Biophysical *Journal*
Biophysical Letter

Biophysical Society

Single-Molecule Specific Mislocalization of Red Fluorescent Proteins in Live *Escherichia coli*

Harshad Ghodke,[1,2] Victor E. A. Caldas,[1] Christiaan M. Punter,[1] Antoine M. van Oijen,[2,*] and Andrew Robinson[1,2]
[1]Zernike Institute for Advanced Materials, Rijksuniversiteit Groningen, Groningen, The Netherlands; and [2]School of Chemistry, University of Wollongong, Wollongong, New South Wales, Australia

ABSTRACT Tagging of individual proteins with genetically encoded fluorescent proteins (FPs) has been used extensively to study localization and interactions in live cells. Recent developments in single-molecule localization microscopy have enabled the dynamic visualization of individual tagged proteins inside living cells. However, tagging proteins with FPs is not without problems: formation of insoluble aggregates and inhibition of native functions of the protein are well-known issues. Previously reported artifacts manifest themselves at all expression levels of the FP-tagged proteins, making the design of control experiments relatively straightforward. Here, we describe a previously uncharacterized mislocalization artifact of *Entacmaea quadricolor* red fluorescent protein variants that is detectable at the single-molecule level in live *Escherichia coli* cells.

Biological imaging, in particular fluorescence microscopy, has revolutionized our understanding of many biological processes. In particular, the ability to endogenously tag proteins in cells using genetic fusions of fluorescent proteins (FPs) has provided important insights into subcellular localizations and structures formed by proteins in live cells as part of normal metabolism (1). While genetic fusions of fluorescent proteins represent the most convenient strategy to image biological phenomena while approximating native conditions, they are not without problems. Besides well-known problems of aggregation and modification of function, recent reports provide evidence for other types of artifacts related to the fusion protein, substantiating the observation that fluorescent proteins are not inert reporters. Notably, it has been reported that Clp proteins in *Escherichia coli* formed artifactual clusters when tagged with particular FPs (2). More recently, it has been demonstrated that photoswitchable FPs can influence the localization of DNA-binding proteins in *E. coli* (3).

Red fluorescent proteins (RFPs) constitute an important part of the fluorescent toolkit, not only because they extend the number of available channels for imaging, but also because cellular autofluorescence tends to be lower in these channels. Dimerization of RFPs is a well-known downside, but the latest generation RFPs promise to be monomeric and possess much improved photophysical properties (4,5). To test the suitability of these RFPs for live cell imaging in *E. coli*, we expressed these RFPs in the absence of any tagged protein and uncovered a mislocalization artifact of RFPs that is detectable at the single molecule level.

We expressed the *Entacmaea quadricolor* RFP derivative, mKate2 (5), from an arabinose-inducible pBAD promoter at single-molecule levels in *E. coli* (MG1655) cells (Fig. 1 *A*; see the Supporting Material). Fluorescence (and hence, copy numbers) of mKate2 could be tuned by varying the concentration of L-arabinose in the growth medium (see Fig. S1) (6). Unexpectedly, fluorescence images of individual cells revealed that mKate2 exhibits distinct, dynamic foci when expressed at very low levels. These foci localize at the cell periphery, suggestive of association with the membrane (Fig. 1 *A*; Movie S1). It should be noted that in the emission window corresponding to the red channel, MG1655 does exhibit cellular autofluorescence (see Fig. S2 *A*). However, this cellular autofluorescence does not exhibit the membrane localization observed in MG1655 expressing mKate2 (compare Fig. S2 A with Fig. S1 A). Further, when imaged and analyzed identically, a greater number of localizations are obtained in the mKate2 strain compared to the WT background (39 vs. 8 peaks/cell for the mKate2 MG1655 strain and WT MG1655, respectively) consistent with leaky expression of mKate2 from the pBAD promoter. This localization is consistent with that of the FP-tagged membrane protein LacY-mCardinal (Fig. 1 *B*) (4) and distinct from that of the nucleoid-associated replisome protein DnaQ-YPet (the E subunit of DNA polymerase III tagged with the YFP variant YPet; Fig. 1 *C*). At higher concentrations of L-arabinose (0.1%), the localization of mKate2 resembles that of a cytosolic protein (Fig. 1 *A*).

Submitted February 17, 2016, and accepted for publication May 31, 2016.
*Correspondence: vanoijen@uow.edu.au
Editor: David Piston.
http://dx.doi.org/10.1016/j.bpj.2016.05.047

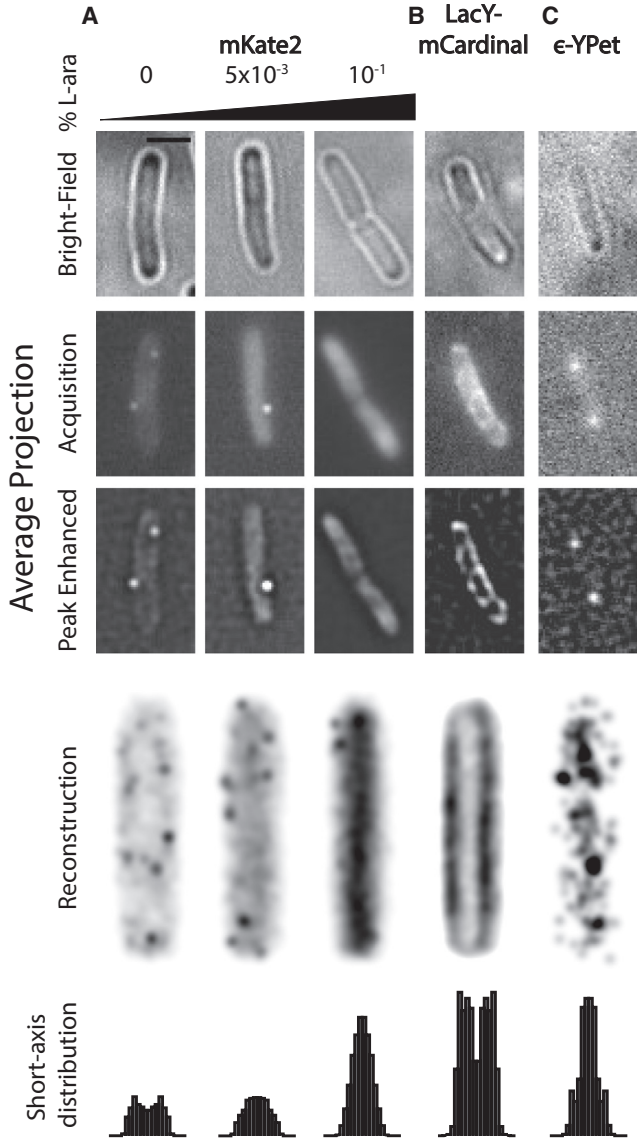


FIGURE 1 Bright-field (*top*), average projections of fluorescence acquisitions (*second row*) and acquisitions filtered to enhance peaks (*third row*), reconstructed images and short-axis projections (*bottom*) of (*A*) mKate2 expressed under a pBAD promoter in the presence of varying amounts of L-arabinose: 0 ($n_{cells} = 142$, $n_{peaks} = 5575$), 5×10^{-3} ($n_{cells} = 137$, $n_{peaks} = 8085$), 10^{-1} ($n_{cells} = 149$, $n_{peaks} = 13{,}283$) percent; and (*B*) mCardinal fusion of LacY expressed under a pBAD promoter on a plasmid in fixed *E. coli* MG1655 ($n_{cells} = 26$, $n_{peaks} = 22{,}954$). (*C*) A chromosomally expressed YPet fusion of dnaQ (ε subunit of polymerase III) in live *E. coli* MG1655 ($n_{cells} = 154$, $n_{peaks} = 277$). Scale bar represents 2.0 μm. Short axis distribution histograms were normalized for number of cells measured.

There are at least two possibilities for why membrane-associated foci may not be visible at higher mKate2 concentrations. The first is that at higher concentrations there is no longer sufficient contrast to distinguish individual membrane-bound mKate2 molecules as foci above the diffuse background of cytosolic molecules. A second possibility is that the distribution of mKate2 molecules between the membrane and cytosol changes as a function of mKate2 concentration, with a greater proportion of molecules bound to the membrane at low concentrations. In our previous work we demonstrated under similar conditions that we could detect a population of monomeric, membrane-bound UmuC-mKate2 molecules against a background of cytosolic mKate2 (7), indicating that our ability to detect foci is not limited by contrast in this concentration regime.

Expressed without fusion partners, YPet did not exhibit clear foci and appeared to be homogenously cytosolic (Fig. S2 *B*), whereas the mKate2 derivative mCardinal exhibited distinct foci at low expression levels (Fig. S2 *C*). The membrane-foci artifact was not detected in MG1655 exhibiting leaky expression of mCherry under the pBAD promoter (Fig. S2 *D*). Together, these observations suggest that artifactual membrane localization may be specific to *E. quadricolor* TagRFP derivatives.

As the artifact is evident when imaging at very low expression levels, analyzing the localization of low-copy-number RFP fusions should still be possible, provided that appropriate control measurements are made. For example, in our recent work on the localization of DNA polymerase V in *E. coli* we observed an interaction between the fusion protein UmuC-mKate2 and the cell inner membrane (7). To demonstrate that membrane interaction is a bona fide property of UmuC, as opposed to an artifact introduced by mKate2, we used an inducible plasmid to increase the concentration of mKate2 to levels where the artifact could no longer be detected and showed that foci of UmuC-mKate2 remained visible on the membrane. As an additional confirmation, electron microscopy revealed the membrane association of wild-type UmuC in fixed cells, using antibody-conjugated gold nanoparticles.

These findings support previous work indicating that fluorescent proteins and their derivatives from different organisms possess characteristics that can differentially influence the behavior of the tagged protein. Control experiments should account for the possibility that some artifacts may manifest at certain expression levels. These findings underscore the importance of verifying observations of fluorescently tagged proteins with orthogonal techniques or multiple tags wherever possible.

SUPPORTING MATERIAL

Supporting Materials and Methods, three figures, and one movie are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(16)30394-0.

AUTHOR CONTRIBUTIONS

H.G. cloned the constructs used in this study; H.G., V.E.A.C., and A.R. performed the microscopy experiments; H.G., V.E.A.C., and C.M.P. wrote the code used to analyze data with inputs from A.R.; and A.M.v.O. and A.R. supervised the work. Article was written with input from all authors.

ACKNOWLEDGMENTS

We thank Dr. Michael Lin for kindly providing the pBAD-mCardinal and pCDNA-mCardinal constructs that were used in this study. We thank Elizabeth Wood and Michael Cox for kindly providing the DnaQ-YPet and UmuC-mKate2 constructs.

REFERENCES

1. Stracy, M., S. Uphoff, …, A. N. Kapanidis. 2014. In vivo single-molecule imaging of bacterial DNA replication, transcription, and repair. *FEBS Lett.* 588:3585–3594.
2. Landgraf, D., B. Okumus, …, J. Paulsson. 2012. Segregation of molecules at cell division reveals native protein localization. *Nat. Methods.* 9:480–482.
3. Wang, S., J. R. Moffitt, …, X. Zhuang. 2014. Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. *Proc. Natl. Acad. Sci. USA.* 111:8452–8457.
4. Chu, J., R. D. Haynes, …, M. Z. Lin. 2014. Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein. *Nat. Methods.* 11:572–578.
5. Shcherbo, D., C. S. Murphy, …, D. M. Chudakov. 2009. Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* 418:567–574.
6. Guzman, L. M., D. Belin, …, J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* 177:4121–4130.
7. Robinson, A., J. P. McDonald, …, A. M. van Oijen. 2015. Regulation of mutagenic DNA polymerase V activation in space and time. *PLoS Genet.* 11:e1005482.

Biophysical Journal, Volume *111*

Supplemental Information

Single-Molecule Specific Mislocalization of Red Fluorescent Proteins in Live *Escherichia coli*

Harshad Ghodke, Victor E.A. Caldas, Christiaan M. Punter, Antoine M. van Oijen, and Andrew Robinson

Supporting Materials and Methods

Cloning

Construction of linker-FP

Fluorescent proteins mKate2 (GeneArt, Life Technologies) (1), YPet(2) (GeneArt, Life Technologies) and mCherry (3) were PCR amplified with primers containing the sequence for the linker (see table) using KOD polymerase (Novagen) using the following cycling conditions: Initial denaturation 95°C for 5 minutes, denaturation 95°C for 15s, annealing at 58°C for 30s and extension at 72°C for 35s (30X), followed by final extension at 72°C for 1 minute. Insert and vector pBAD-myc-HisB (Invitrogen) were both digested with XhoI and XbaI (New England Biolabs) followed by ligation with T4 DNA ligase at 16°C overnight (New England Biolabs) and transformed into E.coli DH5a. Colonies were screened for insertion by PCR, followed by isolation of plasmid and DNA sequencing (GATC biotech).

Primer	Sequence
Linker-mKate2_FW	atc cga gct cga g ATG TCG GCT GGC TCC GCT GCT GGT TCT GGC GAA TTC ATG GTG AGC GAG CTG ATT AAG GAG
Linker-mKate2_Rev	GTT CCT ATT CTC TAG AAA CTA TAG GAA CTT CTC ATC TGT GC
Linker-YPet_FW	atc cga gct cga g ATG TCG GCT GGC TCC GCT GCT GGT TCT GGC GAA TTC ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCC
Linker-YPet_Rev	CGA GGG TAT GAA TGA ATT GTA CAA AGA GCT CTA ATC TAG AAA GCT TCG A
pBAD-mCherry1-F	GGATCCGAGCTCATGGTGAGCAAGGGCGAG
pBAD-mCherry1-R	GCATGTTCTAGATTATTACTTGTACAGCTCGTCCATGC

Construction of LacY-mCardinal

mCardinal insert was amplified from pCDNA-mCardinal using the primers containing KasI and XbaI sites using the PCR program as above with annealing temperature of 70°C and extension time of 20s using KOD polymerase. Both, the insert and pBAD-LacY-eYFP (generous gift from J.T.Mika, University of Groningen) were digested with KasI and XbaI (New England Biolabs) followed by gel purification and ligation with T4 DNA ligase at 16°C overnight. Colonies were screened for the presence of mCardinal (as opposed to eYFP) by digestion of purified plasmid DNA with PstI (New England Biolabs) followed by sequencing.

Primer	Sequence
LacY-mCardinal_F	GGCACTGGAGGCGCC atggtgagcaagggcg
LacY-mCardinal_R	GAGTTTTTGTTCTAGA ttacttgtacagctcgtccatgc

Preparation of cells for imaging

MG1655 cells carrying pBAD plasmids containing the genes for either linker-mKate2, linker-YPet or mCardinal were cultured overnight at 37°C in EZ rich media (Teknova) with glycerol as the carbon source, and the indicated amounts of L-arabinose, and ampicillin. Overnight cultures were reset in fresh EZ media with glycerol with a 1:100 dilution and grown for at least two hours before imaging. Early log phase cells (OD ~ 0.2) were then sandwiched between a silanized glass cover slip and a

clean cover slip, followed by imaging. Each sample was imaged within 10 minutes of preparation. Coverslips were prepared by first sonicating in 5M KOH for 1 h followed by extensive washing with water. Silanization was performed with 2% 3-aminopropyl triethoxy silane in acetone followed by extensive washing with water and drying with nitrogen.

LacY-mCardinal cells were grown at 37°C in EZ rich media with glycerol as carbon source. Cells were fixed by pelleting an early log phase culture followed by resuspension in 100 mM $MgSO_4$ and either 2.8% formaldehyde and 0.04% glutaraldehyde for 30 min or 5.6% formaldehyde and 0.08% glutaraldehyde for 15 min. This was followed by washing twice with 100 mM $MgSO_4$ followed by resuspension in EZ media with glycerol. Resuspended fixed cells were then imaged using the same protocol as for live cells.

Chromosomal fusion of dnaQ-YPet was created by recombination of the YPet-FRT-Kan-FRT cassette into the C-terminus of the dnaQ gene in MG1655 as described previously(4, 5).

Imaging

Samples were imaged on a custom built live-cell imaging microscope (Olympus IX-81) with a 1.49 NA objective and 512x512 $pixel^2$ EM-CCD camera (C9100-13, Hamamatsu) with either the 514 nm or 568 nm laser (Sapphire LP, Coherent). Images were acquired under rapid acquisition mode with the first frame being a bright field image followed by rapid acquisition in the fluorescence channel of a total of 500 frames (mKate2 and mCardinal; Emission filter: 645/75 nm) and 400 frames (Ypet; Emission filter: 540/30 nm) with an integration time of 34 ms. mKate2 was imaged at a laser power of 180 Wcm^{-2}, mCherry and mCardinal were imaged at 18-180 Wcm^{-2}, whereas YPet was imaged at a laser power of 9 Wcm^{-2}.

Image analysis and peak detection

Image analysis was performed in ImageJ (6) and Fiji (7) using custom plugins and macros (available at https://github.com/SingleMolecule, (8)). Images were processed as follows:

1. The fluorescent channel acquisition was first ‘flattened’ to correct the intensity variation arising from the laser beam profile.
2. Flattened images were then Z-projected to obtain the average projection for each acquisition. These averaged projections were used in conjuction with the brightfield images in MicrobeTracker (9) to identify the cell outlines to identify cells that are in focus.
3. Next, peaks were identified in the flattened acquisitions using custom written ImageJ plugins (8). Acquisitions were first averaged to enhance peaks using the peak filter described in ref (10). These peaks were detected using a fluorophore specific signal-to-noise ratio threshold over the cover-slip background, with a minimum distance between adjacent detected peaks of 3 pixels (100 nm/pixel) in each frame. The following SNRs were used: mKate2: 6, LacY-mCardinal:1, Ypet:5, mCherry:5. This method of peak detection is more sensitive to detecting peaks in cells imaged under single molecule imaging conditions, typically with low fluorophore copy numbers, and hence low cytosolic fluorescence, as opposed to imaging cells with high fluorophore copy numbers.
4. Given the shape of the *E.coli* cell, under conditions of high cytosolic fluorescence, the peak detection algorithm falsely detects fluorescence as broad peaks. To eliminate these false-positives, the peak list thus obtained was filtered to include only peaks that possess a full-width at half maximum height of less than 3 pixels or 330 μm. Additionally, foci that lived for longer than 10 consecutive frames were ignored.

5. The peak list was then processed using MicrobeTracker to obtain the relative positions of these peaks within the corresponding cell outline. This peak list consisted of positions (x_i, y_i) for the i^{th} peak such that $x_i \in$ [-D, D] and $y_i \in$ [0, L] where L represents the length of a cell, and 2D corresponds to the width of the cell. In this transformation, (0,0) is assigned to one pole of the cell. For purposes of illustration, the cell length was normalized to 50 pixels. For MG1655, mKate2, mCardinal and LacY-mCardinal data, short-axis distributions (Fig 1 and SI Fig1) were obtained by plotting histograms of x_i values for $y_i \in$ [10, 40] to minimize signal from the poles. For ϵ-Ypet, short-axis distributions were obtained by plotting histograms of x_i for the entire range $y_i \in$ [0, 50]. Short axis distribution histograms were normalized for number of cells measured.

Calculation of number of molecules in a cell as a function of L-arabinose

The average total intensity of a single fluorophore mKate2 molecule was calculated by plotting a histogram of total intensity of a single peak for peaks detected in the 0% arabinose experiment. This histogram was found to fit a sum of two Gaussians with centers at 4.271e+04 (4.102e+04, 4.44e+04) and 7.745e+04 (6.332e+04, 9.159e+04) (R-square: 0.946) (Figure S1A). Of these, the peak with the lower intensity was assigned to monomeric mKate2. The number of fluorophores in each cell was then calculated by dividing the total cellular fluorescence calculated in each cell (corrected for background) by the total intensity of a single fluorophore as obtained from the fit to the histogram. Calculated this way, the average number of molecules for 0, 5 $x10^{-3}$ and 10^{-1} percent L-arabinose are 20 ± 8, 65 ± 39, 158 ± 116 (average ± std. dev) respectively.

Figure legends:

Figure S1

A. Histogram of mean pixel intensity of mKate2 fluorescence in cells as a function of L-arabinose concentration (0, $5x10^{-3}$ and 10^{-1} %) **B.** Histogram of total intensity of mKate2 peaks detected in the 0% ara dataset (bars) and fit to distribution with two gaussian terms and **C.** Histogram of number of molecules per cell in mKate2 expressing cells for varying amounts of arabinose (0, $5x10^{-3}$ and 10^{-1} %).

Figure S2

Bright field, averaged projections of fluorescent acquisitions, reconstructions of the localizations and normalized short-axis distribution of **A**. MG1655 in the red channel (n_{cells} = 34, n_{peaks} = 282). **B.** YPet expressed under an arabinose promoter in the presence of the indicated concentrations of L-arabinose (n_{cells} = 61, 51, 55, 97 and 95; n_{peaks} = 2934, 3540, 5885, 9604 and 9949 for % L-arabinose = 0, 10^{-3}, $5x10^{-3}$, 10^{-2} and 10^{-1} respectively) **C**. mCardinal expressed under a pBAD promoter in the presence of either 0 % L-arabinose (n_{cells} = 51, n_{peaks} = 1369) D. mCherry (% L-arabinose = 0, n_{cells} = 33, n_{peaks} = 635) . Scale bar is 2 μm. Note that intensities of the peaks in the reconstructions are optimized for display purposes and do not reflect the number of cells measured, whereas short-axis distribution histograms are normalized for the number of cells measured with the same scale as in Fig 1A.

Figure S3

Plot of mean background corrected pixel intensity of mKate2 in cells averaged over the population vs. frame (black trace) and histogram of number of peaks detected as a function of frame (grey trace) for L-arabinose **A**. 0, **B**. $5x10^{-3}$ and **C**. 10^{-1} percentage in growth medium.

Movie 1

Video of peak enhanced acquisitions showing dynamic nature of mKate2 foci in cells. Foci appear to diffuse along the cell membrane as indicated by the arrows. Each slice in the movie represents a 34 ms exposure, relayed at 10 fps here.

Supporting References

1. Shcherbo, D., C. S. Murphy, G. V. Ermakova, E. A. Solovieva, T. V. Chepurnykh, A. S. Shcheglov, V. V. Verkhusha, V. Z. Pletnev, K. L. Hazelwood, P. M. Roche, S. Lukyanov, A. G. Zaraisky, M. W. Davidson, and D. M. Chudakov. 2009. Far-red fluorescent tags for protein imaging in living tissues. The Biochemical journal 418:567-574.
2. Nguyen, A. W., and P. S. Daugherty. 2005. Evolutionary optimization of fluorescent proteins for intracellular FRET. Nat Biotechnol 23:355-360.
3. Doherty, G. P., M. J. Fogg, A. J. Wilkinson, and P. J. Lewis. 2010. Small subunits of RNA polymerase: localization, levels and implications for core enzyme composition. Microbiology 156:3532-3543.
4. Huang, L. C., E. A. Wood, and M. M. Cox. 1997. Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the FLP recombinase: the FLIRT system. J Bacteriol 179:6076-6083.
5. Reyes-Lamothe, R., D. J. Sherratt, and M. C. Leake. 2010. Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science 328:498-501.
6. Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671-675.
7. Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9:676-682.
8. Caldas, V. E., C. M. Punter, H. Ghodke, A. Robinson, and A. M. van Oijen. 2015. iSBatch: a batch-processing platform for data analysis and exploration of live-cell single-molecule microscopy images and other hierarchical datasets. Molecular bioSystems.
9. Sliusarenko, O., J. Heinritz, T. Emonet, and C. Jacobs-Wagner. 2011. High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. Mol Microbiol 80:612-627.
10. Hedde, P. N., J. Fuchs, F. Oswald, J. Wiedenmann, and G. U. Nienhaus. 2009. Online image analysis software for photoactivation localization microscopy. Nat Methods 6:689-690.

Figure S1

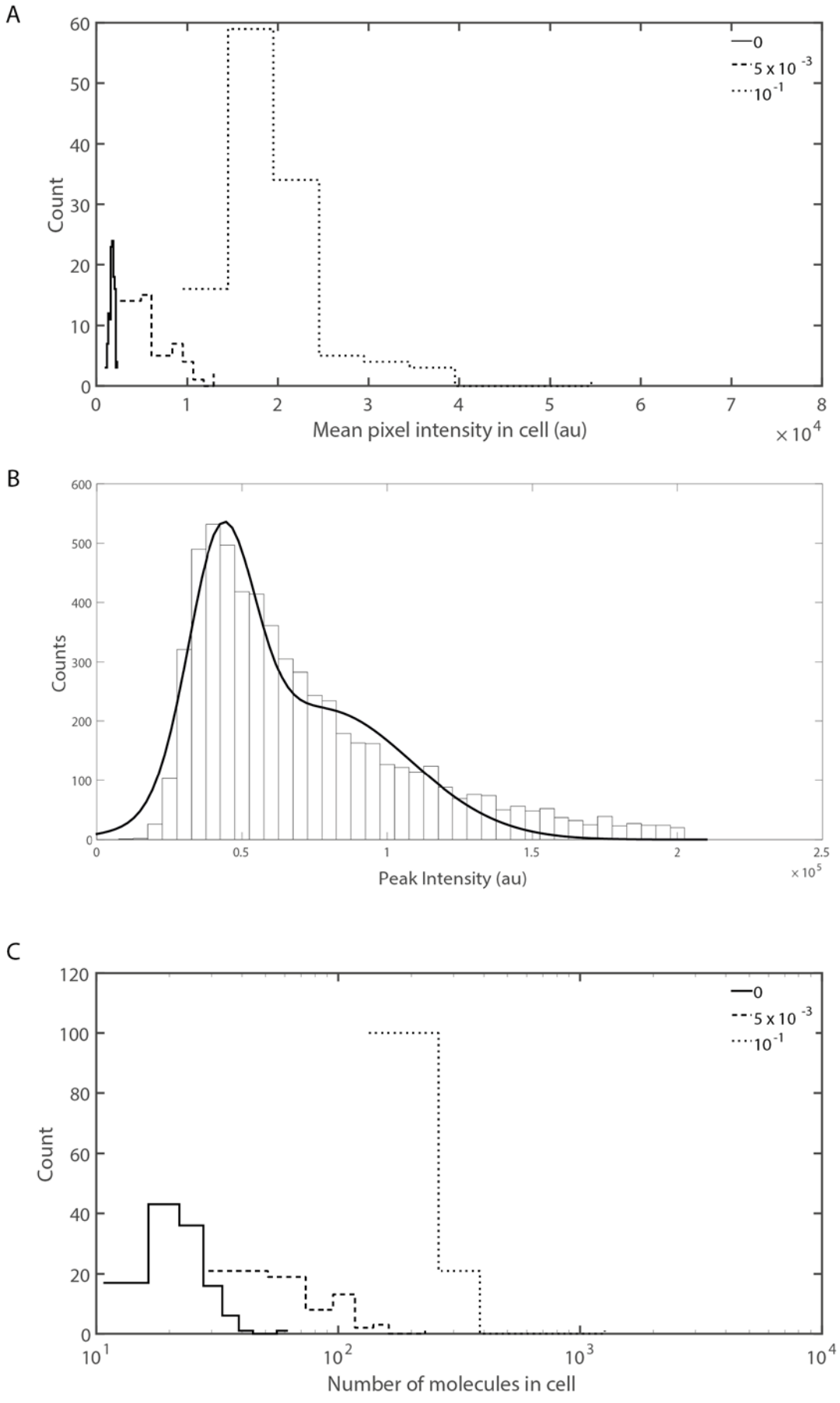

A
Count
Mean pixel intensity in cell (au)
×10^4
0
5 x 10^-3
10^-1
B
Counts
Peak Intensity (au)
×10^5
C
Count
Number of molecules in cell
0
5 x 10^-3
10^-1

Figure S2

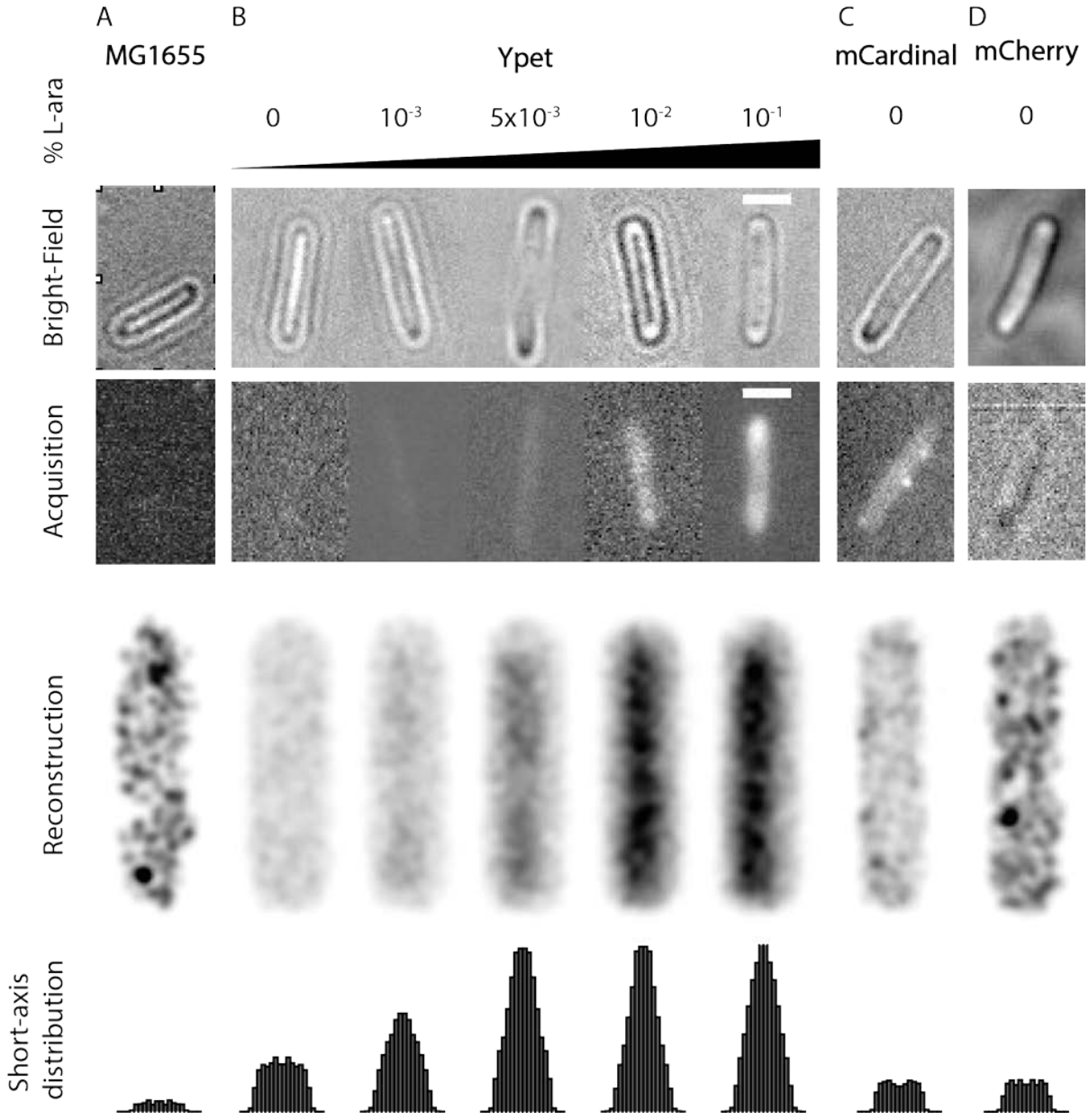
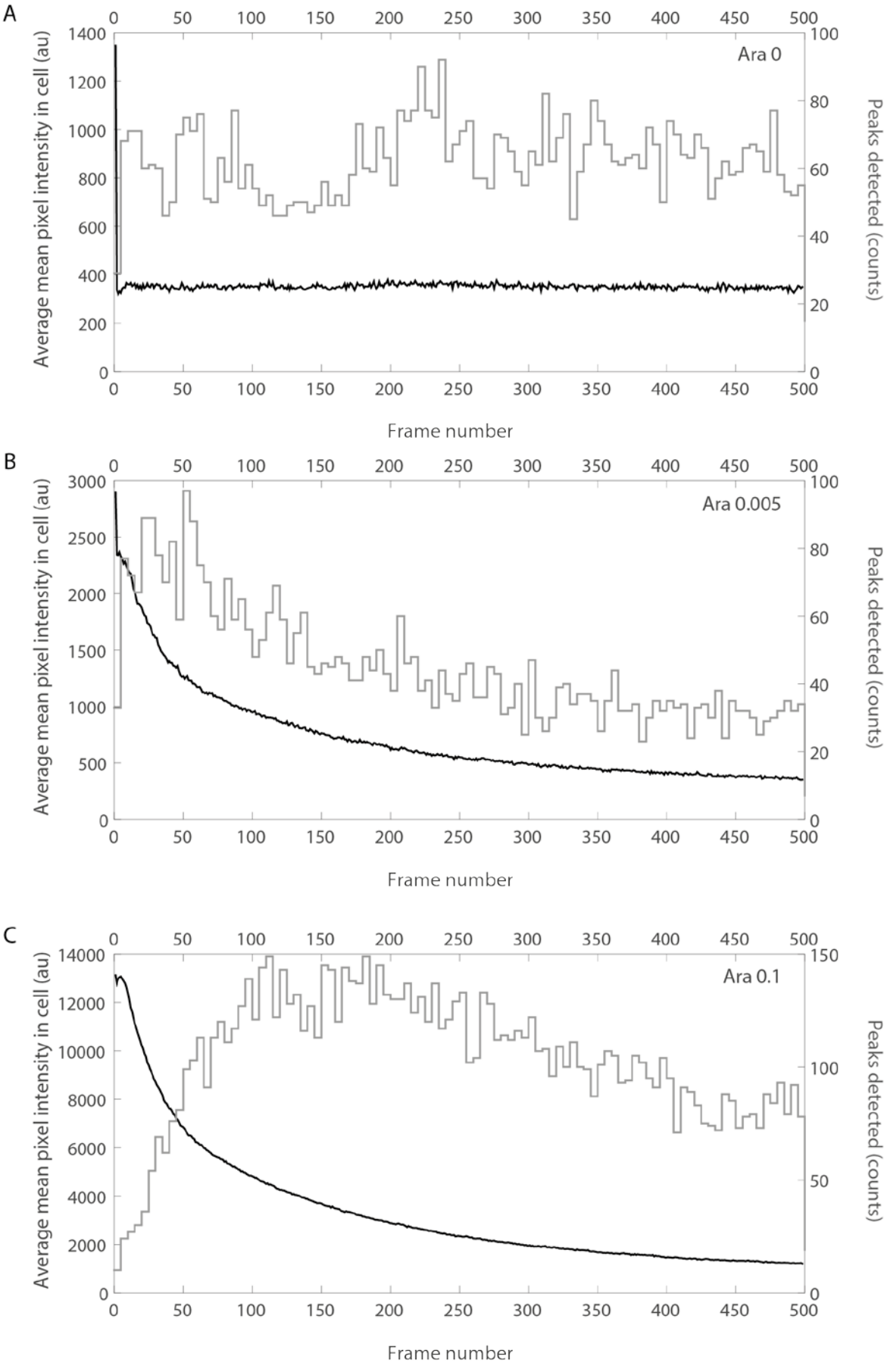

A
MG1655
B
Ypet
C
mCardinal
D
mCherry
% L-ara
0
10-3
5x10-3
10-2
10-1
0
0
Bright-Field
Acquisition
Reconstruction
Short-axis distribution

Figure S3


A
Ara 0
Average mean pixel intensity in cell (au)
Peaks detected (counts)
Frame number
B
Ara 0.005
Average mean pixel intensity in cell (au)
Peaks detected (counts)
Frame number
C
Ara 0.1
Average mean pixel intensity in cell (au)
Peaks detected (counts)
Frame number