

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Absolute quantification of protein copy number in single cells with immunofluorescence microscopy calibrated using single molecule microarrays

Journal:	<i>Analytical Chemistry</i>
Manuscript ID	ac-2020-05177a.R1
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Chatzimichail, Stelios; Imperial College London, Surgery & Cancer Supramaniam, Pashiini; Imperial College London, Chemistry Salehi-Reyhani, Ali; Imperial College London, Surgery & Cancer

SCHOLARONE™
Manuscripts

Absolute quantification of protein copy number in single cells with immunofluorescence microscopy calibrated using single molecule microarrays

Stelios Chatzimichail,¹ Pashiini Supramaniam² and Ali Salehi-Reyhani^{1*}

¹ Department of Surgery & Cancer, Imperial College London, London, W12 0HS, UK

² Department of Chemistry, Imperial College London, London, W12 0BZ, UK

*Correspondence to ali.salehi-reyhani@imperial.ac.uk

Abstract

Great strides toward routine single cell analyses have been made over the last decade, particularly in the field of transcriptomics. For proteomics, amplification is not currently possible and has necessitated the development of ultra-sensitive platforms capable of performing such analyses on single cells. These platforms are improving in terms of throughput and multiplexability but still fall short in relation to more established methods such as fluorescence microscopy. However, microscopy methods rely on fluorescence intensity as a proxy for protein abundance and are not currently capable of reporting this in terms of absolute copy number. Here, a microfluidic implementation of single molecule microarrays for single cell analysis is assessed in its ability to calibrate fluorescence microscopy data. We show that the equivalence of measurements of the steady-state distribution of protein abundance to single molecule microarray data can be exploited to pave the way for absolute quantitation by fluorescence and immunofluorescence microscopy. The methods presented have been developed using GFP but are extendable to other proteins and other biomolecules of interest.

Keywords – single molecule microarray, quantitative immunofluorescence, calibration

Introduction

A major goal in systems biology is to provide a complete and quantitative description of cellular behaviour. This task, however, has been hampered by the difficulty in measuring absolute protein abundances, their variation and exactly which proteins are expressed in a particular cell to begin with. While sufficient in most circumstances, relative quantitation of protein expression or abundance can be misleading. The case may arise whereby the relative abundance of a specific protein to itself or another species in a cell may not change in response to different conditions but may change significantly in terms of absolute quantity. Absolute quantification is attractive for the modelling of cellular networks and, perhaps most importantly, it also allows comparisons of results between different laboratories to be made.

Mass spectrometry (MS) is capable of analysing the proteome to considerable depth where it can catalogue protein expression, protein interactions and identify sites of protein modification.¹ A major limitation to MS and bulk methods of analysis is the averaging of protein concentration within a sample which masks any cellular heterogeneity within a cell population. Great strides toward single cell metabolomics have been made^{2,3}, while single cell proteomics using MS has only very recently been reported. In a landmark achievement for MS, Budnik *et*

1
2
3 *a/* classified cells and explored the relationship between mRNA and protein abundances in
4 single mammalian cells by characterising over a thousand proteins.⁴
5

6 Single cell protein expression is more routinely approached using flow cytometry,
7 predominantly to assess cell surface proteins in heterogeneous cell populations.⁵ Measures
8 of cellular fluorescence, a consequence of the excitation of a fluorescence antibody binding to
9 its antigen or a fusion protein, are generally expressed in arbitrary units. As such, the absolute
10 numbers of the respective cell surface proteins of interest remain unknown. Bead-enhanced
11 quantitation in flow cytometry was initially used to enumerate the absolute numbers of different
12 cell types in complex blood-derived samples.⁶ Methods have since been developed to use
13 beads with a defined number of fluorescent species on their surface to calibrate fluorescent
14 measurements of cell surface markers.^{7,8} Flow cytometry can be used to probe intracellular
15 protein expression; however, there are a lack of calibration approaches for intracellular
16 proteins. Flow cytometry workflows are capable of measuring thousands of cells per second;
17 however, using small samples of a few hundred cells is often not possible due to substantial
18 cell losses. A number of microfluidic chip methods for single cell proteomics have emerged
19 over the last decade, which are inherently more capable in studying rare cells.^{10–15} These
20 single cell platforms rely on the spatial separation of microarrayed spots or barcodes to assay
21 and identify specific proteins. One of the most notable has been the single cell Western Blot
22 from the Herr lab;^{14,16} a single cell analogue of the conventional technique is performed on a
23 device that comprises an array of microwells moulded in a thin layer of a polyacrylamide gel.
24 Single cell microfluidic platforms are rapidly evolving and are beginning to be applied in
25 clinically relevant applications; however, the barriers to entry remain high and the ubiquity and
26 unique advantages of fluorescence microscopy are limiting their adoption.
27

28
29 Analysing cells with fluorescence microscopy imaging has several advantages over integrative
30 measurements such as flow cytometry or microfluidic based methods. Microscopy, of course,
31 can preserve protein localisation such that morphological, spatial and even temporal
32 information can be collected across sub-cellular and supra-cellular length scales.¹⁷
33 Fluorescent labels may be introduced into cells in several ways but specific proteins are
34 typically labelled by either immunofluorescence or the genetic introduction of a fluorescent
35 protein.^{18,19} Quantitative microscopy depends on a consistent ratiometric relationship between
36 the fluorescent signal intensity and the number of target proteins. Fluorescent fusion proteins
37 have a 1:1 stoichiometry between the fluorescent protein and the target, which is either
38 inserted into the genome of the cell or can be expressed from a plasmid. While more
39 challenging, introducing the fusion protein into the genome means that the protein may be
40 expressed from its endogenous promoter at its endogenous level and can be more biologically
41 relevant.²⁰ Of course, fusion proteins can be undesirable since any modification to a protein
42 can interfere with its stability, activity and function.²¹ To address this, Lo *et al.* developed the
43 protein quantitation ratioing technique, whereby a genetic tag separates during translation to
44 produce a stoichiometric ratio of a fluorescent protein and a protein of interest from a single
45 strand of RNA.²² Since they become separated this minimises any interference with the target
46 while also allowing the fluorescent protein to be used as a readout for the number of proteins
47 of interest.
48

49 Immunofluorescence requires cells to be first fixed to conserve their structure and allow their
50 posterior analysis. Depending on the fixative, this preservation may or may not involve the
51 cross-linking of proteins. Once fixed, cells are then permeabilised to allow antibodies, typically,
52 to access the intracellular volume and label proteins. For absolute quantification,
53
54
55
56
57
58
59
60

immunofluorescence is a much greater challenge owing to the problems in performance, particularly specificity, and variability of any antibody (performance, structure, affinity), access to the target protein epitope that may be fully or partially blocked, the presence of any post-translational modifications on the target protein and the cross-reactivity of the antibody for non-target species. The issues that antibodies can present in general have been captured in a number of thought-provoking reports.^{23–25} Furthermore, the choice of fluorophore and their number, and by how much this varies, conjugated to the antibody may also need to be taken into consideration.^{26,27} It quickly becomes apparent, then, that absolute quantification is much harder to achieve in practice.

Toward this goal, Toki et al. took steps to show the quantitative potential of immunofluorescence using mass spectrometry.²⁸ In this manuscript, we demonstrate a method to achieve absolute quantification of single cell intracellular protein copy numbers using immunofluorescence microscopy (**Fig. 1**). An immunofluorescence method is shown to produce data with a consistent ratiometric relationship between the fluorescent signal intensity and the protein copy number of GFP in a human breast cancer cell line. The absolute quantification of protein copy number in these cells was achieved using a single molecule microarray method and used to calibrate the immunofluorescence data. We show that the intensity distribution of GFP protein abundance in these cells as determined by cell fluorescence is reproduced by immunofluorescence and by our calibrated microarrays. The latter technique being capable of quantifying absolute copy number in a cell and so allows us to in turn ‘calibrate’ the immunofluorescence data. The method presented does not rely on a precise measurement of a fluorescent standard but on the equivalence between single cell measurements of the steady-state distributions of protein abundance per cell in a population by two different methods. The approach is conceptually simple and may be easily adopted by other labs.

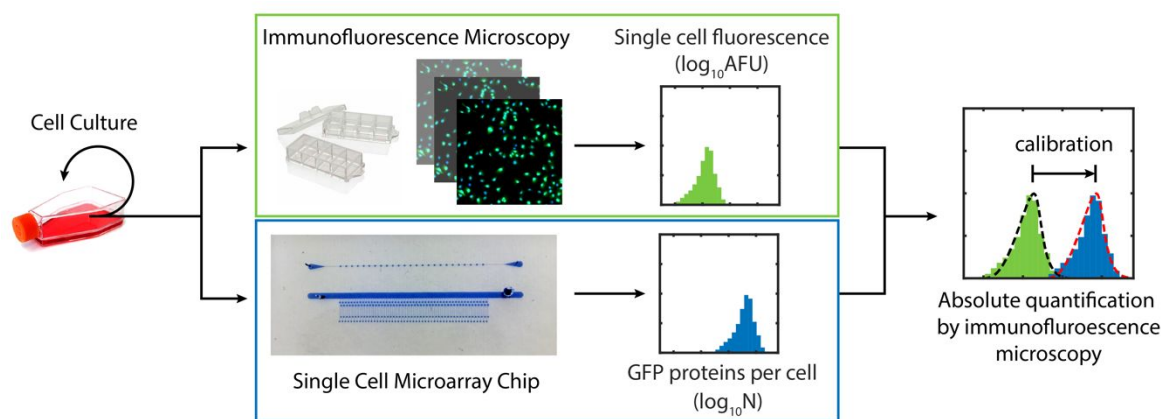


Fig. 1. A summary of the workflow of the proposed protocol to achieve absolute quantification of single cell intracellular protein copy number using immunofluorescence microscopy. Immunofluorescence methods can produce data whereby the protein copy number per cell is assessed using fluorescent signal intensity (AFU; arbitrary fluorescent units). Using the Single Cell Microarray Chip, the absolute copy number (N) of a target protein can be measured. By comparing the summary parameters of the resulting distributions, immunofluorescence microscopy data may then be calibrated from arbitrary to absolute units.

Methods

Single Cell Immunofluorescence

Immunostaining of fixed cells. Cells were fixed, permeabilised then immunostained following methods described in Stadler et al.²⁹. At ~80% confluence cells were removed from the incubator and washed 3× with PBS. For fixation, cells were treated with freshly made 4% paraformaldehyde in cell growth media then washed 3× with PBS. Cells were permeabilised using 0.1% (w/v) Triton X-100, replacing the solution every 5 min for a total exposure of 15 min. Cells were blocked with 4% bovine serum albumin in PBS (PBSA), replacing the solution every 5 min for a total exposure of 10 min. Cells were then incubated overnight at 4°C with 1 µg mL⁻¹ anti-copGFP antibody (Evrogen, Europe) and 1 µg mL⁻¹ goat anti-rabbit labelled with Alexa Fluor 647 (Lot 1910774; Thermo Fisher, Europe) in 4% PBSA. The degree of labelling is 5, i.e. the average number of moles of dye per mole of antibody, and should be taken into account when calibrating, therefore sources which provide the molar ratio and variance of dye to antibody are preferable or, alternatively, must be measured by absorbance. After washing 4× with PBS for 10min, cell nuclei were stained with 300 nM Hoechst 33342 in PBS and washed again 4× with PBS for 10min. Wells were filled with 78% glycerol in PBS, sealed and stored at 4°C before image acquisition.

Immunofluorescence image acquisition. Wide-field fluorescence microscopy was performed on an inverted microscope (Nikon Ti-E, Nikon, Japan) using a mercury lamp (Nikon, Japan) using standard DAPI and FITC filter sets (Nikon, Japan), imaging Hoechst 33342 stained nuclei and GFP, respectively. If required, a Cy5 filter set (Chroma, USA) was used to image dye-labelled antibodies. Images were acquired with an electron-multiplied CCD camera (IXON DU-897E; Andor Technologies, Ireland). The EM-CCD camera settings were as follows; 200 ms acquisition time, electron multiplier gain 10 and the readout mode set to 1MHz at 16-bit. Tile scans of 20 x 20 fields of view (~2.66 mm x 2.66 mm) were acquired using a 60× NA = 1.49 oil immersion objective for each fluorescence channel with an additional brightfield scan using a standard halogen lamp. The focal plane was maintained by a Nikon Perfect Focus System (Nikon, Japan) and a motorised stage ensured good overlap between image channels. The microscope platform was controlled using the NIS Elements software (Nikon, Japan) during acquisitions to automatically acquire large field images.

Immunofluorescence image analysis. CellProfiler (Broad Institute) was used to analyse and segment images to identify cell nuclear and cytoplasmic extents, aided by the nuclear stain or GFP fluorescence. The total fluorescence intensity was determined by summing all pixel intensities for each cell identified. Datasets were exported, and further analysis was performed using MATLAB.

Single Molecule Microarray Chip

Chip fabrication. Based on previous work, a single molecule microfluidic platform to assess protein copy number with single cell resolution was used (**Fig. 4**).^{10,11} See supporting information for further details.

Cell pull-downs. Cells were harvested from 4 wells of the 8 well chambered coverglass used for culturing using trypsin/EDTA. Cells are re-suspended in a solution of 4% bovine serum albumin in PBS and put on ice. The chip is secured on the microscope platform and cells in suspension are flowed through the main channel, to which the analysis chambers are

connected. Cells are manually isolated into chambers using a single beam optical trap (100-200 mW, 1070nm, Ytterbium fibre laser; IPG Laser, Germany) focussed using a 60× NA = 1.49 oil immersion objective. GFP cellular fluorescence is unaffected by the optical trap (**Fig. S1**). Cells are subsequently lysed by optical lysis achieved by focussing a single 6 ns pulse (1064nm, Surelite SLI-10 Nd:YAG; Continuum, USA) into the vicinity of a cell; cells are sheared and lysed by the action of the laser. GFP cellular fluorescence is unaffected by the laser lysis pulse (**Fig. S2**). Upon lysis, cellular protein is free to diffuse into the chamber and able to bind to an antibody spot. Spots are imaged pre- and post-lysis using single molecule total internal reflection fluorescence (TIRF) microscopy (488 nm Versalase; Laser 2000, UK). During TIRF imaging, we maintain a low power of the excitation laser (1.5 mW) and short exposure times (200 ms) in order to minimise photobleaching. When developing the protein assay, each spot is imaged every 10 min for 30 min then every 20 min for a further 60 min, a total of 7 images over 90 min. Each spot is exposed to the excitation laser for a total of 1.4 s. This time series is enough to capture a binding curve for each spot. Equilibrium is achieved within 30-40 min as confirmed by imaging. It is only necessary, then, to capture a single image of each spot once equilibrium has been reached. The copGFP variant of GFP is not particularly susceptible to photobleaching and so this along with the considerations above ensure minimal photobleaching of fluorophores bound to the microarray spot.

Calibration. The microfluidic chip additionally consists of an independently addressed 'Calibration channel' (**Fig. 4A**). Placed along the length of this channel are 25 chambers of the same dimensions as the single cell analysis chambers. Standard solutions of protein of known concentration are flowed through the chambers and allowed to reach binding equilibrium with antibody spots aligned into each chamber. A concentration series is performed to determine the single molecule count at equilibrium for known concentrations of recombinant GFP protein, from 10 to 10^9 molecules per chamber, including buffer only controls to assess the response at zero concentration. The sequence of solutions that are flowed through the calibration channel are from low to high concentrations. Since spots are imaged and analysed at equilibrium there is no requirement for the device to be rinsed between conditions. In each chamber, the antibody spots nominally contain the same concentration of antibody immobilised in each microarray spot. This will enable single molecule counts from lysed single cells to be converted to absolute protein copy numbers per cell. See supporting information for further details.

Single molecule image analysis. Methods have been described in detail elsewhere.³⁰ We perform 2D Gaussian peak fitting by implementing the GDSC plugin by Alex Herbert for ImageJ/Fiji.³¹ The total number of single molecules is defined as the number of peaks in the image which converge to a 2D Gaussian profile, determined by least squares fitting of the data. For images where there is a significant degree of single molecule overlap, the number of single molecules is estimated from dividing the total intensity of a thresholded image by the average single molecule intensity.

Results & discussion

Precision of fixation/permeabilization

Preparative techniques for immunofluorescence are well established and have been extensively characterised. Here, the concern is that the procedures to fix and stain cells may result in a measured distribution of single cell protein expression that significantly differs from

the 'true' distribution. The methods used here have been optimised for cytoplasmic protein staining of mammalian cells.^{29,32} However, the issue remains that, while strong fixation ensures proteins are immobilised and localisation information is preserved, epitope recognition and penetration of antibodies may be impaired to a degree which precludes the overall metrological goal of this work.

To determine if this is the case, a stably transfected MCF7 cell line expressing GFP was used to determine the precision of the fixation, permeabilization and staining process and used alongside a non-transfected MCF7 cell line as a control. MCF7 and MCF7/GFP cells were grown in fibronectin coated chambers of an 8 well chambered coverglass. Live cell microscopy was first performed to determine the distribution of GFP protein abundance in MCF7/GFP cells and would serve as the 'ground truth' distribution (**Fig. 2A**).

Cells were fixed with and without a subsequent permeabilization step as described in the methods. The results show that the distribution of single cell GFP protein expression is preserved (**Fig. 2C & 2D**).

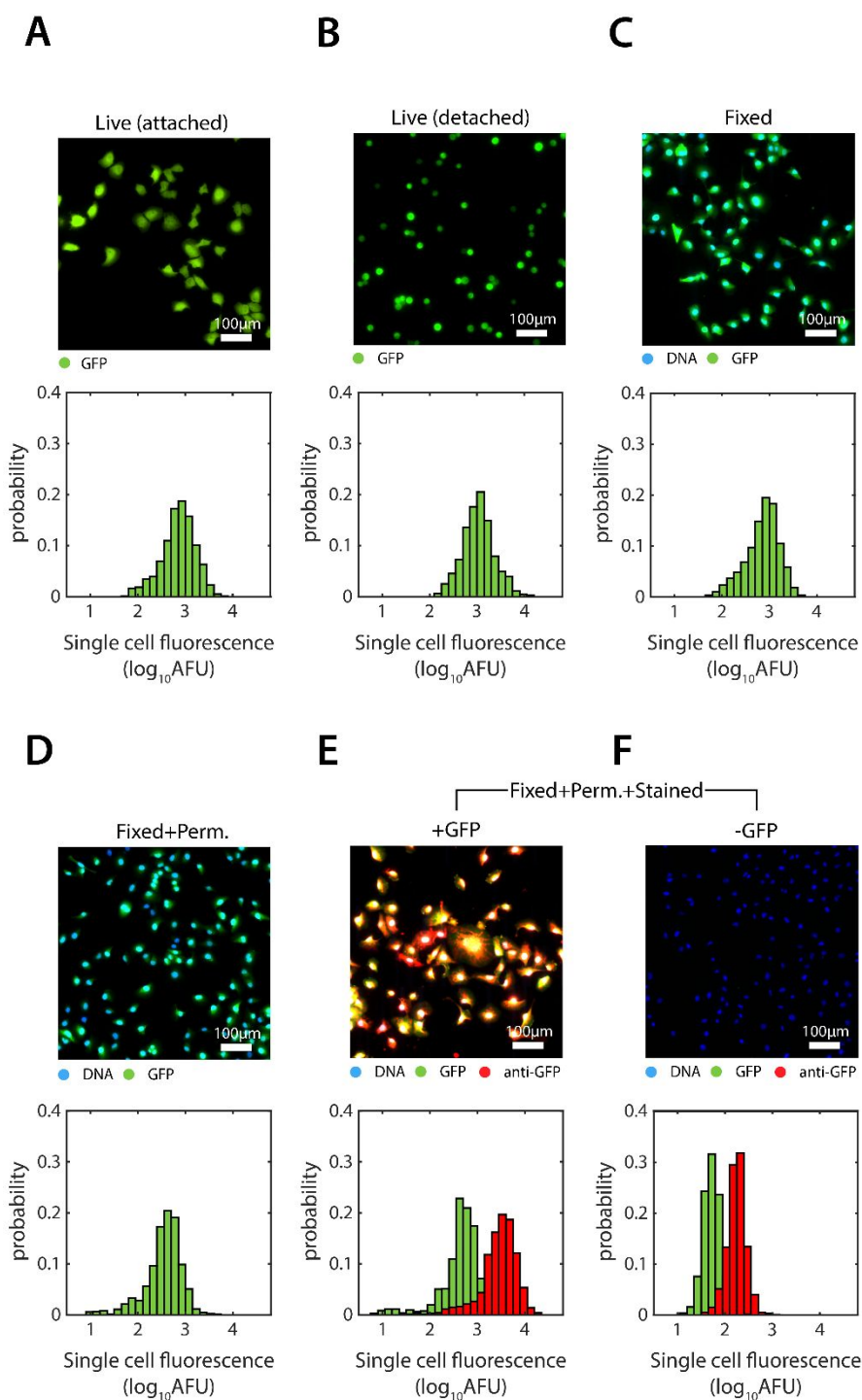


Fig. 2. Measuring single cell protein distributions using fluorescence and immunofluorescence microscopy, units of arbitrary fluorescence units (AFU). GFP abundance in MCF7-GFP cells is quantified for live cells that are (A) attached ($n = 1565$) or (B) detached ($n = 839$), for attached cells that are (C) fixed ($n = 1613$) or (D) fixed and additionally permeabilised ($n = 1450$), and for (E) MCF-GFP ($n = 803$) and (F) MCF7 (no GFP control, $n = 1519$) cells that have been fixed, permeabilised and immunofluorescently stained. Figures show representative fluorescence & immunofluorescence images of cells along with corresponding histograms of single cell fluorescence intensity as a measure of protein abundance.

A comparison of the fluorescence intensities from cell regions and background of both fixed MCF7/GFP and MCF7 control cells confirmed that the fluorescence signal was specific to GFP rather than arising from artefacts of fixation or any other form of endogenous autofluorescence, such as that from NADH or FAD.³³

The steady-state GFP protein distributions are unimodal and asymmetric, a consequence of relative rates of mRNA and protein transcription and translation, respectively, to their degradation.³⁴ Protein distributions may be represented by their mean and variance; however, there are potential problems in doing so for non-symmetric distributions, which have been discussed elsewhere.³⁵ The complete probability distribution of cellular fluorescence or protein abundance is required to characterise certain models of gene expression.³⁶ It has been shown that gamma distributions can well approximate particular protein distributions and is confirmed for GFP in the cell line used here. By fitting, the gamma parameters of shape (k) and scale (θ) were estimated and summarised for all experiments in **Fig. 2** (see **Fig. 3**).

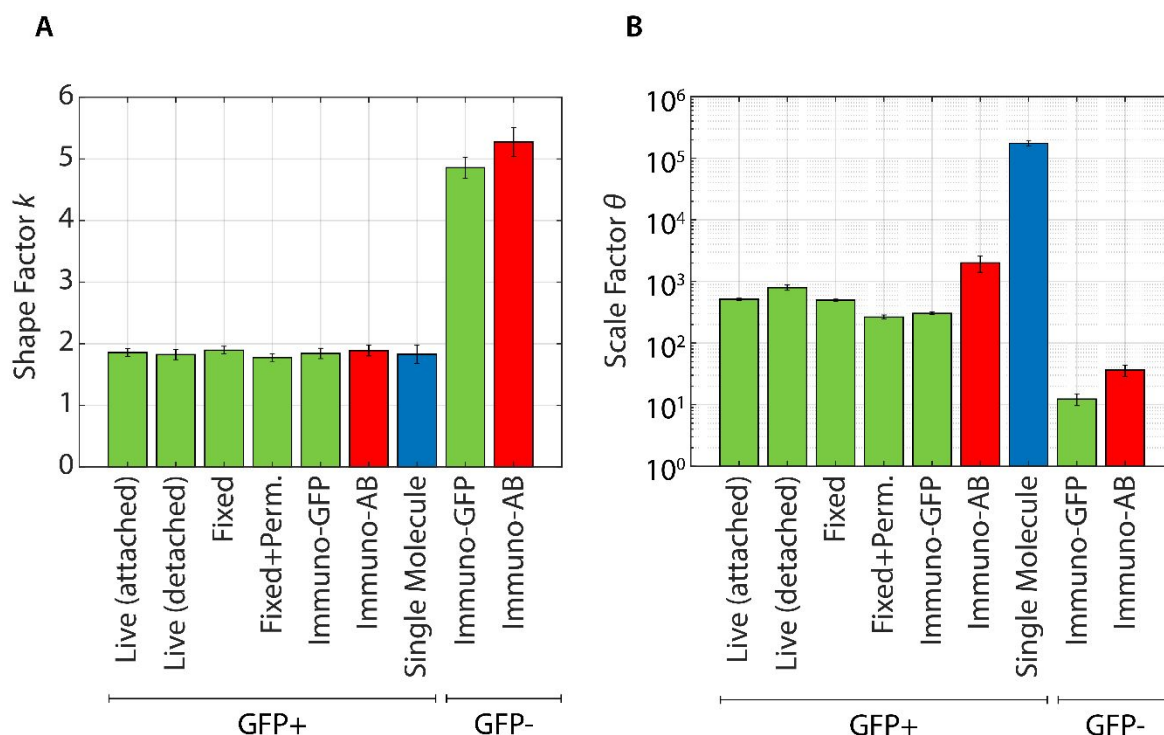


Fig. 3. Summary parameters of gamma fits to fluorescence, immunofluorescence and single molecule microscopy data presented in Fig. 2 and Fig. 4. MCF7-GFP and non-transfected MCF7 cells are denoted by GFP+ and GFP-, respectively. Key: “Live (attached)” denotes cells that are grown upon and imaged attached to coverslips; “Live (detached)” denotes cells which have been enzymatically detached from substrates in culture prior to imaging; “Fixed” denotes cells which have been grown upon coverslips and treated with a fixative (see Methods); “Fixed+Perm” denotes conditions as in “Fixed” where cells are treated subsequently with a permeabilising agent; “Immuno-GFP” and “Immuno-AB” denotes cells which have been treated following the immunofluorescence staining protocol for both GFP+ and GFP- cell lines. “Immuno-GFP” and “Immuno-AB” refers to the fluorescence signal arising from the GFP protein and labelled antibody channels, respectively. “Single Molecule” refers to measurements of GFP+ cells using single molecule microarrays.

Precision of Immunofluorescence

The results so far demonstrate that the probability distribution of the abundance of an intrinsically fluorescent protein in a cell population as measured by direct fluorescence is not significantly altered during fixation or permeabilisation. Next, an anti-GFP antibody would be needed to compare the resulting distribution as determined by immunofluorescence to the live-cell and direct fluorescence fix/permeabilised distributions.

Immunofluorescence of the MCF7 and MCF7/GFP cell lines was repeated where cells would be incubated with a primary anti-GFP antibody and fluorescent secondary antibody. The fluorescence of GFP permits a cell-to-cell comparison between the antibody complex that binds and the GFP protein itself. The staining of the cells by antibody shows that the cells have been comprehensively permeabilised, rendering the pH of the intracellular compartments to that of the buffer i.e. pH 7.4 for PBS. Therefore, we expect the variation in fluorophore intensity due to pH to be minimised or removed altogether.

The distribution of protein abundance as measured by single cell immunofluorescence (**Fig. 2E & 2F**) was then compared to the single cell distributions of live cell fluorescence, for both attached and detached cells, and fluorescence of cells that were fixed and cells that were fixed and permeabilised. Fitting of the distribution parameters showed good agreement with all the distributions measured (**Fig. 3**), particularly the distribution of GFP in live cells (attached), which served as the 'ground truth' (**Fig. 2A**).

Absolute quantification of Cellular GFP

It has been shown that the 'ground truth' distribution of GFP protein abundance as measured by live cell fluorescence is preserved when performing single cell immunofluorescence microscopy. To determine the absolute protein copy numbers of GFP, a single molecule microarray microfluidic-based method was used (**Fig. 4**). The chip permits a reduced volume antibody microarray for the sensitive detection of proteins in single cells.^{10,11} Cells are harvested from culture and introduced into the chip by a syringe pump. An optical trap formed part of the microscope setup and was used to serially isolate individual cells into separate analysis chambers. Cells parked in the chambers were fluorescently imaged before being optically lysed. Upon cell lysis, cellular material diffuses within the chamber and target analytes bind to the immobilised antibody spot. This is where the scaling advantages of using a microfluidic chip are most advantageous, since the nanolitre-scale chamber volume reduces the time to achieve equilibrium and raises the effective concentration of analytes and their capture agents.

Total internal fluorescence (TIRF) microscopy was used to image spots and image datasets were analysed using single molecule detection algorithms to enumerate the amount of GFP bound to each microarray spot. The number of single molecules counted per spot was converted to an absolute protein copy number by way of a calibration curve. Recombinant protein is introduced into the calibration channel of the chip and concentrations spanning 10 to 10⁹ proteins per chamber are tested to establish the calibration curve (**Fig. 4E**). The majority of binding of GFP is complete within 30 min.; to ensure equilibrium has been achieved and the assay is fully completed we image after 60 min. Preparation of pre-fabricated chips (filling with blocking buffer, mounting on microscope stage) can take up to 30 min. The total operation

time from harvesting cells to completing lysis can be achieved within 2 hours ($n = 100$ cells), with an additional hands-off time of 60 min. required for completion of the binding assay and TIRF imaging of the spots. Parallel operation can enable several chip experiments to be completed in a working day by an experienced user.

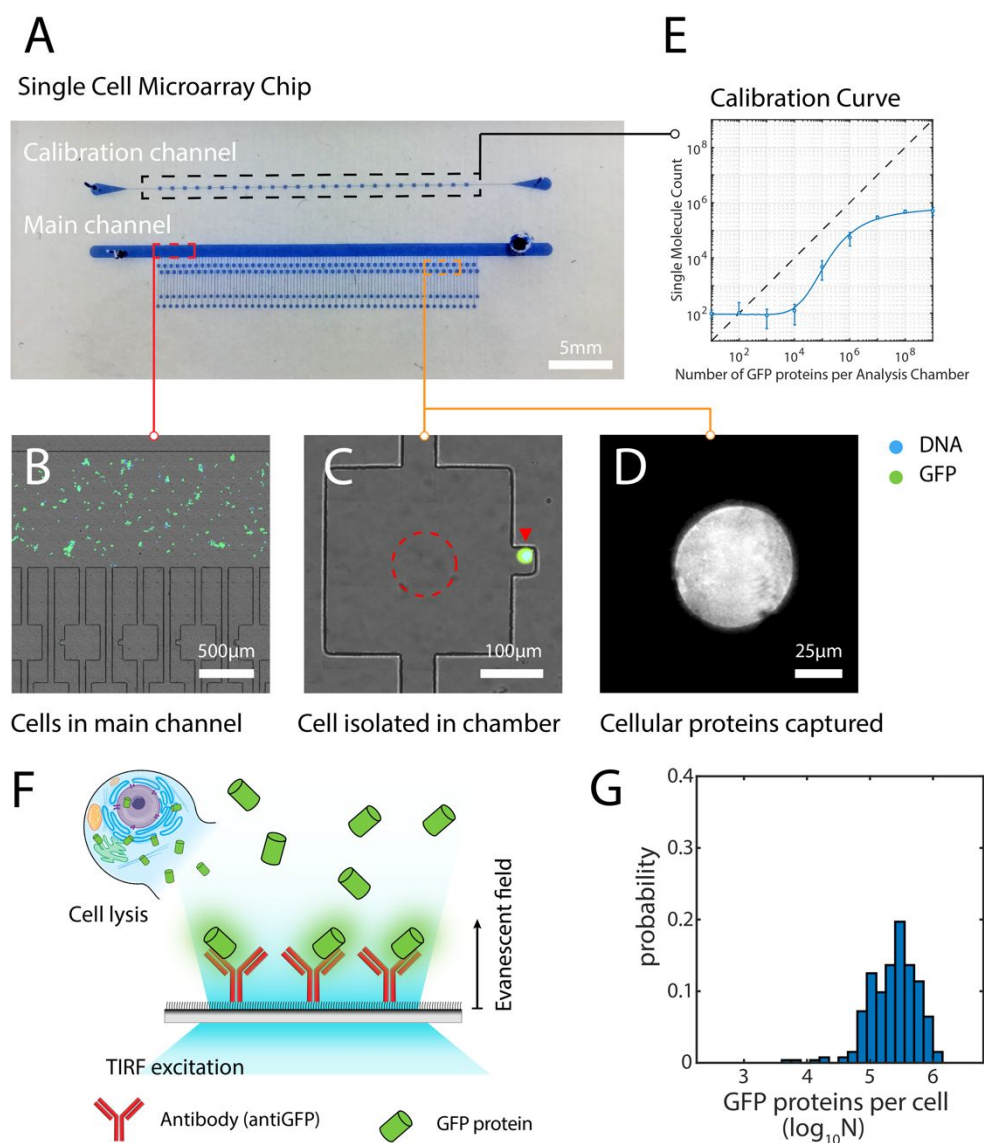


Fig. 4. Using a microfluidic single molecule microarray technique to determine the absolute protein abundance of GFP per single cell. The (A) Single Cell Microarray Chip is formed of a main and calibration channel. (B) Cells are flowed through a main channel from which (C) they are isolated into individual analysis chambers using an optical trap (red arrowhead). The red-dashed circle demarcates (D) a microarrayed antibody spot present in each chamber which captures GFP protein from lysed cells. (E) A calibration curve is established by flowing known concentrations of recombinant GFP protein through the calibration channel; used to convert single molecule counts per spot to number of GFP proteins per chamber and by extension, per single cell. (F) Protein binding to the antibody spot is imaged using TIRF microscopy. The evanescent field extends ~100 nm from the coverslip surface and so only proteins bound to the spot are imaged, with high contrast and strong background rejection enabling single molecule resolution. (G) The histogram is a measure

of the distribution of GFP protein abundance per single cell (N) within a population of MCF7-GFP cells (n = 264).

For the assay presented here, the average background single molecule count and its variation sets the limit of detection (LOD). The LOD at $\mu_{SM}^{bkd} + 3\sigma_{SM}^{bkd}$ (572 molecules) corresponds to $2.22 \pm 0.24 \times 10^4$ GFP proteins, where μ_{SM}^{bkd} is the mean single molecule background count and σ_{SM}^{bkd} is its standard deviation. The dynamic range is defined between the LOD and the limit of saturation (LOS). The LOS at $\mu_{SM}^{sat} - 3\sigma_{SM}^{sat}$ is (3.13×10^5 molecules) corresponds to $1.19 \pm 0.04 \times 10^7$ GFP proteins, where μ_{SM}^{sat} is the mean single molecule count at saturation and σ_{SM}^{sat} is its standard deviation. The dynamic range spans 2.73 ± 0.06 orders of magnitude. The dynamic range of the calibration curve is limited by non-specific binding (NSB) and the saturation limit, or capacity, of the capture antibody spot. The functionalised surfaces used here are optimised in supporting high density microarrays and limiting NSB to ~ 0.0056 molecules μm^{-2} (approx. 100 molecules per field of view). This compares favourably with other single molecule pull down assays.³⁷ From data reported in Jain *et al.*, we estimate NSB on their surfaces to be ~ 0.038 molecules μm^{-2} (approx. 30 molecules per field of view).³⁷ Surfaces that better reject NSB will lead to lower LODs. For example, with reference to **Fig. 4E** reducing NSB to approx. 10 molecules per field of view would result in a LOD of approx. 10^3 GFP proteins per analysis chamber. Modifications to single cell chamber volume (**Fig. 4C**), antibody affinity and capture spot size (**Fig. 4D**) that result in improved LODs and dynamic ranges more suited to low copy number proteins ($<10^3$ proteins) have been explored in detail elsewhere.³⁸ The methods used here may be extended to reject non-specifically bound molecules by the kinetic fingerprinting single molecule binding events.³⁹

A histogram of absolute GFP copy number per single cell is shown in **Fig. 4G** along with the corresponding fitting parameters in **Fig. 3**. The shape parameter of the absolute number distribution is in good agreement with the values measured by fluorescence and immunofluorescence microscopy. The data from which the distributions in **Fig. 2** are derived are measured in arbitrary fluorescence units (AFU) and meaningless beyond any relative assessment of cellular protein expression. The equivalence between the single cell microarray and immunofluorescence data demonstrates in principle how the latter may be calibrated to achieve absolute quantification. The process of calibrating the distributions is demonstrated in **Fig. 5**. By normalising the immunofluorescence distribution to the 'single molecule' distribution, a calibration factor may be calculated for the immunofluorescence microscopy data in order to *calibrate* units of protein abundance from AFU to absolute protein copy number. For the distributions shown here, θ_{SM} ($1.74 \pm 0.16 \times 10^5$ proteins) and θ_{AB} ($2.00 \pm 0.59 \times 10^3$ AFU) are the scale parameters of the fitted gamma distributions for the single molecule and immunofluorescence results, respectively (**Fig. 3**). The number of dyes per antibody must be taken into account when calibrating the immunofluorescence data (see methods) and the overall calibration factor is numerically equal to 17.4 ± 5.4 proteins AFU^{-1} .

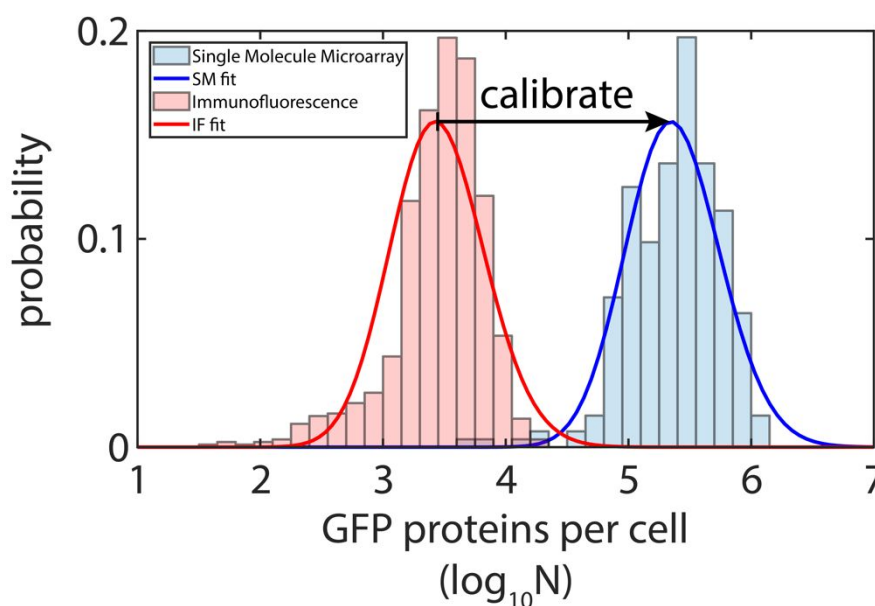


Fig. 5. Calibrating fluorescence and immunofluorescence data. Provided the methods employed to measure single cell protein abundances are accurately measuring the ‘shape’ parameters of the ground truth distribution, a distribution which measures protein abundance in arbitrary fluorescence units may be calibrated to absolute units. Here, GFP protein abundance in these cells is well approximated by a gamma distribution. In this case, calibration is achieved by ratioing the scale parameters of the distributions. The immunofluorescence data presented in Fig. 3 “Immuno-Ab” is represented here by the red histogram. See text for numerical details.

The accuracy of this approach requires the stoichiometry of labelling of the fluorescent protein with the target to be known. The formation of the chromophore of GFP depends on the correct folding of the protein. This makes GFP a good candidate for these goals given that the protein is itself fluorescent and the ‘labelling’ is 1:1 for active GFP. The epitope may remain intact in non-folded GFP; however, will not be counted in our assay. In the introduction, we noted a number of confounding factors when employing immunofluorescence quantitatively. Here we are uncertain as to the exact degree by which each of these factors affect the result. However, we contend that these are not significant enough to preclude the metrological goals of this work as evidenced by the results presented in **Fig. 3** and **Fig. 5**. Of course, we expect these confounding factors to become more important when extending this more broadly to endogenous proteins. In addition, other factors that will need to be considered include post-translational modifications and the presence of binding partners. Generally, proteins of interest may be labelled genetically or chemically with a fluorescent tag. Alternatively, the antibody assay used here would be extended to include a labelled detection antibody in a sandwich format: endogenous proteins (non-fluorescent) would be captured by the immobilised antibody (non-fluorescent) and the binding of a second labelled antibody (fluorescent) from solution would allow the complex to be imaged. For the case of non-fluorescent endogenous proteins, the same principle of equivalence would apply to achieve absolute quantification. Furthermore, since cells, as biological entities, can be unpredictable then apparently unrelated changes, such as culture medium or passage number, could alter their protein expression pattern. To act as either primary or secondary standards, further investigation is required.

Importantly, what we have shown in the case of GFP in these cells is that the shape of the distribution of protein abundance is preserved.

Conclusions

Fluorescence and immunofluorescence microscopy are widely used techniques in measuring the expression and subcellular location of proteins. Here, it's been shown that single molecule microarrays are a suitable calibrant in achieving absolute quantification using fluorescence and immunofluorescence single cell data. The approach was tested using a fluorescent reporter cell line using GFP to determine the accuracy and precision of the method. We believe this is an important step toward absolute quantification of protein copy number using high-content fluorescence microscopy. In the process of developing the method, it was shown that the distributions of fluorescent protein in live cells is equivalent to the calibrated single molecule array data. This would imply that this method has the potential in calibrating time-resolved live cell microscopy data.

The approach relies on the equivalence of each method to faithfully measure the steady state distribution of protein expression within a population of cells. A calibration factor can then be established using the summary parameters of the distributions. The benefit of using single molecule methods is that the detected molecules' fluorescence is digitised and makes it more robust to systematic variation, which is desirable for any standardisation framework achieving instrument independent measurements. Since the distribution of *many* cells in a population is being calibrated and not individual cells, it might be expected that this confers some robustness to the technique to the variation of the imaging parameters of the fluorescence measurements. For instance, microscopy measurements using this cell line on different instrumentation platforms in different laboratories would be expected to produce a single cell distribution of GFP protein abundance that will be identical, or at least insignificantly different, in terms of the fitted gamma distribution *shape* parameter yet likely differ in terms of the numerical value of the *scale* parameter. In that case, a calibration factor may be calculated using the single molecule results reported here. This would be of interest and an important point to consider for anyone wishing to adopt these methods. There will of course be limitations and steps must be made to ensure sufficient signal to noise ratios of cellular fluorescence or single molecule intensity to background. Otherwise, the quantification of low concentrations may result in numerical artefacts if imaging conditions become substantially altered.

Using a cell line expressing free GFP has been useful in developing the method but perhaps of limited interest biologically. Strategies to engineer cells to express fluorescently labelled proteins are becoming ever more accessible. However, it is important that the functionality and expression level of a fluorescent fusion protein be similar to that of the wild-type protein. Encouragingly, Stadler *et al.* undertook a systematic analysis of over 500 human proteins and showed that for 80% of these proteins the same subcellular localisation pattern was obtained using fluorescent protein tagging in live cells or immunofluorescence.¹⁷ Cell lines are regularly employed as standards but not in a purely metrological sense. They may be authenticated using genetic profiling and due to historical cases of contamination have led to researchers annotating, sometimes as a pre-requisite, biologically relevant publications with sequence data.^{40,41} However, cell line genomes can be unstable and cultures will diverge due to accumulated mutations.⁴² While these and other considerations will need to be taken into

account if cells lines are to be used as metrological standards, this is an important technical step forward in achieving absolute protein quantification with high content fluorescence and immunofluorescence microscopy.

Supporting Information

Background: Comparison Methods for Measuring Protein Expression in Single Cells. Methods: Method of cell culture, Single Molecule Microarray Chip Fabrication, Effect of the Optical Trap on Cellular Fluorescence, Effect of the Optical Lysis Laser on Cellular Fluorescence, Single Molecule Count Calibration. Results: Minimum detectable change in copy number.

Data Availability

All relevant data are available from the authors on reasonable request.

Acknowledgements

This work was supported by an Imperial College Fellowship and an Engineering and Physical Science Research Council (EPSRC) Innovation fellowship awarded to ASR.

Author Contributions

S.C. performed single molecule experiments. P.S. aided in microarray and microfluidic chip fabrication. A.S.-R. conceived and designed the project, performed microscopy experiments and wrote the paper. All authors reviewed drafts of the paper.

Competing Interests

The authors declare no competing interests, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Corresponding Author

Correspondence to Ali Salehi-Reyhani (ali.salehi-reyhani@imperial.ac.uk).

References

- (1) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483–490.
- (2) Guillaume-Gentil, O.; Rey, T.; Kiefer, P.; Ibáñez, A. J.; Steinhoff, R.; Brönnimann, R.; Dorwling-Carter, L.; Zambelli, T.; Zenobi, R.; Vorholt, J. A. Single-Cell Mass Spectrometry of Metabolites Extracted from Live Cells by Fluidic Force Microscopy. *Anal. Chem.* **2017**, *89* (9), 5017–5023.
- (3) Masujima, T. Live Single-Cell Mass Spectrometry. *Anal. Sci.* **2009**, *25* (8), 953–960.
- (4) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. SCoPE-MS: Mass Spectrometry of Single Mammalian Cells Quantifies Proteome Heterogeneity during Cell Differentiation. *Genome Biol.* **2018**, *19* (1), 161.
- (5) Adan, A.; Alizada, G.; Kiraz, Y.; Baran, Y.; Nalbant, A. Flow Cytometry: Basic Principles and Applications. *Crit. Rev. Biotechnol.* **2017**, *37* (2), 163–176.
- (6) Montes, M.; Jaensson, E. A.; Orozco, A. F.; Lewis, D. E.; Corry, D. B. A General Method for Bead-Enhanced Quantitation by Flow Cytometry. *J. Immunol. Methods* **2006**, *317* (1–2), 45–55.
- (7) Serke, S.; van Lessen, A.; Huhn, D. Quantitative Fluorescence Flow Cytometry: A

- Comparison of the Three Techniques for Direct and Indirect Immunofluorescence. *Cytometry* **1998**, 33 (2), 179–187.
- (8) Vasilyev, F. F.; Lopatnikova, J. A.; Sennikov, S. V. Optimized Flow Cytometry Protocol for Analysis of Surface Expression of Interleukin-1 Receptor Types I and II. *Cytotechnology* **2013**, 65 (5), 795–802.
- (9) Li, X.; Fan, B.; Cao, S.; Chen, D.; Zhao, X.; Men, D.; Yue, W.; Wang, J.; Chen, J. A Microfluidic Flow Cytometer Enabling Absolute Quantification of Single-Cell Intracellular Proteins. *Lab Chip* **2017**, 17 (18), 3129–3137.
- (10) Salehi-Reyhani, A.; Kaplinsky, J.; Burgin, E.; Novakova, M.; DeMello, A. J.; Templer, R. H.; Parker, P.; Neil, M. a a; Ces, O.; French, P.; et al. A First Step towards Practical Single Cell Proteomics: A Microfluidic Antibody Capture Chip with TIRF Detection. *Lab. Chip.* **2011**, 11 (7), 1256–1261.
- (11) Burgin, E.; Salehi-Reyhani, A.; Barclay, M.; Brown, A.; Kaplinsky, J.; Novakova, M.; Neil, M. A. A.; Ces, O.; Willison, K. R.; Klug, D. R. Absolute Quantification of Protein Copy Number Using a Single-Molecule-Sensitive Microarray. *Analyst* **2014**, 139 (13), 3235–3244.
- (12) Eyer, K.; Stratz, S.; Kuhn, P.; Küster, S. K.; Dittrich, P. S. Implementing Enzyme-Linked Immunosorbent Assays on a Microfluidic Chip to Quantify Intracellular Molecules in Single Cells. *Anal. Chem.* **2013**, 85 (6), 3280–3287.
- (13) Wu, D.; Milutinovic, M. D.; Walt, D. R. Single Molecule Array (Simoa) Assay with Optimal Antibody Pairs for Cytokine Detection in Human Serum Samples. *Analyst* **2015**, 140 (18), 6277–6282.
- (14) Hughes, A. J.; Spelke, D. P.; Xu, Z.; Kang, C.-C.; Schaffer, D. V.; Herr, A. E. Single-Cell Western Blotting. *Nat. Methods* **2014**, 11 (7), 749–755.
- (15) Shi, Q.; Qin, L.; Wei, W.; Geng, F.; Fan, R.; Shin, Y. S.; Guo, D.; Hood, L.; Mischel, P. S.; Heath, J. R. Single-Cell Proteomic Chip for Profiling Intracellular Signaling Pathways in Single Tumor Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109 (2), 419–424.
- (16) Sinkala, E.; Sollier-Christen, E.; Renier, C.; Rosàs-Canyelles, E.; Che, J.; Heirich, K.; Duncombe, T. A.; Vlassakis, J.; Yamauchi, K. A.; Huang, H.; et al. Profiling Protein Expression in Circulating Tumour Cells Using Microfluidic Western Blotting. *Nat. Commun.* **2017**, 8, 14622.
- (17) Stadler, C.; Rexhepaj, E.; Singan, V. R.; Murphy, R. F.; Pepperkok, R.; Uhlén, M.; Simpson, J. C.; Lundberg, E. Immunofluorescence and Fluorescent-Protein Tagging Show High Correlation for Protein Localization in Mammalian Cells. *Nat. Methods* **2013**, 10 (4), 315–323.
- (18) Suzuki, T.; Matsuzaki, T.; Hagiwara, H.; Aoki, T.; Takata, K. Recent Advances in Fluorescent Labeling Techniques for Fluorescence Microscopy. *Acta Histochem. Cytochem.* **2007**, 40 (5), 131–137.
- (19) Day, R. N.; Davidson, M. W. The Fluorescent Protein Palette: Tools for Cellular Imaging. *Chem. Soc. Rev.* **2009**, 38 (10), 2887–2921.
- (20) Verdaasdonk, J. S.; Lawrimore, J.; Bloom, K. Determining Absolute Protein Numbers by Quantitative Fluorescence Microscopy. *Methods Cell Biol.* **2014**, 123, 347–365.
- (21) Scalia, C. R.; Boi, G.; Bolognesi, M. M.; Riva, L.; Manzoni, M.; DeSmedt, L.; Bosisio, F. M.; Ronchi, S.; Leone, B. E.; Cattoretti, G. Antigen Masking During Fixation and Embedding, Dissected. *J. Histochem. Cytochem.* **2016**, 65 (1), 5–20.
- (22) Lo, C.-A.; Kays, I.; Emran, F.; Lin, T.-J.; Cvetkovska, V.; Chen, B. E. Quantification of Protein Levels in Single Living Cells. *Cell Rep.* **2015**, 13 (11), 2634–2644.
- (23) Baker, M. Reproducibility Crisis: Blame It on the Antibodies. *Nature* **2015**, 521 (7552), 274–276.
- (24) Weller, M. G. Quality Issues of Research Antibodies. *Anal. Chem. Insights* **2016**, 11, ACI.S31614.
- (25) O’Kennedy, R.; Fitzgerald, S.; Murphy, C. Don’t Blame It All on Antibodies – The Need for Exhaustive Characterisation, Appropriate Handling, and Addressing the Issues That Affect Specificity. *TrAC Trends Anal. Chem.* **2017**, 89, 53–59.

- (26) Vira, S.; Mekhedov, E.; Humphrey, G.; Blank, P. S. Fluorescent-Labeled Antibodies: Balancing Functionality and Degree of Labeling. *Anal. Biochem.* **2010**, *402* (2), 146–150.
- (27) Zhou, H.; Tourkakis, G.; Shi, D.; Kim, D. M.; Zhang, H.; Du, T.; Eades, W. C.; Berezin, M. Y. Cell-Free Measurements of Brightness of Fluorescently Labeled Antibodies. *Sci. Rep.* **2017**, *7*, 41819.
- (28) Toki, M. I.; Cecchi, F.; Hembrough, T.; Syrigos, K. N.; Rimm, D. L. Proof of the Quantitative Potential of Immunofluorescence by Mass Spectrometry. *Lab. Investig.* **2017**, *97* (3), 329–334.
- (29) Stadler, C.; Skogs, M.; Brismar, H.; Uhlén, M.; Lundberg, E. A Single Fixation Protocol for Proteome-Wide Immunofluorescence Localization Studies. *J. Proteomics* **2010**, *73* (6), 1067–1078.
- (30) Salehi-Reyhani, A. Evaluating Single Molecule Detection Methods for Microarrays with High Dynamic Range for Quantitative Single Cell Analysis. *Sci. Rep.* **2017**, *7* (1), 17957.
- (31) Etheridge, T. J.; Boulineau, R. L.; Herbert, A.; Watson, A. T.; Daigaku, Y.; Tucker, J.; George, S.; Jönsson, P.; Palayret, M.; Lando, D.; et al. Quantification of DNA-Associated Proteins inside Eukaryotic Cells Using Single-Molecule Localization Microscopy. *Nucleic Acids Res.* **2014**, *42* (19), e146–e146.
- (32) Gerdes, M. J.; Sevinsky, C. J.; Sood, A.; Adak, S.; Bello, M. O.; Bordwell, A.; Can, A.; Corwin, A.; Dinn, S.; Filkins, R. J.; et al. Highly Multiplexed Single-Cell Analysis of Formalin-Fixed, Paraffin-Embedded Cancer Tissue. *Proc. Natl. Acad. Sci.* **2013**, *110* (29), 11982–11987.
- (33) Skala, M. C.; Riching, K. M.; Gendron-Fitzpatrick, A.; Eickhoff, J.; Eliceiri, K. W.; White, J. G.; Ramanujam, N. In Vivo Multiphoton Microscopy of NADH and FAD Redox States, Fluorescence Lifetimes, and Cellular Morphology in Precancerous Epithelia. *Proc. Natl. Acad. Sci.* **2007**, *104* (49), 19494–19499.
- (34) Shahrezaei, V.; Swain, P. S. Analytical Distributions for Stochastic Gene Expression. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (45), 17256–17261.
- (35) Iyer-Biswas, S.; Hayot, F.; Jayaprakash, C. Transcriptional Pulsing and Consequent Stochasticity in Gene Expression. **2007**, 12.
- (36) Sherman, M. S.; Cohen, B. A. A Computational Framework for Analyzing Stochasticity in Gene Expression. *PLoS Comput. Biol.* **2014**, *10* (5), e1003596.
- (37) Jain, A.; Liu, R.; Ramani, B.; Arauz, E.; Ishitsuka, Y.; Ragunathan, K.; Park, J.; Chen, J.; Xiang, Y. K.; Ha, T. Probing Cellular Protein Complexes Using Single-Molecule Pull-Down. *Nature* **2011**, *473* (7348), 484–488.
- (38) Salehi-Reyhani, A.; Sharma, S.; Burgin, E.; Barclay, M.; Cass, A.; Neil, M. A. A.; Ces, O.; Willison, K. R.; Klug, D. R. Scaling Advantages and Constraints in Miniaturized Capture Assays for Single Cell Protein Analysis. *Lab Chip* **2013**, *13* (11), 2066–2074.
- (39) Johnson-Buck, A.; Li, J.; Tewari, M.; Walter, N. G. A Guide to Nucleic Acid Detection by Single-Molecule Kinetic Fingerprinting. *Methods* **2019**, *153*, 3–12.
- (40) Nelson-Rees, W.; Daniels, D.; Flandermeyer, R. Cross-Contamination of Cells in Culture. *Science* (80-.). **1981**, *212* (4493), 446–452.
- (41) Almeida, J. L.; Cole, K. D.; Plant, A. L. Standards for Cell Line Authentication and Beyond. *PLoS Biol.* **2016**, *14* (6), e1002476.
- (42) Hughes, P.; Marshall, D.; Reid, Y.; Parkes, H.; Gelber, C. The Costs of Using Unauthenticated, over-Passaged Cell Lines: How Much More Data Do We Need? *Biotechniques* **2007**, *43* (5), 575–586.

For Table of Contents Only

