Circularly permuted fluorogenic proteins for the design of modular biosensors

Alison G Tebo, Frederico M Pimenta, Martha Zoumpoulaki, Carlos Kikuti, Helena Sirkia, Marie-Aude Plamont, Anne Houdusse, Arnaud Gautier

Circularly permuted fluorogenic proteins for the design of modular biosensors

Alison G. Tebo[1,#], Frederico M. Pimenta[1,#], Martha Zoumpoulaki[1], Carlos Kikuti[2], Helena Sirkia[2], Marie-Aude Plamont[1], Anne Houdusse[2] & Arnaud Gautier[1,*]

[1] PASTEUR, Département de Chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France.

[2] Structural Motility, Institut Curie, PSL Research University, CNRS, UMR 144, F-75005 Paris, France.

[#] Equal contributions.

* Correspondence should be addressed to: arnaud.gautier@ens.fr

ABSTRACT

Fluorescent reporters are essential components for the design of optical biosensors able to image intracellular analytes in living cells. Herein, we describe the development of circularly permuted variants of Fluorescence-Activating and absorption-Shifting Tag (FAST) and demonstrate their potential as reporting module in biosensors. Circularly permutated FAST (cpFAST) variants allow one to condition the binding and activation of a fluorogenic ligand (and thus fluorescence) to analyte recognition by coupling them with analyte-binding domains. We demonstrated their use for biosensor design by generating multicolor plug-and-play fluorogenic biosensors for imaging the intracellular levels of Ca^{2+} in living mammalian cells in real-time.

Cells respond to external and internal stimuli by a series of biochemical events that are tightly controlled in space and time. These downstream processes consist of interactions between macromolecules as well as between macromolecules and analytes. Detecting these interactions involves coupling a measurable output to a nanometer-scale molecular recognition event, typically by linking a reporting module to a sensing domain that responds to an input by undergoing a conformational change. Fluorescent proteins (FP) such as the green fluorescent protein (GFP) and its variants have been essential in the development of genetically encoded optical biosensors. FP-based biosensors typically function either by Förster resonance energy transfer (FRET), in which an analyte-binding domain is flanked with two FPs acting as donor and acceptor for FRET studies, or by fluorescence intensity change, in which conformational changes induced by analyte binding change the local environment of the chromophore of circularly permuted FPs.[1-4]

Chemical-genetic fluorogenic reporters provide an attractive alternative to FPs as they display many of the same advantages, while capitalizing on the specific interaction of a fluorogenic chromophore (so-called fluorogen) with its cognate genetically encoded receptor in order to generate fluorescence.[5-7] Fluorogens are chromophores that display no fluorescence until they interact with their cognate tag, resulting in systems with low background signal and little to no maturation time. The modular nature of these reporters opens exciting prospects for the design of genetically encoded biosensors in which fluorogen binding, and thus fluorescence, is conditioned to the recognition of a given analyte. An elegant illustration of this principle was obtained using Spinach, an RNA aptamer that

binds analogs of the GFP chromophore[8]. The coupling of metabolite-specific RNA aptamers to Spinach led to biosensors in which metabolite binding promote Spinach fluorescence[9,10].

To validate this approach with protein-based fluorogenic systems (**Figure 1A**), which have the advantages of being genetically targetable to any subcellular localization and in principle implementable in all model cells and organisms, we evaluated the use of the Fluorescence-Activating and absorption-Shifting Tag (FAST)[11], a variant of the 14 kDa photoactive yellow protein (PYP) evolved to bind and activate (through conformational locking) the fluorescence of various hydroxybenzylidene rhodanine (HBR) derivatives[11,12]. As fluorogen binding affinity depends on the protein structure, we anticipated that conformational change would affect fluorogen binding and thus fluorescence.

To maximize conformational coupling and facilitate insertion of FAST as reporting module into multidomain biosensors, we circularly permuted the protein. Circular permutation is a method used both in the laboratory and by nature whereby the amino acid sequence is altered such that the N and C termini are close in space but the protein retains the same (or similar) 3D structure (and hence function). Relocating the N- and C- termini within the structure of a protein to be spatially close has proved to be effective in increasing conformational sensitivity. This technique has been successfully used in the development of most intensiometric FP-based biosensors[13-15]. We hypothesized that FAST would tolerate circular permutation in much the same way as its parent protein, PYP.[16] Furthermore, we

reasoned that such a development could provide a strategy to create FAST-based biosensors by co-opting previously well-characterized sensor designs. For instance, calcium sensors composed of a circularly permuted FP and a calcium-binding motif comprised of calmodulin and M13 have been extensively explored.[13,14] In these sensors, the calcium-binding moieties calmodulin and M13 are appended to the N and C termini of the circularly permuted fluorescent protein and binding of Ca^{2+} results in the formation of the calmodulin and M13 ternary complex, which in turn induces a conformational change in FP, and results in higher fluorescence. This design, which has been implemented with *Aequorea* GFP or its variants, has not yet been successfully adapted to other fluorescent reporting modules. Herein, we describe the development of a series of circularly permuted FAST (cpFAST) variants with fluorogen affinities spanning an order of magnitude and illustrate using calcium sensors as a proof of principle that these cpFASTs can replace classic FPs for the design of intensiometric biosensors.

We designed seven circularly permuted FAST (cpFAST) variants by linking the original termini with peptide linkers of various lengths, and created new N- and C- termini at positions 115 and 114, respectively, as previously reported for circularly permuted PYP.[16] All cpFASTs showed good expression in *E. coli* (**Figure S1**) and formed green-yellow fluorescent complexes with 4-hydroxy-3-methylbenzylidene rhodanine (HMBR), indicating that the circular permutation did not perturb the three-dimensional topology and function of FAST. Absorption and emission wavelengths, fluorescence quantum yields, and absorption coefficients were nearly identical to those of non-permuted FAST (**Table 1**). Interestingly, we observed that the shorter the linker inserted within cpFAST, the lower

the affinity for HMBR (**Table 1** and **Figure S2**). This trend may result from shorter linkers introducing a physical constraint in the three-dimensional structure of the protein, thus inducing a lower binding affinity. All cpFAST were shown to activate HMBR fluorescence in live cells (**Figure S3**).

To validate the use of our cpFASTs for biosensor design, we created a sensor able to selectively detect Ca^{2+} by taking advantage of the Ca^{2+}-dependent interaction of calmodulin (CaM) and the Ca^{2+}-CaM interacting peptide M13, which has been successfully used in the past to generate genetically encoded FP-based calcium indicators (**Figure S4**)[13,14,17]. Ca^{2+} is a highly versatile secondary messenger that regulates many different cellular functions over a wide temporal range[18]. The concentration Ca^{2+} at resting state and the amplitude of the concentration change can vary significantly as a function of the intracellular localization and the cell type, therefore Ca^{2+} biosensors with various dissociation constants (ideally from nM to mM) are required to fully address the physiological role of Ca^{2+} signaling. Because of the dynamic and transient nature of calcium signaling, Ca^{2+} was an ideal model analyte for us to validate the use of cpFASTs for the design of dynamic fluorogenic biosensors. We connected the N-terminus of cpFASTs to the M13 peptide and their C-terminus to calmodulin (CaM). The corresponding sensors were expressed in *E. coli* with high yields (**Figure S5**), allowing full *in vitro* characterization.

For all sensors, the presence of Ca^{2+} increased HMBR binding affinity (**Figures 1B** and **S6** and **Table 2**), demonstrating that Ca^{2+} promoted fluorogen binding. We observed that

the shorter the linker used within the cpFAST, the larger the Ca^{2+}-induced increase in affinity. In agreement with the circular permutation providing conformational sensitivity, most cpFAST-based sensors displayed a larger Ca^{2+}-induced change in fluorogen affinity than a sensor in which the original FAST was inserted (**Table 2**).

An important feature of FAST is the ability to change its color from green-yellow to orange-red by using 4-hydroxy-3,5-dimethoxybenzylidene rhodanine (HBR-3,5-DOM) instead of HMBR[12]. We therefore tested HBR-3,5-DOM with the cpFAST-based sensors. Ca^{2+}-dependent increases in fluorogen affinity were also observed in that case (**Figures 1C** and **S7** and **Table 2**), demonstrating that the color of the sensors could be tuned at will from green-yellow to orange-red by choosing the appropriate fluorogen. This unprecedented feature provides an experimental versatility not encountered with FP-based biosensors, for which changing the color has required significant reengineering efforts[14]. With FAST-based sensors, sensor optimization is effectively decoupled from color choice, allowing plug-an-play color change.

In intensiometric fluorogen-based sensors, the dynamic range of the sensors is dependent on the conditions of the experiment, notably the fluorogen concentration used (See **SI Text 1** and **Figure S8**). To evaluate the dynamic range of our sensor, we measured the fluorescence fold increase for cpFAST2-based sensor in a range of fluorogen concentrations sampling a large range of conditions and expected fluorescence fold increases (**Figure 1D-G**). The observed fluorescence fold increase was slightly higher (~

1.2) than predicted based solely on the change in fraction of sensor:fluorogen complex, suggesting that the brightness of the sensor was also influenced by Ca^{2+}, which we attributed to an increased rigidity of the sensor in the presence of Ca^{2+}. Similar results were obtained with sensors based on other cpFASTs (**Figure S9**).

Next, we determined the Ca^{2+} affinity of our sensors. We fixed the concentration of fluorogen so that to maximize the change in fraction of sensor:fluorogen complex (see **SI Text 1**). The tested sensors displayed tight Ca^{2+} binding with dissociation constants ($K_{D,Ca}$) of about 60 to 100 nM, in good agreement with that of the CaM and M13 fusion protein linked with two amino acid linkers previously reported ($K_{D,Ca}$ of 80 nM)[19,20], and Hill coefficients (n_H) of about 3 in agreement with cooperative Ca^{2+} binding to CaM (**Figures 1H,I, S10-S13** and **Table 2**). Neither the fluorogen used (HMBR or HBR-3,5-DOM) nor the cpFAST chosen had a significant influence on the affinity for Ca^{2+}. Furthermore, we examined the dependence of $K_{D,Ca}$ on fluorogen concentration for the sensor based on cpFAST2. We found that Ca^{2+} affinity is only slightly dependent on the fluorogen concentration (**Figure S10** and **S11**), indicating that experimental conditions may be tuned at will without affecting the basics of sensor function. The absence of strong dependence of the Ca^{2+} affinity on the fluorogen concentration in this case is linked to the cooperative binding of Ca^{2+}. With their high Ca^{2+} affinity, our cpFAST-based biosensors can allow the detection of subtle changes in Ca^{2+} concentration and complement the current suite of widely used single-FP-based Ca^{2+} indicators[21].

We next evaluated the function of our cpFAST-based Ca^{2+} sensors in living cells. When the sensors were expressed in HeLa cells and labeled with HMBR or HBR-3,5-DOM, the fluorescence was uniformly distributed in cells (**Figure 2**). This observation was in agreement with the sensors being small (33 kDa), monomeric proteins as confirmed by size-exclusion chromatography coupled to multiangle light scattering (**Figure S14**). Treatment with histamine, which triggers Ca^{2+} release from the endoplasmic reticulum to the cytoplasm, induced a significant increase of fluorescence (2 to 3-fold), in accordance with an increase of cytosolic free Ca^{2+} (**Figure 2**, and **movies S1-S3**). This initial peak was followed by oscillations and eventually desensitization (**Figure 2** and **movies S1-S3**). Our sensors did not display the striking oscillations reported with recently, highly-optimized sensors very likely because of their high affinity for Ca^{2+}. However, the observed response was in good agreement with the known behavior of HeLa cells to histamine stimulation[13,14,20,21], demonstrating our ability to image Ca^{2+} levels in living cells in real-time.

In this study, we report the design of functional circularly permuted FAST variants (cpFAST). We demonstrate using Ca^{2+} sensors as a proof-of-principle that cpFASTs can be an alternative to FPs as reporting module in biosensors. In cpFAST-based sensors, the intensiometric response results from conditioning fluorogen binding (and thus fluorescence) to the recognition of an intracellular analyte. We could generate functional Ca^{2+} biosensors with minimal engineering efforts. The successful design of FP-based sensors of the GCaMP family – made of circular permuted GFP flanked with CaM and M13 – results from the serendipitous specific interaction between CaM and GFP in presence of Ca^{2+} that strongly affects GFP conformation and photophysical properties[22]. Engineering

efforts to implement this design to generate genetically encoded calcium indicators with fluorescent reporter modules other than FPs have not been successful so far.[23] In addition, extension of this serendipity-based design principle to analytes other than Ca^{2+} has turned out to be highly challenging.[15,24-27] FAST-based Ca^{2+} sensors display Ca^{2+}-dependent fluorescence without the need for a specific direct interaction between FAST and the Ca^{2+} sensing domains, which should thus facilitate the extension of this design principle to other analytes using various sensing modules. The ability to change the color by changing the fluorogen used allows the design of biosensors with various spectral properties without the need for reengineering, providing an unprecedented experimental versatility for multicolor imaging.

Although conformationally coupling of FAST and the sensing unit allows one to condition fluorogen binding to analyte recognition, it also makes the dynamic range of the sensor and the affinity for the analyte dependent on the fluorogen concentration. Consequently, the fluorogen concentration is a control parameter that must be chosen with care to optimize measurements with the sensors. Despite this, the use of FAST for sensor design opens exciting perspectives. Its robustness and tolerance to insertions make FAST a particularly interesting scaffold for sensor design at once taking advantage of its relatively small size and the ability to change its color at-will. The use of cpFASTs could prove useful in the screening of new sensing modules for which current development is stifled by intrinsic limitations of FPs such as maturation time, tendency to oligomerize, and oxygen sensitivity. The simplicity and tractability of such an approach should allow the design of new plug-and-play biosensors for imaging in real-time analytes, endogenous biomolecules and cellular processes for which no systems currently exist.

SUPPORTING INFORMATION

The supporting information contains Movies S1-S3, SI Text 1, Figures S1-S14 and Materials and Methods.

ACKNOWLEDGEMENTS

This work was supported by the European Research Council (ERC-2016-CoG-724705 FLUOSWITCH), the Agence National de la Recherche (ANR-14-CE09-0002- 01), France BioImaging (ANR-10-INBS-04), the Equipex Morphoscope 2 (ANR-11-EQPX-0029) and the program «Investissements d'Avenir » launched by the French Government and implemented by ANR (ANR-10-IDEX-0001-02 PSL, project Superline).

NOTES

The authors declare the following competing financial interest: A.G. is co-founder and holds equity in Twinkle Bioscience, a company commercializing the FAST technology.

Table 1. **Physico-chemical properties of the circularly permuted FAST (cpFAST) variants bound to HMBR.** Abbreviations are as follows: λ_{abs}, wavelength of maximal absorption; λ_{em}, wavelength of maximal emission; ε, molar absorption coefficient at λ_{abs}; ϕ, fluorescence quantum yield; K_D, thermodynamic dissociation constant.

Variant	Linker size (a.a.)	λ_{abs} (nm)	λ_{em} (nm)	ε ($mM^{-1}cm^{-1}$)	ϕ	K_D (μM)
FAST*	-	481	540	42.5 ± 0.8	0.23 ± 0.03	0.13 ± 0.01
cpFAST1	0	481	540	~ 35 **	~ 0.2**	5.1 ± 0.70
cpFAST2	3	481	540	~ 35 **	~ 0.2**	0.99 ± 0.11
cpFAST3	5	481	540	~ 35 **	~ 0.2**	0.65 ± 0.04
cpFAST4	8	481	540	36.0 ± 0.5	0.19 ± 0.02	0.30 ± 0.04
cpFAST5	11	481	542	36.2 ± 0.9	0.28 ± 0.03	0.22 ± 0.03
cpFAST6	15	481	541	36.9 ± 0.8	0.24 ± 0.03	0.26 ± 0.03
cpFAST7	18	481	540	33.0 ± 2.9	0.20 ± 0.02	0.16 ± 0.01

* previously reported[12].

** Molar absorption coefficients and quantum yields for cpFAST1, cpFAST2 and cpFAST3 are approximate numbers due to restrictions imposed on the maximum amount of complex that can be formed in solution given the reported K_D (i.e., only at appreciably high protein concentrations will ~100% of complex be formed).

Table 2. Physico-chemical properties of the cpFAST-based Ca^{2+} biosensors (M13-Reporter-CaM) labeled with HMBR (A) and HBR-3,5-DOM (B). Abbreviations are as follows: $K_{D,+}$ and $K_{D,-}$ are the thermodynamic dissociation constants of the sensor:fluorogen complexes in presence and in absence of Ca^{2+}; $K_{D,Ca}$ and n_H are the thermodynamic dissociation constant and Hill coefficient associated to Ca^{2+} binding (the fluorogen concentration was fixed to $(K_{D,+}K_{D,-})^{1/2}$).

Reporter		$K_{D,+}$ (μM)	$K_{D,-}$ (μM)	$K_{D,Ca}$ (nM) [n_H]
FAST	A	0.16 ± 0.01	0.36 ± 0.02	
cpFAST1	A	4.0 ± 0.4	18.2 ± 0.8	77 ± 2 [2.8 ± 0.2]
	B	17.4 ± 0.8	37 ±3	63 ± 1 [4.0 ± 0.5]
cpFAST2	A	1.40 ± 0.04	5.2 ± 0.1	99 ± 4 [3.0 ± 0.3]
	B	3.9 ± 0.1	18.0 ± 0.8	82 ± 1 [3.3 ± 0.2]
cpFAST3	A	0.80 ± 0.02	2.9 ± 0.1	64 ± 2 [3.3 ± 0.3]
	B	4.2 ± 0.2	14.8 ± 0.9	57 ± 3 [2.3 ± 0.3]
cpFAST4	A	0.59 ± 0.03	1.39 ± 0.03	
	B	2.8 ± 0.1	9.1 ± 0.4	
cpFAST5	A	0.36 ± 0.01	0.82 ± 0.02	
	B	2.50 ± 0.09	7.5 ± 0.6	

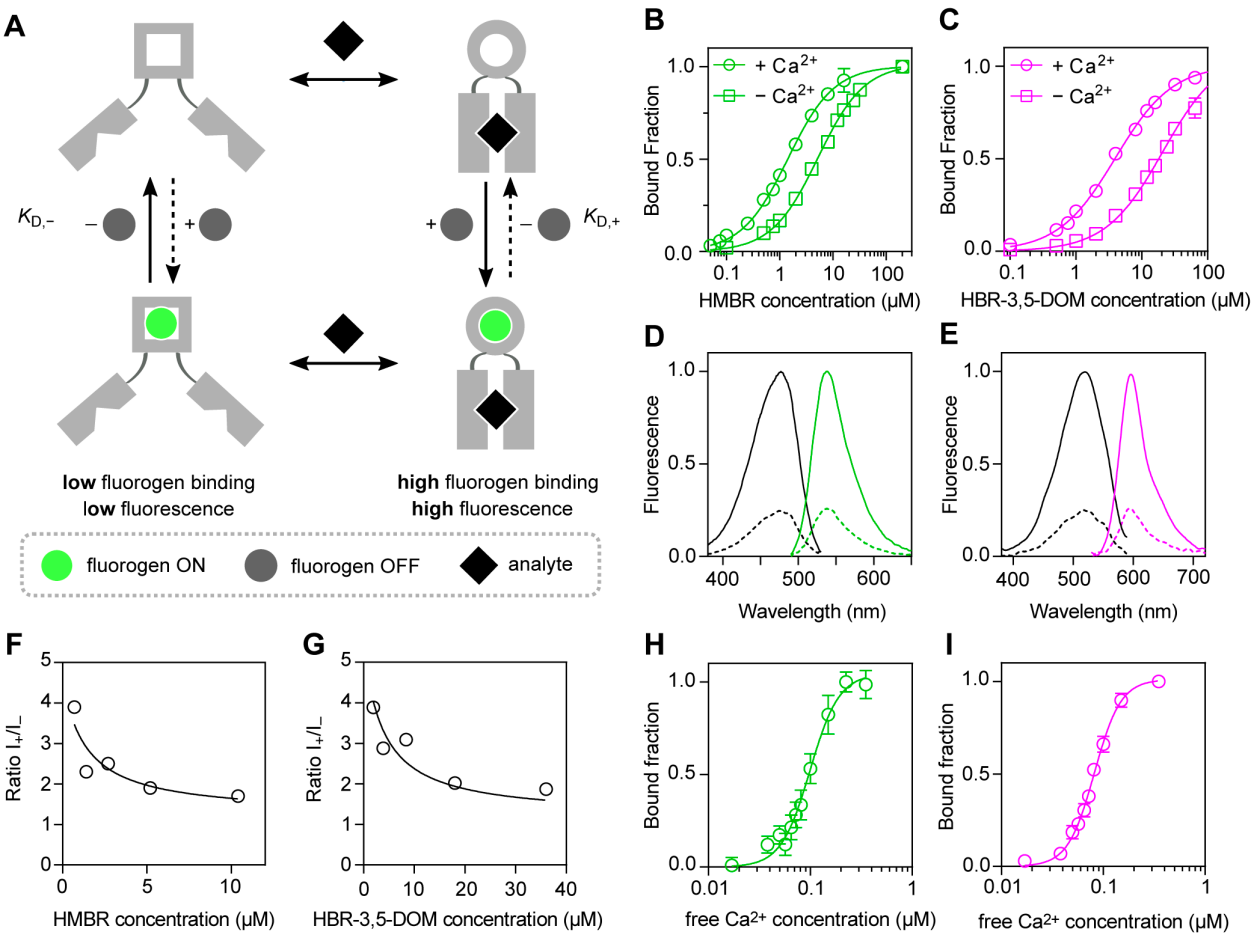

Figure 1. cpFAST-based biosensors. (A) Fluorogenic biosensors composed of a cpFAST variant coupled to analyte-recognition modules in which the binding of the fluorogenic ligand, and thus the fluorescence, is conditioned to the binding of the analyte. $K_{D,-}$ and $K_{D,+}$ are the thermodynamic dissociation constants of the sensor:fluorogen complex in the absence and the presence of analyte, respectively. **(B,C)** Titration curves of M13-cpFAST2-CaM in the absence or in presence of Ca^{2+} with either HMBR **(B)** or HBR-3,5DOM **(C).** Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constants $K_{D,-}$ and $K_{D,+}$ provided in **Table 2**. **(D,E)** Excitation and emission spectra of M13-cpFAST2-CaM:HMBR **(D)** and M13-cpFAST2-CaM:HBR-3,5-DOM **(E)** in the absence (dotted line) and presence (solid line) of Ca^{2+}. The concentration of HMBR was 0.7 μM and that of HBR-3,5-DOM 1.95 μM. Sensor concentration was fixed at 0.1 μM. **(F,G)** Dependence of the I_+/I_- emission ratio on concentration of **(F)** HMBR and **(G)** HBR-3,5-DOM. **(F)** I_+ and I_- were obtained by integrating spectra recorded at a sensor concentration of 100 nM and HMBR concentrations of 0.7 μM, 1.4 μM**,** 2.7 μM**,** 5.2 μM, and 10.4 μM at 25 °C. **(G)** I_+ and I_- were obtained by integrating spectra recorded at a sensor concentration of 100 nM and HBR-3,5-DOM concentrations of 1.95 μM, 3.9 μM, 8.4 μM, 18 μM and 36 μM at 25 °C. **(H,I)** Ca^{2+} titration curve of M13-cpFAST2-CaM in the presence of 2.7 μM HMBR **(H)** and 8.4 μM HBR-3,5DOM **(I)**. Data represent mean ± sem (n =3). Least square fit (line) gave the thermodynamic dissociation constant $K_{D,Ca}$ and Hill coefficient n_H provided in **Table 2**. Sensor concentration was fixed at 0.01 μM.

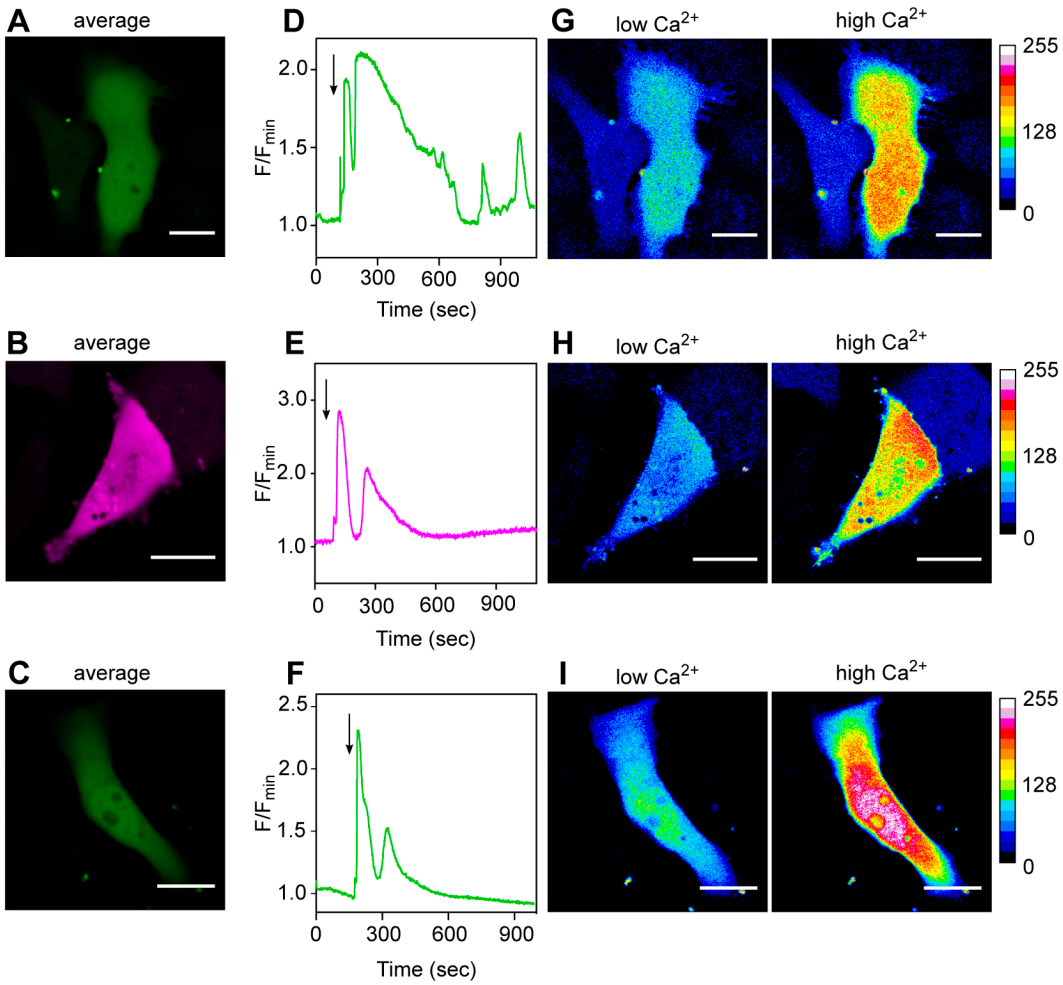

Figure 2. Ca^{2+} imaging in HeLa cells. M13-cpFAST2-CaM was expressed in HeLa cells and labeled with 5 µM HMBR (A,D,G) or 10 µM HBR-3,5DOM (B,E,H). M13-cpFAST3-CaM was expressed in HeLa cells and labeled with 5 µM HMBR (C,F,I). Cells were stimulated with 50 µM histamine. Movies S1, S2 and S3 were recorded by confocal microscopy. (A-C) Average fluorescence images showing homogenous distribution of the sensors. (D-F) Fluorescence intensity changes in response to histamine addition. Time of histamine addition is indicated by the black arrow. (G-I) Low fluorescence was observed in resting state (low Ca^{2+}). Histamine addition (50 µM) induced a significant increase of fluorescence (high Ca^{2+}), in accordance with an increase of cytosolic free Ca^{2+}. (A-I) Scale bars 20 µm.

REFERENCES

(1) Okumoto, S. (2010) Imaging approach for monitoring cellular metabolites and ions using genetically encoded biosensors. *Curr. Opin. Biotechnol. 21*, 45–54.
(2) Palmer, A. E., Qin, Y., Park, J. G., and McCombs, J. E. (2011) Design and application of genetically encoded biosensors. *Trends in Biotechnology 29*, 144–152.
(3) Frommer, W. B., Davidson, M. W., and Campbell, R. E. (2009) Genetically encoded biosensors based on engineered fluorescent proteins. *Chem Soc Rev 38*, 2833–18.
(4) VanEngelenburg, S. B., and Palmer, A. E. (2008) Fluorescent biosensors of protein function. *Curr. Op. Chem. Biol. 12*, 60–65.
(5) Bruchez, M. P. (2015) Dark dyes–bright complexes: fluorogenic protein labeling. *Curr. Op. Chem. Biol. 27*, 18–23.
(6) Jullien, L., and Gautier, A. (2015) Fluorogen-based reporters for fluorescence imaging: a review. *Methods Appl. Fluoresc. 3*, 042007–13.
(7) Li, C., Tebo, A. G., and Gautier, A. (2017) Fluorogenic Labeling Strategies for Biological Imaging. *IJMS 18*, 1473–11.
(8) Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA Mimics of Green Fluorescent Protein. *Science 333*, 642–646.
(9) Paige, J. S., Nguyen-Duc, T., Song, W., and Jaffrey, S. R. (2012) Fluorescence imaging of cellular metabolites with RNA. *Science 335*, 1194–1194.
(10) You, M., Litke, J. L., and Jaffrey, S. R. (2015) Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proc. Natl. Acad. Sci. U. S. A 112*, E2756–65.
(11) Plamont, M.-A., Billon-Denis, E., Maurin, S., Gauron, C., Pimenta, F. M., Specht, C. G., Shi, J., Querard, J., Pan, B., Rossignol, J., Morellet, N., Volovitch, M., Lescop, E., Chen, Y., Triller, A., Vriz, S., Le Saux, T., Jullien, L., and Gautier, A. (2016) Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. *Proc. Natl. Acad. Sci. U. S. A 113*, 497–502.
(12) Li, C., Plamont, M.-A., Sladitschek, H. L., Rodrigues, V., Aujard, I., Neveu, P., Le Saux, T., Jullien, L., and Gautier, A. (2017) Dynamic multicolor protein labeling in living cells. *Chem. Sci. 8*, 5598–5605.
(13) Nakai, J., Ohkura, M., and Imoto, K. (2001) A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. *Nat Biotechnol 19*, 137–141.
(14) Zhao, Y., Araki, S., Wu, J., Teramoto, T., Chang, Y.-F., Nakano, M., Abdelfattah, A. S., Fujiwara, M., Ishihara, T., Nagai, T., and Campbell, R. E. (2011) An expanded palette of genetically encoded Ca^{2+} indicators. *Science 333*, 1888–1891.
(15) Marvin, J. S., Borghuis, B. G., Tian, L., Cichon, J., Harnett, M. T., Akerboom, J., Gordus, A., Renninger, S. L., Chen, T.-W., Bargmann, C. I., Orger, M. B., Schreiter, E. R., Demb, J. B., Gan, W.-B., Hires, S. A., and Looger, L. L. (2013) An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat. Methods 10*, 162–170.
(16) Kumar, A., Burns, D. C., Al-Abdul-Wahid, M. S., and Woolley, G. A. (2013) A circularly permuted photoactive yellow protein as a scaffold for photoswitch design. *Biochemistry 52*, 3320–3331.
(17) Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., and Kim, D. S. (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity.

Nature 499, 295–300.
(18) Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003) Calcium signalling: dynamics, homeostasis and remodelling. *Nat Rev Mol Cell Biol 4*, 517–529.
(19) Porumb, T., Yau, P., Harvey, T. S., and Ikura, M. (1994) A calmodulin-target peptide hybrid molecule with unique calcium-binding properties. *Protein Eng. 7*, 109–115.
(20) Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., and Tsien, R. Y. (1997) Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature 388*, 882–887.
(21) Bootman, M. D., Cheek, T. R., Moreton, R. B., Bennett, D. L., and Berridge, M. J. (1994) Smoothly graded Ca2+ release from inositol 1,4,5-trisphosphate-sensitive Ca2+ stores. *J. Biol. Chem. 269*, 24783–24791.
(22) Akerboom, J., Rivera, J. D. V., Guilbe, M. M. R., Malavé, E. C. A., Hernandez, H. H., Tian, L., Hires, S. A., Marvin, J. S., Looger, L. L., and Schreiter, E. R. (2009) Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem. 284*, 6455–6464.
(23) Shitashima, Y., Shimozawa, T., Asahi, T., and Miyawaki, A. (2018) A dual-ligand-modulable fluorescent protein based on UnaG and calmodulin. *Biochem. Biophys. Res. Comm. 496*, 872–879.
(24) Wu, J., Abdelfattah, A. S., Zhou, H., Ruangkittisakul, A., Qian, Y., Ballanyi, K., and Campbell, R. E. (2018) Genetically Encoded Glutamate Indicators with Altered Color and Topology. *ACS Chem. Biol.*
(25) Qin, Y., Sammond, D. W., Braselmann, E., Carpenter, M. C., and Palmer, A. E. (2016) Development of an Optical Zn 2+Probe Based on a Single Fluorescent Protein. *ACS Chem. Biol. 11*, 2744–2751.
(26) Chen, Z., and Ai, H.-W. (2016) Single Fluorescent Protein-Based Indicators for Zinc Ion (Zn 2+). *Anal. Chem. 88*, 9029–9036.
(27) Sanford, L., and Palmer, A. (2017) Recent Advances in Development of Genetically Encoded Fluorescent Sensors. *Meth. Enzymol. 589*, 1–49.

Supporting Information

Circularly permuted fluorogenic proteins for the design of modular biosensors

Alison G. Tebo[1,#], Frederico M. Pimenta[1,#], Martha Zoumpoulaki[1], Carlos Kikuti[2], Helena Sirkia[2], Marie-Aude Plamont[1], Anne Houdusse[2] & Arnaud Gautier[1,*]

[1] PASTEUR, Département de Chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France.

[2] Structural Motility, Institut Curie, PSL Research University, CNRS, UMR 144, F-75005 Paris, France.

[#] Equal contributions

* Correspondence should be addressed to: arnaud.gautier@ens.fr

Content

Legends of Movies

Movie S1: Ca^{2+} imaging in HeLa cells using M13-cpFAST2-CaM and HMBR. M13-cpFAST2-CaM was expressed in HeLa cells and labeled with 5 μM HMBR. Cells were stimulated with 50 μM of histamine at t = 100 s.

Movie S2: Ca^{2+} imaging in HeLa cells using M13-cpFAST2-CaM and HBR-3,5-DOM. M13-cpFAST2-CaM was expressed in HeLa cells and labeled with 10 μM HBR-3,5-DOM. Cells were stimulated with 50 μM of histamine at t = 80 s.

Movie S3: Ca^{2+} imaging in HeLa cells using M13-cpFAST3-CaM and HMBR. M13-cpFAST3-CaM was expressed in HeLa cells and labeled with 5 μM HMBR. Cells were stimulated with 50 μM of histamine at t = 160 s.

SI Text 1

When the total concentration of fluorogen (F_{tot}) is large compared to the total concentration of sensor (S_{tot}), one can write the fraction of sensor:fluorogen complex (α) as:

$$\alpha = \frac{1}{\frac{K_D}{F_{tot}} + 1}$$

, where K_D is the thermodynamic dissociation constant of the sensor:fluorogen complex.

To characterize the response of a fluorogen-based sensor upon analyte binding, one can either compute the variation of complex fraction upon analyte binding:

$$\alpha_+ - \alpha_- = \frac{1}{\frac{K_{D,+}}{F_{tot}} + 1} - \frac{1}{\frac{K_{D,-}}{F_{tot}} + 1}$$

or the increase of bound fraction upon analyte binding:

$$\alpha_+ / \alpha_- = \frac{\frac{K_{D,-}}{F_{tot}} + 1}{\frac{K_{D,+}}{F_{tot}} + 1}$$

, where α_+ and α_- are respectively the fraction of sensor:fluorogen complex in excess and in absence of analyte, and $K_{D,+}$ and $K_{D,-}$ are respectively the thermodynamic dissociation constants of the sensor:fluorogen complex in excess and in absence of analyte.

One can easily show that $\alpha_+ - \alpha_-$ is optimal when $F_{tot} = (K_{D,+} K_{D,-})^{1/2}$, while α_+ / α_- is maximal at small F_{tot}.

By noting $I_{F,+}$ and $I_{F,-}$ the fluorescence intensities in presence and in absence of the analyte, the fluorescence dynamic range of the sensor is given by:

$$\frac{I_{F,+}}{I_{F,-}} = \frac{B_+ \alpha_+}{B_- \alpha_-}$$

where B_+ and B_- are respectively the brightness of the sensor:fluorogen complex in presence and in absence of analyte. The dynamic range depends thus on the concentration in fluorogen F_{tot} and is maximal at low F_{tot}, approaching the value

$$\frac{B_+}{B_-}\frac{K_{D,-}}{K_{D,+}}$$

However, F_{tot} must be high enough to allow the detection of the fluorescent complex, (See Figure S8). We propose that the ideal compromise to obtain both good dynamic range and satisfactory detection sensitivity is to choose F_{tot} in the range:

$$K_{D,+}/2 \;\leq\; F_{tot} \;\leq\; (K_{D,+}K_{D,-})^{1/2}$$

The fluorogen concentration may be eventually tuned relative to the needs of the experiment — lower, when the signal ratio must be optimized, or higher when overall fluorescence signal is important.

To evaluate the dynamic range of our sensor, we measured the fluorescence fold increase in a range of fluorogen concentrations – $K_{D,+}/2$, $K_{D,+}$, $(K_{D,+}K_{D,-})^{1/2}$, $K_{D,-}$, and $2K_{D,-}$ – chosen so as to sample a large range of conditions and expected fluorescence fold increases

SI Figures

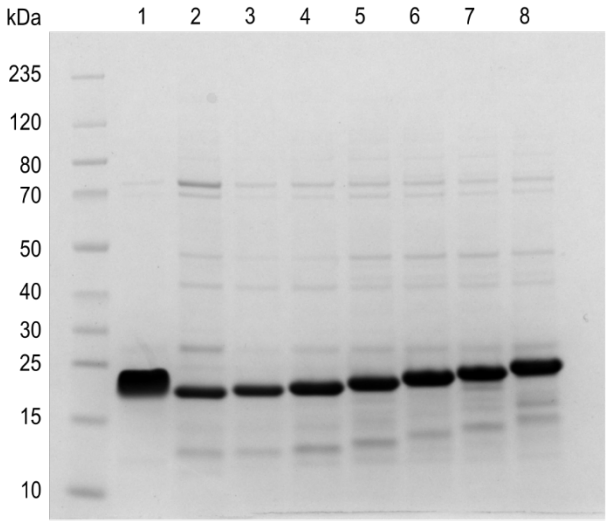

Figure S1. **Bacterial expression and purification of the circularly permuted FAST variants.** Coomassie-stained SDS-PAGE of purified 1: FAST, 2: cpFAST1, 3: cpFAST2, 4: cpFAST3, 5: cpFAST4, 6: cpFAST5, 7: cpFAST6, 8: cpFAST7.

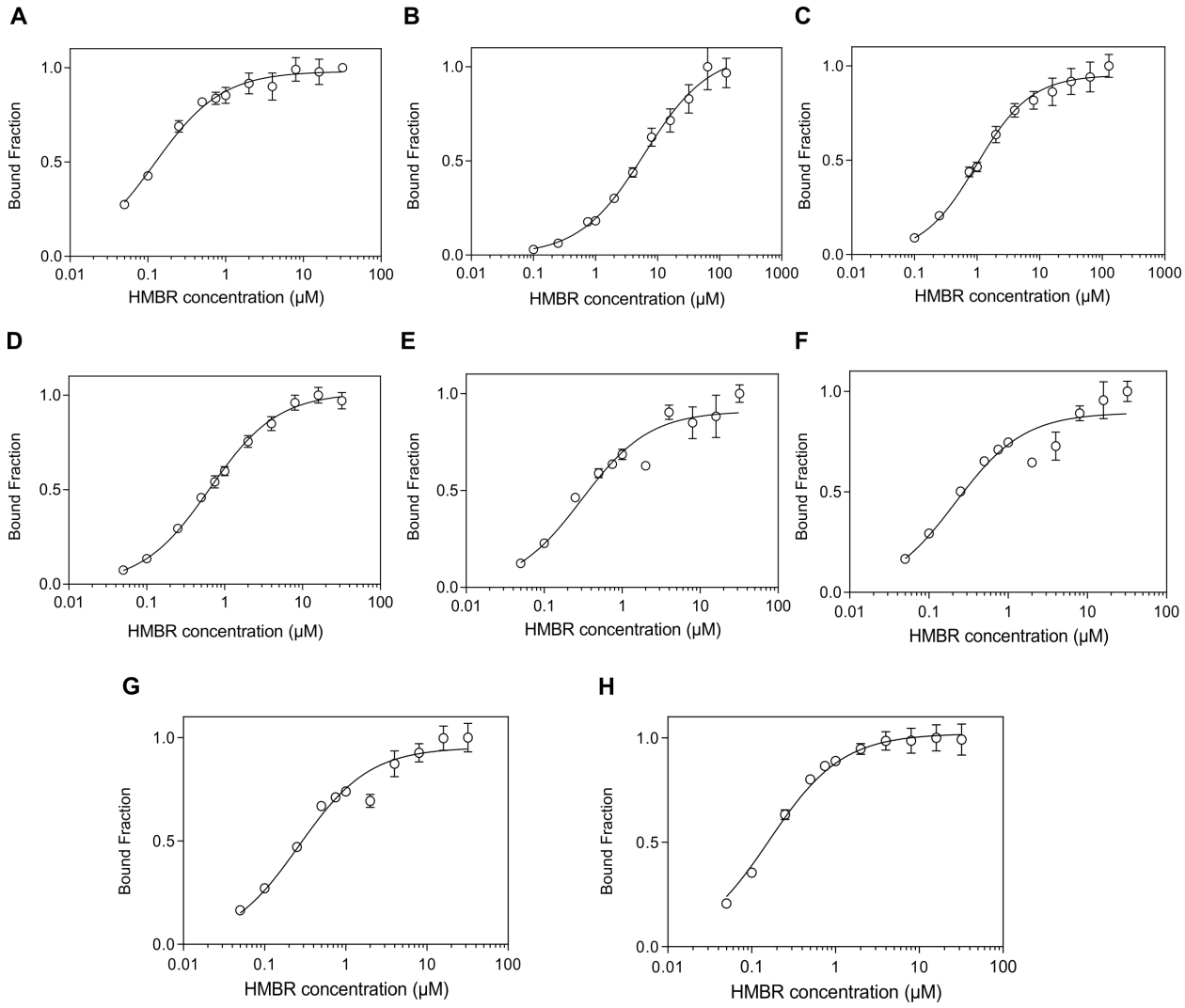

Figure S2. **HMBR titration curves of (A) FAST, (B) cpFAST1, (C) cpFAST2, (D) cpFAST3, (E) cpFAST4, (F) cpFAST5, (G) cpFAST6, (H) cpFAST7.** Titrations were done at a protein concentration of 50 nM in pH 7.4 PBS buffer at 25 °C. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constants K_D provided in **Table 1**.

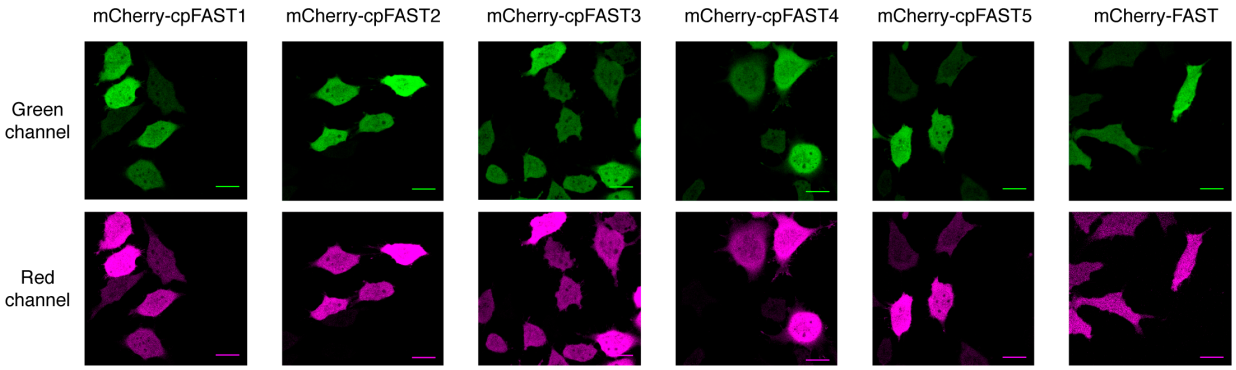

Figure S3. Characterization of cpFAST in mammalian cells. Confocal images of HeLa cells expressing the circularly permuted FAST variants fused to the red fluorescent protein mCherry. The green channel shows the fluorescence signal from the FAST variants labeled with concentrations of HMBR ensuring full labeling of FAST (50 μM for cpFAST1, 40 μM for cpFAST2, 15 μM for cpFAST3, 10 μM for cpFAST4 and 6 μM for cpFAST5, and 5 μM for FAST5), while the red channel shows the fluorescence signal from mCherry. Side-by-side images were recorded with the same settings for direct comparison. Scale bars 20 μm.

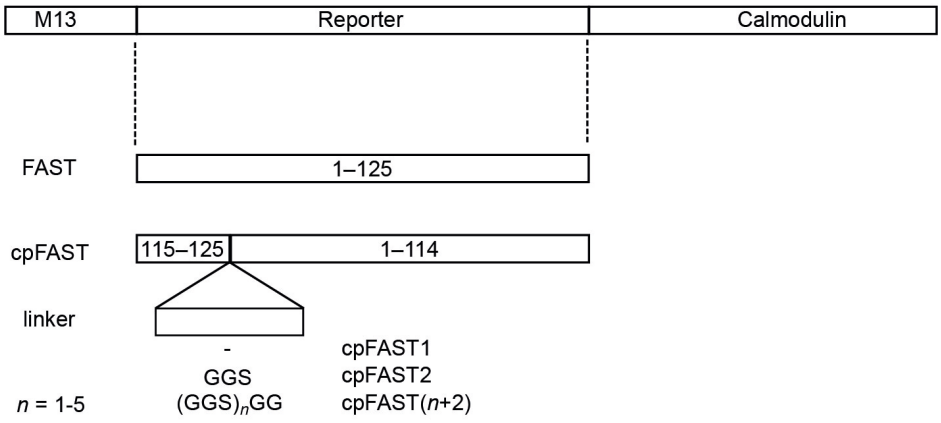

Figure S4. **FAST-based Ca^{2+} biosensors. FAST-based Ca^{2+} biosensors** were constructed by connecting the N-terminus of FAST or cpFAST to the M13 peptide, and its C-terminus to Calmodulin.

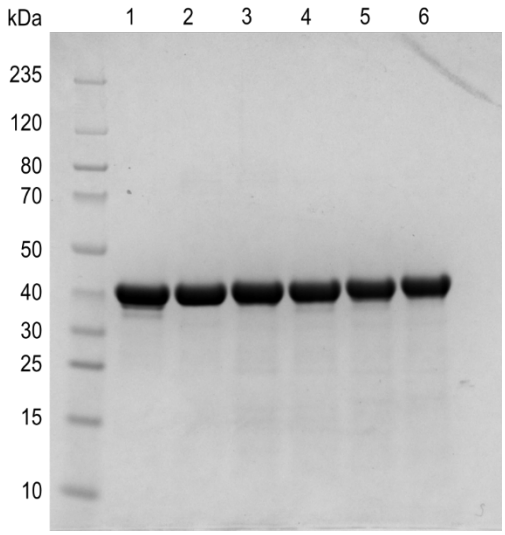

Figure S5. **Bacterial expression and purification of the FAST-based Ca^{2+} biosensors.** Coomassie-stained SDS-PAGE of purified 1: M13-FAST-CaM, 2: M13-cpFAST1-CaM, 3: M13-cpFAST2-CaM, 4: M13-cpFAST3-CaM, 5: M13-cpFAST4-CaM, 6: M13-cpFAST5-CaM.

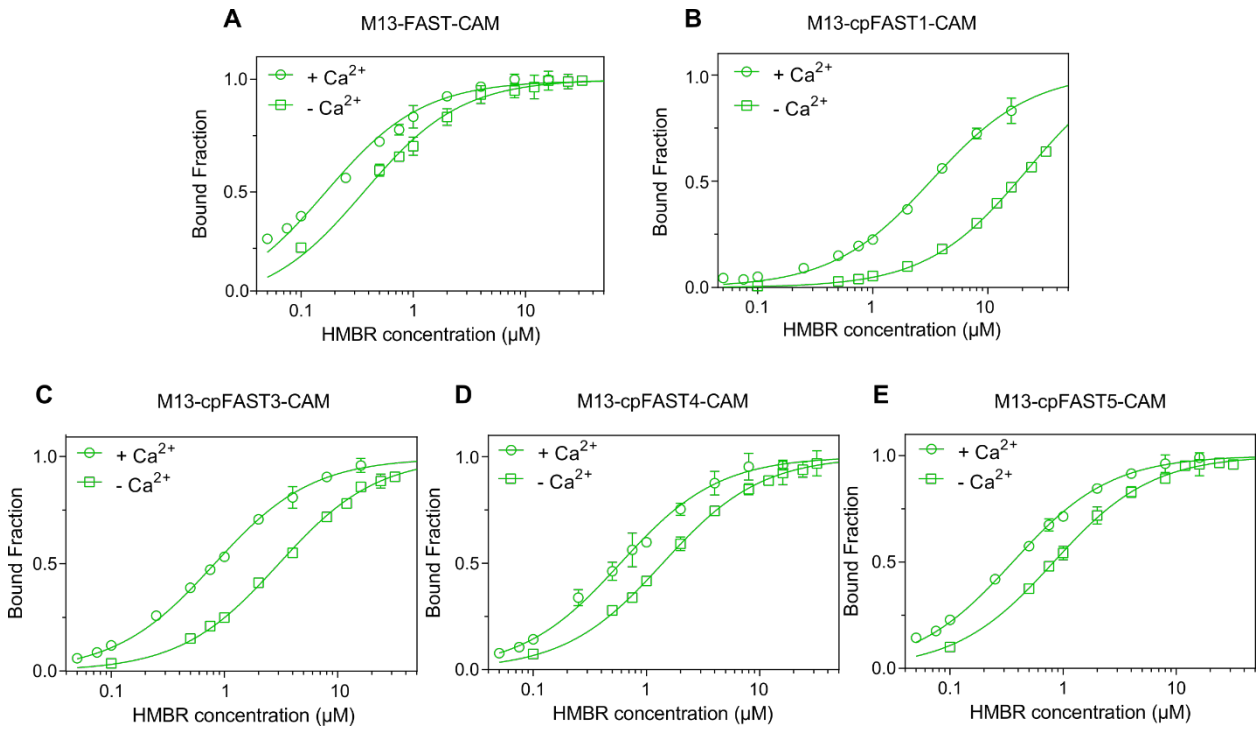

Figure S6. HMBR titration curves of the FAST-based Ca^{2+} biosensors in the presence and absence of Ca^{2+}. (A) M13-FAST-CaM, **(B)** M13-cpFAST1-CaM, **(C)** M13-cpFAST3-CaM, **(D)** M13-cpFAST4-CaM, **(E)** M13-cpFAST5-CaM. Titrations were done at a sensor concentration of 100 nM in pH 7.4 HEPES buffer containing 10 mM $CaCl_2$ (+ Ca^{2+}) or 10 mM EGTA (– Ca^{2+}) at 25 °C. Excitation was fixed at 480 nm and emission at 540 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constants $K_{D,-}$ and $K_{D,+}$ provided in **Table 2**.

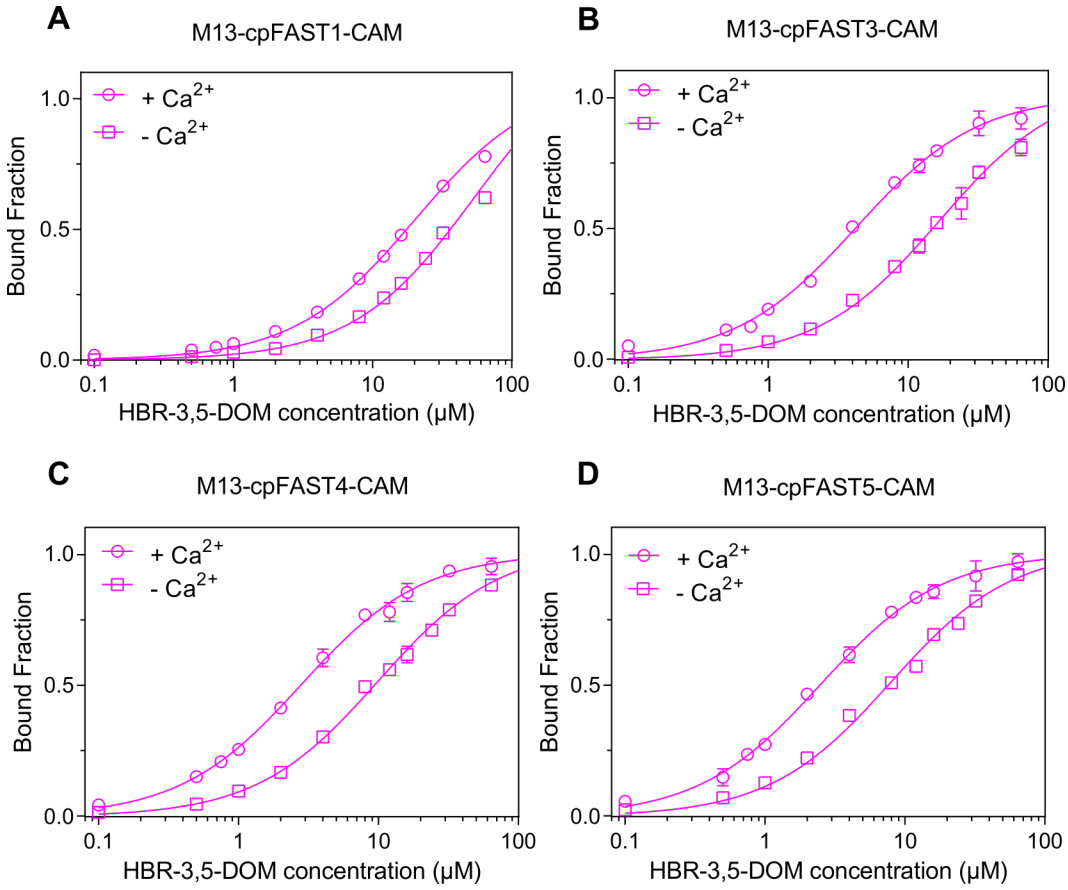

Figure S7. HBR-3,5-DOM titration curves of the FAST-based Ca^{2+} biosensors in the presence and absence of Ca^{2+}. (A) M13-cpFAST1-CaM, **(B)** M13-cpFAST3-CaM, **(C)** M13-cpFAST4-CaM, **(D)** M13-cpFAST5-CaM. Titrations were done at a sensor concentration of 100 nM in pH 7.4 HEPES buffer containing 10 mM $CaCl_2$ (+ Ca^{2+}) or 10 mM EGTA (– Ca^{2+}) at 25 °C. Excitation was fixed at 520 nm and emission at 600 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constants $K_{D,-}$ and $K_{D,+}$ provided in **Table 2**.

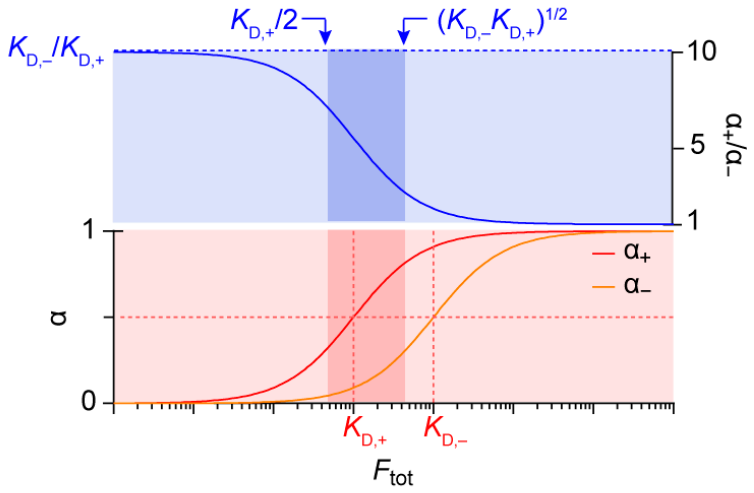

Figure S8. FAST-based optical biosensors. Simulated fluorogen binding curves in the presence and absence of analyte (with $K_{D,-}$ = 10 $K_{D,+}$). α_+ and α_- are the fluorogen bound fractions in presence and absence of the analyte, respectively. Choosing the concentration of fluorogen between $K_{D,+}/2$ and $(K_{D,+}K_{D,-})^{1/2}$ (contrasted rectangles) is a good compromise to obtain both good dynamic range and satisfactory detection sensitivity.

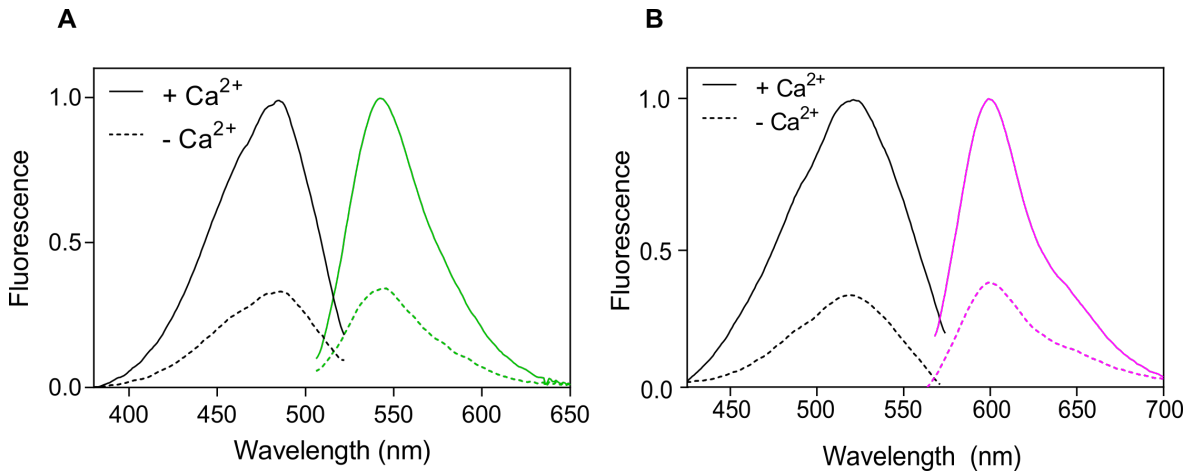

Figure S9. Excitation and Emission spectra of (A) M13-cpFAST3-CaM:HMBR and (B) M13-cpFAST3-CaM:HBR-3,5DOM in the absence (dashed lines) and presence of Ca^{2+} (solid lines). Spectra were recorded at a sensor concentration of 100 nM and a fluorogen concentration of $(K_{D,+}K_{D,-})^{1/2}$ at 25 °C. Excitation was scanned from 380 to 530 nm at 540 nm emission andemission was scanned from 490 to 650 nm at 480 nm excitation.

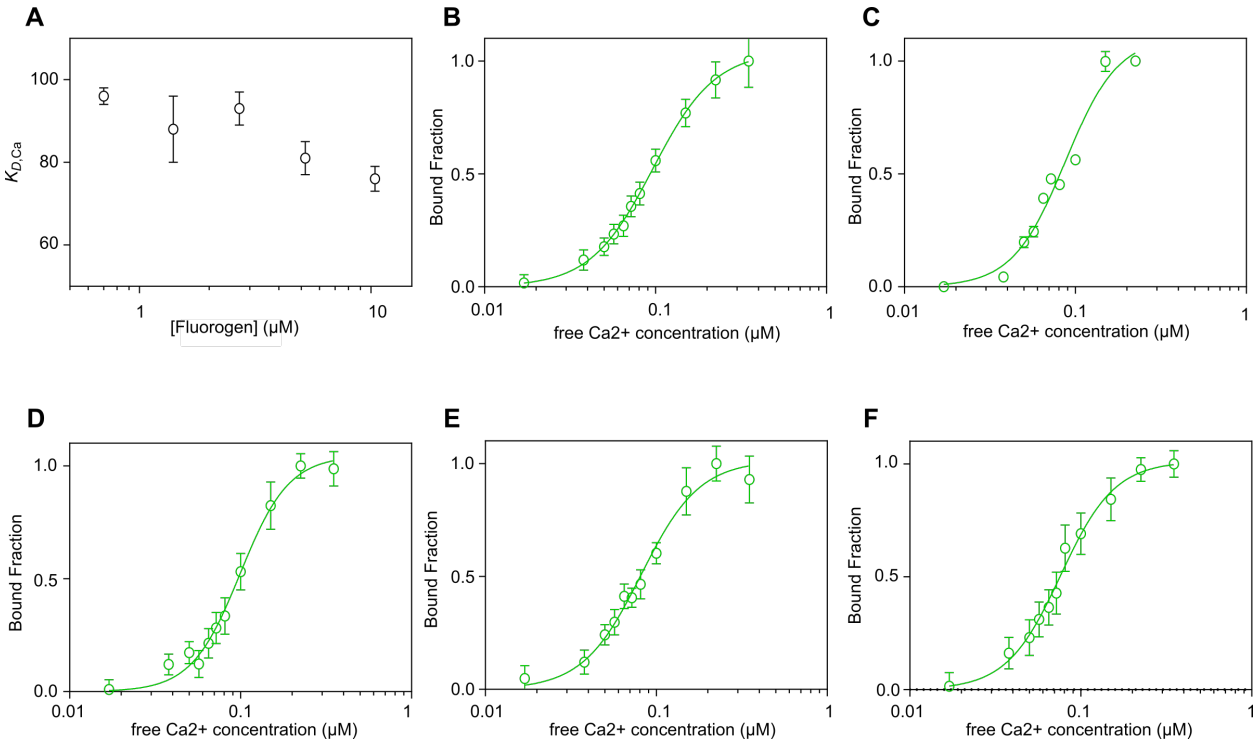

Figure S10. (A) Dependence of $K_{D,Ca}$ on HMBR concentration for M13-cpFAST2-CaM **(B-F).**Titrations were done at a sensor concentration of 10 nM and HMBR concentrations of **(B)** 0.7 µM **(C)** 1.4 µM **(D)** 2.7 µM, **(E)** 5.4 µM and **(F)** 10.4 µM. Excitation was fixed at 480 nm and emission at 540 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constant $K_{D,Ca}$ and Hill coefficient n_H.

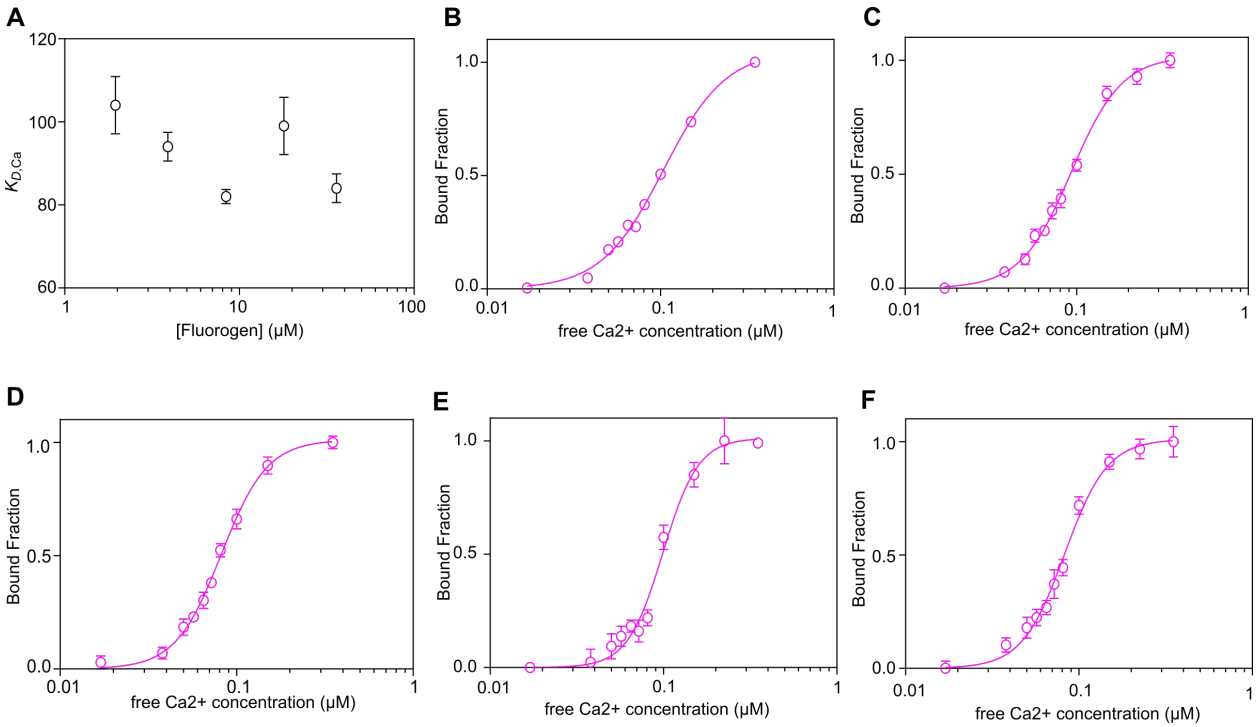

Figure S11. (A) Dependence of $K_{D,Ca}$ on HBR-3,5DOM concentration for M13-cpFAST2-CaM **(B-F).**Titrations were done at a sensor concentration of 10 nM and HBR-3,5DOM concentrations of **(B)** 1.95 μM **(C)** 3.9 μM **(D)** 8.4 μM, **(E)** 18 μM and **(F)** 36 μM. Excitation was fixed at 520 nm and emission at 600 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constant $K_{D,Ca}$ and Hill coefficient n_H.

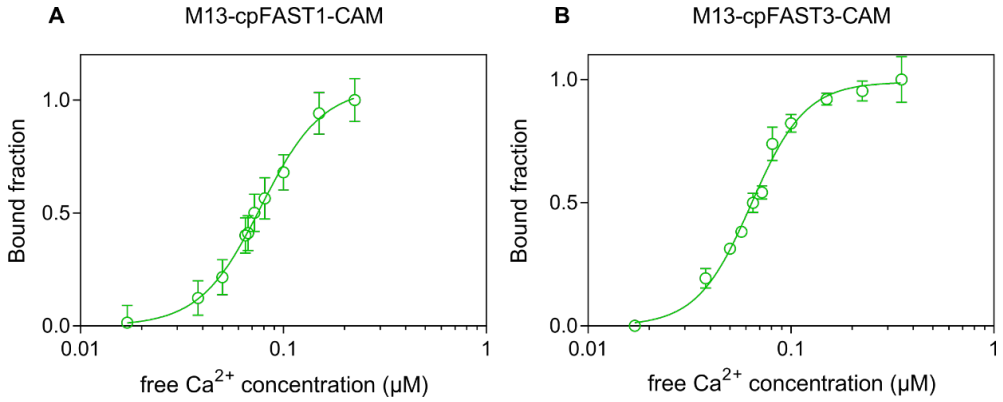

Figure S12. Ca^{2+} titration curves of M13-cpFAST-CaM in presence of HMBR. (A) M13-cpFAST1-CaM. **(B)** M13-cpFAST3-CaM. Titrations were done at a sensor concentration of 100 nM and a HMBR concentration of $(K_{D,+}K_{D,-})^{1/2}$ at 25 °C. Excitation was fixed at 480 nm and emission at 540 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constant $K_{D,Ca}$ and Hill coefficient n_H provided in **Table 2**.

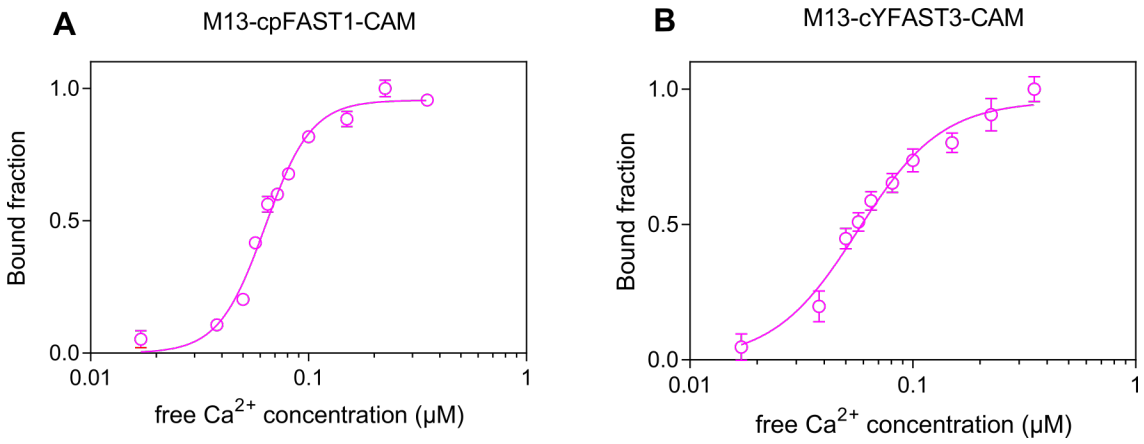

Figure S13. Ca^{2+} titration curves of M13-cpFAST-CaM in presence of HBR-3,5-DOM. **(A)** M13-cpFAST1-CaM. **(B)** M13-cpFAST3-CaM. Titrations were done at a sensor concentration of 100 nM and a HBR-3,5-DOM concentration of $(K_{D,+}K_{D,-})^{1/2}$ at 25 °C. Excitation was fixed at 520 nm and emission at 600 nm. Data represent mean ± sem (n = 3). Least square fit (line) gave the thermodynamic dissociation constant $K_{D,Ca}$ and Hill coefficient n_H provided in **Table 2**.

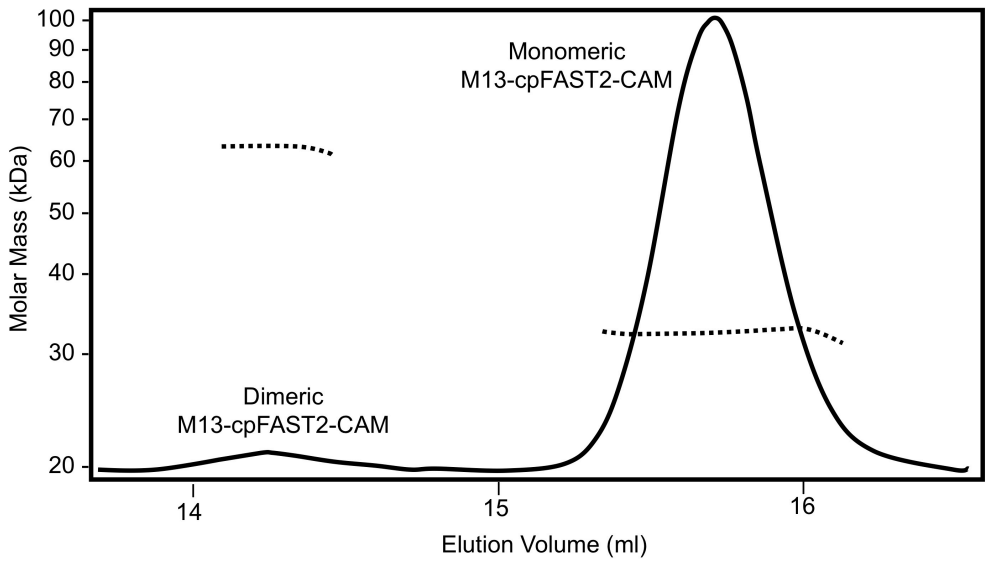

Figure S14. SEC-MALS analysis of recombinant M13-cpFAST2-CaM in 10 mM HEPES, 50 mM NaCl, 2 mM $CaCl_2$, pH 7.5. Continuous line: static light scattering, normalised from 3 angles; dotted line: molecular mass in kDa. Concentration at the monomer peak was ~ 5 mg/ml = 161 µM. In the conditions used, 96% of M13-cpFAST2-CaM (33.4 kDa) was monomeric and 4% was found to be dimeric. Note that no dimers were detected when the SEC column was equilibrated in buffer containing 1 mM EGTA instead of 2 mM $CaCl_2$ (i.e. in absence of Ca^{2+}).

MATERIALS AND METHODS

General FAST is a variant of the photoactive yellow protein (PYP) containing the mutations C69G, Y94W, T95M, F96I, D97P, Y98T, Q99S, M100R, T101G. HMBR and HBR-3,5-DOM were prepared as previously described[1,2].

Molecular Biology The genes encoding the circularly permuted FAST variants were constructed by PCR assembly from the gene encoding FAST optimized for human codon usage[1]. The plasmids for bacterial expression of cpFAST1-7 were generated by inserting the cpFAST1-7 DNA sequence in the plasmid pET18a using the *Nhe* I and *Xho* I restriction sites. The plasmids for mammalian expression of mcherry-cpFAST1-5 were constructed by replacing FAST DNA sequence by the cpFAST1-5 DNA sequence in pAG96 encoding mCherry-GGGS-FAST[1]. The calcium sensors were generated from a synthetic sequence (Eurofins) coding for M13-linker-CaM and inserted in the plasmid pET28a using *Nhe* I and *Xho* I. The sequence for M13 used was MVDSSRRKWNKTGHAVRAIGRLSS. The sequence for CaM was DQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADG DGTIDFPEFLTMMARKMNDTDSEEEIREAFRVFDKDGNGYIGAAELRHVMTNLGEKLT DEEVDEMIRVADIDGDGQVNYEEFVQMMTAK. The various cpFAST DNA sequences were amplified by PCR and inserted using *Mfe* I and *Sac* I restriction sites present in the linker between M13 and CaM. The plasmids for mammalian expression of the circularly permuted FAST variants and the cpFAST-based calcium biosensors were generated by inserting the corresponding DNA sequences into the previously reported pAG104 plasmid[1] using *BamH* I and *Bgl* II

Protein expression and purification Expression vectors were transformed in Rosetta (DE3) pLysS *E. coli* (New England Biolabs). Cells were grown at 37°C in Lysogeny Broth (LB) medium complemented with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol to OD_{600nm} 0.6. Expression was induced for 4 h by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested by centrifugation (4,000 × g for 20 min at 4°C) and frozen. The cell pellet was resuspended in lysis buffer (phosphate buffer 50 mM, NaCl 150 mM, $MgCl_2$ 2.5 mM, protease inhibitor, DNase, pH 7.4) and sonicated (5 min at 20 % of amplitude, 3 sec on, 1 sec off). The

lysate was incubated for 2 h at 4 °C to allow DNA digestion by DNase. Cellular fragments were removed by centrifugation (9200 × g for 1h at 4°C). The supernatant was incubated overnight at 4°C under gentle agitation with Ni-NTA agarose beads in phosphate buffered saline (PBS) (sodium phosphate 50 mM, NaCl 150 mM, pH 7.4) complemented with 10 mM imidazole. Beads were washed with 20 volumes of PBS containing 20 mM Imidazole, and with 5 volumes of PBS complemented with 40 mM Imidazole. His-tagged proteins were eluted with 5 volumes of PBS complemented with 0.5 M Imidazole. The buffer was exchanged to PBS (50 mM phosphate, 150 mM NaCl, pH 7.4) for cpFAST or HEPES buffer (HEPES 50 mM, NaCl 259 mM, pH 7.4) for calcium sensors using PD-10 desalting columns.

Instrumentation Steady state UV-Vis absorption spectra were recorded using a Cary 300 UV-Vis spectrometer (Agilent Technologies), equipped with a Versa20 Peltier-based temperature-controlled cuvette chamber (Quantum Northwest) and fluorescence data were recorded using a LPS 220 spectrofluorometer (PTI, Monmouth Junction, NJ), equipped with a TLC50TM Legacy/PTI Peltier-based temperature-controlled cuvette chamber (Quantum Northwest). Thermodynamic dissociation constants were determined with a Spark 10M plate reader (Tecan).

Thermodynamic analysis The thermodynamic dissociation constants $K_{D,+}$ and $K_{D,-}$ were determined in HEPES buffer (HEPES 50 mM, NaCl 259 mM, pH 7.4) containing either 10 mM $CaCl_2$ or 10 mM EGTA, while the pH in solution was kept constant at 7.4. Normalized fluorescence (or bound fraction) were plotted as a function of fluorogen concentration and fitted using one site specific binding equation to provide $K_{D,+}$ and $K_{D,-}$. The thermodynamic dissociation constants $K_{D,Ca}$ were determined using buffers with precise free Ca^{2+} concentration either purchased commercially (ThermoFisher) or prepared following protocols previously reported by Roger Tsien *et al.*[3] by reciprocal dilution of a Ca^{2+} saturated solution (10 mM CaEGTA in HEPES buffer) and a Ca^{2+} free solution (10 mM K_2EGTA in HEPES buffer) to give a series of solutions with free Ca^{2+} concentration ranging from 0 to 0.351 µM at pH 7.2. In these experiments, the fluorogen concentration was fixed at the optimal fluorogen concentration $(K_{D,+}K_{D,-})^{1/2}$. The normalized

fluorescence intensity of the calcium sensors was plotted as a function of Ca^{2+} concentration and fitted to the Hill equation used for cooperative ligand-receptor binding.

Size Exclusion Chromatography – Multi-Angle Light Scattering (SEC-MALS) Purified M13-cpFAST2-CaM (3.5 mg at 113 µM) was injected in a Superdex 200 Increase (GE Lifesciences) previously equilibrated in 10 mM HEPES, 50 mM NaCl, 2 mM $CaCl_2$, pH 7.5 connected to a TREOS static light detector and to a t-Rex refractometer (both purchased from Wyatt Technologies). Molecular mass was determined by the Astra software (Wyatt).

Mammalian cell culture HeLa cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with phenol red, Glutamax I, and 10% (vol/vol) fetal calf serum (FCS), at 37 °C in a 5% CO_2 atmosphere. For imaging, cells were seeded in µDish IBIDI (Biovalley) coated with poly-L-lysine. Cells were transiently transfected using Genejuice (Merck) according to the manufacturer's protocol for 24 hours prior to imaging.

Fluorescence microscopy Confocal micrographs were acquired on a Zeiss LSM 710 Laser Scanning Microscope equipped with a Plan Apochromat 63×/1.4 NA oil immersion objective. For calcium imaging, cells were washed twice with Dulbecco's Phosphate-Buffered Saline (DPBS) immediately prior to imaging. After rinsing, solution of fluorogen in HHBSS (20 mM HEPES buffered Hanks balanced salt solution (HBSS) containing 2 g/L D-glucose) were added to the cells. Histamine-induced Ca^{2+} dynamics were imaged at intervals of 250 ms for 15-20 min. Approximately 2 min after the start of the experiment, histamine (50 µM final concentration) in HHBSS containing the appropriate amount of fluorogen to maintain it constant was added to allow histamine-induced oscillations to subside. ZEN software was used to collect the data. Raw images were analyzed with Image J using a line ROI with a width of 15 pixels. Presented images are a maximum projection of 10 raw images while the movies were generated as a walking average of 5 frames.

SI References

(1) Plamont, M.-A., Billon-Denis, E., Maurin, S., Gauron, C., Pimenta, F. M., Specht, C. G., Shi, J., Querard, J., Pan, B., Rossignol, J., Moncoq, K., Morellet, N., Volovitch, M., Lescop, E., Chen, Y., Triller, A., Vriz, S., Le Saux, T., Jullien, L., and Gautier, A. (2016) Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. *Proc Natl Acad Sci USA 113*, 497–502.

(2) Li, C., Plamont, M.-A., Sladitschek, H. L., Rodrigues, V., Aujard, I., Neveu, P., Le Saux, T., Jullien, L., and Gautier, A. (2017) Dynamic multicolor protein labeling in living cells. *Chem Sci 8*, 5598–5605.

(3) Tsien, R., and Pozzan, T. (1989) Measurement of cytosolic free Ca^{2+} with quin2. *Methods in Enzymology*, pp 230–262. Elsevier.