



(43) International Publication Date
26 May 2017 (26.05.2017)

- (51) International Patent Classification:
G01N 33/53 (2006.01) *G01N 33/569* (2006.01)
- (21) International Application Number:
PCT/US2016/062958
- (22) International Filing Date:
19 November 2016 (19.11.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/257,850 20 November 2015 (20.11.2015) US
62/257,863 20 November 2015 (20.11.2015) US
62/257,859 20 November 2015 (20.11.2015) US
62/257,796 20 November 2015 (20.11.2015) US
- (71) Applicant: **DUKE UNIVERSITY** [US/US]; Office of Licensing and Ventures, 2812 Erwin Rd., Suite 306, Box 90083, Durham, NC 27705 (US).
- (72) Inventors: **HELLINGA, Homme W.**; 5212 Pine Cone Lane, Durham, NC 27705 (US). **ALLERT, Malin J.**; 7718 San Gabriel Street, Raleigh, NC 27613 (US).
- (74) Agents: **BEATTIE, Ingrid A.** et al.; Mintz Levin Cohn Ferris Glovsky and Popeo. P.C., One Financial Center, Boston, MA 02111 (US).

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

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

Published:

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

(54) Title: RATIOMETRIC BIOSENSORS AND NON-GEOMETRICALLY MODULATED FRET

Indirect

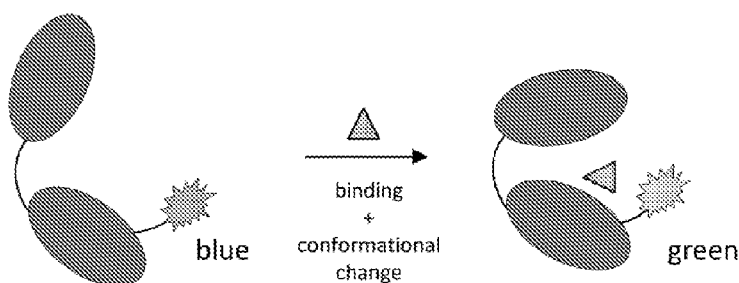


FIG. 1A

(57) Abstract: The present subject matter provides ratiometric biosensors as well as compositions, devices, and methods comprising such biosensors.

RATIOMETRIC BIOSENSORS AND NON-GEOMETRICALLY MODULATED FRET

RELATED APPLICATIONS

5 This application claims benefit of priority to U.S. Provisional Application No. 62/257,850, filed November 20, 2015, U.S. Provisional Application No. 62/257,859, filed November 20, 2015, U.S. Provisional Application No. 62/257,863, filed November 20, 2015, and U.S. Provisional Application No. 62/257,796, filed November 20, 2015, the entire contents of each which are incorporated herein by reference.

INCORPORATION-BY-REFERENCE OF SEQUENCE LISTING

10 The contents of the text file named "35327-521001WO_Sequence_Listing.txt", which was created on November 19, 2016 and is 390 KB in size, is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

15 The present invention relates to compositions and methods for detecting compounds and determining the concentration thereof.

BACKGROUND

20 Determination of analyte concentrations using fluorescent probes is a powerful technique in analytical chemistry. Fluorescent chemosensors have wide-ranging applications in cell biology and analytical chemistry.

25 The majority of fluorescent sensors and biosensors do not undergo changes in emission spectral shape upon analyte binding and accordingly evince monochromatic intensity changes, rather than the dichromatic responses required for ratiometric sensing.

SUMMARY OF THE INVENTION

30 The present subject matter provides methods for converting monochromatic responses into dichromatic responses that enable ratiometric sensing. If the fluorescence emission spectrum changes shape in response to analyte binding such that the ratio of emission intensities at two appropriately chosen wavelengths reports on analyte concentration (dichromatic response), then ratiometric measurements can be used to monitor analyte

concentrations. In embodiments, these methods are based on establishing non-geometrically modulated Förster resonance energy transfer (ngmFRET) between a monochromatic, chemoresponsive fluorophore (a directly responsive partner), and a second fluorophore that neither interacts directly with the ligand, nor is sensitive to ligand-mediated changes in its environment (an indirectly responsive partner). Biosensors that undergo ngmFRET (or altered ngmFRET) upon ligand binding are also provided herein, as well as compositions and devices comprising such biosensors.

Methods, compounds, and compositions provided herein overcome challenges regarding the design of biosensors that produce a ratiometric signal. For example, a biosensor that exhibits a monochromatic response (which does not produce a ratiometric signal) to ligand binding may be converted into a biosensor that produces a dichromatic/ratiometric signal. Moreover, the number of fluorophores that may be utilized in ratiometric biosensors is dramatically increased by the present subject matter. For example, fluorophores that typically do not show a dichromatic response to ligand binding (such as fluorescein and derivatives thereof) may be used together with an additional reporter group (such as another fluorophore) to produce a ratiometric signal. Also included are methods, compounds, and compositions relating to biosensors with multiple reporter groups that have improved ratiometric signals compared to other ratiometric biosensors (*e.g.*, ratiometric biosensors having a single reporter group).

Traditional/conventional geometrically-modulated Fluorescence Resonance Energy Transfer (tgmFRET) is a physical phenomenon that was first described over 50 years ago. In tg mFRET, the transfer of excited state energy from a donor fluorophore to an acceptor fluorophore (*i.e.* energy transfer) is modulated by a ligand-binding event through changes in the distance and/or angle between the donor and acceptor fluorophores. tg mFRET is manifested by opposing changes in the fluorescence emission intensities of the donor and acceptor fluorophores, respectively, in response to ligand binding. For instance, a decrease in distance results in a decrease of the donor fluorescence emission intensity and an increase in the acceptor fluorescence intensity, as energy is transferred from the former to the latter. A ligand-mediated increase in the distance between the partners has the opposite effect (the fluorescence emission intensity of the donor increases, whereas that of the acceptor decreases). In tg mFRET, ligand-mediated modulation of fluorescence intensity arises from global changes in the entire system, and can occur only if both partners are present.

By contrast, in ngmFRET ligand-mediated modulation of fluorescence intensity arises from changes that are localized to the photophysics of the directly responsive fluorophore. Unlike tgmFRET, ligand-mediated changes in fluorescence therefore occur also if only the directly responsive partner is present in isolation by itself. Although the entire ngmFRET system comprising two partners is not required for evincing ligand-mediated changes in fluorescence emission intensity, the response of such a system is qualitatively changed or quantitatively enhanced over the responses of the isolated directly responsive partner (*e.g.* converting a monochromatic into a dichromatic response, thereby enabling ratiometry). Furthermore, unlike tgmFRET, the pattern of fluorescence intensity changes manifested by ligand binding in ngmFRET systems are not limited to opposing changes only. Instead, in ngmFRET almost all combinations of emission intensity changes are possible: opposing changes in the two partners, both partners increase, both decrease, one partner remains unchanged whereas the other increases or decreases. The majority of these responses evince changes that are unequal in magnitude and/or direction (*i.e.* increase, decrease), and accordingly are manifested as ligand-mediated changes in the ratio of the two fluorescence emission intensities. This versatility of ngmFRET system response patterns has great utility in the field of fluorescent biosensors.

The ligand-mediated alteration of the photophysics of the directly responsive partner includes changes to its spectral properties such as the shape of the excitation or emission spectra, and the ratio of radiative to non-radiative emission rates. The fluorescence emission intensity of the indirectly responsive partner in isolation does not change in response to ligand binding; its intensity changes only in the presence of a directly responsive partner in the complete ngmFRET system. In the field fluorescence spectroscopy, the term “quenching” has often been used loosely to refer to a decrease fluorescence emission intensity. However, as used herein, the term “quenching” strictly means a “change in the ratio of radiative to non-radiative emission rates” of a fluorophore.

Aspects of the present subject matter provide biosensors in which ngmFRET occurs between two or more reporter groups (*e.g.*, a donor fluorophore and an acceptor fluorophore) of the biosensor. For example, ngmFRET may change (*e.g.*, increase or decrease) when ligand is bound to the biosensor and a donor fluorophore is contacted with radiation within its excitation wavelength. Effects from tgmFRET and ngmFRET may occur together and be combined into an overall ligand-mediated change in fluorescence emission intensity. In preferred embodiments, less than half or none of the change in overall ligand-mediated

change in fluorescence emission intensity is due to tgmFRET. In embodiments, most of the overall ligand-mediated change in fluorescence emission intensity change is not due to a change in the distance between the donor and acceptor fluorophore or as a result of a change in the orientation between the donor and acceptor fluorophore. In non-limiting examples, less than about 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1%, or 0.5% of the change in overall ligand-mediated change in fluorescence emission intensity is due to tgmFRET. In various embodiments, at least about 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or 99.99% of the ligand-mediated change in fluorescence emission intensity is due to ngmFRET. For example, the change in overall ligand-mediated change in fluorescence emission intensity comprises a spectral change (*e.g.*, in the excitation or emission spectrum) and/or a change in the ratio of the radiative to non-radiative decay rates of one of the fluorophores (by itself and regardless of the presence of any other fluorophore/partner) upon ligand binding.

In some embodiments, ligand binding mediates spectral shifts in the absorption or emission spectrum of the directly responsive partner. In certain embodiments such changes are due at least in part to a switch between different excited states in the ligand-free and ligand-bound biosensor. The two excited states are associated with different transition dipoles. This class of changes is termed “dipole switching” herein. Non-limiting examples of biosensors that show dipole sensing include ttGGBP 17C·Badan- β Zif·Alexa532 and ttGGBP 182C·Acrylodan- β Zif·Alexa532.

In embodiments, the reporter groups include a directly responsive partner (which may be a donor fluorophore or an acceptor fluorophore) and an indirectly responsive partner (which may be a donor fluorophore or an acceptor fluorophore). Depending on context, a “directly responsive” partner is a fluorophore that responds to (i) ligand-induced protein conformational changes upon ligand binding to a ligand-binding protein; or (ii) ligand binding to the directly responsive partner itself. In some embodiments, the directly responsive partner comprises a fluorophore (*i.e.*, it is a directly responsive fluorophore). In various embodiments, the directly responsive fluorophore exhibits a monochromatic or dichromatic spectral change, and/or a change in the ratio of radiative to non-radiative emission rates, upon ligand binding. In certain embodiments relating to ligand binding to the directly responsive partner itself, the directly responsive partner may be a fluorophore such as a fluorescent protein or a small molecule fluorescent compound. An “indirectly responsive” partner is a fluorophore for which no change in emission spectra, excitation spectra, or

change in the ratio of radiative to non-radiative emission rates is caused by ligand binding in the absence of a directly responsive partner. In some embodiments, the indirectly responsive partner comprises a fluorophore (*i.e.*, it is an indirectly responsive fluorophore). When paired with a directly responsive partner with which the indirectly responsive partner is a ngmFRET donor or acceptor, the emission fluorescence intensity of the indirectly responsive partner changes due to a change in energy flow in the ngmFRET pathway upon ligand binding. *See, e.g.*, FIG. 28.

ngmFRET Biosensors

Provided herein are methods, compositions, biosensors, and devices comprising multiple reporter groups, *e.g.* a directly responsive fluorophore and an indirectly responsive fluorophore, between which ngmFRET occurs.

Aspects include a method of detecting a ligand in a sample, comprising contacting a biosensor with a ligand. The biosensor comprises a ligand-binding protein, a directly responsive fluorophore and an indirectly responsive fluorophore. The directly responsive and the indirectly responsive fluorophores are located at two distinct sites of the ligand-binding protein. In some embodiments, the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore. Alternatively, the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore. The method includes contacting the biosensor with radiation comprising a wavelength within the excitation spectrum of the donor fluorophore. When the biosensor is contacted with such radiation, a fluorescence property of the directly responsive fluorophore changes in response to ligand binding. This change in fluorescent property is independent of the indirectly responsive fluorophore, and occurs regardless of whether the indirectly responsive fluorophore is absent or present. The fluorescence properties of the indirectly responsive fluorophore do not change in response to ligand binding in the absence of the directly responsive fluorophore. When the biosensor is contacted with radiation comprising a wavelength within the excitation spectrum of the donor fluorophore, then (i) ngmFRET occurs between the directly responsive fluorophore and the indirectly responsive fluorophore; (ii) fluorescent light is emitted from the biosensor, and the light emitted from the biosensor comprises a combination of light emitted from the directly responsive fluorophore and light emitted from the indirectly responsive fluorophore; and (iii) the ratio of the fluorescence emission intensity emitted from the biosensor at each of two

distinct wavelengths changes in response to ligand binding. In various embodiments, the method further comprises measuring fluorescent light that is emitted from the directly responsive fluorophore and the indirectly responsive fluorophore, and calculating a ratiometric signal to detect the ligand in the sample.

5 The ratiometric signal ($R_{1,2}$) comprises a quotient of two intensities, $I_{\lambda 1}$ and $I_{\lambda 2}$, measured at two independent wavelengths, λ_1 and λ_2 and is calculated according to the following equation:

$$R_{1,2} = I_{\lambda 1} / I_{\lambda 2} .$$

10 In various embodiments, the change in the fluorescent property of the directly responsive fluorophore comprises (i) a bathochromic or hypsochromic shift in the emission or excitation spectrum thereof; and/or (ii) a change in the ratio of radiative to non-radiative emission rates thereof.

15 In some embodiments, the directly responsive fluorophore is Badan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, or 475 nm), and wherein the indirectly responsive fluorophore is 5-iodoacetamidofluorescein (5-IAF) and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 20 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

25 In certain embodiments, the directly responsive fluorophore is Badan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, or 475 nm), and wherein the indirectly responsive fluorophore is Alexa532 and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, or 570 nm).

30 In various embodiments, the directly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, or 465 nm), and wherein the

indirectly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

5 In some embodiments, the directly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, or 475 nm), and wherein the indirectly responsive fluorophore is 5-IAF and emission intensity is
10 measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

 In some embodiments, the directly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and
15 1000nm (*e.g.*, including a wavelength of about 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490 nm), and wherein the indirectly responsive fluorophore is Alexa532 and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556,
20 557, 558, 559, or 560 nm).

 In certain embodiments, the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, or 465 nm), and wherein the indirectly
25 responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

 In various embodiments, the directly responsive fluorophore is Oregon Green and
30 emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, or 465 nm), and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a

wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

In some embodiments, the directly responsive fluorophore is N-(Iodoacetaminoethyl)-
5 1-naphthylamine-5-sulfonic acid (IAEDANS) and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, or 475 nm), and wherein the indirectly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength or range
10 of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

In some embodiments, the directly responsive fluorophore is Alexa532 and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and
15 1000 nm (*e.g.* including a wavelength of about 530, 531, 532, 534, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, or 570 nm), and wherein the indirectly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000 nm (*e.g.* including 470,
20 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 45, 496, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, or 510 nm).

In various embodiments, the directly responsive fluorophore is a yellow fluorescent protein and emission intensity is measured at a wavelength or range of wavelengths between
25 about 400 nm and 1000nm (*e.g.*, including a wavelength of about 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, or 540 nm), and wherein the indirectly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501,
30 502, 503, 504, 505, 506, 507, 508, 509, or 510 nm. In certain embodiments, the directly responsive fluorophore is a yellow fluorescent protein and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533,

534, 535, 536, 537, 538, 539, or 540 nm), and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, or 465 nm).

In embodiments, the directly responsive fluorophore comprises a donor fluorophore and the indirectly responsive fluorophore comprises an acceptor fluorophore. In some embodiments, the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensities of the donor fluorophore and the acceptor fluorophore both decrease upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases, decreases, or remains about the same and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensities of the donor fluorophore and the acceptor fluorophore both increase upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases, decreases, or remains about the same and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases,

decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore.

In embodiments the directly responsive fluorophore comprises an acceptor fluorophore and the indirectly responsive fluorophore comprises a donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore remains about the same and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore remains about the same and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore. In some embodiments, the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore remains about the same, increases, or decreases upon

ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore.

In instances in which an emission intensity increases, the increase may be, *e.g.*, between about 0.1% to 10%, 10% to 50%, or 50% to 100%, or at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 50%, 75%, 100%, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, or 10-fold. In instances in which an emission intensity decreases, the decrease may be, *e.g.*, a decrease of between about at least about 0.1% to 10%, 10% to 50%, or 50% to 100%, or at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, or 90%. In various embodiments in which both the emission intensity of the donor fluorophore and the acceptor fluorophore increases, then the increases are not equal. In certain embodiments in which both the emission intensity of the donor fluorophore and the acceptor fluorophore decreases, then the decreases are not equal.

In various embodiments, the ligand-binding protein comprises the directly responsive fluorophore. For example, the directly responsive fluorophore is formed by an autocatalytic cyclization of an oligopeptide within the ligand-binding protein. In some embodiments, the oligopeptide is located within an interior α helix. In certain embodiments, the oligopeptide comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 consecutive residues. In embodiments, the directly responsive fluorophore is formed by an autocatalytic cyclization of a tripeptide located in an interior α helix of the ligand-binding protein. In various embodiments, ligand-binding protein comprises a yellow fluorescent protein (YFP), *i.e.* the YFP binds to ligand such as a halide anion.

In some embodiments, ligand binding causes a change in signaling by the directly responsive fluorophore.

Also provided is a method of detecting a ligand in a sample, comprising contacting a biosensor with a ligand, wherein the biosensor comprises an amino acid or polypeptide, a directly responsive fluorophore and an indirectly responsive fluorophore. The directly responsive and the indirectly responsive fluorophores are located at two distinct sites of the amino acid or polypeptide, and the directly responsive fluorophore is chemoresponsive. The method may further comprise contacting the biosensor with radiation comprising a wavelength within the excitation spectrum of the donor fluorophore, wherein (i) a fluorescence property of the directly responsive fluorophore changes in response to ligand binding in the absence or presence of the indirectly responsive fluorophore; (ii) a

fluorescence property of the indirectly responsive fluorophore does not change in response to ligand binding in the absence of the directly responsive fluorophore; (iii) ngmFRET occurs between the directly responsive fluorophore and the indirectly responsive fluorophore; (iv) fluorescent light is emitted from the biosensor, wherein the light emitted from the biosensor comprises a combination of light emitted from the directly responsive fluorophore and light emitted from the indirectly responsive fluorophore; and (v) the ratio of the fluorescence emission intensity emitted from the biosensor at each of two distinct wavelengths changes in response to ligand binding. The method may also include measuring fluorescent light that is emitted from the directly responsive fluorophore and the indirectly responsive fluorophore and calculating a ratiometric signal, to detect the ligand in the sample. The ratiometric signal ($R_{1,2}$) comprises a quotient of two intensities, $I_{\lambda 1}$ and $I_{\lambda 2}$, measured at two independent wavelengths, λ_1 and λ_2 and is calculated according to the following equation:

$$R_{1,2} = I_{\lambda 1} / I_{\lambda 2} .$$

As used herein, a “chemoresponsive” fluorophore is a fluorophore to which ligand binds, wherein ligand binding causes a change in signaling by the fluorophore.

As used herein, “signaling” refers to the emission of energy (which may be referred to as a “signal”) by one or more reporter groups. In various implementations, the signal comprises electromagnetic radiation such as a light. In some embodiments, the signal is detected as a complete emission spectrum (or spectra) or a portion (or portions) thereof. For example, a signal may comprise emitted light at a particular wavelength or wavelengths, or range(s) of wavelengths. In some embodiments, a change in signaling comprises a spectral change (*e.g.*, a spectral shift and/or change in intensity). In some embodiments, a change in signaling comprises a dichromatic shift or a monochromatic fluorescence intensity change.

In some embodiments, the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore. Alternatively, the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore.

In various embodiments, the change in the fluorescent property of the directly responsive fluorophore comprises (i) a bathochromic or hypsochromic shift in the emission or excitation spectrum thereof; and/or (ii) a change in the ratio of radiative to non-radiative emission rates thereof.

In some embodiments, the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, or 470 nm), and wherein the indirectly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm).

In certain embodiments, the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, or 530 nm), and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength or range of wavelengths between about 400 nm and 1000nm (*e.g.*, including a wavelength of about 445, 446, 447, 448, 449, 450, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, nm).

Any amino acid or polypeptide may be used to link the chemoresponsive directly responsive fluorophore with the indirectly responsive fluorophore, provided the two fluorophores are close enough for nmFRET to occur. Suitable distances may be determined in part by the distance-dependence of the energy transfer between a given donor-acceptor pair (*see, e.g.* J.R. Lakowicz, 2006, *Principles of Fluorescence Spectroscopy*, *Springer*, incorporated herein by reference). In various embodiments, the amino acid or polypeptide comprises 1 amino acid, or a stretch of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, 750, or 1000 amino acids. In some embodiments, the amino acid or polypeptide comprises at least 1, 2, or 3 thiol groups; at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group; at least 1, 2, or 3 primary amine groups; or at least 1, 2, or 3 lysines that each comprise a primary amine. In certain embodiments, the polypeptide comprise two cysteines, and there is no disulfide bond between the two cysteines. In some embodiments there is no disulfide bond between any pair of cysteines within the amino acid sequence of the polypeptide.

In a non-limiting example, the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of ecTRX (SEQ ID NO: 151). In some embodiments,

the polypeptide comprises a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In certain embodiments, the polypeptide comprises a mutant of ecTRX comprising a D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In various embodiments, the polypeptide comprises a mutant of ecTRX that does not comprise a lysine. In certain embodiments, the polypeptide comprises amino acids in the sequence of any one of SEQ ID NOS: 69-86 or 151.

In certain embodiments, the polypeptide further comprises a hexahistidine tag.

In some embodiments, the ligand comprises a hydrogen ion. For example, the biosensor for pH, wherein the directly responsive fluorophore is pH-sensitive. In various embodiments, the fully excited emission intensity of the directly responsive fluorophore is different at a pH less than about 7.0 (e.g. 6.9, 6.8, 6.7, 6.6, 6.5, 6.4, 6.3, 6.2, 6.1, or 6.0), or about 4.0 to 10.0, compared to a pH of about 7.3, 7.4, 7.5, 7.6, or 7.7.

In various embodiments, the directly responsive fluorophore comprises a pH-sensitive fluorophore comprising fluorescein or a derivative thereof. In embodiments, the directly responsive fluorophore transitions from a monoanion to a dianion at a pH that is less than 7.0 in an aqueous solution.

In certain embodiments, the indirectly responsive fluorophore is attached to the ligand-binding protein via a covalent bond. Various approaches for attaching reporter groups such as directly and indirectly responsive fluorophores to an amino acid or a polypeptide such as a ligand-binding protein are described herein. In some embodiments, the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction.

In some embodiments, the indirectly responsive fluorophore is attached to the ligand-binding protein via a non-covalent bond. In certain embodiments, the indirectly responsive fluorophore is attached to a cysteine or a lysine of the protein.

In various embodiments, the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein. In some embodiments, the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif.

In some embodiments, fluorophore attachment motif comprises an amino acid or polypeptide. Various embodiments may be used to link a fluorophore with a ligand-binding protein. In some embodiments, the amino acid or polypeptide comprises 1 amino acid, or a stretch of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250,
 5 500, 750, or 1000 amino acids. In a non-limiting example, the polypeptide comprises amino acids in the sequence of β Zif (SEQ ID NO: 42). In another non-limiting example, the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).

10 In some embodiments, the directly responsive fluorophore is attached to the ligand-binding protein via a covalent bond. In various embodiments, the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction. In directly responsive fluorophore is attached to a cysteine or a lysine of the protein.

15 In various embodiments, if the acceptor fluorophore comprises palladium, platinum, ruthenium, or osmium, then the acceptor fluorophore is not attached to the amino group of the N-terminus of the ligand-binding protein. In some embodiments, the acceptor fluorophore does not comprise $[\text{Ru}(\text{bpy})_3]^{2+}$, $[\text{Ru}(\text{Ph}_2\text{phen})_3]^{2+}$, $[\text{Ru}(\text{bpy})_2(\text{dcbpy})]^{2+}$, or $[\text{Ru}(\text{bpy})_2(\text{phen-ITC})]^{2+}$, where bpy is 2,2'-bipyridine, phen is 1,10-phenanthroline, dcbpy is
 20 4,4'-dicarboxy-2,2'-bipyridine, and ITC is isothiocyanate. In certain embodiments, the biosensor does not comprise an *E. coli* glutamine-binding protein with Acrylodan attached to 179C. In some embodiments, the biosensor does not comprise *E. coli* glucose-binding protein with Acrylodan attached to 255C.

25 In some embodiments, an overlap of the emission spectrum of the donor fluorophore and the excitation spectrum of the acceptor fluorophore increases upon ligand binding. In certain embodiments, the directly responsive fluorophore comprises the donor fluorophore, and the increase results from a bathochromic shift in the emission spectrum of the donor fluorophore. Alternatively, the directly responsive fluorophore comprises the acceptor fluorophore, and the increase results from a hypsochromic shift in the excitation spectrum of
 30 the acceptor fluorophore.

In various embodiments, an overlap of the emission spectrum of the donor fluorophore and the excitation spectrum of the acceptor fluorophore decreases upon ligand binding. In some embodiments, the directly responsive fluorophore comprises the donor

fluorophore, and the decrease results from a hypsochromic shift in the emission spectrum of the donor fluorophore. In certain embodiments, the directly responsive fluorophore comprises the acceptor fluorophore, and the decrease results from a bathochromic shift in the excitation spectrum of the acceptor fluorophore.

5 In some embodiments, the directly responsive fluorophore has a monochromatic spectral change upon ligand binding. Alternatively, the directly responsive fluorophore has a dichromatic spectral change upon ligand binding.

In certain embodiments, the emission intensity of the donor fluorophore and/or the acceptor fluorophore increases in two phases as ligand concentration increases.

10 In various embodiments, the ratio of radiative to non-radiative emission or intensity of the directly responsive fluorophore increases by at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 50%, 75%, 100%, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, or 10-fold upon ligand binding to the ligand-binding protein. Alternatively, the ratio of radiative to non-radiative emission or intensity of the directly responsive fluorophore
15 decreases by at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 50%, 75%, 90%, 95%, or 99% upon ligand binding to the ligand-binding protein.

In embodiments, the directly responsive fluorophore and the indirectly responsive fluorophore are not a naphthalene derivative. In some embodiments, the directly responsive fluorophore and the indirectly responsive fluorophore are not Prodan, Acrylodan, or Badan.

20 In certain embodiments, the directly responsive fluorophore is not a naphthalene derivative. In some embodiments, the directly responsive fluorophore is not Prodan, Acrylodan, or Badan.

In various embodiments, the directly responsive fluorophore comprises xanthene, a xanthene derivative, fluorescein, a fluorescein derivative, coumarin, a coumarin derivative, cyanine, a cyanine derivative, rhodamine, a rhodamine derivative, phenoxazine, a phenoxazine derivative, squaraine, a squaraine derivative, coumarin, a coumarin derivative, oxadiazole, an oxadiazole derivative, anthracene, an anthracene derivative, a boradiazaindacene (BODIPY) family fluorophore, pyrene, a pyrene derivative, acridine, an acridine derivative, arylmethine, an arylmethine derivative, tetrapyrrole, or a tetrapyrrole
25 derivative. In some embodiments, the directly responsive fluorophore comprises fluorescein or a derivative thereof.
30

In some embodiments, the directly responsive fluorophore and/or the indirectly responsive fluorophore comprises a fluorescent protein. In various embodiments, the directly

responsive fluorophore and/or the indirectly responsive fluorophore comprises an organic compound having a molecular weight less than about 2000 Da (*e.g.*, 5-iodoacetamidofluorescein (5-IAF) or 6-iodoacetamidofluorescein (6-IAF), rhodamine, Oregon Green, eosin, Texas Red, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, Badan, Acrylodan, IAEDANS, comprising 3-cyano-7-hydroxycoumarin, 7-hydroxycoumarin-3-carboxylic acid, 6,8-difluoro-7-hydroxy-4-methylcoumarin, or 7-amino-4-methylcoumarin, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, DRAQ5, DRAQ7, or CyTRAK Orange, cascade blue, Nile red, Nile blue, cresyl violet, oxazine 170, proflavin, acridine orange, acridine yellow, auramine, crystal violet, malachite green, porphyrin, phthalocyanine, bilirubin, pyrene, N,N'-dimethyl-N-(iodoacetyl)-N'-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamide (NBD), N-((2-(iodoacetoxy)ethyl)-N-methylamino)-7-nitrobenz-2-oxa-1,3-diazole (NBDE), JPW4039, JPW4042, JPW4045, Pacific Blue, CPM, N,N'-Dimethyl-N-(Iodoacetyl)-N'-(7-Nitrobenz-2-Oxa-1,3-Diazol-4-yl)Ethylenediamine (IANBD), 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM), BODIPY 499, BODIPY 507/545, BODIPY 499/508, Alexa 432, Alexa488, Alexa532, Alexa546, Cy5, or 1-(2-maleimidylethyl)-4-(5-(4-methoxyphenyl)oxazol-2-yl)pyridinium methanesulfonate (PyMPO maleimide) (PyMPO)). Numerous combinations of directly responsive fluorophores and indirectly responsive fluorophores are possible. For example, in various non-limiting examples, (a) the donor fluorophore comprises Pacific Blue and the acceptor fluorophore comprises 5-IAF or 6-iodoacetamidofluorescein (6-IAF); (b) the donor fluorophore comprises Pacific Blue and the acceptor fluorophore comprises Oregon Green; (c) the donor fluorophore comprises IAEDANS and the acceptor fluorophore comprises 5-IAF or 6-IAF; (d) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises Alexa532; (e) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises 5-IAF or 6-IAF; (f) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises Pacific Blue or YFP; (g) the donor fluorophore comprises 5-IAF or 6-IAF and the acceptor fluorophore comprises Pacific Blue; (h) the donor fluorophore comprises badan and the acceptor fluorophore comprises 5-IAF or 6-IAF; or (i) the donor fluorophore comprises badan and the acceptor fluorophore comprises Alexa532.

Any of the ligand-binding proteins disclosed herein, as well as others, may be included in the biosensors and methods that are provided. In some embodiments, the ligand-binding protein is selected from the group consisting of a glucose-galactose binding protein (GGBP), a glucose-binding protein, a urea-binding protein (UBP), a lactate-binding protein

(LacBP), a calcium-binding protein, a calcium-bicarbonate binding protein (BicarbBP), and an iron-bicarbonate binding protein (FeBP).

Aspects include a biosensor for a ligand comprising a ligand-binding protein, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the ligand-binding-protein, wherein (i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or (ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is an donor fluorophore, and wherein if the acceptor fluorophore comprises ruthenium or osmium, then the acceptor fluorophore is not attached to the amino group of the N-terminus of the ligand-binding protein.

In some embodiments, the ligand-binding protein comprises the directly responsive fluorophore. In certain embodiments, the directly responsive fluorophore is formed by an autocatalytic cyclization of an oligopeptide within the ligand-binding protein.

In various embodiments, the ligand-binding protein comprises a Yellow Fluorescent Protein (YFP; SEQ ID NO: 149) or a fluorescent mutant thereof, and the ligand comprises a halide anion. For example, the halide anion comprises a fluoride (F^-), chloride (Cl^-), a bromide (Br^-), an iodide (I^-), an astatide (At^-) anion, or an ununseptide (Ts^-) anion. In some embodiments, the mutant comprises a mutation that alters the interaction of the mutant with a bound halide anion compared to YFP. In certain embodiments, the mutant comprises a mutation that alters the affinity and/or specificity of the mutant for a halide anion compared to YFP. In various embodiments, the ligand-binding protein comprises 1, 2, 3, 4, or 5 halide anion binding sites.

In some embodiments, at least one amino acid of the YFP or the fluorescent mutant thereof has been substituted with a cysteine. For example, the cysteine is within a first β -strand (β_1), a second β -strand (β_2), a third β -strand (β_3), a fourth β -strand (β_4), a fifth β -strand (β_5), a sixth β -strand (β_6), a seventh β -strand (β_7), an eighth β -strand (β_8), a ninth β -strand (β_9), a tenth β -strand (β_{10}), or an eleventh β -strand (β_{11}) of the YFP or the fluorescent mutant thereof. In certain embodiments, the ligand-binding protein comprises one or more of the following substitutions: E17X, E32X, T43X, F64X, G65X, L68X, Q69X, A72X, H77X, K79X, R80X, E95X, R109X, R122X, D133X, H148X, N149X, V163X, N164X, D173X, Y182X, Q183X, Y203X, Q204X, L221X, and H231X, wherein X is any amino acid, a conservative substitution, or a cysteine, wherein each YFP amino acid position is numbered

as in SEQ ID NO: 150. In non-limiting examples, the ligand-binding protein comprises one or more of the following substitutions: F64L, G65T, L68V, Q69T, A72S, K79R, R80Q, H148Q, H148G, V163A, H231L, H148Q, or Q183A, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150. In various embodiments, the ligand-binding protein
 5 comprises an R at the 96 position, a Y at the 203 position, a S at the 205 position, and an E at the 222 position, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150.

In various embodiments, ligand binding causes a change in signaling by the directly responsive fluorophore. In embodiments, the ligand-binding protein comprises a mutation compared to a naturally occurring protein. For example, at least one amino acid of the
 10 ligand-binding protein has been substituted with a cysteine. In some embodiments, the ligand-binding protein comprises a mutant of a microbial ligand-binding protein. In certain embodiments, the ligand-binding protein comprises a mutant of a microbial periplasmic ligand-binding protein.

In certain embodiments, the ligand comprises glucose, galactose, lactose, arabinose, ribose, maltose, lactate, urea, bicarbonate, phosphate, sulfate, chloride, fluoride, iodide, 15 aspartate, ununseptide, bromide, calcium, a hydrogen ion, a dipeptide, histidine, glutamine, glutamate, aspartate, or iron.

In some embodiments, the ligand-binding protein comprises a GGBP. For example, the GGBP comprises or comprises a mutant of: an *Escherichia* sp. GGBP; a
 20 *Thermoanaerobacter* sp. GGBP; a *Clostridium* sp. GGBP; a *Salmonella* sp. GGBP; a *Caldicellulosiruptor* sp. GGBP; a *Paenibacillus* sp. GGBP; a *Butyrivibrio* sp. GGBP; a *Roseburia* sp. GGBP; a *Faecalibacterium* sp. GGBP; an *Erysipelothrix* sp. GGBP; or an *Eubacterium* sp. GGBP.

In some embodiments, the ligand-binding protein comprises a UBP. For example, the
 25 UBP comprises or comprises a mutant of: an *Marinomas* sp. UBP; a *Marinobacter* sp. UBP; a *Bacillus* sp. UBP; a *Desulfotomaculum* sp. UBP; a *Geobacillus* sp. UBP; a *Clostridium* sp. UBP; a *Caldicellulosiruptor* sp. UBP; a *Thermocrinis* sp. UBP; a *Synechoccus* sp. UBP; a *Paenibacillus* sp. UBP; or a *Thermosynechococcus* sp. UBP.

In some embodiments, the ligand-binding protein comprises a GBP. For example, the
 30 GBP comprises or comprises a mutant of: an *Thermus* sp. GBP; a *Deinococcus* sp. GBP; a *Thermotoga* sp. GBP; a *Kosmotoga* sp. GBP; a *Bacillus* sp. GBP; a *Staphylothermus* sp. GBP; or an *Arthrobacter* sp. GBP.

In some embodiments, the ligand-binding protein comprises a LacBP. For example, the LacBP comprises or comprises a mutant of: a *Thermus* sp. LacBP; a *Thioalkalivibrio* sp. LacBP; a *Roseobacter* sp. LacBP; a *Marinobacter* sp. LacBP; a *Anaeromyxobacter* sp. LacBP; a *Pseudomonas* sp. LacBP; a *Rhodobacter* sp. LacBP; a *Flexistipes* sp. LacBP; or a
 5 *Thermanaerovibrio* sp. LacBP.

In some embodiments, the ligand-binding protein comprises a calcium-binding protein or a BicarbbBP. For example, the ligand-binding protein comprises or comprises a mutant of: a *Synechocystis* sp. BicarbbBP; a *Thermosynechococcus* sp. BicarbbBP; a *Chroococcidiopsis* sp. BicarbbBP; a *Calothrix* sp. BicarbbBP; a *Anabaena* sp. BicarbbBP; or a
 10 *Chamaesiphon* sp. BicarbbBP.

In some embodiments, the ligand-binding protein comprises a FeBP. For example, the ligand-binding protein comprises or comprises a mutant of: a *Mannheimia* sp. FeBP; an *Exiguobacterium* sp. FeBP; a *Thermosynechococcus* sp. FeBP; a *Candidatus Nitrospira* sp. FeBP; a *Thermus* sp. FeBP; a *Meiothermus* sp. FeBP; a *Salinibacter* sp. FeBP; or a
 15 *Halorubrum* sp. FeBP.

Also provide is a biosensor for a ligand comprising an amino acid or a polypeptide, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the amino acid or polypeptide, wherein the directly responsive fluorophore is chemoresponsive,
 20 and wherein (i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or (ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore.

As noted above, any amino acid or polypeptide may be used to link the chemoresponsive directly responsive fluorophore with the indirectly responsive fluorophore,
 25 provided the two fluorophores are close enough for nmFRET to occur. In some embodiments, the amino acid or polypeptide comprises 1 amino acid, or a stretch of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, 750, or 1000 amino acids.

In some embodiments, the polypeptide comprises a stretch of at least 50, 60, 70, 80,
 30 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of ecTRX (SEQ ID NO: 151). In certain embodiments, the polypeptide comprises a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination

thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In some embodiments, the polypeptide comprises a mutant of ecTRX comprising a D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein
5 each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In some embodiments, the polypeptide comprises a mutant of ecTRX that does not comprise a lysine. In various embodiments, the biosensor comprises amino acids in the sequence of any one of SEQ ID NOS: 69-86 or 151.

In some embodiments, the polypeptide further comprises a hexahistidine tag.

10 In certain embodiments, the amino acid or polypeptide comprises at least 1, 2, or 3 thiol groups; at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group; at least 1, 2, or 3 primary amine groups; or at least 1, 2, or 3 lysines that each comprise a primary amine. In some embodiments, there is no disulfide bond between cysteines within the amino acid sequence of the polypeptide.

15 In various embodiments, the ligand comprises a hydrogen ion. In some embodiments, the biosensor is a biosensor for pH, wherein the directly responsive fluorophore is pH-sensitive. In certain embodiments, the fully excited emission intensity of the directly responsive fluorophore is different at a pH less than about 7.0 compared to a pH of 7.5. In some embodiments, the directly responsive fluorophore comprises a pH-sensitive fluorophore
20 comprising fluorescein or a derivative thereof. In some embodiments, the directly responsive fluorophore transitions from a monoanion to a dianion at a pH that is less than 7.0 in an aqueous solution.

In some embodiments, the directly responsive fluorophore is attached to the ligand-binding protein, the amino acid, or the polypeptide via a covalent bond. In some
25 embodiments, the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction. In certain embodiments, the directly responsive fluorophore is attached to a cysteine or a lysine of the protein.

In various embodiments, the indirectly responsive fluorophore is attached to the N-
30 terminus or the C-terminus of the protein. In some embodiments, the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif. In some embodiments, the fluorophore attachment motif comprises an amino acid or a polypeptide. In certain embodiments, the polypeptide comprises amino acids

in the sequence of β Zif (SEQ ID NO: 42). In various embodiments, polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).

5 In certain embodiments, the indirectly responsive fluorophore is attached to the ligand-binding protein via a covalent bond. In some embodiments, the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction. In various embodiments, the indirectly responsive fluorophore is attached to a cysteine or a lysine of the protein.

10 Aspects of the present subject matter further provide a method for assaying the level of a ligand in a subject, comprising contacting a biosensor with a biological sample from the subject. Non-limiting examples of ligands include glucose, galactose, lactose, arabinose, ribose, maltose, lactate, urea, bicarbonate, phosphate, sulfate, chloride, fluoride, iodide, astatide, ununseptide, bromide, calcium, a hydrogen ion, a dipeptide, histidine, glutamine,
15 glutamate, aspartate, and iron.

 In some embodiments, the subject has or is suspected of having abnormal kidney function, abnormal adrenal gland function, diabetes, hypochloremia, bromism, hypothyroidism, hyperthyroidism, cretinism, depression, fatigue, obesity, a low basal body temperature, a goiter, a fibrocystic breast change, lactic acidosis, septic shock, carbon
20 monoxide poisoning, asthma, a lung disease, respiratory insufficiency, Chronic Obstructive Pulmonary Disease (COPD), regional hypoperfusion, ischemia, severe anemia, cardiac arrest, heart failure, a tissue injury, thrombosis, or a metabolic disorder, diarrhea, shock, ethylene glycol poisoning, methanol poisoning, diabetic ketoacidosis, hypertension, Cushing syndrome, liver failure, cancer, or an infection.

25 In various embodiments, the biological sample comprises sweat, tear fluid, blood, serum, plasma, interstitial fluid, amniotic fluid, sputum, gastric lavage, skin oil, milk, fecal matter, emesis, bile, saliva, urine, mucous, semen, lymph, spinal fluid, synovial fluid, a cell lysate, venom, hemolymph, or a fluid obtained from a plant.

 Also provided is a method for assaying the level of ligand in an environmental
30 sample, comprising contacting a biosensor with the environmental sample. In some embodiments, the environmental sample is from an environmental site that is suspected of being polluted. In some embodiments, the environmental sample has been obtained or provided from an environmental substance, fluid, or surface. In various embodiments, the

environmental substance comprises (a) rock, soil, clay, sand, a meteorite, an asteroid, dust, plastic, metal, a mineral, a fossil, a sediment, or wood; (b) the environmental surface comprises the surface of a satellite, a bike, a rocket, an automobile, a truck, a motorcycle, a yacht, a bus, or a plane, a tank, an armored personnel carrier, a transport truck, a jeep, a mobile artillery unit, a mobile antiaircraft unit, a minesweeper, a Mine-Resistant Ambush Protected (MRAP) vehicle, a lightweight tactical all-terrain vehicle, a high mobility multipurpose wheeled vehicle, a mobile multiple rocket launch system, an amphibious landing vehicle, a ship, a hovercraft, a submarine, a transport plane, a fighter jet, a helicopter, a rocket, or an Unmanned Aerial Vehicle, a drone, a robot, a building, furniture, or an organism; or (c) the environmental fluid comprises marine water, well water, drinking well water, water at the bottom of well dug for petroleum extraction or exploration, melted ice water, pond water, aquarium water, pool water, lake water, mud, stream water, river water, brook water, waste water, treated waste water, reservoir water, rain water, or ground water.

Aspects of the present subject matter further provide a method for monitoring the level of a ligand, comprising periodically continuously detecting the level of the ligand, wherein detecting the level of the ligand comprises (a) providing or obtaining a sample; (b) contacting the sample with a biosensor for the ligand; and (c) detecting a signal produced by the biosensor. In some embodiments, the sample is provided or obtained from a subject or from a culture of microbial cells.

Additional embodiments and methods for detecting the presence and/or amount of a ligand are disclosed herein.

Aspects of the present subject matter also provide a method for constructing a biosensor, comprising: (a) providing a ligand-binding protein; (b) identifying at least one putative allosteric, endosteric, or peristeric site of the ligand-binding based a structure of the ligand-binding protein; (c) mutating the ligand-binding protein to substitute an amino acid at the at least one putative allosteric, endosteric, or peristeric site of the second protein with a cysteine; (d) conjugating a donor fluorophore or an acceptor fluorophore to the cysteine to produce single labeled biosensor; (e) detecting whether there is a spectral shift or change in emission intensity of the single labeled biosensor upon ligand binding when the donor fluorophore or the acceptor fluorophore is fully excited; and (f) if a spectral shift or change in emission intensity is detected in (e), attaching a donor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine, and attaching an acceptor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine.

In various embodiments, the ligand-binding protein has been identified by (i) selecting a first protein having a known amino acid sequence (seed sequence), wherein the first protein is known to bind a ligand; (ii) identifying a second protein having an amino acid sequence (hit sequence) with at least 15% sequence identity to the seed sequence; (iii) 5 aligning the seed amino acid sequence and the hit sequence, and comparing the hit sequence with the seed sequence at positions of the seed sequence that correspond to at least 5 primary complementary surface (PCS) amino acids, wherein each of the at least 5 PCS amino acids has a hydrogen bond interaction or a van der Waals interaction with ligand when ligand is bound to the first protein; and (iv) identifying the second protein to be a ligand-binding 10 protein if the hit sequence comprises at least 5 amino acids that are consistent with the PCS.

In some embodiments, the spectral shift comprises a monochromatic fluorescence intensity change or a dichromatic spectral shift.

Also provided is a method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand 15 binding, the method comprising (a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and (b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation 20 spectrum of the first reporter group.

The present subject matter also includes method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising (a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand- 25 binding fluorescent protein; and (b) attaching an acceptor fluorophore or a donor fluorophore to the biosensor, wherein (i) the acceptor fluorophore has an excitation spectrum that overlaps with the emission spectrum of the fluorescent protein; or (ii) the donor fluorophore has an emission spectrum that overlaps with the excitation spectrum of the fluorescent protein.

Also provided is a method of increasing a dichromatic response of a biosensor to 30 ligand binding, the method comprising (a) selecting a biosensor that exhibits a dichromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and (b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum

of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of the first reporter group.

In some embodiments, the second reporter group is within about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, or 200 angstroms (Å) of the first reporter group regardless of whether ligand is bound to the biosensor. Suitable distances may be determined in part by the distance-dependence of the energy transfer between a given donor-acceptor pair (*see, e.g.* J.R. Lakowicz, 2006, Principles of Fluorescence Spectroscopy, *Springer*, incorporated herein by reference). In some embodiments, when the ligand is bound to the biosensor, the average distance between the first reporter group and the second reporter group changes by less than about 5, 4, 3, 2, 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, or 0.01 angstroms (Å) compared to when ligand is not bound to the ligand-binding protein.

The present subject matter further provides a method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising (a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein said biosensor comprises an amino acid or polypeptide and a first reporter group, wherein the first reporter group comprises a chemoresponsive fluorophore; and (b) attaching a second reporter group to said biosensor, wherein said second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of said first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of said first reporter group.

Also included is a method of increasing a dichromatic response of a biosensor to ligand binding, the method comprising (a) selecting a biosensor that exhibits a dichromatic response upon ligand binding, wherein said biosensor comprises an amino acid or a polypeptide and a first reporter group, wherein the first reporter group comprises a chemoresponsive fluorophore; and (b) attaching a second reporter group to said biosensor, wherein said second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of said first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of said first reporter group.

tgmFRET Biosensors

While ngmFRET is preferred to tgmFRET, tgmFRET may be used alternatively or in addition to ngmFRET in certain embodiments.

In various embodiments, the biosensor comprises multiple reporter groups, including a first reporter group and a second reporter group. For example, the first reporter group may comprise a donor fluorophore and the second reporter group may comprise an acceptor fluorophore. In certain embodiments, FRET is detectable by a change in the fluorescence of the acceptor fluorophore or by a decrease in donor fluorophore fluorescence. In various embodiments, the donor fluorophore, and/or the acceptor fluorophore is fluorescent. In some embodiments, both the donor fluorophore and the acceptor fluorophore are fluorescent.

In various embodiments, the angle and/or distance between the donor fluorophore and the acceptor fluorophore changes upon ligand binding. In some embodiments, neither the donor fluorophore nor the acceptor fluorophore is directly responsive to ligand binding. In some embodiments the donor fluorophore and/or the acceptor fluorophore is attached to the N-terminus or the C-terminus of the ligand-binding protein (*e.g.*, directly or via a fluorophore attachment motif). In certain embodiments, the donor fluorophore and/or the acceptor fluorophore is attached to a fluorophore attachment motif. For example, the fluorophore attachment motif may be conjugated to the N-terminus or the C-terminus of the ligand-binding protein.

In some embodiments, the donor fluorophore and/or the acceptor fluorophore comprises a fluorescent protein. In various embodiments, the donor fluorophore and/or the acceptor fluorophore comprises an organic compound having a molecular weight less than about 2000 Da (*e.g.*, 5-iodoacetamidofluorescein (5-IAF) or 6-iodoacetamidofluorescein (6-IAF), rhodamine, Oregon Green, eosin, Texas Red, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, Badan, Acrylodan, IAEDANS, comprising 3-cyano-7-hydroxycoumarin, 7-hydroxycoumarin-3-carboxylic acid, 6,8-difluoro-7-hydroxy-4-methylcoumarin, or 7-amino-4-methylcoumarin, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, DRAQ5, DRAQ7, or CyTRAK Orange, cascade blue, Nile red, Nile blue, cresyl violet, oxazine 170, proflavin, acridine orange, acridine yellow, auramine, crystal violet, malachite green, porphyrin, phthalocyanine, bilirubin, pyrene, N,N'-dimethyl-N-(iodoacetyl)-N'-(7-nitrobenz-2-oxo-1,3-diazol-4-yl)ethylenediamide (NBD), N-((2-(iodoacetoxy)ethyl)-N-methyl-1-amino-7-nitrobenz-2-oxo-1,3-diazole (NBDE), Acrylodan, JPW4039, JPW4042, JPW4045, Oregon Green, Pacific Blue, CPM, N,N'-Dimethyl-N-

(Iodoacetyl)-N'-(7-Nitrobenz-2-Oxa-1,3-Diazol-4-yl)Ethylenediamine (IANBD), 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM), BODIPY 499, BODIPY 507/545, BODIPY 499/508, Alexa 432, Alexa488, Alexa532, Alexa546, Cy5, or 1-(2-maleimidylethyl)-4-(5-(4-methoxyphenyl)oxazol-2-yl)pyridinium methanesulfonate (PyMPO maleimide) (PyMPO)). For example, the organic compound is a fluorophore. Numerous combinations of donor and acceptor fluorophores are possible.

Reporter Group Attachment

Aspects of the present subject matter provide a biosensor that comprises a one or more reporter groups attached to a ligand-binding protein, wherein binding of a ligand to a ligand-binding domain of the ligand-binding protein causes a change in signaling by the reporter group. In various embodiments, the reporter group is attached to an endosteric site, an allosteric site, or a peristeric site of the ligand-binding protein. In embodiments, the reporter group is covalently or noncovalently attached to the ligand-binding protein.

For convenience and depending on context, a reporter group may be referred to by a name of an unattached form of the reporter group regardless of whether the reporter group is attached to a ligand-binding protein. For example, a compound known as "Compound A" when in an unconjugated form may be referred to herein as "Compound A" when in a form that is attached to a ligand-binding protein. In a specific example, the term "Acrylodan" is used to refer to unreacted/unconjugated Acrylodan, as well as Acrylodan that is conjugated to a ligand-binding protein.

In certain embodiments, a biosensor comprises a reporter group that is conjugated to a ligand-binding protein, and the reporter group is conjugated to an amino acid of the protein that is at least about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 angstroms (Å) from the ligand when the ligand is bound to the protein. In embodiments, the reporter group is conjugated to an amino acid of the protein that is about 0.1 Å to about 5 Å, about 5 Å to about 10 Å, about 10 Å to about 20 Å, about 20 Å to about 50 Å, about 50 Å to about 75 Å, or about 75 Å to about 100 Å from the ligand when the ligand is bound to the protein. In some embodiments, the reporter group is conjugated to an amino acid of the protein that is within an α -helix or a β -strand. In some embodiments, the reporter group is conjugated to an amino acid that (i) is not within an α -helix or a β -strand, but is within about 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acids of an amino

acid of the protein's amino acid sequence that is within an α -helix or a β -strand. In some embodiments, the reporter group is conjugated to an amino acid that is in an inter-domain hinge amino acid region between two domains of a protein. In some embodiments, the reporter group is conjugated to an amino acid that is in an inter-domain hinge amino acid region between (i) a α -helix and a β -strand; (ii) two α -helices; or (iii) two β -strands of a protein. In some embodiments, the reporter group is conjugated to an amino acid (*e.g.*, a cysteine such as a cysteine added by substitution compared to a naturally corresponding polypeptide) between positions 1-25, 25-50, 50-75, 75-100, 100-125, 125-150, 150-175, 175-200, 200-225, 225-250, 250-275, 275-350, 275-300, 275-325, 300-325, 300-350, 300-400, or 350-450 (inclusive) of a polypeptide (*e.g.*, not including N-terminal fusion proteins compared to the polypeptide's naturally occurring counterpart).

In certain embodiments, the directly or indirectly responsive fluorophore is conjugated (directly or via a fluorophore attachment motif) to an amino acid that is no more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino acids from the N-terminus or the C-terminus of the ligand-binding protein. In some embodiments, the directly or indirectly responsive fluorophore is conjugated (directly or via a fluorophore attachment motif) to an amino acid that is at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino acids from the N-terminus or the C-terminus of the ligand-binding protein. In some embodiments, about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino acids (including or not including the signal peptide) have been deleted (*e.g.* are absent) from the N-terminus of the protein compared to its naturally occurring counterpart. In some embodiments, less than 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino acids (including or not including the signal peptide) have been deleted (*e.g.* are absent) from the N-terminus of the protein compared to its naturally occurring counterpart. In some embodiments, about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino

acids have been deleted (*e.g.* are absent) from the C-terminus of the protein compared to its naturally occurring counterpart. In some embodiments, less than 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-50, 25-75, or 25-100 amino acids have been deleted (*e.g.* are absent) from the C-terminus of the protein compared to its naturally occurring counterpart.

Periplasmic binding proteins are characterized by two lobes connected by a hinge region; ligand bind at a location at the interface between the two domains. Such proteins or engineered versions thereof (as described herein) can adopt two different conformations: a ligand-free open form and a ligand-bound closed form, which interconvert through a relatively large bending motion around the hinge (FIG. 1A; Dwyer et al., 2004, Current Opinion in Structural Biology 12:495-504).

The remarkable adaptability of this superfamily of ligand-binding proteins is likely to have arisen from positioning the location of binding of the ligand at the interface between the lobes and from the large ligand-mediated conformational change. In this arrangement, ligands are placed within an environment that resembles a protein interior, but the residues forming the contact points or contact sites with the ligand are positioned at the surface of the lobes.

Direct signaling relationships between proteins and reporter groups are readily designed by replacing a residue known to form a ligand contact with a cysteine to which the fluorophore is attached (“endosteric” attachment site). Other, indirect signaling relationships can be established in two ways. The first relies on visual inspection of the ligand complex structure, and identifying residues that are located in the vicinity of the binding site, but do not interact directly with the ligand, and that are likely to be involved in conformational changes. Typically, such “peristeric” sites are located adjacent to the residues that form direct contacts with the bound ligand. In the case of the bPBPs, such residues are located at the perimeter of the inter-domain cleft that forms the ligand binding site location. The environment of these peristeric sites changes significantly upon formation of the closed state. These are examples of positions which are proximal to the ligand-binding pocket/domain. The second, most general, approach identifies sites in the protein structure that are located anywhere in the protein, including locations at some distance away from the ligand-binding site (*i.e.*, distal to the ligand-binding pocket/domain), and undergo a local conformational change in concert with ligand binding. If the structures of both the open and closed states are known, then such “allosteric” sites can be identified using a computational method that

analyzes the conformational changes that accompany ligand binding (Marvin et al., Proc. Natl. Acad. Sci. USA 94:4366-4371, 1997). Alternatively, once allosteric sites have been identified in one bPBP, modeling and structural homology arguments can be invoked to identify such sites in other bPBPs in which only one state has been characterized (Marvin & Hellinga, J. Am. Chem. Soc. 120:7-11, 1998). This generalized conformational analysis also may identify peristeric and endosteric sites, which were identified and classified by visual inspection. The domain or region involved in ligand binding is comprised of a plurality of residues, *e.g.*, non-contiguous amino acids of the ligand-binding protein, which are contact points or sites of contact between the ligand and its cognate ligand-binding protein.

In non-limiting implementations, the reporter group is attached to the ligand-binding protein via a biotin-avidin interaction. The reporter group may be, *e.g.*, conjugated to biotin and the ligand-binding protein is conjugated to avidin. In an example, the avidin is bound to four biotin molecules wherein each biotin molecule is individually conjugated to a reporter group. Alternatively, the reporter group is conjugated to avidin and the ligand-binding protein is conjugated to biotin. For example, the avidin is bound to four biotin molecules, wherein each biotin molecule is individually conjugated to a ligand-binding protein.

As used herein, “conjugated” means covalently attached. One compound may be directly conjugated to another compound, or indirectly conjugated, *e.g.*, via a linker.

In some embodiments, the reporter group is directly attached to the ligand-binding protein. In various embodiments, the reporter group is attached to an amino acid of the ligand-binding protein that is at least about 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 angstroms (Å) from the ligand when the ligand is bound to the ligand-binding protein. In certain embodiments, the reporter group is conjugated to an amino acid having a position within positions 1-25, 25-50, 50-75, 75-100, 100-125, 125-150, 150-175, 175-200, 200-225, 225-250, 250-275, or 275-300 of the ligand-binding protein, wherein position 1 is the N-terminal amino acid of the ligand-binding protein. In non-limiting examples, the reporter group is conjugated to an amino acid of the ligand-binding protein that is (a) within an α -helix or a β -strand of the ligand-binding protein; (b) not within an α -helix; (c) not within a β -strand; (d) within about 5 or 10 amino acids of an amino acid that is within an α -helix or β -strand; (e) within a stretch of consecutive amino acids that links two domains of the ligand-binding protein; (f) within a stretch of consecutive amino acids that links an α -helix and a β -strand; (g) within a stretch of consecutive amino acids that links two α -helices; or (h) within a stretch of consecutive amino

acids that links two β -strands. In some embodiments, the reporter group is directly attached to the N-terminus or the C-terminus of the ligand-binding protein.

The reporter group may be conjugated to the ligand-binding protein a variety of linkers or bonds, including (but not limited to) a disulfide bond, an ester bond, a thioester bond, an amide bond, or a bond that has been formed by a click reaction. In some
5 embodiments, the click reaction is a reaction between (a) an azide and an alkyne; (b) an azide and an alkyne in the presence of Cu(I); (c) an azide and a strained cyclooctyne; (d) an azide and a dibenzylcyclooctyne, a difluorooctyne, or a biarylazacyclooctynone; (e) a diaryl-strained-cyclooctyne and a 1,3-nitrone; (f) an azide, a tetrazine, or a tetrazole and a strained
10 alkene; (g) an azide, a tetrazine, or a tetrazole and a oxanorbornadiene, a cyclooctene, or a trans-cycloalkene; (h) a tetrazole and an alkene; or (i) a tetrazole with an amino or styryl group that is activated by ultraviolet light and an alkene. These exemplary click chemistry reactions have high specificity, efficient kinetics, and occur in vivo under physiological conditions. *See, e.g.*, Baskin et al. Proc. Natl. Acad. Sci. USA 104(2007):16793; Oneto et al.
15 Acta biomaterialia (2014); Neves et al. Bioconjugate chemistry 24(2013):934; Koo et al. Angewandte Chemie 51(2012):11836; Rossin et al. Angewandte Chemie 49(2010):3375, and U.S. Patent Application Publication No. 20160220686, published August 4, 2016, the entire content of each of which is incorporated herein by reference. For a review of a wide variety of click chemistry reactions and their methodologies, *see e.g.*, Nwe K and Brechbiel M W,
20 2009 Cancer Biotherapy and Radiopharmaceuticals, 24(3): 289-302; Kolb H C et al., 2001 Angew. Chem. Int. Ed. 40: 2004-2021. The entire contents of each of the foregoing references are incorporated herein by reference.

As used herein, the term “linker” refers to a molecule or sequence (such as an amino acid sequence), that attaches, as in a bridge, one molecule or sequence to another molecule or
25 sequence. “Linked” means attached or bound by covalent bonds, or non-covalent bonds, or other bonds, such as van der Waals forces. In some embodiments, a linker comprises a chemical structure that has resulted from a reaction used to attach one molecule to another.

In various implementations of the present subject matter, the reporter group is conjugated to a cysteine of the ligand-binding protein. The cysteine may be present on a
30 natural counterpart or version of the ligand-binding protein or added to the ligand-binding protein by a substitution mutation. In some embodiments, the cysteine is at the N-terminus or the C-terminus of the ligand-binding protein.

Non-limiting examples relate to the conjugation of a reporter group to a primary amine of the ligand-binding protein. In certain embodiments, the primary amine is present in a lysine of the ligand-binding protein. The lysine may be present on a natural counterpart or version of the ligand-binding protein or added to the ligand-binding protein by a substitution mutation. In various embodiments, the lysine is at the N-terminus or the C-terminus of the ligand-binding protein.

Aspects of the present subject matter provide a biosensor in which the reporter group is attached to the ligand-binding protein via a linker. In some embodiments, the linker comprises an organic compound that is less than about 30, 20, 15, or 10 Å long. Non-limiting examples of linkers include O, S, NH, PH, and alkyl linkers.

“Alkyl,” as used herein, refers to the radical of saturated or unsaturated aliphatic groups, including straight-chain alkyl, alkenyl, or alkynyl groups, branched-chain alkyl, alkenyl, or alkynyl groups, cycloalkyl, cycloalkenyl, or cycloalkynyl (alicyclic) groups, alkyl substituted cycloalkyl, cycloalkenyl, or cycloalkynyl groups, and cycloalkyl substituted alkyl, alkenyl, or alkynyl groups. Unless otherwise indicated, a straight chain or branched chain alkyl has 30 or fewer carbon atoms in its backbone (*e.g.*, C₁-C₃₀ for straight chain, C₃-C₃₀ for branched chain), more preferably 20 or fewer carbon atoms, more preferably 12 or fewer carbon atoms, and most preferably 8 or fewer carbon atoms. Likewise, preferred cycloalkyls have from 3-10 carbon atoms in their ring structure, and more preferably have 5, 6 or 7 carbons in the ring structure. The ranges provided above are inclusive of all values between the minimum value and the maximum value. The term “alkyl” includes both “unsubstituted alkyls” and “substituted alkyls,” the latter of which refers to alkyl moieties having one or more substituents replacing a hydrogen on one or more carbons of the hydrocarbon backbone. Such substituents include, but are not limited to, halogen, hydroxyl, carbonyl (such as a carboxyl, alkoxycarbonyl, formyl, or an acyl), thiocarbonyl (such as a thioester, a thioacetate, or a thioformate), alkoxyl, phosphoryl, phosphate, phosphonate, a phosphinate, amino, amido, amidine, imine, cyano, nitro, azido, sulfhydryl, alkylthio, sulfate, sulfonate, sulfamoyl, sulfonamido, sulfonyl, heterocyclyl, aralkyl, or an aromatic or heteroaromatic moiety. Unless the number of carbons is otherwise specified, “lower alkyl” as used herein means an alkyl group, as defined above, but having from one to ten carbons, more preferably from one to six carbon atoms in its backbone structure. Likewise, “lower alkenyl” and “lower alkynyl” have similar chain lengths. Preferred alkyl groups are lower alkyls. The alkyl groups may also contain one or more heteroatoms within the carbon backbone. Preferably the

heteroatoms incorporated into the carbon backbone are oxygen, nitrogen, sulfur, and combinations thereof. In certain embodiments, the alkyl group contains between one and four heteroatoms.

5 In some embodiments, the linker comprises a bond formed by a chemical reaction involving a reactive group such as a maleimide group. Alternatively or in addition, the linker comprises a stretch of amino acids. In a non-limiting example, the linker comprises a polyglycine linker. In embodiments, the polyglycine linker comprises 2, 3, 4, 5, or more glycines. Optionally, the polyglycine linker further comprises a serine.

10 In various implementations, the reporter group is attached to a linker via a covalent bond and the linker is attached to a ligand-binding protein via a covalent bond. In embodiments, the covalent bond between the linker and the reporter group and/or the covalent bond between the linker and the ligand-binding protein is a disulfide bond, an ester bond, a thioester bond, an amide bond, a carbamate bond, or a bond that has been formed by a click reaction. Non-limiting examples of click reactions include reactions between an azide
15 and an alkyne; an azide and an alkyne in the presence of Cu(I); an azide and a strained cyclooctyne; an azide and a dibenzylcyclooctyne, a difluorooctyne, or a biarylazacyclooctynone; a diaryl-strained-cyclooctyne and a 1,3-nitrone; an azide, a tetrazine, or a tetrazole and a strained alkene; an azide, a tetrazine, or a tetrazole and a oxanorbornadiene, a cyclooctene, or a trans-cycloalkene; a tetrazole and an alkene; or a
20 tetrazole with an amino or styryl group that is activated by ultraviolet light and an alkene.

The present subject matter also includes biosensors having one or more reporter groups attached to a ligand-binding protein via a fluorophore attachment motif.

Fluorophore Attachment Motifs

25 Aspects of the present subject matter include the use of one or more fluorophore attachment motifs to attach one or more reporter groups to a ligand-binding protein. For example, a reporter group may be attached to a fluorophore attachment motif that is attached to the N-terminus or the C-terminus of the ligand-binding protein.

In various implementations, the fluorophore attachment motif comprises a
30 polypeptide. In some embodiments, the polypeptide comprises amino acids in the β Zif amino acid sequence (SEQ ID NO: 42).

In some embodiments, the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical

to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151). In some embodiments, the polypeptide is a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In certain embodiments, the polypeptide is a mutant of ecTRX comprising a D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

In non-limiting examples, the polypeptide comprises amino acids in the sequence of any one of SEQ ID NOS: 69-86 or 151.

In certain embodiments, the polypeptide comprises (a) at least 1, 2, or 3 thiol groups; (b) at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group; (c) at least 1, 2, or 3 primary amine groups; and/or (d) at least 1, 2, or 3 lysines that each comprise a primary amine. In some embodiments there is no disulfide bond between cysteines within the amino acid sequence of the polypeptide.

In some embodiments, the polypeptide comprises a hexahistidine tag. In some embodiments, the hexahistidine tag is attached to another portion of the polypeptide via a GGS linker.

Reporter Groups

Various types of reporter groups may be used in embodiments of the present subject matter. For example, the reporter group may comprise a fluorophore that produces a fluorescent signal. Biosensors comprising a fluorophore may be referred to herein as fluorescently responsive sensors (FRSs).

Preferably, the binding of ligand to an FRS results in a change in ratiometric ΔR in the signal from a reporter group. A ratiometric signal ($R_{1,2}$) is defined as the quotient of two intensities, $I_{\lambda 1}$ and $I_{\lambda 2}$, measured at two independent wavelengths, λ_1 and λ_2 and may be calculated according to the following equation:

$$R_{1,2} = I_{\lambda 1} / I_{\lambda 2}$$

The two independent wavelengths λ_1 and λ_2 may be from a single fluorophore or from a combination of two or more fluorophores (*e.g.*, a pair of fluorophores between which tgmFRET and/or ngmFRET occurs). In some embodiments, λ_1 falls within the emission

spectrum of a directly responsive fluorophore and λ_2 falls within the emission spectrum of an indirectly responsive fluorophore. In certain embodiments, λ_1 falls within the emission spectrum of an indirectly responsive fluorophore and λ_2 falls within the emission spectrum of a directly responsive fluorophore. In various embodiments, λ_1 falls within the emission spectrum of both a directly responsive fluorophore and an indirectly responsive fluorophore. In various embodiments, λ_2 falls within the emission spectrum of both a directly responsive fluorophore and an indirectly responsive fluorophore.

In some embodiments, intensities are, *e.g.*, integrated, filtered, assessed, detected, or evaluated over a range of wavelengths. In some embodiments, intensities are integrated over a range of wavelengths in a recorded emission spectrum. In some embodiments, a range of wavelengths is selected using a filter. In some embodiments, λ_1 is the intensity over a 1 nm to 60 nm interval centered between 400 and 1000 nm, and λ_2 is the intensity over a 1 nm to 60 nm interval centered between 400 nm and 1000 nm. In some embodiments, intensities are integrated, filtered, assessed, detected, or evaluated over a 1 nm, 2 nm, 10nm, 15nm, 20nm, 25nm, 30nm, 35nm, 40nm, 45nm, 50nm, 55 nm, 60 nm, 75 nm, 100 nm, 10-40 nm, 10-50 nm, 20-50 nm, or 10-100 nm regions, centered between 400-1000 nm, *e.g.* between 420 nm and 520 nm for λ_1 , and 400-1000nm, *e.g.* between 500 nm to 600 nm for λ_2 . In some embodiments, intensities are recorded through a bandpass filter. A non-limiting example of a bandpass filter is a 10 nm, 15 nm, 20 nm, 25 nm, 30 nm, 35 nm, 40 nm, 45 nm, 50 nm, 75 nm, 100 nm, 10-40 nm, 10-50 nm, 20-50 nm, or 10-100 nm bandpass filter, centered between 400-1000 nm, *e.g.* at 452 nm for λ_1 and at 400-1000nm, *e.g.* at 528 nm (λ_2).

Aspects of the present subject matter provide FRSs whose emission spectra change (*e.g.*, the shape of the emission spectra change) in response to ligand binding. In various embodiments, the ratio of intensities at two chosen wavelengths of an FRS's emission spectrum changes upon ligand binding.

In various embodiments, the emission spectra of two or more fluorophores contributes to I_{λ_1} and/or I_{λ_2} . In some embodiments, the emission spectrum of a directly responsive fluorophore contributes to I_{λ_1} and/or I_{λ_2} and the emission spectrum of an indirectly responsive fluorophore contributes to I_{λ_1} and/or I_{λ_2} . In certain embodiments, a directly responsive fluorophore contributes to I_{λ_1} and the emission spectrum of an indirectly responsive fluorophore contributes to I_{λ_2} . In some embodiments, a directly responsive fluorophore contributes to I_{λ_2} and the emission spectrum of an indirectly responsive fluorophore

contributes to $I_{\lambda 1}$. In various embodiments, both the emission spectrum of a directly responsive fluorophore and the emission spectrum of an indirectly responsive fluorophore contributes to $I_{\lambda 1}$. In some embodiments, both the emission spectrum of a directly responsive fluorophore and the emission spectrum of an indirectly responsive fluorophore contributes to

5 $I_{\lambda 2}$.

In some embodiments, the emission wavelength and/or intensity of a fluorophore (*e.g.*, a single fluorophore in a biosensor comprising one reporter group or a directly responsive fluorophore comprising reporter groups between which tgmFRET and/or ngmFRET occurs) changes when the positions of atoms within the fluorophore change with

10 respect to each other (*e.g.*, due to the rotation of bound atoms with respect to each other or a change in the angle of a bond). In non-limiting examples, the emission wavelength and/or intensity of the fluorophore changes when (i) one portion of the fluorophore rotates around a bond axis compared to another portion of the fluorophore and/or (ii) when the angle of a bond between two atoms of the fluorophore changes. In a non-limiting example, the fluorophore is

15 a prodan-derived fluorophore (*e.g.*, Acrylodan or Badan) and binding of ligand alters the orientation of a dimethylamino group, a naphthalene ring, and/or a carbonyl with respect to the ligand-binding protein and/or each other. In a non-limiting example, the degree of polarization of a dipole on the fluorophore changes in response to ligand binding. In various embodiments, the emission wavelength and/or intensity of the fluorophore changes when an

20 atom electrostatically interacts with the fluorophore. For example, the emission wavelength and/or intensity of the fluorophore changes when the source of a positive or negative charge changes its distance with respect to the fluorophore within about 1, 2, 3, 4, 5, or 10 Å of the fluorophore. In some embodiments, the fluorophore exhibits hypsochromicity or bathochromicity upon ligand binding to the ligand-binding domain of the ligand-binding

25 protein. In certain embodiments, the fluorophore has an emission spectrum comprising radiation with a wavelength (*e.g.*, a peak emission wavelength) of about 400 nm, 410 nm, 420 nm, 430 nm, 440 nm, 450 nm, 460 nm, 470 nm, 480 nm, 490 nm, 500 nm, 510 nm, 520 nm, 530 nm, 540 nm, 550 nm, 560 nm, 570 nm, 580 nm, 590 nm, 600 nm, 610 nm, 620 nm, 630 nm, 640 nm, 650 nm, 660 nm, 670 nm, 680 nm, 690 nm, 700 nm, 710 nm, 720 nm, 730 nm,

30 740 nm, 750 nm, 760 nm, 770 nm, 780 nm, 790 nm, 800 nm, 850 nm, 900 nm, 950 nm, or 1000 nm, or about 400 nm to about 450 nm, about 450 nm to about 500 nm, about 500 nm to about 550 nm, about 550 nm to about 600 nm, about 600 nm to about 650 nm, about 650 to

about 700 nm, about 700 nm to about 750 nm, about 750 nm to about 800 nm, or about 800 nm to about 1000 nm.

In some embodiments, the signal comprises the emission intensity of the FRS recorded at a single wavelength or range of wavelengths. The change in signal may be a shift in the single wavelength or range of wavelengths. In some embodiments, the shift in the wavelength is at least about 1 nm, at least about 2 nm, at least about 3 nm, at least about 4 nm, at least about 5 nm, at least about 6 nm, at least about 7 nm, at least about 8 nm, at least about 9 nm, at least about 10 nm, at least about 11 nm, at least about 12 nm, at least about 13 nm, at least about 14 nm, at least about 15 nm, at least about 16 nm, at least about 17 nm, at least about 18 nm, at least about 19 nm, at least about 20 nm, at least about 25 nm, at least about 30 nm, at least about 35 nm, at least about 40 nm, at least about 45 nm, at least about 50 nm, at least about 55 nm, at least about 60 nm, at least about 65 nm, at least about 70 nm, at least about 75 nm, at least about 80 nm, at least about 85 nm, at least about 90 nm, at least about 95 nm, at least about 100 nm, at least about 105 nm, at least about 110 nm, at least about 115 nm, at least about 120 nm, at least about 125 nm, or at least about 130 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 20 nm, about 2 nm to about 20 nm, about 3 nm to about 20 nm, about 4 nm to about 20 nm, about 5 nm to about 20 nm, about 1 nm to about 19 nm, about 1 nm to about 18 nm, about 1 nm to about 17 nm, 1 nm to about 16 nm, about 1 nm to about 15 nm, about 1 nm to about 14 nm, about 1 nm to about 13 nm, about 1 nm to about 12 nm, about 1 nm to about 11 nm, or about 1 nm to about 10 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 20 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 130 nm.

In certain embodiments, the signal comprises the ratio or quotient of the emission intensities recorded at two distinct wavelengths or ranges of wavelengths, *i.e.*, a ratiometric signal. For example, as shown in FIG. 1, ligand binding may be determined by measuring the ratio of blue to green emission intensities. The change in signal may be decreased emission intensity at one wavelength, and no change in emission intensity at the other wavelength. The change in signal may be increased emission intensity at one wavelength, and no change in emission intensity at the other wavelength. The change in signal may be increased emission intensity at one wavelength, and increased emission intensity at the other wavelength. The change in signal may be decreased emission intensity at one wavelength, and decreased emission intensity at the other wavelength. The change in signal may be increased emission intensity at one wavelength, and decreased emission intensity at the other wavelength. In

some embodiments, the change in ratio of the emission intensities recorded at two distinct wavelengths or ranges of wavelengths may be at least about 1.1-fold, at least about 1.2-fold, at least about 1.4-fold, at least about 1.6-fold, at least about 1.8-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3-fold, at least about 3.5-fold, at least about 4-fold, at least about 4.5-fold, at least about 5-fold, at least about 5.5-fold, at least about 6-fold, at least about 6.5-fold, at least about 7-fold, at least about 7.5-fold, at least about 8-fold, at least about 8.5-fold, at least about 9-fold, at least about 9.5-fold, at least about 10-fold, at least about 12-fold, at least about 14-fold, at least about 16-fold, at least about 18-fold, at least about 20-fold, at least about 25-fold, at least about 30-fold, at least about 35-fold, at least about 40-fold, at least about 45-fold, at least about 50-fold, at least about 55-fold, at least about 60-fold, at least about 65-fold, at least about 70-fold, at least about 75-fold, at least about 80-fold, at least about 85-fold, at least about 90-fold, at least about 95-fold, or at least about 100-fold. In some embodiments, the change in ratio of the emission intensities recorded at two distinct wavelengths or ranges of wavelengths may be a decrease of at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, or of 5-25%, 25-50%, 25-75%, 50-75%, 50-90%, or 75-99% or the reciprocal thereof.

The change in signal may be a change in the ratio of the two distinct wavelengths or ranges of wavelengths. The change in signal may be a shift in the two distinct wavelengths or ranges of wavelengths. In some embodiments, one wavelength shifts. In some embodiments, both wavelengths shift. In some embodiments, the shift in the wavelength is at least about 1 nm, at least about 2 nm, at least about 3 nm, at least about 4 nm, at least about 5 nm, at least about 6 nm, at least about 7 nm, at least about 8 nm, at least about 9 nm, at least about 10 nm, at least about 11 nm, at least about 12 nm, at least about 13 nm, at least about 14 nm, at least about 15 nm, at least about 16 nm, at least about 17 nm, at least about 18 nm, at least about 19 nm, at least about 20 nm, at least about 25 nm, at least about 30 nm, at least about 35 nm, at least about 40 nm, at least about 45 nm, at least about 50 nm, at least about 55 nm, at least about 60 nm, at least about 65 nm, at least about 70 nm, at least about 75 nm, at least about 80 nm, at least about 85 nm, at least about 90 nm, at least about 95 nm, at least about 100 nm, at least about 105 nm, at least about 110 nm, at least about 115 nm, at least about 120 nm, at least about 125 nm, or at least about 130 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 20 nm, about 2 nm to about 20 nm, about 3 nm to about 20 nm, about 4 nm to about 20 nm, about 5 nm to about 20 nm, about 1 nm to about 19 nm,

about 1 nm to about 18 nm, about 1 nm to about 17 nm, 1 nm to about 16 nm, about 1 nm to about 15 nm, about 1 nm to about 14 nm, about 1 nm to about 13 nm, about 1 nm to about 12 nm, about 1 nm to about 11 nm, or about 1 nm to about 10 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 20 nm. In some embodiments, the shift in the wavelength is about 1 nm to about 130 nm.

A fluorophore may comprise, *e.g.*, a fluorescent protein or an organic compound having a molecular weight less than about 2000 Daltons (Da). Non-limiting examples of commercially available fluorophores include such as 5-iodoacetamidofluorescein (5-IAF) or 6-iodoacetamidofluorescein (6-IAF), rhodamine, Oregon Green, eosin, Texas Red, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, Badan, Acrylodan, IAEDANS, comprising 3-cyano-7-hydroxycoumarin, 7-hydroxycoumarin-3-carboxylic acid, 6,8-difluoro-7-hydroxy- 4-methylcoumarin, or 7-amino-4-methylcoumarin, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, DRAQ5, DRAQ7, or CyTRAK Orange, cascade blue, Nile red, Nile blue, cresyl violet, oxazine 170, proflavin, acridine orange, acridine yellow, auramine, crystal violet, malachite green, porphyrin, phthalocyanine, bilirubin, pyrene, N,N'-dimethyl-N-(iodoacetyl)-N'-(7-nitrobenz-2-oxo-1,3-diazol-4-yl)ethylenediamide (NBD), N-((2-(iodoacetoxy)ethyl)-N-methyl)-1-amino-7-nitrobenz-2-oxo-1,3-diazole (NBDE), Acrylodan, JPW4039, JPW4042, JPW4045, Oregon Green, Pacific Blue, CPM, N,N'-Dimethyl-N-(Iodoacetyl)-N'-(7-Nitrobenz-2-Oxa-1,3-Diazol-4-yl)Ethylenediamine (IANBD), 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM), BODIPY 499, BODIPY 507/545, BODIPY 499/508, Alexa 432, Alexa488, Alexa532, Alexa546, Cy5, or 1-(2-maleimidylethyl)-4-(5-(4-methoxyphenyl)oxazol-2-yl)pyridinium methanesulfonate (PyMPO maleimide) (PyMPO). In various embodiments, the reporter group was thiol-reactive prior to being conjugated to a polypeptide disclosed herein. In embodiments, the reporter group is linked to a polypeptide disclosed herein via a disulfide bond. Additional non-limiting examples of commercially available fluorophores include fluorescent proteins such as Blue Fluorescent Protein (BFP), TagBFP, mTagBFP2, Azurite, Enhanced Blue Florescent Protein 2 (EBFP2), mKalama1, Sirius, Sapphire, T-Sapphire, Cyan Fluorescent Protein (CFP); Enhanced Cyan Fluorescent Protein (ECFP), Cerulean, SCFP3A, mTurquoise, mTurquoise2, monomeric Midoriishi-Cyan, TagCFP, mTFP1, AmCyan1, Green Fluorescent Protein (GFP), Enhanced Green Fluorescent Protein (EGFP), Emerald, Superfolder GFP, AcGFP1, ZsGreen1, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, Clover, mNeonGreen, Yellow Fluorescent Protein (YFP), Enhanced Yellow Fluorescent Protein

(EYFP), Citrine, Venus, Super Yellow Fluorescent Protein 2 (SYFP2), TagYFP, ZsYellow1, mBanana, Orange Fluorescent Protein (OFP), Monomeric Kusabira-Orange (mKO), mKOκ, mKO2, mOrange, mOrange2, Red Fluorescent Protein (RFP), DsRed-Express, DsRed-Express2, DsRed2, AsRed2, mRaspberry, mCherry, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mRuby2, mPlum, HcRed-Tandem, mKate2, mNeptune, HcRed1, E2-Crimson, NirFP, TagRFP657, IFP1.4, or iRFP.

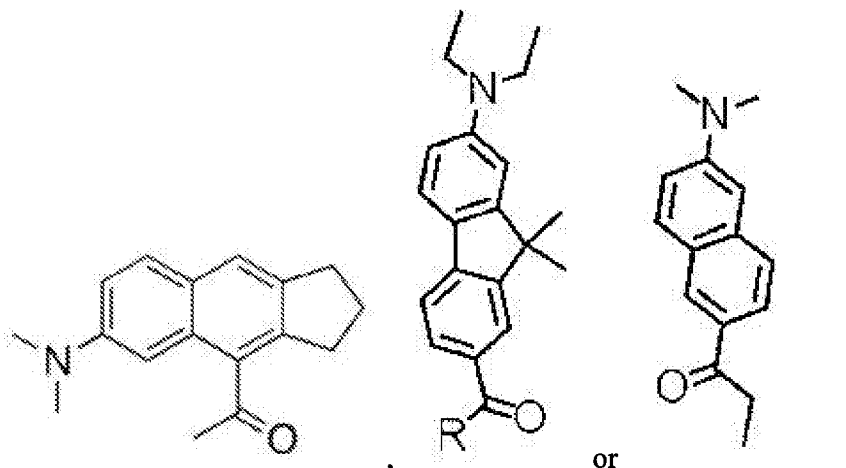
In some embodiments, the fluorophore comprises xanthene, a xanthene derivative, fluorescein, a fluorescein derivative, coumarin, a coumarin derivative, cyanine, a cyanine derivative, rhodamine, a rhodamine derivative, phenoxazine, a phenoxazine derivative, squaraine, a squaraine derivative, coumarin, a coumarin derivative, oxadiazole, an oxadiazole derivative, anthracene, an anthracene derivative, a boradiazaindacene (BODIPY) family fluorophore, pyrene, a pyrene derivative, acridine, an acridine derivative, arylmethine, an arylmethine derivative, tetrapyrrole, or a tetrapyrrole derivative. Non-limiting aspects of fluorophores are discussed in Lavis and Raines (2014) ACS Chem. Biol. 9, 855–866, the entire content of which is incorporated herein by reference. For example, the fluorophore may comprise a xanthene derivative comprising fluorescein or a fluorescein derivative, rhodamine, Oregon Green, eosin, or Texas Red. Non-limiting examples of fluorescein derivatives include 5-fluorescein, 6-carboxyfluorescein, 3'6-carboxyfluorescein, 5(6)-carboxyfluorescein, 6-hexachlorofluorescein, 6-tetrachlorofluorescein, or isothiocyanate. In some embodiments, the fluorophore comprises a cyanine derivative comprising indocarbocyanine, oxacarbocyanine, thiocarbocyanine, or merocyanine. In certain embodiments, the fluorophore comprises a squaraine derivative comprising a ring-substituted squaraine. In various embodiments, the fluorophore comprises a naphthalene derivative comprising a dansyl or prodan naphthalene derivative. In a non-limiting example, the fluorophore comprises prodan or a derivative thereof. In certain embodiments, the fluorophore comprises Badan, Acrylodan, or N-(Iodoacetaminoethyl)-1-naphthylamine-5-sulfonic acid (IAEDANS). In some embodiments, the fluorophore comprises a coumarin derivative such as 3-cyano-7-hydroxycoumarin, 7-hydroxycoumarin-3-carboxylic acid, 6,8-difluoro-7-hydroxy-4-methylcoumarin (DiFMU), or 7-amino-4-methylcoumarin. In various embodiments, the fluorophore comprises an oxadiazole derivative such as pyridyloxazole, nitrobenzoxadiazole, or benzoxadiazole. In certain embodiments, the fluorophore comprises an anthracene derivative comprising an anthraquinone such as DRAQ5, DRAQ7, or CyTRAK Orange. In various embodiments, the fluorophore comprises a pyrene derivative

comprising cascade blue. In non-limiting examples the fluorophore comprises an oxazine derivative such as Nile red, Nile blue, cresyl violet, or oxazine 170. In some embodiments, the fluorophore comprises an acridine derivative such as proflavin, acridine orange, or acridine yellow. In certain embodiments, the fluorophore comprises an arylmethine derivative such as auramine, crystal violet, or malachite green. In various embodiments, the fluorophore comprises a tetrapyrrole derivative comprising porphyrin, phthalocyanine, or bilirubin.

Aspects of the present subject matter relate to the use of fluorophores that may readily be attached to a ligand-binding protein disclosed herein, *e.g.*, at a cysteine residue. For example, a fluorophore may comprise a sulfhydryl group prior to attachment to a ligand-binding protein that is reacted with a moiety of the ligand-binding protein to attach the fluorophore to the ligand-binding protein. In some embodiments, the fluorophore comprised a thiol group prior to attachment to the ligand-binding protein. For example, the fluorophore was thiol reactive prior to attachment to the ligand-binding protein. Non-limiting examples of fluorophores that may readily be attached to ligand-binding proteins using thiol reactions include fluorescein, pyrene, NBD, NBDE, Acrylodan (6-acryloyl 1-2-dimethylaminonaphthalene), Badan (6-bromo-acetyl-2-dimethylamino-naphthalene), JPW4039, JPW4042, or JPW4045.

In certain embodiments, the fluorophore comprises a derivative of a Prodan-based fluorophore such as Acrylodan or Badan. The excitation and emission properties of the Prodan-based fluorophores Acrylodan and Badan can be altered by manipulating the fluorescent ring system, while preserving the dimethylamino donor group, and the twistable carbonyl acceptor (Klymchenko 2013 *Progress in Molecular Biology and Translational Science*, 35-58). Replacement of the two-ring naphthalene with a three-ring anthracene (Lu 2006 *J. Org. Chem.*, 71, 9651-9657), fluorene (Kucharak 2010 *J. Phys. Chem. Lett.*, 1, 616-620), pyrene (Niko 2013 *Chem. Eur. J.*, 19, 9760-9765), or styrene (Benedetti 2012 *J. Am. Chem. Soc.*, 134, 12418-12421) cores significantly red-shift the excitation and emission properties, and in the case of the latter two, improve brightness through improvements in their excitation peak extinction coefficients. The entire content of each of the references cited above (as well as all other references referred to herein including the contents of nucleic acid and amino acid sequence accession number references) are incorporated herein by reference. Non-limiting examples of prodan analogues include 2-cyano-6-

dihexylaminoanthracene and 2-propionyl-6-dihexylaminoanthracene, as well as fluorophores comprising the following structures:



In some embodiments, the fluorophore comprises a fluorescent protein. Fluorescent proteins that emit blue, cyan, green, yellow, orange, red, far-red, or near infrared radiation when contacted with excitation radiation are known in the art and commercially available as proteins and via the expression of vectors that encode the fluorescent protein. Non-limiting examples of fluorescent proteins include Blue Fluorescent Protein (BFP), TagBFP, mTagBFP2, Azurite, Enhanced Blue Florescent Protein 2 (EBFP2), mKalamal, Sirius, Sapphire, T-Sapphire, Cyan Fluorescent Protein (CFP); Enhanced Cyan Fluorescent Protein (ECFP), Cerulean, SCFP3A, mTurquoise, mTurquoise2, monomeric Midoriishi-Cyan, TagCFP, mTFP1, AmCyan1, Green Fluorescent Protein (GFP), Enhanced Green Fluorescent Protein (EGFP), Emerald, Superfolder GFP, AcGFP1, ZsGreen1, Monomeric Azami Green, TagGFP2, mUKG, mWasabi, Clover, mNeonGreen, Yellow Fluorescent Protein (YFP), Enhanced Yellow Fluorescent Protein (EYFP), Citrine, Venus, Super Yellow Fluorescent Protein 2 (SYFP2), TagYFP, ZsYellow1, mBanana, Orange Fluorescent Protein (OFP), Monomeric Kusabira-Orange (mKO), mKO_κ, mKO2, mOrange, mOrange2, Red Fluorescent Protein (RFP), DsRed-Express, DsRed-Express2, DsRed2, AsRed2, mRaspberry, mCherry, mStrawberry, mTangerine, tdTomato, TagRFP, TagRFP-T, mApple, mRuby, mRuby2, mPlum, HcRed-Tandem, mKate2, mNeptune, HcRed1, E2-Crimson, NirFP, TagRFP657, IFP1.4, or iRFP.

In some embodiments, the fluorophore comprises a quantum dot (Medintz et al. 2005) (Sapsford, Berti and Medintz 2006 *Angew Chem Int Ed Engl*, 45, 4562-89; Resch-Genger et al. 2008 *Nat Methods*, 5, 763-75). In some embodiments the emission properties of the

conjugated protein are enhanced by immobilization on or near metallic nanoparticles (Zeng et al. 2014 *Chem Soc Rev*, 43, 3426-52; Shen et al. 2015 *Nanoscale*, 7, 20132-41).

In various embodiments, the peak emission wavelength and/or the emission intensity of the biosensor change when the ligand binds to the ligand-binding protein. In some
5 embodiments, the biosensor exhibits a dichromatic signaling change when the ligand binds to the ligand-binding protein. In various embodiments, the peak emission wavelength of the biosensor shifts by at least about 5, 10, 15, 20, 30, 40, 50, or by about 5-50 nm when the biosensor binds to ligand. In certain embodiments, the emission intensity of the biosensor increases by at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%,
10 150%, 200%, or 300% when the biosensor binds to ligand. In various embodiments, the signal produced by the reporter group persists for at least 1 nanoseconds (ns), 5 ns, 10 ns, 25 ns, 50 ns, 75 ns, 100 ns, 200 ns, 300 ns, 400 ns, 500 ns, 600 ns, 700 ns, 800 ns, 900 ns, 0.001 milliseconds (ms), 0.01 ms, 0.1 ms, 1 ms, 5 ms, 10 ms, 20 ms, 25 ms, 50 ms, 100 ms, or 500 ms when the ligand binds to the ligand-binding protein.

Ligand-Binding Proteins

Aspects of the present subject matter provide biosensors comprising a ligand-binding protein that binds a ligand of interest. Non-limiting examples of ligands include sugars (such as glucose, galactose, lactose, arabinose, ribose, and maltose), lactate, urea, anions (*e.g.*,
20 bicarbonate, phosphate, sulfate, and halide anions such as chloride, fluoride, iodide, astatide, ununseptide, and bromide), cations (*e.g.*, calcium, iron, and hydrogen ions), dipeptides, and amino acids (such as histidine, glutamine, glutamate, aspartate).

The ligand-binding protein may comprise a naturally occurring protein or a protein that is modified compared to a naturally occurring protein. For example, the ligand-binding
25 protein may comprise one or more mutations compared to a naturally occurring protein. In some embodiments, the naturally occurring protein is a naturally occurring counterpart of the ligand-binding protein (*e.g.*, the ligand-binding protein is a mutant of the naturally occurring counterpart).

A “naturally occurring counterpart” of a mutant polypeptide is a polypeptide
30 produced in nature from which the mutant polypeptide has been or may be derived (*e.g.*, by one or more mutations). For example, the naturally occurring counterpart is an endogenous polypeptide produced by an organism in nature, wherein the endogenous polypeptide typically does not have one or more of the mutations present in the mutant polypeptide. For

convenience and depending on context, a naturally occurring counterpart may be referred to herein for the purpose of comparison and to illustrate the location and/or presence of one or more mutations, binding activities, and/or structural features.

As used herein, a “mutation” is a difference between the amino acid sequence of a modified polypeptide/protein and a naturally occurring counterpart. A polypeptide having a mutation may be referred to as a “mutant.” Non-limiting examples of mutations include insertions, deletions, and substitutions. However, the term “mutation” excludes (i) the addition of amino acids to the N-terminus or C-terminus of a polypeptide, and (ii) the omission/deletion/replacement of a polypeptide’s signal peptide (*e.g.*, replacement with another signal peptide or with a methionine).

The addition of amino acids to the N-terminus or C-terminus of a protein via a peptide bond may be referred to herein as a “fusion” of the amino acids to the protein. Similarly, an exogenous protein fused to amino acids (*e.g.*, another protein, a fragment, a tag, or a polypeptide moiety) at its N-terminus or C-terminus may be referred to as a “fusion protein.” The added amino acids may comprise a non-native polypeptide, *e.g.*, a polypeptide reporter group such as a fluorescent protein, a moiety that facilitates the isolation or modification of a polypeptide, or a moiety that facilitates the attachment of a polypeptide to a substrate or surface. As used herein, “non-native” when referring to the added amino acids (*e.g.*, a “polypeptide”) of a fusion protein indicates that the polypeptide is not naturally part of the protein to which it is fused in the fusion protein. For example, the sequence of a non-native polypeptide (“added amino acids”) that is fused to a protein is encoded by an organism other than the organism from which the protein is derived, is not known to be naturally encoded by any organism, or is encoded by a gene other than the wild-type gene that encodes an endogenous version of the protein.

As used herein the term “signal peptide” refers to a short (*e.g.*, 5-30 or 10-100 amino acids long) stretch of amino acids at the N-terminus of a protein that directs the transport of the protein. In various embodiments, the signal peptide is cleaved off during the post-translational modification of a protein by a cell. Signal peptides may also be referred to as “targeting signals,” “leader sequences,” “signal sequences,” “transit peptides,” or “localization signals.” In instances where a signal peptide is not defined for a ligand-binding protein discussed herein, the signal peptide may optionally be considered to be, *e.g.*, the first 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 5-15, 5-20, 5-25, 5-100, 10-15, 10-20, 10-25, 10-50, 10-100, 25-

50, 25-75, or 25-100 amino acids from the N-terminus of the translated protein (compared to a protein that has not had the signal peptide removed, *e.g.*, compared to a naturally occurring protein).

In some embodiments, the ligand-binding protein comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 1-10, 1-15, 1-20, 5-15, 5-20, 10-25, 10-50, 20-50, 25-75, 25-100 or more mutations compared to a naturally occurring protein while retaining at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5%, or about 100% of the activity of the naturally occurring protein. Mutations include but are not limited to substitutions, insertions, and deletions. Non-limiting examples of ligand-binding proteins may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 1-10, 1-15, 1-20, 5-15, 5-20, 10-25, 10-50, 20-50, 25-75, 25-100, or more substitution mutations compared to a naturally occurring protein while retaining at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5%, or about 100% of the activity of the naturally occurring protein. In embodiments, at least one amino acid of the ligand-binding protein has been substituted with a cysteine. Alternatively or in addition, a ligand-binding protein may include one or more mutations that remove a cysteine, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more substitutions or deletions of a cysteine compared to a naturally occurring protein.

Alternatively, the ligand-binding protein is not a mutant. For example, a reporter group is fused to the N-terminus or the C-terminus of the ligand-binding protein.

In various embodiments, a ligand-binding protein may comprise a stretch of amino acids (*e.g.*, the entire length of the ligand-binding protein or a portion comprising at least about 50, 100, 200, 250, 300, or 350 amino acids) in a sequence that is at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, or 99.5% identical to an amino acid sequence of a naturally occurring protein.

In some embodiments, the mutations are conservative, and the present subject matter includes many ligand-binding proteins in which the only mutations are substitution mutations. In non-limiting examples, a ligand-binding protein has no deletions or insertions compared to a naturally occurring protein (*e.g.*, a naturally occurring counterpart). Alternatively, a ligand-binding protein may have (i) less than about 5, 4, 3, 2, or 1 inserted amino acids, and/or (ii) less than about 5, 4, 3, 2, or 1 deleted amino acids compared to a naturally occurring protein.

In various embodiments, a naturally occurring protein to which a ligand-binding protein is compared or has been derived (*e.g.*, by mutation, fusion, or other modification) from a prokaryotic ligand-binding protein such as a bacterial ligand-binding protein. For example, the prokaryotic ligand-binding protein is a mutant, fragment, or variant of a natural (i.e., wild-type) bacterial protein. In various embodiments, the bacterial ligand-binding protein is from a thermophilic, mesophilic, or cryophilic prokaryotic microorganism (*e.g.*, a thermophilic, mesophilic, or cryophilic bacterium).

A microorganism is “thermophilic” if it is capable of surviving, growing, and reproducing at temperatures between 41 and 140 °C (106 and 284 °F), inclusive. In various embodiments, a thermophilic organism has an optimal growth temperature between 41 and 140°C, or that is at least about 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, or 140°C. Many thermophiles are archaea. Thermophilic eubacteria are suggested to have been among the earliest bacteria. Thermophiles are found in various geothermally heated regions of the Earth, such as hot springs and deep sea hydrothermal vents, as well as decaying plant matter, such as peat bogs and compost. Unlike other types of microorganisms, thermophiles can survive at much hotter temperatures, whereas other bacteria would be damaged and sometimes killed if exposed to the same temperatures. Thermophiles may be classified into three groups: (1) obligate thermophiles; (2) facultative thermophiles; and (3) hyperthermophiles. Obligate thermophiles (also called extreme thermophiles) require such high temperatures for growth, whereas facultative thermophiles (also called moderate thermophiles) can thrive at high temperatures, but also at lower temperatures (*e.g.* below 50°C). Hyperthermophiles are particularly extreme thermophiles for which the optimal temperatures are above 80°C. Some microorganisms can live at temperatures higher than 100°C at large depths in the ocean where water does not boil because of high pressure. Many hyperthermophiles are also able to withstand other environmental extremes such as high acidity or radiation levels. A compound (*e.g.*, a protein or biosensor) is “thermotolerant” if it is capable of surviving exposure to temperatures above 41°C. For example, in some embodiments a thermotolerant biosensor retains its function and does not become denatured when exposed to a temperature of about 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, or 140°C for at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 or more minutes. In some embodiments, the thermotolerant compound survives exposure to 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, or 140°C under pressure.

A microorganism is “mesophilic” if it is capable of surviving, growing, and reproducing at temperatures between 20 and 40°C (68 and 104°F), inclusive. “Psychrophiles” or “cryophiles” are microorganisms that are capable of growth and reproduction in cold temperatures. In various embodiments, a psychrophile is capable of
5 growth and reproduction at a temperature of 10°C or less, *e.g.*, between –20 °C and +10 °C.

In some embodiments, the microbial protein is produced by a bacterial microorganism, an archaean microorganism, an algal microorganism, a protozoan microorganism, or a fungal microorganism. In non-limiting examples, the microbial protein is produced by a Gram-positive bacterium or a Gram-negative bacterium. In various
10 embodiments, a biosensor comprises a modified (*e.g.*, mutated, fused, and/or conjugated) periplasmic binding protein or a cytoplasmic binding protein.

Aspects of the present subject matter provide a ligand-binding protein with a mutation that alters the interaction of the ligand-binding protein with a ligand. For example, the ligand-binding protein comprises a mutation that alters the interaction of the ligand-binding
15 protein with the ligand compared to a naturally occurring counterpart. In some embodiments, the ligand-binding protein comprises a mutation that alters the interaction of an amino acid of the ligand-binding protein with a water molecule compared to a naturally occurring counterpart.

In some embodiments, the ligand-binding protein does not comprise a signal peptide. For example, the signal peptide (*e.g.*, that is present in a naturally occurring counterpart) may
20 be replaced with a methionine.

Exemplary implementations relate to a ligand such as sugars (such as glucose, galactose, lactose, arabinose, ribose, and maltose), lactate, urea, anions (*e.g.*, chloride, bicarbonate, phosphate, and sulfate), cations (*e.g.*, calcium and iron), dipeptides, amino acids
25 (such as histidine, glutamine, glutamate, and aspartate). For example, the biosensor may comprise a mutant of, a fragment of, or a fusion protein comprising a microbial ligand-binding protein. In embodiments, the ligand-binding protein is not a mutant or fragment to which a non-native polypeptide has been attached or added.

The ratiometric reagentless biosensors produce precise measurements over extended
30 concentration ranges, *e.g.* from 0.0001 mM to 100 mM, in sample volumes of less than about 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 µl. In some embodiments, the ligand-binding protein comprises a mutation that alters (*e.g.*, increases or decreases) the interaction of the mutant with bound ligand compared to a naturally occurring protein (*e.g.*, a microbial ligand-binding protein). In

non-limiting examples, the ligand-binding protein comprises a mutation that alters (*e.g.*, increases or decreases) the mutant's affinity and/or specificity for ligand compared to an unmutated ligand-binding protein (*e.g.*, a microbial ligand-binding protein). In certain embodiments, the ligand-binding protein comprises a mutation that alters the interaction
 5 between the protein and bound ligand, a mutation that alters the equilibrium between the open and closed states of the ligand-binding protein, a mutation that alters the interaction between the ligand-binding protein and a reporter group (such as a fluorescent conjugate, *e.g.*, the interaction with a carbonyl group or a naphthalene ring of a prodan-derived fluorophore such as Acrylodan or Badan), and/or a mutation that impacts indirect interactions
 10 that alter the geometry of the ligand binding site. In various embodiments, the mutation does not reduce, or negligibly impacts, the thermostability of the ligand-binding protein. In some embodiments, the mutation alters the thermostability of the ligand-binding protein by less than about 1, 2, 3, 4, 5, or 10°C.

The present subject matter provides a glucose-galactose binding protein GGBP that is
 15 or is a mutant of: an *Escherichia* sp. (*e.g.*, *E. albertii*, *E. coli*, *E. fergusonii*, *E. hermannii*, or *E. vulneris*) GGBP; a *Thermoanaerobacter* sp. (*e.g.*, *T. acetoethylicus*, *T. brockii*, *T. ethanolicus*, *T. italicus*, *T. kivui*, *T. mathranii*, *T. pseudethanolicus*, *T. siderophilus*, *T. sulfurigignens*, *T. sulfurophilus*, *T. thermocopriae*, *T. thermohydrosulfuricus*, *T. thermosaccharolyticum*, *T. uzonensis*, or *T. wiegelii*) GGBP; a *Clostridium* sp. (*e.g.*, *C.*
 20 *absonum*, *C. aceticum*, *C. acetireducens*, *C. acetobutylicum*, *C. acidisoli*, *C. aciditolerans*, *C. acidurici*, *C. aerotolerans*, *C. aestuarii*, *C. akagii*, *C. aldenense*, *C. aldrichii*, *C. algidicarni*, *C. algidixylanolyticum*, *C. algifaecis*, *C. algoriphilum*, *C. alkalicellulosi*, *C. aminophilum*, *C. aminovalericum*, *C. amygdalinum*, *C. amylolyticum*, *C. arbusti*, *C. arcticum*, *C. argentinense*, *C. asparagiforme*, *C. aurantibutyricum*, *C. autoethanogenum*, *C. baratii*, *C. barkeri*, *C.*
 25 *bartlettii*, *C. beijerinckii*, *C. bifermentans*, *C. bolteae*, *C. bornimense*, *C. botulinum*, *C. bowmanii*, *C. bryantii*, *C. butyricum*, *C. cadaveris*, *C. caenicola*, *C. caminithermale*, *C. carboxidivorans*, *C. carnis*, *C. cavendishii*, *C. celatum*, *C. celerecrescens*, *C. cellobioparum*, *C. cellulofermentans*, *C. cellulolyticum*, *C. cellulosi*, *C. cellulovorans*, *C. chartatabidum*, *C. chauvoei*, *C. chromiireducens*, *C. citroniae*, *C. clariflavum*, *C. clostridioforme*, *C. coccoides*,
 30 *C. cochlearium*, *C. colletant*, *C. colicanis*, *C. colinum*, *C. collagenovorans*, *C. cylindrosporum*, *C. difficile*, *C. diolis*, *C. disporicum*, *C. drakei*, *C. durum*, *C. estertheticum*, *C. estertheticum estertheticum*, *C. estertheticum laramiense*, *C. fallax*, *C. felsineum*, *C. fervidum*, *C. fimetarium*, *C. formicaceticum*, *C. frigidicarnis*, *C. frigoris*, *C. ganghwense*, *C.*

- gasigenes*, *C. ghonii*, *C. glycolicum*, *C. glycyrrhizinilyticum*, *C. grantii*, *C. haemolyticum*, *C. halophilum*, *C. hastiforme*, *C. hathewayi*, *C. herbivorans*, *C. hiranonis*, *C. histolyticum*, *C. homopropionicum*, *C. huakuii*, *C. hungatei*, *C. hydrogeniformans*, *C. hydroxybenzoicum*, *C. hylemonae*, *C. jejuense*, *C. indolis*, *C. innocuum*, *C. intestinale*, *C. irregulare*, *C. isatidis*, *C. josui*, *C. kluyveri*, *C. lactatifermentans*, *C. lacusfryxellense*, *C. laramiense*, *C. lavalense*, *C. lentocellum*, *C. lentoputrescens*, *C. leptum*, *C. limosum*, *C. litorale*, *C. lituseburense*, *C. ljungdahlii*, *C. lortetii*, *C. lundense*, *C. magnum*, *C. malenominatum*, *C. manganotii*, *C. mayombeii*, *C. methoxybenzovorans*, *C. methylpentosum*, *C. neopropionicum*, *C. nexile*, *C. nitrophenolicum*, *C. novyi*, *C. oceanicum*, *C. orbiscindens*, *C. oroticum*, *C. oxalicum*, *C. papyrosolvans*, *C. paradoxum*, *C. paraperfringens*, *C. paraputrificum*, *C. pascui*, *C. pasteurianum*, *C. peptidivorans*, *C. perenne*, *C. perfringens*, *C. pfennigii*, *C. phytofermentans*, *C. piliforme*, *C. polysaccharolyticum*, *C. populeti*, *C. propionicum*, *C. proteoclasticum*, *C. proteolyticum*, *C. psychrophilum*, *C. puniceum*, *C. purinilyticum*, *C. putrefaciens*, *C. putrificum*, *C. quercicolum*, *C. quinii*, *C. ramosum*, *C. rectum*, *C. roseum*, *C. saccharobutylicum*, *C. saccharogumia*, *C. saccharolyticum*, *C. saccharoperbutylacetonicum*, *C. sardiniense*, *C. sartagoforme*, *C. scatologenes*, *C. schirmacherense*, *C. scindens*, *C. septicum*, *C. sordellii*, *C. sphenoides*, *C. spiroforme*, *C. sporogenes*, *C. sporosphaeroides*, *C. stercorarium*, *C. stercorarium leptospartum*, *C. stercorarium stercorarium*, *C. stercorarium thermolacticum*, *C. sticklandii*, *C. straminisolvans*, *C. subterminale*, *C. sufflavum*, *C. sulfidigenes*, *C. symbiosum*, *C. tagluense*, *C. tepidiprofundii*, *C. termitidis*, *C. tertium*, *C. tetani*, *Clostridium tetanomorphum*, *C. thermaceticum*, *C. thermautotrophicum*, *C. thermoalcaliphilum*, *C. thermobutyricum*, *C. thermocellum*, *C. thermocopriae*, *C. thermohydrosulfuricum*, *C. thermolacticum*, *C. thermopalmarium*, *C. thermopapyrolyticum*, *C. thermosaccharolyticum*, *C. thermosuccinogenes*, *C. thermosulfurigenes*, *C. thiosulfatireducens*, *C. tyrobutyricum*, *C. uliginosum*, *C. ultunense*, *C. villosum*, *C. vincentii*, *C. viride*, *C. xylanolyticum*, or *C. xylanovorans*) GGBP; a *Salmonella* sp. [e.g., *S. bongori*, *S. enterica*, *S. enterica* subspecies *enterica*, *S. enterica* subspecies *salamae*, *S. enterica* subspecies *arizonae*, *S. enterica* subspecies *diarizonae*, *S. enterica* subspecies *houtenae*, *S. enterica* subspecies *indica*, or *S. enterica* subspecies *enterica* serovar *Typhimurium* (*S. typhimurium*)] GGBP; a *Caldicellulosiruptor* sp. (e.g., *C. saccharolyticus*, *C. acetigenus*, *C. bescii*, *C. changbaiensis*, *C. hydrothermalis*, *Caldicellulosiruptor hydrother*, *C. kristjanssonii*, *C. kronotskyensis*, *C. lactoaceticus*, *C. owensensis*, or *C. obsidiansis*) GGBP; a *Paenibacillus* sp. (e.g., *P. agarexedens*, *P. agaridevorans*, *P. alginolyticus*, *P. alkaliterrae*, *P.*

- alvei*, *P. amylolyticus*, *P. anaericanus*, *P. antarcticus*, *P. assamensis*, *P. azoreducens*, *P. azotofixans*, *P. barcinonensis*, *P. borealis*, *P. brasiliensis*, *P. brassicae*, *P. campinasensis*, *P. chinjuensis*, *P. chitinolyticus*, *P. chondroitinus*, *P. cineris*, *P. cookii*, *P. curdlanolyticus*, *P. daejeonensis*, *P. dendritiformis*, *P. durum*, *P. ehimensis*, *P. elgii*, *P. favisporus*, *P.*
- 5 *glucanolyticus*, *P. glycanilyticus*, *P. gordonae*, *P. graminis*, *P. granivorans*, *P. hodogayensis*, *P. illinoisensis*, *P. jamilae*, *P. kobensis*, *P. koleovorans*, *P. koreensis*, *P. kribbensis*, *P. lactis*, *P. larvae*, *P. lautus*, *P. lentimorbus*, *P. macerans*, *P. macquariensis*, *P. massiliensis*, *P. mendelii*, *P. motobuensis*, *P. naphthalenovorans*, *P. nematophilus*, *P. odorifer*, *P. pabuli*, *P. peoriae*, *P. phoenicis*, *P. phyllosphaerae*, *P. polymyxa*, *P. popilliae*, *P. pulvificiens*, *P.*
- 10 *rhizosphaerae*, *P. sanguinis*, *P. stellifer*, *P. terrae*, *P. thiaminolyticus*, *P. timonensis*, *P. tylopili*, *P. turicensis*, *P. validus*, *P. vortex*, *P. vulneris*, *P. wynnii*, *P. xylanilyticus*) GGBP; a *Butyrivibrio* sp. (e.g., *B. proteoclasticus*, *B. crossotus*, *B. fibrisolvens*, or *B. hungatei*) GGBP; a *Roseburia* sp. (e.g., *R. intestinalis*, *R. faecis*, *R. hominis*, or *R. inulinivorans*) GGBP; a *Faecalibacterium* sp. (e.g., *F. prausnitzii*) GGBP; an *Erysipelothrix* sp. (e.g., *E.*
- 15 *rhusiopathiae*, *E. inopinata*, or *E. tonsillarum*) GGBP; or an *Eubacterium* sp. (e.g., *E. rectale*, *E. acidaminophilum*, *E. nodatum*, *E. oxidoreducens*, or *E. foedans*) GGBP.

- The present subject matter provides a urea-binding protein that is or is a mutant of: an *Marinomas* sp. (e.g., *M. posidonica*) urea-binding protein; a *Marinobacter* sp. (e.g., *M.*
- 20 *adhaerens*, *M. algicola*, *M. alkaliphilus*, *M. antarcticus*, *M. arcticus*, *M. aromaticivorans*, *M. bryozoorum*, *M. daepoensis*, *M. daqiaonensis*, *M. excellens*, *M. flavimaris*, *M. gudaonensis*, *M. guineae*, *M. halophilus*, *M. gudaonensis*, *M. hydrocarbonoclasticus*, *M. koreensis*, *M. laciisalsi*, *M. lipolyticus*, *M. litoralis*, *M. lutaoensis*, *M. maritimus*, *M. mobilis*, *M. nitratreducens*, *M. oulmenensis*, *M. pelagius*, *M. persicus*, *M. psychrophilus*, *M. nanhaiticus*, *M. salarius*, *M. salicampi*, *M. salsuginis*, *M. santoriniensis*, *M. sediminum*, *M. segnicrescens*,
- 25 *M. shengliensis*, *M. squalenivorans*, *M. similis*, *M. szutsaonensis*, *M. vinifirmus*, *M. xestospongiae*, *M. zhanjiangensis*, or *M. zhejiangensis*) urea-binding protein; a *Bacillus* sp. (e.g., *B. acidiceler*, *B. acidicola*, *B. acidiproducens*, *B. acidocaldarius*, *B. acidoterrestris*, *B. aeolius*, *B. aerius*, *B. aerophilus*, *B. agaradhaerens*, *B. agri*, *B. ailingensis*, *B. akibai*, *B. alcalophilus*, *B. algicola*, *B. alginolyticus*, *B. alkalidiazotrophicus*, *B. alkalinitrilicus*, *B.*
- 30 *alkalisediminis*, *B. alkalitelluris*, *B. altitudinis*, *B. alveayuensis*, *B. alvei*, *B. amyloliquefaciens*, *B. a. subsp. amyloliquefaciens*, *B. a. subsp. plantarum*, *B. amylolyticus*, *B. andreesenii*, *B. aneurinilyticus*, *B. anthracis*, *B. aquimaris*, *B. arenosi*, *B. arseniciselenatis*, *B. arsenicus*, *B. aurantiacus*, *B. arvi*, *B. aryabhattai*, *B. asahii*, *B. atrophaeus*, *B.*

- axarquiensis*, *B. azotofixans*, *B. azotoformans*, *B. badius*, *B. barbaricus*, *B. bataviensis*, *B. beijingensis*, *B. benzoovorans*, *B. beringensis*, *B. berkeleyi*, *B. beveridgei*, *B. bogoriensis*, *B. boroniphilus*, *B. borstelensis*, *B. brevis* Migula, *B. butanolivorans*, *B. canaveralius*, *B. carboniphilus*, *B. cecembensis*, *B. cellulossilyticus*, *B. centrosporus*, *B. cereus*, *B.*
- 5 *chagannorensis*, *B. chitinolyticus*, *B. chondroitinus*, *B. choshinensis*, *B. chungangensis*, *B. cibi*, *B. circulans*, *B. clarkii*, *B. clausii*, *B. coagulans*, *B. coahuilensis*, *B. cohnii*, *B. composti*, *B. curdlanolyticus*, *B. cycloheptanicus*, *B. cytotoxicus*, *B. daliensis*, *B. decisifrondis*, *B. decolorationis*, *B. deserti*, *B. dipsosauri*, *B. drementensis*, *B. edaphicus*, *B. ehimensis*, *B. eiseniae*, *B. enclensis*, *B. endophyticus*, *B. endoradicis*, *B. farraginis*, *B. fastidiosus*, *B.*
- 10 *fengqiensis*, *B. firmus*, *B. flexus*, *B. foraminis*, *B. fordii*, *B. formosus*, *B. fortis*, *B. fumarioli*, *B. funiculus*, *B. fusiformis*, *B. galactophilus*, *B. galactosidilyticus*, *B. galliciensis*, *B. gelatini*, *B. gibsonii*, *B. ginsengi*, *B. ginsengihumi*, *B. ginsengisoli*, *B. globisporus*, *B. g. subsp.* *globisporus*, *B. g. subsp. marinus*, *B. glucanolyticus*, *B. gordonae*, *B. gottheilii*, *B. graminis*, *B. halmaphilus*, *B. haloalkaliphilus*, *B. halocharis*, *B. halodenitrificans*, *B. halodurans*, *B.*
- 15 *halophilus*, *B. halosaccharovorans*, *B. hemicellulosilyticus*, *B. hemicentroti*, *B. herbersteinensis*, *B. horikoshii*, *B. horneckiae*, *B. horti*, *B. huizhouensis*, *B. humi*, *B. hwajinpoensis*, *B. idriensis*, *B. indicus*, *B. infantis*, *B. infernus*, *B. insolitus*, *B. invictae*, *B. iranensis*, *B. isabeliae*, *B. isronensis*, *B. jeotgali*, *B. kaustophilus*, *B. kobensis*, *B. kochii*, *B. kokeshiiformis*, *B. koreensis*, *B. korlensis*, *B. kribbensis*, *B. krulwichiae*, *B. laevolacticus*, *B.*
- 20 *larvae*, *B. laterosporus*, *B. lautus*, *B. lehensis*, *B. lentimorbus*, *B. lentus*, *B. licheniformis*, *B. ligniniphilus*, *B. litoralis*, *B. locisalis*, *B. luciferensis*, *B. luteolus*, *B. luteus*, *B. macauensis*, *B. macerans*, *B. macquariensis*, *B. macyae*, *B. malacitensis*, *B. mannaniolyticus*, *B. marisflavi*, *B. marismortui*, *B. marmarensis*, *B. massiliensis*, *B. megaterium*, *B. mesonae*, *B. methanolicus*, *B. methylotrophicus*, *B. migulanus*, *B. mojaviensis*, *B. mucilaginosus*, *B. muralis*, *B.*
- 25 *murimartini*, *B. mycoides*, *B. naganoensis*, *B. nanhaiensis*, *B. nanhaiisediminis*, *B. nealsonii*, *B. neidei*, *B. neizhouensis*, *B. niabensis*, *B. niacini*, *B. novalis*, *B. oceanisediminis*, *B. odysseyi*, *B. okhensis*, *B. okuhidensis*, *B. oleronius*, *B. oryzaecorticis*, *B. oshimensis*, *B. pabuli*, *B. pakistanensis*, *B. pallidus*, *B. pallidus*, *B. panacisoli*, *B. panaciterrae*, *B. pantothenicus*, *B. parabrevis*, *B. paraflexus*, *B. pasteurii*, *B. patagoniensis*, *B. peoriae*, *B.*
- 30 *persepoleensis*, *B. persicus*, *B. pervagus*, *B. plakortidis*, *B. pocheonensis*, *B. polygoni*, *B. polymyxa*, *B. popilliae*, *B. pseudalcalophilus*, *B. pseudofirmus*, *B. pseudomycoides*, *B. psychrodurans*, *B. psychrophilus*, *B. psychrosaccharolyticus*, *B. psychrotolerans*, *B. pulvificiens*, *B. pumilus*, *B. purgationiresistens*, *B. pycnus*, *B. qingdaonensis*, *B. qingshengii*,

- B. reuszeri*, *B. rhizosphaerae*, *B. rigui*, *B. ruris*, *B. safensis*, *B. salarius*, *B. salexigens*, *B. saliphilus*, *B. schlegelii*, *B. sediminis*, *B. selenatarsenatis*, *B. selenitireducens*, *B. seohaeanensis*, *B. shacheensis*, *B. shackletonii*, *B. siamensis*, *B. silvestris*, *B. simplex*, *B. siralis*, *B. smithii*, *B. soli*, *B. solimangrovi*, *B. solisalsi*, *B. songklensis*, *B. sonorensis*, *B. sphaericus*, *B. sporothermodurans*, *B. stearothermophilus*, *B. stratosphericus*, *B. subterraneus*, *B. subtilis*, *B. s. subsp. inaquosorum*, *B. s. subsp. spizizenii*, *B. s. subsp. subtilis*, *B. taeaanensis*, *B. tequilensis*, *B. thermantarcticus*, *B. thermoaerophilus*, *B. thermoamylovorans*, *B. thermocatenulatus*, *B. thermocloacae*, *B. thermocopriae*, *B. thermodenitrificans*, *B. thermoglucosidasius*, *B. thermolactis*, *B. thermoleovorans*, *B. thermophilus*, *B. thermoruber*, *B. thermosphaericus*, *B. thiaminolyticus*, *B. thioparans*, *B. thuringiensis*, *B. tianshenii*, *B. trypoxylicola*, *B. tusciae*, *B. validus*, *B. vallismortis*, *B. vedderi*, *B. velezensis*, *B. vietnamensis*, *B. vireti*, *B. vulcani*, *B. wakoensis*, *B. weihenstephanensis*, *B. xiamenensis*, *B. xiaoxiensis*, or *B. zhanjiangensis*) urea-binding protein; a *Desulfotomaculum* sp. (e.g., *D. ruminis*, *D. nigrificans*, *D. australicum*, *D. thermobenzoicum*, *D. geothermicum*, *D. thermocisternum*, *D. aeronauticum*, *D. halophilum*, *D. kuznetsovii*, *D. thermoacetoxidans*, *D. thermosapovorans*, *D. acetoxidans*, *D. reducens*, *D. putei*, *D. luciae*, *D. gibsoniae*, *D. sapomandens*, *D. alkaliphilum*, *D. sp. FSB6*, *D. sp. ASRB-Zg*, *D. sp. 175*, *D. sp. 176*, *D. sp. 171*, *D. sp. C40-3*, *D. sp. TPOSR*, *D. sp. WW1*, *D. sp. SRB-M*, *D. sp. Mechichi-2001*, *D. solfataricum*, *D. sp. ECP-C5*, *D. sp. MPNeg1*, *D. sp. Ox39*, *D. sp. RL50L1*, *D. alcoholivorax*, *D. sp. NC402*, *D. sp. NB401*, *D. sp. NA401*, *D. salinum*, *D. carboxydivorans*, *D. arcticum*, *D. thermosubterraneum*, *D. indicum*, *D. sp. Lac2*, *D. sp. CYP1*, *D. sp. CYP9*, *D. sp. IS3205*, *D. sp. Srb55*, *D. sp. Iso-W2*, *D. sp. 2*, *D. hydrothermale*, *D. sp. ADR22*, *D. sp. Hbr7*, *D. sp. JD175*, *D. sp. JD176*, *D. sp. DSM 7440*, *D. sp. DSM 7474*, *D. sp. DSM 7475*, *D. sp. DSM 7476*, *D. sp. DSM 8775*, *D. sp. cs1-2*, or *D. sp. MJ1*) urea-binding protein; a *Geobacillus* sp. (e.g., *G. thermoglucosidasius*, *G. stearothermophilus*, *G. jurassicus*, *G. toebii*) urea-binding protein; a *Clostridium* sp. (e.g., *C. absonum*, *C. aceticum*, *C. acetireducens*, *C. acetobutylicum*, *C. acidisoli*, *C. aciditolerans*, *C. acidurici*, *C. aerotolerans*, *C. aestuarii*, *C. akagii*, *C. aldenense*, *C. aldrichii*, *C. algidicarni*, *C. algidixylanolyticum*, *C. algifaecis*, *C. algoriphilum*, *C. alkalicellulosi*, *C. aminophilum*, *C. aminovalericum*, *C. amygdalinum*, *C. amyolyticum*, *C. arbusti*, *C. arcticum*, *C. argentinense*, *C. asparagiforme*, *C. aurantibutyricum*, *C. autoethanogenum*, *C. baratii*, *C. barkeri*, *C. bartlettii*, *C. beijerinckii*, *C. bifermentans*, *C. bolteae*, *C. bornimense*, *C. botulinum*, *C. bowmanii*, *C. bryantii*, *C. butyricum*, *C. cadaveris*, *C. caenicola*, *C. caminithermale*, *C.*

- carboxidivorans*, *C. carnis*, *C. cavendishii*, *C. celatum*, *C. celerecrescens*, *C. cellobioparum*, *C. cellulofermentans*, *C. cellulolyticum*, *C. cellulosi*, *C. cellulovorans*, *C. chartatabidum*, *C. chauvoei*, *C. chromiireducens*, *C. citroniae*, *C. clariflavum*, *C. clostridioforme*, *C. coccoides*, *C. cochlearium*, *C. colletant*, *C. colicanis*, *C. colinum*, *C. collagenovorans*, *C.*
- 5 *cylindrosporum*, *C. difficile*, *C. diolis*, *C. disporicum*, *C. drakei*, *C. durum*, *C. estertheticum*, *C. estertheticum estertheticum*, *C. estertheticum laramiense*, *C. fallax*, *C. felsineum*, *C.*
- fervidum*, *C. fimetarium*, *C. formicaceticum*, *C. frigidicarnis*, *C. frigoris*, *C. ganghwense*, *C. gasigenes*, *C. ghonii*, *C. glycolicum*, *C. glycyrrhizinilyticum*, *C. grantii*, *C. haemolyticum*, *C.*
- 10 *halophilum*, *C. hastiforme*, *C. hathewayi*, *C. herbivorans*, *C. hiranonis*, *C. histolyticum*, *C.*
- homopropionicum*, *C. huakuii*, *C. hungatei*, *C. hydrogeniformans*, *C. hydroxybenzoicum*, *C.*
- hylemonae*, *C. jejuense*, *C. indolis*, *C. innocuum*, *C. intestinale*, *C. irregulare*, *C. isatidis*, *C.*
- josui*, *C. kluyveri*, *C. lactatifermentans*, *C. lacusfryxellense*, *C. laramiense*, *C. lavalense*, *C.*
- lentocellum*, *C. lentoputrescens*, *C. leptum*, *C. limosum*, *C. litorale*, *C. lituseburense*, *C.*
- 15 *ljungdahlii*, *C. lortetii*, *C. lundense*, *C. magnum*, *C. malenominatum*, *C. manganotii*, *C.*
- mayombei*, *C. methoxybenzovorans*, *C. methylpentosum*, *C. neopropionicum*, *C. nexile*, *C.*
- nitrophenolicum*, *C. novyi*, *C. oceanicum*, *C. orbiscindens*, *C. oroticum*, *C. oxalicum*, *C.*
- papyrosolvens*, *C. paradoxum*, *C. paraperfringens*, *C. paraputrificum*, *C. pascui*, *C.*
- pasteurianum*, *C. peptidivorans*, *C. perenne*, *C. perfringens*, *C. pfennigii*, *C.*
- phytofermentans*, *C. piliforme*, *C. polysaccharolyticum*, *C. populeti*, *C. propionicum*, *C.*
- 20 *proteoclasticum*, *C. proteolyticum*, *C. psychrophilum*, *C. puniceum*, *C. purinilyticum*, *C.*
- putrefaciens*, *C. putrificum*, *C. quercicolum*, *C. quinii*, *C. ramosum*, *C. rectum*, *C. roseum*, *C.*
- saccharobutylicum*, *C. saccharogumia*, *C. saccharolyticum*, *C. saccharoperbutylacetonicum*, *C.*
- sardiniense*, *C. sartagoforme*, *C. scatologenes*, *C. schirmacherense*, *C. scindens*, *C.*
- septicum*, *C. sordellii*, *C. sphenoides*, *C. spiroforme*, *C. sporogenes*, *C. sporosphaeroides*, *C.*
- 25 *stercorarium*, *C. stercorarium leptospartum*, *C. stercorarium stercorarium*, *C. stercorarium*
- thermolacticum*, *C. sticklandii*, *C. straminisolvens*, *C. subterminale*, *C. sufflavum*, *C.*
- sulfidigenes*, *C. symbiosum*, *C. tagluense*, *C. tepidiprofundii*, *C. termitidis*, *C. tertium*, *C.*
- tetani*, *Clostridium tetanomorphum*, *C. thermaceticum*, *C. thermautotrophicum*, *C.*
- thermoalcaliphilum*, *C. thermobutyricum*, *C. thermocellum*, *C. thermocopriae*, *C.*
- 30 *thermohydrosulfuricum*, *C. thermolacticum*, *C. thermopalmarium*, *C. thermopapyrolyticum*,
- C. thermosaccharolyticum*, *C. thermosuccinogenes*, *C. thermosulfurigenes*, *C.*
- thiosulfatireducens*, *C. tyrobutyricum*, *C. uliginosum*, *C. ultunense*, *C. villosum*, *C. vincentii*,
- C. viride*, *C. xylanolyticum*, or *C. xylanovorans*) urea-binding protein; a *Caldicellulosiruptor*

sp. (e.g., *C. acetigenus*, *C. bescii*, *C. changbaiensis*, *C. hydrothermalis*, *C. kristjanssonii*, *C. kronotskyensis*, *C. lactoaceticus*, *C. owensensis*, or *C. saccharolyticus*) urea-binding protein; a *Thermocrinis* sp. (e.g., *T. ruber*, *T. albus*, or *T. minervae*) urea-binding protein; a *Synechoccus* sp. (e.g., *S. ambiguus*, *S. arcuatus* var. *calciculus*, *S. bigranulatus*, *S. brunneolus*, *S. caldarius*, *S. capitatus*, *S. carcerarius*, *S. elongatus*, *S. endogloeicus*, *S. epigloeicus*, *S. ferruginosus*, *S. intermedius*, *S. koidzumii*, *S. lividus*, *S. marinus*, *S. minutissimus*, *S. mundulus*, *S. nidulans*, *S. rayssae*, *S. rhodobaktron*, *S. roseo-persicinus*, *S. roseo-purpureus*, *S. salinarum*, *S. salinus*, *S. sciophilus*, *S. sigmoideus*, *S. spongiarum*, *S. subsalsus*, *S. sulphuricus*, *S. vantieghemii*, *S. violaceus*, *S. viridissimus*, or *S. vulcanus*) urea-binding protein; a *Paenibacillus* sp. (e.g., *P. agarexedens*, *P. agaridevorans*, *P. alginolyticus*, *P. alkaliterrae*, *P. alvei*, *P. amylolyticus*, *P. anaericanus*, *P. antarcticus*, *P. assamensis*, *P. azoreducens*, *P. azotofixans*, *P. barcinonensis*, *P. borealis*, *P. brasiliensis*, *P. brassicae*, *P. campinasensis*, *P. chinjuensis*, *P. chitinolyticus*, *P. chondroitinus*, *P. cineris*, *P. cookii*, *P. curdlanolyticus*, *P. daejeonensis*, *P. dendritiformis*, *P. durum*, *P. ehimensis*, *P. elgii*, *P. favisporus*, *P. glucanolyticus*, *P. glycanilyticus*, *P. gordonae*, *P. graminis*, *P. granivorans*, *P. hodogayensis*, *P. illinoisensis*, *P. jamilae*, *P. kobensis*, *P. koleovorans*, *P. koreensis*, *P. kribbensis*, *P. lactis*, *P. larvae*, *P. lautus*, *P. lentimorbus*, *P. macerans*, *P. macquariensis*, *P. massiliensis*, *P. mendelii*, *P. motobuensis*, *P. naphthalenovorans*, *P. nematophilus*, *P. odorifer*, *P. pabuli*, *P. peoriae*, *P. phoenicis*, *P. phyllosphaerae*, *P. polymyxa*, *P. popilliae*, *P. pulvifaciens*, *P. rhizosphaerae*, *P. sanguinis*, *P. stellifer*, *P. terrae*, *P. thiaminolyticus*, *P. timonensis*, *P. tylopili*, *P. turicensis*, *P. validus*, *P. vortex*, *P. vulneris*, *P. wynnii*, *P. xylanilyticus*) urea-binding protein; or a *Thermosynechococcus* sp. (e.g., *T. elongatus* or *T. vulcanus*) urea-binding protein.

The present subject matter provides a glucose-binding protein that is or is a mutant of: an *Thermus* sp. (e.g., *T. caldophilus*, *T. eggertssonii*, *T. kawarayensis*, *T. murrieta*, *T. nonproteolyticus*, *T. parvatiensis*, *T. rehai*, *T. yunnanensis*, *T. amyloliquefaciens*, *T. antranikianii*, *T. aquaticus*, *T. arciformis*, *T. brockianus*, *T. caliditerrae*, *T. chliarophilus*, *T. composti*, *T. filiformis*, *T. igniterrae*, *T. islandicus*, *T. oshimai*, *T. profundus*, *T. scotoductus*, *T. tengchongensis*, or *T. thermophilus*) glucose-binding protein; a *Deinococcus* sp. (e.g., *D. aequivivus*, *D. puniceus*, *D. soli*, *D. xibeensis*, *D. aerius*, *D. aerolatus*, *D. aerophilus*, *D. aetherius*, *D. alpinitundrae*, *D. altitudinis*, *D. apachensis*, *D. aquaticus*, *D. aquatilis*, *D. aquiradiocola*, *D. caeni*, *D. cellulosityticus*, *D. claudionis*, *D. daejeonensis*, *D. depolymerans*, *D. deserti*, *D. erythromyxa*, *D. ficus*, *D. frigens*, *D. geothermalis*, *D.*

- gobiensis*, *D. grandis*, *D. hohokamensis*, *D. hopiensis*, *D. indicus*, *D. maricopenensis*, *D. marmoris*, *D. metalli*, *D. misasensis*, *D. murrayi*, *D. navajonensis*, *D. papagonensis*, *D. peraridilitoris*, *D. pimensis*, *D. piscis*, *D. proteolyticus*, *D. radiodurans*, *D. radiomollis*, *D. radiophilus*, *D. radiopugnans*, *D. reticulitermitis*, *D. roseus*, *D. saxicola*, *D. sonorensis*, *D.*
- 5 *wulumuqiensis*, *D. xibeiensis*, *D. xinjiangensis*, *D. yavapaiensis*, or *D. yunweiensis*) glucose-binding protein; a *Thermotoga* sp. (e.g., *T. caldifontis*, *T. elfii*, *T. hypogea*, *T. lettingae*, *T. maritima*, *T. naphthophila*, *T. neapolitana*, *T. petrophila*, *T. profunda*, *T. subterranea*, or *T. thermarum*) glucose-binding protein; a *Kosmotoga* sp. (e.g., *K. olearia*, *K. arenicorallina*, *K. pacifica*, or *K. shengliensis*) glucose-binding protein; a *Bacillus* sp. (e.g., *B. acidicer*, *B.*
- 10 *acidicola*, *B. acidiproducens*, *B. acidocaldarius*, *B. acidoterrestris*, *B. aeolius*, *B. aerius*, *B. aerophilus*, *B. agaradhaerens*, *B. agri*, *B. aidingensis*, *B. akibai*, *B. alcalophilus*, *B. algicola*, *B. alginolyticus*, *B. alkalidiazotrophicus*, *B. alkalinitrilicus*, *B. alkalisediminis*, *B. alkalitelluris*, *B. altitudinis*, *B. alveayuensis*, *B. alvei*, *B. amyloliquefaciens*, *B. a. subsp. amyloliquefaciens*, *B. a. subsp. plantarum*, *B. amylolyticus*, *B. andreesenii*, *B.*
- 15 *aneuriniycticus*, *B. anthracis*, *B. aquimaris*, *B. arenosi*, *B. arseniciselenatis*, *B. arsenicus*, *B. aurantiacus*, *B. arvi*, *B. aryabhattai*, *B. asahii*, *B. atrophaeus*, *B. axarquiensis*, *B. azotofixans*, *B. azotoformans*, *B. badius*, *B. barbaricus*, *B. bataviensis*, *B. beijingensis*, *B. benzoovorans*, *B. beringensis*, *B. berkeleyi*, *B. beveridgei*, *B. bogoriensis*, *B. boroniphilus*, *B. borstelensis*, *B. brevis* Migula, *B. butanolivorans*, *B. canaveralius*, *B. carboniphilus*, *B. cecembensis*, *B.*
- 20 *cellulosilyticus*, *B. centrosporus*, *B. cereus*, *B. chagannorensis*, *B. chitinolyticus*, *B. chondroitinus*, *B. choshinensis*, *B. chungangensis*, *B. cibi*, *B. circulans*, *B. clarkii*, *B. clausii*, *B. coagulans*, *B. coahuilensis*, *B. cohnii*, *B. composti*, *B. curdlanolyticus*, *B. cycloheptanicus*, *B. cytotoxicus*, *B. daliensis*, *B. decisifrondis*, *B. decolorationis*, *B. deserti*, *B. dipsosauri*, *B. drementensis*, *B. edaphicus*, *B. ehimensis*, *B. eiseniae*, *B. enclensis*, *B. endophyticus*, *B.*
- 25 *endoradicis*, *B. farraginis*, *B. fastidiosus*, *B. fengqiensis*, *B. firmus*, *B. flexus*, *B. foraminis*, *B. fordii*, *B. formosus*, *B. fortis*, *B. fumarioli*, *B. funiculus*, *B. fusiformis*, *B. galactophilus*, *B. galactosidilyticus*, *B. galliciensis*, *B. gelatini*, *B. gibsonii*, *B. ginsengi*, *B. ginsengihumi*, *B. ginsengisoli*, *B. globisporus*, *B. g. subsp. globisporus*, *B. g. subsp. marinus*, *B.*
- glucanolyticus*, *B. gordonae*, *B. gottheilii*, *B. graminis*, *B. halmapalus*, *B. haloalkaliphilus*, *B.*
- 30 *halochares*, *B. halodenitrificans*, *B. halodurans*, *B. halophilus*, *B. halosaccharovorans*, *B. hemicellulosilyticus*, *B. hemicentroti*, *B. herbersteinensis*, *B. horikoshii*, *B. horneckiae*, *B. horti*, *B. huizhouensis*, *B. humi*, *B. hwajinpoensis*, *B. idriensis*, *B. indicus*, *B. infantis*, *B. infernus*, *B. insolitus*, *B. invictae*, *B. iranensis*, *B. isabeliae*, *B. isronensis*, *B. jeotgali*, *B.*

- kaustophilus*, *B. kobensis*, *B. kochii*, *B. kokeshiiformis*, *B. koreensis*, *B. korlensis*, *B. kribbensis*, *B. krulwichiae*, *B. laevolacticus*, *B. larvae*, *B. laterosporus*, *B. lautus*, *B. lehensis*, *B. lentimorbus*, *B. lentus*, *B. licheniformis*, *B. ligniniphilus*, *B. litoralis*, *B. locisalis*, *B. luciferensis*, *B. luteolus*, *B. luteus*, *B. macauensis*, *B. macerans*, *B. macquariensis*, *B. macyae*,
5 *B. malacitensis*, *B. mannanyticus*, *B. marisflavi*, *B. marismortui*, *B. marmarensis*, *B. massiliensis*, *B. megaterium*, *B. mesonae*, *B. methanolicus*, *B. methylotrophicus*, *B. migulanus*, *B. mojavenensis*, *B. mucilaginosus*, *B. muralis*, *B. murimartini*, *B. mycoides*, *B. naganoensis*, *B. nanhaiensis*, *B. nanhaiisediminis*, *B. nealsonii*, *B. neidei*, *B. neizhouensis*, *B. niabensis*, *B. niacini*, *B. novalis*, *B. oceanisediminis*, *B. odysseyi*, *B. okhensis*, *B. okuhidensis*,
10 *B. oleronius*, *B. oryzaecorticis*, *B. oshimensis*, *B. pabuli*, *B. pakistanensis*, *B. pallidus*, *B. pallidus*, *B. panacisoli*, *B. panaciterrae*, *B. pantothenicus*, *B. parabrevis*, *B. paraflexus*, *B. pasteurii*, *B. patagoniensis*, *B. peoriae*, *B. persepolensis*, *B. persicus*, *B. pervagus*, *B. plakortidis*, *B. pocheonensis*, *B. polygoni*, *B. polymyxa*, *B. popilliae*, *B. pseudalcalophilus*, *B. pseudofirmus*, *B. pseudomycoides*, *B. psychrodurans*, *B. psychrophilus*, *B.*
15 *psychrosaccharolyticus*, *B. psychrotolerans*, *B. pulvifaciens*, *B. pumilus*, *B. purgationiresistens*, *B. pycnus*, *B. qingdaonensis*, *B. qingshengii*, *B. reuszeri*, *B. rhizosphaerae*, *B. rigui*, *B. ruris*, *B. safensis*, *B. salarius*, *B. salexigens*, *B. saliphilus*, *B. schlegelii*, *B. sediminis*, *B. selenatarsenatis*, *B. selenitireducens*, *B. seohaeanensis*, *B. shacheensis*, *B. shackletonii*, *B. siamensis*, *B. silvestris*, *B. simplex*, *B. siralis*, *B. smithii*, *B.*
20 *soli*, *B. solimangrovi*, *B. solisalsi*, *B. songklensis*, *B. sonorensis*, *B. sphaericus*, *B. sporothermodurans*, *B. stearothermophilus*, *B. stratosphericus*, *B. subterraneus*, *B. subtilis*, *B. s. subsp. inaquosorum*, *B. s. subsp. spizizenii*, *B. s. subsp. subtilis*, *B. taeanensis*, *B. tequilensis*, *B. thermantarcticus*, *B. thermoaerophilus*, *B. thermoamylovorans*, *B. thermocatenulatus*, *B. thermocloacae*, *B. thermocopriae*, *B. thermodenitrificans*, *B.*
25 *thermoglucoasidarius*, *B. thermolactis*, *B. thermoleovorans*, *B. thermophilus*, *B. thermoruber*, *B. thermosphaericus*, *B. thiaminolyticus*, *B. thioparans*, *B. thuringiensis*, *B. tianshenii*, *B. trypoxylicola*, *B. tusciae*, *B. validus*, *B. vallismortis*, *B. vedderi*, *B. velezensis*, *B. vietnamensis*, *B. vireti*, *B. vulcani*, *B. wakoensis*, *B. weihenstephanensis*, *B. xiamenensis*, *B. xiaoxiensis*, or *B. zhanjiangensis*) glucose-binding protein; a *Staphylothermus* sp. (e.g., *S.*
30 *hellenicus* or *S. marinus*) glucose-binding protein; or an *Arthrobacter* sp. (e.g., *A. agilis*, *A. alkaliphilus*, *A. alpinus*, *A. antarcticus*, *A. aurescens*, *A. bambusae*, *A. castelli*, *A. chlorophenolicus*, *A. citreus*, *A. cryoconiti*, *A. cryotolerans*, *A. crystallopoietes*, *A. cumminsii*, *A. cupressi*, *A. defluvii*, *A. enclensis*, *A. flavus*, *A. gandavensis*, *A. globiformis*, *A.*

gyeryongensis, *A. halodurans*, *A. histidinovorans*, *A. humicola*, *A. koreensis*, *A. liuii*, *A. livingstonensis*, *A. luteolus*, *A. methylotrophus*, *A. monumenti*, *A. nanjingensis*, *A. nasiphocae*, *A. nicotinovorans*, *A. nitroguajacolicus*, *A. oryzae*, *A. parietis*, *A. pascens*, *A. pigmenti*, *A. pityocampae*, *A. psychrochitiniphilus*, *A. psychrolactophilus*, *A. ramosus*, *A. rhombi*, *A. roseus*, *A. russicus*, *A. sanguinis*, *A. soli*, *A. stackebrandtii*, *A. subterraneus*, *A. tecti*, *A. tumbae*, *A. viscosus*, or *A. woluwensis*) glucose-binding protein.

The present subject matter provides a lactate-binding protein that is or is a mutant of: a *Thermus* sp. (e.g., *T. caldophilus*, *T. eggertssonii*, *T. kawayensis*, *T. murrieta*, *T. nonproteolyticus*, *T. parvatiensis*, *T. rehai*, *T. yunnanensis*, *T. amyloliquefaciens*, *T.*

- 10 *antranikianii*, *T. aquaticus*, *T. arciformis*, *T. brockianus*, *T. caliditerrae*, *T. chliarophilus*, *T. composti*, *T. filiformis*, *T. igniterrae*, *T. islandicus*, *T. oshimai*, *T. profundus*, *T. scotoductus*, *T. shimai*, *T. tengchongensis*, or *T. thermophilus*) lactate-binding protein, a *Thioalkalivibrio* sp. (e.g., *T. denitrificans*, *T. halophilus*, *T. jannaschii*, *T. nitratreducens*, *T. nitratis*, *T. paradoxus*, *T. sulfidiphilus*, *T. thiocyanodenitrificans*, *T. thiocyanoxidans*, and *T. versutus*)
- 15 lactate-binding protein, a *Roseobacter* sp. (e.g., *R. dentrificans*, *R. litoralis*, *R. pelophilus*, *R. prionitis*, *R. sp. 14III/A01/004*, *R. sp. 1922*, *R. sp. 27-4*, *R. sp. 3008*, *R. sp. 38.98*, *R. sp. 3X/A02/234*, *R. sp. 4318-8/1*, *R. sp. 812*, *R. sp. 8-1*, *R. sp. ANT8230*, *R. sp. ANT9082*, *R. sp. ANT909*, *R. sp. ANT9234*, *R. sp. ANT9240*, *R. sp. ANT9270*, *R. sp. ANT9274*, *R. sp. ANT9276a*, *R. sp. ANT9283*, *R. sp. ARCTIC-P4*, *R. sp. ARK9990*, *R. sp. AzwK-3b*, *R. sp.*
- 20 *AzwLept-1c*, *R. sp. B09*, *R. sp. B-1039*, *R. sp. B11*, *R. sp. Ber2105*, *R. sp. Ber2107*, *R. sp. BS36*, *R. sp. BS90*, *R. sp. C115*, *R. sp. C23*, *R. sp. CCS2*, *R. sp. COL2P*, *R. sp. COLSP*, *R. sp. DG1132*, *R. sp. DG869*, *R. sp. DG889*, *R. sp. DG942*, *R. sp. Do-34*, *R. sp. DSS-1*, *R. sp. DSS-8*, *R. sp. H264*, *R. sp. H265*, *R. sp. H454*, *R. sp. HJ105*, *R. sp. HJ247*, *R. sp. HYL-SA-18*, *R. sp. HZBC52*, *R. sp. HZDC27*, *R. sp. HZDC41*, *R. sp. HZDC42*, *R. sp. HZDC43*, *R. sp.*
- 25 *HZDC7*, *R. sp. J2W*, *R. sp. J356*, *R. sp. J392*, *R. sp. J483*, *R. sp. J486*, *R. sp. J504*, *R. sp. J8W*, *R. sp. JL-126*, *R. sp. JL-129*, *R. sp. JL-131*, *R. sp. JL-132*, *R. sp. JL-135*, *R. sp. JL-137*, *R. sp. JL-351*, *R. sp. JL985*, *R. sp. JLN-A020*, *R. sp. JLN-A036*, *R. sp. JLN-A530*, *R. sp. KAT10*, *R. sp. KAT3*, *R. sp. KT0202a*, *R. sp. KT0917*, *R. sp. KT1117*, *R. sp. LA7*, *R. sp. LA9*, *R. sp. LOB-8*, *R. sp. LZXC12*, *R. sp. LZXC14*, *R. sp. LZXC15*, *R. sp. LZXC16*, *R. sp. LZXC20*, *R.*
- 30 *sp. LZXC23*, *R. sp. LZXC26*, *R. sp. LZXC29*, *R. sp. LZXC4*, *R. sp. LZXC7*, *R. sp. MBT21*, *R. sp. MBT22*, *R. sp. MED001*, *R. sp. MED006*, *R. sp. MED007*, *R. sp. MED008*, *R. sp. MED193*, *R. sp. MED24*, *R. sp. MED26*, *R. sp. MED61*, *R. sp. MED6*, *R. sp. MSI042*, *R. sp. N05I*, *R. sp. NJSS27*, *R. sp. NT N37*, *R. sp. OC-B2-7*, *R. sp. OC-C4-20*, *R. sp. OO-A3-7*, *R.*

- sp. OO-C4-10, *R. sp.* Pht-4, *R. sp.* PIC-68, *R. sp.* PRLIST02, *R. sp.* PRLIST06, *R. sp.* PRLISY01, *R. sp.* PRLISY03, *R. sp.* QSSC9-8, *R. sp.* RED15, *R. sp.* RED1, *R. sp.* RED59, *R. sp.* RED68, *R. sp.* RED85, *R. sp.* S03, *R. sp.* SC-B2-2, *R. sp.* SCB28, *R. sp.* SCB31, *R. sp.* SCB34, *R. sp.* SCB48, *R. sp.* SDBC1, *R. sp.* SDBC6, *R. sp.* SFLA13, *R. sp.* SIO, *R. sp.* SK209-2-6, *R. sp.* SKA26, *R. sp.* SKA44, *R. sp.* SL25, *R. sp.* SO3, *R. sp.* SPO804, *R. sp.* SYOP1, *R. sp.* SYOP2, *R. sp.* TM1035, *R. sp.* TM1038, *R. sp.* TM1040, *R. sp.* TM1042, *R. sp.* TP9, *R. sp.* UAzPsJAC-1b, *R. sp.* UAzPsK-5, *R. sp.* WED10.10, *R. sp.* WED1.1, *R. sp.* WHOI JT-01, *R. sp.* WHOI JT-08, *R. sp.* WHOI JT-22, *R. sp.* WM2, *R. sp.* Y2, *R. sp.* Y3F, *R. sp.* YS-57, *R. sp.* YSCB-1, or *R. sp.* YSCB-3) lactate-binding protein, a *Marinobacter* sp.
- 10 (e.g., *M. adhaerens*, *M. algicola*, *M. alkaliphilus*, *M. antarcticus*, *M. arcticus*, *M. aromaticivorans*, *M. bryozoorum*, *M. daepoensis*, *M. daqiaonensis*, *M. excellens*, *M. flavimaris*, *M. gudaonensis*, *M. guineae*, *M. halophilus*, *M. gudaonensis*, *M. hydrocarbonoclasticus*, *M. koreensis*, *M. lacisalsi*, *M. lipolyticus*, *M. litoralis*, *M. lutaoensis*, *M. maritimus*, *M. mobilis*, *M. nitratreducens*, *M. oulmenensis*, *M. pelagius*, *M. persicus*, *M. psychrophilus*, *M. nanhaiticus*, *M. salarii*, *M. salicampi*, *M. salsuginis*, *M. santoriniensis*, *M. sediminum*, *M. segnicrescens*, *M. shengliensis*, *M. squalenivorans*, *M. similis*, *M. szutsaonensis*, *M. vinifirmus*, *M. xestospongiae*, *M. zhanjiangensis*, or *M. zhejiangensis*) lactate-binding protein, a *Anaeromyxobacter* sp. (e.g., *A. dehalogenans*) lactate-binding protein, a *Polymorphum* sp. (e.g., *P. gilvum*) lactate-binding protein, a *Pseudomonas* sp.
- 15 (e.g., *P. aeruginosa*, *P. alcaligenes*, *P. anguilliseptica*, *P. argentinensis*, *P. borbori*, *P. citronellolis*, *P. flavescens*, *P. mendocina*, *P. nitroreducens*, *P. oleovorans*, *P. pseudoalcaligenes*, *P. resinovorans*, *P. straminea*, *P. chlororaphis*, *P. asplenii*, *P. aurantiaca*, *P. aureofaciens*, *P. chlororaphis*, *P. corrugata*, *P. fragi*, *P. lundensis*, *P. taetrolens*, *P. antarctica*, *P. azotoformans*, *P. blatchfordae*, *P. brassicacearum*, *P. brenneri*,
- 25 *P. cedrina*, *P. corrugata*, *P. fluorescens*, *P. gessardii*, *P. libanensis*, *P. mandelii*, *P. marginalis*, *P. mediterranea*, *P. meridiana*, *P. migulae*, *P. mucidolens*, *P. orientalis*, *P. panacis*, *P. protegens*, *P. proteolytica*, *P. rhodesiae*, *P. synxantha*, *P. thivervalensis*, *P. tolaasii*, *P. veronii*, *P. denitrificans*, *P. pertucinogena*, *P. cremoricolorata*, *P. entomophila*, *P. fulva*, *P. monteilii*, *P. mosselii*, *P. oryzihabitans*, *P. parafulva*, *P. plecoglossicida*, *P. putida*, *P. balearica*, *P. luteola*, *P. stutzeri*, *P. amygdali*, *P. avellanae*, *P. caricapapayae*, *P. cichorii*, *P. coronafaciens*, *P. ficuserectae*, *P. helianthi*, *P. meliae*, *P. savastanoi*, *P. syringae*, *P. tomato*, or *P. viridiflava*) lactate-binding protein, a *Rhodobacter* sp. (e.g., *R. aestuarii*, *R. azotoformans*, *R. blasticus*, *R. capsulatus*, *R. johrii*, *R. maris*, *R. megalophilus*, *R. ovatus*,
- 30

R. sphaeroides, *R. veldkampii*, *R. vinaykumarii*, or *R. viridis*) lactate-binding protein, a *Flexistipes* sp. (e.g., *F. sinusarabici*) lactate-binding protein, or a *Thermanaerovibrio* sp. (e.g., *T. acidaminovorans* or *T. velox*) lactate-binding protein.

The present subject matter provides a ligand-binding protein that is or is a mutant of:

- 5 a *Synechocystis* sp. (e.g., *S. sp.* PCC6803) bicarbonate-binding protein, a *Thermosynechococcus* sp. (e.g., *T. vulcanus*, *T. elongatus*, or *T. elongatus BP-1*) bicarbonate-binding protein, a *Chroococcidiopsis* sp. (e.g., *C. thermalis*, *C. gigantea*, *C. cubana*, or *C. codiicola*) bicarbonate-binding protein, a *Calothrix* sp. (e.g., *C. aberrans*, *C. adscendens*, *C. aeruginea*, *C. africana*, *C. allorgei*, *C. australiensis*, *C. baileyi*, *C. bharadwajae*, *C. borealis*,
- 10 *C. braunii*, *C. breviararticulata*, *C. calida*, *C. castellii*, *C. capitularis*, *C. cavernarum* Copeland, *C. charicola*, *C. clavata*, *C. clavatoidea*, *C. codicola*, *C. columbiana*, *C. compacta*, *C. confervicola*, *C. contarenii*, *C. coriacea*, *C. crustacea*, *C. cylindrica*, *C. desertica*, *C. elsteri*, *C. epiphytica*, *C. evanescens*, *C. estonica*, *C. fasciculata*, *C. feldmannii*, *C. flahaultii*, *C. floccosa*, *C. fritschii*, *C. fuelebornii*, *C. fusca*, *C. fusco-violacea*, *C. geilterii*, *C. geitonos*,
- 15 *C. ghosei*, *C. gigas*, *C. gloeocola*, *C. goetzei*, *C. hunanica*, *C. inaequabilis*, *C. inserta*, *C. javanica*, *C. karnatakensis*, *C. kawraiskyi*, *C. kossinskajae*, *C. kuntzei*, *C. linearis*, *C. minima*, *C. nidulans*, *C. parasitica*, *C. parietina*, *C. parva*, *C. pilosa*, *C. prolifera*, *C. pulvinata*, *C. rectangularis*, *C. reptans*, *C. rodriguezii*, *C. santapaui*, *C. scopulorum*, *C. scytonemicola*, *C. simplex*, *C. simulans*, *C. stagnalis*, *C. subantarctica*, *C. subsimplex*, *C. tenella*, *C. thermalis*,
- 20 *C. turfosa*, *C. viguieri*, *C. vivipara*, *C. violacea*, *C. weberi*, *C. wembaerensis*, *C. aeruginosa*, *C. aestuarii*, *C. antarctica*, *C. atricha*, *C. bossei*, *C. brevissima*, *C. clausa*, *C. conica*, *C. dniprensis*, *C. elenkinii*, *C. fonticola*, *C. galpinii*, *C. gelatinosa*, *C. gracilis*, *C. intricata*, *C. litoralis*, *C. marchica*, *C. nodulosa*, *C. obtusa*, *C. rhizosoleniae*, or *C. schweickertii*) bicarbonate-binding protein, a *Anabaena* sp. (e.g., *A. aequalis*, *A. affinis*, *A. angustumalis*
- 25 *angustumalis*, *A. angustumalis marchita*, *A. aphanizomendoides*, *A. azollae*, *A. bornetiana*, *A. catenula*, *A. cedrorum*, *A. circinalis*, *A. confervoides*, *A. constricta*, *A. cyanobacterium*, *A. cycadeae*, *A. cylindrica*, *A. echinispora*, *A. felisii*, *A. flos-aquae flos-aquae*, *A. flos-aquae minor*, *A. flos-aquae treleasei*, *A. helicoidea*, *A. inaequalis*, *A. lapponica*, *A. laxa*, *A. lemmermannii*, *A. levanderi*, *A. limnetica*, *A. macrospora macrospora*, *A. macrospora*
- 30 *robusta*, *A. monticulosa*, *A. nostoc*, *A. oscillarioides*, *A. planctonica*, *A. raciborskii*, *A. scheremetievi*, *A. sphaerica*, *A. spiroides crassa*, *A. spiroides spiroides*, *A. subcylindrica*, *A. torulosa*, *A. unisporea*, *A. variabilis*, *A. verrucosa*, *A. viguieri*, *A. wisconsinense*, or *A. zierlingii*) bicarbonate-binding protein, or a *Chamaesiphon* sp. (e.g., *C. africanus*, *C.*

amethystinus, *C. britannicus*, *C. carpaticus*, *C. confervicola*, *C. cylindricus*, *C. cylindrosporus*, *C. halophilus*, *C. incrustans*, *C. investiens*, *C. jaoi*, *C. komarekii*, *C. longus*, *C. macer*, *C. major*, *C. minimus*, *C. minutus*, *C. portoricensis*, *C. rostafinskii*, *C. sideriphilus*, *C. tibeticus*, *C. aggregatus*, *C. fallax*, *C. fuscus*, *C. geitleri*, *C. mollis*, *C. niger*, *C. ocobyrsiodes*, *C. polonicus*, *C. polymorphus*, *C. starmachii*, *C. stratosus*, or *C. subglobosus*) bicarbonate-binding protein.

The present subject matter provides a ligand-binding protein that is or is a mutant of: a *Mannheimia* sp. (e.g., *M. caviae*, *M. glucosida*, *M. granulomatis*, *M. haemolytica*, *M. ruminalis*, or *M. varigena*) bicarbonate and iron binding protein, an *Exiguobacterium* sp.

- (e.g., *E. acetylicum*, *E. aestuarii*, *E. alkaliphilum*, *E. antarcticum*, *E. aquaticum*, *E. artemiae*, *E. aurantiacum*, *E. enclense*, *E. indicum*, *E. marinum*, *E. mexicanum*, *E. oxidotolerans*, *E. profundum*, *E. sibiricum*, *E. soli*, or *E. undae*) bicarbonate and iron binding protein, a *Thermosynechococcus* sp. (e.g., *T. vulcanus*, *T. elongatus*, or *T. elongatus BP-1*) bicarbonate and iron binding protein, a *Candidatus Nitrospira* sp. (e.g., *Candidatus Nitrospira defluvii*, *Candidatus Nitrospira nitrificans*, *Candidatus Nitrospira nitrosa*, *Candidatus Nitrospira inopinata*, *Candidatus Magnetobacterium casensis*, *Candidatus Magnetobacterium bavaricum*, *Candidatus Magnetoovum chiemensis*) bicarbonate and iron binding protein, a *Thermus* sp. (e.g., *T. caldophilus*, *T. eggertssonii*, *T. kawarayensis*, *T. murrieta*, *T. nonproteolyticus*, *T. parvatiensis*, *T. rehai*, *T. yunnanensis*, *T. amyloliquefaciens*, *T. antranikianii*, *T. aquaticus*, *T. arciformis*, *T. brockianus*, *T. caliditerrae*, *T. chliarophilus*, *T. composti*, *T. filiformis*, *T. igniterrae*, *T. islandicus*, *T. oshimai*, *T. profundus*, *T. scotoductus*, *T. tengchongensis*, or *T. thermophilus*) bicarbonate and iron binding protein, a *Meiothermus* sp. (*M. chliarophilus*, *M. cerbereus*, *M. granaticius*, *M. rosaceus*, *M. ruber*, *M. rufus*, *M. silvanus*, *M. taiwanensis*, or *M. timidus*) bicarbonate and iron binding protein, a *Salinibacter* sp. (e.g., *S. ruber*, *S. iranicus*, or *S. luteus*) bicarbonate and iron binding protein, or a *Halorubrum* sp. (e.g., *H. aidingense*, *H. alkaliphilum*, *H. arcis*, *H. californiensis*, *H. coriense*, *H. distributum*, *H. ejinorensense*, *H. ezzemoulense*, *H. kocurii*, *H. lacusprofundi*, *H. lipolyticum*, *H. litoreum*, *H. luteum*, *H. orientalis*, *H. saccharovororum*, *H. salsolis*, *H. sodomense*, *H. tebenquichense*, *H. terrestre*, *H. tibetense*, *H. trapanicum*, *H. vacuolatum*, or *H. xinjiangense*) bicarbonate and iron binding protein.

With regard to a defined polypeptide, % identity figures higher or lower than those provided herein will encompass various embodiments. Thus, where applicable, in light of a minimum % identity figure, a polypeptide may comprise an amino acid sequence which is at

least 60%, 65%, 70%, 75%, 76%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to the reference SEQ ID NO or to each of the reference SEQ ID NOs. Where applicable, in light of a maximum % identity to a reference sequence, a polypeptide may comprise an amino acid sequence which is less than 75%, 70%, 65%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, or 15% identical to the reference SEQ ID NO or to each of the reference SEQ ID NOs. In certain embodiments, a polypeptide comprises amino acids in a sequence that is preferably at least about 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, or 30% and less than about 75%, 70%, 65%, 60%, 55%, 50%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, or 30% identical to the reference SEQ ID NO or to each of the reference SEQ ID NOs. Non-limiting examples of reference proteins and amino acid sequences disclosed herein include:

- (i) a glucose-galactose binding protein from *Escherichia coli* (ecGGBP; genome, NC_002695; protein, WP_032329053, SEQ ID NO: 87);
- (ii) a glucose-galactose binding protein from *Thermoanaerobacterium thermosaccharolyticum* (ttGGBP; genome, NC_014410; protein, YP_003852930.1, SEQ ID NO: 88);
- (iii) a glucose-galactose binding protein from *Salmonella typhimurium* (stGGBP; genome, NC_003197; protein, WP_001036943, SEQ ID NO: 89);
- (iv) a glucose-galactose binding protein from *Caldicellulosiruptor hydrothermalis* (chyGGBP; genome, NC_014652; protein identifier, YP_003991244.1, SEQ ID NO: 90);
- (v) a glucose-galactose binding protein from *Caldicellulosiruptor obsidiansis* (cobGGBP; genome, NC_014392; protein, YP_003839461.1, SEQ ID NO: 91);
- (vi) a glucose-galactose binding protein from *Paenibacillus* sp. (pspGGBP; genome, NC_013406; protein, YP_003243743.1, SEQ ID NO: 92);
- (vii) a glucose-galactose binding protein from *Clostridium saccharolyticum* (csaGGBP; genome, NC_014376; protein, YP_003822565.1, SEQ ID NO: 93);

- (viii) a glucose-galactose binding protein from *Butyrivibrio proteoclasticus* (bprGGBP; genome, NC_014387; protein, YP_003830205.1, SEQ ID NO: 94);
- 5 (ix) a glucose-galactose binding protein from *Roseburia intestinalis* (rinGGBP_A; genome, NC_021012; protein, YP_007778116.1, SEQ ID NO: 95);
- (x) a glucose-galactose binding protein from *Faecalibacterium prausnitzii* (fprGGBP; genome, NC_021020; protein, YP_007799070.1, SEQ ID NO: 96);
- 10 (xi) a glucose-galactose binding protein from *Clostridium ljungdahlii* (cljGGBP; genome, NC_014328; protein, CLJU_c08950, SEQ ID NO: 97);
- (xii) a glucose-galactose binding protein from *Clostridium autoethanogenum* (cauGGBP; genome, NC_022592; protein, CAETHG_2989, SEQ ID NO: 98);
- (xiii) a glucose-galactose binding protein from *Roseburia intestinalis* (rinGGBP_B; genome, NC_021012; protein, YP_007778124.1, SEQ ID NO: 99);
- 15 (xiv) a glucose-galactose binding protein from *Erysipelothrix rhusiopathiae* (erhGGBP; genome, NC_015601; protein, YP_004561181.1, SEQ ID NO: 100);
- (xv) a glucose-galactose binding protein from *Eubacterium rectale* (ereGGBP; genome, NC_012781; protein, YP_002936409.1, SEQ ID NO: 101);
- 20 (xvi) a urea-binding protein from *Marinomas posidonica* (mpUBP; genome, NC_015559, protein, YP_004483096.1; SEQ ID NO: 102);
- (xvii) a urea-binding protein from *Marinobacter hydrocarbanoclasticus* (mhUBP; genome, NC_017067, protein, YP_005430828.1; SEQ ID NO: 103);
- 25 (xviii) a urea-binding protein from *Bacillus* sp. (bsUBP; genome, NC_017743, protein, YP_006233530.1; SEQ ID NO: 104);
- (xix) a urea-binding protein from *Desulfotomaculum carboxydvorans* (dsUBP; genome, NC_015565, protein, YP_004496535.1; SEQ ID NO: 105);
- (xx) a urea-binding protein from *Geobacillus thermoglucosidasius* (gtUBP; genome, NC_015660, protein, YP_004588319.1; SEQ ID NO: 106);
- 30 (xxi) a urea-binding protein from *Clostridium thermocellum* (ctUBP; genome, NC_009012, protein, YP_001038237.1; SEQ ID NO: 107);
- (xxii) a urea-binding protein from *Caldicellulosiruptor saccharolyticus* (csUBP; genome, NC_009437, protein, YP_001181243.1; SEQ ID NO: 108);

- (xxiii) a urea-binding protein from *Thermocrinis albus* (taUBP; genome, NC_013894, protein, YP_003473480.1; SEQ ID NO: 109);
- (xxiv) a urea-binding protein from *Geobacillus kaustophilus* (gkUBP; genome, NC_006510, protein, YP_147790.1; SEQ ID NO: 110);
- 5 (xxv) a urea-binding protein from *Paenibacillus* sp. (psUBP; genome, NC_013406, protein, YP_003241723.1; SEQ ID NO: 111);
- (xxvi) a urea-binding protein from *Thermosynechococcus elongatus* (teUBP; genome, NC_004113, protein, YP_681910.1; SEQ ID NO: 112);
- (xxvii) a glucose-binding protein from *Thermus thermophilus* (ttGBP; genome, 10 NC_005835, protein, YP_004303.1 and WP_011172778; SEQ ID NO: 113);
- (xxviii) a glucose-binding protein from *Thermus scotoductus* (tsGBP; genome, NC_014974, protein, YP_004202647.1; SEQ ID NO: 114);
- (xxix) a glucose-binding protein from *Deinococcus maricopenensis* (dmGBP; genome, NC_014958, protein, YP_004171760.1; SEQ ID NO: 115);
- 15 (xxx) a glucose-binding protein from *Thermotoga neapolitana* (tnGBP; genome, NC_011978, protein, YP_002534202.1; SEQ ID NO: 116);
- (xxxi) a glucose-binding protein from *Kosmotoga olearia* (koGBP; genome, NC_012785, protein, YP_0029416871.1; SEQ ID NO: 117);
- (xxxii) a glucose-binding protein from *Bacillus halodurans* (bhGBP; genome, 20 NC_002570, protein, YP_244712.1; SEQ ID NO: 118);
- (xxxiii) a glucose-binding protein from *Staphylothermus marinus* (smGBP; genome, NC_009033, protein, YP_001041152.1; SEQ ID NO: 119);
- (xxxiv) a glucose-binding protein from *Arthrobacter* sp. (asGBP; genome, NC_008541, protein, YP_8313491.1; SEQ ID NO: 120);
- 25 (xxxv) a lactate-binding protein from *Thermus thermophilus* (ttLacBP; genome, NC_006461, protein, YP_144032.1; SEQ ID NO: 121);
- (xxxvi) a lactate-binding protein from *Thermus scotoductus* (tscLacBP; genome, NC_014974, protein YP_004202714.1; SEQ ID NO: 122);
- (xxxvii) a lactate-binding protein from *Thermus oshimai* (toLacBP; genome, 30 NC_019386, protein YP_019386; SEQ ID NO: 123);
- (xxxviii) a lactate-binding protein from *Thioalkalivibrio sulfidophilus* (tsuLacBP; genome, NC_011901, protein YP_002514099.1; SEQ ID NO: 124);

- (xxxix) a lactate-binding protein from *Roseobacter denitrificans* (rdLacBP; genome, NC_008209, protein YP_683924.1; SEQ ID NO: 125);
- (xl) a lactate-binding protein from *Marinobacter* sp. (msLacBP; genome, NC_018268, protein YP_006556686.1; SEQ ID NO: 126);
- 5 (xli) a lactate-binding protein from *Thermus* sp. (tspLacBP; genome, NC_017278, protein YP_005654632.1; SEQ ID NO: 127);
- (xlii) a lactate-binding protein from *Marinobacter adhaerens* (maLacBP; genome, NC_017506, protein YP_005686720.1; SEQ ID NO: 128);
- (xliii) a lactate-binding protein from *Anaeromyxobacter dehalogens* (adLacBP; genome, NC_007760, protein YP_466_099.1; SEQ ID NO: 129);
- 10 (xliv) a lactate-binding protein from *Polymorphum gilvum* (pgLacBP; genome, NC_015259, protein YP_4304976.1; SEQ ID NO: 130);
- (xlv) a lactate-binding protein from *Pseudomonas stutzeri* (psLacBP; genome, NC_018177, protein YP_00652676.1; SEQ ID NO: 131);
- 15 (xlvi) a lactate-binding protein from *Rhodobacter sphaeroides* (rsLacBP; genome, NC_007494, protein RSP_3372; SEQ ID NO: 132);
- (xlvii) a lactate-binding protein from *Flexistipes sinuarabici* (fsLacBP; genome, NC_015672, protein YP_004603455.1; SEQ ID NO: 133);
- (xlviii) a lactate-binding protein from *Thermanaerovibrio acidaminovorans* (taLacBP; genome, NC_013522, protein YP_003317968.1; SEQ ID NO: 134);
- 20 (xlix) a bicarbonate-binding protein from *Synechocystis* sp. (synBicarbBP1; genome, NC_017052, protein YP_005410477.1; SEQ ID NO: 135);
- (l) a bicarbonate-binding protein from *Thermosynechococcus elongatus* (teBicarbBP2; genome, NC_004113, protein NP_682790.1; SEQ ID NO: 136);
- 25 (li) a bicarbonate-binding protein from *Chroococcidiopsis thermalis* (ctBicarbBP3; genome, NC_019695, protein YP_007090308.1; SEQ ID NO: 137);
- (lii) a bicarbonate-binding protein from *Calothrix* sp. (calBicarbBP4; genome, NC_019751, protein YP_007137061.1; SEQ ID NO: 138);
- 30 (liii) a bicarbonate-binding protein from *Anabaena variabilis* (avBicarbBP5; genome, NC_007413, protein YP_321546.1; SEQ ID NO: 139);

- (liv) a bicarbonate-binding protein from *Chamaesiphon minutus* (cmBicarbBP6; genome, NC_019697, protein YP_007099445.1; SEQ ID NO: 140);
- (lv) a bicarbonate and iron binding protein from *Mannheimia haemolytica* (mhFeBP1; genome, NC_0121082, protein, YP_007884192.1; SEQ ID NO: 141);
- (lvi) a bicarbonate and iron binding protein from *Exiguobacterium* sp. (exiFeBP2; genome, NC_012673, protein, YP_002886303.1; SEQ ID NO: 142);
- (lvii) a bicarbonate and iron binding protein from *Thermosynechoccus elongatus* (teFeBP3; genome, NC_004113, protein, NP_681303.1; SEQ ID NO: 143);
- (lviii) a bicarbonate and iron binding protein from *Candidatus nitrospira* (cnFeBP4; genome, NC_014355, protein, YP_003796723.1; SEQ ID NO: 144);
- (lix) a bicarbonate and iron binding protein from *Thermus thermophilus* (ttFeBP5; genome, NC_006461, protein, YP_144894.1; SEQ ID NO: 145);
- (lx) a bicarbonate and iron binding protein from *Meiothermus silvanus* (msFeBP6; genome, NC_014212, protein, YP_003686074.1; SEQ ID NO: 146);
- (lxi) a bicarbonate and iron binding protein from *Salinibacter ruber* (srFeBP7; genome, NC_014032, protein, YP_003572493.1; SEQ ID NO: 147); and
- (lxii) a bicarbonate and iron binding protein from *Halorubrum lacusprofundi* (hlFeBP8; genome, NC_012029, protein, YP_002564837.1; SEQ ID NO: 148).

The ligand-binding proteins disclosed herein may optionally be fused (*e.g.*, at their N-terminal and/or C-terminal ends) to a motif comprising a stretch of amino acids that facilitates the isolation or other manipulation such as conjugation to a moiety or immobilization on a substrate such as a plastic, a cellulose product such as paper, polymer, metal, noble metal, semi-conductor, or quantum dot (*e.g.*, a fluorescent quantum dot). A non-limiting example of such a stretch of amino acids has the sequence: GGSHHHHHH (SEQ ID NO: 152). This motif is not required for, is not believed to influence or affect ligand-binding activity or signal transduction, and may be omitted from any ligand-binding protein or biosensor disclosed herein. Additionally, for every sequence disclosed herein that includes GGSHHHHHH (SEQ ID NO: 152), a corresponding sequence that is identical except that it lacks GGSHHHHHH (SEQ ID NO: 152) is also provided and intended to be disclosed. For example, each of SEQ ID NOs: 1-41 (and the non-limiting examples of other proteins used in the experiments disclosed herein) comprises this motif (SEQ ID NO: 152).

Alternatively or in addition, a ligand-binding protein may be fused to a non-native polypeptide or “added amino acids” that facilitates the attachment thereof to a surface, such as the surface of a device.

In some embodiments, a polypeptide comprises 1, 2, 3, 4, 5, or more substitutions or deletions of a cysteine compared to the naturally occurring counterpart of the polypeptide (*i.e.*, 1, 2, 3, 4, 5, or more native cysteines have been removed), *e.g.*, 1, 2, 3, 4, 5, or more cysteine to alanine substitutions compared to the naturally occurring counterpart of the polypeptide. In some embodiments, all of the cysteines of a polypeptide have been deleted and/or substituted compared to its natural counterpart. In some embodiments, one or more cysteines of a polypeptide have been substituted with an alanine, a serine, or a threonine.

In embodiments, the amino acid sequence of a protein comprises no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 mutations compared to its naturally occurring counterpart. In some embodiments, less than 50, 45, 40, 35, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, or 2 of the mutations is a deletion or insertion of 1, 2, 3, 4, or 5 or no more than 1, 2, 3, 4, or 5 amino acids. In some embodiments, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or more of the mutations is a substitution mutation. In certain embodiments, every mutation to a protein compared to its naturally occurring counterpart is a substitution mutation. In various embodiments, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or more or all of the mutations to a protein compared to its naturally occurring counterpart is a conservative substitution mutation.

In various embodiments, a polypeptide does not have any insertion or deletion compared to its natural counterpart, other than (optionally) the removal of the signal peptide and/or the fusion of compounds such as another polypeptide at the N-terminus or C-terminus thereof.

Exemplary Ligand-Binding Proteins

Various biosensors provided herein comprise ligand-binding proteins, such as proteins that have altered amino acid sequences compared to their naturally occurring counterparts. In embodiments, such proteins are conjugated to reporter groups.

In various embodiments, the C_α root-mean-square deviation (RMSD) between the backbone of a ligand-binding protein and its naturally occurring counterpart is, *e.g.*, between about 0-3 Å, 0-1 Å, 0-1.5 Å, 0-2 Å, 0.1-3 Å, 0.5-1 Å, 0.5-1.5 Å, or 0.5-2 Å, or less than about 0.1 Å, 0.2 Å, 0.3 Å, 0.4 Å, 0.5 Å, 0.6 Å, 0.7 Å, 0.8 Å, 0.9 Å, 1.0 Å, 1.5 Å, 1.6 Å, 1.7 Å, 1.8

Å, 1.9 Å, 2.0 Å, 2.5 Å, or 3 Å. In some embodiments, the C_α RMSD between the N-terminal domain (*i.e.*, the portion of the protein at the N-terminal side of the binding domain hinge) backbone of the ligand-binding protein and the corresponding domain of its naturally occurring counterpart is, *e.g.*, between about 0-3 Å, 0-1 Å, 0-1.5 Å, 0-2 Å, 0.1-3 Å, 0.5-1 Å, 0.5-1.5 Å, or 0.5-2 Å, or less than about 0.1 Å, 0.2 Å, 0.3 Å, 0.4 Å, 0.5 Å, 0.6 Å, 0.7 Å, 0.8 Å, 0.9 Å, 1.0 Å, 1.5 Å, 1.6 Å, 1.7 Å, 1.8 Å, 1.9 Å, 2.0 Å, 2.5 Å, or 3 Å. In certain embodiments, the C_α RMSD between the C-terminal domain (*i.e.*, the portion of the protein at the C-terminal side of the binding domain hinge) backbone of the ligand-binding polypeptide and the corresponding domain of its naturally occurring counterpart is, *e.g.*, between about 0-3 Å, 0-1 Å, 0-1.5 Å, 0-2 Å, 0.1-3 Å, 0.5-1 Å, 0.5-1.5 Å, or 0.5-2 Å, or less than about 0.1 Å, 0.2 Å, 0.3 Å, 0.4 Å, 0.5 Å, 0.6 Å, 0.7 Å, 0.8 Å, 0.9 Å, 1.0 Å, 1.5 Å, 1.6 Å, 1.7 Å, 1.8 Å, 1.9 Å, 2.0 Å, 2.5 Å, or 3 Å. Non-limiting considerations relating to the sequence and structural differences between homologous proteins are discussed in Chothia and Lesk (1986) *The EMBO Journal*, 5(4):823-826, the entire content of which is incorporated herein by reference.

Non-limiting examples of ligand-binding polypeptides that are useful in biosensors provided herein include variants of the naturally occurring proteins disclosed herein that comprise cysteine substitutions and/or N-terminal and/or C-terminal fusions (*e.g.*, to a fluorophore attachment moiety). In embodiments, a biosensor comprises a modified ligand-binding protein having an amino acid substitution compared to its naturally occurring counterpart, such that the polypeptide has a cysteine at one or more of position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259,

260, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, or 400 or any combination of 1, 2, 3, 4, or 5 thereof, wherein the position corresponds a SEQ ID NO disclosed herein. In embodiments, the cysteine is conjugated to a reporter group.

In various embodiments, the dissociation constant of the mutant ligand-binding polypeptide differs by at least about 1 μ M, 5 μ M, 10 μ M, 20 μ M, 25 μ M, 30 μ M, 35 μ M, 40 μ M, 45 μ M, 50 μ M, 75 μ M, 100 μ M, 200 μ M, 300 μ M, 400 μ M, 500 μ M, 600 μ M, 700 μ M, 800 μ M, 900 μ M, 1 mM, 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 7 mM, 8 mM, 9 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, 60 mM, 70 mM, 80 mM, 90 mM, or 100 mM (increase or decrease) compared to its naturally occurring counterpart.

The biosensors and ligand-binding proteins provided herein are robust and useful at a wide range of physical conditions, *e.g.*, pressure, temperature, salinity, osmolality, and pH conditions. For example, biosensors and ligand-binding proteins provided herein may survive substantial periods of time after being dried or exposed to high temperatures. In some embodiments, the biosensor maintains at least about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more of its signal transduction activity after exposure to a temperature of about 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, or 125, or 40-125°C for about 1, 2, 3, 4, 5, 6, 15, 30, 60, 120, 180, 240, or 360 minutes. In certain embodiments, the biosensor maintains at least about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more of its signal transduction activity after 1, 2, 3, 4, or 5 freeze-thaw cycles in an aqueous solution. In various embodiments, the biosensor maintains at least about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more of its signal transduction activity after storage at a temperature of between 20-37°C for about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 18, 24, or 1-24 months in dry form. In some embodiments, the optimal functional temperature of the biosensor is between 41 and 122°C, between 20 and 40°C, or less than about 10°C (*e.g.*, between -20 and +10 °C). Devices, compositions, and biosensors provided herein may be stored, *e.g.*, with or without protection

from exposure to light. In some embodiments, the devices, compositions, and biosensors are stored in the dark, *e.g.*, with protection from light.

Non-limiting examples of glucose-binding proteins include variants of ecGGBP, ttGGBP, stGGBP, chyGGBP, cobGGBP, pspGGBP, csaGGBP, bprGGBP, rinGGBP_A,
 5 rinGGBP_B, fprGGBP, cljGGBP, cauGGBP, erhGGBP, ereGGBP, and chyGGBP.

In embodiments, a biosensor comprises a modified ecGGBP. In non-limiting examples, the modified ecGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Y10X, D14X, N15X, F16X, P70X, N91X, K92X, S112X, S115X, E149X, H152X, P153X, D154X, A155X,
 10 R158X, M182X, W183X, N211X, D212X, D236X, L238X, L255X, N256X, D257X, P294X, and V296X, where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in ecGGBP without including the signal peptide (SEQ ID NO: 153). In some embodiments, the modified ecGGBP comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 of the following substitutions: Y10A, Y10C,
 15 D14C, D14A, D14Q, D14N, D14S, D14T, D14E, D14H, D14L, D14Y, D14F, N15C, F16L, F16A, F16C, F16Y, N91C, N91A, K92A, K92C, E93C, S112A, S115A, E149C, E149K, E149Q, E149S, H152C, H152A, H152F, H152Q, H152N, D154C, D154A, D154N, A155C, A155S, A155H, A155L, A155F, A155Y, A155N, A155K, A155M, A155W, A155Q, R158C, R158A, R158K, M182C, M182W, W183C, W183A, N211C, N211F, N211W, N211K,
 20 N211Q, N211S, N211H, N211M, N211C, D212C, L238C, D236C, D236A, D236N, L255C, N256A, N256D, D257C, P294C, and V293C.

In various embodiments, a biosensor comprises a modified ttGGBP. In non-limiting examples, the modified ttGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Y11X, D15X, T16X,
 25 F17X, G20X, N42X, V67X, R69X, R91X, E92X, A111X, Q148X, H151X, Q152X, A154X, N181X, W182X, D183X, D211X, T237X, T240X, L257X, N258X, D259X, A260X, and K300X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in ttGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 154). In some embodiments, the modified ttGGBP
 30 comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 of the following substitutions: Y11C, D15A, D15E, D15N, D15C, T16S, T16N, T16C, F17C, G20A, G20C, N42C, V67C, R69P, R69C, R91K, E92C, A111C, Q148S, Q148K, Q148E, Q148C, H151Q, H151N, H151F, H151C, Q152P, Q152C, A154S, A154N, A154M, A154F, A154C, N181C, W182C, D183C, D211A,

D211C, T237C, T240A, L257C, N258D, N258S, N258A, N258C, D259C, A260N, A260Q, A260R, A260K, A260W, A260F, A260Y, A260S, A260C, and K300C.

In embodiments, a biosensor comprises a modified stGGBP. In non-limiting examples, the modified stGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Y11X, Y13X, D15X, N16X, F17X, P71X, N92X, K93X, P113X, S116X, E150X, H153X, P154X, D155X, A156X, R159X, M183X, W184X, N211X, N212X, D213X, A214X, D237X, L239X, D258X, P295X, and V297X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in stGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 155). In some embodiments, the modified stGGBP comprises 1, 2 or 3 of the following mutations: Y13C, F17C, and W184C.

In embodiments, a biosensor comprises a modified chyGGBP. In non-limiting examples, the modified chyGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: F12X, D14X, T15X, F16X, R68X, N89X, R90X, A110X, S113X, E147X, H150X, Q151X, D152X, A153X, R156X, M180X, W181X, N207X, N208X, D209X, D210X, D237X, T239X, D258X, V296X, and Y298X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in chyGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 156). In some embodiments, the modified chyGGBP comprises 1, 2, or 3 of the following mutations: F12C, F16C, C39A, W181C, and C206A.

In embodiments, a biosensor comprises a modified cobGGBP. In non-limiting examples, the modified cobGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: F12X, D14X, T15X, F16X, C39X, R68X, N89X, R90X, A110X, S113X, E147X, H150X, Q151X, D152X, A153X, R156X, C173X, M180X, W181X, C206X, N207X, N208X, D209X, D210X, D237X, T239X, D258X, P297X, and Q299X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in cobGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 157). In some embodiments, the modified cobGGBP comprises 1, 2, or 3 of the following mutations: F12C, F16C, C39A, C173A, W181C, and C206A.

In embodiments, a biosensor comprises a modified pspGGBP. In non-limiting examples, the modified pspGGBP may comprise one or more, or any combination of the

following substitutions compared to its naturally occurring counterpart: F9X, D11X, T12X, F13X, R65X, N86X, R87X, A107X, S110X, E144X, H147X, Q148X, D149X, A150X, R153X, M177X, W178X, N204X, N205X, D206X, D207X, D234X, T236X, 255X, A294X, and K296X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in pspGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 158). In some embodiments, the modified pspGGBP comprises 1, 2, or 3 of the following mutations: F9C, F13C, and W178C.

In embodiments, a biosensor comprises a modified csaGGBP. In non-limiting examples, the modified csaGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Y14X, D16X, F18X, C62X, I72X, C82X, N93X, R94X, C113A, S118X, A121X, E152X, N155X, E156X, D157X, S158X, R161X, N185X, W186X, C211X, D241X, L243X, D262X, D290X, I292X, I297X, F299X, Q301X, and T302X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in csaGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 159). In some embodiments, the modified csaGGBP comprises 1, 2, 3, 4, 5, 6, 7, or 8 of the following mutations: Y14C, F18C, C62A, C82A, C113A, W186C, and C211A.

In embodiments, a biosensor comprises a modified bprGGBP. In non-limiting examples, the modified bprGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: C8X, K12X, D14X, N15X, F16X, S72X, N93X, R94X, C112X, C116X, A118X, S121X, A153X, N156X, I157X, D158X, A159X, C179X, N186X, W187X, C211X, N212X, N213X, D214X, A215X, D241X, D243X, K251X, C289X, D290X, and V292X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in bprGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 160). In some embodiments, the modified bprGGBP comprises 1, 2, 3, 4, 5, 6, 7, 8, or 9 of the following mutations: C8A, K12C, F16C, C112A, C116A, C179A, W187C, C211A, and C289A.

In embodiments, a biosensor comprises a modified rinGGBP_A. In non-limiting examples, the modified rinGGBP_A may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: C6X, F10X, D12X, N13X, F14X, S70X, N91X, R92X, C114X, A116X, Q118X, D151X, N154X, V155X, D156X, A157X, R160X, C177X, N184X, W185X, C210X, N211X, N212X, D213X,

A214X, D240X, L242X, L250X, C288X, D289X, and V291X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in rinGGBP_A with the signal peptide replaced with a methionine (SEQ ID NO: 161). In some embodiments, the modified rinGGBP_A comprises 1, 2, 3, 4, 5, 6, 7, or 8 of the following mutations: C6A, F10C, F14C, C114A, C177A, W185C, C210A, and C288A.

In embodiments, a biosensor comprises a modified rinGGBP_B. In non-limiting examples, the modified rinGGBP_B may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Q13X, D15X, T16X, F17X, C66X, C70A, R76X, N97X, R98X, A118X, S121X, E155X, H158X, Q159X, D160X, A161X, R164X, N188X, W189X, N215X, N216X, D217X, D218X, D244X, T246X, D265X, P301X, A303X, and C306X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in rinGGBP_B with the signal peptide replaced with a methionine (SEQ ID NO: 165). In some embodiments, the modified rinGGBP_B comprises 1, 2, 3, 4, 5, or 6 of the following mutations: Q13C, F17C, C66A, C70A, W189C, and C306A.

In embodiments, a biosensor comprises a modified fprGGBP. In non-limiting examples, the modified fprGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: C8A, F12X, D14X, N15X, F16X, T69X, N90X, R91X, C105X, C106X, A113X, S116X, C143X, D146X, N149X, I150X, D151X, A152X, R155X, N179X, W180X, C205A, N206X, N207X, D208X, A209X, D235X, L237X, N243X, D284X, and V286X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in fprGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 162). In some embodiments, the modified fprGGBP comprises 1, 2, 3, 4, 5, 6, or 7 of the following mutations: C8A, F12C, F16C, C105A, C106A, C143A, W180C, and C205A.

In embodiments, a biosensor comprises a modified cljGGBP. In non-limiting examples, the modified cljGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: F11X, N13X, T14X, W15X, V67X, C77X, N88X, R89X, A109X, S112X, E142X, N145X, Q146X, D147X, A148X, R151X, M175X, W176X, C198X, N201X, N202X, D203X, D204X, D231X, T233X, D252X, D291X, and K294X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in cljGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 163). In some

embodiments, the modified cljGGBP comprises 1, 2, 3, 4, or 5 of the following mutations: F11C, W15C, C77A, W176C, and C198A.

In embodiments, a biosensor comprises a modified cauGGBP. In non-limiting examples, the modified cauGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: F12X, N14X, T15X, W16X, V68X, C78X, N89X, R90X, A110X, S113X, E143X, N146X, Q147X, D148X, A149X, R152X, M176X, W177X, C199X, N203X, N204X, D205X, D206X, D233X, T235X, D254X, D293X, and K295X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in cauGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 164). In some embodiments, the modified cauGGBP comprises 1, 2, 3, 4, or 5 of the following mutations: F12C, W16C, C78A, W177C, and C199A.

In embodiments, a biosensor comprises a modified erhGGBP. In non-limiting examples, the modified erhGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: F13X, D15X, N16X, F17X, P76X, N97X, R98X, A119X, S122X, D153X, N156X, V157X, D158X, A159X, R162X, N187X, W188X, N214X, N215X, D216X, G217X, D243X, I245X, D264X, E312X, and V314X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in erhGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 166). In some embodiments, the modified erhGGBP comprises 1, 2, or 3 of the following mutations: F13C, F17C, and W188C.

In embodiments, a biosensor comprises a modified ereGGBP. In non-limiting examples, the modified ereGGBP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: Q13X, D15X, T16X, F17X, C29X, C65X, C69X, R75X, N96X, R97, A117X, S120X, E154X, H157X, Q158X, D159X, A160X, R163X, C183X, N187X, W188X, N214X, N215X, D216X, A217X, D243X, T245X, D264X, P301X, and E303X where X is any amino acid, an amino acid that results in a conservative substitution, or a cysteine, and where each position is counted in ereGGBP with the signal peptide replaced with a methionine (SEQ ID NO: 167). In some embodiments, the modified ereGGBP comprises 1, 2, 3, 4, 5, 6, 7, or 8 of the following mutations: C10A, Q13C, F17C, C29A, C65A, C69A, C183A, and W188C.

Fluorescent Proteins

Various biosensors provided herein comprise fluorescent proteins, such as fluorescent proteins that have altered amino acid sequences compared to their naturally occurring counterparts. In embodiments, such proteins are conjugated to reporter groups. In
5 embodiments, the proteins are not conjugated to a reporter group (*i.e.*, a biosensor comprising the fluorescent protein that does not undergo tgmFRET or ngmFRET is provided).

Aspects of the present subject matter provide a biosensor for ligand, comprising a ligand-binding protein, wherein the ligand-binding protein is a fluorescent protein, and wherein binding of the ligand to a ligand-binding domain of the fluorescent protein causes a
10 change in fluorescence by the fluorescent protein. In some embodiments, the biosensor further comprises a reporter group, *e.g.*, a fluorophore that acts as a ngmFRET donor fluorophore or a ngmFRET acceptor fluorophore with respect to the fluorescent protein.

Green Fluorescent Protein (GFP) and its derivatives such as Yellow Fluorescent Protein (YFP) form their internal fluorophore through an autocatalytic, posttranslational
15 cyclization of a tripeptide from its own amino acid sequence (M. Zimmer, 2002, Chem. Rev. 102, 759-781). This process entails three steps: a nucleophilic attack to create a cyclic peptide, dehydration, and a final oxidation to introduce conjugation (D.P. Barondeau et al., 2003, Proc. Natl. Acad. Sci. USA, 100, 12111-12116). The formation of GFP's or YFP's fluorophore is an autocatalytic process that requires no catalyst external to these proteins.

In some embodiments, the ligand comprises a halide anion such as a fluoride (F^-),
20 chloride (Cl^-), a bromide (Br^-), an iodide (I^-), astatide (At^-) anion, or an ununseptide (Ts^-) anion. In certain embodiments, the fluorescent protein has an affinity (K_d) for the halide anion that is within the concentration range of the halide anion in a subject. Non-limiting examples of fluorescent proteins that bind halide anions include Yellow Fluorescent Protein
25 (YFP; SEQ ID NO: 149) and mutants thereof. In some embodiments, the fluorescent protein comprises a mutation that alters the interaction of the mutant with a bound halide anion compared to YFP. For example, the mutation that may alter the fluorescent protein's affinity and/or specificity for a halide anion compared to YFP. In various embodiments, the fluorescent protein comprises 1 halide anion binding site. Alternatively the fluorescent
30 protein comprises at least 2, 3, 4, or 5 halide anion binding sites. In some embodiments, at least one amino acid of the YFP has been substituted with a cysteine.

YFP is a non-limiting reference protein respect to fluorescence proteins.

In various embodiments, a polypeptide of the present disclosure comprises

(a) amino acids in a sequence that is preferably (i) at least about 10%, 11%, 12%, 13%, 14%, or 15%, and (ii) less than about 75%, 70%, 65%, 60%, 55%, 50%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, or 35% identical to YFP;

5 (b) a stretch of at least 5, 10, or 20 amino acids having at least about 50%, 55%, 60%, 65%, 75%, 80%, 85%, 90%, or 95% identity to a stretch of consecutive amino acids including position 17, 32, 43, 77, 95, 109, 122, 133, 149, 164, 173, 182, 204, and/or 221 of YFP;

10 (c) no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 15 deleted or inserted amino acids compared to YFP, not including added amino acids added to the N-terminus or C-terminus of the polypeptide compared to its natural counterpart, and including or not including the signal peptide of the natural counterpart of the polypeptide;

(d) at least 9, 10, 11, or 12 β -strands or exactly 9, 10, 11, or 12 β -strands.

In some embodiments, the fluorescent polypeptide comprises a β -barrel. In certain
 15 embodiments, the β -barrel comprises 9, 10, 11, or 12 β -strands. In some embodiments, the fluorescent protein comprises a cysteine within the first β -strand (β_1), the second β -strand (β_2), the third β -strand (β_3), the fourth β -strand (β_4), the fifth β -strand (β_5), the sixth β -strand (β_6), the seventh β -strand (β_7), the eighth β -strand (β_8), the ninth β -strand (β_9), the tenth β -strand (β_{10}), or the eleventh β -strand (β_{11}) of a YFP. In some embodiments, the polypeptide
 20 comprises (i) 1, 2, or 3 amino acid substitutions between β_1 and β_2 ; (ii) 1, 2, or 3 amino acid substitutions between β_2 and β_3 ; (iii) 1, 2, or 3 amino acid substitutions between the β_3 and β_4 ; (iv) 1, 2, or 3 amino acid substitutions between the β_4 and β_5 ; (v) 1, 2, or 3 amino acid substitutions between β_5 and β_6 ; (vi) 1, 2, or 3 amino acid substitutions between β_6 and β_7 ; (vii) 1, 2, or 3 amino acid substitutions between the β_7 and β_8 ; (viii) 1, 2, or 3, amino acid
 25 substitutions between β_8 and β_9 ; (ix) 1, 2, or 3 amino acid substitutions between the β_9 and β_{10} ; and/or (x) 1, 2, or 3 amino acid substitutions between β_{10} and β_{11} . In various embodiments, the 1 or more substitutions comprise a substitution with cysteine. In certain embodiments, the cysteine follows β_{11} in the amino acid sequence of the YFP.

Alpha-helical and β -strand segments assignments are calculated from a three-
 30 dimensional protein structure as follows, and as described in C.A.F. Andersen, B. Rost, 2003, *Structural Bioinformatics*, 341-363, P.E. Bourne, ed., Wiley, the entire content of which is incorporated herein by reference. First for a given residue, i , the backbone trace angle, τ , is calculated, defined as the dihedral angle between the four successive C_α atom positions of

residues in the linear protein sequence i , $i+1$, $i+2$, $i+3$. These values are calculated for all residues. Second, the residues that form backbone hydrogen bonds with each other are recorded. A hydrogen bond is scored if the distance between the backbone amide nitrogen and carbonyl oxygen of two different residues in the protein is calculated to be 2.5Å or less, and if the calculated angle between the nitrogen, its amide proton, and the carbonyl is greater than 120°. A residue is deemed to be in an α -helix, if $35 \leq \tau \leq 65$, and it makes a backbone hydrogen bond with its $i+4^{\text{th}}$ neighbor in the linear amino acid sequence. It is deemed to be in a β -strand, if the absolute τ value falls in the interval $120 \leq |\tau| \leq 180$ and if it makes at least one hydrogen bond with another residue with the same τ value range. Alpha-helical segments comprise at least four residues; β -strand residues comprise at least three residues.

In embodiments, a biosensor comprises a modified YFP polypeptide having an amino acid substitution compared to its naturally occurring counterpart, such that the polypeptide has a cysteine at position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, or 250, or any combination of 1, 2, 3, 4, or 5 thereof, wherein the position corresponds a SEQ ID NO disclosed herein for YFP. In embodiments, the cysteine is conjugated to a reporter group.

In embodiments, a biosensor comprises a modified YFP. In non-limiting examples, the modified YFP may comprise one or more, or any combination of the following substitutions compared to its naturally occurring counterpart: E17X, E32X, T43X, F64X, G65X, L68X, Q69X, A72X, H77X, K79X, R80X, E95X, R109X, R122X, D133X, H148X, N149X, V163X, N164X, D173X, Y182X, Q183X, Y203X, Q204X, L221X, and H231X, where X is any amino acid or is an amino acid that results in a conservative substitution. In

some embodiments X is cysteine. In some embodiments, the modified YFP comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 of the following substitutions: F64L, G65T, L68V, Q69T, A72S, K79R, R80Q, H148Q, H148G, V163A, H231L, H148Q, or Q183A, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150. In some embodiments, the
 5 fluorescent protein comprises an R at the 96 position, a Y at the 203 position, a S at the 205 position, and an E at the 222 position compared to YFP, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150. ClBP1 (also referred to as laYFP) comprises L68V, K79R, R80Q, H2131L compared to YFP (as numbered in SEQ ID NO: 150).

10 Adaptor Proteins

Aspects provide biosensor for a ligand, comprising (a) a polypeptide; (b) a directly responsive fluorophore, wherein binding of a ligand to the directly responsive fluorophore causes a change in signaling by the directly responsive fluorophore (*i.e.*, the fluorophore is chemoresponsive); and (b) an indirectly responsive fluorophore. The directly responsive
 15 fluorophore may be a donor fluorophore or an acceptor fluorophore. In some embodiments, the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore. In some embodiments, the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore. ngmFRET occurs between the donor fluorophore and the acceptor fluorophore
 20 when the donor fluorophore is contacted with radiation comprising the excitation wavelength of the donor fluorophore.

Any polypeptide may be used to link a directly responsive fluorophore (*e.g.*, a chemoresponsive fluorophore) with an indirectly responsive fluorophore. Such a polypeptide is referred to as an adaptor protein herein. In some embodiments, the polypeptide comprises
 25 a stretch of at least 2, 3, 4, 5, 6, 47, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, or 500 amino acids.

In some embodiments, the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of ecTRX (SEQ ID NO: 151). In certain embodiments, the
 30 polypeptide comprises at least 1, 2, or 3 thiol groups; at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group; at least 1, 2, or 3 primary amine groups; or at least 1, 2, or 3 lysines that each comprise a primary amine. In various embodiments, there is no disulfide bond between cysteines within the amino acid sequence of the polypeptide.

In some embodiments, the polypeptide comprises a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

5 In various embodiments, the polypeptide comprises a mutant of ecTRX comprising a D3A, D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151. In certain embodiments, the polypeptide comprises a mutant of ecTRX that does not comprise a lysine.

10 In embodiments, the polypeptide further comprises a hexahistidine tag. In some embodiments the polypeptide comprises amino acids in the sequence of any one of SEQ ID NOS:24-41 or 151.

In various embodiments, the biosensor is a pH biosensor and the ligand comprises a hydrogen ion. In embodiments, the directly responsive fluorophore is pH-sensitive. For
15 example, the fully excited emission intensity of the directly responsive fluorophore is different at a pH less than about 7.0 compared to a pH of 7.5. In certain embodiments, the directly responsive fluorophore transitions from a monoanion to a dianion at a pH that is less than 7.0 in an aqueous solution. In some embodiments, the directly responsive fluorophore comprises a pH-sensitive fluorophore comprising fluorescein or a derivative thereof.

20 Exemplary Methods of Using Biosensors Provided Herein

Aspects of the present subject matter provide a method of assaying for a ligand in a sample. The method may include contacting the sample with a biosensor disclosed herein under conditions such that the ligand-binding protein of the biosensor binds to the ligand if
25 ligand is present in the sample. The method also comprises detecting (i) whether a signal is produced by a reporter group of the biosensor; and/or (ii) the a signal produced by a reporter group of the biosensor. In a non-limiting example, a reporter group of the biosensor is fluorescent, and the method further comprises contacting the reporter group with electromagnetic radiation having a wavelength that comprises a wavelength within the band
30 of excitation wavelengths of the reporter group.

In various embodiments, the method further comprises (i) comparing a signal produced by a reporter group of the biosensor when the biosensor is contacted with the sample with a signal produced by a control sample containing a known quantity of ligand;

and (ii) detecting the presence or absence of ligand in the sample based on this comparison. Alternatively or in addition, the method further comprises (i) comparing a signal produced by a reporter group of the biosensor when the biosensor is contacted with the sample with signals produced by a series of control samples containing known quantities of ligand; and
5 (ii) determining the quantity of ligand in the sample based on this comparison. In some embodiments, the series of control samples comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 control samples, and wherein each control sample comprises a different quantity of ligand. Alternatively or in addition, the method further comprises determining the concentration of a ligand in a sample, wherein determining the concentration of the ligand in the sample
10 comprises comparing the signal to a standard hyperbolic ligand binding curve to determine the concentration of the ligand in the test sample, wherein the standard hyperbolic ligand binding curve is prepared by measuring the signal produced by the reporter group of the biosensor when the biosensor is contacted with control samples containing known concentrations of ligand. In various embodiments, the method comprises (i) measuring a
15 ratiometric change (ΔR) and/or an intensity change (ΔI) of a signal produced by the reporter group. In some embodiments, the method includes quantitating the level of ligand present in the sample.

Aspects of the present subject matter also provide a method of assaying for multiple ligands in a sample, wherein the multiple ligands comprise a first ligand and a second ligand.
20 Such a method may include contacting the sample with (i) a first biosensor a first ligand provided herein and (ii) a second biosensor for the second ligand, under conditions such that the ligand-binding protein of the first biosensor binds to the first ligand, if the first ligand is present in the sample, and detecting (i) a signal produced by a reporter group of the first biosensor, or (ii) whether a signal is produced by a reporter group of the first biosensor. In
25 some embodiments, the second biosensor is also a biosensor provided herein, and the second biosensor is contacted with the second ligand under conditions such that the ligand-binding protein of the second biosensor binds to the second ligand it is present in the sample. The method may further comprise detecting (i) a signal produced by a reporter group of the second biosensor, or (ii) whether a signal is produced by a reporter group of the second
30 biosensor.

In some embodiments, the signal produced by the reporter group of the first biosensor is different than the signal produced by the reporter group of the second biosensor. In a non-limiting example, the reporter group of the first biosensor and the reporter group of the

second biosensor are each fluorescent, and the peak emission wavelength of the reporter group of the first biosensor is at least about 10, 25, 50, 75, or 100 nm greater or lower than the peak emission wavelength of the reporter group of the second biosensor.

Non-limiting examples of biosensors include biosensors with ligand-binding proteins comprising a GGBP (*e.g.*, an *E. coli* GGBP) or a derivative or mutant thereof; (ii) an *E. coli* arabinose binding protein (*e.g.*, an *E. coli* arabinose binding protein) or a derivative or mutant thereof; (iii) a dipeptide binding protein (*e.g.*, an *E. coli* dipeptide binding protein) or a derivative or mutant thereof; (iv) a histidine binding protein (*e.g.*, an *E. coli*, histidine binding protein) or a derivative or mutant thereof; (v) a ribose binding protein (*e.g.*, an *E. coli* ribose binding protein) or a derivative or mutant thereof; (vi) a sulfate binding protein (*e.g.*, an *E. coli* sulfate binding protein) or a derivative or mutant thereof; (vii) a maltose binding protein (*e.g.*, an *E. coli* maltose binding protein) or a derivative or mutant thereof; (viii) a glutamine binding protein (*e.g.*, an *E. coli* glutamine binding protein) or a derivative or mutant thereof; (ix) a glutamate/aspartate binding protein (*e.g.*, an *E. coli* glutamate/aspartate binding protein) or a derivative or mutant thereof; (x) a phosphate binding protein (*e.g.*, an *E. coli* phosphate binding protein) or a derivative or mutant thereof; or (xi) an iron binding protein [*e.g.*, a *Haemophilus influenza* (*H. influenzae*) iron binding protein] or a derivative or mutant thereof. For example, the second biosensor comprises an *E. coli* GGBP having a Y10A, Y10C, D14C, D14A, D14Q, D14N, D14S, D14T, D14E, D14H, D14L, D14Y, D14F, N15C, F16L, F16A, F16C, F16Y, N91C, N91A, K92A, K92C, E93C, S112A, S115A, E149C, E149K, E149Q, E149S, H152C, H152A, H152F, H152Q, H152N, D154C, D154A, D154N, A155C, A155S, A155H, A155L, A155F, A155Y, A155N, A155K, A155M, A155W, A155Q, R158C, R158A, R158K, M182C, M182W, W183C, W183A, N211C, N211F, N211W, N211K, N211Q, N211S, N211H, N211M, N211C, D212C, L238C, D236C, D236A, D236N, L255C, N256A, N256D, D257C, P294C, and V293C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or more of these mutations), wherein each amino acid position is numbered as in (SEQ ID NO: 153); (ii) an *E. coli* arabinose binding protein having a D257C, F23C, K301C, L253C, or L298C mutation (*e.g.*, comprising 1, 2, 3, 4, or 5 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681, the entire contents of which are incorporated herein by reference) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681, the entire contents of which are incorporated herein by reference); (iii) an *E. coli* dipeptide binding protein having a D450C, K394C, R141C, S111C, T44C, or W315C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or 6 of these mutations) (*see, e.g.*, U.S. Patent Application Publication

No. 2004/0118681, the entire contents of which are incorporated herein by reference); (iv) an *E. coli*, histidine binding protein having a E167C, K229C, V163C, Y230C, F231C, Y88C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or 6 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681, the entire contents of which are incorporated
5 herein by reference); (v) an *E. coli* ribose binding protein having a T135C, D165C, E192C, A234C, L236C, or L265C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or 6 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681, the entire contents of which are incorporated herein by reference); (vi) an *E. coli* sulfate binding protein having a L65C, N70C, Q294C, R134C, W290C, or Y67C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or 6
10 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated herein by reference); (vii) an *E. coli* maltose binding protein having a D95C, F92C, E163C, G174C, I329C, or S233C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or 6 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated herein by reference); (viii) an *E.*
15 *coli* glutamine binding protein having a N160C, F221C, K219C, L162C, W220C, Y163C, or Y86C mutation (*e.g.*, comprising 1, 2, 3, 4, 5 or more of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated herein by reference); (ix) an *E. coli* glutamate/aspartate binding protein having a A207C, A210C, E119C, F126C, F131C, F270C, G211C, K268C, Q123C, or T129C mutation (*e.g.*,
20 comprising 1, 2, 3, 4, 5 or more of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated herein by reference); (x) an *E. coli* phosphate binding protein having a A225C, N223C, N226C, S164C, or S39C mutation (*e.g.*, comprising 1, 2, 3, 4, or 5 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated
25 herein by reference); or (xi) a *Haemophilus influenza* (*H. influenzae*) iron binding protein having a E203C, K202C, K85C, or V287C mutation (*e.g.*, comprising 1, 2, 3, or 4 of these mutations) (*see, e.g.*, U.S. Patent Application Publication No. 2004/0118681 the entire content of which is incorporated herein by reference). In various embodiments, the sample is suspected of comprising a ligand, such as a ligand disclosed, described, or otherwise
30 mentioned herein.

References and PDB^a files for bPBP structures, genes, and ligand binding

bPBP	crystal structure		DNA sequence	ligand affinity
	open form	closed form		
arabinose BP		Quioco and Vyas, 1984 1ABE	Scripture et al., 1987	Clark et al., 1982; Miller et al., 1983
dipeptide BP	Nickitenko et al., 1995 1DPE	Dunten & Mowbray, 1995 1DPP	Abouhamad et al., 1991	Guyer et al., 1986; Smith et al., 1999
Glu/Asp BP				Barash Halpern, 1975; Willis Furlong, 1975
Fe(III) BP	Bruns et al., 2001 1D9V	Bruns et al., 1997 1MRP	Sanders et al., 1994	Adhikari et al., 1995
glucose BP		Vyas et al., 1988; Vyas et al., 1994 1GLG	Scholle et al., 1987	Anraku, 1968
histidine BP		Yao et al., 1994 1HSL	Joshi & Ames 1996	Miller et al., 1983
maltose BP	Sharff et al., 1992 1OMP	Spurlino et al., 1991; Quioco et al., 1997 1ANF	Duplay et al., 1984	Schwartz et al., 1976
phosphate BP	Ledvina et al., 1996 1OIB	Luecke & Quioco, 1990 1IXH	Mageta et al., 1984	Medveczky & Rosenberg, 1969
glutamine BP	Hsiao et al., 1996 1GGG	Sun et al., 1998 1WDN	Nohno et al., 1986	Weiner et al., 1971
ribose BP	Bjorkman & Mowbray, 1998 1URP	Mowbray & Cole, 1992 2DRI	Groarke et al., 1983	Willis & Furlong, 1974
sulfate BP		Pflugrath & Quioco, 1985; He & Quioco, 1993 1SBP	Hellings & Evans, 1985	Jacobsen & Quioco, 1988

^aProtein Data Bank (Berman et al., 2000)

Abouhamad et al., Molec. Microbiol. 5: 1035–1047 (1991)

Adhikari et al., J. Biol. Chem. 270: 25142–25149 (1995)

Anraku, J. Biol. Chem. 243: 3116–3122 (1968)

Barash & Halpern, Biochim. Biophys. Acta 386: 168–180 (1975)

- Bjorkman & Mowbray, *J. Mol. Biol.* 279: 651–664 (1998)
- Bruns et al., *Biochemistry* 40: 15631–15637 (2001)
- Bruns et al., *Nat. Struct. Biol.* 4: 919–924 (1997)
- Clark et al., *Biochemistry* 21: 2227–2233 (1982)
- Dunten & Mowbray, *Protein Sci.* 4: 2327–2334 (1995)
- Duplay et al., *J. Biol. Chem.* 259: 10606–10613 (1984)
- Groarke et al., *J. Biol. Chem.* 258: 12952–12956 (1983)
- Guyer et al., *J. Bacteriol.* 168: 775–779 (1986)
- He & Quijcho, *Protein Sci.* 2: 1643–1647 (1993)
- Hellings & Evans, *Eur. J. Biochem.* 149: 363–373 (1985)
- Hsiao et al., *J. Mol. Biol.* 262: 225–242 (1996)
- Jacobson & Quijcho, *J. Mol. Biol.* 204: 783–787 (1988)
- Joshi & Ames, *GenBank Accession Number* U47027 (1996)
- Ledvina et al., *Proc. Natl. Acad. Sci. USA* 93: 6786–6791 (1996)
- Luecke & Quijcho, *Nature* 347: 402–406 (1990)
- Magota et al., *J. Bacteriol.* 157: 909–917 (1984)
- Medveczky & Rosenberg, *Biochim. Biophys. Acta* 192: 369–371 (1969)
- Miller et al., *J. Biol. Chem.* 258: 13665–13672 (1983)
- Mowbray & Cole, *J. Mol. Biol.* 225: 155–175 (1992)
- Nickitenko et al., *Biochemistry* 34: 16585–16595 (1995)
- Nohno et al., *Molec. Gen. Genet.* 205: 260–269 (1986)
- Pflugrath & Quijcho, *Nature* 314: 257–260 (1985)
- Quijcho et al., *Structure* 5: 997–1015 (1997)
- Quijcho & Vyas, *Nature* 310: 381–386 (1984)
- Sanders et al., *Infect. Immun.* 62: 4515–4525 (1994)
- Scholle et al., *Molec. Gen. Genet.* 208: 247–253 (1987)
- Scripture et al., *J. Mol. Biol.* 197: 37–46 (1987)
- Schwartz et al., *Eur. J. Biochem.* 71: 167–170 (1976)
- Sharff et al., *Biochemistry* 31: 10657–10663 (1992)
- Smith et al., *Microbiology* 145: 2891–2901 (1999)
- Spurlino et al., *J. Biol. Chem.* 266: 5202–5219 (1991)
- Sun et al., *J. Mol. Biol.* 278: 219–229 (1998)
- Vyas et al., *Biochemistry* 33: 4762–4768 (1994)
- Vyas et al., *Science* 242: 1290–1295 (1988)
- Weiner et al., *Arch. Biochem. Biophys.* 142: 715–717 (1971)
- Willis & Furlong, *J. Biol. Chem.* 249: 6926–6929 (1974)
- Willis & Furlong, *J. Biol. Chem.* 250: 2574–2580 (1975)
- Yao et al., *Biochemistry* 33: 4769–4779 (1994)

Various types of samples may be used in methods provided herein. In non-limiting examples, a sample may comprise a reaction product, a buffer, and/or a solvent. In some embodiments, the solvent is an aqueous solvent. In some embodiments, the solvent comprises a non-polar solvent, a polar aprotic solvent, and/or a polar protic solvent. For example, a sample may comprise water, liquid ammonia, liquid sulfur dioxide, sulfuryl chloride, sulfuryl chloride fluoride, phosphoryl chloride, dinitrogen tetroxide, antimony trichloride, bromine pentafluoride, hydrogen fluoride, dimethyl sulfoxide, hexane, benzene, toluene, 1,4-dioxane, chlorogorm, diethyl ether, dichloromethane, N-methylpyrrolidone, tetrahydrofuran, ethyl acetate, acetone, dimethylformamide, acetonitrile, formic acid, n-butanol, isopropanol, nitromethane, ethanol, methanol, and/or acetic acid.

In embodiments, a sample comprises a Newtonian liquid, a shear thickening liquid, a shear thinning liquid, a thixotropic liquid, a rheopectic liquid, or a Bingham plastic. In some implementations, a sample has a dynamic viscosity of at least about 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, or 2 pascal-seconds (Pa·s) or less than about 2, 1.5, 1.4, 1.3, 1.2, 1.1, 1, 0.9, 0.8, 0.7, 0.6, 0.5 Pa·s; and/or a kinematic viscosity of at least about 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, or 2 centistokes (cSt) or less than about 2, 1.5, 1.4, 1.3, 1.2, 1.1, 1, 0.9, 0.8, 0.7, 0.6, 0.5 cSt.

In various embodiments, the sample comprises a biological sample. The sample may comprise, *e.g.*, a clinical sample (*i.e.*, a sample collected in a clinical or veterinary setting, *e.g.*, by or at the request or supervision or direction of a doctor, nurse, aid worker, or medic) and/or a physiological sample (a sample collected from an organism, *e.g.*, a mammal such as a human). In certain embodiments, the biological sample comprises or has been provided or obtained from a skin surface or a mucosal surface. In some embodiments, the biological sample comprises a biological fluid. Non-limiting examples of biological fluids include sweat, tear fluid, blood, serum, plasma, interstitial fluid, amniotic fluid, sputum, gastric lavage, skin oil, milk, fecal matter, emesis, bile, saliva, urine, mucous, semen, lymph, spinal fluid, synovial fluid, a cell lysate, venom, hemolymph, and fluid obtained from plants such as the fluid transported in xylem cells or phloem sieve tube elements of a plant (*e.g.* sap).

The present subject matter also provides biosensors, methods, compositions, and devices useful for measuring the level of a ligand within a liquid solution or suspension or composition comprising cultured cells or tissue or a supernatant of such a solution or suspension, *e.g.*, a sample of conditioned media or a sample of growth media in which a population of cells was cultured. In some embodiments, the sample is within a culture (*e.g.*,

inserted into a bioreactor) or provided from a media, culture, or reaction, *e.g.*, in a bioreactor.

For example, the sample may be within or provided from a fermenter such as a culture or culture supernatant from a fermentation reaction (*e.g.*, an ongoing fermentation). Thus, the level of a ligand can be assayed at a timepoint of interest or at a series of timepoints over the

5 duration of cell culture, *e.g.* continuously, in or from a reaction or culture. Bioreactors

include devices or systems that support a biologically active environment. For example, a bioreactor may comprise a vessel in which a chemical process is carried out which involves organisms or biochemically active substances derived from such organisms. Such a process can either be aerobic or anaerobic. Organisms growing in bioreactors may be, *e.g.*,

10 submerged or suspended in liquid medium or may be attached to the surface of a solid medium. Submerged cultures may be suspended or immobilized. Suspension bioreactors can use a wider variety of organisms, since special attachment surfaces are not needed, and can operate at much larger scale than immobilized cultures. However, in a continuously operated process the organisms will be removed from the reactor with the effluent. Immobilization is

15 a general term describing a wide variety of cell or particle attachment or entrapment. It can be applied to basically all types of biocatalysis including enzymes, cellular organelles, and cells (*e.g.*, animal cells, plant cells, fungal cells, and bacterial cells). Immobilization is useful for continuously operated processes, since the organisms will not be removed with the reactor effluent, but is limited in scale because the cells are only present on the surfaces of the vessel.

20 A bioreactor may also refer to a device or system meant to grow cells or tissues in the context of cell culture. The interrogation and/or monitoring of ligand levels in such samples permits the evaluation of the status of growth of the cells or production of secreted products by the cells to inform harvest or feeding or other modification of the culture.

Aspects of the present subject matter relate to the use of methods and biosensors
25 provided herein to detect contamination.

In some embodiments, the sample comprises an environmental sample. Depending on context, there are instances in which a biological sample may also be, or may be within, an environmental sample. In certain embodiments, an environmental sample comprises a solute obtained from a biological composition, such as bone, nail, hair, shell, or cartilage. In
30 various embodiments, an environmental sample comprises a solute obtained from an environmental substance and/or an environmental surface. For example, the solute may be dissolved/obtained from the environmental substance and/or an environmental surface using an aqueous or nonaqueous solution. In some embodiments, an aqueous may optionally

comprise a nonaqueous solvent (*e.g.*, mixed with an aqueous solvent). Non-limiting examples of environmental substances include rock, soil, clay, sand, meteorites, asteroids, dust, plastic, metal, mineral, fossils, sediment, and wood. Non-limiting examples of environmental surfaces include the surface of a vehicle such as a civilian vehicle (*e.g.*, a satellite, a bike, a rocket, an automobile, a truck, a motorcycle, a yacht, a bus, or a plane) or a military vehicle (*e.g.*, a tank, an armored personnel carrier, a transport truck, a jeep, a mobile artillery unit, a mobile antiaircraft unit, a minesweeper, a Mine-Resistant Ambush Protected (MRAP) vehicle, a lightweight tactical all-terrain vehicle, a high mobility multipurpose wheeled vehicle, a mobile multiple rocket launch system, an amphibious landing vehicle, a ship, a hovercraft, a submarine, a transport plane, a fighter jet, a helicopter, a rocket, or an Unmanned Aerial Vehicle), a drone, a robot, a building, furniture, or an organism other than a human. In some embodiments, the sample comprises an environmental fluid. Non-limiting examples of environmental fluids include marine water, well water, drinking well water, water at the bottom of well dug for petroleum extraction or exploration, melted ice water, pond water, aquarium water, pool water, lake water, mud, stream water, river water, brook water, waste water, treated waste water, reservoir water, rain water, and ground water. In some embodiments, waste water comprises sewage water, septic tank water, agricultural runoff, water from an area in which chemical or oil spill has or is suspected of having occurred (*e.g.*, an oil spill into a marine environment), water from an area where a radiation leak has or is suspected of having occurred (*e.g.*, coolant from a nuclear reactor), water within the plumbing of a building, water within or exiting a research facility, and/or water within or exiting a manufacturing facility such as a factory.

As used herein, “suspected” with respect to an event means that there has been at least one test (*e.g.*, a test other than a method or assay provided herein), occurrence (*e.g.*, that is likely to or that may cause the event such as an emergency, leak, accident, flood, earthquake, storm, fire, malfunction, sunk vessel, or crash), or report (*e.g.*, by a witness, informant, or observer) that is consistent with the event having occurred.

In certain embodiments, the sample comprises a food or beverage additive and/or a food or beverage composition. In some embodiments, the food or beverage composition comprises a fermented composition. In various embodiments, the sample comprises a fluid obtained from a food composition. Alternatively or in addition, the sample may comprise a solute dissolved from a food composition. In some examples, a solute is or has been dissolved from a food composition with an aqueous or nonaqueous solution. In various

implementations, an aqueous solution may optionally comprise a nonaqueous solvent. In certain embodiments, a sample comprises a food composition in semisolid or liquid form. Non-limiting examples of such compositions include yogurt, soup, ice cream, a broth, a puree, a shake, a smoothie, a batter, a condiment, a sauce, and any combination thereof. In some implementations, a sample is a food engineering process (*e.g.*, obtained from a food design, storage, transport, or production process or from equipment intended to process, transport, or store food). A food composition may comprise, *e.g.*, a plant or a composition isolated from a plant, and/or an animal or a composition isolated from an animal. In various embodiments, a sample comprises a beverage composition. Non-limiting examples of beverage compositions include soft drinks, fountain beverages, water, coffee, tea, milk, dairy-based beverages, soy-based beverages (*e.g.*, soy milk), almond-based beverages (*e.g.*, almond milk), vegetable juice, fruit juice, fruit juice-flavored drinks, energy drinks, sports and fitness drinks, alcoholic products, and beverages comprising any combination thereof. Non-limiting examples of beverage compositions comprising water include purified water (*e.g.*, filtered water, distilled water, or water purified by reverse osmosis), flavored water, mineral water, spring water, sparkling water, tonic water, and any combination thereof. In various embodiments, the sample comprises alcohol. Non-limiting examples of such samples include samples comprising or obtained/provided from beer, malt beverages, liqueur, wine, spirits, and any combination thereof.

In some embodiments, a sample comprises a nutritional or supplement composition. In certain implementations, the nutritional or supplement composition comprises an omega-3 fatty acid, a vitamin, a mineral, a protein powder, or a meal supplement.

In certain embodiments, a biosensor is implanted in a subject's body. For example, a biosensor may be implanted in a subject's blood vessel, vein, eye, natural or artificial pancreas, alimentary canal, stomach, intestine, esophagus, or skin (*e.g.*, within the skin or under the skin). In various embodiments, the biosensor is configured within or on the surface of a contact lens. In some embodiments, the biosensor is configured to be implanted in or under the skin. In non-limiting examples, the biosensor is implanted in a subject with an optode and/or a microbead. In certain embodiments, the biosensor generates a signal transdermally.

Aspects of the present subject matter provide a method for assaying the level of ligand in a subject. The method may comprise contacting a biological sample from the subject with a biosensor for ligand under conditions such that the biosensor binds to ligand present in the

biological sample. The biosensor comprises reporter group that is attached to a ligand binding protein, and binding of ligand to a ligand-binding domain of the ligand binding protein causes a change in signaling by the reporter group.

In various embodiments, the subject has or is suspected of having a disease or disorder, such as abnormal kidney function, abnormal adrenal gland function, diabetes, hypochloremia, bromism, hypothyroidism, hyperthyroidism, cretinism, depression, fatigue, obesity, a low basal body temperature, a goiter, a fibrocystic breast change, lactic acidosis, septic shock, carbon monoxide poisoning, asthma, a lung disease, respiratory insufficiency, Chronic Obstructive Pulmonary Disease (COPD), regional hypoperfusion, ischemia, severe anemia, cardiac arrest, heart failure, a tissue injury, thrombosis, or a metabolic disorder, diarrhea, shock, ethylene glycol poisoning, methanol poisoning, diabetic ketoacidosis, hypertension, Cushing syndrome, liver failure, cancer, or an infection.

As used herein, “suspected” with respect to a subject’s condition (e.g., disease or injury) means that the subject has at least one symptom or test (e.g., a test other than an assay or method provided herein) that is consistent with the condition.

In some embodiments, the biological sample comprises blood, plasma, serum, sweat, tear fluid, or urine. In certain embodiments, the biological sample is present in or on the surface of the subject. In various implementations, the biosensor is applied onto or inserted into the subject. For example, the biosensor may be tattooed into the subject or is in or on a device that is implanted into the subject. In some embodiments, the biosensor may be present in or on a contact lens that is worn by the subject. Methods for determining the level of ligand in a subject who has or is suspected of having a disease that results in or from (or otherwise involves) an altered level of a ligand, may be performed without other testing related to the disease performed as part of a battery of clinical testing.

The present subject matter includes a method for monitoring the level of a ligand, comprising periodically or continuously detecting the level of the ligand, wherein detecting the level of the ligand comprises (a) providing or obtaining a sample; (b) contacting the sample with a biosensor for the ligand under conditions such that the ligand-binding protein of the biosensor binds to the ligand, and (c) detecting a signal produced by the biosensor.

Aspects of the present subject matter also provide a method for monitoring the level of a ligand in a subject, comprising periodically detecting the level of the ligand in the subject. Detecting the level of the ligand in the subject may comprise (a) providing or obtaining a biological sample from the subject; (b) contacting the biological sample with a

biosensor for the ligand provided herein under conditions such that the ligand-binding protein of the biosensor binds to the ligand, if the ligand is present in the biological sample, and (c) detecting (i) a signal produced by a reporter group of the biosensor, or (ii) whether a signal is produced by a reporter group of the biosensor. The level of the ligand may be detected, *e.g.*,
5 at least once every 1, 2, 3, 6, or 12 hours, at least once every 1, 2, 3, or 4 days, at least once every 1, 2, or three weeks, or at least once every 1, 2, 3, 4, 6, or 12 months.

The present subject matter also provides a method for monitoring the level of a ligand in a subject. The method comprises (a) administering a biosensor provided herein or a device comprising a biosensor provided herein to the subject, wherein after administration the
10 biosensor is in contact with a bodily fluid or surface that typically comprises the ligand, and (b) detecting (i) a signal produced by a reporter group of the biosensor continuously or repeatedly at intervals less than about 30 minutes (m), 15m, 10m, 5m, 1m, 30 seconds (s), 15s, 10s, 5s, 1s, 0.1s, 0.001s, 0.0001s, or 0.00001 apart, and/or (ii) whether a signal is produced by a reporter group of the biosensor continuously or repeatedly at intervals less than
15 about 30m, 15m, 10m, 5m, 1m, 30s, 15s, 10s, 5s, 1s, 0.1s, 0.001s, 0.0001s, or 0.00001apart.

Non-limiting aspects of continuously monitoring ligand levels are described in Weidemaier et al. (2011) Biosensors and Bioelectronics 26, 4117-4123 and Judge et al. (2011) Diabetes Technology & Therapeutics, 13(3):309-317, the entire contents of each of which are hereby incorporated herein by reference.

20 Also within the invention is a composition comprising a purified thermostable, ligand-binding fluorescently-responsive sensor protein and a solid substrate, *e.g.*, a particle, a bead such as a magnetic bead, or a planar surface such as a chip or slide, wherein the sensor protein is immobilized onto the solid substrate. An exemplary solid substrate solid substrate comprises a cyclic olefin copolymer.

25 A thermostable ligand sensor protein is one in which the activity (ligand binding) is retained after exposure to relatively high temperatures. For example, the ligand sensor protein comprises a mid-point thermal melt transition greater than 50°C, greater than 60°C, greater than 70°C, greater than 80°C, greater than 90°C, or greater than 100°C. In some embodiments, the sensor protein contains a single cysteine residue. In some embodiments,
30 the single cysteine residue is located in a site of the ligand-binding protein, where it responds to ligand binding. In some examples, the protein comprises the amino acid sequence of SEQ ID NO: 16 (ttGGBP.17C.0.bZif) and 19 (ttGGBP.182C.0.bZif), and in some examples, the single cysteine is conjugated to Badan, Acrylodan, or a derivative thereof. For example, the

derivative comprises a replacement of the two-ring naphthalene of Acrylodan or Badan with a three-ring anthracene, a fluorene, or a styrene. A reporter group is covalently bound to the single cysteine. In some situations, the solid substrate comprises a plurality of sensor proteins, each of which comprises a different dissociation constant (K_d) for ligand, e.g., for
5 detecting and quantifying ligand levels across many ranges of concentrations.

The present subject matter also includes a composition comprising purified ligand sensor protein with less than 65% identity and greater than 27% identity (e.g., 44-48% sequence identity) to any ligand-binding protein disclosed herein, wherein the sensor protein comprises a single cysteine residue, such that the sensor protein is immobilized onto the solid
10 substrate. As described above, a reporter group is covalently bound to the single cysteine. In some example, the solid substrate comprises a plurality of sensor proteins, each of which comprises a different dissociation constant (K_d) for ligand for sensing over a wide range or ranges of ligand concentrations.

In some embodiments, a method of detecting the presence of or the quantity of ligand in a test sample is carried out using the following steps: contacting the test sample with the biosensor or sensor protein/solid support construct to yield a complex of ligand and the ligand-binding protein or biosensor protein; contacting the complex with an excitation light; measuring an emission intensity of the reporter group from at least two wavelengths;
20 computing a ratiometric signal from the two (or more) wavelengths; and comparing the signal to a known ligand binding curve of signals to identify the presence of or calculate the quantity of ligand in the test sample. The test sample may be obtained from a variety of sources. For example, the test sample may be selected from a bodily fluid, a food, a beverage, or a bioreactor culture broth. The testing method may be carried out *in vivo*, e.g.,
25 using an implantable device or dermal patch, or *ex vivo*.

In various embodiments, the subject to be tested is a mammal, e.g., a primate (such as a human, a monkey, a chimpanzee, or a gorilla), a fish, a bird, a reptile, an amphibian, or an arthropod. In some embodiments, the subject is a fish, a cow, a pig, a camel, a llama, a horse, a race horse, a work horse, a goat, a rabbit, a sheep, a hamster, a guinea pig, a cat, a wolf, a
30 dog (e.g., a pet dog, a work dog, a police dog, or a military dog), a rat, a mouse, a seal, a whale, a manatee, a lizard, a snake, a chicken, a goose, a swan, a duck, or a penguin.

In some embodiments, the ligand comprises a halide anion and the ligand-binding protein comprises a fluorescent protein. Aspects of the present subject matter provide a

method for detecting the level of a halide anion in a sample, comprising contacting the sample with a biosensor for a halide anion under conditions such that the biosensor binds to a halide anion present in the sample. In various embodiments the biosensor comprises a halide anion-binding fluorescent protein, and binding of the halide anion to a halide anion-binding domain of the fluorescent protein causes a change in signaling by the fluorescent protein. In some embodiments, the sample is an environmental sample. In a non-limiting example, the sample comprises treated wastewater or drinking water.

Aspects of the present subject matter further provides a method for assaying the level of chloride in a subject, comprising contacting a biological sample from the subject with a biosensor for chloride under conditions such that the biosensor binds to chloride present in the biological sample. The biosensor may comprise, *e.g.*, a chloride-binding fluorescent protein, and binding of chloride to a chloride-binding domain of the fluorescent protein causes a change in signaling by the fluorescent protein. In some embodiments, the subject has or is suspected of having hypochloremia. Alternatively or in addition, the subject has or is suspected of having abnormal kidney or adrenal gland function. In certain embodiments, the biological sample comprises blood, plasma, serum, sweat, tear fluid, or urine. In some embodiments, the method is performed as part of a battery of clinical testing.

Also provided is a method for assaying the level of iodide in a subject, comprising contacting a biological sample from the subject with a biosensor for iodide under conditions such that the biosensor binds to iodide present in the biological sample, wherein the biosensor comprises an iodide-binding fluorescent protein. Binding of iodide to an iodide-binding domain of the fluorescent protein causes a change in signaling by the fluorescent protein. In some embodiments, the subject has or is suspected of having hypothyroidism, hyperthyroidism, cretinism, depression, fatigue, obesity, a low basal body temperature, a goiter, or a fibrocystic breast change. In some embodiments, the biological sample comprises blood, plasma, serum, sweat, tear fluid, or urine. In some embodiments, the method is performed as part of a battery of clinical testing.

The present subject matter further includes a method for assaying the level of bromide in a subject, comprising contacting a biological sample from the subject with a biosensor for bromide under conditions such that the biosensor binds to bromide present in the biological sample, wherein the biosensor comprises a bromide-binding fluorescent protein. Upon binding of bromide to a bromide-binding domain of the fluorescent protein, the signal of the fluorescent protein changes. In some embodiments, the subject has or is suspected of having

bromism. In certain embodiments, the biological sample comprises blood, plasma, serum, sweat, tear fluid, or urine. In some embodiments, the method is performed as part of a battery of clinical testing.

5 Exemplary Devices and Compositions Comprising Biosensors

Aspects of the present subject matter provide a device comprising one or more biosensors provided herein. Such devices may be, *e.g.*, wearable, implantable, portable, or fixed.

10 In some embodiments, the device is a nanoparticle or a microparticle comprising the biosensor. Non-limiting examples of devices include devices comprising a test strip, patch, plate, bead, or chip comprising a biosensor provided herein. In certain embodiments, a device may comprise a desiccated biosensor.

The present subject matter also provides a contact lens or a skin patch comprising a biosensor provided herein. In some embodiments, the biosensor is throughout the contact
15 lens or skin patch or within a particular region or zone of a contact lens or skin patch (*e.g.*, in one or more shapes (*e.g.*, a square, circle, or star), dots, lines, or zones, located at the periphery or a portion of the periphery of a contact lens or patch). In some embodiments, the skin patch comprises an adhesive that facilitates attachment of the patch to the surface of skin.

20 Devices provided herein may include a variety of structural compositions. For example, many polymers (including copolymers), and plastics may be used. Non-limiting examples of compositions useful in certain devices include glass, polystyrene, polypropylene, cyclic olefin copolymers, ethylene-norbornene copolymers, polyethylene, dextran, nylon, amylase, paper, a natural cellulose, a modified cellulose, a polyacrylamide, gabbros, gold,
25 and magnetite (as well as combinations thereof). In some embodiments, the device comprises a hydrogel, a cryogel, or a soluble gel. For example, the biosensor may be incorporated into or onto the hydrogel, cryogel, or soluble gel. In various embodiments, the device comprises a matrix comprising nanopores, micropores, and/or macropores. In certain embodiments, the surface of a device comprises a polymer. In an embodiment, the surface
30 comprises the surface of a particle or a bead having a diameter of about 0.001-1, 0.001-0.1, 0.01-0.1, 0.001-0.01, 0.1-1, 0.1-0.5, or 0.01-0.5 centimeters (cm). For example, the particle comprises a nanoparticle or a microparticle.

Non-limiting examples of polymers include cyclic olefin copolymers, ethylene-norbornene copolymers, polylactic acid, polyglycolic acid, agarose, alginate, poly(lactide-co-glycolide), gelatin, collagen, agarose, natural and synthetic polysaccharides, polyamino acids, poly(lysine), polyesters, polyhydroxybutyrates, polyanhydrides, polyphosphazines, polyvinyl alcohol, polyalkylene oxide, polyethylene oxide, polyallylamines, polyacrylates, modified styrene polymers, poly(4-aminomethylstyrene), pluronic polyols, polyoxamers, polyuronic acid, polyvinylpyrrolidone, hydroxyethyl (meth)acrylate, polyolefins, polyurethane, polystyrene, ethylene/methacrylic acid copolymers, ethylene/methyl methacrylate copolymers, polyester, and polyurethane. In some embodiments, the patch comprises a woven fabric, a knitted fabric, or a nonwoven fabric of a synthetic fiber and/or natural fiber.

Non-limiting examples of temporary tattoo compositions for application to a subject's skin are discussed in U.S. Patent Application Publication No. 20090325221, published December 31, 2009, and U.S. Patent No. 6,428,797, the entire contents of each of which are incorporated herein by reference. Biosensor disclosed herein may be incorporated into any temporary tattoo or other composition for application to the skin. For example, a temporary tattoo decal for application to a subject's skin and configured to detect the presence of a ligand may comprise, e.g., a base paper or plastic; a water-soluble slip layer applied to the base paper or plastic; a temporary tattoo applied to the water-soluble release layer on the base paper, wherein the temporary tattoo comprises a biosensor disclosed herein; an adhesive layer overlying the temporary tattoo; and a protective sheet overlying the adhesive layer.

In some embodiments, the device comprises a plastic polymer comprising cyclic olefin copolymer (COC), such as e.g. TOPAS® COC. Several types of cyclic olefin copolymers are available based on different types of cyclic monomers and polymerization methods. Cyclic olefin copolymers are produced by chain copolymerization of cyclic monomers such as 8,9,10-trinorborn-2-ene (norbornene) or 1,2,3,4,4a,5,8,8a-octahydro-1,4:5,8-dimethanonaphthalene (tetracyclododecene) with ethene (such as TOPAS Advanced Polymer's TOPAS, Mitsui Chemical's APEL), or by ring-opening metathesis polymerization of various cyclic monomers followed by hydrogenation (Japan Synthetic Rubber's ARTON, Zeon Chemical's Zeonex and Zeonor). See, e.g., International Union of Pure and Applied Chemistry (2005) *Purr. Appl. Chem.* 77(5):801-814. These later materials using a single type of monomer may be referred to as cyclic olefin polymers (COPs). A CAS Registry number for COC is 26007-43-2.

In certain embodiments, the device is attached to a surface of a device or is not attached to a surface of the device (*e.g.*, the biosensor is present loosely within the device as a component of a solution or powder).

A biosensor may be attached to a device via a variety of means, *e.g.*, via attachment motif. In some embodiments, the attachment motif is attached to the N-terminus or the C-terminus of the biosensor. In certain embodiments, the biosensor is linked to an attachment motif via a covalent bond. In various embodiments, the biosensor is linked to the attachment motif via a linker. A non-limiting example of a linker is a polyglycine comprising 2, 3, 4, 5, or more glycines and optionally further comprising a serine. In some embodiments, the attachment motif comprises a polypeptide. Non-limiting examples of polypeptides useful in attachment moieties include hexahistidine peptides, hexalysine peptides, zinc-finger domains (ZF-QNKs), and disulfide-containing truncated zinc fingers (β Zifs). An example of a hexalysine peptide comprises amino acids in the sequence of SEQ ID NO: 45, an example of a ZF-QNK comprises amino acids in the sequence of SEQ ID NO: 43, and an example of a β Zif comprises amino acids in the sequence of SEQ ID NO: 42. In some embodiments, the attachment motif comprises a polypeptide that binds to plastic or cellulose.

The hexahistidine, hexalysine, β Zif and QNK-ZF fusions enable FRSs to be immobilized onto chemically functionalized surfaces. Non-limiting aspects of chemically functionalized surfaces are discussed in Biju, V. (2014) *Chem Soc Rev*, 43, 744-64 and McDonagh (2008) *Chem Rev*, 108, 400-422, the entire contents of which are incorporated herein by reference. Directed evolution methods have been used to develop peptides that bind directly to non-functionalized surfaces (Care, Bergquist and Sunna 2015 *Trends Biotechnol*, 33, 259-68; Baneyx 2007 *Curr. Opin. Biotechnol.*, 18, 312-317; Gunay and Klok 2015 *Bioconjug Chem*, 26, 2002-15), including various plastics (Adey et al. 1995 *Gene*, 156, 27-31; Serizawa et al. 2005 *J Am Chem Soc*, 127, 13780-1; Serizawa, Sawada and Kitayama 2007a *Angew Chem Int Ed Engl*, 46, 723-6; Serizawa, Sawada and Matsuno 2007b *Langmuir*, 23, 11127-33; Serizawa, Techawanitchai and Matsuno 2007c *Chembiochem*, 8, 989-93; Matsuno et al. 2008 *Langmuir*, 24, 6399-403; Chen, Serizawa and Komiyama 2011 *J Pept Sci*, 17, 163-8; Kumada 2010 *J. Biosci. and BioEng.*, 109, 583-587; Date et al. 2011 *ACS Appl Mater Interfaces*, 3, 351-9; Kumada 2012, Vodnik, Strukelj and Lunder 2012 *J. Biotech.*, 160, 222-228; Kumada 2014 *Biochem. et Biophys. Acta*, 1844, 1960-1969; Ejima, Matsuno and Serizawa 2010 *Langmuir*, 26, 17278-85), inorganic materials (Hnilova 2012 *Soft Matter*, 8, 4327-4334; Care et al. 2015 *Trends Biotechnol*, 33, 259-68), nanoparticles

(Avvakumova et al. 2014 *Trends Biotechnol*, 32, 11-20), and cellulosic paper (Guo et al. 2013 *Biomacromolecules*, 14, 1795-805). Such peptides, or natural material-binding domains (Oliveira et al. 2015 *Biotechnol Adv*, 33, 358-69), also can be fused to FRSs to direct site-specific, oriented immobilization on their target materials while preserving FRS function.

5 For instance, plastic-binding peptides have been developed that direct immobilization on polystyrene (Adey et al. 1995 *Gene*, 156, 27-31; Serizawa et al. 2007c *Chembiochem*, 8, 989-93; Kumada 2010 *Biochem. et Biophys. Acta*, 1844, 1960-1969; Vodnik et al. 2012 *Anal Biochem*, 424, 83-6), polymethyl acrylate (Serizawa et al. 2005 *J Am Chem Soc*, 127, 13780-1; Serizawa et al. 2007a *Angew Chem Int Ed Engl*, 46, 723-6; Serizawa et al. 2007b
10 *Langmuir*, 23, 11127-33; Kumada 2014 *Biochem. et Biophys. Acta*, 1844, 1960-1969), polycarbonate (Kumada 2012 *J. Biotech.*, 160, 222-228), polylactide (Matsuno et al. 2008 *Langmuir*, 24, 6399-403), and polyphenylene vinylene (Ejima et al. 2010 *Langmuir*, 26, 17278-85). Cellulose-binding peptides (Guo et al. 2013 *Biomacromolecules*, 14, 1795-805) and natural domains (Oliveira et al. 2015 *Biotechnol Adv*, 33, 358-69; Shoseyov, Shani and
15 Levy 2006 *Microbiol Mol Biol Rev*, 70, 283-95) can be used to immobilize fusion proteins on paper. Inorganic material include noble metals (Hnilova 2012 *Soft Matter*, 8, 4327-4334), semi-conductors (Care et al. 2015 *Trends Biotechnol*, 33, 259-68), and fluorescent quantum dots (Medintz et al. 2005 *Nat Mater*, 4, 435-46; Lee et al. 2002 *Science*, 296, 892-5). The entire contents of each of the references above (and all other references herein) is
20 incorporated herein by reference.

In some embodiments, the attachment motif is attached to a device surface and/or within a matrix of the device. In some embodiments, a biosensor is attached to an attachment motif via a covalent bond and the attachment motif is attached to a device via a covalent bond. Non-limiting examples of covalent bonds include disulfide bonds, ester bonds,
25 thioester bonds, amide bonds, and bonds that have been formed by click reactions. Non-limiting examples of a click reaction include a reaction between an azide and an alkyne; an azide and an alkyne in the presence of Cu(I); an azide and a strained cyclooctyne; an azide and a dibenzylcyclooctyne, a difluorooctyne, or a biarylazacyclooctynone; a diaryl-strained-cyclooctyne and a 1,3-nitrone; an azide, a tetrazine, or a tetrazole and a strained alkene; an
30 azide, a tetrazine, or a tetrazole and a oxanorbornadiene, a cyclooctene, or a trans-cycloalkene; a tetrazole and an alkene; or a tetrazole with an amino or styryl group that is activated by ultraviolet light and an alkene.

Alternatively or in addition, a surface of a device may be modified to contain a moiety (e.g. a reactive group) what facilitates the attachment of a biosensor and/or binds to the biosensor. In some embodiments, the biosensor is attached to a surface via a biotin-avidin interaction.

5 In various implementations, the device comprises a first region for receiving a sample and second a region that comprises the biosensor, wherein the first region is separated from the second region by a filter. In some examples, the filter is impermeable to compounds greater than about 1, 2, 3, 4, 5, 10, 50, 200, or 250 kiloDalton (kDa) in size. The sample may comprise, e.g., a tube, such as a tube that is configured for centrifugation. When sample is
10 placed into the first region and the device is centrifuged, then a portion of the sample comprising a ligand flows through the filter into the second region where the biosensor is contacted.

Non-limiting examples of devices provided herein include endoscopy probes and colonoscopy probes.

15 In some embodiments, the device comprises an optode. In non-limiting examples, the optode comprises an optical fiber and a single biosensor or composite biosensor. In certain embodiments, the single biosensor or composite biosensor is immobilized on the surface or at an end of the optical fiber. In some embodiments, the optode is configured for implantation into a subject. Alternatively or in addition, the optode is configured for insertion into a
20 sample.

The devices provided herein may optionally comprise a biosensor panel, a composite sensor, a sensor array, and/or a composition comprising a plurality of biosensors. In various embodiments, a device comprises multiple ligand biosensors that detect a range of different ligand concentrations in a single sample and/or assay run (i.e., each biosensor has a different
25 affinity for ligand). Devices may provide spatial localization of multiple biosensors to provide the necessary addressability of different elements in a multi-sensor array comprising sensors that differ in their engineered affinities for coverage of a wide range of ligand concentrations, or sensors that each detects distinct analytes.

Aspects of the present subject matter provide a biosensor panel comprising a plurality
30 of biosensors, wherein the plurality of biosensors comprises at least one biosensor disclosed herein. In some embodiments, the plurality comprises at least about 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100 biosensors.

The present subject matter also provides a composite sensor. The composite sensor may comprise a sensor element, wherein the sensor element comprises 2 or more biosensors, wherein at least 1 of the 2 or more biosensors is a biosensor disclosed herein. In some embodiments, the biosensors are not spatially separated in the sensor element, *e.g.*, the biosensors are mixed within a solution or on a surface of the sensor element. In various embodiments, the composite sensor comprises a plurality of sensor elements, wherein each sensor element of the plurality of sensor elements comprises 2 or more biosensors, wherein at least 1 of the 2 or more biosensors is a biosensor provided herein. In some embodiments, the plurality of sensor elements comprises at least about 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100 sensor elements.

Also included herein is a sensor array comprising a plurality of biosensors of the present subject matter. The sensor array may include, *e.g.*, multichannel array or a multiplexed array. In some embodiments, the biosensors of the plurality of biosensors are spatially separated from each other. In certain embodiments, the biosensors are arranged linearly or in a grid on a surface of the array.

The present subject matter provides a composition comprising a plurality of biosensors including at least one biosensor disclosed herein. Also provided is a non-human mammal comprising a biosensor or device disclosed herein.

Exemplary Polypeptides and Polynucleotides

The present subject matter provides polynucleotides encoding any one of the polypeptides disclosed herein. The polypeptides are also provided. In various embodiments, the polynucleotides are codon-optimized for expression in a desired host cell, such as bacterial cells (*e.g.*, *E. coli*), yeast, insect cells, plant cells, algal cells, or mammalian cells.

The polypeptides provided herein include polypeptides comprising the amino acid sequence of any one of SEQ ID NOS: 1-41, 87-151, or 153-167. The polynucleotides provided herein include polynucleotides encoding a polypeptide comprising the amino acid sequence of any one of SEQ ID NOS: 1-41, 87-151, or 153-167.

The polypeptides and biosensors provided herein may be in a variety of forms, *e.g.*, purified in solution, dried (*e.g.* lyophilized) such as in the form of a powder, and in the form of a crystal (*e.g.*, a crystal suitable for x-ray crystallography). Thus, aspects of the present subject matter provide crystal structures and crystalized forms of the ligand-binding proteins

and biosensors disclosed herein. Such crystal structures and crystalized proteins are useful for designing and optimizing biosensors using principles and methods discussed herein.

Also provided are expression vectors comprising a polynucleotide of the present subject matter and/or encoding a polypeptide disclosed herein. Non-limiting examples of expression vectors include viral vectors and plasmid vectors. In some embodiments, an expression vector comprises nucleotides in the sequence set forth as any one of SEQ ID NOS: 46-86. In various embodiments, a polynucleotide encoding a ligand-binding protein and/or biosensor is operably linked to a promoter. The promoter may be expressed, *e.g.*, in a prokaryotic and/or a eukaryotic cell.

The subject matter further includes an isolated cell comprising an expression vector provided herein. The isolated cell may be, *e.g.*, a bacterial cell, a yeast cell, an algal cell, a plant cell, an insect cell, or a mammalian cell. Also included is a non-human multicellular organism such as a plant or an animal (*e.g.*, an insect, a mammal, a worm, a fish, a bird, or a reptile) comprising an expression vector disclosed herein.

Exemplary Methods for Designing Biosensors

Aspects of the present subject matter provide method of identifying a candidate ligand-binding protein for use in a biosensor, comprising: (a) selecting a first protein having a known amino acid sequence (seed sequence), wherein the first protein is a ligand binding protein; (b) identifying a second protein having an amino acid sequence (hit sequence) with at least 15% sequence identity to the seed sequence; (c) aligning the seed amino acid sequence and the hit sequence, and comparing the hit sequence with the seed sequence at positions of the seed sequence that correspond to at least 5 primary complementary surface (PCS) amino acids, wherein each of the at least 5 PCS amino acids has a hydrogen bond interaction or a van der Waals interaction with ligand when ligand is bound to the first protein; and (d) identifying the second protein to be a candidate ligand-binding protein if the hit sequence comprises at least 5 amino acids that are consistent with the PCS.

The present subject matter also includes a method for constructing a candidate biosensor, comprising: (a) providing a candidate ligand-binding protein; (b) generating a structure of the second protein; (c) identifying at least one putative allosteric, endosteric, or peristeric site of the second protein based on the structure; (d) mutating the second protein to substitute an amino acid at the at least one putative allosteric, endosteric, or peristeric site of the second protein with a cysteine; and (e) conjugating a fluorescent compound to the

cysteine. In some embodiments, the structure comprises a homology model of the second protein generated using a structure of the first protein. In some embodiments, the structure comprises a structure experimentally determined by nuclear magnetic resonance spectroscopy or X-ray crystallography.

5 Aspects of the present subject matter further provide a method for constructing a biosensor comprising a desired dissociation constant (K_d) for ligand, comprising: (a) providing an initial biosensor that does not comprise the desired K_d for ligand, wherein the initial biosensor is a biosensor provided herein; (b) mutating the initial biosensor to (i) alter a direct interaction in the PCS between the initial biosensor and bound ligand; (ii) manipulate
10 the equilibrium between open and closed states of the initial biosensor; (iii) alter an interaction between the ligand-binding protein and the reporter group of the initial biosensor; or (iv) alter an indirect interaction that alters the geometry of the binding site of the biosensor, to produce a modified biosensor; and (c) selecting the modified biosensor if the modified biosensor comprises the desired K_d for ligand. In some embodiments, the reporter
15 comprises Acrylodan, Badan, or a derivative thereof, and mutating the initial biosensor in (b) comprises altering an interaction between the ligand-binding protein and a carbonyl group of the Acrylodan, Badan, or derivative thereof. In some embodiments, the reporter group comprises Acrylodan, Badan, or a derivative thereof, and mutating the initial biosensor in (b) comprises altering an interaction between the ligand-binding protein and a naphthalene ring
20 of the Acrylodan, Badan, or derivative thereof. In some embodiments, mutating the initial biosensor comprises introducing a substitution mutation into the initial biosensor. In some embodiments, the method further comprises immobilizing the affinity-tuned biosensor on a substrate.

In some embodiments, the second protein comprises (i) amino acids in the sequence
25 of any one of SEQ ID NOS: 1-41, 87-151, or 153-167; (ii) a stretch of amino acids in a sequence that is least about 95, 96, 97, 98, or 99% identical to the sequence of any one of SEQ ID NOS: 1-41, 87-151, or 153-167; (iii) a stretch of at least about 50, 100, 150, 200, 250, 300, 350, or 400 amino acids in a sequence that is at least about 95, 96, 97, 98, or 99% identical to a sequence within any one of SEQ ID NOS: 1-41, 87-151, or 153-167; or (iv) a
30 stretch of at least about 50, 100, 150, 200, 250, 300, 350, or 400 amino acids in a sequence that is identical to a sequence within any one of SEQ ID NOS: 1-41, 87-151, or 153-167. In various embodiments, attaching the reporter group to the putative allosteric, endosteric, or peristeric site of the first protein comprises substituting a cysteine at the site with a cysteine.

For example, the reporter group is conjugated to the cysteine. Preferably, attaching a reporter group to the corresponding amino acid of the second protein produces a functional biosensor.

Aspects also provide method for constructing a biosensor, comprising (a) providing a ligand-binding protein; (b) identifying at least one putative allosteric, endosteric, or peristeric site of the ligand-binding based a structure of the ligand-binding protein; (c) mutating the ligand-binding protein to substitute an amino acid at the at least one putative allosteric, endosteric, or peristeric site of the second protein with a cysteine; (d) conjugating a donor fluorophore or an acceptor fluorophore to the cysteine to produce single labeled biosensor; (e) detecting whether there is a spectral shift or change in emission intensity of the single labeled biosensor upon ligand binding when the donor fluorophore or the acceptor fluorophore is fully excited; and (f) if a spectral shift or change in emission intensity is detected in (g), attaching a donor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine, and attaching an acceptor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine.

In some embodiments, the ligand-binding protein has been identified by (i) selecting a first protein having a known amino acid sequence (seed sequence), wherein the first protein is a ligand-binding protein; (ii) identifying a second protein having an amino acid sequence (hit sequence) with at least 15% sequence identity to the seed sequence; (iii) aligning the seed amino acid sequence and the hit sequence, and comparing the hit sequence with the seed sequence at positions of the seed sequence that correspond to at least 5 primary complementary surface (PCS) amino acids, wherein each of the at least 5 PCS amino acids has a hydrogen bond interaction or a van der Waals interaction with ligand when ligand is bound to the first protein; and (iv) identifying the second protein to be a ligand-binding protein if the hit sequence comprises at least 5 amino acids that are consistent with the PCS. The spectral shift may comprise, *e.g.*, a monochromatic fluorescence intensity change or a dichromatic spectral shift.

Also provided is a method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising (a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and (b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation

spectrum of the first reporter group. In certain embodiments, the second reporter group is within about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, or 200 angstroms (Å) of the first reporter group regardless of whether ligand is bound to the biosensor. In some embodiments, when the ligand is bound to the biosensor, the average distance between the first reporter group and the second reporter group changes by less than about 5, 4, 3, 2, 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, or 0.01 angstroms (Å) compared to when ligand is not bound to the ligand-binding protein.

Aspects include a method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising (a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding fluorescent protein; and (b) attaching an acceptor fluorophore or a donor fluorophore to the biosensor, wherein (i) the acceptor fluorophore has an excitation spectrum that overlaps with the emission spectrum of the fluorescent protein; or (ii) the donor fluorophore has an emission spectrum that overlaps with the excitation spectrum of the fluorescent protein.

Also provided is a method of increasing a dichromatic response of a biosensor to ligand binding, the method comprising (a) selecting a biosensor that exhibits a dichromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and (b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of the first reporter group.

The selected first protein (*e.g.*, the amino acid sequence thereof) may be novel or known. However, in many instances, the function of the first protein will not be known. In a non-limiting example, identifying a protein not previously known to have ligand binding activity may comprise a structurally assisted functional evaluation (SAFE) homolog search method comprising the following steps:

(1) Collecting a sequence homology set using a BLAST sequence alignment tool starting with ligand binding protein sequence disclosed herein as a seed. Permissive settings are used, such that pairwise hits are required to have a minimum of only, *e.g.*, 20%, 25%, 30%, 35% or 40% sequence identity with the seed sequence. The lengths of the hit and seed

are mutually constrained such that the alignment covers at least, *e.g.*, 60%, 65%, 70%, 85%, or 90% within each partner.

(2) Structure-based encoding of biological function: A primary complementary surface (PCS) comprising the protein residues that form hydrogen bonds and van der Waals contacts with a bound ligand is defined using computer-assisted, visual inspection of the three-dimensional structure of the biosensor-ligand complex. This definition specifies residue positions and their permitted amino acid identity. Multiple amino acid identities are permitted at each position to encode functionally equivalent residues. This definition establishes a search filter for the accurate prediction of ligand-binding proteins within the universe of sequence homologs collected in (1).

(3) Accurate sequence alignment: Tools such as ClustalW are used to construct an accurate alignment of all the sequence homologs. The ligand-binding protein seed sequence is included in the alignment. This multiple sequence alignment establishes the equivalent positions of the seed sequence (primary complementary surface) PCS in each sequence homolog.

(4) Function evaluation: The ligand-binding properties of each of the aligned sequence homologs is determined by measuring their compliance with the PCS sequence filter. A "Hamming distance", H , is assigned for each homolog, which specifies the degree of sequence identity of all the residues at the aligned PCS positions. A value of $H=0$ indicates that the identities of all the residues at the aligned PCS positions match the amino acid(s) allowed in the PCS search filter; $H>0$, indicates that one or more aligned positions have disallowed residues. Sequences for which $H=0$ are predicted to encode ligand-binding proteins.

(5) Selection of representative SAFE homologs: The sequence homologs are ordered by (a) identity with the seed PCS, as measured by the Hamming distance, (b) fractional overall sequence identity with the seed sequence. A subset for sequences with $H=0$, sampling the fractional overall sequence identity is selected for experimental verification.

In a non-limiting example, identifying a protein not previously known to have ligand binding activity may comprises the following steps:

(1) performing a computational search of sequence databases to define a broad group of simple sequence or structural homologs of any known, ligand binding protein;

(2) using the list from step (1), deriving a search profile containing common sequence and/or structural motifs shared by the members of the list [*e.g.* by using computer

programs such as MEME (Multiple Em for Motif Elicitation available at meme.sdsc.edu/meme/cgi-bin/meme.cgi) or BLAST];

(3) searching sequence/structural databases, using a derived search profile based on the common sequence or structural motif from step (2) as query (*e.g.*, using computer programs such as BLAST, or MAST (Motif Alignment Search Tool available at meme.sdsc.edu/meme/cgi-bin/mast.cgi), and identifying a candidate sequence, wherein a sequence homology and/or structural similarity to a reference ligand binding protein is a predetermined percentage threshold;

(4) compiling a list of candidate sequences to generate a list of candidate ligand binding proteins;

(5) expressing the candidate ligand-binding proteins in a host organism; and

(6) testing for ligand binding activity, wherein detection of ligand binding in the organism (or the media thereof) indicates that the candidate sequence comprises a novel ligand binding protein.

In non-limiting examples, the MEME suite of sequence analysis tools (meme.sdsc.edu/meme/cgi-bin/meme.cgi) can also be used as an alternative to BLAST. Sequence motifs are discovered using the program “MEME”. These motifs can then be used to search sequence databases using the program “MAST.” The BLAST search algorithm is well-known.

In various embodiments relating to alignments using a ClustalW alignment program, the ClustalW alignment program may be, *e.g.*, ClustalW alignment program version 2.1.

Each embodiment disclosed herein is contemplated as being applicable to each of the other disclosed embodiments. Thus, all combinations of the various elements described herein are within the scope of the invention.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below.

DESCRIPTION OF THE DRAWINGS

FIGS. 1A and B are cartoons showing fluorescent probes. FIG. 1A is a cartoon relating to indirect fluorescent responses. Fluorescent biosensors can be constructed by site-specifically attaching a fluorophore to a protein that undergoes a conformational change upon binding ligand (triangle) in a location between the two lobes of the protein (periplasmic binding protein or engineered derivative thereof), such that the shape and intensities of the fluorescent conjugate emission spectra changes. FIG. 1B is a cartoon relating to direct fluorescent responses. Fluorescent chemosensors based on fluorophores that interact directly with an analyte.

FIGS. 2A-C are graphs illustrating ratiometry. If the fluorescence emission spectrum changes shape in response to binding of an analyte, such as glucose, then the ratio of emission intensities at two appropriately chosen wavelengths reports on analyte concentration. FIG. 2A: In the absence of ligand, the emitted fluorescence color is predominantly blue, whereas the ligand complex fluoresces green. Arrows indicate the direction of change upon ligand addition. FIG. 2B: The ligand dependence of the absolute blue and green intensities. FIG. 2C: The ratio of the blue and green intensities reports enables ligand binding to be determined.

FIGS. 3A-D are graphs and diagrams showing three dominant factors that affect overall ligand-mediated change in fluorescence emission intensity between donor and acceptors in which one partner responds to ligand binding. FIG. 3A: Simplified Jablonski diagram illustrating radiative and non-radiative pathways in the donor and acceptor. The donor excited state (D^*) is formed through illumination by the excitation source (wavy arrow) whereas the acceptor excited state (A^*) is formed by resonance energy transfer (dashed arrow). The fluorescence intensity is determined by the ratio of radiative decay (gray arrows) of the excited states (gray lines) to the ground state (black line) relative to all non-radiative processes (black arrows), and the resonance energy transfer rate, k_r , from donor to acceptor. FIG. 3B: Inter-dipole geometry. Top, energy transfer efficiency ($f = Q_r / (Q_0 - Q_\infty)$, where the Q_r , Q_0 , Q_∞ are the quantum efficiencies at distances r , closest approach, and infinity, respectively) varies as the 6th power of the distance between two dipoles. Bottom, energy transfer efficiency varies as the square of the orientation factor κ , where $\kappa = \sin \theta_D \sin \theta_A \cos \chi - 2 \cos \theta_D \cos \theta_A$ with θ_D and θ_A the angles of the donor (blue) and acceptor (red) electronic transition dipoles with the line connecting them, and χ the angle

between the planes within which they lie. FIG. 3C: Spectral overlap (gray area) between the donor fluorescence emission ($^D I$, blue) and acceptor fluorescence excitation ($^A A$, black) spectra. This overlap increases with bathochromic or hypsochromic shifts of the donor emission (red arrow) and acceptor excitation (dotted blue arrow) spectra, respectively. Shifts in the opposite directions decreases spectral overlap.

FIGS. 4A-C are illustrations of the construction of ngmFRET pairs by combining a fluorophore that responds directly to ligand binding with a non-interacting, indirectly responsive partner. FIG. 4A: A glucose-binding protein in which a directly responsive partner is positioned in the vicinity of the glucose-binding site, and the indirectly responsive partner is fused as the protein C-terminus. FIG. 4B: The internally positioned fluorophore of Yellow Fluorescent Protein responds directly chloride binding by itself and regardless of the presence of any other fluorophore/partner. Its monochromatic response can be converted to a dichromatic one by positioning a second, indirectly responsive fluorophore on the surface of the protein. FIG. 4C: An adaptor protein such as *E. coli* thioredoxin can be used to position an directly responsive chemosensor next to an indirectly responsive partner, thereby converting a monochromatic into a dichromatic signal.

FIGS. 5A-E are cartoons of fusion constructs that enable site-specific labeling of cysteines at two independently addressable sites with distinct, thiol-reactive fluorophores. FIG. 5A: In the first labeling step, an unprotected single thiol (circle) reacts with a fluorophore, while thiols at a second site remain protected within a disulfide bridge. In the second labeling step, the disulfide is deprotected by reduction, and fluorophores are coupled to the second site. FIG. 5B: C-terminal fusion of a β Zif domain (slanted lines) to ttGGBP (solid gray). FIG. 5C: N-terminal fusion of β Zif. FIG. 5D: C-terminal fusion ecTRX (horizontal lines). FIG. 5E: N-terminal fusion of ecTRX.

FIG. 6 is a structural depiction of ionization states affecting fluorescence of the YFP fluorophore (by itself and regardless of the presence of any other fluorophore/partner).

FIG. 7 shows the sequence of laYFP and the locations of cysteine mutations for the construction of semisynthetic chloride sensors. Gray: the tripeptide that forms the fluorophore in the matured protein. Underlined: mutations that tune the YFP wavelength relative to GFP. laYFP retains the wild-type GFP residues H148 and V68 which affect chloride location and affinity (Wang 2015). Standard numbering is used in which the start methionine is 0. Structure taken from Protein Data Bank (PDB) Accession Code: 3sve.

FIGS. 8A and B are structures showing the locations of the cysteine mutations on the surface of laYFP. Arrows indicate β -barrel; central α helix; fluorophore; and chloride-binding site (dark gray sphere). Light gray spheres indicate location of surface cysteine mutations (numbered according to Fig. 7). FIG. 8A: Side-view showing that the cysteine mutations on the surface of the barrel β strands form an annulus that approximately encircles the fluorophore in the hydrophobic core. FIG. 8B: Top view showing the positions of all the annulus cysteine mutations around the barrel (end mutations omitted for clarity).

FIGS. 9A and B are graphs showing the chloride-dependent responses of the emission intensity spectra of representative Acrylodan and Pacific Blue YFP conjugates. Left column, normalized corrected emission spectra (see notes to Table 2): purple line, no chloride; red line, high chloride concentration; thin black lines, intermediate concentrations. YFP emission intensity peak is centered at 530 nm in both conjugates. Arrows indicate direction of change with increased chloride concentrations. Middle column, fit of ratiometric signal (R_{12}) to a Langmuir binding isotherm (yields $^{app}K_d$; see notes to Table 2). Right column: fits of Langmuir binding isotherm to monochromatic intensity (I) signals (yields $^{true}K_d$); gray circles, YFP intensity; black circles, Acrylodan or Pacific Blue intensity. FIG. 9A: ClBP4·Acrylodan ($\lambda_1 = 530$ nm, $\lambda_2 = 500$ nm; $^{app}K_d = 41$ mM; $^{true}K_d = 129$ mM). FIG. 9B: ClBP10·Pacific Blue ($\lambda_1 = 530$ nm, $\lambda_2 = 455$ nm; $^{app}K_d = 260$ mM; $^{true}K_d = 105$ mM).

FIG. 10 is an illustration of a structure showing positions for introducing cysteine mutations in ttGGBP to which fluorophores can be covalently coupled for reagentless biosensor construction. Positions 17, 91, 151, and 182 are endosteric; positions 11, 16, 42, 67, 92, 111, 148, 152, 181, and 183 are peristeric; and positions 257, 259, and 300 are allosteric

FIGS. 11A-P are illustrations of fluorophore structures. Naphthalene family (arrows indicate known or potential internal twists): FIG. 11A shows Acrylodan; FIG. 11B shows Badan; FIG. 11C shows IAEDANS. Xanthene family: FIG. 11D shows Fluorescein (5-IAF and 6-IAF); FIG. 11E shows Oregon Green; FIG. 11F shows Alexa 432; FIG. 11G shows Alexa 532; FIG. 11H shows Alexa 546; FIG. 11I shows Texas Red. Coumarin family: FIG. 11J shows Pacific Blue; FIG. 11K shows CPM. benzoxadiazole family: FIG. 11L shows IANBD. Boradiazaindacine (BODIPY) family: FIG. 11M shows BODIPY 499/508; FIG. 11N shows BODIPY 507/545. Cyanine family: FIG. 11O shows Cy5. Miscellaneous: FIG. 11P shows PyMPO.

FIGS. 12A and B is a pair of graphs showing donor quenching effects. FIG. 12A: The normalized fluorescence intensity ($I(\lambda)$; purple: apo-protein; red, high glucose; thin black line: intermediate concentrations) of the singly labeled F17C-Pacific Blue conjugate increases (blue arrow) in response to glucose binding without significant shifts in the wavelength of the intensity maximum. FIG. 12B: In the doubly labeled fusion protein, the fluorescence emission intensities (color scheme as in FIG. 12A) of the F17C-Pacific Blue directly responsive donor and the β Zif-IAF indirectly responsive acceptor both increase in response to glucose binding. In the absence changes in spectral overlap, the observed intensity response pattern is consistent with decreases in the non-radiative decay rate of the Pacific Blue donor (model $d\phi^0$, Table 1). The relative changes in donor and acceptor emissions are unequal, enabling ratiometric sensing (inset) based on the ratio (R_{12}) of intensities at 456 nm (donor) and 520 nm (acceptor) so that the glucose concentration can be fit to a single-site Langmuir binding isotherm (equations 22 and 24); constant baselines for the apo-protein and ligand complex; $^{app}K_d = 31$ mM).

FIG. 13A is a cartoon and FIG. 13B is a structural illustration relating to ratiometric sensing using Förster resonance energy transfer between pairs of glucose-responsive and non-responsive fluorophores attached site-specifically in a fusion protein that enables orthogonal site-specific cysteine labeling. FIG. 13A: Fusion of a single cysteine ttGGBP mutant (gray line; two alternative cysteine positions indicated, *i.e.*, F17C and W182C) and a disulfide-containing β Zif domain (disulfide indicated), separated by a linker (thin line), enables sequential labeling with two different fluorophores (line sizes indicate relative size of the fusion domains). FIG. 13B: Model showing the positions of the C-terminal β Zif fusion relative to the experimentally determined F17C-Badan and W182C-Acrylodan conjugates of ttGGBP. This domain is connected to ttGGBP via a flexible linker, and is therefore likely to adopt an ensemble of conformations, the approximate extent which is indicated by the oval.

FIGS. 14A-D are graphs showing acceptor dipole switching and quenching effects in fluorescein conjugates. FIG. 14A: W182C-5-IAF (directly responsive acceptor) β Zif-Pacific Blue (indirectly responsive donor). Normalized emission intensities are colored according to glucose concentration range: purple line, apo-protein; dotted gray line, emission at saturated glucose level for higher affinity binding site (phase I response, see main text); red line, emission intensity at highest glucose concentration measured; solid black, intermediate glucose concentrations for phase I; dotted black lines, elevated glucose concentrations in the

phase II response. Directions of signal change: bottom arrow, phase I; top arrow, phase II. Inset, Langmuir binding isotherm of ratiometric signal R_{12} at 456 nm and 520 nm, $^{app}K_d = 2.9$ mM, constant and linear baselines for apo-protein and ligand complex respectively. FIG. 14B: Contour plot of FIG. 14A, indicating phase I and phase II responses (see main text). FIG. 14C: W182C·Oregon Green (directly responsive acceptor) β Zif·Pacific Blue (indirectly responsive donor) coloring as in FIG. 14A. FIG. 14D: Contour plot of FIG. 14C.

FIGS. 15A and B are graphs showing glucose-dependent emission spectra of F17C·Badan and W182C·Acrylodan conjugates of ttGGBP. Corrected spectra (apo-protein, dark red; saturated glucose, purple; intermediate glucose concentrations, black). Insets, fit of the ratiometric signal (equation 1 and 2; 20 nm integration bandwidth): gray circles, experimentally observed ratios; black line, calculated fit (baselines; apo-protein, constant; saturated glucose complex, linear). FIG. 15A: F17C·Badan (hypsochromic; λ_1 , 470 nm; λ_2 , 542 nm; $^{app}K_d$, 0.18 mM); FIG. 15B: F182C·Acrylodan (bathochromic; λ_1 , 475 nm; λ_2 , 545 nm; $^{app}K_d$, 2.2 mM).

FIGS. 16A-D are graphs showing glucose dependence of electronic transitions in the fluorescence emission intensity spectra of F17C·Badan and W182C·Acrylodan ttGGBP conjugates. Columns: left, singular value decomposition (SVD) of the glucose-dependent corrected spectra; right, Gaussian analysis of electronic transitions in the fluorescence emission intensity. FIGS. 16A and B: F17C·Badan; FIGS. 16C and D: F182C·Acrylodan. For SVD analysis, frequency transformations of the spectra (equation 30) were decomposed into principal components (equation 31). The contribution of the first component, C_1 (black), is largely invariant with glucose concentration, whereas the second component, C_2 (red), encode accounts for the glucose-dependent changes in the spectra (inserts: glucose dependence of the fractional contribution of each component, equation 32). In the Gaussian analysis, the spectral emission intensities can be accounted for to a first approximation by two excited state electronic transitions: a low-energy, S_1 (Acrylodan: 521±9 nm; Badan: 530±14 nm) and a high-energy, S_2 (Acrylodan: 477±14 nm; Badan: 477±16 nm) transition. Glucose binding shifts the population of these excited states. At each titration point, the experimentally observed emission intensities were modeled with Gaussians fits (equation 33) for S_1 and S_2 electronic transition. Experimental emission spectra and Gaussian fits are shown only for the apo-protein, and saturated glucose complex. Emission spectra: dashed black line, apo-protein; solid black line, glucose complex. Gaussians (S_1 , lines; S_2 , lines): dashed lines, apo-protein; solid lines, glucose complex. Thin black lines: residuals (equation

35) at each titration point. Inserts show the population fractions (equation 34) of the S_1 and S_2 transitions extracted from the spectra at each titration point (black circles) fit to Langmuir binding isotherms (solid lines) with $^{app}K_d$ values constrained to be the same for both populations. As a first approximation, the wavelengths of S_1 or S_2 transitions are the same in apo-protein and the saturated glucose complex. The residuals indicate that a more extensive treatment is required in which the S_1 and S_2 are split into multiple transitions to fully fit the spectra. Wavelength shifts occur if there is a significant redistribution of the two excited state populations in the apo-protein and the saturated ligand complexes. In the bathochromic ratiometric 182C·Acrylodan conjugate, the S_1 state dominates in the glucose complex (FIG. 16D); in the hypsochromic conjugate 17C·Badan (FIG. 16B) the apo-protein comprises a mixture of the two states, whereas the glucose complex contains almost exclusively the S_2 state.

FIGS. 17A-D are graphs showing glucose dependence of the absorption spectra of ttGGBP Acrylodan and Badan conjugates that undergo wavelength shifts in their fluorescence emission intensities in response to ligand binding. Columns: left, singular value decomposition (SVD) of the glucose-dependent corrected spectra; right, Gaussian analysis of electronic transitions in the absorption spectra. FIGS. 17A and B: F17C·Badan; FIGS. 17C and D: F182C·Acrylodan. For SVD analysis, frequency transformations of the spectra (equation 30) were decomposed into principal components (equation 31). Inserts show glucose dependence of the fractional contribution of each component (equation 32). As with the fluorescence emission spectra (FIG. 16), in the Gaussian analysis, the spectral emission intensities can be accounted for to a first approximation by two ground state electronic transitions (dashed black line, apo-protein; solid black line, glucose complex): G_1 (386 nm) and G_2 (359 nm; apo-protein, dashed; glucose complex, solid). Inset, glucose dependence of population fractions at each glucose titration (black circles) of the G_1 and G_2 transitions fit to Langmuir binding isotherms (solid lines) with $^{app}K_d$ values constrained to be the same for both populations (residuals, thin lines).

FIGS. 18A-F are a set of graphs showing donor dipole switching effects in Acrylodan and Badan conjugates. Four doubly labeled conjugates were constructed, in which Acrylodan or Badan directly responsive donors were combined with Alexa532 or 5-IAF indirectly responsive acceptors. FIG. 18A: F17C·Badan β Zif-Alexa532. $I(\lambda)$, normalized emission intensity: purple line, apo-protein; red line, saturating glucose concentration; thin black lines, intermediate glucose concentrations. Blue arrow, direction of change with increased glucose

concentration. Inset, Langmuir binding isotherm of ratio, R_{12} , of the intensities at 467 nm and 560 nm (equations 23 and 24; constant baselines for apo-protein and ligand complex; $^{app}K_d = 0.16$ mM). FIG. 18B: F17C·Badan β Zif-5-IAF (color scheme as for FIG. 18A). Inset, Langmuir isotherm of ratiometric signal R_{12} at 467 nm and 520 nm, $^{app}K_d = 0.18$ mM, constant and linear baselines for apo-protein and ligand complex respectively. FIG. 18C: W182C·Acrylodan β Zif-Alexa532. Normalized emission intensities are colored according to glucose concentration range: purple line, apo-protein; dotted red line, emission at saturated glucose level for higher affinity binding site (phase I response, see main text); red line, emission intensity at highest glucose concentration measured; solid black, intermediate glucose concentrations for phase I; dotted black lines, elevated glucose concentrations in the phase II response. Directions of signal change: blue arrow, phase I; red arrow, phase II. Inset, Langmuir binding isotherm of ratiometric signal R_{12} at 480 nm and 550 nm, $^{app}K_d = 1.7$ mM, constant and linear baselines for apo-protein and ligand complex respectively. FIG. 18D: W182C·Acrylodan β Zif-5-IAF. Coloring according to FIG. 18C. Inset, Langmuir binding isotherm of ratiometric signal R_{12} at 465 nm and 520 nm, $^{app}K_d = 1.9$ mM, constant and linear baselines for apo-protein and ligand complex respectively. FIG. 18E: W182C·Acrylodan β Zif-Alexa532 contour plot of the glucose dependence of emission intensities, indicating phase I and phase II responses (see main text). FIG. 18F: W182C·Acrylodan β Zif-5-IAF contour plot.

FIGS. 19A and B are graphs providing a comparison of singly-labeled ttGBP17C·Badan (A) and the C-terminal doubly labeled ttGBP17C·Badan-ecTRX·Alexa532 fusion (B). Note the appearance of the approximately invariant emission peak (arrow) of the indirectly responsive Alexa532 acceptor.

FIGS. 20A-D are graphs showing donor dipole switching effects in W182C·IAEDANS. FIG. 20A: Singular value decomposition analysis of the change in the emission intensity of singly labeled W182C·IAEDANS in response to glucose, showing the wavenumber dependence of the invariant (black line) and variant (red line) spectral components. Inset shows change in contribution of the two spectral components with respect to glucose concentration. FIG. 20B: The glucose response is accounted for largely by two electronic transitions (green, 542 nm; blue, 485 nm) which were fits as Gaussians to the experimental emission intensities (purple line, in the absence of glucose; red line, saturating glucose). Black lines show residuals between observed and calculated spectra. Inset shows

the change in the populations of the blue and green transitions in response to glucose. FIG. 20C: W182C·IAEDANS β Zif-5-IAF. Normalized emission intensities, $I(\lambda)$, are colored according to glucose concentration range: purple line, apo-protein; dotted red line, emission at saturated glucose level for higher affinity binding site (phase I response, see main text); red line, emission intensity at highest glucose concentration measured; solid black, intermediate glucose concentrations for phase I; dotted black lines, elevated glucose concentrations in the phase II response. Directions of signal change: blue arrow, phase I; red arrow, phase II. Inset, Langmuir binding isotherm of ratiometric signal R_{12} at 465 nm and 520 nm, $^{app}K_d = 0.09$ mM, constant and linear baselines for apo-protein and ligand complex respectively.

FIG. 20D: Contour plot of the glucose dependence of W182C·IAEDANS β Zif-5-IAF emission intensities, indicating phase I and phase II responses (see main text).

FIG. 21 is a structural illustration of ionization equilibria of the fluorescein carboxylate and phenolic hydroxyl (Martin 1975).

FIG. 22 is an illustration of the structure of *E. coli* thioredoxin (Katti 1990) showing positions of mutations constructed in the adaptor proteins. Disulfide (C32,C35) is indicated. The D2A, D26A, and K57M background mutations constructed in all adaptor proteins are indicated (D2A removes adventitious N-terminal Cu(II)-binding site; D26A and K57M remove charges buried in the hydrophobic core). Large gray spheres: surface lysines mutated to arginine in Adaptor2.0a (K4Q and K18Q in Adaptor2.0b). Structure from PDB accession code 2trx.

FIG. 23 shows amino acid sequences of the engineered adaptor proteins based on *E. coli* thioredoxin. Numbering according the X-ray structure of the mature protein, lacking the initial methionine. Wild-type sequence is shown in full, mutations are given below (blank indicates wild-type residue).

FIGS. 24A-D are graphs showing the pH dependence of the emission spectra of Adaptor1.0 conjugates. Left column: corrected fluorescence emission intensity spectra (see notes to Table 9) of doubly labeled conjugates (purple, pH 4.0; red, pH 9.5; thin black lines, intermediate values at 0.5 pH unit intervals); middle column: ratiometric response, $R_{12} = I_1/I_2$ (black circles), as a function of pH (gray lines, fits to Langmuir binding isotherm give $^{app}pK_a$ values), where I_1 and I_2 are the intensities integrated over a 20 nm band centered at λ_1 and λ_2 ; right column: integrated fluorescence emission intensity changes, I , at the two individual wavelengths as a function of pH (gray lines, fits give $^{true}pK_a$ values). In all cases I_1 (gray circles) corresponds to the Fluorescein integrated intensities at 520 nm (red increases with

pH, arrows), and I_2 (black circles) to 460 nm and 455 nm for Acrylodan and Pacific Blue (PB), respectively. FIG. 24A: Fluorescein attached to C73, and Acrylodan to the disulfide ($^{app}pK_a = 6.22$, $^{true}pK_a = 6.16$). FIG. 24B: Fluorescein attached to the disulfide, and Acrylodan to C73 ($^{app}pK_a = 6.09$, $^{true}pK_a = 6.13$). FIG. 24C: Fluorescein attached to C73, Pacific Blue to the disulfide ($^{app}pK_a = 6.26$, $^{true}pK_a = 6.22$); FIG. 24D: Fluorescein attached to the disulfide, Pacific Blue to C73 ($^{app}pK_a = 6.14$, $^{true}pK_a = 6.17$).

FIGS. 25A-D are graphs showing the pH dependence of the absorption spectra of Adaptor1.0 conjugates. Left column: corrected absorbance, $A(\lambda)$, spectra (see notes to Table 9) of doubly labeled conjugates (purple, pH 4.0; red, pH 9.5; thin black lines, intermediate values at 0.5 pH unit intervals); middle column: ratiometric response, $R_{12} = A_1/A_2$ (black circles), as a function of pH (gray lines, fits to Langmuir binding isotherm give $^{app}pK_a$ values), where A_1 and A_2 are the absorbances integrated over a 20 nm band centered at λ_1 and λ_2 ; right column: integrated absorbance changes, A , at the two individual wavelengths as a function of pH (gray lines, fits give $^{true}pK_a$ values). In all cases A_1 (gray circles) corresponds to the Fluorescein integrated absorbances at 495 nm (increases with pH, arrows), and A_2 (black circles) to 390 nm and 410 nm for Acrylodan and Pacific Blue (PB), respectively (decreases with pH, arrows). FIG. 25A: Fluorescein attached to C73, and Acrylodan to the disulfide ($^{app}pK_a = 6.52$, $^{true}pK_a = 6.40$). FIG. 25B: Fluorescein attached to the disulfide, and Acrylodan to C73 ($^{app}pK_a = 6.70$, $^{true}pK_a = 6.62$). FIG. 25C: Fluorescein attached to C73, Pacific Blue to the disulfide ($^{app}pK_a = 6.72$, $^{true}pK_a = 6.77$); FIG. 25D: Fluorescein attached to the disulfide, Pacific Blue to C73 ($^{app}pK_a = 6.73$, $^{true}pK_a = 6.66$).

FIGS. 26A and B are graphs showing the pH dependence of the fluorescence emission and absorption spectra of Adaptor2.0a conjugate labeled with Fluorescein at the amino terminus, and Acrylodan at the disulfide. Left column, corrected spectra (purple, pH 4.0; red, pH 9.5; thin black lines, intermediate values at 0.5 pH unit intervals). Middle column: ratiometric responses (black circles, ratiometric signals; gray lines, fits to Langmuir binding isotherm give $^{app}pK_a$ values). Right column, integrated changes at the two individual wavelengths gray circles, λ_1 ; black circles, λ_2) as a function of pH (gray lines, fits give $^{true}pK_a$ values). FIG. 26A: pH dependence of fluorescence emission ($\lambda_1 = 520$ nm, $\lambda_2 = 460$ nm; $^{app}pK_a = 5.83$, $^{true}pK_a = 5.82$). FIG. 26B: pH dependence of absorption properties ($\lambda_1 = 498$ nm, $\lambda_2 = 375$ nm; $^{app}pK_a = 6.36$, $^{true}pK_a = 6.08$).

FIGS. 27A-F are graphs showing temperature dependence and pH dependence of Adaptor1.0 and Adaptor 2.0a fluorescent conjugates measured on a Roche LightCycler. Left column: Adaptor1.0 R73C·Pacific Blue, disulfide·Fluorescein, fluorescence ratio recorded at 488 nm and 510 nm; right column: Adaptor 2.0·Fluorescein, disulfide·Pacific Blue, fluorescent ratio recorded at 488 nm and 580 nm. FIGS. 27A, B: pH- and temperature-dependent landscape of the fluorescence ratio (Z axis): dashed-dotted line, approximate mid-point concentrations (pK_a) of the response to pH; dashed line, approximate mid-point temperatures (T_m) of thermal stability (equation 36). FIGS. 27C, D: Thermal melt at pH 4.5, as monitored by $R_{12}=I(488\text{ nm})/I(510\text{ nm})$ emission intensity ratio: blue, experimental data; green, fit to two-state van 't Hoff thermal denaturation ($T_m \sim 360\text{K}$). FIGS. 27E, F: Temperature dependence of pK_a values: crosses, measured pK_a values; line, fit to Gibbs-Helmholtz temperature dependence of the free energy of H^+ binding (equation 37).

FIG. 28 is a diagram relating to directly responsive partners and indirectly responsive partners in ngmFRET pathways.

FIG. 29 shows the sequence of an exemplary chloride-binding protein (YFP) 1 (ClBP1) expression construct (SEQ ID NO: 46).

FIG. 30 shows the sequence of an exemplary ClBP2 expression construct (SEQ ID NO: 47).

FIG. 31 shows the sequence of an exemplary ClBP3 expression construct (SEQ ID NO: 48).

FIG. 32 shows the sequence of an exemplary ClBP4 expression construct (SEQ ID NO: 49).

FIG. 33 shows the sequence of an exemplary ClBP5 expression construct (SEQ ID NO: 50).

FIG. 34 shows the sequence of an exemplary ClBP6 expression construct (SEQ ID NO: 51).

FIG. 35 shows the sequence of an exemplary ClBP7 expression construct (SEQ ID NO: 52).

FIG. 36 shows the sequence of an exemplary ClBP8 expression construct (SEQ ID NO: 53).

FIG. 37 shows the sequence of an exemplary ClBP9 expression construct (SEQ ID NO: 54).

FIG. 38 shows the sequence of an exemplary ClBP10 expression construct (SEQ ID NO: 55).

FIG. 39 shows the sequence of an exemplary ClBP11 expression construct (SEQ ID NO: 56).

5 FIG. 40 shows the sequence of an exemplary ClBP12 expression construct (SEQ ID NO: 57).

FIG. 41 shows the sequence of an exemplary ClBP13 expression construct (SEQ ID NO: 58).

10 FIG. 42 shows the sequence of an exemplary ClBP14 expression construct (SEQ ID NO: 59).

FIG. 43 shows the sequence of an exemplary ttGGBP.11C.0.bZif expression construct (SEQ ID NO: 60).

FIG. 44 shows the sequence of an exemplary ttGGBP.17C.0.bZif expression construct (SEQ ID NO: 61).

15 FIG. 45 shows the sequence of an exemplary ttGGBP.111C.0.bZif expression construct (SEQ ID NO: 62).

FIG. 46 shows the sequence of an exemplary ttGGBP.151C.0.bZif expression construct (SEQ ID NO: 63).

20 FIG. 47 shows the sequence of an exemplary ttGGBP.182C.0.bZif expression construct (SEQ ID NO: 64).

FIG. 48 shows the sequence of an exemplary ttGGBP.17C.3.Trx expression construct (SEQ ID NO: 65).

FIG. 49 shows the sequence of an exemplary Trx.ttGGBP.17C.3 expression construct (SEQ ID NO: 66).

25 FIG. 50 shows the sequence of an exemplary ttGGBP.182C.2.Trx expression construct (SEQ ID NO: 67).

FIG. 51 shows the sequence of an exemplary Trx.ttGGBP.182C.2 expression construct (SEQ ID NO: 68).

30 FIG. 52 shows the sequence of an exemplary Adaptor0 expression construct (SEQ ID NO: 69).

FIG. 53 shows the sequence of an exemplary Adaptor1.0 expression construct (SEQ ID NO: 70).

FIG. 54 shows the sequence of an exemplary Adaptor2.0a expression construct (SEQ ID NO: 71).

FIG. 55 shows the sequence of an exemplary Adaptor2.0b expression construct (SEQ ID NO: 72).

5 FIG. 56 shows the sequence of an exemplary Adaptor3.0 expression construct (SEQ ID NO: 73).

FIG. 57 shows the sequence of an exemplary Adaptor4.0 expression construct (SEQ ID NO: 74).

10 FIG. 58 shows the sequence of an exemplary Adaptor5.0 expression construct (SEQ ID NO: 75).

FIG. 59 shows the sequence of an exemplary Adaptor6.0 expression construct (SEQ ID NO: 76).

FIG. 60 shows the sequence of an exemplary Adaptor7.0 expression construct (SEQ ID NO: 77).

15 FIG. 61 shows the sequence of an exemplary Adaptor8.0 expression construct (SEQ ID NO: 78).

FIG. 62 shows the sequence of an exemplary Adaptor9.0 expression construct (SEQ ID NO: 79).

20 FIG. 63 shows the sequence of an exemplary Adaptor10.0 expression construct (SEQ ID NO: 80).

FIG. 64 shows the sequence of an exemplary Adaptor11.0 expression construct (SEQ ID NO: 81).

FIG. 65 shows the sequence of an exemplary Adaptor12.0 expression construct (SEQ ID NO: 82).

25 FIG. 66 shows the sequence of an exemplary Adaptor13.0 expression construct (SEQ ID NO: 83).

FIG. 67 shows the sequence of an exemplary Adaptor14.0 expression construct (SEQ ID NO: 84).

30 FIG. 68 shows the sequence of an exemplary Adaptor15.0 expression construct (SEQ ID NO: 85).

FIG. 69 shows the sequence of an exemplary Adaptor16.0 expression construct (SEQ ID NO: 86).

DETAILED DESCRIPTION

Biosensors are analytical tools that can be used to measure the presence of a single molecular species in a complex mixture by combining the exquisite molecular recognition properties of biological macromolecules with signal transduction mechanisms that couple ligand binding to readily detectable physical changes (Hall, Biosensors, Prentice-Hall, Englewood Cliffs, N.J.; Scheller et al., Curr. Op. Biotech. 12:35-40, 2001). Ideally, a biosensor is reagentless and, in contrast to enzyme-based assays or competitive immunoassays, does not change composition as a consequence of making the measurement (Hellings & Marvin, Trends Biotech. 16:183-189, 1998). Most biosensors combine a naturally occurring macromolecule such as an enzyme or an antibody, with the identification of a suitable physical signal particular to the molecule in question, and the construction of a detector specific to that system (Meadows, Adv. Drug Deliv. Rev. 21:177-189, 1996). Recently, molecular engineering techniques have been explored to develop macromolecules that combine a wide range of binding specificities and affinities with a common signal transduction mechanism, to construct a generic detection system for many different analytes (Hellings & Marvin, Trends Biotech. 16:183-189, 1998).

Escherichia coli periplasmic binding proteins are members of a protein superfamily (bacterial periplasmic binding proteins, bPBPs) (Tam & Saier, Microbiol. Rev. 57:320-346, 1993). These proteins comprise two domains linked by a hinge region (Quirocho & Ledvina, Molec. Microbiol. 20:17-25, 1996). The ligand-binding site is located at the interface between the two domains. The proteins typically adopt two conformations: a ligand-free open form, and a ligand-bound closed form, which interconvert via a hinge-bending mechanism upon ligand binding. This global, ligand-mediated conformational change has been exploited to couple ligand binding to changes in fluorescence intensity by positioning single, environmentally sensitive fluorophores in locations that undergo local conformational changes in concert with the global change (Brune et al., Biochemistry 33:8262-8271, 1994; Gilardi et al., Prot. Eng. 10:479-486, 1997; Gilardi et al., Anal. Chem. 66:3840-3847, 1994; Marvin et al., Proc. Natl. Acad. Sci. USA 94:4366-4371, 1997, Marvin and Hellings, J. Am. Chem. Soc. 120:7-11, 1998; Tolosa et al., Anal. Biochem. 267:114-120, 1999; Dattelbaum & Lakowicz, Anal. Biochem. 291:89-95, 2001; Marvin & Hellings, Proc. Natl. Acad. Sci. USA 98:4955-4960, 2001; Salins et al., Anal. Biochem. 294:19-26, 2001).

Provided herein are improved biosensors that rapidly, reliably, and accurately detect and quantify ligands with significant advantages over previous systems. The present

disclosure provides a biosensor for ligand, comprising a ligand-binding protein that is attached to one or more reporter group (*e.g.*, 1, 2, 3, or more reporter groups). The binding of a ligand to the ligand-binding domain of the ligand-binding protein causes a change in signaling by the biosensor. In various implementations, the biosensor may produce a signal when a ligand is bound to the ligand binding domain that is not produced (and/or that is different from a signal that is produced) when the ligand is absent from the ligand binding domain. These biosensors have widespread utility including in clinical, industrial, and environmental settings.

Various biosensors provided herein produce a dichromatic, ratiometric signal, *i.e.*, the signal is defined as the quotient of the intensities at two independent wavelengths. The advantage of such a signal is that it provides an internally consistent reference. The self-calibrating nature of a ratiometric measurement removes the necessity for carrying out on-board calibration tests prior to each measurement. The biosensors are reagentless in that their monitoring mechanism requires neither an enzyme nor additional substrates for a signal to develop, nor measurement of substrate consumption or product generation rates to determine ligand concentrations.

Reagentless, fluorescently responsive biosensors present a number of advantages over enzyme-based biosensors, including elimination of chemical transformations, elimination of substrate requirements, and self-calibration, which together lead to rapid response times, continuous monitoring capabilities, simple sample-handling, and lower cost due to simplified manufacturing and distribution processes.

ngmFRET for Ratiometric Measurements Using Reagentless Analyte Sensors

Determination of analyte concentrations using fluorescent probes is a powerful technique in analytical chemistry (FIGS. 1A and B). Fluorescent chemosensors based on small-molecule fluorophores that interact directly with an analyte (Zhang, Yin and Yoon 2014, Lavis and Raines 2008, Lavis and Raines 2014) and fluorescent biosensors based on engineered proteins that couple analyte-binding events to changes in the emission properties of fluorophores (being fluorescent by themselves and regardless of the presence of any other fluorophore/partner) (Okumoto 2012) or semi-synthetically (Wang 2009) incorporated fluorophores have wide-ranging applications in cell biology and analytical chemistry (Borisov and Wolfbeis 2008, Liu 2015, Matzeu 2015, Heo and Takeuchi 2013). If the fluorescence emission spectrum changes shape in response to analyte binding such that the ratio of

emission intensities at two appropriately chosen wavelengths reports on analyte concentration (dichromatic response), then ratiometric measurements can be used to monitor analyte concentrations (FIGS. 2A-C). Ratiometry is essential for devices that rely on quantifying changes in fluorescence emission intensities, because it provides an internally consistent reference (Demchenko 2010, Demchenko 2014). The self-calibrating nature of a ratiometric measurement removes the necessity for carrying out on-board calibration tests prior to each measurement (Choleau et al. 2002), obviating the need for multiple components and fluidic circuitry. Accordingly, reagentless, ratiometric fluorescent sensors have many uses in process engineering, environmental or clinical chemistry, including single-use point-of-care applications (Kozma et al. 2013, Ahmed et al. 2014, Mohammed 2011, Ispas 2012, Rogers and Boutelle 2013, Robinson and Dittrich 2013, Arora et al. 2010, Gubala et al. 2012), wearable devices (Badugu, Lakowicz and Geddes 2005), optodes for continuous monitoring (Weidemaier et al. 2011, Judge et al. 2011), or implanted “tattoos” that are interrogated transdermally (Bandodkar et al. 2015).

The majority of fluorescent chemosensors and biosensors do not undergo changes in emission spectral shape upon analyte binding and accordingly evince monochromatic intensity changes, rather than the dichromatic responses required for ratiometric sensing (de Lorimier et al. 2002). The present subject matter provides methods for converting monochromatic responses into dichromatic responses that enable ratiometric sensing. In embodiments, these methods are based on establishing non-geometrically modulated Förster Resonance Energy Transfer (ngmFRET) between the monochromatic fluorophore (directly responsive partner), and a second fluorophore that neither interacts directly with the ligand, nor is sensitive to ligand-mediated changes in its environment (indirectly responsive partner). Unlike tgmFRET-based chemical sensing systems (Valeur 2012), this arrangement does not rely on analyte-mediated geometrical changes (inter-fluorophore distance or angle) between the donor and acceptor, but instead exploits effects by analyte binding, which alter the photophysics of only the directly responsive partner such as changes in its spectral properties and non-radiative decay rates (FIG. 3).

The exemplary and non-limiting studies described herein demonstrate how these ngmFRET effects were used to convert monochromatic into dichromatic responses and thereby improve the ratiometric properties of dichromatic responses, using three classes of examples that illustrate the application of this technique both to biosensors and chemosensors (FIG. 4):

1. The analyte recognition element is a protein that undergoes an analyte-mediated conformational change that is alters the properties of an environmentally responsive directly responsive partner (FIG. 4A): a glucose-binding protein in which a conjugated directly responsive fluorophore respond via a glucose-induced protein conformational change that alters its emission properties. This fluorophore is paired with a second, indirectly responsive partner attached to a fusion domain (such as a fluorophore attachment motif attached to, *e.g.*, the first or the last amino acid of the ligand-binding protein). The resulting ngmFRET established between the two partners converts a monochromatic response of the directly responsive partner into a dichromatic response, or improves its dichromatic, ratiometric properties.
2. The analyte recognition element is a rigid protein with an analyte-binding site located adjacent to an fluorophore (having fluorescence by itself and regardless of the presence of any other fluorophore/partner; FIG. 4B): the monochromatic response of the fluorophore to chloride ion binding in a yellow fluorescent protein is converted to a dichromatic response using a indirectly responsive extrinsic fluorophore site-specifically attached to the protein surface.
3. The analyte recognition element is a synthetic chemoresponsive fluorophore (FIG. 4C): an adaptor protein is engineered to establish ngmFRET between two, site-specifically attached extrinsic fluorophores. The monochromatic response of the directly responsive partner to proton binding is converted into a dichromatic signal.

The first example represents a large class of protein-based fluorescent biosensors which undergo ligand-mediated conformational changes that alter the local environment of an attached fluorophore. Such conformational changes are found in many proteins; coupling these to fluorescent responses therefore provides a rich source for engineering fluorescent biosensors. For instance, the glucose-binding protein used in this example is a member of the bacterial periplasmic-binding protein (PBP) superfamily which combines a large diversity of ligand specificities with a common structural mechanism (Berntsson et al. 2010) that is well suited to the construction of fluorescent sensors (de Lorimier et al. 2002, Grunewald 2014). However, in these proteins, engineered fluorescent responses are more commonly monochromatic than dichromatic. This first example (FIG. 4A) therefore demonstrates how

ngmFRET can be used to improve the success rate for engineering ratiometric biosensors that exploit ligand-mediated protein motions. Non-limiting examples of ligand-binding proteins include proteins that bind sugars (such galactose-binding proteins, lactose-binding proteins, arabinose-binding proteins, ribose-binding proteins, and maltose-binding proteins), urea-binding proteins, bicarbonate-binding proteins, phosphate-binding proteins, sulfate-binding proteins, calcium-binding proteins, dipeptide-binding proteins, amino acid-binding proteins (such as histidine-binding proteins, glutamine-binding proteins, glutamate-binding proteins, and aspartate-binding proteins), and iron-binding proteins.

The second example (FIG. 4B) represents a smaller set of proteins that contain fluorophores formed by a self-catalyzed cyclization of a peptide within their sequences such as Green Fluorescent Protein, its engineered variants and homologs (Tsien 1998, Zimmer 2002). Some of these fluorophores function as direct natural chemosensors by interacting with ligands such as protons and halides (Miesenbock, De Angelis and Rothman 1998, Grimley et al. 2013), but typically evince only monochromatic responses. The example demonstrates how ratiometric sensing mechanisms can be engineered into these proteins.

The third example (FIG. 4C) illustrates how protein engineering was used to improve the properties of synthetic directly responsive chemosensors, by incorporating them as extrinsic fluorophores into a protein and combining with a second fluorophore to introduce ngmFRET. In this example, the protein therefore functions as an “adaptor”

scaffold/compound that enables facile integration of multiple functionalities. Although many small-molecule fluorescent chemosensors have been developed that measure a wide variety of ligands, prior to the invention it was challenging to engineer dichromatic responses. The use of an adaptor protein provides a facile route to integrate multiple fluorophores to generate a dichromatic sensor construct, enabling the power of monochromatic chemosensors to be harnessed and optimized.

In the third example, a chemoresponsive (directly responsive) fluorophore may be linked to another fluorophore (an indirectly responsive fluorophore) using virtually any polypeptide sequence. Though a pH-sensitive chemoresponsive fluorophore is exemplified in the examples, any other chemoresponsive fluorophore may be used. The principles demonstrated for converting the monochromatic response of fluorescein to protons into a dichromatic signal by incorporation into a dually labeled adaptor protein can be extended to other chemoresponsive fluorophores for the detection of a wide variety of analytes. A well-known thiol- or amine-reactive functional group (imidoesters, NHS esters, carbodiimides,

maleides, aziridines, arcyloyls; *see, e.g.*, G.T. Hermanson, 2013, *Bionjugate Techniques*, Academic Press, incorporated herein by reference) may be incorporated into the chemoresponsive fluorophore such that the chemoresponsive fluorophore can be coupled to the adaptor protein (M.S.T. Goncalves, 2009, *Chem. Rev.*, 190-212). Chemoresponsive sensors can bind specifically to small molecules such as ions, monosaccharides, amino acids, and short peptides, using a variety of molecular recognition units. Such units are then linked to a fluorescent group, the properties of which are altered upon binding the ligand (A. P. Demchenko, 2015, *Introduction to Fluorescence Sensing*, Springer). A variety of schemes can be used to couple proton-binding groups such as amines to fluorescence responses (op. cit.). Fluorescent chemosensors have been developed for toxic metals such as lead, cadmium, and mercury (K.P. Carter et al., 2014, *Chem. Rev.*, 114, 4564-4601). Chemoresponsive fluorophores have been developed for glucose (X.S. Sun, T.D. James, 2015, *Chem. Rev.*, 115, 8001-8037), and other organic analytes including amines, urea, and guanidinium (T.W. Bell and N.M. Hext, 2004, *Chem. Soc. Rev.*, 33, 589-598).

The fluorescent glucose sensor described in the first example has utility in glucose monitoring is essential for the management of diabetes mellitus, a disease that affects at least 366 million people world-wide (Yoo and Lee 2010, Cash and Clark 2010) and is increasing every year. The majority of current glucose-monitoring technologies rely on enzymes for which glucose is one of the substrates (Wang 2008, Bergel, Soupe and Comtat 1989). Glucose concentration measurements therefore are subject to variations in second substrate concentrations consumed in the enzyme reaction, such as oxygen in the case of glucose oxidase (Tang et al. 2001). Additional complications arise in systems where reaction rates are measured for enzymes immobilized on electrodes. In such arrangements, accuracy is compromised by factors that alter the rate at which glucose arrives at the electrode surface interfere with accuracy, such as hematocrit levels (Karon et al. 2008, Tang et al. 2000), or surface "fouling" by deposition of proteins and cells in the foreign body response (Koschwanez and Reichert 2007, Gifford et al. 2006, Wisniewski and Reichert 2000). Ratiometric fluorescent glucose sensors obviate these problems, and accordingly have been incorporated successfully in optodes for continuous glucose monitoring in animals and humans.

Determination of the concentration of chloride, an essential electrolyte, is a routine measurement in clinical chemistry carried out using potentiometric or optical methods (Huber 2001), and has applications to environmental chemistry (Huber 2000) and corrosion

control(de Graaf 2015) as well. Development of a fluorescent, ratiometric chloride biosensor, such as described in the second example enables chloride sensing to be incorporated into point-of-care devices and continuous monitoring systems.

The determination of proton concentrations is necessary for a broad range of applications, ranging from clinical chemistry to environmental science. Accordingly, the development of fluorescent and other optical probes is an active area(Han 2010, Wencel 2014).

Together these examples demonstrate that the ngmFRET mechanism is applicable to the engineering of a wide variety of semi-synthetic protein-based fluorescent biosensors by combining properties unique to proteins with those of synthetic fluorophores. Akin to natural cofactors, the fluorophores endow the proteins with functions that cannot be encoded in amino acids, and the proteins modulate the properties of the fluorophores. As illustrated by the analytes chosen in the examples, such engineered fluorescent biosensors have many potential applications, including medical diagnostics.

Mechanisms for Ligand Sensing using Non-Geometric Modulation of FRET

The subject matter disclosed herein is not limited to or bound by any particular scientific theory. However, the discussion below is provided to facilitate the understanding of possible mechanisms involved with ngmFRET signaling in various embodiments described herein. Equations for calculating various values mentioned herein are also provided.

The total signal, S , of a fluorescent sensor (either single-wavelength emission intensities, I_λ , or ratios of intensities at two wavelengths, R_{12}) is the sum of the fluorescence due to the ligand-free (apo) and ligand-bound states:

$$S = \alpha(1 - \bar{y}) + \beta\bar{y} \quad 1$$

where α and β are the fluorescent baselines in the ligand-free and -bound states, respectively, and $\bar{y}$ is the fractional occupancy of the binding sites. For a single binding site $\bar{y}$ is given by

$$\bar{y} = \frac{[L]}{[L] + K_d} \quad 2$$

where $[L]$ is the ligand (analyte) concentration and K_d the dissociation constant corresponding to an apparent, $^{app}K_d$, or true, $^{true}K_d$, value for fits to R_{12} and I_λ respectively.

Fluorescence quantum yields are the fractions of photons emitted by the excited state relative to the total absorbed, and correspond to the ratio of the radiative decay rate relative to the sum of the rates of all possible decay pathways (FIGS. 3A-D). For a single fluorophore:

$$Q = \frac{k_r}{k_r + k_{nr}} \quad 3$$

5 where k_r and k_{nr} are the radiative and non-radiative decay rates of the excited state, respectively. If we define q as the ratio between the radiative and non-radiative decay rates,

$$q = \frac{k_{nr}}{k_r} \quad 4$$

then the quantum yield can be written as

$$Q = \frac{1}{q + 1} \quad 5$$

10 Chemical sensors exploit the ligand-mediated shift of a fluorescent system between the ligand-free and ligand-bound states which each exhibit distinct quantum yields:

$$Q_{obs} = Q_{apo}(1 - \bar{y}) + Q_{sat}\bar{y} \quad 6$$

where Q_{obs} , Q_{apo} and Q_{sat} are the quantum yield of the total system, the apo-protein, and the ligand-bound complex, respectively. In a system involving energy transfer between a donor and acceptor fluorophore, the Q_{apo} and Q_{sat} quantum yields each are combinations of their
15 respective donor and acceptor quantum yields:

$$Q_{apo} = {}^DQ_{apo} + {}^AQ_{apo} \text{ and } Q_{sat} = {}^DQ_{sat} + {}^AQ_{sat} \quad 7$$

where the superscripts D and A indicate donor and acceptor fluorophores respectively. To understand ngmFRET-based sensors, we therefore need to examine the factors that affect
20 each of these four quantum yields.

The intensity of the light emitted by a donor or its acceptor is determined by the rate of photon emission from their respective excited states (FIG. 3A). The excited state of a donor is formed by the incident light from the excitation source, and there are three pathways by which this state decays: radiative and non-radiative decay and resonance transfer (by itself
25 and regardless of the presence of any other fluorophore/partner). By contrast, the rate of formation of the acceptor excited state is determined by the resonance transfer rate from the donor, and there are only two processes that determine its decay rate: the radiative and non-radiative pathways (by itself and regardless of the presence of any other fluorophore/partner). In an ngmFRET system, the patterns of ligand-mediated fluorescence intensity changes

therefore depend on whether the fluorophore that responds directly to ligand binding functions as a donor or acceptor. To understand these relationships, we analyzed the factors that determine the rates of formation and decay of the donor and acceptor excited states.

The rate of resonance energy transfer, k_t , along a non-radiative pathway between donor and acceptor (FIG. 3A) is a fraction of the donor radiative emission pathway rate (by itself and regardless of the presence of any other fluorophore/partner), ${}^Dk_r$ (the emission rate in the absence of an acceptor) multiplied by the energy transfer coupling factor, ϕ , (Lakowicz 2006, Valeur 2012):

$$k_t = \phi Q_D {}^Dk_r \quad 8$$

where Q_D is the donor quantum yield in the absence of an acceptor.

According to the Förster model of weakly coupled oscillators (Lakowicz 2006, Valeur 2012), the energy transfer coupling factor is dependent on the spectral overlap, J , of the donor emission, ${}^D\lambda_{em}$, and acceptor excitation spectrum, ${}^A\lambda_{ex}$, and the variation of the geometry, G , between the donor and acceptor excited state transition dipoles with distance, r , and

orientation factor, κ :

$$\phi = G(r, \kappa) J({}^D\lambda_{em}, {}^A\lambda_{ex}) \frac{9000 \ln 10}{128 \pi^5 N_A n^4} \quad 9$$

where

$$G(r, \kappa) = \frac{\kappa^2}{r^6} \quad 10$$

and

$$J({}^D\lambda_{em}, {}^A\lambda_{ex}) = \int F({}^D\lambda_{em}) \epsilon({}^A\lambda_{ex}) \lambda^4 d\lambda \quad 11$$

with n the refractive index of medium, N_A Avogadro's number, $F({}^D\lambda_{em})$ the normalized donor emission spectrum, and $\epsilon({}^A\lambda_{ex})$ the absorption coefficient of the acceptor excitation spectrum [this analysis is a re-arrangement of the traditional presentation of the equations describing tmFRET, separating the different contributions (geometry, spectral overlap, quenching)]. Ligand-mediated modulation of r , κ and J therefore affects k_t (Fig. 3b-d), leading to changes in donor and acceptor emission intensities (see below).

At steady state, the concentration of the donor excited state, $[D^*]$, is given by the following rate balance equation (see Fig. 3a):

$$N_0 \alpha k_{ex} - [D^*] ({}^Dk_{nr} + {}^Dk_r + k_t) = 0 \quad 12$$

where N_0 is the population of ground state fluorophores, k_{ex} the rate of excitation photon absorption, α the effective illumination, k_t , the resonance energy transfer rate, ${}^Dk_{nr}$ and ${}^Dk_r$ the radiative and non-radiative decay rates of the donor (by itself and regardless of the presence of any other fluorophore/partner) in the absence of acceptor, respectively.

- 5 Substituting ${}^Dk_r(d+1)$ for ${}^Dk_r+{}^Dk_{nr}$ (using equation 4, with $d \equiv q$, the ratio of non-radiative to radiative decay rates in the donor), and replacing k_t with equation 8 (with $Q_D = 1/(1+d)$), according to equation 5), we obtain

$$N_0 \alpha k_{ex} - [D^*] {}^Dk_r \left(1 + d + \frac{\varphi}{1+d} \right) = 0 \quad 13$$

Hence

$$10 \quad [D^*] = \frac{N_0 \alpha k_{ex}}{{}^Dk_r \left(1 + d + \frac{\varphi}{1+d} \right)} \quad 14$$

The intensity of the emitted donor light, I_D , is

$$I_D = [D^*] {}^Dk_r = \frac{N_0 \alpha k_{ex}}{\left(1 + d + \frac{\varphi}{1+d} \right)} \quad 15$$

The donor quantum yield, Q_D , is this emission intensity relative to the intensity of the excitation, $k_{ex} \alpha N_0$

$$15 \quad Q_D = \frac{1}{\left(1 + d + \frac{\varphi}{1+d} \right)} \quad 16$$

The rate balance equation for the acceptor excited state concentration, $[A^*]$, is given by

$$[D^*] k_t - [A^*] ({}^A k_r + {}^A k_{nr}) \quad 17$$

Consequently, by applying equations 5, 8 and 16, the acceptor quantum yield, Q_A , is

$$20 \quad Q_A = \frac{\varphi}{(1+a)(1+d) \left(1 + d + \frac{\varphi}{1+d} \right)} \quad 18$$

where a is the ratio of the radiative and non-radiative pathways in the acceptor.

The ratio of the acceptor and donor quantum yields therefore is

$$\frac{Q_A}{Q_D} = \frac{\varphi}{(1+d)(1+a)} \quad 19$$

This equation clearly shows that a ligand-mediated change in energy transfer (ϕ) or a change in the ratio of radiative to non-radiative emission rates of either the donor (d) or acceptor (a) leads to a change in the ratio of donor and acceptor emission intensities, thereby enabling ratiometry.

Classical ligand-mediated modulation of tgmFRET (traditional geometric FRET -) is concerned only with ligand-mediated changes in the distance between the donor and acceptor (Clegg 1995, Cheung 1991), and does not take advantage of effects that alter the photophysics of individual fluorophores. By contrast, in ngmFRET systems, the directly responsive partner (DRP) responds to ligand binding through ligand mediated changes that alter the ratio of its radiative and non-radiative pathways (quenching, d or a) or its spectral properties (J), whereas the indirectly responsive partner (IRP) changes only as a consequence of the effect that such change have on the resonance energy transfer rate (k_t). It is important to realize that the DRP can function either as a ngmFRET donor or an acceptor, depending on how the spectral overlap is set up with the IRP. Regardless of whether the DRP is a donor or acceptor, ligand-mediated alteration of its non-radiative to radiative decay rate ratio (parameter d for a DRP donor; a for an acceptor; by itself and regardless of the presence of any other fluorophore/partner) changes its emission intensity. In DRP donors quenching also alters the ngmFRET transfer rate (see equations 8 and 13), thereby changing the emission intensities of not only itself but also its IRP. By contrast, in DRP acceptors quenching does not alter ngmFRET, and hence do not affect its IRP donor intensity. A DRP acceptor therefore can alter intensities of its donor IRP only if ligand binding changes ϕ . If the DRP is a donor, then manipulation of the ngmFRET coupling factor, ϕ , changes the rate of excited state *decay*; if it is an acceptor, the rate of excited state *formation* is altered.

Regardless of whether the DRP is a donor or acceptor, a change in any of the two parameters (ϕ and d or a) alters the ratio of the donor and acceptor quantum yields (equation 19), thereby enabling ratiometry. Ligand-mediated donor DRP quenching affects the quantum yields of both the donor, Q_D , and acceptor, Q_A , quantum yields (equations 16, 18). Quenching of an acceptor DRP alters only Q_A (equation 16). Changes in ϕ affect quantum yields of both fluorophores, regardless whether the DRP functions as the donor or acceptor (equations 9-11, 16, 18). For systems in which there is no ligand-mediated change in the (average) distance between the two fluorophores, ϕ changes only if the DRP switches between two different excited state populations ("dipole switching") in response to ligand

binding and if the two excited states differ in their spectral properties (emission for donor DRPs; absorption for acceptor DRPs). Excited state dipoles usually also differ in their dipole orientations, so it is likely that changes in spectral overlap involve (re-)orientation effects. They are also likely to differ in the relative rates of their radiative and non-radiative decay rates. Dipole switching therefore is likely to involve a combination of changes in ngmFRET and quenching effects.

There are eight possible combinations of ligand-mediated changes in quenching and ngmFRET parameters, which have different outcomes on the two emission intensities and their ratio, depending on whether the DRP is the donor or acceptor. The qualitative behavior of the resulting sixteen possibilities in ngmFRET systems are shown in Table 1. Twelve of these have a predictable outcome on the direction of change in the ratio of the two emission intensities. The effect on the direction of change for both donor and acceptor emission intensities can be predicted for seven models. For the other models, the direction of change of one or both peaks depends on the size of the change in the underlying parameters. Purely geometric effects (changes in inter-dipole distance or orientation) always result in anti-correlated changes in emission intensity changes (*i.e.* one increases and the other decreases, or *vice versa*). Correlated (*i.e.* both intensities increase or decrease) or uncorrelated (one changes, the other remains constant) intensity changes therefore are *prima facie* evidence for an ngmFRET effect.

Table 1. Qualitative analysis of the patterns of donor and acceptor emission intensity changes in ngmFRET^a

Directly responsive partner	Model	Q_A/Q_D	Q_D	Q_A
Donor	$d^0\phi^+$	↑	↓	↑
	$d^0\phi^-$	↓	↑	↓
	$d^+\phi^0$	↓	↓	↓
	$d^+\phi^+$	*	↓	*
	$d^+\phi^-$	↓	*	↓
	$d^-\phi^0$	↑	↑	↑
	$d^-\phi^+$	↑	*	↑
	$d^-\phi^-$	*	↑	*
Acceptor	$a^0\phi^+$	↑	↓	*
	$a^0\phi^-$	↓	↑	*
	$a^+\phi^0$	↓	0	↓
	$a^+\phi^+$	*	↓	*
	$a^+\phi^-$	↓	↑	*
	$a^-\phi^0$	↑	0	↑
	$a^-\phi^+$	↑	↓	↑
	$a^-\phi^-$	*	↑	*

^aThe effects of increasing or decreasing quenching in the directly responsive ngmFRET partner (d for donors, a for acceptors) or the energy transfer coupling (ϕ) between the donor and acceptor are tabulated. The consequences of using a directly responsive donor or acceptor are examined. Changes in quenching and energy transfer coupling parameters can occur singly or in combination, leading to 16 possible models. The models examine the effects of the direction of change in quenching parameters (no change, d^0 or a^0 ; increase d^+ or a^+ ; decrease, d^- or a^-) and the energy transfer coupling factor (no change, ϕ^0 ; increase, ϕ^+ ; decrease, ϕ^-) on the patterns in the direction of change of the donor, Q_D (equation 16) or acceptor, Q_A (equation 18) quantum yields, and their ratio, Q_A/Q_D (equation 19): ↑, increase; ↓, decrease; 0, no change; *, response is dependent on precise quantitation rather than direction of change in the underlying parameter values.

Exemplary Dual Labeling Techniques to Construct Donor-Acceptor Pairs

The directly responsive fluorophore needs to be site-specifically attached to a site in the protein where it can respond to the ligand-binding event (the one exception is the case where the directly responsive fluorophore is integral to the protein itself, as is the case for YFP). Site-specific attachment of the second, passive fluorophore also is desirable, because it establishes a unique ngmFRET interaction, whereas random attachment gives rise to multiple interactions, only some of which are likely to be result in usable signals with the others contributing to background.

Cysteine thiols provide a convenient, chemically unique functionality for site-specific attachments at positions defined by the protein sequence. Furthermore, straightforward, reversible protection strategies have been developed that enable multiple thiols to be labeled orthogonally at independently addressable sites in successive reactions (Smith et al. 2005).

5 In this scheme, two labeling sites can be engineered using fusions that combine two proteins or domains containing a single thiol and disulfide bridge, respectively (FIG. 5). In such constructs, labeling proceeds in two steps. First, the unprotected thiol is modified with the first fluorophore under conditions in which the disulfide bridge is fully formed protecting the second thiols. After this labeling step is complete, unincorporated first fluorophore is
10 removed from the reaction. In the second labeling step the thiols in the disulfide bridge are deprotected by reduction, enabling attachment of the ngmFRET partner fluorophore at their positions.

Typically, the single cysteine is located in a protein that binds the analyte of interest (e.g. the periplasmic glucose-binding protein) for coupling of the directly responsive
15 fluorophore, and the disulfide is located in a small domain or protein fused at the N- or C-terminus for attachment of the indirectly responsive ngmFRET partner. In the examples shown here we have used an 17-residue peptide, β Zif, derived from a zinc finger protein (Smith et al. 2005) as the disulfide-containing domain. The disulfide-containing protein thioredoxin from *Escherichia coli* (Holmgren 1985, Qi 2005, Katti 1990), ecTRX, was
20 also used as a fusion partner. The ecTRX also has been used to construct the adaptor protein by installing a single cysteine mutation that introduces an unprotected thiol on the surface of the protein (see below).

Thiols are not the only means to establish orthogonal chemistries. For example, amines also have a highly selective chemistry. To demonstrate this approach, an unusual
25 version of ecTRX was created in which the amino acid alphabet was limited to 19 residues, eliminating all lysines. This leaves only one reactive primary amine in the protein: the amino terminus. In this manner, constructed a doubly labeled protein by adding the first label to the amino terminus, and the second to the reduced disulfide bridge. A label may be conjugated to ecTRX using a variety of linkers or bonds, including (but not limited to) a disulfide bond,
30 an ester bond, a thioester bond, an amide bond, or a bond that has been formed by a click reaction.

Biosensors

Biosensors are molecular recognition elements that transduce ligand-binding events into physical detectable signals. Biosensors as detailed herein bind at least one ligand and emit a signal. A ligand-bound biosensor results in a signal that is different from the unbound biosensor. This difference facilitates detection of the at least one ligand and/or determination of ligand concentration. The biosensors may be used alone, *i.e.*, without the presence or assistance of other reagents.

Described herein are reagentless biosensors engineered to produce a detectable ratiometric detection signal. These biosensors may have altered ligand-binding affinities, tailored ligand-binding specificities, and/or temperature dependencies of ligand binding or stability compared to corresponding naturally occurring ligand-binding proteins. For example, the herein described engineered ligand biosensors provide high-accuracy information related to extended ligand concentration ranges.

Binding of ligand mediates conformational changes in the biosensor, such as hinge-bending motions of the polypeptide. The conformational changes affect the environment of the reporter such that a change in the reporter-generated signal occurs. That is, without ligand bound, the biosensor results in signal generated from the reporter, and when ligand is bound, the signal generated from the reporter changes. The ligand-bound biosensor results in a reporter-generated signal that is different from the unbound biosensor.

Among the advantages of these fluorophore-containing protein constructs is their high durability. The constructs retain their ability to bind ligand, change shape and thus detect the ligand (a) even when immobilized (directly or indirectly) onto a solid surface such as a bead, plate, or sheet; (b) even after desiccation (and subsequent reconstitution in a physiological buffer solution); (c) even when subjected to ambient conditions, *e.g.*, conditions that can be encountered in storage and/or transportation; and (d) even when aged/stored for extended periods of time, *e.g.*, weeks, months, or even years. Thus, the biosensors do not require refrigeration or a cold chain for distribution, permitting a wider range of applicability such as in-the-field use and reducing the cost of the sensor product.

For clinical applications, microliter volumes, *e.g.*, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, μ l or less of a bodily fluid may be used. In some embodiments, the volume of sample that is applied to a biosensor or a device comprising a biosensor is less than 0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 50, 75, 100, 150, 300, 500, or 1000 μ l. In some embodiments, the volume is about 0.1 μ l to about 1000 μ l, about 0.1 μ l to about 100 μ l, about 1 μ l to about 1000 μ l, about 1 μ l to

about 10 μ l, about 1 μ l to about 100 μ l, about 1 μ l to about 50 μ l, about 10 μ l to about 50 μ l, or about 5 μ l to about 50 μ l. Moreover compared to conventional enzyme-based or antibody based assay systems, the results are achieved virtually instantaneously, *e.g.*, within 30-60 seconds. A further advantage is that the sensors consistently and reliably bind to and detect
5 the analyte in complex fluids such as whole blood. Thus in a clinical setting, whole blood need not be processed, thereby reducing time and cost of the diagnostic procedure.

In non-clinical situations, *e.g.*, industrial or commercial settings such as analysis of waste water, food or beverage production, or bioreactor/fermentation monitoring, the samples to be analyzed can be used directly upon sampling without further purification or processing,
10 similarly reducing time and expense of the test. Moreover, the immobilized sensors need not be washed to remove unbound material following contacting the test sample with the sensors, because the unbound material (“contaminants”) do not materially affect the production of a precise, reliable detectable assay signal.

In some embodiments, the methods and compositions include a plurality of a
15 single type of biosensor. The biosensors may be identical in structure and function. For example, the biosensors of a single type may have the same polypeptide, the same reporter, and the same ligand affinity.

In other embodiments, the methods and compositions include a plurality of different types of biosensors. A plurality of these different types of biosensors may be
20 arranged or incorporated in a panel. As used herein, a “panel” refers to two or more biosensors. The two or more biosensors may be different from each other. The biosensors may differ in structure and/or function. Biosensors may differ in polypeptide sequence, reporter, ligand affinities, or a combination thereof. Accordingly, there may be different types of biosensors. In some embodiments, each biosensor in the panel comprises the same reporter
25 group. In some embodiments, each biosensor in the panel comprises a different reporter group. The panel may include at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least
30 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 90, at least 95, or at least 100 biosensors.

The panel of biosensors includes at least one sensor element. “Sensor element”

refers to a single spot, site, location, or well for the at least one biosensor, to which a sample or aliquot thereof may be applied. The panel may be a composite sensor or an array.

In some embodiments, the panel is a composite sensor. In a composite sensor, each sensor element includes a mixture of two or more different biosensors. In some
5 embodiments, the composite sensor includes one sensor element. In some embodiments, the composite sensor includes two or more sensor elements. In some embodiments, signals are measured from a composite sensor in which the signals arise from one or more biosensors in the sensor element. For example, signals may be measured from a composite sensor in which the signals arise from a subset of the total number of biosensors in the sensor element. For
10 example, signals may be measured from a composite sensor in which the signals arise from two of five biosensors in the sensor element.

In some embodiments, the panel is an array. In an array, each sensor element includes a single type of biosensor. An array comprises a plurality of individually and spatially localized sensor elements. Each sensor element includes a biosensor that is different
15 than or the same as the biosensor of a different sensor element. In some embodiments, signals are measured from an array in which the signals arise separately from two or more selected biosensors in separate sensor elements. An array may comprise a plurality of sensor elements of a variety of sizes and configurations. An array may comprise a plurality of sensor elements arranged linearly. For example, an array may comprise a plurality of micrometer-
20 sized sensor elements arranged in a single row. An array may comprise a plurality of sensor elements arranged in a grid. The grid may be two- or three-dimensional. In some embodiments, the grid is a spatially addressable grid. In some embodiments, the biosensors are incorporated into an array, such as a multichannel or multiplexed array.

The biosensors of the present disclosure can be used in any setting where
25 ligand detection is required or desired, such a medical setting (e.g., determining the level of blood ligand in a subject), environmental setting (e.g., determining the level of ligand in an environmental sample), biological setting (e.g., determining the presence or amount of ligand in a reaction), or in process engineering, such as monitoring the amount of ligand in a fermentation reaction (e.g., beer/wine production, etc.). Other examples include, but are not
30 limited to, uses in the food industry (Suleiman et al, In: Biosensor Design and Application: Mathewson and Finley Eds; American Chemical Society, Washington, DC 1992, vol. 511); in clinical chemistry (Wilkins et al., Med. Eng. Phys. 1996, 18, 273-288; Pickup, Tr. Biotech. 1993, 11, 285-291; Meyerhoff et al., Endricon 1966, 6, 51-58; Riklin et al., Nature 1995, 376,

672-675); Willner et al., J. Am. Chem. Soc. 1996, 118, 10321-10322); as the basis for the construction of a fluorescent flow cell containing immobilized ligand binding protein-FAST conjugates (see, e.g., Wilkins et al., Med. Eng. Phys. 1966, 18, 273-288; Pickup, Tr. Biotech. 1993, 11, 285-291; Meyerhoff et al., Endricon. 1966, 6, 51; Group, New Engl. J. Med. 1993, 329, 977-986; Gough et al., Diabetes 1995, 44, 1005-1009); and in an implantable devices.

The biosensors as detailed herein may be administered in a variety of ways known by those of skill in the art, as appropriate for each application. Biosensors may be provided in a solution. The solution may be buffered. Biosensors may be provided in a solution and mixed directly with a sample. In some embodiments, a biosensor is immobilized onto a surface. Biosensors may be immobilized within a disposable cartridge into which a sample may be introduced or applied. Biosensors may be implanted or incorporated in a wearable device. The biosensor may be provided as an optode.

The biosensor may be attached to or incorporated in a wearable device. Wearable devices may include, for example, adhesive strips, patches, and contact lenses. The biosensor may be configured for placement in contact with a subject's skin or mucosal surface. In some embodiments, the biosensor is configured as an adhesive strip. In some embodiments, the biosensor is configured within or on the surface of a contact lens. In some embodiments, the contact lens is formed from a transparent substrate shaped to be worn directly over a subject's eye, as described in, for example, U.S. Patent No. 8,608,310.

The biosensor may be implanted. The biosensor may be implanted in a subject's body. The biosensor may be implanted in a subject's blood vessel, vein, eye, natural or artificial pancreas, skin, or anywhere in the alimentary canal including the stomach, intestine and esophagus. The biosensor may be implanted in a subject with a microbead. In some embodiments, the biosensor is configured to be implanted in the skin. The biosensor may be implanted in a subject sub-dermally. The biosensor may generate the signal trans-dermally. In some embodiments, the biosensor may be implanted in a subject with transdermal microbeads, wherein the optical signals can be transmitted remotely between the biosensor and detecting device.

In some embodiments, the biosensor is administered as an optode. As used herein, "optode" refers to an optical fiber with a single biosensor, or a composite biosensor, immobilized at the surface or at the end. An "optode" may also be referred to as an "optrode." In some embodiments, the biosensor is implanted in a subject as an optode. The optode may be incorporated with or into a needle. The optode may be incorporated with a

probe such as endoscopy or colonoscopy probes. The optode may be used in a tumor, near a tumor, or at the periphery of a tumor. In some embodiments, the biosensor may be implanted in a subject as an optode, wherein the optical signals can be transmitted between the biosensor and detecting device using physical links. In some embodiments, the biosensor is administered as an optode to a sample or reaction. The optode may be contacted with a sample or reaction. In some embodiments, an optode is used to continuously or episodically monitor a ligand in a sample or reaction.

Methods Of Detecting The Presence Of A Ligand

Provided herein is a method of detecting the presence of a ligand in a sample. The method may include contacting the biosensor with the sample; measuring a signal from the biosensor; and comparing the signal to a ligand-free control. A difference in signal indicates the presence of ligand in the sample.

Also provided herein is a method of detecting the presence of ligand in a sample. The method may include (a) providing a ligand biosensor disclosed herein in which the reporter group is attached the ligand so that a signal transduced by the reporter group when the ligand is bound to ligand differs from a signal transduced by the reporter group when the ligand is not bound to ligand; (b) contacting the biosensor with the test sample under conditions such that the biosensor can bind to ligand present in the test sample; and (c) comparing the signal transduced by the reporter group when the biosensor is contacted with the test sample with the signal transduced by the reporter group when the biosensor is contacted with a ligand-free control sample, wherein a difference in the signal transduced by the reporter group when the biosensor is contacted with the test sample, as compared to when the biosensor is contacted with the control sample, indicates that the test sample contains ligand.

Methods Of Determining The Concentration Of A Ligand

Provided herein is a method of determining the concentration of a ligand in a sample. The method may include contacting the biosensor with the sample; measuring a signal from the biosensor; and comparing the signal to a standard hyperbolic ligand binding curve to determine the concentration of ligand in the test sample. The standard hyperbolic ligand binding curve may be prepared by measuring the signal transduced by the biosensor when contacted with control samples containing known concentrations of ligand.

Another aspect of the present disclosure provides a method of determining the concentration of ligand in a test sample comprising, consisting of, or consisting essentially of: (a) providing a ligand biosensor comprising a ligand biosensor as described herein in which the reporter group is attached the ligand so that a signal transduced by the reporter group when the ligand is bound to ligand differs from a signal transduced by the reporter group when the ligand is not bound to ligand; (b) contacting the biosensor with the test sample under conditions such that the biosensor can bind to ligand present in the test sample; and (c) comparing the signal transduced by the reporter group when the biosensor is contacted with the test sample with a standard hyperbolic ligand binding curve prepared by measuring the signal transduced by the reporter group when the biosensor is contacted with control samples containing known quantities of ligand to determine the concentration of ligand in the test sample.

Methods Of Monitoring The Presence Of A Ligand

The present invention is directed to a method of episodically or continuously monitoring the presence of a ligand in a reaction. In certain embodiments, the biosensors may be used in the continuous monitoring of ligand in a reaction. In certain embodiments, the ligand sensors may be used in episodic monitoring of sample aliquots.

The method of episodically or continuously monitoring the presence of a ligand in a reaction may include contacting the biosensor with the reaction; maintaining the reaction under conditions such that the polypeptide is capable of binding ligand present in the reaction; and episodically or continuously monitoring the signal from the biosensor in the reaction.

The method of episodically or continuously monitoring the presence of a ligand in a reaction may include contacting the biosensor with the reaction; maintaining the reaction under conditions such that the polypeptide is capable of binding ligand present in the reaction; episodically or continuously monitoring the signal from the biosensor in the reaction; and comparing the signal to a standard hyperbolic ligand binding curve to determine the concentration of ligand in the test sample. The standard hyperbolic ligand binding curve may be prepared by measuring the signal transduced by the biosensor when contacted with control samples containing known concentrations of ligand.

In some embodiments, the method further includes comparing the signal to a ligand-free control, wherein a difference in signal indicates the presence of ligand in the reaction.

In some embodiments, the method further includes comparing the signal to a standard hyperbolic ligand binding curve to determine the concentration of ligand in the test sample. The standard hyperbolic ligand binding curve may be prepared by measuring the signal transduced by the biosensor when contacted with control samples containing known
5 concentrations of ligand.

Another aspect of the present disclosure provides a method of continuously monitoring the presence of ligand in a reaction comprising, consisting of, or consisting essentially of: (a) providing a ligand biosensor as described herein in which the reporter group is attached the ligand-binding protein so that a signal transduced by the reporter group
10 when the ligand is bound to ligand-binding protein differs from a signal transduced by the reporter group when the ligand is not bound to the ligand-binding protein; (b) maintaining the biosensor within the reaction and under conditions such that the biosensor can bind to ligand present in the reaction; (c) continuously monitoring the signal transduced by the reporter group when the biosensor is contacted with the ligand present in the reaction; and optionally
15 (d) comparing the signal transduced by the reporter group when the biosensor is contacted with the ligand present in the reaction with the signal transduced by the reporter group when the biosensor is contacted with a ligand-free control sample, wherein a difference in the signal transduced by the reporter group when the biosensor is contacted with the ligand present in the reaction, as compared to when the biosensor is contacted with the control
20 sample, indicates ligand is present in the reaction.

Yet another aspect of the present disclosure provides a method of continuously monitoring the concentration of ligand in a reaction comprising, consisting of, or consisting essentially of: (a) providing a ligand biosensor comprising a ligand biosensor as described herein in which the reporter group is attached the ligand so that a signal transduced by the
25 reporter group when the ligand is bound to ligand differs from a signal transduced by the reporter group when the ligand is not bound to ligand; (b) maintaining the biosensor within the reaction under conditions such that the biosensor can bind to ligand present in the reaction; and (c) continuously monitoring the signal transduced by the reporter group when the biosensor is contacted with the ligand present in the reaction; and (d) comparing the
30 signal transduced by the reporter group when the biosensor is contacted with the ligand present in the reaction with a standard hyperbolic ligand binding curve prepared by measuring the signal transduced by the reporter group when the biosensor is contacted with

control samples containing known quantities of ligand to determine the concentration of ligand in the reaction.

General Definitions

5 Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (*e.g.*, in cell culture, molecular genetics, and biochemistry).

As used herein, the term “about” in the context of a numerical value or range means $\pm 10\%$ of the numerical value or range recited or claimed, unless the context requires a more
10 limited range.

In the descriptions above and in the claims, phrases such as “at least one of” or “one or more of” may occur followed by a conjunctive list of elements or features. The term “and/or” may also occur in a list of two or more elements or features. Unless otherwise implicitly or explicitly contradicted by the context in which it is used, such a phrase is
15 intended to mean any of the listed elements or features individually or any of the recited elements or features in combination with any of the other recited elements or features. For example, the phrases “at least one of A and B;” “one or more of A and B;” and “A and/or B” are each intended to mean “A alone, B alone, or A and B together.” A similar interpretation is also intended for lists including three or more items. For example, the phrases “at least one
20 of A, B, and C;” “one or more of A, B, and C;” and “A, B, and/or C” are each intended to mean “A alone, B alone, C alone, A and B together, A and C together, B and C together, or A and B and C together.” In addition, use of the term “based on,” above and in the claims is intended to mean, “based at least in part on,” such that an unrecited feature or element is also permissible

25 It is understood that where a parameter range is provided, all integers within that range, and tenths thereof, are also provided by the invention. For example, “0.2-5 mg” is a disclosure of 0.2 mg, 0.3 mg, 0.4 mg, 0.5 mg, 0.6 mg etc. up to and including 5.0 mg.

A small molecule is a compound that is less than 2000 daltons in mass. The molecular mass of the small molecule is preferably less than 1000 daltons, more preferably
30 less than 600 daltons, *e.g.*, the compound is less than 500 daltons, 400 daltons, 300 daltons, 200 daltons, or 100 daltons.

As used herein, an “isolated” or “purified” nucleic acid molecule, polynucleotide, polypeptide, or protein, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. Purified compounds are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. For example, a purified compound is one that is at least 90%, 91%, 92%, 93%, 94%, 95%, 98%, 99%, or 100% (w/w) of the desired compound by weight. Purity is measured by any appropriate standard method, for example, by column chromatography, thin layer chromatography, or high-performance liquid chromatography (HPLC) analysis. A purified or isolated polynucleotide (ribonucleic acid (RNA) or deoxyribonucleic acid (DNA)) is free of the genes/nucleic acids or sequences/amino acids that flank it in its naturally-occurring state. Purified also defines a degree of sterility that is safe for administration to a human subject, e.g., lacking infectious or toxic agents.

Similarly, by “substantially pure” is meant a nucleotide or polypeptide that has been separated from the components that naturally accompany it. Typically, the nucleotides and polypeptides are substantially pure when they are at least 60%, 70%, 80%, 90%, 95%, or even 99%, by weight, free from the proteins and naturally-occurring organic molecules with they are naturally associated.

The transitional term “comprising,” which is synonymous with “including,” “containing,” or “characterized by,” is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. By contrast, the transitional phrase “consisting of” excludes any element, step, or ingredient not specified in the claim. The transitional phrase “consisting essentially of” limits the scope of a claim to the specified materials or steps “and those that do not materially affect the basic and novel characteristic(s)” of the claimed invention.

“Subject” as used herein refers to any organism from which a biological sample is obtained. For example, the sample is a biological fluid or tissue. For example, a subject is one who wants or is in need of detecting ligand or determining the concentration of ligand with the herein described biosensors. The subject may be a human or a non-human animal. The subject may be a mammal. The mammal may be a primate or a non-primate. The mammal can be a primate such as a human; a non-primate such as, for example, dog, cat, horse, cow, pig, mouse, rat, camel, llama, goat, rabbit, sheep, hamster, and guinea pig; or

non-human primate such as, for example, monkey, chimpanzee, gorilla, orangutan, and gibbon. The subject may be of any age or stage of development, such as, for example, an adult, an adolescent, or an infant.

As used herein, an “expression vector” is a DNA or RNA vector that is capable of effecting expression of one or more polynucleotides. Preferably, the expression vector is also capable of replicating within the host cell. Expression vectors can be either prokaryotic or eukaryotic, and are typically include plasmids. Expression vectors of the present invention include any vectors that function (*i.e.*, direct gene expression) in host cells of the present invention, including in one of the prokaryotic or eukaryotic cells described herein, *e.g.*, gram-positive, gram-negative, pathogenic, non-pathogenic, commensal, cocci, bacillus, or spiral-shaped bacterial cells; archaeal cells; or protozoan, algal, fungi, yeast, plant, animal, vertebrate, invertebrate, arthropod, mammalian, rodent, primate, or human cells. Expression vectors of the present invention contain regulatory sequences such as transcription control sequences, translation control sequences, origins of replication, and other regulatory sequences that are compatible with the host cell and that control the expression of a polynucleotide. In particular, expression vectors of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control sequences are those which control transcription initiation such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art.

As used herein, the singular forms “a,” “an,” and “the” include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to “a disease,” “a disease state”, or “a nucleic acid” is a reference to one or more such embodiments, and includes equivalents thereof known to those skilled in the art and so forth.

As used herein, “pharmaceutically acceptable” carrier or excipient refers to a carrier or excipient that is suitable for use with humans and/or animals without undue adverse side effects (such as toxicity, irritation, and allergic response) commensurate with a reasonable benefit/risk ratio. It can be, *e.g.*, a pharmaceutically acceptable solvent, suspending agent or vehicle, for delivering the instant compounds to the subject.

The term “diagnosis” refers to a determination that a disease is present in the subject. Similarly, the term “prognosis” refers to a relative probability that a certain future outcome may occur in the subject. For example, in the context of the present disclosure, prognosis can refer to the likelihood that an individual will develop a disease, or the likely severity of the disease (*e.g.*, severity of symptoms, rate of functional decline, survival, *etc.*).

Unless required otherwise by context, the terms “polypeptide” and “protein” are used interchangeably.

A polypeptide or class of polypeptides may be defined by the extent of identity (% identity) of its amino acid sequence to a reference amino acid sequence, or by having a greater % identity to one reference amino acid sequence than to another. A variant of any of genes or gene products disclosed herein may have, *e.g.*, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to the nucleic acid or amino acid sequences described herein. The term “% identity,” in the context of two or more nucleic acid or polypeptide sequences, refers to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using a sequence comparison algorithm or by visual inspection. For example, % identity is relative to the entire length of the coding regions of the sequences being compared, or the length of a particular fragment or functional domain thereof. Variants as disclosed herein also include homologs, orthologs, or paralogs of the genes or gene products described herein. In some embodiments, variants may demonstrate a percentage of homology or identity, for example, at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity conserved domains important for biological function, *e.g.*, in a functional domain, *e.g.* a ligand-binding or catalytic domain.

For sequence comparison, one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters. Percent identity is determined using BLAST. For the BLAST searches, the following parameters were employed: (1) Expect threshold is 10; (2) Gap cost is Existence:11 and Extension:1; (3) The Matrix employed is BLOSUM62; (4) The filter for low complexity regions is “on.”

The present invention also provides for functional fragments of the genes or gene products described herein. A fragment of a protein is characterized by a length (number of amino acids) that is less than the length of the full length mature form of the protein. A fragment, in the case of these sequences and all others provided herein, may be a part of the whole that is less than the whole. Moreover, a fragment ranges in size from a single nucleotide or amino acid within a polynucleotide or polypeptide sequence to one fewer nucleotide or amino acid than the entire polynucleotide or polypeptide sequence. Finally, a fragment is defined as any portion of a complete polynucleotide or polypeptide sequence that is intermediate between the extremes defined above.

For example, fragments of any of the proteins or enzymes disclosed herein or encoded by any of the genes disclosed herein can be 10 to 20 amino acids, 10 to 30 amino acids, 10 to 40 amino acids, 10 to 50 amino acids, 10 to 60 amino acids, 10 to 70 amino acids, 10 to 80 amino acids, 10 to 90 amino acids, 10 to 100 amino acids, 50 to 100 amino acids, 75 to 125 amino acids, 100 to 150 amino acids, 150 to 200 amino acids, 200 to 250 amino acids, 250 to 300 amino acids, or 300 to 350 amino acids. The fragments encompassed in the present subject matter comprise fragments that retain functional fragments. As such, the fragments preferably retain the binding domains that are required or are important for functional activity. Fragments can be determined or generated by using the sequence information herein, and the fragments can be tested for functional activity using standard methods known in the art. For example, the encoded protein can be expressed by any recombinant technology known in the art and the binding activity of the protein can be determined.

As used herein a "biologically active" fragment is a portion of a polypeptide which maintains an activity of a full-length reference polypeptide. Biologically active fragments as used herein exclude the full-length polypeptide. Biologically active fragments can be any size as long as they maintain the defined activity. Preferably, the biologically active fragment maintains at least 10%, at least 50%, at least 75% or at least 90%, of the activity of the full length protein.

Amino acid sequence variants/mutants of the polypeptides of the defined herein can be prepared by introducing appropriate nucleotide changes into a nucleic acid defined herein, or by *in vitro* synthesis of the desired polypeptide. Such variants/mutants include, for example, deletions, insertions or substitutions of residues within the amino acid sequence. A combination of deletion, insertion and substitution can be made to arrive at the final

construct, provided that the final peptide product possesses the desired activity and/or specificity.

Mutant (altered) peptides can be prepared using any technique known in the art. For example, a polynucleotide defined herein can be subjected to *in vitro* mutagenesis or DNA shuffling techniques as broadly described by Harayama (1998). Products derived from mutated/altereD DNA can readily be screened using techniques described herein to determine if they possess, for example, ligand binding activity.

In designing amino acid sequence mutants, the location of the mutation site and the nature of the mutation will depend on characteristic(s) to be modified. The sites for mutation can be modified individually or in series, *e.g.*, by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting other residues adjacent to the located site.

Amino acid sequence deletions generally range from about 1 to 15 residues, more preferably about 1 to 10 residues and typically about 1 to 5 contiguous residues. In some embodiments, a mutated or modified protein does not comprise any deletions or insertions. In various embodiments, a mutated or modified protein has less than about 10, 9, 8, 7, 6, 5, 4, 3, or 2 deleted or inserted amino acids.

Substitution mutants have at least one amino acid residue in the polypeptide molecule removed and a different residue inserted in its place. Sites may be substituted in a relatively conservative manner in order to maintain activity and/or specificity. Such conservative substitutions are shown in the table below under the heading of "exemplary substitutions."

In certain embodiments, a mutant/variant polypeptide has only, or not more than, one or two or three or four conservative amino acid changes when compared to a naturally occurring polypeptide. Details of conservative amino acid changes are provided in the table below. As the skilled person would be aware, such minor changes can reasonably be predicted not to alter the activity of the polypeptide when expressed in a recombinant cell.

Exemplary Substitutions

<u>Original Residue</u>	<u>Exemplary Substitutions</u>
Alanine (Ala)	Val; Leu; Ile; Gly
Arginine (Arg)	Lys
Asparagine (Asn)	Gln; His
Cysteine (Cys)	Ser
Glutamine (Gln)	Asn; His
Glutamic Acid (Glu)	Asp
Glycine (Gly)	Pro; Ala
Histidine (His)	Asn; Gln
Isoleucine (Ile)	Leu; Val; Ala
Leucine (Leu)	Ile; Val; Met; Ala; Phe
Lysine (Lys)	Arg
Methionine (Met)	Leu; Phe
Phenylalanine (Phe)	Leu; Val; Ala
Proline (Pro)	Gly
Serine (Ser)	Thr
Threonine (Thr)	Ser
Tryptophan (Trp)	Tyr
Tyrosine (Tyr)	Trp; Phe
Valine (Val)	Ile; Leu; Met; Phe; Ala

Mutations can be introduced into a nucleic acid sequence such that the encoded amino acid sequence is altered by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. Certain amino acids have side chains with more than one classifiable characteristic. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, tryptophan,

cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tyrosine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a given polypeptide is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a given coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for given polypeptide biological activity to identify mutants that retain activity. Conversely, the invention also provides for variants with mutations that enhance or increase the endogenous biological activity. Following mutagenesis of the nucleic acid sequence, the encoded protein can be expressed by any recombinant technology known in the art and the activity/specificity of the protein can be determined. An increase, decrease, or elimination of a given biological activity of the variants disclosed herein can be readily measured by the ordinary person skilled in the art, *i.e.*, by measuring the capability for binding a ligand and/or signal transduction.

In various embodiments, a polypeptide comprises mutations such that 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, or less than about 10, 9, 8, 7, 6, 5, 4, 3, or 2 amino acids is substituted with a cysteine and/or a lysine.

Polypeptides can be produced in a variety of ways, including production and recovery of natural polypeptides or recombinant polypeptides according to methods known in the art. In one embodiment, a recombinant polypeptide is produced by culturing a cell capable of expressing the polypeptide under conditions effective to produce the polypeptide, such as a host cell defined herein.

Key to the Sequence Listing

SEQ ID NO	Sequence Name
1	CIBP1, also referred to herein as laYFP
2	CIBP2
3	CIBP3
4	CIBP4
5	CIBP5
6	CIBP6
7	CIBP7
8	CIBP8

9	CIBP9
10	CIBP10
11	CIBP11
12	CIBP12
13	CIBP13
14	CIBP14
15	ttGGBP.11C.0.bZif (11C substitution mutant with bZif fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
16	ttGGBP.17C.0.bZif (17C substitution mutant with bZif fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
17	ttGGBP.111C.0.bZif (111C substitution mutant with bZif fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
18	ttGGBP.151C.0.bZif (151C substitution mutant with bZif fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
19	ttGGBP.182C.0.bZif (182C substitution mutant with bZif fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
20	ttGGBP.17C.3.Trx (17C, 16N, 211A substitution mutant with ecTRX fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
21	Trx.ttGGBP.17C.3 (17C, 16N, 211A substitution mutant with ecTRX fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
22	ttGGBP.182C.2.Trx (182C, 69P, 152P substitution mutant with ecTRX fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
23	Trx.ttGGBP.182C.2 (182C, 69P, 152P substitution mutant with ecTRX fusion, signal peptide replaced with M and a GGSHHHHHH at C-terminus)
24	Adaptor0
25	Adaptor1.0
26	Adaptor2.0a
27	Adaptor2.0b
28	Adaptor3.0
29	Adaptor4.0
30	Adaptor5.0
31	Adaptor6.0
32	Adaptor7.0
33	Adaptor8.0
34	Adaptor9.0
35	Adaptor10.0
36	Adaptor11.0
37	Adaptor12.0
38	Adaptor13.0
39	Adaptor14.0
40	Adaptor15.0
41	Adaptor16.0

42	β Zif
43	ZF-QNK
44	Hexahistidine Tag
45	Hexalysine Tag
46	ClBP1 expression construct
47	ClBP2 expression construct
48	ClBP3 expression construct
49	ClBP4 expression construct
50	ClBP5 expression construct
51	ClBP6 expression construct
52	ClBP7 expression construct
53	ClBP8 expression construct
54	ClBP9 expression construct
55	ClBP10 expression construct
56	ClBP11 expression construct
57	ClBP12 expression construct
58	ClBP13 expression construct
59	ClBP14 expression construct
60	ttGGBP.11C.0.bZif expression construct
61	ttGGBP.17C.0.bZif expression construct
62	ttGGBP.111C.0.bZif expression construct
63	ttGGBP.151C.0.bZif expression construct
64	ttGGBP.182C.0.bZif expression construct
65	ttGGBP.17C.3.Trx expression construct
66	Trx.ttGGBP.17C.3 expression construct
67	ttGGBP.182C.2.Trx expression construct
68	Trx.ttGGBP.182C.2 expression construct
69	Adaptor0 expression construct
70	Adaptor1.0 expression construct
71	Adaptor2.0a expression construct
72	Adaptor2.0b expression construct
73	Adaptor3.0 expression construct
74	Adaptor4.0 expression construct
75	Adaptor5.0 expression construct
76	Adaptor6.0 expression construct
77	Adaptor7.0 expression construct
78	Adaptor8.0 expression construct
79	Adaptor9.0 expression construct
80	Adaptor10.0 expression construct
81	Adaptor11.0 expression construct
82	Adaptor12.0 expression construct
83	Adaptor13.0 expression construct
84	Adaptor14.0 expression construct

85	Adaptor15.0 expression construct
86	Adaptor16.0 expression construct
87	ecGGBP [U.S. National Center for Biotechnology Information (NCBI) Accession No. WP_032329053]
88	ttGGBP (NCBI Accession Nos. YP_003852930.1 and WP_013298803.1)
89	stGGBP (NCBI Accession No. WP_001036943)
90	chyGGBP (NCBI Accession Nos. WP_013402088.1 and YP_003991244.1)
91	cobGGBP (NCBI Accession Nos. WP_013289482.1 and YP_003839461.1)
92	pspGGBP (NCBI Accession Nos. WP_015735911.1 and YP_003243743.1)
93	csaGGBP (NCBI Accession Nos. WP_013273028.1 and YP_003822565.1)
94	bprGGBP (NCBI Accession Nos. WP_013280279.1 and YP_003830205.1)
95	rinGGBP_A (NCBI Accession Nos. WP_006855636.1 and YP_007778116.1)
96	fprGGBP (NCBI Accession Nos. WP_015536639.1 and YP_007799070.1)
97	cljGGBP (NCBI Accession No. CLJU_c08950)
98	cauGGBP (NCBI Accession No. CAETHG_2989)
99	rinGGBP_B (NCBI Accession Nos. WP_006855628.1 and YP_007778124.1)
100	erhGGBP (NCBI Accession Nos. WP_003775352.1 and YP_004561181.1)
101	ereGGBP (NCBI Accession Nos. WP_012741392.1 and YP_002936409.1)
102	mpUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004483096.1 and WP_013797647.1]
103	mhUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_005430828.1 and WP_014422383.1]
104	bsUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_006233530.1 and WP_014665698.1]
105	dcUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004496535.1 and WP_013809819.1]
106	gtUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004588319.1 and WP_013877063.1]
107	ctUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_001038237.1 and WP_003515797.1]
108	csUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_001181243.1 and WP_011917972.1]
109	taUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_003473480.1 and WP_012991759.1]
110	gkUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_147790.1 and WP_011231423.1]
111	psUBP; [U.S. National Center for Biotechnology Information (NCBI)

	Accession Nos. YP_003241723.1 and WP_015734090.1]
112	teUBP; [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_681910.1 and WP_011567844.1]
113	ttGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004303.1 and WP_011172778.1]
114	tsGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004202647.1 and WP_015717367.1]
115	dmGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004171760.1 and WP_013557600.1]
116	tnGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_002534202.1 and WP_015919155.1]
117	koGBP [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_002941687.1 and WP_015869326.1]
118	bhGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. NP_244712.1 and WP_010899970.1]
119	smGBP [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_001041152.1 and WP_011839435.1]
120	asGBP [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_831349.1 and WP_011691715.1]
121	ttLacBP1 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_144032.1]
122	tsLacBP2 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004202714.1 and WP_015717434.1]
123	toLacBP3 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_006972155.1 and WP_016329249.1]
124	tsLacBP4 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_002514099.1 and WP_012638591.1]
125	rdLacBP5 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_683924.1 and WP_011569849.1]
126	msLacBP6 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_006556686.1 and WP_014869652.1]
127	tsLacBP7 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_005654632.1 and WP_014515914.1]
128	maLacBP8 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_005886720.1 and WP_014578260.1]
129	adLacBP9 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_466099.1 and WP_011421944.1]
130	pgLacBP10 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_004304976.1 and WP_013653981.1]
131	psLacBP11 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_006522676.1 and WP_014851134.1]
132	rsLacBP12 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. RSP_3372 and YP_354877.1]
133	fsLacBP13 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_004603455.1 and WP_013886373.1]
134	taLacBP14 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_003317968.1]
135	synBicarbBP1 [U.S. National Center for Biotechnology Information

	(NCBI) Accession Nos. YP_005410477.1 and WP_010873997.1]
136	teBicarbBP2 [U.S. National Center for Biotechnology Information (NCBI) Accession No. NP_682790.1]
137	ctBicarbBP3 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_007090308.1 and WP_015152989.1]
138	calBicarbBP4 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_007137061.1 and WP_015197735.1]
139	avBicarbBP5 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_321546.1 and WP_011317875.1]
140	cmBicarbBP6 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_007099445.1 and WP_015162006.1]
141	mhFeBP1 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_007884192.1 and WP_006253500.1]
142	exiFeBP2 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_002886303.1 and WP_015880417.1]
143	teFeBP3 [U.S. National Center for Biotechnology Information (NCBI) Accession No. NP_681303.1]
144	cnFeBP4 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_003796723.1 and WP_013247623.1]
145	ttFeBP5 [U.S. National Center for Biotechnology Information (NCBI) Accession No. YP_144894.1]
146	msFeBP6 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_003686074.1 and WP_013159102.1]
147	srFeBP7 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_003572493.1 and WP_013062602.1]
148	hlFeBP8 [U.S. National Center for Biotechnology Information (NCBI) Accession Nos. YP_002564837.1 and WP_012659409.1]
149	Yellow Fluorescent Protein
150	Yellow Fluorescent Protein without N-terminal Methionine
151	ecTrx
152	GGSHHHHHH Sequence
153	ecGGBP without signal peptide
154	ttGGBP (with signal peptide replaced with M)
155	stGGBP (with signal peptide replaced with M)
156	chyGGBP (with signal peptide replaced with M)
157	cobGGBP (with signal peptide replaced with M)
158	pspGGBP (with signal peptide replaced with M)
159	csaGGBP (with signal peptide replaced with M)
160	bprGGBP (with signal peptide replaced with M)
161	rinGGBP_A (with signal peptide replaced with M)
162	fprGGBP (with signal peptide replaced with M)
163	cljGGBP (with signal peptide replaced with M)
164	cauGGBP (with signal peptide replaced with M)
165	rinGGBP_B (with signal peptide replaced with M)
166	erhGGBP (with signal peptide replaced with M)
167	ereGGBP (with signal peptide replaced with M)

The terms “bZif” and “βZif” are used synonymously herein.

Exemplary amino acid sequences are listed below for convenience:

ecTrx (SEQ ID NO: 151)

MSDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPCKMIAPILDEIADEY
5 QGKLTVAKLNDQNPQTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQL
KEFLDANLA

YFP (SEQ ID NO: 149)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
10 TGKLPVPWPTLVTTFGYGLQCFARYPDHMKRHDFFKSAMPEGYVQERTIF
FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSYQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKGGSN DYKDDDD
K
15

CIBP1 (SEQ ID NO: 1)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
TGKLPVPWPTLVTTFGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
20 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
KGGSHHHHHH**

CIBP2 (SEQ ID NO: 2)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
25 TGKLPVPWPTLVTTFGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
30 KGGSHHHHHH**

CIBP3 (SEQ ID NO: 3)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
TGKLPVPWPTLVTTFGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
35 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
KGGSHHHHHH**

CIBP4 (SEQ ID NO: 4)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
TGKLPVPWPTLVTTFGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
45 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
KGGSHHHHHH**

CIBP5 (SEQ ID NO: 5)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 5 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

CIBP6 (SEQ ID NO: 6)

10 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

15

CIBP7 (SEQ ID NO: 7)

20 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

CIBP8 (SEQ ID NO: 8)

25 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 30 KGGSHHHHHH**

CIBP9 (SEQ ID NO: 9)

35 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

40 CIBP10 (SEQ ID NO: 10)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFAFYDPHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPVLLPDNH
 45 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

CIBP11 (SEQ ID NO: 11)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT

TGKLPVPWPTLVTTFTGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 5 KGGSHHHHHH**

CIBP12 (SEQ ID NO: 12)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
 10 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

CIBP13 (SEQ ID NO: 13)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 TGKLPVPWPTLVTTFTGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
 20 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

CIBP14 (SEQ ID NO: 14)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICT
 25 TGKLPVPWPTLVTTFTGYGVQCFARYPDHMRQHDFFKSAMPEGYVQERTIF
 FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN
 VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
 YLSYQSALS KDPNEKRDHMLLEFVTAAGITLGMDELYKGGSN DYKDDDD
 KGGSHHHHHH**

Glucose-Galactose Binding Protein Fusions

ttGBP.11C.0.bZif (SEQ ID NO: 15)

MKQLNIGVAICKFDDTFMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMCK
 35 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 HQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHDKMAAWLSSFGD
 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 TLLGTVLNDAKNQAKATFNIA YELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSGGSTGEKPYKCPECCKSFSRSGGSHHHHHH**

ttGBP.17C.0.bZif (SEQ ID NO: 16)

MKQLNIGVAIYKFDDTCMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMCK
 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 45 HQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHDKMAAWLSSFGD
 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 TLLGTVLNDAKNQAKATFNIA YELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSGGSTGEKPYKCPECCKSFSRSGGSHHHHHH**

ttGBP.111C.0.bZif (SEQ ID NO: 17)

MKQLNIGVAIYKFDDTFMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMKK
 WDKVYYVGAKCEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 5 HQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHDKMAAWLSSFGD
 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 TLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSGGSTGEKPYKCPECCKSFSRSGGSHHHHHH**

10 ttGBP.151C.0.bZif (SEQ ID NO: 18)

MKQLNIGVAIYKFDDTFMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMKK
 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 CQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHDKMAAWLSSFGD
 15 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 TLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSGGSTGEKPYKCPECCKSFSRSGGSHHHHHH**

ttGBP.182C.0.bZif (SEQ ID NO: 19)

20 MKQLNIGVAIYKFDDTFMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMKK
 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 HQDAILRTQYSIQTVKDAGIKVQELAKDYANCDRVTAHDKMAAWLSSFGD
 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 25 TLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSGGSTGEKPYKCPECCKSFSRSGGSHHHHHH**

ttGBP.17C.3.Trx (SEQ ID NO: 20)

30 MKQLNIGVAIYKFDDNCMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNREPLPEDMKK
 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGVMQYVMLMGQPG
 HQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHDKMAAWLSSFGD
 KIEAVFANNDAMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIQAIKDG
 TLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDITDGKYVWIPYK
 35 KITKDNISDAEQGGSSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPC
 KMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLKFNGE
 VAATKVGALSKGQLKEFLDANLAGGSHHHHHH

Trx.ttGBP.17C.3 (SEQ ID NO: 21)

40 MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLKFNGEVAATKVGALSKGQL
 KEFLDANLAGGSKQLNIGVAIYKFDDNCMTGVRNAMTAEAQGKAKLNMVD
 SQNSQPTQNDQVDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFF
 NREPLPEDMKKWDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHDGV
 45 MQYVMLMGQPGHQDAILRTQYSIQTVKDAGIKVQELAKDYANWDRVTAHD
 KMAAWLSSFGDKIEAVFANNDAMALGAIEALKSAGYFTGNKYIPVVGVD
 TAPGIQAIKDGTTLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDI
 TDGKYVWIPYKKITKDNISDAEQGGSHHHHHH

ttGBP.182C.2.Trx (SEQ ID NO: 22)

MKQLNIGVAIYKFDDTFMTGVRNAMTAEAQGKAKLNMVDSQNSQPTQNDQ
 VDLFITKKMNALAINPVDRTAAGTIIDKAKQANIPVVFFNKEPLPEDMKK
 WDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHMGVVMQYVMLMGEPG
 5 HQDAILRTQYSIQTVKDAGIKVQELAKDYANCDRVTAHDKMAAWLSSFGD
 KIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDATAPGIAIKDG
 TLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDITDGKYVWIPYK
 KITKDNISDAEQGGSSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPC
 KMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNKE
 10 VAATKVGALSKGQLKEFLDANLAGGSHHHHHH

Trx.ttGBP.182C.2 (SEQ ID NO: 23)

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNKEVAATKVGALSKGQL
 15 KEFLDANLAGGSKQLNIGVAIYKFDDTFMTGVRNAMTAEAQGKAKLNMVD
 SQNSQPTQNDQVDLFIKMNALAINPVDRTAAGTIIDKAKQANIPVVFF
 NKEPLPEDMKKWDKVYYVGAKAEQSGILQGQIMADYWKAHPEADKNHMGV
 MQYVMLMGEPGHQDAILRTQYSIQTVKDAGIKVQELAKDYANCDRVTAHD
 KMAAWLSSFGDKIEAVFANNDMALGAIEALKSAGYFTGNKYIPVVGVDAT
 20 TAPGIAIKDGTLLGTVLNDAKNQAKATFNIAYELAQGITPTKDNIGYDI
 TDGKYVWIPYKKITKDNISDAEQGGSHHHHHH

Adaptor Proteins

Adaptor0 (SEQ ID NO: 24)

25 MSAKIIHLTDDSFDTDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAAMLNIDQNPGTAPKYGIRGIPTLLLFKNKEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor1.0 (SEQ ID NO: 25)

30 MSAKIIHLTDDSFDTDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAAMLNIDQNPGTAPKYGICGIPTLLLFKNKEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor2.0a (SEQ ID NO: 26)

35 MSAKIIHLTDDSFDTDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAAMLNIDQNPGTAPKYGICGIPTLLLFKNKEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor2.0b (SEQ ID NO: 27)

40 MSAKIIHLTDDSFDTDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAAMLNIDQNPGTAPKYGICGIPTLLLFKNKEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor3.0 (SEQ ID NO: 28)

45 MSAKIIHLTDCSFDTDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAAMLNIDQNPGTAPKYGIRGIPTLLLFKNKEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor4.0 (SEQ ID NO: 29)

MSAKIIHLTDDSFDCDVLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

5 Adaptor5.0 (SEQ ID NO: 30)

MSAKIIHLTDDSFDTDLKACGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

10 Adaptor6.0 (SEQ ID NO: 31)

MSAKIIHLTDDSFDTDLKADGAILVAFWAECCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

15 Adaptor7.0 (SEQ ID NO: 32)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKCIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

20 Adaptor8.0 (SEQ ID NO: 33)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDCIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

25 Adaptor9.0 (SEQ ID NO: 34)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGCLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

30 Adaptor10.0 (SEQ ID NO: 35)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDCNPGTAPKYGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

35 Adaptor11.0 (SEQ ID NO: 36)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKCGIRGIPTLLL FKN GEVAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

40 Adaptor12.0 (SEQ ID NO: 37)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GECAATKVGALSKGQL
 KEFLDANLAGGSHHHHHH***

45 Adaptor13.0 (SEQ ID NO: 38)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
 QGKLTVAMLNIDQNP GTAPKYGIRGIPTLLL FKN GEVAATKCGALSKGQL
 KEFLDANLAGGSHHHHHH***

Adaptor14.0 (SEQ ID NO: 39)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
QGKLTVAMLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSCGQL
KEFLDANLAGGSHHHHHH***

5

Adaptor15.0 (SEQ ID NO: 40)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
QGKLTVAMLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGCL
KEFLDANLAGGSHHHHHH***

10

Adaptor16.0 (SEQ ID NO: 41)

MSAKIIHLTDDSFDTDLKADGAILVAFWAEWCGPCKMIAPILDEIADEY
QGKLTVAMLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQL
KEFLDCNLAGGSHHHHHH***

15

EXAMPLES

Examples are provided below to facilitate a more complete understanding of the invention. The following examples illustrate the exemplary modes of making and practicing the invention. However, the scope of the invention is not limited to specific embodiments disclosed in these Examples, which are for purposes of illustration only, since alternative methods can be utilized to obtain similar results.

20

EXAMPLE 1: SENSING CHLORIDE USING ngmFRET EFFECTS ENGINEERED INTO YELLOW FLUORESCENT PROTEIN.

25

As the first, simplest, example of ngmFRET, we use YFP, a mutant of Green Fluorescent Protein (GFP) that contains a serendipitous chloride-binding site located next to its fluorophore (Wachter 1999, Jayaraman 2000, Kuner 2000). Chloride binding elicits a monochromatic response of the fluorophore, quenching its fluorescence with increasing chloride concentrations (by itself and regardless of the presence of any other fluorophore/partner). Binding is not accompanied by a protein conformational change (Grimley et al. 2013), therefore clearly ruling out any distance-dependent effects on energy transfer, which limits the number of possible photophysical models.

30

The YFP fluorophore is formed by an autocatalytic cyclization of a tripeptide located in an interior α helix (Zimmer 2002). It exists in four possible states (anionic, cationic, zwitterionic, and neutral), of which only the anionic fluorophore is fluorescent (Elslinger 1999) (FIG. 6). Binding of chloride at a site located adjacent to the fluorophore destabilizes the

35

anionic form, raising the excited state pK_a^* resulting in protonation of the fluorophore and quenching of its fluorescence emission intensity (Grimley et al. 2013). The YFP fluorescence emission spectrum changes intensity but not shape in response to chloride binding (monochromatic response).

5 We have set up an ngmFRET system in which the fluorophore functions as a directly responsive acceptor, with an indirectly responsive donor site-specifically coupled on the YFP surface. In this system both donor and acceptor intensities change in response to chloride binding. Our analysis has shown that in the absence of inter-fluorophore distance changes, as is the case for these YFP conjugates, the emission intensity of the indirectly responsive donor
10 can change only if the efficiency of resonance energy transfer from the donor to the directly responsive acceptor changes as a consequence of alterations in the degree of spectral overlap due to ligand-mediated shifts in the excitation spectrum of the directly responsive acceptor.

To construct this ngmFRET system, we used the laYFP variant (Grimley et al. 2013 (FIG. 7), which binds chloride at a site ~ 14 Å from the YFP phenolic hydroxyl with affinity of
15 ~ 160 mM at neutral pH. A parent construct (ClBP0) was designed for heterologous expression in *Escherichia coli* by optimizing the nucleotide sequence of the open reading frame (Allert, Cox and Hellenga 2010). Thirteen single cysteine mutants (ClBP 2-14) were designed for attachment of the extrinsic indirectly responsive donor fluorophore, located in an annulus around the main YFP β -barrel and at both ends of the barrel (FIG. 8).

20 Acrylodan (Prendergast 1983) and Pacific Blue (Sun 1998) were each attached to purified mutant proteins. In these conjugates both donor and acceptor intensities changed in response to chloride and pH, and the ratio of acceptor to donor intensities decreases (FIG. 9; Table 2). This can happen only if the directly responsive acceptor undergoes dipole switching such that its absorption spectrum undergoes a bathochromic shift and alters the
25 energy transfer coupling factor, thereby decreasing energy transfer efficiency and increasing the intensity of the indirectly responsive donor (equation 16). Chloride binding decreases the fluorescence of the unlabeled YFP, indicating that increased directly responsive acceptor quenching also plays a role in the photophysics of this ngmFRET system. This pattern corresponds to model $a^+\phi$ in Table 1.

Table 2. Chloride affinities and pK_a values of semisynthetic YFP chloride sensors^a.

Name	Mutation	Location	Acrylodan ^b			Pacific Blue ^c		
			$K_d(Cl^-)(mM)$		pK_a	$K_d(Cl^-)(mM)$		pK_a
			$^{app}K_d$	$^{true}K_d$		$^{app}K_d$	$^{true}K_d$	
CIBP2	E17C	β_1	308	52	5.8	30	60	5.6
CIBP3	E32C	β_2	19	46	5.8	76	126	5.8
CIBP4	T43C	β_3	41	129	5.3	96	136	4.4
CIBP5	E95C	β_4	37	77	5.7	220	340	<4
CIBP6	R109C	β_5	77	106	5.4	208	545	<4
CIBP7	R122C	β_6	35	65	5.4	180	360	<4
CIBP8	N149C	β_7	23	34	5.8	122	106	4.4
CIBP9	N164C	β_8	26	70	5.3	259	364	<4
CIBP10	Y182C	β_9	35	88	5.2	105	260	<4
CIBP11	Q204C	β_{10}	72	55	5.2	310	242	<4
CIBP12	L221C	β_{11}	40	140	3.9	167	276	<4
CIBP13	H77C	End	20	55	5.6	150	290	<4
CIBP14	D173C	End	78	110	5.4	217	212	<4

^a Determined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations **20-25**.

5 ^bRatiometry: $\lambda_1=530$ nm, $\lambda_2=500$ nm.

^cRatiometry: $\lambda_1=530$ nm, $\lambda_2=455$ nm.

The collection of thirteen YFP conjugates (Table 2) illustrates that the quantitative details of their ngmFRET behavior have profound effects on sensor utility, even if all constructs evince ratiometric responses to chloride binding. The emission peak of Acrylodan is centered at ~500 nm, whereas Pacific Blue emits at 455 nm. The excitation maximum of YFP is centered at ~515 nm. The spectral overlap with the YFP acceptor therefore is much greater for Acrylodan than for Pacific Blue donors. Consequently the donor emission intensity is weak in Acrylodan conjugates, as most of their excited state energy is transferred to YFP, whereas the corresponding intensities of Pacific Blue donors are much stronger. Furthermore, the Pacific Blue peak is more responsive than Acrylodan, as small shifts in spectral overlap by changes in the directly responsive YFP acceptor have a relatively larger effect on the former than the latter. Even among the Pacific Blue conjugates, there is quantitative difference in the degree of overlap, depending on the attachment position, with concomitant changes in behavior (Table 2). Effective dichromatic chloride sensors therefore are constructed by balancing the degree of spectral overlap such that the donor intensity contributes significantly to the signal.

The response of biosensors based on Langmuir binding isotherms is most sensitive at analyte concentrations that match the apparent K_d value of the protein (Marvin et al. 1997).

The various Acrylodan and Pacific Blue conjugates therefore can be used to measure chloride concentrations over two orders of magnitude, in the 10-1000 mM concentration range. These concentration ranges are relevant to water quality determination, corrosion, and clinical chemistry. The clinical reference range (Burtis 2012) for chloride concentrations is 90-120 mM (normal serum range 97-107 mM). The Pacific Blue conjugates of CIBP3, 6, 8 and 10 exhibit $^{app}K_d$ values centered around 100 mM Cl^- in solutions at pH 7.4 with ionic strengths of ~150 mEq (conditions that match that of serum(Burtis 2012)), and therefore are good candidates for use in point of care (POC) technology. The pK_a values of these variants are less than 4; fluctuations in proton concentrations around pH 7.4 therefore will not interfere with the functioning of these biosensors, which is an important consideration in the development of chloride detection technologies.

EXAMPLE 2: SENSING GLUCOSE USING nmFRET EFFECTS ENGINEERED INTO A THERMOSTABLE GLUCOSE, BINDING PROTEIN.

As a second, more complex example we use a glucose-binding protein in which the directly responsive fluorophore is site-specifically coupled to a cysteine mutant located in the binding pocket of the protein. The indirectly responsive partner is placed either on a C-terminal β Zif or N-terminal thioredoxin fusion, using the multiple thiol-labeling strategy described above. We provide examples of directly responsive donors and acceptors.

A thermostable homolog of *Escherichia coli* glucose-galactose binding protein, ecGGBP (Vyas, Vyas and Quijcho 1988), was identified in the thermophilic bacterium *Thermoanaerobacter thermosaccharolyticum*. Additional non-limiting examples of glucose-galactose binding proteins, as well as other ligand-binding proteins, are disclosed herein. Additionally, non-limiting examples of glucose-galactose binding proteins are described in PCT International Patent Application No. PCT/US2016/050297, filed September 2, 2016. This homolog, ttGGBP, is a member of the bacterial periplasmic-binding protein (PBP) superfamily, which has provided several ligand-responsive reagentless fluorescent sensors (Grunewald 2014). The structure of PBPs comprises two domains connected by a hinge (Berntsson et al. 2010). Ligand binding shifts a conformational equilibrium via a hinge-bending motion from an ensemble of predominantly “open” states to a “closed” state in which the bound ligand is enveloped within the interface between the two protein domains. Using site-specific modification of cysteine mutants, environmentally sensitive fluorophores have been placed strategically to detect these ligand-mediated conformational changes (de

Lorimier et al. 2002). The resulting semi-synthetic, fluorescent PBP conjugates then function as reagentless, fluorescent biosensors for their cognate ligands.

We constructed a series of single cysteine mutants in ttGGBP (FIG. 10) and evaluated the fluorescent responses of Badan and Acrylodan conjugates at these positions. Three of the seventeen positions exhibited dichromatic responses with one or both of these fluorophores (Table 3). The best ratiometric responses were seen for Badan and Acrylodan attached to F17C and W182C, respectively. A series of additional mutations were introduced within these two cysteine mutants to tune their glucose affinities (Table 4). Finally, we tested the responses of a variety of different fluorophores (FIG. 11) at these two positions (Table 5).

Most of these fluorophores evinced monochromatic responses, with the exception of the W182C·IAEDANS conjugate which exhibited a small change in emission spectrum shape upon binding glucose. These singly labeled conjugates form a dataset of directly responsive fluorophore from which examples were drawn to construct glucose-sensing ngmFRET systems.

Table 3. Glucose response of Acrylodan and Badan conjugates in a cysteine scan of the ttGGBP scaffold.

Mutation	Class ^a	Shape ^b	Conjugate ^c	^{apo} T _m (K)	Emission wavelength (nm)		K _d ^{d,e} (mM)	
					λ1	λ2	^{app} K _d	^{true} K _d
Y11C	p	m	A	351	511	470	0.12	0.16
		d	B	349	492	528	0.28	0.24
T16C	p	m	A	351	519	460	8.8	12
		m	B	349	-	-	nb	nb
F17C	e	d	A	349	482	542	0.08	0.06
		d	B	346	467	491	0.087	0.26
N42C	p		A	350	-	-	nb	nb
			B	324	-	-	nb	nb
V67C	p		A	350	-	-	nb	nb
			B	322	-	-	nb	nb
R91C	e	m	A	349	491	540	0.18	0.17
			B	348	-	-	nb ^d	nb ^d
E92C	p		A	350	-	-	nb	nb
			B	346	-	-	nb	nb
A111C	p	m	A	350	515	550	0.19 ^d	0.11 ^d
		m/d	B	348	523	550	0.64	0.55
Q148C	p		A	351	-	-	nb	nb
			B	349	-	-	nb	nb
H151C	e	m	A	351	511	489	0.012	0.025

		m	B	348	523	550	0.018 ^d	0.027 ^d
Q152C	p		A	351	-	-	nb ^d	nb ^d
			B	349	-	-	nb	nb
N181C	p		A	350	-	-	nb	nb
			B	349	-	-	nb	nb
W182C	e	d	A	347	479	526	2.3	2.3
		m	B	347	515	550	27	19
D183C	p		A	348	-	-	nb ^d	nb ^d
			B	348	-	-	nb	nb
L257C	a		A	352	-	-	nb	nb
			B	348	-	-	nb	nb
D259C	a		A	349	-	-	nb	nb
			B	347	-	-	nb	nb
K300C	a		A	349	-	-	nb	nb
			B	348	-	-	nb	nb

^aa, allosteric; e, endosteric; p, peristeric.

^bm, monochromatic; d, dichromatic (i.e. spectral shape change).

^cA, Acrylodan; B, Badan.

^dnoisy data and or bad fit.

5 ^enb; no binding, nd; not determined.

Table 4. Responses of mutant ttGGBP17C and ttGGBP182C conjugates^a.

Protein	Mutation	Class ^b	Conjugate ^c	^{apo} T _m (K)	Emission wavelength (nm)		Glucose ^d		Galactose ^d	S
					λ1	λ2	^{app} K _d (mM)	^{true} K _d (mM)	^{true} K _d (mM)	
ttGGBP17C			B	346	467	519	0.10	0.15	0.19	1.3
„			A	349	487	515	0.08	0.09	3.8	43
ttGGBP17C.1	R91K,Q148E	2	B		463	515	0.6	0.8	0.4	0.46
ttGGBP17C.2	R69P,Q152P	2	B	350	479	523	8.6	10.8	5.2	0.48
„	„	2	A		492	515	15	18	5.0	0.28
ttGGBP17C.3	T16N,D211A	2	B	346	471	531	3.4	4.1	2.5	0.61
„	„	2	A		495	529	1.4	1.5	0.53	0.35
ttGGBP17C.4	H151Q	1	B	347	511	457	16	6.4	14	2.2
„	„	1	A		488	550	50	48	14	0.29
ttGGBP17C.5	D15A	1	B	345	487	530	16	16	4.2	0.26
„	„	1	A	348	483	498	6	6	2.4	0.41
ttGGBP17C.6	D15E	1	B	346	467	525	3.3	3.6	3.6 ^e	1
ttGGBP17C.7	D15N	1	B	347	483	515	0.75	1.0	0.64	0.64
„	„	1	A	348	483	515	0.3	0.3	0.3	0.91
ttGGBP17C.8	T16N	2	A	348	487	529	0.61	0.60	0.26	0.43
ttGGBP17C.9	T16S	2	A		487	520	0.2	0.20	nd	
ttGGBP17C.10	G20A	3	A	351	487	520	0.4	0.4	nd	
ttGGBP17C.11	T240A	3	A	348	487	500	0.04	0.04	nd	
ttGGBP17C.19	N258D	3	B	344			nb ^e	nb ^e	nb ^e	
ttGGBP17C.20	N258S	3	B	344			nb ^e	nb ^e	nb ^e	
ttGGBP17C.21	N258A	3	B	345	532	494	61	69	nb	
ttGGBP17C.22	A260N	3	B	343	523	492	12	15	530 ^e	
ttGGBP17C.23	A260Q	3	B	344	527	490	15	18	nb ^e	
ttGGBP17C.24	A260R	3	B	344	490	515	1.8	1.7	29	17
ttGGBP17C.25	A260K	3	B		515	493	3.2	3.7	nb ^e	
ttGGBP17C.26	A260W	3	B	346	523	509	0.9	0.9	14	16

ttGGBP17C.27	A260F	3	B	346	523	490	0.2	0.3	8.8	33
ttGGBP17C.28	A260Y	3	B	346	523	494	0.06	0.07	2.6	37
ttGGBP17C.29	A260S	3	B	343	527	496	2.1	2.2	250	114
ttGGBP182C			A	347	472	535	2.2	2.3	3.3	1.4
ttGGBP182C.2.0 ^f	R91K Q148E	2,3	A	346	475	545	4.5	6.0	18.5	3.1
ttGGBP182C.2.1 ^g	A154S	1	A	328	480	552	3.0	4.1	13.1	3.2
ttGGBP182C.2.3 ^g	A154N	1	A	346	470	537	16.3	19.0	207	10.9
ttGGBP182C.2.4 ^g	A154M	1	A	346	472	540	0.4	0.4	4.5	11.2
ttGGBP182C.2.5 ^g	H151Q	1	A	346	477	542	10.8	13.0	20.3	1.6
ttGGBP182C.2.6 ^g	H151N	1	A	347	475	537	19.2	17.9	52.6	2.9
ttGGBP182C.2.7 ^g	H151F	1	A	346	475	545	121	124		
ttGGBP182C.2.8 ^g	D15N	1	A	343	474	542	179	209	372	1.8
ttGGBP182C.2.9 ^g	A154F	1	A	345	512	477	0.4	0.3	1.6	4.0
ttGGBP182C.3 ^f	R91K	2	A	346	494	459	1.8	0.9	3.1	3.4
ttGGBP182C.4 ^f	Q148E	2	A	347	475	542	0.6	0.8	6.5	8.1
ttGGBP182C.5 ^f	R69P Q152P	2	A	349	477	535	7.5	8.8	7.4	0.8
ttGGBP182C.6 ^f	T16N D211A	2	A	345	472	530	21.6	27.6	81	2.9
ttGGBP182C.7 ^f	R91K Q148S	2,3	A	346	478	550	0.3	0.4	1.1	3.7
ttGGBP182C.8 ^f	R91K Q148K	2,3	A	346	475	545	28.7	44.2	256	5.8
ttGGBP182C.9 ^f	D15N	1	A	343	475	540	75.1	76.4	282	3.7

^aMeasured on the Nanodrop at room temperature. $\lambda_{\max}$ is the wavelength corresponding to the maximum emission intensity. Optimal ratiometry wavelengths are determined according to the analysis described in Materials and Methods. The $^{\text{true}}K_d$ is determined from monochromatic titration curves; $^{\text{app}}K_d$ from dichromatic ratiometry (equations 20-25).

- 5 Average relative error in the $^{\text{true}}K_d$ values is 5%, in the $^{\text{app}}K_d$ values, 1%. S is the selectivity between glucose and galactose, $S = ^{\text{true}}K_d(\text{galactose}) / ^{\text{true}}K_d(\text{glucose})$; $S > 1$, selective for glucose.

^b1, PCS; 2, inter-domain interaction; 3, contact between protein and fluorophore

^cA, acrylodan; B, badan.

- 10 ^dnb, no bonding; nd, not determined.

^eNoisy data or bad fit.

^fAdditional mutation constructed in ttGGBP182C.

^gAdditional mutation constructed in ttGGBP182C.2.0.

- 15 Table 5. Responses of fluorophores conjugated to F17C and W182C mutants of ttGGBP^a.

Position	Fluorophore ^b	λ_{ex} (nm) ^c	$^{\text{apo}}\lambda_{\text{max}}$ (nm)	$^{\text{apo}}I_{\text{max}}$ (AU x1000)	$^{\text{sat}}\lambda_{\text{max}}$ (nm)	$^{\text{sat}}I_{\text{max}}$ (AU x1000)	$^{\text{true}}K_d$ (mM)
17C	Acrylodan	391	487	15.9	487	20.0	0.2
	Badan	391	519	12.8	467	52.6	0.1
	5-IAF	491	523	64.3	523	70.6	61.3
	Oregon green	496	523	101.9	523	91.7	-
	CPM	384	471	78.5	467	90.9	17.0
	IANBD	478	531	15.3	535	19.8	11.8
	IAEDANS	336	467	12.5	467	14.4	-
	Pacific Blue	410	451	28.6	455	108.6	60.7
182C	Acrylodan	391	479	55.0	515	17.2	6.0
	Badan	391	515	16.8	515	11.9	64.9
	5-IAF	491	519	255.0	519	453.2	5.0
	6-IAF	491	513	50.4	513	63.1	750

	Oregon green	496	519	78.2	519	186.2	20.0
	CPM	384	479	49.5	483	40.3	6.8
	IANBD	478	543	29.1	547	12.5	210
	IAEDANS	336	487	3.7	483	7.9	0.2
	Pacific Blue	410	455	115.0	455	119.4	-
	BODIPY 499	499	519	40.1	515	80.2	30.7
	BODIPY 507	507	531	32.9	535	31.4	11.0
	Alexa 488	495	519	211.6	519	182.2	-
	Alexa 532	532	551	63.0	551	61.0	-
	Alexa 546	546	571	152.1	571	150.2	-
	Texas Red	595	611	23.3	611	23.5	-
	Cy5	646	663	19.2	663	23.3	210
	PyMPO	415	555	3.7	559	4.2	20.6

^a λ_{ex} , preferred excitation wavelength (from supplier); ^{apo} λ_{max} , observed maximum emission wavelength of the apo-protein; ^{apo} I_{max} , observed intensity at ^{apo} λ_{max} ; ^{sat} λ_{max} , observed maximum emission wavelength of the glucose complex; ^{sat} I_{max} , observed intensity at ^{sat} λ_{max} ; ^{true} K_d , affinity determined from fit of equations 20-25 to the monochromatic emission intensities. Emission spectra were measured on the Nanodrop3300, using ~10 μM protein. The observed absolute emission intensities are a rough guide to the relative brightness of the conjugate, because the protein concentration was approximately the same for each experiment.

^bAbbreviations, chemical names and supplier catalogue numbers as follows: Acrylodan (A433); Badan (B6057); 5-IAF (I30451); Oregon Green 488 (O6034); CPM (D346); IANBD (D2004); IAEDANS (I14); Pacific Blue (P30506); BODIPY 499 (D20350); BODIPY 507 (D6004); 6-IAF (I30452); Alexa 488 (A10254); Alexa 532 (A10255); Alexa 546 (A10258); Texas Red (T6008); PyMPO (M6026) from Life Technologies and Cy5 (13080) from Lumiprobe.

^cThe Nanodrop3300 fixed wavelength LED that most closely matched λ_{ex} was used (see Materials and Methods).

Conversion of monochromatic to dichromatic responses.

Case 1: Directly responsive donor, indirectly responsive acceptor; pure quenching (model $d\phi$)

The singly labeled Pacific Blue conjugate at F17C exhibits a strong increase in emission intensity without significant shifts in emission wavelength maxima in response to glucose (FIG. 12A), suggesting that the changes are due primarily to a decrease in the non-radiative decay rate of its excited state in the glucose complex. In the doubly labeled F17C-Pacific Blue $\beta\text{Zif-5-IAF}$ fusion the emission intensities of both the directly responsive donor and indirectly responsive acceptor peaks increase with glucose addition (FIG. 12B). This correlated increase of both donor and acceptor emission intensities is inconsistent with effects confined to changes in the energy transfer coupling factor, ϕ . The lack of observable

changes in the emission spectrum, indicates that the ngmFRET is modulated only by differences in the non-radiative decay rate on the Pacific Blue directly responsive donor in the ligand-free and –bound protein (model $d\phi^0$, Table 1). In model $d\phi^0$, modulation of the indirectly responsive acceptor emission intensity are due to changes in the resonance transfer rate which increases as a consequence of decreased the donor non-radiative decay rate (by itself and regardless of the presence of any other fluorophore/partner).

The β Zif fusion domain is located at the back of the ttGGBP hinge region, close to the attachment point of the directly responsive fluorophore at position F17C or W182C (FIG. 13), forming an ensemble of indirectly responsive fluorophore conformations. In this arrangement, the ligand-mediated protein conformational change that converts the open to the closed state is unlikely to change the average distance or orientation between the single directly responsive fluorophore conformation and the ensemble of indirectly responsive fluorophores. The behavior of the F17C·Pacific Blue β Zif-5-IAF ngmFRET pair is consistent with this interpretation. Distance contributions therefore can be ruled in the interpretation of the intensity change pattern.

Singly labeled Acrylodan attached to position 111C also exhibits a monochromatic increase in emission intensity in response to glucose binding. When paired as a directly responsive donor with β Zif·Alexa532 as the indirectly responsive acceptor, this monochromatic signal is converted to a dichromatic one (Table 7). As with the F17C·Pacific Blue – β Zif·IAF example described above, donor and acceptor emission intensities increase unequally upon binding glucose, enabling ratiometry, and indicating that the $d\phi^0$ model is operational in this construct as well. By contrast, glucose binding evinces only a monochromatic signal in the 111C·Acrylodan - β Zif·IAF system. In this construct the wavelengths of the Acrylodan donor and IAF acceptor emission peaks are so close together that their changes in either cannot be resolved independently. The same effects are observed for the Badan conjugate attached to position 151C both singly labeled and partnered with β Zif·Alexa532 and β Zif·IAF (Table 7). These observations illustrate the importance of carefully tuning donor-acceptor optical properties in order to obtain dichromatic signals, even if ngmFRET between the two partners can be established.

Case 2: Directly responsive donor, indirectly responsive acceptor; pure quenching (model $d^+ \phi^0$)

In contrast to the Badan conjugate described above, at position 151C the singly labeled Acrylodan conjugate exhibits a monochromatic emission intensity decrease. When partnered with an indirectly responsive β Zif·Alexa532 acceptor, both donor and acceptor intensities decrease (Table 6). This behavior is consistent with the pure quenching (change in the ratio of radiative to non-radiative emission rates) model $d^+ \phi^0$ (Table 1). For the reasons described above, the 151C·Acrylodan - β Zif·IAF system evinces a monochromatic response.

Case 3: Indirectly responsive donor, directly responsive acceptor; combined spectral shift and quenching (model $a^- \phi$)

The fluorescence intensities of W182C conjugates singly labeled with closely related members of the fluorescein family, 5-IAF and Oregon Green (FIG. 11D and E), increase in response to glucose binding (Table 5). This observation suggests that the glucose complexes of these two conjugates exhibit a significant decrease in the relative rates of non-radiative decay.

These conjugates were paired as directly responsive acceptors with Pacific Blue as their indirectly responsive donor. The indirectly responsive donor emission intensity increased in concert with that of the directly responsive acceptor in both ngmFRET pairs (FIG. 14). As described for YFP in the first example, such a pattern is inconsistent with pure geometrical effects. The “feedback” to the indirectly responsive donor indicates that ngmFRET efficiency has diminished by a bathochromic shift in the directly responsive acceptor excitation spectrum, resulting in increased donor emission. This can occur only if the directly responsive acceptor switches dipoles, consistent with the presence of multiple excited states in Fluorescein and Oregon Green (Martin 1975). The increase in intensity of the singly labeled conjugates indicates that quenching decreases upon glucose binding. The modulation mechanism therefore combines quenching and ngmFRET coupling effects: model $a^- \phi$ (Table 1).

In both systems, the fluorescent response pattern changes with increased glucose concentrations, shifting from an increase (Phase I) to a decrease (Phase II) in the Q_D/Q_A ratio (Fig. 14), indicating that there are two types of glucose responses. This phenomenon strongly suggests the presence of a second, much weaker affinity glucose-binding site that influences

the environment of the directly responsive acceptor. The intensity of the indirectly responsive donor continues to increase with glucose concentration in Phase II, indicating a further decrease in ϕ with a concomitant bathochromic shift of the acceptor absorption spectrum. The pattern of changes in the Q_D/Q_A ratio is quantitatively dependent on the change in magnitudes of a and ϕ in the $a\phi$ model (Table 1). Accordingly, the change in the direction of the ratiometric signal indicates that the relative sizes of the change in ϕ and a are different in Phases I and II.

Table 6. Glucose affinities of ngmFRET systems in Y11C, A111C, and H151C ttGGBP conjugates.

Conjugate	Emission wavelengths (nm)		Affinity (mM) ^a	
	λ_1	λ_2	^{app} K_d	^{true} K_d
11C·Badan β Zif5-IAF	480	565	0.16	0.16
11C·Badan β Zif·Alexa532	480	565	0.14	0.14
111C·Acrylodan β Zif·Alexa532	495	575	0.015	0.012
151C·Acrylodan β Zif·Alexa532	520	575	0.28	0.11
151C·Badan β Zif·Alexa532	515	580	0.01	0.002

^aDetermined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations 20-25.

Table 7. Ligand affinities of ngmFRET systems of affinity-tuned mutants within the F17C and W182C ttGGBP background.

Conjugate	Emission wavelengths (nm)		^{app} K_d (mM) ^a			
			Glucose		Galactose	
	λ_1	λ_2	R91,Q148	R91K,Q148E	R91,Q148	R91K,Q148E
17C·Badan β Zif5-IAF	467	520	0.2		0.2	
17C·Badan β Zif·Alexa532	467	560	0.2		0.2	
17C·Pacific Blue β Zif5-IAF	456	520	31		19	
182C·Acrylodan β Zif5-IAF	465	520	2.0	4.4	3.7	18.4
182C·Acrylodan β Zif·Alexa532	480	549	1.7		3.6	
182C·5-IAF β Zif·Pacific Blue	455	520	2.9	9.9	nb ^b	
182C·Oregon Green β Zif·Pacific Blue	455	520	1.8	11.6	nb ^b	
182C·IAEDANS β Zif5-IAF	465	520		0.1		5.3

^aDetermined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations 20-25.

^bNo binding. Phase II response only.

Improvements of dichromatic responses.

Case 1: enhancing intensities in strong ratiometric signals (models $d\phi$ and $d^+\phi^+$).

Single Badan and Acrylodan conjugates attached to the F17C and W182C cysteine mutations within the glucose-binding site of ttGGBP exhibit strong dichromatic responses to glucose (FIG. 15). Remarkably, the emission intensity maxima shift in opposite directions at these two positions: glucose binding evinces a bathochromic shift for 182C·Acrylodan, whereas F17C·Badan exhibits a hypsochromic response. Analysis of the emission intensity (FIG. 16) and absorption spectra (FIG. 17) indicated that glucose binding alters the populations of two dominant electronic transitions in both the excited (S_1 and S_2) and ground (G_1 and G_2) states. These two conjugates therefore switch dipoles in response to glucose binding. In the F17C·Badan hypsochromic response, the dominant excited state electronic transition shifts from S_1 (green) in the apo-protein, to S_2 (blue) in the glucose complex; in the 182C·Acrylodan bathochromic response, the opposite redistribution is observed, and the glucose complex is dominated by the S_1 (green) excited state. Similarly, in the ground state, the hypsochromic response shift the electronic transitions in the absorbance spectra from a low- (G_1) to a high-energy state (G_2); whereas in the bathochromic response the shift is $G_2 \rightarrow G_1$.

Doubly-labeled conjugates were prepared, using C-terminal β Zif fusions, in which F17C·Badan and W182C·Acrylodan function as directly responsive donors and β Zif-5-IAF or β Zif-Alexa532 as indirectly responsive acceptors (Table 7). These two acceptors have different spectral characteristics and degrees of spectral overlap with their donors.

All four conjugates exhibit significant changes in fluorescence intensities in response to glucose (FIG. 18). These responses can be divided into two glucose concentration phases. In Phase I which covers the 0-50 mM glucose concentration range, the directly responsive donors exhibited large changes in emission intensities (increases for F17C·Badan, decreases for W182C·Acrylodan conjugates), whereas the indirectly responsive acceptors showed only small, but discernable changes. At >50 mM glucose concentrations in Phase II, the intensities of both directly responsive donor and indirectly responsive acceptor peaks increase in concert.

The singly labeled 17C·Badan and 182C·Acrylodan conjugates are good ratiometric sensors. Nevertheless, their incorporation into an ngmFRET system provides additional advantages. First, it extends the wavelength range over which intensities are sufficiently

bright to be measured precisely. This is particularly clear for the β Zif·Alexa532 acceptors, which provide bright signals around 560 nm whereas neither the singly labeled Badan nor the Acrylodan constructs are particularly bright at this wavelength (compare FIGS. 15 and 18). Second, the ngmFRET systems increase the brightness of the sensors over the entire range.

5 For instance, the glucose complex of the singly labeled 182C·Acrylodan conjugate is approximately 50% less bright compared to the apo-protein (Fig. 15b), whereas these two forms differ by only ~20% when partnered with β Zif·Alexa532.

In Phase I, both the Alexa532 and the 5-IAF indirectly responsive partners decreased in emission intensity in response to the F17C·Badan directly responsive partner, whereas

10 when partnered with W182C·Acrylodan, they showed an increase in intensity. This observation is consistent with the two directly responsive partners undergoing dipole switches in the opposite direction: F17C·Badan, green→blue (hypsochromic shift; dim→bright; model $d^-\phi$, Table 1); 182C·Acrylodan, blue→green (bathochromic shift; bright→dim; model $d^+\phi$, Table 1).

15 In Phase II, at glucose concentrations in excess of ~50 mM, the intensities of both donor and acceptor peaks increase in all four conjugates. This is consistent with a second, low-affinity glucose binding site, as observed for the directly responsive 5-IAF and Oregon Green acceptor system described above. Although both donor and acceptor peak intensities increase, the QA/QD decreases, which is inconsistent with only a decrease in donor

20 quenching; a change in ngmFRET is also implicated (model $d^-\phi$).

In addition to using C-terminal β Zif fusions, we also attached indirectly responsive ngmFRET partners to N- and C-terminal ecTRX fusions (Table 8). For these experiments, we used the affinity-tuned 17C.3 and 182C.2 variants (Table 4). In all cases, we used Alexa352 as the indirectly responsive ngmFRET partner. Surprisingly, ngmFRET was not

25 established in either the N- or C-terminal fusions between 182C.2·Acrylodan and ecTRX (Table 8), nor in the N-terminal ecTRX-17C.3·Badan fusion; only the C-terminal 17C.3·Badan-ecTRX exhibited coupling between the directly responsive donor and indirectly responsive acceptor fluorophores (FIG. 19). The response of this fusion exhibited pattern similar to that observed for its β Zif counter-part, indicating that, once established, the

30 ngmFRET mechanism operates in a manner consistent with the character of the directly responsive fluorophore and the logic of the donor-acceptor arrangement. This results also indicates that the nature of the site-specific attachment point of the indirectly responsive

partner is critical. For instance, the ecTRX fusions could have placed the acceptor too far away, or in the incorrect relative orientation relative to the donor for efficient ngmFRET.

Table 8. Ligand affinities of ttGGBP-Thioredoxin fusion ngmFRET pairs^a.

Conjugate	Emission wavelengths (nm)		Glucose (mM)	
	λ_1	λ_2	^{app} K_d	^{true} K_d
Trx·Alexa532-ttGGBP17C.3·Badan	467	543	3.1	3.1
ttGGBP17C.3·Badan-Trx·Alexa532	467	547	3.6	3.7
Trx·Alexa532-ttGGBP182C.2.0·Acrylodan	475	515	4.7	3.2
ttGGBP182C.2.0·Acrylodan-Trx·Alexa532	475	519	5.8	4.3

^aDetermined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations 20-25.

Case 2: enhancing a weak ratiometric signal (model $d\phi$).

The W182C·IAEDANS- β Zif·5-IAF (fluorescein) construct is another example of a directly responsive donor- indirectly responsive acceptor ngmFRET pair. The singly labeled W182C·IAEDANS conjugate exhibits a modest hypsochromic shift in response to glucose, which, to a first approximation, involves two electronic transitions (FIG. 20). Gaussian fits revealed that the population of a green state dominates in the apo-protein and exchanges for a shorter wavelength (blue) form in the glucose complex (FIG. 20). This behavior is reminiscent to that of the other naphthalene-based fluorophore, Acrylodan, suggesting that IAEDANS also undergoes a glucose-mediated dipole switch.

The W182C·IAEDANS- β Zif·5-IAF ngmFRET pair exhibits a concerted increase in the directly responsive IAEDANS donor and indirectly responsive 5-IAF acceptor emission intensities in response to glucose (FIG. 20C). As with the W182C·Acrylodan conjugate, this response also can be divided into two phases, corresponding to high- and low-affinity glucose binding sites (FIG. 20D). In both phases, the response pattern is consistent with donor dipole switching in which a decrease in spectral overlap due to the hypsochromic shift is accompanied by a decrease in non-radiative decay (model $d\phi$, Table 1).

The singly labeled Badan conjugate attached at position 11C undergoes a green→blue hypsochromic shift in response to glucose binding with a concomitant overall increase in intensity (Table 3). Partnering this directly responsive donor with a β Zif·IAF or

β Zif·Alexa532 indirectly responsive acceptor results in an ngmFRET system that exhibits increases in the emission intensities of both donor and acceptor peaks, and a decrease in their ratio (Table 6), consistent with the $d\phi$ model (Table 1). The wavelength separation of the donor and acceptor peaks is sufficiently large in the 11C·Badan - β Zif·IAF to obtain a
5 dichromatic signal.

EXAMPLE 3: ENGINEERING RATIOMETRIC SIGNALS USING MONOCHROMATIC CHEMOSENSORS AND ADAPTOR PROTEINS.

Here we demonstrate that ngmFRET can be extended to convert monochromatic
10 chemosensors to dichromatic responses, by using an “adaptor” protein to pair directly and indirectly responsive fluorophores via site-specific, orthogonal conjugation chemistries. We used the pH-dependent response of Fluorescein (Martin 1975) as an example of a monochromatic chemosensor. Ionization of the phenolic hydroxyl (FIG. 21) has a pK_a value of 6.7, and is accompanied by a change in emission intensity, with the dianion having a
15 higher quantum yield than the monoanion. Ionization is accompanied by a bathochromic shift of the absorption and excitation spectra, enabling ratiometric measurements based on absorption or fluorescence excitation, but not emission wavelengths (Lin 2000, Han 2010, Wencel 2014). Fluorescein was paired as the chemoresponsive acceptor with Pacific Blue (Sun 1998) or Acrylodan (Prendergast 1983) as pH-insensitive ngmFRET donors. The
20 two fluorophores were attached to adaptor proteins constructed out of mutants of *Escherichia coli* thioredoxin (ecTRX), a member of the thioredoxin protein superfamily (Holmgren 1985, Qi 2005, Yoshioka 2015, Arner 2000), which were engineered to enable orthogonal, site-specific conjugation chemistries at two independently addressable sites.

Two different orthogonal fluorophore attachment chemistries were engineered. In the
25 first (Adaptor1.0 and 3.0-16.0), the existing ecTRX disulfide was combined with an engineered surface cysteines (FIG. 22). This arrangement enables site-specific attachment of one thiol-reactive fluorophore to the surface, keeping the disulfide oxidized, and a second to the disulfide, following reductive deprotection. Adaptor1.0 is remarkably thermostable, with a T_m value of at least 100 °C at neutral pH. In the second adaptor protein (Adaptor2.0), we
30 created a “reduced alphabet” protein that lacks lysines, leaving the N-terminus as the only primary amine. This arrangement enables site-specific attachment of an amine-reactive fluorophore to the adaptor N-terminus, and a second fluorophore to the reduced disulfide.

The thermostability of this protein is more modest at ~60 °C, but sufficient for use in instrumentation.

Construction of adaptor proteins.

5 Wild-type ecTRX is a 108-residue α/β protein that contains a single disulfide linkage within a surface loop (Katti 1990) (FIG. 22). As a starting point for building adaptor proteins (Adaptor0) we introduced several mutations that remove an adventitious Cu(II)-binding site at the N-terminus (D2A) and charged residues buried within the hydrophobic core (D26A, K27M) that tune the redox potential of the disulfide (Ladbury 1993, Hellinga 1992, Langsetmo 1991b, Langsetmo 1991a) (FIG. 23). A hexahistidine purification tag was fused to the C-terminus. In Adaptor1.0 and 3.0-16.0, we further introduced surface cysteine residues. These designs enabled site-specific dual labeling with thiol-reactive probes, the first to the surface thiol, and the second to the disulfide following a reductive deprotection step.

15 Adaptor2.0a and 2.0b are highly unusual proteins in which we limited the amino acid alphabet to 19 residues, and replaced all remaining nine, surface lysines with arginine (2.0a) or a combination of arginine and glutamine (2.0b). In these designs, thiol-reactive probes are attached at the reduced disulfide, and amine-reactive probes at the N-terminus.

The response of Adaptors 1.0 and 3.0-16.0 conjugates to pH.

20 Four dually labeled Adaptor1.0 conjugates were prepared, attaching the thiol-reactive Fluorescein derivative 5-IAF to the surface thiol or reduced disulfide, and Acrylodan or Pacific Blue to the corresponding, orthogonally reactive site. All conjugates were illuminated at 365 nm in a Nanodrop3300 spectrofluorimeter, and their emission intensity spectra recorded as a function of pH. All four conjugates exhibited ngmFRET from the indirectly responsive Acrylodan or Pacific Blue donor to the directly responsive Fluorescein acceptor (FIG. 24A-D, Table 9). In all cases, to a first approximation, the intensity of the Fluorescein acceptor decreased, but not the Acrylodan or Pacific Blue donor changed. This pattern of intensity changes is characteristic of a ngmFRET system in which the directly responsive acceptor quenching changes ($a^+ \phi^0$, Table 1). The increase of the Fluorescein emission intensity with pH is consistent with the higher quantum yield of the dianion compared to the monoanion.

Table 9. pK_a values of conjugates measured by emission and absorbance intensities^a.

Adaptor	Donor	Fluorescein attachment	pK_a	
			Emission	Absorbance
1.0	Acrylodan ^b	R73C	6.2	6.4
		Disulfide	6.1	6.6
	Pacific Blue ^c	R73C	6.2	6.8
		Disulfide	6.2	6.7
2.0a	Acrylodan ^b	N-terminus	5.8	6.1
2.0b	Acrylodan ^b	N-terminus	5.8	

^aDetermined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations 20-25.

5 ^bRatiometry: Absorbance, $\lambda_1=400$ nm and $\lambda_2=496$ nm; Emission, $\lambda_1=465$ nm and $\lambda_2=520$ nm.

^cRatiometry: Absorbance, $\lambda_1=411$ nm and $\lambda_2=496$ nm; Emission, $\lambda_1=455$ nm and $\lambda_2=520$ nm.

10 The choice of fluorophore attachment sites affects the ngmFRET efficiencies between the donors and fluorescein acceptor, as judged by the relative intensity of the donor emission. Labeling of the disulfide rather than the surface R73C thiol with the donor gives better ngmFRET for either the Acrylodan or Pacific Blue donors. Since the distances are approximately the same in these two arrangements, any differences must be due to changes in

15 dipole-dipole orientation (Valeur 2012, Lakowicz 2006, Clegg 1995, Wu 1994, Cheung 1991). ngmFRET is more efficient in the Acrylodan than the Pacific Blue conjugates, presumably reflecting difference in spectral overlap. Fluorescein is maximally excited at 490 nm, which is better matched with the maximum emission intensity of planar Acrylodan at ~500 nm (Allert 2015), than the 450 nm maximum of Pacific Blue.

20 The absorption spectra of the dually-labeled conjugates shift in response to pH (Fig. 25A-B), with a strong peak at ~ 500 nm appearing with the formation of the phenolate dianion at high pH value (Martin 1975).

Adaptors 3.0-16.0 were labeled with Acrylodan/Fluorescein and Pacific Blue/Fluorescein pairs (Table 10). All conjugates demonstrated ngmFRET responses to

25 proton binding.

Table 10. pK_a values of Adaptor conjugates^a.

Name	Cysteine position	Fluorophore Cysteine	Fluorophore disulfide	Emission wavelength (nm)		pK_a	
				λ_1	λ_2	app pK_a	true pK_a
Ada3.0	D11C	Acrylodan	5-IAF	523	460	6.11	6.11
Ada4.0	T15C	Acrylodan	5-IAF	523	460	6.23	6.29
Ada5.0	D21C	Acrylodan	5-IAF	523	460	6.32	6.36
Ada6.0	W32C	Acrylodan	5-IAF	515	460	5.66	5.64
Ada7.0	M38C	Acrylodan	5-IAF	515	460	5.84	5.83
Ada8.0	E45C	Acrylodan	5-IAF	515	480	6.1	5.93
Ada9.0	K53C	Acrylodan	5-IAF	523	460	6.2	6.14
Ada10.0	Q63C	Acrylodan	5-IAF	519	465	5.9	5.98
Ada11.0	Y71C	Acrylodan	5-IAF	515	460	5.95	5.92
Ada12.0	V87C	Acrylodan	5-IAF	523	485	6.22	6.23
Ada13.0	V92C	Acrylodan	5-IAF	523	540	6.78 ^b	7.05 ^b
Ada14.0	K97C	Acrylodan	5-IAF	519	465	6.15	6.16
Ada15.0	Q99C	Acrylodan	5-IAF	523	540	6.54	6.83
Ada16.0	A106C	Acrylodan	5-IAF	523	465	6.26	6.32
Ada3.0	D11C	Pacific Blue	5-IAF	455	520	5.68	6.0
Ada4.0	T15C	Pacific Blue	5-IAF	455	523	5.85	6.21
Ada5.0	D21C	Pacific Blue	5-IAF	455	523	5.88	6.23
Ada6.0	W32C	Pacific Blue	5-IAF	455	520	5.92	6.18
Ada7.0	M38C	Pacific Blue	5-IAF	455	520	6.0	6.21
Ada8.0	E45C	Pacific Blue	5-IAF	455	520	5.89	6.11
Ada9.0	K53C	Pacific Blue	5-IAF	455	520	5.73	6.0
Ada10.0	Q63C	Pacific Blue	5-IAF	455	523	5.76	6.27
Ada11.0	Y71C	Pacific Blue	5-IAF	455	520	5.83	6.13
Ada12.0	V87C	Pacific Blue	5-IAF	455	525	5.78	6.17
Ada13.0	V92C	Pacific Blue	5-IAF	455	523	5.87	6.23
Ada14.0	K97C	Pacific Blue	5-IAF	455	523	5.83	6.23
Ada15.0	Q99C	Pacific Blue	5-IAF	455	523	5.89	6.2
Ada16.0	A106C	Pacific Blue	5-IAF	455	523	5.8	6.26

^aDetermined by fitting ratiometric signal of the intensities measured at λ_1 and λ_2 to equations 20-25.^bPoor fit and/or noisy data.

The response of Adaptor2.0 to pH.

The single primary amine at the amino terminus of each lysine-free Adaptor2.0 proteins was labeled with the amine-reactive Fluorescein derivative FAM. Acrylodan was attached to the disulfide in a second labeling reaction. Adaptor2.0a and 2.0b both exhibited pH-sensitive ngmFRET between the Acrylodan donor and Fluorescein acceptor (FIG. 26A-B). Adaptor2.0a was more soluble than Adaptor2.0b at low pH values, consistent with their difference in number of surface charged groups.

The relatively high Acrylodan donor emission intensity indicates that the ngmFRET efficiency of Adaptor2.0 is less than that of the Adaptor1.0 constructs. This is likely to be a consequence of orientation effects, as the distances between the donors and acceptors are

similar in both adaptor proteins. Furthermore, the donor exhibits two emission peaks, suggesting that the Acrylodan conjugate is located in multiple environments, or adopts both twisted and planar conformations (see above). These results indicate that the environment of the disulfide attachment point differs in these two adaptor proteins, which may be a
 5 consequence of the surface redecoration with arginine residues.

The intensity of the shorter wavelength donor peak (460 nm), which is well-separated from the main Fluorescein emission wavelength (520 nm), does not change in response to pH, suggesting that the ngmFRET efficiency remains constant, consistent with a mechanism based on alteration of the non-radiative decay of the Fluorescein in response to pH, as was
 10 the case for the Adaptor1.0 conjugates. The intensities of neither the acceptor emission, nor its 500 nm absorption band exhibit large dependencies on pH, although the direction of change is the same as for Adaptor1.0. This suggests that the effective charge on the phenolate is reduced in this construct, consistent with formation a hydrogen bond or salt bridge. K36 is located near to the ecTRX N-terminus (FIG. 22), and is a possibly candidate
 15 for forming such an interaction (K57 is adjacent to the N-terminus, but has been mutated to methionine in the adaptor proteins).

Adaptor protein thermostability.

The temperature dependence of the pH responses was determined for the Adaptor1.0 R73C·Pacific Blue, disulfide·Fluorescein conjugate using a Roche LightCycler (FIG. 27).
 20 The thermostability of the conjugate is high, increasing from ~87 °Cat pH 4.5 to >100 °C at neutral pH and above. The temperature dependence of the pK_a value is modest, with a $\Delta C_p = 0.09$ kcal/mol/K.

The thermostability of the Adaptor2.0·Fluorecein, disulfide·Acrylodan is ~60 °Cat neutral pH (FIG. 27). Unlike Adaptor1.0, thermostability does not vary with pH. The
 25 temperature dependence of the pK_a also is modest, with a $\Delta C_p = 0.004$ kcal/mol/K.

The effect of the adaptor protein on Fluorescein pK_a values.

The pK_a values of the excited states, as determined from the emission spectra, are consistently lower than the corresponding ground state values obtained from the absorption spectra (Table 9). This effect is consistent with a Förster cycle in which the excited state base
 30 has a longer emission wavelength than its conjugate acid such that excited state is more acidic than the ground state (Valeur 2012).

The excited and ground state pK_a values of the Adaptor2.0 conjugates are significantly lower (~0.4 pH units) than those of Adaptor1.0 (Table 9). These results indicate

that the phenolate is stabilized by the Adaptor2.0 protein matrix, consistent with a proposed direct interaction with a protein side-chain (see above).

EXAMPLE 4. MATERIALS AND METHODS FOR EXAMPLES 1-3

Gene construction. In all constructs used in this study open reading frames (ORFs) encoding engineered proteins were placed behind an efficient Shine-Dalgarno ribosome-binding site, and flanked by a T7 promoter and terminator at the 5' and 3' ends respectively, using the GeneFab program(Cox et al. 2007). All DNA sequences were first designed *in silico*, and then synthesized by oligonucleotide assembly and cloned into pUC57 by

GeneWiz, Inc. (South Plainfield, New Jersey).

The *E. coli* thioredoxin gene (ecTRX) DNA sequence was taken from the electronic genomic sequence NC_012947, protein identifier YP_003038425.1. Publicly available genome sequences were obtained from the National Center of Biotechnology Information (<ftp://ftvp.ncbi.nih.gov/genomes/Bacteria/all.gbk.tar.gz>). The *T. thermosaccharolyticum* glucose-galactose-binding protein, ttGGBP amino acid sequence was taken from genome NC_014410, protein identifier YP_003852930.1. The putative leader peptide that mediates secretion of ttGGBP into the periplasm was identified by alignment with the mature form of the Escherichia coli glucose-galactose binding protein, ecGGBP, and removing the 49 amino acids N-terminal to the start of the aligned mature ecGGBP amino acid sequence. A single endogenous cysteine residue (C207 in the mature protein) was mutated to alanine, so that further cysteine mutants would install a unique thiol for site-specific labeling with fluorophores. The 18-residue β Zif amino acid sequence is or comprises TGEKPYKCPECGKSFSRS (SEQ ID NO: 42) (Smith et al. 2005). The Yellow Fluorescent Protein amino acid sequence is described in (SEQ ID NO: 149) (Grimley et al. 2013).

The wild-type genomic ecTRX DNA sequence was used to designs the adaptor mutants and the fusion proteins described in this study. The ttGGBP or YFP amino acid sequences were back-translated into DNA optimized in context of the expression construct by the OrfOpt or program designed to predict highly expressed mRNA sequences in *E. coli*. The OrfOpt simultaneously imposes AU-rich nucleotide composition at the 5' and 3' ends of the ORF, low RNA secondary structure content and favorable codon usage (Allert et al. 2010). Subsequent single and multiple point mutations were designed by preparing mutant sequences of the synthetic ORF sequences using GfMutagenesis, an in-house program that introduces point mutations into an ORF using the most prevalent codon in *E. coli* for an

amino acid. To design the sequences of the ttGGBP fusions with the β Zif peptide, the peptide sequence was first placed behind the optimized ttGGBP sequence; the DNA sequence of the resulting fusion protein was then re-optimized with OrfOpt without changing the optimized ttGGBP portion. The N- and C-terminal fusions of ecTRX and ttGGBP mutants
5 were designed by joining the respective wild-type and optimized DNA sequences without further manipulation. All designed protein sequences were terminated with a C-terminal affinity purification tag comprising hexa-histidine peptide placed behind a GGS linker to enable metal-mediated affinity purification(Hengen 1995).

Protein expression, purification, and fluorescent conjugate preparation. Proteins were
10 prepared by fermentation in *Escherichia coli* and purified by immobilized metal affinity chromatography (IMAC) as described(Sameiro 2009). In the final purification step, unlabeled proteins are immobilized on the IMAC beads and labeled overnight (4 °C, rotating end-over-end) with a thiol-reactive fluorophore (five-fold stoichiometric excess over protein). Following two rinses with buffer (100 mM NaCl, 1mM CaCl₂, 20 mM MOPS, pH 6.9) to
15 remove unincorporated label, the proteins were either labeled with a second fluorophore, or eluted from the beads. In experiments that required a second labeling step, the immobilized protein was first extensively washed (5x10 mL), followed by reduction of the cysteine residues in the β Zif disulfide-containing by addition of 1 mM TCEP. After one hour, the protein was washed (5x10 mL), the second fluorophore was added, and incubated as
20 described above. To elute labeled protein from the IMAC beads, 6 mL of elution buffer (400 mM imidazole, 500 mM NaCl, 1mM CaCl₂, 20 mM MOPS, pH 7.8) was added, and the beads removed by centrifugation. Following dialysis of the eluate against three changes of assay buffer (20 mM KCl, 1 mM CaCl₂, 20 mM MOPS, pH 7.4), using 10 kDa semi-permeable membrane (Snake Skin), the fluorescent conjugates were concentrated in a 10
25 kDa cutoff spin concentrator (Vivaspin, GE Healthcare). Their purity was assessed by SDS/PAGE.

Preparation of titration series to measure ligand-binding. 12-, 24-, or 48-point logarithmic titration series were prepared on a Tecan Freedom liquid-handling robot, using an in-house program, 'TitrationPlate', that compiles an abstract description of a multi-
30 component titration series into machine instructions for operating the robot. For glucose titrations in ttGGBP constructs, concentrations were varied from 0-1.7 M in 20 mM KCl, 20 mM MOPS (pH 7.4) supplemented with either 1 mM EGTA or 1 mM CaCl₂. For chloride titrations in ClBP constructs, concentrations were varied from 0-1.95 M in 35 mM potassium

gluconate, 20 mM MOPS (pH 7.4). For pH titrations in adaptor constructs, pH was varied from 4.0-9.5 in 20 mM KCl and 20 mM MOPS.

Determination of emission intensity spectra. Ligand- and wavelength-dependent emission intensities were recorded on a Nanodrop3300 (Thermo Scientific) at room temperature. Using the LED closest to the optimal excitation wavelength of the fluorophore (UV, 365 nm; blue, 470 nm; 'white', 460-550 nm).

Ratiometric analysis of ligand binding. Isothermal urea titrations were extracted from the fluorescent landscape or emission spectra datasets obtained as described above.

Monochromatic emission intensities I_λ (these intensities correspond to a bandpass intensity, recorded either with a physical, or by integrating in the interval $\lambda-\delta$, $\lambda+\delta$ in the case of an emission spectrum), were fit to

$$I_\lambda = {}^{apo}\beta_\lambda (1 - \bar{y}_{true}) + {}^{sat}\beta_\lambda \bar{y}_{true} \quad 12$$

where ${}^{apo}\beta_\lambda$ and ${}^{sat}\beta_\lambda$ are the fluorescence baselines associated with the ligand-free and ligand-bound states of the protein, respectively, and $\bar{y}_{true}$ the fractional saturation of the

protein (Layton and Hellinga 2010). Baseline functions can be constant, linear, or a second-order polynomial. For the ligand- and temperature-dependent fluorescence landscapes, we use a constant value for ${}^{apo}\beta_x$, but ${}^{sat}\beta_x$ is described by a linear dependence on urea concentration, $[L]$:

$${}^{sat}\beta_x = a_x + b_x [L] \quad 21$$

For a single ligand-binding site, the fractional saturation is given by

$$\bar{y} = \frac{[L]}{[L] + K_d} \quad 22$$

where $[L]$ is the ligand concentration and K_d the dissociation constant, ${}^{true}K_d$ for $\bar{y}_{true}$.

A ratiometric signal at a given point in a titration series, $R_{12}(t)$, is given by the ratio of intensities at two wavelengths, ${}^{obs}I(\lambda_1, t)$, ${}^{obs}I(\lambda_2, t)$ in the emission spectrum measured at that point:

$$R_{12}(t) = \frac{a_t {}^{obs}I(\lambda_1, t)}{a_t {}^{obs}I(\lambda_2, t)} \quad 23$$

where a_t is an attenuation factor that describes the effect of variations in sample size (i.e. the amount of observable fluorophore) in the t^{th} sample on the wavelength-independent intensity

of the entire emission spectrum. This signal removes wavelength-independent emission intensity attenuation effects due to variations in conjugate concentration, photobleaching, fluctuations in excitation source intensities, and detection efficiency (Demchenko 2010, Demchenko 2014). It is a key aspect for high-precision sensing using the reagentless fluorescently-responsive sensors described here. The ratiometric signal also can be fit to a binding isotherm:

$$R_{1,2} = {}^{apo}\beta_R (1 - \bar{y}_R) + {}^{sat}\beta_R \bar{y}_R \quad 24$$

where ${}^{apo}\beta_R$ and ${}^{sat}\beta_R$ are the baselines, and $\bar{y}_R$ the apparent fractional saturation of the protein (with ${}^{app}K_d$). In general, ${}^{true}K_d \neq {}^{app}K_d$; if both baselines are constant, a simple

relationship can be derived relating ${}^{app}K_d$ to ${}^{true}K_d$ (Grimley et al. 2013):

$${}^{app}K_d = {}^{true}K_d \frac{{}^{apo}I_{\lambda 2}}{{}^{sat}I_{\lambda 2}} \quad 25$$

where ${}^{apo}I_{\lambda 2}$ and ${}^{sat}I_{\lambda 2}$ are the emission intensities of the monochromatic signal at wavelength λ_2 of the ligand-free and ligand-bound protein, respectively.

Following a fit of the titration series using equations 20-24, a_t values can be recovered by taking the average comparison of the observed and calculated intensities at the two wavelengths:

$$a_t = \frac{1}{2} \left(\frac{{}^{calc}I(\lambda_1, t)}{{}^{obs}I(\lambda_1, t)} + \frac{{}^{calc}I(\lambda_2, t)}{{}^{obs}I(\lambda_2, t)} \right) \quad 26$$

The a_t value can then be applied to all wavelengths to obtain an emission spectrum or integrated intensity of the t^{th} titration point corrected for variations in sample size:

$${}^{corr}I(\lambda) = a_t {}^{obs}I(\lambda) \quad 27$$

where ${}^{corr}I(\lambda)$ and ${}^{obs}I(\lambda)$ are the wavelength-dependent intensities of the corrected and observed emission spectra, respectively.

The fractional error in the chemometric concentration measurement, depends on the first derivative of the binding isotherm as follows (Marvin et al. 1997):

$$\frac{\partial S}{S} = \frac{\varepsilon_{1,2}}{S} \times \left(\frac{dR_{1,2}}{dS} \right)^{-1} \quad 28$$

Where $R_{1,2}$ is the ratiometric signal (equation 23), $\varepsilon_{1,2}$ its experimental error, and ∂S is the resulting chemometric error in the concentration. We can then define a relative precision function

$$P(S) = \frac{S}{\delta S} \times \frac{1}{P_{\max}} \quad 29$$

where $P(S)$ is the relative precision at concentration S , which reaches a maximum value (i.e. lowest error), $P_{\max}$, at the K_d .

For a given isothermal titration, values for $^{app}K_d$ and $^{true}K_d$ were obtained using a non-linear fitting algorithm in which these two parameters were simultaneously fit to the three experimental binding isotherms using equations 20 and 24, with the two monochromatic isotherms sharing the same $^{true}K_d$ value. Three separate pairs of $^{apo}\beta$ and $^{sat}\beta$ were fit in this procedure, corresponding to the two monochromatic and the ratiometric signals, respectively. Two distinct ratiometric response models can be used: coupled (both wavelengths respond to ligand); uncoupled (the second wavelength is non-responsive; i.e. remains constant). Optionally, an attenuation vector, $a(t)$ containing a_t values for each titration point (equation 26), can be refined by iterative fit cycles in which the $a(t)$ vector of a previous cycle is used to adjust the integrated intensities of the next cycle. The program 'Nanodrop3300' was used to analyze ligand binding.

Analysis of emission spectra components. Wavelength-dependent, $I(\lambda)$, emission intensities at were converted to wavenumber-dependent intensities (Valeur 2012, Lakowicz 2006), $I(\nu)$:

$$I(\nu) = \lambda^2 I(\lambda) \quad 30$$

Singular value decomposition was used for model-free identification of regions in the emission spectra that vary with respect to glucose concentration (Henry 1992). An A_{mn} data matrix was constructed by recording $I(\nu)$ values of m frequencies in columns for n titration points in rows. This matrix was decomposed as

$$A_{mn} = U_{mn} S_{nn} V_{nn}^T \quad 31$$

where U_{mn} records n spectral components at m frequencies ranked by the weight of their contribution to the reconstruction of the experimental data, V_{nn} records the contribution of the n^{th} component to the n^{th} titration point, S_{nn} records the weight of the n^{th} component. Decomposition was carried using the in-house Nanodrop3300 program, written in Python. The `linalg.svd` method in the open-source Python `scipy` package (www.scipy.org, version 0.7.2) was used to solve the decomposition. The relative weight of the n^{th} component in U_{mn} , f_n , was calculated from S_{nn} , by normalizing the values in S with its trace:

$$f_n = \frac{S_{nn}(n, n)}{\text{tr}(S_{nn})} \quad 32$$

The fractional states of n individual electronic transitions in a spectrum were determined by fitting n Gaussians (Valeur 2012, Lakowicz 2006) to the emission intensities of the corrected spectra (equation 27) transformed into the frequency domain (equation 30):

$$^{calc}I_i(\nu) = \sum_{i=1}^{i=n} \frac{A_i}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\nu-\mu_i}{\sigma}\right)^2} \quad 33$$

5 where μ_i is the wavenumber corresponding to the peak intensity of the i^{th} transition, A_i the area contributed to the total spectrum by this transition, and σ the spectral width of all transitions. The fraction, f_i , of the i^{th} transition is given by:

$$f_i = \frac{A_i}{\sum_{j=1}^{j=n} A_j} \quad 34$$

Wavelength dependent residuals are given by:

$$10 \quad \Delta(\nu) = ^{obs}I(\nu) - ^{calc}I(\nu) \quad 35$$

Fits were carried out by minimizing the least squares difference between observed and calculated spectra, using simplex and conjugate gradient methods implemented in Nanodrop3300 (scipy package methods optimize.fmin and optimize.leastsq, respectively).

15 For titration series with N spectra, collected as a function of titrant concentration, global fits were used in which, as a first approximation, μ_i values were kept identical in the apo-protein and saturated glucose complex, and σ was universal for all transitions in all spectra. $A_{i,k}$ values were allowed to vary in each k^{th} spectrum. The variation of the fraction for each transition, $f_{i,k}$, was then fit to a binding isotherm (equation 20), constraining the fit $^{app}K_d$ value to be common to all transitions. These analyses revealed the shifts in transition dipole populations in response to ligand binding.

Analysis of ligand-binding properties using thermal melts. Protein thermal stabilities were determined by measuring the temperature-dependence of the fluorescence signal, using a Roche LightCycler (Layton and Hellinga 2010). The total fluorescence intensity, S , is given by

$$25 \quad S = \beta_F f_F + \beta_U f_U \quad 36$$

where f_F and f_U are the fractions of protein in the folded and unfolded states, respectively, and β_F and β_U the fluorescence baselines of these two states. To get the fractions of the two states, we have

$$f_N = \frac{1}{1 + K_U(T)} \text{ and } f_U = 1 - f_N \quad 37$$

where $K_U(T)$ is the temperature-dependent unfolding equilibrium constant, which by the van't Hoff approximation is given by

$$K_U = e^{-\Delta H_U \left(\frac{1}{T} - \frac{1}{T_m} \right) / R} \quad 38$$

- 5 Where T is the temperature, T_m , the unfolding reaction transition mid-point temperature, and ΔH_U the enthalpy of unfolding.

To obtain the temperature dependence of the binding reaction, the K_d values of all the individually determined isotherms were fit the Gibbs-Helmholtz equation (Layton and Hellinga 2010):

$$10 \quad \Delta G_b^*(T) = \Delta^{ref} H_b^* + \Delta C_{p,b} (T - T_{ref}) - T \left(\Delta^{ref} S_b^* + \Delta C_{p,b} \ln \frac{T}{T_{ref}} \right) \quad 39$$

where $\Delta G_b^*(T)$ is the standard free energy of binding at 1 M ligand at temperature T ,

$$\Delta G_b^*(T) = -RT \ln \left(1 + \frac{1}{K_d(T)} \right) \quad 40$$

- $\Delta^{ref} H_b^*$ and $\Delta^{ref} S_b^*$ the molar enthalpy and entropy of binding, respectively, at the reference temperature, T_{ref} , and $\Delta C_{p,b}$ the heat capacity of the binding reaction. This data analysis was
15 carried out using 'TitrationMeltAnalysis'.

RATIOMETRIC ngmFRET BIOSENSORS AND USES THEREOF

- The energy transfer between a donor-acceptor pair can be responsive to ligand-mediated changes in the photophysics of a single partner, without involving a change in the
20 geometry (distance, angle) between the two partners (non-geometrically modulated FRET, ngmFRET). Changes in the individual fluorophore properties include alterations in relative rates of excited state radiative and non-radiative decay rates (quenching) and switching between different excited state dipoles (dipole switching). The latter can alter both the emission intensity spectrum, and the absorption spectrum if the ground state electronic
25 structure is affected by ligand binding. The analysis that we present (equations 8-19) shows that the detailed mechanism by which quenching and dipole switching affects ngmFRET depends on whether the ligand-responsive fluorophore functions as the donor or acceptor in the system (Table 1). Dipole switching affects ngmFRET by altering the spectral overlap

between the two partners. Accordingly, if the directly responsive partner is the donor, then effects are seen only if ligand binding modulates the excited state (emission), whereas if it is the acceptor, only ground state (absorption) effects are observable. Quenching (change in the ratio of radiative to non-radiative emission rates) effects in directly responsive donors alter the ngmFRET transfer rate, thereby affecting the emission intensities of both donor and acceptor; directly responsive acceptor quenching affects only its own intensity. The possible choices in quenching, dipole switching and directly responsive partner functional role together combine into sixteen different scenarios by which ngmFRET occurs (Table 1).

We have constructed over 15 different donor-acceptor pairs in three different protein-based, semi-synthetic, fluorescent biosensors, all of which show ratiometric responses to binding of their cognate ligands (Table 11). Together these conjugates represent five of the sixteen possible ngmFRET scenarios. Pure geometry-based (tgmFRET) effects alter the donor and acceptor emission intensities in opposing directions only. Seven conjugates exhibit parallel ligand-mediated intensity increases or decreases, which clearly indicate the presence of non-geometrical ngmFRET effects. Quantitative differential differences between the donor and acceptor changes the shape of the emission spectrum. In three rigid conjugates, the intensity of a donor changes when partnered with directly responsive acceptor. This effect can occur only if the acceptor ground state absorption spectrum changes in response to ligand binding, providing clear evidence for dipole switching effects.

Table 11. Summary of observed intensity response patterns and ngmFRET mechanisms.

Construct	Ligand	Response pattern ^a			Model ^b
		Donor	Acceptor	A/D ratio	
YFP 17C·Pacific Blue ^c	Chloride	↑	↓	↓	$a^+ \phi$
ttGGBP 11C·Badan β Zif·5-IAF	Glucose	↑	↓	↓	$d \phi$
ttGGBP 11C·Badan β Zif·Alexa532		↑	↓	↓	$d \phi$
ttGGBP 17C·Pacific Blue β Zif·5-IAF		↑	↑	↑	$d \phi^0$
ttGGBP 17C·Badan β Zif·5-IAF ^d		↑	↓	↓	$d \phi$
ttGGBP 17C·Badan β Zif·Alexa532 ^d		↑	↓	↓	$d \phi$
ttGGBP 17C·Badan C-ecTRX·Alexa532		↑	↓	↓	$d \phi$
ttGGBP 111C·Acrylodan β Zif·Alexa532		↑	↑	↓	$d \phi^0$
ttGGBP 151C·Acrylodan β Zif·Alexa532		↓	↓	↑	$d^+ \phi^0$
ttGGBP 151C·Badan β Zif·Alexa532		↑	↑	↑	$d \phi^0$
ttGGBP 182C·Acrylodan β Zif·5-IAF		↓	↑	↑	$d^+ \phi^+$
ttGGBP 182C·Acrylodan β Zif·Alexa532 ^d		↓	↑	↑	$d^+ \phi^+$
ttGGBP 182C·IAEDANS β Zif·5-IAF ^d		↑	↑	↓	$d \phi$
Adaptor 1.0 ^e disulfide·5-IAF C73·Pacific Blue	H ⁺	↑	↓	↓	$a^+ \phi$
Adaptor 1.0 ^e disulfide·5-IAF C73·Acrylodan		↑	↓	↓	$a^+ \phi$

^aChange in magnitude (↑, increase; ↓, decrease; 0, no change) of the donor and acceptor fluorescence intensities and their ratio.

^bSee Table 1

^cResponse pattern and model are the same for all members of the chloride-binding protein series CIBP2-14.

^dPhase I of the response to glucose (see main text).

^eResponse pattern and model are the same for all members of the Adaptor family.

Fluorescently responsive sensors take advantage of ligand-mediated changes in the photophysics of a fluorophore. Analyte binding shifts the population distribution of the two fluorophore states whose photophysical properties are associated with the ligand-free and – bound forms of the sensor, respectively. Ratiometric measurements of differential changes at two distinct wavelengths are optimal for taking advantage of ligand-mediated changes in emission intensities, because these remove unwanted fluctuations in absolute intensity. Ratiometry is possible only if the shape of the fluorescence emission intensity spectrum differs in the ligand-free and –bound states (dichromatic changes). Many sensors are based on fluorescent protein conjugates or chemoresponsive fluorophores. Unfortunately, the

emission intensity spectra of many of these materials change only in intensity and not shape (monochromatic responses), precluding their use in ratiometry. The ngmFRET systems presented here can readily convert monochromatic into dichromatic responses, thereby overcoming the limitations and drawbacks of earlier systems. Furthermore, we have shown that ratiometric responses in a single-fluorophore conjugate can be improved if the emission of one of the switched dipoles is coupled to a significantly brighter acceptor. We have also shown that optimal ratiometry requires careful tuning of the spectral overlap in an ngmFRET system such that both donor and acceptor exhibit significant emission intensities.

The ngmFRET approach can be applied to many important sources of fluorescent sensors. Non-limiting examples include ligand binding proteins for sugars (such as glucose, galactose, lactose, arabinose, ribose, and maltose), lactate, urea, anions (*e.g.*, bicarbonate, phosphate, sulfate, and halide anions such as chloride, fluoride, iodide, astatide, ununseptide, and bromide), cations (*e.g.*, calcium and hydrogen ions), dipeptides, amino acids (such as histidine, glutamine, glutamate, aspartate), and elements (*e.g.*, iron).

REFERENCES

- Ahmed, M. U., I. Saaem, P. C. Wu & A. S. Brown (2014) Personalized diagnostics and biosensors: a review of the biology and technology needed for personalized medicine. *Crit Rev Biotechnol*, 34, 180-96.
- Allert, M., J. C. Cox & H. W. Hellinga (2010) Multifactorial determinants of protein expression in prokaryotic open reading frames. *J Mol Biol*, 402, 905-18.
- Allert, M. J., Kumar, S. Miriyala, J., Bergeron, A., Beese, L.S., Hellinga, H.W. (2015) Conformational coupling between ligand binding and internal fluorophore twisting in reagentless fluorescent glucose biosensors derived from periplasmic binding proteins. *In preparation*.
- Arner, E. S. J., Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.*, 267, 6102-6109.
- Arora, A., G. Simone, G. B. Salieb-Beugelaar, J. T. Kim & A. Manz (2010) Latest developments in micro total analysis systems. *Anal Chem*, 82, 4830-47.
- Badugu, R., J. R. Lakowicz & C. D. Geddes (2005) A glucose-sensing contact lens: from bench top to patient. *Curr Opin Biotechnol*, 16, 100-7.

- Bandodkar, A. J., W. Jia, C. Yardimci, X. Wang, J. Ramirez & J. Wang (2015) Tattoo-Based Noninvasive Glucose Monitoring: A Proof-of-Concept Study. *Anal. Chem.* (Washington, DC, U. S.), 87, 394-398.
- 5 Bergel, A., J. Soupe & M. Comtat (1989) Enzymatic amplification for spectrophotometric and electrochemical assays of NAD⁺ and NADH. *Anal Biochem*, 179, 382-8.
- Berntsson, R. P., S. H. Smits, L. Schmitt, D. J. Slotboom & B. Poolman (2010) A structural classification of substrate-binding proteins. *FEBS Lett*, 584, 2606-17.
- Borisov, S. M. & O. S. Wolfbeis (2008) Optical biosensors. *Chem Rev*, 108, 423-61.
- Burtis, C. A., Ashwood, E.R., Bruns, D.E. 2012. *Tietz Textbook of Clinical Chemistry and*
10 *Molecular Diagnostics*. Elsevier.
- Cash, K. J. & H. A. Clark (2010) Nanosensors and nanomaterials for monitoring glucose in diabetes. *Trends Mol Med*, 16, 584-93.
- Cheung, H. C. (1991) Resonance energy transfer. *Topics in Fluorescence Spectroscopy*, 2, 127-176.
- 15 Choleau, C., J. C. Klein, G. Reach, B. Aussedat, V. Demaria-Pesce, G. S. Wilson, R. Gifford & W. K. Ward (2002) Calibration of a subcutaneous amperometric glucose sensor. Part 1. Effect of measurement uncertainties on the determination of sensor sensitivity and background current. *Biosens Bioelectron*, 17, 641-6.
- Clegg, R. M. (1995) Fluorescence resonance energy transfer. *Curr. Opin. Biotechnol.*, 6, 103-
20 110.
- Cox, J. C., J. Lape, M. A. Sayed & H. W. Hellinga (2007) Protein fabrication automation. *Protein Sci*, 16, 379-90.
- de Graaf, D. B., Abbas, Y., Bomer, J.G., Olthuis, W., van den Berg, A. (2015) Sensor-actuator system for dynamic chloride determination. *Anal. Chim. Acta*, 888, 44-51.
- 25 de Lorimier, R. M., J. J. Smith, M. A. Dwyer, L. L. Looger, K. M. Sali, C. D. Paavola, S. S. Rizk, S. Sadigov, D. W. Conrad, L. Loew & H. W. Hellinga (2002) Construction of a fluorescent biosensor family. *Protein Sci*, 11, 2655-75.
- Demchenko, A. P. (2010) The concept of lambda-ratiometry in fluorescence sensing and imaging. *J Fluoresc*, 20, 1099-128.
- 30 Demchenko, A. P. (2014) Practical aspects of wavelength ratiometry in the studies of intermolecular interactions. *Journal of Molecular Structure*, 1077, 51-67.

- Elsiger, M. A., Wachter, R.M., Hanson, G.T., Kallio, K., Remington, S.J. (1999) Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry*, 38, 5296-5301.
- 5 Gifford, R., J. J. Kehoe, S. L. Barnes, B. A. Kornilayev, M. A. Alterman & G. S. Wilson (2006) Protein interactions with subcutaneously implanted biosensors. *Biomaterials*, 27, 2587-98.
- Grimley, J. S., L. Li, W. Wang, L. Wen, L. S. Beese, H. W. Hellinga & G. J. Augustine (2013) Visualization of synaptic inhibition with an optogenetic sensor developed by cell-free protein engineering automation. *J Neurosci*, 33, 16297-309.
- 10 Grunewald, F. S. 2014. Periplasmic binding proteins in biosensing applications. In *BIOREV*, 205-236. Springer Int., Switzerland.
- Gubala, V., L. F. Harris, A. J. Ricco, M. X. Tan & D. E. Williams (2012) Point of care diagnostics: status and future. *Anal Chem*, 84, 487-515.
- Han, J., Burgess, K. (2010) Fluorescent indicators for intracellular pH. *Chem Rev*, 110, 2709-15 2728.
- Hellinga, H. W., Wynn, R., Richards, F.M. (1992) The hydrophobic core of Escherichia coli thioredoxin shows a high tolerance to nonconservative single amino acid substitutions. *Biochemistry*, 31, 11203-11209.
- Hengen, P. N. (1995) Purification of His-Tag fusion proteins from Escherichia coli. *Trends Biochem Sci*, 20. 20
- Henry, E. R., Hofrichter, J. (1992) Singular value decomposition: application to analysis of experimental data. *Methods Enzymol*, 210, 129-192.
- Heo, Y. J. & S. Takeuchi (2013) Towards smart tattoos: implantable biosensors for continuous glucose monitoring. *Adv Healthc Mater*, 2, 43-56.
- 25 Holmgren, A. (1985) Thioredoxin. *Annu Rev. Biochem.*, 54, 237-271.
- Huber, C., Klimant, I., Krause, C., Werner, T., Mayr, T., Wolfbeis, O. (2000) Optical sensor for seawater salinity. *Fresenius J Anal Chem*, 368, 196-202.
- Huber, C., Klimant, I., Krause, C., Wolfbeis, O. (2001) Dual lifetime referencing as applied to a chloride optical sensor. *Anal. Chem.*, 73, 2097-2103.
- 30 Ispas, C. R. C., G.; Andreescu, S. (2012) Review: Recent Developments in Enzyme-Based Biosensors for Biomedical Analysis. *Anal. Lett.*, 45, 168-186.

- Jayaraman, S., Haggie, P., Wachter, R.M., Remington, S.J., Verkman, A.S. (2000) Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Biol. Chem.*, 275, 6047-6050.
- Judge, K., L. Morrow, A. G. Lastovich, D. Kurisko, S. C. Keith, J. Hartsell, B. Roberts, E. McVey, K. Weidemaier, K. Win & M. Hompesch (2011) Continuous glucose monitoring using a novel glucose/galactose binding protein: results of a 12-hour feasibility study with the becton dickinson glucose/galactose binding protein sensor. *Diabetes Technol Ther*, 13, 309-17.
- Karon, B. S., L. Griesmann, R. Scott, S. C. Bryant, J. A. Dubois, T. L. Shirey, S. Presti & P. J. Santrach (2008) Evaluation of the impact of hematocrit and other interference on the accuracy of hospital-based glucose meters. *Diabetes Technol Ther*, 10, 111-20.
- Katti, S., LeMaster, D.M., Eklund, H. (1990) Crystal structure of thioredoxin from *Escherichia coli* at 1.68Å resolution. *J. Mol. Biol.*, 212, 167-184.
- Koschwanetz, H. E. & W. M. Reichert (2007) In vitro, in vivo and post explantation testing of glucose-detecting biosensors: current methods and recommendations. *Biomaterials*, 28, 3687-703.
- Kozma, P., A. Lehmann, K. Wunderlich, D. Michel, S. Schumacher, E. Ehrentreich-Forster & F. F. Bier (2013) A novel handheld fluorescent microarray reader for point-of-care diagnostic. *Biosens Bioelectron*, 47, 415-20.
- Kuner, T., Augustine, G.J. (2000) A genetically encoded ratiometric indicator for chloride: capturing chloride transients in cultured hippocampal neurons. *Neuron*, 27, 447-459.
- Ladbury, J. E., Wynn, R., Hellinga, H.W., Sturtevant, J.M. (1993) Stability of oxidized *Escherichia coli* thioredoxin and its dependence on protonation of the aspartic acid residue in the 26 position. *Biochemistry*, 32, 7526-7530.
- Lakowicz, J. R. 2006. *Principles of fluorescence spectroscopy*. Springer, New York.
- Langsetmo, K., Fuchs, J.A., Woodward, C. (1991a) The conserved, buried aspartic acid in oxidized *Escherichia coli* thioredoxin has a pKa of 7.5. Its titration produces a related shift in global stability. *Biochemistry*, 30, 7603-7609.
- Langsetmo, K., Fuchs, J.A., Woodward, C., Sharp, K.A. (1991b) Linkage of thioredoxin stability to titration of ionizable groups with perturbed pKa. *Biochemistry*, 30, 7609-7614.
- Lavis, L. D. & R. T. Raines (2008) Bright ideas for chemical biology. *ACS Chem Biol*, 3, 142-55.

- (2014) Bright building blocks for chemical biology. *ACS Chem Biol*, 9, 855-66.
- Layton, C. J. & H. W. Hellinga (2010) Thermodynamic analysis of ligand-induced changes in protein thermal unfolding applied to high-throughput determination of ligand affinities with extrinsic fluorescent dyes. *Biochemistry*, 49, 10831-41.
- 5 Lin, J. (2000) Recent development and applications of optical and fiber-optic pH sensors. *Tr. Anal. Chem.*, 19, 541-551.
- Liu, D., Evans, T., Zhang, F. (2015) Application and advances of metabolite biosensors for metabolic engineering. *Metabolic Engin.*, 31, 35-43.
- Martin, M. M., Lindqvist, L. (1975) The pH dependence of fluorescein fluorescence. *J.*
10 *Luminesc.*, 10, 381-390.
- Marvin, J. S., E. E. Corcoran, N. A. Hattangadi, J. V. Zhang, S. A. Gere & H. W. Hellinga (1997) The rational design of allosteric interactions in a monomeric protein and its applications to the construction of biosensors. *Proc Natl Acad Sci U S A*, 94, 4366-71.
- Matzeu, G., Florea, L., Diamond, D. (2015) Advances in wearable chemical sensor design for
15 monitoring biological fluids. *Sens Actuators B Chem*, 211, 403-418.
- Miesenbock, G., D. A. De Angelis & J. E. Rothman (1998) Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature*, 394, 192-5.
- Mohammed, M. D., M.P.Y. (2011) Lab-on-a-chip based immunosensor principles and
20 technologies for the detection of cardiac biomarkers: a review. *Lab. Chip.*, 11, 569-595.
- Okumoto, S., Jones, A., Frommer, W.B. (2012) Quantitative imaging with fluorescent biosensors. *Annu. Rev. Plant Biol.*, 63, 663-706.
- Prendergast, F. G., Meyer, M., Carlson, G.L., Iida, S., Potter, J.D. (1983) Synthesis, spectral
25 properties, and use of 6-acryloyl-2-dimethylaminonaphthalene (acrylodan). *Journal of Biological Chemistry*, 258, 7541-7544.
- Qi, Y., Grishin, N.V. (2005) Structural classification of thioredoxin-like fold proteins. *Proteins*, 58, 376-388.
- Robinson, T. & P. S. Dittrich (2013) Microfluidic technology for molecular diagnostics. *Adv*
30 *Biochem Eng Biotechnol*, 133, 89-114.
- Rogers, M. L. & M. G. Boutelle (2013) Real-time clinical monitoring of biomolecules. *Annu Rev Anal Chem (Palo Alto Calif)*, 6, 427-53.

- Sameiro, M., and Goncalves, T. (2009) Fluorescent Labeling of Biomolecules with Organic Probes. *Chem. Rev.*, 109, 190-212.
- Smith, J. J., D. W. Conrad, M. J. Cuneo & H. W. Hellinga (2005) Orthogonal site-specific protein modification by engineering reversible thiol protection mechanisms. *Protein Sci*, 14, 64-73.
- Sun, W. C., Gee, K.R., Haugland, R.P. (1998) Synthesis of novel fluorinated coumarins: excellent UV-light excitable fluorescent dyes. *Bioorg Med Chem Lett*, 8, 3107-3110.
- Tang, Z., J. H. Lee, R. F. Louie & G. J. Kost (2000) Effects of different hematocrit levels on glucose measurements with handheld meters for point-of-care testing. *Arch Pathol Lab Med*, 124, 1135-40.
- Tang, Z., R. F. Louie, J. H. Lee, D. M. Lee, E. E. Miller & G. J. Kost (2001) Oxygen effects on glucose meter measurements with glucose dehydrogenase- and oxidase-based test strips for point-of-care testing. *Crit Care Med*, 29, 1062-70.
- Tsien, R. Y. (1998) The green fluorescent protein. *Annu Rev Biochem*, 67, 509-44.
- Valeur, B., Berberan-Santos, M.N. 2012. *Molecular Fluorescence. Principles and Applications*. Weinheim: Wiley.
- Vyas, N. K., M. N. Vyas & F. A. Quijcho (1988) Sugar and signal-transducer binding sites of the Escherichia coli galactose chemoreceptor protein. *Science*, 242, 1290-5.
- Wachter, R. M., Remington, S.J. (1999) Sensitivity of the yellow variant of green fluorescent protein to halides and nitrate. *Current Biology*, 9, R628-R629.
- Wang, H., Nakata, E., Hamachi, I. (2009) Recent progress in strategies for the creation of protein-based fluorescent biosensors. *Chembiochem*, 10, 2560-2577.
- Wang, J. (2008) Electrochemical glucose biosensors. *Chem Rev*, 108, 814-25.
- Wang, W., Grimley, J.S., Augustine, G.J., Beese, L.S., Hellinga, H.W. (2015) Determination of engineered chloride-binding site structures in fluorescent proteins reveals principles of halide recognition. *In preparation*.
- Weidemaier, K., A. Lastovich, S. Keith, J. B. Pitner, M. Sistare, R. Jacobson & D. Kurisko (2011) Multi-day pre-clinical demonstration of glucose/galactose binding protein-based fiber optic sensor. *Biosens Bioelectron*, 26, 4117-23.
- Wencel, D., Abel, T., McDonagh, C. (2014) Optical chemical pH sensors. *Anal. Chem.*, 86, 15-29.
- Wisniewski, N. & M. Reichert (2000) Methods for reducing biosensor membrane biofouling. *Colloids Surf B Biointerfaces*, 18, 197-219.

Wu, P., Brand, L. (1994) Resonance energy transfer: methods and applications. *Anal. Biochem.*, 218, 1-13.

Yoo, E.-H. & S.-Y. Lee (2010) Glucose biosensors: an overview of use in clinical practice. *Sensors (Basel)*, 10, 4558-76.

5 Yoshioka, J. (2015) Thioredoxin superfamily and its effect on cardiac physiology and pathology. *Comprehensive Physiol.*, 5, 513-529.

Zhang, X., J. Yin & J. Yoon (2014) Recent advances in development of chiral fluorescent and colorimetric sensors. *Chem Rev*, 114, 4918-59.

Zimmer, M. (2002) Green fluorescent protein (GFP): applications, structure, and related
10 photophysical behavior. *Chem Rev*, 102, 759-81.

OTHER EMBODIMENTS

While the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the
15 invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

The patent and scientific literature referred to herein establishes the knowledge that is available to those with skill in the art. All United States patents and published or unpublished United States patent applications cited herein are incorporated by reference. All published
20 foreign patents and patent applications cited herein are hereby incorporated by reference. Genbank and NCBI submissions indicated by accession number cited herein are hereby incorporated by reference. All other published references, documents, manuscripts and scientific literature cited herein are hereby incorporated by reference.

While this invention has been particularly shown and described with references to
25 preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

30

What is claimed is:

1. A method of detecting a ligand in a sample, comprising
 - (a) contacting a biosensor with the ligand, wherein the biosensor comprises a ligand-binding protein, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the ligand-binding-protein, wherein
 - (i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or
 - (ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore,
 - (b) contacting the biosensor with radiation comprising a wavelength within the excitation spectrum of the donor fluorophore, wherein
 - (i) a fluorescence property of the directly responsive fluorophore changes in response to ligand binding in the absence or presence of the indirectly responsive fluorophore;
 - (ii) a fluorescence property of the indirectly responsive fluorophore does not change in response to ligand binding in the absence of the directly responsive fluorophore;
 - (iii) non-geometrically modulated Förster resonance energy transfer (ngmFRET) occurs between the directly responsive fluorophore and the indirectly responsive fluorophore;
 - (iv) fluorescent light is emitted from the biosensor, wherein the light emitted from the biosensor comprises a combination of light emitted from the directly responsive fluorophore and light emitted from the indirectly responsive fluorophore; and
 - (v) wherein the ratio of the fluorescence emission intensity emitted from the biosensor at each of two distinct wavelengths changes in response to ligand binding;
 - (c) measuring fluorescent light that is emitted from the directly responsive fluorophore and the indirectly responsive fluorophore; and
 - (d) calculating a ratiometric signal, thereby detecting the ligand in the sample.

2. The method of claim 1, the ratiometric signal ($R_{1,2}$) comprises a quotient of two intensities, $I_{\lambda 1}$ and $I_{\lambda 2}$, measured at two independent wavelengths, λ_1 and λ_2 and is calculated according to the following equation:

$$R_{1,2} = I_{\lambda 1} / I_{\lambda 2} .$$

5

3. The method of claim 1, wherein the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore, and wherein ngmFRET occurs between the donor fluorophore and the acceptor fluorophore when the donor fluorophore is contacted with radiation within its excitation spectrum.

10

4. The method of claim 1, wherein the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore, and wherein ngmFRET occurs between the donor fluorophore and the acceptor fluorophore when the donor fluorophore is contacted with radiation within its excitation spectrum.

15

5. The method of claim 1, wherein the change in the fluorescent property of the directly responsive fluorophore comprises (i) a bathochromic or hypsochromic shift in the emission or excitation spectrum thereof; or (ii) a change in the ratio of radiative to non-radiative emission rates thereof.

20

6. The method of claim 1, wherein

(a) the directly responsive fluorophore is Badan and emission intensity is measured at a wavelength of 467 nm, and wherein the indirectly responsive fluorophore is 5-iodoacetamidofluorescein (5-IAF) and emission intensity is measured at a wavelength of 520 nm;

25

(b) the directly responsive fluorophore is Badan and emission intensity is measured at a wavelength of 467 nm, and wherein the indirectly responsive fluorophore is Alexa532 and emission intensity is measured at a wavelength of 560 nm;

(c) the directly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength of 456 nm, and wherein the indirectly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 520 nm;

30

(d) the directly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength of 465 nm, and wherein the indirectly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 520 nm;

(e) the directly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength of 480 nm, and wherein the indirectly responsive fluorophore is Alexa532 and emission intensity is measured at a wavelength of 549 nm;

(f) the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 455 nm, and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength of 520 nm;

(g) the directly responsive fluorophore is Oregon Green and emission intensity is measured at a wavelength of 455 nm, and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength of 520 nm; or

(h) the directly responsive fluorophore is N-(Iodoacetaminoethyl)-1-naphthylamine-5-sulfonic acid (IAEDANS) and emission intensity is measured at a wavelength of 465 nm, and wherein the indirectly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 520 nm.

7. The method of claim 1, wherein

(a) the directly responsive fluorophore is a yellow fluorescent protein and emission intensity is measured at a wavelength of 530 nm, and wherein the indirectly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength of 500 nm; or

(b) the directly responsive fluorophore is a yellow fluorescent protein and emission intensity is measured at a wavelength of 530 nm, and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength of 455 nm.

8. The method of claim 1, wherein the directly responsive fluorophore comprises a donor fluorophore and the indirectly responsive fluorophore comprises an acceptor fluorophore.

9. The method of claim 8, wherein

- (a) the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 5 (b) the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 10 (c) the emission intensities of the donor fluorophore and the acceptor fluorophore both decrease upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 15 (d) the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 20 (e) the emission intensity of the donor fluorophore increases, decreases, or remains about the same and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 25 (f) the emission intensities of the donor fluorophore and the acceptor fluorophore both increase upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;
- 30 (g) the emission intensity of the donor fluorophore increases, decreases, or remains about the same and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore; or
- (h) the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore.

10. The method of claim 1, wherein the directly responsive fluorophore comprises an acceptor fluorophore and the indirectly responsive fluorophore comprises a donor fluorophore.

5 11. The method of claim 10, wherein

(a) the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

10 (b) the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

15 (c) the emission intensity of the donor fluorophore remains about the same and the emission intensity of the acceptor fluorophore decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

20 (d) the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

25 (e) the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore increases, decreases, or remains about the same upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

(f) the emission intensity of the donor fluorophore remains about the same and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore;

30 (g) the emission intensity of the donor fluorophore decreases and the emission intensity of the acceptor fluorophore increases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore; or

(h) the emission intensity of the donor fluorophore increases and the emission intensity of the acceptor fluorophore remains about the same, increases, or decreases upon ligand binding to the ligand-binding protein when the donor fluorophore is contacted with radiation within the excitation spectrum of the donor fluorophore.

5

12. The method of claim 1, wherein the ligand-binding protein comprises the directly responsive fluorophore.

13. The method of claim 12, wherein the directly responsive fluorophore is formed by an autocatalytic cyclization of an oligopeptide within the ligand-binding protein.

10

14. The method of claim 13 wherein the oligopeptide is located within an interior α helix.

15. The method of claim 13, wherein the oligopeptide comprises three consecutive residues, four consecutive residues, or five consecutive residues.

15

16. The method of claim 12, wherein the directly responsive fluorophore is formed by an autocatalytic cyclization of a tripeptide located in an interior α helix of the ligand-binding protein.

20

17. The method of claim 12, wherein the ligand-binding protein comprises a yellow fluorescent protein (YFP).

18. The method of claim 1, wherein ligand binding causes a change in signaling by the directly responsive fluorophore.

25

19. The method of claim 14, wherein the indirectly responsive fluorophore is attached to the ligand-binding protein via a covalent bond.

20. The method of claim 19, wherein the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction.

30

21. The method of claim 19, wherein the indirectly responsive fluorophore is attached to a cysteine or a lysine of the protein.
22. The method of claim 19, wherein the indirectly responsive fluorophore is attached to
5 the N-terminus or the C-terminus of the protein.
23. The method of claim 19, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif.
- 10 24. The method of claim 23, wherein the fluorophore attachment motif comprises an amino acid or a polypeptide.
25. The method of claim 24, wherein the polypeptide comprises amino acids in the sequence of β Zif (SEQ ID NO: 42).
- 15 26. The method of claim 24, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).
- 20 27. A method of detecting a ligand in a sample, comprising
(a) contacting a biosensor with the ligand, wherein the biosensor comprises an amino acid or a polypeptide, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the amino acid or polypeptide,
25 wherein the directly responsive fluorophore is chemoresponsive, and wherein
(i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or
(ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore,
30 (b) contacting the biosensor with radiation comprising a wavelength within the excitation spectrum of the donor fluorophore, wherein

- (i) a fluorescence property of the directly responsive fluorophore changes in response to ligand binding in the absence or presence of the indirectly responsive fluorophore;
- (ii) a fluorescence property of the indirectly responsive fluorophore does not change in response to ligand binding in the absence of the directly responsive fluorophore;
- (iii) ngmFRET occurs between the directly responsive fluorophore and the indirectly responsive fluorophore;
- (iv) fluorescent light is emitted from the biosensor, wherein the light emitted from the biosensor comprises a combination of light emitted from the directly responsive fluorophore and light emitted from the indirectly responsive fluorophore; and
- (v) wherein the ratio of the fluorescence emission intensity emitted from the biosensor at each of two distinct wavelengths changes in response to ligand binding;
- (c) measuring fluorescent light that is emitted from the directly responsive fluorophore and the indirectly responsive fluorophore; and
- (d) calculating a ratiometric signal, thereby detecting the ligand in the sample.

28. The method of claim 27, the ratiometric signal ($R_{1,2}$) comprises a quotient of two intensities, I_{λ_1} and I_{λ_2} , measured at two independent wavelengths, λ_1 and λ_2 and is calculated according to the following equation:

$$R_{1,2} = I_{\lambda_1} / I_{\lambda_2} .$$

29. The method of claim 27, wherein the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore, and wherein ngmFRET occurs between the donor fluorophore and the acceptor fluorophore when the donor fluorophore is contacted with radiation within its excitation spectrum.

30. The method of claim 27, wherein the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore, and wherein

ngmFRET occurs between the donor fluorophore and the acceptor fluorophore when the donor fluorophore is contacted with radiation within its excitation spectrum.

31. The method of claim 27, wherein the change in the fluorescent property of the directly responsive fluorophore comprises (i) a bathochromic or hypsochromic shift in the emission or excitation spectrum thereof; or (ii) a change in the ratio of radiative to non-radiative emission rates thereof.

32. The method of claim 27, wherein

(a) the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 460 nm, and wherein the indirectly responsive fluorophore is Acrylodan and emission intensity is measured at a wavelength of 523 nm; or

(b) the directly responsive fluorophore is 5-IAF and emission intensity is measured at a wavelength of 520 nm, and wherein the indirectly responsive fluorophore is Pacific Blue and emission intensity is measured at a wavelength of 455 nm.

33. The method of claim 27, wherein the amino acid or the polypeptide comprises 1 amino acid, or a stretch of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, 750, or 1000 amino acids.

34. The method of claim 33, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of ecTRX (SEQ ID NO: 151).

35. The method of claim 33, wherein the amino acid or the polypeptide comprises

(a) at least 1, 2, or 3 thiol groups;

(b) at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group;

(c) at least 1, 2, or 3 primary amine groups; or

(d) at least 1, 2, or 3 lysines that each comprise a primary amine.

36. The method of claim 34, wherein there is no disulfide bond between cysteines within the amino acid sequence of the polypeptide.

37. The method of claim 34, wherein the polypeptide comprises a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

5

38. The method of claim 37, wherein the polypeptide comprises a mutant of ecTRX comprising a D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

10

39. The method of claim 34, wherein the polypeptide comprises a mutant of ecTRX that does not comprise a lysine.

40. The method of claim 34, wherein the polypeptide further comprises a hexahistidine tag.

15

41. The method of claim 34, comprising amino acids in the sequence set forth as any one of SEQ ID NOS: 24-41.

20

42. The method of claim 27, which wherein the ligand comprises a hydrogen ion.

43. The method of claim 42, which is a biosensor for pH, wherein the directly responsive fluorophore is pH-sensitive.

25

44. The method of claim 43, wherein the fully excited emission intensity of the directly responsive fluorophore is different at a pH less than about 7.0 compared to a pH of 7.5.

45. The method of claim 43, wherein the directly responsive fluorophore comprises a pH-sensitive fluorophore comprising fluorescein or a derivative thereof.

30

46. The method of claim 45, wherein the directly responsive fluorophore transitions from a monoanion to a dianion at a pH that is less than 7.0 in an aqueous solution.

47. The method of claim 1, wherein the directly responsive fluorophore is attached to the ligand-binding protein via a covalent bond.

48. The method of claim 47, wherein the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been formed by a click reaction.

49. The method of claim 47, wherein the directly responsive fluorophore is attached to a cysteine or a lysine of the protein.

50. The method of claim 47, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein.

51. The method of claim 27, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif.

52. The method of claim 27, wherein the fluorophore attachment motif comprises an amino acid or a polypeptide.

53. The method of claim 52, wherein the polypeptide comprises amino acids in the sequence of β Zif (SEQ ID NO: 42).

54. The method of claim 52, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).

55. The method of claim 1, wherein if the acceptor fluorophore comprises ruthenium or osmium, then the acceptor fluorophore is not attached to the amino group of the N-terminus of the ligand-binding protein.

56. The method of claim 1, wherein the biosensor does not comprise an *E. coli* glutamine-binding protein with Acrylodan attached to 179C.

57. The method of claim 1, wherein the biosensor does not comprise *E. coli* glucose-binding protein with Acrylodan attached to 255C.

58. The biosensor of claim 1, wherein an overlap of the emission spectrum of the donor fluorophore and the excitation spectrum of the acceptor fluorophore increases upon ligand binding.

59. The method of claim 58, wherein

- (i) the directly responsive fluorophore comprises the donor fluorophore, and the increase results from a bathochromic shift in the emission spectrum of the donor fluorophore; or
- (ii) the directly responsive fluorophore comprises the acceptor fluorophore, and the increase results from a hypsochromic shift in the excitation spectrum of the acceptor fluorophore.

60. The method of claim 1, wherein an overlap of the emission spectrum of the donor fluorophore and the excitation spectrum of the acceptor fluorophore decreases upon ligand binding.

61. The method of claim 60, wherein

- (i) the directly responsive fluorophore comprises the donor fluorophore, and the decrease results from a hypsochromic shift in the emission spectrum of the donor fluorophore; or
- (ii) the directly responsive fluorophore comprises the acceptor fluorophore, and the decrease results from a bathochromic shift in the excitation spectrum of the acceptor fluorophore.

62. The method of claim 1, wherein the directly responsive fluorophore has a monochromatic spectral change upon ligand binding.

63. The method of claim 1, wherein the directly responsive fluorophore has a dichromatic spectral change upon ligand binding.

64. The method of claim 1, wherein the emission intensity of the donor fluorophore and/or the acceptor fluorophore increases in two phases as ligand concentration increases.

65. The method of claim 1, wherein the ratio of radiative to non-radiative emission or
5 intensity of the directly responsive fluorophore increases by at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 50%, 75%, 100%, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, or 10-fold upon ligand binding to the ligand-binding protein.

66. The method of claim 1, wherein the ratio of radiative to non-radiative emission or
10 intensity of the directly responsive fluorophore decreases by at least about 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 50%, 75%, 90%, 95%, or 99% upon ligand binding to the ligand-binding protein.

67. The method of claim 1, wherein the directly responsive fluorophore and the indirectly
15 responsive fluorophore are not a naphthalene derivative.

68. The method of claim 1, wherein the directly responsive fluorophore and the indirectly responsive fluorophore are not Prodan, Acrylodan, or Badan.

20 69. The method of claim 1, wherein the directly responsive fluorophore is not a naphthalene derivative.

70. The method of claim 1, wherein the directly responsive fluorophore is not Prodan, Acrylodan, or Badan.

25 71. The method of claim 1, wherein the directly responsive fluorophore comprises xanthene, a xanthene derivative, fluorescein, a fluorescein derivative, coumarin, a coumarin derivative, cyanine, a cyanine derivative, rhodamine, a rhodamine derivative, phenoxazine, a phenoxazine derivative, squaraine, a squaraine derivative, coumarin, a coumarin derivative,
30 oxadiazole, an oxadiazole derivative, anthracene, an anthracene derivative, a boradiazaindane (BODIPY) family fluorophore, pyrene, a pyrene derivative, acridine, an acridine derivative, arylmethine, an arylmethine derivative, tetrapyrrole, or a tetrapyrrole derivative.

72. The method of claim 71, wherein the directly responsive fluorophore comprises fluorescein or a derivative thereof.

73. The method of claim 1, wherein the ligand-binding protein is selected from the group consisting of a glucose-galactose binding protein (GGBP), a glucose-binding protein, a urea-binding protein (UBP), a lactate-binding protein (LacBP), a calcium-binding protein, a calcium-bicarbonate binding protein (BicarbBP), and an iron-bicarbonate binding protein (FeBP).

74. The method of claim 1, wherein

- (a) the donor fluorophore comprises Pacific Blue and the acceptor fluorophore comprises 5-IAF or 6-IAF;
- (b) the donor fluorophore comprises Pacific Blue and the acceptor fluorophore comprises Oregon Green;
- (c) the donor fluorophore comprises IAEDANS and the acceptor fluorophore comprises 5-IAF or 6-IAF;
- (d) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises Alexa532;
- (e) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises 5-IAF or 6-IAF;
- (f) the donor fluorophore comprises acrylodan and the acceptor fluorophore comprises Pacific Blue or YFP;
- (g) the donor fluorophore comprises 5-IAF or 6-IAF and the acceptor fluorophore comprises Pacific Blue;
- (h) the donor fluorophore comprises badan and the acceptor fluorophore comprises 5-IAF or 6-IAF; or
- (i) the donor fluorophore comprises badan and the acceptor fluorophore comprises Alexa532.

75. A biosensor for a ligand comprising a ligand-binding protein, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the ligand-binding protein, wherein

(i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or

(ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore, and

5 wherein if the acceptor fluorophore comprises ruthenium or osmium, then the acceptor fluorophore is not attached to the amino group of the N-terminus of the ligand-binding protein.

76. The biosensor of claim 75, wherein the ligand-binding protein comprises the directly
10 responsive fluorophore.

77. The biosensor of claim 76, wherein the directly responsive fluorophore is formed by an autocatalytic cyclization of an oligopeptide within the ligand-binding protein.

15 78. The biosensor of claim 76, wherein the ligand-binding protein comprises a Yellow Fluorescent Protein (YFP; SEQ ID NO: 149) or a fluorescent mutant thereof.

79. The biosensor of claim 76, wherein the ligand comprises a halide anion.

20 80. The biosensor of claim 79, wherein the halide anion comprises a fluoride (F^-), chloride (Cl^-), a bromide (Br^-), an iodide (I^-), an astatide (At^-) anion, or an ununseptide (Ts^-) anion.

81. The biosensor of claim 78, wherein the mutant comprises a mutation that alters the
25 interaction of the mutant with a bound halide anion compared to YFP.

82. The biosensor of claim 78, wherein the mutant comprises a mutation that alters the affinity and/or specificity of the mutant for a halide anion compared to YFP.

30 83. The biosensor of claim 79, comprising 1 halide anion binding site.

84. The biosensor of claim 79, comprising at least 2, 3, 4, or 5 halide anion binding sites.

85. The biosensor of claim 78, wherein at least one amino acid of the YFP or the
35 fluorescent mutant thereof has been substituted with a cysteine.

86. The biosensor of claim 85, wherein the cysteine is within a first β -strand (β_1), a second β -strand (β_2), a third β -strand (β_3), a fourth β -strand (β_4), a fifth β -strand (β_5), a sixth β -strand (β_6), a seventh β -strand (β_7), an eighth β -strand (β_8), a ninth β -strand (β_9), a tenth β -strand (β_{10}), or an eleventh β -strand (β_{11}) of the YFP or the fluorescent mutant thereof.

87. The biosensor of claim 78, comprising one or more of the following substitutions: E17X, E32X, T43X, F64X, G65X, L68X, Q69X, A72X, H77X, K79X, R80X, E95X, R109X, R122X, D133X, H148X, N149X, V163X, N164X, D173X, Y182X, Q183X, Y203X, Q204X, L221X, and H231X, wherein X is any amino acid, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150.

88. The biosensor of claim 87, wherein X is C.

89. The biosensor of claim 87, comprising one or more of the following substitutions: F64L, G65T, L68V, Q69T, A72S, K79R, R80Q, H148Q, H148G, V163A, H231L, H148Q, or Q183A, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150.

90. The biosensor of claim 87, comprising an R at the 96 position, a Y at the 203 position, a S at the 205 position, and an E at the 222 position, wherein each YFP amino acid position is numbered as in SEQ ID NO: 150.

91. The biosensor of claim 75, wherein ligand binding causes a change in signaling by the directly responsive fluorophore.

25

92. The biosensor of claim 91, wherein the ligand-binding protein comprises a mutation compared to a naturally occurring protein.

93. The biosensor of claim 92, wherein at least one amino acid of the ligand-binding protein has been substituted with a cysteine.

30

94. The method of claim 92, wherein the ligand-binding protein comprises a mutant of a microbial ligand-binding protein.

95. The method of claim 94, wherein the ligand-binding protein comprises a mutant of a microbial periplasmic ligand-binding protein.

96. The method of claim 91, wherein the ligand comprises glucose, galactose, lactose,
5 arabinose, ribose, maltose, lactate, urea, bicarbonate, phosphate, sulfate, chloride, fluoride, iodide, astatide, ununseptide, bromide, calcium, a hydrogen ion, a dipeptide, histidine, glutamine, glutamate, aspartate, or iron.

97. The biosensor of claim 96, wherein the ligand-binding protein comprises a GGBP.

10

98. The biosensor of claim 97, wherein the GGBP is or is a mutant of: an *Escherichia* sp. GGBP; a *Thermoanaerobacter* sp. GGBP; a *Clostridium* sp. GGBP; a *Salmonella* sp. GGBP; a *Caldicellulosiruptor* sp. GGBP; a *Paenibacillus* sp. GGBP; a *Butyrivibrio* sp. GGBP; a *Roseburia* sp. GGBP; a *Faecalibacterium* sp. GGBP; an *Erysipelothrix* sp. GGBP; or an
15 *Eubacterium* sp. GGBP.

99. A biosensor for a ligand comprising an amino acid or a polypeptide, a directly responsive fluorophore and an indirectly responsive fluorophore, the directly responsive and the indirectly responsive fluorophores being located at two distinct sites of the amino acid or
20 polypeptide, wherein the directly responsive fluorophore is chemoresponsive, and wherein
(i) the directly responsive fluorophore is a donor fluorophore and the indirectly responsive fluorophore is an acceptor fluorophore; or
(ii) the directly responsive fluorophore is an acceptor fluorophore and the indirectly responsive fluorophore is a donor fluorophore.

25

100. The method of claim 99, wherein the amino acid or the polypeptide comprises 1 amino acid, or a stretch of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 500, 750, or 1000 amino acids.

30 101. The biosensor of claim 100, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of ecTRX (SEQ ID NO: 151).

102. The biosensor of claim 100, wherein the amino acid or the polypeptide comprises

(a) at least 1, 2, or 3 thiol groups;

(b) at least 1, 2, or 3 cysteines that each comprise a sulfhydryl group;

(c) at least 1, 2, or 3 primary amine groups; or

5 (d) at least 1, 2, or 3 lysines that each comprise a primary amine.

103. The biosensor of claim 100, wherein there is no disulfide bond between cysteines within the amino acid sequence of the polypeptide.

10 104. The biosensor of claim 101, wherein the polypeptide comprises a mutant of ecTRX comprising a D3X, K4X, K19X, D27X, K37X, K53X, K58X, K70X, R74X, K83X, K91X, K97X, or K101X mutation, or any combination thereof, wherein X is any amino acid, and wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

15 105. The biosensor of claim 104, wherein the polypeptide comprises a mutant of ecTRX comprising a D3A, K4R, K4Q, K19R, K19Q, D27A, K37R, K53M, K53R, K58M, K70R, R74C, K83R, K91R, K97R, or K101R mutation, or any combination thereof, wherein each ecTRX amino acid position is numbered as in SEQ ID NO: 151.

20 106. The biosensor of claim 101, wherein the polypeptide comprises a mutant of ecTRX that does not comprise a lysine.

107. The biosensor of claim 101, wherein the polypeptide further comprises a hexahistidine tag.

25 108. The biosensor of claim 101, comprising amino acids in the sequence set forth as any one of SEQ ID NOS: 24-41.

109. The biosensor of claim 101, which wherein the ligand comprises a hydrogen ion.

30 110. The biosensor of claim 101, which is a biosensor for pH, wherein the directly responsive fluorophore is pH-sensitive.

111. The biosensor of claim 110, wherein the fully excited emission intensity of the directly responsive fluorophore is different at a pH less than about 7.0 compared to a pH of 7.5.

5 112. The biosensor of claim 111, wherein the directly responsive fluorophore comprises a pH-sensitive fluorophore comprising fluorescein or a derivative thereof.

113. The biosensor of claim 110, wherein the directly responsive fluorophore transitions from a monoanion to a dianion at a pH that is less than 7.0 in an aqueous solution.

10

114. The biosensor of claim 75 or 99, wherein the directly responsive fluorophore is attached to the ligand-binding protein, the amino acid, or the polypeptide via a covalent bond.

115. The biosensor of claim 114, wherein the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been
15 formed by a click reaction.

116. The biosensor of claim 114, wherein the directly responsive fluorophore is attached to a cysteine or a lysine of the protein.

20

117. The biosensor of claim 114, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein.

118. The biosensor of claim 114, wherein the indirectly responsive fluorophore is attached
25 to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif.

119. The biosensor of claim 118, wherein the fluorophore attachment motif comprises an amino acid or a polypeptide.

30 120. The biosensor of claim 119, wherein the polypeptide comprises amino acids in the sequence of β Zif (SEQ ID NO: 42).

121. The biosensor of claim 119, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).

5 122. The biosensor of claim 76, wherein the indirectly responsive fluorophore is attached to the ligand-binding protein via a covalent bond.

123. The biosensor of claim 122, wherein the covalent bond comprises a disulfide bond, a thioester bond, a thioether bond, an ester bond, an amide bond, or a bond that has been
10 formed by a click reaction.

125. The biosensor of claim 122, wherein the indirectly responsive fluorophore is attached to a cysteine or a lysine of the protein.

15 126. The biosensor of claim 122, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein.

127. The biosensor of claim 122, wherein the indirectly responsive fluorophore is attached to the N-terminus or the C-terminus of the protein via a fluorophore attachment motif.
20

128. The biosensor of claim 127, wherein the fluorophore attachment motif comprises an amino acid or a polypeptide.

25 129. The biosensor of claim 128, wherein the polypeptide comprises amino acids in the sequence of β Zif (SEQ ID NO: 42).

130. The biosensor of claim 138, wherein the polypeptide comprises a stretch of at least 50, 60, 70, 80, 90, or 100 amino acids in a sequence that is at least about 85%, 90%, 95%, or 99% identical to the amino acid sequence of *E. coli* thioredoxin (ecTRX; SEQ ID NO: 151).
30

131. A method for assaying the level of a ligand in a subject, comprising contacting a biosensor according to claim 75 or 99 with a biological sample from the subject.

132. The method of claim 131, wherein the ligand comprises glucose, galactose, lactose, arabinose, ribose, maltose, lactate, urea, bicarbonate, phosphate, sulfate, chloride, fluoride, iodide, astatide, ununseptide, bromide, calcium, a hydrogen ion, a dipeptide, histidine, glutamine, glutamate, aspartate, or iron.

5

133. The method of claim 131, wherein the subject has or is suspected of having abnormal kidney function, abnormal adrenal gland function, diabetes, hypochloremia, bromism, hypothyroidism, hyperthyroidism, cretinism, depression, fatigue, obesity, a low basal body temperature, a goiter, a fibrocystic breast change, lactic acidosis, septic shock, carbon
10 monoxide poisoning, asthma, a lung disease, respiratory insufficiency, Chronic Obstructive Pulmonary Disease (COPD), regional hypoperfusion, ischemia, severe anemia, cardiac arrest, heart failure, a tissue injury, thrombosis, or a metabolic disorder, diarrhea, shock, ethylene glycol poisoning, methanol poisoning, diabetic ketoacidosis, hypertension, Cushing syndrome, liver failure, cancer, or an infection.

15

134. The method of claim 131, wherein the biological sample comprises sweat, tear fluid, blood, serum, plasma, interstitial fluid, amniotic fluid, sputum, gastric lavage, skin oil, milk, fecal matter, emesis, bile, saliva, urine, mucous, semen, lymph, spinal fluid, synovial fluid, a cell lysate, venom, hemolymph, or a fluid obtained from a plant.

20

135. A method for assaying the level of ligand in an environmental sample, comprising contacting a biosensor according to claim 75 or 99 with the environmental sample.

136. The method of claim 135, wherein the environmental sample is from an
25 environmental site that is suspected of being polluted.

137. The method of claim 135, wherein the environmental sample has been obtained or provided from an environmental substance, fluid, or surface.

30

138. The method of claim 137, wherein

(a) the environmental substance comprises rock, soil, clay, sand, a meteorite, an asteroid, dust, plastic, metal, a mineral, a fossil, a sediment, or wood;

(b) the environmental surface comprises the surface of a satellite, a bike, a rocket, an automobile, a truck, a motorcycle, a yacht, a bus, or a plane, a tank, an armored personnel

carrier, a transport truck, a jeep, a mobile artillery unit, a mobile antiaircraft unit, a minesweeper, a Mine-Resistant Ambush Protected (MRAP) vehicle, a lightweight tactical all-terrain vehicle, a high mobility multipurpose wheeled vehicle, a mobile multiple rocket launch system, an amphibious landing vehicle, a ship, a hovercraft, a submarine, a transport plane, a fighter jet, a helicopter, a rocket, or an Unmanned Aerial Vehicle, a drone, a robot, a building, furniture, or an organism; or

(c) the environmental fluid comprises marine water, well water, drinking well water, water at the bottom of well dug for petroleum extraction or exploration, melted ice water, pond water, aquarium water, pool water, lake water, mud, stream water, river water, brook water, waste water, treated waste water, reservoir water, rain water, or ground water.

139. A method for monitoring the level of a ligand, comprising periodically continuously detecting the level of the ligand, wherein detecting the level of the ligand comprises

(a) providing or obtaining a sample;

(b) contacting the sample with a biosensor for the ligand according to claim 75 or 99, and

(c) detecting a signal produced by the biosensor.

140. The method of claim 139, wherein the sample is provided or obtained from a subject or from a culture of microbial cells.

141. A method for constructing a biosensor, comprising:

(a) providing a ligand-binding protein;

(b) identifying at least one putative allosteric, endosteric, or peristeric site of the ligand-binding based a structure of the ligand-binding protein;

(c) mutating the ligand-binding protein to substitute an amino acid at the at least one putative allosteric, endosteric, or peristeric site of the second protein with a cysteine;

(f) conjugating a donor fluorophore or an acceptor fluorophore to the cysteine to produce single labeled biosensor;

(g) detecting whether there is a spectral shift or change in emission intensity of the single labeled biosensor upon ligand binding when the donor fluorophore or the acceptor fluorophore is fully excited; and

(h) if a spectral shift or change in emission intensity is detected in (g), attaching a donor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine, and attaching an acceptor fluorophore to the second protein if an acceptor fluorophore is attached to the cysteine.

5

142. The method of claim 141, wherein the ligand-binding protein has been identified by

- (i) selecting a first protein having a known amino acid sequence (seed sequence), wherein the first protein is a ligand-binding protein;
- (ii) identifying a second protein having an amino acid sequence (hit sequence) with at least 15% sequence identity to the seed sequence;
- (iii) aligning the seed amino acid sequence and the hit sequence, and comparing the hit sequence with the seed sequence at positions of the seed sequence that correspond to at least 5 primary complementary surface (PCS) amino acids, wherein each of the at least 5 PCS amino acids has a hydrogen bond interaction or a van der Waals interaction with ligand when ligand is bound to the first protein; and
- (iv) identifying the second protein to be a ligand-binding protein if the hit sequence comprises at least 5 amino acids that are consistent with the PCS.

10

15

20

143. The method of claim 141, wherein the spectral shift comprises a monochromatic fluorescence intensity change or a dichromatic spectral shift.

144. A method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising

25

(a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and

30

(b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of the first reporter group.

145. The method of claim 144, wherein the second reporter group is within about 100 angstroms (Å) of the first reporter group regardless of whether ligand is bound to the biosensor.

5

146. The method of claim 144, wherein when the ligand is bound to the biosensor, the average distance between the first reporter group and the second reporter group changes by less than about 1 Å compared to when ligand is not bound to the ligand-binding protein.

10 147. A method of converting a biosensor that shows a monochromatic response upon ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising

(a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding fluorescent protein; and

15 (b) attaching an acceptor fluorophore or a donor fluorophore to the biosensor, wherein (i) the acceptor fluorophore has an excitation spectrum that overlaps with the emission spectrum of the fluorescent protein; or (ii) the donor fluorophore has an emission spectrum that overlaps with the excitation spectrum of the fluorescent protein.

20 148. A method of increasing a dichromatic response of a biosensor to ligand binding, the method comprising

(a) selecting a biosensor that exhibits a dichromatic response upon ligand binding, wherein the biosensor comprises a ligand-binding protein and a first reporter group; and

25 (b) attaching a second reporter group to the biosensor, wherein the second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of the first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of the first reporter group.

149. A method of converting a biosensor that shows a monochromatic response upon
30 ligand binding into a biosensor with a dichromatic response upon ligand binding, the method comprising

(a) selecting a biosensor that exhibits a monochromatic response upon ligand binding, wherein said biosensor comprises an amino acid or a polypeptide and a first reporter group, wherein the first reporter group comprises a chemoresponsive fluorophore; and

5 (b) attaching a second reporter group to said biosensor, wherein said second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of said first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of said first reporter group.

150. A method of increasing a dichromatic response of a biosensor to ligand binding, the
10 method comprising

(a) selecting a biosensor that exhibits a dichromatic response upon ligand binding, wherein said biosensor comprises an amino acid or a polypeptide and a first reporter group, wherein the first reporter group comprises a chemoresponsive fluorophore; and

15 (b) attaching a second reporter group to said biosensor, wherein said second reporter group has (i) an excitation spectrum that overlaps with the emission spectrum of said first reporter group; or (ii) an emission spectrum that overlaps with the excitation spectrum of said first reporter group.

151. A method for monitoring the level of a ligand, comprising periodically detecting the
20 level of the ligand, wherein detecting the level of the ligand comprises

(a) providing or obtaining a sample;

(b) contacting the sample with a biosensor for glucose according to claim 75 or 99 under conditions such that the ligand-binding protein of the biosensor binds to the ligand, and

25 (c) detecting a signal produced by the biosensor.

152. A method for monitoring the level of a ligand, comprising continuously detecting the
level of the ligand, wherein detecting the level of the ligand comprises

(a) providing or obtaining a sample;

30 (b) contacting the sample with a biosensor for the ligand according to claim 75 or 99 under conditions such that the ligand-binding protein of the biosensor binds to the ligand, and

(c) detecting a ratiometric signal produced by the biosensor.

153. A method for monitoring the level of ligand in a subject, comprising
(a) administering a biosensor according to claim 75 or 99 or a device comprising a
biosensor according to claim 75 or 99 to the subject, wherein after administration
5 the biosensor is in contact with a bodily fluid or surface of the subject, and
(b) detecting (i) a signal produced by a reporter group of the biosensor continuously
or repeatedly at intervals less than about 30 minutes apart, and/or (ii) whether a
signal is produced by a reporter group of the biosensor continuously or repeatedly
at intervals less than about 30 minutes apart.

10

154. A device comprising a biosensor according to claim 75 or 99.

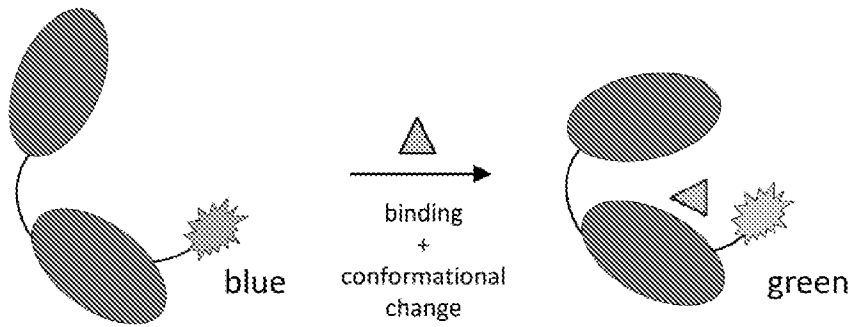
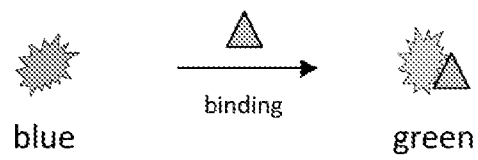
155. The device of claim 154, wherein the biosensor is attached to a surface or matrix of
the device, wherein the surface or matrix comprises a polymer.

15

156. The device of claim 154, wherein the biosensor is attached to a surface or matrix of
the device, wherein the surface or matrix comprises cellulose.

20

157. The device of claim 154, which comprises an optode, a dermal patch, a contact lense,
a bead that is suitable for subcutaneous administration, or a tattoo composition.

Indirect**FIG. 1A****Direct****FIG. 1B**

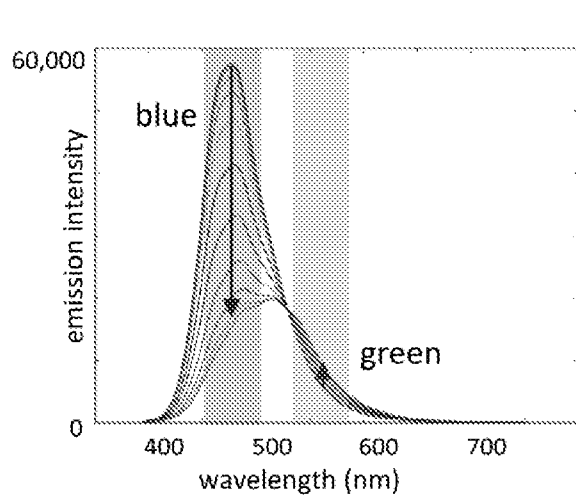
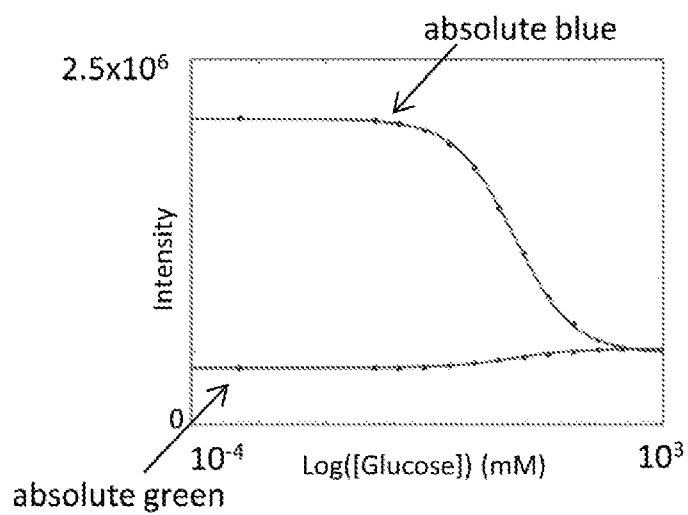
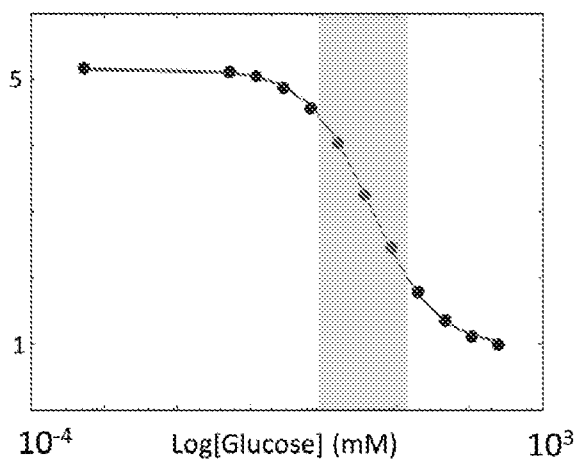
**FIG. 2A****FIG. 2B****FIG. 2C**

FIG. 3A

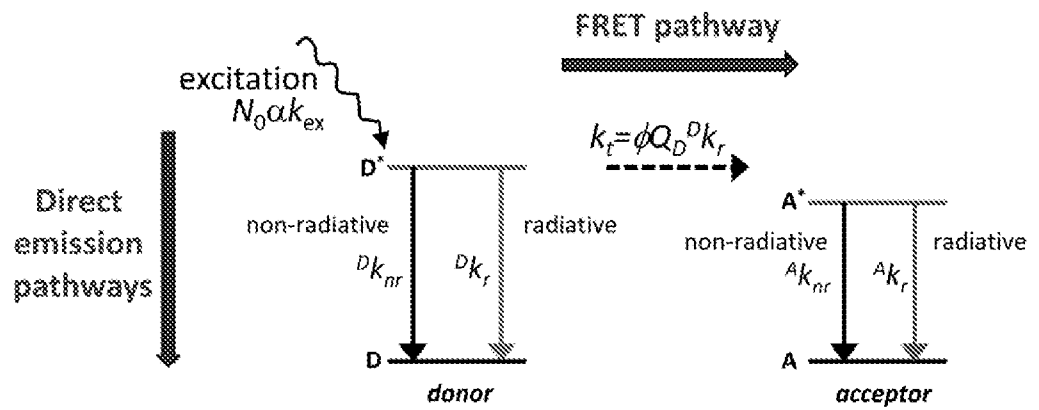


FIG. 3B

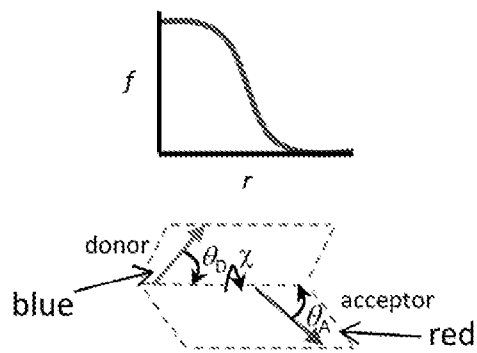


FIG. 3C

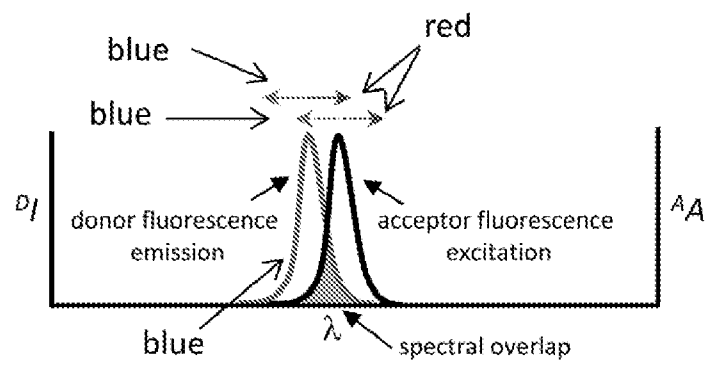


FIG. 3D

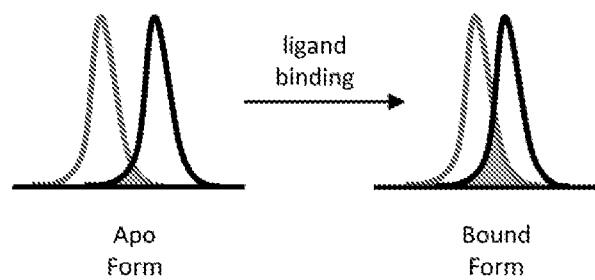
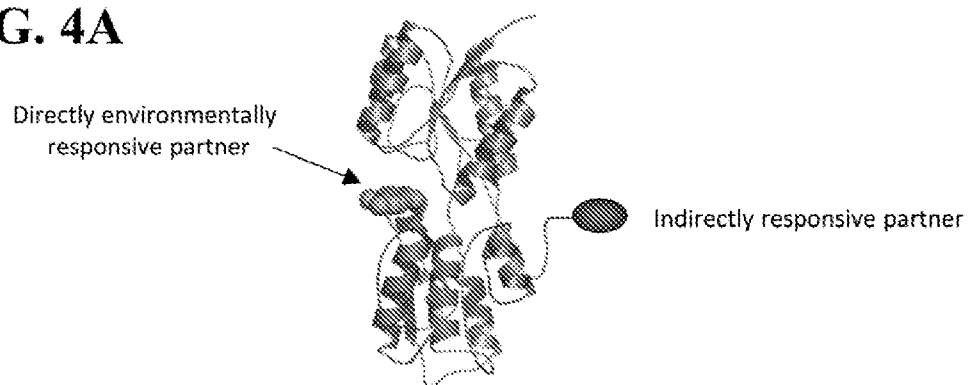
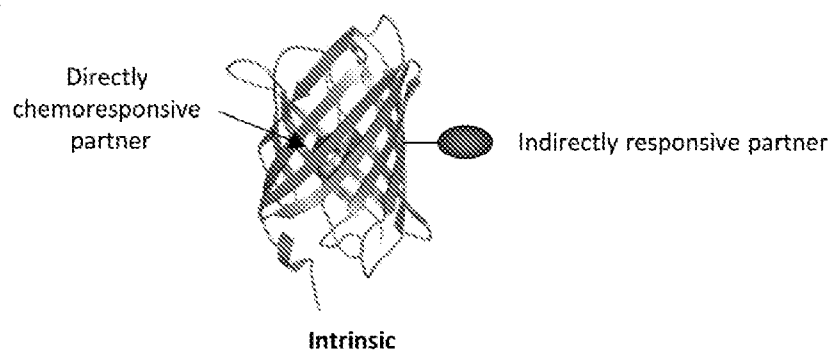
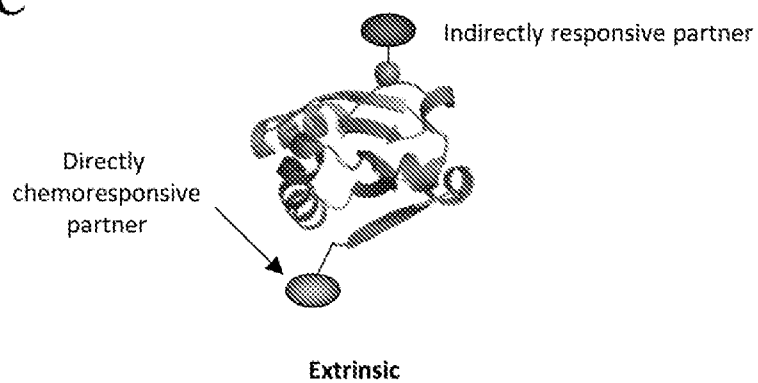
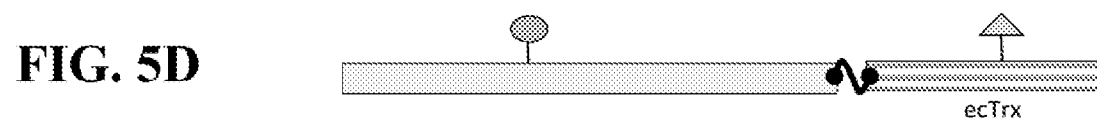
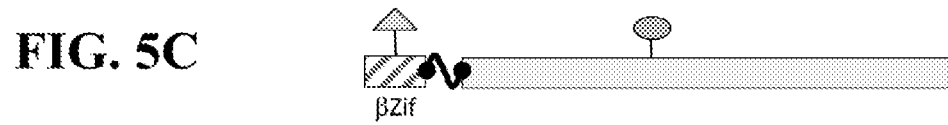
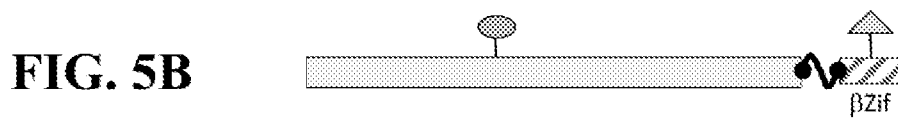
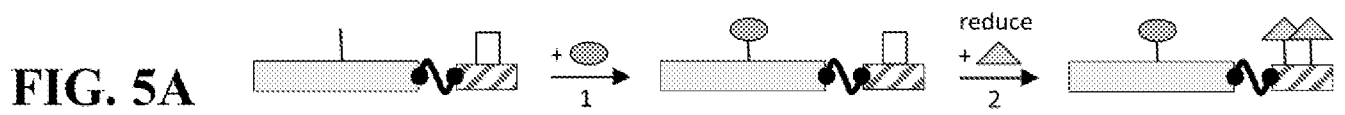
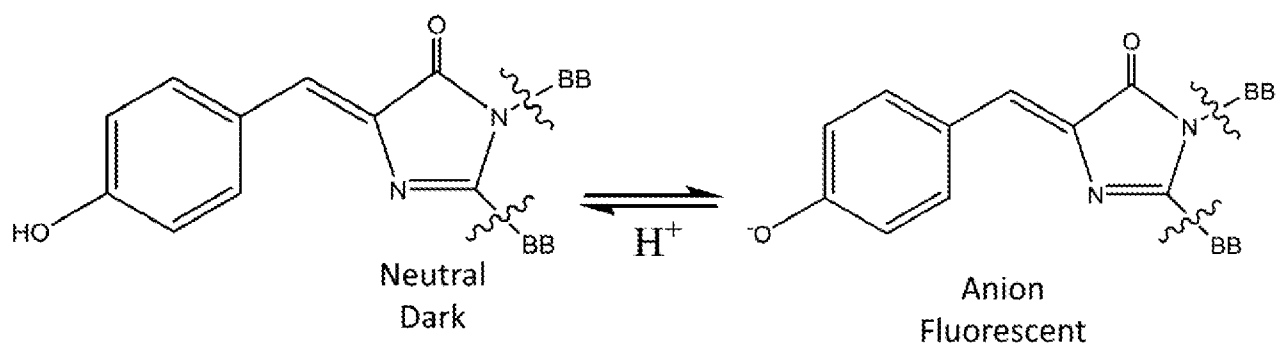
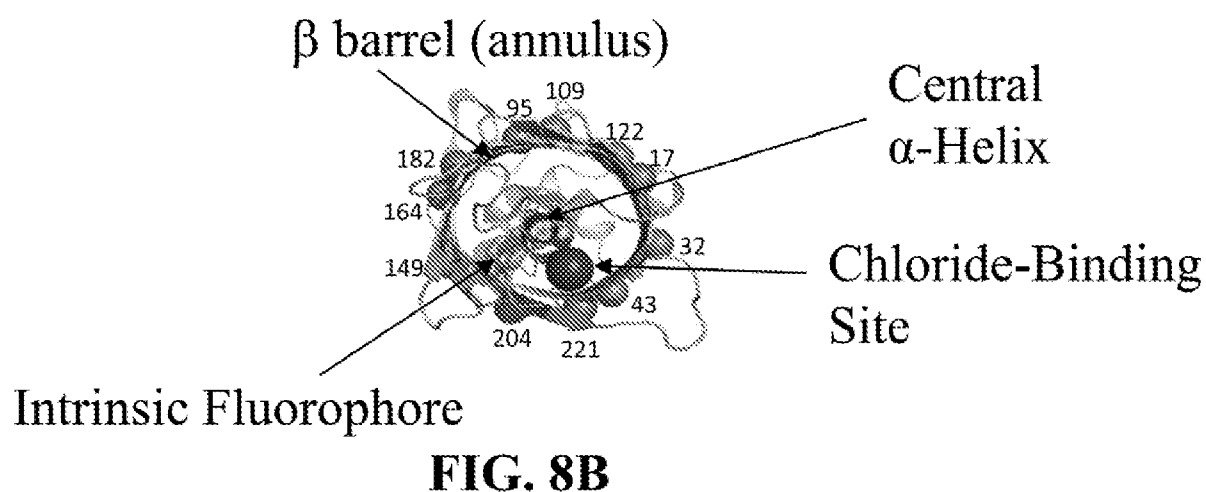
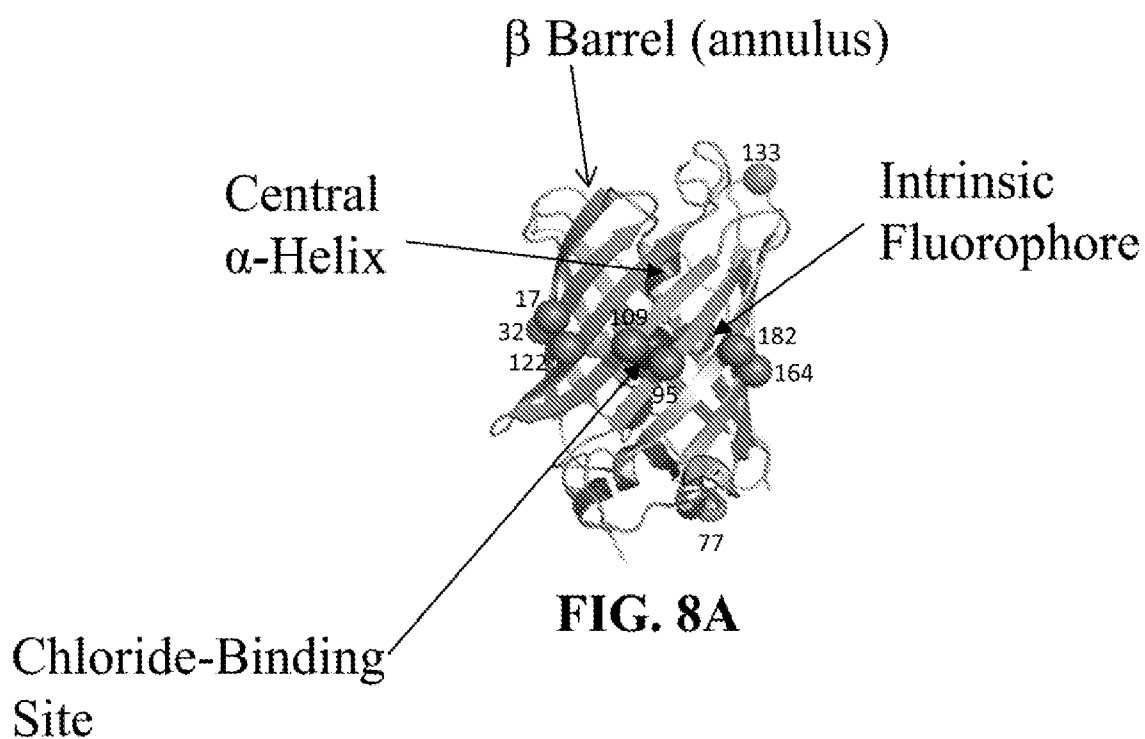


FIG. 4A**FIG. 4B****FIG. 4C**



**FIG. 6**

10	20	30	40	50	60	70	80
MVSKGEELFTGVVPI	LVELDGDVNGHKF	SVSGEGDATY	GKLT	IKTEICTG	KLEVWPTL	VTTF	FGVQCFARYPDHMRQ
90	100	110	120	130	140	150	160
HDFEKSAMPEGYVQ	ERTIEFKDDGN	YKTRA	EVKFE	DTLVNRI	ELKGI	DEFKEDGN	ILGHKLEYNYN
170	180	190	200	210	220	230	240
IKVNEFKIRHNIED	GSVOLADHYQ	ONTPEIGD	GFVLLP	DNHYLS	YQSALS	KDPNEKR	DHVMVILFEFVTAAGITLGMDELYKGS
250							
SNDYKDDDDKGGSHHHHHH							



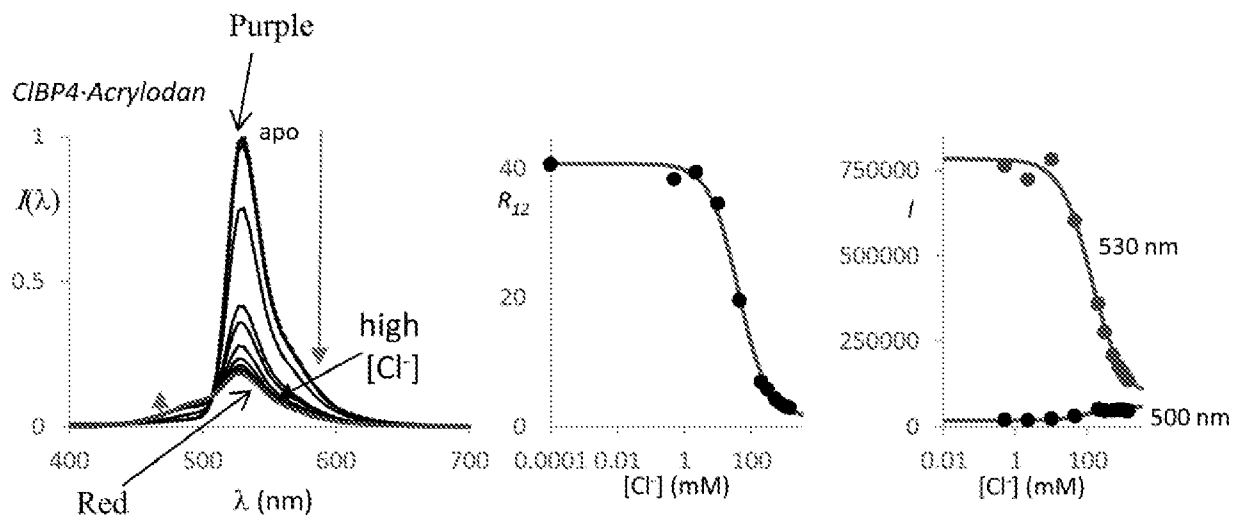


FIG. 9A

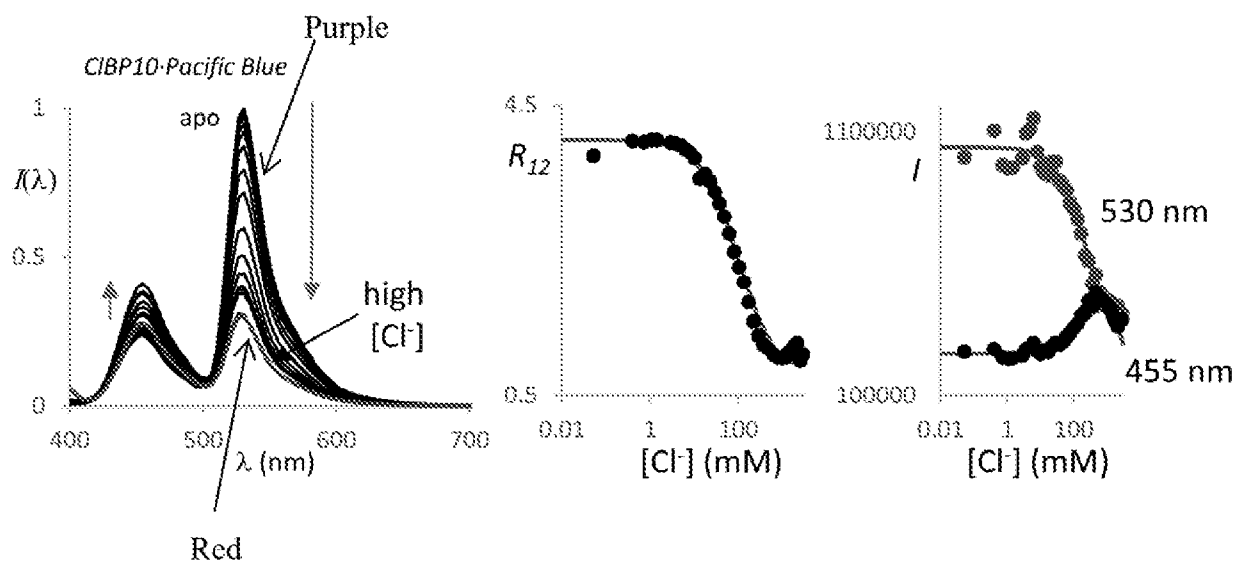
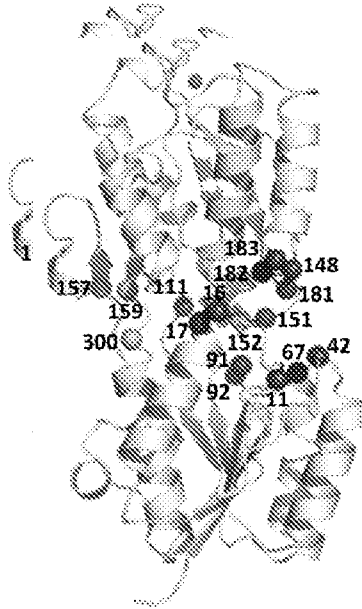
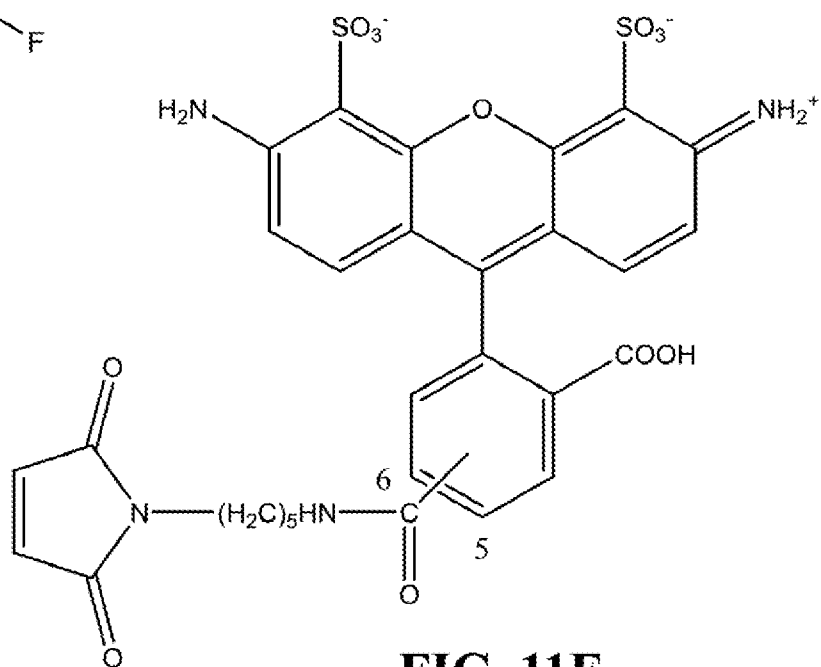
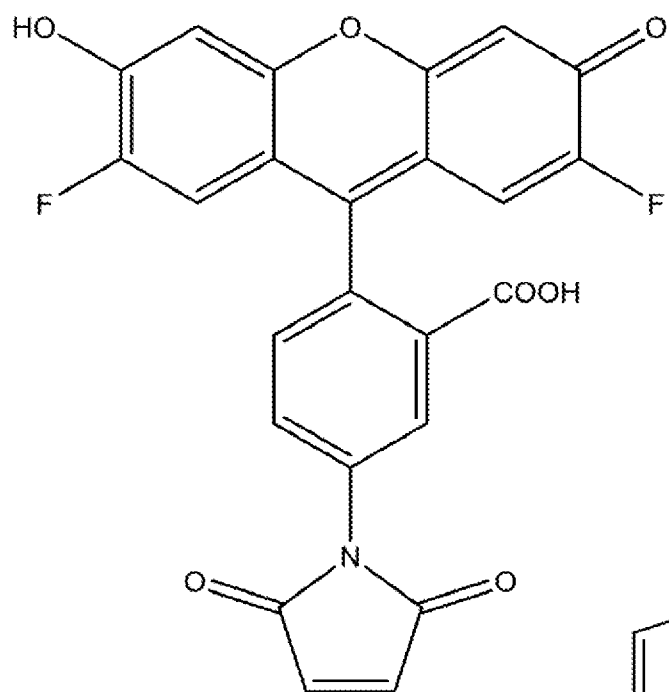
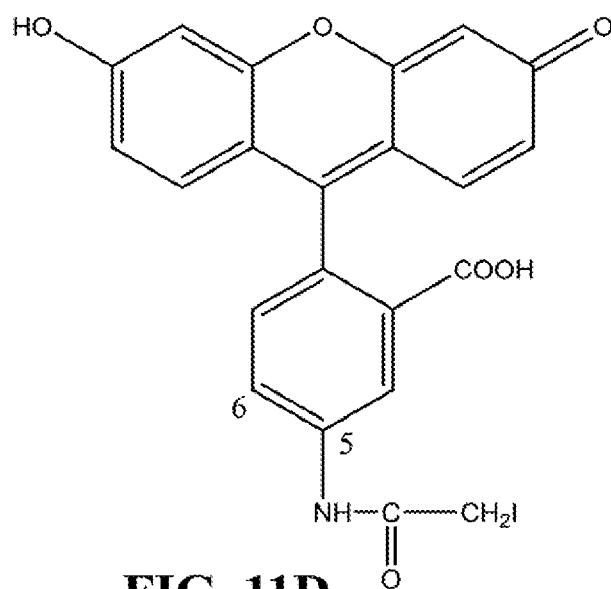
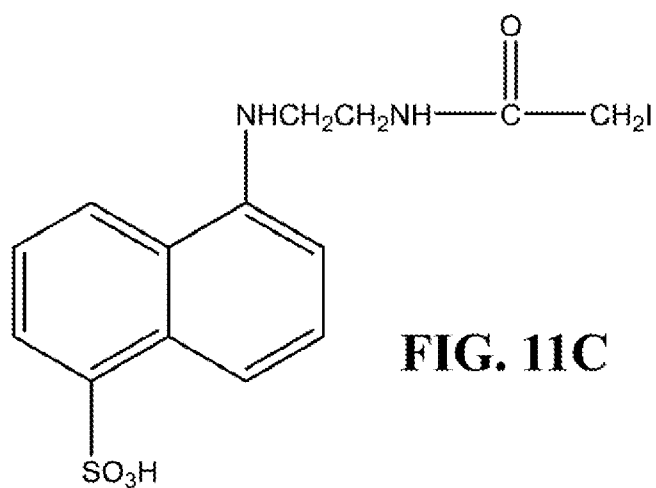
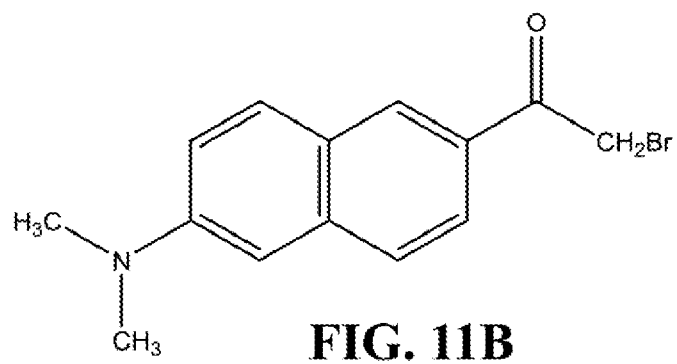
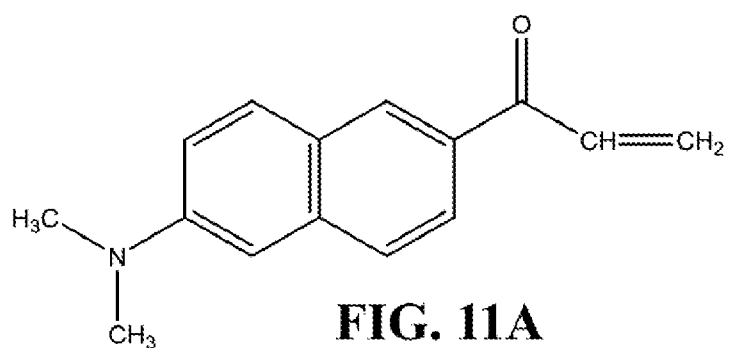


FIG. 9B

**FIG. 10**



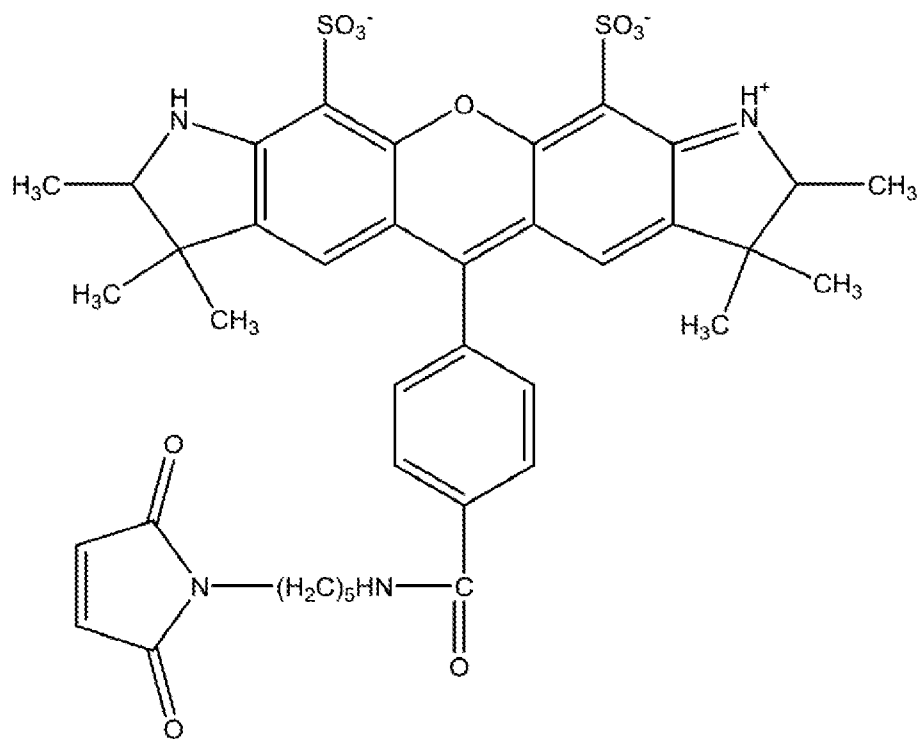


FIG. 11G

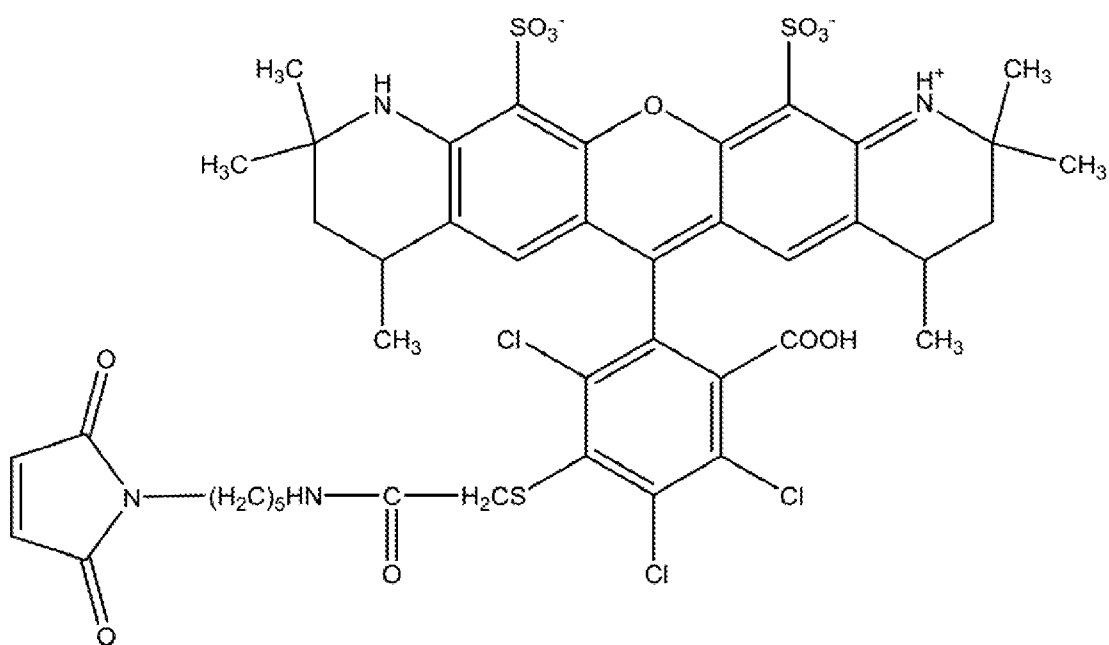
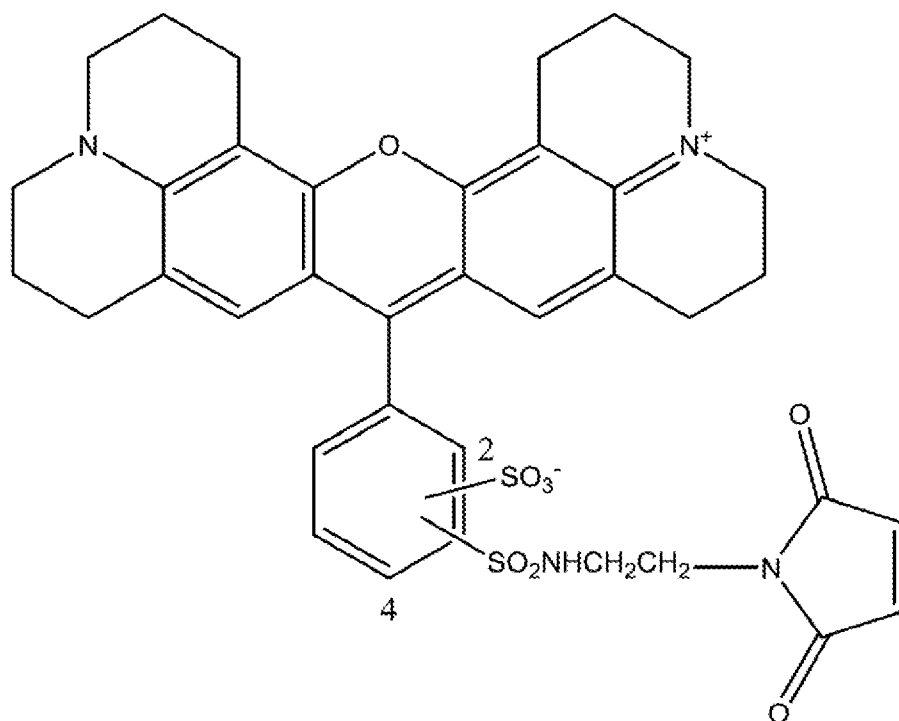
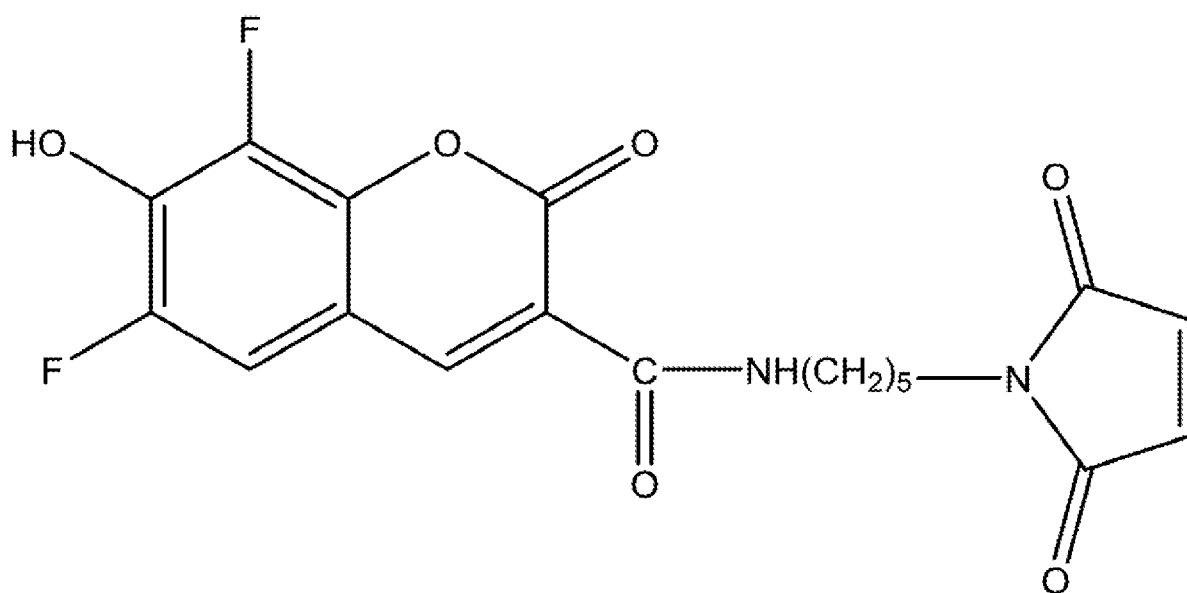
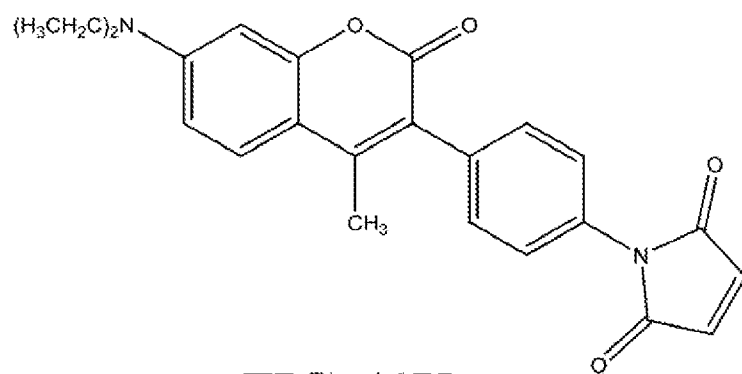
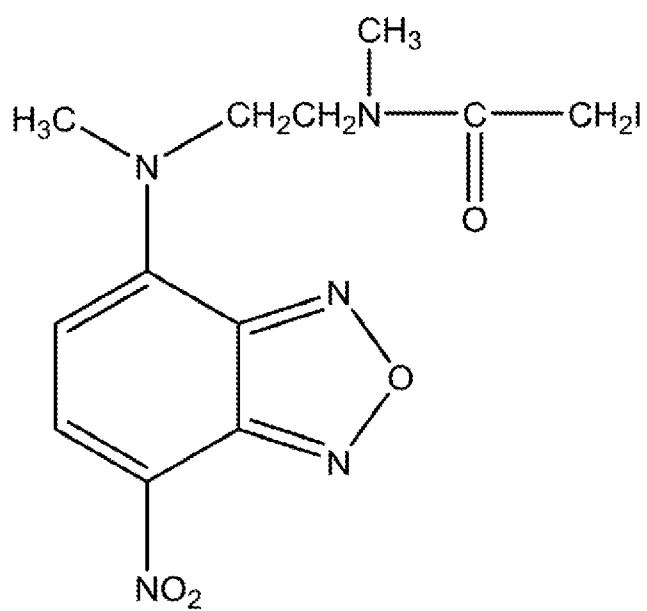
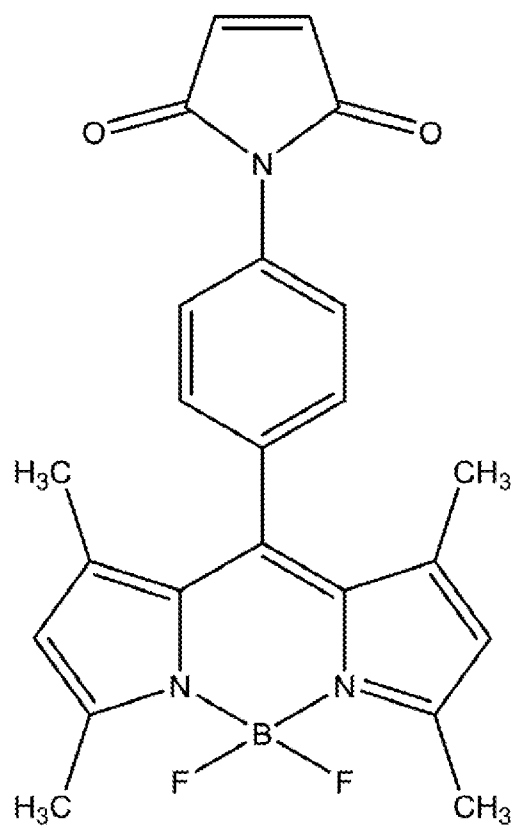
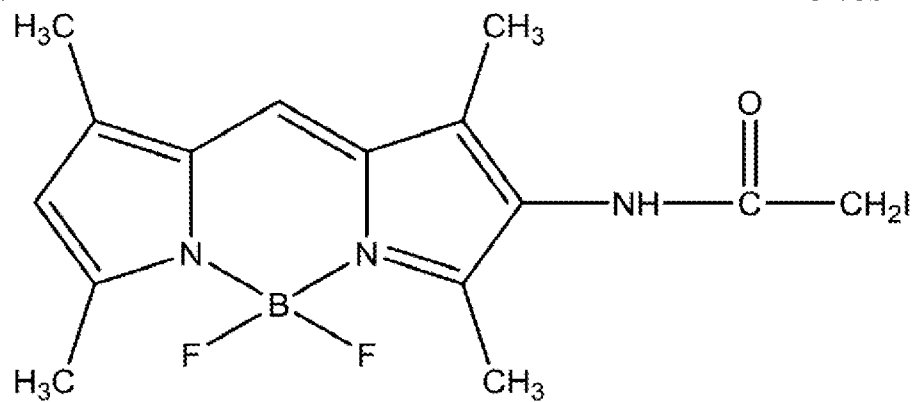
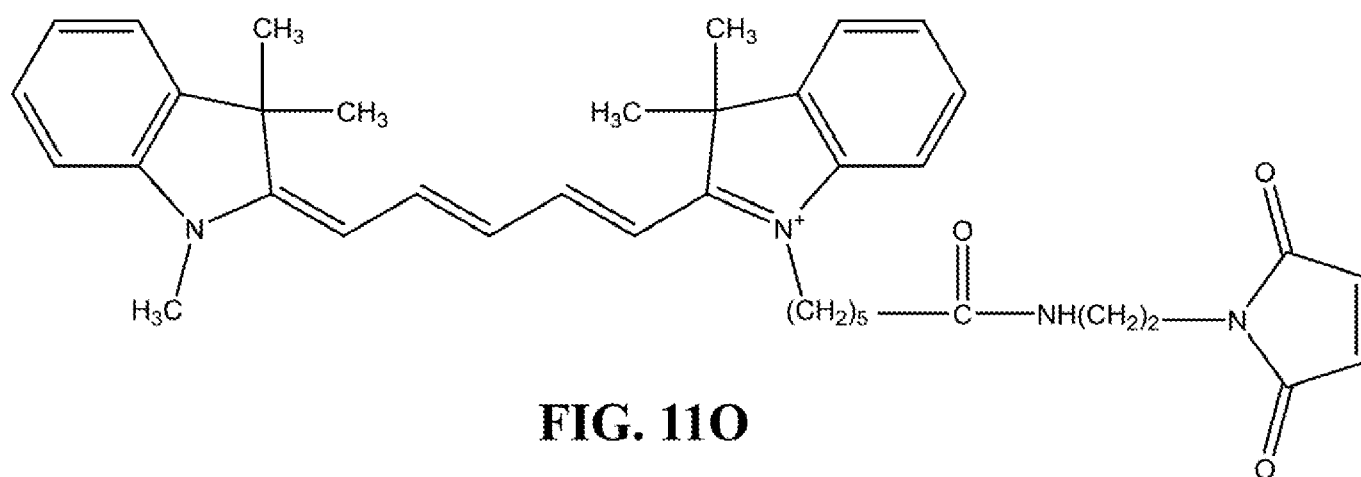
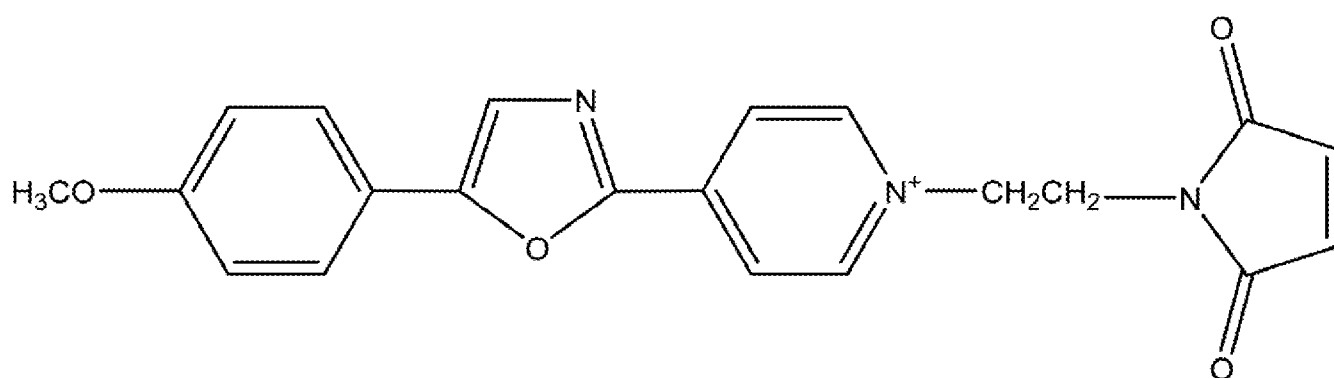


FIG. 11H

**FIG. 11I****FIG. 11J**

**FIG. 11K****FIG. 11L****FIG. 11M**

**FIG. 11N****FIG. 11O****FIG. 11P**

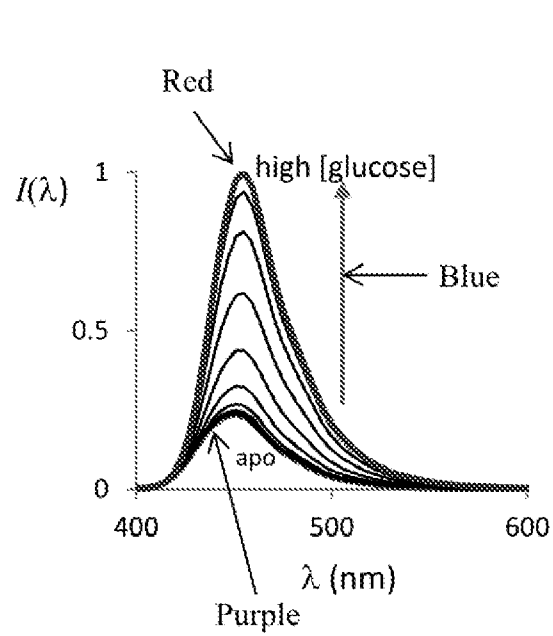
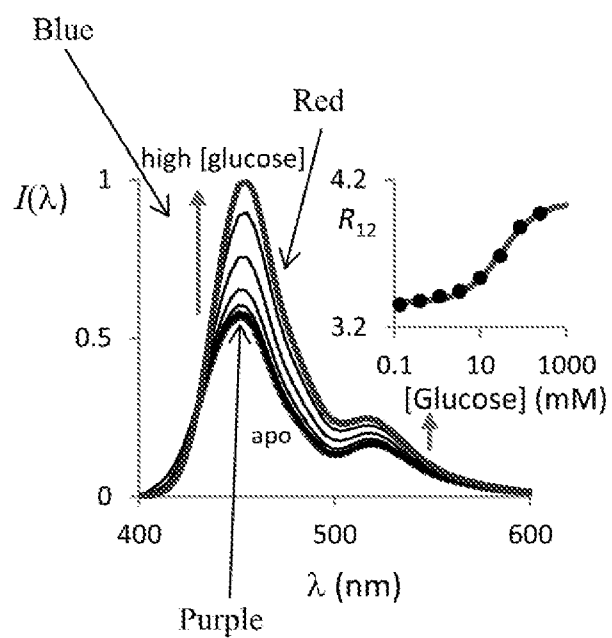
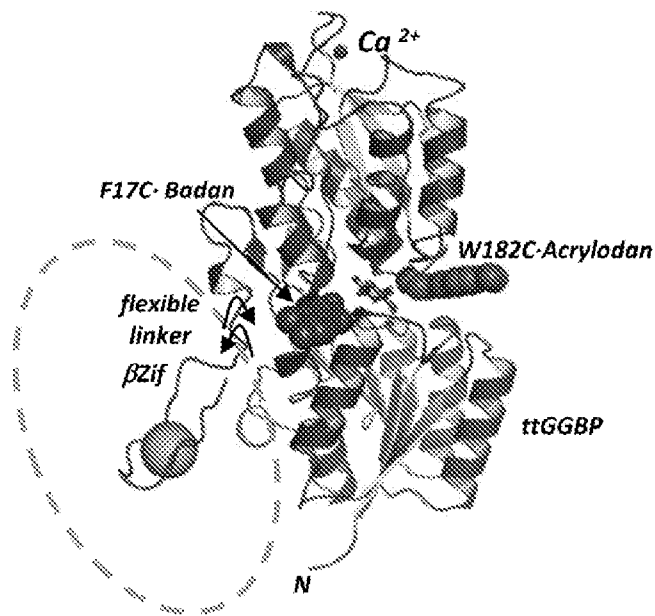
**FIG. 12A****FIG. 12B**

**FIG. 13A****FIG. 13B**

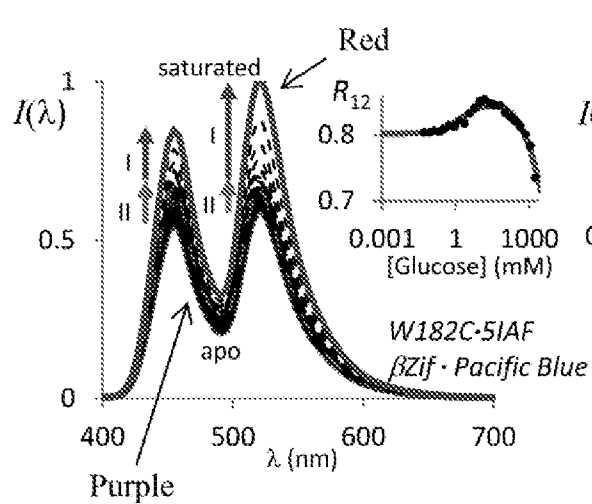


FIG. 14A

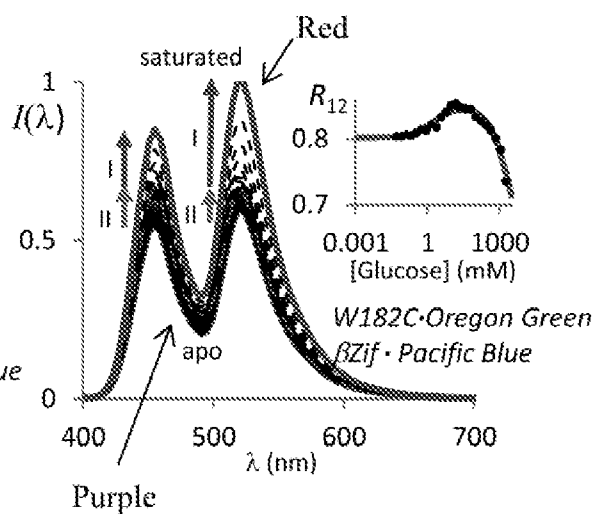


FIG. 14C

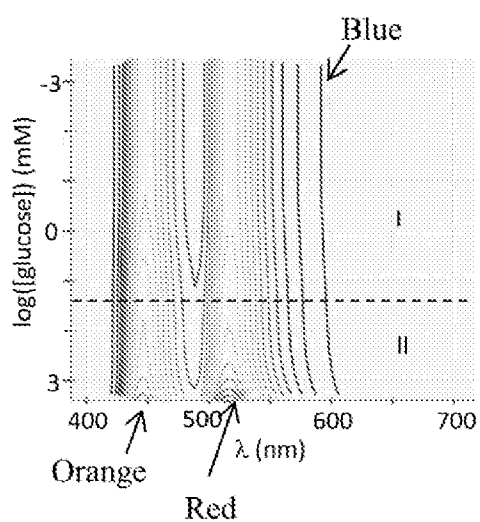


FIG. 14B

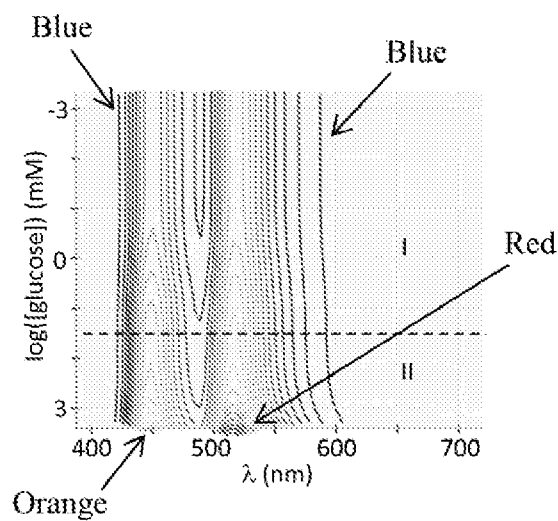
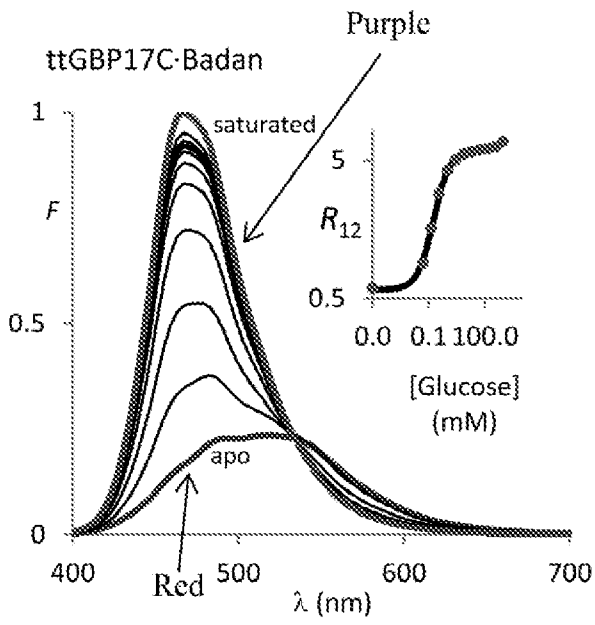
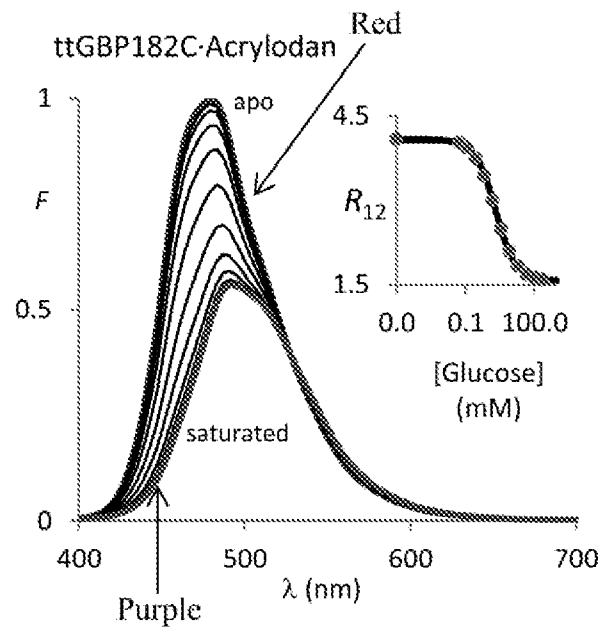


FIG. 14D

**FIG. 15A****FIG. 15B**

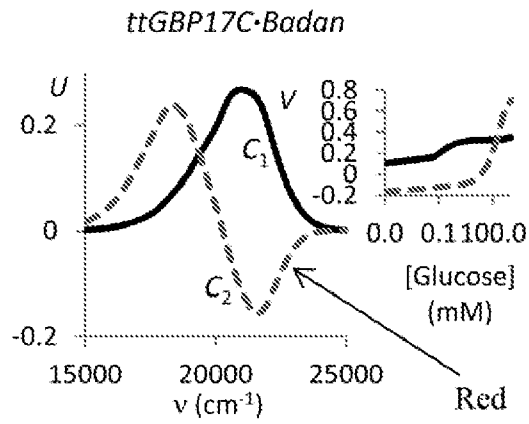


FIG. 16A

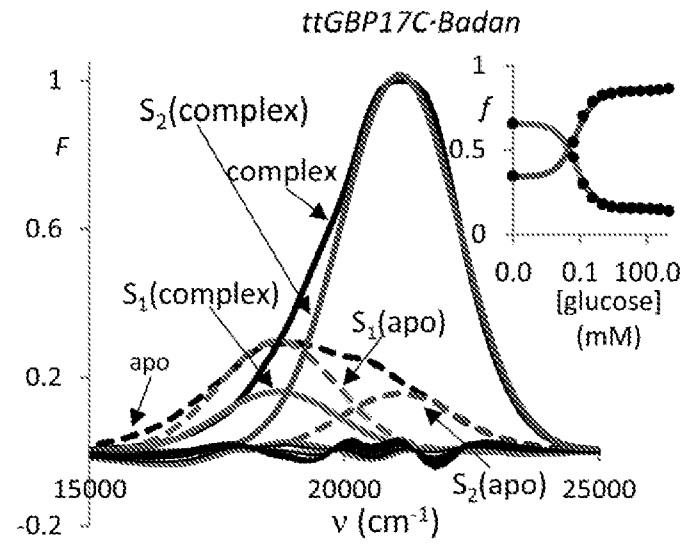


FIG. 16B

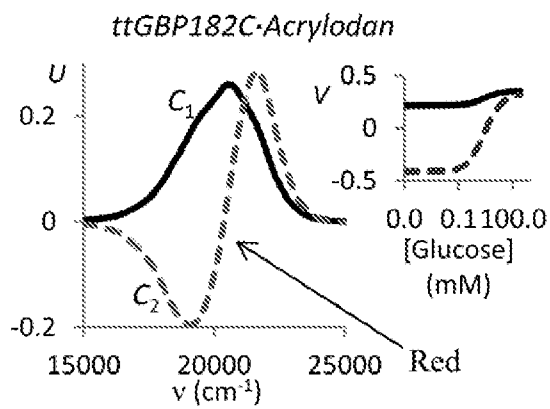


FIG. 16C

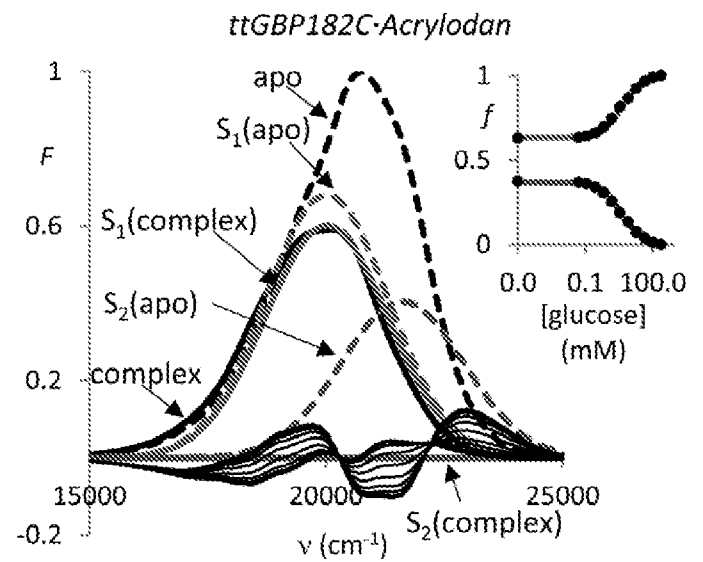


FIG. 16D

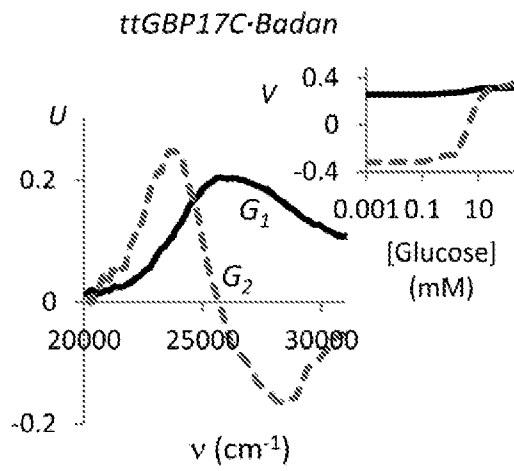


FIG. 17A

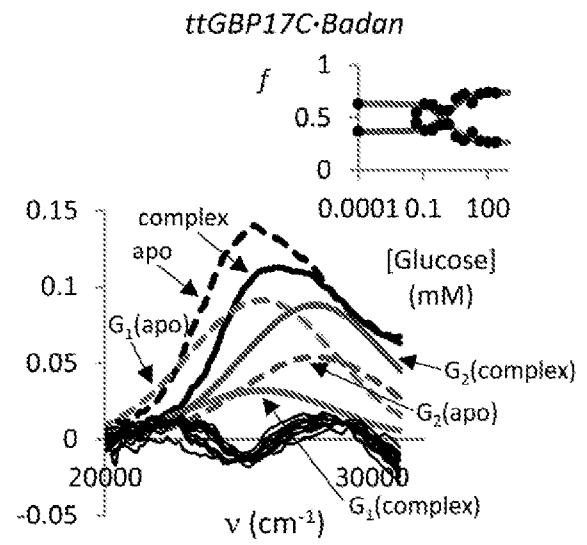


FIG. 17B

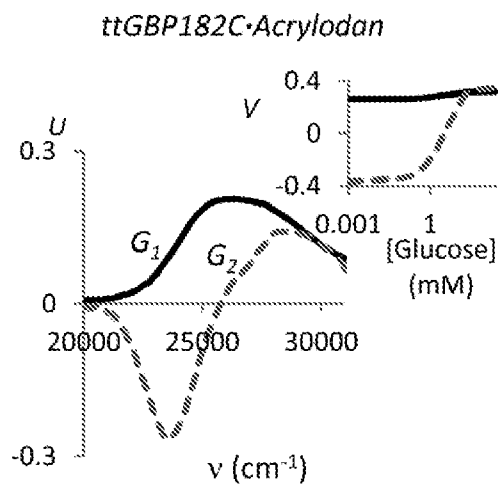


FIG. 17C

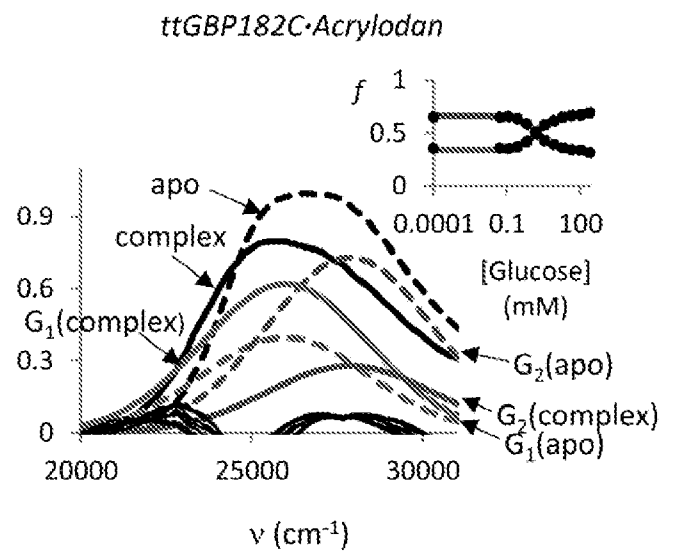


FIG. 17D

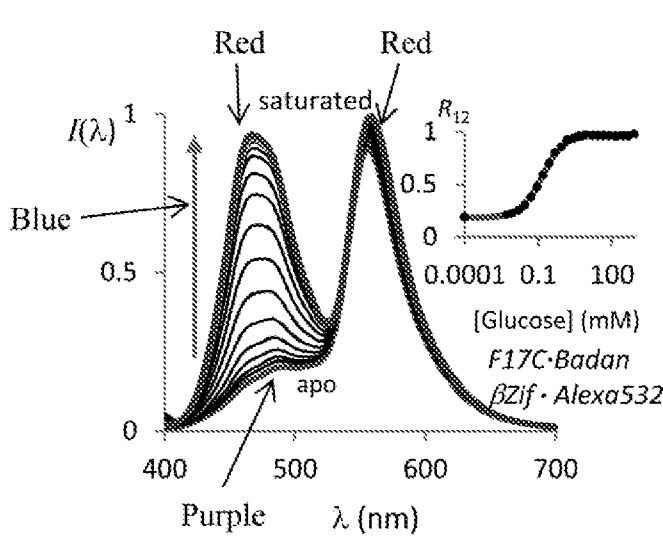


FIG. 18A

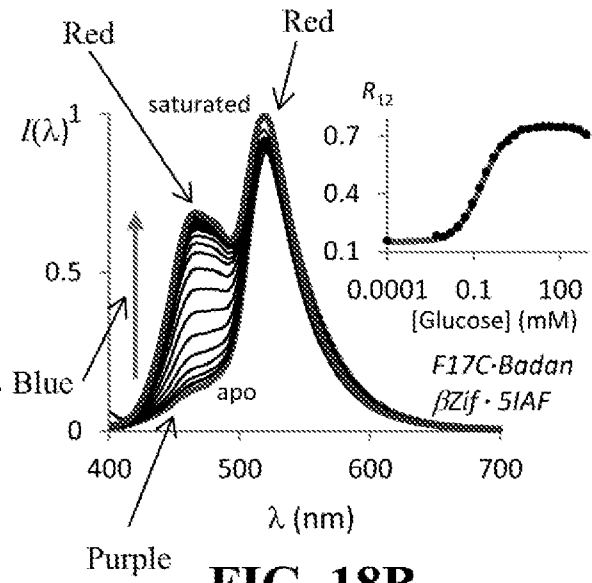


FIG. 18B

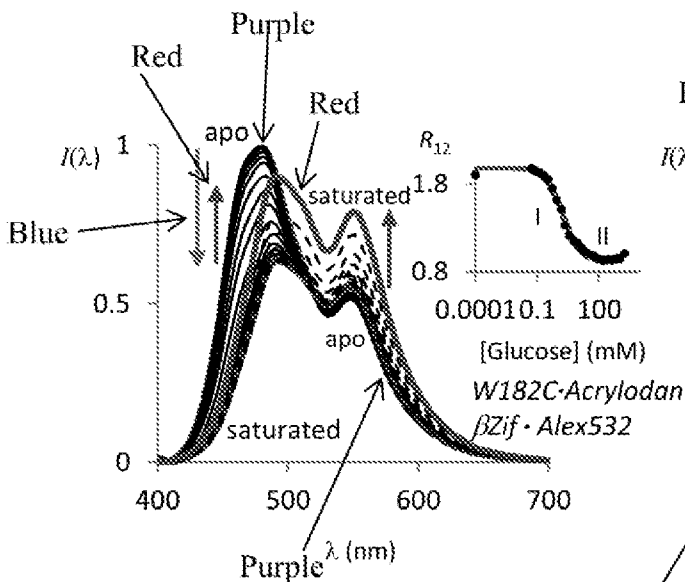


FIG. 18C

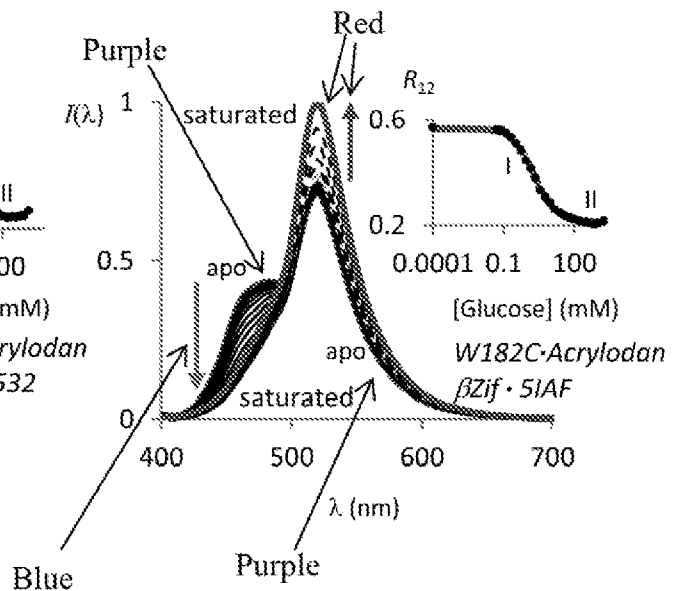
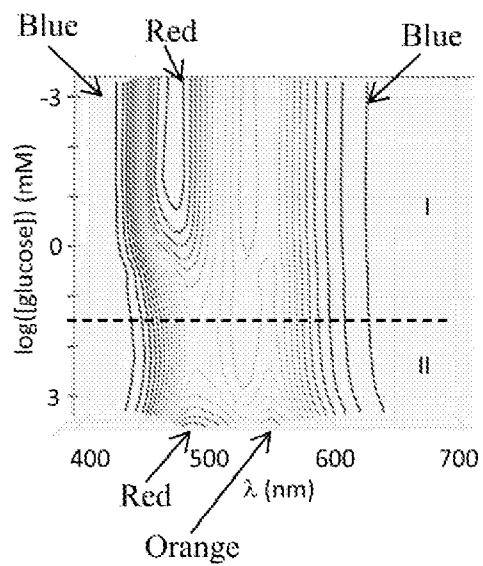
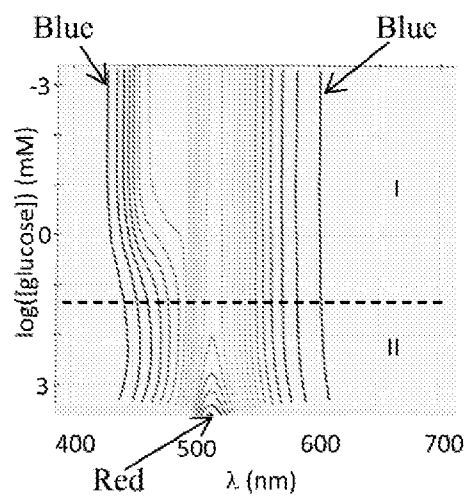


FIG. 18D

**FIG. 18E****FIG. 18F**

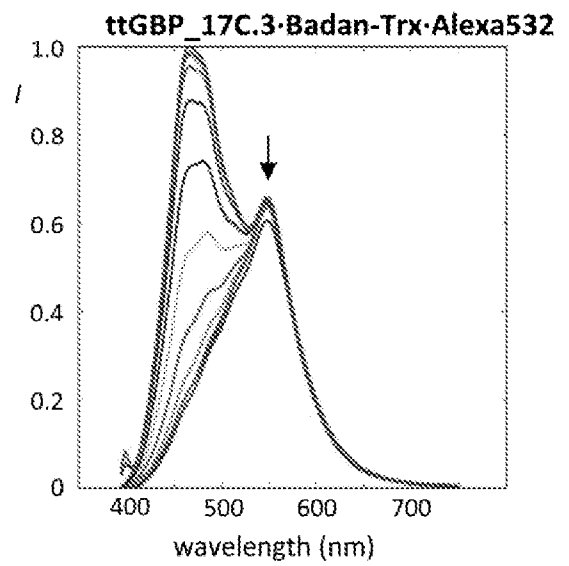
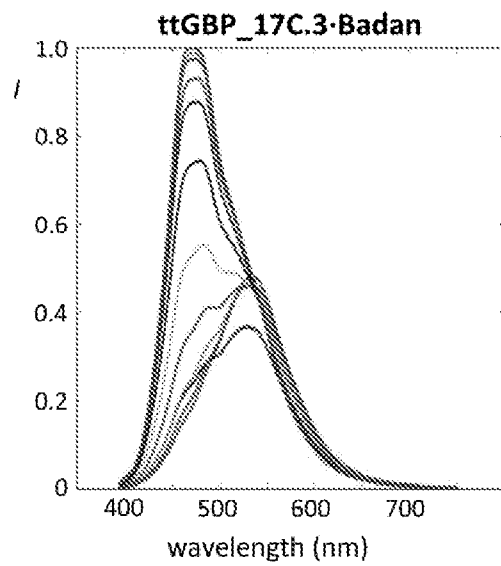


FIG. 20A

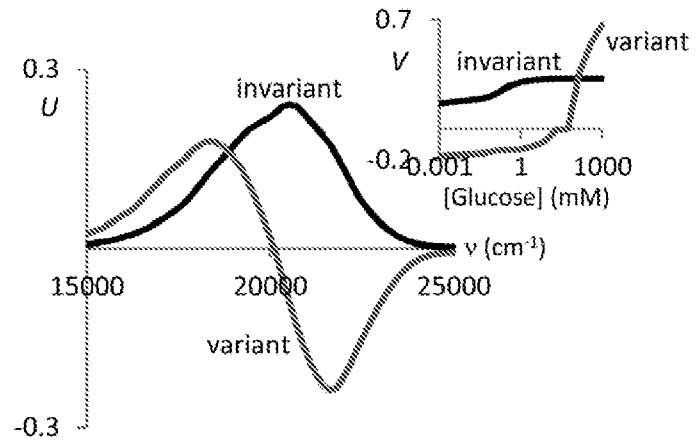


FIG. 20B

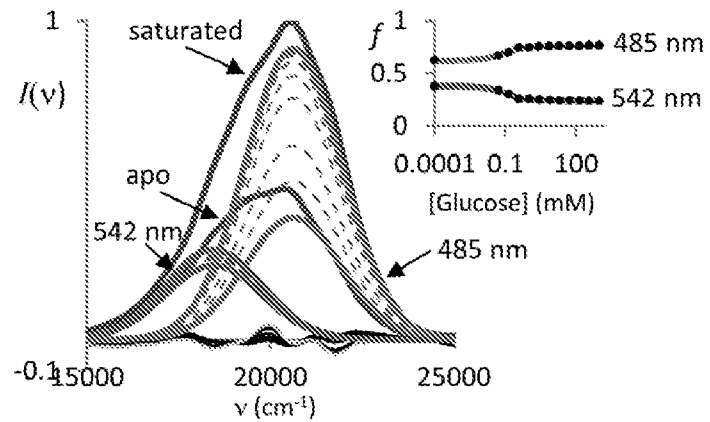


FIG. 20C

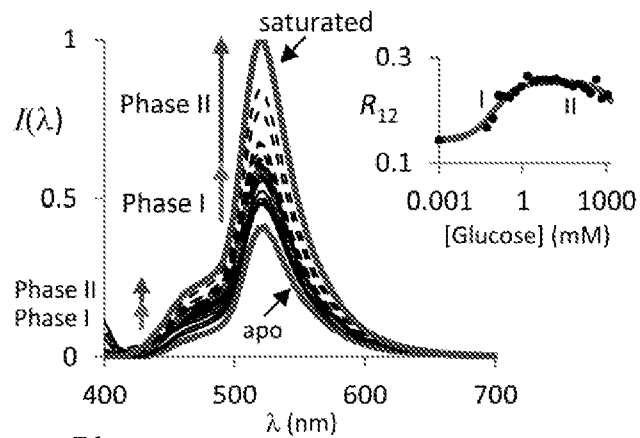
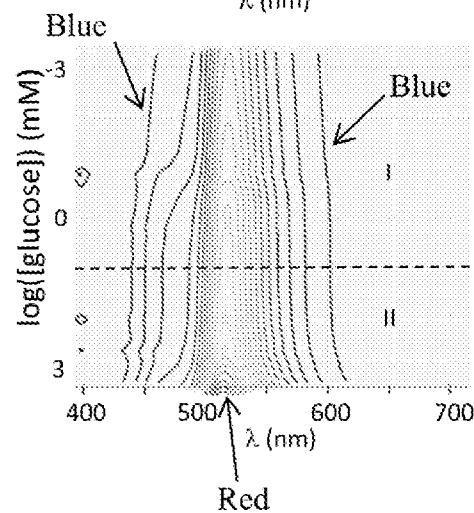
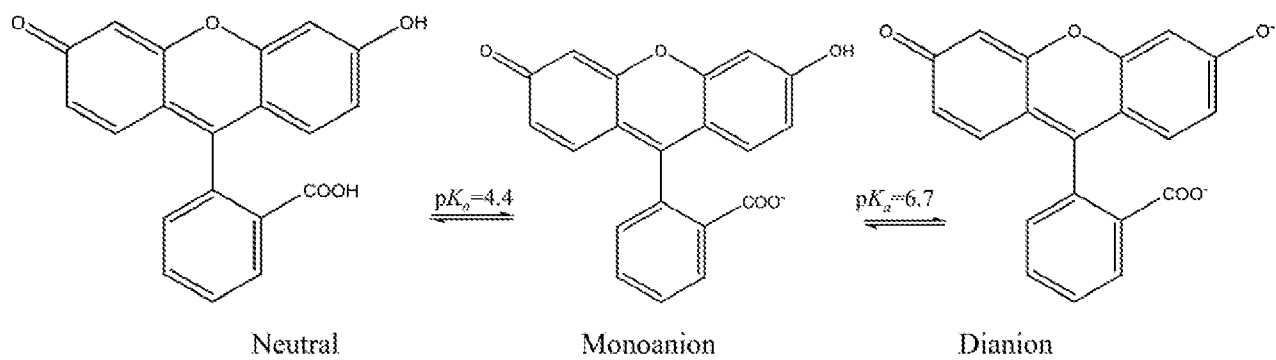


FIG. 20D



**FIG. 21**

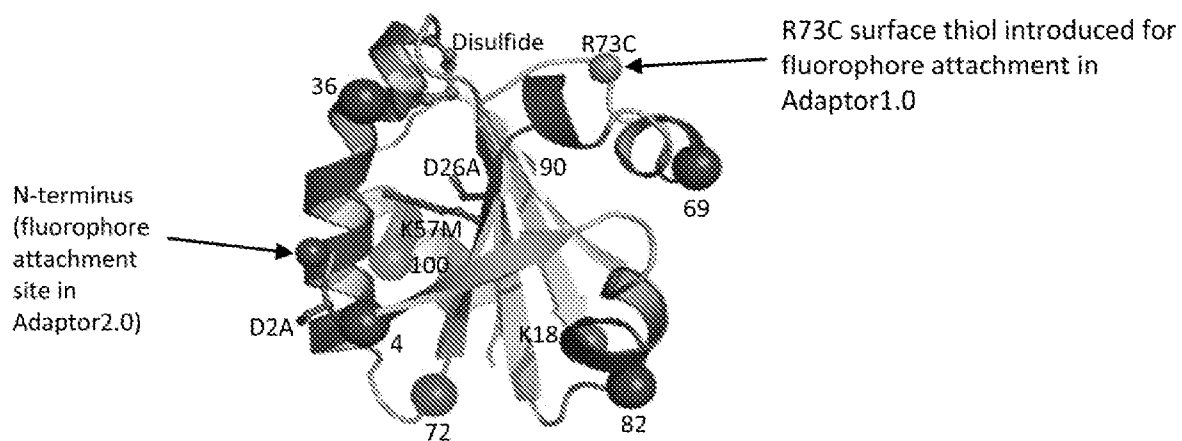
**FIG. 22**

FIG. 23

10	20	30	40	50	60	70	80	90	100	
MSDKTIHLTDDSEDTTVLEKADGAILNDYFWAEWQSPQKMIAPILDEIADEYOGKLTVAFLNIDNPNFGTAPRYGIRGIFTLLFKNQZVAATKVGALSKGJLEFIDANLA										wt
A		A			M					GGSHHHHH Adaptor0
A		A			M					GGSHHHHH Adaptor1.0
AR	R		R	R	M	R	R	R	R	GGSHHHHH Adaptor2.0a
AQ	Q		R	R	M	R	R	R	R	GGSHHHHH Adaptor2.0a

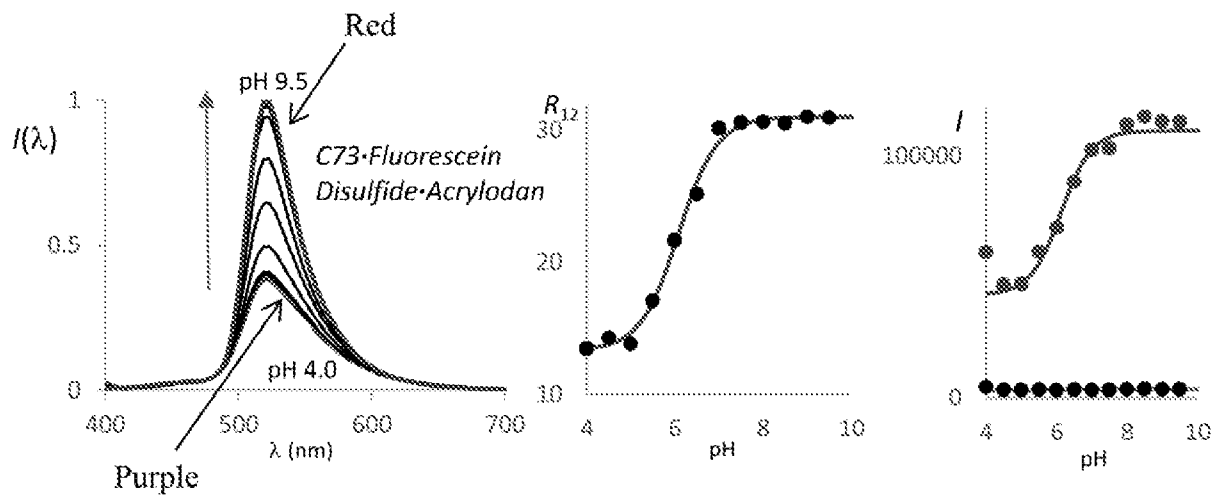


FIG. 24A

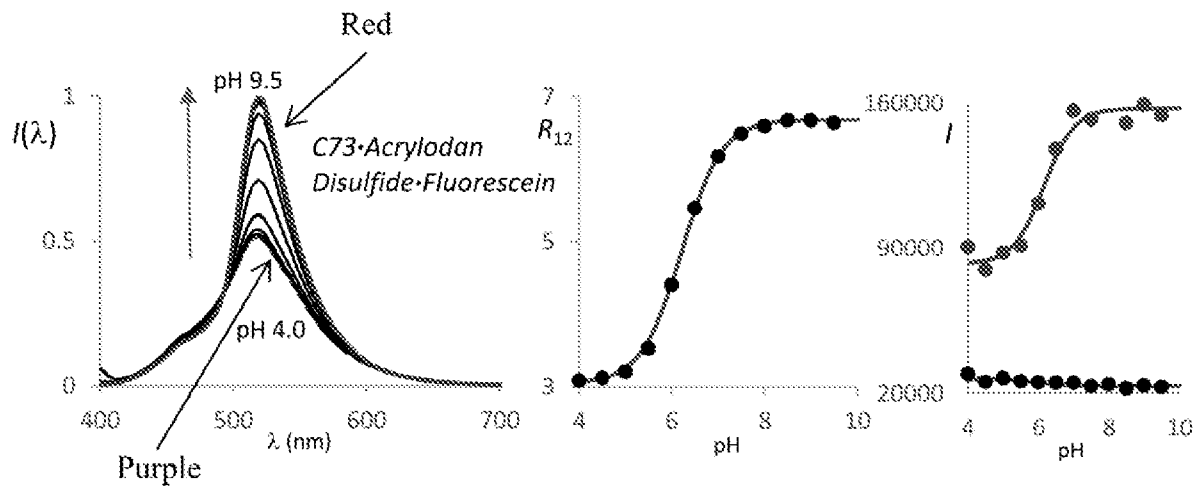


FIG. 24B

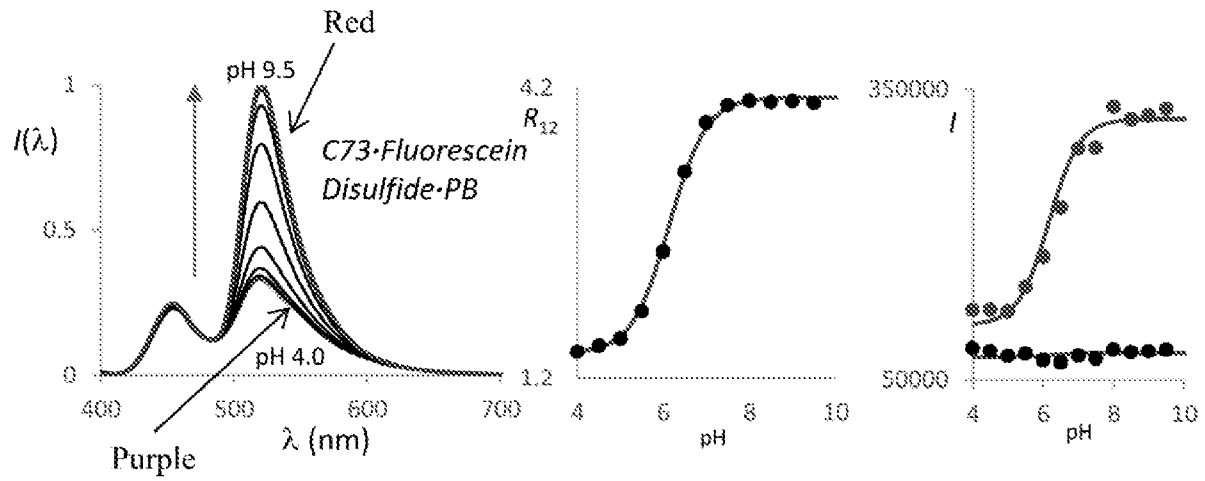


FIG. 24C

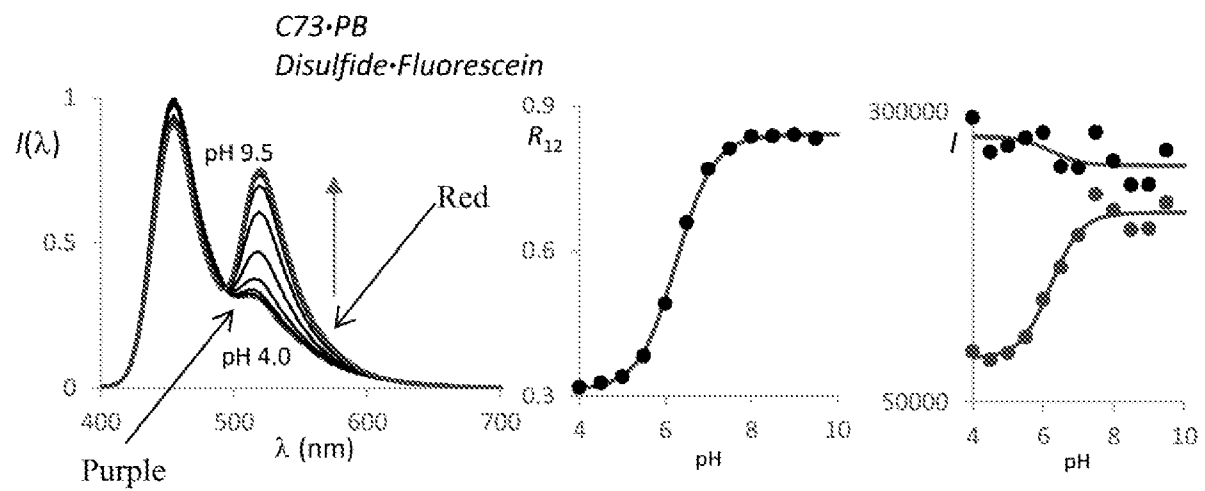


FIG. 24D

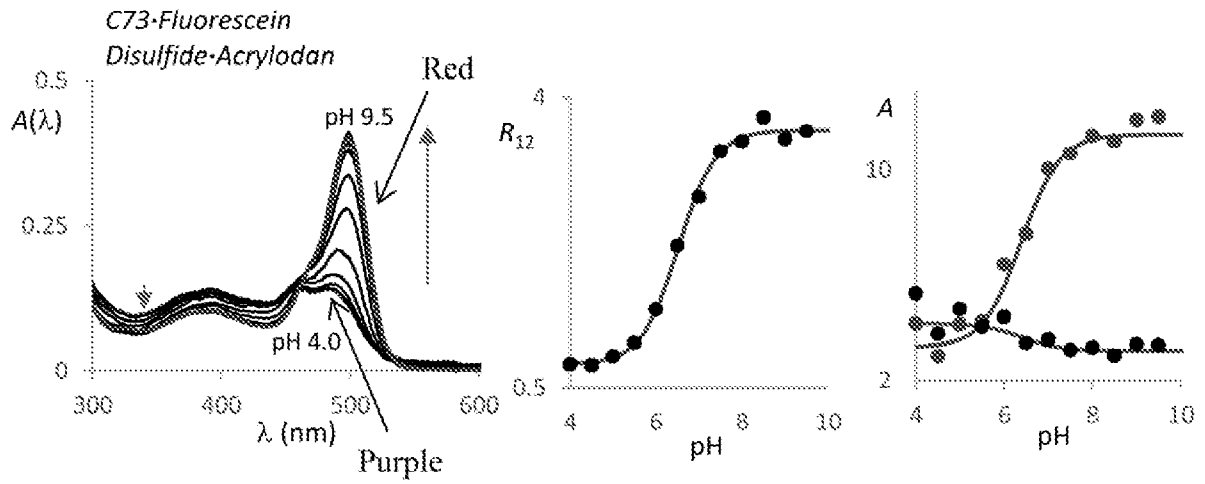


FIG. 25A

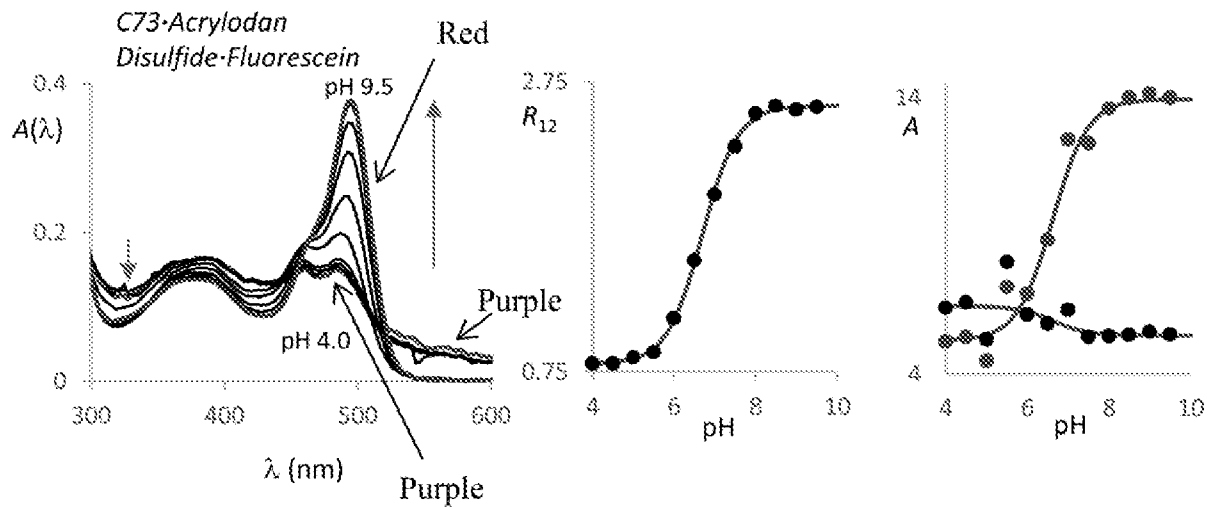


FIG. 25B

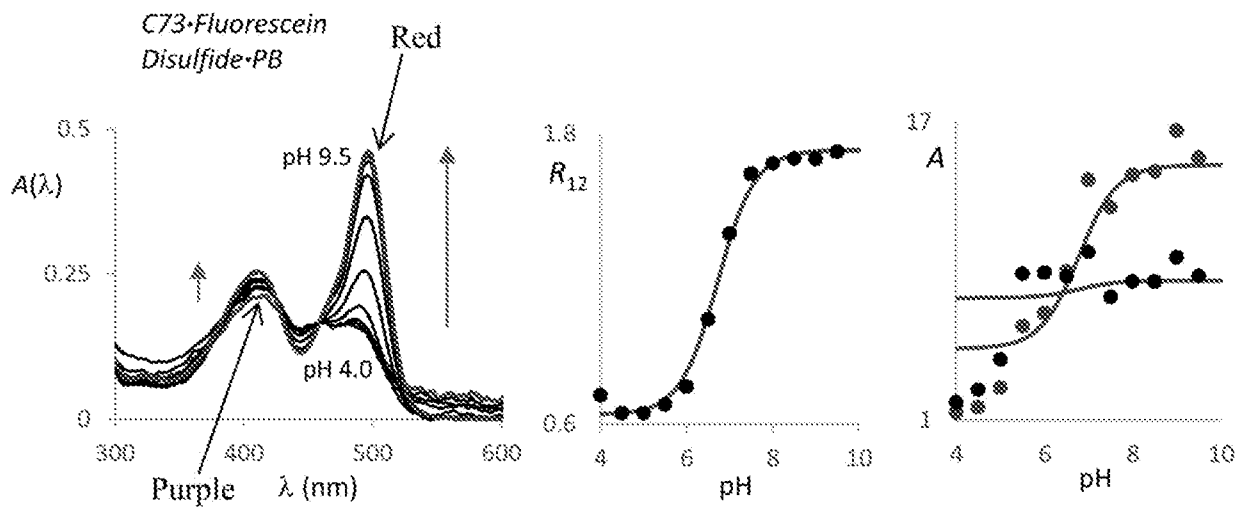


FIG. 25C

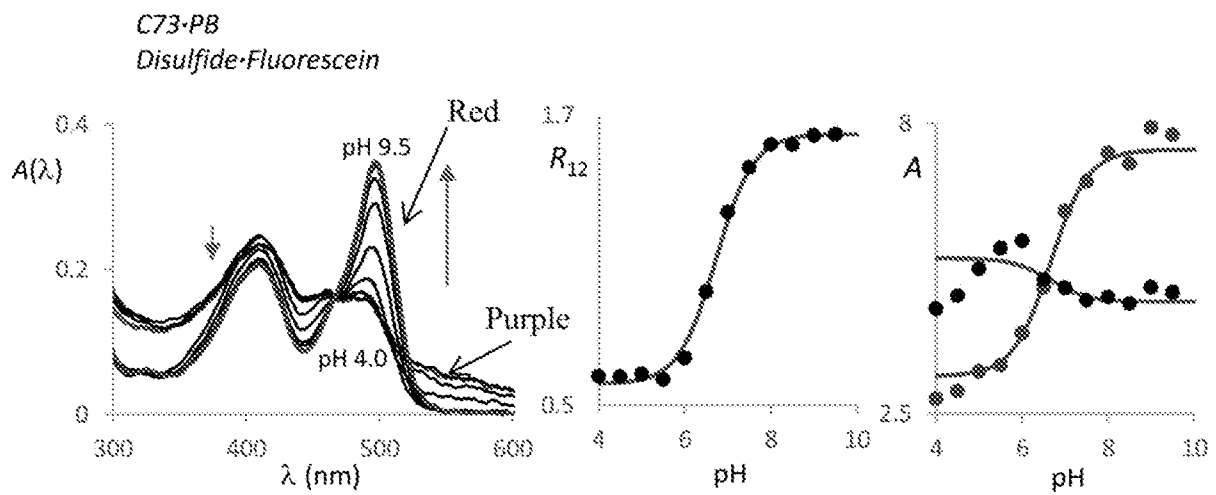
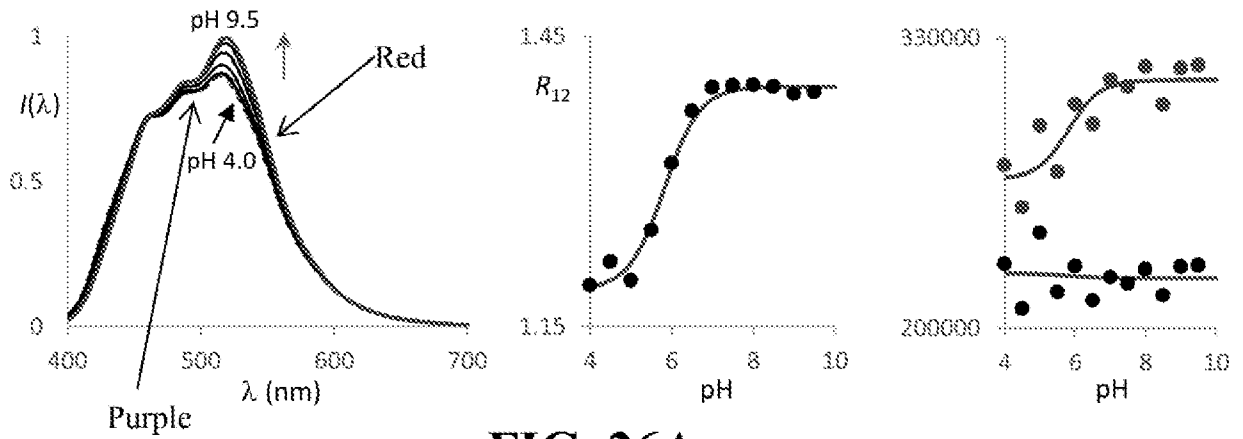
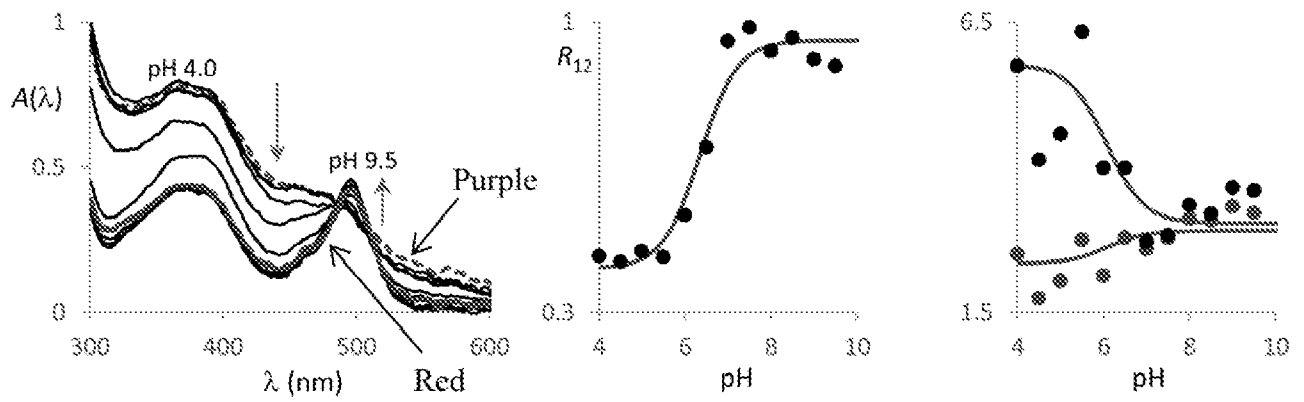


FIG. 25D

**FIG. 26A****FIG. 26B**

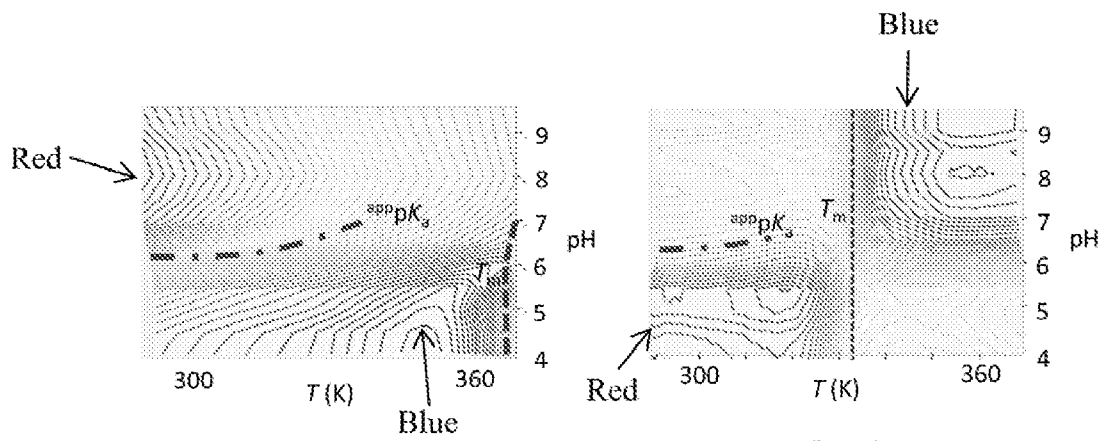
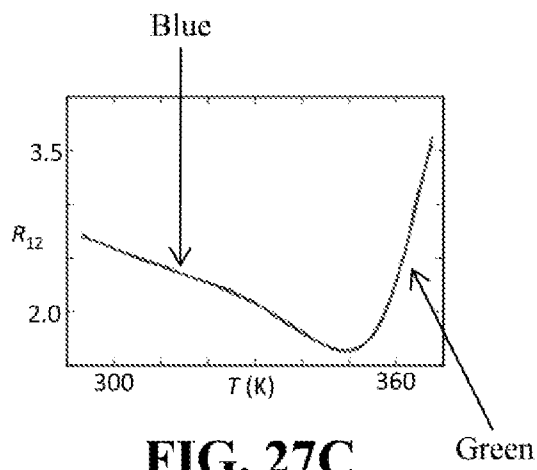
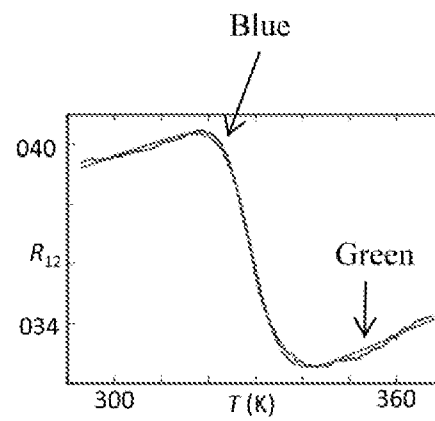
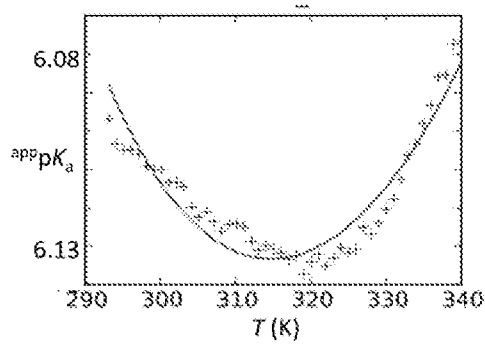
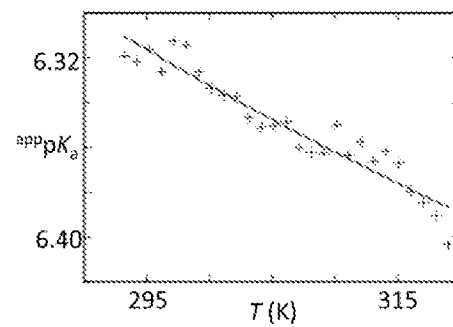
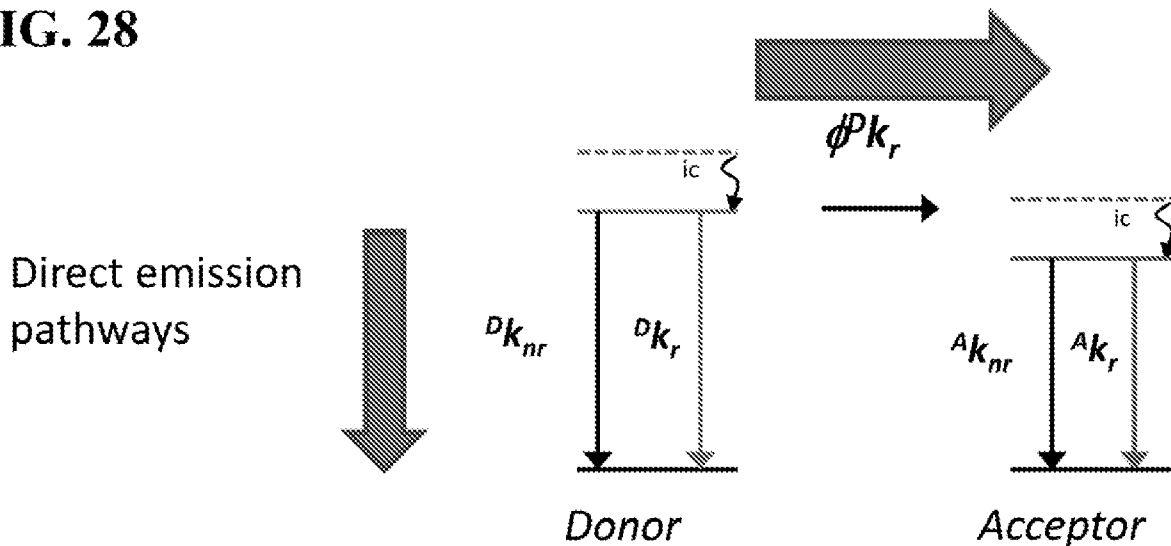
**FIG. 27A****FIG. 27B****FIG. 27C****FIG. 27D****FIG. 27E****FIG. 27F**

FIG. 28**Directly responsive partner**

- Responds directly to ligand-induced protein conformational changes
- Binds ligand (chemosensor)

Changes due to balance on photon flow in FRET and/or direct emission pathways

Indirectly responsive partner

No interactions with ligand or protein conformational changes
Changes only due photon flow in FRET pathway

Effects depend on role of directly responsive partner

Donor: Photon flow through competing **output** pathways

Outputs: Direct emission pathway (quenching) and FRET (spectral overlap) pathway

Acceptor

Balance of photon flow through **input** and **output** pathways

Input: FRET pathway (spectral overlap only)

Output: Direct emission pathway (quenching)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCGCGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
990 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTGCTGAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCAAAACGACCTTCCTCTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 29 (Continued)

970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
1110
1120
1130
1140
1150

GCATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAGCTCG
TTTGCCCAAGAACTCCCAAAAGGACCTTCCTCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 30 (Continued)

FIG. 31 - Exemplary Chloride-binding protein 3 (ClBP3) Expression Construct

CGGTCACTGGAGTCCATAGGTCGCGGTGATGCGGCGGACAGATCGGCGGTAGAGGATCGAGATCTGATCCGCGAAATTAATACGACTCATTATAGGAGACCCAAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTGGGAAACCTGACGGGTATCCGACCGGCGCCACTACGCGCGGTGCTACGCGGCGCATCTCCCTAGCTTAGAGTGTATATATGCTGAGTGATATCCCTCTGGSTGTG
 130 140 150 160 170 180 190 200 210 220 230 240
 GGTTCCTCTAGAAATAATTTGTGTTAACTTTAAGAAAGGAGATATACCATGGTAAGTGAATATATATTCACAGTGTACTCCCAATCTCTGAGAGCTCGCGGATGTCAA
 CCAGAGGAGATCTTTATTAACAAATGAAATCTCTCTATATGATGATACCATTCATCTCCACTTCTTAATAGTGTCCACATCAGGCTTAGAGGATCTTCGAGCTGCGCGTACAGTT
 250 260 270 280 290 300 310 320 330 340 350 360
 GCGCCAAATTTTACGTCGCGGAGCGGAGGAGGAGGACGCAACATATGSCAAACTACCTCGAAATCACTGTGACCCACGCGGAAATGCGCGGTACCATGCTCCTGTAACAC
 GCCGGTGTAAAGATCAGAGCCCTCCCTCCCTCCCTGCGTTGTGTATATACCGTTTGTAGTGGGACTTTAAGTAGACGCGTGGCCCTTTGACGGCCATGGTACCGGATGAGCATTTGGTG
 370 380 390 400 410 420 430 440 450 460 470 480
 AAAACCGATCCACAGGTCACAAAGCGGTGCANTAGGTCGTGTGACGAGTTGACTGAAGAACTTAAAGCCGATACGCTCTCCGATACACGTCCTCGGTCTAAAGAAAGTTCCTACT
 FGTGGCTACGGTGTCCAGTGTTCGCACCGTTATCCAGACACCATCGCTCAACATGACTTCTCAAAATCGGCCATGCCAGAGGCTATGTGCAAGCGCAGACGATTTCTTCAAGGATGA
 490 500 510 520 530 540 550 560 570 580 590 600
 GCAATGATATCTGTGCGCGGCTCCAGTTTAAAGCTCCCTATGGAATCATTTGGCATAGCTCGAGTTCCGTAATCCCTTAATAAATCTCTACCGTTGAGGAGCCAGTATTTGACCCAT
 N Y N S H N V Y I M A D K Q K N G I K V N F K I R H N I E D G S V Q L A D H Y Q
 CAATACAACTCCACAAATGTCTACATATATGCGAGATAAACAAAGAACGCGGATTAAGTCAACTTCAAAATTCGCCATAACATCCAGAGTGGCTCGGTCAGTAGCGGATCATATCA
 610 620 630 640 650 660 670 680 690 700 710 720
 GTTGATGTTGAGGTTGTACAGATGTAATACCGTCTATTTGTTTCTTGCCCTAATTCAGGTTGAAGTTTAAAGCGGCTATTTAGCGGCTTCTACCGGAGGTCATCCGCTAGTATAGT
 Q N T P I G D G P V L L P D N H Y L S Y Q S A L S K D P N E K R D H M V L L E F
 GCAGAACACACCTATCGGTGACGGTCCCTGTACTCTCCCGGATACCCACTATCTCTTACCAAAGTCCCTCAGTAAAGACCCCAATGAGAGCGCGGATCACATGGTATTATCTCGAATT
 730 740 750 760 770 780 790 800 810 820 830 840
 CGTCTGTGTGATAGCCACTGCCAGGACATGAGGAGGCGCTATTGGTGTATAGAGAGATGGTTTCACGGGAGTCAATTTCTGGGTTACTCTTCGCGTAGTGTACCATATAGAGCTTAA
 V T A A G I T L G M D E L Y K G G S N D Y K D D K G G S H H H H *
 TGTGACCGAGCTGCATCACGCTCGGATGGAGCACTCTATATAGGCGGCTAGCAATGATTAACAAGCAGCGGATCAAAAGGTGGCTCTCATCATCATCATCATCATTAATGAAGGG
 850 860 870 880 890 900 910 920 930 940 950 960
 AACTGCGGTGACCGTATGCGAGCCCTACCTGCTTGAGATATTTCCCGGCGCATGCTTAAGTCTTGTGCTACTGTGTCCACCGAGAGTAGTAGTAGTAATTAATTACTTTCC
 CGATATCCAGCACACTGGCGGCGGTACTAGTGGATCCGGGTGCTAAACAAAGCCGGAAGGAGCTGAGTGTGGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGGCCCTCT

970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
GCTATAGGTCGTGTACCCGCGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
AAACGGGCTTTGAGGGGTTTTTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAGCTCG
TTTGCCAGAGACTCCCGCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC
1090
1100
1110
1120
1130
1140
1150

FIG. 31 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATTGTTTCGGGCTTTTCCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
1090 1100 1110 1120 1130 1140 1150
TTTGCCCAAGAACTCCCAAAACGACCTTCCTTCGATATAGGCTCGCTGAGGGTGCCGTCGCAACCGTTGAGC

FIG. 32 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCGCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 33 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTTCCTTCGACTCAACCGACGACGGTGGCCACTCGTTATTGATCGTATTGGGAACCCCGGAGAGA

 AAACGGGCTTTGAGGGGTTTTTGTCTGAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
 1090 1100 1110 1120 1130 1140 1150
 TTGGCCAGAACTCCCCAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 34 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCGCGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAA
1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCAAAAACGACCTTCCTCTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 35 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTTCCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
1090 1100 1110 1120 1130 1140 1150
TTTGCCAGAACTCCCGCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 36 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
1090 1100 1110 1120 1130 1140 1150
TTTGCCCAAGAACTCCCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 37 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
990 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCAAAAGGACCTTCCTCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC
1090 1100 1110 1120 1130 1140 1150

FIG. 38 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTTCCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCGCAAAAACGACATTTCCTTGTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 39 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTTCCTTCGACTCAACCGACGACGSGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA

 AAACGGGCTTTGAGGGGTTTTTCTGTAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
 1090 1100 1110 1120 1130 1140 1150
 TTGGCCAGGAACTCCCGCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCGCAACCGTTGAGC

FIG. 40 (Continued)

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
GCTATAGGTCGTGTACCCGCGCAATGATCACCCTAGGCCGACGATGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
990 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150
AAACGGGCTTTGAGGGGTTTTTGTGCTGAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCAAAACGACCTTCCTCTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 41 (Continued)

970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
GCTATAGGTCGTGTACCCCGGCAATGATCACCCTAGGCCGACGATTGTTTCGGGCTTTCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAGAGA
1090
1100
1110
1120
1130
1140
1150
AAACGGGCTTTGAGGGGTTTTTGTCTGAAAGGAGGAACATAATCCGGAGCGACTCCACGGCAAGTTGGCAAAGCTCG
TTTGCCCAAGAACTCCCGCAAAAACGACATTCTCCTTGATATAGGCTCGCTGAGGGTGCCGTCACACCGTTGAGC

FIG. 42 (Continued)

CGGTCA	CGCTTTGGGA	CTGGCCATAG	AGCTGGCCGGTGA	TCCCGGCCAC	CGATGCTCCGGCTGAG	AGATCGAGAT	CTCGATCCCGGGA	AAATTAATAC	GCAC	CTACTATAG	GGGAGAC	CCACAA	10	20	30	40	50	60	70	80	90	100	110	120
GC	CAGTGC	GAACCC	CTGAC	GGTATCC	GACCGGGCCAC	TACGGCCGGTGTAC	GCAGGCCCATCTC	CTAGCTCTAG	AGCTATTAAT	TATGCTAGATAT	ATCCCTCTG	GTGTG	10	20	30	40	50	60	70	80	90	100	110	120
GGTTTCCCTC	TAGAATAAT	TTTTGTTTAA	CTTTAAGAA	GGAGATATA	CCATGAAACAA	CTCAATA	TCGGTGA	AGCTATTT	GTAAATTC	GACGACAC	ATTTATG	AGATGCG	10	20	30	40	50	60	70	80	90	100	110	120
CCAAAGGGAGAT	CTTTATATAA	CAAAATTC	TCTCTATAT	TGGTACTTT	TGTGATATAG	CCACATCG	ATAACATTA	AGCTGCTGT	GTAAATAC	TGACCCAG	GGCCCTTAC	G	10	20	30	40	50	60	70	80	90	100	110	120
MTAEAQ	QKAKAKL	NMVDS	QNSQP	TQNDQ	VDLFL	ITKKM	NALAI	ITATG	ACCCAGAG	CCCAAGCA	AGCCAGCT	CAGACCTAC	CAAAATG	ATCAGTCA	CGAAAAAT	GAATG	CGTTG	SCAAT	ATATG	AGCTTAA	AGCTTAA	AGCTTAA	AGCTTAA	AGCTTAA
ATACTGGCGT	CTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	CGTTCGGGTTC	
NPVDR	TAAAGT	IAIDKAK	KQANIP	VVFVFN	REPL	EDMK	KWDKV	CAACCC	CTGTAG	ATCGAC	CGGAGCTAT	CAATATCC	TGTAGT	CTTCC	AGCTTCA	CCGCGAG	ACCTTT	CCAGAA	GACATG	AGAAAT	GAATG	CGTTG	SCAAT	
GTG	GGACATC	TAGCGT	GGCGT	CGACCC	CTGATAG	TAGTGT	TCCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	GTTCGTTC	
YVGAK	AKAQ	SEQL	IGILQ	GI	MA	DYWK	AHP	EAD	KN	H	D	GV	MQ	YVM	ATATTA	CGTAG	CGGCAAG	GGCCAA	CAGATG	SGCAT	TTCCAA	SGTCAT	TTCCAA	SGTCAT
HATAAT	GCATCC	CGCTT	CCGCTT	CTCTC	ACCGTA	AGAGTT	CCAGTT	TAGTAC	CGACTA	ATAACCT	TTTCGAG	TAGGCTTC	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	CGCTT	
L	MGQ	P	GHQ	DA	ILR	TA	QY	SI	QT	V	K	DA	GI	K	Y	Q	E	L	A	K	D	Y	A	N
GT	TA	TGG	CCAG	CCCGG	CACT	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	CGCAAT	CTTGA	
CA	ATTA	CCGGT	CGGCG	CCGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	
V	TA	H	DK	M	A	W	L	S	S	F	G	D	K	I	E	A	V	F	A	N	N	D	D	M
TG	TA	CCG	CTCAT	GA	CAAA	TGG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	CTCTG	
AC	AGTGG	GGAG	TA	CTG	TTTAC	CGAC	GAAC	CGCGT	GTCTT	AACTTC	GGCAAA	AAACG	TTTAT	TGCTG	CTAT	ATAC	CGG	CA	CC	CGG	CA	CC	CGG	
Y	F	T	G	N	K	Y	I	P	V	G	V	D	A	T	A	P	G	I	Q	A	I	K	D	M
CT	ATTA	CCG	CGCA	CAAG	TAT	CCAG	TTC	CGAC	CCCA	CGG	GATCC	AGG	GATCC	AGG	GATCC	AGG	GATCC	AGG	GATCC	AGG	GATCC	AGG	GATCC	
GA	TA	AA	TG	CCCG	CTGT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	GTCTG	CTGCT	
K	A	T	F	N	I	A	Y	E	L	A	Q	G	I	T	P	T	K	D	N	I	G	Y	D	
GA	AA	GG	CC	ACT	TTT	CA	CA	AT	GC	AT	GC	AT	GC	AT	GC	AT	GC	AT	GC	AT	GC	AT	GC	
270	280	290	300	310	320	330	340	350	360	370	380	390	400	410	420	430	440	450	460	470	480	490		

CTTCGGTGAAGTTGTACGATGCTGTAACGCTGTCTCCCTAGTCCGGTTGGTTCTCTATTGTAGCCCAATGCTGTAGTGGCTCCCGTTTATGCAACCTTAAGGTATATTTTTTTTAAATGTTT
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 D N I S D A E Q G G S G S T G E K P Y K C P E C G K S F S R S G G S H H H H
 310 320 330 340
 AGACAACTTCGGATGCAAGCAAGTGTTCAGCGCGAGCAACCGGCGAAACCGGTATTAATGTCCGGATGTGCAAAAGCTTTAGCCGCGGCGGCGAGCCCATCATCATCA
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 TCTGTGTAAAGCCCTACGTCCTCGTCCACCAAGTCCCGCTCGTGGCCGCTTTTGGCATATTACAGGCCCTTACACCGTTTTCGAAATCGGCGCTCGCCCGCTCGGTAGTAGTAGT
 H * *
 TCATTAAAGAAAGGCGATATCCAGCAACCTGGCGCGCTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGAAAGCTGAGTTGGTGTGCCACCGCTGAGCAATACATAGCATAA
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 AGTAATTACTTTCCCGCTATAGTTCGTGACCGCGGCAATGATCACCTAGGCGGACGATGTGTTCCGGCTTTCCCTTCGACTCAACCCGACGACGCTGGGCACTCGTTATTGATCCTGATTT
 CCCCTTGGGSCCTCTAAACGGGCTTGAGGGGTTTTTTTGTGAAGAGGAACTATATCCGGAGCGACTCCACGGCACGTTGGCAAGCTCG
 1330 1340 1350 1360 1370 1380 1390 1400 1410
 GGGGAACCCCGGAGATTTGCCAGACTCCCAAAACGACTTTCTCTCTGATATAGGCCCTCGCTGAGGGTSCCGTGCACACCGTTCGAGC

FIG. 43 (Continued)

66/99

CTTTCGGTGAAGTTGTACGATGCTGTAACGCTGTTCCTCCCTAGTCGCGTTGGTTTCTATTGTAGCCCAATGCTGTAGTGGCTCCGTTTATGCAACCTAAGGTATATTTTTTTTAAATGTTT
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 D N I S D A E Q G G S G S T G E K P Y K C P E C G K S F S R S G G S H H H H
 310 320 330 340
 AGACAAATTCGGATGCAGAGCAAGGTGGTTTCAGCGCGGACACCGGCGAAACCGGTATAATGTCCGGATGTGGCAAAAGCTTTAGCCGCGGCGGCGGCAATCATCATCATCA
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 TCTGTGTAAAGCCCTACGTCCTGTTCCACCAAGTCCCGCTGCGTGGCCGCTTTTGGCATATTACAGGCCCTTACACCGTTTTCGAAATCGGCGCTCGCCCGCTCGGTAGTAGTAGT
 H * *
 TCATTAAAGAAAGCGCGATATCCAGSACAATGGCGCGCTTACTAGTGAATCCGGCTGCTAACAAAGCCCGAAAGAAAGCTGAGTTGGCTGCCACCGCTGAGCAATAACTAGCATAA
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 AGTAATATCTTCCCGCTATAGGTCGTGTGACCGCGGCAATGATCACCTAGGCGGACGATGTTTCGGCTTTCCTTCGACTCAACCGGACGACGCTGGGCACTCGTTATTTGATCCTGATTT
 CCCCTTGGGSCCTCTAAACGGGCTTGAGGGGTTTTTTTGTGAAGAGGAACCTATATCCGGAGCGACTCCACGGCACGTTGGGCAAGCTCG
 1330 1340 1350 1360 1370 1380 1390 1400 1410
 GGGGAACCCCGGAGATTTGCCCGAGACTCCCAAAACGACTTTCCTCTTGATATAGGSCCTGCTGAGGGTSCCGTGCACACCGTTCGAGC

FIG. 44 (Continued)

CTTTCGGTGCAAGATTGTAACGATAGCTGATCCCTAGACGGGTGGTGTCTATATTGATAGCAATGCTGTAGTGGCTCCGCTTTATGCAAAACCTTAAGGTATATATTTTAAAGTTT	970	980	990	1000	1010	1020	1030	1040	1050	1060	1070	1080
D N I S D A E Q G G S G G S T G E K P Y K P Y K P E C G K S F S R S G G S H H H H H H	310	320	330	340								
AGACAAACATTTCCGATGACAGACCAAGGTGTTTCAGCGCGACAGCCCGGCGAAAACCGATATAATGTCCGGAATGTGGCAAAAAGCTTTTAGCCGAGCGCGGCGCATCATCATCATCA												
TCTGTTGTAAAGCCTACGTCTCGTTCCACCAAGTCCGCGCTCGTGGCCGCTTTTGGCATATTTCAGGCCCTACACCGCTTTCGAAATCGGGCTCGCCCGCTCGGTAGTAGTAGT	1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200
H * *												
TCATTAATGAAGGGCGGATATCCAGCACACTGGCGGCCCTTACTAGTGGATCCGGCTCTAAACAACCCGAAAGGAAGCTGAGTTGGCTGTCCACCGCTGAGCAATAACTAGCATAA												
AGTAATTACTTTCCCGTATAGGTCTGTGTGACCGCGGCAATGATCATCCTAGCGCGAGCATTTGTTTCGGGTCTTCCTTCGACTCAACCGACGAGTGGCGACTCGTATTTCATCGTATT	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320
CCCTTTGGGSCCTCTAAACGGGCTCTGAGGGGTTTTTGTGTAAGGAGAACTATATCCCGAGCGCACTCCACGCACTGTGGCAAGGTCG												
GGGAAACCCCGGAGATTGGCCAGAACTCCCAAAAAACGACTTTCTCTCTTATATATAGGCTCTGAGGGTSCGTGCAACCCGTTCCAGC	1330	1340	1350	1360	1370	1380	1390	1400	1410			

FIG. 45 (Continued)

[illegible]

CTTTCGGTGAAGTGTGTAACGTAATGCTGTAACGTTGCTCCCTAGTCCGGTTGGTTCTCTATTGTAGCCCAATGCTGTAGTGGCTCCGTTTATGCAACCTAAGGTATATTTTTTTAAATGTTT
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 D N I S D A E Q G G S G S T G E K P Y K C P E C G K S F S R S G G S H H H H
 310 320 330 340
 AGACAAATTCGGATGAGAGCAAGTGTGTTACAGCGGCGAGCACCGGCGAATAACCGTATATAATGTCCGGATGTGCAAAAGCTTTAGCCGCGGCGGCGAGCCCATCATCATCA
 TCTGTGTAAAGCCTACGTCCTGTTCCACCAAGTCCCGCTGCGTGGCCGCTTTTGGCATATTACAGGCCCTTACACCGTTTTCGAAATCGGCCGTCGCCCGCTCGGTAGTAGTAGT
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 H * *
 TCATTAAAGAAAGGCGGATATCCAGSACA CTGGCGCGCTTACTAGTGAATCCGGCTGCTAACAAAGSAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATACTAGCATAA
 AGTAATTACTTTCCCGCTATAGGTCGTGTGACCGCGGCAATGATCACCTAGGCGGACGATGTTTCGGCTTTCTTCGACTCAACCCGACGACGCTGGGCACTCGTTATTGATCCTGATTT
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 CCCCTTGGGSCCTCTAAACGGGCTTGAGGGGTTTTTTTGCTGAAGAGGAAGAACTATATCCGGAGCGACTCCACGGCACGTTGGGCAAGCTCG
 GGGGAACCCCGGAGATTTGCCAGACTCCCAAAACGACTTTCTCTCTGATATAGGSCCTGCTGAGGGTSCCGTGCACACCGTTCGAGC
 1330 1340 1350 1360 1370 1380 1390 1400 1410

FIG. 46 (Continued)

CGCTCA	CGCTTGGGA	CTGGCCATAG	AGCTGGCCG	GATGCCGGTGA	TCCGGCCAC	GATGCGATCG	AGATCGAGAT	CTCGATCC	CGCGAAATTA	TACGACT	CTACTATAG	GGGAGAC	CCACAA
10	20	30	40	50	60	70	80	90	100	110	120		
GC	CAGTGC	GAACCC	CTGAC	GGTATCC	GACCGG	CCACTAG	CGCCGGTGT	CTACG	CGCCGCT	TTTAAAT	TATGCTG	AGTGA	TATCCCTCTG
130	140	150	160	170	180	190	200	210	220	230	240		
CC	AAAGG	GAGATCT	TTTATTA	AAACAA	ATTC	TTCTCT	TATATG	GTACT	TTTGTG	ATATAG	CCACAT	CGATAT	TATGCTAG
250	260	270	280	290	300	310	320	330	340	350	360		
AT	ACTG	CGCTTC	CGGTTC	CGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC
370	380	390	400	410	420	430	440	450	460	470	480		
GT	TGGG	ACATCT	AGCG	TGGG	CGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC	CGGTTC
490	500	510	520	530	540	550	560	570	580	590	600		
TA	TAATG	CA	TCCG	CTTCT	CA	CGTA	AGGTT	CCAG	TTTCT	AG	TTTCT	AG	TTTCT
610	620	630	640	650	660	670	680	690	700	710	720		
CA	ATTA	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
730	740	750	760	770	780	790	800	810	820	830	840		
AC	AGT	G	GGG	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT
850	860	870	880	890	900	910	920	930	940	950	960		
GA	TA	AA	TG	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
970	980	990	1000	1010	1020	1030	1040	1050	1060	1070	1080		
CT	ATTA	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200		
LT	MA	TG	GG	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320		
Y	Y	V	G	A	K	A	E	Q	S	G	I	L	Q
1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440		
AT	ATTA	CG	TAG	GG	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560		
TA	TAATG	CA	TCCG	CTTCT	CA	CGTA	AGGTT	CCAG	TTTCT	AG	TTTCT	AG	TTTCT
1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680		
L	M	G	Q	P	G	H	Q	D	A	I	L	R	T
1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800		
GT	TA	TG	GG	CCG	GTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC	GGTTC
1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920		
V	T	A	H	D	K	M	A	A	W	L	S	S	F
1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040		
TC	TA	CCG	CT	CA	TG	CA	AA	TG	CG	CT	CT	CT	CT
2050	2060	2070	2080	2090	2100	2110	2120	2130	2140	2150	2160		
AC	AGT	G	GGG	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT
2170	2180	2190	2200	2210	2220	2230	2240	2250					

CTTTCGGTGAAGTTGTACGATGCTGTAACGCTGTTCCTCCCTAGTCCGGTTGGTTTCCTATTGTAGCCCAATGCTGTAGTGGCTGCCGTTTATGCAACCTAAGGTATATTTTTTTTAAATGTTT
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 D N I S D A E Q G G S G S T G E K P Y K C P E C G K S F S R S G G S H H H H
 310 320 330 340
 AGACAAATTCGGATGCAGAGCAAGTGTTCAGCGCGGCGACCGCGCAAAACCGTATAATGTCCGGATGTGGCAAAAGCTTATAGCCGCGGCGAGCCCATCATCATCA
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 TCTGTGTAAAGCCTACGTCCTGTTCCACCAAGTCCCGCGTGGCGCTTTTGGCATATTACAGGCCCTTACACCGTTTTCGAAATCGGCGTCCGCCCGTCGGTAGTAGTAGT
 H * *
 TCATTAAAGAAAGCGCGATATCCAGCACACTGGCGCGCTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGAAAGCTGAGTTGGCTGCCACCGCTGAGCAATAACTAGCATAA
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 AGTAATTACTTTCCCGCTATAGTTCGTGACCGCGCGCAATGATCACCCTAGGCGGACGATGTTGTTCCGGCTTTCTTCGACTCAACCGAGCACCGGTGGCGACTCGGTATTGATCCGTATT
 CCCCTTGGGSCCTCTAAACGGGCTTGAGGGGTTTTTTTGTGAAGAGGAACCTATATCCGGAGCGACTCCACCGCACGTTGGCAAGCTCG
 1330 1340 1350 1360 1370 1380 1390 1400 1410
 GGGGAACCCCGGAGATTTGCCAGACTCCCAAAACGACTTTCTCTCTTGATATAGGCTCTGAGGGTSCCGTGCACACCGTTCGAGC

FIG. 47 (Continued)

[illegible]

[illegible]

FIG. 48 (Continued)

ACGCTATGACGCAATGTCATGTCGTAGTGTGCTGCTTACGCTTCTGATGCGTTTGTGACCCCTAGCCCAATGCGCGCTACTGTGTTTACCCGTCG
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 W L S S F G D K I E A V F A N N D A M A L G A I E A L K S A G Y F T G N K Y I P
 310 320 330 340
 GTGGTTGTCAGGCTTGGCGACAAAGATCGAAAGCCGTCTTCGGAACAACACGCGATGCGGTGGGTGCGTGAAGTCTGAGGCTACTTCACCGSCAAACAATACATCCC
 CACCAACAGCTCGAAACCGCTGTCTTAGCTTTCGGCAGAAAGCCGCTTGTGCTGCGTACCGCAACCCACGCGTACCTTCGCGACTTCAGACGTCCTCCGATGAAGTGGCCGTTGTTATGTAGGG
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 V V G V D A T A P G I Q A I K D G T L L G T V L N D A K N Q A K A T F N I A Y E
 350 360 370 380
 GGTGTGGCGTGTGATGCGACCCGACGAGGACAAACATAGGGTACGATCAAGATGTTGGAACCCGTGTTGAACGATGGAAGAACCGGCGAAGCCACTTCAACATCCGCTACGA
 CCAACACCCGACCTACGCTGSGCGTGGCCGTAAGTCCGCTAGTTCCTACCATGGGACAAACCCCTTGGACAACTTGCTACGCTTCTGTTGCTTCTGTTGCTGGAAGTGTAGCGGATGCT
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 L A Q G I T P T K D N I G Y D I T D G K Y V W I P Y K K I T K D N I S D A E Q G
 390 400 410 420
 ACTGGCTCAGGGCATTTACCCGACGAGGACAAACATAGGGTACGATCAAGATGTTGGAACCCGTGTTGAACGATGGAAGAACCGGCGAAGCCACTTCAACATCCGCTACGA
 TGACCGAGTCCCGTAATGGGGCTGCTTCCTGTGTGTAATCCCATGCTGTAGTGTCTGCTTCAATGACACACCTTAAGGCAATGTTTCTAGTGTCTTCTGTTGTTATATAGTCTTACGACTTGTATCC
 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440
 G S H H H H H H
 430
 AGGAAGCCATCATCATCATCATTAATATAAAGSGGATATCCAGCACACTGGGCGCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCCGAAAGGAGAGTGTGGTGTGCTGCC
 TCCTTCGGTAGTAGTAGTAGTAGTAATCATTAATTTTCCCGCTATAGTTCGTGTGACCGCGGCAATGATACCTAGGCGGACGATTTGTTTGGGCTTTCTTTCGACTCAACCGGACGCGG
 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560
 ACCGCTGAGCAATTAAGTACATTAACCCCTTGGGGCCCTTAACGSGGCTTTGAGGGGTTTTTGTGTAAGGAGGAATATATCCGGAGCGCACTCCCAAGGACGCTTGTGSCAAGCTCG
 TSGCGACTCGCTTAATGATCGTAATGGGGAACCCCGGAGATTTGCCAGAACTCCGCCAAAACGACTTTCCTCTTGAATATAGGCTTCGCTGAGGTCGCTGCAACCGCTTCGAGC
 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670

FIG. 49 (Continued)

[illegible]

79/99

FIG. 50 (Continued)

ACGCTATGACGCGATGTCATGTCGTAGGTTTGCCACTTTTACGTCCTGACCGGTTTCTGATGCGTTTGACGCTAGCGCAATGCGCGCTACTGTTTTACCGCTCG
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 W L S S F G D K I E A V F A N N D D M A L G A I E A L K S A G Y F T G N K Y I P 340
 GTGGTTGTCAGGCTTGGCGACAAAGATCGAAAGCCGTCTTCGCGAACAACAGACATGCGGTGGGTGCGTGAAGCTGAGGCTACTTCACCGGCAACAATACATCC
 CACCAACAGCTCGAAACCGCTGTCTTAGCTTTCGGCAGAAAGCGCTTGTGCTGTGATACCGCAACCCACGCTGACCTTCAGACGTCCTCGATGAAGTGGCCGTTTATGTAGGG
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 V V G V D A T A P G I Q A I K D G T L L G T V L N D A K N Q A K A T F N I A Y E 380
 GGTGTGGCGTGTGATGCGACCCGACGAGGACAAATAGGGTACGATCAAGGATGTTGGAACCGGTGGAACCGGTGGAACCGGCGAAGCCACCTTCAACATCGCTACGA
 CCAACACCCGACCTACGCTGCGGTGGCGCCGTAAGTCCGCTAGTTCTTACCATGGGACACCCCTTGGACAACTTGTGACGTTCTGTTGCTTCTGTTGCTGGAAGTTGTAGCGGATGCT
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 L A Q G I T P T K D N I G Y D I T D G K Y V W I P Y K K I T K D N I S D A E Q G 420
 ACTGGCTCAGGCGATTACCCGACGAGGACAAATAGGGTACGATCAAGAGGAGTACGTGTGGATTCGGTACAAAGATCAGAAACATATATACAGATGCTGAACAAGG
 TGACCGAGTCCCGTAATGGGCGTGTCTTCTGTGTAATCCCATGCTGTAGTCTGCGCTTCAATGACACACCTAAGGCAATGTTTCTAGTGTCTTCTGTTATATATAGTCTTACGACTTGTTC
 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440
 G S H H H H H H 430
 AGGAAGCCATCATCATCATATTACTATTAATAAAGGCGATATCCAGCACACTGGGCGCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGAGTGTGGTGTGCTGCC
 TCCTTCGGTAGTAGTAGTAGTAGTAATCATATTATTTTCCCGCTATAGTTCGTGTGACCGCGGCAATGATACCTAGCGCGACGATTTGTTTGGGCTTTCTTTCGACTCAACCGACGACGG
 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560
 ACCGCTGAGCAATTAAGTACATTAACCCCTTGGGCGCTTAACCGGCTCTTGAAGGGTTTCTGCTGAAGAGGAACTATATCCGGAGCGCACTCCACCGCACCTTGTGSCAAGCTCG
 TSGCGACTCGCTTATGATCGTAATGSGGAAACCCGAGAGATTTCAGAGAACTCCCAAAAGACATTTCTCTTGTATATAGGCTTCGTGAGGTCGCTGCAACCGTTTCGAGC
 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670

FIG. 51 (Continued)

FIG. 52 - Exemplary Adaptor protein construct - Adaptor0

CGGTACGCTGGGACTGCCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGACTAGGGCGCTTTAATTATATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGGATGTACTCAAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTTATTAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAACACTGTGCTACATGAGTTTCGCTGCCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGGACGTCTTATAGTCCCGTGTGACTGGCAACCGTTACGACTGTGTAGCTAGTGT
 130 140 150 160 170 180 190 200 210 220 230 240
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTCTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGGTGAGTTGAAGAGTTCCCT
 110 120
 GGGACCGTGAACCGCGCTTTATACCGTAGGCACCATAGGSCGTGAGACGACACAAAGTTTGTGCCACTTCCACCGCGTGTGCTACCCACGTCACAGATTTCCTCAGTCAACTTTCCTCAAGGA
 130 140 150 160 170 180 190 200 210 220 230 240
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCCATCATCATCATATCAATTAATGAAGAGGCGGATATCCAGCACACTGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGGAAAGGAA
 150 160 170 180 190 200 210 220 230 240
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAGTAATTAATTACTTTCCCGCTAAGGTCGTGTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTTCCTT
 250 260 270 280 290 300 310 320 330 340 350 360
 GCTGAGTTGCTGTGCTGCGCAATAACTAGCATAAACCCCTTGGGCGCTTAACGGGTCTTTGAGGGGTTTTTGTGAAAGAGGAACTATATCCGAGCGCACTCCCAAGCA
 370 380 390 400 410 420 430 440 450 460 470 480
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCTCTCTGATATAGGCTCGCTGAGGGTSCCGT
 490 500 510 520 530 540 550 560 570 580 590 600
 610 620 630 640 650 660 670 680 690 700 710 720
 CGTGGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 53 - Exemplary Adaptor protein construct - Adaptor1.0

```

CGGTACGCTTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGGAAATTAATACGACTCATTATAGGAGACCCACAAC
GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCCGCTGCTACGCGCCGCTATCCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
10      20      30      40      50      60      70      80      90      100     110     120

GGTTTCCTCTAGAAAATAATTTTGTATTAACTTTAAGAAAGGAGATATACCATGAGCGCGGAAAATTTATTCACCTGACTGACGACAGTTTGGACACGGATGTACTCAAAGCGGACGGGGGAT
20      30      40      50      60      70      80      90      100     110     120
M S A K I I H L T D D S F D T V L K A D G A I
GGTTTCCTCTAGAAAATAATTTTGTATTAACTTTAAGAAAGGAGATATACCATGAGCGCGGAAAATTTATTCACCTGACTGACGACAGTTTGGACACGGATGTACTCAAAGCGGACGGGGGAT
130     140     150     160     170     180     190     200     210     220     230     240
CCAAAGGGAGATCTTTATTAAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCGCTA
130     140     150     160     170     180     190     200     210     220     230     240

L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
30      40      50      60      70      80      90      100     110     120     130     140
CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCGTGGATGAAATCGCTGAGCAATATCAGGGGCAAACTGACCGTTGCAATGCTGAATCCTGATCAAAA
250     260     270     280     290     300     310     320     330     340     350     360
GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTAGCTAGTATTTT
250     260     270     280     290     300     310     320     330     340     350     360

P G T A P K Y G I C G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
70      80      90      100     110     120     130     140     150     160     170     180
CCCTGGCACATGCGCGGAATATGSCATCTGTGTGTAATCCCGACTCTGTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGTTCAGTTGAAAGAGTTCCCT
370     380     390     400     410     420     430     440     450     460     470     480
GGACCGTSGACCGGCTTTATACCGTACACACCATAGGSGCTGAGACGACACAGTTTTCACCCGCTTGGTTTCAACCCACGTGACAGATTTCAGTCAACTTTCTCAAGGA
370     380     390     400     410     420     430     440     450     460     470     480

D A N L A G G S H H H H H * * *
110     120     130     140     150     160     170     180     190     200     210     220
CGACGTAACCTGGGCGGCGGACCATCATCATCATATCAATTAATGAAGAGGCGGATATCCAGCACATGSGCGCCGTTACTAGTGGATCCGGTCTCTAACAAAGCCCGGAAAGGAA
490     500     510     520     530     540     550     560     570     580     590     600
GCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTCTCTT
490     500     510     520     530     540     550     560     570     580     590     600

GCTGAGTTGCTGTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTGAAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGGACTCCACCGCA
610     620     630     640     650     660     670     680     690     700     710     720
CGACTCAACCGACGCGGTGGGACTCTGTTATTTGATCGTATTTGGGAAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCTCTCTGATATAGGCTCGCTGAGGGTSCCGT
610     620     630     640     650     660     670     680     690     700     710     720

CGTTGGCAAGCTCG
GCAACCGTTCGAGC
730

```


FIG. 55 - Exemplary Adaptor protein construct - Adaptor2.0b

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGCGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCCGCTCTAGCTAGCGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTATTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGCGATGTAATCAAGCGGACGCGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGAGATCTTTATTAAACAAATGAAATCTTCCCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAACACTGTGCTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAATCGATCGATCAAAA
 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCGTTTACTACGCGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTAGCTAGTTT
 370 380 390 400 410 420 430 440 450 460 470 480
 P G T A P K Y G I C G I P T L L L F K N G E V A T K V G A L S K G Q L K E F L
 CCCTGGCACATGCGCGGAATATGSCATCTGTGTATCCCGACTCTGCTGTGTGTTCAAAAACGGTGAATGGCGCAACCAAGTGGTGCACCTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GSGACCGTSGACCGCGCTTTATACCGTACACACCATAGGSGCTGAGACGACACAAGTTTGTGCGCACTTCCACCGCGTGGTTTCAACCGTACAGATTTCCAGTCACTTTCTTCAAGGA
 D A N L A G G S H H H H H * * *
 CGACGTAACCTGGCGGCGGCGAGCCATCATCATCATATTAATGAAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGAATCCGGTCTCTAACAAAGCCCGGAAAGGAA
 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTGACCGCGCGCAATGATCACCCTAGCGCCGAGATTGTTTCGGGCTTTTCCTT
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTCTTTGTTGTGAAGAGGAACTATATCCGAGCGCACTCCACCGCA
 610 620 630 640 650 660 670 680 690 700 710 720
 CGACTCAACCGAGCGGTGGGACTCTGTTATTGATCGTATTGGGGAACCCCGAGATTGTGCCAGACTCCCCAAGAACGACTTTCTCTCTGATATAGSCCTCGCTGAGGGTSCCGT
 CGTGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 56 - Exemplary Adaptor protein construct - Adaptor3.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGCGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCCGCTCCTAGCTCTAGACTAGCGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTTCACCTGACTGCTGCACTGTCAGTTTGGACACGCGATGTACTCAAAGCGGACGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAAAAACAATGAAATCTTCCCTATATGGTACTCGCGCTTTTAAATAGTGGACTGACTGAGCTCAAAACTGTGCTACTATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D C S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAATCGATCGATCAAAA
 70 80 90 100 110 120
 GGAGCAGCGCAAGACCCCGTCTACCCACGCCAGGCACCTTTTACTACGCGGCTAAGACCTACTTTAGGGACAGCTTATAGTCCCGTGGTGGCACTGGCAACCGTTAGCTAGCTAGTTT
 130 140 150 160 170 180 190 200 210 220 230 240
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTACTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 110 120
 GGGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGCTGAGACGACACAAGTTTGTGCTCCACTTCCACCGCGTGGTTCACCCACGTCGACAGATTCCCACTCACTTTCTCAGGA
 130 140 150 160 170 180 190 200 210 220 230 240
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCATCATCATCATTAATGAAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGGAAAGGAA
 150 160 170 180 190 200 210 220 230 240
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGTCGTGTCGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTCTTT
 250 260 270 280 290 300 310 320 330 340 350 360
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATAAACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTCTTTGTTGTAAGAGGAACTATATCCGAGCGCACTCCCAAGCA
 370 380 390 400 410 420 430 440 450 460 470 480
 CGACTCAACCGAGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCTCTGATATAGGCTCGCTGAGGGTGCCT
 490 500 510 520 530 540 550 560 570 580 590 600
 610 620 630 640 650 660 670 680 690 700 710 720
 CGTGGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 57 - Exemplary Adaptor protein construct - Adaptor4.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGAGTCGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGTGCTGCGATGTAATCAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGTCAAGCTACATGAGTTTCGCTGCCCCCGTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D C D V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGTCGCAAAATGATCGCCCGATTCTGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120
 GGAGCAGCGCAAGACCCCGTCTACCCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTAGCTAGTGT
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 P G T A P K Y G I R G I P T L L L F K N G E V A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTATCCCGTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGTTCAGTTGAAAGAGTTCCCT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GSGACCGTSGACCGCGCTTTATACCGTAGGCACCATAGSGCTGAGACGACGACAAAGTTTGTGCCACTTCCACCGCGTTGCTTCCACCGGACAGATTTCCAGTCAACTTTCCTCAAGGA
 490 500 510 520 530 540 550 560 570 580 590 600
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCCATCATCATCATATCAATTAATGAAGAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGAAAGGAA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGSHAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTHGACCGCGCGSCAATGATCACCCTAGGCGCCGAGGATTTGTTTCGGGCTTTCCTT
 610 620 630 640 650 660 670 680 690 700 710 720 730
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGGTTTTTGTGTAAGAGGAACTATATCCGAGCGCACTCCACCGCA
 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCCTCTGATATAGGCTCGCTGAGGGTSCCGT
 960 970 980 990
 CGTTGGCAAGCTCG
 GCAACCGTTCGAGC

FIG. 58 - Exemplary Adaptor protein construct - Adaptor5.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCATTATAGGAGACCCACAAC
 10 GCGATGCGAACCCTGACGGTATCCGACCGGCCACTACGCGCGGTGCTGACGCGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGGTGTG
 20 30 40 50 60 70 80 90 100 110 120
 GGTTCCTCTAGAAAATAATTTGTGTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGGATGTACTCAAAGCGTGGGGGGGAT
 130 140 150 160 170 180 190 200 210 220 230 240
 CCAAAGGGAGATCTTTATTAAAAAATAATGAAATCTTCCCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGACGCCGCCGCTA
 250 260 270 280 290 300 310 320 330 340 350 360
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 CCTCGTCGCTTCTGGGCGAGTGGTGGTCCGTGCGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAATCGATCGATCAAAA
 370 380 390 400 410 420 430 440 450 460 470 480
 GGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGSGCTGAGACGACACAAGTTTGTGCGCTTCCACTTCCACCGCGTTGCTACCCGACGACAGATTCCCACTCACTTTCCTCAAGGA
 490 500 510 520 530 540 550 560 570 580 590 600
 D A N L A G G S H H H H H * * *
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATTAATGAAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGAAAGGAA
 610 620 630 640 650 660 670 680 690 700 710 720
 GCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTTCCTT
 730
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGGTTTTTGTGAAAGGAGGAATAATATCCGAGCGCACTCCACCGCA
 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990
 CGACTCAACCGACGCGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGAGATTGTCAGACTCCCAAAAACGACTTTCCCTCTGATATAGSCCTCGCTGAGGGTGCCT
 960 970 980 990
 CGTTGGCAAGCTCG
 GCAACCGTTCGAGC

FIG. 59 - Exemplary Adaptor protein construct - Adaptor6.0

CGGTACGCTGGGACTGCCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGACTAGCGCGCTTTAATTATATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTTCACCTGACTGACGACAGTTTGGACACGCGATGTACTCAAGCGGACCGGGGGGAT
 130 140 150 160 170 180 190 200 210 220 230 240
 CCAAAGGGAGATCTTTATTAAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAACACTGTGCTACATGAGTTTCGCTGCCCCCGCTA
 MS AK I I H L T D S F D T V L K A D G A I
 20
 L V A F W A E C C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGTGCGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100
 GGAGCAGCGCAAGACCCGCTCACGACGCCACGACCTTTTACTACGCGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGGCAACCGTTACGACTGTGTAGCTAGTATGTTT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTGTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAAGTGGTGCACCTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 110 120
 GGGACCGTGAACCGCGCTTTATACCGTAGGCAACCATAGGCGCTGACGACGACAAAGTTTGGCCACTTCACCGCGGTGGTTTCAACCGACGACAGATTTCACGTCACCTTTCTCAGTTTCTCAAGGA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCCATCATCATCATATTAATGAAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGGAAAGGAA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAAGGTCGTGTGACCGCGCGCAATGATCACCCTAGCGCCGAGATTGTTTCGGGCTTTTCCTT
 610 620 630 640 650 660 670 680 690 700 710 720
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGCACTCCCAACGCA
 630 640 650 660 670 680 690 700 710 720
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCTCTCTGATATAGSCCTCGCTGAGGGTSCCGT
 650 660 670 680 690 700 710 720
 CGTTGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 60 - Exemplary Adaptor protein construct - Adaptor7.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGCGCCACGATGCGTCCGCGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCCCATCTCCTAGCTCTAGACTAGCGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTG
 GGTTCCTCCTAGAAAATAATTTGTGTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGCGATGTAATCAAAGCGGACGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAACACTGTGCTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K C I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAATGCAATCGCCCGATTCGTGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACGTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGGCAACCGTTACGACTGTAGCTAGTTT
 370 380 390 400 410 420 430 440 450 460 470 480
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGACTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAAGTGGTGCACCTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GSGACCGTSGACCGCGCTTTATACCGTAGGCACCATAGGSGCTGAGACGACACAAGTTTTCGCCACTTCACCGCGTGGTTTCAACCCACGTGACAGATTTCAGTCAACTTTCCTCAAGGA
 490 500 510 520 530 540 550 560 570 580 590 600
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCATCATCATCATATCAATTAATGAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGAAAGGAA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGSHAGTAGTAGTAGTAATTAATTACTTTCCCGCTAAGGTCGTTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTCCTT
 610 620 630 640 650 660 670 680 690 700 710 720 730
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTGAAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGCACTCCACCGCA
 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990
 CGACTCAACCGACGCGTGGGACTCGTTATTGATGCTGATTGGGAAACCCCGAGATTGTCAGACTCCCCAAGAACGACTTTCCTCTGATATAGGCTCGCTGAGGGTSCCGT
 960 970 980 990
 CGTGGCAAGCTCG
 GCAACCGTTCGAGC

FIG. 61 - Exemplary Adaptor protein construct - Adaptor8.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAAATACGACTCATTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATTTATTCACCTGACTGACGACAGTTTGGACACGATGTACTCAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAAAAACAAATTGAAATCTTCCCTATATGGTACTCGCGCTTTTAAATAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D C I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATTGCAATCGGAAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTACGCGGCTAAGACCTAACGTAGGCACTGCTTATAGTCCCGTTGACTGGCAACCGTTAGCTAGCTAGTGT
 370 380 390 400 410 420 430 440 450 460 470 480
 P G T A P K Y G I R G I P T L L L F K N G E V A T K V G A L S K G Q L K E F L
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTACTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GSGACCGTSGACCGCGCTTTATACCGTAGGCACCATAGGSGCTGAGACGACACAAGTTTGTGCTCCACTTCCACCGCGTGGTTTCAACCGACGATTTCCAGTCAACTTTTCTCAAGGA
 D A N L A G G S H H H H H * * *
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATTAATGAAAGGCGGATATCCAGCACATGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGGAAAGGAA
 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAAGGTCGTGACCGCGCGGCAATGATCACCCTAGGCGCCGAGGATTTTCGGGCTTTCTCTT
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTAGGGGTTTTTTGTGAAAGGAGGAACATAATCCGAGCGGACTCCACCGCA
 610 620 630 640 650 660 670 680 690 700 710 720
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGCGCAGACTCCCCAAGAACGACTTTCTCTTGTATATAGGCTCGCTGAGGGTSCCGT
 CGTGGCAAGCTCG
 730
 GCAACCGTTGAGC

FIG. 62 - Exemplary Adaptor protein construct - Adaptor9.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGCGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCCTGACGGTATCCGACCGGCCACTACGCGCCGCTGCTACGCGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGCGATGTACTCAAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAACACTGTGCTTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G C L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCTGCTGACCGTTGCAATGCTGAATCGATCGATCAAAA
 70 80 90 100 110 120
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTACGCGGCTAGACCTTACTTTAGGCACTGCTTATAGTCCCGACGACTGSCAACCGTTACGACTGTAGCTAGTGT
 130 140 150 160 170 180 190 200 210 220 230 240
 P G T A P K Y G I R G I P T L L L F K N G E V A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGACTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGGTGAGTTGAAAGAGTTCCCT
 110 120
 GGGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGCTGAGACGACACAAGTTTGTGCTCCACTTCCACCGCGTTTCAACCGTACAGATTTCCAGTCAACTTTTCTCAAGGA
 130 140 150 160 170 180 190 200 210 220 230 240
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATTAATGAAAGGCGGATATCCAGCACATGSCGCGCGCTTACTAGTGAATCCGGTCTCTAACAAAGCCCGGAAAGGAA
 150 160 170 180 190 200 210 220 230 240
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTGTGACCGCGCGCAATGATCACCCTAGGCGCGGATTTGTTTCGGGCTTTCTTT
 250 260 270 280 290 300 310 320 330 340 350 360
 GCTGAGTTGCTGTGCTGCAACGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGCACTCCCAAGCA
 370 380 390 400 410 420 430 440 450 460 470 480
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTGCCAGACTCCCCAAGAACGACTTTTCTCTTGTATATAGGCTCGCTGAGGGTGCCT
 490 500 510 520 530 540 550 560 570 580 590 600
 610 620 630 640 650 660 670 680 690 700 710 720
 CGTGGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 64 - Exemplary Adaptor protein construct - Adaptor11.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGGATGTAATCAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAAAAACAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGGCAACCGTTACGACTGTAGCTAGTATTTT
 130 140 150 160 170 180 190 200 210 220 230 240
 P G T A P K C G I R G I P T L L L F K N G E V A A T K V G A L S K G Q L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATGCGGCATCCCGTGTATCCCGTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAAGTGGTGCACCTGTCTAAAGTTCAGTTGAAGAAGTTCCCT
 110 120
 GGGACCGTGAACCGGCTTTACGCGGTAGGCACCATAGGCTGAGACGACACAAGTTTTCACCCGCTTGGTTTCAACCCACGTCACAGATTTCCAGTCAACTTTCTCAAGGA
 130 140 150 160 170 180 190 200 210 220 230 240
 D A N L A G G S H H H H H * * *
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATCAATTAATAAGAGGCGGATATCCAGCACATGSGCGCCGTTACTAGTGGATCCGGTCTCAACAAAGCCCGAAAGGAA
 250 260 270 280 290 300 310 320 330 340 350 360
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGGTCGTGTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTCTCTT
 370 380 390 400 410 420 430 440 450 460 470 480
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTTTGTGTGAAGAAGGAGAACTATATCCGAGCGCACTCCACCGCA
 490 500 510 520 530 540 550 560 570 580 590 600
 CGACTCAACCGACGCGGTGGGACTCGTTATTGATCGTATTGGGGAACCCCGGAGATTGTCAGACTCCCCAAGAACGACTTTCTCTCTGATATAGGCTCGCTGAGGGTGCCT
 610 620 630 640 650 660 670 680 690 700 710 720
 CGTTGGCAAGCTCG
 730
 GCAACCGTTCGAGC

FIG. 65 - Exemplary Adaptor protein construct - Adaptor12.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGCGCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 GCGATGCGAACCCTGACGGTATCCGACCGGCGCACTACGCGCGGTGATCGGCGCGCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 20 30 40 50 60 70 80 90 100 110 120
 GGTTCCTCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATTTATTCACCTGACTGACGACAGTTTGGACACGATGTACTCAAGCGGACCGGGGGGAT
 130 140 150 160 170 180 190 200 210 220 230 240
 CCAAAGGGAGATCTTTTATTAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCCGCTA
 250 260 270 280 290 300 310 320 330 340 350 360
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 CCTCGTCGCTTCTGGGCGAGTGGTGGTCCGTGTCGCAAAATGATCGCCCGATTCGTGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAATCGATCGATCAAAA
 370 380 390 400 410 420 430 440 450 460 470 480
 GGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGCGCTGACGACGACAAAGTTTGGCCACTTACCGCGCGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTGTAGCTAGTGT
 490 500 510 520 530 540 550 560 570 580 590 600
 P G T A P K Y G I R G I P T L L L F K N G E C A A T K V G A L S K G Q L K E F L
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTATCCCGTGTGTTCAAAAACGGTGAATGCGCGGCAACCAAGTGGTGCACCTGTCTAAAGGTCAGTTGAAAGAGTTCCCT
 610 620 630 640 650 660 670 680 690 700 710 720
 GSGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGCGCTGACGACGACAAAGTTTGGCCACTTACCGCGCGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTGTAGCTAGTGT
 730
 D A N L A G G S H H H H H * * *
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATCAATTAATAAGGCGGATATCCAGCACATGSCGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGGAAAGGAA
 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAGGTCGTGTGACCGCGCGCAATGATCACCCTAGGCGCGGATTTGTTTCGGGCTTTTCCTT
 910 920 930 940 950 960 970 980 990 1000
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTTAACGGGTCTTGAAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGGACTCCACCGCA
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 CGACTCAACCGGACGCGTGGGACTCGTTATTTGATCGTATTTGGGAAACCCCGGAGATTGTCAGACTCCCCAAAACGACCTTTCCTCTGATATAGSCCTCGCTGAGGGTSCCGT
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
 CGTGGCAAGCTCG
 GCAACCGTTCGAGC
 730

FIG. 66 - Exemplary Adaptor protein construct - Adaptor13.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCATAATAGGAGACCAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGCGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTGTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGGATGTAATCAAGCGGACGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAAGTGGACTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCGCTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGTGGTGGTCCGTGCGCAAAATGATCGCCCGATTCTGGATGAAATCGCTGAGCAATATCAGGGCAAACTGACCTTGCATGTGAATCGATCGATCAAAA
 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGGCAACCGTACGACTGTAGCTAGTATTTT
 370 380 390 400 410 420 430 440 450 460 470 480
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K C G A L S K G Q L K E F L
 90 100 110 120
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAATGCGGTGCACTGTCTAAAGGTGAGTTGAAAGAGTTCCCT
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GGGACCGTGAACCGGCTTTATACCGTAGGCACCATAGGSGCTGAGACGACACAAGTTTGTGCGCACTTCCACCGCGTTGGTTACGCCACGTGACAGATTTCACGTCACCTTTCTCAAGGA
 490 500 510 520 530 540 550 560 570 580 590 600
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGAGCCATCATCATCATATTAATGAAAGGCGGATATCCAGCACACTGSGCGCGCTTACTAGTGGATCCGGTCTCAACAAAGCCCGAAAGGAA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGAGTAGTAGTAGTAATTAATTACTTTCCCGCTAAGGTCGTTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTGTTTCGGGCTTTCTTT
 610 620 630 640 650 660 670 680 690 700 710 720
 GCTGAGTTGCTGCTGCCACCGCTGAGCAATAACTAGCATACCCCTTGGGCGCTTAACGGGTCTTGAAGGGTTTTTGTGAAAGGAGGAATAATATCCGAGCGCACTCCACCGCA
 730
 CGTGGGCAAGCTCG
 GCAACCGTTCGAGC

FIG. 68 - Exemplary Adaptor protein construct - Adaptor15.0

CGGTACGCTGGGACTGCATAGGCTGGCCCGGTGATCGGGCCACGATCGCTCCGCGGTAGAGGATCGAGATCTCGATCCCGCAATTAATACGACTCACTATAGGAGACCCACAAC
 10 20 30 40 50 60 70 80 90 100 110 120
 GCCAGTSCGAACCCGTGACGGTATCCGACCGGCCACTACGCGCGGTGCTACGCGGCCCATCTCCTAGCTCTAGAGTGGCGCTTTAATTATGCTGAGTGATATCCCTCTGCTGTTG
 GGTTCCTCTAGAAAATAATTTGTGTTTAACTTTAAGAAAGGAGATATACCATGAGCGCGCAAAATATTACCTGACTGACGACAGTTTGGACACGGATGTAATCAAGCGGACCGGGGGGAT
 20 30 40 50 60 70 80 90 100 110 120
 CCAAAGGGAGATCTTTATTAACAAATGAAATCTTCCTCTATATGGTACTCGCGCTTTTAAATAGAGGAGCTGACTGCTGTCAAAGCTGTGCTACATGAGTTTCGCTGCCCCCGTA
 130 140 150 160 170 180 190 200 210 220 230 240
 M S A K I I H L T D S F D T V L K A D G A I
 L V A F W A E W C G P C K M I A P I L D E I A D E Y Q G K L T V A M L N I D Q N
 30 40 50 60
 CCTCGTCGCTTCTGGGCGAGAGTGGTGGTCCGTGCAAAATGATCGCCCGATTCTGGGATGAATCGCTGAGCAATATCAGGGCAAACTGACCGTTGCAATGCTGAACATCGATCAAAA
 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250
 GGAGCAGCGCAAGACCCGCTCACCACGCCAGGCACCTTTTACTAGCGGGCTAAGACCTACTTTAGGCACTGCTTATAGTCCCGTTGACTGSCAACCGTTACGACTGTAGCTAGTATTT
 260 270 280 290 300 310 320 330 340 350 360
 P G T A P K Y G I R G I P T L L L F K N G E V A A T K V G A L S K G C L K E F L
 70 80 90 100
 CCCTGGCACATGCGCGGAATATGSCATCCCGTGTATCCCGTCTGCTGCTGTGTTCAAAAACGGTGAATGGCGGCAACCAAGTGGTGCACCTGTCTAAAGGTGTGCTTGAAGAGTTCCCT
 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
 GSGACCGTSGACCGGCTTTATACCGTAGGCACCATAGGSGCTGAGACGACACAAGTTTGTGCTCCACTTCCACCGCGTTGCTTCCACCGTGCACAGATTTCCACACGAACCTTTCCTCAAGGA
 D A N L A G G S H H H H H * * *
 110 120
 CGACGTAACCTGGCGGCGGCGACCATCATCATCATATATGAAAGGGCGATATCCAGCACATGSGCGCGCTTACTAGTGAATCCGGTCTCTAACAAAGCCCGAAAGGAA
 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
 SCTGGGATTSAGACCGCCCGCTCGGSHAGTAGTAGTAGTAATATTACTTTCCCGCTAGGGTCGTGTGACCGCGCGCAATGATCACCCTAGGCGCCGAGATTTGTTTCGGGCTTTTCCTT
 610 620 630 640 650 660 670 680 690 700 710 720 730
 GCTGAGTTGCTGTGCTGCTGCAATTAACCTAGCATACCCCTTGGGCGCTCTAAACGGGTCTTTGAGGGTTTGTGTGAAGAGGAACTATATCCGAGCGCACTCCCAAGCA
 CGACTCAACCGACGAGCGGTGGGACTCTGTTATTTGATCGTATTTGGGAAACCCCGGAGATTGTGCCAGACTCCCCAAGAAACGACTTTCTCTTATATAGGCTCGCTGAGGGTGGCGT
 CGTGGCAAGCTCG
 GCAACCGTTGAGC

