nature communications

Article

https://doi.org/10.1038/s41467-024-50956-z

Rapid discovery and evolution of nanosensors containing fluorogenic amino acids

Received: 14 March 2024

Accepted: 24 July 2024

Published online: 05 September 2024

Check for updates

Erkin Kuru[1,2,12] ✉, Jonathan Rittichier[1,2,9,12], Helena de Puig[2,3], Allison Flores[1,2], Subhrajit Rout[1,2], Isaac Han[2], Abigail E. Reese[4], Thomas M. Bartlett[5,10], Fabio De Moliner[4], Sylvie G. Bernier[2], Jason D. Galpin[6], Jorge Marchand[1,11], William Bedell[2], Lindsey Robinson-McCarthy[1], Christopher A. Ahern[6], Thomas G. Bernhardt[5,7], David Z. Rudner[5], James J. Collins[2,3,8], Marc Vendrell[4] ✉ & George M. Church[1,2] ✉

Binding-activated optical sensors are powerful tools for imaging, diagnostics, and biomolecular sensing. However, biosensor discovery is slow and requires tedious steps in rational design, screening, and characterization. Here we report on a platform that streamlines biosensor discovery and unlocks directed nanosensor evolution through genetically encodable fluorogenic amino acids (FgAAs). Building on the classical knowledge-based semisynthetic approach, we engineer ~15 kDa nanosensors that recognize specific proteins, peptides, and small molecules with up to 100-fold fluorescence increases and subsecond kinetics, allowing real-time and wash-free target sensing and live-cell bioimaging. An optimized genetic code expansion chemistry with FgAAs further enables rapid (~3 h) ribosomal nanosensor discovery via the cell-free translation of hundreds of candidates in parallel and directed nanosensor evolution with improved variant-specific sensitivities (up to ~250-fold) for SARS-CoV-2 antigens. Altogether, this platform could accelerate the discovery of fluorogenic nanosensors and pave the way to modify proteins with other non-standard functionalities for diverse applications.

The rapid development of fast, simple, and low-cost biosensors is essential in basic research, diagnostics, and microscopy[1]. Unlike fluorescent probes with always-on emission signals, fluorogenic scaffolds can convert affinity reagents like protein binders, RNA, or DNA aptamers to optical sensors with binding-activated fluorescence[2–5]. Such sensors allow direct and quantitative detection of different targets and find broad use in diagnostics and molecular imaging[2,6–14]. For example, sensors made by the noncovalent binding of fluorogenic probes to genetically encodable binders (e.g., RNA/DNA aptamers or protein-based tags[3–5,15–18]) allow wash-free imaging to study the

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. [3]Institute for Medical Engineering and Science and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]IRR Chemistry Hub and Centre for Inflammation Research, Institute for Regeneration and Repair, The University of Edinburgh, Edinburgh, UK. [5]Department of Microbiology, Harvard Medical School, Boston, MA, USA. [6]Department of Molecular Physiology and Biophysics, The University of Iowa, Iowa City, IA, USA. [7]Howard Hughes Medical Institute, Boston, MA, USA. [8]Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [9]Present address: EnPlusOne Biosciences Inc., Watertown, MA, USA. [10]Present address: Division of Genetics, Wadsworth Center, New York State Department of Health, Albany, NY, USA. [11]Present address: Department of Chemical Engineering, University of Washington, Seattle, WA, USA. [12]These authors contributed equally: Erkin Kuru, Jonathan Rittichier. ✉e-mail: erkin_kuru@hms.harvard.edu; marc.vendrell@ed.ac.uk; gchurch@genetics.med.harvard.edu

https://doi.org/10.1038/s41467-024-50956-z

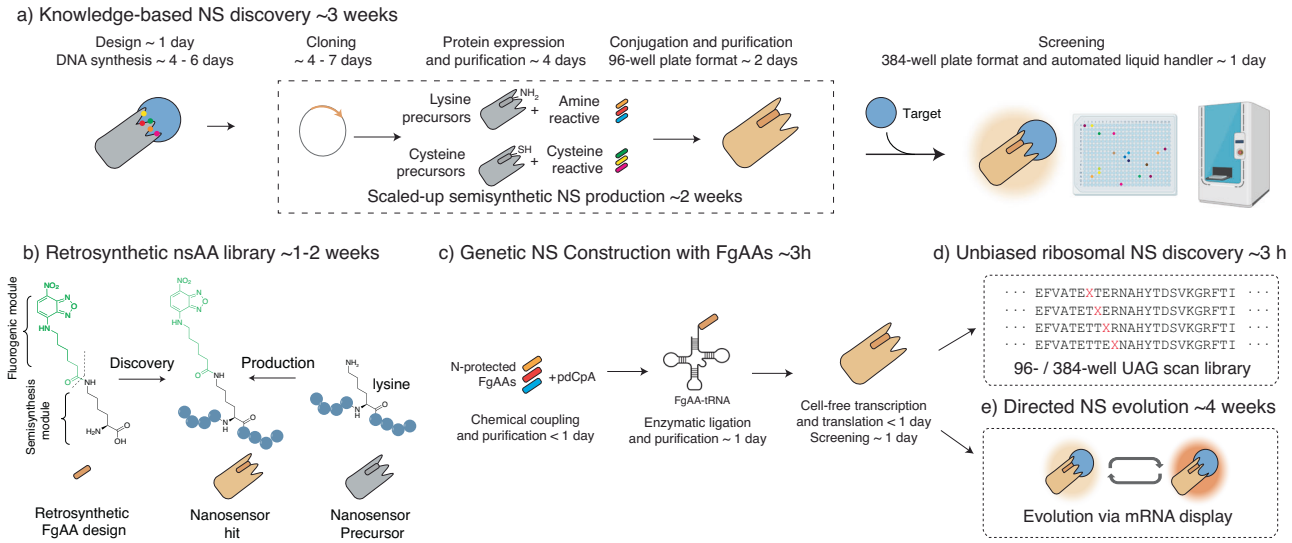


Fig. 1 | A platform to discover, evolve, and produce FgAA-containing NS. The multi-staged NS engineering platform enables rapid discovery, evolution, characterization, and cost-effective manufacturing of optical NS via FgAAs. **a** Hundreds of NS candidates are produced and screened in around 3 weeks by derivatizing protein binders with fluorogens. **b** This process also identifies key structural features for designing and synthesizing retrosynthetic FgAAs. **c** An optimized genetic code expansion chemistry enables cell-free, ribosomal production of functional NS by site-specific protein incorporation of FgAAs and simultaneous characterization of NS candidates without purification. **d** The genetic NS construction offers a rapid (~3 h) discovery strategy utilizing parallelized screening of nanobody scan libraries containing FgAAs (indicated as X) at every position. **e** Cell-free translation of NS also allows the development of a directed evolution pipeline to select improved NS for specific targets. Retrosynthetic nsAA designs allow hits from **a**, **d**, and **e** to be produced at scale in a two-step semisynthetic NS production approach (>20 mg from 1 L *Escherichia coli* culture). Figure created with help from BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en).

spatiotemporal dynamics of cellular components in living cells. However, such tags can be sizable and difficult to adapt for the detection of some targets (e.g., small molecules) due to the limited available binders. As the development of protein-based binders against a plethora of targets is expanding, the transformation of small binders into nanosensors (NS) can address this problem[19,20]. However, the engineering of high-contrast protein biosensors has been limited by suboptimal chemical conjugation of fluorogenic probes and low-throughput screening, both of which also restrict directed evolution approaches[1,2,6–8,10,21].

Building on small protein binders and genetically encodable fluorogenic amino acids (FgAAs), here we design a pipeline to rapidly discover, evolve, and produce protein-based NS (Fig. 1). These NS can detect their targets by emitting binding-activated fluorescence signals in complex environments including live cells. High contrast NS hits can be discovered quickly (e.g., ~3 weeks) by the multiplexed exploration of hundreds of candidates, which can be scaled up by derivatizing <15 kDa protein binders with cysteine- and lysine-reactive fluorogens (Fig. 1). This step also informs on privileged FgAA designs that can be site-specifically encoded to ribosomally construct functional NS via an optimized genetic code expansion chemistry allowing high-throughput characterization of NS candidates without purification. The fully genetic NS construction also unlocks both a rapid (~3 h) discovery approach that relies on ribosomal screening of nanobody scan libraries and the evolution of NS for different targets from large libraries. By streamlining the engineering of optimized and cost-effective sensors, this chemical synthetic biology platform can facilitate the adaptation of NS to address challenges in biomolecule sensing, including microscopy-based analysis of dynamic cellular processes.

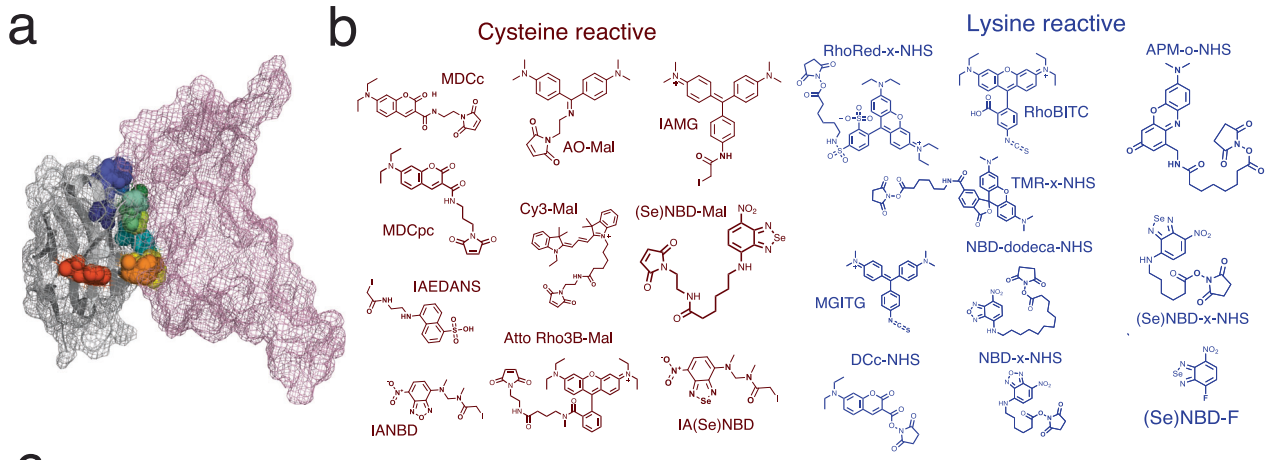
Results and discussion

A streamlined pipeline to discover binding-activated NS via retrosynthetic FgAAs

Building on previous fluorescent/fluorogenic non-standard amino acids (nsAA) and genetic code expansion technologies, we aimed to establish a genetically encoded biosensor discovery and optimization pipeline[9,22–26]. As the first case study, we focused on the well-characterized nanobody VHH72 that binds SARS-CoV-2 spike Receptor Binding Domain from the original CoV-19 strain (RBD_{CoV19}) (Fig. 2a)[27]. Despite its moderate RBD_{CoV19} affinity (~40 nM), VHH72 served as a good starting point for numerous biosensors because its RBD epitope is relatively conserved across sarbecoviruses[28]. We began by substituting a tryptophan residue at the binding interface with a genetically encodable bulky fluorescent amino acid L-(7-hydroxycoumarin-4-yl)ethylglycine (Cou). Although Cou has fluorogenic properties[29,30] and can generate antibody-based biosensors in other contexts[9], the recombinant VHH72 W108Cou variant exhibited negligible fluorescence increases in the presence of saturating RBD_{CoV19} (Supplementary Fig. 1a). This failed attempt highlighted a key challenge in optical biosensor engineering, that is, the need to explore a vast search space for optimal probe, linker, and position combinations that cannot be easily met by the currently limited repertoire of genetically encodable FgAAs.

Inspired by classical knowledge-based semisynthetic biosensor engineering approaches[6–8,10] we decided to establish a modular NS-FgAA discovery platform (Fig. 1). Aided by the crystal structure of VHH72 (Fig. 2a)[27], we designed nanobody variants where the binding interface residues were substituted by cysteines in wild-type VHH72 or by lysines in a lysine-less variant (VHH72 K43R; K65R; K76R; K87R, or VHH72 NoK) with native-like affinity for RBD_{CoV19} (Supplementary Table 1). By functionally reassigning nanobody framework lysines with arginines, lysines can be reintroduced at specific sites for modification with readily available amine-reactive reagents. This extended chemical flexibility to modify either cysteine or lysine variants allowed us to sample a broad range of fluorophores with different chemical reactivities (e.g., maleimides or iodoacetamides for cysteine variants; N-hydrosuccinimidyl (NHS) esters or isothiocyanates for lysine variants), environmental sensitivity (e.g., fluorescent tetramethylrhodamine vs. fluorogenic Malachite Green[18]), variable emission wavelengths (e.g., green fluorescent IANBD vs. blue fluorescent

https://doi.org/10.1038/s41467-024-50956-z

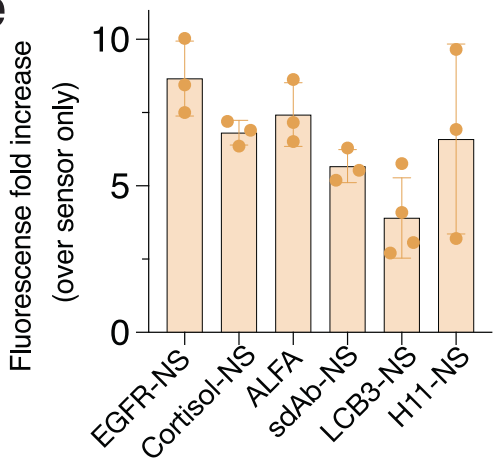


c

VHH72 cysteine substitutions

Thiol reactive probes	wt	I51C	W53C	G56C	Y59C	G102C	T103C	V104C	V105C	W108C	Y110C	W115C
Cy3-Mal	1.3	1.2	1.6	1.7	1.2	1.7	1.4	1.3	1.3	1.4	1.5	1.1
IAEDANS	1.8	1.8	1.	2.0	1.8	2.4	2.6	2.9	1.5	2.1	1.5	1.1
MDCc	1.3	1.7	3.0	8.3	4.5	2.0	1.5	5.6	4.1	1.2	1.8	2.1
IANBD	1.8	1.2	2.3	1.6	4.9	2.6	2.1	2.7	2.2	1.9	2.1	1.2
IA(Se)NBD	1.1	1.3	1.6	2.1	1.2	1.1	1.2	3.6	1.2	1.6	1.2	0.7
(Se)NBD-Mal	1.7	1.9	1.9	3.3	1.6	1.7	1.2	2.5	1.9	1.3	1.8	1.5
AO-Mal	1.0	1.1	1.7	1.3	1.2	1.1	1.1	4.0	1.3	1.4	1.3	1.3
IAMG	2.0	1.9	2.0	1.4	1.6	1.7	1.9	3.0	3.2	4.4	0.9	2.2

e



d

VHH72 lysine substitutions

Amine reactive probes	wt	NoK wt	NoK W53K	NoK V104K	NoK V105K	NoK W108K	NoK V110K	NoK W115K
TMR-x-NHS	1.1	1.4	1.0	1.3	1.0	1.0	1.0	1.8
NBD-x-NHS	1.9	1.4	1.9	13.4	1.5	2.8	1.8	1.4
NBD-dodeca-NHS	1.8	1.9	1.4	1.9	1.8	1.6	1.5	1.5
(Se)NBD-x-NHS	1.6	1.4	1.3	7.1	1.6	2.4	1.6	1.6
(Se)NBD-F	1.4	0.6	1.3	1.4	1.3	0.9	1.2	1.2
APM-o-NHS	1.0	1.1	1.2	1.3	1.3	1.3	1.3	1.4
MGITC	1.7	2.2	3.4	2.0	1.8	0.7	1.8	3.6

f

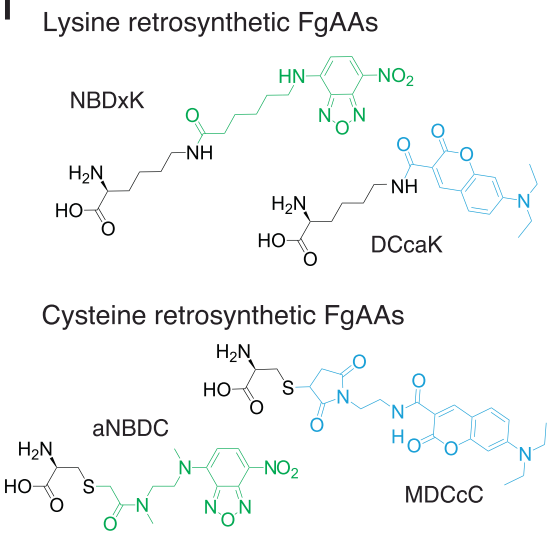


Fig. 2 | A streamlined knowledge-based pipeline to discover and produce binding-activated NS via retrosynthetic FgAAs. a Crystal structure of the VHH72 nanobody (gray) bound to its target, RBD_{CoV19} (purple). The VHH72 cysteine or lysine substitutions are color-coded in **a**, **c**, and **d**. **b** Structures and abbreviated names of reactive fluorogenic probes used in this work. **c**, **d** Fold fluorescence increases from the screening of the VHH72 variants conjugated with the library of cysteine (**c**) and lysine (**d**) reactive probes. **e** This knowledge-based discovery approach is generalizable to other protein scaffolds and molecular targets. Maximum fold increases of different NS in the presence of their corresponding antigens over buffer controls: H11-NS is a biosensor for RBD_{CoV19}, sdAb-NS for SARS-CoV-2 nucleocapsid, LCB3-NS for RBD_{CoV19}, ALFA-NS for ALFA peptide, EGFR-NS for human epidermal growth factor receptor (EGFR), Cortisol-NS for the small molecule cortisol. Bars and error bars represent the average and standard deviation of independent triplicate experiments. **f** Chemical structures of privileged retrosynthetic FgAAs. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-50956-z

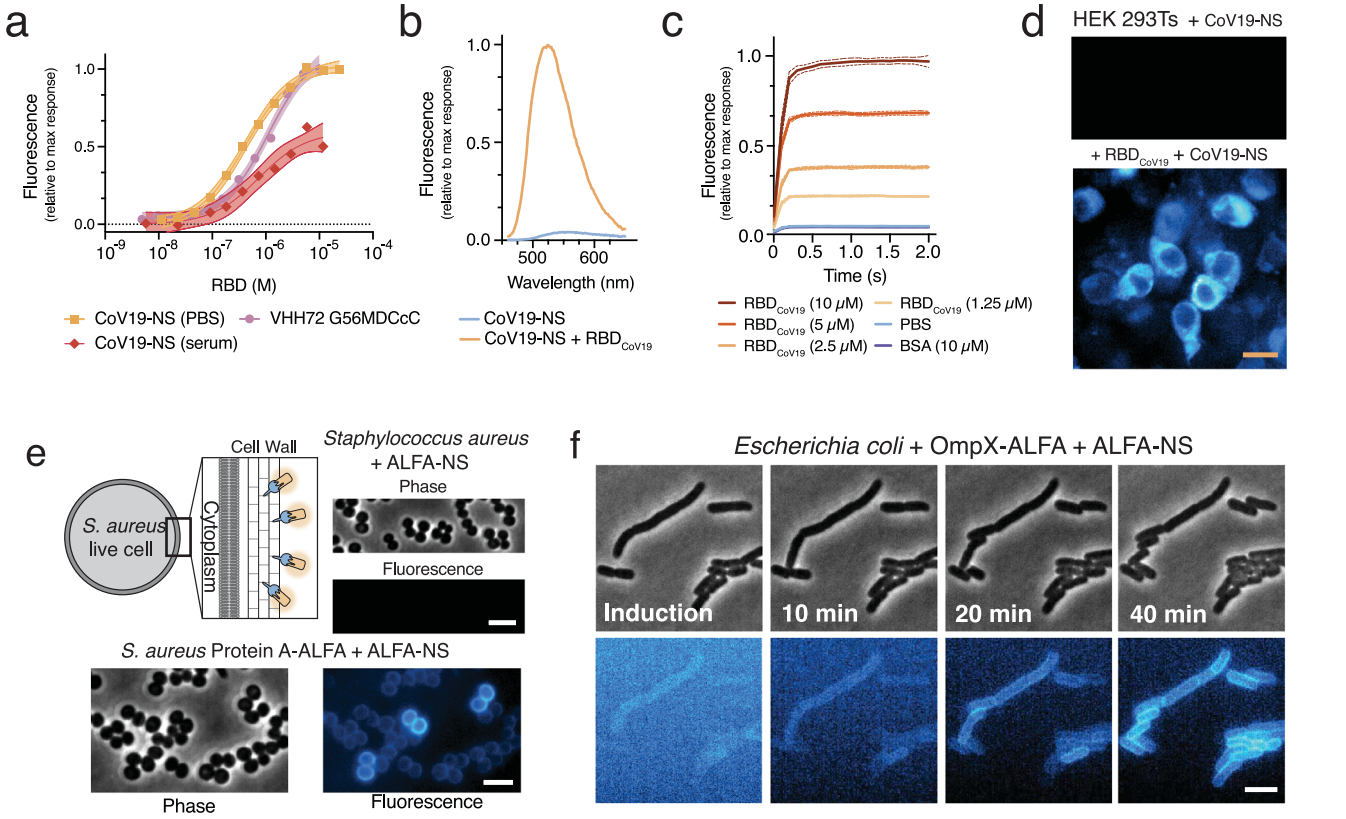


Fig. 3 | Cov19-NS and ALFA-NS enable rapid detection of specific antigens and their imaging in live cells. a CoV19-NS detects RBD_{CoV19} in both serum and PBS. The dose-response curve of the second-best performing CoV19 nanosensor, VHH72 G56MDCcC is like that of CoV19-NS. Each of the relative fluorescence units in the graph was normalized relative to the maximum response of CoV19-NS. **b** Upon subtraction of buffer blank values, CoV19-NS exhibited >100-fold fluorescence increases at $\lambda_{exc}/\lambda_{em}$ 420 nm/492 nm after binding RBD_{CoV19}. Raw fluorescence values are shown in Supplementary Fig. 4. **c** CoV19-NS ratiometrically detects its target in <1 second with no significant fluorescence increases observed in PBS or BSA. Lines and dashed lines represent the average and standard deviation of independent triplicate experiments. **d** CoV19-NS enables wash-free imaging of fixed and permeabilized HEK 293 T cells specifically expressing RBD_{CoV19} in situ. Scale bar: 20 µm. **e**, **f** ALFA-NS can be used for live-cell imaging of ALFA-tag labeled protein A in *Staphylococcus aureus* (**e**) or real-time imaging of OmpX export in *E. coli* (**f**). Scale bar: 2 µm. Dots represent independent measurements. Lines represent a 4PL fit of the dose-response curves. Shaded areas and dashed lines represent the 95% confidence intervals of the fits. Bars and error bars represent the average and standard deviation of independent triplicate experiments. Data from **d**–**f** are representative of two, and three biological replicates, respectively. Data from **d**–**f** are representative of more than triplicate independent measurements. The microscopy images are adjusted so that a quantitative comparison can be made within each panel. Source data are provided as a Source Data file.

MDCc), spacer lengths (e.g., NBD-hexanoate vs. NBD-dodecanoate), or different fluorogenic modes (e.g., photoactivatable NBD fluorophores[31], molecular rotors, IAMG[17] or AO-Mal[32], and the solvatochromic, APM-o-NHS[33], Fig. 2b). Altogether, we designed a workflow to rapidly access hundreds of lysine- and cysteine-substituted combinations and screen them in an automatable, high-throughput format (Fig. 1, Supplementary Tables 2–3). Specifically, we first purified cysteine/lysine-modified nanobody variants from *E. coli* (processing up to 14 variants at once) by autoinduction (i.e., 1–2 mg nanobody from 50 mL culture, sufficient for the entire screening) (Supplementary Fig. 1c). Then, we site-selectively modified the nanobody variants with the above-mentioned fluorescent building blocks and measured their fluorescence fold increase (ΔR_{max}) in the presence of RBD_{CoV19} (Supplementary Figs. 1d and 2b–d). The screening highlighted several VHH72 variant-probe combinations with significant fluorescence fold increases, such as VHH72 G56C modified with MDCc, or VHH72 NoK V104K modified with NBD-hexanoate (NBD-x), which we named as CoV19-NS (Fig. 2c, d).

To demonstrate the generalizability of this workflow, we applied our screening pipeline to other binders, and engineered NS for a wide range of antigens, specifically (1) the nanobody H11-H4 against SARS-CoV-2 RBD[34], (2) the miniprotein LCB3 against the SARS-Cov-2 spike protein RBD[35], (3) the nanobody sdAb-B6 against the SARS-Cov-2 nucleocapsid protein[36], (4) the nanobody NbALFA against the genetically encodable ALFA-tag peptide[37], (5) the nanobody EgA1 against EGFR[38], and (6) the nanobody NbCor against cortisol[39]. The multiplexed screening of ~1000 candidates (Supplementary Tables 2 and 3) resulted in NS with variable dynamic ranges determined by fluorescence dose-response curves (Figs. 2e and 3a, Supplementary Fig. 2, Supplementary Table 4). While the EC_{50} values of these initial NS were poorer than the published K_D values for the parent protein binders (Supplementary Table 1), their binding affinity remained sufficient for successful screening and wash-free live cell microscopy (Fig. 3) This confirms that this workflow remains applicable to a wide range of binders and antigens for initial high-contrast NS discovery. This screen also helped us identify privileged mature fluorogenic entities that we refer to as retrosynthetic FgAA designs that can convert a protein binder into an optical NS when incorporated into the right polypeptide context (Fig. 1). These FgAAs included NBDxK (the FgAA in CoV19-NS), DCcaK (the FgAA in H11-NS), MDCcC (the FgAA in VHH72 G56MDCcC or in previously reported biosensors[21,40]), and aNBDC (the FgAA in LCB3-NS, ALFA-NS, and Cortisol-NS, or in previously reported sensors[8,41,42]) (Fig. 2f).

NS enables rapid detection of specific antigens and their imaging in live cells. Next, we decided to characterize the performance of CoV19-NS. We first corroborated the presence of NBDxK, at the V104K position by mass spectrometry (Supplementary Fig. 3). Given that this analysis revealed significant non-specific N-terminal amine

modification, we enzymatically cleaved N-terminal NBDx to isolate an optimal CoV19-NS preparation that responded to RBD_{CoV19} with substantial raw fluorescence increase values up to ~100-fold when blanked (Fig. 3b, Supplementary Fig. 4). CoV19-NS showed a K_D value of ~300 nM for RBD_{CoV19} (Supplementary Table 1) and EC_{50} values of ~420 nM in buffer and ~790 nM in human serum (and a dynamic range of 10^{-7} M–10^{-5} M) (Fig. 3a, Supplementary Table 4). CoV19-NS detected other SARS-CoV-2/1 RBD variants, but not the RBD from MERS-CoV, indicating that CoV19-NS preserves the specificity of the parent nanobody (Supplementary Fig. 5)[27]. Moreover, CoV19-NS displayed a rapid response for RBD_{CoV19} within 500 milliseconds (Fig. 3c, Supplementary Fig. 6, Supplementary Movies 1–3). The combination of the high fluorogenicity, specificity, rapid binding kinetics, and the small size of NS suggested their application as wash-free imaging reagents for specific cellular targets, which we confirmed first by in situ localization microscopy in fixed cells expressing SARS-CoV-2 spike (Fig. 3d, Supplementary Fig. 7a)[43].

We also tested if NS could serve as reagents for wash-free live cell imaging focusing on cell-envelope proteins. Illuminating where and when dynamic surface proteins are exported in live cells cannot be easily addressed by classical labeling approaches due to poorly secreted fluorescent protein fusions[44] or slow pulse-chase labeling experiments with fluorescent antibodies that prohibit real-time data acquisition[45–47]. As proof of principle, we leveraged ALFA-NS, because its target, the 13 amino acid-long ALFA tag, can be genetically fused to desired proteins with minimal functional perturbation and therefore can be adapted to various systems[37]. First, we showed that ALFA-NS can be used as a live-cell stain to directly image a representative surface protein in the human pathogen *Staphylococcus aureus*, a Gram-positive bacterium, constitutively expressing an ALFA-tagged Staphylococcus protein A fusion (Fig. 3e, Supplementary Fig. 7b). Next, we fused the ALFA-tag to *E. coli* outer-membrane protein X (OmpX), a conserved Gram-negative envelope protein, expressed under an inducible promoter. Wash-free labeling with ALFA-NS showed an induction-dependent, uniform OmpX distribution on the envelope of *E. coli* BL21 cells consistent with previous localization studies (Supplementary Fig. 7c)[48]. Finally, we also added ALFA-NS directly to the growth medium and performed time-lapse fluorescence microscopy with these cells. Upon induction, the uniform ALFA-NS signal on the cell surface gradually increased indicative of the homogenous OmpX export, demonstrating that NS can be utilized to image native protein secretion in real-time with minimal cellular perturbation (Fig. 3f).

Cell-free site-specific translation of FgAAs can accelerate the discovery and functional characterization of NS

Even though this NS discovery strategy showed enhanced modularity, it still displayed limited throughput, given the need for the design, cloning, purification, and screening of individual variants. Therefore, we envisaged that cell-free, genetic incorporation of privileged FgAAs as a pre-conjugated moiety could increase the throughput of NS discovery and characterization as recently shown for the cell-free high-throughput screening of antibodies[49].

To construct NS ribosomally, we decided to chemically acylate FgAAs to an amber-decoding orthogonal $tRNA_{CUA}$ via a key pdCpA dinucleotide intermediate that could be encoded into an in-frame amber-containing nanobody construct (e.g., VHH72 V104UAG) in vitro (Fig. 1). Although various nsAAs have been successfully encoded using similar approaches[50–53], traditional procedures involve multiple purification steps with low recovery yields. We optimized the chemical synthesis of FgAA-pdCpA conjugates (Fig. 4a) with a straightforward protocol, involving activation of Boc-protected FgAAs with 1,1′-carbonyldiimidazole in dry dimethylformamide followed by immediate mixing with pdCpA in water at pH ~8 and solid phase extraction-based purification. We first validated our approach by encoding the standard fluorescent amino acid p-(BODIPY-FL)-aminophenylalanine (BDPaF)[53], into different positions of proteins of various sizes via cell-free protein synthesis (CFPS) (Supplementary Figs. 8a, 9, 10).

Next, we applied the protocol to retrosynthetic FgAAs[31,54,55] (modularly synthesized from Boc-protected lysine or via a Boc-cysteine-pdCpA diversification strategy) and validated their site-specific protein incorporation by MS and/or by in-gel fluorescence (Supplementary Figs. 8b, 11–13)[28,29]. We focused on NBDxK and aNBDC as FgAAs that share the nitrobenzodioxazole fluorogenic core[31], and tested if the target binding of the NS could be detected within the cell-free translation reaction mixture. Notably, the site-specific incorporation of NBDxK to VHH72 V104UAG or of aNBDC to NbCor T53UAG and to NbALFA M63UAG reconstituted the active CoV19-NS or Cortisol-NS, and ALFA-NS respectively. These ribosomally constructed fluorogenic NS identified their targets directly in the cell-free translation mixture, despite the excess FgAA-$tRNA_{CUA}$ species, whereas a control fluorescent nanobody with BDPaF (NbALFA M63BDPaF) lacked a fluorescence response to its target peptide (Supplementary Fig. 14a–c). Moreover, cell-free translated ALFA-NS (NbALFA M63aNBDC) could directly be applied to image OmpX-ALFA in live *E. coli* without purification (Fig. 4b). Finally, ribosomal construction of CoV19-NS or ALFA-NS in the presence of their respective targets resulted in direct target detection in real-time as the nascent NS are being translated (Fig. 4c, Supplementary Fig. 14d).

Next, we tested whether this increased throughput could be utilized to discover NS via unbiased substitution of every residue on a nanobody with a FgAA. For this, we encoded NBDxK into a DNA TAG scan library of VHH72 in the presence of RBD_{CoV19}. Within 3 hours of receiving the synthetic DNA, the real-time screening of 127 nanosensor scan variants (VHH72 Q1NBDxK–VHH72 S127NBDxK, each expressed in comparable levels in separate 5 µL cell-free protein synthesis reactions) recapitulated CoV19-NS (and VHH72 W108NBDxK from Fig. 2d) and identified additional NS hits (Fig. 4d, Supplementary Fig. 15). Surprisingly, the variant with the highest contrast was a new hit (VHH72 S57NBDxK), substituting a serine residue buried within the binding interface that had not been explored at the knowledge-based discovery stage because a large FgAA substitution at that position was initially hypothesized to be detrimental to target binding.

Finally, we also tested if this approach could be used to rapidly prototype NS variants. For this, we generated dose-response curves by utilizing the cell-free NS synthesis reactions without purification and determined the so-called in situ EC_{50} value of ~1740 nM for ribosomally constructed CoV19-NS, which is ~2–4 fold higher than the EC_{50} value determined with CoV19-NS that had been produced semisynthetically (Supplementary Tables 4 and 5, Supplementary Fig. 16). These differences are within the range observed in similar high-throughput in situ antibody screening experiments[49]. In contrast, the RBD_{CoV19} affinity of ribosomally constructed VHH72 S57NBDxK was markedly higher (with an in situ EC_{50} value of ~215 nM, Supplementary Table 5). Taken together, the cell-free genetic construction and screening of NS without purification offers a rapid discovery and characterization route alternative to structure- and purification-dependent classical approaches.

Selection of improved NS from FgAA-containing nanobody libraries

Current directed evolution approaches enable the selection of high-affinity protein binders from protein libraries containing standard amino acids, yet their application to optical sensing often yields poor affinity and low contrast optical biosensors[56–60]. This is due to the fact that the key nsAA component of biosensors (i.e., FgAAs) is easier to deploy to the libraries of small peptides (because of manufacturing challenges) or via the two-step derivatization of protein libraries with fluorogens, which can bias the selection of unreacted protein variants due to long recognized incomplete probe conjugation[8] (Supplementary Fig. 3). We envisioned that one-step genetic construction of NS by site-specific incorporation of FgAAs (e.g., NBDxK for CoV19-NS) would

https://doi.org/10.1038/s41467-024-50956-z

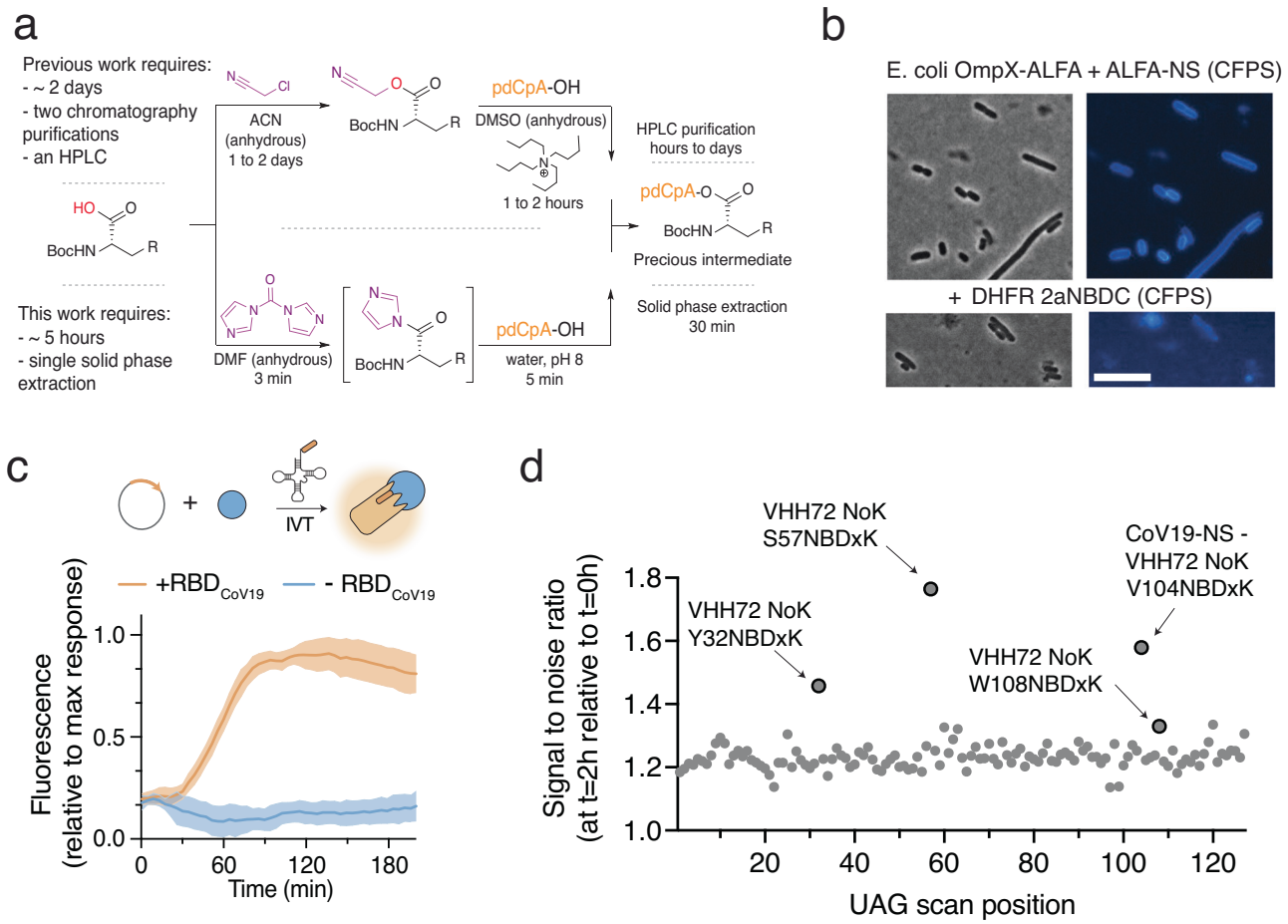


Fig. 4 | Cell-free site-specific translation of FgAAs can accelerate the discovery and functional characterization of NS. a The optimized chemical synthesis of the nsAA-pdCpA intermediate as compared to standard methods. **b** Cell-free translated ALFA-NS (via chemoenzymatically synthesized aNBDC-tRNA$_{CUA}$) enables direct, live-cell imaging of ALFA-tag labeled OmpX *in E. coli* despite the presence of excess aNBDC-tRNA$_{CUA}$. The same with DHFR 2aNBDC does not result in a similar uniform envelope labeling. **c** Ribosomally synthesized CoV19-NS can detect RBD$_{CoV19}$ in real-time in the reaction mixture. Panel created with help from BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **d** Target-dependent fluorescence increase of ribosomally synthesized NS variants. Cell-free translation reactions were set up with a linear DNA library encoding the TAG scan of VHH72, NBDxK-tRNA$_{CUA}$, and RBD$_{CoV19}$. Fluorescence measurements of these variants were carried out without purification, and fluorescence fold increases were calculated by comparing values at the 2-hour time point to those at time zero. This unbiased screen recapitulated CoV19-NS and led to the discovery of additional RBD$_{CoV19}$ nanosensors, such as VHH72 Y32NBDxK or VHH72 S57NBDxK. Lines and shaded areas represent the average and standard deviation of triplicate measurements. Scale bar: 4 µm. Source data are provided as a Source Data file.

accelerate the design of a generic evolution platform, especially when coupled to recombinant and cost-effective semisynthetic production of final hit NS.

To this end, we modified an mRNA/cDNA display protocol for the in vitro selection of nsAA-containing nanobodies, which allows up to 2 rounds of enrichment by magnetic selection per week (Fig. 5a)[61]. Next, we designed and constructed ~10^8 nanobody libraries around VHH72 V104UAG that would present NBDxK in position 104 and sample compensatory mutations at 8 other binding interface locations (Supplementary Fig. 17). First, we tested if improved CoV19-NS variants could be selected from the compensatory CoV19-NS library. After 3 rounds of bench-top selection cycles against RBD$_{CoV19}$, we randomly picked 36 enriched variants and screened them for RBD$_{CoV19}$ sensing via the rapid cell-free NS characterization approach (Fig. 1). From these, 9 variants showed signal-to-noise (SNR) ratios higher than 1.5 and 5 of those (Cov19-NS-2.1 to -2.5) showed significantly higher affinity to RBD$_{CoV19}$ than CoV19-NS (up to 13-fold lower in situ EC$_{50}$ values, Supplementary Tables 5 and 6) indicating that mRNA display can be used to select improved NS for their original target. CoV19-NS showed moderate affinity for RBD$_{CoV19}$ but not for the Omicron B.1.1.529 strain RBD (RBD$_{OB.1}$) (K_D > 9 µM, Supplementary Table 1), likely due to steric clashes between CoV19-NS and this highly mutagenized RBD variant (Supplementary Fig. 17b). Next, we tested if this approach could also be used to evolve NS against a new target like RBD$_{OB.1}$. After 7 rounds of bench-top selection cycles with an Omicron biased mutation library (Supplementary Fig. 17c) against RBD$_{OB.1}$, we observed convergence to distinct variants, and 3 of the top variants strongly responded to RBD$_{OB.1}$ (Fig. 5b, Supplementary Figs. 18 and 19). The scaled-up semi-synthetic production of these NS (final yields of ~2 mg NS from 100 mL *E. coli* culture, Fig. 1) allowed us to determine that these NS (Omicron-NS-1,-2, and -3) exhibited a dramatically improved performance in RBD$_{OB.1}$ sensing when compared to CoV19-NS (K_D values ~ 40 nM and ~10-fold increased brightness; Supplementary Fig. 20, Supplementary Table 1, Fig. 5c). Omicron-NS-1, -2, and -3 also selectively responded to RBD$_{OB.1}$ and RBD$_{OB.3}$ but not to other Omicron variants, such as RBD$_{OB.2}$ and RBD$_{OB.4.5}$ (Supplementary Fig. 21, Supplementary Table 4). Notably, Omicron-NS-1 has S52G, W53P, S57V, Y59G, and V104NBDxK mutations compared to VHH72. In silico alignment of RBD$_{OB.1}$ with RBD$_{CoV19}$/VHH72 suggested that a CoV19-NS Y59G single mutant could recognize RBD$_{OB.1}$ by relieving a predicted steric clash (Supplementary Fig. 17b). To test this hypothesis, we determined in situ EC$_{50}$ values of 10 mutant combinations between CoV19-NS and

https://doi.org/10.1038/s41467-024-50956-z

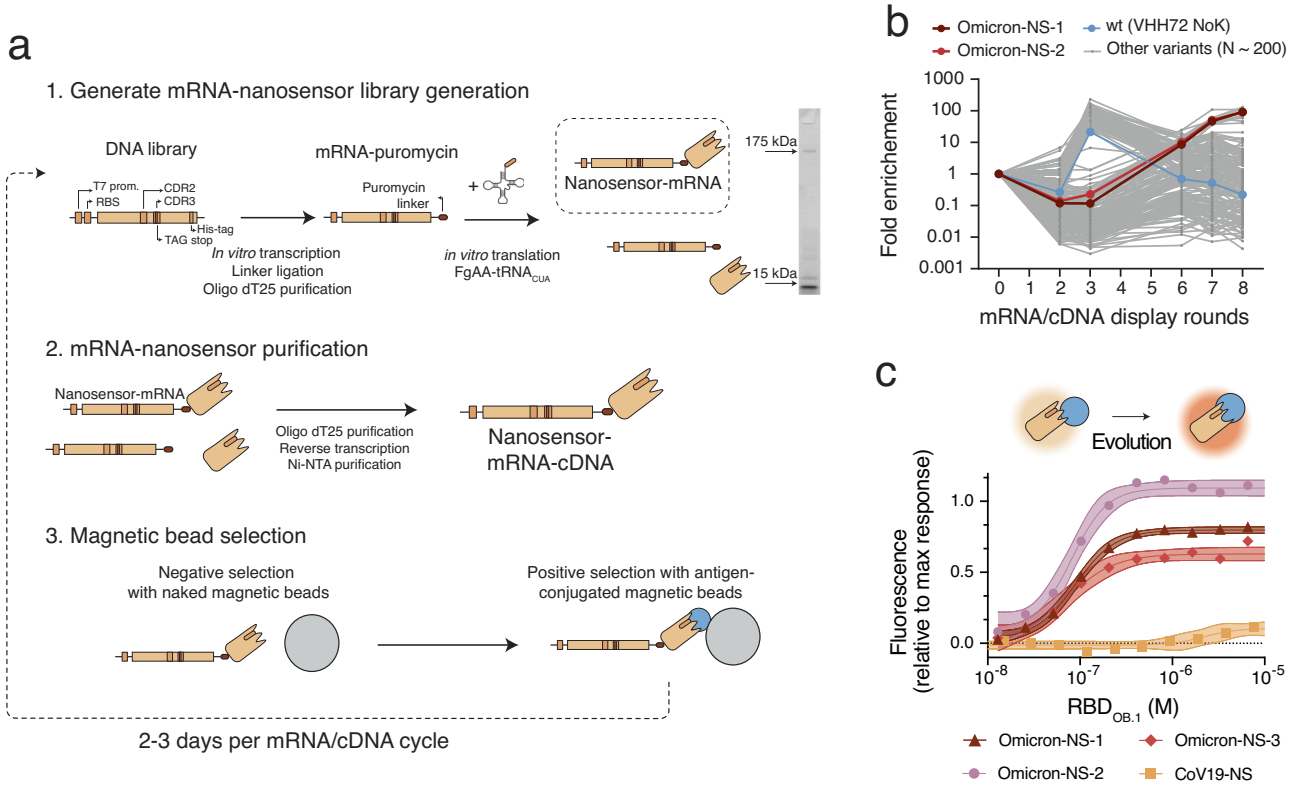


Fig. 5 | Selection of improved NS from FgAA-containing nanobody libraries. **a** The experimental strategy to evolve nanobodies containing nsAAs by mRNA/cDNA display. Panel created with help from BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **b** NS sequences were enriched when selecting for the $RBD_{OB.1}$ antigen. **c** Dose response curves with evolved NS (Omicron-NS-1, 2, and 3) showed high affinity and fluorescence fold improvements relative to CoV19-NS when exposed to $RBD_{OB.1}$. Each of the relative fluorescence units in the graph were normalized relative to the maximum response of Omicron-NS-2 for comparisons of relative NS brightness. Lines represent a 4PL fit of the dose-response curves. Shaded areas and dashed lines represent the 95% confidence intervals of the fits. Bars and error bars represent the average and standard deviation of independent triplicate experiments. Source data are provided as a Source Data file.

Omicron NS-1 against $RBD_{OB.1}$ (Supplementary Table 5). While CoV19-NS Y59G mutation alone showed an in situ EC_{50} value of ~330 nM against $RBD_{OB.1}$, the combination of all 4 mutations found in Omicron-NS-1 still yielded the highest affinity NS (in situ EC_{50} ~ 90 nM, Supplementary Table 5). These results indicate that rational NS design can be supported by unbiased genetic enrichment of improved variants from diverse libraries via directed evolution.

Finally, we tested whether this approach could be used to improve the variant specificity as Omicron-NS still showed a substantial affinity to RBD_{CoV19} (EC_{50} ~ 150 nM for Omicron-NS-1, Supplementary Table 4). For this, we challenged the compensatory CoV19-NS library in each round with a negative selection step against RBD_{CoV19} followed by a positive selection with $RBD_{OB.1}$. After 5 rounds, we found that 3 enriched variants (Omicron-NS-4, -5, and -6) among 8 tested showed a higher SNR in the presence of $RBD_{OB.1}$ than RBD_{CoV19} (Supplementary Table 6). One of these sensors, Omicron-NS-4, demonstrated a 20-fold increase in selectivity with in situ EC_{50} values ~100 nM and ~2200 nM against $RBD_{OB.1}$ and RBD_{CoV19}, respectively (Supplementary Table 5). These results suggest that our evolutionary platform could be, in principle, applied to the discovery and refinement of a wide range of NS against different targets.

In summary, our work describes a strategy to accelerate the design and optimization of optical NS. We have designed a high-throughput workflow to prepare large collections of nanobodies with site-selective reactive fluorophores, and to screen them for fluorescence activation upon recognition of their corresponding antigens. Using this platform, we identified CoV19-NS, a turn-on probe for the SARS-Cov-2 spike protein. CoV19-NS was based on the VHH72 nanobody scaffold and included an amine-reactive nitrobenzodiazole group in position 104 that resulted in strong fluorogenic readouts (up to 100-fold fluorescence increase) upon target binding. Importantly, we demonstrated that this platform could be adapted to NS for multiple and chemically diverse targets—from large nucleocapsid proteins and growth factor receptors (e.g., EGFR) to short peptide sequences (e.g., ALFA-tag peptide) and small molecules (e.g., cortisol)—and with fluorogenic behavior for wash-free microscopy experiments in live cells with minimal perturbation. Next, we optimized the platform towards ribosomal NS construction by the cell-free translation of retrosynthetic FgAAs via an optimized genetic code expansion chemistry and cost-effective NS manufacturing. We showed that this approach is compatible with rapid ribosomal discovery, characterization, and directed evolution of fluorogenic NS. Regarding the latter, we built libraries around Cov19-NS to select for compensatory mutations at the binding interfaces and discovered variant-specific NS variants. Altogether, this multi-staged workflow for the discovery and evolution of NS represents a timely advance for the development of low-cost, rapid, and selective optical sensors with applications in diagnostics, biomolecular sensing and molecular imaging.

Methods

Reagents

The following thiol-reactive probes were purchased from commercial sources: Cyanine3 maleimide (Cy3-Mal, Lumiprobe, Cat. No. 11080), 5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1-Sulfonic Acid (IAEDANS), 7-Diethylamino-3-[N-(2-maleimidoethyl)carbamoyl]coumarin

(MDCc, ThermoFisher, Cat. No. I14) N,N′-Dimethyl-N-(Iodoacetyl)-N′-(7-Nitrobenz-2-Oxa-1,3-Diazol-4-yl)Ethylenediamine (IANBD, ThermoFisher, Cat. No. D2004), 7-Diethylamino-3-[N-(3-maleimidopropyl) carbamoyl]coumarin (MDCpc, Sigma, Cat. No. 07153), N-(6-(diethylamino)-9-(2-((4-((2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)amino)-4-oxobutyl)(methyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-N-ethylethanaminium (Atto Rho3B-Mal, ATTO-Tec, Cat. No. AD Rho3B-41). The following amine-reactive probes were purchased from commercial sources: 6-(Tetramethylrhodamine-5-(and-6)-carboxamido) hexanoic acid, SE (TMR-x-NHS, AnaSpec, Cat. No. AS-81127), succinimidyl 6-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)hexanoate (NBD-x-NHS, AnaSpec, Cat. No. AS-81213), NBD dodecanoic acid N-succinimidyl ester (NBD-dodeca-NHS, chemodex, Cat. No. N0147), Malachite Green isothiocyanate (ThermoFisher, Cat. No. M689), Rhodamine B isothiocyanate RhoBITC, Sigma, Cat. No. 283924), Rhodamine Red™-X, Succinimidyl Ester, 5-isomer (RhoRed-x-NHS, ThermoFisher, Cat. No. R6160), 7-(Diethylamino)coumarin-3-carboxylic acid N-succinimidyl ester (DCc-NHS, Sigma, Cat. No. 36801). The following probes were synthesized as described in the references: 2-iodo-N-(2-((7-nitrobenzo[c][1,2,5]selenadiazol-4-yl)amino)ethyl)acetamide (IA(Se)NBD)[1] was synthesized by suspending (Se)NBD-ethylenediamine (1 eq) in DCM. Then chloroacethyl chloride (1.2 eq) and DIPEA (2 eq) were added. Reaction was allowed to stir at r.t. under N_2 for 1 hr, then solvent was removed under reduced pressure and the crude mixture was purified by column chromatography (DCM/MeOH 95:5) to give ClA(Se)NBD, that was redissolved in acetone and treated with NaI (2 eq) at r.t. overnight. Solvent was then removed under reduced pressure and the crude mixture was purified by column chromatography (DCM/MeOH 95:5) to give IA(Se)NBD as a purple solid (36% yield over 2 steps). N-(2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)-6-((7-nitrobenzo[c][1,2,5] selenadiazol-4-yl)amino)hexanamide ((Se)NBD-Mal)[1] was synthesized by dissolving (Se)NBD-Aminohexanoic acid (1 eq) in DMF. Then, COMU (1.5 eq) was added and the mixture was allowed to stir at r.t. for 10 minutes. N-Aminoethyl maleimide TFA salt (1 eq) in DMF was then added followed by DIPEA (1.5 eq) and reaction was allowed to stir at r.t. for 1 hr, Solvent was then removed under reduced pressure and the crude mixture was purified by column chromatography (DCM/MeOH 98:2) to give (Se)NBD-Mal as a red solid (56% yield). 1-(2-((bis(4-(dimethylamino)phenyl)methylene)amino)ethyl)-1H-pyrrole-2,5-dione (AO-Mal)[2] was synthesized by dissolving bis(4-(dimethylamino)phenyl) methanol (40 mg, 0.15 mmol) in DCM to which tetrafluoroboric acid (48 wt.% in water) was added dropwise until it became a persistent blue color. Then, the aminoethyl maleimide hydrochloride was added (12 mg, 0.07 mmol). After stirring for 5 minutes, DDQ was added (excess) and the reaction continued for several more minutes. The reaction solution was washed with basic water (NaOH, 0.1 M), then acidic water (HCl, 0.1 M), then washed with brine and finally dried over $MgSO_4$. The desired product was then purified by HPLC (method, mobile phase A: H_2O (100%), B: MeCN (100%); linear gradient 5% B to 95% B over 8 minutes at 1 mL/min) to give AO-Mal as a yellow solid (24% overall yield). 2,5-dioxopyrrolidin-1-yl 8-(((7-(dimethylamino)-3-oxo-3H-phenoxazin-1-yl)methyl)amino)-8-oxooctanoate (APM-o-NHS)[3] was synthesized by dissolving disuccinimidyl suberate (250 mg, MW 368.3, 0.68 mmol) in DMF (5 mL) and stirring at room temperature. Then, powdered L-(aminomethyl)-7-(dimethylamino)-3H-phenoxazin-3-one (APO, 50 mg, MW 269.3, 0.64 mmol) was mixed in over 30 minutes. The reaction was stirred overnight, and the crude reaction was concentrated under reduced pressure. The sample was then dissolved (DCM:MeOH 95:5) and purified by column chromatography (silica, DCM:MeOH 95:5) to give the title compound (60 mg, 68% yield) as a dark purple solid, 2,5-dioxopyrrolidin-1-y 6-((7-nitrobenzo[c][1,2,5] selenadiazol-4-yl)amino)hexanoate ((Se)-NBD-x-NHS)[1] was synthesized by dissolving (Se)NBD- Aminohexanoic acid (1 eq) in MeCN. Then, NHS (1.2 eq), DIC (1.2 eq) and DIPEA (1 eq) were added, and reaction was allowed to stir at r.t. for 72 hrs. Solvent was then removed under reduced pressure and the crude mixture was purified by column chromatography (DCM/MeOH 95:5) to give (Se)NBD-x-NHS as a red solid (72% yield). 4-fluoro-7-nitrobenzo[c][1,2,5]selenadiazole ((Se)NBD-F)[1], was syhnthesized by dissolving 4-Fluorobenzoselenadiazole in H_2SO_4 conc. to which HNO_3 conc. was added dropwise at 0 °C. After 15 minutes the mixture was added dropwise into cold water and the resulting precipitate was collected by filtration and dried overnight to give (Se)NBD-F as a yellow solid (70% yield). Cou was purchased from Sigma (Cat. No. 792551). 5-iodoacetemido-malachite green (IAMG) was custom synthesized by TOCRIS. Stock solutions were prepared in anhydrous DMSO, avoiding prolonged exposure to room temperature, and stored at −80 °C. pdCpA was purchased from Dharmacon. PylT tRNA(-CA)$_{CUA}$[4] and Mycoplasma capricolum Trp1(-CA)$_{CUA}$[5] were ordered from Agilent (Supplementary Table 8). The RBD antigens were purchased from Genscript: SARS-CoV-2 Spike RBD, CoV19 (Cat. No. Z03491), SARS-CoV-2 Spike protein RBD, E484Q, L452R (Cat. No. CP0007), SARS-CoV-2 Spike protein S1, del 69-70, N439K (Cat. No. Z03524), SARS-CoV-2 Spike protein RBD, K417N, L452R, T478K (Cat. No. Z03689), SARS-CoV-2 Spike protein RBD, E484K, K417N, N501Y (Cat. No. Z03537); and from Acro Biosystems: SARS-CoV-2 Spike RBD, B.1.1.529/Omicron (Cat. No. SPD-C522e), SARS-CoV-2 Spike RBD, BA.2.12.1/Omicron (Cat. No. SPD-C522q), SARS-CoV-2 Spike RBD, BA.3/Omicron (Cat. No. SPD-C522i), SARS-CoV-2 Spike RBD, BA.4 BA.5/Omicron, (Cat. No. SPD-C522r). SARS-CoV-2 Nucleocapsid protein was purchased from Genscript (Cat. No. Z03480). ALFA elution peptide was purchased from NanoTag Biotechnologies (Cat. No. N1520). Human epidermal growth factor receptor (EGFR) was purchased from Genscript (Cat. No. Z03194). Cortisol sulfate was purchased from Sigma (Cat. No. SMB00980). Graphs were plotted and analyzed in Prism 9 for Windows, GraphPad Software, www.graphpad.com.

Cloning of plasmids expressing protein binders

For PCR, site-directed mutagenesis, and isothermal assembly procedures, Q5® High-Fidelity 2× Master Mix, Q5® Site-Directed Mutagenesis Kit, and Gibson Assembly® Master Mix from New England Biolabs, (NEB Ipswich, MA) were used, respectively, and primers were designed following the manufacturer's instructions. Routinely, new plasmids were constructed by assembling linearized backbones of existing plasmids that are optimized for T7 RNA Polymerase dependent expression, i.e., pET-28a (+), pET-28_TEV, or pPURExpress (Supplementary Table 6), with eBlocks (IDT), and cloning into NEB® 5-alpha Competent *E. coli*. Nanosensor constructs typically contained N-terminal His tag followed by a Thrombin or TEV cleavage tag, the nanobody sequence and the mRNA display tag. Plasmids expressing EgA1, NbCor, and NbALFA variants were synthesized as Clonal Genes (Twist). All these constructs were verified by Sanger sequencing (Azenta/Genewiz) or complete plasmid sequencing (MGH DNA Core). Constructs for in vitro transcription/translation experiments were cloned into linearized backbone of pPURExpress control plasmid lacking an ORF and sequence verified as described above (Supplementary Table 6).

Expression and purification of protein-binder variants

VHH72 W108Cou were expressed essentially as previously described[6] with the following modifications: 1) a pET-28a plasmid was used expressing a VHH72 W108UAG ORF with an in-frame amber (UAG) stop codon. 2) amber suppression was preformed using tRNA$^{Tyr}_{CUA}$ acylated with Cou by CouRS both expressed from pDule-MjCouRS. 3) VHH72 W108Cou was purified as described below (Supplementary Table 6)[7].

Plasmids expressing protein binders were routinely transformed into SHuffle® T7 Competent *E. coli* cells (NEB). Up to 28 different overnight cultures (TB medium, Difco supplemented with 50 μg/mL kanamycin) were used to inoculate 100 mL of TB medium (1:100) supplemented with 50 μg/mL kanamycin at 30 °C and grown to OD600 = 0.4–0.5 with shaking. Cells were cooled to 16 °C before for overnight induction with 0.75 mM IPTG with shaking. Alternatively,

overnight cultures were used to inoculate 50–100 mL of MagicMedia™ *E. coli* Expression Medium (Thermo Fisher) supplemented with 50 μg/mL kanamycin and shaken at 30 °C for 18 h for overnight expression via auto-induction. Cells (50 mL × 2) were harvested by centrifugation (30 min, 5000 × *g*, RT), and pellets were either stored at −20 °C or up to 14 pellets were resuspended in 4 mL BugBuster® Master Mix (EMD Millipore) each and rocked at room temperature for 45 minutes. The lysate was centrifuged (15 min, 5000 × *g*, 4 °C) and the supernatant was added to 0.5 mL HisPur™ Cobalt Resin (Thermo Fisher) that is equilibrated and resuspended in 4 mL Equilibration Buffer (20 mM Tis-HCl pH 8.3, 0.5 M NaCl, 5 mM imidazole) in 15 mL conical tubes. After binding by rotation (35 min, 4 °C), the resin was pelleted (2 min, 700 g, 4 °C) and was washed twice with 1 mL wash Buffer (20 mM Tis-HCl pH 8.3, 0.5 M NaCl, 20 mM imidazole). The protein was eluted by 3 × 0.5 mL Elution Buffer (20 mM Tis-HCl pH 8.3, 0.5 M NaCl, 200 mM imidazole) and buffer was exchanged into 1× Phosphate Buffered Saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) using Zeba™ Spin Desalting Columns, 5 mL, 7 K MWCO (Thermo Fisher) following the manufacturer's instructions. Glycerol was added to final 20% for long-term storage at −80 °C. The protein yields and purity were assessed by running 2 μL samples in Novex™ WedgeWell™ Tris-Glycine Protein Gels (10–20%, 15-well, Invitrogen) following the manufacturer's instructions. This protocol allowed purification of up to 14 protein binder variants at once typically resulting in 1–2 mg protein from 50 mL cultures without the need for further concentration (~1 mg/mL).

Multiplexed modification and scaled-up production of nanosensor variants

Lysine variants of protein-binders were diluted to 50 μM in 1× PBS supplemented with 50 mM Sodium Borate (pH 8.5) and 45 μL aliquots were added into 96-well PCR plates that contained 5 μL stock solutions of amine-reactive probes (typically 2.5 mM in DMSO). The plate was sealed and the conjugation reaction was incubated for 2 h (dark, 25 °C). Cysteine variants of protein-binders were diluted to 50 μM in PBS supplemented with 500 μM Tris(2-carboxyethyl)phosphine (TCEP), sealed and incubated for 2 h (25 °C). 45 μL aliquots were added into 96-well PCR plates that contained 5 μL stock solutions of thiol-reactive probes (typically 2.5 mM in DMSO). The plate was sealed, and the conjugation reaction was incubated overnight (dark, 4 °C). Labelled proteins were separated from the unreactive probes by size-exclusion chromatography, e.g., using Thermo Scientific Zeba Spin Desalting 96-well filter plates (7K MWCO), following the manufacturer's instructions. The degree of labeling was assessed by measuring the ratio of fluorophore to protein from absorbance spectra of the purified conjugate and varied typically between ~0.5–1.5 for amine-reactive probes and ~0.1–0.8 for thiol-reactive probes.

This semisynthetic approach allows for scaled-up production of nanosensor hits with typical yields of ~2 mg of nanosensor from 100 mL *E. coli* culture, and with sufficient purity for downstream assays (Supplementary Fig. 1c, d and Supplementary Fig. 3). The method successfully produced NS containing diverse FgAAs like VHH72 G56MDCcC (MDCcC), CoV19-NS (NBDxK), H11-NS (DCcaK), sdAb-NS (RhoRedxK), LCB3-NS (aNBDC), ALFA-NS (aNBDC), Cortisol-NS (aNBDC), EGFR-NS (MDCpcC) as well as nanosensor hits that were discovered via cell-free nanosensor translation, e.g., Omicron-NS-1 (NBDxK). However, inherent limitations in this approach still exist resulting in heterogenous nanosensor preps due to minor *E. coli* protein impurities and moderately nonspecific or incomplete probe conjugation (Supplementary Figs. 1 and 3)[9]. Indeed, nanosensor performance benefited from optimization of labeling during the scaled-up production. For example, optimized CoV19-NS, which gives up to 100-fold fluorescence change, relied on modification with NBD-x-NHS ester at 150 μM dye and 10% DMSO followed by thrombin or TEV cleave of the N-terminus due to non-specific labeling of the N-terminal free amines (Supplementary Figs. 3a and 4). Ongoing efforts are focused on achieving homogenous NS production by employing higher resolution separation approaches such as anion exchange chromatography with a fast protein liquid chromatography (FPLC) system as shown for a phosphate biosensor[10].

Screening of nanosensor variants and fluorescence dose response curves

2.5–10 μL of NS preparations were transferred into low volume 384-well black flat clear bottom plates (Corning) and equal volumes of antigens (at saturating concentrations of typically ~1 mg/mL or >10 μM) or buffer only (1× PBS) were added. Using a Biotek Synergy H1 plate reader, fluorescence measurements were taken either directly in these plates (reading from the bottom) or transferred into Take3 Microvolume plates (Biotek) at probe-specific optimal excitation/emission wavelengths (Supplementary Table 7).

Fluorescence readings in the Take3 plates were more sensitive than readings in 384-well plates and thus the maximum fluorescence fold increase values (e.g., Supplementary Fig. 4) were calculated in Take3 plates. Dose response curves, e.g., Fig. 3a, or Supplementary Fig. 21, were determined in black 384-well plates by mixing 10 μL of sensor (at final ~2 μM NS) with equal volumes of serial dilutions of their corresponding antigens in 1× PBS unless indicated otherwise, e.g., human serum. The graphs were corrected for the background fluorescence by subtracting the nanosensor signal from the nanosensor plus the antigen signal. Signal values were normalized to peak fluorescence magnitude within an experiment and the graphs were plotted indicating the standard deviation between repeats in shade. Graphs were plotted and parameters such as EC_{50} were calculated in Prism 9 for Windows or Mac, GraphPad Software, www.graphpad.com.

Biolayer interferometry

Biolayer interferometry was performed on an Octet® RED384 instrument (ForteBio) at 30 °C with shaking at 1000 rpm. Signals were collected at the default frequency of 5.0 Hz. First, the Streptavidin (SA) biosensors (ForteBio) were preincubated in PBST [1× PBS containing 0.05% Tween 20 (Sigma-Aldrich)], the assay buffer used throughout the whole procedure, for 15 min. Biotinylated antigen for each nanobody was loaded on the tips at 50 μg/mL. Then, dilutions of the NS and control PBST were associated for 300 s; and dissociated for 600 s. BLI Steady-State analysis was performed in the ForteBio Data Analysis HT Software.

Immunofluorescence microscopy and image acquisition

For microscopy experiments leading to Fig. 3d and Supplementary Fig. 7a, SARS-CoV-2 envelope protein was expressed from a previously described plasmid[11] that is further modified to exclude 21 amino acids believed to be a cryptic ER-retention signal since the original wild-type envelope protein traffics to the ER-GIC. This resulted in the trafficking of Spike protein to the cell membrane, which was used to package SARS-CoV-2 pseudotyped lentivirus. HEK 293 T cells were then transfected with this plasmid vector in the following manner. Chamber slides were coated by applying a solution of 300 μL of 0.5 μg/mL poly-D-lysine in 1× PBS and incubated at 37 °C for 1.5 hours. Chamber slides were then washed once with sterile de-ionized water before cells were seeded into them. Approximately 2.5×10^4 HEK 293Ts were seeded in a volume of 250 μL complete media (DMEM, 10% FBS, 1% Pen/Strep). After overnight growth at 37 °C, 5% CO_2, each well was transfected using PEI. A total of 200 ng SARS-CoV2 plasmid (or the empty plasmid control) was combined with 600 ng PEI in a final volume of 50 μL DMEM (no FBS, no Pen/Strep). DNA-PEI solution was incubated for 10 minutes at room temperature before combining with 300 μL complete media for a total volume of 350 μL. This volume was then swapped with the existing cell culture supernatant in the chamber slides to transfect the cells. The transfected cells were then returned to the incubator for 48 hours, followed by fixation and permeabilization.

Supernatants were removed, and 200 μL 4% paraformaldehyde diluted in 1× PBS was added to each well and incubated at room temperature for 5 minutes. Each well was carefully washed twice with 1× PBS before 200 μL of 0.1% Triton X-100 in 1× PBS was added to each well and incubated for 10 minutes at room temperature. Each well was then washed 3 times with 200 μL 1× PBS, and cells were stored in the same buffer at 4 °C for up to 1 week before staining and imaging experiments. On the day of imaging, the media was exchanged into 1× PBS + CoV19-NS (2.5 μM) and directly imaged. Phase and fluorescence images were acquired using a Nikon Ti2 Eclipse inverted microscope equipped with a Plan Apo Lambda ×20 (0.75 NA, DIC N2) oil objective and Andor Zyla sCMOS camera. NIS-Element AR software was used for image acquisition. Image processing was performed in FIJI. Images were scaled, cropped, and rotated without interpolation. Linear adjustment was performed to optimize the contrast and brightness of the images. Figure construction was performed in Adobe Illustrator.

Data leading to SI Movies 1–2 and Supplementary Fig. 6a were acquired using the same microscope equipped with Plan Apo Lambda DM 4× (0.2 WD, 20) air objective and Countess™ Cell Counting Chambers Slides (Thermo). First 5 μL RBD_{CoV19} or BSA (2 mg/mL) was pipetted at the narrow end. Next, the real-time image acquisition was started and 15 μL CoV19-NS (2.5 μM) was pipetted to the opposite, wider end. NIS-Element AR software was used for image acquisition. For quantitative comparisons, the samples were imaged in the same session with the same image conditions across. Images were analyzed and presented as mentioned above. Data leading to SI Movie 3 and Supplementary Fig. 6b were acquired using an iPhone X (Apple) camera placed on top of the amber cover of a SmartBlue Mini Blue Light Transilluminator. The video was recorded using standard settings while pipetting RBD_{CoV19} into microtubes containing CoV19-NS kept in between the illuminator and the amber cover. Images were analyzed and presented as mentioned above.

For data leading to Fig. 3e and Supplementary Fig. 7b, *S. aureus* RN4220 cells were grown in tryptic soy broth (Becton-Dickinson Bacto-TSB, 30 g/L) at 37 °C with aeration, supplemented with 10 μg/mL erythromycin to maintain the plasmid pTB107 when necessary. pTB107 (Supplementary Table 6) was designed with SnapGene, generated by GenScript using site-directed mutagenesis with pLOW as the template, and PCR-verified. Cells were transformed with either an empty pLOW vector or pTB107 (pLOW_ALFA-spa-LPXTG), containing an in-frame ALFA tag between the native Staphylococcus protein A (SpA) signal sequence and coding sequence. Experiments were conducted from single colonies grown on tryptic soy agar (TSB with 1.5% Difco bacto-agar) plates. Cells were grown in TSB + 10 μg/mL erythromycin overnight into stationary phase, subcultured 1:2000 in fresh TSB, and induced with 50 ng/mL IPTG 1–2 hours prior to mid-log phase (OD_{600} ~ 0.5). 20 μL cells at exponential growth were labelled with 1 μL ALFA-NS (1 μM), then 2 μL cells were immobilized on 1× PBS pads with 2% wt/v agarose (Thermo-Fisher) and directly imaged with a Plan Apo Lambda DM 60× (1.4 NA, Ph3) oil objective. Images were analyzed and presented as mentioned above.

For data leading to Fig. 3f and Supplementary Fig. 7c, One Shot BL21 Star (DE3) Chemically Competent E. coli cells (Invitrogen) were transformed with the plasmid pET28a_TEV_OmpX-ALFA, containing a circularly permuted outer membrane protein OmpX with a C-terminal ALFA tag, and were grown in LB medium at 37 °C with aeration and supplemented with 50 μg/mL kanamycin to maintain the plasmid when necessary (Supplementary Table 6). Experiments were conducted from single colonies grown on LB agar plates. Cells were grown in LB + 50 μg/mL kanamycin overnight into stationary phase, subcultured 1:250 in fresh LB, and induced with 1 mM IPTG 1–2 hours prior to mid-log phase (OD_{600} ~ 0.5). In Supplementary Fig. 7c, 20 μL IPTG induced or uninduced cells at exponential growth were labelled with 1 μL ALFA-NS (1 μM), then 2 μL cells were immobilized on 1× PBS pads with 2% wt/v agarose (Thermo-Fisher) and directly imaged with a Plan Apo Lambda DM ×60 (1.4 NA, Ph3) oil objective. For real-time imaging of OmpX export (Fig. 3f), 3 μL of uninduced exponential cells were immobilized on LB pads containing 2% wt/v UltraPure™ Agarose (Thermo-Fisher) and ALFA-NS (1 μM) and the time-lapse image acquisition (1 fame/min) was started with a with a Plan Apo Lambda DM ×60 (1.4 NA, Ph3) oil objective. After 2-3 frames, concentrated IPTG (5 μL) was spotted on top of the pad to the final concentration of ~1 mM, and the cell growth and OmpX expression were followed via Phase and fluorescence images (GFP/Cy2 channel), respectively. Images were analyzed and presented as mentioned above.

Mass spectrometry methods

To generate the data presented in Supplementary Fig. 3, we conducted LC-MS/MS analysis at the Taplin Biological Mass Spectrometry Facility, Harvard Medical School, following established protocols[62]. Gel bands of interest were excised into ~1 mm^3 pieces. These pieces were subjected to a modified in-gel digestion protocol using either trypsin or chymotrypsin, which included the alkylation of cysteine residues with iodoacetamide. The gel pieces were then dehydrated using acetonitrile for 10 min. After drying in a speed-vac, the gel pieces were rehydrated with a buffer containing 50 mM ammonium bicarbonate and 12.5 ng/μL of proteomics grade trypsin (Promega). After 45 min incubation at 4 °C, the excess trypsin solution was removed, and the gel pieces were covered with fresh ammonium bicarbonate solution and incubated overnight at 37 °C to complete the digestion. The resulting peptides were extracted from the gel pieces by removing the ammonium bicarbonate solution and washing once with a 50% acetonitrile and 1% formic acid solution. The extracted peptides were then dried in a speed-vac and stored at 4 °C until analysis.

On the day of analysis, the dried peptides were reconstituted in a small volume (5–10 μL) of HPLC solvent A (2.5% acetonitrile, 0.1% formic acid). The peptides were then loaded onto a custom-made nano-scale reverse-phase HPLC capillary column packed with 2.6 μm C18 spherical silica beads. After equilibrating the column, a gradient of increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid) was applied to elute the peptides.

As the peptides eluted from the column, they were subjected to electrospray ionization and analyzed by an LTQ Orbitrap Velos Pro ion-trap mass spectrometer. The peptides were further fragmented, and the resulting tandem mass spectra were used to determine their amino acid sequences. The peptide sequences were then compared to a database of known protein sequences using the Sequest software. This analysis allowed us to identify the VHH72 protein in the original sample. To ensure the accuracy of our results, all databases included a reversed version of all the sequences, and the data was filtered to between a 1–2% peptide false discovery rate.

Supplementary Fig. 11 shows a high-resolution UPLC/MS analysis of NBDxK-incorporated peptides. For this first, equal volume of 2% formic acid was added to PURE reactions to precipitate large proteins. Samples were then centrifuged at >12,000 × *g* for 10 min. Samples in 1% formic acid (1 μL injection) were run on an Agilent 1290 UPLC using a Poroshell 120 SB-Aq column (2.7 μm, 2.1 × 50 mm; Agilent) with a linear gradient from 5% to 100% acetonitrile over 3.5 min at a flow rate of 0.6 mL/min with 0.1% formic acid in the mobile phase. Mass spectra were acquired using an Agilent 6530c QTOF with the following source and acquisition parameters: Gas temperature = 300 °C; gas flow = 8 l/min; nebulizer = 35 psig; capillary voltage = 3500 V; fragmentor 175 V; skimmer 65 V; °Ct 1 RF vpp = 750 V; acquisition rate = 3 spectra/s; acquisition time = 333.3 ms/spectrum; collision energy 0 V. Extracted ions for NT-formyl peptides (fM[NBDxK]PVFV and fMFPV[NBDxK]V; [M + H]+ m/z = 1024.4920) were monitored within a 10 ppm window. High Resolution mass spectrometry (HRMS) data for BDPaF-pdCpA and NBDxK-pdCpA leading to Supplementary Fig. 8 were collected and compared to the respective theoretical isotropic patterns of BDPaF-pdCpA: HRMS-ESI (m/z): Calc. for $C_{42}H_{50}BF_2N_{12}O_{15}P_2$ $[M+H]^+$

1073.3049, found 1073.3063 and NBDxK-pdCpA: HRMS-ESI (m/z): Calc. for $C_{37}H_{51}N_{14}O_{18}P_2$ $[M+H]^+$ 1041.2975, found 1041.2985.

tRNA ligation and quantification

The enzymatic esterification of tRNA(-CA) species to nsAA-pdCpAs (resulting nsAA-tRNA $_{CUA}$) was done as previously described[4]. Briefly, 500 μg of PylT tRNA(-CA)$_{CUA}$ or *Mycoplasma capricolum* Trp1 tRNA(-CA)$_{CUA}$ was dissolved in 625 μL 10 mM HEPES + 2.5 mM $MgCl_2$ and folded by heating to 95 °C for 3 min with a subsequent gradual cool-down to 25 °C over 20 min. The aminoacylation reaction to obtain the full length nsAA-tRNA $_{CUA}$ contained the final concentrations of 300 μg/mL folded tRNA(-CA), 0.3 mM nsAA-pdCpA (from 3 mM DMSO stock), 1× of T4 RNA Ligase buffer (from 10x, NEB), 0.125 mM ATP and 600 units/mL of T4 RNA Ligase 1 (NEB). This reaction was incubated at 4 °C for 2 h. The nsAA-tRNA $_{CUA}$ was extracted with acidic phenol chlorophorm (5:1, pH 4.5), ethanol precipitated, washed, air-dried and stored at −80 °C. To determine aminoacylation efficiencies (Supplementary Fig. 9) 1 μg of BDPaF-tRNA $_{CUA}$ was diluted in Novex™ TBE-Urea Sample Buffer (2x) and loaded onto a TBE urea gel (15 %). Electrophoresis was carried out at 120 V for 4 hours in 1× TBE. The gel was scanned for in gel fluorescence using a Fuji FLA-5100 fluorescent image analyzer and subsequently stained with SYBR™ Gold Nucleic Acid Gel Stain and visualized under UV. Charging yield was calculated by quantifying band intensity on the UV scanned gel using FIJI.

CFPS with FgAAs and in situ NS characterization

DNA templates for the in vitro transcription and translation reactions contained optimized sequences of a T7 promoter, a Shine-Dalgarno sequence, the open reading frame with an in-frame amber stop codon (* = TAG) that would be suppressed in the presence of charged full-length tRNA species, e.g., NBDxK-tRNA$_{CUA}$. These templates were typically prepared as circular DNA. Exceptions were mRNA/cDNA display experiments, which relied on linear DNA recovered from the previous selection round, the experiments leading to Supplementary Fig. 11, where peptide expression was conducted from linear DNA for Pep-MFPVFV, Pep-M*PVFV and Pep-PFPV*V sequences, amplified by PriPep-F and PriPep-R primers (Supplementary Table 8). Additionally, data leading to Fig. 4c, d, and Supplementary Figs. 15 and 16 were collected by setting up CFPS reactions directly with linear DNA templates acquired as gBlocks or 96-well eBlocks from IDT.

CFPS reactions were carried out using PURExpress® Δ RF1 or NEBExpress Cell-free E. coli Protein Synthesis System (NEB, Ipswich MA) following the manufacturer's instructions and supplying 20 ng/μL DNA templates with nsAA-tRNA$_{CUA}$ (at final ~8 μM), and 1.5 units/μL RNase Inhibitor Murine (NEB). In experiments where active NS were ribosomally produced, the CFPS reactions were also supplemented with PURExpress® Disulfide Bond Enhancer (NEB). Reactions (5–50 μL) were incubated in 0.2 mL thin wall PCR tubes (Thermo Fisher) at 30–37 °C for 60 min–300 min. CFPS reactions were analyzed by running 1 μL of the reaction in parallel with 10 μL Precision Plus Protein™ fluorescent protein ladder (Bio-Rad) in 1.0 mm Invitrogen™ Novex™ WedgeWell™ 10–20%, Tris-Glycine mini protein gels (Thermo Fischer) following the manufacturer's instructions and materials. The in-gel fluorescence was measured using a Biorad Gel Doc XR+ Imaging System and the gels were Coomassie stained with InstantBlue protein stain (Novus Biologicals), following the manufacturer's instructions. The images were quantified by ImageJ. In this work, CFPS reactions were carried out for platform optimization, rapid NS characterization, or NS discovery. Retrosynthetic FgAA designs allowed cost-effective semisynthetic production of these NS as described above.

For high-throughput screening of ribosomally produced nanosensor variants without purification (Supplementary Fig. 14a–c, and Supplementary Fig. 19), 2.5 μL of CFPS reactions were mixed with equal volume of 1× PBS (negative control), or saturating concentrations of RBD_{CoV19}, $RBD_{OB.1}$, ALFA peptide, or Cortisol in PCR tubes and their fluorescence was quantified by a Biorad Gel Doc XR+ Imaging System and analyzed by ImageJ. For microscopy images leading to Fig. 4b, 2 μL of CFPS reactions were mixed with 2 μL exponential *E. coli* cells expressing OmpX-ALFA from the pET28a_TEV_OmpX-ALFA, and the cells were directly imaged with a Plan Apo Lambda DM ×60 (1.4 NA, Ph3) oil objective on 1× PBS pads as described for Supplementary Fig. 7c above.

For real-time sensing of RBD_{CoV19} or ALFA peptide by nascently translated nanosensor variants (Fig. 4c, d, Supplementary Fig. 14d, and Supplementary Fig. 15a), 5 μL of CFPS reactions were mixed with 0.25 μL 10 mM Tris at pH = 7 (negative control) or concentrated RBD_{CoV19} or ALFA peptides (exchanged to 10 mM Tris at ~20 μM) in low volume 384-well black flat clear bottom plates (Corning). Relative fluorescence units for NBDxK were recorded at excitation/emission wavelengths of 485 nm/528 nm using a Biotek spectrophotometric plate reader at 30 °C over 3 h. The signal values were normalized to peak fluorescence magnitude within an experiment and the graph was plotted indicating the standard deviation between repeats in shade. Graphs were plotted and analyzed in Prism 9 for Windows, GraphPad Software.

High-throughput dose response curves of NS leading to data in Supplementary Fig. 16 and in situ EC_{50} values in Supplementary Table 5 were determined by mixing 2 μL of CFPS reactions with equal volumes of serial dilutions of RBD_{CoV19} or $RBD_{OB.1}$ in 1× PBS in low volume black flat clear bottom 384-well plates (Corning). Liquid transfers were performed using an Echo 525 Acoustic Liquid Handler by Beckman, and fluorescence measurements were taken using a Biotek Synergy H1 plate reader (reading from the bottom of the plate). The graphs were plotted using data from at least 2 repeats and corrected for the background fluorescence by normalizing nanosensor plus the antigen signals against the control nanosensor without antigen signal. Signal values were normalized to maximum fluorescence response within each set and the graphs were plotted indicating the standard deviation between repeats in shade. Graphs were plotted and parameters such *as* in situ EC_{50} were calculated in Prism 9 for Windows or Mac, GraphPad Software, www.graphpad.com.

mRNA/cDNA display

ORF libraries for mRNA/cDNA display were constructed stepwise. The Omicron biased mutation library ORF, LibOmic, containing the fixed in-frame TAG stop codon at position 104 and randomized CDR2 and CDR3 (Supplementary Table 6 also shown in Supplementary Fig. 17a), was built by 4 cycles of overlap extension PCR with Pri1-Pri4 (acquired as PAGE purified Ultramers, IDT) that also allowed the representation of tryptophans in the library. For this, PCR reactions of Pri1-Pri2, Pri1-Pri4, Pri2-Pri3, and Pri3-Pri4 were pooled in 1000:33:33:1 ratio followed by amplification with Pri5&Pri6 (5–10 cycles), and PCR purification (Supplementary Table 8). This library was further diversified by one-piece isothermal assembly using the plasmid library backbone that is linearized by Pri9 & Pri10 and the same cloning strategy. For the general CoV19-NS compensatory library, CompLibCoV (Supplementary Table 6, also shown in Supplementary Fig. 17b), the randomized CDR2 and CDR3 were built by 9 cycles of overlap extension PCR by mixing equal moles of Pri18-Pri21 (acquired as PAGE purified Ultramers, IDT) followed by amplification with Pri5&Pri6 (5–10 cycles) and PCR purification (Supplementary Table 8). This library was further diversified by one-piece isothermal assembly using the plasmid library backbone that is linearized by Pri9 & Pri10 and the same cloning strategy. These library inserts (~500 ng) were assembled into the pPURExpress VHH72 V104UAG plasmid backbone (~800 ng, linearized by the primers Pri7 and Pri8) in a 150 μL Gibson assembly reaction (NEB 50 °C, 1 h). The product was then cleaned and concentrated by ethanol precipitation and the entire product was electro-transformed into *ElectroMAX™ DH10B Cells* (Thermo Fisher) cells. After cells were recovered in SOC for 1 h, overnight cultures were set up for plasmid minipreps by adding 4 mL 2xYT supplemented with carbenicillin (to final 100 μg/mL) at

37 °C with aeration. In parallel, dilutions were plated to estimate the library size. 100 colonies were randomly selected and sequenced (Azenta/Genewiz) to estimate the library quality.

Our mRNA/cDNA display approach involves translating nsAA-containing nanobody libraries and covalently linking them to their cDNA via a puromycin linker. This is achieved by modifying a previously optimized protocol[15] that includes a key step that allows site-specific incorporation of nsAAs at the binding interface as described above (Fig. 5a). The specific deviations from the protocol were the following: LibOmic or CompLibCoV was initially amplified with Pri11-Pri12 that added a 3′ T7 promoter followed by an optimized RBS and a 5′ His-tag followed by a flexible mRNA/cDNA display tag. After this step, the linear DNA library for transcription was amplified by Pri11-Pri13. Alternatively, Pri14-Pri15 was used to amplify the DNA library for isothermal assembly cloning into pPURExpress backbone, which is linearized by using Pri16-17. The CFPS reaction using PURExpress® Δ RF1 contained 8 μM NBDxK-tRNA$_{CUA}$ and was performed at 30 °C for 90 min. After reverse transcription, full-length NS linked to their mRNA and cDNA were enriched via His-Pull-Down using 10 μL Pierce™ Ni-NTA Magnetic Agarose Beads following the manufacturer's protocol. The elution (~150 μL) was subjected to negative and positive selections. Each selection round included a generic negative selection step with 20 μL Magnetic beads™ Streptavidin (from 1 mg/ mL, Acro Biosystems, Cat. No. SMB-B01 for 30 min at room temperature. Omicron RBD-selectivity was achieved by a second negative selection using increasing amounts (e.g., 2.5, 5, 10, 20 μL) of SARS-CoV-2 Spike RBD-coupled Magnetic Beads per round. Positive selection involved incubating the supernatant with decreasing amounts (e.g., 40, 20, 10, 5 μL) of SARS-CoV-2 (Omicron) Spike RBD-coupled Magnetic Beads (Acro Biosystems, Cat. No. MBS-K043) or original SARS-CoV-2 Spike RBD-coupled Magnetic Beads for shorter durations (e.g., 60, 45, 30 min) per round. The supernatant was saved for downstream analysis and the beads were thoroughly washed with Wash/Interact buffer (5 times with a total of 200 μL of buffer containing 20 mM Tris-HCl, 150 mM NaCl, 0.05% Tween 20, and 10 μg/mL BSA at pH 8.0) followed by washes with Tris-Wash buffer (2 times with a total of 200 μL of buffer containing 20 mM Tris-HCl and 0.05% Tween 20 at pH 8.0) as described[15]. The washes were also saved for downstream analysis, which included running protein gels for monitoring in-gel nanosensor fluorescence, qPCR (by signal amplification using Pri11-Pri13), and Illumina next-generation sequencing. The final elution was done using streptavidin elution buffer (G-BIOSCIENCES, Cat. No. 786–549) followed by neutralization with an equal volume of 1 M Tris, pH 8, and ethanol precipitation. The pellet was reconstituted in water or either (i) used as the template for the next cycle by reamplifying with Pri11-Pri13, (ii) cloned into pPURExpress by amplifying with Pri14-Pri15 as mentioned above, or (iii) amplified by Pri5-Pri6 for next generation sequencing (Amplicon-EZ, Azenta/Genewiz).

Illumina next-generation sequencing (NGS) data analysis and read counting

NGS-based amplicon sequencing was performed using the Amplicon-EZ service of Azenta/Genewiz and DNA amplicons from each selection cycle. cDNA in samples from washes and elutions were amplified by Pri5-Pri6) and were prepared following Amplicon-EZ sample submission guidelines. Raw reads (236 bp pair-ended) were merged using BBMerge[16], and filtered for Phred quality scores at or above 20. Resulting reads were forwarded and trimmed using a custom Python script, which identified the first 18 bp of the constant region prior to CDR2. Reads were trimmed such that the forward read started at CDR2 and extended through CDR3 to identify nanosensor variant combinations within both regions, and counts of identical sequences were determined. The read frequency was calculated as the fraction of each unique sequence divided by the total number of trimmed sequences detected within a sample. Almost 200 unique reads were identified in all samples from the mRNA/cDNA display evolution rounds, and fold enrichment was calculated by dividing the read frequencies of subsequent rounds by the read frequency identified in the original library. Scripts are available upon request.

Reporting summary

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

Data availability

NGS raw data generated in this study have been deposited in the NIH Short Read Archive (SRA) database under accession code PRJNA1129853. Protein mass spectrometry raw data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the PRIDE database ID (PXD053593) and https://doi.org/10.6019/PXD053593. Protein structure data used in this study is available in the Worldwide Protein Data Bank under the following accession codes: 6WAQ, 6ZBP, 7JZM, 6I2G, 4KRO and 6ITP. Source data are provided as a Source Data file. Source data are provided with this paper.

Code availability

The custom python script used for NGS data analysis is available through Zenodo [https://doi.org/10.5281/zenodo.12741352].

References

1. Adamson, H. & Jeuken, L. J. C. Engineering protein switches for rapid diagnostic tests. *ACS Sens.* **5**, 3001–3012 (2020).
2. Dong, J. & Ueda, H. Recent advances in quenchbody, a fluorescent immunosensor. *Sensors* **21**, 1223 (2021).
3. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642–646 (2011).
4. VarnBuhler, B. S., Moon, J., Dey, S. K., Wu, J. & Jaffrey, S. R. Detection of SARS-CoV-2 RNA using a DNA aptamer mimic of green fluorescent protein. *ACS Chem. Biol.* **17**, 840–853 (2022).
5. Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).
6. Gulyani, A. et al. A biosensor generated via high-throughput screening quantifies cell edge Src dynamics. *Nat. Chem. Biol.* **7**, 437–444 (2011).
7. de Picciotto, S. et al. Design principles for SuCESsFul biosensors: specific fluorophore/analyte binding and minimization of fluorophore/scaffold interactions. *J. Mol. Biol.* **428**, 4228–4241 (2016).
8. Brient-Litzler, E., Plückthun, A. & Bedouelle, H. Knowledge-based design of reagentless fluorescent biosensors from a designed ankyrin repeat protein. *Protein Eng. Des. Sel.* **23**, 229–241 (2010).
9. Mills, J. H., Lee, H. S., Liu, C. C., Wang, J. & Schultz, P. G. A genetically encoded direct sensor of antibody–antigen interactions. *ChemBioChem* **10**, 2162–2164 (2009).
10. De Lorimier, R. M. et al. Construction of a fluorescent biosensor family. *Protein Sci.* **11**, 2655–2675 (2002).
11. Mellanby, R. J. et al. Tricarbocyanine N-triazoles: the scaffold-of-choice for long-term near-infrared imaging of immune cells in vivo. *Chem. Sci.* **9**, 7261–7270 (2018).
12. Kaplaneris, N. et al. Chemodivergent manganese-catalyzed C–H activation: modular synthesis of fluorogenic probes. *Nat. Commun.* **12**, 3389 (2021).
13. Barth, N. D. et al. A fluorogenic cyclic peptide for imaging and quantification of drug-induced apoptosis. *Nat. Commun.* **11**, 4027 (2020).
14. Fernandez, A., Thompson, E. J., Pollard, J. W., Kitamura, T. & Vendrell, M. A fluorescent activatable AND-gate chemokine CCL2 enables in vivo detection of metastasis-associated macrophages. *Angew. Chem. Int. Ed.* **58**, 16894–16898 (2019).

15. Plamont, M.-A. et al. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. *Proc. Natl. Acad. Sci.* **113**, 497–502 (2016).
16. Strack, R. L., Song, W. & Jaffrey, S. R. Using spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria. *Nat. Protoc.* **9**, 146–155 (2014).
17. Szent-Gyorgyi, C. et al. Malachite green mediates homodimerization of antibody VL domains to form a fluorescent ternary complex with singular symmetric interfaces. *J. Mol. Biol.* **425**, 4595–4613 (2013).
18. Babendure, J. R., Adams, S. R. & Tsien, R. Y. Aptamers switch on fluorescence of triphenylmethane dyes. *J. Am. Chem. Soc.* **125**, 14716–14717 (2003).
19. McMahon, C. et al. Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).
20. Hoogenboom, H. R. Designing and optimizing library selection strategies for generating high-affinity antibodies. *Trends Biotechnol.* **15**, 62–70 (1997).
21. Hirshberg, M. et al. Crystal structure of phosphate binding protein labeled with a coumarin fluorophore, a probe for inorganic phosphate. *Biochemistry* **37**, 10381–10385 (1998).
22. Kuru, E. et al. Release factor inhibiting antimicrobial peptides improve nonstandard amino acid incorporation in wild-type bacterial cells. *ACS Chem. Biol.* **15**, 1852–1861 (2020).
23. Cheng, Z., Kuru, E., Sachdeva, A. & Vendrell, M. Fluorescent amino acids as versatile building blocks for chemical biology. *Nat. Rev. Chem.* **4**, 275–290 (2020).
24. Reese, A. E. et al. Inserting "OFF-to-ON" BODIPY tags into cytokines: a fluorogenic interleukin IL-33 for real-time imaging of immune cells. *ACS Cent. Sci.* **10**, 143–154 (2023).
25. Hsu, Y.-P. et al. Fluorogenic d-amino acids enable real-time monitoring of peptidoglycan biosynthesis and high-throughput transpeptidation assays. *Nat. Chem.* **11**, 335–341 (2019).
26. Kuru, E., Tekkam, S., Hall, E., Brun, Y. V. & Van Nieuwenhze, M. S. Synthesis of fluorescent D-amino acids and their use for probing peptidoglycan synthesis and bacterial growth in situ. *Nat. Protoc.* **10**, 33–52 (2015).
27. Wrapp, D. et al. Structural basis for potent neutralization of betacoronaviruses by single-domain camelid antibodies. *Cell* **181**, 1004–1015.e1015 (2020).
28. Laroche, A. et al. Deep mutational engineering of broadly-neutralizing nanobodies accommodating SARS-CoV-1 and 2 antigenic drift. *mAbs* **14**, 2076775 (2022).
29. Dangerfield, T. L. & Johnson, K. A. Optimized incorporation of an unnatural fluorescent amino acid affords measurement of conformational dynamics governing high-fidelity DNA replication. *J. Biol. Chem.* **295**, 17265–17280 (2020).
30. Wang, J., Xie, J. & Schultz, P. G. A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* **128**, 8738–8739 (2006).
31. Benson, S. et al. Photoactivatable metabolic warheads enable precise and safe ablation of target cells in vivo. *Nat. Commun.* **12**, 2369 (2021).
32. Erez, Y., Amdursky, N., Gepshtein, R. & Huppert, D. Temperature and viscosity dependence of the nonradiative decay rates of auramine-o and thioflavin-T in glass-forming solvents. *J. Phys. Chem. A* **116**, 12056–12064 (2012).
33. Cohen, B. E. et al. A fluorescent probe designed for studying protein conformational change. *Proc. Natl Acad. Sci. USA* **102**, 965–970 (2005).
34. Huo, J. et al. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat. Struct. Mol. Biol.* **27**, 846–854 (2020).
35. Cao, L. et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
36. Ye, Q., Lu, S. & Corbett, K. D. Structural basis for SARS-CoV-2 nucleocapsid protein recognition by single-domain antibodies. **12**, 719037 (2021).
37. Götzke, H. et al. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nat. Commun.* **10**, 4403 (2019).
38. Schmitz, K. R. et al. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. *Structure* **21**, 1214–1224 (2013).
39. Ding, L. et al. Structural insights into the mechanism of single domain VHH antibody binding to cortisol. *FEBS Lett.* **593**, 1248–1256 (2019).
40. Tsai, Y. C., Jin, Z. & Johnson, K. A. Site-specific labeling of T7 DNA polymerase with a conformationally sensitive fluorophore and its use in detecting single-nucleotide polymorphisms. *Anal. Biochem.* **384**, 136–144 (2009).
41. Kim, Y. E., Chen, J., Chan, J. R. & Langen, R. Engineering a polarity-sensitive biosensor for time-lapse imaging of apoptotic processes and degeneration. *Nat. Methods* **7**, 67–73 (2010).
42. de Lorimier, R. M. et al. Construction of a fluorescent biosensor family. *Protein Sci.* **11**, 2655–2675 (2002).
43. Wang, C. et al. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* **11**, 2251 (2020).
44. Meissner, D., Vollstedt, A., van Dijl, J. M. & Freudl, R. Comparative analysis of twin-arginine (Tat)-dependent protein secretion of a heterologous model protein (GFP) in three different Gram-positive bacteria. *Appl. Microbiol. Biotechnol.* **76**, 633–642 (2007).
45. Wang, W. et al. Real-time imaging of cell-surface proteins with antibody-based fluorogenic probes. *Chem. Sci.* **12**, 13477–13482 (2021).
46. He, K., Zeng, S. & Qian, L. Recent progress in the molecular imaging of therapeutic monoclonal antibodies. *J. Pharm. Anal.* **10**, 397–413 (2020).
47. Hori, Y. & Kikuchi, K. Protein labeling with fluorogenic probes for no-wash live-cell imaging of proteins. *Curr. Opin. Chem. Biol.* **17**, 644–650 (2013).
48. Monteiro, R. et al. Specific proteomic identification of collagen-binding proteins in Escherichia coli O157:H7: characterisation of ompa as a potent vaccine antigen. *Cells* **12**, 1634 (2023).
49. Hunt, A. C. et al. A rapid cell-free expression and screening platform for antibody discovery. *Nat. Commun.* **14**, 3897 (2023).
50. Hecht, S. M., Alford, B. L., Kuroda, Y. & Kitano, S. "Chemical aminoacylation" of tRNA's. *J. Biol. Chem.* **253**, 4517–4520 (1978).
51. Robertson, S. A., Ellman, J. A. & Schultz, P. G. A general and efficient route for chemical aminoacylation of transfer RNAs. *J. Am. Chem. Soc.* **113**, 2722–2729 (1991).
52. Leisle, L. et al. Cellular encoding of Cy dyes for single-molecule imaging. *eLife* **5**, e19088 (2016).
53. Kajihara, D. et al. FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids. *Nat. Methods* **3**, 923–929 (2006).
54. Lee, J. et al. Ribosome-mediated incorporation of fluorescent amino acids into peptides in vitro. *Chem. Commun.* **57**, 2661–2664 (2021).
55. Chen, K.-T. et al. Rapid preparation of mycobacterium N-glycolyl lipid I and lipid II derivatives: a biocatalytic approach. *Chemistry* **19**, 834–838 (2013).
56. K.C, T. B. et al. Wash-free and selective imaging of epithelial cell adhesion molecule (EpCAM) expressing cells with fluorogenic peptide ligands. *Biochem. Biophys. Res. Commun.* **500**, 283–287 (2018).
57. Wang, W. et al. A fluorogenic peptide probe developed by in vitro selection using tRNA carrying a fluorogenic amino acid. *Chem. Commun. (Camb.)* **50**, 2962–2964 (2014).

58. Jespers, L., Bonnert, T. P. & Winter, G. Selection of optical biosensors from chemisynthetic antibody libraries. *Protein. Eng. Des. Sel.* **17**, 709–713 (2004).
59. Taki, M., Inoue, H., Mochizuki, K., Yang, J. & Ito, Y. Selection of color-changing and intensity-increasing fluorogenic probe as protein-specific indicator obtained via the 10BASEd-T. *Anal. Chem.* **88**, 1096–1099 (2016).
60. Hashimoto, M., Miki, T., Chang, I. V., Tsutsumi, H. & Mihara, H. Selection of fluorescent biosensors against galectin-3 from an NBD-modified phage library displaying designed α-helical peptides. *Bioorg. Med. Chem. Lett.* **37**, 127835 (2021).
61. Doshi, R. et al. In vitro nanobody discovery for integral membrane protein targets. *Sci. Rep.* **4**, 6760 (2014).
62. Mahlokozera, T. et al. Competitive binding of E3 ligases TRIM26 and WWP2 controls SOX2 in glioblastoma. *Nat. Commun.* **12**, 6321 (2021).

Acknowledgements

This work was supported by the Life Sciences Research Foundation Fellowship awarded to E.K. and Star Friedman Challenge and the Wyss Institute Validation Project awarded to G.M.C. and E.K. Work in the Church laboratory was supported by the US Department of Energy Grant DE-FG02-02ER63445. M.V. acknowledges funds from an ERC Consolidator Grant (DYNAFLUORS, 771443). We thank A. Kruse, M. Schubert, A. Chiappino-Pepe, and T.M. Wannier, for helpful discussions about this project; M. Lipstein and G. Kuziel for their advice on mRNA display; and G. Chao for his help with sensor dynamics experiments. We also thank the technical support from the Center for Macromolecular Interactions at Harvard Medical School, as well as the Harvard Medical School Core facilities for Taplin Mass Spectrometry and the Imaging/Microscopy (MICRON) Center.

Author contributions

E.K., M.V., and G.M.C. designed the study. E.K., H.P., and I.H. designed the nanosensor candidates with initial discussions with W.B., L. R., and J.J.C. E.K., and A.F. carried out nanosensor discovery. J.R., F.M., S.R., and J.D.G. conducted the chemical synthesis of new reactive fluorogens and FgAAs. J.R. optimized the tRNA charging chemistry with feedback from E.K. benchmarking against the conventional tRNA charging chemistry performed by J.D.G. and C.A.A; J.R. and S.R. performed scaled-up synthesis of FgAA-tRNA species; E.K., A.F., and S.G.B. conducted fluorescence screening experiments including dose-response curves. E.K. implemented microscopy experiments with help from I.H., T.M.B., T.G.B., and D.Z.R.; E.K. conducted cell-free assays with guidance from J.M. and A.E.R.; E.K. performed mRNA display. A.F. performed NGS data analysis. E.K. and H.P. conducted data analysis and made the figures with guidance from J.J.C. E.K. wrote the manuscript with feedback from all authors. H.P., A.F., and S.R. contributed equally to this work.

Competing interests

E.K., H.P., J.J.C., and G.M.C. are cofounders of ExtRNA. J.R. and G.M.C. are cofounders of enplusone. G.M.C. is a cofounder of 64-x, EnEvolv, and GRO Biosciences. For a complete list of G.M.C.'s financial interests, please visit arep.med.harvard.edu/gmc/tech.html. J.J.C. is a cofounder and director at Sherlock Biosciences. The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-50956-z.

Correspondence and requests for materials should be addressed to Erkin Kuru, Marc Vendrell or George M. Church.

Peer review information *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.