



(86) Date de dépôt PCT/PCT Filing Date: 1998/03/13 (87) Date publication PCT/PCT Publication Date: 1998/09/17 (45) Date de délivrance/Issue Date: 2010/08/03 (85) Entrée phase nationale/National Entry: 1999/09/14 (86) N° demande PCT/PCT Application No.: US 1998/004978 (87) N° publication PCT/PCT Publication No.: 1998/040477 (30) Priorités/Priorities: 1997/03/14 (US08/818,253); 1997/03/14 (US08/818,252); 1997/08/27 (US08/919,143)	(51) Cl.Int./Int.Cl. <i>C12N 15/00</i> (2006.01), <i>C07H 21/02</i> (2006.01), <i>C07H 21/04</i> (2006.01), <i>C07K 14/435</i> (2006.01), <i>C12N 15/09</i> (2006.01), <i>G01N 33/542</i> (2006.01) (72) Inventeurs/Inventors: TSIEN, ROGER Y., US; MIYAWAKI, ATSUSHI, US (73) Propriétaire/Owner: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA, US (74) Agent: MBM INTELLECTUAL PROPERTY LAW LLP
--	--

(54) Titre : DETECTEURS DE PROTEINES FLUORESCENTES PERMETTANT DE DETECTER DES SUBSTANCES A ANALYSER

(54) Title: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES

(57) **Abrégé/Abstract:**

Fluorescent indicators including a binding protein moiety, a donor fluorescent protein moiety, and an acceptor fluorescent protein moiety are described. The binding protein moiety has an analyte-binding region which binds an analyte and causes the indicator to change conformation upon exposure to the analyte. The donor moiety and the acceptor moiety change position relative to each other when the analyte binds to the analyte-binding region. The donor moiety and the acceptor moiety exhibit fluorescence resonance energy transfer when the donor moiety is excited and the distance between the donor moiety and the acceptor moiety is small. The indicators can be used to measure analyte concentrations in samples, such as calcium ion concentrations in cells.



**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/00, 15/09, C07H 21/02, 21/04		A1	(11) International Publication Number: WO 98/40477 (43) International Publication Date: 17 September 1998 (17.09.98)
(21) International Application Number: PCT/US98/04978 (22) International Filing Date: 13 March 1998 (13.03.98) (30) Priority Data: 08/818,252 14 March 1997 (14.03.97) US 08/818,253 14 March 1997 (14.03.97) US 08/919,143 27 August 1997 (27.08.97) US (71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US). (72) Inventors: TSIEN, Roger, Y.; 8535 Nottingham Place, La Jolla, CA 92037 (US). MIYAWAKI, Atsushi; 3899 Nobel Drive, San Diego, CA 92122 (US). (74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES (57) Abstract Fluorescent indicators including a binding protein moiety, a donor fluorescent protein moiety, and an acceptor fluorescent protein moiety are described. The binding protein moiety has an analyte-binding region which binds an analyte and causes the indicator to change conformation upon exposure to the analyte. The donor moiety and the acceptor moiety change position relative to each other when the analyte binds to the analyte-binding region. The donor moiety and the acceptor moiety exhibit fluorescence resonance energy transfer when the donor moiety is excited and the distance between the donor moiety and the acceptor moiety is small. The indicators can be used to measure analyte concentrations in samples, such as calcium ion concentrations in cells.			

FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES

Background of the Invention

The invention relates to fluorescent protein sensors for detecting and quantifying analytes.

- 10 Measurement of an analyte concentration *in vitro* or *in vivo* by non-invasive techniques can help elucidate the physiological function of the analyte. This can also aid in identifying changes that occur in a cell or organism in response to physiological stimuli. For example, cyclic AMP can be detected by fluorescence resonance energy transfer between a separately labeled proteins that associate with each other but are not
15 covalently attached to each other. See, U.S. Pat. No. 5,439,797.

For example, many effects of Ca^{2+} in cells are mediated by Ca^{2+} binding to calmodulin (CaM), which causes CaM to bind and activate target proteins or peptide sequences. Based on the NMR solution structure of CaM bound to the 26-residue M13 Ca^{2+} -binding peptide of myosin light-chain kinase, Porumb *et al.* fused the C-terminus of CaM via a
20 Gly-Gly spacer to the M13. Ca^{2+} binding switches the resulting hybrid protein (CaM-M13) from a dumbbell-like extended form to a compact globular form similar to the CaM-M13 intermolecular complex. See, Porumb, T., *et al.*, *Prot.Engineering* 7:109-115 (1994).

Fluorescent Ca^{2+} indicators such as fura-2, indo-1, fluo-3, and Calcium-Green have been
25 the mainstay of intracellular Ca^{2+} measurement and imaging. See, for example, U.S. Pat. No. 4,603,209 and U.S. Pat. No. 5,049,673. These relatively low molecular weight indicators can suffer from many technical problems relating to ester loading, leakage of

In preferred embodiments, the indicator further includes the target peptide moiety and a linker moiety that covalently couples the binding protein and the target peptide moiety. The binding protein moiety further includes a peptide-binding region for binding the target peptide moiety. The binding protein moiety can be covalently coupled to the donor fluorescent protein moiety and the target peptide moiety can be covalently coupled to the acceptor fluorescent protein moiety.

- 3 -

The indicator can be a single polypeptide. In preferred embodiments, one of the donor fluorescent protein moiety or the acceptor fluorescent protein moiety is covalently coupled to the carboxy terminus of the single polypeptide and the other of the donor fluorescent protein moiety or the acceptor fluorescent protein moiety is covalently
5 coupled to the amino terminus of the single polypeptide.

The indicator can include a localization sequence. The localization sequence can be a nuclear localization sequence, an endoplasmic reticulum localization sequence, a peroxisome localization sequence, a mitochondrial import sequence, a mitochondrial localization sequence, or a localized protein.

10 In preferred embodiments, the linker moiety is a peptide moiety. The linker moiety can include between about 1 amino acid residue and about 20 amino acid residues. The linker moiety can be -Gly-Gly-.

Preferably, the binding protein moiety is calmodulin, a calmodulin-related protein moiety, cGMP-dependent protein kinase, a steroid hormone receptor, a ligand binding
15 domain of a steroid hormone receptor, protein kinase C, inositol-1,4,5-triphosphate receptor, or recoverin. A calmodulin-related protein moiety is derived from calmodulin that has been modified to have a different binding affinity for calcium or a target peptide moiety.

Most preferably, the binding protein moiety is calmodulin or a calmodulin-related protein
20 moiety. In these embodiments, the target peptide moiety can be a subsequence of a calmodulin-binding domain of M13, smMLCKp, CaMKII, Caldesmon, Calspermin, Calcineurin, PhK5, PhK13, C28W, 59-kDa PDE, 60-kDa PDE, NO-30, AC-28, *Bordetella pertussis* AC, Neuromodulin, Spectrin, MARCKS, F52, β -Adducin, HSP90a, HIV-1 gp160, BBMHBI, Dilute MHC, Mastoparan, Melittin, Glucagon, Secretin, VIP,
25 GIP, or Model Peptide CBP2. Preferably, the target peptide moiety is M13.

- 4 -

In another aspect, the invention features a fluorescent indicator including a target peptide moiety, a binding protein moiety, a linker moiety, a donor fluorescent protein moiety covalently coupled to the binding protein moiety, and an acceptor fluorescent protein moiety covalently coupled to the binding protein moiety. The binding protein moiety has
5 an analyte-binding region which binds an analyte and causes the indicator to change conformation upon exposure to the analyte. The linker moiety covalently couples the binding protein and the target peptide moiety and is a peptide moiety. The binding protein moiety has a peptide-binding region for binding the target peptide moiety. The indicator is a single polypeptide.

10 In another aspect, the invention features a method for determining the concentration of an analyte in a sample. The method includes the steps of contacting the sample with a fluorescent indicator having a donor fluorescent protein moiety, binding protein moiety, and acceptor protein moiety, exciting the donor moiety, and determining the degree of fluorescence resonance energy transfer in the sample corresponding to the concentration
15 of the analyte in the sample.

In preferred embodiments, the step of determining the degree of fluorescence resonance energy transfer in the sample includes measuring light emitted by the acceptor fluorescent protein moiety. In other preferred embodiments, determining the degree of fluorescence resonance energy transfer in the sample includes measuring light emitted
20 from the donor fluorescent protein moiety, measuring light emitted from the acceptor fluorescent protein moiety, and calculating a ratio of the light emitted from the donor fluorescent protein moiety and the light emitted from the acceptor fluorescent protein moiety. In yet other preferred embodiments, determining the degree of fluorescence resonance energy transfer in the sample includes measuring the excited state lifetime of
25 the donor moiety.

The method can further include the steps of determining the concentration of the analyte at a first time after contacting the sample with the fluorescence indicator, determining the concentration of the analyte at a second time after contacting the sample with the

- 5 -

fluorescence indicator, and calculating the difference in the concentration of the analyte at the first time and the second time, whereby the difference in the concentration of the analyte in the sample reflects a change in concentration of the analyte present in the sample.

- 5 In other embodiments, the method can further include the step of contacting the sample with a compound between the first time and the second time, whereby a difference in the concentration of the analyte in the sample between the first time and the second time indicates that the compound alters the presence of the analyte.

In preferred embodiments, the sample includes an intact cell and the contacting step
10 includes incorporating the fluorescent indicator into the cell. The step of incorporating the fluorescent indicator into the cell can include transfecting the cell with an expression vector comprising expression control sequences operably linked to a nucleic acid sequence coding for the expression of the fluorescent indicator. The preferred analyte is calcium.

- 15 In yet another aspect, the invention features an isolated nucleic acid sequence which encodes the fluorescent indicator. In preferred embodiments, an expression vector or a transgenic non-human animal includes the nucleic acid sequence.

In another aspect, the invention features an expression vector including expression control sequences operatively linked to a nucleic acid sequence coding for the expression
20 of the fluorescent indicator. The expression vector can be adapted for function in a prokaryotic cell or a eukaryotic cell.

In another aspect of the invention, a host cell transfected with an expression vector can include an expression control sequence operatively linked to a sequence coding for the expression of the fluorescent indicator. The cell can be a prokaryote, such as *E. coli*, or
25 a eukaryotic cell, such as a yeast cell or mammalian cell.

- 6 -

In another aspect, the invention features a transgenic non-human animal having a phenotype characterized by expression of the nucleic acid sequence coding for the expression of the fluorescent indicator. The phenotype is conferred by a transgene contained in the somatic and germ cells of the animal. The animal can be a mouse.

5 In another aspect, the invention features a method for producing a transgenic non-human animal having a phenotype characterized by expression of the nucleic acid sequence coding for the expression of the fluorescent indicator. The method includes the steps of: (a) introducing a transgene into a zygote of an animal, the transgene comprising a DNA construct encoding the fluorescent indicator; (b) transplanting the zygote into a
10 pseudopregnant animal; (c) allowing the zygote to develop to term; and (d) identifying at least one transgenic offspring containing the transgene. The step of introducing of the transgene into the embryo can be by introducing an embryonic stem cell containing the transgene into the embryo, or infecting the embryo with a retrovirus containing the transgene.

15 "Peptide" refers to a polymer in which the monomers are amino acid residues which are joined together through amide bonds, alternatively referred to as a polypeptide. A "single polypeptide" is a continuous peptide that constitutes the protein. When the amino acids are alpha-amino acids, either the L-optical isomer or the D-optical isomer can be used, the L-isomers being preferred. Additionally, unnatural amino acids such as beta-
20 alanine, phenylglycine, and homoarginine are meant to be included. Commonly encountered amino acids which are not gene-encoded can also be used in the present invention, although preferred amino acids are those that are encodable. For a general review, see, for example, Spatola, A.F., in *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*, B. Weinstein, ed., Marcel Dekker, New York, p. 267 (1983).

25 "Fluorescent protein" refers to any protein capable of emitting light when excited with appropriate electromagnetic radiation. Fluorescent proteins include proteins having amino acid sequences that are either natural or engineered, such as the fluorescent proteins derived from *Aequorea*-related fluorescent proteins.

- 7 -

In FRET, the "donor fluorescent protein moiety" and the "acceptor fluorescent protein moiety" are selected so that the donor and acceptor moieties exhibit fluorescence resonance energy transfer when the donor moiety is excited. One factor to be considered in choosing the donor/acceptor fluorescent protein moiety pair is the efficiency of FRET
5 between the two moieties. Preferably, the efficiency of FRET between the donor and acceptor moieties is at least 10%, more preferably at least 50%, and most preferably at least 80%. The efficiency of FRET can be tested empirically using the methods described herein and known in the art, particularly, using the conditions set forth in the Examples.

10 "Covalently coupled" refers to a covalent bond or other covalent linkage between two moieties. The covalent linkage can include a diradical moiety linking to two moieties.

"Binding protein" refers to a protein capable of binding an analyte. Preferred binding proteins change conformation upon binding the analyte. "Target peptide" refers to a peptide that can bind to the binding protein. The target peptide can be a subsequence of
15 a peptide that binds to the binding protein.

"Analyte" refers to a molecule or ion in solution that binds to the binding protein, causing it to change conformation. Preferably, the analyte binds reversibly to the binding protein.

"Moiety" refers to a radical of a molecule that is attached to another radical of the indicator. Thus, a "fluorescent protein moiety" is the radical of a fluorescent protein
20 coupled to a binding protein moiety or a linker moiety, a "binding protein moiety" is a radical of a binding protein coupled to a fluorescent protein moiety, a "target peptide moiety" is a radical of a target peptide of the binding protein, and a "linker moiety" refers to the radical of a molecular linker that is ultimately coupled to both the donor and acceptor fluorescent protein moieties.

25 "Operatively linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control

- 8 -

sequence operatively linked to a coding sequence is ligated such that expression of the coding sequence is achieved under conditions compatible with the control sequences. "Control sequence" refers to polynucleotide sequences which are necessary to effect the expression of coding and non-coding sequences to which they are ligated. Control
5 sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, components whose presence can influence expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

- 10 "Polynucleotide" refers to a polymeric form of nucleotides of at least 10 bases in length. The nucleotides can be ribonucleotides, deoxynucleotides, or modified forms of either type of nucleotide. The term includes single and double stranded forms of DNA.

The invention can have one or more of the following advantages. Ligand-induced conformational changes can be monitored by FRET if, for example, the amino and
15 carboxy termini of the binding protein are fused to a donor and acceptor GFP. This approach has several advantages over the usual covalent labeling with fluorescent probes. The fluorescent indicator can be generated *in situ* by gene transfer into the cells or organisms. This approach avoids the need to express and purify large quantities of soluble recombinant protein, purify and label it *in vitro*, microinject it back into cells.
20 The fluorescent indicator can be targeted to cellular structures. The sites of fusion between the moieties of the fluorescent indicator are exactly defined, giving a molecularly homogenous product without relying on elaborate protein chemistry. In addition, the chromophore of GFP is fixed in the protein. See, Ormo, M., *et al.*, *Science* 273:1392-1395 (1996). If the GFP donor and acceptor are fused to a host protein rigidly,
25 minor changes in the relative orientation of the ends of the latter would alter FRET. In contrast, most conventional fluorescent labels are attached by flexible linkers that at least partially decouple the fluorophore orientation from

FIG. 5 is a schematic diagram depicting the structures of *cameleon-2* and *cameleon-3*,
 20 and their derivatives, *cameleon-2nu* and *cameleon-3er*, where Kz is Kozak's consensus
 sequence.

- 10 -

FIG. 6a is a digital fluorescence image depicting fluorescence of cameleon-2 localized to the cytosol of HeLa cells, where the bar is 10 μ M.

FIG. 6b is a graph depicting temporal changes in the emission ratio of cameleon-2 for each of the HeLa cells shown in FIG. 6a.

5 FIG. 6c is a digital fluorescence image depicting fluorescence of cameleon-2nu localized to the nuclei of HeLa cells, where the bar is 10 μ M.

FIG. 6d is a graph depicting temporal changes in the emission ratio of cameleon-2nu in the two nuclei shown in FIG. 6c.

FIG. 6e is a digital fluorescence image depicting fluorescence of cameleon-3er in
10 transfected HeLa cells, where the bar is 10 μ M.

FIG. 6f is a graph depicting the time-course of emission ratio of cameleon-3er (average of four cells).

FIG. 7 depicts the nucleotide sequence (SEQ ID NO:1) and amino acid sequence (SEQ ID NO:2) of cameleon-2.

15 FIG. 8 depicts the nucleotide sequence (SEQ ID NO:3) and amino acid sequence (SEQ ID NO:4) of cameleon-2nu.

FIG. 9 depicts the nucleotide sequence (SEQ ID NO:5) and amino acid sequence (SEQ ID NO:6) of cameleon-3.

FIG. 10 is a list depicting the nucleotide sequence (SEQ ID NO:7) and amino acid
20 sequence of cameleon-3er (SEQ ID NO:8).

- 11 -

Detailed Description

A fluorescent indicator that utilizes fluorescent resonance energy transfer ("FRET") to measure the concentration of an analyte includes two fluorescent protein moieties having emission and excitation spectra that render one a donor fluorescent protein moiety and the other an acceptor fluorescent protein moiety. The fluorescent protein moieties are chosen such that the excitation spectrum of one of the moieties (the acceptor fluorescent protein moiety) overlaps with the emission spectrum of the excited protein moiety (the donor fluorescent protein moiety). The donor and acceptor fluorescent protein moieties are bound to a binding protein moiety that changes conformation upon binding the analyte. The change in conformation leads to a change in relative position and orientation of the donor and acceptor fluorescent protein moieties, thereby altering the relative amounts of fluorescence from the two fluorescent protein moieties when the donor is excited by irradiation. In particular, binding of the analyte changes the ratio of the amount of light emitted by the donor and acceptor fluorescent protein moieties. The ratio between the two emission wavelengths provides a measure of the concentration of the analyte in the sample, which is based in part on the binding affinity of the binding protein moiety and the analyte.

Referring to FIG. 1, the donor fluorescent protein moiety is covalently linked to a first region (*e.g.*, the amino terminus) of the binding protein moiety, and the acceptor fluorescent protein moiety is covalently linked to a second region (*e.g.*, the carboxy terminus) of the binding protein moiety such that the donor and acceptor moieties move closer together upon binding the analyte. Alternatively, the donor and acceptor moieties can move farther apart upon binding the analyte. In one embodiment, depicted in FIG. 1, the acceptor moiety is covalently bonded to a target peptide moiety that also binds to the binding protein moiety and the target peptide moiety is covalently bonded to the binding protein moiety by a linker moiety. The linker moiety is flexible enough to allow the target peptide moiety to bind to the binding protein moiety. The donor moiety is excited by light of appropriate intensity within the excitation spectrum of the donor moiety ($\Sigma_{\text{excitation}}$). The donor moiety emits the absorbed energy as fluorescent light (Σ_{emission}). When the acceptor fluorescent protein moiety is positioned to quench the

- 12 -

donor moiety in the excited state, the fluorescence energy is transferred to the acceptor moiety which can emit fluorescent light ($\Sigma_{\text{emission } 2}$). FRET can be manifested as a reduction in the intensity of the fluorescent signal from the donor moiety ($\Sigma_{\text{emission } 1}$), reduction in the lifetime of the excited state of the donor moiety, or emission of
 5 fluorescent light at the longer wavelengths (lower energies) characteristic of the acceptor moiety ($\Sigma_{\text{emission } 2}$). When the conformation of the binding protein moiety changes upon binding the analyte, the fluorescent protein moieties come closer together (or physically separate), and FRET is increased (or decreased) accordingly.

The efficiency of FRET depends on the separation distance and the orientation of the
 10 donor and acceptor fluorescent protein moieties. For example, the Forster equation describes the efficiency of excited state energy transfer, based in part on the fluorescence quantum yield of the donor moiety and the energetic overlap with the acceptor moiety. The Forster equation is:

$$E = (F_0 - F)/F_0 = R_0^6/(R^6 + R_0^6)$$

15 where E is the efficiency of FRET, F and F_0 are the fluorescence intensities of the donor moiety in the presence and absence of the acceptor, respectively, and R is the distance between the donor moiety and the acceptor moiety.

The characteristic distance R_0 at which FRET is 50% efficient depends on the quantum yield of the donor moiety (*i.e.*, the shorter-wavelength fluorophore), the extinction
 20 coefficient of the acceptor moiety (*i.e.*, the longer-wavelength fluorophore), and the overlap between the emission spectrum of the donor moiety and the excitation spectrum of the acceptor moiety. R_0 is given (in D) by

$$R_0 = 9.79 \times 10^3 (K^2 Q J n^4)^{1/6}$$

where K^2 is an orientation factor having an average value close to 0.67 for freely mobile
 25 donors and acceptors, Q is the quantum yield of the unquenched donor moiety, n is the refractive index of the medium separating the donor moiety and the acceptor moiety, and

- 13 -

J is the overlap integral. J can be quantitatively expressed as the degree of spectral overlap between the donor moiety and the acceptor moiety according to the equation:

$$J = \int_0^\infty M_\Sigma F_\Sigma \Sigma^4 d\Sigma / \int_0^\infty F_\Sigma d\Sigma$$

where M_Σ ($M^{-1}cm^{-1}$) is the molar absorptivity of the acceptor and F_Σ is the donor moiety
 5 fluorescence intensity at wavelength Σ . See, for example, Forster, T. *Ann. Physik* 2:55-75
 (1948). Tables of spectral overlap integrals are readily available to those working in the
 field (for example, Berlman, I.B. *Energy transfer parameters of aromatic compounds*,
 Academic Press, New York and London (1973)). FRET is a nondestructive
 spectroscopic method that can monitor proximity and relative angular orientation of
 10 fluorophores in living cells. See, for example, Adams, S.R., *et al.*, *Nature* 349:694-697
 (1991), and Gonzalez, J. & Tsien, R.Y. *Biophys. J.* 69:1272-1280 (1995).

These factors need to be balanced to optimize the efficiency and detectability of FRET
 from the fluorescent indicator. The emission spectrum of the donor fluorescent protein
 moiety should overlap as much as possible with the excitation spectrum of the acceptor
 15 fluorescent protein moiety to maximize the overlap integral J. Also, the quantum yield
 of the donor fluorescent protein moiety and the extinction coefficient of the acceptor
 fluorescent protein moiety should be as large as possible to maximize R_0 . In addition,
 the excitation spectra of the donor and acceptor moieties should overlap as little as
 possible so that a wavelength region can be found at which the donor moiety can be
 20 excited selectively and efficiently without directly exciting the acceptor moiety. Direct
 excitation of the acceptor moiety should be avoided since it can be difficult to distinguish
 direct emission from fluorescence arising from FRET. Similarly, the emission spectra
 of the donor and acceptor moieties should have minimal overlap so that the two
 emissions can be distinguished. High fluorescence quantum yield of the acceptor moiety
 25 is desirable if the emission from the acceptor moiety is to be monitored to determine
 analyte concentration in a sample. In a preferred embodiment, the donor fluorescent
 protein moiety is excited by ultraviolet (<400 nm) and emits blue light (<500 nm), and
 the acceptor fluorescent protein moiety is efficiently excited by blue but not by

- 14 -

ultraviolet light and emits green light (>500 nm), for example, P4-3 and S65T, respectively.

In another preferred embodiment, the donor fluorescent moiety is excited by violet (400-430 nm) and emits blue-green (450-500 nm) and the acceptor fluorescent moiety is
5 efficiently excited by blue-green (450-500 nm) and emits yellow-green light (≈ 520 nm), for example WIB and 10C respectively.

The amount of analyte in a sample can be determined by determining the degree of FRET in the sample. Changes in analyte concentration can be determined by monitoring FRET at a first and second time after contact between the sample and the fluorescent indicator
10 and determining the difference in the degree of FRET. The amount of analyte in the sample can be calculated by using a calibration curve established by titration.

The degree of FRET can be determined by any spectral or fluorescence lifetime characteristic of the excited donor moiety. For example, intensity of the fluorescent signal from the donor, the intensity of fluorescent signal from the acceptor, the ratio of
15 the fluorescence amplitudes near the acceptor's emission maxima to the fluorescence amplitudes near the donor's emission maximum, or the excited state lifetime of the donor can be monitored.

Preferably, changes in the degree of FRET are determined as a function of the change in the ratio of the amount of fluorescence from the donor and acceptor moieties, a process
20 referred to as "ratioing." Changes in the absolute amount of indicator, excitation intensity, and turbidity or other background absorbances in the sample at the excitation wavelength affect the intensities of fluorescence from both the donor and acceptor approximately in parallel. Therefore the ratio of the two emission intensities is a more robust and preferred measure of cleavage than either intensity alone.

25 Fluorescence in a sample is measured using a fluorometer. In general, excitation radiation, from an excitation source having a first wavelength, passes through excitation

- 15 -

optics. The excitation optics cause the excitation radiation to excite the sample. In response, fluorescent proteins in the sample emit radiation which has a wavelength that is different from the excitation wavelength. Collection optics then collect the emission from the sample. The device can include a temperature controller to maintain the sample
5 at a specific temperature while it is being scanned. According to one embodiment, a multi-axis translation stage moves a microtiter plate holding a plurality of samples in order to position different wells to be exposed. The multi-axis translation stage, temperature controller, auto-focusing feature, and electronics associated with imaging and data collection can be managed by an appropriately programmed digital computer.
10 The computer also can transform the data collected during the assay into another format for presentation.

Methods of performing assays on fluorescent materials are well known in the art and are described in, e.g., Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*, New York: Plenum Press (1983); Herman, B., Resonance energy transfer microscopy, in:
15 *Fluorescence Microscopy of Living Cells in Culture, Part B, Methods in Cell Biology*, vol. 30, ed. Taylor, D.L. & Wang, Y.-L., San Diego: Academic Press (1989), pp. 219-243; Turro, N.J., *Modern Molecular Photochemistry*, Menlo Park: Benjamin/Cummings Publishing Co., Inc. (1978), pp. 296-361.

The excited state lifetime of the donor moiety is, likewise, independent of the absolute
20 amount of substrate, excitation intensity, or turbidity or other background absorbances. Its measurement requires equipment with nanosecond time resolution.

Quantum yields of wild-type GFP, S65T, and P4-1 mutants can be estimated by comparison with fluorescein in 0.1 N NaOH as a standard of quantum yield 0.91. J.N. Miller, ed., *Standards in Fluorescence Spectrometry*, New York: Chapman and Hall
25 (1981). Mutants P4 and P4-3 were likewise compared to 9-aminoacridine in water (quantum yield 0.98).

Any fluorescent protein can be used in the invention, including proteins that fluoresce due to intramolecular rearrangements or the addition of cofactors that promote fluorescence. For example, green fluorescent proteins of cnidarians, which act as their energy-transfer acceptors in bioluminescence, are suitable fluorescent proteins for use in the fluorescent indicators. A green fluorescent protein ("GFP") is a protein that emits green light, and a blue fluorescent protein ("BFP") is a protein that emits blue light. GFPs have been isolated from the Pacific Northwest jellyfish, *Aequorea victoria*, the sea pansy, *Renilla reniformis*, and *Phialidium gregarium*. See, Ward, W.W., et al., *Photochem. Photobiol.*, 35:803-808 (1982); and Levine, L.D., et al., *Comp. Biochem. Physiol.*, 72B:77-85 (1982).

A variety of *Aequorea*-related GFPs having useful excitation and emission spectra have been engineered by modifying the amino acid sequence of a naturally occurring GFP from *Aequorea victoria*. See, Prasher, D.C., et al., *Gene*, 111:229-233 (1992); Heim, R., et al., *Proc. Natl. Acad. Sci., USA*, 91:12501-04 (1994); U.S. Patent No. 5,625,048 issued April 29, 1997 or International Publication WO 96/23810 published on August 8, 1996.

The cDNA of GFP can be concatenated with those encoding many other proteins; the resulting fusions often are fluorescent and retain the biochemical features of the partner proteins. See, Cubitt, A.B., et al., *Trends Biochem. Sci.* 20:448-455 (1995). Mutagenesis studies have produced GFP mutants with shifted wavelengths of excitation or emission. See, Heim, R. & Tsien, R.Y. *Current Biol.* 6:178-182 (1996). Suitable pairs, for example a blue-shifted GFP mutant P4-3 (Y66H/Y145F) and an improved green mutant S65T can respectively serve as a donor and an acceptor for fluorescence resonance energy transfer (FRET). See, Tsien, R.Y., et al., *Trends Cell Biol.* 3:242-245 (1993). A fluorescent protein is an *Aequorea*-related fluorescent protein if any contiguous sequence of 150 amino acids of the fluorescent protein has at least 85% sequence identity with an amino acid sequence, either contiguous or non-contiguous, from the wild type *Aequorea* green fluorescent protein. More preferably, a fluorescent protein is an *Aequorea*-related fluorescent protein if any contiguous sequence of 200 amino acids of the fluorescent protein has at least 95% sequence identity with an amino acid sequence, either contiguous or non-contiguous, from the wild type *Aequorea* green fluorescent protein. Similarly, the fluorescent protein

- 17 -

can be related to *Renilla* or *Phialidium* wild-type fluorescent proteins using the same standards. Some *Aequorea*-related engineered versions described in Table I. Other variants or mutants are within the scope of the invention as described, for example, in the Examples.

TABLE I

<u>Clone</u>	<u>Mutation(s)</u>	<u>Excitation max</u> <u>(nm)</u>	<u>Emission</u> <u>max (nm)</u>	<u>Extinction Coefficient</u> <u>(M⁻¹cm⁻¹)</u>	<u>Quantum</u> <u>yield</u>
<i>Wild type</i>	none	395 (475)	508	21,000 (7,150)	0.77
<i>P4</i>	Y66H	383	447	13,500	0.21
<i>P4-3</i>	Y66H;Y145F	381	445	14,000	0.38
<i>W7</i>	Y66W;N146I M153T V163A N212K	433 (453)	475 (501)	18,000 (17,100)	0.67
<i>W2</i>	Y66W;I123V Y145H H148R M153T V163A N212K	432 (453)	480	10,000 (9,600)	0.72
<i>S65T</i>	S65T	489	511	39,200	0.68
<i>P4-1</i>	S65T;M153A K238E	504 (396)	514	14,500 (8,600)	0.53
<i>S65A</i>	S65A	471	504		
<i>S65C</i>	S65C	479	507		
<i>S65L</i>	S65L	484	510		
<i>Y66F</i>	Y66F	360	442		
<i>Y66W</i>	Y66W	458	480		
<i>10c</i>	S65G;V68L S72A;T203Y	513	527		

WO 98/40477

PCT/US98/04978

- 18 -

W1B	F64L;S65T Y66W;N146I M153T V163A N212K	432 (453)	476 (503)
Emerald	S65T;S72A N149K M153T I167T	487	508

- 19 -

<i>Sapphire</i>	S72A;Y145F T203I	395	511
-----------------	---------------------	-----	-----

An additional clone, W1B1 included the following mutations: F64L;S65T; Y66W; F99S; and V163A.

Other fluorescent proteins can be used in the fluorescent indicators, such as, for example, yellow fluorescent protein from *Vibrio fischeri* strain Y-1, Peridinin-chlorophyll *a* binding protein from the dinoflagellate *Symbiodinium* sp. phycobiliproteins from marine cyanobacteria such as *Synechococcus*, e.g., phycoerythrin and phycocyanin, or oat phytochromes from oat reconstructed with phycoerythrobilin. These fluorescent proteins have been described in Baldwin, T.O., *et al.*, *Biochemistry* 29:5509-5515 (1990), Morris, B.J., *et al.*, *Plant Molecular Biology*, 24:673-677 (1994), and Wilbanks, S.M., *et al.*, *J. Biol. Chem.* 268:1226-1235 (1993), and Li *et al.*, *Biochemistry* 34:7923-7930 (1995).

The efficiency of FRET between the donor and acceptor fluorescent protein moieties can be adjusted by changing ability of the two fluorescent proteins to closely associate. The nature of the binding protein moiety, target peptide moiety, and linker moiety each affect the FRET and the response of the indicator to the analyte. Generally, large conformational changes in the binding protein moiety are desired along with a high affinity for the target peptide moiety.

The binding protein moiety is a protein, or part thereof, that changes conformation upon binding an analyte. Proteins that undergo useful conformation change upon binding an analyte include calmodulin (CaM), cGMP-dependent protein kinase, steroid hormone receptors (or ligand binding domain thereof), protein kinase C, inositol-1,4,5-triphosphate receptor, or recoverin. See, for example, Katzenellenbogen, J.A. & Katzenellenbogen, B.S. *Chemistry & Biology* 3:529-536 (1996), and Ames, J.B., *et al.*, *Curr. Opin. Struct. Biol.* 6:432-438 (1996). The binding protein moiety preferably binds target peptides in addition to the analyte. The Ca²⁺-binding affinity of calmodulin can

- 20 -

be tuned as reviewed, for example, in Falke, J.J., *et al.*, *Quart. Rev. Biophys.* 27:219-290 (1994).

The target peptide moiety can contain any of the amino acid sequences in Table II, or a portion thereof with the proviso that the target peptide must bind to the binding protein moiety. The target peptide can be a subsequence of a calmodulin-binding domain. The target peptide moieties listed in Table II are recognized by the binding protein moiety CaM. See, for example, Crivici, A. & Ikura, M. *Annu. Rev. Biophys. Biomol. Struct.* 24:84-116 (1995). The target peptide moiety can be modified to enhance the response of the fluorescent indicator to the analyte. Other target peptide moieties are known in the art for other binding proteins.

- 21 -

TABLE II

Target ^a	Sequence
skMLCK (M13)	KRRWKKNFIAVSAANRFKKISSSGAL
smMLCK (smMLCKp)	ARRKWQKTGHAVRAIGRLSS
5 CaMKII	ARRKLKGAILTTMLATRNFS
Caldesmon	GVRNIKSMWEKGNVFSS
Calspermin	ARRKLKAAVKAVVASSRLGS
PFK (M11)	FMNNWEVYKLLAHIRPPAPKSGSYTV
Calcineurin	ARKEVIRNKIRAIGKMARVFSVLR
10 PhK (PhK5)	LRRLIDAYAFRIYGHVVKKGQQQNRG
(PhK13)	RGKFKVICLTVLASVRIYYQYRRVKPG
Ca ²⁺ -ATPase (C28W)	LRRGQILWFRGLNRIQTQIKVVNAFSSS
59-kDa PDE	RRKHLQRPIFRLRCLVKQLEK
60-kDa PDE	TEKMWQRLKGILRCLVKQLEK
15 NOS (NO-30)	KRRAIGFKKLAEAVKFSAKLMGQ
Type I AC (AC-28)	IKPAKRMKFKTVCYLLVQLMHCRKMFKA
<i>Borderella periussis</i> AC	IDLLWKIARAGARSAVGTEA
Neuromodulin	KAHKAATKIQASFRGHITRKKLKGEKK
Spectrin	KTASPWKSARLMVHTVATFNSIKE
20 MARCKS	KKKKKRFSFKKSFKLSGFSFKKSKK
F52 or MacMARKS	KKKKKFSFKKPFKLSGLSFKRNRK
β-Adducin	KQQKEKTRWLNTPTNTYL RVNVADEVQRNMGS
HSP90a	KDQVANSASFQERLRKHGLEVI
HIV-1 gp160	YHRLRDLLIVKRIVELLGRR
25 BBMHBI	QQLATLIQKTYRGWRCRTHYQLM
Dilute MHC	RAACIRIQKTIRGWLLRKRYLCMQ
Mastoparan	INLKALAALAKKIL
Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ
Glucagon	HSQGTFTTSDYSKYLDSRRAQDFVQWLMNT
30 Secretin	HSDGTFTSELSRLRDSARLQRLQLGLV
VIP	HSDAVFTDNYTRLRKQMAVKKYLNSILN

- 22 -

GIP	YADGTFISDYS AIMNKIRQQDFVNWLLAQQQKS
Model Peptide CBP2	KLWKKLLKLLKLLKLLG

- ^a Abbreviations: AC, adenylyl cyclase; BBMHCI, brush-border myosin heavy chain-I; CaMKII, calmodulin kinase II; CBP2, calmodulin binding peptide-2; GIP, gastrin inhibitory peptide; HIV-1 gp160, human immunodeficiency virus envelope glycoprotein 160; HSP, heat-shock protein; MARCKS, myristoylated alaminte-rich C kinase substrate; MHC, myosin heavy chain; NOS, nitric oxide synthase; PDE, phosphodiesterase; PFK, phosphofructokinase; PhK, phosphorylase kinase; sk-, smMLCK, skeletal muscle- and smooth muscle-myosin light chain kinase; VIP, vasoactive intestinal peptide.
- 10 The length of the linker moiety is chosen to optimize both FRET and the kinetics and specificity of conformational changes induced by analyte binding. The linker moiety should be long enough and flexible enough to allow the binding protein moiety and target peptide moiety to freely interact and respond to analyte concentration. In order to optimize the FRET effect, the average distance between the donor and acceptor
- 15 fluorescent protein moieties should become between about 1 nm and about 10 nm, preferably between about 1 nm and about 6 nm, and more preferably between about 1 nm and about 4 nm, when the analyte is bound (or released). If the linker is too short or too stiff, the donor and acceptor protein moieties cannot readily change position. If the linker moiety is too long, the target peptide moiety might not bind to the binding protein
- 20 moiety effectively. The linker moiety is, preferably, a peptide moiety. The preferred linker moiety is a peptide between about one and 30 amino acid residues in length, preferably between about two and 15 amino acid residues. One preferred linker moiety is a -Gly-Gly- linker.

The linker moiety can include flexible spacer amino acid sequences, such as those known

25 in single-chain antibody research. For example, the linker moiety can be GGGGS (GGGGS)_n, GKSSGSGSESKS, GSTSGSGKSSEGKG, GSTSGSGKSSEGSGSTKG, GSTSGSGKSSEGKG, GSTSGSGKPGSGEGSTKG, or EGKSSGSGSESKF. Linking moieties are described, for example, in Huston, J.S., *et al.*, *PNAS* 85:5879-5883 (1988),

- 23 -

Whitlow, M., *et al.*, *Protein Engineering* 6:989-995 (1993), and Newton, D.L., *et al.*, *Biochemistry* 35:545-553 (1996).

Another embodiment of the invention includes a nucleic acid molecule, comprising a polynucleotide that encodes a fluorescent protein energy transfer partner, comprising: a
5 first chemical entity binding region that binds to a chemical entity and binds to a second protein partner, and a fluorescent protein in frame with the first chemical entity binding region; wherein the fluorescent protein is a first energy transfer partner for a second energy transfer partner attached to the second protein partner. Typically the fluorescent protein is an *Aequorea*-related fluorescent protein.

10 The first chemical entity binding region can comprise a bound conformation of smaller volume when the chemical entity is bound to the first chemical entity binding region compared to the first chemical entity binding region when the chemical entity is not bound to the first chemical entity binding region. The bound conformation can permit increased energy transfer between the first energy transfer partner and the second energy
15 transfer partner.

The bound conformation may also increase the affinity between the first chemical entity binding region and the second protein partner. Often the first chemical entity binding region binds to a protein association region of the second protein partner. In many cases the chemical entity is an analyte, as described herein. The chemical entity could also
20 include other entities such as cellular proteins.

Preferably, the first chemical entity binding region is either calmodulin, a calmodulin-related protein moiety, cGMP-dependent protein kinase, a steroid hormone receptor, a ligand binding domain of a steroid hormone receptor, protein kinase C, inositol-1,4,5-triphosphate receptor, or recoverin. More preferably, the first chemical entity binding
25 region is calmodulin or a calmodulin-related protein.

- 24 -

In some aspects of this embodiment of the invention the fluorescent protein energy transfer partner further comprises a second energy transfer partner in frame with the first energy transfer partner. Such embodiments can be produced by molecular biology techniques as opposed to protein synthesis techniques. Proteins can be produced by either technique to produce fluorescent protein that act as energy transfer partners. For example, the first energy transfer partner is either P4-3, EBFP, or W1B, and the second energy transfer partner is either S65T, EGFP, or 10c.

In addition energy partners can be generated using labelling methods known in the art for tagging proteins with fluorescent labels (e.g., biochemical reagents and non-protein fluorescent tags).

The fluorescent protein energy transfer partner can also further comprise a second protein partner comprising the second energy transfer partner in frame with the first energy transfer partner. The second protein partner can also further comprises a protein association region.

The second protein partner is either a calmodulin-binding domain of skMLCKp, smMLCK, CaMKII, Caldesmon, Calspermin, phosphofructokinase calcineurin, phosphorylase kinase, Ca²⁺-ATPase 59 kDa PDE, 60 kDa PDE, nitric oxide synthase, type I adenylyl cyclase, *Bordetella pertussis* adenylyl cyclase, Neuromodulin, Spectrin, MARCKS, F52, β -Adducin, HSP90a, HIV-1 gp160, BBMHBI, Dilute MHC, Mastoparan, Melittin, Glucagon, Secretin, VIP, GIP, or Model Peptide CBP2. Preferably, the protein association region is M13. The invention can include an amino acid linker in frame with the first chemical entity binding region and fusing the protein association region with the first chemical entity binding region. The invention can also include a fluorescent protein energy transfer partner that further comprises a localization sequence.

- 25 -

Another aspect of the present invention is a nucleic acid molecule, comprising a polynucleotide that encodes 1) a first fluorescent protein energy transfer partner, comprising: a chemical entity binding region that binds to a chemical entity and binds to a second fluorescent protein energy transfer partner, and a first fluorescent protein in frame with the first chemical entity binding region; and 2) the second fluorescent protein energy transfer partner, comprising: a protein association region that binds to the chemical entity binding region, and a second fluorescent protein in frame with the protein association region; wherein the first fluorescent protein and the second fluorescent protein are energy transfer partners and the first fluorescent protein energy transfer partner is in frame with the second first fluorescent protein energy transfer partner. The first and second fluorescent proteins can be *Aequorea*-related fluorescent proteins.

A further aspect of the present invention is this nucleic acid molecule, wherein the chemical entity binding region comprises a bound conformation of smaller volume when the chemical entity is bound to the chemical entity binding region compared to the chemical entity binding region when the chemical entity is not bound to the chemical entity binding region and the bound conformation permits increased energy transfer between the first fluorescent protein energy transfer partner and the second fluorescent protein energy transfer partner. Another aspect of the present invention is this nucleic acid, wherein the bound conformation increases the affinity between the chemical entity binding region and the second fluorescent protein energy transfer partner. The chemical entity is an analyte.

Another aspect of the present invention is a system for monitoring protein-protein association, comprising: 1) a first fluorescent protein energy transfer partner, comprising: a chemical entity binding region that binds to a chemical entity and binds to a second fluorescent protein energy transfer partner, and a first energy transfer partner; and 2) the second fluorescent protein energy transfer partner, comprising: a protein association region that binds to the chemical entity binding region, and a second energy transfer partner; wherein the first energy transfer partner and the energy transfer partner are energy transfer partners.

- 26 -

The fluorescent indicators can also include a localization sequence to direct the indicator to particular cellular sites by fusion to appropriate organellar targeting signals or localized host proteins. A polynucleotide encoding a localization sequence, or signal sequence, can be ligated or fused at the 5' terminus of a polynucleotide encoding the
5 fluorescence indicator such that the signal peptide is located at the amino terminal end of the resulting fusion polynucleotide/polypeptide. In the case of eukaryotes, the signal peptide is believed to function to transport the fusion polypeptide across the endoplasmic reticulum. The secretory protein is then transported through the Golgi apparatus, into secretory vesicles and into the extracellular space or, preferably, the external
10 environment. Signal peptides which can be utilized according to the invention include pre-pro peptides which contain a proteolytic enzyme recognition site. Other signal peptides with similar properties to pro-calcitonin described herein are known to those skilled in the art, or can be readily ascertained without undue experimentation.

The localization sequence can be a nuclear localization sequence, an endoplasmic
15 reticulum localization sequence, a peroxisome localization sequence, a mitochondrial localization sequence, or a localized protein. Localization sequences can be targeting sequences which are described, for example, in "Protein Targeting", chapter 35 of Stryer, L., *Biochemistry* (4th ed.). W.H. Freeman, 1995. The localization sequence can also be a localized protein. Some important localization sequences include those targeting the
20 nucleus (KKKRRK), mitochondrion (amino terminal MLRTSSLFTRRVQPSLFRNLRRLQST-), endoplasmic reticulum (KDEL at C-terminus, assuming a signal sequence present at N-terminus), peroxisome (SKF at C-terminus), prenylation or insertion into plasma membrane (CaaX, CC, CXC, or CCXX at C-terminus), cytoplasmic side of plasma membrane (fusion to SNAP-25), or the Golgi
25 apparatus (fusion to furin).

The fluorescent indicators can be produced as fusion proteins by recombinant DNA technology. Recombinant production of fluorescent proteins involves expressing nucleic acids having sequences that encode the proteins. Nucleic acids encoding fluorescent proteins can be obtained by methods known in the art. For example, a nucleic acid

- 27 -

encoding the protein can be isolated by polymerase chain reaction of cDNA from *A. victoria* using primers based on the DNA sequence of *A. victoria* green fluorescent protein. PCR methods are described in, for example, U.S. Pat. No. 4,683,195; Mullis, *et al. Cold Spring Harbor Symp. Quant. Biol.* 51:263 (1987), and Erlich, ed., *PCR Technology*, (Stockton Press, NY, 1989). Mutant versions of fluorescent proteins can be made by site-specific mutagenesis of other nucleic acids encoding fluorescent proteins, or by random mutagenesis caused by increasing the error rate of PCR of the original polynucleotide with 0.1 mM MnCl₂ and unbalanced nucleotide concentrations. See, *e.g.*, U.S. Patent No. 5,625,048 issued April 29, 1997 or International Publication
5
10 WO 96/23810 published on August 8, 1996.

The construction of expression vectors and the expression of genes in transfected cells involves the use of molecular cloning techniques also well known in the art. Sambrook *et al.*, *Molecular Cloning -- A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, (1989) and *Current Protocols in Molecular Biology*, F.M. Ausubel
15 *et al.*, eds., (1998).

Nucleic acids used to transfect cells with sequences coding for expression of the polypeptide of interest generally will be in the form of an expression vector including expression control sequences operatively linked to a nucleotide sequence coding for
20 expression of the polypeptide. As used, the term "nucleotide sequence coding for expression of" a polypeptide refers to a sequence that, upon transcription and translation of mRNA, produces the polypeptide. This can include sequences containing, *e.g.*, introns. As used herein, the term "expression control sequences" refers to nucleic acid sequences that regulate the expression of a nucleic acid sequence to which it is
25 operatively linked. Expression control sequences are operatively linked to a nucleic acid sequence when the expression control sequences control and regulate the transcription and, as appropriate, translation of the nucleic acid sequence. Thus, expression control sequences can include appropriate promoters, enhancers, transcription terminators, a start codon (*i.e.*, ATG) in front of a protein-encoding gene, splicing signals for introns,

- 28 -

maintenance of the correct reading frame of that gene to permit proper translation of the mRNA, and stop codons.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing the fluorescent indicator coding sequence and appropriate transcriptional/translational control signals. These methods include *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination. (See, for example, the techniques described in Maniatis, *et al.*, *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory, N.Y., 1989).

Transformation of a host cell with recombinant DNA may be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *E. coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the CaCl_2 method by procedures well known in the art. Alternatively, MgCl_2 or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell or by electroporation.

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used. Eukaryotic cells can also be cotransfected with DNA sequences encoding the fusion polypeptide of the invention, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein. (*Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982). Preferably, a eukaryotic host is utilized as the host cell as described herein.

- 29 -

Techniques for the isolation and purification of either microbially or eukaryotically expressed polypeptides of the invention may be by any conventional means such as, for example, preparative chromatographic separations and immunological separations such as those involving the use of monoclonal or polyclonal antibodies or antigen.

- 5 A variety of host-expression vector systems may be utilized to express fluorescent indicator coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing a fluorescent indicator coding sequence; yeast transformed with recombinant yeast expression vectors containing the fluorescent indicator coding sequence; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing a fluorescent indicator coding sequence; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing a fluorescent indicator coding sequence; or animal cell systems infected with recombinant virus expression vectors (e.g., retroviruses, adenovirus, vaccinia virus) containing a fluorescent indicator coding sequence, or transformed animal cell systems engineered for stable expression.

Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, *etc.* may be used in the expression vector (see, e.g., Bitter, *et al.*, Methods in Enzymology 153:516-544, 1987). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage S, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like may be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus 7.5K promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted fluorescent indicator coding sequence.

- 30 -

In bacterial systems a number of expression vectors may be advantageously selected depending upon the use intended for the fluorescent indicator expressed. For example, when large quantities of the fluorescent indicator are to be produced, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Those which are engineered to contain a cleavage site to aid in recovering fluorescent indicator are preferred.

In yeast, a number of vectors containing constitutive or inducible promoters may be used. For a review see, Current Protocols in Molecular Biology, Vol. 2, Ed. Ausubel, *et al.*, Greene Publish. Assoc. & Wiley Interscience, Ch. 13, 1988; Grant, *et al.*, Expression and Secretion Vectors for Yeast, in Methods in Enzymology, Eds. Wu & Grossman, 31987, Acad. Press, N.Y., Vol. 153, pp.516-544, 1987; Glover, DNA Cloning, Vol. II, IRL Press, Wash., D.C., Ch. 3, 1986; and Bitter, Heterologous Gene Expression in Yeast, Methods in Enzymology, Eds. Berger & Kimmel, Acad. Press, N.Y., Vol. 152, pp. 673-684, 1987; and The Molecular Biology of the Yeast *Saccharomyces*, Eds. Strathern *et al.*, Cold Spring Harbor Press, Vols. I and II, 1982. A constitutive yeast promoter such as ADH or LEU2 or an inducible promoter such as GAL may be used (Cloning in Yeast, Ch. 3, R. Rothstein In: DNA Cloning Vol.11, A Practical Approach, Ed. DM Glover, IRL Press, Wash., D.C., 1986). Alternatively, vectors may be used which promote integration of foreign DNA sequences into the yeast chromosome.

In cases where plant expression vectors are used, the expression of a fluorescent indicator coding sequence may be driven by any of a number of promoters. For example, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV (Brisson, *et al.*, *Nature* 310:511-514, 1984), or the coat protein promoter to TMV (Takamatsu, *et al.*, *EMBO J.* 6:307-311, 1987) may be used; alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi, *et al.*, 1984, *EMBO J.* 3:1671-1680; Broglie, *et al.*, *Science* 224:838-843, 1984); or heat shock promoters, *e.g.*, soybean hsp17.5-E or hsp17.3-B (Gurley, *et al.*, *Mol. Cell. Biol.* 6:559-565, 1986) may be used. These constructs can be introduced into plant cells using Ti plasmids, Ri plasmids, plant virus vectors,

- 31 -

direct DNA transformation, microinjection, electroporation, *etc.* For reviews of such techniques see, for example, Weissbach & Weissbach, *Methods for Plant Molecular Biology*, Academic Press, NY, Section VIII, pp. 421-463, 1988; and Grierson & Corey, *Plant Molecular Biology*, 2d Ed., Blackie, London, Ch. 7-9, 1988.

- 5 An alternative expression system which could be used to express fluorescent indicator is an insect system. In one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The fluorescent indicator coding sequence may be cloned into non-essential regions (for example, the polyhedrin gene) of the virus and placed
10 under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the fluorescent indicator coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed,
15 see Smith, *et al.*, *J. Viol.* 46:584, 1983; Smith, U.S. Patent No. 4,215,051.

Eukaryotic systems, and preferably mammalian expression systems, allow for proper post-translational modifications of expressed mammalian proteins to occur. Eukaryotic cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, phosphorylation, and, advantageously secretion of the gene
20 product should be used as host cells for the expression of fluorescent indicator. Such host cell lines may include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, Jurkat, HEK-293, and WI38.

Mammalian cell systems which utilize recombinant viruses or viral elements to direct expression may be engineered. For example, when using adenovirus expression vectors,
25 the fluorescent indicator coding sequence may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of

- 32 -

expressing the fluorescent indicator in infected hosts (*e.g.*, see Logan & Shenk, *Proc. Natl. Acad. Sci. USA*, 81: 3655-3659, 1984). Alternatively, the vaccinia virus 7.5K promoter may be used. (*e.g.*, see, Mackett, *et al.*, *Proc. Natl. Acad. Sci. USA*, 79: 7415-7419, 1982; Mackett, *et al.*, *J. Virol.* 49: 857-864, 1984; Panicali, *et al.*, *Proc. Natl. Acad. Sci. USA* 79: 4927-4931, 1982). Of particular interest are vectors based on bovine papilloma virus which have the ability to replicate as extrachromosomal elements (Sarver, *et al.*, *Mol. Cell. Biol.* 1: 486, 1981). Shortly after entry of this DNA into mouse cells, the plasmid replicates to about 100 to 200 copies per cell. Transcription of the inserted cDNA does not require integration of the plasmid into the host's chromosome, thereby yielding a high level of expression. These vectors can be used for stable expression by including a selectable marker in the plasmid, such as the *neo* gene. Alternatively, the retroviral genome can be modified for use as a vector capable of introducing and directing the expression of the fluorescent indicator gene in host cells (Cone & Mulligan, *Proc. Natl. Acad. Sci. USA*, 81:6349-6353, 1984). High level expression may also be achieved using inducible promoters, including, but not limited to, the metallothionine IIA promoter and heat shock promoters.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the fluorescent indicator cDNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, *etc.*), and a selectable marker. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, *Cell*, 11: 223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, *Proc. Natl. Acad. Sci. USA*, 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, *et al.*, *Cell*, 22: 817, 1980) genes

- 33 -

can be employed in tk^- , $hgprt^-$ or $aprt^-$ cells respectively. Also, antimetabolite resistance can be used as the basis of selection for $dhfr$, which confers resistance to methotrexate (Wigler, *et al.*, *Proc. Natl. Acad. Sci. USA*, 77: 3567, 1980; O'Hare, *et al.*, *Proc. Natl. Acad. Sci. USA*, 8: 1527, 1981); gpt , which confers resistance to mycophenolic acid
5 (Mulligan & Berg, *Proc. Natl. Acad. Sci. USA*, 78: 2072, 1981; neo , which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, *J. Mol. Biol.*, 150:1, 1981); and $hygro$, which confers resistance to hygromycin (Santerre, *et al.*, *Gene*, 30: 147, 1984) genes. Recently, additional selectable genes have been described, namely $trpB$, which allows cells to utilize indole in place of tryptophan; $hisD$, which allows cells
10 to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. USA*, 85:8047, 1988); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue L., In: *Current Communications in Molecular Biology*, Cold Spring Harbor Laboratory, ed., 1987).

15 DNA sequences encoding the fluorescence indicator polypeptide of the invention can be expressed *in vitro* by DNA transfer into a suitable host cell. "Host cells" are cells in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However,
20 such progeny are included when the term "host cell" is used. Methods of stable transfer, in other words when the foreign DNA is continuously maintained in the host, are known in the art.

Recombinant fluorescent protein can be produced by expression of nucleic acid encoding the protein in prokaryotes, such as *E. coli* or in eukaryotes, such as yeast cells or
25 mammalian cells. The fluorophore of *Aequorea*-related fluorescent proteins results from cyclization and oxidation of residues 65-67.

- 34 -

The construct can also contain a tag to simplify isolation of the fluorescent indicator. For example, a polyhistidine tag of, *e.g.*, six histidine residues, can be incorporated at the amino terminal end of the fluorescent protein. The polyhistidine tag allows convenient isolation of the protein in a single step by nickel-chelate chromatography.

- 5 In a preferred embodiment, the fluorescent indicator is a fusion protein produced by recombinant DNA technology in which a single polypeptide includes a donor moiety, a peptide linker moiety and an acceptor moiety. The donor moiety can be positioned at the amino-terminus relative to the acceptor moiety in the polypeptide. Such a fusion protein has the generalized structure: (amino terminus) donor fluorescent protein moiety
10 --peptide linker moiety --acceptor fluorescent protein moiety (carboxy terminus). Alternatively, the donor moiety can be positioned at the carboxy-terminus relative to the acceptor moiety within the fusion protein. Such a fusion protein has the generalized structure: (amino terminus) acceptor fluorescent protein moiety --peptide linker moiety --donor fluorescent protein moiety (carboxy terminus). The invention also envisions
15 fusion proteins that contain extra amino acid sequences at the amino and/or carboxy termini, for example, polyhistidine tags.

Thus, fluorescent indicators encoded by a recombinant nucleic acid include sequences coding for expression of a donor fluorescent protein moiety, an acceptor fluorescent protein moiety and a peptide linker moiety. The elements are selected so that upon
20 expression into a fusion protein, the donor and acceptor moieties exhibit FRET when the donor moiety is excited. The recombinant nucleic acid can be incorporated into an expression vector comprising expression control sequences operatively linked to the recombinant nucleic acid. The expression vector can be adapted for function in prokaryotes or eukaryotes by inclusion of appropriate promoters, replication sequences,
25 markers, *etc.*

The expression vector can be transfected into a host cell for expression of the recombinant nucleic acid. Host cells can be selected for high levels of expression in order to purify the fluorescent indicator fusion protein. *E. coli* is useful for this purpose.

- 35 -

Alternatively, the host cell can be a prokaryotic or eukaryotic cell selected to study the activity of an enzyme produced by the cell. In this case, the linker peptide is selected to include an amino acid sequence recognized by the protease. The cell can be, *e.g.*, a cultured cell or a cell *in vivo*.

- 5 A primary advantage of fluorescent indicator fusion proteins is that they are prepared by normal protein biosynthesis, thus completely avoiding organic synthesis and the requirement for customized unnatural amino acid analogs. The constructs can be expressed in *E. coli* in large scale for *in vitro* assays. Purification from bacteria is simplified when the sequences include polyhistidine tags for one-step purification by
10 nickel-chelate chromatography. Alternatively, the substrates can be expressed directly in a desired host cell for assays *in situ*.

In another embodiment, the invention provides a transgenic non-human animal that expresses a nucleic acid sequence which encodes the fluorescent indicator.

- The "non-human animals" of the invention comprise any non-human animal having
15 nucleic acid sequence which encodes a fluorescent indicator. Such non-human animals include vertebrates such as rodents, non-human primates, sheep, dog, cow, pig, amphibians, and reptiles. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse. The "transgenic non-human animals" of the invention are produced by introducing "transgenes" into the germline of
20 the non-human animal. Embryonal target cells at various developmental stages can be used to introduce transgenes. Different methods are used depending on the stage of development of the embryonal target cell. The zygote is the best target for micro-injection. In the mouse, the male pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2 pl of DNA solution.
25 The use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host genome before the first cleavage (Brinster *et al.*, *Proc. Natl. Acad. Sci. USA* 82:4438-4442, 1985). As a consequence, all cells of the transgenic non-human animal will carry the incorporated transgene. This will

- 36 -

in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene. Microinjection of zygotes is the preferred method for incorporating transgenes in practicing the invention.

The term "transgenic" is used to describe an animal which includes exogenous genetic material within all of its cells. A "transgenic" animal can be produced by cross-breeding two chimeric animals which include exogenous genetic material within cells used in reproduction. Twenty-five percent of the resulting offspring will be transgenic *i.e.*, animals which include the exogenous genetic material within all of their cells in both alleles. 50% of the resulting animals will include the exogenous genetic material within one allele and 25% will include no exogenous genetic material.

Retroviral infection can also be used to introduce transgene into a non-human animal. The developing non-human embryo can be cultured *in vitro* to the blastocyst stage. During this time, the blastomeres can be targets for retro viral infection (Jaenich, R., *Proc. Natl. Acad. Sci USA* 73:1260-1264, 1976). Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida (Hogan, *et al.* (1986) in *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). The viral vector system used to introduce the transgene is typically a replication-defective retro virus carrying the transgene (Jahner, *et al.*, *Proc. Natl. Acad. Sci. USA* 82:6927-6931, 1985; Van der Putten, *et al.*, *Proc. Natl. Acad. Sci USA* 82:6148-6152, 1985). Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells (Van der Putten, *supra*; Stewart, *et al.*, *EMBO J.* 6:383-388, 1987). Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele (D. Jahner *et al.*, *Nature* 298:623-628, 1982). Most of the founders will be mosaic for the transgene since incorporation occurs only in a subset of the cells which formed the transgenic nonhuman animal. Further, the founder may contain various retro viral insertions of the transgene at different positions in the genome which generally will segregate in the offspring. In

- 37 -

addition, it is also possible to introduce transgenes into the germ line, albeit with low efficiency, by intrauterine retro viral infection of the midgestation embryo (D. Jahner *et al.*, *supra*).

A third type of target cell for transgene introduction is the embryonal stem cell (ES). ES
5 cells are obtained from pre-implantation embryos cultured *in vitro* and fused with
embryos (M. J. Evans *et al.* *Nature* 292:154-156, 1981; M.O. Bradley *et al.*, *Nature* 309:
255-258, 1984; Gossler, *et al.*, *Proc. Natl. Acad. Sci USA* 83: 9065-9069, 1986; and
Robertson *et al.*, *Nature* 322:445-448, 1986). Transgenes can be efficiently introduced
into the ES cells by DNA transfection or by retro virus-mediated transduction. Such
10 transformed ES cells can thereafter be combined with blastocysts from a nonhuman
animal. The ES cells thereafter colonize the embryo and contribute to the germ line of
the resulting chimeric animal. (For review see Jaenisch, R., *Science* 240: 1468-1474,
1988).

"Transformed" means a cell into which (or into an ancestor of which) has been
15 introduced, by means of recombinant nucleic acid techniques, a heterologous nucleic acid
molecule. "Heterologous" refers to a nucleic acid sequence that either originates from
another species or is modified from either its original form or the form primarily
expressed in the cell.

"Transgene" means any piece of DNA which is inserted by artifice into a cell, and
20 becomes part of the genome of the organism (*i.e.*, either stably integrated or as a stable
extrachromosomal element) which develops from that cell. Such a transgene may
include a gene which is partly or entirely heterologous (*i.e.*, foreign) to the transgenic
organism, or may represent a gene homologous to an endogenous gene of the organism.
Included within this definition is a transgene created by the providing of an RNA
25 sequence which is transcribed into DNA and then incorporated into the genome. The
transgenes of the invention include DNA sequences which encode which encodes the
fluorescent indicator which may be expressed in a transgenic non-human animal. The
term "transgenic" as used herein additionally includes any organism whose genome has

- 38 -

been altered by *in vitro* manipulation of the early embryo or fertilized egg or by any transgenic technology to induce a specific gene knockout. The term "gene knockout" as used herein, refers to the targeted disruption of a gene *in vivo* with complete loss of function that has been achieved by any transgenic technology familiar to those in the art.

5 In one embodiment, transgenic animals having gene knockouts are those in which the target gene has been rendered nonfunctional by an insertion targeted to the gene to be rendered non-functional by homologous recombination. As used herein, the term "transgenic" includes any transgenic technology familiar to those in the art which can produce an organism carrying an introduced transgene or one in which an endogenous

10 gene has been rendered non-functional or "knocked out."

Without further elaboration, it is believed that one skilled in the art can, based on the description herein, utilize the present invention to its fullest extent.

A fluorescent indicator for Ca^{2+} was produced by sandwiching CaM-M13 fusion, described in Porumb, T., *et al.*, *Prot. Engineering* 7:109-115 (1994), between a blue (P4-3) and a green (S65T) GFP mutant, as illustrated in FIG. 2a. The chimeric cDNA was cloned at *Bam*HI/*Eco*RI sites of pRSETB (Invitrogen), HIS is the amino-terminal tag peptide containing the polyhistidine sequence and XCaM is *Xenopus* calmodulin. Chimeric proteins incorporating a polyhistidine tag were expressed in *Escherichia coli*, purified by

15 nickel-chelate and size-exclusion chromatography, and their fluorescence characterized. Referring to FIG. 2a, the fluorescent CaM-based indicator ("cameleon-1") readily changes emission color by retracting and extending a long tongue (M13) into and out of the mouth of the CaM.

20

The amino acid composition of the boundary regions between the CaM-M13 hybrid and GFPs can be important to optimize protein folding and the Ca^{2+} -sensitivity of FRET.

25 One particularly sensitive indicator is shown in FIG. 2b. Referring to FIG. 2b, the amino acid and nucleotide sequences of the boundaries between P4-3 and CaM and between M13 and S65T in cameleon-1 are shown. Cameleon-1 has an 11 amino acid deletion at the C-terminus of P4-3 and a 5 amino acid deletion at the C-terminus of M13. Two

- 39 -

restriction sites *Sph*I and *Sac*I are underlined, which were utilized to connect the genes of P4-3 and S65T to the CaM-M13 gene, respectively. To facilitate the folding of the GFP that is fused to any other protein, a few glycine residues are usually inserted into the boundary. See, for example, Porumb, T., *et al.*, *Prot. Engineering* 7:109-115 (1994).
5 Further, glycine residues were not introduced so that P4-3 and S65T were fused rigidly to the CaM and M13, respectively. The rigid fusion leads to more effective transduction of the conformational change of CaM-M13, causing a greater change in FRET efficiency.

The fluorescent indicator was efficiently expressed and folded in bacteria and increased its ratio of UV-excited 510 nm to 445 nm emissions by 70% upon binding Ca^{2+} , as shown
10 in FIG. 3. The emission spectra of cameleon-1 were measured in 100 mM KCl, 20 mM MOPS, 1 mM EDTA, KOH to pH 7.20, before and after addition of 2 mM CaCl_2 to give 1 mM free Ca^{2+} . The Ca^{2+} binding to EDTA caused a local acidification of the solution, and a small fraction of the protein was denatured. Thus the spectrum after the Ca^{2+} addition dropped down slightly (compare with FIG. 4a). The decrease in blue and
15 increase in green emission indicated that Ca^{2+} increased the efficiency of FRET from P4-3 to S65T, consistent with the expected decrease in distance between the two ends of the protein. The Ca^{2+} response was fully reversible upon chelation of Ca^{2+} .

The Ca^{2+} -specificity of the response of cameleon-1 was examined. Mg^{2+} , pH, and ionic strength did not alter the emission spectra of either the Ca^{2+} -saturated and
20 Ca^{2+} -unsaturated forms. The emission spectra of saturated and unsaturated cameleon-1 were also not affected by hydrophobic proteins such as bovine serum albumin. Isolated CaM saturated with Ca^{2+} typically becomes sticky with hydrophobic amino acids exposed to the surface. The CaM in cameleon-1, on the other hand, appears to interact preferentially with its intramolecularly-adjacent M13 peptide. The self-contained nature
25 of the system minimizes the possibility that the protein might interact with endogenous CaM-binding sequences in eukaryotic cells.

- 40 -

The Ca^{2+} binding behavior of cameleon-1 was examined. The CaM-M13 hybrid protein without GFPs displayed a biphasic Ca^{2+} binding with two dissociation constants (80 nM and 2 μM). See, Porumb, T., *et al.*, *Prot. Engineering* 7:109-115 (1994). Titration experiments revealed that the emission ratio of cameleon-1 has a biphasic Ca^{2+} dependency, as shown in FIGS. 4a and 4b. FIG. 4a shows the change in emission spectrum of cameleon-1 on titration with Ca^{2+} when excited at 380 nm. The titration was done using KHHEDTA/KCaHEDTA solutions at pH 7.47. For clarity only two intermediate concentrations of Ca^{2+} are shown.

FIG. 4b shows Ca^{2+} titration curves of cameleon-1 (open circles) and cameleon-1/E104Q (solid circles). Data points were from 4 independent experiments at different pH using Ca^{2+} /EGTA and Ca^{2+} /HEEDTA systems for each protein. In each experiment, the emission ratio (510/445 nm) change was normalized to the value of full saturation with Ca^{2+} , which increased by 60-75% over the value of zero Ca^{2+} . The data of cameleon-1/E104Q were fitted to a four parameter logistic function curve (dotted line). The data of cameleon-1 were analysed using a linear combination of 2 four parameter logistic function fits (solid line). The apparent dissociation constants (K_d 's) for cameleon-1 were 68 nM and 11 μM , and the Hill coefficients were 1.8 and 1.0, respectively. The binding curve can be used to quantify the concentration of Ca^{2+} present in the sample. Because of simulated negative cooperativity, cameleon-1 covers a very wide range of Ca^{2+} concentration, from $<10^{-7}$ to $\sim 10^{-3}$ M.

The affinity of the CaM binding protein moiety can be modified. Many site-directed mutations have been studied for their effects on the Ca^{2+} binding and Ca^{2+} -induced conformational changes of CaM. See, Maune, J.F., *et al.*, *J. Biol. Chem.* 267:5286-5295 (1992), and Gao, Z.H., *et al.*, *J. Biol. Chem.* 268:20096-20104 (1993). For example, a mutant chimera protein with the conserved bidentate glutamic acid at position 12 of the third Ca^{2+} binding loop of the CaM mutated to glutamine (cameleon-1/E104Q) was constructed. The mutation eliminated the high-affinity response of cameleon-1, as indicated in FIG. 4b, (solid circles). Cameleon-1/E104Q showed a monophasic response (K_d , 4.4 μM ; Hill coefficient, 0.76), which corresponds closely to the low affinity

component of the cameleon-1 FRET response. Other modifications of CaM can be made to tune the Ca^{2+} affinities for particular applications of the fluorescent indicator.

HeLa cells were transfected with the recombinant plasmid (cameleon-1/pCDNA3) to determine whether cameleon-1 can work as a Ca^{2+} indicator in live cells. When the cells
5 were excited with UV, however, the fluorescence of cameleon-1 was hardly detectable, mainly because of the dim fluorescence of the P4-3 component. Expression and folding of GFP at 37°C in mammalian cells was improved by introducing mammalian codon bias into the cDNA and mutating Phe64 to Leu, as in the commercially available construct "EGFP" (Clontech), which encodes F64L/S65T with mammalian codons.
10 The same changes were introduced into P4-3 (Y66H/Y145F), which did not change its final fluorescence properties but did improve expression in the HeLa cells. The improved blue mutant ("EBFP") and EGFP substituted P4-3 and S65T, respectively, in cameleon-1 to make cameleon-2, shown in FIG. 5, where Kz is Kozak's consensus sequence (M. Kozak, J. Cell Biol. 108:229-241 (1989). The nucleotide sequence (SEQ
15 ID NO:1) and amino acid sequence (SEQ ID NO:2) of cameleon-2 are shown in FIG. 7.

The E104Q mutation which afforded low Ca^{2+} affinity in cameleon-1 was also introduced into cameleon-2; the resulting chimera protein (cameleon-2/E104Q) cameleon-3, also shown in FIG. 5. The nucleotide sequence (SEQ ID NO:5) and amino acid sequence (SEQ ID NO:6) of cameleon-3 are shown in FIG. 9.

20 Significant emission signals were observed from both the EBFP and EGFP in the cells expressing cameleon-2 or cameleon-3. Referring to FIG. 6a, fluorescence of cameleon-2 localized to the cytosol of HeLa cells. The fluorescence was uniformly distributed in the cytosolic compartment but excluded from the nucleus, as expected for a protein of 74 kDa without localization sequences or targeting signals. The image was taken using a
25 330WB80 (excitation filter) and a 535DF45 (emission filter). The bar is 10 μM .

FIG. 6b shows time courses of the spatially averaged green:blue emission ratios from two individual HeLa cells expressing cameleon-2. The two cells shown in FIG. 6a were

- 42 -

excited by UV (330WB80) and monitored every 15 seconds by digital imaging microscopy. See, for example, Tsien, R.Y. & Harootunian, A.T. *Cell Calcium*, 11:93-109 (1990). The emission bands (440DF40 and 535DF45) over the cytoplasmic regions were alternately sampled. *In situ* calibration was performed for each of the cells. The
5 pre-stimulation ratio (arrowhead) was assumed to be 50 nM, and $R_{\max}$ (arrow) the value after saturation with Ca^{2+} . The calculated values for pCas 7, 6, 5, 4 are indicated by horizontal bars on the right side of the panel.

Elevation of cytosolic Ca^{2+} concentration by saturating doses of ATP (as a purinergic agonist) and histamine produced significant increases in the emission ratio. Blockage
10 of the histamine receptor by the antagonist cyproheptadine caused a rapid decrease in ratio, indicating the reversible behavior of the indicator. Addition of ionomycin followed by a high concentration (15 mM) of extracellular Ca^{2+} gave a large increase of the ratio (70-80% increase of the initial ratio value), which should correspond to the maximal ratio $R_{\max}$. Assuming the lowest ratio observed before stimulation represents the $R_{\min}$,
15 calibration for free Ca^{2+} concentration can be performed. See, Adams, S.R., *et al.*, in *Fluorescent and Luminescent Probes for Biological Activity* (ed. Mason, W.T.) (Academic Press, 1993).

By contrast, cameleon-3, which lacks the high affinity component of Ca^{2+} binding to CaM, did not detect the changes in cytosolic Ca^{2+} concentration signals due to ATP or
20 histamine, but gave a similar $R_{\max}$ in response to ionomycin and 20 mM extracellular Ca^{2+} . Cameleon-3 is less sensitive to and buffers cytosolic Ca^{2+} to a lesser extent than does cameleon-2. The probable high Ca^{2+} dissociation rate of cameleon-3 is advantageous for tracking rapid Ca^{2+} release kinetics. The *in vitro* study revealed that the cameleon indicators show a relatively fast cellular response to Ca^{2+} concentration
25 changes.

Addition of a nuclear localization sequence to cameleon-2 yielded a Ca^{2+} indicator, cameleon-2nu, shown in FIG. 5. The nucleotide sequence (SEQ ID NO:3) and amino acid sequence (SEQ ID NO:4) of cameleon-2nu are shown in FIG. 8. The fluorescence

of cameleon-2nu was localized to nuclei, as depicted in FIG. 6c. The time course of Ca^{2+} concentrations in nuclei was followed (FIG. 6d) and was similar to the results obtained in the cytosol (FIG. 6b). Agonist-induced changes in the free Ca^{2+} concentration inside the endoplasmic reticulum in intact cells were similarly monitored. The low-affinity indicator cameleon-3 was engineered to reside in the lumen of endoplasmic reticulum (ER) (cameleon-3er) (FIG. 5) by addition of a localization sequence at the amino terminus and a KDEL signal for ER retention at the carboxy terminus of the fluorescent indicator. The nucleotide sequence (SEQ ID NO:7) and amino acid sequence of cameleon-3er (SEQ ID NO:8) are shown in FIG. 10. Reticular patterns of fluorescence were seen in HeLa cells expressing the protein. FIG. 6e is a digital fluorescence image of cameleon-3er in transfected HeLa cells. The image was obtained with a cooled CCD camera system with a 480DF30 (excitation filter) and a 535DF45 (emission filter). The bar is 10 μM .

FIG. 6f shows a time course of the average $\text{Ca}^{2+}_{\text{er}}$ concentration of four cells obtained with a video-rate confocal microscope. Digital fluorescence images were the result of simultaneous acquisition of two confocal single-wavelength emission images at 450 nm (65 nm bandpass) and 535 nm (45 nm bandpass). After background subtraction, the ratio of the long wavelength image over the short wavelength one was calculated. Cells were illuminated for 66 msec (2 frames) for each time point. The interference filters and dichroics, and the sensitivity of the detectors are different between the CCD and confocal microscope systems. Therefore, the ratios obtained in the two systems differ quantitatively.

The pre-stimulus $\text{Ca}^{2+}_{\text{er}}$ concentration reported by cameleon-3er was consistently higher than cytosolic or nuclear Ca^{2+} concentrations reported by cameleon-2 and cameleon-2nu, respectively. Histamine reproducibly decreased the $\text{Ca}^{2+}_{\text{er}}$ concentration in all of 15 cells, whereas it always increased the cytosolic and nuclear Ca^{2+} concentrations. Receptor blockade by cyproheptadine reversed the decrease $\text{Ca}^{2+}_{\text{er}}$ concentration, indicating refill of the Ca^{2+} pools. In FIG. 6f, the ratio did not reach the value of the resting state, whereas complete reversion was observed in five other experiments. Addition of

- 44 -

ionomycin and 20 mM extracellular Ca^{2+} increased $\text{Ca}^{2+}_{\text{er}}$ concentration to a saturating value above the starting level. Changes in $\text{Ca}^{2+}_{\text{er}}$ concentration were generally slower than those of the cytosol and nuclei.

The targetability of the fluorescent indicators can permit Ca^{2+} measurements at previously inaccessible sites such as the immediate vicinity of synaptic vesicles or Ca^{2+} channels, and in genetically tractable organisms that do not load well with esters of small-molecule indicators.

The Examples described below are illustrative of the disclosed method; however, many alternatives, modifications and variations will be clear to those skilled in the art.

10 Examples

Gene construction

The cDNA of the GFP mutant P4-3 was amplified by PCR with a sense primer containing a *Bam*HI site, and a reverse primer containing an *Sph*I site and eliminating the GFP stop codon. See, for example, Heim, R. & Tsien, R.Y. *Current Biol.* 6:178-182 (1996). Likewise, the cDNA of S65T was amplified with a *Sac*I site and an *Eco*RI site introduced to the 5' and 3' ends of the gene, respectively. Two restriction sites (*Sph*I and *Sac*I) were introduced by PCR into 5' and 3' ends of the CaM-M13 gene, respectively, using the pHY1 as a template. See, Porumb, T., *et al.* *Prot. Engineering* 7:109-115 (1994). All the amplification reactions were done by Pfu polymerase (Stratagene). The restricted products were ligated and cloned in-frame into the *Bam*HI/*Eco*RI sites of pRSETB (Invitrogen). The modifications of the boundary regions between P4-3 and CaM and between M13 and S65T were performed by PCR or by a combined use of restriction enzymes, Klenow fragment of DNA polymerase I, T4 DNA polymerase, mung bean exonuclease, and T4 DNA ligase as described, for example, in *Molecular Cloning, A Laboratory Manual* (eds. Sambrook, J., Fritsch, E.F. & Maniatis, T.) (CSH Laboratory Press, 1989). The phEGFP plasmid was commercially available from Clontech. Two amino acid substitutions (Y66H and Y145F) were made in hEGFP to construct EBFP. Oligonucleotide-directed mutageneses were carried out using the Muta-Gene Phagemid

in vitro kit (Bio-Rad) at the codons for Y66H and Y145F of EGFP, and for E104Q of the calmodulin. The 5' end of the EBFP gene was modified by PCR to have a *HindIII* restriction site followed by a Kozak's consensus sequence (ACCGCC-ATG). The *HindIII/EcoRI* fragment encoding the entire chimeric protein was subcloned in the
 5 mammalian expression vector pCDNA3 (Invitrogen).

For cameleon-2nu, the cameleon-2 DNA was extended by PCR at the 3' end with the sequence encoding the nuclear localization sequence (PKKKRKVEDP). See, Forbes, D.J. *Ann.Rev.Cell Biol.* 8:495-527 (1992). Cameleon-3er DNA was likewise obtained by extending the cameleon-3 DNA at the 5' end with the sequence encoding the
 10 localization sequence peptide from calreticulin (MLLPVPLLLGLLGLAAAD), and at the 3' end with the sequence encoding the ER retention signal (KDEL). See, Kendall, J.M. *et al.*, *Biochem. Biophys. Res. Commun.* 189:1008-1016 (1992).

Protein expression and spectroscopy

The expression of chimera proteins in bacteria was performed using the T7 expresson
 15 system (pRSETB/JM109(DE3)). Cultures were grown at room temperature, and protein expression was induced by isopropyl B-D-thiogalactoside. Cells were lysed by a French press. The polyhistidine-tagged chimera proteins were purified from the cleared lysates on nickel-chelate columns (Qiagen). The protein samples in the eluates were concentrated by Centricon™ 30 (Amicon), and were further purified by gel-filtration
 20 column to remove abortive chimera proteins which resulted from proteolysis or misfolding. Emission spectra of the purified proteins were measured using a fluorometer (Spex Industries, Edison, NJ) at excitation 380 nm.

Ca²⁺ titration and calibration

The titration experiments were performed by the "pH-metric method" as described in
 25 Grzegorz, G., *et al.*, *J. Biol. Chem.* 260:3440-3450 (1985). *In situ* calibration for cytosolic Ca²⁺ concentration utilized the equation:

$$[\text{Ca}^{2+}]_c = K'_d \left((R - R_{\min}) / (R_{\max} - R) \right)^{1/nH}$$

where K'_d is the apparent dissociation constant corresponding to the Ca^{2+} concentration at which R is midway between $R_{\max}$ and $R_{\min}$, and nH is the Hill coefficient.

Imaging

Two to five days after the cDNA transfection with lipofectin (Gibco BRL), the cells were
5 imaged on a Zeiss AxiovertTM microscope with a cooled CCD camera (Photometrics, Tucson, AZ) interfaced to a personal computer. The program MetaFluor 2.75 (Universal Imaging) was used for controlling data acquisition and analysis. Dual-emission ratio imaging was carried out by manually switching the two emission filters (440DF40 for EBFP, 535DF45 for EGFP) in front of a single imaging camera. The excitation filter
10 (330WB80) was used with a 420DRLP dichroic mirror. Digital fluorescence imaging with a video-rate confocal microscope was performed as described in Tsien, R.Y. & Backskai, B.J. *Handbook of Biological Confocal Microscopy* (ed. Pawley, J.B.) (Plenum Press, New York, 1995) p. 459-478. Cells were illuminated with wide band UV (351-364 nm) from an Ar^+ ion laser. The primary dichroic (380DRLP) reflects UV and
15 transmits light emitted from the specimen, which is subsequently split by a secondary dichroic (505DRLP) into two broad bands: EBFP emission (450DF65) and EGFP emission (535DF45), and counted by photomultiplier tubes.

-47-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: The Regents of the University of California
- (ii) TITLE OF INVENTION: Fluorescent Protein Sensors for
Detection of Analytes
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MBM & CO.
 - (B) STREET: P.O. BOX 809, STATION B
 - (C) CITY: OTTAWA
 - (D) PROVINCE: ONTARIO
 - (E) COUNTRY: CANADA
 - (F) POSTAL CODE: K1P 5P9
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 2,283,629
 - (B) FILING DATE: 13-MAR-1998
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/818,252
 - (B) FILING DATE: 14 March 1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER US 08/818,253
 - (B) FILING DATE: 14 MARCH 1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/919,143
 - (B) FILING DATE: 27 AUGUST 1997
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Swain, Margaret
 - (B) REGISTRATION NUMBER: 10926
 - (C) REFERENCE/DOCKET NUMBER: 581-122
- (ix) TELECOMMUNICATION INFORMATION:

-48-

(A) TELEPHONE: 613/567-0762

(B) TELEFAX: 613/563-7671

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1929 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
				5					10					15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
			35				40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
CTG	ACC	CAT	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	His	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65					70				75						80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85				90						95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	432
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
	130					135					140					

-49-

AAC	TTC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165					170					175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
		195					200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
	210					215				220						
GTG	ACC	GCC	GCC	CGC	ATG	CAT	GAC	CAA	CTG	ACA	GAA	GAG	CAG	ATT	GCA	720
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala	
225					230					235					240	
GAG	TTC	AAA	GAA	GCC	TTC	TCA	TTA	TTC	GAC	AAG	GAT	GGG	GAC	GGC	ACC	768
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr	
				245					250					255		
ATC	ACC	ACA	AAG	GAA	CTT	GGC	ACC	GTT	ATG	AGG	TCG	CTT	GGA	CAA	AAC	816
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn	
			260					265					270			
CCA	ACG	GAA	GCA	GAA	TTG	CAG	GAT	ATG	ATC	AAT	GAA	GTC	GAT	GCT	GAT	864
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp	
		275					280					285				
GGC	AAT	GGA	ACG	ATT	TAC	TTT	CCT	GAA	TTT	CTT	ACT	ATG	ATG	GCT	AGA	912
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg	
	290					295					300					
AAA	ATG	AAG	GAC	ACA	GAC	AGC	GAA	GAG	GAA	ATC	CGA	GAA	GCA	TTC	CGT	960
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg	
305					310					315					320	
GTT	TTT	GAC	AAG	GAT	GGG	AAC	GGC	TAC	ATC	AGC	GCT	GCT	GAA	TTA	CGT	1008
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Glu	Leu	Arg	
				325					330					335		
CAC	GTC	ATG	ACA	AAC	CTC	GGG	GAG	AAG	TTA	ACA	GAT	GAA	GAA	GTT	GAT	1056
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp	
			340					345					350			

-50-

GAA	ATG	ATA	AGG	GAA	GCA	GAT	ATC	GAT	GGT	GAT	GGC	CAA	GTA	AAC	TAT	1104
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr	
		355					360					365				
GAA	GAG	TTT	GTA	CAA	ATG	ATG	ACA	GCA	AAG	GGG	GGG	AAG	AGG	CGC	TGG	1152
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp	
	370					375					380					
AAG	AAA	AAC	TTC	ATT	GCC	GTC	AGC	GCT	GCC	AAC	CGG	TTC	AAG	AAG	ATC	1200
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile	
385					390					395					400	
TCC	GAG	CTC	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	1248
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	
				405					410					415		
CCC	ATC	CTG	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	1296
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	
			420					425					430			
GTG	TCC	GGC	GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	1344
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	
		435					440					445				
AAG	TTC	ATC	TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	1392
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	
	450					455					460					
GTG	ACC	ACC	CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	1440
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	
465					470					475					480	
CAC	ATG	AAG	CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	1488
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	
				485					490					495		
GTC	CAG	GAG	CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	1536
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	
			500					505					510			
CGC	GCC	GAG	GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	1584
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	
		515					520					525				
CTG	AAG	GGC	ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	1632
Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	
	530					535					540					
CTG	GAG	TAC	AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	1680
Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	
545					550					555					560	

-51-

CAG	AAG	AAC	GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	1728
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	
			565						570					575		
GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	1776
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	
			580					585					590			
GGC	GAC	GGC	CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	1824
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	
		595					600					605				
TCC	GCC	CTG	AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	1872
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	
	610					615					620					
CTG	GAG	TTC	GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	1920
Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	
625					630					635					640	
TAC	AAG	TAA														1929
Tyr	Lys															

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 642 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (C) STRANDEDNESS:
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
				5					10					15		
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
			35				40					45				
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55			60							
Leu	Thr	His	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65					70				75						80	
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85				90						95		

-52-

Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
			100					105					110		
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
		115					120					125			
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	130					135					140				
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn
145					150					155					160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser
				165					170					175	
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly
			180					185					190		
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu
		195					200					205			
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe
	210					215					220				
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala
225					230					235					240
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr
				245					250					255	
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn
			260					265					270		
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp
		275					280					285			
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg
	290					295					300				
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg
305					310					315					320
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Glu	Leu	Arg
				325					330					335	
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp
			340					345					350		
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr
		355					360					365			
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp

[illegible]

-54-

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 1959 BASE PAIRS
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
				5					10					15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
		35					40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
CTG	ACC	CAT	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	His	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
	65				70				75						80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85					90					95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
		115					120					125				
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	432
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
	130					135					140					
AAC	TTC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	

-55-

145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165					170					175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
		195					200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
	210					215					220					
GTG	ACC	GCC	GCC	CGC	ATG	CAT	GAC	CAA	CTG	ACA	GAA	GAG	CAG	ATT	GCA	720
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala	
225					230					235					240	
GAG	TTC	AAA	GAA	GCC	TTC	TCA	TTA	TTC	GAC	AAG	GAT	GGG	GAC	GGC	ACC	768
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr	
				245					250					255		
ATC	ACC	ACA	AAG	GAA	CTT	GGC	ACC	GTT	ATG	AGG	TCG	CTT	GGA	CAA	AAC	816
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn	
			260					265					270			
CCA	ACG	GAA	GCA	GAA	TTG	CAG	GAT	ATG	ATC	AAT	GAA	GTC	GAT	GCT	GAT	864
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp	
		275					280					285				
GGC	AAT	GGA	ACG	ATT	TAC	TTT	CCT	GAA	TTT	CTT	ACT	ATG	ATG	GCT	AGA	912
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg	
	290					295					300					
AAA	ATG	AAG	GAC	ACA	GAC	AGC	GAA	GAG	GAA	ATC	CGA	GAA	GCA	TTC	CGT	960
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg	
305					310					315					320	
GTT	TTT	GAC	AAG	GAT	GGG	AAC	GGC	TAC	ATC	AGC	GCT	GCT	GAA	TTA	CGT	1008
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Glu	Leu	Arg	
				325					330					335		
CAC	GTC	ATG	ACA	AAC	CTC	GGG	GAG	AAG	TTA	ACA	GAT	GAA	GAA	GTT	GAT	1056
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp	
			340					345					350			
GAA	ATG	ATA	AGG	GAA	GCA	GAT	ATC	GAT	GGT	GAT	GGC	CAA	GTA	AAC	TAT	1104
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr	

-56-

355	360	365	
GAA GAG TTT GTA CAA ATG ATG ACA GCA AAG GGG GGG AAG AGG CGC TGG Glu Glu Phe Val Gln Met Met Thr Ala Lys Gly Gly Lys Arg Arg Trp 370 375 380			1152
AAG AAA AAC TTC ATT GCC GTC AGC GCT GCC AAC CGG TTC AAG AAG ATC Lys Lys Asn Phe Ile Ala Val Ser Ala Ala Asn Arg Phe Lys Lys Ile 385 390 395 400			1200
TCC GAG CTC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG Ser Glu Leu Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val 405 410 415			1248
CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser 420 425 430			1296
GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu 435 440 445			1344
AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu 450 455 460			1392
GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp 465 470 475 480			1440
CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr 485 490 495			1488
GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr 500 505 510			1536
CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu 515 520 525			1584
CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys 530 535 540			1632
CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys 545 550 555 560			1680
CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu			1728

-57-

565								570				575				
GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	1776
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	
			580										590			
GGC	GAC	GGC	CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	1824
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	
		595					600					605				
TCC	GCC	CTG	AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	1872
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	
	610					615					620					
CTG	GAG	TTC	GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	1920
Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	
625					630					635					640	
TAC	AAG	CCA	AAA	AAG	AAG	AGA	AAG	GTG	GAA	GAC	GCT	TAA				1959
Tyr	Lys	Pro	Lys	Lys	Lys	Arg	Lys	Val	Glu	Asp	Ala					
			645						650							

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 652 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS:

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu
				5					10					15	
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly
			20					25					30		
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile
		35					40					45			
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr
	50					55					60				
Leu	Thr	His	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys
65					70					75					80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu

-58-

85								90				95			
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
			100					105					110		
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
		115					120					125			
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	130					135					140				
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn
145					150					155					160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser
				165					170					175	
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly
			180					185					190		
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu
		195					200					205			
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe
	210					215					220				
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala
225					230					235					240
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr
				245					250					255	
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn
			260					265					270		
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp
		275					280					285			
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg
	290					295					300				
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg
305					310					315					320
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Glu	Leu	Arg
				325					330					335	
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp
			340					345					350		
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr
		355					360					365			

-59-

Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp
370						375					380				
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile
385					390					395					400
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val
				405					410					415	
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser
			420					425					430		
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu
		435					440					445			
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu
	450					455					460				
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp
465					470					475					480
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr
				485					490					495	
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr
			500					505					510		
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu
		515					520					525			
Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys
	530					535					540				
Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys
545					550					555					560
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu
				565					570					575	
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile
			580					585					590		
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln
		595					600					605			
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu
	610					615					620				
Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu
625					630					635					640

-61-

AAC	TTC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165					170					175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
		195					200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
	210					215					220					
GTG	ACC	GCC	GCC	CGC	ATG	CAT	GAC	CAA	CTG	ACA	GAA	GAG	CAG	ATT	GCA	720
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala	
225					230					235					240	
GAG	TTC	AAA	GAA	GCC	TTC	TCA	TTA	TTC	GAC	AAG	GAT	GGG	GAC	GGC	ACC	768
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr	
				245					250					255		
ATC	ACC	ACA	AAG	GAA	CTT	GGC	ACC	GTT	ATG	AGG	TCG	CTT	GGA	CAA	AAC	816
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn	
			260					265					270			
CCA	ACG	GAA	GCA	GAA	TTG	CAG	GAT	ATG	ATC	AAT	GAA	GTC	GAT	GCT	GAT	864
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp	
		275					280					285				
GGC	AAT	GGA	ACG	ATT	TAC	TTT	CCT	GAA	TTT	CTT	ACT	ATG	ATG	GCT	AGA	912
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg	
	290					295					300					
AAA	ATG	AAG	GAC	ACA	GAC	AGC	GAA	GAG	GAA	ATC	CGA	GAA	GCA	TTC	CGT	960
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg	
305					310					315					320	
GTT	TTT	GAC	AAG	GAT	GGG	AAC	GGC	TAC	ATC	AGC	GCT	GCT	CAG	TTA	CGT	1008
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Gln	Leu	Arg	
				325					330					335		
CAC	GTC	ATG	ACA	AAC	CTC	GGG	GAG	AAG	TTA	ACA	GAT	GAA	GAA	GTT	GAT	1056
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp	

-62-

340				345				350								
GAA	ATG	ATA	AGG	GAA	GCA	GAT	ATC	GAT	GGT	GAT	GGC	CAA	GTA	AAC	TAT	1104
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr	
		355					360					365				
GAA	GAG	TTT	GTA	CAA	ATG	ATG	ACA	GCA	AAG	GGG	GGG	AAG	AGG	CGC	TGG	1152
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp	
	370					375					380					
AAG	AAA	AAC	TTC	ATT	GCC	GTC	AGC	GCT	GCC	AAC	CGG	TTC	AAG	AAG	ATC	1200
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile	
385					390					395					400	
TCC	GAG	CTC	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	1248
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	
				405					410					415		
CCC	ATC	CTG	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	1296
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	
			420					425					430			
GTG	TCC	GGC	GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	1344
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	
		435					440					445				
AAG	TTC	ATC	TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	1392
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	
	450					455					460					
GTG	ACC	ACC	CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	1440
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	
465					470				475						480	
CAC	ATG	AAG	CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	1488
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	
				485				490						495		
GTC	CAG	GAG	CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	1536
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	
			500					505					510			
CGC	GCC	GAG	GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	1584
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	
		515					520					525				
CTG	AAG	GGC	ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	1632
Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	
	530					535					540					
CTG	GAG	TAC	AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	1680

-63-

Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	
545					550					555					560	
CAG	AAG	AAC	GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	1728
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	
				565					570					575		
GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	1776
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	
			580					585					590			
GGC	GAC	GGC	CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	1824
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	
		595					600					605				
TCC	GCC	CTG	AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	1872
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	
	610					615					620					
CTG	GAG	TTC	GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	1920
Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	
625					630					635					640	
TAC	AAG	TAA														1929
Tyr	Lys															

(7) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 642 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (C) STRANDEDNESS:
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
				5					10					15		
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
		35					40					45				
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
Leu	Thr	His	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	

-64-

65					70						75				80
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu
				85					90					95	
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
			100					105					110		
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
		115					120					125			
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	130					135					140				
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn
145					150					155					160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser
				165					170					175	
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly
			180					185					190		
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu
		195					200					205			
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe
	210					215					220				
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala
225					230					235					240
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr
				245					250					255	
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn
			260					265					270		
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp
		275					280					285			
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg
	290					295					300				
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg
305					310					315					320
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Gln	Leu	Arg
				325					330					335	
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp
			340					345					350		

-65-

Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr
	355						360					365			
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp
	370					375					380				
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile
385					390					395					400
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val
				405					410					415	
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser
			420					425					430		
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu
		435					440					445			
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu
	450					455					460				
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp
465					470					475					480
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr
				485					490					495	
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr
			500						505				510		
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu
		515					520					525			
Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys
	530					535					540				
Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys
545					550					555					560
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu
				565					570					575	
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile
			580					585					590		
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln
		595					600					605			
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu
	610					615					620				

-66-

Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu
 625 630 635 640

Tyr Lys

(8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 1971 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG CTG CTG CCC GTC CCC CTG CTG CTG GGC CTG CTG GGC GCC GCC GCC	48
Met Leu Leu Pro Val Pro Leu Leu Leu Gly Leu Leu Gly Ala Ala Ala	
5 10 15	
GAC GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	96
Asp Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
20 25 30	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	144
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
35 40 45	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	192
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
50 55 60	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	240
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
65 70 75 80	
CTG ACC CAT GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	288
Leu Thr His Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
85 90 95	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	336
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
100 105 110	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	384
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
115 120 125	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	432

-67-

Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
130						135					140					
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	480
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
145					150					155					160	
AAC	TTC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	528
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
				165					170						175	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	576
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
			180					185					190			
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	624
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
		195					200					205				
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	672
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
	210					215					220					
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	720
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
225					230					235					240	
GTG	ACC	GCC	GCC	CGC	ATG	CAT	GAC	CAA	CTG	ACA	GAA	GAG	CAG	ATT	GCA	768
Val	Thr	Ala	Ala	Arg	Met	His	Asp	Gln	Leu	Thr	Glu	Glu	Gln	Ile	Ala	
				245					250					255		
GAG	TTC	AAA	GAA	GCC	TTC	TCA	TTA	TTC	GAC	AAG	GAT	GGG	GAC	GGC	ACC	816
Glu	Phe	Lys	Glu	Ala	Phe	Ser	Leu	Phe	Asp	Lys	Asp	Gly	Asp	Gly	Thr	
			260					265					270			
ATC	ACC	ACA	AAG	GAA	CTT	GGC	ACC	GTT	ATG	AGG	TCG	CTT	GGA	CAA	AAC	864
Ile	Thr	Thr	Lys	Glu	Leu	Gly	Thr	Val	Met	Arg	Ser	Leu	Gly	Gln	Asn	
		275					280					285				
CCA	ACG	GAA	GCA	GAA	TTG	CAG	GAT	ATG	ATC	AAT	GAA	GTC	GAT	GCT	GAT	912
Pro	Thr	Glu	Ala	Glu	Leu	Gln	Asp	Met	Ile	Asn	Glu	Val	Asp	Ala	Asp	
		290				295					300					
GGC	AAT	GGA	ACG	ATT	TAC	TTT	CCT	GAA	TTT	CTT	ACT	ATG	ATG	GCT	AGA	960
Gly	Asn	Gly	Thr	Ile	Tyr	Phe	Pro	Glu	Phe	Leu	Thr	Met	Met	Ala	Arg	
305					310					315					320	
AAA	ATG	AAG	GAC	ACA	GAC	AGC	GAA	GAG	GAA	ATC	CGA	GAA	GCA	TTC	CGT	1008
Lys	Met	Lys	Asp	Thr	Asp	Ser	Glu	Glu	Glu	Ile	Arg	Glu	Ala	Phe	Arg	
				325					330					335		
GTT	TTT	GAC	AAG	GAT	GGG	AAC	GGC	TAC	ATC	AGC	GCT	GCT	CAG	TTA	CGT	1056

-67.1-

Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Gln	Leu	Arg	
			340					345					350			
CAC	GTC	ATG	ACA	AAC	CTC	GGG	GAG	AAG	TTA	ACA	GAT	GAA	GAA	GTT	GAT	1104
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp	
		355					360					365				
GAA	ATG	ATA	AGG	GAA	GCA	GAT	ATC	GAT	GGT	GAT	GGC	CAA	GTA	AAC	TAT	1152
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr	
	370					375					380					
GAA	GAG	TTT	GTA	CAA	ATG	ATG	ACA	GCA	AAG	GGG	GGG	AAG	AGG	CGC	TGG	1200
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp	
385					390					395					400	
AAG	AAA	AAC	TTC	ATT	GCC	GTC	AGC	GCT	GCC	AAC	CGG	TTC	AAG	AAG	ATC	1248
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile	
				405					410					415		
TCC	GAG	CTC	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	1296
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	
			420					425					430			
CCC	ATC	CTG	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	1344
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	
		435					440					445				
GTG	TCC	GGC	GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	1392
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	
	450					455					460					
AAG	TTC	ATC	TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	1440
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	
465					470				475						480	
GTG	ACC	ACC	CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	1488
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	
				485				490						495		
CAC	ATG	AAG	CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	1536
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	
			500					505					510			
GTC	CAG	GAG	CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	1584
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	
		515						520				525				
CGC	GCC	GAG	GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	1632
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	
	530					535					540					
CTG	AAG	GGC	ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	1680

-67.2-

Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	
545					550					555					560	
CTG	GAG	TAC	AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	1728
Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	
				565					570					575		
CAG	AAG	AAC	GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	1776
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	
			580					585					590			
GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	1824
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	
		595					600					605				
GGC	GAC	GGC	CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	1872
Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	
	610					615					620					
TCC	GCC	CTG	AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	1920
Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	
625					630					635					640	
CTG	GAG	TTC	GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	AAG	GAC	GAG	CTG	1968
Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Lys	Asp	Glu	Leu	
				645					650					655		
TAA																1971

(9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 656 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS:

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Leu	Leu	Pro	Val	Pro	Leu	Leu	Leu	Gly	Leu	Leu	Gly	Ala	Ala	Ala	
				5					10					15		
Asp	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
			20					25					30			
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			35				40					45				
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	

-67.3-

50		55		60
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr				
65		70		75
				80
Leu Thr His Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys				
	85		90	95
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu				
	100		105	110
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu				
	115		120	125
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly				
	130		135	140
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr				
145		150		155
				160
Asn Phe Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn				
	165		170	175
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser				
	180		185	190
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly				
	195		200	205
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu				
	210		215	220
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe				
225		230		235
				240
Val Thr Ala Ala Arg Met His Asp Gln Leu Thr Glu Glu Gln Ile Ala				
	245		250	255
Glu Phe Lys Glu Ala Phe Ser Leu Phe Asp Lys Asp Gly Asp Gly Thr				
	260		265	270
Ile Thr Thr Lys Glu Leu Gly Thr Val Met Arg Ser Leu Gly Gln Asn				
	275		280	285
Pro Thr Glu Ala Glu Leu Gln Asp Met Ile Asn Glu Val Asp Ala Asp				
	290		295	300
Gly Asn Gly Thr Ile Tyr Phe Pro Glu Phe Leu Thr Met Met Ala Arg				
305		310		315
				320
Lys Met Lys Asp Thr Asp Ser Glu Glu Glu Ile Arg Glu Ala Phe Arg				

-67.4-

325					330					335					
Val	Phe	Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Gln	Leu	Arg
			340					345					350		
His	Val	Met	Thr	Asn	Leu	Gly	Glu	Lys	Leu	Thr	Asp	Glu	Glu	Val	Asp
		355					360					365			
Glu	Met	Ile	Arg	Glu	Ala	Asp	Ile	Asp	Gly	Asp	Gly	Gln	Val	Asn	Tyr
	370					375					380				
Glu	Glu	Phe	Val	Gln	Met	Met	Thr	Ala	Lys	Gly	Gly	Lys	Arg	Arg	Trp
385					390					395					400
Lys	Lys	Asn	Phe	Ile	Ala	Val	Ser	Ala	Ala	Asn	Arg	Phe	Lys	Lys	Ile
				405					410					415	
Ser	Glu	Leu	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val
			420					425					430		
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser
		435					440					445			
Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu
	450					455					460				
Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu
465					470					475					480
Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp
				485					490					495	
His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr
			500					505					510		
Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr
		515					520					525			
Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu
	530					535					540				
Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys
545					550					555					560
Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys
				565					570				575		
Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu
			580					585					590		
Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile
		595					600					605			

-67.5-

Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln
610 615 620

Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu
625 630 635 640

Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Lys Asp Glu Leu
645 650 655

WO 98/40477

PCT/US98/04978

- 68 -

From the above description, the essential characteristics of the present invention can be ascertained. Without departing from the spirit and scope thereof, various changes and modifications of the invention can be made to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

THE EMBODIMENTS OF THE INVENTION FOR WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A fluorescent indicator fusion protein for detecting an analyte which is a ligand of a binding protein, wherein the fusion protein comprises:
 - a binding protein moiety having the ligand binding region of the binding protein and which undergoes a conformational change upon the binding of the ligand;
 - a donor fluorescent protein moiety wherein the donor fluorescent protein moiety is fused to the binding protein moiety;
 - an acceptor fluorescent protein moiety wherein the acceptor fluorescent protein moiety is fused to the binding protein moiety; and
 - wherein the donor fluorescent protein moiety has an emission spectrum which overlaps with the absorption spectrum of the acceptor fluorescent protein moiety and the excitation spectrum of the donor fluorescent moiety comprises a wavelength of excitation at which the donor fluorescent protein moiety can be excited effectively and selectively with respect to the acceptor fluorescent protein moiety; and the emission spectrum of the donor fluorescent protein moiety and the emission spectrum of the acceptor fluorescent protein moiety allow the two emissions to be distinguished, and fluorescence resonance energy transfer (FRET) between the donor fluorescent protein moiety and the acceptor fluorescent protein moiety can be detected when the donor fluorescent protein moiety is excited at the wavelength; and
 - wherein the acceptor fluorescent protein moiety and the donor fluorescent protein moiety change position relative to each other when the ligand binds to the ligand-binding region, altering the FRET.
2. The fusion protein of claim 1, wherein the donor fluorescent protein moiety and the acceptor fluorescent protein moiety are Aequorea-related fluorescent protein moieties.

3. The fusion protein of claim 1, wherein the binding protein moiety additionally has a domain capable of binding a target peptide moiety and the target peptide moiety is fused to the binding protein moiety by a polypeptide linker, wherein the linker is from 1 to 20 amino acids in length, and the binding of the ligand to the ligand binding region of the binding protein moiety alters the binding of the target peptide moiety to the target peptide binding domain of the binding protein moiety to alter the FRET.
4. The fusion protein of claim 1, wherein one of the donor fluorescent protein moiety or the acceptor protein moiety is fused at the carboxy terminus of the fusion protein and the other one of the donor fluorescent protein moiety or the acceptor protein moiety is fused at the amino terminus of the fusion protein.
5. The fusion protein of claim 1, wherein the binding protein moiety is calmodulin, a calmodulin-related protein moiety, cGMP-dependent protein kinase, a steroid hormone receptor, a ligand binding domain of a steroid hormone receptor, protein kinase C, inositol-1, 4,5-triphosphate receptor, or recoverin.
6. The fusion protein of claim 3, wherein the binding protein moiety is calmodulin or a calmodulin-related protein moiety.
7. The fusion protein of claim 1, wherein the donor fluorescent protein moiety is a green fluorescent protein mutant selected from blue-shifted mutant P4- 3, improved blue mutant EBFP, and blue-green fluorescent mutant W1B, and the acceptor fluorescent protein moiety is a green fluorescent protein mutant selected from improved green mutant S65T, green fluorescent protein mutant EGFP, and yellow-green-shifted mutant I0c.
8. The fusion protein of claim 3, wherein the target peptide moiety is a calmodulin-binding domain of skeletal muscle myosin light chain kinase (skMLCKp), smooth muscle-myosin light chain kinase (smMLCK), calmodulin

kinase II (CaMKII), Caldesmon, Calspermin, phosphofructokinase, calcineurin, phosphorylase kinase, Ca^{2+} -ATPase 59 kDa phosphodiesterase (PDE), 60 kDa PDE, nitric oxide synthase, type I adenylyl cyclase, *Bordetella pertussis* adenylyl cyclase, Neuromodulin, Spectrin, myristoylated alanine-rich C kinase substrate (MARCKS), MARCKS-related protein (F52), B-Adducin, heat shock protein 90a (HSP90a), human immunodeficiency virus-1 envelope glycoprotein 160 (HIV-1 gp160), brush-border myosin heavy chain-1 (BBMHBI), myosin heavy chain (MHC), Mastoparan, Melittin, Glucagon, Secretin, vasoactive intestinal peptide (VIP), gastrin inhibitory peptide (GIP), or calmodulin binding peptide-2 (Model Protein CBP2).

9. The fusion protein of claim 8, wherein the target peptide moiety is the calmodulin-binding peptide of skeletal muscle myosin light-chain kinase.
10. The fusion protein of claim 9, wherein the polypeptide linker is -Gly-Gly-.
11. The fusion protein of claim 1, wherein the fusion protein further comprises a localization moiety and the localization moiety is a localization moiety for the nucleus, the endoplasmic reticulum, a peroxisome, the Golgi apparatus, the plasma membrane or the mitochondria.
12. The fusion protein of claim 1, wherein the binding protein moiety comprises a conformation of smaller volume when the ligand is bound to the ligand binding region of the binding protein moiety compared to the volume of the binding protein moiety when the ligand binding region is not bound to the ligand, whereby the FRET is increased by the binding of the ligand to the ligand binding region of the binding protein moiety.
13. The fusion protein of claim 3, wherein the binding protein moiety has increased affinity for the target peptide moiety when the ligand binding region of the binding protein moiety is bound by the ligand.

14. The fusion protein of claim 1, wherein the donor fluorescent protein moiety is capable of being excited by ultraviolet light to emit blue light and the acceptor fluorescent protein moiety is capable of being efficiently excited by the emitted blue light but not by the exciting ultraviolet light and of emitting green light when excited by the emitted blue light.
15. The fusion protein of claim 1, wherein the donor fluorescent protein moiety is capable of being excited by violet light to emit blue-green light and the acceptor fluorescent protein moiety is capable of being efficiently excited by the emitted blue-green light but not by the exciting violet light and of emitting yellow-green light when excited by the emitted blue-green light.
16. The fusion protein of claim 1, wherein the donor and acceptor fluorescent protein moieties are selected from among the green fluorescent proteins of cnidarians.
17. The fusion protein of claim 16, wherein the donor and acceptor fluorescent protein moieties are selected from or related to the fluorescent proteins of *Renilla Reniformis*, *Phialidium gregarium*, and *Aequoria Victoria*.
18. The fusion protein of claim 1, wherein the donor and acceptor fluorescent protein moieties are selected from the green fluorescent proteins of cnidarians; the yellow fluorescent proteins of *Vibrio fischeri* strain Y-1, peridinin-chlorophyll *a* binding protein from the dinoflagellate *Symbiodinium* sp.; phycobiliproteins from marine cyanobacteria, and oat phytochromes from oat reconstructed with phycoerythrobilin.
19. An isolated nucleic acid molecule encoding the fluorescent indicator fusion protein of any one of claims 1 to 18.

20. An expression vector comprising a nucleic acid encoding the fluorescent indicator fusion protein of any one of claims 1 to 18.
21. The expression vector of claim 20, wherein the nucleic acid encoding the fluorescent indicator fusion protein is operatively linked to an expression control sequence.
22. The expression vector of claim 21, wherein the vector is adapted for function in a prokaryotic cell or a eukaryotic cell.
23. A host cell transfected with the expression vector of claim 20.
24. The cell of claim 23, wherein the cell is a prokaryote.
25. The cell of claim 23, wherein the cell is a eukaryote.
26. The cell of claim 23, wherein the cell is a yeast cell.
27. The cell of claim 23, wherein the cell is a mammalian cell.
28. A method for determining the concentration in a sample of an analyte which is a ligand of a binding protein, comprising:
 - contacting the sample with the fusion protein of any one of claims 1 to 18;
 - exciting the donor fluorescent protein moiety; and
 - determining the degree of FRET in the sample by measuring light emitted by the acceptor fluorescent protein moiety.
29. The method of claim 28, wherein the analyte is calcium.
30. A method for determining the concentration of an analyte in the cell of claim 23, wherein the analyte is ligand of the binding protein, comprising:

- 74 -

expressing the fusion protein in the cell;
exciting the donor fluorescent protein moiety; and
determining the degree of FRET in the sample by measuring light emitted
by the acceptor fluorescent protein moiety.

31. The method of claim 30, wherein the cell is a mammalian cell.
32. The method of claim 30, wherein the analyte is calcium.
33. The method of claim 30, wherein the fusion protein comprises a localization moiety.

1/18

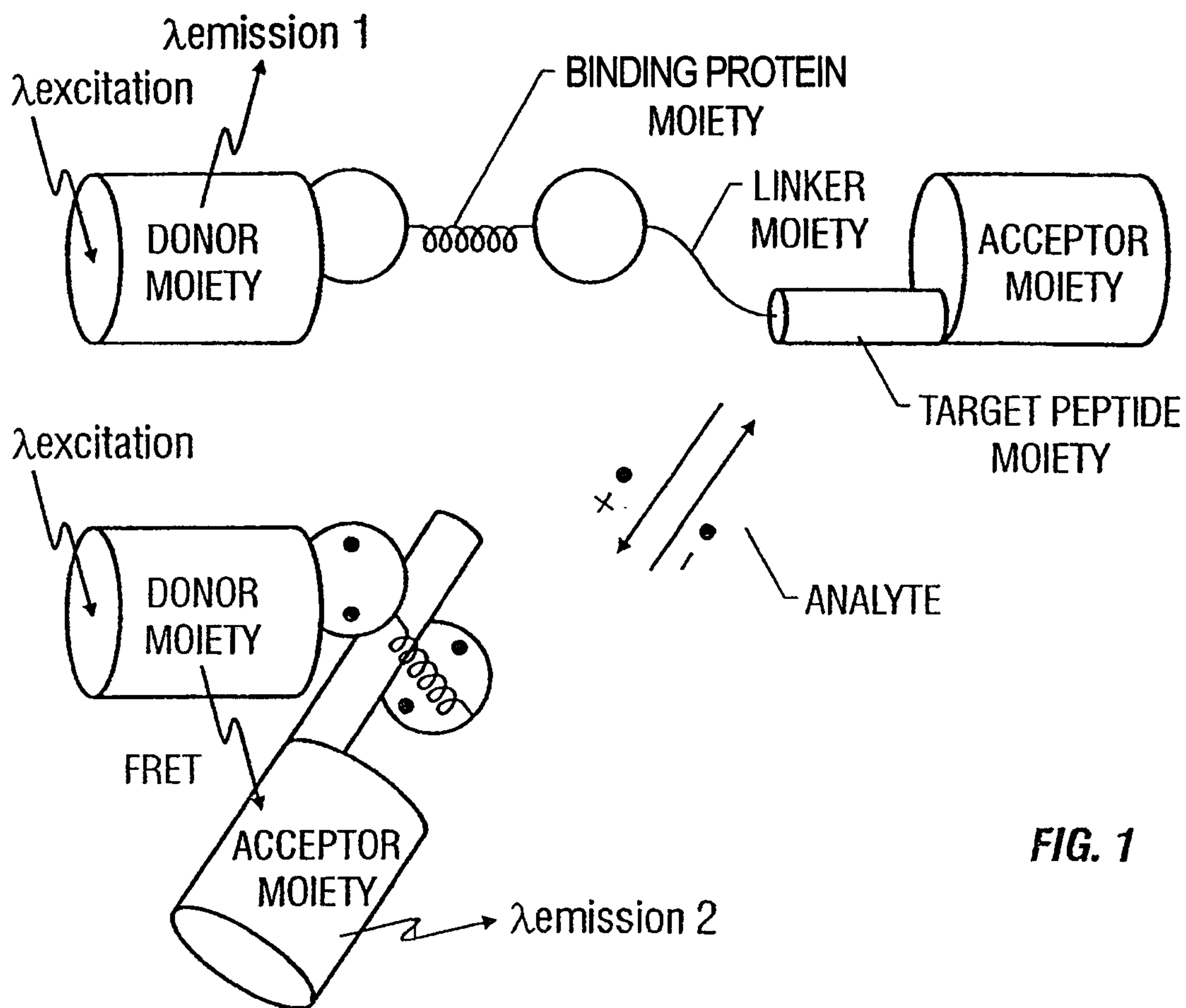
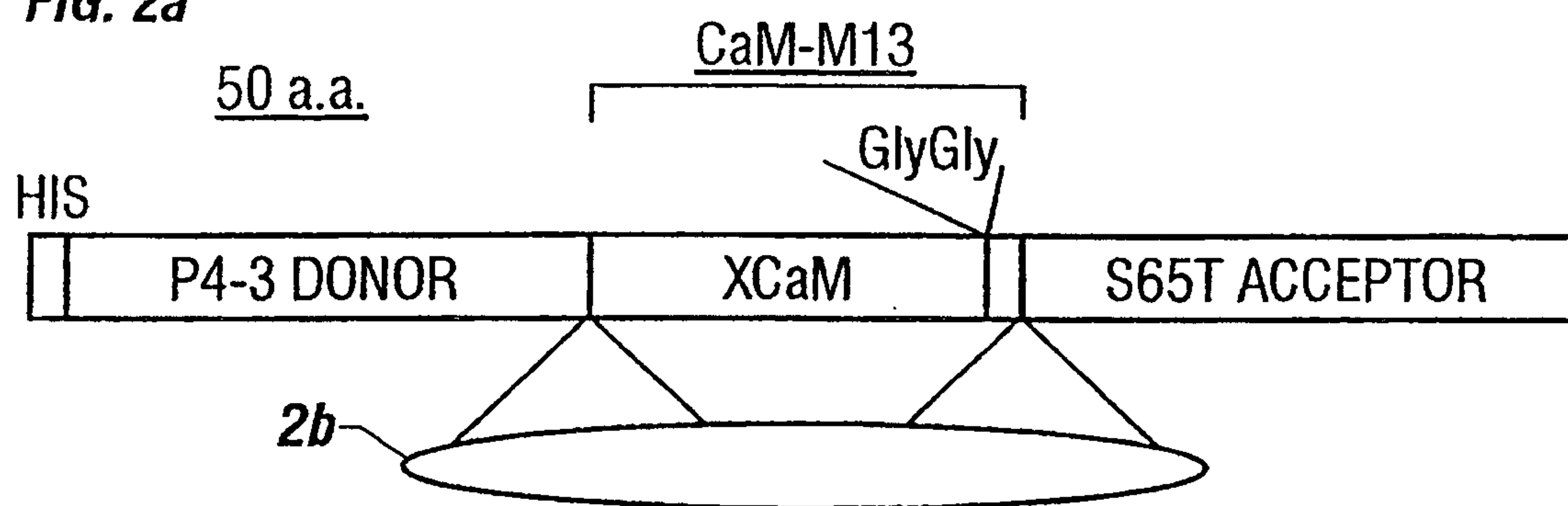


FIG. 1

FIG. 2a



226	227		1	2		20	21		1	2	
Ala	Ala	Arg	Met	His		Ile	Ser	Glu	Leu	Met	Ser
GCT	GCT	CGC	ATG	CAT		ATC	TCC	GAG	CTC	ATG	AGT
			<u>SphI</u>						<u>SacI</u>		

FIG. 2b

2/18

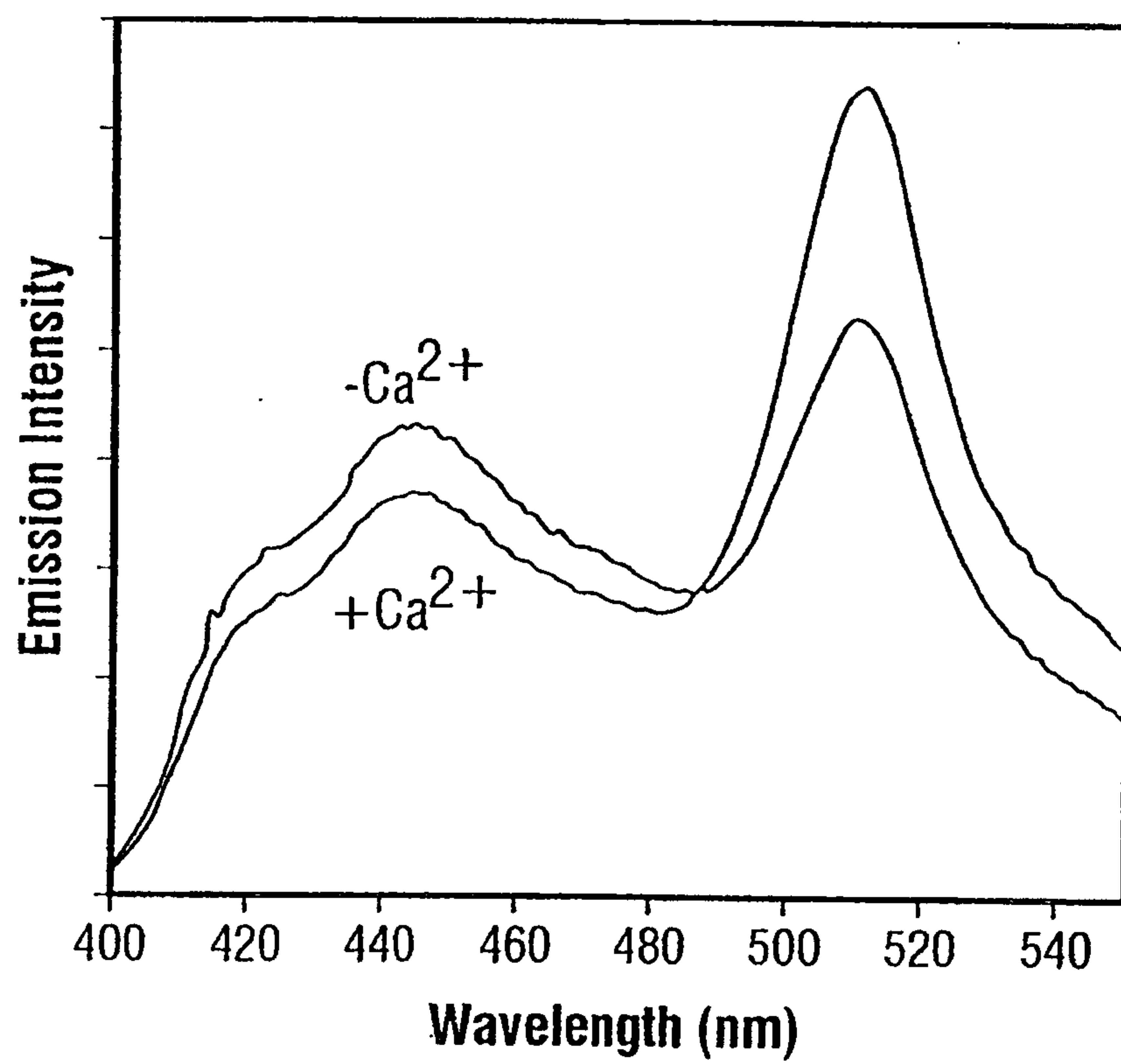


FIG. 3

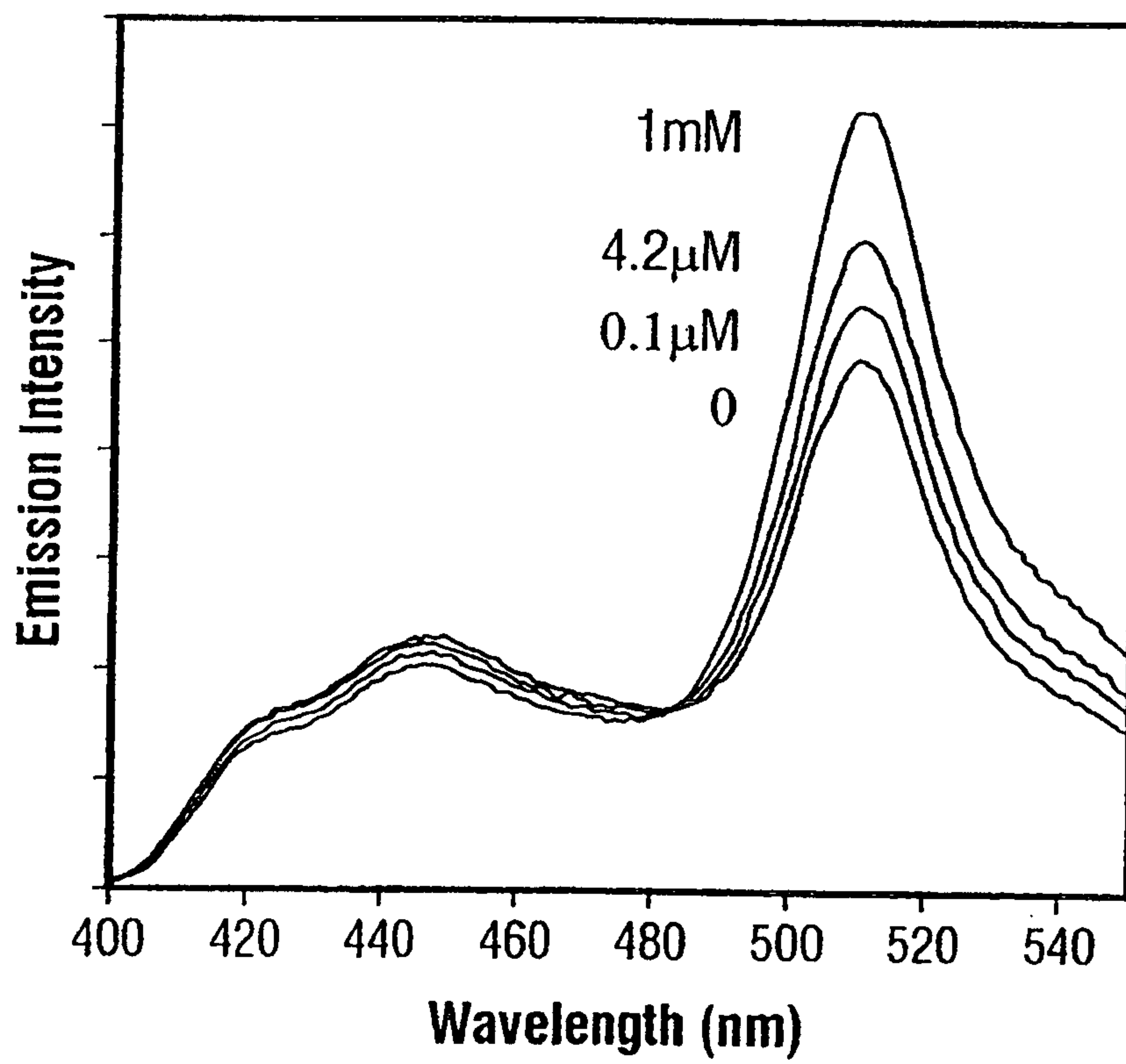


FIG. 4a

3/18

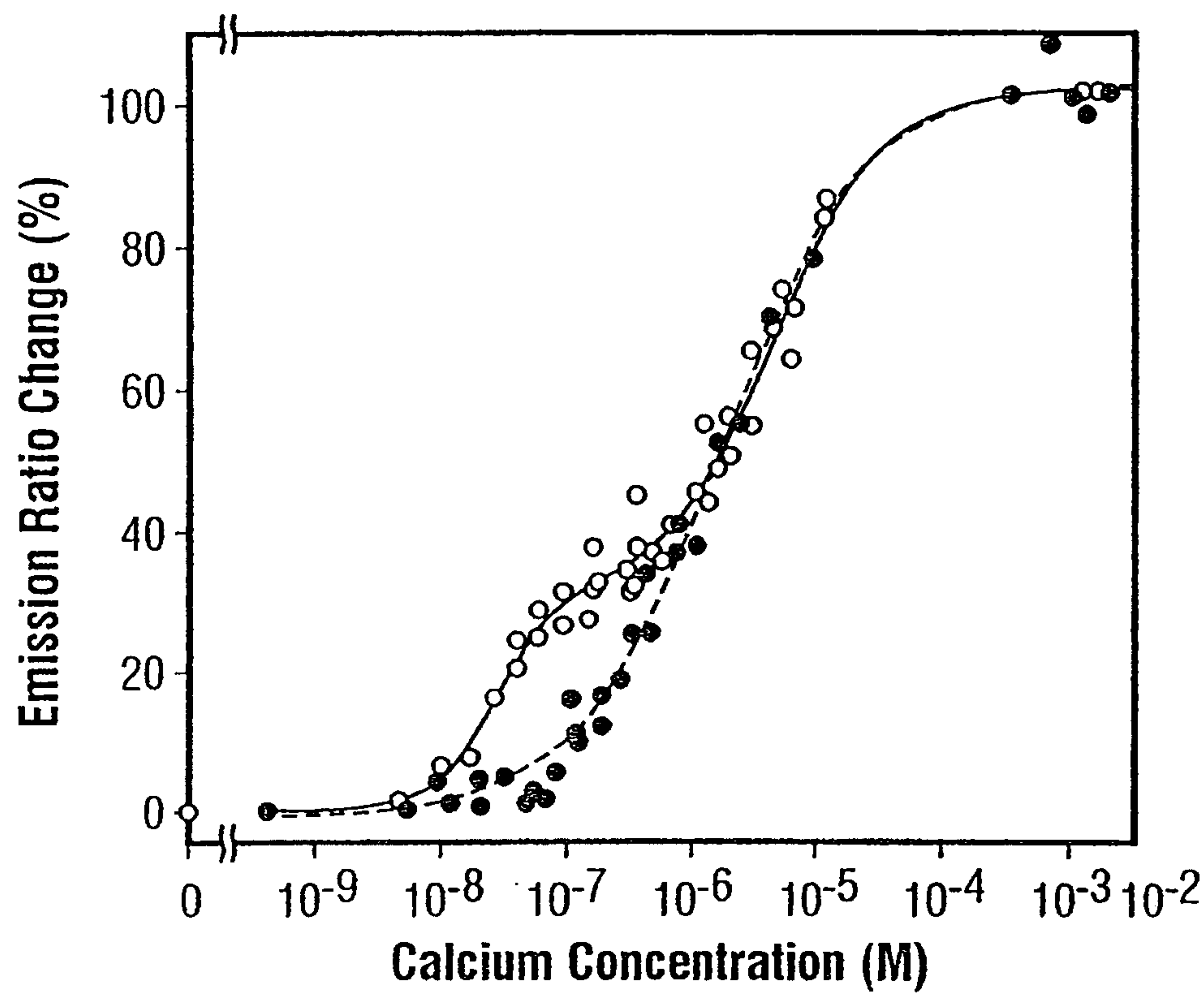


FIG. 4b

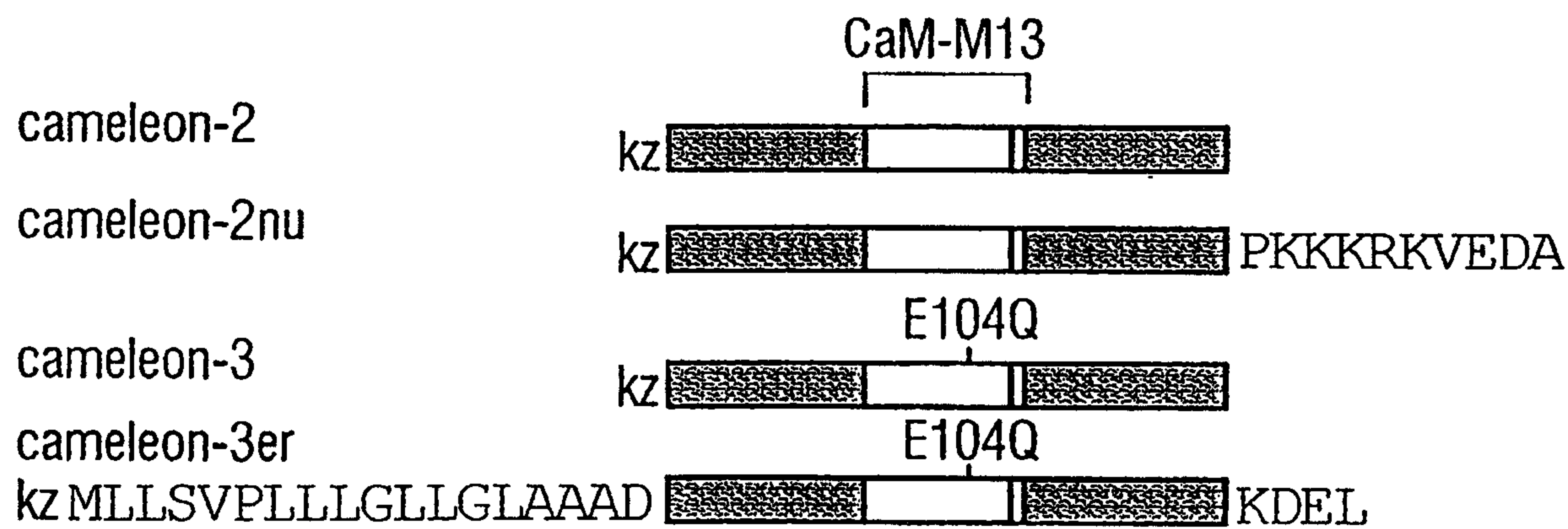


FIG. 5

4/18



FIG. 6a

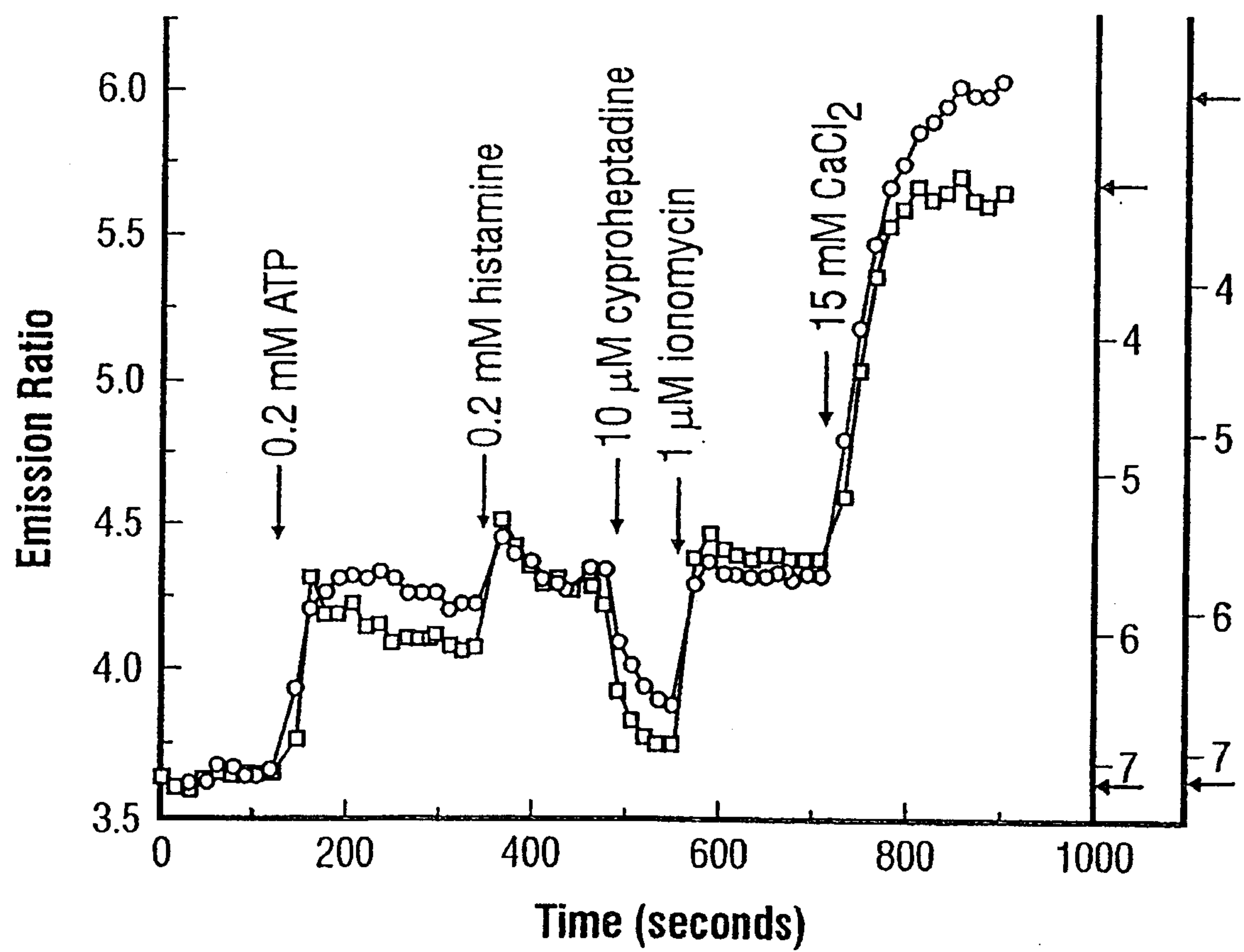


FIG. 6b

5/18

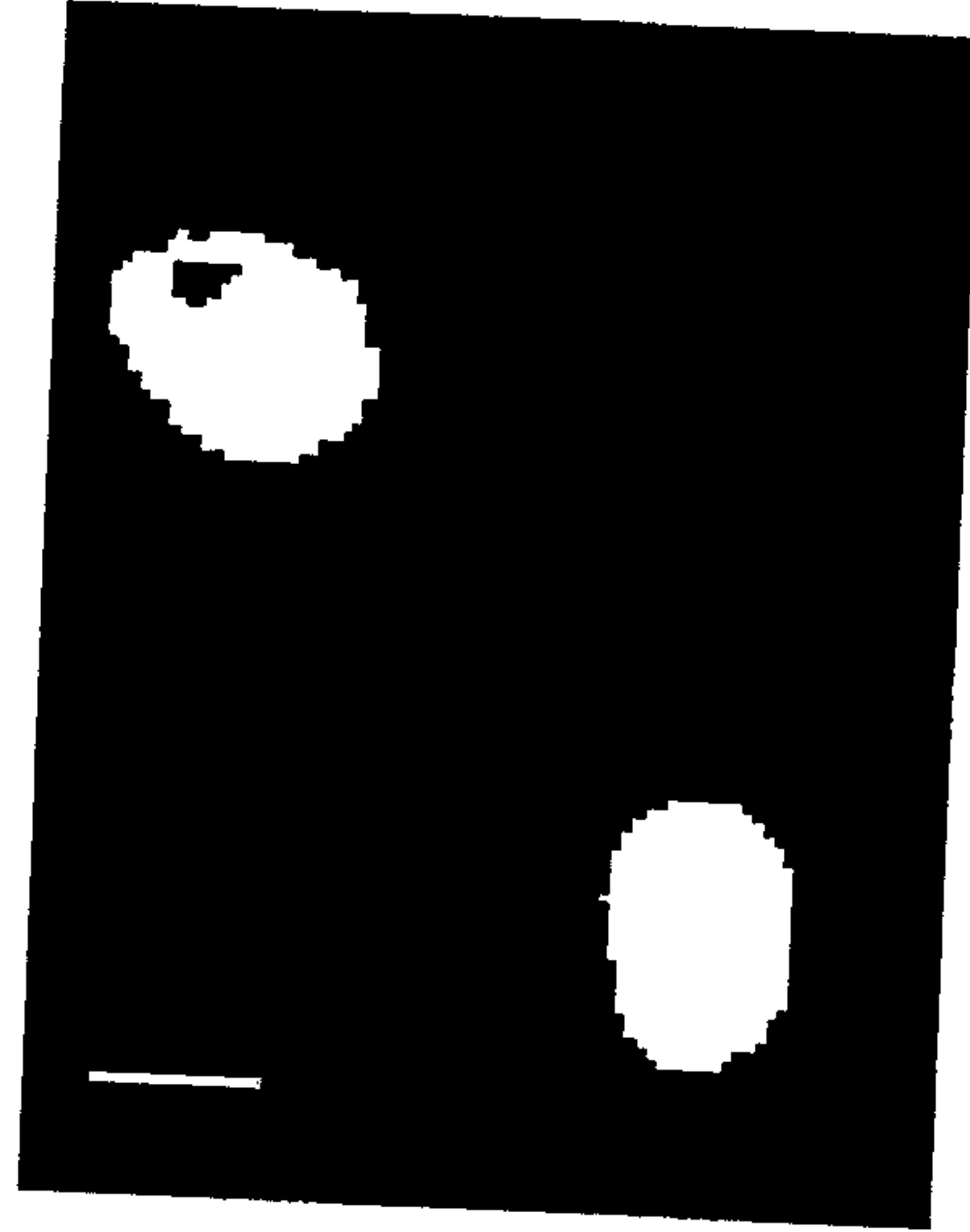


FIG. 6c

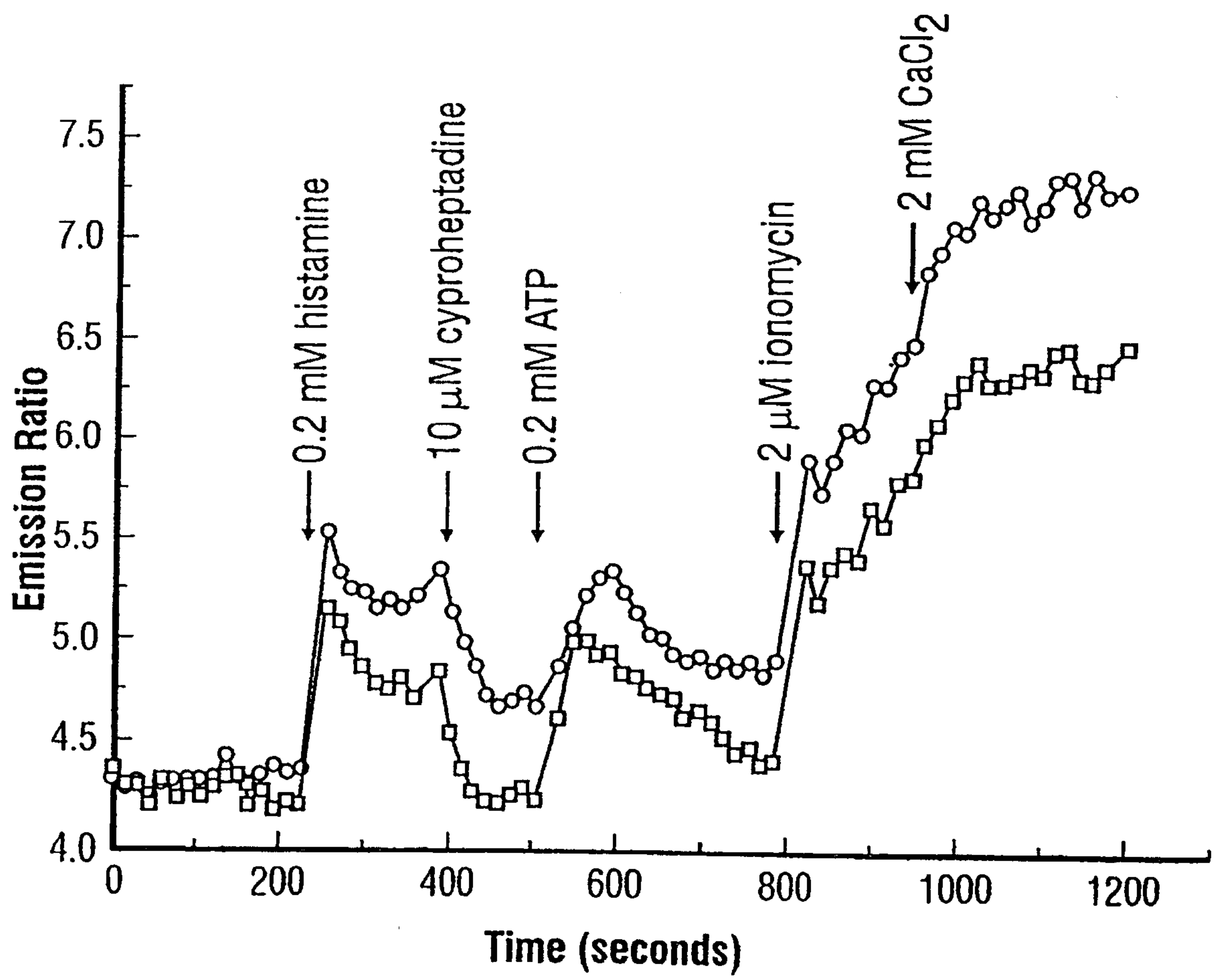


FIG. 6d

6/18



FIG. 6E

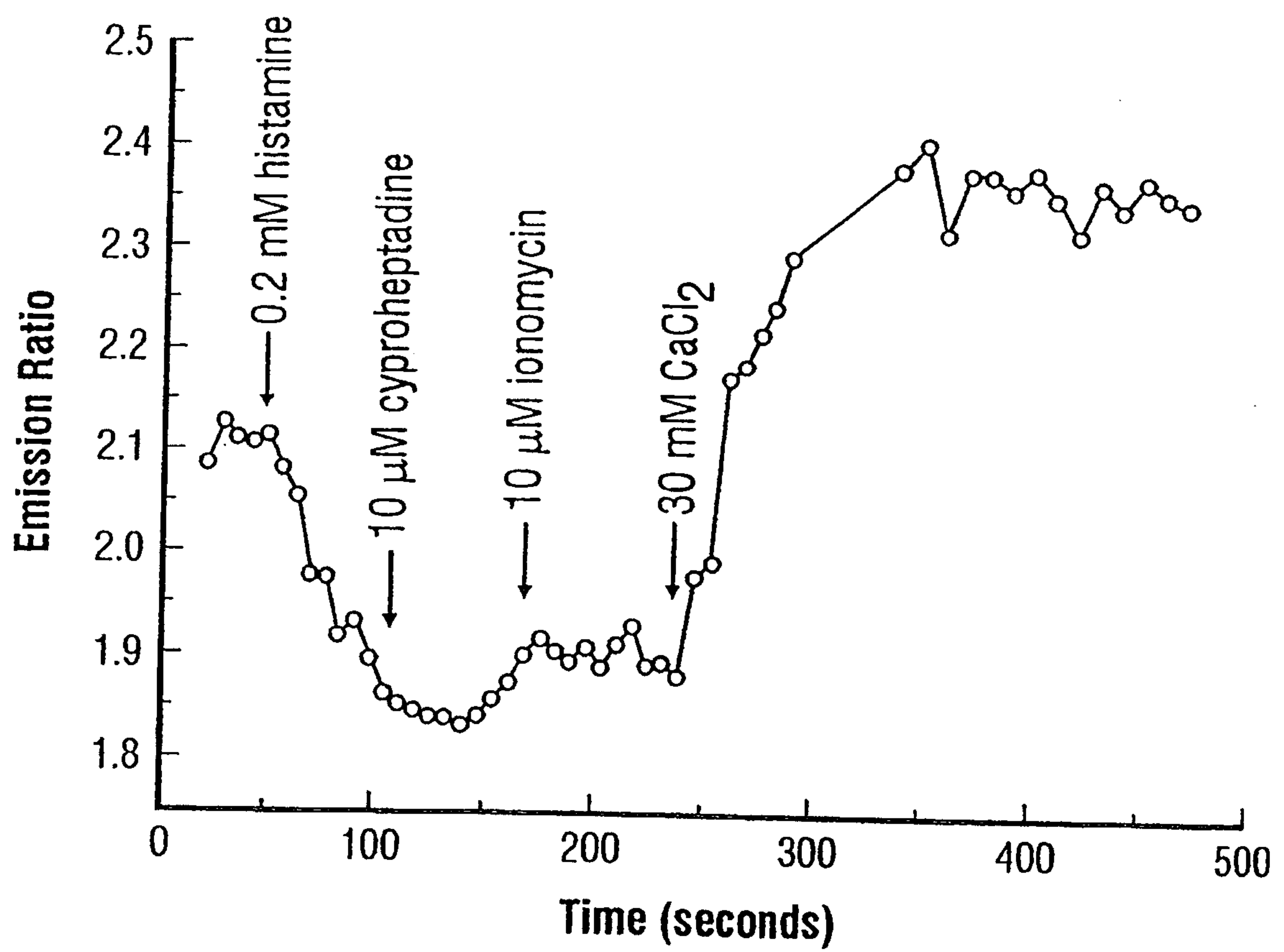


FIG. 6F

7/18

(SEQ ID. NO:1)
ATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGGA
CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGGATGCCACCTACGGCAAGCTGA
CCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTG
ACCCATGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTC
CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGA
CCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACCTTCAACAGCCACAACGTCTA
TATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAAACACCCCATCGGCGACGGCCCCGTGCTG
CTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCGCCGCATGCATGACCAACTGACAGAAAGAGCAGA
TTGCAGAGTTCAAAGAAGCCTTCTCATTTATTCGACAAGGATGGGACGGCACCATCACCAAAAG
GAACTTGGCACCCGTTATGAGGTCGCTTGGAACAACCCAAACGGAAGCAGAAATTGCAGGATATGAT
CAATGAAGTCGATGCTGATGGCAATGGAACGATTTACTTTCCCTGAATTTCTTACTATGATGGCTA
GAAAATGAAGGACACAGACAGCGAAGAGGAAATCCGAGAAGCATTCCTGTGTTTTTGACAAGGAT
GGGAACGGCTACATCAGCGCTGCTGAATTACGTCACGTCATGACAAACCTCGGGAGAAGTTAAC
AGATGAAGAAGTTGATGAAATGATAAGGAAGCAGATATCGATGGTGATGCCCAAGTAAACTATG

Fig. 7-1

8/18

(SEQ ID. NO:1)
AAGAGTTTGTACAAATGATGACAGCAAAGGGGGGGAAGAGCGCTGGAAGAAACTTCATTGCC
GTCAGCGCTGCCAACCGGTTCAAGAAGATCTCCGAGCTCATGGTGAGCAAGGGCGAGAGCTGTT
CACCGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAACCGGC
AAGCTGCCCGTGGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCAGTGCTTCAGCCG
CTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGG
AGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCCGAGGTGAAGTTCGAGGGC
GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCGACAGCAGAAGAACG
GCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCACTCGCCGACCAC
TACCAGCAGAACACCCCATCGGCGACGGCCCGTGCTGCTGCCCGACAACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCTGTGGAGTTCGTGA
CCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA

FIG. 7-2

9/18

(SEQ ID. NO:2)
MVSKEEELFTGVVPILVELDGVNGHKFSVSGEGEDATYGKLTILKFICTTGKLPVPWPTLVTTLL
THGVQCFSRYPDHMKQHDFEKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGID
FKEDGNILGHKLEYNFSHNYYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVL
LPDNHYLSTQSALSKDPNEKRDHMLLEFVTAARMHDQLTEEQIAEFKEAFSLEDKDGDTITTK
ELGTVMRSLGQNPTAEELQDMINEVDADGNGTIYFPEFLTMARKMKDTSDEEIREAFRVFDKD
NGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGVNYYEEFVQMMTAKGKRWRWKNFIA
VSAANRFKKISELMVSKGEELEFTGVVPILVELDGVNGHKFSVSGEGEDATYGKLTILKFICTTG
KLPVPWPTLVTTLYGVQCFSRYPDHMKQHDFEKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEED
DTLVNRIELKGIDFKEDGNILGHKLEYNFSHNYYIMADKQKNGIKVNFKIRHNIEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK*

FIG. 7-3

10/18

(SEQ ID. NO:3)
ATGGTGAGCAAGGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA
CGTAAACGGCCACAAGTTCAGCGTGTCGGCGAGGGCGGATGCCACCTACGGCAAGCTGA
CCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCTGGCCACCCCTCGTGACCACCCCTG
ACCCATGGCGTGCAGTGCTTCAGCCGCTACCCGACCATCATCTTCTTCAAGGACGACGACTTCTTCAAGTC
CGCCATGCCGAAGGCTACGTCCAGGAGCGCACCCCTGGTGAACCGCATCGAGCTGAAGGCATCGAC
CCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACCTTCAACAGCCACAACGTCTA
TATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAAACACCCCATCGGCGACGGCCCGTGCTG
CTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGTGGAGTTCGTGACCGCCCGCATGCATGACCAACTGACAGAAAGAGCAGA
TTGCAGAGTTCAAAGAAGCCTTCTCATTTATTCGACAAGGATGGGACGGCACCATCACCAAAAG
GAACTTGGCACCGTTATGAGGTCGCTTGACAATAACCAACGGAAGCAGAATTGCAGGATATGAT
CAATGAAGTCGATGCTGATGGCAATGGAACGATTTACTTTCTCTGAATTCTTACTATGATGGCTA
GAAAAATGAAGGACACAGACAGCGAAGAGGAATCCGAGAAGCATTCGCTGTTTGTGACAAGGAT
GGGAACGGCTACATCAGCGCTGCTGAATTACGTACGTCAAGCAAAACCTCGGGGAGAAGTTAAC
AGATGAAGAAGTTGATGAATGATAAGGGAAGCAGATATCGATGGTGATGGCCAAGTAAACTATG

Fig. 8-1

11/18

(SEQ ID. NO:3)

AAGAGTTTGTACAAATGATGACAGCAAAGGGGGAAGAGCGCTGGAAGAAAACTTCATTGCC
GTCAGCGCTGCCAACCGGTTCAAGAAAGATCTCCGAGCTCATGTTGAGCAAGGCGAGGAGCTGTT
CACCGGGTGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGC
AAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCAGTGCTTCAGCCG
CTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGG
AGCGACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGGAGGTGAAGTTCGAGGGC
GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACACTACAAGCCACAACGTCCTATATCATGGCCGACAAGCAGAAGAACG
GCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCAC
TACCAGCAACAACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAGACCCCAACGAGAAGCGGATCACATGCTCCTGCTGGAGTTCGTGA
CCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGCCAAAAAGAGAGAAAGGTGGAA
GACGCTTAA

FIG. 8-2

12/18

(SEQ ID. NO:4)
MVSKEEELFTGVVPILVELDGVNKGKFSVSGEGDATYGKLTILKFICTTGKLPVPWPTLVTTTL
THGVQCFSRYPDHMKQHDFEKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEEDTLVNRIELKGID
FKEDGNILGHKLEYNFSHNYYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVL
LPDNHYLSTQSALSKDPNEKRDHMLLEFVTAARMHDQLTEEQIAEFKEAFSLFDKDGDTITTK
ELGTVMRSLGQNPTAEELQDMINEVDADGNGTIYFPEFLTMARKMKDTSDEEIREAEFRVFDKD
NGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGVNYYEEFVQMTAKGKRWRWKNFIA
VSAANREKKISELMVSKGEEELFTGVVPILVELDGVNKGKFSVSGEGDATYGKLTILKFICTTG
KLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFEKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEED
DTLVNRIELKGIDFKEDGNILGHKLEYNYSHNYYIMADKQKNGIKVNFKIRHNIEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKPKKKRVE
DA*

FIG. 8-3

13/18

(SEQ ID. NO:5)
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA
CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGGATGCCACCTACGCAAGCTGA
CCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCAACCCTG
ACCCATGGCGTGAGTGCTTCAGCCGCTACCCCGACCATCTTCTTCAAGGACGACGACTTCTTCAAGTC
CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCCATCTTCTTCAAGGACGACGGCAACTACAAGA
CCCGCGCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAAGCTGGAGTACAACCTTCAACAGCCACAACGTCTA
TATCATGCCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTG
CTGCCCGACAACCACTAGCTGAGCACCCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCGCCCGCATGCATGACCAACTGACAGAAGAGCAGA
TTGCAGAGTTCAAAGAAGCCTTCTCATTTATTCGACAAGGATGGGACGGCACCATCACCAAAAG
GAACTTGGCACCCGTTATGAGGTCGCTTGGAACAACCCAAACGGAAGCAGAATTGCAGGATATGAT
CAATGAAGTCGATGCTGATGGCAATGGAACGATTTACTTTCTGAATTTCTTACTATGATGGCTA
GAAAATGAAGGACACAGACAGCGAAGAGGAAATCCGAGAAGCATTCGCTGTTTGGACAAGGAT
GGGAACGGCTACATCAGCGCTGCTCAGTTACGTACGTCACTGACAAACCTCGGGAGAAGTTAAC
AGATGAAGAAGTTGATGAAATGATAAGGGAAGCAGATATCGATGGTGATGGCCAAGTAAACTATG

Fig. 9-1

14/18

(SEQ ID. NO:5)

AAGAGTTTGTACAAATGATGACAGCAAAAGGGGGAAGAGGCGCTGGAAGAAACTTCATTGCC
GTCAGCGCTGCCAACCGGTTCAAGAAGATCTCCGAGCTCATGGTGAGCAAGGCGAGGAGCTGTT
CACCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAACCGGC
AAGCTGCCCGTGCCCTGGCCCCACCCTCGTGACCAACCTGACCTACGGCGTGCAGTGCTTCAGCCG
CTACCCCGACCATGAAGCAGCAGCAGCTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGG
AGCGACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCGAGGTGAAGTTCGAGGGC
GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACACTACAAGCCACAACGTCTATATCATGGCCGACACAAGCAGAAGAACG
GCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCACTCGCCGACCAC
TACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGGAGTTCGTGA
CCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA

FIG. 9-2

15/18

(SEQ ID. NO:6)
MVSKEEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTTL
THGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAENVKFEEDTLVNRIELKGID
FKEDGNILGHKLEYNFNSHNVYIMADKQKNGIKVNEKIRHNIEDGSVQLADHYQQNTPIGDGPVL
LPDNHYLSTQSALS KDPNEKRDHMLLEFVTAARMHDQLTEEQIAEFKEAFSLFDKDGDTITTK
ELGTVMRSLGQNPTEAELQDMINEVDADGNGTIYFPFELTMMARKMKDTDSEEEIREAFRVFDKD
NGYISAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGVNYYEEFVQMMTAKGKRWRKKNFIA
VSAANRFKKISELMVSKGEEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTG
KLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAENVKFEED
DTLVNRIELKGIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNEKIRHNIEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYK*

FIG. 9-3

16/18

(SEQ ID. NO:7)
ATGCTGCTGCCCGTCCCCCTGCTGCTGGCCCTGCTGGCGCCGCCGACGTGAGCAAGGGCGA
GGAGCTGTTACCGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGGATGCCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGC
ACCACGGCAAGCTGCCCGTGCCCTGGCCCACTCGTGACCACTGACCCATGGCGTGCAAGTG
CTTACGCCGTACCCCGACCATGATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCGAAGGCT
ACGTCCAGGAGCGACCATCTTCTTCAAGGACGACGGCACTACAAGACCCGCCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGACAAGCTGGAGTACAACCTTCAACAGCCACAACGTCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGTGAACTTCAAGATCCGCCACAACATCGAGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGG
AGTTCGTGACCGCCCGCATGCATGACCAACTGACAGAAGAGCAGATTGCAGAGTTCAAAAGAA
GCCTTCTCATTATTCGACAAGGATGGGACGGCACCATCACACAAGGAACCTTGGCACCGTTAT
GAGGTCGCTTGGACAAAACCAACGGAAGCAGAAATTGCAGGATATGATCAATGAAGTCGATGCTG
ATGGCAATGGAACGATTTACTTTCCTGAATTTCTTACTATGATGGCTAGAAAAATGAAGACACA
GACAGCGAAGAGGAATCCGAGAAGCATTCCTGTGTTTTTGACAAGGATGGGAACGGCTACATCAG
CGCTGCTCAGTTACGTACGTCATGACAAAACCTCGGGGAGAAGTTAACAGATGAAGAAGTTGATG

Fig. 10-1

17/18

(SEQ ID. NO:7)

AAATGATAAGGAAGCAGATATCGATGGTGATGCCAAGTAACATATGAAGAGTTTGTACAAATG
ATGACAGCAAAGGGGGAAGAGCGCTGGAAGAAACTTCATTGCCGT CAGCGCTGCCAACCG
GTTCAAGAAGATCTCCGAGCTCATGGTGAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCCCA
TCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTG
GCCACCCCTCGTGACCAACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCGACCACATGA
AGCAGCAGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTC
AAGGACGACGGCAACTACAAGACCCGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGCAACAAGCTGGAGTACA
ACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGACGGCATCAAGTGAACTTC
AAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCC
CATCGGCGACGGCCCCGTGTGTGCTGCCCGACAACCCTACCTGAGCACCCAGTCCGCCCTGAGCA
AAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGATCACT
CTCGGCAAGGACGAGCTGTAA

FIG. 10-2

18/18

(SEQ ID. NO:8)
MLLPVPLLLGLLGAADVSKGEELEFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLLKFIC
TTGKLPVPWPTLLVTTLTHGVQCFSRYPDHMKQHDEFFKSAMPEGYVQERTIFFKDDGNYKTRAENVK
FEGDTLVNRIELKGIDFKEDGNILGHKLEYNFSHNVYIMADKQKNGIKVNEFKIRHNIEDGSVQL
ADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAARMHDQLTEEQIAEFKE
AFSLFDKDGDTITTKELGTVMRSLGQNPTAEELQDMINEVDADGNGTIYFPEFLTMMARKMKDT
DSEEEIREAFRVFDKDGNGYISAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQM
MTAKGKRWRWKNFIAVSAANRFKKISELMVSKGEELEFTGVVPIILVELDGVNGHKFSVSGEGEG
DATYGKLTLLKFICTTGKLPVPWPTLLVTTLTYGVCFSRYPDHMKQHDEFFKSAMPEGYVQERTIFF
KDDGNYKTRAENVKFEEDTLVNRIELKGIDFKEDGNILGHKLEYNFSHNVYIMADKQKNGIKVNF
KIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLLEFVTAAGIT
LGKDEL*

FIG. 10-3