FPCountR: Absolute quantification of fluorescent proteins for synthetic biology

Eszter Csibra and Guy-Bart Stan

Department of Bioengineering, Imperial College Centre for Synthetic Biology (IC-CSynB), Imperial College London, London, UK, SW7 2AY

Correspondence:
Eszter Csibra (e.csibra@imperial.ac.uk) and Guy-Bart Stan (g.stan@imperial.ac.uk)

Keywords: absolute quantification, calibration, engineering biology, fluorescent proteins, microplate reader, synthetic biology

Abbreviations: BCA, bicinchoninic acid; BSA, bovine serum albumin; EC, extinction coefficient; EDTA, ethylenediaminetetraacetic acid; FP, fluorescent protein, GFP, green fluorescent protein, MEFL, molecules of equivalent fluorescein, MEFP, molecules of equivalent fluorescent protein, OD, optical density; PEMS, particles of equivalent microspheres, SDS-PAGE, sodium dodecyl sulphate polyacrylamide gel electrophoresis; SEVA, Standardised European Vector Architecture

Abstract

This paper presents a generalisable method for the calibration of fluorescence readings on microplate readers, in order to convert arbitrary fluorescence units into absolute units. FPCountR relies on the generation of bespoke fluorescent protein (FP) calibrants, assays to determine protein concentration and activity, and a corresponding analytical workflow. We systematically characterise the assay protocols for accuracy, sensitivity and simplicity, and describe a novel 'ECmax' assay that outperforms the others and even enables accurate calibration without requiring the purification of FPs. To obtain cellular protein concentrations, we consider methods for the conversion of optical density to either cell counts or alternatively to cell volumes, as well as examining how cells can interfere with protein counting via fluorescence quenching, which we quantify and correct for the first time. Calibration across different instruments, disparate filter sets and mismatched gains is demonstrated to yield equivalent results. It can also reveal that mCherry absorption at 600nm does not confound cell density measurements unless expressed to over 100,000 proteins per cell. FPCountR is presented as pair of open access tools (protocol and R package) to enable the community to use this method, and ultimately to facilitate the quantitative characterisation of synthetic microbial circuits.

Introduction

There is a growing awareness that tackling the challenge of synthetic circuit design requires the synthesis of empirical characterisation data on genetic parts with mathematical modelling approaches for predicting and realising desired behaviours (Delvigne *et al.*, 2018; Beal *et al.*, 2020; Perrino *et al.*, 2021). However, there are numerous challenges in integrating experimental data with quantitative frameworks, as experimental data is typically acquired in relative or arbitrary units (specific to instruments and their respective settings), which cannot be converted into useful units and therefore limits our ability to make comparisons between experiments and models.

Fluorescent proteins (FPs) are our most versatile tools for the assessment of synthetic genetic element performance. Since their discovery, FPs have rightly been recognised as uniquely valuable reporter proteins for quantitative characterisation (Chalfie *et al.*, 1994; Rodriguez *et al.*, 2017), since they do not require the addition of exogenous components to fluoresce. This makes their use easy and cost effective. Various laboratory instruments allow the characterisation of fluorescent systems in a wide range of dimensions and scales – through direct visualisation (using fluorescence microscopy), via single-cell fluorescence analysis (using flow cytometry), or via timecourse kinetic data acquisition (using microplate readers).

The ‘protein quantification problem’ of reporting GFP levels acquired by such instruments in ‘relative fluorescent units’ (RFU) has been widely recognised in synthetic biology (Beal *et al.*, 2018). This recognition has led to the adoption of calibration standards such as fluorescein,

66 a small molecular fluorophore with similar excitation and emission characteristics to GFP.
67 Fluorescein can be used to calibrate a given instrument by converting the instrument's
68 arbitrary RFU output into units of 'molecules of equivalent fluorescein' (MEFL). This
69 technique has been demonstrated to enable the comparison of GFP expression data
70 gathered from different instruments (Fedorec *et al.*, 2020) as well as across laboratories
71 (Beal *et al.*, 2018).
72
73 While it is now approaching mainstream usage in synthetic biology, the conversion of green
74 fluorescence values into MEFL is arguably not the most important type of quantification
75 required for building synthetic circuits. Three aspects of the protein quantification problem
76 remain elusive. First, fluorescein is only a good calibrant for green fluorescent proteins,
77 leaving blue, yellow, orange and red FPs uncalibrated. Second, fluorescein can only provide
78 a conversion to units of fluorescein, whereas what is actually needed is a conversion to units
79 of protein. Currently, most experiments cannot even reveal the order of magnitude at which
80 FPs are being expressed (ie. 10 molecules per cell vs 100,000) – in contrast to what is
81 possible with RNA sequencing (Gorochowski *et al.*, 2019) – which limits our ability to
82 understand how protein abundance contributes to cellular burden (Ceroni *et al.*, 2015; Boo
83 *et al.*, 2019). Third, while fluorescein can allow the comparison of GFP levels between
84 instruments and laboratories, it cannot address the comparison between two different FPs
85 in the same circuit. This is only possible if RFUs from both FPs can be separately converted
86 into molecular units of protein. Some have attempted to tackle this by attempting to predict
87 the relative brightness of FPs using theoretical values (Boada *et al.*, 2019; Vignoni *et al.*,
88 2019), but such calibrations make a number of unvalidated assumptions, for instance about

89 translation rate equivalence across constructs, that require rigorous testing before they can
90 be adopted.
91
92 Fortunately, there is a reasonably simple solution. The ideal calibrant in molecular biology is
93 considered to be a purified sample of the molecule to be measured – in this case, the
94 fluorescent protein itself. While purified FPs are not commercially available, they can be
95 produced 'in-house', thereby providing that crucial direct link between relative fluorescence
96 units and molecules of protein. Indeed, others have proposed protocols for such calibrations
97 (Taniguchi *et al.*, 2010; Finan *et al.*, 2015; Hirst *et al.*, 2015; Tie *et al.*, 2016), although such
98 practice remains rare in the literature. We suspect this is due to (a) an underappreciation
99 that absolute quantification is possible without 'omics', (b) a reticence to try unfamiliar
100 biochemical protocols that are not usually part of the synthetic biology repertoire, and (c)
101 doubt that such protocols could be accurate or sensitive enough for general usage. We
102 suspected that many of these assertions were due to misunderstandings about protein
103 purification methodologies and the biochemical characteristics of fluorescent proteins in
104 general, and this motivated us to develop a general, yet simple-to-use calibration protocol
105 of this nature.
106
107 In what follows, we outline our optimised calibration method and present it as a pair of
108 resources: a wet lab protocol, called FPCount (available on protocols.io) and an
109 accompanying analysis package, called FPCountR (available on GitHub). We present data
110 showing the development of this protocol, and systematically characterise the biochemical
111 and analytical requirements of an accurate and sensitive calibration. We also present a
112 'hack', which acts both to simplify the method to remove the requirement for protein

113 purification, and to make it more sensitive and robust. Using FPCountR, we show that
114 conversion to molecular units can be used to calibrate across different instruments,
115 disparate filter sets and mismatched gains to yield equivalent results. We demonstrate that
116 conversion to absolute units allows the user of our method to compare the protein
117 production efficiency of different fluorescent proteins expressed from an otherwise
118 identical vector in molecules per cell, or as a molar concentration. Finally, we demonstrate
119 that this method can be used to quantitatively evaluate the experimental protocols
120 themselves, such as the extent to which red fluorescent proteins confound optical density
121 readings in timecourse assays.
122
123 **Results**
124
125 Our aim for this work was to develop a generalisable method for FP calibration that could be
126 used by any group wishing to calibrate fluorescence readings on microplate readers to
127 molecular units. To do this, we defined a number of key aims for our proposed method.
128 First, it should be accurate and sensitive, as we need the method to correctly estimate
129 molecule numbers within cells, and as protein yields from small-scale purifications are
130 typically modest. Second, the calibration protocol should be as simple as possible and
131 adapted ideally such that each respective assay may be carried out in 96-microwell plate
132 format using the same plate reader that is being calibrated. This way, multiple fluorescent
133 proteins may be calibrated at once, and end users do not require any additional
134 instrumentation. Third, the method should be suitable for the particular characteristics of
135 fluorescent proteins. These proteins are smaller and structurally distinct from typical
136 protein calibrants such as bovine serum albumin (BSA), and are known to present certain

137 challenges for quantification due to light absorption by their chromophores. Thus, any assay
138 developed for non-fluorescent proteins requires a separate validation on FPs to
139 demonstrate that they are also adequate for this class of proteins. Finally, we wanted to
140 enable the easy analysis of the data, by (i) enabling easy conversion of raw calibration data
141 into a conversion factor that links the arbitrary fluorescence output of a protein with its
142 quantity in molecular units, and (ii) allowing easy conversion of data from all future
143 timecourse data from that instrument to produce outputs in (e.g. GFP) molecules, rather
144 than relative (e.g. green) fluorescence units. An overview of the FPCount fluorescent protein
145 calibration protocol is illustrated in Fig. 1.
146
147 **Design of purification protocol for obtaining protein calibrants**
148 In order to obtain our fluorescent protein calibrants, they must first be produced by
149 overexpression, and purified. Protein purification methodologies can often be highly
150 complex, requiring specialist expertise and instrumentation. To alleviate this, our protocol
151 was explicitly designed to be as straightforward as possible, involving only the minimum
152 number of required steps, and using commonly available reagents. It was also designed to
153 be amenable to small-volume purifications to enable the calibration of multiple proteins in
154 parallel. The protocol is summarised in Fig. 2A. A standardised FP expression vector was
155 constructed from an arabinose-inducible His-tagged FP construct in a high-copy SEVA vector
156 (Fig. 2B). Overnight expression was used to maximise protein production, and cells were
157 lysed using sonication to avoid the requirement to add chemical components that may
158 interfere with downstream processes, such as EDTA (with His-tag purification) or detergents
159 (with protein quantification). Insoluble proteins were removed via centrifugation and
160 proteins were purified using cobalt resin to minimise the co-isolation of impurities.

161 Expression yields and solubility were checked by SDS-PAGE (Fig. 2C) and the quality of
162 purified FP calibrants was verified by SDS-PAGE and fluorescence excitation and emission
163 scanning (Fig. 2D-E, Supplementary Fig. 1). Purified calibrants were of consistently good
164 purity and yield.
165
166 **Conducting a plate reader calibration**
167 The calibration of plate readers with fluorescein has traditionally been conducted using a
168 dilution series of known concentrations of fluorescein, subjected to a fluorescence assay in
169 the plate reader whose calibration is desired (measurement of relative fluorescence units,
170 RFU). The results can be used to relate fluorescein molecule number to RFU to obtain a
171 conversion factor, which can in turn be used to convert RFU readouts from experimental
172 data into MEFL units (Beal *et al.*, 2018). For protein calibrants, one additional step is
173 required: protein concentration determination.
174
175 In our initial protocols, we opted for the bicinchoninic acid (BCA) assay due to its sensitivity,
176 ease of use and low protein-to-protein variability (Noble *et al.*, 2007). In addition,
177 microplate-optimised reagents for a 'microBCA' were available from ThermoFisher, with
178 excellent reported sensitivities (to 2 ng/µl). Pilot tests showed an inhibitory effect of the Tris
179 and NaCl in the elution buffer (Supplementary Fig. 2A) suggesting that buffer exchange
180 would be necessary for high assay sensitivity. Background signal was also observed from
181 protease inhibitors in the buffer, but this didn't seem to disrupt quantification accuracy.
182 Normalised BSA values were fitted to a polynomial equation to obtain a standard curve
183 (Supplementary Fig. 2D, ii). This was then used to predict the concentrations of FP dilutions
184 by first removing values under the reported threshold of sensitivity and fitting a linear

185 model through the rest of the values (Supplementary Fig. 2D, ii and iii). An extra step was
186 added to the recommended protocol (Supplementary Fig. 2A, ii and Supplementary Fig. 2C)
187 to account for baseline absorbance of red FPs in the A562 range. The calculations are
188 handled by the get_conc_bca() function. Using these predicted protein concentrations, an
189 adapted version of the generate_cfs() function from flopR (Fedorec *et al.*, 2020) was used to
190 generate conversion factors for mCherry in a Tecan Spark plate reader for the RFP-typical
191 filter set (ex 560/20, em 620/20), showing that conversion factors from RFU to molecules
192 may be obtained using the FPCount protocol analogously to obtaining conversion factors
193 from RFU to MEFL for fluorescein. Example outputs are shown in Supplementary Fig. 3.
194
195 **Development of a second assay for FP concentration – the A280 assay**
196 We sought to verify the accuracy of the BCA assay by re-quantifying our FPs with a second
197 method that is likely to give reliable concentration estimates. The most commonly used
198 assays for protein quantification include SDS-PAGE-based Coomassie assays and Bradford
199 assays (Bradford, 1976; Noble, 2014), both based on the same reaction, which has been
200 reported to possess more protein-to-protein variability than the BCA assay (Tal *et al.*, 1985;
201 Noble *et al.*, 2007; Thermo Fisher Scientific, 2017). It has also been suggested that
202 proteomics could be used for absolute quantification. However, proteomic data is also not
203 inherently 'absolute' – to obtain accurate absolute protein counts users typically need to
204 provide a calibrant of known concentration that is treated identically to the sample
205 (Wasinger *et al.*, 2013; Calderón-Celis *et al.*, 2018). Once again, the calibrant should ideally
206 be identical to the protein we wish to count, i.e. purified FP. Using proteomics to validate
207 our BCA results for our FP calibrants would therefore be a logically circular process that
208 wouldn't allow us to validate our method.

While a wide variety of protein assays exist, the only widely-used 'absolute' assay that does not require a calibrant is the A280 assay. As the name suggests, it quantifies protein concentration via light absorbance at 280 nm, where three amino acid residues are known to absorb light in a way that has been shown to be approximately additive (Pace *et al.*, 1995). This finding means that a reasonable prediction of light absorbance at 280 nm can be made for any pure protein of known primary sequence by way of an extinction coefficient (EC; expected light absorption for a given concentration of protein). As sample absorbance relates to molecular concentration according to Beer's law, i.e., $A = EC * C * L$, (where A is the absorbance, EC is the extinction coefficient ($M^{-1}cm^{-1}$), C is the concentration (M), and L is the path length (cm)), the protein concentration may be calculated from absorbance using only the extinction coefficient and the path length. The most common formats for A280 measurements are laborious, single-throughput cuvette- or Nanodrop-type measurements, requiring the adaptation of the standard A280 protocols for use in 96-well microplates. While many of our findings about how to achieve this adaptation successfully are available in the literature if one knows where to look, we have summarised the requirements in Supplementary Note 1, in the hope these will be useful to readers unfamiliar with these sources (see also Supplementary Fig. 4-8). In brief, the best results for A280 assays are obtained by using UV-clear plastic, removing additives such as imidazole, correcting for path length variation, and correcting for light scatter via analytical methods. This requires the collection of an absorbance spectrum from 200-1000 nm rather than just one reading at 280 nm, and is processed by two consecutive R functions, plot_absorbance_spectrum() and get_conc_a280() (Supplementary Fig. 6-7).

233 **Systematic testing of the calibration protocol allows method validation**

234 We sought to conduct a systematic assessment of the BCA and A280 methods by testing
235 three spectrally distinct FPs in two buffers, assessed with both assays in parallel (Fig. 3). The
236 chosen FPs (mTagBFP2, mGFPmut3 and mCherry) are widely used, monomeric, reasonably
237 fast-maturing and bright. All three are almost identical on the protein level to their FPbase
238 entries (Fig. 2B; Materials & Methods; Supplementary Data), with the exception that they all
239 have a His_6 tag N-terminal extension, and mGFPmut3 includes a well-defined monomeric
240 A206K mutation (Zacharias *et al.*, 2002; Shaner *et al.*, 2005). The chosen buffers (T5N15
241 (5mM Tris-HCl pH 7.5, 15 mM NaCl) and T5N15 with protease inhibitors) were both
242 compatible with the microBCA assay (Supplementary Fig. 2B), however, pilot studies
243 suggested they might have different effects on the A280 assay. Following purification, FPs
244 were initially dialysed to remove additives, then re-dialysed into the assay buffer (Fig. 3A).
245 Each FP:buffer combination was then serially diluted, and subjected to an absorbance scan
246 (200-1000 nm measurement, for the A280), a fluorescence assay (fluorescence
247 measurement with appropriate filters for each FP) and a microBCA assay (reagent addition,
248 incubation and A562 nm measurement). The results of this comparative test are shown in
249 Fig. 3B-C, Supplementary Fig. 9 and Supplementary Tables 1-2. Broadly, the results from
250 each assay validate those of the other assay: the measured concentration of each FP using
251 the microBCA and A280 assays are within 2-fold of each other for most samples (Fig. 3C) and
252 apparent linear ranges reach ~1 ng/µl for each assay (Fig. 3B), an impressive result for the
253 A280 assay considering that the reported sensitivity on the Nanodrop is ~100-fold higher
254 (Thermo Fisher Scientific, 2010). However, it can be noted that there is some buffer
255 sensitivity for both assays – the microBCA produces more linear results in the buffer
256 containing protease inhibitors, whereas the A280 does better in the buffer without them.

257 Conversion factors obtained from different purification batches gave similar estimates
258 where concentration estimates were made using optimal assay:buffer pairings, varying
259 between 1.02 and 1.63-fold (Supplementary Fig. 9B).
260
261 **FPbase-enabled ECmax method performs better than conventional assays**
262 We trialled a third protein assay during this experiment, designated here as the 'ECmax'
263 method. The principal idea behind this assay is that the A280 extinction coefficient,
264 recorded in units of $M^{-1}cm^{-1}$, that relates a protein's absorbance at 280 nm to its
265 concentration, is not the only known extinction coefficient for FPs. FPs also possess an
266 extinction coefficient ('EC') corresponding to their light absorption at their peak ('max')
267 excitation wavelength. As the 'ECmax' of most FPs is available on FPbase, we can automate
268 its retrieval using the FPbase API (Lambert, 2019). The analytical processing steps for the
269 ECmax assay (in get_conc_ecmax()) are similar to those of the A280 assay (Supplementary
270 Fig. 8), and require no further readings other than those completed for the A280 assay. As
271 the maximal absorbance peaks for all FPs tested were higher than those at 280 nm
272 (Supplementary Fig. 1B, Supplementary Table 3), we anticipated that the ECmax assay
273 would be more sensitive. Further, as protease inhibitors absorb at wavelengths under 300
274 nm, we hypothesised this assay may be less buffer-sensitive than the A280. Compellingly,
275 both appear to be true: we consistently found that the ECmax assay produced larger linear
276 ranges and lower limits of detection than the other assays (approaching 0.1 ng/µl, 10-fold
277 better than the A280 assay and 1000-fold better than a Nanodrop), and that it produced
278 almost indistinguishable results whether or not the buffer contained protease inhibitors
279 (Fig. 3B-C, Supplementary Fig. 9, Supplementary Table 1). In addition, predictions from the
280 ECmax assay closely match those from the A280, typically predicting concentrations

281 matching at 80-100 % those of the expected result (rather than 170-220% for microBCA; Fig.
282 3C, Supplementary Table 3). For these reasons, we propose that the ECmax assay would be
283 the most robust assay to include in a simple calibration protocol and will proceed using
284 concentrations calculated from ECmax assays in what follows.
285
286 **Fluorescence assays are not linear in all buffers**
287 Along with protein concentration, fluorescence assays were carried out for each FP to
288 obtain relative fluorescence units to relate to FP molecules in each well. One unexpected
289 observation from this experiment was that fluorescence assays produce textbook linear
290 responses to concentration in buffers containing protease inhibitors, but without these, a
291 steep fluorescence drop is observed, particularly below ~10^{12} molecules per well (Fig. 3D).
292 On first encounter, this result can appear as though it were caused by pipetting error.
293 However, for this experiment, microplates containing dilution series using both buffers were
294 prepared in parallel by the same person, at the same time, using the same equipment, and
295 the result was reproducible across different days and FPs. The lack of linearity is suspected
296 to be due to FP destabilisation in T5N15 buffer that doesn't contain any stabilising agents. In
297 such case, we suspect protease inhibitors could act to provide stabilisation in the T5N15+pi
298 buffer. Alternatively, it may be due to protein adsorption to the plastic at very low
299 concentrations (Aitken and Learmonth, 2002), which can be successfully avoided by using a
300 buffer containing protease inhibitors. In either case, the observed non-linearity results in an
301 underestimation of the conversion factor in these buffers that 'scales' with gain and may
302 result in an underestimation of conversion factor values of up to 8-fold (Fig. 3E), suggesting
303 that the removal of protease inhibitors is inadvisable for an accurate calibration.
304

305 **ECmax assay enables FP calibration without protein purification**

306 Having established that the ECmax assay, a protein quantification assay that relies only on
307 the peak light absorbance of each FP, is highly accurate and sensitive for purified proteins
308 measured using trusted methods (Fig. 3), we asked whether this method could enable us to
309 drop the purification step altogether. Dropping the purification step is not possible for the
310 other assays as they are designed to quantify any and all proteins, but the ECmax should in
311 principle be specific for the considered FP and may therefore be used to quantify FPs in
312 crude lysates. This requires that cell lysates, complex mixtures of all the soluble proteins
313 from a lysed cell, do not significantly absorb at the relevant wavelengths. To investigate this,
314 we harvested and lysed cells expressing our three FPs, separated the soluble fraction, and
315 concentrated it (~16-fold). Putting these through an ECmax assay and fluorescence assay,
316 we observed almost identical conversion factor values as from our purified proteins
317 (Supplementary Fig. 10). Of two lysis methods – sonication and use of a commercial lysis
318 reagent – the former produces more accurate results. It therefore appears that by
319 establishing the ECmax assay for FP quantification, it is possible to remove the purification
320 step altogether without compromising calibration accuracy.

321

322 **Calibration allows comparison across instruments for FPs other than GFP**

323 While fluorescein enables the comparison of experimental results across laboratories and
324 instruments by converting arbitrary units into 'molecules of equivalent fluorescein' (MEFL)
325 units, FP calibration in principle offers the same capability by converting arbitrary units into
326 units of 'molecules of equivalent FP' (MEFP), which is carried out using our R package's
327 process_plate() function (Fig. 4A-B). This allows us to quantify the number of FP molecules
328 in each well of our microplates. To calculate uncalibrated 'per cell' values, typical studies

329 will divide the RFU values by the optical density (OD600 or OD700) of the culture, which
330 quantifies cell density. The calibration of optical density to particle number can be achieved
331 through a similar calibration process using microspheres of similar size to *E. coli* (Stevenson
332 *et al.*, 2016; Beal *et al.*, 2020). Using both calibrations, it is possible to quantify molecule
333 number per cell in 'molecules of equivalent FP per particles of equivalent microspheres'
334 (MEFP/PEMS, Fig. 4C), units which should allow cross comparison between different
335 instruments, gains and filter sets. To test this, overnight cultures of *E. coli* containing
336 mCherry expression vectors were split into separate but identical microplates containing
337 arabinose, and were grown in two plate readers using a range of settings. The results show
338 that normalised values of relative fluorescence differ by ~1.5, ~3 and ~130-fold without
339 calibration, whereas such values become reliably comparable after calibration, even for
340 experiments conducted using instrument settings that produce values that cannot be legibly
341 plotted on the same axis (Fig. 4D).
342
343 **Calibration allows estimation of absolute cellular protein concentration**
344 We next asked if calibration to units of MEFP/PEMS is a reasonable approximation for
345 molecule number per cell (Fig. 5A). We carried out microsphere calibration using 1 cm
346 cuvettes and a standard spectrophotometer and obtained conversion factors
347 (Supplementary Table 4) that fell within the range quantified by empirical OD600-specific
348 cell counts (Volkmer and Heinemann, 2011). In addition, other authors have confirmed that
349 values of fluorescent protein per cell using fluorescein and microsphere calibrants
350 (MEFL/PEMS) are approximately equal to those obtained using fluorescein-calibrated single
351 cell data on a flow cytometer (Beal *et al.*, 2020; Fedorec *et al.*, 2020), suggesting particle
352 counts are likely to approximate actual cell numbers. However, we observed a major caveat

353 to the use of microspheres as calibrants, which is that their absorbance profiles differ from
354 that of cells. This can frustrate their ability to provide accurate conversions between 'per
355 cell' data calculated using OD700 versus OD600 measurements (Supplementary Fig. 11 and
356 Supplementary Tables 4-5) for which we discuss solutions in Supplementary Note 2.
357
358 The question of whether the measured fluorescence of FPs in cells is the equivalent of
359 measured fluorescence of the same number of FPs *in vitro* is less clear. Some authors have
360 found that cells attenuate (or 'quench') fluorescence (Zhang *et al.*, 2009; Hirst *et al.*, 2015),
361 but the magnitude of the effect has not been systematically investigated, particularly for
362 modest cell concentrations found in a typical *E. coli* growth assay. We quantified the
363 quenching properties of *E. coli* cells on our three FPs by mixing an increasing concentration
364 of non-fluorescent cells with purified FPs, and quantifying the difference in apparent
365 fluorescence with added cells (Fig. 5B and Supplementary Fig. 12). Our results suggest that
366 this 'quenching' effect amounts to less than 20 % of the fluorescence signal for moderate
367 cell densities (OD600/cm under 0.5), but increases to ~30 % for the highest cell densities
368 typically observed in microplate-scale cultures (OD600/cm around 2). This information was
369 used to add a correction step into the process_plate() calculations so as to compensate for
370 the expected percentage loss of fluorescence with increasing cell density (Fig. 5B, right
371 panel). The complete analytical workflow from calibration to experimental data processing
372 is illustrated in Supplementary Fig. 13.
373
374 Using these amendments, we can show that it is possible to convert response curve assay
375 data into molecules per cell. Figure 5C shows one experiment using mCherry expression
376 construct in two vectors with different origins of replication. Using these vectors, we obtain

377 figures for mCherry abundance that vary between ~3,000 to 80,000 molecules per cell for
378 p15A, but ~1000 to 600,000 molecules per cell for colE1. Protein abundance information,
379 available from proteomics and ribosome profiling studies, suggests that the typical *E. coli*
380 protein is present in the order of 100s-1000s copies per cell, and the most abundant are
381 present in the order of 10^5 copies per cell or higher (Ishihama *et al.*, 2008; Arike *et al.*, 2012;
382 Li *et al.*, 2014). Over-induction using the high-copy (colE1) vector therefore appears to allow
383 synthetic protein expression to reach the level of the most abundant proteins in the cell.
384 This is supported by the fact that these vectors reliably overexpress FPs to a level
385 observable by SDS-PAGE in unpurified lysates (Fig. 1, Supplementary Fig. 1). Modest
386 expression (in the order of 10^2-10^3 per cell) can be achieved by combining low arabinose
387 concentrations with either vector. Leaky expression can be quantified in the colE1 vector as
388 contributing hundreds of mCherry molecules per cell even without induction, whereas in
389 the p15A vector it was merely ten copies per cell. In other words, using an average colE1
390 vector allows us to utilise the full spectrum of protein abundances from modestly expressed
391 enzymes such as the RecBCD helicase (~100 copies per cell) through to the most abundant
392 ribosomal proteins (~100,000 copies per cell, Li *et al.*, 2014). We can also use this to
393 compare the number of molecules produced from the same vectors but two different FPs.
394 Interestingly, measurements from identical SEVA vectors revealed that while the FP
395 abundances are in the same order of magnitude, mTagBFP2 accumulated to higher levels
396 per cell than mCherry by 1.83-fold on average, despite sharing the same promoter, 5'
397 untranslated region, ribosome binding site and N-terminal protein sequences (Fig. 5C-D).
398 Larger discrepancies were observed with colE1 which induced a relatively high leak in
399 expression observable at low arabinose levels, and resulted in higher mCherry accumulation
400 than mTagBFP2 at maximum induction.

401

402 In 2011, Volkmer and colleagues noted that while OD-specific *E. coli* cell counts varied with
403 growth rate, the OD-specific total cell volume was approximately 3.6 µl per OD600/cm,
404 regardless of strain or growth condition. Using OD as a measure of the cumulative cellular
405 volume in a culture can therefore be used to convert fluorescence and OD measurements
406 into concentrations in molar units, instead of 'per cell' values, and such conversions may be
407 more appropriate for comparing experimental results with quantitative modelling of cellular
408 reaction networks, since it is unaffected by growth rate differences. Using this method, we
409 find that FP abundances using the same vectors populate a range of concentrations
410 between 0.01-100 µM (Fig. 5D).

411

412 **Absolute quantification reveals hidden properties of fluorescent proteins**

413 Finally, we were interested in testing whether the quantification of fluorescent proteins in
414 cells could shed more light on the question of how to avoid quantification errors in a typical
415 experiment. The presence of red FPs has been suggested to interfere with bacterial cell
416 density estimations at 600 nm since red FPs typically absorb well at this wavelength (Hecht
417 *et al.*, 2016) and has led to the conclusion that circuits using RFP must be quantified at 700
418 nm, which is unaffected by their presence. However, the number of molecules of red FPs
419 that might be required for this effect to occur has never been quantified.

420

421 Calibrated timecourse data of mCherry overexpression in *E. coli* (Fig. 6) was examined to
422 quantify these effects, with mGFPmut3 and mTagBFP2 used as negative controls. The ratio
423 between OD600 and OD700 measurements was used to identify errors caused by red FP
424 absorbance. Linear models fitted to the relationship between measured OD600 and OD700

425 values confirmed that this relationship was very similar for all uninduced cells (Fig. 6A;
426 OD600 = 1.30 * OD700 – 0.02), but mCherry induction resulted in a measurable deviation
427 (OD600 = 1.37 * OD700 – 0.03, Fig. 6A-C): similar to what Hecht and colleagues measured
428 (Hecht *et al.*, 2016). However, the relationship between this shift and cellular protein copy
429 number was not previously investigated. Looking at this relationship (Fig. 6D), our results
430 indicate that the error due to measuring cell density using OD600 values for mCherry is only
431 apparent when mCherry levels per cell are high (over 100,000 per cell), and that the
432 magnitude of this difference reaches only ~5 %. In other words, the effect is smaller in
433 magnitude than errors from cell quenching, and only affects systems where mCherry is
434 overexpressed. In contrast, mGFPmut3 expression to similar levels has no effect on cell
435 density estimation using OD600, as expected. Surprisingly, we observed the opposite trend
436 for mTagBFP2, in which, assuming OD700 produces a correct estimate, OD600 appears to
437 underestimate cell density where mTagBFP2 is expressed at high levels per cell. The reasons
438 for this are currently unclear and beyond the scope of this paper, but our results suggest
439 that this could be an interesting avenue for future work. Generally, this experiment
440 confirmed that the presence of low to moderate FP levels per cell (under 100,000) do not
441 perturb cell density estimates, and errors are of a lower magnitude in all cases than those
442 from cellular fluorescence quenching.
443
444 **Discussion**
445
446 Our aim for this work was to develop a generalisable method that allows fluorescence
447 readings on microplate readers to be calibrated to molecular units of fluorescent protein.
448 The method ought to be (1) accurate and sensitive, (2) as simple as possible, (3) suitable for

449 any fluorescent protein, and (4) easily analysed. To develop the method, we adopted the
450 principles of redundant experimental design, including the validation of multiple assay
451 types, characterisation of the method's consistency, and the assessment of its generality for
452 three different FPs (Casadevall and Fang, 2016).
453
454 Our initial method using a simple purification protocol (Fig. 2) and a commercial protein
455 assay allowed us to develop an analysis pipeline to obtain conversion factors from purified
456 FP calibrants. To demonstrate the accuracy and validity of our calibrations, we verified that
457 the absorbance and fluorescence spectra of the calibrants matched their counterparts on
458 FPbase (Fig. 1, Supplementary Fig. 2; Raynal *et al.*, 2014; Lambert, 2019), and validated our
459 initial protein assay measurements by cross comparison with two further methods. To do
460 this, we adapted the low-throughput A280 assay into an accurate, high-throughput assay
461 format, and showed that these were suitable for use with FPs even though some absorb in
462 the near-UV range. This type of assay for FP quantification has not, to our knowledge, been
463 demonstrated in the literature before, and we contend that it is likely to be of particular
464 interest since it requires no calibrant or commercial reagent, no expensive quartz-based
465 consumables and exhibits a sensitivity that exceeds that of commercial systems such as the
466 Nanodrop.
467
468 We also discovered a methodological 'hack' to obtaining FP concentrations using the
469 extinction coefficients at their maximum excitation wavelength, the ECmax assay, which was
470 both the simplest and the most robust of all the assays tested. Specifically, the ECmax assay
471 was the least affected by buffer conditions, and had the largest linear range (to ~ 0.1 ng/µl;
472 Fig. 3). We note that the assay is limited by the fact that it requires the used FP to be

473 documented on FPbase, and assumes that the documented ECmax measured by other
474 laboratories is accurate. Promisingly, our results suggest good inter-lab agreement for these
475 measurements (compare A280 and ECmax estimates). In addition, an analysis of all FPs on
476 FPbase, comparing the proteins' extinction coefficients at 280 nm versus at their maximum
477 excitation wavelength, supports the idea that the ECmax is a more sensitive assay for most
478 FPs (Supplementary Fig. 14, Supplementary Table 3). While the EC(280) values are highly
479 uniform (median: ~27,400 $M^{-1}cm^{-1}$), likely because most FPs are very similar in size, the
480 ECmax values are mostly considerably larger (median ~64,200 $M^{-1}cm^{-1}$). Impressively, as the
481 ECmax assay specifically quantifies FP concentration rather than total protein concentration,
482 we were also able to show that it is possible to do these calibrations in crude lysate without
483 compromising on accuracy, demonstrating that calibrants may be produced without affinity
484 purification (Supplementary Fig. 10). Overall, we expect calibrations carried out using the
485 ECmax method, using either purified calibrants or (sonicated) cell lysates, to be equally
486 accurate. (Alternatively, if necessary, users can purify FPs and cross-check quantifications
487 using multiple assays.) We provide both rigorous and expedient protocols for prospective
488 users on protocols.io.
489
490 Such calibrations can then be used, first, to enable the comparison of experimental results
491 from different plate readers or across different settings, in molecules of fluorescent protein
492 per particles of equivalent microspheres (MEFP/PEMS; Fig. 4B). This is akin to using
493 fluorescein but with a broader application range, since using bespoke calibrants for each FP
494 allows us to calibrate instruments for any FP regardless of its spectral characteristics. If used
495 merely as comparative units, the precision (repeatability) of each calibration is important,

496 but their accuracy (whether conversions predict molecule numbers as closely as possible) is
497 not.
498
499 Second, they can be used to express protein abundance as 'molecules per cell' (Fig. 5C).
500 Accuracy here is an important consideration, and will not only depend on the accuracy of
501 the FP calibration (discussed above), but also on the microsphere calibration, and the
502 removal of any interactions between absorbance and fluorescence characteristics of cells
503 expressing FPs. As to microsphere calibration, cross-comparison with flow cytometry data
504 suggests cell count estimates from microsphere calibrations are reasonable (Beal *et al.*,
505 2020; Fedorec *et al.*, 2020), although Beal and colleagues used 0.961 µm microspheres
506 whose size is closer to *E. coli* than those used in this study (0.890 µm; the larger type are
507 now unavailable). Our protein abundance estimates (10^2-10^5 proteins per cell; Fig. 5B) are
508 also within reasonable bounds (Ishihama *et al.*, 2008; Li *et al.*, 2014), corroborating their use
509 and indicating that FP calibrations may enable protein abundance comparisons between
510 microplate assays and proteomics experiments.
511
512 The consideration of whether the presence of cells interferes with fluorescence
513 quantification (or vice versa) is multifaceted. It is well known that cells interfere with
514 fluorescence measurements through autofluorescence. Cellular autofluorescence is known
515 to largely impact GFP quantification accuracy (Lichten *et al.*, 2014), and is corrected for in
516 FPCountR by normalising to the background fluorescence of control cells at a similar OD
517 (akin to Hirst *et al.*, 2015; Fedorec *et al.*, 2020). The 'quenching' of apparent FP fluorescence
518 by the presence of cells is more rarely considered (Hirst *et al.*, 2015). We have found that
519 the effect size is comparable for different FPs (Supplementary Fig. 12), unlike for

520 autofluorescence, supporting previous observations by others (Zhang *et al.*, 2009). A
521 correction for this attenuation in FPCountR compensates fluorescence according to the
522 expected percentage quenched at the measured OD (Fig. 5B). Both of these corrections are
523 included in the process_plate() function.
524
525 Certain sources of error cannot be adequately addressed by calibration alone. While some
526 authors note that pH can affect the molecular brightness of certain FPs (Hirst *et al.*, 2015), it
527 could not be compensated for analytically without user input detailing both the pH response
528 profile of the included FPs and the pH of their cells. Fortunately, since the cellular pH in *E.*
529 *coli* is limited between pH 7.2-7.8 (Wilks and Slonczewski, 2007), and even pH-sensitive FPs
530 exhibit only mild (<10%) variation in molecular brightness between pH 7-8 (Kneen *et al.*,
531 1998; Roberts *et al.*, 2016), pH-dependent changes in molecular brightness are unlikely to
532 have a large effect on quantifications.
533
534 Overall, it seems likely that the accuracy of calibrated molecule counts per cell may be
535 limited by the accuracy of cell count calibration. It is known that OD-specific cell counts
536 change with growth rate, and that this is due to the positive relationship between growth
537 rate and cell size for *E. coli* (Volkmer and Heinemann, 2011). We therefore believe these
538 calculations of protein counts per cell will be more accurate where cell size and microsphere
539 size are as similar as possible, and inaccuracies will arise if the differences become
540 significant. Thus, these values are likely to be approximately accurate so long as the
541 following assumptions are true: that *E. coli* absorbance (scatter) is well represented by
542 microsphere properties, and that it doesn't change significantly over time or between
543 samples. These assumptions are likely to be true for typical *E. coli* experiments since cell

544 shrinkage has been shown to take place only after several hours in stationary culture, and
545 only large growth rate differences maintained for prolonged periods were observed to make
546 a significant difference to cell size in exponentially growing cells (Volkmer and Heinemann,
547 2011). In addition, our finding that microsphere absorbance profiles differ from *E. coli* ones
548 suggest the use of caution when comparing between OD600 and OD700 measurements
549 calibrated using microspheres (Supplementary Note 2 and Supplementary Fig. 11). In
550 contrast, using a conversion from OD to total cell volume in a given culture well allows us to
551 remove the requirement for counting cells (Fig. 5D) and allows the expression of protein
552 concentrations as ~0.01-100 µM. These figures should be approximately accurate under the
553 assumption that the OD-specific cell volume doesn't vary significantly between samples or
554 over time. This is strongly supported by the results of Volkmer and colleagues (Volkmer and
555 Heinemann, 2011) whose data suggests this variation is within ~2-fold across a wide range
556 of growth conditions, but also by others who have shown that as cell volumes increase, their
557 OD-specific cell counts decrease by approximately the same magnitude (Schaechter *et al.*,
558 1958; Basan *et al.*, 2015). Units of concentration may also be more meaningful for reaction
559 modelling since ultimately it is molecular concentration that is critical for binding and
560 kinetics (Bloom *et al.*, 2014; Siegal-Gaskins *et al.*, 2014; Boada *et al.*, 2019).
561
562 Absolute FP quantification should allow us to make systematic comparisons between
563 different FPs in synthetic circuits. While such a comparison was beyond the scope of this
564 paper, we did observe that while our standardised SEVA vectors produced mCherry and
565 mTagBFP2 on a comparable order of magnitude, there were observable differences
566 between the levels of each FP. This supports the idea that the use of identical
567 transcriptional and translational elements cannot always be assumed to guarantee

568 equivalent translation (or degradation) rates of different fluorescent proteins (Tuller *et al.*,
569 2010; Tuller and Zur, 2015; Cambray *et al.*, 2018), and is not a good assumption on which to
570 base FP calibrations, as suggested elsewhere (Vignoni *et al.*, 2019).
571
572 We used absolute protein quantification to investigate the problem first described by Hecht
573 and colleagues (Hecht *et al.*, 2016) in which an association between mCherry
574 overexpression and deviations in cell density measurements were made. Repeating these
575 assays showed a clear effect, but this effect was of modest magnitude (<5%) and was only
576 apparent at very high mCherry levels per cell (over 10^5 molecules per cell; Fig. 6). This
577 suggests that for most circuits that use moderate expression levels to minimise cellular
578 burden, OD600 values would remain an accurate way to quantify cell density. We also
579 unexpectedly observed the opposite effect in mTagBFP2 expressions. As this analysis was
580 technically a quantification of whether OD600 and OD700 measurements deviate from each
581 other in the presence of different FPs, these results might not suggest that OD600-based
582 cell density readings in the presence of high mTagBFP2 are inaccurate. It might instead
583 suggest that, conversely, for mTagBFP2, the OD700 readings are inaccurate. We do not
584 currently have an explanation for this finding. One possible contributing factor may be that
585 while it is often described as one of the best blue FPs, mTagBFP2 has a higher propensity to
586 aggregate than mCherry and most GFPs (Cranfill *et al.*, 2016). Cell stress is known to induce
587 *E. coli* to elongate (Justice *et al.*, 2008; Basan *et al.*, 2015), which may affect its scattering
588 properties. It is possible that very high FP levels may frequently cause a small but significant
589 error in cell density estimates due to combinations of effects from light absorption and
590 scatter that warrants more study in order to allow us to further improve molecular
591 quantifications of FPs under those conditions.

592

593 Absolute quantification need not be limited to fluorescent proteins. The last few years have
594 seen a fantastic expansion of fluorogenic molecules, tools that have allowed the specific
595 quantification of localised proteins (Svendsen *et al.*, 2008; Li *et al.*, 2018; Csibra *et al.*, 2020),
596 proteins in anaerobic environments (Streett *et al.*, 2019; Charubin *et al.*, 2020), and the
597 fluorescent quantification of RNAs (Pothoulakis *et al.*, 2014; Siegal-Gaskins *et al.*, 2014;
598 Yerramilli and Kim, 2018). Calibration of these molecules would be more complex to achieve
599 but no less valuable. Calibration of alternative instruments such as flow cytometers are also
600 of interest, but would require a very different approach, requiring calibrants to be attached
601 to particles of cell-like sizes.

602

603 While flow cytometry and mass spectrometry allow us to probe single-cell measurements or
604 the protein complement of an entire cell, respectively, microplate readers remain an
605 important screening platform in the Design-Build-Test-Learn cycle due to ease of use, low
606 cost and high iterative capabilities. This necessitates the development of methods for
607 extracting informative numbers from such data. We hope that our demonstration of how to
608 achieve absolute FP quantifications using FPCountR can contribute to the effort to develop
609 more quantitative approaches for the analysis of circuit behaviour in synthetic biology.

610

611 **Materials and Methods**

612

613 **Materials**

614 Primers and gblocks were obtained from IDT, and *E. coli* strain DH5-alpha (Invitrogen,
615 18265-017) was used for molecular cloning. Chemicals and protein reagents were purchased

616 from Merck, Sigma, ThermoFisher Scientific, and Bio-Rad, molecular biology reagents from
617 NEB and Life Technologies and general laboratory reagents from Corning, Greiner Bio-One,
618 Starlab and VWR.
619
620 **Fluorescent proteins**
621 The mCherry (Shaner *et al.*, 2004) protein sequence was based on the FPbase entry for
622 mcherry (FPbase ID: ZERB6) with the following changes: M1(MVHHHHHHGSG). The
623 mGFPmut3 (Cormack, Valdivia and Falkow, 1996) protein sequence was based on the
624 FPbase entry for gfpmut3 (FPbase ID: A20WC) with the following changes:
625 M1(MVHHHHHHGSG), A206K (a substitution to make the protein monomeric; (Zacharias *et*
626 *al.*, 2002; Shaner, Steinbach and Tsien, 2005)). The mTagBFP2 (Subach *et al.*, 2011) protein
627 sequence was based on the FPbase entry for mtagbfp2 (FPbase ID: ZO7NN) with the
628 following changes: M1(MVHHHHHHGSG). Full protein sequences are provided in
629 Supplementary Table 6.
630
631 **DNA assembly**
632 Vectors for fluorescent protein purification and growth curve assays were constructed
633 according to standard protocols, via Golden Gate and Gibson assembly techniques using *E.*
634 *coli* strain DH5α (Invitrogen, 18265-017). Constructs were assembled into Standardised
635 European Vector Architecture (SEVA) backbones (Silva-Rocha *et al.*, 2013): pS381 was
636 generated from pS181 with chloramphenicol substitution; pS361 was generated similarly
637 from pS161. In an effort to approximately equalise expression levels between different
638 proteins, the 5' untranslated region (including ribosome binding site) and 5' region of each
639 construct was set to be identical up to residue 11 (coding region begins: *DNA*:

640 ATGGTTCACCATCATCATCACcacGGTtcgggc, *protein*: MVHHHHHHGSG). The second residue
641 was set to valine to reduce the effects of N-end mediated degradation (Mogk, Schmidt and
642 Bukau, 2007). The affinity tag chosen for purification was the His_6 tag, which was followed
643 by a short unstructured linker (GSG). Full DNA sequences of vectors are provided in
644 Supplementary Table 6.
645
646 **FPCount (wet lab) protocol**
647 The protocols are available on protocols.io. In total, we have described three FPCount
648 calibration protocols: (1) the full protocol using FP purification and three protein assays for
649 cross-validation, as conducted in Fig. 3 available at https://www.protocols.io/view/fpcount-
650 protocol-full-protocol-bztsp6ne (Csibra and Stan, 2021a), (2) a short protocol requiring
651 purification but using the ECmax assay only, available at
652 https://www.protocols.io/view/fpcount-protocol-short-protocol-bzt6p6re (Csibra and Stan,
653 2021c), and (3) the purification-free protocol that uses FPs in lysates for calibration,
654 available at https://www.protocols.io/view/fpcount-protocol-in-lysate-purification-free-
655 proto-bzudp6s6 (Csibra and Stan, 2021b). It is anticipated that users will have differing
656 requirements concerning protocol time and validation, so we have provided all three
657 protocols.
658
659 **Protein expression and harvesting**
660 Fluorescent proteins were produced using pS381 (SEVA) expression vectors in *E. coli*
661 BL21(DE3) strains. Glycerol stocks were inoculated into 50 ml Luria Broth (Miller)
662 supplemented with 50 µg/ml chloramphenicol and 0.02 % arabinose and were grown
663 overnight at 30 °C at 250 rpm. Cells were harvested after ~16 h by transferring them to

664 prechilled containers on ice. All further steps were conducted on ice. OD600 readings were
665 taken, and 40 OD of cells was transferred to fresh falcon tubes, washed once in T50N300 (50
666 mM Tris-HCl pH 7.5, 300 mM NaCl) and resuspended in lysis buffer (T50N300, 1X protease
667 inhibitors (EDTA-free, Pierce A32955), filter sterilised, supplemented with lysozyme 100
668 µg/µl). Cells were separated into 20 OD (2 ml) fractions and sonicated (Qsonica Q125
669 sonicator, 50% amplitude, 10s on, 10s off, 2 min). Lysates were supplemented with 5 mM
670 $CaCl_2$, 50 mM $MgCl_2$ and treated with DNase I (50 U/ml, bovine pancreas, MP Biomedicals,
671 219006210) for 30 min at 4 °C. Soluble fractions were isolated by isolation of the
672 supernatant after centrifugation (16,000 xg, 30 min, 4 °C), and both fractions were checked
673 by SDS-PAGE followed by staining with Coomassie-based dye according to the
674 manufacturer's instructions (Instant Blue Protein Stain, Sigma ISB1L-1L).
675
676 **Protein purification**
677 Fluorescent proteins were purified in batch using His-tag affinity chromatography at room
678 temperature according to the resin manufacturer's instructions (Thermo Fisher). Lysates
679 were supplemented with 10 mM imidazole and 600 µl was applied to HisPur Cobalt resin
680 (300 µl, ThermoFisher) equilibrated in Binding Buffer (T50N300+pi, 10 mM imidazole),
681 mixed and incubated at room temperature for ~15 min before removal (1,000 xg, 1 min).
682 This was repeated four times, before the resin was washed with 10 column volumes of
683 Binding Buffer. Protein was eluted in Elution Buffer (T50N300+pi, 150 mM imidazole). All
684 protein fractions for calibration were stored protected from light at 4 °C.
685
686
687

688 **Preparation for calibration assays**

689 Elution fractions were combined and concentrated approximately 10-fold using Amicon
690 centrifugal filter columns (Merck, UFC5010), followed by buffer exchange (1000x) into
691 T5N15 (5mM Tris-HCl pH 7.5, 15 mM NaCl) or T5N15+pi (T5N15, 1x protease inhibitors, filter
692 sterilised).

693

694 **Microplate reader assays.** All assays were carried out using a Tecan Spark microplate reader
695 except the fluorescence spectra assays, which were carried out using a BMG Clariostar Plus
696 microplate reader.

697

698 **Calibration assays.** For each FP calibration, both concentration and fluorescence assays
699 were carried out on the same dilution series of protein. Concentrated, buffer exchanged FP
700 (100 µl) was diluted in 900 µl buffer, then diluted 1:2 into 500 µl buffer in 1.5 ml
701 eppendorfs. A total of 11 dilutions were prepared in this way, distributed into UV-
702 transparent microplates (Greiner, 655801) as duplicates (225 µl). Bovine serum albumin
703 (BSA) standards (from Micro BCA Protein Assay kit, ThermoFisher, 23235) were prepared in
704 parallel with the same buffer(s). This dilution set was then subjected to the protein
705 concentration and fluorescence assays.

706

707 **Protein concentration assays: A280 assay and ECmax assay**

708 Absorbance assays were carried out on 225 µl protein in UV-transparent plates, using the
709 Spark absorbance scan method (see Supplementary Methods).

710

711

712 **Fluorescence assays**

713 Following absorbance scans, 200 µl from each well of the original plate was transferred into
714 clear polystyrene plates (Corning, 3370). This plate was sealed (Eppendorf Masterclear real-
715 time PCR film adhesive, 30132947) and used to run the Spark fluorescence methods (see
716 Supplementary Methods) on all relevant instruments, channels and gains.

717

718 **Protein concentration assays: microBCA assay**

719 BCA assays were carried out using the Micro BCA Protein Assay kit (ThermoFisher, 23235)
720 according to the manufacturer's instructions (microplate protocol). Briefly, 150 µl of
721 working reagent was dispensed into a clean microplate, and 150 µl from each well of the
722 fluorescence assay plate was mixed into the reagent with a multichannel pipettor. Reactions
723 were covered with a plate seal (BreatheEasy sealing membrane, Sigma, Z380059), and
724 subjected to the Spark microBCA method (see Supplementary Methods).

725

726 **Calibration of OD600 and OD700 values using microspheres**

727 Calibration of optical density readings used to quantify cell number (OD600 and OD700) was
728 carried out according to published protocols (Fedorec *et al.*, 2020). The microspheres used
729 were monodisperse silica microspheres (Cospheric, SiO2MS-2.0, 2.0g/cc, d50=0.890um,
730 CV=3.2%, <1% Doubles).

731

732 **Bacterial timecourse assays**

733 DH10B *E. coli* transformants were grown overnight in M9 medium (M9 salts (1X, Sigma
734 M6030), casamino acids (0.2%), fructose (0.8%), thiamine HCl (0.25 mg/ml), $MgSO_4$ (2 mM),
735 $CaCl_2$ (0.1 mM)) supplemented with 50 µg/ml chloramphenicol, in a deep-well plate (30 ^{o}C,

736 700 rpm), and diluted the following morning into fresh M9 with antibiotic (deep-well plate,
737 30 °C, 700 rpm) to an OD600 (cm^{-1}) of 0.05. After 1 hour, cultures were transferred into
738 clear 96-well microplates (Corning, 3370) with pre-loaded arabinose (5 µl). Plates were
739 sealed (BreatheEasy sealing membrane, Sigma, Z380059) and grown in a Tecan Spark plate
740 reader in kinetic mode (see Supplementary Methods, Spark growth curve method).
741
742 **Analytical methods**
743 All data was analysed using R (R Core Team, 2021). The FPCountR package that was
744 developed for the FP calibrations is available on GitHub at
745 https://github.com/ec363/fpcountr (Csibra, 2021). For a summary of the functions, see the
746 Fig. 1 and Supplementary Fig. 13.
747
748 **Fluorescence scans**
749 Excitation and emission spectra of the fluorescent proteins were conducted using a BMG
750 Clariostar Plus microplate reader in sealed plates (Corning, 3370; Eppendorf Masterclear
751 real-time PCR film adhesive, 30132947) at wavelengths appropriate to each FP (see
752 Supplementary Methods).
753
754 **Figures**
755 Figures were created using RStudio and Biorender.com.
756
757 **Data availability**
758 Computer code produced in this study is available on the following repositories:
759 - R package for data analysis: GitHub (https://github.com/ec363/fpcountr)

760 This study includes no data deposited in external repositories.

761

762 **Acknowledgements**

763 GBS and EC acknowledge funding from the Royal Academy of Engineering (RAEng CiET

764 1819\5).

765

766 **Author contributions**

767 EC and GBS conceived the study. EC conducted the experiments, analysed the data and

768 wrote the R package. EC and GBS wrote the manuscript.

769

770 **Conflict of Interest**

771 The authors declare no conflict of interest.

772

773 **References**

774 Aitken, A. and Learmonth, M.P. (2002) 'Protein Determination by UV Absorption', in Walker,
775 J. M., *Protein Protocols Handbook, The*. New Jersey: Humana Press, pp. 3–6. doi:10.1385/1-
776 59259-169-8:3.

777 Arike, L. *et al.* (2012) 'Comparison and applications of label-free absolute proteome
778 quantification methods on Escherichia coli', *Journal of Proteomics*, 75(17), pp. 5437–5448.
779 doi:10.1016/j.jprot.2012.06.020.

780 Basan, M. *et al.* (2015) 'Inflating bacterial cells by increased protein synthesis', *Molecular*
781 *Systems Biology*, 11(10), p. 836. doi:10.15252/msb.20156178.

782 Beal, J. *et al.* (2018) 'Quantification of bacterial fluorescence using independent calibrants',
783 *PLOS ONE*, 13(6), p. e0199432. doi:10.1371/journal.pone.0199432.

784 Beal, J., Farny, N.G., *et al.* (2020) 'Robust estimation of bacterial cell count from optical
785 density', *Communications Biology*, 3(1), p. 512. doi:10.1038/s42003-020-01127-5.

786 Beal, J., Goñi-Moreno, A., *et al.* (2020) 'The long journey towards standards for engineering
787 biosystems: Are the Molecular Biology and the Biotech communities ready to standardise?',
788 *EMBO reports*, 21(5). doi:10.15252/embr.202050521.

789 Bloom, R.J., Winkler, S.M. and Smolke, C.D. (2014) 'A quantitative framework for the
790 forward design of synthetic miRNA circuits', *Nature Methods*, 11(11), pp. 1147–1153.
791 doi:10.1038/nmeth.3100.

792 Boada, Y. *et al.* (2019) 'Characterization of Gene Circuit Parts Based on Multiobjective
793 Optimization by Using Standard Calibrated Measurements', *ChemBioChem*, 20(20), pp.
794 2653–2665. doi:10.1002/cbic.201900272.

795 Boo, A., Ellis, T. and Stan, G.-B. (2019) 'Host-aware synthetic biology', *Current Opinion in*
796 *Systems Biology*, 14, pp. 66–72. doi:10.1016/j.coisb.2019.03.001.

797 Bradford, M.M. (1976) 'A rapid and sensitive method for the quantitation of microgram
798 quantities of protein utilizing the principle of protein-dye binding', *Analytical Biochemistry*,
799 72(1), pp. 248–254. doi:10.1016/0003-2697(76)90527-3.

800 Calderón-Celis, F., Encinar, J.R. and Sanz-Medel, A. (2018) 'Standardization approaches in
801 absolute quantitative proteomics with mass spectrometry', *Mass Spectrometry Reviews*,
802 37(6), pp. 715–737. doi:10.1002/mas.21542.

803 Cambray, G., Guimaraes, J.C. and Arkin, A.P. (2018) 'Evaluation of 244,000 synthetic
804 sequences reveals design principles to optimize translation in Escherichia coli', *Nature*
805 *Biotechnology* [Preprint]. doi:10.1038/nbt.4238.

806 Casadevall, A. and Fang, F.C. (2016) 'Rigorous Science: a How-To Guide', *mBio*, 7(6).
807 doi:10.1128/mBio.01902-16.

808 Ceroni, F. *et al.* (2015) 'Quantifying cellular capacity identifies gene expression designs with
809 reduced burden', *Nature Methods*, 12(5), pp. 415–418. doi:10.1038/nmeth.3339.

810 Chalfie, M. *et al.* (1994) 'Green fluorescent protein as a marker for gene expression',
811 *Science*, 263(5148), pp. 802–805. doi:10.1126/science.8303295.

812 Charubin, K., Streett, H. and Papoutsakis, E.T. (2020) 'Development of Strong Anaerobic
813 Fluorescent Reporters for Clostridium acetobutylicum and Clostridium ljungdahlii Using
814 HaloTag and SNAP-tag Proteins', *Applied and Environmental Microbiology*. Edited by R.M.
815 Kelly, 86(20). doi:10.1128/AEM.01271-20.

816 Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) 'FACS-optimized mutants of the green
817 fluorescent protein (GFP)', *Gene*, 173(1), pp. 33–38. doi:10.1016/0378-1119(95)00685-0.

818 Cranfill, P.J. *et al.* (2016) 'Quantitative assessment of fluorescent proteins', *Nature Methods*,
819 13(7), pp. 557–562. doi:10.1038/nmeth.3891.

820 Csibra, E. (2021) *FPCountR: Fluorescent protein calibration for plate readers*. Zenodo.
821 doi:10.5281/zenodo.5760028.

822 Csibra, E., Renders, M. and Pinheiro, V. (2020) 'Bacterial cell display as a robust and versatile
823 platform for the engineering of low affinity ligands and enzymes', *ChemBioChem*, p.
824 cbic.202000203. doi:10.1002/cbic.202000203.

825 Csibra, E. and Stan, G.-B. (2021a) 'FPCount protocol - Full protocol', *protocols.io* [Preprint].
826 doi:10.17504/protocols.io.bztsp6ne.

827 Csibra, E. and Stan, G.-B. (2021b) 'FPCount protocol - in-lysate (purification free) protocol',
828 *protocols.io* [Preprint]. doi:10.17504/protocols.io.bzudp6s6.

829 Csibra, E. and Stan, G.-B. (2021c) 'FPCount protocol - Short protocol', *protocols.io* [Preprint].
830 doi:10.17504/protocols.io.bzt6p6re.

831 Delvigne, F. *et al.* (2018) 'Improving control in microbial cell factories: from single cell to
832 large-scale bioproduction', *FEMS Microbiology Letters* [Preprint].
833 doi:10.1093/femsle/fny236.

834 Fedorec, A.J.H. *et al.* (2020) 'FlopR: An Open Source Software Package for Calibration and
835 Normalization of Plate Reader and Flow Cytometry Data', *ACS Synthetic Biology*, 9(9), pp.
836 2258–2266. doi:10.1021/acssynbio.0c00296.

837 Finan, K., Raulf, A. and Heilemann, M. (2015) 'A Set of Homo-Oligomeric Standards Allows
838 Accurate Protein Counting', *Angewandte Chemie International Edition*, 54(41), pp. 12049–
839 12052. doi:10.1002/anie.201505664.

840 Gorochowski, T.E. *et al.* (2019) 'Absolute quantification of translational regulation and
841 burden using combined sequencing approaches', *Molecular Systems Biology*, 15(5), p.
842 e8719. doi:10.15252/msb.20188719.

843 Hecht, A. *et al.* (2016) 'When Wavelengths Collide: Bias in Cell Abundance Measurements
844 Due to Expressed Fluorescent Proteins', *ACS Synthetic Biology*, 5(9), pp. 1024–1027.
845 doi:10.1021/acssynbio.6b00072.

846 Hirst, C.D. *et al.* (2015) 'Protocol for the Standardisation of Transcriptional Measurements',
847 in McGenity, T.J., Timmis, K.N., and Nogales, B. (eds) *Hydrocarbon and Lipid Microbiology*
848 *Protocols*. Berlin, Heidelberg: Springer Berlin Heidelberg (Springer Protocols Handbooks),
849 pp. 9–26. doi:10.1007/8623_2015_148.

850 Ishihama, Y. *et al.* (2008) 'Protein abundance profiling of the Escherichia coli cytosol', *BMC*
851 *Genomics*, 9(1), p. 102. doi:10.1186/1471-2164-9-102.

852 Justice, S.S. *et al.* (2008) 'Morphological plasticity as a bacterial survival strategy', *Nature*
853 *Reviews Microbiology*, 6(2), pp. 162–168. doi:10.1038/nrmicro1820.

854 Kneen, M. *et al.* (1998) 'Green Fluorescent Protein as a Noninvasive Intracellular pH
855 Indicator', *Biophysical Journal*, 74(3), pp. 1591–1599. doi:10.1016/S0006-3495(98)77870-1.

856 Lambert, T.J. (2019) 'FPbase: a community-editable fluorescent protein database', *Nature*
857 *Methods*, 16(4), pp. 277–278. doi:10.1038/s41592-019-0352-8.

858 Li, C. *et al.* (2018) 'Fluorogenic Probing of Membrane Protein Trafficking', *Bioconjugate*
859 *Chemistry*, 29(6), pp. 1823–1828. doi:10.1021/acs.bioconjchem.8b00180.

860 Li, G.-W. *et al.* (2014) 'Quantifying Absolute Protein Synthesis Rates Reveals Principles
861 Underlying Allocation of Cellular Resources', *Cell*, 157(3), pp. 624–635.
862 doi:10.1016/j.cell.2014.02.033.

863 Lichten, C.A. *et al.* (2014) 'Unmixing of fluorescence spectra to resolve quantitative time-
864 series measurements of gene expression in plate readers', *BMC Biotechnology*, 14(1), p. 11.
865 doi:10.1186/1472-6750-14-11.

866 Mogk, A., Schmidt, R. and Bukau, B. (2007) 'The N-end rule pathway for regulated
867 proteolysis: prokaryotic and eukaryotic strategies', *Trends in Cell Biology*, 17(4), pp. 165–
868 172. doi:10.1016/j.tcb.2007.02.001.

869 Noble, J.E. *et al.* (2007) 'A Comparison of Protein Quantitation Assays for Biopharmaceutical
870 Applications', *Molecular Biotechnology*, 37(2), pp. 99–111. doi:10.1007/s12033-007-0038-9.

871 Noble, J.E. (2014) 'Quantification of Protein Concentration Using UV Absorbance and
872 Coomassie Dyes', in *Methods in Enzymology*. Elsevier, pp. 17–26. doi:10.1016/B978-0-12-
873 420070-8.00002-7.

874 Pace, C.N. *et al.* (1995) 'How to measure and predict the molar absorption coefficient of a
875 protein', *Protein Science*, 4(11), pp. 2411–2423. doi:10.1002/pro.5560041120.

876 Perrino, G. *et al.* (2021) 'Control engineering and synthetic biology: working in synergy for
877 the analysis and control of microbial systems', *Current Opinion in Microbiology*, 62, pp. 68–
878 75. doi:10.1016/j.mib.2021.05.004.

879 Pothoulakis, G. *et al.* (2014) 'The Spinach RNA Aptamer as a Characterization Tool for
880 Synthetic Biology', *ACS Synthetic Biology*, 3(3), pp. 182–187. doi:10.1021/sb400089c.

881 R Core Team (2021) *R: The R Project for Statistical Computing*. Available at: https://www.r-
882 project.org/ (Accessed: 5 December 2021).

883 Raynal, B. *et al.* (2014) 'Quality assessment and optimization of purified protein samples:
884 why and how?', *Microbial Cell Factories*, 13(1), p. 180. doi:10.1186/s12934-014-0180-6.

885 Roberts, T.M. *et al.* (2016) 'Identification and Characterisation of a pH-stable GFP', *Scientific*
886 *Reports*, 6(1), p. 28166. doi:10.1038/srep28166.

887 Rodriguez, E.A. *et al.* (2017) 'The Growing and Glowing Toolbox of Fluorescent and
888 Photoactive Proteins', *Trends in Biochemical Sciences*, 42(2), pp. 111–129.
889 doi:10.1016/j.tibs.2016.09.010.

890 Schaechter, M., MaalØe, O. and Kjeldgaard, N.O.Y. 1958 (1958) 'Dependency on Medium
891 and Temperature of Cell Size and Chemical Composition during Balanced Growth of
892 Salmonella typhimurium', *Microbiology*, 19(3), pp. 592–606. doi:10.1099/00221287-19-3-
893 592.

894 Shaner, N.C. *et al.* (2004) 'Improved monomeric red, orange and yellow fluorescent proteins
895 derived from Discosoma sp. red fluorescent protein', *Nature Biotechnology*, 22(12), pp.
896 1567–1572. doi:10.1038/nbt1037.

Shaner, N.C., Steinbach, P.A. and Tsien, R.Y. (2005) 'A guide to choosing fluorescent proteins', *Nature Methods*, 2(12), pp. 905–909. doi:10.1038/nmeth819.

Siegal-Gaskins, D. *et al.* (2014) 'Gene Circuit Performance Characterization and Resource Usage in a Cell-Free "Breadboard"', *ACS Synthetic Biology*, 3(6), pp. 416–425. doi:10.1021/sb400203p.

Silva-Rocha, R. *et al.* (2013) 'The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes', *Nucleic Acids Research*, 41(Database issue), pp. D666–D675. doi:10.1093/nar/gks1119.

Stevenson, K. *et al.* (2016) 'General calibration of microbial growth in microplate readers', *Scientific Reports*, 6(1), p. 38828. doi:10.1038/srep38828.

Streett, H.E., Kalis, K.M. and Papoutsakis, E.T. (2019) 'A Strongly Fluorescing Anaerobic Reporter and Protein-Tagging System for *Clostridium* Organisms Based on the Fluorescence-Activating and Absorption-Shifting Tag Protein (FAST)', *Applied and Environmental Microbiology*. Edited by M. Kivisaar, 85(14). doi:10.1128/AEM.00622-19.

Subach, O.M. *et al.* (2011) 'An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore', *PLOS ONE*, 6(12), p. e28674. doi:10.1371/journal.pone.0028674.

Svendsen, S. *et al.* (2008) 'Spatial separation and bidirectional trafficking of proteins using a multi-functional reporter', *BMC Cell Biology*, 9(1), p. 17. doi:10.1186/1471-2121-9-17.

Tal, M., Silberstein, A. and Nusser, E. (1985) 'Why does Coomassie Brilliant Blue R interact differently with different proteins? A partial answer.', *Journal of Biological Chemistry*, 260(18), pp. 9976–9980. doi:10.1016/S0021-9258(17)39198-6.

Taniguchi, Y. *et al.* (2010) 'Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells', 329, p. 8.

Thermo Fisher Scientific (2010) *NanoDrop 1000 Spectrophotometer, V3.8 User's Manual*. Available at: https://tools.thermofisher.com/content/sfs/manuals/nd-1000-v3.8-users-manual-8%205x11.pdf (Accessed: 26 November 2021).

Thermo Fisher Scientific (2017) *Protein assay technical handbook: Tools and reagents for improved quantitation of total or specific proteins*. Available at: https://assets.thermofisher.com/TFS-Assets/LSG/brochures/protein-assay-technical-handook.pdf (Accessed: 26 November 2021).

Tie, H.C., Madugula, V. and Lu, L. (2016) 'The development of a single molecule fluorescence standard and its application in estimating the stoichiometry of the nuclear pore complex', *Biochemical and Biophysical Research Communications*, 478(4), pp. 1694–1699. doi:10.1016/j.bbrc.2016.09.005.

Tuller, T. *et al.* (2010) 'An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation', *Cell*, 141(2), pp. 344–354. doi:10.1016/j.cell.2010.03.031.

934 Tuller, T. and Zur, H. (2015) 'Multiple roles of the coding sequence 5' end in gene expression
935 regulation', *Nucleic Acids Research*, 43(1), pp. 13–28. doi:10.1093/nar/gku1313.

936 Vignoni, A. *et al.* (2019) 'Fluorescence calibration and color equivalence for quantitative
937 synthetic biology.**This work was partially supported by MINECO/AEI, EU DPI2017-82896-
938 C2-1-R.', *IFAC-PapersOnLine*, 52(26), pp. 129–134. doi:10.1016/j.ifacol.2019.12.247.

939 Volkmer, B. and Heinemann, M. (2011) 'Condition-Dependent Cell Volume and
940 Concentration of Escherichia coli to Facilitate Data Conversion for Systems Biology
941 Modeling', *PLoS ONE*. Edited by J. Langowski, 6(7), p. e23126.
942 doi:10.1371/journal.pone.0023126.

943 Wasinger, V.C., Zeng, M. and Yau, Y. (2013) 'Current Status and Advances in Quantitative
944 Proteomic Mass Spectrometry', *International Journal of Proteomics*, 2013, pp. 1–12.
945 doi:10.1155/2013/180605.

946 Wilks, J.C. and Slonczewski, J.L. (2007) 'pH of the Cytoplasm and Periplasm of *Escherichia*
947 *coli* : Rapid Measurement by Green Fluorescent Protein Fluorimetry', *Journal of*
948 *Bacteriology*, 189(15), pp. 5601–5607. doi:10.1128/JB.00615-07.

949 Yerramilli, V.S. and Kim, K.H. (2018) 'Labeling RNAs in Live Cells Using Malachite Green
950 Aptamer Scaffolds as Fluorescent Probes', *ACS Synthetic Biology*, 7(3), pp. 758–766.
951 doi:10.1021/acssynbio.7b00237.

952 Zacharias, D.A. *et al.* (2002) 'Partitioning of Lipid-Modified Monomeric GFPs into Membrane
953 Microdomains of Live Cells', *Science*, 296(5569), pp. 913–916. doi:10.1126/science.1068539.

954 Zhang, C. *et al.* (2009) 'Correcting for the inner filter effect in measurements of fluorescent
955 proteins in high-cell-density cultures', *Analytical Biochemistry*, 390(2), pp. 197–202.
956 doi:10.1016/j.ab.2009.04.029.

957

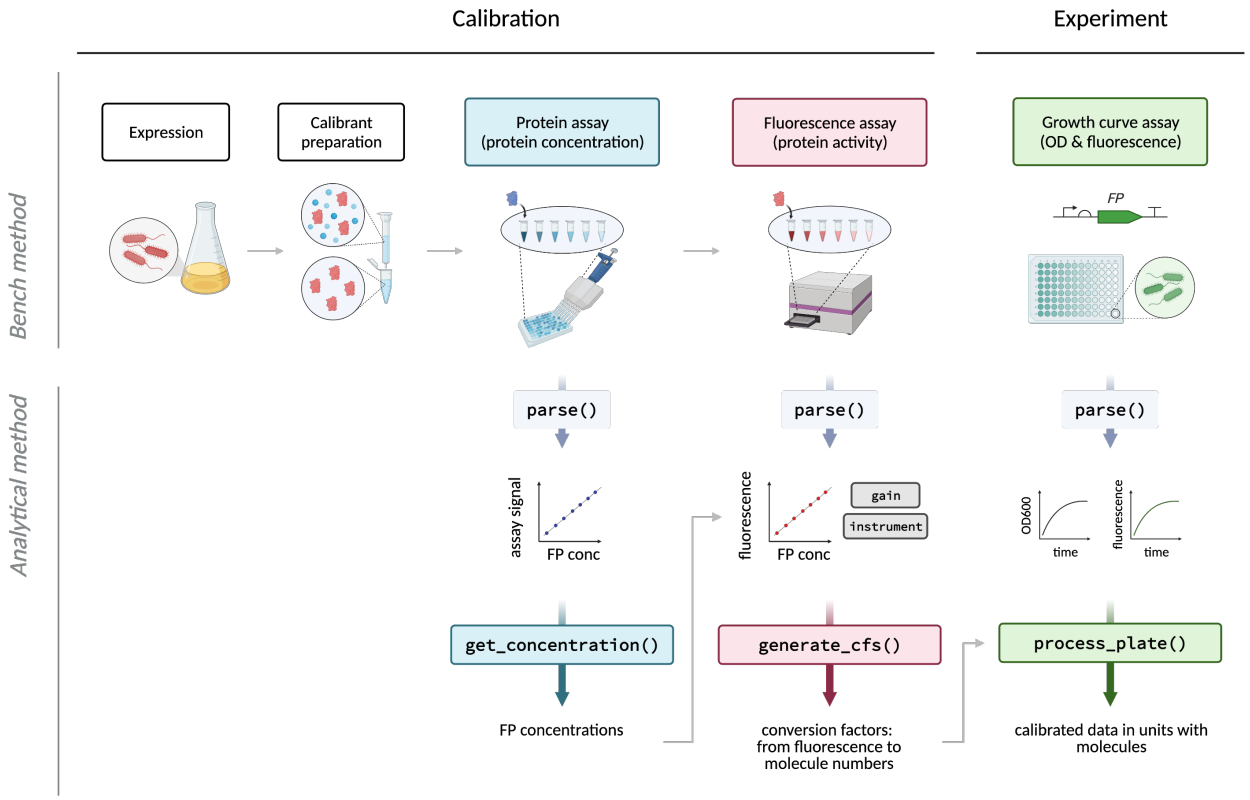

Figure 1. Overview of fluorescent protein calibration workflow using FPCountR. A calibration workflow is described *(left)*, followed by a demonstration of how this calibration can be used to convert experimental data from arbitrary fluorescence units per optical density into molecules per cell *(right)*. The calibration workflow consists of a wet lab protocol (*top, available on protocols.io*) and an analysis package (*bottom, available on GitHub*). In brief, the protocol describes how to prepare fluorescent protein calibrants by expression and purification, though the latter step is optional as lysates allow accurate calibration without the need for purification. The protocol also describes how to collect data for the calibration for both the protein assay to determine protein concentration, as well as the fluorescence assay to determine protein activity. The analytical workflow is provided as an open source R package, complete with functions that enable the extraction of protein concentrations from protein assay data, conversion factors (arbitrary fluorescence units per molecule) from a combination of protein and fluorescence assay data, and functions that allow users to convert experimental data into absolute units.

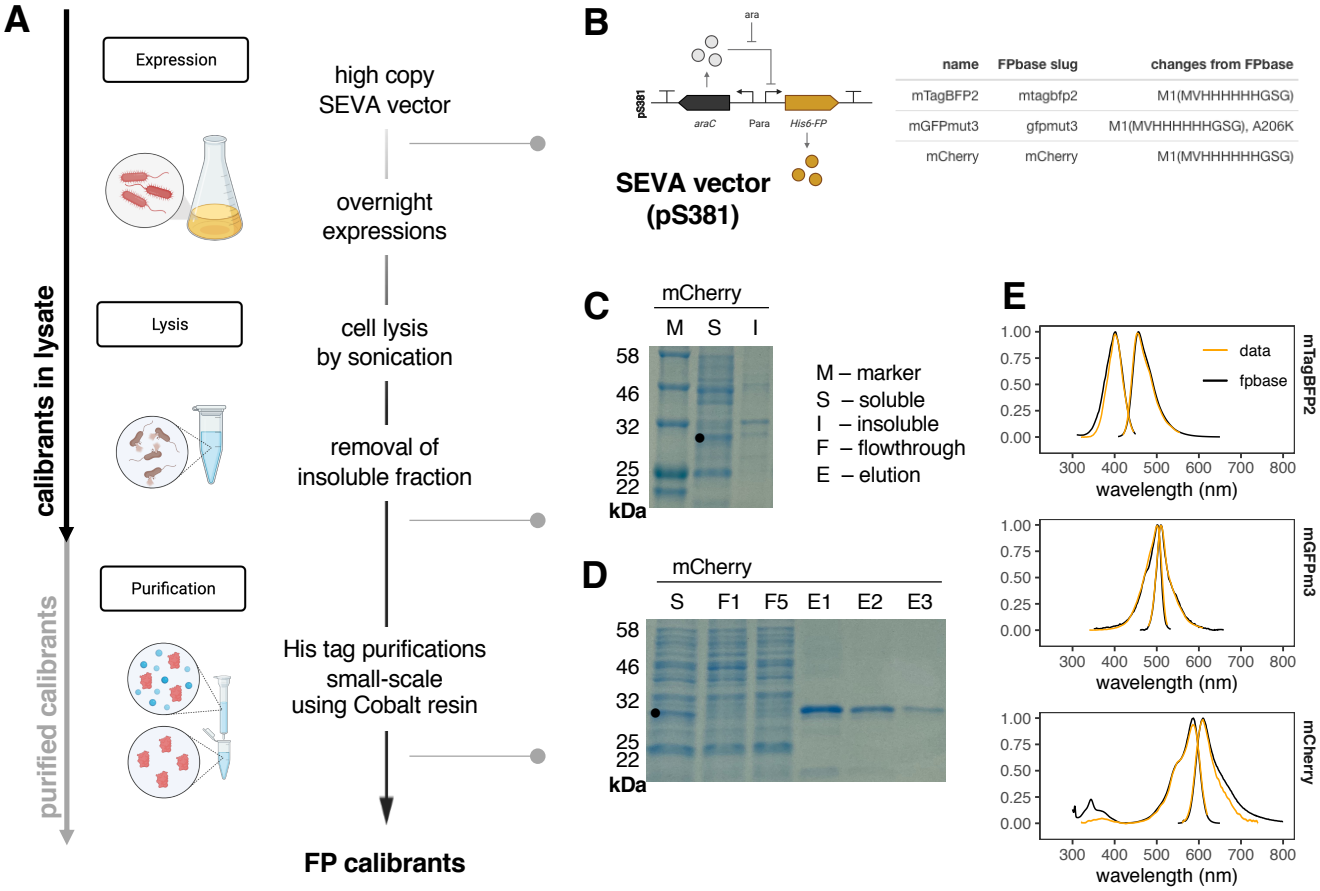
A
calibrants in lysate
purified calibrants
Expression
high copy
SEVA vector
overnight
expressions
Lysis
cell lysis
by sonication
removal of
insoluble fraction
Purification
His tag purifications
small-scale
using Cobalt resin
FP calibrants
B
ara
pS381
araC
Para
His6-FP
SEVA vector
(pS381)
name
FPbase slug
changes from FPbase
mTagBFP2
mtagbfp2
M1(MVHHHHHHGSG)
mGFPmut3
gfpmut3
M1(MVHHHHHHGSG), A206K
mCherry
mCherry
M1(MVHHHHHHGSG)
C
mCherry
M S I
58
46
32
25
22
kDa
M – marker
S – soluble
I – insoluble
F – flowthrough
E – elution
D
mCherry
S F1 F5 E1 E2 E3
58
46
32
25
22
kDa
E
data
fpbase
mTagBFP2
mGFPm3
mCherry
1.00
0.75
0.50
0.25
0.00
300 400 500 600 700 800
wavelength (nm)

Figure 2. Preparation of fluorescent protein calibrants.
A. Protocol summary. The use of high-copy vectors and overnight expression was designed to maximise protein production, and the temperature was dropped to 30 °C to minimise misfolding. Cells were lysed using sonication, to avoid the requirement to add chemical components that may interfere with downstream processes, such as EDTA (with His-tag purification) or detergents (with protein quantification), or with unknown components of 'black box' commercial lysis reagents. Insoluble proteins were removed via centrifugation and SDS-PAGE was used to confirm that the majority of the expressed FP was in the soluble fraction. Proteins were purified using His-tag affinity purification. His tags are popular affinity tags that are known for their reliability and are small in size, making them unlikely to compromise fusion protein function. Cobalt resin was used as it has higher specifity than nickel resins and is expected to co-isolate fewer impurities. The arrows on the left represent the steps required to prepare purified calibrants (grey) vs. calibrants in lysate (black). **B. Vector and FP design.** A standardised FP expression vector was constructed from an arabinose-inducible His-tagged FP construct in a high-copy SEVA vector. Three commonly-used FPs from across the spectral range were chosen for testing this protocol: mTagBFP2, mGFPmut3, and mCherry. A table of the three proteins illustrates any changes in protein sequence compared to their FPbase counterparts, showing identical matches with the exception of affinity tags and a monomerising mutation for GFPmut3. **C. Expression and solubility verification.** SDS-PAGE analysis of lysates after separation of the insoluble fraction was used to make sure that most of the fluorescent protein was soluble. The displayed SDS-PAGE is from an mCherry purification, showing the separation of the soluble (S) vs. insoluble (I) fraction, next to the protein marker (M) on a 12 % gel. **D. Purification verification.** SDS-PAGE analysis after purification was used to confirm the success of purifications. The displayed SDS-PAGE is from an mCherry purification, showing the separation of the soluble (S) fraction, next to two flowthrough (F) fractions from the binding steps showing efficient FP binding to the cobalt resin, and three elution (E) fractions. **E. Fluorescence spectra.** Fluorescence spectral scans were used to confirm that the purified FP behaves as expected. The figure shows obtained spectra (normalised such that the highest value = 1) fitted to a loess model with a 95% confidence interval (orange) overlaid with FPbase spectra (black) for each FP. Displayed spectra represent one sample measured in duplicate, that is representative of at least 2 independently purified batches of calibrant.

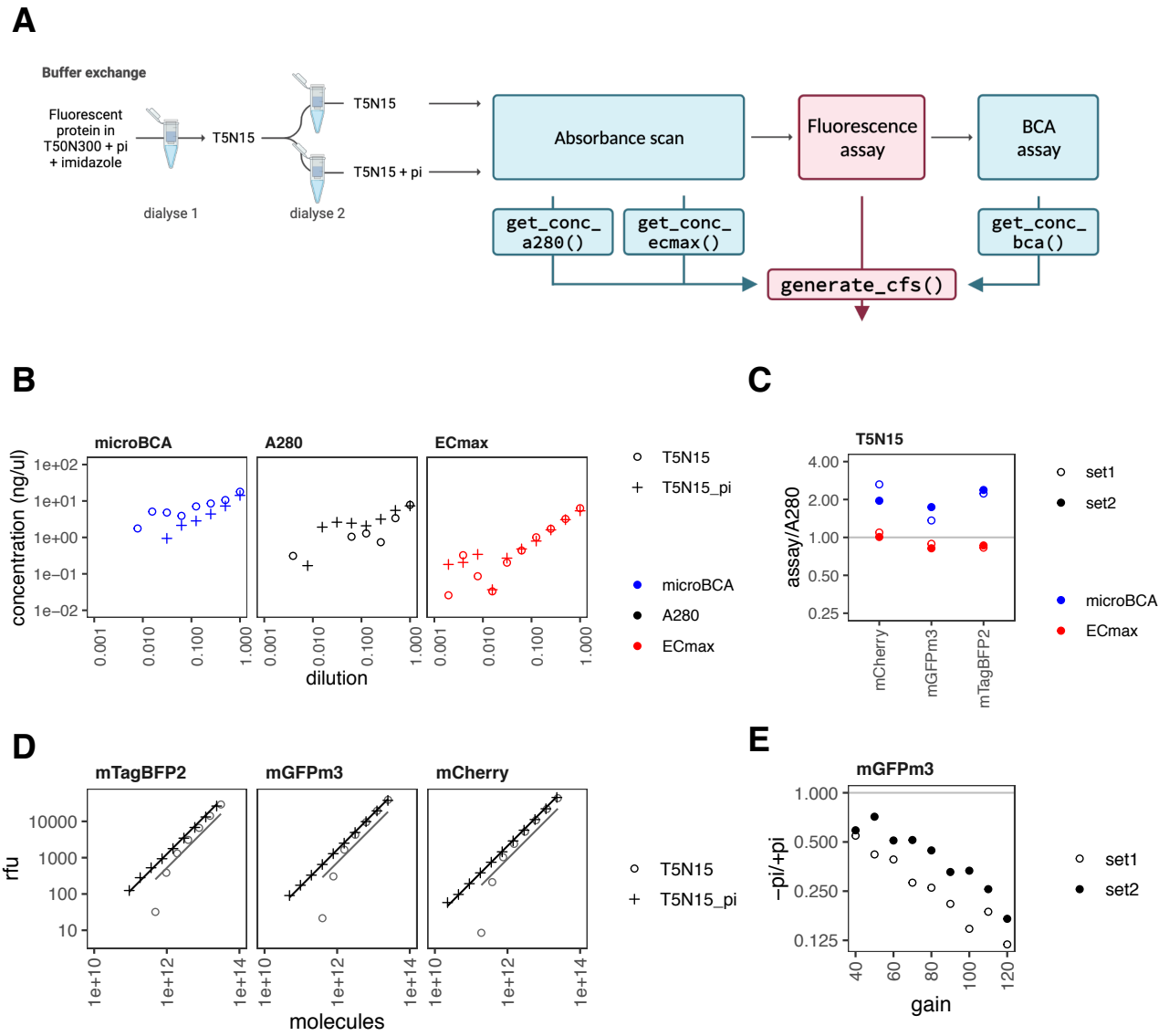
A
Buffer exchange
Fluorescent protein in T50N300 + pi + imidazole
T5N15
dialyse 1
dialyse 2
T5N15
T5N15 + pi
Absorbance scan
Fluorescence assay
BCA assay
get_conc_a280()
get_conc_ecmax()
get_conc_bca()
generate_cfs()
B
microBCA
A280
ECmax
concentration (ng/ul)
dilution
T5N15
T5N15_pi
microBCA
A280
ECmax
C
T5N15
assay/A280
mCherry
mGFPm3
mTagBFP2
set1
set2
microBCA
ECmax
D
mTagBFP2
mGFPm3
mCherry
rfu
molecules
T5N15
T5N15_pi
E
mGFPm3
−pi/+pi
gain
set1
set2

Figure 3. Protein assay comparison reveals that both assay linearity and limit of detection are sensitive to buffer components

A. Sample preparation and assay workflow. FPs in elution buffer (T50N300 with protease inhibitors (pi) and imidazole) were dialysed (1) into T5N15 and then again (2) into T5N15 or T5N15 with pi. Serial dilutions of each of these FPs in both buffers were then prepared, and each dilution was subjected to three protein assays (microBCA, A280 and ECmax assay) as well as the fluorescence assay. **B. Effect of buffer and assay on measured protein concentration of each dilution.** An example of how the buffer and choice of assay effect the measured protein concentrations using a dilution series of mTagBFP2 protein (for full results see Supplementary Fig. 9). The serial dilution was a 2-fold dilution, where 11 dilutions were prepared and measured in duplicate. Points represent the mean of the duplicate values. All dilutions were measured with the A280 and ECmax assays but only the top 8 dilutions were tested with the microBCA assay. The values for the A280 and ECmax assays were normalised for scatter as indicated in Supplementary Table 2. Any missing data points had concentrations recorded as being below 0.01 ng/μl. **C. Effect of assay on calculated protein concentration compared to the A280.** The raw data of each serial dilution, as displayed in B, was fitted to a linear model and used to estimate the concentration of the first sample in the series (where the dilution factor = 1), according to methods for each assay as set out in Supplementary Fig. 2, 7 and 8. For a given FP batch (set), the microBCA and ECmax assay error was calculated by taking the fold difference in concentration predicted by the named assay, versus that predicted by the A280 assay using T5N15 buffer. Each point therefore represents one value for each FP batch. The full data for this figure can be found in Supplementary Table 1. **D. Fluorescence assay sensitivity to buffer.** Combining the protein concentration predictions from the ECmax assay (x axis) with relative fluorescence units of the instrument to be calibrated from the fluorescence assay, conversion factors may be calculated by fitting a linear model of relative fluorescence units vs. molecules of protein. Points represent the mean of duplicate values at each concentration, and the line represents the fit found by the generate_cfs function. Patterns of linearity evident in these plots are representative of all fluorescence assays carried out in these two buffers. **E. Effect of buffer on conversion factor calculation.** Conversion factors were calculated for each protein, set and buffer using the modelled fits between relative fluorescence units and protein concentration from the ECmax assay, as in D. The fold difference in conversion factor prediction in buffers without (-) and with (+) protease inhibitors (pi) are displayed for mGFPm3, across both batches (set1 and set2) of protein. Each point therefore represents one value for each FP batch.

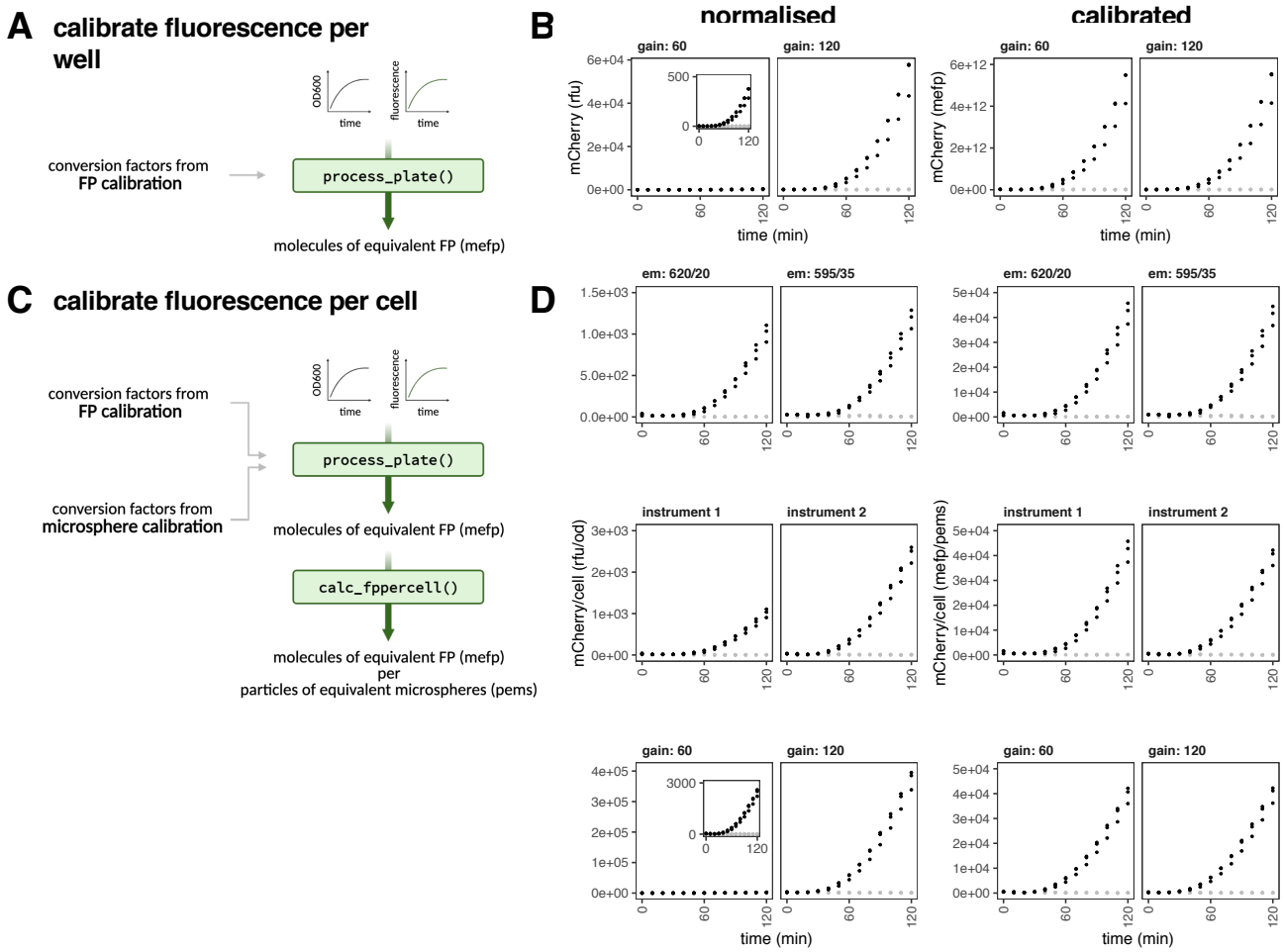

Figure 4. FP calibration allows comparison across gains and instruments for FPs other than GFP. **A. Calibration of fluorescence per well in units of MEFP.** Timecourse experimental data of *E. coli* protein expression may be processed using the process_plate() function using conversion factors obtained from FP calibration. This allows the conversion of normalised data (in relative fluorescence units, rfu) to be converted into calibrated units, of 'molecules of equivalent fluorescent protein' or 'MEFP'. **B. Comparison of normalised vs calibrated data in MEFP.** Starter cultures of *E. coli* DH10B containing pS381_ara_mCherry were transferred into a 96-well plate. mCherry expression was uninduced (grey) or induced 0.1% (black) arabinose at 0 minutes. Absorbance at OD700 and fluorescence was monitored every ten minutes. Data was collected from three biological replicates, each of which is plotted. *Left panel*: normalised mCherry in units of RFU. *Right panel*: calibrated mCherry in units of MEFP. Inset plot shows the same data as the parent plot on a zoomed axis. **C. Calibration of fluorescence per cell in units of MEFP/PEMS.** By combining the FP conversion factors with conversion factors from a microsphere calibration the data can be further processed using the calc_fppercell() function into 'per cell' data with units of MEFP per 'particles of equivalent microspheres'' (PEMS). **D. Comparison of normalised vs calibrated data in MEFP/PEMS.** Expressions were carried out as in B. Normalised and calibrated values are shown when compared across different filter sets (*top,* notation: emission wavelength/bandwidth), instruments (*middle*) and gains (*bottom*). Data for the filter and gain comparisons were taken using the same instrument.

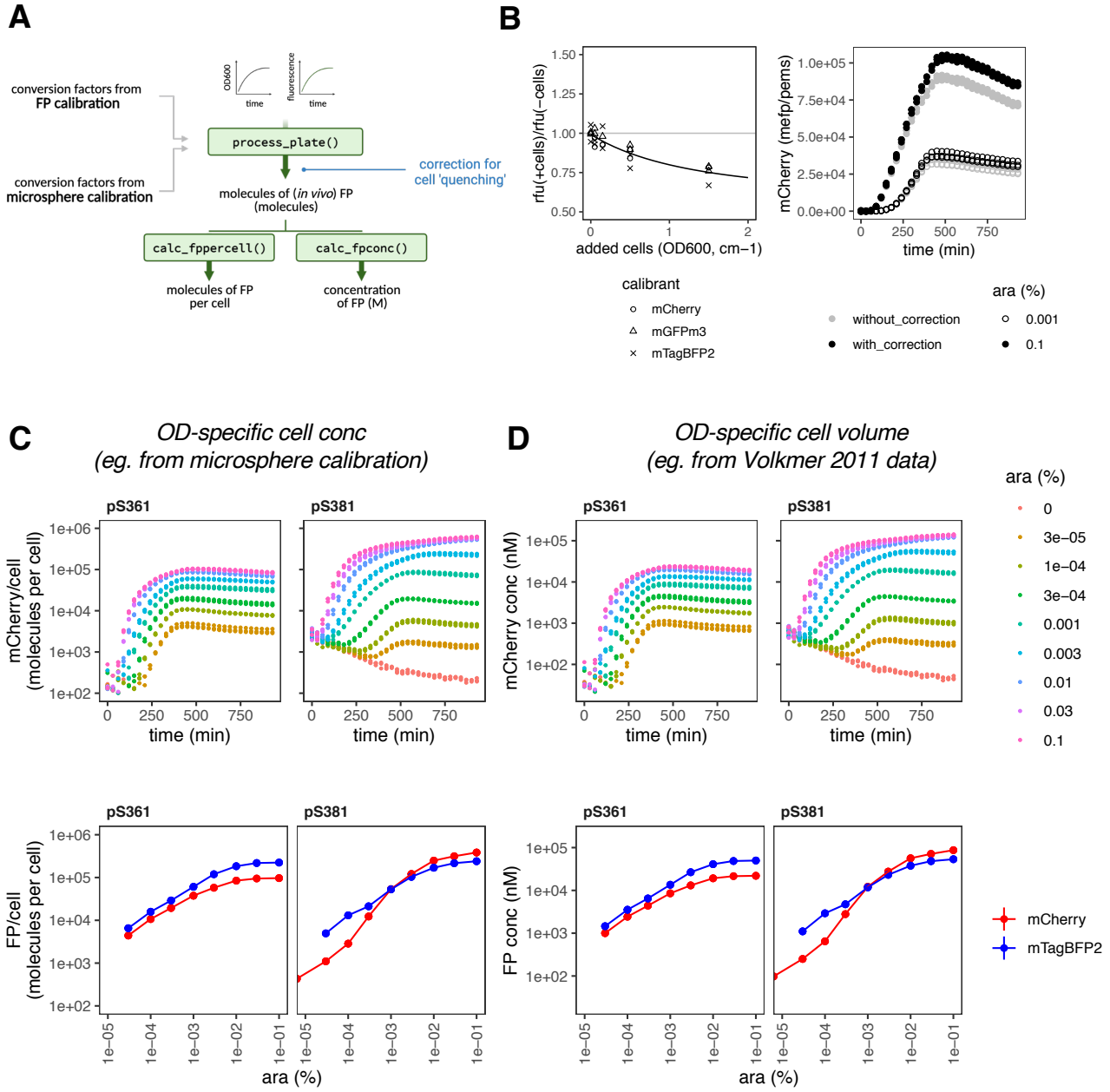
A
conversion factors from
FP calibration
conversion factors from
microsphere calibration
OD600
time
fluorescence
time
process_plate()
correction for
cell 'quenching'
molecules of (in vivo) FP
(molecules)
calc_fppercell()
calc_fpconc()
molecules of FP
per cell
concentration
of FP (M)
B
rfu(+cells)/rfu(−cells)
added cells (OD600, cm−1)
mCherry (mefp/pems)
time (min)
calibrant
mCherry
mGFPm3
mTagBFP2
without_correction
with_correction
ara (%)
0.001
0.1
C
OD-specific cell conc
(eg. from microsphere calibration)
pS361
pS381
mCherry/cell
(molecules per cell)
time (min)
FP/cell
(molecules per cell)
ara (%)
D
OD-specific cell volume
(eg. from Volkmer 2011 data)
mCherry conc (nM)
FP conc (nM)
ara (%)
0
3e−05
1e−04
3e−04
0.001
0.003
0.01
0.03
0.1
mCherry
mTagBFP2

Figure 5. Absolute quantification of *E. coli* timeseries data in molecules per cell.
A. Functions to convert experimental data to absolute units. Diagram of modifications to the process_plate() function to (i) incorporate a compensation step based on a quantitative understanding of the impact of cell density on apparent fluorescence (this allows the units to be recorded as molecules per cell), and (ii) to calculate molecular concentration of each FP instead, in molar units. **B. Quantification of the quenching effect on fluorescence on three FPs.** Purified FPs were mixed with non-fluorescent *E. coli* at a range of concentrations, and OD600 and fluorescence intensity were recorded. After normalising for cellular autofluorescence, the fold differences between relative fluorescence intensity (rfu) with (+) and without (-) added cells was quantified *(left)*. Data was collected in duplicates, with both points plotted. A model was fit through this data to enable prediction of expected fluorescence quenching for a given cell density on experimental data. An example of the effect of the correction (*right panel*). An mCherry expression vector induced with low (open circles) and high (closed circles) concentration arabinose is presented without (grey) and with (black) correction. Data was collected from three biological replicates and all points are plotted. **C-D. Absolute protein quantification in molecules per cell (C) and molar concentration (D).** mCherry expression (*top*) from medium (pS361, p15A) and high (pS381, colE1) copy vectors was induced at a range of arabinose concentrations and quantified in a timecourse assay in a calibrated plate reader. Data was processed as described in (A) and cell estimates based on microsphere calibrations were used to calculate per cell values (C), or OD-specific cell volume data from Volkmer et al., 2011 was used to calculate molar concentrations (D). Data was collected from three biological replicates, each of which is plotted. (*Bottom*) mCherry expression from the top panel is compared with mTagBFP2 expression from an identical assay, plotted against arabinose concentration at the 420 min post induction. Data was collected from three biological replicates. Displayed points show the mean, and error bars indicate standard deviations.

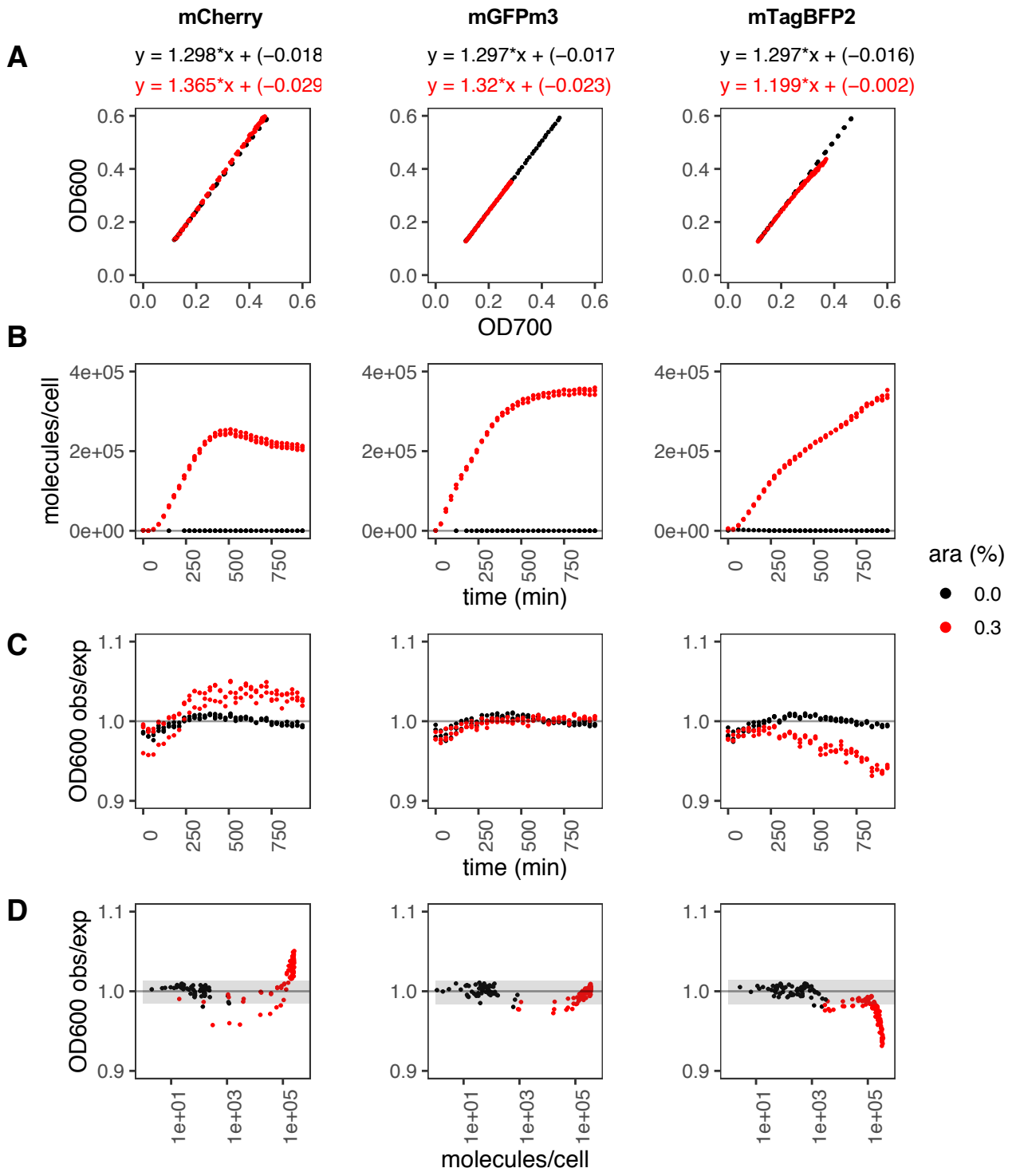

Figure 6. Evaluation of OD600:OD700 ratios in cell growth assays.
Timecourse expression assays of three FPs conducted in a plate reader were monitored for OD600, OD700 and fluorescence intensity every 30 minutes, without (black) and with (red) arabinose. Data was collected from three biological replicates, and all points are plotted. This is a representative experiment of at least two independent experiments for each FP. **A. Relationship between OD700 and OD600 values.** Linear models were fitted to data from three FPs, and shown above each plot. **B-C. Timecourse assay data.** Molecules of FP per cell calculated using calibrated FP and calibrated OD700 (B) and observed vs. expected OD600 values (C), plotted against time. Expected values were calculated from the OD700 values and the measured OD600~OD700 relationship of the 0% arabinose sample (see A). **D. Effect of FP abundance on OD600 error.** The OD600 error was plotted against abundance of each FP in molecules/cell. Grey shading indicates the mean of the samples without arabinose ± 2*standard deviations.