Recombination-stable multimeric green fluorescent protein for characterization of weak promoter outputs in Saccharomyces cerevisiae

Rugbjerg, Peter; Knuf, Christoph; Förster, Jochen; Sommer, Morten Otto Alexander

Published in:
FEMS Yeast Research

Link to article, DOI:
10.1093/femsyr/fov085

Publication date:
2015

Document Version
Peer reviewed version

Recombination-stable multimeric green fluorescent protein for characterization of weak promoter outputs in *Saccharomyces cerevisiae*

Research article

Running title: Recombination-stable multimeric GFP

Authors:

Peter Rugbjerg[1] – petru@biosustain.dtu.dk

Christoph Knuf[1] – chrik@biosustain.dtu.dk

Jochen Förster[1] – jfor@biosustain.dtu.dk

Morten O. A. Sommer[1 *] - msom@bio.dtu.dk

1) Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark

Kogle Allé 6, DK-2970 Hørsholm, Denmark

* Corresponding author

Abstract

Green fluorescent proteins (GFPs) are widely used for visualization of proteins to track localization and expression dynamics. However, phenotypically important processes can operate at too low expression levels for routine detection, i.e. be overshadowed by autofluorescence noise. While GFP functions well in translational fusions, the use of tandem GFPs to amplify fluorescence signals is currently avoided in *Saccharomyces cerevisiae* and many other microorganisms due to the risk of loop-out by direct-repeat recombination. We increased GFP fluorescence by translationally fusing three different GFP variants, yeast-enhanced GFP, GFP+ and superfolder GFP to yield a sequence-diverged triple GFP molecule 3vGFP with 74-84 % internal repeat identity. Unlike a single GFP, the brightness of 3vGFP allowed characterization of a weak promoter in *S. cerevisiae*. Utilizing 3vGFP, we further engineered a less leaky Cu^{2+}-inducible promoter based on *CUP1*. The basal expression level of the new promoter was approx. 61 % below the wild-type *CUP1* promoter, thus expanding the absolute range of Cu^{2+}-based gene control. The stability of 3vGFP towards direct-repeat recombination was assayed in *S. cerevisiae* cultured for 25 generations under strong and slightly toxic expression after which only limited reduction in fluorescence was detectable. Such non-recombinogenic GFPs can help quantify intracellular responses operating a low copy number in recombination-prone organisms.

Keywords: signal amplification, synthetic biology, promoter engineering, protein multimerization

Introduction

Green fluorescent protein (GFP) is an invaluable tool for real-time visualization of intracellular proteins. Since the initial cloning, numerous improvements, variants and applications have been developed (Snapp 2009; Miyawaki 2011). GFP is particularly useful for quantification of intracellular events, localizations and populations at single-cell resolution. However, a minimal expression level is required such that the fluorescent output exceeds the cell autofluorescence and produces detectable signals. Still, biologically important processes occur through the interaction of a few molecules per cell, which is hard to quantify using existing fluorescent proteins and non-specialized experimental setups (Raj and van Oudenaarden 2009; Li and Xie 2011; Gahlmann and Moerner 2014). Further, the engineering of synthetic cell functionalities can depend on fine characterization and balancing of low gene expression levels (Ajikumar et al. 2010; Harton et al. 2013).

The strategies for improving fluorescent output signals include the design of new GFP variants such as GFP+, yeast-enhanced GFP (yEGFP) and superfolder GFP (sfGFP) (Cormack et al. 1997; Scholz et al. 2000; Pédelacq et al. 2006). Still, monitoring of single-molecule events such as chromosome movements in *Escherichia coli* has e.g. required multimerization of 96 DNA binding sites to localize enough fluorescent protein to produce a distinguishable signal (Xie et al. 2008). Artificial tethering of a bright yellow fluorescent protein (Venus YFP) to the inside *E. coli* cell membrane allowed a microscope-detectable signal from a single YFP-tagged protein (Yu et al. 2006). Thus without techniques for single-molecule GFP sensitivity, the full-genome mapping of subcellular protein localization in *Saccharomyces cerevisiae* (yeastGFP) did not produce signals above background for 361 open reading frames (8 pct. of

total) otherwise shown to be expressed in the growth phase assayed (Ghaemmaghami et al. 2003; Huh et al. 2003). Equivalently, the issue of not detecting all low-expressing *S. cerevisiae* proteins was also observed when the GFP library was applied to flow cytometry (Newman et al. 2006).

In some contexts, simple overexpression may shed light over the lacking information, but since the location of many proteins is a result of interactions with other cell components, a radical change in copy number could easily result in artificial observations. In other situations, the target output is the activity of specific weak promoters, e.g. in synthetic biological circuits, fluorescence-coupled biosensors or when developing promoter libraries. Several technologies permit the engineering of new promoters, e.g. responsive to other inducer molecules by hybridizing with upstream TF-binding sites (Blazeck and Alper 2013) or tuned to match fine, desirable transcription levels through mutagenesis of a strong native promoter (Nevoigt et al. 2006). Difficulties in GFP detection may have been a limitation in these developments for weaker promoter levels, though low expression may be phenotypically important for a wide range of synthetic biology purposes. In synthetic circuit designs, any concealed information on the shape of dose-response curves inhibits the analysis of mechanistic clues otherwise given by the response curvature (Ang et al. 2013). In applications of metabolite biosensors, background-covered signal levels means that the full regulatory capability cannot be utilized, e.g. limiting subsequent fluorescence-activated cell sorting (FACS). Ultimately, such autofluorescence could conceal properly functional GFP completely (Billinton and Knight 2001).

94 The efforts aimed at reducing the autofluorescence target two phenomena: Simple
95 medium autofluorescence arises from measuring fluorescence without isolating cells
96 from medium, e.g. in continuously growing cultures. These effects can be reduced by
97 the choice of medium or spectral unmixing by correcting for autofluorescence from a
98 wavelength representing effects of the culture medium (Lichten et al. 2014).
99 However, the cell autofluorescence is a more central issue, i.a. resulting from the
100 fluorescence of flavins and NAD(P)H (Billinton and Knight 2001). Cellular
101 autofluorescence also impacts techniques such as flow cytometry and microscopy and
102 the weak signal intensity must be amplified intrinsically to the cell.
103
104 Previous studies in mammalian cell lines have tackled the obstacle of cell
105 autofluorescence using directly repeated GFPs typically fused three to six times in
106 tandem using a small translational linker (Genové et al. 2005). By such approaches, it
107 has been possible to achieve good linear increments in fluorescence signals. However,
108 tandem repeats are problematic in organisms with proficient homologous
109 recombination such as *Escherichia coli* or *S. cerevisiae* where recombination between
110 DNA can happen within windows of identity at around 25 nucleotides (Ahn et al.
111 1988). This could explain why tandem GFP methods are avoided in these organisms.
112 However, even slight sequence divergence between repeats substantially decreases the
113 rate of recombination as seen in the case of recombination between 350 bp inverted
114 repeats, which was 4,600-fold reduced when sequence identity was reduced from 100
115 % to 74 % in *S. cerevisiae* (Datta et al. 1997). Similar effects occur in *E. coli* where
116 up to 1,000-fold reduction was observed following a reduction in repeat identity to 80
117 % (Rayssiguier et al. 1989).

Thus, in this study we present a simple methodology to take advantage of the ability to add sequence divergence to tandem proteins while maintaining function through variation in amino acid sequence as well as synonymous codon usage. By fusing three different GFP variants that vary mainly at nucleotide-level, we produce a new triple tandem GFP (3vGFP) stabilized towards direct-repeat recombination. We demonstrate the utility of 3vGFP through a genetically triggered promoter (ON/OFF) and developing and characterizing a new version of a Cu^{2+}-responsive promoter with reduced leakiness. Application of 3vGFP allowed visualization of weak signals that could not be separated from autofluorescence levels using the brightest individual GFP variant, superfolder GFP. Lastly, we test the stability towards recombination after culturing of the strain harboring 3vGFP through 25 generations.

Materials and methods

Materials

Unless otherwise stated, reagents were purchased from Sigma-Aldrich. Synthetic complete (SC) medium was prepared from 1.4 g/L synthetic complete drop-out mix lacking uracil, tryptophan, leucine and histidine (Y2001), 6.7 g/L yeast nitrogen base without amino acids (Y0626) and 20 g/L D-glucose, pH standardized to 5.6. When SC was supplemented with additional amino acids, 60 mg/L leucine, 20 mg/L uracil, 20 mg/L histidine-HCl and 20 mg/L tryptophan was added. Yeast Peptone Dextrose medium contained 20 g/L D-glucose.

Oligonucleotides were purchased from Integrated DNA Technologies.

Plasmids

The plasmids employed in this study are listed in Table 1.

Table 1 Plasmids employed in this study, describing whether they lead to chromosomal integration or propagate autonomously in *S. cerevisiae*.

Plasmid	**Expression cassette (promoter-ORF-terminator)**	**Maintenance in *S. cerevisiae* through**	**Reference**
pPR4-3vGFP	pSPAL10-3vGFP-tURA3	*CEN/ARS, HIS3*	This study
pPR4-sfGFP	pSPAL10-sfGFP-tURA3	*CEN/ARS, HIS3*	This study
pCU2-3vGFP	pCUP1dim -3vGFP-tURA3	*CEN/ARS, URA3*	This study
pCfB258-CUP1-3vGFP	pCUP1-3vGFP-tCYC1	Chromosomal integration	This study
pCfB258-CUP1-SPO13-3vGFP	pCUP1dim -3vGFP-tCYC1	Chromosomal integration	This study
pDS1U-X2-3vGFP	pTEF1-3vGFP	Chromosomal integration	This study
pEXP22	pADH1-GAL4AD-RalGDS-tADH1	*TRP1*	Life Technologies
pEXP32	pADH1-GAL4DBD-Krev1-tADH1	*LEU2*	Life Technologies
pRS413	-	*LEU2*	(Sikorski and Hieter, 1989)
pRS415	-	*HIS3*	(Sikorski and Hieter, 1989)

Strains

The strains analyzed in this study are listed in Table 2.

The following background strains were used to construct the strains:

Saccharomyces cerevisiae MaV203 (MATα*, leu2-3,112, trp1-901, his3Δ200, ade2-101, gal4Δ, gal80Δ,SPAL10::URA3, GAL1::lacZ, HIS3UAS GAL1::HIS3@LYS2,*

can1^{R}*, cyh2*^{R}) (Purchased from Life Technologies).

Saccharomyces cerevisiae PRa18 (MATα*, leu2-3,112, trp1-901, his3Δ200, ade2-101, gal4Δ, gal80Δ, GAL1::lacZ, can1*^{R}*, cyh2*^{R}) Derived from *S. cerevisiae* MaV203.

Saccharomyces cerevisiae PRa26*:* MATα*, leu2-3,112, trp1-901, his3Δ200, ade2-101, gal4Δ, gal80Δ, GAL1::lacZ, rad16::KanMX, can1*^{R}*, cyh2*^{R}. Derived from *S. cerevisiae* PRa18.

Saccharomyces cerevisiae CfB1010 (MAT*a; ura3-52; his3Δ1; leu2-3/112; MAL2-8*^{c}; *SUC2; are2Δ::loxP-KanMX; X-3::tHMG1*-P$_{TEF1}$-P$_{PGK1}$*-AtATR2*). Derived from *S. cerevisiae* CEN.PK 102-5B.

Table 2 *S. cerevisiae* strains analyzed in this study, indicating which plasmids or chromosomal integrations were introduced into the respective parental strains.

Strain name	Promoter	GFP	Plasmid #1	Plasmid #2	Plasmid #3	Integrative plasmid	Parent strain
PRa106	ON	3vGFP	pPR4-3vGFP	pEXP32	pEXP22	-	PRa26
PRa107	OFF	3vGFP	pPR4-3vGFP	pRS415	pEXP22	-	PRa26
PRa108	-	-	pRS413	pRS415	pEXP22	-	PRa26
PRa109	ON	sfGFP	pPR4-sfGFP	pEXP32	pEXP22	-	PRa26
PRa110	OFF	sfGFP	pPR4-sfGFP	pRS415	pEXP22	-	PRa26
CK24	pCUP1	3vGFP	-	-	-	pCfB258-CUP1-3vGFP	CfB1010
CK28	pCUP1dim	3vGFP	-	-	-	pCfB258-CUP1-SPO13-3vGFP	CfB1010
PRa114	pTEF1	3vGFP	-	-	-	pDS1U-X2-3vGFP	PRa18

Construction of 3vGFP plasmids

Plasmids were constructed by uracil-excision (USER) cloning. The general method for USER cloning was based on agarose gel-purification of the PCR products amplified using DNA polymerase X7 (Nørholm 2010). These were mixed in an equimolar 20 µL reaction with 0.5 µL USER enzyme (New England Biolabs) and 0.5

μL DpnI FastDigest (Thermo Scientific) in FastDigest buffer at 37 degrees C for 1-2 hours. Following 25 minutes at room temperature, 2.5 μL reaction was transformed into *E. coli*. Correctly cloned plasmids were identified using restriction analysis and DNA sequencing. The detailed use of oligonucleotides for assembly of all plasmids is described in Supplementary data.

Construction of strains

Plasmids and DNA for chromosomal targeting was introduced in *S. cerevisiae* by methods described previously (Gietz and Schiestl 2007). The PRa18 strain was constructed from the MaV203 strain by deletion of *SPAL10::URA3* through replacement with a *kanMX* gene deletion cassette flanked by loxP recombination sites from the pUG6 plasmid as described before (Güldener et al. 1996). DNA flanks to direct homologous recombination of the cassette to the chromosomal locus were generated by PCR on *S. cerevisiae* MaV203 gDNA spanning a fragment from 5'-CCATTCAACTAACATCACAC to 5'-CCTTCACCATAAATATGCC (upstream flank) and from 5'-CTCACAAATTAGAGCTTC to 5'-CCCATATCCAACTTCCAA (downstream flank). These flanks were cloned to the *kanMX* gene deletion cassette and transformed into yeast. The *kanMX* cassette was looped out by heterologous expression of Cre recombinase from the pSH47 plasmid (Güldener et al. 1996). To construct PRa26 subsequently, the chromosomal *HIS3* gene within the *rad16* locus was deleted using the same *kanMX* approach. The targeting flanks spanned regions from 5'- AGTTGGTACACCAGTTATACGG to 5'-AAAGCATAGGATACCGAGAAAC (upstream flank) and 5'-TGACATCACCCGAAAAGAAGC to 5'- GATTATGGTTACGATGTCGA (downstream flank).

To construct PRa114, the pTEF1-3vGFP construct was chromosomally integrated into the PRa18 strain using divisible selection (Rugbjerg et al. 2015). DNA fragments for integration was liberated from the vector pDS1U-X2-3vGFP by digestion with *Smi*I and transformed into yeast along with empty divisible selection plasmids pDS2 and pDS3 in order to reconstitute the selectable Ura^+ phenotype.

To construct respectively CK24 and CK28 from the CfB1010 strain, the pCUP1-3vGFP and pCUP1dim-3vGFP was chromosomally integrated by cloning into the EasyClone integrative vectors (Jensen et al. 2013). The DNA fragments for integration were obtained through *Not*I digestion of the vectors pCfB258-CUP1-3vGFP and pCfB258-CUP1-SPO13-3vGFP respectively, followed by agarose gel purification.

Estimation of TEF1-3vGFP fitness cost

Microtiter cultures of 200 µL YPD was inoculated by 100x backdilution of overnight YPD pre-cultures of PRa114 and PRa108, each inoculated from single colonies. The cultures were cultivated in a 96-well plate at 30 deg. C and continuous shaking in an ELx808 plate reader (BioTek), set to measure optical density every 15 minutes at OD_{630}. The plate was covered with a BreathSeal (Greiner Bio-one) and plastic lid. Growth rates were calculated for all three biological replicates by exponential regression between OD_{630} and time (hours) during the same OD_{630} span of exponential growth phase. All OD_{630} values were initially standardized to the time zero reading to account for differences in seal absorbance.

Cultivations for stability tests

The PRa114 strain was cultured from a single colony inoculated in 25 mL YPD medium and cultured at 30 deg. C and 250 rpm horizontal shaking in three parallel

lineages. By measuring OD_{600}, the number of generations passed was calculated. 2 % of the culture was passed to fresh medium and grown again until totally 25 generations had passed. For comparison between cultured population and reference strain, approx. 25 µL of each cell population was inoculated in YPD medium at the same time and cultured at 30 deg. C for 16 hours with 250 rpm horizontal shaking.

Fluorescence measurements

Pre-cultures in selective SC medium were inoculated from single colonies and cultures overnight at 30 deg. C. From these, 200 µL microtiter cultures of selective SC medium were inoculated and cultured at 30 deg. C with 300 rpm horizontal shaking in an Innova shaking incubator for 16 hours. As cover, the microtiter plates were covered with a BreathSeal (Greiner Bio-one) and a plastic lid.

The cell cultures were diluted approx. 1:100 in FACS flow buffer (BD Biosciences) and analyzed on a LSR Fortessa flow cytometer (BD Biosciences) equipped with a blue laser (488 nm) and set to measure 10,000 cells within a gate defined by forward and side scatter to capture all yeast cells. A FITC filter (530/30 nm) was used to measure GFP fluorescence reporting the area of the measured peaks. The laser voltage was adjusted to optimally utilize the dynamic range of detection. Data was processed and visualized as histograms with FlowJo version 10 (default settings) by overlaying the populations for each particular comparison.

Sequence alignment

Simple nucleotide and protein sequence alignment was performed using the ClustalO algorithm (Sievers et al. 2011).

Results and discussion

Amplification of fluorescence by tandems of differently encoded GFPs

To amplify the fluorescence signal of a GFP molecule while keeping transcription strength constant, the new 3vGFP protein was engineered by fusion of nucleotide sequences encoding yEGFP, GFP+ and sfGFP (Cormack et al. 1997; Pédelacq et al. 2006) (Fig. 1A). Two glycine residues were introduced as translational linker in each junction. The fluorescence of 3vGFP was evaluated when expressed from a weak *S. cerevisiae* hybrid promoter (p*SPAL10*) (Vidal et al. 1996) based on p*SPO13* to mimic low-expression applications (Huang and Schreiber 1997; Harton et al. 2013). The low-level strength of p*SPAL10* is attained by utilizing the UME6 repressor binding site naturally present within the *SPO13* promoter, which allows very low expression levels e.g. useful for control of cell growth. Further, GAL4-binding sites fused 179 bp upstream of start codon provide an upstream activating sequence, allowing transcription factor-based ON/OFF inputs.

The output fluorescence was first evaluated with single sfGFP (Fig. 1B), which is the individually brightest of the three GFPs tested. However, the fluorescence levels could not be distinguished from the control strain devoid of genes encoding GFP (PRa108). In contrast, the fluorescence of a strain (PRa106) carrying the gene encoding 3vGFP controlled by the same promoter was 3-fold higher than the background level and thus the level of the single sfGFP strain (Fig. 1B).

To test the utility of 3vGFP as output signal in a synthetic biology setting, we constructed versions of the strain with the p*SPAL10* promoter turned OFF. The

promoter is activated (ON) when a hybrid GAL4 activation domain binds a cognate hybrid GAL4 DNA-binding domain, which interacts with GAL4-binding sites of p*SPAL10*. The protein-protein interaction domains were based on the known Krev1 and RalGDS interaction domains (Herrmann et al. 1996). However omitting the DNA-binding domain prevents reconstitution of a functional transactivator (OFF). These ON/OFF effects of present DNA-binding domain remained hidden below the background levels of the sfGFP strains, while observable in strains with 3vGFP as output (Fig. 1B).

Figure 1

Stability towards recombination

Direct-repeat recombination in mitotic *S. cerevisiae* is reported to occur at rates between $5.8 \cdot 10^{-5}$ and $12 \cdot 10^{-5}$ per cell generation for repeats of several kilo base pair identity (Dornfeld and Livingston 1992). This recombination rate is linearly dependent on identity length at such long segments, however the rate drops rapidly below the minimal efficient processing segment (MEPS) length at around 250 bp in *S. cerevisiae* (Jinks-Robertson et al. 1993). While internal identity of 3vGFP ranges 74-84 % (Fig. 2B), the identical segments are maximally at a ten-fold shorter length than the MEPS.

To test the recombination stability of 3vGFP, we wanted to measure whether the fluorescence levels originating from 3vGFP would attenuate following repeated culturing. While the 3vGFP molecule is engineered to limit direct-repeat recombination, long-term cultivation could potentially still lead to this especially if favored by a concurrent fitness advantage. To test stability at high expression level, we therefore also chromosomally integrated *3vGFP* under control of the strong

promoter from *TEF1* i.e. at a level surpassing the intended use of 3vGFP. Expressing 3vGFP from the *TEF1* promoter caused a considerate cost in fitness of approx. 15 % in YPD, reducing the growth rate from an average of 0.35 hr^{-1} to 0.30 hr^{-1} compared to the negative control strain PRa108. Following culturing by serial passing (2 %) of liquid cultures for 25 generations of three parallel lineages, single-cell level analysis revealed that the average fluorescence level of the cell population had diminished by 7 percent, perhaps due to spontaneous direct-repeat recombination. The single cell-level visualization indicated a slight left-shift of the population (Fig 2A). These results exemplify that direct-repeat recombination can occur within 3vGFP in *S. cerevisiae* and if selected for, these effects can become significant. However, since 3vGFP is intended for use at levels of low expression, a fitness advantage is not likely to further drive diminished fluorescence at a typical utility of 3vGFP.

Figure 2.

Application of 3vGFP to construct an inducible promoter with reduced leakiness

Inducible promoters are important for development of e.g. synthetic genetic circuits, but the leakiness levels can be problematic in certain uses. To demonstrate the utility of 3vGFP, we therefore wanted to use it as output for genetic re-engineering of the popular Cu^{2+}-responsive promoter of *S. cerevisiae CUP1*. p*CUP1* has been employed in many different biotechnological cases (Labbé and Thiele 1999; Scholz et al. 2000; Rugbjerg et al. 2013), but displays considerable baseline activity (leakiness). p*CUP1* induction results from elevated Cu^{2+} concentrations mediated through binding of Cu^{2+} to the ACE1 transcription factor, which in turn binds to upstream activating sequence (UAS) elements of p*CUP1* (Huibregtse 1989; Evans et al. 1990) (elements schematically depicted in Fig. 3A). The leakiness level of p*CUP1* measured with 3vGFP corresponded to 2.5-fold the cell autofluorescence (Fig. 3B). Based on the regulatory mechanism of ACE1, we anticipated that trace levels of Cu^{2+} in the growth medium did not cause this leakiness, but rather assumed this basal transcriptional activity to be ACE1-independent. Accordingly, as strategy we hypothesized that swapping the promoter region downstream of ACE1 UASs for a transcriptionally repressed promoter could provide attenuation, while maintaining the response to ACE1-dependent induction. We therefore combined the upstream region of p*CUP1* (-149 to -454) containing three ACE1-binding sites, with part of the *S. cerevisiae* p*SPO13* (-1 to -157) including its UME6 repressor-binding site (Fig. 3A). This new promoter (p*CUP1dim*) controlling 3vGFP resulted in fluorescence that was reduced approx. 61 % (before background-subtraction) to levels close to the cell autofluorescence (Fig. 3B), while the promoter remained responsive to addition of Cu^{2+} (Fig. 3C).

Figure 3

The recombination-stabilized tandem GFP described in this study can enable characterization of minimally expressed genes in recombination-efficient organisms such as *S. cerevisiae* and other yeasts. As shown in this study, 3vGFP allowed characterization of the activation of a weak promoter and accordingly characterization of manipulations taking place at such low expression levels. Further, this particular approach of recombination-stabilizing GFPs with different protein and nucleotide sequences can be scaled in number. Recent brighter fluorescent proteins could be applied such as mNeonGreen (Shaner et al. 2013).

In principle, sequence divergence could be generated strictly at nucleotide level through codon optimization of segments encoding the same protein. Codon optimization can however introduce significant effects on the translation efficiencies (Goodman et al. 2013). Another concern may be spurious promoter/RBS activities, which could theoretically cause transcription and translation initiation from locations within the tandem GFP, thus producing truncated tandem proteins. Such situations would complicate the isolation of promoter responses and might require alleviation of the second and third GFP start codon.

An alternative method for assessment of promoter activities could be the use of the fluorescent RNA of the Spinach family, which bypasses the step of translation since the RNA forms the fluorescent signal (Paige et al. 2012; Pothoulakis et al. 2014). However, while the technology has potential for synthetic biological use, its general applicability remains to be seen, such as the detection limits for low expression levels. Further relevant, fluorescent *in* situ hybridization for RNA (RNA FISH) is a

technique allowing sensitive detection of transcripts at single-cell level (Zenklusen et al. 2008). This alleviates genetic engineering, but entails more sample treatment than for detection of GFP fluorescence.

In this study, a new simple strategy for engineering tandem fluorescent proteins was employed to produce brighter GFP signals with improved stability towards loop-out recombination. GFPs with sequence variation mainly at nucleotide level were translationally linked to form a recombination-stabilized tandem GFP molecule 3vGFP. Such GFPs could be useful for characterizing promoter activities in the range where normal single GFP signals fall below the cell autofluorescence levels. We specifically applied the 3vGFP molecule to characterize the ON/OFF levels of a weak promoter, which was not possible using a single sfGFP, and to develop a new hybrid Cu^{2+}-responsive promoter pCUP1dim with lower leakiness level. The plasmid pCU2-3vGFP encompassing the nucleotide sequence of 3vGFP and pCUP1dim will be deposited at the Addgene repository.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Novo Nordisk Foundation, the European Union Seventh Framework Programme (FP7-KBBE-2013-7-single-stage) under Grant agreement no. 613745, Promys, and Deutsche Bundesstiftung Umwelt,.

Acknowledgement

372 George Church is acknowledged for sfGFP encoded on pJ251-GERC (AddGene
373 plasmid 47441).
374

375

376

377

References

Ahn BY, Dornfeld KJ, Fagrelius TJ, Livingston DM. 1988. Effect of limited homology on gene conversion in a Saccharomyces cerevisiae plasmid recombination system. Mol. Cell. Biol. 8:2442–2448.

Ajikumar PK, Xiao W-H, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science 330:70–74.

Ang J, Harris E, Hussey B. 2013. Tuning Response Curves for Synthetic Biology. ACS Synth. Biol. [Internet]:547–567. Available from: http://pubs.acs.org/doi/abs/10.1021/sb4000564

Billinton N, Knight a W. 2001. Seeing the wood through the trees: a review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence. Anal. Biochem. [Internet] 291:175–197. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11401292

Blazeck J, Alper HS. 2013. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. Biotechnol. J. 8:46–58.

Cormack B, Bertram G, Egerton M. 1997. Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in Candida albicans. Microbiology [Internet] 143:303–311. Available from: http://mic.sgmjournals.org/content/143/2/303.short

Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S. 1997. Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc. Natl. Acad. Sci. [Internet] 94:9757–9762. Available from: http://www.pnas.org/content/94/18/9757.short

Dornfeld KJ, Livingston DM. 1992. Plasmid recombination in a rad52 mutant of Saccharomyces cerevisiae. Genetics 131:261–276.

Evans C, Engelke D, Thiele D. 1990. ACE1 Transcription Factor Produced in Escherichia coli Binds Multiple Regions within Yeast Metallothionein Upstream Activation site Sequences. Mol. Cell. Biol. [Internet] 10:426–429. Available from: http://mcb.asm.org/content/10/1/426.short

Gahlmann A, Moerner WE. 2014. Exploring bacterial cell biology with single-molecule tracking and super-resolution imaging. Nat. Rev. Microbiol. [Internet] 12:9–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24336182

Genové G, Glick B, Barth A. 2005. Brighter reporter genes from multimerized fluorescent proteins. Biotechniques [Internet] 39:814–822. Available from: http://www.biotechniques.com/article/05396BM02

Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. Nature [Internet] 425:737–741. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14562106

Gietz RD, Schiestl RH. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. [Internet] 2:31–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17401334

Goodman D, Church G, Kosuri S. 2013. Causes and effects of N-Terminal Codon Bias in Bacterial Genes. Science (80-.). [Internet] 342:475–479. Available from: http://www.sciencemag.org/content/342/6157/475.short

Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH. 1996. A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res. [Internet] 24:2519–2524. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=145975&tool=pmcentrez&rendertype=abstract

Harton M, Wingler L, Cornish V. 2013. Transcriptional Regulation Improves the Throughput of Three-Hybrid Counter Selections in Saccharomyces cerevisiae. Biotechnol. J. [Internet]:1–23. Available from: http://onlinelibrary.wiley.com/doi/10.1002/biot.201300186/abstract

Herrmann C, Horn G, Spaargaren M, Wittinghofer A. 1996. Differential Interaction of the Ras Family GTP-binding Proteins H-Ras, Rap1A, and R-Ras with the Putative Effector Molecules Raf Kinase and Ral-Guanine Nucleotide Exchange Factor. J. Biol. Chem. [Internet] 271:6794–6800. Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.271.12.6794

Huang J, Schreiber SL. 1997. A yeast genetic system for selecting small molecule inhibitors of protein-protein interactions in nanodroplets. Proc. Natl. Acad. Sci. U. S. A. [Internet] 94:13396–13401. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=28315&tool=pmcentrez&rendertype=abstract

Huh W-K, Falvo J V, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature [Internet] 425:686–691. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14562095

Huibregtse J. 1989. Copper-induced binding of cellular factors to yeast metallothionein upstream activation sequences. Proc. Natl. Acad. Sci. [Internet] 86:65–69. Available from: http://www.pnas.org/content/86/1/65.short

Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, Forster J, Nielsen J, Borodina I. 2013. EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. [Internet]:1–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24151867

Jinks-Robertson S, Michelitch M, Ramcharan S. 1993. Substrate length requirements for efficient mitotic recombination in Saccharomyces cerevisiae. Mol. Cell. Biol. 13:3937–3950.

Labbé S, Thiele D. 1999. Copper ion inducible and repressible promoter systems in yeast. Methods Enzymol. [Internet] 306:145–153. Available from: http://www.sciencedirect.com/science/article/pii/S0076687999060103

Li G-W, Xie XS. 2011. Central dogma at the single-molecule level in living cells. Nature [Internet] 475:308–315. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3600414&tool=pmcentrez&rendertype=abstract

Lichten C a, White R, Clark IBN, Swain PS. 2014. Unmixing of fluorescence spectra to resolve quantitative time-series measurements of gene expression in plate readers. BMC Biotechnol. [Internet] 14:11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3917901&tool=pmcentrez&rendertype=abstract

Miyawaki A. 2011. Proteins on the move: insights gained from fluorescent protein technologies. Nat. Rev. Mol. Cell Biol. [Internet] 12:656–668. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21941275

462 Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. 2006. Engineering of promoter
463 replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. Appl.
464 Environ. Microbiol. 72:5266–5273.

465 Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. 2006.
466 Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature
467 441:840–846.

468 Nørholm MHH. 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA
469 engineering. BMC Biotechnol. [Internet] 10:21. Available from:
470 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2847956&tool=pmcentrez&renderty
471 pe=abstract

472 Paige JS, Nguyen-Duc T, Song W, Jaffrey SR. 2012. Fluorescence imaging of cellular metabolites with
473 RNA. Science [Internet] 335:1194. Available from:
474 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3303607&tool=pmcentrez&renderty
475 pe=abstract

476 Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization
477 of a superfolder green fluorescent protein. Nat. Biotechnol. [Internet] 24:79–88. Available from:
478 http://www.ncbi.nlm.nih.gov/pubmed/16369541

479 Pothoulakis G, Ceroni F, Reeve B, Ellis T. 2014. The spinach RNA aptamer as a characterization tool
480 for synthetic biology. ACS Synth. Biol. [Internet] 3:182–187. Available from:
481 http://www.ncbi.nlm.nih.gov/pubmed/23991760

482 Raj A, van Oudenaarden A. 2009. Single-Molecule Approaches to Stochastic Gene Expression. Annu
483 Rev Biophys:255–270.

484 Rayssiguier C, Thaler DS, Radman M. 1989. The barrier to recombination between Escherichia coli
485 and Salmonella typhimurium is disrupted in mismatch-repair mutants. Nature 342:396–401.

486 Rugbjerg P, Myling-Petersen N, Sommer M. 2015. Flexible metabolic pathway construction using
487 modular and divisible selection gene regulators. Metab. Eng.:in press. Available from:
488 http://www.sciencedirect.com/science/article/pii/S1096717615001019

489 Rugbjerg P, Naesby M, Mortensen UH, Frandsen RJ. 2013. Reconstruction of the biosynthetic pathway
490 for the core fungal polyketide scaffold rubrofusarin in Saccharomyces cerevisiae. Microb. Cell
491 Fact. [Internet] 12:31. Available from:
492 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3654996&tool=pmcentrez&renderty
493 pe=abstract

494 Scholz O, Thiel A, Hillen W, Niederwieser M. 2000. Quantitative analysis of gene expression with an
495 improved green fluorescent protein. Eur. J. Biochem. [Internet] 267:1565–1570. Available from:
496 http://onlinelibrary.wiley.com/doi/10.1046/j.1432-1327.2000.01170.x/full

497 Shaner NC, Lambert GG, Chammas A, Ni Y, Cranfill PJ, Baird M a, Sell BR, Allen JR, Day RN,
498 Israelsson M, et al. 2013. A bright monomeric green fluorescent protein derived from
499 Branchiostoma lanceolatum. Nat. Methods [Internet] 10:407–409. Available from:
500 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3811051&tool=pmcentrez&renderty
501 pe=abstract

502 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M,
503 Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence
504 alignments using Clustal Omega. Mol. Syst. Biol. [Internet] 7:539. Available from:
505 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3261699&tool=pmcentrez&renderty
506 pe=abstract

507 Snapp EL. 2009. Fluorescent proteins: a cell biologist's user guide. Trends Cell Biol. [Internet]
508 19:649–655. Available from:
509 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2784028&tool=pmcentrez&renderty
510 pe=abstract

511 Vidal M, Brachmann RK, Fattaey a, Harlow E, Boeke JD. 1996. Reverse two-hybrid and one-hybrid
512 systems to detect dissociation of protein-protein and DNA-protein interactions. Proc. Natl. Acad.
513 Sci. U. S. A. [Internet] 93:10315–10320. Available from:
514 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=38381&tool=pmcentrez&rendertype
515 =abstract

516 Xie XS, Choi PJ, Li G-W, Lee NK, Lia G. 2008. Single-molecule approach to molecular biology in
517 living bacterial cells. Annu. Rev. Biophys. [Internet] 37:417–444. Available from:
518 http://www.ncbi.nlm.nih.gov/pubmed/18573089

519 Yu J, Xiao J, Ren X, Lao K, Xie X. 2006. Probing Gene Expression in Live Cells, One Protein
520 Molecule at a Time. Science (80-.). [Internet] 311:1600–1603. Available from:
521 http://www.sciencemag.org/content/311/5767/1600.short

522 Zenklusen D, Larson DR, Singer RH. 2008. Single-RNA counting reveals alternative modes of gene
523 expression in yeast. Nat. Struct. Mol. Biol. 15:1263–1271.

524

525

Figures

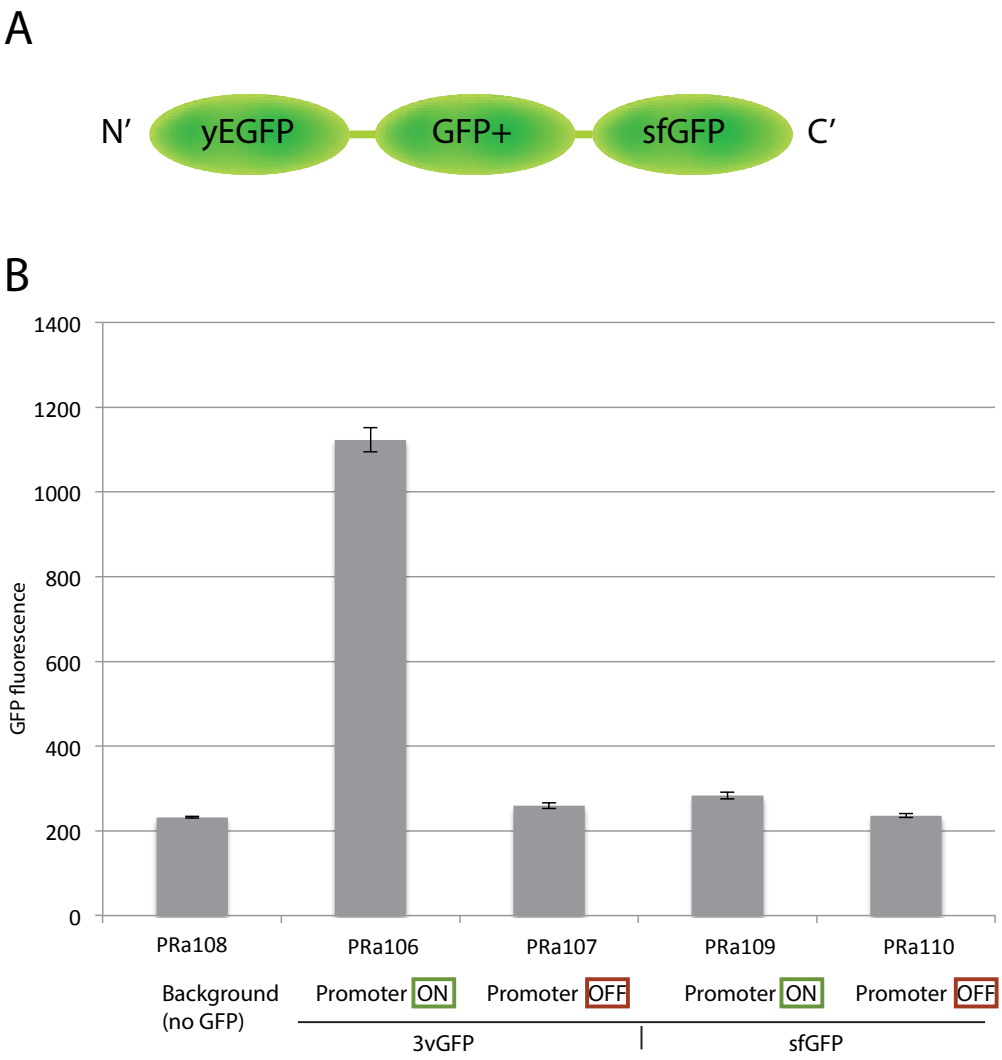
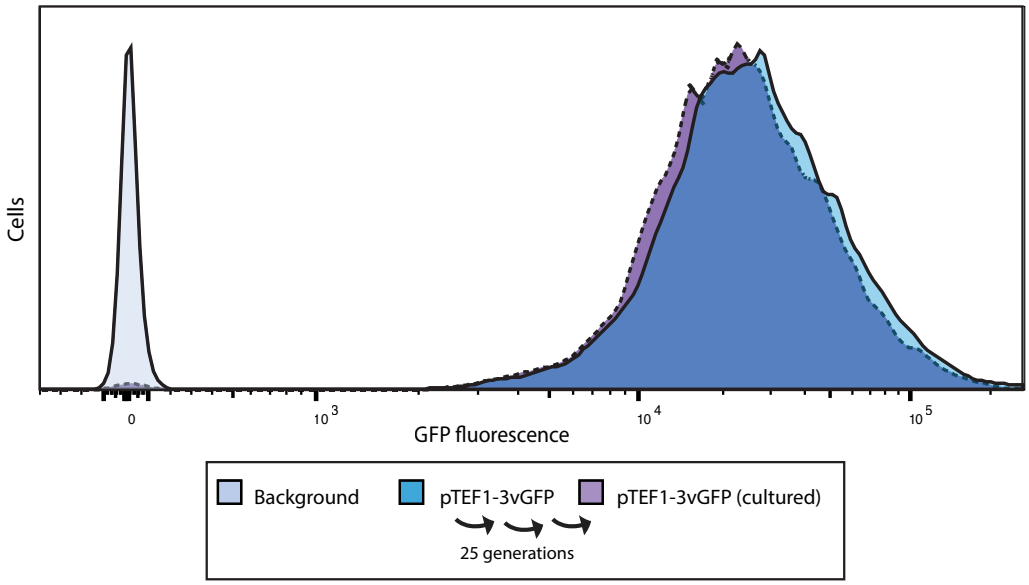


Figure 1 Increased GFP fluorescence signal above autofluorescence level by triple tandem GFP (3vGFP). A) Internal organization of individual GFP molecules fused as 3vGFP. 3vGFP consists of yeast-enhanced GFP (yEGFP), GFP+ and superfolder GFP. B) The *S. cerevisiae* strains carrying 3vGFP allowed the capture of the weak, ON/OFF promoter pSPAL10 unlike strains carrying a single sfGFP. The ON levels with single sfGFP corresponded to the background level of the empty control strain without GFP. The strains are described in detail in Table 2. Error bars depict standard error from biological replicates (n = 3).

A



B

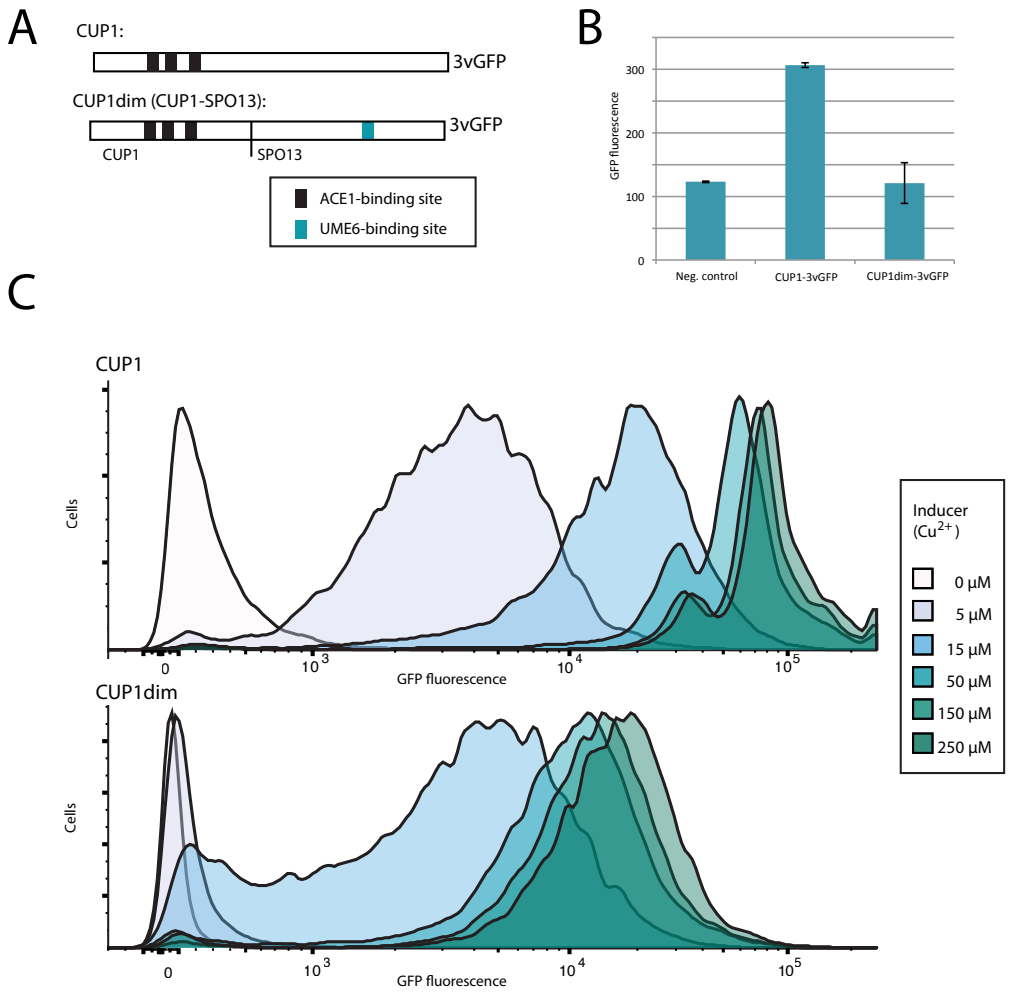
Nucleotide-level identity

1: sfGFP	100.00		
2: yEGFP	74.23	100.00	
3: GFP+	76.33	84.45	100.00

Protein-level identity

1: sfGFP	100.00		
2: yEGFP	94.12	100.00	
3: GFP+	94.96	96.64	100.00

Figure 2 Stability of the triple tandem GFP (3vGFP) towards loop-out recombination. A) Parallel lineages of a pTEF1-3vGFP *S. cerevisiae* strain was cultured for 25 generations and re-measured to verify stability towards loop-out recombination, compared to a background strain without GFP. Flow cytometry of representative example shown. Each sample contained 10,000 cells. The maxima of the samples are standardized to an equal top point. B) Sequence identities between the three direct repeats of sequences encoding GFP variants, as calculated by ClustalO.



546

547 **Figure 3 Development of weak Cu^{2+} -responsive promoter through**
548 **characterization with 3vGFP.** A) Organization of DNA-binding sites for the Cu^{2+}-
549 responsive ACE1 activator and UME6 repressor in the wildtype *CUP1* promoter and
550 the new dimmed, hybrid promoter p*CUP1dim*. B) OFF-level fluorescence measured
551 in absence of Cu^{2+} demonstrating the lower activity of the new hybrid promoter as
552 captured with 3vGFP. Error bars depict standard error from biological replicates (n =
553 3). C) Fluorescence of strain populations in response to addition of Cu^{2+}. Flow
554 cytometry of representative example shown. Each sample contained 10,000 cells. The
555 maxima of the samples are standardized to an equal top point.

556