PERMUTED AND NONPERMUTED LUCIFERASE BIOSENSORS

Cross-Reference to Related Applications

This application claims benefit under 35 U.S.C. § 119(e) of U.S. application Serial No. 60/788,608, filed April 3, 2006, U.S. application Serial No. 60/879,771, filed January 10, 2007 and U.S. application Serial No. 60/901,133, filed February 14, 2007, the disclosures of which are incorporated herein by reference in their entirety.

Field of the Invention

This invention relates to the field of biochemical assays and reagents. More specifically, this invention relates to modified luciferases and to methods for their use.

Background

Luciferases are enzymes that catalyze the oxidation of a substrate (e.g., luciferin) with the concomitant release of photons of light. Luciferases have been isolated from numerous species, including Coleopteran arthropods and many sea creatures. Because it is easily detectable and its activity can be quantified with high precision, luciferases have been used widely to study gene expression and protein localization. Unlike green fluorescent protein (GFP), which requires up to 30 minutes to form chromophore, the products of luciferases can be detected immediately upon completion of synthesis of the polypeptide chain (if substrate and oxygen are also present). In addition, no post-translational modifications are required for enzymatic activity, and the enzyme contains no prosthetic groups, bound cofactors, or disulfide bonds. Luciferases are useful reporters in numerous species and in a wide variety of cells.

Luciferases possess additional features that render them particularly useful as reporter molecules for biosensing, i.e., molecules which reveal molecular properties of a system. Biosensors (i.e., sensors which comprise a biological component) generally function by means of a two-step process: signal generation

mediated through a biological component, and signal transduction and/or amplification through an electrical component. Signal generation is typically achieved through binding, energy transfer or catalysis. Signal generation by enzymatic catalysis can be particularly useful due to the inherent efficiency and specificity of these chemical processes. Most catalytic reactions generate less than the energy of hydrolysis for two molecules of ATP, or about 70 kJ/mole. However, the luminescence elicited by luciferases has much higher energy content. For instance, the reaction catalyzed by firefly luciferase (560 nm) emits 214 kJ/mole of energy. Furthermore, luciferases are also highly efficient at converting chemical energy into photons, i.e., they have high quantum yields. Luciferases are thus extremely efficient for generating detectable signals.

Luciferase biosensors have been described. For example, Sala-Newby et al. (1991) disclose that a *Photinus pyralis* luciferase cDNA was modified to generate cyclic AMP-dependent protein kinase phosphorylation sites. In particular, a valine at position 217 was mutated to arginine to generate a site, RRFS (SEQ ID NO:117), and the heptapeptide kemptide, the phosphorylation site of the porcine pyruvate kinase, was added at the N- or C-terminus of the luciferase. Sala-Newby et al. relate that the proteins carrying phosphorylation sites were characterized for their specific activity, pI, effect of pH on the color of the light emitted, and effect of the catalytic subunit of protein kinase A in the presence of ATP. They found that only one of the recombinant proteins (RRFS; SEQ ID NO:117) was significantly different from wild-type luciferase and that the RRFS (SEQ ID NO:117) mutant had a lower specific activity, lower pH optimum, emitted greener light at low pH and, when phosphorylated, decreased its activity by up to 80%. It is disclosed that the latter effect was reversed by phosphatase.

Waud et al. (1996) engineered protein kinase recognition sequences and proteinase sites into a *Photinus pyralis* luciferase cDNA. Two domains of the luciferase were modified by Waud et al.; one between amino acids 209 and 227 and the other at the C-terminus, between amino acids 537 and 550. Waud et al. disclose that the mutation of amino acids between residues 209 and 227 reduced bioluminescent activity to less than 1% of wild-type recombinant, while engineering

peptide sequences at the C-terminus resulted in specific activities ranging from 0.06%-120% of the wild-type recombinant luciferase. Waud et al. also disclose that addition of a cyclic AMP dependent protein kinase catalytic subunit to a variant luciferase incorporating the kinase recognition sequence, LRRASLG (SEQ ID NO:1), with a serine at amino acid position 543, resulted in a 30% reduction activity. Alkaline phosphatase treatment restored activity. Waud et al. further disclose that the bioluminescent activity of a variant luciferase containing a thrombin recognition sequence, LVPRES (SEQ ID NO:2), with the cleavage site positioned between amino acids 542 and 543, decreased by 50% when incubated in the presence of thrombin.

Ozawa et al. (2001) describe a biosensor based on protein splicing-induced complementation of rationally designed fragments of firefly luciferase. Protein splicing is a posttranslational protein modification through which inteins (internal proteins) are excised out from a precursor fusion protein, ligating the flanking exteins (external proteins) into a contiguous polypeptide. It is disclosed that the N- and C-terminal intein DnaE from *Synechocystis sp.* PCC6803 were each fused respectively to N- and C-terminal fragments of a luciferase. Protein-protein interactions trigger the folding of DnaE intein, resulting in protein splicing, and thereby the extein of ligated luciferase recovers its enzymatic activity. Ozawa et al. disclose that the interaction between known binding partners, phosphorylated insulin receptor substrate 1 (IRS-1) and its target N-terminal SH2 domain of PI 3-kinase, was monitored using a split luciferase in the presence insulin.

Paulmurugan et al. (2002) employed a split firefly luciferase-based assay to monitor the interaction of two proteins, i.e., MyoD and Id, in cell cultures and in mice using both complementation strategy and an intein-mediated reconstitution strategy. To retain reporter activity, in the complementation strategy, fusion proteins need protein interaction, i.e., via the interaction of the protein partners MyoD and Id, while in the reconstitution strategy, the new complete beetle luciferase formed via intein-mediated splicing maintains it activity even in the absence of a continuing interaction between the protein partners.

A protein fragment complementation assay is disclosed in Michnick et al. (U.S. Patent Nos. 6,270,964, 6,294,330 and 6,428,951). Specifically, Michnick describe a split murine dihydrofolate reductase (DHFR) gene-based assay in which an N-terminal fragment of DHFR and a C-terminal fragment of DHFR are each fused to a GCN4 leucine zipper sequence. DHFR activity was detected in cells which expressed both fusion proteins. Michnick et al. also describe another complementation approach in which nested sets of S1 nuclease generated deletions in the aminoglycoside kinase (AK) gene are introduced into a leucine zipper construct, and the resulting sets of constructs introduced to cells and screened for AK activity.

What is needed is an improved recombinant luciferase for use as a biosensor, e.g., in detecting cellular events such as protein-protein interactions, intracellular signal transduction, or physiological transformations, with a high degree of specificity and a high signal sensitivity.

Summary of the Invention

The invention provides an improved gene product, e.g., a modified luciferase such as a modified beetle luciferase, such as a firefly or click beetle luciferase, an anthozoan luciferase such as a *Renilla* luciferase, or a crustacean luciferase, which, in the presence of one or more molecules of interest, such as cAMP, cGMP, a kinase, a phosphatase, or calcium, has one or more altered activities. In one embodiment, the amino acid sequence of the modified luciferase is different than the amino acid sequence of a corresponding unmodified (native, wild-type or parental, e.g., a mutant luciferase with one or more substitutions) luciferase as a result of one or more modifications at a site (residue) or in a region which is tolerant to modification, e.g., tolerant to an insertion, a deletion, circular permutation, or any combination thereof. In one embodiment, the regions which are tolerant to modification include surface loops between secondary structures, such as beta sheets or alpha helices, found on the native, wild-type luciferase. One or more modifications may be internal relative to the N- or C-terminus of the unmodified luciferase, and/or may be at the N- and/or C-terminus of the unmodified luciferase,

e.g., a deletion of luciferase sequences and/or insertion of one or more amino acid residues optionally including luciferase sequences at the modification site, thereby yielding a modified luciferase. A deletion within the scope of the invention includes a deletion of one or more amino acid residues at a site or in a region of a luciferase sequence that is tolerant to a deletion. The modification(s) may include circular permutation and the introduction (insertion) of one or more discreet (isolated) heterologous amino acid sequences, at least one of which directly or indirectly interacts with a molecule of interest, and optionally may include the deletion of one or more amino acids, e.g., at a site(s) or in a region(s) tolerant to modification including the N- and/or C-terminus of the unmodified luciferase, so long as the resulting modified luciferase has bioluminescent activity before and/or after the interaction with the molecule of interest, e.g., bioluminescent activity is altered after interaction with the molecule of interest. In one embodiment, the modification may be the absence of a peptide bond in the modified luciferase between two amino acids which are linked via a peptide bond in the corresponding unmodified luciferase, in conjunction with a peptide bond in the modified luciferase between residues found at or near the N-terminal and C-terminal residues of the corresponding unmodified luciferase, yielding a circularly permuted luciferase, which optionally includes one or more isolated heterologous amino acid sequences, at least one of which directly or indirectly interacts with a molecule of interest. In one embodiment, the one or more heterologous amino acid sequences, which directly or indirectly interact with a molecule of interest, which sequences are in a circularly permuted luciferase at or near sequences corresponding to the N-terminal and/or C-terminal residues of the corresponding unmodified luciferase. In another embodiment, the one or more heterologous amino acid sequences which directly or indirectly interact with a molecule of interest are at or near the N-terminal and/or C-terminal residues of the circularly permuted or noncircularly permuted luciferase. In one embodiment, the one or more heterologous amino acid sequences which directly or indirectly interact with a molecule of interest in a circularly permuted luciferase are at site(s) or in a region(s) tolerant to modification which is/are not at or near the N-terminal and/or C-terminal residues of the circularly permuted

luciferase, i.e., the heterologous sequences are internal to the N- and C-termini. In one embodiment, the circularly permuted luciferase is modified to include two or more heterologous amino acid sequences, which heterologous amino acid sequences are independently at or near sequences corresponding to the N-terminal and/or C-terminal residues of the corresponding unmodified luciferase, at or near the N-terminal and/or C-terminal residues of the circularly permuted luciferase, at site(s) or in a region(s) tolerant to modification which is/are not at or near the N-terminal and/or C-terminal residues of the circularly permuted or noncircularly permuted luciferase, or any combination thereof. In one embodiment, the heterologous amino acid sequences each interact directly or indirectly with a different molecule of interest. In a further embodiment, a circularly permuted luciferase includes at least two heterologous amino acid sequences which interact with each other in the presence or absence of particular exogenous agents. The two heterologous amino acid sequences may contain the same or different sequences. Moreover, the modified luciferase may include deletions at the N- and C-terminus of 1 to about 10 or about 30, residues, or any integer in between, e.g., 15 residues, corresponding to the N- or C-terminus of the unmodified luciferase. The length of the deletion may be greater than 30 residues depending on the particular luciferase and the length of a desirable deletion may be determined by routine deletion analysis. The modified luciferase may be employed to detect reversible interactions, e.g., binding of two or more molecules, formation of disulfide bonds or other conformational changes, changes in conditions, such as pH, temperature or solvent hydrophobicity, or irreversible interactions, via an alteration in the activity of the modified luciferase, such as an alteration in light intensity, color or kinetic profile. The modified luciferase may also be employed to detect interactions that result in structural modifications of the modified luciferase, e.g., phosphorylation by a kinase or bond cleavage by a protease.

As described below, in-frame insertions resulting in modified click beetle luciferases with detectable activity were at residue 21, 25, 117, 358, 376, 379, 398, 399, 400, 401, 402, 403, 405, 406, 407, 409 or 490 of click beetle luciferase, i.e., those residues and/or regions near those residues are tolerant to modification.

As also described below, in-frame insertions resulting in modified firefly luciferases with detectable activity were at residue 7, 121, 233, 267, 294, 303, 361, 540 or 541 of firefly luciferase, i.e., those residues and/or regions near those residues are tolerant to modifications. Additional residues or regions tolerant to modification are also described herein below.

Thus, a beetle luciferase may be modified at a residue, for instance, residue 21, 25, 117, 358, 376, 379, 398, 399, 400, 401, 402, 403, 405, 406, 407, 409 or 490, or in a region corresponding to residue 15 to 30, e.g., residue 21 or 25, residue 112 to 122, e.g., residue 117, residue 352 to 362, for instance, residue 358, residue 371 to 384, e.g., residue 379, residue 393 to 414, or residue 485 to 495, of a click beetle luciferase, or at residue 7, 37, 47, 75, 83, 107, 121, 144, 160, 174, 188, 198, 205, 225, 233, 242, 255, 268, 308, 316, 358, 377, 403, 435, 490 or 540, or in a region corresponding to residue 2 to 12, residue 32 to 53, e.g., residue 32 to 43 or residue 42 to 52, residue 70 to 88, e.g., residue 70 to 80 or residue 78 to 88, residue 102 to 126, e.g., residue 102 to 112 or residue 116 to 126, residue 139 to 165, residue 183 to 203, residue 220 to 247, e.g., residue 228 to 238, residue 262 to 273, residue 303 to 313, residue 353 to 408, residue 485 to 495, or residue 535 to 546 of a firefly luciferase. Corresponding positions may be identified by aligning luciferase sequences using, for instance, sequence alignment programs. Residues or regions in a luciferase tolerant to modification may be employed as sites to circularly permute the luciferase, for an insertion, or to "split" the luciferase into two molecules that may be employed in protein complementation or protein splicing assays.

The invention further includes a modified anthozoan luciferase having at least one modification at a site or in a region which is tolerant to modification, including but not limited to at a residue corresponding to residue 2, 30, 31, 42, 45, 46, 68, 69, 90, 91, 92, 110, 111, 150, 151, 168, 169, 193, 207, 208, 223, 224, 251, 259, 274, or 311 or in a region corresponding to residue 2 to 12, residue 26 to 36, residue 37 to 47, residue 64 to 74, residue 86 to 97, e.g., residue 90 or 91, residue 96 to 116, residue 147 to 157, residue 218 to 234, e.g., residue 223, 234, 228, 229 or

230, or residue 301 to 311 of a *Renilla* luciferase (Genbank ID AF025843). Corresponding positions may be identified by aligning luciferase sequences using, for instance, sequence alignment programs. Residues or regions in a luciferase tolerant to modification may be employed as sites to circularly permute the luciferase, for an insertion, or to "split" the luciferase into two molecules that may be employed in protein complementation or protein splicing assays.

Further included is a modified crustacean luciferase, e.g., a copepod luciferase, having at least one modification at a site or in a region which is tolerant to modification, including but not limited to in a region corresponding to residue 43 to 53, residue 63 to 73, residue 79 to 89, residue 95 to 105, residue 105 to 115, residue 109 to 119, residue 121 to 131 or residue 157 to 168 of a *Gaussia* luciferase, e.g., see Figure 41, or in a region corresponding to residue 45 to55 or residue 79 to 89 of a mature *Oplophorus* luciferase.Corresponding positions may be identified by aligning luciferase sequences using, for instance, sequence alignment programs. Residues or regions in a luciferase tolerant to modification may be employed as sites to circularly permute the luciferase, for an insertion or to "split" the luciferase into two molecules that may be employed in protein complementation or protein splicing assays.

In one embodiment, the modified luciferase has a detectable activity and includes an insertion of one or more amino acids relative to a corresponding unmodified luciferase at a site or in a region which is tolerant to modification, which insertion includes an amino acid sequence which directly interacts with a molecule of interest, e.g., an insertion which includes a recognition sequence for the molecule of interest, or indirectly interacts with the molecule of interest, e.g., via another molecule. In one embodiment, a modified luciferase comprises an insertion of 2 or more, e.g., 3, 4, 5, 10, 20, 50, 100, 200, 300 or more, but less than about 1000, or any integer in between, amino acid residues. For instance, an insertion of an IP3 sequence may include about 700 amino acid residues. In one embodiment, the modified luciferase with an insertion further comprises a deletion of luciferase sequences, e.g., a deletion of 1 or more, but less than about 100, for instance less than 50, 40, 30, 20, 10 or 5, or any integer in between, residues.

In one embodiment, the invention provides circularly permuted luciferases further modified to include an insertion of an amino acid sequence which directly interacts with a molecule of interest, e.g., an insertion which includes a recognition sequence for the molecule of interest, or indirectly acts with the molecule of interest, e.g., via another molecule. For example, as described hereinbelow, luciferases having a N- and/or C- terminus as well as an internal residue or region which are tolerant to modification were circularly permuted at tolerant residues or regions and at different tolerant residues or regions, and one or more heterologous amino acid sequences were inserted, at least one of which directly or indirectly interacts with a molecule of interest. The resulting modified luciferase was shown to have an alteration in detectable activity in the presence of the molecule of interest.

In one embodiment, circularly permuted beetle luciferases, circularly permuted decapod crustecean luciferases (e.g., *Oplophorus* luciferase), or circularly permuted *Renilla* luciferases having a cAMP or cGMP binding site were shown to have altered luciferase activity in the presence of a cyclic nucleotide, e.g., cAMP or cGMP. Cyclic nucleotide binding sites useful in the luciferases of the invention may have G(E/Q/K)(L/K/S/I)(A/I/C/G)(L/I)X(P/V/T/R/E)R(A/T/H/S)(A/S)(V/T/S/N/W) (SEQ ID NO:118), where X is 2 to 6 amino acids. cAMP binding sites (domains) useful in the circularly permuted luciferases of the invention include but are not limited to cAMP binding sites in exchange protein directly activated by cAMP (Epac) (Bos et al., 2003; and see, for instance, NCBI Accession No. AF115480), including Epac 2B, Epac 1, and Epac IIA, cyclic nucleotide gated ion channels such as hyperpolarization-activated cyclic nucleotide modulated channel (Zagotta et al., 2003), neuropathy target esterase (Dremier et al., 2003), PKA regulatory type IIβ subunit (see, e.g., NCBI Accession No. M124921), e.g., PKA IIβA and PKA IIβB, PKA regulatory type Iα subunit, e.g., PKA IαA and PKA IαB, PKG IIA, PKG IIB, and catabolite activating protein. Also described herein, a noncircularly permuted *Renilla* luciferase and a non-circularly permuted decapod crustecean luciferase having a cAMP binding site had altered luciferase activity in the present of cAMP.

cGMP binding sites useful in the circularly permuted luciferases of the invention include but are not limited to cGMP binding sites in a cGMP dependent protein kinase (GK), e.g., GK I, or a GAF regulatory region in phosphodiesterases (PDEs), e.g., PDE2 or PDE5, adenyl cyclases, or FnlA. In one embodiment, the cyclic nucleotide binding domain containing luciferase of the invention further includes a subcellular localization signal, which is useful to detect subcellular localization and/or concentration of cyclic nucleotides.

As described hereinbelow, luciferase biosensors were prepared with insertions of various sequences representing at least four different structural fold classes. In particular, one of the fold classes participates in the modulation of numerous enzymes through different small molecule interactions. Moreover, insertion of an allosteric domain, i.e., one that changes structural conformation upon binding another molecule, into a luciferase of the invention may be used to detect conformational changes, e.g., phosphorylation or protease cleavage.

Hence, in one embodiment, a modified luciferase of the invention comprises an amino acid sequence which is circularly permuted relative to the amino acid sequence of a corresponding luciferase, such as an unmodified wild type luciferase, resulting in a new N- and C-terminus in the circularly permuted luciferase, at least one of which is at a site or in a region which is tolerant to modification, and is engineered to have functionality by introducing a heterologous amino acid sequence which directly or indirectly interacts with, for instance, a cyclic nucleotide. In another embodiment, the circularly permuted luciferase includes other modifications, including but not limited to insertions and/or deletions internal to the N- or C-terminus of the circularly permuted luciferase, for instance, another insertion and/or a deletion, e.g., at or near the N- and C-terminus of the corresponding unmodified luciferase such as at residues corresponding to residues 1 to about 10 or about 30, or any integer in between, of the N-terminus and/or corresponding to the last residue or about the last 30, e.g., last 15, or any integer in between 1 and 30, residues of the C-terminus of the corresponding unmodified luciferase.

In one embodiment, in the absence of the molecule of interest, the activity of a modified luciferase of the invention is less than the activity of a corresponding unmodified luciferase, e.g., the reporter activity of the modified luciferase is about 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 50%, 70% or more, but less than 100% that of a corresponding unmodified luciferase, the activity of which modified luciferase is optionally detectable. In another embodiment, in the absence of the molecule of interest, the activity of a modified luciferase of the invention is substantially the same or greater than the activity of a corresponding unmodified luciferase, e.g., the reporter activity of the modified luciferase of the invention is about 1.5-fold, e.g., at least 2-, 3- or 5-fold or more, that of a corresponding unmodified luciferase. In the presence of the molecule of interest, the activity of the modified luciferase of the invention is detectably altered. For instance, a detectable alteration in activity of a modified luciferase in the presence of the molecule of interest is an alteration of at least 0.001%, 0.01%, 0.1%, 1%, 10%, or 100%, and up to 2-fold, 4-fold, 10-fold, 100-fold, 1,000-fold, 10,000-fold or more, relative to the activity of the modified luciferase in the absence of the molecule of interest. Thus, the physical proximity of the molecule of interest which interacts with a modification present in the modified luciferase but not the corresponding unmodified luciferase, alters, e.g., decreases, eliminates or increases, the activity of the modified luciferase. For example, a modified beetle, anthozoan luciferase or decapod crustecean may be a circularly permuted beetle, anthozoan or decapod crustecean luciferase with a cAMP binding site. The luminescent signal of such a modified luciferase in the presence of cAMP may be decreased, eliminated or increased relative to the luminescent signal of the modified luciferase in the absence of cAMP or the luminescent signal of the corresponding unmodified beetle, anthozoan or decapod crustecean luciferase in the presence or absence of cAMP.

Accordingly, a modified luciferase of the invention may be employed as a biosensor.

The invention also provides an isolated nucleic acid molecule (polynucleotide) comprising a nucleic acid sequence encoding a modified luciferase of the invention. Further provided is an isolated nucleic acid molecule comprising a

nucleic acid sequence encoding fusion protein comprising a modified luciferase and one or more amino acid residues at the N-terminus (a N-terminal fusion partner) and/or C-terminus (a C-terminal fusion partner) of the modified luciferase. Thus, as used herein, a "fusion protein" is a polypeptide which includes one or more amino acids at the N-terminus and/or C-terminus of a modified luciferase of the invention. Preferably, the presence of one or more fusion partners in the fusion protein does not substantially alter the detectable activity of the fusion protein relative to a corresponding modified luciferase. The N- or C-terminal fusion partner may be a sequence used for purification, e.g., a glutathione S-transferase (GST) or a polyHis sequence, a sequence intended to alter a property of the modified luciferase, e.g., a protein destabilization sequence, a protein or nucleic acid interaction sequence (e.g, a binding sequence), a subcellular localization sequence, or a sequence which has a property which is distinguishable from one or more properties of the luciferase in the fusion protein. In one embodiment, the fusion protein comprises a modified luciferase and a fusion partner which is a reporter protein that is different than the luciferase, which reporter protein is useful as an intramolecular control, e.g., a fluorescent protein or another luciferase. In another embodiment, the invention includes a vector comprising a nucleic acid sequence encoding a fusion protein comprising a modified luciferase of the invention and a nucleic acid fragment which encodes a reporter protein that is different than the luciferase in the modified luciferase. Optionally, optimized nucleic acid sequences, e.g., human codon optimized sequences, encoding at least the luciferase, and preferably the modified luciferase or a fusion protein comprising a modified luciferase, are employed in the nucleic acid molecules of the invention, as those optimized sequences can increase the strength of the signal for luciferase. The optimization of nucleic acid sequences is known to the art, see, for example, WO 02/16944.

The invention also includes a stable cell line that expresses a modified luciferase, or fusion protein of the invention, as well as an expression cassette comprising a nucleic acid molecule encoding the modified luciferase or fusion protein of the invention, and a vecto(e.g., a plasmid, virus, or defective viral particles) capable of expressing the nucleic acid molecule of the invention in a host

cell. Preferably, the expression cassette comprises a promoter, e.g., a constitutive or regulatable promoter, operably linked to the nucleic acid sequence. In one embodiment, the expression cassette contains an inducible promoter. Also provided is a host cell, e.g., a prokaryotic cell or an eukaryotic cell such as a plant or vertebrate cell, e.g., a mammalian cell, including but not limited to a human, non-human primate, canine, feline, bovine, equine, ovine or rodent (e.g., rabbit, rat, ferret or mouse) cell, which comprises the expression cassette or vector of the invention, and a kit which comprises the nucleic acid molecule, expression cassette, vector, host cell or modified luciferase or fusion protein of the invention.

A modified luciferase of the invention may be employed in applications where unmodified luciferases cannot, such as, as a functional reporter to measure or detect various conditions or molecules of interest, e.g., steroids via insertion of a hormone receptor binding site, for instance, an estrogen binding domain, a calcium binding domain, a protease via insertion of a protease recognition site, or cyclic nucleotides via insertion of a cyclic nucleotide binding site. For instance, a vector encoding a modified luciferase comprising an insertion of a cAMP binding site, or a modified luciferase comprising an insertion of a cAMP binding site, is mixed with a sample, e.g., a cell, cell lysate, *in vitro* transcription/translation mixture, or supernatant, and the activity of the modified luciferase in the sample detected or determined, e.g., optionally at one or more time points, and optionally relative to a corresponding unmodified luciferase, or similarly modified luciferase having reduced interaction with cAMP (e.g., further modified by mutations to specific amino acids to reduce the binding affinity with cAMP), or a control sample without cAMP or having a differing amount of cAMP. An alteration in luminescent activity in the sample, for instance, over time, and/or relative to a control, e.g., a cell having a specified amount of cAMP, indicates the presence or amount of cAMP in the sample, or change in amount of cAMP related to experimental condition. In one embodiment, a cell is contacted with a vector comprising a promoter, e.g., a regulatable or constitutive promoter, and a nucleic acid sequence encoding a modified luciferase of the invention which comprises an insertion which interacts with the cyclic nucleotide. In one embodiment, a transfected cell is cultured under

conditions in which the promoter induces transient expression of the modified luciferase, and the presence or amount of luminescence determined. In another embodiment, a modified luciferase of the invention which comprises an insertion which interacts with the cyclic nucleotide and a sample suspected of having a cyclic nucleotide are mixed. Then the amount of luminescence is determined. The invention thus provides a method of detecting the amount of a cyclic nucleotide.

In one embodiment, the modified luciferase is a modified anthozoan luciferase such as a modified *Renilla* luciferase. In one embodiment, the modified anthozoan luciferase is a circularly permuted anthozoan luciferase such as a circularly permuted *Renilla* luciferase. In another embodiment, the modified anthozoan luciferase is not circularly permuted. The modified anthozoan luciferase has one or more heterologous amino acid sequences, including at least one which directly or indirectly interacts with a molecule of interest. In one embodiment, the amino acid sequence is one which, during or after interaction with the molecule of interest, undergoes a conformational change, which in turn alters the activity of the luciferase, e.g., a modified *Renilla* luciferase with such an amino acid sequence is useful to detect allosteric interactions.

In one embodiment, the modified luciferase is a modified decapod crustecean luciferase such as a modified *Oplophorus* luciferase. In one embodiment, the modified decapod crustecean luciferase is a circularly permuted decapod crustecean luciferase such as a circularly permuted *Oplophorus* luciferase. In another embodiment, the modified decapod crustecean luciferase is not circularly permuted. The modified decapod crustecean luciferase has one or more heterologous amino acid sequences, including at least one which directly or indirectly interacts with a molecule of interest. In one embodiment, the amino acid sequence is one which, during or after interaction with the molecule of interest, undergoes a conformational change, which in turn alters the activity of the luciferase, e.g., a modified *Oplophorus* luciferase with such an amino acid sequence is useful to detect allosteric interactions.

Exemplary amino acid sequences of interest to fuse to a modified anthozoan luciferase or a modified decapod crustacean luciferase of the invention include but

are not limited to an enterokinase site, a protease cleavage site, e.g., a site for a caspase, for instance, a caspase 3 cleavage site, a caspase 8 cleavage site, PSA, or a viral protease such as a Rhinovirus protease cleavage site, a SARS protease cleavage site, or a TEV protease cleavage site (NLYFQG; SEQ ID NO:119), a cyclic nucleotide binding site, a hormone binding site, a calcium binding domain such as calmodulin which is regulated by EGTA and $CaCl_2$, or a double fusion with sequences that interact with each other and optionally are modulated by an exogenous agent, e.g., FKBP and FRB, where rapamycin induces binding and FK506 promotes dissociation of binding; a domain from PKA-R and a domain from PKA-C, which may be regulated by cAMP; a domain from SH2 and a domain that is capable of being phosphorylated, which may be regulated by for instance a tyrosine kinase or a phosphatase; a domain from 14-3-3t and a domain that is capable of being phosphorylated, which may be regulated by for example, cAMP-PKA; a domain from WW and a domain that is capable of being phosphorylated, which may be regulated by for example a Ser-Thr kinase; a domain from dihydrofolate reductase (DHFR), which may be regulated by methotrexate (MTX) or BisMTX; a domain from gyrase B (GyrB), which may be regulated by coumermycin or novobiocin; or a double fusion with sequences from the same domain. Thus, in one embodiment, the circularly permuted anthozoan luciferase or a modified decapod crustacean luciferase is modified to include two or more heterologous sequences, which heterologous sequences are independently at or near sequences corresponding to the N-terminal and/or C-terminal residues of the corresponding unmodified luciferase, at or near the N-terminal and/or C-terminal residues of the circularly permuted luciferase, at site(s) or in a region(s) tolerant to modification which is not at or near the N-terminal and/or C-terminal residues of the circularly permuted luciferase, or any combination thereof, wherein the two heterologous amino acid sequence may interact with different molecules of interest.

Further provided are methods of identifying one or more agents that directly or indirectly modulate a molecule of interest.

In one embodiment, the invention provides a method to detect, or determine the activity of, a molecule of interest in a cell. The method includes providing a

luminogenic reaction mixture comprising a cell with a vector having a nucleic acid sequence comprising an open reading frame for a modified luciferase, e.g., a modified beetle luciferase. The modified luciferase has an insertion relative to a corresponding unmodified luciferase, which insertion is at a residue or in a region in a luciferase sequence which is tolerant to modification. The insertion includes an amino acid sequence which directly or indirectly interacts with a molecule of interest relative to the corresponding unmodified luciferase. The mixture is at about 20°C to about 47°C, e.g., about 37°C to about 45°C. Luminescence in the mixture is then detected or determined, thereby detecting or determining the presence, amount or activity of the molecule in the cell. As described hereinbelow, incubating a luminogenic reaction mixture with cells encoding a luciferase that is a biosensor for cAMP at physiological temperatures and/or conditions, e.g., about 37°C and/or about 5% CO_2, for a period of time prior to addition of a test agent provided faster responses and a greater dynamic range.

Also provided is the use of a biosensor of the invention for imaging in cells or multicellular organisms, e.g., living mammals.

Brief Description of the Figures

Figure 1. Positions of Tn5 insertions (bolded) in a click beetle luciferase (amino acid sequence corresponds to SEQ ID NO:3).

Figure 2. Amino acid sequence of a parental (unmodified) firefly luciferase (luc+) (SEQ ID NO:210).

Figure 3. Schematic of a luminescent cAMP binding assay with a circularly permuted luciferase.

Figure 4. PKA regulatory subunit type IIβ (RIIβB). X-ray crystal structure of rat RIIβB amino acids 264-412 (PDB 1CX4). RIIβB is rendered as a red ribbon; cAMP is rendered as ball and stick. The primary sequence similarity between rat (amino acids 264-412) and human RIIβB (amino acids 266-414) is 96.6% (program Megallign, DNAStar).

Figure 5A. Circularly permuted firefly luciferase (CPM-FF Luc) expression plasmid. HSV-TK or T7 promoters were utilized to express the circularly permuted firefly luciferase in mammalian cells or in lysates, respectively. Amino acids 544 and 4 of firefly luciferase are linked by a Gly/Ser rich 42 amino acid peptide (SEQ ID NO:196).

Figure 5B. Expression plasmids for CPM-FF Luc fusions to RIIβB (CPM-FF Luc/RIIβB). Unique combinations of restriction enzymes allowed DNA encoding RIIβB to be ligated in-frame to generate plasmids that encode CPM-FF Luc/RIIβB fusion proteins with various X/Y peptide linker lengths (GSTG corresponds to SEQ ID NO:122; GSSG corresponds to SEQ ID NO:197; GSSGGSGGSG corresponds to SEQ ID NO:198, GSGGSGGSSG corresponds to SEQ ID NO:199; GSSGGSGGSGGGSGGSGGSG corresponds to SEQ ID NO:200; and GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:201).

Figure 5C. Epac DNA sequence (SEQ ID NO:15) modified for *E. coli* expression.

Figure 5D. Circularly permuted *Renilla* luciferase (CPM-hRL) expression plasmid and constructs expressing fusions of CPM-hRL to RIIβB (CPM-hRL/RIIβB). Unique combinations of restriction enzymes allowed DNA encoding RIIβB to be ligated in-frame to generate plasmids that encode CPM-hRL/RIIβB fusion proteins with various X/Y peptide linker lengths (GSTG corresponds to SEQ ID NO:122; GSSG corresponds to SEQ ID NO:197; GSSGGSGGSG corresponds to SEQ ID NO:198; GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:201). The Gly/Ser rich 42 amino acid peptide corresponds to SEQ ID NO:196.

Figure 6. SDS-PAGE analysis of *in vitro* transcription/translation products of circularly permuted beetle luciferases with cAMP binding sites. Expression of CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4), (X=10,Y=10), and (X=20,Y=20) amino acid residues.

Figure 7. Functional characterization of CPM-FF Luc/RIIβB based cAMP sensors with X/Y linker lengths of (X=4,Y=4), (X=10,Y=10), and (X=20,Y=20) amino acid residues.

Figure 8A. Dose response experiment using CPM-FF Luc/RIIβB based cAMP sensors with X/Y linker lengths of (X=4,Y=4), (X=10,Y=10), and (X=20,Y=20) amino acid residues.

Figure 8B. Selectivity of the CPM-FF Luc/RIIbB based cAMP sensor with X/Y linker lengths of (X=10,Y=10) amino acid residues.

Figures 9. Homogeneous cAMP assay data from reactions with CPM-FF Luc/RIIβB cAMP biosensor with X/Y linker lengths of (X=10, Y=10).

Figures 10A-B. Comparison of RLU activity for cAMP binding site containing circularly permuted *Renilla* luciferases.

Figures 11A-B. Measurement of cAMP concentrations in lysates of forskolin treated HEK293 cells with two different CPM-FF Luc/RIIβB cAMP biosensors.

Figure 11C. RLU over time in HEK293 cells transiently transfected with DNA encoding CPM-FF Luc/RIIβB based cAMP luciferase biosensor with X/Y linker lengths of (X=10, Y=0).

Figure 12. Functional characterization of the CPM-FF Luc/RIIβB cAMP biosensors with X/Y linker lengths in the set [2x (x = 0-5), 2y (y = 0-5)] amino acid residues. Luciferase activity in the presence and absence of 100 μM cAMP. Linker combinations (10, 2) and (10, 6) not shown.

Figure 13. Functional characterization of the CPM-FF Luc/RIIβB cAMP sensors with X/Y linker lengths in the set [2x (x = 0-5), 2y (y = 0-5)] amino acid residues. Fold induction in luciferase activity in the presence of 100 μM cAMP. Linker combinations (10, 2) and (10, 6) not shown.

Figure 14. Functional characterization of the CPM-FF Luc/RIIβB cAMP sensors with X/Y linker lengths in the sets [10, -2n (n = 1-7)], [10, 2n (n = 1-5)], and [10 + 2n (n = 1-5), 0] amino acid residues. Luciferase activity in the presence or absence of 100 μM cAMP.

Figure 15. Functional characterization of the CPM-FF Luc/RIIβB cAMP sensors with X/Y linker lengths in the sets [10, -2n (n = 1-7)], [10, 2n (n = 1-5)], and [10 + 2n (n = 1-5), 0] amino acid residues. Fold induction in luciferase activity in the presence of 100 μM cAMP.

Figures 16A-B. Comparison of dose response experiment using the CPM-click beetle Luc/RIIβB cAMP sensors with X/Y linker lengths of (X=4, Y=4) and (X=10, Y=4) amino acid residues and the corresponding CPM-FF luciferases.

Figures 17A-B. Comparison of dose response experiment using the CPM-FF Luc/RIαB cAMP sensors with X/Y linker lengths of (X=4, Y=4) and (X=20, Y=20) amino acid residues and the corresponding CPM-FF Luc/RIIβB.

Figures 18A-B. Comparison of dose response experiment using the CPM-thermal stable Luc/RIIβB cAMP sensors with X/Y linker lengths of (X=4, Y=4) and (X=20, Y=20) amino acid residues and the corresponding CPM-FF luciferases.

Figure 19. Monitoring changes in cAMP concentration in HEK293 cells using a CPM-hRL/RIIβB cAMP biosensor with X/Y linker lengths of (X=4, Y=20).

Figure 20. Sequence of CPM-FF Luc (SEQ ID NO:16).

Figure 21. Schematic of CPM-FF Luc GAF constructs. GSTG corresponds to SEQ ID NO:122; GSSG corresponds to SEQ ID NO:197; GSSGGSGGSG corresponds to SEQ ID NO:198; GSGGSGGSSG corresponds to SEQ ID NO:199; GSSGGSGGSGGGSGGSGGSG corresponds to SEQ ID NO:200; GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:201; and the 42 RT control peptide corresponds to SEQ ID NO:196.

Figure 22. RLU for various CPM-FF Luc GAF constructs in the presence and absence of cGMP.

Figure 23. Fold induction with increasing concentrations of cGMP or cAMP for a CPM-FF Luc GAF construct.

Figure 24. Schematic of CPM-FF Luc calcium biosensors. GSTG corresponds to SEQ ID NO:122; GSSG corresponds to SEQ ID NO:197; GSSGGSGGSG corresponds to SEQ ID NO:198, GSGGSGGSSG corresponds to SEQ ID NO:199; GSSGGSGGSGGGSGGSGGSG corresponds to SEQ ID NO:200; and GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:201; the 42 RT control peptide corresponds to SEQ ID NO:196; LEGSGGGG corresponds to SEQ ID NO:202; and GGGGSGPW corresponds to SEQ ID NO:203.

Figure 25. RLU for various CPM-FF Luc calcium biosensors in the presence of $CaCl_2$ or EDTA and EGTA.

Figure 26. Additional sites for modification of a firefly luciferase.

Figure 27. RLU and fold induction *in vitro* for CPM-FF Luc cAMP biosensors at various sites.

Figure 28. RLU and fold induction *in vivo* for CPM-FF Luc cAMP biosensors at various sites.

Figure 29. Constructs with an insertion of RIIβB in a noncircularly permuted *Renilla* luciferase.

Figure 30. RLU and fold induction for the constructs shown in Figure 29.

Figure 31. Constructs with RIIβB in a circularly permuted *Renilla* luciferase and varying linker lengths.

Figure 32. RLU and fold induction for the constructs in Figure 31.

Figure 33. Constructs with RIαB in a circularly permuted *Renilla* luciferase.

Figure 34. RLU and fold induction for the constructs in Figure 33.

Figure 35. Activity test *in vitro*. Construct pBFB287 was used for the 91 site. Following expression using the TnT T7 Coupled Rabbit Reticulocyte Lysate System, 8.5 µL of TNT reaction was mixed with 8.5 µL of 300 mM HEPES/200 mM thiourea (pH about 7.5) supplemented with 1.7 µL of 1 mM cAMP stock or dH_2O; reactions were allowed to incubate at room temperature for approximately 10 minutes. Five µL of each sample was added to a 96 well plate well in triplicate and luminescence was measured using 100 µL of *Renilla* luciferase assay reagent on a Glomax luminometer.

Figure 36. Activity test *in vitro*. Construct 201325.44.H6 was used for the 91 site. Following expression using the TnT T7 Coupled Wheat Germ Extract System, 15 µL of TNT reaction was supplemented with 1.5 µL of 1 mM cAMP stock or dH_2O; reactions were allowed to incubate at room temperature for approximately 10 minutes. 15 µL of this mixture was then added to 75 ul 1X Renilla Lysis Buffer and 20 µL of each sample was added to a 96 well plate well in triplicate and luminescence was measured using 100 µL of *Renilla* luciferase assay reagent on a Glomax luminometer for the 91 and 223 constructs. For the 229 construct, cAMP induction was measured as described in Figure 35.

Figure 37. Transient transfection data for CPM RLuc cAMP biosensors.

Figure 38A. Schematic of a single step assay for GPCR with a CPM FF Luc cAMP biosensor.

Figure 38B. RLU versus increasing forskolin concentration in a CPM FF Luc cAMP assay.

Figure 39. Data from a screen of a library of compounds with a CPM FF Luc cAMP biosensor.

Figure 40. Dose response of particular compounds using a CPM FF Luc cAMP biosensor.

Figure 41. Amino acid sequence of an exemplary copepod luciferase (SEQ ID NO:204; Genbank ID AAG54095).

Figure 42. Comparison of the relative response of a CPM-FF Luc/RIIβ cAMP biosensor at room temperature and 37°C over time.

Figures 43A-B. RLU for a CRE reporter and a CPM-FF Luc/RIIβ cAMP biosensor in the presence of various agonists (A) or antagonists (B) at room temperature and 37°C.

Figure 44. Fold induction over time with cells stably transfected CPM-FF Luc/RIIβ and exposed to different amounts of dopamine at 37°C.

Figure 45. RLU versus log M dopamine at 37°C.

Figure 46. Potency ranking for various agonists at 37°C.

Figure 47. Potency ranking for various antagonists at 37°C.

Figure 48. Potency ranking of agonists of beta2-adrenergic receptor using HEK293/CPM-FF Luc/RIIβ. HEK293 cells stably expressing CPM-FF Luc/RIIB were stimulated with agonists of the endogenous beta-2 adrenergic receptor. Luminescence was measured after 26 minutes incubation at room temperature.

Figure 49. Potency ranking of agonists of beta2-adrenergic receptor using HEK293. HEK293 cells stably expressing CPM-FF Luc/RIIB were incubed with antagonists in the presence of 0.033 μM isopreterenol. Luminescence was measured after 31 minute incubation at room temperature.

Figure 50. Comparison of bioluminescent GPCR assays with various agonists.

Figure 51. Comparison of bioluminescent GPCR assays with various antagonists.

Figure 52. Detection of intracellular changes in cAMP with a CPM RLuc/ RIIβB cAMP biosensor. A) Comparison of detection with different promoters. B) Forskolin induction. C) SK38393 induction. D) Dopamine induction.

Figure 53. Detection of intracellular changes in cAMP in cells with a CPM RLuc/ RIIβB cAMP biosensor

Figure 54. RLU for FLuc contructs with RIIβB and various linker lengths.

Figure 55. Nucleic acid sequences for *Oplophorus* luciferase and fusions constructs thereof (SEQ ID NOs:205, 206, 207, 208, 209)

Figure 56. RLU for *Oplophorus* luciferase fusions in an *in vitro* protein complementation assay (PCA). Fold induction was determined after background subtraction.

Figure 57. SDS-PAGE analysis of *Oplophorus* luciferase (OpLuc) fusions. Lane 1) full length OpLuc; lane 2) co-expressed 50-FRB and FKBP-51; lane 3) 50-FRB; lane 4) FKBP-51; and lane 5) no DNA control.

Figure 58. RLU for *Oplophorus* luciferase fusions in an *in vitro* PCA. Fold induction was determined after background subtraction.

Figure 59. SDS-PAGE analysis of *Oplophorus* luciferase fusions. Lane 1) full length OpLuc; lane 2) co-expressed 84-FRB and FKBP-85; lane 3) 84-FRB; lane 4) FKBP-85; and lane 5) no DNA control.

Figure 60. RLU for *Oplophorus* luciferase fusions in a cell based PCA. N=3. ss=split site. Fold induction was determined after background subtraction.

Figure 61. RLU for *Oplophorus* luciferase circular permuted-like fusions in an *in vitro* PCA. Fold induction was determined after background subtraction.

Figure 62. SDS-PAGE analysis of *Oplophorus* luciferase circular permuted-like fusions. Lane 1) full length OpLuc; lane 2) co-expressed 51-FKBP and FRB-50; lane 3) co-expressed 85-FKBP and FRB-84; lane 4) FRB-50; lane 5) FRB-84; lane 5) 51-FKBP; lane 7) 85-FKBP; and lane 8) no DNA control.

84-FRB; lane 4) FKBP-85; and lane 5) no DNA control.

Figure 63. CP *Oplophorus* luciferase based vector.

Figure 64. Results with RIIbetaB CP *Oplophorus* luciferase based vector. Left column indicates activity of the intact luciferase (control). Second column from left: corresponding construct with 4 aa linkers. Third column from left: -''- with 10 aa linker. Fourth column from left: -''- with 20 aa linker.

Figure 65. Serine/threonine kinase/phosphatase constructs. Peptide sequences specifically identified in the table are: EIYGEFGSSG (SEQ ID NO:267), EIYGEFGSSGGSGGSG (SEQ ID NO:268), EIYGEFGSSGGSGGSGGGSGGSGGSG (SEQ ID NO:269), GSSG (SEQ ID NO:270), GSTSGSGKPGSGEGSEIYGEFGSSG (SEQ ID NO:271), GSTSGSGKPGSGEGSEIYGEFGSGGSGGSSG (SEQ ID NO:272), GSTSGSGKPGSGEGSEIYGEFGSGGSGGSGGGSGGSGGSSG (SEQ ID NO:273), GSTSGSGKPGSGEGSEIYGEFGSGSGGSGGSSG (SEQ ID NO:274), GSTG (SEQ ID NO:275), GSSGGSGGSG (SEQ ID NO:276), GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277), GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:278), GSSGRKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:279), GGSGGSGSSGRKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:280), GSGGSGGSGG (SEQ ID NO:281), GSSGGSGGSGGGSGGSGSSGRKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:282), RKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:283), GGSSGRKRDRLGTLGIGGSSG (SEQ ID NO:284), GGSSGRKRDRLGTLGIGSSGSGGSGG (SEQ ID NO:285), GGSSGRKRDRLGTLGIGSGGSGGSGGTSGGSGGSSG (SEQ ID NO:287), GSSGGSGGSGGGSGGSG (SEQ ID NO:288), GGSSGRKRDRLGTLGIGSSGSGGSGGTSGGSGGSSG (SEQ ID NO:289), GSSGGSGGSGGGRKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:290), GSGGSGGSSG (SEQ ID NO:291), GSSGGSGGSGGGSGGSGGSGRKRDRLGTLGIGGSSGGGSGGGGSGG (SEQ ID NO:292), GSGG (SEQ ID NO:293), and GGSGGGGSGG (SEQ ID NO:294).

Figure 66. *In vitro* Fluc serine/threonine kinase assay.

Figure 67A-DDDDDDDDDDD. Representative nucleic acid sequences.

Detailed Description of the Invention

Definitions

The term "nucleic acid molecule", "polynucleotide", or "nucleic acid sequence" as used herein, refers to nucleic acid, DNA or RNA, that comprises coding sequences necessary for the production of a polypeptide or protein precursor. The encoded polypeptide may be a full-length polypeptide, a fragment thereof (less than full-length), or a fusion of either the full-length polypeptide or fragment thereof with another polypeptide, yielding a fusion polypeptide.

A "nucleic acid", as used herein, is a covalently linked sequence of nucleotides in which the 3′ position of the pentose of one nucleotide is joined by a phosphodiester group to the 5′ position of the pentose of the next, and in which the nucleotide residues (bases) are linked in specific sequence, i.e., a linear order of nucleotides. A "polynucleotide", as used herein, is a nucleic acid containing a sequence that is greater than about 100 nucleotides in length. An "oligonucleotide" or "primer", as used herein, is a short polynucleotide or a portion of a polynucleotide. An oligonucleotide typically contains a sequence of about two to about one hundred bases. The word "oligo" is sometimes used in place of the word "oligonucleotide".

Nucleic acid molecules are said to have a "5′-terminus" (5′ end) and a "3′-terminus" (3′ end) because nucleic acid phosphodiester linkages occur to the 5′ carbon and 3′ carbon of the pentose ring of the substituent mononucleotides. The end of a polynucleotide at which a new linkage would be to a 5′ carbon is its 5′ terminal nucleotide. The end of a polynucleotide at which a new linkage would be to a 3′ carbon is its 3′ terminal nucleotide. A terminal nucleotide, as used herein, is the nucleotide at the end position of the 3′- or 5′-terminus.

DNA molecules are said to have "5′ ends" and "3′ ends" because mononucleotides are reacted to make oligonucleotides in a manner such that the 5′ phosphate of one mononucleotide pentose ring is attached to the 3′ oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotides referred to as the "5′ end" if its 5′ phosphate is not linked to the 3′

oxygen of a mononucleotide pentose ring and as the "3′ end" if its 3′ oxygen is not linked to a 5′ phosphate of a subsequent mononucleotide pentose ring.

As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide or polynucleotide, also may be said to have 5′ and 3′ ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5′ of the "downstream" or 3′ elements. This terminology reflects the fact that transcription proceeds in a 5′ to 3′ fashion along the DNA strand. Typically, promoter and enhancer elements that direct transcription of a linked gene (e.g., open reading frame or coding region) are generally located 5′ or upstream of the coding region. However, enhancer elements can exert their effect even when located 3′ of the promoter element and the coding region. Transcription termination and polyadenylation signals are located 3′ or downstream of the coding region.

The term "codon" as used herein, is a basic genetic coding unit, consisting of a sequence of three nucleotides that specify a particular amino acid to be incorporated into a polypeptide chain, or a start or stop signal. The term "coding region" when used in reference to structural gene refers to the nucleotide sequences that encode the amino acids found in the nascent polypeptide as a result of translation of a mRNA molecule. Typically, the coding region is bounded on the 5′ side by the nucleotide triplet "ATG" which encodes the initiator methionine and on the 3′ side by a stop codon (e.g., TAA, TAG, TGA). In some cases the coding region is also known to initiate by a nucleotide triplet "TTG".

The term "gene" refers to a DNA sequence that comprises coding sequences and optionally control sequences necessary for the production of a polypeptide from the DNA sequence.

As used herein, the term "heterologous" nucleic acid sequence or protein refers to a sequence that relative to a reference sequence has a different source, e.g., originates from a foreign species, or, if from the same species, it may be substantially modified from the original form.

Nucleic acids are known to contain different types of mutations. A "point" mutation refers to an alteration in the sequence of a nucleotide at a single base position from the wild-type sequence. Mutations may also refer to insertion or

deletion of one or more bases, so that the nucleic acid sequence differs from a reference, e.g., a wild-type, sequence.

As used herein, the terms "hybridize" and "hybridization" refer to the annealing of a complementary sequence to the target nucleic acid, i.e., the ability of two polymers of nucleic acid (polynucleotides) containing complementary sequences to anneal through base pairing. The terms "annealed" and "hybridized" are used interchangeably throughout, and are intended to encompass any specific and reproducible interaction between a complementary sequence and a target nucleic acid, including binding of regions having only partial complementarity. Certain bases not commonly found in natural nucleic acids may be included in the nucleic acids of the present invention and include, for example, inosine and 7-deazaguanine. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length of the complementary sequence, base composition and sequence of the oligonucleotide, ionic strength and incidence of mismatched base pairs. The stability of a nucleic acid duplex is measured by the melting temperature, or "T_m". The T_m of a particular nucleic acid duplex under specified conditions is the temperature at which on average half of the base pairs have disassociated.

The term "recombinant DNA molecule" means a hybrid DNA sequence comprising at least two nucleotide sequences not normally found together in nature.

The term "vector" is used in reference to nucleic acid molecules into which fragments of DNA may be inserted or cloned and can be used to transfer DNA segment(s) into a cell and capable of replication in a cell. Vectors may be derived from plasmids, bacteriophages, viruses, cosmids, and the like.

The terms "recombinant vector" and "expression vector" as used herein refer to DNA or RNA sequences containing a desired coding sequence and appropriate DNA or RNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism. Prokaryotic expression vectors include a promoter, a ribosome binding site, an origin of replication for autonomous replication in a host cell and possibly other sequences, e.g. an optional operator sequence, optional restriction enzyme sites. A promoter is defined as a DNA

sequence that directs RNA polymerase to bind to DNA and to initiate RNA synthesis. Eukaryotic expression vectors include a promoter, optionally a polyadenlyation signal and optionally an enhancer sequence.

A polynucleotide having a nucleotide sequence encoding a protein or polypeptide means a nucleic acid sequence comprising the coding region of a gene, or in other words the nucleic acid sequence encodes a gene product. The coding region may be present in either a cDNA, genomic DNA or RNA form. When present in a DNA form, the oligonucleotide may be single-stranded (i.e., the sense strand) or double-stranded. Suitable control elements such as enhancers/promoters, splice junctions, polyadenylation signals, etc. may be placed in close proximity to the coding region of the gene if needed to permit proper initiation of transcription and/or correct processing of the primary RNA transcript. Alternatively, the coding region. Other regulatory elements include, but are not limited to, transcription factor binding sites, splicing signals, polyadenylation signals, termination signals and enhancer elements.

Transcriptional control signals in eukaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription. Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in yeast, insect and mammalian cells. Promoter and enhancer elements have also been isolated from viruses and analogous control elements, such as promoters, are also found in prokaryotes. The selection of a particular promoter and enhancer depends on the cell type used to express the protein of interest. Some eukaryotic promoters and enhancers have a broad host range while others are functional in a limited subset of cell types. For example, the SV40 early gene enhancer is very active in a wide variety of cell types from many mammalian species and has been widely used for the expression of proteins in mammalian cells. Two other examples of promoter/enhancer elements active in a broad range of mammalian cell types are those from the human elongation factor 1 gene and the long terminal repeats of the Rous sarcoma virus; and the human cytomegalovirus.

The term "promoter/enhancer" denotes a segment of DNA containing sequences capable of providing both promoter and enhancer functions (i.e., the functions provided by a promoter element and an enhancer element as described above). For example, the long terminal repeats of retroviruses contain both promoter and enhancer functions. The enhancer/promoter may be "endogenous" or "exogenous" or "heterologous." An "endogenous" enhancer/promoter is one that is naturally linked with a given gene in the genome. An "exogenous" or "heterologous" enhancer/promoter is one that is placed in juxtaposition to a gene by means of genetic manipulation (i.e., molecular biological techniques) such that transcription of the gene is directed by the linked enhancer/promoter.

The presence of "splicing signals" on an expression vector often results in higher levels of expression of the recombinant transcript in eukaryotic host cells. Splicing signals mediate the removal of introns from the primary RNA transcript and consist of a splice donor and acceptor site. A commonly used splice donor and acceptor site is the splice junction from the 16S RNA of SV40.

Efficient expression of recombinant DNA sequences in eukaryotic cells requires expression of signals directing the efficient termination and polyadenylation of the resulting transcript. Transcription termination signals are generally found downstream of the polyadenylation signal and are a few hundred nucleotides in length. The term "poly(A) site" or "poly(A) sequence" as used herein denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript. Efficient polyadenylation of the recombinant transcript is desirable, as transcripts lacking a poly(A) tail are unstable and are rapidly degraded. The poly(A) signal utilized in an expression vector may be "heterologous" or "endogenous." An endogenous poly(A) signal is one that is found naturally at the 3′ end of the coding region of a given gene in the genome. A heterologous poly(A) signal is one which has been isolated from one gene and positioned 3′ to another gene. A commonly used heterologous poly(A) signal is the SV40 poly(A) signal. The SV40 poly(A) signal is contained on a 237 bp *Bam*H I/*Bcl* I restriction fragment and directs both termination and polyadenylation.

Eukaryotic expression vectors may also contain "viral replicons "or "viral origins of replication." Viral replicons are viral DNA sequences that allow for the extrachromosomal replication of a vector in a host cell expressing the appropriate replication factors. Vectors containing either the SV40 or polyoma virus origin of replication replicate to high copy number (up to 10^4 copies/cell) in cells that express the appropriate viral T antigen. In contrast, vectors containing the replicons from bovine papillomavirus or Epstein-Barr virus replicate extrachromosomally at low copy number (about 100 copies/cell).

The term "*in vitro*" refers to an artificial environment and to processes or reactions that occur within an artificial environment. *In vitro* environments include, but are not limited to, test tubes and cell lysates. The term "*in vivo*" refers to the natural environment (e.g., an animal or a cell) and to processes or reaction that occur within a natural environment.

The term "expression system" refers to any assay or system for determining (e.g., detecting) the expression of a gene of interest. Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used. A wide range of suitable mammalian cells are available from a wide range of source (e.g., the American Type Culture Collection, Rockland, MD). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are well known to the art. Expression systems include *in vitro* gene expression assays where a gene of interest (e.g., a reporter gene) is linked to a regulatory sequence and the expression of the gene is monitored following treatment with an agent that inhibits or induces expression of the gene. Detection of gene expression can be through any suitable means including, but not limited to, detection of expressed mRNA or protein (e.g., a detectable product of a reporter gene) or through a detectable change in the phenotype of a cell expressing the gene of interest. Expression systems may also comprise assays where a cleavage event or other nucleic acid or cellular change is detected.

The term "wild-type" as used herein, refers to a gene or gene product that has the characteristics of that gene or gene product isolated from a naturally

occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "wild-type" form of the gene. In contrast, the term "mutant" refers to a gene or gene product that displays modifications in sequence and/or functional properties (i.e., altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

The term "isolated" when used in relation to a nucleic acid, as in "isolated oligonucleotide" or "isolated polynucleotide" refers to a nucleic acid sequence that is identified and separated from at least one contaminant with which it is ordinarily associated in its source. Thus, an isolated nucleic acid is present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids (e.g., DNA and RNA) are found in the state they exist in nature. For example, a given DNA sequence (e.g., a gene) is found on the host cell chromosome in proximity to neighboring genes; RNA sequences (e.g., a specific mRNA sequence encoding a specific protein), are found in the cell as a mixture with numerous other mRNAs that encode a multitude of proteins. However, isolated nucleic acid includes, by way of example, such nucleic acid in cells ordinarily expressing that nucleic acid where the nucleic acid is in a chromosomal location different from that of natural cells, or is otherwise flanked by a different nucleic acid sequence than that found in nature. The isolated nucleic acid or oligonucleotide may be present in single-stranded or double-stranded form. When an isolated nucleic acid or oligonucleotide is to be utilized to express a protein, the oligonucleotide contains at a minimum, the sense or coding strand (i.e., the oligonucleotide may single-stranded), but may contain both the sense and anti-sense strands (i.e., the oligonucleotide may be double-stranded).

By "peptide," "protein" and "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation). The nucleic acid molecules of the invention may also encode a variant of a naturally-occurring protein or polypeptide fragment thereof, which has an amino acid sequence that is at least 85%, 90%, 95% or 99% identical to the

amino acid sequence of the naturally-occurring (native or wild-type) protein from which it is derived. The term "fusion polypeptide" or "fusion protein" refers to a chimeric protein containing a reference protein (e.g., luciferase) joined at the N- and/or C-terminus to one or more heterologous sequences (e.g., a non-luciferase polypeptide). In some embodiments, a modified polypeptide, fusion polypeptide or a portion of a full-length polypeptide of the invention, may retain at least some of the activity of a corresponding full-length functional (nonchimeric) polypeptide. In other embodiments, in the absence of an exogenous agent or molecule of interest, a modified polypeptide, fusion polypeptide or portion of a full-length functional polypeptide of the invention, may lack activity relative to a corresponding full-length functional polypeptide. In other embodiments, a modified polypeptide, fusion polypeptide or portion of a full-length functional polypeptide of the invention in the presence of an exogenous agent may retain at least some or have substantially the same activity, or alternatively lack activity, relative to a corresponding full-length functional polypeptide.

Polypeptide molecules are said to have an "amino terminus" (N-terminus) and a "carboxy terminus" (C-terminus) because peptide linkages occur between the backbone amino group of a first amino acid residue and the backbone carboxyl group of a second amino acid residue. The terms "N-terminal" and "C-terminal" in reference to polypeptide sequences refer to regions of polypeptides including portions of the N-terminal and C-terminal regions of the polypeptide, respectively. A sequence that includes a portion of the N-terminal region of polypeptide includes amino acids predominantly from the N-terminal half of the polypeptide chain, but is not limited to such sequences. For example, an N-terminal sequence may include an interior portion of the polypeptide sequence including bases from both the N-terminal and C-terminal halves of the polypeptide. The same applies to C-terminal regions. N-terminal and C-terminal regions may, but need not, include the amino acid defining the ultimate N-terminus and C-terminus of the polypeptide, respectively.

The term "recombinant protein" or "recombinant polypeptide" as used herein refers to a protein molecule expressed from a recombinant DNA molecule. In

contrast, the term "native protein" is used herein to indicate a protein isolated from a naturally occurring (i.e., a nonrecombinant) source. Molecular biological techniques may be used to produce a recombinant form of a protein with identical properties as compared to the native form of the protein.

The terms "cell," "cell line," "host cell," as used herein, are used interchangeably, and all such designations include progeny or potential progeny of these designations. By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced a nucleic acid molecule of the invention. Optionally, a nucleic acid molecule of the invention may be introduced into a suitable cell line so as to create a stably-transfected cell line capable of producing the protein or polypeptide encoded by the gene. Vectors, cells, and methods for constructing such cell lines are well known in the art. The words "transformants" or "transformed cells" include the primary transformed cells derived from the originally transformed cell without regard to the number of transfers. All progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Nonetheless, mutant progeny that have the same functionality as screened for in the originally transformed cell are included in the definition of transformants.

The term "homology" refers to a degree of complementarity between two or more sequences. There may be partial homology or complete homology (i.e., identity). Homology is often measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group. University of Wisconsin Biotechnology Center. 1710 University Avenue. Madison, WI 53705). Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, insertions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

The term "isolated" when used in relation to a polypeptide, as in "isolated protein" or "isolated polypeptide" refers to a polypeptide that is identified and

separated from at least one contaminant with which it is ordinarily associated in its source. Thus, an isolated polypeptide is present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated polypeptides (e.g., proteins and enzymes) are found in the state they exist in nature.

The term "purified" or "to purify" means the result of any process that removes some of a contaminant from the component of interest, such as a protein or nucleic acid. The percent of a purified component is thereby increased in the sample.

As used herein, "pure" means an object species is the predominant species present (i.e., on a molar basis it is more abundant than any other individual species in the composition), and preferably a substantially purified fraction is a composition wherein the object species comprises at least about 50 percent (on a molar basis) of all macromolecular species present. Generally, a "substantially pure" composition will comprise more than about 80 percent of all macromolecular species present in the composition, more preferably more than about 85%, about 90%, about 95%, and about 99%. Most preferably, the object species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single macromolecular species.

The term "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of sequences encoding amino acids in such a manner that a functional (e.g., enzymatically active, capable of binding to a binding partner, capable of inhibiting, etc.) protein or polypeptide is produced.

As used herein, the term "poly-histidine tract" or (His tag) refers to a molecule comprising two to ten histidine residues, e.g., a poly-histidine tract of five to ten residues. A poly-histidine tract allows the affinity purification of a covalently linked molecule on an immobilized metal, e.g., nickel, zinc, cobalt or copper, chelate column or through an interaction with another molecule (e.g., an antibody reactive with the His tag).

A "protein destabilization sequence" includes, but is not limited to, a PEST sequence, for example, a PEST sequence from cyclin, e.g., mitotic cyclins, uracil permease or ODC, a sequence from the C-terminal region of a short-lived protein such as ODC, early response proteins such as cytokines, lymphokines, protooncogenes, e.g., c-myc or c-fos, MyoD, HMG CoA reductase, or S-adenosyl methionine decarboxylase, CL sequences, a cyclin destruction box, or N-degron.

As used herein, a "marker gene" or "reporter gene" is a gene that imparts a distinct phenotype to cells expressing the gene and thus permits cells having the gene to be distinguished from cells that do not have the gene. Such genes may encode either a selectable or screenable marker, depending on whether the marker confers a trait which one can 'select' for by chemical means, i.e., through the use of a selective agent (e.g., a herbicide, antibiotic, or the like), or whether it is simply a "reporter" trait that one can identify through observation or testing, i.e., by 'screening'. Elements of the present disclosure are exemplified in detail through the use of particular marker genes. Of course, many examples of suitable marker genes or reporter genes are known to the art and can be employed in the practice of the invention. Therefore, it will be understood that the following discussion is exemplary rather than exhaustive. In light of the techniques disclosed herein and the general recombinant techniques which are known in the art, the present invention renders possible the alteration of any gene. Exemplary reporter proteins are encoded by nucleic acid molecules comprising modified reporter genes including, but are not limited to, modifications of a *neo* gene, a β-gal gene, a *gus* gene, a *cat* gene, a *gpt* gene, a *hyg* gene, a *hisD* gene, a *ble* gene, a *mprt* gene, a *bar* gene, a nitrilase gene, a galactopyranoside gene, a xylosidase gene, a thymidine kinase gene, an arabinosidase gene, a mutant acetolactate synthase gene (ALS) or acetoacid synthase gene (AAS), a methotrexate-resistant *dhfr* gene, a dalapon dehalogenase gene, a mutated anthranilate synthase gene that confers resistance to 5-methyl tryptophan (WO 97/26366), an R-locus gene, a β-lactamase gene, a *xyl*E gene, an α-amylase gene, a tyrosinase gene, a luciferase (*luc*) gene, (e.g., a *Renilla reniformis* luciferase gene, a firefly luciferase gene, or a click beetle luciferase

(*Pyrophorus plagiophthalamus*) gene), an aequorin gene, a red fluorescent protein gene, or a green fluorescent protein gene.

All amino acid residues identified herein are in the natural L-configuration. In keeping with standard polypeptide nomenclature, abbreviations for amino acid residues are as shown in the following Table of Correspondence.

TABLE OF CORRESPONDENCE

1-Letter	3-Letter	AMINO ACID
Y	Tyr	L-tyrosine
G	Gly	L-glycine
F	Phe	L-phenylalanine
M	Met	L-methionine
A	Ala	L-alanine
S	Ser	L-serine
I	Ile	L-isoleucine
L	Leu	L-leucine
T	Thr	L-threonine
V	Val	L-valine
P	Pro	L-proline
K	Lys	L-lysine
H	His	L-histidine
Q	Gln	L-glutamine
E	Glu	L-glutamic acid
W	Trp	L-tryptophan
R	Arg	L-arginine
D	Asp	L-aspartic acid
N	Asn	L-asparagine
C	Cys	L-cysteine

I. Methods to Identify Residues or Regions of a Luciferase Which Are Tolerant to Modification

Numerous methods are available to identify sites and/or regions in a luciferase gene which may be modified, e.g., disrupted, yet when transcribed and translated, yield a desirable, for instance, a readily detectable, gene product. For instance, amplification reactions may be employed to delete and/or insert nucleotides for one or more amino acid residues in a luciferase gene. Alternatively, transposons may be employed to prepare libraries of insertional mutations. Transposons are mobile DNA sequences found in the genomes of prokaryotes and eukaryotes. Transposon tagging has long been recognized as a powerful research tool for randomly distributing primer binding sites, creating gene "knockouts," and introducing a physical tag or a genetic tag into large target DNAs. Insertions in a reporter gene useful to prepare the modified luciferases of the invention are those which are internal, in frame insertions in the coding region for the luciferase.

One frequently used transposition system is the Tn5 system isolated from gram-negative bacteria. The Tn5 transposase is a small, single subunit enzyme that has been cloned and purified to high specific activity, and carries out transposition without the need for host cell factors. Moreover, Tn5 transposon insertions into target DNA are highly random, and proceed by a simple process. Tn5 transposase will transpose any DNA sequence contained between its short 19 basepair Mosaic End (ME) Tn5 transposase recognition sequences.

The GPS-M Mutagenesis System uses TnsABC* Transposase to insert a Tn7-based transposon randomly into a DNA target. Target DNA may be a plasmid, cosmid, BAC or purified chromosomal DNA. If the insertion site is within a translated gene segment, this will normally result in a null (loss of function) mutation. There is minimal site preference for insertion, so disruption of any open reading frame is possible. Due to target immunity, only one insertion occurs per DNA molecule *in vivo* over a distance of about 190 kb. Therefore, the *in vitro* reaction produces a population of target DNA molecules each containing the transposable element at a different position.

The transposon donor can be modified by adding to or replacing the antibiotic, e.g., kanamycin, resistance marker. The donor plasmid may be grown in standard laboratory *E. coli* strains, and the vector backbone carries a different

antibiotic marker, e.g., Amp^r, than the transposon and an origin of replication. To destroy unreacted donor molecules and avoid undesirable reaction products, the donor can be destroyed by digestion with a rare-cutting enzyme, for instance, PI-*Sce*I (VDE). For applications in which the mutagenized DNA is transformed into naturally-competent organisms (which take up single DNA strands), the gaps are filled-in and ligated.

Once sites tolerant to modification in a luciferase sequence are identified, insertions, deletions and permutations, or any combination thereof, of the sequences may be prepared. With regard to permuted sequences, Plainkum et al. (2003) reported that circularly permuted forms of ribonuclease A having new N- and C-termini and a peptide linker containing a protease recognition site linking the original N- and C-termini had reduced ribonuclease activity due to steric occlusion of the active site. Plainkum et al. found that cleavage of the circularly permuted ribonuclease A with the protease increased the activity of the protein, presumably by removing the block to the active site. In the case of luciferase, the N- and C-termini are separated by about 40 angstroms, a distance equivalent to 5-6 amino acids. Circularly permuted firefly luciferases were prepared, one of which had a new N-terminus at Asp(234) and a new C-terminus at Pro(233) and a recognition site for the protease enterokinase which cleaves on the carboxyl terminal side of Asp(4)Lys (see U.S. published application 20050153310 and PCT/US2004/032705). The activity of the fused mutant protein was increased about 90- to about 150-fold by treatment with enterokinase (Figure 3). Other biosensors included a caspase-3 DEVD cleavage site (Figure 3), a PSA cleavage site, e.g., Ala-Asn-Lys-Ile-Ser-Tyr-Gln-Ser-Ser-Ser-Thr-Glu (SEQ ID NO:17), a Rhinovirus protease site, e.g., Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro (SEQ ID NO:19), and a SARS virus protease site, e.g., TSAVLQSGFR (SEQ ID NO:20), inserted into the circularly permuted firefly luciferase or a click beetle luciferase (CP1: R = Asn401 and CP2: R=Arg223) (see U.S. published application 20050153310 and PCT/US2004/032705). CP2 has an insertion at a position in click beetle luciferase which corresponds to position 234 in firefly luciferase. As described hereinbelow, circularly permuted *Renilla* lucferases were prepared.

The biosensors of the invention include but are not limited to those in which the heterologous amino acid sequence includes a protein binding domain, such as one that binds IL-17RA, e.g., IL-17A, or the IL-17A binding domain of IL-17RA, Jun binding domain of Erg, or the EG binding domain of Jun; a potassium channel voltage sensing domain, e.g., one useful to detect protein conformational changes, the GTPase binding domain of a Cdc42 or rac target, or other GTPase binding domains, domains associated with kinase or phosphotase activity, e.g., regulatory myosin light chain, PKCδ, pleckstrin containing PH and DEP domains, other phosphorylation recognition domains and substrates; glucose binding protein domains, glutamate/aspartate binding protein domains, PKA or a cAMP-dependent binding substrate, InsP3 receptors, GKI, PDE, estrogen receptor ligand binding domains, apoK1-er, or calmodulin binding domains.

In one embodiment, the biosensor is useful to detect a GTPase, e.g., binding of Cdc42 or Rac to a EBFP, EGFP PAK fragment, Raichu-Rac, Raichu-Cdc42, integrin alphavbeta3, IBB of importin-a, DMCA or NBD-Ras of CRaf1 (for Ras activation), binding domain of Ras/Rap Ral RBD with Ras prenylation sequence. In one embodiment, the biosensor detects PI(4,5)P2 (e.g., using PH-PCLdelta1, PH-GRP1), PI(4,5)P2 or PI(4)P (e.g., PH-OSBP), PI(3,4,5)P3 (e.g., using PH-ARNO, or PH-BTK, or PH-Cytohesin1), PI(3,4,5)P3 or PI(3,4)P2 (e.g., using PH Akt), PI(3)P (e.g., using FYVE-EEA1), or Ca2+ (cytosolic) (e.g., using calmodulin, or C2 domain of PKC.

In one embodiment, the domain is one with a phosphorylated tyrosine (e.g., in Src, Ab1 and EGFR), that detects phosphorylation of ErbB2, phosphorylation of tyrosine in Src, Ab1 and EGFR, activation of MKA2 (e.g., using MK2), cAMP induced phosphorylation, activation of PKA, e.g., using KID of CREG, phosphorylation of CrkII, e.g., using SH2 domain pTyr peptide, binding of bZIP transcription factors and REL proteins, e.g., bFos and bJun ATF2 and Jun, or p65 NFkappaB, or microtubule binding, e.g., using kinesin.

Thus, the invention includes luciferase biosensors including circularly permuted luciferases, which luciferase sequence may include deletions of residues at the original (wild type) N- or C-termini, or both, e.g., deletion of 1 to 3 or more

residues at the N-terminus and 1 to 6 or more residues at the C-terminus, as well as sequences that directly or indirectly interact with a molecule of interest.

II. Exemplary Polynucleotides and Proteins

The invention includes a modified luciferase encompassing any amino acid sequence which provides a polypeptide having a detectable activity, e.g., luminescent activity, as well as protein fragments thereof, which are recombinantly or synthetically synthesized. The luciferase sequences of a modified luciferase are the same or are substantially the same as the amino acid sequence of a corresponding unmodified luciferase. A polypeptide or peptide having substantially the same sequence means that an amino acid sequence is largely, but may not entirely be, the same and retains a functional activity of the sequence to which it is related. In general, two amino acid sequences are substantially the same or substantially homologous if they are at least 70% identical, e.g., have at least 80%, 90%, 95% or more identity.

Homology or identity is often measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Such software matches similar sequences by assigning degrees of homology to various deletions, substitutions and other modifications. The terms "homology" and "identity" in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same when compared and aligned for maximum correspondence over a comparison window or designated region as measured using any number of sequence comparison algorithms or by manual alignment and visual inspection.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then

calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

Methods of alignment of sequence for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted by the local homology algorithm of Smith et al. (1981), by the homology alignment algorithm of Needleman et al. (1970), by the search for similarity method of Person et al. (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection.

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins et al. (1988); Higgins et al. (1989); Corpet et al.(1988); Huang et al. (1992); and Pearson et al. (1994). The ALIGN program is based on the algorithm of Myers and Miller (1988). The BLAST programs of Altschul et al. (1990), are based on the algorithm of Karlin and Altschul (1990).

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., 1990). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative

alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul (1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a test nucleic acid sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid sequence to the reference nucleic acid sequence is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al. (1997). Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al., *supra*. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g. BLASTN for nucleotide sequences, BLASTX for proteins) can be used. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, 1989). See http://www.ncbi.nlm.nih.gov.

In particular, a polypeptide may be substantially related but for a conservative variation. A conservative variation denotes the replacement of an

amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. Other illustrative examples of conservative substitutions include the changes of: alanine to serine; arginine to lysine; asparagine to glutamine or histidine; aspartate to glutamate; cysteine to serine; glutamine to asparagine; glutamate to aspartate; glycine to proline; histidine to asparagine or glutamine; isoleucine to leucine or valine; leucine to valine or isoleucine; lysine to arginine, glutamine, or glutamate; methionine to leucine or isoleucine; phenylalanine to tyrosine, leucine or methionine; serine to threonine; threonine to serine; tryptophan to tyrosine; tyrosine to tryptophan or phenylalanine; valine to isoleucine to leucine.

In one embodiment, a polynucleotide of the invention is optimized for expression in a particular host. As used herein, optimization includes codon optimization as well as, in eukaryotic cells, introduction of a Kozak sequence, and/or one or more introns. Thus, a nucleic acid molecule may have a codon composition that differs from that of a wild-type nucleic acid sequence encoding an unmodified luciferase at more than 30%, 35%, 40% or more than 45%, e.g., 50%, 55%, 60% or more of the codons. Preferred codons for use in the invention are those which are employed more frequently than at least one other codon for the same amino acid in a particular organism and, more preferably, are also not low-usage codons in that organism and are not low-usage codons in the organism used to clone or screen for the expression of the nucleic acid molecule. Moreover, preferred codons for certain amino acids (i.e., those amino acids that have three or more codons,), may include two or more codons that are employed more frequently than the other (non-preferred) codon(s). The presence of codons in the nucleic acid molecule that are employed more frequently in one organism than in another organism results in a nucleic acid molecule which, when introduced into the cells of the organism that employs those codons more frequently, is expressed in those cells

at a level that is greater than the expression of the wild-type or parent nucleic acid sequence in those cells.

In one embodiment of the invention, the codons that are different are those employed more frequently in a mammal, while in another embodiment the codons that are different are those employed more frequently in a plant. A particular type of mammal, e.g., human, may have a different set of preferred codons than another type of mammal. Likewise, a particular type of plant may have a different set of preferred codons than another type of plant. In one embodiment of the invention, the majority of the codons which differ are ones that are preferred codons in a desired host cell. Preferred codons for mammals (e.g., humans) and plants are known to the art (e.g., Wada et al., 1990). For example, preferred human codons include, but are not limited to, CGC (Arg), CTG (Leu), TCT (Ser), AGC (Ser), ACC (Thr), CCA (Pro), CCT (Pro), GCC (Ala), GGC (Gly), GTG (Val), ATC (Ile), ATT (Ile), AAG (Lys), AAC (Asn), CAG (Gln), CAC (His), GAG (Glu), GAC (Asp), TAC (Tyr), TGC (Cys) and TTC (Phe) (Wada et al., 1990). Thus, preferred "humanized" synthetic nucleic acid molecules of the invention have a codon composition which differs from a wild type nucleic acid sequence by having an increased number of the preferred human codons, e.g. CGC, CTG, TCT, AGC, ACC, CCA, CCT, GCC, GGC, GTG, ATC, ATT, AAG, AAC, CAG, CAC, GAG, GAC, TAC, TGC, TTC, or any combination thereof. For example, the nucleic acid molecule of the invention may have an increased number of CTG or TTG leucine-encoding codons, GTG or GTC valine-encoding codons, GGC or GGT glycine-encoding codons, ATC or ATT isoleucine-encoding codons, CCA or CCT proline-encoding codons, CGC or CGT arginine-encoding codons, AGC or TCT serine-encoding codons, ACC or ACT threonine-encoding codon, GCC or GCT alanine-encoding codons, or any combination thereof, relative to the wild-type nucleic acid sequence. Similarly, nucleic acid molecules having an increased number of codons that are employed more frequently in plants, have a codon composition which differs from a wild-type nucleic acid sequence by having an increased number of the plant codons including, but not limited to, CGC (Arg), CTT (Leu), TCT (Ser), TCC (Ser), ACC (Thr), CCA (Pro), CCT (Pro), GCT (Ser), GGA (Gly), GTG (Val), ATC

(Ile), ATT (Ile), AAG (Lys), AAC (Asn), CAA (Gln), CAC (His), GAG (Glu), GAC (Asp), TAC (Tyr), TGC (Cys), TTC (Phe), or any combination thereof (Murray et al., 1989). Preferred codons may differ for different types of plants (Wada et al., 1990).

The modified luciferase proteins or fusion proteins of the invention may be prepared by recombinant methods or by solid phase chemical peptide synthesis methods. Such methods have been known in the art since the early 1960's (Merrifield, 1963) (See also Stewart et al., Solid Phase Peptide Synthesis, 2nd ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) and have recently been employed in commercially available laboratory peptide design and synthesis kits (Cambridge Research Biochemicals). Such commercially available laboratory kits have generally utilized the teachings of Geysen et al. (1984) and provide for synthesizing peptides upon the tips of a multitude of "rods" or "pins" all of which are connected to a single plate. When such a system is utilized, a plate of rods or pins is inverted and inserted into a second plate of corresponding wells or reservoirs, which contain solutions for attaching or anchoring an appropriate amino acid to the pin' s or rod' s tips. By repeating such a process step, e.g., inverting and inserting the rod' s and pin' s tips into appropriate solutions, amino acids are built into desired peptides. In addition, a number of available FMOC peptide synthesis systems are available. For example, assembly of a polypeptide or fragment can be carried out on a solid support using an Applied Biosystems, Inc. Model 431A automated peptide synthesizer. Such equipment provides ready access to the peptides of the invention, either by direct synthesis or by synthesis of a series of fragments that can be coupled using other known techniques.

III. Fusion Partners Useful with the Modified Luciferase of the Invention

The polynucleotide of the invention which encodes a modified luciferase may be employed with other nucleic acid sequences, e.g., a native sequence such as a cDNA or one which has been manipulated *in vitro*, e.g., to prepare N-terminal, C-terminal, or N- and C-terminal fusion proteins, e.g., a fusion with a protein encoded by a different reporter gene including a selectable marker. Many examples of

suitable fusion partners are known to the art and can be employed in the practice of the invention.

Fusion partners include but are not limited to affinity domains or other functional protein sequences, such as those having an enzymatic activity. For example, a functional protein sequence may encode a kinase catalytic domain (Hanks and Hunter, 1995), producing a fusion protein that can enzymatically add phosphate moieties to particular amino acids, or may encode a Src Homology 2 (SH2) domain (Sadowski et al., 1986; Mayer and Baltimore, 1993), producing a fusion protein that specifically binds to phosphorylated tyrosines.

Affinity domains are generally peptide sequences that can interact with a binding partner, e.g., such as one immobilized on a solid support. DNA sequences encoding multiple consecutive single amino acids, such as histidine, when fused to the expressed protein, may be used for one-step purification of the recombinant protein by high affinity binding to a resin column, such as nickel sepharose. Sequences encoding peptides, such as the chitin binding domain (which binds to chitin), glutathione-S-transferase (which binds to glutathione), biotin (which binds to avidin and strepavidin), and the like, can also be used for facilitating purification of the protein of interest. The affinity domain can be separated from the protein of interest by methods well known in the art, including the use of inteins (protein self-splicing elements (Chong et al., 1997). Exemplary affinity domains include HisV5 (HHHHH) (SEQ ID NO:4), HisX6 (HHHHHH) (SEQ ID NO:5), C-myc (EQKLISEEDL) (SEQ ID NO:6), Flag (DYKDDDDK) (SEQ ID NO:7), SteptTag (WSHPQFEK) (SEQ ID NO:8), hemagluttinin, e.g., HA Tag (YPYDVPDYA) (SEQ ID NO:9), GST, thioredoxin, cellulose binding domain, RYIRS (SEQ ID NO:10), Phe-His-His-Thr (SEQ ID NO:11), chitin binding domain, S-peptide, T7 peptide, SH2 domain, C-end RNA tag, WEAAAREACCRECCARA (SEQ ID NO:12), metal binding domains, e.g., zinc binding domains or calcium binding domains such as those from calcium-binding proteins, e.g., calmodulin, troponin C, calcineurin B, myosin light chain, recoverin, S-modulin, visinin, VILIP, neurocalcin, hippocalcin, frequenin, caltractin, calpain large-subunit, S100 proteins, parvalbumin, calbindin D_{9K}, calbindin D_{28K}, and calretinin, inteins, biotin,

streptavidin, MyoD, Id, leucine zipper sequences, and maltose binding protein. In one embodiment, the fusion partner is a sequence useful to purify a fusion protein, e.g., a His or GST tag, and in one embodiment the purification tag is fused to the N- or C-terminus of a circularly permuted luciferase.

Another class of fusion partners includes a protein encoded by a reporter gene, including, but are not limited to, a *neo* gene, a β-gal gene, a *gus* gene, a *cat* gene, a *gpt* gene, a *hyg* gene, a *hisD* gene, a *ble* gene, a *mprt* gene, a *bar* gene, a nitrilase gene, a galactopyranoside gene, a xylosidase gene, a thymidine kinase gene, an arabinosidase gene, a mutant acetolactate synthase gene (ALS) or acetoacid synthase gene (AAS), a methotrexate-resistant *dhfr* gene, a dalapon dehalogenase gene, a mutated anthranilate synthase gene that confers resistance to 5-methyl tryptophan (WO 97/26366), an R-locus gene, a β-lactamase gene, a *xyl*E gene, an α-amylase gene, a tyrosinase gene, an anthozoan luciferase (*luc*) gene, (e.g., a *Renilla reniformis* luciferase gene), an aequorin gene, a red fluorescent protein gene, or a green fluorescent protein gene. Included within the terms selectable or screenable marker genes are also genes which encode a "secretable marker" whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected by their catalytic activity. Secretable proteins fall into a number of classes, including small, diffusible proteins detectable, e.g., by ELISA, and proteins that are inserted or trapped in the cell membrane.

IV. Vectors and Host Cells Encoding the Modified Luciferase or Fusions Thereof

Once a desirable nucleic acid molecule encoding a modified luciferase or a fusion thereof is prepared, an expression cassette encoding the modified luciferase or a fusion protein comprising the modified luciferase is prepared. For example, a nucleic acid molecule comprising a nucleic acid sequence encoding a modified luciferase is optionally operably linked to transcription regulatory sequences, e.g., one or more enhancers, a promoter, a transcription termination sequence or a combination thereof, to form an expression cassette. The nucleic acid molecule or

expression cassette may be introduced to a vector, e.g., a plasmid or viral vector, which optionally includes a selectable marker gene, and the vector introduced to a cell of interest, for example, a prokaryotic cell such as *E. coli*, *Streptomyces* spp., *Bacillus* spp., *Staphylococcus* spp. and the like, as well as eukaryotic cells including a plant (dicot or monocot), fungus, yeast, e.g., *Pichia*, *Saccharomyces* or *Schizosaccharomyces*, or a mammalian cell. Preferred mammalian cells include bovine, caprine, ovine, canine, feline, non-human primate, e.g., simian, and human cells. Preferred mammalian cell lines include, but are not limited to, CHO, COS, 293, Hela, CV-1, SH-SY5Y, HEK293, and NIH3T3 cells.

The expression of an encoded modified luciferase may be controlled by any promoter capable of expression in prokaryotic cells or eukaryotic cells. Preferred prokaryotic promoters include, but are not limited to, SP6, T7, T5, *tac*, *bla*, *trp*, *gal*, *lac* or maltose promoters. Preferred eukaryotic promoters include, but are not limited to, constitutive promoters, e.g., viral promoters such as CMV, SV40 and RSV promoters, as well as regulatable promoters, e.g., an inducible or repressible promoter such as the tet promoter, the hsp70 promoter and a synthetic promoter regulated by CRE. The nucleic acid molecule, expression cassette and/or vector of the invention may be introduced to a cell by any method including, but not limited to, calcium-mediated transformation, electroporation, microinjection, lipofection and the like.

V. Exemplary Uses

The modified luciferases or fusions thereof are useful for any purpose including, but not limited to, detecting the amount or presence of a particular molecule (a biosensor), isolating a particular molecule, detecting conformational changes in a particular molecule, e.g., due to binding, phosphorylation or ionization, facilitating high or low throughput screening, detecting protein-protein, protein-DNA or other protein-based interactions, or selecting or evolving biosensors. For instance, a modified luciferase or a fusion thereof, is useful to detect, e.g., in an *in vitro* or cell-based assay, the amount, presence or activity of a particular kinase (for example, by inserting a kinase site into a reporter protein), RNAi (e.g., by inserting a sequence suspected of being recognized by RNAi into a coding sequence for a

reporter protein, then monitoring reporter activity after addition of RNAi), or protease, such as one to detect the presence of a particular viral protease, which in turn is indicator of the presence of the virus, or an antibody; to screen for inhibitors, e.g., protease inhibitors; to identify recognition sites or to detect substrate specificity, e.g., using a modified luciferase with a selected recognition sequence or a library of modified luciferases having a plurality of different sequences with a single molecule of interest or a plurality (for instance, a library) of molecules; to select or evolve biosensors or molecules of interest, e.g., proteases; or to detect protein-protein interactions via complementation or binding, e.g., in an *in vitro* or cell-based approach. In one embodiment, a modified luciferase which includes an insertion is contacted with a random library or mutated library of molecules, and molecules identified which interact with the insertion. In another embodiment, a library of modified luciferases having a plurality insertions is contacted with a molecule, and modified luciferases which interact with the molecule identified. In one embodiment, a modified luciferase or fusion thereof, is useful to detect, e.g., in an *in vitro* or cell-based assay, the amount or presence of cAMP or cGMP (for example, by inserting a cAMP or cGMP binding site into a circularly permuted luciferase), to screen for inhibitors or activators, e.g., inhibitors or activators of cAMP or cGMP, inhibitors or activators of cAMP binding to a cAMP binding site or inhibitors or activators of G protein coupled receptors (GPCR), to identify recognition sites or to detect substrate specificity, e.g., using a modified luciferase with a selected recognition sequence or a library of modified luciferases having a plurality of different sequences with a single molecule of interest or a plurality (for instance, a library) of molecules, to select or evolve cAMP or cGMP binding sites, or in whole animal imaging.

The invention also provides methods to monitor the expression, location and/or trafficking of molecules in a cell, as well as to monitor changes in microenvironments within a cell, using a modified luciferase or a fusion protein thereof. In one embodiment, a modified luciferase comprises a recognition site for a molecule, and when the molecule interacts with the recognition site, that results in an increase in activity, and so can be employed to detect or determine the presence

or amount of the molecule. For example, in one embodiment, a modified luciferase comprises an internal insertion containing two domains which interact with each other under certain conditions. In one embodiment, one domain in the insertion contains an amino acid which can be phosphorylated and the other domain is a phosphoamino acid binding domain. In the presence of the appropriate kinase or phosphatase, the two domains in the insertion interact and change the conformation of the modified luciferase resulting in an alteration in the detectable activity of the modified luciferase. In another embodiment, a modified luciferase comprises a recognition site for a molecule, and when the molecule interacts with the recognition site, results in an increase in activity, and so can be employed to detect or determine the presence of amount or the other molecule.

Two-hybrid systems are extremely powerful methods for detecting protein:protein interactions *in vivo* as well as identifying residues/domains involved in protein:protein interactions. The basis of two-hybrid systems is the modular domains found in some transcription factors: a DNA-binding domain, which binds to a specific DNA sequence, and a transcriptional activation domain, which interacts with the basal transcriptional machinery (Sadowski, 1988). A transcriptional activation domain in association with a DNA-binding domain may promote the assembly of RNA polymerase II complexes at the TATA box and increase transcription. In the CheckMate™ Mammalian Two-Hybrid System (Promega Corp., Madison, WI), the DNA-binding domain and the transcriptional activation domain, produced by separate plasmids, are closely associated when one protein ("X") fused to a DNA-binding domain interacts with a second protein ("Y") fused to a transcriptional activation domain. In this system, interaction between proteins X and Y results in transcription of either a reporter gene or a selectable marker gene. In particular, the pBIND Vector contains a yeast GAL4 DNA-binding domain upstream of a multiple cloning region, and a pACT Vector contains the herpes simplex virus VP16 activation domain upstream of a multiple cloning region. In addition, the pBIND Vector expresses the *Renilla reniformis* luciferase. The two genes encoding the two potentially interactive proteins of interest are cloned into pBIND and pACT Vectors to generate fusion proteins with the DNA-binding

domain of GAL4 and the activation domain of VP16, respectively. The pG5*luc* Vector contains five GAL4 binding sites upstream of a minimal TATA box, which in turn, is upstream of the firefly luciferase gene (*luc*+). The pGAL4 and pVP16 fusion constructs are transfected along with pG5*luc* Vector into mammalian cells. Two to three days after transfection the cells are lysed, and the amount of *Renilla* luciferase and firefly luciferase can be quantitated using the Dual-Luciferase® Reporter Assay System (Promega Cat.# E1910). Interaction between the two test proteins, as GAL4 and VP16 fusion constructs, results in an increase in firefly luciferase expression over the negative controls. A modified luciferase of the invention, e.g., one which is deleted at a site or region which is tolerant to modification (a N-terminal fragment), is fused to a DNA binding domain while the remainder of the luciferase (the C-terminal fragment) is fused to a transcriptional activator domain.

The invention also provides methods of screening for agents ("test" agents) capable of modulating the amount of a molecule of interest such as a cyclic nucleotide. "Modulation" refers to an alteration of a property; such enhancement or inhibition of a biological or chemical activity, where the alteration may be contingent on the occurrence of a specific event, such as activation of a signal transduction pathway, and/or may be manifest only in particular cell types. A "modulator" refers to an agent (naturally occurring or non-naturally occurring), such as, for example, a biological macromolecule (e.g., nucleic acid, protein, non-peptide, or organic molecule), small molecules, an extract made from biological materials such as bacteria, plants, fungi, or animal (particularly mammalian) cells or tissues, or any other agent. Modulators are evaluated for potential activity as inhibitors or activators (directly or indirectly) of a biological process or processes (e.g., agonist, partial antagonist, partial agonist, or antagonist) by inclusion in the screening assays described herein. The activities (or activity) of a modulator may be known, unknown or partially known. Such modulators can be screened using the methods of the invention. The term "test agent" refers to an agent to be tested by one or more screening method(s) of the invention as a putative modulator. Usually, various predetermined concentrations are used for screening such as 0.01 μM, 0.1

μM, 1.0 μM, and 10.0 μM. Controls can include the measurement of a signal in the absence of the test agent, comparison to an agent known to modulate the target, or comparison to a sample (e. a cell, tissue or organism) before, during and/or after contacting with the test agent.

In one embodiment, the method includes screening for agents that modulate protease activity. For example, in one embodiment, a method of identifying an agent capable of modulating apoptosis is provided. Caspase family proteases have been associated with apoptosis. Thus, the method includes contacting a sample suspected of containing a caspase-family protease with an agent suspected of modulating the caspase activity, and a modified luciferase having a cleavage site cleavable by the caspase. The activity of the modified luciferase is detected in the sample before and after contacting with the test agent. An increase in activity after contacting with the agent is indicative of an agent that inhibits apoptosis and a decrease is indicative of an agent that activates apoptosis.

Accordingly, the invention provides a screening system useful for identifying agents which modulate the cleavage of recognition sequence present in a modified luciferase protein of the invention and detecting its activity. This allows one to rapidly screen for protease activity modulators. Utilization of the screening system described herein provides a sensitive and rapid means to identify agents which modulate (e.g., inhibit or activate) a protease, for example, a caspase family protease.

A modified luciferase protein of the invention is thus useful as a substrate to study agents or conditions that modulate an interaction between an insertion in the modified luciferase protein and a molecule of interest. In particular, the invention contemplates modified luciferase proteins in which the insertion includes an amino acid sequence that is a cleavage site for an enzyme of interest. Thus, when the molecule of interest is a protease, the insertion comprises a peptide containing a cleavage recognition sequence for the protease. A cleavage recognition sequence for a protease is a specific amino acid sequence recognized by the protease during proteolytic cleavage. Accordingly, the invention provides methods to determine the amount of a protease in a sample by contacting the sample with a modified

luciferase polypeptide of the invention and measuring changes in luciferase activity. The modified luciferase protein of the invention can be used for, among other things, monitoring the activity of a protease inside a cell that expresses the modified luciferase.

In one embodiment, a modified luciferase of the invention is thus useful as a substrate to study agents or conditions that modulate an interaction between a cyclic nucleotide binding site in the modified luciferase and a molecule of interest such as a cyclic nucleotide, agents or conditions that modulate the presence or amount of a cyclic nucleotide, or agents or conditions that modulate molecules such as receptors that are associated with intracellular cyclic nucleotide concentrations. In particular, the invention contemplates modified luciferase proteins in which the insertion includes a cAMP or cGMP binding site. Thus, when the molecule of interest is cAMP or cGMP, the invention provides a method to determine the presence or the amount of cAMP or cGMP in a sample by contacting the sample with a modified luciferase polypeptide of the invention and measuring changes in luciferase activity. The modified luciferase protein of the invention can be used for, among other things, monitoring the amount or presence of cAMP or cGMP or molecules that alter the amount or presence of cAMP or cGMP inside a cell that has the modified luciferase.

The assays of the invention can be used to screen drugs to identify compounds that alter the amount, for example, of cyclic nucleotide or alter the binding of a cyclic nucleotide to a cyclic nucleotide binding site. In one embodiment, the assay is performed on a sample *in vitro* containing cAMP. A sample containing a known amount of cAMP is mixed with a modified luciferase of the invention and with a test agent. The amount of the luciferase activity in the sample is then determined. Then the amount of activity per mole of cAMP in the presence of the test agent may be compared with the activity per mole of cAMP in the absence of the test agent. A difference indicates that the test agent alters the amount of cAMP or binding of cAMP to the cAMP binding site.

In one embodiment, cells are conditioned or contacted with an agent suspected of directly or indirectly modulating, for instance, cAMP amount or

binding. The cells or cells in culture are lysed and cAMP amount measured. For example, a lysed cell sample containing a known or unknown amount of cAMP is mixed with a modified luciferase of the invention. The amount of cAMP in the sample is then determined as above by determining the degree of modified luciferase activity in a control or non-treated sample and the treated lysed cellular sample. The activity or inhibition can be calculated based on a per microgram or milligram protein in the sample. Typically, the difference is calibrated against standard measurements to yield an absolute amount of cAMP.

The materials and composition for use in the assay of the invention are ideally suited for the preparation of a kit. Such a kit may comprise a carrier means containing one or more container means such as vials, tubes, and the like, each of the container means comprising one of the separate elements to be used in the method. One of the containers comprises a modified luciferase or polynucleotide (e.g., in the form of a vector) of the invention. A second container may contain a substrate for the modified luciferase.

The invention will be further described by the following non-limiting examples.

Example I

Sites Tolerant to Modification in Click Beetle and Firefly Luciferases

Positions in click beetle and firefly luciferases that are tolerant to modification and certain modified luciferases are disclosed in U.S. published application 20050153310 and PCT/US2004/032705, the disclosures of which are incorporated by reference herein (see also Figure 1 and Table 1.)

Table 1

Inserted after Amino Acid in Firefly Luciferases	% Activity
7	10
121	5-10
233	50-75
267	2
294	3
303	5-10
361	3-5
540	15
541	75

Example II

Circular Permuted Firefly Luciferase Fusion to cAMP Binding Site

cAMP is one of the most important second messengers for cellular signal transduction. cAMP assays are extremely important for G-protein coupled receptor (GPCR) drug discovery. To identify biosensors for cAMP, cAMP binding sites were fused to circularly permuted firefly luciferases (CPM-FF Luc) (Figures 5A-B) (pBFB8, pBFB9, pBFB10, pBFB11, pBFB22, pBFB40, pBFB41, pBFB42). One CPM-FF Luc cAMP binding site fusion employed the cAMP binding site from human Epac1 (Exchange protein directly activated by cAMP) (Bos, 2003). Previous studies showed that a single chain fragment from human Epac1 (residues 157 to 316) binds cAMP (Nikolaev, J. Biol. Chem., 279, 37215 (2004)). A second CPM-FF Luc /cAMP binding site fusion employed the B domain from the human PKA regulatory subunit type IIB (CPM-FF Luc /RIIβB).

Materials and Methods

A DNA fragment encoding residues 157-316 of human Epac1 was synthesized, which included some silent nucleotide changes to potentially increase the expression in *E. coli* (Figure 5C). Two primers were used to generate a PCR fragment of EPAC1 with *Xho*I and *Nco*I sites at the 5' and 3' ends, respectively:

5' primer: atgcctcgagGAAGAAGAACTTGCTGAAGCTG (SEQ ID NO:22)

3' primer: atgccatggAACTCCATGTTCTTCTAAACGC (SEQ ID NO:23)

The resulting PCR fragment was digested and cloned into *Xho*I and *Nco*I sites of a circularly permuted beetle luciferase construct. The resulting plasmids expressed a

modified firefly (pSPLuc+, Promega Corporation) luciferase with EPAC1 inserted between the original N- and C- termini. The correct size of the fusion protein was verified by TnT cell-free expression and SDS-PAGE (Figure 6). This construct was identified as FF105.

DNA encoding RIIβB was inserted into a novel expression vector encoding CPM-FF Luc /RIIβB fusions [Luc2.0 (234-544)-linker X-human RIIβ (residues 266-414)-linker Y-Luc (4-233)]. By using unique combinations of restriction enzymes, various constructs were generated with RIIβB fused to CPM-FF Luc with a variety of X/Y peptide linker lengths.

Synthesis of a CPM-FF Luc expression plasmid for subsequent insertion of RIIβB

A synthetic 1816 bp fragment encoding CPM-FF Luc (DNA 2.0; SEQ ID NO:16, see Figure 20) was digested with *Hind*III/*Xba*I and ligated to the 3265 bp *Hind*III/*Xba*I fragment of pGL4.74 (Promega Corp.). The resultant plasmid encodes a circularly permuted mutant of synthetic luciferase (Luc2.0; Promega Corp.) with amino acids 544 and 4 of firefly luciferase connected by a 42 amino acid Gly/Ser rich peptide [Luc2.0 (234-544)-42 aa Gly/Ser rich peptide-Luc2.0 (4-233)] (pBFB8). Figures 5A depicts this parent CPM-FF Luc expression plasmid (pBFB8) and the unique restriction sites used to create various linker lengths and to insert the cAMP domain. This fusion protein can be expressed *in vitro* or *in vivo* using T7 or HSV-TK promoters, respectively. In addition, *Sgf*I and *Pme*I restriction enzyme sites were included at the 5' and 3' ends to facilitate subsequent transfer of this open reading frame to additional plasmids (Flexi vector system; Promega Corp.).

Synthesis of plasmids encoding CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues

The plasmid DNA construct described above was digested with unique restriction enzymes present in the multiple cloning site (MCS) linking the DNA fragments encoding Luc2.0 (233-544) and Luc2.0 (4-233) to allow synthesis of CPM-FF Luc/RIIβB expression constructs with X/Y linker lengths of (X=4,Y=4), (X=10,Y=10), and (X=20,Y=20) amino acid residues. Figure 5B depicts the linkers lengths flanking the RIIβB domain to create pBFB9 (X=4, Y=4), pBFB10 (X=10, Y=10) and pBFB11 (X=20, Y=20).

To synthesize the construct with (X=4,Y=4) linker lengths, primers 5'-AAA AAA GTC GAC CGG AAT GTA TGA AAG CTT TAT TGA GTC ACT GCC-3' (SEQ ID NO:25; BFB51) and 5'-AAA AAA GAG CTC CCA ACA ATA TCC ATG TTC GTT CCA AAC-3' (SEQ ID NO:26; BFB20) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *SalI*/*Sac*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI*/*Sac*I. This new construct was identified as pBFB9.

To synthesize the construct with (X=10,Y=10) linker lengths, primers 5'-AAA AAA TCC GGA ATG TAT GAA AGC TTT ATT GAG TCA CTG CC-3' (SEQ ID NO:211; BFB21) and 5'-AAA AAA AGG CCT ACA ATA TCC ATG TTC GTT CCA AAC-3' (SEQ ID NO:27; BFB22) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *Bsp*EI/*Stu*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Bsp*EI/*Zra*I. This new construct was identified as pBFB10.

To synthesize the construct with (X=20,Y=20) linker lengths, primers 5'-AAA AAA CCC GGG ATG TAT GAA AGC TTT ATT GAG TCA CTG CC-3' (SEQ ID NO:28; BFB23) and 5'- AAA AAA TCC GGA CCC AAC AAT ATC CAT GTT CGT TCC AAA C-3' (SEQ ID NO:29; BFB24) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant

product was digested with *BspEI*/*Sma*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Age*I/*Nru*I. This new construct was identified as pBFB11.

Expression of CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4), (X=10,Y=10), and (X=20,Y=20) amino acid residues.

The synthesis of fusion proteins of the predicted size was confirmed for the CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues using the TNT® T7 Coupled Wheat Germ Extract System (Promega Corp.) together with the FluoroTect GreenLys *in vitro* Translation Labeling System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

400 ng plasmid DNA
10 µL TnT Wheat Germ Extract
0.8 µL TNT reaction buffer
0.4 µL T7 polymerase
0.4 µL amino acid mixture
0.4 µL rRNasin
0.4 µL FluoroTect GreenLys label
dH_2O to 20 µL total volume

Following incubation at 30 °C for 1.5 hours, 5 µL of TNT reaction was resolved via SDS-PAGE following the manufacturer's protocol (NuPAGE Novex 4-12% bis-tris gel, Invitrogen Corp.). Translated proteins were subsequently visualized via fluorimager (Typhoon Variable Mode Imager, Amersham Biosciences). Densitometry analysis (ImageQuant, GE Healthcare) indicated that the CPM-FF Luc/RIIβB fusion proteins with variable X/Y linker lengths were expressed similarly to the CPM-FF Luc fusion proteins having the 42 amino acid Gly/Ser rich peptide (pBFB8) and Epac1 (FF105).

Functional characterization of CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues

Luciferase activity in the presence and absence of 100 μM cAMP was measured for the CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

400 ng plasmid DNA
10 μL Rabbit Retic Extract
0.8 μL TNT reaction buffer
0.4 μL T7 polymerase
0.4 μL amino acid mixture
0.4 μL rRNasin
dH_2O to 20 μL total volume

Following incubation at 30°C for 1.5 hours, the respective fusion proteins were incubated in the presence or absence of 100 μM cAMP by combining 9 μL of TNT® reaction with 1 μL of 1 mM cAMP stock or dH_2O. Following incubation for 10 minutes at room temperature, 1 μL of sample was added to 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.) solution +/- 100 μM cAMP (90 μL LAR + 10 μL 1 mM cAMP stock or dH_2O). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo).

Dose response experiment using CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues

The cAMP dose response of CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues was measured following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following

components were assembled according to the manufacturer's recommended protocol:

2000 ng plasmid DNA
50 µL Rabbit Retic Extract
4 µL TNT reaction buffer
2 µL T7 polymerase
2 µL amino acid mixture
2 µL rRNasin
dH_2O to 100 µL total volume

Following incubation at 30°C for 2 hours, the respective fusion proteins were incubated with varying concentrations of cAMP by combining 9 µL of TnT® reaction with 1 µL of cAMP stock solution (final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, 25, or 100 µM cAMP). Following incubation at room temperature for ≥ 25 minutes, 1 µL of sample was added to 100 µL of Luciferase Assay Reagent (LAR; Promega Corp.) with the respective concentration of cAMP (90 µL LAR + 10 µL cAMP stock solution). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo).

Selectivity of the CPM-FF Luc/RIIβB fusion protein with X/Y linker length of (X=10,Y=10; pBFB10) amino acid residues

The selectivity of the CPM-FF Luc/RIIβB fusion protein with X/Y linker length of (X=10,Y=10; pBFB10) amino acid residues for cAMP activation relative to other cyclic nucleotides was measured following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

6000 ng plasmid DNA
150 µL Rabbit Retic Extract
12 µL TNT reaction buffer
6 µL T7 polymerase
6 µL amino acid mixture

6 μL rRNasin

dH_2O to 300 μL total volume

Following incubation at 30°C for 2.3 hours, the fusion protein was incubated with varying concentrations of cAMP, cGMP, or N6-benzoyl cAMP by combining 9 μL of TNT® reaction with 1 μL of cyclic nucleotide stock solution (final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, 25, or 100 μM cAMP). Following incubation at room temperature for ≥ 29 minutes, 1 μL of sample was added to 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.) with the respective concentration of cyclic nucleotide (90 μL LAR + 10 μL cyclic nucleotide stock solution). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo).

Results

Protein kinase A regulatory subunit type IIβ (PRKAR2B), has two cAMP binding sites, A and B. The cAMP binding site from the B domain (RIIβB) was used to prepare a circularly permutated luciferase (CPM-FF Luc) with RIIβB (CPM-FF Luc/RIIβB). CPM-FF Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB9), (X=10,Y=10; pBFB10), and (X=20,Y=20; pBFB11) amino acid residues each showed an induction of luciferase activity in the presence of 100 μM cAMP of 23-, 58-, and 39-fold, respectively (Figure 7). As expected, no cAMP regulation was seen for the CPM-FF Luc fusion protein having the 42 amino acid Gly/Ser rich peptide (pBFB8). In addition to RIIβB, the cAMP binding site from Epac1 was used to generate a cAMP sensor (FF105). However, the fold induction in luciferase activity was less than the RIIβB based sensor (Figure 7).

Each CPM-FF Luc/RIIβB fusion protein showed a unique dose response with variable values for the effective concentration for 50% maximal fold induction (Figure 8A). The CPM-FF Luc/RIIβB fusion protein with X/Y linker length of (X=10,Y=10; pBFB10) amino acid residues showed enhanced selectivity for binding to cAMP relative to other cyclic nucleotides (Figure 8B).

Example III

Circularly Permuted *Renilla* Luciferases with cAMP Binding Sites

Materials and Methods

Four humanized *Renilla* luciferase DNA fragments were amplified from either pF5RK or phRL-null vectors (Promega Corp.) and cloned into the CPM-FF Luc fusion protein construct = [Luc2.0 (234-544)-42 amino acid Gly/Ser rich peptide-Luc2.0 (4-233)] (pBFB8; Figure 5A) to generate a circularly permuted *Renilla* luciferase open reading frame (CPM-hRL) split either between positions Ser91/Tyr92 or Ile223/Pro224 (Figure 5D). The sequencing primers used to generate the four humanized *Renilla* luciferase DNA fragments were: 5'-ATGGGCGATCGCCatgtatcgcctcctggatcactacaag–3' (hRL92 *Sgf*I; FF273; SEQ ID NO:110); 5'–ATGGGCGATCGCCatgcctctcgttaagggaggcaagc–3' (hRL224 *Sgf*I; FF277; SEQ ID NO:111); 5'–gcatCTCGAGccctgctcgttcttcagcacgcgc–3' (hRL311 / *Xho*I; FF294; SEQ ID NO:112); 5'–atgcGAGCTCaggagcttccaaggtgtacgacccg–3' (hRL2 *Sac*I; FF295; SEQ ID NO:113); 5'-TTGTGTTTAAACtgagccattcccgctcttgccg–3 '(hRL91/*Pme*I; FF276; SEQ ID NO:114); and 5'–TTGTGTTTAAACgatctcgcgaggccaggagagg–3' (hRL223 *Pme*I; FF278; SEQ ID NO:115). Primer pairs FF273/FF294 and FF277/FF294 were used to amplify the C terminal fragment of the humanized *Renilla* luciferase DNA (hRL 92-311 and hRL 224-311, respectively). The resultant products were digested with *SgfI/XhoI* restriction enzymes and ligated into the parent CPM-FF Luc fusion protein construct = [Luc2.0 (234-544)-42 amino acid Gly/Ser rich peptide-Luc2.0 (4-233)], pBFB8, digested with *SgfI/XhoI*. Primer pairs FF276 / FF295 and FF278 / FF295 were used to amplify the N terminal fragments of the humanized *Renilla* luciferase DNA (hRL 2-91 and hRL 2-223, respectively). The resultant products were digested with *SacI/PmeI* restriction enzymes and ligated into the intermediate CPM-FF Luc / hRL plasmid encoding [hRL (92-311 or 224-311)-42 amino acid Gly/Ser rich peptide-Luc2.0 (4-233)] digested with *SacI/PmeI*. This resulted in the generation of CPM-hRL expression vectors where the circularly permuted hRL luciferase fragments are fused by a 42 amino acid Gly/Ser rich peptide (identical to the Gly/Ser rich peptide of Figure 5A, 201325.15.A1 (CPM91); 201325.15.B6 (CPM223)). The sequence encoding human RIIβB amino acids 266-414 (Genbank

ID BC075800) was cloned into subsets of the unique restriction enzyme sites that encode amino acids present in the Gly/Ser rich peptide as previously described for the CPM-FF Luc/RIIβB cAMP sensors (Figure 5D). The resulting constructs encode CPM-hRL/RIIβB fusions with either X=4, Y=20 (201325.44.H6 (CPM91); 201325.33.C9 (CPM223)), X=10, Y=4 (201325.50.D12 (CPM91); 201325.54.E2 (CPM223)) or X=10, Y=20 (201325.58.E11 (CPM91); 201325.54.E12 (CPM223)) Gly/Ser rich linkers fused to the N- and C-termini of RIIβB, respectively (Figure 5D). In addition, the full length hRL open reading frame was cloned into the *SgfI/PmeI* sites of the CPM-FF Luc expression plasmid encoding Luc2.0 (234-544)-42 amino acid Gly/Ser rich peptide-Luc2.0 (4-233) (201325.50.A7, Figure 5A).

One μg purified plasmid DNA per 50 μl Wheat Germ TnT® (Promega cat# L4140) reaction was used to express the protein products. Wheat Germ TnT® reactions were carried out at 30°C for 1 hour in the presence of FluoroTect™ $Green_{Lys}$ tRNA (Promega cat#L5001). The CPM-hRL constructs were expressed together with the following controls: CPM-FF Luc/RIIβB with X=10, Y=4 (pBFB41), full length *Renilla* luciferase (201325.50.A7), and a "no DNA" (negative control). Fifteen μl of each lysate was mixed with either 1.5 μl 1 mM cAMP (Promega cat# V642A, 100 μM final concentration) or water (Promega cat#P119C) and incubated for 10 minutes at room temperature. Seventy five μl of 1X *Renilla* Luciferase Assay Lysis Buffer (5X *Renilla* Luciferase Assay Lysis Buffer (Promega cat#E291A) plus water (Promega cat#P119C) was added to the *Renilla* luciferase reaction and "no DNA" samples, mixed, and 20 μl of each mixture was added in triplicate to a 96 well white flat bottom plate. Two μl of the CPM-FF Luc/RIIβB with X=10, Y=4 linkers sample (pBFB41) was added in triplicate to a 96 well white flat bottom plate. One hundred μl of *Renilla* Luciferase Assay Buffer plus 1X *Renilla* Luciferase Assay Substrate (Promega Corp.; cat# E2820) was added to each of the *Renilla* luciferase and "no DNA" wells. One hundred μl of Luciferase Assay Buffer plus Luciferase Assay Substrate (Promega Corp.; cat# E1500) was added to each well containing the CPM-FF Luc/RIIβB with X=10, Y=4 linkers (pBFB41). Luminescence was measured using a Veritas Luminometer. Prior to cAMP

incubation, 10 µl of each lysate was size fractionated on an SDS-PAGE gel. Fluorescent protein products were visualized on a Typhoon imager.

Results

Wheat Germ TnT® reactions resulted in approximately equal amounts of each construct protein. There was no visible protein product from the "no DNA" sample. The full length *Renilla* luciferase construct (201325.50.A7) resulted in about 100-fold more luminescence than the CPM-hRL91-42aa construct (201325.15.A1) and about 100,000-fold more luminescence than the CPM-hRL223-42aa construct (201325.15.B6). The RIIβB constructs CPM-hRL91-4aa-RIIβB-20aa (201325.44.H6) and CPM-hRL91-10aa-RIIβB-20aa (201325.58.E11) gave more luminescence when incubated with 100 µM cAMP than water (115- to 146-fold and 100-fold, respectively). The RIIβB constructs CPM-hRL223-4aa-RIIβB-20aa (201325.33.C9), CPM-hRL223-10aa-RIIβB-4aa (201325.54.E2) and CPM-hRL223-10aa-RIIβB-20aa (201325.54.E12) gave 1.7- to 2.1-fold more luminescence when incubated with 100 µM cAMP than water. The full length *Renilla* luciferase (201325.50.A7), CPM-hRL91-42aa (201325.15.A1), and CPM-hRL223-42aa constructs (201325.15.B6) did not change with cAMP incubation more than 1.3-fold as compared to water. The CPM-FF Luc/RIIβB sensor with X=10, Y=4 linkers construct (pBFB41) gave 85-90-fold more luminescence in the presence of cAMP. The "no DNA" reaction had low luminescence (1,000,000-fold less than full length *Renilla* luciferase) and did not change with cAMP incubation (see Figure 10).

Example IV

In vitro Detection of cAMP with CPM-FF Luc/RIIβB cAMP Biosensors

Materials and Methods

To demonstrate the efficacy of cAMP measurement in cell lysates and in the presence of cell lysis detergents, CPM-FF Luc/RIIβB fusion protein with X/Y linker lengths of (X=10,Y=10; pBFB10) was expressed using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components

were assembled according to the manufacturers recommended protocol and incubated at 30 °C for 1.5 hours:

1000 ng plasmid DNA
25 µL Rabbit Retic Extract
2 µL TNT reaction buffer
1 µL T7 polymerase
1 µL amino acid mixture
1 µL rRNasin
dH_2O to 50 µL total volume

To simulate the experimental conditions of cAMP measurement following detergent mediated lysis of cells, the following components were mixed at room temperature with final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, and 25µM cAMP:

0.5 µL TNT® expressed cAMP sensor
19.5 µL Wheat Germ Extract (Promega Corp.; cat# L4140, part# L411A)
5 µL cAMP stock solution
25 µL Bright-Glo assay reagent (Promega Corp., cat# E2610)

The assembled reaction was immediately mixed and the luciferase activity was measured continuously using a Turner 20/20N luminometer at 1 measurement per second (Turner Biosystems).

In some experiments, to enhance signal stability and luminescence, the reaction mixture includes 4 mM luciferin (Promega Bioscience), 2 mM Coenzyme A (Sigma), 10 mM ATP (Pharmacia), 10 mM DTT (Promega), 16 mM magnesium sulfate, 150 mM HEPES, pH 8.0 (Fisher), 1% Tergitol N101 (Sigma), 1% Mazu DF101, and 1 mM CDTA (Sigma). *In vitro* translated CPM-FF Luc/RIIβB cAMP biosensors were synthesized using TnT® Coupled Rabbit Reticulocyte System (Promega) using 1 µg of plasmid DNA for 50 µl total reaction volume and added to the reaction mixture immediately prior to assaying for cAMP (addition of 1 µl of translated product per 100 µl of assay reagent). 100 µl of assay reagent plus sensor

was then added to either 100 µl of cell culture or 100 µl of cAMP diluted in complete media (DMEM/F12 + 10% FBS).

Cell Culture

For the *in vitro* analyses, HEK-293 cells were plated in a 96 well plate and grown to 50-90% confluency in 100 µl DMEM/F12 (Invitrogen) with 10% FBS (Hyclone) at 37°C with 5% CO_2. Cells were stimulated with 0.02 to 250 µM forskolin (Sigma) where the forskolin was diluted by 2-fold dilutions in the complete media.

Standard curve with cAMP

1 mM cAMP (Promega) was diluted into complete DMEM/F12 media with 10% FBS using a concentration range of 0.005 to 50 µM cAMP, where cAMP is serially diluted by 2-fold dilutions. 100 µl of cAMP was mixed with 100 µl of Homogeneous cAMP Luminescent Assay Reagent.

Results

CPM-FF Luc/RIIβB cAMP sensors functioned in a variety of lysis buffers and with a variety of luciferase reagents. Furthermore, the CPM-FF Luc/RIIβB cAMP sensors were employed in homogenous assay formats for detection of cAMP *in vitro* (Figures 9, 11A and 11B). For example, using wheat germ extract, dose dependent values of luciferase activity developed within approximately three minutes with a dynamic range of cAMP detection between 0.025 to 25 μM cAMP (Figure 9). In an additional example using an optimized reagent formulation, *in vitro* detection of cAMP showed a signal to background ratio of 20 and an EC_{50} of 1.28 μM for the CPM-FF Luc/RIIβB cAMP sensor with X/Y linker lengths of X=10, Y=4 (pBFB41) (Figure 11A). Similarly, using the same optimized reagent formulation, *in vitro* detection of cAMP showed a signal to background ratio of 11 and an EC_{50} of 0.64 μM for the CPM-FF Luc/RIIβB cAMP sensor with X/Y linker lengths of (X=10, Y=10) (pBFB10) (Figure 11A). The present cAMP assay has the following advantages: a bioluminescence readout, which reduces compound interference; a homogenous one-step format; and the specificity that requires both binding and the capability of inducing a conformational change.

Example V

Intracellular Detection of Changes in cAMP Concentration using CPM-FF Luc/RIIβB cAMP Biosensors

Cell Culture

Cells were cultured in 60 ml in DMEM/F12 with HEPES buffer (Invitrogen) with 10% FBS at 37°C with 5% CO_2.

Plasmids

The ORF encoding the CPM-FF Luc/RIIβB based cAMP biosensor with X/Y linker lengths of (X=10, Y=0) was transferred to Flexi vector pF4K (Flexi vector system; Promega Corp.). The resultant plasmid construct (pBFB141) utilizes an upstream CMV promoter for expression of the associated cAMP biosensor in mammalian cells.

Transfections

Cells were transfected with *Trans*It®-LT1 Reagent (MIRUS) using 0.3 µl *Trans*It®-LT1 reagent and 0.15 µg DNA per well of a 96 well plate. Cells were allowed to grow overnight and were assayed the next day.

Modulation of biosensor

Approximately 1 day after transfection, cells were removed from the incubator and equilibrated to room temperature. A 5 µl aliquot of 100 mM Luciferin EF was added to a total of 90 µl of cell culture plus transfection reagent to give a final concentration of approximately 5 mM luciferin. Cells were then incubated at room temperature for at least 90 minutes. After 90 minutes at room temperature, baseline measurements of luminescence were measured using a 96 well Veritas Luminometer (Turner Biosystems; integration time of 0.5 seconds per well). Cells were then induced with 10 µM isopreterenol (CalBiochem), 50 mM forskolin (Sigma) or not induced (0.1% DMSO, Sigma) and luminescence was measured continuously for about 30 minutes. After 30 minutes, 10 mM propranolol (Sigma) was added to cells with isopreterenol and 0.1% DMSO was added to all other samples. Luminescence was then measured continuously for the next 30 minutes. A final addition of 50 µM forskolin was added to the isopreterenol/propranolol sample and 0.1% DMSO was added to all other samples. Luminescence was then measured continuously for the next half hour. Samples were measured in sets of 12 replicates. 10x stocks of isopreterenol, propranolol, forskolin and DMSO were made in 1x PBS (Invitrogen).

Results

To measure changes in the intracellular concentration of cAMP, HEK 293 cells were transiently transfected with the CPM-FF Luc/RIIβB (X=10, Y=0, pBFB141) construct followed by treatment with compounds known to increase the intracellular cAMP concentration through GPCR activation (isopreterenol, β-adrenergic receptor agonist), decrease intracellular cAMP concentration through GPCR inhibition (propranolol, β-adrenergic receptor antagonist), or increase intracellular cAMP concentration through activation of adenylate cyclase (forskolin). Both isopreterenol and forskolin treatment alone increased light output

from transfected cells approximately 2-fold, reflecting an increase in intracellular cAMP concentration (Figure 11C). In addition, the temporal response of changes in cAMP concentration was followed by treating transfected cells with isoperterenol, propranolol, followed by forskolin (Figure 11C). Wild type luciferase and the CPM-FF Luc/RIIβB fusion protein expressing the 42 amino acid Gly/Ser rich peptide (pBFB8) were also tested and showed no specific response to addition of known modulators of intracellular cAMP concentration.

Example VI

Light Output and Fold Induction Vary as a Function of X/Y Peptide Linker Lengths for CPM-FF Luc/RIIβB Based cAMP Sensors

A. Synthesis of plasmids encoding CPM-FF Luc/RIIβB based cAMP sensors with variable X/Y peptide linker lengths

To generate a set of CPM-FF Luc/RIIβB based cAMP sensors with variable X/Y peptide linker lengths of [2x (x = 0-5), 2y (y = 0-5)], plasmids encoding sensors of (X=0,Y=0, pBFB89), (X=2,Y=2, pBFB96), (X=6,Y=6, pBFB108), and (X=8,Y=8, pBFB115) were synthesized using splice overlap extension PCR (SOE PCR). Once acquired, standard molecular cloning techniques were used to exchange DNA fragments between plasmids encoding CPM-FF Luc/RIIβB based cAMP sensors with (X=0,Y=0), (X=2,Y=2), (X=4,Y=4), (X=6,Y=6), (X=8,Y=8), and (X=10,Y=10) peptide linkers to generate all remaining clones in this set. In addition, SOE PCR was used to synthesize clones in [10 + 2n (n = 0-5), 0] and [10, -2n (n = 1-7)] sets (Table 2).

Table 2

pBFB89	X=0, Y=0	**RIIβB** (SEQ ID NO:124)
pBFB90	X=0, Y=2	**RIIβB-SG** (SEQ ID NO:125)
pBFB91	X=0, Y=4	**RIIβB-GSSG** (SEQ ID NO:126)
pBFB92	X=0, Y=6	**RIIβB-SGGSSG** (SEQ ID NO:127)
pBFB93	X=0, Y=8	**RIIβB-GGSGGSSG** (SEQ ID NO:128)
pBFB94	X=0, Y=10	**RIIβB-GSGGSGGSSG** (SEQ ID NO:129)
pBFB95	X=2, Y=0	**GS-RIIβB** (SEQ ID NO:130)
pBFB96	X=2, Y=2	**GS-RIIβB-SG** (SEQ ID NO:131)

pBFB97	X=2, Y=4	**GS-RIIβB-GSSG** (SEQ ID NO:132)
pBFB98	X=2, Y=6	**GS-RIIβB-SGGSSG** (SEQ ID NO:133)
pBFB99	X=2, Y=8	**GS-RIIβB-GGSGGSSG** (SEQ ID NO:134)
pBFB100	X=2, Y=10	**GS-RIIβB-G SGGSGGSSG** (SEQ ID NO:135)
pBFB101	X=4, Y=0	**GSTG-RIIβB** (SEQ ID NO:136)
pBFB102	X=4, Y=2	**GSTG-RIIβB-SG** (SEQ ID NO:137)
pBFB9	X=4, Y=4	**GSTG-RIIβB-GSSG** (SEQ ID NO:138)
pBFB103	X=4, Y=6	**GSTG-RIIβB-SGGSSG** (SEQ ID NO:139)
pBFB104	X=4, Y=8	**GSTG-RIIβB-GGSGGSSG** (SEQ ID NO:140)
pBFB39	X=4, Y=10	**GSTG-RIIβB-GSGGSGGSSG** (SEQ ID NO:141)
pBFB105	X=6, Y=0	**GSTGGS-RIIβB** (SEQ ID NO:142)
pBFB106	X=6, Y=2	**GSTGGS-RIIβB-SG** (SEQ ID NO:143)
pBFB107	X=6, Y=4	**GSTGGS-RIIβB-GSSG** (SEQ ID NO:144)
pBFB108	X=6, Y=6	**GSTGGS-RIIβB-SGGSSG** (SEQ ID NO:145)
pBFB109	X=6, Y=8	**GSTGGS-RIIβB-GGSGGSSG** (SEQ ID NO:146)
pBFB110	X=6, Y=10	**GSTGGS-RIIβB-GSGGSGGSSG** (SEQ ID NO:147)
pBFB111	X=8, Y=0	**GSTGGSGG-RIIβB** (SEQ ID NO:148)
pBFB112	X=8, Y=2	**GSTGGSGG-RIIβB-SG** (SEQ ID NO:149)
pBFB113	X=8, Y=4	**GSTGGSGG-RIIβB-GSSG** (SEQ ID NO:150)
pBFB114	X=8, Y=6	**GSTGGSGG-RIIβB-SGGSSG** (SEQ ID NO:151)
pBFB115	X=8, Y=8	**GSTGGSGG-RIIβB-GGSGGSSG** (SEQ ID NO:152)
pBFB116	X=8, Y=10	**GSTGGSGG-RIIβB-GSGGSGGSSG** (SEQ ID NO:153)
pBFB117	X=10, Y=0	**GSSGGSGGSG-RIIβB** (SEQ ID NO:154)
pBFB118	X=10, Y=2	**GSSGGSGGSG-RIIβB-SG** (SEQ ID NO:155)
pBFB41	X=10, Y=4	**GSSGGSGGSG-RIIβB-GSSG** (SEQ ID NO:156)
pBFB119	X=10, Y=6	**GSSGGSGGSG-RIIβB-SGGSSG** (SEQ ID NO:157)
pBFB120	X=10, Y=8	**GSSGGSGGSG-RIIβB-GGSGGSSG** (SEQ ID NO:158)
pBFB 10	X=10, Y=10	**GSSGGSGGSG-RIIβB-GSGGSGGSSG** (SEQ ID NO:159)
pBFB128	X=10, Y=-2	**GSSGGSGGSG-RIIβB (266-412)** (SEQ ID NO:160)
pBFB129	X=10, Y=-4	**GSSGGSGGSG-RIIβB (266-410)** (SEQ ID NO:161)
pBFB130	X=10, Y=-6	**GSSGGSGGSG-(266-408)** (SEQ ID NO:162)
pBFB131	X=10, Y=-8	**GSSGGSGGSG-RIIβB (266-406)** (SEQ ID NO:163)
pBFB132	X=10, Y=-10	**GSSGGSGGSG-RIIβB (266-404)** (SEQ ID NO:164)
pBFB133	X=10, Y=-12	**GSSGGSGGSG-RIIβB (266-402)** (SEQ ID NO:165)

pBFB134	X=10, Y=-14	**GSSGGSGGSG-RIIβB (266-400)** (SEQ ID NO:166)
pBFB135	X=12, Y=0	**GSSGGSGGSGGG-RIIβB** (SEQ ID NO:167)
pBFB136	X=14, Y=0	**GSSGGSGGSGGGSG-RIIβB** (SEQ ID NO:168)
pBFB137	X=16, Y=0	**GSSGGSGGSGGGSGGS-RIIβB** (SEQ ID NO:169)
pBFB138	X=18, Y=0	**GSSGGSGGSGGGSGGSGG-RIIβB** (SEQ ID NO:170)
pBFB139	X=20, Y=0	**GSSGGSGGSGGGSGGSGGSG-RIIβB** (SEQ ID NO:171)

(RIIβB corresponds to amino acids 266-414 of Genbank ID AAH75800)

i. Synthesis of a plasmid encoding a CPM-FF Luc/RIIβB based cAMP sensor lacking peptide linkers (X=0,Y=0; pBFB89)

To synthesize the construct lacking peptide linkers (X=0,Y=0), three separate primer pairs were used to amplify RIIβB DNA to generate three separate PCR products. Primer pair 5'- CCT CGA ACA CCG AGC GAC C-3' (SEQ ID NO:31) and 5'-GCA GTG ACT CAA TAA AGC TTT CAT ACA TCT TCT TGG CCT TAA TGA GAA TCT CG-3' (SEQ ID NO:18) were used to generate product #1; primer pair 5'-CGA GAT TCT CAT TAA GGC CAA GAA GAT GTA TGA AAG CTT TAT TGA GTC ACT GC-3' (SEQ ID NO:32) and 5'-GGC CCT TCT TAA TGT TTT TGG CTA CAA TAT CCA TGT TCG TTC CAA ACA G-3' (SEQ ID NO:33) were used to generate product 2; and primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA GCC AAA AAC ATT AAG AAG GGC C-3' (SEQ ID NO:34) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3 (SEQ ID NO:35) were used to generate product 3. SOE PCR of the three products yielded the full-length PCR product, which was subsequently digested with *SgfI/XbaI* restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *SgfI/XbaI*.

ii. Synthesis of a plasmid encoding a CPM-FF Luc/RIIβB based cAMP sensor with (X=2,Y=2; pBFB96) peptide linker lengths

To synthesize the construct with peptide linkers (X=2,Y=2), three separate primer pairs were used to amplify RIIβB to generate three separate PCR products. Primer pair 5'- CCT CGA ACA CCG AGC GAC C-3' (SEQ ID NO:36; BFB31) and 5'-CAA TAA AGC TTT CAT ACA TCG AGC CCT TCT TGG CCT TAA TGA GAA TCT CG-3' (SEQ ID NO:37; BFB120) were used to generate product 1;

primer pair 5'-CGA GAT TCT CAT TAA GGC CAA GAA GGG CTC GAT GTA TGA AAG CTT TAT TG-3' (SEQ ID NO:38; BFB119) and 5'-CTT CTT AAT GTT TTT GGC ACC GGA TAC AAT ATC CAT GTT CGT TCC AAA CAG-3' (SEQ ID NO:39; BFB122) were used to generate product 2; and primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA TCC GGT GCC AAA AAC ATT AAG AAG-3' (SEQ ID NO:40; BFB122) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:41; BFB34) were used to generate product 3. SOE PCR of the three products yielded the full-length PCR product, which was subsequently digested with *SgfI/XbaI* restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *SgfI/XbaI*.

iii. Synthesis of a plasmid encoding a CPM-FF Luc/RIIβB based cAMP sensor with (X=6,Y=6; pBFB108) peptide linker lengths

To synthesize the construct with peptide linkers (X=6,Y=6), primers 5'-AAA AAA AAA GTC GAC CGG AGG TTC AAT GTA TGA AAG CTT TAT TGA GTC ACT GC-3' (SEQ ID NO:42; BFB123) and 5'-AAA AAA GAG CTC CCT CCA GAT ACA ATA TCC ATG TTC GTT CCA AAC AG-3' (SEQ ID NO:43; BFB124) were used to PCR amplify RIIβB DNA. The resultant product was digested with *SalI/SacI* restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/SacI*.

iv. Synthesis of a plasmid encoding a CPM-FF Luc/RIIβB based cAMP sensor with (X=8,Y=8; pBFB115) peptide linker lengths

To synthesize the construct with peptide linkers (X=8,Y=8), primers 5'-AAA AAA GTC GAC CGG AGG TTC AGG CGG TAT GTA TGA AAG CTT TAT TGA GTC ACT GC-3' (SEQ ID NO:44; BFB125) and 5'-AAA AAA GAG CTC CCT CCA GAT CCA CCT ACA ATA TCC ATG TTC GTT CCA AAC AG-3' (SEQ ID NO:116; BFB126) were used to PCR amplify RIIβB DNA. The resultant product was digested with SalI/SacI restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/SacI*.

v. Synthesis of plasmids encoding the remaining CPM-FF Luc/RIIβB based cAMP sensors with peptide linker lengths in the set [2x (x = 0-5), 2y (y = 0-5)]

*Xho*I/*Xba*I or *Xmn*I/*Xba*I restriction enzyme digests were performed on plasmids encoding CPM-Luc/RIIβB based cAMP sensors with peptide linker lengths of (X=0,Y=0), (X=2,Y=2), (X=4,Y=4), (X=6,Y=6), (X=8,Y=8), and (X=10,Y=10). In each case, the restriction enzyme digest generates two fragments: a smaller fragment encoding a C-terminal portion of RIIβB, linker Y, and the Luc2.0 4-233 fragment; and a larger fragment containing all remaining elements of the original plasmid, including the sequences encoding Luc2.0 234-544, linker X, and an N-terminal portion of RIIβB. To generate all 36 clones in the [2x (x = 0-5), 2y (y = 0-5)] set, the smaller fragments were ligated to the larger fragments from the various restriction enzyme digests.

vi. Synthesis of plasmids encoding CPM-FF Luc/RIIβB based cAMP sensors with peptide linker lengths in the set [10 + 2n (n = 1-5), 0]

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with peptide linker length (X=12,Y=0; pBFB135), two separate primer pairs were used to amplify RIIβB to generate two separate PCR products. Primer pair 5'-AAA AAA TCC GGA GGA GGT ATG TAT GAA AGC TTT ATT GAG TCA CTG C-3' (SEQ ID NO:46 BFB142) and 5'-GGC CCT TCT TAA TGT TTT TGG CTA CAA TAT CCA TGT TCG TTC CAA ACA G-3' (SEQ ID NO:47; BFB118) were used to generate product #1; primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA GCC AAA AAC ATT AAG AAG GGC C-3' (SEQ ID NO:48; BFB117) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:49; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *Bsp*EI/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Bsp*EI/*Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with peptide linker length (X=14,Y=0; pBFB136), two separate primer pairs were used to amplify RIIβB to generate two separate PCR products. Primer pair 5'-AAA AAA TCC GGA GGA GGT TCT GGC ATG TAT GAA AGC TTT ATT GAG TCA CTG C-3' (SEQ ID NO:45; BFB143) and 5'-GGC CCT TCT TAA TGT TTT TGG CTA CAA TAT CCA TGT TCG TTC CAA ACA G-3' (SEQ ID NO:21;

BFB118) were used to generate product 1; primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA GCC AAA AAC ATT AAG AAG GGC C-3' (SEQ ID NO:24; BFB117) and 5'- GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:30; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *Bsp*EI/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Bsp*EI/*Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with peptide linker length (X=16,Y=0; pBFB137), two separate primer pairs were used to amplify RIIβB to generate two separate PCR products. Primer pair 5'-ATA AAT TCC GGA GGA GGT TCT GGC GGA TCA ATG TAT GAA AGC TTT ATT GAG TCA CTG C-3' (SEQ ID NO:50; BFB144) and 5'-GGC CCT TCT TAA TGT TTT TGG CTA CAA TAT CCA TGT TCG TTC CAA ACA G-3' (SEQ ID NO:51; BFB118) were used to generate product 1; primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA GCC AAA AAC ATT AAG AAG GGC C-3' (SEQ ID NO:52; BFB117) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:53; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *Bsp*EI/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Bsp*EI/*Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with peptide linker length (X=18,Y=0; pBFB138), two separate primer pairs were used to amplify RIIβB to generate two separate PCR products. Primer pair 5'-AAA AAT TCC GGA GGA GGT TCT GGC GGA TCA GGC GGT ATG TAT GAA AGC TTT ATT GAG TCA CTG C-3' (SEQ ID NO:54; BFB145) and 5'-GGC CCT TCT TAA TGT TTT TGG CTA CAA TAT CCA TGT TCG TTC CAA ACA G-3' (SEQ ID NO:55; BFB118) were used to generate product 1; primer pair 5'-CTG TTT GGA ACG AAC ATG GAT ATT GTA GCC AAA AAC ATT AAG AAG GGC C-3' (SEQ ID NO:56; BFB117) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:57; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was

subsequently digested with *Bsp*EI/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Bsp*EI/*Xba*I.

vii. Synthesis of plasmids encoding CPM-FF Luc/RIIβB based cAMP sensors with peptide linker lengths in the set [10, -2n (n = 1-7)]

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with RIIβB residues 266-412 (10, -2; pBFB128), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAA AAA GTC GAC CGG AGG TTC AGG CGG TTC-3' (SEQ ID NO:58; BFB127) and 5'-GGC CCT TCT TAA TGT TTT TGG CAT CCA TGT TCG TTC CAA ACA GG-3' (SEQ ID NO:59; BFB128) were used to generate product 1; primer pair 5'-CCT GTT TGG AAC GAA CAT GGA TGC CAA AAA CAT TAA GAA GGG CC-3' (SEQ ID NO:60; BFB129) and 5'-GTA TCT TAT CAT GTC TGC TCG AAG CG-3' (SEQ ID NO:61; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *Sal*I/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Xho*I/*Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with RIIβB residues 266-410 (10, -4; pBFB129), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:62; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCGTTCGTTCCAAACAGGGCAACTAAC-3' (SEQ ID NO:63; BFB130) were used to generate product #1; primer pair 5'-GTTAGTTGCCCTGTTTGGAACGAACGCCAAAAACATTAAGAAGGGCC-3' (SEQ ID NO:64; BFB131) and 5'-GTATCTTATCATGTCTGCTCGAAGCG-3' (SEQ ID NO:65; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *Sal*I/*Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *Xho*I/*Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with RIIβB residues 266-408 (10, -6; pBFB130), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:66; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCTCCAAACAGGGCAACTAACTGTTCTTC -3' (SEQ ID NO:67; BFB132) were used to generate product 1; primer pair 5'-GAAGAACAGTTAGTTGCCCTGTTTGGAGCCAAAAACATTAAGAAGGG CC -3' (SEQ ID NO:68; BFB133) and 5'-GTATCTTATCATGTCTGCTCGAAGCG-3' (SEQ ID NO:69; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *SalI/Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with RIIβB residues 266-406 (10, -8; pBFB131), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:70; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCCAGGGCAACTAACTGTTCTTCATAGG-3' (SEQ ID NO:71; BFB134) were used to generate product 1; primer pair 5'-CCTATGAAGAACAGTTAGTTGCCCTGGCCAAAAACATTAAGAAGGGC C -3' (SEQ ID NO:72; BFB135) and 5'-GTATCTTATCATGTCTGCTCGAAGCG-3' (SEQ ID NO:73; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *SalI/Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal

peptide linker, with RIIβB residues 266-404 (10, -10; pBFB132), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:74; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCAACTAACTGTTCTTCATAGGTAGCGATG -3' (SEQ ID NO:75; BFB136) were used to generate product 1; primer pair 5'-CATCGCTACCTATGAAGAACAGTTAGTTGCCAAAAACATTAAGAAGGGCC-3' (SEQ ID NO:76; BFB137) and 5'-GTATCTTATCATGTCTGCTCGAAGCG -3' (SEQ ID NO:77; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *SalI/Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with RIIβB residues 266-402 (10, -12; pBFB133), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:78; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCCTGTTCTTCATAGGTAGCGATGTTCC-3' (SEQ ID NO:79; BFB138) were used to generate product 1; primer pair 5'-GGAACATCGCTACCTATGAAGAACAGGCCAAAAAACATTAAGAAGGGCC -3' (SEQ ID NO:80; BFB139) and 5'-GTATCTTATCATGTCTGCTCGAAGCG-3' (SEQ ID NO:81; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *SalI/Xba*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/Xba*I.

To synthesize the plasmid encoding the CPM-FF Luc/RIIβB based cAMP sensor with an N-terminal peptide linker length of (X=10), lacking a C-terminal peptide linker, with R2βB residues 266-400 (10, -14; pBFB134), two separate primer pairs were used to amplify RIIβB DNA to generate two separate PCR

products. Primer pair 5'- AAAAAAGTCGACCGGAGGTTCAGGCGGTTC-3' (SEQ ID NO:82; BFB127) and 5'-GGCCCTTCTTAATGTTTTTGGCTTCATAGGTAGCGATGTTCCTTTTC -3' (SEQ ID NO:83; BFB140) were used to generate product 1; primer pair 5'-GAAAAGGAACATCGCTACCTATGAAGCCAAAAAACATTAAGAAGGGCC-3' (SEQ ID NO:84; BFB141) and 5'-GTATCTTATCATGTCTGCTCGAAGCG-3' (SEQ ID NO:85; BFB34) were used to generate product 2. SOE PCR of the two products yielded the full-length PCR product, which was subsequently digested with *SalI/XbaI* restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/XbaI*.

B. Functional characterization of CPM-FF Luc/RIIβB based cAMP sensors with variable X/Y peptide linker lengths

i. Functional characterization of CPM-FF Luc/RIIβB based cAMP sensors with X/Y peptide linkers in the set [2x (x = 0-5), 2y (y = 0-5)]

Luciferase activity in the presence and absence of cAMP was measured for the CPM-FF Luc/RIIβB cAMP sensors with X/Y linker lengths in the set [2x (x = 0-5), 2y (y = 0-5)] following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

400 ng plasmid DNA
10 µL Rabbit Retic Extract
0.8 µL TNT reaction buffer
0.4 µL T7 polymerase
0.4 µL amino acid mixture
0.4 µL rRNasin
dH_2O to 20 µL total volume

Following incubation at 30°C for 1.5 hours, the respective fusion proteins were incubated in the presence or absence of 100 µM cAMP by combining 9 µL of TNT® reaction with 1 µL of 1 mM cAMP stock or dH_2O. Following incubation for ≥ 15 minutes at room temperature, 1 µL of sample was added to 100 µL of Luciferase Assay Reagent (LAR; Promega Corp.) solution +/- 100 µM cAMP (90

μL LAR + 10 μL 1 mM cAMP stock or dH_2O). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). Overall, a trend was observed with CPM-FF Luc/RIIβB fusions with X/Y linker lengths in the set [2x (x = 0-5), 2y (y = 0-5)] where increasing luciferase activity was measured in the presence or absence of 100 μM cAMP with increasing peptide linker length (Figure 12). In addition, a second trend was observed where the fold induction of luciferase activity in the presence of 100 μM cAMP increased with increasing peptide linker length (Figure 13).

ii. Functional characterization of CPM-FF Luc/RIIβB based cAMP sensors with X/Y peptide linkers in the sets [10, -2n (n = 1-7)], [10, 2n (n = 1-5)], and [10 + 2n (n = 0-5), 0] amino acid residues

Luciferase activity in the presence and absence of 100 μM cAMP was measured for the CPM-FF Luc/RIIβB cAMP sensors with X/Y linker lengths in the sets [10, -2n (n = 1-7)], [10, 2n (n = 1-5)], and [10 + 2n (n = 0-5), 0] amino acid residues following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

400 ng plasmid DNA
10 μL Rabbit Retic Extract
0.8 μL TNT reaction buffer
0.4 μL T7 polymerase
0.4 μL amino acid mixture
0.4 μL rRNasin
dH_2O to 20 μL total volume

Following incubation at 30°C for 1 hour, the respective fusion proteins were incubated in the presence or absence of 100 μM cAMP by combining 9 μL of TNT® reaction with 1 μL of 1 mM cAMP stock or dH_2O. Following incubation at room temperature for ≥ 9 minutes, 1 μL of sample was added to 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.) solution +/- 100 μM cAMP (90 μL LAR + 10 μL 1 mM cAMP stock or dH_2O). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). In

general, luciferase activity in the presence or absence of 100 μM cAMP decreased with increasing C-terminal truncation of RIIβB for CPM-FF Luc/RIIβB cAMP sensors lacking C-terminal peptide linkers (Figure 14). In addition, the maximal fold induction in the presence of 100 μM cAMP for CPM-FF Luc/RIIβB cAMP sensors of the set [10, -2n (n = 1-7)] and [10, 2n (n = 1-5)] was the sensor with peptide linkers of (X=10, Y=0; pBFB117). Moreover, CPM-FF Luc/RIIβB cAMP sensors of the set [10 + 2n (n = 0-5), 0] showed a maximal fold induction for the sensor with peptide linkers of (X=10, Y=0; pBFB117) amino acid residues (Figure 15).

Example VII

A cAMP Biosensor With Circularly Permuted Click Beetle Luciferase and the B Domain from the PKA Regulatory Subunit Type IIβ

A. Synthesis of a CPM-Click Beetle Luc expression plasmid for subsequent insertion of RIIβB (pBFB53)

To synthesize a click beetle variant of the plasmid synthesized in Example X, part A, primers 5'-TATAATGCTAGCGATCGCCATGGGCGTGACTGTGCTGGTGTATC-3' (SEQ ID NO:86; BFB94) and 5'-TTTTTTCTCGAGCCGCCGCCAGCTTTTTCGAGG-3' (SEQ ID NO:87; BFB95) were used to amplify the click beetle equivalent of the firefly luciferase fragment encoding residues 234-544 (click beetle luciferase amino acids 231-542) from plasmid pCBG68-basic (Genbank Acc# AY258593; Promega Corp). The resultant product was digested with *NheI*/*XhoI* restriction enzymes and ligated into the parent CPM-FF Luc (pBFB8) expression plasmid digested with *NheI*/*XhoI* to give plasmid intermediate 1. Subsequently, primers 5'-AAAAAAGAGCTCCGGTGAAAAGAACGTGATCTACGGCC-3' (SEQ ID NO:88; BFB96) and 5'-AAAAAATCTAGAGTTTAAACAGGGATCAATTGAGTACCCACAC-3' (SEQ ID NO:89; BFB97) were used to amplify the click beetle equivalent of the firefly luciferase fragment encoding residues 4-233 (click beetle luciferase amino acids 5-230) from plasmid pCBG68-basic (Genbank Acc# AY258593; Promega Corp). The

resultant product was digested with *SacI/XbaI* restriction enzymes and ligated into plasmid intermediate 1 described above digested with *SacI/XbaI*.

B. Synthesis of plasmids encoding CPM-click beetle Luc/RIIβB fusion proteins with peptide linkers of (X=4, Y=4; pBFB54) and (X=10, Y=4; pBFB55) amino acid residues.

To synthesize the construct with (X=4,Y=4) linker lengths, primers 5'-AAA AAA GTC GAC CGG AAT GTA TGA AAG CTT TAT TGA GTC ACT GCC-3' (SEQ ID NO:90; BFB51) and 5'-AAA AAA GAG CTC CCA ACA ATA TCC ATG TTC GTT CCA AAC-3' (SEQ ID NO:91; BFB20) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *SalI/SacI* restriction enzymes and ligated into the parent CPM-Click Beetle Luc (pBFB53).

To synthesize the construct with (X=10,Y=4) linker lengths, primers 5'-AAA AAA GAG CTC CCA ACA ATA TCC ATG TTC GTT CCA AAC-3' (SEQ ID NO:92; BFB20) and 5'-AAA AAA TCC GGA ATG TAT GAA AGC TTT ATT GAG TCA CTG CC-3' (SEQ ID NO:93; BFB21) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *BspEI/SacI* restriction enzymes and ligated into the parent CPM-Click Beetle Luc (pBFB53) expression plasmid digested with *BspEI/SacI*.

B. Functional characterization of CPM-click beetle Luc/RIIβB fusion proteins with peptide linkers of (X=4, Y=4; pBFB54) and (X=10, Y=4; pBFB55) amino acid residues.

The cAMP dose response of CPM-click beetle Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB54) and (X=10,Y=4; pBFB55) amino acid residues was measured following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturers recommended protocol:

2400 ng plasmid DNA
60 µL Rabbit Retic Extract
4.8 µL TNT reaction buffer
2.4 µL T7 polymerase

2.4 µL amino acid mixture
2.4 µL rRNasin
dH_2O to 120 µL total volume

Following incubation at 30°C for 1.5 hours, the respective fusion proteins were incubated with varying concentrations of cAMP by combining 9 µL of TNT® reaction with 1 µL of cAMP stock solution (final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, and 25 µM cAMP). Following equilibration at room temperature for approximately 20 minutes, 1 µL of sample was added to 100 µL of Luciferase Assay Reagent (LAR; Promega Corp.) solution containing the respective concentration of cAMP (90 µL LAR + 10 µL cAMP stock solution). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). CPM-click beetle Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4,Y=4; pBFB54) and (X=10,Y=4; pBFB55) amino acid residues showed fold inductions in luciferase activity at 25 µM cAMP of 4.0 and 5.5, respectively. However, the fold induction for the click beetle luciferase based cAMP sensors was less than the fold induction of the firefly luciferase based sensors at all concentrations tested (Figure 16).

Example VIII

A cAMP Biosensor Utilizing Circularly Permuted Firefly Luciferase and the B domain from the PKA Regulatory Subunit Type Iα

DNA encoding the B domain from the human PKA regulatory subunit type Iα (RIαB) was ligated into an expression vector encoding CPM-FF Luc/RIαB fusions [Luc2.0 (234-544)-linker X-human RIα (residues 245-381)-linker Y-Luc2.0 (4-233)].

A. Synthesis of CPM-FF Luc/RIαB fusion proteins with peptide linkers of (X=4, Y=4; pBFB56) and (X=20, Y=20; pBFB58) amino acid residues

To synthesize the construct with (X=4,Y=4) linker lengths, primers 5'-ATATAACTCGAGCGGAATGTATGAGGAATTCCTTAGTAAAGTCTCTATTTTAG-3' (SEQ ID NO:94; BFB98) and 5'-AAAAAAGAGCTCCCGACAGACAGTGACACAAAACTGTTGTAC-3' (SEQ

ID NO:95; BFB99) were used to amplify RIαB DNA (Genbank Acc# BC036285). The resultant product was digested with *Xho*I/*Sac*I restriction enzymes and ligated into the parent CPM-FF Luc (pBFB8) expression plasmid digested with *Xho*I/*Sac*I.

To synthesize the construct with (X=20,Y=20) linker lengths, primers 5'-ATTAAACCCGGGATGTATGAGGAATTCCTTAGTAAAGTCTCTATTTTAG-3' (SEQ ID NO:96; BFB102) and 5'-AAAAAATCCGGACCCGACAGACAGTGACACAAAACTGTTGTAC-3' (SEQ ID NO:97; BFB103) were used to amplify RIαB DNA from (Genbank Acc# BC036285. The resultant product was digested with *Sma*I/*Bsp*EI restriction enzymes and ligated into the parent CPM-FF Luc (pBFB8) expression plasmid digested with *Nru*I/*Age*I.

B. Functional characterization of CPM-FF Luc/RIαB fusion proteins with peptide linkers of (X=4, Y=4; pBFB56) and (X=20, Y=20; pBFB58) amino acid residues

The cAMP dose response of CPM-FF Luc/RIαB fusion proteins with X/Y linker lengths of (X=4, Y=4; pBFB56) and (X=20, Y=20; pBFB58) amino acid residues was measured following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

2400 ng plasmid DNA
60 µL Rabbit Retic Extract
4.8 µL TNT reaction buffer
2.4 µL T7 polymerase
2.4 µL amino acid mixture
2.4 µL rRNasin
dH_2O to 120 µL total volume

Following incubation at 30 °C for 1.5 hours, the respective fusion proteins were incubated with varying concentrations of cAMP by combining 9 µL of TNT® reaction with 1 µL of cAMP stock solution (final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, 25, and 100 µM cAMP). Following equilibration at room temperature for ≥ 10 minutes, 1 µL of sample was added to 100 µL of Luciferase Assay Reagent (LAR; Promega Corp.) solution containing the respective

concentration of cAMP (90 µL LAR + 10 µL cAMP stock solution). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). CPM-FF Luc/RIαB fusion proteins with X/Y linker lengths of (X=4, Y=4; pBFB56) and (X=20, Y=20; pBFB58) amino acid residues showed fold inductions in luciferase activity at 100 µM cAMP of 1.8. However, the fold induction for the RIαB based cAMP sensors was less than the fold induction of the RIIβB based sensors at concentrations ≥ 0.025 µM (Figure 17).

Example IX

A cAMP Biosensor Utilizing a Circularly Permuted Thermal Stable Luciferase and the B domain from the PKA Regulatory Subunit Type IIβ

A. Synthesis of a CPM-Thermal Stable Luc expression plasmid for subsequent insertion of RIIβB (pBFB45)

To synthesize a thermal stable luciferase (UltraGlo luciferase, Promega Corp.) primers 5'-AATTAAGCTAGCGATCGCCATGACGTCAGCAATTTTAACGGTAATACC-3' (SEQ ID NO:98; BFB88) and 5'-TTTTTTCTCGAGCCATTGGTGTGTTTTTCTAACATTTGTCTTAAC-3' (SEQ ID NO:99; BFB89) were used to amplify the UltraGlo luciferase equivalent of the firefly luciferase fragment encoding residues 234-544 (UltraGlo luciferase residues 233-543). The resultant product was digested with *NheI*/*XhoI* restriction enzymes and ligated into the parent CPM-FF Luc (pBFB8) expression plasmid digested with *NheI*/*XhoI* to give plasmid intermediate 1. Subsequently, primers 5'-AATTTTGAGCTCCGGTGATAAGAATATTTTATATGGGCCCGAAC-3' (SEQ ID NO:100; BFB90) and 5'-AAAAAATCTAGAGTTTAAACGGGATTAATTGCGTTACCAAAAGTAG-3 (SEQ ID NO:101; BFB91) were used to amplify the click beetle equivalent of the firefly luciferase fragment encoding residues 4-233 (UltraGlo luciferase residues 3-232). The resultant product was digested with *Sac*I/XbaI restriction enzymes and ligated into plasmid intermediate 1 described above digested with *SacI*/*Xba*I.

B. Synthesis of plasmids encoding CPM-Thermal Stable Luc/RIIβB fusion proteins with peptide linkers of (X=4, Y=4; pBFB51) and (X=20, Y=20; pBFB52) amino acid residues

To synthesize the plasmid encoding the CPM-Thermal Stable Luc/RIIβB fusion protein with (X=4,Y=4) linker lengths, primers 5'-AAA AAA GTC GAC CGG AAT GTA TGA AAG CTT TAT TGA GTC ACT GCC-3' (SEQ ID NO:102; BFB51) and 5'-AAA AAA GAG CTC CCA ACA ATA TCC ATG TTC GTT CCA AAC-3' (SEQ ID NO:103; BFB20) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *SalI*/*Sac*I restriction enzymes and ligated into the parent CPM-Thermal Stable Luc expression plasmid (pBFB45) described above digested with *Xho*I/*Sac*I.

To synthesize the plasmid encoding the CPM-Thermal Stable Luc/RIIβB fusion protein with (X=20,Y=20) linker lengths, primers 5'-AAA AAA CCC GGG ATG TAT GAA AGC TTT ATT GAG TCA CTG CC-3' (SEQ ID NO:104; BFB23) and 5'-AAA AAA TCC GGA CCC AAC AAT ATC CAT GTT CGT TCC AAA C-3' (SEQ ID NO:105; BFB24) were used to amplify RIIβB DNA from ATCC 10625233 (Genbank ID BC075800). The resultant product was digested with *Bsp*EI/*Sma*I restriction enzymes and ligated into the parent CPM-Thermal Stable Luc expression plasmid described above digested with *Age*I/*Nru*I.

C. Functional characterization of CPM-Thermal Stable Luc/RIIβB fusion proteins with peptide linkers of (X=4, Y=4; pBFB51) and (X=20, Y=20; pBFB52) amino acid residues

The cAMP dose response of CPM-Thermal Stable Luc/RIIβB fusion proteins with X/Y linker lengths of (X=4, Y=4; pBFB51) and (X=20, Y=20; pBFB52) amino acid residues was measured following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

2400 ng plasmid DNA
60 µL Rabbit Retic Extract
4.8 µL TNT reaction buffer

2.4 µL T7 polymerase
2.4 µL amino acid mixture
2.4 µL rRNasin
dH_2O to 120 µL total volume

Following incubation at 30°C for 1.5 hours, the respective fusion proteins were incubated with varying concentrations of cAMP by combining 9 µL of TNT® reaction with 1 µL of cAMP stock solution (final concentrations of 0, 0.01, 0.025, 0.1, 0.25, 1, 2.5, 10, 25, and 100 µM cAMP). Following equilibration at room temperature for ≥ 19 minutes, 1 µL of sample was added to 100 µL of Luciferase Assay Reagent (LAR; Promega Corp.) solution containing the respective concentration of cAMP (90 µL LAR + 10 µL cAMP stock solution). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). The CPM-Thermal Stable Luc/RIIβB fusion protein with X/Y linker lengths of X=4, Y=4 amino acid residues (pBFB51) showed a fold induction in luciferase activity at 100 µM cAMP of 1.5 (Figure 18). However, the CPM-Thermal Stable Luc/RIIβB fusion protein with X/Y linker lengths of X=20, Y=20 amino acid residues (pBFB52) was unresponsive to cAMP (Figure 18). In both cases, the fold induction in luciferase activity for CPM-Thermal Stable Luc/RIIβB based cAMP sensors was less than the fold induction of the firefly luciferase based sensors at concentrations ≥ 0.025 µM (Figure 18).

Example X

Intracellular Detection of Changes in cAMP Concentration using CPM *Renilla* Luciferase/RIIβB Biosensor (Forskolin Titration)

Cell Culture

100 µl HEK-293 cells were plated in a 96 well plate and grown to 70-90% confluency in DMEM/F12 with HEPES buffer (Invitrogen) with 10% FBS at 37°C with 5% CO_2.

Transfections

Cells were transfected with *Trans*It®-LT1 Reagent (MIRUS) using 0.3 µl *Trans*It®-LT1 reagent and 0.15 µg DNA (CPM-hRL/RIIβB cAMP biosensor with

X/Y peptide linker lengths of (X=4, Y=20) (201325.78.E5)) per well of a 96-well plate. Cells were allowed to grow overnight and were assayed the next day.

Modulation of Biosensor

Approximately 1 day after transfection, cells were removed from incubator and equilibrated to room temperature. A 10 µl aliquot of 600 µM EnduRen Live Cell Substrate (Promega) was added to a total of 100 µl of cell culture to give a final concentration of approximately 60 µM coelentrazine. Cells were then incubated at room temperature for at least 15 minutes. After 15 minutes at room temperature, baseline measurements of luminescence were measured using a 96-well Veritas Luminometer (Turner) at 0.5 seconds per well. Cells were then induced with 0.025 µM – 250 µM forskolin (Sigma) or not induced (0.1% DMSO (Sigma)) and luminescence was measured continuously for about 30 minutes (Figure 19). Samples were measured in sets of 5 replicates per concentration of forskolin. EC_{50}s were calculated using GraphPad Prism for Windows, Version 4.

Results

Light output increased from cells transfected with DNA encoding the CPM-hRL/RIIβB cAMP biosensor with X/Y peptide linker lengths of (X=4, Y=20) (201325.78.E5) following stimulation with forskolin (Figure 19). Maximal levels of forskolin induced light output 3.6-fold above that of untreated cells. In addition, the EC_{50} of the forskolin response was 0.059 µM (Figure 19).

Example XI

A cGMP Biosensor Utilizing Circularly Permuted Firefly Luciferase and the B domain from the cGMP activated protein kinase (GKI-B) or human phosphodiesterase 2A (PDE2A)

cGMP is an important cellular second messenger with a variety of physiological functions, particularly in the cardiovascular and nervous systems. A series of cGMP sensors were prepared by fusing a circularly permuted firefly luciferase to a cGMP binding domain.

A. Synthesis of plasmids encoding CPM-FF Luc/GKI-B fusion proteins with peptide linkers of (X=4, Y=4) and (X=10, Y=10) amino acid residues.

To synthesize the construct with (X=4,Y=4) linker lengths, primers 5'-AAAAAACTCGAGCGGATTAAAAAGCGTTCCAACATTCCAG -3' (SEQ ID NO:106; BFB151) and 5'-AAAAAAGAGCTCCCAGACAGCTTCAGGTTGGCGAAG -3' (SEQ ID NO:107; BFB163) were used to amplify human GKI-B DNA (Origene, cat # TC116252; Genbank Acc # NM_006258), for instance, DNA corresponding to residues 231 to 350 (pBFB164, pBFB165) or 231 to 373 (pBFB171, pBFB172). The resultant product was digested with *XhoI/Sac*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *XhoI/Sac*I.

To synthesize the construct with (X=10,Y=10) linker lengths, primers 5'-AAAAAATCCGGATTAAAAAGCGTTCCAACATTCCAG -3' (SEQ ID NO:108; BFB153) and 5'- AAAAAAAGGCCTGACAGCTTCAGGTTGGCGAAG -3' (SEQ ID NO:109; BFB164) were used to amplify human GKI-B DNA (Origene, cat # TC116252; Genbank Acc # NM_006258). The resultant product was digested with *Bsp*EI/*Stu*I restriction enzymes and ligated into the parent CPM-FF Luc expression plasmid (pBFB8) digested with *BspE*I/*Zra*I.

B. Functional characterization of CPM-FF Luc/GKI-B fusion proteins with X/Y linker lengths of (X=4,Y=4) and (X=10,Y=10) amino acid residues

Luciferase activity in the presence and absence of 100 μM cGMP was measured for the CPM-FF Luc/GKI-B fusion proteins with X/Y linker lengths of (X=4,Y=4) and (X=10,Y=10) amino acid residues following expression using the TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

400 ng plasmid DNA
10 μL Rabbit Retic Extract
0.8 μL TNT reaction buffer
0.4 μL T7 polymerase
0.4 μL amino acid mixture
0.4 μL rRNasin
dH_2O to 20 μL total volume

Following incubation at 30°C for 1 hour, the respective fusion proteins were incubated in the presence or absence of 100 μM cGMP by combining 9 μL of TNT® reaction with 1 μL of 1 mM cGMP stock or dH_2O. Following incubation for ≥ 10 minutes at room temperature, 1 μL of sample was added to 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.) solution +/- 100 μM cGMP (90 μL LAR + 10 μL 1 mM cGMP stock or dH_2O). Luminescence was measured using a Veritas Microplate Luminometer (Turner Biosystems; program Bright-Glo). The CPM-FF Luc/GKI-B fusion protein with (X=4, Y=4) linker lengths (pBFB171) showed a 2-fold decrease in luciferase activity in the presence of 100 μM cGMP. In addition, the CPM-FF Luc/GKI-B fusion protein with (X=10, Y=10) linker lengths (pBFB172) showed a 1.5-fold decrease in luciferase activity in the presence of 100 μM cGMP.

Table 2

pBFB	Linker combination	RLU with 100 μM cGMP	RLU with No cGMP
pBFB171	(X=4, Y=4)	247,801	497,938
pBFB172	(X=10, Y=10)	1,148,496	1,707,449

C. Synthesis of plasmids encoding CPM-FF Luc/human phosphodiesterase 2A (PDE2A; Genbank NM_002599; amino acid residues 416-549)

DNA sequences encoding circularly permuted firefly luciferase constructs with engineered N- and C-termini at residues 234 and 233, respectively, were fused to a sequence encoding human PDE2A, which has a different protein fold relative to the RIIβB domain [Met-(Luc2.0 234-544)-(Linker X)-(human PDE2A 416-549)-(Linker Y)-(Luc2.0 4-233)-Val]. The cGMP binding domain from human PDE2A belongs to a large family of small molecule binding units called GAF domains. Constructs were made with X/Y linker lengths of (pBFB167; X=4, Y=4) (pBFB168; X=10, Y=10), and (pBFB169; X=20, Y=20) amino acid residues (Figure 21). PDE2A based biosensors were identified with 2 and 11 fold induction in luminescence activity in the presence of 100 μM cGMP for constructs with (pBFB168; X=10, Y=10) and (pBFB169; X=20, Y=20) amino acid linkers,

respectively, following expression *in vitro* using the T7 Coupled Reticulocyte Lysate System (Figure 22). Moreover, activation of these biosensors by cGMP was found to be dose dependent and selective over cAMP in separate experiments following expression using the T7 Coupled Reticulocyte Lysate System (pBFB169; Figure 23).

Thus, these cGMP sensors are useful for the detection of changes in cGMP concentration *in vitro*, and these biosensors will likely be useful for detecting changes in cGMP concentration in living cells for use in cell culture experiments or for whole animal imaging studies.

Example XII

Luciferase Calcium Biosensors

Calcium biosensors were prepared by fusing sequences encoding a circularly permuted firefly luciferase having engineered N- and C-termini at residues 234 and 233, respectively, to sequences encoding protein domains that bind calcium. One type of calcium biosensor utilized a mutant of fast chicken skeletal muscle troponin C (TnC) (amino acids 15-163; N109D, D111N, N145D, D147N; Genbank NM_205450) [Met-(Luc2.0 234-544)-(Linker X)-(TnC)-(Linker Y)-(Luc2.0 4-233)-Val], and the second type of calcium biosensor utilized human calmodulin (CaM) (amino acids 5-148; Genbank BC005137) [Met-Luc+ (234-544)-(Linker X)-human Calmodulin (5-148)-(Linker Y)-Luc+ (4-233)].

CPM-FF Luc/TnC and CPM-FF Luc/CaM constructs with varying X/Y peptide linker lengths were expressed *in vitro* using the T7 Coupled Reticulocyte Lysate System (pBFB225, pBFB226, pBFB227, pBFB7; Figure 24). Reactions were then supplemented with 10 mM $CaCl_2$ or 10 mM EDTA plus 2.5 mM EGTA. A maximal response was seen for a CPM-FF Luc/CaM biosensor with (X=8, Y=8; pBFB7), where X = LEGSGGGG (SEQ ID NO:306) and Y = GGGGSGPW (SEQ ID NO:307), with a greater than 60 fold reduction in luminescence activity in the presence of calcium. Similar responses, although of lower magnitude, were seen for CPM-FF Luc/CaM constructs with different X/Y peptide linker lengths (pBFB225, pBFB226, pBFB227). No response was seen for a control construct having a

random 42 amino acid linker or for wild-type firefly luciferase (pBFB8 and pBFB22; Figure 25).

These biosensors will likely be useful for the detection of changes in calcium concentration both *in vitro* and inside living cells.

Example XIII

cAMP Biosensors Using Multiple Sites of Modification in Firefly Luciferase

Additional sites of modification, such as circular permutation, can be used for the development of a firefly luciferase biosensor, e.g., a cAMP biosensor. Above, a cAMP biosensor was prepared using a circularly permuted mutant of firefly luciferase with the primary structure Met-(Luc2.0 residues 234-544)-GSSGGSGGSGGG-RIIβB-(Luc2.0 residues 4-233)-Val (SEQ ID NO:184; RIIβB is the B cAMP binding domain from human PKA regulatory domain type IIβ amino acids 266-414). Analogous constructs were prepared using firefly luciferase mutants circularly permuted at additional residues. Overall, twenty-three independent constructs were tested that encoded fusion proteins of the following type: Met-(Luc2.0 residues X-544)-GSSGGSGGSGGG-RIIβB-(Luc2.0 residues 4-Y)-Val (GSSGGSGGSGGG corresponds to SEQ ID NO:121; Figure 26 lists X/Y values for the various constructs). For each of these constructs, excluding the construct with circular permutation at residue 255, a site was chosen in a solvent exposed surface loop bounded by secondary structures such as a beta sheet or alpha helix, for circular permutation using PDB file 1LCI (http://www.rcsb.org/pdb/home/home.do). Solvent exposed surface loops may be more amenable as sites of modification, such as circular permutation, than sites buried in the protein core or sites that are involved in alpha or beta structures. This is supported by the lack of activity seen for the construct with circular permutation at 255, where Tyr255 is a component of an alpha helix that is buried in the protein core. This collection of constructs represents the majority, but not all, surface turns seen in the 1LCI crystal structure.

Following expression using the TNT T7 Coupled Reticulocyte Lysate System, a number of different sites of circular permutation were identified where luminescence activity exceeded the background detection levels of the luminometer and fold inductions in the presence of 100 μM cAMP were greater than two-fold (CPM sites: 37, 47, 75, 83, 107, 144, 160, 188, 225, 233, 242, 268, 308, 358, 377, 403, and 490). In addition, constructs were identified where the fold induction in luminescence activity was greater than CPM 233, with maximal fold activation values greater than 200 fold in this experiment. DNA encoding select constructs was transferred to a mammalian expression vector containing a CMV promoter (pF9A; Promega Corp.). The constructs were: pBFB317 (CPM site 268), pBFB318 (CPM site 358), pBFB319 (CPM site 47), pBFB321 (CPM site 225), pBFB322 (CPM site 233), pBFB325 (CPM site 308), pBFB326 (CPM site 377), pBFB327 (CPM site 403), pBFB328 (CPM site 75), and pBFB329 (CPM site 83) (see Figure 26 for X, Y values). Following transient transfection with DNA encoding the various Met-(Luc2.0 residues X-544)-GSSGGSGGSGGG-RIIßB-(Luc2.0 residues 4-Y)-Val (GSSGGSGGSGGG corresponds to SEQ ID NO:121) constructs, HEK293 cells were treated with 50 μM forskolin to activate endogenous adenylate cyclase. Following incubation for 16 minutes, luminescence was measured from the live cell populations. As predicted, the various constructs functioned as cAMP biosensors inside living cells. Interestingly, constructs that showed the highest fold induction inside cells were not the same constructs with the highest fold induction *in vitro* (compare Figures 27-28).

Example XIV

A Nonpermuted *Renilla* Luciferase cAMP Biosensor

As described herein, circularly permuted *Renilla* luciferase constructs can be employed as a biosensor. Nonpermuted *Renilla* luciferase constructs having RIIβB inserted into sites tolerant to modification, e.g., between residues 91/92, 223/224 or 229/230, were prepared. Constructs were generated as described above. They are: hRL(1-91)-4 amino acid peptide linker-RIIBetaB-4 amino acid peptide linker-hRL (92-311) (201360.17.A3), hRL(1-91)-4 amino acid peptide linker-RIIBetaB-20

amino acid peptide linker-hRL992-311) (201360.17.A12), hRL(1-91)-10 amino acid peptide linker-RIIBetaB-4 amino acid linker-hRL(92-311) (201360.17.D7), hRL(1-91)-42 amino acid peptide linker-hRL(92-311) (201325.165.A2), hRL(1-223)-4 amino acid peptide linker-RIIBetaB-4 amino acid linker-hRL(224-311) (201360.24.A1), hRL(1-223)-4 amino acid peptide linker-RIIBetaB-20 amino acid linker-hRL(224-311) (201360.24.A10), hRL(1-223)-10 amino acid peptide linker-RIIBetaB-4 amino acid linker-hRL(224-311) (201360.24.C5), hRL(1-223)-10 amino acid peptide linker-RIIBetaB-20 amino acid linker-hRL(224-311) (201360.24.E11), hRL(1-223)-42 amino acid peptide linker-hRL(224-311) (201325.177.B7), hRL(1-229)-4 amino acid peptide linker-RIIBetaB-4 amino acid linker-hRL(230-311) (201360.19.E9), hRL(1-229)-4 amino acid peptide linker-RIIBetaB-20 amino acid linker-hRL(230-311) (201360.54.A1), hRL(1-229)-42 amino acid peptide linker -hRL(230-311) (201325.165.C5) (Figure 29).

Protein was expressed from the constructs using the TnT T7 Coupled Wheat Germ Lysate System, 17 µL of TNT reaction was mixed with 17 µL of 300 mM HEPES/200 mM thiourea (pH about 7.5) supplemented with 3.4 µL of 1 mM cAMP stock or dH_2O; reactions were allowed to incubate at room temperature for approximately 10 minutes. Ten µL of each sample was added to a 96 well plate well in triplicate and luminescence was measured using 100 µL of *Renilla* luciferase assay reagent on a Glomax luminometer.

The hRL(1-91)-linker-RIIBetaB-linker-hRL(92-311) proteins were induced by about 12 to 23 fold, the hRL(1-223)-linker RIIBetaB-linker-hRL(224-311) proteins were not induced and the hRL(1-229)-linker-RIIBetaB-(230-311) proteins were induced by about 2 to 9 fold. None of the 42 amino acid linker constructs were induced, nor were the full length *Renilla* luciferase construct (201325.50.A7) or the "no DNA" controls (Figure 30).

Example XV

Light Output and Fold Induction Vary as a Function of X/Y Peptide Linker Lengths for CPM-hRL91 Luc/RIIβB Based cAMP Sensors

Constructs encoding CPM-hRL91 Luc/RIIβB based cAMP sensors with variable X/Y peptide linker lengths were generated (Figure 31). Protein was expressed from the constructs using the TnT T7 Coupled Wheat Germ Lysate System, 17 µL of TNT reaction was mixed with 17 µL of 300mM HEPES/200 mM thiourea (pH about 7.5) supplemented with 3.4 µL of 1 mM cAMP stock or dH_2O; reactions were allowed to incubate at room temperature for approximately 10 minutes. Ten µL of each sample was added to a 96 well plate well in triplicate and luminescence was measured using 100 µL of *Renilla* luciferase assay reagent on a Glomax luminometer. As shown in Figure 32, light output and fold induction varied with linker length. Fold induction ranged from about 87 to 331. The 42 amino acid linker construct, the full length *Renilla* luciferase construct and the "no DNA" control were not induced (Figure 32).

Example XVI

A cAMP Biosensor Utilizing Circularly Permuted *Renilla* Luciferase and the B domain from the PKA Regulatory Subunit Type Iα or a GAF domain

DNA encoding the B domain from the human PKA regulatory subunit type Iα (RIαB) was ligated into an expression vector encoding CPM-hRL91 Luc/RIαB fusions [hRL (92-311)-linker X-human RIα (residues 245-381)-linker Y-hRL (1-91)]; (X=4, Y=20; pBFB210), (X=4, Y=4; pBFB211), (X=10, Y=10; pBFB212) and (X=20, Y=20; pBFB213) (Figure 33). Protein was expressed from the constructs using the TnT T7 Coupled Wheat Germ Lysate System, 17 µL of TNT reaction was mixed with 17 µL of 300mM HEPES/200 mM thiourea (pH about 7.5) supplemented with 3.4 µL of 1 mM cAMP stock or dH_2O; reactions were allowed to incubate at room temperature for approximately 10 minutes. Ten µL of each sample was added to a 96 well plate well in triplicate and luminescence was measured using 100 µL of *Renilla* luciferase assay reagent on a Glomax luminometer. As shown in Figure 34, light output and fold induction varied with linker length. Fold induction ranged from about 2.8 to 6.8. The 42 amino acid linker construct (201325.15.A1), the full length *Renilla* construct (201325.50.A7) and the "no DNA" control were not induced (Figure 34).

An additional type of cAMP biosensor was constructed using a circularly permuted *Renilla* luciferase (hRL) and a GAF domain. The plasmid DNA construct encoding the following fusion protein: Met-(hRL 92-311)-GSSGGSGGSGGGSGGSGGSG-(GAF A domain from Trypanosoma brucei PDE; Genbank AF192755 amino acids 241-375)-GSGGSGGSGGTSGGSGGSSG-A-(hRL 3-91)-Val (SEQ ID NO:185) [clone pBFB232]. Following expression using the T7 Coupled Reticulocyte Lysate System, luminescence activity was measured in the presence or absence of exogenous cAMP. In the presence of cAMP, the measured activity was 7595 RLU; in the absence of cAMP, the measured activity was 298 RLU (about a 25 fold change). These results indicate that additional domains can be used in CPM hRL constructs in the generation of biosensors. This type of reagent may allow the monitoring of changes in cAMP concentration in living cells, and it also may provide distinct advantages over existing FRET-based cAMP biosensors in that assay format. Moreover, since the GAF domain is a highly conserved fold in nature responsible for binding a wide range of molecules, it is likely that additional types of CPM hRL biosensors could be made using this fold.

Example XVII

cAMP Biosensors Using Multiple Sites of Modification in *Renilla* Luciferase

A cAMP biosensor having a circularly permuted mutant of *Renilla* luciferase with the primary structure Met-(hRL 92-311)-GSTG-RIIßB-GSGGSGGSGGTSGGSGGSSG (hRL 2-91)-Val (SEQ ID NO:186; RIIßB is the B cAMP binding domain from human PKA regulatory domain type IIBeta amino acids 266-414) showed an increase in luminescence activity upon binding to cAMP. Analogous constructs, either "split" proteins or circularly permuted proteins, can be generated using *Renilla* luciferase mutants modified at additional residues. Overall, fourteen independent circularly permuted constructs were tested encoding fusion proteins of the following type: Met-(hRL X-311)- GSTG-RIIßB-GSGGSGGSGGTSGGSGGSSG(hRL 2-Y)-Val (GSTG corresponds to SEQ ID NO:122; GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:123). The following table provides X/Y values for the fourteen constructs.

Table 3

CPM site	X value	Y value	Clone ID
31	32	30	pBFB276
42	43	41	pBFB277
69	70	68	pBFB278
111	112	110	pBFB279
151	152	150	pBFB280
169	170	168	pBFB281
193	194	192	pBFB282
208	209	207	pBFB283*
251	252	250	pBFB284
259	260	258	pBFB285
274	275	273	pBFB286
91	92	91	pBFB287 and 201325.44.H6
223	224	223	201325.33.C9
229	230	229	201325.86.B1

*Note: for construct pBFB283, the last amino acids at the C terminal were PFSEFKPD (SEQ ID NO:120) instead of PFK and no Val was inserted prior to the stop codon.

For all but four of these constructs, a site was chosen in a solvent exposed surface loop for circular permutation using a homology model of *Renilla* luciferase using 1BN6 (*Rhodococcus sp.*) and 2DHD (*Xanthobacter autotrophicus*) haloalkane dehalogenase crystal structures as templates. Solvent exposed surface loops may be more amenable as sites of modification, e.g., circular permutation, than sites buried in the protein core or sites that are involved in alpha or beta structures. This hypothesis is supported by the lack of activity seen for the firefly luciferase construct with circular permutation at 255, where Tyr255 is a component of an alpha helix that is buried in the protein core. This collection of constructs represents some, but not all, of the surface turns seen in the homology model structure. Four CPM sites: 91, 111, 223 and 229, were chosen based on previous reports (Kaihara et al., 2003, Remy et al., 2005 and Paulmurugan et al., 2003). The constructs were expressed using the TNT T7 Coupled Reticulocyte Lysate System or TnT T7 Coupled Wheat Germ Extract System and tested *in vitro* (Figures 35 and 36).

The results indicate that a number of different sites of circular permutation can be used to generate a biosensor such as a cAMP biosensor. Alternative sites of circular permutation were identified with uninduced/induced levels of activity

greater than the initial construct with circular permutation at 91 (CPM 91). In addition, constructs were identified where the fold induction in luminescence activity was greater than CPM 91. In addition, owing to the very low solubility of CPM 91 when expressed in *E. coli*, the additional constructs will be tested for increased solubility compared to this construct. Increased solubility may facilitate the development of an *in vitro* biosensor such as a cAMP detection reagent.

The results also indicate that a number of sites are not useful for circular permutation. All the sites between residues 169 and 274 had low induced and uninduced activities and the fold induction in luminescence activity was about 2 fold or lower.

Constructs were designed in a vector backbone (pF5A; Promega Corp.) which allows for both *in vitro* expression (T7 promoter) as well as mammalian expression (CMV promoter). Following transient transfection with DNA encoding the various Met-(hRL residues X-311)-GSTG-RIIβB-GSGGSGGSGGTSGGSGGSSG-(hRL residues 2-Y)-Val (GSTG corresponds to SEQ ID NO:122; GSGGSGGSGGTSGGSGGSSG corresponds to SEQ ID NO:123) constructs (pBFB276, pBFB277, pBFB278, pBFB279, pBFB280, pBFB287), HEK293 cells were treated with 100 μM forskolin to activate endogenous adenylate cyclase. Following incubation for 14 minutes, luminescence was measured from the live cell populations. As predicted, the various constructs functioned as cAMP biosensors inside living cells. Interestingly, construct CPM 31 showed the highest fold induction *in vitro*, however, this was not the case inside cells. However, in general, the light output and fold inductions showed similar trends *in vitro* and *in vivo* (Figure 37).

Example XVIII

A number of different genetic constructs were prepared to test the possibility of creating biosensors using *Gaussia* luciferase (Gluc) lacking the seventeen amino acid N-terminal peptide that acts as a secretion signal (Genbank AAG54095; amino acids 18-185). *Gaussia* luciferase with or without the N-terminal signal peptide has been reported to give greater light intensity relative to other luciferases when

measured from living cells (Tannous et al., 2005; Remy et al., 2006). In addition, fragments of Gluc have been used in systems of protein complementation (Gluc split at amino acid residue 110; Remy et al., 2006); thus, it is likely that Gluc will also be amenable to circular permutation at this site or other sites.

To prepare a Gluc cAMP biosensor, predictions of protein secondary structure were used to choose various sites of Gluc circular permutation: Met-(Gluc A- 185)-(Linker X)-(human RIIbetaB Genbank BC075800 amino acid residues 266-414)-(Linker Y)-(Gluc 18-B).

Table 4

CPM site	A residue	B residue	Length Linker X	Length Linker Y	pBFB#
100	101	99	4	4	pBFB290
100	101	99	10	10	pBFB291
100	101	99	20	20	pBFB292
110	111	109	4	4	pBFB293
110	111	109	10	10	pBFB294
110	111	109	20	20	pBFB295
48	49	47	4	4	pBFB296
48	49	47	10	10	pBFB297
48	49	47	20	20	pBFB298
68	69	67	4	4	pBFB299
68	69	67	10	10	pBFB300
68	69	67	20	20	pBFB301
84	85	83	4	4	pBFB302
84	85	83	10	10	pBFB303
84	85	83	20	20	pBFB304
91	92	90	4	4	pBFB305
91	92	90	10	10	pBFB306
91	92	90	20	20	pBFB307
114	115	113	4	4	pBFB308
114	115	113	10	10	pBFB309
114	115	113	20	20	pBFB310
126	127	125	4	4	pBFB311
126	127	125	10	10	pBFB312
126	127	125	20	20	pBFB313
162	163	161	4	4	pBFB314
162	163	161	10	10	pBFB315
162	163	161	20	20	pBFB316

Where the various linker combinations have the sequences:

Table 5

Linker combination	Sequence
(X=4,Y=4)	GSTG-RIIbetaB-GSSG (SEQ ID NO:187)
(X=10, Y=10)	GSSGGSGGSG-RIIbetaB-GSGGSGGSSG (SEQ ID NO:188)
(X=20,Y=20)	GSSGGSGGSGGGSGGSGGSG-RIIbetaB-GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:189)

Sites useful for a Gluc cAMP may be substituted to generate biosensors for other molecules using this site of circular permutation. Moreover, sites amenable to circular permutation in one copepod luciferase are likely useful in other copepod luciferases, such as the luciferase from *Metridia longa*.

Example XIX

Methods for cell-based GPCR assays can involve direct detection of intracellular signal transduction events. Among the most successful are methods using fluorescent dyes or aequorin for real-time monitoring of intracellular calcium. However, analogous technologies have been lacking for the detection of intracellular cAMP dynamics. A circularly permuted firefly luciferase with the allosteric RIIβB cAMP binding domain of Protein Kinase A is a sensor capable of emitting luminescence in proportion to the concentration of cAMP. Live cell, zero-step GPCR assays using this sensor allow the dynamic detection of changes in cAMP concentration using stable or transiently transfected cell lines. In addition, it is possible to develop a single-step homogenous assay format for detection of cAMP *in vitro* (Figure 38).

The ORF from pBFB135, under the class of biosensors called "CPM-FF Luc/RIIβB," was used to generate the transient and stable cells lines described below. These cell lines are called "CP234-Luc/RIIB," "cAMP LucSensor," "LucSensor," and "FF cAMP Sensor."

HEK293 cells stably expressing CP234-Luc/RIIB (ORF derived from pBFB135) were resuspended in complete media and mixed with 5 mM luciferin-EF. Cells were plated at 1 x 10^5 cells per well in a 96 well plate and equilibrated to room temperature for 1.5 hours. After stimulation with forskolin, luminescence was

measured at 15 minutes using a GloMax™ Luminometer. The results showed that this assay generate EC_{50} values of 0.36 μM for forskolin (Figure 38).

For a Z' measurement, 2 x 10^4 cells were aliquoted per well to a 384-well plate and equilibrated using a similar protocol. Half the plate was induced with 20 μM forskolin, whereas the other half remained uninduced. Luminescence was captured 15 minutes after induction using a TECAN GENios Pro™ luminometer. The fold of induction was 6.1 and Z' was 0.83. Since assays with Z' greater than 0.5 are considered good quality for high-throughput screening (HTS), the cAMP biosensor-based assay is amenable for HTS.

HEK293 cells stably expressing the dopamine D1 receptor were transiently transfected with plasmid DNA encoding CP234-Luc/RIIB or the R361K mutant (a mutation in the cAMP binding domain) (ORFs derived from pBFB135 and pBFB147, respectively). Cells were plated and equilibrated with luciferin-EF as described above, and compounds from a LOPAC library (plate 6) were added to each well (10 μM). Following incubation for 50 minutes, the plates were read on a TECAN GENios Pro™ luminometer. Hits that also were identified using a luciferase reporter gene assay (CRE response element) are shown in red (Figure 39). Most hits identified by the cAMP biosensor assay correlated with hits identified by the CRE-Luc reporter assay, validating the biological relevance of the cAMP biosensor GPCR assay.

Cells were also plated and equilibrated with luciferin-EF, and then after compound addition, luminescence was measured at 40 minutes using a GloMax™ Luminometer. The pharmacokinetic parameters of EC_{50} and IC_{50} values generated using the cAMP biosensor assay correlated well with those reported in the literature using other methods, again validating the biological relevance of the cAMP biosensor GPCR assay (Figure 40).

Responses were also tested in cells incubated at different temperatures (Figure 42) and with a variety of agonists and antagonists (Figure 43). HEK293 cells expressing the cAMP LucSensor (ORF derived from pBFB135) and a dopamine D1 receptor were incubated with luciferin for 1.5 hours at room temperature or 37°C, then contacted with agonist or antagonist. Responses were

measured on a luminometer. There was a more rapid and dynamic response to compounds when cells were incubated under physiological conditions, e.g., 37°C and CO_2. The results at 37°C were qualitatively similar to those expected for intracellular cAMP dynamics. At room temperature, there was a slower response with a lower dynamic range, which may be useful for large scale screening.

Figure 44 shows a time course for fold induction in cells stably transfected with the cAMP LucSensor and contacted with different amounts of dopamine. The results show that the system allows for monitoring of cAMP dynamics in live cells in real time. Moreover, the results in Figure 45 show that the system permits evaluation of compound potency, which is relatively consistent at different time points. Figure 46 provides potency rankings (EC_{50}) and results for various agonists and shows that some compounds are partial agonists. Data for antagonist potency (IC_{50}) is shown in Figure 47.

The cAMP LucSensor can also be used measure modulations of GPCR already expressed in the host cell (endogenous GPCR). An example is shown using HEK293 cells which expressed beta2-adrenergic receptor and stably transfected with the cAMP LucSensor. Following similar protocols as described for the dopamine receptor, Figures 48-49 showed the potency ranking of various agonists and antagonists, respectively.

A comparison of three bioluminescent GPCR assays was conducted. The results for those assays and agonists are shown in Figure 50. The results for the three bioluminescent assays with antagonists are shown in Figure 51. The rankings for the tested compounds were the same in all three assays.

The increased in luminescence of the cAMP LucSensor in the presence of cAMP may be the result of an increased efficiency in a conformational change from "open" to "closed".

HEK293 cells stably expressing the dopamine D1 receptor were also transiently transfected with plasmid DNA encoding CPM-hRL Luc/ RIIβB X=4, Y=20 under the CMV promoter (201325.78.E5) or TK promoter (201325.44.H6) and then stimulated with either forskolin, SKF38393 or dopamine. Wild type *Renilla* luciferase and CPM-hRL Luc without the RIIβB domain were also tested

and showed no specific response to cAMP modulation (data not shown). Cells were transfected in a T75 flask with *Trans*It®-LT1 Reagent (MIRUS) using 60 µL *Trans*It®-LT1 reagent and 30 µg DNA per flask, allowed to grow over night and assayed the next day. Approximately one day after transfecting, cells were removed from incubator and trypsinized, counted and 10,000 cells per well were plated in a 96 well plate in DMEM/F12 (HEPES buffer, Invitrogen) with 10% FBS and 60 µM EnduRen Live Cell Substrate. EnduRen Live Cell Substrate (Promega) was reconstituted in 100 µL DMSO and was added to pre-warmed complete media to a final concentration of 60 µM. Cells were then incubated for at least 1 hour at 37°C and then cooled to room temperature. After 15 minutes at room temperature, baseline measurements of luminescence were measured using a 96 well GloMax™ Luminometer at 0.5 seconds per well. Cells were then induced with 10x stocks, made in complete media, of Forskolin (Sigma), SKF38393 (Sigma), Dopamine (Sigma) or not induced (0.1% DMSO (Sigma)) and luminescence was measured continuously for about 30 minutes. Samples were measured in sets of four replicates per concentration of Forskolin, Dopamine or SKF38393. EC_{50} data represents 15 minutes after induction and were calculated using GraphPad Prism for Windows, Version 4.

Similar to the CPM-FF Luc/RIIβB biosensor, the EC_{50} values generated using the CPM-hRL Luc/RIIβB X=4, Y=20 biosensor (201325.44.H6 and 201325.78.E5) correlated well with those reported in the literature using other methods, again validating the biological relevance of the cAMP biosensor GPCR assay (Figure 52A-D).

Example XX

Intracellular Detection of Changes in cAMP Concentration using CPM-hRL Luc/RIIβB cAMP Biosensors

Cell Culture

Cells were cultured in 2 mL DMEM/F12 with HEPES buffer (Invitrogen) and 10% FBS at 37°C with 5% CO_2 in a 6 well plate.

Plasmids

Three of the constructs described in Example XVII were used to detect intracellular changes of cAMP concentrations. The constructs used were: pBFB277, pBFB279 and pBFB287. HEK293 cells stably expressing CPM91-hRL/RIIβB (ORF derived from 201325.44.H6 were also used in these experiments.

Transfections

HEK293 cells were transfected with *Trans*It®-LT1 Reagent (MIRUS) using 6 μL *Trans*It®-LT1 reagent and 2 μg DNA (pBFB277, pBFB279 and pBFB287) per well of a 6 well plate. Cells were allowed to grow overnight and were assayed the next day.

Modulation of biosensor

Approximately 1 day after transfection, cells were trypsinized, resuspended in fresh DMEM/F12 with HEPES buffer (Invitrogen) with 1% FBS and plated in a 96 well plate at approximately 10,000 cells per well. Alternatively, a HEK293 cell line stably expressing CP91-hRL/RIIβB was plated in a 96 well plate at approximately 10,000 cells per well. A 10 μL aliquot of 600 μM EnduRen was added to a total of 100 μL of cell culture to give a final concentration of approximately 5.5 μM EnduRen. Cells were then incubated at 37°C with 5% CO_2. After 5 hours, the plate was removed from the incubator and allowed to cool to room temperature for at least 20 minutes. After 20 minutes, baseline measurements of luminescence were measured using a 96 well Veritas Luminometer (Turner Biosystems; integration time of 0.5 seconds per well). Cells were then induced with 10 μM isopreterenol (CalBiochem), 50 μM forskolin (Sigma) or not induced (0.1% DMSO, Sigma) and luminescence was measured continuously for about 30 minutes. After 30 minutes, 10 μM propranolol (Sigma) was added to cells already induced with isopreterenol and 0.1% DMSO was added to all other samples. Luminescence was then measured continuously for the next 30 minutes. A final addition of 50 μM forskolin was added to the isopreterenol/propranolol sample and 0.1% DMSO was added to all other samples. Luminescence was then measured continuously for the next half hour. Samples were measured in sets of 4-6 replicates. 10x stocks of isopreterenol, propranolol, forskolin and DMSO were made in DMEM/F12 with HEPES buffer (Invitrogen) and 1% FBS.

Results

To measure changes in the intracellular concentration of cAMP, HEK 293 cells were transiently transfected with three CPM-hRL Luc/RIIβB (X=4, Y=20) constructs (circularly permuted at different positions within *Renilla* luciferase) followed by treatment with compounds known to increase the intracellular cAMP concentration through GPCR activation (isopreterenol, a β-adrenergic receptor agonist), decrease intracellular cAMP concentration through GPCR inhibition (propranolol, a β-adrenergic receptor antagonist), or increase intracellular cAMP concentration through activation of adenylate cyclase (forskolin). Both isopreterenol and forskolin treatment alone increased light output from transfected cells approximately 2-fold, reflecting an increase in intracellular cAMP concentration (Figure 53). In addition, a temporal response to changes in cAMP concentration was observed by treating the cells with isoperterenol, followed by propranolol, followed by forskolin (Figure 53). Detection of cAMP modulation using the *Renilla* luciferase biosensor was also demonstrated in HEK293 cells stably expressing CPM91-hRL/RIIβB. These data showed an about 5-fold increase in light output in response to isopreterenol and forskolin treatment (Figure 53). Similar to the transiently transfected cells, a temporal response to changes in cAMP concentration was observed by treating the cells with isoperterenol, followed propranolol, followed by forskolin (Figure 53).

Example XXI

Nonpermuted Firefly Luciferase cAMP biosensors

Various nonpermuted firefly luciferase constructs having RIIβB directly inserted into sites tolerant to modification, e.g., between residues 233/234, 355/359, 82/83, and 307/308, were prepared. DNA encoding the following fusion proteins was cloned into vector pF9A:

Table 6

pBFB403	Met-(Luc2.0 4-233)-GSTG-RIIbetaB-GSSG-(Luc2.0 234-544) (SEQ ID NO:172)
pBFB404	Met-(Luc2.0 4-233)-GSSGGSGGSG-R2betaB-GSGGSGGSSG-(Luc2.0 234-544) (SEQ ID NO:173)
pBFB405	Met-(Luc2.0 4-233)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(Luc2.0 234-544) (SEQ ID NO:174)

pBFB406	Met-(Luc2.0 4-355)-GSTG-RIIbetaB-GSSG-(Luc2.0 359-544) (SEQ ID NO:175)
pBFB407	Met-(Luc2.0 4-355)-GSSGGSGGSG-R2betaB-GSGGSGGSSG-(Luc2.0 359-544) (SEQ ID NO:176)
pBFB408	Met-(Luc2.0 4-355)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(Luc2.0 359-544) (SEQ ID NO:177)
pBFB409	Met-(Luc2.0 4-82)-GSTG-RIIbetaB-GSSG-(Luc2.0 83-544) (SEQ ID NO:178)
pBFB410	Met-(Luc2.0 4-82)-GSSGGSGGSG-R2betaB-GSGGSGGSSG-(Luc2.0 83-544) (SEQ ID NO:179)
pBFB411	Met-(Luc2.0 4-82)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(Luc2.0 83-544) (SEQ ID NO:180)
pBFB412	Met-(Luc2.0 4-307)-GSTG-RIIbetaB-GSSG-(Luc2.0 308-544) (SEQ ID NO:181)
pBFB413	Met-(Luc2.0 4-307)-GSSGGSGGSG-R2betaB-GSGGSGGSSG-(Luc2.0 308-544) (SEQ ID NO:182)
pBFB414	Met-(Luc2.0 4-307)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(Luc2.0 308-544) (SEQ ID NO:183)

Luc2.0 = *Photinus pyralis* luciferase encoded by the *luc2.0* gene (see Genbank ID AY738222); RIIbetaB = residues 266-414 of human PKA regulatory subunit type II beta (Genbank BC075800)

Protein was expressed from these constructs using the TnT T7 Coupled Reticulocyte Lysate System. Following expression, 9 µL of TNT reaction was mixed with 1 µL 1 mM cAMP stock or H_2O, and the reactions were allowed to incubate at room temperature for approximately 15 minutes. Following incubation, 2 µL of solution was aliquoted to individual wells of a 96 well plate in triplicate. Luminescence was measured using a Glomax luminometer following injection of 100 µL of Luciferase Assay Reagent (0.5 second integration time).

The results indicate that cAMP biosensors can be generated by direct insertion of RIIβB into any of the four chosen insertion sites (see Figure 54). The results also indicate that sites that are tolerant to circular permutation also appear to be tolerant to direct insertion to generate viable biosensors.

Example XXII

Nonpermuted and Permuted *Oplophorus* Luciferase cAMP biosensors

Oplophorus gracilirostris luciferase (OpLuc) catalizes oxidation of coelentrazine to emit blue light. The mature form of the enzyme is 18.7 kD (169 aa). The original ORF includes 27 extra residues which represent a putative signal

peptide for secretion. Removal of the putative 27 aa signal peptide resulted in about 50 fold increase in the luciferase activity. Due to its small size, OpLuc is particularly amenable to use as a biosensor or in PCA.

OpLuc is active and stable if it is present in TnT cell free extract or *E.coli* cell lysate. However, it immediately inactivates upon purification. Gel filtration showed that the luciferase (expressed in *E. coli* without the 35 kD protein found in the native organism) eluted between 13.7 and 29 kD protein standards. MW of the enzyme is 18.7 kD. Therefore, it appears that, if expressed without the 35 kD protein, the luciferase is maintained as a monomer. The enzyme remains active at pH 7.5-9 and the activity begins to decrease at pH 9.5.

Various nonpermuted *Oplophorus* luciferase (OpLuc) constructs having RIIβB directly inserted into sites tolerant to modification, e.g., between residues 50/51 and 84/85, were prepared. DNA encoding the following fusion proteins was cloned into vector pF5K:

Table 7

pBF B397	Met-(OpLuc 1-50)-GSTG-R2betaB-GSSG-(OpLuc 51-169) (SEQ ID NO:190)
pBF B398	Met-(OpLuc 1-50)-GSSGGSGGSG-R2betaB-GSSGGSGGSG-(OpLuc 51-169) (SEQ ID NO:191)
pBF B399	Met-(OpLuc 1-50)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(OpLuc 51-169) (SEQ ID NO:192)
pBF B400	Met-(OpLuc 1-84)-GSTG-R2betaB-GSSG-(OpLuc 85-169) (SEQ ID NO:193)
pBF B401	Met-(OpLuc 1-84)-GSSGGSGGSG-R2betaB-GSSGGSGGSG-(OpLuc 85-169) (SEQ ID NO:194)
pBF B402	Met-(OpLuc 1-84)-GSSGGSGGSGGGSGGSGGSG-R2betaB-GSGGSGGSGGTSGGSGGSSG-(OpLuc 85-169) (SEQ ID NO:195)

Residue '1' in the above table indicates the first residue in the mature form of the protein (lacking the signal peptide for secretion, residue 28 in Genbank AB030246); RIIbetaB = residues 266-414 of human PKA regulatory subunit type II beta (Genbank BC075800).

Protein was expressed from these constructs using the TnT T7 Coupled Reticulocyte Lysate System. Following expression, 9 µL of TNT reaction was mixed with 1µL 1mM cAMP stock or H_2O, and the reactions were allowed to incubate at room temperature for approximately 15 minutes. Following incubation, 10 µL of 2× buffer (300 mM HEPES, pH = 8.0, 200 mM thiourea) was added to each reaction, and luminescence was measured from the resulting 20µL of solution

following addition of 100 µL of Renilla Assay Reagent using a Turner 20/20N luminometer (1 second integration time). The results are listed in the following table:

Table 8

pBFB397 +	918
pBFB397 -	225
pBFB398 +	4,917
pBFB398 -	291
pBFB399 +	38,051
pBFB399 -	356
pBFB400 +	10,369
pBFB400 -	6,387
pBFB401 +	6,124
pBFB401 -	2,304
pBFB402 +	62,264
pBFB402 -	8,568
FL Opluc +	25,225,870
FL Opluc -	23,231,428
No DNA +	120
No DNA -	116

FL Opluc = expression of residues 28-169 of Genbank BC075800; '+' = addition of exogenous cAMP to 50 µM final concentration; '-' = no exogenous cAMP was added.

Other vectors include circularly permutated mutants of *Oplophorus* luciferase (OpLuc) with a RIIβB domain cloned into a pF4K-CMV plasmid to enable expression under T7 and CMV promoters. Various circularly permuted *Oplophorus* luciferase (OpLuc) constructs having RIIβB inserted into sites tolerant to modification were also prepared (CPM OpLuc/RIIβB). See Figure 63. Numbers in brackets correspond to amino acid residues in the mature form of *Oplophorus* luciferase. The integers "4", "10" and "20" indicate the linkers of corresponding

length. Note that Met and Val residues were added to N-terminus of the luciferase. Therefore, the position of each split in circularly permutated mutants is shifted for two amino acid residues. For example, the split marker "50-51" (referred to the residue order in the native mature form of the enzyme) occurred in-between residues 52 and 53 in the actual luciferase version used.

pF4K-CMV-[51-169]-4-RIIβB-4-[1-50]-OpLuc

pF4K-CMV-[51-169]-10-RIIβB-10-[1-50]-OpLuc

pF4K-CMV-[51-169]-20-RIIβB-20-[1-50]-OpLuc

pF4K-CMV-[85-169]-4-RIIβB-4-[1-84]-OpLuc

pF4K-CMV-[85-169]-10-RIIβB-10-[1-84]-OpLuc

pF4K-CMV-[85-169]-20-RIIβB-20-[1-84]-OpLuc

pF4K-CMV-[113-169]-4-RIIβB-4-[1-112]-OpLuc

pF4K-CMV-[113-169]-10-RIIβB-10-[1-112]-OpLuc

pF4K-CMV-[113-169]-20-RIIβB-20-[1-112]-OpLuc

pF4K-CMV-[135-169]-4-RIIβB-4-[1-134]-OpLuc

pF4K-CMV-[135-169]-10-RIIβB-10-[1-134]-OpLuc

pF4K-CMV-[135-169]-20-RIIβB-20-[1-134]-OpLuc

pJ15:4809-OgLuc- 2.7 kb plasmid with cloned full-size *Oplophorus* luciferase ORF (by DNA 2.0)

pJ15:4810-2.6 kb plasmid with the ORF of the mature *Oplophorus* luciferase ORF (27 aa signal peptide was deleted) (by DNA 2.0)

pF1K-OgLucS-3.7 kb. The full-size luciferase ORF was cloned into pF1K (FL OpLuc)

pF1K-OgLuc-3.6 kb. ORF of the mature luciferase was cloned into pF1K

pF1K-OpLucDN-3.6 kb. Identical to pF1K-OgLuc except that first four N-terminal residues were deleted

pF1K-OpLucDC-3.6 kb. Identical to pF1K-OgLuc except that last three C-terminal residues were deleted

pF1K-OpLucDNDC-3.6 kb. Identical to pF1K-OgLuc except that first four N-terminal and last three C-terminal residues were deleted

pFVDnK-OgLucS-4.4 kb. HaloTag was fused with the full-size luciferase ORF

pFVDnK-OgLuc-4.5 kb. HaloTag was fused with ORF of the mature luciferase

pFN6K-opLuc-3.6 kb. HQ-tag was introduced into N-terminus of ORF of the mature luciferase

Equal amounts of CPM OpLuc/RIIβB constructs (0.1 μg of plasmid per 50 μl of reaction mixture; Figure 64) were expressed in a rabbit reticulocyte TnT system (Promega #L1170). After the TnT reactions were complete, cAMP was added to the final concentration of 0.1 mM and the mixtures were additionally incubated at room temperature for 15 minutes. The reactions were diluted ten fold with Renilla lysis buffer and luciferase activity was measured in Renilla reagent as recommended (Renilla Luciferase Assay System, #E2810, Promega Corp.).

Induction of luciferase activity was observed with all four circularly permutated luciferase constructs (Figure 64). The construct with the luciferase split between residues 84 and 85 demonstrated the highest induction (about 250 fold). The 20 amino acid linker supported the most efficient folding.

The results indicate that cAMP biosensors can be generated by either circular permutation or direct insertion of RIIβB into any of the above chosen insertion sites.

Example XXIII

Protein Complementation with *Oplophorus* Luciferase

To determine sites in *Oplophorus* luciferase useful for protein complementation, N- and C-terminal fusions were prepared. Vector backbones included pF3A for *in vitro* experiments and pF5K for in cell experiments. The following constructs were prepared: "N term–FRB", i.e., OpLuc (1-50 or 1-84) 10 aa G/S linker–FRB, "FKBP – C term", i.e., FKBP–(G4S)2 linker–OpLuc (51-170 or 85-170), "FRB-N term," i.e., FRB –(G4S)2 linker–OpLuc (1-50 or 1-84), and "C term – FKBP," i.e., OpLuc (51-170 or 85-170) - 10aa G/S linker – FKBP. See table below.

Construct	Vector	Type	Description	Figure legend
201518.54.06	pF5K	Full length	FL-OpLuc	FL OpLuc
201518.57.E6	pF5K	FRB-N term	FRB-OpLuc (1-50)	FRB-50
201518.57.G3	pF5K	FRB-N term	FRB-OpLuc (1-84)	FRB-84
201518.101.04	pF5K	FKBP – C term	FKBP-OpLuc (51-170)	FKBP-51
201518.57.H12	pF5K	FKBP – C term	FKBP-OpLuc (85-170)	FKBP-85
pBFB395	pF5K	N term–FRB	OpLuc (1-50) - FRB	50-FRB
pBFB396	pF5K	N term–FRB	OpLuc (1-84) - FRB	84-FRB
pBFB415	pF5K	C term – FKBP	OpLuc (51-170) - FKBP	51-FKBP
pBFB416	pF5K	C term – FKBP	OpLuc (85-170) - FKBP	85-FKBP
201518.45.08	pF3A	Full length	FL-OpLuc	FL OpLuc
201518.57.A2	pF3A	FRB-N term	FRB-OpLuc (1-50)	FRB-50
201518.57.A11	pF3A	FRB-N term	FRB-OpLuc (1-84)	FRB-84
201518.57.D9	pF3A	FKBP – C term	FKBP-OpLuc (51-170)	FKBP-51
201518.61.H3	pF3A	FKBP – C term	FKBP-OpLuc (85-170)	FKBP-85
201518.110.4-1	pF3A	N term–FRB	OpLuc (1-50)-FRB	50-FRB
201518.104.04	pF3A	N term–FRB	OpLuc (1-84)-FRB	84-FRB
201518.129.03	pF3A	C term – FKBP	OpLuc (51-170) - FKBP	51-FKBP
201518.129.06	pF3A	C term – FKBP	OpLuc (85-170) - FKBP	85-FKBP

Proteins were either singly expressed or co-expressed using the TnT® SP6 High-Yield Protein Expression System at 30 °C for 2 hours (as per the manufacturer's protocol; Promega Corp.). Twenty µL lysate was incubated +/- 1 µM rapamycin for 15 minutes at room temperature. Ten µL lysate was diluted 1:1 in 2× HEPES/thiourea and 5 µL was placed in a 96-well plate well, in triplicate. Luminescence was measured by addition of 100 µL Renilla Luciferase Assay Reagent (R-LAR) by injectors. The in vitro results for a split at positions 50/51 (50-FRB + FKBP-51) are shown in Figure 56 and those for 84/85 (84-FRB + FKBP-85) are shown in Figure 58. Figures 57 and 59 show the results for the respective SDS-PAGE analyses. Five µL -/+ rapamycin lysate was size fractionated on 4-12% SDS-PAGE. Figures 61-62 show in vitro results for the 51-FKBP + FRB-50 and 85-FKBP + FRB-85 orientations. For the data in Figure 62, 7.5 µL +/- rapamycin lysate was size fractionated on 4-12% SDS-PAGE.

Figure 60 shows the in-cell results using HEK-293 cells. HEK-293 cells were transiently transfected with complimenting fragments or with the individual fragments of *Oplophorus* luciferase in a 6-well plate and incubated overnight. The next day cells were trypsinized and plated in a 96-well plate at 20,000 cells per well. At the same time 1 µM rapamycin or vehicle (DMSO) was added to the cells and

they were allowed to recover overnight at 37°C with 5% CO_2. The next day media was removed and 20 µL of 1 x Renilla Luciferase Assay Lysis buffer was added to each sample and the plate was shaken for 15 minutes at 500 rpm. 100 µL of Renilla Luciferase Assay Reagent was injected into each well and samples were measured for 3 sec/well with a 0.5 second delay.

Example XXIV

The CPM firefly luciferase (FF Luc) and *Renilla* luciferase (hRL Luc) were also used as biosensors to assay kinase/phosphatase activities. In a manner similar to previous biosensors of cAMP, cGMP, and calcium, various circularly permuted (CPM) FF Luc and hRL Luc constructs were made to detect phosphorylation by tyrosine or serine/threonine kinases (phosphorylation on the underlined Tyr or Thr residues, respectively, in the constructs described below). The conformational change, caused by the binding of the phosphorylated peptide sequence with the tethered phosphopeptide recognition domain, may cause a modulation of the fused biosensor luciferase activity. This represents a novel class of reagents able to measure the activity of kinases, perhaps with enhanced performance characteristics relative to existing FRET-based biosensors.

The peptide sequences and recognition domains used for the tyrosine kinase and serine/threonine kinase were, respectively: peptide GSTSGSGKPGSGEGSEIYGEF (SEQ ID NO:295) or EIYGEF (SEQ ID NO:296) with phosphopeptide recognition domain human Src SH2 domain (Genbank NM_005417; aa residues 151-248) and RKRDRLGTLGI (SEQ ID NO:297) with phosphopeptide recognition domain FHA2 from Rad53p (codon optimized version of the nucleic acid sequence Genbank accession # AY693009 which aligns to bases 1717- 2186; aa residues 573-730 of accession #AAT93028).

The multiple sites for CPM that were previously identified as functional for generating biosensors in FF Luc and hRL Luc were used for the construction of kinase biosensors. These constructs were either made using PCR products ligated into unique restriction sites or Splicing by Overlapping Extension PCR (SOE-PCR). The FF Luc constructs were made in the pF9A backbone and the hRL Luc

constructs were made in the pF5A backbone, except for plasmids pBFB174, 175, 176, 178, 179, 180, 181, 182, 228, 229 and 230 which were made in the modified pGL4.74 backbone described in Example II.

The following constructs were made: Met-(Luc2.0 or hRL C-terminal fragment)-(Linker X)-(peptide phosphorylated by kinase)-(linker)-(phosphopeptide recognition domain)-(Linker Y)-(Luc2.0 or hRL N-terminal fragment)-Val. Constructs were also made in which the order of the peptide phosphorylated by the kinase and the phosphopeptide recognition domain were switched. In addition, the following constructs were made for the tyrosine kinase FF Luc biosensor: Met-(short peptide phosphorylated by kinase)-(linker X)-(Luc2.0)-(linker)-(phosphopeptide recognition domain)-Val. See Figure 65.

Tyrosine Kinase Constructs

1) Met-(Luc2.0 234-544)-GSSG-(human Src SH2 domain)-GSG-GSTSGSGKPGSGEGSEIYGEF-(Linker Y)-(Luc2.0 4-233)-Val, where Y = GSGGSGGSSG (SEQ ID NO:291), or GSGGSGGSGGGSGGSGGSSG (SEQ ID NO:286). (GSSG corresponds to SEQ ID NO:270; GSGGSTSGSGKPGSGEGSEIYGEF corresponds to SEQ ID NO:298). Clones pBFB180, 181, 182, 365, 366, 367.

2) Met-EIYGEF-(Linker X)-(Luc2.0 4-544)-GSSG-(human Src SH2 domain), where X = GSSG (SEQ ID NO:270), GSSGGSGGSG (SEQ ID NO:276), or GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277). (EIYGEF corresponds to SEQ ID NO:296). Clones pBFB174, 175, 176.

3) Met-(hRL 92-311)-GSG-(human Src SH2 domain)-GSG-GSTSGSGKPGSGEGSEIYGEF-(Linker X)-GSSG-(hRL 2-91)-Val, where X = GSSG (SEQ ID NO:270), GSGGSGGSSG (SEQ ID NO:291), or GSGGSGGSGGGSGGSGGSSG (SEQ ID NO:286). (GSGGSTSGSGKPGSGEGSEIYGEF corresponds to SEQ ID NO:298). Clones pBFB228, 229, 230.

4) Met-(Luc2.0 A-544)-(Linker X)-(human Src SH2 domain)-GSTSGSGKPGSGEGSEIYGEF-(Linker Y)-(Luc2.0 4-B)-Val, where X = GSTG (SEQ ID NO:275), GSSGGSGGSG (SEQ ID NO:276), or

GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277) and Y = GSSG (SEQ ID NO:270), GSGSGGSGGSSG (SEQ ID NO:299), or GSGGSGGSGGGSGGSGGSSG (SEQ ID NO:286). (GSTSGSGKPGSGEGSEIYGEF corresponds to SEQ ID NO:295). CPM sites [A, B] = [235, 233], [359, 355], [84, 82], [309, 307]. Clones pBFB368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379.

5) Met-(hRL A-311)-(Linker X)-(human Src SH2 domain)-GSTSGSGKPGSGEGSEIYGEF-(Linker Y)-(hRL 3-B)-Val, where X = GSSG (SEQ ID NO:270), GSSGGSGGSG (SEQ ID NO:276), or GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277) and Y = GSSG (SEQ ID NO:270), GSGSGGSGGSSG (SEQ ID NO:299), or GSGGSGGSGGGSGGSGGSSG (SEQ ID NO:286). (GSTSGSGKPGSGEGSEIYGEF corresponds to SEQ ID NO:295). CPM sites [A, B] = [92, 91], [42, 41], [111, 110], [31, 30], [69, 68]. Clones pBFB380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394.

Serine/Threonine Kinase Constructs

1) Met-(Luc2.0A-544)-(linker X)- RKRDRLGTLGI-(GGSSGGGSGGGGSGG)-(Rad53p FHA2 domain)-(linker Y)-(Luc2.04-B), where X = GSSG (SEQ ID NO:270), GGSGGSGSSG (SEQ ID NO:300), or GSSGGSGGSGGGSGGSGSSG (SEQ ID NO:301), Y = GSSG (SEQ ID NO:270), GSGGSGGSGG (SEQ ID NO:281), or GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:278). (RKRDRLGTLGIGGSSGGGSGGGGSGG corresponds to SEQ ID NO:283) CPM sites were [A, B] = [235, 233], [359, 355], [84, 82], [309, 307]. Clones pBFB335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346.

2) Met-(hRL A-311)-(linker X)- RKRDRLGTLGI-(GGSSGGGSGGGGSGG)-(Rad53p FHA2 domain)-(linker Y)-(hRL 3-B), where X = GSSG (SEQ ID NO:270), GSSGGSGGSGGG (SEQ ID NO:302), or GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277), Y = GSSG (SEQ ID NO:270), GSGGSGGSSG (SEQ ID NO:291), or GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:278). (RKRDRLGTLGIGGSSGGGSGGGGSGG corresponds to SEQ ID NO:283). CPM sites were [A, B] = [92, 91], [42, 41], [111, 110], [31, 30], [69, 68].

Clones pBFB350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364.

3) Met-(Luc2.0A-544)-(linker X)- (Rad53p FHA2 domain)-GGSSG-RKRDRLGTLGI-(linker Y)-(Luc2.04-B), where X = GSGG (SEQ ID NO:293), GGSGGGGSGG (SEQ ID NO:294), or GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277), Y = GGSSG (SEQ ID NO:304), GSSGSGGSGG (SEQ ID NO:305), or GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:278). (GGSSGRKRDRLGTLGI corresponds to SEQ ID NO:303). CPM sites were [A, B] = [235, 233], [359, 355]. Clones pBFB417, 418, 419, 420, 421, 422.

4) Met-(hRL A-311)-(linker X)- (Rad53p FHA2 domain)-GGSSG-RKRDRLGTLGI-(linker Y)-(hRL 3-B), where X = GSGG (SEQ ID NO:293), GGSGGGGSGG (SEQ ID NO:294), or GSSGGSGGSGGGSGGSGGSG (SEQ ID NO:277), Y = GGSSG (SEQ ID NO:304), GSSGSGGSGG (SEQ ID NO:305), or GSGGSGGSGGTSGGSGGSSG (SEQ ID NO:278). (GGSSGRKRDRLGTLGI corresponds to SEQ ID NO:303). CPM sites were [A, B] = [42, 41], [111, 110]. Clones pBFB423, 424, 425, 426, 427, 428.

In vitro testing of a subset of Serine/Threonine Kinase sensors

Constructs pBFB335, 336, 338, 339, 340, 417, 418, 419, 422, 22 and 8 were tested in vitro using TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components were assembled according to the manufacturer's recommended protocol:

1 µg plasmid DNA
25 µL Rabbit Retic Extract
2 µL TNT reaction buffer
1 µL T7 polymerase
1 µL amino acid mixture
1 µL rRNasin
dH_2O to 50 µL total volume

Following incubation at 30°C for 1 hour, the respective fusion proteins were incubated in the presence or absence of 10 ng Akt1/PKB alpha recombinant enzyme (Upstate Biotechnology) by combining 2 µL of TNT® reaction with 8 µL water + 4

μL 5x Reaction Buffer (40 mM MOPS/NaOH pH 7.0, 1mM EDTA) + 4 μL 5x Mg-ATP (50 mM Mg acetate, 0.5 mM ATP) + 2 μL 5 ng/ μL enzyme (diluted from 100 ng/ul stock diluted in PKB dilution buffer [20 mM MOPS (7.0), 1mM EDTA, 5% glycerol, 0.05% DTT, 1 mg/ml BSA]) or 2 μL PKB dilution buffer only. Samples were then incubated at 30°C for 20 minutes. Five μL of sample was added to 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.) solution and pipetted up and down 4X rapidly to mix. Luminescence was measured using a Turner 20/20N luminometer (Turner Biosystems; 1 second integration time). All samples were measured in triplicate.

Results

Construct pBFB340 showed a 50% decrease in luminescence plus Akt1/PKB as compared to no Akt1/PKB. The control constructs pBFB22 and pBFB8 did not change with Akt1/PKB addition (Figure 66).

The protocol for other Serine/Threonine kinase sensors is identical to the one above except that for the CPM hRL Luc samples, 5 μL of sample are added to a 96 well plate + 5 μL 2x Renilla lysis buffer without detergents (150mM HEPES, 100mM Thiourea) and 100 μL Renilla Assay Reagent (Promega Corp.) is added by injectors using a Veritas Microplate Luminometer and luminescence is measured (Turner Biosystems; Bright-Glo program; 3 second integration time). FF Luc samples are measured by adding 100 μL of Luciferase Assay Reagent (LAR; Promega Corp.), to 5 μL of sample in a 96 well pate, by injectors using a Veritas Microplate Luminometer and luminescence measured (Turner Biosystems; Bright-Glo program; 3 second integration time).

The tyrosine kinase sensors are tested as follows: Proteins are expressed in *vitro* using TNT® T7 Coupled Reticulocyte Lysate System (Promega Corp.). Briefly, the following components are assembled according to the manufacturer's recommended protocol:

1 μg plasmid DNA
25 μL Rabbit Retic Extract
2 μL TNT reaction buffer
1 μL T7 polymerase

1 µL amino acid mixture
1 µL rRNasin
dH_2O to 50 µL total volume

Following incubation at 30 °C for 1 hour, the respective fusion proteins are used in 50 µl kinase reactions as follows: 1X ProFlour reaction buffer (Promega Corp.) + 10 µl RR TnT reaction + 100 µM sodium vanadate + 1 mM $MnCl_2$ + 1 mM MgATP + 0.5 µl c-Src Kinase or water. At 0, 30 and 60 minutes after addition of Src Kinase, 10 µl aliquots are taken and stored at -20 °C until assayed. For the CPM FF Luc samples, 5 µl is transferred to a 96 well plate and 100 ul Luciferase Assay Reagent (LAR; Promega Corp.) is added by injectors using a Veritas Microplate Luminometer and luminescence will be measured (Turner Biosystems; Bright-Glo program; 3 second integration time). For the CPM hRL Luc samples, 5 µL of sample is added to a 96 well plate + 5 µL 2x Renilla lysis buffer without detergents (150mM HEPES, 100mM Thiourea) and 100 µL Renilla Assay Reagent (Promega Corp.) is added by injectors using a Veritas Microplate Luminometer and luminescence was measured (Turner Biosystems; Bright-Glo program; 3 second integration time).

To test kinase sensors in cells, the FF Luc and CPM hRL Luc serine/threonine kinase biosensors are tested as follows: HEK293 and NIH /3T3 cells are plated in 96 well plates at a cell density of 1-1.5 x 10^4 cells per well in CO_2-independent media (Invitrogen) + 10% FBS. They are then transfected with *Trans*It®-LT1 Reagent (MIRUS) using 4.2 µL *Trans*It®-LT1 reagent and 1.4 µg DNA per well. Cells are allowed to grow overnight at 37°C / 10% CO_2. The next day the media is changed to CO_2-independent media + 0.2% FBS to serum-starve the cells. The cells are then allowed to grow overnight at 37 °C/10% CO_2. Approximately 2 days after transfection, the cells are equilibrated with a final concentration of 5 mM Luciferin-EF (Promega Corp.) for the FF Luc sensors or 60 µM EnduRen (Promega Corp.) for the CPM hRL Luc sensors. All cells are allowed to equilibrate for 1.5 hours at 37 °C/ 10%CO_2. After 1.5 hours, baseline measurements of luminescence are measured using a Mithras LB 940 Luminometer (Berthold Technologies; integration time of 1 second per well) at 37 °C. Next, half

of the cells are treated with a kinase activator such as Platelet-Derived Growth Factor (PDGF, 50 ng/ml final concentration). Luminescence will then be measured continuously for the next 30 minutes at 37 °C.

Example XXV

Determination of suitable split points for creating circularly permuted proteins in the absence of three-dimensional protein structure information.

Method

1) Obtain the amino acid sequence of the protein of interest.
2) Use one or more computer programs to predict protein structure features that aid in the determination of suitable split points. Suitable split points are likely exposed on the protein surface. Split points that lie outside of regular secondary structure elements such as helices and sheets are less likely to disrupt protein structure and function.

   Predict surface exposed protein regions: exposed regions are likely to be hydrophilic. The distribution of hydrophilic and hydrophobic residues along a protein sequence (hydrophobicity plot/score) can be computed based on commonly used hydrophobicity scales using programs available at open access websites (e.g. ProtScale from the ExPASy proteomics server of the Swiss Institute of Bioinformatics http://www.expasy.ch/cgi-bin/protscale.pl) and as part of commercial sequence analysis packages (e.g. Lasergene from DNASTAR).

   Predict protein secondary structure: such programs are available at open access websites (see list on ExPASy Proteomics Tools website http://www.expasy.org/tools/#secondary) and as part of commercial sequence analysis packages (e.g. Lasergene from DNASTAR).
3) Select split points based on the results from one or more prediction methods.

Example

1) Protein sequence: *Oplophorus gracilorostris* mature luciferase sequence (Genbank accession BAB13776, residues 28-196).

2) **Predict surface exposed protein regions**: calculate per-residue hydrophobicity score based on the Kyte-Doolittle hydrophobicity scale using window sizes of 5 and 7, which specify the range recommended for finding putative surface-exposed regions (Kyte J and Doolittle RF: A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105, 1982).

Predict protein secondary structure: use five different prediction algorithms:

a. PSIPRED (Jones DT. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 292: 195-202. McGuffin LJ, Bryson K, Jones DT).

b. JPRED (Cuff JA, Clamp ME, Siddiqui AS, Finlay M and Barton GJ. 1998. Jpred: A Consensus Secondary Structure Prediction Server, Bioinformatics 14:892-893).

c. PORTER (G Pollastri, A McLysaght. "Porter: a new, accurate server for protein secondary structure prediction". Bioinformatics, 21(8),1719-20, 2005).

d. SCRATCH (G Pollastri, D Przybylski, B Rost, P Baldi: Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. Proteins, 47, 228-335, 2002).

e. PROF (M Ouali, R King: Cascaded multiple classifiers for secondary structure prediction. Protein Science, 9, 1162-1176, 1999).

3) Compile results of protein structure feature predictions in a table for comparison. Select suitable split points in areas that are hydrophilic (low hydrophobicity score) and lie outside of predicted regular secondary structure elements (helices and sheets). See Table 9 (in three sections below).

Table 9: Compiled structure feature prediction results for *Oplophorus gracilorostris* mature luciferase. Secondary structure prediction results code is H=helix, E=sheet, C=coil, blank=coil. Hydrophobicity prediction score is >0 for hydrophobic and <0 for hydrophilic areas. Suitable split point examples are marked xxx in rightmost column.

Seq #	Mature Seq	PSIPRED	JPRED	PORTER	SCRATCH	PROF	Hydrophobicity K-D		
							win=7	win=5	
1	F	C		C	C	C			
2	T	C		C	C	E			
3	L	H		C	H	E		0.84	
4	A	H		H	H	E		0.84	
5	D	H		H	C	E	0.81	1.82	
6	F	H		C	C	E	0.40	0.98	
7	V	H		C	C	C	0.09	-0.08	
8	G	C		C	C	C	-0.72	0.44	
9	D	C		C	C	C	-1.00	-0.82	
10	W	H		H	H	C	-0.41	-2.36	
11	Q	H		H	H	H	-0.77	-2.42	
12	Q	H		H	H	C	-1.38	-1.36	
13	T	H		H	H	C	-1.72	-1.26	
14	A	C		C	C	C	-1.72	-0.82	
15	G	C		C	C	C	-2.01	-0.82	
16	Y	C		C	C	C	-2.01	-1.38	
17	N	C		C	C	C	-1.16	-2.44	
18	Q	H		H	H	C	-0.66	-3.06	
19	D	H	H	H	H	E	-1.24	-1.96	
20	Q	H	H	H	H	E	-1.59	-0.5	
21	V	H	H	H	H	E	-1.49	-0.5	
22	L	H	H	H	H	E	-1.14	-0.5	
23	E	H	H	H	H	E	-0.33	0.12	
24	Q	C	H	C	C	C	-0.03	-0.8	
25	G	C		C	C	C	0.27	-0.8	
26	G	C		C	C	C	0.22	-0.26	
27	L	H		H	H	C	0.11	0.28	
28	S	H	H	H	H	H	0.11	1.12	
29	S	H	H	H	H	H	0.70	1.76	
30	L	H	H	H	H	H	1.17	0.3	
31	F	H	H	H	H	H	1.17	0.82	
32	Q	H	H	H	H	H	1.21	1.74	
33	A	H	H	H	H	H	1.21	0.9	
34	L	H	H	H	H	H	1.77	1.18	
35	G	C		C	C	C	1.27	1.72	
36	V	C		C	C	C	0.78	2.2	
37	S	C		C	C	E	1.67	1.3	
38	V	C		C	C	E	1.08	1.06	
39	T	C		C	C	E	0.22	1.12	
40	P	C		C	C	C	0.73	0.58	
41	I	H	E	C	C	C	0.73	-1.04	
42	Q	H	E	E	E	E	1.24	-0.06	
43	K	H	E	E	E	E	0.69	1.1	
44	V	H	E	E	E	E	0.72	0.96	
45	V	E	E	E	E	E	0.51	1.5	
46	L	E	E	E	E	E	-0.38	2.2	
47	S	C		C	C	E	-0.03	0.66	
48	G	C		C	C	C	0.82	-0.88	
49	E	C		C	C	C	-0.08	-1.72	
50	N	C		C	C	C	-0.34	-0.8	
51	G	C		C	C	C	-1.16	-1.5	xxx
52	L	C		C	C	C	-0.57	-0.44	xxx
53	K	E	E	C	C	C	-0.88	-0.44	
54	A	E		C	C	C	-0.02	0.54	
55	D	E	E	E	C	E	0.87	-0.86	
56	I	E	E	E	E	E	1.41	0.76	
57	H	E	E	E	E	E	0.81	1.3	
58	V	E	E	E	E	E	1.10	2.9	
59	I	E	E	E	E	E	0.51	1.68	
60	I	E	E	E	E	E	0.86	2.06	

61	P	E	E	E	E	E	0.78	0.52	
62	Y	C		C	C	E	1.04	-0.46	
63	E	C		C	C	C	0.53	-0.6	
64	G	C		C	C	C	0.34	-0.44	
65	L	C		C	C	C	-0.54	-0.26	
66	S	C		C	C	C	-0.16	1	
67	G	C		H	H	C	-0.06	0.38	
68	F	H	H	H	H	C	0.76	0	
69	O	H	H	H	H	C	1.30	0.08	
70	M	H	H	H	H	C	0.49	0.92	
71	G	H	H	H	H	C	0.79	1.26	
72	L	H	H	H	H	H	1.33	1.26	
73	I	H	H	H	H	H	1.33	1.26	
74	E	H	H	E	H	E	1.29	2.24	
75	M	H	H	E	H	E	1.54	2.04	
76	I	H	E	E	H	E	2.06	0.36	
77	F	H	E	E	H	E	1.49	1.9	
78	K	H	E	E	E	E	0.81	2.36	
79	V	H	E	E	E	E	1.67	1.2	
80	V	H	E	E	E	E	1.07	0.32	
81	Y	C	E	C	C	E	0.18	1.94	
82	P	C	E	C	C	E	-0.49	0.4	
83	V	C		C	C	C	-0.41	-1.14	1st choice
84	D	C		C	C	C	-0.57	-1.52	xxx
85	D	C		C	C	C	-1.17	-1.84	xxx
86	H	C		C	C	C	-0.82	-2.12	
87	H	C		C	C	E	-0.14	-2.2	
88	F	E	E	E	E	E	-0.19	-0.6	
89	K	E	E	E	E	E	-0.16	0.94	
90	I	E	E	E	E	E	0.09	2.34	
91	I	E	E	E	E	E	0.40	1.14	
92	L	E	E	E	E	E	0.68	1.66	
93	H	E	E	E	E	E	0.79	0.68	
94	Y	E		E	C	C	1.69	-0.36	
95	G	C		C	C	C	1.69	-0.36	
96	T	E	E	E	E	E	0.80	1.12	
97	L	E	E	E	E	E	0.33	2.28	
98	V	E	E	E	E	E	1.16	1.66	
99	I	E	E	E	E	E	1.22	1.72	
100	D	C	E	C	C	E	1.09	1.8	
101	G	C		C	C	C	0.78	0.82	
102	V	C		C	C	C	0.57	-0.4	
103	T	C		C	C	C	0.60	-0.4	
104	P	C		C	C	C	-0.29	0.06	
105	N	H		C	C	C	-0.04	0.12	
106	M	H		H	H	C	0.31	-0.44	
107	I	H		H	H	H	-0.20	-0.38	
108	D	H		H	H	E	-0.62	0.88	
109	Y	H		C	H	C	-0.62	0.42	
110	F	C		C	C	C	-0.38	-1.38	
111	G	C		C	C	C	-0.77	-1	
112	R	C		C	C	C	-1.31	-1	xxx
113	P	C		C	C	C	-0.42	-1.88	xxx
114	Y	C		C	C	C	-0.08	-1.88	
115	P	C		C	C	C	0.08	-0.08	
116	G	C		C	C	C	0.43	0.6	
117	I	E	E	E	E	E	0.54	1.7	
118	A	E	E	E	E	E	0.68	2.58	
119	V	E	E	E	E	E	0.39	1.96	
120	F	E	E	E	E	E	0.18	0.98	

121	D	C		C	C	C	0.72	-0.16	
122	G	C		C	C	C	0.14	-1.7	
123	K	C		C	C	C	0.41	-1.36	
124	Q	E	E	E	E	E	-0.13	-0.8	
125	I	E	E	E	E	E	-0.49	0.12	
126	T	E	E	E	E	E	-0.18	0.76	
127	V	E	E	E	E	E	0.29	1.38	
128	T	E	E	E	E	E	0.62	0.34	
129	G	E	E	E	E	E	0.62	1.24	
130	T	E	E	E	E	E	0.08	0.22	
131	L	E	E	E	C	E	-0.23	-0.34	
132	W	E		C	C	C	-1.13	-0.34	
133	N	C		C	C	C	-0.56	-0.9	
134	G	C		C	C	C	-0.66	-2.44	
135	N	C		C	C	C	-0.97	-1.36	
136	K	E		C	C	E	-1.78	-0.92	
137	I	E	E	E	E	E	-2.18	-1.54	
138	Y	E		E	E	E	-1.37	-1.54	
139	D	C		E	C	E	-0.82	-1.66	
140	E	C	E	E	C	E	-0.82	-1.8	
141	R	E	E	E	E	E	-0.57	-0.64	
142	L	E	E	E	C	E	-1.46	-0.64	
143	I	C		C	E	E	-1.36	-0.26	
144	N	C		C	C	C	-1.06	-0.06	
145	P	C		C	C	C	-0.24	-0.9	
146	D	C		C	C	C	0.68	-1.96	
147	G	C		C	C	C	0.57	-0.5	
148	S	C		C	C	E	-0.43	0.58	
149	L	E	E	E	E	E	0.42	1.84	
150	L	E	E	E	E	E	0.52	1.02	
151	F	E	E	E	E	E	1.41	2.02	
152	R	E	E	E	E	E	1.07	1.12	
153	V	E	E	E	E	E	1.11	1.26	
154	T	E	E	E	E	E	1.16	0	
155	I	E	E	E	E	E	0.66	0.82	
156	N	C		C	C	C	0.30	0.82	
157	G	C		C	C	C	0.70	0.82	
158	V	C		C	C	C	-0.27	-0.16	
159	T	C		C	C	C	0.23	0.36	
160	G	H		H	H	C	0.01	-0.46	
161	W	H	H	H	H	C	0.01	-0.54	
162	R	H	H	H	H	H	-0.33	0.1	
163	L	H	H	H	H	H	-0.30	-0.52	
164	C	H	H	H	H	H	0.20	-1.04	
165	E	H	H	H	H	H	0.44	0.76	
166	N	H	H	H	H	H		0.76	
167	I	H		H	H	H		0.62	
168	L	C		C	H	C			
169	A	C		C	C	C			

Thus, split sites for any protein, e.g., one to be used in PCA or one to be used as a biosensor (insertion of domain directly in-between the split sites or into a circular permuted mutant, circularly permuted at the split sites), in the absence of a three-dimensional structure, can be selected.

References

Altschul et al., J. Mol. Biol., 215:403 (1990).

Altschul et al., Nuc. Acids Res., 25:3389 (1977).

Bos, Nat. Rev. Mol. Cell. Biol., 4:733 (2003)

Chong et al., Gene, 192:271 (1997).

Corpet et al., Nucl. Acids Res., 16:1088 (1988).

Dremier et. al., FEBS Lett., 546:163 (2003).

Geysen et al., Proc. Natl. Acad. Sci. USA, 3998 (1984).

Hanks and Hunter, FASEB J, 9:576-595 (1995).

Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA, 89:10915 (1989).

Higgins et al., Gene, 73:237 (1988).

Higgins et al., CABIOS, 5:157 (1989).

Huang et al., CABIOS, 8:155 (1992).

Kaihara et al., Anal. Chem., 75:41 (2003).

Karlin & Altschul, Proc. Natl. Acad. Sci. USA, 90:5873 (1993).

Lee et al., Anal. Biochem., 316:162 (2003).

Liu et al., Gene, 237:153 (1999).

Mayer and Baltimore, Trends Cell. Biol., 3:8 (1993).

Merrifield, J. Am. Chem. Soc., 2149 (1963).

Murray et al., Nucl. Acids Res., 17:477 (1989)

Myers and Miller, LABIOS, 4:11 (1988).

Ozawa et al, Analytical Chemistry, 73:2516 (2001).

Needleman et al., J. Mol. Biol., 48:443 (1976).

Paulmurugan et al., Anal. Chem., 75:1584 (2003).

Paulmurugan et al., Proc. Natl. Acad. Sci. USA, 99:3105 (2002).

Paulmurugan et al., PNAS, 24:15603 (1999).

Pearson et al., Methods Mol. Biol., 24:307 (1994).

Pearson et al., Proc. Natl. Acad. Sci. USA, 85:2444 (1985.

Plainkum et al., Nat. Struct. Biol., 10:115 (2003).

Remy et al., Biotechniques, 38:763 (2005).

Remy et al., Nat. Methods, 3:977 (2006).

Sadowski, et al., Mol. Cell. Bio., 6:4396 (1986).

Sala-Newby et al., Biochem J., 279:727 (1991).

Sala-Newby et al., FEBS, 307:241 (1992).

Sambrook et al., In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor (1989).

Smith et al., Adv. Appl. Math., 2:482 (1981).

Stewart et al., Solid Phase Peptide Synthesis, 2 ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12).

Tannous et al., Mol. Thera., 11:435 (2005).

Wada et al., Nucl. Acids Res., 18:2367 (1990).

Wang et al., BBRC, 282:28 (2001).

Waud et al, BBA, 1292:89 (1996).

Zagotta et al., Nature, 425:730 (2003).

All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification, this invention has been described in relation to certain preferred embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details herein may be varied considerably without departing from the basic principles of the invention.

WHAT IS CLAIMED IS:

1. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a circularly permuted luciferase which comprises a cyclic nucleotide binding site, which luciferase is permuted at a site or in a region which in a corresponding nonpermuted luciferase is tolerant to modification, and wherein the activity of the circularly permuted luciferase is detectable.

2. The polynucleotide of claim 1 wherein the circularly permuted luciferase further comprises a tag of at least one amino acid at the N-terminus, C-terminus, or both.

3. The polynucleotide of claim 2 wherein the tag is a PEST sequence, a GST sequence, or a polyhistidine sequence.

4. The polynucleotide of claim 1 wherein the circularly permuted luciferase further comprises a deletion of luciferase sequences at sequences corresponding to the N-terminus and/or C-terminus of the nonpermuted luciferase.

5. The polynucleotide of claim 4 wherein the deletion at the N-terminus or the C-terminus is no more than 15 residues of luciferase sequence.

6. The polynucleotide of claim 1 or 5 wherein the cyclic nucleotide binding site is at sequences corresponding to the N-terminus and/or C-terminus of the nonpermuted luciferase.

7. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site is about 4 to about 200 amino acid residues in length.

8. The polynucleotide of claim 1 wherein the luciferase is a Coleopteran luciferase.

9. The polynucleotide of claim 1 wherein the luciferase is a click beetle luciferase or a firefly luciferase.

10. The polynucleotide of claim 1 wherein the luciferase is an anthozoan luciferase.

11. The polynucleotide of claim 10 wherein the luciferase is a *Renilla* luciferase.

12. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site is in a region corresponding to residue 2 to 12, residue 32 to 53, residue 70 to 88, residue 102 to 126, residue 139 to 165, residue 183 to 203, residue 220 to 247, residue 262 to 273, residue 303 to 313, residue 353 to 408, residue 485 to 495, or residue 535 to 546 of a firefly luciferase.

13. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site is in a region corresponding to residue 15 to 30, residue 112 to 122, residue 352 to 362, residue 371 to 384, residue 393 to 414, or residue 485 to 495 of a click beetle luciferase.

14. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site is in a region corresponding to residue 2 to 12, residue 26 to 47, residue 64 to 74, residue 86 to 116, residue 147 to 157, residue 223 to 234, or residue 301 to 311 of a *Renilla* luciferase.

15. The polynucleotide of claim 1 wherein the permutation is in a region corresponding to residue 2 to 12, residue 32 to 53, residue 70 to 88, residue 102 to 126, residue 139 to 165, residue 203 to193, residue 220 to 247,

residue 262 to 273, residue 303 to 313, residue 353 to 408, residue 485 to 495, or residue 535 to 546 of a firefly luciferase.

16. The polynucleotide of claim 1 wherein the permutation is in a region corresponding to residue 15 to 30, residue 112 to 122, residue 352 to 362, residue 371 to 384, residue 393 to 414, or residue 485 to 495 of a click beetle luciferase.

17. The polynucleotide of claim 1 wherein the permutation is in a region corresponding to residue 26 to 47, residue 64 to 74, residue 85 to 116, residue 147 to 157, residue 223 to 234, or residue 301 to 311 of a *Renilla* luciferase.

18. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site binds cAMP.

19. The polynucleotide of claim 1 wherein the cyclic nucleotide binding site binds cGMP.

20. A vector comprising the polynucleotide of any one of claims 1 to 19.

21. A host cell comprising the polynucleotide in the vector of claim 20.

22. A circularly permuted luciferase encoded by the polynucleotide of any one of claims 1 to 19.

23. A method to detect or determine a cyclic nucleotide in a cell, comprising:
a) providing a mixture comprising the host cell of claim 21 or a lysate thereof, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the mixture, thereby detecting or determining the presence or amount of cyclic nucleotide in the cell.

24. A method to detect or determine a cyclic nucleotide in a sample, comprising:
a) providing a mixture comprising a sample suspected of having cyclic nucleotide, the circularly permuted luciferase of claim 22, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the mixture.

25. The method of claim 23 or 24 wherein the luciferase is a firefly luciferase.

26. The method of claim 23 or 24 wherein the luciferase is a click beetle luciferase.

27. The method of claim 23 or 24 wherein the luciferase is a *Renilla* luciferase.

28. The method of claim 23 or 24 wherein binding of cyclic nucleotide to the cyclic nucleotide binding site results in an increase in luminescence.

29. The method of claim 24 wherein the sample comprises cells.

30. The method of claim 24 wherein the sample comprises a cell lysate or cell fraction.

31. A method to detect one or more modulators of a G protein coupled receptor, comprising:
a) providing a sample comprising one or more test agents, the host cell of claim 21 or a lysate thereof, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the sample.

32. A method to detect one or more modulators of a G protein coupled receptor, comprising:

a) providing a sample comprising one or more test agents, the circularly permuted luciferase of claim 22, and reagents for a luminescence reaction; and

b) detecting or determining luminescence in the sample.

33. A method to detect one or more modulators of a G protein coupled receptor, comprising:

a) comparing luminescence from a first luminogenic reaction mixture comprising one or more test agents, the host cell of claim 21 or a lysate thereof, and reagents for a luminescence reaction to luminescence from a corresponding luminogenic reaction mixture that does not contain the one or more test agents, but includes the host cell of claim 21 or a lysate thereof, and reagents for a luminescence reaction; and

b) detecting or determining whether the one or more test agents in the first luminogenic reaction mixture alter the luminescence in the first luminogenic reaction mixture relative to the corresponding luminogenic reaction mixture.

34. A method to detect one or more modulators of a G protein coupled receptor, comprising:

a) comparing luminescence from a first luminogenic reaction mixture comprising one or more test agents, the circularly permuted luciferase of claim 22, and reagents for a luminescence reaction to luminescence from a corresponding luminogenic reaction mixture that does not contain the one or more test agents, but includes the circularly permuted luciferase of claim 22 and the reagents; and

b) detecting or determining whether the agents in the first luminogenic reaction mixture alter the luminescence in the first luminogenic reaction mixture relative to the corresponding luminogenic reaction mixture.

35. The method of any one of claims 31 to 34 wherein the one or more agents enhance G protein coupled receptor activity.

36. The method of any one of claims 31 to 34 wherein the one or more agents inhibit G protein coupled receptor activity.

37. The method of claim 31 or 34 wherein the one or more agents are contacted with a solid substrate prior to addition of the host cell or lysate thereof and reagents.

38. The method of claim 32 or 34 wherein the one or more agents are contacted with a solid substrate prior to addition of the circularly permuted luciferase and reagents.

39. The method of claim 31 or 33 wherein the host cell expresses a recombinant G protein coupled receptor.

40. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a circularly permuted anthozoan luciferase optionally comprising an insertion comprising an amino acid sequence which directly or indirectly interacts with a molecule of interest, wherein the permutation is at a residue or in a region in an anthozoan luciferase sequence which is tolerant to modification, and wherein the insertion is at a different residue or region in an anthozoan luciferase sequence that is tolerant to modification.

41. The polynucleotide of claim 40 wherein the nucleic acid sequence encodes a fusion protein comprising the circularly permuted anthozoan luciferase and a tag of at least one amino acid at the N-terminus, C-terminus, or both.

42. The polynucleotide of claim 40 wherein the circularly permuted anthozoan luciferase further comprises a deletion of anthozoan luciferase sequences corresponding to N-terminal and/or C-terminal sequences of a corresponding noncircularly permuted anthozoan luciferase.

43. The polynucleotide of claim 42 wherein the deletion is no more than 15 residues of anthozoan luciferase sequence.

44. The polynucleotide of claim 40 wherein the circularly permuted anthozoan luciferase further comprises a deletion of anthozoan luciferase sequences N-terminal and/or C-terminal to the insertion.

45. The polynucleotide of claim 44 wherein the deletion is no more than 15 residues of anthozoan luciferase sequence.

46. The polynucleotide of claim 40 wherein the insertion is at sequences corresponding to the N-terminus and/or C-terminus of a corresponding noncircularly permuted anthozoan luciferase.

47. The polynucleotide of claim 40 wherein the insertion is about 4 to about 200 amino acid residues.

48. The polynucleotide of claim 40 wherein the insertion is in a region corresponding to residue 2 to 12, residue 26 to 47, residue 64 to 74, residue 86 to 116, residue 147 to 157, residue 223 to 234, or residue 301 to 311 of a noncircularly permuted *Renilla* luciferase.

49. The polynucleotide of claim 40 wherein the permutation is in a region corresponding to residue 2 to 12, residue 26 to 47, residue 64 to 74, residue 86 to 116, residue 147 to 157, residue 223 to 234, or residue 301 to 311 of a noncircularly permuted *Renilla* luciferase.

50. A vector comprising the polynucleotide of any one of claims 40 to 49.

51. A host cell comprising the polynucleotide in the vector of claim 50.

52. A modified anthozoan luciferase encoded by the polynucleotide of any one of claims 40 to 49.

53. A method to detect or determine a molecule of interest in a cell, comprising:
a) providing a mixture comprising the host cell of claim 51 or a lysate thereof, and reagents for a luminescence reaction, wherein the circularly permuted luciferase comprises the insertion; and
b) detecting or determining luminescence in the mixture, thereby detecting or determining the presence or amount of the molecule in the cell.

54. A method to detect or determine a molecule of interest in a sample, comprising:
a) providing a mixture comprising a sample suspected of having cyclic nucleotide, the modified luciferase of claim 52 which comprises the insertion, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the mixture.

55. A method to detect one or more modulators of a molecule of interest, comprising:
a) providing a sample comprising one or more test agents, the host cell of claim 51 or a lysate thereof, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the sample.

56. A method to detect one or more modulators of a molecule of interest, comprising:
a) providing a sample comprising one or more test agents, the luciferase of claim 52 which comprises the insertion, and reagents for a luminescence reaction; and
b) detecting or determining luminescence in the sample.

57. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified beetle luciferase, wherein the modified beetle luciferase comprises a cyclic nucleotide binding site relative to a corresponding unmodified beetle luciferase, wherein the cyclic nucleotide binding is at a residue or in a region in a beetle luciferase sequence which is tolerant to modification, and wherein the activity of the modified beetle luciferase is detectable.

58. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site is at the C-terminus of the modified beetle luciferase.

59. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site is internal to the N-terminus and C-terminus.

60. The polynucleotide of claim 57 wherein the nucleic acid sequence further encodes a tag of at least one amino acid at the N-terminus, C-terminus, or both, of the modified beetle luciferase.

61. The polynucleotide of claim 57 wherein the modified beetle luciferase further comprises a deletion of beetle luciferase sequences N-terminal and/or C-terminal to the cyclic nucleotide binding site.

62. The polynucleotide of claim 61 wherein the deletion is no more than 15 residues of beetle luciferase sequence.

63. The polynucleotide of claim 57 wherein the modified beetle luciferase further comprises a deletion of beetle luciferase sequences at sequences corresponding to the N-terminus and/or C-terminus of the unmodified beetle luciferase.

64. The polynucleotide of claim 63 wherein the deletion at the N-terminus or the C-terminus is no more than 15 residues of beetle luciferase sequence.

65. The polynucleotide of claim 57 wherein the beetle luciferase is a click beetle luciferase.

66. The polynucleotide of claim 57 wherein the beetle luciferase is a firefly luciferase.

67. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site is in a region corresponding to residue 2 to 12, residue 32 to 53, residue 70 to 88, residue 102 to 126, residue 139 to 165, residue 183 to 203, residue 220 to 247, residue 262 to 273, residue 303 to 313, residue 353 to 408, residue 485 to 495, or residue 535 to 546 of a firefly luciferase

68. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site is in a region corresponding to residue 15 to 30, residue 112 to 122, residue 352 to 362, residue 371 to 384, residue 393 to 414, or residue 485 to 495 of a click beetle luciferase.

69. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site binds to cAMP.

70. The polynucleotide of claim 57 wherein the cyclic nucleotide binding site binds to cGMP.

71. A vector comprising the polynucleotide of any one of claims 57 to 70.

72. A host cell comprising the vector of claim 71.

73. A modified beetle luciferase encoded by the polynucleotide of any one of claims 57 to 70.

74. A method to detect a cyclic nucleotide in a cell, comprising: a) contacting a cell with the vector of claim 71; and b) detecting or determining the activity of the modified beetle luciferase encoded by the vector.

75. A method to detect a cyclic nucleotide in a sample, comprising: a) contacting a sample with the modified beetle luciferase of claim 73; and b) detecting or determining the activity of the modified beetle luciferase encoded by the vector.

76. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified anthozoan luciferase, wherein the modified anthozoan luciferase comprises an internal insertion relative to a corresponding unmodified anthozoan luciferase, which insertion is at a residue or in a region which in a corresponding wild-type anthozoan luciferase is tolerant to modification, wherein the insertion comprises an amino acid which directly or indirectly interacts with a molecule of interest, and wherein the activity of the modified anthozoan luciferase is detectable.

77. The polynucleotide of claim 76 wherein the insertion is in a region corresponding to residue 2 to 12, residue 26 to 47, residue 64 to 74, residue 86 to 116, residue 147 to 157, residue 223 to 234, or residue 301 to 311 of a *Renilla* luciferase.

78. A vector comprising the polynucleotide of claim 76 or 77.

79. A host cell comprising the vector of claim 78.

80. A modified anthozoan luciferase encoded by the polynucleotide of claim 76 or 77.

81. A method to detect a molecule of interest in a cell, comprising: a) contacting a sample having cells or *in vitro* transcription/translation mixture and the vector of claim 78, wherein the insertion is recognized by the molecule of interest; and b) detecting or determining the activity of the modified anthozoan luciferase encoded by the vector, thereby detecting or determining the presence or amount of the molecule in the sample.

82. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified firefly luciferase, wherein the modified firefly luciferase comprises an internal insertion relative to a corresponding unmodified firefly luciferase, which insertion is at a residue or in a region in a firefly luciferase sequence which is tolerant to modification, wherein the insertion comprises an amino acid sequence which directly or indirectly interacts with a molecule of interest relative to the corresponding unmodified firefly luciferase, and wherein the activity of the modified firefly luciferase is detectable, and wherein the insertion is not in a region corresponding to residue 2 to 12, residue 116 to 126, residue 228 to 238, residue 262 to 272, residue 289 to 308, residue 356 to 366, residue 432 to 442, or residue 535 to 546 of a firefly luciferase.

83. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified firefly luciferase, wherein the modified firefly luciferase comprises an insertion and a fragment of firefly luciferase sequence relative to a corresponding unmodified firefly luciferase, wherein the fragment has at least 50 contiguous amino acid residues the corresponding unmodified firefly luciferase, wherein the insertion is at a residue or in a region in a firefly luciferase sequence which is tolerant to modification, wherein the activity of the modified firefly luciferase is

increased by a second fragment of firefly luciferase sequence which corresponds to a different at least 50 contiguous amino acid residues of the corresponding unmodified firefly luciferase, wherein the insertion comprises an amino acid sequence which directly or indirectly interacts with a molecule of interest, and wherein the C-terminus or N-terminus of the fragment does not correspond to a residue in a region corresponding to residue 116 to 126, residue 228 to 238, residue 262 to 272, residue 289 to 308, residue 356 to 366, or residue 432 to 442 of a firefly luciferase.

84. The polynucleotide of claim 82 or 83 wherein the modified firefly luciferase further comprises a deletion of firefly luciferase sequences N-terminal and/or C-terminal to the insertion.

85. The polynucleotide of claim 84 wherein the deletion is no more than 15 residues of firefly luciferase sequence.

86. The polynucleotide of claim 82 or 83 wherein the modified firefly luciferase further comprises a deletion of firefly luciferase sequences at sequences corresponding to the N-terminus and/or C-terminus of the unmodified firefly luciferase.

87. The polynucleotide of claim 86 wherein the deletion at the N-terminus or the C-terminus is no more than 15 residues of firefly luciferase sequence.

88. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a circularly permuted copepod luciferase optionally comprising an insertion comprising an amino acid sequence which directly or indirectly interacts with a molecule of interest, wherein the permutation is at a residue or in a region in a copepod luciferase sequence which is tolerant to modification, and wherein the insertion is at a different residue or region in a copepod luciferase sequence that is tolerant to modification.

89. The polynucleotide of claim 88 wherein the permutation is in a region corresponding to residue 43 to 53, residue 63 to 73, residue 79 to 89, residue 95 to 105, residue 105 to 115, residue 109 to 119, residue 121 to 131 or residue157 to 168 of a *Gaussia* luciferase.

90. The method of claim 23 or 33 wherein the mixture, host cell, detection or determination of luminescence, or any combination thereof, is at about 30°C to about 47°C.

91. The method of claim 31 wherein the host cell, sample, detection or determination of luminescence, or any combination thereof, is at about 30°C to about 47°C.

92. The method of claim 74 wherein the cell, detection or determination of activity, or any combination thereof, is at about 30°C to about 47°C.

93. The method of claim 23, 31, or 33 wherein the host cell is stably transfected with the polynucleotide.

94. The method of claim 74 wherein the cell is stably transfected with the vector.

95. A method to detect or determine the presence or activity of a molecule of interest in a cell, comprising: a) providing a luminogenic reaction mixture comprising a cell with a vector having a nucleic acid sequence comprising an open reading frame for a modified luciferase, wherein the modified luciferase comprises an insertion relative to a corresponding unmodified luciferase, which insertion is at a residue or in a region in a luciferase sequence which is tolerant to modification, wherein the insertion comprises an amino acid sequence which directly or indirectly interacts with a

molecule of interest relative to the corresponding unmodified luciferase, wherein the activity of the modified luciferase is detectable, and wherein the mixture is at about 30°C to about 47°C; and

b) detecting or determining luminescence in the mixture, thereby detecting or determining the presence, amount or activity of the molecule in the cell.

96. The method of claim 95 wherein the modified luciferase is a beetle luciferase.

97. The method of claim 96 wherein the modified beetle luciferase is a firefly luciferase.

98. The method of claim 96 wherein the modified beetle luciferase is a click beetle luciferase.

99. The method of claim 95 wherein the modified luciferase is a circularly permuted luciferase.

100. The method of claim 95 wherein the insertion is a cyclic nucleotide binding site.

101. The method of claim 95 wherein the mixture further comprises one or more test agents.

102. The method of claim 101 wherein the luminescence in the mixture is compared to the luminescence in a corresponding mixture which lacks the one or more test agents.

103. The method of claim 95 wherein the mixture is subjected to about 5% CO_2.

104. A method to identify an agent that directly or indirectly alters the activity of a heterologous sequence in modified beetle luciferase, comprising:

a) contacting a luminogenic reaction mixture comprising a cell expressing a modified luciferase with one or more agents, wherein the modified luciferase is encoded by a vector having a nucleic acid sequence comprising an open reading frame for a modified luciferase, wherein the modified luciferase comprises a heterologous sequence relative to a corresponding unmodified luciferase, which heterologous sequence is at a residue or in a region in a luciferase sequence which is tolerant to modification, wherein the heterologous sequence comprises an amino acid sequence which directly or indirectly interacts with a molecule of interest relative to the corresponding unmodified luciferase, wherein the activity of the modified luciferase is detectable, and wherein the mixture is at about 30°C to about 47°C; and

b) identifying whether the one or more agents alter the activity of the modified beetle luciferase.

105. The method of claim 104 wherein the modified luciferase is a beetle luciferase.

106. The method of claim 105 wherein the beetle luciferase is a firefly luciferase.

107. The method of claim 105 wherein the beetle luciferase is a click beetle luciferase.

108. The method of claim 95 or 105 wherein the cell is stably transfected with the vector.

109. The method of claim 104 wherein the heterologous sequence is a cyclic nucleotide binding site.

110. The method of claim 104 wherein the mixture is subjected to about 5% CO_2.

111. The method of claim 95 or 104 wherein the modified luciferase is a modified anthozoan luciferase.

112. The polynucleotide of claim 40 wherein the insertion includes a peptide substrate for a serine, threonine or tyrosine kinase.

113. The polynucleotide of claim 40 wherein the insertion includes a phosphoserine peptide binding domain, a phosphothreonine peptide binding domain, or a phosphotyrosine peptide binding domain.

114. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a circularly permuted luciferase which comprises a heterologous amino acid sequence comprising a peptide substrate for a serine, threonine or tyrosine kinase, which luciferase is permuted at a site or in a region which in a corresponding nonpermuted luciferase is tolerant to modification, and wherein the activity of the circularly permuted luciferase is detectable.

115. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a circularly permuted luciferase which comprises a heterologous amino acid sequence comprising a phosphoserine peptide binding domain, a phosphothreonine peptide binding domain, or a phosphotyrosine peptide binding domain, which luciferase is permuted at a site or in a region which in a corresponding nonpermuted luciferase is tolerant to modification, and wherein the activity of the circularly permuted luciferase is detectable.

116. The polynucleotide of claim 114 or 115 wherein the luciferase is an anthozoan luciferase.

117. The polynucleotide of claim 114 or 115 wherein the luciferase is a beetle luciferase.

118. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified decapod luciferase, wherein the modified decapod luciferase comprises an insertion and a first fragment of decapod luciferase sequence relative to a corresponding unmodified decapod luciferase, wherein the first fragment has at least 35 contiguous amino acid residues of the corresponding unmodified mature decapod luciferase, wherein the insertion is at a residue or in a region in a mature decapod luciferase sequence which is tolerant to modification, wherein the activity of the modified decapod luciferase is increased by a second fragment of a decapod luciferase sequence which corresponds to a different at least 35 contiguous amino acid residues of the corresponding unmodified mature decapod luciferase, wherein the insertion comprises an amino acid sequence which directly or indirectly interacts with a molecule of interest.

119. A polynucleotide comprising a nucleic acid sequence comprising an open reading frame for a modified decapod luciferase, wherein the modified decapod luciferase comprises an internal insertion relative to a corresponding unmodified decapod luciferase, which insertion is at a residue or in a region in a mature decapod luciferase sequence which is tolerant to modification, wherein the insertion comprises an amino acid sequence which directly or indirectly interacts with a molecule of interest relative to the corresponding unmodified mature decapod luciferase, and wherein the activity of the modified decapod luciferase is detectable.

120. The polynucleotide of claim 118 wherein the first fragment has at least 45 to at least 85 residues.

121. The polynucleotide of claim 119 wherein the insertion is in a region corresponding to residue 45 to 55 or residue 79 to 89 of an *Oplophorus* luciferase.

122. A method to determine one or more sites tolerant to modification in a native protein, comprising:

identifying in a protein likely to have a native structure that includes secondary structure, one or more residues in the amino acid sequence of the protein that are in a region having a majority of hydrophilic residues, which region does not likely form a secondary structure in the native protein structure.