Developing a Method for Tracking RNA in Live *E. coli* Cells using fluorogenic DNA oligonucleotide probes.

Edward Wheeler
Wadham College
University of Oxford

A thesis submitted in fulfilment of the requirements for MSc by Research at the University of Oxford

September 2024

Acknowledgements

I would like to thank Achillefs Kapanidis for his guidance and supervision throughout the course of the Masters. I am very grateful for the opportunity to get involved with some biophysics over the past year.

Thank you for those who worked on *in vivo* FISH before and with me; namely Emma Lalande, Hafez el Sayyed and Rebecca Andrews. Particularly to Emma, for teaching me basic lab skills and providing plenty of data for analysis. Hafez and Emma provided all the strains to me throughout the year, without which the project would not have been possible. Hafez provided constant experimental guidance and some much needed positivity. Mirjam Kümmerlin helped me with understanding and design of fluorogenic probes. Piers Turner provided some great software and coding tips. Thank you to Tom Brown and Afaf El-Sagheer for synthesis of the 2xATTO647N probes.

I am grateful to my fellow lab members for constant entertainment; Oliver Pambos, Jacob Wright, Rasched Haidari, Alison Farrar, Emma Lalande, Mirjam Kümmerlin, Jagadish Prasad Hazra, Sammi Ta, Qing Zhao, Piers Turner, Stelios Chatzimichail, Hafez El Sayyed and Nicolas Shiaelis.

Figures were created with BioRender.com and OriginPro 2021.

Abstract

Super-resolution microscopy and the development of single molecule Fluorescence *In Situ* Hybridisation (smFISH) allows detection and localisation of individual molecules inside cells, deciphering intricate mechanistic details of various biological processes. FISH is well established in fixed cells, but there has been limited success in live cells, particularly in bacteria owing to their smaller size and lack of well-defined partitions. We aim to develop an *in vivo* smFISH technique which uses short single stranded fluorogenic DNA probes which are electroporated into live *Escherichia coli.* The spatial information from our study can potentially lead to visualising mRNA molecules throughout their lifetime, from transcription to translation to degradation.

Our proof of principle target consists of a short plasmid-expressed RNA, that contains six binding sites for the fluorogenic probe. The observed foci can be tracked and exhibit a wide range of behaviour; from immobility to rapidly diffusing spots, with many transitions between these states. A range of fluorogenic probes were tried with their *in vivo* behaviour characterised.

Despite the large proportion of non-specific interactions, it was observed that on a population level the mobility distributions of the tracks vary between the control and test samples. For this reason, it was concluded that the method has some degree of success and improvements to the design are suggested.

1. Introduction

RNA is a ubiquitous molecule to biochemistry. Modern research has revealed RNA to be a multifaceted molecule, playing a fundamental role in protein synthesis and catalysing reactions. In the past, scientists have studied RNA and its interactions on an ensemble level, but development of single molecule microscopy has the potential to reach a far higher understanding of the behaviour of RNA molecules. For example, single molecule tracking of messenger RNA (mRNA) will make it possible to visualise mRNA throughout its lifetime from transcription initiation, translation, transport and RNA degradation.

These processes require spatial and temporal control that, particularly in prokaryotes, remains poorly understood. Unlike eukaryotes, they do not contain membrane-bound organelles but there is increasing evidence that bacteria utilise spatial separation for a variety of functions (Guo *et al.*, 2024). This thesis will describe the development of a method to obtain this spatio-temporal information through tracking of mRNA using fluorogenic DNA oligonucleotide probes by *in vivo* single molecular fluorescence *in situ* hybridisation (smFISH).

1.1. Introduction to basic concepts

1.1.1. Fluorescence

Fluorescence occurs when a photon is absorbed, exciting an electron to a higher energy state. This excited state undergoes vibrational losses where a proportion of this energy is lost through non-radiative processes. The electron then relaxes to the ground state, emitting a photon. The shift between the absorbed wavelength and the emitted wavelength is called the Stokes' shift and is due to vibrational energy losses. Fluorescence has become a key biochemical tool, firstly with the harnessing of fluorescent proteins and now the development of organic fluorescent dyes. More details on the photophysics of fluorescence, can be found here (Lakowicz, 2006).

1.1.2. Organic Fluorophores

There has been significant progress developing small organic fluorophores over the past two decades. Organic fluorophores can be bonded to biological molecules to permit imaging of biological systems. Several organic fluorophores were used throughout this thesis; ATTO647N, ATTO655 and ATTO643. A related group of molecules is organic quenchers. These are organic molecules that are designed to interact with fluorophores to prevent fluorescence through a variety of mechanisms, discussed in **section 1.1.4**. Fluorophores have a lifetime which represents the average number of photons they absorb and emit before becoming permanently damaged, known as photobleaching. The fluorophores used in this thesis are all organic molecules with different photochemistry. This is particularly important when considering *in vivo* imaging, as different structures, functional groups or hydrophobicity may have different propensities for hybridisation or non-specific interactions. **Figure 1-1** shows the spectra of one such fluorophore, ATTO647N.

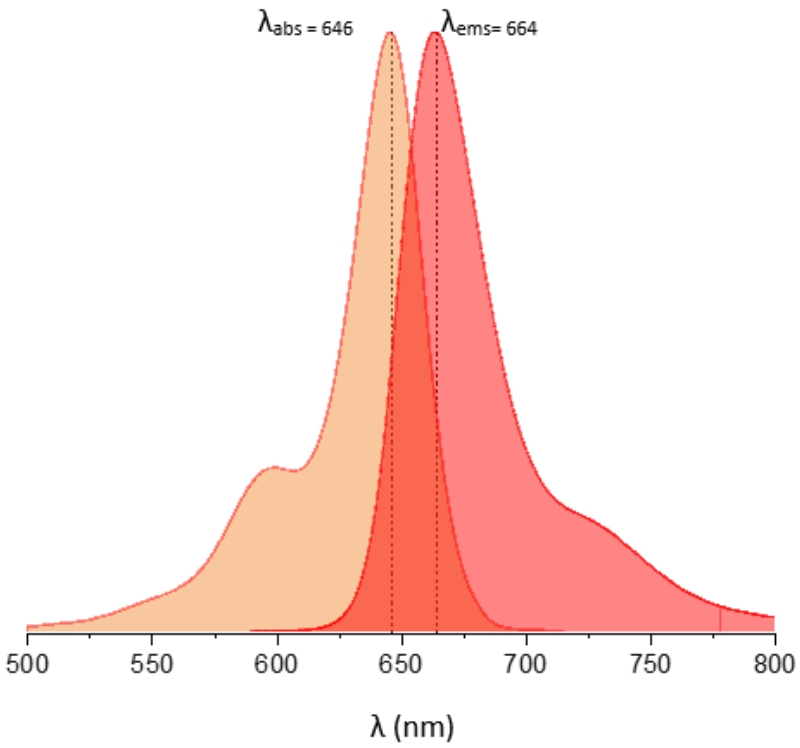

Figure 1-1: The absorption and emission spectra of ATTO647N, the dye used most frequently in this thesis. The absorption peak is at 646 nm, but due to Stokes' shift the emission peak is at 664 nm. The data are taken from (*ATTO 647N*).

1.1.3. Förster Resonance Energy Transfer (FRET)

Förster Resonance Energy Transfer (FRET) occurs between two molecules, through dipole interactions. Energy is transferred between a donor molecule that has been excited through absorbing a photon, and a nearby acceptor molecule. This interaction is highly length dependent, falling off with the sixth power of distance between the donor and the acceptor. Measurement of the FRET efficiency (*E*), as shown in equation (1), allows measurement of nanometer level distances. The acceptor may be a fluorophore itself and so emit a photon of a certain wavelength. This is often a green-red FRET pair, such as Cy3-ATTO647N. However, in this thesis, FRET was utilised between a fluorophore and a quencher which dissipates the energy non-radiatively.

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \qquad (1)$$

E is the FRET efficiency, *R* is the distance between the donor and acceptor and R_0 is the Förster radius, i.e. the distance where $E = 0.5$. R_0 depends on the absorption and emission spectra of the donor and acceptor, through what is known as the overlap integral. In simplified terms, for a fluorophore-quencher pair, the greater the overlap between the emission spectra and the absorption spectra, the greater the FRET efficiency and therefore quenching. Full details can be found here (Wu and Brand, 1994).

1.1.4. Quenching

Quenching is a process which reduces the photons emitted by a fluorophore. This is essential in single molecule imaging as a mechanism to prevent background signal when the fluorophore is not interacting with its target. The quenching mechanisms relevant to this thesis are static, collisional and FRET quenching. Static quenching occurs when a fluorophore forms an interaction with another molecule, forming a non-fluorescent complex. This molecule may be another fluorophore, or a quencher. The non-fluorescent complex is not able to absorb the incident photon and so no signal is observed (Lakowicz, 2006). Collisional

quenching is a similar process, where an excited fluorophore interacts directly with another molecule which leads to the excited electron entering a triplet state. This then relaxes to the ground state without emitting the expected photon. Finally, FRET quenching occurs between a FRET pair; a fluorophore and a quencher. As discussed, the quenching efficiency is distance dependent but is also dependent on overlap between the spectra of the fluorophore and the quencher. Informed choice of a fluorophore-quencher pair is essential to ensure effective imaging (Lakowicz, 2006).

1.1.5. Imaging

Imaging was conducted on an Oxford Nanoimaging microscope known as the Nanoimager. Throughout this thesis only Highly Inclined and Laminated Optical sheet (HILO) and bright-field microscopy was used. HILO involves illuminating the sample slightly below the total internal reflection angle. HILO reduces background signal when compared to epi-fluorescence microscopy, but also illuminates a larger region of the sample than Total Internal Reflection Fluorescence microscopy (TIRF). *Escherichia coli* extend > 1 μm in the z-direction, so TIRF microscopy would not provide illumination throughout the cell. Bright field microscopy involves direct illumination of the sample and was used simply to observe the positions of the cells before excitation of the sample.

1.2. Existing methods for RNA Localisation and Tracking

Live tracking of RNA in cells provides spatio-temporal information of RNA dynamics, which in turn informs on transcription and translation, two of the most fundamental biochemical processes. In recent decades there have been increasing attempts to characterise RNA localisation and dynamics using a variety of different techniques. However, they all follow the same basic pathway: a molecular probe, a complementary target, imaging methods and analysis of the images produced. Effective methods and choices are required at each level of this pipeline to reach accurate conclusions. In this section I shall discuss the methods adopted in the literature and the conclusions they were able to make as a result.

1.2.1. Fluorescent Protein Systems: MS2 and Cas

MS2-GFP system:

The most widely used RNA labelling method uses the MS2 bacteriophage coat protein (MCP). MCP contains an RNA binding site with high affinity to a stem loop secondary structure in the RNA (Peabody, 1993). Typically, multiple MCP binding hairpins are engineered into the RNA of interest in the 3' untranslated region in order to limit the effect of the labelling on native mRNA behaviour. Fusing a fluorescent protein, such as GFP, to MCP allows imaging of the RNA. This method was developed by Singer and colleagues in yeast (Bertrand *et al.*, 1998). This particular hairpin motif is not common in mammalian cells and therefore results in MCP binding with high specificity. This system has been used successfully to study eukaryotic physiology, one example is the role of RNA transport in oocyte development, revealing the role of the cytoskeleton in RNA transport (Forrest and Gavis, 2003).

This system was first developed in bacterial cells with some modifications, using 96 hairpin repeats and slightly modified MCP and GFP proteins (Golding and Cox, 2004). They were able to localise and track MCP-GFP fusions in live cells, studying mobility and chain elongation of the RNA. Further studies have improved on the tracking and spatial distributions of the RNA, (Gupta *et al.*, 2014, Nevo-Dinur *et al.*, 2011, Gijtenbeek *et al.*, 2016). This has allowed inference into the physiology of bacteria, such as the localisation of translation, degradation and the related machinery within the cell. Since 2004, the improvements in

imaging techniques and fluorophore development have allowed for fewer MS2 hairpins required to resolve single RNA molecules, being reduced from 96 to only two (Sattler and Graumann, 2021).

Despite the success of the MS2-GFP system, one major drawback is the presence of a high background from the unbound fluorescent proteins. A method to improve specificity is to use a split fluorophore that only fluoresces when binding in tandem with another split fluorophore. This is achieved by fusing two non-fluorescent elements of a fluorophore to two different RNA binding proteins. When these proteins bind to adjacent motifs on the RNA they recover fluorescence. This idea was utilised to study binding of proteins, and was subsequently adapted for RNA using the PP7 bacteriophage coat protein in tandem with MCP (Wu, Chen and Singer, 2014, Park, Moon and Park, 2020). This system and the MS2 system are shown in a schematic in **Figure 1-2**.

There have been several derivatives of this method emerging over the past few decades. For example, an orthogonal system using the N-protein of the λ bacteriophage which only requires short 15 nucleotide hairpins and a 22 amino acid peptide (Daigle and Ellenberg, 2007, Lange *et al.*, 2008). Advances in protein labelling to include chemical tags has also been harnessed, which may circumvent various issues with using fluorescent proteins (Carrocci and Hoskins, 2014).

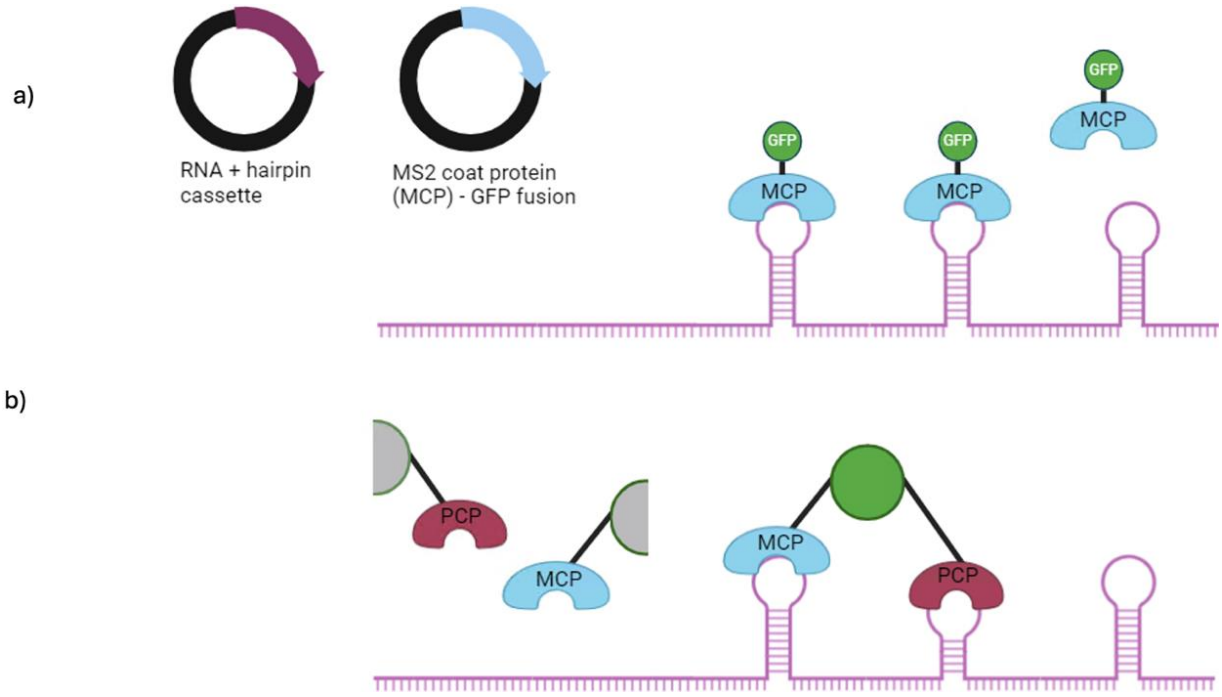

Figure 1-2: RNA tracking techniques using bacteriophage coat proteins. **a)** MS2 coat protein (MCP) fused to a fluorescent protein is expressed from a plasmid. The RNA of interest is expressed from another plasmid and the MCP binds to the engineered hairpins. This has been implemented successfully for RNA tracking. **b)** A more involved system using alternating hairpins with affinities to MCP and PP7 coat protein (PCP) respectively. Each coat protein is fused to a non-fluorescent region of a fluorescent protein which recovers fluorescence when brought into close proximity with its counterpart on the RNA. This will require three plasmids (not shown), one for each of the coat proteins and one for the RNA.

Cas Protein system:

CRISPR-Cas systems have huge potential for DNA targeting, editing and engineering. An aspect of this is the Cas proteins, which recognise and bind to a specific target. Catalytically inactive Cas9 (dCas9) recognises and binds to a specific RNA motif, but lacks the usual function of cleaving the RNA. A dCas9-GFP fusion was harnessed to bind to a single guide RNA (sgRNA) which allowed tracking of the sgRNA without effecting the native physiology (Nelles *et al.*, 2016; Sun *et al.*, 2020). This has been done in conjunction with the MS2 system for confirmation. The complexity of developing the sgRNA motifs may have limited the development of this technique (George *et al.*, 2018).

Other Cas proteins have been used for similar purposes. Cas13 can bind directly to single stranded RNA (ssRNA), unlike Cas9, which significantly reduces the level of engineering required to track the RNA (Abudayyeh *et al.*, 2017, Wang, Yang and Chen, 2020). As with the MS2 system, the presence of freely diffusing fluorescent proteins will produce a large background signal. This may be overcome by engineering multiple binding sites into the target RNA. Similar to the MS2/PP7 system described earlier a switchable fluorescent protein has been developed. A Cas6 protein fused to two non-fluorescent elements of the fluorescent protein mVenus was shown to recover fluorescence upon Cas6 binding to the RNA and bringing the mVenus elements into close proximity due to a conformational change in the Cas protein (Gao *et al.*, 2022). These methods highlight the potential of using Cas proteins to track RNA in a relatively non-obstructive way, with low background signal. This system has been focused on eukaryotic cells so far but should be transferable to prokaryotes. A schematic showing the fundamentals of the Cas protein system is shown below, in **Figure 1-3**.

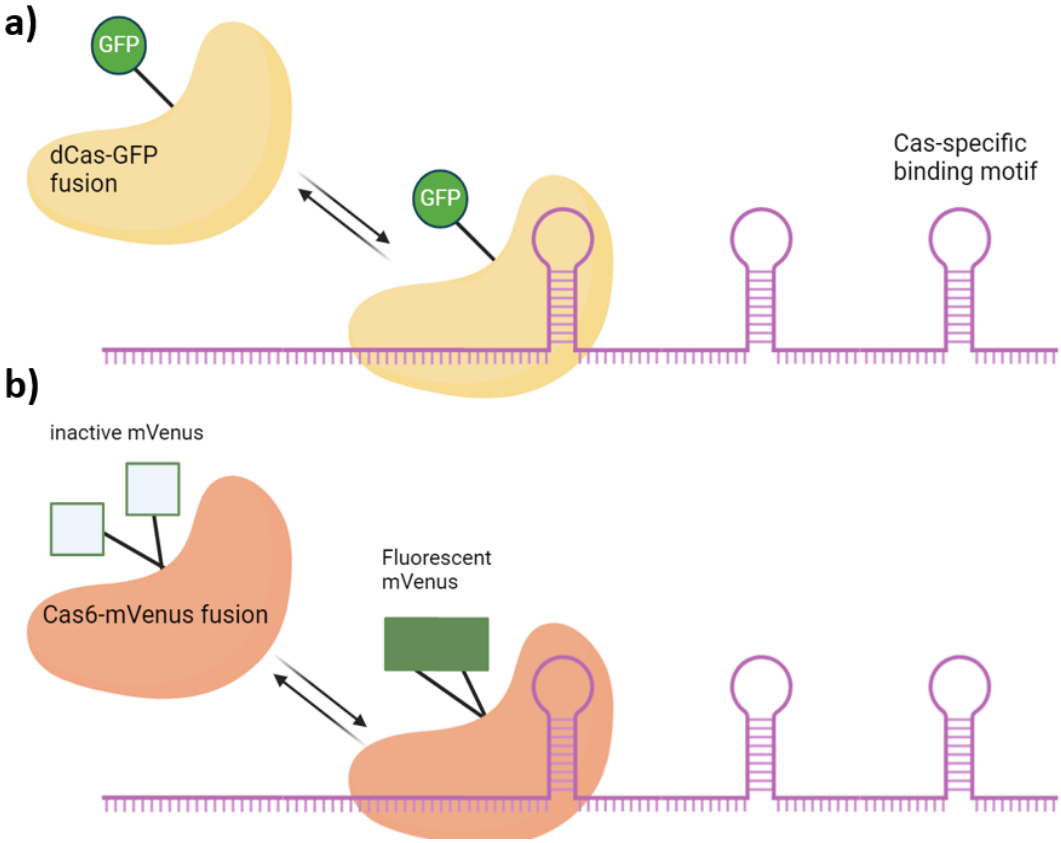

Figure 1-3: **a)** Binding of a Cas-fluorescent protein fusion to a specific RNA motif has potential for tracking of the RNA with high specificity. **B)** A mechanism to reduce the background uses non-fluorescent mVenus elements that recover fluorescence when in close proximity due to a conformational change in the Cas protein which occurs during binding to the RNA.

Together, these fluorescent protein systems provide a robust and repeatable way of localising and tracking RNA in both eukaryotic and prokaryotic cells and have led this field since its development. The Cas based systems are less common in the literature. However, there are several limitations. When designing the experiments, the requirement for engineering hairpins into each RNA of interest may be limiting in terms of throughput. There is some element of fine-tuning in terms of the dual plasmids required to produce the RNA binding protein-fluorescent protein fusion, and the target RNA itself at the right levels. Although recent developments have allowed the number of hairpins to be reduced significantly, they may still interfere with native processes particularly through repetitive RNA motifs and binding of proteins to an RNA may reduce mobility and hinder natural RNA physiology. This may happen by preventing the binding of other RNA binding proteins which are crucial in translation, localisation and degradation. Additionally, the signal to noise ratio (SNR) is typically low due to unbound fluorescent proteins. Finally, to image multiple RNAs, different fusions and hairpins will be required which limits multiplexing.

1.2.2. Organic Fluorophore-based techniques

A great benefit of using fluorescent proteins is the high specificity they achieve. However, they have limited photophysical properties when compared to modern organic fluorophores, e.g. they typically have a lower quantum yield and are more susceptible to bleaching and blinking (Elf and Barkefors, 2019).

There has been significant progress in characterisation and quality of organic fluorophores with a huge range of photochemistry. They can also be readily combined with specifically designed quenchers. A combination of highly effective fluorescent molecules and quenching mechanisms means organic fluorophores are versatile biological imagers.

<u>Fluorescent Aptamer Systems:</u>

An aptamer is a region of single stranded RNA that forms a specific secondary structure which has an affinity to bind to a range of targets. Aptamers can be designed to bind to fluorescent molecules. This will allow for localisation and tracking of an RNA that has been engineered to contain such an aptamer, typically in the 3' untranslated region. The Spinach system was the first to be used effectively in live cells (Paige, Wu and Jaffrey, 2011). The choice of fluorophore

was 4-hydroxybenzlidene imidazolinone (HBI), a fluorescent molecule chemically similar to GFP. This was chosen due to its specificity and quenching behaviour. When HBI is freely diffusing or loosely bound to a cellular component the fluorophore can dissipate energy through molecular motions rather than photon emission. Only when tightly bound to the aptamer is fluorescence seen. An aptamer had to be designed which has a high affinity to bind the fluorophore. This was done using a directed evolution method: Systematic Evolution of Ligands by Exponential Enrichment (SELEX). This success was a proof of concept for aptamer systems in live cells.

Since Spinach there have been a multitude of different fluorophores produced with various enhanced properties. These include Peppers (Chen *et al.*, 2019) and a selection of Mango aptamers (Autour *et al.*, 2018), among many more. For a recent review on the myriad of fluorophores, see (Chen *et al.*, 2023).

A similar method makes use of a fluorophore-quencher system. This allows a larger range of fluorophores to be used with aptamers. Contact quenching was used effectively for a range of fluorophores with a remarkably high SNR (Sunbul and Jäschke, 2013). A major issue with the use of aptamers is that a lengthy selection process must be completed for each fluorophore. If an aptamer can be designed that has a high affinity to a quencher, which is an effective quencher for many fluorophores, then this could allow imaging with minimal engineering (Arora, Sunbul and Jäschke, 2015). This system and a general aptamer system schematic are shown in **Figure 1-4**.

Despite the obvious potential for aptamer systems, they have faced some limitations, notably the high level of engineering/selection required and cytotoxicity of fluorophores for *in vivo* imaging (Fei and Sharma, 2018). This also makes multiplexing more challenging, if several orthogonal systems are required simultaneously.

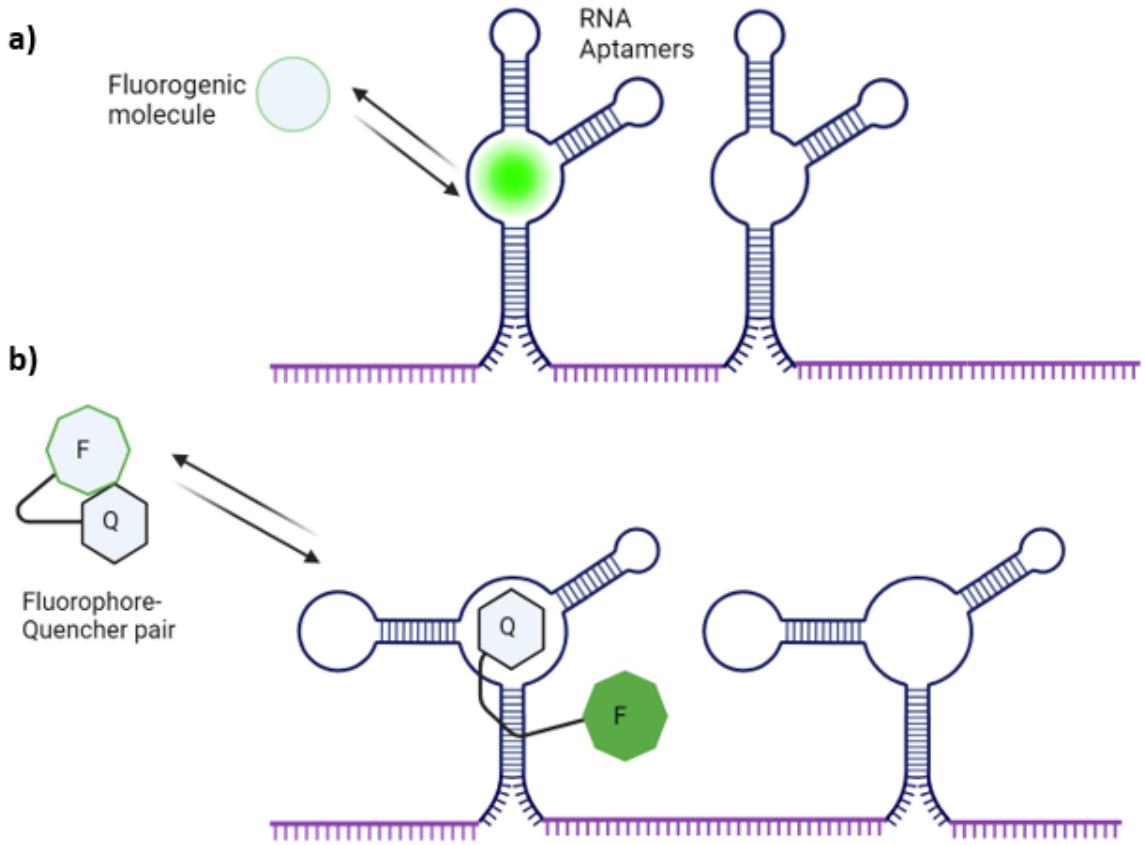

Figure 1-4: A generic aptamer design schematic. **a)** A fluorogenic molecules which binds tightly to the aptamer, recovering fluorescence. This is because the fluorophore cannot undergo certain non-radiative transitions when bound to the aptamer. **b)** A similar method using linkage of a fluorophore-quencher pair opens this method to a wide range of organic dyes. The dye becomes fluorescent when the contact quenching is removed due to the high affinity of the quencher to the aptamer. When the quencher is bound to the aptamer the average distance between the fluorophore and quencher increases, reducing the FRET efficiency.

Fluorescent *In Situ* Hybridisation (FISH):

FISH has been the gold standard for localisation of DNA and RNA sequences, and was first developed over 40 years ago (Bauman *et al.*, 1980). FISH is the hybridisation of DNA or RNA probes, conjugated to a fluorescent molecule, to their target sequences. This allows for fluorescent imaging of the target sequence. FISH proved a remarkably versatile technique, used across many cell types and applications. For a comprehensive review, see (Shakoori, 2017).

The development of single molecule FISH (smFISH) allowed individual RNA molecules to be observed in cells through binding of multiple probes to the same molecule (Femino *et al.*, 1998). Information on the initiation, elongation and localisation of single RNA transcripts was available for the first time. However, FISH requires fixation of the cells, typically with strong chemicals such as formaldehyde. This makes live cell tracking of RNA impossible with FISH, and so the precise dynamics of the RNA are lost. Furthermore, the fixation process may introduce artefacts. The demand to overcome these issues lead to the development of live cell FISH, using probes called molecular beacons.

Molecular Beacons

Molecular beacons (MBs) are short, doubly labelled oligonucleotides first introduced in 1996 (Tyagi and Kramer, 1996). The 5' end was labelled with an organic dye, and the 3' is labelled with a quencher. When in solution, the MBs form a hairpin/stem structure which brings the dye and quencher into close proximity, reducing fluorescence. See **Figure 1-5** for an example structure of a MB. When the MB hybridises to its complementary target the rigidity of the double stranded nucleic acid increases the average distance between the dye and quencher. Contact and FRET quenching are reduced, and fluorescence recovered. This simple design has potential for specific, high signal to noise imaging and has been implemented in a variety of settings, cell types and fields. Some examples *in vivo* successes include: monitoring gene expression levels in live cells (Medley *et al.*, 2005, Peng *et al.*, 2005) , RNA tracking (Vargas *et al.*, 2005, Turner-Bridger *et al.*, 2018), virus detection (Santangelo *et al*, 2006, Yeh *et al.*, 2008), developmental biology (Kang et al., 2011, Bratu *et al.*, 2003) and cell type detection (Jha *et al.*, 2015).

Like standard FISH, molecular beacons can be used to image RNA on the single molecule level. This is achieved through the binding of multiple probes to the same transcript

through engineering: addition of repeat target sequences to the RNA, or by designing multiple MBs that hybridise to different, endogenous regions on the RNA. This was initially explored with addition of 96 MB binding sites to a GFP mRNA (Vargas *et al.*, 2005). This allowed for tracking of the mRNA and novel insight into the transport of RNA in the nucleus. Advances in microscopy and fluorophore design have allowed for imaging with fewer than 8 binding sites (Chen *et al.*, 2017). It is important to reduce the number of sites required in order to limit the effect of the hybridisation and/or engineered tag on the behaviour of the transcript. For example, the same paper showed a significant (~ 1.5x) increase in transcript mobility when moving from 32 to 8 repeat targets.

The main limitation of MBs is the high tendency for non-specific interactions with intracellular components. This may be through interactions with similar nucleic acid sequences, proteins, or other cytoplasmic components. It is essential that the signal from the target RNA is distinguishable from these non-specific signals and that rigorous controls are carried out to ensure so. It has become standard practice to design MBs that are at least partially constructed from modified nucleic acids. The chemical modifications reduce the binding affinity to a variety of cellular components, with a particular focus on DNAses and RNAses (Zhao *et al.*, 2016; Bratu *et al.*, 2003; Chen *et al.*, 2017; Chan, Lim and Wong, 2006). I will not discuss the chemistry of the modifications here but they are covered clearly in (Mao *et al.*, 2020). Careful design of the sequence, hairpin structure and dye are essential for effective imaging. For example, the annealing temperature of the hairpin stem much be such that the probe will unquench and hybridise at physiological temperatures, and the hybridising region must be as unique as possible. This last point can be confirmed with a Basic Local Alginment Search Tool (BLAST). Another approach uses dual-FRET MBs. These probes are designed to bind adjacent regions of the target RNA, meaning that a change in FRET efficiency can be seen when both are bound. This improves the ability to detect the RNA from non-specific interactions.

Due to their small size, MBs are simple to deliver into live cells without causing damage to the membrane. Typical methods include microporation, microinjection or pore-producing bacterial toxins. Again, for more information on this see (Mao *et al.*, 2020).

To summarise, MBs have a huge potential for RNA imaging in live cells, provided the non-specific interactions are not too numerous. Developing resolvable single molecule MBs has the potential for accurate imaging of mRNA throughout its life cycle. Furthermore, MBs

have not yet been used in bacteria, to my knowledge. An MB and dual-FRET MB schematic is shown in **Figure 1-5**.

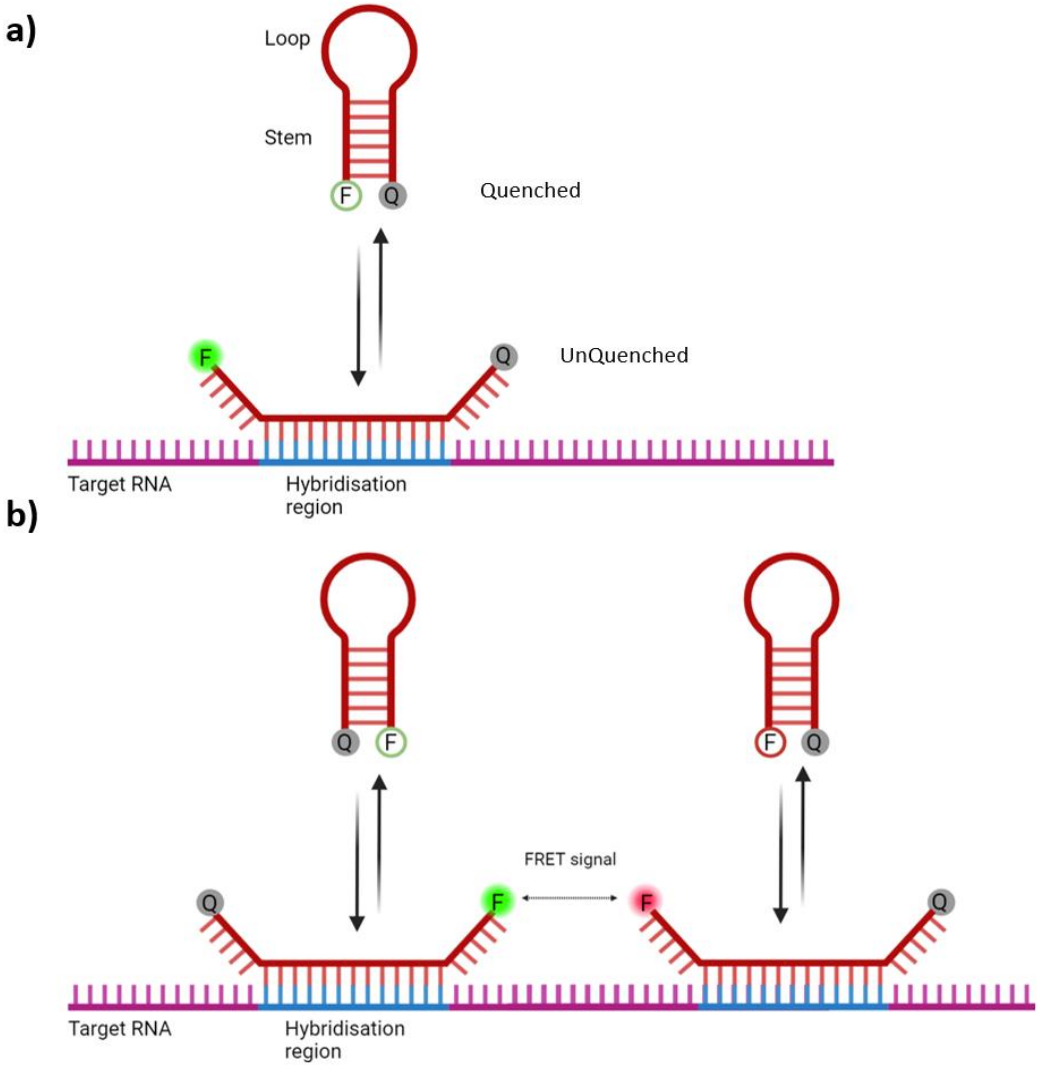

Figure 1-5: a) An MB is shown in a hairpin structure due to the annealing stem. The fluorophore is quenched due to close proximity with the quencher due to the hairpin secondary structure. The MB binds to a complementary region in a target RNA, recovering fluorescence. The specifics of whether the stem regions are also complementary are variable, but will affect specificity and hybridisation dynamics. **b)** A FRET system is utilised with two MBs that bind at adjacent sites on the mRNA. A characteristic FRET efficiency will be measurable when they are simultaneously bound

Another notable fluorescent oligonucleotide method is the use of Exciton-Controlled Hybridization-Sensitive Fluorescent Oligonucleotide (ECHO) probes. These are fluorogenic probes that utilise excitonic coupling (a quenching mechanism) between two thiazole orange dyes that is broken upon hybridisation, leading to high fluorescence (Oomoto *et al.*, 2015). This gives an idea of the versatility of the photophysics of fluorogenic probes.

1.2.3. Summary

I have briefly summarised the most common methods for RNA in live cells. This is, however, a limited review and a fast-moving field, with new developments taking place each year. An example of this is a recent paper that showed uptake of fluorescent nucleosides into transcribed RNA (Pfeiffer *et al.*, 2024). The fact that so many varied techniques are being developed indicates the biological value in RNA imaging. There is a clear requirement for a simple, unobtrusive method for RNA imaging. Although the mechanism for RNA tracking are so varied, they all require a focus on specificity and reducing background signal. The imager needs to be designed carefully to ensure that the imager binds to the correct target, and limits the signal from the non-specific interactions. These techniques have been applied primarily in eukaryotic cells and rarely on the single molecule level.

1.3. Current system

A protocol for *in vivo* FISH already exists, designed and optimised in the Kapanidis group by Dr Rebecca Andrews, Miss Emma Lalande and Dr Hafez El Sayyed. The method is most similar to the MB approach and uses short, single stranded DNA (ssDNA) fluorogenic probes. Like MBs, these are designed to quench in solution, but fluoresce upon binding to a complementary target through base pairing hybridisation. Specifically, the probes are 8 base long ssDNA, sequence TCCACCGT, labelled at both ends with ATTO647N fluorophores. This probe was used successfully in an *in vitro* hybridisation project by Dr Rebecca Andrews.

To test this system *in vivo*, a synthetic target cassette was designed, consisting of six complementary eight base targets, interspaced by random sequences. These random sequences differ slightly between strain RNAs, ranging from 6 to 15 nt. This is introduced as a synthetic gene on a plasmid, as well as two genomic inserts. When the probes are delivered, the transcribed mRNA is detected through the hybridisation of the probes to their complementary targets. The purpose of engineering six target sites is to increase the SNR through the binding of multiple probes to the same mRNA. Crucially, this makes the mRNA detectable despite the likely non-specific interactions the probe may have with a range of intracellular components. A schematic that illustrates this design is shown in **Figure 1-6**.

This method allows imaging of mRNA in *E. coli.* The proposed probe is small and bright, meaning the imaging process has a low risk of affecting the native behaviour of the mRNA. There are numerous challenges that need to be overcome to implement this system effectively. Simply put, the method is effective if the mRNA can be accurately tracked over the non-specific interactions. This will be the main focus of this thesis.

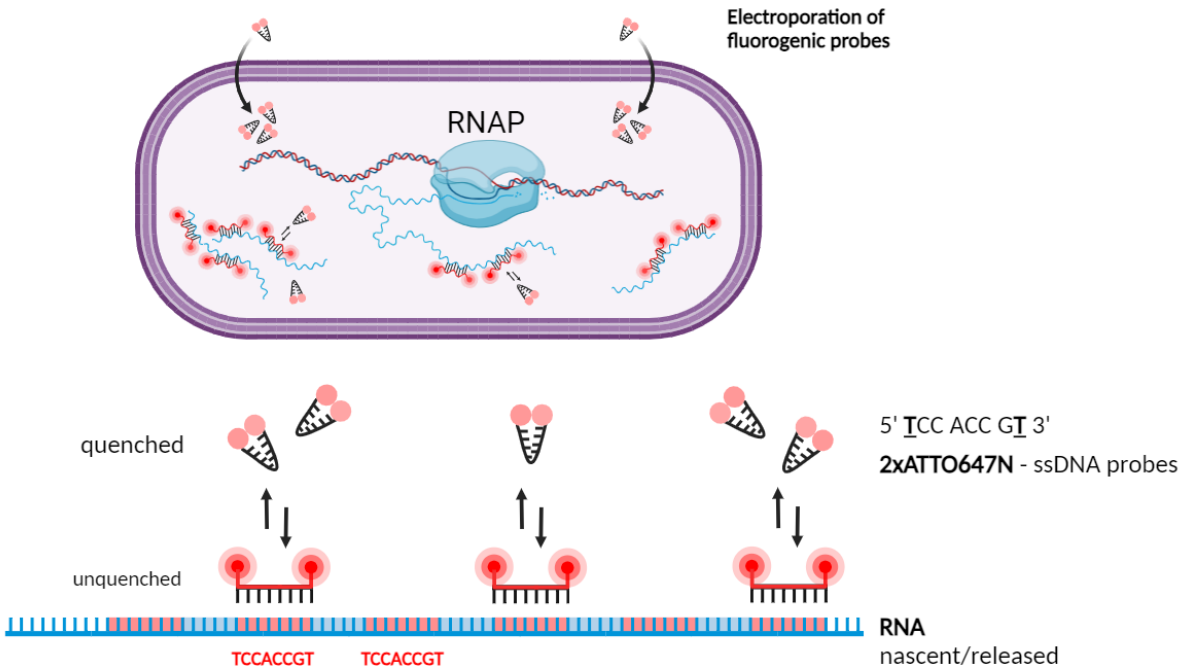

Figure 1-6: A schematic showing the proposed mechanism to study RNA. When the RNA containing the target cassette is transcribed, binding sites for a quenched 2xATTO647N probe are available for transient binding. Hybridisation of the probe leads to fluorescence of both fluorophores, as the contact quenching is relinquished. This will allow localisation and therefore tracking of the RNA with low background. The gene may be transcribed from a plasmid or a genomic insert. Schematic adapted from work by Miss Emma Lalande.

1.3.1. Probe delivery

Electroporation is an effective method for internalisation of fluorescent proteins and DNAs, based on previous work in the lab (Crawford *et al.*, 2013; Plochowietz, Crawford and Kapanidis, 2014). *E. coli* are gram-negative bacteria so are small (2-3 μm long) with a cell wall. This makes probe uptake techniques such as microinjection unrealistic. However, chemical internalisation techniques, such as using buffers containing Dimethyl Sulfoxide (DMSO) and TSS were also tested. The probe internalisation and the effect of electroporation will be discussed further in **section 3.7.**

1.3.2. Strains and Probes

Throughout the thesis, a range of different strains of *E. coli* are referred to. These can be separated into three categories. Firstly, wild type MG1655 *E. coli.* These will be referred to as the WT strain for the duration of the thesis. This is used throughout as a control, as these cells do not have an engineered target. Secondly, there are a variety of strains with introduced plasmids. These contain synthetic genes, with a cassette of interspaced binding sites under either an arabinose or IPTG-induced promoter. Another category is genomic inserts. A genomic insert with six binding sites for the 8mer was inserted into two locations in the *E. coli* MG1655 genome. This is done as part of a non-native arabinose-induced T7 polymerase. It is also introduced into the 3' UTR of the rpoB operon. On top of this, two colocalisation strains were engineered. This includes fluorescent protein binding sites on a target-carrying plasmid, and near the rpoB gene on the chromosome. These strains, referred to as pysynth and rpoB-ParS in this thesis, were constructed to identify colocalisation events to test the specificity of the method. The details of the colocalisation technique can be found in **section 3.2.** The complement of strains used is described in **Table 1-2**.

A variety of different probes were tested of varying lengths and fluorophore/quencher modifications. These are shown in **Table 1-1**, and their properties are discussed individually throughout the thesis.

Probe Name	Sequence (5' to 3')	5' label	3' label
8mer 2xATTO647N	**T**CCACCG**T**	ATTO647N	ATTO647N
8mer 2xATTO655	**T**CCACCG**T**	ATTO655	ATTO655
8mer ATTO643-BHQ1	**T**CCACCG**T**	ATTO643	BHQ1
13mer 2xATTO647N	**T**CCACCGTCGAT**A**	ATTO647N	ATTO647N
18mer ATTO643-BHQ2	**A**TCATGATCCCTAGATC**T**	ATTO643	BHQ2

Table 1-1: Details of the probes used in this thesis. Bold letters represent the positions where a fluorophore/quencher modification was made.

Strain name	Genomic or Plasmid?	Induction	Complementary probe	Binding site per RNA
WT MG1655	N/A	N/A	N/A	N/A
pB	Plasmid	Arabinose	8mer TCCACCGT	6
Pysynth-ParS	Plasmid	Arabinose	8mer TCCACCGT	6
pB24	Plasmid	IPTG	8mer TCCACCGT	24
gt7	Genomic	Arabinose	8mer TCCACCGT	6
rpoB	Genomic	N/A	8mer TCCACCGT	6
rpoB-ParS	Genomic	N/A	8mer TCCACCGT	6
pB24x13	Plasmid	IPTG	13mer TCCACCGTCGATA	24
pB TG repeat	Plasmid	IPTG	TGTGTGTGTGTGTGTG	24
pB24x18	Plasmid	IPTG	ATCATGATCCCTAGATCT	24

Table 1-2 : Details of the strains and targets used in this thesis.

2. Analysis of existing method

This section will briefly describe the experimental protocol of the existing technique as well as analysis based on in-house software. The strains that were used in this section were the plasmid-based six-target 8mer cassette, pB, and the two genomic targets, rpoB and gt7.

2.1. Protocol

This protocol is based around electroporation of the ssDNA probes into *E. coli.* The *E. coli* are then recovered, washed and imaged.

2.1.1. Preparation of electrocompetent cells.

Firstly, the *E. coli* need to be made electrocompetent. This follows the protocol developed previously in the lab, e.g. (Plochowietz, Crawford and Kapanidis, 2014). *E. coli* were grown in an overnight LB culture to saturation. 1 ml of this culture was taken and grown in a 100ml LB culture up to OD_{600} = 0.2, at 37 °C. For many of the strains, particularly those with engineered plasmids, an antibiotic selection marker was added during culturing. After reaching OD_{600} = 0.2, the cells were centrifuged at 4700 rpm for 10 min. The supernatant was removed and the pellet resuspended in 50 mL MillliQ water. This washing step was repeated at 25 mL and 15 mL also. After spinning down the 15 mL sample, the cells were resuspended in 1 mL milliQ and centrifuged at 13,000 rpm. Finally, the pellet was resuspended in 450 µL of 15% glycerol and aliquoted, 20 µL at a time, into PCR tubes. Throughout the washing process the samples/centrifuges were kept on ice. The aliquots were then stored at -80 °C.

The purpose of the washing process is to remove the salts from the LB culture. This improves the efficiency of electroporation process and increases probe internalisation. In previous work commercial electrocompetent cells were purchased but this is not possible for the purposes of this project due to the requirement for engineered cells. The process of making cells electrocompetent introduces experimental variation. Different degrees of washing will change the probe internalisation between samples.

2.1.2. Preparation for Imaging

The 20 µL cell stocks of interest were taken from the -80 °C freezer and placed on ice to thaw for 5-10 min. The cells were thoroughly mixed. 5 µL of 2xATTO647N ssDNA probe stock, diluted in MillQ water, was added to a final volume of 25 µL. Final probe concentrations were in the range of 10-100 µM. Immediately after addition of the probes, the solution was thoroughly mixed, placed in a 0.1 cm GenePulser cuvette, and electroporated once at 1.4 kV. Immediately after electroporation, the cells were suspended in 500 µL of S.O.C. media for 20 min at 37 °C for recovery.

Once the cells had been recovered, they were centrifuged at 6,000 rpm at 4 °C for 2 minutes. The pellet was resuspended in 500 µL of 1x PBS. The cells were washed further a minimum of 3 more times in 1x PBS. After washing, the cells were resuspended in 20 µL PBS, depending on the cell density after washing. A maximum of 5 µL of the cell solution was added to a 1x rich M9 media (M9 salts supplemented with 1x MEM vitamins, 1x MEM amino acids, 0.2 mM $MgSO_4$, 0.01 mM and 2% glucose) and 1x Agarose pad for imaging. If a sample required arabinose induction, arabinose was added to the pad for a final concentration of 1-2%. The pad was then left at room temperature for a minimum of 30 min before imaging, if induction was required.

The samples were imaged using HILO microscopy on the Nanoimager. The wavelength of excitation was 640 nm, ideal for the ATTO fluorophore family. A range of exposure times were tested, from 20 ms to 500 ms. It was found that 50 ms gives a sensible range of behaviours with high SNR and long lifetimes. Choice of exposure time is subtly important for accurate tracking. There will be motion blur of molecules that are diffusing rapidly relative to the exposure time. It was thought that the biologically relevant RNA target would be tracked at 50 ms, and artefact such as free dyes would not. A laser power of 8% was used throughout as this was the minimum that allowed localisation.

2.2. Analysis and Tracking

Firstly, the videos were analysed in ImageJ, to get an idea of the amount of signal in each field of view and if there are any interesting events obvious to the eye. Then, more detailed analysis was carried out using in-house MatLab and Python software.

A simple test used throughout this thesis is to evaluate the intensity of the cells throughout the video. Cells that contain more events should have a higher intensity above background on average. The intensity curve above background is integrated, to give the excess signal due to the probes (and autofluorescence) over the course of the video.

Existing single molecule tracking techniques were used to localise and track the foci seen in the cells. This infrastructure was developed over the past decade in the Kapanidis lab, but was focused on Photoactivated Localisation Microscopy (PALM). Adaptations were taken to adapt for organic fluorophore tracking, such as reducing tracks errors by removing frames where several interfering tracks overlapped.

A 2D Gaussian fitting was applied to each spot above a certain threshold. The exact parameters of the tracking required optimisation. Due to the nature of the method the number of foci observed in the cells is not easily controllable, especially with variable probe uptake between cells. Therefore, to reduce errors from crowding, frames where there were more than two spots in a cell were removed. A memory parameter of three frames was applied. The majority of videos were taken at an exposure time of 50 ms. This is too high for accurate tracking of fast-moving species, perhaps representing freely diffusing dyes, but is sufficient to track biologically relevant behaviours.

The typical parameter used as an indication of mobility is the apparent diffusion coefficient, *D**. This is a common mechanism in the literature but its meaning is far from simple, given the complex nature of diffusion in the cytoplasm. A comprehensive review of the mathematical and biological background can be found here (Bohrer and Xiao, 2020).

With the specific imager and imaging characteristics involved here, the majority of events are demonstrating sub-diffusive behaviour. This is expected as the molecules, whatever they may be, are moving in the viscous cytoplasm with a multitude of interactions possible with cellular components. Furthermore, many of molecules are seen to exhibit constrained behaviour. This could represent anything from a freely diffusing molecule constrained within the cell as a whole to plasmid motion constrained within a cell pole. In general, there are a huge range of feasible behaviours to be observed. Although it may be possible to characterise these using this method it will not be the focus of this thesis. Instead,

*D** will simply be used as a general indication of mobility which is sufficient to detect the wide range of mobilities observed here.

The position in the cell where the different mobility populations are found may also be biologically interesting. For example, mRNA that is being transcribed will be bound to the plasmid. It is thought that plasmids are typically constrained to the polar regions of the cell and so it would be expected to find a population of tracks with low mobility in the polar regions (Reyes-Lamothe *et al.*, 2014). To do this, the outline of a cell is projected onto the shape of a 'standard' cell of rectangular shape. Each position within the cell of a certain track is then projected into this rectangle to build up a 2D distribution. The distribution can be separated into sub-distributions of tracks that have mobility within certain thresholds.

Fundamental to the analysis is the use of cell masks. A custom CellPose model was trained to segment the cell outlines from the brightfield images (Stringer *et al.*, 2021). Any mistakes were fixed through manual annotation. The segmentation process was packaged into a single software by Dr Piers Turner (*piedrro - Overview*).

2.3. Experimental Results

Here I will present the data that were collected using the 2xATTO647N probe. This is mainly focused on testing the specificity of the method by comparing the cellular intensity traces, and characterising typical tracks.

2.3.1. Standard Field of View

The results take the form of a brightfield image followed by a video, typically of over 1000 frames at 50 ms. There is a large increase in intensity when the laser is turned on due to fluorescence of the dye and some autofluorescence in the cells. The autofluorescence is typically bleached within 10 frames at 50 ms. Throughout the initial stages of the video, the number of foci reduce rapidly due to photobleaching, but there are events that begin late in the video. This is to be expected if there are a pool of quenched probes in the cell that have some probability of unquenching in the course of the video.

To some extent the method is working, with obvious internalisation and visualisation of bright foci. This is shown in **Figure 2-1** , where a typical cell and its mean fluorescent intensity curve is shown throughout the 50 s video. The cell brightness distribution is heterogeneous. There are many cells that are so bright that individual foci are not clear, and many cells that show no obvious foci. This is due to the unpredictable nature of the electroporation, where cells with differing membrane permeabilities are exposed to the probe solution. Furthermore, there will be varying expression of the target gene in each cell due to stochastic gene expression.

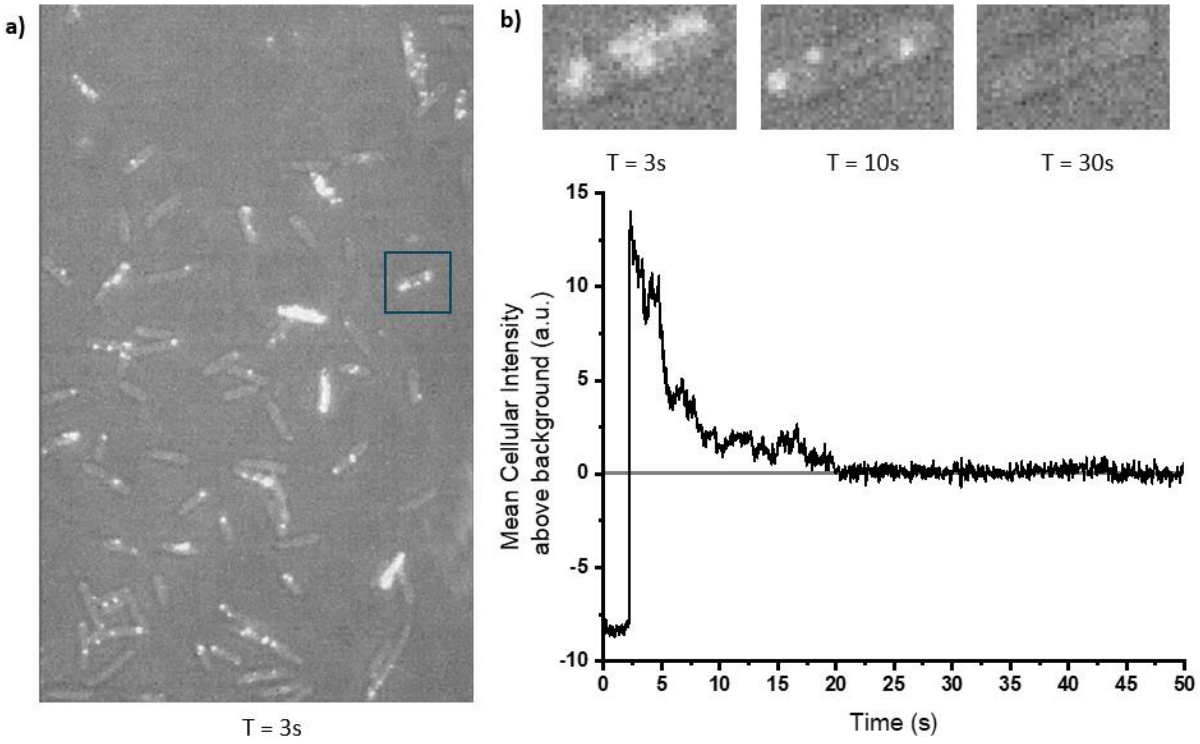

Figure 2-1: A representative field of view for WT cells that were electroporated in the presence of 25 nM 2xATTO647N probes, immediately after the laser is turned on at T = 3 s. The brightfield image is overlaid on the red channel video to demonstrate the location of the cells. It is clearly visible that there is heterogeneity in signal: many cells show no signal and a few are completely bright. However, many cells show clear foci that can be localised and tracked. **b)** A zoomed in view of a cell shown in the box throughout the video. Initially we see a large amount of signal due to a combination of fluorophores and autofluorescence. During this phase, localisation and tracking will be challenging due to the crowding. However, this quickly bleaches to reveal three clear foci that bleach themselves shortly afterwards. Signal is present until around the 20 s mark.

2.3.2. MilliQ and WT Control

A key control is to follow the same protocol, but instead electroporate only MilliQ water, rather than no probes. All the signal seen in this MilliQ control will be the result of cellular autofluorescence. It is clear visually that the autofluorescence is negligible compared to the bright foci seen in the samples electroporated with the 2xATTO647N probes, as shown in **Figure 2-2**. This means that localisation of the foci will be unaffected by the autofluorescence. However, when comparing cell intensities, the autofluorescence may have a significant contribution so must be kept in mind. Characterising cells that have no probe is important to understand how many cells are taking up the probe in the test samples.

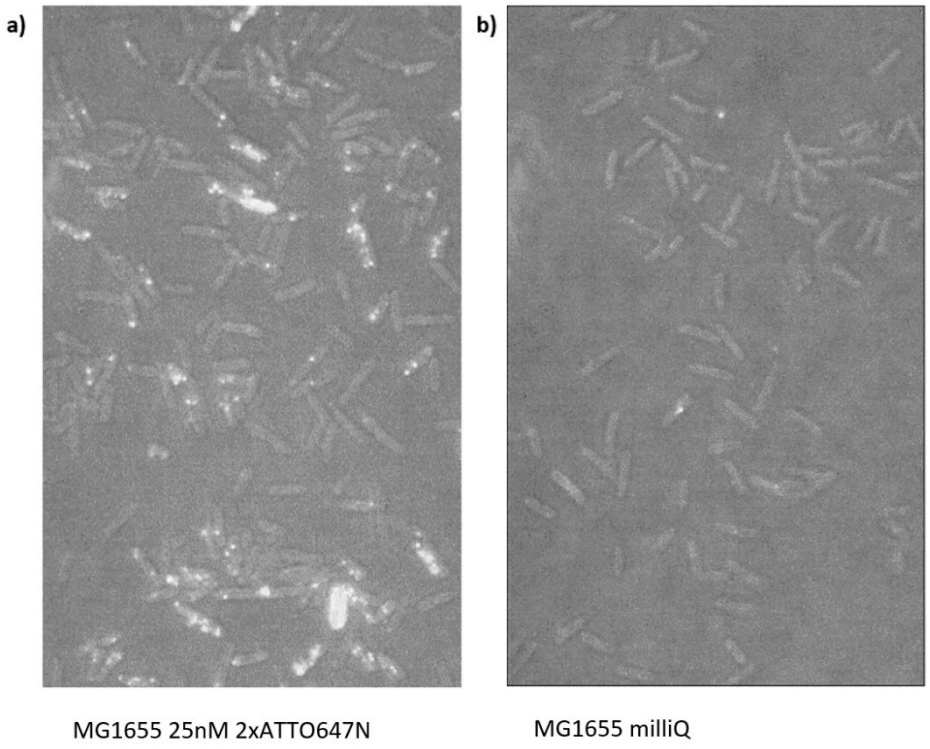

Figure 2-2: **a**) WT MG1655 cells electroporated with 25 nM 2xATTO647N probes. **b)** WT MG1655 cells electroporated with 5 µl MilliQ water instead. There are a small number of foci visible during the movie. They are few in number and short lived, meaning that the cellular autofluorescence will play a negligible role in localisation of the 2xATTO647N foci. However, there is whole-cell fluorescence which may play a role if comparing the intensities of cells.

The primary control for *in vivo* FISH is to observe the brightness of the WT MG1655 cells. The signal seen here can be assumed to come from four sources: probe bound to target RNA that happens to exist in the *E. coli* chromosome, binding to non-specific cellular components, ineffective quenching, and damaged/singly labelled probes. In the cells containing the pBEC plasmid there should be an additional source of signal: the probe bound to the specific target mRNA. In order to have an accurate, reliable mRNA tracking method, the signal from the specific mRNA must be identifiable from the other interactions.

It is inevitable that there will be a significant number of these non-specific interactions by the nature of using a short ssDNA probe in a complex cellular environment. One way of overcoming this issue is population analysis. Even if it is ambiguous whether an individual track is specific, it may be possible to make conclusions on the population level about the frequency or behaviour of the tracks. Furthermore, the RNA of interest has six binding sites. This may allow for the binding of multiple probes to the same site which would allow the RNA to be separated from non-specific interactions which will consist of only single probes. In fact, due to the probe being an 8mer, hybridisation will be weak with low dwell times allowing on-off transient binding of many probes to the same molecule.

In order to obtain accurate localisation of the target mRNA over long periods of time, there must be a large pool of probes in the cell to access the target. Upon excitation of the sample, the unquenched fluorophores will fluoresce. The degree of fluorescence can be used as an estimate for specificity. If the number of probes bound to the target is similar or greater than the number of fluorophores non-specifically fluorescing then there will be a discernible difference between a sample with the target (e.g. pB) and the control WT MG1655 sample. This assumes the cell autofluorescence is negligible.

Although this is an approximation, it should give an idea of the degree of specificity of the probe. A more detailed method was also implemented, involving the integration of the signal from each cell over the whole video. This will consider the lifetime of the signal, as well as spots that emerge later in the video. This is important if there exists a pool of quenched probes that stochastically bind and fluoresce to a target. In terms of background subtraction, the mean of the final 400 frames of the video that did not contain any localisable spots were taken. The first 10 frames of the video were excluded to limit the impact of the cellular autofluorescence on the whole cell intensities. This method was used predominantly to determine the difference in intensity of cells between samples.

As seen in **Figure 2-3**, there is no obvious visual difference between the pB and WT MG1655 samples. This is confirmed by a general intensity analysis on a cellular basis. This suggests that the signal from the probe binding to the specific RNA is minimal compared to the host of non-specific interactions. There are a number of options that could be causing the large proportion of non-specific signal, which will be discussed in future sections. It should be noted that there were variations in intensity across repeats. The source of the variation is likely due to the efficiency of the electroporation leading to different levels of probe internalisation. Despite this, it was consistently seen that there were similar intensity levels between the pB and WT strains and it was ensured that the cells were made electrocompetent identically, side by side.

2.3.3. The rpoB strain

Similar experiments were conducted using the rpoB tagged strains. These were prepared almost identically to the other strains. Consistently there was some evidence that the rpoB strain showed of a marginally higher population of high brightness cells than the WT and pB strains, as seen. This suggests that the targets on the genome are being bound to at least some degree. Perhaps, the mRNA molecules, being up to 9 kb, have a long lifetime which allows more time for probes to hybridise. Furthermore, as the cells are harvested at exponential phase (OD_{600} = 0.2) in rich media, expression of the rpoB gene is expected to be high. **Figure 2-4** shows the corresponding cell brightness distributions, with exemplar field of views FOVs shown.

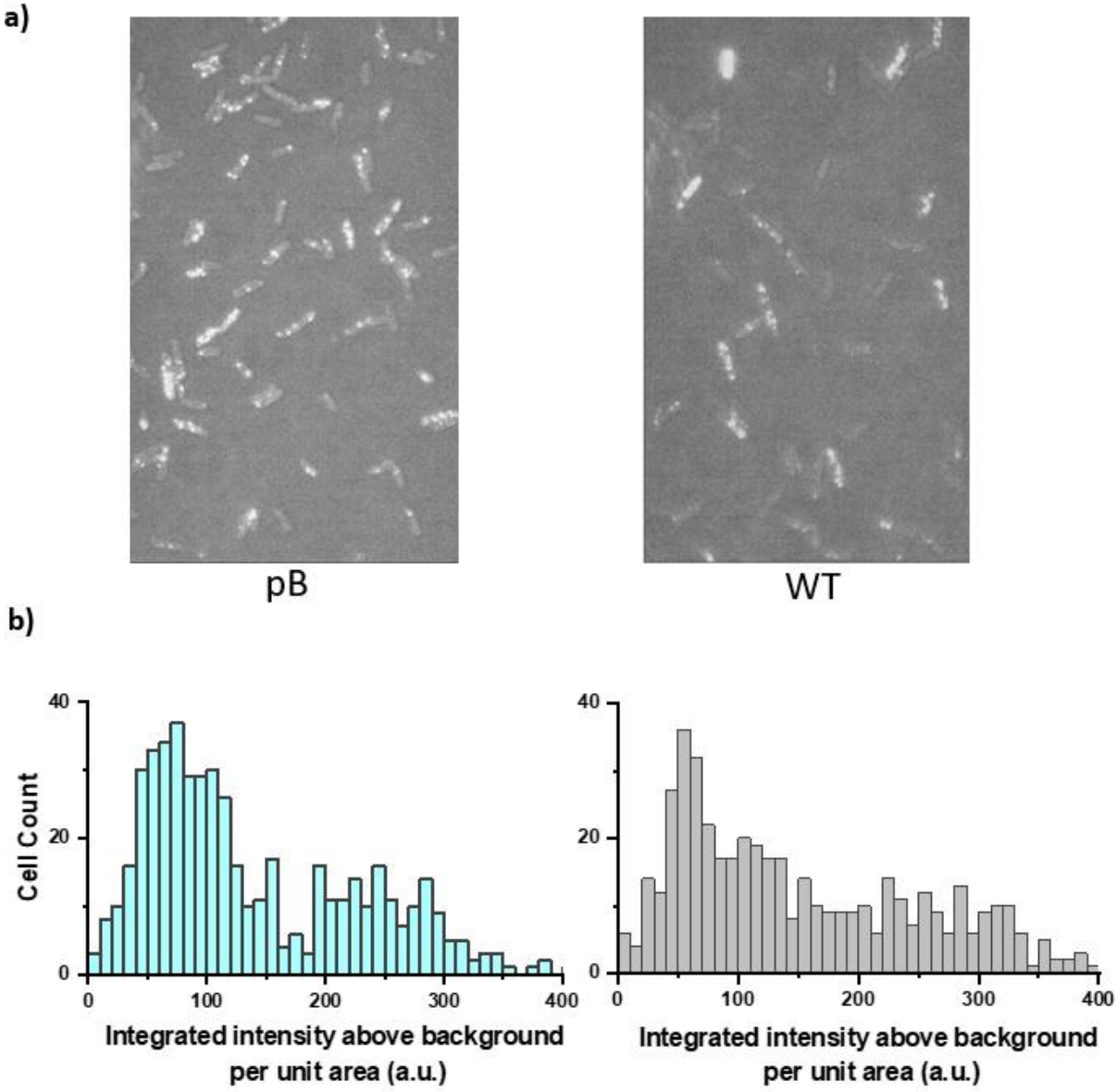

Figure 2-3: a) Two representative field of views (FOVs) from pB and WT strains, each electroporated with 50 nM 2xATTO647N probes immediately after the laser was turned on. As visible from the images, both samples have cells that are very bright and cells that are clearly devoid of signal. It is not clear visually which of the cell populations are brighter. **b)** Here, the cell intensities above background for five FOVs are calculated and plotted on a population level. A MilliQ sample population is shown for reference (FOV not shown in **a)**). There are two main populations: those that show no signal, in line with the milliQ distribution, and a long distribution tail of cells with varying probe uptake. From both the FOV visualisation and the intensity distributions it is clear than the samples have similar amounts of signal. N = 400 cells across multiple FOVs for pB and WT, and 150 for the milliQ sample.

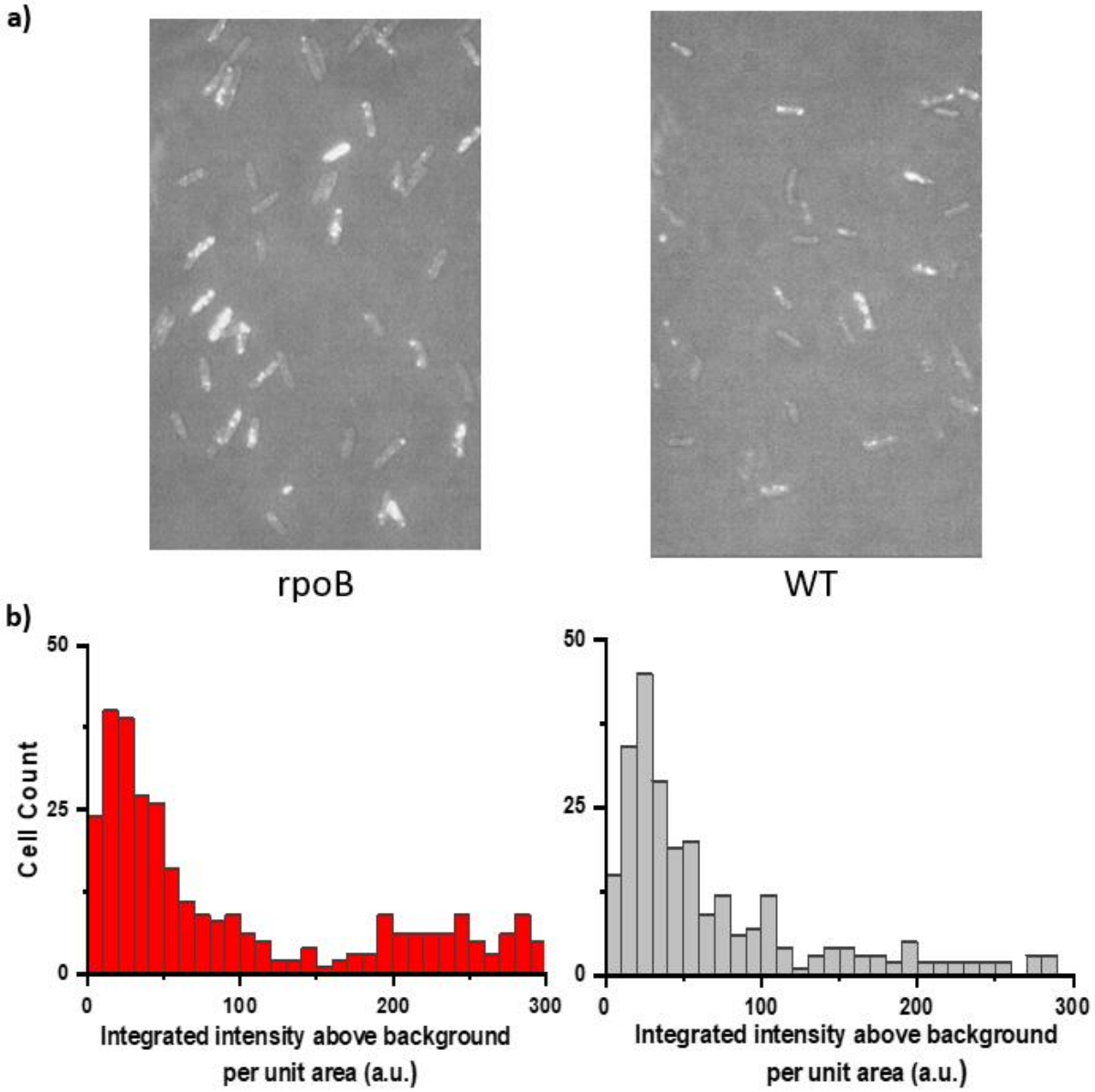

Figure 2-4: a) Typical FOVs for the rpoB and WT samples, electroporated with 25 nM 2xATTO647N probe. Again, there is no clear difference, but it seems that there are more cells that are completely full of signal in the rpoB signal, however this is just from one FOV, so could be due to stochastic variation. **b)** The cell intensity above background distributions, calculated as in **Figure 2-3**. Again, there is a clear population with little signal, and a long tail of brighter cells. As observed in the images, there is slightly higher population of cells that are very bright in the rpoB sample. These data are collated across 5 FOVs.

2.3.4. Rifampicin

A further test of the specificity is to perturb the system using rifampicin. Rifampicin is an antibiotic which prevents transcription through binding to the β subunit of RNA polymerase (RNAP) (Adams *et al.*, 2021). Therefore, cells exposed to rifampicin will have a lower concentration of mRNA. This is because the pool of mRNA in the cell will be degraded but is not replenished. Rifampicin application has been used to measure the half-life of an mRNA, typically through bulk RNA measurements, and gives a typical RNA half-life of around 5 minutes (Bernstein *et al.*, 2002). Firstly, consider the plasmid strain, pB. There are no measurements of the half-life of the synthetic target cassette of 120 nt, but there must be some degree of degradation. In theory, cells exposed to rifampicin should show less signal due to the expected lower concentration of target mRNA. The chromosomal strains should show the same pattern. However, they are many kilobases in length and the rpoB tagged strain contains ribosomal binding sites. The impact of the binding of ribosomes on the degradation rate is not clear. Furthermore, transcripts can be protected from degradation by addition of various functional groups to the 5' end (Mohanty and Kushner, 2022). In short, the stability of the labelled transcript is hard to predict.

Cells containing the pB plasmid were added to an agarose pad which contained a final concentration of 100 ug/mL rifampicin. Various FOVs were taken over the course of an hour and the brightness of the cells quantified. The stability of the probes is also being tested here. If the probes are vulnerable to digestion over the hour time period, it would be expected that the amount of signal would rise. On the other hand, the opposite may be seen if the fluorophores themselves are not robust against degradation processes.

It was found that there was minimal change in the intensity, if any, shown in **Figure 2-5**. Assuming that the transcript is degraded, this is evidence that a significant proportion of the probes are interacting with components inside the cell that are not RNA.

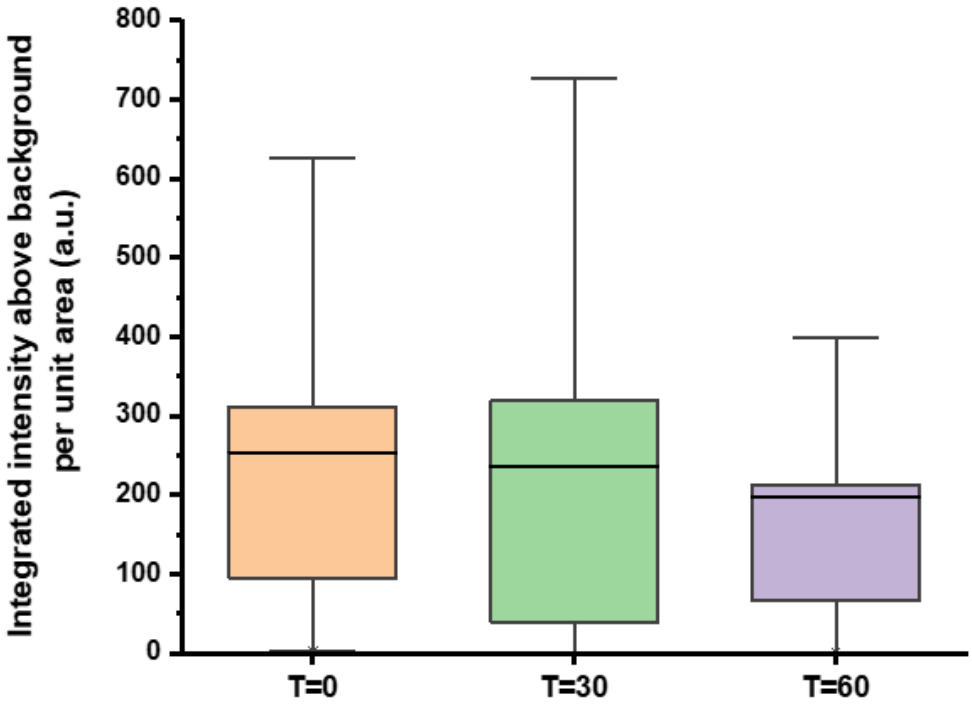

Figure 2-5: 100 µg/mL of rifampicin was applied to the agarose pad containing a pB sample that has been electroporated in the presence of 100 nM of 2xATTO647N probes. Different FOVs were taken roughly every 4 min and the cell integrated intensity distribution calculated. This resulted in a small change in the intensity across an hour period. This suggests that both the probes are stable for long periods of time and that reduction in RNA levels due to rifampicin treatment was minimal.

2.4. Localisation and Tracking

Although the initial frames provide a useful insight into the degree of probe internalisation and quenching, the tracking of the foci reveals many rich behaviours. There are tracks that show a wide range of mobilities and positions with the cells, including tracks which demonstrate transitions between mobility states. In addition, the population of *D** values will show which behaviours are most common.

If specific binding to the mRNA is being achieved, then there should be biologically relevant events covering a range of mobilities. **Figure 2-6** shows three example tracks across the spectrum of *D** values, taken from a rpoB sample. This ranges from spots with very little motion at all, with $D^* \approx 3 \times 10^{-4}$ μm^2 s^{-1}. There are various interpretations of these spots. It is possible that it represents a probe binding to a chromosomally tethered RNA. Values have been reported for chromosomal diffusion of $D^* \approx 10^{-4}$ μm^2 s^{-1} (Javer *et al.*, 2013). However, another explanation is that the probe is bound a stationary object, perhaps a component of the cytoskeleton or membrane. The localisation error is such that a stationary object will be reported as having some non-zero *D** value. It is possible that events such as these are moving within the localisation error, which is around 40 nm for our setup. The tracks with higher mobilities could represent the probe bound to various released transcripts. Larger transcripts will not be able to percolate through the nucleoid, so are typically confined to the polar regions. It should be noted that the tracks shown in **Figure 2-6** are considerably longer than average. This is a selection bias as the behaviours can be more easily illustrated. However, it is unclear what the dwell time for an 8mer fluorogenic probe would be *in vivo*.

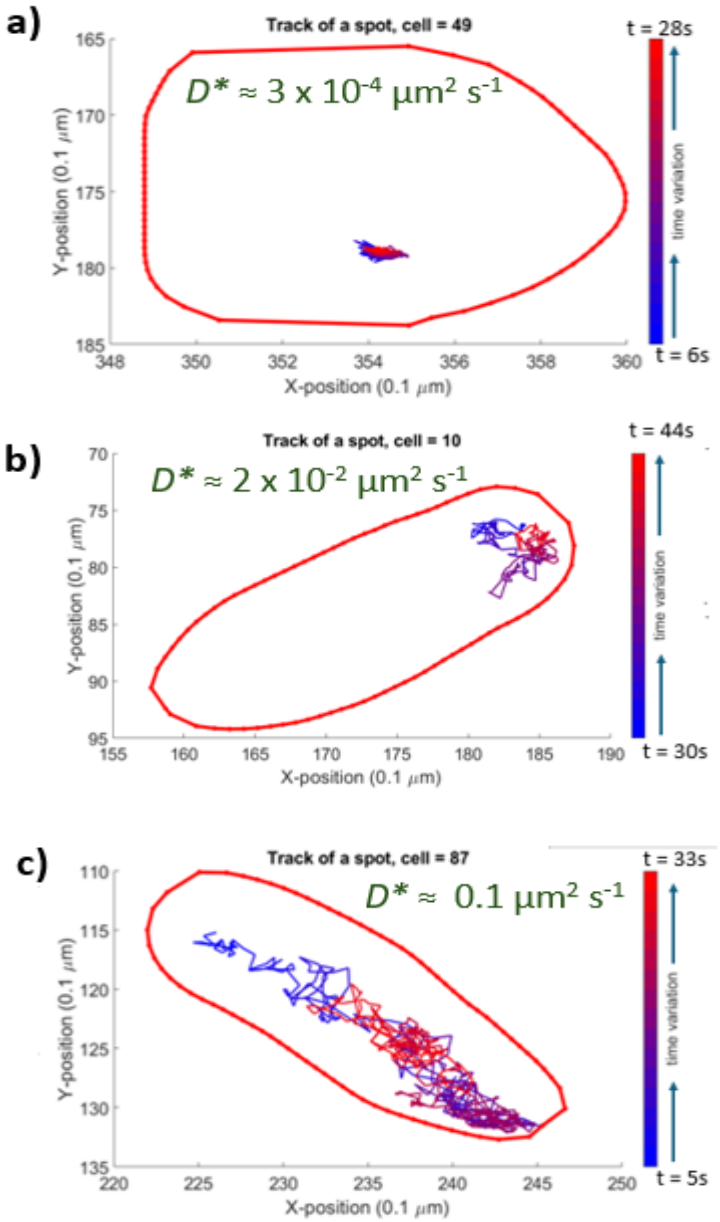

Figure 2-6: Three tracks from a rpoB sample that show a range of mobilities. It should be noted that tracks of this length are rare, occurring approximately 1-2 per FOV. The time taken through the track is represented by a colour scheme of blue to red. **a)** A stationary spot, moving within the localisation error. $D^* \approx 3 \times 10^{-4}$ µm^2 s^{-1}. This is consistent with values given in the literature for chromosome diffusion (Javer *et al.*, 2013), but could also be consistent with a probe embedded in the membrane or other immobile structure. **b)** This spot explored the polar region of the cell, with $D^* \approx 2 \times 10^{-2}$ µm^2 s^{-1}. This could be a released transcript undergoing translation in the ribosome rich polar regions. **c)** A faster moving track that explored the whole cell, $D^* \approx 0.1$ µm^2 s^{-1}. It is not clear what this could represent, but could be a non-specific interaction or a transcript that is partially degraded and can therefore percolate the nucleoid. Importantly, these behaviours are observable in WT samples which prevents clear biological inference. The FOVs where these tracks were taken from were collected by Miss Emma Lalande.

The mRNA molecules will contain six binding sites for the 2xATTO647N probe. If the local concentration of the probe is high enough, transient binding of multiple probes to the same spot can be expected. To estimate the number of fluorophores per foci, the unitary intensity of the fluorophore *in vivo* must be calculated. To achieve this, the whole cell intensity trace was calculated (above background) and imported into in-house HMM fitting software, developed by Dr Piers Turner. The hidden states were fitted and the differences between adjacent states were found, assuming they represent the steps of single or double fluorophore intensities. The single intensity steps are due to the photobleaching of a single fluorophore. It is possible that this may happen in two ways: the probe is in a contact quenched state but one fluorophore bleaches leaving the other unquenched and fluorescent, or the probe is in an unquenched state and one fluorophore simply photobleaches. The double intensity steps are due to a probe binding and leaving the contact quenched state, or the reverse, a probe unbinding and contact quenching is resumed. An exemplar cell showing these steps in shown in **Figure 2-7**.

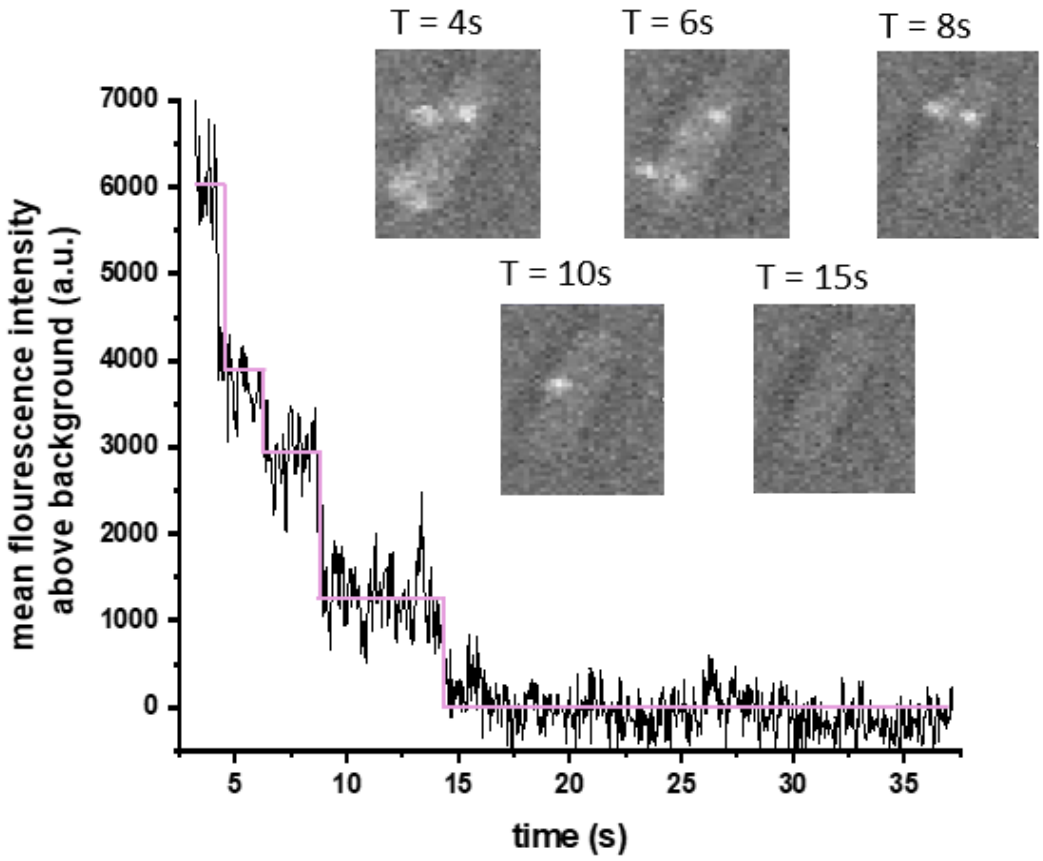

Figure 2-7: Shows a trace of a WT cell electroporated with 50nM 2xATTO647N probes. The state mean found using Hidden Markov Model fitting overlay the intensity trace in pink. Initially, the cell contains many spots and some autofluorescence, before undergoing several bleaching steps to end with zero signal. There is significant fluctuation around each state and the steps are of varying heights, which makes it difficult to calculate an accurate unitary intensity. However, the step from T = 6-8 s is approximately half of that from T = 4- 6 s and T=8-10 s, which could represent the difference between single and double fluorophore steps. The area under this curve is what is calculated when comparing the integrated intensity distributions of cells.

The calculation of the unitary intensity proved difficult. The imaging in HILO means that there is a significant intensity variation if the spot moves in the z-plane. Combining this with uneven illumination across a field of view and fluctuating laser intensity makes it very hard to pick out the unitary intensity of an individual fluorophore on a population level. However, discrete steps in the intensity trace coupled with the disappearance and reappearance of foci can give a rough estimate on a cell by cell basis.

At every frame of the video, the spot intensity is calculated by integrating under the fitted 2D Gaussian, and subtracting the background. If there are enough discrete steps, i.e. beyond blinking and photobleaching of two fluorophores, then it suggests that multiple probes have bound to the same mRNA molecule. If these molecules can be identified from the non-specific interactions consistently then progress can be made in spite of the significant noise.

One such trace is shown below in **Figure 2-8**. These traces proved rare, with <1 per FOV. This is likely due to the probability of two, non-photobleached probes binding to the same mRNA is very low. The high local concentration required to make these feasible would require a very low rate of non-specific reactions to extract accurate traces. It would be useful to compare the rate that these traces occur between strains with the target and the WT control. However, they were too rare make robust conclusions.

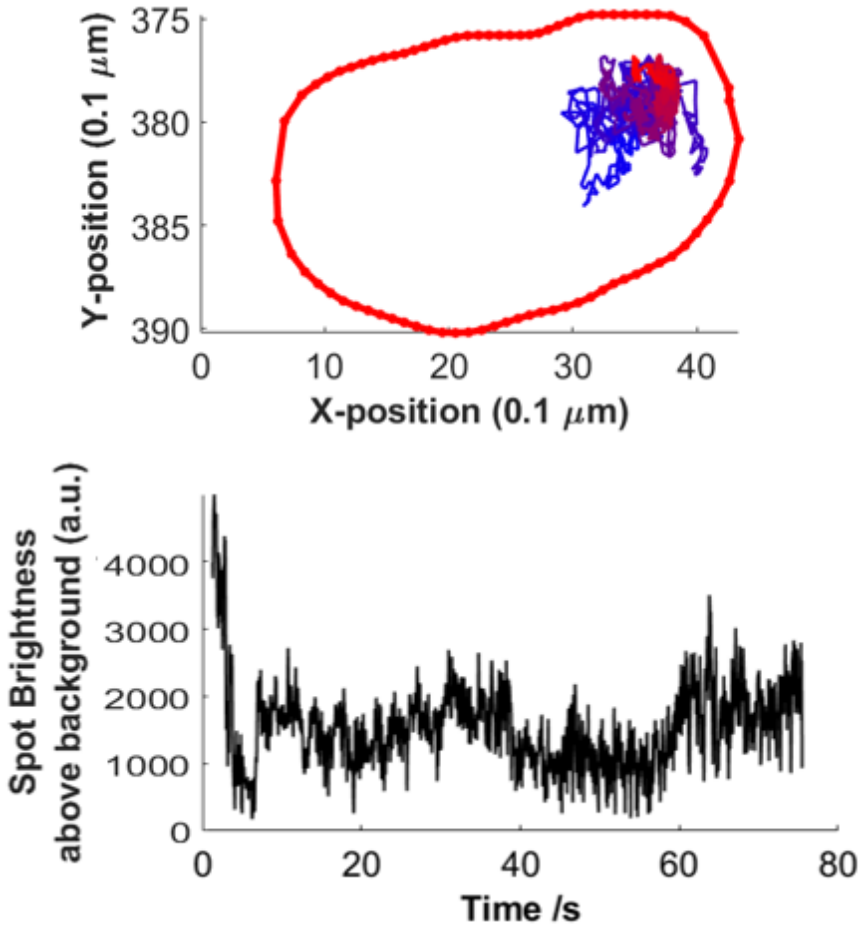

Figure 2-8: This spot was identified in a rpoB sample as it shows a wide range of intensities. Initially the spot begins very bright, which suggests that multiple probes are bound to the same object, but they quickly bleach or dissociate. It is possible here than there was interference from other spots. At around 5 s there is an intensity jump, which could be either a single or double fluorophore increase. For roughly 75 s, the spot varies in intensity whilst exploring the pole of a cell with $D^* \approx 1 \times 10^{-2}$ μm^2 s^{-1}. This may be due to transient binding of other probes to a released transcript.

2.4.1. Transitions

A particularly interesting feature of the tracks is transitions between different mobility states. This is more noticeable if the tracks are longer. One method of achieving this is to utilise dark interval imaging where the laser is turned off for much of the cycle. A typical cycle would be to excite for 50ms and then turn off for 100-200 ms. This cycle is repeated. This is a simple way of tracking longer events as photobleaching is the limiting factor. **Figure 2-9** shows an extended track taking using dark interval imaging that displayed a range of behaviours and transitions.

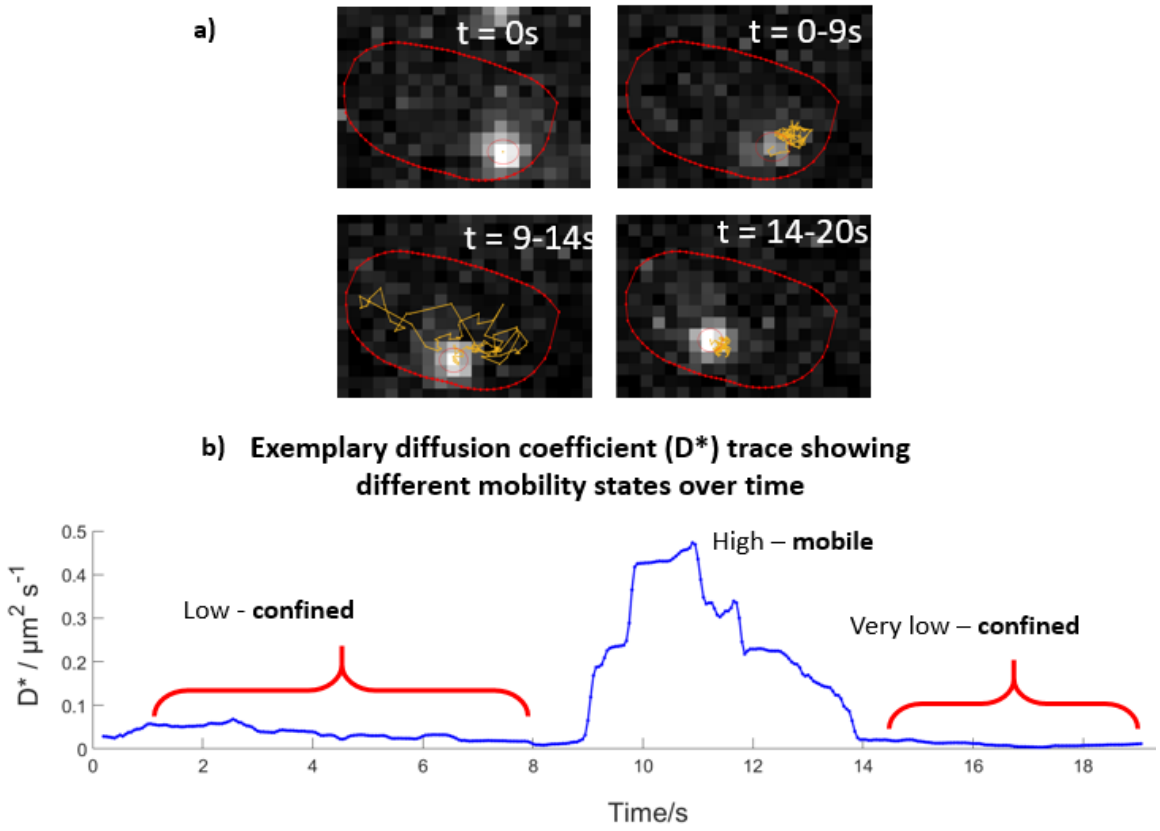

Figure 2-9: Dark interval microscopy was used here to observe tracks over longer periods. **a)** Four stills from a track taken in a pB sample. The yellow line shows the path taken by the fluorophore in each of the time intervals. **B)** A rolling *D** value is calculated throughout the track. The track is initially confined with $D^* \approx 2 \times 10^{-2}$ µm^2 s^{-1} and is exploring the polar region of a cell. Then there is a huge increase in mobility for approximately 5 s, before the spot becomes even further constrained for the remaining duration of the event. Transition such as these are uncommon, with typically fewer than one per FOV.

2.4.2. Population data

Although the videos have similar numbers of spots and similar brightness', it is possible that the spot behaviours between samples may be different. The best way to assess this is to plot a distribution of the mobilities of the tracks observed per sample. Each sample will contain a wide range of mobilities, and inevitably contain tracking errors. However, there may be a discernible difference in the populations if the tracks have a higher probability of a certain behaviour over another. Predictions can be made on the mobility populations expected given the nature of the fluorophore and the target. For example, the rpoB strain should show a larger population of slow-moving events, e.g. $D^* = 10^{-3}\,\mu m^2\,s^{-1}$, compared to the pB strain due to the slower movement of the genome vs the plasmid and the longer transcript.

The tracks were split into segments of 5 frames long, or 250 ms. This was chosen by trial and error as many tracks were confined on about this timescale. Importantly, this does weight the distributions towards longer tracks, as they will produce more segments. It was decided that this bias was not an issue, due to the longer tracks exhibiting a large range of mobilities. The mean square displacement (MSD) was plotted against time for this 5-frame segment, and then D^* calculated from the gradient. The parameters chosen here, such as exposure time, segment length and so on will lead to a wide range of different mobilities calculated. However, if these parameters are kept controlled, then differences in populations will be detectable, even if the exact D^* values are not entirely robust. Any segments that were labelled with a D^* value < 10^{-5} $\mu m^2\,s^{-1}$ were set to the value $D^* = 10^{-5}$ $\mu m^2\,s^{-1}$. This is only likely to occur with molecules that are bound within the localisation error. **Figure 2-10** and **Figure 2-11** show the diffusion histograms produced from this method. D^* is plotted on a log-scale as a wide range of diffusion coefficients are observed and the populations are clearer.

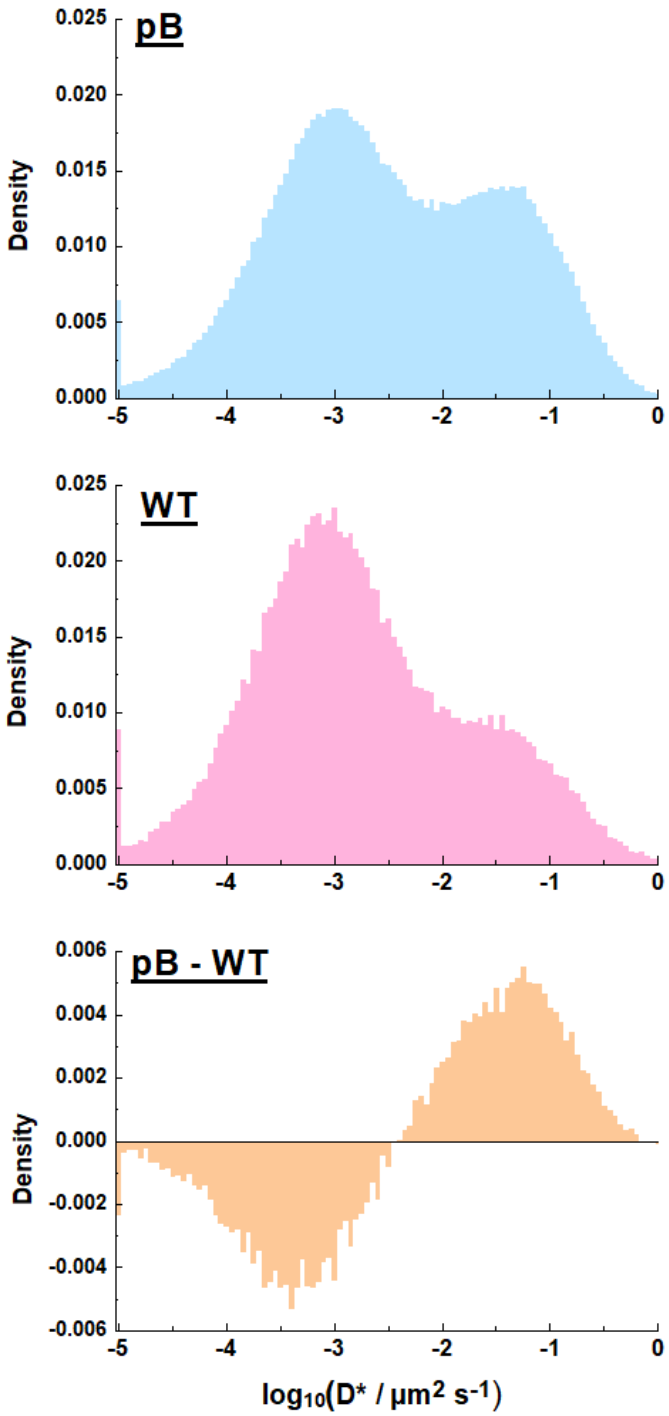

Figure 2-10: The tracks were split up into smaller tracks of five frames in length. This was chosen as it is roughly the point at which many tracks become confined.

The first panel shows the population of tracks for a pB sample. There is a peak at around $D^* \approx 1 \times 10^{-3}$ µm^2 s^{-1}. This represents immobile tracks which are confined in a region which is of the order of the localisation error. The second peak, at around $D^* \approx 1\text{x } 10^{-1}$ µm^2 s^{-1}, represents mobile tracks diffusing throughout the cell.

The second panel shows the same plot but for the control WT sample taken alongside the first panel. The first population of immobile tracks is present, but the mobile population is reduced relative to the pB sample.

Finally, the third panel shows the difference between the two density populations. It can be seen that the immobile population is more pronounced in the WT sample, at expense of the faster population at $D^* \approx 1 \times 10^{-2}$ µm^2 s^{-1}.

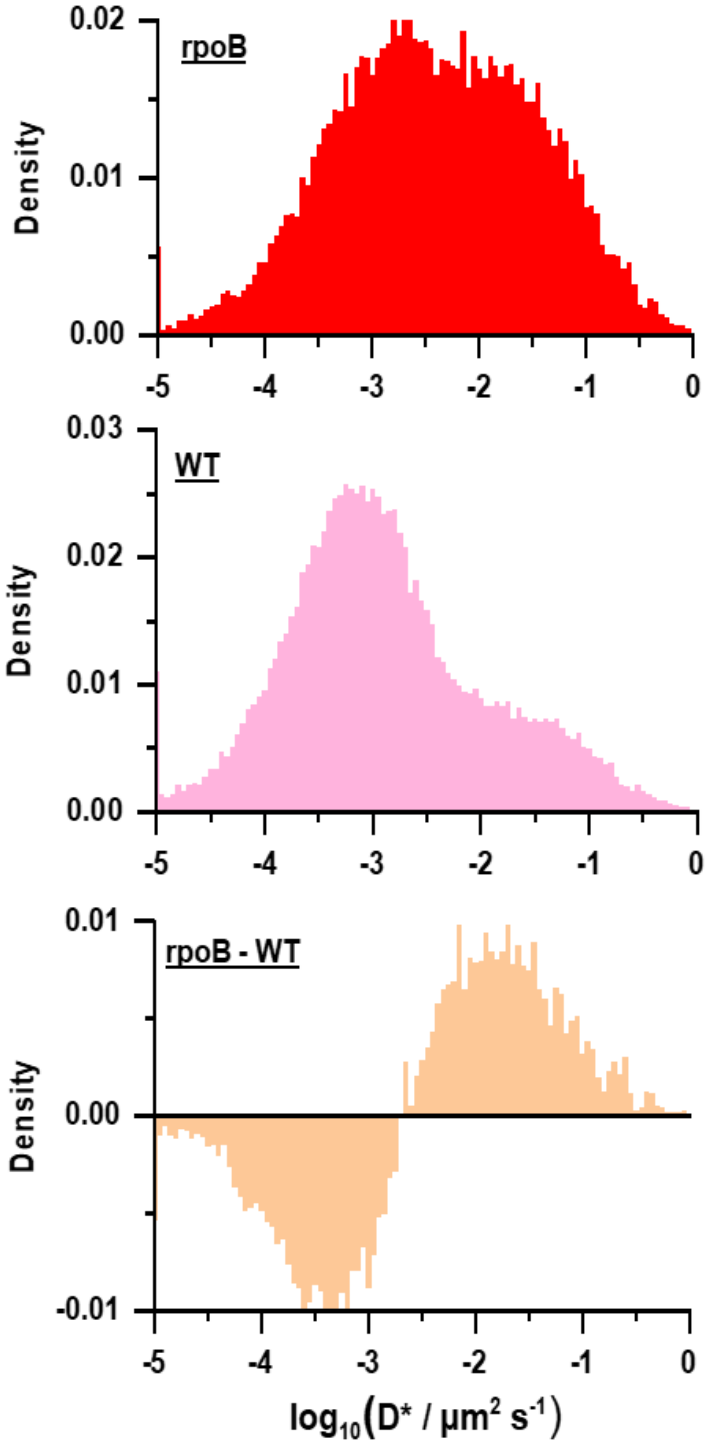

Figure 2-11: An identical plot to **Figure 2-10** but from an experiment with a rpoB strain sample rather than a pB sample.

The first panel shows the mobility distribution for rpoB cells. The populations are less clear, but it seems that there remains a significant proportion of immobile tracks. The mobile population, as seen in the pB sample, is shifted to lower mobilities.

The WT population is very similar to **Figure 2-10**, which shows consistency between biological replicates.

The difference between the populations again shows the large immobile population in the WT sample, with rpoB showing a greater proportion of molecules with $D^* > 2 \times 10^{-2}$ µm^2 s^{-1}.

Relative to the pB sample, the rpoB distribution shows a lower proportion of mobile molecules. This can be interpreted as due to the larger target transcript in the rpoB sample displaying more confined motion.

2.5. Summary

Figure 2-1 introduces a typical FOV and shows the bleach-down curve of a cell electroporated with 2xATTO647N 8mer probes. The first control of comparing to a MilliQ sample shown in **Figure 2-2** showed that the fluorescent probes will be localisable relative to cell autofluorescence. However, the WT control showed brightness distributions to both the pB and rpoB sample, **Figure 2-3** and **Figure 2-4** which suggests a significant proportion of the signal derives from fluorophores that are not bound to the desired target. Cells with rifampicin treatment, designed to prevent the transcription of the target mRNA, did not have significantly different brightness levels as seen in **Figure 2-5**.

Despite the brightness populations suggesting only limited specificity a wide range of spot behaviours are presented in **Figure 2-6**, **Figure 2-8** and **Figure 2-9** with some indication of multiple probes bound to the same target. This is supported by HMM fitting to find the unitary intensity of the fluorophore. This was challenging on a population level but is shown in **Figure 2-7** for a single cell. Forming a distribution of the *D** values across many FOVs indicated that their may be a difference in the mobility populations between the different strains. Furthermore, the WT control consisted mainly of low mobility spots, seen in **Figure 2-10** and **Figure 2-11**.

As previously reported ATTO647N is a hydrophobic dye (Kolmakov *et al.*, 2010) (Plochowietz, Crawford and Kapanidis, 2014). It is therefore likely that the non-specific interactions are driven by interactions of the fluorophore with other hydrophobic intracellular components. **Chapter 3** will discuss the attempts to improve the specificity, both by increasing the brightness of the target through adding more binding sites and reducing the number of non-specific interactions by exploring the behaviour of other fluorophores.

3. Further method development:

A major challenge with this method is that the non-specific interactions occur at a high rate. In this chapter I will discuss many attempts made to reduce these interactions and increase the signal from the specific target. This began initially with work on the plasmid strains increasing the number of binding sites from 6 to 24. A further test of specificity was to use the colocalisation strains to identify genuine events. Then, a 13mer was used for hope of higher specificity than the 8mer. After this, it was decided to test other fluorophores from the ATTO family, ATTO643 and ATTO655, which have different chemistries. ATTO643 was paired with a Black Hole Quencher, BHQ1, which gave higher quenching efficiencies.

3.1. 24 target plasmid – pB24

Analysis of the unitary intensity suggests that the vast majority of events are single or double fluorophore events. Both of these probably represent a single probe. It was thought that the six binding sites on the standard target were not sufficient. Additional binding sites would provide higher specificity, with longer and brighter events. This was tested with a plasmid was provided by Dr Hafez El Sayyed which was identical to the previous plasmid, but contained 24 evenly spaced, 8 nt targets. It was hoped that the additional binding sites would provide higher specificity, with longer and brighter events. This is referred to as the pB24 strain. This could provide the specificity needed to consistently detect the mRNA from the non-specific signal. **Figure 3-1** shows three typical FOVs and the cell brightness distributions for the pB24, pB and WT strains. The track time distribution for the three sample is also shown. Overall it seems the pB sample showed greater brightness values but the pB24 sample was indistinguishable from the WT. The reason for this is unclear but could be due to variation in probe uptake.

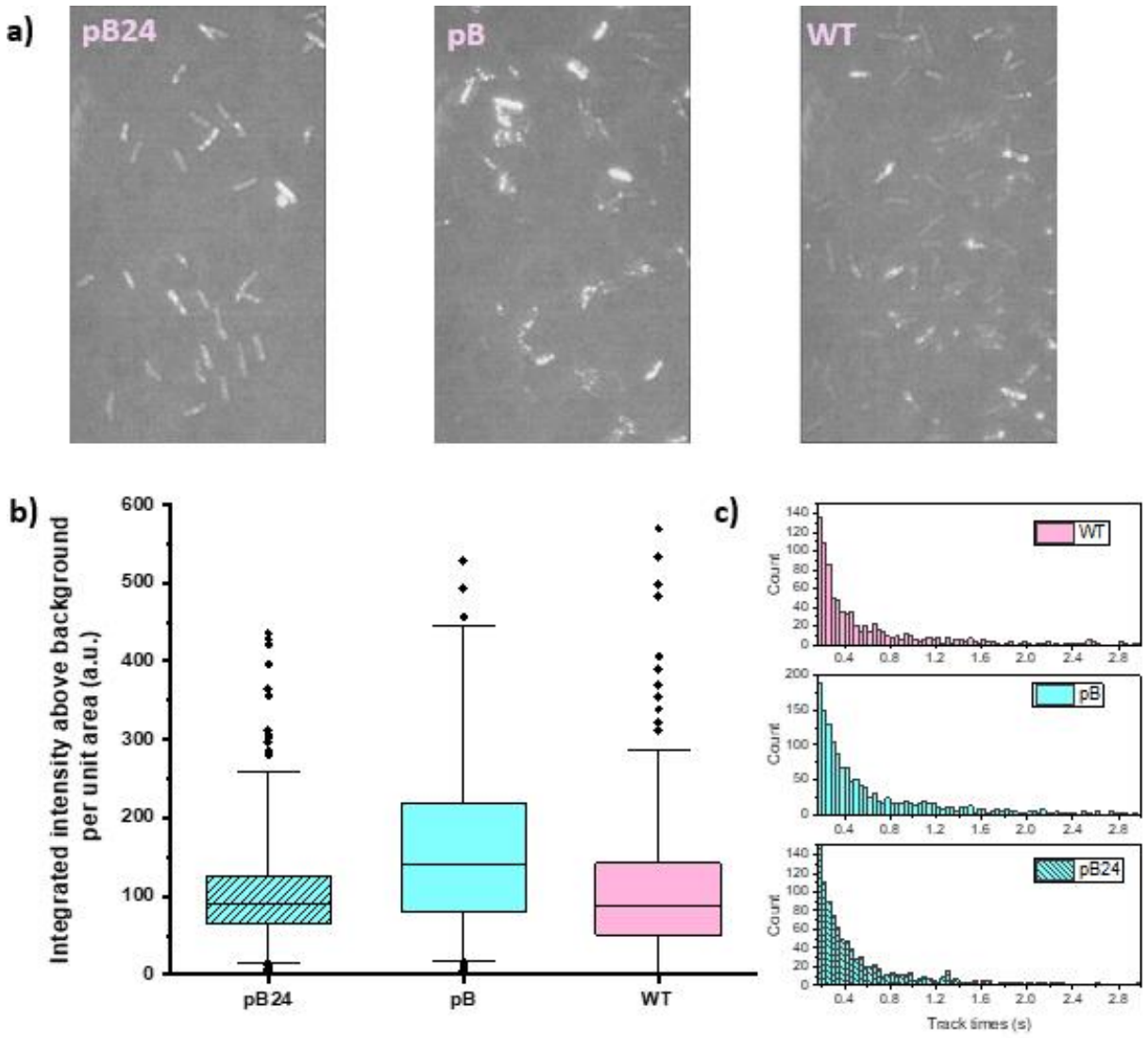

Figure 3-1: **a)** Three FOVs, representing pB24, pB and WT samples. Visually they look similar, with high probe uptake. These cells were electroporated with 50nM 2xATTO647N probes. **b)** A box plot showing the intensity distributions. In this experiment the pB sample had cells that were in general brighter than the pB24 and WT samples. This could be chance, or a small degree of specificity. However, the pB24 sample showed signal of the same level as the WT. It is possible that the 24 repeats in the plasmid disrupt the binding of the probe through transcript secondary structure, or induction of the transcript is too low. **c)** Three histograms showing the track lengths for the three samples. The histograms are all very similar, suggesting that the interactions with the target are not limited by the ability for further probes to bind, or that the majority of the interactions are non-specific.

3.2. Colocalisation strains

The colocalisation strains were engineered as a further test of specificity. A strain was engineered for a plasmid strain (pysynth) and a genomic strain (rpoB-ParS). Firstly, the pysynth plasmid contains a target cassette consisting of 6 evenly spaced 8mer targets at the 3' end of a random 1 kb sequence. 12 TetO sequences were added upstream of the promoter that are specific to a TetR-GFP protein produced from the genome. The rpoB-ParS strain contains the same target cassette added to the 3' UTR of the rpoB gene, as with the rpoB strain. About 18 kb away from the rpoB site on the genome a ParS site was inserted, which is specific to ParB-EGFP. The ParB-EGFP fusion is produce on a plasmid. Colocalisation events, between the ParS/TetO binding sites and the ATTO647N channel, would suggest that the foci are nascent RNAs undergoing transcription. If the spot could be tracked until the end of transcription, it would be expected to diffuse away from the ParS/TetO site when the transcript is released.

It was difficult to tell whether the pysynth strain showed any colocalisation. Many cells did not show any foci in the green channel. Candidate colocalisation events were identifiable, but were rare with << 1 per FOV. This is not surprising, as the copy number of the pysnyth plasmid as high as 50. The probability of being able to detect the TetO site and observe a transcript event is low. A two-channel image from the pysynth data is shown in **Figure 3-2a.**

The rpoB-ParS strain did show some convincing colocalisation events that occurred once every 2-3 FOVs. This is evidence that there is some degree of specificity. However, it is hard to make any concrete conclusions without a statistical analysis of the colocalisation rate as there were many more foci that did not colocalise. It was decided instead to pursue further modifications of the method, to try and improve specificity. **Figure 3-2b** shows some colocalisation examples, but it should be noted these events are rare.

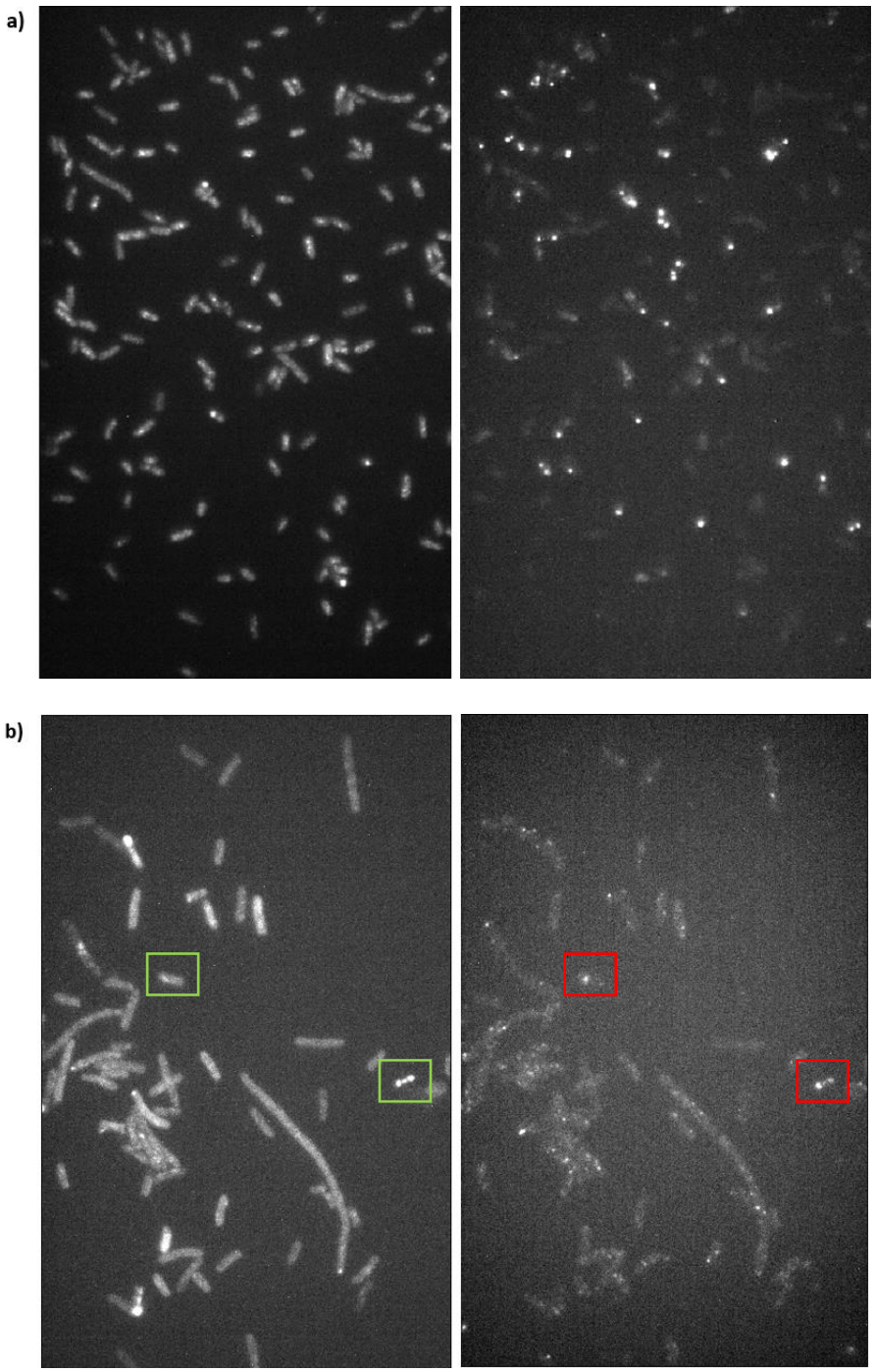

Figure 3-2: a) Left: green channel image from video of the pysynth strain. Many cells show clear foci, presumably at least one plasmid. Right: the corresponding red channel 2xATTO647N image. Although clear foci are visible in both, few obvious colocalization events were detected. **b)** Left: green channel image from the rpoB-ParS strain. A few stationary spots are visible, likely to be the chromosomal site, although could possibly be ParS aggregates. Right: corresponding red channel 2xATTO647N video. Some colocalization events are visible, as demonstrated by the boxes on the images. Several FOVs were needed to see a colocalisation event on average.

3.3. 13mer target plasmid – pB24x13

Other than photobleaching, the specific events should be limited by the transient nature of the 8mer hybridisation. For this reason, it was thought that a longer probe could be more effective, as it would have a longer dwell time than an 8mer. A 13mer was chosen, sequence *TCCACCGTCGATA*. A plasmid was ordered containing a 24-target cassette under an IPTG induced promoter. Experiments were carried out as before, but IPTG was added to the agarose pad to a final concentration of 10 nM and incubated for at least 30 min before imaging.

The use of a transcript with 24 targets for a 13mer should increase the probability of the probe to bind the transcript, rather than take part in some non-specific interaction. Also, the hybridisation events will be longer, as the free energy of the interaction is more negative. **Figure 3-3** shows the data taken from this strain. Although this strain did show tracks with steps at a higher rate than the pB strain (with 8 targets), it was still limited by a high proportion of immobile events, i.e. $D^* \approx 10^{-3}$ μm^2 s^{-1}. For this reason, it was decided to move towards other fluorophores altogether.

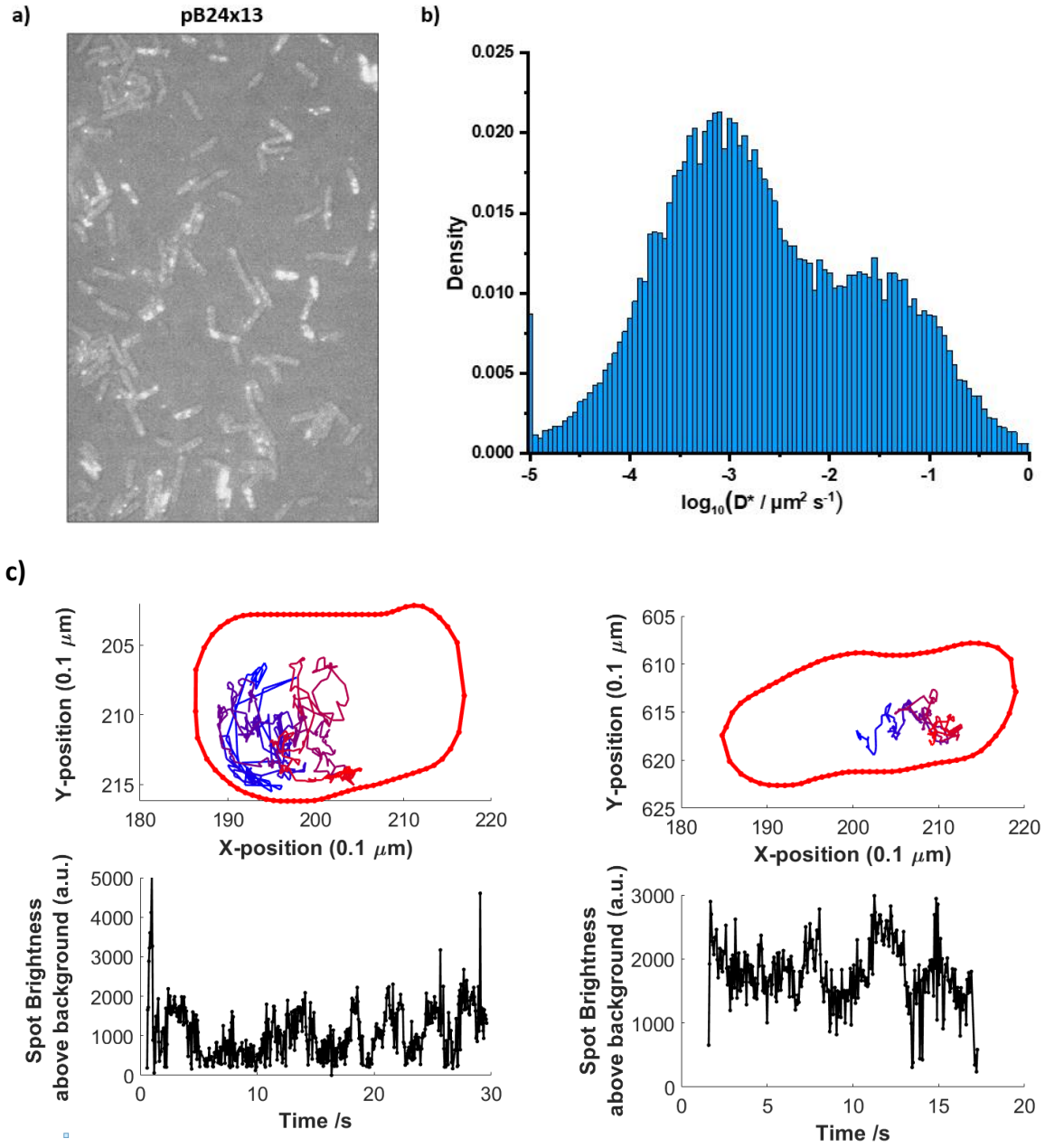

Figure 3-3: a) A FOV of *E. coli* containing the pB24x13 plasmid, after electroporation with 100nM of the 13mer 2xATTO647N probe. A similar distribution of cell intensities is visible as with the WT and pB samples shown previously. **b)** A diffusion population for the pB24x13 strain. The populations are very similar to the previous 2xATTO647N samples. This is likely due to the hydrophobicity of the dye leading to a significant immobile population. **c)** Interestingly, there are many tracks that show clear intensity steps, as the two exemplar tracks shown here demonstrate. These are much more common than in the 8mer samples, which could suggest a higher proportion of targets with multiple probes bound. However, the control dataset needs to be fully analysed to determine whether this could be a result of different quenching states that are relevant for the longer probe.

3.4. In vitro Fluorophore characterisation

Other fluorophores were tested to improve on specificity from the 2xATTO647N probes. There are a range of organic fluorophores that have been used *in vivo,* and some fluorophore chemistries may be more effective than others. There was a selection of different fluorophores in the lab, as well as several batches of the 2xATTO647N 8mer. Notably, two other 8mer probes were tested, with the same sequence as the 2xATTO647N probe, *TCCACCGT*. One was doubly labelled with ATTO655, the other was labelled with ATTO643 at the 3' end and a quencher at the 5' end, Black Hole Quencher 1 (BHQ1). Firstly, these new probes were tested to determine their degree of quenching and therefore suitability for *in vivo* FISH. These probes were taken from another *in vitro* project. A first step is to ensure that the probes quench and hybridise *in vitro*, using a fluorimeter.

The fluorimeter determines the degree of quenching and hybridisation for each probe. Initial experiments were carried out in 50 mM Hepes, 200 µM NaCl, 200 mM $MgCl_2$. Firstly, a background reading of 50 µL of the buffer was taken. Then 0.5 µL of 5µM probe was added to reach a final concentration of 100 nM. This gives the background where any fluorescence is due to imperfect quenching in solution. Then, 0.5 µl of 100 µM complementary target was added. This was repeated until the fluorescence reached a maximum, always less than 400 nM final concentration of target. All probes had the same sequence and therefore target, and were diluted in MilliQ. The probes were tested sequentially, and the sequence was repeated 3 times, ensuring the cuvette was thoroughly washed each time.

The data from the fluorimeter is shown in **Figure 3-4**. The original 2xATTO647N stock, produced by a collaborator in the Brown group (department of chemistry, University of Oxford) in 2017 showed an Fluorogenic Factor (FF) of ~ 4. This was slightly less than the newly synthesised probes, also in the Brown group but in July 2024. This suggests that there has been some degradation over time. The Biomers stock shows a poor FF <3. This is likely due to a lower labelling efficiency. Furthermore, the 2017 and 2024 probes from Brown group were doubly purified using HPLC but only a single run of HPLC was used by Biomers. The double labelled 2x ATTO655 probes shows a low FF < 2. This is likely due to the fluorophores being hydrophilic, meaning they do not associate closely in solution, and therefore contact quench, as tightly as the hydrophobic 2xATTO647N probes. Finally, the ATTO643-BHQ1 probe had an

FF ~ 14. This is due to the BHQ1 and undergoing FRET quenching. It seems from this data that FRET quenching is a more effective mechanism than contact quenching.

Interestingly, further addition of the complementary target did not lead to saturation of signal, but a consistent decrease. This can be explained to some extent by dilution, as 0.5 µL of solution was added each time. This would be expected to decrease signal by a few percent, but the decrease was up to 25%. For this reason, the peak of the signal was taken, usually around 200 nM of complementary target.

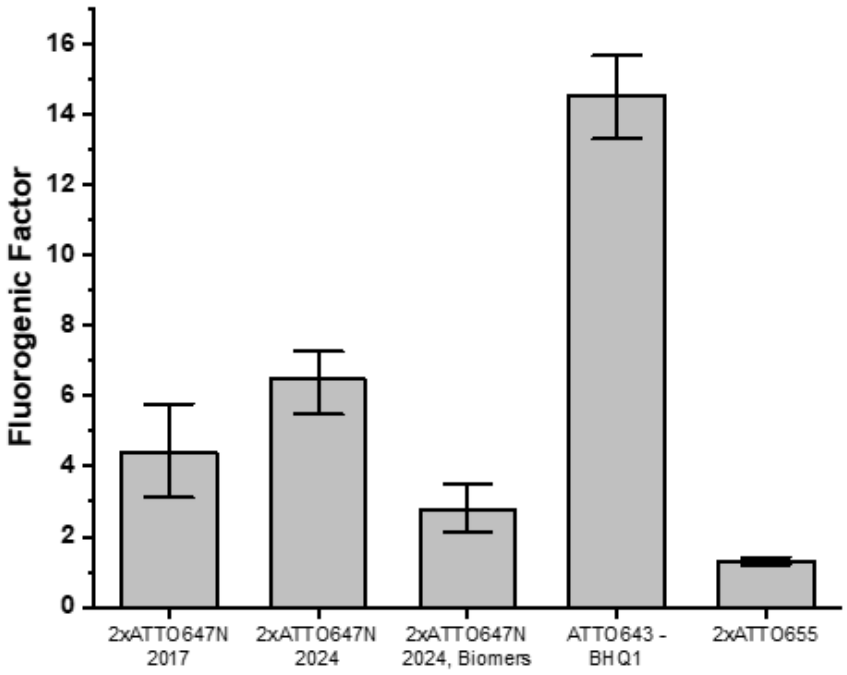

Figure 3-4: *In vitro* measurements of the Fluorogenic Factor (FF), defined as the ratio between the fluorescence measured in the absence of target and the maximum fluorescence obtained when saturated with the complementary target. This was measured for each of the 8mer probes. This includes three 2xATTO647N probes from various ages and sources. The contact quenched probes all performed poorly when compared to the FRET quenched ATTO643xBHQ1 probe.

The fluorimeter measures steady state ensemble fluorescence, rather than a single molecule approach. It is not necessarily clear whether the fluorophores are emitting in their quenched state, or momentarily de-quenching. It is also worth looking at the absorbance

spectra to get an idea of the probe behaviour. The Nanodrop can give an indication of the ratio of species within the solution and absorption spectra from a variety of probes used are shown in **Figure 3-5**. For example, relative heights of the two peaks in the 2xATTO spectrum inform qualitatively on the ratio of free dye/singly labelled DNAs to quenched species. Roughly, the higher the left peak relative the right indicates more quenching. This indicates that the low FF seen for the 2xATTO647N Biomers sample is due to a higher proportion of singly labelled DNAs.

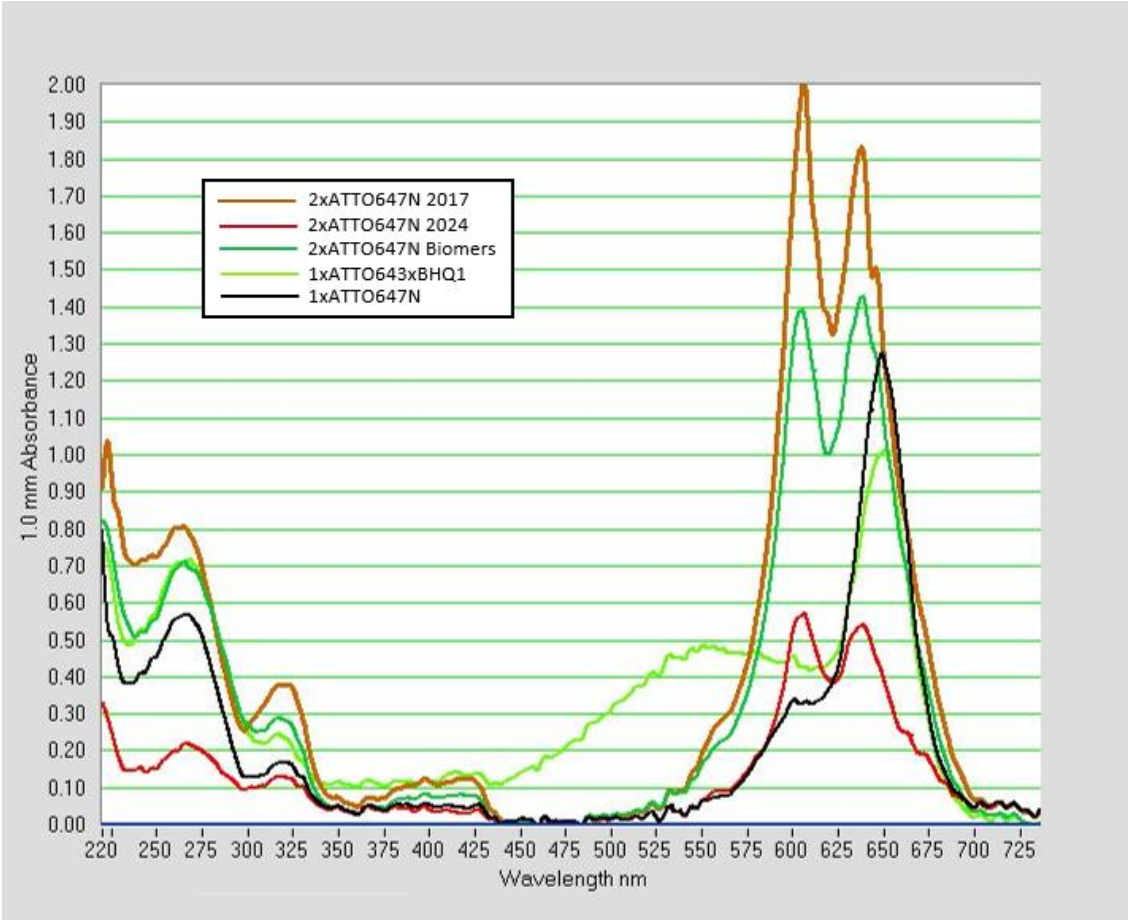

Figure 3-5: A screenshot of the nanodrop results from a selection of probes used in the previous section. These absorbance spectra were used to make predictions regarding the quenching efficiencies of the probes,

3.5. 2xATTO655

The first other fluorophore to be tested *in vivo* was the double labelled 2xATTO655 probe. As previously mentioned, the probe is an ineffective quencher. For this reason, it is unlikely that the 2xATTO655 probe will be useful for tracking RNA. However, it will be interesting to assess the background signal as it will give an idea of the dependence of the non-specific interactions of the fluorophore. **Figure 3-6** shows a typical FOV and the mobility distributions that resulted from the tracking. Unexpectedly, the uptake of the fluorophore is significantly lower, as evidenced by much lower intensity values seen in the cells. It is possible that there was an issue with the electroporation for this sample. Despite this, there are many trackable foci throughout the video. The mobility distributions show a clear shift towards the mobile population, but the immobile population is still significant. Comparing this data to the 2xATTO647N data shows that events are dependent on the choice of fluorophore.

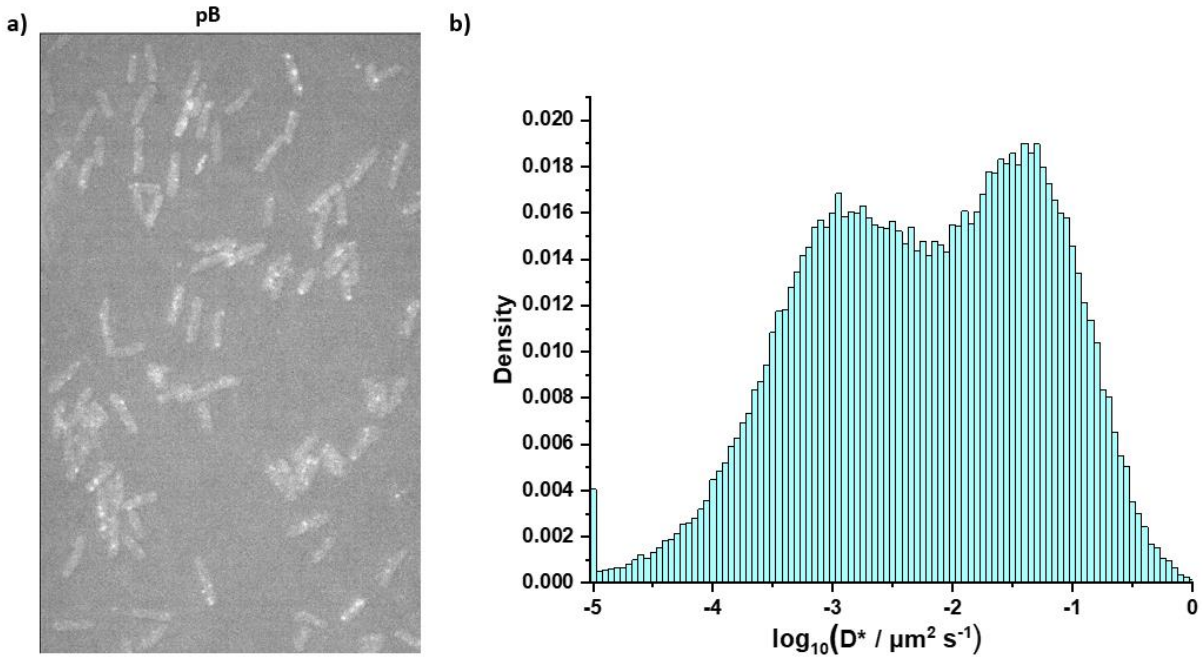

Figure 3-6: a) A FOV of pB positive *E. coli* incubated (but not electroporated) with 50 nM 2xATTO655 probes. Notice that despite the apparent lack of quenching, the cells are less bright than that of the 2xATTO647N probes. This may be attributed to fewer non-specific interactions than the 2xATTO647N probe as a result of lower hydrophobicity. Alternatively, there may be variations in probe uptake. This issue is discussed for ATTO643 in **section 3.7**. **b)** The diffusion histograms for the pB show a smaller proportion of the immobile population at $D^* \approx 1 \times 10^{-3}$ µm^2 s^{-1}, compared to the mobile population at $D^* \approx 1 \times 10^{-1}$ µm^2 s^{-1}.

3.6. ATTO643-BHQ1

Due to the promising quenching results in steady state ensemble fluorescence, the ATTO643-BHQ1 probe was deemed a promising imager. There should be a lower amount of background signal as a result of the high quenching. Furthermore, ATTO643 is a hydrophilic dye, and so may be less likely to take part in non-specific interactions (*ATTO 643*). If this is true, then a higher concentration of probes can be tolerated inside a cell before the cell becomes flooded with non-specific signal. This may allow for a higher local concentration to be reached in the vicinity of the RNA which would increase the probability of achieving multiple probes on the RNA.

Indeed, this was seen, as concentrations of up to 500 nM were routinely used. However, the samples were not electroporated. This is evidence that the probes quench well *in vivo*, however it does need to be shown that the probe uptake is unaffected by the change in fluorophore, which is discussed in **section 3.7**. It is possible that the hydrophobic ATTO647N fluorophore aided the probe to penetrate the membrane. At 500 nM ATTO643-BHQ1, the cells show clear foci but are visually less bright than their 2xATTO647N counterparts at 50 nM. **Figure 3-7** shows a FOV from each of the strains used here. As a result of this, the autofluorescence plays a larger role in the cellular integrated intensity. To factor this in, each cell was classed as containing signal if it had at least one track in at some point in the video. Cells that showed no tracks were designated 'empty'. This is classified in **Figure 3-8**, where it is apparent that the target carrying strains contained a higher proportion of cells of higher brightness containing tracks. This is positive indication of specificity but requires further validation due to potential variation in probe uptake.

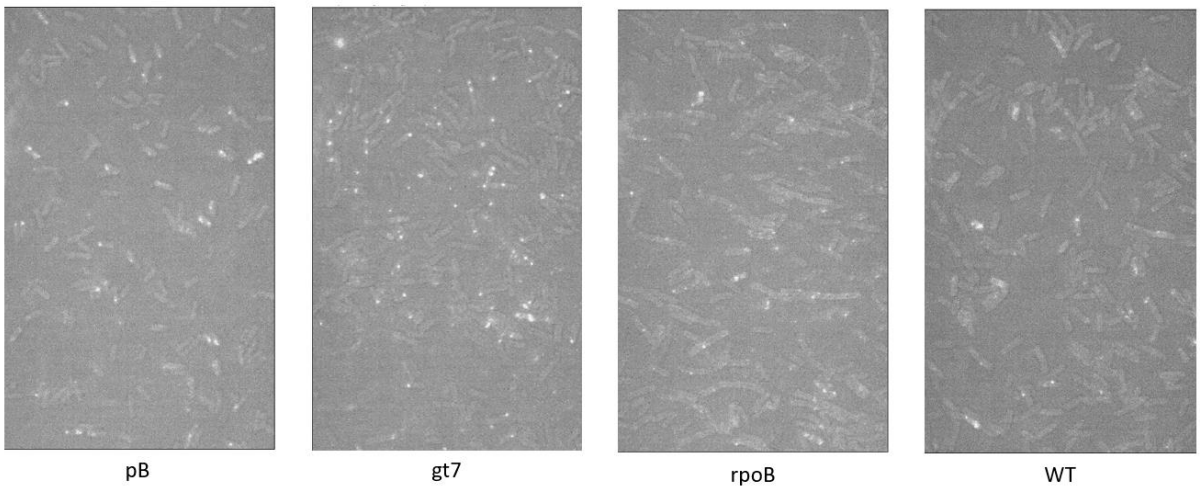

Figure 3-7: Typical FOVs for the four 8mer target strains, incubated (but not electroporated) in the presence of 500nM ATTO643-BHQ1 probes. It is clear that the cells are overall less bright than the 2xATTO647N samples. Despite this, clear foci are visible which produce many tracks. It is hypothesised that the significantly reduced signal is due to lower probe uptake. There is also heterogeneity, where many cells are empty but a few are relatively very bright

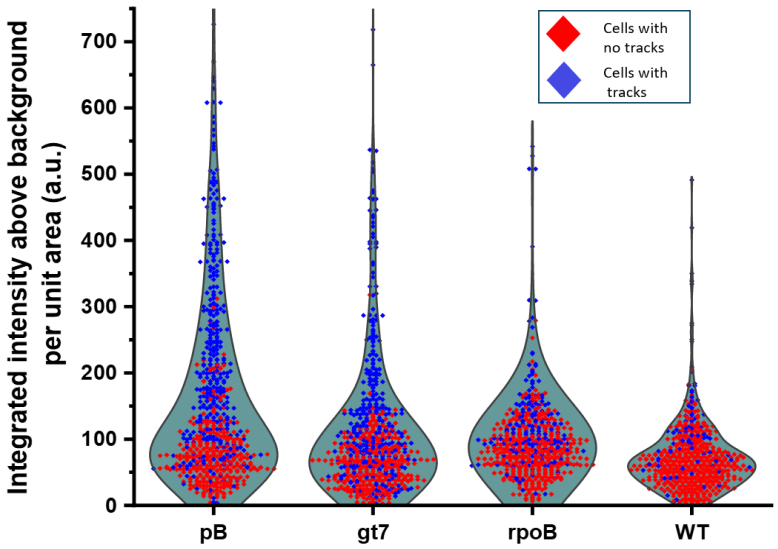

Figure 3-8: Violin plots of the cellular integrated intensity for the four strains are shown. The data points are overlaid and showed in red for cells with no tracks observed at any point in the video. Blue data points show cells that did contain at least one track. As with the 2xATTO647N data there is a long tail of bright cells whilst most show minimal signal. The fact that the cells with no tracks give non-zero signal is due to the autofluorescence of the cells

In addition to whole cell intensity measurements, the foci were localised and tracked as with the 2xATTO647N and 2xATTO655 probes. As before, intensity steps showing steps is a suggestion of probes binding to the same mRNA. In fact, a much greater proportion of tracks did indeed show tracks. However, these were present in the WT control in similar numbers. As a result it was concluded that a single ATTO643-BHQ1 probe has multiple intensity states that it can explore whilst bound to the same object. Some tracks in the WT control that illustrate this behaviour are shown in **Figure 3-9**.

Mobility distributions were calculated identically to the 2xATTO647N data. The distributions for the pB, rpoB and gt7 strains plus the WT control are shown in **Figure 3-10**. This is consistent with ATTO643 being more hydrophilic. Overall, the populations are shifted towards more mobile tracks, with the significant $D^* \approx 1 \times 10^{-3}$ μm^2 s^{-1} populations diminished. The pB and gt7 populations show stark differences compared to their corresponding WT control data, but the rpoB is less clear. This is perhaps due to a lower specificity to what may be a low copy number transcript. Alternatively, it may be due to the slow-moving transcript D^* population having a greater overlap with the WT non-specific interactions. However, it should be noted that this data remains preliminary and requires significant verification.

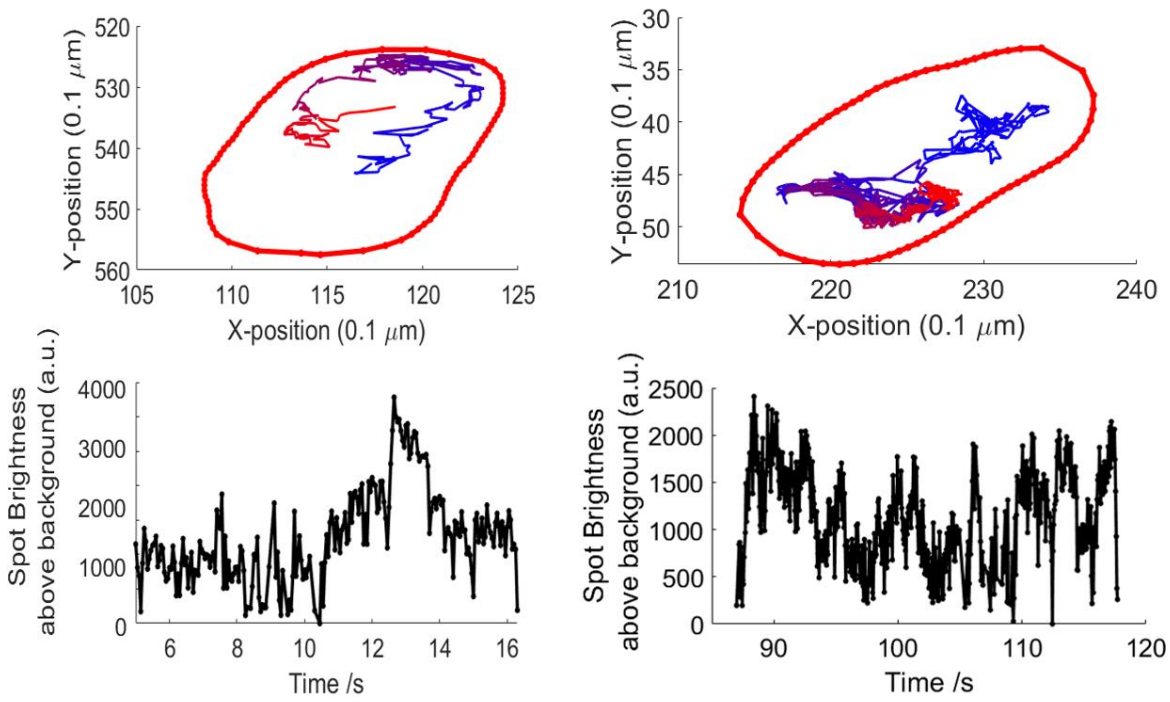

Figure 3-9: Exemplar tracks shown from the WT ATTO643-BHQ1 data. These tracks were selected due to the steps visible in their intensity traces. Tracks showing such steps are common in all samples, with multiple occurrences per FOV. Because these tracks were selected from WT data it would be unexpected to contain spots that had multiple probes binding to the same places. If it is assumed that these tracks are formed of a single fluorophore this must be a result of different probe photophysics. One possible explanation is due with the quencher. When the probe is hybridised, then the average distance between the fluorophore and the quencher will increase, leading to an increase in brightness. However, as this is FRET quenching, and the distances involved are still small (length of 8mer is ≈ 2.5nm), the FRET efficiency will not be zero. However, if the quencher blinks or photobleaches, then the FRET efficiency could become zero temporarily. This would lead to an additional increase in intensity, which would produce steps in the intensity trace. This builds an extra level of complexity to the data and may make it more difficult to extract genuine mRNA traces.

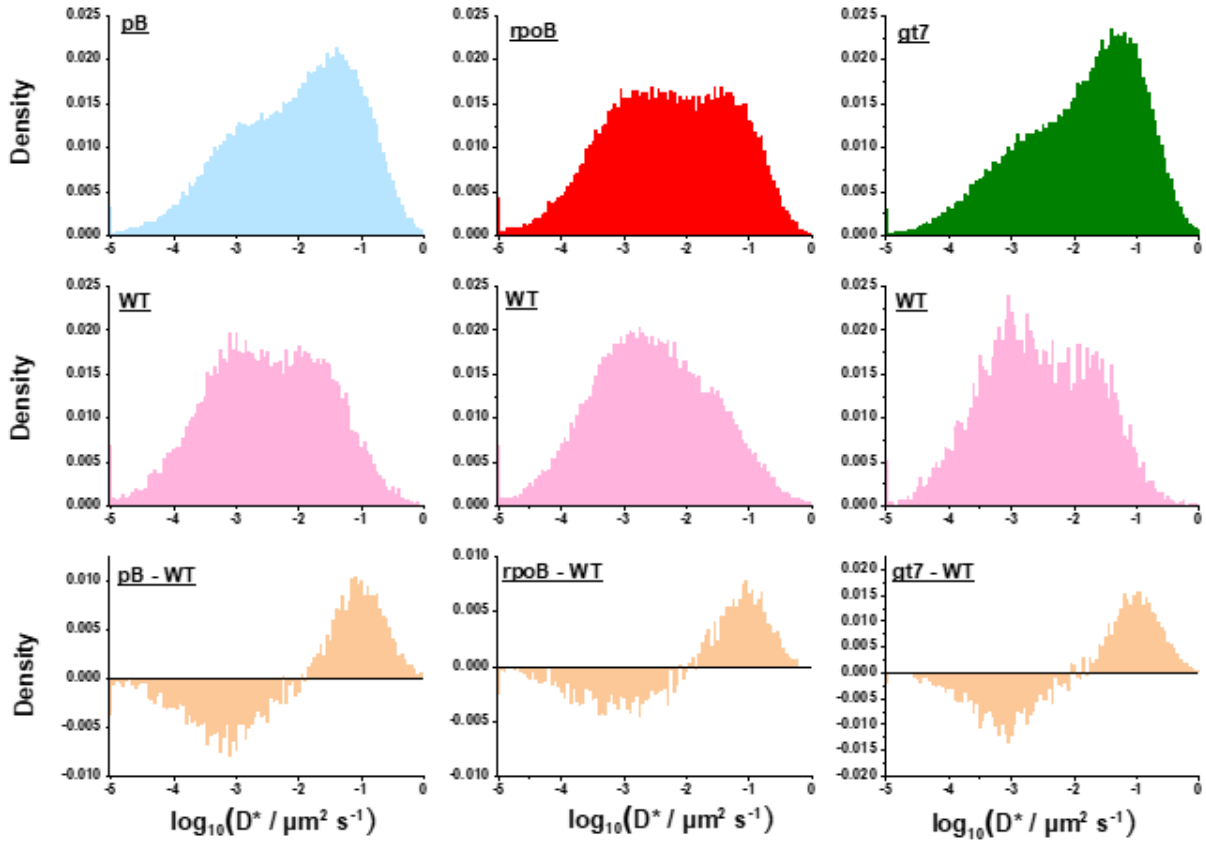

Figure 3-10: Mobility populations from three experiments that covered the 8mer strains. The plasmid strain, pB, shows a significant peak at around $D^* \approx 0.05$ μm^2 s^{-1} and a minor peak at $D^* \approx 1 \times 10^{-3}$ μm^2 s^{-1}. This is in contrast to the 2xATTO647N data, where the $D^* \approx 1 \times 10^{-3}$ μm^2 s^{-1} was the predominant population. The rpoB data is similar to the 2xATTO647N, showing a broad peak across the range of mobilities. The third strain, gt7, shows a very similar population to pB, with a dominant population at $D^* \approx 1 \times 10^{-1}$ μm^2 s^{-1}. For each of the experiments, the WT data shows a broad peak with less obvious subpopulations. This is again different from the 2xATTO647N data, as the peak at $D^* \approx 1 \times 10^{-3}$ μm^2 s^{-1} is no longer dominant. All three strains have the result that the WT data is more populated at the lower mobilities, $D^* \approx 1 \times 10^{-3}$ μm^2 s^{-1}, and less so for higher mobilities $D^* \approx 1 \times 10^{-1}$ μm^2 s^{-1}.

It is clear that despite the minimal differences in intensities between the control and target containing strains, it seems there may be differences in the mobility populations. This suggests that the probe is binding to the complementary target but not quenching effectively when unbound. It is worth comparing the *D** values obtained here to values reported in the literature. Encouragingly, a recent study suggested that a 4 kb transcript had $D^* \approx 0.9\ \mu m^2\ s^{-1}$, when freely diffusing and not bound by ribosomes (Sattler and Graumann, 2021). This is similar to the higher mobility gt7 population in **Figure 3-10**. Notably, the gt7 strain transcript has no ribosome binding site. Less encouraging is the pB strain which has similar populations but a transcript of only 120 bases. It is possible that a lower exposure time is needed to track this population and the dominant population at around $D^* \approx 0.3\ \mu m^2\ s^{-1}$ is in fact the tracks of the plasmid. Values in the literature are typically much lower for plasmid *D** but these involve binding of fluorescent proteins to larger plasmids (Reyes-Lamothe *et al.*, 2014).

3.7. Probe uptake

The electroporation protocol was developed previously in the lab, involving internalisation of fluorescent dyes and proteins (Crawford *et al.*, 2013; Plochowietz, Crawford and Kapanidis, 2014). However, the internalisation of short, ssDNA probes in *E. coli* has not been fully characterised. It was found previously that there was significant probe uptake by non-electroporated cells when using the 2xATTO647N 8mer. The mechanism for this is unclear, but proposed to be due to the small size of the probes and the heat shock from removing the cells from the -80 °C freezer. The degree of which electroporation increases probe uptake is unclear. The media the cells are cultured in, the process of making the cells electrocompetent, probe concentration and degree of heat shock all play a role in probe internalisation.

To test the alternative internalisation of the probe, LB-grown cells that were cultured on the day (and therefore not frozen) were chosen. Two aliquots of these MG1655 cells were incubated with a final concentration of 100 nM of the 13mer 2xATTO647N probe. This probe was chosen as it is the longest 2xATTO647N probe available and therefore should minimise alternative routes through the membrane. One sample was electroporated in the presence

of the probes and the other purely incubated. This process was repeated for the ATTO643-BHQ1 8mer probe also. The FOVs are shown in **Figure 3-11**.

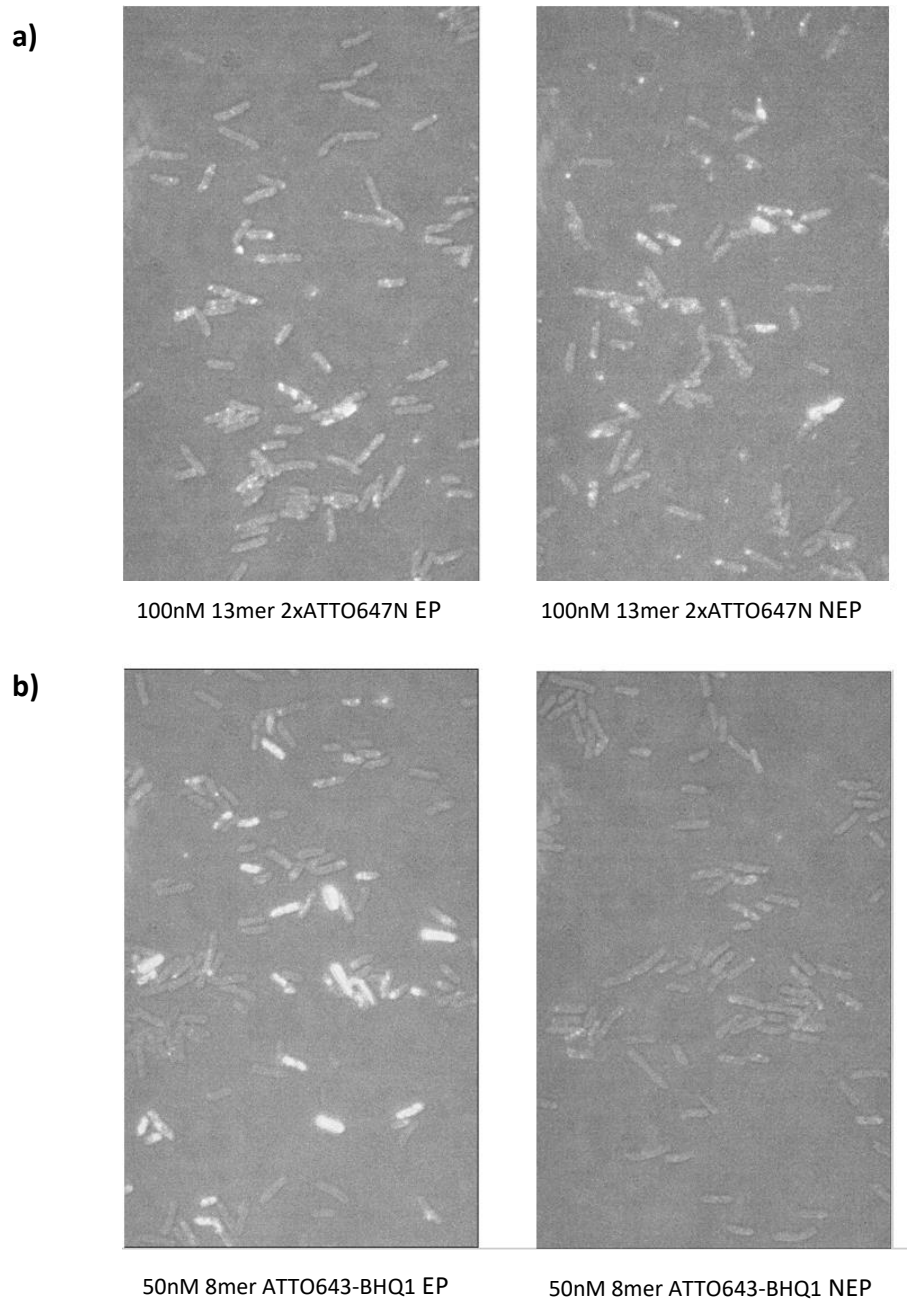

Figure 3-11: Four representative panels of an electroporation test are shown**. a)** WT cells in the presence of 100 nM 2xATTO647N probes. The FOV on the left was electroporated (EP) and the FOV on the right was not. Visually there is minimal difference between the FOVs with a high degree of heterogeneity and similar probe internalisation. It was consistently observed that electroporation only played a minor role for the 2xATTO647N probe internalisation. **b)** WT cells in the presence of ATTO643-BHQ1 probes. Here the difference between the samples is obvious meaning that EP is essential for probe internalisation.

Simple estimates of the brightness of cells with and without electroporation suggest that the electroporated population is marginally brighter for the 13mer 2xATTO647N. However, it is clear that electroporation makes a very significant difference for the ATTO643-BHQ1 8mer. It is therefore likely that the hydrophobicity of the fluorophore is playing a role in probe internalisation. ATTO643 (and BHQ1) are hydrophilic and so cannot pass through the membrane passively, like ATTO647N.

This leads to an important question as to whether electroporation is a sensible choice. Electroporation damages cells and therefore may reduce the accuracy of biological findings due to the cells exhibiting a stressed state. After electroporation at 1.4 kV, only two-thirds of the cells continue dividing (Plochowietz, Crawford and Kapanidis, 2014). For the ATTO647N probes electroporation is not recommended for this reason. Even for ATTO643-BHQ based probes it may be preferable to incubate with a high concentration of probes, say >500 nM which achieves similar loading efficiencies as electroporation at 10 nM, but without damaging the cell physiology.

Other methods for probe internalisation were explored, including chemical proration and heat shock using DMSO and TSS buffer. These methods show a lower rate of internalisation and due to the existing, published protocols regarding electroporation it was decided that the issue did not need to be pursued further for the purposes of this thesis.

3.8. Summary

The preliminary data shown in **Figure 3-1** suggests that the addition of further binding sites is not beneficial to the system in general. This is evidenced by the 24-target plasmid strain, pB24, performing similarly to the pB strain with just six binding sites. Strains with GFP binding sites inserted in the proximity of the target gene, pysnyth and rpoB-ParS were designed to test specificity. Interestingly there were colocalisation events visible with the rpoB-ParS every few FOVs, demonstrated in **Figure 3-2**. Although a 13mer 2xATTO647N target showed potential to improve on the shortcomings of the 8mer 2xATTO647N, due to the tracks shown in **Figure 3-3**, it was not pursued further due to the desire to move away from the ATTO647N fluorophore. The importance of *in vitro* characterisation was introduced in **Figure 3-4** and

Figure 3-5. Steady state fluorescence and absorption spectra were useful to predict how the probes may behave *in vivo.*

The 2xATTO655 and ATTO643-BHQ1 probes produced different behaviour in terms of both cell intensities and mobilities. **Figure 3-6 - Figure 3-10** present the data on these fluorophores. The mobility distributions of the ATTO643-BHQ1 are suggesting a range of biological behaviours not seen in the WT control. This has opened up an avenue of research based on selection of the optimum fluorophore-quencher system. The data presented in this chapter is in the preliminary phase and needs to be verified. However, the conclusion is that the ATTO643-BHQ1 probe leads to different behaviours and is a promising avenue of further work. Due to questions surrounding probe uptake and the intensity steps seen in the WT control there is insufficient evidence to conclude that the probe is specific to the target at this stage.

Finally, it was shown in **Figure 3-11** that electroporation has a minimal impact on 2xATTO647N probes, but significantly increases the probe uptake for the ATTO643-BHQ1 8mer probe.

It was decided that the best course of action was to redesign the probes for the project considering the data from the past two sections. This is to achieve clear specificity both in terms of mobility and brightness.

4. Further work

The key to achieving an effective method for *in vivo* smFISH is an effective probe. Learning the lessons from the previous work, new probes were designed which aim to improve on the method further. This section will describe the reasoning behind the design and the results they aim to obtain.

4.1. Probe Design

An important consideration is the number of off-target matches for the probe complementary target in the *E. coli* genome. An online Basic Local Alignment Search Tool (BLAST) provided by the National Centre for Biotechnology Information (BLAST: Basic Local Alignment Search Tool (nih.gov)) was used to check this for the MG1655 *E. coli* strain. 82 exact matches and around 2000 single base mismatches were found for the 8mer TCCACCGT probe. It is possible that these genes are poorly expressed, but in such large numbers they may form a significant proportion of the non-specific interactions. For this reason, it is sensible to choose a probe with as few of these off-target matches as possible. This is of course easier for longer probes.

It is also desirable for the probe to have as high a fluorogenic factor as possible. To achieve this, the fluorophore and quencher (or two fluorophores) need to be in as close proximity as possible when the probe is not hybridising. It was thought that utilising a molecular beacon type approach, i.e. a stem loop system, would be effective. This is also based off work by Kapanidis group member Miss Mirjam Kümmerlin, where a ribosomal RNA probe shows higher than expected fluorogenic factor due to formation of a loop. The stem of this loop was kept, and the rest of the probe was iterated through to find the sequence that has the fewest matches in the MG1655 genome. The probe sequence was the tweaked to modify the annealing temperature by changing the GC content, for example. If the stem is too tightly bound, then the on-rate of hybridisation will be too low and there will be insufficient binding. On the other hand, if the stem is too weakly bound there will be insufficient quenching. Another factor that will affect the on-rate for hybridisation is the inclusion of a region before the stem that is free to hybridise. DNA and RNA folding simulations were done

online, using the Mfold web server (Zuker, 2003). This was done mainly to aid the choice of the stem-loop structure. The output from this process is shown in **Figure 4-1**.

Finally, a suitable fluorophore-quencher pair needs to be chosen. This is not simple, as there is little understanding on how fluorophores behave *in vivo* in the literature. However, given the results of the previous section it was decided to continue using ATTO643. ATTO643 also performed well in several metrics compared including specificity *in vivo,* albeit in eukaryotic cells (Steen *et al.*, 2024). Given the high fluorogenic factor of the ATTO643 - BHQ1 pair relative to the 2xATTO probes, it was decided to use a quencher. It was decided to use BHQ2, as it has a higher absorption spectrum overlap with ATTO643 than BHQ1. This means that the FRET quenching will be more efficient. It is possible that the probe will not fluoresce upon hybridisation due to the FRET efficiency remaining high despite the spatial separation induced by hybridisation. In this case, it would be sensible to redesign the system to use a longer probe, or switch to BHQ1. In fact, steady state ensemble fluorescence data showed promising results, as seen in **Figure 4-2** where a fluorogenic factor of > 25 was measured.

An alternative method is to use a repetitive sequence. A 200 base AC repeat, for example, would provide a large surface for a TG repeat probe to probe to anneal too. This method was used successfully for imaging of membrane protein dynamics, albeit on the surface of a eukaryotic cell (Niederauer *et al.*, 2023). An AC repeat was chosen to avoid self-binding of the transcript leading to secondary structures that left the target region inaccessible. Particularly, uracil was excluded due its ability to bond with guanosine (Hermann and Westhof, 1999) . Similar to the hairpin probe, an 18mer TG repeat ATTO643-BHQ2 probe was chosen. As before, secondary structure simulations of the probe and the target were done on Mfold to check (Zuker, 2003).

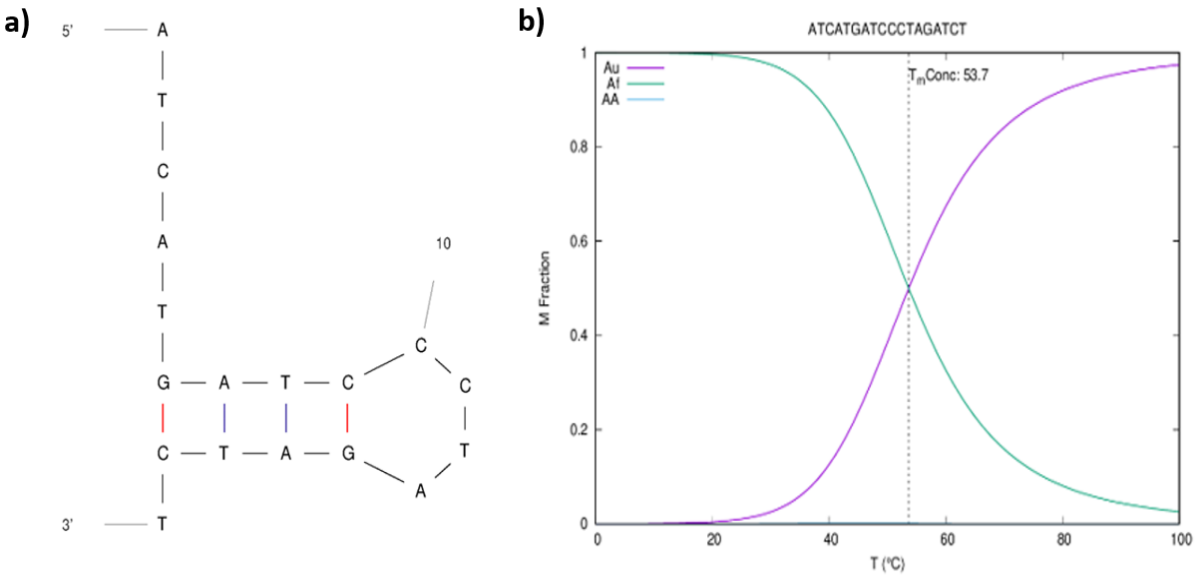

Figure 4-1: Two plots produced by Mfold for a proposed stem-loop probe (Zuker, 2003). **a)** Predicted secondary structure. There is a four base pairing which forms a stem, and a short four base loop. There has been a region of five bases (ATCAT) left before the stem. In theory, this will aid hybridisation as there is a significant complementary region not contained in the rigid structure of the stem. **b)** A secondary structure state graph, plotted against temperature. The green curve represents the fraction of probes that are in the stem-loop formation, whilst the purple curve is the fraction of probes that show no secondary structure. The imaging is typically done at around room temperature, T = 20-25°C. At this temperature, the probes at the point where almost all are in the stem-loop formation, but some are exploring a disordered state. This should ensure a high degree of quenching whilst still permitting hybridisation.

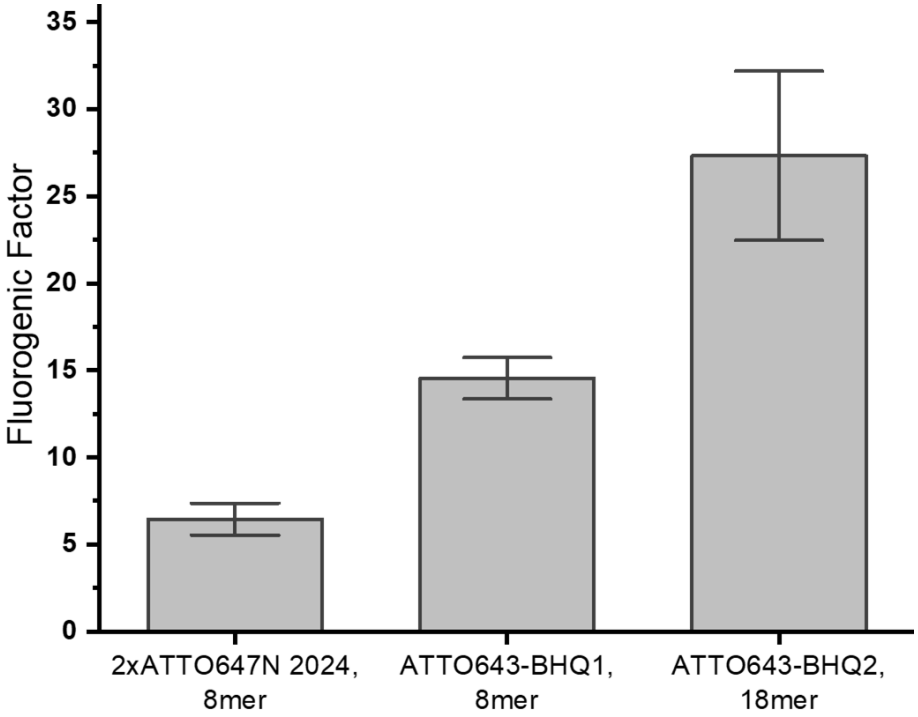

Figure 4-2: Steady state ensemble fluorescence measurements of the fluorogenic factor; defined as the ratio between the background signal with no target and the maximum signal observed when complementary target added. The newly designed ATTO643-BHQ2 18mer probe significantly outperforms the doubly labelled 2xATTO647N 8mer. It is also an improvement on the ATTO643-BHQ1 8mer probe. This indicates that there is a high degree of quenching and importantly the BHQ2 does not prevent fluorescence upon hybridisation. However, this is no guarantee that the probe will perform well *in vivo* as the degree of non-specific opening due to interactions with intracellular components is unknown.

4.2. Endonucleases and SSB

An important consideration is whether probes will be robust and stable inside cells for sufficient periods of time. It was chosen to explore the possibility that probes are degraded by endonucleases, as this would have a detrimental effect on specificity. Endonucleases could degrade the backbone of the probe, removing the quenching and causing fluorescence. This was tested using DNAse1 obtained from New England Biolabs.

The 2xATTO647N 8mer was tested using steady state bulk fluorescence. Firstly, a DNase1 buffer was made: 10 mM Tris-HCL, 2.5 $MgCl_2$, 0.5 mM $CaCl_2$. The background fluorescence of 50µl of this buffer was measured. Next, 0.5µL of 2xATTO probe was added to a final concentration of 100 nM. This leads to a significant increase in fluorescence due to incomplete quenching. Then 1 µL of DNase1 was added and the fluorescence measured for 60 min, where the fluorescence had stopped increasing. This was assumed to be the point where degradation was complete. Finally, the complementary target was added for a final concentration of 100nM, and no increase in signal was seen. The final signal was approximately threefold higher than before the DNAse1 was added, showing that the probes are vulnerable to degradation by DNAse1, and therefore other endonucleases.

This assay was done *in vitro* and the details cannot be transferred to the *in vivo* system, but the increase in signal shows that the probes are vulnerable to binding to endonucleases in general. This is a potential source of free dye that needs to be investigated further.

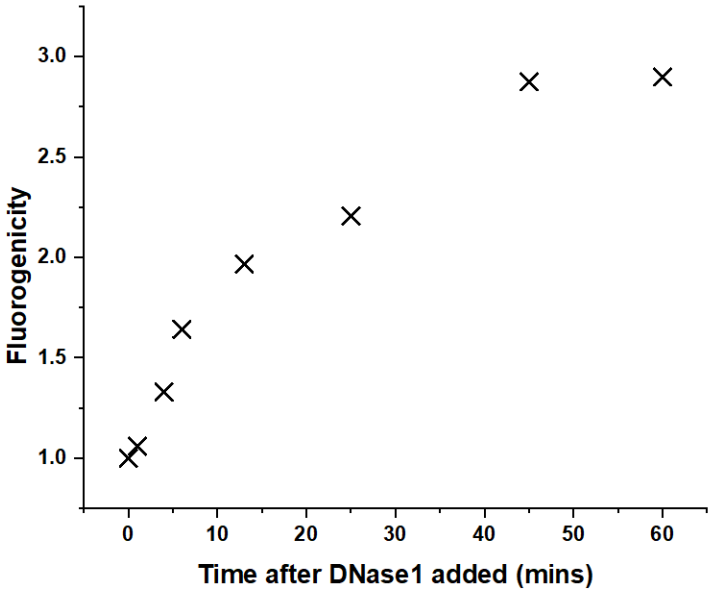

Figure 4-3: A time assay for a 2xATTO647N probe exposed to one enzyme unit of DNAse1 from New England Biolabs. It is clear that the DNAse1 has some nuclease activity on the probe, leading to digestion and the prevention of contact quenching. After 60 min the signal plateaus, suggesting that the probes are all degraded. Finally, 100 nM of complementary target was added to the system, and no signal increase was seen.

Dr Hafez El Sayyed has *E. coli* that are deficient of seven nucleases. This would provide an opportunity to test whether endonuclease activity is a significant factor in the non-specific interactions. Finally, this finding does suggest that the probes are redesigned using modified nucleic acids. These have been shown to be resistant to nucleases and other enzymes that may lead to probe degradation (Bratu *et al.*, 2003).

In a similar vein, it was hypothesised that single stranded DNA binding protein (SSB), which exists with high copy number in the cell, at least several thousand (Zhao *et al.*, 2019), may be binding to the probe inducing fluorescence. It was possible to test this *in vitro* also. ET SSB (Extremely Thermostable Single-Stranded DNA Binding protein) was obtained from New England Biolabs. Preliminary experiments suggest negligible interaction but further verification is required.

5. Conclusion

Firstly, this thesis presents a brief review of the current mechanisms for RNA tracking *in vivo,* and demonstrates the importance of understanding RNA behaviour in general. **Chapter 2** shows a selection of experiments testing a method previously designed in the lab to track RNA in *E. coli*, called *in vivo* single molecule Fluorescent *In Situ* Hybridisation (smFISH). The probes chosen were 8 base DNA probes doubly labelled with ATTO647N. These probes were expected to bind to a synthetic gene consisting of 6 binding sites, expressed from a plasmid. This target cassette has also been introduced into two genomic targets: the 3' UTR of the rpoB gene and a non-native T7 RNA polymerase. The existing method showed promise, but was limited by the non-specific interactions, which are driven by the characteristics of the 2xATTO647N probe sticking to intracellular components.

As a result of this, the existing method was extended to include RNA targets with additional binding sites, 13mer 2xATTO647N probes and strains with GFP colocalisation, all with the hope of improving specificity. Some clear tracks and colocalisation events were obtained, but the method was still limited. It was decided to test other fluorophores and quenching mechanisms through a 2xATTO655 and an ATTO643-BHQ1 8mer probe. The different fluorophores showed a range of different photophysics and *in vivo* behaviours. There is evidence that the mobility populations vary between the strains with an introduced target and the control, a WT MG1655 sample with no target introduced.

Finally, the conclusions obtained from characterising the variety of probes and target mRNAs were used to design a new system, using ATTO643-BHQ2 18mer probes. It is hoped that these new constructs will allow for accurate *in vivo* smFISH in future work. It is likely that future work will move away from the idea of brightness comparisons and focus more directly on the mobility populations.

Abudayyeh, O.O. *et al.* (2017) 'RNA targeting with CRISPR–Cas13', *Nature*, 550(7675), pp. 280–284. Available at: https://doi.org/10.1038/nature24049.

Adams, R.A. *et al.* (2021) 'Rifamycin antibiotics and the mechanisms of their failure', *The Journal of Antibiotics*, 74(11), pp. 786–798. Available at: https://doi.org/10.1038/s41429-021-00462-x.

Arora, A., Sunbul, M. and Jäschke, A. (2015) 'Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers', *Nucleic Acids Research*, 43(21), p. e144. Available at: https://doi.org/10.1093/nar/gkv718.

ATTO 643 (no date) *ATTO-TEC GmbH*. Available at: https://www.atto-tec.com/ATTO-643.html?language=en (Accessed: 17 September 2024).

Autour, A. *et al.* (2018) 'Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells', *Nature Communications*, 9(1), p. 656. Available at: https://doi.org/10.1038/s41467-018-02993-8.

Bauman, J.G.J. *et al.* (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), pp. 485–490. Available at: https://doi.org/10.1016/0014-4827(80)90087-7.

Bernstein, J.A. *et al.* (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), pp. 9697–9702. Available at: https://doi.org/10.1073/pnas.112318199.

Bertrand, E. *et al.* (1998) 'Localization of ASH1 mRNA Particles in Living Yeast', *Molecular Cell*, 2(4), pp. 437–445. Available at: https://doi.org/10.1016/S1097-2765(00)80143-4.

Bohrer, C.H. and Xiao, J. (2020) 'Complex Diffusion in Bacteria', *Advances in experimental medicine and biology*, 1267, pp. 15–43. Available at: https://doi.org/10.1007/978-3-030-46886-6_2.

Bratu, D.P. *et al.* (2003) 'Visualizing the distribution and transport of mRNAs in living cells', *Proceedings of the National Academy of Sciences*, 100(23), pp. 13308–13313. Available at: https://doi.org/10.1073/pnas.2233244100.

Carrocci, T.J. and Hoskins, A.A. (2014) 'Imaging of RNAs in Live Cells with Spectrally Diverse Small Molecule Fluorophores', *The Analyst*, 139(1), p. 10.1039/c3an01550e. Available at: https://doi.org/10.1039/c3an01550e.

Chan, J.H., Lim, S. and Wong, W.F. (2006) 'Antisense Oligonucleotides: From Design to Therapeutic Application', *Clinical and Experimental Pharmacology and Physiology*, 33(5–6), pp. 533–540. Available at: https://doi.org/10.1111/j.1440-1681.2006.04403.x.

Chen, M. *et al.* (2017) 'A molecular beacon-based approach for live-cell imaging of RNA transcripts with minimal target engineering at the single-molecule level', *Scientific Reports*, 7(1), p. 1550. Available at: https://doi.org/10.1038/s41598-017-01740-1.

Chen, W. *et al.* (no date) 'Single mRNA Imaging with Fluorogenic RNA Aptamers and Small-molecule Fluorophores'. Available at: https://doi.org/10.1002/anie.202209813.

Chen, X. *et al.* (2019) 'Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs', *Nature Biotechnology*, 37(11), pp. 1287–1293. Available at: https://doi.org/10.1038/s41587-019-0249-1.

Crawford, R. *et al.* (2013) 'Long-Lived Intracellular Single-Molecule Fluorescence Using Electroporated Molecules', *Biophysical Journal*, 105(11), pp. 2439–2450. Available at: https://doi.org/10.1016/j.bpj.2013.09.057.

Daigle, N. and Ellenberg, J. (2007) 'λN-GFP: an RNA reporter system for live-cell imaging', *Nature Methods*, 4(8), pp. 633–636. Available at: https://doi.org/10.1038/nmeth1065.

Elf, J. and Barkefors, I. (2019) 'Single-Molecule Kinetics in Living Cells', *Annual Review of Biochemistry*, 88(1), pp. 635–659. Available at: https://doi.org/10.1146/annurev-biochem-013118-110801.

Fei, J. and Sharma, C.M. (2018) 'RNA localization in bacteria', *Microbiology spectrum*, 6(5), p. 10.1128/microbiolspec.RWR-0024–2018. Available at: https://doi.org/10.1128/microbiolspec.RWR-0024-2018.

Femino, A.M. *et al.* (1998) 'Visualization of single RNA transcripts in situ', *Science (New York, N.Y.)*, 280(5363), pp. 585–590. Available at: https://doi.org/10.1126/science.280.5363.585.

Forrest, K.M. and Gavis, E.R. (2003) 'Live Imaging of Endogenous RNA Reveals a Diffusion and Entrapment Mechanism for nanos mRNA Localization in Drosophila', *Current Biology*, 13(14), pp. 1159–1168. Available at: https://doi.org/10.1016/S0960-9822(03)00451-2.

Gao, F. *et al.* (2022) 'A Cas6-based RNA tracking platform functioning in a fluorescence-activation mode', *Nucleic Acids Research*, 50(8), p. e46. Available at: https://doi.org/10.1093/nar/gkac014.

George, L. *et al.* (2018) 'Intracellular RNA-tracking methods', *Open Biology*, 8(10), p. 180104. Available at: https://doi.org/10.1098/rsob.180104.

Gijtenbeek, L.A. van *et al.* (2016) 'On the Spatial Organization of mRNA, Plasmids, and Ribosomes in a Bacterial Host Overexpressing Membrane Proteins', *PLOS Genetics*, 12(12), p. e1006523. Available at: https://doi.org/10.1371/journal.pgen.1006523.

Golding, I. and Cox, E.C. (2004) 'RNA dynamics in live *Escherichia coli* cells', *Proceedings of the National Academy of Sciences*, 101(31), pp. 11310–11315. Available at: https://doi.org/10.1073/pnas.0404443101.

Guo, D. *et al.* (2024) 'Liquid-Liquid phase separation in bacteria', *Microbiological Research*, 281, p. 127627. Available at: https://doi.org/10.1016/j.micres.2024.127627.

Gupta, A. *et al.* (2014) 'In vivo kinetics of segregation and polar retention of MS2-GFP-RNA complexes in Escherichia coli', *Biophysical Journal*, 106(9), pp. 1928–1937. Available at: https://doi.org/10.1016/j.bpj.2014.03.035.

Hermann, T. and Westhof, E. (1999) 'Non-Watson-Crick base pairs in RNA-protein recognition', *Chemistry & Biology*, 6(12), pp. R335–R343. Available at: https://doi.org/10.1016/S1074-5521(00)80003-4.

Javer, A. *et al.* (2013) 'Short-time movement of E. coli chromosomal loci depends on coordinate and subcellular localization', *Nature Communications*, 4(1), p. 3003. Available at: https://doi.org/10.1038/ncomms3003.

Jha, R. *et al.* (2015) 'Molecular beacon-based detection and isolation of working-type cardiomyocytes derived from human pluripotent stem cells', *Biomaterials*, 50, pp. 176–185. Available at: https://doi.org/10.1016/j.biomaterials.2015.01.043.

Kang et al. (2011) 'Molecular beacon-based bioimaging of multiple microRNAs during myogenesis - ScienceDirect'. Available at: https://www.sciencedirect.com/science/article/pii/S0142961210014298 (Accessed: 12 August 2024).

Kolmakov, K. *et al.* (2010) 'Red-Emitting Rhodamine Dyes for Fluorescence Microscopy and Nanoscopy', *Chemistry – A European Journal*, 16(1), pp. 158–166. Available at: https://doi.org/10.1002/chem.200902309.

Lakowicz, J.R. (ed.) (2006) *Principles of Fluorescence Spectroscopy*. Boston, MA: Springer US. Available at: https://doi.org/10.1007/978-0-387-46312-4.

Lange, S. *et al.* (2008) 'Simultaneous Transport of Different Localized mRNA Species Revealed by Live-Cell Imaging', *Traffic*, 9(8), pp. 1256–1267. Available at: https://doi.org/10.1111/j.1600-0854.2008.00763.x.

Mao, S. *et al.* (2020) 'Recent Advances in the Molecular Beacon Technology for Live-Cell Single-Molecule Imaging', *iScience*, 23(12), p. 101801. Available at: https://doi.org/10.1016/j.isci.2020.101801.

Medley, C.D. *et al.* (2005) 'Simultaneous Monitoring of the Expression of Multiple Genes Inside of Single Breast Carcinoma Cells', *Analytical Chemistry*, 77(15), pp. 4713–4718. Available at: https://doi.org/10.1021/ac050881y.

Mohanty, B.K. and Kushner, S.R. (2022) 'Regulation of mRNA decay in E. coli', *Critical reviews in biochemistry and molecular biology*, 57(1), pp. 48–72. Available at: https://doi.org/10.1080/10409238.2021.1968784.

Nelles, D.A. *et al.* (2016) 'Programmable RNA tracking in Live Cells with CRISPR/Cas9', *Cell*, 165(2), pp. 488–496. Available at: https://doi.org/10.1016/j.cell.2016.02.054.

Nevo-Dinur, K. *et al.* (2011) 'Translation-Independent Localization of mRNA in E. coli', *Science* [Preprint]. Available at: https://doi.org/10.1126/science.1195691.

Niederauer, C. *et al.* (2023) 'Dual-color DNA-PAINT single-particle tracking enables extended studies of membrane protein interactions', *Nature Communications*, 14(1), p. 4345. Available at: https://doi.org/10.1038/s41467-023-40065-8.

Oomoto, I. *et al.* (2015) 'ECHO-liveFISH: *in vivo* RNA labeling reveals dynamic regulation of nuclear RNA foci in living tissues', *Nucleic Acids Research*, 43(19), pp. e126–e126. Available at: https://doi.org/10.1093/nar/gkv614.

Paige, J.S., Wu, K.Y. and Jaffrey, S.R. (2011) 'RNA Mimics of Green Fluorescent Protein', *Science*, 333(6042), pp. 642–646. Available at: https://doi.org/10.1126/science.1207339.

Park, S.Y., Moon, H.C. and Park, H.Y. (2020) 'Live-cell imaging of single mRNA dynamics using split superfolder green fluorescent proteins with minimal background', *RNA (New York, N.Y.)*, 26(1), pp. 101–109. Available at: https://doi.org/10.1261/rna.067835.118.

Peabody, D.S. (1993) 'The RNA binding site of bacteriophage MS2 coat protein', *The EMBO journal*, 12(2), pp. 595–600. Available at: https://doi.org/10.1002/j.1460-2075.1993.tb05691.x.

Peng, X.-H. *et al.* (2005) 'Real-time Detection of Gene Expression in Cancer Cells Using Molecular Beacon Imaging: New Strategies for Cancer Research', *Cancer Research*, 65(5), pp. 1909–1917. Available at: https://doi.org/10.1158/0008-5472.CAN-04-3196.

Pfeiffer, P. *et al.* (2024) 'Metabolic RNA labeling in non-engineered cells following spontaneous uptake of fluorescent nucleoside phosphate analogues', *Nucleic Acids Research*, p. gkae722. Available at: https://doi.org/10.1093/nar/gkae722.

piedrro - Overview (no date) *GitHub*. Available at: https://github.com/piedrro (Accessed: 17 September 2024).

Plochowietz, A., Crawford, R. and Kapanidis, A.N. (2014) 'Characterization of organic fluorophores for *in vivo* FRET studies based on electroporated molecules', *Phys. Chem. Chem. Phys.*, 16(25), pp. 12688–12694. Available at: https://doi.org/10.1039/C4CP00995A.

Reyes-Lamothe, R. *et al.* (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), pp. 1042–1051. Available at: https://doi.org/10.1093/nar/gkt918.

Santangelo, P. (2006) 'Live-Cell Characterization and Analysis of a Clinical Isolate of Bovine Respiratory Syncytial Virus, Using Molecular Beacons', *Journal of Virology* [Preprint]. Available at: https://doi.org/10.1128/jvi.80.2.682-688.2006.

Sattler, L. and Graumann, P.L. (2021) 'Real-Time Messenger RNA Dynamics in Bacillus subtilis', *Frontiers in Microbiology*, 12. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.760857 (Accessed: 10 October 2023).

Shakoori, A.R. (2017) 'Fluorescence In Situ Hybridization (FISH) and Its Applications', *Chromosome Structure and Aberrations*, pp. 343–367. Available at: https://doi.org/10.1007/978-81-322-3673-3_16.

Steen, P.R. *et al.* (2024) 'The DNA-PAINT palette: a comprehensive performance analysis of fluorescent dyes', *Nature Methods*, 21(9), pp. 1755–1762. Available at: https://doi.org/10.1038/s41592-024-02374-8.

Stringer, C. *et al.* (2021) 'Cellpose: a generalist algorithm for cellular segmentation', *Nature Methods*, 18(1), pp. 100–106. Available at: https://doi.org/10.1038/s41592-020-01018-x.

Sun, N.-H. *et al.* (2020) 'CRISPR-Sunspot: Imaging of endogenous low-abundance RNA at the single-molecule level in live cells', *Theranostics*, 10(24), pp. 10993–11012. Available at: https://doi.org/10.7150/thno.43094.

Sunbul, M. and Jäschke, A. (2013) 'Contact-Mediated Quenching for RNA Imaging in Bacteria with a Fluorophore-Binding Aptamer', *Angewandte Chemie International Edition*, 52(50), pp. 13401–13404. Available at: https://doi.org/10.1002/anie.201306622.

Turner-Bridger, B. *et al.* (2018) 'Single-molecule analysis of endogenous β-actin mRNA trafficking reveals a mechanism for compartmentalized mRNA localization in axons', *Proceedings of the National Academy of Sciences*, 115(41), pp. E9697–E9706. Available at: https://doi.org/10.1073/pnas.1806189115.

Tyagi, S. and Kramer, F.R. (1996) 'Molecular Beacons: Probes that Fluoresce upon Hybridization', *Nature Biotechnology*, 14(3), pp. 303–308. Available at: https://doi.org/10.1038/nbt0396-303.

Vargas, D.Y. *et al.* (2005) 'Mechanism of mRNA transport in the nucleus', *Proceedings of the National Academy of Sciences*, 102(47), pp. 17008–17013. Available at: https://doi.org/10.1073/pnas.0505580102.

Wang, Y., Yang, L.-Z. and Chen, L.-L. (2020) 'Protocol for Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 System', *STAR Protocols*, 1(1), p. 100037. Available at: https://doi.org/10.1016/j.xpro.2020.100037.

Wu, B., Chen, J. and Singer, R.H. (2014) 'Background free imaging of single mRNAs in live cells using split fluorescent proteins', *Scientific Reports*, 4(1), p. 3615. Available at: https://doi.org/10.1038/srep03615.

Wu, P.G. and Brand, L. (1994) 'Resonance Energy Transfer: Methods and Applications', *Analytical Biochemistry*, 218(1), pp. 1–13. Available at: https://doi.org/10.1006/abio.1994.1134.

Yeh, H.-Y. *et al.* (2008) 'Visualizing the dynamics of viral replication in living cells via Tat peptide delivery of nuclease-resistant molecular beacons', *Proceedings of the National Academy of Sciences*, 105(45), pp. 17522–17525. Available at: https://doi.org/10.1073/pnas.0807066105.

Zhao, D. *et al.* (2016) 'Single-molecule detection and tracking of RNA transcripts in living cells using phosphorothioate-optimized 2'-O-methyl RNA molecular beacons', *Biomaterials*, 100, pp. 172–183. Available at: https://doi.org/10.1016/j.biomaterials.2016.05.022.

Zhao, T. *et al.* (2019) 'Super-resolution imaging reveals changes in Escherichia coli SSB localization in response to DNA damage.', *Genes to cells : devoted to molecular & cellular mechanisms*, 24(12), pp. 814–826. Available at: https://doi.org/10.1111/gtc.12729.

Zuker, M. (2003) 'Mfold web server for nucleic acid folding and hybridization prediction', *Nucleic Acids Research*, 31(13), pp. 3406–3415.