High-performance probes for light and electron microscopy

Howard Hughes Medical Institute

Published as: *Nat Methods*. 2015 June ; 12(6): 568–576.

High-performance probes for light and electron microscopy

Sarada Viswanathan[1], **Megan E. Williams**[2], **Erik B. Bloss**[1], **Timothy J. Stasevich**[1,6], **Colenso M. Speer**[3], **Aljoscha Nern**[1], **Barret D. Pfeiffer**[1], **Bryan M. Hooks**[1,4,6], **Wei-Ping Li**[1], **Brian P. English**[1], **Teresa Tian**[1], **Gilbert L. Henry**[1], **John J. Macklin**[1], **Ronak Patel**[1], **Charles R. Gerfen**[4], **Xiaowei Zhuang**[3,5], **Yalin Wang**[1,6], **Gerald M. Rubin**[1], and **Loren L. Looger**[1]

Loren L. Looger: loogerl@janelia.hhmi.org

[1]Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA

[2]Department of Neurobiology & Anatomy, University of Utah, Salt Lake City, UT

[3]Howard Hughes Medical Institute, Department of Chemistry & Chemical Biology, Harvard University, Cambridge, MA

[4]Laboratory of Systems Neuroscience, National Institute of Mental Health, Bethesda, MD

[5]Howard Hughes Medical Institute, Department of Physics, Harvard University, Cambridge, MA

Abstract

We describe an engineered family of highly antigenic molecules based on GFP-like fluorescent proteins. These molecules contain numerous copies of peptide epitopes and simultaneously bind IgG antibodies at each location. These “spaghetti monster” fluorescent proteins (smFPs) distribute well in neurons, notably into small dendrites, spines and axons. smFP immunolabeling localizes weakly expressed proteins not well resolved with traditional epitope tags. By varying epitope and scaffold, we generated a diverse family of mutually orthogonal antigens. In cultured neurons and mouse and fly brains, smFP probes allow robust, orthogonal multi-color visualization of proteins,

Correspondence to: Loren L. Looger, `loogerl@janelia.hhmi.org`.

[6]Present addresses: Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO (T.J.S.); Department of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, PA (B.M.H.); Advanced Microscopy Facility, University of Virginia School of Medicine, Charlottesville, VA (Y.W.)

ACCESSION CODES AND CONSTRUCT AVAILABILITY

Sequences for all constructs have been deposited in Gen Bank: KR107019-KR107024. Constructs have been deposited at Add gene (addgene.org). AAV viruses expressing the smFP probes are available from the University of Pennsylvania Vector Core (http://www.med.upenn.edu/gtp/vectorcore/Catalogue.shtml). Fly stocks are available from the Bloomington Stock Center. Other constructs are available upon request.

AUTHOR CONTRIBUTIONS

S.V., G.M.R. and L.L.L. conceived the project. L.L.L. performed molecular modeling and designed sequences. S.V. and B.D.P. constructed the clones. S.V. and M.E.W. performed experiments in cultured neurons. M.E.W. performed hippocampal neuron work. B.M.H. and C.R.G. performed 4-color labeling experiments. E.B.B. performed array tomography experiments. C.M.S. and X.Z. designed STORM experiments and C.M.S. performed STORM imaging and analyzed data. J.J.M. and R.P. performed biophysical characterization. A.N. performed fly experiments. W.L. and Y.W. performed electron microscopy. T.J.S. and B.P.E. performed single-molecule imaging. T.T. and G.L.H. performed pull-down experiments. S.V. and L.L.L. led the project.

SUPPLEMENTARY INFORMATION

Supplemental information is available in the online version of this paper.

METHODS

Complete methods are available with the online version of the paper.

COMPETING FINANCIAL INTERESTS

The authors have no competing financial interests.

cell populations and neuropil. smFP variants complement existing tracers, greatly increase the number of simultaneous imaging channels, and perform well in advanced preparations such as array tomography, super-resolution fluorescence imaging and electron microscopy. In living cells, the probes improve single-molecule image tracking and increase yield for RNA-Seq. These probes facilitate new experiments in connectomics, transcriptomics and protein localization.

INTRODUCTION

Protein tags are ubiquitous tools in all areas of biology[1]. Although many types of tags exist, the two most commonly used are peptide antigens ("epitopes")[2] and fluorescent proteins (FPs). Epitope tags are short antigenic peptide sequences that facilitate immunohistochemistry (IHC) with tag-specific antibodies when attached to a protein of interest (POI). The principal advantage of epitope tags for IHC is the availability of reliable primary antibodies for detection, particularly when antibodies to the POI are non-specific, raised in the same species as antibodies to other targets, or unavailable entirely. Almost all epitope tagging experiments draw upon a small set of validated peptide antigens, including influenza hemagglutinin (HA)[3], myelocytomatosis viral oncogene (myc)[4], simian virus 5-derived epitope (V5)[5], the synthetic peptide FLAG[6], the synthetic streptavidin-binding strep-tag[7], and more recently OLLAS (*Escherichia coli* OmpF linker and mouse langerin)[8] and Sun Tag[9]. The small size of epitope tags (typically 8–12 amino acids) enables their attachment to POIs, even in multiple copies, without affecting protein folding, targeting or protein-protein interactions. However, the affinity of antibodies for small tags can be low; single or even multimeric tags are frequently insufficient for detection when the POI is weakly expressed. Furthermore, peptide epitopes are not stably expressed in cells without fusion to a scaffold protein[10].

Alternatively, FPs may be used in fusions to visualize POI localization, or expressed alone as cell-filling tracers. *Aequorea victoria* green fluorescent protein (GFP), for example, is soluble, bright, stable, and generally well tolerated by cells for protein localization, isolation and tracking[11]. The existing FP toolkit offers fluorescence across the visible spectrum[12] and compared to peptide antigens, FPs can offer higher affinity for IHC, as well as pre-IHC live fluorescence imaging.

Despite these advantages, endogenously fluorescent FPs are not suitable in many applications. The broad excitation/emission spectra of FPs hinder native imaging in combinations of more than 2 or 3, and many anti-FP antibodies cross-react with related probes, severely limiting options for IHC with multiple FP channels. Additionally, low FP expression levels may be insufficient for target localization while over-expression of most coral-derived FPs can result in aggregation and cytotoxicity, while failing to uniformly label neurites and other small structures.

To overcome the limitations of existing FP and peptide epitopes, we developed new molecular tags combining the advantages of both. Specifically, an ideal probe should combine the solubility, cell tolerance and optional endogenous fluorescence of FPs (FPs can easily be rendered dark, while retaining their 3-dimensional structure), together with orthogonal antibody recognition and tagging of POIs with multiple epitope copies. Here, we

describe a new family of extremely antigenic protein tags called “spaghetti monster” fluorescent proteins (smFPs). smFPs have 10–15 copies of single epitope tags strategically inserted into an FP scaffold with either an intact or darkened chromophore. smFPs permit robust, multi-color tracing of neurons and processes in multiple independent channels easily separable by conventional epifluorescence filter sets. This expands options for labeling and following defined populations of neurons and other cell types through brain tissue, where experiments are typically limited to a single excellent channel (GFP), with a handful of inferior options for second and third channels. The modular construct design facilitates further expansion of this toolkit, and a common scaffold helps to normalize tracer expression level, sub-cellular localization and half-life. In a range of advanced sample preparations and imaging strategies we show that smFPs are high-performance probes for light and electron microscopy applications as well as for molecular biology and biochemistry.

RESULTS

Molecular design and preliminary characterization

To create hyperantigenic labels, we chose protein scaffolds that would accommodate numerous peptide tag insertions while retaining their proper folding and cellular trafficking. GFP is soluble, stable, well tolerated by cells and accommodates the addition of peptide epitopes to its N- and C-termini and internal loops[13]. “Superfolder” GFP (sfGFP) is a hyperstable variant[14] that accepts large insertions into its loops while retaining folding and fluorescence[15]. Thus, we reasoned that sfGFP would be an ideal scaffold for antigen presentation. Peptide tags were designed into sfGFP to optimize their antigenicity and permit simultaneous binding of multiple antibodies.

We chose six epitope tags (HA, myc, V5, FLAG, OLLAS, strep II) primarily based on the commercial availability of high-affinity antibodies with widespread validation. Epitopes were inserted in sets of 4 into the internal 172–173 loop of sfGFP[15] and 3–4 epitopes were added to each terminus of sfGFP (Fig. 1a, Supp Fig. 1, Supp. Table 1). Fluorescence correlation spectroscopy (FCS) was used to quantify the number of antibodies bound to smFPs in solution, based on the change in diffusion properties with increasing molecular weight of the smFP-antibody complex. The smFP containing 10 FLAG peptides and an intact, fluorescent GFP chromophore (“smFP_FLAG_bright”) was expressed in bacteria, purified, and titrated with monoclonal IgG anti-GFP or monoclonal IgG anti-FLAG primary antibody. Then, the diffusion time (τ_D) of the GFP chromophore of smFP_FLAG_bright was determined (Fig. 1b, c). Titration with the monoclonal anti-GFP antibody yielded a τ_D of 0.57 ± 0.02 msec (std. dev., n = 5), consistent with a single binding event with a dissociation constant K_d< 10 nM. (K_d accuracy is limited to the smFP concentration, which was 10 nM in these experiments. See Online Methods for details.) Titration with the anti-FLAG antibody M2 yielded a τ_D of 1.31 ± 0.04 msec (n = 5), with saturation occurring at 100 nM antibody, consistent with the 100 nM concentration of epitopes (Fig. 1c). This measured value of τ_D corresponds to a molecular weight of 1700 kDa, consistent with 11.3 ± 1 bound anti-FLAG antibodies, based on a calibration series (Fig. 1b), where τ_D scales as the power law (complex MW)$^{0.39}$, and assuming an individual antibody MW of 150 kDa

(Online Methods). A similar conclusion about the number of bound antibodies was found when the 10xFLAG smFP was replaced by a 3xFLAG version (Fig. 1c), where the bound antibody number was determined to be 3.4 ± 0.3. Thus, the smFP format displays FLAG epitopes with full M2 antibody accessibility and high affinity. FCS and tests in *Drosophila* and mammalian cells showed that the smFPs are strongly stained by their corresponding antibodies, monomeric, with no aggregation or cytotoxicity (Fig. 1d, Supp. Figs. 1–4, Supp. Table 2). We also created smFPs based on mRuby2[16] and mWasabi[17] scaffolds, which are sequence-divergent from GFP and not detected with anti-GFP antibodies (Supp. Figs. 5–6).

Use as multi-channel connectomic tracers

GFP is widely used in tracing studies but provides only a single channel of fluorescence labeling. Red fluorescent proteins (RFPs) such as tdTomato are typically used as a second color channel, but exhibit low photostability and green emission. Antibody amplification of these probes typically weakly enhances signal and greatly increases background fluorescence. Furthermore, many RFP variants are cytotoxic, prone to aggregation and do not diffuse readily into fine processes. Other FP super family members, such as cyan and blue proteins, can suffer similar defects, and often cross-react with anti-GFP antibodies.

To demonstrate the increased performance of smFPs as connectomic tracers, we performed a set of neuronal labeling experiments and compared the results with published data from experiments with conventional FPs. smFPs were delivered by adeno-associated virus (AAV) serotype 2/1 into primary mouse vibrissal somatosensory cortex (S1). Opposite hemispheres of adult mice were infected with AAV expressing smFP_myc or smFP_FLAG (Fig. 2a, b). Intense cytoplasmic staining was observed at the injection sites. Both smFP probes labeled long-range projections with no signs of label aggregation or axonal blebbing. Axon projections were traced to several cortical and subcortical regions (Fig. 2c, Supp. Figs. 7 and 8a–c). Projection patterns observed here are consistent with previous tracing studies using GFP[18] and demonstrate that the smFPs are excellent long-distance neural tracers. A third probe, smFP_V5, showed similarly good efficacy in tracing long-distance projections from S1 (Supp. Fig. 8d). Thus the smFPs provide multiple orthogonal, high-contrast channels, suitable for labeling cells, neurites or specific proteins.

We tested smFP labels alongside EGFP, tdTomato and the blue fluorescent protein mTagBFP2[19]. Cre-dependent virus (AAV-FLEX-*CAG*) expressing either smFP_myc or mTagBFP2 was injected into layer 5 S1 of *Rbp4_KL100_Cre* transgenic animals (Supp. Fig. 9), which express Cre in a subset of L5 neurons[20]. While the injection site was prominent with mTagBFP2, contralateral projections were not clearly visible. With smFP_myc, long-distance projections (*e.g.* to contralateral S1 and piriform cortex) were bright and the filling complete (Supp. Fig. 9).

To demonstrate compatibility with existing tracers, we performed a four-color labeling experiment with EGFP, tdTomato, Ruby2_FLAG and smFP_myc. We injected four adjacent cortical areas (vibrissal primary motor cortex (vM1), forelimb M1 (fM1), vibrissal S1 (vS1) and lower limb S1 (llS1), unilaterally) in *Rbp4_KL100_Cre* mice with Cre-dependent AAV viruses encoding the 2 smFPs and 2 FPs (Fig. 2d–f, Supp. Fig. 10). Long-range projections could be traced throughout the brain, including projections to thalamic nuclei, ectorhinal and

perirhinal cortex (Fig. 2g–i), superior colliculus (Fig. 2j–l) and brainstem nuclei including the spinal trigeminal ganglia (Fig. 2m, n). Axons from each of the four projections could be visualized independently in the same tissue, *e.g.* within the cortical peduncle and medial posterior nucleus of the thalamus (POm).

Visualization of sub-cellular structures

smFPs containing multiple, high-affinity binding sites for primary antibodies should enable IHC labeling of fine neuronal structures with higher fidelity at lower concentrations than GFP. To test this, limiting concentrations of plasmids (~0.5 μl of 0.25 μg/μl per brain) encoding either EGFP or smFP_FLAG were delivered to hippocampal CA1 and immunostained with primary monoclonal antibodies against GFP or FLAG. Sections were treated with secondary antibodies conjugated to Alexa488 and imaged under identical conditions. Although GFP appeared similar or slightly brighter in somata, neuronal processes were brighter and more completely labeled with smFP_FLAG as compared to GFP (Fig. 3a–d). In particular, basal dendrites in the *stratum oriens*, the most distal dendrites in the *stratum lacunosum moleculare* and spines throughout the neuron were better resolved with smFP_FLAG than GFP (Fig. 3b, d).

Hippocampal CA3 pyramidal neurons receive specialized input from dentate gyrus mossy fiber axons[21–23] at “thorny excrescence” (TE) spines. Fluorescent visualization of individual thorns arising from a single spine neck has historically proven problematic, typically requiring microinjection of dye with subsequent immunohistochemical amplification[24]. While sustained high-level FP expression can sometimes resolve TEs[25], this precludes many experiments, including developmental analysis. Here we demonstrate that even the notoriously difficult-to-label CA3 TEs can be visualized with low levels of genetically encoded smFPs.

As before, limiting amounts of DNA encoding either EGFP or smFP_FLAG were delivered, this time to hippocampal CA3 neurons. The complex, multi-headed thorns were much more clearly resolved with smFP_FLAG than GFP (Fig. 3e, f). For comparison, Lucifer Yellow-filled neurons (Fig. 3g) had labeling density similar to that of smFP_FLAG. In a separate experiment, we delivered viruses encoding tdTomato and Ruby2_FLAG *in utero* into CA3. Again, thorns were more clearly resolved with Ruby2_FLAG than tdTomato (Supp. Fig. 11). Antibody amplification of the tdTomato signal failed to appreciably improve labeling (Supp. Fig. 12) and Ruby2_FLAG detected significantly more ($p = 0.0074$) spines than tdTomato (Supp. Fig. 12). Together, our results suggest that smFP constructs label fine neuronal processes at much lower expression levels than are needed for detection with conventional FPs.

Enhanced detection of low-abundance proteins

For proteins lacking suitable antibodies, epitope tagging provides a way to reveal sub-cellular distributions. High-affinity tags that provide specific labeling upon low exogenous expression, or from chromosomal knock-in, would be ideal for investigating proteins for which good primary antibodies are unavailable. We focused on N-cadherin (cadherin-2), a postsynaptic cell adhesion protein that plays a critical role in neural development[26–28]. Due

to the high sequence similarity between cadherin family members, anti-cadherin antibodies often recognize multiple species in tissue, precluding unambiguous localization of N-cadherin. We used N-cadherin fusion proteins to compare detection efficiency between smFPs and conventional epitope tags. The smFPs performed much better than 1x, 2x or 3x-HA tags (Fig. 3h–k, Supp. Figs. 13 and 14).

Single-molecule tracking in living cells

Single-molecule tracking using FP fusions is a powerful method for monitoring protein dynamics in live cells in real-time[29, 30]. However, FPs emit relatively few photons before photobleaching, which hampers localization precision and experimental timeframes in single-molecule imaging[31, 32]. Alternatively, enzymatic tags such as Halo[33], SNAP[34] and Clip[35] tags allow the covalent coupling of reactive small molecule dyes to POI-enzyme fusions[36]. However, such enzymatic tags only couple one dye molecule per POI, and as such provide limited signal amplification over direct FP fusions. smFPs are ideal alternatives for single-molecule tracking of specific POIs in living cells since their hyperantigenicity combined with loading of dye-coupled antibodies into the cell would enable visualization of the POI with a substantially brighter signal from multiple antibodies bound to the target epitope.

As proof of principle, we separately fused the histone subunit H2B to EGFP, Halo tag and the smFP Ruby2_FLAG, and expressed the chimeric proteins in HeLa cells *via* transient transfection for 3–6 hrs. Halo tag-expressing cells were bead-loaded with Alexa488-conjugated Halo-tag substrate. Ruby2_FLAG-expressing cells were bead-loaded with Alexa488-conjugated Fab antibody fragments against FLAG. smFP-labeled cells showed the brightest images from single-molecule tracking, clearly resolving histone punctae (Fig. 3l) that were difficult to observe with other probes (Fig. 3m, n). Background labeling of the anti-FLAG antibody in the nucleus was minimal (Fig. 3o). Thousands of single molecules were tracked for each label and quantified. Signals from Ruby2_FLAG were several-fold brighter than those for both Halo tag and GFP, with a long tail of bound antibodies (up to ~7–10) (Fig. 3p, Supp. Movie 1). The smFPs thus provide much brighter alternatives to existing methods of single-molecule tracking, and facilitate multi-color tracking.

Utility in high-resolution microscopy

The above results highlight the utility of the smFP probes in traditional confocal and wide-field microscopy. Techniques that offer dramatic improvements in imaging resolution, including array tomography (AT)[37], super-resolution fluorescence imaging such as Stochastic Optical Reconstruction Microscopy (STORM), and electron microscopy (EM) may each benefit from the advantages of smFPs compared with conventional FP fusions or IHC against endogenous targets. Array tomography relies on embedding samples in acrylic resin, which may compromise sample antigenicity for subsequent IHC. Similarly, the strong fixation required for maximal ultrastructural preservation in material prepared for EM can weaken or destroy target antigenicity. For STORM and other localization-based super-resolution approaches, image resolution is in part dependent on the molecular density of target labeling and smFPs could provide superior antigenicity for enhanced density and

resolution. To examine the efficacy of smFPs for high-resolution imaging, we tested the smFP antigens under each modality.

Array tomography

To test the utility of smFPs for AT[37], we examined synaptic connectivity between interneurons expressing Cre recombinase under control of the somatostatin promoter (*Sst*-IRES-Cre mice)[38] and hippocampal CA1 pyramidal cells expressing EGFP (*Thy-1* line M)[39]. AAVs expressing Cre-dependent tdTomato or smFP_FLAG were injected into the dorsal hippocampus of double heterozygous mice, and AT was used to visualize putative synapses between interneuron axons and CA1 dendrites using an antibody against the presynaptic protein synapsin.

Anti-FLAG immunolabeling enabled visualization of the entire somato-dendritic axis of SSt^+ neurons, including small-diameter axonal fibers in the *stratum lacunosum moleculare* (SLM) located up to 500 μm from the cell soma (Fig. 4a, Supp. Fig. 15a, b). smFP_FLAG-labeled axons in the SLM exhibited varicosities but not blebs (Supp. Fig. 15c) and were found to form putative synaptic (*i.e*., $synapsin^+$) contacts onto the distal dendrites of GFP^+ pyramidal cells throughout serial ultra-thin sections (Fig. 4b), suggesting that the smFP_FLAG epitope yielded a complete fill without any obvious toxicity or disruption of synaptic connectivity. Both monoclonal and polyclonal anti-FLAG antibodies worked exceptionally well with secondary antibodies conjugated to several commonly used fluorophores, including imaging channels such as 405 nm (Supp. Fig. 15d–f), which often yield dim signal in IHC. The bright smFP fluorescence yielded high signal-to-noise wide-field images in as little as 80 msec/field.

In contrast, images from animals expressing tdTomato yielded signals that were substantially weaker, requiring wide-field exposure times of up to 5 sec (*i.e*. 2 orders of magnitude longer), and could not be enhanced with either commercial or in-house anti-RFP antibodies (data not shown). The smFP toolkit has substantial advantages over existing FP reagents in terms of image quality and acquisition speed, and will be ideal in AT connectomic studies where multiple genetically- or anatomically-defined cell types need to be reconstructed.

Super-resolution STORM imaging

STORM is a wide-field super-resolution fluorescence imaging approach that overcomes the diffraction limit in lateral and axial resolution *via* stochastic activation of individual photoswitchable molecules and precise localization of the activated molecules[40, 41]. Based on Nyquist sampling theory, the final resolution of the STORM image depends not only on the localization precision of each fluorophore, but also on probe density. By providing highly antigenic tags to label proteins for which antibody labeling is weak or nonexistent, smFPs could benefit super-resolution imaging by increasing the probe density relative to conventional FP or epitope fusions.

To examine the efficacy of smFPs for STORM, we co-injected Cre-dependent viruses (AAV-FLEX-*CAG*) driving expression of tdTomato and smFP_myc into L5 S1 of

Rbp4_KL100_Cre transgenic mice. Ultrathin cryosections of excised cortex were immunolabeled with primary antibodies for RFP and myc, and secondary antibodies coupled to photoswitchable organic dyes (Online Methods). Dendrites expressing tdTomato and smFP_myc showed dense labeling with both anti-dsRed (magenta) and anti-myc (green) probes (Fig. 5a). The localization density was qualitatively similar between the tdTomato and myc channels, and individual spine necks with diameters of ~100 nm were well labeled with smFP_myc, demonstrating the utility of this smFP as a dense cell filling probe for neurite imaging in STORM experiments (Fig. 5b).

The availability of multiple orthogonal smFPs each with comparable antigenicity and highly specific antibodies offers the potential for dense labeling of several targets simultaneously in multi-color super-resolution experiments. To demonstrate this, we co-electroporated N-cadherin-smFP_HA fusion and smFP_V5 constructs into mouse embryos *in utero* at E16.5. Ultrathin cryosections were immunolabeled with antibodies coupled to the photoswitchable dyes Atto488, Cy3B, and Alexa647 for three-color imaging (Supp. Fig. 16). smFP_V5 provided a dense cytoplasmic fill (Supp. Fig. 16c) that facilitated super-resolution imaging of dendritic spines in a manner comparable to anti-GFP labeling of the smFP scaffold (Supp. Fig. 16b, c). N-cadherin-smFP_HA showed strong labeling of postsynaptic spine heads and dendritic shafts, with much weaker labeling of spine necks (Supp. Fig. 16a), consistent with the regulated subcellular localization of N-cadherin by MAGUK proteins[42]. The density of each smFP label was sufficient for STORM and allowed for multi-color super-resolution imaging of both fine dendritic neuropil and N-cadherin cluster localization within spines. From these experiments, we conclude that smFPs provide alternate channels with effective probe densities for super-resolution imaging. smFPs will be particularly useful in multi-color STORM applications, where the smFP platform offers more flexibility for high-quality labeling of multiple channels compared with conventional FPs.

Electron microscopy

Electron microscopy remains the gold standard for unambiguous identification of synapses. However, the identification of specific cells and proteins within EM samples remains challenging because the antigenicity of most proteins fails to survive EM processing protocols. Given that the smFPs robustly enhance labeling in light microscopy experiments, we tested the efficacy of these probes by immuno EM in resin-embedded tissue.

We performed immunogold staining on 60-nm serial sections cut from resin-embedded mouse cortical tissue transfected with smFP_FLAG_bright and smFP_myc (Fig. 5c, Online Methods). Both probes showed specific labeling of transfected cell bodies and neurites with minimal background (Fig. 5d, f and Supp. Fig. 17a–d). Labeling was consistently strong in neurites spanning serial sections, allowing 3D reconstruction (Fig. 5e, g and Supp. Movies 2–5). Both probes were seen even in the finest neuronal processes with label density equivalent to EGFP (Supp. Fig. 17e–f). We also tested the ability of the smFP tags to survive osmium tetroxide (OsO_4) fixation. The smFP_HA label persisted even in samples fixed with 4% PFA / 0.2% glutaraldehyde/1% OsO_4, providing strong gold particle labeling (Fig. 5h, i). This enables immuno EM detection of defined cells and proteins in samples with good ultrastructure preservation. The availability of multiple smFPs and use of additional

distinguishable gold particle sizes, or other complementary detection methods, should allow expansion of the number of simultaneously detected proteins and/or traced cell populations.

In addition to these applications, smFPs also facilitate enhanced efficiency for nucleic acid profiling (Supp. Fig. 18).

DISCUSSION

smFPs expand the toolkit for multi-channel imaging and biochemistry to at least 6 strongly antigenic, mutually orthogonal labels. This increased number of color channels facilitates multi-color mapping of somata, dendritic arbors and axonal projections, as well as target proteins, in brain tissue. The tags work well in confocal and super-resolution STORM imaging, array tomography and immuno EM, facilitating a wide variety of new experiments.

Stochastic expression of the labels by multi-color FLP-out, a new method for multicolor stochastic labeling in *Drosophila*[43], gives rise to "Brainbow-like"[44] labeling with high signal-to-noise[43, 45, 46]. In addition to robust labeling, the use of a single protein scaffold for sets of antigenic tags increases sub-cellular co-localization, which aids in assigning label combinations and dosages in cells expressing multiple labels. Stochastic smFP expression can likely be adapted to other model organisms as well.

The high single-molecule tracking photon counts from orthogonal tags and antibodies will facilitate robust simultaneous multi-color molecule tracking in cells and *in vivo*. The high avidity of the smFP-antibody interactions will increase signal-to-noise in RNA-Seq and ChIP-Seq experiments targeting sparse cell types, increasing reliability of brain transcriptomics.

The success of the smFPs in the experiments shown here validates the design strategy. Biophysical characterization showed that smFP_FLAG bound the maximum possible number of IgG antibodies. Recently, the Sun Tag was shown to allow avid binding to linear epitope repeats, but suffered from low thermodynamic stability and poor expression levels of long repeats, leading to selection of much shorter repeats for optimal use in cells[9]. The smFP design format suffers from no such problems. The fluorescent protein backbone renders the antigens readily expressed and diffusible in cells, as evidenced by their penetration into small structures like axons and spines; expression levels were high under every tested setting, with no evidence of aggregation or cytotoxicity.

The modular nature of the design strategy and the compatibility with diverse FP scaffolds imply that the toolkit can be systematically expanded; more epitope repeats could likely be added to the scaffold format without affecting cellular stability and trafficking. Several FP scaffolds make the antigens orthogonal to anti-GFP antibodies, adding an extra imaging channel (particularly useful in animals already transgenically expressing GFP). Rendering FP chromophores dark, while preserving folding and stability, preserves spectral bandwidth for small molecule dyes. Alternatively, keeping FP chromophores intact permits the use of the labels in live imaging followed by *post hoc* immunohistochemistry to locate small, labeled regions for EM reconstruction.

ONLINE METHODS

Molecular biology

DNA encoding smFPs were ordered from DNA2.0. Genes encoding smFPs were sub-cloned into pRSETa (Life Technologies) for protein expression and purification in *Escherichia coli* BL21 (this adds an N-terminal His tag for purification, and increases the MW by 4 kDa). Genes encoding smFP variants were sub-cloned into the pCAGGS vector with a *CAG* promoter (*CMV* enhancer, β-actin promoter) and regulatory element from the woodchuck hepatitis virus (WPRE)[50] for expression in HeLa cells and *in utero* electroporation[51, 52]. For expression in flies, R59A05-GAL4[49] was used to drive expression of UAS-smFP reporter constructs. For expression in mice, GFP and smFP variants were expressed using an adeno-associated virus serotype 2/1 (AAV2/1) driving the probe under control of the human *synapsin-1* promoter or a Cre-dependent (FLEX) version of the *CAG* promoter; live virus was produced (Janelia Farm Viral Vector Core). All constructs were verified by sequencing.

Cell and neuronal cultures

Cells were obtained from the American Type Culture Collection (ATCC) and cultured according to their protocol. All cell lines used in these experiments have been verified to be free from contaminating mycoplasma, viruses and other cells. smFP variants were transfected using an Amaxa (Lonza) Nucleofector 96w shuttle device. 7e05 live HeLa cells were transfected with 1 μg DNA per shuttle well and plated onto two 35 mm MatTek plates. Cells were immunostained 24–48 hrs post transfection. Primary hippocampal neurons were obtained from P0 rat pups by dissection, dissociated with papain and plated onto coverslips coated with Poly-D-Lysine (PDL) at a density of 80–100,000 per coverslip and cultured in NBActiv4 medium (Brain Bits).

Fluorescence correlation spectroscopy (FCS)

The number of antibodies bound to an smFP was found from solution measurements of diffusion time of antibody-bound smFP_FLAG_bright using two-photon FCS. Calibration of diffusion time versus molecular weight was obtained using the following markers: hydrolyzed Alexa488, 534 Da (A-20000, Invitrogen); hydrolyzed Alexa546, 963 Da (A-20002, Invitrogen); epidermal growth factor (EGF)-FITC, 6.5 kDa (E-3478, Invitrogen); EGFP, 32.7 kDa (4999-100 Biovision); RSET-smFP_FLAG_bright, 42.3 kDa; bovine serum albumin (BSA)-Alexa488, 69 kDa (A13100, Invitrogen); M2 anti-FLAG mAb-FITC, 153 kDa. Protein solutions were prepared in PBS buffer containing 0.2 mg/ml BSA. For antibody titrations, unlabeled anti-FLAG and anti-GFP antibodies were purchased (Supp. Table 2), and their concentrations were estimated from manufacturers' mg/ml specifications. Both smFP_FLAG_bright (10 FLAG epitopes) and smFP_FLAG_bright_3x (3 FLAG epitopes at the C-terminus) were used at 10 nM protein concentration for the titration, yielding 100 nM and 30 nM of antibody binding sites, respectively. Antibody-antigen solutions were incubated 30 minutes before measurements, then pipetted into coverslip-bottom dishes (MatTek) that had been pre-treated with 0.2 mg/ml BSA in PBS for 5 min, rinsed and dried, to block the surface. All measurements were taken at 25 °C on an inverted microscope (IX-81; Olympus) with a 1.2 NA water-immersion objective. Focused laser excitation at the sample was 2 mW of 940 nm light from a Ti:sapphire laser (Chameleon

Ultra II; Coherent), characterized by a beam radius w_0 of 430 nm at the focus. Details of the experimental setup and methods have been described elsewhere[53]. The diffusion time was found by fitting the fluorescence autocorrelation data to a diffusion model using a custom fitting program (Vijay Iyer, Janelia Farm) running on Matlab (Mathworks). The calibration of diffusion time versus molecular weight was obtained from fitting performed in Origin Pro 8 (Origin Lab). Diffusion coefficient D was related to diffusion time τ_D by the equation:

$$D = w_0^2 / 8\tau_D.$$

In utero electroporation

All procedures were performed according to the guidelines set by the Institutional Animal Care and Use Committees and Institutional Biosafety Committees of the University of Utah and Janelia Farm. Pregnant mothers (pups E14-E18) were deeply anesthetized with isoflurane (2%). The uterine horns were exposed and plasmid DNA (0.5 µl of ~5 µg/µl for most experiments; 0.5 µl of ~0.25 µg/µl for limiting expression) (Endo Q-prepped DNA mixed with 0.03% Fast Green dye in phosphate buffer), injected into the ventricle of 3–4 embryos through a micropipette (~0.1 µl per embryo) and electroporated using custom forceps electrodes (5 pulses, 100 ms, 40 V each).

Immunohistochemistry

Mice were perfused with 4% PFA and post fixed for 2 hours at room temperature. Brains were rinsed in 1x PBS (3×15 mins) and 50 µm thick coronal sections were cut on a vibratome. Sections were blocked in 3% BSA + 0.3% Triton in PBS for 1–2 hrs and incubated with primary antibody diluted in block overnight at 4 °C. Sections were rinsed in 0.3% Triton (3× 15 mins) and incubated in secondary antibody diluted in blocking buffer for 2 hrs at room temperature. Sections were rinsed as before, mounted on glass slides and cover-slipped with Vectashield (Vectashield). All antibody dilutions are shown in Supp. Table 2. Images were contrast-adjusted in ImageJ.

Intracranial viral injections

All procedures were performed according to the guidelines set by the Janelia Farm Research Campus Institutional Animal Care and Use Committee and Institutional Biosafety Committee. Animals (adult C57/BL6J, Charles River; either sex) mice were anesthetized under isoflurane and AAV virus encoding smFPs, serotype 2/1 was injected with a custom-made volumetric injection system (based on a Narishige MO-10 manipulator). Glass pipettes (Drummond) were pulled and beveled to a sharp tip (30 µm outer diameter), back-filled with mineral oil and front-loaded with viral suspension immediately before injection.

Injection coordinates for 2-color labeling experiment

Somatosensory cortex (S1) was injected in both hemispheres: smFP_myc into the left hemisphere and smFP_FLAG into the right hemisphere. Stereotactic coordinates for this region were: −0.59 mm anterior-posterior, 3 mm medial-lateral, and between −0.6 and −0.4 mm dorso-ventral, all relative to the Bregma suture.

Injection coordinates for 4-color labeling experiment

Four cortical regions were injected, all in the right hemisphere: tdTomato into vibrissal motor cortex (vM1), EGFP into vibrissal somatosensory cortex (vS1), smFP_myc into forelimb motor cortex (fM1), and smFP_FLAG into lower limb somatosensory cortex (llS1). Stereotactic coordinates for these regions were: vM1, 1.1 mm anterior-posterior, 0.9 mm medial-lateral, and between −0.5 and −0.8 mm dorso-ventral; vS1, −0.6 mm anterior-posterior, 2.0 mm medial-lateral, and between −0.5 and −0.8 dorso-ventral; fM1, 0.5 mm anterior-posterior, 2.0 mm medial-lateral, and between −0.5 and −0.8 dorso-ventral; and llS1, −1.0 anterior-posterior, 1.5 medial-lateral, and between −0.6 and −0.65 mm dorso-ventral, all relative to the Bregma suture.

4-color labeling

Two weeks following AAV injection (Intracranial viral injections), mouse brains were perfused, sectioned at 80 μm, and immunostained as described above (Immunohistochemistry). Sections were imaged on a Zeiss or Olympus microscope with a motorized stage with a 10x objective[20]. Tiled images of each section (~80–200 tiles per section) were taken and merged using Neurolucida software (MBF Bioscience).

Electron microscopy fixation

For HeLa cells transfected with pCAG-smFP_HA, fixation was carried out after 24–48 hrs of expression with 4% paraformaldehyde (PFA) plus 0.2% glutaraldehyde (glut) in 100 mM sodium cacodylate buffer on ice for 1 hour. Some samples were then post-fixed in 1% OsO_4. For mouse brains expressing smFP_myc and smFP_FLAG_bright (*in utero* electroporated), fixation was carried out by perfusion with 4% PFA. The brains were dissected out and cut into 150 μm thick slices with a Leica vibratome. The slices were kept in 2% PFA. Regions with smFP_FLAG_bright fluorescence were punched out with a 2 mm biopsy punch under a fluorescent dissecting microscope.

HM20 embedding

The punched-out regions from the brain vibratome sections were transferred to 20% BSA, then to the well of a Type B 0.15/0.15 mm specimen carrier (Techno Trade International), which was capped by the flat side. The carrier assembly was high-pressure frozen with an HPF Compact 01 high pressure-freezing machine (M. Wohlwend GmbH). The frozen carrier assembly was forced open in liquid nitrogen and freeze substituted in acetone with 0.1% uranyl acetate (UA) in a Leica EM AFS2 freeze substitution unit at −90 °C for 72 hours. Afterwards, the temperature was brought up slowly (1 °C/hour) to −45 °C; samples were rinsed with cold acetone three times. Samples were infiltrated with 30%, 50%, and 70% HM20 resin in acetone for 4 hours each and then 100% resin overnight, followed by infiltration in fresh 100% resin for 2 hours. Samples were then transferred to precooled BEEM capsules and UV polymerized at −45 °C.

Immunogold labeling

Ultrathin sections (60 nm) were cut from embedded samples with a Leica UC6 ultra-microtome and picked up on nickel EM grids for immunolabeling. Sections were incubated

with PBS (pH 7.4) for 5 minutes, 50 mM glycine in PBS for 15 minutes, and 1% BSA in PBS for 30 minutes. After incubation with primary antibodies in 1% BSA in PBS overnight at 4 °C, the sections were rinsed in PBS three times, followed by further washing in PBS for 4× 10 minutes. Sections were incubated with secondary immunogold-labeled antibodies for two hours at room temperature followed by three quick rinses in PBS and further washing in PBS for 4× 10 minutes. Immunogold labeling was stabilized by incubation with 1% glut in PBS for 10 minutes. Afterwards, sections were washed sequentially in PBS for 3× 5 minutes and water for 4× 5 minutes and let air-dry. The gold labels were amplified with silver enhancement (intensification kit from Nanoprobes) for 3 minutes. Finally, the sections were contrast-stained with 3% UA and Sato's triple lead.

For single immunolabeling of smFP_HA, antibodies were a mouse anti-HA primary and an immunogold (5 nm) goat anti-mouse secondary. For double immunolabeling of smFP_myc and smFP_FLAG, primary antibodies were goat anti-myc and rabbit anti-FLAG. Secondary antibodies were immunogold (6 nm) donkey anti-goat and 12 nm donkey anti-rabbit. All antibody dilutions were in 1% BSA in PBS. Antibody details are shown in Supp. Table 2.

TEM imaging

Sections were examined using an FEI Tecnai 20 TEM operated at 80kV. Images were acquired with a Gatan Ultrascan 4k × 4k camera. Serial section images were acquired as montages of 4k × 4k micrographs using the automated acquisition software Leginon. Montaging and 3D reconstruction were carried out using ImageJ-TrakEM2[54, 55].

Array tomography

In these experiments, AAV-FLEX-*CAG*-smFP_FLAG constructs were injected into dorsal hippocampus of heterozygous *Sst*-IRES-Cre mice bred with the *Thy-1*:EGFP mice using a stereotactic apparatus under isoflurane anesthesia. AAVs were allowed to express for at least two weeks, after which animals were perfused transcardially with ice-cold 4% paraformaldehyde and 0.125% glutaraldehyde in 100 mM phosphate buffer, pH 7.2. Following a 6-hour postfix in the same perfusate at 4 °C, 200 μm-thick coronal sections were made on a Vibratome and the CA1 region was manually micro-dissected under a fluorescent dissecting microscope. Tissue blocks containing CA1 were dehydrated in a series of ascending ethanol washes at 4 °C, infiltrated with LR White resin/Benzoyl peroxide, and polymerized by the addition of LR White accelerator (EMS). Following polymerization, ribbons of serial ultrathin (100 nm) sections were collected onto gelatin-coated (0.01% gelatin with 0.001% chromium potassium sulfate, Sigma) coverslips and labeled with anti-synapsin (Cell Signaling) and anti-FLAG antibodies (M2, Sigma) (Supp. Table 2) according to[37]. Arrays were imaged with a 63× 1.4 NA objective on a modified Zeiss Axio Observer inverted microscope equipped with standard filter sets and a custom-built infrared skew-beam autofocus system. Images were assembled into aligned z-stacks with custom software and putative synaptic contacts were visualized in TrakEM2.

Sample preparation for STORM

Brains injected or electroporated with smFP/FP constructs were removed following perfusion and post-fixed by submersion in 4% PFA overnight at 4 °C. Small cubes of cortex

and hippocampus were excised and infiltrated in a graded series (3/6/9/12%) of gelatin (Bovine skin type-B, Sigma) at 37 °C for 30 minutes each. Gelatin-embedded tissue was cryo-protected overnight in 2.3 M sucrose at 4 °C prior to plunge-freezing in liquid nitrogen. Tissue was sectioned at 100–500 nm on an ultracryotome (Leica) and sections were retrieved in a loop with a 10% methylcellulose/2.3M sucrose solution (Tokayasu method) and mounted on gelatin (0.5%) chromium potassium sulfate (0.05%) subbed coverglass or polylysine-coated (0.1%) coverglass. Sections were incubated in blocking buffer (10% normal donkey serum, Jackson Immunoresearch, with 0.3% Triton-X 100 and 0.3 M glycine) for 1 hr at room temperature (RT) prior to overnight incubation in primary antibodies diluted in blocking buffer. Antibodies used are shown in Supp. Table 2 (all at 1:100). Primary antibodies for Fig. 5 were rabbit anti-RFP (recognizes dsRed) and mouse anti-myc; primary antibodies for Supp. Fig. 16 were rat anti-HA, chicken anti-GFP and mouse anti-V5. Following primary incubation, tissue was rinsed 6 × 20 minutes in 1x PBS at RT and incubated in secondary antibodies in blocking buffer (1:100 dilution) at RT for 1 hour. Secondaries used were donkey anti-rabbit (DyLight750/Alexa405), donkey anti-mouse (Alexa647/Alexa405; reporter/activator), donkey anti-chicken (Cy3B), and donkey anti-rat (Atto488). Purchased secondary antibodies were dye-labeled using standard amine-reactive chemistry. Tissue was washed in 1x PBS 6 × 20 minutes and post-fixed 10 minutes at RT in 3% PFA/0.1% glut. Coverslips were transferred to 1x PBS for storage prior to imaging.

STORM imaging

Coverslips were mounted on flow chambers, filled with imaging buffer (17.5 μM glucose oxidase (Sigma), 10 mM cysteamine (Sigma), 200 mM Tris, 10 mM NaCl, 710 nM catalase (Sigma), 10% glucose, pH 7.4), and sealed with epoxy. Images were collected on a custom STORM system with laser excitation powers of 1–3 kW/cm^2 at 752, 647, 561 and 488 nm. A Quad view emission splitter (Photometrics) was used to separate individual fluorescent channels and images were collected on an EMCCD camera (Andor). 10–20,000 STORM frames at 30 or 60 Hz were acquired for each channel and 405 nm activation laser power was ramped slowly to maintain optimal single-molecule density for Alexa reporters. Single-molecule fitting[56] was performed offline using custom software (Insight3) and four-color image registration was performed in Matlab.

Nuclear antigen labeling and isolation (INTACT)

Antibody-adsorbed Dynal Protein-G beads (Invitrogen: 100-03D) were generated as described previously[57]. Briefly, 300 μl of beads were bound to 1μg of anti-GFP (Invitrogen: G10362), anti-FLAG (Sigma: F7425) or anti-V5 antibody (Abdserotec: MCA1360) in 1xPBS-0.01%Tween-20.

10XUAS-unc84-smFP_FLAG and 10XUAS-unc84-smFP_V5 reporter lines were generated by phiC31-mediated transgenesis[58]. *Tdc*-GAL4[59] was used to drive expression of both tags in octopaminergic neurons and nuclei were purified by the INTACT procedure[57]. Approximately 200 frozen heads were added to 20 ml homogenization buffer (10 mM β-glycerophosphate pH 7.0, 2 mM $MgCl_2$, 0.5% NP-40) and passed over a Yamato continuous flow homogenizer, set at 100 rpm, 6 times. The homogenate was then successively filtered through first a 20 μm and then a 10 μm nylon filter (Partec: 040042325, 0400422314). 60 μl

of Dynal Protein-G magnetic beads adsorbed to anti-GFP, anti-FLAG or anti-V5 antibody were then added to the filtered homogenate, which was subjected to constant agitation at 4 °C for 30 minutes. Bead-bound nuclei were captured on a magnet, washed 3x with homogenization buffer, filtered through a 10 μm nylon filter and analyzed by both light and fluorescence microscopy.

Antigen-binding fragment (Fab) preparation for single-molecule image tracking

The Pierce Fab Preparation kit (Thermo Fisher Scientific) was used to generate Fab from 5 mg of anti-DYKDDDDK (anti-FLAG) tag monoclonal antibody (Wako; 012-22384). Anti-FLAG Fabs were then concentrated up to ~2 mg/ml in PBS using an Ultra free 0.5 filter (10K cut-off; Millipore). Fab purity and integrity were tested by SDS-PAGE using a 10–20% gradient gel (Wako). To fluorescently conjugate Fab, Alexa488 tetrafluorophenyl ester (for labeling 1 mg protein; Invitrogen) was dissolved in 50 μl dimethyl sulphoxide (DMSO; Wako) and stored at −20 °C. On the day of conjugation, a 5 μl aliquot of the ester was mixed into a 100 μl solution containing 100 μg of Fab and 100 mM $NaHCO_3$ (pH 8.3). The mixture was then incubated for 1 h at room temperature with gentle rotation. A PD-mini G-25 desalting column (GE Healthcare) pre-equilibrated with PBS was then used to remove unconjugated dye. Fab were finally concentrated up to ~1 mg/ml using an Ultrafree 0.5 filter (10K cut-off; Millipore) and the dye:protein ratio (~1) was calculated from the absorbance at 280 and 494 using the extinction coefficient of IgG and the correction factor at 280 nm provided by the manufacturer (0.11).

Cell culture and live-cell imaging for single-molecule image tracking

HeLa cells were grown and imaged in Dulbecco's modified Eagle medium (DMEM) without Phenol Red (Invitrogen) with 10% fetal bovine serum on pre-cleaned coverslips. Approximately three hours before imaging, H2B-smFP or H2B-Halo tag[60] (a gift from J. McNally, NCI, NIH) was transiently transfected into cells with Opti-mem medium (Invitrogen) using the Lipofectamine 2000 reagent (Invitrogen). About one hour before imaging, cells expressing H2B-smFP or H2B-Halo tag were loaded with, respectively, Alexa488 conjugated anti-FLAG Fab or Alexa488 Halo ligand (Promega) using a bead loading method[61–63]. Briefly, the medium was removed from the dish and saved, 1 μl of fluorescent Fab concentrated at ~1/10 mg/ml (or, for H2B-Halo tag cells, 1 μl of 10 nM Alexa488 Halo ligand in PBS) was pipetted onto the coverslip center, and glass beads (106 μm; Sigma-Aldrich; G-4649) were sprinkled on top. After tapping the dish four to eight times, the original DMEM was added back to the dish and the cells were returned to an incubator. Just before imaging, the media was replaced (and glass beads washed out) with fresh, phenol red-free DMEM. As a control, untransfected cells were also bead-loaded with Alexa488-conjugated anti-FLAG Fab in the same way as cells transfected with H2B-smFP.

During imaging, cells were maintained at 37 °C, 5% CO_2, and constant humidity using a Tokai-hit stage top incubator (INUP-PPZI-F1). Live-cell single particle tracking experiments were recorded on a custom-built three camera RAMM frame (ASI) microscope using an Olympus 1.4 NA PLAPON 60x OSC objective, and a 300 mm focal length tube lens (LAO-300.0, Melles Griot), resulting in 100x overall magnification. Stroboscopic 488 nm excitation (laser on for 9 ms of the 20 ms camera exposure) of a Stradus 488–150 laser

(Vortran) operating at 20 mW power was achieved using a NI-DAQ-USB-6363 acquisition board (National Instruments) *via* Lab view 2012. Peak power levels were estimated at 1 kW/cm^2 at the sample. A 2 mm-thick quad-band excitation dichroic (ZT405/488/561/640rpx, Chroma), a 2 mm-thick emission dichroic (T490pxrxt, Chroma), and a band-pass emission filter (FF01-525/45-25, Semrock) filtered the emitted light. Fluorescence was detected with a back-illuminated EMCCD camera (Andor Technology, Ixon Ultra DU-897U-CS0-EXF, 17 MHz)[64]. All imaging experiments were performed on cells prepared independently on two separate days.

Particle tracking

Localizer[65] was used to track single H2B particles in acquired live-cell movies. The following settings were chosen for particle detection and track linking: 1 pixel maximum jump distance, 5 frame minimum track length, 8-way adjacent tracking, 1.3 standard deviation, and 20 GLRT sensitivity. Resulting tracks were then exported as text files using code written in Igor 6.3.4 (WaveMetrics), and Mathematica (Wolfram Research) was used to make a histogram of the intensities of particles from the first frame of each track.

Reproducibility statement

Sample sizes were chosen to be sufficient to show the magnitude of effects of the different probes. Most effects were large and qualitative and did not require statistical analysis of significance. Where statistical analysis was performed, all information on tests is provided. No data were excluded.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We would like to thank the Cell Culture, Vivarium, Fly Facility, Histology, Electron Microscopy, Media and Molecular Biology Shared Resources, and the Fly Light Project Team, at Janelia. N. Betley, A. Hantman, J. Colonell, J.-C Rah provided helpful discussions. H. Su and H. Kimura helped with reagents, B. Karsh assisted with image alignment for immuno-EM and array tomography and H. Rouault helped with statistical analysis. Members of the Looger lab and M. Jefferies provided helpful feedback during the project. A. Hantman and K. Ritola (Janelia) provided the mTagBFP2 virus.

REFERENCES

1. Waugh DS. Making the most of affinity tags. Trends Biotechnol. 2005; 23:316–320. [PubMed: 15922084]
2. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biot. 2003; 60:523–533.
3. Wilson IA, et al. The structure of an antigenic determinant in a protein. Cell. 1984; 37:767–778. [PubMed: 6204768]
4. Evan GI, Lewis GK, Ramsay G, Bishop JM. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. Mol. Cell. Biol. 1985; 5:3610–3616. [PubMed: 3915782]
5. Southern JA, Young DF, Heaney F, Baumgartner WK, Randall RE. Identification of an epitope on the P-Proteins and V-Proteins of simian-virus 5 that distinguishes between 2 isolates with different biological characteristics. J. Gen. Virol. 1991; 72:1551–1557. [PubMed: 1713260]

6. Hopp TP, et al. A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio-Technol. 1988; 6:1204–1210.
7. Schmidt TGM, Koepke J, Frank R, Skerra A. Molecular interaction between the strep-tag affinity peptide and its cognate target, streptavidin. J. Mol. Biol. 1996; 255:753–766. [PubMed: 8636976]
8. Park SH, et al. Generation and application of new rat monoclonal antibodies against synthetic FLAG and OLLAS tags for improved immunodetection. J Immunol. Methods. 2008; 331:27–38. [PubMed: 18054954]
9. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell. 2014; 159:635–646. [PubMed: 25307933]
10. Reits E, et al. A major role for TPPII in trimming proteasomal degradation products for MHC class I antigen presentation. Immunity. 2004; 20:495–506. [PubMed: 15084277]
11. Rizzo MA, Davidson MW, Piston DW. Fluorescent protein tracking and detection: fluorescent protein structure and color variants. Cold Spring Harb. Protoc. 2009; 2009 pdb top63.
12. Shaner NC, Patterson GH, Davidson MW. Advances in fluorescent protein technology. J. Cell Sci. 2007; 120:4247–4260. [PubMed: 18057027]
13. Abedi MR, Caponigro G, Kamb A. Green fluorescent protein as a scaffold for intracellular presentation of peptides. Nucleic Acids Res. 1998; 26:623–630. [PubMed: 9421525]
14. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 2006; 24:79–88. [PubMed: 16369541]
15. Kiss C, et al. Antibody binding loop insertions as diversity elements. Nucleic Acids Res. 2006; 34:E132–E132. [PubMed: 17023486]
16. Lam AJ, et al. Improving FRET dynamic range with bright green and red fluorescent proteins. Nat. Methods. 2012; 9:1005–1012. [PubMed: 22961245]. [PubMed: 22961245]
17. Ai HW, Olenych SG, Wong P, Davidson MW, Campbell RE. Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. BMC Biol. 2008; 6:13. [PubMed: 18325109]. [PubMed: 18325109]
18. Mao T, et al. Long-range neuronal circuits underlying the interaction between sensory and motor cortex. Neuron. 2011; 72:111–123. [PubMed: 21982373]
19. Subach OM, Cranfill PJ, Davidson MW, Verkhusha VV. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. PLoS One. 2011; 6:e28674. [PubMed: 22174863]
20. Gerfen CR, Paletzki R, Heintz N. GENSAT BAC cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. Neuron. 2013; 80:1368–1383. [PubMed: 24360541]
21. Ramón, y; Cajal, S. Histologie du Système Nerveux de l'Homme et des Vértebrés. Azoulay, L., translator. Instituto Ramón y Cajal, Madrid, 1952, translated from Spanish by
22. Amaral DG, Dent JA. Development of the mossy fibers of the dentate gyrus: I A light and electron microscopic study of the mossy fibers and their expansions. J. Comp. Neurol. 1981; 195:51–86. [PubMed: 7204652]
23. Chicurel ME, Harris KM. Three-dimensional analysis of the structure and composition of CA3 branched dendritic spines and their synaptic relationships with mossy fiber boutons in the rat hippocampus. J. Comp. Neurol. 1992; 325:169–182. [PubMed: 1460112]
24. Williams ME, et al. Cadherin-9 regulates synapse-specific differentiation in the developing hippocampus. Neuron. 2011; 71:640–655. [PubMed: 21867881]
25. McAuliffe JJ, et al. Altered patterning of dentate granule cell mossy fiber inputs onto CA3 pyramidal cells in limbic epilepsy. Hippocampus. 2011; 21:93–107. [PubMed: 20014385]
26. Redies C. Cadherin expression in the developing vertebrate CNS: from neuromeres to brain nuclei and neural circuits. Exp. Cell Res. 1995; 220:243–256. [PubMed: 7556431]
27. Fannon AM, Colman DR. A model for central synaptic junctional complex formation based on the differential adhesive specificities of the cadherins. Neuron. 1996; 17:423–434. [PubMed: 8816706]

28. Uchida N, Honjo Y, Johnson KR, Wheelock MJ, Takeichi M. The catenin cadherin adhesion system is localized in synaptic junctions bordering transmitter release zones. J. Cell Biol. 1996; 135:767–779. [PubMed: 8909549]
29. Ritchie K, Kusumi A. Single-particle tracking image microscopy. Method. Enzymol. 2003; 360:618–634.
30. Seefeldt B, et al. Fluorescent proteins for single-molecule fluorescence applications. J. Biophot. 2008; 1:74–82.
31. Ha T, Tinnefeld P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annu. Rev. Phys. Chem. 2012; 63:595–617. [PubMed: 22404588]
32. Martin-Fernandez ML, Clarke DT. Single molecule fluorescence detection and tracking in mammalian cells: the state-of-the-art and future perspectives. Int. J. Mol. Sci. 2012; 13:14742–14765. [PubMed: 23203092]
33. Los GV, et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem. Biol. 2008; 3:373–382. [PubMed: 18533659]
34. Kolberg K, Puettmann C, Pardo A, Fitting J, Barth S. SNAP-tag technology: a general introduction. Curr. Pharm. Design. 2013; 19:5406–5413.
35. Gautier A, et al. An engineered protein tag for multiprotein labeling in living cells. Chem. Biol. 2008; 15:128–136. [PubMed: 18291317]
36. Gebhardt JC, et al. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. Nat. Methods. 2013; 10:421–426. [PubMed: 23524394]
37. Micheva KD, Smith SJ. Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. Neuron. 2007; 55:25–36. [PubMed: 17610815]
38. Lovett-Barron M, et al. Regulation of neuronal input transformations by tunable dendritic inhibition. Nat. Neurosci. 2012; 15:423–430. [PubMed: 22246433]
39. Feng G, et al. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. Neuron. 2000; 28:41–51. [PubMed: 11086982]
40. Rust MJ, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods. 2006; 3:793–795. [PubMed: 16896339]
41. Huang B, Bates M, Zhuang X. Super-resolution fluorescence microscopy. Annu. Rev. Biochem. 2009; 78:993–1016. [PubMed: 19489737]
42. Wang SH, et al. Dlg5 regulates dendritic spine formation and synaptogenesis by controlling subcellular N-cadherin localization. J. Neurosci. 2014; 34:12745–12761. [PubMed: 25232112]
43. Nern A, Pfeiffer BD, Rubin GM. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc. Natl. Acad. Sci. USA. 2015; 112:E2967–E2976. [PubMed: 25964354]. [PubMed: 25964354]
44. Livet J, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature. 2007; 450:56–62. [PubMed: 17972876]
45. Aso Y, et al. The neuronal architecture of the mushroom body provides a logic for associative learning. Elife. 2014:e04577. [PubMed: 25535793]
46. Wolff T, Iyer NA, Rubin GM. Neuroarchitecture and neuroanatomy of the *Drosophila* central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits. J. Comp. Neurol. 2014; 523:997–1037. [PubMed: 25380328]
47. Tyn MT, Gusek TW. Prediction of diffusion coefficients of proteins. Biotechnol. Bioeng. 1990; 35:327–338. [PubMed: 18592527]
48. Petrášek Z, Schwille P. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. Biophys. J. 2008; 94:1437–1448. [PubMed: 17933881]
49. Jenett A, et al. A GAL4-driver line resource for *Drosophila* neurobiology. Cell Rep. 2012; 2:991–1001. [PubMed: 23063364]
50. Gray NW, Weimer RM, Bureau I, Svoboda K. Rapid redistribution of synaptic PSD-95 in the neocortex in vivo. PLoS biology. 2006; 4:e370. [PubMed: 17090216]
51. Saito T, Nakatsuji N. Efficient gene transfer into the embryonic mouse brain using *in vivo* electroporation. Dev. Biol. 2001; 240:237–246. [PubMed: 11784059]

52. Tabata H, Nakajima K. Efficient *in utero* gene transfer system to the developing mouse brain using electroporation: Visualization of neuronal migration in the developing cortex. Neuroscience. 2001; 103:865–872. [PubMed: 11301197]
53. Mütze J, et al. Excitation spectra and brightness optimization of two-photon excited probes. Biophys. J. 2012; 102:934–944. [PubMed: 22385865]
54. Rah JC, et al. Thalamocortical input onto layer 5 pyramidal neurons measured using quantitative large-scale array tomography. Frontiers in Neural Circuits. 2013; 7
55. Cardona A, et al. TrakEM2 software for neural circuit reconstruction. PLoS One. 2012; 7:e38011. [PubMed: 22723842]
56. Bates M, Huang B, Dempsey GT, Zhuang X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. Science. 2007; 317:1749–1753. [PubMed: 17702910]
57. Henry GL, Davis FP, Picard S, Eddy SR. Cell type-specific genomics of *Drosophila* neurons. Nucleic Acids Res. 2012; 40:9691–9704. [PubMed: 22855560]
58. Groth AC, Fish M, Nusse R, Calos MP. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. Genetics. 2004; 166:1775–1782. [PubMed: 15126397]
59. Cole SH, et al. Two functional but noncomplementing *Drosophila* tyrosine decarboxylase genes: distinct roles for neural tyramine and octopamine in female fertility. J. Biol. Chem. 2005; 280:14948–14955. [PubMed: 15691831]
60. Mazza D, Abernathy A, Golob N, Morisaki T, McNally JG. A benchmark for chromatin binding measurements in live cells. Nucleic Acids Res. 2012; 40:e119. [PubMed: 22844090]
61. Hayashi-Takanaka Y, et al. Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. Nucleic Acids Res. 2011; 39:6475–6488. [PubMed: 21576221]
62. Mcneil PL, Warder E. Glass-beads load macromolecules into living cells. J. Cell Sci. 1987; 88:669–678. [PubMed: 2459146]
63. Stasevich TJ, et al. Regulation of RNA polymerase II activation by histone acetylation in single living cells. Nature. 2014; 516:272–275. [PubMed: 25252976]
64. Edelstein, A.; Amodaj, N.; Hoover, K.; Vale, R.; Stuurman, N. Computer control of microscopes using μManager. John Wiley & Sons, Inc; 2010.
65. Dedecker P, Duwe S, Neely RK, Zhang J. Localizer: fast, accurate, open-source, and modular software package for superresolution microscopy. J. Biomed. Opt. 2012; 17:126008. [PubMed: 23208219]

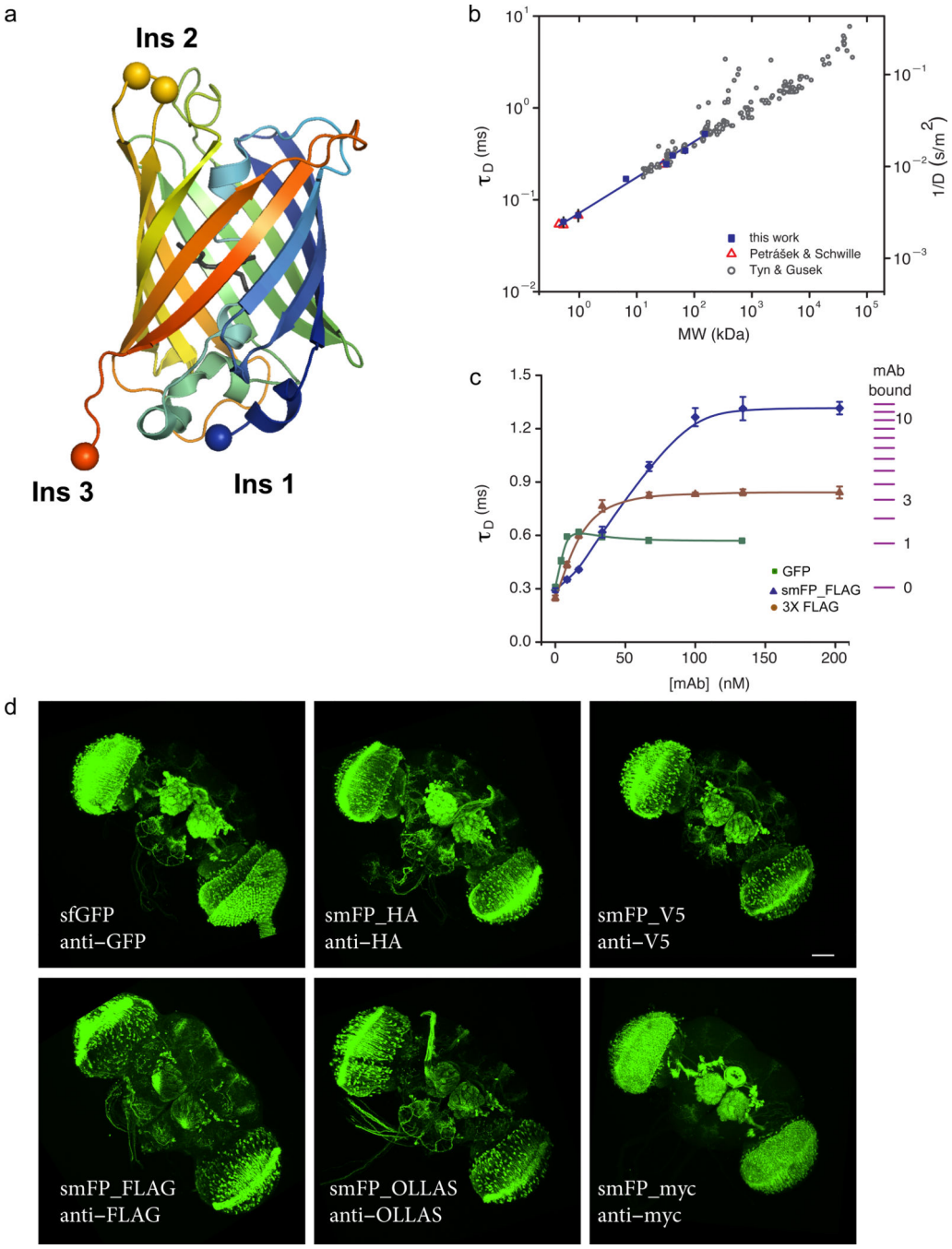

Figure 1.

Probe development and preliminary characterization. (**a**) Schematic of modular smFP construction. sfGFP shown in cartoon; chromophore either left intact or rendered invisible (grey). Epitope tags are inserted at the N- (Ins 1, blue sphere) and C- (Ins 3, red) termini and into a loop (Ins 2, yellow). (**b–c**) Fluorescence correlation spectroscopy (FCS) measurements to quantify antibody binding to FLAG epitopes of smFP, based on changes in diffusion with molecular weight. (**b**) Calibration of solution-phase diffusion time τ_D and diffusion coefficient D versus molecular weight (MW) measured by FCS. Molecular weight

markers were: hydrolyzed Alexa488, 534 Da; hydrolyzed Alexa546, 963 Da; epidermal growth factor (EGF)-FITC, 6.5 kDa; EGFP, 32.7 kDa; smFP_FLAG_bright, 42.3 kDa; bovine serum albumin (BSA)-Alexa488, 69 kDa; anti-FLAG M2-FITC, 153 kDa. n = 5 experimental replicates; mean ± SD shown. Data was fit to a power-law: $\tau_D = (0.072 \pm 0.006) \times MW^{(0.39\pm0.02)}$ (± SEM, $R^2 = 0.99$). Diffusion time and diffusion coefficient are related by $D = w_o^2/8\tau_D$, where w_o is the e^{-2} laser-beam radius, here w_o = 430 nm at 940 nm excitation. Also shown are literature values of diffusion coefficients of proteins[47], and dyes and proteins measured using scanning FCS[48]. **c**) FCS-determined diffusion times of 10 nM smFP_FLAG_bright (with 10 FLAG copies per protein yielding 100 nM of FLAG epitopes) titrated with monoclonal antibodies against GFP (Invitrogen rabbit monoclonal) or FLAG (Sigma M2), which shows saturated binding above 100 nM antibody; and 10 nM smFP_3xFLAG_bright (30 nM epitopes) with anti-FLAG antibody, which shows saturated binding above 30 nM antibody. n = 5 experimental replicates for titrations; mean ± SD shown. Ladder at right is the calculated number of antibodies bound to smFP (42.3 kDa) for the corresponding τ_D, based on the calibration obtained in **c**). **d**) Fly brains showing expression and staining of smFP probes. R59A05-GAL4[49] crossed with UAS-myr-smFP flies, with myr-sfGFP as a control. All probes, including sfGFP, lack chromophores. Dissected fly brains were stained with anti-tag antibodies or anti-GFP and Alexa488-conjugated secondaries, in small volumes (~10 μl); slight differences in staining are likely due to variation in antibody penetration. Scale bar, 50 μm.

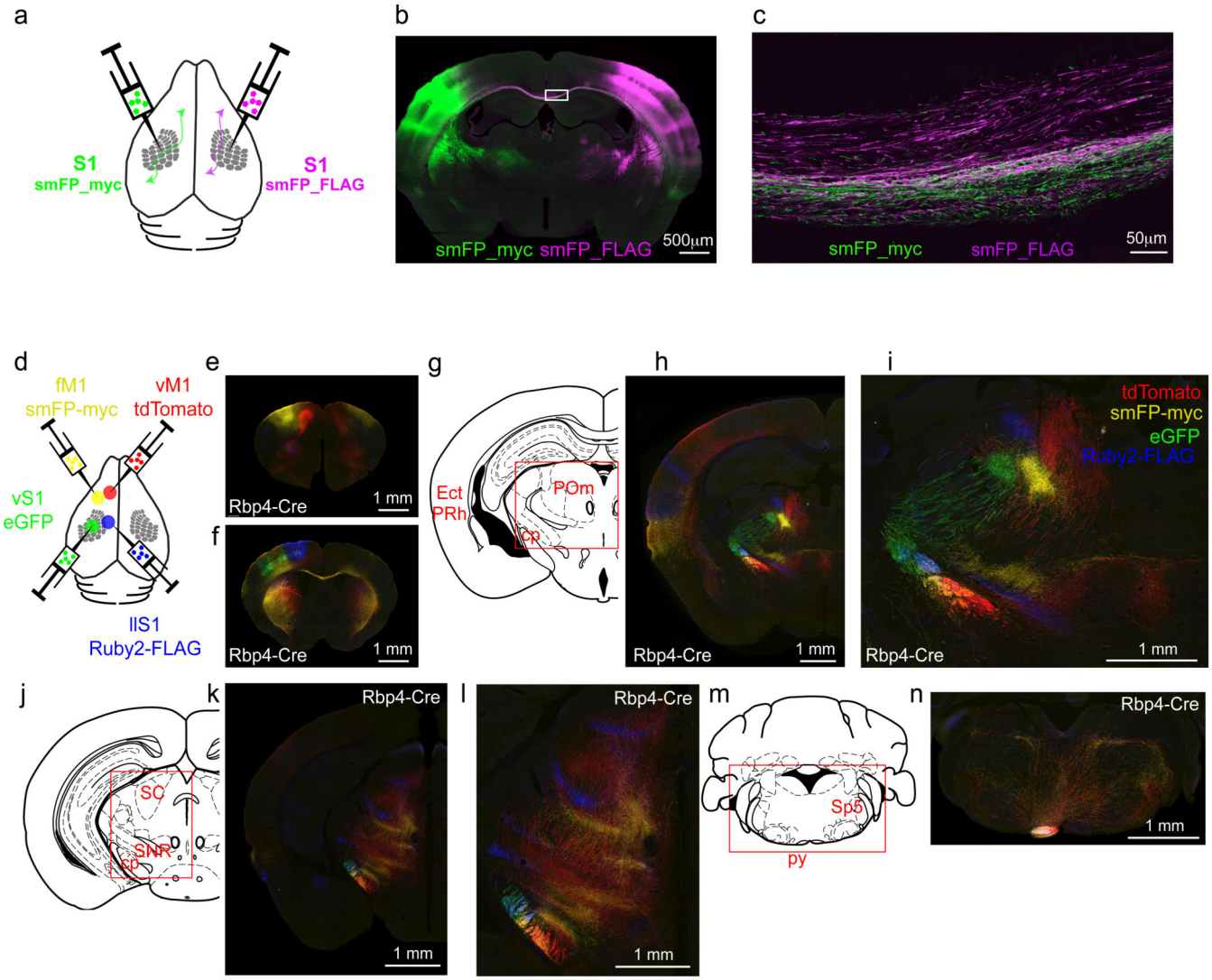

Figure 2.

Multi-channel projection labeling with smFP probes and FPs. (**a**) Injection schematic. Right hemisphere injected with smFP_FLAG and left hemisphere injected with smFP_myc. smFP_FLAG and smFP_myc are detected with tag-specific primary antibodies and corresponding secondary antibodies conjugated with Alexa488 (myc) and Alexa594 (FLAG). (**b**) Representative (n = 2 replicates) injection site image. (**c**) Confocal image showing a zoomed view of boxed area in panel **b**; long-range axonal projections from S1 in both hemispheres crossing the corpus callosum. (**d**) Injection schematic for the 4-color tracing. Left hemisphere of *Rbp4_KL100_Cre* mouse injected (all AAV-FLEX-*CAG*) with two fluorescent proteins (tdTomato and GFP) as well as two smFP constructs (smFP_myc and Ruby2_FLAG). Injections targeted four topographic areas: vibrissal sensory (vS1) and vibrissal motor cortex (vM1) as well as limb motor (fM1) and sensory areas (llS1) as indicated. (**e–n**) Schematics and images of 80 μm coronal sections. (**e–f**) Injection sites. Fluorescent neuronal somata in L5 at injection sites in fM1/vM1 (**h**) and vS1/llS1 (**i**) are clearly visible, as are long-range axonal projections. Schematics of long-range targets of L5

neurons are provided (**g, j, m**) alongside images of coronal sections from these planes. Red boxes indicate enlarged areas. (**h, i**) show cortico-cortical and cortico-thalamic axons in ectorhinal (Ect) and perirhinal (PRh) cortex as well as thalamus (POm). These descending axons continue in adjacent tracts to the cortical peduncle (cp; **h–l**) and target midbrain regions including *substantia nigra reticulata* (SNR) and superior colliculus (SC; **k, l**), as well as the spinal trigeminal nucleus (Sp5; **m, n**), >5 mm from the injection site. Intermingling of axons of all four colors in the pyramidal tract (py) results in the bright white color in **n**. Drawing are adapted with permission from Paxinos and Franklin, Mouse Brain in Stereotaxic Coordinates, 2nd Ed. (2001).

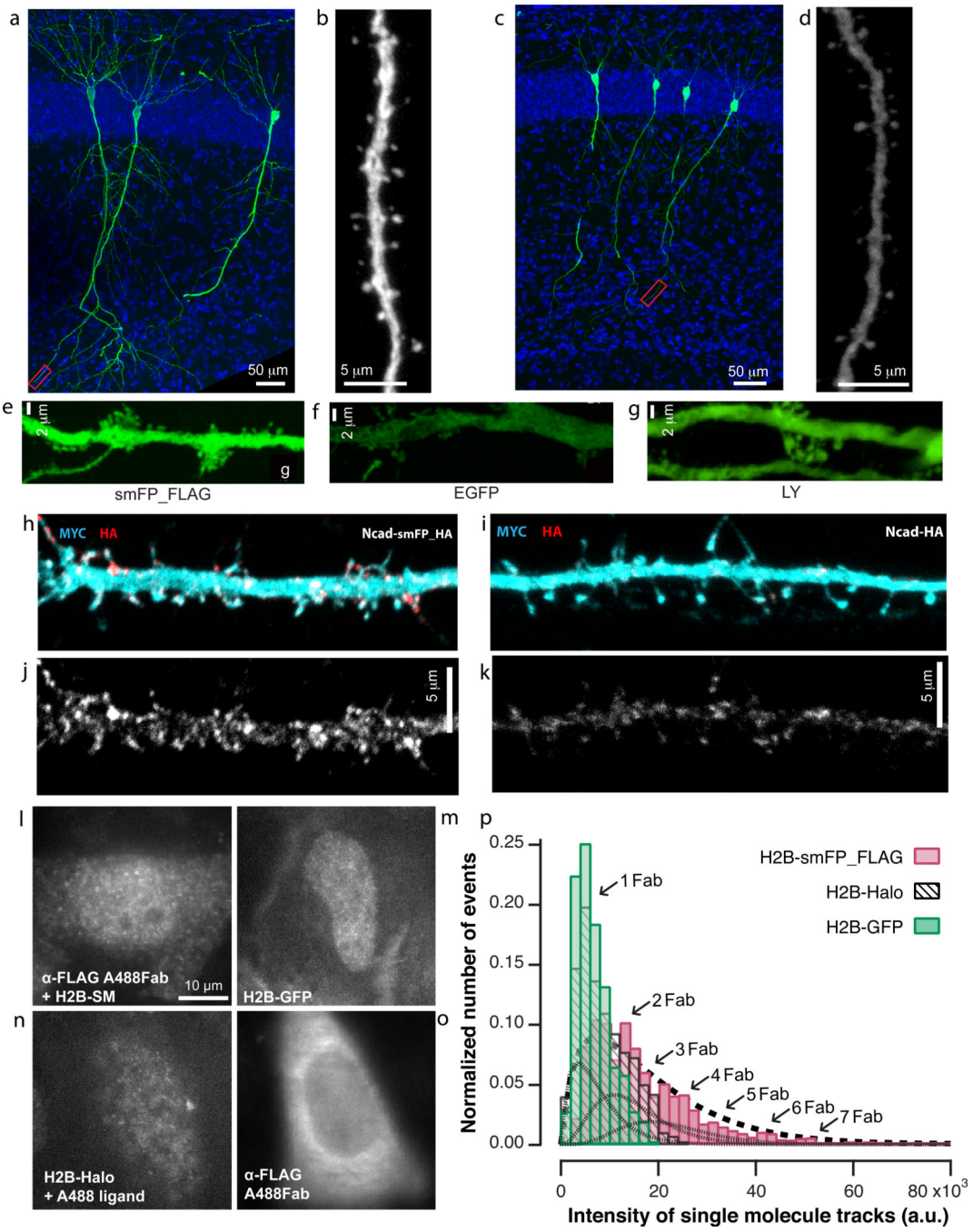

Figure 3.

Improved labeling of cells and single molecule tracking efficiency of proteins in fixed and live preparations. Comparison of smFP_FLAG (**a, b**) and EGFP (**c, d**) labeling of CA1 hippocampal pyramidal cells. (**a, c**) Fields of view showing sparsely labeled cells, from apical dendrites through soma to distal dendrites. (**b, d**) Zoom of distal dendrite in boxed areas shown in **a** and **c**. All images were taken under identical confocal settings. Blue shows Hoechst nuclear counterstain. (**e**–**g**) Images of CA3 “thorns” in P15 mouse brain slices, expressing either smFP_FLAG (**e**) or EGFP (**f**) or filled with Lucifer Yellow (LY) (**g**).

Samples were amplified with anti-FLAG, anti-GFP or anti-LY primary antibodies and secondary antibodies conjugated to Alexa488. (**h–k**) Single dendrites of cultured rat hippocampal neurons co-electroporated with smFP_myc and either N-cadherin-smFP_HA (**h, j**) or N-cadherin-HA (**i, k**). The smFP_myc fills cells cytoplasmically and serves as a control for normalizing expression levels. Blue: anti-myc primary + secondary-Alexa488, pseudocolored blue. Red: anti-HA primary + secondary-Alexa555. **h, i**: combined imaging channels; **j, k**: N-cadherin channel alone. (**l–o**) HeLa cells with tracked labeled histone molecules. All images were acquired and displayed with identical parameters. (**l**) H2B-smFP_FLAG transfected cell with Alexa488-labeled anti-FLAG antibody Fab fragments. (**m**) H2B-EGFP transfected cell with Alexa488-labeled anti-GFP antibody. (**n**) H2B-Halo tag transfected cell with Alexa488 Halo substrate added. (**o**) Background labeling of anti-FLAG antibody. Alexa488-labeled anti-FLAG antibody added to untransfected cells. (**p**) Histogram of the intensity of single molecule tracks from EGFP (green), smFP_FLAG (magenta) and Halo tag (black). Arrowheads point to the smFP_FLAG histogram peaks (intensities ~7k; 14k; 21k; 28k; 35k; 42k; 49k) that give an estimate of the number of bound Fab molecules per tag. The smFP_FLAG histogram is fit by

$$f(I)=\sum_{n=1}^{10} C_n^{10} p^n (1-p)^{10-n} \frac{I^{2n-1} a^{2n}}{(2n-1)!} e^{-aI}$$

with *a* determined from the fit to the Halo tag histogram ($n = 1$), and with $p = 0.2$, consistent with stochastic antibody binding with an average Fab occupancy of 20% (dashed curve). The three dotted curves depict the first three components of f(I). Movies corresponding to **l–o** are provided in Supp. Movie 1.

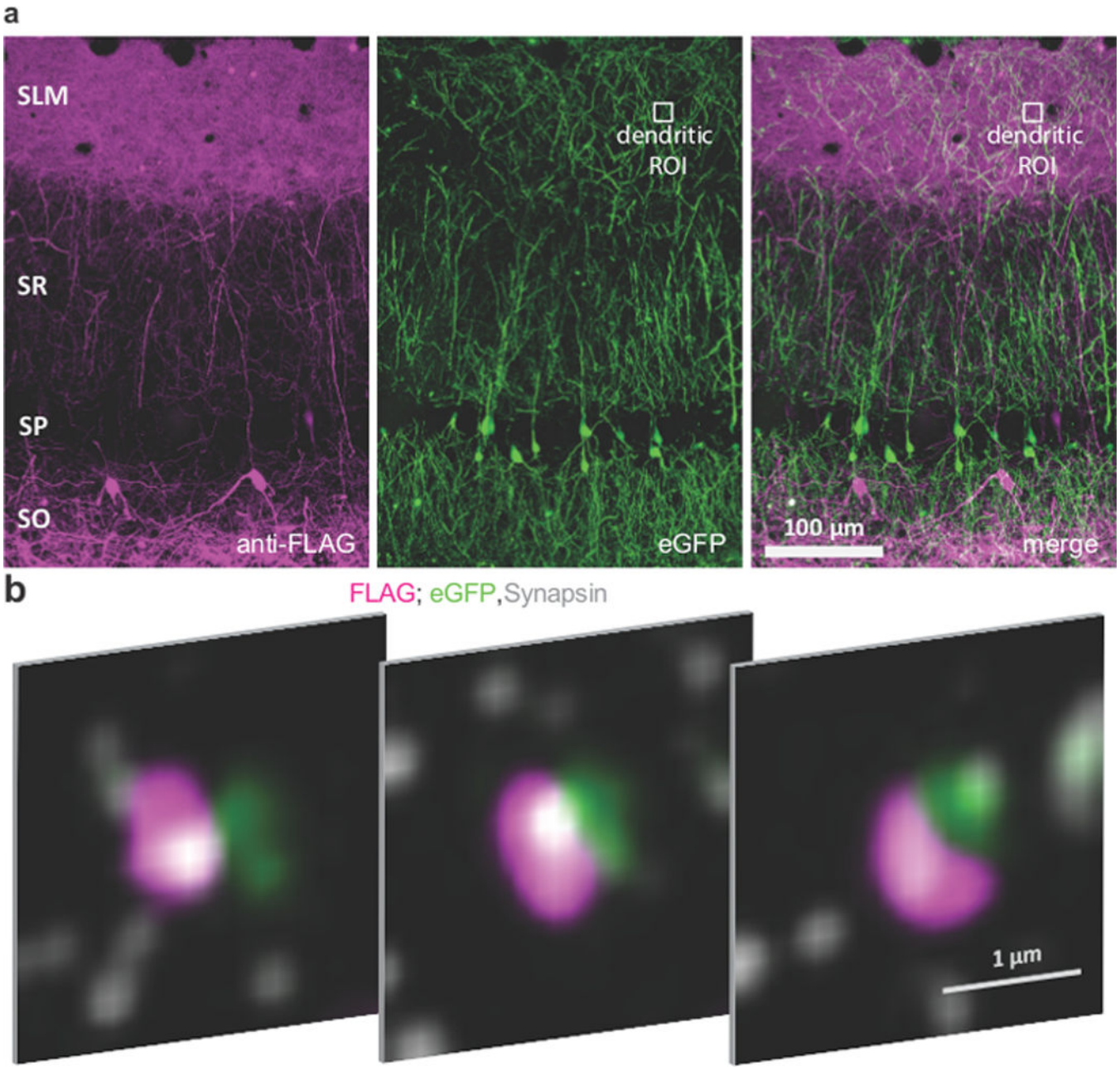

Figure 4.

Utility of smFP_FLAG for array tomography. (**a**) Left: large-scale AT z-projections showing smFP_FLAG labeling of somatostatin-Cre interneurons with cell bodies in CA1 hippocampus *stratum oriens* (SO) and axonal projections in *stratum lacunosum moleculare* (SLM). Center: EGFP+ pyramidal cell signal from the same volume. Right: merged projection. (**b**) A series of 3 consecutive ultra-thin sections (100 nm) showing an smFP_FLAG-labeled varicosity (magenta) forming a putative synapsin (gray) immunopositive synapse onto a distal pyramidal cell dendrite (green) in SLM. Note the labeling consistency across serial sections for each channel. Bottom: zoomed images

showing each channel from section two (white box), with arrows showing pixels overlaying between the synapsin and smFP_FLAG channels.

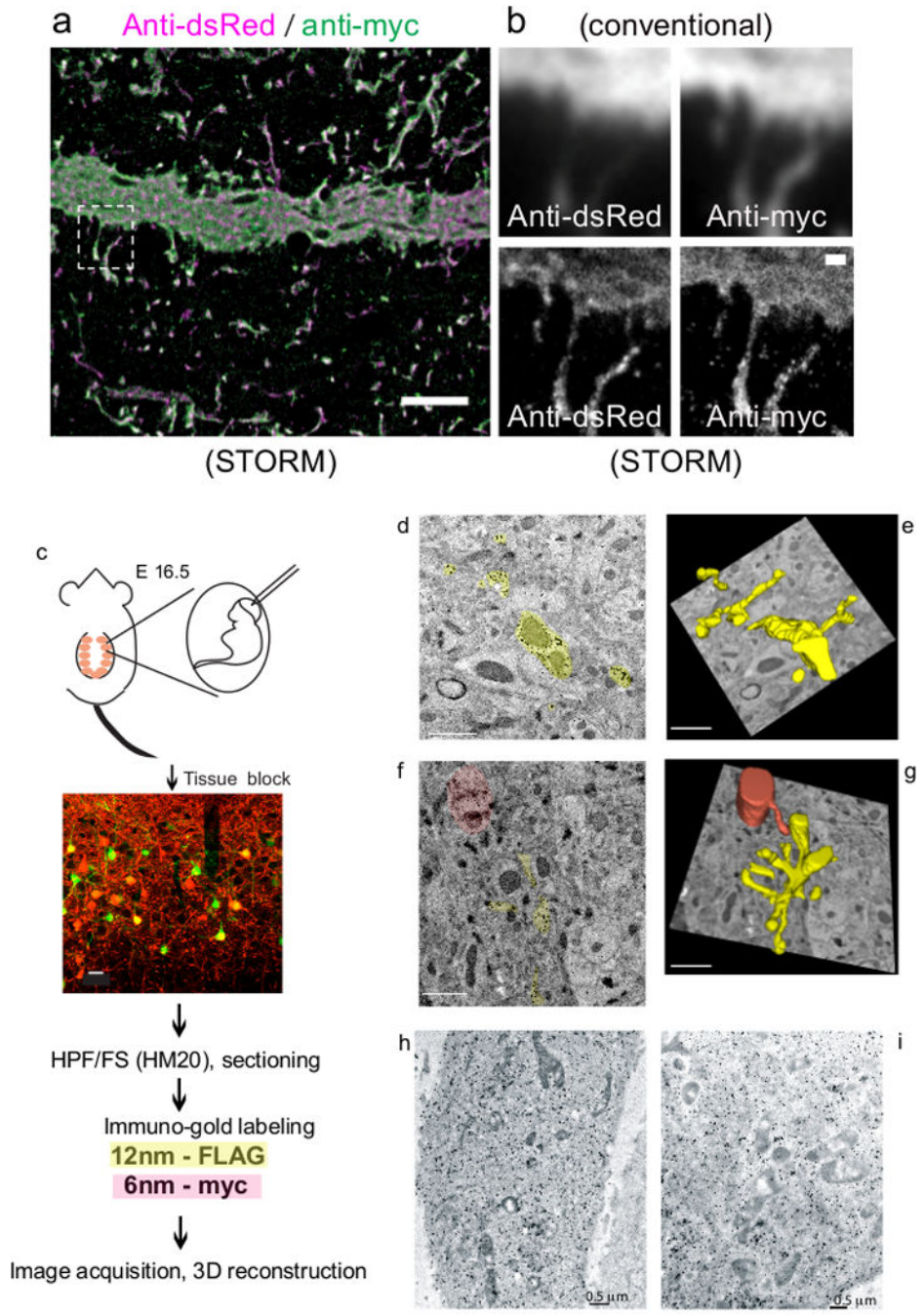

Figure 5.

Utility of smFPs for STORM and immuno EM. (**a–d**) Multi-color STORM imaging of smFPs in mouse brain slice. (**a**) STORM image of IHC-labeled neurites expressing tdTomato and smFP_myc constructs from an ultracryosection of mouse cortex. (**b**) smFP_myc and tdTomato yield comparable STORM images of individual dendritic spines on this layer V pyramidal neuron (bottom panels correspond to dashed boxed region in **a**. The diffraction-limited conventional image counterparts are shown on the top for comparison. Scale bars: **a**, 5 µm; **b**, 500 nm. (**c**) Schematic of the double immunogold

labeling experiments in mouse brain tissue. Confocal image shows a representative vibratome section expressing green smFP_FLAG_bright fluorescence and immunolabeled with anti-myc (red) to verify expression of both constructs (scale bar 20 μm). smFPs were double-labeled with 6 nm (myc) and 12 nm (FLAG) colloidal gold particles followed by silver enhancement. HPF-FS, high-pressure freezing followed by freeze substitution. (**d–e**) smFP_FLAG immunogold labeling (yellow pseudocolor) and the resulting serial reconstruction (41 serial sections) of dendrites from a layer 2/3 cortical neuron. Scale bar 1 μm. (**f–g**) Double immuno EM labeling and the resulting serial reconstruction of dendrites from two layer 2/3 cortical neurons (30 serial sections). smFP_FLAG pseudocolored yellow, smFP_myc pseudocolored red. Note high density of stain in all images, even in spines, with minimal background labeling. Scale bar 1 μm. (**h–i**) Silver enhanced anti-HA immunogold labeling in HeLa cells: (**h**) 1% OsO_4, (**i**) No secondary fix. Notice the improved structure preservation with 1% OsO_4, with no observable difference in label density. Movies showing the 3D reconstruction of the data in **d–g** are provided in Supp. Movies 2–5.