

We claim

1. A cage protein comprising a helical bundle, wherein the cage protein comprises a structural region and a latch region, wherein the latch region comprises one or more target binding polypeptide, wherein the cage protein further comprises a first reporter protein domain, wherein the first reporter protein domain undergoes a detectable change in reporting activity when bound to a second split reporter protein domain, and wherein the structural region interacts with the latch region to prevent solution access to the one or more target binding polypeptide.
2. The cage protein of claim 1, further comprising the second reporter protein domain, wherein one of the first reporter protein domain and the second reporter domain is present in the latch region and the other is present in the structural region, wherein an interaction of the first reporter protein domain and the second reporter protein domain is diminished in the presence of target to which the one or more target binding polypeptide binds.
3. The cage protein of claim 1, wherein the second reporter protein domain is not present in the cage protein.
4. The cage protein of any one of claims 1-3, wherein the helical bundle comprises between 2-9, 2-8, 2-7, 3-9, 3-8, 3-7, 4-9, 4-8, 4-7, 5-9, 5-8, 5-7, 6-9, 6-8, 6-7, 2-6, 3-6, 4-6, 5-6, 2-5, 3-5, 4-5, 2-4, 3-4, 2-3, 2, 3, 4, 5, 6, 7, 8, or 9 alpha helices.
5. The cage protein of any one of claims 1-4, wherein each helix in the structural region is independently between 18-60, 18-55, 18-50, 18-45, 22-60, 22-55, 22-50, 22-45, 25-60, 25-55, 25-50, 25-45, 28-60, 28-55, 28-50, 28-45, 32-60, 32-55, 32-50, 32-45, 35-60, 35-55, 35-50, 35-45, 38-60, 38-55, 38-50, 38-45, 40-60, 40-58, 40-55, 40-50, or 40-45 amino acids in length.
6. The cage protein of any one of claims 1-5, comprising amino acid linkers connecting adjacent alpha helices.
7. The cage protein of claim 6, wherein the amino acid linkers are independently between 2 and 10 amino acids in length, not including any further functional sequences that may be fused to the linker, or independently 3-10, 4-10, 5-10, 6-10, 7-10, 8-10, 9-10, 2-9, 3-

9, 4-9, 5-9, 6-9, 7-9, 8-9, 2-8, 3-8, 4-8, 5-8, 6-8, 7-8, 2-7, 3-7, 4-7, 5-7, 6-7, 2-6, 3-6, 4-6, 5-6, 2-5, 3-5, 4-5, 2-4, 3-4, 2-3, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids in length.

8. The cage protein of any one of claims 1-7, wherein the latch region is at the C-terminus of the cage protein. In other embodiments, the latch region may be at the N-terminus of the cage protein

9. The cage protein of any one of claims 1-8, wherein the first reporter protein domain is present in the latch region.

10

10. The cage protein of claim 9, wherein the second reporter protein, when present, is present in the structural region.

11. The cage protein of claim 10, wherein the second reporter protein is at the N-terminus of the structural region, or is within 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the N-terminus of the structural region.

12. The cage protein of claim 9, wherein the one or more target binding polypeptide and the first reporter protein domain are separated by at least 10 amino acids in the latch region.

20

13. The cage protein of any one of claims 1-12, wherein the one or more target binding polypeptide is at or within 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the C-terminus of the latch region.

14. The cage protein of any one of claims 1-12, wherein the first reporter protein domain is at the C-terminus of the latch region or within 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the C-terminus of the latch region, or wherein the first reporter protein domain is at the N-terminus of the latch region or within 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the N-terminus of the latch region.

30

15. The cage protein of any one of claims 1-14, wherein the first reporter protein domain, and the second reporter domain when present, comprise reporter protein domains selected from the group consisting of luciferase (including but not limited to firefly, Renilla, and

- Gaussia luciferase), bioluminescence resonance energy transfer (BRET) reporters, bimolecular fluorescence complementation (BiFC) reporters, fluorescence resonance energy transfer (FRET) reporters, colorimetry reporters (including but not limited to β -lactamase, β -galactosidase, and horseradish peroxidase), cell survival reporters (including but not limited to dihydrofolate reductase), electrochemical reporters (including but not limited to APEX2), radioactive reporters (including but not limited to thymidine kinase), and molecular barcode reporters (including but not limited to TEV protease).
16. The cage protein of any one of claims 1-15 wherein the latch region is at the C-terminus of the cage protein.
17. The cage protein of any one of claims 1-15, wherein the latch region is at the N-terminus of the cage protein.
18. The cage protein of any one of claims 1-8, wherein the first reporter protein domain is present in the latch region.
19. The cage protein of claim 18, wherein the second reporter protein, when present, is present in the structural region.
20. The cage protein of claim 19, wherein the second reporter protein is at the N-terminus of the structural region, or is within 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the N-terminus of the structural region.
21. The cage protein of claim 18, wherein the one or more target binding polypeptide and the first reporter protein domain are separated by at least 10 amino acids in the latch region.
22. The cage protein of any one of claims 1-21, wherein the one or more target binding polypeptide is at or within 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the C-terminus of the latch region.
23. The cage protein of any one of claims 1-21, wherein the first reporter protein domain is at the C-terminus of the latch region.

24. The cage protein of any one of claims 1-23, wherein the first reporter protein domain, and the second reporter domain when present, comprise a split reporter protein domain from a reporter selected from the group consisting of luciferase (including but not limited to
 5 firefly, Renilla, and Gaussia luciferase), bimolecular fluorescence complementation (BiFC) reporters, colorimetry reporters (including but not limited to β -lactamase, β -galactosidase, and horseradish peroxidase), cell survival reporters (including but not limited to dihydrofolate reductase), electrochemical reporters (including but not limited to APEX2), radioactive reporters (including but not limited to thymidine kinase), and molecular barcode
 10 reporters (including but not limited to TEV protease).

25. The cage protein of any one of claims 1-24, wherein the cage does not include the second split reporter domain, wherein the first split reporter protein domain comprises:

(a) an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%,
 15 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27359 and 27664-27672;

(b) an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS: 27360-27361:

20 VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRP EERFPMSTF KVLICGAVLS
 RVDAGQEQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTII
 GGPKELTAFI HNMGDHVTRL DRWEPELNEA IPNDERDTT PAAMATTLRK LLTGNGR
 (split β -lactamase A; SEQ ID NO:27360) and

25 LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
 GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW (Split beta lactamase B; SEQ ID
 NO:27361);

(c) an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
 30 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS:27362-27378, wherein underlined residues are amino acid linkers or other optional residues that may be present or absent, and when present may be any amino acid sequence, and wherein any N-terminal methionine residues may be present or absent:

35 VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSENGLKIDIHVIIPYEGLSG
 DQMGQIEKIFKVVYPVDNHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGN
 KIIDERLINPDGSLFLFRVTINGVTGWRLLHERILA (TeLuc; SEQ ID NO:27362) (full
 luminescent or fluorescent protein that can be used to create FRET and/or BRET sensors);

- LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
5 PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSLNKGMD ELYK (CyOFF variant; SEQ ID
NO: 27363) (full luminescent or fluorescent protein that can be used to create FRET and/or
BRET sensors);
- VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
10 KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSLNKGMD ELYK (CyOFF variant SEQ ID
NO: 27364) (full luminescent or fluorescent protein that can be used to create FRET and/or
BRET sensors);
- EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
15 PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSLNKGMD ELYK (CuOFF variant; SEQ ID
20 NO: 27365) (full luminescent or fluorescent protein that can be used to create FRET and/or
BRET sensors);
- KVFTLGDFVGDWRQTAGYNQAQVLEQGGTSLFQNLGVSVTPIQRIVLSGENGLKIDIHV
IIPYEGLSQDMAQIEKIFKVVPVDDHHFKAILHYGTLVIDGVTNMDYFGQPYEGIA
25 KFDGKKITVTGTLWNGNTIIDERLINPDGSLLFRTVINGVTGWRLHERILA (LumiLuc; SEQ ID
NO: 27366) (full luminescent or fluorescent protein that can be used to create FRET and/or
BRET sensors);
- MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGQGTGNP NDGYEELNLK STKGDLQFSP W
ILVPHIGYG FHQYLPYPDG MSPFQAAMVD GSGYQVHRM QFEDGASLTV NYRYTYEGSH IKG
30 EAQVKG T GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTY
TFAKP MAANYLKNQF MYVFRKTELK HSKTELNFKE WQKAFTDVMG MDLYK
(mNeonGreen; SEQ ID NO: 27367) (full luminescent or fluorescent protein that can
be used to create FRET and/or BRET sensors);
- MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
35 ILSPQFMYGS RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGGAVTVTQ DTSLEDGTL
I YKVKLRGTNF PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRLADP
KT TYKAKKPVQM PGAYNVDRKL DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK
(mScarlet-i; SEQ ID NO: 27368) (full luminescent or fluorescent protein that can be
used to create FRET and/or BRET sensors);
- SGKSYPTVSADYQKAVEKAKKRLGGFIAEKRCAPLMRLAWHSAGTFDKRTKTGGPFGTIRYPALAH
40 SANSGLDIAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKPELPPEGRLPDATK
GSDHLRDVFGKAMGLTDQDIVALSGGHTLGAHKERSGFEGPWTSNPLVFDNSYFTELLSGEKEGGG
SGGGGS (APEX2-1-200; SEQ ID NO: 27369);
- GGGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFADA
45 (APEX2-201-250; SEQ ID NO: 27370);
- MGSHHHHHHSGSENLYFQSGGS

- VRPLNCIVA VSQNMIGIGN GDLWPPLRN ESKYFQRMFT TSSVEGKQNL
VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRGHFL AKSLDDALRL
IEQPELGGGGSGGGGS (DHFR_A (1-105); SEQ ID NO:27371);
- 5 SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRE
GGGGSGGGGS ASKV DMVWIVGGSS VYQAMNQPG HLRLFVTRIM QEFESDTFFP
EIDLKGYKLL PEYPGVLSEV QEEKGIKYKF EVYEKKD (DHFR_B (106-186); SEQ ID
NO:27372);
- 10 QLTPTFYDNSCPNVSNIVRDIIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAF
GNANSARGFSVIDR
MKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFTLPLQKLD
SFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGSGGS
(sHRPa is the large split HRP fragment. It consists of amino acids
1-213 of horseradish peroxidase (HRP) with the following 4
15 mutations: T21I, P78S, R93G, N175S) (SEQ ID NO:27373);
- NLSALVDFDLRTPTIFDNKYVNLLEEQGLIQSDQELFSSPDATDTIPLVRSEFANSTQTFNFAFVEAM
DRMGNTPLTGTQGQIRNCRVNSNGSGGS (sHRPb is the small split HRP
fragment. It consists of amino acids 214-308 of horseradish
peroxidase (HRP) with the following 2 mutations: N255D, L299R) (SEQ
20 ID NO:27374);
- GESLFKGRDYNPISSTICHLTNESDGHSTSLYGIGFGPFITNKHLFRNNGTLLVQSLHGVFKVKN
TTTLQQLHLDGRDMIIIRMPKDFPPFPQKLKFREPQREERICLVTFNQTTGGGGSGGGGS (N_Tev
(1-118) (SEQ ID NO:27375);
- 25 GGGGSGGGGSKSMSSMVSDTSCTFPSSDGIFWKHWIQTGDGQCGSPLVSTRDGFIVGIHSASNFNTNTN
NYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMDKP C_Tev (119-221) (SEQ
ID NO:27376);
- 30 MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQGEATEV RLEQKMPDLL
RVYIDGPHGM GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV LGASETIANI
YTTQHRDLQD EISAGDAVV MTSAQITMGM PYAVTDAVLA PHIGGEAGSS
HAPPPALTLI FDRHPAALL CYPAAARYLMG SMTTPQAVLAF VALIPPTLPG
TNIVLGALPE DRHIDRLAKR QRPGERDLA MLAAIRRVYG LLANTVRYLQ
GGGSWREDWG QLSGT GGGGSGGGGS (thymidine kinase_TK_A (1-265) (SEQ ID
NO: 27377); and/or
- 35 GGGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP
NGDLYNVFAW ALDVLAKRLR PMHVFILDYD QSPAGCRDAL LQLTSGMVQT
HVTTPGSIPT ICDLARTFAR EMGEAN (thymidine kinase_TK_B (266-376) (SEQ
ID NO: 27378).
26. The cage protein of any one of claims 1-24, wherein the cage comprises the second
40 split reporter protein domain, wherein
- (a) one of the first reporter protein domain and the second reporter protein
domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%,

90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NOS: 27359, and 27664-27672;

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27379, wherein the N-terminal methionine residue may be present or absent:
 MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRIVRSGENALKIDIHIIIPYEGLSADQMAQIE
 EVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSM
 LFRVTINS (LgBit) (SEQ ID NO: 27379);

(b) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27360

VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRP EERFPMSTF KVLCCGAVLS
 RVDAGQEQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTIT
 GGPKELTAF LHMMDHVT RLDRWEPELNEA IPNDERDTTT PAAMATTLRK LLTGNGR (split β -
 lactamase A) (SEQ ID NO: 27360),

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27361:
 LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD GKPSRIVVIY
 TTGSQATMDE RNRQIAEIGA SLIKHW (Split beta lactamase B) (SEQ ID NO: 27361);

(c) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27362:

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSG
 DQMGQIEKIFKVVPVDNHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGN
 KIIDERLINPDGSL LFRVTINGVTGWRLHERILA (TeLuc) (SEQ ID NO: 27362), (full
 luminescent or fluorescent protein that can be used to create FRET and/or BRET sensors)

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS: 27363-27365:

- LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
 KVFIIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
 PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
 PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CyOFP variant) (SEQ ID
 5 NO: 27363) (full luminescent or fluorescent protein that can be used to create FRET
 and/or BRET sensors);
- VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
 KVFIIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
 10 PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
 PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK(CyOFP variant) (SEQ ID
 NO: 27364) (full luminescent or fluorescent protein that can be used to create FRET
 and/or BRET sensors); and
- 15 EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
 KVFIIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSIQDGELI YNVKVRGVNF
 PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
 PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK(CyOFP variant) (SEQ ID
 NO: 27365) (full luminescent or fluorescent protein that can be used to create FRET
 20 and/or BRET sensors);
- (d) one of the first reporter protein domain and the second reporter protein domain
 comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:
 25 27366:
 KVFTLGDFVGDWRQTAGYNQAQVLEQGGLTSLFQNLGVSVTPIQRIVLSGENGLKIDIHV
 IIPYEGLSCDQMAQIEKIFKVVPVDDHHFKAILHYGTLVIDGVTPNMIDYFGQPYEGIA
 KFDGKKITVTGTLWNGNTIIDERLINPDGSLFRVTVINGVTGWRLHERILA (LemiLuc) (SEQ ID
 NO: 27366) (full luminescent or fluorescent protein that can be used to create FRET and/or
 30 BRET sensors),
 and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%,
 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence
 of SEQ ID NO: 27368, wherein the N-terminal methionine residue may be present or absent:
 MVSKEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD ILSPQFMYG
 35 S RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGGAVTVTQ DTSLEDGTLI YVKLGRGNF PPDGPVM
 QKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFET TYKAKKPVQM PGAYNVDRKL DITSH

NEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i) (SEQ ID NO:27368) (full

luminescent or fluorescent protein that can be used to create FRET and/or BRET sensors);

- (e) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27367, wherein the N-terminal methionine residue may be present or absent:

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGQGTGNP NDGYEELNLK STKGDQLQFSP WILVPHIGY
G FHQYLPYPDG MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT GFPADGP
VMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP MYVFR

- 10 KTELK HSKTELNFKE WQKAFTDVMG MDELYK (mNeonGreen) (SEQ ID NO:27367), (full
luminescent or fluorescent protein that can be used to create FRET and/or BRET sensors),

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27368, wherein the N-terminal methionine residue may be present or absent:

- 15 MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD ILSPQFMYG
S RAFIKHPADI PDYYKQSFPE GFKWERVMNF EDGGAIVTVTQ DTSLEDGTLI YKVKLRGTNF PPDGPVM
QKK TMGWEASTER LYPEDGV LKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL DITSH
NEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i) (SEQ ID NO: 27368) (full

luminescent or fluorescent protein that can be used to create FRET and/or BRET sensors);

- 20 (f) one of the first reporter protein domain and the second reporter protein domain
comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence SEQ ID NO:
27369, wherein underlined residues are optional residues that may be present or absent, and
when present may be any amino acid sequence

- 25 SGKSYPTVSADYQKAVEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTFDKRTKTGGPFGTIRYPAELAHSA NSGLD
IAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKPELPPEGRLPDATKGS DHLRDVFGKAMG
LTDQDIVALSGGHTLGAAHKERSGFEGPWTSNPLVFNDNSYFTELLSGEKEGGGSGGGGS (APEX2-1-200)

(SEQ ID NO: 27369) (split engineered variant of soybean ascorbate peroxidase protein
for chemiluminescent and colorimetric detection system);

- 30 and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%,
91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence
of SEQ ID NO: 27370, wherein underlined residues are optional residues that may be present
or absent, and when present may be any amino acid sequence

GGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFADA (APEX2-201-250) (SEQ ID NO: 27370) (split engineered variant of soybean ascorbate peroxidase protein for chemiluminescent and colorimetric detection system);

- (g) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27371, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

MGSHHHHHHSGSGSENLYFQSGSGGS
 10 VRPLNCIVA VSQNMIGIGN GDLWPPLRN ESKYFQRM TT TSSVEGKQNL
VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRGHFL AKSLDDALRL
IEQPELGGGSGGGGS (DHFR_A (1-105)); (SEQ ID NO: 27371) (split dihydrofolate reductase protein reporter for cell survival or fluorescence)

- and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27372, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

SGSG DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAEERLIREAAAASEKISRE
GGGSGGGGS ASKV DMVWIVGGSS VYQEAMNQPGLRLFVTRIM QEFESDTFFP
 20 EIDLKGYKLL PEYPGVLSEV QEEKGIKYKF EVYEKDD (DHFR_B (106-186)); (SEQ ID NO: 27372) (split dihydrofolate reductase protein reporter for cell survival or fluorescence);

- (h) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27373, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

QLTPTFYDNSCPNVSNIVRDIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAFGNANSA
RGFSVIDRMKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFTLPLQK
DSFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGGSGS (sHRP
 30 is the large split HRP fragment. It consists of amino acids 1-213 of horseradish peroxidase (HRP) with the following 4 mutations: T21I, P78S, R93G, N175S; plasmid 73147 (SEQ ID NO: 27373);

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence

of SEQ ID NO:27374, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

NLSALVDFDLRTPITFDNKYYVNL EEQKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAMDRMGNIT
PLTGTQGQIRRNCRVNSNGGSGS (sHRPb is the small split HRP fragment. It

5 consists of amino acids 214-308 of horseradish peroxidase (HRP) with the following 2 mutations: N255D, L299R: plasmid 73148) (SEQ ID NO: 27374);

(i) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:

10 27375, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

GESLFKGPRDYNPISSITICHLTNESDGHTTSLYGIGFGFPFIITNKHLFRRNNGTLLVQSLHGVFKVKNTTTLQQH
LIDGRDMIIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTGGGSGGGGSN_Tev (1-118) (SEQ
ID NO: 27375) (Split TEV protease);

15 and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27376, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

20 GGGGSGGGGSKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHSASNFNTNNYFTSV
PKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMDKP (C_Tev (119-221)); (SEQ ID NO:
27376) (Split TEV protease);

(j) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:

25 27377, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and wherein the N-terminal methionine residue may be present or absent:

MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQQEATEV RLEQKMPTLL
RVYIDGPHGM GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV LGASETIANI
30 YTTQHRLDQG EISAGDAVV MTSAQITMGM PYAVTDAVLA PHIGGEAGSS
HAPPPALTLL FDRHPAAALL CYPAAARYLMG SMTPQAVLAF VALIPPTLPG
TNIVLGALPE DRHIDRLAKR QRPGERLDLA MAAIRRVYG LLANTVRYLQ
GGGSWREDWG QLSGT GGGGSGGGGS (thymidine kinase_TK_A (1-265)) (SEQ ID NO:
27377);

and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27378, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence

5 GGGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP
 NGDLYNVFAW ALDVLAKRLR PMHVFILDYD QSPAGCRDAL LQLTSGMVQT
 HVTTPGSIPT ICDLARTFAR EMGEAN (thymidine kinase_TK_B (266-376) (SEQ ID NO:
 27378) .

10 27. The cage protein of any one of claims 1-26, wherein the one or more target binding polypeptide is capable of binding to a target including but not limited to an antibody, a toxin, a diagnostic biomarker, a viral particle, a disease biomarker, a metabolite or a biochemical analyte.

15 28. The cage protein of any one of claims 1-26, wherein the one or more target binding polypeptide is capable of binding to an antibody target.

29. The cage protein of claim 28, wherein the one or more target binding polypeptide comprises one or more epitope recognized by antibodies against a viral target.

20

30. The cage protein of claim 28 or 29, wherein the one or more target binding polypeptide comprises one or more epitope recognized by antibodies against SARS-Cov-2.

31. The cage protein of any one of claims 1-26, wherein the one or more target binding
 25 polypeptide is capable of binding to a disease marker or toxin.

32. The cage protein of any one of claims 1-26, wherein the one or more target binding polypeptide is capable of binding to Bcl-2, Her2 receptor, Botulinum neurotoxin B, albumin, epithelial growth factor receptor, prostate-specific membrane antigen (PSMA), citrullinated
 30 peptides, brain natriuretic peptides, and/or cardiac Troponin I.

33. The cage protein of any one of claims 1-32, wherein the one or more target binding polypeptide comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%,

80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS: 27380-27430.

34. The cage protein of claim 33, wherein the one or more target binding polypeptide
5 comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%,
90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid
sequence selected from the group consisting of SEQ ID NOS:27397-27430, or selected from
SEQ ID NOS:27397-27406, 27409-27416, and 27427-27430.

10 35. The cage protein of claim 34, wherein amino acid substitutions relative to the
reference target binding polypeptide amino acid are selected from the allowable amino acid
substitutions provided in Table 3

15 36. The cage protein of claim 34 or 35, wherein interface residues are identical to those in
the reference target binding polypeptide or are conservatively substituted relative to interface
residues in the reference target binding polypeptide as detailed in Table 2.

37. The cage protein of any one of claims 34-36, wherein the one or more target binding
polypeptide comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%,
20 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the
amino acid sequence selected from the group consisting of SEQ ID NOS: 27397-27406 and
27431-27466.

25 38. The cage protein of claim 37, wherein the one or more target binding polypeptide
comprises an amino acid substitution relative to the amino acid sequence of SEQ ID
NO:27397 at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or all 18 residues selected
from the group consisting of 2, 4, 5, 14, 15, 17, 18, 27, 28, 32, 37, 38, 39, 41, 42, 49, 52, and
55.

30 39. The cage protein of claim 38, wherein the substitutions in the one or more target
binding polypeptide are selected from the substitutions listed in Table 5, either individually or
in combinations in a given row.

40 The cage protein of any one of claims 34-36, wherein the one or more target binding polypeptide comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS: 27409-27416 and
5 27467-27493.

41. The cage protein of claim 40, wherein the target binding comprises an amino acid substitution relative to the amino acid sequence of SEQ ID NO:27409 at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or all 20 residues selected from the group consisting
10 2, 6, 8, 9, 13, 14, 19, 22, 25, 26, 28, 29, 34, 35, 37, 40, 43, 45, 49, and 62.

42. The cage protein of claim 41, wherein the substitutions are selected from the substitutions listed in Table 7, either individually or in combinations in a given row.

15 43. The cage protein of any one of claims 34-36, wherein the target binding comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS: 27427-27430 and 27494.

20 44. The cage protein of claim 44, wherein the one or more target binding polypeptide comprises an amino acid substitution relative to the amino acid sequence of SEQ ID NO:27429 at or both residues selected from the group consisting 63 and 75.

45. The cage protein of claim 44, wherein the substitutions comprise R63A and/or K75T.

25 46. The cage protein of any one of claims 1-45, comprising the amino acid sequence at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid sequence of a cage protein selected from the group consisting of SEQ ID NOS:1-49, 51-52, 54-59, 61, 65, 67-91, 92 -
30 2033, 2034-14317, 27094-27117, 27120-27125, 27,278 to 27,321, even-numbered SEQ ID NOS between SEQ ID NOS: 27126 and 27276, and cage proteins listed in Tables 8 and 9, not including optional amino acid residues (, and not including amino acid residues in the latch region, and wherein the N-terminal and/or C-terminal 60 amino acids of each cage protein may be optional

47. The cage protein of any one of claims 1-47, comprising the amino acid sequence at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, not including optional amino acid residues, to the amino acid sequence of a cage protein selected from the consisting of SEQ ID NOS: 27497-27620, wherein the N-terminal protein purification tag (*MGSHHHHHHGSSENLYFQSGG* (SEQ ID NO:27624); or *MGSHHHHHHGSSENLYFQG* (SEQ ID NO:27625); or *GSHHHHHHGSSENLYFQG* (SEQ ID NO:27626)) is optional, are not considered in the percent identity comparison, and can be present or absent and preferably are absent.
48. A key protein capable of binding to the structural region of a cage protein of any one of claims 1-47 that does not include the second reporter protein domain, wherein binding of the key protein to the cage protein only occurs in the presence of a target to which the cage protein one or more target binding polypeptide can bind, wherein the key protein comprises a second reporter protein domain, wherein interaction of the key protein second reporter protein domain and the cage protein first reporter protein domain causes a detectable change in reporting activity from the first reporter protein domain .
49. The key protein of claim 48, wherein the second reporter protein domain is at the N-terminus or the C-terminus of the key protein, or is within 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid of the N-terminus or the C-terminus of the key protein.
50. The key protein of any one of claims 48-49, wherein the second reporter protein domain comprises a reporter protein domain selected from the group consisting of luciferase (including but not limited to firefly, Renilla, and Gaussia luciferase), bioluminescence resonance energy transfer (BRET) reporters, bimolecular fluorescence complementation (BiFC) reporters, fluorescence resonance energy transfer (FRET) reporters, colorimetry reporters (including but not limited to β -lactamase, β -galactosidase, and horseradish peroxidase), cell survival reporters (including but not limited to dihydrofolate reductase), electrochemical reporters (including but not limited to APEX2), radioactive reporters (including but not limited to thymidine kinase), and molecular barcode reporters (including but not limited to TEV protease).

51. The key protein of any one of claims 48-50, wherein the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NOS:27360-23379, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and wherein any N-terminal methionine residue may be present or absent.

52. The key protein of any one of claims 48-51, wherein the key protein, not including the second split reporter protein domain, comprises an amino acid sequence at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, not including optional amino acid residues, to the amino acid sequence selected from the group consisting of SEQ ID NOS:14318-26601, 26602-27015, 27016-27050, 27,322 to 27,358, and key polypeptides with an odd-numbered SEQ ID NO between SEQ ID NOS: 27127 and 27277), key proteins in Table 8, and key proteins in Table 9.

53. The key protein of any one of claims 48-52, wherein the key protein comprises an amino acid sequence at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, not including optional amino acid residues in parentheses, to the amino acid sequence of a key protein selected from the group consisting of SEQ ID NOS: 27621-27623, wherein residues in parentheses are optional and may be present or absent.

> lucKey: MGS-(His)6-TEV site-linker-LgBit-linker-latch sequence
(MGSHHHHHHSGSGSENLYFQG) SGMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQIRIVRSGE
NALKIDIHIIPYEGLSADQMAQIEEVFKVVPVDDHHFKVILPYGTLVIDGVTNMLNYFGRPYEGIAVFDGKKI
TVTGTLWNGNKIIDERLITPDGSMFLFRVTINSGGSGGGSGGGSGGSDEARKAIARVKRESKRIVEDAERLIREA
AAASEKISREAERLIREAAAASEKISRE (SEQ ID NO:27621)

Key-2GGSGG-CyOFF (CyOFF sequence in bold/underline):
(M) DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISREGGSGG **GGVSK**
GEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLODGEI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGGMD ELYK (SEQ ID NO: 27622)

Key-LacB (split β -lactamase B in bold/underline):
SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRESGGGGSGGGSGG
GG **LLTLASRQQLIDWME ADKVAGPLL R SALPAGWFIA DKSGAGERGS RGIIAALGPD GKPSRIVVIY**
TTGSQATMDE RNRQIAEIGA SLIKHW (SEQ ID NO: 27623)

54. A biosensor, comprising

(a) the cage protein of any one of claims 1-47 wherein the cage does not include
5 the second reporter protein domain; and

(b) the key protein of any one of claims 48-53;
wherein the key protein can only bind to the cage protein in the presence of a target to
which the cage protein one or more target binding polypeptide can bind; and wherein binding
of the first reporter protein domain of the cage protein to the second reporter protein domain
10 of the key protein causes a detectable change in reporting activity from the first reporter
protein domain.

55. The biosensor of claim 54, wherein

(a) the first reporter protein domain comprises an amino acid sequence at least
15 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or
100% identical to the amino acid sequence SEQ ID NO: 27359, and 27664-27672;
and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%,
80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the
amino acid sequence of SEQ ID NO 27379, wherein the N-terminal methionine residue may
20 be present or absent

(b) one of the first reporter protein domain and the second reporter protein domain
comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID
NO:27360, and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%,
25 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid
sequence of SEQ ID NO: 27361;

(c) one of the first reporter protein domain and the second reporter protein domain
comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,
94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID
30 NO:27362, and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%,
90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid
sequence selected from the group consisting of SEQ ID NOS:27363-27365;

(d) one of the first reporter protein domain and the second reporter protein domain
comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27366, and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO 27368:

- 5 (e) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27367, wherein the N-terminal methionine residue may be present or absent, and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 10 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO 27368, wherein the N-terminal methionine residue may be present or absent;

- (f) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27369, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence; and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid of SEQ ID NO:27370, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid 15 sequence;

- (g) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27371, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and the other comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO: 27372, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence; 20

- (h) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27373, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and the other comprises an amino acid 25

sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27374, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence;

5 (i) one of the first reporter protein domain and the second reporter protein domain comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27375, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and the other comprises an amino acid
10 sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27376, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence;

(j) one of the first reporter protein domain and the second reporter protein domain
15 comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27377, wherein the N-terminal methionine residue may be present or absent, and wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence, and the other comprises an amino acid sequence at
20 least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27378, wherein underlined residues are optional residues that may be present or absent, and when present may be any amino acid sequence.

25 56. The biosensor of claim 54, wherein the cage protein comprises a cage protein of claim 47 and the key protein comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical, not including optional amino acid residues in parentheses, to the amino acid sequence of SEQ ID NO: 27621.

30 > lucKey: MGS(His)6-TEV site-linker-LgBit-linker-latch sequence
(MGSHHHHHGSGSENLYFQG) SGMVFTLEDFVGDWETAAYNLDQVLEQGGVSSLLQNLAVSVTPIQIRIVRSGE
NALKIDIHIIPYEGLSADQMAQIEEVFKVVPVDDHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKI
TVTGTLWNGNKIIDERLITPDGSMFLFRVTINSGGSGGGSGGGSGGSDEARKAIARVKRESKRIVEDAERLIREA
AAASEKISREAERLIREAAAASEKISRE

35

57. The biosensor of claim 54, wherein the cage protein and the key protein comprise a protein pair comprising

- (i) a cage protein comprising an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27620, wherein the residues in parentheses are optional and may be present or absent:

LacATrop (split β -lactamase A in bold; underline cTnT and cTnC):

(MGSHHHHHGSGSENLYFQG SGGS)**VFAHPETLVK VKDAEDQLGA RVGYIELDLN**

10 **SGKILESFRP EERFPMSTF KVLCCGAVLS RVDAGQEQLG RRIHYSQNDL**

VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL

HNMGDHVTRL DRWEPELNEA IPNDERDTT PAAMATTLRK LLTGNGR

SGGGSGGGSGGGG (SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIARLQELNLELVYLAVELT

DPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAIRDAAESRKILEEGSGSGSDALDELQKLN

15 ELAKLLKAIETQDLNLRRAAKAFLEAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA

AKRESERIEEARLIEKAKEESERIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALA

QIQELNLDLLRLASEL) TDPDEARKAIAVTGYRLFEEILDAERLIREAAAASEDQLREAAKELWQTIYNLEAEKF

DLQEKFKQOKYEINVLRNRINDNQVSKTKDDSKGSEELSDLFRMFDKNADGYIDLEELKIMLQATGETITED

DIEELMKDGDKNNDGRIDYDEFLEFMKGVE; and

- 20 (ii) a key protein comprising an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% identical to the amino acid sequence of SEQ ID NO:27631.

LLTLASRQQIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD GKPSRIVVIY

TTGSQATMDE RNRQIAEIGA SLIKHW (SEQ ID NO:27361)

25

58. A method for detecting a target, comprising

- (a) contacting the cage protein of any one of claims 47 where the cage protein comprises the second reporter protein domain, or the biosensor of any one of claims 40-43 with a biological sample under conditions to promote binding of the cage protein one or more target binding polypeptide to a target present in the biological sample, causing a detectable change in reporting activity from the first reporter protein domain; and

- (b) detecting the change in reporting activity from the first reporter protein domain, wherein the change in reporting activity identifies the sample as containing the target.

35

59. The method of claim 58, wherein the target is selected from the group including but not limited to an antibody, a toxin, a diagnostic biomarker, a viral particle, a disease biomarker, a metabolite or a biochemical analyte.
- 5 60. The method of any one of claims 58-59, wherein the target is an antibody.
61. The method of claim 60, wherein the target comprises antibodies selective for a virus.
62. The method of claim 58, wherein the one or more target binding polypeptide
10 comprises the amino acid sequence selected from the group consisting of SEQ ID NO: 27292-27394 and 27547-27548, and a polypeptide comprising an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid sequence selected from the group consisting of SEQ ID NO: 27397-27494.
- 15 63. The method of any one of claims 58-62, wherein the cage polypeptide comprises the cage protein of claim 47.
64. The method of any one of claims 58-59, wherein the target is a disease marker or
20 toxin.
65. The method of claim 64, wherein the disease marker or toxin comprises Bcl-2, Her2 receptor, Botulinum neurotoxin B, albumin, epithelial growth factor receptor, prostate-specific membrane antigen (PSMA), citrullinated peptides, brain natriuretic peptides, and/or
25 cardiac Troponin I.
66. The method of claim 64 or 65, wherein the one or more target binding polypeptide comprises an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the amino acid
30 sequence selected from the group consisting of SEQ ID NO: 27380-27390, wherein any N-terminal amino acid is optional and may be present or absent.
67. The method of any one of claims 64-66, wherein the cage protein comprises the cage
35 protein of claim 47.

68. A method for designing a biosensor, cage protein, or key protein comprising the steps of any method described herein.

5

69. A nucleic acid encoding the cage protein or key protein of any of the preceding claims.

70. An expression vector comprising the nucleic acid of claim 69 operatively linked to a suitable control element, such as a promoter.

10

71. A cell comprising the cage protein, key protein, composition, nucleic acid, or expression vector of any preceding claim.

72. A pharmaceutical composition comprising the cage protein, key protein, composition, nucleic acid, expression vector, or cell of any preceding claim, and a pharmaceutically acceptable carrier.

15

73. An epitope, comprising or consisting of the amino acid sequence of SEQ ID NO: 27384

20

74. A method for detecting Troponin I in a sample, comprising contacting a biological sample with the epitope of claim 73 under conditions suitable to promote binding of Troponin I in the sample to the epitope to form a binding complex, and detecting binding complexes that demonstrate presence of Troponin I in the sample.

25