In vivo single molecule fluorescent *in situ* hybridisation – tracking single RNAs in the live *Escherichia coli* cell

Emma Isaura Lalande

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford



Lincoln College

University of Oxford

Hilary Term 2025

‘Learn how to see. Realize that everything connects to everything else.’

~ Leonardo da Vinci ~

Disclaimer

I hereby declare that the work in this thesis is that of the candidate alone, except where indicated in the text, and as described here:

- The original 8mer-2x647N probe aliquot used in Chapter 3 was synthesised and purified by Afaf El-Sagheer and Tom Brown, originally for use in Gap-Seq by Rebecca Andrews.
- The HPLC second-pass purification of commercially synthesised 8mer-2x647N (Biomers.net), used in Chapter 3, was collaboratively performed with Rana Abdul-Razak. Subsequent SDS-PAGE for probe comparison was performed with Jagadish Prasad Hazra.
- The BW27784 and MG1655 *E. coli* p6 and gT7 tag system strains were originally generated by Hafez El Sayyed.
- The TetR-YPet plasmid used in p6 co-localisation studies was constructed and donated by Alfredas Bukys.
- The Chapter 5 probes were generously shared Mirjam Kümmerlin (EUB338-1x647N, EUB338-ATTO643/BHQ1), Alison Farrar (EUB338-1xCy3B) and Nicolas Shiaelis (8-40 nt x647N T-repeats).
- All ensemble mean intensity measurements were generated through a custom Python code generously co-developed with Rasched Haidari.
- In Chapters 4 & 5, the spatial distribution analysis by localisation density subject to cell length was performed using Python code by Jacob Wright.
- In Chapter 5, the apparent diffusion analyses were performed using MATLAB code written by Jacob Wright.
- Across all Chapters, the open source Napari *BacSeg* and *Moltrack* platforms were generated by Piers Turner.
- Biorender.com was used to produce all schematics in this thesis.

Emma Lalande

April 2025

Abstract

Transcription is fundamental to the determination of gene expression and consequent live cell phenotype. In the *E. coli* bacterium – where subcellular structuring and complex gene expression processes are increasingly elucidated – RNAs are transcribed via RNA polymerase (RNAPs), translated via ribosomes, and eventually degraded via the multiprotein RNA degradosome. The careful balance of RNAP, ribosome and degradosome stoichiometry and spatial distributions thereby mediates the copy number, lifetimes and localization of RNAs in the cell – alongside the diverse roles of the RNAs themselves.

Visualizing real-time single mRNA molecules *in vivo* enables deeper exploration of single gene expression rates, transcription dynamics, mRNA spatiotemporal distributions and mobility behaviours. The smaller size, higher structural integrity and lower mRNA copy numbers expressed in bacterial cells, however, present a consistent challenge in the live cell tagging of mRNAs. Whilst fluorescent protein and fluorogenic aptamer systems have been seminal in prokaryotic single molecule mRNA localisation and tracking, these methods hold several caveats: the bulky engineered labels can perturb mRNA physiological processing, often with limited fluorophore brightness and photostability. To visualize target RNAs *in vivo* with reduced impact on mobility and processing, this thesis develops a novel permutation of single molecule fluorescent *in situ* hybridisation (FISH) in live *E. coli*. Using fluorogenic, single-stranded DNA probes complementary to artificial and native target RNAs, the spatial localization, lifetimes and mobility behaviours of single RNAs were observed via superresolution microscopy down to a possible single fluorophore resolution.

Corroborating reported observations in live cell RNA mobilities, RNAs were shown to rapidly diffuse through the whole cell, and exhibit static confinement when proximal to a transcription focus or a ribosomal translation hotspot. Whilst there are improvements to be made around probe specificity, the endogenous labelling of target RNAs has been proven via both electroporation and passive internalisation of short FISH-like probes. This thesis presents a promising avenue for the endogenous labelling of any single RNA in the live *E. coli* cell, applicable across numerous fields.

Acknowledgments

One of my favourite parts of a thesis is the acknowledgements section. The gratitude, joy and support that always shine through are to me, synonymous with the collaboration, creativity and inspiration so innate to Science and Discovery. I could write a whole volume to all the people, places and learnings that have inspired, taught and led me to where I am now; I shall save those stories for another time.

I would first like to thank my Supervisor, Professor Achillefs Kapanidis. From complex idea discussions to navigating experimental pipelines, I have been able to develop all sorts of skills necessary for an independent scientist and thinker with your help. I have particularly appreciated your advice and support in scientific writing – communication being close to my heart. I have learnt much from your wisdom; thank you.

A deep thank you to everyone in the Kapanidis group who were friendly faces, supporting shoulders and motivational speakers throughout: Stelios Chatzimichail, Suchintak Dash, Hafez El Sayyed, Alison Farrar, Rasched Haidari, Jagadish Prasad Hazra, Christof Hepp, Mirjam Kümmerlin, Andrew MacMahon, Abhishek Mazumder, Oliver Pambos, Sanna Piippo-Henderson, Nic Shiaelis, Sammi Ta, Piers Turner, Heesoo Uhm, Anna Wang, Ed Wheeler, Jacob Wright, Aleksander Zagajewski & Qing Zhao. Many contributed to the progress of this research project, and will continue to; thank you particularly to Miri for fluorescence ideas, Ed for his MRes work and Rasched, Jacob and Piers for analysis. The Kavli has been a wonderful place and community – thank you so much to Rishi, Désirée, Dominika, Martin and Sebastian, and the rest of the smiley faces across this amazing building. Your cheerful and generous presence has been consistently uplifting.

Thank you to the BBSRC for funding my DPhil.

I wanted to add some special lab ‘thank yous’: Hafez, thank you for your constant belief, skills training and storytelling; I always appreciated the colour you brought.
Nic – thank you for telling me to ‘just do it’; you were, and still are, right!
Ali & Miri – your strength, wisdom and laughter were sunshine, on happy days and gloomier days alike. I am so lucky to have such wonderful Women in STEM around me; keep shining.
Stelios & Piers – your down-to-earth, perfectly-timed and thoughtful advice I will carry with me always! Your perspectives motivated and grounded me – thank you.
Jagadish – thank you for taking me in all ways, teaching me new experiments so enjoyably, and bringing such bright generosity and pragmatism every day; keep dreaming.

My acknowledgements would be incomplete without a short ode to Oxford. I have had the privilege of living within your walls longer than many; your dreaming spires and time-worn stone see everything, in eloquent silence. My journey started in Lincoln –

amidst all the ardour, tumult and follies of undergrad, I have felt at home ever since. Thank you to Carmie and Joe for infallibly being there for me; David & Nigel for my intellectual growth and subsequent path to this DPhil. LMH – you have since captured my heart, and become my second home here; thank you especially to Ed, who has been the best co-worker I could ask for.

So, a huge thank you to Lincoln and LMH, two stunning colleges that have shared so much with me over the years, given me a lifetime of friends and family. A shoutout to Jungle for all the vibes: you continue to keep me moving, inspired and energized; here's to more dancing through life (and the lab, libraries and city streets where possible). The years pass ever so quickly in this city where time seems to stand still. I am still, and will always be, inspired by you.

To my Tribe – we may be flung across all four corners of this wild and wonderful world, but I hold you in my heart every day. Your craziness, thoughtful musings and storytelling have enabled me to insatiably dream, push further, think more and keep loving. Never change, always grow, and I cannot wait for when we can share the fleeting moments of life next. You know who you are, beautiful people; I could not have done this without you.

To my Family – you have stoically supported me from the background as long as I remember. Thank you for telling me what I needed to hear when you knew I needed but didn't want it, forever telling me I can do it, wryly appreciating my stubbornness otherwise, and celebrating the small wins. To Papa, thank you for supporting me whenever you could, and making sure I was watching the rugby!

Mum – none of this would have been possible without your unconditional support. Your confidence, (sometimes stern!) advice and delicious food have made this journey all the more doable – and I'm sure this will all continue. Thank you for all you do, always.

Nana – this Thesis is for you. You were always so determined, so absolutely certain of me; I will keep pushing. We miss you, and are all continuously inspired by you, as ever.

Luke – you are the most beautiful soul; thank you for always believing in me.
You have been one of the greatest joys of this journey, and you teach and show me so much every day. You have been the endless sky to my deep ocean; here's to many more adventures, discoveries and shared successes.

Thank you so much to everyone who has been a part of this for me; people are everything, Everything is connected. This is the ultimate gift.

Talks

- (Upcoming) *In vivo* single molecule FISH for RNA detection in live *Escherichia coli* cells; European Biophysical Societies' Association, Rome, June 2025

- *In vivo* single molecule FISH for RNA detection in live *Escherichia coli* cells; Early Career Microbiologists Symposium, John Innes Centre, September 2024

- *In vivo* single molecule FISH for RNA detection in live *Escherichia coli* cells; 35th RNA Polymerase Workshop, Oxford, March 2024

- *In vivo* single molecule FISH for RNA detection in live *Escherichia coli* cells; Interdisciplinary Biosciences DTP Annual Symposium, Oxford, June 2024

- *In vivo* single molecule FISH for RNA detection in live *Escherichia coli* cells; 19th Nucleic Acids Forum, London, July 2023

Contents

Chapter 1 – Introduction 1

1.1 The *E. coli* bacterium: the King of prokaryotic model organisms 2

1.2 Prokaryotic transcription & translation: the RNA polymerase and ribosome 2

1.3 An organised system – compartmentalization of the *E. coli* cell 4

1.4 'The importance of being earnest' – understanding more about RNAs in the live cell.... 6

1.4.1 Methodological insights from the whole cell 7

1.4.2 Methodological insights from fixed cells: fluorescent *in situ* hybridization (FISH) 7

1.4.3 Methodological insights into RNA localization in live cells 10

1.4.3.1 Engineered labelling 10

1.4.3.2 Endogenous labelling 14

1.5 I spy with my little eye: current perspectives on RNA spatiotemporal dynamics in the *E. coli* cell 16

1.6 Overview 19

Chapter 2 – Materials & Methods 20

2.1 Chapters 3 & 4 – targeting mRNAs via engineered *in vivo* smFISH 20

2.1.1 Preparation of reagents and strains 20

2.1.2 *E. coli* electrocompetency procedure 25

2.1.3 *In vivo* smFISH experimental pipeline 25

2.1.4 Cell viability assessment 27

2.1.5 Data analysis 28

2.2 Chapter 5 – Passive internalisation of fluorescently labelled ssDNAs 31

Chapter 3 – Establishing engineered *in vivo*, single molecule FISH via short, transient fluorogenic probes 35

3.1 A whole new world: super-resolution 35

3.1.1 Single molecule fluorescence microscopy; a brief overview of applied methods. 35

3.1.2 Fluorophores light the way 37

3.2 Engineered *in vivo* smFISH: an origin story 43

3.2.1 *In vivo* single molecule FISH via short fluorogenic oligos – the premise 43

3.2.2 Continuing methodology development 46

3.2.3 Preliminary findings 47

3.3 Chapter summary 47

3.4 Recapitulating preliminary findings: *in vivo* smFISH 48

3.4.1 Engineering systems to evaluate probe-RNA hybridisation and probe transcription 48
3.4.2 Investigating mRNA tagging in plasmid and chromosomal systems 52
3.4.3 Controlling for specificity – wildtype comparisons 58
3.4.4 Controlling for transcription specificity – blocking active transcription with rifampicin 61
3.5 Technical methodological optimisation 65
3.5.1 BW27784 versus MG1655 65
3.5.2 Colocalising p6 plasmids 70
3.5.3 A question of probe specificity – comparing probes for specific signal 72
3.5.4 Technical Optimisation 81
3.5.5 Cell viability 87
3.6 Passive internalisation of labelled biomolecules 89
3.6.1 Technical exploration of fluorescently-labelled ssDNA passive internalisation 90
3.7 Discursive summary 98
Supplement 3 100
Chapter 4 – Implications on transcription initiation: applying engineered *in vivo* smFISH 103
4.1 Introduction 103
4.2 Molecular design – tagging the native *rpoB* promoter 105
4.3 Comparing 8mer-2x647N hybridisation to the *rpoB* transcript 108
4.4 Applying the 8mer-ATTO643/BHQ1 probe for engineered *in vivo* smFISH in tag strains 111
4.4.1 Populational cell intensity comparisons 111
4.4.2 Single-cell photo-bleaching – 8mer-643/q1 photostability and brightness across tag strains 114
4.4.3 Comparing apparent diffusive populations across strains 118
4.4.4 Exploring populations of tracked molecules: counts per cell, durations & spatial distributions 122
4.5 Behavioural characteristics of single RNA molecules bound by 8mer-643/q1 127
4.6 Discussion 142
4.6.1 *rpoB* transcription – a complex machine 142

4.6.2 On the fundamental importance of growth conditions 143
4.6.3 Discursive summary 145
Supplement 4 149
Chapter 5 – Prospects of permeability: applying *in vivo* smFISH via passive internalisation 158
5.1 Introduction 158
5.2 Experimental approach 163
5.3 Internalisation efficiency: comparing existing methods 164
5.4 EUB338-ATTO643/BHQ1 as a suitable probe for targeting individual ribosomes 170
5.4.1 Comparing the photostability of EUB-647N and EUB-643/q1 probes 170
5.4.2 Comparing the unitary intensities of EUB-647N and EUB-643/q1 probes 172
5.5 Exploring single ribosomes *in vivo* 174
5.5.1 Single-particle tracking of EUB-643/q1-bound 16S rRNA 174
5.5.2 Ribosomal populations show subdiffusive properties 174
5.5.3 Populational distributions of tracked molecules 177
5.5.5 Observing behaviours of individual ribosomal particles 182
5.6 Discursive summary 188
Supplement 5 192
Chapter 6 – Exploring single target RNAs in the live cell via short fluorogenic ssDNA probes 199
6.1 It's the little things in life 199
6.2 All the questions, and not so many answers 199
6.3 Future progress – endless parameters, boundless opportunities? 201
6.3.1 Multi-pronged approaches 201
6.3.2 The bigger picture 204
Appendix 1 – Optimisation experiments for *in vivo* smFISH: figures 207
1.1 Electroporation parameters 207
1.2 Microscopy parameters 212
1.3 Passive internalisation of labelled ssDNAs: figures for technical experiments 216
Appendix 2 – Molecule distributions, behaviours and visuals in MG1655 *E. coli* with 50 nM 8mer-2x647N 220
2.1 Evaluating the photostability and brightness of 8mer-2x647N across p6, gT7 and *rpoB*-tag strains 220
2.2 Comparing apparent diffusive populations across strains 222

2.3 Populational tracked molecules – counts per cell, durations & spatial distributions 225
2.4 Behavioural characteristics of single RNA molecules bound by 8mer-2x647N..... 229
Bibliography 235

Chapter 1 – Introduction

The goal of this thesis is to make headway in exploring the complex fundamental process of bacterial transcription, at a single molecule level in real time.

Bacteria, as one of the primary life forms on Earth, are ubiquitous, vastly biodiverse, and essential to life as known by humans. The *Escherichia coli* (*E. coli*) bacterium, a favourite model organism across the life sciences, is the most well-established prokaryotic model organism, and has formed the basis of much of what is known today about conserved prokaryotic processes. Transcription, concerned with ribo-nucleic acid (RNA) production and regulation as part of the molecular central dogma, is fundamental to defining the life cycle, phenotype and purpose of any particular cell (Brenner et al., 1961). The prokaryotic RNA polymerase (RNAP) is highly elucidated structurally and mechanically, as are its populations and perturbations in the live cell; however, at the level of live cell single molecule transcription and subsequent RNA transcripts, there remains more to clarify.

Current tools in probing RNA spatial distributions primarily rely on ensemble measurements – which often have a sensitivity of several 100 RNA copies, significantly surpassing the common 10 or fewer copies at any one time for most *E.coli* genes. Conversely, of the established single molecule tools to explore specific mRNA roles in cells, most tools require fixed and permeabilised cells, or even extracted cell lysates, losing the value of representing temporal sensitivity, spatial landscapes and cooperative/coordinated intracellular processes. As such, the problems addressed in further developing an *in vivo*, single molecule means of visualising target RNAs are the finite and temporary concentrations of physiologically relevant bacterial RNAs, regulated and performing specific (and potentially unknown) roles intracellularly.

The employment of fluorescence techniques to achieve single molecule resolution has enabled single-transcript tagging in live *E. coli*. The implications of specifically visualising RNAs, spatially- and temporally-defined, in live *E. coli* are substantial: exploring previously technically unfeasible target genes' impacts on the single cell; investigating non-coding RNAs for phenotypic regulation; redefining co-regulation of gene expression processes; and untangling the extent of process/pathway coordination as just a few. Translationally down the line, the bigger picture of improving understanding of all these intercalated processes could be influential in antimicrobial resistance spaces, opening up new avenues for bacterial infection control via pharmaceutical interventions, or contributing to synthetic biology developments in climate change and bioremediation. The excitement is to come.

1.1 The *E. coli* bacterium: the King of prokaryotic model organisms

Escherichia coli as a bacterium is interesting from myriad research perspectives: from structural, mechanistic and adaptive angles, this bacillus holds decades of contributions towards medical, pharmacological, biotechnological and synthetic innovation (Blount, 2015). A Gram-negative, facultative anaerobe with a versatile and resilient metabolic capacity, these coliforms optimally have a doubling time of ~20 minutes at 37°C (Son & Taylor, 2021) – making these bacteria a prime case study for environmental selection pressure, evolutionary adaptation, and holding clinical relevance for humans both as gastrointestinal commensals and pathogens (uropathogenic *E. coli*, extra-intestinal pathogenic *E. coli*) (Escherich, 1988). Exploring the interplay of all these different factors in the survival, efficiency and infectivity of *E. coli*, via laboratory and clinical strains, continues the development of prokaryotic, particularly Gram-negative, bacterial function and life cycle, and can open many translational avenues.

Considering molecular biology and genetics, the *E. coli* bacterium has been a favourite model organism for over 80 years, holding prokaryotic pride of place in contributing to establishing the central dogma of molecular biology via genetic flow and expression (Crick, 1958), genetic recombination (Lederberg and Tatum, 1946) and transposition, operonic structures (Jacob & Monod, 1961) and latterly, the development of core genetic engineering techniques and applications of synthetic biology (Blount, 2015). The laboratory K-12 strain of *E. coli* is the godfather of *E.coli* strains used in fundamental research: fully genome-sequenced in 1997 (Blattner et al., 1997), the K-12 MG-1655 *E. coli* bacterium has a genomic circular chromosome of ~4.6 Mb, comprising 4,703 genes of which 4,328 are protein-encoding in 3742 transcriptional units, and leaving ~8% of the genome as non-coding, according to EcoCyc[1] (Keseler et al., 2021).

1.2 Prokaryotic transcription & translation: the RNA polymerase and ribosome

[1] EcoCyc – a globally curated database for the *E. coli* K-12 genome, including information on individual gene function, (co-)regulation and related pathways.
Annotations may be submitted throughout the research community, and alongside RegulonDB, these two databases are largely regarded as the most up-to-date resources for K-12 *E. coli* genomic information.

The bacterial RNA polymerase

In light of *E. coli* metabolic versatility and corresponding precise, nuanced gene expression regulation, the exact genomic and phenotypic qualities of a single *E. coli* cell are fundamentally moderated by the *E.coli* RNAP.

The workhorse of prokaryotic transcription, the *E. coli* RNAP has been the subject of decades of protein studies, *in vitro* characterisation, and synthetic biology ventures since its discovery, purification and subunit structure/function elucidation in the 1960-70s (Burgess, 1969; Chamberlin, 1974; Furth et al., 1962; Hurwitz et al., 1960). A hetero-pentamer in its core enzyme form, RNAP comprises α1, α2, β, β' and ω subunits, associating with a σ-factor to form the ~450 kDa RNAP holoenzyme (Burgess et al., 1969) and initiate active transcription in binding to DNA promoter regions. σ-factor binding is essential for specific DNA recognition at the -35 and -10 promoter elements, generating a closed promoter complex. Transcription may then proceed: the β/β' "crab-claw" clamp, holding the catalytic active site and a positive charge for DNA-binding, enables the entry of single-stranded DNA in the DNA-RNA open complex in its closed, stably binding conformation. A secondary channel in the β/β' clamp allows access to ribonucleotides for RNA transcript elongation, and an RNA exit channel enables nascent RNA release.

There is high evolutionary conservation of RNAP structure and function across the domains – the clamp structure, the catalytic identity of the active site, core invariant motifs are strikingly similar in *E. coli* RNAP and the eukaryotic RNAP II (Ebright, 2000). Considering this conservation, the fundamental essentiality of RNAP is further emphasised; as, therefore, is the importance of precise and nuanced transcriptional regulation.

There are multitudinous ways of regulating the RNAP holoenzyme in specific gene transcription. The major σ factor, σ^{70}, dominates transcriptional activity in the live cell with ~2000 specific promoters out of ~4500 genes available (Shimada et al., 2014); but there are, in fact, six other σ factors for RNAP holoenzyme formation. These six σ factors are all specific stress-related components (i.e. nutrient starvation, envelope damage, temperature shock), redirecting RNAP transcription and cellular energy to essential genes required for stress evasion and survival (Ishihama, 2017). There are additionally ~300 transcription factors for the variable enhancement and repression of effective RNAP binding and elongation, alongside chemical actuators and attenuators which can synergistically coordinate transcription under stress (Ishihama, 2017).

The ribosome

Ribosomes are the actuators of translation – protein production – in living cells. Also highly conserved across all kingdoms, the ~2.5 MDa prokaryotic 70S ribosome is characterized by two subunits, the 50S large and 30S small subunits, comprising a whole host of rRNAs and small ribosomal proteins (Benjamin et al., 1998).

The components of the ribosome in *E. coli* are generated by 7 *rrn* operons across the genome, making up >50% total cellular RNAs during rapid growth (Condon et al., 1995). Indeed, ~40% cell energy is dedicated to ribosome biogenesis and translation (Kaczanowska & Rydén-Aulin, 2007). Acknowledging such intensive applications of finite energy sources, the regulation of complex maturation, assembly and functionality of the ribosome army is intrinsic to efficient gene expression in response to cell needs and environmental reactivity. Control of ribosome biogenesis is exerted at the level of rRNA transcription, rRNA nucleolytic processing in maturation, posttranslational modification of small ribosomal proteins, pre- and during assembly, as well as coordinated support from protein RNA chaperones and maturation enzymes (Kaczanowska & Rydén-Aulin, 2007). Post-assembly, translation is coordinated by essential translation initiation factors, tRNA kinetics during peptide chain elongation, and specific "SOS" system mediation when under stress (Wilson & and Nierhaus, 2007). Bidirectional regulation of *rrn* genes by repression of circulating ribosomal proteins stoichiometrically adds to control of ribosome biogenesis (Kaczanowska & Rydén-Aulin, 2007).

1.3 An organised system – compartmentalization of the *E. coli* cell

The *E. coli* cell is a densely packed, crowded unit with ~30% of the cell populated by macromolecules (Zimmerman & Trach, 1991) – yet simultaneously shows "remarkable spatial organization" (Bakshi et al., 2012). Such macromolecular crowding alongside entropic compacting forces exert distinct molecular spatial distributions (Woldringh et al., 1995; Zimmerman, 2006).

The *E. coli* nucleoid, ~1.5 mm in contour length but constrained to <1 μm in length at the centre of the cell, comprises the bacterial chromosome DNA and a large number of DNA-binding proteins (DBPs) (Gerber et al., 2023; Jun, 2015). The central compaction of the nucleoid is likely driven by a variety of interactions, including macromolecular exclusion at the cell periphery, positive entropy leading to higher folding with additional supercoiling by topoisomerases, and strand linking by DBPs. These forceful compactions excludes larger

macromolecules from the nucleoid (Mondal et al., 2011), allowing for free diffusion of smaller entities, such as unengaged RNAPs and free ribosomal subunits.

By this reasoning, several seminal live cell studies have determined that of the ~55000 ribosomes (Bakshi et al., 2012), only ~20% ribosomes are unassembled subunits (Bremer et al., 2003), free for nucleoid perfusion (Bakshi et al., 2012). Of the ~5000 RNAPs, ~80% are found in the nucleoid (Bakshi et al., 2012; Klumpp & Hwa, 2008; Stracy et al., 2015). Considering that ~85% of RNAP interactions with DNA are "searching" interactions (Stracy et al., 2015), often taking much time to successfully collide and activate transcription (Elf et al., 2007), RNAPs residing predominantly within the nucleoid increases transcription success amid a crowded environment – and speaks to the role of RNAPs in co-regulating nucleoidal structure (Bakshi et al., 2015; Gerber et al., 2023).

The picture of the cell thus appears: a central dense nucleoid, perfused with small proteins and RNAPs, and ribosomes and bulkier cellular components excluded and segregated to the periphery and poles of the cell. Notwithstanding the growth rate dependency of intracellular crowding – leading to nucleoidal remodelling for higher expression accessibility of active genes (Yang et al., 2019) – several theories for the connection of transcription and translation have been proposed (Figure 1.1).

Transcription-translation coupling (Miller et al., 1970) proposes the clustering of active genes in the periphery and closer to the cell poles, for efficient trafficking of mRNAs to ribosomal hotspots. This coupling potentially enables simultaneous polysomal translation of a nascently elongating transcript; albeit, unlikely the predominant mechanism for cellular protein manufacture (Gerber et al., 2023; Stracy et al., 2015; Yang et al., 2019). The transertion hypothesis proposes the physical linking of co-transcriptional-translation to both the nucleoidal DNA and membrane-linked polypeptide chains from actively translating ribosomes (Woldringh, 2002). Theorising that physical linkage of dogmatic components would act as an "expanding force" on the nucleoid, there are reports of nucleoid expansion on the inhibition of transcription (Bakshi et al., 2014) and translation (van Helvoort J M et al., 1996). Finally, renewed interest in liquid-liquid phase condensates has suggested viscosity-mediated compartmentalization of macromolecules in the cell (Gao et al., 2021).

It so is likely that all these subcellular structures and mechanisms occur somewhat concurrently, in a growth-dependent and environmentally-sensitive manner. Most highlighted here, is the precise, nuanced and exquisite control that proceeds in the life and regulation of the *E. coli* genotype and phenotype.

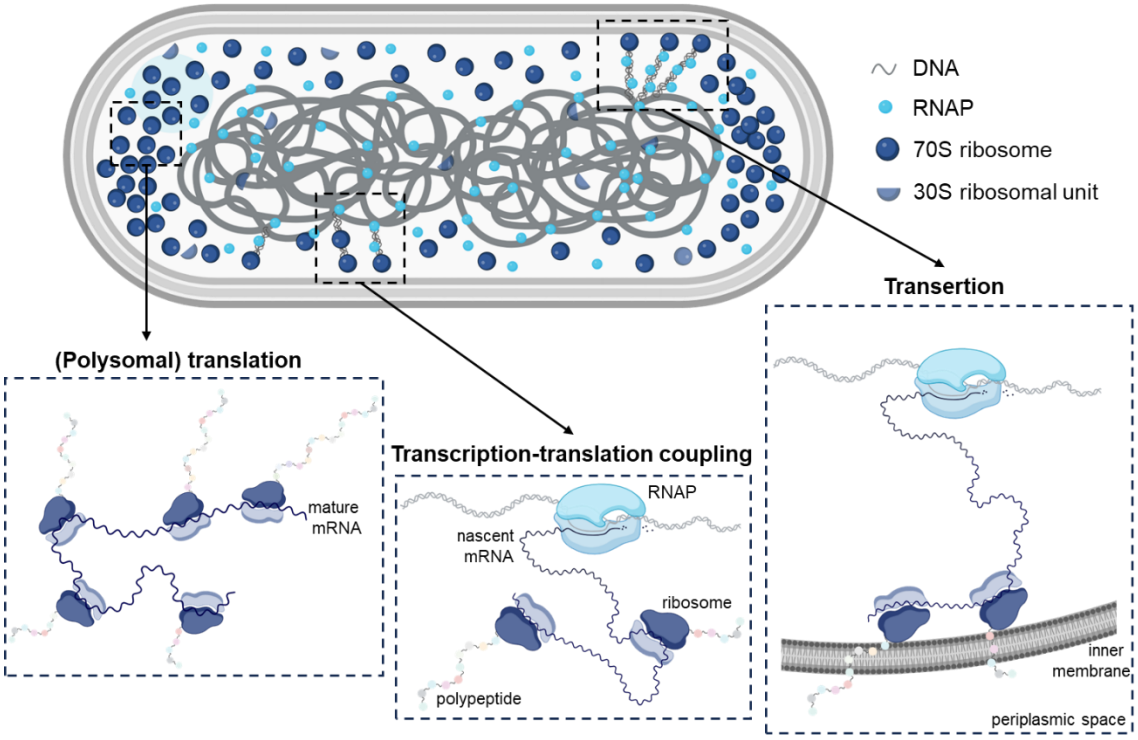


Figure 1.1: Schematic of the macromolecularly crowded *E. coli* bacterium.
Regional crowding is dictated by the centralised DNA nucleoid, nucleoid-segregated ribosome populations, and ubiquitous RNA polymerases. Much of mRNA translation proceeds via high efficiency polysomal translation, but the interaction of transcription and translation is thought to further coordinate core cellular processes and cellular crowding. The predominant theories for the interaction of transcription and translation are coupling of translation during nascent RNA transcription, and transertion (chain-linking between DNA-RNAP-ribosome-inner membrane proteins). Regional influences may also be exerted by phase condensates (light blue circle, left). Adapted from Bakshi et al., 2015.

1.4 'The importance of being earnest' – understanding more about RNAs in the live cell

Since discarding the "amorphous bags of enzymes" view at the turn of the century (Jacobs & Shapiro, 1999), the tight and nuanced spatiotemporal control and structural regulation of the bacterial cell is a universe open to exploration. As evidenced by the growing pool of knowledge and theories around the complex, refined management of gene expression in the live cell, structure, space and timing are all intercalated and finely tuned.

Here-in-lies the interest in the bacterial world of RNA: characterising the expression of individual genes, defining the relative concentrations and functionalities of RNAs and circulating core machinery, partially determining the phenotypic state of the cell. RNAs seem to play expansive and diverse roles in seemingly all levels in the life of a single *E. coli*. Increasingly apparent, the effects of RNAs are not constrained to their sequence and corresponding function; indeed, their spatial localization, molecular co-localisation and regional behaviours are just as important.

As such, a steadily emergent field is that of determining RNA localization in the cell – fixed and live, at the single molecule and at the whole-cell levels – and exploring consequent biological implications. The following subsections of this Introduction describe and evaluate current methods for bacterial RNA localization[2].

1.4.1 Methodological insights from the whole cell

In the boom of single-cell technologies in the last 30 years, much insight into time-constrained gene expression relating to phenotypic variation has been achieved through the global profiling of RNA/protein distributions, and genomic association studies.

The quantitative and computationally correlative power of system-wide techniques has led to a greater understanding of gene expression sensitive to the single molecule level. In *E. coli*, RNA and protein concentrations for the same gene are seen to be largely uncorrelated (Taniguchi et al., 2010), RNA degradation rates are gene-dependent (H. Chen et al., 2015), and genome-wide expression rates are dependent on growth phase, media and environment (Lamoureux et al., 2023; Sastry et al., 2019). Many of these studies rely on RNA-sequencing as a premise for detecting timepoint-specific RNA populations, and are an excellent starting point to understanding holistic interactions in the single cell.

1.4.2 Methodological insights from fixed cells: fluorescent *in situ* hybridization (FISH)

Relying upon the faithful Watson-Crick base pairing and predictability of the B-DNA structure (Chakraborty et al., 2016), the premise of *in situ* hybridization (ISH) is the antisense hybridization of synthetic (fluorescent) nucleic acid probes to their complementary nucleic acid target. First applied to spatially visualize DNAs in *Drosophila* salivary glands by Purdue & Gall in 1969 with radio-labelled probes, radio-labelled probes were quickly applied to various RNA targets in eukaryotic cells (Rudkin & Stollar, 1977), highlighting non-random mRNA subcellular localization.

ISH, however, relied on immunohistochemical indirect visualization; the introduction of directly fluorescently-labelled oligonucleotide probes in the 1980s (Bauman et al., 1980; Lawrence et al., 1989; Lawrence & Singer, 1986) led to improved resolution into the nm range, reduced experimental time and multi-colour (hence multi-target) potential: the nascence of RNA-FISH (Gerber et al., 2023). There are now various means of designing, manipulating and multiplexing nucleic acid probes and their fluorescent signal, and amid the realization of other

[2] A note here: all of the methods described here were first elucidated for the purpose of eukaryotic mRNA localization. Indeed, there are far more extensive method permutations for eukaryotes, owing to their singular cell membrane and hence higher workability than bacteria. Notable mentions are made.

methods for RNA visualization in single cells, RNA-FISH remains the stalwart for transcriptomic cell-cell heterogeneity (van Gijtenbeek & Kok, 2017).

Single molecule FISH (smFISH)

With rapid developments in fluorescence imaging techniques came the advent of smFISH in 1998 (Femino et al., 1998), adapted for application in *E. coli* by Golding and co. (Skinner et al., 2013; So et al., 2011) – one of the first techniques for single cell visualization of RNAs.

The general protocol follows as depicted in Figure 1.2: cells are fixed and permeabilized, enabling entry of the ~20-50 nucleotide (nt) fluorescently-labelled DNA oligo probes into the cell for complementary binding; probe-cell samples are thoroughly washed to remove unbound probes; and fluorescence imaging may then take place. These initial FISH experiments often experienced poor signal-to-noise (SNR) from high background fluorescence due to unbound or non-specific probes, and poor labelling efficiencies and fluorophore signal hindering resolution. Diffraction-limited single RNA visualization was subsequently achieved via several strategies: the tiling of many short oligo probes to the target RNA multiplied fluorescent readout; probe duplexing at the target RNA for augmented fluorescence; engineering of a repetitive array sequence into the target of interest for multiple consecutive probe-binding events.

These methods have all been applied in the quantitation of endogenous coding and non-coding RNAs (Fei et al., 2015; Sepúlveda et al., 2016; Skinner et al., 2013; So et al., 2011),

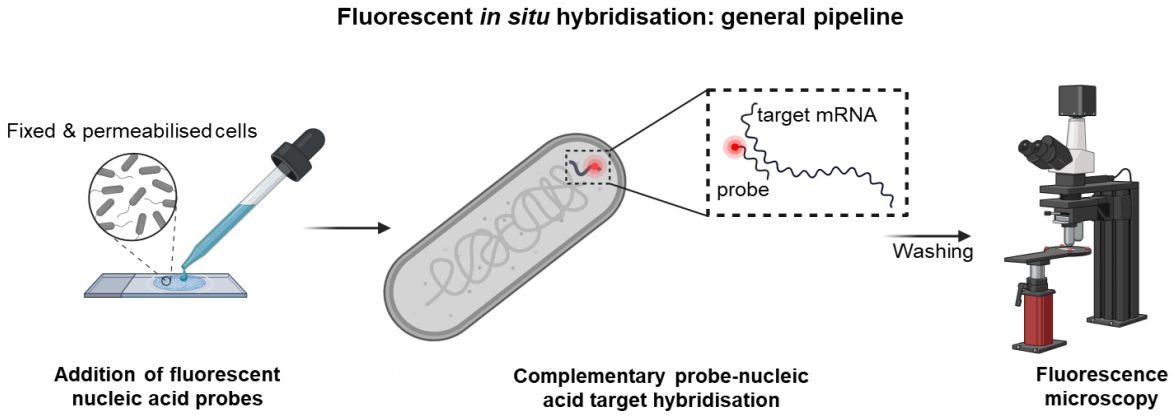


Figure 1.2: Schematic of the fluorescent *in situ* hybridisation (FISH) pipeline.
Bacterial cells are fixed and permeabilised, incubated with fluorescently-labelled nucleic acid FISH probes, sequentially washed and imaged via fluorescence microscopy. Complementary binding of the FISH probe to target nucleic acids (DNA/RNA) give fluorescent signals.

alongside the probing of the transcript lifecycle from initiation through to degradation (Iyer et al., 2016). The many further permutations of RNA-FISH (Branched DNA, Rolling Circle Amplification, Split-FISH) concerning signal amplification have so far only been robustly optimised for eukaryotic systems (Gerber et al., 2023). These would likely require much adaptation for prokaryotic purposes, accounting for the much smaller size of bacterial cells, lower copy number gene expression and well-established challenges of probe internalisation.

Aiming to reduce the duration of RNA-FISH procedures, the release of Turbo FISH has since optimised for hybridization and washing procedures such that quantifiable single RNA foci can be obtained within 5 minutes (Shaffer et al., 2013). This has opened a more accessible avenue for RNA-FISH (particularly for diagnostic purposes), and indicates that there is more room for innovation in the RNA-smFISH sphere.

Multiplexed FISH

The above FISH methods are all relatively low-throughput, enabling expression visualization of up to ~5 genes simultaneously (Shaffer et al., 2013). Multiplexing smFISH via the application of optical barcoding has enabled high-throughput localization of thousands of RNAs, giving a transcriptomic view of the cell at any given time.

By way of multiplexed optical barcoding, two predominant methods are Sequential FISH (SeqFISH)(Eng et al., 2019; Lubeck & Cai, 2012) and Multiplexed Error Robust FISH (MERFISH)(K. H. Chen et al., 2015; Moffitt & Zhuang, 2016). SeqFISH applies a library of readout probes utilising a small number of fluorophores, and 10s of consecutive rounds of hybridization to obtain uniquely barcoded signals for single mRNA identification within the single cell transcriptome. MERFISH, comparatively, uses the concept of 'hamming distance' for quantised separation of barcoded probe signals with a similar small number of fluorophores but less than half the number of consecutive rounds of hybridisation (thus reducing coding errors).

Notably, both SeqFISH and MERFISH have been applied in prokaryotic systems. Cai and co. have subsequently adapted SeqFISH to par-seqFISH, which enabled the transcriptomic characterisation of *Pseudomonas aeruginosa* in planktonic and biofilm cultures, sharing insight into the spatiotemporal metabolic heterogeneity of pathogenic subpopulations (Dar et al., 2021). Zhuang and co. applied MERFISH with superresolution microscopy (SRM) to *E. coli*, remarkably showing transcriptomic-level spatial patterning of RNAs: mRNA localization appeared dependent on the subcellular locations of their proteins, and genomic organization

appeared to have little impact on transcriptome organization (Moffitt et al., 2016). More extensively-applied, MERFISH has also been used in the characterisation of antibiotic resistance phenotypes of clinical pathogens (Chatzimichail et al., 2024).

Fixed FISH methods are thereby the principal means of RNA and gene expression quantitation in a single cell, consistently giving robust spatial information as to RNA locality and subcellular organization. With the small size, cellular robustness and low copy RNA copy number (often <10 copies for endogenous genes (Bernstein et al., 2002)), the rapid progress in SRM has further advanced the capacity of smFISH to visualize RNAs in bacterial cells. No genetic engineering is necessary for classical fixed FISH, and since cells are fixed, functional perturbation is improbable (Fei & Sharma, 2018); however, the chemical fixation/permeabilization procedures can artefactually affect native spatial patterning and probe-target accessibility (Schnell et al., 2012).

1.4.3 Methodological insights into RNA localization in live cells

The precise quantitation of single RNA and transcriptomic-level RNAs in fixed cells holds much relevance for understanding numeric gene expression in the cell. However, with increasing appreciation of the nuanced regulation, low expression rates and import of targeted subcellular localization, the active dynamics and temporal resolution of the life cycles of RNAs are key to understanding more about the live cell. mRNAs, alongside their variable and often ill-defined roles, have been shown to hold half-lives between 2-9 minutes, stability and decay rates often dependent on growth rate (Bernstein et al., 2002; Esquerré et al., 2014) – yet another factor in the life cycle and subcellular localization of mRNAs. Considering the many factors impacting mRNA localization and spatial distribution, visualising and tracking RNAs in the live cell can give further insight into the intercalation and regulation of gene expression in the *E. coli* cell.

There are two main players in live cell RNA localization: genetic engineering for target visualization, and direct endogenous targeting. As alluded to in 1.4.2, the leaps and bounds made in SRM, combinatorial methods and fluorescence chemistry (for brighter, higher photostable and less toxic fluorophores) have enabled processive development of unique live cell methods.

1.4.3.1 Engineered labelling

Fluorescent protein (FP) aptamer systems

FP aptamer systems have been widely used across eukaryotic, prokaryotic and archaic cell systems, leveraging orthogonal protein-RNA interactions for target-specific fluorescence labelling. The Figure 1.3A schematic displays the process of FP-aptamer labelling of target RNAs: an RNA binding motif (stem-loop) is engineered into the locus of interest and an FP sequence fused to the cognate RNA-binding protein (RBP); complementary binding of the FP-RBP to the target RNA stem-loop labels the target transcript. Arraying of RNA binding motifs as repetitive sequences allows for the recruitment of multiple labelled RBPs, amplifying fluorescent signal. The exact locus of binding motif insertion can be used to probe different stages of transcription (5' untranslated region (UTR) for transcription initiation, 3' UTR for termination).

Of numerous RNA-RBP pairs, the MS2 aptamer loop system is the most established. The MS2 phage coat protein (MS2), thoroughly structurally characterized since the 1990s, is the major coat subunit of the *E. coli* MS2 bacteriophage, and dually functions as a dimerized repressor of the viral replicase gene on binding to its cognate RNA motif (Peabody, 1990; Peabody & Ely, 1992; Peabody & Lim, 1996). With high specificity, Singer and co. fused the MS2 protein with GFP, enabling live cell RNA visualization in yeast via RNA-binding motif 3' UTR insertion (Bertrand et al., 1998). Wide-ranging systems and applied targets have since been reported, with varying arrayed repeats (6-128 stem-loop sites)(Nevo-Dinur et al., 2011; Tantale et al., 2016; Vera et al., 2019). The MS2 aptamer system can be parallelized with like systems, such as the PP7 bacteriophage pair (Chao et al., 2008), for observing multiple target transcripts simultaneously, as well as applied with combinatorial methods such as endogenous protein labelling to explore connections between target RNAs and their protein outputs (Katz et al., 2016).

Whilst SNR can be amplified in eukaryotic cells by encoding for site-specific trafficking away from RNA regions of interest (Bertrand et al., 1998), the overall smaller cell size, lesser defined subcellular infrastructure and lower transcript copy numbers in bacteria make the application of FP-aptamer systems more challenging. Golding and Cox first successfully adapted the MS2 aptamer system to live *E. coli*, enabling the tagging of the LacZ gene with 96 motif sites with MS2-GFPs (Golding & Cox, 2004). However, as seen across all applications of the MS2 system in *E. coli* (Golding et al., 2005; Kannaiah et al., 2019 e.g.), *B. Subtilis* (Sattler & Graumann, 2021) and *C. crescentus* (Montero Llopis et al., 2010), the diffuse background fluorescence from circulating unbound and potentially aggregated MS2-FP fusions renders single fluorophore resolution of tagged transcripts inaccessible (Fei & Sharma, 2018). Of course, the necessary arraying and cognate binding also leads to an unnaturally longer transcript, associated with a

heavy, bulky complex (Fei & Sharma, 2018; Gerber et al., 2023; van Gijtenbeek & Kok, 2017). Noted in both prokaryotic and eukaryotic systems, this unphysiological complexation undoubtedly generates artefacts in the transcript's mobility behaviours, and if in the 5' UTR, can further lead to functional disruption (Gerber et al., 2023).

Nevertheless, the MS2 aptamer system has been foundational in enabling the faithful observation of target RNAs in the live cell, generating distinct insights into promoter transcription dynamics (Kandhavelu et al., 2011; Muthukrishnan et al., 2012), and spatial patterning of mRNAs through the live cell (Gijtenbeek et al., 2016; Montero Llopis et al., 2010; Nevo-Dinur et al., 2011).

With the lower photostability, brightness and bulkier nature of FPs as labels, one avenue in recent years has been the use of engineered fluorophore-binding peptide tags over FPs. Instead of encoding a FP for an MS2 fusion, a peptide tag is encoded; on addition of permeable fluorescent ligands to the cell, the ligands can covalently bond to available MS2-fused peptide tags, enabling irreversible fluorescent tagging of MS2 and any complementarily tagged RNAs post-washing of unbound ligand. Many peptide tags are available (Q. Yan & Bruchez, 2015), and two notable options for *E. coli* experimentation are the SNAP-tag (Keppler et al., 2003) (human DNA repair protein derivative, benzyl-guanine ligand) and HALO-tag (Los et al., 2008)(bacterial haloalkane dehydrogenase derivative, chloroalkane dye ligands). With the caveat of bulky fusions as with FPs, the principal improvement comes with prolonged photostability and brightness of the conjugated fluorophores, ideal for extended observation of *in vivo* protein temporal dynamics (Ke et al., 2016).

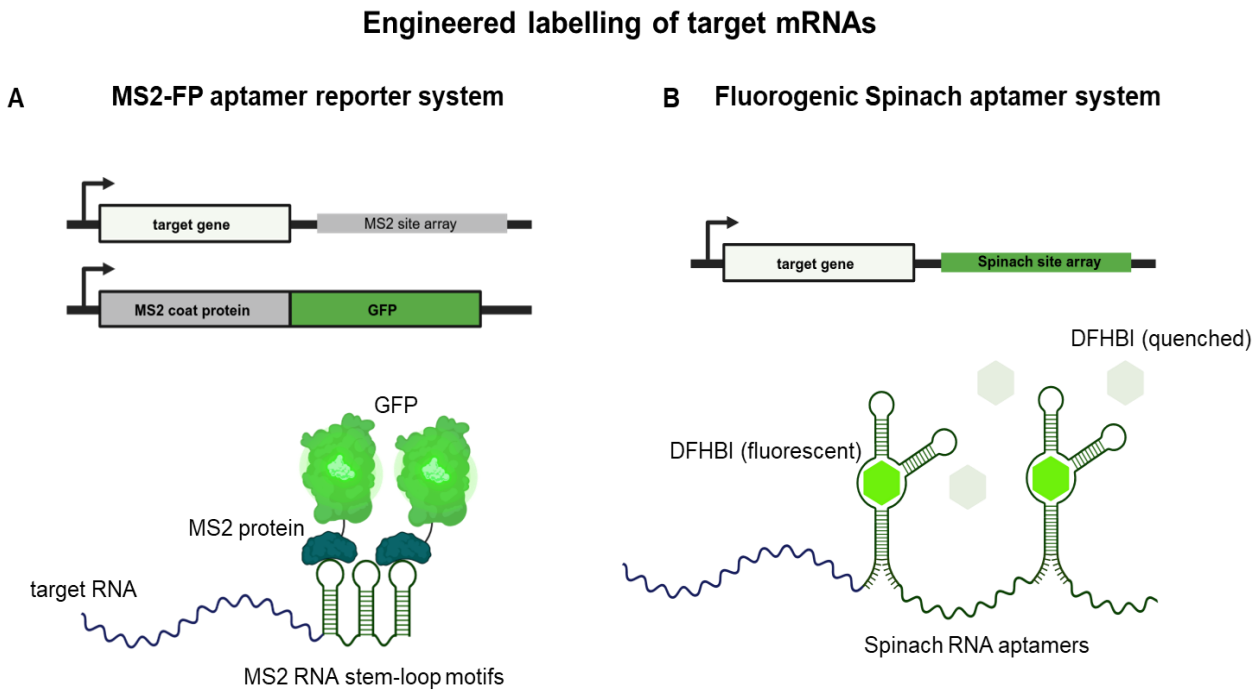


Figure 1.3: Schematic of engineered methods for labelling mRNAs in live bacterial cells.
(A) Engineering of orthogonal RNA-binding protein reporter systems via creating a fluorescent protein fusion that selectively binds to stem-loop motifs enables fluorescent tagging of single target RNAs. The viral MS2 phage protein with a GFP fusion is a seminal example of FP aptamer systems.
(B) Averting heavy FP labels, a similar RNA-motif system utilises fluorogenic chemical labelling of RNA aptamers instead. Spinach as a key example, when adding its cognate fluorophore DFHBI into the cell, binding of DFHBI by the Spinach RNA aptamers elicits fluorescence.

Fluorogenic RNA aptamer systems

By way of avoiding bulky, likely disruptive fluorescent labels with lower SNR, fluorogenic aptamer reporter systems have been developed which make use of various small, diffusive and bright photostable fluorescent molecules. Using a soluble fluorophore mimetic of GFP, binding to a target RNA aptamer enables fluorescence, exhibiting minimal background when unbound (Figure 1.3B). The Spinach aptamer was the first reported pair tractable for live cell applications (binding GFP mimic DFHBI)(Paige et al., 2011).

Several derivatives with improved photostability, folding kinetics and brightness have since been developed (Trachman & Ferré-D'Amaré, 2019), as have refined versions of the DFHBI dye for enhanced yield and fluorogenic efficiency (Ilgu et al., 2016; Song et al., 2014). As above, whilst single molecule resolution has been consistently demonstrated with fluorogenic aptamer systems, single fluorophore resolution has not been achieved predominantly due to insufficient fluorophore brightness, and proximity to single molecule readout requiring tandem arraying of aptamer sites (Zhang et al., 2015). Similarly to the FP reporter systems, this limits

target RNA populations to endogenously abundant RNAs, less suitable for the typical low expression of bacteria (Gerber et al., 2023).

Overall, the capability of the aptameric systems is most significant in high target specificity by endogenous encoding of protein fusions. However, such endogenous constructs are non-native, weighty and perturb the natural behaviours and processing of the target molecules under study. With lower SNR, despite continuing rapid developments, single fluorophore resolution has not been consistently realised, and bacterial cells present a more challenging target holistically through size and structure.

1.4.3.2 Endogenous labelling

Live cell FISH-based methods

With FISH and indeed smFISH so highly elucidated, the ideal scenario would be to specifically target endogenous RNAs in the live cell, without the need for genetic engineering and consequent possibly perturbational reporter constructs. Live cell smFISH would thus require the delivery of exogenous probes complementary to an endogenous target into the live cell, with high photostability and bright fluorophores, low cytotoxicity and low degradational potential (Figure 1.4).

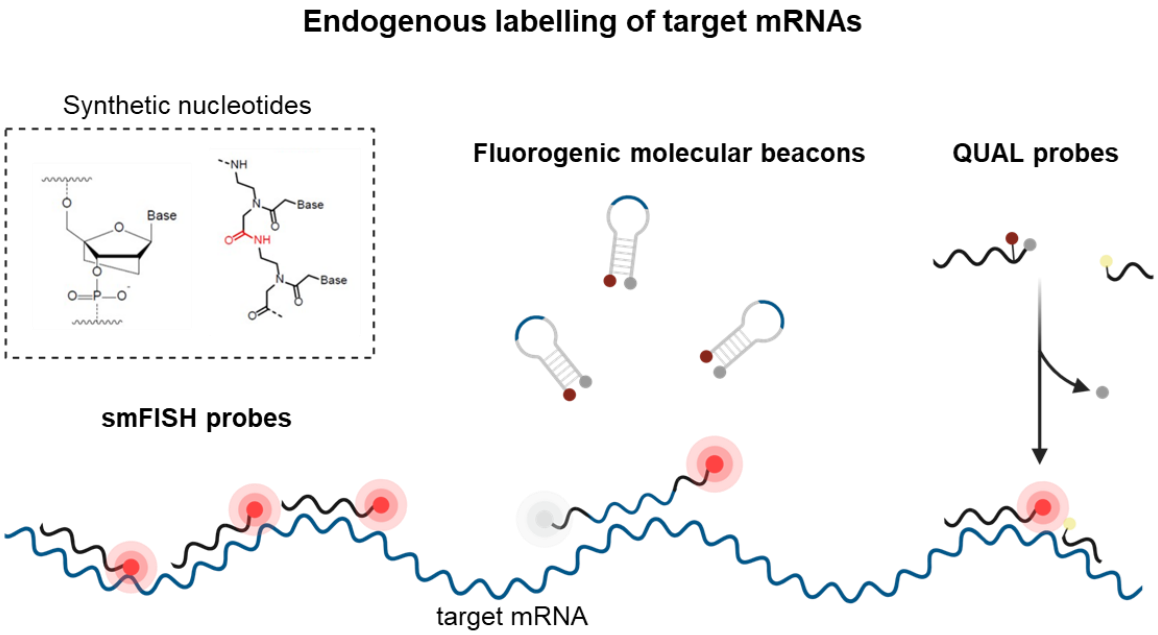


Figure 1.4: Schematic of endogenous labelling methods for mRNAs in live bacterial cells. Endogenous labelling relies on the successful internalisation of FISH-related probes into the bacterial cell. Probes may be designed as fluorescent smFISH-compatible (organic dye-labelled oligonucleotide probe), molecular beacons (FRET-quenched until linearization through complementary target binding enables fluorescence), and QUAL probes (chemical dequenching of the fluorescent probe on complementary binding of the reactive probe pair) as core examples. Improving specificity as a key goal, the use of synthetic nucleotide bases can reduce off-target binding. Adapted from van Gijtenbeek & Kok, 2017.

One seminal approach in constructing endogenously targeting probes is Molecular Beacons (MBs). First introduced in 1996 by Tyagi & Kramer, MBs are modified FISH-like probes comprising a core complementary target sequence, flanked by two self-hybridising arms that are 5'-3'-labelled with a fluorophore and quencher (Tyagi & Kramer, 1996). The dark state exerted by fluorescence resonant energy transfer (FRET) quenching of the fluorophore is relieved on complementary target hybridisation and hence linearization of the MB probe, denoting fluorescence (Santangelo et al., 2004). MBs have been applied in multiple eukaryotic systems (Mao et al., 2020), internalised via microinjection, microporation or endotoxin-mediated poration, but have yet to be successfully tested in prokaryotes. Considering the viable internalisation of labelled biomolecules by electroporation into live *E. coli* cells (Aigrain et al., 2015; Sustarsic et al., 2014), for the study of DNA FRET standards (Plochowietz et al., 2014), tRNA subcellular localization (Plochowietz et al., 2017) and chemotactic response protein CheY behaviour (Di Paolo et al., 2016), it is possible that MBs could be internalized into *E. coli* in a similar manner (van Gijtenbeek & Kok, 2017).

Nevertheless, MBs exhibit many relevant FISH probe-related challenges: imperfect quenching can still lead to spurious fluorescent signal (false positives), and with some propensity to aggregate can generate excessive background fluorescence. With continual technological prospecting in the field of fluorophore chemistry, high efficiency quenchers, like the Black Hole Quencher family (*BLACK HOLE QUENCHER™*), can be selected alongside organic dye fluorophores in line with experimental purpose, alongside wider characterisation of suitable FRET pairings. From a specificity perspective, whilst MBs have been shown capable of differentiating single base mismatches in their target sequence *in vitro*, it is highly likely that MBs in a complex *in vivo* environment could bind to similar sequences to the target (Mao et al., 2020). Target RNAs also ought to be devoid of secondary structures for MB binding, and indeed, MBs could disrupt cognate protein-RNA interactions *in vivo* (Mao et al., 2020).

A potential improvement to target-specific hybridisation is the use of synthetic nucleotides, in the form of peptide nucleic acids (PNAs), locked nucleic acids (LNAs) and 2'-O-methylated oligonucleotides, all of which have been shown to exhibit higher nuclease resistance. PNAs show increased RNA hybridisation specificity versus DNA, due to their inert peptide backbone (Egholm et al., 1993; Nielsen et al., 1991; Socher et al., 2008), and LNAs, higher rigidity and specificity in base pairing due to linking of the ribonucleotidic 2' oxygen and 4' carbon (Koshkin et al., 1998; Obika et al., 1997). Notably, PNAs and LNAs have been substantially characterised in the 16S rRNA tagging for bacterial species identification using RNA-FISH (van Gijtenbeek & Kok, 2017). Quenched autoligating (QUAL) probes by Kool and co. have also been shown to

give highly specific fluorescence on complementary 16S rRNA binding (via adjacent probe binding leading to nucleophilic attack and loss of a 3' DABSYL quencher moiety), with 9mer probes (Sando et al., 2004; Silverman & Kool, 2005). However, a hybridization surfactant buffer was required for QUAL internalisation and sufficient bright signal, which dose-dependently reduced cell viability (Sando et al., 2004; Silverman & Kool, 2005); single molecule capacity is yet uncharacterised.

From this perspective, there is potential to combine learnings from the MBs and QUAL probe designs with the synthetic higher specificity of PNAs/LNAs, towards optimisation of endogenous labelling of RNAs in live bacterial cells by exogenous probes.

CRISPR: a notable mention

CRISPR/Cas systems have been incredibly well-elucidated in a range of systems, and particularly using *E. coli* as a "chassis microbe" for the purposes of biotechnological metabolite production (Lim et al., 2024). Since 2016, CRISPR/Cas systems have also been applied by Yeo and co. for RNA visualization in human cells: utilizing single guide RNAs (sgRNAs) complementary to the endogenous target RNA, labelled Cas9 proteins are specifically guided and label endogenous targets on binding (Batra et al., 2017; Nelles et al., 2016). Despite remarkable observations, and the potential for combination with aptamer systems for signal multiplexing (Wang et al., 2016), these systems did not give rise to single molecule resolution; but, a very recent and promising *in vivo* smFISH methodology has been released by Doudna and co. (Xia et al., 2025). The 'smLiveFISH' method fuses both CRISPR-Csm targeting of specific RNAs with multiplexed sgRNAs, enabling the single molecule resolution of target mRNA behaviours in live human cells. So far, CRISPR/Cas systems have not been realized in prokaryotes for single molecule RNA detection.

It so seems that endogenous labelling offers the most promising prospect in the exploration of RNAs of interest in the live *E. coli* cell, whilst endogenously encoded reporter systems may offer a suitable baseline comparison. As is always the case with novel and combinatorial methodologies, exploring unknown territories requires back-up; and applying these methods in parallel ought to help identify more tractable paths along the way.

1.5 I spy with my little eye: current perspectives on RNA spatiotemporal dynamics in the *E. coli* cell

mRNA localization

With the consistent progress in techniques for the real-time visualization of RNAs and proteins in bacterial cells, a deeper understanding of gene expression and phenotypic governance of the cell is constantly developing. The subcellular localization of these core components in the cell appears non-random, navigating the crowded cellular milieu and even experiencing compartmentalization (Bakshi et al., 2015; Gerber et al., 2023; Irastortza-Olaziregi & Amster-Choder, 2021). In the complex interplay of transcription, translation and degradation, the stoichiometric balance of key actors – the RNAPs, ribosomes and degradosome machinery – physically and functionally determines the processivity of transcription, which is also intrinsically determined by the environment and subsequent cellular growth rates (Bremer & Dennis, 2008; Klumpp & Hwa, 2008).

The current state of play suggests that the majority of mRNAs localize to specific regions of protein requirement. At a more holistic level: following MS2/smFISH observations of translation-independent but protein locale-related localization of select cytoplasmic and membrane protein-encoding RNAs by Amster-Choder and co. (Nevo-Dinur et al., 2011), transcriptomic insights from Moffitt et al. via MERFISH, and Kannaiah et al. using RNA-seq-related procedures have shown whole-cell spatial RNA patterning, with discrete regions of the cell poles, the cytoplasm and the membrane (Kannaiah et al., 2019; Moffitt et al., 2016). Concerning the ~27% of *E. coli* mRNAs visualized by Moffitt et al., further examination of membrane protein-encoding transcripts did appear translation-dependent; in co-labelling signalling peptides involved in envelope-targeted trafficking, enriched mRNAs with signal sequences were observed at the bacterial membrane (Moffitt et al., 2016). Despite a more corroborative perspective of distinct spatial patterning for multiple RNA species in the *E. coli* cell, this is by no means an exhaustive conclusion. Indeed there have been some notable MS2 observations indicating little mobility from the transcription locus, perhaps indicating RNA tethering during nascent transcription (Golding & Cox, 2004; Montero Llopis et al., 2010).

The interplay of transcription, translation and degradation in determining RNA localization

Further exploration by Golding and co. using the MS2 system led to their observation of quantized transcription bursting at the inducible Lac promoter of their target RNA, where both target and MS2 gene transcription were tightly controllable through inducible systems (Golding et al., 2005). This followed a gene activation-inactivation model (ON/OFF states), where transcription activation relies on RNAP holoenzyme binding (Golding et al., 2005; Le et al., 2005). Subsequent findings corroborated that genes are mainly in their OFF state, that burst

size is a key determinant in mRNA counts per gene, and more active genes are “burstier” (i.e. low copy number genes tend to be non-bursty)(So et al., 2011).

Now, considering more recent findings of the global RNAP and ribosome population spatial layouts, the interconnected regulation of gene expression can be seen from single molecule and live cell data alike. Strong nucleoid segregation of ribosomes to the nucleoid periphery and cell poles has been observed by Bakshi et al. (2012), alongside growth-dependent population distributions of RNAPs by Kapanidis and co., with dense RNAP clusters typically forming in the nucleoid periphery during rapid growth (Endesfelder et al., 2013; Stracy et al., 2015). During slower growth rates, both ribosomal subunits and inactive RNAPs seem to leisurely diffuse through the nucleoid (Bakshi et al., 2012; Sanamrad et al., 2014), which appears relatively unstructured (Hadizadeh Yazdi et al., 2012). Comparatively, on transition to fast growth, the nucleoid shows extensive region-specific restructuring (Hadizadeh Yazdi et al., 2012), which coincides with significant upticks in rRNA transcription and number of active ribosomes (Bremer & Dennis, 2008). These upticks would be expected if high transcription activity corresponds to nucleoidal remodelling to locate active genes at the periphery for easy RNAP access, and hence rapid ribosomal interactions to nascent transcripts (Stracy et al., 2015; Valencia-Burton et al., 2009).

Accounting for these complex interactions: transcription by RNAPs bidirectionally influences the structure of the nucleoid; redistributing RNAPs to the higher-mobility nucleoid periphery, circulating ribosomes ready for translation of accessible mRNAs are constitutively available. Ergo, the coupling of transcription and translation for high efficiency turnover of mRNAs during fast growth is *a*. extremely processive to ensure the cell maintains essential components required for growth, and *b*., in the context of RNA localization, suggestive of most RNAs being localized to regions of high translation and protein requirement.

The RNA degradosome represents the final stage in the life cycle of an mRNA, degradation. Endonuclease RNase E forms the core of the complex, providing the mRNA catalytic binding site and acting as a scaffold for degradosome component assembly (Carpousis, 2007); RNase E processing is proposed as the rate-limiting step in mRNA degradation (Kushner R., 2002). A large multienzyme complex comprising endonucleases, exoribonucleases, a helicase and a glycolytic enzyme, RNase E also stabilizes the degradosome via tethering to the inner membrane (Khemici et al., 2008) and interacting with the bacterial cytoskeletal proteins (MinD/MreB system)(Taghbalout & Rothfield, 2007). RNase E-dependent degradation of nascent, free and polysomal mRNAs thus predominantly occurs in the membrane-proximal

cell periphery. Whilst RNase E-mediated degradation of mRNAs is redundant, RNase E is essential and a principal actor of tRNA/rRNA maturation, mRNA degradation, and even selective rRNA degradation in *E. coli* (Arraiano et al., 2010; Carpousis, 2007).

Concerning the delicate interplay of transcription, translation and degradation, growth-dependent RNase E activity acts as a further regulator of mRNA concentration in the milieu. Seminal studies have shown reduced mRNA lifetimes at faster growth (Bernstein et al., 2002; Esquerré et al., 2014), suggesting that the dense co-localized regions, 'hotspots', of proximal transcription and translation machinery aid to regulate essential protein expression over non-essential genes when RNase E-dependent degradation is also upregulated (Bakshi et al., 2015; Le Nguyen et al., 2022; Zhu et al., 2019).

1.6 Overview

Of the formidable techniques in play for visualizing RNA localization, mobility and lifetimes in the live cell, this thesis hopes to contribute to furthering understanding of specific RNA behaviours in live *E. coli*, with as little perturbation to physiological states as possible.

The key aims were:

1. To reproducibly demonstrate the capacity to visualize single mRNAs in the live *E. coli* cell using exogenously internalised, fluorogenic short ssDNA probes that specifically hybridise to a short 5' UTR array site.
2. To apply exogenous probe tagging in the context of a native *E. coli* gene, demonstrating tractability of visualizing native mRNAs from genomic target genes of interest.
3. To optimise for signal specificity and physiological relevance, towards achieving single fluorophore resolution of RNA targets.

This introductory Chapter 1 establishes the problem, current tools and levels of understanding in both prokaryotic and eukaryotic applications. Chapter 2 details materials and methods developed and used throughout the thesis. Chapter 3 briefly introduces the essential background to fluorescence techniques and single molecule imaging, then provides the proof-of-concept of FISH applied in live *E. coli* via electroporation of fluorescent ssDNA probes, and methodological technical optimisations. In Chapter 4, the targeting of *in vivo* smFISH to a native chromosomal site is discussed, and biological implications of all strain observations suggested. In Chapter 5, the passive internalisation of short fluorescent DNAs into live cells is explored. Chapter 6 concludes, with future work.

Chapter 2 – Materials & Methods

2.1 Chapters 3 & 4 – targeting mRNAs via engineered *in vivo* smFISH

2.1.1 Preparation of reagents and strains

Nucleic acids - probes

DNA oligonucleotides were acquired from Biomers.net, IDT and Sigma-Aldrich; further purifications of commercial labelled ssDNA probes were supported by Dr Afaf El-Sagheer (Tom Brown group), and by Dr Rana Abdul Razak (Andrew Turberfield group). The labelled DNA oligonucleotides used in Dr Andrews' experiments and preliminary recapitulations were synthesised by Dr El-Sagheer and Professor Tom Brown.

Oligo sequences and modifications are stated in Table 1:

	Construct sequence and modifications	*Characteristic*	*Source*
2xATTO647N 8mer probe (8mer-2x647N)	5'-(ATTO647N)TCCACCGT(ATTO647N)-3'	Net charge: +1 Quantum yield: 65% (*Atto 647 N, Sigma-Aldrich*)	1. Original: Dr Afaf El-Sagheer, Tom Brown group 2. Commercial: Biomers.net 3. In-house-purified: Biomer.net commercial aliquot, collaboration with Dr Rana Abdul Razak.
ATTO643-BHQ1 8mer probe (8mer-643/q1)	5'-(ATTO643)TCCACCGT(BHQ1)-3'	Net charge: -1 Quantum yield: 62% (*ATTO 643*, ATTO-TEC)	Biomers.net

Oligonucleotides were dissolved in sterile ultrapure Millipore Milli-Q water (upH_2O) to a final concentration of 100 µM, stored at -20°C.

The Biomers' commercial 2xATTO647N 8mer probe (TCCACCGT) aliquot was further purified

in-house to remove any majority free-dye remaining in solution, using reverse-phase HPLC chromatography via an Agilent 1200 HPLC system with Waters XBridge® Oligonucleotide BEH C18 column (4.6 mm x 50 mm, 130 Å pore size, 2.5 µm particle size) and gradient buffers (A: 0.1 M triethylammonium (TEAA), 5% acetonitrile; B: 0.1 M TEAA, 75% acetonitrile) at a flow rate of 1 ml/min over a total run time of 25 min (denoted from now on, in-house-purified). Buffer B in collected fractions was removed *in vacuo*, and selected fractions resuspended in upH_2O, verified by Nanodrop spectrophotometry and diluted to necessary stock concentrations for *in vivo* experimentation. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was further performed collaboratively with Dr Jagadish Prasad Hazra, to check the separation of different aliquots of the 8mer probes.

Plasmids and DNA construction sequences

Plasmids and synthetic sequences were commercially acquired from Thermo Fisher, having been designed in SnapGene, via resources from EcoCyc (Karp et al., 2023) and RegulonDB (Gama-Castro et al., 2016). The TetR-YPet plasmid was generously shared by Alfredas Bukys (Ben Berks group); all cloning was supported by Dr Hafez El-Sayyed.

Plasmids, purpose and origin are stated in Table 2:

Plasmids & Vector constructs	*Key characteristics*	*Purpose*	*Contributors*
Plasmid strain constructs			
p6 (6x tag site array)	pBEC1 – pACBSR backbone (Herring et al., 2003), f1 origin; 6x site array with 6-bp spacers under arabinose-inducible pBAD, *rrnB* terminator and chloramphenicol selectivity (Rebecca Andrews, 2021).	*In vivo* smFISH plasmid system	Dr H. El-Sayyed (plasmid construction)
p24 intermediary (24x tag site array)	p6 plasmid, with a 24x site array with 15-bp spacers exchanged for the 6x array.	Control for fluorescent foci observations with multiple probe binding events.	-

Co-localisation p6	pSynth – 6x site array with 15-bp spacers inserted adjacent to a 1 kB random fragment, under a lac-inducible pTRC promoter and T7 phage RNA polymerase terminator. 25x *tetO* array upstream, alongside *lacI* insertion and kanamycin selectivity.	*tetO* sister plasmid with 6x site array for p6 co-localisation – generating the target tag transcript.	Dr H. El-Sayyed (backbone construction)
TetR-YPet	TetR fusion reporter plasmid following the design of (Reyes-Lamothe et al., 2014); fusion with YPet fluoroscent protein derivative under a lac promoter, low copy number pSC101 origin; chloramphenicol selective.	TetR fusion reporter system plasmid for co-localisation experiments.	A. Bukys (plasmid construction & donation)
Insertion vectors			
pEL-*rpoB*	pHAFcfp backbone; kanamycin selectivity.	Base for phage recombination of tagging native gene *rpoB;* pHAFcfp vector.	Dr H. El-Sayyed (backbone construction)
PKD46	Temperature-dependent expression vector, ampicillin selectivity.	λ-Red phage protein expression vector for λ-Red phage recombination of genomic target sequences.	(Datsenko & Wanner, 2000)
pCP20	Temperature-dependent expression vector, ampicillin and chloramphenicol selectivity.	FLP synthase expression vector, for FLP-FRT excision of recombineered cassette sites.	(Cherepanov & Wackernagel, 1995; Datsenko & Wanner, 2000)

Bacterial strain construction

All strains used in this thesis are derivatives of the laboratory K-12 *E. coli* strain; these are listed in Table 3. The proof-of-concept K-12 MG1655 and BW27784 *E. coli* strains (chloramphenicol-selectable) – transformed with plasmid pBEC1 (termed p6 here) for inducible extra-chromosomal mRNA target production, and for the chromosomally-tagged strains, P1-transduced under a T7 phage RNA polymerase promoter – were constructed by Dr Hafez El Sayyed and tested previously by Dr El Sayyed and Dr Andrews (Rebecca Andrews, 2021).

Strains in use are stated in Table 3:

E.coli Strain	*Origin*
Wildtype K-12 BW27784	Berks group
Wildtype K-12 MG1655	Kapanidis group
BW27784 p6	Dr H. El Sayyed
BW27784 chromosomal T7	Dr H. El Sayyed
MG1655 p6	Dr H. El Sayyed, EIL
MG1655 chromosomal T7	Dr H. El Sayyed
MG1655 psynth-p6	EIL
MG1655 *rpoB*-tag	EIL

Vector molecular cloning and plasmid strain construction

All PCRs performed towards plasmid constructions and amplifications utilised the high-fidelity Q5 DNA Polymerase (NEB) according to manufacturer's reagents and protocol, or the Phanta/Ghost DNA Polymerases (Vazyme) with corresponding manufacturer instructions. HiFi Q5 reactions (NEB) or DreamTaq DNA Polymerase and partner reagents (NEB) were used for verification PCRs in checking for successful transformation and transduction of recipient *E. coli* cells with target insertions.

For plasmid constructions and transduction vectors, Gibson assemblies were conducted using the HiFi Q5 DNA Polymerase (NEB) and the corresponding NEBuilder HiFi DNA Assembly mix (NEB), as per manufacturer's directions. Resultant PCR amplicons were analysed via agarose gel electrophoresis (AGE) (1-3% agarose gel pre-stained with 1:100 SYBR Safe (Invitrogen)). Desired fragments were excised, and purified via the GeneJet Gel Extraction Kit (Thermo Fisher Scientific), or purified directly from the PCR mixture via the GeneJet PCR Purification Kit (Thermo Fisher Scientific).

Multisite plasmid-target strains were constructed by transforming K-12 MG1655 strains via the chosen target site number hosted on a modular vector. The modular vector was created from the chloramphenicol-selective pBEC1 plasmid (pBAD, low-intermediate copy number), such that any promoter of interest and any multisite tag sequence could be incorporated via flanking primer-overhang sequences and Gibson assembly. Tag sequences may also be inserted/excised via restriction enzyme digestion as an alternative approach.

Commercially-obtained 6-site and 24-site tag sequences (GeneArt, Thermo Fisher Scientific) were subsequently PCR-amplified and incorporated into the modular backbone via Gibson assembly, verified through AGE and purified to generate p6 and p24 (pEL) plasmids. Recipient MG1655 cells were transformed through heat-shock, 42°C for 43 seconds, and recovered for 1 hr at 37°C in Super Optimal broth with Catabolite repression (SOC) medium. Transformed cells were selected by lawn-plating on low-salt LB agar plates supplemented with the necessary antibiotics (50 µg/mL chloramphenicol for all plasmid strains, with additional kanamycin for the double-transformant co-localised p6 strain).

In the event of needing further starting plasmid volumes, plasmid DNA was extracted from existing MG1655 strains using the Plasmid DNA Miniprep Kit (Thermo Fisher Scientific).

Constructions were confirmed via Nanopore sequencing (Source Bioscience, UK).

Chromosomal strain construction

The 6-site-tagged *rpoB* promoter (rpoBp) strain was constructed via the replacement of the endogenous rpoBp with a tag-fused rpoBp of the same sequence. The rpoBp and 6-site tag sequence, linked by a 20-bp linker sequence, were incorporated into the pEL-*rpoB* plasmid, and the target cassette (upstream *frt*-flanked kanamycin selectivity cassette, promoter and tag array) were amplified using primers with 50-nt overhangs homologous to the target chromosomal insertion site (5' untranslated region (UTR) of the *rpoBC* transcriptional unit), complementary with Lambda red phage processing. Purified via the GeneJet PCR Purification Kit (Thermo Fisher Scientific), the DNA fragments were electroporated at 1.8 kV/cm field strength into PKD46 MG1655 *E. coli* cells overexpressing Lambda red proteins. Cells were recovered for 3 hr at 30°C and lawn-plated on 50 µg/mL kanamycin, low-salt LB plates for overnight growth at 37°C.

Target site-specific insertion of the fragment into the chromosome was verified via colony PCR using primers flanking the *rpoB*C transcriptional unit. The verified fusion was transferred into recipient wildtype MG1655 cells via P1 transduction and ampicillin selectivity; further strain

clean-up to excise the kanamycin selectivity cassette was conducted via the site-directed FLP-FRT recombination system, on transformation of the pCP20 plasmid encoding a FLP recombinase. Seemingly, *kan* cassette excision was not 100% successful, and it should be assumed that the *rpoB/rpoBC*-tagged chromosomal strain has a maintained/partial kanamycin selectivity cassette transduced into the chromosome. This strain will be referred to as the *rpoB*-tag strain henceforth.

2.1.2 *E. coli* electrocompetency procedure

Overnight cultures of select *E. coli* strains were made using inocula from frozen glycerol stocks (final 25% glycerol), grown in 5 mL LB at 37°C, shaking at 125 rpm. The overnight cultures were diluted 1:100 the following day into a volume (ranging from 25 mL to 200 mL) of fresh low-salt LB or glucose-supplemented M9 (M9Glu) medium (2x recipe in 50 mL: 5 mL 20x M9 salts, 1M casamino acids, 1M vitamin supplements, 1M $CaCl_2$, 1M $MgSO_4$, 0.2% glucose, upH_2O) with the 50 µg/mL kanamycin, 35 µg/mL chloramphenicol as necessary, in a 250 mL Erlenmeyer flask. The batch cultures were grown at 37°C, 125 rpm, and regularly sampled for optical density (OD). Upon reaching OD_{600} = 0.20 ± 0.05, cell cultures were removed from incubation and allowed to chill on ice in 50 mL aliquots in pre-chilled 50 mL Falcon tubes for 15 min.

For arabinose-induced strains, batch culture volumes were split in half approximately 15 min before OD_{600} = 0.20 in accordance with approximate bacterial doubling time, and one of the two halved cultures was induced via the addition of 1% L^+-arabinose, incubated at 37°C, shaking at 125 rpm.

Following chilling, the 50 mL culture aliquots were spun down for 10 min at 7000 rcf, 4°C, and sequentially washed with chilled sterile upH_2O volumes of 50 mL, 25 mL and 15 mL, down to final combined resuspension of ~1 mL. After this final wash, cells were spun for 2 min at 8000 rpm, resuspended in 15% glycerol solution (900 µL for a 100 mL initial batch for optimal cell density at imaging), and aseptically aliquoted into 20 µL volumes in PCR tube strips and immediately stored at -80°C until use.

Individual experiments were optimized for the use of -80°C cell aliquots, having checked comparisons of electrocompetent cells made on the day from fresh cultures and used immediately, from freshly-made -80°C electrocompetent cell aliquots, and from variable-age -80°C aliquots.

2.1.3 *In vivo* smFISH experimental pipeline

Electroporation & sample preparation

The pipeline as described follows the published protocol documentation in Aigrain et al., 2015 (Aigrain et al., 2015).

For each 20 µL cell aliquot, 5 µL of labelled-ssDNA (stored at -20°C in sterile upH_2O, stock concentrations to give final concentrations of 5 nM-1 µM in a 25 µL total volume) were added, gently mixed and incubated on ice for 5 min. Each competent cell and labelled ssDNA mixture was then transferred into a prechilled 0.1 cm electroporation cuvette, and placed in a Bio-Rad Micropulser Electroporator to receive a 1.4 kV/cm field strength shock. 500 µL SOC (Thermo Fisher Scientific) was added immediately after electroporation under flame, and all samples were recovered at 37°C for 20 min, shaking at 125 rpm. Post-recovery, cells were harvested by centrifugation for 3 min at 3300x*g*, 4°C, and sequentially washed 4x times with 500 µL 1x phosphate-buffered saline (PBS, Thermo Fisher Scientific) with 2 min centrifugation intervals. Cells were finally resuspended for in 10-30 µL fresh 1xPBS for imaging, and kept on ice.

For positive controls (corresponding to testing for electrocompetent cell autofluorescence), cell aliquots were electroporated with 5 µL chilled sterile upH_2O instead of ssDNAs.

For control experiments exploring RNAP block, 100-200 µg/mL rifampicin (Rif) were added to the SOC medium in recovery, recovered and washed 4x times with 1xPBS and resuspended in 30 µL 1xPBS as standardised above.

In Chapter 3.5.4 (Appendix 1.1), testing parameters in recovery from electroporation, the above protocol was performed with adaptations to the probe incubation time duration, recovery time duration and temperature, and the recovery media in use.
For probe incubation, recovery duration and temperature experiments:
Cells were mixed with 5 µL of probe stock for a final concentration of 50 nM for a period of 5 or 10 min on ice. Cell samples were then recovered in 500 µL SOC for either 5 or 20 min, at room temperature or at 37°C (125 rpm); subsequent steps as above.
For the recovery media experiment: cells were mixed with 5 µL of probe stock for a final concentration of 50 nM for a period of 10 min on ice. Cell samples were then recovered in 500 µL M9Glu, SOC or rich-defined medium (RDM, recipe as (Neidhardt et al., 1974)) for 20 min, at 37°C, 125 rpm; subsequent steps as above.

Agarose pad preparation

During the recovery incubation, 1% agarose pads were made by mixing a 1:1 ratio of molten 2x agarose solution (1:1000 BioRad Certified Molecular Biology Agarose powder in 5 mL upH_2O) and 2x M9Glu medium. 150-200 µL was pipetted circularly onto each cover slide, a burnt slide

placed on top of the agarose pad, and allowed to cool to stably set. Pads were kept in the dark until use, and moist during longer experiments via a upH_2O reservoir in the slide holders. Burnt slides were generated by Sanna Piippo-Henderson, via heating glass microscopy slides incrementally to 500°C (by 100°C over 5 hours), burnt at 500°C for 1 hour, and cooled incrementally by 150° per hour for sterilisation and removal of optical debris, ready for supperesolution microscopy.

Immediately prior to imaging each sample, each sample aliquot was rapidly mixed by pipetting or brief vortexing, and 3-4 µL per sample added across the pad. The burnt slide was replaced onto the pad surface, and placed burnt side-down towards the immersion-oil covered objective for imaging.

For the Rif control experiments, the agarose pads were prepared to comprise a final concentration of 100-200 µg/mL Rif.

Superresolution microscopy

Superresolution microscopy was performed using the commercial superresolution microscope, the Oxford Nanoimager (NIM).

For single molecule RNA live-cell imaging, single-channel data were acquired with a 50 ms exposure time per frame (standardized, ranging from 5 ms to 500 ms frame times) using 638 nm (red channel) constant laser activation between 5 mW to 25 mW, corresponding to the organic fluorophores in use. All samples were imaged in highly inclined and laminated optical sheet (HILO) mode, the 51° illumination angle the same across all experiments. Brightfield images were taken in the green channel, using white LED illumination.

Alongside constant laser exposure, dark interval video acquisition was performed as a means of extending photon budget and lifetime. Dark interval imaging data were acquired through a 5:1 ratio exposure regime, alternating 638 nm laser activation at 5-10 mW for 50 ms, every 250 ms (dark), in HILO.

For co-localization experiments, where organic fluorophores of different excitation-emission wavelengths were in use, single-channel data were acquired sequentially as: i. red channel data for RNA imaging as above, and ii. green channel data for TetR-YPet marker visualization via 50 ms exposure at 5-20 mW 450 nm laser power in HILO, up to 100 frames.

2.1.4 Cell viability assessment

Visual cell division in a field of view (FOV)

Cell viability experiments were prepared and performed using the above procedures. In live-cell imaging, brightfield images of cell populations in a set region of the agarose pad were acquired at times 0, 20, 40 and 60 min. Multiple FOVs were obtained across timepoints for a single slide by utilising the ONI NIM user software interface to programme FOV coordinates, enabling manual selection of specific regions of the agarose pad to consecutively image.

Cell viability was then determined by manually comparing timepoint brightfields per pad region, and classifying cells as 'growing', 'divided', 'static' or 'compromised', as in (Plochowietz et al., 2014; Sustarsic et al., 2014).

Plate reader – ensemble growth curve measurements (baseline verification)

Cell samples were generated from freshly-grown MG1655 *E. coli* strains, single competent cell aliquots or post-experimental cell samples (originally from frozen competent cell aliquots).

Cell samples were centrifuged at 7000 rpm for 2 min at ambient temperature; pellets were resuspended in 1 mL corresponding media of interest (low-salt LB or M9Glu), and optical density measured at OD_{600}. Each sample was then diluted accordingly to reach a time 0 starting OD_{600} = 0.01, aliquoted as 200 µL triplicates into a 96 Greiner BIO well plate. Control triplicate wells contained 200 µL of the corresponding base media.

Absorbance readings were taken on a CLARIOStar plate reader: every 5 min up to 30 hr (orbital mode), with continuous 300 rpm shaking between readings. Samples were blanked against media controls. All strains grew to give standardized growth curves; data are not shown in this thesis, but the method is documented here for sake of completeness.

2.1.5 Data analysis

Cell segmentation & fluorescence intensity – populational measurements

All experimental FOVs documented were segmented using the Napari plug-in, 'BagSeg' (Farrar et al., 2024); using self-curated, pre-trained machine-learning models to segment *E. coli* cells from brightfield images (>2000 FOVs). The BacSeg Napari plug-in is available open source at https://github.com/piedrro/naparibacseg.

Resultant segmentations per FOV were curated to remove any anomalous segmentations or segmentations of compromised cells, and exported in file formats suitable for ImageJ (TIFF, open source access via https://imagej.net/ij/) and MATLAB processing (OUFTI, accessible via https://oufti.org/ (Paintdakhi et al., 2016); MathWorks, https://uk.mathworks.com/).

Representative FOVs throughout the thesis are depicted as false-coloured maximum intensity Z-stack projections (red or green according to dye wavelength; no colour for upH_2O), overlaid with the segmented cell outlines; stacked projections are all adjusted to the same brightness and contrast range. Maximum intensity projections are shown for representative FOVs presenting the brightest (I_0) and 100th (I_{100}) frames' visuals, exact number of projected frames defined per figure.

Mean fluorescence intensities per cell area were then obtained using a custom python script: utilising the OUFTI cell masks, complementary brightfield and video data, cell masks were overlaid onto the video and mean fluorescence intensity per cell area calculated at I_0, I_{100} (defined 100 frames after I_0) and the mean intensity per cell area over the course of the video. Background fluorescence intensity was calculated as the background intensity per pixel, from the regions classified as outside cell masks.

Cell intensity values were corrected for autofluorescence by subtracting the background intensity per pixel from mean cell intensities. Cells were then classified as loaded if the intensity values were non-zero following the subtraction of the averaged mean cell intensity per cell area (from I_0) of "empty cells", (cells electroporated with upH_2O). The median of all background-corrected mean intensity per cell area values for loaded cells was calculated as the representative value of intensity between samples, given the median is more robust to outlier skew. The whole cell population, i.e. background-subtracted mean intensities per cell area, were used to plot histograms denoting the distribution of intensities acquired across a heterogeneous cell population.

In Chapters 3-4, a third classified population, loaded tag-strain cells corrected for off-target signal, was calculated by further subtracting the averaged mean intensity per cell area value from wildtype cells, as a proxy for specific signal.

Photobleaching time traces and the unitary intensity of a focus

For select representative FOVs, the mean intensity per cell area across all cell masks for all frames per video were extracted using the TIFF segmentation masks in ImageJ.

To obtain individual cell bleaching traces and an averaged bleaching trace across a FOV, an average photobleaching time trace was generated from all visually loaded cells measured and fit with a single or double exponential. Taking the y0 as baseline, the y0 value was used to equivalate the background intensity value – enabling a correction for autofluorescence per cell area by subtracting the per-FOV y0 from all mean intensity per cell area values, and the

averaged intensity trace. Individual cell intensity traces could then be plotted over time alongside the averaged intensity trace per FOV, all background corrected.

To obtain the unitary intensity of a single fluorescent focus – in theory, a probe-hybridised mRNA – photobleaching time traces of individual cells loaded with 1-10 molecules per cell were plotted. Traces with 1-10 visible step changes were then analysed through *AutoStepfinder* (Loeff et al., 2021), an automated, bias-robust step detection method for single molecule kinetic states, curated per trace (parameters: 500 iteration range, threshold 0.1-1.0, median fit) for optimal fits to each trace. Step sizes above 3 were retained (as anything below likely correlated to slight overfitting) and from >100 cells per condition, a step-height histogram was plotted. Fitting of the step height histograms with a single-fit Gaussian peak function gave a peak centre, which was taken as the approximate molecular unitary intensity of bright spots per sample condition. Noted, step sizes up to 30 were retained, but perhaps for histogram and peak plotting a shorter step size range would be optimal for better fitting of a single Gaussian peak function.

To calculate the approximate number of molecules across number of cells in a sample, the I_0 mean fluorescence intensities per cell area were divided by the corresponding unitary intensity for cells under the same conditions. The spread of molecule counts across the cells was then plotted as a histogram in OriginPro 2021.

Single molecule localisation and tracking

Custom-written MATLAB software, Locoli and StormTracker, were used to perform single molecule localisation and tracking analyses of fluorescent foci in live *E. coli* cells (Crawford et al., 2013; Uphoff et al., 2013). Point-spread functions (PSFs) by elliptical Gaussian fittings of fluorescent foci were localised to a precision of 40 nm, and localized PSFs were linked to form tracks if observed 4 frames consecutively. A memory parameter of 1 was used to account for fluorescent blinking in forming tracks, and a threshold of 18 used to limit off-target signal inclusion in analyses.

The apparent diffusion coefficient (D*) of fluorescent foci was then calculated via the above custom MATLAB software, using the mean squared displacement (MSD) of each track (minimum 4 linked segments, and truncated at 5 linked segments). The D* is an approximation of the actual particle diffusion, acknowledging the likely impact of foci motion blurring and localization precision error (Michalet & Berglund, 2012; Uphoff et al., 2013).

D* histograms were generated in MATLAB with a two- or three-species peak function (free two species, three gamma two constrained at 0.1 $\mu m^2/s$ bound, 0.3 $\mu m^2/s$ intermediary (Chapter 4)) fit to the D* histograms. MSD by time plots were generated in OriginPro 2021.

Analysis of molecular behaviours

To analyse and depict the behaviours of individual molecules and subsegments of molecular trajectories in live cells, cells from the presented representative FOVs were localised, tracked and analysed through the Napari python plug-in, *Moltrack*, generated by Dr Piers Turner. The *MolTrack* Napari plug-in is available open source at https://github.com/piedrro/napari-moltrack.

Tracking was performed through Picasso (Strauss, 2022), and similarly used a minimum consecutive frame number of 4 for linking trajectories, and truncating tracks at 5 segments for D* calculation from the per-molecule MSD. Individual molecule and track statistics were extracted, and plotted accordingly in OriginPro 2021.

2.2 Chapter 5 – Passive internalisation of fluorescently labelled ssDNAs

DNA oligonucleotide preparation

Acquired oligonucleotides were dissolved in sterile upH_2O to a final concentration of 100 µM, stored at -20°C. Oligo sequences and modifications are noted in Table 4:

	5' modification	Sequence	3' modification	Origin
2xATTO647N 8mer probe	ATTO647N	TCCACCGT		Dr R. Andrews
EUB338-1x ATTO647N (EUB-647N)	ATTO647N	GCTGCCTCCCGT AGGAGT		Dr M. Kümmerlin
EUB338-1xCy3B (EUB-Cy3B)	Cy3B	GCTGCCTCCCGT AGGAGT		Dr A.Farrar
ATTO643-EUB338-BHQ1 (EUB-643/q1)	ATTO643	GCTGCCTCCCGT AGGAGT	BHQ1	Dr M. Kümmerlin
8mer T-repeat	ATTO647N	8xT		Dr N. Shiaelis
20mer T-repeat	ATTO647N	20xT		Dr N. Shiaelis

40mer T-repeat	ATTO647N	40xT		Dr N. Shiaelis
75mer T-repeat	ATTO647N	75xT		Dr N. Shiaelis

K-12 MG1655 E. coli cell preparations – fresh cell experiments

Overnight cultures of wildtype *E. coli* (5 mL LB seeded from fresh colonies) were diluted 1:500 or 1:1000 in low-salt LB or M9Glu (2x recipe in 50 mL: 5 mL 20x M9 salts, 1M casamino acids, 1M vitamin supplements, 1M $CaCl_2$, 1M $MgSO_4$, 0.2% glucose, upH_2O) media and grown at 37°C, 250 rpm to an OD_{600} = 0.25 ± 0.5.

Fresh cells were then harvested by centrifuging 5mL for 2 min at 7000 rpm, room temperature; pellets were resuspended in 100 µL 1xPBS. Per aliquot, 5 µL of the recipient concentration probe solution were added, gently mixed by pipetting and left for 10 min incubation on the bench (standard procedure). Each cell aliquot was then suspended in 500 µL 1xPBS (room temperature), centrifuged at 3300x*g* for 3 min at 19°C, and then washed a further 3x with 500 µL 1xPBS. Cells were finally resuspended in 30 µL 1xPBS.

For temperature control experiments – as above, except cells were incubated for 10 min with 100 nM EUB338-1xATTO647N probes at 4°C, 19°C or 37°C.

K-12 MG1655 E. coli cell preparations – competent cell experiments

Competent cells were prepared in the same manner as outlined in 2.1.2, with the exception that cells were grown to OD_{600} = 0.25 ± 0.5 (further into early log phase, exponential).

Technical experiments were conducted as outlined above, using competent cells thawed on the bench at room temperature.

For temperature control experiments – as above, except cells were incubated for 10 min with 100 nM EUB338-1xATTO647N probes at 4°C, 19°C or 37°C.

For length dependence of internalisation at room temperature, cells were incubated with 100nM probes of T-repeats tagged with 5' 1xATTO647N, of 8 nt, 20 nt, 40 nt and 75 nt lengths.

Cell viability assessment

Cell viability by observed growth was performed as documented in 2.1.4.
Competent LB and M9Glu wildtype MG1655 aliquots were incubated with 100 nM EUB338-1xATTO647N for 10 min at room temperature, then washed 4x times consecutively with 500 µL

1xPBS (centrifugation 3, 2 min at 3300x*g*, 19°C). Cells were resuspended in 30 µL 1xPBS and ~4.5 µL cell solution added to M9Glu agarose pads.

Per cell sample, a single agarose pad was imaged for 60 min: recording the positional locations of 3-5 FOVs using the NIM coordinate programming, BF images were taken every 20 min. Fluorescence images were acquired at T0 and T60 across all FOVs.

Microscopy & Analysis

Superresolution microscopy was performed using a second Oxford Nanoimager (NIM); with slightly different alignment and brightfield illumination, this NIM was used for all Chapter 5 experiments.

For single molecule ribosome live-cell imaging, single-channel data were acquired with a 50 ms exposure time per frame (standardized, ranging from 20 ms to 250 ms frame times) using 638 nm (red channel) constant laser activation of ~10 mW (between 5 mW to 25 mW), corresponding to the organic fluorophores in use. All samples were imaged in HILO, the 53.5° illumination angle the same across all experiments. Brightfield images were taken in the green channel, using LED illumination at 465 nm (blue light to avoid photobleaching the fluorophores).

The analysis pipeline is the same as described in 2.1.5, with some amendments: firstly, the loading categorisation in ensemble fluorescence intensity comparisons was based on the averaged mean intensity per cell area of empty cells (where for passive internalisation, empty is defined as wildtype cells incubated with upH_2O only).

In comparing different internalisation methods, a statistical analysis of ensemble mean intensities per cell area was also performed via the non-parametric Kruskal-Wallis test, followed by a Dunn's post-hoc test for multiple pairwise comparisons, in OriginPro 2021(https://www.originlab.com/).

Contrary to diffusion analyses using custom MATLAB software, all diffusion analyses in Chapter 5 were performed using *Moltrack*. As in obtaining individual track statistics in 2.1.5, *Moltrack* localisation and tracking was performed via Picasso, linking trajectories with a minimum consecutive frame number of 4, truncating tracks at 5 segments for D* calculation. A memory parameter of 1 was employed, and localisations thresholded for within cell segmentation areas only. Deviating from common practice: whilst all retained tracks were truncated into 5-frame segments, all truncations were kept (as opposed to just the first 5-segment trajectory). This compensated for fewer datapoints, enabling the generation of a D*

histogram, but is over-representative of longer molecular trajectories and serves as an initial readout of a viable single molecule experimentation pipeline.

Chapter 3 – Establishing engineered *in vivo*, single molecule FISH via short, transient fluorogenic probes

The Introduction gives an insight into the various approaches for single molecule probing of target RNAs in live bacterial cells. However, many of the well-elucidated live cell approaches involve bulky RNA arrays and fluorescent fusions for target RNA visualization, which can perturb the physiology and function of the target RNA mobility, behaviour and processing. Additionally, where single molecule fluorescence measurements have been achieved, visualizing single fluorophores has yet to be reproducibly described. The objective of Chapter 3 is hence to explore the feasibility of taking a FISH approach to tagging single target RNAs in the live *E. coli* cell, with minimal perturbation to the target RNA via genetic constructs, and down to a possible single fluorophore resolution.

3.1 A whole new world: super-resolution

3.1.1 Single molecule fluorescence microscopy; a brief overview of applied methods

From the first aqueous single fluorophore observations by the Keller group in 1990, a single molecule boom took place over the next 15 years, leading to cumulative innovations in the fields of fluorescent protein (FP) and organic dye developments, and applied microscopy. Such was the progress, Abbe's long-standing diffraction limit (Abbe, 1873) was definitively overcome by 2006, via a number of possible "smart" manipulations of fluorophore stochastic or selective activation/deactivation as evidenced by the respective Hell (Hell & Wichmann, 1994; Klar & Hell, 1999), Betzig (Betzig et al., 2006) and Zhuang (Rust et al., 2006) groups, or signal-decoding by the Heintzmann Group (Heintzmann & Cremer, 1999).

Further rapid developments led to the now-standardised super-resolution imaging techniques of Stimulated Emission-Depletion microscopy (STED, Donnert et al., 2006), Superresolution Structured Illumination Microscopy (SR-SIM, Heintzmann & Huser, 2017), and a multitude of single molecule localization-based (SMLM) techniques: Photoactivated Localisation Microscopy (PALM, Manley et al., 2008; Shroff et al., 2008), Stochastic Optical Reconstruction Microscopy (STORM, Bates et al., 2007), direct STORM (dSTORM, Wombacher et al., 2010) and Point Accumulation for Imaging in Nanoscale Topography (PAINT, Sharonov & Hochstrasser, 2006). All gave rise to nm spatial resolution *in vitro* and *in vivo* alike. A whole new world of exploration for the fluorescence microscopy community, the 2014 Nobel Prize in Chemistry for the "development of super-resolved fluorescence microscopy" was duly awarded to Eric Betzig, Stefan Hell and William Moerner for pioneering single molecule and super-resolution microscopy methods and fluorophore technologies.

With such dynamical resolution developments, single molecule interests expanded from highly controllable *in vitro* systems to the litany of intertwined processes and variables occurring in the live cell. Labelled protein fusions have since paved the way for single molecule gene expression, spatiotemporal protein dynamics and adaptive regulation studies (Choi et al., 2008; Elf et al., 2007; English et al., 2011; Reyes-Lamothe et al., 2014; Yu et al., 2006) – and all in the context of a crowded, (often) phase-separated and exquisitely regulated cellular milieu (Ellis, 2001; Gao et al., 2021).

SMLM is a diverse methodology for visualising single species and dynamic motion over time. Applying conventional widefield imaging with precise spatial localization processing, hundreds of single fluorescent species can be resolved to ~20-50 nm over large fields of view (FOVs) to varying depths (Lelek et al., 2021). For all experiments conducted in this thesis, highly inclined and laminated optical sheet (HILO) mode[3] was applied, common in live cell imaging to achieve a greater penetration depth and reduced light scattering for higher resolution of intracellular targets, comparative to epifluorescence (Tokunaga et al., 2008).

The basis of SMLM relies on the definition of a particular fluorescent particle by the pixel coordinates of its imaged fluorescence point-spread-function (PSF), iteratively imaging and localizing the single particle to obtain a localisation precision per particle over time, then cumulatively processing localizations into a single image (Lelek et al., 2021)(Figure 3.1A). Since the accurate localizations are reliant on precise spatial determination, the core assumption is that PSFs are non-overlapping – hence the critical importance of excellent signal-to-noise ratio (SNR), low autofluorescence background, and experimentally-optimised fluorescent molecule density (Lelek et al., 2021). Any additional boosts to precision by way of fluorophore photostability and temporal separation also promote higher accuracy localization.

[3] HILO is one of three principal widefield modalities, the other two being epi-fluorescence and total internal reflection fluorescence (TIRF). Epi-fluorescence is more appropriate for larger samples due to the wider illumination region achieved, and TIRF with low penetration depth and high SNR optimal for surface-level imaging of live samples, such as membranal or cell wall studies. HILO for intracellular targets gives ~8x better SNR than epi-fluorescence mode, and is hence preferable for this use-case.

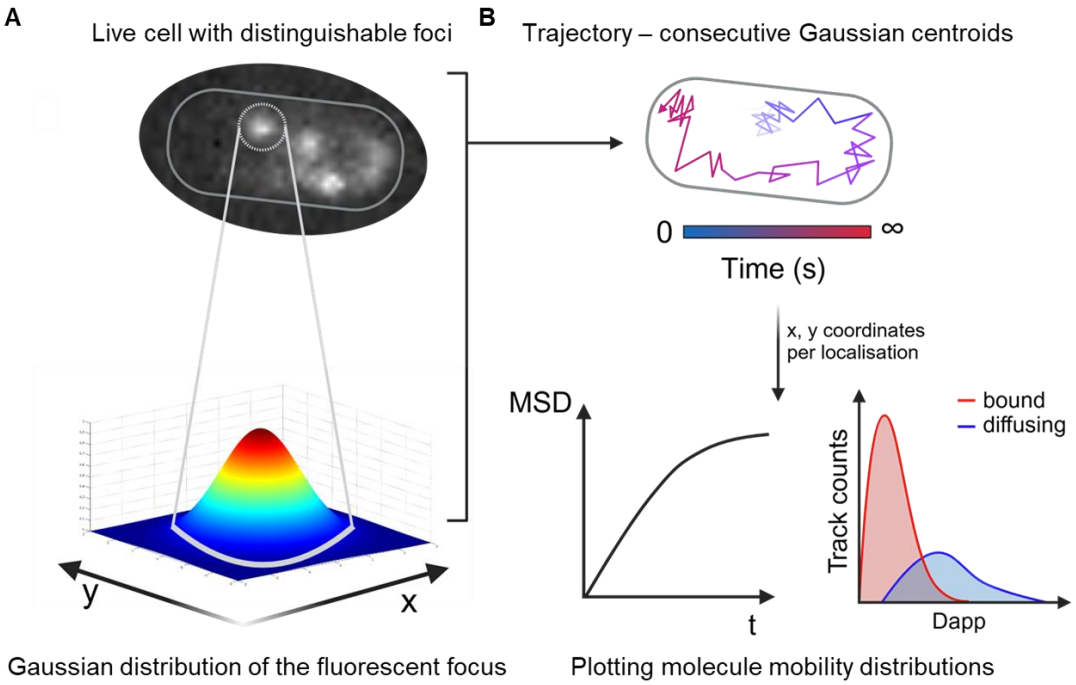


Figure 3.1 – Schematic displaying the SMLM process from acquisition of a live-cell single molecule to analysed mobility distributions.
(A) Localisation of single molecules in SMLM by fitting a 2D-Gaussian to each bounded molecule, acquiring x,y coordinates to define the centroid of the molecule as molecular identity (peak of the Gaussian) – occurring for each localised molecule across time.
(B) For a given localised molecule, consecutive centroids are linked together to form molecular tracks over time. The x,y coordinates of each centroid may then be taken to calculate the mean squared displacement (MSD) for each molecular trajectory, enabling apparent diffusion analysis from the MSDs to give mobility distributions of tracked molecules in a live cell.
Adapted from Anne Plochowietz, 2015; Sanamrad et al., 2014; English et al., 2011.

After taking the high-throughput acquisitions from widefield imaging with SMLM, further computational processing of localized particles can be used to link consecutive frame localization to generate single particle trajectories, and subsequently calculate particle mobility via their mean-squared displacement over time and comparison to apparent diffusion coefficients (Anne Plochowietz, 2015)(Figure 3.1B). Note that there are several ways of processing localizations (Deschout et al., 2014) and tracking (Chenouard et al., 2014; Michalet & Berglund, 2012), and the optimal choice is dependent on the experimental design and purpose.

3.1.2 Fluorophores light the way

Earlier *in vivo* single molecule exploration predominantly relied upon genetically encoded FP fusions for studying the expression, regulation and dynamics of labelled proteins in the live cell, and FP fusions indeed remain a tenet of single molecule biomolecule labelling methods today (particularly in PALM). Encoding a FP adjacent to the target protein of interest, the cell

endogenously generates a FP-labelled protein, ideally expressed at physiological rates. However, the large molecular sizes of the FP labels – such as the 27 kDa GFP, a favourite – and typical C-terminus fusions have been shown to be able to perturb target protein function (Rodriguez et al., 2017), and likely impact mobility such that exact physiological representation is uncertain.

Organic dyes, on the other hand, are significantly smaller (on the scale of <1.5 kDa for typical conjugates), and ~10x brighter and more photostable than FPs (Dempsey et al., 2011; Fernández-Suárez & Ting, 2008). Although recent FP developments have generated ever more photostable and bright FP options (Bajar et al., 2016; Saurabh et al., 2016), combining the likely perturbational effects of large FPs on their target with the more generally higher brightness and smaller size of organic dyes make these small synthetic fluorophores an ideal choice for systems looking to transiently label small biomolecules.

3.1.2.1 Fluorescence – an overview

As fluorescent signals from fluorophore-labelled biomolecules are central to the interpretation of the labelled biomolecules' behaviours and spatiotemporal properties, the photophysical characteristics of specific fluorophores are paramount to the experimental design.

Fluorescence is defined as the emission of photons, in this case from a fluorophore, on the absorption of different wavelengths of light, when electrons fall back to ground state from an excited state. These excitation-emission wavelengths shifts can be displayed as spectra for a fluorophore of interest (Figure 3.2, excitation-emission spectra for ATTO647N). Highlighting the Stokes shift as the wavelength separation between the excitation-emission spectra peaks enables appropriate selection of spectral overlaps, for example, to avoid spectral overlap in dual-colour systems.

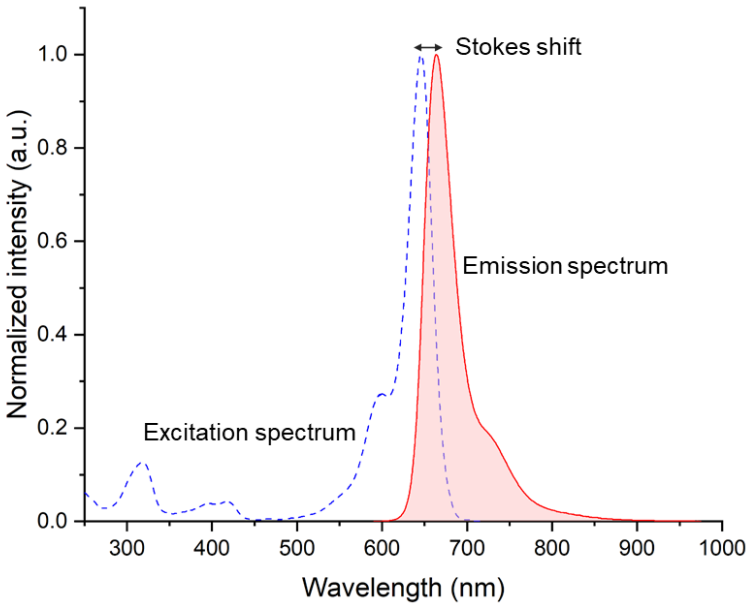


Figure 3.2 – Excitation-emission spectra for ATTO647N.
Normalized intensity of the excitation spectrum in blue, emission and hence fluorescence signal in red; Stokes shift as the wavelength distance between absorption (excitation) and emission peaks: 25 nm for ATTO647N (Kolmakov et al., 2010).
Spectral data obtained from Biomers.net GmbH.

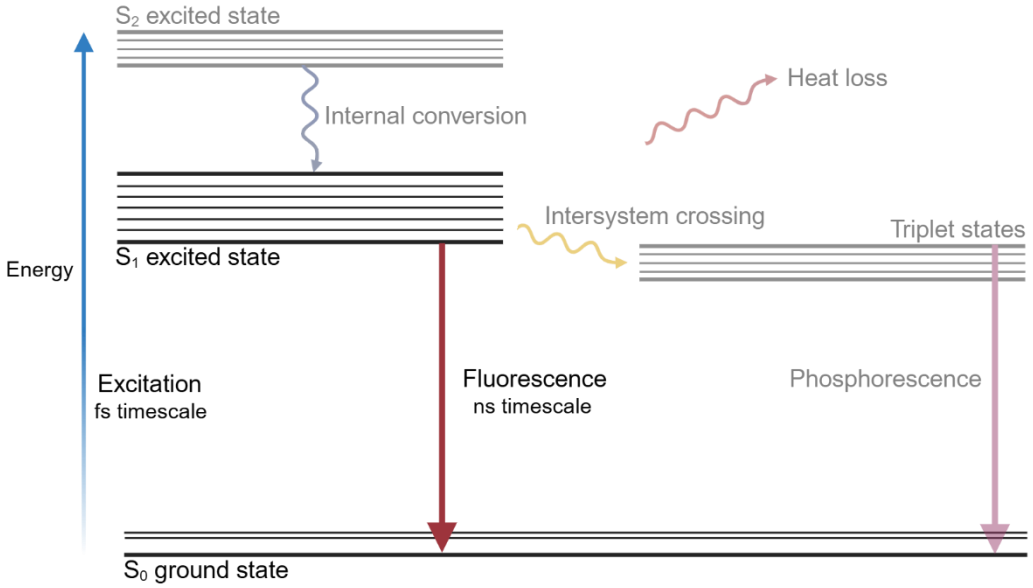


Figure 3.3 – Jablonski diagram showing ground (S_0), excited singlet states ($S_{1,2}$) and mixed triplet states, alongside possible transitions and expected timescales for optimal fluorescence conversion ($S_0 \leftrightarrow S_1$).

The exact quantum states and likely transitions occurring in a fluorophore excitation-emission cycle, and hence fluorescence observed, can be theoretically depicted using the Jablonski diagram (Figure 3.3). Fluorophores may hold either singlet or triplet quantum states; optimally for single molecule observations, a fluorophore would solely exhibit excitation to its singlet state, de-excitation and falling back to ground state to give a 100% quantum yield (fluorescence signal) on the timescale of a few nanoseconds. However, intermediary transitions into triplet state, and higher vibrational singlet states are also possible, both of which deplete fluorescent quantum yield and can occur on longer timescales.

As such, the ideal fluorophore for single molecule experiments would have high quantum yield (bright fluorescent signal), low fluorescence lifetimes (reducing likelihood of alternative states), a wide Stokes shift (reducing spectral overlap and hence signal clarity) and high photostability (reducing photo-bleaching propensity to give longer observations)(Markus Sauer et al., 2011).

3.1.2.2 Transiently binding organic dyes

As alluded to above (3.1.2), chemical dyes cannot be synthesised in cells, and hence must be conjugated to either the target of interest, or a complementarily hybridising biomolecule. As such, pursuant to *in vitro* labelling – giving a higher level of specific site-labelling control and higher efficiency for doubly-labelled constructs in particular – the dye-labelled biomolecules must be exogenously introduced into cells. Whilst there are numerous means of internalising labelled biomolecules into live cells, extensive work by the Kapanidis group has demonstrated that electroporation gives appreciable internalisation efficiencies (Aigrain et al., 2015; Plochowietz et al., 2014, 2017; Sustarsic et al., 2014); thus electroporation was selected as the method of internalisation for short dye-labelled ssDNAs here.

Of the different chemically structured organic dyes widely available, all classes (cyanines, rhodamines, carbopyronines, and oxazines) fit the general optimal requirements for a suitable fluorophore for single molecule fluorescence studies (Dempsey et al., 2011; Ha & Tinnefeld, 2012). Considering photostability, however, the radical pushing of fluorophore lifetimes in an *in vitro* setting via careful chemical stabilization (Widengren et al., 2007) is not so readily conducive to live cell systems, where many such stabilising methods would be cytotoxic. As such, suitable *in vivo* fluorophores need to have high innate photostability.

ATTO647N is one such choice fluorophore: a dye in the red spectral region with high quantum yield, high thermal and photostability (pH-independent)(*Atto 647 N, Sigma-Aldrich*), ATTO647N has been shown to enable long acquisition times apt for single molecule tracking

(Hovius et al., 2006). Noted, the lack of structural protection as conveyed via polypeptide folding in FPs increase the off-target sticking propensity of the organic dyes, and ATTO647N does indeed show non-specific sticking to cell membranes (Zanetti-Domingues et al., 2013). Recent developments have led to the ATTO647N relative, ATTO643 – showing a similar spectral range, high quantum yield and stability to ATTO647N, but with a lesser propensity for unspecific binding (*ATTO 643*, ATTO-TEC) due to its structure-conveyed high hydrophilicity. These two red dyes are the primary fluorophores selected for use across experiments in this body of work.

3.1.2.3 Fluorescence quenching and fluorogenicity

Uniting the premise of FISH with the SMLM based on PAINT for the transient labelling of short ssDNAs *in vivo*, an *in vivo* smFISH probe would ideally host a bright, highly photostable and small fluorophore, which only fluoresces on specific complement binding, i.e. to a target RNA.

FISH is reliant upon complementary hybridisation of a labelled ssDNA probe with its DNA or RNA target of interest. In PAINT, the same premise of transient hybridization is applied in an SMLM context, whereby cumulative single molecule localisations are achieved through temporally-segregated fluorescence signals (as opposed to photophysical properties as in other SMLM techniques)(Lelek et al., 2021)(Figure 3.4).

As such, as much as the very bright and photostable ATTO647N/ATTO643 are ideal fluorophores for *in vivo* applications, they won't do on their own; a process by which to selectively limit fluorescence emission is required: quenching. Quenching may be seen between two of the same (self-quenching)(Kessler et al., 2023, 2024) or different molecules, and broadly categorise into static or dynamic quenching processes (Lakowicz, 2006b).

Briefly, static quenching denotes ground-state complexation such that excitation to emit fluorescence cannot occur (Lakowicz, 2006a). Dynamic quenching occurs through the interaction of a quenching molecule with a fluorophore in its excited state – subsequently open to actuation through a variety of interactions, including collisional processes, intermediary transitions to triplet state, and fluorescence resonance energy transfer (FRET) targeted transferral of excited electrons onto the quencher molecule, where energy is released as heat (Lakowicz, 2006a).

ATTO647N has been shown to exhibit self-quenching properties, thought to occur via possible contact-mediated interactions at very short lengths (≤7 nt)(Schröder et al., 2019), but the exact mechanism is unknown. As for ATTO643, effective FRET-quenching can be achieved via

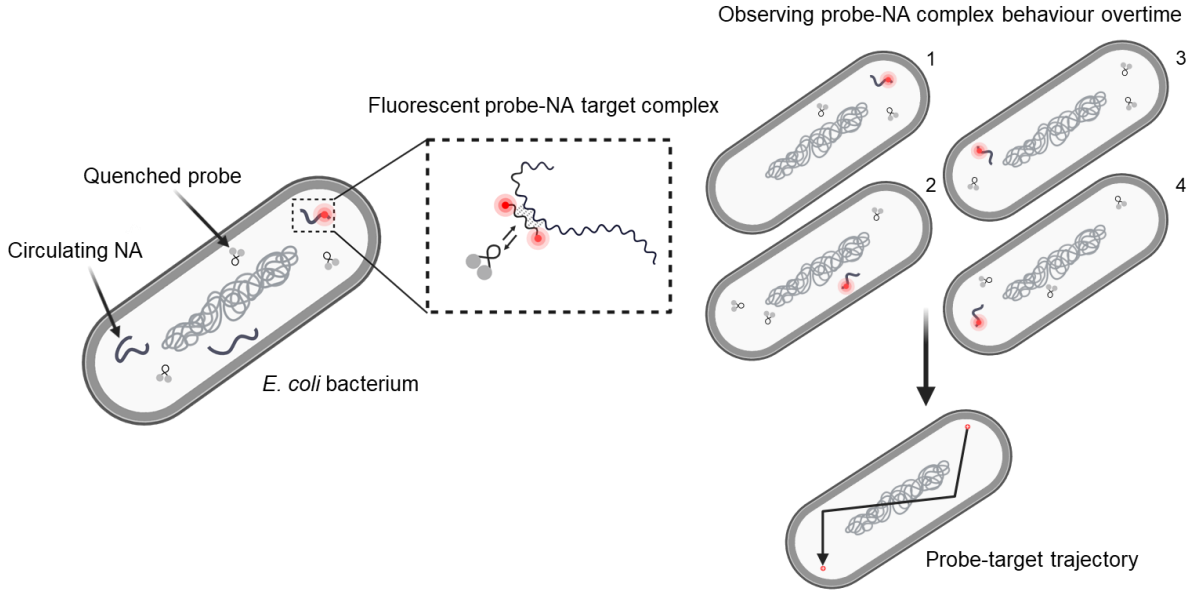


Figure 3.4 – Schematic depicting the combination of live cell FISH and PAINT-SMLM premises for the application of live single-particle tracking of nucleic acids. Transiently hybridising fluorogenic probes to target nucleic acids can be localised and tracked over time to obtain single molecule spatiotemporal behaviours.

complexation with commercial non-fluorescent quenchers, in this case, by the Black Hole

Quencher family (BHQs; *BLACK HOLE QUENCHER™*). The BHQs are well established in a number of systems studying molecular conformational dynamics *in vitro* (DNA hairpins (Schwartz & Quake, 2009), ribosome-tRNA interactions (J. Chen et al., 2012), viral genome segment structure (Kümmerlin et al., 2025)). There are significant advantages of applying FRET-quenching systems: efficient energy transfer in FRET can be selected for by electing different spectral overlaps with a fluorophore of choice, and the non-fluorescent nature of the BHQs improve SNR (lesser background)(J. Chen et al., 2012). Both advantageously combine to produce brighter localisable signals on de-quenching, ideal for SMLM.

Although there is no published documentation on BHQs for quenching in live bacterial cells, the use of fluorogenic probes for taxonomic categorisation of pathogenic bacteria has been demonstrated in live cells (Franzini & Kool, 2011). As such, fluorogenic probes may be generated by doubly-labelling short ssDNAs with 2xATTO647N, or with a pair of ATTO643-BHQ1 dyes.

3.2 Engineered *in vivo* smFISH: an origin story

The golden chalice of the capability to directly observe, localise and track target RNAs, to resolution of a single molecule, in live and healthy cells has many significant implications, as reviewed in the Introduction. Notably, combining fluorescence/genetic engineering technological developments with single molecule resolution capacity greatly expands the current repertoire of gene targets accessible for RNA exploration with minimal intracellular perturbation, across tangible cellular lifecycles. Whilst there has been much innovation around single RNA detection in live cells, most current methods (MS2 fusions, fluorogenic aptamer systems, CRISPR/Cas9 systems) for live-cell imaging of circulating transcripts enable single molecule resolution, but so far, without reliable single fluorophore resolution. With the single exception of recent developments by Doudna and co. around the CRISPR/Csm system in eukaryotic cells (Xia et al., 2025), achieving single molecule resolution for *any* target RNA in an equivalent prokaryotic system has yet to be realized.

3.2.1 *In vivo* single molecule FISH via short fluorogenic oligos – the premise

The conceptual premise underlying engineered *in vivo*, single molecule FISH (*in vivo* smFISH) with fluorogenic probes follows as such: in a cell with circulating, sequence-specific fluorogenic probes, when there is no transcription of the target gene and hence no circulating target mRNA, there is no probe-RNA hybridisation and signal (Figure 3.5A). In the instance of transcription, releasing target transcripts into the cell milieu, freely-diffusing fluorogenic probes may then specifically hybridise to their target mRNA, de-quench and fluoresce.

Continuing the above concept of specific hybridisation of quenched probes to their target RNA sites in the live cell, a transcription life cycle of hypothetical observations can be mapped (Figure 3.5B):

1. On initial probe internalisation, when transcription of the target gene has yet to take place, there are no complementary transcripts for probe binding – therefore no dequenching and no fluorescence signals observed.
2. On transcription initiation and the progressive synthesis of the complementary RNA, probes can transiently bind and de-quench – visualised as largely immobile fluorescent signals as the locus of hybridisation is contained to the RNAP-DNA open complex and nascent RNA.
3. On release of the target RNA following transcription termination, the target RNA may rapidly diffuse through the cell, any constraint possibly related to transcript-specific compartmentalisation or complexation – this could be observed as fast-diffusing fluorescent spots.
4. On degradation of the target RNA by the degradosome or other cellular degradational systems, the RNA-probe binding sites are sequentially lost – leading to the rapid loss of fluorescence signals.

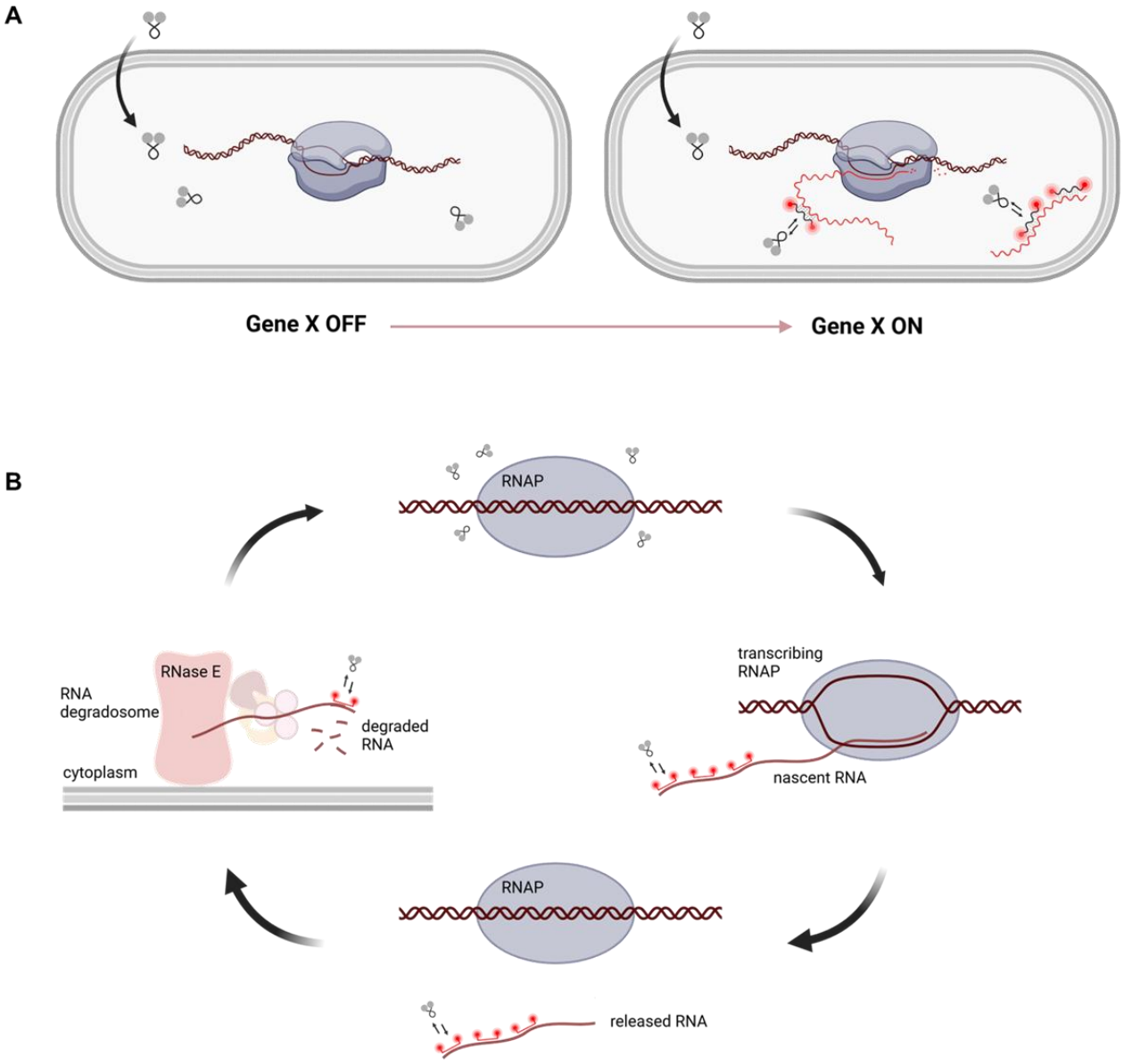


Figure 3.5 – Schematic depicting the hypothesised fluorescence observations in specific tagging of target RNAs in live *E. coli* using fluorogenic, short ssDNA probes.
(A) Without the presence of the target RNA in the cell, circulating fluorogenic probes remain quenched; when nascent target RNAs and released target RNAs are present in the cell milieu, fluorogenic probes may complementarily hybridise and de-quench to generate an observable fluorescent signal.
(B) Hypothetical visual possibilities at each stage of a transcript's life cycle with complementary and specific probe-RNA hybridisation. Translation can occur during nascent transcription and during transcript release via (poly)ribosomal binding (omitted for simplicity on the schematic).
Not to scale.

Here, transcript hybridisation specificity is enabled by engineering a probe-complementary "tag cassette", upstream of the target gene transcription start site and downstream from the gene promoter (in the 5' untranslated region, UTR). The "cassette" refers to the arraying of probe-specific sites, such that there may be any theoretical number of sites available for multiple probe hybridisation simultaneously – counterbalancing the expected transient hybridisation time with a confluent signal over longer periods generated through multiple RNA probe-site occupancy at any one time. This design relates to the arraying of tandem sites, in MS2 systems for example, by way of amplifying fluorescent signal, but without substantial size and mass burden of a long additional sequence due to the short probe and intercalating spacer sequences.

As such, this engineered *in vivo* smFISH method is technically as tractable as the capacity to engineer the tag cassette into any promoter region of interest, in a chosen prokaryotic model – with the *E. coli* genome considerably well-described, a promising prospect.

3.2.2 Continuing methodology development

With a design rationale of single molecule fluorescent foci upon complementary hybridisation of short, fluorogenic probes to mRNA, the consistent and reliable internalisation of the exogenously supplied fluorogenic probes into the live cell is a basic component.

Building on robustly developed methodologies established in the Kapanidis group for the internalisation of biomolecules into live *E. coli* (Crawford et al., 2013; Sustarsic et al., 2014), the primary experimental process applied is as documented in Aigrain et al., 2015.

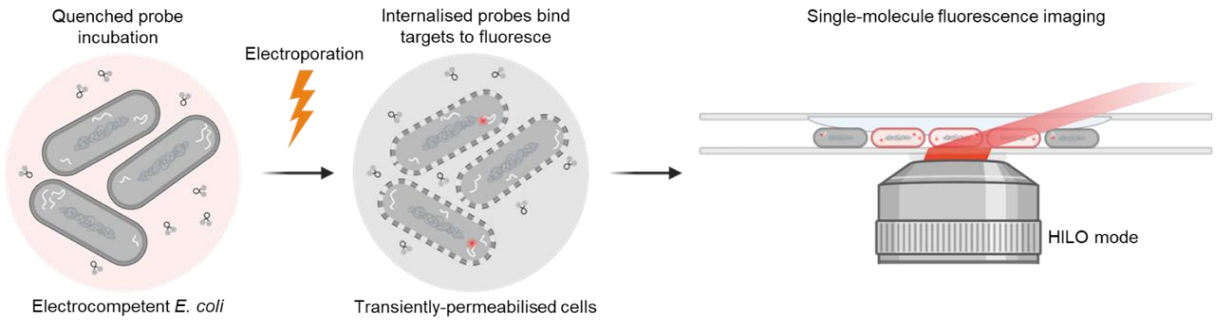


Figure 3.6 – Schematic of the experimental pipeline of electroporated-internalisation of labelled ssDNA probes and live-cell imaging. Live *E. coli* cells are incubated with quenched fluorogenic probes in solution; electroporation enables internalisation of probes, allowing for complementary probe-target hybridisation and de-quenching (shown in red). Widefield imaging in HILO mode using a 638 nm laser allows for fluorescent signal acquisition from probe-target complexation in live cells.

Briefly: *E. coli* cells are made competent; extracellular fluorophore-labelled biomolecules are introduced; an electric shock is applied across the cell-biomolecule mixture (electroporation) to allow a few probes to enter cells; electroporated cells are recovered; and finally plated on agarose pads to enable superresolution imaging (Aigrain et al., 2015) (Figure 3.6). Having demonstrated the ability to internalise a diverse range of biomolecules including to entire protein of ~100 KDa, this methodology has subsequently been applied to explore the spatiotemporal behaviours of exogenously labelled tRNAs in real time in live *E. coli*, showing that tRNA distribution was translation-dependent and predominantly ribosome-associated, giving further indication of ribosomal cellular distributions (Plochowietz et al., 2017).

With the electroporation of fluorophore-labelled biomolecules as a foundation, this methodology was tested and further optimised for the purpose of internalising very short labelled ssDNAs (≥8 nts), particularly towards lower copy number and more transient targets: intracellular mRNAs.

3.2.3 Preliminary findings

The conceptual premise of engineered *in vivo* smFISH as elucidated in 3.2.1 was first tested by Dr Rebecca Andrews and Dr Hafez El-Sayyed, using short, self-quenching DNA oligonucleotide probes, doubly-labelled with the organic dye, 2xATTO647N. Originally developed for Gap-Seq (Andrews et al., 2022), these probes were trialled in *E. coli* strains engineered by Dr Hafez El-Sayyed, as per their fluorogenicity and short size for transient binding to dedicated complementary sites. Dr Andrews and Dr El-Sayyed preliminarily showed localised fluorescent signals in live cells on electroporation of the 2xATTO647N ssDNA probes, in both plasmid-system and chromosomal-system strains generating complementary RNAs via plasmid or chromosomal transcription respectively.

Henceforth, the basis of this project was founded – the aims being to:
i. establish a robust system for localising single target mRNAs in real time via exogenously-delivered fluorogenic ssDNA probes; and
ii. to extend this methodology in probing the transcription life cycle of theoretically any gene of interest in the live *E. coli* cell.

3.3 Chapter summary

In this Chapter, existing K-12 *E. coli* strains comprising probe-binding arrays via plasmid and chromosomal systems were characterised *pro* proof-of-concept, and further K-12 MG1655 strains were designed and developed, including a co-localisable plasmid system.

Extensive fluorescence SRM was applied via *in vivo* smFISH to determine the specificity of the ground methodology, and parameters for optimisation. Comparing probe signal in wildtype cells to signal in the induced plasmid and chromosomal tag strain systems indicated majority signal specificity of probe-RNA target hybridisation, requiring further confirmation by single molecule analysis. Transcription specificity was probed and largely confirmed via perturbing transcription elongation using rifampicin. Hitherto partially established foundations of engineered *in vivo* smFISH, technical aspects of the methodology were explored to optimise for internalisation of exogenous probes and for signal specificity. In conducting technical optimisation, a notable observation of passive probe internalisation, without electroporation, revealed a potential alternative route for future *in vivo E. coli* single molecule experimentation.

3.4 Recapitulating preliminary findings: *in vivo* smFISH

3.4.1 Engineering systems to evaluate probe-RNA hybridisation and probe transcription

To test the initial feasibility of visualising single RNAs in live cells using transient binding of very short fluorogenic ssDNA probes (8mers), artificial inducible plasmid and chromosomal systems were designed to generate target RNAs with 78-nt tag array sites for probe hybridisation (designed and constructed by Dr El-Sayyed). These tag-carrying strains comprised downstream transcribable but not translatable DNA sequences under the control of an arabinose-inducible promoter and a strong terminator.

Figure 3.7A presents a schematic of the tag-array construct, where 6x 8 nt probe binding sites were designed with spaced intervals (spacers) of 5x 6 nt random repeated sequences, alongside the actual sequence of 2x sites and corresponding spacers (DNA, RNA, Figure 3.7B). Once transcribed from the DNA, the RNA sequence and 6x sites are theoretically available for transient 8mer probe binding and de-quenching to give fluorescent signal. With 6 possible sites, there is a higher likelihood of occupancy of any site despite possible secondary structures, steric occlusion and unknown binding kinetics by the probes.

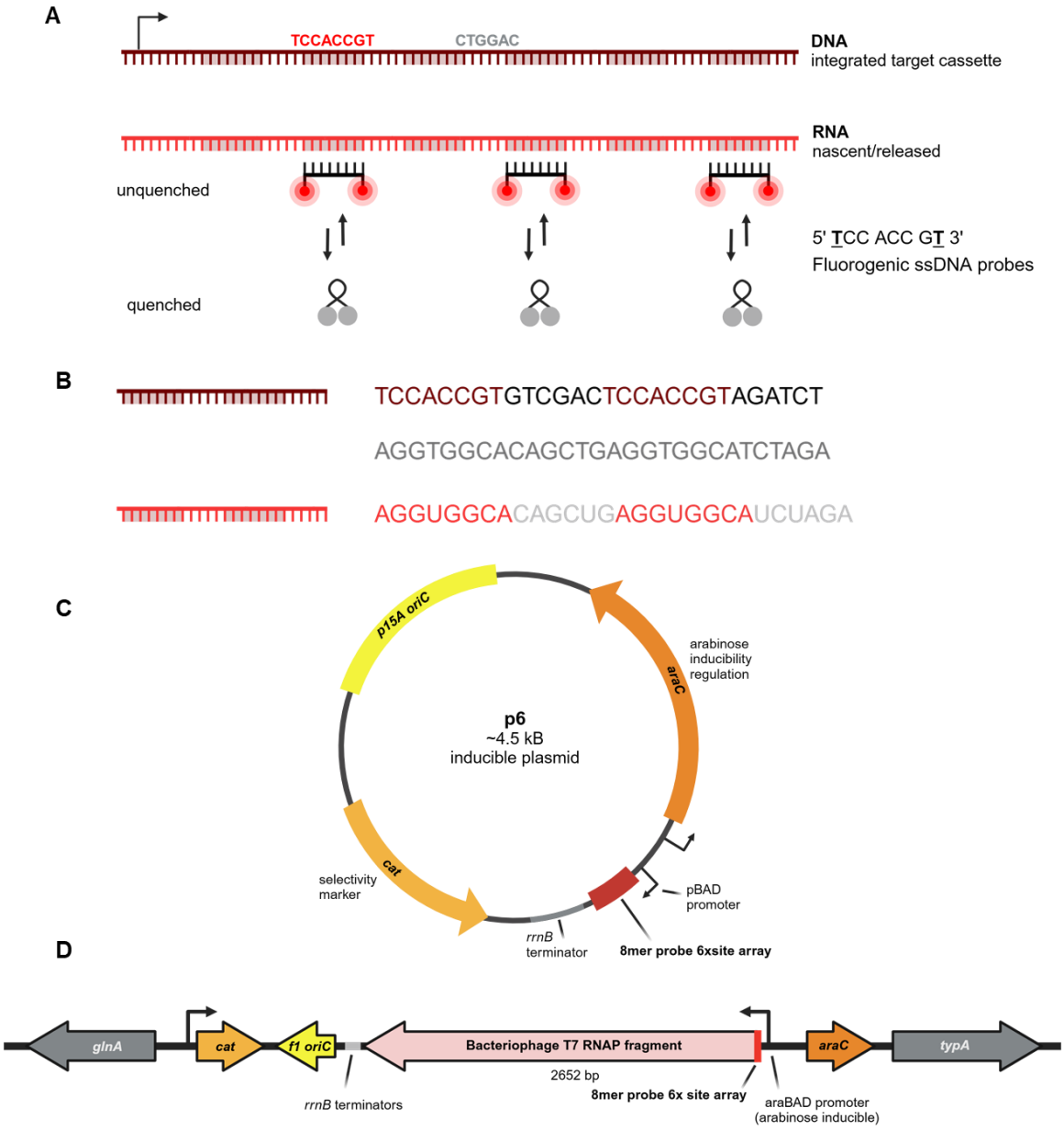


Figure 3.7 – Schematic of the 6-site array for complementary probe hybridisation, and whole strain constructs for p6 and gT7 tag *in vivo* smFISH strains.
(A) Representation of DNA cassette, parallel RNA output and transient on-off binding of 8-nt fluorogenic ssDNA probes, where probe hybridisation to its complementary RNA site leads to dequenching and fluorescent signal. The 6x 8-nt sites are separated by randomized 6-nt spacers.
(B) Two-site DNA sequence from the constructs, and corresponding RNA.
(C) p6 construct design displaying the core 6-site array with an arabinose-inducible promoter and strong terminator, alongside the entire plasmid construct transformed into cells (chloramphenicol-selective).
(D) Representation of the gT7 strain chromosomal insert – where a 2652 nt bacteriophage T7 RNAP fragment (non-native to *E. coli*) is inserted under the control of the same arabinose-inducible promoter and strong terminator as in p6, selectable by chloramphenicol. The ~7 kb insert is site-specific, between native *glnA* and *typA* gene loci.
Not to scale.

The p6 plasmid construct was first designed, to generate high copy numbers of short 128 nt RNA targets comprising a tag array of 6 complementary sites for probe hybridisation (Figure 3.7C). A classical expression pET-28a vector with a p15A origin (low-intermediate copy number) was chosen to give 1-10 plasmids (Sathiamoorthy & Shin, 2012; B. Shao et al., 2021) per cell with localisation preference for the cell poles – with the aim of producing a high transcript copy number, localisable and most visible at the poles. The araBAD operon promoter was chosen to enable induction of plasmid expression via the addition of L-arabinose during growth (Guzman et al., 1995), as a means of further augmenting high transcription of the probe target site and fragment: in the presence of arabinose, pBAD transcription is activated by a conformational change in the transcriptional regulator AraC (Lee et al., 1987). No ribosomal binding site (RBS) is included, so minimal transcript interaction with translational machinery would be anticipated, and strong transcription termination would be exerted by the well-characterised *rrnB* terminators (Brosius et al., 1981; Orosz et al., 1991). Selectivity for successful transformation of recipient cells was conferred by chloramphenicol resistance (via *cat*).

The gT7 chromosomal construct followed the same premise: the insertion of a cassette comprising a chloramphenicol resistance gene, arabinose inducibility augmented by the *araC* gene, and the 6x 8mer tag array inserted at the 5' of a 2652 bp bacteriophage T7 RNAP fragment, under the control of pBAD and the *rrnB* terminators (Figure 3.7D). An overall ~7 kb cassette, the construct was site-specifically inserted via Lambda-red recombination to sit between the *glnA* (Magasanik, 1989) and *typA* (Freestone et al., 1998) gene loci. Approximately 132 kb away from the genomic origin, this locus would be unlikely to affect essential pathways, and could be further probed via changing growth conditions as *glnA*/*typA* expression is regulated by nutrient availability. The T7 phage RNAP gene sequence is particularly well-characterized (Tabor & Richardson, 1985), serving as a useful longer target sequence associated with the tag array, as a non-native RNA with no downstream translation without the inclusion of a RBS.

Evaluating the potential for probe spectral and steric interactions, the 6-site tag array was adapted to incorporate spacers of 15 bases (Figure 3.8). To maximise the likelihood of effective transcription of the array and attached target gene, each spacer was designed as a random nucleotide sequence, giving no identity between spacers. Whilst tandem repeats are naturally-occurring in *E. coli,* functionally regulating various aspects of gene expression (Stern et al., 1984; Suenaga et al., 2022; Yoshida et al., 2000), the construction and incorporation of highly repetitive sequences in molecular cloning remains a perennial challenge.

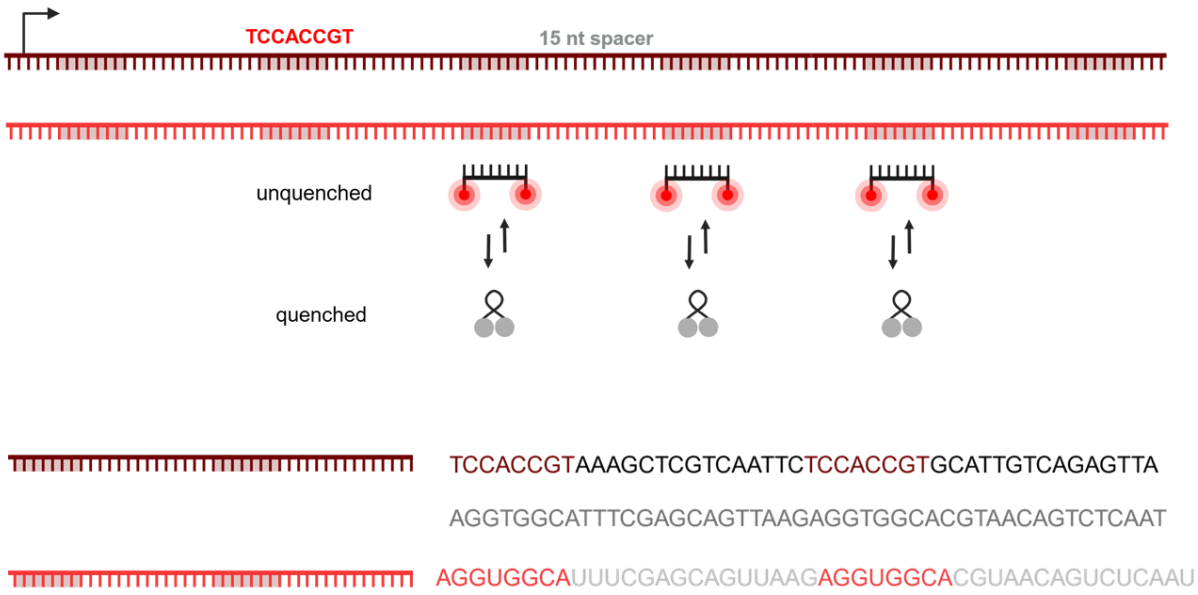


Figure 3.8 – Schematic of the 6-site 8mer array with 15 nt spacers for complementary probe hybridisation.
Representation of DNA cassette, parallel RNA output and transient on-off binding of 8-nt fluorogenic ssDNA probes with 15 nt randomized unique spacers (upper), alongside two-site DNA sequence from the constructs and corresponding RNA (lower).
Not to scale.

As such, alongside the increased spacer length, each consecutive spacer sequence was fully randomized; spacer sequences were chosen when sequence complementarity in the K-12 genome had a hit rate between 0-3 identities. Cross-comparison with the probe sequence, TCCACCGT, showed up to 60% binding potential with the spacers but with no more than 3 consecutive nucleotides at any one time – thereby deemed suitable. Whilst RNA secondary structure predictions were taken into account – observing some propensity for stem-loop structures between repetitive sequences – it was noted that the predictive tools were unlikely to reliably recapitulate structures and circumstances that would relate to an *in vivo* context (Bliss et al., 2020; Leamy et al., 2016), and therefore were not weighted determinants.

There are also 150 sequence hits in the K-12 genome for the TCCACCGT probe sequence. Of these hits, 57 are in transcribed gene regions and could hence act as specific RNA off-targets, enabling probe binding and dequenching. Probe binding to complementary DNA sequences is highly unlikely.

Noted – the 5' UTR of specific gene loci often comprise regulatory sequences involved in translation initiation and mRNA stability, modulating mRNA accessibility and degradation rates. Whilst ribosomal binding will not be affected in the p6 and gT7 tag strains, where no

target mRNA translation occurs, it is possible that a 5' UTR insertion in native gene loci could impact downstream mRNA processing, stability and degradation rates.

3.4.2 Investigating mRNA tagging in plasmid and chromosomal systems

In attempting to recapitulate Dr Andrews' findings that the electroporation-mediated internalisation of labelled DNAs, particularly the 2xATTO647N 8nt (8mer-2x647N) probe complementary to its RNA hybridising sequence, led to observable fluorescent foci in tagged strains, similar observations were indeed realized in electroporating tagged cells with the same 2xATTO647N probe.

M9Glu competent, arabinose-induced p6 and gT7 BW27784-background cells were electroporated with 50 nM of 8mer-2x647N, and imaged respectively in HILO at ~10 mW 638 nm using 50 ms exposure. Figure 3.9A presents maximum intensity projections of representative fields of view (FOVs) for both p6 and gT7 cells with 50 nM probes, alongside representative FOVs of the same cell types without the probe (electroporated with upH_2O). In the absence of the probe, neither strain showed signal, confirming that fluorescence observed was specific to probe de-quenching in the region of the cell. Due to the significantly higher copy number of plasmid-transcribed mRNAs, it was anticipated that p6 cells would generally exhibit a wider spread of fluorescence heterogeneity across the cell population, such that some cells would be "flooded" with signal and some already at the level of distinguishable foci, pursuant to variable internalised probe concentrations.

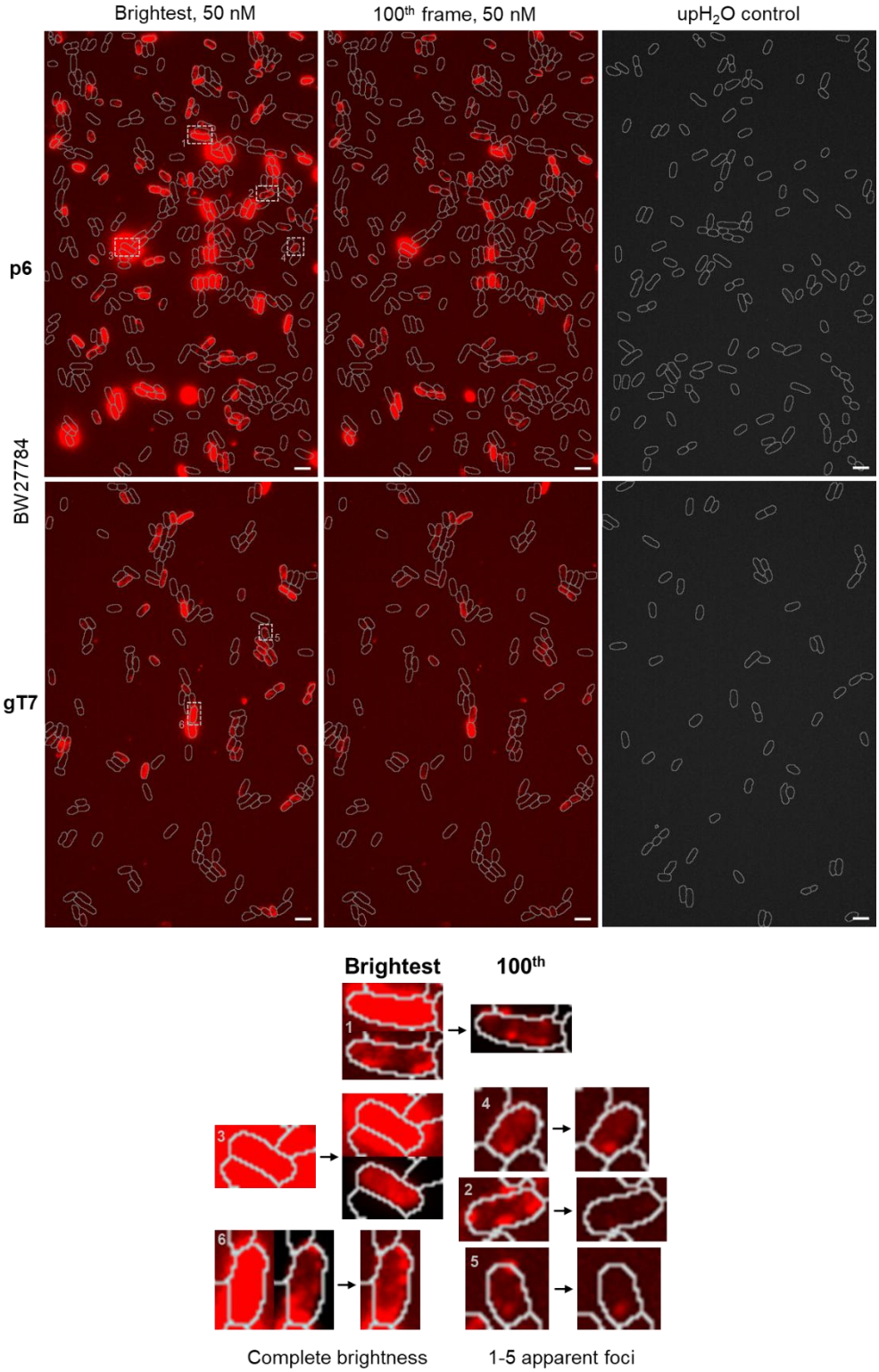


Figure 3.9 – Recapitulating fluorescence observations from probe-RNA hybridisation in live p6 and gT7 BW27784 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N: representative visuals.
(A) Maximum intensity projections of frames 1-1000 (brightest, ~50 s), 100-1000 (100th frame) of probe-electroporated cells, and of frames 1-1000 for cells electroporated with upH$_2$O, for both BW27784 arabinose-induced p6 and gT7 tag strains.
(B) Exemplar cells, highlighted as 1-6 from the brightest frame FOVs in (A), presenting cells with a range of fluorescent signal on laser activation (I_0) and 5 seconds later (I_{100}). Some cells give a "halo effect" with the appearance of a cell "flooded" with signal; others have 1-5 distinguishable foci from the start of acquisition. Where there is a higher contrast still, the same maximum intensity cell still has been adjusted from the standard brightness and contrast settings to allow for clearer observation of signal distribution.
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

Ensemble intensity data

Analysing the ensemble background-corrected mean intensities of cells from both strains classified as 'loaded' with 50 nM probe from the brightest frames (I_0) and the 100th frames (I_{100}) of video data, the p6 cells appeared to have a marginally higher median mean intensity per cell area than the gT7 cells (I_0: 21.5 units p6; 13.4 units gT7; I_{100}: 6.9 units for p6, 5.5 units for gT7) (Figure 3.10A). This is likely explained by a greater population of p6 cells with any signal over gT7 cells, and more pertinently, the propensity for much higher copy number of target sites for probe hybridisation in p6 cells. A higher p6 copy number would be anticipated due to the transcription rapidity of a ~128 nt sequence, the accessibility of plasmids for circulating RNAPs and the possibility of multiple plasmid copies per cell.

Whilst gathering intensity data from I_0 (i.e. first full laser activation of the cells) is a good indicator of probe uptake heterogeneity and a potential proxy for cell phenotypic heterogeneity across a population, it was reproducibly observed that some cells were extremely bright – a kind of 'halo effect' extending beyond the cell boundaries was observed (exemplar cells, lefthand side in Figure 3.9B; righthand side, cells with distinguishable foci). An imaging artefact corresponding to overexposure as well as diffraction effects from the probes in the cell periphery, this is most likely explainable by high intracellular transcript load and complementary probe binding.

There is a possibility of ATTO647N sticking to the cellular membrane, leading to spurious signal, although this is unlikely to substantially contribute to such observations. Observations of very bright cells were nonetheless consistent at I_0, and could be confounding in calculating mean intensity distributions, particularly with dense clusters of cells. As such, calculating the mean intensities per cell area at I_{100} of each video was elected as a more sustainable means of comparison, based on the consistent observation that ~5 seconds was sufficient to photo-bleach down most signal to leave probable probe-mRNA complexes, and hence specific fluorescent signal. Cells were classified as loaded if intensities were above zero after a subtraction of the averaged mean intensity of empty cells (cells electroporated with upH_2O) .

To observe the intensity distribution across the whole cell populations in BW27784 p6 and gT7 cells with 50 nM 8mer-2x647N, background-corrected mean intensities per cell area were plotted as histograms (Figure 3.10B). This showed that across both I_0 and I_{100}, p6 cells presented a larger spread of mean intensity per cell area values than in gT7, and that

interestingly, by far the predominant population of cells across both strains had mean intensities below 15 units. This suggests most cells internalise a concentration of probe such that distinguishable single fluorescent focus events are present from the start, which is ideal for single molecule analyses. Notably, whilst p6 cells appeared to have a slightly higher ensemble mean intensity than gT7 cells, gT7 actually appeared to have a higher rate of loading: in p6 cells, ~77% and 58.5% cells across the I_0 and 100th frames respectively were classified as loaded, compared to 88% and 74% of gT7 cells.

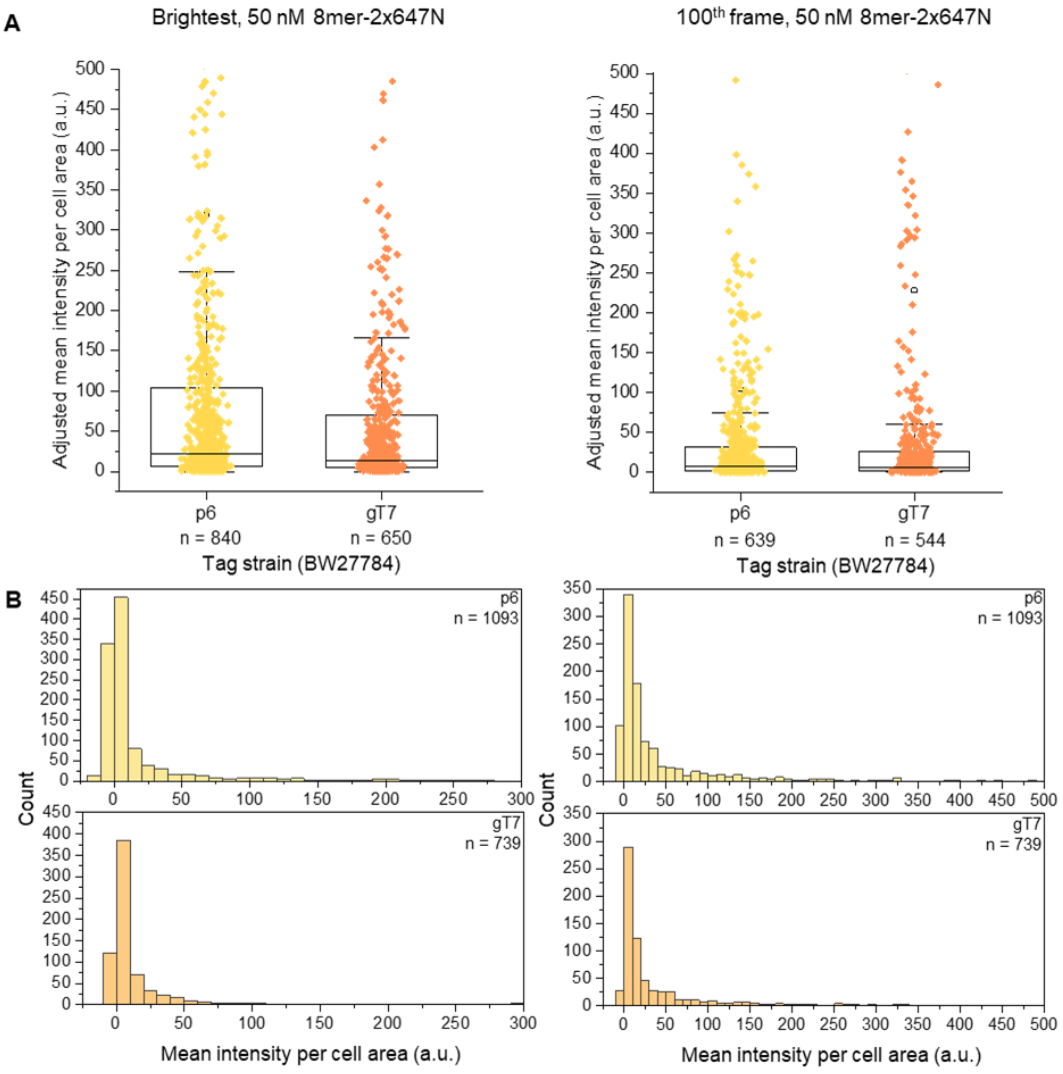


Figure 3.10 – Recapitulating fluorescence observations from probe-RNA hybridisation in live p6 and gT7 BW27784 *E. coli* cells electroporated with 50 nM 8mer-2x647N: population-level intensity distributions.
(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 50 nM 8mer-2x647N-electroporated induced p6 and gT7 cells, from I_0 and I_{100}. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
I_0 medians: 21.5 units p6, 13.4 units gT7; I_{100}: 6.9 units p6, 5.5 units for gT7.
(B) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 10.
All mean intensities per cell area shown are background corrected and n gives the number of cells per sample (N = 3, 2 experiments for p6 and gT7 respectively).

For the remainder of Chapter 3, ensemble mean intensity plots correspond to the background-corrected mean intensities of *loaded* cells and mean intensity histogram distributions to the background-corrected mean intensities of the *whole cell population*, calculated at the *100th frame*, unless explicitly specified otherwise.

Temporal representations

In further recapitulating preliminary findings, both p6 and gT7 cells with 50 nM 8mer-2x647N were demonstrated to have likely probe-mRNA hybridisation events up to 50 second timescales.

ImageJ temporal projections of representative FOVs give approximations of the duration signals are observed in cells. Taking the representative FOVs above, temporal projections of the same videos displayed signals towards the end of the video, suggesting that there was the capacity to have hybridisation events for much longer timespans, even beyond the 50 seconds imaged. This is explainable by the quenched nature of the 8mer-2x647N probe, and a likely pool of quenched probes internalised in the cell milieu (Figure 3.11A). Further corroborated by individual photo-bleaching time traces, single events are shown to be possible throughout the entire imaging time of 50 seconds, highlighted as quantized steps in mean intensity (black arrows) above background in Figure 3.11B.

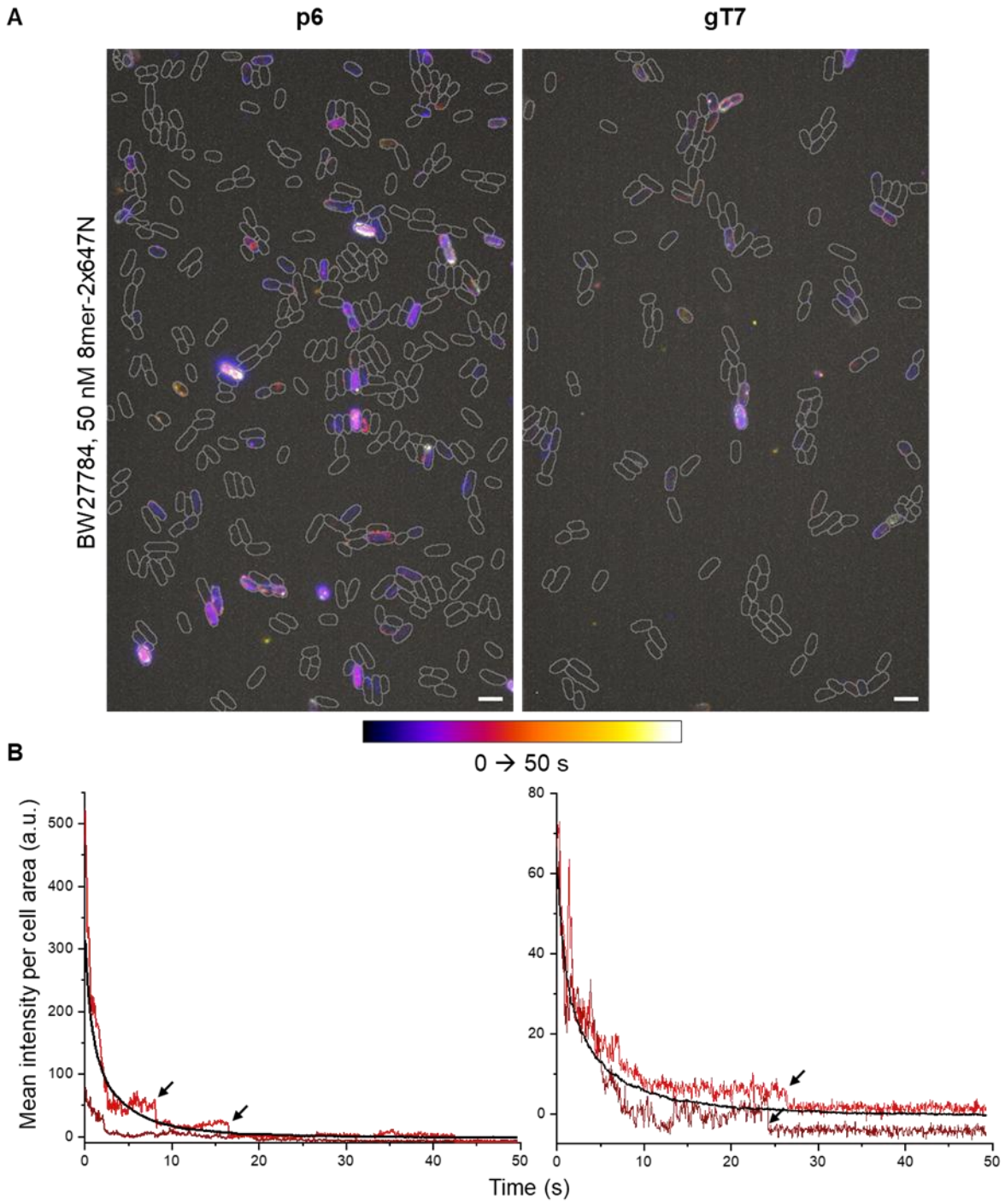


Figure 3.11 – Recapitulating fluorescence observations from probe-RNA hybridisation in live p6 and gT7 BW27784 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N: temporal visuals and photostability time-traces.
(A) Temporal projections of 1-1000 frames of the same representative FOVs of probe-electroporated induced p6 and gT7 BW27784 cells in Figure 3.9. The colour bar denotes time, where darkest is time 0 (50 ms exposure, 50 s acquisition).
(B) Corresponding averaged background-corrected mean intensity per frame time traces for the representative FOVs (black), overlaid with two exemplar bleaching time traces for individual visually-loaded cells (red shades). Black arrows denote possible quantized bleaching steps. Scale bars – 3 μm.

3.4.3 Controlling for specificity – wildtype comparisons

Having demonstrated that the tag strains do indeed give fluorescent signals that are only observable in the presence of the 8mer-2x647N probe, it was vital to determine the relative specificity of the probe de-quenching on binding its target sequence.

To explore the specificity of the 8mer-2x647N to its complementary tag sequence, AGGUGGCA, M9Glu-competent BW27784-background wildtype *E. coli* cells (without any engineering) were electroporated with 50 nM 8mer-2x647N probes and imaged under the same conditions as the cells in 3.4.2. In a perfect system, the wildtype cells with the 50 nM probe would exhibit very low fluorescence, much like the probe-absent control. However, accounting for the multiple possible states of the probe itself – partial de-quenching, imperfect labelling efficiencies – and the fact that the short TCCACCGT sequence means that there 57 potential off-target RNA sites, there was indeed variable fluorescence observed in wildtype cells (Figure 3.12A). Nevertheless, without arrayed targets and with unknown copy numbers of each of these off-targets at any one time, it is likely that overall fluorescent signals in the control strain would be shorter and more transient.

Ensemble intensity measurements

Comparing the mean intensities per cell area of competent BW27784 wildtype cells with 50 nM probes from the I_0 and 100^{th} frames alongside the same plots for p6 and gT7 cells, the median mean intensity per cell area for wildtype cells was certainly lower (I_0: 8.7 units, 100^{th}: 3.4 units)(Figure 3.12B). Interestingly, classifying loaded cells using the same method – non-zero intensities after a second subtraction for averaged empty cell mean intensity – ~72% and 68% wildtype cells were classified as containing fluorescent signal.

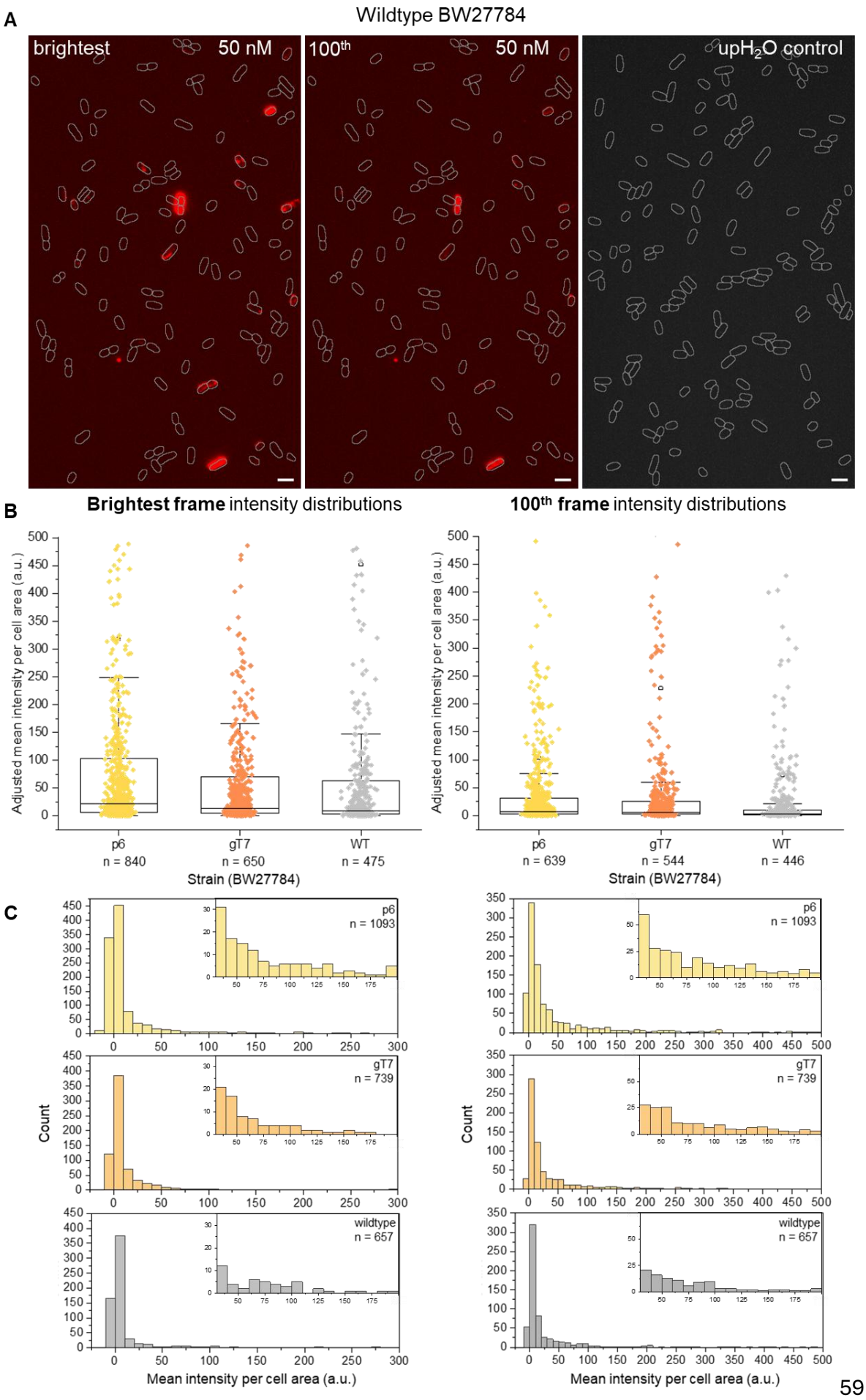

A
Wildtype BW27784
brightest
50 nM
100th
50 nM
upH2O control
B
Brightest frame intensity distributions
100th frame intensity distributions
Adjusted mean intensity per cell area (a.u.)
p6
n = 840
gT7
n = 650
WT
n = 475
Strain (BW27784)
p6
n = 639
gT7
n = 544
WT
n = 446
Strain (BW27784)
C
Count
p6
n = 1093
gT7
n = 739
wildtype
n = 657
Mean intensity per cell area (a.u.)

Figure 3.12 – Specificity control in recapitulating fluorescence observations: wildtype BW27784 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N: representative visuals and population-level intensity distributions.
(A) Maximum intensity projections of frames 1-1000 (brightest, ~50 s), 100-1000 (100th frame) of probe-electroporated cells, and of frames 1-1000 for cells electroporated with upH_2O for wildtype BW27784.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 50 nM 8mer-2x647N-electroporated wildtype cells compared with the 3.5.2 plots for induced p6 and gT7 cells, from I_0 and I_{100}. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Wildtype loaded cell medians: 8.7 units I_0, 3.4 units I_{100}.
(C) Comparative histogram distributions of the whole cell populations (where unloaded classify with sub-zero values) between wildtype, induced p6 and gT7 cells; bin size = 10.
Scale bars – 3 μm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected and n gives the number of cells per sample (N = 3 for wildtype).

Plotting the spread of the whole cell populations, most cells had mean intensities up to 15 units at I_{100}, as in the p6 and gT7 cells, but notably fewer wildtype cells with intensities above the 100-unit mark compared to the tag strains (Figure 3.12C). This suggested that one way of ensemble thresholding tag cells for specific signal could relate to higher probe internalisation and observable signal. Note that the histogram distribution at I_0 shows some very high intensities, but this tail is all but gone in the equivalent I_{100} plot, further exemplifying the observation that high initial signal often photobleached within the first 5 seconds of acquisition.

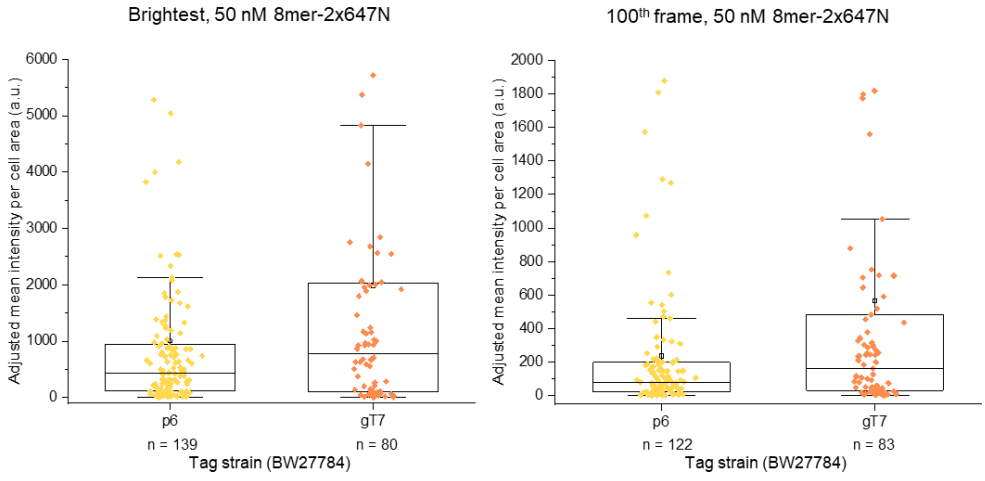


Figure 3.13 – A proxy for specific signals in BW27784 tag strains electroporated with 50 nM8mer-2x647N probes: populational mean intensity distributions for wildtype-corrected loaded p6 and gT7 cells. Mean intensities per cell with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for classified loaded p6 and gT7 cells from 3.12, further to subtraction of the averaged mean intensity of loaded wildtype cells. 50 nM 8mer-2x647N-electroporated wildtype cells compared with the 3.12 plots for induced p6 and gT7 cells, from the I_0 and I_{100}.
Wildtype-subtracted loaded cell medians: I_0 : 436.7 units p6, 781 units gT7; I_{100} 79.4 units p6, 160.8 units gT7.
n gives the number of cells per sample (N = 3, 2 experiments for both strains respectively).

As a proxy for thresholding tag cells with specific signal, loaded p6 and gT7 cells were corrected for a third time by subtracting the averaged value for the mean intensities of the loaded wildtype cells (Figure 3.13). This gave potential 'specifically-loaded' populations of 17%, 19% for p6 cells (I_0 , I_{100}) and ~12%, 15% for gT7 cells (I_0, I_{100}); with median population intensities of 79.4 units for p6 and 160.8 units for gT7 cells at I_{100}. These "loaded populations" are likely an underestimate of loaded cells with specific probe-target RNA complexes, as the wildtype cells characterise non-specific signal; hence in tag strain cells with specific binding, the subtraction overestimates non-specific interactions. Corresponding histograms are shown in Supplement 3.1.

Whilst there were likely substantial numbers of tag cells with definite specific binding events, this high overlap between wildtype and tag strain mean intensities relating to probe dequenching suggested that the ensemble identification of specific signal would be challenging. Now, considering that the goal is ultimately the exploration of single molecule behaviours, and at that, single mRNAs and their respective transcription life cycles, ensemble fluorescence signal in wildtype cells is not so problematic. With careful thresholding and visual comparison between wildtype and tag strain molecule behaviours and timestamps, individual molecules of interest can be identified as specific probe-mRNA hybridisation events, and analysed as such.

3.4.4 Controlling for transcription specificity – blocking active transcription with rifampicin

One standard tool for probing prokaryotic transcription is the antibiotic rifampicin. Rifampicin has been shown to bind at the β' subunit of the core RNAP enzyme with high affinity (Lill & Hartmann, 1973; Naryshkina et al., 2001). Conformational changes on binding leads to inhibited catalysis (Hartmann et al., 1967), such that transcription does not proceed beyond the addition of the third nucleotide. Initiation complexation of DNA, RNAP and ribonucleotides is also subsequently reduced (Mosteller & Yanofsky, 1970), although interestingly, actively elongating RNAPs bound by rifampicin can continue in their elongation (Yarbrough et al., 1976).

To investigate the transcription specificity of specific probe-RNA binding events, M9Glu competent wildtype, arabinose-induced p6 and gT7 cells were electroporated with 50 nM 8mer-2x647N probes, recovered with SOC rich medium containing 100 µg/mL rifampicin. The cells were then imaged on M9Glu agarose pads also containing 100 µg/mL rifampicin via HILO at ~10 mW 638 nm laser and 50 ms exposure, at 30 min intervals. The null hypothesis was that

there would be no change with exposure to rifampicin if signals were not related to the transcription cycle; conversely, with an effective block on the available circulating RNAPs, fewer target mRNAs should be generated overtime, and hence the ensemble mean intensity per cell should plateau and decrease over time with constant exposure to rifampicin. Since the minimum inhibitory concentration of rifampicin has been reported between 8-25 μg/mL (Elitas et al., 2023; Weinstein & Zaman, 2018), an upper bound 4-fold higher concentration was applied to ensure effective RNAP block.

Figure 3.14A shows tiled maximum intensity projections of representative FOVs for all three strains, from 0 to 60 min (T0-T60), where T0 is defined as the addition of cells to the rifampicin agarose-pad. All FOVs seem quite bright – which is in part due to the autofluorescent quality of rifampicin itself – but visually there is an overall decrease over time in the number of cells with fluorescent signal, as well as an apparent decrease in the abundance of signal per cell, i.e. fewer super bright cells. Curiously, T30 appears to show equally bright cells, if not slightly brighter cells, than the T0 sample. By T60 though, all samples show proportionally reduced signal across the cell population, suggesting that RNAP block by rifampicin was effective in inhibiting transcription, and thereby the tag transcripts are increasingly fewer in induced tag strain cells.

These observations are corroborated via plotting the whole cell intensity distribution to visualize cell intensity trends overtime (Figure 3.14B, lower panel). There is an overall trend of consistently lower signal at T60 comparative to T0, but the T30 populations unanimously exhibit more cells with much higher brightnesses, and an apparently brighter population generally. Perplexingly, mean intensity box plots of individual cell mean intensities, whilst showing some decrease in intensity for both tag strains, show an increase in intensity in wildtype cells. These observations require further testing for a more robust determination of transcription specificity.

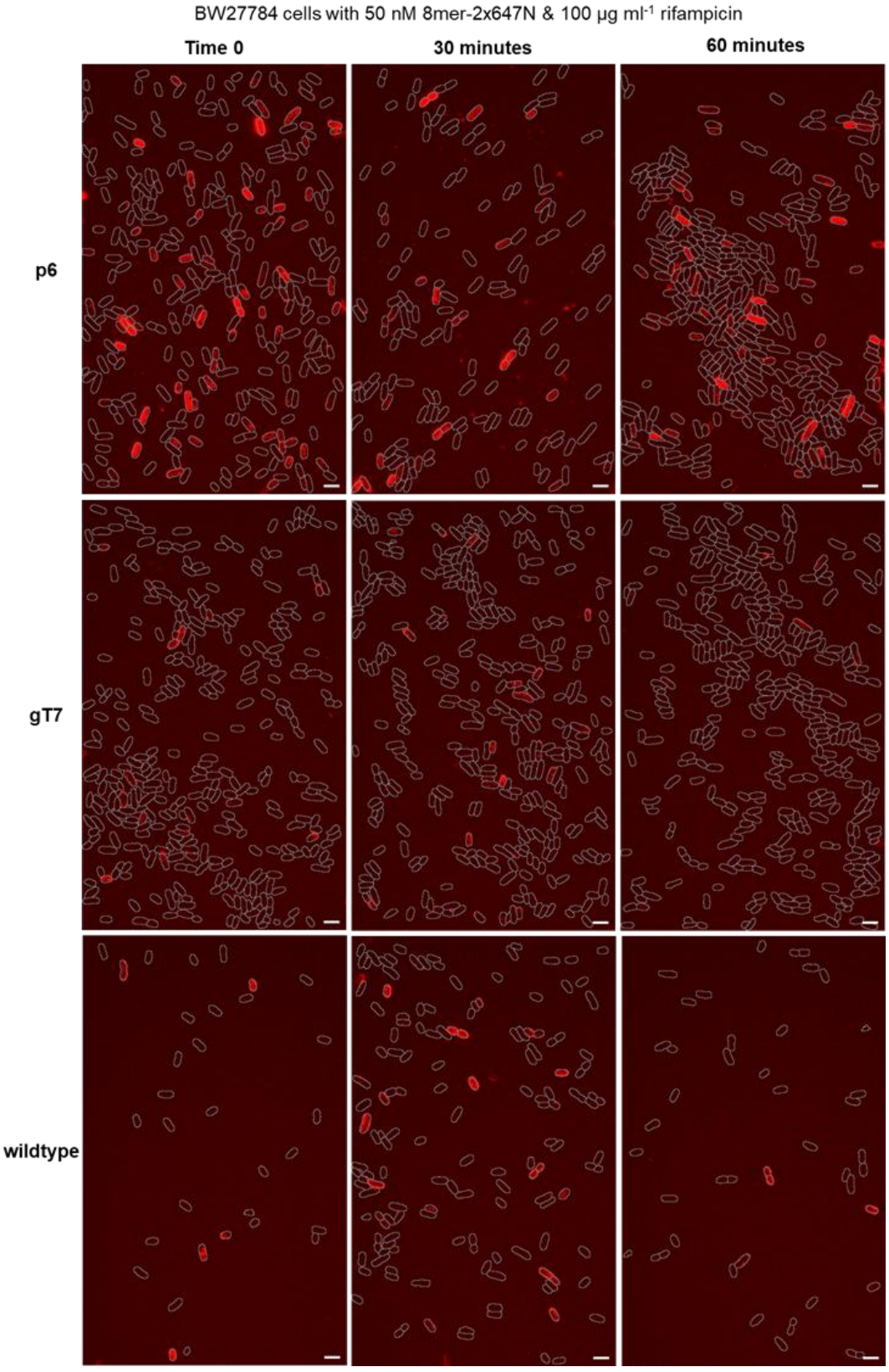


Figure 3.14A – Testing transcription specificity – rifampicin block on transcription in live wildtype, induced p6 and gT7 BW27784 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N: representative visuals.
Maximum intensity projections of frames 100-1000 (~45 s) of probe-electroporated cells recovered in 100 µg/mL rifampicin and imaged on 100 µg/mL rifampicin M9Glu agarose pads at 0, 30 and 60 min.
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

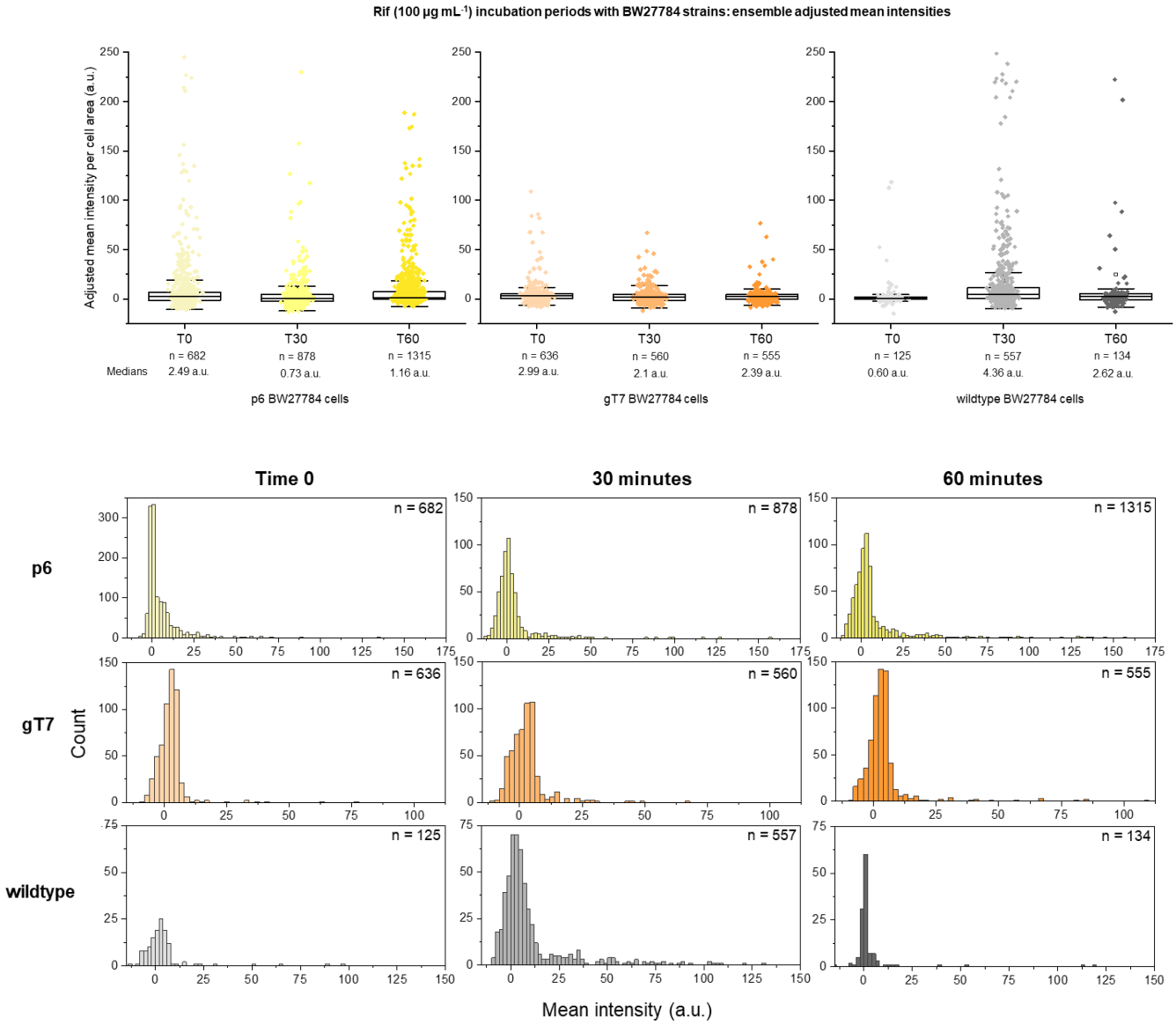


Figure 3.14B –Testing transcription specificity with rifampicin block: population-level intensity distributions.
(Upper) Whole population mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for p6, gT7 and wildtype cells with 50 nM 8mer-2x647N exposed to 100 µg/mL rifampicin (Rif) during recovery and imaging (0-60 min period). Intensities are calculated from I_{100}, corrected for background (but not classified for loading).
p6 medians: 2.49 units T0, 0.73 units T30, 1.16 units T60; gT7: 2.99 units T0, 2.1 units T30, 2.39 units T60; wildtype: 0.60 units T0, 4.36 units T30, 2.62 units T60.
(Lower) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 2.5, n gives the number of cells per sample (N = 1).
All mean intensities per cell area shown are background corrected, taken at the 100th frame.

3.5 Technical methodological optimisation

The initial evidence towards *in vivo* smFISH is based on fluorescent signals specific to probe de-quenching, although exact target mRNA specificity so far appears quite variable and challenging to immediately categorise. Although well-established and broadly applicable in live cell single molecule work, the methodological pipeline thus far has not been optimised for the precise function in question: internalising exogenous labelled short ssDNAs for the specific tagging of target mRNAs.

As such, a series of technical methodological experiments were conducted with the aims of:

- Improving the probe-mRNA complexation specificity (as represented by fewer ensemble signals and single molecule characterisation in wildtype cells)
- Improving the internalisation efficiency of the exogenous labelled ssDNA probes
- Improving the distribution of cells in a population with internalised probe

From 3.5.3 onwards, all experiments use the original 8mer-2x647N probe unless otherwise specified.

3.5.1 BW27784 versus MG1655

The K-12 MG1655 *E. coli* strain is the most commonly used and documented laboratory *E. coli* strain, comprising the fewest genomic mutations. MG1655 is hence accepted as the most reliable laboratory strain for physiological relevance as a prokaryotic model.

The BW27784 *E. coli* strain, also of K-12 background, originates from the Lambda red recombined BW25113 K-12 strain pioneered by Wanner & Datsenko (Datsenko & Wanner, 2000), the parent strain of the Keio collection of ~4000 single-gene knockout mutants (Baba et al., 2006; N. Yamamoto et al., 2009). Complete genome sequencing of the BW25113 highlighted three key differences to MG1655: the deletion of the arabinose and rhamnulose operons (araBAD, rhaDAB), a *lacZ* fragment replacement with 4x tandem *rrnB* terminators, and an *hsdR* frameshift mutation generating premature termination (Baba et al., 2006). The daughter strain, BW27784, was first generated by Khlebnikov et al. in constructing arabinose-inducible expression systems using pBAD and constitutive expression of the low-affinity arabinose transporter *araE* (Khlebnikov et al., 2001). As such, the BW27784 strain hosts only the *araE* gene of the arabinose transporters, as well as the parent deletions.

To determine whether MG1655-background strains showed a more apparent difference in mean cell intensity per area and fluorescence intensity distributions between tag strains and

wildtype cells with 50 nM 8mer-2x647N probes compared to the BW27784 cells, M9Glu competent MG1655 cells of all three strains were treated under the same conditions.

Side-by-side comparisons of representative FOV maximum intensity projections for wildtype cells with 50 nM 8mer-2x647N probes show, first and foremost, that far fewer cells have any measurable signal in MG1655 cells compared to BW27784 cells (Figure 3.15A). These FOVs even show that the MG1655 signals are shorter and more quickly photo-bleached than in BW27784 cells (Figure 3.15B). Of the signals observed in the MG1655 cells, most cells appear highly bright, either fully flooded with signal or with distinct foci that are membrane-proximal, suggesting that probe de-quenching and consequent fluorescence could be due to membrane implantation of the probes or membrane trapping, over sequence-mediated de-quenching.

Comparing the MG1655 loaded mean intensity per cell area to those of the BW27784 populations, whilst p6 and gT7 MG1655 population intensities are lower than in BW27784 (MG1655 I_{100} medians: 7.8 units p6, 8.2 units gT7), the wildtype MG1655 loaded population is considerably lower than the equivalent BW27784 (I_{100}: 2.0 units), with a much tighter distribution of cell intensities in the I_{100} measurement (Figure 3.15C). Displaying the distribution of mean intensities of cells across the whole population from I_{100}, both p6 and gT7 cells showed a spread of intensities up to 50 units, and even higher for gT7 cells. Comparatively, the majority of the wildtype population was left-skewed, with sub-zero values (~25% cells classified as unloaded, showing negative values from background correction in calculation) or 0-5 units (~36% cells).

Interestingly, gT7 cells showed a much wider spread of intensity values; this could represent a variable copy number of chromosomal transcripts under the T7 promoter, as well as variable internalised probe concentrations. Considering p6 cells also had most cells with 0-5 unit signals (~49% cells), and hence more overlapping with the wildtype MG1655 populations, it is apparent that thresholding and defining of specific probe-mRNA hybridisation events are again more dependent on single molecule characterisation than ensemble thresholding.

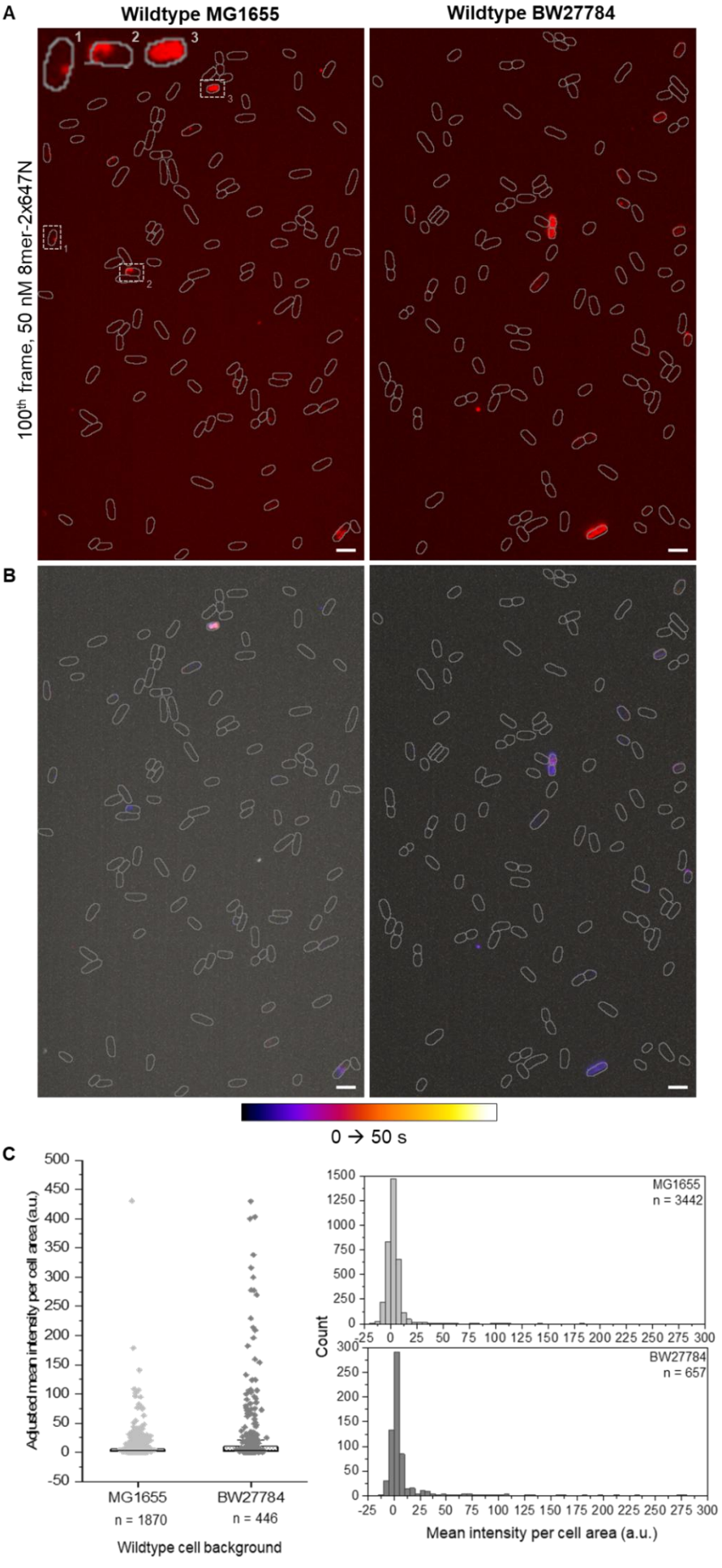

A
Wildtype MG1655
Wildtype BW27784
100th frame, 50 nM 8mer-2x647N
B
0 → 50 s
C
Adjusted mean intensity per cell area (a.u.)
MG1655
n = 1870
BW27784
n = 446
Wildtype cell background
Count
MG1655
n = 3442
BW27784
n = 657
Mean intensity per cell area (a.u.)

Figure 3.15 – Comparing the influence of strain background in signal specificity: MG1655 versus BW27784.
(A) Maximum intensity projections of 100-1000 frames (~45 s) representative FOVs of MG1655 and BW27784 wildtype M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes. The MG1655 representative FOV shows an upper left inset of numbered exemplar fluorescent cells, diffuse with signal or with specifically membrane proximal signal.
(B) Corresponding temporal projections of 1-1000 frames; the colour bar denotes time, where darkest is time 0 (50 ms exposure, 50 s acquisition).
(C) Loaded cell mean intensities per cell area with corresponding boxplots (left) (25%, median and 75% quartiles with outlier + 1.5 whiskers) median and 75% quartiles with outlier + 1.5 whiskers) for 50 nM 8mer-2x647N-electroporated wildtype MG1655/BW27784 cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Medians: 2.0 units MG1655 wildtype, BW27784 as in 3.4.3.
Histogram distributions of the whole cell populations (right); bin size = 5.
Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken at I_{100}, and n gives the number of cells per sample (N = 3 experiments for both backgrounds respectively).

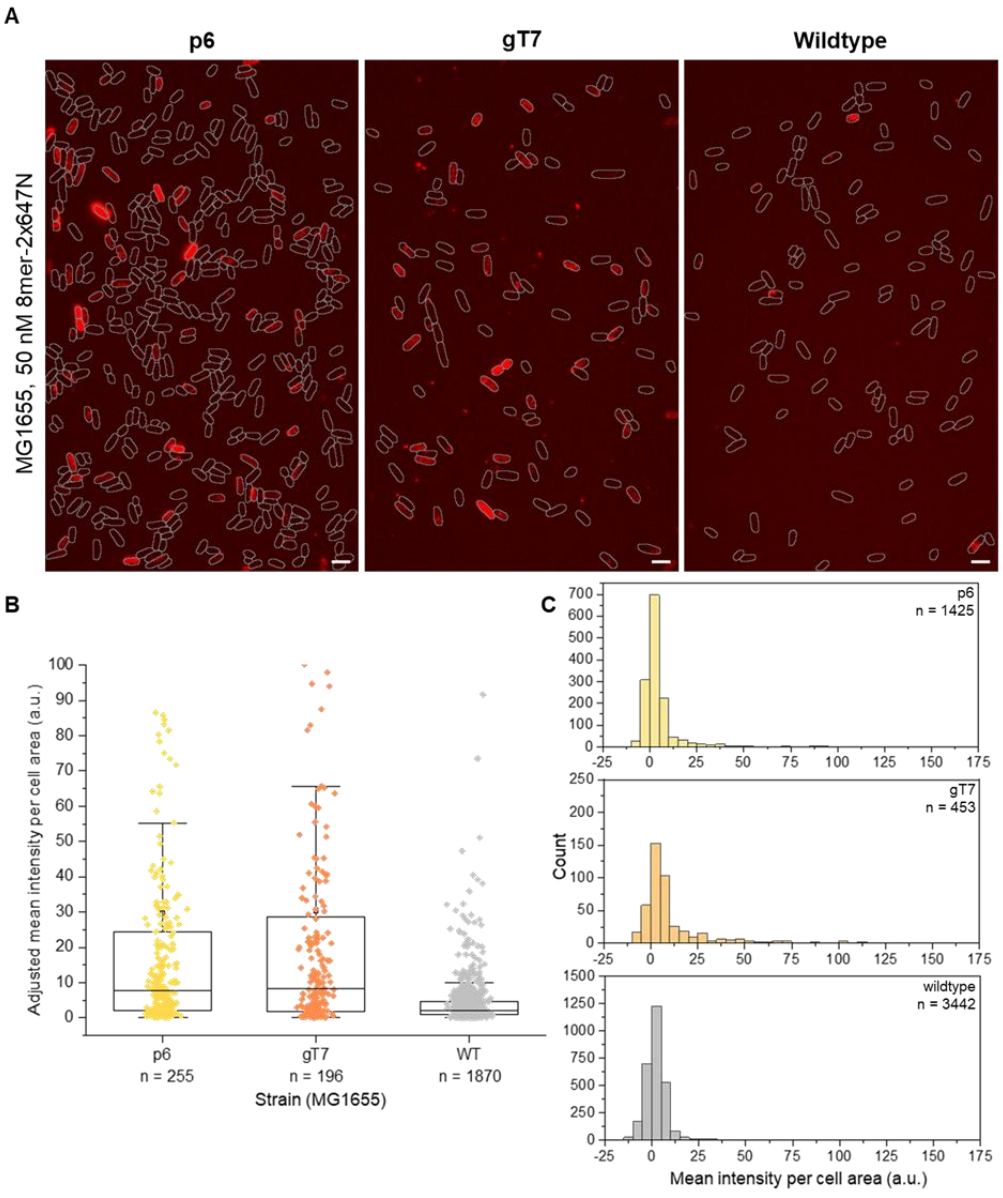


Figure 3.16 – Fluorescence observations in live wildtype, induced p6 and gT7 MG1655 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N.
(A) Maximum intensity projections of 100-1000 frames (~45 s) representative FOVs of MG1655 wildtype, induced p6 and gT7 M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes (false-colour in red).
(B) Loaded cell mean intensities per cell area with corresponding boxplots (left) (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 50 nM 8mer-2x647N-electroporated wildtype MG1655/BW27784 cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Medians: 7.8 units p6, 8.2 units gT7; wildtype as 3.5.1.
(C) Histogram distributions of the whole cell populations (right); bin size =5.
Scale bars – 3 µm, FOVs overlaid with cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken at I_{100}, and n gives the number of cells per sample (N = 3, 2, 3 experiments for p6, gT7, wildtype respectively).

As a proxy for thresholding tag cells with specific signal, loaded p6 and gT7 MG1655 cells were corrected for a third time by subtracting the averaged mean intensities of the 'loaded' wildtype cells (Figure 3.18). This gave potential 'specifically-loaded' populations of ~11%, 26% for p6 cells (I_0, I_{100}) and ~18%, 39% for gT7 cells (I_0, I_{100}); with median population intensities of 20.7 units for p6 and 22.1 units for gT7 cells at I_{100}. Corresponding histograms are shown in Supplement 3.2.

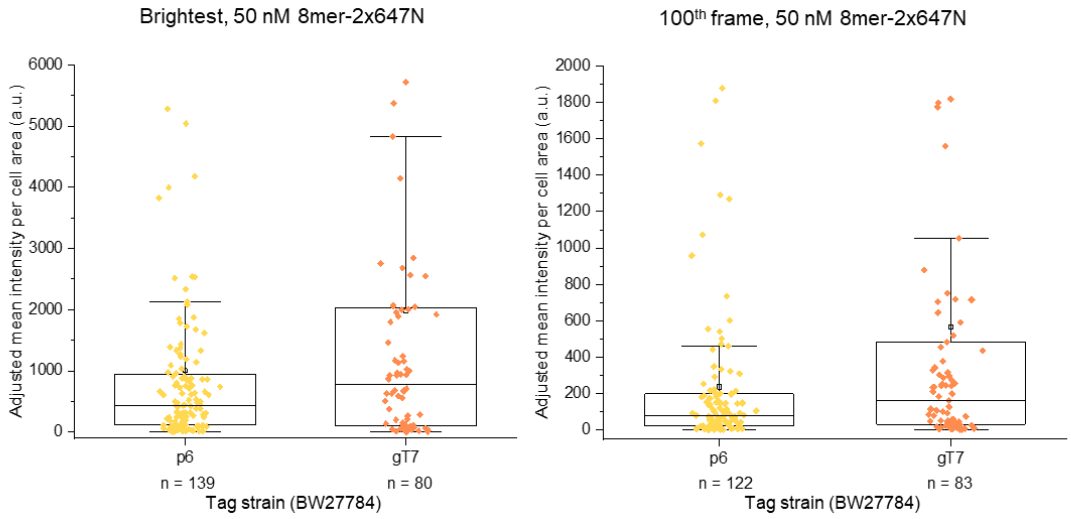


Figure 3.17 – A proxy for specific signals in MG1655 tag strains electroporated with 50 nM8mer-2x647N probes: populational mean intensity distributions for wildtype-corrected loaded p6 and gT7 cells.
Mean intensities per cell area (I_0, I_{100}) with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for classified loaded p6 and gT7 cells from 3.12, further to subtraction of the averaged mean intensity of loaded wildtype cells.
Wildtype-subtracted loaded cell medians: I_0 72.3 units p6, 50.2 units gT7; I_{100}: 20.7 units p6, 22.1 units gT7.
n gives the number of cells per sample (N = 3, 2 experiments for both strains respectively).

Whilst the arabinose-inducible expression systems are generally considered to be less "leaky" than other systems (such as LacZ) – presenting more all-or-nothing expression phenotypes – it is possible that transcription can occur without proper arabinose induction at pBAD. This appeared to be the case with the p6 and gT7 strains, where little difference is observed regardless of arabinose addition during growth phases – suggesting lack of efficacy of the arabinose system overall. As for brighter signal observations in wildtype BW27784 over MG1655 cells, one potential explanation could be that the transporter mutations in the BW27784 strain have led to membrane composition differences giving an overall increase in porosity and propensity to internalise small molecules, such as 8mer ssDNAs. Curiously, a comparison of a "palette of fluorophores" by Salcedo-Sora et al. indeed found significant differences in the uptake of different dyes between BW25113 and MG1655 *E. coli*, and noting that the BW25113 mutations do have major implications as to membrane transport systems (Salcedo-Sora et al., 2021), the levels of careful experimental design and heuristics required for feasible interpretation are highlighted. There is insufficient evidence to substantiate this here, however, and further mechanistic evidence as to the reason behind increased signal in BW27784 wildtype cells has so far been out of scope.

As such, MG1655-background cells were deemed a more suitable actor for *in vivo* smFISH regarding consistent identification of hybridisation specificity, and all further experiments presented in this thesis were conducted using MG1655 *E. coli* cells.

3.5.2 Colocalising p6 plasmids

By way of attempting to define further controls for probe-specific observations *in vivo*, a *tetO*/TetR reporter dual plasmid system was designed to attempt to give a proxy for the number of p6 plasmids per cell.

Figure 3.18 presents a schematic for the co-localisable p6 construct – here, the 8mer tag array (15 nt spacers) is associated with a 1000 nt fragment, under the control of the high expression Lac-inducible pTRC promoter (E. Amann et al., 1988) and strong T7 phage terminator. The 25x *tetO* operator sites are half of the reporter system, whereby TetR-YPet fusions generated from a second plasmid are able to bind to the *tetO* sites and fluoresce to localise the p6 plasmids. On addition of tetracycline, the TetR-YPet may dissociate from the *tetO* sites, and contribute to high expression of the 1078 nt target transcript. Co-localisation of red and green foci would theoretically overlay, and could be seen as yellow signals in false colour representative FOVs.

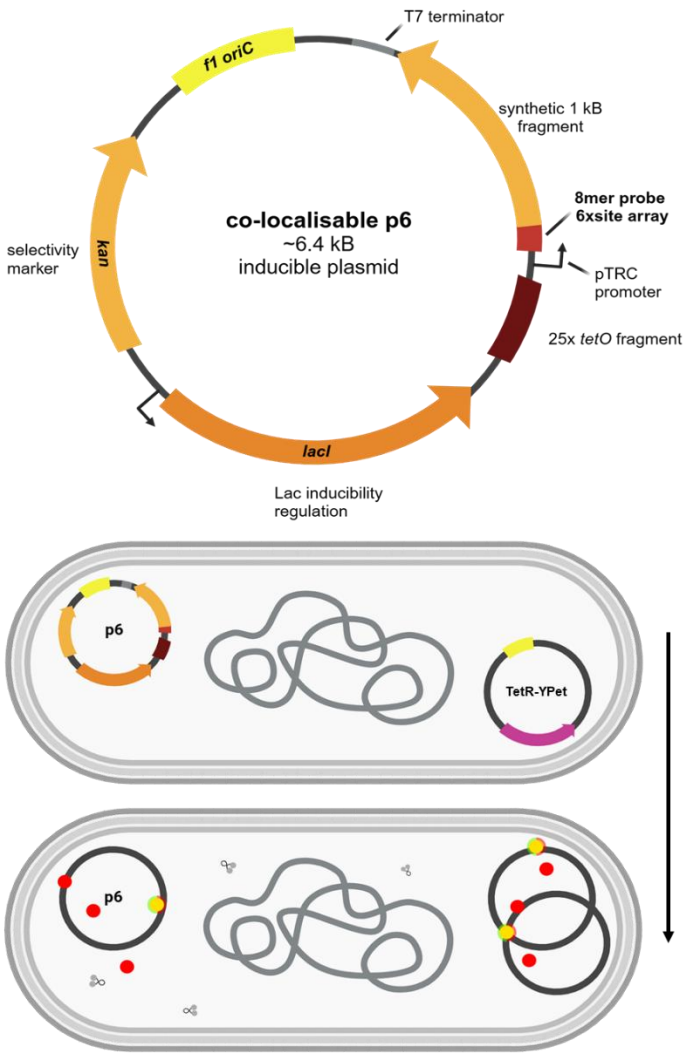


Figure 3.18 – Gauging p6 plasmid copies per cell: a co-localisable *tetO/R* reporter p6 strain. (Upper) Plasmid schematic denoting the 6-site array in conjunction with a 1 kb synthetic sequence for tagged RNA output, from a low copy number p15a hosting a 25x *tetO* array. (Lower) Schematic of possible observations in MG1655 *E. coli* transformed with the *tetO*-p6 plasmid and a TetR-YPet strain; where the *tetO*/TetR interaction leading to TetR-YPet expression would be anticipated to show both red (fluorogenic ssDNA probe de-quenching) and green (YPet) signals co-localised to regions of high plasmid density, leading to yellow foci.

Maximum intensity projections of a representative FOV for co-localisable p6 cells present the 8mer 2x647N signal in red, YPet in green and a combined overlay of the two (Figure 3.19). Most cells appear to contain diffuse YPet signal, challenging to specifically localise to a plasmid region; some cells present with distinguishable foci, as presented by the exemplar YPet cells in the lower left of the YPet FOV. These distinguishable foci could correlate to plasmid-dense regions, and indeed overlaying the ATTO647N and YPet signals gave some yellow foci, indicating co-localisation of YPet and probe signals, hence the probe-RNA complexes *in situ* with their plasmid. More analysis is needed for quantification and interpretations as to number

of generative plasmids, which will likely be challenging due to the predominantly diffuse YPet FP signal that floods most cells.

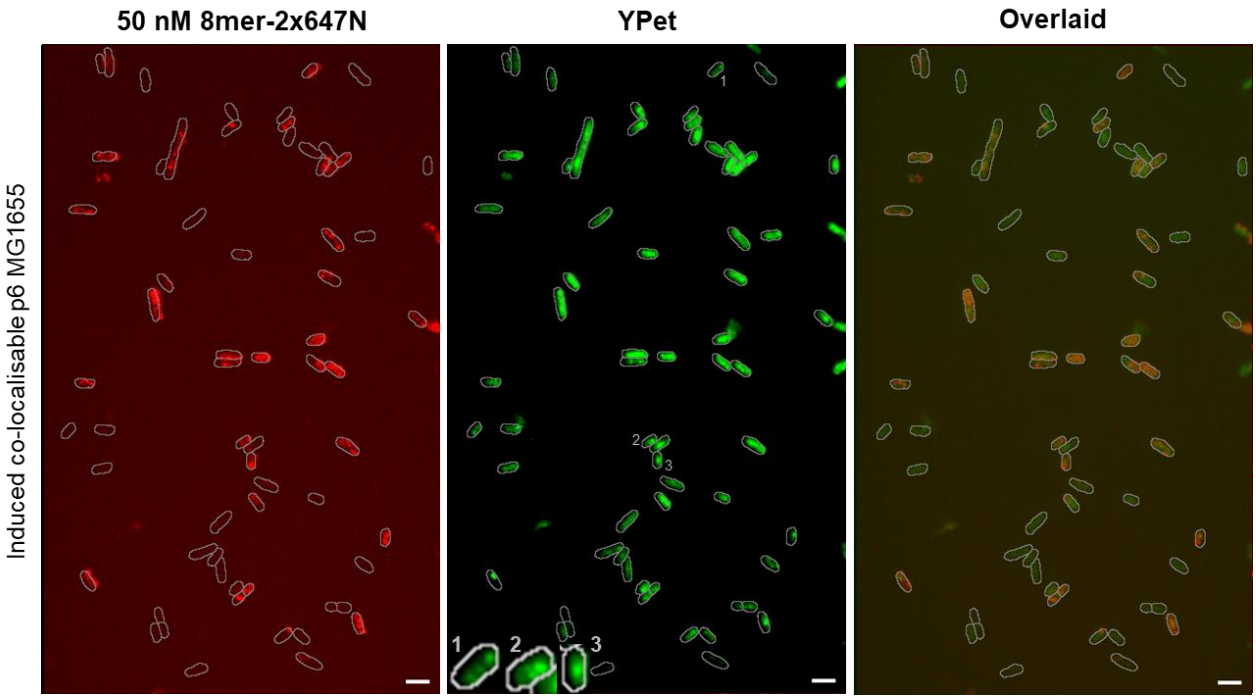


Figure 3.19 – Gauging p6 plasmid copies per cell: a co-localisable *tetO*/TetR reporter p6 strain. Maximum intensity projections of a representative FOV of the co-localisable p6 competent M9Glu cells electroporated with 50 nM 8mer-2x647N showing probe fluorescent signal (red false-colour), YPet signal (green false-colour) and co-localised foci (overlaid red and green false-coloured projections). Insets of exemplar cells with distinguishable YPet foci numbered in the lower left of the YPet FOV.
Scale bars – 3 µm; overlaid cell segmentation outlines.

3.5.3 A question of probe specificity – comparing probes for specific signal

With the acknowledgement of strain background as a contributor to potential signal confoundment, and the choice of MG1655 in presenting lower off-target signal in wildtype cells, the next questionable constituent in signal specificity was the probes themselves.

To attempt to improve signal specificity, three different forms of the 8mer-2x647N probe were trialled – the "original" 8mer-2x647N, a newly-synthesised "commercial" 8mer-2x647N, and an aliquot of the commercial 8mer-2x647N run through a second HPLC purification process in-house ("in-house-purified"). A new fluorogenic probe, 8mer-ATTO643/BHQ1 (8mer-643/q1) was also designed. All probe samples were tested in wildtype MG1655 cells to establish which probes showed least off-target signal.

As relayed in section 3.1.2.2, ATTO647N is only moderately hydrophilic, lending its lower water solubility to a higher likelihood for aggregation and hydrophobic interactions – often as "sticking" to the bilayer membranes of the cell (Zanetti-Domingues et al., 2013). This stickiness is more likely to lead to background and off-target fluorescence, thereby confounding or overshadowing specific signals (Kolmakov et al., 2010). Of course, in the synthesis and purification of such short ssDNA constructs, a 100% dual labelling efficiency and "pure" product is also unrealistic. As such, there is a possibility of background signal contributions from half species (singly-labelled ssDNAs, for example), and free dye in solution. Nevertheless, free dye in solution should be sustainably removed via the standardized washing steps, and it would be expected that half species – singly-labelled ssDNAs – would photo-bleach on a shorter timescale than the quenched and thereby inherently self-photo-protected 8mer-2x647N.

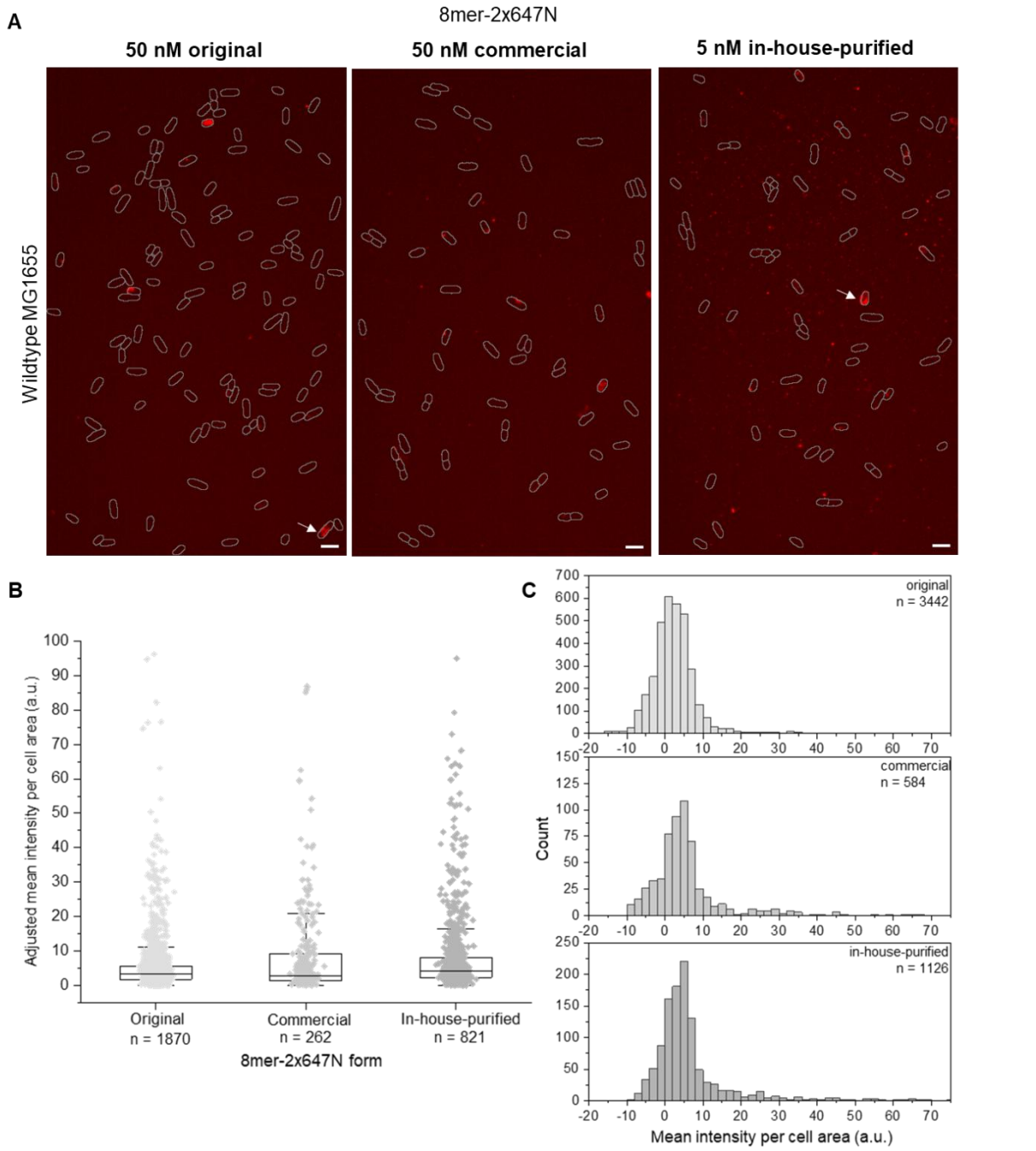


Figure 3.20 – Probing signal specificity: competent M9Glu wildtype MG1655 electroporated with different forms of the 8mer-647N probes

(A) Maximum intensity projections of frames 100-1000 of wildtype MG1655 cells electroporated with 50 nM original, 50 nM commercial and 5 nM in-house-purified 8mer-2x647N probes.

(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for corresponding 8mer-2x647N-electroporated wildtype cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.

Wildtype loaded cell medians: 3.4 units original, 2.8 units commercial and 4.1 units in -house-purified.

(C) Comparative histogram distributions of the whole cell populations (where unloaded classify with sub-zero values) between wildtype, induced p6 and gT7 cells; bin size = 2, n gives the number of cells per sample (N = 3 for wildtype).

Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from I_{100}, and n gives the number of cells per sample (N = 3, 2, 3 per probe form).

Comparing representative FOVs maximum intensity projections of wildtype MG1655 cells with 50 nM original and commercial, 5nM in-house-purified 8mer-2x647N (Figure 3.20A), it seems that the original probe aliquot has the lowest number of cells containing any signal. Though promising, this could also be due to the age of the probe sample, multiple freeze-thawing events and general use. The commercial 8mer-2x647N seems to show slightly higher signal in wildtype cells, perhaps due to lesser degradation and reduced propensity to rapid photo-bleaching comparative to the original 8mer-2x647N sample.

The double-HPLC purified 8mer-2x647N shows the highest wildtype signal. In-house HPLC chromatograms indicated that most peaks observed had both DNA- and ATTO647N-associated signals; where black arrows denote collected peak fractions (Supplement 3.3). According to reverse-phase HPLC, where hydrophobic molecules take longer to elute, the different peak fractions suggested 3-4 possible species with variable charge interactions (free dye, singly-labelled ssDNAs, doubly-labelled DNAs). The 8mer-2x647N doubly-labelled construct, expected to be net negatively-charged from the complexation of 2xATTO647N with the ssDNA backbone, would therefore have a longer elution time. Considering the in-house-purified probes exhibit the fastest bleaching within the first 5 seconds (further demonstrated in Figure 3.23), and few recurrent signals afterwards, primarily extracted free dye may be in observation as opposed to intact fluorogenic probes.

These visual observations were also reflected in the "loaded" cell I_{100} mean intensity plots, whereby the medians of the original, commercial and in-house-purified probes are 3.4 units, 2.8 units and 4.1 units respectively (Figure 3.20B). Notably, the respective loaded populations per 8mer-2x647N probe sample were ~54% for the original, ~45% for the commercial and ~73% for the in-house-purified 8mer-2x647N. The corresponding histogram distributions of the whole cell populations show that the original probe indeed has the highest population of cells with negative (~ 35% cells) and 0-5 units (~42% cells) mean intensities compared to the two newer 8mer-2x647N samples, both of which had a consistently longer tail of higher intensity cells (up to ~95 units for the commercial probe, ~65 units for in-house-purified 8mer-2x647N)(Figure 3.20C). Although few cells had very high intensities, classed as outliers – potentially dead and fully porous cells – both newer probe samples showed over 100 counts of cells with intensities beyond 12.5 units, suggesting a need for more rigorous thresholding or alternative designs.

Strikingly, visible on all of the wildtype cell samples are cells which appear to have a highlighted membrane, just at the periphery of the cell (Figure 3.20A, arrows). These

observations would suggest off-target trapping of the 8mer-2x647N probes (Salcedo-Sora et al., 2021), perhaps in the periplasmic space and between the inner and outer bilayer membranes, leading to a "racetrack" appearance. There are studies characterising periplasmic proteins and inner/outer membrane compositions with FP fusions or dye-labelled constructs (Ke et al., 2016), but few studying the artefacts of labelled construct delivery. Dye/probe trapping could explain a higher intensity value for certain cells, through effective localised sequestration of probes and potentially membrane implantation or photochemical dequenching, but this has not been mechanistically explored and is purely a hypothetical explanation.

Accounting for the photochemistry of ATTO643-BHQ1 fluorophore pairings, whereby effective quenching of an emissive fluorophore can be achieved by using a dark quencher, the higher efficiency quenching and fluorogenicity of this particular pairing meant that no overexposure on laser activation was reproducibly seen. Indeed, fewer fluorescent signals in wildtype cells observed overall (Figure 3.21A). Using the I_0 mean intensity per cell area, and in direct comparison with the three forms of 8mer-2x647N probes, the 8mer-643/q1 probes showed a considerably lower median mean intensity per cell area than the three forms of 8mer-2x647N in wildtype cells (3.1 units for 8mer-643/q1, versus 6.2 units, 7.8 units and 13.5 units respectively), as well as a much tighter distribution and shorter outlier tail (Figure 3.21B).

Of the wildtype cell populations, ~42% cells with 8mer-643/q1 were classified as loaded at I_0 compared to 65%, 75% and 89% of wildtype cells with 50 nM original/commercial and 5 nM in-house-purified 8mer-2x647N respectively. This determines the 8mer-643/q1 probe as the optimal choice for *in vivo* smFISH experiments thus far, and is suggestive of ATTO647N membrane-sticking in contributing to non-specific fluorescence.

This tighter distribution is further corroborated by the histogram distribution of the whole cell population I_0 mean intensities, showing that for the 8mer-643/q1 probe in wildtype cells, there are negligible mean intensities higher than 30 units. All other probes tested give wildtype cells with intensities up to at least 50 units and often beyond (very bright) (Figure 3.21C). Whilst none are statistically significant at $p = 0.05$, the tighter distribution of intensities with 50 nM 8mer-643/q1, higher proportion of unloaded cells at I_0 and overall lower proportion of loaded cells compared to the three 8mer-2x647N probes, indicates that **the 8mer-643/q1 probe is the best choice for increased signal specificity,** at least at the populational level.

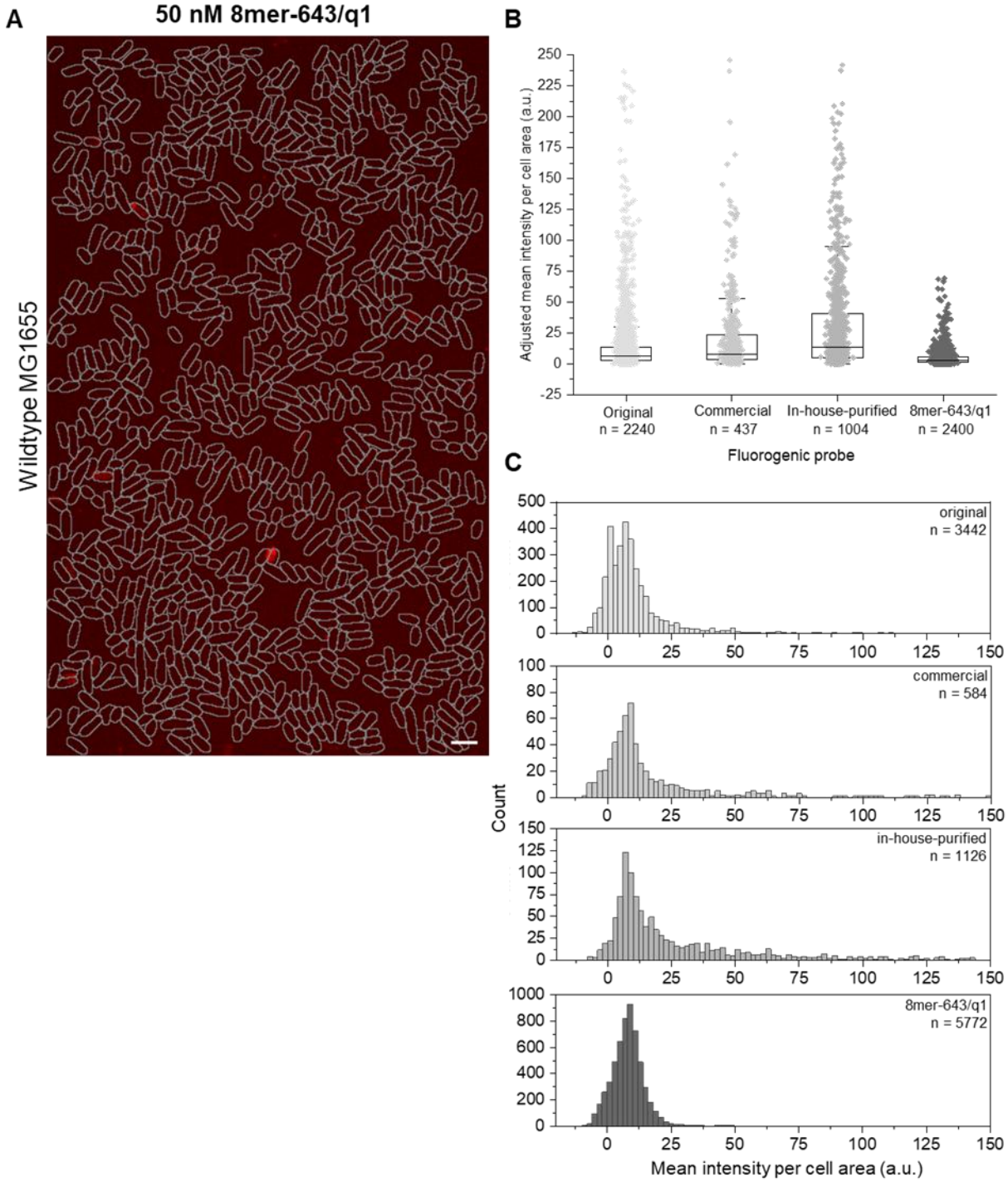


Figure 3.21 – Probing signal specificity: competent M9Glu wildtype MG1655 electroporated with 50 nM 8mer-643/q1 probes

(A) Maximum intensity projection of frames 100-1000 of wildtype MG1655 cells electroporated with 50 nM 8mer-643/q1 probes.

(B) Loaded cell I_0 mean intensities per cell area with corresponding boxplot (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 8mer-643/q1-electroporated wildtype cells compared against the loaded 8mer-2x647N populations. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.

Wildtype loaded cell medians: units 6.2 units original, 7.8 units commercial, 13.5 units in-house-purified and 3.1 units 8mer-643/q1.

(C) Comparative histogram distributions of the whole cell populations between all four probes in wildtype cells; bin size = 2.

Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from the brightest frame, and n gives the number of cells per sample (N = 3, 2, 3 per probe form).

Ensemble diffusion behaviours

As discussed in 3.4.3, whilst mean intensity per cell area measurements can give an insight into probe internalisation and off-target binding in wildtype cells – a proxy for specificity – there is less insight into the characteristics of those events. One way of gaining an overview into the single event characteristics is through single molecule localisation and tracking: to obtain the distribution of apparent diffusive states (D*) of molecules thresholded for tracking.

Figure 3.22 shows the D* histograms of the same wildtype cell populations electroporated with the various trialled probes. Although there are insufficient data to fit Gaussian curves to the commercial 8mer-2x647N population, all four probes trialled show a left-skewed distribution: the red fit corresponding to molecules classified as "bound" particles with all probes giving bound populations with a D* between 0.03-0.075 $\mu m^2/s$. Although the two species fit is approximated, the blue curve ("diffusing" molecules) is not particularly closely corresponding to the populations, indicating that the population of diffusing particles is likely smaller than suggested here.

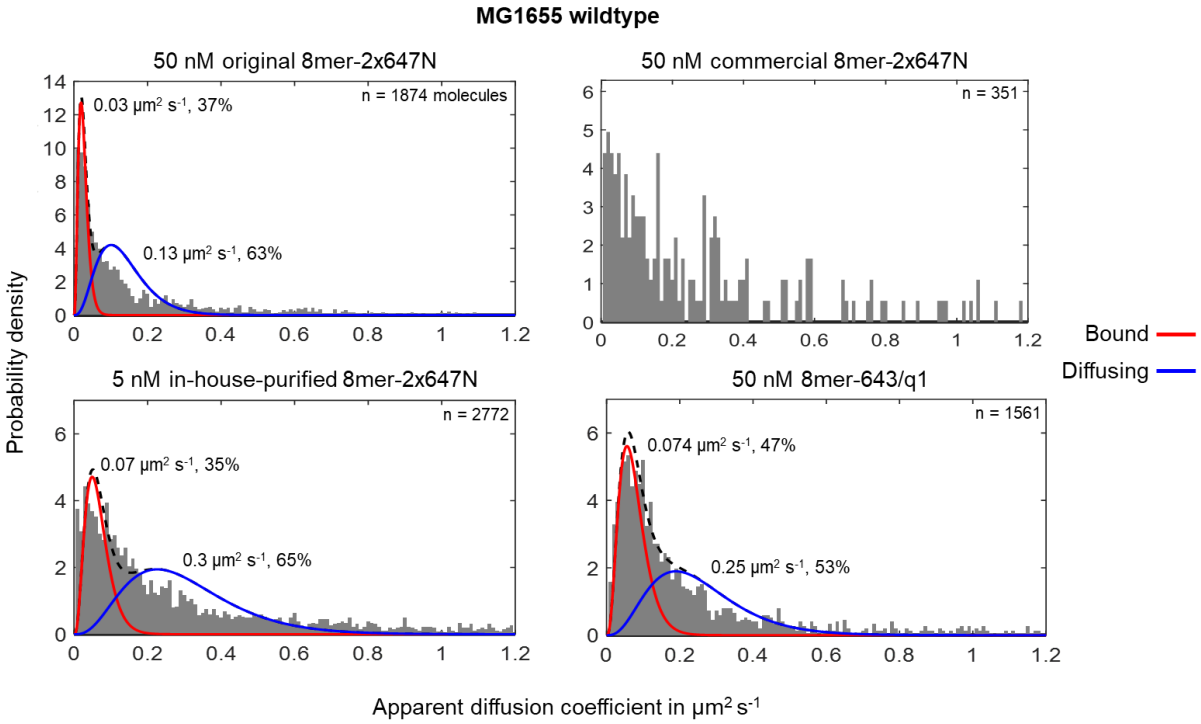


Figure 3.22 – Exploring the diffusive properties of tracked fluorescent molecules in wildtype MG1655 competent M9Glu cells with different probe forms: apparent diffusion coefficients. Apparent diffusion coefficient distributions fit with two-free species curves (red as bound, blue as diffusing), with the exception of the commercial probe sample (too few molecules, free fit); n gives the number of tracked molecules per sample.

Combining this insight into the diffusive states of the fluorescent particles in wildtype cells with observations of representative photo-bleaching time traces per probe tested (Figure 3.23), it is clear that in the majority of cases, loaded cells bleach quickly to baseline, and some outlier events occur such that fluorescence may arise late into an acquisition. From the average traces per probe sample in MG1655 wildtype cells, 88.6% original, 90.8% commercial, and 93.4% in-house-purified signal are bleached by 5 seconds; 98.1%, 98.4% and 99% by 10 seconds respectively. The 8mer-643/q1 probes appear to have negligible signal at the start of acquisition compared to the 8mer-2x647N probes samples, and show 73% and 81.7% signal bleaching by 5 and 10 seconds.

Thus, taking the original 8mer-2x647N as the most robust 8mer-2x647N option, further comparisons at the ensemble diffusive level can be made, after the 5 second benchmark for rapid photo-bleaching of off-target signal, followed by single molecule analyses per particle of interest. In the case of the ultimate aim: the 8mer-643/q1 appears to be the most robust probe of those trialled so far, towards contributing to engineered *in vivo* single molecule understanding of transcription initiation in this context.

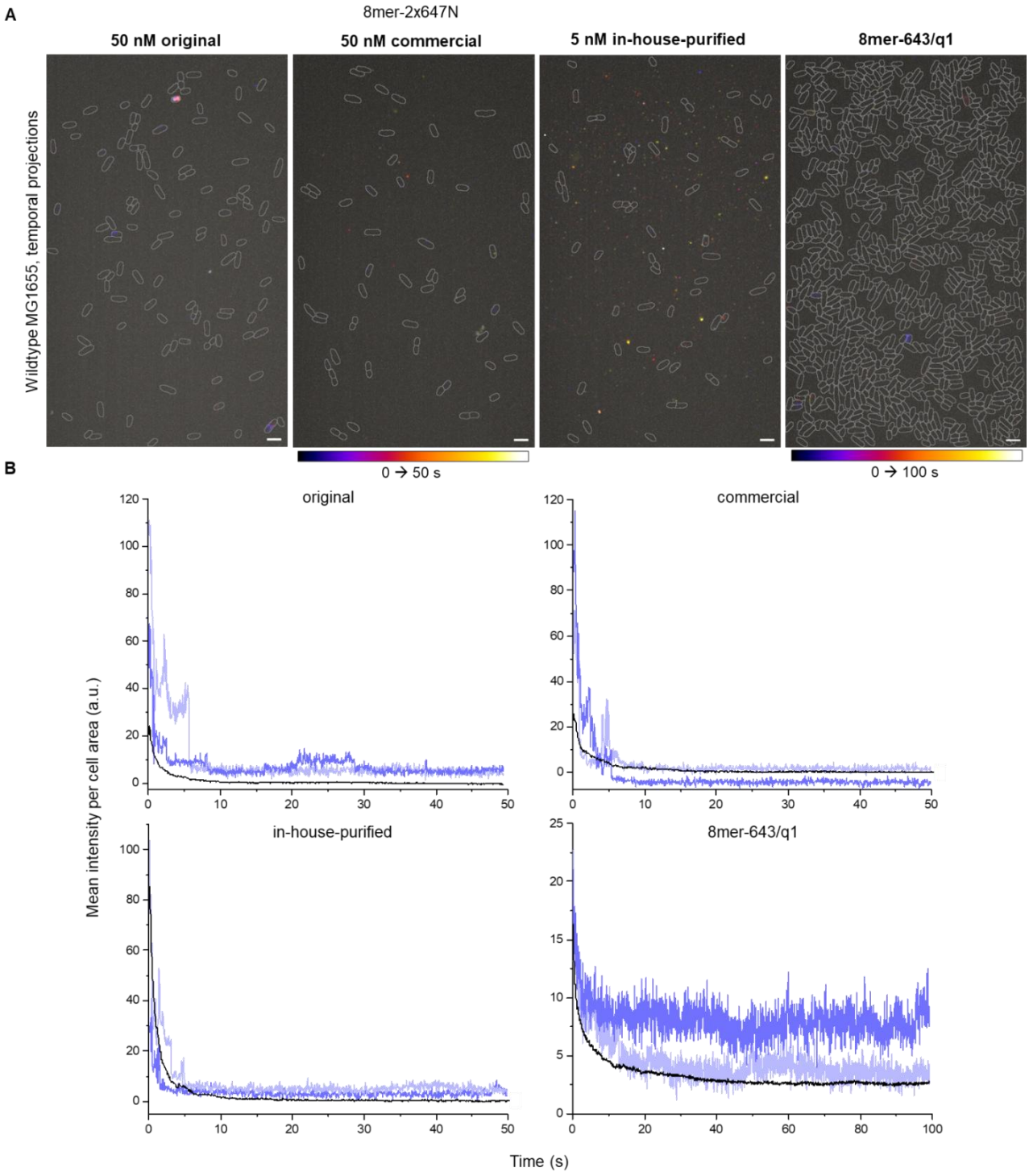


Figure 3.23 – Visualising the photostability of different probe forms in wildtype MG1655 *E. coli* M9Glu competent cells.
(A) Temporal projections of 1-1000 frames of the same representative FOVs of probe-electroporated wildtype MG1655 cells in 3.5.2, 1-2000 frames for the 8mer-643/q1 probes. The colour bar denotes time, where darkest is time 0 (50 ms exposure, 50-100 s acquisition).
(B) Corresponding averaged background-corrected mean intensity per frame time traces for the representative FOVS (black), overlaid with two exemplar bleaching time traces for individual visually-loaded cells (blue shades).
Scale bars – 3 µm.

3.5.4 Technical Optimisation

In 3.5, there were three stated aims for technical experimentation: improving hybridisation specificity, probe internalisation efficiency and probe distribution across a cell population.

A series of experiments were hence conducted with either MG1655 wildtype or induced p6 cells, particularly with internalisation efficiency and distribution homogeneity in mind. These concerned the comparison of LB and M9Glu growth media in competency, exploration of electroporation parameters (recovery duration, temperature and media), and imaging parameters. Figure 3.24, as a technical representation, shows representative FOVs and ensemble analyses for the comparison of growth media in competency – to try to establish the least background autofluorescence, increasing specific signal resolution – and to optimize for the number of signals observed across a population.

Representative FOV maximum intensity projections of wildtype cells electroporated with 100 nM 8mer-2x647N (original) show that M9Glu wildtype competent cells have fewer cells proportionally that have any signal compared to LB competent cells (Figure 3.24A). Moreover, LB-grown samples tested appeared to have some spurious background signal; some very faint signals in and around cells can also be seen in the upH_2O representative control (righthand side), although this could be due to insufficiently effective washing away of the probe.

Plotting loaded cell mean intensities and the whole population distributions for the wildtype cells, the M9Glu appeared to have a wider spread of loaded cells and median intensity (4.6 units) than the LB wildtype cells (4.1 units)(Figure 3.24B-C). Notably, though, the overall population distribution was much more skewed towards cells with low or no signal in M9Glu cells over LB cells, (though this could partly be sampling bias).

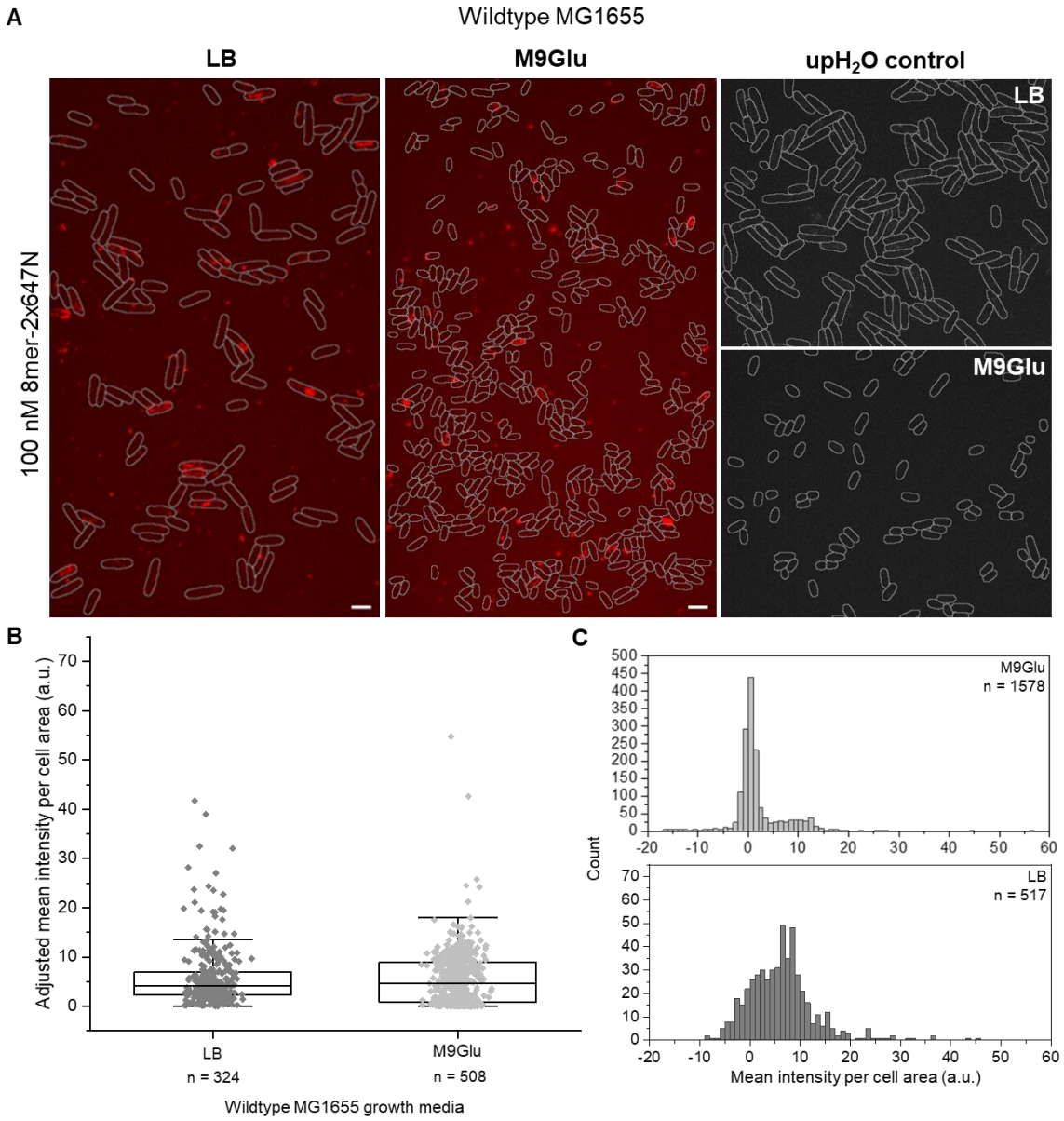


Figure 3.24 – Comparing LB versus M9Glu for growth pre-competency.
(A) Maximum intensity projections of frames 100-1000 of wildtype MG1655 cells electroporated with 100 nM original 8mer-2x647N probes, grown in LB or M9Glu; corresponding control upH$_2$O representative FOVs alongside.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 8mer-2x647N-electroporated LB and M9Glu competent wildtype cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH$_2$O-electroporated (empty) cells.
Wildtype loaded cell medians: 4.1 units LB, 4.6 units M9Glu.
(C) Comparative histogram distributions of the whole cell populations (where unloaded classify with sub-zero values) between LB and M9Glu wildtype cells; bin size = 1.
Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from I_{100} , and n gives the number of cells per sample (N = 1, 3 LB and M9Glu respectively).

This difference between LB- and M9Glu-grown competent cells is apparent and interesting: LB as an established, common rich medium for standardized cell growth would seem optimal for observing single molecule events, particularly relating to processes like transcription. In richer media, stress states are less probable, cell fitness more likely maintained, a physiological representation of fast cellular growth. However, the competency procedure itself is a likely inductor of stress, as is the electroporation step, despite the subsequent rich media recovery period; therefore it is unclear how pertinent pre-competency growth is in securing representative cellular states. There is also the question, of course, of the compromise of optimal reagents/procedures for clean and highly distinguishable signal at single molecule resolution – with richer media come more autofluorescent components, decreasing such photo-clarity.

As such, M9Glu-grown cells seem a more preferable choice for *in vivo* smFISH experiments: closer physiological representation to native states with unpredictable nutrient availability, and optically cleaner for single molecule imaging. Interestingly, the higher probe internalisation rates of M9Glu cells may be explained by the poorer nutrient content of M9Glu media[4], in that with partial nutrient starvation, reduced synthesis for cell membrane components and robustness and increased probability to the environment for optimised uptake may inevitable lead to greater short ssDNA uptake too (Sastry et al., 2019; Tao et al., 1999). This phenomenon of media dependency for cell wall robustness is not thoroughly described in published literature, and will be discussed in more detail further on.

The choice of M9Glu competent cells formed the basis for all further experimentation as follows for the single molecule and potentially single fluor observations of probe-mRNA hybridisation events *in vivo*. As per other specific optimisation parameters, corresponding figures and text can be found in Appendix 1.

Electroporation parameters

To optimize for increased numbers of cells with internalised probes (as any fluorescent signal), induced p6 cells were incubated with 50 nM original 8mer-2x647N probes for 5 or 10 min on ice, electroporated, and recovered for 5 or 20 min at room temperature or 37°C. Summarising the mean intensity distributions by taking the median of loaded cell mean intensities (Figure 3.25A), conditions were overall negligibly different, albeit with 20 min recovery duration

[4] M9Glu composition (2x): 5 mL 20x M9 salts, 1M casamino acids, 1M vitamin supplements, 1M $CaCl_2$, 1M $MgSO_4$, 0.2% glucose, upH_2O
Doubling time for MG1655 strains 2-4 hours (2 hour lower bound for wildtype cells).

appearing to give higher intensity signal at all recovery temperatures. This is corroborated via the histogram distributions (Appendix 1.1). Given the ~20 min "standard" doubling time of *E. coli* (Son & Taylor, 2021), a 20 min recovery period would seem most appropriate. Subsequently, a higher probe internalisation and fluorescent event observation subsequent to 20 min recovery is unsurprising; indeed, a longer recovery period would also increase pore-healing post-1.4 kV/cm shock. Similarly, 37°C as a recovery temperature was expected to be optimal, due to the proclivity for *E. coli* well-elucidated growth phases at the 30-37°C range – appreciating that *E. coli* are enterococcal prokaryotes, and adapted primarily to cohabiting mammalian digestive tracts. These observations further confirm the reproducibility of the published protocol by Aigrain et al., 2015; as such, a 10 min probe incubation and 20 min recovery period at 37°C was elected for all *in vivo* smFISH experiments.

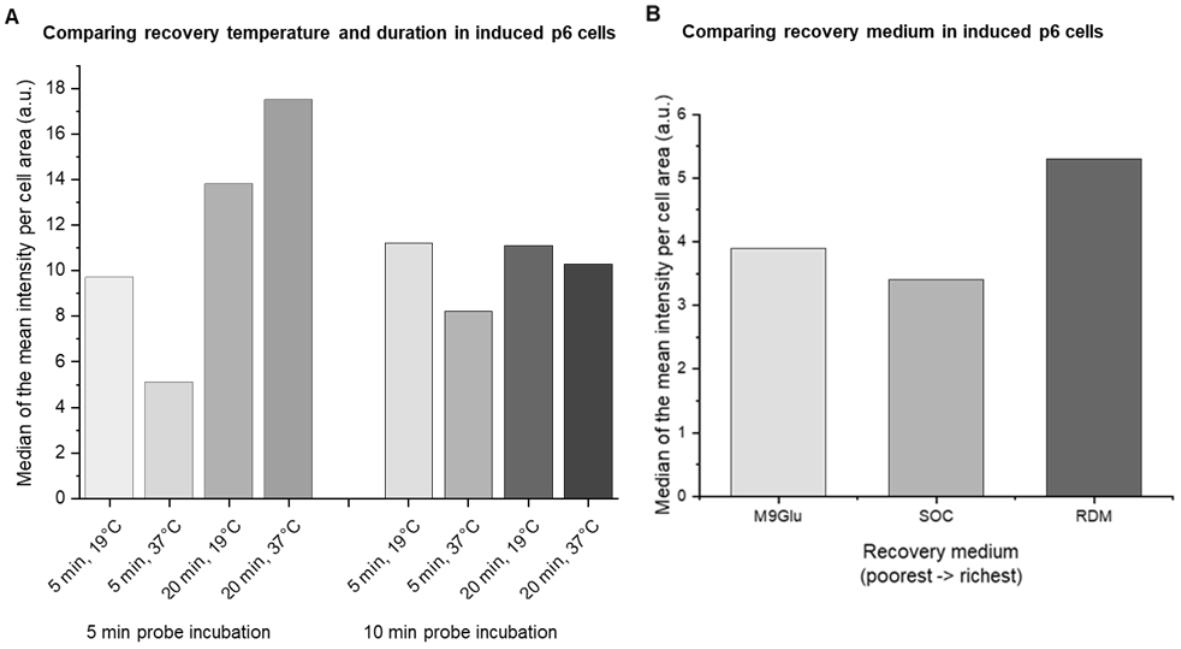


Figure 3.25 – Summary comparisons of recovery parameters post-electroporation in induced p6 cells with 50 nM 8mer-2x647N.
(A) Column plot of the median populational mean intensity per cell area, for cells incubated with probes for 5 (left series) and 10 min (right series), and then recovered for either 5 or 20 min at 19°C or 37°C.
5 min incubation, loaded cell medians: 9.7 units 5 min/room temperature, 5.1 units 5 min/ 37°C, 13.8 units 20 min/room temperature, 17.5 units 20 min/ 37°C.
10 min incubation, loaded cell medians: 11.2 units 5 min/room temperature, 8.2 units 5 min/ 37°C, 11.1 units 20 min/room temperature, 10.3 units 20 min/ 37°C.
(B) Column plot of the median populational mean intensity per cell area, for cells treated under standard protocol but recovered in either M9Glu, SOC or RDM media. Loaded cell medians: 3.9 units M9Glu, 3.4 units SOC, 5.3 units RDM.

Considering the preference for longer recovery post-electroporation, richer media in recovery were also anticipated to further improve post-electroporation cell viability, possibly probe retention and spread. Interestingly of the three media tested, M9Glu (poorest) and rich defined medium (RDM, richest) both showed slightly higher median mean intensities (Figure 3.25B) and loaded cell populations (Appendix 1.1) than the more commonly used SOC medium in induced p6 cells. Despite highest visualised loading in RDM cells, SOC was the most conducive medium experimentally – much easier to retain a sizeable cell pellet during post-recovery washing phases, and most consistent in pellet sizes and retention across multiple experiments. Since the overall effects were not significant, but experimentally SOC was preferable, SOC was selected as the recovery medium of choice.

Imaging parameters

Imaging parameter optimisation constitutes a further final aspect of technical refinement, with the goal of extending photon budget and maximising signal timescales. Key means of extending photon budget and signal acquisition are through modulating and optimising laser power and exposure time (Appendix 1.2).

Using the ONI NanoImager (NIM) for all *in vivo* smFISH experiments, a number of laser powers and HILO angles were tested; a HILO angle of 51° was quickly determined as suitable for sufficient imaging intracellular imaging depth. As for laser power, induced p6 cells with 50 nM 8mer-2x647N were imaged at 50 ms exposure, at 5%, 8% or 10% laser powers, corresponding to ~2, 4 and 10 mW. Foci only became highly resolvable from 8% laser power upwards. With negligible difference between 8-10% and the non-linearity of the laser activation systems, 10% was deemed the most robust choice for consistent illumination, despite the higher propensity for rapid bleach down.

Determining an appropriate exposure acquisition time is important for gathering single molecule information of interest; exploring different exposure times can shift the balance of single molecules observed, localised and tracked in cells.

From the underpinnings of superresolution microscopy as outlined in the Introduction, the exposure time in frame acquisition is a key determinant in the ability to fit PSFs to bright foci, hence to the different types of molecules that may be reliably observed. By this premise, a shorter exposure time is required to resolve faster diffusing molecules, and conversely, a longer exposure time is required to more accurately localise and track the same slower-diffusing or bound molecule. Indeed, if too high an exposure time is used for a sample that has

a high population of fast-diffusing molecules, the lack of molecular resolution would simply appear as higher background autofluorescence (motion blurring), which is also undesirable.

Exposure times between 20 and 250 ms were tested; 20 ms was kept as a lower bound for the capture of fast diffusing species above exposure times which could potentially capture free dye species. Following the premise that a longer and bulkier target, such as chromosomal mRNAs, would diffuse more slowly and hold several complex conformational states, 100 ms exposure and over were able to capture fluorescent signals in more cells later into acquisition (>50 seconds). 50 ms exposure time was selected as a standardized exposure time across most experiments: enabling a sound balance between capturing faster diffusive species without excessive blurring, allowing for high enough resolution of slower-moving/'bound', and in line with published protocols (Aigrain et al., 2015; Plochowietz et al., 2017; Sustarsic et al., 2014).

Stroboscopy, or alternating laser activation with dark interval periods, is one further option for extending photon budget and the real-time durations of tagged molecules in the live cell. In stroboscopy, coupling laser excitation to different camera exposure times enables much longer acquisition timescales by reducing laser exposure to the fluorophores: for example, a 10 ms laser 'ON' combined with a consecutive 200 ms laser 'OFF'. Since it was not possible to have alternating, different exposure times on the NIM, dark interval acquisitions were trialled as a proxy: for example, a cell sample electroporated with fluorogenic probes could be imaged for 50 ms with a 10 mW 638 nm activated laser, and then dark for 250 ms (1:5 ratio). This form of alternating dark interval acquisition enabled substantially longer acquisitions and indeed extended ATTO dye photon budget, but subsequent analytical processing was much more laborious. In the longer term, further dark interval and stroboscopy imaging of samples would yield useful insights in working towards reproducibly observing full transcription cycles, but remains beyond scope for the time being.

Overall, the majority of technical experimentation demonstrated the robustness of existing standardized protocols for the internalisation of labelled biomolecules into live cells for single molecule study. Main experimental adjustments related to the choice of growth medium in competency, and the imaging parameters applied for the capture of more transient molecules, mRNAs, with variable mobilities in the live cell.

3.5.5 Cell viability

Beyond procedural parameters, no level of optimisation is helpful if the cells are unable to survive. To test the viability of M9Glu cells, with and without electroporation, cells from the different strains were viability-tested at the single cell level.

To elucidate the cell fitness on a single cell level, brightfield microscopy growth experiments were performed for wildtype, induced p6 and gT7 strains, in line with published electroporation-mediated single molecule viability studies (Aigrain et al., 2015; Anne Plochowietz, 2015; Sustarsic et al., 2014).

Cell samples were treated as standardized for an *in vivo* smFISH experiment, but when added to the agarose pad were left unperturbed for 60 min, and imaged every 15 min to allow for observations of replication. Tiled brightfields of representative FOVs from 0-60 min at 15 min intervals show that, indeed, cell growth and division is sustained across all strains tested (Figure 3.26). Cell phenotypes were classified as growing and divided (red), in stasis (no apparent growth, but structurally sound, orange), and damaged (ruptured, fragmented, any characteristics suggesting cell death, black), by comparing the time 0 FOV with the 60 min FOVs. Correspondingly, over 90% cells were classified as growing and dividing across all strains.

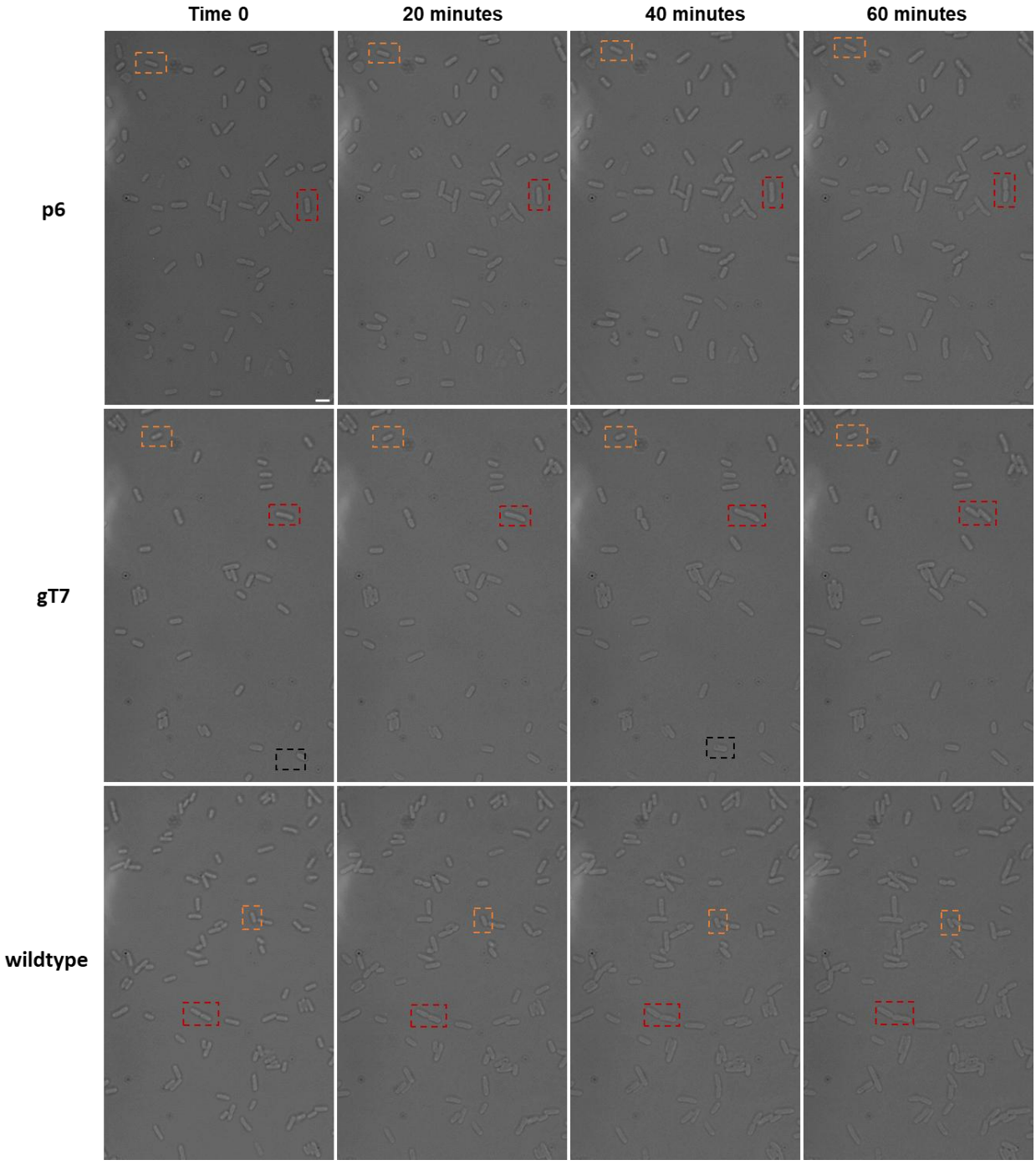


Figure 3.26 – Observing electroporated cell fitness by single cell growth time series. Representative FOVs of time points, 0, 20, 40 and 60 min for p6, gT7 and wildtype M9Glu competent cells electroporated with 50 nM 8mer-643/q1 probes.
Growing/dividing cells are highlighted in red, static in orange and compromised in black; scale bar – 3 µm.

3.6 Passive internalisation of labelled biomolecules

A key control established in the various publications utilising the method published in Aigrain et al., 2015 was the non-electroporated control: where sample cells were incubated with the fluorescently-labelled target molecule, but no electroporation was performed. The goal of this control was to demonstrate the efficacy of electroporation in internalising labelled constructs, particularly biomolecules of a larger scale – henceforth, an anticipated low to no fluorescent signal in non-electroporated cell samples.

Interestingly, initial attempts to demonstrate such non-electroporated controls were unsuccessful: even without electroporation, there seemed to be significant signal present in cells indicating internalised probes, apparent across all strains, wildtype and tag strains alike (Figure 3.27). Fewer cells compared to electroporated samples comprised fluorescent signals, but there were signals nonetheless, and even the molecule characteristics appeared alike, albeit on shorter timescales.

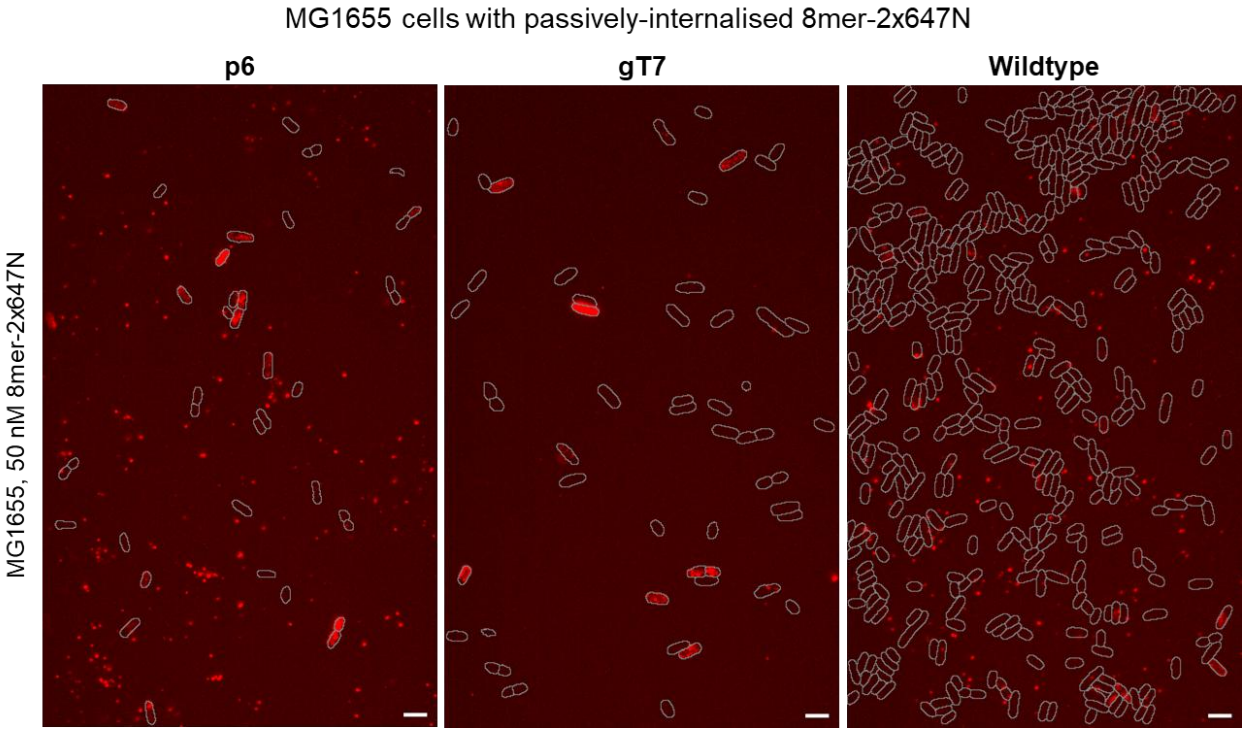


Figure 3.27 – The non-electroporated control – a reproducible observation of apparent passive internalisation.
Maximum intensity projections of 100-1000 frames (~45 s) representative FOVs of wildtype,, induced p6 and gT7 MG1655 M9Glu competent cells incubated but not electroporated with 50 nM 8mer-2x647N probes, prepared and imaged under the same conditions as the electroporated samples.
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

There-in-lay a secondary focus of interest: what was causing cells to passively internalise labelled ssDNAs, and could this be an effective, less stressful means of conducting single molecule *in vivo* studies?

3.6.1 Technical exploration of fluorescently-labelled ssDNA passive internalisation

To identify determinants of passive internalisation, technical experiments into the effects of competency, probe incubation temperature and length dependence of fluorescently labelled ssDNAs were conducted. As in 3.5.4, technical experiments are verbally summarised here with some exemplar figures, following the same ensemble analytical pipeline; Appendix 1.3 presents corresponding figures for data described but not shown here.

Establishing a suitable probe for technical parameter comparison

Preliminary observations of passive internalisation occurred with 8mer-2x647N – unsurprisingly as very short, single-stranded sequences, even when labelled with small fluorophores. It is known that *E. coli* do have various pathways for extracellular DNA uptake (Sun, 2018), particularly when under stress, although *E. coli* themselves are considered to not possess natural competence. The 8mer-2x647N probe, however, would not constitute a suitable target for robust exploration of passive internalisation variables, due to need for an engineered sequence, the likely transience of hybridisation and probe fluorogenicity. In selecting a more appropriate probe for passive internalisation comparisons, the 8mer-2x647N acted as a baseline, ensuring observations were analogous to those reproducibly made during *in vivo* smFISH experimentation.

To select an appropriate probe for passive internalisation experiments, LB and M9Glu competent cells were incubated with 100 nM 8mer-2x647N, EUB338-647N or EUB338-Cy3B probes for 10 min at room temperature, washed at room temperature and imaged as standardized for *in vivo* smFISH experiments (Appendix 1.3). Note that no recovery period was conducted following probe incubation, as no electroporation was performed.

Very few LB cells (~10% per probe sample) showed internalised signal generally for all three probes, even the very short 8mer-2x647N, suggesting that those cells comprising signal could either form a specific permissive subpopulation capable of small molecule uptake, or could be compromised, internalising probes due to ruptured membranes. M9Glu cells comparatively appeared to have ~90% cells containing some signal with 8mer-2x647N, ~50% EUB338-647N-internalised cells and ~90% EUB338-Cy3B-internalised cells, corroborated by

the median mean intensity measurements, indicating that M9Glu competency is likely a key determinant of passive internalisation.

Singly-labelled EUB338 probes with 1x647N or 1xCy3B fluorophores were chosen as possible suitable standards by which to gauge the extent of passive internalisation, because of their consistent use case as fixed FISH probes, and their well-elucidated binding target on the prokaryotic 16S rRNA, the 30S subunit of the ribosome (R. I. Amann, 1990). At 18-nt-long, the single-stranded EUB338 probe stably hybridises with its 16S rRNA complementary target, and the nature of the single label means that all signals are visible until they are bleached. Of course, irrespective of target binding, any internalised singly-labelled EUB probe would likely be visible, which was ideal for the purpose of holistically quantifying the extent of passive internalisation; but the prospect of a specific target opened potential avenues for further, more precise study later on.

The EUB-647N probe was selected as the probe for standardized use in passive internalisation technical experiments, due to its more distinguishable bright foci than the other tested probes. Although EUB-Cy3B is a commonly used probe (Chatzimichail et al., 2024; Farrar et al., 2024), its more generally diffuse fluorescent appearance in cells presented a greater challenge in identifying "background" versus probe-related signal. Preferable to EUB-Cy3B consequently, and with ATTO647N-labelled probes optimised in various ways for *in vivo* smFISH experiments, the ATTO647N probe was the most appropriate choice.

Passive internalisation – cell viability

Although cells which have passively internalised labelled ssDNAs have not undergone stress perturbations for internalisation (especially chemical, electrical), their state of competency, consecutive washes and the lack of a recovery period may bring into question the longer-term fitness of the cells.

The cell viability of competent M9Glu and LB wildtype cells with 100 nM passively-internalised EUB-647N probe was tested in the same manner as applied for the *in vivo* smFISH strain samples in 3.5.5 (Appendix 1.3). Comparative to electroporated cell observations, over 90% cells appeared to be growing and dividing, although notably, LB grown cells were much harder to stably characterise.

Cell state and temperature dependency of passive internalisation

Having selected a standard probe for the comparison of condition dependency on passive internalisation, it was possible to explore the impact of two key proponents in passive internalisation: competency state, and temperature of probe incubation.

To test the relative importance of the competency state and probe incubation temperature on passive internalisation of 100 nM EUB-647N probe, wildtype M9Glu and LB freshly-grown versus competent cells were incubated with 100 nM probe at either 4°C, 19°C or 37°C.

Fresh cells

Fresh cells were generally impervious to probe internalisation, likely due to a combination of high membrane/cell wall robustness from fresh growth and supplied nutrients, alongside higher autofluorescent background from fewer overall washing steps (representative FOV maximum intensity projections, Figure 3.28).

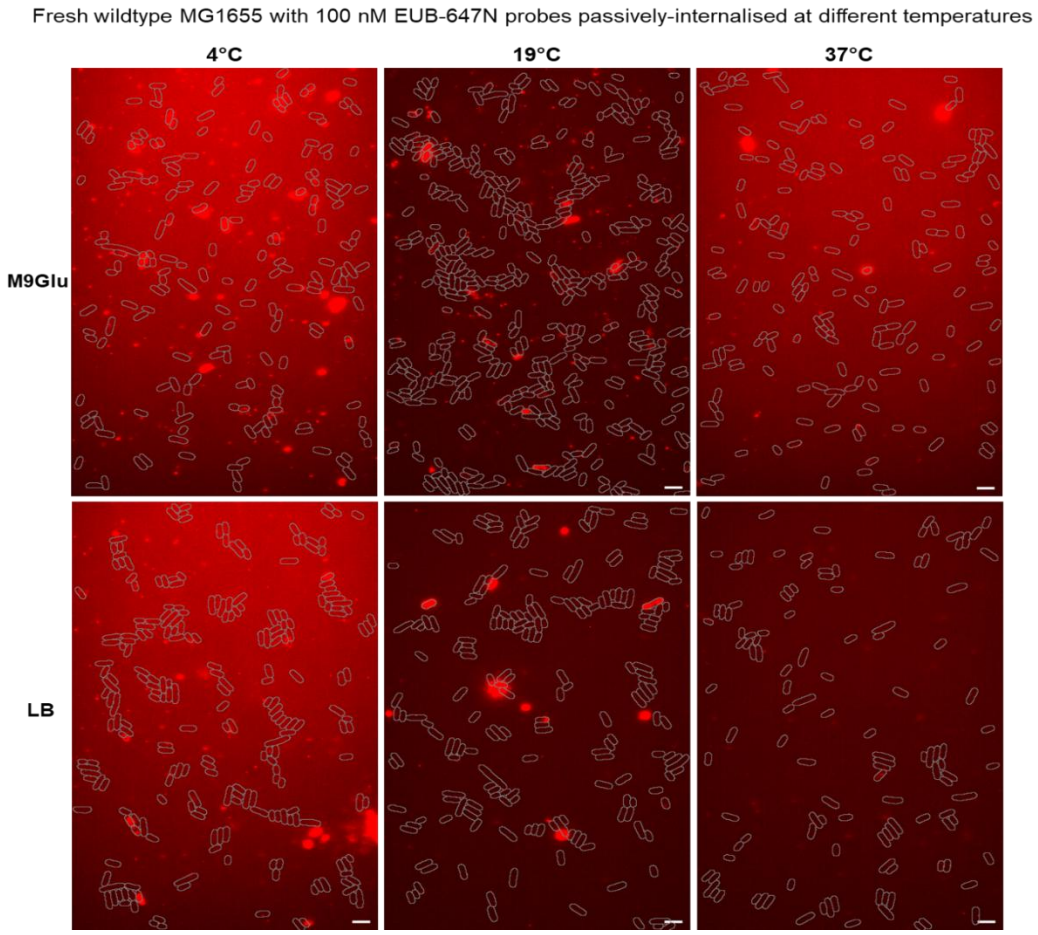


Figure 3.28 – Testing passive internalisation propensity of 100nM EUB-1x647N probes incubated at different temperatures with M9Glu and LB fresh wildtype MG1655 cells. Maximum intensity projections of 1-2000 frames (~100 s) representative FOVs of wildtype MG1655 fresh cells incubated with 100 nM EUB-1x647N incubated for 10 min at 4°C, 19°C and 37°C.
Top panel shows M9Glu competent cells, and the lower panel shows LB competent cells. Scale bars – 3 µm, false-colour in red, and overlaid cell segmentation outlines.

Intensity measurements are hence likely confounded by measurements related to freely diffusing probes in solution and background autofluorescence overlaying cells (corresponding plots in Supplement 3.3). Higher extracellular probe concentrations were thought to be due to higher cell robustness meaning fewer probes internalised, and potentially ineffective washing.

Competent cells

M9Glu competent cells comparatively showed much higher median mean intensities per cell area across all temperatures for probe incubation tested – 8.6 units for 4°C, 6.1 units for 19°C and 7.3 units for 37°C – and very similar whole cell population distributions of mean intensities. Loaded cell populations corresponded to ~84% for 4°C, ~79% for 19°C and 83% for 37°C.

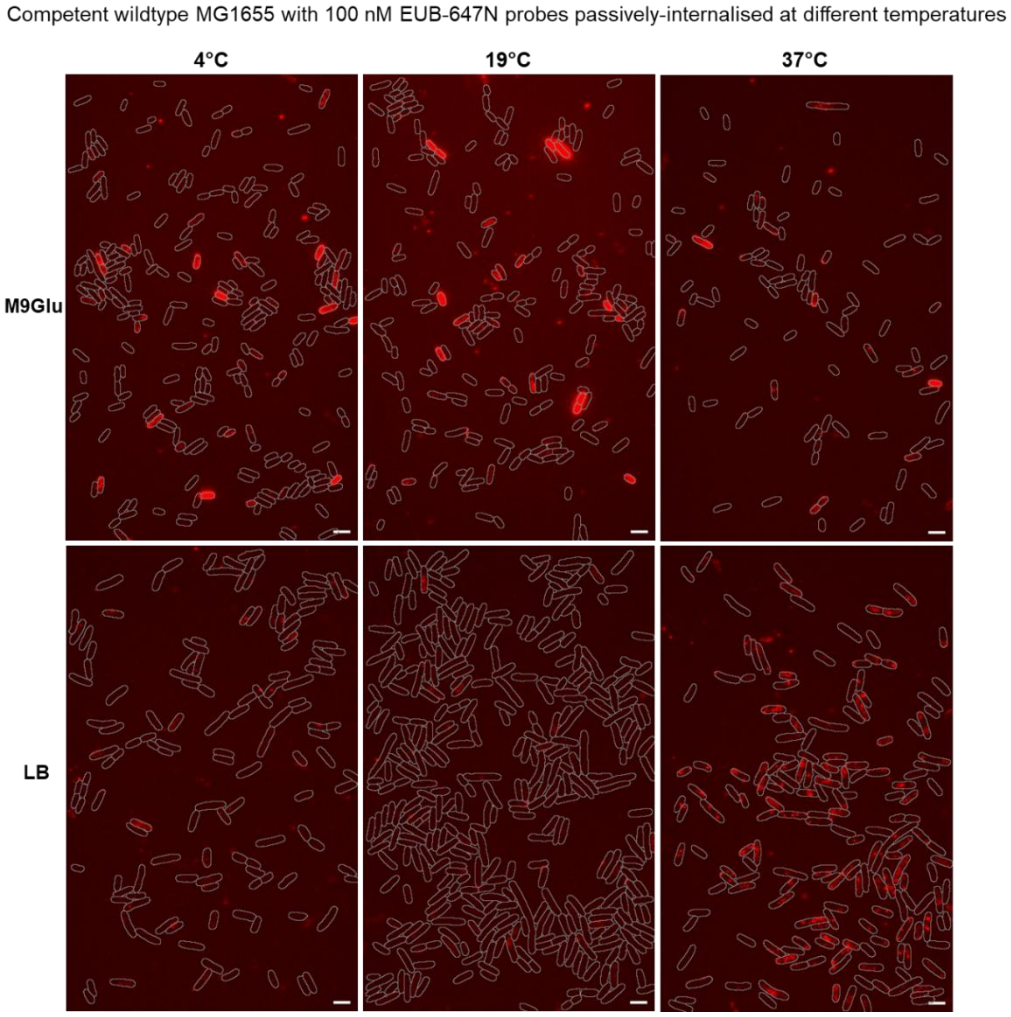


Figure 3.29A – Testing passive internalisation propensity of 100nM EUB-1x647N probes incubated at different temperatures with M9Glu and LB competent wildtype MG1655 cells. Maximum intensity projections of 1-2000 frames (~100 s) representative FOVs of wildtype MG1655 competent cells incubated with 100 nM EUB-1x647N incubated for 10 min at 4°C, 19°C and 37°C.
Top panel shows M9Glu competent cells, and the lower panel shows LB competent cells. Scale bars – 3 µm, false-colour in red, and overlaid cell segmentation outlines.

Despite the 19°C having the lowest median mean intensity of the three temperatures tested, somewhat perplexingly, the 19°C sample also had the longest tail in whole cell distributions, pertaining to higher numbers of cells with higher individual intensities. The respective up-shifted 4°C and 19°C mean intensity populations could be due to stark temperature gradient formation in the 37°C case (as competent cells are rapidly thawed and then heated with probes present); and at 4°C, perhaps a more ambient molecular dynamic exchange with 4°C acting as a slight cold shock for thawed cells. Nevertheless, **the defining factor in passive probe uptake appears to be the state of the cell, i.e. competent and therefore preconditioned for permeability and membrane poration**, over temperature gradient formation ("accidental heat-shock") and variation during probe incubation.

Interestingly, competent LB cells also appeared largely impervious to the passive internalisation of the EUB-647N probe across different temperatures, with the exception of 37°C giving ~50% visually loaded cells in the representative FOV (Figure 3.29A). M9Glu competent cells seemed much more permissive to probe internalisation across all temperatures (Figure 3.29B,A). Corroborating the observation in the representative FOV, very tight loaded cell mean intensity distributions in 4°C and 19°C LB competent cell samples (~75% 4°C and ~72% 19°C loaded; medians 2.9 units and 3.5 units respectively), showed negligible internalisation overall, but the 37°C sample had a median mean intensity of 10.9 units, and ~93% loaded cells (Figure 3.29B,B). This was further supported by the whole cell population distribution of mean intensities, where for 4-19°C, all cell intensities were clustered around the 0-20 units mark, whereas for the 37°C sample, a broad cell population had intensities between 0-30 units, and smaller populations up to as high as 65 counts (Figure 3.29B).

For ease of experimentation, all further passive internalisation experiments were conducted at room temperature with competent M9Glu and LB cells unless otherwise specified.

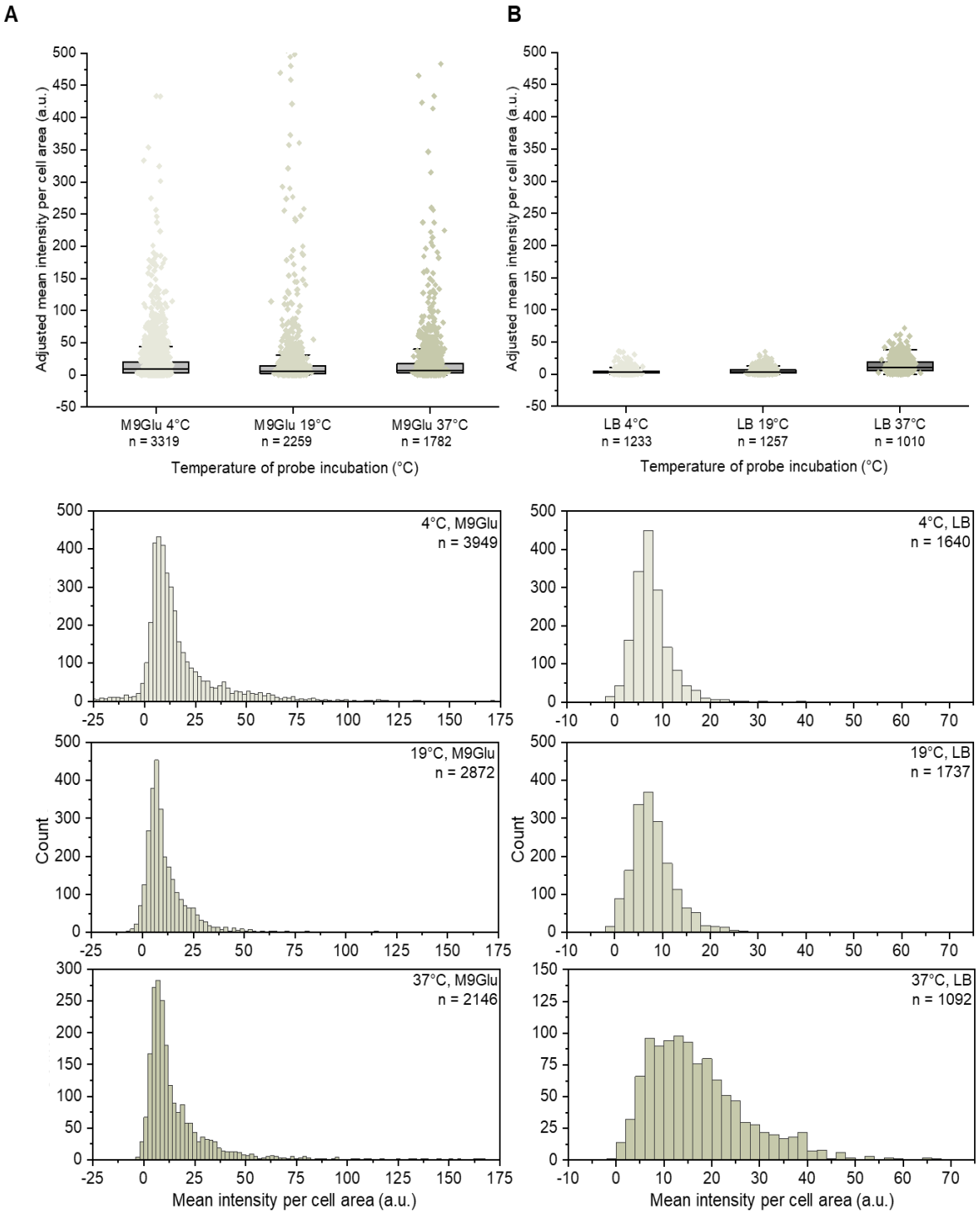


Figure 3.29B – Testing passive internalisation propensity of 100nM EUB-1x647N probes incubated at different temperatures with M9Glu and LB competent wildtype MG1655 cells: population-level intensity distributions.
(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for M9Glu (left) and LB (right) competent wildtype cells with 100 nM passively internalised EUB-1x647N incubated at 4°C, 19°C and 37°C. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of empty cells.
Loaded cell medians: M9Glu: 8.6 units 4°C, 6.1 units 19°C, 7.3 units 37°C; LB: 2.9 units 4°C, 3.5 units 19°C, 10.9 units 37°C.
(B) Comparative histogram distributions of the whole cell populations with different probe incubation temperatures for M9Glu and LB competent cells; bin size = 2.
n gives the number of cells per sample (N = 2 experiments).

DNA-Length dependence of passive internalisation

As priorly observed, the 8mer-647N probes were much more readily internalised in both LB and M9Glu competent cells than the 18 nt EUB338 probes. It was of subsequent interest whether there was an upper bound for passive internalisation of ssDNAs.

To test the length dependence of ssDNA passive internalisation, 100nM TTT-multimers labelled with a 5' ATTO647N only of lengths 8 nt, 20 nt, 40 nt and 75 nt were incubated for 10 min at 19°C with M9Glu competent cells, washed and imaged as standardised. Interestingly, the passive internalisation efficiency seemed to trend upwards in M9Glu competent wildtype cells as the ssDNAs gained length. Loaded cell populations across all tested FOVs were ~75%, ~94%, ~96% and ~97% respectively (Figure 3.30), corroborated by median mean intensity measurements. In the case of LB competent wildtype cells (Supplement 3.4), very few molecules were internalised, consistent across all lengths; again, indicating the media dependency and competency states as defining factors in passive internalisation.

At the upper bound of 75nt, M9Glu competent wildtype cells were highly permissive to uptake; it would therefore be of great interest to see if this extends up to small protein fragments, such as sigma factors.

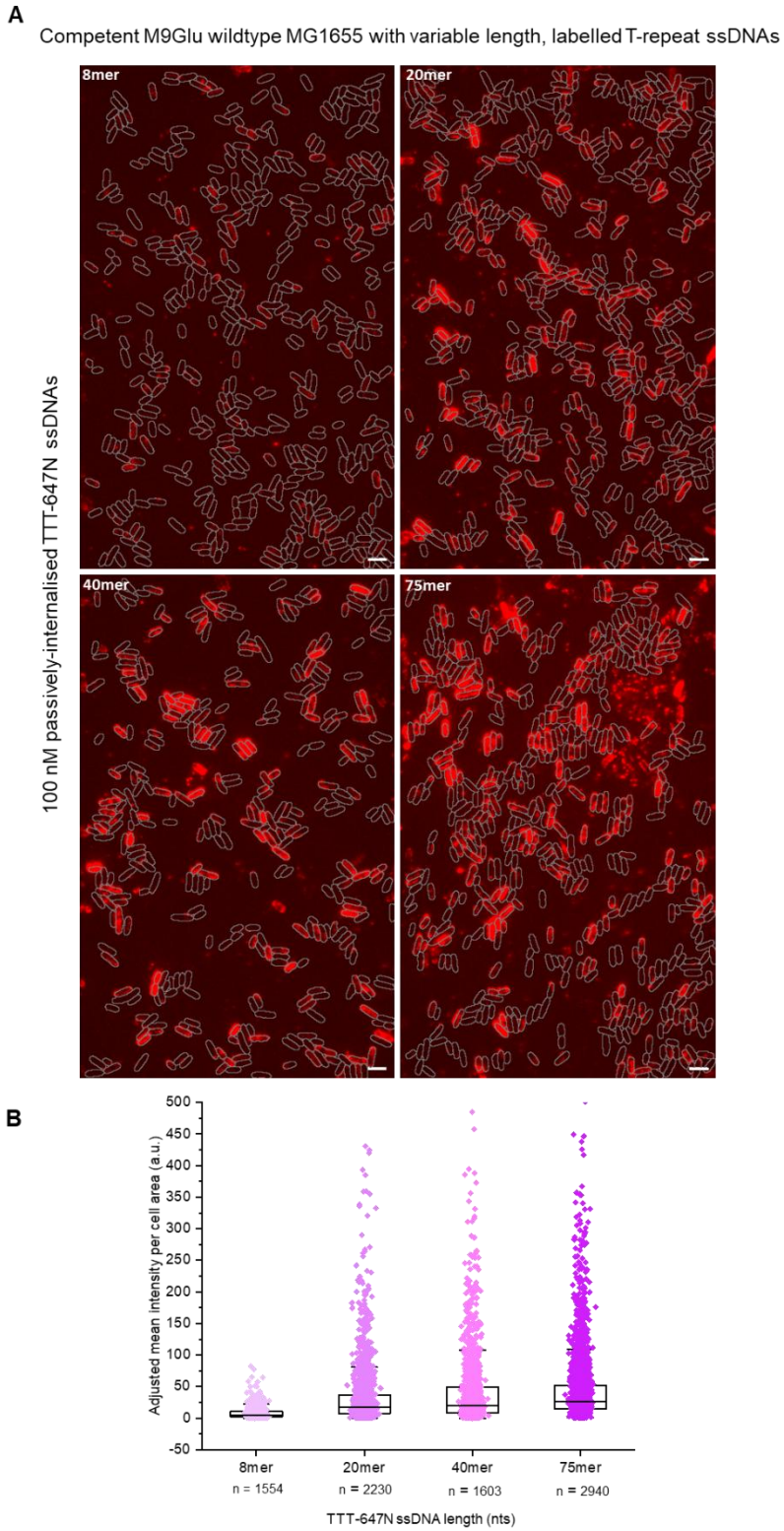


Figure 3.30 – Testing passive internalisation propensity of different lengths of TTT-647N ssDNAs with M9Glu competent wildtype MG1655 cells.
(A) Maximum intensity projections of 1-2000 frames (~100 s) representative FOVs of wildtype MG1655 competent cells incubated with 100 nM 8mer, 20mer, 40mer and 75mer TTT-647N ssDNAs.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for M9Glu competent wildtype cells with 100 nM passively internalised variable length TTT-647N ssDNAs. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of empty cells.
Loaded cell medians: 4.7 units 8mer, 16.8 units 20mer, 19.7 units 40mer, 26.5 units 75mer.
Scale bars – 3 µm, false-colour in red, and overlaid cell segmentation outlines; n gives the number of cells per sample (N = 1).

3.7 Discursive summary

The work in this Chapter confirmed the ability to exogenously deliver short, fluorogenic ssDNA probes into live *E.coli* for the purpose of single molecule tagging. Both with and without induced internalisation, targeted ssDNA probes were internalised into cells, with the majority population of cells remaining viable. The design of fluorogenicity for the short probes enabled a PAINT-like capacity in acquisition of distinguishable single molecules, which in turn allowed for single molecule localisation and tracking, and downstream molecular mobility analyses.

Aside from positive results observed in tag strain cells electroporated with fluorogenic probes – and a decrease in fluorescence at 60 min with rifampicin exposure, indicating that the probes mainly hybridise to circulating RNAs – there is the caveat of lacking a robust control for probe-complex specificity. Whilst the absence of the engineered tag site in cells, hence wildtype cells, make for a necessary control, the complete quenching of probes would be hypothetically unlikely, surplus to *in vivo* observations of fluorescent foci in wildtype cells with internalised probes.

There are a number of likely explanations for this: for one, an 8-nt length is very short and hence highly prone to having numerous off-target sites, providing possible specific binding to off-target complementary RNAs. More likely, though, are two separate possibilities: the implantation and trapping of the ATTO dyes in the cell membrane, and the sequestration of ssDNA probes by the single-stranded binding protein (SSB).

Well-characterised in *E. coli* for the high-affinity binding of circulating ssDNAs, essential SSB stabilizes ssDNA intermediates in manifold genomic maintenance processes, including DNA replication, recombination and repair, and degradation (Antony & Lohman, 2019; Meyer & Laine, 1990). It is possible that the 8mer probes may transiently bind to SSB (and potentially SSB-cooperative actors (Shereda et al., 2008) like the Rec protein family in homologous recombination repair mechanisms (Han et al., 2006; Pham et al., 2023)), but it is unlikely that the off-target longer and more static foci are a result of protein binding. Homotetrameric SSB has been shown to exhibit a number of reversible binding modes enabling high-affinity capture of 35-70 nt-long ssDNAs (and 40 nt intermediaries particularly at 37°C)(Bujalowski & Lohman, 1986), but no stable binding of ssDNAs sub-20 nts, and all with complex energetic dynamics (Suksombat et al., 2015). Considering this transience of protein-related sequestration, it is hence also unlikely that cellular degradation machinery definitively acts upon the probe pool on the time scale of the experiments, and potentially even beyond. Probe longevity is probably

also augmented by the potentially protective capacity of the 5'/3' dye labels in preventing ssDNA termini recognition.

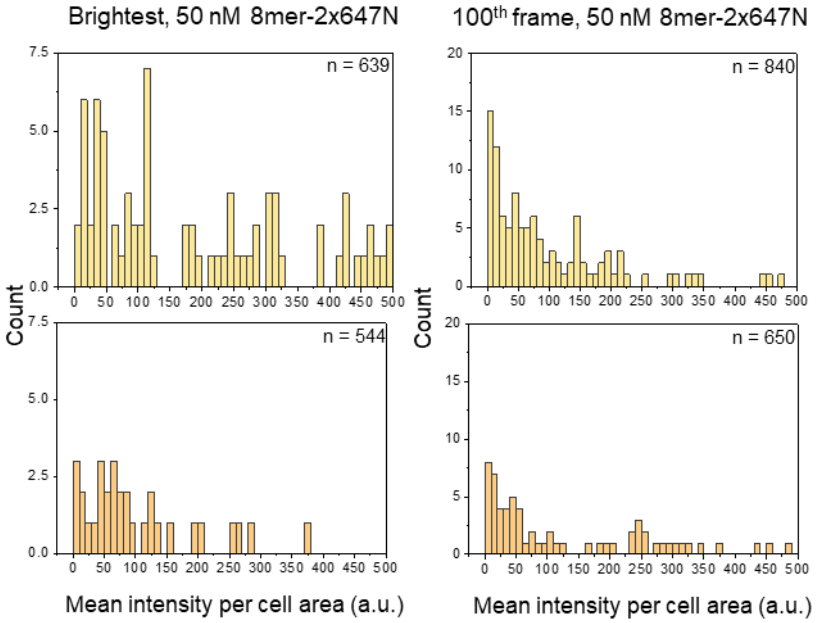
As such, the key "sticking point" when it comes to unspecific signal in wildtype cells most likely relates to membrane interactions by the ATTO dyes, particularly ATTO647N known for its well-typed 'stickiness' (Salcedo-Sora et al., 2021; Zanetti-Domingues et al., 2013). The ATTO dye family designed to reduce aggregation propensity (*How to Choose a Label*, ATTO-TEC), and in the 2xATTO647N self-quenching form for fluorogenicity, off-target signals are most anticipated to be incomplete probe photo-bleaching states for rapidly diffusing signals, and membrane-binding for static foci. The ATTO643 dye, thus, developed to retain the photostability and brightness characteristics of ATTO647N but reduce off-target binding propensities, makes for a more robust dye in probe labelling, and indeed has been shown here to be the best option for engineered *in vivo* smFISH experimentation.

Notwithstanding the specificity challenges, there are, of course, all the other hurdles to jump through that come with SRM: trade-offs to be made between fluor choice for specificity/density/brightness with photon budget for spatial resolution/variable depth/acquisition speed/photodamage and detection efficiency (microscope constraints)(Schermelleh et al., 2019). For the purpose of engineered *in vivo* smFISH, interpretations of molecular observations are further explored using the original 8mer-2x647N as a baseline and 8mer-643/q1 for more robust readouts (Chapter 4). But, there is always room for improvement; and the region of greatest interest for increasing signal specificity is to continue probe and dye-label development.

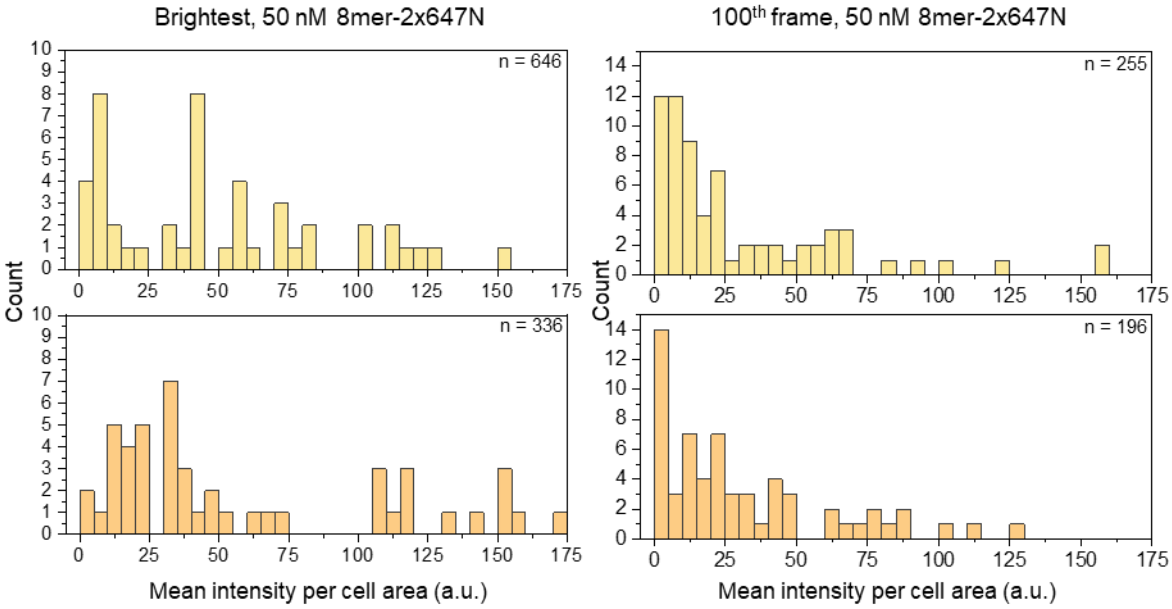
Overall, iterative technical experimentation demonstrated that the standardised method documented by Aigrain et al. holds true for this use-case (Aigrain et al., 2015). Variations in probe incubation duration, recovery durations, temperature and media all made negligible impact on the overall experimental outcome, but certainly aided towards fine-tuning and verification of methodological parameters. The sole exception to observations across the board (Aigrain et al., 2015; Crawford et al., 2013; Sustarsic et al., 2014) was the determinative influence of cell background, growth parameters and competency state – leading to the observation of passive internalisation of labelled ssDNAs. A reproducible finding, and most defined by competency, passive internalisation of labelled ssDNAs opens a potentially promising avenue for further *in vivo* single molecule studies with even less perturbation to the live cell – explored further in Chapter 5.

Supplement 3

3.1 – A proxy for specific signals in BW27784 tag strains electroporated with 50 nM8mer-2x647N probes: histogram distributions for wildtype-corrected loaded p6 and gT7 cells. Bin size = 10; n gives the number of cells per sample (N = 3, 2 experiments for p6 and gT7 respectively).



3.2 – A proxy for specific signals in MG1655 tag strains electroporated with 50 nM8mer-2x647N probes: histogram distributions for wildtype-corrected loaded p6 and gT7 cells. Bin size = 10; n gives the number of cells per sample (N = 3, 2 experiments for p6 and gT7 respectively).

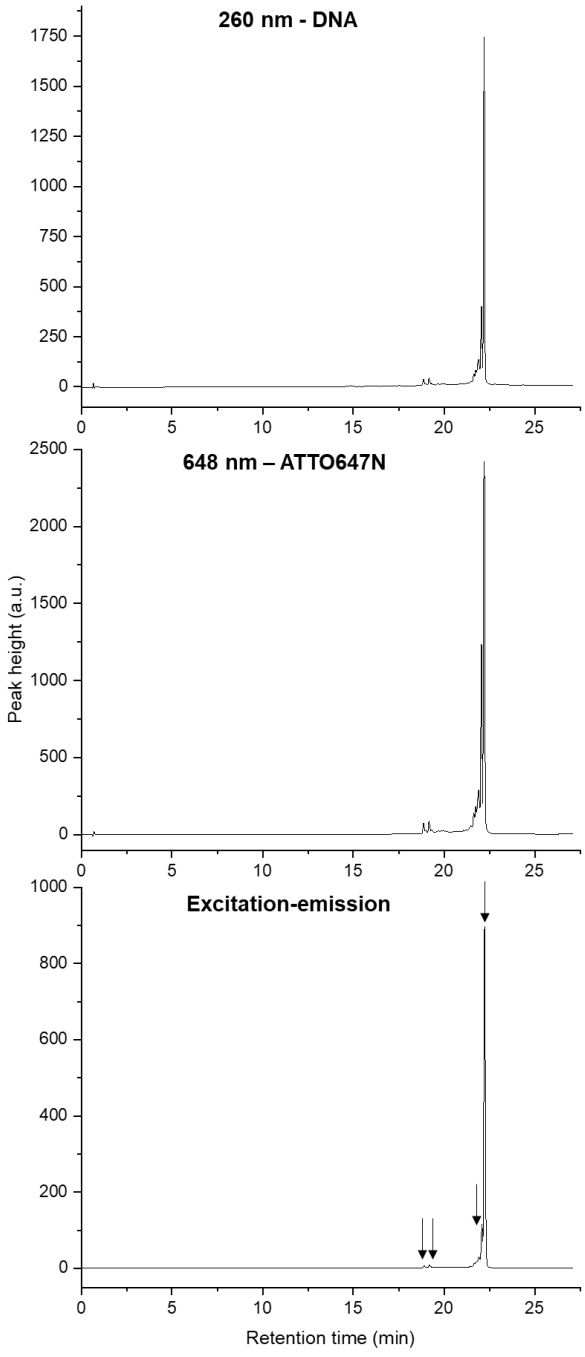


3.3 – Reverse-phase HPLC chromatograms of the in-house purification of commercial 8mer-2x647N.

(A) Fractionation of 8mer-2x647N at 260 nm wavelength, corresponding to visualizing ssDNA.

(B) Fractionation of 8mer-2x647N at 648 nm wavelength, corresponding to the excitation of ATTO647N.

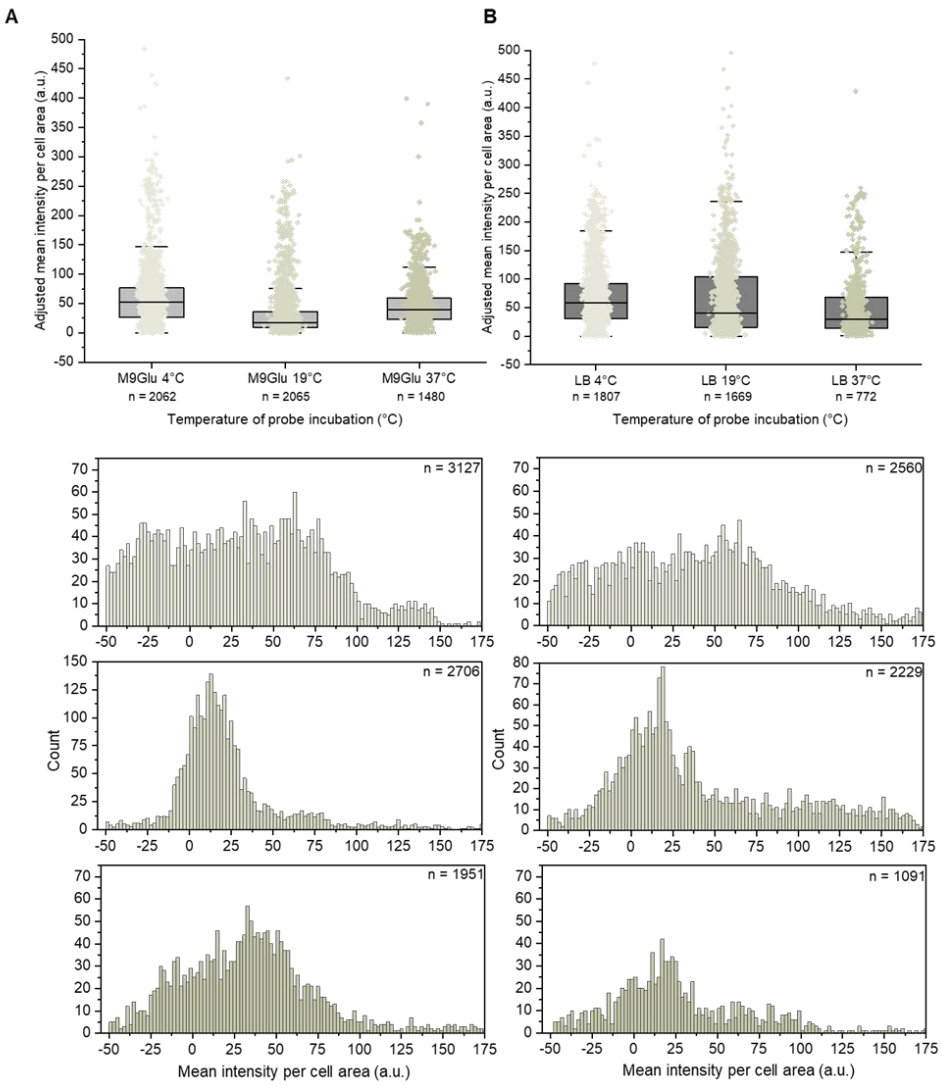
(C) Combined excitation-emission curve for the 8mer-2x647N sample.



3.4 – Testing passive internalisation propensity of 100nM EUB-1x647N probes incubated at different temperatures with M9Glu and LB fresh wildtype MG1655 cells: population-level intensity distributions.

(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for M9Glu (left) and LB (right) fresh wildtype cells with 100 nM passively internalised EUB-1x647N incubated at 4°C, 19°C and 37°C. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of empty cells. Loaded cell medians: M9Glu: 42.4 units 4°C, 17.6 units 19°C, 39.1 units 37°C; LB: 59 units 4°C, 40 units 19°C, 29.6 units 37°C.

(B) Comparative histogram distributions of the whole cell populations with different probe incubation temperatures for M9Glu and LB competent cells; bin size = 2.
n gives the number of cells per sample (N = 2 experiments).



Chapter 4 – Implications on transcription initiation: applying engineered *in vivo* smFISH

This Chapter demonstrates the application of engineered *in vivo* smFISH in tagging a native promoter site in the *E. coli* genome, under the *rpoB* gene promoter, and assesses observations at the single molecule level of all three tag strains, p6, gT7 and *rpoB*-tag as per their biological implications.

4.1 Introduction

The *E. coli* 4.6 Mb genome (Blattner et al., 1997), extensively mapped and so far with 4623 genes annotated (Gama-Castro et al., 2016; Ghatak et al., 2019; Hu et al., 2009; O'Leary et al., 2016; J. Zhou & Rudd, 2013), is the best elucidated prokaryotic model, and a core organismal model more broadly. Yet despite such dedicated description of gene networks and functionality, 34.6% unique genes remain experimentally uncharacterised (Ghatak et al., 2019). At the transcriptomic level, iterative progress across the fields of intracellular organisation, phase condensates, mRNA trafficking and single molecule transcription make the characterisation of lesser-known genes and their mRNA products an enticing opportunity – towards developing a more holistic understanding of gene expression in the live *E. coli* cell.

Having defined a usable baseline in engineered non-native plasmid and chromosomal systems with engineered *in vivo* smFISH, one such potential avenue for furthering the understanding of single transcriptional events and post-transcriptional outcomes is by tagging a target gene endogenously in the live cell. Design rationales for candidate target genes are presented in Figure 4.1, whereby the target gene product is depicted alongside the target gene genomic location. The candidates in 4.1 are all essential genes: the RNAP β clamp core subunit, σ^S stress response factor, and the 16S rRNA of the 30S ribosome subunit. These are of interest for their implications in cellular gene regulation, particularly in understanding the spatiotemporal profiling and expression of essential products in co-regulation of transcription and translation (Bakshi et al., 2015; Gerber et al., 2023; Le Nguyen et al., 2022; Surovtsev & Jacobs-Wagner, 2018). The *rpoB* candidate is the focus of this Chapter.

As reviewed in the Introduction, there are many methods that enable the visualization and behavioural characterisation of mRNAs in the live cell, typically using arrayed aptamer constructs and fluorescent fusions. Whilst reliable site-specificity can be achieved with these methods, the engineered components have been shown to disrupt transcription of the target RNAs, have largely undefinable impacts on downstream translational processing, and available fluorophore photostability/brightness often constrain single molecule resolution and acquisition duration (Gerber et al., 2023).

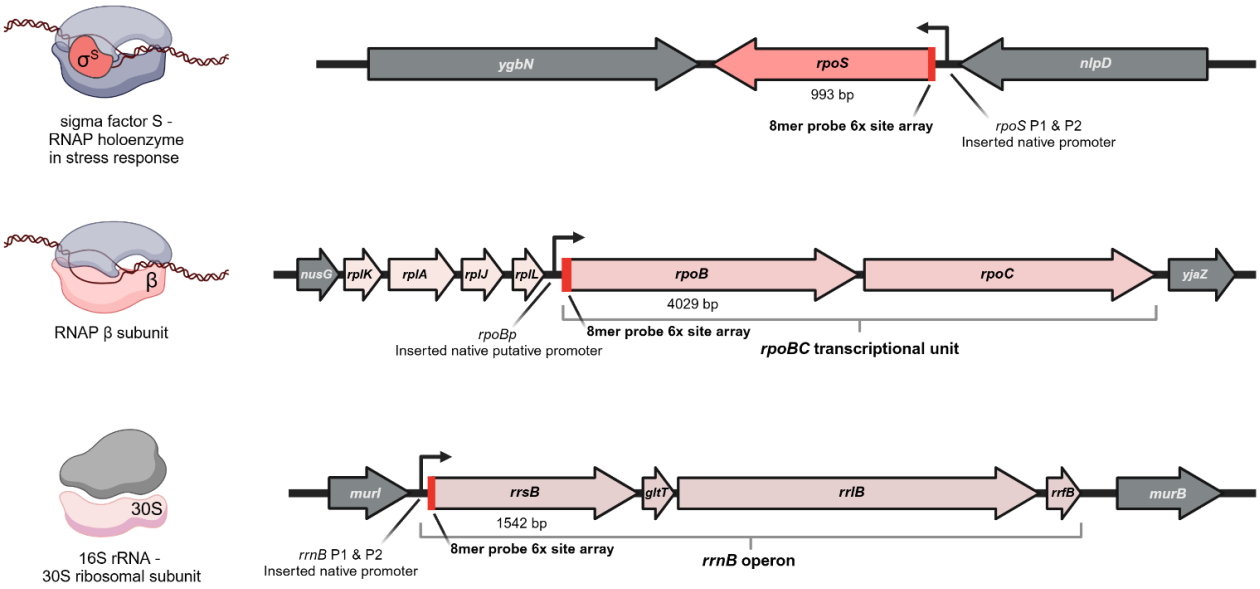


Figure 4.1 – Schematic presenting three candidate genes of interest for *in vivo* tagging of RNAs transcribed in real time using fluorogenic ssDNA probe hybridisation to an engineered 5' tag array.
Candidate constructions designed for tagging:
i. The stress sigma factor, sigma S – *rpoS*
ii. The β subunit of RNAP – *rpoB*
iii. The 16S rRNA forming the 30S ribosome subunit – *rrsB*
The genomic DNA schematics present the candidate gene (shades of red), with the bright red bar denoting the engineered tag array site at the 5' of the gene, between the transcription start site and the RBS. Insertion of the tag array occurs alongside insertion of the native promoter sequence of the candidate gene to ensure no sequence disruption in cloning. Corresponding transcriptional unit/operonic genes highlighted, and surrounding genes shown in grey.
Not to scale.

With the goal of observing the physiological transcription and downstream processing of native, essential gene RNAs, a system with minimal perturbations to both the target RNA and cellular machinery involved with tight fluorescent control for single mRNA resolution is key. As such, the transient hybridisation of short fluorogenic probes to a small 5' UTR site in the target RNA presents a promising avenue for the single molecule appraisal of target, and here essential, mRNAs at low copy numbers in the live *E. coli* cell.

Overview

In this Chapter, the MG1655 K-12 *E. coli* strain comprising the 6x 8mer tag array inserted in the 5' UTR of the *rpoB* gene (*rpoB*-tag strain) was designed and constructed. Initial characterisation of the *rpoB*-tag strain with the different forms of the 8mer-2x647N probes was performed for comparison to non-native tag system observations. The main comparative study between the tag strains and wildtype control cells was then conducted with the more robust 8mer-643/q1 probes, from ensemble intensity and photostability characteristics to molecular

mobility distributions. Finally, representative single cell observations of single molecule behaviours were made.

The equivalent pipeline for observations with the original 8mer-2x647N probes can be seen in Appendix 2. Used as a partial baseline for observations, there are some references to comparing 8mer-2x647N results with those of the 8mer-643/q1 probes.

4.2 Molecular design – tagging the native *rpoB* promoter

rpoB as a native target – the rationale

RNAP is the core enzyme conducting transcription, highly conserved across both prokaryotes and eukaryotes. There are considered to be ~1000-5000 RNAPs per cell (Bakshi et al., 2012; Endesfelder et al., 2013; Grigorova et al., 2006), with more recent estimates suggesting that in *E. coli* with a doubling time of ~30 min, there are ~1000 free RNAP complexes, ~2400 DNA-searching RNAPs (non-specific associations)(Patrick et al., 2015), and ~20% actively elongating nascent mRNAs (Grigorova et al., 2006). A core determinant in all hierarchies of intracellular regulation – specific gene expression, coordinated translation rates, environmental adaptation – probing the transcription life cycle of genes encoding for RNAP subunits is of interest in gauging regulation of transcription within the live cell.

A polycistronic mRNA

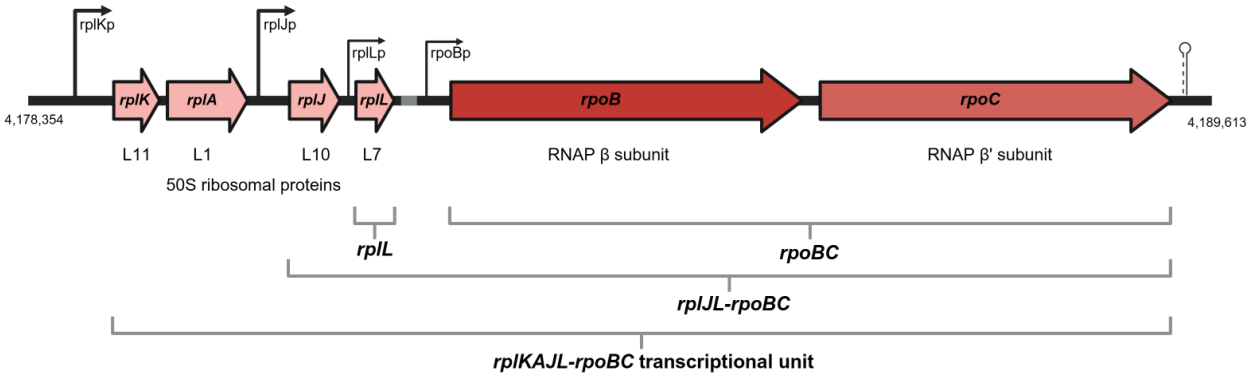


Figure 4.2 – Schematic denoting the structure of the *rplKAJL-rpoBC* operon.
Candidate *rpoB* is shown in dark red, and closely linked and co-expressed *rpoC* shown in a similar shade; 50S ribosomal protein genes shown in light pink. Dominant promoters rplKp and rplJp denoted by thick black arrows, minor rplLp and putative rpoBp shown in grey. Operonic terminator depicted as a stem loop (right), and attenuator site 3' to *rplL* shown in light grey. Possible transcriptional units labelled accordingly; genomic location (bp) as the transcription start site of rplKp to the end of the operonic terminator.
Not to scale; adapted from EcoCyc (Karp et al., 2023).

As highlighted by Figure 4.2, the *rpoB* gene is one of 6 within the essential *rplKAJL-rpoBC* operon, and *rpoB* is actually part of a linked and co-expressed pair, *rpoBC* – reconsolidating the highly systematic and efficient programming of gene expression in *E. coli*. As the major β and β' subunits of RNAP, it is imperative that there is stoichiometric balance in expression for regulated formation of RNAPs as and when, hence *rpoBC* tight, coordinated regulation and expression.

It is well established that transcription and translation are two highly intercalating processes, so far as they are co-regulatory; this is increasingly evidenced across the macro and micro scales, from circulating RNAP/ribosomal populations to the regulation of respective subunit expression (Bakshi et al., 2012; Le Nguyen et al., 2022). In the case of *rpoBC*, this crossover has been observed for some time – classical studies from the 1970-90s identified the 'rif region' of the *E. coli* genome (*rplK*-, comprising *rpl* and *rpo* gene clusters for small ribosomal subunit proteins and the RNAP clamp subunits respectively (Zengel & Lindahl, 1994). Through *in vitro* hybridization assays, lambda phage constructions and *in vivo* gene fusions, these studies identified two major promoters and subsequent transcription units (Barry et al., 1979; Downing & Dennis, 1987; Linn & Scaife, 1978; Ma et al., 1981; Newman et al., 1982; Petersen, 1989; Ralling & Linn, 1984; Steward & Linn, 1991; M. Yamamoto & Nomura, 1979). The first major promoter, designated rplKp here, was found to be the dominant promoter, and without specific termination sequences, was found to enhance the consecutive *rplJ* unit expression to an almost doubled transcription rate versus the upstream unit (Barry et al., 1979; Downing & Dennis, 1987; Ma et al., 1981; Newman et al., 1982; Ralling & Linn, 1984; Steward & Linn, 1991).

Now, these dominant promoters concern the ribosomal protein transcription units; but as the major promoters, were also found to enable continuous transcription until end-termination at the C-terminus of *rpoC*. Notably, activation of both promoters was only seen to increase *rpoB* transcription by ~20%, primarily due to the attenuator locus directly upstream of *rpoB*C (Barry et al., 1979; Post et al., 1979; Steward & Linn, 1991) – which was shown to attenuate ~70-80% of transcript read-out before RNAP progression onto *rpoB*C (Downing & Dennis, 1987; Ralling & Linn, 1984). Corroborating the dominance of these promoter sites, more recent studies have found tight expression control through these promoters via the alarmone, ppGpp pathway under amino acid starvation (Lemke et al., 2011). More nuanced control in high growth rates and adaptive states seems possible via intragenic *rpoB*C mutations modifying RNAP slippage capacity in transcription elongation (Y. N. Zhou et al., 2013) and long-range co-translational coupling (Baughman & Nomura, 1983, 1984; Petersen, 1989).

Alongside myriad regulatory elements, the presence of additional promoter or promoter-like sequences in polycistronic regions of the bacterial chromosome is highly likely for promoting nuanced expression. To this end, recent studies have identified internal promoters initiating antisense transcription or variable transcription (Tutukina et al., 2007). In the case of the *rplKAJL -rpoB*C operon, Barry et al. putatively found a weak promoter in the *rplJ-rplL* intergenic region (Barry et al., 1979). Considering this possibility for the transcription of solely *rpoB*C within the polycistronic structure, Ralling & Linn reported no evidence for such a promoter. However, observations were made that the addition of rifampicin, contrary to blocking transcription as is the case with most transcribable complexes, actually appeared to induce *rpoB*C higher transcription rates (Newman et al., 1982) by suppressing the *rpl* attenuator effect.

These reports together suggested the presence of a weak, putative promoter 5' of *rpoB*, subsequently enabling multiple *in vivo* mRNA species (Fukuda & Nagasawa-Fujimori, 1983). Notably, G-quadruplexes have increasingly been implicated as transcriptional mediators, enriched in transcription-translation regulatory sequences (Holder & Hartig, 2014; X. Shao et al., 2020); and the 5' UTR of *rpoB* (after the transcription start site, TSS) has been identified as a possible G-quadruplex structure. Hence, in addition to the putative weak promoter sequence, there are even more potential regulatory mechanisms involved with the delicate expression profiles of *rpoB*C.

Design and construction of the tagged-rpoBp (*rpoB*-tag) *MG1655 E. coli strain*

The weak, poorly elucidated rpoBp was chosen as the candidate "promoter" for tagging. The primary rationale behind rpoBp was 5' proximity to the *rpoBC* transcriptional unit, and redundancy in expression. A schematic of the engineered chromosomal site, alongside the molecular cloning fragment sequence is given in Supplement 4.1.

To generate a lambda-phage targetable insert at the rpoBp locus of the *E. coli* MG1655 genome, the rpoBp sequence (taken from EcoCyc) was inserted into the modular plasmid system developed (Supplement 4.1). The rpoBp sequence was verified against the K-12 MG1655 whole genome sequence (*Escherichia coli Str. K-12 Substr. MG1655, Complete Genome*, 2025). As follows, the target promoter was followed by the six-site array of 8mer-complementary TCCACCGT, linked by a 20nt linker. A slight modification from the original array design, the randomized spacer sequences were increased to 15 nt from 6 nt, by way of compensating for possible steric occlusion by consecutive site occupancy by probes (Griffith, 1978). Lambda red homology arms were designed to overlap with the kanamycin selectivity

cassette (5') and the 8mer tag array to enable P1 transduction of recipient MG1655 cells with the tag cassette ~150 nts upstream of the *rpoB* start codon, leaving the ribosomal binding site (RBS) intact and unperturbed. FLP/FRT reductase was attempted to remove the antibiotic selectivity cassette, but complete removal was not verified.

Following the premise of engineered *in vivo* smFISH, the initiation of transcription would lead to a nascent mRNA, to which the fluorogenic 8mer ssDNAs could hybridise and de-quench, leading to fluorescent foci. Engineering the tag immediately 5' of the *rpoB*C transcriptional unit would result in fluorescent foci, dependent solely on *rpoBC* expression and unperturbed by expression of the upstream polycistron. The redundancy of the system is that, even if the weak promoter is in fact non-functional, *rpoBC* expression would still be enacted by RNAP initiated at the *rplK* and/or *rplJ* promoters, preserving overall essential expression. Any lethality in molecular construction served as a partial control for site-directed engineering: off-targets in this essential operon (major promoter disruption, deleterious gene mutations) would not enable cell survival.

4.3 Comparing 8mer-2x647N hybridisation to the *rpoB* transcript

To establish comparable probe observations as in Chapter 3.5.2, the different 8mer probe samples were trialled in M9Glu competent *rpoB*-tag cells. M9Glu competent *rpoB*-tag cells were electroporated with 50 nM original, commercial, 5 nM in-house-purified 8mer-2x647N respectively, and imaged respectively in HILO at ~10 mW 638 nm, 50 ms exposure.

Figure 4.3A presents maximum intensity projections of representative FOVs for *rpoB*-tag cells electroporated with the various 8mer-2x647N probes, alongside representative FOVs of the same cell types without the probe (electroporated with upH_2O). The *rpoB*-tag cells with commercial 8mer-2x647N appeared to show highest heterogeneity, and overall most uniform and highest spread of signal in the cells with the original 8mer-2x647N probes. Whilst cells with the in-house-purified 8mer-2x647N do appear loaded and individual foci are visible, there appears to be overall fewer fluorescent foci, indicating as priorly discussed that in all likelihood, much of the collected probe fraction contained rapidly photo-bleaching free dye or partially-labelled ssDNAs, and few intact doubly-labelled ssDNAs – thereby leading to fewer specific signals overall.

With an anticipated transcript of 4000-8329 nt depending on *rpoB* or *rpoBC* full transcription and endogenous gene expression, fewer copies per cell in real time would be expressed from the physiological locus. Subsequently, it was hypothesised that comparative to the inducible tag strains, fewer hybridisation events would occur, and hence fewer signals seen in a given

cell at a given time. Irrespective of the probe used, this observation was reproducibly confirmed; a side-by-side FOV comparison for all strains electroporated with 50 nM original 8mer-2x647N is shown in Supplement 4.2.

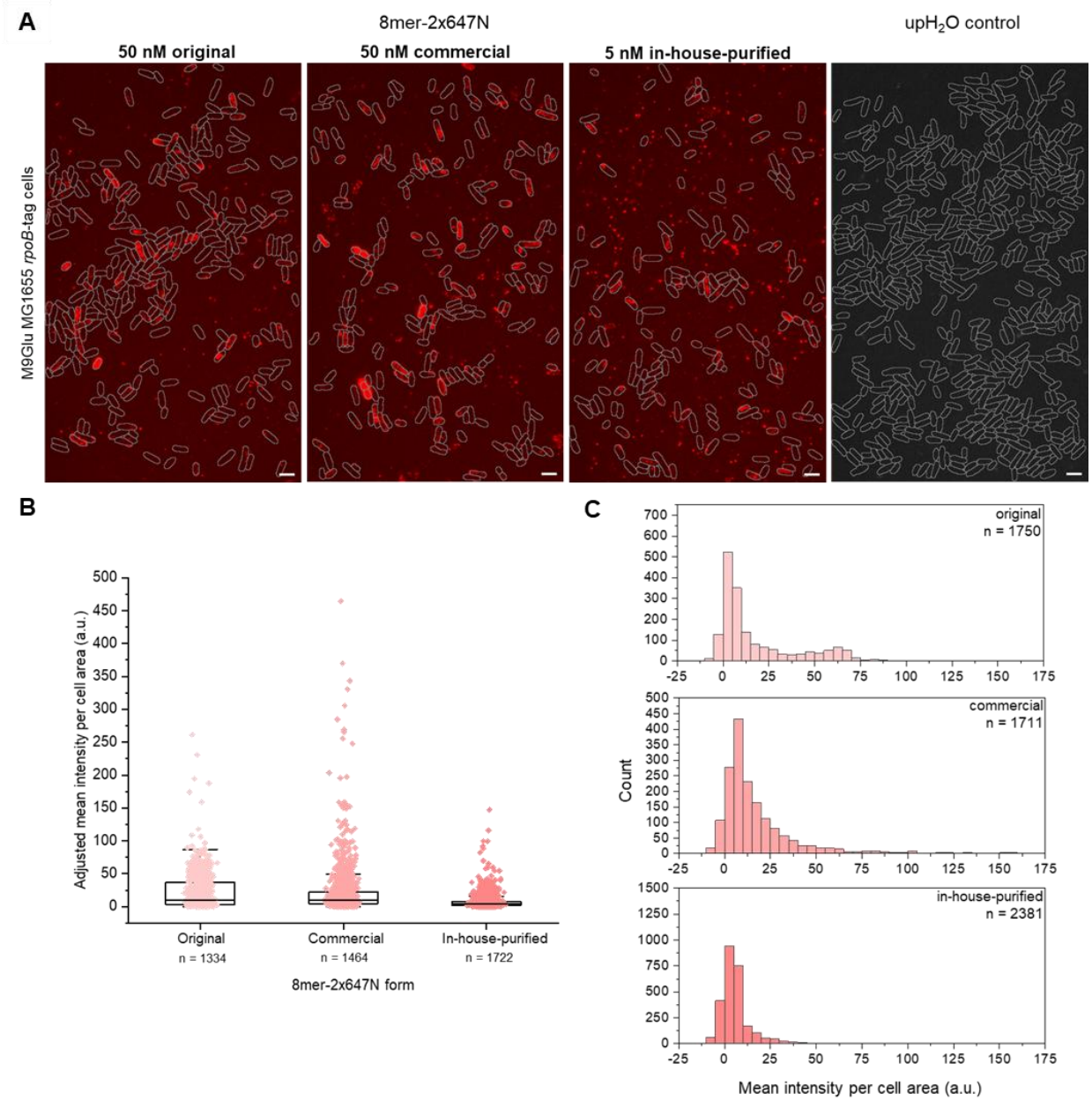


Figure 4.3 – Testing the different forms of 8mer-2x647N in *rpoB*-tag *E. coli* M9Glu competent cells.
(A) Maximum intensity projections of frames 100-2000 (100th frame) of probe-electroporated cells. (B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for *rpoB*-tag cells electroporated with 50 nM original, commercial and 5 nM in-house-purified 8mer-2x647N. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH$_2$O-electroporated (empty) cells. Medians: 21.5 units original, 13.4 units commercial, 6.9 units in-house-purified.
(C) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 5.
All mean intensities per cell area shown are background corrected and n gives the number of cells per sample (N = 3, 2,3 experiments respectively).
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

Ensemble intensity data

Analysing the 100th frame[5] background-corrected mean intensities of cells classified as loaded with the three 8mer-2x647N, the original/commercial 8mer-2x647N gave the highest median mean intensities per cell area (original: 9.5 units, commercial: 10.4 units), with negligible difference between the two. Both probes showed ~3 times higher median mean intensity per cell area than cells electroporated with in-house-purified 8mer-2x647N (3.8 units)(Figure 4.3B).

Plotting the intensity distributions across the whole cell populations for the different 8mer-2x647N probes, the in-house-purified probe gave the highest proportion of sub-zero valued, "empty", cells (~145 cells, versus 140 for original and 125 for commercial), and the least heterogenously loaded cell population (Figure 4.3C). Notably, the commercial 8mer-2x647N showed the widest cell loading heterogeneity (by a magnitude of >2), despite having the smallest overall number of cells, with ~71% cells with intensities between 0-25 units – likely corresponding to cells with 1-5 distinguishable foci. Whilst ~ 66% cells containing the original 8mer-2x647N probe also had between 0-25 unit intensities, there was less of a tail of highly-loaded, "super-bright" cells, suggesting that cells were more uniformly loaded with fewer, probably individually resolvable, fluorescent foci – optimal for single molecule analyses. By way of loading classification, ~76% cells were loaded with the original 8mer-2x647N , 86% cells with the commercial 8mer-2x647N, and ~72% cells with the in-house-purified 8mer-2x647N.

As approximated for the induced p6 and gT7 tag strains electroporated with 50 nM original 8mer-2x647N, a proxy for thresholding out spurious signal can be achieved by performing a third correction, subtracting the average mean intensity of loaded M9Glu competent wildtype cells per probe. This increased the median mean intensity per cell area for all probes in electroporated *rpoB*-tag cells (original: 29.2 units, commercial: 12.0 units, in-house-purified: 8.0 units), corresponding to 36% original, 52% commercial and 4% in-house-purified, loaded cells respectively (Supplement 4.3).

[5] The brightest frame results for the median mean intensity per cell area loaded cell boxplots and whole population distributions are in Supplement 4.3, alongside the equivalent plots for induced p6 and gT7 cell populations from 3.5, for direct visual comparison.

Replotting the histograms to show the distribution of these corrected populations reveals that the original probe leads to more highly loaded cells, the commercial probe gives the most uniform distribution of probes across cells, and that barely any loaded cells with the in-house-purified probes likely contain real signal (Supplement 4.3). This suggests that the commercial probe may actually be the most robust 8mer-2x647N probe for producing a more ubiquitously loaded cell population, with a spread of loaded cells with variable foci. Indicating the ability to form a substantial intracellular probe pool, this would be ideal in the case of longer, less frequently transcribed mRNAs: extending the photon budget, thus the possible acquisition time and the probability of visualising majority proportions of a single molecule transcription life cycle.

4.4 Applying the 8mer-ATTO643/BHQ1 probe for engineered *in vivo* smFISH in tag strains

As established in Chapter 3, the 8mer-643/q1 appears as the most suitable probe for single molecule event capture. The 8mer-643/q1 had the least fluorescent signal of all trialled probes at 42% cells classified as loaded from a population of over 5000 cells, and with most signals giving a low mean intensity (between 0-5 counts). Combining this lower off-target signal and reduced membrane-sticking propensity of the ATTO643, the overall fewer occurring and tighter intensity signals of the 8mer-643/q1 make this probe the most robust option yet to investigate specific mRNA-probe complexation and the exploration of target mRNA behaviours *in vivo* in engineered tag strains.

4.4.1 Populational cell intensity comparisons

Visually comparing all three tag strains – induced p6 and gT7, and *rpoB*-tag – with the M9Glu wildtype competent cells (shown in 3.6.2) electroporated with 50 nM 8mer-643/q1 as standardized showed that all tag strains had significantly higher signal than in the wildtype cells, both in individual cells and heterogeneously spread across the populations. From the maximum intensity projections of representative FOVs, wildtype cells have ~4% visually-loaded cells, versus ~18% in p6 cells, 25% in gT7 cells, and ~9% in *rpoB*-tag cells (Figure 4.4A).

Although not generating the apparent overexposure as observed with the 8mer-2x647N probes, extremely highly loaded cells were still present across all tag strains – around 18% of the loaded population in p6 cells, 9% in gT7 cells and 4% in *rpoB*-tag cells. It is unclear what is the cause of such high internalisation. Although cell death or near-fatal membrane porosity is one possible explanation, none of the super bright cells in these FOVs appear membrane damaged, and at the molecular level, show foci with representative behaviours (sporadic bound states, rapid diffusion). Since probe internalisation remains heterogeneous across a

cell population, it is more likely that high probe internalisation is caused by a combination of single cell phenotypic variation in membrane composition, proximity to high probe concentration before and during electroporation, and single cell variation in the rate of membrane re-stabilisation post-electroporation. Of note, is the apparent "racetrack" appearance of many cells across all strains, once again suggestive of trapping of the labelled ssDNAs in the intermembranal or periplasmic space (Salcedo-Sora et al., 2021).

As anticipated, the median mean intensities per cell area of classified loaded cells from the brightest frame follows as above: median 6.0 unit intensity for p6 cells, 9.6 units for gT7 cells, 5.7 units for *rpoB*-tag cells and 3.1 units for the wildtype (Figure 4.4B). Classified loaded populations (as all cells with mean intensities above zero, post-subtracting the brightest averaged mean intensity of empty cells) were 67% for p6 cells, 77% for gT7 cells, 71% for *rpoB*-tag cells and 42% for the wildtype cells.

Saliently, whilst the p6 and gT7 cells showed whole cell distributions mirroring the observations in the representative FOVs and mean intensity spreads – a smaller population of highly bright cells, 77% p6 and 69% gT7 cells with intensities between 0-25 units respectively – the *rpoB*-tag distribution effectively entirely overlaps with the wildtype population distribution (Figure 4.4C). From an ensemble perspective, there is no difference between the *rpoB*-tag and wildtype populations, demonstrating a necessary reliance on single cell and molecule observations for further classification of signal specificity.

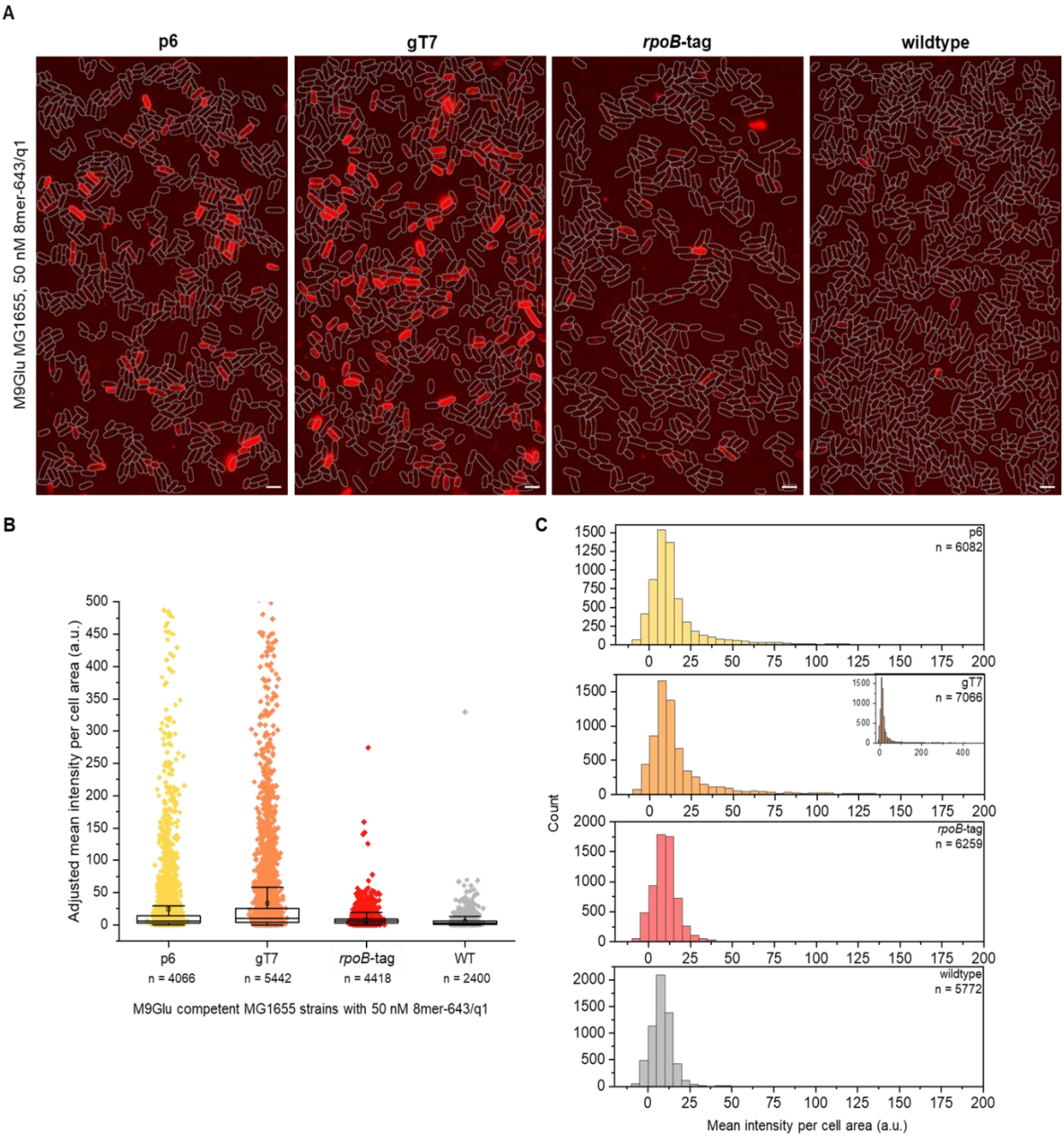


Figure 4.4 – Comparing 8mer-643/q1 observations in tag-system *E. coli* M9Glu competent cells.
(A) Maximum intensity projections of frames 1-2000 (brightest frame) for tag strain and wildtype cells electroporated with 50 nM 8mer-643/q1.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers). Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells. Medians: 6.0 units p6, 9.6 units gT7, 5.7 units *rpoB*-tag, 3.1 units wildtype.
(C) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 5. gT7 shows an inlay depicting whole spread, giving counts up to 400.
All mean intensities per cell area shown are background-corrected, and n gives the number of cells per sample (N = 2 experiments).
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

Applying a third correction, subtracting the brightest averaged mean intensity of the wildtype cells, as a proxy for specific binding events: the median mean intensity per cell area for p6 cells was 30.3 units, 25.5 units for gT7 and 6.4 units for *rpoB*-tag cells (Supplement 4.5). These values equated to loaded populations of 19% p6, 29% gT7 and 5% *rpoB*-tag cells respectively from the whole cell populations. The corresponding wildtype intensity-corrected p6 and gT7 populations align with the population histogram distributions above (Figure 4.4C). Notably, wildtype intensity correction of the *rpoB*-tag cells does seem to reveal a sub-population with possibly specific signal, indicating that those cells with possible signal are likely to appear brighter, and almost immediately show distinguishable foci – optimal for further single molecule identification and analyses, essential to appropriate molecular behavioural classification.

4.4.2 Single-cell photo-bleaching – 8mer-643/q1 photostability and brightness across tag strains

To determine the photostability and approximate brightnesses of single probes from 50 nM 8mer-643/q1 electroporated into tag strain cells, single cell photo-bleaching studies were conducted.

Temporal projections of representative FOVs give approximations of the duration of signals observed in loaded cells (Figure 4.5A). From >175 single loaded cell[6] time traces across the three tag strains, averaged time traces (black) overlaid with two exemplar single cell traces were plotted (red)(Figure 4.5B). Many loaded cells appear to comprise signals across the whole acquisition, some up until the end of the 100 second duration, and the average time traces suggest that initial high fluorescence bleaches down quickly (within ~5 seconds). Beyond ~5 seconds, it appears that highly resolvable step-wise events can be observed, indicating the efficient quenching and high photostability of the fluorogenic 8mer-643/q1 when taken alongside the temporal projections.

[6] Where loaded cell is defined by visual categorisation from maximum intensity projections; >500 cells per strain visually categorised in ImageJ. Mean intensity per frame obtained in ImageJ via overlaying BacSeg masks for region determination. Average lifetime trace generated by taking the mean of each per frame intensity for all cells, fitting a single or double exponential to determine the y0 for baseline autofluorescence, and subtracting the y0 baseline from each per frame value for background correction.

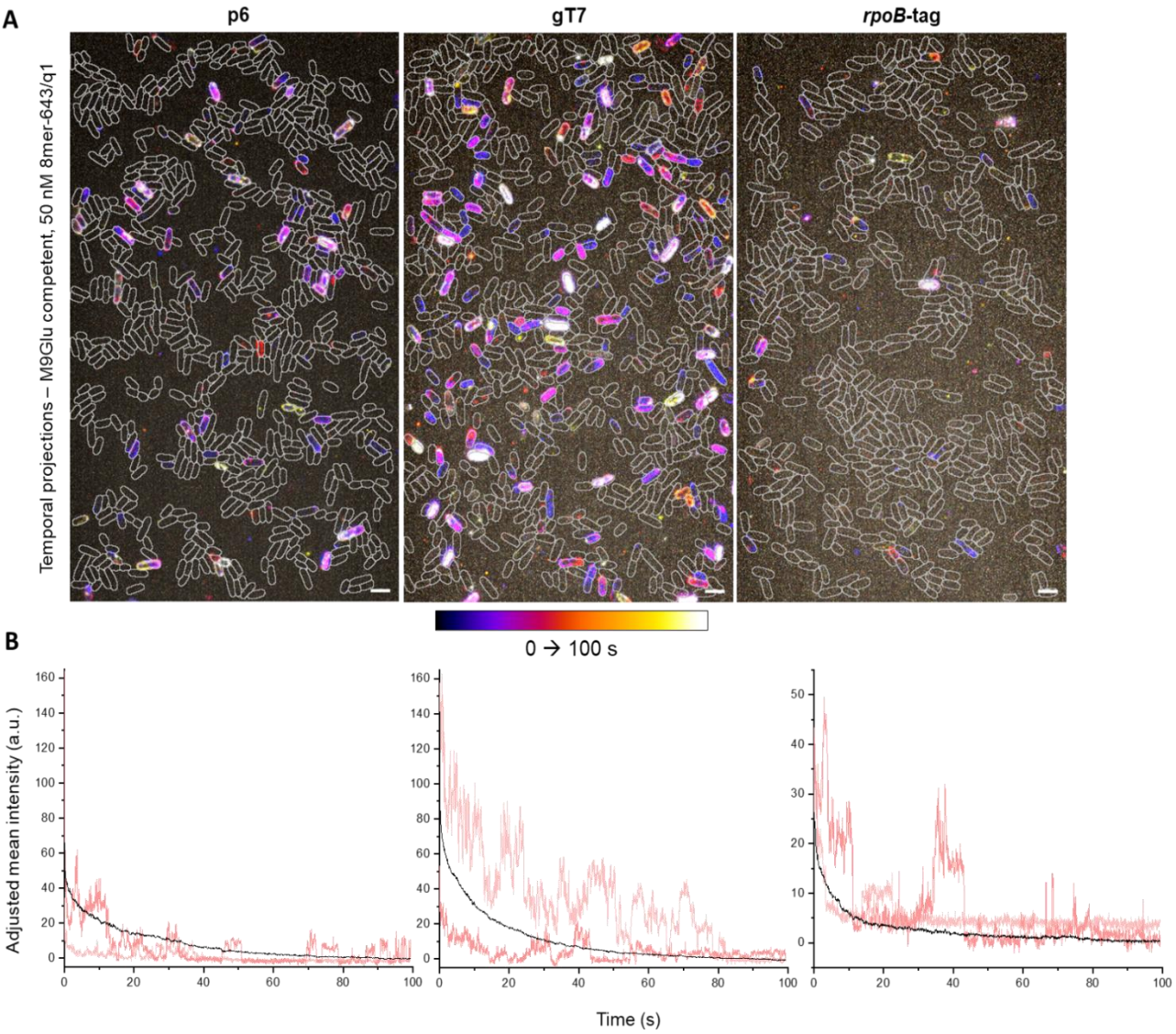


Figure 4.5 – Visualising the photostability of tag strain MG1655 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-643/q1.
(A) Temporal projections of 1-2000 frames of the same representative FOVs of probe-electroporated tag strain MG1655 cells in 4.4.1. The colour bar denotes time, where darkest is time 0 (50 ms exposure, 50 s acquisition).
(B) Corresponding averaged background-corrected mean intensity per frame time traces for the representative FOVs (black), overlaid with two exemplar bleaching time traces for individual visually-loaded cells (red shades).
Scale bars – 3 µm, overlaid cell segmentation outlines.

To evaluate the brightnesses of the 8mer-643/q1 across the representative tag strains FOVs, visually-identified loaded single cell time traces were fitted using step height analysis via *AutoStepFinder* (Loeff et al., 2021). As step changes in the photo-bleaching intensity traces are changes in the measured mean intensity – due to fluorescing foci bleaching and disappearing, or hybridisation events leading to fluorescence – the step heights can be correlated to individual fluorescent foci, and hence give an approximation of the intensity of a single fluor, the unitary intensity (Figure 4.6A).

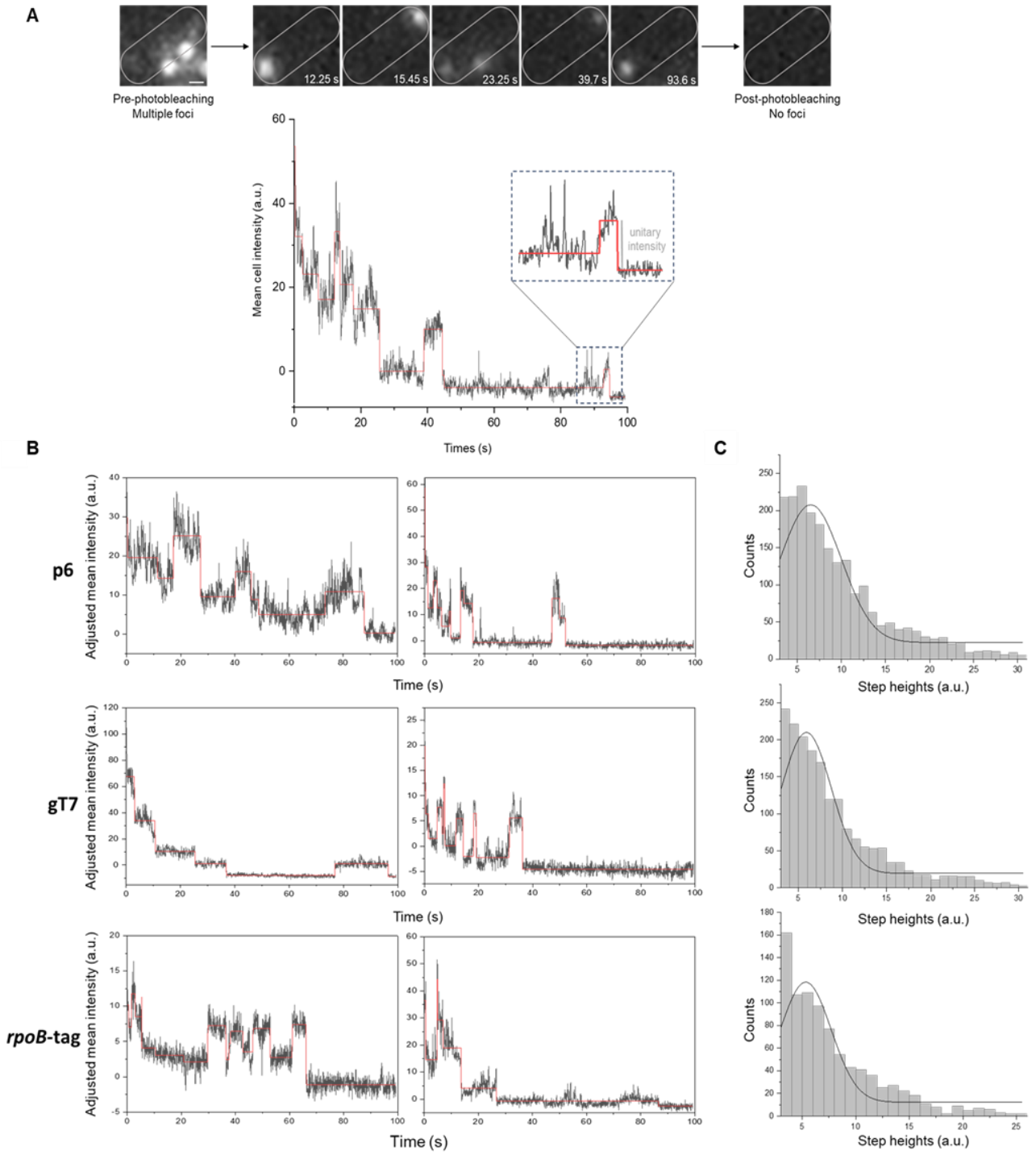


Figure 4.6 – Characterising the unitary intensity of the original 8mer-643/q1 probes in M9Glu competent tag strain cells.
(A) Exemplar single cell trace of a live gT7 cell with distinguishable bright foci from ~10 seconds onwards, from 2 clear very bright foci and ~4 diffuse foci at the start of acquisition; associated step analysis of bleaching over time presented below for central montage.
(B) For loaded p6, gT7 and *rpoB*-tag cells electroporated with 50 nM 8mer-643/q1, photo-bleaching step height analysis. Exemplar single cell photo-bleaching traces (black) fit for step heights (red) using *AutoStepFinder*. Steps correspond to the consecutive bleaching of fluorescent molecules, which can be correlated via the video acquisition.
(C) Corresponding histograms distributions of step height intensities from fitted single cell traces, where the peak denotes the unitary intensity of a single fluor, i.e. the average brightness of a probe-RNA hybridisation and de-quenching event.
p6: 6.5 ± 0.41; gT7: 5.9 ± 0.35; *rpoB*-tag: 5.3 ± 0.41units.

From >175 traces per strain, 115, 110 and 120 time traces were fitted for p6, gT7 and *rpoB*-tag cells respectively. Highly loaded traces were discarded due to overfitting of very high initial fluorescence and hence inability to resolve individual foci; traces with negligible steps were also discarded. Step height histograms were then manually thresholded for step heights measurements between 3-30 units, to account for overfitting (< 3) and underfitting by the algorithm, or loading beyond focal resolution (>30).

Representative traces with overlaid step height fittings (red) for each of the tag strains containing 1-5 foci from the start of acquisition are shown below (Figure 4.6B). These traces exemplify sequences of fluorescent foci observed, as well as time to bleach to baseline. The corresponding step height histograms for 8mer-643/q1 foci in all three tag strains are shown in Figure 4.6C, where a single gaussian peak fit denotes an approximated unitary intensity of 6.5 units for p6 cells, 5.9 units for gT7 cells and 5.3 units for *rpoB*-tag cells. Since the same concentration of probes is in use, same preparation and imaging conditions, it was thought that the unitary intensity value across strains would be effectively the same. Indeed, the differences are small, inclusive for the *rpoB*-tag cells. In this case, this suggests that the molecular engineering between the strains does not appear to significantly affect the phenotypic or genotypic states of the MG1655 cells extensively. Differences in approximated unitary intensities per probes hence more likely speak to probe photokinetics than any particular cell genetic and molecular states.

Applying the unitary intensity values obtained per strain, the distribution of molecules per cell may be approximated by dividing the brightest frame mean intensity per cell area values by the unitary intensity (Supplement 4.6). Taking counts of 1 focus upwards from the whole cell populations, ~31% of p6 cells, ~50% of gT7 and ~38% *rpoB*-tag cells were classified as having distinct countable foci. Corroborating with the representative temporal projections presented in 4.5A, gT7 cells appear to have the highest proportion of loaded cells, and p6 and gT7 cells both showing a wide spread of higher loading (~28% p6 and ~34% gT7 cells with 10 or more foci).

For *rpoB*-tag *cell*s, most cells appear to have distinguishable foci from the start, likely indicative of the lower copy number of tagged mRNAs present at any one given time. This once again denotes the need for analysis at the level of an individual event, as ensemble analyses overlap highly with wildtype equivalents and would erroneously suggest no specific signal.

Noted, the appearance and indeed measured intensity of a fluorescent focus can change depending on the depth of the complex in the cell, the imaging angle and the exposure time.

For this reason, all imaging conditions are consistent across comparisons. Although this does not affect ensemble approximations of the numbers of molecules per cell, it does impact potential read-outs of single molecule behavioural observations.

4.4.3 Comparing apparent diffusive populations across strains

As applied in 3.5.3, single molecule localisation and tracking give insight into the different diffusive states of identified fluorescent foci. To compare the distribution of apparent diffusive states (D*) of molecules thresholded for tracking, the apparent diffusive states of all MG1655 tag system and wildtype strains with 50 nM electroporated 8mer-643/q1 were compared (Figure 4.7).

Using the BacSeg cell segmentation masks, fluorescent foci were localised and linked across consecutive frames for track formation, from frames 1 to 2000 frames, using the StormTracker and Locoli platforms. For n single molecule trajectories in the p6 (n=35057), gT7 (n= 92440), *rpoB*-tag (n=6545) and wildtype (n=2917) cells, the MSD of individual probe-bound mRNAs was calculated for all trajectories across 4 or more consecutive frames, and the D* calculated per strain from trajectories truncated to 4 steps (Uphoff et al., 2013).

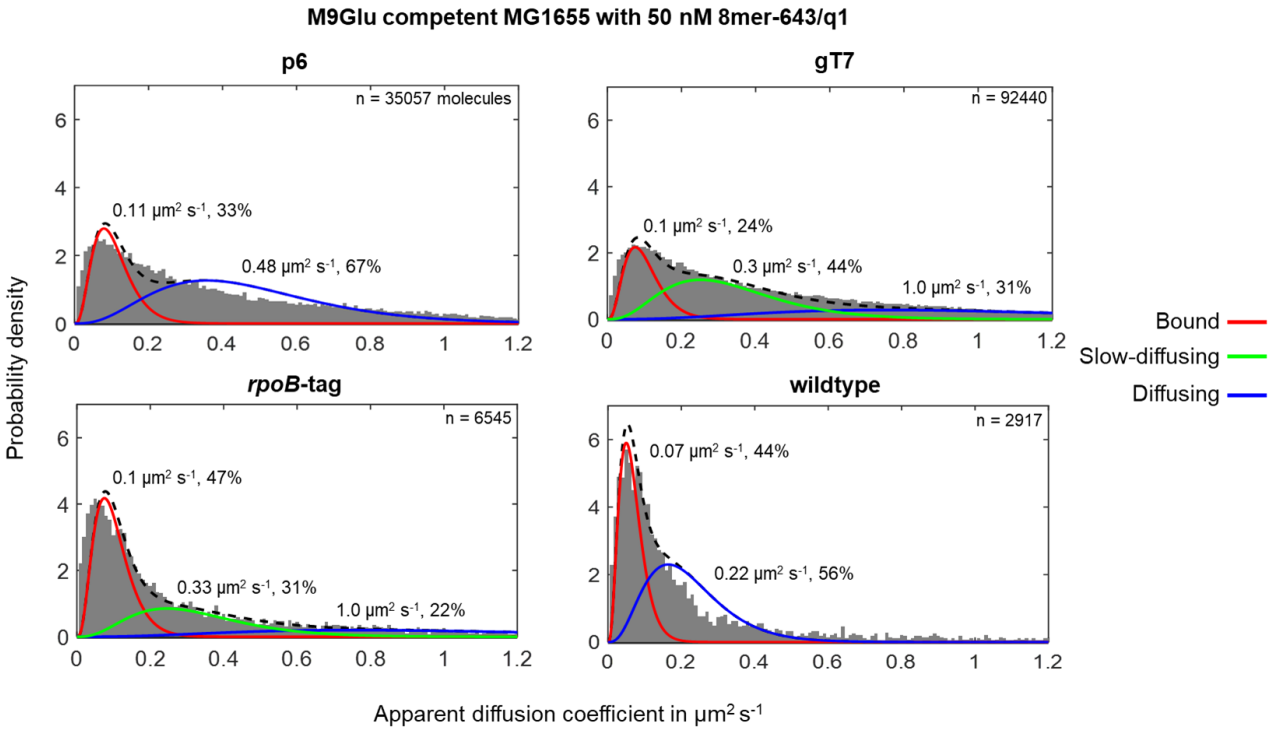


Figure 4.7 – Exploring the diffusive properties of tracked fluorescent molecules in MG1655 competent M9Glu tag strain and wildtype cells with 8mer-643/q1: apparent diffusion coefficients. Apparent diffusion coefficient distributions fit with two-free species curves (p6 and wildtype) or three gamma two-constrained fit (gT7, *rpoB*-tag) (red as bound, blue as diffusing, intermediary slow-diffusing as green).
n gives the numbers of tracked molecules per sample; N=2 (16 FOVs per sample).

Applying a two free species fit to the wildtype and p6 distributions, the wildtype population gave a 44% highly bound population at 0.07 $\mu m^2/s$ and a 56% "diffusive" population at 0.22 $\mu m^2/s$; the p6 distributions were 33% bound with 0.11 $\mu m^2/s$ and 67% diffusing with 0.48 $\mu m^2/s$ D*. A three-gamma, two-constrained species fit was applied to both chromosomal tag strains, where the bound and slow-diffusing fits were constrained (0.1 $\mu m^2/s$ red, 0.3 $\mu m^2/s$ green respectively). A three-species fit was selected owing to the longer lengths of the gT7 and *rpoB*-tag transcripts, and intact translation for *rpoB/rpoBC* – indicating the possibility of multiple conformational and hence mobility states (confined to a representation of three for analysis).

For gT7 cells, there was a 24% bound population with 0.1 $\mu m^2/s$ D*, a 44% slow-diffusing population at 0.34 $\mu m^2/s$ and a 31% fast-diffusing population at 1.0 $\mu m^2/s$. For *rpoB*-tag cells, there was a 47% bound molecule population at 0.1 $\mu m^2/s$, a 31% slow-diffusing population at 0.3 $\mu m^2/s$, and a 31% fast-diffusing population at 1.0 $\mu m^2/s$. Considering the constraint of the bound population at 0.1 $\mu m^2/s$ for *rpoB*-tag cells, the histogram plot suggests that a D* of 0.1 $\mu m^2/s$ may be an overestimate, bound molecules likely therefore even slower than 0.1 $\mu m^2/s$ – which could correspond to a long, RNAP or ribosome-engaged mRNA.

Both bound and diffusing populations of molecules in wildtype cells have smaller D* values than the bound populations for all tag strains. Since both wildtype mobility classifications are slower than those of the tag strain cells with 8mer-2x647N, the hypothesis that the hydrophobicity of ATTO647N dyes encourage membrane sticking and insertion leading to highly static foci is supported, as well as suggesting that the likely charge-neutral ATTO-643/BHQ1 fluor combination does not exhibit extensive off-target binding (*ATTO 643*, ATTO-TEC). The wildtype diffusing population is also relatively ill-fit with a free species assignment, suggesting that the free fit shows an upper bound for faster moving species. Most fluorescent foci in wildtype cells appear highly static, and may be visually identifiable as so, alongside some rapidly-diffusing, non-specific signals.

With the short length of the p6 transcript (128 nt), two species are a rational outcome when considering the two likely states occurring: rapid transcription, where the nascent transcript is bound by the RNAP albeit on the scale of ~3.5 seconds, and freely diffusing once released. Considering Bakshi et al. measured a D* for the diffusing 28 kDa Kaede protein at ~7.3 $\mu m^2/s$ (Bakshi et al., 2011), and accounting for the molecular weight of 3-4 ribonucleotides as ~ 1 kDa (*DNA and RNA Molecular Weights and Conversions, Thermo Fisher Scientific*), the resultant 128 nt transcript would have an estimated molecular weight of ~41 kDa. Even nearly double the weight, the short length, higher flexibility and linearity of the p6 transcript would make an

even faster diffusion rate more than acceptable; although it is unlikely that even a short mRNA would diffuse by itself, instead complexed by different cellular machinery. Indeed, there are consistent counts of much faster diffusive species, within reasonable fast diffusion and possibly whole cell periphery-covering speeds. The fit is not so clean, however, indicating that there are likely more than two species present even with very short transcripts, and that the diffusing population at 0.48 $\mu m^2/s$ is a lower bound estimate.

For gT7 cells with a ~2700 nt transcript, slower diffusion via stochastic Brownian diffusion or otherwise would be anticipated. Likewise, larger bound populations for longer transcription time were hypothesised, of the order of 67.5 seconds if taking the rate of transcription elongation as 40 nt/s (Vogel & Jensen, 1994). Slightly perplexingly, the 24% 0.1 $\mu m^2/s$ population is smaller than the equivalent bound population in p6 cells, although this observation is rectified by the much longer likely transcription rates of gT7, hence fewer overall transcription events than for plasmid-mediated p6 transcription. The intermediary population for gT7 seems most prominent, which could correspond to actively-transcribing complexes. Noted, there is no RBS in the gT7 construct, hence no translation would occur.

Applying the same transcription considerations to *rpoB*-tag cells: a 4000-8329 nt transcript would be anticipated, corresponding to molecular weights of 1.28-2.67 MDa, on the order of 100-200 seconds' transcription. These are huge molecular constructs, very unwieldy and in all likelihood – peripheralized mid-cell nucleoidal locus of *rpoBC* – would be almost immediately be engaged in translation (Mayer et al., 2021; Sanamrad et al., 2014; Sattler & Graumann, 2021), probably polysomally. The *rpoB*-tag strain, being the only natively-engaging system, is theoretically the only one of the three to exhibit tagged-mRNA translation – there are no RBS intact associated with the p6 and gT7 tag-strain constructs. Due to the essentiality of the *rpoBC* operon, and inevitable lethality should the engineered tag disrupt core processing, it is assumed that functionality is intact. The highest proportion of tracked molecules as 'bound', 47% at 0.1 $\mu m^2/s$, is therefore feasible.

Whilst all tag strains exhibit a bound population around 0.1 $\mu m^2/s$, likely synonymous with a nascently-tethered mRNA in early transcription elongation, the complete overlap with the wildtype population signifies that specific probe-mRNA complex behaviours must be identified visually, per molecule. Nevertheless, the >50% tracked molecules as intermediary or fast-diffusing species with a $D^* \geq 0.3\ \mu m^2/s$ across all tag strains indicate likely real signal, as such populations are not seen in wildtype cells. This is promising, considering recent

evidence that mRNAs may diffuse freely throughout the cell, nucleoid-inclusive (Bakshi et al., 2012; Sanamrad et al., 2014; Sattler & Graumann, 2021).

Taking the same estimation metric as Sattler & Graumann, assuming a fast-diffusing molecular D* of 0.4 $\mu m^2/s$ in a 3 µm-wide spherical cell ($t = r^2/6D$), a single mRNA may cover the whole cell in <1 second. This appears in agreement with Sattler & Graumann's mRNA estimates of an ~4 second whole-cell diffusion time, with similar supportive observations in eukaryotic cells (Katz et al., 2016; Yan et al., 2016); especially considering all use the MS2 systems, adding weight and diffusion dynamic complexity due to the arraying/bulkiness of the reporter system.

Finally, in all cases, it is unlikely that the fast diffusive species is being confounded by capturing of free dye signal. Sattler & Graumann in their MS2-mVenus system for mRNA tracking opted for 75 ms exposure based on prior evidence that an exposure time of <20 ms would be required to best track freely-diffusing mVenus (duly tested by the authors using 8 ms exposure). As organic dyes are 25-80x smaller than FPs like mVenus[7], and can be estimated to diffuse on the order of >10 $\mu m^2/s$, acquisitions with a frame rate of 50 ms will not resolve freely diffusing ATTO dyes, and even if still conjugated to its 8mer ssDNA, would likely blur beyond trackability. Of course, with the aforementioned stickiness of a dye like ATTO647N, static appearance from sticking would indeed be trackable; but most likely highly membrane proximal, thereby enabling thresholding to limit false positives.

The D* histograms indicate a significant difference in tracked molecule acquisitions between the wildtype and tag strain cells with 8mer-643/q1. These suggest that further analysis at the individual molecule level is required.

[7] mVenus is a 26.9 kDa fluorescent protein (Lambert, FPbase); comparatively, organic dyes are in the region of 25-80x smaller, and alongside their higher compaction, often planar structure, have an overall lower hydrodynamic radius. Much of the experimental diffusion coefficient data for free organic dyes has been obtained from aqueous measurements *in vitro* (P. Kapusta, PicoQuant GmbH; Xu & Yeung, 1997) – unrepresentative of cell states – but nevertheless, acknowledging rhodamine dye free diffusion in water at 25°C of ~ 330 $\mu m^2/s$ (P. Kapusta, PicoQuant GmbH), the 800-1200 Da ATTO647N (*Atto 647 N, Sigma-Aldrich*) could be estimated to diffuse when accounting for the macromolecular crowding of the cell milieu.

4.4.4 Exploring populations of tracked molecules: counts per cell, durations & spatial distributions

Following diffusive state observations at the ensemble level, exploring the populational distributions of the tracked molecules themselves (counts per cell, molecular durations, spatial distribution) was of interest.

To enable the quantitation of the spatiotemporal aspects of single molecules in tag strains with 50 nM 8mer-643/q1 probes, >2500 cells from 8 FOVs per strain were analysed via single molecule localisation and tracking. Cell video data was processed using *MolTrack*, a single molecule localisation and tracking platform, implementing BacSeg Cellpose for cell segmentation and Picasso for molecular localisation. Thresholding trajectories to an equivalent stringency as used in Locoli and Stormtracker processing – linking consecutive frames of 5 or more to form tracks – single molecule statistics were extracted. Per 8 FOVs analysed, 23190 tracks were found in p6 cells, 51041 tracks in gT7 cells, 5459 tracks in *rpoB*-tag cells, and 1939 tracks in wildtype cells[8]. Representative FOVs from *Moltrack* for each strain are shown in Supplement 4.7.

Counts per cell

Investigating the normalised counts of tracked molecules within cells from tag strains (versus the wildtype controls) showed that the majority of cells across all strains comprised just one fluorescent signal, corroborative of the signal spread seen across experiments so far (Figure 4.8). The wildtype cells had the lowest proportion (~5%) of cells loaded with 10 or more tracked molecules compared to the tag strains, for which p6 cells had ~40% cells, gT7 ~47% cells and *rpoB*-tag cells ~10% cells with over 10 molecules per cell. This suggests that the more highly loaded cells may actually represent specific hybridisation complexes, and may offer greater insight into different molecular behaviours; appreciating that some of these are potentially spuriously higher numbers, as mentioned above. Overall, fewer cells with tracked molecules were found in wildtype cells compared to cells with tagged mRNAs: 688 wildtype cells had tracked molecules, compared to 1105 p6, 1958 gT7 and 828 *rpoB*-tag cells.

[8] Notably, whilst the track numbers per sample give a distributional indication, they cannot be directly correlated to exact numbers of molecules due to the prospect of fluorophore blinking and subsequent possible misassignment of tracks to molecules (i.e. the same molecule with a blinking event could be classified as two molecules). This is accounted for via the memory parameter (1 frame) and the consecutive frame linking for trajectory formation, but is not entirely fail-safe.

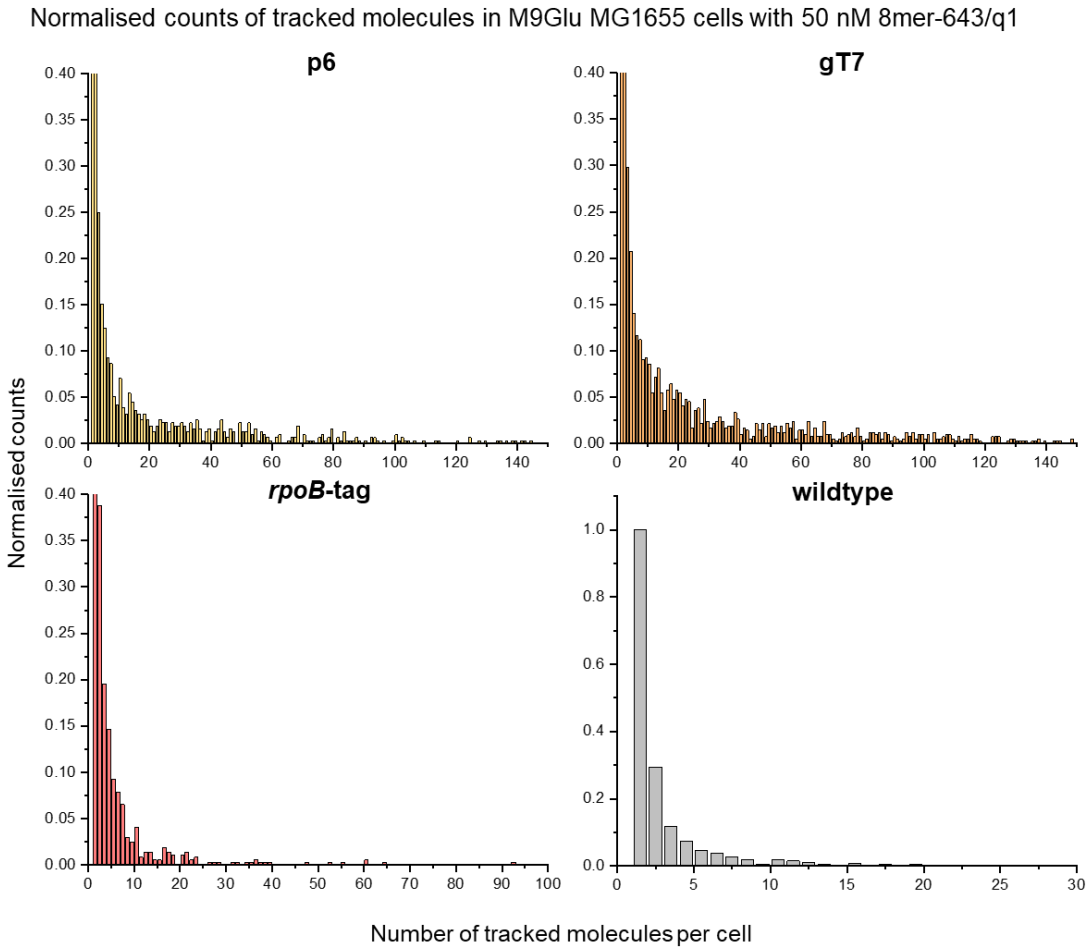


Figure 4.8 – Tracked molecule distributions comparing normalised counts per cell in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-643/q1.
Bin size = 1.

The disparity in the number of more highly loaded cells between the wildtype and inducible tag strains was even greater than with the 8mer-2x647N probes (Appendix 2.3), corroborating higher 8mer-643/q1 probe specificity. Across all tag strains, there were also proportionally more few-foci cells than in the wildtype, a further indication of probe specificity. However, the 10% *rpoB-tag cell*s with higher numbers of tracked molecules was much lower than the 34% measured in *rpoB*-tag cells with 8mer-2x647N. This could be reasoned by the higher efficacy quenching of the 8mer-643/q1 and expected low copy number of the *rpoB/rpoBC* mRNA *in vivo* requiring a higher electroporated concentration of probe than was used.

Tracked molecule durations

Considering the differing lengths of the tagged mRNAs, it was hypothesised that the gT7 and *rpoB*-tag cells would exhibit foci with much longer durations than the p6 cells, but also that fewer and more transient wildtype signals would be observed. Taking the rate of transcription elongation as 40 nt/s (Vogel & Jensen, 1994), estimated maximum durations would be ~3.2

seconds for p6 transcripts, ~67 seconds for gT7, and between 100-225 seconds for *rpoB/rpoBC*.

Figure 4.9 shows the normalised distributions of tracked particle durations, clipped at 16 seconds (Supplement 4.8 displays the full range per strain). All distributions are highly overlapping at the lower time distributions – all strains demonstrated 74-85% molecules have durations between 0.2-0.75 seconds – suggesting that most of these signals are highly transient, either rapidly-diffusing particles or misassigned and erroneously unconnected trajectories for the same molecule. Notably, 85% of wildtype cells had tracked molecules between 0.2-0.75 seconds, versus 29% in cells with the 8mer-2x647N (Appendix 2.3); although influenced by different n, this indicates the higher specificity and lower occurrence of hypothesised membrane sticking/implantation by the 8mer-643/q1.

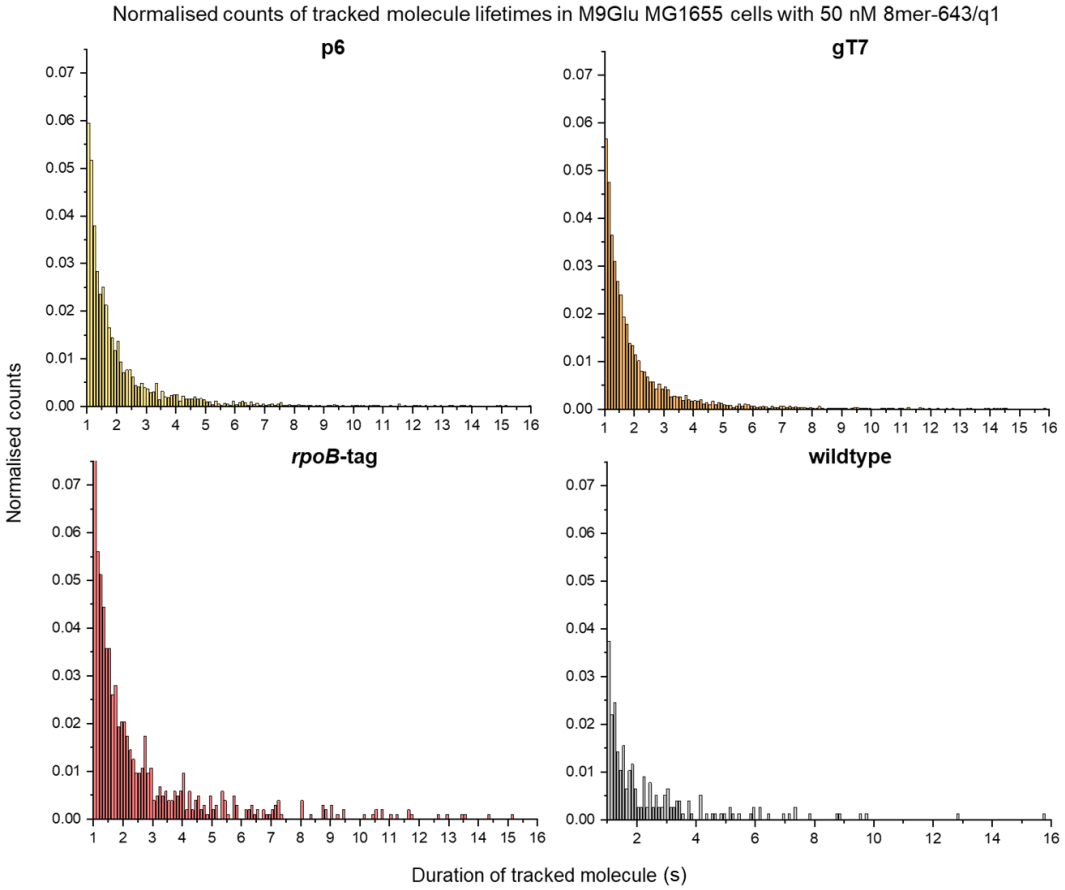


Figure 4.9 – Normalised duration distributions for tracked molecules in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-643/q1. Bin size = 0.1, distribution thresholded at 1 second and cropped to 0.07 counts to show spread of tracked molecule counts with longer durations.
Proportion of molecule lifetimes up to 1 second: 85.6% p6, 85.8% gT7, 80.7% *rpoB*-tag, 89.2% wildtype.

Taking 3 seconds onwards as a more insightful gauge of particle durations based on the approximate p6 transcription time, p6, gT7 and wildtype cells all had ~2.5% molecules with durations above 3 seconds, whereas *rpoB*-tag cells had an ~4% population. 11 *rpoB*-tag molecules and 26 gT7 molecules had durations between 20-72 seconds, and one gT7 molecule even lasted for 99 seconds. Considering that the gT7 mRNA has an estimated transcription time of ~85 seconds, these observations seem highly feasible and indicative of partial visualisations of the lengthy chromosomal tag transcription cycles, and possibly even the majority of a cycle for the one 99 second gT7 molecule. These results indicate that the chromosomal tag mRNAs indeed had the widest distributions of tracked molecule duration, and exhibited the longest durations.

Spatial distributions of tracked molecules

A final useful populational distribution is the locality of the cell in which molecules are most present. Considering there is some debate around the preferential regional localisation of types of mRNAs, and further debate around mRNA nucleoidal subdiffusion, an insight into membrane proximity is of interest (Fei & Sharma, 2018).

Across wildtype and *rpoB*-tag cell populations, tracked molecules appear to have very little consistent spatial proximity to the membrane (Figure 4.10A) – variable distances of 10- 600 nms away, the latter of which could be halfway across a cell (assuming a bacillus of ~1 μm width). The p6 and gT7 molecule populations appear more consistently left-skewed, closer to the membrane, which could be explained by nascently-transcribing complexes considering their respective transcription loci – typically polar plasmids and the peripheral gT7 chromosomal locus. The p6 and gT7 localisations with 8mer-643/q1 are also more membrane-proximal than indicated by the 8mer-2x647N results (Appendix 2.3).

Regardless, a wide distance distribution is still observed, aligning with the 8mer-2x647N observation and interpretation of rapidly-diffusing mRNAs traversing much of the cell, nucleoid subdiffusion inclusive (Montero Llopis et al., 2010; Sattler & Graumann, 2021). This variable distance of *rpoB* from the membrane could indicate a diffusing *rpoB* mRNA to polar and peripheral hotspots in the cell, ready for processive transcription-translation processes; given the overlap with the wildtype, however, this is purely theoretical.

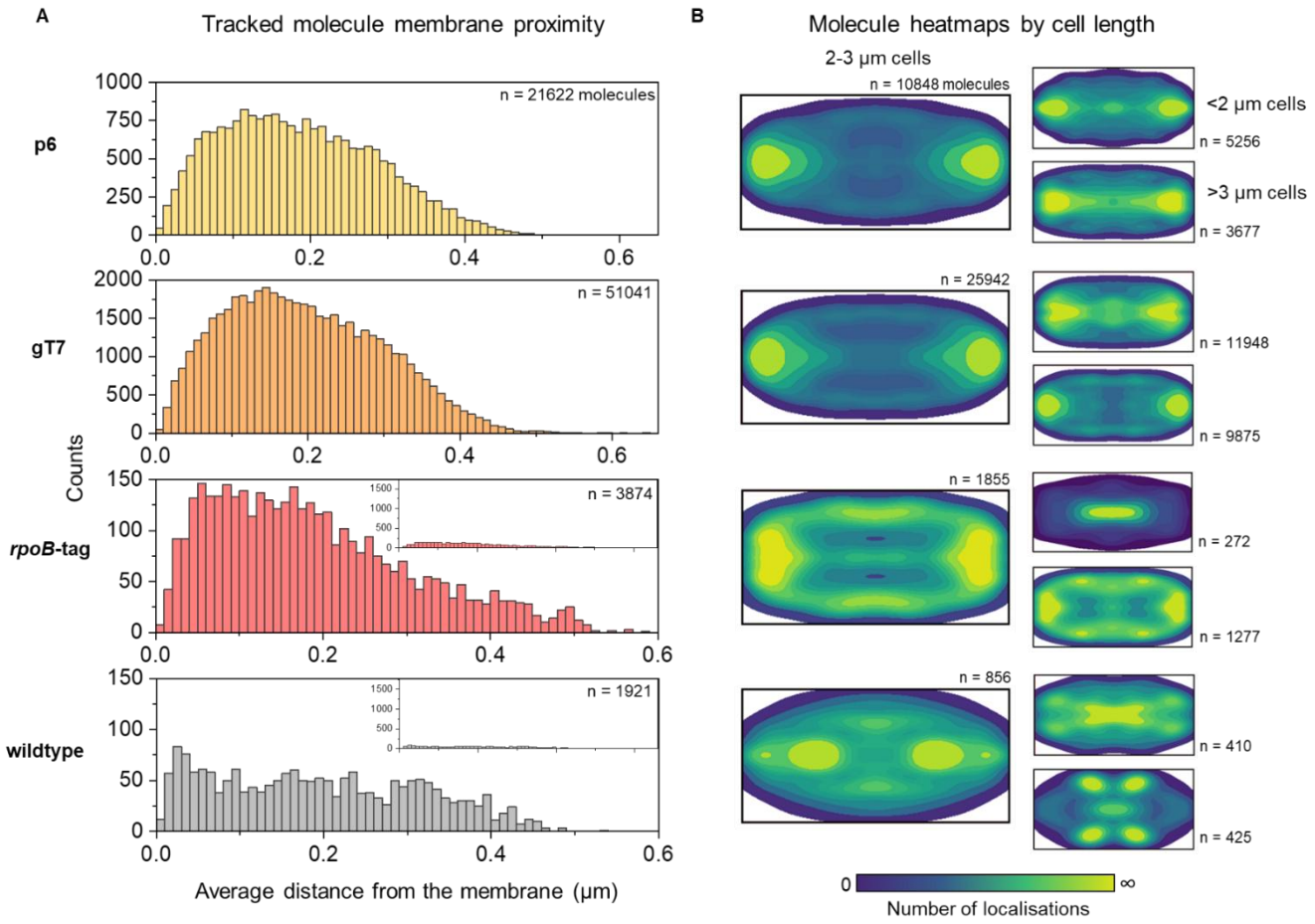


Figure 4.10 – Spatial distributions of tracked molecules in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-2x647N.
(A) Histogram distributions of average molecule distance from the membrane; bin size = 0.05. Insets on *rpoB*-tag and wildtype plots show spread at same y spread as p6 and gT7.
(B) Heatmaps showing the density of molecule localisations in different regions of the cell; the central plot denotes the average unit cell (length 2-3 μm), and the two corresponding righthand heatmaps denote cell lengths of <2 μm and >3 μm per strain sample.
n gives the number of molecules.

The average number of localisations was regionally projected across a unit cell, subcategorised by length into <2 μm, 2-3 μm and >3 μm cells to approximately represent the average E. coli cell (1 μm wide, 3 μm long) and capture shorter/perhaps dividing cells otherwise. Projecting the localisation density of molecules regionally in the cell (Figure 4.10B), very similar observations as with the 8mer-2x647N probes were made (Appendix 2.3). The majority of p6 and gT7 localised molecules were polar across all cell lengths, likely corresponding to transcription site proximity. The wildtype cells had a 2-30x lower number of localisations than the tag system strains for an average 2-3 μm-long cell and no consistent localisation pattern.

The *rpoB*-tag cell heatmaps also exhibit polar tendencies for localisations – these could indeed correspond to the midcell transcription locus of *rpoBC*, and the trafficking of the very

long mRNAs to the poles for efficient polysomal translation. However, the heatmaps additionally show more extensive midcell localisations, especially for the <2 μm (showing exclusively midcell localisations). Whilst the shorter cell heatmap signal could be spurious (low sampling, different cell morphology), the higher proportion of midcell localisations in 2-3 μm cells and longer suggests an indication of regional transcription sites.

Overall, there are clear spatial distributions for the p6 and gT7 cells: predominant polar and peripheral localisation. Whilst there are no consistent spatial distributions in wildtype cells – indicating spurious signal – *rpoB*-tag cells also show a proclivity for polar localisation, and some midcell peripheral hotspots. This consistent polar localisation across tag strains is suggestive mainly of mRNA trafficking, and particularly *rpoB/rpoBC* translation.

4.5 Behavioural characteristics of single RNA molecules bound by 8mer-643/q1

Further insight can be garnered from directly observing single molecules and their behaviours, aiding the specificity challenge at the ensemble level. To ascertain exemplary behaviours of fluorescent foci in live cells, representative loaded cells were taken, *Moltrack* tracking and the original video data compared.

Figures 4.11-20 depict representative observations made across M9Glu competent cells with 50 nM 8mer-643/q1, highlighted on the *Moltrack* representative FOVs in Supplement 4.7. In the tagged strains particularly, the behaviours observed and captured here in the exemplar cells represent ~10-20% behaviours seen across cells with signal. The process of presenting the exemplar cells is standardized across the strains, such that: exemplar cells are shown as a montage of stills taken from the real-time video, correlated to text descriptions of individual foci. For each cell, the *Moltrack* output of all tracked molecule trajectories in the cell is presented. Alongside this, *Moltrack* images of specific track segments (focus tracks) with corresponding plots for each focus tracked molecule's step size over time and proximity to the cell membrane, by way of representing diffusive and spatial behaviours respectively. Some exemplar cell specific focal intensity traces are also presented.

Wildtype MG1655

Starting with wildtype cell samples to classify off-target species, the predominant behaviour observed was a highly static focus, very close to the membrane. These were often very short-lived (<1.5 seconds) and most of these behaviours ceased past the first 5 seconds – suggesting transient off-target binding and rapid bleach-downs. Three sparse but consistent behavioural

observations beyond 5 seconds were: static-very slowly diffusing, regionally localised diffusing, and some peripheral (possibly membranal) rapidly diffusing.

Exemplar wildtype cell A represents the most commonly observed signal behaviour in wildtype cells: highly static to slow-diffusing signal, confined to the periphery and proximal to the cell membrane. Whereas most off-target and rapidly-bleached signal disappears by 5 seconds into acquisition, this cell showed no fluorescent foci until 12.35 seconds into imaging (Figure 4.11A).

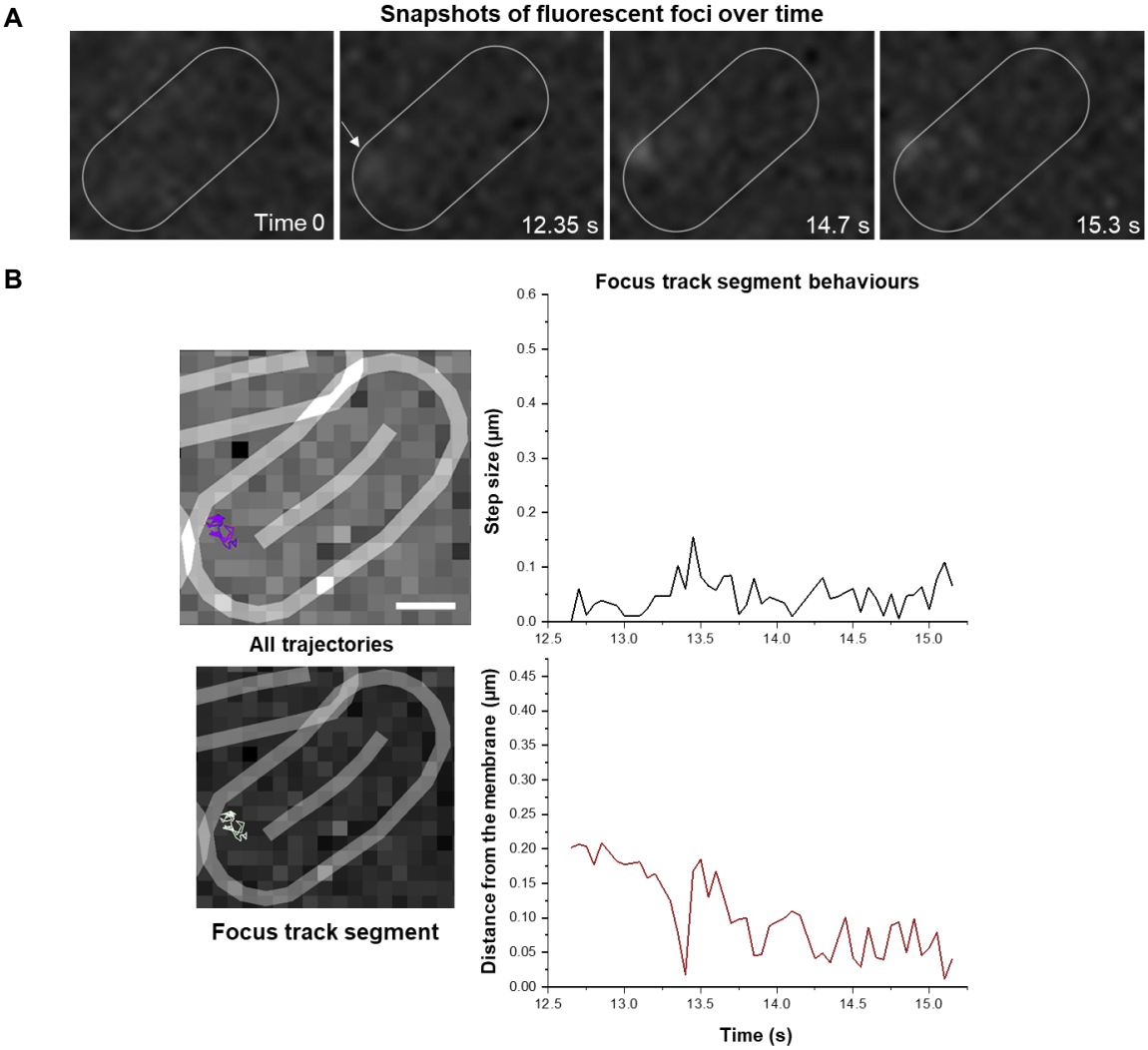


Figure 4.11 – 8mer-643/q1 visuals over time: wildtype exemplar cell A.
(A) Montage of cell stills, showing the appearance of fluorescent foci over time during a 100 second acquisition (50 ms exposure, ~10 mW 638 nm laser; still time given in seconds, lower corner of each still.
(B) *Moltrack* visuals for wildtype exemplar cell A: focus track segments and behaviours. *Moltrack* 'all trajectories' showing the trajectories for all tracked molecules in the cell (left); track colour denotes speed, where yellow is fastest and dark blue, slowest.
The 'focus track segment' denotes a specific track segment analysed in *Moltrack*; the righthand side plots, molecule step size (black) and distance from the membrane (burgundy) over time, are representations of the behaviours of the molecule in the focus track segment. Scale bar – 500 nm.
This flow applies to all further exemplar molecule figures in 4.5.

From 12.35-15.3 seconds, exemplar cell A generated a diffusely-bright focus, initially static at the cell periphery before appearing to slowly diffuse towards the midcell, then back to the cell periphery and disappearing. The high confinement and low diffusion around the cell membrane is confirmed by measuring the step size of this focus and its membrane proximity: for the ~3 second period, the consistent sub-0.10 μm step size (with the exception of 0.15 μm at 13.5 seconds, likely the slow diffusion motion) denotes a very slow-moving and close to static focus overall, increasingly membrane-proximal (within 100 nm of the membrane from 14.25 seconds onwards)(Figure 4.11B).

This was the only molecule that appeared in cell A (as denoted by the 'all trajectories' *Moltrack* output in Figure 4.11B), suggesting that whilst sufficient 8mer-643/q1 probes had been internalised, an off-target event only occurred once. This suggests that the probability of off-target hybridisation events is quite low across the cell population generally.

Exemplar cell B represents an example of the fast-diffusing focal behaviour observed behaviour in wildtype cells, one that is also observed in tag-strain cells. Excepting the single-frame-long bright midcell focus at 11.55 seconds, most fluorescent signal observed is highly diffuse and poorly distinguishable from time 0 to 16.7 seconds, synonymous with very rapidly-diffusing molecules (Figure 4.12A). The 'all trajectories' *Moltrack* output denotes tracked molecules that seemingly diffuse across the whole length and width of the cell, sometimes outlining the whole cell perimeter – rapidly, since most tracks seen here are yellow. Indeed, focus track segments between 0-8 seconds show very large steps sizes over time, corroborating very fast diffusion; focus track segment 2 even shows a molecule traversing the length of the cell in 0.7 seconds (Figure 4.12B). The focus track segments varied from 200-400 nm from the membrane, indicating little consistency or specific correlation (Figure 4.12B).

The combination of the diffuse signal, high speed diffusion and highly variable membrane proximity suggests the presence of multiple molecules, and perhaps momentary aggregation for the single bright focus. The erratic motion that appears to cover the whole surface and length of the cell could correspond to prior hypotheses about probe trapping within the membrane or the periplasm.

Following the higher potential for spurious signal pre-5 seconds, and evidence from populational studies that signals occurring later in acquisition are more likely to be specific complex hybridisation, the exemplar molecules from the tag strains and their focus track segments shown are mainly >5 seconds, often well into the acquisition period.

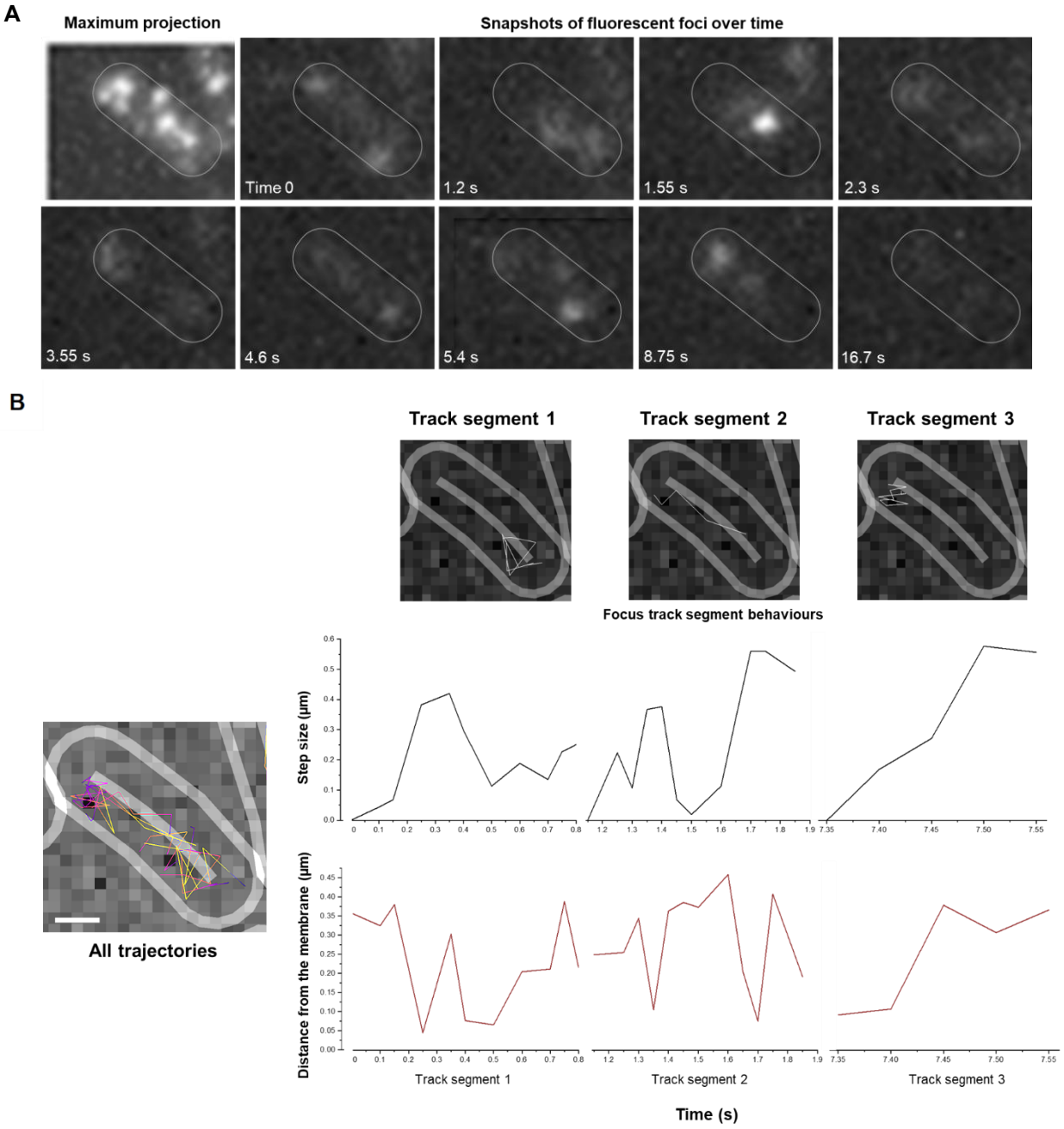


Figure 4.12 – 8mer-643/q1 visuals over time: wildtype exemplar cell B.

Induced p6

The p6 cells principally exhibited extremely bright foci, suggesting that multiple probes may be able to bind to the six possible complementary binding sites on the mRNA; this brightness could also be explained by proximity to the cell surface, within the region of the HILO illumination angle. Most loaded cells presented bright foci that diffused across large portions of the cell, consistent with fast diffusion of released transcripts, but often showed periods of slower diffusion and confinement, particularly at the cell poles and periphery. This could correlate to diffusing through denser subcellular compartments, or transient capture by proteins in the cell periphery, particularly membrane-associated degradation machinery.

Exemplar cell A (Figure 4.13) represents all possible diffusion behaviours observed in p6 cells: rapidly diffusing foci, slower diffusing foci constrained regionally, and highly static, very constrained signals. A cell with much diffuse signal at time 0 (possibly a "racetrack" too), individual intermediary-diffusing bright foci are visible post-5 second bleach-down, alongside consistent rapidly-diffusing, diffuse molecules, until there was no remaining fluorescent signal was visible at ~30 seconds. This is suggestive of released mRNAs (or potential non-specific signal) for the diffuse, rapidly diffusing signals, and potential transcription events for the transient bright foci.

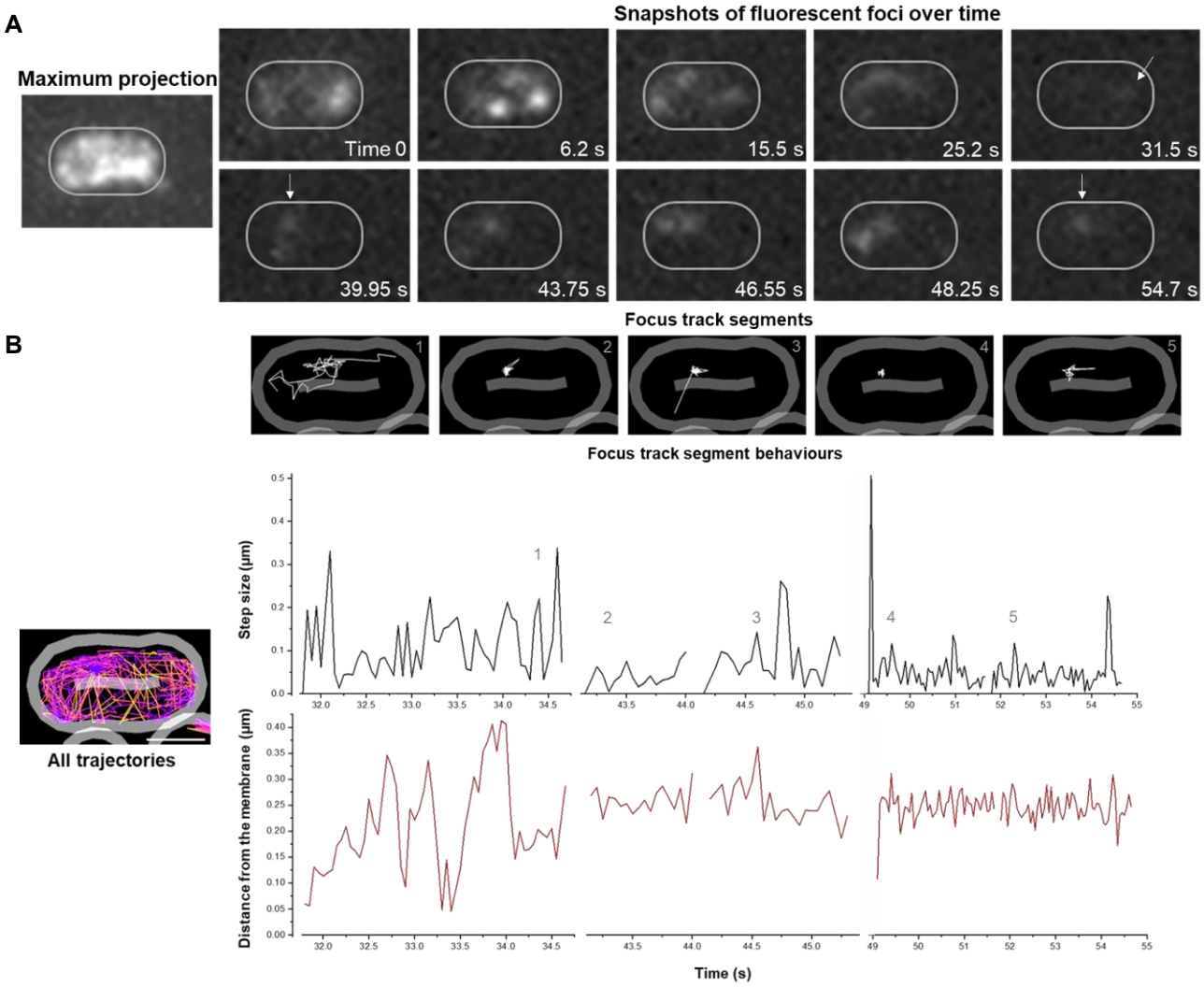


Figure 4.13 – 8mer-643/q1 visuals over time: p6 exemplar cell A.

Interestingly, 31.5 seconds into the 100 second acquisition, a very subtle focus appeared (Figure 4.13A white arrow, focus track segment 1 in 4.13B). This focus then gained brightness and diffused from the right upper cell pole across the length of the cell, as shown by the 0.1-0.3 µm variable step sizes from 31.5-35 seconds, before becoming highly bound around the midcell region – as shown by the dim focus at 39.95 seconds. Over the next ~15 seconds, the

dim static focus remained completely bound, highly confined ~250 nm from the membrane, before disappearing at ~54.8 seconds (Figure 4.13B).

Notably, the bound focus appeared to be releasing other dim foci between 32.25-54.8 seconds – as if perhaps a central transcription site had formed, and nascent and released transcripts were budding off and around this transcription core. Should there be a transcription focus and hence multiple mRNAs, taking a mean intensity profile of the focus should show variations in intensities as mRNAs are transcribed (increased intensity) and released (decreased intensity)(Figure 4.14). Indeed, correlating the approximate p6 unitary intensity with the focal trace suggests a core focus with >10 interacting molecules over the 20 second period if assuming one probe binding each molecule – with an estimated ~3.2 second transcription time for p6 mRNAs, at least 6 mRNAs could be transcribed over this period.

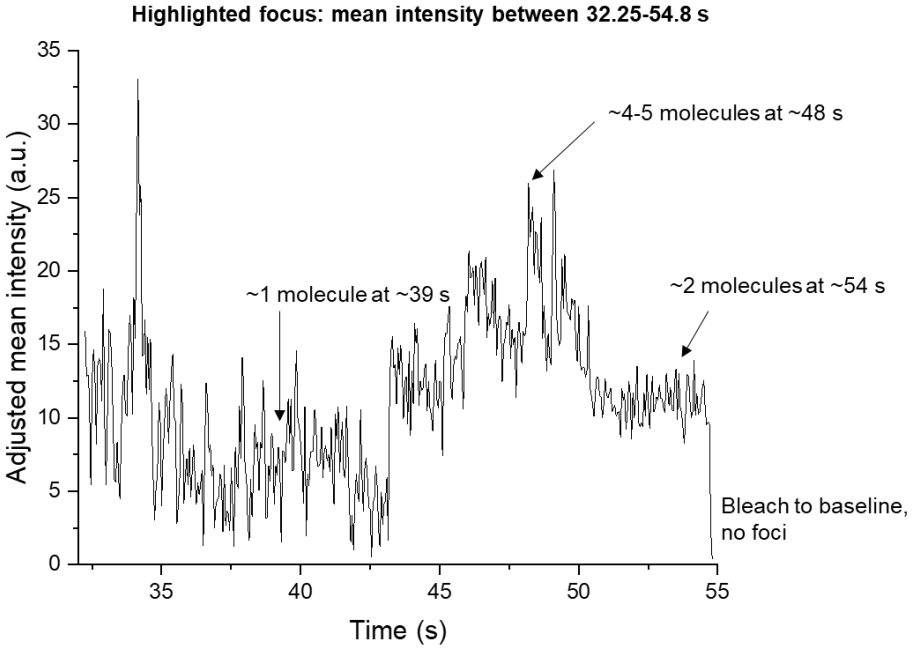


Figure 4.14 – Adjusted mean intensity trace of the upper midcell static focus in p6 exemplar cell A, between 32.25-54.8 seconds.
Annotations estimate the number of possible molecules observed, correlated to the intensity variations by approximating the unitary intensity at ~6 a.u.

Focus track segments 2-5 shown here present the tight focus observed, alongside the relatively consistent <0.1 μm molecular steps (Figure 4.13B). Spikes in the step sizes could correspond to the additional foci observed near the central focus interacting with the core focus PSF and subsequent incorporation in the measurement. It is quite remarkable to see ~3 foci around the static focus over the 15 second duration: taking the p6 mRNA transcription

time of 3.2 seconds, the 10 seconds needed to processively transcribe 3 mRNAs within this time period is feasible, suggesting that the dim focus could indeed be a core transcription site.

Exemplar cell B (Figure 4.15A) is another example of the continuum between diffusing and bound foci. No fluorescent signal is seen in the cell until 4.7 seconds into acquisition, where a bright focus appears at the upper cell pole. This bright focus diffuses around the upper midcell for the next ~2.5 seconds – as shown by the large step sizes and highly variable 50-450 nm membrane proximity – before becoming highly confined at the left upper cell periphery. This static confinement can be seen at 7.5 seconds, overall lasting for ~1.5 seconds between 6.75-8.25 seconds where the molecular step size is ~0.05 µm, within an average of 150 nm of the membrane (Figure 4.15B). The bright focus resumes its diffusion across the whole upper cell region as can be seen from stills 13-18.85 seconds, before disappearing.

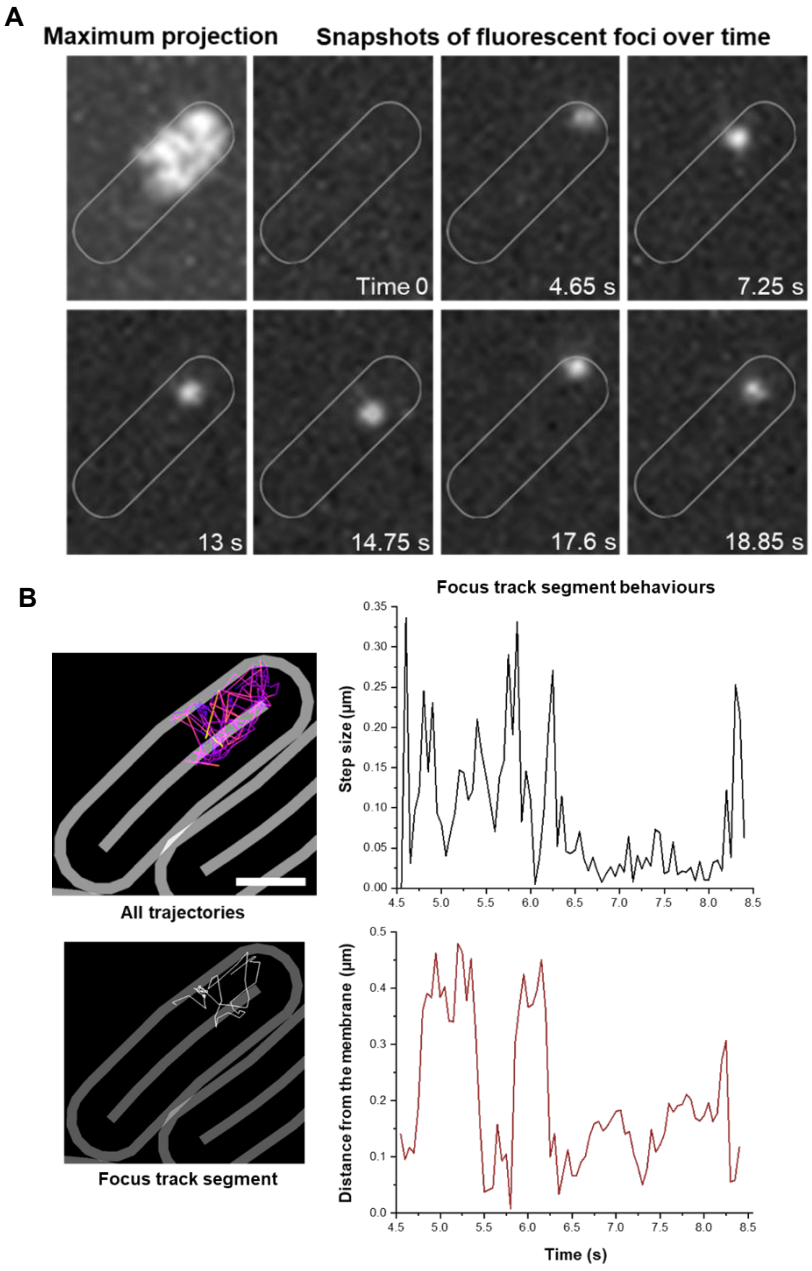


Figure 4.15 – 8mer-643/q1 visuals over time: p6 exemplar cell B.

Induced gT7

With a longer mRNA but no intact RBS, the gT7 mRNA was anticipated to exhibit several states: a tight bound state during early transcription elongation, followed by slower, localised diffusion that could correspond to slow subdiffusion of a tethered mRNA, followed by faster diffusion associated with a released or trafficked mRNA.

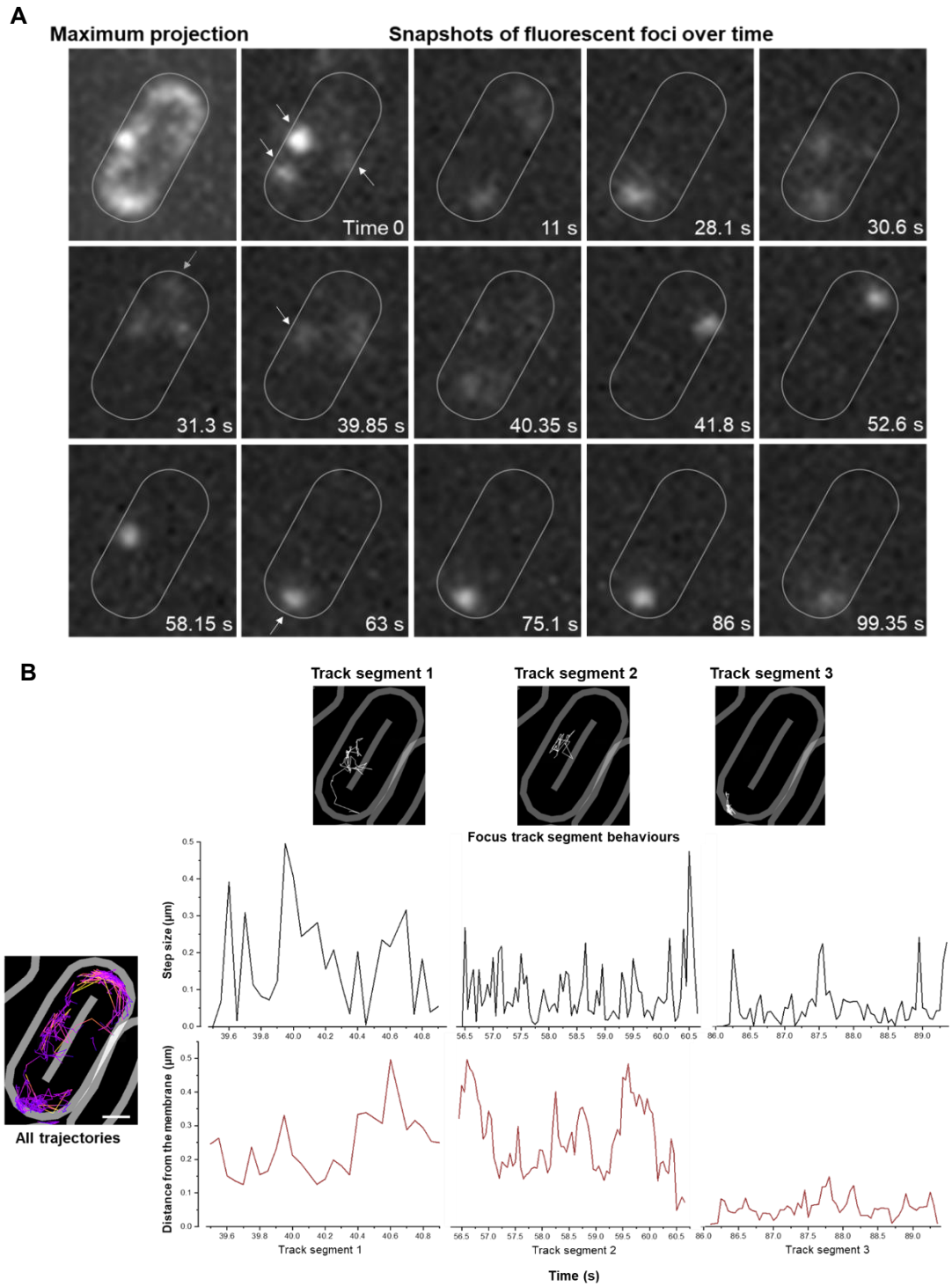


Figure 4.16 – 8mer-643/q1 visuals over time: gT7 exemplar cell A.

Exemplar cell A (Figure 4.16A) demonstrates all these behaviours (representatively – but these behaviours together are not so often seen in a single cell, partial capture more common). Much like the relatively loaded p6 cell A, a similar phenotype is apparent for gT7 cell A, with ~ 1-3 distinguishable fluorescent foci (highlighted by arrows on 4.16A) – some localising to the membrane – up to as long as 35 seconds into acquisition. Notably, there appear to be three confined regions in the gT7 cell: bound foci at both cell poles, and at the midcell. This is concurrent with the peripheral but more central chromosomal transcription locus of the gT7 insert, and ribosomal hotspots to which mRNAs are trafficked – despite the lack of a RBS and hence no translation, it is still most likely that the gT7 mRNAs would localise to the cell poles.

From time 0 to 39.85 seconds, multiple diffuse and bright foci can be seen around the midcell and cell poles, exhibiting variable mobilities. Focus track segment 1 denotes one of these left midcell foci, as indicated at 39.85-40.35 seconds. Presenting a similar behaviour as described for p6 cell B: over ~1.5 seconds, a molecule appears to break away from a subtle apparent core focus in the left midcell, diffuse down the periphery of the cell and around the bottom pole, before rapidly diffusing off over an ~1.5 second period (Figure 4.16B). This diffusion behaviour can be seen by the 0.025-0.5 μm widely variable step sizes, 100-500 nm distances from the membrane.

From 40.9 seconds on, rapid diffusion of foci appears to continue, where foci seem to consistently localise at a regionally-localised dim focus transiently seen at the upper cell pole, as highlighted at 31.3 seconds (grey arrow). This focal colocalization continues at the bright focus as seen at 52.6 seconds, with transient colocalization over ~15 seconds. At ~55 seconds, this bright focus detaches, slowly hopping down the midcell as can be seen at 58.15 seconds. At the midcell, slow regional diffusion occurs – predominantly 0.05-0.2 μm steps centrally (200-500 nm from the membrane) as shown by focus track segment 2 – before diffusing to the lower left cell pole as seen at 63 seconds.

A sustained bright focus, it slowly diffuses back and forth along the cell pole for the next ~17 seconds displaying variations in brightness and resolution. At ~86 seconds, the focus regains brightness and becomes bound at the lower left cell pole for ~6 seconds – denoted by focus track segment 3, where the focus is consistently within 50 nm of the membrane, and averages steps of 0.1 μm for ~3 seconds. For the remaining ~28 seconds of acquisition, slow regional subdiffusion with variable brightness is visible around the lower cell pole, as seen at 99.35 seconds; the focus remains visible until the end of the 100 second acquisition. Interestingly, no mRNA release-like foci are seen, solely transitions between dim-very bright of the singular

slow-diffusing focus from 60.45-99.35 seconds (Figure 4.17). Taking the gT7 unitary intensity of ~6 a.u., the intensity trace variation between ~10-30 a.u. for a single visible focus could suggest the binding of 1-5 probes to the same mRNA, sequestered at the cell pole. This is possible on this ~30 second timescale, since the gT7 mRNA, with no RBS, would not be translated but would likely localise to the poles before degradation (on the scale of >2 min (Esquerré et al., 2014)). This visible brightness variation could also, however, simply be HILO mode imaging at different penetration depths across the cell length.

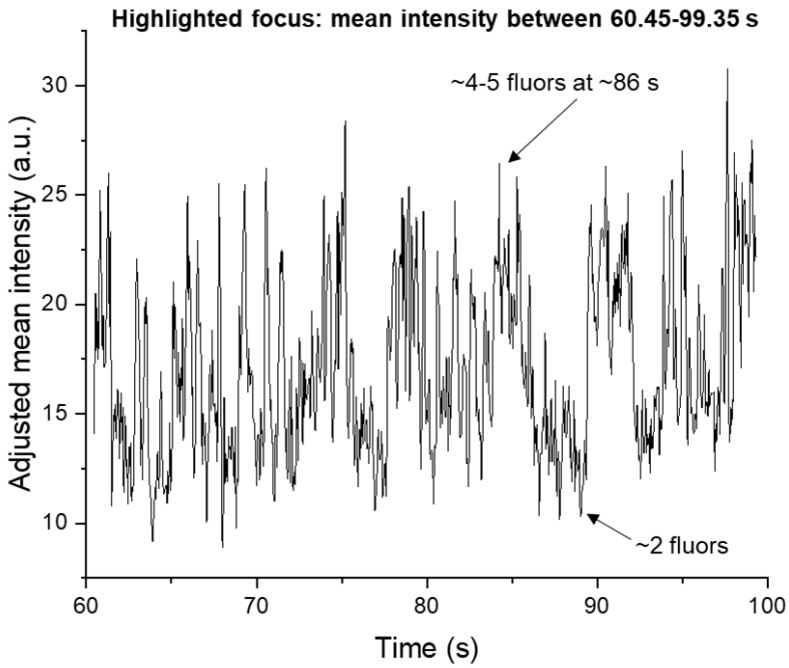


Figure 4.17 – Adjusted mean intensity trace of the upper midcell static focus in gT7 exemplar cell A, between 60.45-99.35 seconds.
Annotations estimate the number of possible probe binding events observed, correlated to the intensity variations by approximating the unitary intensity at ~6 a.u.

The recurrent foci at both cell poles – the upper right focus from repeatedly as seen from 11-52.6 second stills, and the lower 63-99.35 second regional focus – could indeed represent mRNA hotspots, in line with ribosomal hotspots for coordinated transcription-translation. An opposing interpretation could be that these are simply those probes stuck in the lipid membrane – but the ATTO643 fluorophore, even with improved photostability, is unlikely to sustain localised bright signals with such defined characteristics for so long, and free dye would not be resolvable. Indeed, these long, highly regionally localised and interconverting bound-diffusing mobilities are not observed in the wildtype control, indicating that these are real complex observations. Accounting for a likely gT7 transcription site around the midcell region (and hence midcell bright foci), and the usual trafficking of mRNAs to the cell periphery

and poles for translation, it is likely that these observations do correlate with gT7 mRNA transcription and trafficking.

Exemplar cell B (Figure 4.18) presents another noteworthy phenomenon: mid-division cells, with shared bright foci. Dividing cells are commonly observed across all cell samples, and it has been challenging as to how to classify them: one cell, or two?

In the case of classification as one cell – visually, by the extent of an observable septum separating the daughter cells – the cells often are highly loaded with visible fluorophores from laser activation, and exhibit multiple focal characteristics.

From a saturated cell at time 0, 3 distinct fluorescent foci are distinguishable at 14.25 seconds, where the 2 midcell foci are very bright. The lower bright midcell focus as indicated is then visible, highly confined between 14.25-29.75 seconds (Figure 4.18A); and almost seems as though a second focus is overlapping at 22.55 seconds, indicative of foci appearing released at the same locus – likely a transcription locus.

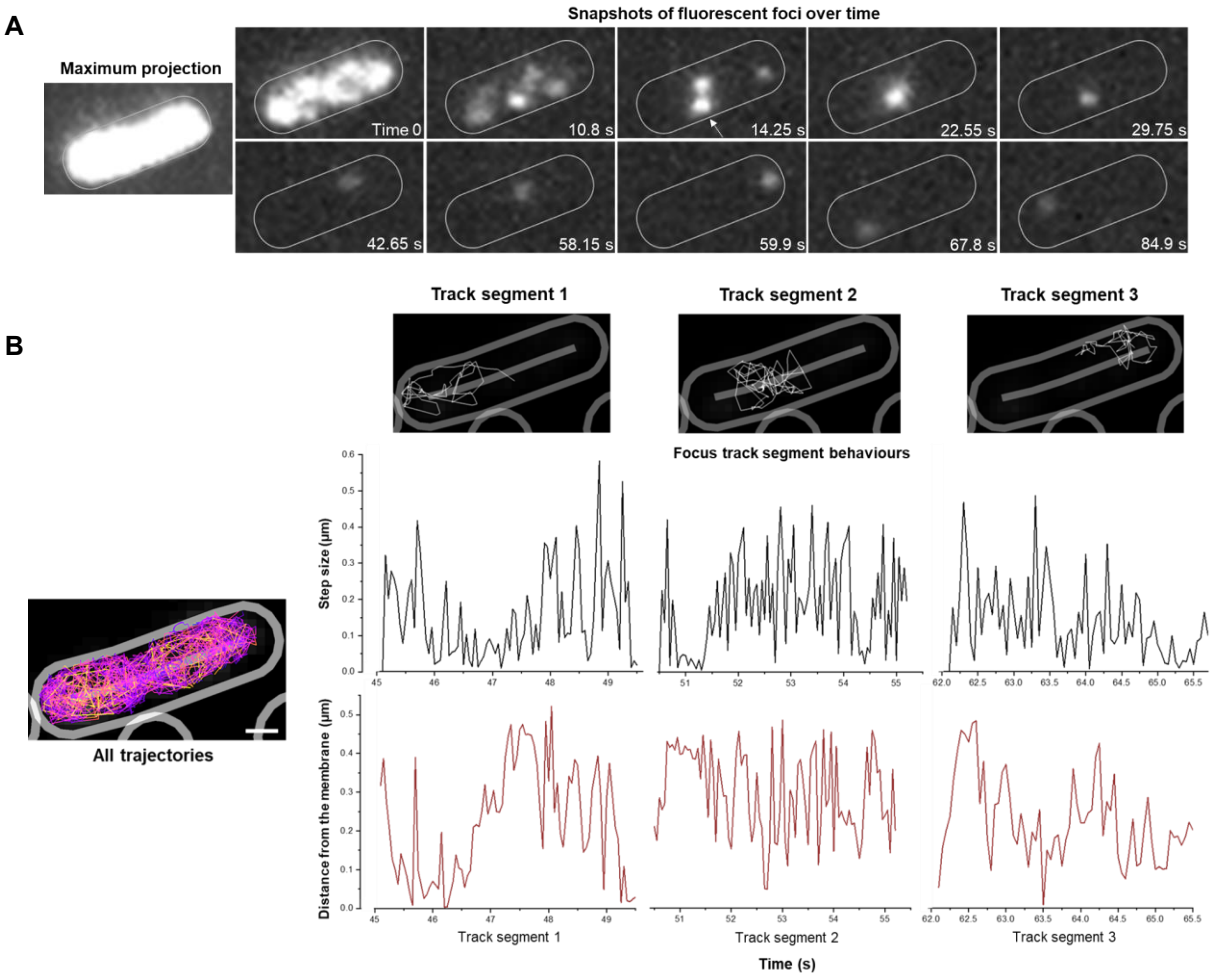


Figure 4.18 – 8mer-643/q1 visuals over time: gT7 exemplar cell B.

From 29.75 seconds onwards, a single molecule can be seen rapidly diffusing across the whole length of the conjoined, dividing cells, before bleaching at 86.5 second: consistent with RNA release. Between 42.65 seconds and 84.9 seconds of the acquisition, the bright focus can be seen to diffuse widely, become transiently more tightly regionally localised three times (lower cell pole, midcell at 58.15 seconds, and upper cell pole at 59.9 seconds in the stills) before diffusing around the cell again. These focal mobility behaviours, all variably distant from the membrane, represented by the three track segments, all denoting predominant regional diffusion behaviours with typically 0.2-0.5 μm steps, and transient binding as shown by 0.05 μm steps for ~0.5-1 second (Figure 4.18B).

rpoB-tag

Far fewer *rpoB*-tag cells exhibited clear bright foci with 50 nM 8mer-643/q1 probes; further, many cells seemed to have a diffuse, almost autofluorescence-like appearance, which was absent in all other strains, wildtype inclusive. The cause of the latter is unknown – perhaps an artefact of membrane interactions – but certainly, more molecular events may be observed with a higher initial probe concentration.

Much like the gT7 cells, there are three consistent loci observed in *rpoB*-tag cells: localisations at the midcell and cell poles – corresponding to the transcription locus, and ribosomal translational hotspots. In the case of *rpoB/rpoBC*, a much longer, essential mRNA, it would be anticipated that translation could occur at the midcell transcription locus and more expressly at the cell poles. There are a number of consistent behaviours that can be observed for foci in *rpoB*-tag cells.

Exemplar cell A (Figure 4.19) excitingly denotes a possible representation of the active transcription of an *rpoB* mRNA. As priorly estimated, full transcription of a 4000-8329 nt *rpoB/rpoBC* transcript would take on the order of 100-225 seconds – unlikely to be fully captured without a higher quenched probe pool available. Nevertheless, the majority of trajectories in cell A shows a very slowly diffusing, increasingly static particle at the upper pole (Figure 4.19B).

From time 0 to 25 seconds, recurrent foci are visible at the cell poles, and a particularly bright focus static at the midcell for ~9 seconds, whereby more diffuse foci seemed to diffuse in and out of it as seen in the gT7 cells. The diffuse focus highlighted at 25.1 seconds in the upper pole appeared statically localised until ~40 seconds through the acquisition, transiently disappearing and becoming visible again at the ~52 seconds (Figure 4.19A). For ~4.5 seconds

from ~52-56 seconds, this polar focus appeared to detach and very slowly diffuse around the pole periphery, constrained to ≤300 nm of the membrane with sub-0.15 μm steps as shown by focus track segments 1-2 (Figure 4.19B), before disappearing. This very slow polar diffusion over the course of ~5 seconds is consistent with chromosomal diffusion (Javer et al., 2013).

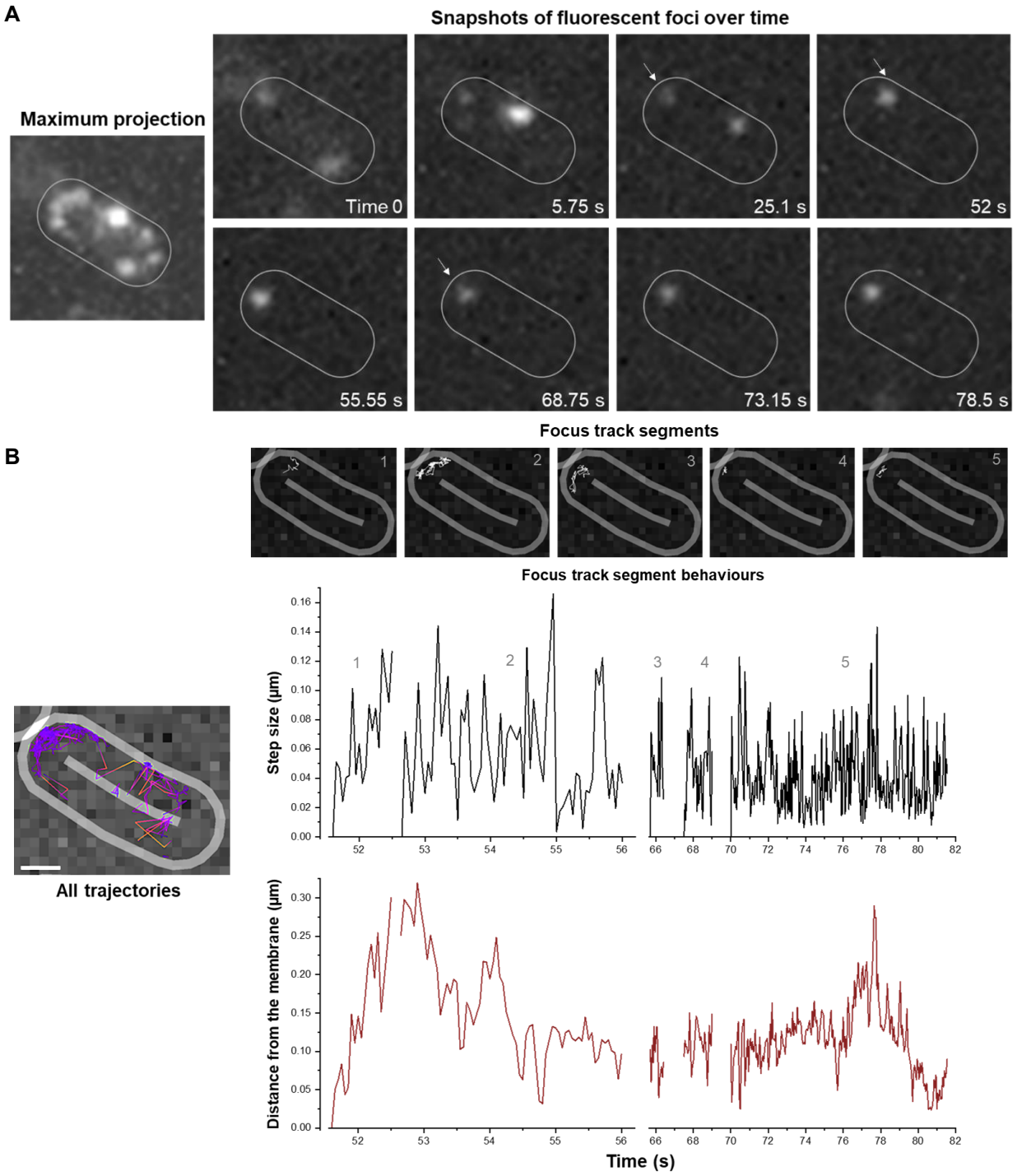


Figure 4.19 – 8mer-643/q1 visuals over time: *rpoB*-tag exemplar cell A.

At ~66 seconds of the 100 second acquisition, a bright focus appeared at the same upper pole locus, visibly constrained albeit with some apparent blinking. Over the 66-82 second period shown in focus track segments 3-5, the bright focus remained principally bound, with occasional detachment events leading to slow localised diffusion back to the original locus. The small step sizes of 0.01-0.1 μm, the smallest seen yet, mostly with 150 nm of the membrane, corroborate a tightly localised, largely immobilised focus, possibly corresponding to an mRNA engaged in polysomal translation.

These small subdiffusive rearrangements and predominant static constrainment continued until the focus disappeared at ~83 seconds; this high confinement is shown in stills 68.75-78.5 seconds. Taken together, the video and tracking analyses show ~30 seconds' constrained polar localisation, and indeed possibly longer if accounting for earlier foci at the same locus. This could indeed be indicative of a translation hotspot, whereby multiple ribosomes in the cell pole are available to engage in processive mRNA translation, particularly in the case of essential gene mRNAs such as *rpoB*.

Exemplar cell B presents all mobility behaviours: rapid diffusion, slow regional diffusion and static foci. Indeed, static foci can be seen at time 0, as indicated by the second midcell bright focus, shown to be tightly bound at a relatively consistent ~250 nm from the membrane with <0.1 μm steps from ~3.5-6.5 seconds (focus track segment 1, Figure 4.20B). Following the premise of focusing on exemplar molecules after the first 5 seconds, 1-3 bright foci can be consistently seen from 19.75 seconds onwards. Three recurrent foci of variable brightness appear to diffuse around the cell poles and periphery for the next ~30 seconds, from 21.85-39.8 seconds. The notably highly bright foci at the midcell (24.95, 33.75 and 39.8 seconds) could correspond to transcription hotspots.

Given the L-macrodomain location of the essential operon in the genome (highly accessible for rapid expression), and Sattler & Graumann's observation of ribosomal L1 protein localising near its transcription locus (Sattler & Graumann, 2021), i.e. within the same polycistronic operon, it is likely that both transcription and translation of *rpoB/rpoBC* occurs at the same midcell region. The transient high brightnesses observed could also represent the superposition of multiple molecules at the same core locus, particularly as this localised variable brightness behaviour was not observed in wildtype cells.

At 43.8 seconds through acquisition, a distinct bright focus appears proximal to the upper cell pole. This bright focus proceeds to slowly diffuse and explore the upper polar region of the cell for the next ~20 seconds as shown between 43.8-65.3 seconds, intermittently static at the cell

pole, before disappearing. Focus track segment 2 represents ~5 seconds of this polar exploration: with step sizes mostly between 0.1-0.2 μm and anywhere between 25-400 nm distance from the membrane, this slow polar diffusion is most characteristic of a tethered *rpoB/rpoBC* mRNA engaging with ribosomes for translation (Figure 4.20B).

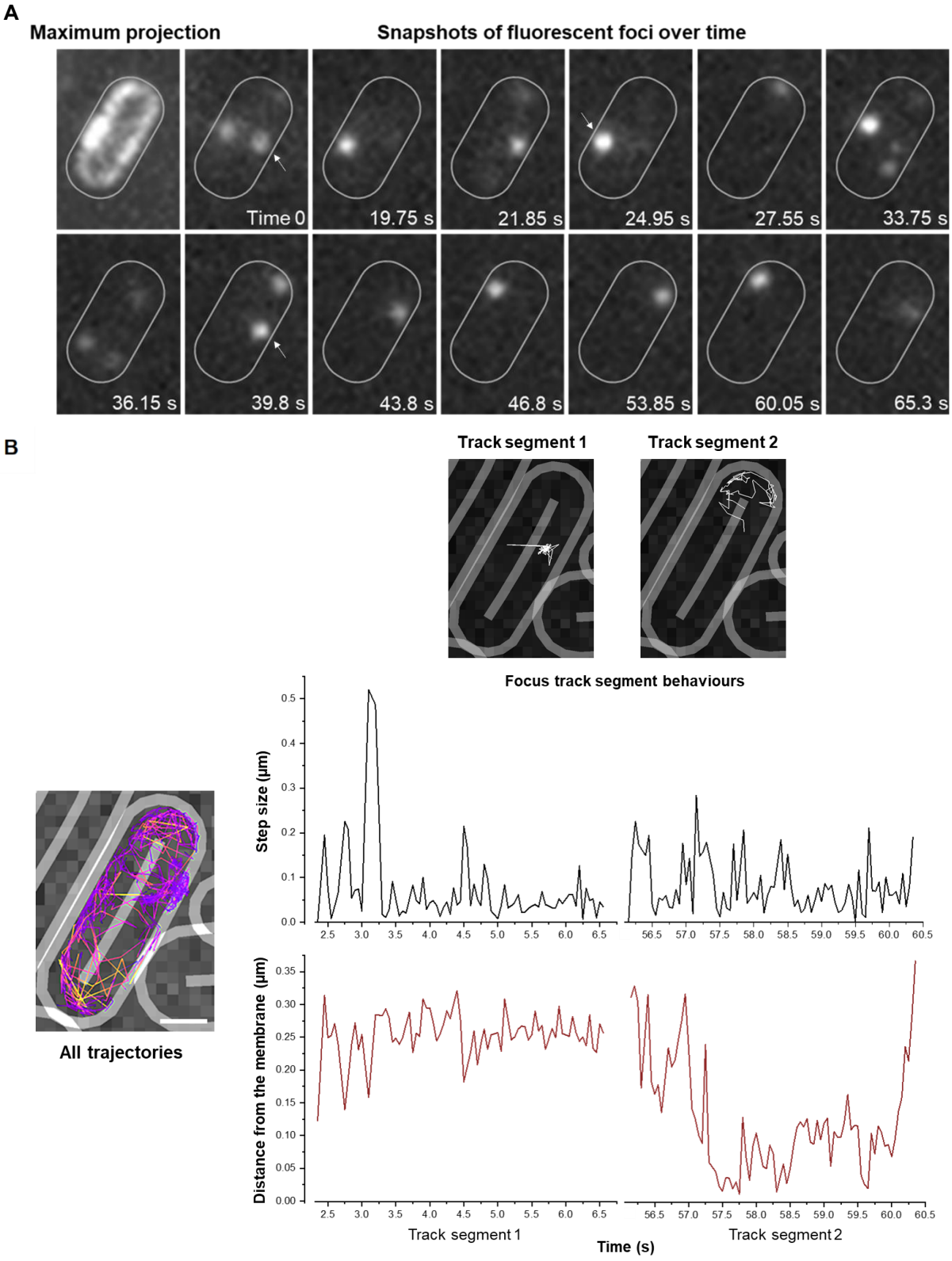


Figure 4.20 – 8mer-643/q1 visuals over time: *rpoB*-tag exemplar cell B.

Within a 100 second acquisition, it would not be possible to visualize the entire lifecycle of an *rpoB/rpoBC* mRNA, from nascence to full translation. Taking 40 nt/s as the rate of transcription to give 100-225 transcription time (Vogel & Jensen, 1994), and 60 nt/s translation rate for a corresponding ~67-139 second translation time for *rpoB/rpoBC* (Bremer & Dennis, 2008; Young & Bremer, 1976), the entire mRNA life cycle would be on the order of ~180-365 seconds, give or take trafficking time. It is hence most likely that partial observations of both transcription and translation elongation are made – and indeed, possibly coupled transcription-translation. As such, the midcell observations most probably correspond to visualization of *rpoB/rpoBC* transcription, potentially coupled with translation; polar foci most likely translation, including dense polysomal engagement of the mRNA.

It must be noted that accurate time correlations for *rpoB*, aside from all the evident possible confounders already discussed, are complicated further by the operonic structure and essentiality of the mRNAs produced. Actually, since the *rpoB* non-dominant promoter is poorly described(M. Yamamoto & Nomura, 1979), it is unclear whether *rpoB* mRNA is taken solely, whether it is still linked to the upstream ribosomal protein mRNAs, or whether it is primarily transcribed alongside *rpoC* (Barry et al., 1980; Downing & Dennis, 1987; Ralling & Linn, 1984). Indeed, it is highly likely that all three mRNA variants (Chow & Dennis, 1994), and more, are possible, and the proportionalities of each determined by the exquisite regulation enacted by the cell in controlling transcription, translation and environmental interaction.

4.6 Discussion

4.6.1 *rpoB* transcription – a complex machine

As alluded to in the rationale, it turns out that this essential gene target is not the simplest, nor most easily explicable regulated site.

Polycistronic mRNA processing

The dominant promoters in the *rplKAJL-rpoBC* operon have been well-documented; Downing & Dennis further identified polycistronic mRNA processing via attenuation and RNase III processing (Downing & Dennis, 1987), corroborated via *rplL* attenuation by Barry et al. (Barry et al., 1980). Whilst there appears to be higher transcription frequencies and post-processing of the small ribosomal protein mRNAs, less seems to be known about *rpoB-rpoC*. One study by Severinov et al. into the tethering of the large β and β' RNAP subunits showed that, amazingly, an *rpoBC* fusion polypeptide gives rise to a functional RNAP (Severinov et al., 1997). With *rpoB* and *rpoC* separated only by a short linker sequence and actively co-transcribed, the

implication therefore is that no post-transcriptional processing of the digenic mRNA is necessary, and translation can commence rapidly (Severinov et al., 1997).

Accounting for the multiple promoter sites, processing of upstream *rplK-rplJ* cistrons, and the possible lack of cleavage of *rpoB*C – note that there are few reports specifically addressing this operon in the advent of higher resolution omic technologies – it seems that an 8329 nt transcript would often be anticipated. Irrespective of the strength or even actuality of the rpoBp, the engineered tag array would likely remain intact and transcribed in the 5' UTR of *rpoB* (out of range of RNase III cleavage of *rplL*).

As such, any arising signal probably does correspond to the transcription initiation of *rpoBC* – but, no reliable correlations can be drawn yet between the length of time a focus is observed for, particularly when "bound". This static behaviour could correspond to two likely scenarios: a continuous and lengthy transcription elongation event, or, a polysomal complex as processive translation of the ββ' mRNAs has begun.

4.6.2 On the fundamental importance of growth conditions

In another return to the discussion about the importance of the most fundamental elements of experimental design: the growth medium and growth phase of cells have been shown to significantly change the rates of transcription elongation, mRNA half-lives and decay interactions (Bernstein et al., 2002), specifically for highly accessible and transcribed regions of the nucleoid. The transcription elongation estimates applied throughout this Chapter correspond to an approximation of media-dependent doubling time and consequent transcription rates (Vogel et al., 1992; Vogel & Jensen, 1994). Notably, a 40 nt/s estimate is most suitable here, when also considering that elongation rates are much faster for certain essential operons (*rrn* operons as a classical example (Gotta et al., 1991)) and for intergenic regions (Dennis et al., 2009); but no specific profiling for *rpoBC* or for non-native insertions has been conducted.

Esquerré et al. reported that for MG1655 M9Glu-grown *E. coli*, with a doubling time of ~1 hour, mRNAs exhibited half-lives nearly half those of slower-growing *E. coli* (2.4 min versus 4.2 min at ~6 hour doubling time); and that mRNA degradation machinery expression is transcriptionally regulated by way of cellular mRNA concentration (Esquerré et al., 2014). In competent cell growth to early log phase, with a doubling time of ~2-3 hours for tag strains, it is possible that upregulated degradation within the cell could impact the observation time of probe-mRNA hybridisation complexes. If so, the 2.4 mins' lower bound half-life would occur

within the ~5 mins' feasible acquisition time with the fluorogenic probes currently trialled for probe-labelled single molecule mRNA imaging.

Here-in-lies another open question: alongside their observation of lower mRNA stability at higher growth rates, Esquerré et al. did further confirm stabilization of specific mRNAs (Esquerré et al., 2014; Lin et al., 2012; Nilsson et al., 1984). Now, this largely relates to membrane protein clusters, and the essential operon of *rpoB*C is not discussed. But, acknowledging the implications of managing cellular expression levels by regulating RNAP and ribosome concentrations – for which the *rplKAJL-rpoB*C operon is fundamental – it is possible that *rpoB* mRNA could be stabilized by ribosomal binding in translation initiation (Kaberdin & Bläsi, 2006; Richards et al., 2012). Moreover, there is a possibility that the 5' tag site near the RBS in the *rpoB*-tag mRNA could disrupt effective RNase E accessibility for 5'-dependent RppH-related degradation. This may not be so consequential, though, as essential mRNA degradation is likely feasible via RppH-independent RNase E and exoribonuclease pathways as well (Deana et al., 2008).

Continuing consolidation of *rpoB* as a serious target of interest – essentiality confirmed, but the pleiotropic bidirectional effects of *rpoB* are even wider than the implications discussed here. Mentioned in 4.2, rifampicin, contrary to its common inhibitory impact on cellular transcription, at low concentrations acts as an activator for *rpoBC* transcription (Howe et al., 1982; Newman et al., 1982). This is unsurprising from the purview of resistance mutations – since rifampicin affinity is for the β clamp subunit (RpoB) – and the aforementioned careful stoichiometric regulation of transcriptional/translational subunits in determining circulating machinery concentrations. As such, rifampicin binding to RNAPs and inhibiting autorepression of *rpoBC* transcription is a deliberate cell mechanism to replenish RNAP concentrations (Fukuda & Nagasawa-Fujimori, 1983). Although an observation of the more convoluted effect of rifampicin on *rpoBC* transcription was unlikely to be seen here due to the high rifampicin concentrations utilised, the extent of *rpoBC* intricate expression regulation is emphasized again.

Indeed, *rpoB* is one of the most highly mutable genomic regions, a "mutational hotspot" (Chattopadhyay et al., 2009), whether in relation to chemical stressors (like rifampicin) or growth-mediated effects (like M9Glu). This has been demonstrated through various adaptive laboratory evolution studies (LaCroix et al., 2015; Reynolds, 2000) – for *rpoB*, all to the end of increasing cell fitness, and consolidating the importance of this essential gene.

4.6.3 Discursive summary

There are limited reports of tracking mRNAs in the live bacterial cell, and even fewer reported experimental diffusion coefficients for most species in general – indeed, many seminal works are based off essential protein diffusion behaviours (Bakshi et al., 2012; Sanamrad et al., 2014; Stracy et al., 2015), and many mRNA mobility studies have linked their findings to these observations. This Chapter has enabled the contribution of further *in vivo* mRNA spatiotemporal localisation and mobility behaviour observations, for artificial and native *rpoB/rpoBC* mRNAs via engineered *in vivo* smFISH in the *E. coli* cell.

Spatial localization of mRNAs

As for key findings around mRNA diffusive properties, consistent observations using the MS2 aptamer system have shown two main mRNA subpopulations, a significant bound subpopulation primarily at the cell poles and periphery, and a mobile population appearing to more readily diffuse across the cell. Golding & Cox first demonstrated these mRNA localization trends in *E. coli* via their plasmid-mediated MS2-GFP and *lacZ*-96xMS2 tandem arrayed target (Golding & Cox, 2004). Sattler & Graumann, using their chromosomal MS2-mVenus FP aptameric system with only 2 tandem MS2 sites in the 3' UTR of target operons in *B. subtilis*, demonstrated that mRNAs can diffuse rapidly, throughout the cell (including nucleoidal regions), either as free mRNAs or in complexation with ribosomes. They also concluded that the more static populations are polarly and peripherally localised, ready for processive polysomal translation of mRNAs (Sattler & Graumann, 2021). Montero-Llopis et al. likewise showed these subcellular localisation propensities with MS2-GFP (48x tandem sites) in *C. crescentus* (Montero Llopis et al., 2010).

Promisingly, the observations made in this Chapter, particularly for the endogenous *rpoB/rpoBC* mRNA, corroborate the spatial localisation and visible mobility populations of these seminal works. Static or predominantly immobilised populations were mostly observed at the cell poles and periphery, in agreement with the above, further indicating the likelihood of active 70S and polysomal translation at the periphery and poles as shown by Bakshi et al., Sanamrad et al., and Mayer et al. (Bakshi et al., 2012; Mayer et al., 2021; Sanamrad et al., 2014). Immobilised populations at the midcell were indicative of longer cells, potentially in mid-replication, with increased peripheral space accumulating ribosomes for rapid essential protein synthesis (Chai et al., 2014; Sattler & Graumann, 2021) – which would include RNAP subunits, hence *rpoBC* mRNAs.

Diffusion behaviours of tracked mRNAs

Across all tag strain cells, a highly mobile population was observed, seemingly able to traverse the entire length of the cell over milliseconds-seconds, freely diffusing around the cell periphery and apparently through the nucleoid too. This further corresponds to the seminal observations noted that mRNAs can freely diffuse throughout the cell on the scale of a few seconds, despite their long lengths and large sizes. Additionally, Sattler & Graumann indicated that artificial mRNAs behave similarly to native mRNAs, suggesting that observations across p6, gT7 and *rpoB*-tag cells are likely to be comparable to specific transcription locus, transcript size, and translation in the case of *rpoB/BC*.

The topic of apparent diffusion coefficients and quantified mobility states is a bit more complex. Interestingly, whilst the apparent diffusion coefficient data obtained here do show similar trends to published mRNA mobility populations, these D* values are comparably faster across all subpopulations. Golding & Cox first reported an estimated D* of ~0.03 $\mu m^2/s$ for diffusing plasmid mRNAs in *E. coli*, albeit ranging from 0.001-0.03 $\mu m^2/s$ across the 7 cells tracked (Golding & Cox, 2004). Montero-Llopis et al. estimated mRNA diffusion rates from their FISH approach, hybridising LNA probes to a 120x *lacO* array adjacent to a native mRNA, in *C. crescentus* – estimating length-dependent D*s for free or ribosomal-bound mRNAs, where 6.3 kb target mRNAs were freely diffusing at ~0.5 $\mu m^2/s$, and at 0.1 $\mu m^2/s$ when loaded with ribosomes. However, these models were calculated for application with their fixed FISH observations, where single molecule sensitivity was not guaranteed.

Sattler & Graumann performed substantial diffusion coefficient analyses for their 4 operonic chromosomal mRNAs in *B. subtilis*, obtaining 3 populations: a ~30% static population at ~0.003 $\mu m^2/s$, a ~40% medium-mobile fraction at ~0.02 $\mu m^2/s$ and a ~30% fast diffusing population at ~0.1 $\mu m^2/s$. The authors also indicated that a 3-populational model best explained the mobility populations, though without definite proof of populations beyond bound and freely diffusing populations.

Although the Chapter 4-reported D* for the mRNAs here are an order of magnitude higher than those in the literature, multiple populations beyond just bound and freely diffusing appear to best describe the mobility fractions, and indeed similar fractions are observed. Considering the chromosomal gT7 and *rpoB/rpoBC* mRNAs, both present a bound fraction (24-47%) at 0.1 $\mu m^2/s$, an intermediary diffusing fraction (31-44%) at ~0.3 $\mu m^2/s$, and a fast-diffusing fraction (22-31%) at 1.0 $\mu m^2/s$. This certainly follows the trends observed by Sattler & Graumann, particularly by subpopulation percentages.

Notably, Sattler & Graumann also reported that their diffusion rates were similar across distinct mRNAs, with considerably slower diffusion than large proteins even at upper bound freely diffusing behaviours (Bakshi et al., 2011; Mayer et al., 2021; Sattler & Graumann, 2021; Schibany et al., 2018). The mobility populations seen in tag-strain cells seem to corroborate this likeness, although overlap more with reported protein diffusion rates. One thought as to non-specific signal: Sattler & Graumann also noted static and fast-diffusing fraction differences in cells when measuring free mVenus versus mRNA-mVenus complexes – much like free probes sporadically non-specifically binding in wildtype cells. With wildtype off-target signals in *in vivo* smFISH holding lower-shifted D* fractions values than in tag strain cells, and most wildtype signals absent beyond the initial 5-10 second rapid bleachdown, analysis of molecular mobilities after 5-10 seconds could perhaps better represent specific mRNA expression events.

Indeed, the D* values obtained here are also quite aligned with the Montero-Llopis models (with gT7 as ~2.7 kb and *rpoB/rpoBC* as ~4-8.3 kb), considering that the FISH-estimated model by Montero-Llopis could be more representative of a probe-hybridised mRNA. There are therefore certain caveats to these comparisons: many of the cited literature span a number of different bacterial species – where Sattler & Graumann specifically chose *B. subtilis* for its spatial segregation of transcription and translation (Lewis et al., 2000) – and these primary *in vivo* bacterial reports mainly utilise the MS2 aptamer system. With the tandem arraying and binding of multiple FPs, adding significant bulk to the target mRNA could certainly slow all observed mobility fractions; and single molecule resolution may be less easily distinguishable with the diffuse FP background fluorescence than with organic dyes. Achieving a suitable comparator for defining the accuracy and indeed implications of transiently-tagged single mRNAs by small, organic-dye labelled probes, thus remains a challenge.

Complex interactions in the cell milieu

Considering the context of the crowded cytoplasm and increasing evidence of macromolecular crowding kinetics, subcellular domains interplaying with gene expression (Norred et al., 2018; Tabaka et al., 2014), there are many possible interactions at both the level of the probes (membrane implantation, transient protein interactions) and of the target RNAs especially.

Take *rpoB/rpoBC* as an example: an extremely lengthy essential RNA, it is likely to be nascently bound for much of its lifetime, and experience polysomal translation for rapid protein output. On top of this, regional macromolecular crowding could exert specific thermodynamic folding

effects on the RNAs (Dupuis et al., 2014), not to mention a whole host of possible transient interactions leading up to RNase E complexation in RNA degradation. As such, determining biological implications of target RNA subdiffusion requires further complementary approaches to quantitatively verify these observations.

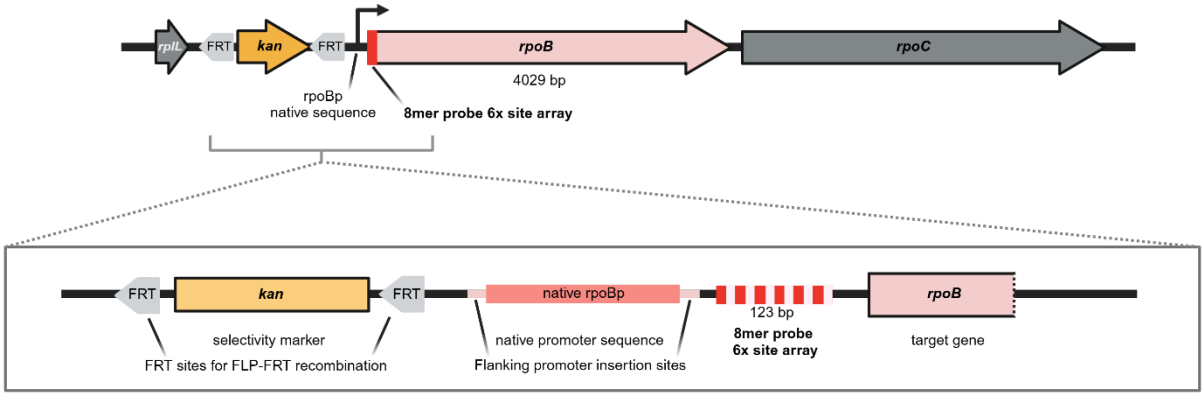
Nonetheless, the apparent polar localisation of the majority of localisations for all tag strains and visuals of rapid diffusion over the whole cell length corroborates with established reports of mRNA subcellular localization, perhaps in association with their protein output (Bakshi et al., 2015; Castellana et al., 2016; Moffitt et al., 2016; Nevo-Dinur et al., 2011). In the case of *rpoBC*, this could be of notable interest around evidencing of co-transcription/translation hotspots.

Other than the photon-budget limitations of the fluorophores in use and components of the imaging process, there are many cellular parameters *in vivo* that could influence particle lifetimes and their behaviours. Thinking about the organic dye-labelled ssDNAs specifically, little is known about their degradation rates, proclivity to aggregate or be sequestered within the cell. With such short length, it would be unsurprising should the probes survive longer in the cell before degradation, due to projected transient/unstable SSB and hence exonuclease interactions (Han et al., 2006); the terminal dye labels may also confer some protection.

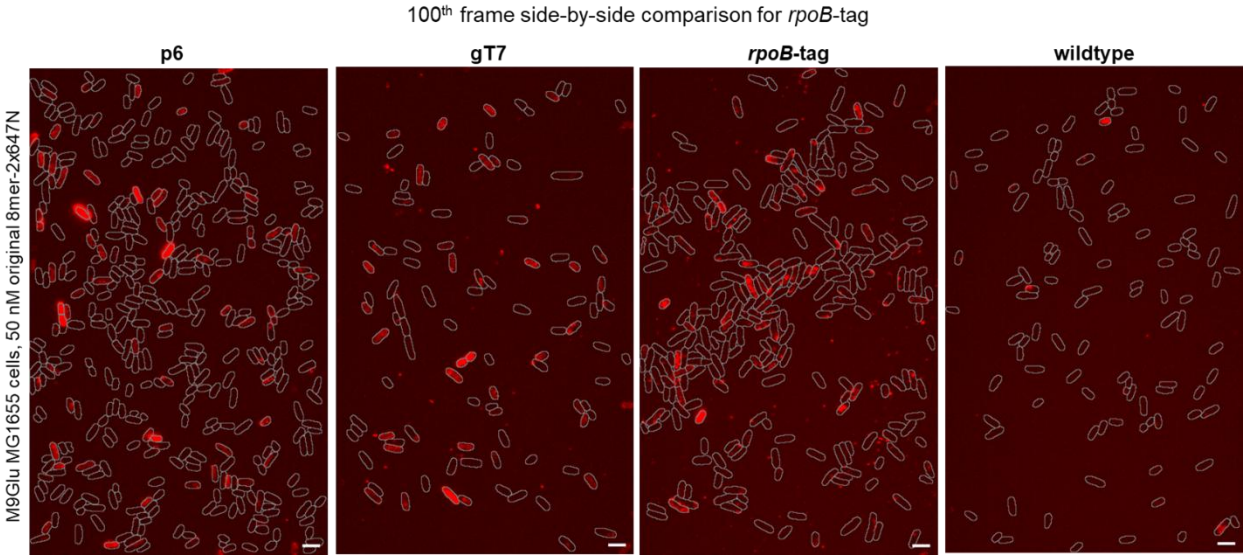
All-in-all, interpreting the molecular characteristics of probe-bound *rpoB/rpoBC* mRNAs is challenging. Beyond estimations of transcription durations and regional localisation, correlating the lifetime of an observed particle to initiation, elongation, spatial diffusion on release is limited, even with mobility information. Further concurrent reports of signal co-localisation, using key proteins (RNAP, RNase E) in the cell, could help to further classify labelled mRNA states; as could the development of a dually-tagged mRNA (such as the dual-labelled *rpoB*C described in Supplement 4.11).

Supplement 4

4.1 Schematic of modular system design for engineering in any target promoter and array sequence of choice, using *rpoB* as an exemplar.



4.2 Side-by-side representation of 50 nM original 8mer-2x647N-electroporated tag-strain and wildtype cells, 100th frame. Scale bar – 3 µm.



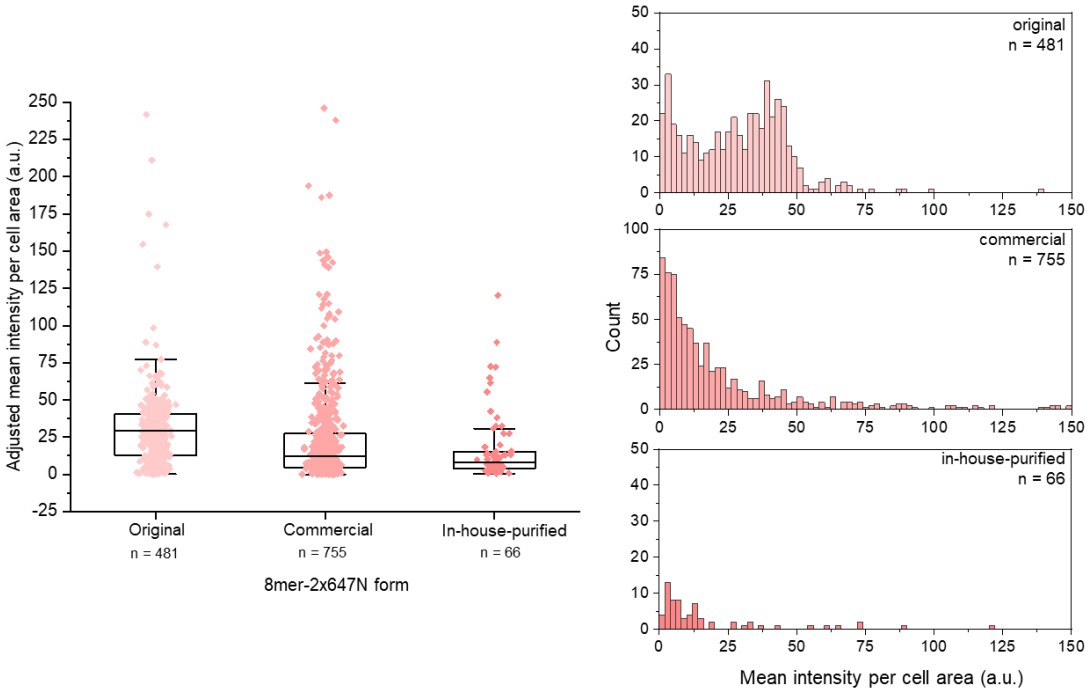
4.3 A proxy for specific signals in *rpoB*-tag cells electroporated with various 8mer-2x647N probes: populational mean intensity distributions for wildtype-corrected loaded *rpoB*-tag cells.

(Left) Mean intensities per cell area (100th frame) with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for classified loaded *rpoB*-tag cells, further to subtraction of the averaged mean intensity of loaded wildtype cells.

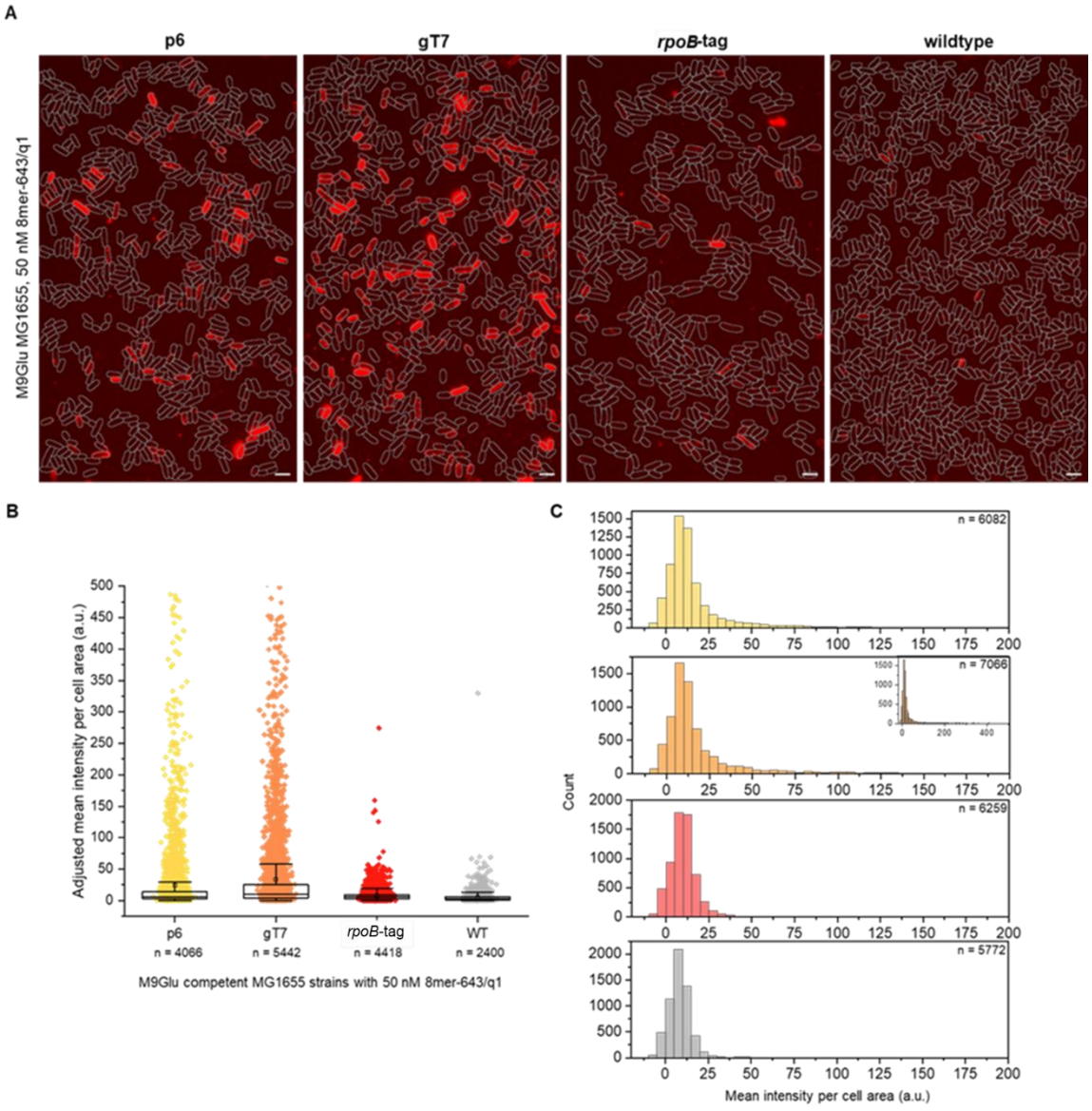
Wildtype-subtracted loaded cell medians: 79.4 units p6, 160.8 units.

(Right) Whole population histogram distributions of wildtype mean-subtracted tag cell mean intensities, adjusted by background correction. Bin size = 2.

n gives the number of cells per sample (N = 3, 2, 3 experiments for both strains respectively).



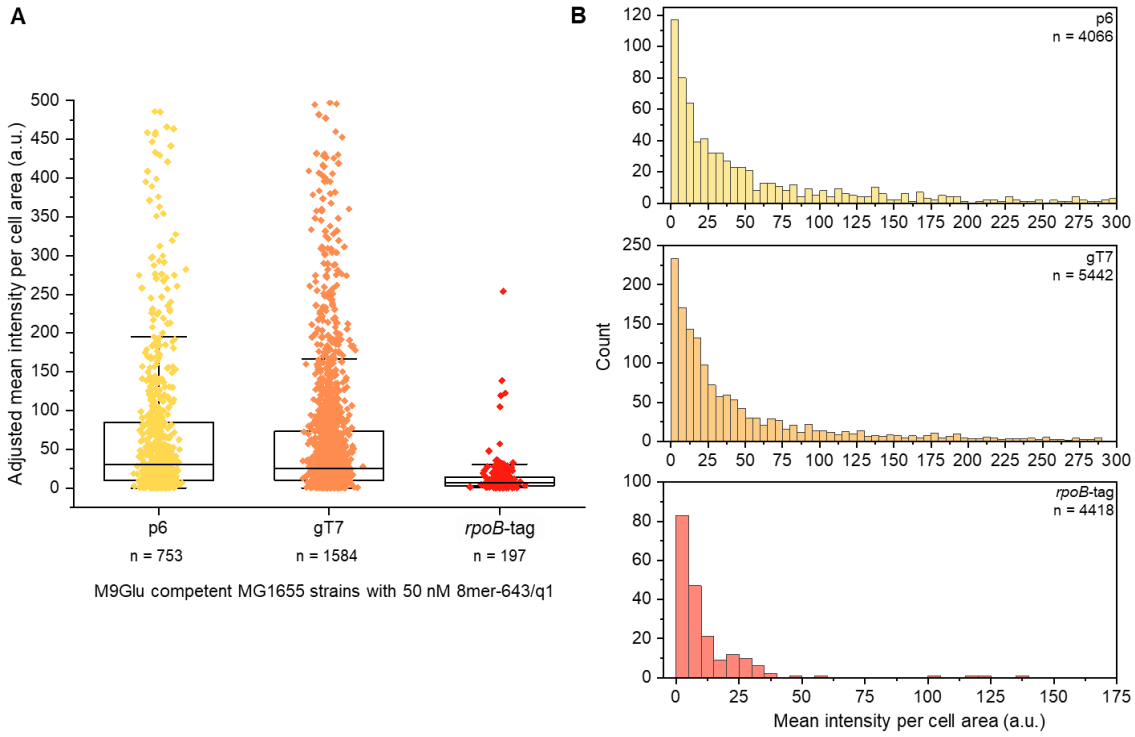
4.4 Comparison of *rpoB*-tag cell mean intensity distributions and representative FOV against p6, gT7 and wildtype results from the 100th frame, as presented in Chapter 3.



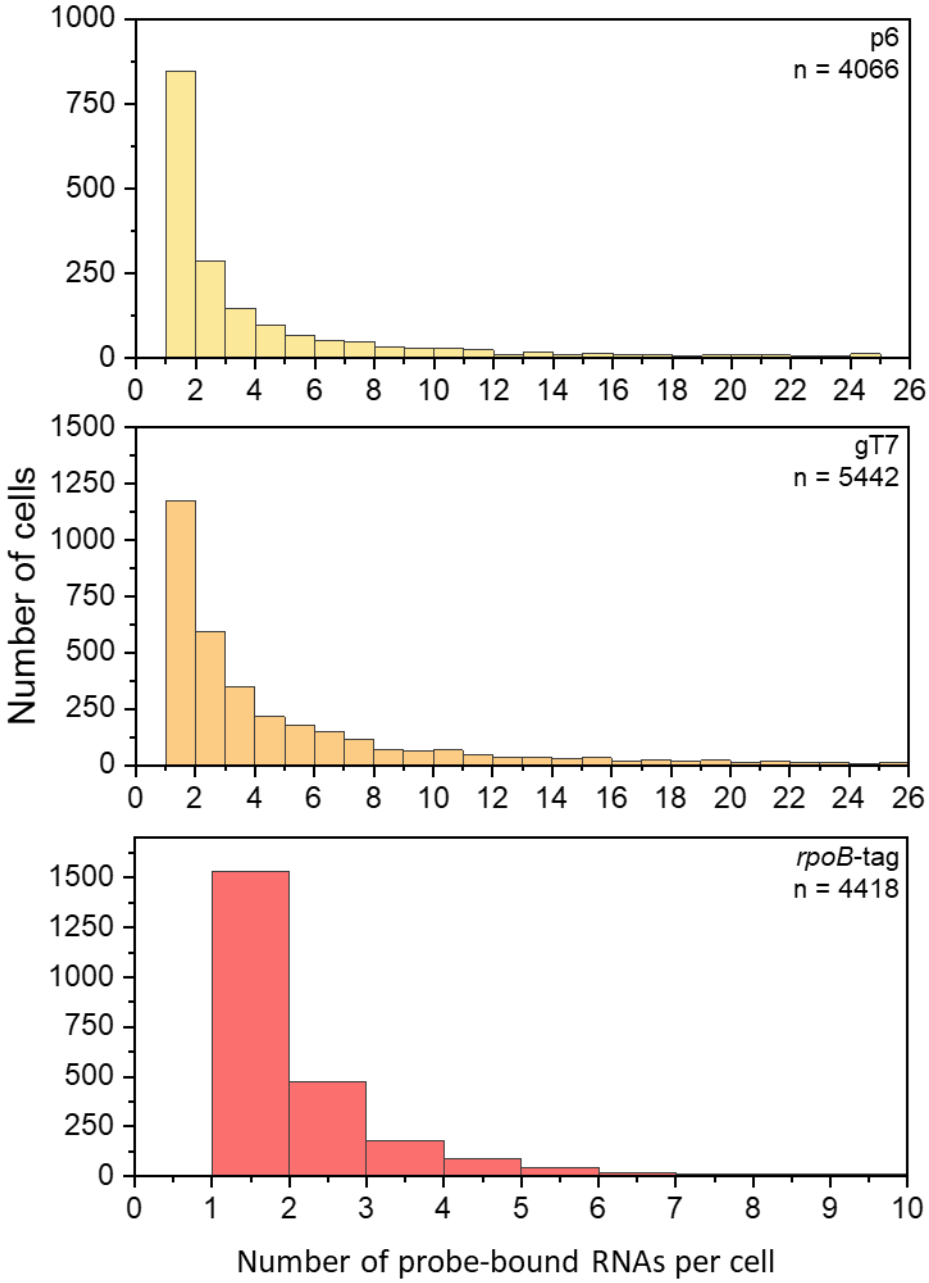
4.5 A proxy for specific signals in tag strains electroporated with 50 nM 8mer-643/q1 probes: populational mean intensity distributions for wildtype-corrected loaded p6 and gT7 cells. (Left) Mean intensities per cell area (brightest frame) with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for classified loaded p6, gT7 and *rpoB*-tag cells from Figure 4.4, further to subtraction of the averaged mean intensity of loaded wildtype cells. Wildtype-subtracted loaded cell medians: 30.3 units p6, 25.5 units gT7, 6.4 units rpoB-tag cells.

(Right) Whole population histogram distributions of wildtype mean-subtracted tag cell mean intensities, adjusted by background correction. Bin size = 5.

n gives the number of cells per sample (N = 2 experiments).

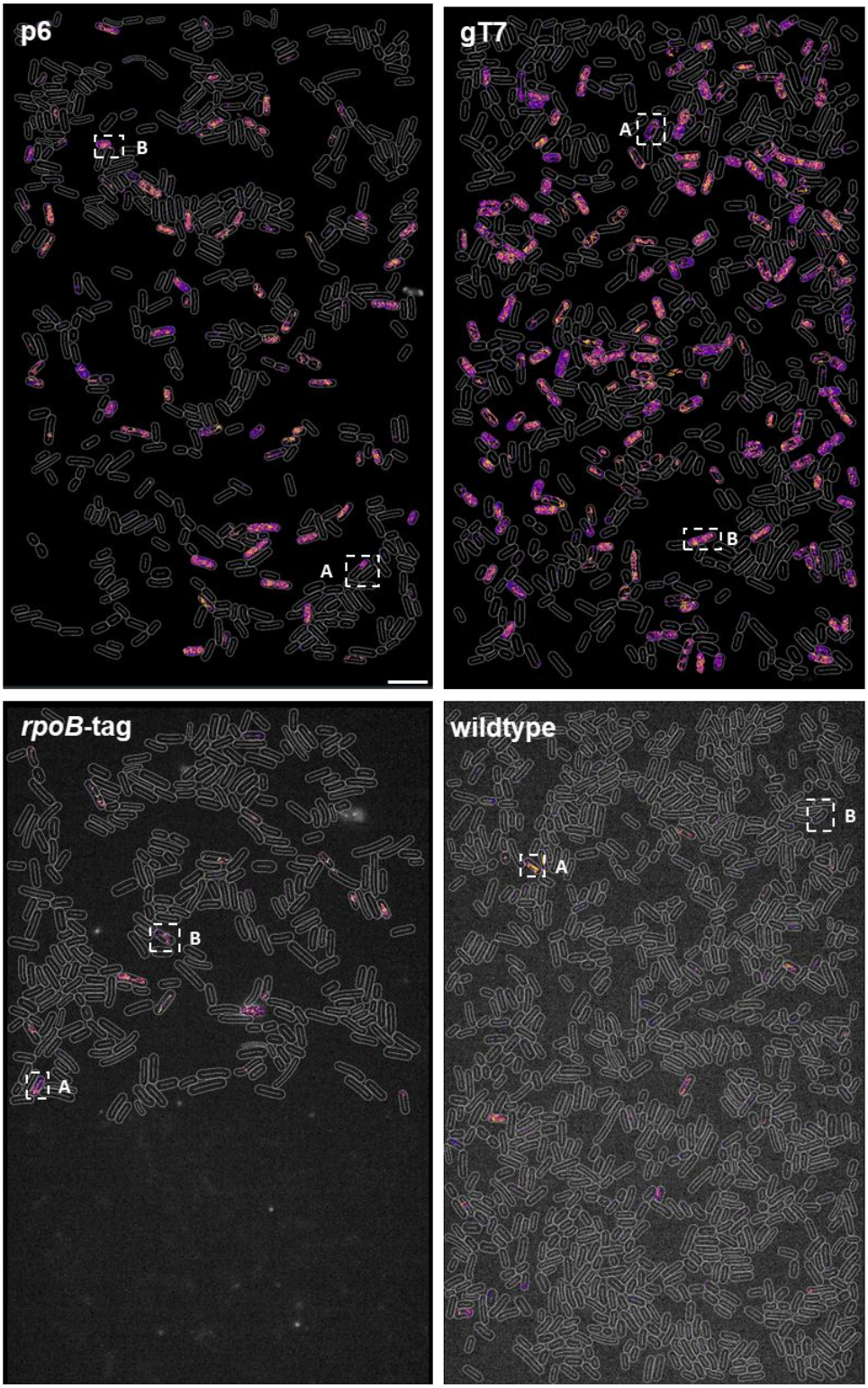


4.6 Approximated distribution of molecule counts per cell, calculated per strain with 50 nM 8mer-643/q1 by dividing the brightest frame adjusted mean intensities by the unitary intensity obtained per strain.

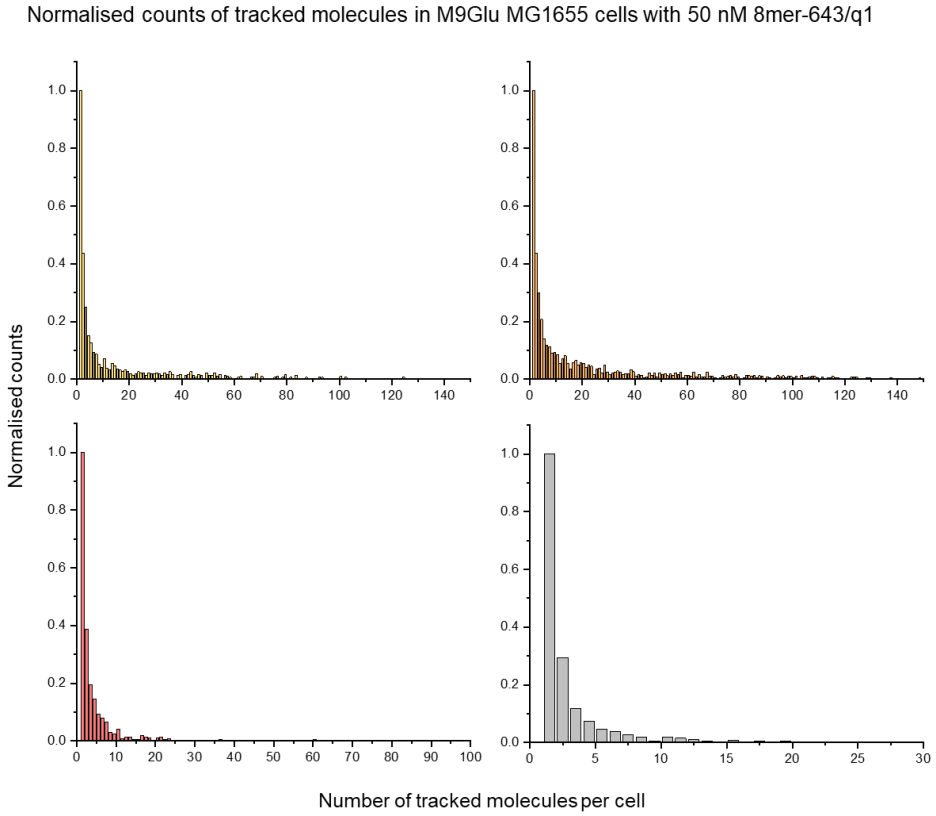


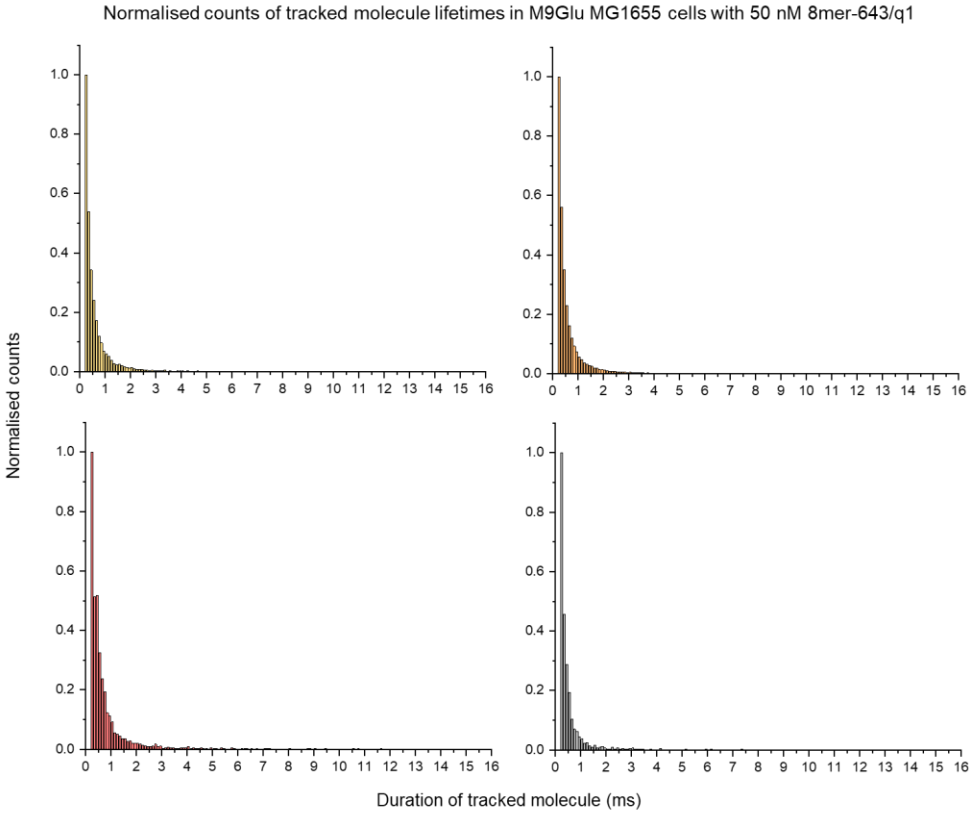
4.7 *Moltrack* processing of the representative FOVs for tag-strain and wildtype cells electroporated with 50 nM 8mer-643/q1, showing all trajectories per cells. A/B labelled boxes correspond to the exemplar cells discussed per strain; scale bar – 5 μm.

MolTrack-tracked molecules in M9Glu competent MG1655 cells with 50 nM 8mer-643/q1

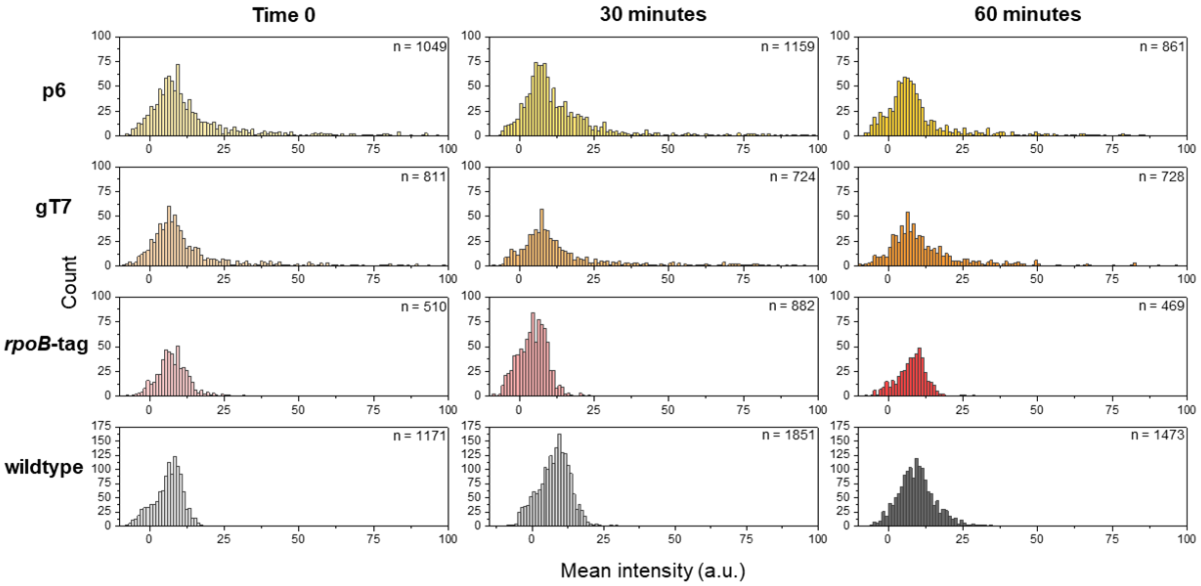


4.8 Full range counts' range for 8mer-643/q1 tracked foci counts per cell and lifetimes.

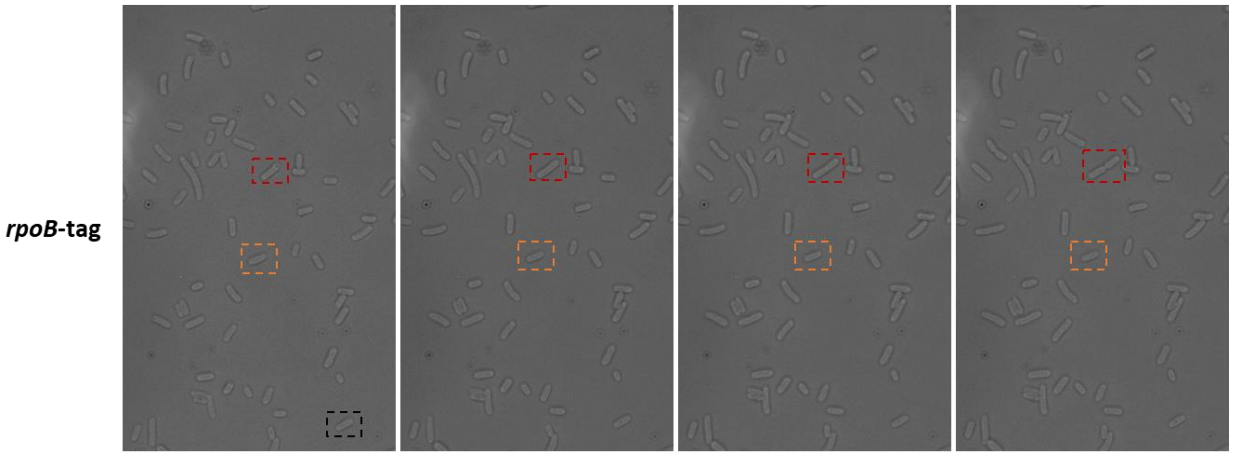




4.9 Rifampicin time course (100 µg/mL) for all strains with 8mer-643/q1; whole cell population intensity distributions from time 0 to time 60 at 30 min intervals. Bin size = 1.

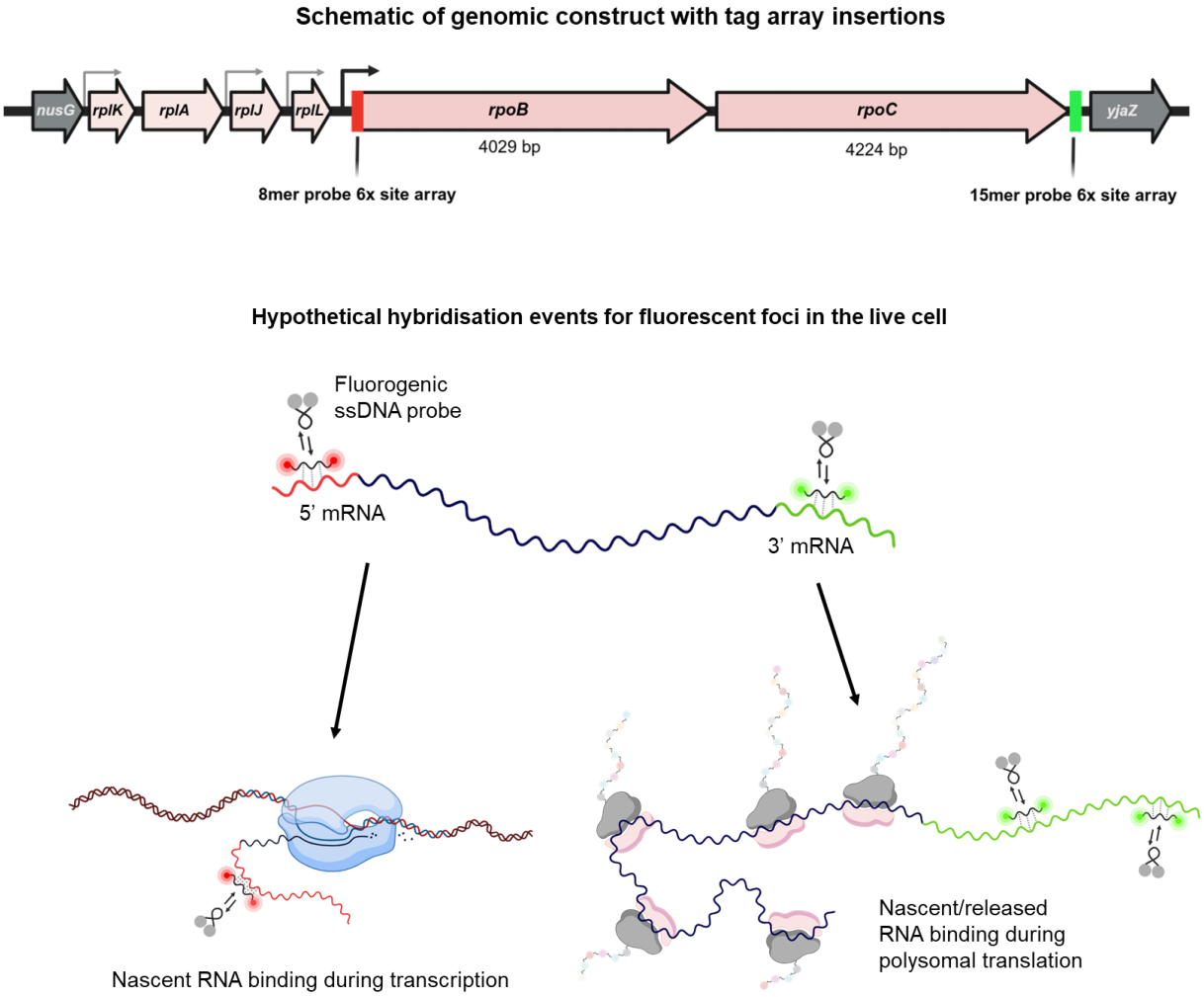


4.10 Cell viability imaging for *rpoB*-tag cells, using the same approach as in Chapter 3.



4.11 Schematic of a dual-labelled endogenous *rpoBC* target RNA.

With a 5' UTR red label upstream of *rpoB* and a C-terminus green label after *rpoC*, there are several possible hypothetical observations: just red signal during early transcription, just green if polysomal translation/degradation has occurred upstream, and potentially concurrent red and green labels if an entire rpoBC transcript is present in the cell.



Chapter 5 – Prospects of permeability: applying *in vivo* smFISH via passive internalisation

5.1 Introduction

The membranes of a cell, eukaryotic and prokaryotic alike, are sites of diverse activity: selectively permeable, the uptake and efflux of pervasive particles can be regulated by membrane lipid and protein composition, passive and active transport mechanisms, all influenced by local thermodynamics (Balantič et al., 2021). There is much value in externally modulating this permeability – drug delivery, bio-engineering and synthetic biology as examples – and so the ability to transiently form pores in membranes has been extensively methodologically characterised (electrical field, ultrasound sonication, light, chemical and mechanical manipulation; albeit primarily in eukaryotic cells)(Rems et al., 2022).

With the added complexity of unique lipopolysaccharide components, increased robustness from cell wall peptidoglycan, and the comparative small cell size of bacteria, there are not so many viable induced internalisation methods. Whilst many bacteria exhibit natural competence (Blokesch, 2016), *E. coli* cells are not considered to be naturally competent despite holding some competence homologs (Palchevskiy & Finkel, 2006; Sinha & Redfield, 2012). Nevertheless, numerous of the above internalisation methods have been adapted and substantially elucidated for prokaryotes, particularly in the realms of molecular cloning and synthetic biology.

The most established internalisation methods for bacteria are chemical induction, heat-shock and electroporation, briefly summarised in Table 5.1; additional purpose-specific internalisation methods are out of scope for discussion here. Adapted for prokaryotic application in 1970-1980s (Dower et al., 1988; Hanahan, 1983; Mandel & Higa, 1970; Taketo, 1988; Van Die et al., 1983), primarily for the transformation of plasmid DNA into prokaryotic cells, these methods all use the same premise of transient pore formation in the bacterial cell membrane to allow the uptake of external molecules, down a concentration gradient.

Table 5.1 – Brief overview of the three dominant induced internalisation methods for bacteria; all methods here would be followed by a recovery period at 37°C.

Internalisation method	**General mechanism**	**Appropriate procedures for *E. coli***
Thermal (heat-shock)	Increased thermodynamic mobility leading to the uptake of exogenous molecules, and thermal modification of membrane components increasing pore size and membrane depolarisation (Asif et al., 2017).	Competent bacteria kept at ~4°C, mixed with exogenous biomolecules, then heated at 42°C for 30-45 seconds. Heat-shock is typically applied to (chemically) competent cells.
Chemical	Charge stabilization of lipopolysaccharides, the membrane and exogenous nucleic acids (Asif et al., 2017).	Chemical competency of cells by washing with Ca^{2+}/Mg^{2+} (mM) in solution; can be followed by a heat-shock. Chemical competency can also be achieved via the use of dimethyl sulfoxide (DMSO) and polyethylene glycol (PEG)(transformation/storage solution, TSS)(Chung & Miller, 1993).
Electrical (electroporation)	Anisotropic pulsing of a high-intensity electrical field strength causes an induced transmembrane voltage, leading to local modification of the membrane (Balantič et al., 2021).	Cells made competent by ice-cold washing and snap-freezing; application of a 0.5-10 kV/cm pulsed electric field.

The charged and variably diffusive state of the cell membrane, determined predominantly by phospholipid and protein content, enables the membrane to act as a capacitor. The efficiency of electroporation- and chemically-induced internalisation thus rely on cellular permeabilization through unbalancing the electrostatics of the cell membrane to form transient pores (Balantič et al., 2021).

Although core tenets of molecular biology, the application of these internalisation methods for single molecule experimentation is few-and-far-between. Only electroporation has been established as an efficient internalisation method for incorporating organic dye-labelled biomolecules into live bacterial cells for single molecule studies (Aigrain et al., 2015; Crawford et al., 2013; Sustarsic et al., 2014): probing the *in vivo* behaviour and diffusive properties of internalised DNAs (Crawford et al., 2013; Plochowietz et al., 2014); tRNAs (Plochowietz et al., 2017; Volkov et al., 2019); and chemotaxis regulator protein CheY (Di Paolo et al., 2016).

Whilst electroporation of fluorescent biomolecules has been a significant step forward in developing functional understanding of key cellular components in the single cell, closer to the physiological state than before, the impact of an electric charge across a membrane should not be underestimated. Applying reversible electroporation[9] to bacterial cells has been shown to induce lipid peroxidation of membrane lipid components (Benov et al., 1994), leading to significant physiological alterations and possible cell death. Electroporated eukaryotic cells have been shown to suffer membrane protein denaturation (Tsong, 1990), cytoskeletal network and telomere disruption (Stacey et al., 2011). Although generally lacking evidence and understanding of molecular mechanisms (Balantič et al., 2021), there are almost certainly similar structural effects in bacteria. Acknowledging the fundamental importance of an integral membrane in regulation of transcription, translation and degradation processes in the live cell, a closer physiological appraisal of vital activity *in vivo* would be attained by reducing any perturbation to the membranes.

In Chapter 3.6, short ssDNAs were shown to be capable of traversing the bacterial cell membranes and thus being internalised. Building on these technical observations, the potential for the introduction of fluorescently-labelled biomolecules into live cells without induced transformation could be an exciting step closer towards a more "natural" physiological cell state post-permeabilisation. Furthermore, in considering the extensive engineering required for the proposed methodology of *in vivo* smFISH, using faithfully-hybridising, fluorescently labelled ssDNA to specifically target endogenous RNAs in the live cell would be a highly attractive complementary approach.

As observed in previous chapters, there are many challenges associated with targeting mRNAs; to test the possibility of internalising ssDNA probes to target physiological behaviours

[9] Irreversible electroporation – characterised by an extended electrical field pulses, leading to permanent structural damage of the cell membrane and typically cell death (Ho & Mittal, 1996); applied in controlled contexts for eukaryotic cell ablation, for example.

in the live cell, a more robust target was required. Prokaryotic rRNA genes represent a particular RNA group of interest due to their essentiality in ribosome formation in the live cell. Since the small 30S subunit of the prokaryotic 70S ribosome comprises 16S rRNA, complementary hybridisation of the universal eubacterial 338 (EUB338) probe is an effective marker of the ribosome and its 30S subunits (R. I. Amann, 1990)(Figure 5.1).

The considerable conservation of the 16S rRNA core structures in particular, and high intracellular rRNA copy numbers, have meant that rRNA-targeting fluorescent probes have long been used for bacterial phylogenetic typing of bacterial populations (Amann et al., 1990; Pernthaler *et al.*, 2001) using fixed FISH. More recently, 16S rRNAs have been applied in probing fixed cell ribosomal distributions, and live bacterial FISH and cell sorting (Batani et al., 2019).

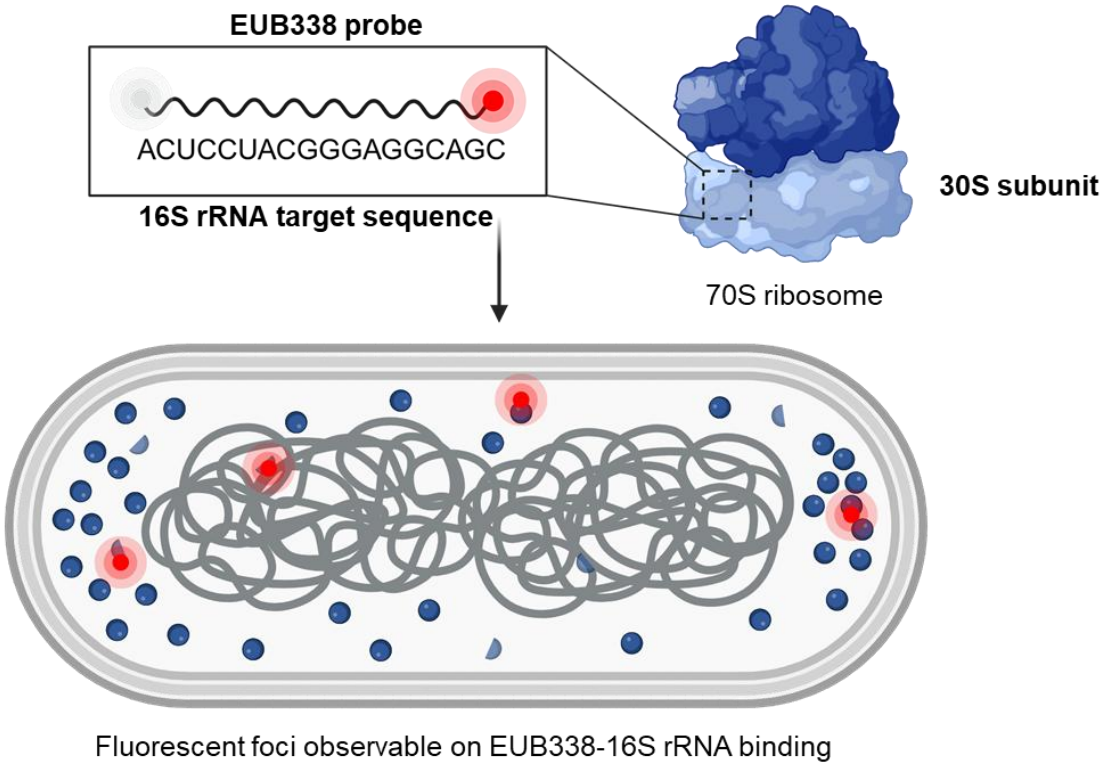


Figure 5.1 – Schematic of EUB338 probe binding, and potential visuals when applied in live cells.
The EUB338 probe binds specifically to an 18-rnt sequence in the essential 16S rRNA, which is the platform for forming the 30S free ribosomal unit. Fluorogenically labelling the 30S subunit and 70S whole ribosome with smFISH probes could give single fluorescent foci observable in the live cell.

The EUB338 probe, targeting the 16S rRNA of most bacteria from positions 338 to 355 as defined by Amann et al., 1990 (see RNA sequence in Figure 5.1), has been consistently used in fixed FISH and fluorescence associated cell sorting (FACS). A subsequent negative control, the nonsense probe, NonEUB338, has confirmed the EUB338 probe as reliable for 16S rRNA hybridisation (Wallner et al., 1993). Whilst the exact binding loci and kinetics of the EUB338 probe remain largely uncharacterised, doubly-labelled EUB338 has been shown to have increased binding tendencies and FISH signal over singly-labelled alternatives (Stoecker et al., 2010), suggesting that combinatorial organic dye labelling of the EUB338 probe may be suitable for *in vivo*, single molecule use.

With such an abundance of ribosomes in live cells (upwards of 20,000 at any one time (Bakshi et al., 2012; Mohapatra & Weisshaar, 2018)), accurate understandings of populational mobilities and spatial patterns are highly insightful. The capability to elucidate the real-time spatiotemporal behaviours of individual ribosomes and RNAPs via endogenous labelling and PALM has enabled greater understanding of the diffusive properties of the ribosome and its subunits, spatial localisation and co-transcriptional interactions – towards efficiently and bidirectionally regulating single cell expression profiles (Bakshi et al., 2012, 2013; Endesfelder et al., 2013). Bakshi et al. estimated an ~55,000 circulating ribosomes in an average fast-growing *E. coli*, where ~75% of 70S ribosomes are polarly localised and 10-15% involved in co-transcription-translation; supporting the circulation model of ribosomes and RNAPs in co-regulated transcription and translation. More recent work by Mohapatra & Weisshaar (Mohapatra & Weisshaar, 2018) indicated ~20% of 30S subunits freely diffuse through the nucleoid and wider cell.

Following the establishment of EUB338 as a suitable FISH probe in fixed *E. coli* cells to observe ribosomal distributions for rapid antibiotic susceptibility testing (Chatzimichail et al., 2024; Farrar et al., 2024), recent work in the Kapanidis group has shown that EUB338 probes can robustly bind to ribosomes in the live cell to obtain single molecule ribosomal behaviours (Mirjam Kümmerlin, 2024). In particular, an EUB338 probe labelled with an ATTO643 dye and Black Hole Quencher 1 has been shown to bind specifically and exhibit fluorogenicity such that fluorescent background suppression gives a much higher signal-to-noise ratio and dynamic range (Mirjam Kümmerlin, 2024). The EUB388-ATTO643/BHQ1 (EUB-643/q1) probe is thus a suitable choice for continued *in vivo*, single molecule ribosomal spatiotemporal characterisation.

Chapter summary

In this Chapter, a permutation of the established protocol for live cell biomolecule internalisation as described in Chapter 3 was applied, where fluorogenic probes were passively internalised: to assess ribosomal diffusion behaviours in a single molecule regime, with reduced membrane perturbation. A comparison of classical internalisation methods with passive internalisation showed negligible differences in internalisation of EUB338 1xATTO647N (EUB-647N) probes with and without recovery of the cells. Using the quenched EUB-643/q1 probe, experiments by passive internalisation were then conducted in competent wildtype *E. coli* to explore ribosomal behavioural characteristics. The results showed similar mobility distributions and spatiotemporal behaviours to key literature, indicating that endogenous RNA targeting with fluorogenic ssDNA probes in M9Glu competent cells via passive internalisation is tractable.

5.2 Experimental approach

Experiments were performed using the same approach as in Chapter 3.6.
Briefly: EUB338 probes were added to thawed competent wildtype *E. coli* and allowed to incubate for 10 min with the cells; the probe-cell mixtures were then iteratively washed and finally resuspended for superresolution imaging on the NIM in HILO, using 10% (~10 mW) 638 nm laser power. All experiments were conducted at room temperature (referred to as 19°C).

Ensemble measurements comparing the internalisation efficiency of different internalisation methods was assessed via the uptake of the EUB-647N, in line with the Chapter 3.6 technical experiments, to avoid any further variables in uptake being introduced by a second label on the 18mer ssDNA sequence. Mean intensity per cell area of the brightest frame was deemed a suitable ensemble comparator to show cell fluorescence distributions as a proxy for fluorescent probe internalisation. “Loaded” subpopulations were categorized as any brightness higher than the averaged mean intensity per cell area of the brightest frame for empty cells, corrected for background.

For EUB-643/q1 experiments, probe photostability, brightness and unitary intensity were determined as documented in Chapter 4. Averaged time traces were generated from the averaged mean intensity per cell per frame, the unitary intensity via step-height fittings of single loaded cell traces using *AutoStepfinder* (Loeff et al., 2021), and an approximation for populational spread of molecule counts per cell by dividing brightest cell intensities by the unitary intensity.

The apparent diffusion of ribosome-related populations was found by the localisation and tracking of independent fluorescent foci in *Moltrack*. Extracted particle metrics were used to calculate the apparent diffusion coefficients of tracked particles, observe the distributions of tracked molecule counts per cell and durations, and generate unit cell renders of molecule spatial localisations across the cell. At a single-cell level, individual foci were observed over time, enabling extraction of focus track step sizes and distance from the membrane to give single molecule insight into the behaviours of ribosomal units.

5.3 Internalisation efficiency: comparing existing methods

To compare the tractability of passive internalisation as an alternative to standardized stimulated internalisation, the relative loading efficiency of passive versus established inductive internalisation methods was evaluated via ensemble intensity measurements.

Comparing internalisation without recovery in competent cells

Figure 5.2 shows the comparison of passive internalisation, electroporation at 1.4 kV/cm, DMSO-mediated chemical internalisation (0.1% DMSO) and heat-shock (~45 s at 43°C) on the loading of EUB-647N probes into competent wildtype *E. coli* cells. Efficient internalisation of the EUB-647N probe was shown across the four different internalisation methods in both M9Glu and LB competent wildtype cells (Figure 5.2A, A-B), and all samples presented a higher intracellular fluorescence compared to control cells with 5 µL upH_2O instead of probe solution (Supplement 5.1A). Per internalisation method, the loading efficiencies of M9Glu cells were ~82%, ~66%, ~77% and ~66%, and ~71%, ~70%, ~89%, ~74% for LB cells, for probes internalised passively, via DMSO treatment, via electroporation and via heat-shock respectively, without recovery.

Visually, the passive internalisation and electroporation cell populations look most similar in both M9Glu and LB cells, with the highest spread of signal across cells, albeit heterogeneous (Figure 5.2A). Interestingly, the “halo effect” of saturatingly-bright signal in cells, as seen with the 8mer-2x647N probes in Chapter 3, was seen across the DMSO and heat shock samples more than the passive and electroporation samples. Any cells that appear fluorescent but were suspected to have ruptured membranes or immobilised in a different plane on the agarose pad were not included in analysis, and thus are seen as fluorescence without cell outlines overlaid.

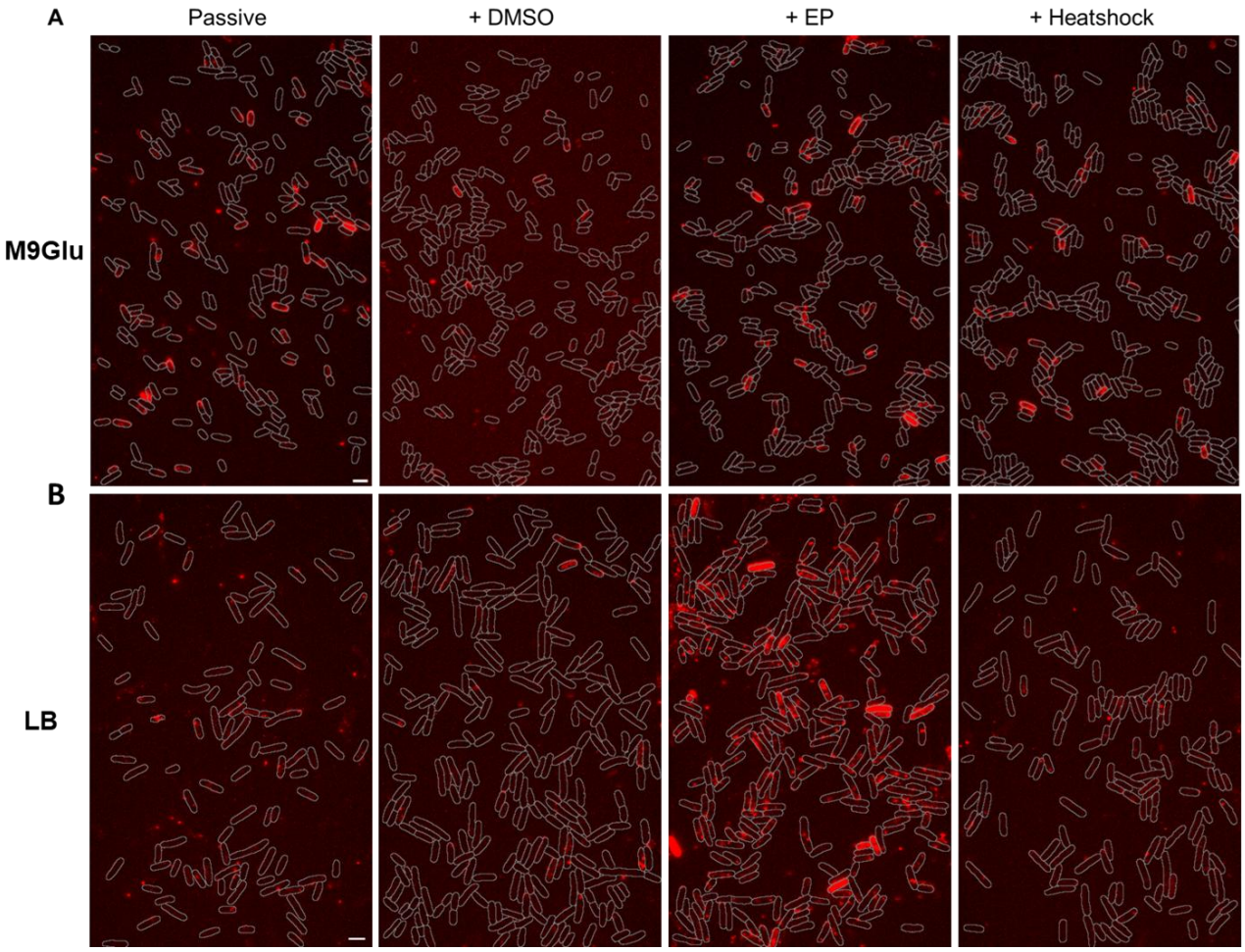


Figure 5.2A – Comparison of live-cell probe internalisation methods without recovery. Maximum intensity projections of frames 1-2000 for representative FOVs of competent M9Glu (A) and LB-grown (B) MG1655 *E. coli* with 100nM EUB-647N probes, internalised passively, chemically-stimulated via 0.1% DMSO, via 1.4 kV/cm electroporation (EP) or via heat shock at 42°C.
Scale bar – 3 µm, overlaid cell segmentation outlines.

Analysing the brightest frame background-corrected mean intensities of M9Glu cells classified as loaded with 100 nM EUB-647N, passive internalisation and electroporation gave the highest median mean intensities per cell area (4.7 units passive, 4.3 units electroporation), with negligible difference between the DMSO and heat-shock (2.6 units DMSO, 3.0 units heat-shock)(Figure 5.2B,A).

Similar observations were made for the LB cells, with median mean intensities 2.83 units for passive, 7.92 units for electroporation, 3.15 units DMSO and 3.92 units heat-shock for cells with 100 nM EUB-647N.

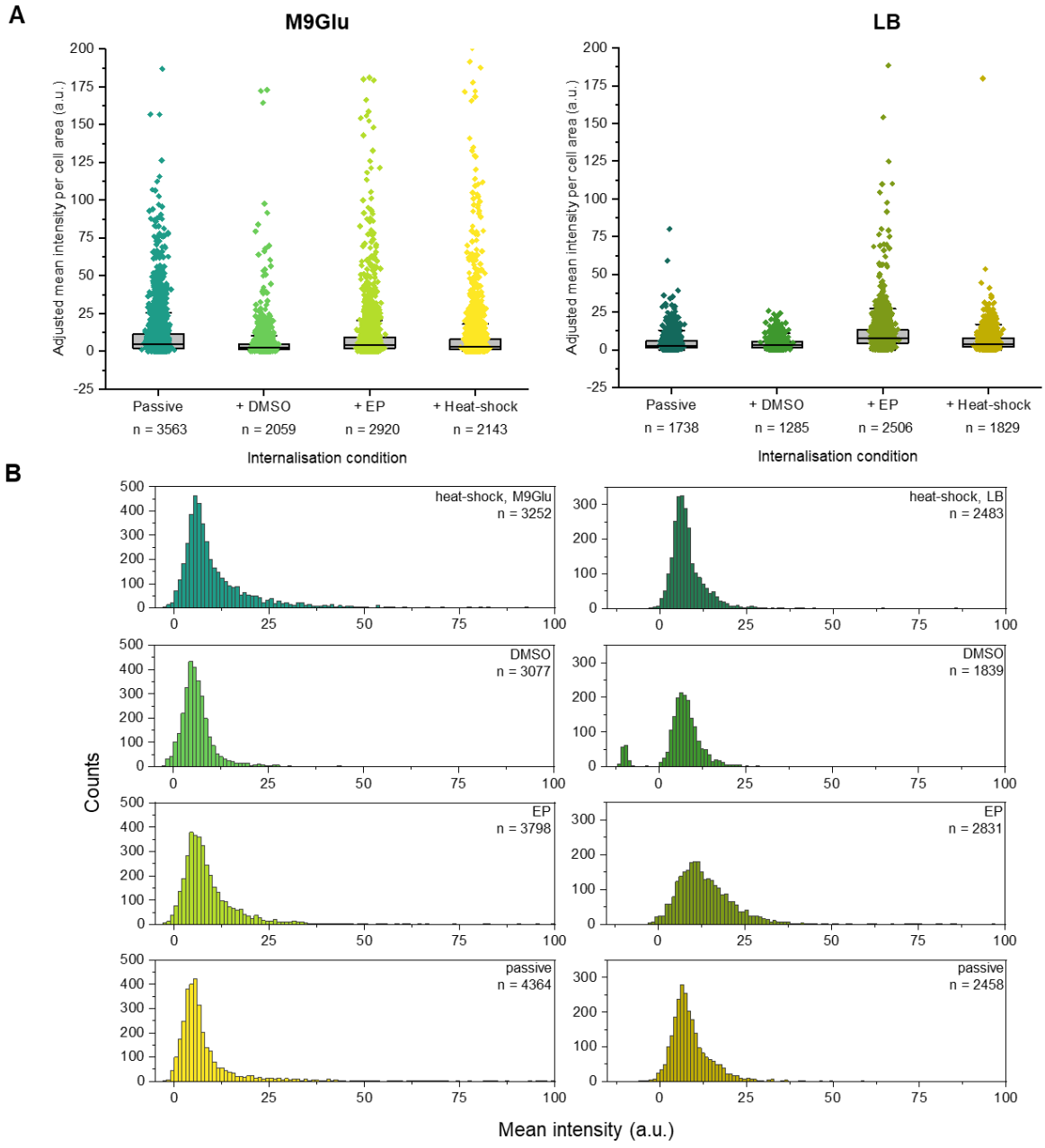


Figure 5.2B – Comparison of live-cell probe internalisation methods without recovery.
(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 100 nM EUB-647N-loaded M9Glu and LB cells via passive, DMSO, electroporation or heat shock-mediated internalisation. Cell classification as 'loaded' defined as mean intensity above the mean intensity value averaged across empty cells (incubated with upH_2O only).
M9Glu medians: 4.67 units passive, 4.28 units electroporation, 2.61 units DMSO, 3.0 units heat-shock.
LB medians : 2.83 units passive, 7.92 units electroporation, 3.15 units DMSO, 3.92 units.
(B) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 1.
All mean intensities shown are background corrected, and n is the number of cells per condition from 5 FOVs across N=2 experiments. Any negative values observed in (B) are due to global background subtraction from originally empty cells.

Interestingly, the whole cell distributions do not seem to show much difference in M9Glu cells, but a much wider range of intensity characteristics are seen in LB cells. As observed in previous chapters, the population of loaded cells remains highly heterogenous across all internalisation methods, and indeed many cells are not loaded at all. Cells with passively internalised probes, though, appeared to have the most similar internalisation distribution across M9Glu and LB cells, suggesting that signal spread was more homogenous across cells showing fluorescence. Nonethless, since heterogeneity is observed across all four methods consistently, and has been documented widely for biomolecule internalisation, it is likely that this heterogeneity represents single cell differences in membrane composition, robustness and structure, despite clonal origins.

Comparing internalisation with recovery

It is possible that the absence of a recovery period post-stimulus in the above may have affected the cell viability and hence artificially increased loading capacity. The equivalent experiment, including a recovery step post-stimulation, was performed to compare internalisation rates in a manner more customary with perturbational internalisation methods such as electroporation for M9Glu and LB competent wildtype cells (Figure 5.3A; supplement 5.1B for empty cell controls).

Similarly to observations in Figure 5.2, all four internalisation methods visually presented a wide spread of loaded cells, with similar overall loading populations: the loading efficiencies of cells were ~79%, ~64%, ~78% and ~76% in M9Glu cells and ~66%, ~31%, ~85%, ~67% for LB cells, for probes internalised passively, via DMSO treatment, via electroporation and via heat-shock respectively, with recovery.

The purpose of recovery post-perturbation is to promote strengthening and reduce poration of the cell membrane, as well as re-acclimatizing the cell over the course of a single division time. In theory, this would increase cell fitness and viability, particularly for perturbational internalisation methods (Plochowietz et al., 2017; Sustarsic et al., 2014) – although notably, the observed viability of cells undergoing passive internalisation (Chapter 3.6) suggests that recovery is not necessary for unperturbed competent cells with signal.

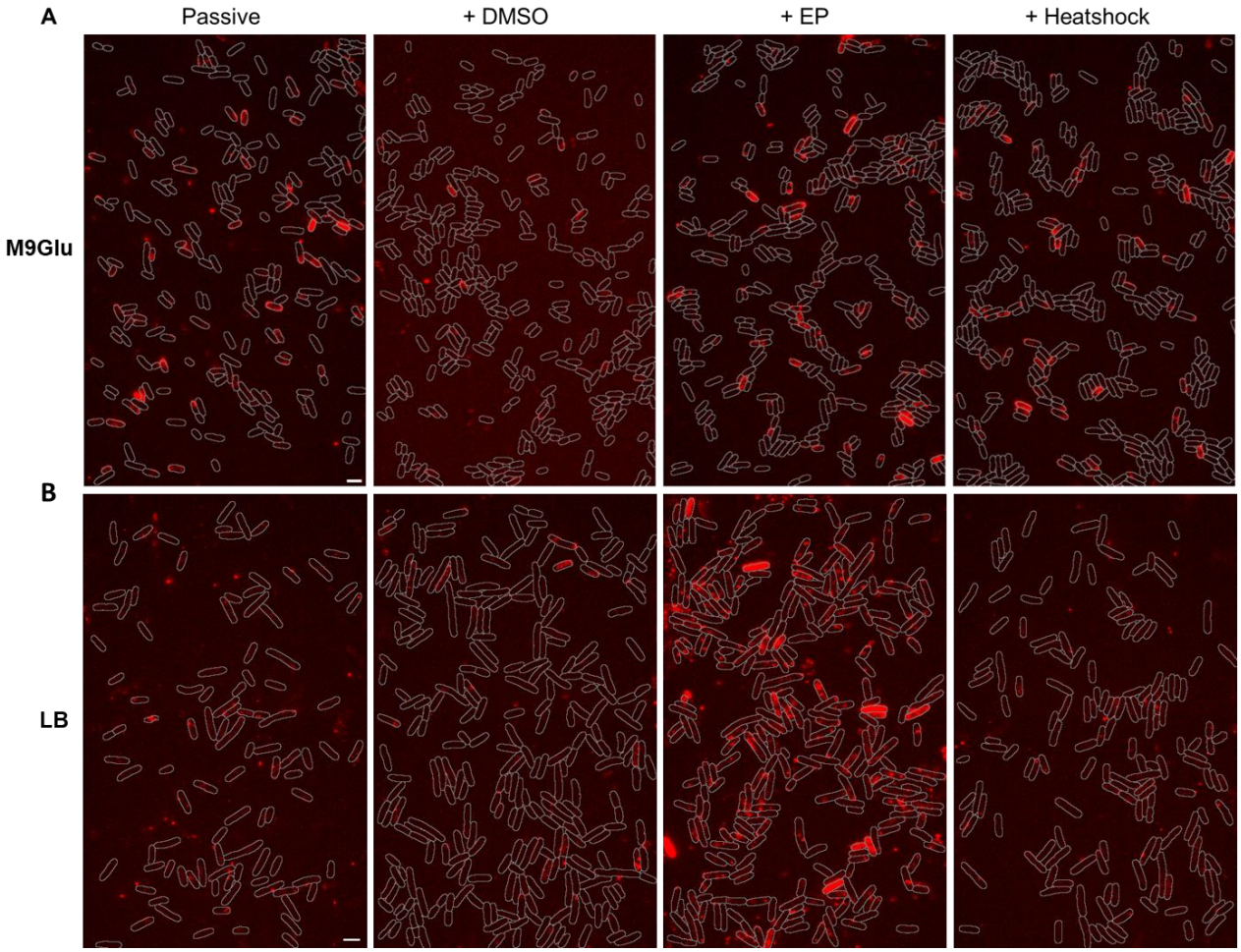


Figure 5.3A – Comparison of live-cell probe internalisation methods with recovery.
Maximum intensity projections of frames 1-2000 for representative FOVs of competent M9Glu (A) and LB-grown (B) MG1655 *E. coli* with 100nM EUB-647N probes, internalised passively, chemically-stimulated via 0.1% DMSO, via 1.4 kV/cm electroporation (EP) or via heat shock at 42°C, post 20-min recovery in SOC medium.
Scale bar – 3 µm, overlaid cell segmentation outlines.

The median mean intensities of M9Glu and LB cells show that the recovery period appears to have negligible impact on the internalisation efficiency of 100 nM EUB-647N (M9Glu: 4.62 units passive, 3.63 units electroporation, 1.70 units DMSO, 3.21 units heat-shock; LB: 3.16 units passive, 6.54 units electroporation, 1.57 units DMSO, 2.76 units heat-shock)(Figure 5.3B). A populational distribution pattern similar to that of the unrecovered samples is observed, with wider range in intensities with LB cells; this again suggests that the additional recovery step has negligible impact on internalisation overall.

A non-parametric, Kruskal-Wallis test for ranked medians performed on both unrecovered and recovered, loaded cells showed no significant difference between the four internalisation methods. Electroporation is undeniably more characterised than any other internalisation

method for single molecule studies in fixed and live cells, and shows the highest internalisation efficiency overall.

Notably, the lack of substantial differences between passive and perturbational stimuli for labelled ssDNA uptake shows that applying no stimulus is a viable option, and may correspond to less disruption to the homeostatic cell state, even if fewer cells are loaded compared to other methods.

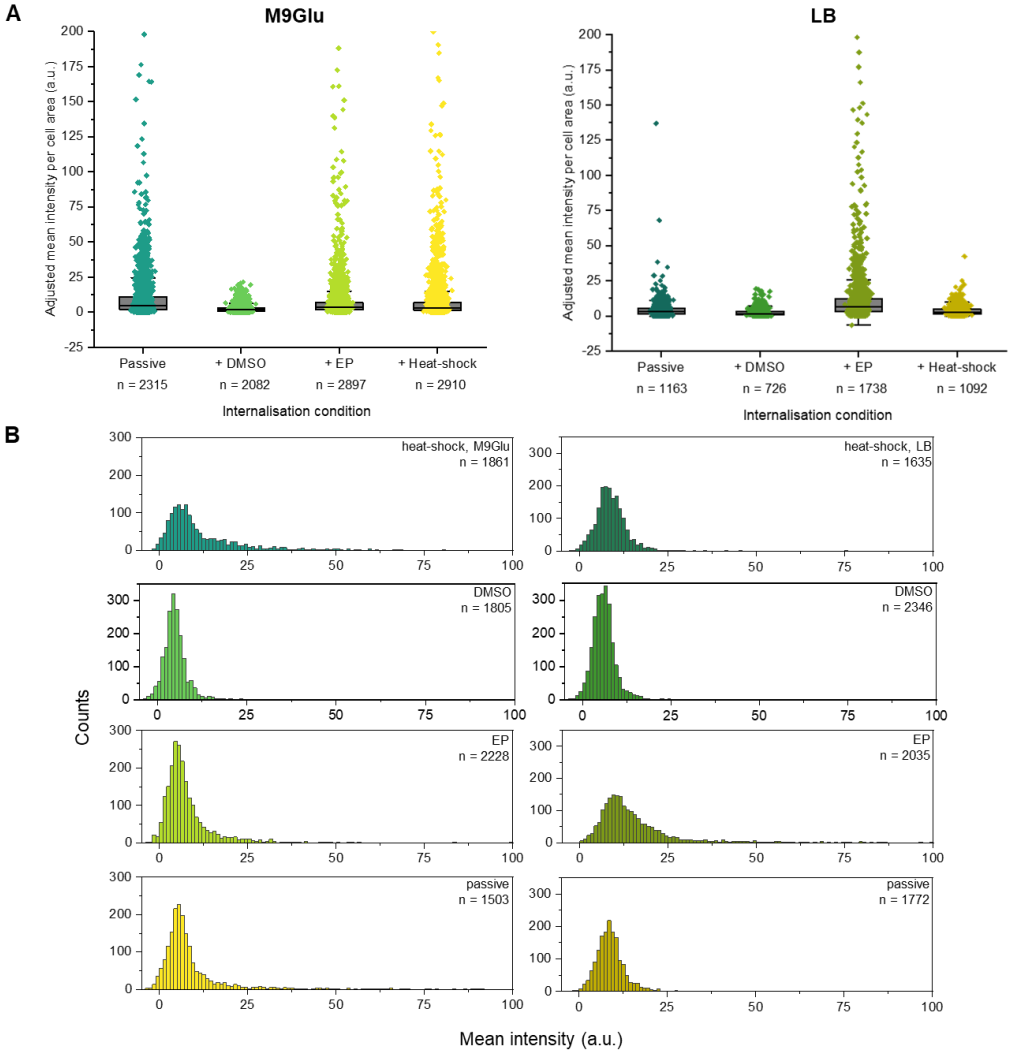


Figure 5.3B – Comparison of live-cell probe internalisation methods with recovery.
(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 100 nM EUB-647N-loaded M9Glu and LB cells via passive, DMSO, electroporation or heat shock-mediated internalisation. Cell classification as 'loaded' defined as mean intensity above the mean intensity value averaged across empty cells (incubated with upH_2O only).
Recovered M9Glu medians: 4.62 units passive, 3.63 units electroporation, 1.70 units DMSO, 3.21 units heat-shock Recovered LB medians: 3.16 units passive, 6.54 units electroporation, 1.57 units DMSO, 2.76 units heat-shock
(B) Histogram distributions of the whole cell populations (where unloaded classify with sub-zero values); bin size = 1.
All mean intensities shown are background corrected, and n is the number of cells per condition from 5 FOVs across N=2 experiments.

5.4 EUB338-ATTO643/BHQ1 as a suitable probe for targeting individual ribosomes

EUB-647N was an optimal probe for the direct comparison of changing experimental variables on the propensity of *E. coli* cells to passively internalise labelled ssDNA, due to its ability to present as bright spots – specifically bound or otherwise – in the cell. Taking the same premise as in previous chapters – whereby an ideal probe fluoresces only on complementary hybridisation to its target – the labelling of EUB338 with ATTO643 and BHQ1 have made for a reliable quenched probe, demonstrated to show bright fluorescence on unquenching with rRNA target hybridisation by preliminary efforts in the Kapanidis group (Mirjam Kümmerlin, 2024).

Whilst the impact of a second label on the passive internalisation efficiency is unknown, the limited effect of ssDNA length up to 75 nts long on passive internalisation (Chapter 3.6) suggested that passive uptake of EUB-643/q1 would be possible. The hydrophilicity of ATTO643 and hence reduced propensity for membrane sticking, as discussed in Chapter 4, would suggest EUB-643/q1 as a desirable probe for avoiding off-target signal. In any case, the use-case for EUB-643/q1 was to obtain clear, localisable bright foci corresponding to single ribosomes in live cells, once passively internalised – high internalisation rates were therefore not the goal, rather a sufficient label concentration for tracking single ribosomes.

5.4.1 Comparing the photostability of EUB-647N and EUB-643/q1 probes

To characterise EUB-643/q1 photostability comparative to EUB-647N for single molecule fluorescence *in vivo*, temporal projections of representative FOVs and the average and single cell photobleaching time traces decay of specific signal fluorescence were measured (Figure 5.4).

Visually, both probes show high fluorescence intensities heterogeneously spread across the cell population in representative intensity projections; the EUB-647N FOV shows a higher proportion of cells with signal than EUB-643/q1, as would be expected with successful quenching of ATTO643 by BHQ1.

Comparing the temporal projections of these two FOVs (Figure 5.4B), the majority of the EUB-647N visually loaded cells exhibit signal early in acquisition, whereas loaded EUB-643/q1 cells give signal across the whole acquisition period, even up to 100 seconds. From the average photobleaching decay traces for EUB-647N and EUB-643/q1 respectively (Figure 5.4C), EUB-643/q1 probes exhibit a brightest intensity of ~50% of the initial brightness of EUB-647N signal.

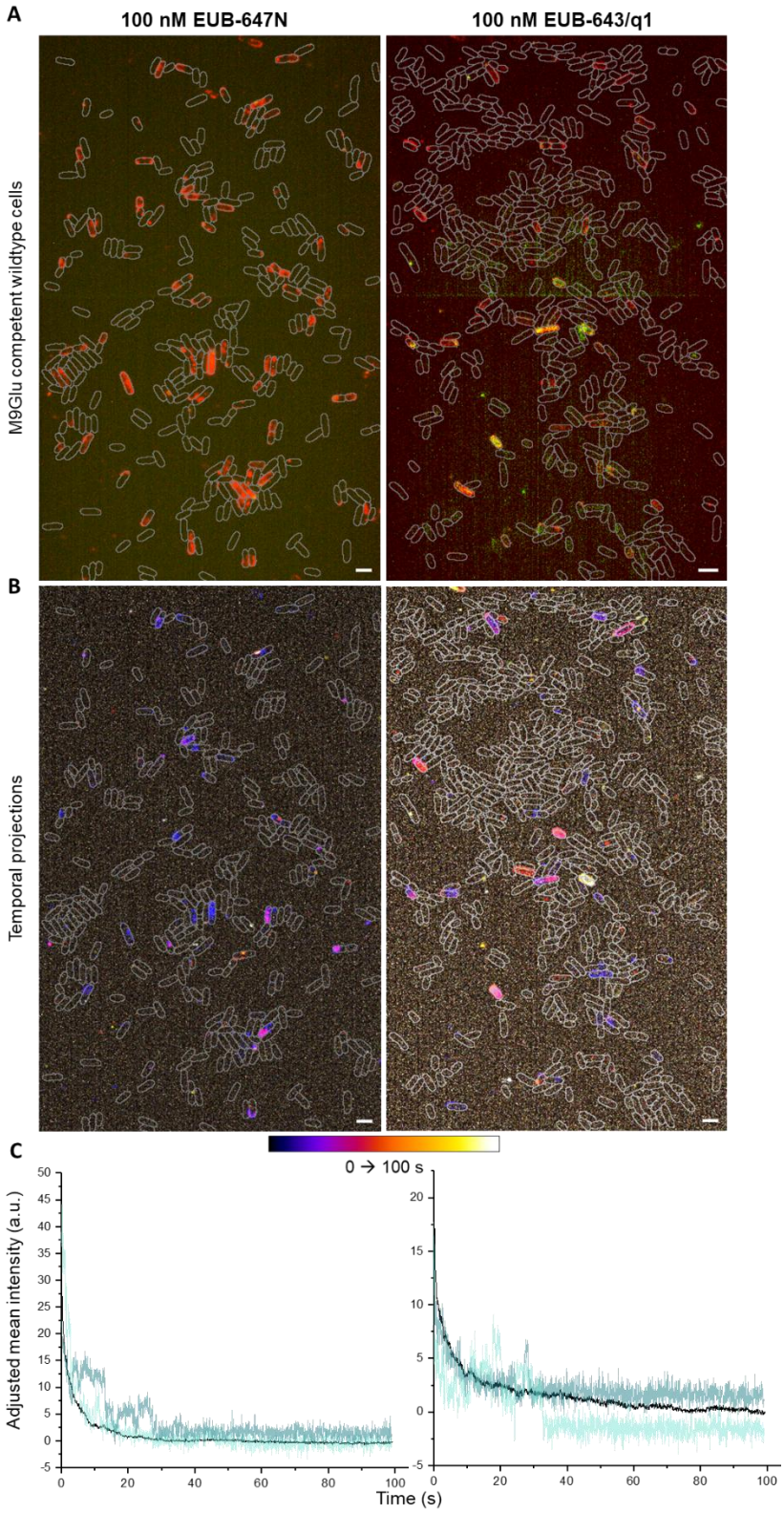


Figure 5.4 – Visualising the photostability of *E. coli* M9Glu competent cells with 100 nM passively internalised EUB-647N and EUB-643/q1 probes.
(A) Representative composite FOVs of M9Glu competent MG1655 *E. coli* with passively internalised 100nM EUB-647N and 100 nM EUB-643/q1 probes; where composites are overlays of maximum intensity projections (red) and summed intensity projections (green).
(B) Temporal projections of 1-2000 frames of the above representative FOVs. The colour bar denotes time, where darkest is time 0 (50 ms exposure, 100 s acquisition).
(C) Corresponding averaged background-corrected mean intensity per frame time traces for the representative FOVs (black), overlaid with two exemplar bleaching time traces for individual visually-loaded cells (light blue).
Scale bars – 3 µm, overlaid cell segmentation outlines.

Considering the approximate half-life of ~2.5 seconds and the temporal projection contrast, it appears that the quenching of the EUB-643/q1 probes is effective and allows for stable signal appearance over the course of longer acquisitions, providing there is a sufficient internalised probe pool.

5.4.2 Comparing the unitary intensities of EUB-647N and EUB-643/q1 probes

Attributing the step-changes to the bleaching and appearance of fluorescent foci, as in Chapter 4, *AutoStepfinder* step detection was applied to single cell time traces to approximate the unitary intensity for a gauge on probe loading in cells. Figure 5.5A shows the process of an exemplar cell, bright with multiple foci at the start of acquisition. Undergoing bleaching over time until no signal can be seen, individual foci can be distinguished – the corresponding cell intensity trace over time can be correlated with these fluorescent foci, and the intensity of one focus approximated (here, to ~5 a.u.).

Of the 171 and 179 time traces fit for cells with passively-internalised EUB-647N and EUB-643/q1, several representative single cell traces (black) fitted for step detection (blue) are presented in Figure 5.5B. Notably, the appearance of fluorescent signals later into acquisition were much more prevalent in EUB-643/q1 traces, as would be anticipated with effective quenching supplying photoprotection and increasing photostability. Plotting obtained step heights, thresholded from 3 units upwards, and fitting with a single Gaussian peak gave unitary intensity of ~4.5 units for both probes (Supplement 5.2).

Attributing bright fluorescent foci in live cells to be probe-hybridised ribosomes, the determined unitary intensity for EUB-643/q1 was used to approximate the spread of probe-bound ribosomes in loaded cells (Figure 5.5C). By dividing brightest cell mean intensities cell area by a unitary intensity of 4.5 units, the majority of cells with 100 nM, 200 nM and 500 nM were found to comprise 1-5 foci (98% 100 nM, 90% 200 nM, 92% 500 nM). Probe concentrations of 200-500 nM showed a higher number of cells with more than 5 foci, as can be seen via the representative composite FOVs.

Notably, at 500 nM the background autofluorescence particularly from free dye in solution was often high despite washing, making accurate intensity determinations and clear single molecule studies challenging; as such, concentrations between 100-200 nM were optimal for single molecule observations.

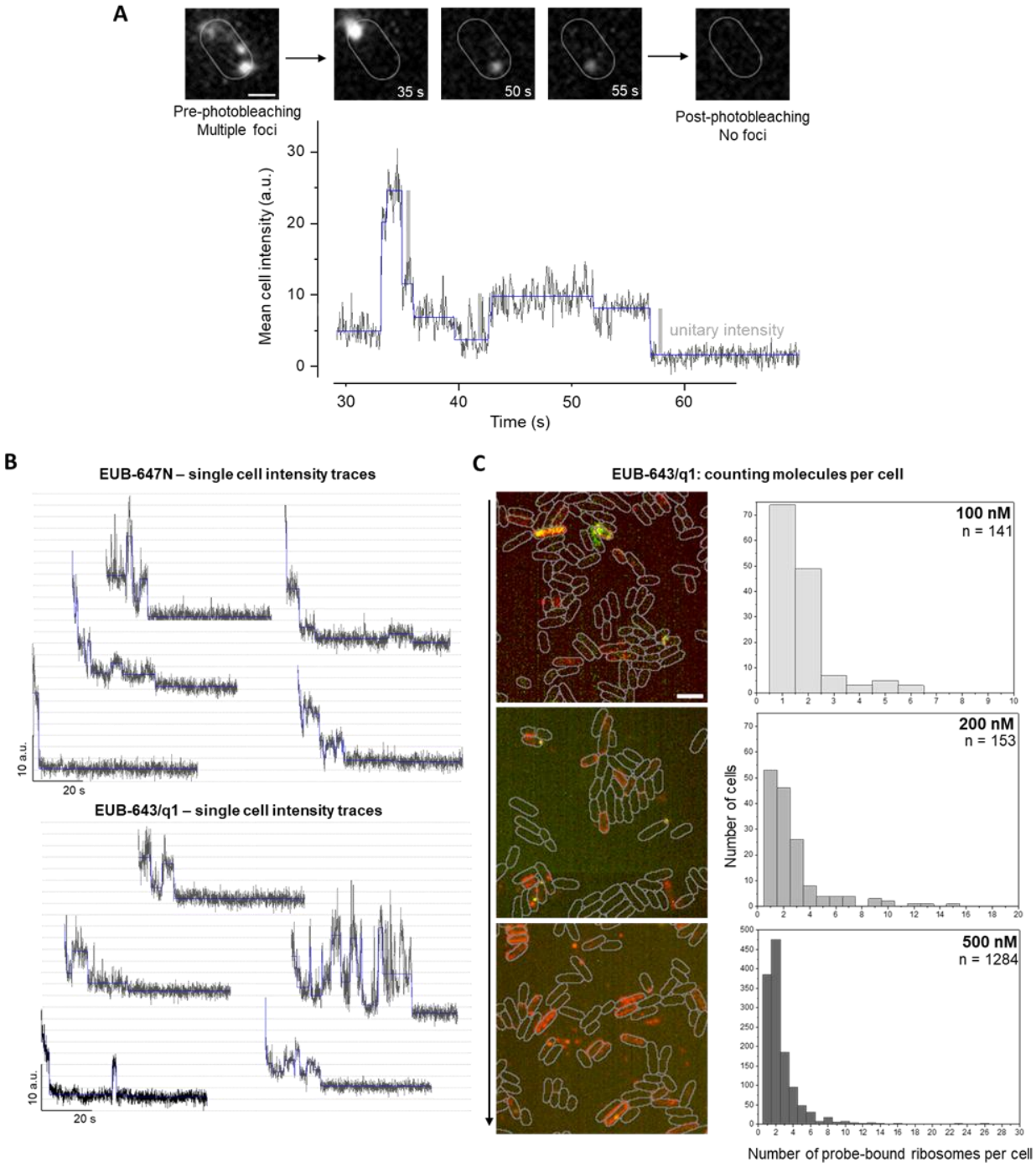


Figure 5.5 – Characterising the internalisation of the EUB-647N and EUB-643/q1 probes in M9Glu competent wildtype cells.
(A) Exemplar single cell trace of a live cell with passively internalised EUB-643/q1 giving 1-5 possible bright foci from the start of acquisition, down to 1-3 clear foci later in acquisition; associated step analysis of bleaching over time presented below for central montage.
Scale bar – 1 µm.
(B)Representative bleach-down traces and *AutoStepFinder* step fits where step heights correspond to single focus bleaching (unitary intensity) for cells with 100 nM EUB-647N and EUB-643/q1 probes.
(C) Representative composite FOVs of M9Glu competent wildtype cells with 100 nM, 200 nM and 500 nM EUB-643/q1 (left), alongside corresponding histograms of probe-bound molecule counts across numbers of cells (right). Scale bar – 3 µm.

5.5 Exploring single ribosomes *in vivo*

5.5.1 Single-particle tracking of EUB-643/q1-bound 16S rRNA

To capture the localisation and diffusion of single 16S rRNAs in M9Glu competent wildtype *E.coli* cells, cell samples with 100 nM EUB-643/q1 were imaged at ~10 mW 638 nm laser power in HILO mode, at 20 ms, 50 ms and 100 ms exposure times.

Previous single molecule studies of labelled ribosomes and small proteins have utilised exposure times of >20 ms to definitively capture variable subdiffusive properties of freely-diffusing 30S subunits and intact whole ribosomes, in the case of 16S rRNA (Bakshi et al., 2011, 2012; Mohapatra & Weisshaar, 2018; Sanamrad et al., 2014). Having established that resolving free dye in the cell would likely require <10 ms exposure time, imaging at 20 ms hoped to capture the faster diffusive subspecies of the free 30S ribosomal subunit (Sanamrad et al., 2014) and overcome motion blurring in higher exposure times. Considering multiple mobility behaviours have been demonstrated through several analogous methods (Bakshi et al., 2013; English et al., 2011; Mohapatra & Weisshaar, 2018; Sanamrad et al., 2014), but high binding behaviour ascribed to actively transcribing ribosomes seen across all, 100 ms exposure was applied alongside 50 ms exposure to attempt to capture any slow, potentially static or tethered molecules.

As in Chapter 4, individual bright foci were localised in *Moltrack* within their unique cell segmentations, and localisations linked to form tracked trajectories. Trajectories were retained from a minimum track length of 5 consecutive frames, giving n single molecule trajectories in the 20 ms (n=3736), 50 ms (n=2890) and 100 ms (n=987) samples. Supplement 5.3 presents representative *Moltrack* FOVs of tracked particles in M9Glu competent cells with passively internalised EUB-643/q1 (100 nM starting concentration), imaged at 20 ms, 50 ms and 100 ms exposure.

5.5.2 Ribosomal populations show subdiffusive properties

To compare the apparent diffusion coefficient (D*) of the tracked molecules across cell sample and exposure times, the mean-squared displacement (MSD) of the ribosomes was calculated as previously established (Stracy et al., 2015; Uphoff et al., 2013) for trajectories constructed from 5 consecutive frames and more (Supplement 5.4A). Normally, the D* would be derived from the MSDs of retained trajectories truncated to the first 5 steps. Here, however, due to low sampling size, all trajectories have been truncated into 5-step segments, a rolling window applied and all 5-step truncations for all trajectories retained (Mirjam Kümmerlin,

2024). This represents an oversampling of longer trajectories versus shorter ones, but still gives an approximation of diffusive ribosomal subspecies in the live cell.

Considering the ability of EUB-643/q1 to bind both 30S freely-diffusing subunits and intact 70S ribosomes via their 16S rRNA target hybridisation site, a complex distribution would be expected for D* populations. To agree with published reports, a predominantly bimodal population for 30S and 70S ribosomal units corresponding to approximately 0.3 μm s^{-1} and 0.1 μm s^{-1} respectively would be anticipated (Sanamrad et al., 2014). With the caveat of insufficient sampling, the D* distributions of tracked molecules in M9Glu competent wildtype cells with 100 nM EUB-643/q1 at 20 ms, 50 ms and 100 ms exposure were fitted with a two free species model (Figure 5.6).

Most reports have captured the mobility behaviours of ribosomal subunits with exposure times of 20 ms – enabling higher resolution of faster-moving molecules, and reduced motion blurring. Figure 5.6 shows that, at 20 ms exposure, the 48% faster diffusing species was attributed a D* of 0.37 $\mu m^2/s$ and a 52% slower moving population at D* 0.04 $\mu m^2/s$.

These D* values correspond to previous reports of ribosomal species: Bakshi et al. found an ~20-30% population of 30S diffusing at D* of ~0.14 $\mu m^2/s$ and 70S-polysomal population at ~0.02 $\mu m^2/s$ (Bakshi et al., 2012). Mohapatra & Weisshaar subsequently found fast-growing cells' 30S subunit D* of 0.2 $\mu m^2/s$ and the 70S species to have a D* of 0.02 $\mu m^2/s$ (Mohapatra & Weisshaar, 2018). More recently, using the EUB-643/q1 probes via electroporation in *E. coli* to capture 30S and 70S ribosomal units at 20 ms exposure, Dr Kümmerlin found a 13% 30S freely diffusing population at ~0.34 $\mu m^2/s$, and an 87% bound population at 0.06 $\mu m^2/s$, predominantly localised to the cell periphery (Mirjam Kümmerlin, 2024). The 20 ms bound and fast diffusing populations certainly have comparable diffusion rates to those documented in Bakshi et al., Sanamrad et al., and Dr Kümmerlin's works, particularly for the freely diffusing 30S subunit fraction. However, the estimates for the fractions of each species bear little resemblance with previous studies – this is likely due to the oversampling of longer trajectories, where if more fast trajectories are generally captured at 20 ms, would likely overrepresent diffusing 30S subunits over bound 70S ribosomes.

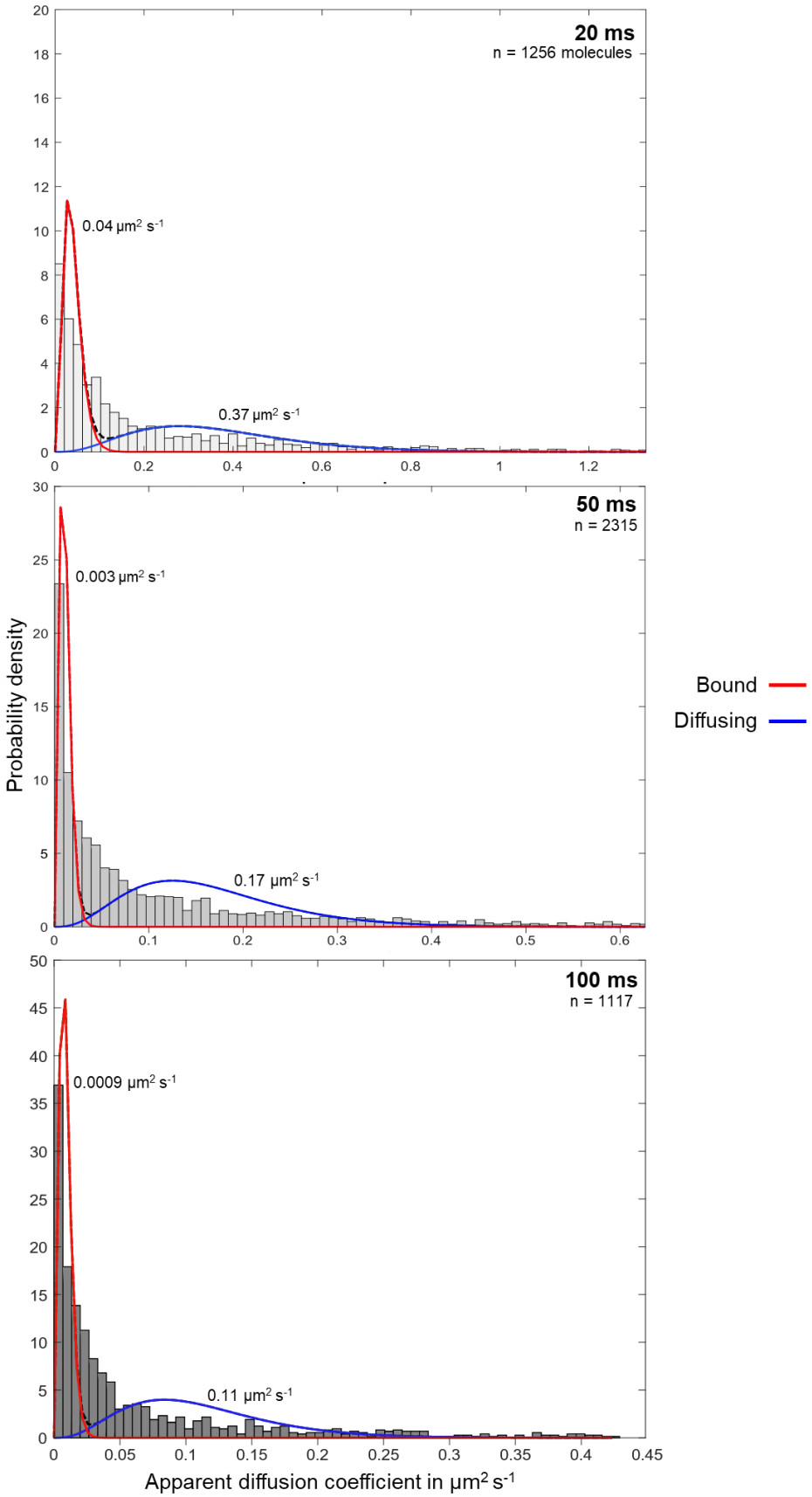


Figure 5.6 – Exploring the diffusive properties of tracked fluorescent molecules in M9Glu competent wildtype cells with 100 nM EUB-643/q1: apparent diffusion coefficients. Apparent diffusion coefficient distributions fit with two-free species curves for cells imaged at 20 ms, 50 ms and 100 ms exposure times (red as bound, blue as diffusing). n gives the numbers of tracked molecules per sample post-thresholding; N=1.

With increased motion blurring and better resolution of confined molecules at 50 ms and 100 ms exposures, it would be anticipated that the diffusion states observed were increasingly static. Indeed, from the two-state fits, at 50 ms ~61% bound molecules showed a D* of ~0.003 $\mu m^2/s$, further slowing to 0.0009 $\mu m^2/s$ for 50% molecules at 100 ms. This left diffusing fractions of 49% free at 0.17 $\mu m^2/s$ captured at 50 ms, and 50% free molecules diffusing at 0.11 $\mu m^2/s$ at 100 ms. The substantial left-shift in bound diffusion coefficients brings the confinement observed into the region of chromosomal diffusion – exceptionally slow. These bound populations could therefore represent polysomes at the nucleoid periphery, perhaps engaged in transertion, whereby the ribosomes would be effectively membrane-tethered by their nascent polypeptides, or potentially confined to dense actively-translating ribosomal hotspots in the cell poles (Bakshi et al., 2015).

Noted, all the two-free fits across 20-100 ms are not ideal fits, and indeed deviate substantially from the histogram, explaining the non-comparable fraction percentages. This likely partially arises from the oversampling occurring via D* calculation; although, corresponding MSD over time plots suggest that oversampling effect is only noticeably prominent in the 100 ms sample (Supplement 5.5). An attempt at an improved fit via a three species state can be seen in Supplement 5.6, where an intermediary "slow-diffusing" population is touted for freely-diffusing 70S ribosomes (not engaged in translation); however, the bound populations are also poorly fit across 20 and 100 ms samples. Overall, higher statistical power is required, and further experimentation for better sampling.

5.5.3 Populational distributions of tracked molecules

Exploring the distributions of tracked molecule counts per cell and molecule durations across the cell can give a more holistic view of the subpopulations of probe-hybridised particles and cell loading.

Counts per cell

The internalisation rates and subsequent spread of individual tracked molecules across a population of cells would be anticipated to be similar, irrespective of different exposure times – unless the faster 20 ms exposure significantly enabled greater capture of fast-diffusing molecules.

Investigating the normalised counts of tracked molecules from M9Glu competent wildtype cells imaged at 20 ms, 50 ms and 100 ms showed that the majority of cells exhibited 1-2 fluorescent foci (Figure 5.7). Interestingly, the 20 ms sample showed a much higher frequency of cells with 2-10 foci (58% 20ms, versus 32% for 50 ms, 33% for 100 ms), and indeed the highest frequency of cells with >20 foci (12% 20ms, 8% 50 ms, 4% 100 ms). This suggests that the lower exposure time does enable the capture of more of the faster-diffusing population, which may be lost in longer exposure times.

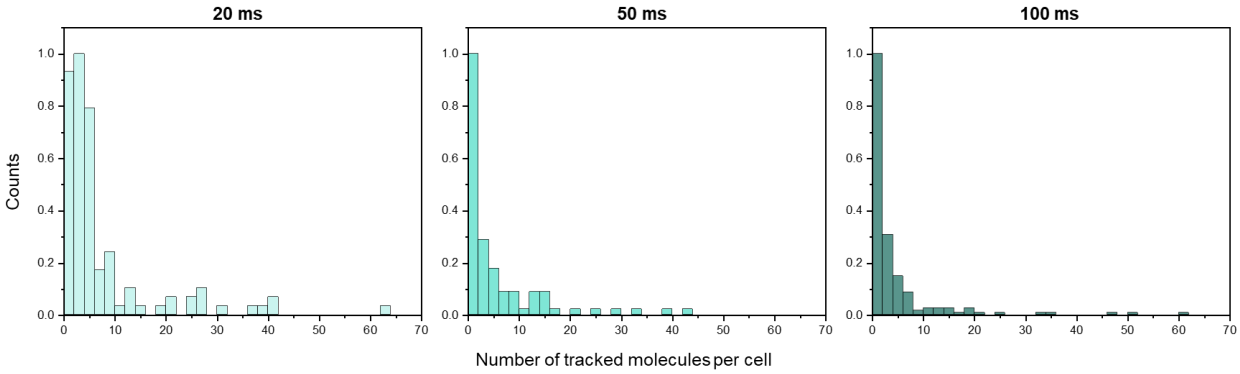


Figure 5.7 – Tracked molecule distributions comparing normalised counts per cell in M9Glu competent wildtype cells electroporated with 100 nM EUB-643/q1 imaged at 20 ms, 50 ms and 100 ms exposure. Bin size = 2.

Track durations

With the ability to capture faster-diffusing, likely more transient molecules, it was hypothesised that the duration of molecular tracks would increase with increasing exposure time. Figure 5.8 shows the normalised distributions of tracked particle durations, clipped at 0.4 counts (Supplement 5.7 displays the full range per condition). All distributions are highly overlapping at the lower time distributions – all exposure times demonstrated 74-85% molecules have durations between 0.2-2 seconds – suggesting that most of these signals are highly diffusive, potentially correlating to free 30S subunits. 50 ms and 100 ms acquisitions showed 3% and 12% molecule durations above 4 seconds, denoting the likely presence of immobile molecules, perhaps tethered or translating ribosomes.

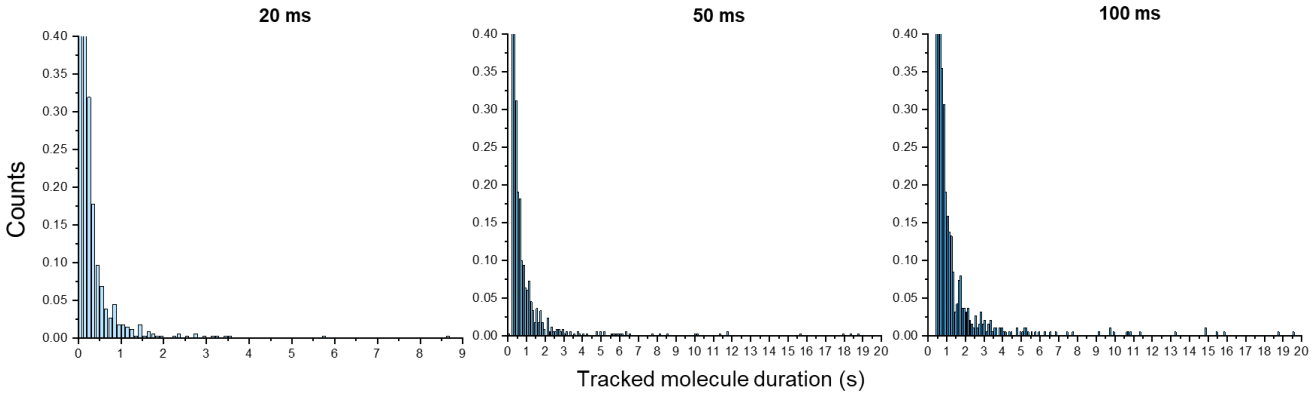


Figure 5.8 – Tracked molecule duration distributions comparing normalised counts per cell in M9Glu competent wildtype cells electroporated with 100 nM EUB-643/q1 imaged at 20 ms, 50 ms and 100 ms exposure. Bin size = 0.1 (seconds).

5.5.4 Tracked ribosomes are predominantly peripherally localised

Many studies have established the major population of 70S ribosomes to be membrane-proximal and nucleoid-excluded, from cryo-electron microscopy, live cell fluorescence microscopy (Lewis et al., 2000) and single molecule studies (Bakshi et al., 2012, 2015). This peripheral and polar localisation of whole ribosome units is in line with observations of high transcription rates of rRNA operons clustered near the chromosomal origin of replication (Gaal et al., 2016), generating "transcription foci" and proposed hotspots for polysome formation and coupled transcription-translation in fast-growing cells (Jin et al., 2017). More recently, the free 30S subunit has been reported to freely diffuse across the cellular milieu including through the dense nucleoid (Mohapatra & Weisshaar, 2018; Sanamrad et al., 2014).

Figure 5.9A presents histogram distributions of the tracked molecules distance from the cell membrane, across the samples imaged with different exposure times. 20 ms tracked molecules showed the widest spread of distances from the membrane, further corroborating the increased capture of faster diffusing molecules, which would likely be able to traverse the whole cell, including nucleoidal perfusion as 30S free subunits (Sanamrad et al., 2014). Both 50ms and 100 ms molecules show a higher proportion of molecules within 100-200 nm of the membrane, suggestive of whole 70S ribosomes which could be actively translating, in ribosomal hotspots, or even transerted into the membrane (Bakshi et al., 2015).

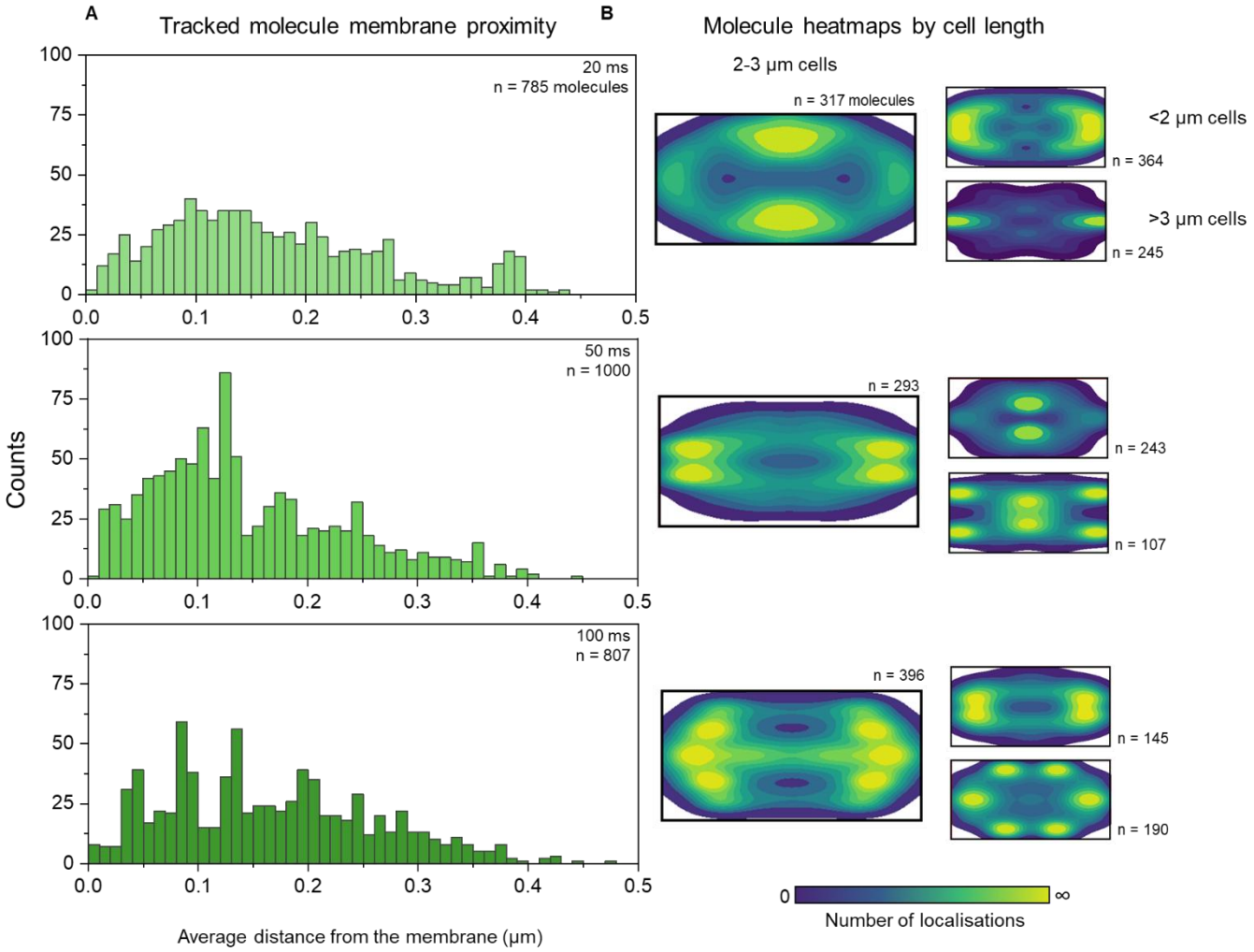


Figure 5.9 – Spatial distributions of tracked molecules in M9Glu competent wildtype cells with 100 nM EUB-643/q1, imaged at 20 ms, 50 ms and 100 ms exposure.
(A) Histogram distributions of average molecule distance from the membrane; bin size = 0.01.
(B) Heatmaps showing the density of molecule localisations in different regions of the cell; the central plot denotes the average unit cell (length 2-3 µm), and the two corresponding righthand heatmaps denote cell lengths of <2 µm and >3 µm per strain sample.
n gives the number of molecules.

Plotting heatmaps of molecular localisations by cell length likewise corroborate these observations (Figure 5.9B). From the average cell (2-3 µm long), most tracked molecules across all samples remain membrane-proximal, with increasing shifts towards highly localised membrane proximity. This could reflect transertion of polysomes and membrane architecture in enabling efficient transcription-translation to proceed in exponentially growing cells. Noted, there is also a small possibility that highly static trajectories could represent membrane-inserted and de-quenched probes, thereby false positives[10]. All heatmaps indicate

[10] Notably in the case of false positives through probe-membrane implantation, the likelihood is low considering the hydrophilicity of the EUB-643/q1 probe, and even if dequenching were to occur, it is

that the majority of localised molecules are peripheral, with the 20 ms unit cell showing higher nucleoid localisations, as would be expected with decreased motion blurring enabling the resolution of the faster-diffusing 30S subunit population. Observations around 2-3 µm average length cells are caveated by lower sampling size.

By way of observing dense 70S ribosomal populations particularly in the cell poles, where high mRNA concentrations ready for processive translation accumulate, Bakshi et al. also demonstrated the anti-correlation between ribosome and DNA densities in *E. coli* via endogenous staining in widefield microscopy (Bakshi et al., 2012). This anticorrelation was recapitulated by Farrar et al. in fixed *E. coli* (Farrar et al., 2025), and notably is suggested here via maximum intensity projections of live *E. coli* with EUB-643/q1 probes (Figure 5.10). Axial line profiles across the length of the exemplar cells depict the bimodal and trimodal intensity distributions across shorter and longer cells respectively, as would be expected of wildtype *E. coli* nucleoidal architecture and peripheralized majority ribosomal populations. The trimodal line plot from the cell presenting 3 bright foci is of particular interest (Figure 5.10A), as this likely corresponds to a dividing cell – in agreement with previous observations that in replication, ribosomal populations accumulate around the midcell FtsZ ring (Chai et al., 2014).

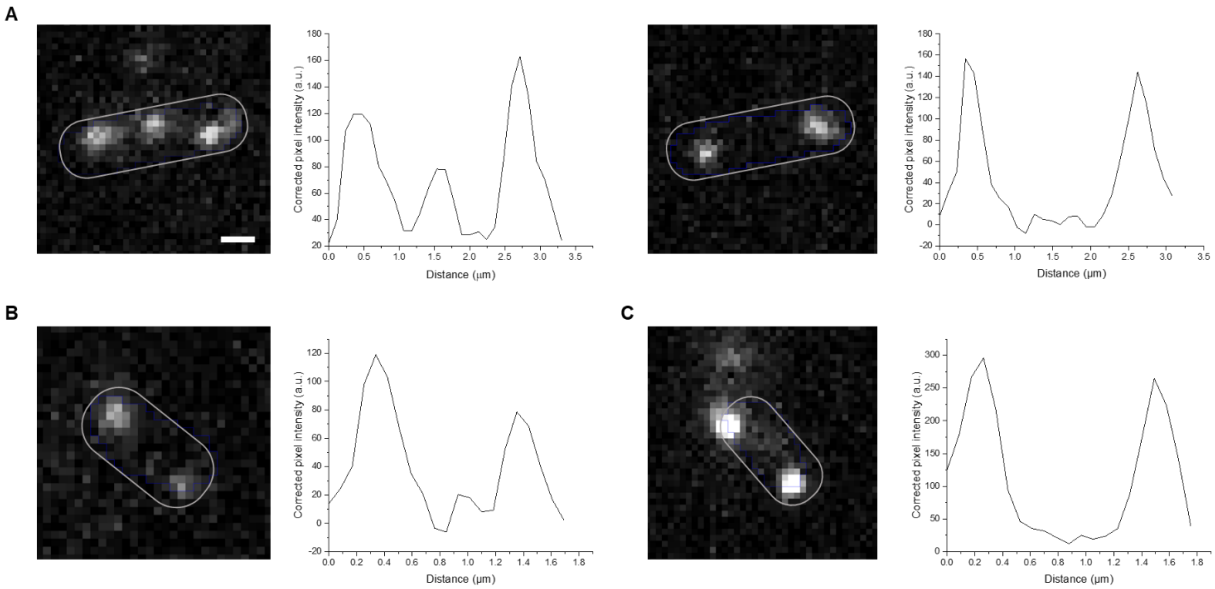


Figure 5.10 – Exemplar stills of M9Glu competent wildtype cells (A-C) showing polar-proximal fluorescent foci, suggestive of ribosomal 'hotspots', with corresponding line profiles along the length of the cell. (A) depicts the same cell, initially with 3 foci suggestive of a mid-replication cell, then the 2 polar foci; other exemplars shown in (B, C).
Scale bar – 500 nm.

unlikely that particularly long localisations would be generated due to photobleaching – thereby decreasing the probability of retaining false positive localisations into later analysis.

5.5.5 Observing behaviours of individual ribosomal particles

As with the mRNAs visualised with transiently binding exogenous fluorogenic probes in Chapter 4, the greatest insight often comes from direct observation of the imaged fluorescent foci themselves. Across all exposure times, three key behaviours were consistently observed in ~50% cells containing signal: fast-diffusing foci that were able to traverse the length of the cell, slow-diffusing and regionally-constrained foci, and finally, highly static, membrane- or polarly-localised signal. Based on the key reports discussed and evidence presented so far, these three behaviours were thought to relate to freely-diffusing 30S subunits, free 70S ribosomes or confined 30S subunits, and whole engaged and actively-translating 70S ribosomes.

Figures 11-13 depict representative observations made across M9Glu competent cells imaged at 20-100 ms exposure. As in Chapter 4, the representative cells are shown as a montage of stills taken from the real-time video acquisition, correlated to text descriptions of real-time individual foci. For each cell, the *Moltrack* output of all tracked molecule trajectories in the cell is presented. Alongside the full trajectory *Moltrack* image, are *Moltrack* images of specific track segments (focus tracks) with corresponding plots for each focus tracked molecule's step size over time and proximity to the cell membrane, by way of representing diffusive and spatial behaviours respectively.

20 ms exposure

Most foci observed in 20 ms-imaged cells were fast-diffusing (Supplement 5.8 as an exemplar), but there were also many molecules that interchanged between fast-diffusing and slow-diffusing/static states. The exemplar cell in Figure 5.11 presents foci that exhibit all these diffusive states.

From time 0 to ~34 seconds later, tight polar localisation is seen for foci at both cell poles, exhibiting bound behaviours and slow localised diffusion – these are likely indicative of 70S actively engaged ribosomes, confined at the cell poles and interacting with mRNAs. Indeed, the *Moltrack* 'all trajectories' cell image shows three regions of high density localisations: one at each of the poles, and a third at the midcell – all of which relate to the visible polar foci and subsequent midcell focus. The selected focus track segment corresponds to the highlighted bright focus in the upper polar region seen at 28.2 seconds and subsequently at 32.5 seconds. Characteristic of the bright polar foci, this track segment shows step sizes primarily constrained to 0.1-0.2 μm over the ~5 second timescale plotted, remaining within 250 nm of the membrane at most – hence denoting a highly constrained molecule. Only at ~31 seconds

does the step size spike to ~0.45 µm, corresponding to the observed slow diffusion of the polar focus from the left to righthand side of the cell pole as shown at 28.2 and 32.5 seconds' stills.

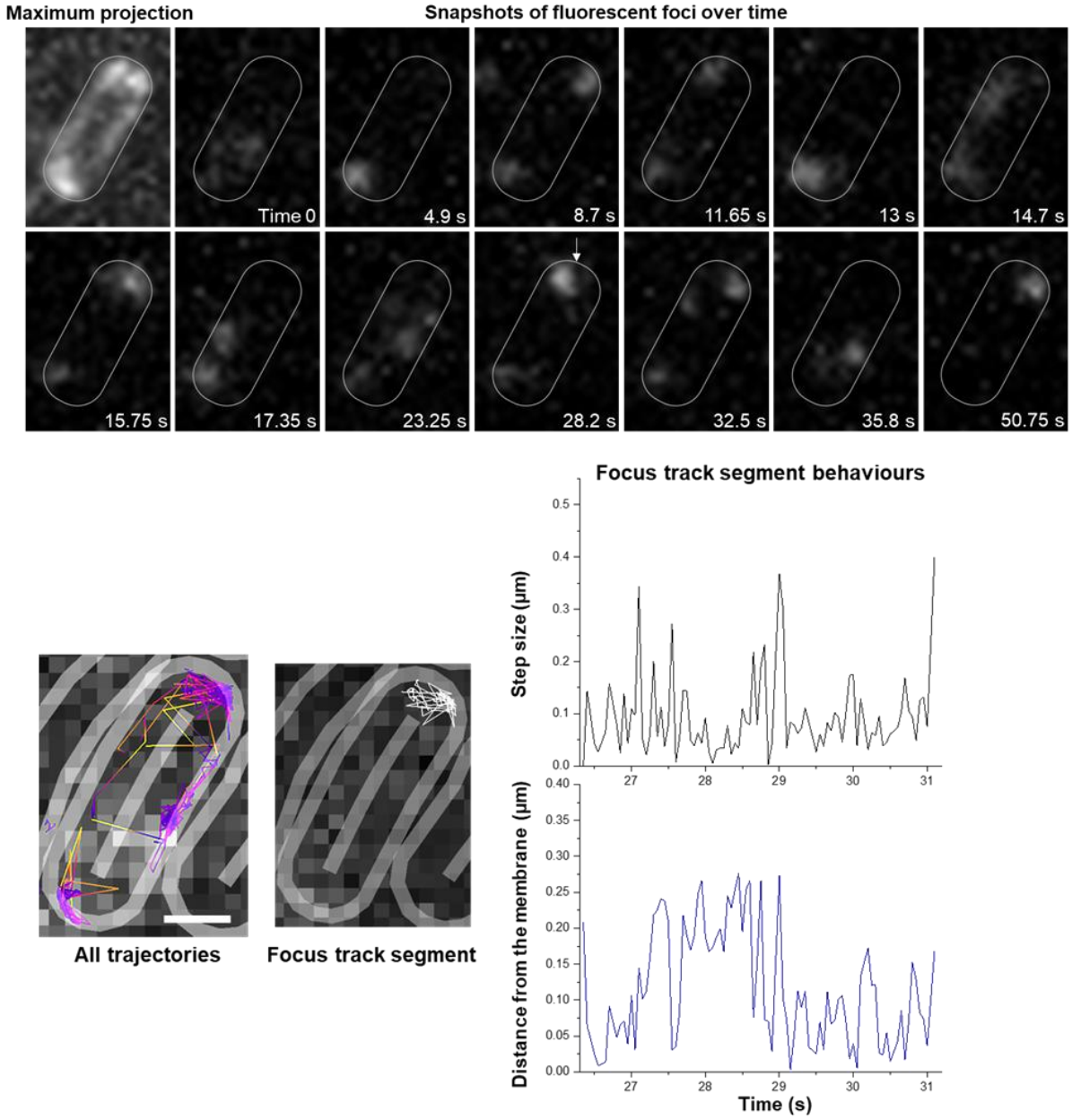


Figure 5.11 – EUB-643/q1 visuals over time: exemplar cell at 20 ms exposure.
(Upper) Montage of cell stills, showing the appearance of fluorescent foci over time during a 100 second acquisition (20 ms exposure, ~10 mW 638 nm laser; still time given in seconds, lower corner of each still.
(Lower) *Moltrack* visuals: focus track segments and behaviours.
Moltrack 'all trajectories' showing the trajectories for all tracked molecules in the cell (left); track colour denotes speed, where yellow is fastest and dark blue slowest.
The 'focus track segment' denotes a specific track segment analysed in *Moltrack*; the righthand side plots, molecule step size (black) and distance from the membrane (blue) over time, are representations of the behaviours of the molecule in the focus track segment. Scale bar – 500 nm.
This flow applies to all further exemplar molecule figures in 5.5.5.

Shown at 35.8 seconds is a midcell static focus; this midcell bound focus is visible in the acquisition for ~7 seconds, before diffusing off to the upper cell pole as shown in the final still at 50.75 seconds. The upper cell focus bleaches at ~51.5 seconds. Although not shown here, a further polarly localised and low mobility focus at the lower cell pole is seen intermittently from 85 to ~139 seconds, demonstrating that with a sufficient concentration of internalised quenched probes, complementary target hybridisation is tractable on the minutes timescale.

Supplement 5.8A shows a second example of a cell tracked at 20 ms exposure, whereby tracked fluorescent foci are rapidly diffusing, even traversing the whole cell (denoted cell A, Supplement 5.4). This is represented in the montage as fluorescent foci predominantly appearing at either cell poles, as shown by the two focus track segments. One at each pole, with track segment 1 highlighted as the fluorescent focus at 26.1 seconds in the montage, the step sizes of the highlight focus varying between 0.1-0.5 μm over milliseconds and little consistent association with the membrane denotes a highly transient molecule – likely freely diffusing 30S subunit.

50 ms exposure

50 ms exposure seemingly a sweet spot in exposure time for capturing both fast-diffusing, likely free 30S subunits, and more immobile, potentially tethered or actively translating 70S ribosomes, all types of behaviours are exhibited in the exemplar cell presented below (Figure 5.12).

At time 0, three bright and distinctive foci can be seen at the upper cell pole and closer to the midcell. The highlighted left midcell focus is represented by track segment 1 – over the first 3 seconds, this molecule localises at the midcell, diffuses down the cell midline to the lower pole where it briefly binds (as shown by ~ 50 milliseconds at 0.05 μm step size) and then diffuses back to its original midcell position. The generally lower membrane proximity, even during the brief binding event, is synonymous with a largely freely diffusing molecule, suggesting that this could be a diffusing whole 70S ribosome that is yet to be engaged in translation.

From 3.6-6.8 seconds, extremely rapid diffusion of bright foci such that they are less consistently resolvable, can be seen; noting, that the bright polar focus at 3.6 seconds is not the same molecule as represented by midcell focus, track segment 1. These rapidly diffusing foci are most likely freely diffusing 30S subunits, with lesser resolution than can be obtained

at 20 ms exposure, but that are well represented by the long trajectories depicted in the *Moltrack* 'all trajectories' cell image.

From 7.35-12.55 seconds, a more constrained focus appears, slowly diffuses across the upper cell pole and becomes highly confined (~ 3.5 second period) before bleaching. Focus track segment 2 corroborates this observation, whereby short and rapid changes in step size between ~7.5-10 seconds, variably distant from the membrane, show the molecule's slow polar diffusion, only to become highly static with a step size of ~0.025 µm, around 200 nm from the membrane, between 10-12.55 seconds. The specific high confinement locus is highlighted on the *Moltrack* focus segment 2 image – and is highly indicative of a 70S ribosome bound and actively translating at a polar hotspot.

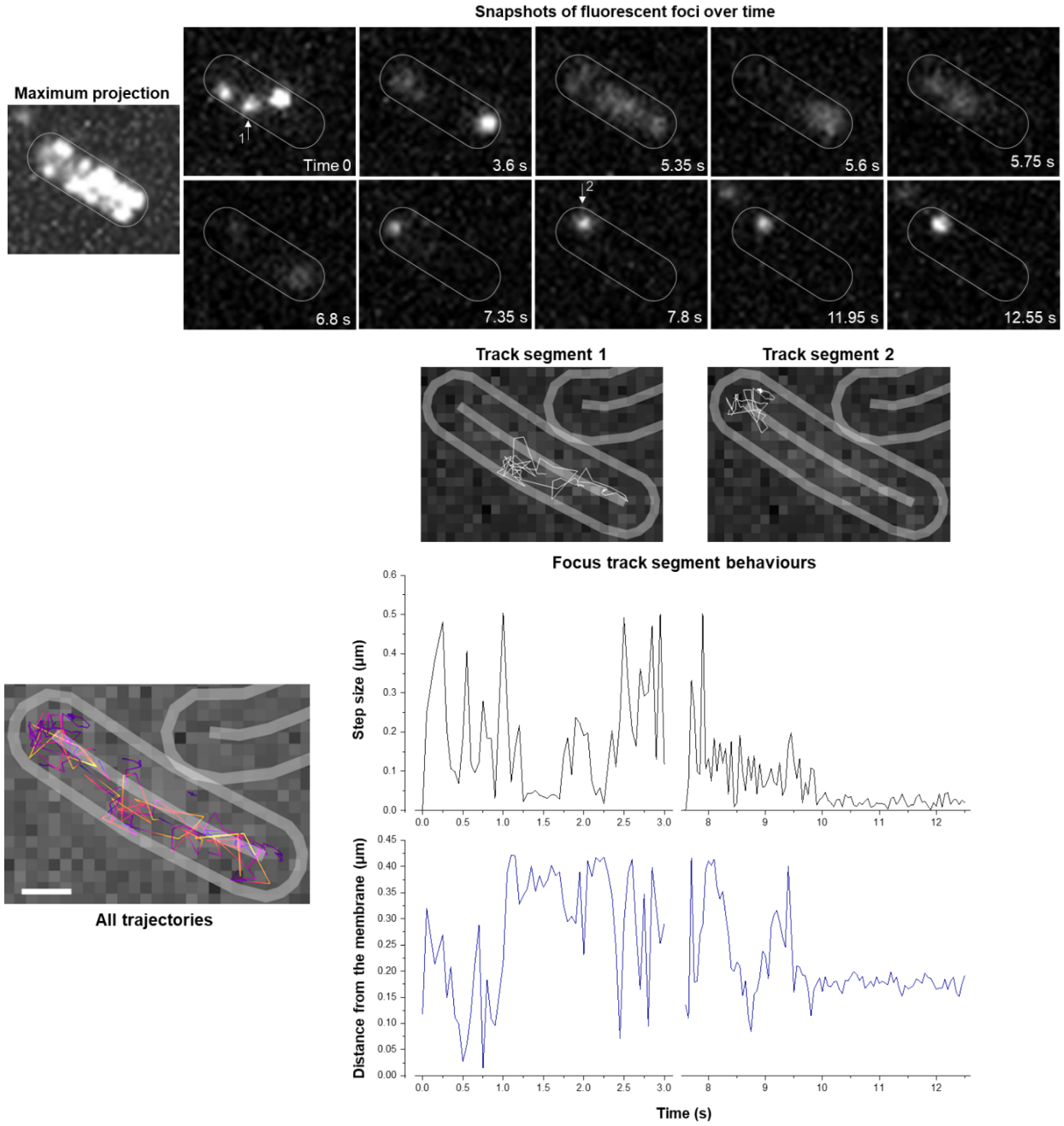


Figure 5.12 – EUB-643/q1 visuals over time: 50 ms exposure, exemplar cell.

100 ms exp

With a variety of behaviours equally visible at 100 ms exposure (Supplement 5.8B showing a rapid diffusion exemplar, denoted cell A on Supplement 5.4), the exemplar cell here denotes a possible mid-division event.

Whilst there are bright and fast-diffusing foci which appear to transit along the whole length of the cell (as seen by light tracks in the 'all trajectories' cell image) – and perhaps then, are shared between the two near-daughter cells – a single static focus in visible at the midcell from time 0 to for 23.65 seconds later, possibly FtsZ ring site (Figure 5.13). The focus track segment shows that this confined focus has a step size ~0.05 μm across the whole 11-23.65 seconds tracked, emphasizing that the same molecule is being tracked and is highly bound at the midcell.

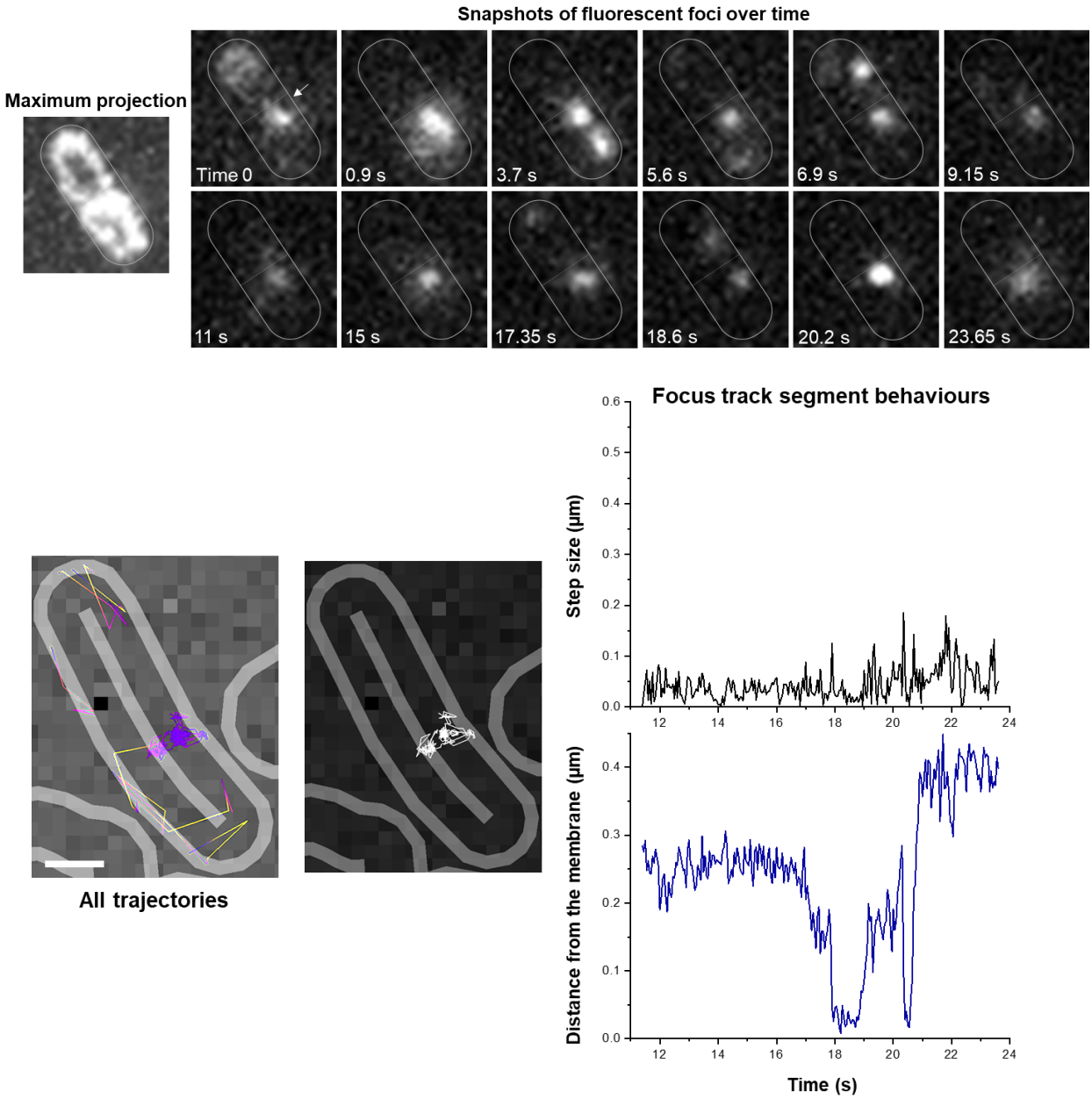


Figure 5.13 – EUB-643/q1 visuals over time: 100 ms exposure, exemplar cell.

Interestingly, this bound molecule appears confined ~250 nm to the cell membrane for ~5.5 seconds, only to move closer to the membrane between 18-19 seconds and at ~20.2 seconds – with the varying brightness as seen in the montage at equivalent timepoints. Indeed, taking the mean intensity per frame of the highlighted midcell focus further demonstrates variations correlating to multiple individual molecules at the same locus, suggesting multiple ribosomes are present at the same locus, perhaps polysomal translation or transertion (Figure 5.14). Correlating the focus intensity trace with the montage and an approximated unitary intensity of ~ 5 a.u., there could even be as many as >15 overlaid molecules at the same hotspot.

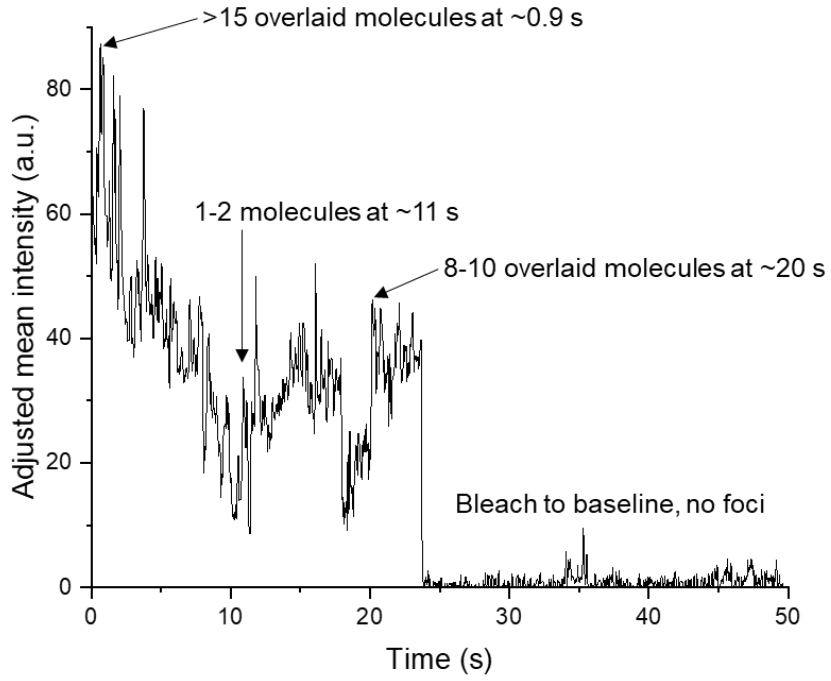


Figure 5.14 – Adjusted mean intensity trace of the core midcell static focus across the entire 50 second acquisition at 100 ms exposure.
Annotations estimate the number of possible molecules observed, correlated to the intensity variations by approximating the unitary intensity at ~5 a.u.

Just before disappearing, this static focus detaches, diffuses into the centre of the cell and then to the membrane, and further into the midcell centre again, as can be seen towards the end of the track segment for membrane proximity, increasing up to 400 nm from the membrane.

As such, it is likely that this midcell static focus could indeed be a translational hotspot for whole 70S ribosomes, engaged in active translation for much of this observation – especially as the midcell periphery hosts many essential gene loci, upregulated during cell division.

5.6 Discursive summary

Endogenous RNA labelling via passively-internalised probe hybridisation – a parallel approach

One complementary avenue to engineered probe hybridisation sites as seen in engineered *in vivo* smFISH is, using the same FISH-PAINT premise, to use probes that are sequence-specific to endogenous RNAs, no engineering necessary. Of course, exogenous probes must still be internalised – but the ability to passively internalise such labelled ssDNAs, i.e. without induction and extensive membrane disruption, could much improve physiological functioning and readout of target RNA behaviours.

The 16S rRNA of the 30S ribosomal subunit and its corresponding well-established EUB338 probe thereby act as an excellent proof-of-concept for endogenous probing of single rRNAs *in vivo*. As the viability of cells appears to be unaffected by passive internalisation and no recovery period, the similar intensity distributions across all compared internalisation methods indeed demonstrate that passive probe internalisation is a tractable, and physiologically preferable, option.

Using the fluorogenic EUB-643/q1 probe, the mobility behaviours of freely diffusing 30S subunits and 70S whole, likely actively-translating, ribosomes were able to be captured at different exposure times. Considering that 20 ms has most commonly been applied in suitably capturing the freely diffusing 30S subunit in a number of published works, the 0.04 $\mu m^2/s$ bound and 0.37 $\mu m^2/s$ freely diffusing fractions are comparable rates to those reported by Bakshi et al., Sanamrad et al., and recently, Dr Kümmerlin. These hence likely correspond to 70S ribosomes and freely diffusing 30S subunits respectively. As exposure time increased, both fractions consecutively decreased in diffusion rates – and indeed, tighter bound fractions could even be capturing polysomal translation, in agreement with Sattler & Graumann's observations of mRNA mobilities in ribosomal polar hotspots (Sattler & Graumann, 2021). Excitingly, spatial observations further recapitulated seminal observations of ribosomal mobility and loci: through spatial localisation heatmaps and single molecule acquisitions, most cells exhibited highly polar fluorescent foci – indicative of nucleoid exclusion, and high-density ribosome hotspots. Across all exposure times, foci were seen to rapidly diffuse across the cell – again, corroborating seminal observations that the smaller 30S subunits can freely diffuse through the nucleoid. Further recapitulative observations foci constrained at the midcell are suggestive of ribosomal hotspots (Bakshi et al., 2012; Castellana et al., 2016; Sanamrad et al., 2014), most notably in longer cells with reduced nucleoidal density at the midcell, and mid-division cells (supplying protein at the FtsZ ring)(Chai et al., 2014).

It is likely, however, that more than just two mobility states (bound and freely-diffusing) are occurring: freely-diffusing 30S subunits through the nucleoid, in the periphery, whole diffusing 70S ribosomes in the periphery, phase-constrained 30S and 70S, and highly constrained 70S ribosomes actively translating mRNAs, are all possible. Indeed, translating ribosomes could be complexed with mRNAs, could be in a polysomal complex or could be transerted via the membrane – all of which would change the apparent diffusion readouts of labelled molecules. Whilst a 3 free species fit was attempted to capture more diffusive states, there are insufficient data for the D* fits here – further experimentation is required for more reliable representations of 30S and 70S mobility fractions in live *E. coli*.

Concerning mobility fractions and behaviours of ribosomal subunits and whole ribosomes, there are many opportunities to extend exploration. Whilst 30S subunits appear to distribute uniformly across the cytoplasm and nucleoidal regions irrespective of growth condition, the extent of 70S nucleoid exclusion is more impacted by growth phase and rate. With passive internalisation, further optimisation of different growth media and competency states could enable wider appraisal of growth states on ribosomal behaviours in the live cell.

Indeed, a clear caveat here is the lack of control verification for wildtype ribosomal behaviours. It would be hence interesting to see if Sanamrad et al.'s observations of free 30S behaviours during rifampicin-mediated mRNA depletion and translation inhibition via erythromycin (Sanamrad et al., 2014), or chloramphenicol (Siibak et al., 2009), in reducing 70S assembly could be recapitulated. Taken together, the mobilities, behaviours and proportions of 70S ribosomes and 30S subunits could be probed across various growth, stress and adaptive states – of potential impact in further understanding gene expression regulation, and even antibiotic mechanism of action.

On the fundamental importance of growth conditions

The tractability of passive internalisation is promising, in both maintaining physiological states of the cell as much as possible, and towards further elucidating the impact of electroporation as a primary internalisation method in single molecule studies.

It is speculated that, since the predominant determinants of passive internalisation appear to be media outgrowth and competency, the Gram-negative membrane physicochemical composition defines the passive internalisation capabilities of a cell. The impacts of growth variables have long been discussed (Jones et al., 1981; Schaechter et al., 1958), and chemical and thermal competency procedures have repeatedly been shown to influence phospholipid

composition and membranal phase transitions (Harlos & Eibl, 1980; Jacobson & Papahadjopoulos, 1975; Overath et al., 1975; Reusch et al., 1986).

With complex envelope and membrane biogenesis pathways in *E. coli*, the growth medium of cells has significant impact on the global expression of essential genes. For example, glucose-supplemented LB-grown *E. coli* have been shown to exhibit much higher rRNA and transcriptional/translational accessory factor expression, via gene expression profiling (Tao et al., 1999). Contrastingly, *E. coli* grown in minimal media show a significant uptick in the *rpoS* stress-response expression network (Tao et al., 1999), and have more recently been shown to exhibit stress-related restructuring, particularly regarding outer membrane, cell wall integrity and protein-lipid composition (Babu et al., 2011). These reports speak to the importance of membrane integrity and regulation in determining adaptive cell genotype and phenotype – further defining passive internalisation in reducing membrane perturbation as a useful avenue for physiological appraisal of single biomolecule behaviours.

Of course, if growth medium and competency states affect global expression networks of the cell, core machinery like RNAPs and ribosomes are intrinsically impacted: at the levels of expression rates, intracellular concentrations, and functional activity (Fan et al., 2023; Gyorfy et al., 2015). In agreement with Tao et al.'s gene expression profiling, Bakshi et al. (2014) also demonstrated upregulated synthesis of rRNAs in rich media-grown *E. coli*, hence more ribosomes for higher global translation rates; further corroborated by Mohapatra and Weisshaar, showing 70S ribosome distributions varying in fast and slow growth conditions (Mohapatra & Weisshaar, 2018). RNAPs show similar trends, where larger clusters and circulating RNAPs in rich media-grown *E. coli* have been seen at a single molecule resolution in fixed cells (Endesfelder et al., 2013). And so stands the importance of nutrient availability (ergo growth medium) on the principal actors of gene expression in the *E. coli* bacterium: in the crowded cell, the careful balance of RNAP and ribosome expression, stoichiometry and functional coupling define the cell itself.

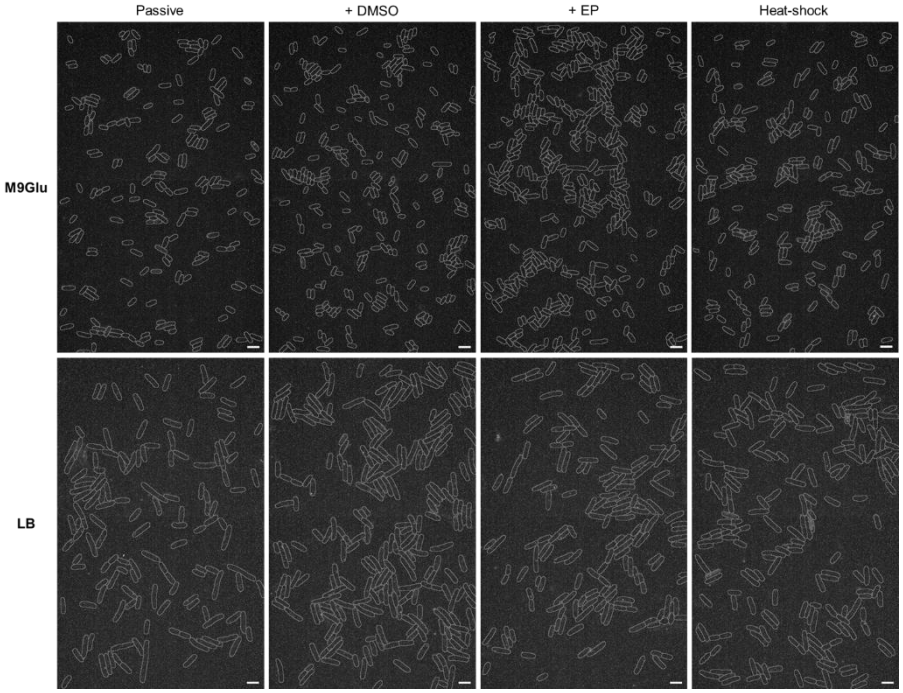
From this Chapter, two key points emerge in contributing to understanding the complex processes of gene expression and actions of core machinery in the cell: i. the growth state of the cell is fundamentally important; and ii. the least perturbation to cell membrane integrity is ideal.

Passive internalisation of fluorogenic probes complementary to a target RNA (rRNA and ribosome in this case) certainly contributes to the second, and opens up many possibilities for single molecule study. Many dual fluorescent systems – endogenously labelled proteins and

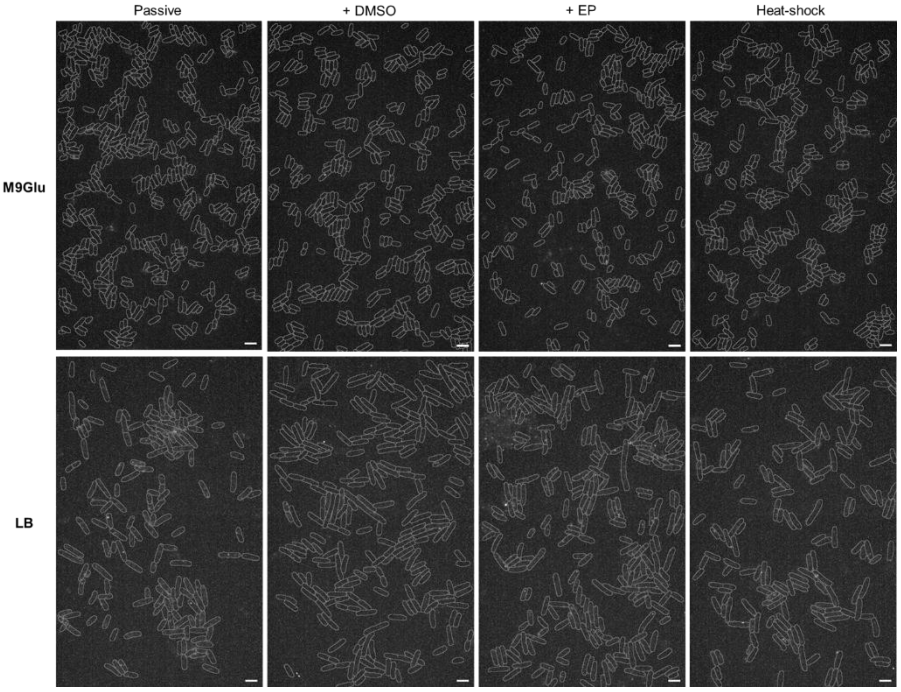
FISH probe-labelled RNAs – could be employed to probe the interactions of core machinery with specific RNAs and spatiotemporal behaviours of different molecules (ribosomes, RNAPs, degradosome components). Pushing passive internalisation with larger labelled biomolecules (sigma factors, for example) could enable better appraisal of RNA behaviours and core machinery interactions in states of stress. As for deepening knowledge around growth medium, phase and competency states – omics-level insights are most likely to give a solid baseline towards experimental design, utilising growth state at a single molecule level to probe cellular adaptation and fitness. Combining knowledge gained at these different levels of the cell will, in the longer term, greatly contribute to the nuanced understanding of gene expression in *E. coli*.

Supplement 5

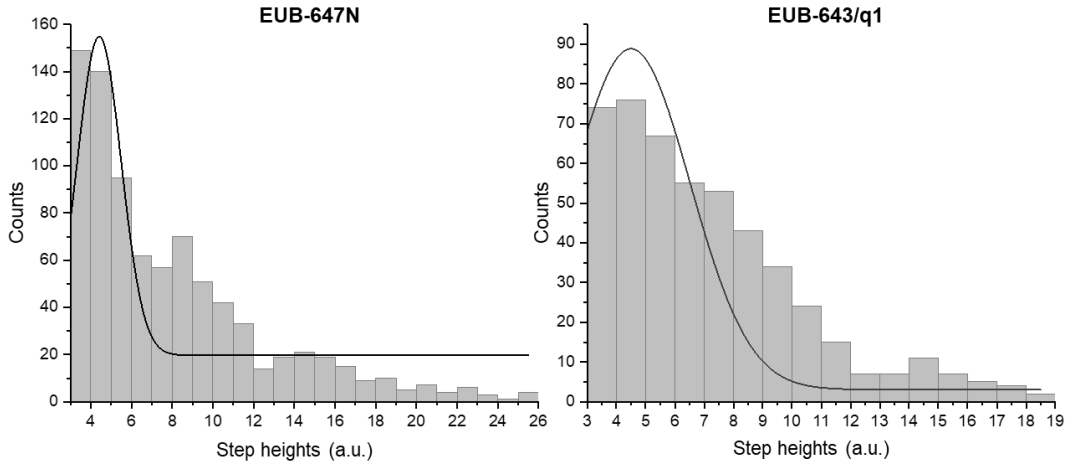
5.1A – Empty cell representative FOVs, per internalisation condition for M9Glu and LB competent wildtype cells without recovery; where 'empty' refers to the incubation of upH_2O to cells instead of probe solution, prior to stimulation/washing. Scale bar – 3 µm.



5.1B – Empty cell representative FOVs, per internalisation condition for M9Glu and LB competent wildtype cells with recovery; where 'empty' refers to the incubation of upH_2O to cells instead of probe solution, prior to stimulation/recovery/washing. Scale bar – 3 µm.

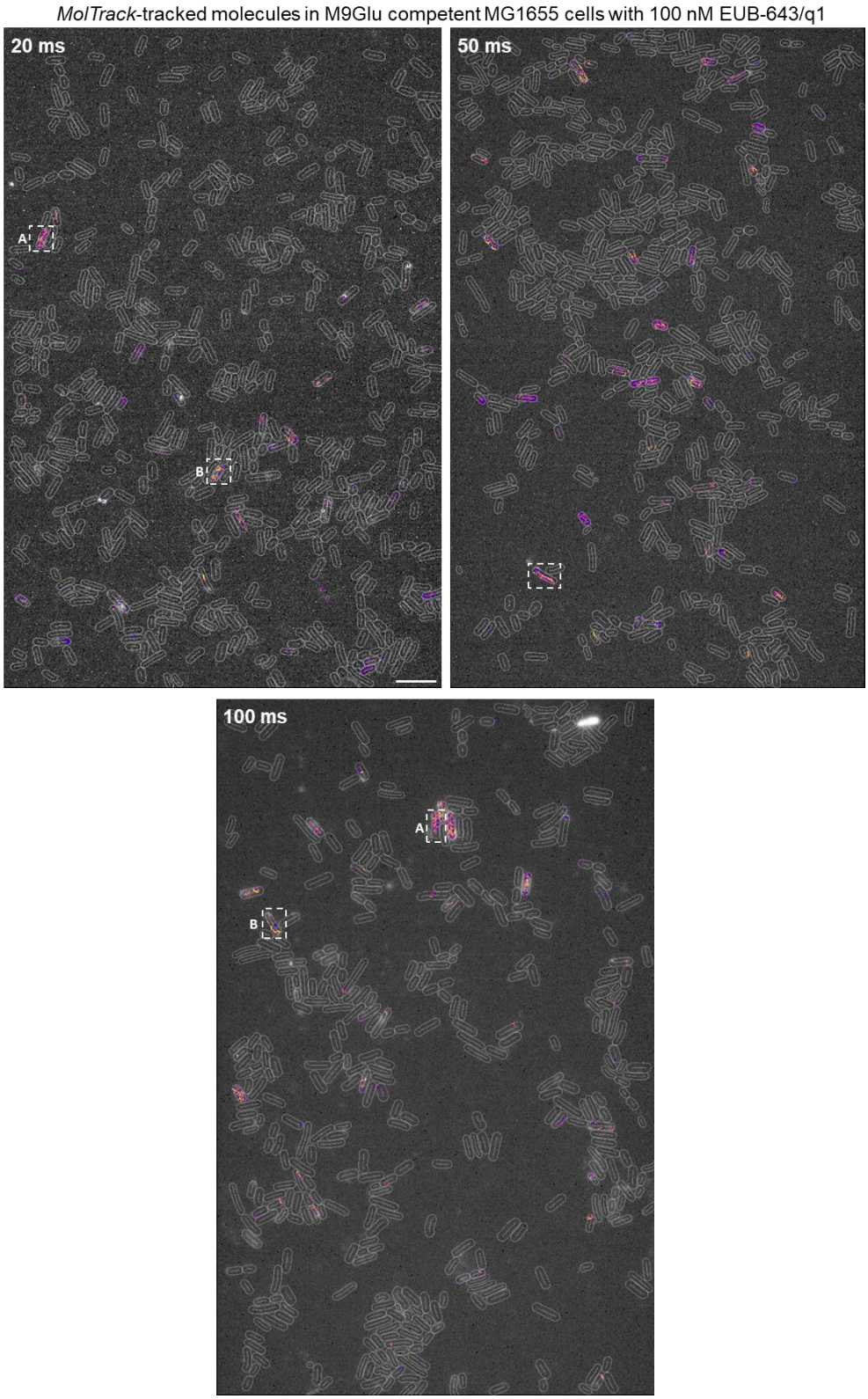


5.2 *AutoStepFinder* step height histograms for cells with 100 nM EUB-647N and EUB-643/q1, fit with a single Gaussian peak (black).

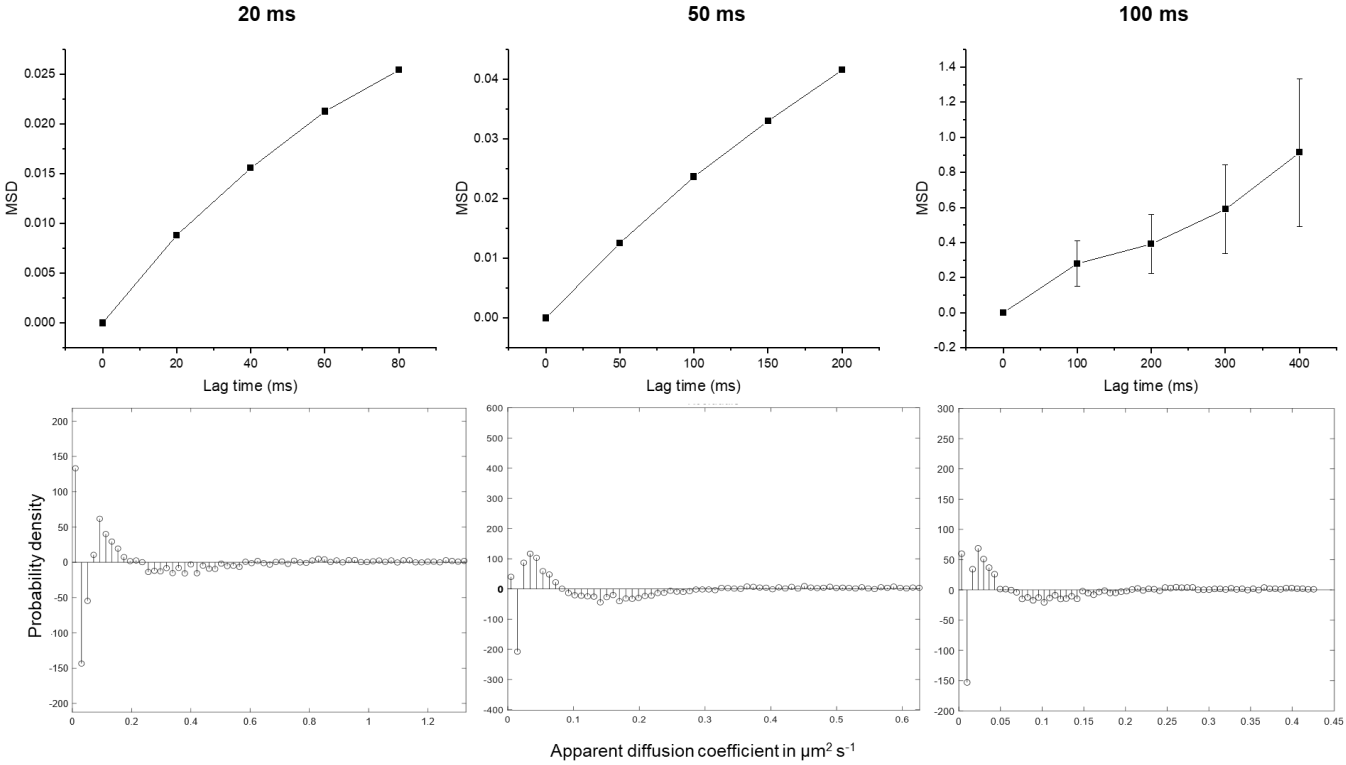


The corresponding step-height distributions for EUB-647N and EUB-643/q1 gave unitary intensities of 4.4 ± 0.2 a.u. (n = 853 fitted steps from 171 traces) and 4.5 ± 0.45 a.u. (n = 488 fitted steps from 179 traces) respectively. However, the poor fitting across the x axis suggests that fitting may show poor accuracy for single bleaching steps – large bleaching steps may suggest 2-3 foci bleaching simultaneously – and that a multiple peak fit may be more appropriate.

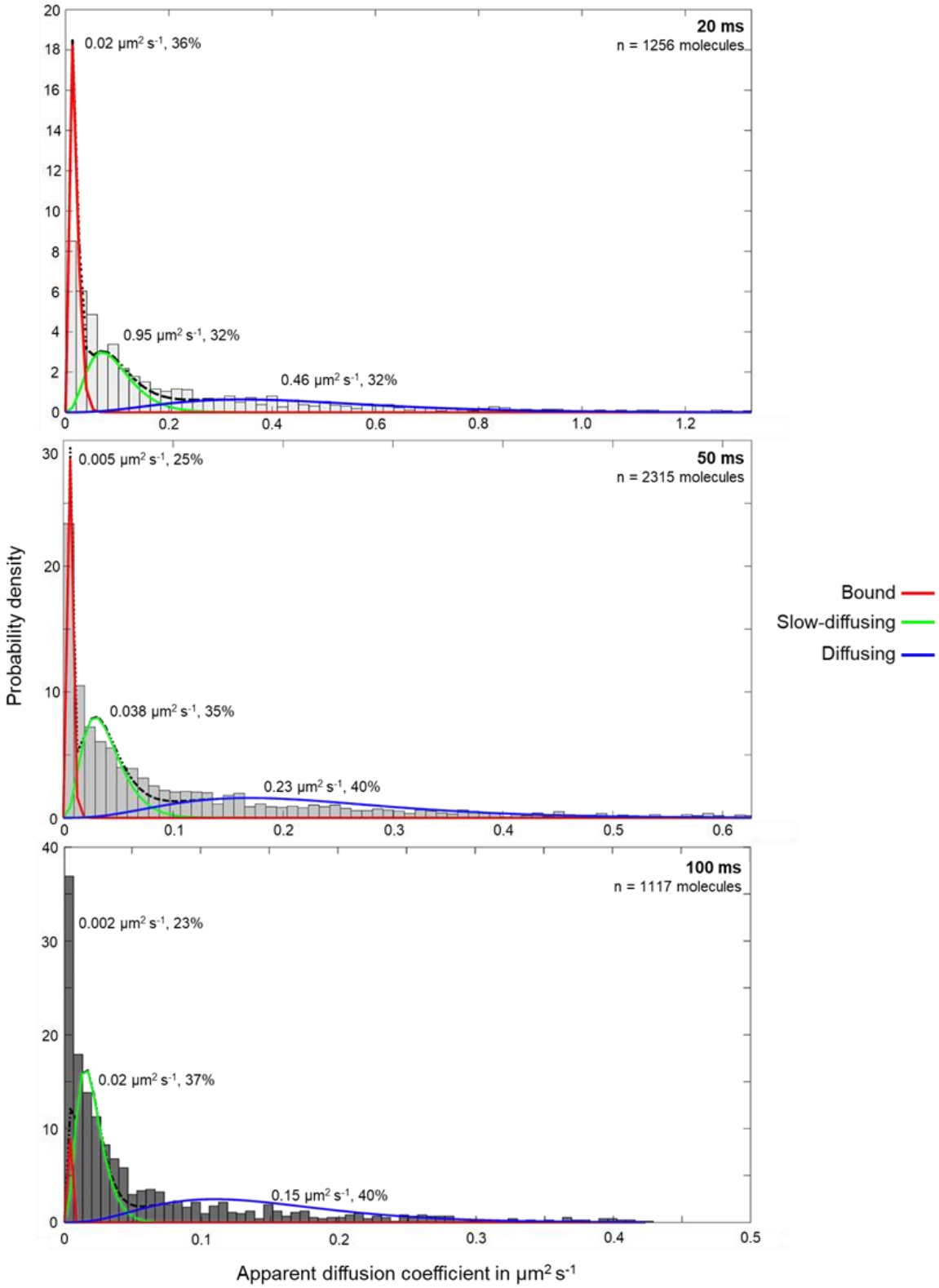
5.3 Moltrack processing of representative FOVs for M9Glu competent wildtype cells with 100 nM passively internalised EUB-643/q1 imaged at 20, 50 and 100 ms exposure. All trajectories captured in cells are shown, coloured by speed (yellow fastest); labelled boxes denote cells discussed. Scale bar – 5 µm.



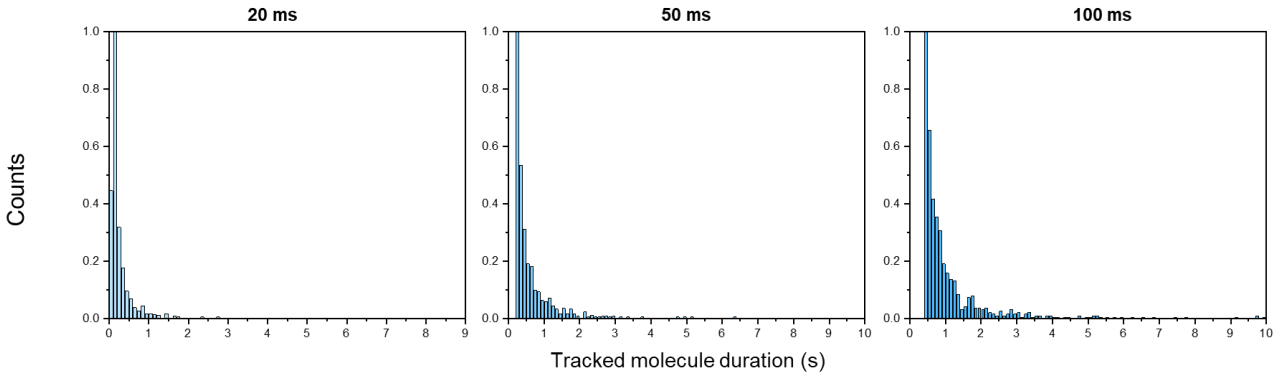
5.5 For each apparent diffusion distribution for cells imaged at 20 ms, 50 ms and 100 ms exposure – MSD by time (±SEM)(A) the variability of localisations used to determine the apparent diffusion coefficients; residual plots (B) show the accuracy of curve fittings.



5.6 Apparent diffusion coefficient analyses of the same tracked molecule populations, fit with a three-free species fit. Population fits as bound (red), slow-diffusing (green) and fast-diffusing (blue), with peak diffusion rate and approximate molecule population annotated.

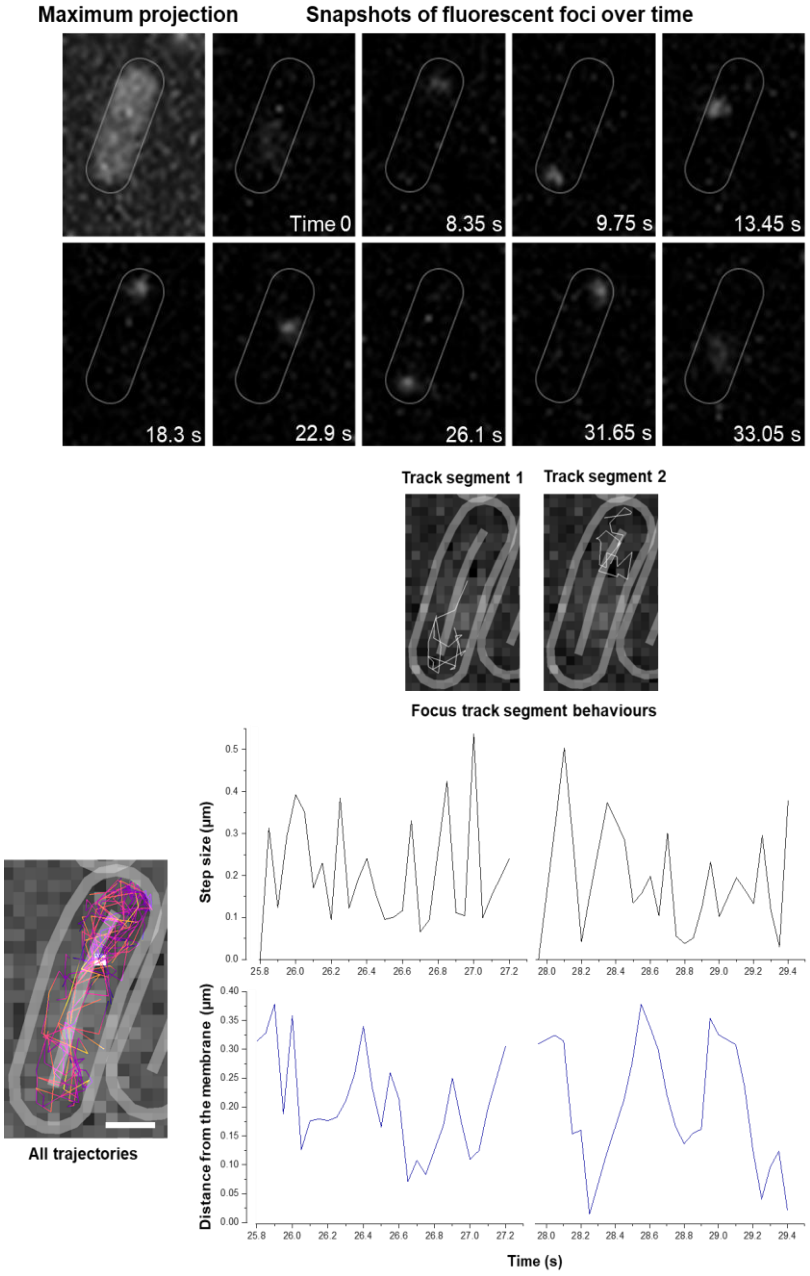


5.7 Full range of normalised tracked molecule durations, bin size = 0.01.



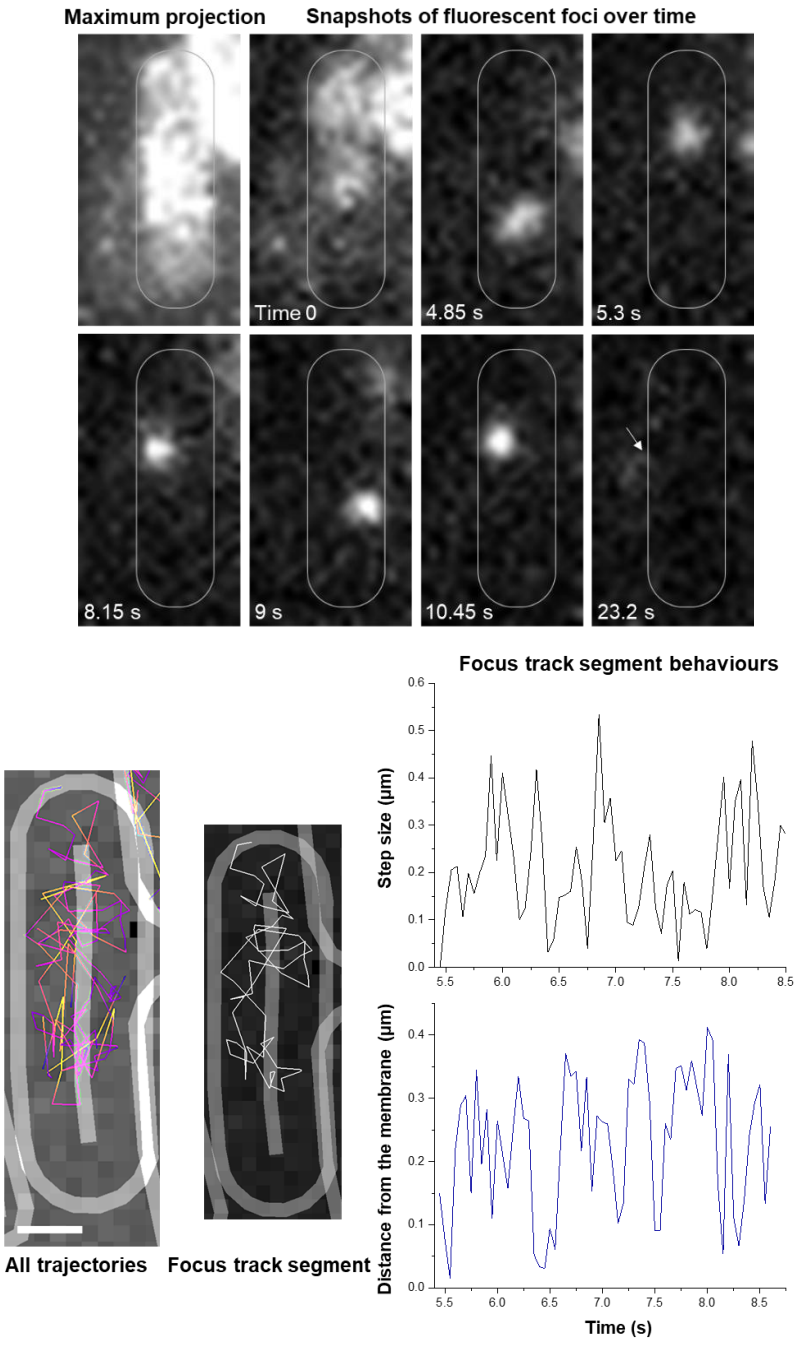
5.8 Rapid diffusion molecule exemplars

5.8A – 20 ms exposure, exemplar cell A



5.8B – 100 ms, exemplar cell A

The exemplar cell shown here is full of diffuse fluorescence at the start of acquisition, but quickly bleaches down to reveal a single focus from around 4 seconds onwards. The bright focus slowly diffuses up and down the midcell and around the upper pole for the next ~7 seconds before disappearing. One more dim, static focus is briefly seen at the midcell at ~23 seconds through the acquisition. The variable diffusion speeds and initially rapid diffusion is suggestive of freely diffusing molecules, perhaps 30S subunits – this corresponds to the large and seemingly random step sizes and membranal distances seen.



Chapter 6 – Exploring single target RNAs in the live cell via short fluorogenic ssDNA probes

6.1 It's the little things in life

The advent and progressive development of single molecule superresolution microscopy, fluorophore chemistry and genetic engineering techniques have revolutionised the capability to visualise the phenomena of cellular life, down to the level of a single copy of RNA. Yet, understanding these phenomena is enduringly complex, and highly dependent on the methodologies applied.

The body of work in this thesis has reproducibly demonstrated the ability to tag single RNAs in the live *E. coli* cell through the internalisation of exogenous short fluorogenic ssDNA probes, either passively or via electroporation as a keystone of induced internalisation. Target RNAs may be non-native, or indeed may be native RNAs site-specifically engineered to incorporate the complementary probe binding array: opening a tractable and exciting opportunity to study endogenous RNAs in *E. coli* with minimal perturbation to the cell. Notably, genetic engineering of the probe binding array enables any cloning-tractable gene target to be selected, whilst internalisation of probes that are specific to endogenous RNAs without any cloning is also achievable.

Detailed technical experimentation confirmed the robustness of a well-characterised method for electroporation-mediated internalisation of labelled biomolecules. Further experimentation highlighted the importance of growth media and competency procedures in increasing overall cellular permeability, leading to the ability of viable competent cells to passively internalise sufficient fluorogenic probes for single molecule experimentation. To this end, molecular observations corresponding to perceived probe-target RNA complexes have been characterised for two non-native RNAs, the native mRNA from the *rpoBC* transcriptional unit, and the 16S rRNA of the 30S ribosomal subunit. Further progress in improving methodological specificity could lead to an openly-accessible and widely applicable methodology for the study of target RNA behaviours in the context of the live *Escherichia coli* cell.

6.2 All the questions, and not so many answers

Much of the experimentation and discussion in this thesis has underscored time and time again the level to which a panoply of fundamental parameters affects even the most basic of observations and implications.

It is probable that the remarkable representative molecular behaviours for fluorescent foci in Chapter 4 with the 8mer-643/q1 probes are specific signals – wildtype bright foci were simply not visible on the same timescale, unless outlining the membrane, and with simpler focal and step characteristics, suggesting that off-target signals would be unlikely to arise like those seen in the tagged cells. Taking therefore the behaviours as "real" – there are some fascinating observations of regional subdiffusive properties, particularly at the midcell and poles, all membrane-peripheral.

There has been much debate about the patterns of RNA spatiotemporal localization – several accord membrane-proximal and often polar localisation of mRNAs ready for translation and protein function (Benhalevy et al., 2015; Moffitt et al., 2016; Nevo-Dinur et al., 2011; Sattler & Graumann, 2021; Stracy et al., 2015); others, that mRNAs localize near their site of transcription (limited diffusion)(Kannaiah et al., 2019; Montero Llopis et al., 2010), and others again, that mRNAs may diffuse through the bacterial nucleoid relatively freely (Bakshi et al., 2015). Noted, many of these studies use different model prokaryotes, and different systems for visualizing single mRNAs.

From the molecular observations here, such exposition of membrane-proximal signals aligns with reports observing translation-independent mRNA targeting to the membrane (using the MS2-PP7 system)(Nevo-Dinur et al., 2011) and spatiotemporal patterning of mRNAs along the cell periphery (using an aptamer system, plasmid-mediated)(Valencia-Burton et al., 2009). The subdiffusive behaviours of chromosomal mRNAs, portraying diffusion that could cover the length of the cell, but generally owing slower-diffusing species likely due to extensive length and molecular weight, appears to be in accordance with the chromosomal MS2 observations in Gram-negative *Caulobacter crescentus* (Montero Llopis et al., 2010) and Gram-positive *Bacillis subtilis* (Sattler & Graumann, 2021).

mRNA visuals are few and far between in single live cells, particularly in bacteria where chromosomal mRNA copy numbers are low, and bulky labelling constructs perturb native functions. Considering the apparent accordance with existing literature, towards mRNA subcellular diffusive localisation and molecular behaviours, it seems that the exchange of engineered, bulkier constructs (MS2-FP, fluorogenic aptamer systems) with short hybridising labelled ssDNAs may be tractable. One key noteworthy caveat here: since the other cited methods involve endogenous FP expression, the cells studied would not experience the electroporation-mediated membrane and likely nucleoidal disruption that is experienced with the work presented in this thesis.

As increasingly evident, the transcriptomic world is wild and complex, with particular involvement of the membrane. Compromising the membrane, hence, renders these observations less physiologically relevant; and little is experimentally validated about nucleoid stability after electroporation, nor indeed much about the impact of electroporation generally, on cellular processes (Ho & Mittal, 1996). Considering chromosomal mRNAs under native promoters are of prominent interest, and within apparent reach here, establishing a take with passive probe internalisation may open further opportunities for extending observations into physiological closeness.

Certainly, with improvements in fluorophore technologies, rational design for hybridisation specificity – towards diminishing imaging artefacts and obtaining clear and specific molecular tracks – and accounting for the core impacts of cellular growth state, engineered *in vivo* smFISH present a promising avenue for reliable native mRNA tracking in real time.

6.3 Future progress – endless parameters, boundless opportunities?

6.3.1 Multi-pronged approaches

There are several low-hanging fruit for progress to begin with, from the outlooks of understanding growth metrics for end-observations, and for the continuation of ongoing projects. A number of these are briefly documented:

Omics and in silico modelling in understanding cellular phenotypes

From the perspective of understanding the determinants of growth media, competency procedures and growth rates on observing single mRNAs, applying existing knowledge from omics studies is likely to give great predictive insight into cell phenotypic variation consequent to growth.

The compendium of RNA-seq data for MG1655 *E. coli* grown in different media by Palsson and co. is an ideal starting point for the comparison of media on the cell transcriptome (Lamoureux et al., 2023; Sastry et al., 2019), particularly by way of implications on the structure and composition of the cell membrane. There seems to be extremely limited reference to competency procedures in general, hence performing an in-house RNA-seq study to gather the transcriptomic profiling of cells post-competency would be very useful in determining adaptive impact. RNA-seq could also offer a greater understanding of passive internalisation: whether in the process of making cells competent, RNAs relating to membrane composition and structural integrity show different expression levels, and perhaps further, which

expression networks are up- and down-regulated during changes in cell competency, growth medium-dependent.

Adjunct to RNA-seq, there are further existing data from genome wide expression profiling, ChIP-seq (Cho et al., 2009; Großmann et al., 2017) and RNET-seq (Yakhnin et al., 2023), all which contribute to the understanding of global RNA elongation and decay rates. These serve to expand the appraisal of transcription at specific sites, and are very helpful in determining candidates of interest and correlative observations. Additional *in silico* analyses of experimental data, such as Kalwarczyk et al.'s methodology for the determination of *in vivo* protein diffusion coefficients (Kalwarczyk et al., 2012), could be applied to generate predictive baselines for smFISH live cell observations.

In the meantime, higher control of growth parameters – such as using continuous bulk culturing via turbidostats (Steel et al., 2020) – could aid to reduce cell-cell variability and generate a reliable starting point across all experiments, whether *in vivo* smFISH or omics-level data acquisitions.

Probe design

The nuances of fluorogenic probe optimisation in the quest for specificity warrant further study. The question of probe specificity is arguably the greatest limitation in interpreting the work presented in this thesis.

There are several possible probe design permutations that could work towards alleviating specificity issues. A molecular beacon probe design as used by Sattler & Graumann are one such possible route for optimisation – utilising a transient stem-loop structure to achieve FRET and de-quenching only by stem-loop opening on complementary target hybridisation. With low autofluorescence and a longer target sequence design possible, there could be in theory fewer off-target binding sites. smFRET probes could similarly be designed to enable dual-colour signal observations on de-quenching – spatial co-localisation in predicted regions could be helpful in understanding transient de-quenching. However, transient half species leading to spurious off-target fluorescence are still possible with both of these methods, as is protein-trapping and unknown aggregation propensities.

By way of improving probe stability and specificity, the application of peptide nucleic acids (PNAs), synthetic DNA mimics with a peptide polymer backbone, may be the best option. With higher characterized binding stabilities above 10 nt and the ability to bind to bidirectionally bind to target nucleic acids, PNAs have also been shown to bind to mRNAs irrespective of

secondary structures – suggesting that PNAs could be a particularly promising prospect for *in vivo* smFISH specific probe-RNA complexation (Pellestor & Paulasova, 2004).

Considering thoughtful probe design, it is noted that the actual binding kinetics of the 8mer probe sequence to its target remains unknown. With the precise control exerted in *in vitro* environments, probe characterisation of on-off kinetics and transitions could be useful in further explaining *in vivo* observations.

The question of a suitable baseline

As with any progressive steps to establishing permutations of methodologies, there are often few appropriate comparators by which to establish a baseline for observations.

For *in vivo* engineered smFISH, one potential baseline as per specific RNA targets could be fixed FISH, to evaluate the spatial distributions of the target RNA (or even multiple targets) at given times (Chatzimichail et al., 2024; K. H. Chen et al., 2015). However, fixed FISH results with transiently binding probes may be difficult to interpret due to poor stabilisation of fluorescent signals, perhaps from free dye in the cell, actively diffusing unbound probes or potentially unfixed RNAs complexed with complementarily bound probes likely diffusion of probes, even after washing of fixed samples. Nucleic acid interactions could be stabilised via the use of nucleic acid fixatives like psoralen; but much optimisation of fixed FISH methodologies would be necessary for readable results.

Similarly, for the case of a given target, since many single mRNA observations have been achieved via aptameric systems, it would be interesting to directly compare observations for the same target across both methods. Of course, the caveats mentioned persist and behavioural comparisons could only be loosely drawn, but could nevertheless aid in understanding method robustness if corroborative observations are made.

Finally, mathematical and *in silico* modelling of target RNA interactions at the levels of the fluorogenic probe and within the cellular environment, as is often characterized in synthetic biology (Sechkar & Steel, 2025), could greatly enhance predictions of observed fluorescent foci prior to experimentation.

Extending native candidate gene observations

As evidenced in discussion, essential genes typically exhibit complex regulation, and there are few singularly well-characterised targets. *lacZ* appears to be well studied through a variety of *in vitro* and *in vivo* systems, and hence may provide a more workable native starting point than

some of the ongoing candidates. Nonetheless, adjunct information from omics-level studies as described above could contribute significantly to understanding and hypothesising current observations *in vivo*.

As for understanding observations from *rpoB*-tag cells, two ongoing attempts at clarifying transcription behaviours are the generation of a dual-colour system for *rpoBC*, and *in vitro* transcription assays using the rpoBp putative promoter. The goal of a C-terminal tag array (6x 15mer), much the same as the 5' UTR 6x 8mer site, would be to try to decode the co-transcription of *rpoBC* – co-localisation of 5' and 3' signal could suggest full mRNA release before 5' degradation. As for the *in vitro* transcription assay – preliminary observations show consistent RNAP initiation and elongation, indicating that rpoBp does indeed seem to act as a promoter; further incorporation of probe-binding dynamics could be explored in tandem.

6.3.2 The bigger picture

From a more holistic perspective, with tighter specificity control and understanding of probe-RNA molecular behaviours, there would be boundless applications, across numerous fields. With the tractability of passive internalisation of short fluorogenic ssDNA probes in M9Glu competent *E. coli* cells, providing the media choice and competency are not confounding as per experimental purpose, mRNA tagging could be applied to observe any number of target mRNA spatiotemporal behaviours in the cell. This could be applied in appraising the mRNA characterisation of antibiotic resistance phenotypes, to mRNA interactions with the degradosome machinery, to exploring the roles and regulation of untranslated RNAs. The combination of omics, *in silico*, genetic engineering and single molecule techniques can be applied for applying correlative observations at both the RNA and protein levels. Taking advantage of the rich tapestry of such techniques would build an exciting bigger picture of the regulation, coordination and dependencies of the most fundamental processes of the cell and beyond.

P.S. There is always more to learn, and even more to reflect on; on coming across (the Father of mRNA) Sydney Brenner's contemplations of Biology and development during this time, so many of this thoughts and words ring true as ever. Inspiring, perhaps a little daunting – and most certainly something to keep in mind for the future of Science.

Brenner, S. (2002). Life sentences: Detective Rummage investigates. *Genome Biology*, *3*(9), comment 1013.1-2. (Epub 2002 Aug 21.)

‘Out beyond ideas of wrongdoing and rightdoing,
there is a field. I’ll meet you there.’

~ Rumi ~

Appendices

Appendix 1 – Optimisation experiments for *in vivo* smFISH: figures

1.1 Electroporation parameters

Probe incubation and cell recovery time

To optimize for increased numbers of cell with internalised probes (as any fluorescent signal), induced p6 cells were incubated with 50 nM 8mer-2x647N for 5 or 10 min on ice, electroporated as standardized, and then recovered for 5 or 20 min in SOC medium at room temperature or 37°C.

5 min probe incubation

At 5 min probe incubation, maximum intensity projection representative FOVs show highest visual loading for cells recovered for 20 min at 37°C (Figure A1.1A). The loaded cell mean intensity per cell area plots likewise corroborated the slightly higher signal observation, with median 17.5 units for cells recovered for 20 min at 37°C, versus 9,7 units (5 min, room temperature), 5.1 units (5 min, 37°C), and 13.8 units (20 min, room temperature) (Figure A1.1B). Interestingly, no difference was seen in the loading classification, where ~78% cells were loaded at 20 min, 37°C, compared to ~77% (5 min, room temperature), ~70% (5 min, 37°C), and ~77% (20 min, room temperature). The cells recovered for 20 min at 37°C also showed the widest spread of cells comprising fluorescent signal, and 37°C recovery temperature more generally appears to be the optimal temperature (Figure A1.1C).

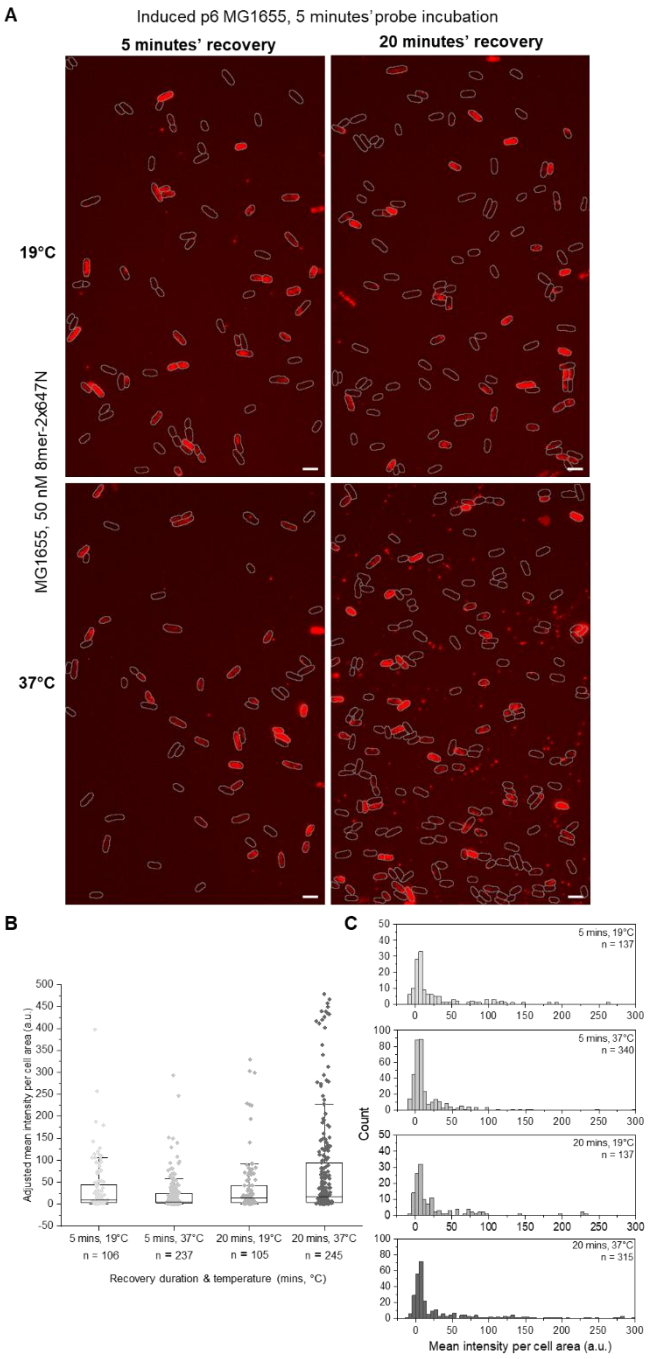


Figure A1.1 – Optimising signal: testing recovery duration and recovery temperature for induced p6 competent M9Glu cells probe-incubated for 5 min prior to electroporation.
(A) Maximum intensity projections of frames 100-2000 of induced p6 MG1655 cells electroporated with 50 nM original 8mer-2x647N probes following a 5-min incubation period, recovered for 5 or 20 min in SOC at either room temperature or 37°C.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 8mer-2x647N-electroporated M9Glu competent wildtype cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Loaded cell medians: 9.7 units 5 min/room temperature, 5.1 units 5 min/ 37°C, 13.8 units 20 min/room temperature, 17.5 units 20 min/ 37°C.
(C) Comparative histogram distributions of the whole cell populations across different recovery durations and temperatures; bin size = 5.
Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from the 100th frame, and n gives the number of cells per sample (N = 1).

10 min probe incubation

At 10 min probe incubation, maximum intensity projection representative FOVs again showed highest visually loaded subpopulations for cells recovered for 20 min (Figure A1.2A). The loaded cell mean intensity per cell area plots show negligible differences between mean intensity populations, with median 10.3 units for cells recovered for 20 min at 37°C, versus 11.2 units (5 min, room temperature), 8.2 units (5 min, 37°C), and 11.1 units (20 min, room temperature) (Figure A1.2B). There was little difference, although slightly higher for 20 min' at room temperature, in the loading classification, where ~78% cells were loaded at 20 min, 37°C, compared to 70% (5 min, room temperature), ~76% (5 min, 37°C), and ~85% (20 min, room temperature). The cells recovered for 20 min at 37°C showed the most consistent wide spread of cells comprising fluorescent signal, with highest counts of all four conditions at 10 min' probe incubation for cells comprising 5-15 units of signal (Figure A1.2C).

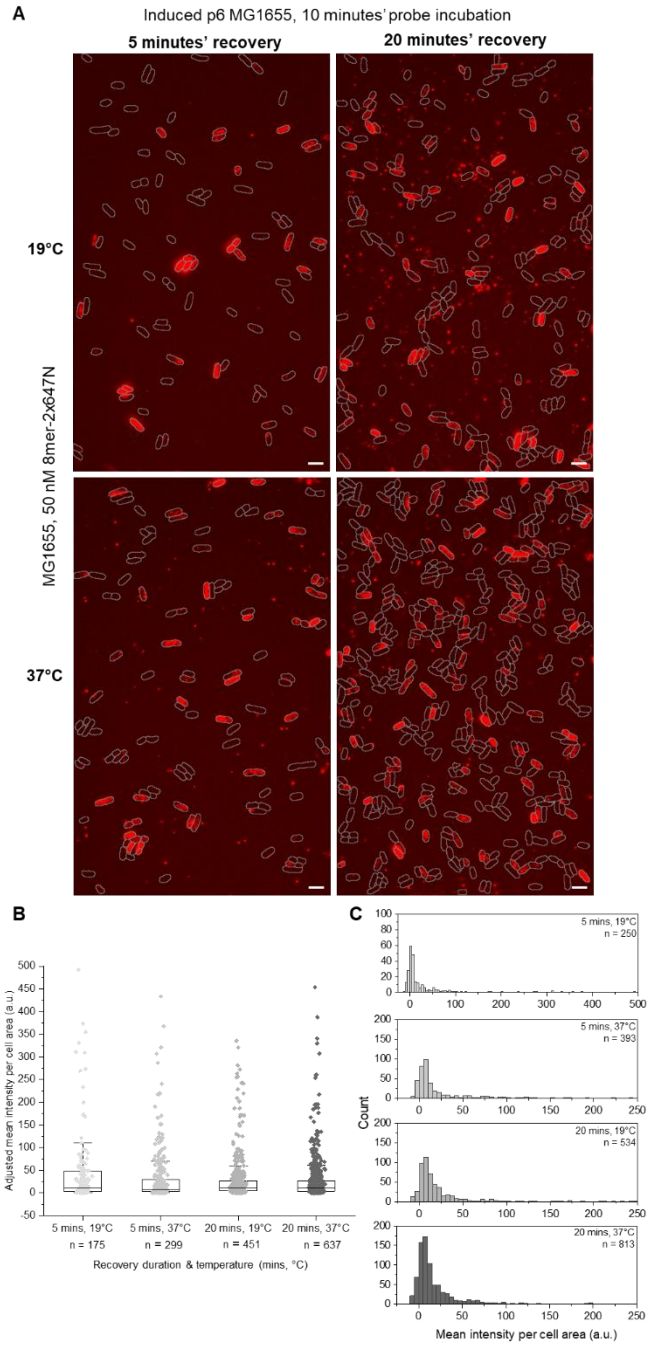


Figure A1.2 – Optimising signal: testing recovery duration and recovery temperature for induced p6 competent M9Glu cells probe-incubated for 10 min prior to electroporation.
(A) Maximum intensity projections of frames 100-2000 of induced p6 MG1655 cells electroporated with 50 nM original 8mer-2x647N probes following a 10-min incubation period, recovered for 5 or 20 min in SOC at either room temperature or 37°C.
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 8mer-2x647N-electroporated M9Glu competent wildtype cells. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Loaded cell medians: 11.2 units 5 min/room temperature, 8.2 units 5 min/ 37°C, 11.1 units 20 min/room temperature, 10.3 units 20 min/ 37°C.
(C) Comparative histogram distributions of the whole cell populations across different recovery durations and temperatures; bin size = 5.
Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from the 100th frame, and n gives the number of cells per sample (N = 1).

Recovery media

To explore whether probe internalisation efficiency and the proportion of cells comprising fluorescent signal could be improved by selecting for a particular recovery medium, induced p6 cells were recovered in either M9Glu (poorest), SOC (rich) or RDM (richest) for the standardized 20 min.

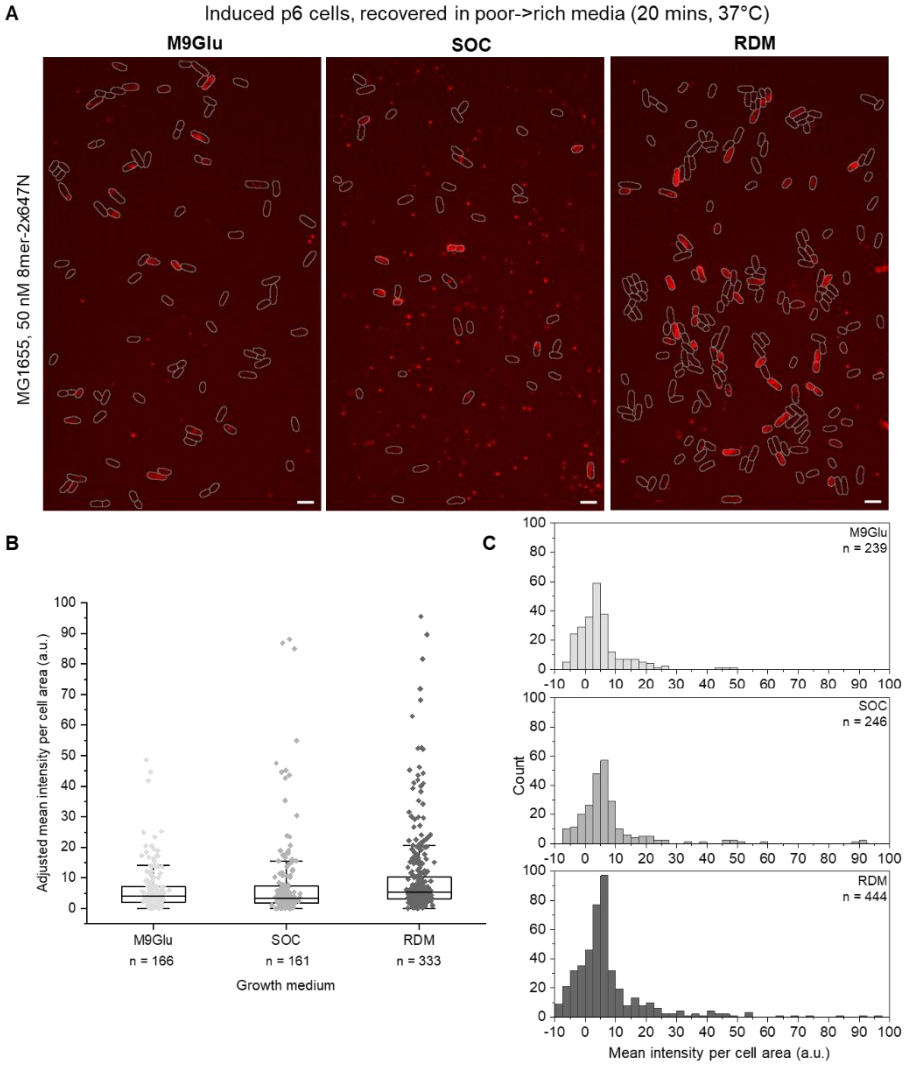


Figure A1.3 – Optimising signal: testing recovery media for induced p6 competent M9Glu cells post-50 nM 8mer-2x647N electroporation.
(A) Maximum intensity projections of frames 100-2000 of induced p6 MG1655 cells electroporated with 50 nM original 8mer-2x647N probes and recovered for 20 min, 37°C, in either M9Glu, SOC or RDM media (poorest->richest).
(B) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for 8mer-2x647N-electroporated M9Glu competent p6 cells with different recovery media. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of upH_2O-electroporated (empty) cells.
Loaded cell medians: 3.9 units M9Glu, 3.4 units SOC, 5.3 units RDM.
(C) Comparative histogram distributions of the whole cell populations across different recovery durations and temperatures; bin size = 2.
Scale bars – 3 µm , false-colour in red and overlaid cell segmentation outlines; all mean intensities per cell area shown are background corrected, taken from the 100th frame, and n gives the number of cells per sample (N = 1).

It would be expected that the richer the medium, the better the cell fitness and growth propensity, and hence recovery post-electroporation – which is the observed trend in maximum intensity projections of representative FOVs (Figure A1.3A). Per sample, classed loaded subpopulations were ~70% for M9Glu, ~66% for SOC and 75% for RDM-mediated recovery. Although negligible differences in the mean intensities of the loaded cells overall (medians: 3.9 units M9Glu, 3.4 units SOC, 5.3 units RDM)(Figure A1.3B), the whole cell distributions corroborate that RDM-recovered cells had the widest and most consistent spread of signal (~250 cells with 0-10 units, versus ~170 cells for SOC and ~150 cells for M9Glu)(Figure A1.3C).

1.2 Microscopy parameters

Laser power

To explore the compromise between resolution and bleaching propensity, induced p6 cells with 50 nM 8mer-2x647N were imaged at 50 ms exposure, at 5%, 8% or 10% laser powers on the ONI NanoImager, corresponding to ~2, 4 and 10 mW. Maximum intensity projections of representative FOVs demonstrate that at 5%, foci were blurry and insufficiently focused (Figure A1.4) an 8% power may be a good compromise, but there seemed little difference consistently between 8-10% laser powers on the NIM, and with the non-linearity of the laser activation systems, 10% seemed the most robust choice for consistency in laser illumination of the samples, despite the higher propensity for rapid bleach down.

Notably, the use of high laser powers for rapid bleach-down of off-target/spurious fluorescence could also be used advantageously: one touted method of rapidly achieving single molecule resolution, appropriate for individual analysis, was to perform a high intensity bleach-down at 30% laser power (~25 mW) for ~30 seconds, and then acquire as normal at 10%. Although quickly substituted for optimised probes, initial attempts displayed feasible recurrent fluorescent signal after high-intensity bleach-down, suggesting that quenching of the probes enabled a reserved pool in the cell milieu, available for de-quenching on transient hybridisation should a target mRNA be transcribed and released.

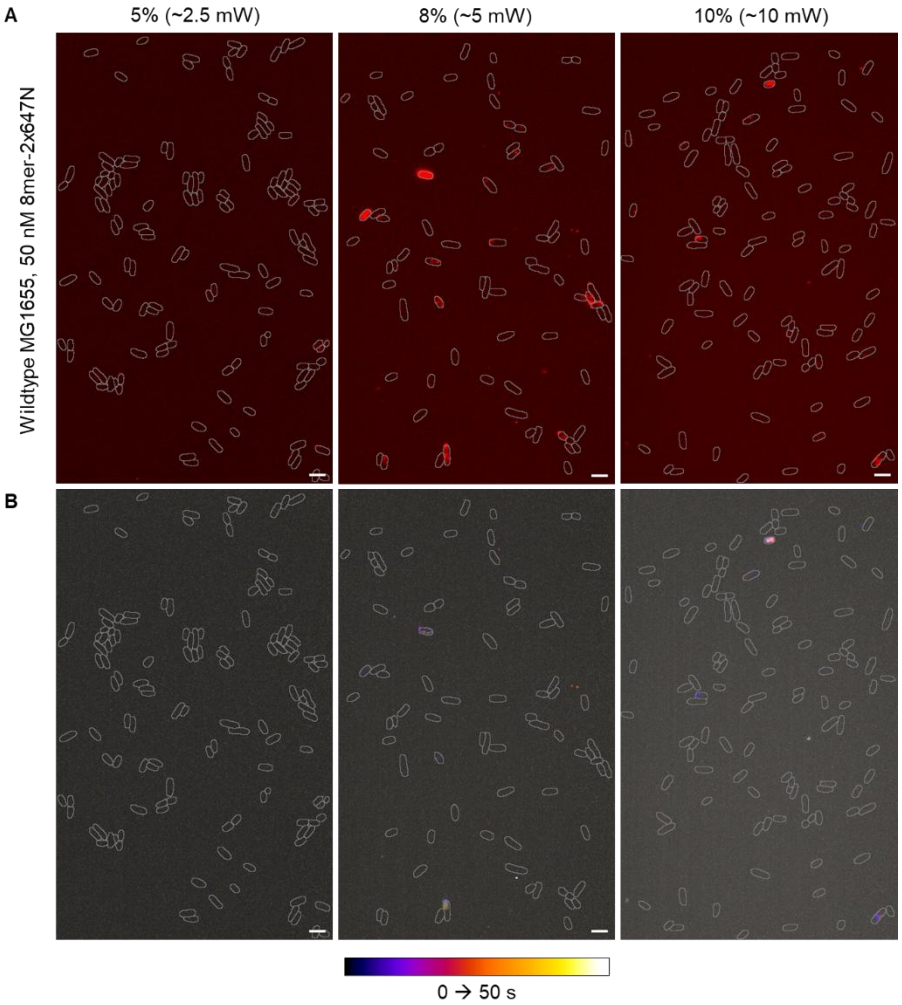


Figure A1.4 – Testing laser power for optimal resolution.
(A) Maximum intensity projections of 100-1000 frames (~45 s) representative FOVs of wildtype MG1655 M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes, imaged at 5%, 8% and 10% 638 nm laser powers.
(B) Corresponding temporal projections of 1-1000 frames; the colour bar denotes time, where darkest is time 0 (50 ms exposure, 50 s acquisition).
Scale bars – 3 μm, false-colour in red and overlaid cell segmentation outlines.

Dark interval acquisition as a proxy for stroboscopy

A representative FOV for an induced gT7 sample with 50 nM 8mer-2x647N with 50:250 ms ON:OFF dark interval illumination (Figure A1.5A) compared to equivalent cells with constant laser activation (Figure A1.5B) depicts the level of signal retained at an equivalent 90-100 second timepoints. Corresponding temporal projections (righthand side) indicate molecule durations seen through the acquisition period. The dark interval acquisition does appear to give much better photostability for the same equivalent durations, showing that photon budget can be suitably increased for longer acquisitions through this modality.

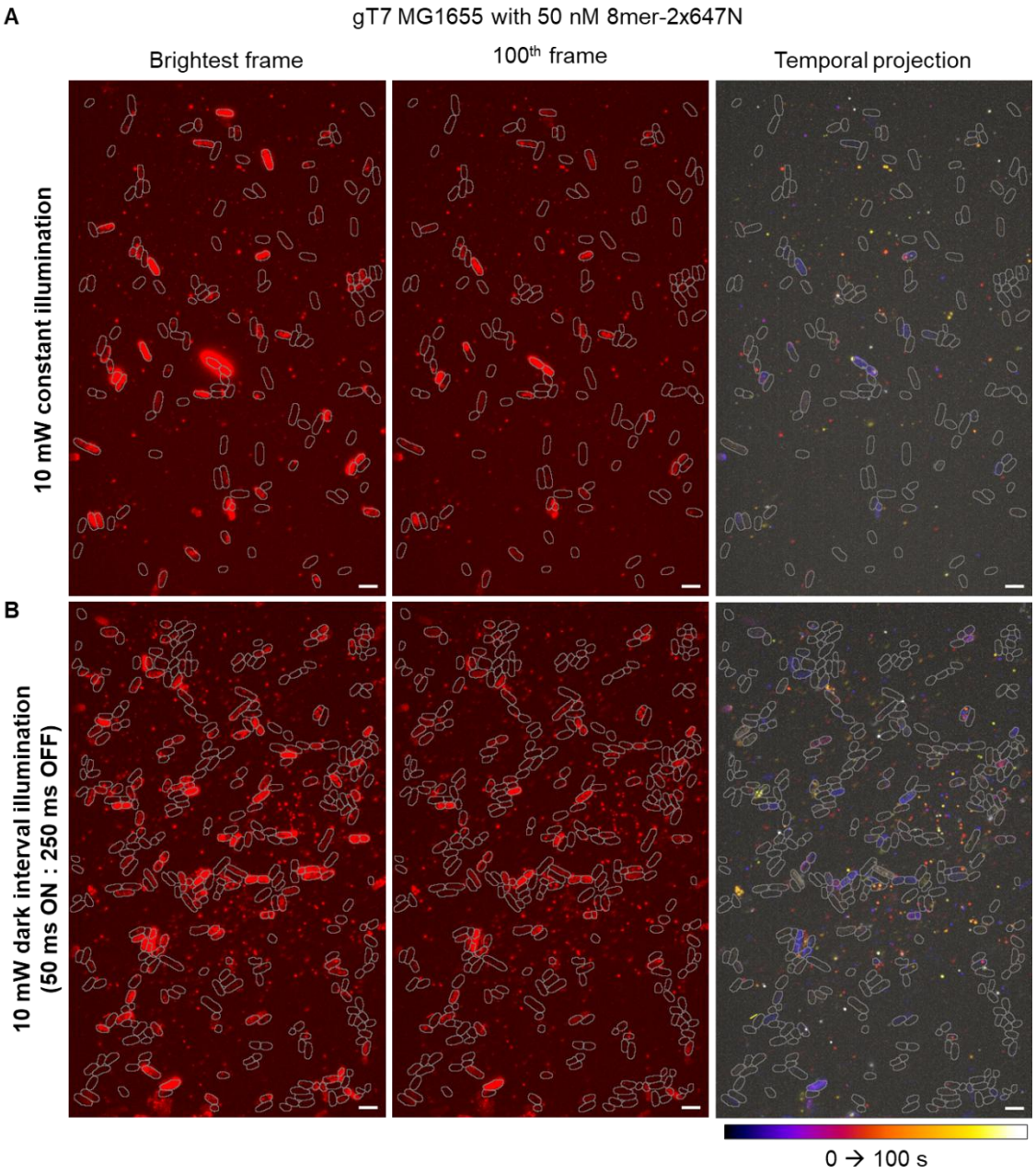


Figure A1.5 – Exploring dark interval acquisitions as an alternative to stroboscopy.
(A) Maximum intensity projections of 1-2000 frames (brightest, ~100s) and 100-2000 frames (~90 s) and corresponding temporal projection of a representative FOV of induced gT7 MG1655 M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes, imaged at 10% 638 nm constant laser excitation.
(B) Maximum intensity projections of 1-2000 frames (brightest, ~100s) and 100-2000 frames (~90 s) and corresponding temporal projection of a representative FOV of induced gT7 MG1655 M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes, imaged with 10% 638 nm constant laser excitation for 50 ms every 250 ms (1:5 excitation:dark).
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines; for temporal projections, the colour bar denotes time, where darkest is time 0 (50 ms exposure, 100 s acquisition).

Exposure time

Figure A1.6A presents representative FOVs for *rpoB*-tag system cells with 50 nM 8mer-2x647N, imaged for ~90 seconds at 50 ms, 75 ms and 100 ms exposure times alongside corresponding temporal projections (A1.6B). All are similar, but the 100 ms temporal projection shows more cells with fluorescence later into acquisition.

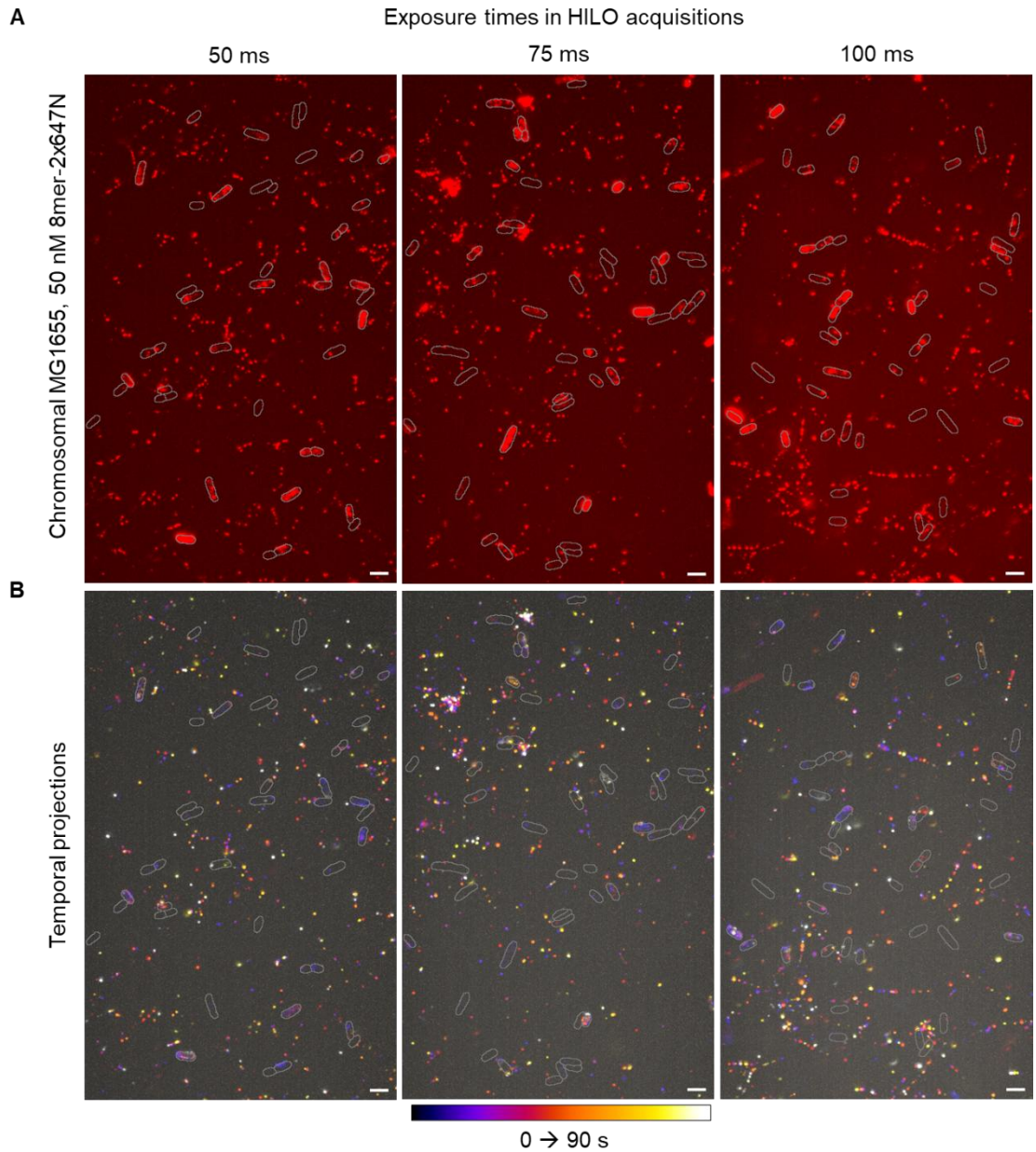


Figure A1.6 – Testing different exposure times in optimising resolution.
(A) Maximum intensity projections of 100-2000 frames (~90 s) representative FOVs of chromosomal system MG1655 M9Glu competent cells electroporated with 50 nM 8mer-2x647N probes, imaged at 10% 638 nm constant laser excitation with 50 ms, 75 ms and 100 ms exposure.
(B) Corresponding temporal projections of 1-2000 frames; the colour bar denotes time, where darkest is time 0 (50 ms exposure, 100 s acquisition).
Scale bars – 3 µm, false-colour in red and overlaid cell segmentation outlines.

1.3 Passive internalisation of labelled ssDNAs: figures for technical experiments

Establishing a probe for parameter comparison

Maximum intensity projections of the first 5 seconds (100 frames)[11] of representative FOVS of M9Glu and LB competent cells incubated with 100 nM of 8mer-2x647N, EUB-647N or EUB3-Cy3B (Figure A1.7). M9Glu cells showed between 50-90% visual internalisation of probes; ~10% LB cells only showed apparent internalisation.

Comparing the mean intensity per cell area from the brightest frame across all samples, there appeared to be little difference between the median mean intensity per cell area of the 647N probes but higher intensities for Cy3B in M9Glu samples (medians: 3.7 units 8mer-647N, 5.0 units EUB-647N, 11.0 units EUB-Cy3B), nor with LB samples with the exception of the 8mer-2x647N, somewhat unsurprisingly (medians: 33.4 units 8mer-647N, 5.9 units EUB-647N, 8.2 units EUB-Cy3B) (Figure A1.8A). M9Glu cell populations classified as loaded corresponded to ~77% for 8mer-2x647N-internalised cells, ~68% for EUB338-647N-internalised cells and ~88% EUB338-Cy3B-internalised cells. Interestingly, LB cell populations classified as loaded were also very high (~74% 8mer-2x647N, ~88% EUB338-647N, ~86% EUB338-Cy3B), but this is potentially caveated by spurious extracellular signals confounding the mean intensity calculations.

[11] Note that for passive internalisation experiments, optimised for the singly-labelled, 18-nt EUB probes, high photonic signal at the immediate activation of the laser during acquisition was less common, and therefore the first 5s of all acquisitions are used comparatively, and only brightest frame mean intensities calculated as a comparator for condition-dependent passive internalisation of probes.

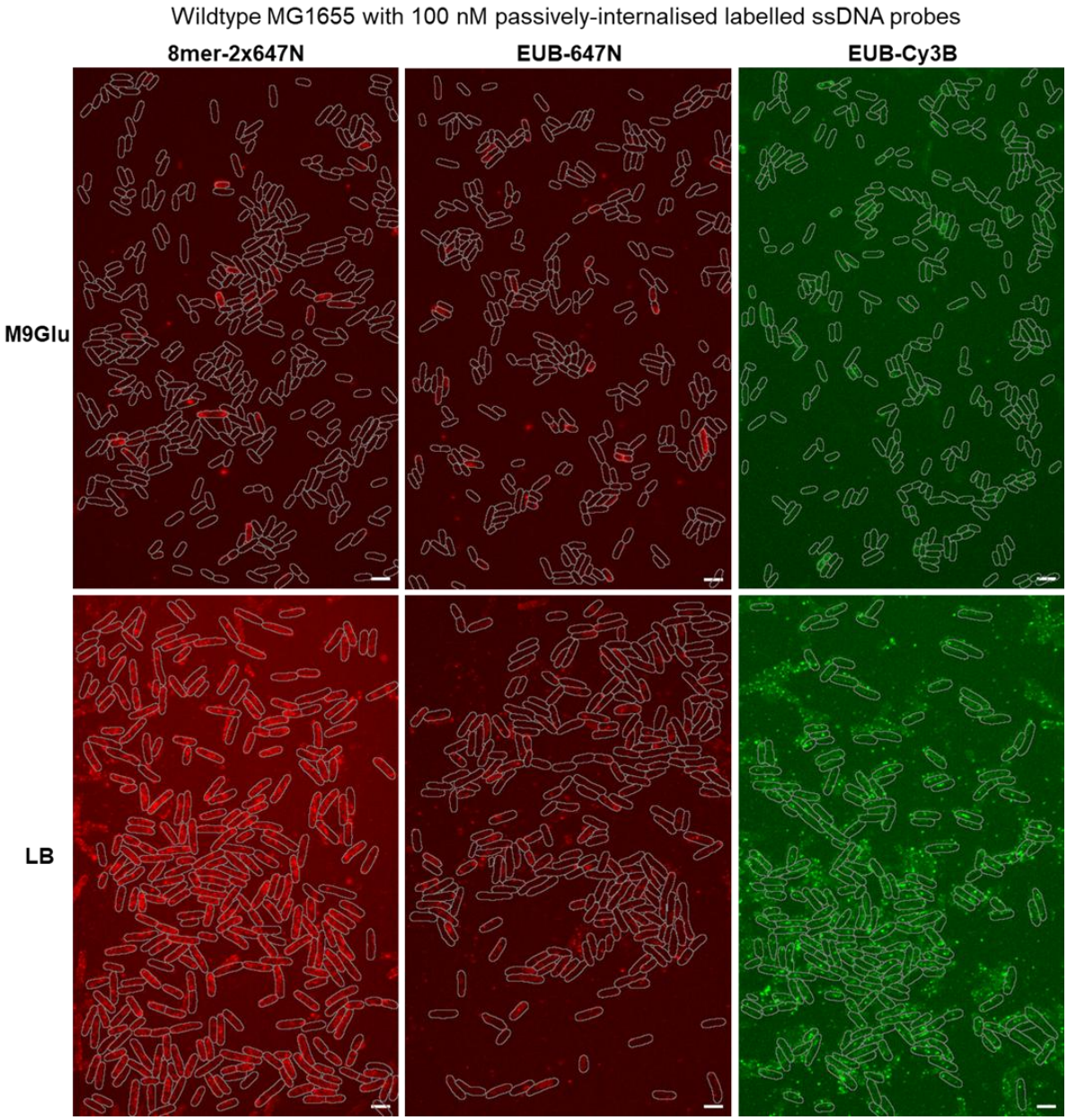


Figure A1.7 – Testing the passive internalisation of different labelled ssDNA probes in M9Gluand LB competent wildtype MG1655 cells.
Maximum intensity projections of 1-2000 frames (~100 s) representative FOVs of wildtype MG1655 competent cells incubated with 100 nM 8mer-2x647N (original), EUB-1x647N and EUB-1xCy3B probes, imaged at 10% 638 nm constant laser excitation. Top panel shows M9Glu competent cells, and the lower panel shows LB competent cells.
Scale bars – 3 µm, false-colour in red for 647N-labelled probes, green for Cy3B probes and overlaid cell segmentation outlines.

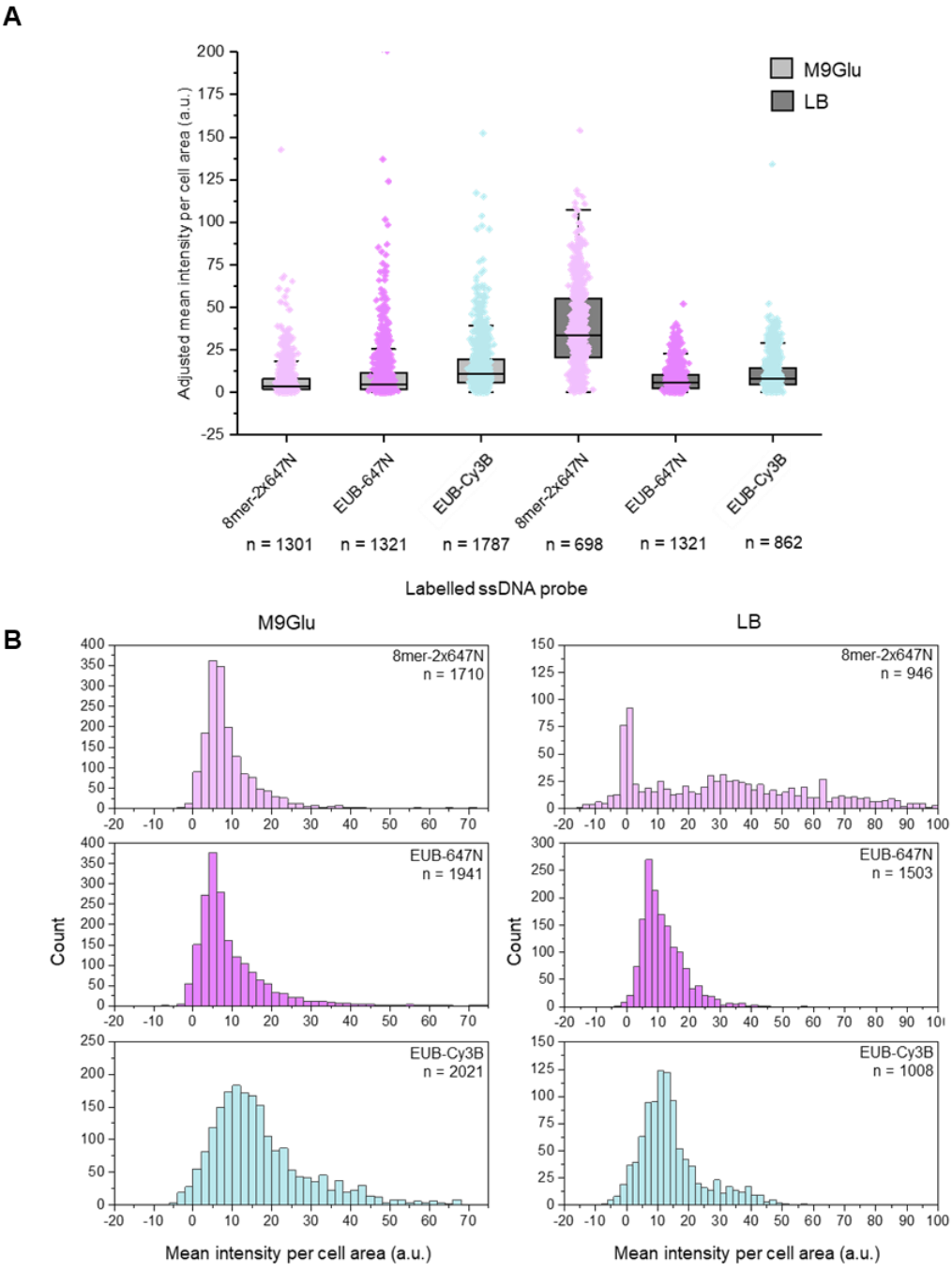


Figure A1.8 – Testing the passive internalisation of different labelled ssDNA probes in M9Glu and LB competent wildtype MG1655 cells: population-level intensity distributions.
(A) Loaded cell mean intensities per cell area with corresponding boxplots (25%, median and 75% quartiles with outlier + 1.5 whiskers) for M9Glu and LB competent wildtype cells with 100 nM passively internalised 8mer-2x647N, EUB-1x647N and EUB-1xCy3B. Cell classification as 'loaded' defined as mean intensity above the mean intensity value of empty cells.
Loaded cell medians: M9Glu: 3.7 units 8mer-2x647N, 5.0 units EUB-1x647N , 11.0 units EUB-1xCy3B; LB: 33.4 units 8mer-2x647N, 5.9 units EUB-1x647N , 8.2 units EUB-1xCy3B.
(C) Comparative histogram distributions of the whole cell populations with different probes for M9Glu and LB competent cells; bin size = 2.
n gives the number of cells per sample (N = 2 experiments).

The EUB-647N probe gave the tightest distribution of cell intensities and the lowest negative cell populations (empty cells) across both growth media cell samples (Figure A1.8B). A tighter distribution is helpful in gauging passive internalisation of fluorescent probes, since further changes in intensity distributions depicted as histograms could be more analogous to condition-dependent effects on cell phenotypic heterogeneity, as opposed to differences due to the probe itself.

Passive internalisation – cell viability

To test the cell viability of competent M9Glu wildtype cells with 100 nM passively-internalised EUB-647N probe, brightfield imaging of cells was performed from 0 to 60 min, at 15 min intervals. Cell phenotypes were classified as growing and divided (red), in stasis (orange), and damaged (black), by comparing the time 0 FOV with the 60 min representative FOV (Figure A1.9). Comparative to electroporated cell observations, over 90% cells appeared to be growing and dividing, although notably, LB grown cells were much harder to stably characterise.

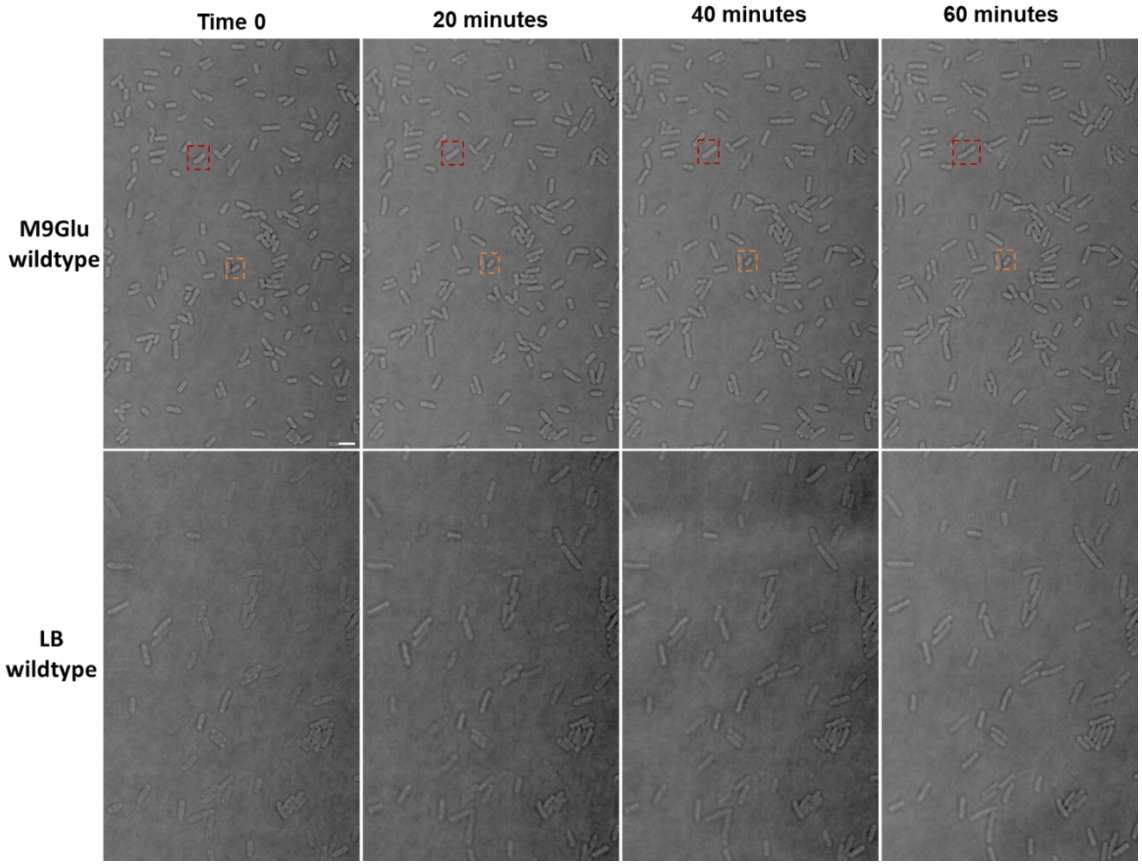


Figure A1.9 – Observing electroporated cell fitness by single cell growth time series.
Representative FOVs of time points, 0, 20, 40 and 60 min for M9Glu and LB competent wildtype cells with 100 nM passively internalised EUB-647N probes.
For M9Glu cells, growing/dividing cells are highlighted in red, static in orange and compromised in black; scale bar – 3 µm.

Appendix 2 – Molecule distributions, behaviours and visuals in MG1655 *E. coli* with 50 nM 8mer-2x647N

Comparing tag strain RNA-probe complexes with the original 8mer-2x647N

Using the original 8mer-2x647N probes as a baseline, it was noted that the capacity to interrogate specific RNA-probe complexes would become more apparent at the level of the single cell, individual events and focal molecular characteristics.

2.1 Evaluating the photostability and brightness of 8mer-2x647N across p6, gT7 and *rpoB*-tag strains

To determine the photostability of 50 nM 8mer-2x647N electroporated into induced p6, gT7 and *rpoB*-tag cells, single cell photo-bleaching studies were conducted.

The temporal projections in Figure A2.1A depict that much brighter cells sustaining multiple foci over a much longer period than for cells with distinguishable foci from the start, as would be anticipated with a higher available number of probes for possible binding. To obtain representative photo-bleaching time traces for each strain, over 200 visually-loaded single cells were taken, the per frame mean intensity per cell area measured, and an average loaded cell trace calculated (background corrections from single/double exponential fits to give a y0 baseline). Figure A2.1B presents the average lifetime trace (in black) per strain, overlaid onto the average time trace plot are two further exemplar single cell traces; the half-time fluorescence decay for de-quenched 8mer-2x647N can hence be approximated to ~10 seconds. Most signals across the 3 strains appear to bleach by ~30 seconds; individual cells may then exhibit distinguishable events later in acquisition. These >30 second signals appear more prevalent in the rpoB-tag cells particularly, indicative of a longer mRNA present for longer timescales in the cell – and indicative of real signal.

To evaluate the brightnesses of the same of 8mer-2x647N across all tag strains, visually-identified loaded single cell time traces were fitted using step height analysis via *AutoStepFinder* (Loeff et al., 2021).

From 145 traces for p6, 88 traces for gT7 and 208 single cell time traces for *rpoB*-tag loaded cells with 50 nM 8mer-2x647N probes, 111, 74 and 154 time traces were fitted for p6, gT7 and *rpoB*-tag cells respectively. Step height histograms were then manually thresholded for step heights measurements between 3-30 units.

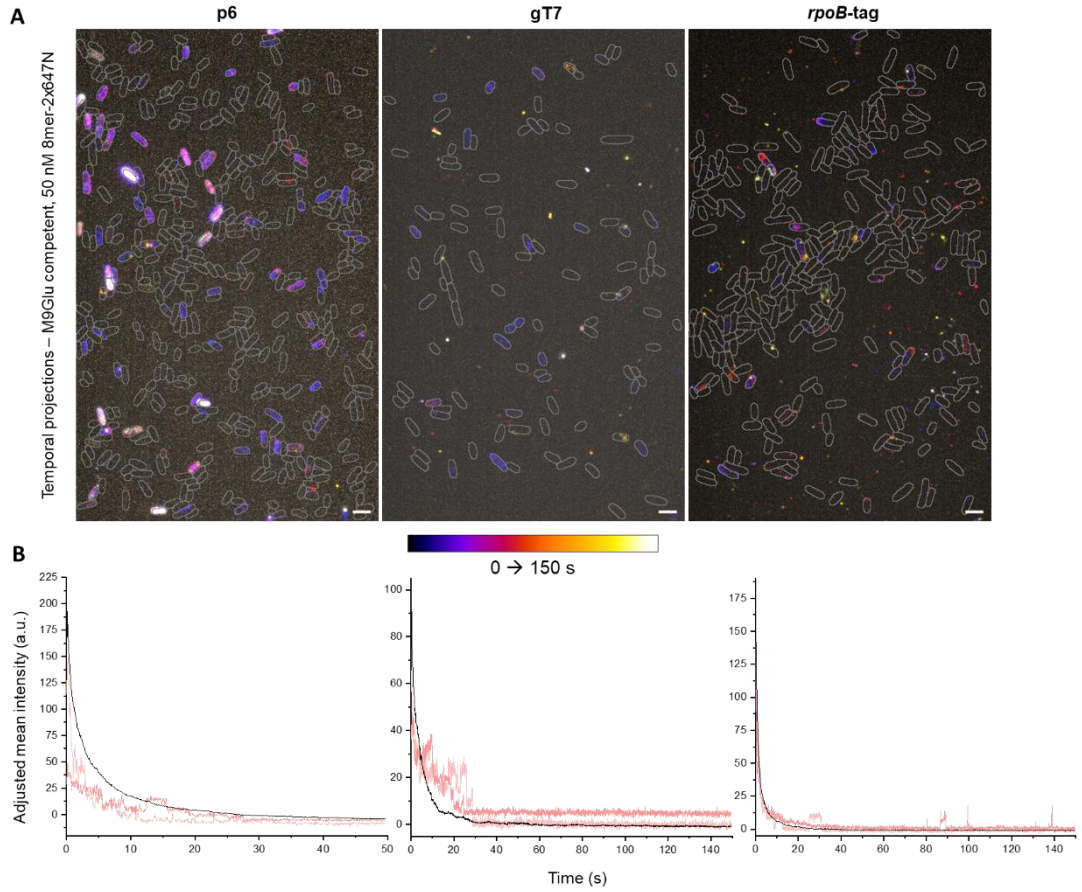


Figure A2.1 – Visualising the photostability of tag strain MG1655 *E. coli* M9Glu competent cells electroporated with 50 nM 8mer-2x647N.
(A) Temporal projections of the same representative FOVs of probe-electroporated tag strain MG1655 cells in 4.3. 100-1000 frames for p6, 100-2000 for gT7 and *rpoB*-tag cells. The colour bar denotes time, where darkest is time 0 (50 ms exposure, 50 s acquisition).
(B) Corresponding averaged background-corrected mean intensity per frame time traces for the representative FOVS (black), overlaid with two exemplar bleaching time traces for individual visually-loaded cells (red shades).
Scale bars – 3 µm, overlaid cell segmentation outlines.

Representative traces (black) with overlaid step height fittings (red) for each of the tag strains containing 1-5 foci from the start of acquisition are shown below to exemplify sequences of fluorescent foci observed, as well as time to bleach to baseline (Figure A2.2A). The corresponding step height histograms for 8mer-2x647N foci in all three tag strains are shown on the righthand side, where the single gaussian peak fit denotes an approximated unitary intensity of 6.1 units for p6 cells, 5.9 units for gT7 cells and 3.5 units for *rpoB*-tag cells (Figure A2.2B). Notably, the single Gaussian peak fits do not best fit the histograms, (a two Gaussian fit likely more appropriate), which likely explains the lower approximated unitary intensity for *rpoB*-tag cells over the p6/gT7 cells. Since the same concentration of probes is in use, and same preparation and imaging conditions too, it was anticipated that the unitary intensity value across strains would be effectively the same.

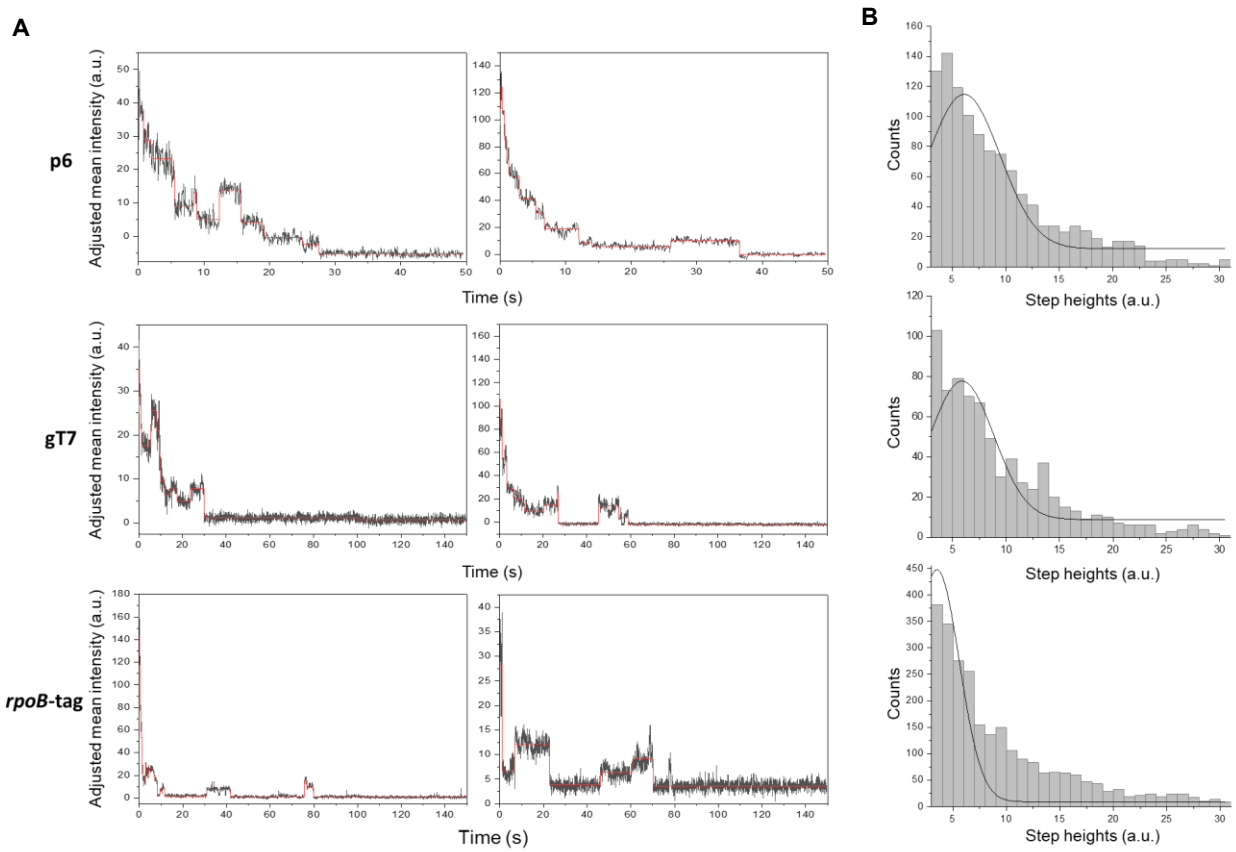


Figure A2.2 – Characterising the unitary intensity of the original 8mer-2x647N probes in M9Glu competent tag strain cells.
(A) For loaded p6, gT7 and rpoB-tag cells electroporated with 50 nM 8mer-2x647N, photo-bleaching step height analysis. Exemplar single cell photo-bleaching traces (black) fit for step heights (red) using *AutoStepFinder*. Steps correspond to the consecutive bleaching of fluorescent molecules, which can be correlated via the video acquisition.
(B) Corresponding histograms distributions of step height intensities from fitted single cell traces, where the peak denotes the unitary intensity of a single fluor, i.e. the average brightness of a probe-RNA hybridisation and de-quenching event.
p6: 6.1; gT7: 5.9; *rpoB*-tag: 3.5 units.

2.2 Comparing apparent diffusive populations across strains

To compare the distribution of apparent diffusive states (D*) of molecules thresholded for tracking between the three tag strains with 50 nM 8mer-2x647N. Using the curated cell segmentation masks, identified fluorescent foci were localised and linked across consecutive frames for track formation using the StormTracker and Locoli platforms, from frames 100 to 1000 frames – attempting to reduce the capture of non-specific signal in the first 5 second rapid bleach-down. For n single molecule trajectories in the p6 cells (n=3736), gT7 (n=2890) and *rpoB*-tag (n=987) cells, the MSD of individual probe-bound mRNAs was calculated for all trajectories across 4 or more consecutive frames, and the apparent diffusion coefficient (D*) calculated per strain from trajectories truncated to 4 steps (Uphoff et al., 2013).

The D* distributions for both the p6 and gT7 cells appeared subtly bimodal, enabling a two free diffusive species fit (Figure A2.3A). With the short length of the p6 transcript (128 nt), two species is a rational outcome when considering the two likely states occurring: rapid transcription, where the nascent transcript is bound by the RNAP albeit on the scale of ~3.5 seconds, and freely diffusing once released. The corresponding 34% "bound" population (red curve) at 0.083 $\mu m^2/s$ and 66% "diffusing" population at 0.4 $\mu m^2/s$ thereby seem appropriate – with the expected high copy number from the arabinose-induced plasmid system, the tag transcript would likely be in the order of 100s of copies, the majority of which would be able to rapidly diffuse from the cell poles around the cell. A two free species fit to another subtle bimodal distribution, of truncated trajectories from the gT7 cells, gave the same proportionality split populations with slightly different values – 0.12 $\mu m^2/s$ bound particle D* and 0.5 $\mu m^2/s$ diffusing. This is perhaps a little perplexing as per the transcript length (~2700 rnt) and estimated transcription time (~67.5 seconds, (Vogel & Jensen, 1994)).

Considering the transcription time estimated, a 50 second acquisition would not be long enough to enable visualisation of a full transcription initiation and elongation event; but imaging for 150 seconds potentially would. Figure A2.3B shows the D* distribution for gT7 trajectories from 100-3000 frames, with a three gamma two constrained fit. A much better fit of the overall distribution, 13% bound populations are classed as having a D* = 0.02 $\mu m^2/s$, 23% slowly-diffusing (green) with D* = 0.11 $\mu m^2/s$, and a 65% faster-diffusing species at D = 0.6 $\mu m^2/s$. The three-species fit seems much more appropriate: the intermediary diffusing population could correspond to actively-transcribing complexes, where the nascent mRNA and 5' tag site would be increasingly mobile as elongation proceeded. Noted, the diffusing population is not confluent to the histogram tail for any of the two-free fit strain distributions, suggesting that there are many more complex behaviours occurring than represented here.

The *rpoB*-tag cell D* distribution fit with a three-gamma two-constrained species curves presented alongside the wildtype distribution denotes a 31% bound population with D* = 0.03 $\mu m^2/s$, a 35% slow-diffusing population with D* = 0.1 $\mu m^2/s$, and a 34% fast-diffusing population with D* = 0.5 $\mu m^2/s$ diffusion rate. The left-skew in diffusion rates seems appropriate given the extensive expected 4000-8329 nt transcript and potential almost immediate/simultaneous translation.

Across all tag strains, there is much overlap with the wildtype D* distribution for populations between D* = 0.02-0.13 $\mu m^2/s$. Whilst there is probably much specific, "real" signal within the tag strain populations at these diffusion rates (especially as the wildtype diffusing population

is poorly fit, and probably more ~0.08 μm²/s), it is not possible to make any such conclusions at a populational level.

As such, the key population across all strains that is not visible in the wildtype distribution is the fast-diffusing population, which seems to be the majority population for all despite poor fitting. This is promising, considering recent evidence that mRNAs may diffuse freely throughout the cell, nucleoid-inclusive (Bakshi et al., 2012; Sanamrad et al., 2014; Sattler & Graumann, 2021).

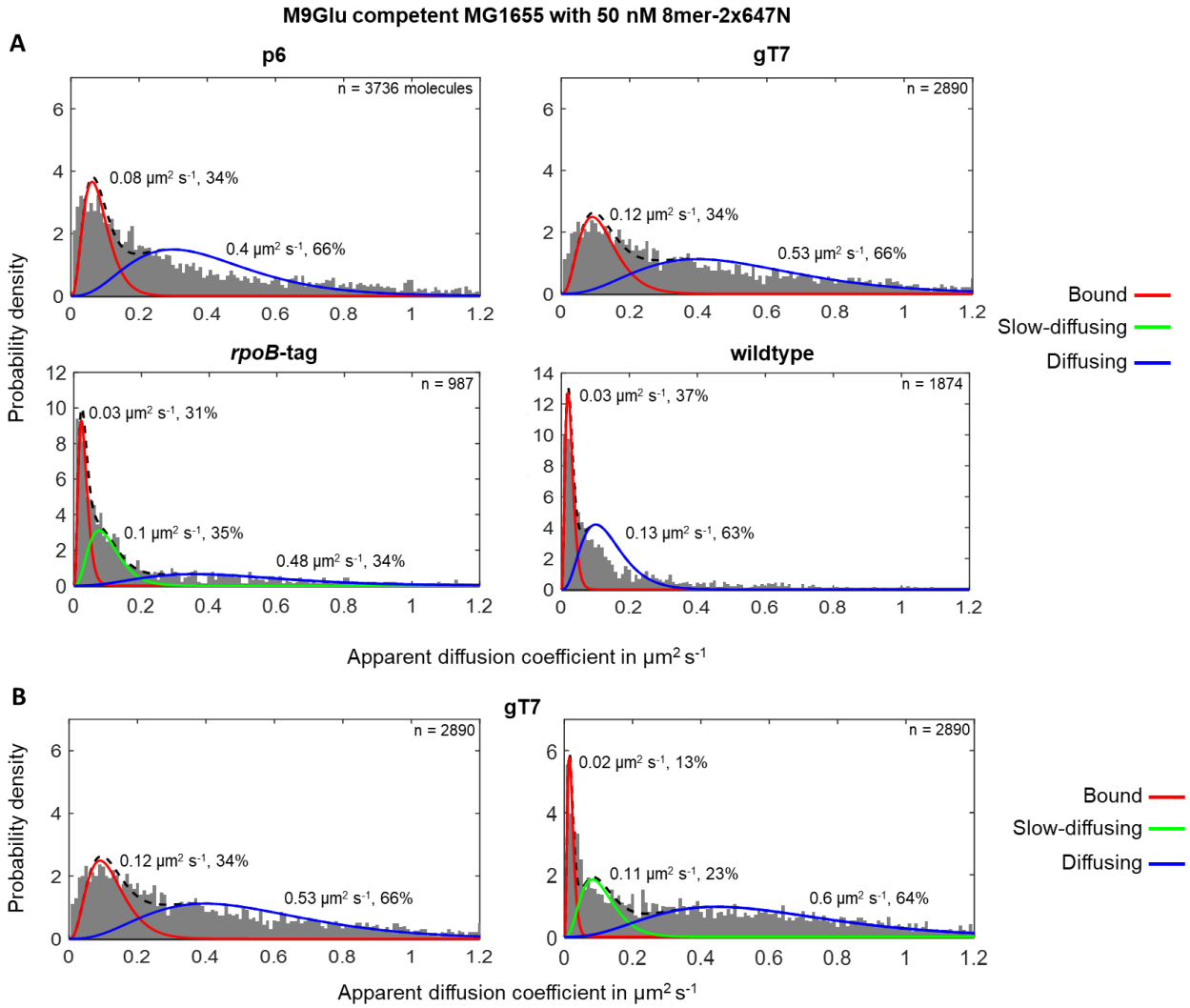


Figure A2.3 – Exploring the diffusive properties of tracked fluorescent molecules in MG1655 competent M9Glu tag strain and wildtype cells with original 8mer-2x647N: apparent diffusion coefficients.
(A) Apparent diffusion coefficient distributions fit with two-free species curves (p6, gt7 and wildtype) and three gamma two-constrained fit (rpoB-tag) for the same population of gT7 cells (red as bound, blue as diffusing and intermediary green as slow-diffusing).
(B) Apparent diffusion coefficient distributions for gT7 cells, comparing a two-free species and three gamma two-constrained fit for the same population of gT7 cells.
n gives the numbers of tracked molecules per sample.

2.3 Populational tracked molecules – counts per cell, durations & spatial distributions

To enable the quantitation of the spatiotemporal aspects of single molecules in tag strains with 50 nM 8mer-2x647N probes, four FOVs (>800 cells) per strain were processed using *MolTrack*. Per sample, 3997 were found in p6 cells, 3124 tracks in gT7 cells, 7118 tracks in *rpoB*-tag cells, and 276 tracks in wildtype cells, after thresholding for spurious trajectories. Representative FOVs for each strain are shown in Figure A2.4.

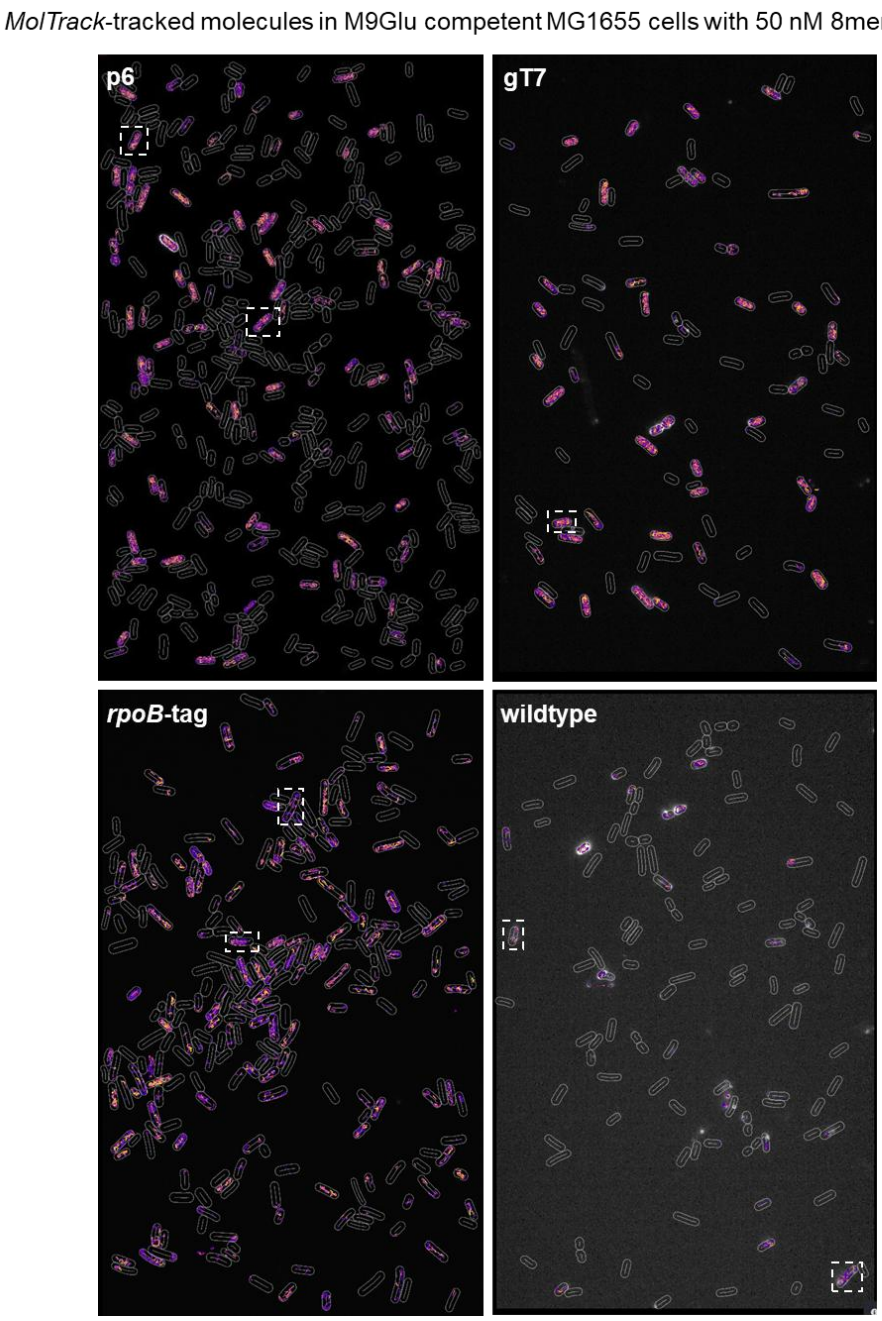


Figure A2.4 – *Moltrack* processing of the representative FOVs for tag-strain and wildtype cells electroporated with 50 nM original 8mer-2x647N, showing all trajectories per cells. Scale bar – 5 µm.

Counts per cell

Investigating the normalised counts of tracked molecules within cells from tag strains (versus the wildtype controls) showed that the majority of cells across all strains comprised just one fluorescent signal, corroborative of the signal spread seen across experiments so far (Figure A2.5). The wildtype cells had the lowest proportion (~9%) of cells loaded with 10 or more tracked molecules compared to the tag strains, for which p6 cells had 40% cells, gT7 ~47% cells and *rpoB*-tag cells ~34% cells with over 10 molecules per cell. Overall, far fewer cells with tracked molecules were found in wildtype cells compared to cells with tagged mRNAs: 65 wildtype cells had tracked molecules, compared to 255 p6 cells, 176 gT7 cells and 564 *rpoB*-tag cells. This suggests that the more highly loaded cells may actually represent specific hybridisation complexes, and may offer greater insight into different molecular behaviours; appreciating that some of these are potentially spuriously higher numbers, as mentioned above.

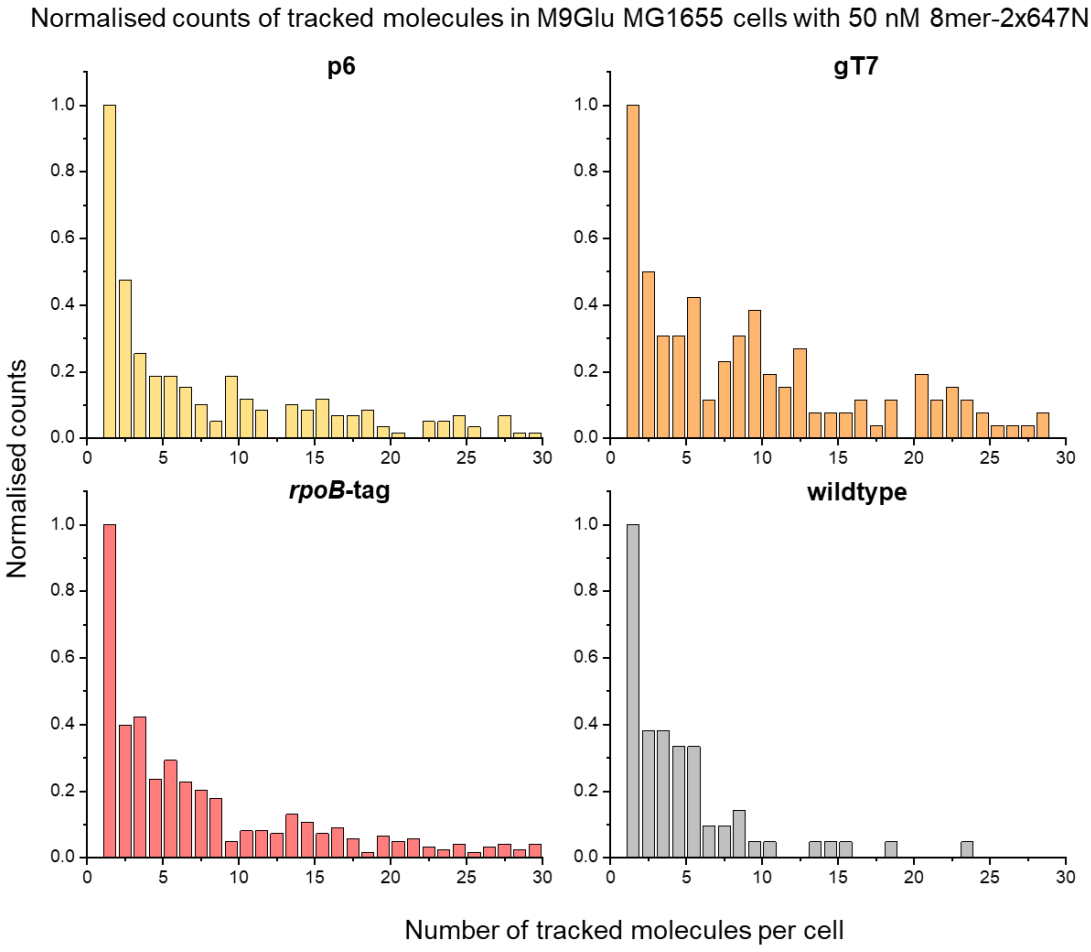


Figure A2.5 – Tracked molecule distributions comparing normalised counts per cell in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-2x647N.
Bin size = 1.

Tracked molecule durations

Figure A2.6 shows the normalised distributions of tracked particle lifetimes, clipped at 10 seconds. Interestingly, all distributions are highly overlapping at the lower time distributions – between 73-77% tag strain cell molecules have lifetimes between 0.2-0.75 seconds – suggesting that most of these signals are highly transient, either rapidly-diffusing particles or misassigned and erroneously unconnected trajectories for the same molecule. Only 29% of wildtype cells had tracked molecules between 0.2-0.75 seconds, with higher longer-lived proportions, but this more reflective overall of the small total tracked molecules in wildtype cells (276 versus ~3000-7000 tracked molecules in tag strain cells). Indeed, the chromosomal tag mRNAs had the widest distributions of lifetimes; overall, particles with >3 seconds' lifetimes likely constitute mRNA-probe binding events.

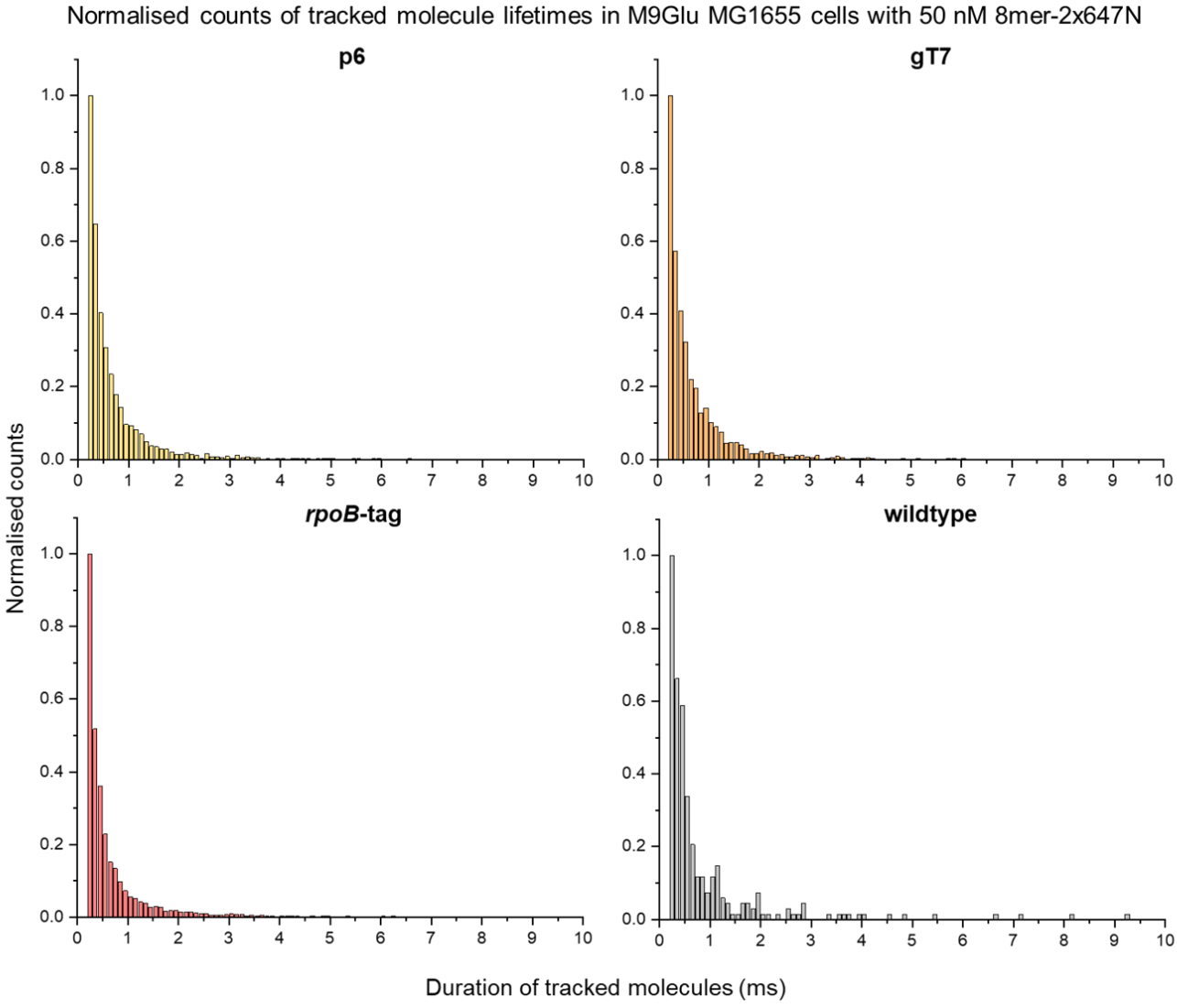


Figure A2.6 – Normalised lifetime distributions for tracked molecules in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-2x647N. Bin size = 0.1.

Spatial distribution of tracked molecules

Whilst the small population of tracked molecules in wildtype cells show no trend in membrane proximity, all three tag strains seem to have distinct populations around the 200-300 nm distance from the membrane (Figure A2.7). Despite a somewhat uniform distribution, each tag-strain has molecules apparently diffusing between ~10-550 nm from the membrane – suggesting that in rapid diffusion, it is likely that mRNAs can traverse most of the cell (assuming a bacillus of ~1 µm width). More membrane-proximal species here could be nascently-transcribing complexes – considering the polar localisation of the p6 plasmids and the nucleoid peripheral localisation of the gT7 locus – but equally could be spurious signal from membrane-bound probes.

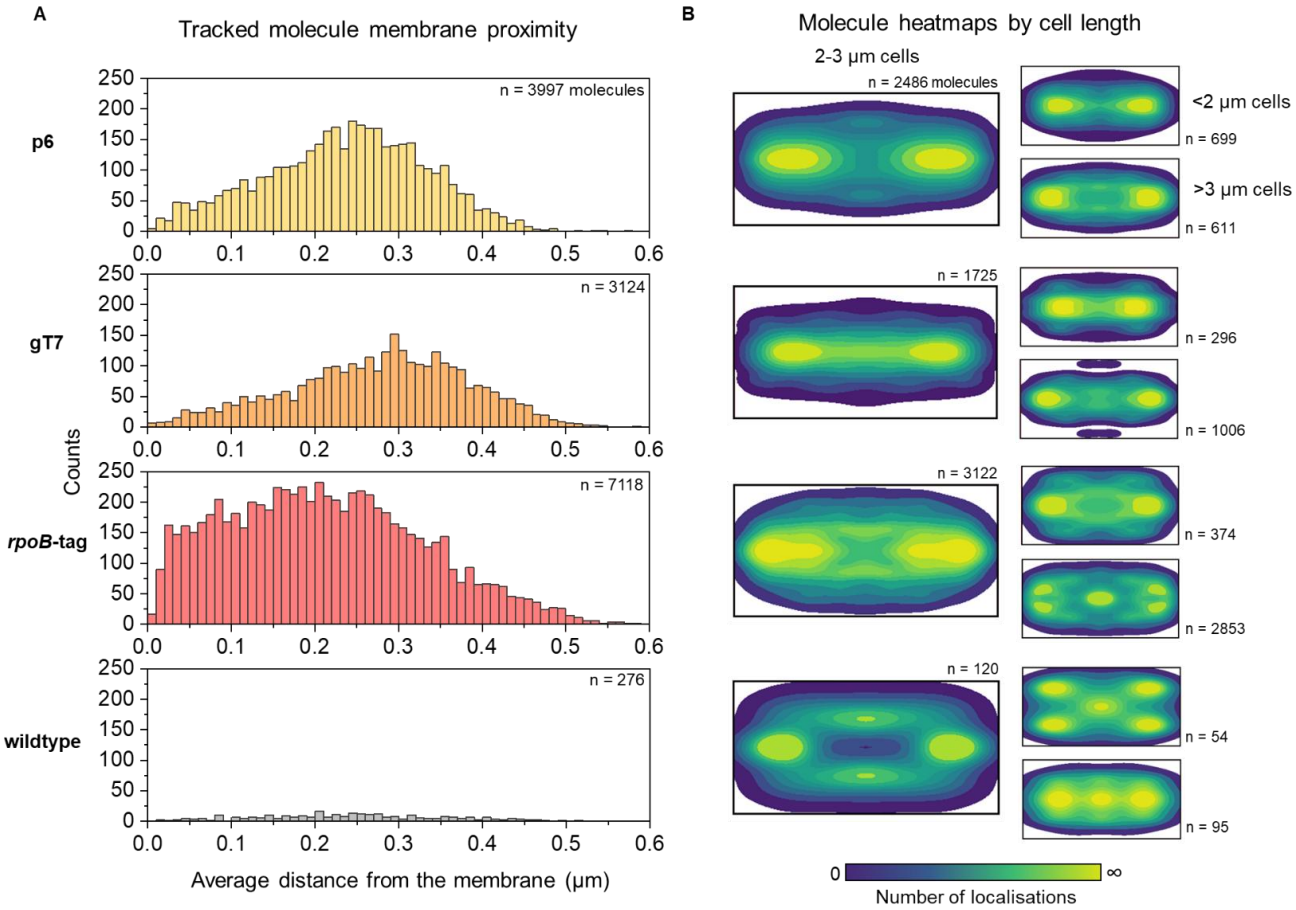


Figure A2.7 – Spatial distributions of tracked molecules in M9Glu competent tag strain and wildtype cells electroporated with 50 nM 8mer-2x647N.
(A) Histogram distributions of average molecule distance from the membrane; bin size = 0.05.
(B) Heatmaps showing the density of molecule localisations in different regions of the cell; the central plot denotes the average unit cell (length 2-3 µm), and the two corresponding righthand heatmaps denote cell lengths of <2 µm and >3 µm per strain sample.
n gives the number of molecules.

Plotting the localisation density of tracked molecules by cell length, localisation density was highest at the poles, across all cell lengths for induced p6 and gT7 cells, particularly for the p6 cells (Figure A2.7B). This corroborates with polarly-localised plasmid-generated RNAs and peripheral genomic location of the gT7 construct respectively. There are higher midcell localisations for gT7, which may also correspond to genomic transcription and lateral diffusion of the released gT7 transcripts.

Interestingly in *rpoB*-tag cells, there also appears this polar tendency, alongside a higher proportion of midcell localisations – these could indeed correspond to the midcell transcription locus of *rpoBC*, and the trafficking of the very long mRNAs to the poles for efficient polysomal translation. Negligibly consistent spatial patterning in wildtype cells across all lengths may be more indicative of the very low comparative localisation n – showing that wildtype signals are much fewer generally, and likely do not have definitive off-target loci.

2.4 Behavioural characteristics of single RNA molecules bound by 8mer-2x647N

Accounting for the lesser specificity of the 8mer-2x647N probes in all strains, the following exemplar molecules presented are for an initial overview of likely characteristics, with brief speculations of biological implications.

From the representative FOVs in several exemplar cells (highlighted): exemplar cells are shown as a montage of stills taken from the real-time video, correlated to text descriptions of individual foci. For each cell, the *Moltrack* output of all tracked molecule trajectories in the cell is presented; track colouring denotes speed, where dark blue is static and yellow fastest.

Wildtype

The majority of cells with fluorescent signals were extremely highly loaded, and bleached down very quickly, well within the first 5 seconds of acquisition.

Exemplar cell A (Figure A2.8) denotes a cell with two distinct foci from laser activation: the first focus, around the midcell is completely static, suggestive of membrane implantation of the ATTO647N dye; the second, rapidly diffusing down the cell midline. Both are bleached by 1.75 seconds into the acquisition. A third diffuse static focus appears at the upper cell pole 1 second later, bleaching within 6 seconds; again, suggestive of probe-membrane non-specific interaction. No further signals arise.

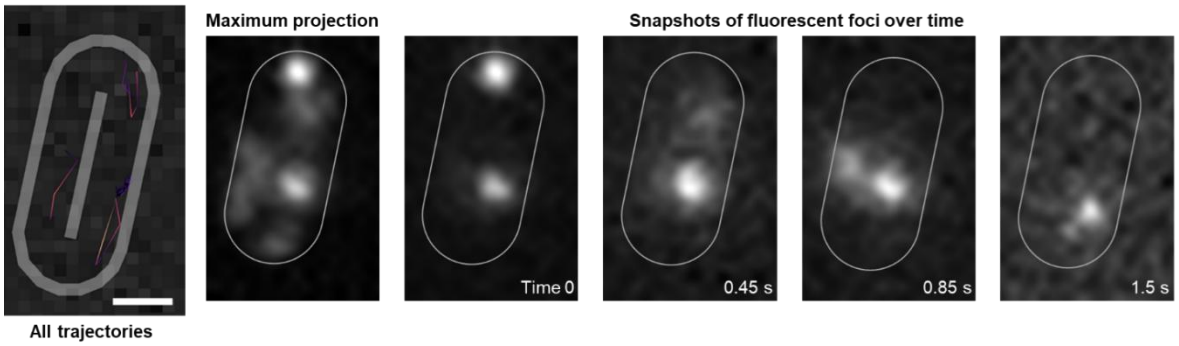


Figure A2.8 – Visuals over time: wildtype exemplar cell A.
Left to right: *Moltrack* 'all trajectories' output of all tracked molecules in the cell (left); maximum intensity projection of cell signal; montage of single frame stills of fluorescent foci over time (time in seconds). Scale bar – 500 nm.

Figures A2.9-A2.14 follow this same flow.

Cell B (Figure A2.9) is representative of a highly loaded cell – the ~6 foci from the start all rapidly diffuse across the second region of the cell, and disappear within 9.5 seconds of acquisition. One more bright focus arises, interestingly, at 12.4 seconds – this fluor appears to explore the midcell for ~10 seconds with varying diffusion speeds, before extinguishing. This is a surprisingly long period of time, albeit with some brightness and resolution variations indicating different depths of diffusion as well. One suggestion is that the 8mer-2x647N may have been bound by a small diffusive protein like SSB, which is known to sequester ssDNAs and transport them through the cell.

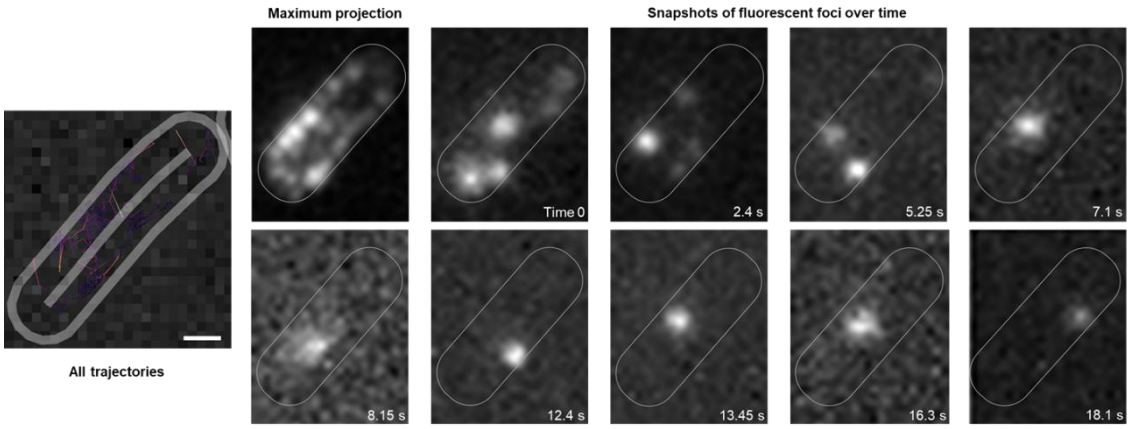


Figure A2.9 – Visuals over time: wildtype exemplar cell B.

Overall, with most wildtype signal coming from highly loaded cells, and within the first ~5 seconds, fluorescent molecule behaviours from the tag strains were investigated with molecules occurring beyond the 5 second time point, and avoiding very bright cells.

Induced p6

As expected for the short 128 rnt transcript, loaded p6 cells typically exhibited multiple very fast-diffusing foci, which were able to cover the whole length of the cell several times before extinguishing. Many cells had foci up until the end of the 50 second acquisition.

Exemplar cell A (Figure A2.10)is a representative exemplar of rapidly-diffusing foci; comprising 1-3 bright and fast-diffusing foci at the start of the video, the bright foci were seen to diffuse up and down the cell midline and periphery, occasionally becoming bound, and consecutively bleaching until no signals were visible by ~26 seconds.

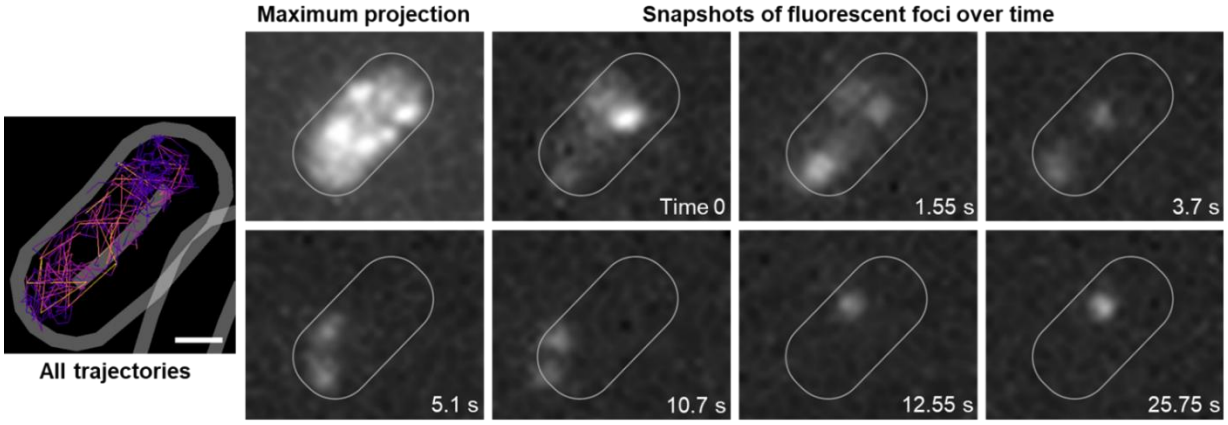


Figure A2.10 – Visuals over time: p6 exemplar cell A.

Exemplar cell B (Figure A2.11) exhibits a possible transcription site focus: a single intermediary diffusing focus was visible from the start until 20 seconds into acquisition. At the same locus (indicated at 19.1 seconds) , a diffuse focus reappeared at 20.1 seconds (represented by 23.05-25.2 second stills here) and remained static for the next 25 seconds, until disappearance. The rapid diffusion could represent released RNAs diffusing around the cell, versus a potential transcription hotspot with the static focus, considering the polar localisation of the p15a plasmid.

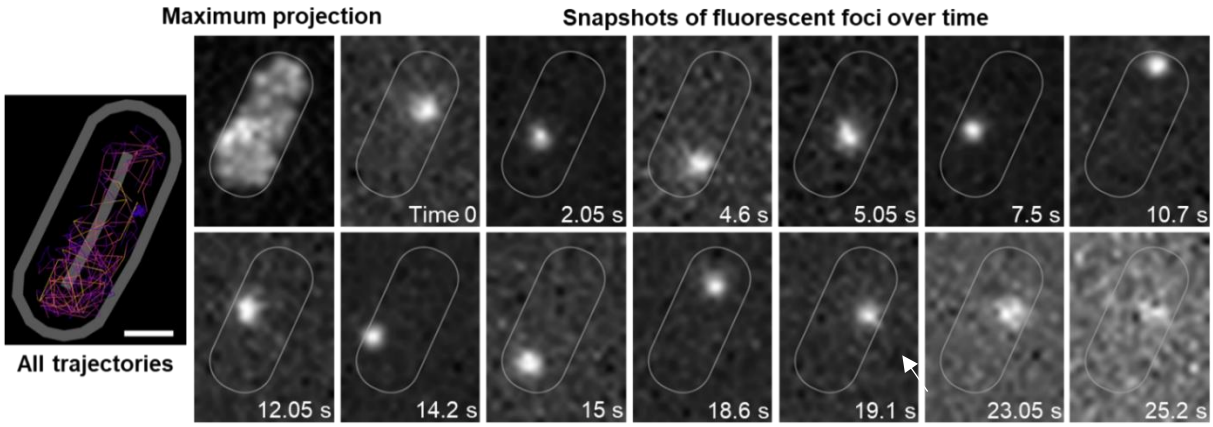


Figure A2.11 – Visuals over time: p6 exemplar cell B.

Induced gT7

Surprisingly for an anticipated 2736 rnt transcript, very few cells showed particularly static or highly immobile foci which could correspond to a core transcription site. Instead, most cells containing probes and fluorescence signal showed fast-diffusing distinguishable foci of variable brightness, which could be seen all the way from laser activation to ~25-35 seconds of 150 seconds' acquisition. This is exactly the case in the exemplar cell below (Figure A2.12), where two bright foci can be seen from the start, bleaching and then transitorily reappearing over a 25 second period, all while rapidly "dancing" around the perimeter of the cell.

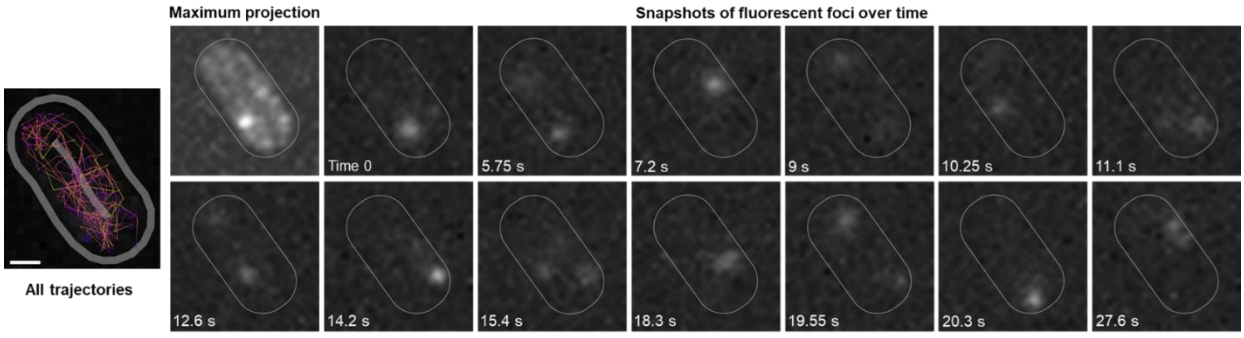


Figure A2.12 – Visuals over time: gT7 exemplar cell.

rpoB-tag

With an anticipated 4000-8329 nt transcript, the estimated transcription time would be well over 100 seconds – unlikely to be fully coherently captured unless a sufficiently concentrated quenched probe pool is available. Nevertheless, considering the extended timescales for both transcription and consecutive translation, it was hypothesised that much longer, and more bound, foci would be observable. Indeed recurrent signals visualised longer into an acquisition period after majoritive photo-bleaching.

Exemplar cell A is one such case of recurrent signal following an extended dark period (Figure A2.13). With multiple bright foci at time 0, bleaching to 5-6 distinguishable foci by ~2 seconds into acquisition, the remaining foci diffused around the midcell and consecutively bleach down by ~ 8 seconds. There were then no further signals observed until 40.5 seconds into the 150 second acquisition, where a single bright focus appears, proximal to the lower midcell membrane. This focus diffuses into the midcell (43.3-47.95 seconds), then over to the left cell pole (48.85 seconds), exhibiting slow diffusion along the pole membrane for ~4 seconds before appearing to move back to the midcell and disappearing around 55 seconds. No other foci arise. This hence suggests that many cells do sustain a population of quenched probes in the cell milieu.

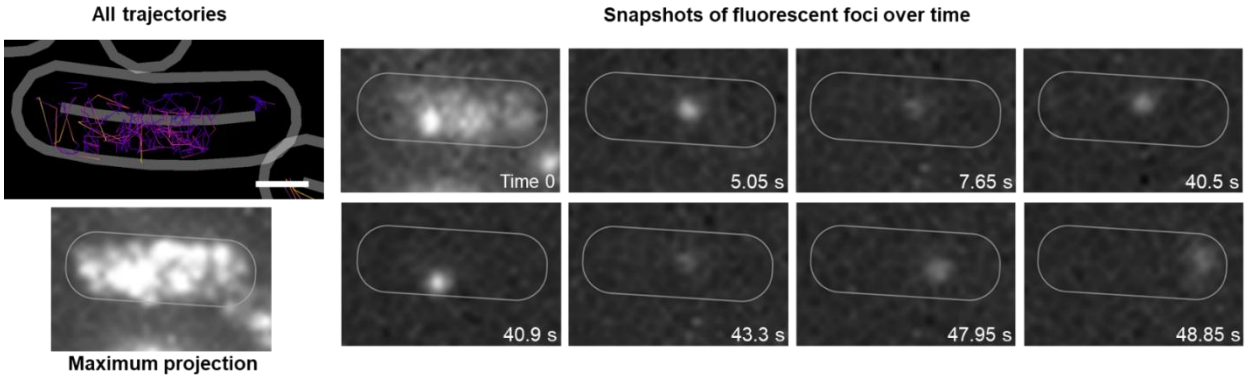


Figure A2.13 – Visuals over time: *rpoB*-tag exemplar cell A.

A potential translation focus may be visible in exemplar cell B, also an example of one of the very bright cells (high probe loading)(Figure A2.14). Flooded with signal at time 0, rapid photo-bleaching of de-quenched probes gave ~6 distinguishable bright foci and some diffuse background foci by 3.95 seconds; 5 clear foci are seen at 8.5 seconds. Up to 5 foci are then consistently observed up to 21.85 seconds, most notably with two highly static foci at the right cell pole. The highlighted lower right polar focus remains static for the ~55 seconds documented here.

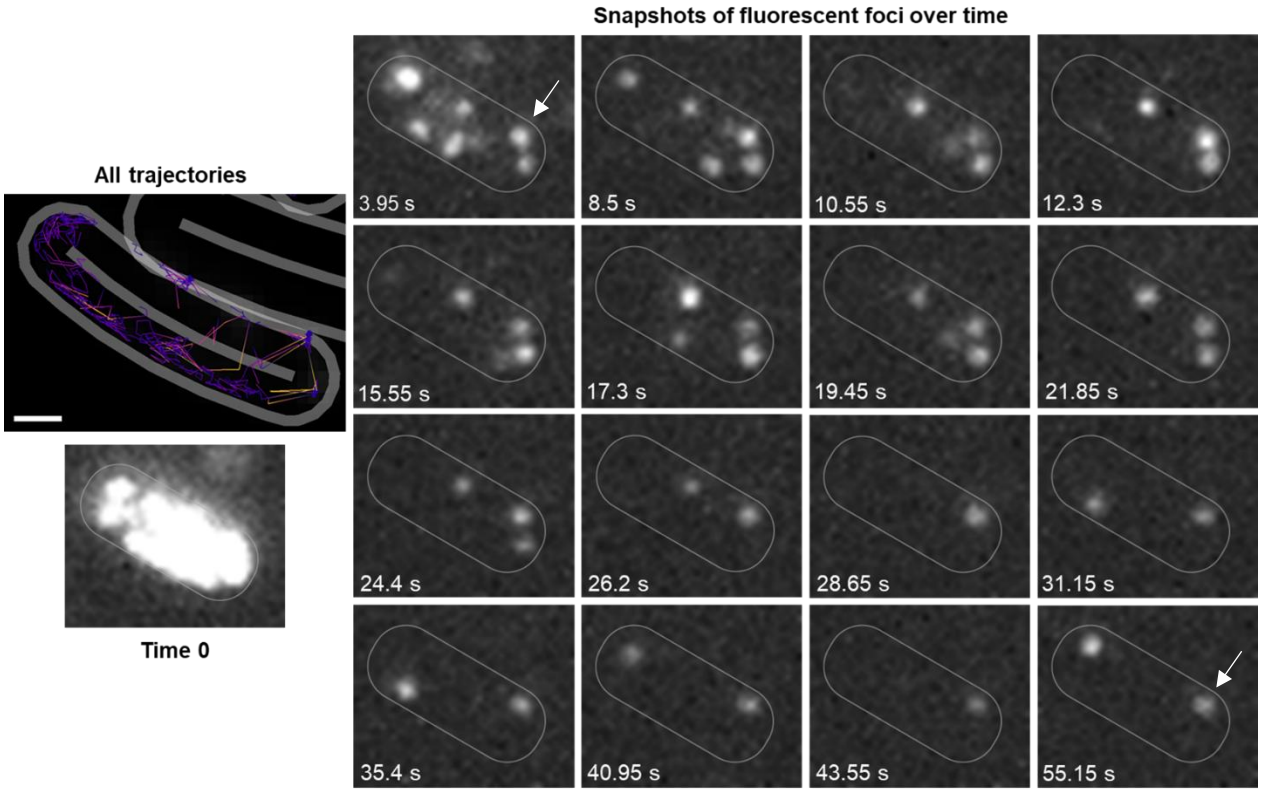


Figure A2.14– Visuals over time: *rpoB*-tag exemplar cell B.

With the majority of translation nucleoid-excluded and localised to the cell poles where high ribosome density has been shown (Bakshi et al., 2012; Sanamrad et al., 2014), a static focus on the order of 55 seconds is suggestive of a localised and dedicated region of translation, perhaps polysomal.

Bibliography

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, *9*(1), 413–468. https://doi.org/10.1007/BF02956173

Aigrain, L., Sustarsic, M., Crawford, R., Plochowietz, A., & Kapanidis, A. N. (2015). Internalization and Observation of Fluorescent Biomolecules in Living Microorganisms via Electroporation. *JoVE*, *96*, e52208. https://doi.org/10.3791/52208

Amann, E., Ochs, B., & Abel, K. J. (1988). Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. *Gene*, *69*(2), 301–315. https://doi.org/10.1016/0378-1119(88)90440-4

Amann, R. I. (1990). *Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations*. https://doi.org/10.1128/aem.56.6.1919-1925.1990

Andrews, R., Steuer, H., El-Sagheer, A. H., Mazumder, A., Sayyed, H. El, Shivalingam, A., Brown, T., & Kapanidis, A. N. (2022). *Transient DNA binding to gapped DNA substrates links DNA sequence to the single molecule kinetics of protein-DNA interactions* (p. 2022.02.27.482175). bioRxiv. https://doi.org/10.1101/2022.02.27.482175

Anne Plochowietz. (2015). *Single molecule FRET studies in live bacteria*. University of Oxford.

Antony, E., & Lohman, T. M. (2019). Dynamics of E. coli single stranded DNA binding (SSB) protein-DNA complexes. *Seminars in Cell & Developmental Biology*, *86*, 102–111. https://doi.org/10.1016/j.semcdb.2018.03.017

Arraiano, C. M., Andrade, J. M., Domingues, S., Guinote, I. B., Malecki, M., Matos, R. G., Moreira, R. N., Pobre, V., Reis, F. P., Saramago, M., Silva, I. J., & Viegas, S. C. (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiology Reviews*, *34*(5), 883–923. https://doi.org/10.1111/j.1574-6976.2010.00242.x

Asif, A., Mohsin, H., Tanvir, R., & Rehman, Y. (2017). Revisiting the Mechanisms Involved in Calcium Chloride Induced Bacterial Transformation. *Frontiers in Microbiology*, *8*, 2169. https://doi.org/10.3389/fmicb.2017.02169

ATTO 643. ATTO-TEC GmbH. Retrieved 27 February 2025, from https://www.atto-tec.com/ATTO-643.html?language=en

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, *2*(1), 2006.0008. https://doi.org/10.1038/msb4100050

Babu, M., Díaz-Mejía, J. J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., Yousif, F., Ding, H., Xiong, X., Nazarians-Armavil, A., Alamgir, M., Ali, M., Pogoutse, O., Pe'er, A., Arnold, R., Michaut, M., Parkinson, J., Golshani, A., Whitfield, C., … Emili, A. (2011). Genetic Interaction Maps in Escherichia coli Reveal Functional Crosstalk among Cell Envelope Biogenesis Pathways. *PLoS Genetics*, *7*(11), e1002377. https://doi.org/10.1371/journal.pgen.1002377

Bajar, B. T., Wang, E. S., Lam, A. J., Kim, B. B., Jacobs, C. L., Howe, E. S., Davidson, M. W., Lin, M. Z., & Chu, J. (2016). Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Scientific Reports*, *6*(1), 20889. https://doi.org/10.1038/srep20889

Bakshi, S., Bratton, B. P., & Weisshaar, J. C. (2011). Subdiffraction-Limit Study of Kaede Diffusion and Spatial Distribution in Live *Escherichia coli*. *Biophysical Journal*, *101*(10), 2535–2544. https://doi.org/10.1016/j.bpj.2011.10.013

Bakshi, S., Choi, H., Mondal, J., & Weisshaar, J. C. (2014). Time-dependent effects of transcription- and translation-halting drugs on the spatial distributions of the Escherichia coli chromosome and ribosomes. *Molecular Microbiology*, *94*(4), 871–887. https://doi.org/10.1111/mmi.12805

Bakshi, S., Choi, H., & Weisshaar, J. C. (2015). The spatial biology of transcription and translation in rapidly growing Escherichia coli. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00636

Bakshi, S., Dalrymple, R. M., Li, W., Choi, H., & Weisshaar, J. C. (2013). Partitioning of RNA Polymerase Activity in Live Escherichia coli from Analysis of Single molecule Diffusive Trajectories. *Biophysical Journal*, *105*(12), 2676–2686. https://doi.org/10.1016/j.bpj.2013.10.024

Bakshi, S., Siryaporn, A., Goulian, M., & Weisshaar, J. C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. *Molecular Microbiology*, *85*(1), 21–38. https://doi.org/10.1111/j.1365-2958.2012.08081.x

Balantič, K., Miklavčič, D., Križaj, I., & Kramar, P. (2021). The good and the bad of cell membrane electroporation. *Acta Chimica Slovenica*, *68*(4), Article 4. https://doi.org/10.17344/acsi.2021.7198

Barry, G., Squires, C. L., & Squires, C. (1979). Control features within the rplJL-rpoBC transcription unit of Escherichia coli. *Proceedings of the National Academy of Sciences*, *76*(10), 4922–4926. https://doi.org/10.1073/pnas.76.10.4922

Barry, G., Squires, C., & Squires, C. L. (1980). Attenuation and processing of RNA from the rplJL–rpoBC transcription unit of Escherichia coli. *Proceedings of the National Academy of Sciences*, *77*(6), 3331–3335. https://doi.org/10.1073/pnas.77.6.3331

Batani, G., Bayer, K., Böge, J., Hentschel, U., & Thomas, T. (2019). Fluorescence in situ hybridization (FISH) and cell sorting of living bacteria. *Scientific Reports*, *9*(1), 18618. https://doi.org/10.1038/s41598-019-55049-2

Bates, M., Huang, B., Dempsey, G. T., & Zhuang, X. (2007). Multicolor Super-Resolution Imaging with Photo-Switchable Fluorescent Probes. *Science*, *317*(5845), 1749–1753. https://doi.org/10.1126/science.1146598

Batra, R., Nelles, D. A., Pirie, E., Blue, S. M., Marina, R. J., Wang, H., Chaim, I. A., Thomas, J. D., Zhang, N., Nguyen, V., Aigner, S., Markmiller, S., Xia, G., Corbett, K. D., Swanson, M. S., & Yeo, G. W. (2017). Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell*, *170*(5), 899-912.e10. https://doi.org/10.1016/j.cell.2017.07.010

Baughman, G., & Nomura, M. (1983). Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in Escherichia coli. *Cell*, *34*(3), 979–988. https://doi.org/10.1016/0092-8674(83)90555-x

Baughman, G., & Nomura, M. (1984). Translational regulation of the L11 ribosomal protein operon of Escherichia coli: Analysis of the mRNA target site using oligonucleotide-directed mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(17), 5389–5393. https://doi.org/10.1073/pnas.81.17.5389

Bauman, J. G. J., Wiegant, J., Borst, P., & van Duijn, P. (1980). A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA. *Experimental Cell Research*, *128*(2), 485–490. https://doi.org/10.1016/0014-4827(80)90087-7

Benhalevy, D., Bochkareva, E. S., Biran, I., & Bibi, E. (2015). Model Uracil-Rich RNAs and Membrane Protein mRNAs Interact Specifically with Cold Shock Proteins in Escherichia coli. *PLOS ONE*, *10*(7), e0134413. https://doi.org/10.1371/journal.pone.0134413

Benjamin, D. R., Robinson, C. V., Hendrick, J. P., Hartl, F. U., & Dobson, C. M. (1998). Mass spectrometry of ribosomes and ribosomal subunits. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7391. https://doi.org/10.1073/pnas.95.13.7391

Benov, L. C., Antonov, P. A., & Ribarov, S. R. (1994). Oxidative damage of the membrane lipids after electroporation. *General Physiology and Biophysics*, *13*(2), 85–97.

Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S., & Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proceedings of the National Academy of Sciences*, *99*(15), 9697–9702. https://doi.org/10.1073/pnas.112318199

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., & Long, R. M. (1998). Localization of ASH1 mRNA Particles in Living Yeast. *Molecular Cell*, *2*(4), 437–445. https://doi.org/10.1016/S1097-2765(00)80143-4

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., & Hess, H. F. (2006). Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, *313*(5793), 1642–1645. https://doi.org/10.1126/science.1127344

BLACK HOLE QUENCHER™ (BHQ™). Retrieved 7 March 2025, from https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/genomics/qpcr/black-hole-quenchers

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., & Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science (New York, N.Y.)*, *277*(5331), 1453–1462. https://doi.org/10.1126/science.277.5331.1453

Bliss, N., Bindewald, E., & Shapiro, B. A. (2020). Predicting RNA SHAPE scores with deep learning. *RNA Biology*, *17*(9), 1324–1330. https://doi.org/10.1080/15476286.2020.1760534

Blokesch, M. (2016). Natural competence for transformation. *Current Biology*, *26*(21), R1126–R1130. https://doi.org/10.1016/j.cub.2016.08.058

Blount, Z. D. (2015). The unexhausted potential of E. coli. *eLife*, *4*, e05826. https://doi.org/10.7554/eLife.05826

Bremer, H., Dennis, P., & Ehrenberg, M. (2003). Free RNA polymerase and modeling global transcription in *Escherichia coli*. *Biochimie*, *85*(6), 597–609. https://doi.org/10.1016/S0300-9084(03)00105-6

Bremer, H., & Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*, *3*(1). https://doi.org/10.1128/ecosal.5.2.3

Brenner, S., Jacob, F., & Meselson, M. (1961). An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature*, *190*(4776), 576–581. https://doi.org/10.1038/190576a0

Brosius, J., Dull, T. J., Sleeter, D. D., & Noller, H. F. (1981). Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *Journal of Molecular Biology*, *148*(2), 107–127. https://doi.org/10.1016/0022-2836(81)90508-8

Bujalowski, W., & Lohman, T. M. (1986). Escherichia coli single-strand binding protein forms multiple, distinct complexes with single-stranded DNA. *Biochemistry*, *25*(24), 7799–7802. https://doi.org/10.1021/bi00372a003

Burgess, R. R. (1969). Separation and Characterization of the Subunits of Ribonucleic Acid Polymerase. *Journal of Biological Chemistry*, *244*(22), 6168–6176. https://doi.org/10.1016/S0021-9258(18)63521-5

Burgess, R. R., Travers, A. A., Dunn, J. J., & Bautz, E. K. F. (1969). Factor Stimulating Transcription by RNA Polymerase. *Nature*, *221*(5175), 43–46. https://doi.org/10.1038/221043a0

Carpousis, A. J. (2007). The RNA Degradosome of Escherichia coli: An mRNA-Degrading Machine Assembled on RNase E. *Annual Review of Microbiology*, *61*(Volume 61, 2007), 71–87. https://doi.org/10.1146/annurev.micro.61.080706.093440

Castellana, M., Hsin-Jung Li, S., & Wingreen, N. S. (2016). Spatial organization of bacterial transcription and translation. *Proceedings of the National Academy of Sciences*, *113*(33), 9286–9291. https://doi.org/10.1073/pnas.1604995113

Chai, Q., Singh, B., Peisker, K., Metzendorf, N., Ge, X., Dasgupta, S., & Sanyal, S. (2014). Organization of Ribosomes and Nucleoids in Escherichia coli Cells during Growth and in Quiescence *. *Journal of Biological Chemistry*, *289*(16), 11342–11352. https://doi.org/10.1074/jbc.M114.557348

Chakraborty, K., Veetil, A. T., Jaffrey, S. R., & Krishnan, Y. (2016). Nucleic Acid–Based Nanodevices in Biological Imaging. In *Annual Review of Biochemistry* (Vol. 85, Issue Volume 85, 2016, pp. 349–373). Annual Reviews. https://doi.org/10.1146/annurev-biochem-060815-014244

Chamberlin, M. J. (1974). The Selectivity of Transcription. *Annual Review of Biochemistry*, *43*(Volume 43, 1974), 721–775. https://doi.org/10.1146/annurev.bi.43.070174.003445

Chao, J. A., Patskovsky, Y., Almo, S. C., & Singer, R. H. (2008). Structural basis for the coevolution of a viral RNA–protein complex. *Nature Structural & Molecular Biology*, *15*(1), 103–105. https://doi.org/10.1038/nsmb1327

Chattopadhyay, S., Weissman, S. J., Minin, V. N., Russo, T. A., Dykhuizen, D. E., & Sokurenko, E. V. (2009). High frequency of hotspot mutations in core genes of Escherichia coli due to short-term positive selection. *Proceedings of the National Academy of Sciences*, *106*(30), 12412–12417. https://doi.org/10.1073/pnas.0906217106

Chatzimichail, S., Turner, P., Feehily, C., Farrar, A., Crook, D., Andersson, M., Oakley, S., Barrett, L., Sayyed, H. E., Kyropoulos, J., Nellåker, C., Stoesser, N., &

Kapanidis, A. N. (2024). Rapid identification of bacterial isolates using microfluidic adaptive channels and multiplexed fluorescence microscopy. *Lab on a Chip*, *24*(20), 4843–4858. https://doi.org/10.1039/D4LC00325J

Chen, H., Shiroguchi, K., Ge, H., & Xie, X. S. (2015). Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. *Molecular Systems Biology*, *11*(1), 781. https://doi.org/10.15252/msb.20145794

Chen, J., Tsai, A., Petrov, A., & Puglisi, J. D. (2012). Nonfluorescent Quenchers To Correlate Single molecule Conformational and Compositional Dynamics. *Journal of the American Chemical Society*, *134*(13), 5734–5737. https://doi.org/10.1021/ja2119964

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, *348*(6233), aaa6090. https://doi.org/10.1126/science.aaa6090

Chenouard, N., Smal, I., de Chaumont, F., Maška, M., Sbalzarini, I. F., Gong, Y., Cardinale, J., Carthel, C., Coraluppi, S., Winter, M., Cohen, A. R., Godinez, W. J., Rohr, K., Kalaidzidis, Y., Liang, L., Duncan, J., Shen, H., Xu, Y., Magnusson, K. E. G., … Meijering, E. (2014). Objective comparison of particle tracking methods. *Nature Methods*, *11*(3), 281–289. https://doi.org/10.1038/nmeth.2808

Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, *158*(1), 9–14. https://doi.org/10.1016/0378-1119(95)00193-A

Cho, B.-K., Zengler, K., Qiu, Y., Park, Y. S., Knight, E. M., Barrett, C. L., Gao, Y., & Palsson, B. Ø. (2009). The transcription unit architecture of the Escherichia coli genome. *Nature Biotechnology*, *27*(11), 1043–1049. https://doi.org/10.1038/nbt.1582

Choi, P. J., Cai, L., Frieda, K., & Xie, X. S. (2008). A stochastic single molecule event triggers phenotype switching of a bacterial cell. *Science (New York, N.Y.)*, *322*(5900), 442. https://doi.org/10.1126/science.1161427

Chow, J., & Dennis, P. P. (1994). Coupling between mRNA synthesis and mRNA stability in Escherichia coli. *Molecular Microbiology*, *11*(5), 919–931. https://doi.org/10.1111/j.1365-2958.1994.tb00371.x

Chung, C. T., & Miller, R. H. (1993). [43] Preparation and storage of competent *Escherichia coli* cells. In R. Wu (Ed.), *Methods in Enzymology* (Vol. 218, pp. 621–627). Academic Press. https://doi.org/10.1016/0076-6879(93)18045-E

Condon, C., Liveris, D., Squires, C., Schwartz, I., & Squires, C. L. (1995). rRNA operon multiplicity in Escherichia coli and the physiological implications of rrn inactivation. *Journal of Bacteriology*, *177*(14), 4152–4156. https://doi.org/10.1128/jb.177.14.4152-4156.1995

Crawford, R., Torella, J. P., Aigrain, L., Plochowietz, A., Gryte, K., Uphoff, S., & Kapanidis, A. N. (2013). Long-Lived Intracellular Single molecule Fluorescence Using Electroporated Molecules. *Biophysical Journal*, *105*(11), 2439–2450. https://doi.org/10.1016/j.bpj.2013.09.057

Crick, F. H. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology*, *12*, 138–163.

Dar, D., Dar, N., Cai, L., & Newman, D. K. (2021). Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science (New York, N.Y.)*, *373*(6556), eabi4882. https://doi.org/10.1126/science.abi4882

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences*, *97*(12), 6640–6645. https://doi.org/10.1073/pnas.120163297

Deana, A., Celesnik, H., & Belasco, J. G. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature*, *451*(7176), 355–358. https://doi.org/10.1038/nature06475

Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M., & Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods*, *8*(12), 1027–1036. https://doi.org/10.1038/nmeth.1768

Dennis, P. P., Ehrenberg, M., Fange, D., & Bremer, H. (2009). Varying Rate of RNA Chain Elongation during rrn Transcription in Escherichia coli. *Journal of Bacteriology*, *191*(11), 3740–3746. https://doi.org/10.1128/jb.00128-09

Deschout, H., Zanacchi, F. C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S. T., & Braeckmans, K. (2014). Precisely and accurately localizing single emitters in fluorescence microscopy. *Nature Methods*, *11*(3), 253–266. https://doi.org/10.1038/nmeth.2843

Di Paolo, D., Afanzar, O., Armitage, J. P., & Berry, R. M. (2016). Single molecule imaging of electroporated dye-labelled CheY in live *Escherichia coli*. *Philosophical*

Transactions of the Royal Society B: Biological Sciences, *371*(1707), 20150492. https://doi.org/10.1098/rstb.2015.0492

DNA and RNA Molecular Weights and Conversions—UK. Retrieved 27 February 2025, from https://www.thermofisher.com/uk/en/home/references/ambion-tech-support/rna-tools-and-calculators/dna-and-rna-molecular-weights-and-conversions.html

Donnert, G., Keller, J., Medda, R., Andrei, M. A., Rizzoli, S. O., Lührmann, R., Jahn, R., Eggeling, C., & Hell, S. W. (2006). Macromolecular-scale resolution in biological fluorescence microscopy. *Proceedings of the National Academy of Sciences*, *103*(31), 11440–11445. https://doi.org/10.1073/pnas.0604965103

Dower, W. J., Miller, J. F., & Ragsdale, C. W. (1988). High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Research*, *16*(13), 6127–6145. https://doi.org/10.1093/nar/16.13.6127

Downing, W. L., & Dennis, P. P. (1987). Transcription products from the rplKAJL-rpoBC gene cluster. *Journal of Molecular Biology*, *194*(4), 609–620. https://doi.org/10.1016/0022-2836(87)90238-5

Dupuis, N. F., Holmstrom, E. D., & Nesbitt, D. J. (2014). Molecular-crowding effects on single molecule RNA folding/unfolding thermodynamics and kinetics. *Proceedings of the National Academy of Sciences*, *111*(23), 8464–8469. https://doi.org/10.1073/pnas.1316039111

Ebright, R. H. (2000). RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II. *Journal of Molecular Biology*, *304*(5), 687–698. https://doi.org/10.1006/jmbi.2000.4309

Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., & Nielsen, P. E. (1993). PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature*, *365*(6446), 566–568. https://doi.org/10.1038/365566a0

Elf, J., Li, G.-W., & Xie, X. S. (2007). Probing Transcription Factor Dynamics at the Single molecule Level in a Living Cell. *Science (New York, N.Y.)*, *316*(5828), 1191. https://doi.org/10.1126/science.1141967

Elitas, M., Kalayci Demir, G., & Vural Kaymaz, S. (2023). Mathematical Model for Growth and Rifampicin-Dependent Killing Kinetics of Escherichia coli Cells. *ACS Omega*, *8*(41), 38452–38458. https://doi.org/10.1021/acsomega.3c05233

Ellis, R. J. (2001). Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, *26*(10), 597–604. https://doi.org/10.1016/s0968-0004(01)01938-7

Endesfelder, U., Finan, K., Holden, S. J., Cook, P. R., Kapanidis, A. N., & Heilemann, M. (2013). Multiscale Spatial Organization of RNA Polymerase in Escherichia coli. *Biophysical Journal*, *105*(1), 172–181. https://doi.org/10.1016/j.bpj.2013.05.048

Eng, C.-H. L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., & Cai, L. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature*, *568*(7751), 235–239. https://doi.org/10.1038/s41586-019-1049-y

English, B. P., Hauryliuk, V., Sanamrad, A., Tankov, S., Dekker, N. H., & Elf, J. (2011). Single molecule investigations of the stringent response machinery in living

bacterial cells. *Proceedings of the National Academy of Sciences*, *108*(31), E365–E373. https://doi.org/10.1073/pnas.1102255108

Escherich, Th. (1988). The Intestinal Bacteria of the Neonate and Breast-Fed Infant. *Reviews of Infectious Diseases*, *10*(6), 1220–1225. https://doi.org/10.1093/clinids/10.6.1220

Escherichia coli str. K-12 substr. MG1655, complete genome (545778205; Version 3). (2025). [Dataset]. NCBI Nucleotide Database. http://www.ncbi.nlm.nih.gov/nuccore/U00096.3

Esquerré, T., Laguerre, S., Turlan, C., Carpousis, A. J., Girbal, L., & Cocaign-Bousquet, M. (2014). Dual role of transcription and transcript stability in the regulation of gene expression in Escherichia coli cells cultured on glucose at different growth rates. *Nucleic Acids Research*, *42*(4), 2460–2472. https://doi.org/10.1093/nar/gkt1150

Fan, J., El Sayyed, H., Pambos, O. J., Stracy, M., Kyropoulos, J., & Kapanidis, A. N. (2023). RNA polymerase redistribution supports growth in E. coli strains with a minimal number of rRNA operons. *Nucleic Acids Research*, *51*(15), 8085–8101. https://doi.org/10.1093/nar/gkad511

Farrar, A., Feehily, C., Turner, P., Zagajewski, A., Chatzimichail, S., Crook, D., Andersson, M., Oakley, S., Barrett, L., El Sayyed, H., Fowler, P. W., Nellåker, C., Kapanidis, A. N., & Stoesser, N. (2024). Infection Inspection: Using the power of citizen science for image-based prediction of antibiotic resistance in Escherichia coli treated with ciprofloxacin. *Scientific Reports*, *14*(1), 19543. https://doi.org/10.1038/s41598-024-69341-3

Farrar, A., Turner, P., El Sayyed, H., Feehily, C., Chatzimichail, S., Ta, S., Crook, D., Andersson, M., Oakley, S., Barrett, L., Nellåker, C., Stoesser, N., & Kapanidis, A. (2025). Ribosome phenotypes for rapid classification of antibiotic-susceptible and resistant strains of Escherichia coli. *Communications Biology*, *8*(1), 1–13. https://doi.org/10.1038/s42003-025-07740-6

Fei, J., & Sharma, C. M. (2018). RNA Localization in Bacteria. *Microbiology Spectrum*, *6*(5), 10.1128/microbiolspec.rwr-0024–2018. https://doi.org/10.1128/microbiolspec.rwr-0024-2018

Fei, J., Singh, D., Zhang, Q., Park, S., Balasubramanian, D., Golding, I., Vanderpool, C. K., & Ha, T. (2015). Determination of in vivo target search kinetics of regulatory noncoding RNA. *Science*, *347*(6228), 1371–1374. https://doi.org/10.1126/science.1258849

Femino, A. M., Fay, F. S., Fogarty, K., & Singer, R. H. (1998). Visualization of Single RNA Transcripts in Situ. *Science*, *280*(5363), 585–590. https://doi.org/10.1126/science.280.5363.585

Fernández-Suárez, M., & Ting, A. Y. (2008). Fluorescent probes for super-resolution imaging in living cells. *Nature Reviews Molecular Cell Biology*, *9*(12), 929–943. https://doi.org/10.1038/nrm2531

Franzini, R. M., & Kool, E. T. (2011). Improved Templated Fluorogenic Probes Enhance the Analysis of Closely Related Pathogenic Bacteria by Microscopy and Flow Cytometry. *Bioconjugate Chemistry*, *22*(9), 1869. https://doi.org/10.1021/bc2003567

Freestone, P., Trinei, M., Clarke, S. C., Nyström, T., & Norris, V. (1998). Tyrosine phosphorylation in *Escherichia coli*1. *Journal of Molecular Biology*, *279*(5), 1045–1051. https://doi.org/10.1006/jmbi.1998.1836

Fukuda, R., & Nagasawa-Fujimori, H. (1983). Mechanism of the rifampicin induction of RNA polymerase beta and beta' subunit synthesis in Escherichia coli. *Journal of Biological Chemistry*, *258*(4), 2720–2728. https://doi.org/10.1016/S0021-9258(18)32985-5

Furth, J. J., Hurwitz, J., & Anders, M. (1962). The Role of Deoxyribonucleic Acid in Ribonucleic Acid Synthesis. *Journal of Biological Chemistry*, *237*(8), 2611–2619. https://doi.org/10.1016/S0021-9258(19)73796-X

Gaal, T., Bratton, B. P., Sanchez-Vazquez, P., Sliwicki, A., Sliwicki, K., Vegel, A., Pannu, R., & Gourse, R. L. (2016). Colocalization of distant chromosomal loci in space in E. coli: A bacterial nucleolus. *Genes & Development*, *30*(20), 2272–2285. https://doi.org/10.1101/gad.290312.116

Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D., … Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Research*, *44*(D1), D133–D143. https://doi.org/10.1093/nar/gkv1156

Gao, Z., Zhang, W., Chang, R., Zhang, S., Yang, G., & Zhao, G. (2021). Liquid-Liquid Phase Separation: Unraveling the Enigma of Biomolecular Condensates in

Microbial Cells. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.751880

Gerber, A., Otterdijk, S. van, Bruggeman, F. J., & Tutucci, E. (2023). Understanding spatiotemporal coupling of gene expression using single molecule RNA imaging technologies. *Transcription*, *14*(3–5), 105. https://doi.org/10.1080/21541264.2023.2199669

Ghatak, S., King, Z. A., Sastry, A., & Palsson, B. O. (2019). The y-ome defines the 35% of *Escherichia coli* genes that lack experimental evidence of function. *Nucleic Acids Research*, *47*(5), 2446–2454. https://doi.org/10.1093/nar/gkz030

Gijtenbeek, L. A. van, Robinson, A., Oijen, A. M. van, Poolman, B., & Kok, J. (2016). On the Spatial Organization of mRNA, Plasmids, and Ribosomes in a Bacterial Host Overexpressing Membrane Proteins. *PLOS Genetics*, *12*(12), e1006523. https://doi.org/10.1371/journal.pgen.1006523

Golding, I., & Cox, E. C. (2004). RNA dynamics in live Escherichia coli cells. *Proceedings of the National Academy of Sciences*, *101*(31), 11310–11315. https://doi.org/10.1073/pnas.0404443101

Golding, I., Paulsson, J., Zawilski, S. M., & Cox, E. C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell*, *123*(6), 1025–1036. https://doi.org/10.1016/j.cell.2005.09.031

Gotta, S. L., Miller, O. L., & French, S. L. (1991). rRNA transcription rate in Escherichia coli. *Journal of Bacteriology*, *173*(20), 6647–6649. https://doi.org/10.1128/jb.173.20.6647-6649.1991

Griffith, J. (1978). DNA structure: Evidence from electron microscopy. *Science (New York, N.Y.)*, *201*(4355). https://doi.org/10.1126/science.663672

Grigorova, I. L., Phleger, N. J., Mutalik, V. K., & Gross, C. A. (2006). Insights into transcriptional regulation and σ competition from an equilibrium model of RNA polymerase binding to DNA. *Proceedings of the National Academy of Sciences*, *103*(14), 5332–5337. https://doi.org/10.1073/pnas.0600828103

Großmann, P., Lück, A., & Kaleta, C. (2017). Model-based genome-wide determination of RNA chain elongation rates in Escherichia coli. *Scientific Reports*, *7*(1), 17213. https://doi.org/10.1038/s41598-017-17408-9

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, *177*(14), 4121–4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995

Gyorfy, Z., Draskovits, G., Vernyik, V., Blattner, F. F., Gaal, T., & Posfai, G. (2015). Engineered ribosomal RNA operon copy-number variants of E. coli reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Research*, *43*(3), 1783–1794. https://doi.org/10.1093/nar/gkv040

Ha, T., & Tinnefeld, P. (2012). Photophysics of Fluorescence Probes for Single Molecule Biophysics and Super-Resolution Imaging. *Annual Review of Physical Chemistry*, *63*, 595. https://doi.org/10.1146/annurev-physchem-032210-103340

Hadizadeh Yazdi, N., Guet, C. C., Johnson, R. C., & Marko, J. F. (2012). Variation of the folding and dynamics of the Escherichia coli chromosome with growth conditions. *Molecular Microbiology*, *86*(6), 1318–1333. https://doi.org/10.1111/mmi.12071

Han, E. S., Cooper, D. L., Persky, N. S., Sutera, V. A., Jr, Whitaker, R. D., Montello, M. L., & Lovett, S. T. (2006). RecJ exonuclease: Substrates, products and interaction with SSB. *Nucleic Acids Research*, *34*(4), 1084–1091. https://doi.org/10.1093/nar/gkj503

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, *166*(4), 557–580. https://doi.org/10.1016/S0022-2836(83)80284-8

Harlos, K., & Eibl, H. (1980). Influence of calcium on phosphatidylglycerol. Two separate lamellar structures. *Biochemistry*, *19*(5), 895–899. https://doi.org/10.1021/bi00546a011

Hartmann, G., Honikel, K. O., Knüsel, F., & Nüesch, J. (1967). The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, *145*(3), 843–844. https://doi.org/10.1016/0005-2787(67)90147-5

Heintzmann, R., & Cremer, C. G. (1999). Laterally modulated excitation microscopy: Improvement of resolution by using a diffraction grating. *Optical Biopsies and Microscopic Techniques III*, *3568*, 185–196. https://doi.org/10.1117/12.336833

Heintzmann, R., & Huser, T. (2017). Super-Resolution Structured Illumination Microscopy. *Chemical Reviews*, *117*(23), 13890–13908. https://doi.org/10.1021/acs.chemrev.7b00218

Hell, S. W., & Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, *19*(11), 780–782. https://doi.org/10.1364/OL.19.000780

Herring, C. D., Glasner, J. D., & Blattner, F. R. (2003). Gene replacement without selection: Regulated suppression of amber mutations in *Escherichia coli*. *Gene*, *311*, 153–163. https://doi.org/10.1016/S0378-1119(03)00585-7

Ho, S. Y., & Mittal, G. S. (1996). Electroporation of Cell Membranes: A Review. *Critical Reviews in Biotechnology*, *16*(4), 349–362. https://doi.org/10.3109/07388559609147426

Holder, I. T., & Hartig, J. S. (2014). A Matter of Location: Influence of G-Quadruplexes on Escherichia coli Gene Expression. *Chemistry & Biology*, *21*(11), 1511–1521. https://doi.org/10.1016/j.chembiol.2014.09.014

Hovius, R., Guignet, E., Segura, J.-M., Joachim Piguet, & Horst Vogel. (2006). *Single Molecule Detection with Atto 647N NTA* [BioFiles: Sigma Aldrich]. https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/protein-biology/protein-quantitation/single molecule-detection

How to Choose a Label. Retrieved 11 March 2025, from https://www.atto-tec.com/support/how-to-choose-a-label-59/?utm_source=chatgpt.com&language=en

Howe, K. M., Newman, A. J., Garner, I., Wallis, A., & Hayward, R. S. (1982). Effect of rifampicin on expression of lacZ fused to promoters or terminators of the E.coli rpoBC operon. *Nucleic Acids Research*, *10*(22), 7425–7438. https://doi.org/10.1093/nar/10.22.7425

Hu, P., Janga, S. C., Babu, M., Díaz-Mejía, J. J., Butland, G., Yang, W., Pogoutse, O., Guo, X., Phanse, S., Wong, P., Chandran, S., Christopoulos, C., Nazarians-Armavil, A., Nasseri, N. K., Musso, G., Ali, M., Nazemof, N., Eroukova, V., Golshani, A., … Emili, A. (2009). Global Functional Atlas of Escherichia coli

Encompassing Previously Uncharacterized Proteins. *PLOS Biology*, *7*(4), e1000096. https://doi.org/10.1371/journal.pbio.1000096

Hurwitz, J., Bresler, A., & Diringer, R. (1960). The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. *Biochemical and Biophysical Research Communications*, *3*(1), 15–19. https://doi.org/10.1016/0006-291X(60)90094-2

Ilgu, M., Ray, J., Bendickson, L., Wang, T., Geraskin, I. M., Kraus, G. A., & Nilsen-Hamilton, M. (2016). Light-up and FRET aptamer reporters; evaluating their applications for imaging transcription in eukaryotic cells. *Methods*, *98*, 26–33. https://doi.org/10.1016/j.ymeth.2015.12.009

Irastortza-Olaziregi, M., & Amster-Choder, O. (2021). RNA localization in prokaryotes: Where, when, how, and why. *WIREs RNA*, *12*(2), e1615. https://doi.org/10.1002/wrna.1615

Ishihama, A. (2017). Building a complete image of genome regulation in the model organism *Escherichia coli*. *The Journal of General and Applied Microbiology*, *63*(6), 311–324. https://doi.org/10.2323/jgam.2017.01.002

Iyer, S., Park, B. R., & Kim, M. (2016). Absolute quantitative measurement of transcriptional kinetic parameters in vivo. *Nucleic Acids Research*, *44*(18), e142. https://doi.org/10.1093/nar/gkw596

Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*(3), 318–356. https://doi.org/10.1016/S0022-2836(61)80072-7

Jacobs, C., & Shapiro, L. (1999). Bacterial cell division: A moveable feast. *Proceedings of the National Academy of Sciences*, *96*(11), 5891–5893. https://doi.org/10.1073/pnas.96.11.5891

Jacobson, K., & Papahadjopoulos, D. (1975). Phase transitions and phase separations in phospholipid membranes induced by changes in temperature, pH, and concentration of bivalent cations. *Biochemistry*, *14*(1), 152–161. https://doi.org/10.1021/bi00672a026

Javer, A., Long, Z., Nugent, E., Grisi, M., Siriwatwetchakul, K., Dorfman, K. D., Cicuta, P., & Cosentino Lagomarsino, M. (2013). Short-time movement of E. coli chromosomal loci depends on coordinate and subcellular localization. *Nature Communications*, *4*(1), 3003. https://doi.org/10.1038/ncomms3003

Jin, D. J., Mata Martin, C., Sun, Z., Cagliero, C., & Zhou, Y. N. (2017). Nucleolus-like compartmentalization of the transcription machinery in fast-growing bacterial cells. *Critical Reviews in Biochemistry and Molecular Biology*, *52*(1), 96–106. https://doi.org/10.1080/10409238.2016.1269717

Jones, I. M., Primrose, S. B., Robinson, A., & Ellwood, D. C. (1981). Effect of growth rate and nutrient limitation on the transformability of Escherichia coli with plasmid deoxyribonucleic acid. *Journal of Bacteriology*, *146*(3), 841. https://doi.org/10.1128/jb.146.3.841-846.1981

Jun, S. (2015). Chromosome, Cell Cycle, and Entropy. *Biophysical Journal*, *108*(4), 785–786. https://doi.org/10.1016/j.bpj.2014.12.032

Kaberdin, V. R., & Bläsi, U. (2006). Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiology Reviews*, *30*(6), 967–979. https://doi.org/10.1111/j.1574-6976.2006.00043.x

Kaczanowska, M., & Rydén-Aulin, M. (2007). Ribosome Biogenesis and the Translation Process in Escherichia coli. *Microbiology and Molecular Biology Reviews*, *71*(3), 477–494. https://doi.org/10.1128/mmbr.00013-07

Kalwarczyk, T., Tabaka, M., & Holyst, R. (2012). Biologistics—Diffusion coefficients for complete proteome of Escherichia coli. *Bioinformatics*, *28*(22), 2971. https://doi.org/10.1093/bioinformatics/bts537

Kandhavelu, M., Mannerström, H., Gupta, A., Häkkinen, A., Lloyd-Price, J., Yli-Harja, O., & Ribeiro, A. S. (2011). In vivo kinetics of transcription initiation of the lar promoter in Escherichia coli. Evidence for a sequential mechanism with two rate-limiting steps. *BMC Systems Biology*, *5*(1), 149. https://doi.org/10.1186/1752-0509-5-149

Kannaiah, S., Livny, J., & Amster-Choder, O. (2019). Spatiotemporal Organization of the E. coli Transcriptome: Translation Independence and Engagement in Regulation. *Molecular Cell*, *76*(4), 574-589.e7. https://doi.org/10.1016/j.molcel.2019.08.013

Kapusta, P. GmbH, PicoQuant. *Absolute Diffusion Coefficients: Compilation of Reference Data for FCS Calibration*. Retrieved 27 February 2025, from https://www.picoquant.com/images/uploads/page/files/7353/appnote_diffusioncoefficients.pdf

Karp Pd, Paley S, Caspi R, Kothari A, Krummenacker M, Midford Pe, Moore Lr, Subhraveti P, Gama-Castro S, Tierrafria Vh, Lara P, Muñiz-Rascado L, Bonavides-Martinez C, Santos-Zavaleta A, Mackie A, Sun G, Ahn-Horst Ta, Choi H, Covert Mw, ... Paulsen I. (2023). The EcoCyc Database (2023). *EcoSal Plus*, *11*(1). https://doi.org/10.1128/ecosalplus.esp-0002-2023

Katz, Z. B., English, B. P., Lionnet, T., Yoon, Y. J., Monnier, N., Ovryn, B., Bathe, M., & Singer, R. H. (2016, January 13). *Mapping translation 'hot-spots' in live cells by tracking single molecules of mRNA and ribosomes*. eLife; eLife Sciences Publications Limited. https://doi.org/10.7554/eLife.10415

Ke, N., Landgraf, D., Paulsson, J., & Berkmen, M. (2016). Visualization of Periplasmic and Cytoplasmic Proteins with a Self-Labeling Protein Tag. *Journal of Bacteriology*, *198*(7), 1035–1043. https://doi.org/10.1128/jb.00864-15

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., & Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology*, *21*(1), 86–89. https://doi.org/10.1038/nbt765

Keseler, I. M., Gama-Castro, S., Mackie, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Kothari, A., Krummenacker, M., Midford, P. E., Muñiz-Rascado, L., Ong, W. K., Paley, S., Santos-Zavaleta, A., Subhraveti, P., Tierrafría, V. H., Wolfe, A. J., Collado-Vides, J., Paulsen, I. T., & Karp, P. D. (2021). The EcoCyc Database in 2021. *Frontiers in Microbiology*, *12*, 711077. https://doi.org/10.3389/fmicb.2021.711077

Kessler, L. F., Balakrishnan, A., Deußner-Helfmann, N. S., Li, Y., Mantel, M., Glogger, M., Barth, H.-D., Dietz, M. S., & Heilemann, M. (2023). Self-quenched Fluorophore Dimers for DNA-PAINT and STED Microscopy. *Angewandte Chemie (International Ed. in English)*, *62*(39), e202307538. https://doi.org/10.1002/anie.202307538

Kessler, L. F., Balakrishnan, A., Menche, T., Wang, D., Li, Y., Mantel, M., Glogger, M., Dietz, M. S., & Heilemann, M. (2024). Self-Quenched Fluorophore-DNA Labels

for Super-Resolution Fluorescence Microscopy. *The Journal of Physical Chemistry. B*, *128*(28), 6751–6759. https://doi.org/10.1021/acs.jpcb.4c02065

Khemici, V., Poljak, L., Luisi, B. F., & Carpousis, A. J. (2008). The RNase E of Escherichia coli is a membrane-binding protein. *Molecular Microbiology*, *70*(4), 799–813. https://doi.org/10.1111/j.1365-2958.2008.06454.x

Khlebnikov, A., Datsenko, K. A., Skaug, T., Wanner, B. L., & Keasling, J. D. (2001). Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology*, *147*(12), 3241–3247. https://doi.org/10.1099/00221287-147-12-3241

Klar, T. A., & Hell, S. W. (1999). Subdiffraction resolution in far-field fluorescence microscopy. *Optics Letters*, *24*(14), 954–956. https://doi.org/10.1364/OL.24.000954

Klumpp, S., & Hwa, T. (2008). Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proceedings of the National Academy of Sciences*, *105*(51), 20245–20250. https://doi.org/10.1073/pnas.0804953105

Kolmakov, K., Belov, V. N., Bierwagen, J., Ringemann, C., Müller, V., Eggeling, C., & Hell, S. W. (2010). Red-emitting rhodamine dyes for fluorescence microscopy and nanoscopy. *Chemistry (Weinheim an Der Bergstrasse, Germany)*, *16*(1), 158–166. https://doi.org/10.1002/chem.200902309

Koshkin, A. A., Singh, S. K., Nielsen, P., Rajwanshi, V. K., Kumar, R., Meldgaard, M., Olsen, C. E., & Wengel, J. (1998). LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic

acid recognition. *Tetrahedron*, *54*(14), 3607–3630. https://doi.org/10.1016/S0040-4020(98)00094-5

Kümmerlin, M., Zhao, Q., Hazra, J., Hepp, C., Farrar, A., Turner, P., & Kapanidis, A. N. (2025). *Tunable fluorogenic DNA probes drive fast and high-resolution single molecule fluorescence imaging* (p. 2025.01.21.634148). bioRxiv. https://doi.org/10.1101/2025.01.21.634148

Kushner Sidney R. (2002). mRNA Decay in Escherichia coli Comes of Age. *Journal of Bacteriology*, *184*(17), 4658–4665. https://doi.org/10.1128/jb.184.17.4658-4665.2002

LaCroix, R. A., Sandberg, T. E., O'Brien, E. J., Utrilla, J., Ebrahim, A., Guzman, G. I., Szubin, R., Palsson, B. O., & Feist, A. M. (2015). Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of Escherichia coli K-12 MG1655 on Glucose Minimal Medium. *Applied and Environmental Microbiology*, *81*(1), 17–30. https://doi.org/10.1128/AEM.02246-14

Lakowicz, J. R. (Ed.). (2006a). Mechanisms and Dynamics of Fluorescence Quenching. In *Principles of Fluorescence Spectroscopy* (pp. 331–351). Springer US. https://doi.org/10.1007/978-0-387-46312-4_9

Lakowicz, J. R. (Ed.). (2006b). Quenching of Fluorescence. In *Principles of Fluorescence Spectroscopy* (pp. 277–330). Springer US. https://doi.org/10.1007/978-0-387-46312-4_8

Lambert, T. *mVenus at FPbase*. FPbase. Retrieved 27 February 2025, from https://www.fpbase.org/protein/mvenus/

Lamoureux, C. R., Decker, K. T., Sastry, A. V., Rychel, K., Gao, Y., McConn, J. L., Zielinski, D. C., & Palsson, B. O. (2023). A multi-scale expression and

regulation knowledge base for Escherichia coli. *Nucleic Acids Research*, *51*(19), 10176–10193. https://doi.org/10.1093/nar/gkad750

Lawrence, J. B., & Singer, R. H. (1986). Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell*, *45*(3), 407–415. https://doi.org/10.1016/0092-8674(86)90326-0

Lawrence, J. B., Singer, R. H., & Marselle, L. M. (1989). Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell*, *57*(3), 493–502. https://doi.org/10.1016/0092-8674(89)90924-0

Le Nguyen, H., Duviau, M.-P., Laguerre, S., Nouaille, S., Cocaign-Bousquet, M., & Girbal, L. (2022). Synergistic regulation of transcription and translation in Escherichia coli revealed by co-directional increases in mRNA concentration and translation efficiency. *Microbiology Spectrum*, *10*(1), e02041-21. https://doi.org/10.1128/spectrum.02041-21

Le, T. T., Harlepp, S., Guet, C. C., Dittmar, K., Emonet, T., Pan, T., & Cluzel, P. (2005). Real-time RNA profiling within a single bacterium. *Proceedings of the National Academy of Sciences*, *102*(26), 9160–9164. https://doi.org/10.1073/pnas.0503311102

Leamy, K. A., Assmann, S. M., Mathews, D. H., & Bevilacqua, P. C. (2016). Bridging the gap between in vitro and in vivo RNA folding. *Quarterly Reviews of Biophysics*, *49*, e10. https://doi.org/10.1017/S003358351600007X

Lee, N., Francklyn, C., & Hamilton, E. P. (1987). Arabinose-induced binding of AraC protein to araI2 activates the araBAD operon promoter. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(24), 8814–8818. https://doi.org/10.1073/pnas.84.24.8814

Lelek, M., Gyparaki, M. T., Beliu, G., Schueder, F., Griffié, J., Manley, S., Jungmann, R., Sauer, M., Lakadamyali, M., & Zimmer, C. (2021). Single molecule localization microscopy. *Nature Reviews Methods Primers*, *1*(1), 1–27. https://doi.org/10.1038/s43586-021-00038-x

Lemke, J. J., Sanchez-Vazquez, P., Burgos, H. L., Hedberg, G., Ross, W., & Gourse, R. L. (2011). Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5712–5717. https://doi.org/10.1073/pnas.1019383108

Lewis, P. J., Thaker, S. D., & Errington, J. (2000). Compartmentalization of transcription and translation in Bacillus subtilis. *The EMBO Journal*, *19*(4), 710–718. https://doi.org/10.1093/emboj/19.4.710

Lill, U. I., & Hartmann, G. R. (1973). On the Binding of Rifampicin to the DNA-Directed RNA Polymerase from Escherichia coli. *European Journal of Biochemistry*, *38*(2), 336–345. https://doi.org/10.1111/j.1432-1033.1973.tb03066.x

Lim, X., Zhang, C., & Chen, X. (2024). Advances and applications of CRISPR/Cas-mediated interference in *Escherichia coli*. *Engineering Microbiology*, *4*(1), 100123. https://doi.org/10.1016/j.engmic.2023.100123

Lin, H.-H., Lin, C.-H., Hwang, S.-M., & Tseng, C.-P. (2012). High Growth Rate Downregulates fumA mRNA Transcription but Is Dramatically Compensated by Its mRNA Stability in Escherichia coli. *Current Microbiology*, *64*(5), 412–417. https://doi.org/10.1007/s00284-012-0087-6

Linn, T., & Scaife, J. (1978). Identification of a single promoter in E. coli for rplJ, rplL and rpoBC. *Nature*, *276*(5683), 33–37. https://doi.org/10.1038/276033a0

Loeff, L., Kerssemakers, J. W. J., Joo, C., & Dekker, C. (2021). *AutoStepfinder*: A fast and automated step detection method for single molecule analysis. *Patterns*, *2*(5), 100256. https://doi.org/10.1016/j.patter.2021.100256

Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., & Wood, K. V. (2008). HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology*, *3*(6), 373–382. https://doi.org/10.1021/cb800025k

Lubeck, E., & Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nature Methods*, *9*(7), 743–748. https://doi.org/10.1038/nmeth.2069

Ma, J.-C., Newman, A. J., & Hayward, R. S. (1981). Internal promoters of the rpoBC operon of Escherichia coli. *Molecular and General Genetics MGG*, *184*(3), 548–550. https://doi.org/10.1007/BF00352538

Magasanik, B. (1989). Regulation of transcription of the *glnALG* operon of *Escherichia coli* by protein phosphorylation. *Biochimie*, *71*(9), 1005–1012. https://doi.org/10.1016/0300-9084(89)90104-1

Mandel, M., & Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, *53*(1), 159–162. https://doi.org/10.1016/0022-2836(70)90051-3

Manley, S., Gillette, J. M., Patterson, G. H., Shroff, H., Hess, H. F., Betzig, E., & Lippincott-Schwartz, J. (2008). High-density mapping of single molecule

trajectories with photoactivated localization microscopy. *Nature Methods*, *5*(2), 155–157. https://doi.org/10.1038/nmeth.1176

Mao, S., Ying, Y., Wu, R., & Chen, A. K. (2020). Recent Advances in the Molecular Beacon Technology for Live-Cell Single molecule Imaging. *iScience*, *23*(12), 101801. https://doi.org/10.1016/j.isci.2020.101801

Markus Sauer, Johan Hofkens, & Jorg Enderlein. (2011). *Handbook of Fluorescence Spectroscopy and Imaging*. Wiley-VCH Verlag GmbH & Co. KGaA. DOI:10.1002/9783527633500

Mayer, B., Schwan, M., Oviedo-Bocanegra, L. M., Bange, G., Thormann, K. M., & Graumann, P. L. (2021). Dynamics of Bacterial Signal Recognition Particle at a Single Molecule Level. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.663747

Meyer, R. R., & Laine, P. S. (1990). The single-stranded DNA-binding protein of Escherichia coli. *Microbiological Reviews*, *54*(4), 342–380. https://doi.org/10.1128/mr.54.4.342-380.1990

Michalet, X., & Berglund, A. J. (2012). Optimal diffusion coefficient estimation in single-particle tracking. *Physical Review E*, *85*(6), 061916. https://doi.org/10.1103/PhysRevE.85.061916

Miller, O. L., Hamkalo, B. A., & Thomas, C. A. (1970). Visualization of Bacterial Genes in Action. *Science*, *169*(3943), 392–395. https://doi.org/10.1126/science.169.3943.392

Mirjam Kümmerlin. (2024). *Fluorogenic DNA Probes for Single molecule Fluorescence Imaging in vitro and in vivo*. University of Oxford.

Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., & Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. *eLife*, *5*, e13065. https://doi.org/10.7554/eLife.13065

Moffitt, J. R., & Zhuang, X. (2016). Chapter One—RNA Imaging with Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH). In G. S. Filonov & S. R. Jaffrey (Eds.), *Methods in Enzymology* (Vol. 572, pp. 1–49). Academic Press. https://doi.org/10.1016/bs.mie.2016.03.020

Mohapatra, S., & Weisshaar, J. C. (2018). Functional mapping of the E. coli translational machinery using single molecule tracking. *Molecular Microbiology*, *110*(2), 262–282. https://doi.org/10.1111/mmi.14103

Mondal, J., Bratton, B. P., Li, Y., Yethiraj, A., & Weisshaar, J. C. (2011). Entropy-Based Mechanism of Ribosome-Nucleoid Segregation in E. coli Cells. *Biophysical Journal*, *100*(11), 2605–2613. https://doi.org/10.1016/j.bpj.2011.04.030

Montero Llopis, P., Jackson, A. F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T., & Jacobs-Wagner, C. (2010). Spatial organization of the flow of genetic information in bacteria. *Nature*, *466*(7302), 77–81. https://doi.org/10.1038/nature09152

Mosteller, R. D., & Yanofsky, C. (1970). Transcription of the tryptophan operon in Escherichia coli: Rifampicin as an inhibitor of initiation. *Journal of Molecular Biology*, *48*(3), 525–531. https://doi.org/10.1016/0022-2836(70)90064-1

Muthukrishnan, A.-B., Kandhavelu, M., Lloyd-Price, J., Kudasov, F., Chowdhury, S., Yli-Harja, O., & Ribeiro, A. S. (2012). Dynamics of transcription driven by the tetA promoter, one event at a time, in live Escherichia coli cells. *Nucleic Acids Research*, *40*(17), 8472–8483. https://doi.org/10.1093/nar/gks583

Naryshkina, T., Mustaev, A., Darst, S. A., & Severinov, K. (2001). The β' Subunit of *Escherichia coli* RNA Polymerase Is Not Required for Interaction with Initiating Nucleotide but Is Necessary for Interaction with Rifampicin*. *Journal of Biological Chemistry*, *276*(16), 13308–13313. https://doi.org/10.1074/jbc.M011041200

Neidhardt, F. C., Bloch, P. L., & Smith, D. F. (1974). Culture Medium for Enterobacteria. *Journal of Bacteriology*, *119*(3), 736. https://doi.org/10.1128/jb.119.3.736-747.1974

Nelles, D. A., Fang, M. Y., O'Connell, M. R., Xu, J. L., Markmiller, S. J., Doudna, J. A., & Yeo, G. W. (2016). Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell*, *165*(2), 488–496. https://doi.org/10.1016/j.cell.2016.02.054

Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S., & Amster-Choder, O. (2011). Translation-Independent Localization of mRNA in E. coli. *Science*, *331*(6020), 1081–1084. https://doi.org/10.1126/science.1195691

Newman, A. J., Ma, J.-C., Howe, K. M., Garner, I., & Hayward, R. S. (1982). Evidence that rifampicin can stimulate readthrough of transcriptional terminators in Escherichia coli, including the attenuator of the rpoBC operon. *Nucleic Acids Research*, *10*(22), 7409–7424. https://doi.org/10.1093/nar/10.22.7409

Nielsen, P. E., Egholm, M., Berg, R. H., & Buchardt, O. (1991). Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science*, *254*(5037), 1497–1500. https://doi.org/10.1126/science.1962210

Nilsson, G., Belasco, J. G., Cohen, S. N., & von Gabain, A. (1984). Growth-rate dependent regulation of mRNA stability in Escherichia coli. *Nature*, *312*(5989), 75–77. https://doi.org/10.1038/312075a0

Norred, S. E., Caveney, P. M., Chauhan, G., Collier, L. K., Collier, C. P., Abel, S. M., & Simpson, M. L. (2018). Macromolecular Crowding Induces Spatial Correlations That Control Gene Expression Bursting Patterns. *ACS Synthetic Biology*, *7*(5), 1251–1258. https://doi.org/10.1021/acssynbio.8b00139

Atto 647 N BioReagent, suitable for fluorescence, ≥80% (HPCE) | Sigma-Aldrich. Retrieved 27 February 2025, from https://www.sigmaaldrich.com/GB/en/product/supelco/02175

Obika, S., Nanbu, D., Hari, Y., Morio, K., In, Y., Ishida, T., & Imanishi, T. (1997). Synthesis of 2'-*O*,4'-*C*-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -*endo* sugar puckering. *Tetrahedron Letters*, *38*(50), 8735–8738. https://doi.org/10.1016/S0040-4039(97)10322-7

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733-745. https://doi.org/10.1093/nar/gkv1189

Orosz, A., Boros, I., & Venetianer, P. (1991). Analysis of the complex transcription termination region of the Escherichia coli rrn B gene. *European Journal of Biochemistry*, *201*(3), 653–659. https://doi.org/10.1111/j.1432-1033.1991.tb16326.x

Overath, P., Brenner, M., Gulik-Krzywicki, T., Shechter, E., & Letellier, L. (1975). Lipid phase transitions in cytoplasmic and outer membranes of *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *389*(2), 358–369. https://doi.org/10.1016/0005-2736(75)90328-4

Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA Mimics of Green Fluorescent Protein. *Science*, *333*(6042), 642–646. https://doi.org/10.1126/science.1207339

Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., & Jacobs-Wagner, C. (2016). Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular Microbiology*, *99*(4), 767–777. https://doi.org/10.1111/mmi.13264

Palchevskiy, V., & Finkel, S. E. (2006). Escherichia coli Competence Gene Homologs Are Essential for Competitive Fitness and the Use of DNA as a Nutrient. *Journal of Bacteriology*, *188*(11), 3902–3910. https://doi.org/10.1128/jb.01974-05

Patrick, M., Dennis, P. P., Ehrenberg, M., & Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, *119*, 80–91. https://doi.org/10.1016/j.biochi.2015.10.015

Peabody, D. S. (1990). Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. A system for genetic analysis of a protein-RNA interaction. *Journal of Biological Chemistry*, *265*(10), 5684–5689. https://doi.org/10.1016/S0021-9258(19)39417-7

Peabody, D. S., & Ely, K. R. (1992). Control of translational repression by protein—Protein interactions. *Nucleic Acids Research*, *20*(7), 1649–1655. https://doi.org/10.1093/nar/20.7.1649

Peabody, D. S., & Lim, F. (1996). Complementation of RNA Binding Site Mutations in MS2 Coat Protein Heterodimers. *Nucleic Acids Research*, *24*(12), 2352–2359. https://doi.org/10.1093/nar/24.12.2352

Pellestor, F., & Paulasova, P. (2004). The peptide nucleic acids (PNAs), powerful tools for molecular genetics and cytogenetics. *European Journal of Human Genetics*, *12*(9), 694–700. https://doi.org/10.1038/sj.ejhg.5201226

Pernthaler, J., Glöckner, F.-O., Schönhuber, W., & Amann, R. (2001). Fluorescence *in situ* hybridization (FISH) with rRNA-targeted oligonucleotide probes. In *Methods in Microbiology* (Vol. 30, pp. 207–226). Academic Press. https://doi.org/10.1016/S0580-9517(01)30046-6

Petersen, C. (1989). Long-range translational coupling in the rplJL-rpoBC operon of Escherichia coli. *Journal of Molecular Biology*, *206*(2), 323–332. https://doi.org/10.1016/0022-2836(89)90482-8

Pham, P., Wood, E. A., Cox, M. M., & Goodman, M. F. (2023). RecA and SSB genome-wide distribution in ssDNA gaps and ends in Escherichia coli. *Nucleic Acids Research*, *51*(11), 5527–5546. https://doi.org/10.1093/nar/gkad263

Plochowietz, A., Crawford, R., & Kapanidis, A. N. (2014). Characterization of organic fluorophores for in vivo FRET studies based on electroporated molecules. *Physical Chemistry Chemical Physics*, *16*(25), 12688–12694. https://doi.org/10.1039/C4CP00995A

Plochowietz, A., Farrell, I., Smilansky, Z., Cooperman, B. S., & Kapanidis, A. N. (2017). In vivo single-RNA tracking shows that most tRNA diffuses freely in live bacteria. *Nucleic Acids Research*, *45*(2), 926–937. https://doi.org/10.1093/nar/gkw787

Post, L. E., Strycharz, G. D., Nomura, M., Lewis, H., & Dennis, P. P. (1979). Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit beta in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(4), 1697–1701. https://doi.org/10.1073/pnas.76.4.1697

Ralling, G., & Linn, T. (1984). Relative activities of the transcriptional regulatory sites in the rplKAJLrpoBC gene cluster of Escherichia coli. *Journal of Bacteriology*, *158*(1), 279–285. https://doi.org/10.1128/jb.158.1.279-285.1984

Rebecca Andrews. (2021). *Development of Single molecule DNA Hybridisation Tools for High-Content Analysis of Protein-DNA Interactions and Gene Expression*. University of Oxford.

Rems, L., Tang, X., Zhao, F., Pérez-Conesa, S., Testa, I., & Delemotte, L. (2022). Identification of electroporation sites in the complex lipid organization of the plasma membrane. *eLife*, *11*, e74773. https://doi.org/10.7554/eLife.74773

Reusch, R. N., Hiske, T. W., & Sadoff, H. L. (1986). Poly-beta-hydroxybutyrate membrane structure and its relationship to genetic transformability in Escherichia coli. *Journal of Bacteriology*, *168*(2), 553–562. https://doi.org/10.1128/jb.168.2.553-562.1986

Reyes-Lamothe, R., Tran, T., Meas, D., Lee, L., Li, A. M., Sherratt, D. J., & Tolmasky, M. E. (2014). High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. *Nucleic Acids Research*, *42*(2), 1042–1051. https://doi.org/10.1093/nar/gkt918

Reynolds, M. G. (2000). Compensatory Evolution in Rifampin-Resistant Escherichia coli. *Genetics*, *156*(4), 1471–1481. https://doi.org/10.1093/genetics/156.4.1471

Richards, J., Luciano, D. J., & Belasco, J. G. (2012). Influence of translation on RppH-dependent mRNA degradation in Escherichia coli. *Molecular Microbiology*, *86*(5), 1063–1072. https://doi.org/10.1111/mmi.12040

Robert J. Trachman Iii & Adrian R. Ferré-D'Amaré. (2019). Tracking RNA with light: Selection, structure, and design of fluorescence turn-on RNA aptamers. *Quarterly Reviews of Biophysics*, *52*, e8. https://doi.org/10.1017/S0033583519000064

Rodriguez, E. A., Campbell, R. E., Lin, J. Y., Lin, M. Z., Miyawaki, A., Palmer, A. E., Shu, X., Zhang, J., & Tsien, R. Y. (2017). The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends in Biochemical Sciences*, *42*(2), 111–129. https://doi.org/10.1016/j.tibs.2016.09.010

Rudkin, G. T., & Stollar, B. D. (1977). High resolution detection of DNA–RNA hybrids in situ by indirect immunofluorescence. *Nature*, *265*(5593), 472–473. https://doi.org/10.1038/265472a0

Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, *3*(10), 793–796. https://doi.org/10.1038/nmeth929

Salcedo-Sora, J. E., Jindal, S., O'Hagan, S., & Kell, D. B. (2021). A palette of fluorophores that are differentially accumulated by wild-type and mutant strains of Escherichia coli: Surrogate ligands for profiling bacterial membrane transporters. *Microbiology*, *167*(2). https://doi.org/10.1099/mic.0.001016

Sanamrad, A., Persson, F., Lundius, E. G., Fange, D., Gynnå, A. H., & Elf, J. (2014). Single-particle tracking reveals that free ribosomal subunits are not excluded from the *Escherichia coli* nucleoid. *Proceedings of the National Academy of Sciences*, *111*(31), 11413–11418. https://doi.org/10.1073/pnas.1411558111

Sando, S., Abe, H., & Kool, E. T. (2004). Quenched Auto-Ligating DNAs: Multicolor Identification of Nucleic Acids at Single Nucleotide Resolution. *Journal of the American Chemical Society*, *126*(4), 1081–1087. https://doi.org/10.1021/ja038665z

Santangelo, P. J., Nix, B., Tsourkas, A., & Bao, G. (2004). Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Research*, *32*(6), e57–e57. https://doi.org/10.1093/nar/gnh062

Sastry, A. V., Gao, Y., Szubin, R., Hefner, Y., Xu, S., Kim, D., Choudhary, K. S., Yang, L., King, Z. A., & Palsson, B. O. (2019). The Escherichia coli transcriptome mostly consists of independently regulated modules. *Nature Communications*, *10*(1), 5536. https://doi.org/10.1038/s41467-019-13483-w

Sathiamoorthy, S., & Shin, J. A. (2012). Boundaries of the Origin of Replication: Creation of a pET-28a-Derived Vector with p15A Copy Control Allowing Compatible Coexistence with pET Vectors. *PLOS ONE*, *7*(10), e47259. https://doi.org/10.1371/journal.pone.0047259

Sattler, L., & Graumann, P. L. (2021). Real-Time Messenger RNA Dynamics in Bacillus subtilis. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.760857

Saurabh, S., Perez, A. M., Comerci, C. J., Shapiro, L., & Moerner, W. E. (2016). Super-resolution Imaging of Live Bacteria Cells Using a Genetically Directed, Highly

Photostable Fluoromodule. *Journal of the American Chemical Society*, *138*(33), 10398–10401. https://doi.org/10.1021/jacs.6b05943

Schaechter, M., MaalØe, O., & Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *Microbiology*, *19*(3), 592–606. https://doi.org/10.1099/00221287-19-3-592

Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., & Drummen, G. P. C. (2019). Super-resolution microscopy demystified. *Nature Cell Biology*, *21*(1), 72–84. https://doi.org/10.1038/s41556-018-0251-8

Schibany, S., Kleine Borgmann, L. A. K., Rösch, T. C., Knust, T., Ulbrich, M. H., & Graumann, P. L. (2018). Single molecule tracking reveals that the bacterial SMC complex moves slowly relative to the diffusion of the chromosome. *Nucleic Acids Research*, *46*(15), 7805–7819. https://doi.org/10.1093/nar/gky581

Schnell, U., Dijk, F., Sjollema, K. A., & Giepmans, B. N. G. (2012). Immunolabeling artifacts and the need for live-cell imaging. *Nature Methods*, *9*(2), 152–158. https://doi.org/10.1038/nmeth.1855

Schröder, T., Scheible, M. B., Steiner, F., Vogelsang, J., & Tinnefeld, P. (2019). *Interchromophoric Interactions Determine the Maximum Brightness Density in DNA Origami Structures* (No. arXiv:1912.05261). arXiv. https://doi.org/10.48550/arXiv.1912.05261

Schwartz, J. J., & Quake, S. R. (2009). Single molecule measurement of the 'speed limit' of DNA polymerase. *Proceedings of the National Academy of Sciences of*

the United States of America, *106*(48), 20294–20299. https://doi.org/10.1073/pnas.0907404106

Sechkar, K., & Steel, H. (2025). Model-guided gene circuit design for engineering genetically stable cell populations in diverse applications. *Journal of The Royal Society Interface*, *22*(223), 20240602. https://doi.org/10.1098/rsif.2024.0602

Sepúlveda, L. A., Xu, H., Zhang, J., Wang, M., & Golding, I. (2016). Measurement of gene regulation in individual cells reveals rapid switching between promoter states. *Science*, *351*(6278), 1218–1222. https://doi.org/10.1126/science.aad0635

Severinov, K., Mooney, R., Darst, S. A., & Landick, R. (1997). Tethering of the Large Subunits of Escherichia coli RNA Polymerase. *Journal of Biological Chemistry*, *272*(39), 24137–24140. https://doi.org/10.1074/jbc.272.39.24137

Shaffer, S. M., Wu, M.-T., Levesque, M. J., & Raj, A. (2013). Turbo FISH: A Method for Rapid Single Molecule RNA FISH. *PLoS ONE*, *8*(9), e75120. https://doi.org/10.1371/journal.pone.0075120

Shao, B., Rammohan, J., Anderson, D. A., Alperovich, N., Ross, D., & Voigt, C. A. (2021). Single-cell measurement of plasmid copy number and promoter activity. *Nature Communications*, *12*(1), 1475. https://doi.org/10.1038/s41467-021-21734-y

Shao, X., Zhang, W., Umar, M. I., Wong, H. Y., Seng, Z., Xie, Y., Zhang, Y., Yang, L., Kwok, C. K., & Deng, X. (2020). RNA G-Quadruplex Structures Mediate Gene Regulation in Bacteria. *mBio*, *11*(1), 10.1128/mbio.02926-19. https://doi.org/10.1128/mbio.02926-19

Sharonov, A., & Hochstrasser, R. M. (2006). Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, *103*(50), 18911–18916. https://doi.org/10.1073/pnas.0609643104

Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M., & Keck, J. L. (2008). SSB as an Organizer/Mobilizer of Genome Maintenance Complexes. *Critical Reviews in Biochemistry and Molecular Biology*, *43*(5), 289–318. https://doi.org/10.1080/10409230802341296

Shimada, T., Yamazaki, Y., Tanaka, K., & Ishihama, A. (2014). The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of Escherichia coli. *PLOS ONE*, *9*(3), e90447. https://doi.org/10.1371/journal.pone.0090447

Shroff, H., Galbraith, C. G., Galbraith, J. A., & Betzig, E. (2008). Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nature Methods*, *5*(5), 417–423. https://doi.org/10.1038/nmeth.1202

Siibak, T., Peil, L., Xiong, L., Mankin, A., Remme, J., & Tenson, T. (2009). Erythromycin- and Chloramphenicol-Induced Ribosomal Assembly Defects Are Secondary Effects of Protein Synthesis Inhibition. *Antimicrobial Agents and Chemotherapy*, *53*(2), 563–571. https://doi.org/10.1128/aac.00870-08

Silverman, A. P., & Kool, E. T. (2005). Quenched autoligation probes allow discrimination of live bacterial species by single nucleotide differences in rRNA. *Nucleic Acids Research*, *33*(15), 4978–4986. https://doi.org/10.1093/nar/gki814

Sinha, S., & Redfield, R. J. (2012). Natural DNA Uptake by Escherichia coli. *PLoS ONE*, *7*(4), e35620. https://doi.org/10.1371/journal.pone.0035620

Skinner, S. O., Sepúlveda, L. A., Xu, H., & Golding, I. (2013). Measuring mRNA copy number in individual Escherichia coli cells using single molecule fluorescent in situ hybridization. *Nature Protocols*, *8*(6), 1100–1113. https://doi.org/10.1038/nprot.2013.066

So, L., Ghosh, A., Zong, C., Sepúlveda, L. A., Segev, R., & Golding, I. (2011). General properties of transcriptional time series in Escherichia coli. *Nature Genetics*, *43*(6), 554–560. https://doi.org/10.1038/ng.821

Socher, E., Bethge, L., Knoll, A., Jungnick, N., Herrmann, A., & Seitz, O. (2008). Low-Noise Stemless PNA Beacons for Sensitive DNA and RNA Detection. *Angewandte Chemie International Edition*, *47*(49), 9555–9559. https://doi.org/10.1002/anie.200803549

Son, M. S., & Taylor, R. K. (2021). Growth and Maintenance of Escherichia coli Laboratory Strains. *Current Protocols*, *1*(1), e20. https://doi.org/10.1002/cpz1.20

Song, W., Strack, R. L., Svensen, N., & Jaffrey, S. R. (2014). Plug-and-Play Fluorophores Extend the Spectral Properties of Spinach. *Journal of the American Chemical Society*, *136*(4), 1198–1201. https://doi.org/10.1021/ja410819x

Stacey, M., Fox, P., Buescher, S., & Kolb, J. (2011). Nanosecond pulsed electric field induced cytoskeleton, nuclear membrane and telomere damage adversely impact cell survival. *Bioelectrochemistry*, *82*(2), 131–134. https://doi.org/10.1016/j.bioelechem.2011.06.002

Steel, H., Habgood, R., Kelly, C. L., & Papachristodoulou, A. (2020). In situ characterisation and manipulation of biological systems with Chi.Bio. *PLOS Biology*, *18*(7), e3000794. https://doi.org/10.1371/journal.pbio.3000794

Stern, M. J., Ames, G. F.-L., Smith, N. H., Clare Robinson, E., & Higgins, C. F. (1984). Repetitive extragenic palindromic sequences: A major component of the bacterial genome. *Cell*, *37*(3), 1015–1026. https://doi.org/10.1016/0092-8674(84)90436-7

Steward, K. L., & Linn, T. (1991). In vivo analysis of overlapping transcription units in the rplKAJLrpoBC ribosomal protein—RNA polymerase gene cluster of Escherichia coli. *Journal of Molecular Biology*, *218*(1), 23–31. https://doi.org/10.1016/0022-2836(91)90870-C

Stoecker, K., Dorninger, C., Daims, H., & Wagner, M. (2010). Double Labeling of Oligonucleotide Probes for Fluorescence In Situ Hybridization (DOPE-FISH) Improves Signal Intensity and Increases rRNA Accessibility. *Applied and Environmental Microbiology*, *76*(3), 922–926. https://doi.org/10.1128/AEM.02456-09

Stracy, M., Lesterlin, C., Garza de Leon, F., Uphoff, S., Zawadzki, P., & Kapanidis, A. N. (2015). Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proceedings of the National Academy of Sciences*, *112*(32), E4390–E4399. https://doi.org/10.1073/pnas.1507592112

Strauss, M. T. (2022). Picasso-server: A community-based, open-source processing framework for super-resolution data. *Communications Biology*, *5*(1), 1–3. https://doi.org/10.1038/s42003-022-03909-5

Suenaga, H., Matsuzawa, T., & Sahara, T. (2022). Discovery by metagenomics of a functional tandem repeat sequence that controls gene expression in bacteria. *FEMS Microbiology Ecology*, *98*(4), fiac037. https://doi.org/10.1093/femsec/fiac037

Suksombat, S., Khafizov, R., Kozlov, A. G., Lohman, T. M., & Chemla, Y. R. (2015). Structural dynamics of E. coli single-stranded DNA binding protein reveal DNA wrapping and unwrapping pathways. *eLife*, *4*, e08193. https://doi.org/10.7554/eLife.08193

Sun, D. (2018). Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02154

Surovtsev, I. V., & Jacobs-Wagner, C. (2018). Subcellular Organization: A Critical Feature of Bacterial Cell Replication. *Cell*, *172*(6), 1271–1293. https://doi.org/10.1016/j.cell.2018.01.014

Sustarsic, M., Plochowietz, A., Aigrain, L., Yuzenkova, Y., Zenkin, N., & Kapanidis, A. (2014). Optimized delivery of fluorescently labeled proteins in live bacteria using electroporation. *Histochemistry and Cell Biology*, *142*(1), 113–124. https://doi.org/10.1007/s00418-014-1213-2

Tabaka, M., Kalwarczyk, T., Szymanski, J., Hou, S., & Holyst, R. (2014). The effect of macromolecular crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics*, *2*. https://doi.org/10.3389/fphy.2014.00054

Tabor, S., & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of*

the National Academy of Sciences, *82*(4), 1074–1078. https://doi.org/10.1073/pnas.82.4.1074

Taghbalout, A., & Rothfield, L. (2007). RNaseE and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. *Proceedings of the National Academy of Sciences*, *104*(5), 1667–1672. https://doi.org/10.1073/pnas.0610491104

Taketo, A. (1988). DNA transfection of *Escherichia coli* by electroporation. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, *949*(3), 318–324. https://doi.org/10.1016/0167-4781(88)90158-3

Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., & Xie, X. S. (2010). Quantifying E. coli Proteome and Transcriptome with Single molecule Sensitivity in Single Cells. *Science*, *329*(5991), 533–538. https://doi.org/10.1126/science.1188308

Tantale, K., Mueller, F., Kozulic-Pirher, A., Lesne, A., Victor, J.-M., Robert, M.-C., Capozi, S., Chouaib, R., Bäcker, V., Mateos-Langerak, J., Darzacq, X., Zimmer, C., Basyuk, E., & Bertrand, E. (2016). A single molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nature Communications*, *7*(1), 12248. https://doi.org/10.1038/ncomms12248

Tao, H., Bausch, C., Richmond, C., Blattner, F. R., & Conway, T. (1999). Functional Genomics: Expression Analysis of Escherichia coli Growing on Minimal and Rich Media. *Journal of Bacteriology*, *181*(20), 6425–6440. https://doi.org/10.1128/jb.181.20.6425-6440.1999

Technologies, B. *Dual-labeled BHQ probes—Performance to “dye” for—LGC*. Retrieved 7 March 2025, from

https://blog.biosearchtech.com/thebiosearchtechblog/bid/35814/dual-labeled-bhq-probes-performance-to-dye-for

Tokunaga, M., Imamoto, N., & Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single molecule imaging in cells. *Nature Methods*, *5*(2), 159–161. https://doi.org/10.1038/nmeth1171

Tsong, T. Y. (1990). On electroporation of cell membranes and some related phenomena. *Bioelectrochemistry and Bioenergetics*, *24*(3), 271–295. https://doi.org/10.1016/0302-4598(90)80028-H

Tutukina, M. N., Shavkunov, K. S., Masulis, I. S., & Ozoline, O. N. (2007). Intragenic promotor-like sites in the genome of Escherichia coli discovery and functional implication. *Journal of Bioinformatics and Computational Biology*, *5*(2B), 549–560. https://doi.org/10.1142/s0219720007002801

Tyagi, S., & Kramer, F. R. (1996). Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nature Biotechnology*, *14*(3), 303–308. https://doi.org/10.1038/nbt0396-303

Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D. J., & Kapanidis, A. N. (2013). Single molecule DNA repair in live bacteria. *Proceedings of the National Academy of Sciences*, *110*(20), 8063–8068. https://doi.org/10.1073/pnas.1301804110

Valencia-Burton, M., Shah, A., Sutin, J., Borogovac, A., McCullough, R. M., Cantor, C. R., Meller, A., & Broude, N. E. (2009). Spatiotemporal patterns and transcription kinetics of induced RNA in single bacterial cells. *Proceedings of the National Academy of Sciences*, *106*(38), 16399–16404. https://doi.org/10.1073/pnas.0907495106

Van Die, I. M., Bergmans, H. E. N., & Hoekstra, W. P. M. (1983). Transformation In Escherichia coli: Studies On The Role Of The Heat Shock In Induction Of Competence. *Microbiology*, *129*(3), 663–670. https://doi.org/10.1099/00221287-129-3-663

van Gijtenbeek, L. A., & Kok, J. (2017). Illuminating Messengers: An Update and Outlook on RNA Visualization in Bacteria. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.01161

van Helvoort J M, Kool J, & Woldringh C L. (1996). Chloramphenicol causes fusion of separated nucleoids in Escherichia coli K-12 cells and filaments. *Journal of Bacteriology*, *178*(14), 4289–4293. https://doi.org/10.1128/jb.178.14.4289-4293.1996

Vera, M., Tutucci, E., & Singer, R. H. (2019). Imaging Single mRNA Molecules in Mammalian Cells Using an Optimized MS2-MCP System. In Y. Shav-Tal (Ed.), *Imaging Gene Expression: Methods and Protocols* (pp. 3–20). Springer. https://doi.org/10.1007/978-1-4939-9674-2_1

Vogel, U., & Jensen, K. F. (1994). The RNA chain elongation rate in Escherichia coli depends on the growth rate. *Journal of Bacteriology*, *176*(10), 2807–2813. https://doi.org/10.1128/jb.176.10.2807-2813.1994

Vogel, U., Sørensen, M., Pedersen, S., Jensen, K. F., & Kilstrup, M. (1992). Decreasing transcription elongation rate in Escherichia Coli exposed to amino acid starvation. *Molecular Microbiology*, *6*(15), 2191–2200. https://doi.org/10.1111/j.1365-2958.1992.tb01393.x

Volkov, I. L., Seefeldt, A. C., & Johansson, M. (2019). Tracking of single tRNAs for translation kinetics measurements in chloramphenicol treated bacteria. *Methods*, *162–163*, 23–30. https://doi.org/10.1016/j.ymeth.2019.02.004

Wallner, G., Amann, R., & Beisker, W. (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry*, *14*(2), 136–143. https://doi.org/10.1002/cyto.990140205

Wang, S., Su, J.-H., Zhang, F., & Zhuang, X. (2016). An RNA-aptamer-based two-color CRISPR labeling system. *Scientific Reports*, *6*, 26857. https://doi.org/10.1038/srep26857

Weinstein, Z. B., & Zaman, M. H. (2018). Evolution of Rifampin Resistance in Escherichia coli and Mycobacterium smegmatis Due to Substandard Drugs. *Antimicrobial Agents and Chemotherapy*, *63*(1), e01243. https://doi.org/10.1128/AAC.01243-18

Widengren, J., Chmyrov, A., Eggeling, C., Löfdahl, P.-Å., & Seidel, C. A. M. (2007). Strategies to Improve Photostabilities in Ultrasensitive Fluorescence Spectroscopy. *The Journal of Physical Chemistry A*, *111*(3), 429–440. https://doi.org/10.1021/jp0646325

Wilson, D. N., & and Nierhaus, K. H. (2007). The Weird and Wonderful World of Bacterial Ribosome Regulation. *Critical Reviews in Biochemistry and Molecular Biology*, *42*(3), 187–219. https://doi.org/10.1080/10409230701360843

Woldringh, C. L. (2002). The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Molecular Microbiology*, *45*(1), 17–29. https://doi.org/10.1046/j.1365-2958.2002.02993.x

Woldringh, C. L., Jensen, P. R., & Westerhoff, H. V. (1995). Structure and partitioning of bacterial DNA: Determined by a balance of compaction and expansion forces? *FEMS Microbiology Letters*, *131*(3), 235–242. https://doi.org/10.1111/j.1574-6968.1995.tb07782.x

Wombacher, R., Heidbreder, M., van de Linde, S., Sheetz, M. P., Heilemann, M., Cornish, V. W., & Sauer, M. (2010). Live-cell super-resolution imaging with trimethoprim conjugates. *Nature Methods*, *7*(9), 717–719. https://doi.org/10.1038/nmeth.1489

Xia, C., Colognori, D., Jiang, X. S., Xu, K., & Doudna, J. A. (2025). Single molecule live-cell RNA imaging with CRISPR–Csm. *Nature Biotechnology*, 1–8. https://doi.org/10.1038/s41587-024-02540-5

Xu, X.-H., & Yeung, E. S. (1997). Direct Measurement of Single molecule Diffusion and Photodecomposition in Free Solution. *Science*, *275*(5303), 1106–1109. https://doi.org/10.1126/science.275.5303.1106

Yakhnin, A. V., Bubunenko, M., Mandell, Z. F., Lubkowska, L., Husher, S., Babitzke, P., & Kashlev, M. (2023). Robust regulation of transcription pausing in Escherichia coli by the ubiquitous elongation factor NusG. *Proceedings of the National Academy of Sciences*, *120*(24), e2221114120. https://doi.org/10.1073/pnas.2221114120

Yamamoto, M., & Nomura, M. (1979). Organization of genes for transcription and translation in the rif region of the Escherichia coli chromosome. *Journal of Bacteriology*, *137*(1), 584–594. https://doi.org/10.1128/jb.137.1.584-594.1979

Yamamoto, N., Nakahigashi, K., Nakamichi, T., Yoshino, M., Takai, Y., Touda, Y., Furubayashi, A., Kinjyo, S., Dose, H., Hasegawa, M., Datsenko, K. A., Nakayashiki, T., Tomita, M., Wanner, B. L., & Mori, H. (2009). Update on the Keio collection of Escherichia coli single-gene deletion mutants. *Molecular Systems Biology*, *5*(1), 335. https://doi.org/10.1038/msb.2009.92

Yan, Q., & Bruchez, M. P. (2015). Advances in chemical labeling of proteins in living cells. *Cell and Tissue Research*, *360*(1), 179–194. https://doi.org/10.1007/s00441-015-2145-4

Yan, X., Hoek, T. A., Vale, R. D., & Tanenbaum, M. E. (2016). Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell*, *165*(4), 976–989. https://doi.org/10.1016/j.cell.2016.04.034

Yang, S., Kim, S., Kim, D.-K., Jeon An, H., Bae Son, J., Hedén Gynnå, A., & Ki Lee, N. (2019). Transcription and translation contribute to gene locus relocation to the nucleoid periphery in E. coli. *Nature Communications*, *10*(1), 5131. https://doi.org/10.1038/s41467-019-13152-y

Yarbrough, L. R., Wu, F. Y. H., & Wu, C.-W. (1976). Molecular mechanism of the rifampicin-RNA polymerase interaction. *Biochemistry*, *15*(12), 2669–2676. https://doi.org/10.1021/bi00657a029

Yoshida, T., Obata, N., & Oosawa, K. (2000). Color-coding reveals tandem repeats in the *Escherichia coli* genome1. *Journal of Molecular Biology*, *298*(3), 343–349. https://doi.org/10.1006/jmbi.2000.3667

Young, R., & Bremer, H. (1976). Polypeptide-chain-elongation rate in Escherichia coli B/r as a function of growth rate. *Biochemical Journal*, *160*(2), 185–194. https://doi.org/10.1042/bj1600185

Yu, J., Xiao, J., Ren, X., Lao, K., & Xie, X. S. (2006). Probing Gene Expression in Live Cells, One Protein Molecule at a Time. *Science*, *311*(5767), 1600–1603. https://doi.org/10.1126/science.1119623

Zanetti-Domingues, L. C., Tynan, C. J., Rolfe, D. J., Clarke, D. T., & Martin-Fernandez, M. (2013). Hydrophobic Fluorescent Probes Introduce Artifacts into Single Molecule Tracking Experiments Due to Non-Specific Binding. *PLOS ONE*, *8*(9), e74200. https://doi.org/10.1371/journal.pone.0074200

Zengel, J. M., & Lindahl, L. (1994). *Diverse Mechanisms for Regulating Ribosomal Protein Synthesis in Escherichia coli* (W. E. Cohn & K. Moldave, Eds.; Vol. 47, pp. 331–370). Academic Press. https://doi.org/10.1016/S0079-6603(08)60256-1

Zhang, J., Fei, J., Leslie, B. J., Han, K. Y., Kuhlman, T. E., & Ha, T. (2015). Tandem Spinach Array for mRNA Imaging in Living Bacterial Cells. *Scientific Reports*, *5*(1), 17295. https://doi.org/10.1038/srep17295

Zhou, J., & Rudd, K. E. (2013). EcoGene 3.0. *Nucleic Acids Research*, *41*(D1), D613–D624. https://doi.org/10.1093/nar/gks1235

Zhou, Y. N., Lubkowska, L., Hui, M., Court, C., Chen, S., Court, D. L., Strathern, J., Jin, D. J., & Kashlev, M. (2013). Isolation and Characterization of RNA Polymerase rpoB Mutations That Alter Transcription Slippage during Elongation in Escherichia coli*. *Journal of Biological Chemistry*, *288*(4), 2700–2710. https://doi.org/10.1074/jbc.M112.429464

Zhu, M., Mori, M., Hwa, T., & Dai, X. (2019). Disruption of transcription–translation coordination in Escherichia coli leads to premature transcriptional termination. *Nature Microbiology*, *4*(12), 2347–2356. https://doi.org/10.1038/s41564-019-0543-1

Zimmerman, S. B. (2006). Shape and compaction of *Escherichia coli* nucleoids. *Journal of Structural Biology*, *156*(2), 255–261. https://doi.org/10.1016/j.jsb.2006.03.022

Zimmerman, S. B., & Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of Molecular Biology*, *222*(3), 599–620. https://doi.org/10.1016/0022-2836(91)90499-v