Circular permuted red fluorescent proteins and calcium ion indicators based on mCherry

Haley J.Carlson and Robert E.Campbell

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

1To whom correspondence should be addressed.
E-mail: robert.e.campbell@ualberta.ca

Received June 15, 2013; revised September 4, 2013; accepted September 22, 2013

Red fluorescent indicators for calcium ion (Ca$^{2+}$) are preferable, relative to blue-shifted alternatives, for biological imaging applications due to the lower phototoxicity, lower autofluorescent background and deeper tissue penetration associated with longer wavelength light. Accordingly, we undertook the development of a genetically encoded Ca$^{2+}$ indicator based on the popular and widely utilized Discosoma-derived red fluorescent protein, mCherry. Starting from a promising but dimly fluorescent circular permuted variant of mCherry, we first engineered a 13-fold brighter variant (cp196V1.2) through directed evolution. This bright cp196V1.2 was then used as the scaffold for creation of eight distinct libraries of potential Ca$^{2+}$ indicators via permutation at different sites within the 7th and 10th β-strands, and fusion of calmodulin and M13 to the new termini. Screening of these libraries led to the conclusion that, consistent with previous investigations of homologous fluorescent proteins, the 146–145 site in β-strand 7 is the most promising permutation site for construction of useful Ca$^{2+}$ indicators. Further rounds of directed evolution ultimately led to an indicator that exhibits a 250% change in intrinsic brightness in response to Ca$^{2+}$ and an exceptionally high affinity ($K_d = 6$ nM) for Ca$^{2+}$.

Keywords: calcium ion/calmodulin/circular permutation/fluorescent protein/imaging

Introduction

As a universal second messenger, the calcium ion (Ca$^{2+}$) is a central player in numerous intracellular signaling pathways. Accordingly, there is a persistent need to develop improved methods for detection and imaging of Ca$^{2+}$ concentrations in the context of single cells and tissues. Due to its high sensitivity and high spatial and temporal resolution, fluorescence-based detection using Ca$^{2+}$ responsive synthetic dyes has traditionally been the most important strategy for detecting and imaging of intracellular Ca$^{2+}$ dynamics (Tsien, 1999). With the advent of the Aequorea green fluorescent protein (GFP) and its numerous variants and homologs, synthetic Ca$^{2+}$ dyes have been gradually displaced by protein-based indicators that provide the added advantage of being genetically encodable (Miyawaki et al., 1997; Nagai et al., 2001; Nakai et al., 2001). Genetically encodable indicators are minimally invasive, amenable to subcellular targeting, and can be used for the creation of transgenic animals with tissue-specific expression of the indicator.

From a protein engineering perspective, the challenge has been to develop protein-based indicators that meet or surpass the performance characteristics of synthetic dye-based indicators while providing a similar diversity of color choices. One desirable feature of fluorescent proteins (FPs) for live cell imaging, whether they are for imaging of fusion protein localization or imaging of Ca$^{2+}$ dynamics with an engineered indicator, is red-shifted excitation and fluorescence emission. Generally speaking, more red-shifted fluorophores are associated with less phototoxic excitation light, less autofluorescent background and deeper tissue penetration (Shu et al., 2009). An additional benefit of red fluorescent probes is that they are spectrally orthogonal to blue light-excitable chromophores, and can therefore be used in conjunction with GFP-based probes for multi-parameter imaging (Carlson and Campbell, 2009) or with various optogenetic actuators for manipulation of cell function (Alford et al., 2013).

A proven strategy for the creation of single FP-based Ca$^{2+}$ indicators is to genetically fuse calmodulin (CaM), and the Ca$^{2+}$–CaM-interacting peptide M13, to the N- and C-termini of a circular permuted (cp) FP in which the new termini are located in close proximity to the chromophore. Indeed, this design was first applied to Aequorea GFP-derived enhanced GFP (Nagai et al., 2001) and enhanced yellow fluorescent protein (Nagai et al., 2001) in 2001 and has since been applied to several FP variants from different species (Zhao et al., 2011; Akerboom et al., 2013; Hoi et al., 2013). Accordingly, there is now a substantial and growing selection of FP-based Ca$^{2+}$ indicators with fluorescent hues spanning the visible spectrum (Wu et al., 2013). In all examples reported to date, the new termini are located in the vicinity of residue 145 (GFP numbering) which is located in β-strand 7 in close proximity to the chromophore (Fig. 1A).

Included in the currently available selection of FP-based Ca$^{2+}$ indicators are red fluorescent variants derived from mApple (Zhao et al., 2011; Wu et al., 2013) and mRuby (Akerboom et al., 2013). To date there have been no reports of a Ca$^{2+}$ indicator based on the Discosoma-derived red FP mCherry, which remains one of the most widely used red fluorescent proteins (RFPs) due to its numerous favorable properties which include: red-shifted emission, monomeric structure, fast maturation, good photostability and relative insensitivity to physiological pH changes (Shaner et al., 2004). Some years ago, we initiated an effort to engineer an mCherry-based Ca$^{2+}$ indicator. As a first step towards this goal, we determined locations within the mCherry structure that could tolerate peptide insertion (Li et al., 2008). However, cp mCherry variants with new termini at these same locations generally exhibited poor folding efficiency in Escherichia coli and were effectively insoluble (Li et al., 2008). Based on this work, we later engineered a cp variant of mCherry (designated...
cp193g7) in which the original C- and N-termini are linked by a six amino acid linker (GAGGGS) and new termini have been introduced at position 193 (Carlson et al., 2010) (Fig. 1A). Relative to mCherry, the cp193g7 variant also has six point mutations that were accumulated during directed evolution for improved fluorescent brightness. The cp193g7 variant has 60% of the intrinsic brightness (i.e. the product of the quantum yield ($\Phi$) and the extinction coefficient ($e$)) of mCherry. Unfortunately, it is not a particularly promising candidate for conversion into a Ca$^{2+}$ indicator since the 193 position is relatively far from the chromophore at the N-terminal end of $\beta$-strand 10. Further exploration of circular permutation sites within $\beta$-strand 10 at positions closer to the chromophore led to the discovery of the dimly fluorescent cp196V1.0 variant (Fig. 1B and Supplementary Fig. S1) (Carlson et al., 2010). A number of cp mCherry variants have also been reported by Shui et al., 2011 although none of those variants have improved brightness or expression relative to cp193g7.

Here, we describe our subsequent and ultimately successful efforts to develop an mCherry-based Ca$^{2+}$ indicator. Despite our best efforts to develop a Ca$^{2+}$ indicator in which the new termini were located in $\beta$-strand 10, the design which ended up being successful had the new termini located in $\beta$-strand 7.

Materials and methods

General

Gel extraction kits for DNA fragment purification and restriction enzymes were purchased from Fermentas. Plasmid miniprep kits were obtained from Fermentas or Epoch Life Sciences. Taq DNA polymerase (New England Biolabs), Pfu polymerase (Fermentas) and T4 DNA ligase (Invitrogen) were used according to the manufacturers protocol unless otherwise noted. DNA sequencing was done using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosciences) and analyzed by the University of Alberta, Molecular Biology Services Unit. All E. coli transformations were performed by electroporation.

Directed evolution of cp RFPs

Error prone polymerase chain reaction (PCR) was carried out using Taq polymerase in the presence of MnCl$_2$ and deoxycytidine triphosphate or deoxythymidine triphosphate deficiency (Cadwell and Joyce, 1992). Staggered extension process (StEP) shuffling was performed as previously described (Zhao and Zha, 2006). Primers for saturation mutagenesis contained the degenerate codon NNK at the target amino acid position. Briefly, overlap extension PCR was used to generate two DNA fragments, which were recombined using the 5’ and 3’ PCR primers. All gene libraries were doubly digested with XhoI and EcoRI and ligated into similarly digested pBAD/His B plasmid (Invitrogen).

Plasmid libraries were expressed in DH10B cells on Luria–Bertani (LB) agar plates supplemented with ampicillin and 0.2% arabinose. The colonies were screened with a home-built colony screening system that has been previously described (Ai et al., 2006). Plates were illuminated with 535–550 nm light and emission collected with a CCD camera equipped with a 630–660 nm emission filter. If no red fluorescence developed overnight, the plates were left on the bench top for several days and rescreened. The brightest red colonies were picked and grown overnight in 5 ml LB culture tubes (37°C, 250 rpm). The plasmid DNA was isolated, pooled and used as the template DNA for subsequent error prone PCR or StEP shuffling rounds. Colony brightness was measured by acquiring and analyzing fluorescence images with the home-built imaging system. Briefly, an Image Pro (Media Cybernetics) macro automatically identifies colonies in a background corrected fluorescent image of the plate and exports the brightness and size of each colony to spreadsheet. Percentage brightness was calculated relative to mCherry colonies, which were used as a control and imaged on plates concurrently with the cp RFPs.

Circularly permuted RFP purification and characterization

A 5 ml overnight starter culture expressing the gene for the protein of interest was used to inoculate 500 ml of LB-amp and the flasks were shaken for 3–4 h until the optical density was 0.6. Arabinose was then added to 0.02% and the culture was shook for another 16 h. The next day the bacteria were pelleted by centrifugation at 10 000 rpm for 10 min and resuspended in phosphate-buffered saline (PBS) before being lysed by a cell disrupter (Constant Systems). The lysates were then centrifuged at 14 000 g for 45 min, and the proteins were...
purified from the supernatants using Ni-NTA chromatography (Amersham). Proteins were eluted with 300 mM imidazole, which was then removed by buffer exchange with 10 mM PBS or 10 mM Tris, pH 7.4 using Amicon centrifuge tubes (MWCO 10 000). Proteins were stored at 4°C prior to spectral characterization.

Protein fluorescence was measured using a Quantamaster spectrophotometer (PTI) and the absorbance spectra were collected on a DU800 UV-Vis spectrometer (Beckman). For determination of the quantum yield, solutions of the proteins were diluted to three absorbance values between 0.01 and 0.05 at 550 nm and prepared in triplicate. The proteins were excited at 550 ± 5 nm and the fluorescence emission was integrated under the curve. Quantum yields were calculated as described by Lakowicz (1999), with mCherry used as a quantum yield standard. Extinction coefficients were determined as previously described (Carlson et al., 2010). Briefly, the protein absorbance was matched at 280 nm such that absorbance at 587 nm was between 0.2 and 0.6. Three separate absorbance values were measured in triplicate.

Construction and screening of libraries of potential Ca^{2+} indicators

The gene encoding cp196V1.2 was used as the template for construction of eight distinct libraries of potential Ca^{2+} mCherry, and also append XhoI and MluI sites. DNA fragments were purified by agarose gel electrophoresis and extracted using the GeneJet gel extraction kit. The purified DNA was digested with XhoI and MluI and ligated into a modified pBAD/His B plasmid with a TorA export sequence at the 5’ end as previously described (Zhao et al., 2011). The ligation products were used to transform E. coli, which was then plated out onto LB agar plates with 0.02% arabinose and incubated overnight at 37°C.

Plates of E. coli expressing these libraries were imaged for red fluorescence using the home-built system. If no red fluorescence was immediately apparent, the plates were stored at 4°C and checked daily. The brightest colonies were picked off of plates and used to inoculate 5 ml LB-amp-arabinose culture tubes. After overnight culturing at 37°C, an osmotic shock protocol was used to extract the periplasmic protein fraction (Zhao et al., 2011). Briefly, the bacteria was pelleted by spinning at 12 000 rpm and 4°C. The pellet was resuspended with 500 μl of pH 8.0, 30 mM Tris-Cl, 1 mM ethylenediaminetetraacetic acid, 20% sucrose and shaken on ice for 10 min. The pellet is spun down at 9000 g, 4°C and then resuspended with 500 μl of 5 mM MgSO4, 1 mM ethylene glycol tetraacetic acid (EGTA), 10 mM Tris-Cl, pH 7.4. For pellets where osmotic shock extraction was inefficient, the remaining bacterial pellet was extracted with B-PER (Pierce). For routine measurements of the Ca^{2+} response, protein extracts were aliquoted (100 μl) into three replicate wells in a 96-well microplate (Corning) along with either 1 mM EGTA or 1 mM CaCl2. The fluorescence emission was measured with 550 nm excitation on a Tecan Safire2 microplate reader.

For further evolution of the Ca^{2+} indicator, the gene was PCR amplified under error prone conditions with primers that contained an XbaI and a HindIII site. PCR products were purified and ligated into the periplasmic expression plasmid mentioned above. The ligation products were used to transform E. coli DH10B, and the colonies were screened as described above. Quikchange lightning mutagenesis (Agilent) was used to randomize the Gly–Gly linker of various targeted positions within the gene.

Ca^{2+} indicator purification and characterization

A single colony of the desired mutant was picked and grown overnight in a 5 ml LB-amp culture tube at 37°C. The 5 ml culture tube was used to inoculate 500 ml of LB-amp. After 4 h of shaking at 37°C arabinose was added to a final concentration of 0.02% and continued shaking overnight. For poorly folding mutants, the flasks were transferred to a shaker at 30°C for two nights after induction. To obtain high yields of expressed protein, a modified Terrific broth growth media was used (11 water with 20 g LB; 14 g tryptone; 7 g yeast; 9.4 g K2HPO4; 2.2 g KH2PO4; 0.8% w/v glycerol). Bacteria were pelleted at 10 000 rpm, 4°C for 10 min and the pellet was resuspended at 4°C in 10 mM Tris-Cl, 150 mM NaCl, pH 7.4 and then lysed using a cell disruptor. Proteins were purified as described above for the cp RFPs.

To determine the pHKₐ, a series of buffers was prepared by first adjusting a solution of 30 mM trisodium citrate and 30 mM sodium borate to pH 11.5. The pH of the solution was then adjusted with HCl (12 and 1 M) in 0.5 pH units and 10–15 ml was removed at each desired pH value. For each pH, a Ca^{2+}-free solution (10 mM EGTA) and a Ca^{2+}-containing (10 mM CaCl2) version was prepared. In triplicate, 5 μl of the protein solution was mixed with 50 μl of the desired pH buffer in a 396-well plate format. Fluorescence emission was recorded using a fluorescence microplate reader.

The apparent Kₐ for Ca^{2+} was determined by mixing the appropriate protein with buffers containing various amount of Ca^{2+}. The buffers were prepared following the recipe from the Invitrogen Ca^{2+} calibration buffer kits. The Ca^{2+}-free buffer (30 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 100 mM KCl, 10 mM EGTA, pH 7.2) and Ca^{2+}-saturated buffer (30 mM MOPS, 100 mM KCl, 10 mM CaEGTA, pH 7.2) were mixed in different ratios to generate buffers with Ca^{2+} concentrations ranging from 0 to 39 μM Ca^{2+}. Similar to the pH titrations, 5–10 μl of the protein was mixed with 150–200 μl of each Ca^{2+} buffer. Fifty microliters of each solution was aliquotted in triplicate into a 396-well plate and the fluorescence emission was recorded using the plate reader. Each emission peak was integrated and plotted against the log of the calculated free Ca^{2+} concentration. Each Ca^{2+} titration curve was fit with an appropriate sigmoidal curve to determine the values of the Kₐ and the Hill coefficient.

Quantum yields and extinction coefficients in the presence and the absence of Ca^{2+} were determined as described for the cp RFPs. The Ca^{2+} association kinetics of CH-GEC02.1 were determined using a SX20 stopped-flow spectrometer (Applied Photophysics). The Ca^{2+} indicator was diluted 400× in buffer (30 mM MOPS, 100 mM KCl, 1 mM EGTA, pH 7.2) and mixed (1:1) under stopped-flow conditions with a series of Ca^{2+} buffers that were prepared by mixing a buffered solution (30 mM MOPS, 100 mM KCl) with different ratios of 10 mM EGTA and 10 mM CaEGTA. The change in fluorescence signal upon mixing was used to determine the relaxation coefficient (k_obs) for the Ca^{2+} association reactions at various Ca^{2+} concentrations (700 pM to 1300 nM) by fitting the curve to a single-phase decay.

Live cell imaging

CH-GEC02.0 and CH-GEC02.1 were each ligated between the BamHI and EcoRI sites of mammalian expression plasmid
PCDNA3.1(+) Plasmids were isolated as described above, and the integrity of the inserted gene confirmed by DNA sequencing. HeLa cells were maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum and glutamax (Invitrogen) at 37°C and 5% CO2. Cells at ∼70% confluency were transfected with 750 ng of plasmid DNA using Turbofect (Invitrogen) according to the manufacturer’s protocol. On the day of imaging the complete medium was removed and the dishes were washed twice with PBS and left with 1 ml of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-buffered hanks balanced saline solution (HHBSS).

An inverted Nikon Eclipse Ti microscope equipped with a 200 W metal halide lamp (PRIOR Lumen), 60× oil objective, and a QuantEM; 512SC 16-bit cooled CCD camera (Photometrics) was used for fluorescence microscopy. The protocol for Ca2+ imaging was adapted from the literature (Palmer and Tsien, 2006). Briefly, the rinsed dishes containing the transfected cells were imaged using 300 ms exposure every 30 s for 2.5 min to obtain the baseline fluorescence, after which 1 ml of 5 mM histamine was added and images were captured every 5 s for 12 min. After 12 min, the dish was rinsed three times with HHBSS and images were taken every 30 s for 1.5 min to reestablish a baseline signal. Cells were washed three times with Ca2+-free HHBSS and 1 ml of 1 mM EGTA with 5 μM ionomycin was added and images are captured every 10 s for 5 min. Finally, the cells were rinsed three times again with HHBSS before 1 ml of 2 mM Ca2+ with 5 μM ionomycin was added and images were taken every 30 s for 2.5 min.

Results

Directed evolution of cp196V1.0 for brighter fluorescence

In an effort to improve the fluorescent brightness of the dimly fluorescent cp196V1.0, we undertook a process of directed evolution using fluorescence-based screening of randomly mutated gene libraries expressed in E. coli colonies. At both ends of the gene, a Gly–Gly dipeptide was genetically added as an eventual linker region to M13 and CaM, which would each later be fused to the cp FP gene to create the Ca2+ indicator-type construct. Colonies expressing those members of the gene library that exhibited the brightest red fluorescence were picked, cultured and the plasmid DNA used as the template for the subsequent round of library creation and screening. Library creation by error prone PCR (Cirino et al., 2003) and recombination of improved variants by the StEP method (Zhao and Zha, 2006) were used in alternating rounds. In each round, we screened 2000–4000 colonies, of which ∼100 were picked, cultured overnight and subjected to osmotic shock or detergent extraction to provide soluble protein. The fluorescence intensity both with and without Ca2+ was then determined using a multi-well plate reader and the top variants (<10) were used as the template for the subsequent round of library creation and screening. The brightest variant after four rounds, designated cp196V1.1, contained mutations Met150Leu, Tyr193His, Gly219Val, Thr223Ile and a Gly–Val mutation in the first of two Gly residues that had been appended to the N-terminus of the protein (Fig. 1B and Supplementary Fig. S1). Whereas the colonies expressing the V1.0 versions of cp196 was only 2% as bright as mCherry after 3 days at 4°C, the V1.1 versions was 81% as bright. As summarized in Table I, the quantum yield of cp196V1.1 remained similar to cp196V1.0 while the extinction coefficient increased 10-fold from 4600 to 46 500 M⁻¹ cm⁻¹, likely due to improved folding and maturation.

Sequencing of additional improved variants isolated during directed evolution led us to identify six additional sites where beneficial mutations had occurred: Ser147, Met150, Gln188, Tyr193, Asn194 and His221. Each of these positions was subjected to site-directed saturation mutagenesis to create a small library, which was then screened in colony format for variants with improved brightness. This iterative procedure eventually led us to cp196V1.2 with three new mutations: Ser147Ile, Gln188Ser and a Gly–Cys mutation (likely introduced during PCR amplification) in the second of the two Gly residues appended to the N-terminus of the protein (Fig. 1B and Supplementary Fig. S1). Compared with cp196V1.1 the quantum yield of cp196V1.2 increased from 0.14 to 0.21, while the extinction coefficient improved from 46 500 to 49 000 M⁻¹ cm⁻¹ (Table I). The colony brightness relative to mCherry after 24 h at 37°C also improved significantly from 8 to 40%, suggesting that the efficiency of protein folding at physiological temperature was substantially improved.

Generation of potential Ca2+ indicators

Using cp196V1.2 as template, a series of potential Ca2+ indicators were generated at various positions within the β-barrel (Fig. 2). Positions 197–200 were chosen in β-strand 10 because they are closest to the chromophore. We have previously reported that permutation at these sites was not tolerated in the context of the cp196V1.0 backbone (Carlson et al., 2010). However, we hoped that the improved folding efficiency of cp196V1.2 would allow insertion of termini closer to improved folding and maturation. Table I. Spectral characteristics of cp196 variants

<table>
<thead>
<tr>
<th>Variant</th>
<th>Φ</th>
<th>e (M⁻¹ cm⁻¹)</th>
<th>Intrinsic brightness (mM⁻¹ cm⁻¹)</th>
<th>Relative brightness in E. coli (24 h at 37°C)</th>
<th>Relative brightness in E. coli (72 h at 4°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCherry</td>
<td>0.22</td>
<td>72 000</td>
<td>15.8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>cp196V1.0</td>
<td>0.15</td>
<td>4600</td>
<td>0.8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>cp196V1.1</td>
<td>0.14</td>
<td>46 500</td>
<td>6.5</td>
<td>8</td>
<td>81</td>
</tr>
<tr>
<td>cp196V1.2</td>
<td>0.21</td>
<td>49 500</td>
<td>10.4</td>
<td>40</td>
<td>98</td>
</tr>
</tbody>
</table>

Fig. 2. Circular permutation sites in cp196V1.2 used for the construction of eight different libraries of potential Ca2+ indicators. For each insertion site, M13 was fused to the N-terminus, CaM was fused to the C-terminus, and the N- and C-terminal amino acids of the cp FP were randomized to create a library.
to the chromophore. Sites 144–147 in β-strand 7 were also chosen, as these approximate the split points in the Aequorea GFP-derived indicators (Nagai et al., 2001; Nakai et al., 2001) and the mApple-based indicators (Zhao et al., 2011). For each insertion site, the Gly–Gly linkers at the M13-cp FP and the cp FP–CaM junctions were retained. In addition, for each insertion the first and last amino acids of the cp FP (i.e. the two amino acids in closest proximity to the split point) were randomized to create a library of variants. A similar strategy was previously used to identify cp196V1.0 (Carlson et al., 2010).

The libraries were expressed in bacterial colonies using the TorA export system, as previously reported for the screening of the GECO series of Ca$^{2+}$ indicators (Zhao et al., 2011). Since the periplasmic Ca$^{2+}$ concentration (>100 μM) is much higher than the expected Kd of these indicators (~0.1–1 μM), the potential indicators are expected to be in the Ca$^{2+}$-bound state (generally the bright state) rather than the Ca$^{2+}$-free state. After imaging the colony libraries on LB-agar-arabinose plates, bright red fluorescent colonies were picked and cultured overnight. Proteins were extracted using non-denaturing detergents or osmotic shock and fluorescence was measured in the presence of both 1 mM EGTA and 1 mM Ca$^{2+}$ in microplate format. Percentage response was calculated by using the equation:

$$\% \text{Change} = \left( \frac{F_{Ca^{2+}}}{F_{EGTA}} - 1 \right) \times 100$$

where $F_{Ca^{2+}}$ and $F_{EGTA}$ are the peak fluorescence intensities in the presence and the absence of Ca$^{2+}$, respectively.

It is reported that the 198–197 permutation site is tolerated in mCherry, while the 199–198 site is not (Shui et al., 2011). We did observe some 199–198 variants with dim fluorescence, likely due to the improved folding efficiency of our starting template. In cases where few or no fluorescent colonies were observed, non-fluorescent colonies were picked and the plasmid DNA subjected to sequencing to confirm the integrity of the library. In contrast to our poor success rate in mCherry, while the 199–198 site is not (Shui et al., 2011), in mCherry, the 199–198 site is not (Shui et al., 2011). We did observe some 199–198 variants with dim fluorescence upon binding Ca$^{2+}$.

Interestingly, the 146–145 site that gave the highest proportion of behaviors including: fluorescence decreases upon binding Ca$^{2+}$, fluorescence increases upon binding Ca$^{2+}$, no change in fluorescence upon binding Ca$^{2+}$ and even no fluorescence when expressed in culture (Supplementary Table S1). Interestingly, the 146–145 site that gave the highest proportion of red fluorescent colonies also tended to result in the Ca$^{2+}$ indicators with the largest fluorescence changes (up to a %change of +72%). Of the relatively few red fluorescent variants with new termini in β-strand 10, only the 199–198 variants gave a modest inverse response to Ca$^{2+}$ (up to a %change of −20%).

Given the fact that our best mCherry-based Ca$^{2+}$ indicator had new termini in β-strand 7, plus the ample precedent that effective Ca$^{2+}$ indicators could be created from permutations at analogous sites of other FP variants (Nagai et al., 2001; Nakai et al., 2001; Zhao et al., 2011; Akerboom et al., 2013; Hoi et al., 2013), we decided to narrow the focus of our efforts to only the most promising variants. Specifically, we designated the variant in which the FP domain started with residue S146L and ended with residue A145L as mCherry-derived genetically encoded Ca$^{2+}$ indicator for optical imaging version 1.0 (CH-GECO1.0) and used it as the template for further optimization.

**Directed evolution and optimization of CH-GECO1.0**

In an effort to improve the performance of CH-GECO1.0, we used its gene as the template for library generation by error prone PCR, and screened the resulting library for improved brightness and/or response to Ca$^{2+}$ as described above. The most improved variant, designated CH-GECO1.1, identified after one round of library generation and screening exhibited similar colony brightness to CH-GECO1.0, had an improved response of 100%, and contained the single mutation Asp23Ala in the CaM domain (Table II and Supplementary Fig. S2). Notably, Asp23 one of the coordinating groups that bind Ca$^{2+}$ in the first EF hand of CaM (Ikura et al., 1992).

As additional rounds of error prone PCR did not produce any further improved variants, we turned our focus to the Gly–Gly linker that connects M13 and the cp FP. As revealed in the crystal structures of the Ca$^{2+}$-bound states of GCaMP2 (Wang et al., 2008; Akerboom et al., 2009), RCaMP (Akerboom et al., 2013), and R-GECO1 (Akerboom et al., 2013), this linker (Leu–Glu in GCaMP2, Pro–Val–Val in R-GECO1 and Ala–Ile in RCaMP) plays important roles in the fluorescence response. To explore sequence diversity at this linker, we employed Quickchange mutagenesis to mutate both residues of the Gly–Gly linker to a subset of alternative residues (i.e. Thr, Ser, Ala, Val, Pro, Leu, Ile, Phe and Met). Following extensive library screening, one particularly bright colony (3.5-fold brighter than colonies expressing CH-GECO1.1 after 24 h) was identified. The corresponding variant, designated CH-GECO2.0, had a response of 170% to Ca$^{2+}$ and, upon sequencing, was revealed to be an unintentional genetic hybrid of R-GECO1 and CH-GECO1.1 (Table II and Supplementary Fig. S2). This gene hybrid appears to have resulted from seven template switching events during PCR (Supplementary Fig. S3). Starting from the N-terminus, there is a region of M13 from R-GECO1, then a region from CH-GECO1.1 (up to ~166), then a region from R-GECO1 (from ~174 to the cp linker), then a region from CH-GECO1.1 (from ~1 to ~41), then a region from R-GECO1 (from ~65 to ~73), then a region from CH-GECO1.1 (~83 through to ~23 of CaM), then a region of R-GECO1 (to ~61 of CaM) and finally a region from CH-GECO1.1 (~77 of CaM to the C-terminus). In addition, the original Gly–Gly linker has been mutated to Ser–Leu and there is one additional mutation, Asp132Asn, that is not present in either gene template, bringing the total number of mutations separating CH-GECO1.1 and CH-GECO2.0 to 16. Despite the numerous mutations, CH-GECO2.0 retains

**Table II. Summary of relative colony brightness and fluorescence signal change of CH-GECO variants**

<table>
<thead>
<tr>
<th>Variant</th>
<th>Relative brightness in E. coli (24 h at 37 °C)</th>
<th>Relative brightness in E. coli (72 h at 4°C)</th>
<th>%change with Ca$^{2+}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH-GECO1.0</td>
<td>4</td>
<td>25</td>
<td>72%</td>
</tr>
<tr>
<td>CH-GECO1.1</td>
<td>4</td>
<td>14</td>
<td>100%</td>
</tr>
<tr>
<td>CH-GECO2.0</td>
<td>14</td>
<td>49</td>
<td>170%</td>
</tr>
<tr>
<td>CH-GECO2.1</td>
<td>13</td>
<td>75</td>
<td>250%</td>
</tr>
</tbody>
</table>
excitation ($\lambda_{\text{ex,Ca}^2+} = 588$ nm) and emission ($\lambda_{\text{em,Ca}^2+} = 604$ nm) maxima that are very similar to those of CH-GECO1.1 ($\lambda_{\text{ex,Ca}^2+} = 584$ nm and $\lambda_{\text{em,Ca}^2+} = 604$ nm) and distinct from R-GECO1 ($\lambda_{\text{ex,Ca}^2+} = 561$ nm and $\lambda_{\text{em,Ca}^2+} = 589$ nm) (Zhao et al., 2011).

Following several additional rounds of library creation by error prone PCR and screening, we identified an improved variant, designated CH-GECO2.1, with a 250% response to Ca$^{2+}$ and five mutations compared with CH-GECO2.0 (Table II and Supplementary Fig. S2). Three of these mutations occur in the CaM domain (Asp21Gly, Phe61Leu and Thr77Ser) and two mutations are in the cp FP domain (Ile147Thr and Gly191Asp). Both Asp21Gly and Phe61Leu are in EF hands of CaM, while Thr77Ser is in a loop between the second and third EF hand. Ile147Thr is located on $\beta$-strand 7 right after the M13 to cp FP linker, while Gly191Asp is located in the loop between $\beta$-strands 9 and 10 and is relatively far from both the chromophore and the termini of the cp FP. Further screening of libraries produced by error prone PCR did not result in the identification of further improved variants.

Characterization of CH-GECO2.0 and CH-GECO2.1

In vitro characterization of CH-GECO2.0 revealed that binding to Ca$^{2+}$ causes the quantum yield to increase from 0.07 to 0.16 and the extinction coefficient to increase from 21 700 to 23 800 M$^{-1}$ cm$^{-1}$ (Table III). For CH-GECO2.1, binding to Ca$^{2+}$ causes the quantum yield to increase from 0.05 to 0.17 and the extinction coefficient to undergo a 30% increase from 37 000 to 48 000 M$^{-1}$ cm$^{-1}$. Notably, the extinction coefficient of the Ca$^{2+}$-bound state of CH-GECO2.1 is somewhat less than mCherry itself ($\Phi = 0.22; \epsilon = 72 000$ M$^{-1}$ cm$^{-1}$) (Shaner et al., 2004), but compares favorably with the cp196V1.2 template used in this work ($\Phi = 0.21; \epsilon = 49 500$ M$^{-1}$ cm$^{-1}$). Notably, the apparent increase in intrinsic brightness for CH-GECO1 is 330%, which is in poor agreement with the value of 250% provided in Table II. We attribute this discrepancy to the different buffer conditions used for routine measurements of overall Ca$^{2+}$ response and the in vitro characterization.

Typically, the mechanism of single FP-based Ca$^{2+}$ indicators is best described as either a Ca$^{2+}$-dependent shift in the effective $pK_a$ of the chromophore (i.e. the pH at which fluorescence is 50% of its maximum) (Zhao et al., 2011), or a Ca$^{2+}$-dependent change in the brightness of the chromophore (Akerboom et al., 2013) (Fig. 3). To gain further insight into the mechanism of our CH-GECO indicators, we determined the fluorescence intensity as a function of pH, both with and without Ca$^{2+}$ (Fig. 4). The resulting curves exhibited an unexpected level of complexity that is suggestive of a mechanism that is more complex than those previously reported.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$+/-$ Ca$^{2+}$</th>
<th>$\lambda_{\text{ex}}$ (nm)</th>
<th>$\lambda_{\text{em}}$ (nm)</th>
<th>$\Phi$</th>
<th>$\epsilon$ (M$^{-1}$ cm$^{-1}$)</th>
<th>Brightness (M$^{-1}$ cm$^{-1}$)</th>
<th>pK$_a$</th>
<th>$K_d$ (nM)</th>
<th>Hill coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH-GECO2.0</td>
<td>$-Ca^{2+}$</td>
<td>588</td>
<td>602</td>
<td>0.07</td>
<td>21 700</td>
<td>1.5</td>
<td>5.3</td>
<td>28</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>$+Ca^{2+}$</td>
<td>588</td>
<td>604</td>
<td>0.16</td>
<td>23 800</td>
<td>3.8</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH-GECO2.1</td>
<td>$-Ca^{2+}$</td>
<td>588</td>
<td>604</td>
<td>0.05</td>
<td>37 000</td>
<td>1.9</td>
<td>5.1</td>
<td>6</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>$+Ca^{2+}$</td>
<td>588</td>
<td>604</td>
<td>0.17</td>
<td>48 000</td>
<td>8.2</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Possible mechanisms for single FP-based Ca$^{2+}$ indicators. (A) As observed with most members of the GECO series, binding of Ca$^{2+}$ can induce a shift in the chromophore pK$_a$ which increases the proportion of the protein in the fluorescent anionic state (Zhao et al., 2011; Wu et al., 2013). Alternatively, Ca$^{2+}$ binding can cause a change in the intrinsic brightness of the chromophore (proportional to $\Phi \times \epsilon$), as observed for RCaMP (Akerboom et al., 2013). (B) Expected pH vs. fluorescence profile for a pK$_a$ modulation mechanism. (C) Expected pH vs. fluorescence profiles for an intrinsic brightness modulation mechanism.
Specifically, in CH-GECO1.0, 2.0 and 2.1, the $pK_a$ of the chromophore remains relatively constant at $\approx 5$–5.5 in both the absence and the presence of Ca$^{2+}$. Close examination of the CH-GECO2.1 pH titration curve in the absence of Ca$^{2+}$ reveals that there are two additional ionization events occurring: one with a $pK_a \approx 7$ and one a $pK_a \approx 9.5$. Analysis of the crystal structure of mCherry led Shu et al. (2006) to propose that Glu 215 is a titratable group with $pK_a$ properties (Shu et al., 2006), and we expect that it is playing an analogous role in CH-GECOs. Our efforts to identify the other ionizable group and to gain further insight into the mechanism of CH-GECO2.1 have been reported elsewhere (Carlson and Campbell, 2013).

To determine the affinity of the CH-GECO indicators for Ca$^{2+}$, we performed in vitro measurements of the indicator fluorescence in Ca$^{2+}$-buffered solutions (Fig. 5). These experiments revealed that, relative to other indicators of similar design, the mCherry-based have unexpectedly low $K_d$s of 28 nM for CH-GECO2.0 and 6 nM for CH-GECO2.1 (Table III). In contrast, the $K_d$ for R-GECO1 is 480 nM (Zhao et al., 2011). To determine the Ca$^{2+}$-association kinetics of the CH-GECO2.1 indicator, the rates of fluorescence increase for protein diluted into buffered solutions containing various amounts of Ca$^{2+}$ were determined using a stopped-flow spectrometer (Supplementary Fig. S4). Kinetic parameters extracted from the fitted data are in good agreement with the steady state $K_d$ (Supplementary Table SII). The CaM domains of CH-GECO2.0 and R-GECO1 differ by just the Asp23Ala substitution, while the CaM domains of CH-GECO2.1 and R-GECO1 differ by the additional Asp21Gly, Phe61Leu and Thr77Ser substitutions. To determine if one of these residues was responsible for the dramatic decrease in $K_d$, we reverted each mutation in the context of CH-GECO2.1. To our surprise the Ala23Asp, Gly21Asp and Ser77Thr variants of CH-GECO2.1 had essentially unchanged $K_d$s of 7, 7 and 8 nM, respectively. The Leu61Phe variant could not be purified as a soluble protein. These results indicate that the low $K_d$ of the CH-GECO variants is a result of an interaction between the CaM and the cp FP domain. Further mutagenesis studies to determine the nature of this interaction are ongoing.

**Live cell imaging**

To determine if CH-GECO2.0 and CH-GECO2.1 could be used for imaging of Ca$^{2+}$ dynamics in the cytoplasm of mammalian cells, HeLa cells were transiently transfected with appropriate expression vectors and the cells were stimulated with histamine to induce Ca$^{2+}$ oscillations (Palmer and Tsien, 2006). Representative intensity vs. time traces for individual cells are shown in Fig. 6. As expected for indicators with such low $K_d$s and relatively small in vitro responses to Ca$^{2+}$, histamine-, Ca$^{2+}$/ionomycin- and EGTA/ionomycin-induced fluorescence changes were rather muted and limited to $\approx 10$–30%. The low $K_d$ values mean that the majority of the protein should be in the Ca$^{2+}$-bound state even at resting Ca$^{2+}$ concentrations. EGTA/ionomycin washes were not able to lower the fluorescence below resting levels indicated that the residual Ca$^{2+}$ concentration is high enough for these indicators to remain largely in the bound state. Imaging results averaged over multiple cells is provided in Supplementary Table SIII.
Improved cp mCherry variants

Circular permutated FP variants with new termini in close proximity to the chromophore have proven to be an exceptionally useful scaffold for the construction of genetically encoded Ca\(^{2+}\) indicators (Nagai et al., 2001; Nakai et al., 2001). Unfortunately, red FPs such as the Discosoma-derived mCherry variant have proven to be somewhat less amenable to circular permuta-tion than their green fluorescent Aequorea-derived homologs (Li et al., 2008). This substantial technical hurdle meant that the first red fluorescent Ca\(^{2+}\) indicators (Zhao et al., 2011) were not reported until approximately a decade after the first green fluorescent Ca\(^{2+}\) indicators (Nagai et al., 2001; Nakai et al., 2001). One strategy we have taken to circumvent this problem has been to use directed evolution to rescue the folding and chromophore maturation efficiency in RFP variants with new termini that are still relatively distant from the chromophore (Carlson et al., 2010). In this way, we had hoped to create engineered variants that are more tolerant of permutation and effectively bootstrap our way to variants that can tolerate new termini in close proximity to the chromophore. Precedent for this approach comes from a report that folding-enhanced GFP showed increased tolerance of circular permutation (Pedelacq et al., 2006).

Building upon our previous efforts, we have successfully used a combination of error prone PCR, saturation mutagenesis and site-directed mutagenesis, coupled with high throughput image-based screens of bacterial colonies, to rescue the brightness of a cp mCherry variant with new termini at position 196. Our brightest variant, designated cp196V1.2, has 66% of the intrinsic brightness of mCherry and E. coli colonies expressing the gene are 40% as bright as those expressing mCherry after 24 h at 37°C. As illustrated in Fig. 1A, position 196 is relatively close to the chromophore of mCherry and is, seemingly, a very promising template for the construction of Ca\(^{2+}\) indicator. However, as we later found, permutation at position 145 in β-strand 7 was actually the most effective for construction of Ca\(^{2+}\) indicators. Notably, previous attempts to introduce circular permutation sites in β-strand 7 of mCherry (Shui et al., 2011) and mKate (Gautam et al., 2009; Shui et al., 2011) have been unsuccessful, suggesting that the additional mutations to create cp196V1.2 were critical to the success of this effort.

Fig. 6. Single cell Ca\(^{2+}\) imaging traces obtained with CH-GECO2.0 and CH-GECO2.1 in transiently transfected HeLa cells. (A) Intensity vs. time traces for cytoplasmic CH-GECO2.0 in single cells following addition of histamine, then EGTA plus ionomycin and finally Ca\(^{2+}\) plus ionomycin. (B) Intensity vs. time traces for cytoplasmic CH-GECO2.1 with the same treatments as in (A). Results from multiple cells are summarized in Supplementary Table SIII.

mCherry-based Ca\(^{2+}\) indicators

Using the optimized cp196V1.2 as a template, we generated eight libraries of potential Ca\(^{2+}\) indicator constructs with new termini in either β-strand 10 or β-strand 7. Functional Ca\(^{2+}\) indicators were identified only at certain permutation positions, with the only positive responses (i.e. a fluorescence increase upon binding Ca\(^{2+}\)) associated with proteins permutated at positions 146–145 and 147–146. To identify these indicators, we performed colony-based screening with the indicators exported to the high Ca\(^{2+}\) environment of the periplasm. This screening format would have biased our selection of variations to those that are bright in the Ca\(^{2+}\)-bound state. It is possible that variants with large inverted responses (i.e. those that are quenched in the bound state and bright in the apo state) would have been overlooked, but variants with modest inverted responses would have been picked if they retained substantial red fluorescence in the bound state.

With the benefit of hindsight, the identification of functional indicators at permutation sites 146–145 and 147–146 is perhaps not very surprising, since we effectively rediscovered the same permutation position that has been exploited in all previous Ca\(^{2+}\) indicator designs to date (Nagai et al., 2001; Nakai et al., 2001; Zhao et al., 2011; Akerboom et al., 2013).

We had initiated this work with the hypothesis that it should be possible to use permutation sites other than 146–145 for the creation of effective Ca\(^{2+}\) indicators. To the best of our knowledge, there have been no dedicated efforts to engineer indicators based on other permutation sites, so perhaps the general reliance on the 146–145 site is simply because it was the first site that was shown to work. Our original hypothesis was confirmed, in part, by the identification of some modestly functional inverse response Ca\(^{2+}\) indicators based on permutation at the 199–198 site in β-strand 10. It is possible that, with further engineering and optimization, indicators based on this site could be engineered to have large fluorescent responses. Overall, the results of this work provide strong support for the conclusion that 146–145 is a privileged site within FPs with respect to the creation of circularly permuted indicators.

The best indicator identified in our initial screen, a variant designated CH-GECO1.0, was permutated at 146–145 and exhibited a fluorescent response to Ca\(^{2+}\) of 72%. This response is much less than the 1600% change of R-GECO1 (Zhao et al., 2011), the 400–1200% change of RCaMP1.
variants (Akerboom et al., 2013) and the 2700–14 600% change of recently reported R-GECO1 variants (Wu et al., 2013). Accordingly, we sought to increase the response using directed evolution. While random mutation and screening did lead to a slightly improved variant, a more substantial improvement came in the form of an unintentional hybrid of R-GECO1 and CH-GECO1.1, designated CH-GECO2.0. Remarkably, this hybrid appears to have resulted from at least seven template switching events between the CH-GECO1.1 and a contaminating amount of the R-GECO1 template. Despite the fact that it contains a substantial number of R-GECO1-derived mutations, CH-GECO2.0 retains spectral properties that are very similar to that of CH-GECO1.1, and increased fluorescence response of 170%. A final round of directed evolution led to the identification of a further improved variant, CH-GECO2.1 with several more mutations and a 250% increase in fluorescence upon binding to Ca²⁺. While this response remains much less than other red fluorescent variants, we were unable to identify further improved variants through the screening of gene libraries created by error prone PCR or saturation mutation of particular residues.

**Prospects for mCherry-based Ca²⁺ indicators**

The two major limitations of CH-GECO2.1 are its modest 250% fluorescent response and its exceptionally high affinity for Ca²⁺ (Kₐ = 6 nM). A response of this magnitude could potentially be useful, and is actually quite good in comparison with many Ca²⁺ indicators based on Förster resonance energy transfer (Ding et al., 2011), but it falls well short of other currently available red fluorescent Ca²⁺ indicators (Zhao et al., 2011; Akerboom et al., 2013; Wu et al., 2013). This means that, even under the best case scenario where the Ca²⁺ concentration of interest is similar to the Kₐ (i.e. in the low nanomolar range), the fluorescence change will be modest. However, most biological Ca²⁺ imaging is concerned with concentrations in the range of ~1 μM (for the cytoplasm) to 10 s of μM (for the endoplasmic reticulum) (Palmer and Tsien, 2006), at which CH-GECO2.1 will exist mostly in the bound state. Consistent with our expectations, imaging of histamine-induced Ca²⁺ oscillations in HeLa cells with cytoplasmic CH-GECO2.1 revealed that the dynamic range was decreased several fold compared with the in vitro determined range. While further engineering will be required to tune the Kₐ of CH-GECO2.1 to be better matched with typical cytoplasmic or endoplasmic reticulum concentrations, the high affinity could prove to be useful for cell types with very low resting levels of Ca²⁺ (Horikawa et al., 2010).

In conclusion, we have engineered bright circularly permuted variants of mCherry and used the best one of these as a template for the development of a series of new Ca²⁺ indicators. Unfortunately, we cannot currently recommend any of our CH-GECO variants for routine live cell imaging applications due to the availability of alternative green (Akerboom et al., 2012) and red (Zhao et al., 2011; Akerboom et al., 2013; Wu et al., 2013) variants with much larger signal changes and Kₐs that are better tuned to physiologically relevant Ca²⁺ concentrations. Rather, we suggest that the CH-GECO series may serve as promising templates for further engineering to create improved high-affinity Ca²⁺ indicators or indicators for biological analytes other than Ca²⁺. For example, single FP-based indicators for hydrogen peroxide (Belousov et al., 2006) and ATP: ADP ratio (Berg et al., 2009) have been reported, and it is possible that the cp RFP portion of CH-GECO could be incorporated into such designs to create red fluorescent versions. For the immediate future, our efforts will focus on probing the unique mechanism of CH-GECO2.1 with particular emphasis on identifying the key ionizable group with a pKₐ ~7, and the key interactions between the CaM and the FP domain.

**Supplementary data**

Supplementary data are available at PEDS online.

**Acknowledgements**

We thank the University of Alberta Molecular Biology Services Unit and Yongxin Zhao for technical support and advice. In addition, we thank Andy Holt and Christopher W. Cairo for providing access to instrumentation.

**Funding**

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC). H.J.C. was supported by postgraduate scholarships from NSERC and Alberta Ingenuity. R.E.C. holds a Tier II Canada Research Chair in Bioanalytical Chemistry.

**References**


