

De novo design of modular and tunable protein biosensors

<https://doi.org/10.1038/s41586-021-03258-z>

Received: 13 June 2020

Accepted: 19 January 2021

Published online: 27 January 2021

 Check for updates

Alfredo Quijano-Rubio^{1,2,8}, Hsien-Wei Yeh^{1,8}, Jooyoung Park^{1,6}, Hansol Lee³, Robert A. Langan^{1,7}, Scott E. Boyken^{1,7}, Marc J. Lajoie^{1,7}, Longxing Cao¹, Cameron M. Chow¹, Marcos C. Miranda¹, Jimin Wi⁴, Hyo Jeong Hong⁴, Lance Stewart¹, Byung-Ha Oh^{1,3,8} & David Baker^{1,5,8}

Naturally occurring protein switches have been repurposed for the development of biosensors and reporters for cellular and clinical applications¹. However, the number of such switches is limited, and reengineering them is challenging. Here we show that a general class of protein-based biosensors can be created by inverting the flow of information through de novo designed protein switches in which the binding of a peptide key triggers biological outputs of interest². The designed sensors are modular molecular devices with a closed dark state and an open luminescent state; analyte binding drives the switch from the closed to the open state. Because the sensor is based on the thermodynamic coupling of analyte binding to sensor activation, only one target binding domain is required, which simplifies sensor design and allows direct readout in solution. We create biosensors that can sensitively detect the anti-apoptosis protein BCL-2, the IgG1 Fc domain, the HER2 receptor, and Botulinum neurotoxin B, as well as biosensors for cardiac troponin I and an anti-hepatitis B virus antibody with the high sensitivity required to detect these molecules clinically. Given the need for diagnostic tools to track the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)³, we used the approach to design sensors for the SARS-CoV-2 spike protein and antibodies against the membrane and nucleocapsid proteins. The former, which incorporates a de novo designed spike receptor binding domain (RBD) binder⁴, has a limit of detection of 15 pM and a luminescence signal 50-fold higher than the background level. The modularity and sensitivity of the platform should enable the rapid construction of sensors for a wide range of analytes, and highlights the power of de novo protein design to create multi-state protein systems with new and useful functions.

Protein-based biosensors have important roles in synthetic biology and clinical applications, but the design of biosensors has so far been mostly limited to reengineering natural proteins¹. Finding specific analyte-binding domains that undergo conformational changes upon binding is challenging, and even when available, extensive protein engineering efforts are generally required to effectively couple them to a reporter domain^{5,6}. It is therefore desirable to construct modular biosensor platforms that can be easily repurposed to detect different protein targets of interest. Modular systems have been developed to detect antibodies^{7–9} and small molecules^{10,11}, but general protein sensors are a bigger challenge given the great diversity of protein structures, sizes and oligomerization states, and approaches such as semisynthetic protein platforms^{12–14}, or calmodulin switches^{15,16}, usually require considerable screening to find potential candidates owing to limited predictability¹⁷.

A protein biosensor can be constructed from a system with two nearly isoenergetic states, the equilibrium between which is modulated by

the analyte being sensed. Desirable properties in such a sensor are: (i) the conformational change triggered by an analyte should be independent of the details of the analyte, so the same overall system can be used to sense many different targets; (ii) the system should be tunable so that analytes with different binding energies and different typical concentrations can be detected over a large dynamic range; and (iii) the conformational change should be coupled to a sensitive output. We hypothesized that these attributes could be attained by inverting the information flow in de novo designed protein switches in which binding to a target protein of interest is controlled by the presence of a peptide actuator². We developed a system that consisted of two protein components: first, a ‘lucCage’ that comprises a cage domain and a latch domain that contains a target-binding motif and a split luciferase fragment (small BiT (SmBiT) I14)¹⁸; and second, a ‘lucKey’ that contains a key peptide that binds to the open state of lucCage and the complementary split luciferase fragment (large BiT (LgBiT) I1S)¹⁸ (Fig. 1a). lucCage has

¹Department of Biochemistry, Institute for Protein Design, University of Washington, Seattle, WA, USA. ²Department of Bioengineering, University of Washington, Seattle, WA, USA. ³Department of Biological Sciences, KAIST Institute for the Biocentury, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. ⁴Department of Systems Immunology, College of Biomedical Science, Kangwon National University, Chuncheon, South Korea. ⁵Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. ⁶Present address: Sana Biotechnology, Inc, Seattle, WA, USA. ⁷Present address: Outpace Bio, Inc., Seattle, WA, USA. ⁸These authors contributed equally: Alfredo Quijano-Rubio, Hsien-Wei Yeh. [✉]e-mail: bhoh@kaist.ac.kr; dabaker@uw.edu

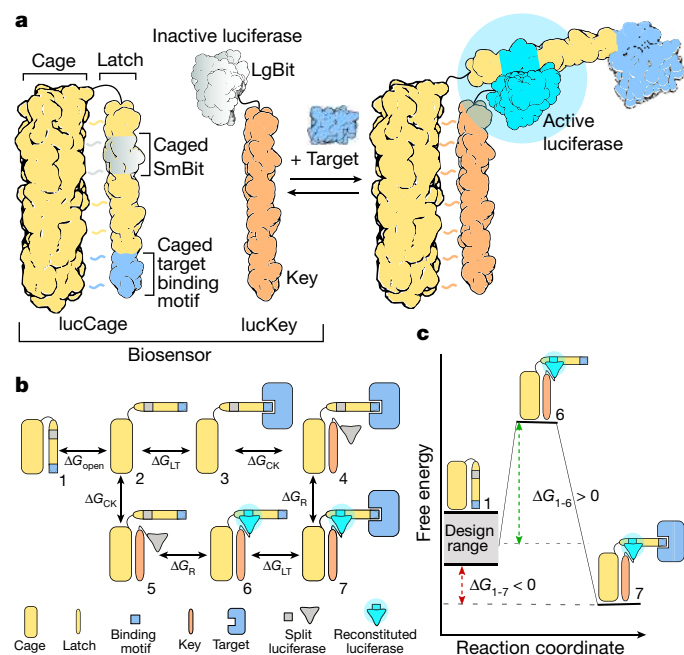


Fig. 1 | De novo design of multi-state biosensors. **a**, Sensor schematic mechanism. The closed form of lucCage (left) cannot bind to lucKey, thus preventing the split luciferase SmBiT fragment from interacting with LgBiT. The open form (right) can bind to both the target and the key, enabling the reconstitution of SmBiT and LgBiT for luciferase activity. **b**, Thermodynamics of biosensor activation. The free energy cost (ΔG_{open}) of the transition from closed cage (species 1) to open cage (species 2) disfavours the association of key (species 5) and reconstitution of luciferase activity (species 6) in the absence of target. In the presence of the target, the combined free energies of target binding ($2 \rightarrow 3$; ΔG_{LT}), key binding ($3 \rightarrow 4$; ΔG_{CK}), and SmBiT–LgBiT association ($4 \rightarrow 7$; ΔG_{R}) overcome the unfavourable ΔG_{open} , driving the opening of the lucCage and reconstitution of luciferase activity. **c**, Thermodynamics of biosensor design. The designable parameters are ΔG_{open} and ΔG_{CK} ; ΔG_{R} is the same for all targets, and ΔG_{LT} is pre-specified for each target. For sensitive but low background analyte detection, ΔG_{open} and ΔG_{CK} must be tuned such that the closed state (species 1) is substantially lower in free energy than the open state (species 6) in the absence of target, but higher in free energy than the open state in the presence of target (species 7).

two states: a closed state, in which the cage domain binds to the latch and sterically occludes the binding motif from binding the target and SmBiT from combining with LgBiT to reconstitute luciferase activity, and an open state, in which these binding interactions are not blocked and lucKey can bind to the cage domain. The association of lucKey with lucCage results in the reconstitution of luciferase activity (Fig. 1a, right). The thermodynamics of the system are tuned such that the binding free energy of lucKey to lucCage (ΔG_{CK}) is insufficient to overcome the free energy cost of lucCage opening (ΔG_{open}) in the absence of target ($\Delta G_{\text{open}} - \Delta G_{\text{CK}} \gg 0$), but in the presence of the target, the additional binding free energy of the latch to the target (ΔG_{LT}) drives latch opening and luciferase reconstitution ($\Delta G_{\text{open}} - \Delta G_{\text{CK}} - \Delta G_{\text{LT}} < 0$) (Fig. 1b, c). This system satisfies properties (i) and (ii) above, as a wide range of binding activities can be caged, and because the switch is thermodynamically controlled, the lucKey and target binding energies can be adjusted to achieve activation at the relevant target concentrations. Because lucKey and lucCage are always the same, the system is modular—the same molecular association can be coupled to the binding of many different targets. Bioluminescence provides a rapid and sensitive readout of the analyte-driven lucCage–lucKey association, satisfying property (iii).

The states of this biosensor system are in thermodynamic equilibrium, with the tunable parameters ΔG_{open} and ΔG_{CK} governing the populations of the possible species, along with the free energy of association

of the analyte to the binding domain ΔG_{LT} (Fig. 1b). We simulated the dependence of the sensor system on ΔG_{open} (Extended Data Fig. 1a), ΔG_{LT} (Extended Data Fig. 1b), and the concentration of analyte and the sensor components (Extended Data Fig. 1c, d). The sensitivity of analyte detection is a function of ΔG_{LT} , with a lower limit of roughly one-tenth of the dissociation constant (K_d) for analyte binding (Extended Data Fig. 1b). Above this lower limit, varying the concentration of lucCage and lucKey enables the system to respond to different ranges of target concentration (Extended Data Fig. 1c, d). Sensitivity can be further modulated by tuning the strength of the intramolecular cage–latch interaction and the intermolecular cage–key interaction (ΔG_{open} and ΔG_{CK} , respectively); for example, too tight cage–latch interaction results in a low signal in the presence of target, and too weak an interaction results in a high background signal in the absence of target (Extended Data Fig. 1a, e). Our design strategy aims to find this balance by modulating ΔG_{open} and ΔG_{CK} by varying the length of the latch (and key) helix and by introducing either favourable hydrophobic or unfavourable buried polar interactions at the cage–latch or cage–key interfaces² (Extended Data Fig. 1f, g).

Designing tunable lucCage sensors

To design sensors based on these principles, we developed a ‘GraftSwitchMover’ Rosetta-based method to identify placements of target binding peptides within the latch such that the resulting protein is stable in the closed state and the interactions with the target are blocked (Supplementary Methods). As a first test, we grafted the SmBiT peptide and the BIM peptide in the closed state of the previously described optimized asymmetric LOCKR switch² (Extended Data Fig. 2). SmBiT adopts a β -strand conformation within the luciferase holoenzyme, but we assumed that it could adopt a helical secondary structure in the context of the helical bundle scaffold, because secondary structure can be context-dependent¹⁹. We sampled different placements for the two peptide sequences across the latch, selected the lowest energy solutions (Extended Data Fig. 2a) and expressed 12 designs in *Escherichia coli*. We mixed the designs with lucKey in a 1:1 ratio, then added BCL-2, which binds to BIM with nanomolar affinity²⁰, and observed a rapid increase in luminescence (Extended Data Fig. 2b, f; we refer to the best of these as ‘lucCageBIM’), which demonstrates that the LOCKR actuator² operated in reverse can function as a biosensor. The detection range of the analyte could be tuned by varying the concentration of the sensor (lucCage plus lucKey) (Extended Data Fig. 2g), as anticipated in our model simulations (Extended Data Fig. 1c). lucCageBIM has SmBiT at position 312 in the latch (SmBiT312) (Extended Data Fig. 2d); the cage with this placement (lucCage) was used as the base scaffold for the biosensors described below.

lucCage sensors with miniprotein sensing domains

We next investigated the incorporation of a range of binding modalities for analytes of interest within lucCage by developing methods for computationally caging target-binding proteins, rather than peptides, in the closed state (Supplementary Methods). As a test case, we caged the de novo designed influenza A H1 haemagglutinin (HA)²¹ binding protein HB1.9549.2 into a shortened version of the LOCKR switch²² (sCage), optimized to improve stability and facilitate crystallization efforts (Fig. 2a). Two out of the five designs were functional, and bound HA in the presence but not the absence of key (Extended Data Fig. 3b). The crystal structure of the best design, sCageHA_267-1S, determined to 2.0 Å resolution (Supplementary Table 1, Protein Data Bank (PDB) code 7CBC), showed that all HA-binding interface residues except one (Phe273) interact with the cage domain (blocking binding of the latch to the target) as intended by design (Fig. 2a, Extended Data Fig. 3a–c).

With this structural validation of the design concept, we next sought to develop sensors for Botulinum neurotoxin B (BoNT/B),

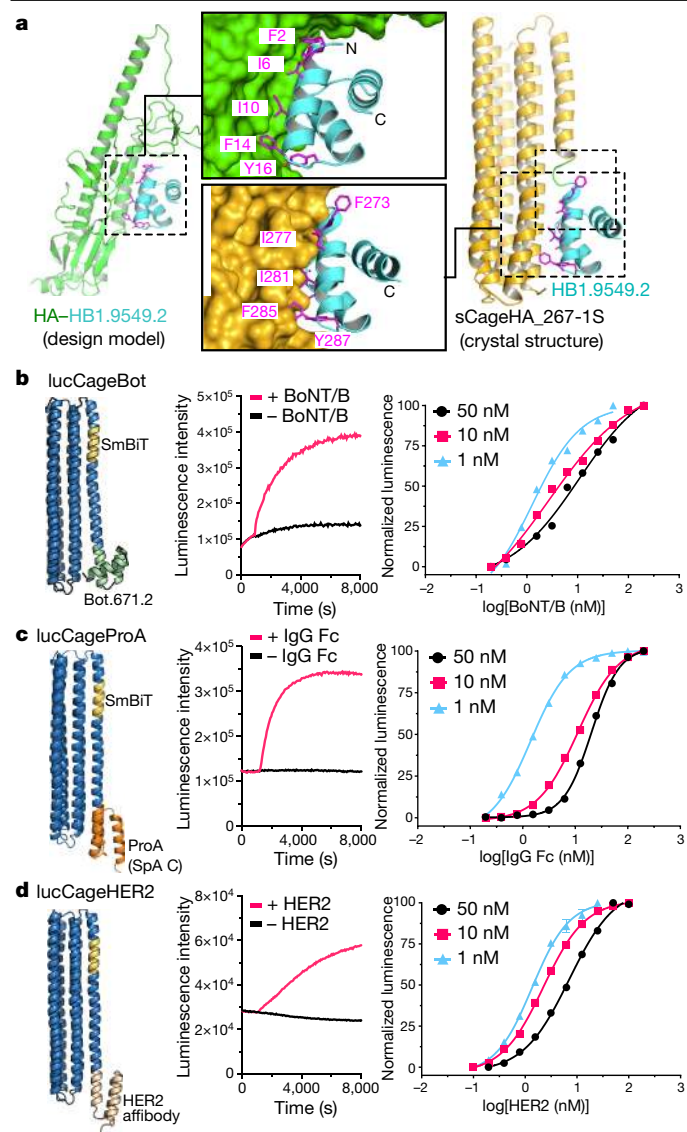


Fig. 2 | Design and characterization of de novo biosensors incorporating small proteins as sensing domains. **a**, Structural validation of sCageHA_{267-1S}, which cages a designed influenza binding protein inside a LOCKR switch. Left, design model of the de novo binder HB1.9549.2 (cyan ribbon) bound to the stem region of influenza haemagglutinin (HA, green ribbon)²¹. Right, crystal structure (PDB code 7CBC) of sCageHA_{267-1S}, comprising HB1.9549.2 (cyan) grafted into a shortened and stabilized version of the LOCKR switch²² (sCage, yellow ribbon). Middle, all residues of HB1.9549.2 involved in binding to HA (magenta, top) except for F273 are buried in the closed state of the switch (bottom). The magenta labels indicate the same set of amino acids in the two panels (for example, F2 in the top panel corresponds to F273 in the bottom panel). **b–d**, Functional characterization of lucCageBot (**b**), lucCageProA (**c**) and lucCageHER2 (**d**). Left, structural models incorporate a de novo designed binder for BoNT/B (Bot.671.2)²¹ (**b**) the C domain of protein A (SpA C)²³ (**c**) or a HER2-binding affibody²⁴ (**d**) into lucCage (blue ribbon) with caged SmBIT fragment (gold ribbon). Middle, measurement of luminescence intensity after the addition of 50 nM of analyte (BoNT/B (**b**), IgG Fc (**c**) or HER2 (**d**)) to a mixture of 10 nM of each lucCage and 10 nM of lucKey. Right, detection over a wide range of analyte concentrations by changing the biosensor (lucCage plus lucKey) concentration (coloured lines). All experiments were performed in triplicate, representative data are shown, and data are mean ± s.d.

the immunoglobulin Fc domain and the HER2 receptor. We grafted a de novo designed binder for Botulinum neurotoxin (Bot.0671.2)²¹, the C domain of the generic antibody-binding protein A²³ and a

HER2-binding affibody²⁴ into lucCage. After screening a few designs for each target (Extended Data Figs. 4, 5), we obtained highly sensitive lucCages (lucCageBot, lucCageProA and lucCageHER2) that can detect BoNT/B (Fig. 2b, Extended Data Fig. 4), human IgG Fc domain (Fig. 2c, Extended Data Fig. 5a–d), and HER2 receptor (Fig. 2d, Extended Data Fig. 5e–h), respectively, demonstrating the modularity of the platform. The designed sensors respond within minutes after the addition of the target, and their sensitivity can be tuned by changing the concentration of lucCage and lucKey (Fig. 2). With further development, these sensors could enable the rapid and low-cost detection of botulinum neurotoxins in the food industry²⁵, and detection of serological levels of soluble HER2 (>15 ng ml⁻¹; within the detection range of lucCageHER2) associated with metastatic breast cancer²⁶.

lucCage sensor for cardiac troponin

We next designed sensors for cardiac troponin I, which is the standard early diagnostic biomarker for acute myocardial infarction²⁷. We took advantage of the high-affinity interactions between cardiac troponins T, C and I (cTnT, cTnC and cTnI, respectively) (Fig. 3a) and designed 11 biosensor candidates by inserting 6 truncated cTnT sequences at different latch positions (Extended Data Fig. 6a). The best candidate, lucCageTrop627, was able to detect cTnI but not at sufficiently low levels for clinical use as the rule-in and rule-out levels of cTnI assay for the diagnosis of patients with acute myocardial infarction are in the low picomolar range²⁷. Because the limit of detection (LOD) of our sensor platform is about $0.1 \times K_d$ of the latch–target affinity (K_{LT}), we sought to improve the sensitivity of lucCageTrop627 by increasing the cTnI binding affinity. We fused cTnC to the C terminus of the sensor to take advantage of the high-affinity interaction between the three cardiac troponins (Extended Data Fig. 6b–d). The resulting sensor, lucCageTrop, has a single-digit picomolar LOD that is suitable for the quantification of clinical samples (Fig. 3b, Extended Data Fig. 6e, f).

lucCage sensors for HBV and SARS-CoV-2 antibodies

The detection of specific antibodies is important for monitoring the spread of a pathogen in a population²⁸, the success of vaccination²⁹, and levels of therapeutic antibodies⁹. To adapt our system for serological antibody analyses, we sought to incorporate linear epitopes recognized by the antibodies of interest into lucCage. We first developed a sensor for antibodies against the preS1 domain of the hepatitis B virus (HBV) surface protein L³⁰. The best of eight designs tested, termed 'lucCageHBV', had an approximately 150% increase in luciferase activity after the addition of the anti-HBV antibody HzKR127-3.2³¹ (Extended Data Fig. 7a–d). To further improve the dynamic range and LOD of lucCageHBV (Extended Data Fig. 7e), we introduced a second copy of the peptide at the end of the latch to increase latch affinity with the bivalent antibody (K_{LT}) (Fig. 3c, d). The resulting design, termed 'lucCageHBV α ', had a LOD of 260 pM and a dynamic range of 225% (Fig. 3e, Extended Data Fig. 7g–i), with a luminescence intensity easily detectable with a camera (Extended Data Fig. 7j). Because the concentrations of most therapeutic antibodies in serum are in the low micromolar to nanomolar range⁹, this platform should be useful for monitoring the concentrations of therapeutic antibodies in circulation³².

We next sought to use the lucCageHBV sensor to detect HBV surface antigen. Because our sensors are under thermodynamic control, we hypothesized that the pre-assembled sensor–antibody complex would re-equilibrate in the presence of the target HBV surface antigen protein preS1, with antibody redistributing to bind free preS1 instead of the epitope on lucCageHBV (Fig. 3f). The luminescence of the lucCageHBV plus HzKR127-3.2 mixture decreased shortly after the addition of the preS1 domain (Fig. 3g); the sensitivity of this readout enabled quantification of the preS1 concentration in the clinically relevant range³³ (Fig. 3h, Extended Data Fig. 7f).

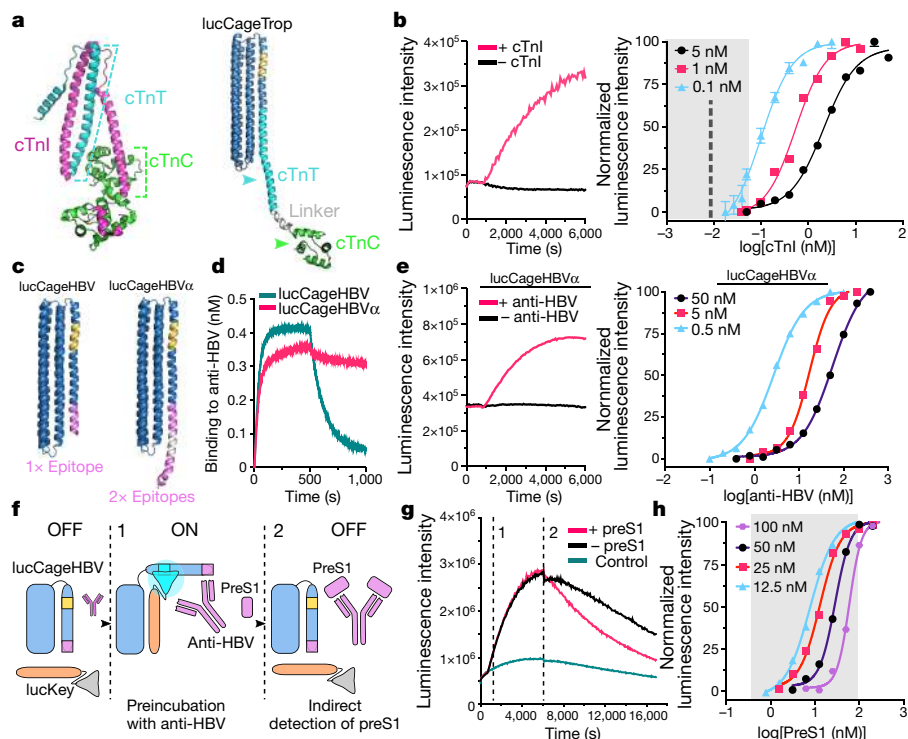


Fig. 3 | Design and characterization of biosensors for cTnI and an anti-HBV antibody. **a**, Design of cTnI sensor. Left, structure of cardiac troponin (PDB code 4Y99). cTnT, cTnC and cTnI are shown in cyan, green and magenta, respectively. Right, design model of lucCageTrop. **b**, Left, luminescence signal increases after the addition of 1 nM cTnI to 0.1 nM lucCageTrop plus lucKey. Right, wide detection range accessible by changing the concentration of the sensor components (coloured lines). Grey area indicates the cTnI concentration range relevant to the diagnosis of acute myocardial infarction²⁷; the dotted line indicates the clinical cut-off for acute myocardial infarction defined by the World Health Organization (WHO) (0.6 ng ml^{-1} , 25 pM). **c**, HBV sensor design models (gold, SmBit; grey, linker; magenta, HBV preS1 epitope). **d**, lucCageHBV α with two epitope copies has higher affinity by biolayer

interferometry for the anti-HBV antibody HzKR127-3.2 ($K_d = 0.68 \text{ nM}$) than lucCageHBV ($K_d = 20 \text{ nM}$). **e**, Left, luminescence signal increases after the addition of 50 nM anti-HBV antibody to 1 nM lucCageHBV α plus lucKey. Right, sensitive anti-HBV antibody detection over a wide concentration range. **f**, Mechanism for the detection of preS1 using lucCageHBV. **g**, Kinetics of bioluminescence after the addition of the anti-HBV antibody ('1') and subsequently preS1 ('2'), which decreases bioluminescence by competing with the sensor for the antibody. **h**, The detection of preS1 can be achieved over the relevant post-HBV infection concentration levels (grey area) by varying the concentration of antibody (indicated by coloured labels). All experiments were performed in triplicate, representative data are shown, and data are mean \pm s.d.

The COVID-19 pandemic has created an urgent need for diagnostic tools for both the SARS-CoV-2 virus and antiviral antibodies³. To design sensors for anti-SARS-CoV-2 antibodies, we first identified from the literature highly immunogenic linear epitopes in the proteomes of SARS-CoV^{34,35} and SARS-CoV-2³⁶ that are not present in 'common' strains of *Coronaviridae*. Among these, we focused on two epitopes in the membrane (M) and nucleocapsid (N) proteins that are recognized by sera from patients with SARS and COVID-19^{35,36} and for which cross-reactive animal-derived antibodies are commercially available (Methods). We designed sensors for each epitope and identified designs that specifically responded to anti-membrane and anti-nucleocapsid antibodies (Extended Data Fig. 8a, b). These sensors reached full signal in 2–5 min and had an approximately 50–70% dynamic range in response to low nanomolar amounts of antibodies (Fig. 4a, b, Extended Data Fig. 8c, d).

lucCage sensors for SARS-CoV-2 spike protein

To create sensors that can detect SARS-CoV-2 viral particles directly, we integrated a de novo designed picomolar affinity binder to the receptor-binding domain (RBD) of the SARS-CoV-2 spike protein named LCB1⁴ into the lucCage format (Fig. 4c). Of 13 candidates tested, the best, which we refer to as 'lucCageRBD', could detect both monomeric RBD and the full trimeric SARS-CoV-2 spike protein³⁷ with 15 pM and 47 pM LOD, respectively, and a more than 1,700% dynamic range for the RBD detection (Fig. 4c, Extended Data Fig. 9). We further increased the dynamic range

of lucCageRBD to 5,300% by tuning the cage–key affinity (K_{CK}) through shortening lucKey (Extended Data Fig. 10a–c). In addition to virus detection, the RBD sensor could also be used to monitor antibody generation in response to vaccination using a competition format analogous to that described above for the detection of HBV antibodies (Fig. 3f)—the ability to quantify responses over a wide dynamic range and to distinguish neutralizing antibodies binding at the ACE2-binding site on the RBD (and hence competing with LCB1 in the sensor) from non-neutralizing antibodies binding elsewhere are potential advantages over lateral flow assays.

To evaluate the ability of our sensor platform to function in complex biological matrices, we compared RBD detection by lucCageRBD in buffer, simulated nasal matrix³⁸, and human serum, and observed only a minor reduction in the latter two conditions (Fig. 4c). Following a suggestion by M. Merckx³⁹, we controlled for variation in absolute luminescence signal in spiked serum samples from four different donors and spiked simulated nasal matrix using a BRET reference⁴⁰ for internal calibration, and found that with such calibration the RBD could be accurately quantified without compromising the sensor dynamic range (Extended Data Fig. 11). These results suggest that the lucCage system could be used in point-of-care diagnostic devices.

Sensor specificity

To test the specificity of the designed biosensors, we measured the activation kinetics of each lucCage in response to each of the targets

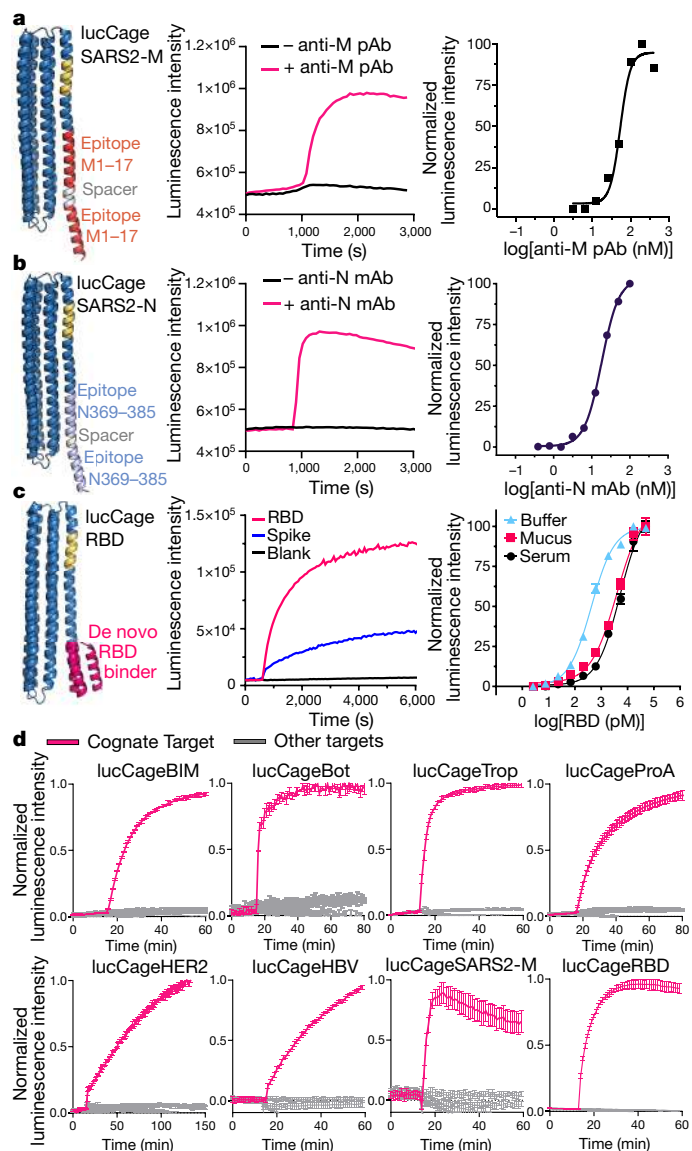


Fig. 4 | Design of high-specificity biosensors for anti-SARS-CoV-2 antibodies and SARS-CoV-2 viral proteins. **a**, Left, lucCageSARS2-M sensor incorporates two copies of the SARS-CoV-2 membrane protein 1–17 epitope (red) connected with a flexible spacer. Middle, kinetics of luminescence activation of 50 nM lucCageSARS2-M plus lucKey after the addition of 100 nM anti-SARS-CoV-1-M rabbit polyclonal antibodies (pAb) that cross-react with residues 1–17 of the SARS-CoV-2 membrane protein. Right, response of 5 nM lucCageSARS2-M plus lucKey to varying concentrations of target anti-M polyclonal antibody. **b**, Left, lucCageSARS2-N incorporates two copies of the SARS-CoV-2 nucleocapsid protein 369–382 epitope (light blue). Middle, kinetics of luminescence activation of 50 nM lucCageSARS2-N plus lucKey after the addition of 100 nM anti-SARS-CoV-1-N mouse monoclonal antibody (clone 18F629.1) that recognizes the epitope. Right, response of 50 nM lucCageSARS2-N plus lucKey to varying concentrations of target anti-N monoclonal antibody. **c**, Left, lucCageRBD incorporates a de novo SARS-CoV-2 RBD binder⁴ (LCB1, magenta). Middle, luminescence intensities increase after the addition of 16.7 nM SARS-CoV-2 RBD or trimeric spike protein to a mixture of 1 nM lucCageRBD plus lucKey. Right, detection over a range of analyte concentrations in buffer, 10% synthetic nasal matrix³⁸ or 10% serum. **d**, Biosensor specificity. Each sensor at 1 nM was incubated with 50 nM of its cognate target (magenta lines) and the targets for the other biosensors (grey lines). Targets are BCL-2, BoNT/B, human IgG Fc, HER2, cTnI, anti-HBV antibody (HzKR127-3.2), anti-SARS-CoV-1-M polyclonal antibody and SARS-CoV-2 RBD. All experiments were performed in triplicate, representative data are shown, and data are mean \pm s.d.

one at a time. Each sensor responded rapidly and sensitively to its cognate target, but not to any of the others (Fig. 4d). For the most part, the actual sensors (Supplementary Tables 2, 3) performed as predicted by the simple thermodynamic model; for example, experiments at varying key and sensor concentrations suggest little coupling between parameters. However, there is considerable variation between different sensors in the level of activation at saturating target concentrations or high lucKey concentrations, which for most is lower than that expected for the complete luciferase reconstitution predicted by the model (Extended Data Fig. 10d–g, Supplementary Table 4). This may be a consequence of steric interference between target binding to the latch and luciferase reconstitution as the target binding motif and the luciferase SmBiT are adjacent to each other in the latch; such interference could be resolved by increasing the separation between the two in the switch. The potential of the lucCage system is illustrated by the high dynamic range (5,300%) and picomolar sensitivity of the lucCageRBD sensor: the near optimal K_{open} value results in a very low background in the absence of target without compromising the extent of activation at low target concentrations.

Discussion

It is instructive to put our sensors in the context of the many protein-based biosensor platforms that have been developed over the years with considerable success (Supplementary Discussion, Supplementary Table 5). Our sensor platform is based on the thermodynamic coupling between defined closed and open states of the system, and thus, its sensitivity depends on the free energy change that occurs after the sensing domain binds to the target but not the specific geometry of the binding interaction (the semi-synthetic small molecule sensors^{10,11} also have this property). This enables the incorporation of various binding modalities—including small peptides, globular mini-proteins, antibody epitopes and de novo designed binders—to generate sensitive sensors for a wide range of protein targets with little or no optimization. For point of care applications, our system, similar to other bioluminescence-based protein biosensor platforms⁸, has the advantages of being homogeneous, no-wash, and a nearly instantaneous readout; the quantification of luminescence can be performed with inexpensive and accessible devices such as the camera of a mobile phone⁸. In hospital settings, the ability to modularly design sensors with identical readouts for diverse targets could enable the quick readout of large numbers of different compounds using an array of hundreds of different sensors.

Until recently, the focus of de novo protein design was the design of proteins with new structures that correspond to single deep free energy minima; our results highlight the progress in the field that now enables more complex multistate systems to be readily generated. Similar to other de novo designed proteins, our sensors are expressed at high levels in cells and are very stable⁴¹, which should considerably facilitate their manufacturing and distribution. As highlighted by the outstanding performance of the lucCageRBD sensor, there is a strong synergy between the general ‘molecular device’ architecture of our platform and de novo designed high-affinity mini-protein binders^{4,21} (these de novo mini-proteins are also effective with other platforms⁴²). As the power of computational design continues to increase, it should become possible to detect an ever wider range of targets with higher sensitivity using lucCage sensors. Beyond biosensors, our results highlight the potential of de novo protein design to create more general solutions for current day challenges than can be achieved by repurposing native proteins that have evolved to solve completely different challenges.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03258-z>.

- Stein, V. & Alexandrov, K. Synthetic protein switches: design principles and applications. *Trends Biotechnol.* **33**, 101–110 (2015).
- Langan, R. A. et al. De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
- Udugama, B. et al. Diagnosing COVID-19: the disease and tools for detection. *ACS Nano* **14**, 3822–3835 (2020).
- Cao, L. et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
- Yeh, H.-W. & Ai, H.-W. Development and applications of bioluminescent and chemiluminescent reporters and biosensors. *Annu. Rev. Anal. Chem. (Palo Alto Calif.)* **12**, 129–150 (2019).
- Greenwald, E. C., Mehta, S. & Zhang, J. Genetically encoded fluorescent biosensors illuminate the spatiotemporal regulation of signaling networks. *Chem. Rev.* **118**, 11707–11794 (2018).
- Banala, S., Aper, S. J. A., Schalk, W. & Merckx, M. Switchable reporter enzymes based on mutually exclusive domain interactions allow antibody detection directly in solution. *ACS Chem. Biol.* **8**, 2127–2132 (2013).
- Arts, R. et al. Detection of antibodies in blood plasma using bioluminescent sensor proteins and a smartphone. *Anal. Chem.* **88**, 4525–4532 (2016).
- van Rosmalen, M. et al. Dual-color bioluminescent sensor proteins for therapeutic drug monitoring of antitumor antibodies. *Anal. Chem.* **90**, 3592–3599 (2018).
- Yu, Q. et al. Semisynthetic sensor proteins enable metabolic assays at the point of care. *Science* **361**, 1122–1126 (2018).
- Yu, Q. et al. A biosensor for measuring NAD⁺ levels at the point of care. *Nat. Metab.* **1**, 1219–1225 (2019).
- Schena, A., Griss, R. & Johnsson, K. Modulating protein activity using tethered ligands with mutually exclusive binding sites. *Nat. Commun.* **6**, 7830 (2015).
- Arts, R. et al. Semisynthetic bioluminescent sensor proteins for direct detection of antibodies and small molecules in solution. *ACS Sens.* **2**, 1730–1736 (2017).
- 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).
- Guo, Z. et al. Generalizable protein biosensors based on synthetic switch modules. *J. Am. Chem. Soc.* **141**, 8128–8135 (2019).
- Edwardraja, S. et al. Caged activators of artificial allosteric protein biosensors. *ACS Synth. Biol.* **9**, 1306–1314 (2020).
- 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).
- Dixon, A. S. et al. NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
- Minor, D. L., Jr & Kim, P. S. Context-dependent secondary structure formation of a designed protein sequence. *Nature* **380**, 730–734 (1996).
- Kale, J., Osterlund, E. J. & Andrews, D. W. BCL-2 family proteins: changing partners in the dance towards death. *Cell Death Differ.* **25**, 65–80 (2018).
- Chevalier, A. et al. Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).
- Lajoie, M. J. et al. Designed protein logic to target cells with precise combinations of surface antigens. *Science* **369**, 1637–1643 (2020).
- Deis, L. N. et al. Suppression of conformational heterogeneity at a protein-protein interface. *Proc. Natl Acad. Sci. USA* **112**, 9028–9033 (2015).
- Eigenbrot, C., Ultsch, M., Dubnovitsky, A., Abrahmsén, L. & Härd, T. Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc. Natl Acad. Sci. USA* **107**, 15039–15044 (2010).
- Hobbs, R. J., Thomas, C. A., Halliwell, J. & Gwenin, C. D. Rapid detection of Botulinum neurotoxins—a review. *Toxins (Basel)* **11**, 418 (2019).
- 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).
- 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).
- Collins, M. H. Serologic tools and strategies to support intervention trials to combat Zika virus infection and disease. *Trop. Med. Infect. Dis.* **4**, 68 (2019).
- Pondé, R. A. 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).
- 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. USA* **104**, 9230–9235 (2007).
- 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).
- Ovacik, M. & Lin, K. Tutorial on monoclonal antibody pharmacokinetics and its considerations in early development. *Clin. Transl. Sci.* **11**, 540–552 (2018).
- Locarnini, S. & Bowden, S. Hepatitis B surface antigen quantification: not what it seems on the surface. *Hepatology* **56**, 411–414 (2012).
- 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).
- He, Y., Zhou, Y., Siddiqui, P., Niu, J. & Jiang, S. Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. *J. Clin. Microbiol.* **43**, 3718–3726 (2005).
- Wang, H. et al. SARS-CoV-2 proteome microarray for mapping COVID-19 antibody interactions at amino acid resolution. *ACS Cent. Sci.* **6**, 2238–2249 (2020).
- Hsieh, C.-L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
- Panpradist, N. et al. Swab sample transfer for point-of-care diagnostics: characterization of swab types and manual agitation methods. *PLoS ONE* **9**, e105786 (2014).
- Ni, Y. et al. RAPPID: a platform of ratiometric bioluminescent sensors for homogeneous immunoassays. Preprint at <https://doi.org/10.1101/2020.10.31.363044> (2020).
- Yeh, H.-W. et al. Red-shifted luciferase-luciferin pairs for enhanced bioluminescence imaging. *Nat. Methods* **14**, 971–974 (2017).
- Baker, D. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 (2019).
- Yang, C. et al. Bottom-up de novo design of functional proteins with complex structural features. *Nat. Chem. Biol.* <https://doi.org/10.1038/s41589-020-00699-x> (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Design of the sensor system (lucCage and lucKey)

The low affinity SmBit 114 (VTGYRLFEEIL)¹⁸ was grafted into the latch of the asymmetric LOCKR switch previously described² using GraftSwitchMover, a RosettaScripts-based protein design algorithm (see Supplementary Methods for details). The grafting sampling range was assigned between residues 300 and 330. The resulting designs were energy-minimized, visually inspected and selected for subsequent gene synthesis, protein production and biochemical analyses. The best SmBit position on the latch was experimentally determined to be an insertion at residue 312, as described in Extended Data Fig. 2. This design was named lucCage. lucKey was assembled by genetically fusing the LgBit of NanoLuc¹⁸ to the key peptide previously described². All protein sequences are listed in Supplementary Table 6.

Computational grafting of sensing domains into lucCage

For peptides and epitopes, the amino acid sequence for each sensing domain was grafted using Rosettascripts⁴³ GraftSwitchMover into all α -helical registers between residues 325 and 359 of lucCage. In the cases in which the desired sequence to be inserted exceeded the length of the lucCage latch, we made use of Rosetta Remodel⁴⁴ to model the C terminus extension of lucCage (see Supplementary Methods for details). The resulting lucCages were energy-minimized using Rosetta fast relax⁴⁵, visually inspected and typically less than ten designs were selected for subsequent protein production and biochemical characterization.

For protein domains, the main secondary structure element forming the interface of the binding protein domain with the target was identified. The amino acid sequence was extracted and grafted into lucCage using the GraftSwitchMover or Rosetta Remodel as described above. Then, we used MergePDBMover and Pymol 2.0 to align, model and visualize the full-length binding domain in the context of the switch (see Supplementary Methods for details). The designs were energy-minimized using Rosetta fast relax and visually inspected for selection.

Synthetic gene construction

The designed protein sequences were codon-optimized for *E. coli* expression and ordered as synthetic genes in pET21b+ or pET29b+ *E. coli* expression vectors. The synthetic gene was inserted at the NdeI and XhoI sites of each vector, including an N-terminal hexahistidine tag followed by a TEV protease cleavage site and a stop codon was added at the C terminus.

General procedures for bacterial protein production and purification

The *E. coli* Lemo21(DE3) strain (NEB) was transformed with a pET21b+ or pET29b+ plasmid encoding the synthesized gene of interest. Cells were grown for 24 h in LB medium supplemented with carbenicillin or kanamycin. Cells were inoculated at a 1:50 ml ratio in the Studier TBM-5052 autoinduction medium supplemented with carbenicillin or kanamycin, grown at 37 °C for 2–4 h, and then grown at 18 °C for an additional 18 h. Cells were collected by centrifugation at 4,000g at 4 °C for 15 min and resuspended in 30 ml lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole, 1 mM PMSF, 0.02 mg ml⁻¹ DNase). Cell resuspensions were lysed by sonication for 2.5 min (5 s cycles). Lysates were clarified by centrifugation at 24,000g at 4 °C for 20 min and passed through 2 ml of Ni-NTA nickel resin (Qiagen, 30250) pre-equilibrated with wash buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole). The resin was washed twice with 10 column volumes (CV) of wash buffer, and then eluted with 3 CV of elution buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole). The eluted proteins were concentrated using Ultra-15 Centrifugal Filter Units (Amicon) and further purified by using a Superdex 75 Increase 10/300 GL (GE Healthcare) size exclusion column in TBS (25 mM Tris-HCl pH 8.0,

150 mM NaCl). Fractions containing monomeric protein were pooled, concentrated, and snap-frozen in liquid nitrogen and stored at –80 °C.

In vitro bioluminescence characterization

A Synergy Neo2 Microplate Reader (BioTek) was used for all in vitro bioluminescence measurements. Assays were performed in 50% DPBS with calcium (Gibco) plus 50% Nano-Glo (Promega) assay buffer for cTnI sensors and 50% HBS-EP (GE Healthcare Life Sciences) plus 50% Nano-Glo assay buffer was used for other sensors. 10 \times lucCage, 10 \times lucKey and 10 \times target proteins of desired concentrations were first prepared from stock solutions. For each well of a white opaque 96-well plate, 10 μ l of 10 \times lucCage, 10 μ l of 10 \times lucKey and 20 μ l of buffer were mixed to reach the indicated concentration and ratio. The lucCage and lucKey components were incubated for 60 min at room temperature to enable pre-equilibration. The plate was centrifuged at 1,000g for 1 min and incubated at room temperature for a further 10 min. Then, 50 μ l of 50 \times diluted furimazine (Nano-Glo luciferase assay reagent, Promega) was added to each well. For assays containing serum or simulated nasal matrix (110 mM NaCl, 1% (w/v) mucin, 10 μ g ml⁻¹ human genomic DNA³⁸), buffer composition was replaced by the biological matrix. Bioluminescence measurements in the absence of target were taken every 1 min after injection (0.1 s integration and 10 s shaking during intervals). After around 15 min, 10 μ l of serially diluted 10 \times target protein plus a blank was injected and bioluminescence kinetic acquisition continued for a total of 2 h. To derive half-maximal effective concentration (EC₅₀) values from the bioluminescence-to-analyte plot, the top three peak bioluminescence intensities at individual analyte concentrations were averaged, subtracted from blank, and used to fit the sigmoidal 4PL curve. To calculate the LOD, the linear region of bioluminescence responses of sensors to its analyte was extracted and a linear regression curve was obtained. It was used to derive the standard deviation of the response and the slope of the calibration curve (*S*). The LOD was determined as: 3 \times (s.d./*S*).

Detection of spiked RBD in human serum specimens

Serum specimens were derived from excess plasma or sera from adults (>18 years) of both genders provided by the Director of the Clinical Chemistry Division, the hospital of University Washington. All anonymized donor specimens were provided de-identified. Because the donors consented to have their excess specimens be used for other experimental studies, they could be transferred to our study without additional consent. All samples were passed through 0.22- μ m filters before use. Ten microlitres of 10 \times serially diluted monomeric RBD (167–0.69 nM), 5 μ l of 20 \times lucCage (20 nM), 5 μ l of 20 \times lucKey (20 nM), 5 μ l of 20 \times Antares2 (2 nM), and 10, 20, 25 or 50 μ l of human donor serum or simulated nasal matrix were mixed with 1:1 HBS:Nano-Glo assay buffer to reach a total volume of 75 μ l. The plate was centrifuged at 1,000g for 1 min. Then, 25 μ l of 25 \times diluted furimazine in buffer was added to each well. Bioluminescence signals were recorded from both 470/40 nm and 590/35 nm channels every 1 min for a total of 1 h. The ratio at each time point was calculated by the equation described in Extended Data Fig. 11b. Monomeric SARS-CoV-2 RBD was expressed and purified as previously described⁴⁶.

Biolayer interferometry

Protein–protein interactions were measured by using an Octet RED96 System (ForteBio) using streptavidin-coated biosensors (ForteBio). Each well contained 200 μ l of solution, and the assay buffer was HBS-EP+ buffer (GE Healthcare Life Sciences, 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% (v/v) surfactant P20) plus 0.5% non-fat dry milk blotting grade blocker (BioRad). The biosensor tips were loaded with analyte peptide or protein at 20 μ g ml⁻¹ for 300 s (threshold of 0.5 nm response), incubated in HBS-EP+ buffer for 60 s to acquire the baseline measurement, dipped into the solution containing cage and/or key for 600 s (association step) and dipped into the HBS-EP+ buffer for 600 s

(dissociation steps). The binding data were analysed with the ForteBio Data Analysis Software version 9.0.0.10.

Design and characterization of lucCageBIM

The BIM peptide sequence (EIWIAQELRRIGDEFNAYYA) was threaded into the lucCage scaffold as described in 'Computational grafting of sensing domains into lucCage'. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The bioluminescence detection signal was measured for each design lucCage at 20 nM mixed with lucKey at 20 nM, in the presence or absence of target BCL-2 protein at 200 nM. Recombinant BCL-2 was produced as previously described⁴⁷.

Design and characterization of lucCageHER2, lucCageProA, lucCageBot and lucCageRBD

The main binding motifs of the Bot.0671.2 de novo binder, *Staphylococcus aureus* protein A domain C (SpaC), the HER2 antibody and the de novo RBD binder LCB1 were threaded into lucCage as described in 'Computational grafting of sensing domains into lucCage' (see Supplementary Tables 3 and 6 for sequences). The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The designs were screened by measuring bioluminescence signal for each design lucCage at 20 nM mixed with lucKey at 20 nM, in the presence or absence of 200 nM target protein. The target proteins used were: Botulinum neurotoxin B HcB expressed as previously described⁴⁸, human IgG1 Fc-HisTag (AcroBiosystems, IG1-H5225) and human HER2-HisTag (AcroBiosystems, HE2-H5225). Monomeric SARS-CoV-2 RBD and the trimeric SARS-CoV-2 spike protein (Hexapro pre-stabilized version³⁷) were expressed and purified as previously described⁴⁶.

Design and characterization of lucCageTrop

The cTnT binding motif sequence was truncated into fragments of different length (Extended Data Fig. 6) and threaded into the lucCage scaffold as described in 'Computational grafting of sensing domains into lucCage'. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The designs were screened by measuring bioluminescence signal for each design lucCage at 20 nM mixed with lucKey at 20 nM in the presence or absence of 100 nM cTnI (Genscript, Z03320-50). Subsequently, lucCageTrop, an improved version by fusion to cTnC, was created by genetically fusing the following sequence to the C terminus of lucCageTrop627.

Design and characterization of lucCageHBV and lucCageHBV α

The binding motif (GANSNPDWDFN) of the preS1 domain was threaded into the lucCage scaffold at every position after residues 336 using the Rosetta GraftSwitchMover. Following the Rosetta FastRelax protocol, eight designs were selected for protein production. The designs were screened by measuring bioluminescence signal for each design lucCage (20 nM) and lucKey (20 nM) in the presence or absence of the anti-HVB antibody HzKR127-3.2 (100 nM) to select lucCageHBV. Subsequently, lucCageHBV α was constructed by genetically fusing a sequence containing a second antigenic motif (GGSGGGSSGFGANSNPDWDFNPN) to lucCageHBV.

Design and characterization of lucCageSARS2-M and lucCageSARS2-N

Antigenic epitopes of the SARS-CoV-2 membrane protein (amino acids 1–31, 1–17 and 8–24) and the nucleocapsid protein (amino acids 368–388 and 369–382) were computationally grafted into lucCage as described in 'Computational grafting of sensing domains into lucCage'. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. All designs at 50 nM were mixed with 50 nM lucKey and experimentally screened for an increase in luminescence in the presence of rabbit anti-SARS-CoV membrane polyclonal antibodies (ProSci, 3527) at 100 nM or mouse anti-SARS-CoV

nucleocapsid monoclonal antibody (clone 18F629.1, NovusBio NBP2-24745) at 100 nM.

Design and characterization of sCageHA variants

HB1.9549.2 was embedded into the parental six-helix bundle for sCage design at different positions along the latch helix of the scaffold. To promote more favourable intramolecular interactions, three consecutive residues on the latch were intentionally substituted with glycine to allow for conformational freedom. The five designs were produced in *E. coli*. Biolayer interferometry analysis was performed with purified cages (1 μ M) and biotinylated influenza A H1 HA²¹ loaded onto streptavidin-coated biosensor tips (ForteBio) in the presence or absence of the key (2 μ M) using an Octet instrument (ForteBio).

Production and purification of HzKR127-3.2

The synthetic V_H and V_L DNA fragments were subcloned into the pdCMV-dhfrC-cA10A3 plasmid containing the human C γ 1 and C γ DNA sequences. The vector was introduced into HEK 293F cells using Lipofectamine (Invitrogen), and the cells were grown in FreeStyle 293 (GIBCO) in 5% CO₂ in a 37 °C humidified incubator. The culture supernatant was loaded onto a protein A-sepharose column (Millipore), and the bound antibody was eluted by the addition of 0.2 M glycine-HCl (pH 2.7), followed by immediate neutralization with 1 M Tris-HCl (pH 8.0). The solution was dialysed against 10 mM HEPES-NaOH (pH 7.4), and the purity of the protein was analysed by SDS-PAGE.

Production and purification of the preS1 domain

The DNA fragment encoding the preS1 domain (residues 1–56) was cloned into the pGEX-2T (GE Healthcare) plasmid, and the protein was produced in the *E. coli* BL21(DE3) strain (NEB) at 18 °C as a fusion protein with glutathion-S-transferase (GST) at the N terminus. The cell lysates were prepared in a buffer solution (25 mM Tris-HCl pH 8.0, 300 mM NaCl), and clarified supernatant was loaded onto GSTBind Resin (Novagen). The GST-preS1 domain was eluted with the same buffer containing additional 10 mM reduced glutathione, further purified using a Superdex 75 Increase 10/300 GL (GE Healthcare) size-exclusion column, and concentrated to 34 μ M.

Production of sCageHA_{267-1S} and its variants

sCageHA_{267-1S} and sCageHA_{267-1S(E99Y/T144Y)} were expressed at 18 °C in the *E. coli* LEMO21(DE3) strain (NEB) as a fusion protein containing a (His)₁₀-tagged cysteine protease domain (CPD) derived from *Vibrio cholerae*⁴⁹ at the C terminus. The protein was purified using HisPur nickel resin (Thermo), a HiTrap Q anion exchange column (GE Healthcare) and a HiLoad 26/60 Superdex 75 gel filtration column (GE Healthcare). For selenomethionine (SelMet)-labelling, an I30M mutation was introduced additionally to generate a sCageHA_{267-1S(E99Y/T144Y/I30M)} variant. This protein was expressed in the *E. coli* B834 (DE3) RIL strain (Novagen) in the minimal medium containing SeMet, and purified according to the same procedure for purifying the other variants.

Crystallization and structure determination of sCageHA_{267-1S}

Two point mutations (Glu99Tyr and Thr144Tyr) were introduced in an attempt to induce favourable crystal packing interactions. Good-quality single crystals of sCageHA_{267-1S(E99Y/T144Y/I30M)} were obtained in a hanging-drop vapour-diffusion setting by micro-seeding in a solution containing 11% (v/v) ethanol, 0.25 M NaCl, 0.1 M Tris-HCl (pH 8.5). The crystals required strict maintenance of the temperature at 25 °C. For cryoprotection, the crystals were soaked briefly in the crystallization solution supplemented with 15% 2,3-butanediol and flash-cooled in the liquid nitrogen. A single-wavelength anomalous dispersion dataset was collected at the Se absorption peak and processed with HKL2000⁵⁰. Se positions and initial electron density map were calculated using the AutoSol module in PHENIX⁵¹. The model building and structure refinement were performed by using COOT⁵² and PHENIX.

Statistical analysis

No statistical methods were used to pre-determine the sample size. No sample was excluded from data analysis, and no blinding was used. De-identified clinical serum samples were randomly used for spiking in target proteins. Results were successfully reproduced using different batches of pure proteins on different days. Unless otherwise indicated, data are shown as mean \pm s.d., and error bars in figures represent s.d. of technical triplicate. Biolayer interferometry data were analysed using ForteBio Data Analysis Software version 9.0.0.10. All data were analysed and plotted using GraphPad Prism 8.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The atomic coordinates of sCageHA_267-1S have been deposited in the PDB under accession code 7CBC. The original experimental data that supports the findings of this work are available from the corresponding authors upon request. Plasmids encoding the biosensor proteins described in this Article are available from the corresponding authors upon request.

Code availability

The design models and RosettaScripts code used in the manuscript have been deposited to http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/designcode_and_models.zip. The code for the numerical simulations shown in this manuscript are available at http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/model_simulation.py

- Fleishman, S. J. et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS ONE* **6**, e20161 (2011).
- Huang, P.-S. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS ONE* **6**, e24109 (2011).
- Khatib, F. et al. Algorithm discovery by protein folding game players. *Proc. Natl Acad. Sci. USA* **108**, 18949–18953 (2011).
- Walls, A. C. et al. Elicitation of potent neutralizing antibody responses by designed protein nanoparticle vaccines for SARS-CoV-2. *Cell* **183**, 1367–382.e17 (2020).
- Berger, S. et al. Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *eLife* **5**, e20352 (2016).
- Jin, R., Rummel, A., Binz, T. & Brunker, A. T. Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. *Nature* **444**, 1092–1095 (2006).

- Shen, A. et al. Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin. *Nat. Chem. Biol.* **5**, 469–478 (2009).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
- Potterton, L. et al. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D* **60**, 2288–2294 (2004).
- Yeh, H.-W. et al. ATP-independent bioluminescent reporter variants to improve in vivo imaging. *ACS Chem. Biol.* **14**, 959–965 (2019).

Acknowledgements We acknowledge funding from HHMI (D.B.), the LG Yonam Foundation (B.-H.O.), the BK21 PLUS project of Korea (H.L.), the United World Antiviral Research Network (UWARN) one of the Centers Researching Emerging Infectious Diseases 'CREIDs', NIAID 1 U01 AI151698-01 (D.B., L.S. and H.-W.Y.), The Audacious Project at the Institute for Protein Design (D.B., H.-W.Y., C.M.C. and M.C.M.), Eric and Wendy Schmidt by recommendation of the Schmidt Futures (A.Q.-R. and H.-W.Y.), the Washington Research Foundation (J.P. and M.J.L.), the Nordstrom Barrier Institute for Protein Design Directors Fund (R.A.L.), The Open Philanthropy Project Improving Protein Design Fund (D.B. and S.E.B.), the gift support from Gree Real Estate (A.Q.-R.), 'la Caixa' Foundation (A.Q.-R., ID 100010434 under grant LCF/BQ/AN15/10380003), support 1U19AG065156-01 (D.B.), and Air Force Office of Scientific Research FA9550-18-1-0297 (D.B.). We thank M. Wener for collecting de-identified human sera specimens, W. C. Van Voorhis for advice and support with the anti-SARS-CoV-2 antibody sensors, N. Panpradist and B. Lutz for providing simulated nasal matrix, S. Berger for sharing the BCL-2 protein target, D. A. Silva Manzano for providing Botulinum neurotoxin B, A. Kang for setting up screening crystal trays, and L. Carter, B. Fiala and the Institute for Protein Design for providing SARS-CoV-2 RBD and spike protein. We thank M. Merx for suggesting the internal BRET referencing to control for sample fluctuations in luciferase activity. The X-ray data were collected on the Beamline 5C at the Pohang Accelerator Laboratory, Korea. All protein structure and model images were generated using PyMOL 2.0.

Author contributions D.B. directed the work. D.B., A.Q.-R., H.-W.Y., B.-H.O., J.P. and L.S. designed and further conceptualized the research. A.Q.-R., J.P., B.-H.O. and H.-W.Y. performed the computational design of the sensors. A.Q.-R., J.P. and B.-H.O. explored the early sensor designs. H.-W.Y. and A.Q.-R. optimized the design of sensors to improve performance and conducted the experimental characterizations. H.-W.Y. and A.Q.-R. performed the experimental validation. B.H.O. directed and H.L. performed the crystallographic work. H.-W.Y. and R.A.L. wrote the thermodynamic model and performed the simulations. R.A.L. wrote GraftSwitchMover. S.E.B. and M.J.L. designed the parental cage and key protein scaffolds. L.C. designed the RBD binder LCB1. C.M.C., M.C.M., J.W. and H.J.H. performed protein purification. D.B., B.-H.O., A.Q.-R. and H.-W.Y. wrote the original draft. All authors reviewed and accepted the manuscript.

Competing interests D.B., A.Q.-R., H.-W.Y. and J.P. are co-inventors in a provisional patent application (application number 63/030,836 submitted by the University of Washington) covering the biosensors described in this Article. D.B., A.Q.-R., H.-W. Y., L.C. and L.S. are co-inventors in a provisional patent application (application number 63/051549 submitted by the University of Washington) covering the SARS-CoV-2 RBD biosensor described in this Article. The rest of the authors declare no competing interests.

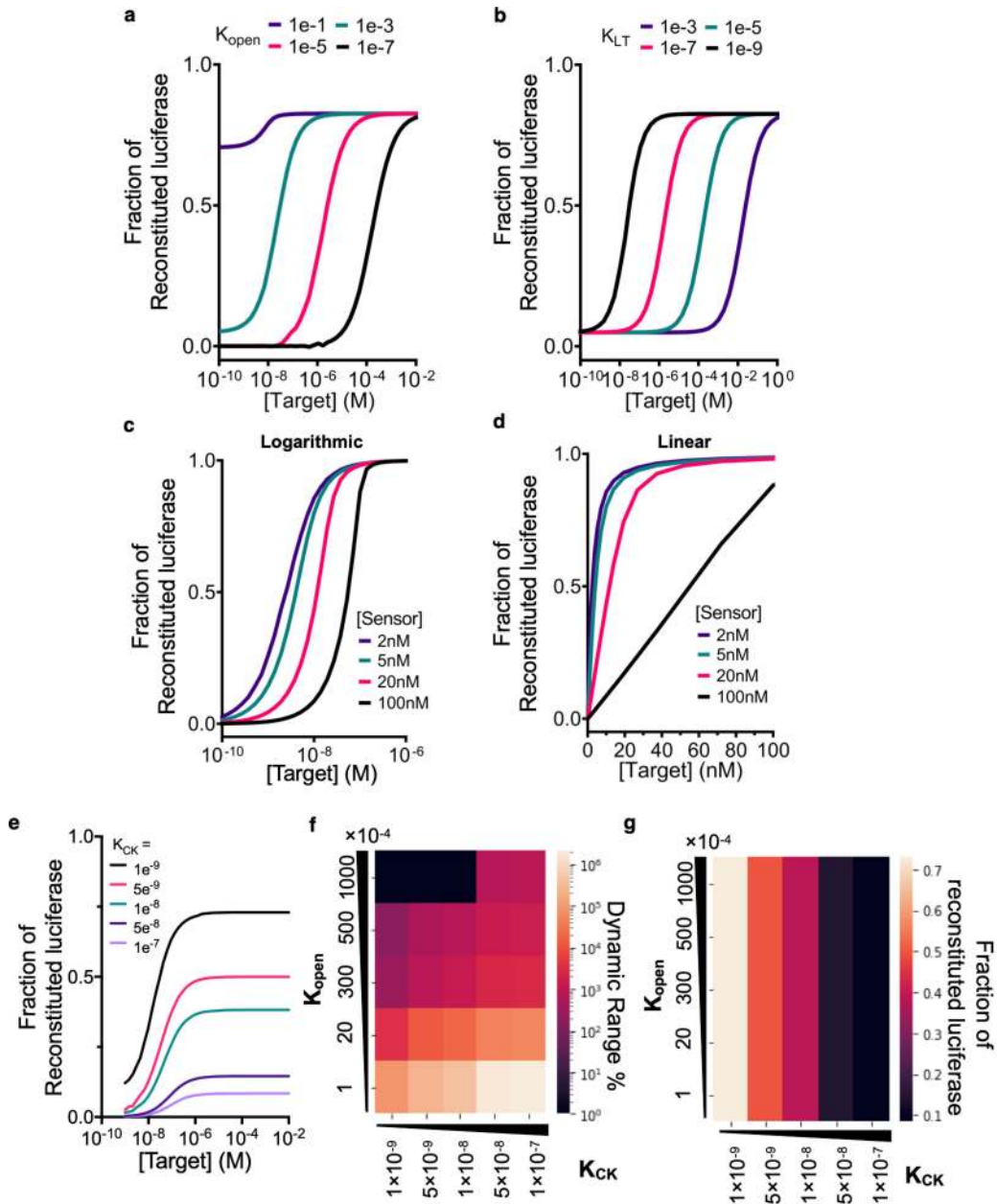
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-03258-z>.

Correspondence and requests for materials should be addressed to B.-H.O. or D.B.

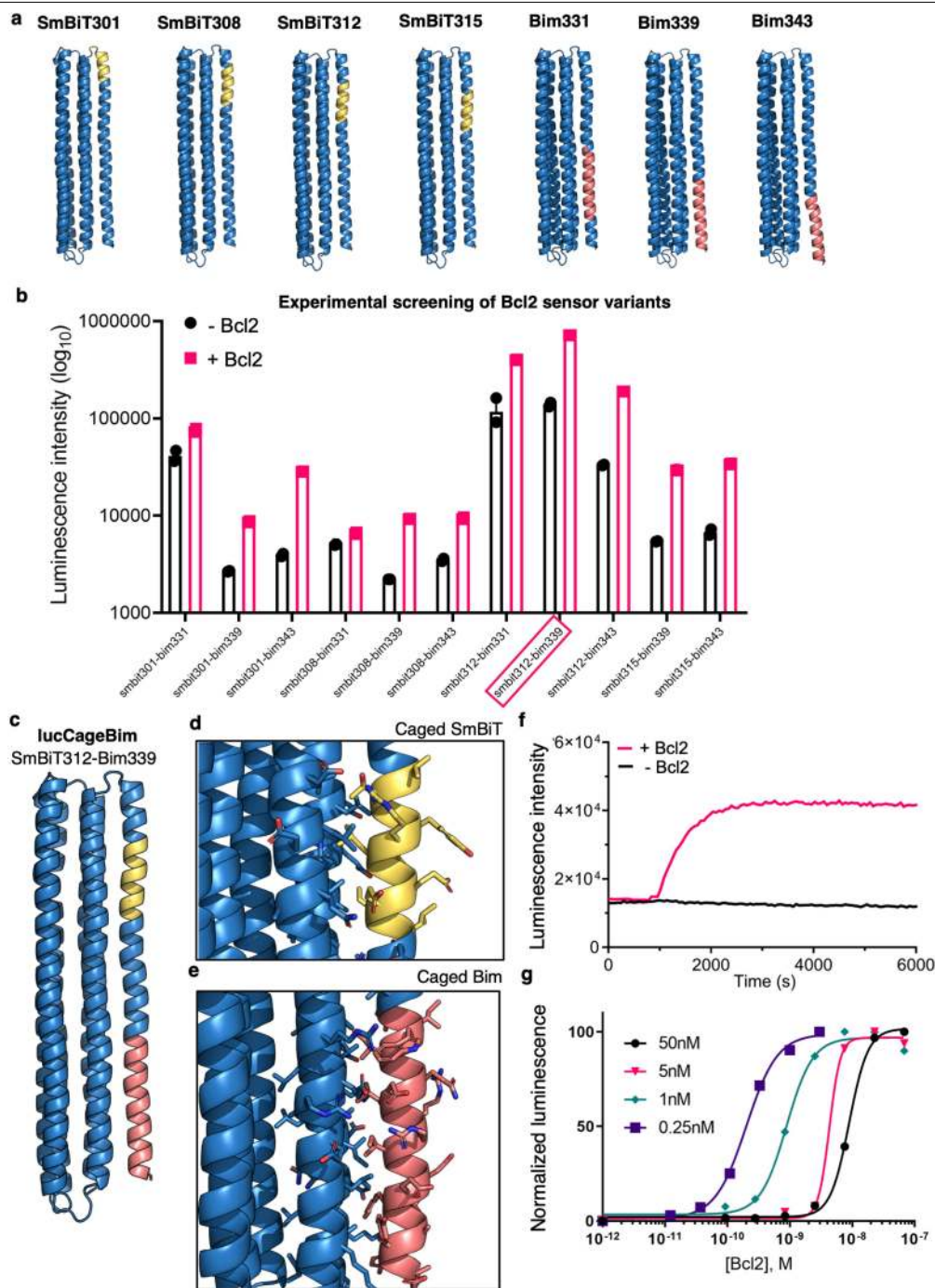
Peer review information Nature thanks Caryn Bern, Vincent Hilser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Numerical simulations of the sensor thermodynamic equilibria showing the tunability of the lucCage platform to optimize sensitivity and dynamic range. a–e, Numerical simulations of the coupled equilibria shown in Fig. 1b for different values of K_{open} (a), K_{LT} (b), $[LucKey]_{tot}$ (c), $[LucCage]_{tot}$ (d) and K_{open} (e). Unless indicated otherwise, the simulations were performed with fixed values for $K_{open} = 1 \times 10^{-3}$ M, $K_{LT} = 10^{-9}$ M, and $K_{CK} = 10^{-8}$ M, and the concentration of the sensor components to 10:100 nM ($[LucCage]_{tot}:[LucKey]_{tot}$). **a**, Increasing ΔG_{open} (smaller K_{open}) shifts the sensor response to higher analyte concentrations. **b**, The sensor LOD is approximately $0.1 \times K_{LT}$; the driving force for opening the switch becomes too weak below this concentration. **c, d**, The effective target detection range can be tuned by changing the concentrations of the two sensor components. Simulation results

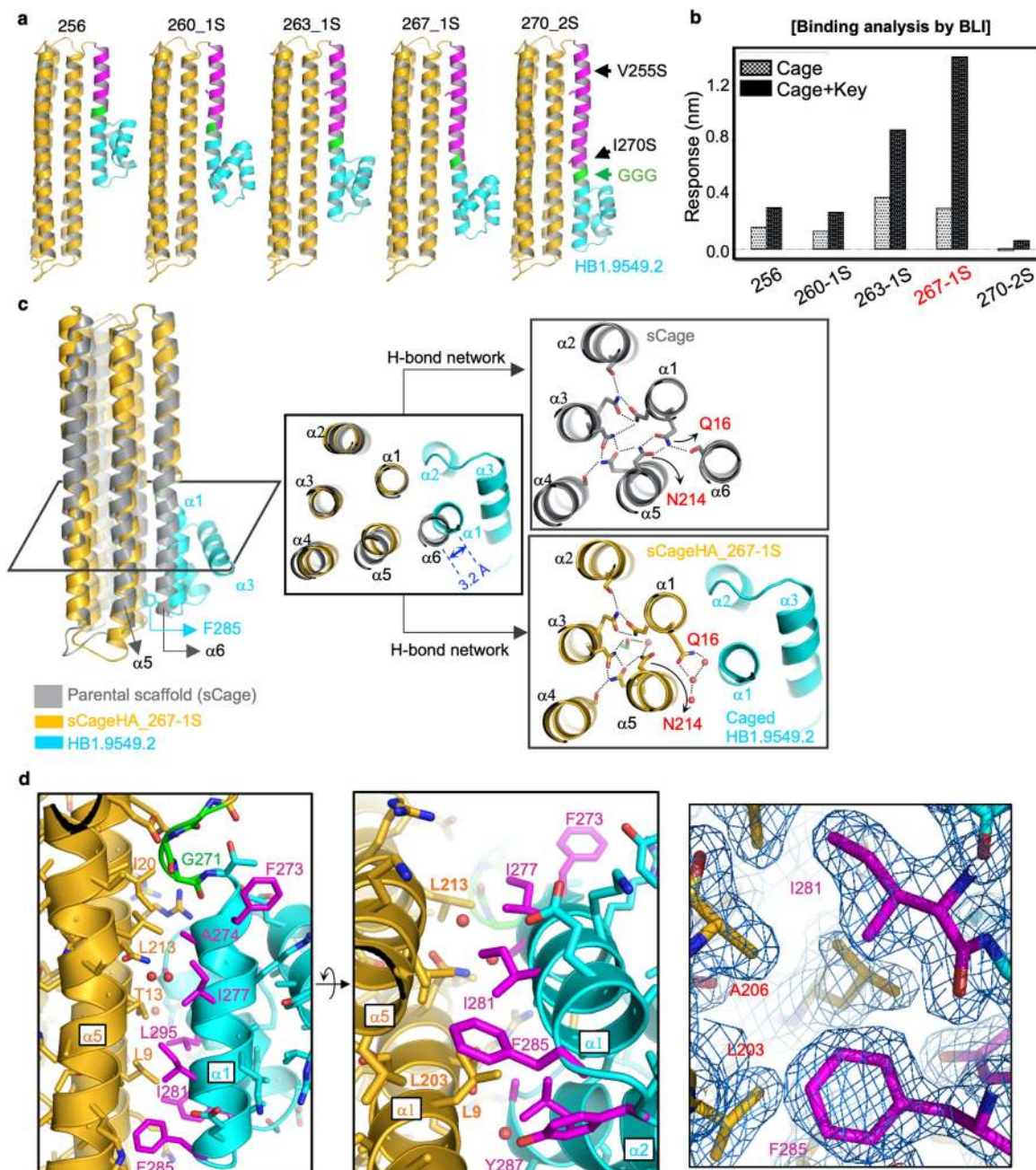
shown in a logarithmic scale (c) or linear scale (d) for target concentration illustrate that the steepness of the response depends on the ratio of the sensor concentration to the latch–target binding interaction (K_{LT}). **e**, K_{CK} values affect both species responsible for background and signal (species 6 and 7 in Fig. 1b, respectively), leading to different sensor dynamic ranges. **f, g**, Simulations with various K_{open} and K_{CK} values. The formation of species 6 (in Fig. 1b) is increased by large K_{open} values and strong lucCage–lucKey interactions (K_{CK}). **f**, A heat map representing the calculated sensor dynamic range according to the K_{open} and K_{CK} values. K_{open} exerts a predominant effect on the dynamic range, and K_{CK} provides an additional one-order of tunability. **g**, A heat map showing the fraction of reconstituted luciferase (sensitivity) at saturating target concentration, indicating a trade-off of K_{CK} tuning.



Extended Data Fig. 2 | Determination of the optimal SmBiT position in lucCage and characterization of the BCL-2 biosensor lucCageBIM.

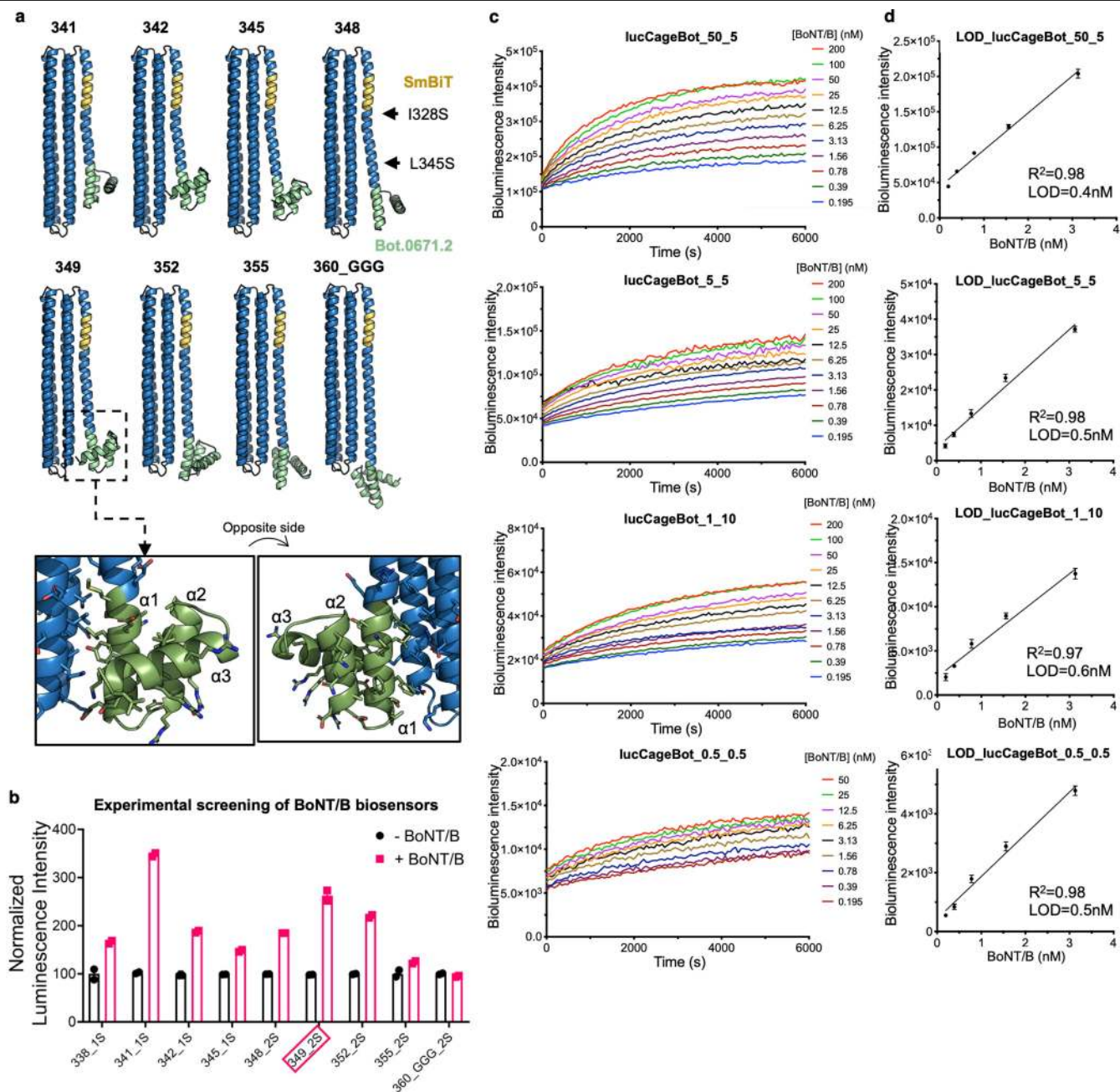
a, Protein models showing the different threading positions of SmBiT (gold) and the BIM peptide (salmon) on the latch helix of the de novo LOCKR switch (blue). **b**, Experimental screening of 11 de novo BCL-2 sensors. Eleven variants were generated by combining the SmBiT and BIM positions in **a** and characterized by activation of their luminescence after the addition of BCL-2. Luminescence measurements were performed with each design (20 nM) and lucKey (20 nM) in the presence or absence of BCL-2 (200 nM). SmBiT312-BIM339

(hence lucCageBIM) was selected for posterior characterization owing to its higher brightness, dynamic range and stability. **c–g**, Characterization of lucCageBIM. **c**, Structural design model in ribbon representation. **d**, Close-up view showing the predicted interface of SmBiT (gold) and cage (blue). **e**, Close-up view showing the predicted interface of BIM (salmon) and cage (blue). **f**, Kinetic luminescence measurements after the addition of BCL-2 (200 nM) to a mixture of lucCageBIM (20 nM) and lucKey (20 nM). **g**, Tunable sensitivity of lucCageBIM to BCL-2 by changing the concentrations of sensor (lucCageBIM and lucKey) components (coloured curves).



Extended Data Fig. 3 | Functional screening of sCageHA designs and crystal structure of sCageHA_267-1S. **a**, Structural models of sCageHA designs with the embedded de novo binder HB1.9549.2. The HB1.9549.2 protein (cyan) was grafted into a parental six-helix bundle (sCage, yellow) at different positions along the latch helix (magenta), including three consecutive glycine residues (green). The black arrows indicate the additionally introduced single V255S (1S) or double V255S/I270S (2S) mutation(s) on the latch. **b**, Experimental validation of five sCageHA designs binding to HA in the presence or absence of the key by biolayer interferometry. The concentration of the sCages and the key were 1 μ M and 2 μ M, respectively. Each experiment was performed once. sCageHA_267-1S exhibited the highest fold of activation. **c**, Structural comparison showing the flexible nature of sCage to enable caging of HB1.9549.2. The structural model of sCage (grey) and the crystal structure of sCageHA_267-1S (gold) are

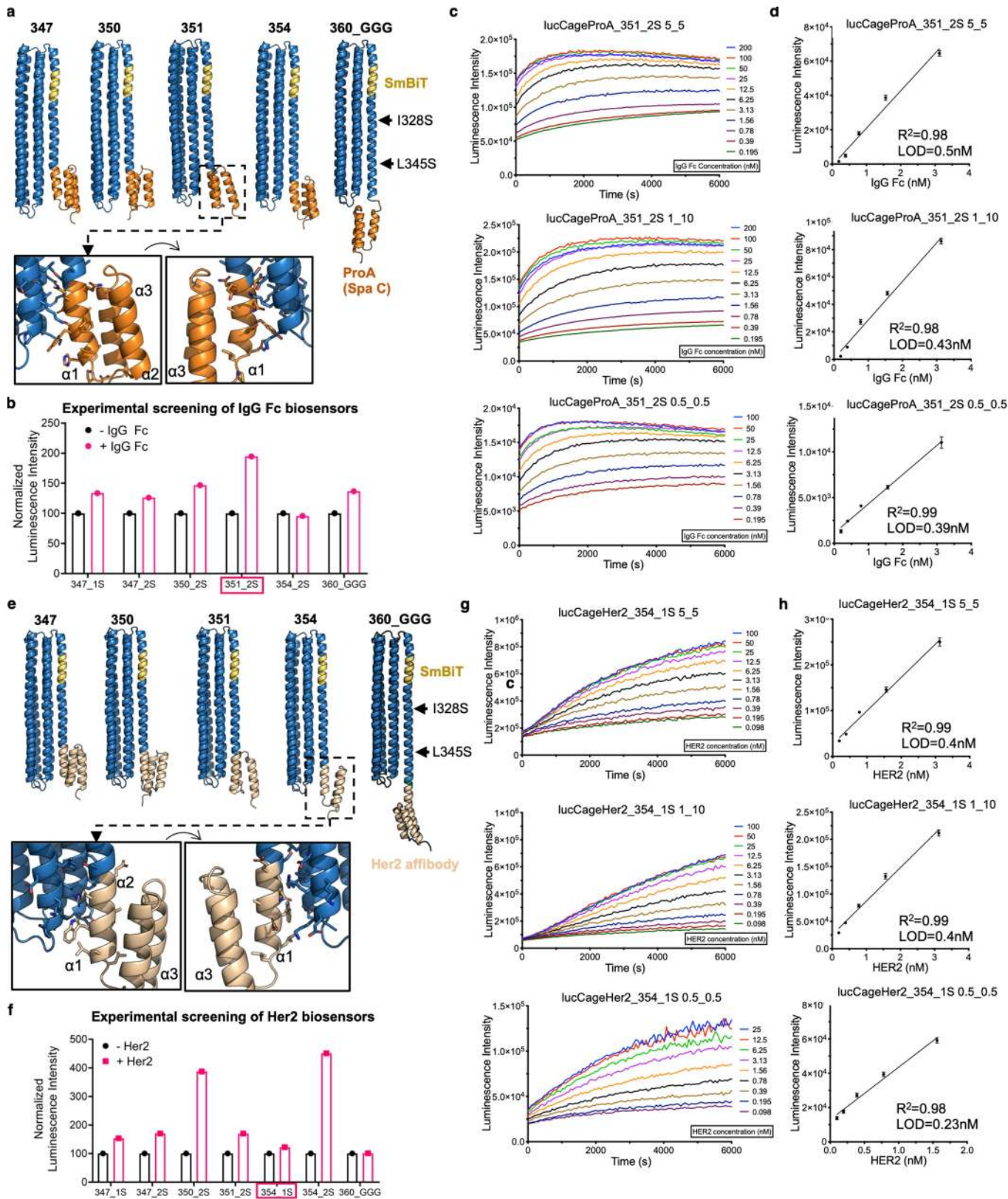
superposed, and a narrow section (black box) is shown in an orthogonal view for detail. The N-terminal helix of HB1.9549.2 is displaced from the latch helix (α 6) by 3.2 \AA (middle) with a concomitant displacement of α 5 and partial disruption of a hydrogen-bond network involving Q16 and N214 of sCage (right). **d**, A close-up view of the intramolecular interactions of sCageHA_267-1S. The HA-binding residues are highlighted in magenta. Both the N-terminal helix (cyan α 1) and the following helix (cyan α 2) of HB1.9549.2 interact with the cage. The intramolecular interactions are all hydrophobic. The bulky hydrophobic side chain of F285 tightly abuts against the backbone atoms of α 5 of sCage, which is unlikely to happen without a bending of α 5. Unfavourable interactions are also found: F273 is solvent-exposed, and the Y287 hydroxyl group is buried in the apolar environment. The rightmost panel shows the quality of the electron density map.



Extended Data Fig. 4 | Design and characterization of a BoNT/B sensor.

a, Structural models of the BoNT/B sensor designs showing the different threading positions of Bot.0671.2 (green, PDB code 5VID) on the latch of lucCage (blue). The SmBiT peptide is shown in gold ribbon representation. I328S and L345S indicate mutations introduced to tune the latch-cage interface ('IS' denotes I328S, '2S' denotes I328S/L345S)², and 'GGG' indicates the presence of three consecutive glycine residues between the latch and the grafted protein. The black box shows a close-up view of the interface of the cage (blue) and Bot.0671.2 (green) in the 349_2S design. **b**, Experimental screening of nine de novo BoNT/B sensors. Luminescence measurements were performed for each design (20 nM) and lucKey (20 nM) in the presence or

absence of the BoNT/B protein (200 nM). The luminescence values for each design were normalized to 100 in the absence of BoNT/B. Design 349_2S was selected as the best candidate owing to high sensitivity and stability, and was named lucCageBot. **c**, Determination of lucCageBot sensitivity. Bioluminescence was measured over 6,000 s in the presence of serially diluted BoNT/B protein. The concentrations of lucCageBot:lucKey (in nM) were 50:5, 5:5, 1:10 and 0.5:0.5 (from top to bottom). **d**, LOD calculations for the sensor at different concentrations. The concentrations of lucCageBot:lucKey (in nM) were 50:5, 5:5, 1:10 and 0.5:0.5 (from top to bottom). All experiments were performed in triplicate, representative data are shown, and data are mean \pm s.d.



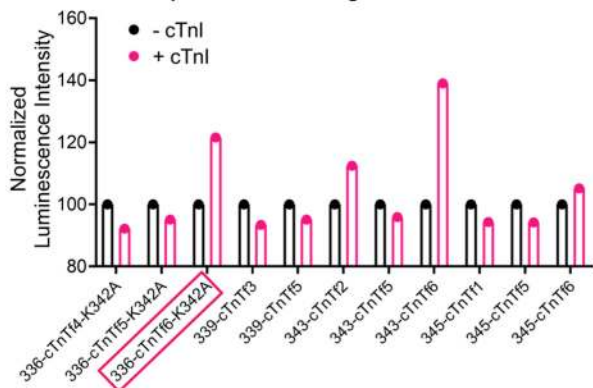
Extended Data Fig. 5 | See next page for caption.

Article

Extended Data Fig. 5 | Design and characterization of an Fc domain sensor (lucCageProA) and a HER2 sensor (lucCageHER2). **a**, Structural models of the Fc sensor designs showing the different threading positions of the *S. aureus* protein A domain C (orange, PDB code 4WWI) on the latch of lucCage (blue). The SmBit peptide is shown in gold ribbon representation. I328S and L345S indicate mutations introduced to tune the latch–cage interface as in Extended Data Fig. 4a. **b**, Experimental screening of six de novo Fc domain sensors. Luminescence measurements were performed for each design (20 nM) and lucKey (20 nM) in the presence or absence of recombinant human IgG1 Fc (200 nM). The luminescence values were normalized to 100 in the absence of Fc. Design 351_2S was selected as the best candidate owing to high sensitivity, and was named lucCageProA. This experiment was performed using single replicates in two independent instances, representative data are shown. **c**, Determination of the sensitivity of lucCageProA. Bioluminescence was measured over 6,000 s in the presence of serially diluted Fc protein. The lucCageBot:lucKey concentrations (in nM) were 5:5 (top), 1:10 (middle) and 0.5:0.5 (bottom). **d**, LOD calculations for the sensor at different concentrations. The lucCageBot:lucKey concentrations (in nM) were 5:5 (top), 1:10 (middle) and 0.5:0.5 (bottom). **e**, Structural models of the HER2 sensor

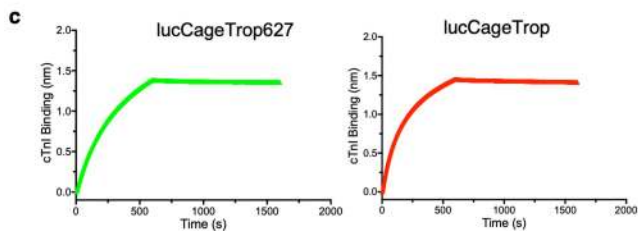
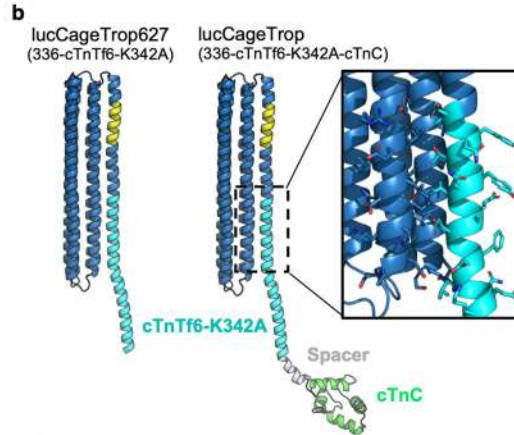
designs showing the different threading positions of the HER2 affibody protein (PDB code 3MZW, beige) on the latch of lucCage (blue), as in **a**. The black boxes show a close-up view of the interface of the cage (blue) and the HER2 affibody (beige) in the 354_2S design. **f**, Experimental screening of seven de novo HER2 sensors. Luminescence measurements were taken for each design (20 nM) and lucKey (20 nM) in the presence or absence of the ectodomain of HER2 (200 nM). The luminescence values were normalized to 100 in the absence of HER2 ectodomain. This experiment was performed using single replicates in two independent instances, representative data are shown. Design 354_2S was selected as the best candidate owing to high sensitivity and stability, and was named lucCageHER2. **g**, Determination of the sensitivity of lucCageHER2. Bioluminescence was measured over 6,000 s in the presence of serially diluted HER2 ectodomain protein. The lucCageBot:lucKey concentrations (in nM) were 5:5 (top), 1:10 (bottom) and 0.5:0.5 (bottom). **h**, LOD calculations for the sensor at different concentrations. The lucCageBot:lucKey concentrations (in nM) were 5:5 (top), 1:10 (middle) and 0.5:0.5 (bottom). All experiments were performed in triplicate unless specifically indicated, representative data are shown, and data are mean \pm s.d.

a Experimental screening of cTnI biosensors

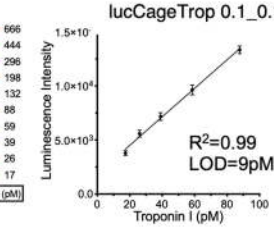
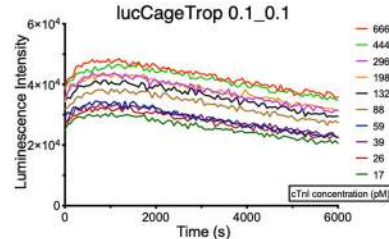
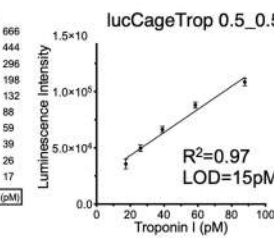
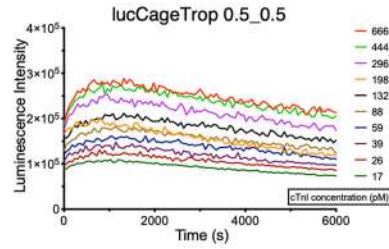
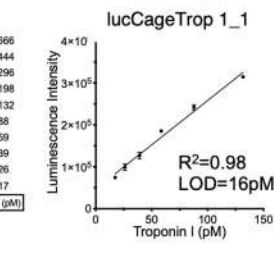
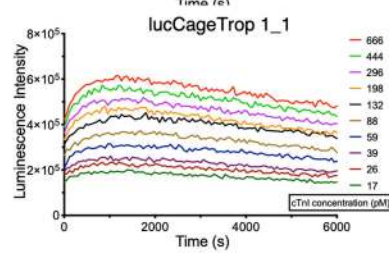
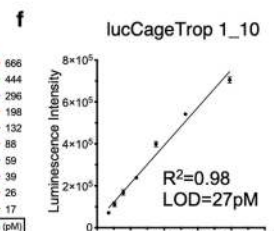
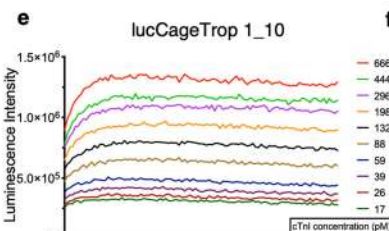
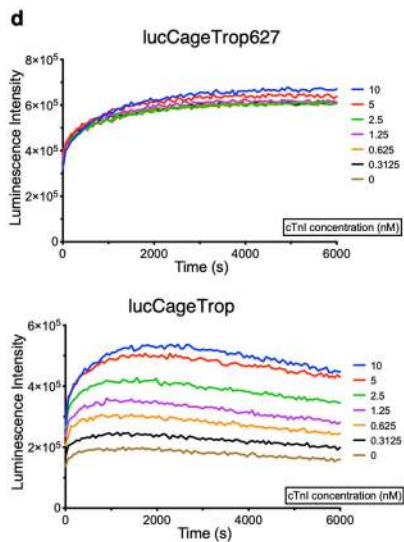


cTnTf1: 226-EDQLREKAKELWQTI-240
 cTnTf2: 226-EDQLREKAKELWQTIYN-242
 cTnTf3: 226-EDQLREKAKELWQTIYNLEAE-246

cTnTf4: 226-EDQLREKAKELWQTIYNLEAEKFD-249
 cTnTf5: 226-EDQLREKAKELWQTIYNLEAEKFDLQE-252
 cTnTf6: 226-EDQLREKAKELWQTIYNLEAEKFDLQEKFKQKYEINVLRRNRINDNQ-272



Sensor Name	K_{on} (1/Ms)	K_{off} (1/s)	K_D (nM)
lucCageTrop627	6.62×10^4	1.22×10^{-5}	0.19
lucCageTrop	1.03×10^5	2.78×10^{-6}	0.027

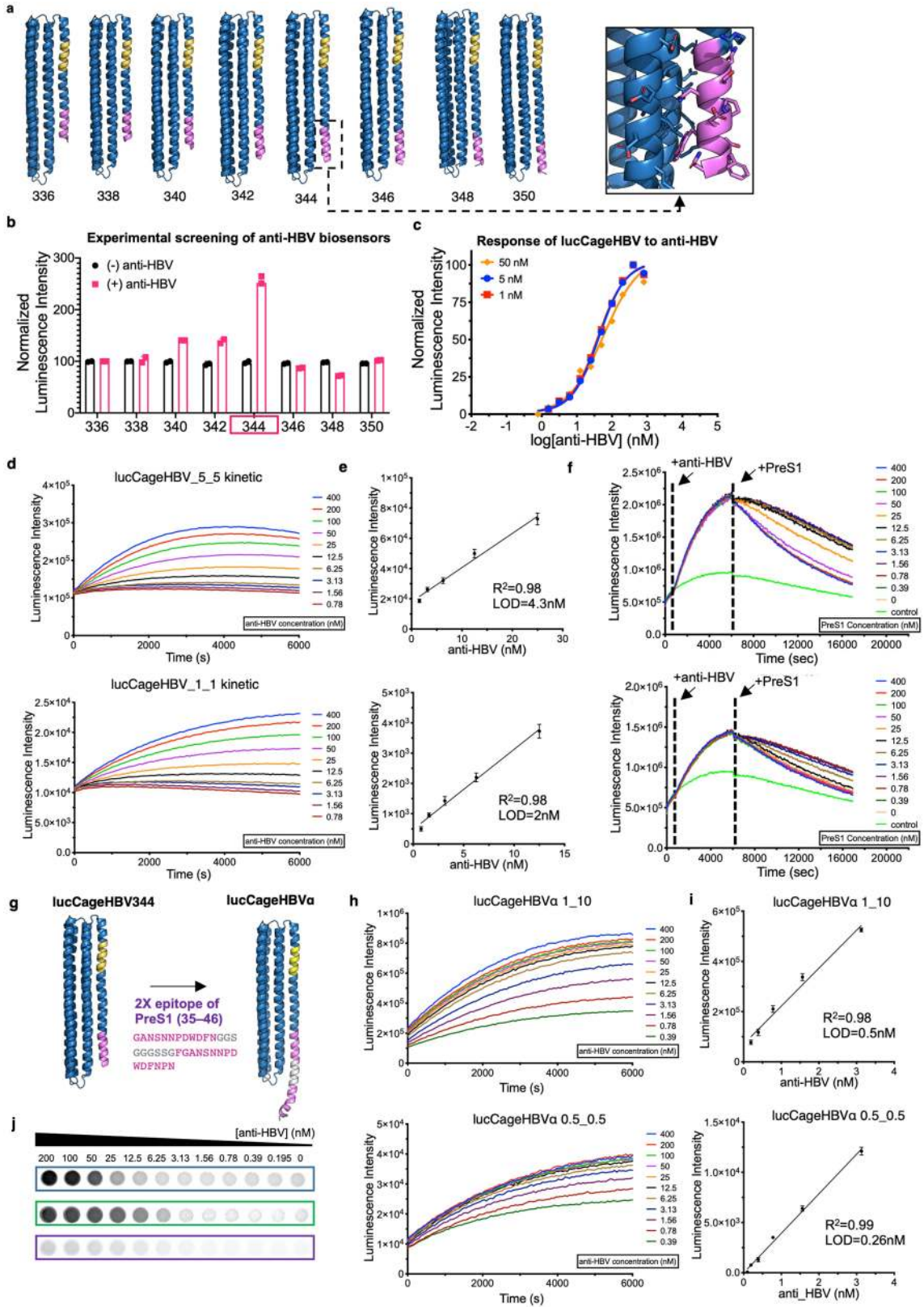


Extended Data Fig. 6 | See next page for caption.

Article

Extended Data Fig. 6 | Design, selection and engineering of lucCageTrop for cTnI detection. **a**, Experimental screening of designed sensors for cTnI. Fragments of cTnT, namely cTnTf1-f6, were computationally grafted into lucCage at different positions of the latch. All designs were produced in *E. coli* and experimentally screened at 20 nM and 20 nM lucKey for an increase in luminescence in the presence of 100 nM cTnI. The luminescence values were normalized to 100 in the absence of cTnI. This experiment was performed using single replicates in two independent instances, representative data are shown. Design 336-cTnTf6-K342A was selected as the best candidate (named lucCageTrop627) on the basis of its sensitivity, activation fold change and stability. **b**, Models of lucCageTrop627 and lucCageTrop—an improved version produced by fusion of cTnC at the C terminus of lucCageTrop627. The models are shown in ribbon representation comprising SmBit (gold) a fragment of cTnT (cyan) (PDB code 4Y99), and cTnC (green) (PDB code 4Y99). The black box

shows a close-up view of the interface of the cage (blue) and cTnT (cyan) in the lucCageTrop design. **c**, The binding affinity of lucCageTrop627 and lucCageTrop to cTnI was measured by biolayer interferometry. lucCageTrop showed sevenfold higher affinity to cTnI than lucCageTrop627. **d**, Comparison of bioluminescence kinetics between lucCageTrop627 (top) and lucCageTrop (bottom) in the presence of serially diluted cTnI. Higher binding affinity leads to improved dynamic range and sensitivity of the sensor. **e**, Determination of the sensitivity of lucCageTrop. Bioluminescence was measured over 6,000 s in the presence of serially diluted cTnI. The lucCageTrop:lucKey concentrations (in nM) were 1:10, 1:1, 0.5:0.5 and 0.1:0.1 (from top to bottom). **f**, LOD calculations for the sensor at different concentrations. The lucCageTrop:lucKey concentrations (in nM) were 1:10, 1:1, 0.5:0.5 and 0.1:0.1 (top to bottom). All experiments were performed in triplicate unless otherwise indicated, representative data are shown, and data are mean \pm s.d.



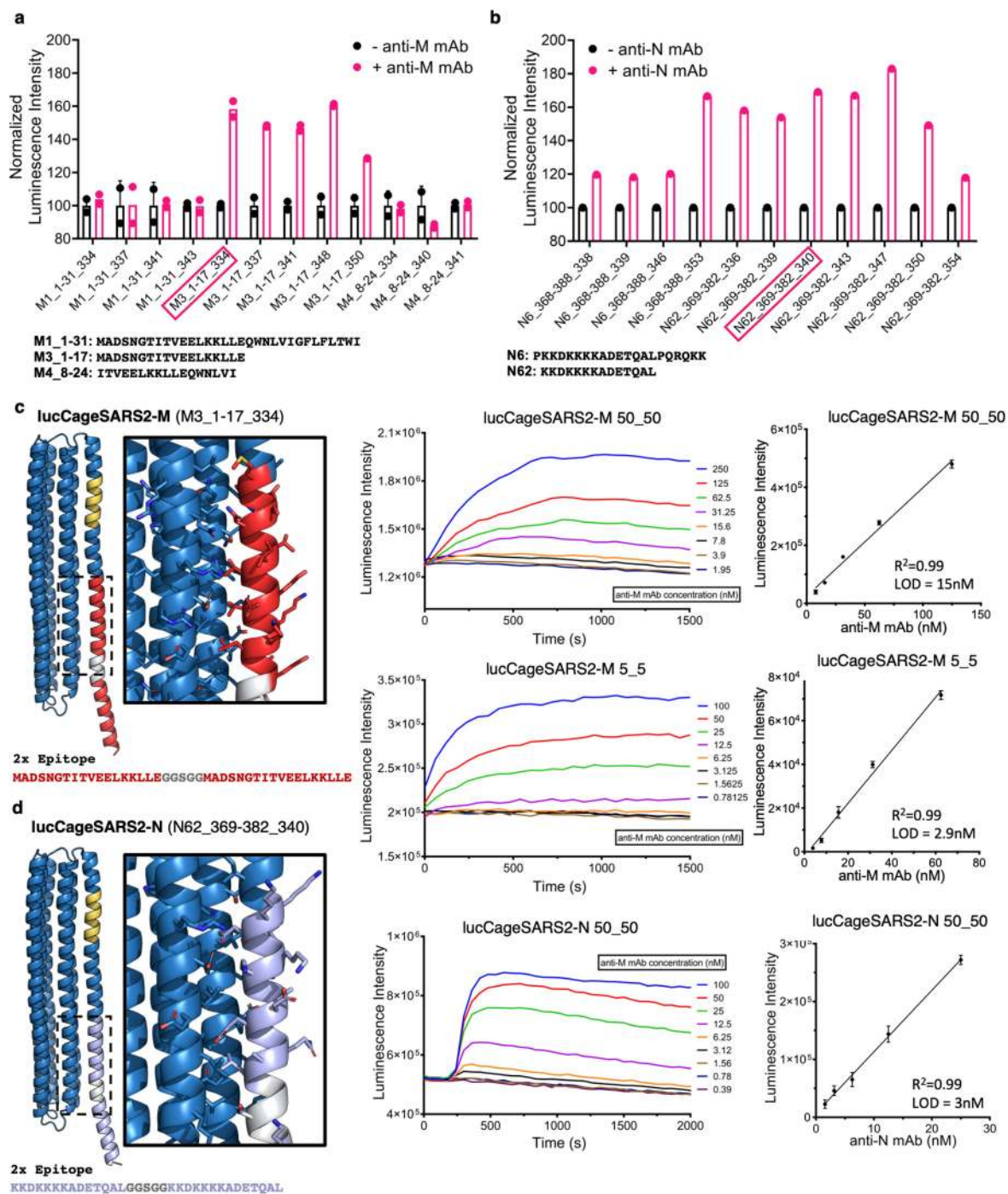
Extended Data Fig. 7 | See next page for caption.

Article

Extended Data Fig. 7 | Design and characterization of an anti-HBV antibody sensor.

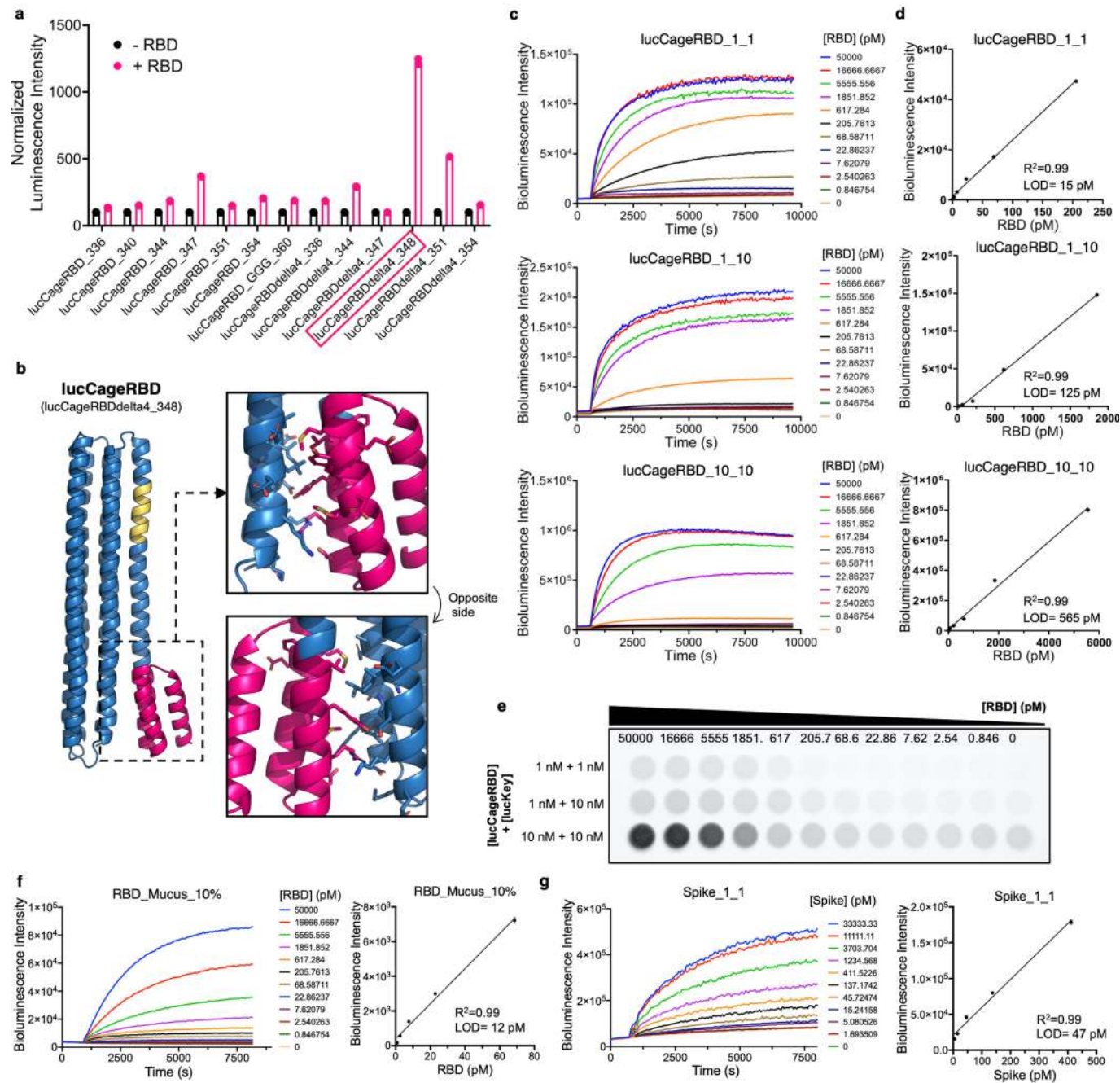
a, The energy-minimized models of lucCage designs are shown with the threaded segments of SmBiT (gold) and the antigenic motif of preS1 (magenta). The right box shows a close-up view of the cage-motif interface of the HBV344 design. **b**, Experimental screening of all designs performed by monitoring the luminescence of each lucCage (20 nM) and lucKey (20 nM) in the presence or absence of the anti-HBV antibody HzKR127-3.2 (100 nM). The luminescence values were normalized to 100 in the absence of anti-HBV. This experiment was performed in duplicate in two independent instances, and representative data are shown. The design HBV344 was selected owing to its better performance and was named lucCageHBV. **c, d**, Determination of lucCageHBV sensitivity. Bioluminescence was measured over 6,000 s in the presence of serially diluted HzKR127-3.2. The lucCageHBV:lucKey concentrations were 5:5 nM (top) and 1:1 nM (bottom). The maximum values of the curves in **d** are used to obtain the curves in **c**. **e**, LOD calculations for the sensor at different concentrations. The lucCageHBV:lucKey concentrations were 5:5 nM (top) and 1:1 nM (bottom). **f**, Detection of preS1 by competition of lucCageHBV344 and HzKR127-3.2 shown in Fig. 3f. Luminescence kinetics after the addition of the antibody (anti-HBV, first arrow). The anti-HBV antibody

concentrations were 50 nM (top) and 12.5 nM (bottom). At 6,000 s, different concentrations of the preS1 domain were injected into the wells (second arrow), and the decreased luminescence signals were used to detect preS1. **g**, Design of lucCageHBV α for improved detection of an anti-HBV antibody. The structural model of lucCageHBV α is shown with a close-up detail of the predicted interface between the preS1 epitope (magenta) and lucCage (blue). The design comprises two copies of the epitope preS1 (amino acids 35–46), spaced by a flexible linker (grey) to enable bivalent interaction with the antibody. The SmBit peptide is shown in gold. **h**, Determination of lucCageHBV α detection sensitivity to the presence of the antibody HzKR127-3.2 (anti-HBV). Bioluminescence was measured over 6,000 s in the presence of serially diluted HzKR127-3.2. The lucCageHBV α :lucKey concentrations (in nM) were 1:10 (top) and 0.5:0.5 (bottom). **i**, The linear region of a calibration curve was used to determine the LOD of antibody detection. **j**, Bioluminescence images acquired with a BioRad ChemiDoc imaging system. The lucCageHBV α :lucKey concentrations (in nM) are 50:5 (top), 5:5 (middle) and 1:10 (bottom). Changes in bioluminescence levels were detected as a function of the concentration of HzKR127-3.2. All experiments were performed in triplicate unless specifically indicated, and representative data are mean \pm s.d.



Extended Data Fig. 8 | Design and characterization of sensors for anti-SARS-CoV-2 antibodies. **a, b**, Experimental screening of de novo sensors for antibodies against the SARS-CoV-2 membrane protein (**a**) and the nucleocapsid protein (**b**). Selected epitopes of the membrane protein (M1, M3 and M4) and the nucleocapsid protein (N6 and N62) were computationally grafted into lucCage at different positions of the latch. Each design comprised two tandem copies of each epitope, separated by a flexible linker, to take advantage of the bivalent binding of antibodies. All designs were experimentally screened for an increase in luminescence at 20 nM of each lucCage design and 20 nM of lucKey in the presence of anti-M rabbit polyclonal antibodies (**a**) or anti-N mouse monoclonal antibody at 100 nM (clone 18F629.1) (**b**). These experiments were performed in duplicate (**a**) or single replicates (**b**) in two independent instances, and representative data are shown. The luminescence values were normalized to 100 in the absence of antibodies. Designs M3_1-17_334 and N62_369-382_340 were selected as the best candidates owing to high sensitivity and stability, and were named lucCageSARS2-M and

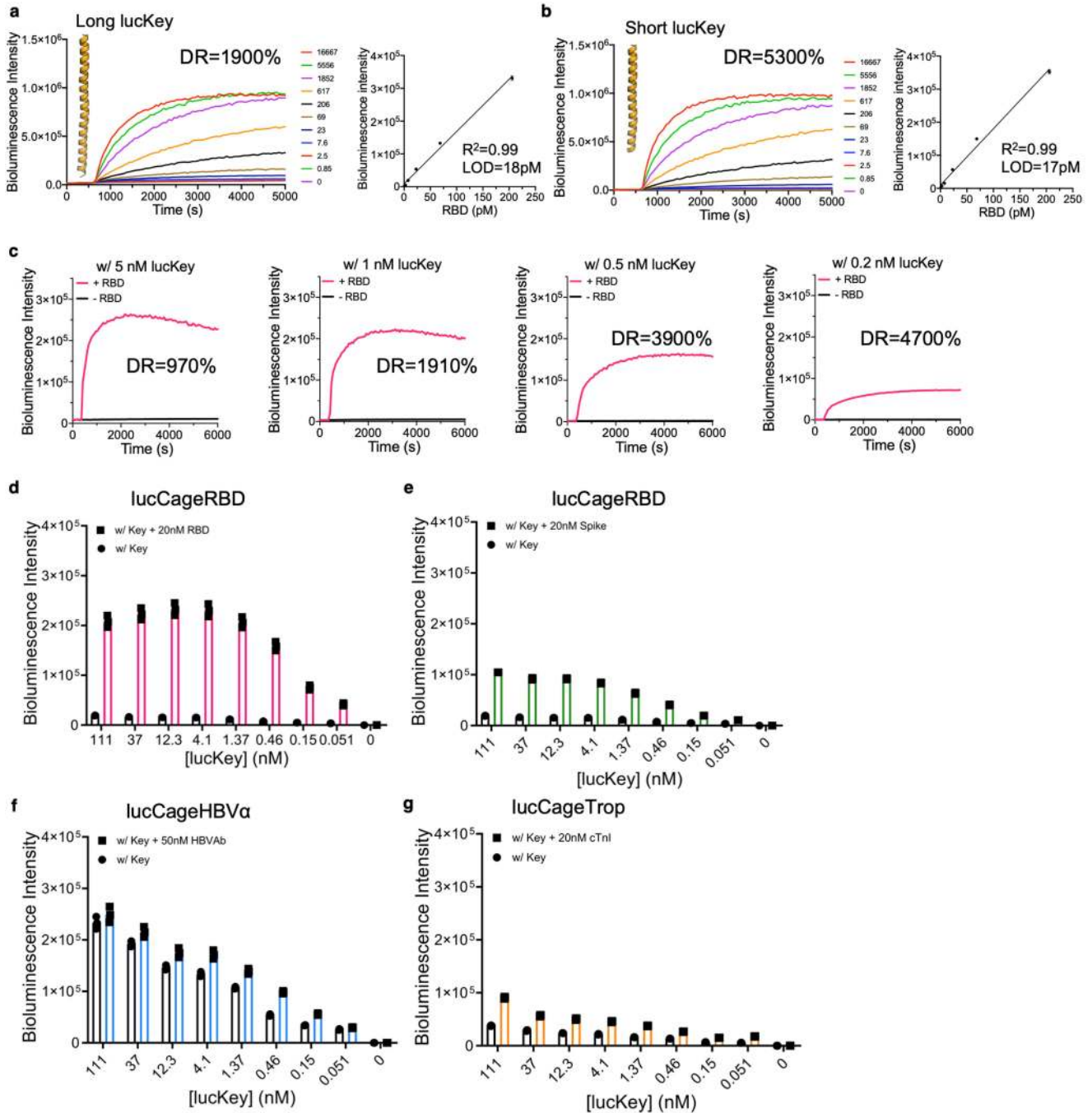
lucCageSARS2-N, respectively. **c**, Left, structural model of lucCageSARS2-M, showing a close-up view of the predicted interface between the M3 epitope (red) and lucCage (blue). Middle, determination of lucCageSARS2-M sensitivity to anti-M polyclonal antibody. Bioluminescence was measured over 4,000 s in the presence of serially diluted anti-M polyclonal antibody. The lucCageSARS2-M:lucKey concentrations (in nM) were 50:50 (top) and 5:5 (bottom). Right, LOD calculations for the sensor at different concentrations. **d**, Left, structural model of lucCageSARS2-N, showing a close-up view of the predicted interface between the N62 epitope (purple) and lucCage (blue). Middle, determination of lucCageSARS2-N sensitivity to anti-N monoclonal antibody. Bioluminescence was measured over 4,000 s for lucCageSARS2-N plus lucKey at 50 nM in the presence of serially diluted anti-N antibody. Right, LOD calculations for the sensor. All experiments were performed in triplicate unless specifically indicated, representative data are shown, and data are mean \pm s.d.



Extended Data Fig. 9 | Design and characterization of SARS-CoV-2 RBD sensors.

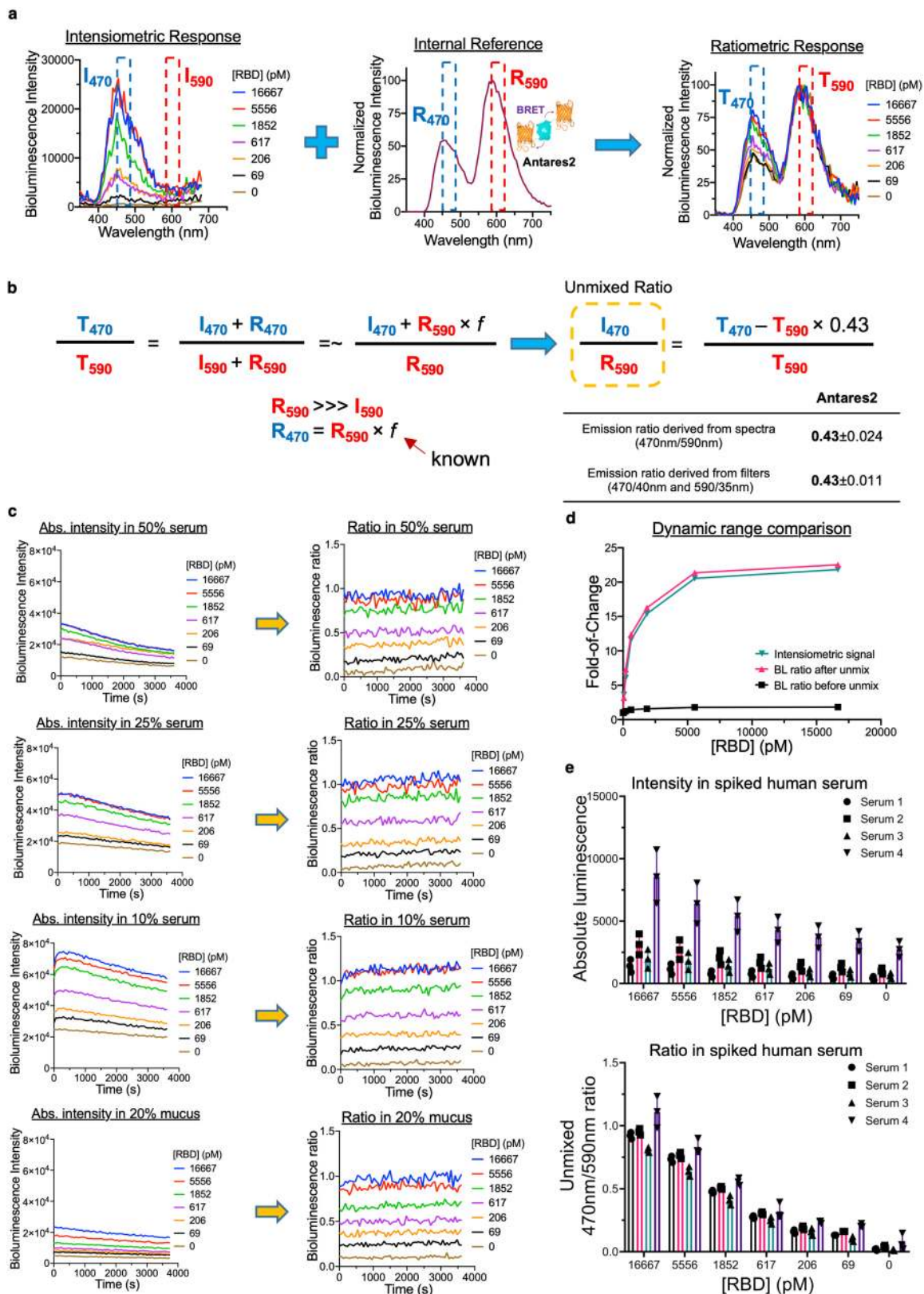
a, Experimental screening of de novo sensors for the RBD of the SARS-CoV-2 spike protein. All designs were experimentally screened for increases in luminescence at 20 nM of each lucCage design and 20 nM of lucKey in the presence of 200 nM RBD. The luminescence values were normalized to 100 in the absence of RBD. This experiment was performed in duplicate in two independent instances, and representative data are shown. Design lucCageRBDdelta4_348 was selected as the best candidate owing to high sensitivity and stability, and was named lucCageRBD. **b**, Structural model of lucCageRBD composed of the LCBI binder (magenta) grafted into lucCage (blue) comprising a caged SmBiT fragment (gold). The black boxes show a close-up view of the interface of cage (blue) and the LCBI binder (magenta) in the lucCageRBD design. **c**, Determination of the sensitivity of lucCageRBD. Bioluminescence was measured over 10,000 s in the presence of serially

diluted RBD protein. The concentrations of lucCageRBD:lucKey concentration (in nM) were 1:1 (top), 1:10 (middle) and 10:10 (bottom). **d**, LOD calculations for the sensor at different concentrations. The lucCageRBD:lucKey concentrations (in nM) were 1:1 (top), 1:10 (middle) and 10:10 (bottom). **e**, Bioluminescence images acquired with a BioRad ChemiDoc imaging system. Changes in bioluminescence levels were detected as a function of the concentration of RBD with 1 nM lucCageRBD and 10 nM lucKey. **f**, Detection of RBD in 10% simulated nasal matrix. Left, bioluminescence was measured overtime in the presence of serially diluted RBD protein. Right, LOD was calculated to be 12 pM. **g**, Detection of spike protein in a 20% diluted pooled serum. Left, bioluminescence was measured overtime in the presence of serially diluted HexaPro spike protein. Right, LOD was calculated to be 47 pM. All experiments were performed in triplicate unless otherwise indicated, representative data are shown, and data are mean ± s.d.



Extended Data Fig. 10 | lucCageRBD tunability by varying the lucKey length and concentration and the comparison of bioluminescent signals over a range of lucKey concentrations in the presence of target at saturating concentration. **a, b**, Experimental evaluation of the effect of lucKey length (K_{CK}) on the dynamic range (DR) of lucCageRBD to detect monomeric SARS-CoV-2 RBD. A truncated lucKey (short lucKey), 14 residues shorter than the full-length key at its C terminus (**b**), provides better dynamic range than the full-length lucKey (**a**) owing to reduced background signal, as predicted by the simulation in Extended Data Fig. 1f, whereas the LOD remains the same. **c**, The effect of lucKey concentration on the dynamic range. Decreasing the concentration of lucKey increases the dynamic range of lucCageRBD owing to reduced background signal, but with accompanying reduced maximum

bioluminescence signal. **d, e**, lucCageRBD (1 nM) was incubated with 20 nM RBD (**d**) or 20 nM spike protein (**e**), which are expected to result in full reconstitution of the luciferase activity. In the presence of spike protein, the same sensor was unable to yield the maximal bioluminescent signal, which suggests the effect of factors not captured by the simulations such as steric hindrance against complete luciferase reconstitution. **f**, lucCageHBV α (1 nM) incubated with 50 nM of the HBV antibody HzKR127-3.2 shows almost complete activation, but high background signal. **g**, lucCageTrop (1 nM) shows non-ideal background signal and moderate target-driven activation in the presence of 20 nM cTnI. All experiments were performed in triplicate, representative data are shown, and data are mean \pm s.d.



Extended Data Fig. 11 | See next page for caption.

Extended Data Fig. 11 | Integration of Antares2 as the internal reference for calibration of lucCageRBD in different biological matrices. **a**, The bioluminescent emission spectra of lucCageRBD (left) in response to varying concentrations of RBD. Antares2 is an efficient CyOFPI-teLuc-CyOFPIBRET system⁴⁰ with a peak emission at 590 nm (middle). The emission spectra were recorded from a mixture of lucCageRBD and lucKey (both at 1 nM), Antares2 (0.1 nM) and RBD at varying concentrations (right). By acquiring the individual signal from 470/40 nm and 590/35 nm channels, the intensimetric responses from lucCageRBD were converted into ratiometric readouts. **b**, Equations to calculate the spectrally unmixed ratio. The total signal from the 470/40 nm channel (T_{470}) is the sum of the signals from the lucCageRBD sensor (I_{470}) and the Antares2 reference (R_{470}), and the total signal from the 590/35 nm channel (T_{590}) is equal to the sensor signal (I_{590}) plus the reference signal (R_{590}). Because lucCageRBD gives negligible emission at the 590/35 nm channel, T_{590} is approximately equal to R_{590} ($R_{590} \gg I_{590}$). R_{470} is $R_{590} \times f$, a predetermined constant for Antares2, and therefore the unmixed ratio (I_{470}/R_{590}) could be calculated in real time during signal acquisition. The constant f for Antares2 was consistently determined to be 0.43 by either recording the full spectra or

from the filter set. **c**, Varying concentrations of RBD were spiked in 50%, 25% or 10% pooled serum or in 20% simulated nasal fluid. Absolute bioluminescence intensities and emission kinetics were different across the matrices owing to matrix inhibition effect and substrate turnover³³. By contrast, calibration with Antares2 resulted in stable ratiometric signals (I_{470}/R_{590}). **d**, The bioluminescence intensity of lucCageRBD at saturating RBD concentration (green curve) is approximately 20-fold higher than the background level. Reporting the raw ratio (T_{470}/T_{590}) as a function of the RBD concentration compromises the sensor dynamic range (black curve) owing to a notable emission at the 470/40 nm channel (R_{470}) from Antares2. After calculation and conversion of the unmixed ratio, the dynamic range becomes 20-fold higher than the background level with ratiometric readouts (magenta curve). **e**, Detection of spiked RBD in four different anonymized human sera (50%) shows that calibration using spectrally resolved Antares2 as an internal reference can minimize the variations of the intensimetric bioluminescence in these matrices. Bioluminescent signals and s.d. were measured in triplicate, and a representative one is shown for emission spectra and emission kinetics, respectively. Data are mean \pm s.d.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Bioluminescence data acquired on a Synergy Neo2 multi-mode microplate reader (Biotek). Biolayer interferometry data acquired on an Octet RED96 (ForteBio).

Data analysis Bioluminescence data were analyzed and plotted using GraphPad Prism 8. Target response curves were fitted using a Sigmoidal 4PL fit in GraphPad Prism 8. Limit of detection calculations were performed using Simple Linear regression in GraphPad Prism 8. BLI data was analyzed using ForteBio Data Analysis Software version 9.0.0.10 and plotted using GraphPad Prism 8. The model building and structure refinement were performed by using COOT and PHENIX. The design models and RosettaScripts code used in the manuscript have been deposited to http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/designcode_and_models.zip. The code for the numerical simulations shown in this manuscript are available at http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/model_simulation.py. All protein structure images were generated using PyMOL 2.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The atomic coordinates of sCageHA_267-1S have been deposited in the Protein Data Bank (<http://www.rcsb.org>) under an accession code 7CBC. The design models and RosettaScripts code used in the manuscript have been deposited to http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/designcode_and_models.zip. The code for the numerical simulations shown in this manuscript are available at http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/model_simulation.py. The original experimental data that support the findings of this work are available from the corresponding authors upon request. Plasmids encoding the biosensor proteins described in this article are available from the corresponding authors upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine the sample size. in vitro experiments were done in triplicate.
Data exclusions	No sample was excluded from data analysis
Replication	The results were successfully replicated using different batches of pure proteins on different days.
Randomization	De-identified clinical serum samples were randomly used for the detection of spiked target proteins.
Blinding	no blinding was employed since all experiments are in vitro.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

1. SARS-CoV Matrix Antibody (ProSci, Cat. No.: 3527)
2. SARS Nucleocapsid Protein Antibody (18F629.1) (NovusBio Cat. No. NBP2-24745)
3. HzKR127-3.2

Validation

1. SARS-CoV Matrix Antibody (ProSci, Cat. No.: 3527) is a rabbit IgG polyclonal antibody raised against a peptide corresponding to 13 amino acids near the amino-terminus of human SARS-CoV Matrix protein. The antibody is proven to bind the immunogen by ELISA. by the manufacturer. This antibody has predicted crossreactivity with SARS-CoV-2 Matrix protein based on immunogen sequence: human SARS-CoV2 Matrix protein: (identity 77%, homology 93%) by the manufacturer, which is confirmed in this work.
2. SARS Nucleocapsid Protein Antibody (18F629.1) is validated by Western Blot by the manufacturer. The antibody was developed by immunizing mice with a synthetic peptide corresponding to amino acids 354-385 from the N (SARS Nucleocapsid) for the Human

SARS coronavirus. Immunogen Percent Identity to SARS-CoV-2 Nucleocapsid Protein predicted to be 100% by the manufacturer and cross-reactivity confirmed in this work.

3. Validation of the antibody function is thoroughly described here: 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). The antibody was produced by Wi and Hong (Department of Systems Immunology, College of Biomedical Science, Kangwon National University, Chuncheon 200-701, Republic of Korea).

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HEK293F (Invitrogen; No. K9000-01)
Authentication	Cells were not further authenticated in the laboratory.
Mycoplasma contamination	HEK293F cells were tested negative for Mycoplasma by the provider, and it was not further confirmed in the laboratory
Commonly misidentified lines (See ICLAC register)	<i>Name any commonly misidentified cell lines used in the study and provide a rationale for their use.</i>

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Serum specimens were derived from excess plasma or sera from adults (>18 yo) of both genders kindly provided by the Director of the Clinical Chemistry Division, the hospital of University Washington. All anonymized donor specimens were provided de-identified.
Recruitment	the donors consented to have their excess specimens be used for other experimental studies, they could be transferred to our study without additional consent.
Ethics oversight	the Clinical Chemistry Division, the hospital of University Washington.

Note that full information on the approval of the study protocol must also be provided in the manuscript.