



(12) 发明专利申请

(10) 申请公布号 CN 116368156 A

(43) 申请公布日 2023.06.30

(21) 申请号 202180049976.6

(22) 申请日 2021.05.25

(30) 优先权数据

63/030,836 2020.05.27 US

63/051,549 2020.07.14 US

63/067,643 2020.08.19 US

(85) PCT国际申请进入国家阶段日

2023.01.28

(86) PCT国际申请的申请数据

PCT/US2021/034104 2021.05.25

(87) PCT国际申请的公布数据

W02021/242780 EN 2021.12.02

(71) 申请人 华盛顿大学

地址 美国华盛顿

(72) 发明人 阿尔弗雷多·基亚诺·卢比奥

朴周永 页先伟 大卫·贝克

曹龙兴 布赖恩·考文垂

因娜·戈瑞斯尼克

丽莎·科佐多伊

兰斯·约瑟夫·斯图尔特

(74) 专利代理机构 北京高沃律师事务所 11569

专利代理师 王立普

(51) Int.Cl.

G07K 19/00 (2006.01)

权利要求书16页 说明书130页

序列表(电子公布) 附图18页

(54) 发明名称

基于从头设计的蛋白开关的模块化和通用化生物传感器平台

(57) 摘要

本公开提供了包括螺旋束的笼蛋白,其中所述笼蛋白包括结构区和门锁区,其中所述门锁区包括一种或多种靶标结合多肽,其中当所述第一报告蛋白结构域结合到第二拆分报告蛋白结构域时,报告活性发生可检测的变化,并且其中所述结构区与所述门锁区相互作用以防止溶液进入所述一种或多种靶标结合多肽。

1. 一种笼蛋白,包含螺旋束,其中所述笼蛋白包括结构区和闩锁区,其中所述闩锁区包含一种或多种靶标结合多肽,其中所述笼蛋白还包括第一报告蛋白结构域,其中当结合到第二拆分报告蛋白结构域时,所述第一报告蛋白结构域报告活性发生可检测的变化,并且其中所述结构区与所述闩锁区相互作用以防止溶液进入所述一种或多种靶标结合多肽。

2. 根据权利要求1所述的笼蛋白,其还包含第二报告蛋白结构域,其中所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个存在于所述闩锁区中,而另一个存在于所述结构区中,其中在所述一种或多种靶标结合多肽结合的靶标存在下,所述第一报告蛋白结构域和所述第二报告蛋白结构的相互作用减弱。

3. 根据权利要求1所述的笼蛋白,其中所述第二报告蛋白结构域不存在于所述笼蛋白中。

4. 根据权利要求1-3中任一项所述的笼蛋白,其中所述螺旋束包含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或9个 α 螺旋。

5. 根据权利要求1-4中任一项所述的笼蛋白,其中所述结构区中的每个螺旋长度独立地在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或40-45个氨基酸之间。

6. 根据权利要求1-5中任一项所述的笼蛋白,其包含连接相邻 α 螺旋的氨基酸接头。

7. 根据权利要求6所述的笼蛋白,其中所述氨基酸接头的长度独立地为2至10个氨基酸之间,不包括可与所述接头融合的任何其他功能序列,或长度独立地为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或10个氨基酸。

8. 根据权利要求1-7中任一项所述的笼蛋白,其中所述闩锁区位于所述笼蛋白的C端;在其他实施方案中,闩锁区可以位于笼蛋白的N端。

9. 根据权利要求1-8中任一项所述的笼蛋白,其中所述第一报告蛋白结构域存在于所述闩锁区中。

10. 根据权利要求9所述的笼蛋白,其中所述第二报告蛋白在存在时存在于所述结构区中。

11. 根据权利要求10所述的笼蛋白,其中所述第二报告蛋白位于所述结构区的N端,或位于所述结构区N端的10、9、8、7、6、5、4、3、2或1个氨基酸内。

12. 根据权利要求9所述的笼蛋白,其中所述一种或多种靶标结合多肽和所述第一报告蛋白结构域被所述闩锁区中的至少10个氨基酸分开。

13. 根据权利要求1-12中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽位于所述闩锁区C端的10、9、8、7、6、5、4、3、2或1个氨基酸处或其内。

14. 根据权利要求1-12中任一项所述的笼蛋白,其中所述第一报告蛋白结构域位于所述闩锁区的C端处或在所述闩锁区C端的20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸内,或其中所述第二报告蛋白结构域位于所述闩锁区N端或在闩锁区N端20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸内。

15. 根据权利要求1-14中任一项所述的笼蛋白,其中所述第一报告蛋白结构域和当存在时的所述第二报告蛋白结构域包括报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子、荧光共振能量转移(FRET)报告分子、比色报告分子(包括但不限于 β -内酰胺酶、 β -半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶)和分子条形码报告分子(包括但不限于TEV蛋白酶)。

16. 根据权利要求1-15中任一项所述的笼蛋白,其中所述锁区位于所述笼蛋白的C端。

17. 根据权利要求1-15中任一项所述的笼蛋白,其中所述锁区位于所述笼蛋白的N端。

18. 根据权利要求1-8中任一项所述的笼蛋白,其中所述第一报告蛋白结构域存在于所述锁区中。

19. 根据权利要求18所述的笼蛋白,其中所述第二报告蛋白在存在时存在于所述结构区中。

20. 根据权利要求19所述的笼蛋白,其中所述第二报告蛋白位于所述结构区的N端,或位于所述结构区N端的10、9、8、7、6、5、4、3、2或1个氨基酸内。

21. 根据权利要求18所述的笼蛋白,其中所述一种或多种靶标结合多肽和所述第一报告蛋白结构域被所述锁区中的至少10个氨基酸分开。

22. 根据权利要求1-21中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽位于所述锁区C端的10、9、8、7、6、5、4、3、2或1个氨基酸处或其内。

23. 根据权利要求1-21中任一项所述的笼蛋白,其中所述第一报告蛋白结构域位于所述锁区的C端。

24. 根据权利要求1-23中任一项所述的笼蛋白,其中所述第一报告蛋白结构域和当存在时的所述第二报告结构域包括来自报告蛋白的拆分报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、双分子荧光互补(BiFC)报告蛋白、比色报告分子(包括但不限于 β -内酰胺酶、 β -半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶),和分子条形码报告(包括但不限于TEV蛋白酶)。

25. 根据权利要求1-24中任一项所述的笼蛋白,其中所述笼不包括第二拆分报告蛋白结构域,其中第一拆分报告蛋白结构域包括:

(a) 与SEQ ID NO:27359和27664-27672的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

(b) 与选自SEQ ID NO:27360-27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRPEERFPMMSTF KVLLCGAVLS
RVDAGQEQLG RRIHYSQNDLVEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTIGPKELTAFL
HNMGDHVTRL DRWEPELNEA IPNDERDTTTPAAMATTLRK LLTGENGR (拆分 β -内酰胺酶A; SEQ ID

NO:27360) 和LLTLASRQQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGSRGIIAALGPD GKPSRIVVIYTTGSQATMDE RNRQIAEIGA SLIKHW(拆分 β -内酰胺酶B;SEQ ID NO:27361);

(c) 与选自SEQ ID NO:27362-27378的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中,带下划线的残基是氨基酸接头或可以存在或不存在的其他任选残基,且当存在时,可以是任何氨基酸序列,并且其中任何N端蛋氨酸残基可以存在或不存在:

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMG QIEKIFKVVYPVDNHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDG SLLFRVTINGVTGWRLHERILA(TeLuc;SEQ ID NO:27362)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVVVEGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPEGFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNFPANGPVMQKK TLGWEPSTETMYPADGGLEG RCDKALKLVGGGHLHVNFKT TYKSKKPVKM PGVHYVDRRL ERIKEADNETYVEQYEHAVA RYSNLGGMD ELYK(CyOFF变体;SEQ ID NO:27363)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYEGKQTNRIKVV EGGPLPFAFD ILATHFMYGS KVFYKYPADLPDYFKQSFPE GFTWERVMVFEDGGVLTATQ DTSLQDGELIYNVKVRGVNFPANGPVMQKK TLGWEPSTET MYPADGGLEGRCDKALKLVG GHLHVNFKT TYKSKKPVKM PGVHYVDRRLERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK(CyOFF变体SEQ ID NO:27364)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVVVEGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPEGFTWERVMVFEDGGVLTATQ DTSLQDGELI YNVKVRGVNFPANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVGGGHLHVNFKT TYKSKKPVKM PGVHYVDRRL ERIKEADNETYVEQYEHAVA RYSNLGGMD ELYK(CuOFF变体;SEQ ID NO:27365)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

**KVFTLGDFVGDWRQTAGYNQAQVLEQGGTSLFQNLGVSVTPIQRIV
LSGENGLKIDIHV
IIPYEGLSCDQMAQIEKIFKVVYPVDDHHFKAILHYGTLVIDGVTPNMI
DYFGQPYEGIA**

KFDGKKITVTGTLWNGNTIIDERLINPDGSLLFRVTINGVTGWRLHERILA(LumiLuc;SEQ ID NO:27366)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGQGTGNP NDGYEELNLK STKGDQLQFSP WILVPHIGYG FHQYLPYPDG MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP MYVFRKTELK HSKTELNFK E WQKAFTDVMG MDELYK(mNeonGreen;SEQ ID NO:27367)(可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYEGTQTAKLKVT KGGPLPFSWD ILSPQFMYGS RAFIKHPADIPDYKQSFPE GFKWERVMNF EDGGAVTVTQ DTSLEDGTLIYKVKLRGTNF PPDGPVMQKK TMGWEASTER LYPEDGVKGDIAKMALRLKD GGRYLADFKT TYKAKKPVMQ

PGAYNVDRKLDITSHNEDYT VVEQYERSEG RHSTGGMDEL YK(mScarlet-i;SEQ ID NO:27368) (可用于创建FRET和/或BRET传感器的全发光或荧光蛋白);

**SGKSYPTVSADYQKA VEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTF
DKRKTGGPFGTIRYPAELAHSANGLDIAVRLLEPLKAEFPILSYADF
YQLAGVVAVEVTGGPEVVPFHPGREDKPELPPEGRLPDATKGS DHLRD
VFGKAMGLTDQDIVALSGGHTLGAAHKERSGFEGPWTSNPLVFDNSY
FTELLSGEKEGGGGSGGGGS (APEX2-1-200; SEQ ID NO:27369);**

**GGGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAE
AHQKLSLGFADA (APEX2-201-250; SEQ ID NO:27370);**

MGSHHHHHHGSSENLYFQSGGS

**VRPLNCIVA VSQNMIGIGKN GDLPWPPLRN ESKYFQRM TT
TSSVEGKQNL
VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRG AHFL
AKSLDDALRL
IEQPELGGGGSGGGGS (DHFR_A (1-105); SEQ ID NO:27371);**

**SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIR
EAAAASEKISRE**

**GGGGSGGGGS ASKV DMVWIVGGSS VYQEAMNQPG HLRLFVTRIM
QEFESDTFFP
EIDLGKYKLL PEYPGVLSEV QEEKGIKYKF EVYEKKD
(DHFR_B (106-186); SEQ ID NO:27372);**

QLTPTFYDNPCPNVSNIVRDIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAFGNAN
SARGFSVIDRMKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFTL PQL
KDSFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGGSGS (sHRPa是
大的HRP拆分片段,它由具有以下4种突变的辣根过氧化物酶 (HRP) 的1-213个氨基酸组成:
T21I,P78S,R93G,N175S) (SEQ ID NO:27373);

NLSALVDFDLRTPTIFDNKYVNVLEEQGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAMDRMG
NITPLTGTQGGIRRNCRVVNSNGGSGS (sHRPb是小的HRP拆分片段,它由具有以下2种突变的辣根
过氧化物酶 (HRP) 的214-308个氨基酸组成:N255D,L299R) (SEQ ID NO:27374);

GESLFKGPRDYNPISSSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRN
 NGTLLVQSLHGVFKVKNTTTLQQHLIDGRDMIIRMPKDFPPFPQKLF
 REPQREERICLVTTNFQTGGGGSGGGGS (N_Tev (1-118) (SEQ ID
 NO:27375);

GGGGSGGGGSSKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPL
 VSTRDGFIVGIHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWR
 LNADSVLWGGHKVFMDKP C_Tev (119-221) (SEQ ID NO:27376);

MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQQEATEV
 RLEQKMPTLL
 RVYIDGPHGM GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV
 LGASETIANI
 YTTQHRLDQG EISAGDAAVV M TSAQITMGM PYAVTDAVLA
 PHIGGEAGSS
 HAPPALTLI FDRHPAALL CYPARYLMG SMTPQAVLAF
 VALIPPTLPG
 TNIVL GALPE DRHIDRLAKR QRPGERLDA MLAAIRRVYG
 LLANTVRYLQ

GGGSWREDWG QLSGT GGGGSGGGGS(胸苷激酶_TK_A(1-265) (SEQ ID NO:27377) ;和/或
GGGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP
 NGDLYNVFAW ALDVLAKRLR PMHVFILDYD QSPAGCRDAL
 LQLTSGMVQT

HVTTPGSIPT ICDLARTFAR EMGEAN(胸苷激酶_TK_B(266-376) (SEQ ID NO:27378)。

26. 根据权利要求1-24中任一项所述的笼蛋白,其中所述笼包含第二拆分报告蛋白结构域,其中(a)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27359和27664-7672的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,;

另一个包含与SEQ ID NO:27379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸残基可以存在或不存在:

MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIV
 RSGENALKIDIHIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG
 TLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLIT
 PDGSMLFRVTINS (LgBiT) (SEQ ID NO:27379);

(b)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27360的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列

VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRPEERFPMMSTF KVLLCGAVLS

RVDAGQEQLG RRIHYSQNDLVEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTIGGPKELTAFL
HNMGDHSVTRL DRWEPELNEA IPNDERDTTTPAAMATTLRK LLTGNGR (拆分 β -内酰胺酶A) (SEQ ID
NO:27360),

另一个包含与SEQ ID NO:27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、
92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGSRGIIAALGPD GKPSRIVVIY
TTGSQATMDE RNRQIAEIGA SLIKHW (拆分 β -内酰胺酶B) (SEQ ID NO:27361);

(c) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:
27362的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、
97%、98%或100%相同的氨基酸序列:

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMG
QIEKIFKVVYPVDNHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDG
SLLFRVTINGVTGWRLLHERILA (TeLuc) (SEQ ID NO:27362) (可用于创建FRET和/或BRET传感器的
全发光或荧光蛋白),

另一个包含与选自SEQ ID NO:27363-27365的氨基酸序列至少70%、75%、80%、85%、
90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVVVEGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPEGFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNFPANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVGGGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNETYVEQYEHAVA RYSNLGGMD ELYK (CyOFF变体) (SEQ ID NO:27363) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYEGKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFIKYPADLPDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELIYNVKVRGVNFP
PANGPVMQKK TLGWEPSTET MYPADGGLEGRCDKALKLVG GGHHLHVNFKT TYKSKKPVKM
PGVHYVDRRLERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CyOFF变体) (SEQ ID NO:27364) (可
用于创建FRET和/或BRET传感器的全发光或荧光蛋白); 和

EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVVVEGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPEGFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNFPANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVGGGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNETYVEQYEHAVA RYSNLGGMD ELYK (CyOFF变体) (SEQ ID NO:27365) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

(d) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:
27366的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、
97%、98%或100%相同的氨基酸序列:

KVFTLGDFVGDWRQTAGYNQAQVLEQGGTSLFQNLGVSVTPIQRIV
LSGENGLKIDIHV
IIPYEGLSCDQMAQIEKIFKVVYPVDDHHFKAILHYGTLVIDGVTPNMI^{K F D}
DYFGQPYEGIA

GKKITVTGTLWNGNTIIDERLINPDGSLLFRVTINGVTGWRLLHERILA (LemiLuc) (SEQ ID NO:27366) (可

用于创建FRET和/或BRET传感器的全发光或荧光蛋白),

另一个包含与SEQ ID NO:27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸残基可以存在或不存在:

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYYKQSFPE GFKWERVMNF EDGGAIVTQ DTSLEDGLI YVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL
DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i) (SEQ ID NO:27368) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

(e) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27367的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在:

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGGTGNP NDGYEELNLK STKGDQLQFSP
WILVPHIGYG FHQYLPYDGM MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT
GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP
MYVFRKTELK HSKTELNFKE WQKAFTDVMG MDELYK (mNeonGreen) (SEQ ID NO:27367), (可用于
创建FRET和/或BRET传感器的全发光或荧光蛋白),

另一个包含与SEQ ID NO:27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸残基可以存在或不存在:

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYYKQSFPE GFKWERVMNF EDGGAIVTQ DTSLEDGLI YVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL
DITSHNEDYT VVEQYERSEG RHSTGGMDELYK (mScarlet-i) (SEQ ID NO:27368) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

(f) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与氨基酸序列SEQ ID NO:27369至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

SGKSYPTVSADYQKAVEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTFDKRKTGGPFGTIRYPAELAHSANS
GLDIAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKPELPPEGRLPDATKGSDDLDRDVF
GKAMGLTDQDIVALSGGHTLGAHKERSGFEGPWTSNPLVFDNSYFTELLS~~GEKEGGGGSGGGGS~~ (APEX2-1-200)
(SEQ ID NO:27369) (用于化学发光和比色检测系统的大豆抗坏血酸过氧化物酶蛋白的拆
分工程化变体);

另一个包含与SEQ ID NO:27370的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGGSGGGGSGLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFADA (APEX2-201-
250) (SEQ ID NO:27370) (用于化学发光和比色检测系统的大豆抗坏血酸过氧化物酶蛋白

的拆分工程化变体)；

(g) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27371的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

MGSHHHHHGSGSENLYFQSGGS

VRPLNCIVA VSQNMIGIGKN GDLPWPPLRN ESKYFQRMTT

TSSVEGKQNL

VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRGAHFL

AKSLDDALRL

IEQPELGGGGSGGGGS (DHFR_A(1-105)) ; (SEQ ID NO:27371) (细胞生存或荧光的拆分二氢叶酸还原酶蛋白报告分子)

另一个包含与SEQ ID NO:27372的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

SGSG

DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAA

AASEKISRE

GGGGSGGGGS ASKV DMVWIVGGSS VYQEAMNQPG HLRLFVTRIM

QEFESDTFFP

EIDLKGYKLL PEYPGVLSEV QEEKGIKYKF EVYEKKD

(DHFR_B(106-186)) ; (SEQ ID NO:27372) (细胞生存或荧光的拆分二氢叶酸还原酶蛋白报告分子)；

(h) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27373的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

QLTPTFYDNPCPNVSNIVRDIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAFGNAN SARGFSVIDRMKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFTLPLQL KDSFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGGSGS (sHRPa是大的HRP拆分片段,它由具有以下4种突变辣根过氧化物酶(HRP)的1-213个氨基酸组成: T21I,P78S,R93G,N175S:质粒73147 (SEQ ID NO:27373) ;

另一个包含与SEQ ID NO:27374的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

NLSALVDFDLRTPTIFDNKYVYNLEEQKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAMDRMG NITPLTGTQQGQIRRNCRVNSNGGSGS (sHRPb是小的HRP拆分片段,它由具有以下2种突变的辣根过氧化物酶(HRP)的214-308个氨基酸组成:N255D,L299R:质粒73148) (SEQ ID NO:27374) ;

(i) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27375的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GESLFKGRDYNPISSTICHLTNEVDGHTTSLYIGIFGPFIIITNKHLFRRNNGTLLVQSLHGVFKVKNNTTL
QQHLIDGRDMI IIRMPKDFPPFPQKLKFRPQREERICLVTTNFQTGGGGSGGGGSN_Tev (1-118) (SEQ ID NO:27375) (拆分的TEV蛋白酶);

另一个包含与SEQ ID NO:27376的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGGSGGGGSKSMSSMVSDTSCTFPSSDGFIFWKHWIQTQDQCGSPLVSTRDGFIVGIHSASNFTNTNNYFT
SVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMKDP (C_Tev (119-221)); (SEQ ID NO:27376) (拆分的TEV蛋白酶);

(j) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27377的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且其中N端蛋氨酸残基可以存在或不存在:

MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQQEATEV
RLEQKMPTLL
RVYIDGPHGM GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV
LGASETIANI
YTTQHRLDQG EISAGDAAVV M TSAQITMGM PYAVTDAVLA
PHIGGEAGSS
HAPPPALTLI FDRHPIAALL CYP AARYLMG SMTPQAVLAF
VALIPPTLPG
TNIVLGALPE DRHIDRLAKR QRPGERLDA MLAAIRRVYG
LLANTVRYLQ

GGGSWREDWG QLSGT GGGGSGGGGS(胸苷激酶_TK_A(1-265)) (SEQ ID NO:27377);

另一个包含与SEQ ID NO:27378的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP
NGDLYNVFAW ALDVLAKRLR PMHVFILDYD QSPAGCRDAL
LQLTSGMVQT

HVTTPGSIPT ICDLARTFAR EMGEAN(胸苷激酶_TK_B(266-376)) (SEQ ID NO:27378)。

27. 根据权利要求1-26中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽能够与靶标结合,所述靶标包括但不限于抗体、毒素、诊断生物标志物、病毒颗粒、疾病生物标

志物、代谢物或生化分析物。

28. 根据权利要求1-26中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽能够结合抗体靶标。

29. 根据权利要求28所述的笼蛋白,其中所述一种或多种靶标结合多肽包括一种或多种由针对病毒靶的抗体识别的表位。

30. 根据权利要求28或29所述的笼蛋白,其中所述一种或多种靶标结合多肽包含一种或多种由针对SARS-Cov-2的抗体识别的表位。

31. 根据权利要求126中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽能够结合疾病标志物或毒素。

32. 根据权利要求1-26中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽能够结合Bcl-2、Her2受体、肉毒杆菌神经毒素B、白蛋白、上皮生长因子受体、前列腺特异性膜抗原(PSMA)、瓜氨酸肽、脑钠肽和/或心肌钙蛋白I。

33. 根据权利要求1-32中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27380-27430的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

34. 根据权利要求33所述的笼蛋白,其中所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27397-27430或选自SEQ ID NO:27397-27406、27409-27416和27427-27430的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

35. 根据权利要求34的笼蛋白,其中相对于参考靶标结合多肽氨基酸的氨基酸取代选自表3中提供的允许氨基酸取代。

36. 根据权利要求34或35所述的笼蛋白,其中界面残基与参考靶标结合多肽中的界面残基相同,或相对于参考靶标结合肽中的表面残基保守地被取代,如表2所示。

37. 根据权利要求34-36中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27397-27406和27431-27466的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

38. 根据权利要求37所述的笼蛋白,其中所述一种或多种靶标结合多肽包含相对于SEQ ID NO:27397的氨基酸序列1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17或全部18个在选自2、4、15、14、17、18、27、28、32、37、38、39、41、42、49、52和55残基的氨基酸取代。

39. 根据权利要求38所述的笼蛋白,其中所述一种或多种靶标结合多肽中的取代选自表5中列出的取代,可以是单独的,也可以是给定行中的组合。

40. 权利要求34-36中任一项所述的笼蛋白,其中所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27409-27416和27467-27493的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

41. 根据权利要求40的笼蛋白,其中所述靶标结合包含相对于SEQ ID NO:27409的氨基酸序列1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或全部20个在选自2、6、8、

14、19、22、25、26、28、29、34、35、37、40、43、45、49和62残基的氨基酸取代。

42. 根据权利要求41所述的笼蛋白,其中所述取代选自表7中列出的取代,可以是单独的,也可以是给定行中的组合。

43. 根据权利要求34-36中任一项所述的笼蛋白,其中所述靶标结合包含与选自SEQ ID NO:27427-27430和27494的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

44. 根据权利要求44的笼蛋白,其中所述一种或多种靶标结合多肽在选自63和75的残基或两个残基处包含相对于SEQ ID NO:27429的氨基酸序列的氨基酸取代。

45. 根据权利要求44的笼蛋白,其中所述取代包括R63A和/或K75T。

46. 根据权利要求1-45中任一项所述的笼蛋白,其包含与选自SEQ ID NO:1-49、51-52、54-59、61、65、67-91、92-2033、2034-14317、27094-27117、27120-27125、27278至27321,SEQ ID NO:27126至27276之间的偶数SEQ ID NO,以及表8和表9中列出的笼蛋白,不包括任选的氨基酸残基并且不包括锁区中的氨基酸残基,并且其中每种笼蛋白的N端和/或C端60个氨基酸可以是任选的。

47. 根据权利要求1-47中任一项所述的笼蛋白,其包含与选自SEQ ID NO:27497-27620的笼蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,其中N端蛋白纯化标签(MGSHHHHHGSGSENLYFQSGG (SEQ ID NO:27624);或MGSHHHHHGSENLYFQG (SEQ ID NO:27625);或GSHHHHHGSGSENLYFQG (SEQ ID NO:27626))是任选的,在百分比同一性比较中不考虑,并且可以存在或不存在,优选不存在。

48. 一种钥匙蛋白,其能够与权利要求1-47中任一项不包括第二报告蛋白结构域的笼蛋白的结构区结合,其中所述钥匙蛋白包括第二报告蛋白结构域,其中所述钥匙蛋白第二报告蛋白结构域和所述笼蛋白第一报告蛋白结构的相互作用引起来自所述第一报告蛋白结构的报告活性的可检测变化。

49. 根据权利要求48所述的钥匙蛋白,其中所述第二报告蛋白结构域位于所述钥匙蛋白的N端或C端,或位于所述钥匙蛋白的N端或C端的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或1个氨基酸内。

50. 根据权利要求48-49中任一项所述的钥匙蛋白,其中所述第二报告蛋白结构域包括报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子,荧光共振能量转移(FRET)报告分子、比色报告分子(包括但不限于β-内酰胺酶、β-半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶),和分子条形码报告分子(包括但不限于TEV蛋白酶)。

51. 根据权利要求48-50中任一项所述的钥匙蛋白,其中所述第二报告蛋白结构域包含与选自SEQ ID NO:27360-23379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且其中任何N端氨基酸残基可以存在或不存在。

52. 根据权利要求48-51中任一项所述的钥匙蛋白,其中不包括第二拆分报告蛋白结构域的钥匙蛋白包含与选自SEQ ID NO:14318-26601、26602-27015、27016-27050、27322至27358和在27127和27277之间具有奇数SEQ ID NO的钥匙多肽的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,不包括任选的氨基酸残基,表8中的钥匙蛋白以及表9中的钥匙蛋白。

53. 根据权利要求48-52中任一项所述的钥匙蛋白,其中所述钥匙蛋白包含与选自SEQ ID NO:27621-27623的钥匙蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,不包括括号中的任选氨基酸残基,其中括号中的残基是任选的并且可以存在或不存在;

>lucKey:MGS-(His)6-TEV位点-接头-LgBit-接头-闪烁序列

(MGSHHHHHHGGSGSENLYFQG)SGMVFTLEDFVGDWEQTAAYNLDQV
LEQGGVSSLLQNLAVSVTPIQRIVRSGENALKIDIHIIPYEGLSADQMAQ
IEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFD
GKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSGGSGGGGSGGGSG
GSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAA
AASEKISRE (SEQ ID NO:27621)

钥匙-2GGSGG-Cy0FP (Cy0FP序列加粗/加下划线):

(M)DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIRE
AAAASEKISREGGSGG **GGVSK GEELIK ENMRSKLYLE**
GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFIKYPADL PDYFKQSFPE GFTWERVMVF
EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT
TYKSKKPVKM PGVHYVDRRL ERIKEADNET YVEQYEHAVA
RYSNLGGGMD ELYK (SEQ ID NO: 27622)

钥匙-LacB (拆分 β -内酰胺酶B加粗/加下划线):

SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIR
EAAAASEKISRESGGGGSGGGGSGGGG **LLTLASRQOLIDWME**
ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW (SEQ ID NO:
27623)。

54. 一种生物传感器,包括

- (a) 权利要求1-47中任一项所述的笼蛋白,其中所述笼不包括第二报告蛋白结构域;和
- (b) 权利要求48-53中任一项所述的钥匙蛋白;

其中所述钥匙蛋白只能在所述笼蛋白一种或多种靶标结合多肽能够结合的靶标存在下与所述笼蛋白结合;并且其中所述笼蛋白的第一报告蛋白结构域与所述钥匙蛋白的第二

报告蛋白结构域的结合引起来自所述第一报告蛋白结构域的报告活性的可检测变化。

55. 根据权利要求54所述的生物传感器, 其中

(a) 第一报告蛋白结构域包含与氨基酸序列SEQ ID NO: 27359和27664-27672至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

并且所述第二报告蛋白结构域包含与SEQ ID NO 27379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中所述N端蛋氨酸残基可以存在或不存在

(b) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27360的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 另一个包含与SEQ ID NO: 27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

(c) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27362的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 另一个包含与选自SEQ ID NO: 27363-27365的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

(d) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27366的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 另一个包含与SEQ ID NO 27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

(e) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27367的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中所述N端蛋氨酸残基可以存在或不存在, 另一个包含与SEQ ID NO 27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中N端蛋氨酸残基可以存在或不存在;

(f) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27369的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中带下划线的残基是可以存在或不存在的任选残基, 并且当存在时可以是任何氨基酸序列; 另一个包含与SEQ ID NO: 27370的氨基酸至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中带下划线的残基是可以存在或不存在的任选残基, 并且当存在时可以是任何氨基酸序列;

(g) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO: 27371的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列, 其中带下划线的残基是可以存在或不存在的任选残

基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27372的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

(h)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27373的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27374的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

(i)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27375的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27376的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

(j)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27377的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在,并且其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且另一个包含与SEQ ID NO:27378的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列。

56.根据权利要求54的生物传感器,其中所述笼蛋白包含权利要求47的笼蛋白,并且所述钥匙蛋白包含与SEQ ID NO:27621的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,不包括括号中的任选氨基酸残基;

>lucKey:MGS-(His)6-TEV位点-接头-LgBit-接头-锁序列

(MGSHHHHHHGSSENLYFQG)SGMVFTLEDFVGDWEQTAAYNL
DQVLEQGGVSSLLQNLAVSVTPIQRIVRSGENALKIDIHIIPYEGLSADQ
MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGI
AVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSGGSGGGGSG
GGSGGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIR
EAAAASEKISRE。

57.根据权利要求54所述的生物传感器,其中所述笼蛋白和所述钥匙蛋白包括蛋白对,所述蛋白对包括

(i)笼蛋白,其包含与SEQ ID NO:27620的氨基酸序列至少70%、75%、80%、85%、

90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中括号中的残基是任选的并且可以存在或不存在:

LacATrop (拆分 β -内酰胺酶A加粗;cTnT和cTnC加下划线):

**(MGSHHHHHGGSGSENLYFQG SGGG)VFAHPETLVK VKDAEDQLGA
RVGYIELDLN
SGKILESFRP EERFPMMSTF KVLLCGAVLS RVDAGQEQLG
RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI
GGPKELTAFL
HNMGDHSVTRL DRWEPENEA IPNDERDTT PAAMATTLRK
LLTGENGR
SGGGGSGGGGSGGGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIR
LAEALLEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRA
EKEIDDAAKESKKILEEARKAIRDAAEESRKILEEGSGSGSDALDELQK
LNLELAKLLLKAI AETQDLNLRAAKAFLEAAAKLQELNIRAVELLVKL
TDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLIEKAKE
ESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELL
RALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLIR
EAAAASEDQLREAAKELWQTIYNLEAEKFDLQEKFKQKYEINVLRN
RINDNQKVSCTKDDSKGKSEBELSDLFRMFDKNADGYIDLEELKIMLQ
ATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE; 以及**

(ii) 钥匙蛋白,其包含与SEQ ID NO:27631的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列。

**LLTLASRQQLIDWME ADKVAGPLL SALPAGWFIA DKSGAGERGS
RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW
(SEQ ID NO:27361)**

58. 一种用于检测靶标的方法,包括

(a) 在促进笼蛋白一种或多种靶标结合多肽与生物样品中存在的靶标结合的条件下,将权利要求47中任一项所述的笼蛋白或权利要求40-43中任一项所述的生物传感器与生物样品接触,引起来自第一报告蛋白结构域的报告活性的可检测变化;和

(b) 检测来自第一报告蛋白结构域的报告活性的变化,其中报告活性的改变将样品识别为含有靶。

59. 根据权利要求58所述的方法,其中所述靶标选自包括但不限于抗体、毒素、诊断生物标志物、病毒颗粒、疾病生物标记、代谢物或生化分析物的组。

60. 根据权利要求58-59中任一项所述的方法,其中所述靶标是抗体。

61. 根据权利要求60所述的方法,其中所述靶标包括对病毒具有选择性的抗体。

62. 根据权利要求58所述的方法,其中所述一种或多种靶标结合多肽包含选自SEQ ID NO:27292-27394和27547-27548的氨基酸序列,以及包含与选自SEQ ID NO:27397-27494的

氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、98%和98%或100%相同的氨基酸序列的多肽。

63. 根据权利要求58-62中任一项所述的方法,其中所述笼多肽包含权利要求47的笼蛋白。

64. 根据权利要求58-59中任一项所述的方法,其中所述靶标是疾病标志物或毒素。

65. 根据权利要求64的方法,其中所述疾病标志物或毒素包括Bcl-2、Her2受体、肉毒杆菌神经毒素B、白蛋白、上皮生长因子受体、前列腺特异性膜抗原(PSMA)、瓜氨酸化肽、脑钠肽和/或心肌钙蛋白I。

66. 权利要求64或65的方法,其中所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27380-27390的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,其中任何N端氨基酸是任选的并且可以存在或不存在。

67. 根据权利要求64-66中任一项所述的方法,其中所述笼蛋白包括权利要求47的笼蛋白。

68. 一种用于设计生物传感器、笼蛋白或钥匙蛋白的方法,包括本文所述的任何方法的步骤。

69. 编码前述权利要求中任一项所述的笼蛋白或钥匙蛋白的核酸。

70. 一种表达载体,其包含权利要求69的核酸,所述核酸可操作地连接到诸如启动子的合适控制元件。

71. 一种细胞,其包含前述权利要求任一项所述的笼蛋白、钥匙蛋白、组合物、核酸或表达载体。

72. 一种药物组合物,其包含任何前述权利要求的笼蛋白、钥匙蛋白、组合物、核酸、表达载体或细胞,以及药学上可接受的载体。

73. 一种表位,其包含SEQ ID NO:27384的氨基酸序列,或其由组成。

74. 一种用于检测样品中的肌钙蛋白I的方法,包括在适合于促进样品中肌钙蛋白I与表位结合的条件下使生物样品与权利要求73的表位接触以形成结合复合物,并对证明样品中存在肌钙蛋白II的结合复合物进行检测。

基于从头设计的蛋白开关的模块化和通用化生物传感器平台

交叉引用

本申请要求于2020年5月27日提交的序列号63/030,836、2020年7月14日提交的序列号63/051,549和2020年8月19日提交的序列号63/067,643的美国临时专利申请的优先权,通过引用将其全部并入本文。

联邦基金声明

本发明是在国防高级研究计划局(DARPA)授予的第FA8750-17-C-0219号拨款的政府支持下完成的。政府对这项发明有一定的权利。

序列表声明

[0003] 序列表的计算机可读形式通过电子提交方式与本申请一起提交,并通过引用将其整体并入本申请。序列表包含在2021年5月25日创建的文件中,文件名为“20-1075-WO_Sequence-Listing_ST25.txt”,大小为32,910kb。

背景技术

[0004] 传感器蛋白已成为一个活跃的研究领域。传统的ELISA方法需要多个液体处理步骤,阻碍了在临床的使用。侧流免疫色谱法快速、廉价,但灵敏度、重现性有限,定量性能差。ELISA和侧流法还需要两个结合模块来检测靶标,一个用于捕获,另一个用于读出。蛋白传感器构建的一个主要障碍是寻找经历足够构象变化的分析物结合结构域。最常用的结合结构域(例如抗体)在配体结合时仅经历环的微小结构变化。将合适的报告分子与最佳几何形状耦合以放大构象变化也是成功的生物传感器的关键。然而,在蛋白界面中计算设计小分子结合位点和生成半合成蛋白传感器都是当前颇具挑战性的问题。因此,需要使用简单且稳健的计算方案设计生物传感器的通用方法,该方案几乎不需要经验优化。

发明内容

[0005] 在一个方面,本公开提供了包括螺旋束的笼蛋白,其中所述笼蛋白包括结构区和锁区,其中所述锁区包括一种或多种靶标结合多肽,其中当所述第一报告蛋白结构域结合到第二报告蛋白结构域时,报告活性发生可检测的变化,并且其中所述结构区与所述锁区相互作用以防止溶液进入所述一种或多种靶标结合多肽。在一个实施方案中,笼蛋白还包括第二报告蛋白结构域,其中第一报告蛋白结构域和第二报告蛋白结构域中的一个存在于锁区中,另一个存在于结构区中,其中在所述一种或多种靶标结合多肽结合的靶标存在下,所述第一报告蛋白结构域和所述第二报告蛋白结构域的相互作用减弱。在另一个实施方案中,第二报告蛋白结构域不存在于笼蛋白中。在另一个实施方案中,第一报告蛋白结构域和第二报告蛋白结构域(当存在时)包括报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子、荧光共振能量转移(FRET)报告分子、比色报告分子(包括但不限于 β -内酰胺酶、 β -半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于

APEX2)、放射性报告分子(包括但不限于胸苷激酶),和分子条形码报告分子(包括但不限于TEV蛋白酶)。在一个实施方案中,所述一种或多种靶标结合多肽能够与靶标结合,所述靶标包括但不限于抗体、毒素、诊断生物标志物、病毒颗粒、疾病生物标志物、代谢物或生化分析物。

[0006] 在另一方面,本公开提供了能够结合到本公开的任何实施方案的笼蛋白的结构区的钥匙蛋白,该结构区不包括第二报告蛋白结构域,其中所述钥匙蛋白包括第二报告蛋白结构域,其中所述钥匙蛋白第二报告蛋白结构域和所述笼蛋白第一报告蛋白结构的相互作用引起来自所述第一报告蛋白结构的报告活性的可检测变化,第二报告蛋白结构域包括选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子、比色报告分子(包括但不限于 β -内酰胺酶、 β -半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶),和分子条形码报告分子(包括但不限于TEV蛋白酶)。

[0007] 在另一方面,本公开提供了生物传感器,包括

[0008] (a)本发明实施方案的笼蛋白,其中所述笼不包含第二报告蛋白结构域;和

[0009] (b)本发明任何实施方案的钥匙蛋白;

[0010] 其中所述钥匙蛋白只能在所述笼蛋白一种或多种靶标结合多肽能够结合的靶标存在下与所述笼蛋白结合;并且其中所述笼蛋白的第一报告蛋白结构域与所述钥匙蛋白的第二报告蛋白结构域的结合引起来自所述第一报告蛋白结构域的报告活性的可检测变化。

[0011] 在另一方面,本公开提供了用于检测靶标的方法,包括

[0012] (a)在促进笼蛋白的一种或多种靶标结合多肽与生物样品中存在的靶标结合的条件下,将其中笼蛋白包含第二报告蛋白结构域的本发明任何实施方案的笼蛋白或本发明任何一个实施方案的生物传感器与生物样品接触,引起来自第一报告蛋白结构域的报告活性的可检测变化;和

[0013] (b)检测来自第一报告蛋白结构域的报告活性的变化,其中报告活性的改变将样品鉴定为含有所述靶标。

[0014] 在进一步的方面,本公开提供了包括本文所述的任何方法步骤的用于设计生物传感器、笼蛋白或钥匙蛋白的方法;编码本公开的任何实施方案的笼蛋白或钥匙蛋白的核酸;包含可操作地连接到合适的控制元件(例如启动子)的本公开的实施方案的核酸的表达载体;包含本公开任何实施方案的笼蛋白、钥匙蛋白、组合物、核酸或表达载体的细胞(例如重组细胞);包含本发明任何实施方案中的笼蛋白、钥匙蛋白、组合物、核酸、表达载体或细胞、和药学上可接受的载体的药物组合物;包含或由SEQ ID NO:27384的氨基酸序列组成的表位;以及检测样品中的肌钙蛋白I的方法,包括在适合于促进样品中肌钙蛋白I与表位结合的条件下使生物样品与所述表位接触以形成结合复合物,并对证明样品中存在肌钙蛋白II的结合复合物进行检测。

附图说明

[0015] 图1(a-f),多状态变构生物传感器的从头设计。a.传感器示意图。生物传感器由两种蛋白组成:lucCage和lucKey,它们以封闭(关闭)和开放(打开)状态存在。lucCage的封闭

形式(左)无法与lucKey结合,因此,阻止了拆分的荧光素酶SmBit片段与LgBit相互作用。开放的形式(右)可以结合靶标和钥匙,并允许SmBit与lucKey上的LgBit结合以重建荧光素酶活性。

b. 生物传感器激活的热力学。从封闭笼(类别1)到开放笼(类别2)的转变的自由能成本 ΔG_{open} 不利于钥匙(类别5)的结合和在没有靶标的情况下荧光素酶活性的重建(类别6)。在靶标存在的情况下,组合的靶标结合(2 \rightarrow 3. ΔG_{LT})自由能、钥匙结合(3 \rightarrow 4. ΔG_{CK})自由能和SmBit LgBit关联(4 \rightarrow 7. ΔG_{R})自由能克服不利的 ΔG_{open} ,驱动lucCage的打开和荧光素酶活性的重建。

c. 基于热力学的生物传感器设计策略。对于每种生物传感器,可设计参数为 ΔG_{open} 和 ΔG_{CK} ; ΔG_{R} 对于所有靶标都是相同的, ΔG_{LT} 是为每种靶标预先指定的。对于敏感但低背景的分析物检测, ΔG_{open} 和 ΔG_{CK} 必须设计为:在没有靶标的情况下,封闭状态(类别1)的自由能显著低于开放状态(类别6),但在有靶标的情况(类别7)下,自由能高于开放状态。

d-f, (d) K_{open} 、(e) K_{LT} 和(f) $[\text{lucKey}]_{\text{tot}}$ 和 $[\text{lucCage}]_{\text{tot}}$ 的不同值下的b中所示的耦合平衡的数值模拟。 K_{open} 、 K_{LT} 、 K_{CK} 分别设置为 1×10^{-3} nM、1 nM和10 nM,传感器组分的浓度为10:100 nM (lucCage:lucKey),除非明确指出。d. 增加 ΔG_{open} 会改变对较高分析物浓度的响应。e. 传感器检测限约为 $0.1 \times K_{\text{LT}}$;用于打开开关的驱动力在该浓度以下变得太弱。f. 可以通过改变传感器组分浓度来调节有效靶标检测范围。

[0016] 图2(a-d). 对小蛋白进行整合作为传感结构域的从头生物传感器设计和表征。a. 将小蛋白结构域限制在LOCKR开关中的一般策略和结构验证。左图:与流感血凝素(HA,带状图像)茎区结合的从头结合域(binder)HB1.9549.2的设计模型¹⁵。右图:sCageHA_267_1S的晶体结构,包括接枝到LOCKR开关的缩短和稳定版中的HB1.9549。中间图:除F273外,参与与HA结合的HB1.9549.2的所有残基(顶部)都埋在开关(底部)的闭合状态下,以阻止其相互作用。标签指示两幅图中相同的氨基酸组(上图中的F2对应于下图中的F73)。

b-d, 3种变构生物传感器的功能表征:lucCageBot(肉毒杆菌神经毒素b(BoNT/b)检测)、lucCagerProA(Fc结构域检测)和lucCage Her2(Her2受体检测)。左图:所示生物传感器的结构模型(带状图像),所述生物传感器整合了BoNT/B(Bot.671.2)、通用抗体结合蛋白A(SpaC)的C结构域和Her2结合抗体的从头设计的结合域,分别接枝到包含笼状SmBit片段的lucCage中。中间图:将50 nM分析物(BoNT/B、IgG-Fc或Her2)加入到10 nM每个lucCage和10 nM lucKey的混合物中时发光强度的动力学测定。右图:在广泛的分析物浓度范围内通过改变生物传感器浓度(50 nM、5 nM和1 nM lucCage和lucKey;分别为青色、品红色和黑色线)进行的检测。

[0017] 图3(a-h). 心肌钙蛋白I和抗HBV抗体生物传感器的设计和表征。a. 心肌钙蛋白传感器lucCageTrop的设计I。左图:心肌钙蛋白的结构(PDB ID: 4Y99);右图:lucCageTrop的设计模型,一种处于闭合状态的cTnI传感器,包含cTnT和cTnC段。

b. 左图:将1 nM cTnI添加到0.1 nM lucCageTrop传感器+0.1 nM lucKey后,发光动力学增加。右图:通过改变传感器组件(线)的浓度,可以实现宽的分析物(cTnI)检测范围。灰色区表示与急性心肌梗死(AMI)诊断相关的cTnI浓度范围;虚线表示由WHO(0.6 ng/mL, 25 pM)定义的临床AMI截止值。

c. 设计lucCageHBV和lucCageHBV α 的模型,分别含有SmBit和一种或两种来自乙型肝炎病毒(HBV)PreS1蛋白的串联抗原表位。

d. lucCageHBV α (两个表位拷贝)对抗HBV抗体HzKR127-3.2(Kd=0.68 nM)的亲合力高于lucCageHBV(一个表位拷贝)(Kd=20 nM),如生物膜层干涉分析所示。

e. 左图:将10 nM抗HBV抗体添加到1 nM lucCageHBV α +1 nM lucKey后,生物发光信号的动力学增加。右图:通过改变传感器成分的浓度,可以在较宽的浓度范围内实现灵敏的抗HBV抗

体检测。f. 使用lucCageHBV检测HBV蛋白PreS1的机制示意图。g. 添加抗HBV抗体(步骤1)和随后的PreS1(步骤2)后的生物发光动力学。当PreS1与传感器竞争抗体时,添加PreS1后生物发光降低。h. 在相关的HBV感染后浓度水平(灰色区)上可以实现PreS1的敏感检测。传感器与抗HBV抗体预混合;PreS1检测范围可以通过改变抗体的浓度(由彩色标签指示)来调节。

[0018] 图4(a-d).设计用于检测抗SARS-CoV-2抗体和SARS-CoV-2 RBD的生物传感器。a. SARS-CoV-2病毒结构表示,显示主要结构蛋白:包膜蛋白(E)、膜蛋白(M)、核衣壳蛋白(N)和含有受体结合域(RBD)的刺突蛋白(S)。根据公布的免疫原性数据选择M和N蛋白的线性表位。b. 左图:lucCageSARS2-M的结构模型。将SARS-CoV-2膜蛋白a.a.1-17表位的两个拷贝接枝到与柔性间隔区连接的lucCage中。中间图:添加100nM的抗SARS-CoV-1膜蛋白兔多克隆抗体后lucCageSARS2-M(50nM)+lucKey(50nM)的发光激活动力学(ProSci,3527)。这些最初针对对应于SARS-CoV基质蛋白的氨基酸末端附近的13个氨基酸产生的抗体与SARS-CoV-2膜蛋白的残基1-17交叉反应。右图:lucCageSARS2-M(5nM)+lucKey(5nM)对不同浓度的靶抗M pAb的反应。c. 左图:lucCageSARS2-N的结构模型。将SARS-CoV-2核衣壳蛋白369-382表位的两个拷贝接枝到与柔性间隔区连接的lucCage中。中间图:添加100nM抗SARS-CoV-1-N小鼠单克隆抗体(克隆18F629.1)后lucCageSARS2-N(50nM)+lucKey(50nM)的发光激活动力学。这种最初针对SARS-CoV-1核衣壳蛋白的残基354-385产生的抗体与SARS-CoV-2核衣壳蛋白的残基369-382交叉反应。右图:lucCageSARS2-N(50nM)+lucKey(50nM)对不同浓度靶标(抗N单抗)的反应。d. SARS-CoV-2 RBD传感器lucCageRBD的功能表征。左图:lucCageRBD的结构模型,模型示出了包含被关住的SmBiT片段的LCB1结合域。第二幅图:向1nM lucCageRBD和1nM lucKey的混合物中加入16.7nM RBD后的发光强度的动力学测定。第三幅图:通过改变生物传感器的浓度(10和1nM lucCage和lucKey),在广泛的分析物浓度范围内进行检测。右图:检测溶液中RBD时,lucCageRBD和lucKey的检测限(LOD)分别为1nM。测定LOD为15pM。

[0019] 图5.生物传感器特异性。将1nM的每种传感器与50nM的同源靶标(黑线)和其他生物传感器的靶标(灰色线)孵育。靶标是Be1-2、BoNT/B、人IgG-Fc、Her2、心肌钙蛋白I、抗HBV抗体(HzKR127-3.2)、抗SARS-CoV-1-M多克隆抗体和SARS-CoV-2RBD。所有实验一式三份,显示了代表性数据,数据以平均值 \pm 标准偏差表示。

[0020] 图6(a-g).lucCage中最佳SmBit位置的确定和lucCageBim(Bcl-2生物传感器)的表征。a. 蛋白模型显示了SmBiT和Bim肽在从头LOCKR开关的闩锁螺旋上的不同螺旋位置。b. 11种从头Bcl-2传感器的实验筛选。通过组合(a)中的SmBit和Bim位置产生11种变体,并通过添加Bcl-2后激活其发光来表征。在存在或不存在Bcl-2(200nM)的情况下,对每种设计物(20nM)和lucKey(20nM)进行发光测定。SmBiT312-Bim339(因此称为lucCageBim)因其较高的亮度、动态范围和稳定性而被选择用于后验表征。c-g,lucCageBim的表征。c. 带状图像中的结构设计模型。d. 放大显示SmBiT和笼的预测界面。e. 放大显示Bim和笼的预测界面。f. 向lucCageBim(20nM)和lucKey(20nM)的混合物中加入Bcl-2(200nM)后的动力学发光测定。g. 通过改变传感器(lucCageBim和lucKey)成分(曲线)的浓度,可调节lucCage Bim对Bcl-2的灵敏度。

[0021] 图7(a-d).sCageHA设计物的功能筛选和sCageHA_267-1S的晶体结构。a. 使用嵌入

的从头的结合域HB1.9549.2的sCageHA设计的结构模型。在锁螺旋的不同位置(包括三个连续的甘氨酸残基),将HB1.95492蛋白接枝到亲本六螺旋束(sCage)中。黑色箭头表示锁上额外引入的单V255S(1S)或双V255S/I270S(2S)突变。b. 无论是否存在钥匙,都通过生物膜层干涉法对五种结合HA的sAgeHA设计进行实验验证。sCage和钥匙的浓度分别为1 μ M和2 μ M。sCageHA_267-1S的活化倍数最高。c. 结构比较显示了sCage的柔性特性,以实现关注HB1.9549.2。将sCage结构模型和sCageHA_267-1S晶体结构叠加在一起,并在正交视图中显示了窄截面(黑框)的细节。HB1.9549.2的N端螺旋从锁螺旋移位(α 6) **3.2Å**(中间图),同时 α 5伴随位移,涉及sCage的Q16和N214的氢键网络的部分破坏(右图)。d. sCageHA_267-1S分子内相互作用的放大图。HA结合残基突出显示。HB1.9549.2的N端螺旋(α 1)和随后的螺旋(α 2)与笼相互作用。分子内的相互作用都是疏水的。F285的庞大疏水侧链紧紧地邻接sCage的 α 5骨架原子,这在没有 α 5弯曲的情况下不太可能发生。还发现了不利的相互作用:F273暴露在溶剂中,Y287羟基埋藏在非极性环境中。最右边的图示出了电子密度图的质量。

[0022] 图8(a-d). 肉毒神经毒素B传感器的设计和表征。a. 肉毒杆菌神经毒素B(BoNT/B)传感器设计的结构模型显示了lucCage锁上Bot.0671.2(PDB ID:5VID)的不同螺旋位置。SmBit肽以带状表示。I328S和L345S表示引入的突变以调节锁-笼界面(1S=I328S,2S=I328S/L345S)²,"GGG"表示锁和接枝蛋白之间存在三个连续的甘氨酸残基。黑框显示了349_2S设计中笼和Bot.0671.2界面的特写视图。b. 9种从头BoNT/b传感器的实验筛选。在存在或不存在BoNT/B蛋白(200nM)的情况下,对每种设计物(20nM)和lucKey(20nm)进行发光测定。在不存在BoNT/B的情况下,将每种设计物的发光值归一化为100。设计物349_2S因其高灵敏度和稳定性被选为最佳候选,并命名为lucCageBot。c. lucCagerBot灵敏度的测定。在连续稀释的BoNT/B蛋白存在下,在6000秒内测定生物发光。从上到下lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。d. 不同浓度下传感器的检测限(LOD)计算。从上到下-lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。误差条表示SD。

[0023] 图9(a-d). Fc结构域传感器的设计和表征。a. Fc传感器设计的结构模型示出了lucCage锁上金黄色葡萄球菌(*S. aureus*)蛋白A结构域C(PDB ID:4WWI)的不同螺旋位置。SmBit肽以带状表示。I328S和L345S表示引入的突变以调节锁-笼界面,(1S=I328S,2S=I328S/L345S)²,"GGG"表示锁和接枝蛋白之间存在三个连续的甘氨酸残基。b. 6种从头Fc结构域传感器的实验筛选。在存在或不存在重组人IgG1-Fc(200nM)的情况下,对每种设计物(20nM)和lucKey(20nM)进行发光测定。在不存在Fc的情况下,将发光值归一化为100。设计物351_2S因其高灵敏度和稳定性被选为最佳候选,并命名为lucCageProA。c. lucCageProA灵敏度的测定。在连续稀释的Fc蛋白存在下,在6000秒内测定生物发光。从上到下lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。d. 不同浓度下传感器的检测限(LOD)计算。从上到下-lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。误差条表示SD。

[0024] 图10(a-d). Her2传感器的设计和表征。a. Her2传感器设计的结构模型示出了lucCage锁上Her2亲和抗体蛋白(PDB ID:3MZW)的不同螺旋位置。SmBit肽以带状表示。I328S和L345S表示引入的突变以调节锁-笼界面,"GGG"表示锁和接枝蛋白之间存在三个连续的甘氨酸残基。黑框示出了在存在或不存在Her2的胞外域(200nM)的情况下,354_2S设计物中Cage和Her2亲和抗体界面的放大图。在不存在Her2胞外域的情况下,发光值归一

化为100。由于高灵敏度和稳定性,354_2S设计物被选为最佳候选,并命名为lucCageHer2。c,lucCagerHer2灵敏度的测定。在连续稀释的Her2胞外域蛋白存在下,在6000秒内测定生物发光。从上到下-lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。d,不同浓度下传感器的检测限(LOD)计算。从上到下-lucCageBot:lucKey浓度(nM)=50:5、5:5、1:10、0.5:0.5。误差条表示SD。

[0025] 图11(a-f)。用于心肌钙蛋白I检测的lucCageTrop的设计、选择和工程化。a.心肌钙蛋白I(cTnI)设计传感器的实验筛选。心肌钙蛋白T的片段,即cTnTf1-f6,在门锁的不同位置通过计算接枝到lucCage中。所有设计物均在大肠杆菌中产生,在20nM进行试验筛选,20nM lucKey用于cTnI(100nM)存在下的发光增加。在不存在cTnI的情况下,发光值归一化为100。设计物336-cTnTf6-K342A根据其敏感性、激活倍数变化和稳定性被选为最佳候选(命名为lucCageTrop627)。

cTnTf1:226-EDQLREKAKELWQTI-240(SEQ ID NO:27385)

cTnTf2:226-EDQLREKAKELWQTIYN-242(SEQ ID NO:27386)

cTnTf3:226-EDQLREKAKELWQTIYNLEAE-246(SEQ ID NO:27387)

cTnTf4:226-EDQLREKAKELWQTIYNLEAEKFD-249(SEQ ID NO:27388)

cTnTf5:226-EDQLREKAKELWQTIYNLEAEKFDLQE-252(SEQ ID NO:27389)

cTnTf6:226-EDQLREKAKELWQTIYNLEAEKFDLQEKFKQKYEINVLNRNRINDNQ-272(SEQ ID NO:27390) b.lucCageTrop627和lucCage Trop的模型,通过在lucCageTrop627的C端融合心肌钙蛋白C(cTnC)的改进版本。模型以带状表示形式显示,包括SmBit cTnT片段(PDB ID:4Y99)和cTnC(PDBID:4Y99)。黑框显示了lucCageTrop设计中Cage和cTnT界面的特写视图。c.通过生物膜层干涉法测定lucCageTrop627和lucCage Trop对cTnI的结合亲和力。lucCageTrop对cTnI的亲和力比lucCageTrop627高7倍。d.在连续稀释的cTnI存在下,lucCage Trop627(上图)和lucCageTrop(下图)之间的生物发光动力学比较。更高的结合亲和力导致传感器的动态范围和灵敏度的提高。e.lucCageTrop灵敏度的测定。在连续稀释的cTnI存在下,在6000秒内测定生物发光。从上到下-lucCageTrop:lucKey浓度(nM)=1:10、1:1、0.5:0.5、0.1:0.1。f,不同浓度下传感器的检测限(LOD)计算。从上到下-lucCageTrop:lucKey浓度(nM)=1:10、1:1、0.5:0.5、0.1:0.1。误差条表示SD。

[0026] 图12(a-f)。抗HBV抗体传感器的设计和表征。a.lucCage设计物的能量最小化模型分别示出了SmBit的螺旋段和PreS的抗原基序。黑框示出了HBV344设计物的笼-基序界面的放大视图。b.通过在存在或不存在抗HBV抗体HzKR127-3.2(100nM)的情况下监测每种lucCage(20nM)和lucKey(20nM)的发光来进行所有设计的实验筛选。在没有抗HBV的情况下,发光值归一化为100。HBV344设计因其更好的性能而被选择,并被命名为lucCageHBV。c、d.lucCageHBV敏感性的测定。在连续稀释的HzKR127-3.2存在下,在6000s内测定生物发光。从上到下-lucCageHBV:lucKey浓度(nM)=50:5、5:5、1:1。c中曲线的最大值用于获得d、e不同浓度下传感器的检测限(LOD)计算中的曲线。从上到下-lucCageHBV:lucKey浓度(nM)=50:5、5:5、1:1。f,添加抗体后的发光动力学(抗HBV,第一个箭头)。从上到下-抗HBV抗体浓度=100、50、12.5nM。在6000秒时,将不同浓度的PreS1结构域注入孔中,并使用降低的发光信号检测PreS1。误差条表示SD。

[0027] 图13(a-d)。用于改进抗HBV抗体检测的lucCageHBV α 的实验表征。a.lucCageHBV α

的结构模型,带有PreS1表位和lucCage之间预测界面的放大细节。该设计物包括表位PreS1 (a.a.35-46) GANSNPDWDFNGGSGGGSSGFGANSNPDWDFNPN (SEQ ID NO:27630) 的两个拷贝,由柔性接头隔开,以实现与抗体的二价相互作用。SmBit肽以带状表示。b. 检测lucCageHBV α 对抗体HzKR127-3.2 (抗HBV) 存在的敏感性。在连续稀释的HzKR127-3.2存在下,在6000s内测定生物发光。从上到下-lucCageHBV α :lucKey浓度 (nM) = 50:5、5:5、1:10、0.5:0.5。c. 校准曲线的线性区用于确定检测限 (LOD) 和抗体检测的动态范围。d. 使用BioRad ChemiDoc成像系统采集的生物发光图像。从上到下-lucCageHBV α :lucKey浓度 (nM) = 50:5、5:5、1:10。生物发光强度水平的变化被检测为HzKR127-3.2浓度的函数。

[0028] 图14(a-d). 抗SARS-CoV-2抗体传感器的设计和表征。a-b, 针对SARS-CoV-2膜蛋白 (a) 和核衣壳蛋白 (b) 的抗体的从头传感器的实验筛选。膜蛋白的选定表位 (M1、M3和M4;

M1_1-31: MADSNGTITVEELKLLLEQWNLVIGFLFTWI (SEQ ID NO:27659);

M3_1-17: MADSNGTITVEELKLLLE (SEQ ID NO:27660);

M4_8-24: ITVEELKLLLEQWNLVI (SEQ ID NO:27661)) 和核衣壳蛋白 (N6单体 (PKKDKKKKADETQALPQRQKK; SEQ ID NO:27662) 和N62单体 (KKDKKKKADETQAL; SEQ ID NO:27663) 在锁定的不同位置计算接枝到lucCage中。每种设计物包括每种表位的两个串联拷贝,通过柔性连接体分离,以利用抗体的二价结合。在存在抗M兔多克隆抗体 (ProSci, 3527) (a) 或100nM抗N小鼠单克隆抗体 (克隆18F629.1) (b) 的情况下,对所有设计进行实验筛选,以确定每种lucCage设计和20nM lucKey的发光增加。在没有抗体的情况下,发光值归一化为100。由于高灵敏度和稳定性, M3_1-17_334和N62_369-382_340设计物被选为最佳候选,分别命名为lucCageSARS2-M和lucCageSARS2-N。c. 左图: lucCageSARS2-M的结构模型,示出了M3表位和lucCage之间预测界面的放大图。中间图: 测定lucCageSARS2-M (MADSNGTITVEELKLLLEGGSGGMADSNGTITVEELKLLLE (SEQ ID NO:27392)) 对抗M pAb的敏感性。在连续稀释的抗M pAb存在下,在4000秒内测定生物发光。从上到下-lucCageSARS2-M:lucKey浓度 (nM) = 50:50、5:5。右图: 不同浓度下传感器的检测限 (LOD) 计算。d. 左图: lucCageSARS2-N的结构模型,示出了N62表位和lucCage之间预测界面的放大图。中间图: 测定lucCageSARS2-N (KKDKKKKADETQALGGSGGKKDKKKKADETQAL; SEQ ID NO:27548) 对抗N单抗的敏感性。在存在连续稀释的抗N抗体的情况下,在50nM下对lucCageSARS2-N+lucKey在4000s内测定生物发光。右图: 传感器的LOD计算。误差条表示SD。

[0029] 图15(a-e) . a. SARS-CoV-2刺突蛋白受体结合域 (RBD) 从头传感器的实验筛选。在200nM RBD的存在下,对所有设计进行实验筛选,以在20nM的每个lucCage设计物和20nM lucKey下增加发光。在没有RBD的情况下,将发光值归一化为100。由于高灵敏度和稳定性,设计物lucCageRBD Δ 4_348被选为最佳候选,并命名为lucCageRBD。b. lucCageRBD的结构模型,其由接枝到lucCage中的LCB1结合域组成。黑框示出了lucCageRBD设计中笼和LCB1结合域界面的放大图。c. lucCageRBD灵敏度的测定。在连续稀释的RBD蛋白存在下,在10000s内测定生物发光。从上到下-lucCageRBD:lucKey浓度 (nM) = 1:1、1:10、10:10。d. 不同浓度下传感器的检测限 (LOD) 计算。从上到下-lucCageRBD:lucKey浓度 (nM) = 1:1、1:10、10:10。e. 使用BioRad ChemiDoc成像系统采集的生物发光图像。用1nM的lucCageRBD和10nM的lucKey检测作为RBD浓度的函数的生物发光强度水平的变化。

[0030] 图16. 基于LOCKR的生物传感器的一般原理和通过各种拆分的蛋白组件的扩展读

数。

[0031] 图17(a-c). (a) 分子间HBV抗体BRET传感器(S0512)的原理图、发射光谱和BRET比值的剂量依赖性变化。(b) 分子内HBV抗体BRET传感器(B0622)的原理图、发射光谱和标准曲线。为了获得最佳的BRET效率,进行了接头优化。(c) B0622_6的发射光谱和BRET比值对HBV抗体存在的剂量依赖性变化(DFISREVSKGEELIKENMRSK是SEQ ID NO:27655; DFISREEELIKENMRSK是SEQ IDNO:27656; DFISREELIKENMRSK是SEQ ID:27657; DFISREKENMRSK是SEQ ID NO:27658)。使用2nM传感器浓度和20、5、0nM(从左到右)MBP_钥匙。

[0032] 图18. 比色肌钙蛋白I传感器(LacATrop)的示意图、头孢硝噻吩(比色底物)的水解机制以及 β -内酰胺酶活性对人心肌钙蛋白I(cTnI)的剂量依赖性变化。在OD490监测 β -内酰胺酶活性。每个cTnI中 β -内酰胺酶的初始速率计算为 β -内酰胺酶活性。下图显示,在cTnI存在下,溶液中的剂量依赖性颜色从黄色变为红色。

[0033] 图19(a-d). CoV LOCKR诊断。A. 说明了阴性和阳性对照的策略。阴性对照将接受额外过量的合成线性肽表位,以占据可用抗体上的所有表位结合位点。阳性对照样品将含有 lucCage-ProA/lucKey 成分,以测定 IgG 或 IgM 抗体的存在,其中 lucCage 的锁组包含 Fc 结构域抗体结合蛋白 A。B. 功能性阳性对照 lucCage-ProA 组分(已被鉴定且能够在 **Nano-Glo®** 试剂(Promega)的存在下与 lucKey 添加后几分钟内以及与仅含有 lucKey 的缓冲液(黑线)一起来检测多克隆兔 IgG 抗体(中间图)与。右图示出了该系统对低至 10nM IgG 的灵敏度,在 1、10 和 5nM 的不同浓度传感器(lucCage+lucKey)下,用不同浓度的 IgG 孵育时的归一化发光。C. LOCKR 生物传感器特异性评估。将 10nM 的传感器(50nM 的 LucCage SARS2-N)与 50nM 同源靶标(其他生物传感器或缓冲液的靶标)孵育。仅对同源靶标观察到强烈反应。D. POCD CoV LOCKR 装置。预填充在无菌包装中的装置(左)-在一个通道中包括(+) 阳性对照 lucCage ProA/lucKey 试剂,其设计用于在结合 IgG 时激活,(s) 测试样品 lucCage-冠状病毒-表位/lucKey 试剂和(-) 阴性对照试剂,所述阴性对照试剂为 lucCage-冠状病毒-表位/lucKey 加上过量肽表位[~1mM]。

[0034] 图20(a-c). CoV LOCKR 诊断。设计的 LOCKR 配置提供了一种动态“全溶液”测定,以检测表位特异性抗体的存在。A. 一开始, lucCage 表位和 lucKey 蛋白存在于处于“关闭”状态的黑暗溶液中。B. 添加含有能够结合目标表位的抗体的流体后,暴露 lucCage 的锁结合界面,允许 lucKey 结构域结合,定位融合的大块拆分荧光素酶与小块拆分荧光素酶结合。这导致荧光素酶发光的重组(“打开”)。C. 添加含有目标表位的重组抗原将改变抗体从锁到抗原的结合平衡,导致拆分荧光素酶活性的重组减少,从而导致昏暗的发光度(“Bim”)。

[0035] 图21. 间接检测。本公开的传感器平台可以重新调整用途,以适应“间接检测”方法,其中通过将生物传感器与靶标结合多肽的靶标(例如抗 HBV 抗体)预孵育,重组拆分报告蛋白(分子间或分子内实施方案;分子间实施方案如图21所示),在该示例中导致荧光激活。然后将激活的生物传感器与样品一起孵育,以检测抗体结合的抗原(在本例中为乙型肝炎病毒抗原(PreS1))的存在,导致抗体与抗原结合,分离的报告蛋白组分之间的相互作用丧失,报告活性降低/消除(在这种情况下,荧光活性丧失)。

[0036] 图22. CoV LOCKR 诊断的对照样品。A. 说明了阴性和阳性对照的策略。阴性对照将接受额外过量的合成线性肽表位,以占据样品中可用抗体上的所有表位结合位点。而阳性对照样品将含有 lucCage ProA/lucKey 组分,以测定 IgG 或 IgM 抗体的存在。其中 lucCage 的

锁组分包含Fc结构域抗体结合蛋白A.B.功能性阳性对照lucCage-ProA组分已被鉴定(中间图),并且能够在**Nano-Glo®**试剂(Promega)存在下与lucKey在添加后几分钟内以及与仅含有lucKey的缓冲液(黑线)一起来检测多克隆兔IgG抗体。右图示出了该系统对低至10nM IgG的灵敏度,在1、10和5nM的不同浓度传感器(lucCage+lucKey)下,用不同浓度的IgG孵育时的归一化发光。

具体实施方式

[0037] 所引用的所有参考文献通过引用全部并入本文。在本申请中,除非另有说明,否则所使用的技术可以在以下几种已知参考文献中找到:Molecular Cloning:A Laboratory Manual (Sambrook, et al., 1989, Cold Spring Harbor Laboratory Press), Gene Expression Technology (Methods in Enzymology, Vol. 185, edited by D. Goeddel, 1991, Academic Press, San Diego, CA), “Guide to Protein Purification” in Methods in Enzymology (M.P. Deutscher, ed., (1990) Academic Press, Inc.); PCR Protocols: A Guide to Methods and Applications (Innis, et al. 1990, Academic Press, San Diego, CA), Culture of Animal Cells: A Manual of Basic Technique, 2nd Ed. (R.I. Freshney, 1987, Liss, Inc. New York, NY), Gene Transfer and Expression Protocols, pp. 109-128, ed. E.J. Murray, The Humana Press Inc., Clifton, N.J.), and the Ambion 1998 Catalog (Ambion, Austin, TX)。

[0038] 如本文所用,除非上下文另有明确规定,否则单数形式“a”、“an”和“the”包括复数指称。

[0039] 如本文所用,氨基酸残基缩写如下:丙氨酸(Ala;A)、天冬酰胺(Asn;N)、天冬氨酸(Asp;D)、精氨酸(Arg;R)、半胱氨酸(Cys;C)、谷氨酸(Glu;E)、谷氨酰胺(Gln;Q)、甘氨酸(Gly;G)、组氨酸(His;H)、异亮氨酸(Ile;I)、亮氨酸(Leu;L)、赖氨酸(Lys;K)、蛋氨酸(Met;M)、苯丙氨酸(Phe;F)、脯氨酸(Pro;P)、丝氨酸(Ser;S)、苏氨酸(Thr;T)、色氨酸(Trp;W)、酪氨酸(Tyr;Y)和缬氨酸(Val;V)。

[0040] 在本文公开的多肽的所有实施方案中,任何N端蛋氨酸残基都是任选的(即:可以存在或可以不存在)。

[0041] 本公开的任何方面的所有实施方案都可以组合使用,除非上下文另有明确规定。

[0042] 除非上下文另有明确要求,否则在整个说明书和权利要求中,词语“包括(comprise)”、“包括(comprising)”等应被解释为包容性,而不是排他性或穷尽性;也就是说,为“包括但不限于”的意义。使用单数或复数的单词也分别包括复数和单数。此外,在本申请中使用的词语“此处”、“上面”和“下面”以及具有类似含义的词语应指本申请的整体,而非本申请的任何特定部分。

[0043] 在第一方面,本公开提供了包括螺旋束的笼蛋白,其中所述笼蛋白包括结构区和锁区,其中所述锁区包括一种或多种靶标结合多肽,其中笼蛋白还包括第一报告蛋白结构域,其中所述第一报告蛋白结构域在结合到第二报告蛋白结构时经历可检测的报告活性变化,并且其中所述结构区与所述锁区相互作用以防止溶液进入所述一种或多种靶标结合多肽。

[0044] 笼蛋白及其在蛋白开关中的应用一般描述于美国专利申请公开

No.US20200239524中,其全文通过引用并入本文。本公开通过结合报告分子和一种或多种靶标结合多肽,对此类笼蛋白和蛋白开关提供了显著改进,允许用作模块化和通用化生物传感器平台,所述平台能够实现如本文所公开的用于不同传感目的的广泛读数。

[0045] 笼多肽包括闩锁区和结构区(即,不是闩锁区的笼多肽的剩余部分)。闩锁区可以存在于笼多肽的任一端附近。在一个实施方案中,闩锁区位于C端螺旋线处。在各种实施方案中,闩锁区可以包括笼状多肽中N端或C端部分的单个 α -螺旋的一部分或全部。在各种其他实施方案中,闩锁区可以包括笼状多肽中的第一、第二、第三、第四、第五、第六或第七 α 螺旋的一部分或全部。在其他实施方案中,闩锁区可以包括笼多肽中两个或多个不同 α 螺旋的全部或部分;例如,一个 α 螺旋的C端部分和下一个 α 螺旋的N端部分、所有两个连续的 α 螺旋等。

[0046] 这些实例提供了关于示例性笼蛋白和报告活性的广泛细节。可以使用任何合适的报告蛋白结构域,所述报告蛋白结构域涉及两个单独的蛋白组分(例如,如本文所述的BRET和FRET格式),或可拆分成两个(或多个)蛋白结构域的报告蛋白,当两个(或者多个)拆分蛋白结构域结合时,其活性可被重组。

[0047] 可检测的变化可以是相关报告活性的任何增加或减少,视为适合预期目的。以下在讨论本公开的生物传感器时以及在示例中描述了可利用的报告活性的可检测变化的各种非限制性实施方案。

[0048] 在一个实施方案中,笼蛋白还包括第二报告蛋白结构域,其中第一报告蛋白结构域和第二报告结构域中的一个存在于闩锁区中,另一个存在于结构区中,其中在所述一种或多种靶标结合多肽结合的靶的存在下,所述第一报告蛋白结构域和所述第二报告蛋白结构的相互作用减弱。

[0049] 在另一个实施方案中,第二报告蛋白结构域不存在于笼蛋白中,而是存在于另一种组分(即:“钥匙”,如下所述)中,或者可以存在于其他地方。

[0050] 在一个实施方案中,笼蛋白螺旋束包括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或9个 α 螺旋。

[0051] 在另一个实施方案中,笼蛋白结构区中的每个螺旋长度可以独立地在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或40-45个氨基酸之间。

[0052] 在另一个实施方案中,由于在闩锁区内存在一种或多种靶标结合多肽,闩锁区可以在本公开的设计中延伸,因此一个 α -螺旋/多个 α -螺旋在闩锁区中可以显著长于结构区中的 α -螺旋,仅受闩锁中存在的靶标结合多肽的长度的限制。

[0053] 在任何这些实施方案中,笼蛋白中的相邻 α 螺旋可以任选地通过氨基酸接头连接。连接每个 α -螺旋的氨基酸接头可以具有任何适合的长度或适合预期用途的氨基酸组成。在一个非限制性实施方案中,每个氨基酸接头的长度独立地在2至10个氨基酸之间,不包括可与接头融合的任何其他功能序列。在各种非限制性实施方案中,每个氨基酸接头的长度独立地为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或10个氨基酸。在所有实施方案中,接头可以是结构化的或柔性的(例如聚GS)。这些接头可编码其他功能序列,如认为适合预期用途。

[0054] 闩锁区可以存在于笼蛋白上被认为适合于预期目的的任何合适位置。在一个实施方案中,闩锁区位于笼蛋白的C端。在另一个实施方案中,闩锁区可以位于笼蛋白的N端。

[0055] 类似地,第一报告蛋白结构域可以存在于笼蛋白上被认为适合于预期目的的任何合适位置。在一个实施方案中,第一报告蛋白结构域存在于闩锁区中。在一个实施方案中,第一报告蛋白结构域位于闩锁区的C端或闩锁区C端的20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸内。在另一个实施方案中,第一报告蛋白结构域或位于闩锁区N端处或在闩锁区N端的20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸内。

[0056] 在另一个实施方案中,第二报告蛋白可以存在于笼蛋白中;在该实施方案中,第二报告蛋白结构域可以存在于结构区中。在一个这样的实施方案中,第二报告蛋白可以存在于结构区的N端处,或者可以在结构区N端的10、9、8、7、6、5、4、3、2或1个氨基酸内。

[0057] 笼蛋白包含一种或多种(即1、2、3等)靶标结合多肽。在一个实施方案中,笼蛋白包含一种靶标结合多肽。在另一个实施方案中,笼蛋白包括两种靶标结合多肽。在一个实施方案中,一种或多种靶标结合多肽和第一报告蛋白结构域在笼蛋白的闩锁区中被至少10个氨基酸分开。在另一个实施方案中,一种或多种靶标结合多肽位于闩锁区C端的10、9、8、7、6、5、4、3、2或1个氨基酸处或其内。

[0058] 可以使用任何合适的报告蛋白结构域,所述报告蛋白结构域涉及两种单独的蛋白组分(例如,如本文所述的BRET和FRET形式),或可拆分成两种(或多种)蛋白结构域的报告蛋白,当两种(或者多种)拆分蛋白结构域结合时,其活性可被重组。在一个实施方案中,当存在于笼蛋白中时,第一报告蛋白结构域和第二报告蛋白结构域包括报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子,荧光共振能量转移(FRET)报告分子、比色报告分子(包括但不限于 β -内酰胺酶、 β -半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶)、和分子条形码报告分子(包括但不限于TEV蛋白酶)。

[0059] 在一个实施方案中,笼蛋白不包括第二报告结构域。在一个这样的实施方案中,第一报告蛋白结构域包括:

(a) 与SEQ ID NO:27359和27664-27672的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:VTGYRLFEEIL(SmBit)(SEQ ID NO:27359),VTGYRLFEEKIL(SEQ ID NO:27664),VTG.YRLFEEKIS(SEQ ID NO:27665),VSGWRLFEEKIS(SEQ ID NO:27666),VEGYRLFEEKIS(SEQ ID NO:27667),VTGYRLFEEKES(SEQ ID NO:27668),VTGWRLFEEKIL(SEQ ID NO:27669),VTGWRLFKEIL(SEQ ID NO:27670),VTGYRLFKEIL(SEQ ID NO:27671),LAGWRLFEEKIS(SEQ ID NO:27672);

[0060] (b) 与选自SEQ ID NO:27360-27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRP EERFPMSTF KVLCCGAVLS

RVDAGQEQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKEFLAFL
HNMGDHSVTRL DRWEPELNEA IPNDERDTT PAAMATTLRK LLTGENGR (拆分 β -内酰胺酶A; SEQ ID
NO:27360) 和LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW (拆分 β -内酰胺酶B; SEQ ID NO:27361);

(c) 与选自SEQ ID NO:27362-27378的氨基酸序列至少70%、75%、80%、85%、
90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中下划
线的残基是氨基酸接头或其他可以存在或不存在的可选残基,并且当存在时,可以是氨基
酸序列,并且其中任何N端蛋氨酸残基可以存在或不存在:

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSG
DQMGIKIFKVVYPVDNHFKVILHYGTLVIDGVTNMIYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLI
NPDGSLFRVTINGVTGWRHERILA (TeLuc; SEQ ID NO:27362) (可用于创建FRET和/或BRET传感
器的全发光或荧光蛋白);

LIK ENMRKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFDILATHFMYGS
KVFYKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL ERIKEADNET
YVEQYEHAVARYSNLGGMD ELYK (CyOFP变体; SEQ ID NO:27363) (可用于创建FRET和/或BRET传
感器的全发光或荧光蛋白);

VSKGEELIK ENMRKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFYKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CyOFP变体; SEQ ID NO:27364) (可用于创建FRET
和/或BRET传感器的全发光或荧光蛋白);

EELIK ENMRKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFYKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CuOFP变体; SEQ ID NO:27365) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

KVFTLGDVFDWRQTAGYNQAQVLEQGGTSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLS
CDQMAQIEKIFKVVYPVDDHFKAILHYGTLVIDGVTNMIYFGQPYEGIAKFDGKKITVTGTLWNGNTIIDERL
INPDGSLFRVTINGVTGWRHERILA (LumiLuc; SEQ ID NO:27366) (可用于创建FRET和/或BRET
传感器的全发光或荧光蛋白);

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGGGTGNP NDGYEELNLK STKGDLQFSP
WILVPHIGYG FHQYLPYDPG MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT
GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP
MYVFRKTELK HSKTELNFKE WQKAFTDVMG MDELYK (mNeonGreen; SEQ ID NO:27367) (可用于创
建FRET和/或BRET传感器的全发光或荧光蛋白);

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGAVTQT DTSLEDGLTI YKVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL

DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i; SEQ ID NO:27368) (可用于创建 FRET和/或BRET传感器的全发光或荧光蛋白);

SGKSYPTVSADYQKAVEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTFDKRKTGGPFGTI
RYPaelahsAnsgLDIAVRLLEPLKAefPILSYADfyQLAGVVAVEVTGGPEVPFHPGREDK
PELPPEGRLPDATKGSdHLRDVFGKAMGLTDQDIVALSGGHTLGAaHKERSGFEGPWTSN
PLVFDNSYFTELLSgeKEGGGGSGGGGS (APEX2-1-200; SEQ ID NO:27369);

GGGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSelGFAD
 A (APEX2-201-250; SEQ ID NO:27370);

MGSHHHHHHGSGSENLYFQSGGS

VRPLNCIVA VSQNMIGIKN GDLWPPLRN ESKYFQRMTT TSSVEGKQNL
 VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRGaHFL AKSLDDALRL
IEQPELGGGGSGGGGS (DHFR_A (1-105); SEQ ID NO:27371);

SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRE
GGGGSGGGGS ASKV DMVWIVGGSS VYQEAMNQPG HLRLFVTRIM QEFESDTFFP
 EIDLgKYKLL PEYpGVlSEV QEEKGiKYKF EVYEKKD (DHFR_B (106-186); SEQ ID
 NO:27372);

QLTPTFYDnSCPnVSNIVRDIIvNELRSDPRIAASILRLHFHDCfVNGCDASILLDNTTSFRT
EKDAFGNANSARGFSVIDR

MKAAVESACPGTVSCADLLTIAAQQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFTLpQLKD
 SFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGGSGS (sHRPa是大的
 HRP拆分片段,它由具有以下4种突变的辣根过氧化物酶 (HRP) 的1-213个氨基酸组成:
 T21I,P78S,R93G,N175S) (SEQ ID NO:27373);

NLSALVDFDLRTPTIFDNKYVNLeeQKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFfNAFVEAM
 DRMGnITPLTGTQgQIRNCRVvNSNGGSGS (sHRPb是小的HRP拆分片段,它由具有以下2种突变辣
 根过氧化物酶 (HRP) 的214-308个氨基酸组成:N255D,L299R) (SEQ ID NO:27374);

GESLfkGPRDYNpISStICHlTnESDGHtTSLYgIGfGPFiITNKHLfRRNNGtLLVQSLHGv
FKVKNTTTTLQqHLIDGRDMIIRMPKDFPPFPQKLkFREpQREERICLVtTNfQTGGGGSGG
GGs (N_Tev (1-118) (SEQ ID NO:27375);

GGGGSGGGGSKSMSSMVSDTSCTFPSSDGIFWKHWIQTkdGQCgSPLVSTRDGFIVGIHS
ASNfTNTNNYfTsvPKnfMELLTnQEAQQWvSGWRLNADSVLWGGHKVFMDKP C_Tev
 (119-221) (SEQ ID NO:27376);

MAsYPCHQHA SAFDQAARSr GHSNRRTALR PRRQqEATEV RLEqKMPTLL RvYIDGPHGM
 GKTTTTQLLV ALGSRDDIVY VPEPMtyWQV LGASetiANI YTTQHRLDQg EISAGDAAVV MTSaQITMGM
 PYAVTDAVLA PHIGGEAGSS HAPPPALTLI FDRHPiAALL CyPAARYLMG SMTpQAVLAF VALIPPTLPG
 TNIVLgALPE DRHIDRLAKR QRpGERLdLA MLaAIRRVYG LLANTVRYLQ GGGSwREDWG QLSGT

GGGSGGGGS(胸苷激酶_TK_A(1-265)_(SEQ ID NO:27377);和/或

GGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP NGDLYNVFAW
ALDVLAKRLR PMHVFILDYD QSPAGCRDAL LQLTSGMVQT HVTTPGSIPT ICDLARTFAR EMGEAN(胸苷
激酶_TK_B(266-376)_(SEQ ID NO:27378)

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSG
DQMGIKIEKIFKVVYPVDNHHFKVILHYGTLVIDGVTNPMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLI
NPDGSLFRVTINGVTGWRHERILA(TeLuc;SEQ ID NO:27362)(可用于创建FRET和/或BRET传感
器的全发光或荧光蛋白);

LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL ERIKEADNET
YVEQYEHAVA RYSNLGGMD ELYK(CyOFP变体;SEQ ID NO:27363)(可用于创建FRET和/或BRET
传感器的全发光或荧光蛋白);

VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK(CyOFP变体SEQ ID NO:27364)(可用于创建FRET
和/或BRET传感器的全发光或荧光蛋白);

EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK(CuOFP变体;SEQ ID NO:27365)(可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

KVFTLGDVFDWRQTAGYNQAQVLEQGGTSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLS
CDQMAKIEKIFKVVYPVDDHHFKAILHYGTLVIDGVTNPMIDYFGQPYEGIAKFDGKKITVTGTLWNGNTIIDERL
INPDGSLFRVTINGVTGWRHERILA(LumiLuc;SEQ ID NO:27366)(可用于创建FRET和/或BRET
传感器的全发光或荧光蛋白);

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVQGTGNP NDGYEELNLK STKGDLQFSP
WILVPHIGYG FHQYLPYPDG MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT
GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP
MYVFRKTELK HSKTELNFKE WQKAFTDVMG MDELYK(mNeonGreen;SEQ ID NO:27367)(可用于创
建FRET和/或BRET传感器的全发光或荧光蛋白);

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGAVTVTQ DTSLEDGTLI YKVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL
DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK(mScarlet-i;SEQ ID NO:27368)(可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

SGKSYP TVSADYQKAVEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTFDKRKTGGPFGTI
RYP AELAH SANSGLDI AVRLL EPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDK
PELPPEGRLPDATKGS DHLRDVFGKAMGLTDQDIVALS GGHTLGA AHKERSGFEGPWTSN
PLVFDNSYFTELLS GEKEGGGGSGGGGS (APEX2-1-200; SEQ ID NO:27369);

GGGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFAD
 A (APEX2-201-250; SEQ ID NO:27370);

MGSHHHHHHGSGSENLYFOGSGGS

VRPLNCIVA VSQNMIGIGN GDLPPWPLRN ESKYFQRMTT TSSVEGKQNL
 VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRG AHFL AKSLDDALRL
IEQPELGGGGSGGGGS (DHFR_A (1-105); SEQ ID NO:27371);
 SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRE
GGGGSGGGGS ASKV DMVWIVGGSS VYQEAMNQPG HLRLFVTRIM QEFESDTFFP
 EIDLKGYKLL PEYPGVLSEV QEEKGIKYKF EVYEKKD (DHFR_B (106-186); SEQ ID
 NO:27372);

QLTPTFYDN SCPNVSNIVRDIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRT
 EKDAFGNANSARGFSVIDR

MKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFLLANANLPAPFFTLPLQKD
 SFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDP TLNTTYLQTLRGLCPLNGGSGS (sHRPa是大的
 HRP拆分片段。它由具有以下4种突变的辣根过氧化物酶 (HRP) 的1-213个氨基酸组成：
 T21I, P78S, R93G, N175S) (SEQ ID NO:27373) ;

NLSALVDFDLRTP TIFDNKYVYNLEEKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAVVEAM
 DRMGNITPLTGTGGQIRRNCRVNSNGGSGS (sHRPb是小的HRP拆分片段。它由具有以下2种突变的
 辣根过氧化物酶 (HRP) 的214-308个氨基酸组成：N255D, L299R) (SEQ ID NO:27374) ;

GESLFK GPRDYNPISSTICHLT NESDGHTTSLYGIGFGFPFIITNKHLFRRNNGTLLVQSLHGV
FKVKNTTTLQQHLIDGRDMIIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTGGGGSGG
GGG (N_Tev (1-118) (SEQ ID NO:27375);

GGGGSGGGGSKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHS
ASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMDKP C_Tev
(119-221) (SEQ ID NO:27376);

MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQQEATEV RLEQKMPTLL RYVIDGPHGM
 GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV LGASETIANI YTTQHRLDQG EISAGDAAVV M TSAQITMGM
 PYAVTDAVLA PHIGGEAGSS HAPPPALTLI FDRHPAALL CYP AARYLMG SMTPQAVLAF VALIPPTLPG
 TNIVLGALPE DRHIDRLAKR QRPGERDLA MLAAIRRVYG LLANTVRYLQ GGGSWREDWG
 QLSGTGGGGSGGGGS (胸苷激酶_TK_A (1-265) (SEQ ID NO:27377) ; 和/或

GGGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP NGDLYNVFAW
 ALDVLAKRLR PMHVFILDYD QSPAGCRDAL LQLTSGMVQT HVTTPGSIPT ICDLARTFAR EMGEAN (胸苷

激酶_TK_B(266-376)_(SEQ ID NO:27378)

[0061] 包含报告蛋白结构域的笼蛋白的这一实施方案将在靶标分析物的存在下与包含第二报告结构域的第二生物传感器组分“钥匙”蛋白(下文讨论)相互作用。

[0062] 在另一个实施方案中,笼包含第二报告蛋白结构域,其中(a)第一报告蛋白结构域和第二报告蛋白结构域之一包含与SEQ ID NO:27359和27664-27672的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

[0063] 另一个包含与SEQ ID NO:27379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸残基可以存在或不存在:

**MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRIVRSGENALKIDIHIP
YEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFD
GKKITVTGTLWNGNKIIDERLITPDGSMFLFRVTINS (LgBiT) (SEQ ID NO:27379);**

[0064] (b)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27360的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列

VFAHPETL VKVKDAEDQLGA RVGYIELDLN SGKILESFRP EERFPMSTF KVLICGAVLS
RVDAGQEQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
HNMGDHVTRL DRWEPNELNEA IPNDERDTT PAAMATTLRK LLTGNGR (拆分β-内酰胺酶A) (SEQ ID
NO:27360),

[0065] 另一个包含与SEQ ID NO:27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW (拆分β-内酰胺酶B) (SEQ ID NO:27361);

[0066] (c)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27362的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVPPIQRIVLSGENGLKIDIHVIIPYEGLSG
DQMGQIEKIFKVVYPVDNHFKVILHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLI
NPDGSLFRVTINGVTGWRLHERILA (TeLuc) (SEQ ID NO:27362) (可用于创建FRET和/或BRET传
感器的全发光或荧光蛋白),

[0067] 另一个包含与选自SEQ ID NO:27363-27365的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

LIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD ILATHFMYGS
KVFYKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL ERIKEADNET
YVEQYEHAVA RYSNLGGMD ELYK (CyOFF变体) (SEQ ID NO:27363) (可用于创建FRET和/或BRET
传感器的全发光或荧光蛋白);

VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD

ILATHFMYGS KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPKVM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CyOFP变体) (SEQ ID NO:27364) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);和

EELIK ENMRKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV EGGPLPFAFD
ILATHFMYGS KVFIKYPADL PDYFKQSFPE GFTWERVMVF EDGGVLTATQ DTSLQDGELI YNVKVRGVNF
PANGPVMQKK TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPKVM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (CyOFP变体) (SEQ ID NO:27365) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

[0068] (d)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID
NO:27366的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、
96%、97%、98%或100%相同的氨基酸序列:

KVFTLGDVFGDWRQTAGYNQAQVLEQGGLTSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLS
CDQMAQIEKIFKVVYPVDDHHFKAILHYGTLVIDGVTNPMIDYFGQPYEGIAKFDGKKITVTGTLWNGNTIIDERL
INPDGSLFRVTINGVTGWRLHERILA (LemiLuc) (SEQ ID NO:27366) (可用于创建FRET和/或BRET
传感器的全发光或荧光蛋白),

[0069] 另一个包含与SEQ ID NO:27368的氨基酸序列至少70%、75%、80%、85%、90%、
91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸
残基可以存在或不存在:

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGGAVTVTQ DTSLEDGTLI YKVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL
DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i) (SEQ ID NO:27368) (可用于创建
FRET和/或BRET传感器的全发光或荧光蛋白);

[0070] (e)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID
NO:27367的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、
96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在:

MVSKGEEDNM ASLPATHELH IFGSINGVDF DMVGGGTGNP NDGYEELNLK STKGDLQFSP
WILVPHIGYG FHQYLPYPDG MSPFQAAMVD GSGYQVHRTM QFEDGASLTV NYRYTYEGSH IKGEAQVKGT
GFPADGPVMT NSLTAADWCR SKKTYPNDKT IISTFKWSYT TGNGKRYRST ARTTYTFAKP MAANYLKNQP
MYVFRKTELK HSKTELNFKE WQKAFTDVMG MDELYK (mNeonGreen) (SEQ ID NO:27367) (可用于创
建FRET和/或BRET传感器的全发光或荧光蛋白),

[0071] 另一个包含与SEQ ID NO:27368的氨基酸序列至少70%、75%、80%、85%、90%、
91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸
残基可以存在或不存在:

MVSKGEAVIK EFMRFKVHME GSMNGHEFEI EGEGERPYE GTQTAKLKVT KGGPLPFSWD
ILSPQFMYGS RAFIKHPADI PDYKQSFPE GFKWERVMNF EDGGAVTVTQ DTSLEDGTLI YKVKLRGTNF
PPDGPVMQKK TMGWEASTER LYPEDGVLKG DIKMALRLKD GGRYLADFKT TYKAKKPVQM PGAYNVDRKL
DITSHNEDYT VVEQYERSEG RHSTGGMDEL YK (mScarlet-i) (SEQ ID NO:27368) (可用于创建

FRET和/或BRET传感器的全发光或荧光蛋白)；

[0072] (f)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与氨基酸序列SEQ ID NO:27369至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

SGKSYPTVSADYQKAVEKAKKRLGGFIAEKRCAPLMLRLAWHSAGTFDKRKTGGPFGTIRYPAELAH
SANSGLDIAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKPELPPEGRLPDATKGS DHLRDV
FGKAMGLTDQDIVALSGGHTLGAAHKERSGFEGPWTSNPLVFDNSYFTELLSGEKEGGGGSGGGGS (APEX2-1-
200) (SEQ ID NO:27369) (用于化学发光和比色检测系统的大豆抗坏血酸过氧化物酶蛋白
的拆分工程化变体)；

[0073] 另一个包含与SEQ ID NO:27370的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGGSGGGGS GLLQLPSDKALLSDPVFRPLVDKYAAEDDAFFADYAEAHQKLSLGFADA (APEX2-
201-250) (SEQ ID NO:27370) (用于化学发光和比色检测系统的大豆抗坏血酸过氧化物酶
蛋白的拆分工程化变体)；

[0074] (g)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27371的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

MGSHHHHHGSGSENLYFQSGGS VRPLNCIVA VSQNMGIGKN GDLWPPLRN ESKYFQRMTT
TSSVEGKQNL VIMGRKTWFS IPEKNRPLKD RINIVLSREL KEPPRGAHFL AKSLDDALRL
IEQPELGGGGSGGGGS (DHFR_A(1-105))；(SEQ ID NO:27371) (用于细胞存活或荧光的拆分二
氢叶酸还原酶蛋白报告分子)

[0075] 另一个包含与SEQ ID NO:27372的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

SGSG DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISREGGGGS
GGGGS ASKV DMVWIVGGSS VYQEAMNQPGLHRLFVTRIM QEFESDTFFP EIDLKGYKLL PEYPGVLSEV
QEEKGIKYKF EVYEKKD (DHFR_B(106-186))；(SEQ ID NO:27372) (用于细胞存活或荧光的拆
分二氢叶酸还原酶蛋白报告分子)；

[0076] (h)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27373的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

QLTPTFYDNSCPNVSNIVRDIIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAF
GNANSARGFSVIDRMKAAVESACPGTVSCADLLTIAAQSVTLAGGPSWRVPLGRRDSLQAFDLANANLPAPFFT
LPQLKDSFRNVGLNRSSDLVALSGGHTFGKSQCRFIMDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGSGS
(sHRPa是大的HRP拆分片段。它由具有以下4种突变的辣根过氧化物酶(HRP)的1-213个氨基

酸组成:T21I,P78S,R93G,N175S:质粒73147(SEQ ID NO:27373);

[0077] 另一个包含与SEQ ID NO:27374的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

NLSALVDFDLRPTIFDNKYVNLLEEQKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAM
DRMGNITPLTGTQGQIRRNCRVVNSNGGSGS(sHRPb是小的HRP拆分片段。它由具有以下2种突变的辣根过氧化物酶(HRP)的214-308个氨基酸组成:N255D,L299R:质粒73148)(SEQ ID NO:27374);

[0078] (i)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27375的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GESLFKGPDRDYNPISSTICHLTNE SDGHTTSLYIGIGFPGPIITNKHLFRRNNGTLLVQSLHGVFKVKN
TTTTLQQHLIDGRDMIIRMPKDFPPFPQKLFREPQREERICLVTTNFQTGGGSGGGGSN_Tev(1-118)(SEQ ID NO:27375)(拆分TEV蛋白酶);

[0079] 另一个包含与SEQ ID NO:27376的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGSGGGGSKSMSSMVSDTCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHSASNFTNTN
NYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMDKP(C_Tev(119-221));(SEQ ID NO:27376)(拆分TEV蛋白酶);

[0080] (j)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27377的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且其中N端蛋氨酸残基可以存在或不存在的:

MASYPCHQHA SAFDQAARSR GHSNRRTALR PRRQQEATEV RLEQKMPTLL RVIYIDGPHGM
GKTTTTQLLV ALGSRDDIVY VPEPMTYWQV LGASETIANI YTTQHRLDQG EISAGDAAV MTSAQITMGM
PYAVTDAVLA PHIGGEAGSS HAPPPALTLI FDRHPAALL CYPAAARYLMG SMTPQAVLAF VALIPPTLPG
TNIVL GALPE DRHIDRLAKR QRPGERDLA MAAIRRVYV LLANTVRYLQ GGGSWREDWG QLSGT
GGGSGGGGS(胸苷激酶_TK_A(1-265))(SEQ ID NO:27377);

[0081] 另一个包含与SEQ ID NO:27378的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列

GGGSGGGGS AVPPQ GAEPQSNAGP RPHIGDTLFT LFRAPPELLAP NGDLYNVFAW
ALDVLAKRLR PMHVFILDYD QSPAGCRDAL LQLTSGMVQT HVTTPGSIPT ICDLARTFAR EMGEAN(胸苷
激酶_TK B(266-376))(SEQ ID NO:27378)

[0082] 包含两种报告蛋白结构域的笼蛋白的这些实施方案在靶标分析物存在下与第二生物传感器组分“钥匙”相互作用。这种相互作用引起的构象变化使得笼蛋白中的两种报告

蛋白接近(或远离),从而允许通过测定报告信号的增加(或减少)来定量分析物。

[0083] 结合目标靶标的任何合适的靶标结合多肽可用于本公开的笼蛋白中,如认为适合预期用途。如上所述,笼蛋白可包含1、2、3、4或更多种靶标结合多肽,如本文所例示。在一个实施方案中,笼蛋白包含1种靶标结合多肽。在另一个实施方案中,笼蛋白包含2种或更多种靶标结合多肽,每种靶标结合多肽可以相同或不同。

[0084] 类似地,一种或多种靶标结合多肽的靶标可以是适合于一种或多种靶标结合多肽可用的预期目的的任何靶标。在一个非限制性实施方案中,所述一种或多种靶标结合多肽能够与靶标结合,所述靶标包括但不限于抗体、毒素、诊断生物标志物、病毒颗粒、疾病生物标志物、代谢物或目标生化分析物。在存在2种或多种靶标结合多肽的实施方案中,每种靶标结合多肽可以结合相同的靶标,或者可以独立地结合不同的靶。在2种或多种靶标结合多肽结合到同一靶标的实施方案中,它们可以结合到靶的相同区(例如,为相互作用增加亲和力),或者可以结合到靶标的不同区。

[0085] 如本领域技术人员将理解的,一种或多种靶标结合多肽可以包括任何类型的多肽,包括但不限于从头设计的蛋白、亲和抗体、Affimer、锚蛋白重复蛋白(天然存在或设计的)、纳米抗体等。

[0086] 在一个实施方案中,一种或多种靶标结合多肽能够结合抗体靶。在另一个实施方案中,所述一种或多种靶标结合多肽包括由针对病毒靶标的抗体识别的一种或多种表位。在另一个实施方案中,一种或多种靶标结合多肽包括一种或多种被抗SARS-Cov-2抗体识别的表位。在本文描述的各种其他实施方案中,所述一个或多个靶标结合多肽能够结合到疾病标志物或毒素、Bcl2、Her2受体、肉毒杆菌神经毒素B、心肌钙蛋白I、白蛋白、表皮生长因子受体、前列腺特异性膜抗原(PSMA)、瓜氨酸肽(citrullinated peptide)、脑钠肽或任何其他合适的靶标。

[0087] 在各种非限制性实施方案中,一种或多种靶标结合多肽包含与选自SEQ ID NO: 27380-27430的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

[0088] 表1. 示例性靶标结合多肽合多肽

生物传感器	传感结构域	传感结构域序列
lucCageBim	Bim	EIWIAQELRRIGDEFNAYYAAA (SEQ ID NO:27380)
lucCageBoT	Bot.0671.2	MFAELKAKFFLEIGDRDAARNALRKA GYSDEEAERIIRKYELE (SEQ ID NO:27381)
lucCageProA	蛋白 A 结合域 C (SpaC)	EQQNAFYEILHLPNLTEEQRNGFIQSLK DDPSVSKEILAEAKKLNDAQAPK (SEQ ID

[0089]

		NO:27382)
lucCagHer2	Her2 亲和抗体	EMRNAYWEIALLPNLNNQKRAFIRSL YDDPSQSANLLAEAKKLNDAPK (SEQ ID NO:27383)
lucCageTrop	cTnI + cTnC	EDQLREKAKELWQTIYNLEAEKFDLQE KFKQQKYEINVLNRINDNQKVSCTKDDSK GKSEEELSDLFRMFDKNADGYIDLEELKIM LQATGETITEDDIEELMKDGDKNNDGRIDY DEFLEFMKGVE (SEQ ID NO:27384) - cTnTf1:226-EDQLREKAKEL WQTI-240 (SEQ ID NO:27385) - cTnTf2:226-EDQLREKAKEL WQTIYN-242 (SEQ ID NO:27386) - cTnTf3:226- EDQLREKAKELWQTIYNLEAE-24 6 (SEQ ID NO:27387) - cTnTf4:226-EDQLREKAKEL WQTIYNLEAEKFD-249 (SEQ ID NO:27388) - cTnTf5:226-EDQLREKAKEL WQTIYNLEAEKFDLQE-252 (SEQ ID NO:27389) - cTnTf6:226-EDQLREKAKEL WQTIYNLEAEKFDLQEKFKQQKY EINVLNRINDNQ-272 (SEQ ID NO:27390) - EDQLREAAKELWQTIYNLEAEK FDLQEKFKQQKYEINVLNRINDNQKVS KTKDDSKGKSEEELSDLFRMFDKNADG YIDLEELKIMLQATGETITEDDIEELMKD GDKNNDGRIDYDEFLEFMKGVE (SEQ ID NO:27391)

lucCageSARS 2-M	SARS-CoV-2 核 衣壳蛋白(a.a. 369-382) 2x	MADSNGTITVEELKKLLEGGSGGMAD SNGTITVEELKKLLE (SEQ ID NO:27392) - MADSNGTITVEELKKLL EQWNLVIGFLFLTWIGGSGGM ADSNGTITVEELKKLLEQWNL VIGFLFTWI (SEQ ID NO:27393) - ITVEELKKLLEQWNLVI GGSGGITVEELKKLLEQWNLV I (SEQ ID NO:27394)
lucCageSARS 2-N	SARS-CoV-2 膜 蛋白(a.a. 1-17) 2x	N62:KKDKKKKADETQALGGSGGKKD KKKKADETQAL (SEQ ID NO:27548 N6:PKDKKKKADETQALPQRQKKGGS GGPKDKKKKADETQALPQRQKK (SEQ ID NO:27547)
sCageHA	HB1.9549.2	TFACRIAARIAAEFGYSEEQIKELLKNA GCSEDEARDAVEYLR (SEQ ID NO:27396)

>LCB1-1

DKEWILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEE (SEQ ID NO:27397)

>LCB1-2

DKEEILNKIYEIMRLLDELGNAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEE (SEQ ID NO: 27398)

>LCB1-3

DKEWILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKQGDERLLEEAERLLEEVEE (SEQ ID NO: 27399)

>LCB1-4

DKENILQKIYEIMKTLDQLGHAEASMQVSDLIYEFMKQGDERLLEEAERLLEEVEE (SEQ ID NO: 27400)

>LCB1-5

DKENILQKIYEIMKTLDQLGHAEASMNVSVDLIYEFMKQGDERLLEEAERLLEEVEE (SEQ ID NO: 27401)

LCB1_v1.1_Cys

DKENILQKIYEIMKTLDQLGHAEASMQVSDLIYEFMKQGDERLLEEAERLLEEVEERC (SEQ ID NO: 27402)

>LCB1_v1.2

DKENILQKIYEIMKTLDQLGHAEASMYVSDLIYEFMKQGDERLLEEAERLLEEVER (SEQ ID NO: 27403)

>LCB1_v1.3

DKENILQKIYEIMKTLEQLGHAEASMQVSDLIYEFMKQGDERLLEEAERLLEEVER (SEQ ID NO: 27404)

>LCB1_v1.4

DKENILQKIYEIMKTLEQLGHAEASMQVSDLIYEFMKQGDENLLEEAQQLQEVEER (SEQ ID NO: 27405)

>LCB1_v1.5 (LCB1_v1.3 with N-link Glycosylation)

DKENILQKIYEIMKTLEQLGHAEASMNVSDLIYEFMKQGDERLLEEAERLLEEVER (SEQ ID NO: 27406)

>LCB2-1

SDDEDSVRYLLYMAELRYEQGNPEKAKKILEMAEFIKRNNNEELERLVREVKKRL (SEQ ID NO: 27407)

>LCB2-2

SDDEDAVRYLLYMAELLYKQGNPEEAKKLELAEFIKRNNNEELERLVREVKKRL (SEQ ID NO: 27408)

>LCB3-1

NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27409)

>LCB3-2

NDDELLMLVTDLVAEALLFAKDEEIKKRVFTLFELADKAYKNNDRDTLSKVVSSELKELLE RLQ (SEQ ID NO: 27410)

>LCB3_v1.2

NDDELHMQMTDLVYEALHFAKDEEIQKHVFQLFEKATKAYKNKDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27411)

>LCB3-4

NDDELHMQMTDLVYEALHFAKDEEIQKHVFQLFENATKAYKNKDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27412)

>LCB3_v1.1

NDDELHMQMTDLVYEALHFAKDEEFQKHVFQLFEKATKAYKNNDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27413)

>LCB3_v1.3

NDDELHMQMTDLVYEALHFAKDEEFQKHVFQLFEKATKAYKNKDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27414)

>LCB3_v1.4

NDDELHMQMTDLVYEALHKAKDEEFQKHVFQLFEKATKARKNKDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27415)

>LCB3_v1.5

NDELHMQMTDLVYEALHKAKDEEMQKRVFQLFEQADKAYKTKDRQKLEKVVEELKE
LLERLLS (SEQ ID NO: 27416)

>LCB4-1

QREKRLKQLEMLLEYAIERNDPYLMFDVAVEMLRRLAEENNDERIIERAKRILEEYE (SEQ
ID NO: 27417)

>LCB4-2

DREERLKYLEMILLELAVERNDPYLIFDVAIELLRRLAEENNDERIYERAKRILEEVE (SEQ ID
NO: 27418)

>LCB5-1

SLEELKEQVKELKKELSPERMRLIEEALRFLEEGNPAMAMMVLSDLVYQLGDPRVIDLY
MLVTKT (SEQ ID NO: 27419)

>LCB5-2

SLEEVKEILKELKKELSPEDRRLIEEALRLLEEGNPAMASMVLSDLVFLGDPRVIELMLV
TKT (SEQ ID NO: 27420)

>LCB6-1

DREQRLVRFLVRLASKFNLSPEQILQLFEVLEELLERGVSEEEIRKQLEEVAKELG (SEQ
ID NO: 27421)

>LCB6-2

DREQRLVRFLVRLASKFNLSMEQILILFDVLEELLERGVSEEEIRKILEEVAKEL (SEQ ID
NO: 27422)

>LCB7-1

DDDIRYLIYMAKLRLEQGNPEEAKEKVLEMARFLAERLGMEELLKEVRELLRKIEELR
(SEQ ID NO: 27423)

>LCB7-2

DDDVRYLIYMAKLLLEQGNPEEAKEKVLESARFAAELLGNEELLKEVRELLRKIEELR (SEQ
ID NO: 27424)

>LCB8-1

PIIELLREAKEKNDEFAISDALYLVNELLQRTGDPRLEEVLFLIWRALKEKDPRLLDRAIE
LFE (SEQ ID NO: 27425)

>LCB8-2

PVTELLREAKEKNDPMAISDALFLVFELAQRTGDPRLEEVLFLIWRALKEKDPRLLDRAIE
LFE (SEQ ID NO: 27426)

>AHB1-1

DEDLEELERLYRKAEEVAKAKEKASRRGDDERAKEQMERAMRLFDQVFELAQELQEKQ
TDGNRQKATHLDKAVKEAADELYQRVR

(SEQ ID NO: 27427)

>AHB1-2

DEDLEELERLYRKAEEVAKAEAEASRRGDKERAKELLERLALHLFDQVFELAQELQEKLTD
EKRQKATHLTKAVHEAADELYQRVR

(SEQ ID NO: 27428)

>AHB2-1

ELEERVMLLDQVSELAHELLHKLTGEELQRATHFDKWANEAILIKSDDEREIREIEE
EARRILEHLEELARK (SEQ ID NO: 27429)

>AHB2-2_

ELEEQVMHVLDDQVSELAHELLHKLTGEELERAAYFNWWATEMMLLEIKSDDEREIREIEE
EARRILEHLEELARK (SEQ ID NO: 27430)

[0090] SEQ ID NO:27397-27430的多肽以高亲和力结合SARS-CoV-2刺突糖蛋白受体结合结构域(RBD)。已对SEQ ID NO:27397-27430的多肽进行了广泛的突变分析,允许确定多肽内每个残基的允许的取代。允许的取代如表3所示(数字表示残基编号,字母表示该残基上可能存在的单字母氨基酸)。

[0091] 因此,在一个实施方案中,所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27397-27430或选自SEQ ID NO:2797-27406、27409-27416、27427-27430的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,相对于参考靶标结合多肽氨基酸序列(即SEQ ID NO:27397-27430之一)的氨基酸取代选自表1中提供的允许的氨基酸取代。

[0092] 与RBD靶标在8A范围内的界面残基的残基数列于下表2中。

表2

'LCB1': [3,6,7,10,13,17,20,22,23,25,26,29,32,33,36],

'LCB2': [1,2,5,6,9,12,13,16,20,32,35,39],

'LCB3': [1,3,4,6,7,10,11,13,14,15,18,27,30,33,34,37],

'LCB4': [8,11,12,15,23,24,26,27,28,30,31,34,56],

'LCB5': [35,37,38,40,41,44,47,48,53,56,60,63],

'LCB6': [3,4,7,8,11,12,14,15,21,24,25,28,31,32,35],

'LCB7': [2,3,6,7,9,10,13,17,29,32,33,36],

'LCB8': [14,15,16,19,22,23,26,29,30,38,41,42,45],

'AHB1', [34,38,41,45,48,49,52,63,64,67,68,70,71,74,78,81,82,85],

'AHB2', [4,7,11,14,15,18,21,26,29,30,33,34,36,37,40,43,44,47,48].

[0093] 在另一个实施方案中,界面残基与参考靶标结合多肽中的那些残基相同(即:SEQ ID NO:27397-27430中的一个或相对于参考靶标结合肽中的界面残基保守地被取代,如表2所示)。

表3

LCB1 (SEQ ID NO:27397-27406)

1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

2--A,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

3--A,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y

4--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

5--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
6--A,C,I,L,M,Q,T,V
7--A,C,D,E,F,G,H,M,N,P,Q,R,S,V,W,Y
8--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
9--C,I,L,M,N,Q,T,V
10--C,F,V,W,Y
11--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
12--A,C,D,H,I,L,M,N,S,T,V,Y
13--C,I,M,Q
14--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
15--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
16--C,F,I,L,M,T,V
17--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
18--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
20--A,C,D,E,F,G,H,K,L,M,N,Q,R,S,T,W
21--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
22--A,C,D,F,G,H,I,L,M,N,P,Q,S,T,V,W,Y
23--C,E,M,N,P,Q,S,T,V
24--A,C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y
25--A,C,G,M,N,Q,S,T,V
26--M,N,V
27--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
28--A,C,G,I,L,S,T,V
29--A,C,S,V,W
30--D
31--A,C,D,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
32--C,F,H,I,L,M,N,P,T,V
33--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
34--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
35--A,C,D,F,H,M,Q,V,W,Y
36--A,C,D,E,G,H,I,L,M,N,Q,R,S,T,V,W,Y
37--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
39--A,C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y
40--D,E,G,H,N,P,Q
41--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

44--A,C,D,E,F,G,H,I,K,L,M,Q,R,S,V,W,Y
45--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
46--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
47--A,C,G,P,S,T,V
48--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
49--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
50--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
51--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
52--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB2 (SEQ ID NO: 27407-27408)
1--A,C,D,E,G,N,P,S,T
2--D,M,P,Q,Y
3--A,D,E,N,Q
4--C,D,E,V
5--D
6--A,C,D,E,G,N,Q,S,T,V
7--A,C,G,I,L,M,P,S,T,V
8--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
9--D,N,Y
10--I,L,T
11--C,E,G,I,L,M,W
12--F,H,Y
13--E,M,Q,R,V
14--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
15--A,C,D,E,G,H,I,K,L,M,N,Q,R,S,T,V
16--C,H,L,T
17--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
18--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
20--A,C,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,Y
21--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
22--A,C,D,E,G,I,K,L,N,P,Q,R,S,T,V
23--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
24--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
25--A,C,E,G,H,I,K,N,P,Q,R,S,T,Y

26--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
27--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
28--H,K,R,T,Y
29--C,D,E,H,I,K,L,M,N,P,Q,R,S,T,V,Y
30--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,S,T,V,W,Y
31--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,Y
32--F,H,I,K,L,M,P,Q,R,Y
33--A,C,G,P,S,T
34--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
35--F,H,Y
36--A,C,E,H,I,L,M,S,V
37--A,C,E,G,H,L,M,Q,R,S,T,V,W
38--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
39--A,C,D,E,G,H,I,K,L,M,N,P,Q,R,S,T,V
40--A,C,D,E,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
41--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
44--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
45--A,C,E,F,I,L,M,P,S,T,V,W,Y
46--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
47--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
48--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
49--A,C,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
50--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
51--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
52--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB3 (SEQ ID NO: 27409-27416)
1--C,E,F,I,M,N,T,W
2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
3--D,G,L,M,N,S,Y
4--A,C,E,F,H,K,Q,T
5--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
6--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
7--A,C,D,F,I,L,M,P,R,S,V,W

8--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
9--A,C,E,F,G,H,I,L,M,N,Q,R,S,T,V,Y
10--A,C,F,G,H,K,M,N,Q,R,S,T,Y
11--D,F,H,L,M,N,Q
12--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
13--A,F,I,L,M,N,Q,S,T,V
14--A,C,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
15--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
16--A,C,D,E,F,G,H,I,L,M,N,P,Q,R,S,T,V,W,Y
17--A,C,D,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W
18--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
21--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
24--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
25--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
26--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
27--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
28--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
29--A,C,D,E,F,G,I,L,M,N,P,S,T,V,W,Y
30--C,E,F,H,L,N,S,W,Y
31--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
32--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
33--A,C,E,F,I,K,P,Q,S,V,W,Y
34--A,D,E,F,G,H,M,N,P,Q,R,S,V,W,Y
35--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
36--A,C,E,G,H,I,M,N,Q,S,T,V
37--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
39--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
40--A,C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y
41--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
44--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
45--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
46--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

- 47--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 48--A,C,E,F,G,I,K,L,M,N,P,Q,S,T,V,W
- 49--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 50--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 51--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 52--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 55--A,C,E,F,G,H,I,K,L,M,N,Q,S,T,V,W,Y
- 56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 57--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 58--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 59--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 60--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 61--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 62--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 63--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 64--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- LCB4(SEQ ID NO:27417-27418)
- 1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 3--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 4--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 5--C,D,H,K,N,Q,R,Y
- 6--A,C,F,G,I,K,L,M,P,Q,R,S,T,V,Y
- 7--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 8--A,C,H,I,M,N,Q,R,S,T,V,Y
- 9--A,C,D,G,H,I,K,L,M,N,Q,R,S,T,V,Y
- 10--A,C,D,E,M,N,P,Q,S,T,V
- 11--C,D,G,H,I,K,L,M,N,P,R,S,T,V
- 12--F,G,I,L
- 13--F,I,L,M,S,V,Y
- 14--A,C,D,E,G,L,M,N,Q,R,S,T,V
- 15--C,E,F,G,H,I,L,M,S,V,W,Y
- 16--A,G,T,Y
- 17--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
- 18--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
- 19--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
- 20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

21--C,D,Q,Y
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--E,F,H,Y
24--A,F,G,I,L,M,W
25--A,C,E,G,H,I,K,L,M,N,Q,R,S,T,V,Y
26--C,F,H,I,L,N,S,T,V,W
27--D,Q,W,Y
28--A,C,D,I,L,V,Y
29--A,C,E,G,K,L,N,Q,R,S,T
30--C,I,L,M,P,T,V
31--C,D,E
32--A,C,E,I,L,M,Q,S,T,V,Y
33--A,C,E,F,G,H,I,K,L,M,Q,R,S,T,V,Y
34--C,D,F,G,H,L,M,N,P,R,S,T,W,Y
35--A,C,E,F,G,H,I,K,L,N,P,R,T,V,W
36--A,C,G,S,T,V
37--A,C,D,E,G,H,I,K,L,M,N,P,Q,R,S,T,V,Y
38--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
39--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
40--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,Y
41--A,C,D,E,G,H,K,N,Q,S,W
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,Y
44--A,E,F,G,H,I,K,L,M,N,Q,R,S,T,V
45--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
46--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
47--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
48--A,C,M,S,T,V
49--A,H,I,K,L,M,N,Q,R,S,T,V,W,Y
50--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
51--A,F,I,K,L,M,R,T,V,W,Y
52--F,I,K,L,M,V
53--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB5(SEQ ID NO:27419-27420)
1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

3--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
4--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
5--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
6--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
7--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
8--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
9--A,C,E,F,G,H,I,L,M,N,Q,S,T,V,W,Y
10--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
11--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
12--A,C,D,E,F,G,H,I,L,M,N,P,Q,R,S,T,V,W,Y
13--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
14--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
15--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
16--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
17--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
18--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
21--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--A,C,D,E,F,G,H,I,L,M,N,P,Q,R,S,T,W,Y
24--A,C,D,E,F,G,H,I,L,M,N,P,Q,S,T,V,W,Y
25--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
26--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
27--A,C,G,H,I,S,T,V
28--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
29--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
30--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
31--A,C,E,F,H,I,K,L,M,N,Q,S,T,V,W,Y
32--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
33--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
34--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
35--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
36--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
37--A,C,D,E,F,G,H,I,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,D,E,G,I,L,M,N,P,Q,S,T,V,W
39--A,C,F,G,L,M,N,S,T,V,W
40--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,Y
41--C,H,I,L,M,P,R

42--A,C,E,G,H,I,L,M,P,T,V,Y
43--C,I,L,M,Q,T,V
44--A,C,D,F,G,H,I,M,S,T
45--D,Y
46--A,C,D,F,I,L,R,V
47--C,E,G,I,V
48--F,I,V,W,Y
49--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
50--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
51--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
52--C,D,E,H,I,K,N,P,Q,R,S,T,Y
53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
56--F,I,L,M,T,V,W
57--A,C,D,E,F,G,H,N,P,Q,R,S,T,W,Y
58--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
59--A,C,F,I,L,M,T,V,Y
60--C,F,M,N,V,Y
61--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
62--A,C,F,G,I,L,M,S,T,V,W
63--A,C,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
64--A,C,E,F,G,H,K,L,N,P,R,S,T,W,Y
65--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB6 (SEQ ID NO:27421-27422)
1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
3--E,W
4--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
5--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
6--F,L,M,R,S
7--H,T,V
8--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
9--F,M
10--A,K,L,W
11--D,E,G,V,Y
12--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
13--E,L
14--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

15--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
16--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
17--F,N,P,S
18--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
19--L,N,Q,V
20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
21--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--C,D,P,Q,R,W
24--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
25--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
26--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
27--D,H,L,S,W
28--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
29--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
30--L,Q,V,W
31--I,K,L,S
32--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
33--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
34--A,F,L,T,V
35--C,D,G,H,K,L,N,T
36--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
37--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
39--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
40--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
41--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
44--F,I
45--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
46--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
47--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
48--L,Q,R,T
49--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
50--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
51--C,V,Y
52--A,E,H,K
53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--C,F,H,L,P,W,Y
56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB7 (SEQ ID NO:27423-27424)
1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
3--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
4--I,T,V
5--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
6--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
7--L,P,Y
8--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
9--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
10--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
11--A
12--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
13--A,L,P
14--H,L,R,T,Y
15--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
16--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
17--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
18--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
21--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--A,S
24--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
25--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
26--C,G,S,V,Y
27--K,L,M,W
28--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
29--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
30--A,Y
31--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
32--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
33--A,C,F,I,K,L,V,W
34--A,H,L
35--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

36--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
37--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
39--A,C,K,L,M,N
40--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
41--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
42--A,C,D,L,V
43--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
44--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
45--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
46--Q,S,V
47--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
48--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
49--E,L
50--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
51--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
52--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
53--I
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
56--L,M,N,R
57--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
LCB8 (SEQ ID NO:27425-27426)
1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
2--C,F,I,L,M,S,V,W,Y
3--A,C,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
4--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
5--A,C,F,G,I,K,L,M,Q,S,T,V,W,Y
6--H,I,K,L,M
7--A,H,I,K,L,M,N,P,Q,R,W,Y
8--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
9--A,C,F,G,I,L,M,S,Y
10--A,F,H,K,L,M,Q,R,S
11--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
12--A,C,D,E,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
13--A,C,D,E,F,G,H,M,N,Q,S,W,Y
14--C,D,E,H,N,Q,S
15--A,D,E,F,H,I,L,M,N,P,Q,S,T,V,W,Y
16--C,F,M,N,R,Y

17--A,C,I,L,M,Q,R,V
18--A,C,F,H,I,L,M,T,V,Y
19--I,Q,S
20--D,N
21--A,C,G,S,V
22--A,C,I,L,M,V
23--C,F,R,T,W,Y
24--A,C,D,E,F,G,H,I,L,M,N,Q,R,S,T,V,W,Y
25--C,E,S,T,V,Y
26--A,C,D,E,F,G,H,N,Q,S,T
27--A,C,D,E,G,H,I,K,L,M,N,Q,R,S,T,V
28--C,E,F,G,H,I,K,L,M,Q,R,W,Y
29--A,C,F,G,H,I,K,L,M,N,Q,R,S,T,V,Y
30--A,C,E,G,H,K,M,N,P,Q,R,T
31--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,Y
32--A,C,D,E,G,H,I,K,N,Q,R,S,T,W
33--A,C,E,G,H,K,M,N,P,Q,R,S,W,Y
34--C,D,E,F,H,M,N,W,Y
35--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,Y
36--A,C,D,E,F,G,H,K,L,M,N,Q,R,S,T,V,W,Y
37--F,G,H,I,L,M,S,T,Y
38--D,E,H,Q,W,Y
39--C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y
40--A,C,E,G,H,I,K,M,P,V,Y
41--C,F,H,I,K,L,M,R,S,T,V
42--E,F,I,T,W,Y
43--A,C,D,E,F,H,I,L,M,N,Q,R,S,T,V,W,Y
44--C,G,I,K,L,M,T,V,Y
45--G,S,W,Y
46--C,I,K,L,M,N,Q,R,S,T
47--A,C,E,N,Q,S,T,V
48--C,D,E,F,H,I,L,M,W
49--C,D,F,H,K,L,M,N,Q,R,T
50--A,C,D,E,N,Y
51--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
52--A,C,D,E,G,H,K,L,M,N,Q,R,S,T
53--A,C,D,E,F,G,H,I,L,M,N,P,Q,S,T,V,W,Y
54--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
55--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,S,T,V,W,Y

56--C,I,L,M
57--A,C,D,E,G,I,N,Q,S,T
58--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
59--A,C,G,P,S
60--A,C,E,F,G,I,L,M,N,Q,S,T,V
61--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
62--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
63--A,C,E,F,G,H,I,L,M,N,Q,S,T,V,W,Y
64--A,C,D,E,G,H,I,K,L,M,N,P,Q,S,T,V
65--A,C,D,E,G,H,I,K,L,M,N,P,Q,R,S,T,W,Y
AHB1 (SEQ ID NO:27427-27428)
1--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
2--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
3--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
4--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
5--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
6--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
7--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
8--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
9--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
10--A,C,F,H,I,K,L,M,N,Q,R,S,T,V,W,Y
11--F,N,Y
12--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
13--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
14--A,D,G
15--A,C,D,E,G,H,I,K,L,M,N,Q,R,S,T,V
16--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
17--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
18--A,C,D,E,F,G,H,I,L,M,N,Q,S,T,V,W,Y
19--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
20--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
21--A,C,E,G,S,V
22--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
23--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
24--A,C,D,E,F,H,K,L,M,N,Q,R,S,T,V,Y
25--A,C,D,F,G,H,L,M,N,Q,R,S,T,V,W,Y
26--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
27--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
28--A,C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,Y

29--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
30--A,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
31--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
32--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
33--A,G,S
34--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
35--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
36--A,C,D,E,F,G,H,K,L,M,N,P,Q,R,S,T,V,W,Y
37--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
38--A,C,E,G,H,M,P,Q
39--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
40--A,C,D,E,G,K,N,Q,R,S,T
41--A,C,D,E,F,G,H,I,L,M,N,P,Q,S,T,V,W,Y
42--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
43--A,C,D,E,F,G,H,I,K,L,M,N,Q,S,T,V,W,Y
44--E,F,H,Q,S,W,Y
45--D,N
46--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
47--C,T,V
48--F,S,W,Y
49--A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y
50--A,C,F,H,I,K,L,M,N,Q,R,S,T,V,W,Y
51--A,D,G,H,N,S
52--H,K,Q,R
53--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
54--A,C,H,I,K,L,M,N,P,Q,R,S,T,V
55--A,C,E,G,H,K,N,Q,R,S,T
56--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
57--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
58--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
59--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
60--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
61--A,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
62--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
63--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
64--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
65--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
66--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
67--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y

- 68--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 69--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 70--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 71--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 72--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 73--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 74--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 75--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 76--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 77--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 78--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 79--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 80--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 81--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 82--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 83--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 84--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- 85--A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
- AHB2 (SEQ ID NO:27429-27430)
- 1--C,G,A,V,F,Y,W,S,Q,D,E,R,K
- 2--C,P,G,V,I,M,L,F,Y,W,S,N,Q,D,E,R,H
- 3--C,G,A,V,I,F,S,T,D,E,K
- 4--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
- 5--C,P,G,A,V,M,L,Y,W,S,N,Q,D,E,R,K,H
- 6--G,A,V,I,F,S,T,D,H
- 7--C,P,G,V,I,M,L,F,W,S,T,N,Q,E,R,K,H
- 8--C,P,G,A,V,M,L,Y,W,S,T,N,Q,D,E,R,K,H
- 9--C,P,G,A,V,I,M,L,F,W,S,T,N,Q,D,E,R,K,H
- 10--C,P,G,A,V,I,L,Y,W,S,T,N,E,R,K
- 11--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,H
- 12--C,P,G,A,V,I,L,F,Y,W,S,T,N,Q,D,E,R,K,H
- 13--C,G,A,V,M,L,F,W,S,T,N,E,H
- 14--C,P,G,A,V,I,Y,S,T,N,D,E,R,H
- 15--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K
- 16--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
- 17--C,P,G,A,V,L,Y,W,S,T,Q,D,E,R
- 18--C,P,A,V,I,M,F,Y,N,Q,R,K,H
- 19--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
- 20--C,P,G,A,V,M,L,Y,W,N,Q,E,R,K,H

21--C,P,G,A,V,I,M,L,F,Y,W,S,N,Q,E,R,K,H
22--C,P,G,A,V,M,L,F,Y,S,T,N,Q,D,E,R,K,H
23--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,E,R,K
24--C,P,G,A,V,I,M,L,F,Y,W,S,Q,E,R,H
25--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,R,H
26--C,G,A,V,L,Y,S,N,D,R,K,H
27--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
28--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
29--C,P,G,V,I,M,L,F,Y,W,S,T,N,Q,D,R,K,H
30--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
31--C,G,A,V,I,M,L,F,Y,W,S,T,Q,D,E,R,K,H
32--P,G,A,V,I,L,W,S,T,D,R,H
33--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,E,R,K,H
34--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
35--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
36--C,P,G,A,V,I,L,F,Y,S,T,N,Q,D,E,R,H
37--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
38--C,P,G,A,V,I,M,L,F,Y,W,S,T,Q,E,R,K
39--C,P,G,A,V,I,W,S,Q,E,R,H
40--C,P,G,A,V,I,L,Y,W,S,T,N,D,E,R,K,H
41--C,P,G,A,V,I,M,L,Y,W,S,T,N,Q,D,E,R,K,H
42--C,P,G,A,V,M,L,Y,W,S,T,N,Q,D,E,R,K,H
43--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
44--C,P,G,A,V,I,M,L,F,W,S,T,Q,D,E,R,H
45--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
46--C,P,G,A,V,I,M,L,F,S,T,Q,E,R,K
47--C,G,A,V,I,M,L,F,W,S,T,N,Q,D,E,R,H
48--C,P,G,A,V,I,M,L,F,Y,W,S,N,Q,E,R,K
49--C,P,G,A,V,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
50--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
51--C,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
52--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
53--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,D,E,R,K,H
54--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
55--C,P,G,A,V,I,M,L,F,Y,S,T,N,Q,D,E,R,K,H
56--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
57--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
58--C,G,A,V,I,M,L,F,Y,W,S,T,N,E,R,K,H
59--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H

60--C,G,A,V,I,M,L,F,Y,W,S,T,Q,D,E,R,K
61--C,P,G,A,V,I,M,L,F,Y,W,S,N,Q,D,E,R,K,H
62--C,G,A,V,L,S,T,N,D,E,K,H
63--C,P,G,A,V,I,L,F,Y,W,S,T,N,Q,D,E,R,K,H
64--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,H
65--C,G,A,V,I,M,L,F,Y,S,T,N,R,K,H
66--C,P,G,A,V,I,M,L,W,T,Q,E,R
67--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
68--C,P,G,A,V,I,L,F,Y,W,S,T,N,Q,D,E,R,H
69--P,G,V,I,M,L,Y,W,S,T,Q,R,K
70--C,P,G,A,V,I,M,L,F,Y,W,S,T,N,Q,D,E,R,K,H
71--C,G,A,V,L,F,W,S,Q,D,E,R,K
72--C,V,I,L,S
73--P,G,A,V,S,T,E
74--C,A,L,F,Y,S,T,R,H
75--C,P,G,V,I,L,F,W,S,N,D,E,R,K

[0094] 在一个实施方案中,所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27397-27406和27431-27466的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

表4:示例性LCB1变体

名称	结合域蛋白
LCB1_4N	DKENILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO:27431)
LCB1_4K	DKEKILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO:27432)
LCB1_14K	DKEWILQKIYEIMKLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO:27433)
LCB1_15T	DKEWILQKIYEIMRTLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO:27435)
LCB1_18Q	DKEWILQKIYEIMRLLDQLGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO:27436)
LCB1_18K	DKEWILQKIYEIMRLLDKLGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO: 27437)
LCB1_27Q	DKEWILQKIYEIMRLLDELGHAEASMQVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO: 27438)
LCB1_27Y	DKEWILQKIYEIMRLLDELGHAEASMYVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO: 27439)
LCB1_17E	DKEWILQKIYEIMRLLLEELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO: 27440)
LCB1_17R	DKEWILQKIYEIMRLLRELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEER (SEQ ID NO: 27441)
LCB1_42N	DKEWILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDENLLEEAEERLLEEVEER (SEQ ID NO: 27442)
LCB1_49Q	DKEWILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEQLLEEVEER (SEQ ID NO: 27443)
LCB1_52Q	DKEWILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLQEVEER (SEQ ID NO: 27444)

LCB1_32L	DKEWILQKIYEIMRLLDELGHAEASMRVSDLLYEFMKKGDERLLEEAERLLEEVEER (SEQ ID NO: 27445)
LCB1_28A	DKEWILQKIYEIMRLLDELGHAEASMRASDLIYEFMKKGDERLLEEAERLLEEVEER (SEQ ID NO: 27446)
LCB1_v1.3_ACH1	DKENILQKIYEIMKTLEQLGHAEASMYVSDLIYEFMKQGDERLLEEAERLLEEVEER (SEQ ID NO: 27447)
LCB1_v1.3_ACH2	DKENILQKIYEIMKTLEQLGHAEASMQVSDLIYEFMKQGDENLLEEAERLLEEVEER (SEQ ID NO: 27448)
LCB1_v1.3_ACH3	DKENILQKIYEIMKTLEQLGHAEASMQVSDLIYEFMKQGDERLLEEAEQLLLEEVEER (SEQ ID NO: 27449)
LCB1_v1.3_ACH4	DKENILQKIYEIMKTLEQLGHAEASMYVSDLIYEFMKQGDENLLEEAEQLLLEEVEER (SEQ ID NO: 27450)
LCB1_v1.3_ACH5	DKENILQKIYEIMKTLEQLGHAEASMQVSDLIYEFMKQGDENLLEEAEQLLLEEVEER (SEQ ID NO: 27451)
LCB1_v1.3_1	DRENILQKIYEIMKELEKLGHAEASMQVSDLIYEFMQDKDERLLEEAERLLEEVEKVR (SEQ ID NO: 27452)
LCB1_v1.3_2	DRENILQKIYEIMKELRQLGHAEASMQVSDLIYEFMKTDKDKRLLEEAERLLEEVEKVR (SEQ ID NO: 27453)
LCB1_v1.3_3	DRENILQKIYEIMKTLRRLGHAEASMQVSDLIYEFMQDKDKRLLEEAERLLEEVEQR (SEQ ID NO: 27454)
LCB1_v1.3_4	DKENVLQKIYEIMKELERLGHAEASMQVSDLIYEFMKTDKDERLLEEAERLLEEVEKVR (SEQ ID NO: 27455)
LCB1_v1.3_5	DRENILQKIYEIMKTLEKLGHAEASMQASDLIYEFMKTDKDERLLEEAERLLEEVEQR (SEQ ID NO: 27456)
LCB1_v1.3_6	DKENILQKIYEIMKTLRALGHAEASMQVSDLIYEFMQTKDERLLEEAERLLEEVEKVR (SEQ ID NO: 27457)
LCB1_v1.3_7	DKENVLQKIYEIMKTLEKLGHAEASMQVSDLIYEFMQTKDKRLLEEAERLLEEVEQR (SEQ ID NO: 27458)
LCB1_v1.3_15	DRENILQKIYEIMKELEKLGHAEASMQVSDLIYEFMQDKDENLLEEAERLLEEVEKVR (SEQ ID NO: 27459)
LCB1_v1.3_16	DRENILQKIYEIMKELRQLGHAEASMQVSDLIYEFMKTDKDKNLEEAERLLEEVEKVR (SEQ ID NO: 27460)
LCB1_v1.3_17	DRENILQKIYEIMKTLRRLGHAEASMQVSDLIYEFMQDKDKNLEEAERLLEEVEQR (SEQ ID NO: 27461)

LCB1_v1.3_19	DRENILQKIYEIMKTLEKLGHAEASMQASDLIYEFMKTKDENLLEEAEERLLEEVQR (SEQ ID NO: 27462)
LCB1_v1.3_20	DKENILQKIYEIMKTLRALGHAEASMQVSDLIYEFMQTKDENLLEEAEERLLEEVKR (SEQ ID NO: 27463)
LCB1_v1.3_21	DKENVLQKIYEIMKTLEKLGHAEASMQVSDLIYEFMQTKDKNLLEEAEERLLEEVQR (SEQ ID NO: 27464)
LCB1_v2.2	DKENVLQKIYEIMKELERLGHAEASMQVSDLIYEFMKTKDENLLEEAEERLLEEVKR (SEQ ID NO: 27465)
LCB1_v2.2_ompT	DKENVLQKIYEIMKELERLGHAEASMQVSDLIYEFMKTKDENLLEEAEERLLEEVTR (SEQ ID NO: 27466)

[0095] 在另一个实施方案中,所述一种或多种靶标结合多肽包含在相对于SEQ ID NO: 27397的氨基酸序列的、选自2、4、15、17、18、27、28、32、37、38、39、41、42、49、52和55处的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17或全部18个残基的氨基酸取代。在另一个实施方案中,一种或多种靶标结合多肽中的取代选自表5中所列的取代,在给定行中单独或组合。

表5. 示例性LCB1突变

名称	亲本	来自野生型的突变												
LCB1_1	LCB1	W4N	R14K	L15T	E18Q	R27Q	K38Q							
LCB1_2	LCB1	W4N	R14K	L15T	E18Q	R27Y	K38Q							
LCB1_3	LCB1	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q						
LCB1_4	LCB1	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q	R42N	R49Q	E52Q			
LCB1_4N	LCB1	W4N												
LCB1_4K	LCB1	W4K												
LCB1_14K	LCB1	R14K												
LCB1_15T	LCB1	L15T												
LCB1_18Q	LCB1	E18Q												
LCB1_18K	LCB1	E18K												
LCB1_27Q	LCB1	R27Q												
LCB1_27Y	LCB1	R27Y												
LCB1_38Q	LCB1	K38Q												
LCB1_17E	LCB1	D17E												

LCB1_17R	LCB1	D17R											
LCB1_42N	LCB1	R42N											
LCB1_49Q	LCB1	R49Q											
LCB1_52Q	LCB1	E52Q											
LCB1_32L	LCB1	I32L											
LCB1_28A	LCB1	V28A											
LCB1_v1.3	LCB1	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q					
LCB1_v1.3_ACH1	LCB1_v1.3	W4N	R14K	L15T	D17E	E18Q	R27Y	K38Q					
LCB1_v1.3_ACH2	LCB1_v1.3	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q	R42N				
LCB1_v1.3_ACH3	LCB1_v1.3	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q	R49Q				
LCB1_v1.3_ACH4	LCB1_v1.3	W4N	R14K	L15T	D17E	E18Q	R27Y	K38Q	R42N	R49Q			
LCB1_v1.3_ACH5	LCB1_v1.3	W4N	R14K	L15T	D17E	E18Q	R27Q	K38Q	R42N	R49Q			
LCB1_v1.3_1	LCB1_v1.3	K2R	W4N	R14K	L15E	D17E	E18K	R27Q	K37Q	K38D	G39K	E55K	
LCB1_v1.3_2	LCB1_v1.3	K2R	W4N	R14K	L15E	D17R	E18Q	R27Q	K38T	G39K	E41K	E55K	
LCB1_v1.3_3	LCB1_v1.3	K2R	W4N	R14K	L15T	D17R	E18R	R27Q	K37Q	K38D	G39K	E41K	
LCB1_v1.3_4	LCB1_v1.3	W4N	I5V	R14K	L15E	D17E	E18R	R27Q	K38T	G39K	E55K		
LCB1_v1.3_5	LCB1_v1.3	K2R	W4N	R14K	L15T	D17E	E18K	R27Q	V28A	K38T	G39K	E55Q	
LCB1_v1.3_6	LCB1_v1.3	W4N	R14K	L15T	D17R	E18A	R27Q	K37Q	K38T	G39K	E55K		
LCB1_v1.3_7	LCB1_v1.3	W4N	I5V	R14K	L15T	D17E	E18K	R27Q	K37Q	K38T	G39K	E41K	
LCB1_v1.3_15	LCB1_v1.3	K2R	W4N	R14K	L15E	D17E	E18K	R27Q	K37Q	K38D	G39K	R42N	
LCB1_v1.3_16	LCB1_v1.3	K2R	W4N	R14K	L15E	D17R	E18Q	R27Q	K38T	G39K	E41K	R42N	
LCB1_v1.3_17	LCB1_v1.3	K2R	W4N	R14K	L15T	D17R	E18R	R27Q	K37Q	K38D	G39K	E41K	
LCB1_v1.3_19	LCB1_v1.3	K2R	W4N	R14K	L15T	D17E	E18K	R27Q	V28A	K38T	G39K	R42N	
LCB1_v1.3_20	LCB1_v1.3	W4N	R14K	L15T	D17R	E18A	R27Q	K37Q	K38T	G39K	R42N	E55K	
LCB1_v1.3_21	LCB1_v1.3	W4N	I5V	R14K	L15T	D17E	E18K	R27Q	K37Q	K38T	G39K	E41K	
LCB1_v2.2	LCB1_v1.3	W4N	I5V	R14K	L15E	D17E	E18R	R27Q	K38T	G39K	R42N	E55K	
LCB1_v2.2_ompT	LCB1_v1.3	W4N	I5V	R14K	L15E	D17E	E18R	R27Q	K38T	G39K	R42N	E55T	

[0096] 在另一个实施方案中,所述一种或多种靶标结合多肽包含与选自SEQ ID NO: 27409-27416和27467-27493的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

表6. 示例性LCB3变体

名称	结合域蛋白
LCB3_8Q	NDDELHMQMTDLVYEALHFAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27467)
LCB3_8T	NDDELHMTMTDLVYEALHFAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27468)
LCB3_19K	NDDELHMLMTDLVYEALHKAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27469)
LCB3_19I	NDDELHMLMTDLVYEALHIAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVEELKELLE RLLS (SEQ ID NO: 27470)
LCB3_25F	NDDELHMLMTDLVYEALHFAKDEEFKRVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27471)
LCB3_25M	NDDELHMLMTDLVYEALHFAKDEEMKRVFQLFELADKAYKNNDRQKLEKVVEELKEL LERLLS (SEQ ID NO: 27472)
LCB3_26Q	NDDELHMLMTDLVYEALHFAKDEEIQKRVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27473)
LCB3_28H	NDDELHMLMTDLVYEALHFAKDEEIKKHVFQLFELADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27474)
LCB3_35K	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFEKADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27475)
LCB3_37T	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFELATKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27476)
LCB3_40R	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFELADKARKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27477)
LCB3_43K	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFELADKAYKNKDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27478)
LCB3_34G	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFGLADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27479)
LCB3_34Y	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFYLDKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27480)
LCB3_34T	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFTLADKAYKNNDRQKLEKVVEELKELL ERLLS (SEQ ID NO: 27481)
LCB3_49K	NDDELHMLMTDLVYEALHFAKDEEIKKRVFQLFELADKAYKNNDRQKLEKVVVEELKELL ERLLS (SEQ ID NO: 27482)
LCB3_v1.2_A	NDDELHMQMTDLVYEALHFAKDEEIQKHVFQLFGKATKAYKNKDRQKLEKVVEELKELL

CH1	ERLLS (SEQ ID NO: 27483)
LCB3_v1.2_A CH2	NDELHMQMTDLVYEALHFAKDEEIQKHVFQLFYKATKAYKNKDRQKLEKVV EELKELL ERLLS (SEQ ID NO: 27484)
LCB3_v2.2	NLDELHMQMTDLVYEALHFAKDEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKELL ERLLS (SEQ ID NO: 27485)
LCB3_v1.3_2	NDELHMQMTDLVYEALHFAKTEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKELL ERLLS (SEQ ID NO: 27486)
LCB3_v1.3_3	NDELHMQMTDLVYEALHFAKDEEFQKHVFQLFEKARKAYKNKDRQKLEKVV EELKEL LERLLS (SEQ ID NO: 27487)
LCB3_v1.3_4	NDELHMQMTDLVWEALHFAKDEEFQKHVFQLFEKARKAYKNKDRQKLEKVV EELKEL LERLLS (SEQ ID NO: 27488)
LCB3_v1.3_5	NDELHMQMTDLVWEALHFAKDEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKEL LERLLS (SEQ ID NO: 27489)
LCB3_v1.3_6	NEDELHMQMTDLVWEALHFAKDEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKEL LERLLS (SEQ ID NO: 27490)
LCB3_v1.3_7	NDELHMQMTDLVWEALHFAKTEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKEL LERLLS (SEQ ID NO: 27491)
LCB3_v1.3_15	NLDELHMQMTDLVYEALHFAKTEEFQKHVFQLFEKATKAYKNKDRQKLEKVV EELKELL ERLLS (SEQ ID NO: 27492)
LCB3_v2.3	NIDELLMQVTDLIYEALHFAKDEEFQKHAFQLFEKATKAYKNKDKQKLEKVV EELKELLE RILS (SEQ ID NO: 27493)

[0097] 在一个实施方案中,靶标结合包含在相对于SEQ ID NO:27409的氨基酸序列的、选自2、6、8、19、13、16、19、22、25、26、28、29、34、35、37、40、43、45、49和62处的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或全部20个氨基酸取代,所述取代选自表7中所列的取代,在给定行中单独或组合。

表7. 示例性LCB3突变

名称	亲本	来自野生型的突变								
		L8Q	I25F	K26Q						
LCB3_1	LCB3	L8Q	I25F	K26Q	R28H	L35K	D37T			
LCB3_2	LCB3	L8Q	K26Q	R28H	L35K	D37T	N43K			
LCB3_3	LCB3	L8Q	I25F	K26Q	R28H	L35K	D37T	N43K		

LCB3_4	LCB3	L8Q	F19K	I25F	K26Q	R28H	L35K	D37T	Y40R	N43K
LCB3_8Q	LCB3	L8Q								
LCB3_8T	LCB3	L8T								
LCB3_19K	LCB3	F19K								
LCB3_19I	LCB3	F19I								
LCB3_25F	LCB3	I25F								
LCB3_25M	LCB3	I25M								
LCB3_26Q	LCB3	K26Q								
LCB3_28H	LCB3	R28H								
LCB3_35K	LCB3	L35K								
LCB3_37T	LCB3	D37T								
LCB3_40R	LCB3	Y40R								
LCB3_43K	LCB3	N43K								
LCB3_34G	LCB3	E34G								
LCB3_34Y	LCB3	E34Y								
LCB3_34T	LCB3	E34T								
LCB3_49K	LCB3	E49K								
LCB3_v1.2	LCB3	L8Q	K26Q	R28H	L35K	D37T	N43K			
LCB3_v1.2_ACH1	LCB3_v1.2	L8Q	K26Q	R28H	E34G	L35K	D37T	N43K		
LCB3_v1.2_ACH2	LCB3_v1.2	L8Q	K26Q	R28H	E34Y	L35K	D37T	N43K		
LCB3_v2.2	LCB3_v1.3	D2L	L8Q	I25F	K26Q	R28H	L35K	D37T	N43K	
LCB3_v1.3_2	LCB3_v1.3	L8Q	D22T	I25F	K26Q	R28H	L35K	D37T	N43K	
LCB3_v1.3_3	LCB3_v1.3	L8Q	I25F	K26Q	R28H	L35K	D37R	N43K		
LCB3_v1.3_4	LCB3_v1.3	L8Q	Y14W	I25F	K26Q	R28H	L35K	D37R	N43K	
LCB3_v1.3_5	LCB3_v1.3	L8Q	Y14W	I25F	K26Q	R28H	L35K	D37T	N43K	
LCB3_v1.3_6	LCB3_v1.3	D2E	L8Q	Y14W	I25F	K26Q	R28H	L35K	D37T	N43K

LCB3_v1.3_7	LCB3_v1.3	L8Q	Y14W	D22T	I25F	K26Q	R28H	L35K	D37T	N43K
LCB3_v1.3	LCB3_v1.2	L8Q	I25F	K26Q	R28H	L35K	D37T	N43K		
LCB3_v1.3_15	LCB3_v1.3	D2L	L8Q	D22T	I25F	K26Q	R28H	L35K	D37T	N43K
										V29A, L35K, D37T, N43K, R45K, L62I
LCB3_v2.3	LCB1_v2.1	D2I	H6L	L8Q	M9V	V13I	I25F	K26Q	R28H	L62I

[0098] 在一个实施方案中,靶标结合包含与选自SEQ ID NO:27427-27430和27494的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列。

AHB2 ELEEQVMHVLDQVSELAHELLHKL TGEELERAAYFNWWATEMMLLELIKSSD
v2 EREIREIEEEAARILEHLEELART (SEQ ID NO:27494)

[0099] 在一个这样的实施方案中,所述一种或多种靶标结合多肽包含在相对于氨基酸序列SEQ ID NO:27430的选自63和75的残基处或在两个残基处的氨基酸取代。在另一个实施方案中所述取代包括R63A和/或K75T。

[0100] 在另一个实施方案中,笼蛋白包含与US20200239524(或W02020/018935)中公开的笼多肽的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,不包括任选的氨基酸残基并且不包括锁区中的氨基酸残基。这些笼蛋白氨基酸序列不包括一种或多种靶标结合多肽或第一报告蛋白结构域(或存在时的第二报告蛋白结构),因此可以将其添加到本实施方案的笼蛋白中。

[0101] 示例性的这样的实施方案是US20200239524中公开的笼多肽的SEQ ID NO:1-49、51-52、54-59、61、65、67-91、92-2033、2034-14317、27094-27117、27120-27125、27278至27321,以及SEQ ID NO在SEQ ID NO:27126和27276之间的偶数编号的笼多肽、表3(当前申请中的表8)和/或表4(当前申请的表9),并在本文和序列列表中复现。

[0102] 在每个实施方案中,每种笼蛋白的N端和/或C端60个氨基酸可以是任选的,因为末端60个氨基酸残基可以包括可以被修饰的锁区,例如通过用一种或多种靶标结合多肽和第一报告蛋白结构域替换锁的全部或部分。在一个实施方案中,N端60个氨基酸残基是任选的;在另一个实施方案中,C端60个氨基酸残基是任选的;在另一个实施方案中,N端60个氨基酸残基和C端60个残基中的每一个都是任选的。在一个实施方案中,在确定序列同一性百分比时,不包括这些任选的N端和/或C端60残基。在另一个实施方案中,在确定序列同一性百分比时可以包括任选残基。

表8

行数	笼 (第一列)	钥匙 (第二列)
1	LOCKR_extend18 (SEQ ID NO:6), BimLOCKR_extend18 (SEQ ID NO:22), , 1fix-long-Bim-t0 (SEQ ID NO: 54), 1fix-long-GFP-t0 (SEQ ID NO: 55), 1fix-short-BIM-t0 (SEQ ID NO: 56), 1fix-short-GFP-t0 (SEQ ID NO: 57),	p18_MBP (SEQ ID NO:27020), p76-long (SEQ ID NO:27027), p76-short (SEQ ID NO:27028),
2	LOCKRb (SEQ ID NO:7),	key_b (SEQ ID NO:27022)
3	LOCKRc (SEQ ID NO:8),	key_c (SEQ ID NO:27023)

表9

笼名称	笼序列	钥匙名称	钥匙序列
2加1_笼_C端_2406	SEVDEVVKEVEDLVRNEELVEE VVRVVEKVVTDDRRLVEEVVREI RKIVKDVEDLARKLDKEELKRVL DEMREIERLLEKLRRHSKKLDDE LKRLLLEELREHSRRVEKRLEDLLK ELRREGVDEKVLKLEKVIREVRE	2加1_钥匙_C端_2406	EKVLKLEKVIREVRERSTRALRKV EEVIRRVREESERALRDLERVVKEV EKRMREAAAR(SEQ ID NO:27127)

	RSTRALRKVEEVIRRVREESERAL RDLERVVKEVEKRMREAAR (SEQ ID NO:27126)		
2 加 1_ 笺_ C 端 _5398	SVEELLRKLEEVLRKIREENERSL KELRDRAREIVKRNRETNRELEEV IKELEKRLSGADKEKVEELVRRIR RIVERVVEEDRRTVEEIEKIAREV VKRDRDSADRVRTVEDVLRKAT GSEDIVRKIERIVETIEREVRESVK KVEEIARDIRRKVDESVKVNEKLL RDVDKKARDRKK(SEQ ID NO:27128)	2 加 1_ 钥匙 _C 端_5398	EDIVRKIERIVETIEREVRESVKKVEE IARDIRRKVDESVKVNEKLLRDVDK KARDRKK(SEQ ID NO:27129)
2 加 1_ 笺_ C 端 _5405	SESDDVIRKLELLEELRTHVEKSI RDLRKILEDSTRHAKRSIEELERLL EEVRKKPGDEEVRKTVEEISRRVA ENVKRLEDLYRRMEEVKKNLDR LRKRVEDIIREVEEARKKGVDEEK LKDLIRKLRDILRRAAEAHKKLID DARESLERAKREHEKLIDRLKKIL EELER(SEQ ID NO:27130)	2 加 1_ 钥匙 _C 端_5405	EEKLKDILRKLDRDILRRAAEAHKKLI DDARESLERAKREHEKLIDRLKKILE ELER(SEQ ID NO:27131)
2 加 1_ 笺_ C 端 _5406	DREREVKKRLDEVREIERLLRRV EEESRRVAEEIRRLIEVRRRNKK VTEEIRELLKGLKDKEVRRVLER LRKLNAESDELLERILERLRLRVE ATNRLVKAIEELRRLVEKIVREVP DSEELREELKKLERKIEKVAKEIH DHDKEVTERLEDLLRRITEHARKS DREIETAR(SEQ ID NO:27132)	2 加 1_ 钥匙 _C 端_5406	EELREELKKLERKIEKVAKEIHDHD KEVTERLEDLLRRITEHARKSDREIE ETAR(SEQ ID NO:27133)
2 加 1_ 笺_ C 端 _5409	SEAEELLKRLEDRAEEILRRLEEIL RTSRKLAEDVLRELEKLLRESERR IREVLEELRGIKDKKELEDVIREVE KELDESLSRSELLKDVLLKLLDD NLKESERLVEDIDRELAKILEDLK KAGVPKEVVDEIKRIVDEVRERLK	2 加 1_ 钥匙 _C 端_5409	KEVVDEIKRIVDEVRERLKRIVDEN AKIVEDARRALEKIVKENEELRRLK KELRELK(SEQ ID NO:27135)

	RIVDENAKIVEDARRALEKIVKEN EEILRRLKKELRELRK(SEQ ID NO:27134)		
2加1_笼 _5414_GFP11_ C端	SEIEKILKEIEDLARRDEEVSKKIV EDIRRLAKEVEDTSRDIRVKIEELA KRVLDRLRKDGSKEELEKEVREV VKTLEELVKDNHRLIRRAVEEMK RLVEENHRHSREVVKELEDLVRE LRKGSSEDSESRDHMVLHEYVNA AGITSEKSNEEVKRTVEELVRRME ESNDRVRDLVRRLVEELKRAVD(S EQ ID NO:27140)	2加1_钥匙 _C端_5414	EDSERLVREVEDLVRRLVRRSEKSN EEVKRTVEELVRRMEESNDRVRDL VRRLVEELKRAVD(SEQ ID NO:27141)
2加1_笼_C端 _5421	SVDEVLKEIEDALRRLKEEVERVL KENEDELRRLEEEVRRVLKEDEEL LESLKRGVGSDEVDRVVDEIAK LSAEILEKVKKVVKEIRDSLETVK RRVDDVRRRLKELLDEIKRGSDE KAIRDVAKEIRDRLKEEEEIEEVT RRNLKLLADVVEEIRRVHEKTRRL LETVLRRAT(SEQ ID NO:27146)	2加1_钥匙 _C端_5421	EKAIRDVAKEIRDRLKEEEEIEEVT RRNLKLLADVVEEIRRVHEKTRRL ETVLRRAT(SEQ ID NO:27147)
2加1_笼_C端 _5432	DEIRKVVKEITDLLKASNDKNRK VVEEIRDLLRKSKKLADELVERLR ALVEDLRRRIDKSGDKETAEDIVR RIIEELKRILKEIEDLARRINREIER LVEEVERDNRDVNRAIEELLKDIA RRGSEDLKRVEERAREVSRNE ESMRRVKEDADVSEANKEVLDR VREEVKRLIEEVRETLR(SEQ ID NO:27148)	2加1_钥匙 _C端_5432	SEDLKRVEERAREVSRNEESMRRV KEDADVSEANKEVLDRVREEVKR LIEEVRETLR(SEQ ID NO:27149)
2加1_笼_C端 _5435	STAETVAEEVERVLKHSDDLIKEV EDVNRVVEEIKRVIRELEENER LVAEVRKGVKGEILAEIEKRLADN SEKVREVAERAKKLEENTARVK DILRESRKLVKDLLDEVRTGTGSEE	2加1_钥匙 _C端_5435	EAAAREIIKRLREVNKRTKEKLDLI KHSEEVLERVKRLIDELRKHSEEV EDLRRRAK(SEQ ID NO:27151)

	AAREIHKRLREVNKRTKEKLDELI KHSEEVLERVKRLIDELRKHSEEV LEDLRRRAK(SEQ ID NO:27150)		
2加1_笼_C端 _5439	SRVEEIIEDLRRLLLEEIRKENEDSIR RSKELLDRVKEINDTIAELERLLK DIEKEVREKGESESEEVKKALRAVL EELEKLLRRVAEINEEVLRRNSKL VEEDERRNREVLKELARLVEELIR EIGDEDKARKVAEVAEKVLRDID KLDRESKEAFRATNEEIAKLDEDT ARVAERVKKAIEDLAK(SEQ ID NO:27154)	2加1_钥匙 _C端_5439	EDKARKVAEVAEKVLRDIDKLDRE SKEAFRATNEEIAKLDEDTARVAER VKKAIEDLAK(SEQ ID NO:27155)
2加1_笼_C端 _5447	SEADDVLKCLAETVKRIIERLKKL TDDSRRLVEEVHRRNDKLSKESA EAVRKAERGIDEKDVRKLLLEDL KKKSEEVAERNKRILDTLREISKR AEDEVKVLKELEKTLKELEDRR PDSEELSAEVKLLDEVKALAR HKDENDKLLKEIEDSLRRHKEEN DRLLEKLKESTR(SEQ ID NO:27156)	2加1_钥匙 _C端_5447	EELSAEVKLLDEVKALARHKDE NDKLLKEIEDSLRRHKEENDRLEK LKESTR(SEQ ID NO:27157)
2加1_笼_C端 _5456	SAEELLREVAELVKRVDEDLRRL LEEVRASNEEVIRLEEILKRIEEE NRKVVEELRRGGVSEDLVRESKR LVDESRRVIEKLVKESADSVETR ETVDRLREELKRLVEEIAKMVKG GSSEETVKRLLDELRELLERLKRTI EELLKRNRDLLADAEEKARLLE ENRLLKAARDTAT(SEQ ID NO:27158)	2加1_钥匙 _C端_5465	EETVKRLLDELRELLERLKRTIEELL KRNRDLLADAEEKARLLEENRKL LKAARDTAT(SEQ ID NO:27159)
2加1_笼_N端 _2406	SEVDEVVKEVEDLVRNEELVEE VVRVVEKVVTDDRRLVEEVVREI RKIVKDVEDLARKLDKEELKRVL DEMREIERLLEKLRRHSSKLLDDE	2加1_钥匙 _N端_2406	SEVDEVVKEVEDLVRNEELVEEV VVRVVEKVVTDDRRLVEEVVREIRKI VKDVEDLARK(SEQ ID NO:27163)

	LKRLEELREHSRRVEKRLEDLLK ELRREGVDEKVLKLEKVIREVRE RSTRALRKVEEVIRRVREESERAL RDLERVVKEVEKRMREAAAR(SEQ ID NO:27162)		
2 加 1_ 笼_N 端 _5406	DRREVKKRLDEVREIERLLRRV EESRRVAEEIRRLIEVRRRNKK VTEIRELKGLKDKEEVRRVLER LRKLNAESELLERLERLRLRVE ATNRLVKAIEELRRLVEKIVREVP DSEELREELKKLERKIEKVAKEIH DHDKEVTERLEDLLRRITEHARKS DREIEETAR(SEQ ID NO:27164)	2 加 1_ 钥匙 _N 端_5406	DRREVKKRLDEVREIERLLRRVE EESRRVAEEIRRLIEVRRRNKKVTE EIRELLKGL(SEQ ID NO:27165)
2 加 1_ 笼_N 端 _5409	SEAEELKRLDRAEEILRRLEEIL RTSRKLAEDVLELEKLLRESERR IREVLEELRGIKDKKELEDVIREVE KELDESLERSRELLKDVLKLLDD NLKESERLVEDIDRELAKILEDLK KAGVPKEVVDEIKRIVDEVRERLK RIVDENAKIVEDARRALEKIVKEN EEILRRLKKELELRLK(SEQ ID NO:27166)	2 加 1_ 钥匙 _N 端_5409	SEAEELKRLDRAEEILRRLEEILRT SRKLAEDVLELEKLLRESERRIREV LEELRGI(SEQ ID NO:27167)
3 加_ 笼 _529_GFP11_C 端	SEAEDLEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL EESDRLEELKRRLEEILRKSHELL RRLEEVLREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLLEELVEVIKEILEKHLRLEELV RVIERILREVGKDKDEAERRDHM VLHEYVNAAGITEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27178)	3 加_ 钥匙_C 端_529	KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27179)
3 加 1_ 笼_C 端	SEKEELKRLDKLLKELKRLSDEL	3 加 1_ 钥匙	EDELKVEEDLKRLEDKLLKLLLEDY

<p>_500</p>	<p>KATIDKILKILKEVSEEVKRTADEL LDAIRRGVDEEVLREIKREIEEIE KKLRKVNKEIEDEIREIKKKLDEV DDKITKEVEKIKEALDKGGVDAK EVIKALKEILKEHADVFEDVLRRL KEIIRHRDVVKEVLEELRKILEK VAEVLKRQGRSEDELKVEEDLK RLEDKLLKLEDYEKKVRELEET LDDLRLKYETLRRLEKELEEAER (SEQ ID NO:27184)</p>	<p>_C 端_500</p>	<p>EKKVRELEETLDDLRLKYETLRRLE EKELEEAER(SEQ ID NO:27185)</p>
<p>3 加 1_笼_C 端 _510</p>	<p>SEKEELLKLIKRVIELLKRVL EEHL RLVEDVIRRLKELLSNEKIVREVI EDLKRLLEVRGDKEELDRIKEKL EEVLERYKRRLEIEIKRDLERMLED YKRELKRIEEDLRRVLEEVERIAT RGECPAEALIDKLRKILERALREL DKLSKKLDELKVKVLEEELEKSNR EIDKLLKDVLRVVEEGGASEDLLR KAKKVITEVREKLKRNLEDVRRVI EDVKRKSARILEEARRLIEEVERE LEKIRK(SEQ ID NO:27190)</p>	<p>3 加 1_钥匙 _C 端_510</p>	<p>EDLLRKAKKVITEVREKLKRNLEDV RRVIEDVKRKSARILEEARRLIEEVE RELEKIRK(SEQ ID NO:27191)</p>
<p>3 加 1_笼 _528_GFP11_C 端</p>	<p>SEAEDLEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL EESDRLLEELKRRLEEILRKSHELL RRLEEVLRLEIKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLLEELVEVIKEILEKHLRLEELV RVIERILREVGRDHMVLHEYVNA AGITLDRLEEEHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27192)</p>	<p>3 加 1_钥匙 _C 端_528</p>	<p>KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27193)</p>
<p>3 加 1_笼 _528_GFP11_C</p>	<p>SEAEDLEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE</p>	<p>3 加 1_钥匙 _C 端_528</p>	<p>KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL</p>

端	ELRKGTVEDKDILRDLERRLEIL EESDRLLEELKRRLEEILRKSHELL RRLEEVLREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLLEELVEVIKEILEKHLRLLEELV RVIERILREVGKDKRDHMLHEVY VNAAGITLREEHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27194)		KRVDRST(SEQ ID NO:27195)
3 加 1_笼 _528_GFP11_C 端	SEAEDEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL EESDRLLEELKRRLEEILRKSHELL RRLEEVLREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLLEELVEVIKEILEKHLRLLEELV RVIERILREVGKDKDEAERDHMV LHEYVNAAGITHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27196)	3 加 1_钥匙 _C 端_528	KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27197)
3 加 1_笼 _529_GFP11_C 端	SEAEDEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL EESDRLLEELKRRLEEILRKSHELL RRLEEVLREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLLEELVEVIKEILEKHLRLLEELV RVIERILREVGKRDHMLHEYVN AAGITDRLREEHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27198)	3 加 1_钥匙 _C 端_529	KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27199)
3 加 1_笼 _529_GFP11_C 端	SEAEDEELIKELAELLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL	3 加 1_钥匙 _C 端_529	KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27201)

	EESDRLLEELKRRLEEILRKSHELL RRLEEVLRREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLEELVEVIKEILEKHLRLEELV RVIERILREVGKDRDHMLHEVY NAAGITRLREEHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27200)		
3 加 1_笼 _529_GFP11_C 端	SEAEDLEELIKELAECLKDVIRKLE KINRRLVKILEDIIRRLKEISKEAEE ELRKGTVEDKDILRDLERRLEIL EESDRLLEELKRRLEEILRKSHELL RRLEEVLRREILKRAEEVKRSNLPK EELIKEIVKLEELLRVIEKILEDNI RLEELVEVIKEILEKHLRLEELV RVIERILREVGKDKDRDHMLHE YVNAAGITREEHEEVKRRLEEELT RLRETHKKIEKELREALKRVDRS T(SEQ ID NO:27202)	3 加 1_钥匙 _C 端_529	KDEAERRRRELKDKLDRLREEHEEV KRRLEEELTRLRETHKKIEKELREAL KRVDRST(SEQ ID NO:27203)
3 加 1_笼 _534_GFP11_C 端	DEDRIIEIARLLEELLRELELIKK LIETNRRLNEEHERAVRELARLLE ELLDRLVKKGISDEKLRIRERLK RALDDLRLHREINKRLEDLVREL EKLVRREILKELKDALEELRRASAR AGGEEVLRREEIVKLLDLVRI LERLKEIHKDNVRLRELNERLTR IVEDLVRLIREILREAGVDERDHM VLHEYVNAAGITIKRLHEDLERKL KESEDELREIARLEEKIRRLLEKL ERKRR(SEQ ID NO:27206)	3 加 1_钥匙 _C 端_534	EKIAEEIERELEELRRMIKRLHEDLE RKLKESEDELREIARLEEKIRRLLEE KLERKRR(SEQ ID NO:27207)
3 加 1_笼 _534_GFP11_C 端	DEDRIIEIARLLEELLRELELIKK LIETNRRLNEEHERAVRELARLLE ELLDRLVKKGISDEKLRIRERLK RALDDLRLHREINKRLEDLVREL	3 加 1_钥匙 _C 端_534	EKIAEEIERELEELRRMIKRLHEDLE RKLKESEDELREIARLEEKIRRLLEE KLERKRR(SEQ ID NO:27209)

	<p>EKLVREILKELKDALEELRRASAR AGGEEVLRRLLEEIVKLLDLVRRRI LERLKEIHKDNVRLRELNERLTR IVEDLVRLIREILREAGVDEKIRDH MVLHEYVNAAGITRLHEDLERKL KESEDELREIEARLEEKIRRLLEEKL ERKRR(SEQ ID NO:27208)</p>		
<p>3 加 1_笼 _534_GFP11_C 端</p>	<p>DEDRIIEIARLLEELLRELLELIKK LIETNRRLNEEHERAVRELARLLE ELLDRLVKKGISDEKLRIRERLKL RALDDLERLHREINKRLEDLVREL EKLVREILKELKDALEELRRASAR AGGEEVLRRLLEEIVKLLDLVRRRI LERLKEIHKDNVRLRELNERLTR IVEDLVRLIREILREAGVDEKIAEEI ERDHMVLHEYVNAAGITLERKCLK ESEDELREIEARLEEKIRRLLEEKLE RKRR(SEQ ID NO:27210)</p>	<p>3 加 1_钥匙 _C 端_534</p>	<p>EKIAEEIERELEELRRMIKRLHEDLE RKLKESDELREIEARLEEKIRRLLEE KLERKRR(SEQ ID NO:27211)</p>
<p>3 加 1_笼_C 端 _539</p>	<p>SEKEKLLKESEEEVRRRLRRTLEEL LRKYREVLERLRKELREIEERVD VVRRLKEVLDKGLDIDTIIKEVE DLLKTVLDRLRELLDKIRRLTKEA IEVVREIIRIVRHAERVKDELKKE GGDKEKLDKDRVDRLIKENTRHLKEI LDRIEDLVRSEKLRDIIEVRRL IEELRKAEBIKKGPDERLVKTLIE DVEAVIKRILELITRVAEDNERVL ERIIRELTDNLERHLKIVREIVK(SE Q ID NO:27212)</p>	<p>3 加 1_钥匙 _C 端_539</p>	<p>ERLVKTLIEDVEAVIKRILELITRVAE DNERVLERIIRELTDNLERHLKIVREI VK(SEQ ID NO:27213)</p>
<p>3 加 1_笼_C 端 _548</p>	<p>DKAEVLREALKLLKDLLEELIKIH EESLKRILDIDTLVKVHEDALRA LKELLERSGLDERELRKVERMAT ESLRTIAKLKEELRDLARRSLEKL REDLKRVDLTLRKVEEKVVRTGP</p>	<p>3 加 1_钥匙 _C 端_548</p>	<p>EEIDRELKRVVEELRRLHEEIKERLD DVARRSEELRRIKKLKEVVKEIRK KKLK(SEQ ID NO:27215)</p>

	SEELIEELIRTIEKLLKEIVRINEEV LKAVRELLKTLKLSQDVVRRIEE ILRKGVPPEIDRELKRVVEELRR LHEEIKERLDDVARRSEEELRRIK KLKEVVKEIRKLLK(SEQ ID NO:27214)		
3加1_笼_C端 _556	SERELIERWLELHKEILRLIRELVE RLLKLHREILDITIKKLIRELLELLE DIARKLGLDKEAKDELREIAKRV DKLEKLERESRKVEEDLKRKLKE LTDESQVVEKRVQDVVRRGTQSR EEIAEELLRLDRKLLKAVEELLKEI LDLNKLLDDVRAILEETRRVLEK LLDRVRRGERTDDERRTLTELLK RMEDILEKVERTLKKLLDSSARM AEEVKKTLKELLERSEKVAEDVR K(SEQ ID NO:27216)	3加1_钥匙 _C端_556	DDERRTLTELLKRMEDILEKVERTL KKLLDSSARMAEEVKKTLKELLER SEKVAEDVRK(SEQ ID NO:27217)
3加1_笼_C端 _560	SKELLEEVVRAIELLKRHLEKL KRILEEIVRLLEEHLKVERVLEAI LSLLDDLLRRGGDERAIRTLEDVK RRLREILERLADENAKAIKRLADL LDKLEKRNKEAIERLEEILEELKR VRRDEELLRVLETLLKIIEDILREN TKVLEDLLRLVEEILEANLRVVEE LLRLAREILTEIVGDEDKLEIEDE LRRLEELRRDKAIKDRLRELKK DLDEANRRIKETLKKLLREVEK(S EQ ID NO:27218)	3加1_钥匙 _C端_560	EDKLKEIEDELRRLEELRRDKAIK DRLRELKKDLDEANRRIKETLKKLL REVEK(SEQ ID NO:27219)
3加1_笼 _568_GFP11_C 端	KEIEETLKELEDLNREMVETNRRV LEETRRLNKETVDRVKATLDELA KMLKKLVDDVRKGPTSEELKRL AELEELLARVRRVEELLKSTD LERAVKDSADALRRSHEVLKEVA SRVKRAKDEGLPREEVLRLRELL	3加1_钥匙 _C端_568	KAVEELEKALEEIKRRLKEVIDRYE DELRLRKEYKEKIDKYERKLEEIE RRERT(SEQ ID NO:27221)

	ERHAKVLKDIVRVSEKLLREHLK VLREIVEVLEELLERILKVILDTR DHMVLHEYVNAAGITKRRLKEVI DRYEDELRLKLRKEYKEKIDKYER KLEEIERRERT(SEQ ID NO:27220)		
3 加 1_笼 _568_GFP11_C 端	KEIEETLKELEDLNREMVETNRRV LEETRRLNKETVDRVKATLDELA KMLKKLVDDVRKGPTSEELKRL AELEELLARVRRVEELLKSTD LERAVKDSADALRRSHEVLKEVA SRVKRAKDEGLPREEVLRLRELL ERHAKVLKDIVRVSEKLLREHLK VLREIVEVLEELLERILKVILDTTG GDRDHMVLHEYVNAAGITLKEVI DRYEDELRLKLRKEYKEKIDKYER KLEEIERRERT(SEQ ID NO:27222)	3 加 1_钥匙 _C 端_568	KAVEELEKALEEIKRRLKEVIDRYE DELRKLKLRKEYKEKIDKYERKLEEIE RRERT(SEQ ID NO:27223)
3 加 1_笼_C 端 _581	SALETVKKLEDSSEKIERIVEEDE RVAKESSDRIRRLVEEDKRVAD LDIEKIGD TD TLLKLV EEW SRTS KKLDDV LK LHKDWS D SRR LLE EILRVHEELIRRVKEILDREGKPEE VVRELEKVLKESLDTLEEIIRRLDE ANAATVKRVADVIRELEDINRKY LEEIKRGSDDAEAVIKVIEKLIRAN KRVWDALLKINEDLVRVNKT VW KELLRVNEKLARDLERVVK (SEQ ID NO:27226)	3 加 1_钥匙 _C 端_581	AEAVIKVIEKLIRANKRVWDALLKI NEDLVRVNKT VWKELLRVNEKLAR DLERVVK(SEQ ID NO:27227)
3 加 1_笼_C 端 _585	SKEEKLKDDVRAVLEDLDRVLKE LEKLS EDNLRELKRVLD RITDLHR RILDELKRGIGSEELLRRVEKVLK DNLDLLRKLVEEHKESSE RDLKR VEDLVREI KEVLRKLELED RGTD IRKIEEIEIRLLRKIRKAVEESKDL NRRNSERIEEVARRSEELARRLLK	3 加 1_钥匙 _C 端_585	SKAAEDILRVLEKLVKVSREAIKLIL ELSEHHVRVSTRIARLLLDVARKLA EVIKEAER(SEQ ID NO:27229)

	EIRERGDSKAAEDILRVLEKLVKV SREAIKLILELSEHHVRVSTRIARL LLDVARKLAEVIKEAER(SEQ ID NO:27228)		
3 加 1_笼_C 端 _587	SEIEDVIRRLRKILEDLERVSEKLL REIKKILDEARRLNEEVIKEIKRVL EDAVRVFRDGSKEELAKLVEE LIRELAKLAKEVDEIHKRIVERLK ALVEDAERIHRKIVETLEEIVRGVP SEELKRVVEAIVEVIKEHLKVLAD VIRRIKAIBENAETIKRVLEDIVRV LELVLRGEGSIEDLVREVERLIKRI EDSLRELEKTVRELLKRIKEASDK VREDVDRLIKELKEAAD(SEQ ID NO:27230)	3 加 1_钥匙 _C 端_587	IEDLVREVERLIKRIEDSLRELEKTV RELLKRIKEASDKVREDVDRLIKEL KEAAD(SEQ ID NO:27231)
3 加 1_笼_C 端 _605	SREELDRILEAIAKILEDLKRLLIDE NLARLEEVVRELERIIDRNKLIRE ILDELKKGSGSEEILEKIKKVDKEL EDLIRRLKLEDLIRETERRLREI LKRIRDLLKEVKDRDKDLERLLE VLEEVLRVIAELAKELDSLKVL KVVEEVLRLNNEVNKEVLDVIRE LAKDGGSEIIRKLDLLKEVEKV HKEVKDRIRKLEDHKRSLDEVK KKLERLLERAKEVVEREKK(SEQ ID NO:27232)	3 加 1_钥匙 _C 端_605	DEIRKLDLLEKVEKVHKEVKDRIR KLEDHKRSLDEVKKKLERLLERAK EVVEREKK(SEQ ID NO:27233)
3 加 1_笼_C 端 _607	SEREELLERIKEILKRVKDKLDED LKRLKEILEKLKEKADRDLLELRR RIEEVREKLERTGRTDLVKEVLD TVRRNLENLKRLEDILRKLLEENV KNLTDLVREILKLITELIKRLEDGG LPKEVLDALRRVLEKLEELLREIL ERLKRSLAVKRKIEELLKELERS LDELRRALERIRKEIGDSETAVRAI	3 加 1_钥匙 _C 端_607	SETAVRAIRVLEKHLEAVRRVLEEL LKVLAEHLETVRELIERLKRVLLEEAI EVVERVAR(SEQ ID NO:27235)

	<p>IRVLEKHLEAVRRVLEELLKVLAE HLETVRELIERLKRVL EEAEIEVVE RVAR(SEQ ID NO:27234)</p>		
<p>3 加 1_笼 _611_GFP11_C 端</p>	<p>SLEEITKRILLEVEENLARHEEILR ELLELAKRLAKEDRDILEEVLKLI EELLKILLEDNGSSEEDLKRLIKEV IEELRAVVKRVKDKWDEVVKRIE DLVKKLKELHDDTLRKLRELVRK IVTDISESGGEAEKV KRVVEKILEL VERLAKVVKESVEKLLEILRELAE VSKRVAEALLRLLLEELVRVIRIKD ERDHMVLHEYVNAAGITLLDELE EVHKRVKKELEDIIEENRRVVKRV RDELREIKRELDE(SEQ ID NO:27238)</p>	<p>3 加 1_钥匙 _C 端_611</p>	<p>ERTLREVVRKVLEEAKRLLDELEEV HKRVKKELEDIIEENRRVVKRVRDE LREIKRELDE(SEQ ID NO:27239)</p>
<p>3 加 1_笼_C 端 _632</p>	<p>SEKELVDDIRILEEILRLLRSLEE VIRLLEENEKLVRRHLKTVIDILRR VAKLLDENGIRTDEADRVLERLE KAHRELLEDYKRALEKIKETLERV LREAEVVKKIDDALRKLGGKSKE VLKRLEELLRLVEKIAEEIKRLLS ELVRVTEELVRTNKELLEEA VRVI RKEVGDDSLVREVEELIKRLEKHI DDLKTSRDLVKRVLDLVDEVVK RVEDLVERVKEKIDT(SEQ ID NO:27240)</p>	<p>3 加 1_钥匙 _C 端_632</p>	<p>DSL VREVEELIKRLEKHIDLLKTSR DLVKRVLDLVDEVVKRVEDLVERV KEKIDT(SEQ ID NO:27241)</p>
<p>3 加 1_笼 _646_GFP11_C 端</p>	<p>DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKA EKILERLN RVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTL VKEYA DLSKEVLER VERIVREYVKLSDEV VKSLAEIVEELIRIIEDLLRKGNRD</p>	<p>3 加 1_钥匙 _C 端_646</p>	<p>EDVKRAEELVSRLRKLEDVKKAS EDIVREVERIVRELAKRSDEILK KLE DIVEKLRE(SEQ ID NO:27245)</p>

	HMVLHEYVNAAGITRKLLEDVKK ASEDIVREVERIVRELAKRSDEILK KLEDIVEKLRE(SEQ ID NO:27244)		
3 加 1_笼 _646_GFP11_C 端	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKAEKILERLNRVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTLKEYA DLSKEVLERVERIVREYVKLSDEV VKSLAEIVEELIRIIEDLLRKGND EDRDHMLHEYVNAAGITEDVK KASEDIVREVERIVRELAKRSDEIL KKLEDIVEKLRE(SEQ ID NO:27246)	3 加 1_钥匙 _C 端_646	EDVKRALEELVSRLRKLEDVKKAS EDIVREVERIVRELAKRSDEILKLE DIVEKLRE(SEQ ID NO:27247)
3 加 1_笼 _646_GFP11_C 端	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKAEKILERLNRVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTLKEYA DLSKEVLERVERIVREYVKLSDEV VKSLAEIVEELIRIIEDLLRKGND EDVRDHMLHEYVNAAGITDVK KASEDIVREVERIVRELAKRSDEIL KKLEDIVEKLRE(SEQ ID NO:27248)	3 加 1_钥匙 _C 端_646	EDVKRALEELVSRLRKLEDVKKAS EDIVREVERIVRELAKRSDEILKLE DIVEKLRE(SEQ ID NO:27249)
3 加 1_笼 _647_GFP11_C 端	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKAEKILERLNRVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTLKEYA DLSKEVLERVERIVREYVKLSDEV	3 加 1_钥匙 _C 端_647	EDVKRALEELVSRLRKLEDVKKAS EDIVREVERIVRELAKRSDEILKLE DIVEKLRE(SEQ ID NO:27251)

	VKSLAEIVEELIRIIEDLLRKG NLR DHMVLHEYVNAAGITKLEDVK KASEDIVREVERIVRELAKRSDEIL KKLEDIVEKLRE(SEQ ID NO:27250)		
3 加 1_笼 _647_GFP11_C 端	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKAEKILERLNRVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTLKEYA DLSKEVLERVERIVREYVKLSDEV VKSLAEIVEELIRIIEDLLRKG NLD EDVKRALERDHMVLHEYVNAAG ITSEDIVREVERIVRELAKRSDEIL KKLEDIVEKLRE(SEQ ID NO:27252)	3 加 1_钥匙 _C 端_647	EDVKRALEELVSRLRKLEEDVKKAS EDIVREVERIVRELAKRSDEILKKLE DIVEKLRE(SEQ ID NO:27253)
3 加 1_笼_C 端 _647	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWKKLVEDIA EILRRIVELLRRGGVPEELLDRLA KVVKSIVEKAEKILERLNRVSKAI AEKLTIVDELNEVSKEIVKRAED ILRKGKDKETVLRALRTLKEYA DLSKEVLERVERIVREYVKLSDEV VKSLAEIVEELIRIIEDLLRKG NLD EDVKRALEELVSRLRKLEEDVKK ASEDIVREVERIVRELAKRSDEILK KLEDIVEKLRE(SEQ ID NO:27254)	3 加 1_钥匙 _C 端_647	EDVKRALEELVSRLRKLEEDVKKAS EDIVREVERIVRELAKRSDEILKKLE DIVEKLRE(SEQ ID NO:27255)
3 加 1_笼_C 端 _653	DEEETLRLLERKVELAKEYLDVS KEVIDRTTKLLDEYLKTSKRIVDA TVELLERGDLPDELKRLAEELE RSLRELEEEIKRLKRELEESLKKLK EIIDRLAEEAEKLLAVLKRGESE EEALRALASLVRELIEVLENDER	3 加 1_钥匙 _C 端_653	SREALEEARRLEELLRELNEITKDL EAKLEKLLRDLNELTKALEEELKRL LDELKKRTD(SEQ ID NO:27257)

	LRDVLRRLLIEALRKNNEILERVLR KLVRAAEERGRDESSREALEEAR RRLEELLRELNEITKDLEAKLEKL LRDLNELTKALEEELKRLLDELKK RTD(SEQ ID NO:27256)		
3 加 1_笼_C 端 _658	DEERIKTLEDINAKLVEDIKRILD KVAELNERLADAIKILEETKRILE ATTRKVRKDGEISEELLRRLEEKL RKLEDLERVLAHEHEDESRRILEE VERLLKRHADASKELDRARVA RGVKSDELVDRLKKLIDDSLESV RELIERLKELLDRLVKSVEDLIRTI KELLDRLVEVLREGVSDKDTLRT VEKLVEDVKRRLDKLEEDYKRLI EEVKKELDKLLKEYEDALREIKK RIDE(SEQ ID NO:27258)	3 加 1_钥匙 _C 端_658	KDTLRTVEKLVEDVKRRLDKILLED YKRLIEEVKKELDKLLKEYEDALRE IKKRIDE(SEQ ID NO:27259)
3 加 1_笼_N 端 _263	SLVDELRKSLERNVRVSEEVARRL KEALKRWVDVVRKVVEDLIRLNE DVVRVVEKVTVDESAIERVRRRIE ELNRKLDVAVLKKNEDLVRRLTEL LDKLEENRRLVEEDEDLKRKG GTEEVIDTILELIERSIERLKRLLE LLRIVREALKDNKRVA DENLKKL KEILDELKRGVEEELKRVLEK AADLHRRLKDRHRKLEDLERIR ELKKKLDEVVEENKRSVDELKR(S EQ ID NO:27262)	3 加 1_钥匙 _N 端_263	SLVDELRKSLERNVRVSEEVARRL EALKRWVDVVRKVVEDLIRLNEDV VRVVEKV(SEQ ID NO:27263)
3 加 1_笼 _647_GFP11_ N 端	DAEEVVKRLADVLRENDETIRKV VEDLVRIAENDRLWRDHMVLHE YVNAAGITLLRRGGVPEELDRL AKVVK SIVEKA EKILERLNRVSKA IAEKLKTIVDELNEVSKEIVKRAE DILRKGKDKETVLRALRTLKEY ADLSKEVLERVERIVREYVKLSDE	3 加 1_钥匙 _N 端_647	DAEEVVKRLADVLRENDETIRKVVE DLVRIAENDRLWKKLVEDIAEILR RIVELLRRG(SEQ ID NO:27277)

	VVKSLAEIVEELIRIIEDLLRKGNL DEDVKRALEELVSRLRKLEEDVK KASEDIVREVERIVRELAKRSDEIL KKLEDIVEKLRE(SEQ ID NO:27276)		
--	---	--	--

[0103] 在各种具体实施方案中,笼蛋白包含与选自SEQ ID NO:27497-27620的笼蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,其中所述N端蛋白纯化标签(MGSHHHHHHGGSGSENLYFQGG (SEQ ID NO:27624);或MGSHHHHHHGGSENLYFQG (SEQ ID NO:27625);或GSHHHHHHGGSGSENLYFQG (SEQ ID NO:27626))是任选的,在同一性百分比比较中不予考虑,并且可以存在或不存在。在一个实施方案中,N端蛋白纯化标签不存在。

表10. 氨基酸序列

(以下序列包含用于蛋白纯化目的的6His TEV标签MGSHHHHHHGGSGSENLYFQG (SEQ ID NO:27495)或其变体。结构区的N端氨基酸是可选的,在与所要求保护的笼蛋白相关的同一性百分比比较中不考虑(结构区在括号内)括号内的C端区构成门锁区。

SmBit序列(VTGYRLFEEIL) (SEQ ID NO:27359) 加下划线。

传感结构域以粗体显示

lucCageBim变体(Bcl2传感器)

-SmBit序列:VTGYRLFEEIL (SEQ ID NO:27359)

-BIM序列:**EIWIAQELRRIGDEFNAYYA** (SEQ ID NO:27496)

>nluc301_bim331

**MGSHHHHHHGGSGSENLYFQGG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAARKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL**

QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLsREEIWIAQELRRIGDEF
NAYYAAASEKISRE (SEQ ID NO:27497)

>nLuc308_bim331

MGSHHHHHHGGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARLL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLsREEIWIAQELRRIGDEFNA
YYAAASEKISRE (SEQ ID NO: 27498)

>nLuc312_bim331

MGSHHHHHHGGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARLL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREEIWIAQELRRIGDEFNA
YYAAASEKISRE (SEQ ID NO: 27499)

>nLuc315_bim331

MGSHHHHHHGGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARLL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIARVKVTGYRLFEEILRLsREEIWIAQELRRIGDEFNA
YYAAASEKISRE (SEQ ID NO: 27500)

>nLuc301_bim339

MGSHHHHHHGGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA

AAKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLsREAAAASEKIEIWIAQE
LRRIGDEFNAYYAE (SEQ ID NO: 27501)

>nLuc308_bim339

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEI KEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLsREAAAASEKIEIWIAQELR
RIGDEFNAYYAE (SEQ ID NO: 27502)*

>nLuc312_bim339

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEI KEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKIEIWIAQELR
RIGDEFNAYYAE (SEQ ID NO: 27503)*

>nLuc315_bim339

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEI KEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIARVKVTGYRLFEEILRLsREAAAASEKIEIWIAQELR
RIGDEFNAYYAE (SEQ ID NO: 27504)*

>nLuc301_bim343

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLsREAAAASEKISREAEIWI
AQELRRIGDEFNAYYA (SEQ ID NO: 27505)*

>nLuc308_bim343

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLsREAAAASEKISREAEIWI
AQELRRIGDEFNAYYA (SEQ ID NO: 27506)*

>nLuc312_bim343

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEIWI
AQELRRIGDEFNAYYA (SEQ ID NO: 27507)*

>nLuc315_bim343

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIARVKVTGYRLFEEILRLsREAAAASEKISREAEIWI
AQELRRIGDEFNAYYA (SEQ ID NO: 27508)*

lucCageTrop变体(心肌钙蛋白I传感器)

-SmBit序列:VTGYRLFEEIL(SEQ ID NO:27359)

-使用的心肌钙蛋白T(cTnT)变体序列:

- cTnTf1:226-**EDQLREKAKELWQTI**-240 (SEQ ID NO:27385)
- cTnTf2:226-**EDQLREKAKELWQTIYN**-242 (SEQ ID NO:27386)
- cTnTf3:226-**EDQLREKAKELWQTIYNLEAE**-246 (SEQ ID NO:27387)
- cTnTf4:226-**EDQLREKAKELWQTIYNLEAEKFD**-249 (SEQ ID NO:27388)
- cTnTf5:226-**EDQLREKAKELWQTIYNLEAEKFDLQE**-252 (SEQ ID NO:27389)
- cTnTf6:226-
EDQLREKAKELWQTIYNLEAEKFDLQEKFKQKYEINVLRNRINDNQ-
272 (SEQ ID NO:27390)

-cTnC:

**KVSKTKDDSKGKSEEELSDLFRMFDKNADGYIDLEELKIMLQATGETITED
DIEELMKDGDKNNDGRIDYDEFLEFMKGV** (SEQ ID NO:27627)

>336-cTnTf4-K342A (jp625_1fix_nluc312_cTnT336_K342A_359end)

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASE**EDQLREaAKEL
WQTIYNLEAEKFD*** (SEQ ID NO: 27509)

>336-cTnTf6-K342A (jp626_1fix-nluc312_cTnT336_K342A_362end)

*MGSHHHHHHGSSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASE**EDQLREaAKEL
WQTIYNLEAEKFDLQE*** (SEQ ID NO: 27510)

>336-cTnTf6-K342A (jp627_1fix-nluc312_cTnT336_K342A_0001_382end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEDQLREaAKEL
WQTIYNLEAEKFDLQEKFKQQKYEINVLRNRINDNQ (SEQ ID NO: 27511)

>339-cTnTf3 (jp628_1fix-nluc312_cTnT339_359end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKIEDQLREKAKEL
WQTIYNLEAE (SEQ ID NO: 27512)

>339-cTnTf5 (jp629_1fix-nluc312_cTnT339_0001_365end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKIEDQLREKAKEL
WQTIYNLEAEKFDLQEKFKQQKYEINVLRNRINDNQ (SEQ ID NO: 27513)

>339-cTnTf6 (jp630_1fix-nluc312_cTnT339_0001_385end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKIEDQLREKAKEL

WQTIYNLEAEKFDLQEKFKQQKYEINVLNRNRINDNQ
(SEQ ID NO: 27514)

>343-cTnTf2 (jp631_1fix-nluc312_cTnT343_359end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEDQLREKA
KELWQTIYN (SEQ ID NO: 27515)

>343-cTnTf5 (jp632_1fix-nluc312_cTnT343_0001_369end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEDQLREKA
KELWQTIYNLEAEKFDLQEKFKQQKYEINVLNRNRINDNQ (SEQ ID NO: 27516)

>343-cTnTf6 (jp633_1fix-nluc312_cTnT343_0001_389end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREA
EDQLREKAKELWQTIYNLEAEKFDLQEKFKQQKYEINVLNRNRINDNQKFKQQKYEI
NVLNRNRINDNQ (SEQ ID NO: 27517)

>345-cTnTf1 (jp634_1fix-nluc312_cTnT345_359end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI

RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEREDQLRE
KAKELWQTI (SEQ ID NO: 27518)

>345-cTnTf5(jp635_1fix-nluc312_cTnT345_0001_371end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEREDQLRE
KAKELWQTIYNLEAEKFDLQEKFKQKQKYEINVLNRINDNQ (SEQ ID NO: 27519)

>345-cTnTf6(jp636_1fix-nluc312_cTnT345_0001_391end)

MGSHHHHHHGSSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIA
RLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAI
RDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEKISREAEREDQLRE
**KAKELWQTIYNLEAEKFDLQEKFKQKQKYEINVLNRINDNQKFKQKQKYEINVLNR
INDNQ** (SEQ ID NO: 27520)

>lucCageTrop

MGSHHHHHHGSSENLYFQGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLsREAAAASEDQLREaAKEL
**WQTIYNLEAEKFDLQEKFKQKQKYEINVLNRINDNQKVSCTKDDSKGKSEEEELSDLF
RMFDKNADGYIDLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFM
KGVE** (SEQ ID NO: 27521)

lucCageBot变体(肉毒杆菌神经毒素B传感器)

-Bot.0671.2序列:

MFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE (SEQ ID NO:
27381)

>BoNTB_338_1S

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKMF
E

LKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO: 27522)

>BoNTB_341_1S

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRM
FAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO: 27523)

>BoNTB_342_1S

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
MFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO: 27524)

>BoNTB_345_1S

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
AERMFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO:
 27525)

>BoNTB_348_2S

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIRMF AELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO:
 27526)

>BoNTB_349_2S

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIREMFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ ID NO:
 27527)

>BoNTB_352_2S

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIREAAAMFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE* (SEQ
 ID NO: 27528)

>BoNTB_355_2S

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIREAAAASEMFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE*
 (SEQ ID NO: 27529)

>BoNTB_GGG_2S

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIREAAAASEKISREGGGMFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIR
K

YELE* (SEQ ID NO: 27530)

>BoNTB_GGG_2S_fullBotBinder

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRE
 AERSIREAAAASEKISREGGGS~~HM~~QPMFAELKAKFFLEIGDRDAARNALRKAGYSDEE
AE

RIIRKYELE* (SEQ ID NO: 27531)

lucCageProA变体 (Fc结构域生物传感器)

-金黄色葡萄球菌蛋白A结构域C (SpaC) 序列:

EQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAQAPK

(SEQ ID NO:27382)

>SpaC_360GGG

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERSIREAAAASEKISREGGGFNKEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVS
KEILAEAKKLNDAQAPK* (SEQ ID NO: 27532)

>SpaC_354-2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
 AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
 ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
 EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
 EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
 LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
 REAERSIREAAAASEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLN
DAQAPK*

(SEQ ID NO: 27533)

>SpaC_351_2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
 AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
 ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
 EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
 EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
 LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
 REAERSIREAAEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAQ
APK*

(SEQ ID NO: 27534)

>SpaC_350_2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
 AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
 ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
 EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
 EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
 LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
 REAERSIREAEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAQA
PK*

(SEQ ID NO: 27535)

>SpaC_347_2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
 AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
 ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
 EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
 EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
 LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
 REAERSIEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAQAPK*

(SEQ ID NO: 27536)

>SpaC_347_1S

**MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERLIEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLND AQAPK***
(SEQ ID NO: 27537)

lucCageHer2变体 (Fc结构域生物传感器)

-Her2亲和抗体序列:

EMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLND AQAPK
(SEQ ID NO:27383)

>AffiHer2_347_1S

**MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERLIEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLND AQAP
K***

(SEQ ID NO: 27538)

>AffiHer2_347_2S

**MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERSIEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLND AQAP
K***

(SEQ ID NO: 27539)

E

>AffiHer2_350_2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKIS
REAERSIREAEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKLNDQA
APK*

(SEQ ID NO: 27540)

>AffiHer2_351_2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKIS
REAERSIREAAEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKLND
QAPK*

(SEQ ID NO: 27541)

>AffiHer2_354-2S

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKIS
REAERSIREAAAASEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQPK*

(SEQ ID NO: 27542)

>AffiHer2_360GGG

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIE
EARRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKIS
REAERSIREAAAASEKISREGGGVDNKNFNKEMRNAYWEIALLPNLNNQKRAFIRSLYD
DPSQSANLLAEAKKLNDQAQPK* (SEQ ID NO: 27543)

>AffiHer2_354-2S_2x1

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERSIREAAAASEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPKGGGNKEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPK* (SEQ ID NO: 27544)

>AffiHer2_354-2S_2x2

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERSIREAAAASEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPKGGGNKEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPK* (SEQ ID NO: 27545)

>AffiHer2_354-2S_3x

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEE
ARKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFL
EAAAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIE
EARLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLREL
LRLAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKIS
REAERSIREAAAASEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPKGGGNKEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPKGGGNKEMRNAYWEIALLPNLNNQKRAFIRSLYDDPSQSANLLAEAKKL
NDAQAPK* (SEQ ID NO: 27546)

lucCageSARS2N变体(anti-SARS-CoV-2 Nucleocapsid protein antibodies传感

器)

-使用的SARS-Cov-2核衣壳蛋白表位肽:

- N6:

**PKKDKKKKADETQALPQRQKKGGSGGPKDKKKKKADETQALPQRQK
K (SEQ ID NO:27547)**

- N62: **KKDKKKKKADETQALGGSGGKKDKKKKKADETQAL (SEQ ID
NO:27548)**

>lucCageSARS2-N6_368-388_339

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEYKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEK**PKKD
KKKKADETQALPQRQKKGGSGGPKDKKKKKADETQALPQRQKK*** (SEQ ID NO:
27549)

>lucCageSARS2-N6_368-388_346

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEYKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISRE
AERS**PKDKKKKKADETQALPQRQKKGGSGGPKDKKKKKADETQALPQRQKK***
(SEQ ID NO: 27550)

>lucCageSARS2-N6_368-388_353

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEYKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISREA
ERSIREAAAAP**PKDKKKKKADETQALPQRQKKGDRADLRKTKRRKPTKPKHCRNVK
KS** (SEQ ID NO: 27551)

>lucCageSARS2-N62_369-382_336

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASKKDKK
KKADETQALGGSGGKKDKKKKADETQAL* (SEQ ID NO: 27552)

>lucCageSARS2-N62_369-382_340

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKIKKD
KKKKKADETQALGGSGGKKDKKKKADETQAL* (SEQ ID NO: 27553)

>lucCageSARS2-N62_369-382_343

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISREA
KKDKKKKADETQALGGSGGKKDKKKKADETQAL* (SEQ ID NO: 27554)

>lucCageSARS2-N62_369-382_347

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISREA
ERSIKDKKKKADETQALGGSGGKKDKKKKADETQAL* (SEQ ID NO: 27555)

>lucCageSARS2-N62_369-382_350

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISREA
 ERSIREAKKDKKKKKADETQALGGSGGKKDKKKKKADETQAL* (SEQ ID NO: 27556)

>lucCageSARS2-N62_369-382_354

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISREA
 ERSIREAAAASKKDKKKKKADETQALGGSGGKKDKKKKKADETQAL* (SEQ ID NO:

27557)

lucCageSARS2M变体(抗SARS-Cov-2膜蛋白抗体传感器)

-使用的SARS-Cov-2膜蛋白表位肽:

- M1_1-31:**MADSNGTITVEELKKLLEQWNLVIGFLFTWIGGSGGMAD
 SNGTITVEELKKLLEQWNLVIGFLFTWI** (SEQ ID NO:27393)

- M3_1-17:**MADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE**
 (SEQ ID NO:27392)

- M4_8-24:**ITVEELKKLLEQWNLVIGGSGGITVEELKKLLEQWNLVI**
 (SEQ ID NO:27394)

>lucCageSARS2-M3_1-17_341

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
 KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPA TIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
 RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
 ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILD AERLSREAAAASEKISRM
 ADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE* (SEQ ID NO: 27558)

>lucCageSARS2-M3_1-17_343

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREA
MADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE* (SEQ ID NO: 27559)

>lucCageSARS2-M3_1-17_348

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREA
ERSIRMADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE* (SEQ ID NO:
27560)

>lucCageSARS2-M3_1-17_350

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREA
ERSIREAMADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE* (SEQ ID NO:
27561)

>lucCageSARS2-M4_8-24_334

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAITVEELKK
LLEQWNLVIGGSGGITVEELKKLLEQWNLVI* (SEQ ID NO: 27562)

>lucCageSARS2-M4_8-24_340

GSHHHHHHGSSENLYFQG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISITV
EELKKLLEQWNLVIGSGGITVEELKKLLEQWNLVI* (SEQ ID NO: 27563)

>lucCageSARS2-M4_8-24_341

GSHHHHHHGSSENLYFQG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRIT
VEELKKLLEQWNLVIGSGGITVEELKKLLEQWNLVI* (SEQ ID NO: 27564)

>lucCageSARS2-M4_8-24_348

GSHHHHHHGSSENLYFQG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEA
RRLIEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLR
ALAQLQELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREA
ERSRITVEELKKLLEQWNLVIGSGGITVEELKKLLEQWNLVI* (SEQ ID NO: 27565)

>lucCageM3_334_SmBit_位置301

GSHHHHHHGSSENLYFQGG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEALL
EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLSREAAAMADSNGTITVEE
LKKLLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27566)

>lucCageM3_334_SmBit_位置308

GSHHHHHHGSSENLYFQGG(SKEAAKQLQDLNIELARKLLEASTKLQRLNIRLAEALL
EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL

QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLSREAAAMADSNGTITVEELK
KLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27567)

>lucCageM3_334_7环

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDGGGSGGGPDEARKAIAVTGYRLFEEILDAERLSREAAAMADSNGTIT
VEELKKLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27568)

>lucCageM3_334_3环

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDGGGPDEARKAIAVTGYRLFEEILDAER
LSREAAAMADSNGTITVEELKKLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO:

27569)

>lucCageM3_341_SmBit_位置301

GSHHHHHHGSGSENLYFQSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALL
EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLSREAAAASEKISRMADSN
GTITVEELKKLEGGSGGMADSNGTITVEELKKLLE(SEQ ID NO: 27570)

>lucCageM3_341_SmBit_位置308

GSHHHHHHGSGSENLYFQSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALL
EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLSREAAAASEKISRMADSN
ITVEELKKLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27571)

>lucCageM3_341_7环

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDGGGSGGGPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRMA
DSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27572)

>lucCageM3_341_3环

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDGGGPDEARKAIAVTGYRLFEEILDAER

LSREAAAASEKISRMADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE (SEQ ID
NO: 27573)

>LUCCAGEM3_334_4拷贝

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAMADSNGTITVEE
LKKLLEGGSGGMADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLEGGSGGMA
DSNGTITVEELKKLLE* (SEQ ID NO: 27574)

>LUCCAGEM3_337_4拷贝

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEMADSNGTIT
VEELKKLLEGGSGGMADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLEGGSGG
MADSNGTITVEELKKLLE (SEQ ID NO: 27575)

>LUCCAGEM3_341_4拷贝

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISRMADSN
 GTITVEELKKLLEGGSGGMADSN GTITVEELKKLLEGGSGGMADSN GTITVEELKKLLEGG
 GSGGMADSN GTITVEELKKLLE (SEQ ID NO: 27576)

>LUCCAGEM3_348_4拷贝

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISREAERSIR
 MADSN GTITVEELKKLLEGGSGGMADSN GTITVEELKKLLEGGSGGMADSN GTITVEELK
 KLEGGSGGMADSN GTITVEELKKLLE (SEQ ID NO: 27577)

>LUCCAGEM3_334_2拷贝无接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAMADSN GTITVEE
 LKKLLEMADSN GTITVEELKKLLE (SEQ ID NO: 27578)

>LUCCAGEM3_337_2拷贝无接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEMADSN GTIT
 VEELKKLLEMADSN GTITVEELKKLLE (SEQ ID NO: 27579)

>LUCCAGEM3_341_2拷贝无接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISRMADSN
GTITVEELKKLLEMADSNGTITVEELKKLLE (SEQ ID NO: 27580)

>LUCCAGEM3_348_2拷贝无接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISREAERSIR
MADSNGTITVEELKKLLEMADSNGTITVEELKKLLE (SEQ ID NO: 27581)

>LUCCAGEM3_334_4拷贝_接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAMADSNGTITVEE
LKKLLEGGSGGMADSNGTITVEELKKLLEGGSGGGSGGSGGMADSNGTITVEELKK
LLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27582)

>LUCCAGEM3_337_4拷贝_接头

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEMADSNGTIT
VEELKKLLEGGSGGMADSNGTITVEELKKLLEGGSGGGSGGSGGMADSNGTITVEEL
KKLLEGGSGGMADSNGTITVEELKKLLE (SEQ ID NO: 27583)

>LUCCAGEM3_341_4拷贝_接头

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISRMADSN
 GTTITVEELKKLLEGGSGGMADSN GTTITVEELKKLLEGGSGGGSGGSGGSGGMADSN GTTIT
 VEELKKLLEGGSGGMADSN GTTITVEELKKLLE (SEQ ID NO: 27584)

>LUCCAGEM3_348_4拷贝_接头

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEKISREAERSIR
 MADSN GTTITVEELKKLLEGGSGGGSGGSGGSGGMADSN GTTITVEELKKLLEGGSGGMAD
 SN GTTITVEELKKLLEGGSGGMADSN GTTITVEELKKLLE (SEQ ID NO: 27585)

>LUCCAGEM3_334_2拷贝_接头_SpaC_Z

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAMADSN GTTITVEE
 LKKLLEGGSGGMADSN GTTITVEELKKLLEGGGGSGGGSGGSGGSGGSGGNKFNKEQQNA
 FYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKL NDAQAPKGGVDNKNFNKEQQNA
 FYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKL NDAQAPK (SEQ ID NO: 27586)

>LUCCAGEM3_337_2拷贝_接头_SpaC_Z

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERDAERLSREAAAASEMADSN GTTIT
 VEELKKLLEGGSGGMADSN GTTITVEELKKLLEGGGGSGGGSGGSGGSGGSGGNKFNKEQ
 QNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKL NDAQAPKGGVDNKNFNKEQ
 QNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKL NDAQAPK (SEQ ID NO:
 27587)

>LUCCAGEM3_341_2拷贝_接头_SpaC_Z

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISRMADSN
GTITVEELKKLLEGGSGGMADSNGTITVEELKKLLEGGGGSGGGSGGSGGSGGSGGNKFN
KEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLND AQAPKGGVDNKNFN
KEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLND AQAPK (SEQ ID NO:
27588)

>LUCCAGEM3_348_2拷贝_接头_SpaC_Z

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIR
MADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLEGGGGSGGGSGGSGGSGGS
GGNKNFNKEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLND AQAPKGG
VDNKNFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLND AQAPK
(SEQ ID NO: 27589)

lucCageRBD变体(SARS-CoV2刺突蛋白受体结合结构域(RBD)生物传感器)

-LCB1:

**DKEWILQKIYEIMRLLDDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLE
EVER (SEQ ID NO: 27397)**

-LCB1_Δ4:

**ILQKIYEIMRLLDDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEE
EVER (SEQ ID NO: 27590)**

>lucCageRBD_336

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASDKEWILQKIYEI
**MRLLDDELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEE
EVER* (SEQ ID NO:
27591)**

>lucCageRBD_340

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISDKEWILQKI
YEIMRLLELGHAEASMRVSDLIYEFMKKGDERLLEEAERLLEEVER* (SEQ ID NO:
27592)

>lucCageRBD_344

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAEDKEWI
LQKIYEIMRLLELGHAEASMRVSDLIYEFMKKGDERLLEEAERLLEEVER* (SEQ ID
NO: 27593)

>lucCageRBD_347

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIDKE
WILQKIYEIMRLLELGHAEASMRVSDLIYEFMKKGDERLLEEAERLLEEVER*
(SEQ ID NO: 27594)

>lucCageRBD_351

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
ADKEWILQKIYEIMRLLELGHAEASMRVSDLIYEFMKKGDERLLEEAERLLEEVER
***** (SEQ ID NO: 27595)

>lucCageRBD_354

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
AAASDKEWILQKIYEIMRLLELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEE
VER* (SEQ ID NO: 27596)

>lucCageRBD_GGG_360

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
AAASEKISREGGDKIEWILQKIYEIMRLLELGHAEASMRVSDLIYEFMKKGDERLL
EEAERLLEEVEVER* (SEQ ID NO: 27597)

>lucCageRBD Δ 4_336

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASILQKIYEIMRLLE
ELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEVER* (SEQ ID NO: 27598)

>lucCageRBD Δ 4_340

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISILQKIYEIM
RLLELGHAEASMRVSDLIYEFMKKGDERLLEEAEERLLEEVEVER* (SEQ ID NO: 27599)

>lucCageRBD Δ 4_344

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAEILQKIY
EIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEVER* (SEQ ID NO:
27600)

>lucCageRBD Δ 4_347

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAEERSIILQ
KIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEVER* (SEQ ID
NO: 27601)

>lucCageRBD Δ 4_348

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAEERSIRIL
QKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEVER* (SEQ ID
NO: 27602)

>lucCageRBD Δ 4_351

MGSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAIAETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAEERSIREA
AILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAERLLEEVEVER* (SEQ
ID NO: 27603)

>lucCageRBD Δ 4_354

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
AAASILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAEERLLEEVE*
(SEQ ID NO: 27604)

>lucCageRBD Δ 4_357

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
AAASEKIILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAEERLLEEVE
R* (SEQ ID NO: 27605)

>lucCageRBD Δ 4_GGG_360

MGSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEAR
KAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAA
AKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLI
EKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQ
ELNLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIREA
AAASEKISREGGILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKGDERLLEEAEER
LLEEVE* (SEQ ID NO: 27606)

>lucCageRBD_348_d4LCB1v1.3

GSHHHHHHSGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIRILQKI
YEIMKTLEQLGHAEASMQVSDLIYEFMKGQDERLLEEAEERLLEEVE* (SEQ ID NO:
27607)

>lucCageRBD_ Δ 4_348

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLLELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27608)

>lucCageRBD_smbit128

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEKILDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLLELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27609)

>lucCageRBD_smbit99

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLFEEKsDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLLELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27610)

>lucCageRBD_smbit86

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVsGwRLFkIsDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLLELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27611)

>lucCageRBD_smbit104

GSHHHHHHGGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVeGYRLFekIsDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLLELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27612)

>lucCageRBD_smbit101

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLF_{kes}DAERLSREAAAASEKISREAERSIRILQKI
YEIMRLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27613)

>lucCageRBD_smbit_Y315W_E320K

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTG_wRLF_{kes}ILDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27614)

>lucCageRBD_smbit_Y315W_E319K

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTG_wRLF_{kes}ILDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27615)

>lucCageRBD_smbit_E319K

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKA
IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
LQELNIRAVELLVKLTPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAIAVTGYRLF_{kes}ILDAERLSREAAAASEKISREAERSIRILQKI
YEIMRLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO: 27616)

>lucCageRBD_SmBit_位置301

GSHHHHHHGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALL
 EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
 RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
 IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
 QELNLDLLRLASEL)ggsVTGYRLFEEILRVKRESKRIVEDAERLsREAAAASEKISREAERSI
 RILQKIYEIMRLLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO:
 27617)

>lucCageRBD_SmBit_位置308

GSHHHHHHGSGSENLYFQGSGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALL
 EAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
 RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRL
 IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
 QELNLDLLRLASEL)TDPDEARVTGYRLFEEILRIVEDAERLsREAAAASEKISREAERSIRIL
 QKIYEIMRLLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID NO:
 27618)

>lucCageRBD_环

GSHHHHHHGSGSENLYFQG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAI
 ARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKA
 IRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAK
 LQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARRLIEK
 AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
 NLDLLRLASEL)TDGGSGGPDEARKAIAVTGYRLFEEILDAERLSREAAAASEKISREAERS
 IRILQKIYEIMRLLDELGHAEASMRVSDLIYEFMCKGDERLLEEAERLLEEVEER (SEQ ID
 NO: 27619)

LacATrop (拆分 β -内酰胺酶A加粗;cTnT和cTnC加下划线):

MGSHHHHHHSGSGENLYFOG (SGGSVFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLCCGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
HNMGDHSVTRL DRWEPELNEA IPNDERDTTT PAAMATTLRK LLTGNGR
SGGGSGGGSGGGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIARLQ
ELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIRRAEKEIDDAAKESKKILEEARKAIRDA
AEESRKILEEGSGSGSDALDELQKLNLELAKLLKAI AETQDLNLRAAKAFLEAAKLQEL
NIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARRLIEKAKEE
SERIIEGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQELNLDL
LRLASEL)TDPDEARKAIAVTGYRLFEEILDAERLIREAAAASEDQLREAAKELWQTIYNLE
AEKFDLQEKFKQOKYEINVLNRINDNQVSKTKDDSKGKSEEELSDLFRMFDKNADGYI
DLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE (SEQ ID NO:
 27620)

[0104] 在另一方面,本公开提供了能够结合到本文公开的任何实施方案或实施方案组合的笼蛋白结构区的钥匙蛋白,其中所述钥匙蛋白与所述笼蛋白的结合仅在所述笼蛋白一种或多种靶标结合多肽能够结合的靶标的存在下发生,其中所述钥匙蛋白第二报告蛋白结构域和所述笼蛋白第一报告蛋白结构域的相互作用引起来自所述第一报告蛋白结构域的报告活性的可检测变化。

[0105] 如本文所公开的,该方面的钥匙蛋白可以例如与笼状多肽结合使用,通过诱导构象变化的竞争性分子内结合替换锁,使钥匙蛋白第二报告蛋白结构域和笼蛋白第一报告蛋白结构域相互作用,引起第一报告蛋白结构域的报告活性发生可检测的变化。

[0106] 在一个实施方案中,其中第二报告蛋白结构域位于钥匙蛋白的N端或C端,或位于钥匙蛋白N端或C-端的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或1个氨基酸内。

[0107] 在另一个实施方案中,第二报告蛋白结构域包括报告蛋白结构域,所述报告蛋白结构域选自荧光素酶(包括但不限于萤火虫荧光素酶、海肾荧光素酶和高斯荧光素酶)、生物发光共振能量转移(BRET)报告分子、双分子荧光互补(BiFC)报告分子、比色报告分子(包括但不限于β-内酰胺酶、β-半乳糖苷酶和辣根过氧化物酶)、细胞存活报告分子(包括但不限于二氢叶酸还原酶)、电化学报告分子(包括但不限于APEX2)、放射性报告分子(包括但不限于胸苷激酶)、和分子条形码报告分子(包括但不限于TEV蛋白酶)。在各种非限制性实施方案中,第二报告蛋白结构域包含与选自SEQ ID NO:27360-23379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中下划线残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且其中任何N端蛋氨酸残基可以存在或不存在。

[0108] 在另一个实施方案中,不包括第二报告蛋白结构域的钥匙蛋白包含与US20200239524(或W02020/018935)中公开的钥匙多肽的氨基酸序列、或选自SEQ ID NO:14318-26601、26602-27015、27016-27050、27322至27358的钥匙多肽,以及W02020/018935的表3(本文表8)和/或表4(本文表9)中SEQ ID NO在SEQ ID NO:27127和27277之间的奇数钥匙多肽至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、

93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,不包括任选的氨基酸残基,。

[0109] 在另一个实施方案中,钥匙蛋白包含与选自SEQ ID NO:27621-27623的钥匙蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,其中括号中的残基是任选的并且可以存在或不存在。

>lucKey:MGS-(His)6-TEV位点-接头-LgBit-接头-闪烁序列

(MGSHHHHHHGSGSENLYFQG)SGMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQN
LAVSVTPIQIRIVRSGENALKIDIHIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTL
VIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSGG
SGGGSGGGSGGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAA
ASEKISRE (SEQ ID NO:27621)

Key-2GGSGG-Cy0FP (Cy0FP序列加粗/加下划线) :

(M)DPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISREGG
SGG GGVSK GEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV
EGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPE GFTWERVMVF
EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK TLGWEPSTET
MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGGMD ELYK (SEQ ID NO: 27622)

Key-LacB (拆分 β -内酰胺酶B加粗/加下划线) :

SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRES
GGGGSGGGSGGGG LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA
DKSGAGERGS RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNROIAEIGA SLIKHW
(SEQ ID NO: 27623)

[0110] 在另一个方面,本公开提供了一种生物传感器,其包括(a)本文任何实施方案或实施方案组合的笼蛋白,其中所述笼不包括第二报告蛋白结构域;和(b)本文实施方案或实施方案组合的钥匙蛋白;其中所述钥匙蛋白只能在所述笼蛋白一种或多种靶标结合多肽能够结合的靶标存在下与所述笼蛋白结合;并且其中所述笼蛋白的第一报告蛋白结构域与所述钥匙蛋白的第二报告蛋白结构域的结合引起来自所述第一报告蛋白结构域的报告活性的可检测变化。

[0111] 如本文所述,发明人开发了一种反向LOCKR系统,以及包含与钥匙肽连接的第二报告蛋白结构域的钥匙蛋白。该系统至少有三种重要状态(图1C)。状态1是封闭的关闭状态,其中结构区与闪烁区相互作用,在空间上阻止一种或多种靶标结合多肽与其靶标结合,阻止第一报告蛋白结构域与第二报告蛋白结构域结合以重建报告蛋白活性。状态2或3是开放状态,其中这些结合相互作用不受阻,并且钥匙蛋白可以结合笼蛋白结构域。当钥匙蛋白与笼蛋白结构域以及带有多肽靶标的一种或多种靶标多肽的三分子结合导致报告蛋白活性的重建时,状态7是稳定的打开状态。将笼蛋白与钥匙蛋白或靶蛋白单独混合不足以激活报告活性。钥匙蛋白和靶蛋白与笼蛋白一起在同一溶液中导致报告蛋白活性的重组。闪烁区-靶标强相互作用提供驱动力,以在状态6(背景)上增加(populate)打开状态7(信号)。下面

的示例提供了更多详细信息。

[0112] 如上所述,可检测的变化可以是相关报告活性的任何增加或减少,视为适合预期目的。在各种非限制性实施方案中,报告活性的可检测变化可以包括但不限于:

- 第一报告蛋白结构域是一种拆分的荧光或发光蛋白结构域,当结合到第二拆分的报告蛋白结构域时,该结构域不发出荧光/光,或可检测地发出更少的荧光/光。

- 第一和第二报告蛋白结构域是BRET或FRET对,当彼此结合时与当彼此不结合时,它们在不同波长下发出可检测信号。

- 当第一和第二报告蛋白结构域重建DHFR活性时,在选择靶标存在下,通过二氢叶酸还原酶(DHFR)互补进行细胞存活选择。

- 当第一和第二报告蛋白结构域重组TEV蛋白酶活性以用作分子条形码时,二代测序作为描述靶选择性途径上的化学或遗传干扰的读数。

- 当第一和第二报告蛋白结构域重组胸苷激酶时进行的正电子发射断层扫描(PET)。

- 当第一和第二报告蛋白结构域重建APEX2活性时的电化学读数。

- 当第一和第二报告蛋白结构域重建 β -内酰胺酶或辣根过氧化物酶活性时的比色读数。

[0113] 在本公开的生物传感器的各种实施方案中:

[0114] (a) 第一报告蛋白结构域包含与氨基酸序列SEQ ID NO:27359和27664-27672至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列;

[0115] 并且所述第二报告蛋白结构域包含与SEQ ID NO 27379的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在;

[0116] (b) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27360的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,另一个包含与SEQ ID NO:27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%,或100%相同的氨基酸序列;

[0117] (c) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27362的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,另一个包含与选自SEQ ID NO:27363-27365的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%,或100%相同的氨基酸序列;

[0118] (d) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27366的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,另一个包含与SEQ ID NO 27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%,或100%相同的氨基酸序列;

[0119] (e) 所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID

NO:27367的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在,另一个包含与SEQ ID NO 27368的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中N端蛋氨酸残基可以存在或不存在;

[0120] (f)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27369的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;另一个包含与SEQ ID NO:27370的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

[0121] (g)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27371的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27372的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

[0122] (h)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27373的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27374的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

[0123] (i)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27375的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,另一个包含与SEQ ID NO:27376的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列;

[0124] (j)所述第一报告蛋白结构域和所述第二报告蛋白结构域中的一个包含与SEQ ID NO:27377的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中所述N端蛋氨酸残基可以存在或不存在,并且其中带下划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列,并且另一个包含与SEQ ID NO:27378的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中带下

划线的残基是可以存在或不存在的任选残基,并且当存在时可以是任何氨基酸序列。

[0125] 在生物传感器的一个具体实施方案中,笼蛋白包括一种笼蛋白,该种笼蛋白包含与表10中列出的笼蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,不包括任选的氨基酸残基,其中所述N端蛋白纯化标签(MGSHHHHHHGSGSENLYFQSGG (SEQ ID NO:27624);或MGSHHHHHGSENLYFQG (SEQ ID NO:27625);或GSHHHHHHGSGSENLYFQG (SEQ ID NO:27626))是任选的,并且可以存在或不存在,并且钥匙蛋白包含与SEQ ID NO:27621的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,不包括括号中的任选氨基酸残基。

>lucKey:MGS- (His)6-TEV位点-接头-LgBit-接头-闪烁序列

(MGSHHHHHHGSGSENLYFQG)SGMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQN
LAVSVTPIQRIVRSGENALKIDIHIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTL
VIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSGG
SGGGGSGGGSGGSDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAA
ASEKISRE (SEQ ID NO: 27621)

[0126] 在生物传感器的另一个具体实施方案中,笼蛋白和钥匙蛋白包括蛋白对,该蛋白对包括:

[0127] (i) 笼蛋白,其包含与SEQ ID NO:27620的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列,其中括号中的残基是任选的并且可以存在或不存在:

LacATrop (拆分 β -内酰胺酶A加粗;cTnT和cTnC加下划线):

(MGSHHHHHHGSGSENLYFQG SGG)**VFAHPETLVK VKDAEDQLGA RVGYIELDLN**
SGKILESFRP EERFPMMSTF KVLLCGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
HNMGDHSVTRL DRWEPELNEA IPNDERDTTT PAAMATTLRK LLTGENGR
SGGGGSGGGGSGGGGSKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIARLQ
ELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARKAIRDAA
EESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLR AAKAFLEAAAKLQELN
IRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRA AKRESERIIEEARLIEKAKEESE
RIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQELNLDLLR
LASELTDPEARKIAVGTGYRLFEEILDAERLIREAAAASEDQLREAAKELWQTIYNLEAE
KFDLQEKFKQOKYEINVLRNRINDNQVSKTKDDSKGKSEEELSDLFRMFDKNADGYIDL
EELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE (SEQ ID NO:27620);
和

[0128] (ii) 钥匙蛋白,其包含与SEQ ID NO:27361的氨基酸序列至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或100%相同的氨基酸序列:

LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW (SEQ ID NO:27361)

[0129] 在另一方面,本公开提供了用于检测靶标的方法,包括

[0130] (a) 使本文公开的任何实施方案的笼蛋白(其中所述笼蛋白包含第二报告蛋白结构域),或本文公开的任何实施方案的生物传感器与生物样品在促进笼蛋白一种或多种靶标结合多肽与生物样品中存在的靶标结合的条件下接触,引起来自第一报告蛋白结构域的报告活性的可检测变化;和

[0131] (b) 检测来自报告蛋白结构域的报告活性的变化,其中报告活性的改变将样品识别为含有靶标。

[0132] 如上所述,本发明人已经开发了一种反向LOCKR系统,该系统以笼蛋白为例,该笼蛋白包括结构区和含有第一报告蛋白结构域和一种或多种靶标结合多肽(在实施例中有时称为分析物结合基序/靶表位)的锁区,以及包含与钥匙肽连接的第二报告蛋白结构域的钥匙蛋白。如上所述,可检测的变化可以是相关报告活性的任何增加或减少,视为适合预期目的。上文描述了报告活性中可检测的变化的各种非限制性实施方案,并且在下面的实施例中详细举例说明了用于检测这种可检测的改变的方法。基于本文的教导,本领域技术人员可以确定用于测定可检测的目标变化的适当技术。

[0133] 如图19所示和实施例3所述,这些方法可以适应“间接检测”方法,其中报告蛋白(分子间(笼蛋白中的第二报告结构域)或分子内(钥匙上的第二报告蛋白)实施方案通过将生物传感器与靶标结合多肽的靶标预孵育而重组,导致报告活性的恢复。然后将激活的生物传感器与样品一起孵育,以检测一种或多种靶标结合多肽结合的靶标的存在,使靶标与一种或多种靶标结合多肽结合,报告蛋白组分之间的相互作用丧失,报告活性下降/消失。

[0134] 可以使用任何合适的生物样品,包括但不限于血液、血清、唾液、尿液、精液、阴道液、淋巴液、组织液、消化液、汗液、眼泪、鼻涕、羊水和母乳。

[0135] 任何靶标都可以被检测为适合预期用途,并且一种或多种靶标结合多肽可用于包含在笼蛋白中。在非限制性实施方案中,所述靶标选自包括但不限于抗体、毒素、诊断生物标志物、病毒颗粒或疾病生物标志物。在一个具体实施方案中,所述靶标是抗体。在另一个实施方案中,所述靶标包括对病毒具有选择性的抗体。在各种这样的实施方案中,一种或多种靶标结合多肽可包含选自SEQ ID NO:27292-27394和27547-27548的氨基酸序列,以及包含与选自SEQ ID NO:27397-27494的氨基酸序列至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列的多肽。在这些实施方案中,所述方法可用于检测针对SARS冠状病毒(包括SARS-CoV或SARS-CoV-2)的抗体的存在。

[0136] 在各种进一步的实施方案中,笼多肽包含与表10中列出的笼蛋白的氨基酸序列至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同(不包括任选氨基酸残基)的氨基酸序列。

[0137] 在另一个实施方案中,所述靶标是疾病标志物或毒素。在一个这样的实施方案中,疾病标志物或毒素包括Bcl-2、Her2受体、肉毒杆菌神经毒素B、白蛋白、上皮生长因子受体、前列腺特异性膜抗原(PSMA)、瓜氨酸肽、脑钠肽和/或心肌钙蛋白I。在另一个实施方案中,所述一种或多种靶标结合多肽包含与选自SEQ ID NO:27380-27390的氨基酸序列65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的氨基酸序列,其中任何N端氨基酸是任选的并且可以存在或不存在。

[0138] 在各种进一步的实施方案中,笼多肽包含与表10中列出的笼蛋白的氨基酸序列至

少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同(不包括任选氨基酸残基)的氨基酸序列。

[0139] 本公开还提供了用于设计/制造生物传感器、笼蛋白或钥匙蛋白的方法,包括本文所述的任何方法的步骤,例如在以下实施例中的步骤。

[0140] 在另一方面,本公开提供了编码本公开的笼蛋白、钥匙蛋白或表位的核酸。所述核酸序列可以包括RNA(例如mRNA)或DNA。这样的核酸序列可以包括用于促进编码蛋白的表达和/或纯化的额外序列,包括但不限于polyA序列、修饰的Kozak序列和编码表位标签、输出信号和分泌信号、核定位信号和质膜定位信号的序列。基于本文的教导,对于本领域技术人员来说,什么核酸序列将编码本发明的蛋白是显而易见的。

[0141] 在另一个方面,本公开提供了表达载体,其包含与合适的控制序列可操作地连接的本公开的任何实施方案或实施方案的组的核酸。“表达载体”包括可操作地将核酸编码区或基因连接到能够影响基因产物表达的任何控制序列的载体。可操作地连接到本公开的核酸序列的“控制序列”是能够影响核酸分子表达的核酸序列。控制序列不需要与核酸序列邻接,只要它们具有指导其表达的功能即可。因此,例如,在启动子序列和核酸序列之间可以存在插入的未翻译但转录的序列,并且启动子序列仍然可以被认为与编码序列“可操作地连接”。其他此类控制序列包括但不限于聚腺苷酸化信号、终止信号和核糖体结合位点。这种表达载体可以是本领域已知的任何类型,包括但不限于基于质粒和病毒的表达载体。用于驱动所公开的核酸序列在哺乳动物系统中表达的控制序列可以是组成性的(由多种启动子中的任何一种驱动,包括但不限于CMV、SV40、RSV、肌动蛋白、EF)或诱导性的(通过多种诱导性启动子中任何一种,包括但不限于四环素、蜕皮素、类固醇响应启动子)。

[0142] 在一个方面,本公开提供了包含本公开的任何实施方案或实施方案的组的笼蛋白、钥匙蛋白、表位、生物传感器、核酸和/或表达载体的细胞,其中所述细胞可以是原核的或真核的,例如哺乳动物细胞。在一个实施方案中,可以用本公开的核酸或表达载体瞬时或稳定地转染细胞。可通过本领域已知的任何技术将表达载体转染到原核和真核细胞中。根据本发明生产多肽的方法是本发明的附加部分。该方法包括以下步骤:(a)在有利于多肽表达的条件下培养根据本发明这一方面的宿主,和(b)任选地回收所表达的多肽。

[0143] 在另一方面,本公开提供了药物组合物,其包括

[0144] (a) 本文任何实施方案或实施方案组的笼蛋白、钥匙蛋白、生物传感器、表位、重组核酸、表达载体和/或细胞;和

[0145] (b) 药学上可接受的载体。

[0146] 组合物可进一步包括(a)冻存稳定剂;(b)表面活性剂;(c)填充剂;(d)张力调节剂;(e)稳定剂;(f)防腐剂和/或(g)缓冲液。在一些实施方案中,药物组合物中的缓冲液是Tris缓冲液、组氨酸缓冲液、磷酸盐缓冲液、柠檬酸盐缓冲液或乙酸盐缓冲液。该组合物还可以包括冻干保护剂,例如蔗糖、山梨醇或海藻糖。在某些实施方案中,组合物包括防腐剂,例如苯扎氯铵、苯乙氧基铵、氯己定、苯酚、间甲酚、苜醇、对羟基苯甲酸甲酯、对羟基苯丙酯、氯丁醇、邻甲酚、对甲酚、氯甲酚、硝酸苯汞、硫柳汞、苯甲酸及其各种混合物。在其他实施方案中,组合物包括填充剂,如甘氨酸。在另一些实施方案中,该组合物包括表面活性剂,例如聚山梨醇酯-20、聚山梨醇酯-40、聚山梨醇酯-60、聚山梨醇酯-65、聚山梨醇酯-80、聚山梨醇酯-85、聚氧乙烯醚-188、山梨醇单月桂酸酯、山梨醇单硬脂酸酯、山梨醇单油酸酯、

山梨醇三硬脂酸酯、山梨醇三油醇酯或其组合。该组合物还可以包括张力调节剂,例如使制剂与人体血液基本等渗或等渗的化合物。示例性的张力调节剂包括蔗糖、山梨醇、甘氨酸、蛋氨酸、甘露醇、葡萄糖、肌醇、氯化钠、精氨酸和精氨酸盐酸盐。在其他实施方案中,组合物还包括稳定剂,例如,基本上防止或减少冻干或液体形式的纳米结构的化学和/或物理不稳定性的分子。示例性稳定剂包括蔗糖、山梨醇、甘氨酸、肌醇、氯化钠、蛋氨酸、精氨酸和精氨酸盐酸盐。

[0147] 在另一方面,本公开提供了一种表位,其包含SEQ ID NO:27384的氨基酸序列,或由其组成。

lucCageTrop	cTnI + cTnC	EDQLREKAKELWQTIYNLEAEKFDLQ
		EKFKQQKYEINVLNRNRINDNQKVSKT
		KDDSKGKSEEELSDLFRMFDKNADGY
		IDLEELKIMLQATGETITEDDIEELMKD
		GDKNNDGRIDYDEFLEFMKGVE (SEQ
		ID NO:27384)

[0148] 该表位可以例如用于本公开的生物传感器中。在一个方面,本公开提供了检测样品中肌钙蛋白I的方法,包括在适合于促进样品中的肌钙蛋白I与表位结合的条件下使生物样品与表位接触以形成结合复合物,以及检测证明样品中存在肌钙蛋白II的结合复合物。本文公开的生物样品和检测的所有实施例也可用于这些方法中。

实施例

[0149] 这里,我们表明,通过从头设计的蛋白开关将信息流反转,可以创建一类非常通用的基于变构蛋白的生物传感器,其中肽键的结合触发目标生物输出。使用广泛适用的设计原则,我们通过设计的锁蛋白和钥匙蛋白的结合,将目标蛋白分析物的结合与荧光素酶活性的重组和生物发光读出进行变构耦合。因为传感器纯粹基于分析物结合到开关激活的热力学耦合,所以只需要一种靶标结合结构域,这简化了传感器设计并允许在溶液中直接读出。我们通过创建生物传感器(该生物传感器在几乎没有优化的情况下敏感地检测抗凋亡蛋白Bcl-2、hIgG1-Fc结构域、Her2受体和肉毒神经毒素B),以及用于心肌钙蛋白I和抗乙型肝炎病毒(HBV)抗体的生物传感器(其达到检测这些分子的临床相关浓度所需的亚毫摩尔灵敏度),证明了该平台的模块性。我们还使用该方法设计针对SARS-CoV-2蛋白表位的抗体和SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的传感器。后者结合了从头设计的RBD结合域,其检测限为15pM,添加RBD后发光增加了17倍。该平台的模块化和灵敏度使得能够快速构建用于各种分析物的传感器,并突出了从头蛋白设计的力量,以创建具有新的有用功能的多状态蛋白系统。

[0150] 蛋白生物传感器可以由具有两种几乎能量相等的状态的系统构成,这两种状态之间的平衡由被检测的分析物调节。这种传感器中的期望特性是(i)分析物触发的构象变化应该独立于分析物的细节(因此相同的整体系统可以用于传感许多不同的化合物)(ii)系统应该是可调的,和(iii)构象变化应与敏感的输出耦合。我们假设,这些属性可以通过反转从头设计的蛋白开关中的信息流来实现,其中与靶标蛋白的结合由肽促动器的存在控制。这些开关由一个恒定的“笼”区组成,该区封锁了(sequester)一个与目标靶标结合的

“开锁”；添加一个肽“钥匙”将锁从笼中移开，导致靶标结合和相关的下游事件。然而，从热力学角度来看，钥匙和靶标是等价的：钥匙和靶标与笼的结合是热力学耦合的，因为开锁必须打开，需要自由能成本 ΔG_{open} (图1b)，以便两者结合。因此，结合靶标和钥匙的自由能比结合两者的自由能之和更有利(图1c)。钥匙和靶标的区别在于它们的可变性；钥匙是恒定的，而靶标可以是任何期望的相互作用。对于促动器，需要有恒定的输入驱动广泛的可定制响应，因此在我们之前的工作中，输入是(恒定的)钥匙，输出与蛋白降解、核输出等相关的各种靶标结合。我们推断，系统的输入可以反转，以创建具有恒定读数的生物传感器-添加(可变)靶标可以诱导(恒定)钥匙与(恒定)笼的结合，并且这种结合可以与酶读数耦合。这样的系统将满足上述性质(i)和(ii)，因为可以限制广泛的结合活性，并且由于开关是热力学控制的，因此可以直接调节钥匙和靶标结合的相对能量以实现在相关靶标浓度下的激活。因为钥匙和笼总是相同的，所以系统是模块化的：相同的分子连接可以与许多不同靶标的结合耦合。

[0151] 为了实现特性(iii)，我们推断生物发光可以提供分析物驱动的笼钥匙连接的快速灵敏读出，并探索了可逆拆分荧光素酶互补系统的使用。我们开发了一个由两个蛋白组分组成的系统：一个“lucCage”，包括一个笼结构域和一个含有短拆分荧光素酶片段(SmBiT)的开锁结构域，以及一个选择的分析物结合基序；以及“lucKey”，其包含较大的拆分荧光素酶片段(LgBit)和钥匙肽(图1a)。lucCage有两种状态：封闭状态，其中cage结构域结合开锁，并在空间上阻止分析物结合基序与其靶标结合，并阻止SmBiT与LgBit结合以重建荧光素酶活性；以及开放状态，其中这些结合相互作用不受阻，lucKey可以结合笼结构域。lucKey与lucCage的结合导致荧光素酶活性的重建(图1a, 右)。靶标可以被视为变构调节荧光素酶活性，因为与传感器的结合位于远离酶活性位点的位置。

[0152] 这种系统的状态处于热力学平衡，可调参数 ΔG_{open} 和 ΔG_{CK} 控制可能类别的群体，以及分析物与结合域 ΔG_{LT} 的结合自由能(图1b)。为了实现高灵敏度，在没有靶标(类别6)的情况下，封闭状态(类别1)的自由能必须比开放状态低得多，以避免背景信号 $\Delta G_{1-6} > 0$ ，但在有靶标(类别7, $\Delta G_{1-7} < 0$)的情况下，封闭状态的自由能要比开放状态高得多，因此靶标检测在能量上是有利的(图1c)。为了指导生物传感器灵敏度的优化，我们模拟了传感器系统对 ΔG_{open} (图1d)、 ΔG_{LT} (图1e)、分析物浓度和传感器组分的依赖性(图1f) (详见补充方法)。正如预期的那样，分析物检测的灵敏度是 ΔG_{LT} 的函数，分析物结合的下限约为Kd的十分之一(图1e；低于该浓度，结合的自由能太小，无法打开开关)。因此，对其靶标具有高亲和力的传感结构域将产生更灵敏的生物传感器。通过改变lucCage和lucKey的浓度，可以将系统的灵敏度进一步调整到该下限以上，从而使传感系统对不同的靶标浓度范围做出响应(图1f)。调节分子内笼-开锁相互作用的强度(ΔG_{open})会影响催化活性类别(类别6和7, 图1d)的平衡群体，进而影响灵敏度：过强的相互作用会导致靶存在时的低信号，而过弱的相互作用则会导致靶不存在时的高背景。我们的设计策略旨在通过设计具有 ΔG_{open} 值范围的闭合状态(类别1)传感器来找到这种平衡：通过增加(减少)开锁螺旋的长度，并通过在笼-开锁界面引入有利的疏水相互作用或不利的空间碰撞和包埋的极性原子，可以增加(减小) ΔG_{open} ；我们使用这两种策略来调整下面描述的传感器(ΔG_{CK} 也可以调整，但我们发现这对于这里描述的传感器没有必要)。

[0153] 为了简化基于这些原理的新传感器的设计，我们开发了一种基于Rosetta™的计算

方法,用于将不同的传感结构域整合到名为GraftSwitchMover的LOCKR开关中。该方法确定了将靶标结合肽包埋在锁内的最合适位置,使得所得蛋白在封闭状态下稳定。这是通过最大化肽和笼之间有利的疏水嵌合作用(packing interaction)和最小化不利的包埋亲水性残基的数量来实现的。该方法将开关的三维模型、结合目标靶标的肽序列以及该肽中与靶标相互作用的残基列表(界面残基)作为输入,并返回一组设计,其中预测肽与靶标的结合将通过与笼的结合而被阻断(见补充方法)。最后一组设计涵盖了一系列 ΔG_{open} 值(图1c),可以通过在锁中引入失稳突变来进一步调整:I328S(“1S”)或I328S/L345S(“2S”)。然后对这些设计物进行实验表征,以找到最灵敏的生物传感器。

[0154] 我们首先开始通过在Langan等人,2020²(图6)中描述的优化不对称LOCKR开关的闭合状态下接枝SmBiT肽和Bim肽来测试我们的假设。SmBiT在荧光素酶全酶中自然采用 β 链构象,但我们假设它将在螺旋束骨架的背景下采用螺旋二级结构,这与一些肽序列可以以背景依赖的方式采用不同二级结构的观察结果一致。我们对锁中两个肽序列的不同螺线进行采样,建立三维模型,选择最低能量的溶液(SmBiT的3个位置,Bim肽的4个位置)(图6a),并在大肠杆菌中表达12种设计物。我们将设计物与lucKey以1:1的比例混合,然后添加与Bim具有纳米摩尔亲和力的Bcl-2,并监测荧光素酶活性(图6b)。我们发现,在含有新笼的设计物、lucKey和furimazine底物的溶液中加入Bcl-2后,发光迅速增加(图6f),表明反向LOCKR系统确实可以作为生物传感器发挥作用。最佳Bcl-2传感器候选lucCageBim的进一步表征表明,可以通过改变传感器的浓度(lucCage+lucKey)来调整分析物检测范围(图6g),如我们的模型模拟中所预期的(图1f)。不同设计的实验表征表明,将SmBiT插入LOCKR笼(SmBiT312)的位置312产生了最高的稳定性和亮度(图6b),因此我们使用该设计(此后称为“lucCage”)作为下述生物传感器的基础骨架。

[0155] 为了探索我们新的生物传感器平台的多功能性,我们接下来研究了lucCage中目标分析物的一系列结合模式的结合。首先,我们开始探索如何计算封闭状态下的靶标结合蛋白,而不是肽。我们确定了结合蛋白与其靶标的主要相互作用表面,提取了其中涉及的主要二级结构元素,以将其用于上述计算方案,并从生成的许多螺线中选择了最佳设计。然后,我们使用RosettaTMRemodel在所述开关和选定设计物的背景中对全长结合结构域进行建模,所述设计物中该界面以最小的空间冲突埋入笼中(参见补充方法)。作为测试案例,我们将从头设计的蛋白HB1.9549.2(其结合甲型流感H1血凝素(HA)¹⁵)笼化进缩短版的LOCKR开关(sCage)中,进行优化以提高稳定性并促进结晶工作(图2a)。五种设计物中有两种是功能性的,在存在而不是不存在钥匙的情况下结合HA(图7b)。最佳设计sCageHA_267-1S的晶体结构被确定为**2.0Å**的分辨率(表11),表明除一种残基(F273)外,所有HA结合残基都与笼结构域相互作用(阻断锁与开关的结合),如设计物所示(图2a,图7a-c)。通过对设计概念的结构验证,我们接下来试图开发新的传感器,使用小蛋白作为检测肉毒杆菌神经毒素、免疫球蛋白Fc结构域和Her2受体的传感结构域。为此,我们将肉毒杆菌神经毒素B(BoNT/B)¹⁵的从头设计结合域、通用抗体结合蛋白A¹⁶的C结构域和Her2结合亲和抗体¹⁷接枝到lucCage中。在为每种靶标筛选了一些设计物后(图8-10),我们获得了高度敏感的lucCage(lucCageBot、lucCagerProA和lucCageHer2),它们分别可以检测BoNT/B(图2b、图8)、hIgG Fc结构域(图2c、图9)和Her2受体(图2d;图10),证明了平台的模块性。所设计的传感器在添加靶标后几分钟内做出响应,并且可以通过改变lucCage和lucKey的浓度来调整其灵敏度(图2),正

如我们的模型模拟所预测的那样(图1f)。这些传感器可用于多种应用,例如食品工业中高毒性肉毒杆菌神经毒素的快速和低成本检测,该检测目前严重依赖于活体动物生物测定,或检测与转移性乳腺癌相关的高血清水平的可溶性Her2($> 15\text{ng/mL}$),其水平可用lucCageHer2的当前灵敏度检测。

[0156] 接下来,我们为临床环境中相关的其他靶标设计了传感器。由于生物发光传感器不需要光来激发,因此与荧光相比,高灵敏度和低背景读数更适合直接测定生物介质(如血液和血清)中的分析物,用于即时护理应用。我们首先瞄准了心肌钙蛋白I(cTnI),它是急性心肌梗死(AMI)的标准早期诊断生物标志物。我们利用cTnT、cTnC和cTnI之间的高亲和力相互作用(图3a),并通过在不同的锁位置插入6种截短的cTnT序列来设计11种生物传感器候选(图11a)。最佳候选lucCageTrop627能够检测到cTnI,但其水平不足以用于临床(图11d)。由于用于诊断患者AMI的cTnI测定的排除和排除水平在低pM范围内,并且如上所述,我们的传感器平台的检测限(LOD)约为锁靶亲和力(K_{LT})的 $0.1 \times K_d$,我们通过将cTnC融合到其末端,进一步提高了传感器对cTnI的亲和力(图3a,图11b,c)。所产生的传感器lucCageTrop具有适合于临床样品定量的一位数pM LOD(图3b,图11e,f)。

[0157] 特异性抗体的检测对于监测病原体在人群中的传播(在病原体被消灭后很长时间内抗体仍然存在)、疫苗接种的成功率和治疗性抗体的水平非常重要。为了使我们的系统适用于这种抗体血清学分析,我们试图将目标抗体识别的线性表位结合到lucCage中,这样抗体的结合将打开开关,允许lucKey结合以及荧光素酶活性的重建。我们首先基于中和抗体(HzKR127)的晶体结构开发了一种抗乙型肝炎病毒(HBV)抗体传感器,该中和抗体与来自病毒表面蛋白L的PreS1结构域的肽结合²⁵。测试的8种设计中的最佳设计物lucCageHBV(HBV344)在添加HzKR127-3.2(HzKR127的改进版本²⁶)后,荧光素酶活性增加了约150%(图12a,B)。为了进一步提高lucCageHBV的动态范围和LOD($\sim 2\text{nM}$,图12c-e),我们通过锁末端引入额外的肽拷贝来提高锁靶亲和力(K_{LT}),以利用抗体与其表位的二价相互作用(图3c,d)。所得设计物,命名为lucCageHBV α ,LOD为260pM,动态范围为225%(图3e;图13a-c),发光强度易于用相机检测(图13d)。因此,该平台可检测LOD在监测治疗性抗体范围内的特异性抗体。接下来,我们演示了lucCageHBV传感器用于检测乙型肝炎表面抗原(HBsAg)。由于我们的传感器处于热力学控制下,我们假设传感器-抗体复合物的预组装将在靶标HBsAg蛋白PreS1存在下重新平衡,抗体重新分布以结合游离的PreS1,而不是lucCageHBV上的表位(图3f)。事实上,添加PreS1结构域后不久,lucCageHBV加HzKR127-3.2混合物的发光降低(图3g);该读数的灵敏度使PreS1浓度能够在临床相关范围内定量28(图3h,图12f)。HBsAg血清清除率是监测肝炎诊断和疫苗接种效果后治疗进展的主要生物标志物之一,但目前的商业HBsAg检测无法区分三种HBsAg蛋白亚型。我们的PreS1传感器(检测HBsAg L抗原)表明该系统可以实现亚型特异性识别。

[0158] 新冠肺炎大流行表明,迫切需要开发新的诊断工具,通过检测SARS-CoV-2病毒本身来跟踪活动感染,并检测抗病毒抗体,以评估病毒在人群中的传播程度,并确定未来感染风险较低的个体。为了设计抗SARS-CoV-2抗体的传感器,我们首先从文献中鉴定了SARS-CoV^{31,32}和SARS-CoV-2蛋白组^{33,34}中的高度免疫原性线性表位,这些表位不存在于冠状病毒科的“常见”毒株中(即HCoV-OC43、HCoV-HKU1、HCoV-229E和HCoV-NL63;我们没有排除对SARS-CoV或MERS的反应性,因为它们的分布范围要小得多)。其中,我们重点研究了SARS和

新冠肺炎患者血清中发现的膜蛋白和核衣壳蛋白中的两种表位,其交叉反应动物源性抗体是商用的(见图4图例以及表位和抗体描述的材料和方法)。我们为每种表位设计了传感器(图14a,b),并鉴定了对纯抗M和抗N蛋白抗体的存在特异性反应的设计(图4b,c)。这些传感器速度快(2-5分钟达到完全信号),对低纳摩尔量抗体的响应具有约50-70%的动态范围(图4b,c,图14c,d)。

[0159] 为了创建能够直接检测SARS-CoV-2病毒颗粒的传感器,我们在LucCage格式中整合了一种设计的与SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的皮摩尔亲和性结合域,命名为LCB1(图4d)。在测试的13种候选中,最好地,我们称为lucCageRBD,具有最小的背景,出色的动态范围(1700%),易于用相机检测,LOD低(15pM)(图4d,图15)。该传感器优越的动态范围和灵敏度是LCB1对RBD(K^{LT})的高亲和力的结果,与我们的热力学模型一致,突出了LucCage传感器平台和从头结合域设计的协同作用。

[0160] 由于LucCage系统的模块化和可工程性,设计SARS-CoV-2抗体和RBD传感器、获得合成基因、表达和纯化蛋白以及评估传感器性能只用了三周时间。

[0161] 为了测试本工作中开发的生物传感器的特异性(不包括lucCageHBV和lucCageRBD对PreS1的间接检测),我们测定了每一种对所有靶标(Bcl-2、肉毒神经毒素B、IgG-Fc、Her2、心肌钙蛋白I、单克隆抗HBV抗体(HzKR127-3.2)、抗SARS-CoV-1-M多克隆抗体(克隆3527)、抗SARS-CoV-1-N单克隆抗体((克隆18F629.1)和PreS1)。如图5所示,每种传感器都对其同源靶标作出快速而敏感的反应,但对其他任何一种都没有反应。表12和表13分别列出了每种lucCage传感器特性和所用传感域的汇总。

[0162] 大多数以前的基于蛋白的生物传感器平台依赖于靶标-传感器相互作用的特定几何形状来触发报告成分的构象变化,因此专门用于检测挑战的子集。由于这种靶标依赖性,可能需要相当大的优化来实现新靶标的高灵敏度检测。我们的传感器平台基于系统定义的闭合和开放状态之间的热力学耦合,因此,其灵敏度取决于与靶标结合的传感域的自由能变化,而不是结合相互作用的特定几何结构。这使得能够结合各种结合方式,包括小肽、球状迷你蛋白、抗体表位和从头设计的结合物,以产生用于广泛蛋白靶标的敏感传感器,而几乎没有优化或没有优化。对于即时护理(POC)应用,我们的系统具有均匀、无需清洗、全溶于水、几乎瞬时读出的优点,并且可以通过廉价且可访问的设备(如手机摄像头)来实现发光的量化。在医院环境中,在相同原理下可预测地制造多种传感器的能力可以使用例如384孔板上的数百种不同传感器阵列快速读出大量不同的化合物。

[0163] 直到最近,从头蛋白设计的重点放在了设计具有对应于单个密集(deep)自由能最小值的新结构的蛋白;我们的结果突出了该领域的进展,使得更复杂的多状态系统能够容易地生成。我们的传感器在细胞中高水平表达,并且非常稳定,这大大促进了进一步的制造过程。我们平台的通用“分子设备”架构与高亲和力微蛋白结合域从头设计的互补进展的协同作用特别好,可以设计为易于与lucCage平台兼容的三维结构。LucCageRBD强调了这种完全从头开始的方法的潜力,具有直接从计算机中出来的传感器1700%的动态范围和15pM的LOD,无需任何实验优化。

参考文献

1. Stein, V. & Alexandrov, K. Synthetic protein switches: design principles and applications. *Trends Biotechnol.* 33, 101-110 (2015).

2.Langan,R.A.et al.De novo design of bioactive protein switches.Nature 572,205-210(2019) .

3.Adams,E.R.et al.Antibody testing for COVID-19:A report from the National COVID Scientific Advisory Panel.medRxiv 2020.04.15.20066407(2020) .

4.Yeh,H.-W.&Ai,H.-W.Development and Applications of Bioluminescent and Chemiluminescent Reporters and Biosensors.Annu.Rev.Anal.Chem.12,129-150 (2019) .

5.Greenwald,E.C.,Mehta,S.&Zhang,J.Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks.Chem.Rev.118,11707-11794(2018) .

6.Schena,A.,Griss,R.&Johnsson,K.Modulating protein activity using tethered ligands with mutually exclusive binding sites.Nat.Commun.6,7830 (2015) .

7.Arts,R.et al.Semisynthetic Bioluminescent Sensor Proteins for Direct Detection of Antibodies and Small Molecules in Solution.ACS Sens 2, 1730-1736(2017) .

8.Xue,L.,Prifti,E.&Johnsson,K.A General Strategy for the Semisynthesis of Ratiometric Fluorescent Sensor Proteins with Increased Dynamic Range.J.Am.Chem.Soc.138,5258-5261(2016) .

9.Guo,Z.et al.Generalizable Protein Biosensors Based on Synthetic Switch Modules.J.Am.Chem.Soc.141,8128-8135(2019) .

10.Edwardraja,S.et al.Caged activators of artificial allosteric protein biosensors.ACS Synth.Biol. (2020)doi:10.1021/acssynbio.9b00500.

11.Ribeiro,L.F.,Warren,T.D.&Ostermeier,M.Construction of Protein Switches by Domain Insertion and Directed Evolution.Methods Mol.Biol.1596,43-55(2017) .

12.Dixon,A.S.et al.NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells.ACS Chem.Biol.11,400-408(2016) .

13.Minor,D.L.,Jr&Kim,P.S.Context-dependent secondary structure formation of a designed protein sequence.Nature 380,730-734(1996) .

14.Huang,P.-S.et al.RosettaRemodel:a generalized framework for flexible backbone protein design.PLoS One 6,e24109(2011) .

15.Chevalier,A.et al.Massively parallel de novo protein design for targeted therapeutics.Nature 550,74-79(2017) .

16.Deis,L.N.et al.Suppression of conformational heterogeneity at a protein-protein interface.Proc.Natl.Acad.Sci.U.S.A.112,9028-9033(2015) .

17.Eigenbrot,C.,Ultsch,M.,Dubnovitsky,A.,Abrahmsén,L.&**Hård**, T.Structural basis for high-affinity HER2 receptor binding by an engineered

protein. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15039-15044 (2010).

18. Hobbs, R.J., Thomas, C.A., Halliwell, J. & Gwenin, C.D. Rapid Detection of Botulinum Neurotoxins-A Review. *Toxins* 11, (2019).

19. Perrier, A., Gligorov, J., Lefèvre, G. & Boissan, M. The extracellular domain of Her2 in serum as a biomarker of breast cancer. *Lab. Invest.* 98, 696-707 (2018).

20. Yu, Q. et al. Semisynthetic sensor proteins enable metabolic assays at the point of care. *Science* 361, 1122-1126 (2018).

21. Rubini Gimenez, M. et al. One-hour rule-in and rule-out of acute myocardial infarction using high-sensitivity cardiac troponin I. *Am. J. Med.* 128, 861-870. e4 (2015).

22. Collins, M.H. Serologic Tools and Strategies to Support Intervention Trials to Combat Zika Virus Infection and Disease. *Trop Med Infect Dis* 4, (2019).

23. Pond é, R.A. de A. Expression and detection of anti-HBs antibodies after hepatitis B virus infection or vaccination in the context of protective immunity. *Arch. Virol.* 164, 2645-2658 (2019).

24. van Rosmalen, M. et al. Dual-Color Bioluminescent Sensor Proteins for Therapeutic Drug Monitoring of Antitumor Antibodies. *Anal. Chem.* 90, 3592-3599 (2018).

25. Chi, S.-W. et al. Broadly neutralizing anti-hepatitis B virus antibody reveals a complementarity determining region H3 lid-opening mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9230-9235 (2007).

26. Kim, J.H. et al. Enhanced humanization and affinity maturation of neutralizing anti-hepatitis B virus preS1 antibody based on antigen-antibody complex structure. *FEBS Lett.* 589, 193-200 (2015).

27. Ovacik, M. & Lin, K. Tutorial on Monoclonal Antibody Pharmacokinetics and Its Considerations in Early Development. *Clin. Transl. Sci.* 11, 540-552 (2018).

28. Locarnini, S. & Bowden, S. Hepatitis B surface antigen quantification: Not what it seems on the surface. *Hepatology* vol. 56 411-414 (2012).

29. Cornberg, M. et al. The role of quantitative hepatitis B surface antigen revisited. *Journal of Hepatology* vol. 66 398-411 (2017).

30. Perera, R.A. et al. Serological assays for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), March 2020. *Euro Surveill.* 25, (2020).

31. Chow, S.C.S. et al. Specific epitopes of the structural and hypothetical proteins elicit variable humoral responses in SARS patients. *J. Clin. Pathol.* 59, 468-476 (2006).

32. He, Y., Zhou, Y., Siddiqui, P., Niu, J. & Jiang, S. Identification of immunodominant epitopes on the membrane protein of the severe

acuterespiratory syndrome-associated coronavirus. *J.Clin.Microbiol.* 43,3718-3726 (2005) .

33.Wang,H.et al.SARS-CoV-2 proteome microarray for mapping COVID-19 antibody interactions at amino acid resolution. (2020) doi:10.1101/2020.03.26.994756.

34.Dahlke,C.et al.Distinct early IgA profile may determine severity of COVID-19 symptoms:an immunological case series.medRxiv 2020.04.14.20059733 (2020) .

35.Yu,Q.et al.A biosensor for measuring NAD levels at the point of care. *Nature Metabolism* vol.1 1219-1225 (2019) .

36.Arts,R.et al.Detection of Antibodies in Blood Plasma Using Bioluminescent Sensor Proteins and a Smartphone. *Anal.Chem.* 88,4525-4532 (2016) .

37.Tenda,K.et al.Paper-Based Antibody Detection Devices Using Bioluminescent BRET-Switching Sensor Proteins. *Angewandte Chemie* vol.130 15595-15599 (2018) .

38.Adamson,H.et al.Affimer-Enzyme-Inhibitor Switch Sensor for Rapid Wash-free Assays of Multimeric Proteins. *ACS Sens.* 4,3014-3022 (2019) .

39.Schena,A.,Griss,R.&Johnsson,K. Corrigendum: Modulating protein activity using tethered ligands with mutually exclusive binding sites. *Nat.Commun.* 7,10758 (2016) .

40.Berger,S.et al.Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *Elife* 5, (2016) .

41.Jin,R.,Rummel,A.,Binz,T.&Brunger,A.T. Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. *Nature* 444,1092-1095 (2006) .

42.Shen,A.et al.Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin. *Nat.Chem.Biol.* 5,469-478 (2009) .

43.Otwinowski,Z.&Minor,W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276,307-326 (1997) .

44.Liebschner,D.et al.Macromolecular structure determination using X-rays,neutrons and electrons:recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75,861-877 (2019) .

45.Potterton,L.et al.Developments in the CCP4molecular-graphics project. *Acta Crystallogr.D Biol.Crystallogr.* 60,2288-2294 (2004) .

方法

传感器系统的设计:lucCage和lucKey

[0164] SmBit (VTGYRLFEEIL;SEQ ID NO:27359) 使用GraftSwitchMover (一种基于

RosettaScripts™的蛋白设计算法)接枝到Langan等人,2019中描述的不对称LOCKR开关的闭环中(详见补充方法)。接枝取样范围指定在残基300-330之间。所得设计是能量最小化的,目视检查并选择用于随后的基因合成、蛋白生产和生化分析。实验确定闭环上的最佳SmBit位置为残基312处的插入,如图6所示。lucKey是通过将NanoLuc¹²的LgBit与Langan等人,2019中描述的钥匙肽进行基因融合而组装的。(完整序列列表见表10)

传感结构域到lucCage的计算接枝

[0165] 肽和表位:使用Rosetta™ GraftSwitchMover将每种传感结构域的氨基酸序列接枝到lucCage的325-360残基之间的所有 α -螺旋寄存器(register)中(详见补充方法)。所得lucCages的能量最小化,目视检查,通常选择少于10种设计用于随后的蛋白生产和生化表征。

[0166] 蛋白结构域:首先,鉴定形成结合蛋白相互作用表面的主要二级结构元件,提取其氨基酸序列,并使用如上所述的GraftSwitchMover将其接枝到lucCage中。然后,我们使用Rosetta™ Remodel¹⁴在开关的背景下对全长结合域进行建模,在开关中该界面被掩埋在笼子上(详见补充方法)。这些设计将能量最小化,并进行目视检查以供选择。通常,选择少于十种设计用于生化表征。

合成基因构建

[0167] 设计的蛋白序列被密码子优化用于大肠杆菌表达(IDT密码子优化工具),并在pET21b+或pET29b+大肠杆菌表达载体(IDT)中作为合成基因排序。将合成基因插入每种载体的NdeI和XhoI位点,包括N端六组氨酸标签,随后是TEV蛋白酶切割位点,并在C端添加终止密码子。

细菌蛋白生产和纯化的一般程序

[0168] 用编码目标合成基因的pET21b+或pET29b+质粒转化大肠杆菌LEM021(DE3)菌株(NEB)。将细胞在补充有羧苄青霉素或卡那霉素的LB培养基中生长24小时。将细胞以1:50mL的比例接种在补充有羧苄青霉素或卡那霉素的Studier TBM-5052自诱导培养基中,在37℃下生长2-4小时,然后在18℃下再生长18小时。通过在4000g、4℃下离心15分钟收集细胞,并重新悬浮在30mL裂解缓冲液(20mM Tris-HCl pH 8.0、300mM NaCl、30mM咪唑、1mM PMSF、0.02mg/ml DNA酶)中。通过超声处理2.5分钟(5秒循环)裂解细胞重悬液。通过在24000g 4℃下离心20分钟澄清裂解物,并穿过用洗涤缓冲液(20mM Tris HCl pH 8.0,300mM NaCl,30mM咪唑)预平衡的2ml Ni-NTA镍树脂(Qiagen,30250)。用10柱体积(CV)的洗涤缓冲液洗涤树脂两次,然后用3CV的洗脱缓冲液(20mM Tris HCl pH 8.0,300mM NaCl,300mM咪唑)洗脱。使用Ultra-15离心过滤单元(Amicon)浓缩洗脱的蛋白,并使用Superdex™ 75 Increase 10/300 GL(GE Healthcare)尺寸排除柱在Tris缓冲盐水(TBS;25mM Tris HCl pH 8.0,150mM NaCl)中进一步纯化。将含有单体蛋白的级分合并、浓缩并在液氮中速冻,并储存于-80℃。

体外生物发光表征

[0169] 用Synergy™ Neo2微孔板阅读器(BioTek)进行所有体外生物发光测定。在1:1=HBS-EP:Nano-Glo检测缓冲液中进行检测,用于抗HBV和RBD传感器,而1:1=DPBS:Nano-Glo检测缓冲液用于其他传感器。首先从储备溶液中制备所需浓度的10X lucCage、10X luc Key和10X靶蛋白。对于白色不透明96孔板的每个孔,将10 μ L 10X lucCage、10 μ L 10 X

lucKey和20 μ L缓冲液混合以达到指定的浓度和比例。将平板以1000 \times g离心1分钟,并在室温下再孵育10分钟。然后,将50 μ L稀释的furimazine (Nano-GloTM荧光素酶测定试剂, Promega)添加到每个孔中。注射后每1分钟进行一次无靶生物发光测定(0.1秒积分和10秒振荡)。约15分钟后,注入10 μ L连续稀释的10X靶蛋白加空白,并持续生物发光动力学采集总共2小时。为了从生物发光到分析物图(bioluminescence-to-analyte plot)中得出EC₅₀值,将各个分析物浓度下的前三个峰值生物发光强度取平均值,从空白中减去,并用于拟合S形4PL曲线。为了计算LOD,提取传感器对其分析物的生物发光响应的线性区,并获得线性回归曲线。它用于导出响应的标准偏差(SD)和校准曲线的斜率(S)。检测限为3 \times (SD/S)。实验测定结果一式三份,适用时显示平均值。使用不同批次的纯蛋白在不同的日期成功地对结果进行了重复。

生物膜层干涉法(BLI)

[0170] 使用**Octet®RED96**系统(ForteBio)和链霉亲和素涂层生物传感器(ForteBio)测定蛋白-蛋白相互作用。每个孔含有200 μ L溶液,测定缓冲液为HBS-EP+缓冲液(10mM HEPES pH 7.4,150mM NaCl,3mM EDTA,0.05%v/v表面活性剂P20,0.5%脱脂奶粉)。将生物传感器尖端以20 μ g/mL的分析物肽/蛋白上样300s(0.5nm响应阈值),在HBS-EP+缓冲液中孵育60s以获得基线测定值,浸入含有Cage和/或Key的溶液中600s(关联步骤),并浸入HBS-EP-缓冲液中600s。使用ForteBio数据分析软件9.0.0.10版分析结合数据。

lucCageBim的设计与表征

[0171] 将Bim肽序列(EIWIAQELRRIGDEFNAYYAA;SEQ ID NO:27380)穿入lucCage骨架中,如“将传感结构域设计到lucCage中”一节所述。所选设计在大肠杆菌中表达、纯化并对发光激活进行表征。在存在或不存在200nM的靶Bcl-2蛋白的情况下,测定测定与20nM lucKey混合的20nM每种设计物lucCage的生物发光检测信号。Bcl-2在其他地方描述⁴⁰。

lucCageHer2、lucCagerProA、lucCageBot和lucCageRBD的设计和表征

[0172] Bot.0671.2从头结合域、金黄色葡萄球菌蛋白A结构域C(SpaC)、Her2亲和抗体和从头RBD结合域LCB1的主要结合基序被穿入lucCage,如“将传感结构域设计到lucCage中”一节所述(传感结构域的序列见表13)。所选设计在大肠杆菌中表达、纯化并对发光激活进行表征。在存在或不存在200nM靶蛋白的情况下,测定与20nM lucKey混合的20nM每种设计物lucCage的生物发光检测信号。使用的靶蛋白是:如前所述⁴¹表达的肉毒杆菌神经毒素B HcB、人IgG1 Fc HisTag(AcroBiosystems,目录号IG1-H5225)和人Her2 HisTag(Acrobiosystem,目录号HE2-H5225)。

lucCageTrop的设计与表征

[0173] 心肌钙蛋白T(cTnT)结合基序(EDQLREKAKELWQTIYNLEAEKFDLQEKFKQKYEINVLNRN INDNQ;SEQ ID NO:27390)被拆分成不同长度的片段(见图11),并穿入lucCage支架,如“将传感结构域设计到lucCage”部分所述。所选设计物在大肠杆菌中表达、纯化并对发光激活进行表征。在存在或不存在100nM心肌钙蛋白I(Genscript,目录号Z03320-50)的情况下,测定与20nM lucKey混合的20nM每种设计物lucCage的生物发光检测信号。随后,通过将以下序列与lucCageTrop627的C端进行基因融合,创建了lucCage Trop,这是一种通过与心肌钙蛋白C(cTnC)融合的改进型(KVSKTKDSDSKGKSEEELSDLFRMFDKNADGYIDLEELKIMLQATGETITEDD IEELMKDGDKNNDGR[DYDEFLEFMKGVE;SEQ ID NO:27627)。

lucCageHBV和lucCage HBV α 的设计与表征

[0174] 使用Rosetta™ GraftSwitchMover将PreS1结构域的结合基序(GANSNNPDWDFN(SEQ ID NO:27629))在残基336之后的每个位置处穿入lucCage骨架。按照Rosetta™FastRelax方案,选择八种设计物用于蛋白生产。在存在或不存在抗HVB抗体HzKR127-3.2(100nM)的情况下,用设计的lucCage(20nM)和lucKey(20nm)测定生物发光,以选择lucCage HBV。随后,通过将含有第二抗原基序的序列(GGSGGGSSGFGANSNNPDWDFNPN;SEQ ID No:27628)与lucCageHBV基因融合,构建了lucCage HBV α 。

lucCageSARS2-M和lucCage SARS2-N的设计与表征

[0175] SARS-CoV-2膜蛋白(a.a.1-31、1-17和8-24)和核衣壳蛋白(a.a.368-388和369-382)的抗原表位通过计算接枝到lucCage中,如“将传感结构域设计到lucCage中”一节所述。所选设计在大肠杆菌中表达、纯化并表征发光激活。将50nM的所有设计物与50nM lucKey混合,并在100nM下在兔抗SARS-CoV膜多克隆抗体(ProSci,目录号:3527)或小鼠抗SARS-CoV核衣壳单克隆抗体(克隆18F629.1,NovusBio目录号:NBP2-24745)的存在下进行发光增加的实验筛选。

sCageHA变体的设计和表征

[0176] HB1.9549.2被嵌入亲本六螺旋束中,用于在骨架的闭环螺旋的不同位置进行骨架设计。为了促进更有利的分子内相互作用,特意用甘氨酸取代闭环上的三个连续残基,以允许构象自由。这五种设计物是在大肠杆菌中产生的。在存在或不存在钥匙(2 μ M)的情况下,使用Octet™仪器(ForteBio)将纯化的笼(1 μ M)和生物素化的甲型H1流感血凝素(HA)¹⁵上样到链霉亲和素涂层的生物传感器探头(ForteBio)上,进行生物膜层干涉分析。

HzKR127-3.2的生产和纯化

[0177] 将合成的V_H和V_L DNA片段亚克隆到含有人C γ 1和C γ κ DNA序列。使用Lipofectamine™(Invitrogen)将载体导入HEK 293T细胞中,并在37℃湿度培养箱中在5% CO₂的FreeStyle™ 293(GIBCO)中培养细胞。将培养上清液上样到蛋白A-Sepharose™柱(Millipore)上,通过添加0.2M甘氨酸-HCl(pH 2.7)洗脱结合的抗体,然后立即用1M Tris-HCl中和(pH 8.0)。用10mM HEPES NaOH(pH 7.4)透析溶液,并通过SDS-PAGE分析蛋白的纯度。

PreS1结构域的产生和纯化

[0178] 将编码PreS1结构域(残基1-56)的DNA片段克隆到pGEX-2T(GE Healthcare)质粒中,并在18℃下在大肠杆菌BL21(DE3)菌株(NEB)中产生该蛋白,作为N端谷胱甘肽-S-转移酶(GST)的融合蛋白。在缓冲溶液(25mM Tris-HCl pH 8.0,300mM NaCl)中制备细胞裂解物,并将澄清的上清液上样到GSTBind™树脂(Novagen)。用含有额外10mM还原型谷胱甘肽的相同缓冲液洗脱GST-PreS1结构域,使用Superdex™ 75 Increase 10/300 GL(GE Healthcare)尺寸排除柱进一步纯化,并浓缩至34 μ M。

sCageHA_267-1S及其变体的生产

[0179] sCageHA_267-1S和sCageHA_267-1S(E99Y/T144Y)在18℃下在大肠杆菌LEM021(DE3)菌株(NEB)中表达,作为C端包含(His)₁₀-标签标记的衍生自霍乱弧菌(Vibrio cholerae)⁴²的半胱氨酸蛋白酶结构域(CPD)融合蛋白。使用HisPur™镍树脂(Thermo)、HiTrap™ Q阴离子交换柱(GE Healthcare)和HiLoad 26/60 Superdex™75凝胶过滤柱(GE

Healthcare) 纯化蛋白。对于硒蛋氨酸 (SeMet) 标记, 额外引入 I30M 突变以产生 sCageHA_{267-1S} (E99Y/T144Y/I30M) 变体。该蛋白在含有 SeMet 的最小培养基中在大肠杆菌 B834 (DE3) RIL 菌株 (Novagen) 中表达, 并根据纯化其他变体的相同程序进行纯化。

sCageHA_{267-1S} 的结晶和结构测定

[0180] 引入两点突变 (Glu99Tyr 和 Thr144Tyr) 试图诱导有利的晶体堆积相互作用。通过在含有 11% (v/v) 乙醇、0.25M NaCl、0.1M TrisHCl (pH 8.5) 的溶液中进行微晶接种, 在悬滴气相扩散设置中获得了高质量的 sCageHA_{267-1S} 单晶 (E99Y/T144Y/I30M)。晶体需要将温度严格保持在 25°C。为了低温保护, 将晶体短暂浸泡在补充有 15% 2,3-丁二醇的结晶溶液中, 并在液氮中快速冷却。在 Se 吸收峰处收集单波长异常色散 (SAD) 数据集, 并用 HKL2000⁴³ 进行处理。使用 PHENIX⁴⁴ 中的 AutoSolTM 模块计算 Se 位置和初始电子密度图。使用 COOT⁴⁵ 和 PHINIX 进行模型构建和结构修正。

补充信息

补充讨论:

[0181] 我们基于从头开关的通用蛋白传感系统依赖于定义的封闭状态 (K_{open}) 和定义的开放状态 (K_{LT} 和 K_{CK}) 之间的热力学耦合 (见图 1a-c)。使用我们的系统, 不仅可以通过结合已知的结合结构域, 还可以通过结合从头结合域来实现对任意靶标的靶标特异性, 在从头结合域中, 我们可以完全控制蛋白折叠和几何结构。由于我们的系统中没有柔性或半柔性接头, 并且我们能够设计不同类型的笼结合结构域相互作用, 因此构象变化与结合靶标相互作用解耦, 这使得该系统在开放状态下结构上更可预测。RosettaTM 中最新开发的 GraftSwitchMover 可以一步完成传感器设计, 而无需使用其他形式来根据经验重新设计传感器配置。LucKey 与传感器的开放形式的分子间结合产生发光信号, 提供额外的可调参数 K_{CK} , 可与 K_{open} 一起优化, 以最大化传感器动态范围、分析范围、特异性和灵敏度。

补充方法

1. 热力学模型

[0182] 平衡常数定义为开锁打开的 K_{open} (等式 1), lucCage 和 lucKey 的解离常数为 K_{CK} (等式 2 和 3), 开锁和靶标的解离系数为 K_{LT} (等式 4 和 5)。 K_R 描述了重组荧光素酶的平衡, 该平衡由报告的 NanoBit 系统的解离常数 ($190\mu\text{M}^{19}$) 和拆分对应物的有效局部浓度 (C_{eff}) 确定 (等式 6 和 7)。我们在这里将 C_{eff} 设置为 1mM, 因为文献建议了针对分子内相互作用伴侣的高微摩尔至低微摩尔的范围²⁰, 并且我们的模块化开关应比柔性接头跨越更短的距离。每种组分的总量是恒定的, 因此引入了等式 8、9 和 10。给定四个平衡常数 (K_{open} 、 K_{CK} 、 K_{LT} 和 K_R) 和三个总浓度 ([lucCage] 总量、[lucKey] 总量和 [target] 总量), python 模块 sympy.nsolve 用于数值求解这十个方程, 并找出平衡状态下各类别的浓度。从溶液中提取发光类别 6 和 7 的总浓度, 除以 [lucCage] 总量, 并绘制相应的图, 图 1d 为各种 K_{open} , 图 1e 为 K_{LT} , 图 1f 为 [lucCage] 总量和 [lucKey] 总量。图 1f 的数字在 0-1 之间归一化。

公式1:

$$K_{\text{open}} = \frac{[2]}{[1]}$$

公式2:

$$K_{\text{CK}} = \frac{[2] \times [\text{lucKey}]_{\text{free}}}{[5]}$$

公式3:

$$K_{\text{CK}} = \frac{[3] \times [\text{lucKey}]_{\text{free}}}{[4]}$$

公式4:

$$K_{\text{LT}} = \frac{[6] \times [\text{target}]_{\text{free}}}{[7]}$$

公式5:

$$K_{\text{LT}} = \frac{[2] \times [\text{target}]_{\text{free}}}{[3]}$$

公式6:

$$K_{\text{R}} = \frac{190 \mu\text{M}}{C_{\text{eff}}} = \frac{[5]}{[6]}$$

公式7:

$$K_{\text{R}} = \frac{[4]}{[7]}$$

公式8:

$$[\text{lucCage}]_{\text{total}} = [1] + [2] + [3] + [4] + [5] + [6] + [7]$$

公式9:

$$[\text{lucKey}]_{\text{total}} = [\text{lucKey}]_{\text{free}} + [4] + [5] + [6] + [7]$$

公式10:

$$[\text{target}]_{\text{total}} = [\text{target}]_{\text{free}} + [3] + [4] + [7]$$

2. 传感结构域到lucCage的计算接枝

[0183] lucCage传感器的结构模型是通过将每种传感域接枝到lucCage骨架的闩锁上来创建的(见表13)。使用RosettaScriptsTM方案(GraftSwitch_relax.xml,参见编码可用性)进行设计,将带有注释界面残基(sensing_domains.fasta,参见编码可用性)的传感域列表穿入lucCage模型(lucCage.pdb,参见编码可用性)。bash脚本(run_GraftSwitch.sh,参见编码可用性)用于调用RosettaScriptsTM。该方案使用两个连续的RosettaTM移动器(mover):(i)GraftSwitchMover将所需的传感域序列穿入lucCage闩锁的定义区(氨基酸325-359),并选择具有埋在笼/闩锁界面中定义的“重要残基”的设计;(ii)和MultiplePoseMover,以松弛(FastRelax,在给定来自前一个移动器的突变的情况下,找到最低能量结构),过滤和评分来自前一移动器的每个输出模型。通过目视在PyMol 2.0中对所得设计进行了进一步评估。这是通过选择在新螺纹序列和笼之间表现出良好疏水嵌合作用的设计,并舍弃具有可能破坏传感器闭合状态的不利包埋亲水残基的设计(除非这些残基被注释为“重要残基”)。

[0184] 为了接枝具有预定义三级结构(即Bot.671.2、SpaC和Her2亲和抗体)的迷你蛋白结合域,我们首先确定了结合蛋白与其靶的主要相互作用表面,并确定了参与其中的主要二级结构元素。我们将这些元素的氨基酸序列添加到了sensing_domains.fasta文件以在上述方案中使用它们。输出是带有接枝界面元素的lucCage设计模型。然后,我们使用RosettaTM Remodel结构域插入²¹对开关背景中的全长传感域进行建模(Remodel_domain_insertion.sh,参见编码可用性),然后使用Relax来找到最低能量结构(Relax.sh,见编码可用性),通过目视在PyMol 2.0中选择最佳设计物。

补充信息参考文献

1. **Bahadır, E.B. & Sezgintürk, M.K.** Lateral flow assays: Principles, designs and labels. Trends Analyt. Chem. 82, 286-306 (2016).

2. Yeh, H.-W. & Ai, H.-W. Development and Applications of Bioluminescent and Chemiluminescent Reporters and Biosensors. *Annu. Rev. Anal. Chem.* 12, 129-150 (2019).
3. Greenwald, E.C., Mehta, S. & Zhang, J. Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks. *Chem. Rev.* 118, 11707-11794 (2018).
4. Glasgow, A.A. et al. Computational design of a modular protein sense-response system. *Science* 366, 1024-1028 (2019).
5. Guo, Z. et al. Generalizable Protein Biosensors Based on Synthetic Switch Modules. *J. Am. Chem. Soc.* 141, 8128-8135 (2019).
6. Yu, Q. et al. Semisynthetic sensor proteins enable metabolic assays at the point of care. *Science* 361, 1122-1126 (2018).
7. van Rosmalen, M. et al. Dual-Color Bioluminescent Sensor Proteins for Therapeutic Drug Monitoring of Antitumor Antibodies. *Anal. Chem.* 90, 3592-3599 (2018).
8. Adamson, H. et al. Affimer-Enzyme-Inhibitor Switch Sensor for Rapid Wash-free Assays of Multimeric Proteins. *ACS Sens.* 4, 3014-3022 (2019).
9. Tenda, K. et al. Paper-Based Antibody Detection Devices Using Bioluminescent BRET-Switching Sensor Proteins. *Angewandte Chemie* vol. 130 15595-15599 (2018).
10. Griss, R. et al. Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat. Chem. Biol.* 10, 598-603 (2014).
11. Arts, R. et al. Detection of Antibodies in Blood Plasma Using Bioluminescent Sensor Proteins and a Smartphone. *Anal. Chem.* 88, 4525-4532 (2016).
12. Lopez-Ruiz, N. et al. Smartphone-based simultaneous pH and nitrite colorimetric determination for paper microfluidic devices. *Anal. Chem.* 86, 9554-9562 (2014).
13. Leippe, D.M. et al. A bioluminescent assay for the sensitive detection of proteases. *Biotechniques* 51, 105-110 (2011).
14. Troy, T., Jekic-McMullen, D., Sambucetti, L. & Rice, B. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging* 3, 9-23 (2004).
15. Yeh, H.-W., Wu, T., Chen, M. & Ai, H.-W. Identification of Factors Complicating Bioluminescence Imaging. *Biochemistry* 58, 1689-1697 (2019).
16. Yeh, H.-W. et al. ATP-Independent Bioluminescent Reporter Variants To Improve in Vivo Imaging. *ACS Chem. Biol.* 14, 959-965 (2019).
17. Edwardraja, S. et al. Caged activators of artificial allosteric protein biosensors. *ACS Synth. Biol.* (2020) doi:10.1021/acssynbio.9b00500.
18. Langan, R.A. et al. De novo design of bioactive protein

switches.Nature 572,205-210(2019) .

19.Dixon,A.S.et al.NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells.ACS Chem.Biol.11,400-408(2016) .

20.Krishnamurthy,V.M.,Semetey,V.,Bracher,P.J.,Shen,N.& Whitesides,G.M.Dependence of Effective Molarity on Linker Length for an Intramolecular Protein-Ligand System.Journal of the American Chemical Society vol.129 1312-1320(2007) .

21.Huang,P.-S.et al.RosettaRemodel:a generalized framework for flexible backbone protein design.PLoS One 6,e24109(2011)

[0185] 表11.X射线数据收集和结构修正统计。

数据收集	SelMET-sCageHA_267-1S(E99Y/T144Y/I30M)
空间组	<i>C2</i>
晶胞尺寸	

a, b, c (Å)	178.993 60.127 71.799
a, β , γ (°)	90, 112.463, 90
波长(Å)	0.9794
分辨率 (Å)	50-2.03 (2.03-2.00) ^a
R_{sym}	6.6 (16.5) ^a
$I/\sigma(I)$	24.0 (3.5) ^a
完整性 (%), >1 σ	70.6 (33.8) ^a
冗余度	2.5 (1.6) ^a
修正	
分辨率 (Å)	46.09-1.99 (2.06-1.99) ^a
反射数	37603
$R_{\text{work}} / R_{\text{free}}$	0.2078/0.2515
R.m.s 偏差	
键(Å)/角 (°)	0.007/0.910
平均 B 值 (Å ²)	38.19
拉氏图 (Ramachandran plot) (%)	
有利/额外允许的	97.7/2.3
通常允许的	0.0

^a括号中的数字是最高分辨率shell的统计数据。

表12. 本工作中的生物传感器概述

生物传感器	分析靶标	动态范围 ^a	LOD (nM)	检测范围 (nM)
-------	------	-------------------	----------	-----------

lucCageBim	Bcl-2	200%	0.2	0.2-12.5
lucCageBoT	肉毒杆菌神经毒素 B	130%	0.4	0.4-50
lucCageProA	Fc 结构域	350%	0.39	0.39-12.5
lucCagHer2	Her2 受体	380%	0.23	0.23-25
lucCageTrop	肌钙蛋白 I	250%	0.009	0.009-0.3
lucCageHBV	抗 HBV 抗体 (HzKR127-3.2)	98%	2	2-100
lucCageHBV α	抗 HBV 抗体 (HzKR127-3.2)	215%	0.26	0.26-100
lucCageHBV+ HzKR127-3.2	PreS1	80%	0.6	0.6-100
lucCageSARS2-M	抗 SARS-Cov-M	50%	2.9	2.9-250
lucCageSARS2-N	抗 SARS-Cov-NP	70%	3.0	3.0-100
lucCageRBD	SARS-CoV-2 RBD	1700%	0.015	0.015-6

^a定义为强度变化 ($\Delta E/E_{\min}$)。 ΔE 是饱和靶标浓度下总生物发光发射的最大变化， E_{\min} 是没有分析靶标时的发射。

表13. 本工作中使用的传感域列表

生物传感器	传感结构域	传感结构域序列
lucCageBim	Bim	EIWIAQELRRIGDEFNAYYAAA (SEQ ID NO:27380)
lucCageBoT	Bot.0671.2	MFAELKAKFFLEIGDRDAARNALRKA GYSDEEAERIIRKYELE (SEQ ID NO:27381)
lucCageProA	蛋白 A 结构域 C (SpaC)	EQQNAFYEILHLPNLTEEQRNGFIQSL KDDPSVSKEILAEAKKLNDQAQPK

		(SEQ ID NO:27382)
lucCagHer2	Her2 亲和抗体	EMRNAYWEIALLPNLNNQKRAFIRS LYDDPSQSANLLAEAKKLNDAQAPK (SEQ ID NO:27383)
lucCageTrop	cTnI + cTnC	EDQLREKAKELWQTIYNLEAEKFDLQ EKFKQQKYEINVLNRNRINDNQKVSKT KDDSKGKSEEELSDLFRMFDKNADGY IDLEELKIMLQATGETITEDDIEELMKD GDKNNDGRIDYDEFLEFMKGVE (SEQ ID NO:27384)
lucCageHBV	preS1 (a.a. 35-46)	GANSNNPDWDFN (SEQ ID NO:27629)
lucCageHBV α	preS1 (a.a. 35-46) 2x	GANSNNPDWDFNGGSGGGSSGFGAN SNNPDWDFNPN (SEQ ID NO:27630)
lucCageSARS2-M	SARS-CoV-2 核衣壳 蛋白 (a.a. 369-382) 2x	MADSNGTITVEELKKLLEGGSGGMAD SNGTITVEELKKLLE (SEQ ID NO:27392)
lucCageSARS2-N	SARS-CoV-2 膜蛋白 (a.a. 1-17) 2x	KKDKKKKADETQALGGSGGKKDKKK KADETQAL (SEQ ID No:27548)
lucCageRBD	LCB1_Δ4	ILQKIYEIMRLLDELGHAEASMRVSDL IYEFMKKGDERLLEEAERLLEEVER (SEQ ID NO:27590)

实施例2

扩展基于LOCKR的生物传感器的通用读数

[0186] 上述传感器平台可以重新调整用途以容纳几乎所有的拆分报告分子,其中一个互补报告分子片段在遗传上融合到笼的N端,另一个片段融合到锁(分子内)或钥匙(分子间)的C端。各种类型的拆分蛋白对或RET对(图16)可以实现广泛的读数,例如生物发光(萤火虫荧光素酶¹、海肾荧光素酶²和高斯荧光素酶³)、生物发光共振能量转移⁴⁻⁶(BRET)、双分子荧光互补^{7,8}(BiFC)、荧光共振能量转移(FRET)、比色报告分子酰胺酶⁹、 β -半乳糖苷酶¹⁰和辣根过氧化物酶¹¹)、细胞存活报告分子(二氢叶酸还原酶¹²)、电化学报告分子(APEX2¹³)、放射性报告分子(胸苷激酶¹⁴)和分子条形码报告分子(TEV蛋白酶¹⁵)。

[0187] 本公开的从头开关平台可以是通用的和定制的,以检测任意目标靶标,但也可以用宽范围的读数重新编程,以用于不同的传感目的。对于细胞成像,具有BiFC或FRET读数的传感器可以提供优异的时空分辨率来监测细胞内靶标的动态。在广泛的合成生物学领域,传感器可以,例如,1)促进使用遗传生物传感器进行药物发现的基于细胞的多重分析;2)使用分子条形码(TEV蛋白酶)和下一代测序(NGS)作为读数技术,分析靶标选择途径上的化学

或遗传干扰;和3) 在选择的目标存在下,通过二氢叶酸还原酶(DHFR)互补进行细胞存活选择。对于体内成像,可以通过拆分发光蛋白或使用拆分胸苷激酶的正电子发射断层扫描(PET)来监测生物活性和蛋白靶标,从而允许在深层组织中成像。对于即时护理应用,比色读数提供了最方便的设置,因为信号采集不需要仪器。此外,电化学读数与最成功的POC装置——血糖仪兼容,后者可以读取电化学信号,以检测低丰度靶标。总的来说,我们预计,我们从头传感器设计、结合域设计和拆分蛋白重组的结合可以导致用户定义输入的应用程序的真正爆发。

[0188] 为了提供概念证明,我们设计了一种分子间BRET传感器(S0512)来检测HBV抗体,其中teLuc基因融合到笼中,CyOFP连接到钥匙的C端(图17A)。表位序列的两个拷贝被穿在锁上。在存在HBV抗体的情况下,观察到约5%的比值变化(580/450nm),检测限为约11nM。同时,我们还设计了一种分子内BRET传感器(S0622),在笼的N端和C端含有teLuc(BRET供体)和CyOFP(BRET受体)。该设计导致高初始BRET效率。在存在HBV抗体的情况下,抗体-锁驱动力将打破笼-锁的相互作用,然后增加BRET供体和受体的距离,导致BRET效率降低(图17B)。S0622的检测限为约11nM,比值变化为约20%的450/580nm。为了改善比值变化,我们优化了钥匙和CyOFP之间的连接长度。B0622_6显示出最高的初始BRET效率。当传感器保持低nM灵敏度时,观察到高达约207%的450/580nm比值变化(图17C)。同样,传感器的动态范围和灵敏度可以通过钥匙浓度来调节,这是我们模块化传感器平台中的可调因素之一。

[0189] 为了扩大即时护理应用的读数,我们利用拆分β-内酰胺酶报告了启动时笼和钥匙的组装情况。重组的β-内酰胺酶能够催化比色底物-头孢硝噻吩的水解,从而产生红色产物(OD 490)。对于即时护理应用,这种比色读数优于光学读数,因为人眼可以直接区分颜色变化。与闪光型生物发光(闪光型生物发光通常显示出突发发射导致时间依赖性信号采集的显著复杂性)相比,由此产生的比色产物在溶液中会随着时间的推移而累积。因此,这是一种终点测定(活性更高的β-内酰胺酶更快到达终点)。值得注意的是,β-内酰胺酶可以在生物流体(如血清和尿液¹⁹)中保持活性。这里的关键设计理念是尽可能降低背景活性,以减少假阳性的几率。我们通过简单地进行LacA-CageTrop融合和Key-LacB融合来证明lucCageTrp转化为LacATrop(图18)。β-内酰胺酶活性在人心肌钙蛋白I(cTnI)存在下启动。以低nM灵敏度获得了良好的标准曲线,并且人眼可以容易地确定从黄色到红色的颜色变化。

设计物序列:

S0512(粗体为teLuc序列;下划线为HBV表位):

MGSHHHHHHGSSENLYFQGS**GGVFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQN**
LGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDNHHFKVIL
HYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLF
RVTINGVTGWRLHERILAGS(SELARKLLEASTKLQRLNIRLAEALLEAIARLQELNLELV
 YLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESEKILEEAREAISGSGSELAKLL
 LKAI AETQDLNLR AAKAFLEAAAKLQELNIRAVELLVKLTDPATIREALEHAKRRSKEIIDE
 AERAIRAAKRESERIEEARRLIEKSGSGSELARELLRAHAQLQRLNLELLRELLRALAQL
 QELNLDLLRLASEL)TDPDEARKAANSNNPDWDFIVEDAERLIREAAAAANSNNPDWDFL
 IR (SEQ ID NO:27651)

Key-2GGSGG-Cy0FP (粗体为Cy0FP序列) :

MDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISREGGS
GG GGVSK **GEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE GKQTNRIKVV**
EGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPE GFTWERVMVF
EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK TLGWEPSTET
MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM PGVHYVDRRL
ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (SEQ ID NO: 27622)

B0622 (粗体为teLuc序列;粗体和下划线为Cy0FP序列;下划线为HBV表位) :

MGSHHHHHHSGSGSENLYFQSGGVFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQ
LGVSVTPIQRIVLSGENGLKIDIHVIPYEGLSGDQMGQIEKIFKVVYPVDNHHFKVIL
HYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLF
RVTINGVTGWRLHERILAGS(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLE
AIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEARK
AIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAA
KLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARLIEK
AKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQEL
NLDLLRLASEL)TDPDEARKAANSNPDWDFIVEDAERLIREAAAASEKISREAERLANSN
NPDWDFISRE VSKGEELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE
GKQTNRIKVV EGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPE
GFTWERVMVF EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (SEQ ID NO:27652)

B0622_4:

MGSHHHHHHSGSGSENLYFQSGGVFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQ
NLGVSVTPIQRIVLSGENGLKIDIHVIPYEGLSGDQMGQIEKIFKVVYPVDNHHFKVI
LHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSL
LFRVTINGVTGWRLHERILAGS(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
LEAIARLQELNLELVYLAVELTDPKRIRDEIKEVKDKSKEIIRRAEKEIDDAAKESKKILEEA
RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIIEEARL
IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
QELNLDLLRLASEL)TDPDEARKAANSNPDWDFIVEDAERLIREAAAASEKISREAERLA
NSNPDWDFISRE EELIK ENMRSKLYLE GSVNGHQFKC THEGEGKPYE
GKQTNRIKVV EGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPE
GFTWERVMVF EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK
TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM
PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (SEQ ID
NO:27653)

B0622_6:

MGSHHHHHHSGSGSENLYFQSGGVFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQ
 NLGVSVTPIQRIVLSGENGLKIDIHVIIPIYEGLSGDQMGQIEKIFKVVYPVDNHHFKVI
 LHYGTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSL
 LFRVTINGVTGWRLHERILAGS(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEAL
 LEAIARLQELNLELVYLAVELTDPKRIRDEIKEYKDKSKEIIRRAEKEIDDAAKESKKILEEA
 RKAIRDAAEESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEA
 AAKLQELNIRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIEEARLL
 IEKAKEESERIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQL
 QELNLDLLRLASEL)TDPDEARKAANSNPDWDFIVEDAERLIREAAAASEKISREAERLA
 NSNPDWDFISRE LIK ENMRSKLYLE GSVNGHOFKC THEGEGKPYE

GKQTNRIKVV EGGPLPFAFD ILATHFMYGS KVFYKYPADL PDYFKQSFPE

GFTWERVMVF EDGGVLTATQ DTSLODGELI YNVKVRGVNF PANGPVMQKK

TLGWEPSTET MYPADGGLEG RCDKALKLVG GGHLHVNFKT TYKSKKPVKM

PGVHYVDRRL ERIKEADNET YVEQYEHAVA RYSNLGGMD ELYK (SEQ ID

NO:27654)

KeyLacB(拆分 β -内酰胺酶B,粗体):

SGSGDPDEARKAIARVKRESKRIVEDAERLIREAAAASEKISREAERLIREAAAASEKISRES
 GGGSGGGGSGGGG LLTLASRQQLIDWME ADKVAGPLLR SALPAGWFIA
 DKSGAGERGS RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW
 (SEQ ID NO: 27623)

LacATrop(粗体为拆分 β -内酰胺酶A;下划线为cTnT和cTnC):

MGSHHHHHHSGSGSENLYFQSGGVSFAHPETLVK VKDAEDQLGA RVGYIELDLN
 SGKILESFRP EERFPMSTF KVLLCGAVLS RVDAGQEQLG RRIHYSQNDL
 VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
 HNMGDHVTRL DRWPELNEA IPNDERDTT PAAMATTLRK LLTGENGR
 SGGGSGGGGSGGGG(SKEAAKKLQDLNIELARKLLEASTKLQRLNIRLAEALLEAIARLQ
 ELNLELVYLAVELTDPKRIRDEIKEYKDKSKEIIRRAEKEIDDAAKESKKILEEARKAIRDAA
 EESRKILEEGSGSGSDALDELQKLNLELAKLLLKAI AETQDLNLRAAKAFLEAAAKLQELN
 IRAVELLVKLTDPATIRRALEHAKRRSKEIIDEAERAIRAAKRESERIEEARLLIEKAKEESE
 RIIREGSGSGDPDIKKLQDLNIELARELLRAHAQLQRLNLELLRELLRALAQLQELNLDLLR
 LASEL)TDPDEARKAIAVTGYRLFEEILDAERLIREAAAASEDQLREAAKELWQTIYNLEAE
KFDLQEKFKQKYEINVLNRNRINDNOKVSKTKDDSKGKSEEELSDLFRMFDKNADGYIDL
EELKIMLQATGETTIEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE (SEQ ID NO:27620)

参考文献

(1) Luker, K. E.; Smith, M. C.; Luker, G. D.; Gammon, S. T.; Piwnica-Worms, H.;
 Piwnica-Worms, D. Proc Natl Acad Sci U S A 2004, 101, 12288-12293.

(2) Kaihara, A.; Kawai, Y.; Sato, M.; Ozawa, T.; Umezawa, Y. Anal Chem 2003, 75,

4176-4181.

(3) Remy, I.; Michnick, S. W. *Nat Methods* 2006, 3, 977-979.

(4) Chu, J.; Oh, Y.; Sens, A.; Ataie, N.; Dana, H.; Macklin, J. J.; Laviv, T.; Welf, E. S.; Dean, K. M.; Zhang, F.; Kim, B. B.; Tang, C. T.; Hu, M.; Baird, M. A.; Davidson, M. W.; Kay, M. A.; Fiolka, R.; Yasuda, R.; Kim, D. S.; Ng, H. L., et al. *Nat Biotechnol* 2016, 34, 760-767.

(5) Yeh, H. W.; Karmach, O.; Ji, A.; Carter, D.; Martins-Green, M. M.; Ai, H. W. *Nat Methods* 2017, 14, 971-974.

(6) Yeh, H. W.; Xiong, Y.; Wu, T.; Chen, M.; Ji, A.; Li, X.; Ai, H. W. *ACS Chem Biol* 2019, 14, 959-965.

(7) Zhou, J.; Lin, J.; Zhou, C.; Deng, X.; Xia, B. *Acta Biochim Biophys Sin (Shanghai)* 2011, 43, 239-244.

(8) Ohashi, K.; Kiuchi, T.; Shoji, K.; Sampei, K.; Mizuno, K. *Biotechniques* 2012, 52, 45-50.

(9) Galameau, A.; Primeau, M.; Trudeau, L. E.; Michnick, S. W. *Nat Biotechnol* 2002, 20, 619-622.

(10) Wehrman, T. S.; Casipit, C. L.; Gewertz, N. M.; Blau, H. M. *Nat Methods* 2005, 2, 521-527.

(11) Martell, J. D.; Yamagata, M.; Deerinck, T. J.; Phan, S.; Kwa, C. G.; Ellisman, M. H.; Sanes, J. R.; Ting, A. Y. *Nat Biotechnol* 2016, 34, 774-780. (12) Remy, I.; Michnick, S. W. *Proc Natl Acad Sci U S A* 1999, 96, 5394-5399.

(13) Han, Y.; Branon, T. C.; Martell, J. D.; Boassa, D.; Shechner, D.; Ellisman, M. H.; Ting, A. *ACS Chem Biol* 2019, 14, 619-635.

(14) Massoud, T. F.; Paulmurugan, R.; Gambhir, S. S. *Nat Med* 2010, 16, 921-926.

(15) Wehr, M. C.; Holder, M. V.; Gailite, I.; Saunders, R. E.; Maile, T. M.; Ciirdaeva, E.; Instrell, R.; Jiang, M.; Howell, M.; Rossner, M. J.; Tapon, N. *Nat Cell Biol* 2013, 15, 61-U132.

(16) Landry, C. R.; Levy, E. D.; Abd Rabbo, D.; Tarassov, K.; Michnick, S. W. *Cell* 2013, 155, 983-989.

(17) Bowes, J.; Brown, A. J.; Hamon, J.; Jarolimek, W.; Sridhar, A.; Waldron, G.; Whitebread, S. *Nat Rev Drug Discov* 2012, 11, 909-922.

(18) Geyer, P. E.; Holdt, L. M.; Teupser, D.; Mann, M. *Mol Syst Biol* 2017, 13, 942.

(19) Adamson, H.; Ajayi, M. O.; Campbell, E.; Brachi, E.; Tiede, C.; Tang, A. A.; Adams, T. L.; Ford, R.; Davidson, A.; Johnson, M.; McPherson, M. J.; Tomlinson, D. C.; Jeuken, L. J. C. *ACS Sens* 2019, 4, 3014-3022.

实施例3

[0190] 如图20(小图A)所示,上述传感器平台可以重新调整用途,以适应“间接检测”方法,其中拆分的报告蛋白(分子间或分子内实施方案;分子间实施方案如下所示)通过生物

传感器与靶标结合多肽的靶标(例如抗体)的预孵育而重组,从而在本实施例中使发光激活。然后将激活的生物传感器与样品一起孵育以检测抗体结合的抗原的存在,使抗体与抗原结合,分离的报告蛋白组分之间的相互作用丧失,报告活性降低/消失(在这种情况下,发光活性丧失)。基于本文的公开将清楚,该实施方案可用于间接检测任何目标分析物。这种方法不限于使用抗体及其同源抗原。在另一个实施方案(小图B)中,通过将生物传感器与靶标结合多肽的靶标(例如SARS-CoV-2刺突蛋白)预孵育,对拆分报告蛋白进行重组(分子间或分子内实施方案;分子间实施方案如下所示),从而在本实施例中产生发光激活。然后将激活的生物传感器与样品一起孵育,以检测刺突结合的抑制剂(例如LCB1抑制剂)的存在,使抗体与抗原结合,拆分的报告蛋白组分之间的相互作用丧失,报告活性降低/消失(在这种情况下,发光活性丧失)。该方法可用于检测抑制剂,也可作为评估多种变体抑制效力的工具。基于本文的公开将清楚,该实施方案可用于间接检测任何目标分析物。另一个示例如图21所示。

示例性用途

[0191] 本文中的诊断传感器(lucCageBim、lucCage Bot、lucCageTrop、lucCageProA、lucCageHer2、lucKegeHBV、lucCageSARS2-M、lucCageSARS2-N),并测定每种传感器响应于其所有靶标(Bcl-2、肉毒神经毒素B、心肌钙蛋白I、IgG-Fc、Her2)、抗HBV(HzKR127-3.2)、抗SARS-M多克隆抗体(3527)、抗SARS-N单克隆抗体(18F629.1))。每种传感器都对其同源靶标快速而敏感地做出反应,但对其他任何传感器都没有反应(图19C)。通过SeroNet,可以将最佳CoV LOCKR诊断格式化为各种POCD(图19D)。

[0192] LOCKR诊断组合在血液或血清中存在抗冠状病毒“抗表位”特异性抗体的情况下激活化学发光,并可通过添加含有目标表位的抗原而关闭,如图22所示。

实施例4

[0193] SARS-CoV-2感染被认为通常从鼻子开始,病毒在那里复制数轮,然后传播到更广泛的呼吸系统。因此,将高浓度的病毒抑制剂递送到鼻子和呼吸系统通常可能在感染早期提供预防性保护和治疗效果,并且对于经常接触感染者的卫生保健工作者和其他人特别有用。许多单克隆抗体作为全身性SARS-CoV-2治疗药物正在开发中,但这些化合物不适合鼻内递送,因为抗体是大的且通常不是非常稳定的分子,并且结合位点的密度低(两个/150Kd抗体);Fc结构域提供了很少的附加益处。更理想的是对单克隆病毒具有非常高亲和力的蛋白抑制性,但具有更高的稳定性和非常小的尺寸,以最大化抑制结构域的密度,并能够通过雾化直接输送到呼吸系统中。

[0194] 我们着手重新设计与Ace2结合竞争的RBD的高亲和力结合物。我们探索了两种策略:首先,我们试图在Ace2中构建 α -螺旋,该 α -螺旋使大部分与RBD的相互作用发生在一种小的设计蛋白中,该蛋白与RBD进行额外的相互作用,以获得更高的亲和力;其次,我们试图从头开始设计不与RBD结合的结合物。第二种方法的优点是设计的可能性范围大得多,因此可以识别潜在的更高亲和力的结合模式。对于第一种方法,我们使用RosettaTM蓝图构建器生成包含Ace2螺旋的小蛋白,对于第二种方法,使用大的迷你蛋白文库进行RIF对接和设计。这些设计与围绕Ace2结合位点的RBD表面的不同区相互作用。方法1和方法2的设计物以长寡核苷酸编码,并筛选与酵母细胞表面荧光标记的RBD结合。深度测序确定了3种Ace2螺旋骨架设计(方法1)和150种从头界面设计(方法2),这些设计在FACS分选RBD结合后明显富

集。设计物在大肠杆菌中表达并纯化,在生物膜层干涉实验中发现许多设计具有可溶性表达并结合RBD,并且可以与ACE-2有效竞争结合RBD(示例如图2所示)。根据BLI数据,微结合域的RBD结合亲和力为:LCB1<1nM,LCB3<1nM。LCB2、LCB4、LCB5、LCB6、LCB7、LCB8的亲和力范围为1~20nM,不同结合域的相对强度为LCB4>LCB2>LCB9=LCB5>LCB6>LCB7。

[0195] 为了确定设计是否通过设计的界面结合RBD,构建了位点饱和文库,其中每种设计物中的每个残基一次一个地被20个氨基酸中的每一个取代,并对其进行FACS分类以进行RBD结合。深度测序表明,结合界面残基和蛋白核心残基在构建此类位点饱和文库(SSM)的许多设计中是保守的(SSM用于定义表3中氨基酸变化的允许位置)。对于大多数设计物,FACS分选中富集了少量取代,这表明它们增加了对RBD的结合亲和力。为了方法1设计物和8种方法2设计物的最高亲和力,构建包含这些取代的组合文库,并再次用FACS筛选结合;由于结合亲和力非常高,分选中使用的浓度低至20pM。每个文库都聚集在少量密切相关的序列上,对于每种设计物,其中一种优化的变体在大肠杆菌中表达并纯化。

[0196] 通过生物膜层干涉法研究了具有不同结合模式的8种优化设计物与RBD的结合。对于许多设计物,Kd的范围为1-20nM,而对于其余设计物,Kd低于1nM;太强,无法用该技术可靠测定。设计物的圆二色性光谱与设计模型一致,并且设计在室温下数天后保持完全结合活性。

[0197] 我们研究了这些设计阻断活病毒感染人类细胞的能力。在不同量的设计的结合域存在下,将100FFU的SARS-CoV-2添加到 $2.5-3 \times 10^4$ 非洲绿猴肾细胞(vero细胞)中。我们观察到了所有设计物对感染的有效抑制,IC50在1nM至0.02nM范围内。

[0198] 与抗体相比,设计的结合域作为潜在的治疗剂具有几个优点。总之,它们跨越了一系列的结合模式,结合在一起病毒逃逸的可能性很小。在高温下长时间保持活性表明它们不需要冷链。这些设计比完整的抗体分子小20倍,因此在同等质量下,潜在的中和位点多20倍,提高了局部给药的潜在疗效。对于简单得多的小蛋白来说,商品成本和规模化到非常高产量的能力应该更低,而与抗体不同,它们不需要在哺乳动物细胞中表达才能正确折叠。小尺寸和高稳定性应使其易于通过雾化直接输送到呼吸系统。免疫原性 is 任何外来分子的一个潜在问题,但对于先前表征的从头设计的小蛋白,几乎没有观察到或没有观察到免疫反应,这可能是因为高溶解度和稳定性以及小尺寸使其不太可能在树突状细胞上呈现。

参考文献

1. Yuan M,Wu NC,Zhu X, Lee CD, So RTY, Lv H, Mok CKP, Wilson IA:

A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* 2020,368(6491):630-633.

2. Case JB, Rothlauf PW, Chen RE, Liu Z, Zhao H, Kim AS, Bloyet L-M, Zeng Q, Tahan S, Droit L et al: Neutralizing antibody and soluble ACE2 inhibition of a replication-competent VSV-SARS-CoV-2 and a clinical isolate of SARS-CoV-2. *bioRxiv* 2020:2020.2005.2018.102038.

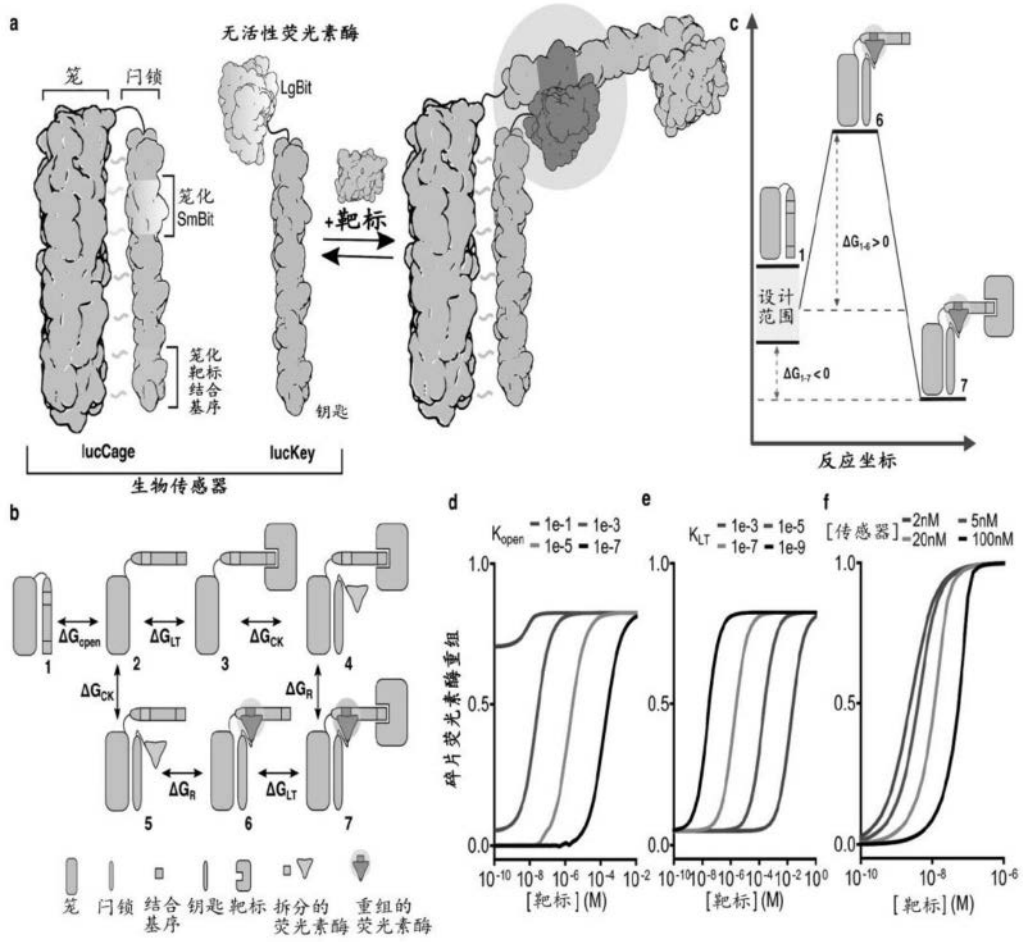


图1

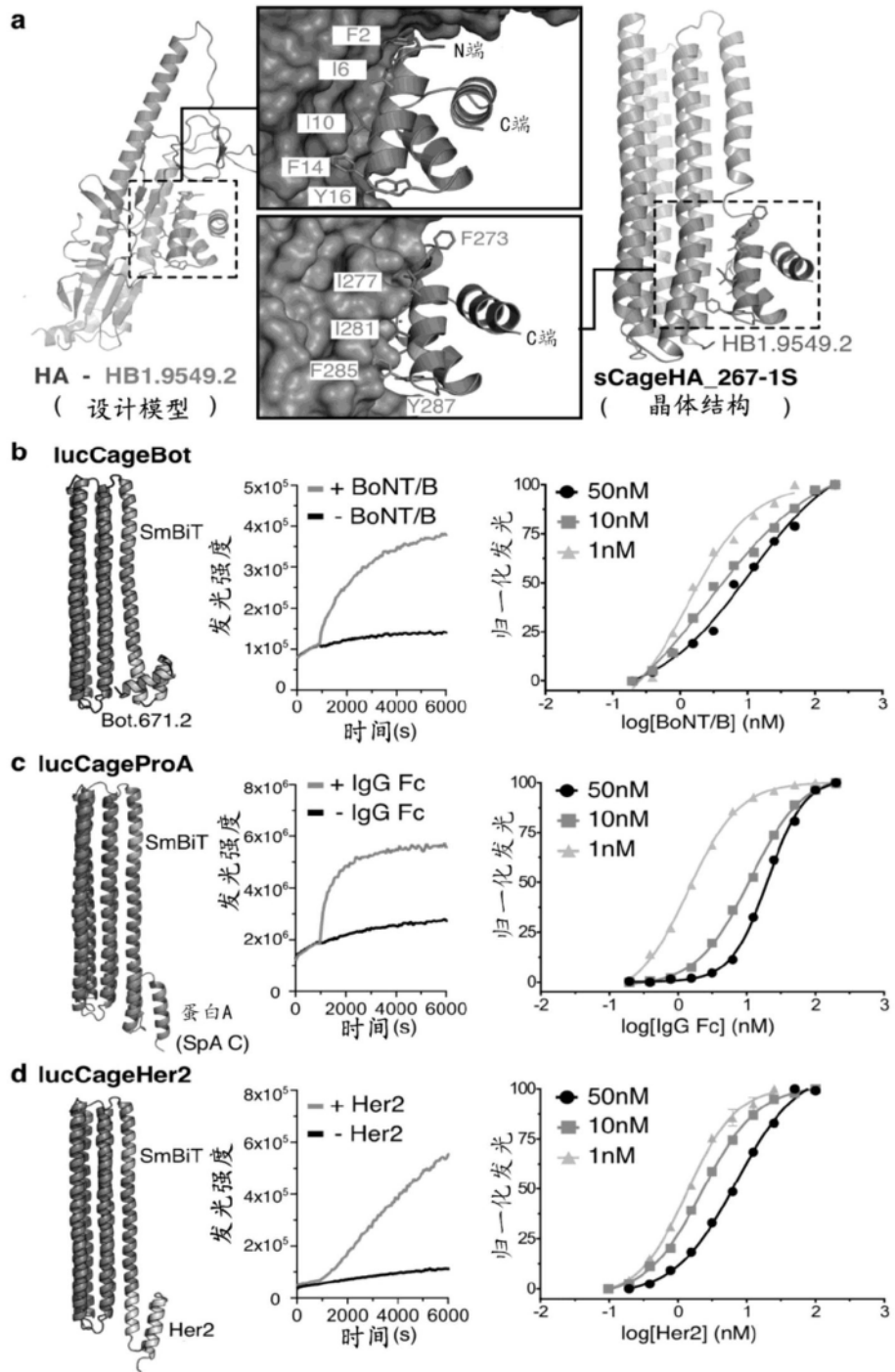


图2

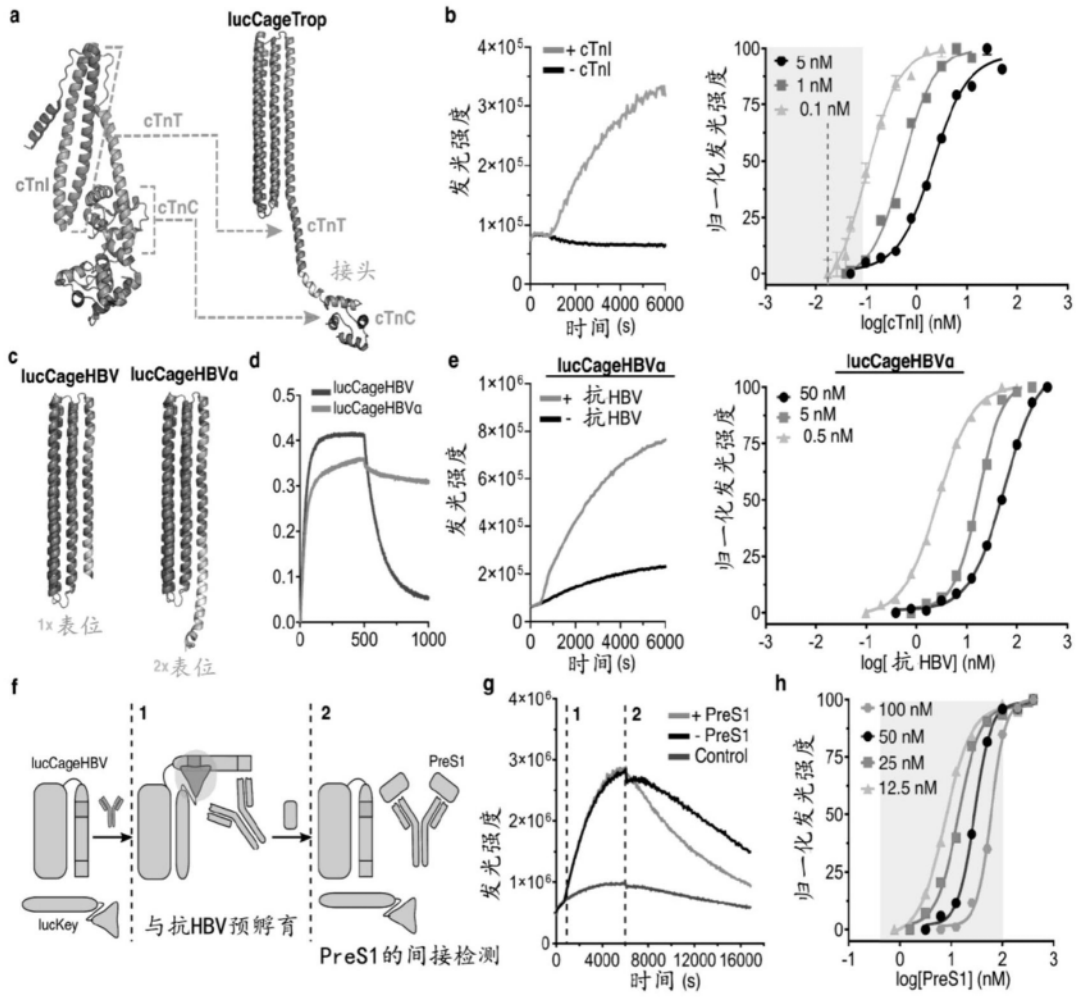


图3

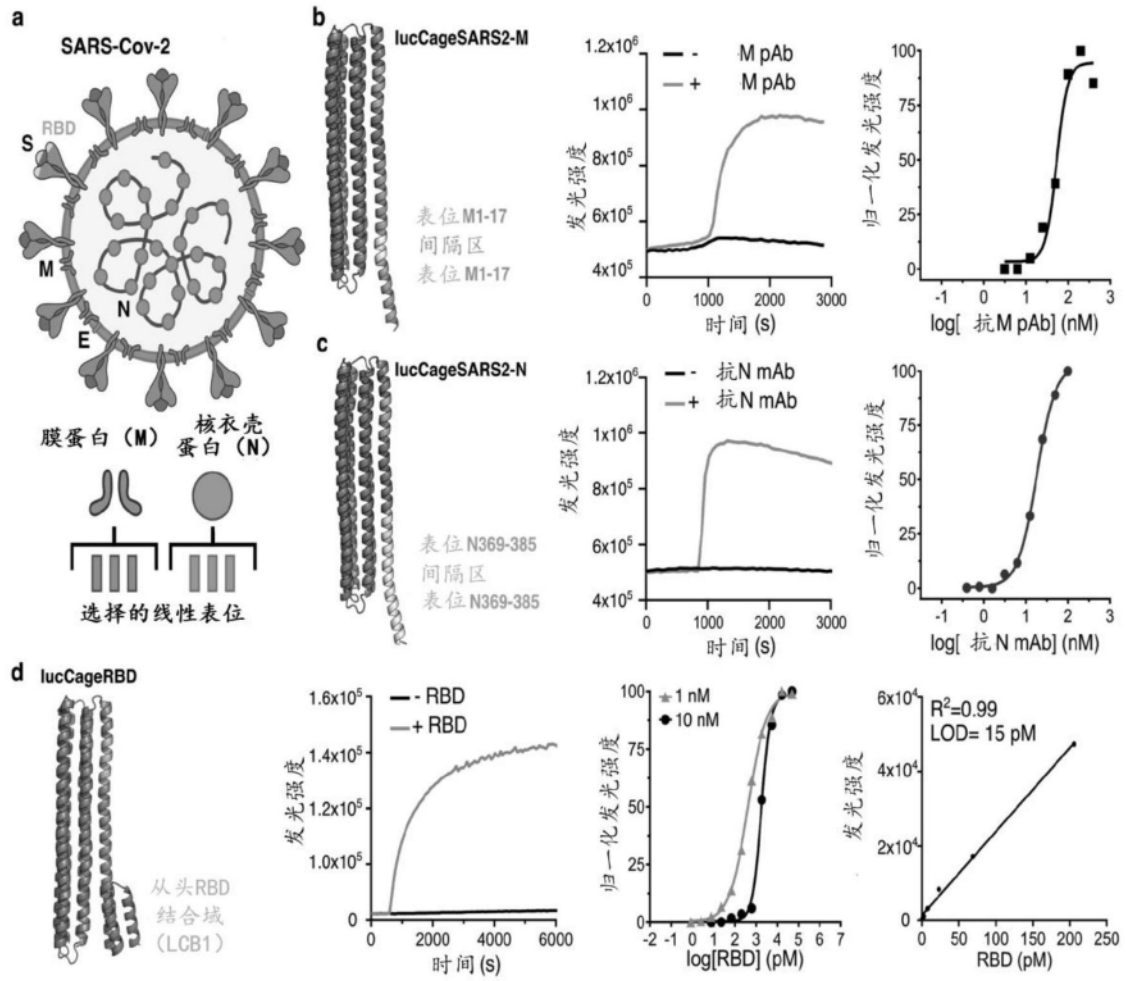


图4

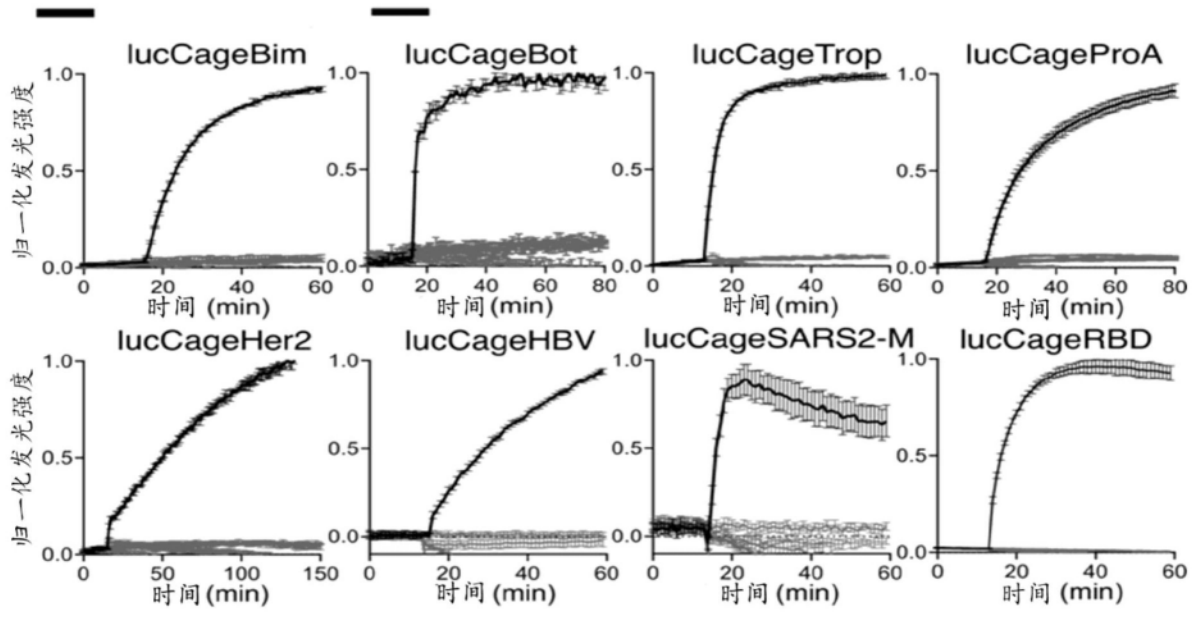


图5

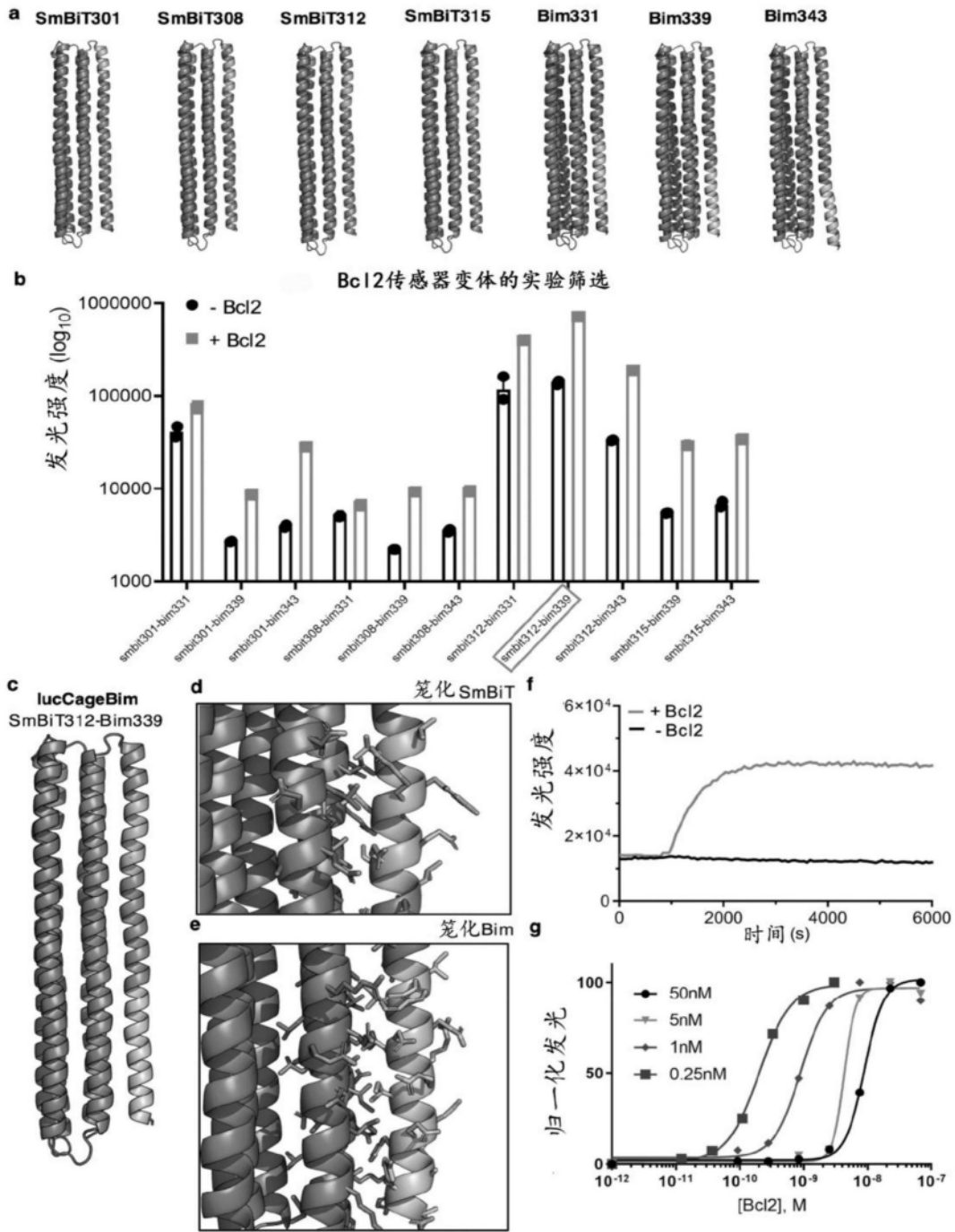


图6

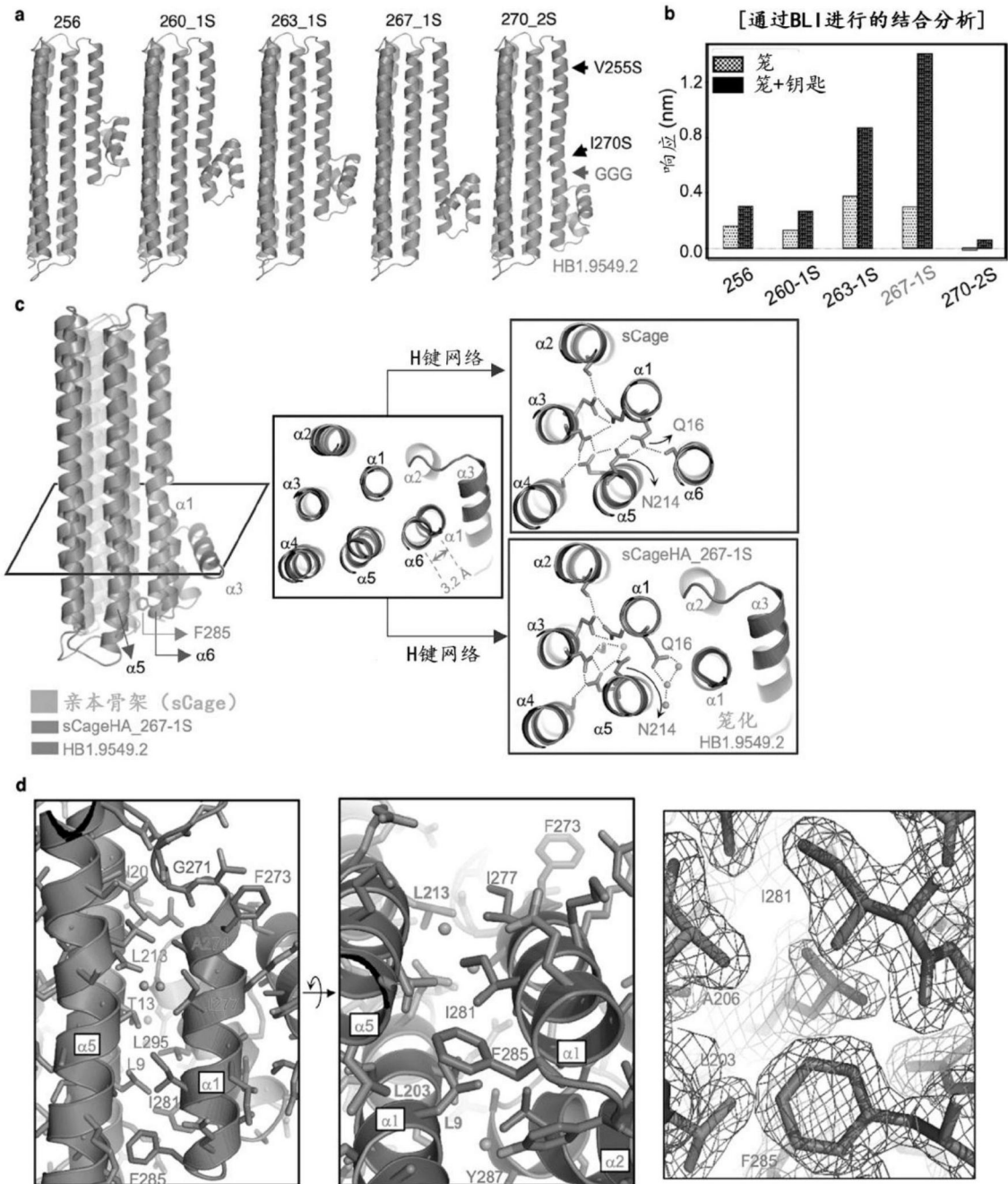


图7

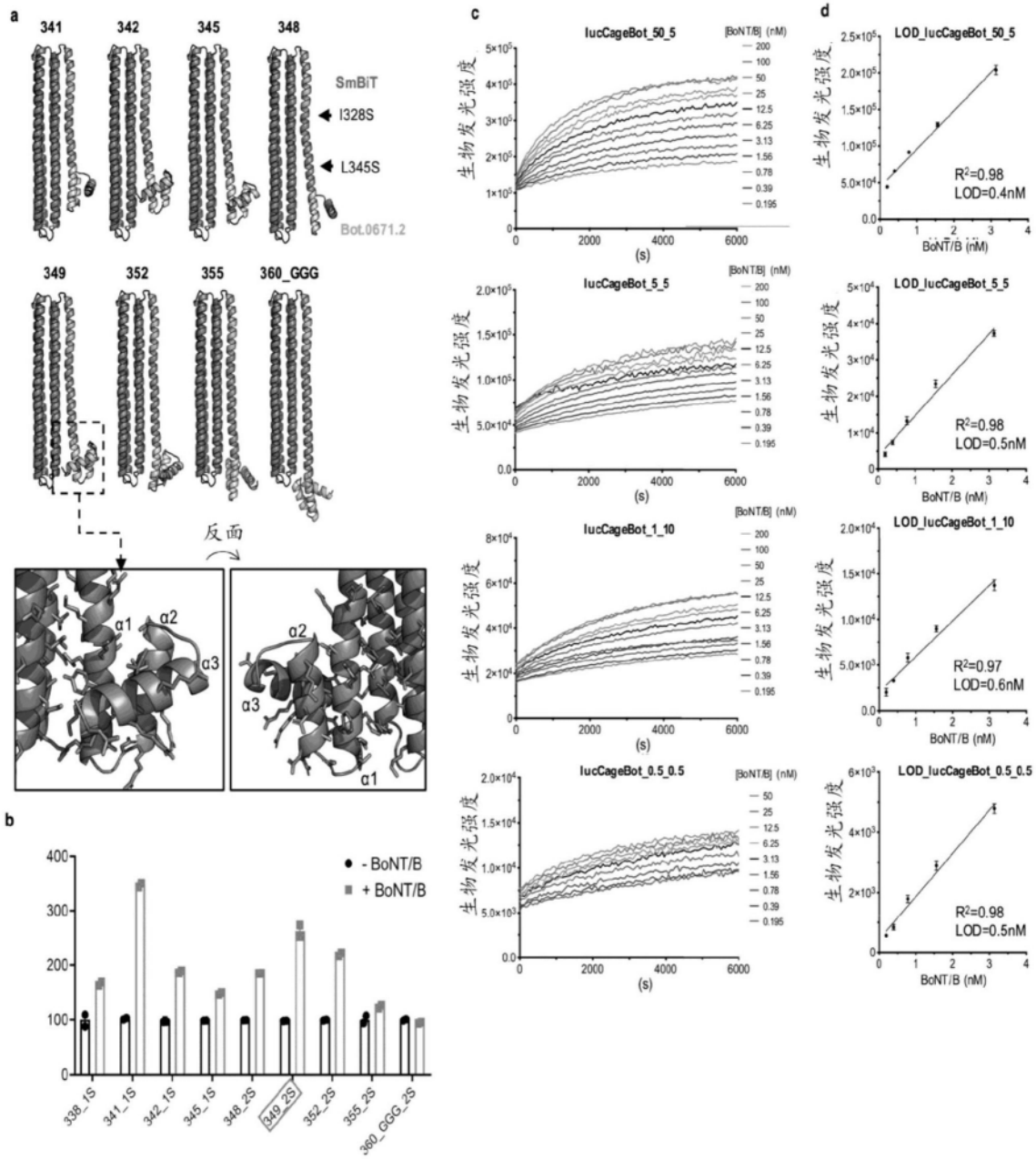


图8

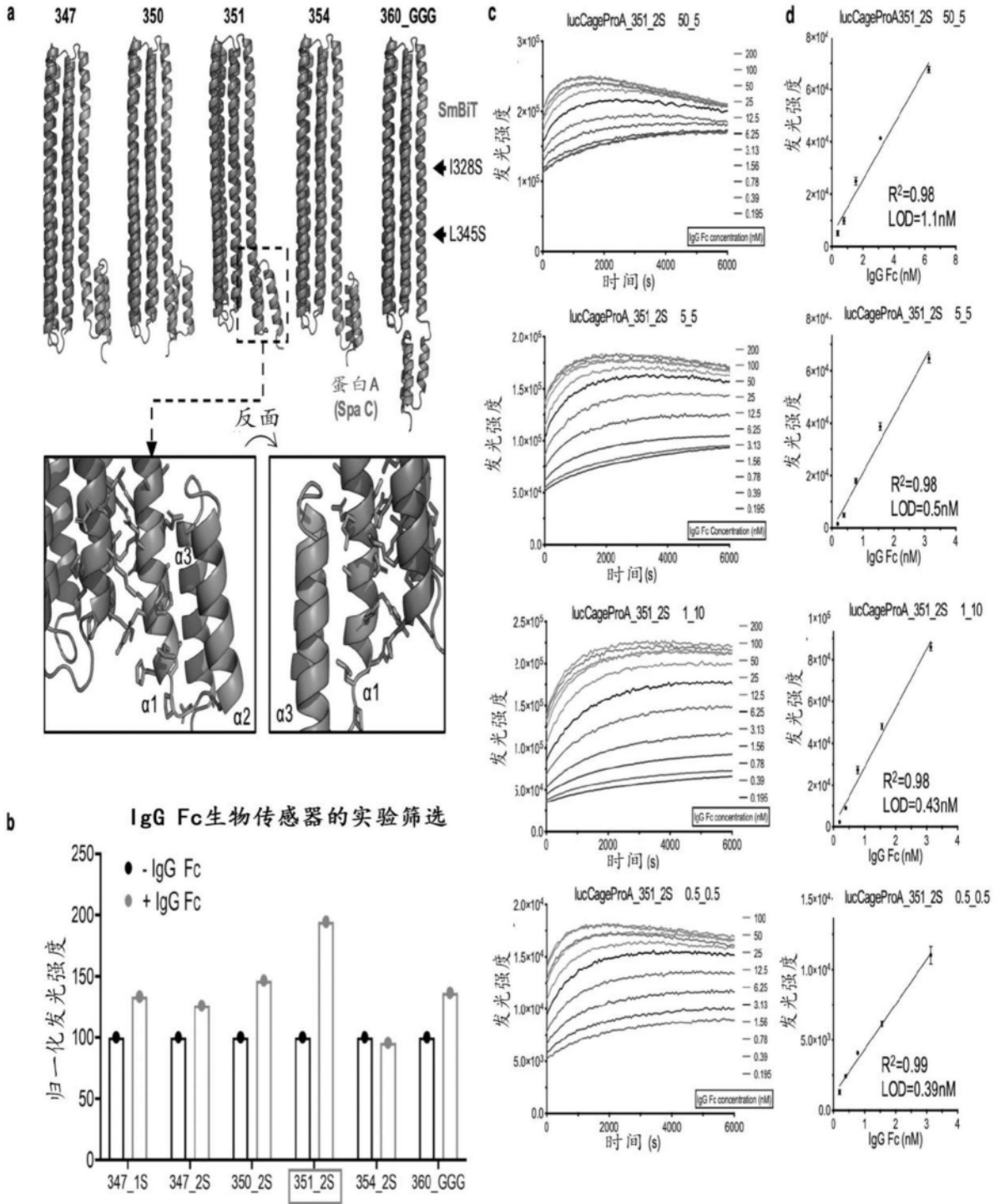


图9

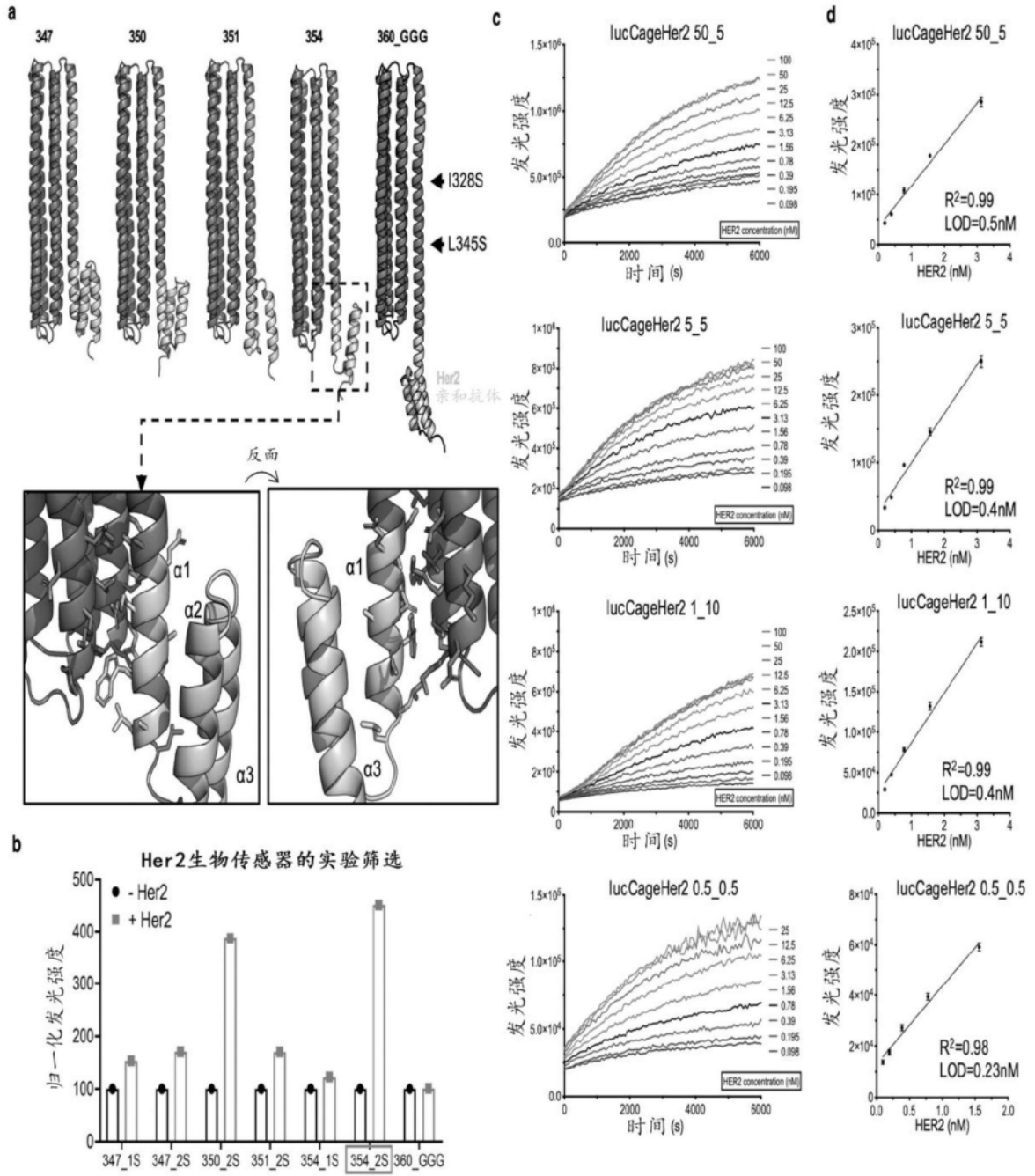


图10

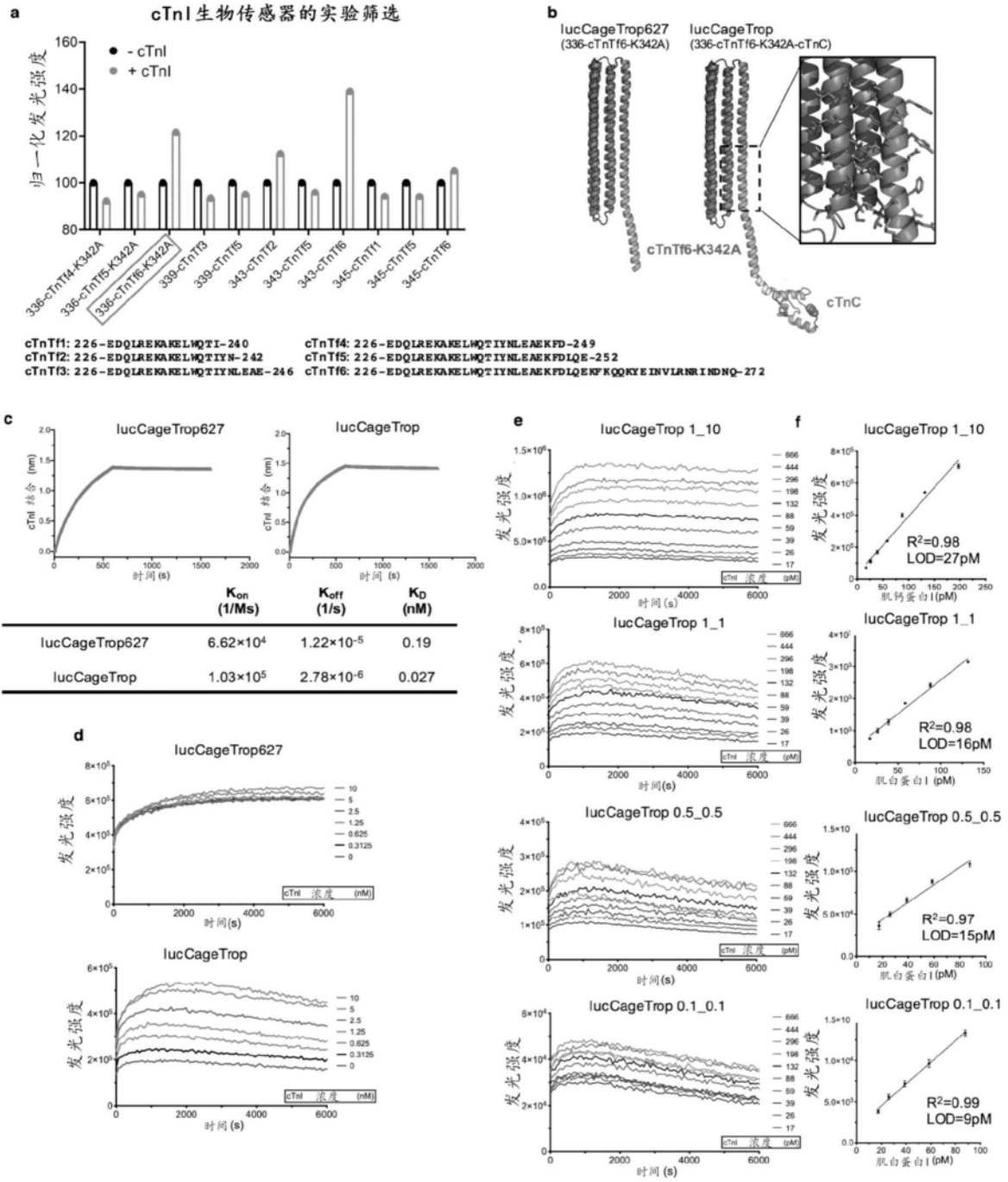


图11

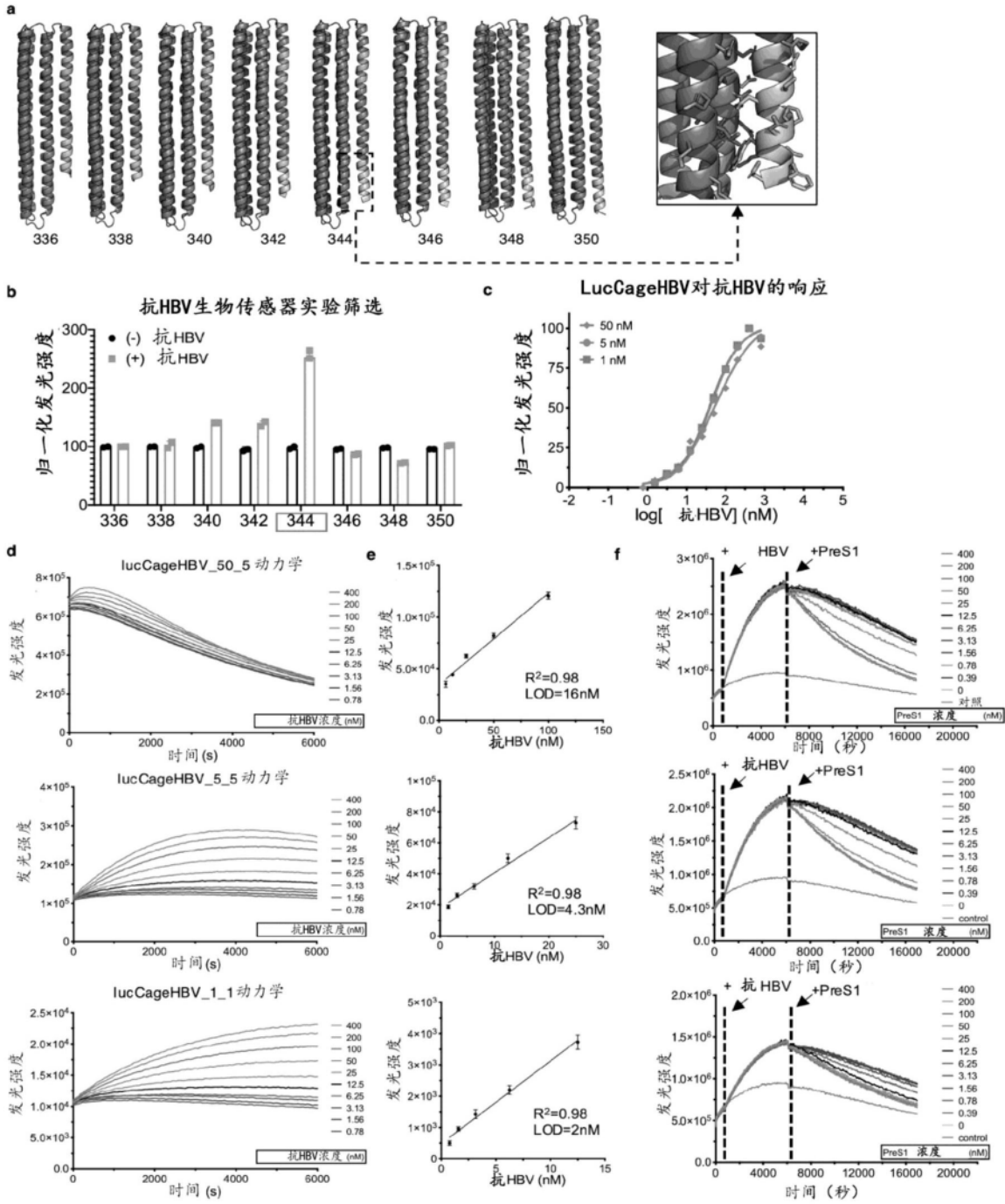


图12

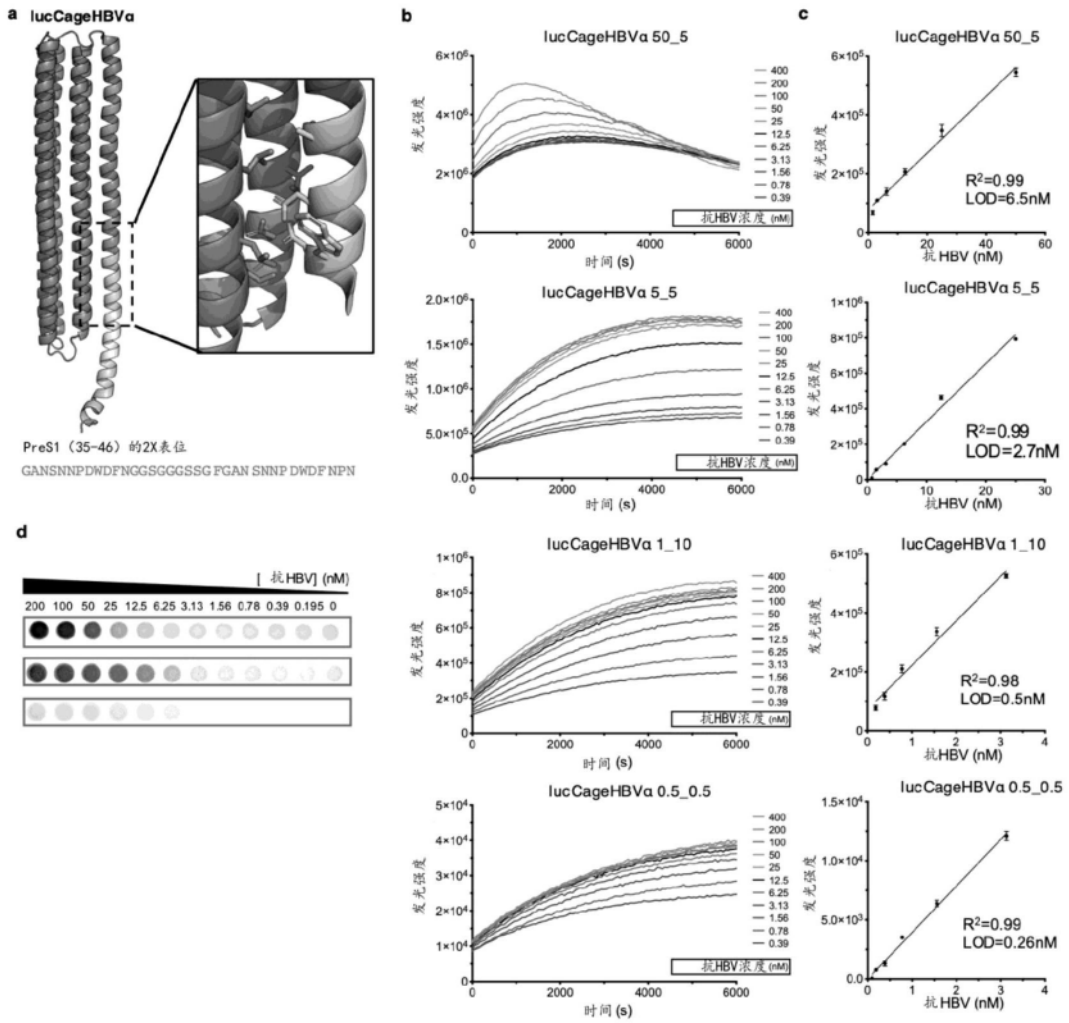


图13

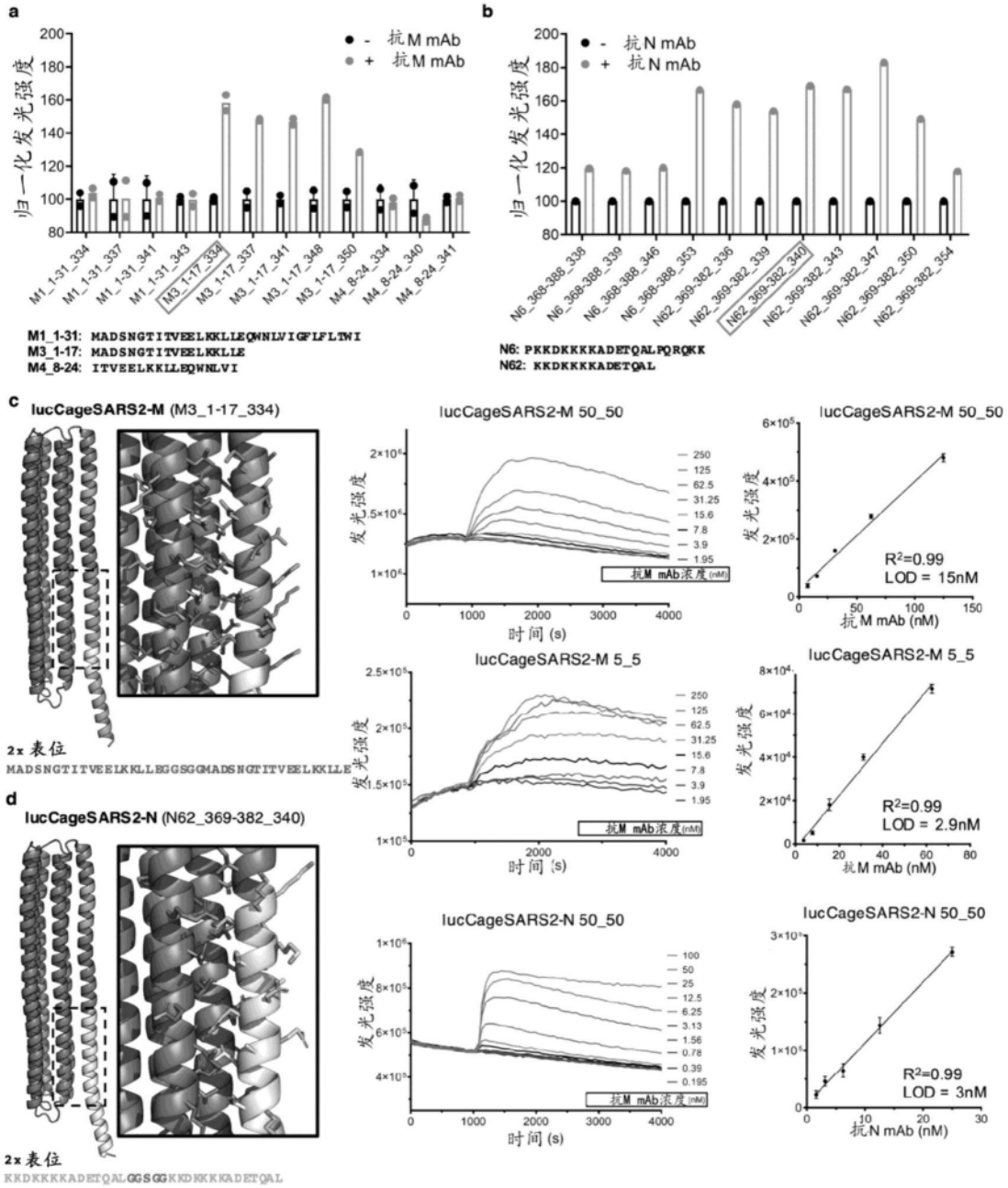


图14

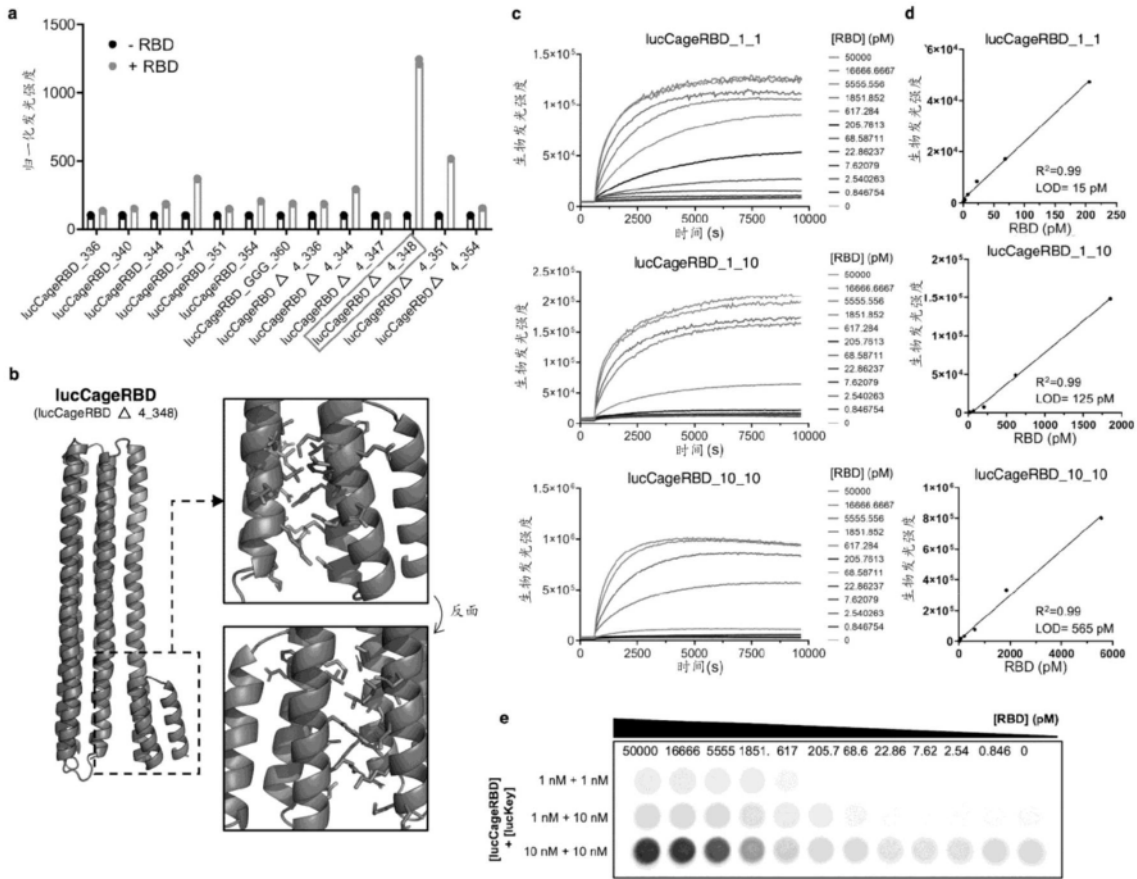


图15

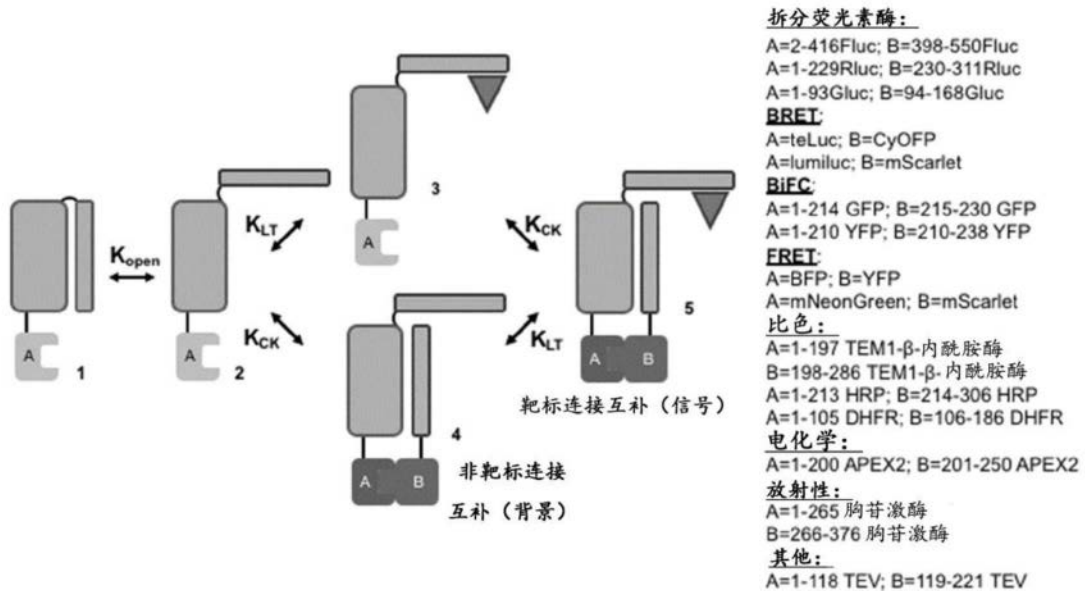


图16

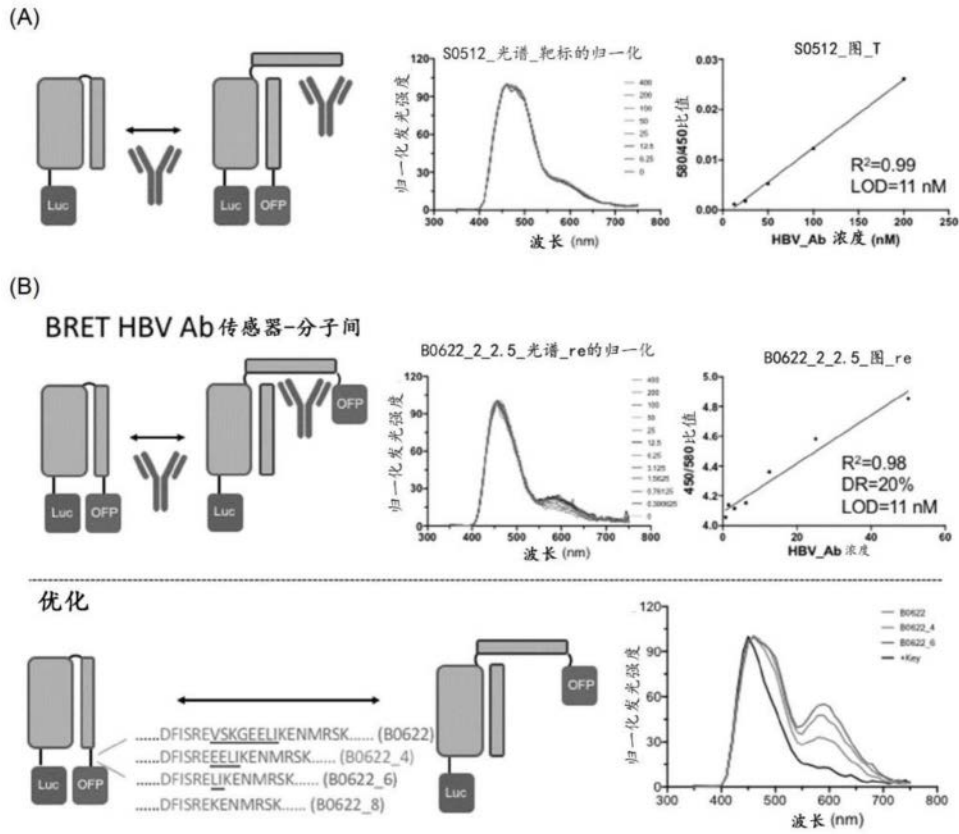


图17

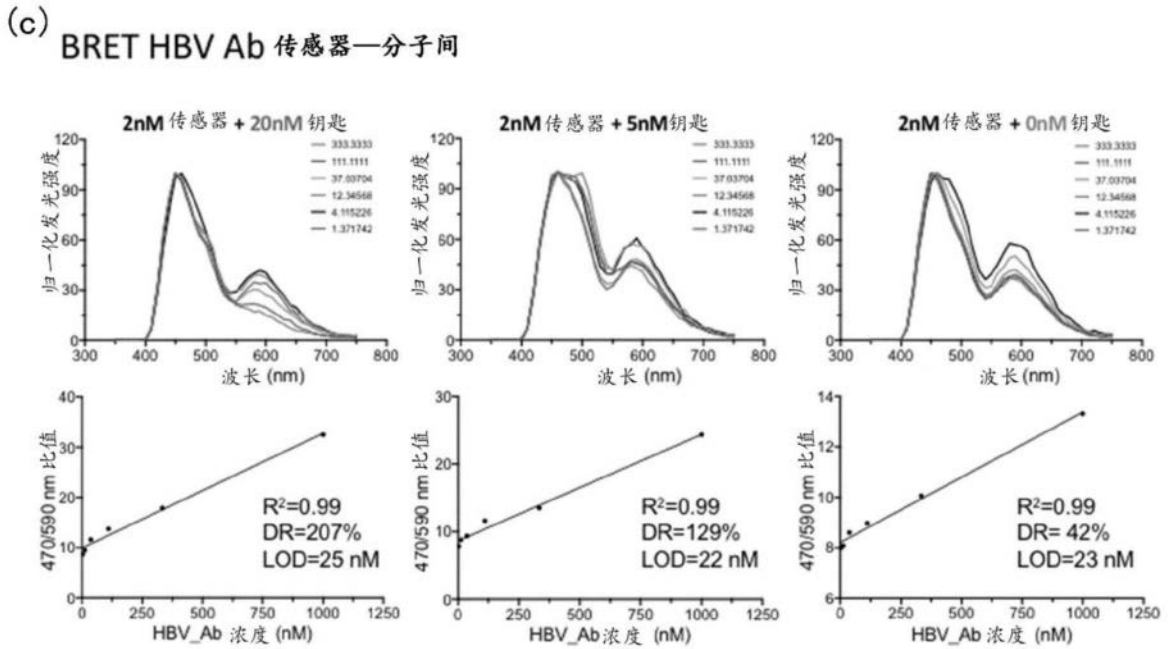


图17续

肌钙蛋白I比色传感器

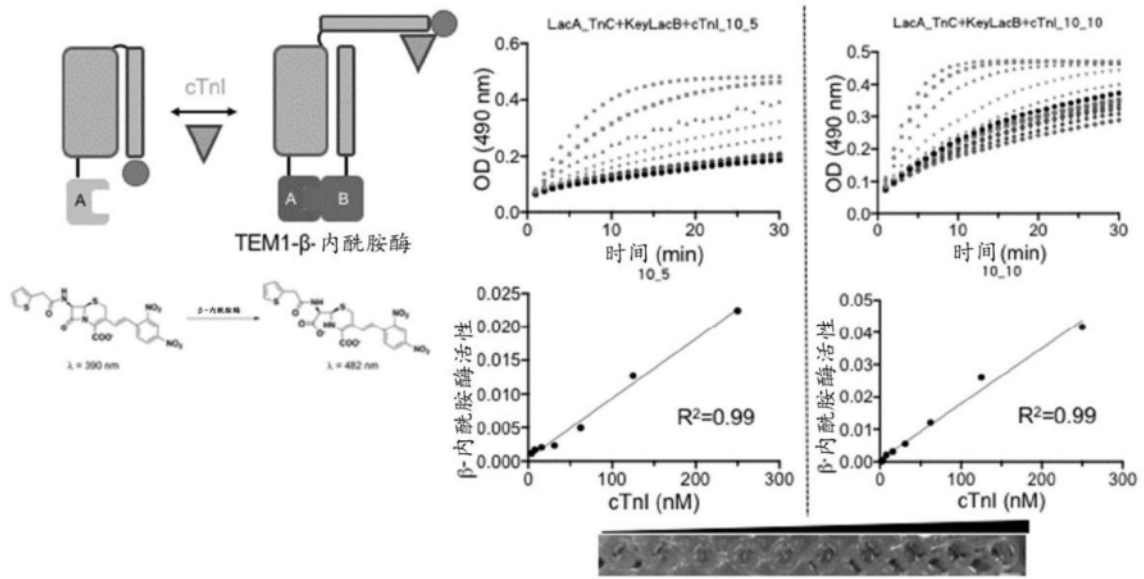
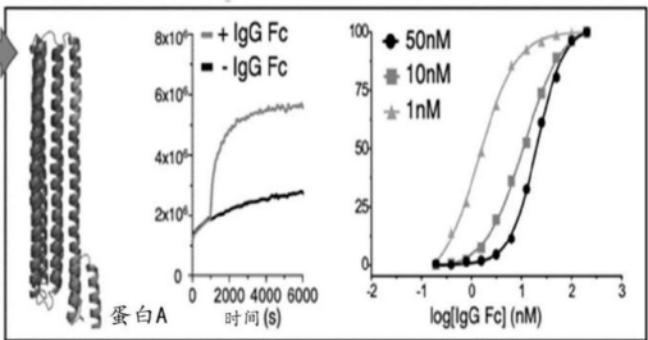


图18

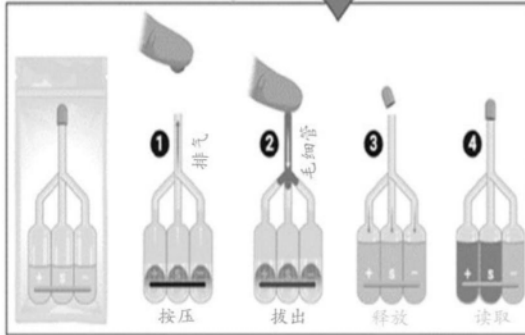
A. 试验和对照



B. IgG的阳性对照



D. POCD CoV LOCKR 诊断



C. 多重LOCKR诊断

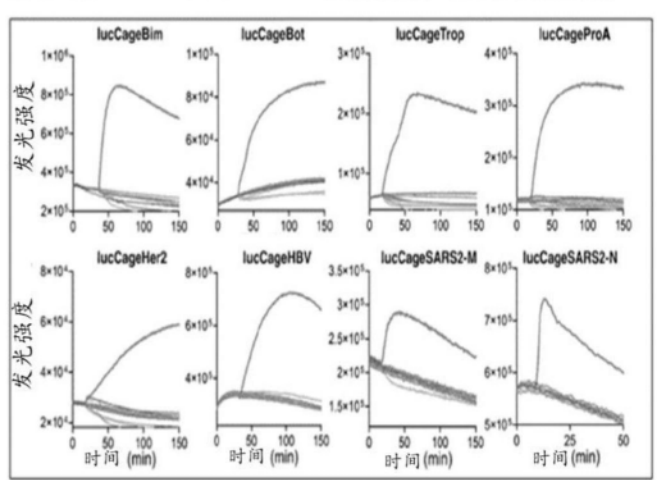


图19

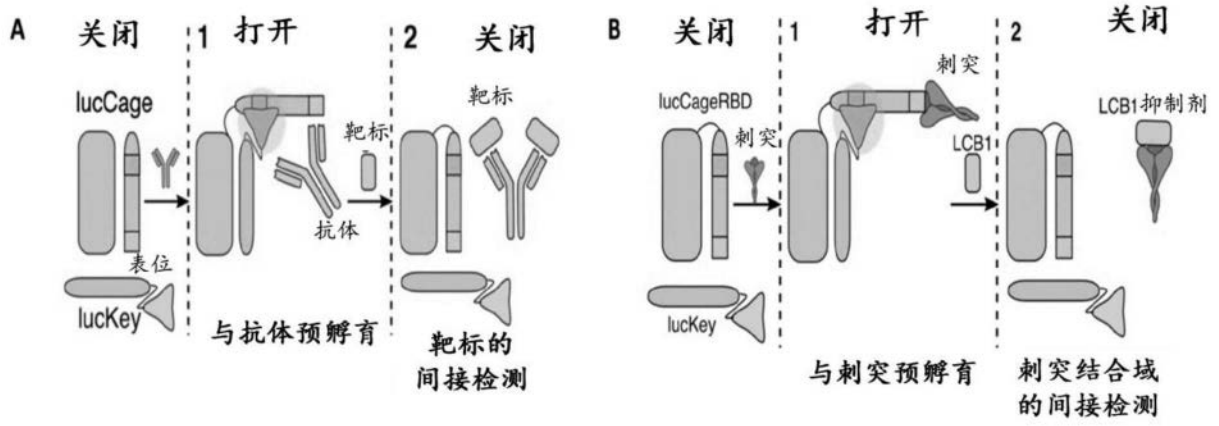


图20

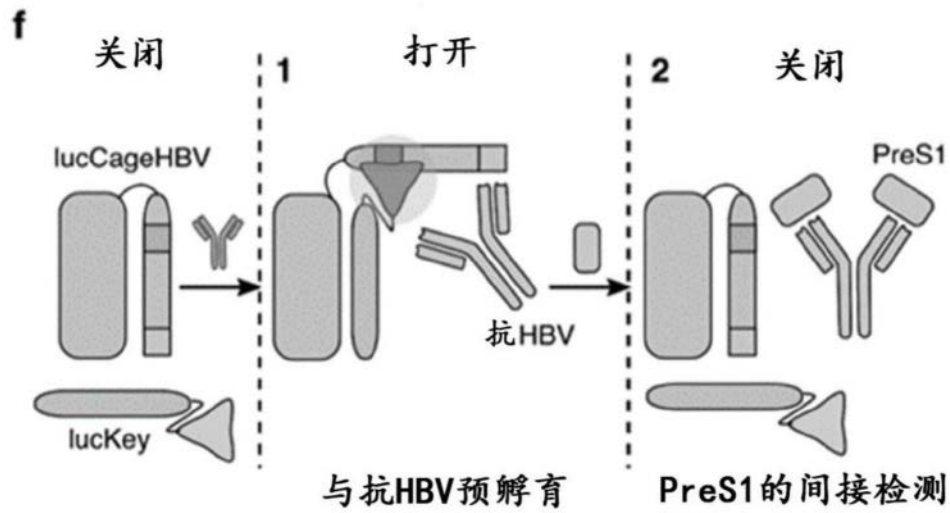


图21

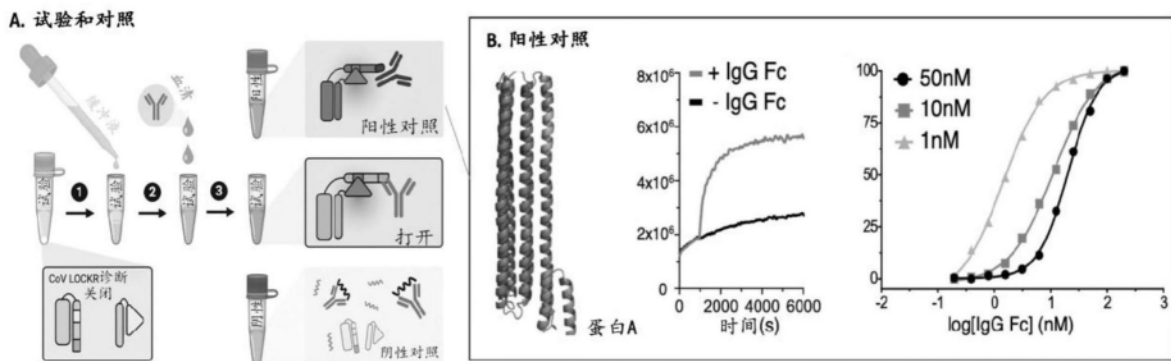


图22