Overcoming chromoprotein limitations by engineering a red fluorescent protein

Letian Bao, P. Navaneeth K. Menon, Josefine Liljeruhm, Anthony C. Forster *

Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

ARTICLE INFO

Keywords:
Chromoprotein
Fluorescent protein
mRFP1
Toxicity
Fitness
E. coli

ABSTRACT

Chromoproteins (CPs) are widely-used visual reporters of gene expression. We previously showed that, for coloration in Escherichia coli, CPs had to be overexpressed and that this caused large fitness costs with the most useful (darkly colored) CPs. These fitness costs were problematic because passage of plasmids encoding darkly colored CPs in liquid culture frequently resulted in loss of color due to mutations. Unexpectedly, an early variant of the monomeric red fluorescent protein 1 (mRFP1) gene that was codon-optimized for E. coli (abbreviated mRFP1E) was found here to be an ideal replacement for CP genes. When we subcloned mRFP1E in the same way as our CP genes, it produced a similarly dark color, yet affected E. coli fitness minimally. This finding facilitated testing of several hypotheses on the cause of CP cytotoxicities by gel electrophoresis and size-exclusion chromatography: toxicities correlated with the combination of amounts of expression, oligomerization and inclusion bodies, not isoelectric point. Finally, a semi-rational mutagenesis strategy created several mRFP1 protein variants with different colors without altering the fitness cost. Thus, these mutants and mRFP1E are suitable for comparative fitness costs between different strains of E. coli. We conclude that our new mRFP1E series overcomes prior limitations of CPs.

1. Introduction

The family of eukaryotic fluorescent proteins (FPs; [1]), chromoproteins (CPs; [2]) and their engineered variants are widely used as reporters for gene expression. They derive from jellyfish and corals, sharing a homologous three-amino-acid fluorophore/chromophore which self-matures post-translationally within a β-barrel by reacting with oxygen [1].

CPs have certain advantages over FPs, such as having dark colors under ambient light that enable inexpensive, instrument-free analysis. CPs also circumvent problems due to background fluorescence of the sample, ultraviolet-light-induced FP photobleaching and cell damage, and the need for eye and skin protection. Applications of CPs include markers in living organisms for cloning [3,4], teaching [5] and biosensors ([6] and references therein).

We recently expressed constitutively 14 CPs from plasmids in E. coli and compared their properties [7]. One conclusion was that, like FPs [8], CPs needed to be heavily overexpressed for easy detection. Although high-copy plasmids gave strong coloration, most of the CP colors were frequently lost in liquid culture due to mutations in the plasmid that silenced CP expression without affecting antibiotic resistance. Thus, most CPs were quite toxic to their E. coli hosts, despite having being codon-optimized for E. coli, resulting in strong selection pressure for loss of expression. Unfortunately, the two CPs that affected the growth rates the least, amilGFP and fwYellow, exhibited the lowest color contrast to yellowish wild-type E. coli and LB agar. So, all of the really useful (darkly colored and fast-maturing) CPs often lost their colors in liquid culture, which was challenging for maintenance of stocks and suboptimal when choosing a reporter [9], especially for strain competition studies. However, loss of CP expression was negligible after re-streaking on LB agar plates from fresh plates because solid culturing alone (with each colony starting from a single cell bottleneck) reduces bacterial and plasmid generations compared with solid then liquid culturing (with each liquid culture starting from up to 10^8 cells/colony), and because better selection occurs in liquid culture due to free bacterial movement in three dimensions [10].

Our initial solution to the instability of CP genes was to dramatically reduce their copies by integrating them into the E. coli chromosome and amplifying expression with maximum-strength promoters [7]. However, that solution was more complicated compared with cloning in plasmids,

* Corresponding author.
E-mail address: a.forster@icm.uu.se (A.C. Forster).

https://doi.org/10.1016/j.ab.2020.113936
Received 22 May 2020; Received in revised form 28 August 2020; Accepted 30 August 2020
Available online 3 September 2020
0003-2697/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
2. Material and Methods

### 2.1. Strains

Unless otherwise stated, subcloning was in *E. coli* DH5α and CPs/FPs were overexpressed constitutively in *E. coli* MG1655, the exception being an IPTG induction in *E. coli* BL21 (DE3).

### 2.2. Plasmids

BioBrick plasmid vectors and parts specified with BBA numbers below were obtained from the Registry of Standard Biological Parts (http://parts.igem.org/Main_Page). pSB1C3, pSB1K3, pSB3K3 (backbones in Fig. S1) and pSB1K3 contained the *E. coli* codon-optimized, BioBrick restriction-enzyme-free (RFC10 standard) mRFP1 Bba_E1010 marker gene (abbreviated to mRFP1E here) under the control of a LacI-repressible promoter (which leaks because LacI is only encoded on the chromosome, not the plasmid; see bottom construct of Fig. 13.2 of [11]; the ccdB positive selection marker at the top right was absent). Initial related versions of the vectors were described by [12] using the nomenclature pSB1, high copy = 100–300 for origin pMB1; pSB3, medium copy = 20–30 for origin P15A; K, kanamycinR; C, chloramphenicolR; R = resistance. Vectors were used according to the international Genetically Engineered Machine (iGEM) competition’s version of BioBrick 3A cloning [5] that drops out the mRFP1E marker insert (http://parts.igem.org/Main_Page) rather than the primary published version of BioBrick 3A cloning [11] that drops out the ccdB marker. Note that colonies expressing mRFP1E can be safely and inexpensively distinguished from non-fluorescent red/pink CP colonies by illuminating with blue light and observing fluorescence through orange glasses (Safe Imager, Invitrogen).

The 14 BioBrick CP genes relevant to this study have been described [7] (N.B. gfasPurple is BBa_K1033919). However, upon submission to Addgene, this non-profit company determined that 5/14 contained unexpected sequences upstream of the genes. We therefore reconstructed the five different sequences inferred from our publication by PCR mutagenesis as follows:

- tsPurple, fwYellow and scOrange: ACCTTAGAGGTAAAACAT was changed to AAAGAGGAGAAATACTAG.
- spisPink: ATTTAATAGGAGGTAGTTG was changed to AAAGAGGAGAAATACTAG.
- meffBlue: GCTCTTTAAACATTATCATGACTAATAGGAGTACAC was changed to AAAGAGGAGAAATACTAG.

In side-by-side comparisons on LB agar plates, the reconstructed CPs showed very similar color intensities to the versions deposited at the five different sequences inferred from our publication by PCR (http://parts.igem.org/Main_Page) rather than the primary published version of BioBrick 3A cloning [11] that drops out the ccdB marker. Note that colonies expressing mRFP1E can be safely and inexpensively distinguished from non-fluorescent red/pink CP colonies by illuminating with blue light and observing fluorescence through orange glasses (Safe Imager, Invitrogen).

### 2.3. Subcloning of mRFP1E

BBA_J3502 (*E. coli*-codon-optimized mRFP1 coding sequence Bba_E1010 plus RBS Bba_B0034) was assembled with constitutive medium-strength promoter Bba_J23110 or constitutive maximal promoter apFB46 [13] in plasmids pSB1C3, pSB1K3 and/or pSB3K3 by iGEM 3A assembly (see above). The apFB46 promoter flankned by EcoRI, XbaI and SpeI restriction sites was synthesized by hybridization of 5′-AATTCGGCGGCCTTCTAGAAGGGCGGCAAAAGATATTGACTGCAGTCTTTTGTACATTAATAGTTCGCTATA-3′ and 5′-CTGTTATAGCAATCTATTTATAGTACCAAAGATGGGAGACTCATTCTTTTGGCGCGCCCTCTAGAAGCGGCCGCG-3′ purchased from Integrated DNA Technologies, Inc. The coding sequence and flanking regions of pSB1C3-B0034-E1010 are given in Fig. S1.

### 2.4. Mutagenesis of amino acids 64–65 of the mRFP1 fluorophore region and “superfolder” mutation of CPs

Inverse PCR mutagenesis was performed as described [5,14] using the phosphorylated primers 5′-NNNTCCTGGGGAGAACCCATGTCC-3′ and 5′-NNNTACGGGTCTCAAAAGCTTACGT-3′ (Fig. S1). mRFP1E mutant sequences were: mRFP1_Yellow: ATTCCT; mRFP1_Pink: GTCTCT; mRFP1_Amber: TTATAT; mRFP1_Green: TTTTGTT.

Known superfolder/mutation-accelerating mutations in dsRed1 (amino acids 41–44 = TQNV [15,16]) were introduced in homologous positions of amICP (wild-type EQTV -> TQNV and TQTV) and asPink (wild-type TQEM -> TQNV, TQTV, TQTA and TQNA) in BioBrick plasmids.

### 2.5. Synthesis and subcloning of BioBrick mammalian-codon-optimized mRFP1 and dsRed

The new BioBrick genes were synthesized (Twist Bioscience) to contain the original mammalian-codon-optimized mRFP1 and dsRed gene sequences [17] except with one synonymous mutation each (Q114: CAG to CAA) to remove the only BioBrick site (PstI). These were then subcloned with exactly the same flanking sequences as mRFP1E in pSB1C3, and further subcloned into pET24a vector (Novagen) for IPTG induction under control of T7 RNA polymerase in *E. coli* BL21 (DE3).

### 2.6. Plasmid stability assay in liquid culture

Bacteria transformed with different plasmids were inoculated into 1 mL LB medium with chloramphenicol (25 µg/mL) or kanamycin (50 µg/mL) in 10 biological replicates. Dilution by 1000-fold was performed with LB chloramphenicol/kanamycin every 24 h to allow ~10 generations of growth of bacteria. ~40 Generations were continuously grown unless all color was lost in all replicates. All visualizations were done in ambient light.

### 2.7. Growth rate assay of mRFP1 variants and CPs

Plasmids were transformed into *E. coli* MG1655. Bacterial growth rate assays were performed in LB medium with chloramphenicol (25 µg/mL) or kanamycin (50 µg/mL). Overnight liquid cultures were diluted 1000-fold with LB chloramphenicol/kanamycin and 200 µl of each sample was transferred to 96-well plates (BRAND, Germany). The cultures (4 biological replicates of each sample in duplicate technological replicates) were grown with continuous shaking for 16 h and optical density (OD) at 600 nm was measured every 5 min using an Infinite M200 Pro plate reader (Tecan, Switzerland). Maximum growth rates were calculated from the linear part of the lnOD vs. t curve using

\[
G = \frac{\text{ln}2 \times (t_2 - t_1)}{\text{ln}\text{OD}_{600(t_2)} - \text{ln}\text{OD}_{600(t_1)}},
\]

where G = generation time, t = time. Controls were blank LB medium, wild-type *E. coli* MG1655 and bacteria with promoter-less mRFP1E (pSB1C3-B0034-mRFP1E). The promoter-less mRFP1 generation time was normalized to 1.
2.8. Competition fitness cost assay

The protocol was adapted from an FP publication [18]. Starter cultures of individual colonies of E. coli MG1655 transformed with CP and mRFP1 mutant plasmids (at least 5 biological replicas) were grown overnight in LB medium. Then cultures were mixed 1:1 and serially passaged with a 1000-fold dilution every 24 h (1 μL culture in 1 mL LB), resulting in 10 generations of growth per passage. The ratios between the two competing strains were measured by plating and scoring for different colored colonies. The strains were competed for 30–50 generations and selection coefficients were calculated as described [19], [31] using the regression model

\[ s = \ln \left( \frac{R_i}{R_0} \right) / t \]

where R is the ratio of CP expressing strain to wild type, t is the number of generations.

2.9. Preparation of soluble fraction and inclusion bodies (IBs) for SDS-PAGE

Overnight cultured cells (2 mL with 3 biological replicates) were collected by centrifugation at 13,000 rpm for 1 min, resuspended in 300 μL non-ionic detergent lysis buffer (0.5% v/v Triton X-100, 100 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl pH 8.0) and 0.8 mg/mL (final conc.) lysozyme in 10 mM Tris-HCl pH 8.0 and incubated at 37 °C for 1 h. After three cycles of freezing with liquid nitrogen and thawing, the same amount of lysozyme was added again, the samples further incubated at 37 °C for 1.5 h, then refrozen and thawed to ensure thorough lysis.

IBs were obtained as follows [32]. Lysates were centrifuged at 14,000 rpm for 5 min and the supernatant (soluble fraction) stored at −20 °C. Pellets (10 μL) were resuspended with 1 mL 50 mM Tris-HCl containing 4 M urea (pH adjusted to 8.0), centrifuged at 14,000 rpm for 5 min and the supernatant (containing solubilized membrane proteins) discarded. The washed pellets were dissolved (330 μL for 2 mL overnight culture) in 50 mM Tris-HCl containing 8 M urea (pH adjusted to 8.0) and shaken at 220 rpm at 37 °C for ~16 h. Samples were centrifuged at 14,000 rpm for 5 min and 7.5 μL supernatant (IBs) was mixed with 2.5 μL 4 × SDS loading buffer (6.4% SDS, 16% glycerol, 4% β-mercaptoethanol, 63 mM Tris-HCl pH 8.0, 0.004% bromophenol blue) and loaded without boiling on a 12% SDS polyacrylamide gel. Soluble fraction and total cells were boiled for 10 min with 1× SDS loading buffer before loading. The marker ladder was PageRuler Prestained Protein Ladder 10–180 kDa (ThermoFisher).

2.10. "Semi-native" lysis and loading on SDS-PAGE

Our method was adapted from FP publications [17,20,21] by using a different lysis buffer. The stored supernatant from the above non-ionic detergent lysis protocol was thawed on ice and mixed with normal 4× SDS loading buffer just before loading without boiling on a 12% SDS polyacrylamide gel. The gel was run with 1× SDS running buffer (40 mM Tris base, 0.2 M glycine, 0.03% SDS, pH adjusted to 8.5). After electrophoresis, the gel was observed under ambient light, then UV light (254 and 366 nm, with the latter showing the bands better), then stained with Coomassie Blue.

2.11. Native lysis and PAGE

Overnight cultures of E. coli MG1655 bearing CP plasmids grown at 37 °C (25 °C for cJbBlue potentially to enable soluble expression) were harvested by centrifugation at 4000 rpm for 5 min at 4 °C, resuspended in 2 mL 50 mM Tris-Cl pH 8.0 containing 300 mM NaCl, lysed by sonication (cycles of sonication for 5 s and pausing 5 s, with total time of 40 min) on ice-water mixture. After sonication, cell lysates were centrifugated at 1400 rpm for 10 min at 4 °C, supernatants were collected and filtered through 0.2 μm filters and kept on ice. PAGE (12% acrylamide, 0.4 M Tris-Cl pH 8.0, 0.3% APS, 0.08% TEMED) used 40 mM Tris-Cl containing 0.2 M glycine pH 8.0 as running buffer. Samples were made to 5% glycerol for loading and run at 90 V.

2.12. Size exclusion chromatography

A Superdex® 75 10/300 GL column (GE Healthcare Biosciences AB, Sweden) was pre-equilibrated with two column volumes of 50 mM Tris-HCl pH 8.0, 300 mM NaCl. 100 μL of CP prepared as in 2.11 was loaded into the column at 0.5 mL/min. UV detection was at each CP’s maximal absorbance (except cJblue was detected at 600 nm). 100 μL mixture of purified standards was loaded on the same day and detected at 280 nm.

2.13. Fluorescence spectroscopy

Soluble fractions of overnight cultures, freshly prepared as in 2.9, were analyzed immediately without dilution on a Hitachi F-7000 fluorescence spectrophotometer. mRFP1_Yellow and promoter-less mRFP1 had emissions <5% of mRFP1 and thus were considered non-fluorescent. mRFP1_Yellow had Absmax = 498 nm.

2.14. Homologous modeling of the mRFP1 3D structure

The structure was predicted using dsRed (PDB: 1ZGO) as the template for homologous modeling by SWISS-MODEL (https://swissmodel.expasy.org/).

3. Results and discussion

3.1. Strategies for stabilizing darkly colored CP expression

We reasoned that several approaches might potentially yield a CP gene that could be overexpressed from a high-copy plasmid in E. coli to give a darkly colored CP with a low fitness cost.

First, we tried to darken the colors of the two CPs that had low fitness costs: the yellow amilGFP and fwYellow [7]. The strategy was based on our successful changing of the color of amilCP by randomly mutating the first amino acid of the chromophore tripeptide and the immediately-preceding amino acid (positions 64 and 65, GFP numbering). Although inverse PCR mutagenesis of amilGFP [7] and fwYellow worked, unfortunately the colonies did not have different colors and almost all lost their fluorescence. This was not unexpected, given the low sequence homology between the CPs and our similar disappointing results earlier using the same mutagenesis strategy with spisPink and t5Purple (unpublished). Thus, we concluded at this stage that our color-change method seemed specific to amilCP (but see below).

Second, we tried to make superfolder variants of darkly colored CPs to see if they would be darker due to folding more efficiently and/or maturing the chromophore faster, thus requiring less overexpression for visualization. Known superfolder/maturation-accelerating mutations [15,16] were introduced at homologous positions in plasmid- and chromosomally-encoded CPs (see Material and Methods for details). For the former, maturation times were compared by opening to air plates that had been grown anaerobically overnight in bags of nitrogen gas, but maturation times were not hastened. For the latter, expression was subsequently optimized by random mutagenesis, with screening for darker colonies being conducted visually. Although darker colonies were sometimes produced, the most promising was unfortunately due to synthesis of more, not darker colored, CP based on stained gel bands (not shown).

Third, we tried expressing darkly colored CPs from a medium-copy plasmid (pSB3K3) by using a maximal promoter, instead of our standard medium-strength constitutive promoter (J23110) in a high-copy...
plasmids we had been amplifying contained an mRFP1 gene mRFP1E here. We thus shifted focus to comparing mRFP1E with darkly re-codon-optimized for mammals by CLONTECH Laboratories, the equal-lowest fitness cost [7], amiilGFP, indicating that the promoter was too strong (see below) and that balancing promoter strength and plasmid copy number is not straightforward with CPs.

Fourth, we evaluated a particular red fluorescent protein (RFP) gene (Fig. 13.2 of [11]) that is often used as a visible marker for plasmid cloning in E. coli (e.g. [5]). This was because, when amplifying the plasmid vectors in liquid cultures, we noticed that its red color seemed surprisingly stable. The gene encodes an extensively-engineered, monomeric protein derivative (mRFP1 [17]) of the tetrameric dsRed FP from Discosoma sp. coral [20]. Although the starting dsRed gene had been codon-optimized for mammals by CLONTECH Laboratories, the plasmids we had been amplifying contained an mRFP1 gene re-codon-optimized for E. coli called mRFP1 BBa_E1010, abbreviated to mRFP1E here. We thus shifted focus to comparing mRFP1E with darkly colored CPs with the aim of circumventing CP toxicities.

3.2. mRFP1E produces comparable color intensity with much lower toxicity than genes for darkly colored CPs

To enable direct comparison with our CP plasmids, the mRFP1E coding region was subcloned so that it had exactly the same flanking sequences: high-copy vector pSB1C3, constitutive medium-strength promoter J23110 and strong RBS B0034. pSB1C3-J23110-mRFP1E produced visible color comparable in darkness to CPs on LB agar plates (Fig. 1A and S2A), despite mRFP1 not being regarded as a CP. Although CPs had proven too toxic for subcloning downstream of the maximal promoter apFAB46 in medium-copy pSB3K3 (see above), such a subcloning for mRFP1E surprisingly succeeded by giving large red colonies (Fig. 1B and S2B, right half). These colonies were 60% darker than the combination of medium-copy plasmid pSB3K3 and medium promoter J23110 (Fig. 1B and S2B) and were even 15% darker than our standard combination of high-copy plasmid and medium promoter.

Fig. 1. Color intensities and toxicities of mRFP1E and CP plasmid expression. Effects of promoter strength (medium J23110 versus maximal apFAB46) and plasmid copy number (high pSB1- versus medium pSB3-) are compared under appropriate antibiotic selection (chloramphenicol [C] versus kanamycin [K]). (A) and (B) Plates incubated at 37 °C for 20 h before photography under ambient light. Average color intensities of the colony centres ± standard deviations were determined as in Fig. S2. (C) Exponential growth rates ± standard errors (~3 h after dilution) versus promoter-less mRFP1E control. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Third, given that all native CPs examined are tetramers or higher-order oligomers [2,26] while mRFP1 was engineered to be monomeric, we measured CP and mRFP1 oligomeric states in the soluble fraction of E. coli by "semi-native" SDS-PAGE (see Material and Methods). For all cases, except the slowly-maturing cjBlue, a single prominent band of the expected color was visualized in ambient light or under ultraviolet light (Fig. 2 top and middle, respectively). Staining (Fig. 2 bottom) showed that all of the bands observed in the upper two panels were indeed the major overexpressed soluble proteins in the cells. A major cjBlue band was absent due to most residing in IBs (see below). The mobilities of all of the native CP proteins were consistent with tetramers or higher oligomeric forms as expected, and only the three engineered protein sequences ran much faster: scOrange (dimeric mobility), mRFP1 (monomeric mobility) and fwYellow (monomeric mobility; result not shown). We concluded that our "semi-native" lysis and loading conditions, despite containing SDS in the gel, preserved the CP/FP β barrel tertiary structure (based on its requirement for color/fluorescence and tetramerization) and preserved the tetrameric structure of all the native (unengineered) CPs (possibly a requirement for their color/fluorescence). Completely detergent-free lysis/PAGE [27], where mobility is affected by pl in addition to size, surprisingly gave very similar relative mobilities to the "semi-native" SDS PAGE despite the large variation in pls (Fig. S5). However, size estimates by non-denaturing lysis/size exclusion chromatography [27], which should be more reliable due to independence from influence of pl and detergent exposure, differed in showing that fwYellow was actually a dimer and scOrange was actually a tetramer (Fig. S6). The discrepancies in sizes for these two proteins were presumably due to them having the next lowest pls to mRFP1, assuming that "semi-native" SDS PAGE is more affected by pl than we originally expected.

As their associated growth rates (Fig. S3) ranged from fastest rates for mRFP1 monomer, fwYellow dimer and amilGFP tetramer to medium rates for amilCP, asPink and scOrange tetramers to slowest rates for cjBlue, amajLime and gfasPurple tetramers, they could not be explained simply by the extent of oligomerization. Nevertheless, the two smaller than tetramer were the fastest growing.

Fourth, we measured the IBs and total cellular amounts of mRFP1 and the CPs. Protein overexpression in E. coli often forms IBs, and early FP studies reported non-fluorescent IBs of GFP [28]. Regarding total cellular proteins and the soluble fraction (Fig. 3A and B), asPink stood out with apparent boiling-dependent cleavage at the chromophore of the soluble (matured) protein in line with [29]. amilGFP was the least expressed, probably because it was one of only two proteins lacking codon optimization (the other being amilCP); least expression potentially provided a trivial explanation for its equal-least toxicity of the CPs (Fig. S3). However, this correlation did not generalize to the other proteins (Fig. 3A). Only when also examining the IB measurements (Fig. 3C) could the growth rates finally be explained. The slower rates produced IBs while the faster rates lacked significant IBs. In addition, the three fastest of the two smaller-than-tetramers (mRFP1 and fwYellow) and the lowest expressing (amilGFP). In the case of gsOrange, the protein for IB formation and greater toxicity can be traced to just its four amino-acid differences from amilCP [2] which do not significantly change its pl. Note that cjBlue was mostly IBs (compare Fig. 3B and C), explaining the lack of detection in the soluble fraction (Figs. 2 and 3B) and likely its slow color formation [7]. Interestingly, expression at 25 °C instead of 37 °C abolished IB production for mRFP1E and enabled blue color and the other CPs (results not shown). Although growth at 25 °C is too slow for production of dark colors overnight, it is likely advantageous for preparation of correctly-folded versions of CPs that gave IBs at 37 °C. We also note that we preferred the pSB1C3 backbone for our experiments because, surprisingly, the pSB1K3 backbone at 37 °C produced IBs from its kanamycin resistance gene.

Fifth, we wondered how the toxicities of the E. coli-codon-optimized mRFP1E gene compared with the mammalian-codon-biased mRFP1 gene [17]. Thus, the mRFP1 gene was made by gene synthesis (with one synonymous mutation to remove the BioBrick restriction site) and subcloned with exactly the same flanking sequences as mRFP1E in pSB1C3. Surprisingly, E. coli cells transformed with the mRFP1 gene were uncolored due to a lack of overexpressed mRFP1 protein (no overexpression band on SDS-PAGE). We thus speculated that the purple-color of the E. coli colonies in Fig. 8D of [17] might have been due to the use of a T7 RNA polymerase inducible promoter instead of an E. coli RNA polymerase constitutive promoter, so we subcloned the
mRFP1 gene into the inducible pET24a vector. Indeed, this transformed pET construct, upon induction with IPTG, did yield red-colored BL21 (DE3) E. coli cells (not shown). Thus, codon optimization for E. coli was necessary for coloring E. coli by constitutive overexpression of the mRFP1 protein, in contrast with amilCP and amilGFP [7].

In summary, FP/CP toxicities in E. coli correlated well with combined total expression, oligomerization and IB data. This was not necessarily expected because, while more protein expression and oligomerization are likely to promote aggregation and IB formation in general, IBs can be non toxic in bacteria.

3.4. mRFP1 fluorophore-region mutagenesis creates different-colored variants with equivalent fitness costs

The last step needed to overcome the limitations of CPs was to develop mRFP1 mutants with different dark colors for competition fitness-cost assays. The fitness cost or advantage of a mutation of interest (e.g. antibiotic resistance) is measured by competition assays between isogenic cell lines differing in as few other genetic changes as possible [19]. Although yellow, orange and purple mutants of the mRFP1 gene existed, these contained several amino-acid changes (which may or may not differentially affect bacterial fitness) and mammalian codon bias [30]. We previously obtained color variants of amilCP for competition assays by randomly mutating positions 64 and 65 [7] but thought it unlikely to work with mRFP1 because: (i) mRFP1 and amilCP share low sequence homology, (ii) it did not change the colors of four other CPs (see results above), and (iii) randomly mutating position 65 alone did not generate very different-colored mRFP1 variants, although one gave an orange gel band (Fig. 1d of [21]). Nevertheless, we tried the same strategy for mRFP1 and, to our surprise, it succeeded, although a much smaller fraction of mutants was colored (Fig. 4A; Fig. S7A). Yellow, pink, purple-red and purple mRFP1 variants were created, and we also recreated the orange variant [21] (but with E. coli codon bias), with useful darkneses exhibited by the non-yellow ones (Figs. S7B and S8).

All yellow variants had lost fluorescence (see 2.13), while the other variants had significant shifts in their absorption/excitation/emission spectra (Fig. 4B). As fluorophore region mutations might be expected to affect maturation times, we compared these by opening to air plates that had been grown anaerobically overnight in bags of nitrogen gas (Fig. S9). Only the pink mutant had a substantially slower rate of maturation. Faster maturation or additional colors might be engineered in the future by mutagenesis of other amino acids neighboring [33] the fluorophore (Fig. S10).

The characterized mRFP1 and amilCP mutant fluorophore regions (see amino acid sequences in Fig. 4A and [7]) had related pairs of amino-acid substitutions: mostly hydrophobic at 64 and often methionine at 65. There was some correlation with the known conservation of F64 in FPs, not CPs [7], and with visible colors of identical 64-67 fluorophore regions. More specifically, mRFP1_Magenta (FMYG) shares the same fluorophore region with similarly-colored mCherry [30], while mRFP_Violet (CMYG) shares the same fluorophore region with tsPurple, but CMYG is also present in differently-colored asPink, scOrange and aeBlue [7], and mRFP1_Orange (FCYG) shares the same fluorophore region with pink-colored mTangerine [30]. The fluorophore regions of mRFP1_Pin (VTYG) and non-fluorescent mRFP1_Yellow (ILYG) are absent from the FP database (https://www.fpbase.org/).

Fitness costs of mRFP1E and its mutants were measured by competition with promoter-less mRFP1E (Fig. 4C). As expected and desired, the colored variants had indistinguishable toxicities. Thus, the plasmid-encoded mRFP1E variants are highly suitable markers for visual competition assays. Given their low fitness costs in comparison with darkly colored CPs, they have a major advantage over CPs in these and other colorimetric assays.

4. Conclusions

Cytotoxicities of the mRFP1E and CP genes in E. coli correlated with combined expression level, oligomerization and inclusion bodies data, suggesting protein aggregation as the molecular cause. While the toxicities have been problematic for the handling and applications of the most useful (darkly colored) plasmid-encoded CPs, we circumvented the problem here with the mRFP1E gene and our new different-colored variants that had equivalent fitness costs. Furthermore, our monomorphic variants may have advantages over oligomerizing CPs in fusion-tagging experiments. In conclusion, the mRFP1E series overcomes limitations of CPs as visual reporters in E. coli for teaching, research and commercial applications, and we are making the plasmids available without restriction.

Author contribution

JL and ACF conceptualized the study, PNKM and JL performed pilot mutagenesis studies, PNKM performed pilot gel studies, LB gathered all data shown in the figures, all authors analyzed the data, LB and ACF

Fig. 3. SDS-12% PAGE of protein fractions from mRFP1E and CP constructs. Denaturing gels stained with Coomassie Blue from a representative experiment showing (A) total proteins, (B) soluble fractions and (C) inclusion bodies. Fractions from equivalent cell weights were loaded from left to right in order of decreasing CP growth rate [7]. The calculated MW of mRFP1E protein is 25.4 kDa. Note that the appearance of the smaller asPink band in A and B was dependent upon boiling.
Fig. 4. Fluorophore region mutants of mRFP1 protein and their colors, fluorescence spectra and fitness costs in competition assays. (A) F64 and Q65 of mRFP1 (GFP numbering) were mutated randomly, different-colored bacterial colonies were selected and sequenced (mutated amino acid pairs are given under names) and pSB1C3 derivatives were streaked on a chloramphenicol plate then incubated at 37 °C for 18 h. (B) Soluble fractions of mRFP1 mutants photographed in ambient light and analyzed for fluorescence immediately after preparation. Excitation spectra (dashed line; emission collected at 640 nm) and emission spectra (solid lines, excitation at 540 nm) were individually normalized. (C) Fitness costs of the mutants measured by direct competition in liquid LB medium with isogenic E. coli transformed with the promoter-less pSB1C3–B0034-mRFP1 plasmid (except amilGFP was competed with mRFP1E). Analysis was by plating and counting of colored versus white colonies. Negative values designate fitness costs and error bars are standard errors. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

wrote the manuscript.

Subject category

Protein structure and analysis.

Declaration of competing interest

None.

Acknowledgements

We are very grateful to Erik Wistrand-Yuen for advice on fitness costs, Emil Marklund for advice on fluorescence spectroscopy, Victoria Karpenko for subcloning assistance and Robert Campbell and Yi Shen for comments on the manuscript. A.C.F. was supported by the Swedish Research Council (grant no. 2017–04148).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ab.2020.113936.

References


