nature chemical biology

PERSPECTIVE

https://doi.org/10.1038/s41589-020-00718-x



# Structure- and mechanism-guided design of single fluorescent protein-based biosensors

Yusuke Nasu[1,3], Yi Shen[2,3], Luke Kramer[1] and Robert E. Campbell[1,2] ✉

**Intensiometric genetically encoded biosensors, based on allosteric modulation of the fluorescence of a single fluorescent protein, are powerful tools for enabling imaging of neural activities and other cellular biochemical events. The archetypical example of such biosensors is the GCaMP series of $Ca^{2+}$ biosensors, which have been steadily improved over the past two decades and are now indispensable tools for neuroscience. However, no other biosensors have reached levels of performance, or had revolutionary impacts within specific disciplines, comparable to that of the $Ca^{2+}$ biosensors. Of the many reasons why this has been the case, a critical one has been a general black-box view of biosensor structure and mechanism. With this Perspective, we aim to summarize what is known about biosensor structure and mechanisms and, based on this foundation, provide guidelines to accelerate the development of a broader range of biosensors with performance comparable to that of the GCaMP series.**

Many biological experiments that would have once been performed with purified proteins or tissue lysates can now be done, in principle, in the context of live model organisms. The fact that such in vivo experiments are possible is a testament to numerous advances in molecular imaging over the past several decades. Among molecular imaging modalities, fluorescence imaging stands out due to its sensitivity and versatility and the widespread accessibility of microscopy instrumentation.

A driving force behind advances in fluorescence molecular imaging has been the development of genetically encoded biosensors based on fluorescent proteins (FPs), including the color variants and homologs of *Aequorea* green FP (GFP)[1]. While alternative fluorophores such as synthetic dyes[2] and fluorescent nanoparticles[3] have their associated advantages, the genetically encoded nature of FPs makes them readily shareable between laboratories and the technology of choice for most biological imaging experiments. Indeed, transgenic animals that express FP-based biosensors (in a tissue, cell type or organelle of interest) enable minimally invasive and effectively reagent-free in vivo assays. For an exhaustive listing of published genetically encoded biosensor variants, see ref. [1] and biosensorDB.ucsd.edu.

In this Perspective, we aim to introduce a wider community to the basic principles, and best practices, of single FP-based biosensor design (Box 1). Due to the universal role of $Ca^{2+}$ in cell signaling, there may never be another series of biosensors with the revolutionary impact of the GCaMP series. However, countless other fields of biological research may yet be revolutionized by the advent of high-performance biosensors that enable practical in vivo molecular imaging and redefine the scope of biological questions to be asked, and answered.

## Types of genetically encoded biosensors

The two main classes of genetically encoded biosensors are those based on Förster resonance energy transfer (FRET) between two hues of FP[4] and those based on the allosteric modulation of fluorescence intensity from a single FP[5–7]. The advantages of FRET-based biosensors include greater ease of development and an inherently ratiometric fluorescent response that is more amenable to calibration for quantitative imaging. A disadvantage is the requirement for two emission color bands (limiting the possibilities for multicolor, multiparameter imaging)[8]. A second disadvantage stems from the fact that different wavelengths of light are absorbed and scattered differently within tissue, so the observed emission ratio may change as a function of tissue depth[9].

Single FP-based biosensors (Fig. 1a) are substantially more difficult to engineer than FRET-based biosensors and are typically intensiometric, making them less appropriate for quantitative imaging applications. In contrast, they are more user-friendly than FRET-based biosensors due to their typically larger, single-wavelength, intensiometric responses. As they require only a single color band for excitation and emission, single FP-based biosensors are well suited for multiparameter imaging[10] and avoid the problem of differential absorbance and scatter of different wavelengths as a function of tissue depth. Notably, some single FP-based biosensors are inherently ratiometric (see examples in refs. [10–13]) and others can be converted to ratiometric biosensors by fusion to a second color of FP[14,15]. As the tremendous success of the GCaMP series of single FP-based biosensors[16–21] illustrates, biologists (especially within neuroscience) have collectively decided that the advantages of intensiometric single FP-based biosensors outweigh the need for quantitation. This is probably in large part due to the all-or-none binary nature of many aspects of neuronal signaling, which diminish the need for quantitative measurements. This trend is evident in the flurry of recent papers describing intensiometric single FP-based biosensors of neurotransmitters and neuromodulators[22–32]. As quantitation is critical to many other areas of cell biology, the development of ratiometric biosensors should not be abandoned as the field moves forward.

The main challenge associated with single FP-based biosensors has been the sophisticated protein design and engineering required to develop them. Through the first decade of this field (1999–2009), only a handful of researchers were able to meet this challenge and develop functional biosensors[5,16,33–37]. However, over the past decade, an increasing number of successful examples (Supplementary Tables 1 and 2 and refs. [1,6]) have helped protein engineers to converge on a standardized, yet versatile, design strategy. In this design strategy, a

[1]Department of Chemistry, Graduate School of Science, The University of Tokyo, Tokyo, Japan. [2]Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada. [3]These authors contributed equally: Yusuke Nasu, Yi Shen. ✉e-mail: campbell@chem.s.u-tokyo.ac.jp

**Box 1 | Biosensor development: take-home messages**

- *Aequorea* GFP is highly amenable to conversion into a biosensor, and all new efforts should start with it. Other color variants should only be attempted after achieving success with GFP.
- To create a GFP-based prototype biosensor, a cpGFP domain (topology in Fig. 1b) or GFP domain (topology in Fig. 1c,d) should be linked to a permissive and conformationally mobile site in the sensing domain via minimal linkers connected to the purple and orange gate post residues. The sensing domain should effectively replace the two bulge residues.
- The strategy to create an RFP-based biosensor is essentially the same as for GFP-based biosensors, but the sensing domain should be connected to the orange gate post residue and the Glu residue immediately following the purple gate post residue. The sensing domain should effectively replace only the second bulge residue.
- As many insertion sites as possible within the sensing domain should be tested. A direct response (fluorescence increase) is generally preferred to an inverse response (fluorescence decrease), but variants with flipped responses can arise during subsequent optimization.
- Systematic deletion of linker residues should be attempted to find the minimal linkers that allow the protein to fold efficiently and be soluble. If an initial screen produces few fluorescent variants, a longer linker may be required. It is advisable to copy linkers from high-performance biosensors.
- For linker optimization, pairs of residues should be systematically randomized to all possible amino acids (or a subset thereof). The gate post positions should be screened first, with subsequent libraries progressing stepwise towards the sensing domain.
- Substantial improvements in biosensor $\Delta F/F_{min}$ and folding efficiency are likely to arise during directed evolution using libraries in which the complete gene sequence has been randomly mutated.

sensing domain is connected to an FP at a position close to the center of the FP β-barrel, where the FP's intrinsic chromophore resides (Fig. 1a). This connection is most commonly achieved by insertion of a circularly permuted FP (cpFP) into a sensing domain (Fig. 1b). A cpFP is a topologically mutated FP that has the original amino (N) and carboxy (C) termini connected with a polypeptide linker, and new termini introduced elsewhere in the protein. Less commonly, a sensing domain (Fig. 1c), which could itself be circularly permuted (Fig. 1d), is inserted into an FP[38,39].

Mechanistically, analyte binding to the sensing domain induces a conformational change of the polypeptide backbone and/or amino acid side chains at the FP insertion site, which is proximal to the chromophore. These changes alter the chromophore protonation state, extinction coefficient or quantum yield, leading to changes in fluorescence intensity[40]. Following optimization using protein engineering methods such as linker optimization and directed evolution[11,41], single FP-based biosensors can be made to exhibit analyte-dependent fluorescence intensity changes as high as two orders of magnitude (for example, GCaMP5F and jGCaMP7c, with $\Delta F/F_{min} = \sim 150$, where $\Delta F$ is the change in fluorescence intensity, and $F_{min}$ is the minimum fluorescence intensity)[19,21].

## Defining the FP connection site: gate posts and bulge

The field of single FP-based biosensor development started with a serendipitous discovery. In 1998, Geoffrey Baird, then a PhD student in the laboratory of Roger Y. Tsien, was screening a library of randomly mutated cyan FP variants when he discovered a variant in which Tyr145 had been replaced with the FKTRHN hexapeptide[5]. Baird and Tsien went on to demonstrate that insertion of calmodulin (CaM) or a zinc finger at that same site of yellow FP rendered the fluorescence intensity dependent on $Ca^{2+}$ and $Zn^{2+}$, respectively. The CaM insertion into yellow FP was designated as camgaroo1—the first example of a single FP-based biosensor. Now, 20 years after Baird's discovery, we know in no uncertain terms that the region immediately adjacent to Tyr145 is a uniquely privileged site for sensing domain connection.

We find it instructive to define this connection site as two gate post residues that flank a compliant two-residue bulge in the protein structure where insertions and fusions are well tolerated (Fig. 2a). The bulge is a region where the regular hydrogen bond pattern of the FP β-barrel is disrupted and the backbone bulges towards the exterior, with two consecutive residues having their side chains pointed towards the solvent. The gate post and bulge motif correspond to residues 145–148 as numbered in GFP[42,43], where 145 and 148 are the gate posts and 146 and 147 are the bulge, and this is conserved across FP homologs (Figs. 2b–d and 3). In this work, we define the residues that are structurally aligned with Tyr145 and His148 of GFP as the purple and orange gate posts, respectively.

In the context of biosensors, positions in the primary sequence that come before position 145 or after position 148 (that is, positions further into the GFP domain) are rarely mutated (Supplementary Tables 1 and 2). In contrast, positions that come after position 145 or before position 148 represent the linkages to the sensing domain and are highly diverse among biosensors. The gate post residues themselves exhibit an intermediate level of diversity, and X-ray crystal structures of biosensors reveal that these residues are frequently (for the purple gate post) or occasionally (for the orange gate post) in conformations similar to those in the parental FPs (as in Fig. 2). Note that the numbering of these positions will differ between FPs from different species due to differences in protein length, and the numbers are quite different in cpFPs and biosensors. For proteins that have had their structure determined by X-ray crystallography, we number these residues as in the specific Protein Data Bank (PDB) structure.

## Replacing the bulge with a sensing domain

Regardless of FP species of origin, biosensor topology or structure of the sensing domain, X-ray structures of biosensors reveal that the sensing domain fusion or insertion is always located in the bulge region between the gate posts (Fig. 4). Indeed, with the benefit of hindsight, Baird and Tsien's cyan FP variant with Tyr145 replaced with FKTRHN should better be described as 'Tyr145Phe with KTRHN inserted in the bulge region'[5]. Notably, Tyr145Phe is known to improve the folding efficiency of GFP[44]. We expect that this insight extends to essentially all reported biosensors, including those that were designed to have the sensing domain fusion (or insertion) adjacent to the bulge. The sole example of a structurally characterized biosensor with the sensing domain connected outside of the bulge is mRuby-based RCaMP (Fig. 4f), which has the sensing domain effectively located immediately after the orange gate post (between residues 143 and 144) of mRuby (Fig. 3) in the $Ca^{2+}$-bound state[45].

There are two reasons why the gate posts and bulge constitute a privileged site. The first is that this is where the chromophore most closely approaches the β-barrel shell that isolates it from the environment. More specifically, it is where the Tyr-derived protonated phenol (in equilibrium with the deprotonated phenolate with a $pK_a$ that is typically within a few units of the physiological pH) is in closest proximity to the surrounding β-barrel. Permutation or sensing domain insertion between the gate posts creates a gap that exposes the phenol(ate) moiety, making its $pK_a$ sensitive to the local

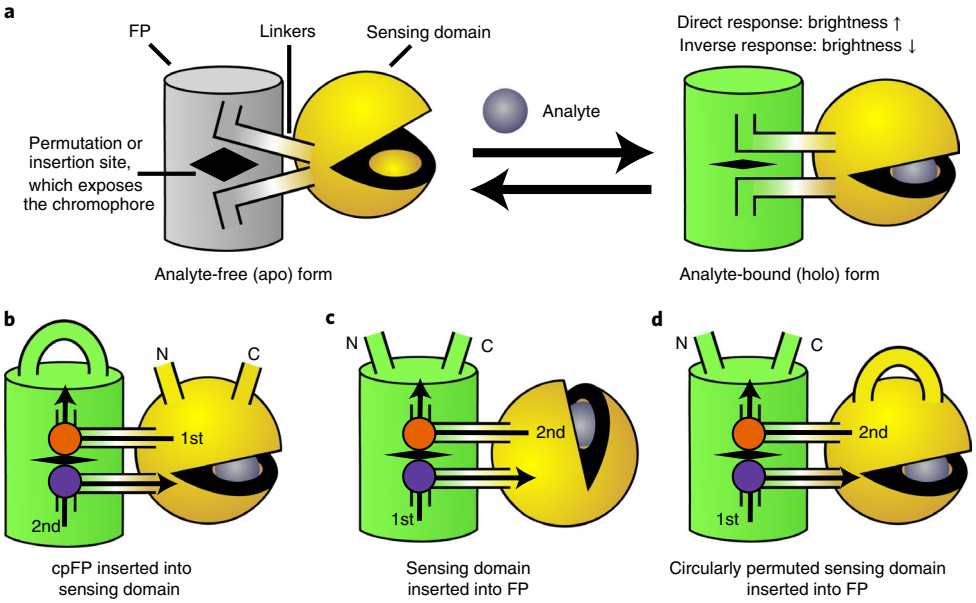


**Fig. 1 | Single FP-based biosensors and topologies used in biosensor creation. a**, Schematic of the structure and generic mechanism of a biosensor. **b**, A biosensor constructed by insertion of a cpFP into a sensing domain. The dark orange and purple circles represent the gate post residues that flank the insertion site. In this topology, the linker that ends at the orange gate post comes first in the primary sequence and the linker that starts from the purple gate post comes second. **c**, A biosensor constructed by insertion of a sensing domain (connected via its normal N and C termini) into an FP. In this topology, the linker that starts at the purple gate post comes first in the primary sequence and the linker that ends at the orange gate post comes second. **d**, A circularly permuted sensing domain inserted into GFP. The order of linkers is the same as in **c**. Representation inspired by ref. [5].

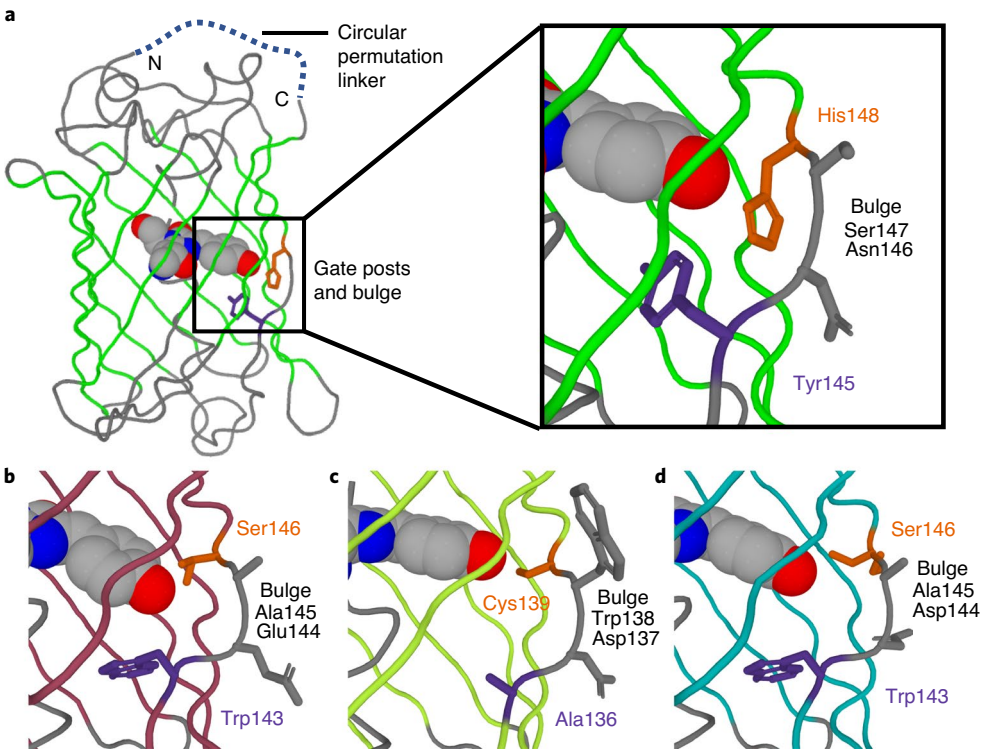


**Fig. 2 | Representation of the gate post and bulge residues in FPs derived from different species. a**, Structure of *Aequorea* GFP (PDB ID: 1EMA)[42], highlighting the position of the gate posts and the bulge. In cpFP-based biosensors (as in Fig. 1b), a polypeptide linker (dashed line) connects the original N and C termini. **b**, Analogous region of RFP[95]. The structure of mCherry[53] (PDB ID: 2H5Q)[95] is used to represent *Discosoma* species RFP-derived mApple[55], *Entacmaea quadricolor* eqFP611-derived mRuby[56] and *E. quadricolor* eqFP578-derived FusionRed[57]. **c**, Analogous region of *Branchiostoma lanceolatum*-derived mNeonGreen (PDB ID: 5LTR)[65]. **d**, Analogous region of *Clavularia* coral-derived mTFP1 (PDB ID: 2HQK)[62]. For both mCherry and mTFP1, Ser146 was modeled as a mixture of two conformations. β-sheet secondary structures are represented in green, red, green-yellow and teal for GFP-, RFP-, mNeonGreen- and mTFP1-derived biosensors, respectively.

| FP | Sequence in vicinity of bulge | PDB ID |
|---|---|---|
| EGFP | 145 148<br>NILGHKLEYNYNSHNVYIMADKQK | 1EMA |
| mTFP1 | 143 146<br>GPVMQKKTTGWDASTERMYVRDGV | 2HQK |
| mNeonGreen | 136 139<br>GPVMTNSLTAADWCRSKKTYPNDK | 5LTR |
| mPapaya | 145 148<br>GPVMKKMTTNWEASTEKIVPVPKQ | 1XAE (zFP538) |
| mCherry | 143 146<br>GPVMQKKTMGWEASSERMYPEDGA | 2H5Q |
| mApple | 143 146<br>GPVMQKKTMGWEASEERMYPEDGA | 2H5Q (mCherry) |
| mRuby | 140 143<br>GAVMQKKTKGWEPNTEMMYPADGG | 3U0N |
| FusionRed | 140 143<br>GPVMQKKTLGWEASTETMYPADGG | 6U1A |
| mEos2 | 139 142<br>GPVMQKKTLKWEPSTEKMYVRDGV | 5DTL |

**Fig. 3 | Gate posts, bulges and adjacent sequences in selected FPs.** For each FP, the sequence (colored according to the approximate color of the FP fluorescence) is shown from ten residues before the first gate post (purple and underlined) to ten residues after the second gate post (orange and underlined). The bulge residues are shown in black. The structures used were 1EMA[42], 2HQK[62], 5LTR[65], 1XAE (zFP538)[66], 2H5Q[95], 3U0N[45], 6U1A[96] and 5DTL[97]. EGFP, enhanced GFP.

microenvironment. For both GFP and red FP (RFP), the phenol form absorbs shorter-wavelength light and is either much less fluorescent or emits long Stokes shift fluorescence[46,47] due to excited-state proton transfer and emission from the phenolate state. The phenolate form absorbs longer-wavelength light and is typically the brighter fluorescent state. In the case of R-GECO1 (Fig. 4e), $Ca^{2+}$ binding causes a shift in $pK_a$ from 8.9 (that is, mostly the dark phenol form at physiological pH) to 6.6 (that is, mostly the bright phenolate form at physiological pH)[11]. A similar change in $pK_a$ from 8.0 to 7.1 has been reported for GCaMP6m[48]. These $pK_a$ changes, combined with changes in quantum yield and extinction coefficient, fully explain the fluorescence response of $Ca^{2+}$ biosensors[40,48]. An unfortunate consequence of these $pK_a$-dependent mechanisms is that biosensors may artifactually respond to pH changes in the physiological range.

The second reason why the gate posts and bulge are privileged is that this region, of the otherwise monolithic 11-stranded FP β-barrel, is permissive towards insertion and introduction of new termini in cpFP variants. Indeed, the bulge region—the one spot in the β-barrel where the near-perfect hydrogen bonding pattern is interrupted—is one of the few locations where new termini can be introduced without disrupting the ability of the cpFP to become fluorescent[5,49,50]. Plausibly, the bulge is a universal feature of FPs as it enables the chromophore, which is the largest protuberance from the central α-helix, to fit within a β-barrel (of fixed radius) that is slightly too small to accommodate it.

While the gate post plus bulge region seems to constitute a privileged site across all classes of FP, there appear to be (based on the limited examples currently available) small but critical differences in the details of exactly where the sensing domain should be inserted. While past efforts have, necessarily, narrowed in on the vicinity of the bulge region as the ideal insertion site, a range of approaches have been taken with respect to precisely where within this region the protein should be permuted, whether bulge and gate post residues should or should not be deleted, and whether or not to add in additional linker residues. As the published examples are the ones that have worked to varying degrees, we can begin to build on this foundation and provide guidelines for creating effective biosensors based on different classes of FP.

## Bulge and gate posts in different classes of FP

In wild-type GFP, the purple gate post is Tyr145, which is directly below the chromophore with its side chain directed towards the interior of the barrel (Fig. 2a). One surprising revelation from examining the sequences of all GFP-based biosensors is that there is not one example of a biosensor that retains a Tyr at this position. Much more frequent are Phe (a superfolder substitution)[44] and other non-polar amino acids such as Leu, with numerous exceptions (Supplementary Tables 1 and 2). Based on the available crystal structures (for example, the iGABASnFR prototype)[29], Phe at this position can be packed similarly to Tyr in wild-type GFP. However, the structures of early GCaMP variants reveal that a polar residue (such as Thr in GCaMP2 and GCaMP3)[19,51,52] at this position can allow the main chain to be pulled away from its wild-type β-sheet-like conformation, further exposing the chromophore to the external environment. In later-generation GCaMP variants, this residue was mutated to a non-polar Leu, but the chain remains pulled away from the β-barrel (Fig. 4a). The particularly exposed chromophores of GCaMP variants are probably contributing factors to the particularly large fluorescent responses of this family of biosensors.

The orange gate post residue of GFP is His148, which has its imidazole side chain in the plane of the β-sheet with one edge directed towards the solvent and a nitrogen on the other edge forming a hydrogen bond with the chromophore phenol(ate) (Fig. 2a). His is often retained at this position in biosensors, but polar and charged residues such as Glu and Asp are also frequently observed, with many exceptions. As will be discussed in the section 'Mechanisms', crystal structures reveal that this residue is probably a key player in the fluorescent response mechanism of many biosensors.

Four different RFPs have been incorporated into biosensors to date: mCherry[53,54], mApple[11,55], mRuby[45,56] and FusionRed[57,58] (Fig. 3). Like GFP, the purple gate post residue (Trp143 using mCherry numbering) is buried in the interior of the protein, directly below the chromophore (Fig. 2b). All biosensors based on RFP strictly conserve Trp at this position, presumably because burying this large hydrophobic side chain provides a great deal of stability to the protein. The residue that immediately follows the purple gate post Trp143, Glu144 (that is, the first residue of the bulge), is highly conserved in the roughly ten RFP-based biosensors reported to date. Inspection of the structures of mCherry, R-GECO1, RCaMP and K-GECO1 (Fig. 5) reveals that the carboxylate side chain of this Glu forms a hydrogen bond to a His residue (His172 using mCherry numbering) two β-strands away. This hydrogen bond probably contributes to protein stability and Glu should probably be retained when constructing prototype biosensors. The CH-GECO2.1 $Ca^{2+}$-biosensor has a Leu at this position, revealing that, in some cases, an RFP can tolerate the absence of this hydrogen bond, and that this residue is a candidate for mutation during later optimization[54].

In many of the RFP biosensors reported to date, the orange gate post residue is Val (a legacy of the cpmApple domain from R-GECO1), with a few examples of polar residues such as Asn and Thr. In the R-GECO1 structure, this residue is effectively pulled away from the body of the FP and the side chain interacts with the solvent (Figs. 4e and 5b). It is probable that polar or charged

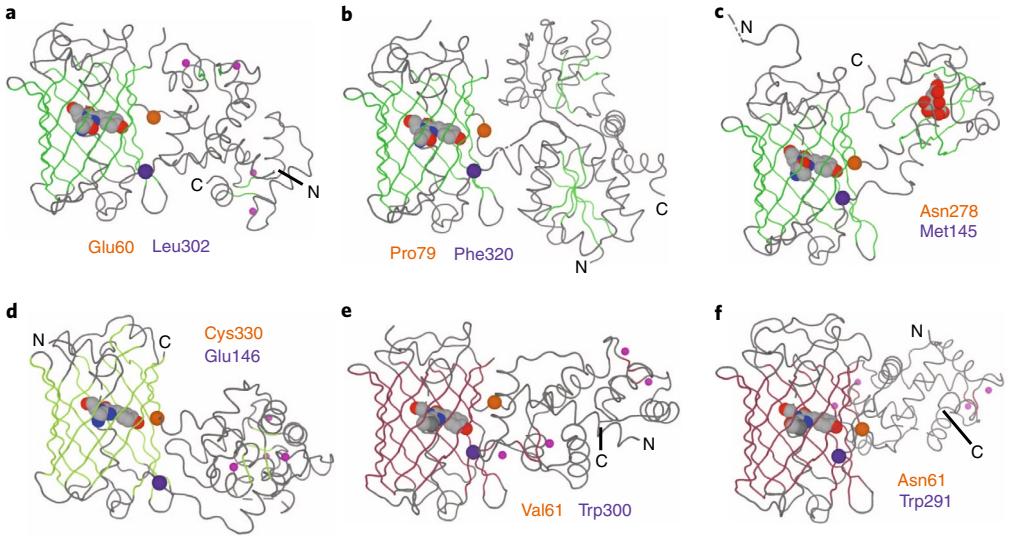


**Fig. 4 | Representative X-ray crystal structures of biosensors. a**, $Ca^{2+}$-bound biosensor GCaMP6m (PDB ID: 3WLD) based on insertion of cpGFP between the RS20 peptide and CaM[20,98]. $Ca^{2+}$ ions are represented as magenta spheres. **b**, Nicotine biosensor iNicSnFR1 (PDB ID: 6EFR) based on an engineered choline-/betaine-binding protein-sensing domain[26]. **c**, Citrate-bound biosensor Citron1 (PDB ID: 6LNP) based on insertion of the periplasmic citrate-binding domain of a bacterial sensor histidine kinase into GFP[82]. **d**, $Ca^{2+}$-bound biosensor NCaMP7 (PDB ID: 6XW2) based on insertion of CaM–RS20 into mNeonGreen[68]. **e**, $Ca^{2+}$-bound biosensor R-GECO1 (PDB ID: 4I2Y) based on insertion of cpmApple[55] between the RS20 peptide and CaM[11,45]. **f**, $Ca^{2+}$-bound biosensor RCaMP (PDB ID: 3U0K) based on insertion of cpmRuby[56] between RS20 and CaM[45]. The biosensors in **a**, **b**, **e** and **f** are examples of cpFPs inserted into a sensing domain (Fig. 1b). The biosensors in **c** and **d** are examples of sensing domains inserted into an FP (Fig. 1c). For all of the structures, the Cα of the gate post residue that aligns with Tyr145 of GFP (equivalent to Ala136 of mNeonGreen or Trp143 of RFP) is shown as a purple sphere and the Cα of the gate post residue that aligns with His148 of GFP (equivalent to Cys139 of mNeonGreen or Ser146 of RFP) is shown as an orange sphere. All residues are numbered according to their respective X-ray crystal structures.

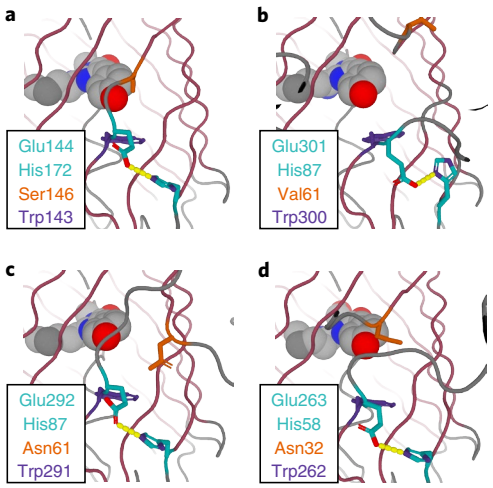


**Fig. 5 | Conservation of bulge residue Glu144 in RFP- and cpRFP-based biosensors. a**, Parent RFP, here represented using mCherry (PDB ID: 2H5Q)[95]. **b**, $Ca^{2+}$-bound R-GECO1 (PDB ID: 4I2Y)[11,45]. **c**, $Ca^{2+}$-bound RCaMP (PDB ID: 3U0K)[45]. **d**, $Ca^{2+}$-bound K-GECO1 (PDB ID: 5UKG)[58].

residues at this position could, by analogy with the GFP-based biosensors, serve to modulate interactions with the chromophore and provide large changes in fluorescence.

Of the various RFPs, the cpmApple scaffold from R-GECO1 has proven to be particularly amenable to engineering of new biosensors with substantial fluorescence responses (Supplementary Tables 1 and 2). However, a limitation of cpmApple-based biosensors is that they tend to exhibit blue-light-induced photoswitching behavior and tend to accumulate as fluorescent puncta in neurons during long-term expression[38,59]. mRuby- and FusionRed-based biosensors do not suffer from photoswitching or puncta formation to the same extent as cpmApple[58,60], but have not yet proven to be as versatile for the creation of biosensors for analytes other than $Ca^{2+}$.

Other FPs that have been converted into biosensors include mNeonGreen (Fig. 2c)[61], monomeric teal fluorescent protein 1 (mTFP1; Fig. 2d)[62], mPapaya (Fig. 3)[63] and the photoconvertible mEos2 (Fig. 3)[64]. The crystal structures of these parental FPs reveal the presence of the characteristic bulge region[62,65,66]. We suggest that future efforts to create biosensors from mNeonGreen, mTFP1, mPapaya or mEos2 should incorporate the FP domains of mNG-GECO1 (or NCaMP7)[67,68], ZnGreen1 (ref. [69]), Y-GECO2 (ref. [70]) or CaMPARI[50] biosensors, respectively. For mNeonGreen-, mTFP1- and mEos2-based biosensors, the two bulge residues should be replaced with the sensing domain (as recommended for GFPs), but mPapaya-based biosensors should probably retain the Glu that follows the purple gate post (as recommended for RFPs).

## Position within the sensing domain to attach the FP

A key challenge in constructing a biosensor is to identify a location in the sensing domain that both undergoes a substantial conformational change and also tolerates cpFP insertion (Fig. 1b) or FP fusion (Fig. 1c,d). This challenge is most easily overcome in cases where the normal termini of the sensing domain are close together and undergo a substantial change in separation distance upon analyte binding[71]. In such a situation, the sensing domain can be directly inserted into the FP via its normal termini (Fig. 1c). However, when the termini are not close together in space, or do not change their distance upon analyte binding, a suitable position within the sensing domain must be identified.

The archetypical analyte-responsive sensing domains are the solute-binding proteins (SBPs; also commonly known as

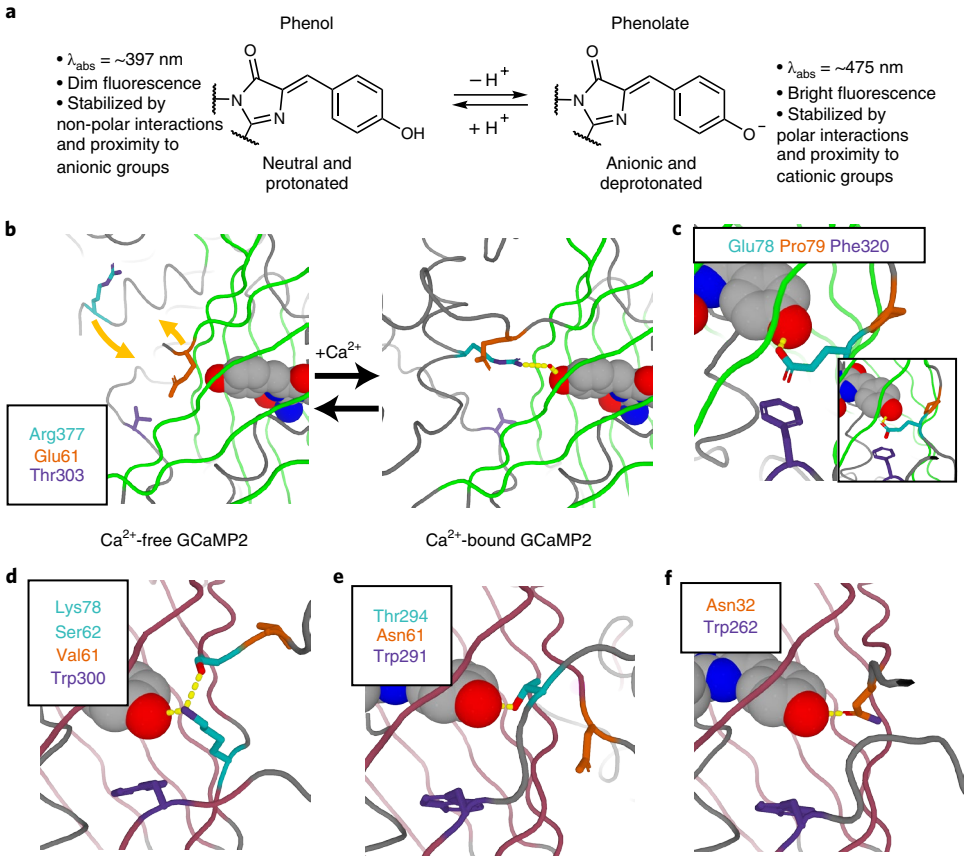


**Fig. 6 | GFP- and RFP-based biosensor structures provide insight into the mechanism of the fluorescence response. a**, A major contributing factor to the fluorescence response of many biosensors is a shift in the equilibrium between the phenol and phenolate forms of the FP chromophore due to changes in the environment of the ionizable phenol(ate) moiety. A shift towards the phenolate form upon analyte binding produces an increase in fluorescence (direct response), whereas a shift towards the phenol form produces a decrease in fluorescence (inverse response). **b**, GCaMP2 in the $Ca^{2+}$-free (left; PDB ID: 3EKJ) and $Ca^{2+}$-bound forms (right; PDB ID: 3EVR)[51]. The residue numbers differ by one from the GCaMP6m structure (Fig. 3a). The orange arrows represent the movement of Arg377 and the displacement of the first linker (RS20 to the orange gate post). **c**, iNicSnFR1 (PDB ID: 6EFR)[26], probably in the unbound state (see main text). Inset: iAChSnFR0.4 (PDB ID: 6URU)[31] in the unbound state. **d**–**f**, Structures of $Ca^{2+}$-bound R-GECO1 (PDB ID: 4I2Y)[45] (**d**), RCaMP (PDB ID: 3U0K)[45] (**e**) and K-GECO1 (PDB ID: 5UKG)[58] (**f**), RFP-based $Ca^{2+}$ biosensors showing interactions that probably contribute to stabilization of the chromophore in the brightly fluorescent phenolate state. In **b**–**f**, the purple gate post residue is shown with purple sticks and the orange gate post residue is shown with orange sticks. Other residues proposed to be involved in the mechanism are shown with teal carbon atoms, with key hydrogen bonds shown as dashed yellow lines.

periplasmic-binding proteins or PBPs), which are near-universal binding proteins used by prokaryotic organisms to capture small molecules[72]. SBPs undergo a characteristic clam-shell (or Venus flytrap)[73] motion that has been exploited for the engineering of single FP-based biosensors including ones for maltose[74,75], glucose[76,77], trehalose[75], phosphonates[78], glutamate[79], GABA[29], acetylcholine[31] and nicotine[26]. The most effective of these biosensors tend to have circularly permuted GFP (cpGFP) inserted into the hinge region where the conformational changes associated with the clam-shell motion are greatest. Similarly, G protein-coupled receptor (GPCR)-based biosensors are based on insertion of a cpFP into the third intracellular loop of the GPCR[22–25] where activation-dependent conformational changes are the greatest[80].

In cases where there are structures of a sensing domain in both the ligand-bound and ligand-free forms, a residue-by-residue analysis of the binding-dependent changes in inter-residue dihedral angles can reveal the locations that undergo the largest conformational changes[32,74]. In the case of the phosphonate biosensor, empirical screening led to the identification of an insertion site that produced substantial allosteric coupling between the SBP and the inserted cpGFP[78]. Only later, once the structure was solved, did it become clear that this insertion point was in the region with the greatest change in inter-residue dihedral angle. The failure of other insertion sites to produce functional biosensors was rationalized by the absence of substantial changes in the dihedral angles.

If the structure or a homology model of only one form of the sensing domain (that is, liganded or unliganded) is available, possible insertion points can be chosen using intuition and visual inspection. Positions that are close to the hinge region, close to the ligand-binding site or have higher than average B-factors (a proxy for disorder and mobility in X-ray crystal structures) have the greatest chance of tolerating FP insertion and also producing functional biosensors. Regardless of the rationale used to choose insertion sites, it is recommended to test as many different variants as possible. Even when screening large numbers of insertion variants, discovery of a prototype biosensor with $\Delta F/F_{min} = \sim 0.25$ should be considered a success and a promising starting point for future engineering. Discovery of a prototype with $\Delta F/F_{min} > 1$ is a home run. A prototype with this level of response can almost certainly be greatly improved, in terms of $\Delta F/F_{min}$, by subsequent rounds of linker optimization and directed evolution.

In one illustrative example that takes the idea of testing as many sites as possible close to its logical extreme, a cleverly designed transposon was used to generate libraries in which cpGFP was randomly and exhaustively inserted into *Escherichia coli* D-maltose-binding protein[75]. Having generated 210 of 370 possible insertions, it was found that the best-performing variant had $\Delta F/F_{min} = 2.4$, which was improved to $\Delta F/F_{min} = 8$ following linker optimization. The exceptional performance of the initially discovered variant emphasizes the importance of screening as many insertion sites as possible when engineering a new biosensor.

## Linker length and composition

When screening initial insertion or fusion sites, researchers typically add a few additional residues of linker between the FP domain and sensing domain to minimize steric clashes and allow each domain to fold properly. Although this undoubtedly increases the chances of discovering sites in the sensing domain that are permissive to insertion, longer linkers are antithetical to the goal of developing high-performance biosensors. Biosensors work by communicating a conformational change in the sensing domain, which may be relatively subtle, to a change in the chromophore environment (Fig. 6a). The longer the linkers, the more likely it is that this conformational change will be dampened or lost due to bond rotations in linker residues that are not in the vicinity of the chromophore. When no linkers or very short linkers are used, regions of the FP domain, or regions of the sensing domain, can essentially become partially unraveled and serve as a linker.

In summary, linkers should be as short as possible (to maximize coupling between the sensing domain the FP), but not so short that the folding of the protein is adversely affected. For discovery of initial prototypes, linker lengths of zero, one, two (and perhaps three) residues should be attempted. The variant with the best compromise of brightness (higher is better) and linker length (shorter is better), or with the largest $\Delta F/F_{min}$ value, should be carried on to subsequent rounds of linker optimization. Brightness is determined by the intrinsic properties of the FP chromophore (quantum yield, extinction coefficient and p$K_a$), the chromophore formation efficiency and the concentration of properly folded protein (in *E. coli* or otherwise). In the case of membrane-associated biosensors (for example, ones that use a GPCR[22,23] or voltage-sensitive domain[81]), brightness, membrane localization and $\Delta F/F_{min}$ must be assessed in mammalian cells. Linker optimization should be done by screening libraries in which residues are randomized to all possible amino acids (or a subset thereof), either individually or in pairs. Due to their high probability of being integral to the fluorescence response mechanism, gate post positions should be screened first and subsequent libraries should start with the next closest residues, and progress towards the sensing domain.

## Mechanisms

The first glimpses of a biosensor's molecular mechanism were provided by the X-ray crystal structures of GCaMP2 in 2008 and 2009 (refs. [51,52]) and, to date, GCaMP2 remains the only biosensor for which both analyte-free and analyte-bound structures are available (Fig. 6b)[51,52]. In the $Ca^{2+}$-free structure, electron density was not observed for much of the CaM domain and all of the RS20 peptide, indicating that those regions are conformationally mobile or unstructured. Arg377 of CaM is distant from the chromophore, which is predominately in the phenol (protonated) form. Upon binding of $Ca^{2+}$, compaction of the CaM–RS20-sensing domain brings the guanidinium side chain of Arg377 into a water-bridged interaction with the chromophore. This interaction stabilizes the phenolate (anionic) form of the chromophore, resulting in a greater fraction of the protein in the cyan light-excitable, brightly fluorescent state[48]. Similarly, the structure of $Ca^{2+}$-bound NCaMP7 (PDB ID: 6XW2) suggests that a residue of CaM—specifically the phenol moiety of Tyr225—stabilizes the phenolate form of the chromophore through a water-bridged interaction (Fig. 4d)[68].

Examining the X-ray crystallographic structures of other biosensors reveals the diversity of mechanistically relevant interactions that can occur between the chromophore and the surrounding protein microenvironment. For example, in the maltose-bound structure of a maltose biosensor (with cpGFP insertion at position 175), His177 (equivalent to the orange gate post residue of GFP; Fig. 2a) is hydrogen bonded with the chromophore, probably stabilizing the fluorescent phenolate form[74]. In the structure of the nicotine biosensor iNicSnFR1, the phenol group of the chromophore is hydrogen bonded to the carboxylic acid side chain of Glu78. This would be expected to stabilize the protonated (phenol) form of the chromophore (Fig. 6c). However, this structure was originally interpreted as representing the nicotine-bound form, although nicotine was not clearly resolved in the binding site[26]. We suggest that this structure is more likely to represent the apo dark-state structure. Support for this suggestion comes from the structure of the apo state of the acetylcholine biosensor iAChSnFR0.4 (PDB ID: 6URU; Fig. 6c, inset), in which Glu78 is in an identical conformation[31]. Possibly, in both iAChSnFR and iNicSnFR, ligand binding induces a conformational change that moves the Glu78 side chain away from the chromophore, leading to formation of the phenolate form and an increase in fluorescence. A similar line of reasoning was used to guide the rational engineering of the long Stokes shift $Ca^{2+}$ biosensor REX-GECO1 (ref. [12]).

Further insight into the mechanism of biosensors can be obtained by examining the structures for the $Ca^{2+}$-bound states of the RFP-based biosensors. In the $Ca^{2+}$-bound state of R-GECO1, the phenolate form of the chromophore is stabilized by an electrostatic interaction with the amine group of Lys78 (Fig. 6d). Lys78 is located on an adjacent β-strand of cpmApple and is presumably brought into this conformation only when the CaM–RS20-sensing domain is in its compacted $Ca^{2+}$-bound state. In the $Ca^{2+}$-bound state of RCaMP, the phenolate moiety of the chromophore is engaged in a hydrogen bonding interaction with Thr294, which is on the second linker (cpRFP to CaM) (Fig. 6e). Finally, in the $Ca^{2+}$-bound state of K-GECO1, the phenolate moiety of the chromophore is engaged in a hydrogen bonding interaction with orange gate post residue Asn32 at the end of the first linker (Fig. 6f). In citrate-bound Citron1 (PDB ID: 6LNP), orange gate post residue Asn278 makes an essentially identical interaction with the chromophore (Fig. 4c)[82].

Taken together, these structures reveal the diversity of interactions that may contribute to the modulation of biosensor fluorescence. Specifically, interactions can occur between the chromophore and a side chain from an adjacent strand of the β-barrel (that is, as in R-GECO1) or a side chain from the first (that is, as in K-GECO1, iNicSnFr and MBP175-cpGFP.L1-HL) or second linker (that is, as in RCaMP), or even from the sensing domain itself (that is, as in GCaMP).

## Outlook

Following a first decade (1999–2009) of slow and intermittent progress in the field of single GFP-based biosensors, the second decade (2009–2019) was a renaissance. In the second decade, we have learned about the heights of performance that biosensors can reach, we have learned that the color palette can be expanded to enable multiparameter imaging and we have learned much more about the range of sensing domains that can be exploited. As discussed in this Perspective, we are also beginning to gain insight into the structures and mechanisms of biosensors. Undoubtedly, these insights will lead us towards a wider range of high-performance, user-friendly biosensors.

One trend for the near future will be expansion of the biosensor color palette into the near-infrared through the use of biliverdin-binding FPs (BV-FPs)[83]. A first step in this direction was a report

of the inverse-response NIR-GECO1 $Ca^{2+}$ biosensor[84,85] based on insertion of a CaM–RS20 into the monomeric infrared fluorescent protein (mIFP)[86] BV-FP (topology as in Fig. 1e). For in vitro applications, NIR-GECO1 creates new opportunities for crosstalk-free multiparameter imaging with red, green or cyan-yellow FRET-based biosensors. Furthermore, the wavelengths required for NIR-GECO1 excitation are well separated from the wavelengths used for optogenetic activation with channelrhodopsin-2. It remains unclear whether BV-FPs will be versatile building blocks for biosensor development, and structural characterization of NIR-GECO1 will probably be required before its fluorophore module can be effectively transplanted into other sensing domains.

Another expected trend is the utilization of a broader range of sensing domains. There are >140,000 protein structures in the PDB (rcsb.org)[87], demonstrating that there are an abundance of building blocks to work with. Even within well-established classes of sensing domains, researchers have only scratched the surface in terms of biosensor creation. The number of SBP structures in the PDB increased from 114 in 2010 to ~500 in 2016 (ref. [88]). Approximately 400 of these SBPs are known to bind to ~150 different target ligands, with another ~100 SBPs having yet-undetermined ligand specificity. SBPs are amenable to computational protein design to change the ligand-binding specificity, further expanding the range of possible biosensors[89]. Another promising, yet largely unexploited, source of sensing domains are the gluconate operon repressor family transcription factors[90]. These proteins are composed of an N-terminal DNA-binding domain that is allosterically controlled by a C-terminal molecular recognition domain that can bind a variety of ligands of interest[91]. A recently reported pyruvate biosensor, PyronicSF, was constructed by insertion of cpGFP into the gluconate operon repressor family transcription factor PdhR[92]. Another promising source of sensing domains are lipocalins[93], a family of proteins that bind to a variety of small molecules and have recently been exploited as the sensing domain of a serotonin biosensor[32]. Perhaps the most revolutionary new class of biosensors are those based on insertion of a cpFP into a GPCR[22–25]. This highly modular design may be applicable to many of the hundreds of GPCRs that are in the human genome alone[94].

As the example of GCaMP demonstrates, it is possible to develop biosensors that have large fluorescence changes ($\Delta F/F_{min}$ = ~10–100), analyte affinity tuned to the physiologically relevant range, and brightness that rivals that of the parent FP. Indeed, precedent suggests that $\Delta F/F_{min} > 10$ and brightness > 50% of the parent FP may represent a critical threshold in terms of user-friendliness and robust utility for in vivo molecular imaging experiments. The first biosensor to cross this threshold was GCaMP3, which was widely regarded as the first broadly useful $Ca^{2+}$ biosensor (with $\Delta F/F_{min} = 12$ and 60% of the brightness of enhanced GFP in the $Ca^{2+}$-bound state)[19]. With this Perspective, we hope to have provided guidelines, and inspiration, that will soon lead to a greater number of biosensors that pass this performance threshold and revolutionize entire research disciplines, just as $Ca^{2+}$ biosensors have done for neuroscience.

Received: 10 October 2019; Accepted: 2 December 2020;
Published online: 8 February 2021

## References

1. Greenwald, E. C., Mehta, S. & Zhang, J. Genetically encoded fluorescent biosensors illuminate the spatiotemporal regulation of signaling networks. *Chem. Rev.* **118**, 11707–11794 (2018).
2. Lavis, L. D. Teaching old dyes new tricks: biological probes built from fluoresceins and rhodamines. *Annu. Rev. Biochem.* **86**, 825–843 (2017).
3. Smith, B. R. & Gambhir, S. S. Nanomaterials for in vivo imaging. *Chem. Rev.* **117**, 901–986 (2017).
4. Miyawaki, A. et al. Fluorescent indicators for $Ca^{2+}$ based on green fluorescent proteins and calmodulin. *Nature* **388**, 882–887 (1997).
5. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl Acad. Sci. USA* **96**, 11241–11246 (1999).
**This landmark paper describes the discovery of the GFP insertion site adjacent to residue 145, describes the first examples of single FP-based biosensors ($Ca^{2+}$ and $Zn^{2+}$) and summarizes the various possible topologies for single FP-based indicators.**
6. Kostyuk, A. I., Demidovich, A. D., Kotova, D. A., Belousov, V. V. & Bilan, D. S. Circularly permuted fluorescent protein-based indicators: history, principles, and classification. *Int. J. Mol. Sci.* **20**, 4200 (2019).
7. Kaczmarski, J. A., Mitchell, J. A., Spence, M. A., Vongsouthi, V. & Jackson, C. J. Structural and evolutionary approaches to the design and optimization of fluorescence-based small molecule biosensors. *Curr. Opin. Struct. Biol.* **57**, 31–38 (2019).
8. Carlson, H. J. & Campbell, R. E. Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. *Curr. Opin. Biotechnol.* **20**, 19–27 (2009).
9. Boffi, J. C., Knabbe, J., Kaiser, M. & Kuner, T. KCC2-dependent steady-state intracellular chloride concentration and pH in cortical layer 2/3 neurons of anesthetized and awake mice. *Front. Cell. Neurosci.* **12**, 7 (2018).
10. Mehta, S. et al. Single-fluorophore biosensors for sensitive and multiplexed detection of signalling activities. *Nat. Cell Biol.* **20**, 1215–1225 (2018).
11. Zhao, Y. et al. An expanded palette of genetically encoded $Ca^{2+}$ indicators. *Science* **333**, 1888–1891 (2011).
12. Wu, J. et al. A long Stokes shift red fluorescent $Ca^{2+}$ indicator protein for two-photon and ratiometric imaging. *Nat. Commun.* **5**, 5262 (2014).
13. Barykina, N. V. et al. FGCaMP7, an improved version of fungi-based ratiometric calcium indicator for in vivo visualization of neuronal activity. *Int. J. Mol. Sci.* **21**, 3012 (2020).
14. Su, S. et al. Genetically encoded calcium indicator illuminates calcium dynamics in primary cilia. *Nat. Methods* **10**, 1105–1107 (2013).
15. Ast, C. et al. Ratiometric Matryoshka biosensors from a nested cassette of green- and orange-emitting fluorescent proteins. *Nat. Commun.* **8**, 431 (2017).
16. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise $Ca^{2+}$ probe composed of a single green fluorescent protein. *Nat. Biotechnol.* **19**, 137–141 (2001).
**Together with Nagai et al.[33], this paper describes the development of the first cpFP-based $Ca^{2+}$ genetically encoded calcium indicators (GECIs).**
17. Tallini, Y. N. et al. Imaging cellular signals in the heart in vivo: cardiac expression of the high-signal $Ca^{2+}$ indicator GCaMP2. *Proc. Natl Acad. Sci. USA* **103**, 4753–4758 (2006).
18. Tian, L. et al. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* **6**, 875–881 (2009).
**This paper reports the first demonstration of $Ca^{2+}$ imaging of neuronal activity in awake behaving mice using the single FP-based GECI GCaMP3.**
19. Akerboom, J. et al. Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
20. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
21. Dana, H. et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).
22. Patriarchi, T. et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360**, eaat4422 (2018).
**Together with Sun et al.[23], this paper describes the development of the first single FP-based dopamine biosensors using a GPCR as the sensing moiety.**
23. Sun, F. et al. A genetically encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice. *Cell* **174**, 481–496.e19 (2018).
24. Jing, M. et al. A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nat. Biotechnol.* **36**, 726–737 (2018).
25. Feng, J. et al. A genetically encoded fluorescent sensor for rapid and specific in vivo detection of norepinephrine. *Neuron* **102**, 745–761.e8 (2019).
26. Shivange, A. V. et al. Determining the pharmacokinetics of nicotinic drugs in the endoplasmic reticulum using biosensors. *J. Gen. Physiol.* **151**, 738–757 (2019).
27. Helassa, N. et al. Ultrafast glutamate sensors resolve high-frequency release at Schaffer collateral synapses. *Proc. Natl Acad. Sci. USA* **115**, 5594–5599 (2018).
28. Marvin, J. S. et al. Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. *Nat. Methods* **15**, 936–939 (2018).
29. Marvin, J. S. et al. A genetically encoded fluorescent sensor for in vivo imaging of GABA. *Nat. Methods* **16**, 763–770 (2019).
30. Sun, F. et al. Next-generation GRAB sensors for monitoring dopaminergic activity in vivo. *Nat. Methods* **17**, 1156–1166 (2020).
31. Borden, P. M. et al. A fast genetically encoded fluorescent sensor for faithful in vivo acetylcholine detection in mice, fish, worms and flies. Preprint at *bioRxiv* https://doi.org/10.1101/2020.02.07.939504 (2020).
32. Zhang, S., Li, X., Drobizhev, M. & Ai, H. A fast high-affinity fluorescent serotonin biosensor engineered from a tick lipocalin. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.18.048397 (2020).

33. Nagai, T., Sawano, A., Park, E. S. & Miyawaki, A. Circularly permuted green fluorescent proteins engineered to sense $Ca^{2+}$. *Proc. Natl Acad. Sci. USA* **98**, 3197–3202 (2001).
34. Belousov, V. V. et al. Genetically encoded fluorescent indicator for intracellular hydrogen peroxide. *Nat. Methods* **3**, 281–286 (2006).
35. Souslova, E. A. et al. Single fluorescent protein-based $Ca^{2+}$ sensors with increased dynamic range. *BMC Biotechnol.* **7**, 37 (2007).
36. Nausch, L. W. M., Ledoux, J., Bonev, A. D., Nelson, M. T. & Dostmann, W. R. Differential patterning of cGMP in vascular smooth muscle cells revealed by single GFP-linked biosensors. *Proc. Natl Acad. Sci. USA* **105**, 365–370 (2008).
37. Berg, J., Hung, Y. P. & Yellen, G. A genetically encoded fluorescent reporter of ATP:ADP ratio. *Nat. Methods* **6**, 161–166 (2009).
38. Wu, J. et al. Genetically encoded glutamate indicators with altered color and topology. *ACS Chem. Biol.* **13**, 1832–1837 (2018).
**Together with Qian et al.[39], this paper demonstrates that it is possible to convert a biosensor based on 'insertion of a cpFP into a sensing domain' into a biosensor based on 'insertion of a permuted sensing domain into an FP'.**
39. Qian, Y., Rancic, V., Wu, J., Ballanyi, K. & Campbell, R. E. A bioluminescent $Ca^{2+}$ indicator based on a topological variant of GCaMP6s. *ChemBioChem* **20**, 516–520 (2019).
40. Molina, R. S. et al. Understanding the fluorescence change in red genetically encoded calcium ion indicators. *Biophys. J.* **116**, 1873–1886 (2019).
41. Wardill, T. J. et al. A neuron-based screening platform for optimizing genetically-encoded calcium indicators. *PLoS ONE* **8**, e77728 (2013).
42. Ormö, M. et al. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).
43. Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
44. Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88 (2006).
45. Akerboom, J. et al. Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics. *Front. Mol. Neurosci.* **6**, 2 (2013).
**This paper reports a multi-colored set of GECIs and the X-ray crystal structures of R-GECO1 and RCaMP1, which provided the first mechanistic insight into RFP-based GECIs.**
46. Chattoraj, M., King, B. A., Bublitz, G. U. & Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl Acad. Sci. USA* **93**, 8362–8367 (1996).
47. Kogure, T. et al. A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat. Biotechnol.* **24**, 577–581 (2006).
48. Barnett, L. M., Hughes, T. E. & Drobizhev, M. Deciphering the molecular mechanism responsible for GCaMP6m's $Ca^{2+}$-dependent change in fluorescence. *PLoS ONE* **12**, e0170934 (2017).
49. Topell, S., Hennecke, J. & Glockshuber, R. Circularly permuted variants of the green fluorescent protein. *FEBS Lett.* **457**, 283–289 (1999).
50. Fosque, B. F. et al. Neural circuits. Labeling of active neural circuits in vivo with designed calcium integrators. *Science* **347**, 755–760 (2015).
51. Akerboom, J. et al. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem.* **284**, 6455–6464 (2009).
**Together with Wang et al.[52], this paper reports the X-ray crystal structures of GCaMP, providing some of the first mechanistic insights into the $Ca^{2+}$-induced fluorescence response of a GECI.**
52. Wang, Q., Shui, B., Kotlikoff, M. I. & Sondermann, H. Structural basis for calcium sensing by GCaMP2. *Structure* **16**, 1817–1827 (2008).
53. Shaner, N. C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
54. Carlson, H. J. & Campbell, R. E. Circular permutated red fluorescent proteins and calcium ion indicators based on mCherry. *Protein Eng. Des. Sel.* **26**, 763–772 (2013).
55. Shaner, N. C. et al. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545–551 (2008).
56. Kredel, S. et al. mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS ONE* **4**, e4391 (2009).
57. Shemiakina, I. I. et al. A monomeric red fluorescent protein with low cytotoxicity. *Nat. Commun.* **3**, 1204 (2012).
58. Shen, Y. et al. A genetically encoded $Ca^{2+}$ indicator based on circularly permutated sea anemone red fluorescent protein eqFP578. *BMC Biol.* **16**, 9 (2018).
59. Abdelfattah, A. S. et al. A bright and fast red fluorescent protein voltage indicator that reports neuronal activity in organotypic brain slices. *J. Neurosci.* **36**, 2458–2472 (2016).
60. Dana, H. et al. Sensitive red protein calcium indicators for imaging neural activity. *eLife* **5**, e12727 (2016).
61. Shaner, N. C. et al. A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods* **10**, 407–409 (2013).
62. Ai, H.-W. W., Henderson, J. N., Remington, S. J. & Campbell, R. E. Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **400**, 531–540 (2006).
63. Hoi, H. et al. An engineered monomeric *Zoanthus* sp. yellow fluorescent protein. *Chem. Biol.* **20**, 1296–1304 (2013).
64. McKinney, S. A., Murphy, C. S., Hazelwood, K. L., Davidson, M. W. & Looger, L. L. A bright and photostable photoconvertible fluorescent protein. *Nat. Methods* **6**, 131–133 (2009).
65. Clavel, D. et al. Structural analysis of the bright monomeric yellow-green fluorescent protein mNeonGreen obtained by directed evolution. *Acta Crystallogr. D Struct. Biol.* **72**, 1298–1307 (2016).
66. Remington, S. J. et al. zFP538, a yellow-fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. *Biochemistry* **44**, 202–212 (2005).
67. Zarowny, L. et al. Bright and high-performance genetically encoded $Ca^{2+}$ indicator based on mNeonGreen fluorescent protein. *ACS Sens.* **5**, 1959–1968 (2020).
68. Subach, O. M. et al. Novel genetically encoded bright positive calcium indicator NCaMP7 based on the mNeonGreen fluorescent protein. *Int. J. Mol. Sci.* **21**, 1644 (2020).
69. Chen, Z. & Ai, H.-W. Single fluorescent protein-based indicators for zinc ion ($Zn^{2+}$). *Anal. Chem.* **88**, 9029–9036 (2016).
70. Zhao, Y. et al. Inverse-response $Ca^{2+}$ indicators for optogenetic visualization of neuronal inhibition. *Sci. Rep.* **8**, 11758 (2018).
71. Arai, S. et al. RGB-color intensiometric indicators to visualize spatiotemporal dynamics of ATP in single cells. *Angew. Chem. Int. Ed.* **57**, 10873–10878 (2018).
72. De Boer, M. et al. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **8**, e44652 (2019).
73. Mao, B., Pear, M. R., McCammon, J. A. & Quiocho, F. A. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J. Biol. Chem.* **257**, 1131–1133 (1982).
74. Marvin, J. S., Schreiter, E. R., Echevarría, I. M. & Looger, L. L. A genetically encoded, high-signal-to-noise maltose sensor. *Proteins* **79**, 3025–3036 (2011).
75. Nadler, D. C., Morgan, S.-A., Flamholz, A., Kortright, K. E. & Savage, D. F. Rapid construction of metabolite biosensors using domain-insertion profiling. *Nat. Commun.* **7**, 12266 (2016).
**This paper describes a systematic method for the discovery of single FP-based indicators through the use of transposons to create large libraries in which cpGFP is randomly inserted into the sensing domain.**
76. Mita, M. et al. Green fluorescent protein-based glucose indicators report glucose dynamics in living cells. *Anal. Chem.* **91**, 4821–4830 (2019).
77. Keller, J. P. et al. In vivo glucose imaging in multiple model organisms with an engineered single-wavelength sensor. Preprint at *bioRxiv* https://doi.org/10.1101/571422 (2019).
78. Alicea, I. et al. Structure of the *Escherichia coli* phosphonate binding protein PhnD and rationally optimized phosphonate biosensors. *J. Mol. Biol.* **414**, 356–369 (2011).
79. Marvin, J. S. et al. An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat. Methods* **10**, 162–170 (2013).
**This paper describes the development of iGluSnFr, a single FP-based glutamate indicator based on an SBP. iGluSnFr is the archetype for the ever-expanding SnFr series of biosensors.**
80. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR dynamics: structures in motion. *Chem. Rev.* **117**, 139–155 (2017).
81. Chamberland, S. et al. Fast two-photon imaging of subcellular voltage dynamics in neuronal tissue with genetically encoded indicators. *eLife* **6**, e25690 (2017).
82. Zhao, Y., Shen, Y., Wen, Y. & Campbell, R. E. High-performance intensiometric direct- and inverse-response genetically encoded biosensors for citrate. *ACS Cent. Sci.* **6**, 1441–1450 (2020).
83. Shcherbakova, D. M., Stepanenko, O. V., Turoverov, K. K. & Verkhusha, V. V. Near-infrared fluorescent proteins: multiplexing and optogenetics across scales. *Trends Biotechnol.* **36**, 1230–1243 (2018).
84. Qian, Y. et al. A genetically encoded near-infrared fluorescent calcium ion indicator. *Nat. Methods* **16**, 171–174 (2019).
85. Qian, Y. et al. Improved genetically encoded near-infrared fluorescent calcium ion indicators for in vivo imaging. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.08.032433 (2020).
86. Yu, D. et al. A naturally monomeric infrared fluorescent protein for protein labeling in vivo. *Nat. Methods* **12**, 763–765 (2015).
87. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
88. Scheepers, G. H., Lycklama a Nijeholt, J. A. & Poolman, B. An updated structural classification of substrate-binding proteins. *FEBS Lett.* **590**, 4393–4401 (2016).

89. Vongsouthi, V. et al. A computationally designed fluorescent biosensor for D-serine. Preprint at *bioRxiv* https://doi.org/10.1101/2020.08.18.255380 (2020).
90. Haydon, D. J. & Guest, J. R. A new family of bacterial regulatory proteins. *FEMS Microbiol. Lett.* **63**, 291–295 (1991).
91. Jain, D. Allosteric control of transcription in GntR family of transcription regulators: a structural overview. *IUBMB Life* **67**, 556–563 (2015).
92. Arce-Molina, R. et al. A highly responsive pyruvate sensor reveals pathway-regulatory role of the mitochondrial pyruvate carrier MPC. *eLife* **9**, e53917 (2020).
93. Schiefner, A. & Skerra, A. The menagerie of human lipocalins: a natural protein scaffold for molecular recognition of physiological compounds. *Acc. Chem. Res.* **48**, 976–985 (2015).
94. Lagerström, M. C. & Schiöth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* **7**, 339–357 (2008).
95. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y. & Remington, S. J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639–9647 (2006).
96. Muslinkina, L. et al. Two independent routes of post-translational chemistry in fluorescent protein FusionRed. *Int. J. Biol. Macromol.* **155**, 551–559 (2020).
97. Berardozzi, R., Adam, V., Martins, A. & Bourgeois, D. Arginine 66 controls dark-state formation in green-to-red photoconvertible fluorescent proteins. *J. Am. Chem. Soc.* **138**, 558–565 (2016).
98. Ding, J., Luo, A. F., Hu, L., Wang, D. & Shao, F. Structural basis of the ultrasensitive calcium indicator GCaMP6. *Sci. China Life Sci.* **57**, 269–274 (2014).

## Acknowledgements

We thank F. Subach, K. Horikawa, Y. Yang, T. Nagai, Y. Qin, Y. Li, F. Sun and H. Ai for providing gene sequences, and G. Baird for comments. R.E.C. acknowledges the Japan Society for the Promotion of Science (JSPS), Natural Sciences and Engineering Research Council of Canada (NSERC) and Canadian Institutes of Health Research (CIHR) for funding support.

## Author contributions

Y.N., Y.S. and R.E.C. wrote and edited the manuscript. Y.N., Y.S., L.K. and R.E.C. performed the literature survey that informed the conclusions and opinions expressed in this Perspective.

## Competing interests

R.E.C. is listed as an inventor on patents related to the development of various monomeric FPs. R.E.C. and Y.S. are listed as inventors on patent applications that describe single FP-based $Ca^{2+}$ biosensors.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41589-020-00718-x.

**Correspondence** should be addressed to R.E.C.

**Reprints and permissions information** is available at www.nature.com/reprints.

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