgttacaaacatccggacaaccggctttagctgaaagtggaaa
HMSGYVNNPEATNALIDKDG
Table 20. Cfluc-PDGFRB (see also SEQ ID NOs:36-37).

```
atggtatgcaggttcatgattaaacactgcaacgagcggctcttgataagatctaggtgagctgtc
agtgctgagcttgacaggttcatcgtcactgccgttgatcaggtttgccttcaggtgtgttgcct
agtgagcttgacaggtttccttgacaggcttgacgcttgtgcgctgtgttgccttcaggtgtgt
tgacgcttgacaggtttccttgacaggcttgacgcttgtgcgctgtgttgccttcaggtgtgt
```
Table 22. Cfluc-FYN (see also SEQ ID NOs:40-41).

atggtatagtctcccggttaaaccacaaatcgcggtattggtgccgcgtatagagatgtm
mgmgsygvnnpewhatnaldk
ggatggctgatattctggagatacagcttactggaagagaacgcacactttttcgtattg
pgvptomsdypksfpsfpkwarqd
ppsttttaaataaacaagagatatacaggtggcccccctctgtaattggaa
rdrlkslikykgqvapaelec
tcgatattggtacacaccccaaacatctgcgaagcgggtgcggttgctcttcccagcat
sillnpifdasgvgalpdd
hggacgcgccgggtacccctgcggttggtttttttgagcgacggaaagaacagatgacgagaa
dagelpavvvvlehgkmtm
aaagagatgtggattacgctgcagctcaagaaaccgaaataagttgcgcgggaagga
keiivdyvasqvttakkklrrg
gttgtggtgtctggagagattaccgcggaggtctattaccgggaacttcgagcgaagaatc
vfvfvpvkgglgtgkl达尔ki
agagagatctccatattccagcggccagagtt gccgcagggcgtt
relikakkkgkskglqoggt
tcaggcgtggtgggtctctgcggggtgaggctgctgccacactggcgagttgccacacagcttcgggaagtt
sgggsgsgggggsvdamaglcccrc
tagtagtctctgccaaaggaggtccggagtactaccatctgtctgtctcttgcaaaaccaaa
lvvpcchkgmprltdlsvkkt
agttcgggaaatactccctgaggactttccctcctctgactgtacagcttacagtagtggaatgtgacgagggcag
dvweipreslqlikrllgnq
Table 23. DHFR-NLuc(2-416) (see also SEQ ID Nos: 42-43).

atggccagcaggatcatttcaggtgaatcaggaacaaacaaagtttgctacaagcagctttctttgevwmgtwnwngntkvaiktlaaaccaagcacaattcagctgtgagcagcagatcatcaggaagaagtctgkpgtmspeshflleeaqimkkkaaacccagcacaaatgcgccggaatctccaggtgtctgagagtcgccatcattactcagttctttcaaataggtggaaggaagagctcteymknksslddflkkgdegraatgtagattaccaaatctctgagcctagctgcagacacctttgtaagttgatgacagttactcctgagctgtcaccf RTIKLQVYLVAQVAAAGMAYIgagcgcagcactatatctaatctatagagcagatcagctgcagacacctttgtaagttgacagttactcctgagctgtcaccf

CWWKDPDEERPTEFELYQLQSFELGACTCTTTACCGCGAGACGGCCCAATCCACCTGTGGAACAACTGTAAC

D Y F T A T E P Q Y Q P G E N L -
Table 24. Fos-Nluc (see also SEQ ID NOs:44-45)

atgggcagcagcagccatccaccatctcaccaccaagcagacgatcagctgcattact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
agactttctcaacagttgaaactaacatctggaaatcatctgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
dllqtaetdqldekealsqte
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
atgtcataatcctttcctacaatgtacatctgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
ianlkekkeklekplagaaggg
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
ggaggttcggttggttggttggttcaacaggtggagcatctgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
pqpylpdedgtageqqlhkam
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
aagacgtacacacctcgtgctcgtgaacacatctgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
kpjaflvpigtaftpdaehinv
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
atcagctccagcagcagctacatctgataacaggtgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
tyaevefmsvrlaesahkry
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
ggacgataaatcataaatacagacaggacttactgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
glntmhrivvcsensnlqlpmm
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
ccagggtgggtgggttgattatcactgcaggtgcaggtgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
pvlgalfigyavapandiyng
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
gaacgtgatcccacagctgatcagcttcacagtgcctggatgtctccaaaa
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
ereellenmsnqpttvvvs
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
aaggggttcagcaaatggatccacagtgcctggatgtctccaaaa
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
klqkiligkklpiqki
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
atccagctaccacacctcgtgacgacttactgcacggacttact
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
imsdtyqfgsmithvts
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
catacctccccgcgttttaaatggatagtacgtgcctggatgtctccaaaa
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL 
hlppgfncyedydpvpesfdrdx
MGSSHHHHSSSLTVKIGLYRQPPDPNQESSLYLYRQPPDPNQPQTVKPSQQGGVYPLPGL
Proteins useful for the detection of amyloid oligomers with designed protein conjugated to luciferase halves are given in Tables 25-28.

Table 25. CFuic-TJ10. See also SEQ ID NO:92-93.

atgatgtcaggattatgttaaacaataccggaaagccgaccaacgcccttggattgacaaggtgaga
M S G Y V N N P E A T N A L I D K D G
M S G Y V N N P E A T N A L I D K D G
gttgctacattcggagacataagcttacttgggacagacagacagacacttcctctcatagtgac
W L H S G I A Y W D E D E H F F I V D
cggctgagcttttaaatattaaataaagaagatatacggtgggccccgctgataatttggattctc
R L S K I R Y K G Y Q V A P A E L S
tttgttacaaacaccccaacatcttcgacgaggcggagctggctccccgacgagctagac
I L L Q H P N I P D A G V A G L P D D
gccggtaaactccgccgccgccgctggtgtttttggagcacaggaaacgcgtggagagtt
e Iv D Y V a S Q V T T A K K L R G G V
gatcgtggagtacgctgcacgtcaagtaaacaacccgccaagaaagtggcggagagtt
gagratatccattaggcacagatccggaaacactccgacgcaagaaaatcaga
E I L I K A K K G G S K S K L G L Q G G S
ggcggattggggttctggcgcgggttgaggacccggg
G G G G C G G G G G G S P C
gggccagacattcggcttagatatcaggggtaaa
M A Q T W L S I Q G K
acccctgatttgccagatcaggctattgctcatagcgcgaagcgtggaaacatcttc
T L Y W Q I R Y A I D A A A E E K I F
aaacagtagctcaacgcaaaaaggtatccgacgagcgtggataaggactacagcagcgtaccaaa
K Q Y A N D N G I D G E W T Y D D A T K
acccctacggctacccgaa
T F T V T
Table 26. TJ10-NFluc. See also SEQ ID NO:94-95.

catcagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagctagCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGC

92
Table 27. CFLuc-HTB1. See also SEQ ID NO:96-97.

```
catatgatgtcccgtgttacataaacaatccggaa
| GCGGAAACAGCCTTTGTTGAAAGATGGATATGGCTACATTCGTTTGGAGAATATTCGTTTG
| MGSTGYFVYNNPE
| ACGCAGGAACACCACTTCTCATAGTTGCCGCGCAGCACTTTCAATTAACTAAGAAGA
| DSADEHFIIVDRKLKSLIKYKG
| TATCGGTGGGCCGCTGATATTGGAAATGATTGTCTACACCCCAACACTTCGCAC
| YQVAFALESILQHPNIFD
| GGCGGGTTGGCAAGTCGTTTCCGCGAGTTGCGCGTCGGAATCTTCGGCGGCGTGGTGT
| ACGVAAFFDPDDAGELPAAVV
| TGGAGACCGAAAGCAGATGACGCGAAAGAGAGACTGCGTTACGTGCAGCTAAGTAT
| EUHGKTMTKEIVSYVASQV
| ACAACCGGAAAGGTGCGGCGAGGTTGTTGGACCGAATACCGGAAGCTTCTTT
| TAKKLRLGGRVVVDVFKGL
| ACCGGGAAACCTCAGCAGAAGAATACAGAGAGACCTCCATAAGGCCAAGAAGGCCGA
| TGKLRARKIREIILIKAKKGG
| AAGTCGAAATGCCGCTGAGGGTGCGCCGCTGAGGCTGGTCTGCGGGGCGGAC
| KSKRGLQGSGSGGSGSGSGSGGGGGG
| CGCGGATGCGCCGACCTCAACCTATACATCAACGTCGAACCCCTGAAGGTTAAAC
| PGMRAFTFKLIIINGKTLKGEI
| ACCATCGAGAGCTGACGCTGACAGATGACGCGAGACCGCGATACCGTACACGAC
| AIEAVDAAAAEAEKIFPQKYAND
| AAGCTGGTATCGCGGTAATGGACCCCAGACGACGCTACAACACCTCCCGTCCGGAA
| NGIDGEWTYDADATKTFVTET
| CTCGAGTCTGTTAAGAAGACCGCTGCTGGAATATTGACGCCAGCAGCATG
| LESGKETAAAARKFERQHM
```

Table 28. HTB1-NFluc. See also SEQ ID NO:98-99.

```
atggcgcggatcgcattcaccatcacac
| MGGSSHHHH
| CATCAGGATCCGACGCTCATTATATGCGCGAGCGCCACCTCAAGTATCATCAACCGGTAAA
| HHSAAHMAQTFKLIIINGK
| ACCCTGAAAGGTGAATCCATCATGCGAGCTTGCCAGCGACTGAAAGAAATCCCT
| TLKGEIITEAIVDAAAAEAEKIF
| AAACAGCTCGTCAACGCCAACCGTGATCGCTGTAAATGGACCCCGACGCTACAACAA
| KQYÄNDNGIDGEWTYDDATK
| ACCCTGACCTACCCGAC
| TFTVTET
| GGTTGGGGTGGCGGTTAGGCTTCCGGCGGGTTCTGCTGCTGGTGTAACGGCAGCGCAAA
| GGZGGGGSGGSGSGSGSGGTEDAK
| AACATAAGAAGACGCCGGCGGCATTATACCTTATCTTCCAGGAGTTGCGAAGGGCGAGGA
| NIKKGPAFFYPLEDEGTAGEQ
| CTCGATAAGTGTTAGAGATTATCGCGCCCGTGGTCTGAAATATTGCTTATACAGTAC
| LHAKMRYALVPGTIAPTD
| CATACGGGATGAAACATCGACGCGGAAATATCGCTTGCCTGGTTGGCCGAGAA
| HIEVNTYAEYFEMSVRLEAGCTTGGAAACGATAGGGGCTTAATACACACAGACATCGCTGATACCGTAAAGAAACTCT
| AMKRYLGTNTNHRIIVVCSENSCTCAATCTCCTGCCTGGCGCTTTATATTCGAGATTTCGATTGCCGCCCG
| LPQFPMPVLGALFIGVAVAAPA
| AACGACATATTATGCAACCTTAATTGTCAACACAGATGAAACATTCTCCCGCTACGGGTA
| NDINYNEELNLSMNISQPTVGTGTTGTCTTCAAAGGTTTGGAAAAAATTGTGACCGAAAAAAATATACCAATA
| VFSKXKGRLQXILNVQKXKLPIATCAGAAGATAATTACCTCAGGCTTCCCACCGGTTACACGGGATTCGATAC
| IQKRIIIDSDKTDYYQGPSMNYACGTCGTCACATCTCATCTCAACCCCGGTTTTAATGTAATACGATTTTGACGACTC
```

[00192] Examples of methods employing living cells or transgenic organisms are provided in US Patent Publications 2005/0144661, 2004/0235064; 2007/0161067; 2006/0224331; and US Patents 6,897,017; 6,872,871; 7,166,424; 7,160,691; 6,828,099; 6,428,951; 6,929,916; 7,062,219; and 7,176,287. See also Kim et al. (130); Porter et al. (23); Porter et al. (58); Paulmurugan et al. (131). These references provide numerous examples of split reporters useful in the practice of the present invention in addition to those particular examples taught herein, although certain split reporters may be preferred over others.

[00193] All references throughout this application, for example patent documents including issued or granted patents or equivalents; patent application publications; and non-patent literature documents or other source material; are hereby incorporated by reference herein in their entireties, including any Supporting Information, addenda, as though individually incorporated by reference, to the extent each reference is at least partially not inconsistent with the disclosure in this application (for example, a reference that is partially inconsistent is incorporated except for the partially inconsistent portion of the reference).
[00194] All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains and these references cited herein reflect the state of the art as of their filing and publication dates, it is intended that this information can be employed herein, if needed, to exclude (or disclaim) specific embodiments that are in the prior art. For example, when a compound is claimed, it should be understood that compounds known in the prior art, including certain compounds disclosed in the references cited herein (for example, in referenced patent documents), are not intended to be included in the claim.

[00195] When a group of substituents is disclosed herein, it is understood that all individual members of those groups and all subgroups, including any isomers and enantiomers of the group members, and classes of compounds that can be formed using the substituents are disclosed separately. When a compound is claimed, it should be understood that compounds known in the art including the compounds disclosed in the references disclosed herein are not intended to be included. When a Markush group or other grouping is used herein, all individual members of the group and all combinations and subcombinations possible of the group are intended to be individually included in the disclosure.

[00196] Every formulation or combination of components described or exemplified can be used to practice the invention, unless otherwise stated. Specific names of compounds are intended to be exemplary, as it is known that one of ordinary skill in the art can name the same compounds differently. When a compound is described herein such that a particular isomer or enantiomer of the compound is not specified, for example, in a formula or in a chemical name, that description is intended to include each isomer and enantiomer of the compound described individually or in any combination. One of ordinary skill in the art will appreciate that methods, device elements, starting materials, synthetic methods, and amino acid and protein sequences other than those specifically exemplified but functionally equivalent to those specifically disclosed herein can be employed in the practice of the invention without resort to undue experimentation. All art-known functional equivalents, of any such methods, device elements, starting materials, synthetic methods, and sequences with equivalent function to those specifically disclosed are intended to be
included in this invention. Whenever a range is given in the specification, for example, a temperature range, a time range, or a composition range, all intermediate ranges and subranges, as well as all individual values included in the ranges given are intended to be included in the disclosure.

[00197] As used herein, "comprising" is synonymous with "including," "containing," or "characterized by," and is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. As used herein, "consisting of" excludes any element, step, or ingredient not specified in the claim element. As used herein, "consisting essentially of" does not exclude materials or steps that do not materially affect the basic and novel characteristics of the claim. Any recitation herein of the term "comprising", particularly in a description of components of a composition or in a description of elements of a device, is understood to encompass those compositions and methods consisting essentially of and consisting of the recited components or elements. The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein.

[00198] The terms and expressions which have been employed are used as for description and not for limitation, and there is no intention in the use of such terms and expressions to exclude any equivalents of the features or portions thereof shown and described, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by certain embodiments and optional features, modification and variation of the concepts disclosed herein may be made by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

[00199] In general the terms and phrases used herein have their art-recognized meanings which can be found by reference to standard texts, journal publication and contexts known to those skilled in the art.

[00200] Although the description herein contains certain specific information and examples, these should not be construed as limiting the scope of the invention but
as merely providing illustrations of some of the presently preferred embodiments of the invention. For example, thus the scope of the invention should be determined by the appended claims and their equivalents, rather than by the examples given.

[00201] An active reporter protein is one for which activity can be detected directly or indirectly upon reassembly of the first and second fragments. In the case of a fluorescent protein, fluorescence can be detected at the relevant wavelength. In the case of a reassembled enzyme, the enzymatic activity is detected via generation of a signal upon action of a substrate to generate a product which can be measured spectrophotometrically at a particular wavelength of light, for example. In the case of luciferase, generation of light is measured using a luminometer, for example.
Bibliography

10. Michnick, S. W.; Ear, P. H.; Manderson, E. N.; Remy, I.; Stefan, E. *Nat Rev Drug Discov* 2007, 6, 569-82.
CLAIMS

We claim:

1. A method for detecting interactions of a ligand molecule with an interacting protein having a ligand binding site using a split monomeric reporter protein system in a cell-free assay, said method comprising the steps of:

(a) providing (i) a cell-free system comprising coupled or uncoupled transcription and translational machinery; (ii) a first nucleic acid molecule encoding a first interacting fusion protein, said first interacting fusion protein comprising a first fragment of a reporter protein covalently linked to a first interacting domain comprising a ligand binding site, and (iii) a second interacting fusion protein, said second interacting fusion protein comprising a second fragment of a reporter protein covalently linked to a second interacting domain, wherein neither the first nor second fragment of the reporter protein are active alone but are active when reassembled to form an active reporter protein;

(b) allowing the expression of the nucleic acid molecule in the cell-free system in step (a) to produce the first interacting protein and providing a second interacting protein;

(c) contacting the first and second interacting proteins either (i) with a composition where the second interacting protein comprises a ligand of the first interacting protein such that when the ligand and the first interacting protein are present, the first and second fragments of the reporter reassemble to form an active reporter protein; or (ii) with a composition that contains a ligand of both interacting proteins such that when a ligand to both of the first and second interacting proteins is present, the ligand binds to the first and second interacting proteins and the first and second fragments of the reporter reassemble to form an active reporter protein;

(d) detecting the active reporter protein when the ligand is present in the composition and the first and second fragments have reassembled to form an active reporter protein.
2. The method of claim 1 wherein the monomeric reporter protein is a fluorescent protein, a β-lactamase or a luciferase.

3. The method of claim 1 wherein the fluorescent protein is a green, blue, yellow, red or enhanced green fluorescent protein.

4. The method of claim 1 wherein the first nucleic acid molecule comprises at least one stabilizing element region.

5. The method of claim 4, wherein the stabilizing element region is a stem-loop structure.

6. The method of claim 5, wherein the stem and loop structure is a 5’ stem-loop structure and/or a 3’ stem-loop structure.

7. The method of claim 5, wherein the stem-loop structure is a 3’ stem-loop structure of bacteriophage T3 and/or is a 5’ stem-loop structure of bacteriophage T7.

8. The method of claim 1, wherein the first interacting protein and the second interacting protein comprise antibody (Ab), single chain Ab, an antigen-binding fragment of an Ab, or an antigen-binding fragment of a single chain Ab.

9. The method of claim 1, wherein the ligand is a double stranded DNA molecule, a single stranded DNA molecule, an RNA molecule, a methylated DNA molecule, a peptide, a protein, an antigen, an amyloid protein, or a small molecule ligand.

10. The method of claim 9, wherein the small molecule ligand is an enzyme binding molecule, the enzyme selected from a group consisting of transferases, hydrolases, ligases, oxidoreductases, lyases and isomerases.

11. The method of claim 10, wherein the enzyme is a kinase.

12. The method of claim 9, wherein the ligand is an RNA molecule and the interacting protein is a pumilio domain, a KH domain, a RRM domain, Argonaut, bacteriophage MS2 coat protein, or eukaryotic initiation factor 4a.
13. The method of claim 1, wherein the second interacting fusion protein is translated in the cell-free system from a second nucleic acid molecule encoding the second fragment of reporter protein covalently linked in frame to the nucleic acid encoding the second interacting domain.

14. The method of claim 1, wherein the ligand molecule and the interacting protein are p53 protein and HDM2; Bcl and Bak; FKBP and FRAP; BAD and BCL-XL; p38a MAPK and MAPK activated protein kinase 2; cMyc and Max; HIF1α and p300; Fos and Jun; PIN1 and Jun; PKA and PKI, or JT10 and HTB10.

15. The method of claim 9, wherein the ligand is a double stranded DNA molecule, a single stranded DNA molecule, an RNA molecule or a methylated DNA molecule, DNA or RNA of a pathogen or a diseased cell, whereby assessment of disease markers, diagnosis and prognosis is achieved.

16. The method of claim 15, wherein the ligand is a double stranded DNA molecule, a single stranded DNA molecule, an RNA molecule or a methylated DNA molecule, DNA or RNA of a pathogen, whereby confirmation of presence or absence of the pathogen in a biological, environmental, food, beverage, meat, poultry, fish or water sample is achieved.

17. The method of claim 9, wherein the ligand is a peptide or protein derived from aa pathogen or diseased cell, whereby disease markers are measured with that result that prediction, diagnosis or assessment of therapeutic success is achieved.

18. The method of claim 9, wherein the ligand is a protein of a pathogen, whereby confirmation of presence or absence of the pathogen in a biological, environmental, pharmaceutical, commercial, food or water sample is achieved.

19. A method for detecting an agonist or an antagonist of an interaction of a ligand molecule with a biomolecule using a split monomeric reporter protein system in a cell-free assay, said method comprising the steps of:
(a) providing (i) a cell-free system comprising coupled or uncoupled transcription and translational machinery; (ii) a first nucleic acid molecule encoding a first fragment of reporter protein covalently linked in frame to nucleic acid encoding a first interacting protein, said interacting protein comprising a ligand binding site, (iii) a second nucleic acid molecule encoding a second fragment of reporter protein covalently linked in frame to nucleic acid encoding a second interacting protein, whereby the first and second RNA or DNA molecules, when transcribed into mRNA are translated into the first and second proteins in either coupled or uncoupled translation systems;

(b) allowing the cell-free system provided in step (a) to produce the first and second fragments of the reporter protein;

(c) contacting the first and second fragments of the reporter protein produced in step (b) with either (i) with a composition where the second interacting protein comprises a ligand of the first interacting protein such that when the ligand and the first interacting protein are present, the first and second fragments of the reporter reassemble to form an active reporter protein; or (ii) with a composition that comprises a ligand of both interacting proteins such that when a ligand to both of the first and second interacting proteins is present, the ligand binds to the first and second interacting proteins and the first and second fragments of the reporter reassemble to form an active reporter protein;

(d) simultaneously or subsequently with respect to step (c) contacting with a composition which may or may not contain an antagonist or an agonist of ligand binding; and

(e) detecting the signal generated by the active reporter protein when the ligand is present in the composition and comparing the signal in the presence and absence of the composition, whereby an antagonist of ligand binding is identified or detected when the signal is less in the presence of the composition than in the absence of the composition or an agonist of ligand binding is detected or identified when the signal is greater in the presence of the compositions than in the absence of the composition.
20. The method of claim 19, wherein the ligand molecule and the biomolecule comprise but are not limited to, p53 protein and HDM2; Bcl and Bak; FKBP and FRAP; BAD and BCL-XL; p38a MAPK and MAPK activated protein kinase 2; cMyc and Max; HIF1α and p300; Fos and Jun; PIN1 and Jun; or PKA and PKI.

21. The method of claim 19, wherein the ligand binding protein is an enzyme selected from a group consisting of transferases, hydrolases, ligases, oxidoreductases, lyases and isomerases.

22. The method of claim 21, wherein the enzyme is a protein kinase.

23. The method of claim 19, wherein the first interacting protein and the second interacting protein comprise (i) antibody (Ab), (ii) single chain Ab, (iii) antigen-binding fragment of an Ab or antigen-binding fragment of a single chain Ab.

24. The method of claim 9, wherein the ligand is an amyloidogenic protein comprising beta-amyloid(1-40, 1-41, 1-42, 1-43), prion protein, alpha-synuclein, tau, immunoglobulin, islet amyloid polypeptide or huntington protein.

25. The method of claim 24, wherein the amyloidogenic protein is in a sample extracted from or is present in a biological sample, whereby assessment of disease markers, diagnosis and prognosis is achieved.
Figure 2
Figure 3
Figure 4

**a**

**b**

Relative ALU

![Graph](image1)

**c**

<table>
<thead>
<tr>
<th>K_{d}^{app}</th>
<th>IC_{50}</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGG</td>
<td>0.56</td>
</tr>
<tr>
<td>TAG</td>
<td>2.61</td>
</tr>
<tr>
<td>TTG</td>
<td>47.28</td>
</tr>
<tr>
<td>TCG</td>
<td>105.57</td>
</tr>
</tbody>
</table>

**d**

\[
y = 0.5955x + 0.1947 \\
R^2 = 0.9958
\]

**e**

![Graph](image2)

**Figure 4**

Apparent K_{d} (nM)

<table>
<thead>
<tr>
<th>IC_{50} (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>80</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>120</td>
</tr>
</tbody>
</table>

Relative ALU

<table>
<thead>
<tr>
<th>[Inhibitor] (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>100</td>
</tr>
</tbody>
</table>
Figure 7

(a) Enzymatic reaction: \( \text{ATP} + \text{O}_2 + \text{Fluorescein} \rightarrow \text{Luciferase} \)

(b) Protein interaction:
- Recognition Element A
- Extracellular Protein
- VEGF
- Flt-1
- CFuc

(c) Bar graph: Relative ALU vs. No VEGF, with 50 nM VEGF Dimer.
Figure 9
Figure 10
Figure 11
Figure 12

Early intermediates

Monomer → Low molecular weight oligomers ← Structural transition

Small/Intermediate oligomers → Protofibrils (~100 nm) → Fibriils (>1 μm)

Late intermediates and Fibrils (ThT and Congo red positive, also imaged by AFM and TEM)

Class I
- Recognized by TJ10
  - Aβ40
  - IAPP
  - Insulin

Class II
- Recognized by TJ10 and HTB1
  - Aβ42
  - Prion 106-126

Class III
- Not recognized by either TJ10 or HTB1
  - Lysozyme

Recognized by polyclonal antibody isolated by Glabe and coworkers

- Aβ40
- IAPP
- Insulin
- Aβ42
- Prion 106-126
- Lysozyme
- polyQ40
- α-synuclein

Recognized by B10AP
- Aβ40

Recognized by monoclonal antibodies WO1 & WO2 isolated by Wetzel and coworkers

- Aβ40
- IAPP
- Lysozyme
- polyQ42
- TTR
- β2microglobulin
- IgV L domain, T10S
Figure 13
Figure 14

Inactive Split-Luciferase
DHFR
Kinase

Complex Driven Light Production
N-Luc
C-Luc
Trt-Staur
Inhibitor
Staurosporine

Luciferase Activity Inhibited

II

Trimethoprim

Figure 15

<table>
<thead>
<tr>
<th>Kinase</th>
<th>50 µM Sunitinib</th>
<th>50 µM PP1</th>
<th>50 µM Roscovitine</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKA</td>
<td>149</td>
<td>8632.5</td>
<td>5623</td>
</tr>
<tr>
<td></td>
<td>125 nM jun-st.</td>
<td>5623</td>
<td>5102</td>
</tr>
<tr>
<td></td>
<td>125 nM jun-st.</td>
<td></td>
<td>7539.5</td>
</tr>
<tr>
<td>PDGFRB</td>
<td>110</td>
<td>294</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>125 nM jun-st.</td>
<td></td>
<td>259.5</td>
</tr>
<tr>
<td>CDK2</td>
<td>131.5</td>
<td>260.5</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>267.5</td>
</tr>
<tr>
<td></td>
<td>125 nM jun-st.</td>
<td></td>
<td>157</td>
</tr>
<tr>
<td>FYN</td>
<td>88</td>
<td>196.5</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>131.5</td>
</tr>
<tr>
<td></td>
<td>125 nM jun-st.</td>
<td></td>
<td>194</td>
</tr>
</tbody>
</table>
Figure 17
**Figure 20**

AaR-NFLuc  CFLuc-E2C

Add 36-mer dsDNA target containing zinc finger sites

**Figure 21**

![Graphs and diagrams representing experimental results]
Figure 22
Figure 23
Figure 24
Figure 25
## INTERNATIONAL SEARCH REPORT

### A. CLASSIFICATION OF SUBJECT MATTER

**IPC(8)**: G01N 33/566; C12Q 1/68 (2009.01)  
**USPC**: 436/501; 435/8  
According to International Patent Classification (IPC) or to both national classification and IPC

### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
USPC: 436/501; 435/8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
436/501; 435/8

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
PubMed, USPTO, EPAB, JPAB; Google Scholar

- cell free system, transcription and translational machinery, nucleic acid, fusion protein, fragment, reporter protein, interacting domain, ligand binding site, active reporter protein, contacting, composition, detect$, monomeric reporter protein, (fluorescent or 13-lactamase or

### C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 2004/0170970 A1 (VARSHAVSKY et al.), 02 September 2004 (02.09.2004). para [0019] [0043] [0077]. [0080] [0083] [0121]. [0122] [0142] [0174]</td>
<td>1-3, 9, 13, 19</td>
</tr>
<tr>
<td>Y</td>
<td>US 6,428,961 B1 (MICHNICK et al.) 06 August 2002 (06.08.2002). col ,22, in 18-22</td>
<td>16, 18</td>
</tr>
</tbody>
</table>

* Further documents are listed in the continuation of Box C.

---

Date of the actual completion of the international search: 05 February 2009 (05.02.2009)  
Date of mailing of the international search report: 17 FEB 2009

Name and mailing address of the ISA/US  
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents  
P.O. Box 1450, Alexandria, Virginia 22313-1450  
Facsimile No.: 571-273-3201  
Authorized officer: Lee W. Young

PCT Helpdesk: 571-272-4300  
PCT DSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)