Investigating RNA Localized to the Cytoskeleton

A Thesis

Presented to

The Faculty of the Department of Molecular Biology

The Colorado College

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Arts

By

Pedro Tirado Velez

May, 2020

Abstract

Subcellular RNA localization can directly impact the intracellular interactions, morphology, and cellular functions of any given eukaryotic cell. The importance of maintaining proper delivery of mRNA transcripts has previously been studied, with mutations in any component of subcellular localization resulting in abnormal cell developments or functions. Interestingly, mutations in critical cytoskeletal components of microtubules, such as tubulin alpha 1, have been implicated in diseases characterized as “tubulinopathies”, a wide range of brain malfunctions. To gain a deeper understanding of the elements involved in RNA localization and to identify interactions from a multilevel perspective, a combination of techniques were used to identify novel localized RNAs found nearby cytoskeletal structures. To achieve this, we used spatially restricted nucleobase-oxidation to identify novel RNA interactions in close proximity to tubulin, a component of microtubules. In this technique, a localized light reactive fluorophore can be used to produce oxygen radicals with the ability to tag RNA transcripts within a short diffusion radius, thus, allowing for the extraction and identification of RNA and their interactions. As a means for further understanding tubulinopathies and novel RNA interactions within tubulin, fluorophore J646 was localized to tubulin alpha 1 and microtubule-acting cross-linking factor 43 proteins and used to tag nearby RNA. To verify these novel RNA interactions, single-molecule fluorescence *in Situ* hybridization (smFISH), a technique that allows for the study of localized RNA, DNA, or protein, was to be used. Although this technique allows for visualization of *in vivo* RNA localization, it involves the production of expensive fluorescently labeled oligonucleotide probes that may not possess high sensitivity to detect RNA in low abundance. Therefore, to verify novel RNA interactions utilizing microtubules for localization with increased detection, a procedure for generating smFISH probes from readily available oligonucleotides with more than one attached dye was created. Although further optimizations are required before usage, I have successfully attached more than one dye to the unlabeled smFISH probes. By pairing both nucleobase oxidation and smFISH, a greater understanding of novel RNA localization utilizing microtubules can be achieved. Furthermore, by developing a protocol to generate highly sensitive smFISH probes by increasing the attached dye counts from one to several, a greater fluorescence can be achieved allowing for higher sensitivity of even the smallest quantities of localized RNA.

Introduction

Eukaryotes possess several mechanisms for the regulation of gene expression that strongly contribute to the cell’s stability, morphology, development, and function. As described through the central dogma, mRNAs are created from the transcription of DNA while proteins are created from the translation of mRNA. This description, however, does not address the underlying complexities involved in gene expression such as promoters, repressors, activators, transcription initiating factors, and mRNA processing to understand how cell-specific gene expression gives rise to the features that define the function of the cell. Most views of gene expression fail to appreciate how dynamic mRNAs are during gene expression. Not only can mRNAs be found traversing the nuclear envelope to enter the cytoplasm, but mRNAs can also be trafficked to membrane bound organelles such as the endoplasmic reticulum or the mitochondria for localized expression.[1] The mechanisms facilitating RNA transport are critical for proper cellular function as diseases such as spinal muscular atrophy can arise.[1] This points towards RNA localization and the mechanisms that facilitate localization of extreme importance for human health.

RNA localization provides ranging benefits

The localization of RNA, suggested to be an evolutionarily conserved process, has been hypothesized to provide several benefits for the cell.[1,2] First, given that a single RNA molecule

can be translated into many protein-products, it would seemingly be more energetically favorable to rapidly produce high levels of a particular protein at the site where it is required rather than in the cytosol or ER prior to transport throughout the cell. In line with this, ribosomal occupancy studies in live cells suggests that roughly 10-25 ribosomes coat an individual mRNA at a given time, further highlighting the capability of mass protein production.[1] Furthermore, mRNA localization can provide a faster means for responding to stimuli unlike the active transport of proteins to specific cellular compartments.[1,2] For example, during axon directional migration, localized translation of targeted mRNAs is required for proper extension of growth cones in response to external stimuli.[3] According to Chin et al, 2017, it may prove much more favorable to produce the required protein, at a faster rate, on site by using the RNA template as opposed to actively transporting copious amounts of nascent proteins. Secondly, RNA localization may provide a favorable means to regulate the function of noncoding RNA such as microRNAs. For example, in neuronal cells, microRNAs can be localized to specific regions to prevent the translation of mRNA.[4] Third, nascent proteins produced from localized RNA transcripts are also readily more available for post-translational modifications that dictate and alter their roles in localized areas.[1,3,5] For example, RNAs localized and translated on the surface of the ER allow for the nascent protein's proper folding and post-translational modification.[3] Lastly and more importantly, protein production through localized RNA transcripts may also limit incorrect trafficking of proteins to subcellular compartments, which can result in toxic effects.[1] For example, improper trafficking of embryonic polarity regulators can lead to unwanted developments and improper patterning.[1] Overall, RNA localization provides an energetically favorable method for controlling gene expression, thus, allowing for proper protein production at required sites and a faster means for responding to external stimuli.

RNA localization leads to compartmentalization of gene expression

RNA localization, a complex regulatory process that leads to the organizational compartmentalization of protein expression, is mediated by RNA sequences and RNA binding proteins (RBP). In its most basic form, a fully processed mRNA molecule consists of the following: an upstream 5' cap, 5'UTR, a coding region, 3'UTR, and a poly-(A) tail. Within these RNA features exist smaller elements such as *cis*-acting elements that play crucial roles in RNA trafficking and spatial arrangement. *Cis*-acting elements, otherwise known as "zip-codes", consists of nucleotide sequences or motifs within the 3'UTR, 5'UTR, or coding region of mRNA that are recognized by *trans*-acting elements or RBPs for compartmentalization.[6] These zip-codes, varying from a few nucleotides to over 1000kb long, facilitate transport of target mRNA by providing binding sites for specific RBPs to build competent RNA-RBP complexes (RNP) that are recognized and trafficked by molecular motors.[6] The identification of these zip-code sequences has been difficult as many factors, such as intrinsic RNA structures, may play a role in initiating RNA-Protein interactions necessary for localization.[6,7] Previously, as a means for identifying RNA sequences that may act as zip-codes, cellular locations such as extended axons/dendrites have been segmented then fractionated for localized RNA.[6,7,8] In these studies, fractionated RNA samples are identified through RNA sequencing and an enrichment of sequences similarity is searched for.[6,7,8] Then, these identified sequences are tested *in vivo* by addition or deletion of the identified sequence in an RNA to test if its removal or addition affects that RNA's localization. This strategy has been used to validate if that deleted/added sequence can function as a "zip-code".[8] These studies illuminated that the presence of a zip-code is sufficient to cause localization of injected mRNA, assuming that *trans*-acting RBPs involved in the localization are expressed. For example, injection of capped protein-free transcripts into the basal cytoplasm of *Drosophila* embryos caused localization of the RNAs to the apical cytoplasm, signifying zip-codes within these RNA to be sufficient for localization.[9] Although these studies highlighted the importance of *cis*-acting elements, *trans*-acting factors also play a

significant role in mediating RNA localization. For example, mutations in an RBP involved in the localization of glutelin mRNA in developing rice endosperm significantly altered correct localization.[10] All together, these highlight the importance of the interaction between *cis*-acting elements and trans-acting elements in mediating RNA localization.

Along with *cis*- and *trans-acting* elements, cytoskeletal fibers provide the cell with a means for spatially organizing RNAs.[1,7] The cytoskeleton, which is involved in cellular processes ranging from cell contraction to cell stability, is composed of three different types: actin filaments, intermediate filaments, and microtubules.[11] Microfilaments, made up of identical actin proteins fibers, help with various cellular functions such as muscle contractions or cell morphology.[11] In comparison, intermediate filaments, which vary in subunit composition, primary provide mechanical function for the cell as well as strength and support to tubulin structures. Similarly, microtubules, cylindrical tubes composed of tubulin heterodimers termed alpha- and beta- tubulin, are involved in intercellular transport, neuronal division, and cellular morphology. Microtubules function as a dynamic skeleton for the cell and are made up of two ends: a quick polymerizing plus-end and an anchored minus-end.[9,10,11] Tubulin alpha 1a, a component of the tubulin heterodimer, has been suggested to be primarily expressed in differentiated neuronal cells. Macf43 has been shown to associate at the plus-ends of microtubules. Altogether, cytoskeletal components provide a highway for transport of various cargo such as the RNA-RBP complex involved in RNA localization.

To facilitate localization, motor proteins, such as kinesin, mediate RNA transport through ATP mechanical action along cytoskeletal structures.[11] Each of these proteins utilize one of two cytoskeletal components, microfilaments or microtubules, to facilitate the transport of cargo such as RNA to specific compartments.[11] Kinesins are motor proteins that primarily travel using polymerizing microtubules and all contain the following: a highly conserved globular motor domain containing a microtubule/ATP binding site and a diverse tail domain involved in cargo recognition and interaction.[11,12] These motor domains are found in diverse regions of the protein with its location within the protein defining whether the kinesin travels towards the plus end of microtubules or towards the minus end of microtubules.[11] Given the complexities of motor protein, understanding how each is involved in RNA localization is understudied and may present further implications in RNA mislocalization resulting in disease.

The specific mechanism of action for RNA localization is dependent on *trans*-acting factors that recognize *cis*-features to mediate the necessary interactions needed for localization. Currently, two mechanisms for RNA localization mediated by the cytoskeleton, each utilizing *cis*-acting and *trans*-acting elements, have been characterized. The first of two mechanisms, called diffusion coupled local entrapment, involves the diffusion of mRNA transcripts followed by localized anchoring to a specific compartment.[13] The second and most predominant localization mechanism, directed transport, involves localization of RNP through polarized cytoskeletal structures previously addressed. Travel is facilitated by cytoskeletal motor proteins, which often utilize *trans*-acting elements such as adaptor proteins to mediate interactions between motor proteins and mRNAs. For example, She2p and She3p act as adaptor proteins and link *ASH1 m*RNA to myosin motor protein, Myo4p, for transport in *Saccharomyces cerevisiae*.[14] These mechanisms can further be subdivided into two models of localization, a translationally dependent and translationally independent model. In the translationally driven model, mRNA transport is thought to occur from a translational dependent process involving the recognition of signal peptides. In this model, as the ribosome-mRNA complex produces a nascent protein, specific peptide sequences within the protein are recognized by the signal recognition particle (SRP).[14, 15] The ribonuclear complex formed is translationally repressed and upon deliver, becomes translationally competent. A second model involves recognition of “zip-code” sequences within the mRNA transcript by RBPs, which links the mRNA to motor proteins for further localization.[15] This model does not require signal peptides from within the nascent protein but uses mRNA nucleotides exposed from intrinsic RNA structures to interact with RNA binding

proteins. Often, these mRNAs are translationally repressed until activated in specific compartments.[15] Within the mechanisms addressed, there are further underlying complexities involved to ensure proper RNA localization, thus, additional studies must be undertaken to gain a greater understanding.

Malfunction in RNA localization can result in disease

Due to the plethora of mechanisms involved in RNA localization, mutations found at several points during the process can have massive downstream effects that lead to disease. Improper RNA localization has been implicated in several diseases such as Alzheimer's or fragile X-syndrome and has been shown to play a major role in several neurodegenerative diseases.[16] For example, mutations found in myosin VIIa, a cytoskeletal motor protein utilizing microfilaments for travel, have been implicated in a neuronal disease characterized by visual impairment and deafness called Usher syndrome I.[11] Structural studies of this altered motor protein have identified the region to be involved in interactions with the central domain of adaptor protein Sans, which plays a role in hormonin mRNA targeting to the correct compartments of the inner ear.[11] Because the cytoskeleton plays a significant role in facilitating localization by providing a highway for motor proteins to travel, mutations in the cytoskeletal makeup may significantly disrupt proper interactions, which may lead to mislocalization. Further investigations may therefore implicate cytoskeletal subunit mutations in novel or known diseases.

Interactions from components that do not play a crucial role in RNA localization can also result in disease. For example, in neurological and neuromuscular disorders, there is often a presence of unstable microsatellites or repetitive nucleotide repeats in protein-coding genes that give rise to mRNAs with the ability to interfere with cellular components.[1] These microsatellites can lead to disease by introducing the following loss or gain of function scenarios: disruption of normal localization pathways through sequestering of *trans*-acting elements necessary for transport or microsatellite mRNAs translation into toxic dipeptide repeat proteins that may go on to form aggregates with RBPs necessary for the localization of other mRNAs.[1] Interestingly, recent evidence has also implicated RNA localization in cancer. In a study involving 26 cancer types ranging from thyroid to adrenocortical carcinomas, at least one cancer sample had 281 RBPs found to be enriched for mutations, possibly indicating irregular RNA localization pathways.[17] Given the extensive role RNA localization has in gene expression and cell function, this mechanism may yet be identified to play a larger and more significant role in known diseases.

It is not surprising that mutations in mechanisms directly involved in RNA localization, such as mutations in cytoskeleton components, can result in significant abnormal cell function or development. This is illustrated by mutations within the tubulin family proteins that give rise to tubulinopathies, an autosomal dominant disorder characterized by a wide range of brain malfunctions such as developmental delay and epilepsy. For example, this can be seen by mutations within the tubulin alpha 1 protein, which often results in *TUBA1A*-associated tubulinopathy.[18] Since *TUBA1A* accounts for a large portion of expressed α-tubulin mRNA in cells, with up to 95% of α-tubulin in embryonic neuronal cells, mutations in this gene can significantly alter numerous tubulin/microtubule functions.[19] Although *TUBA1A* mutations are highly associated with these neuronal defects, they are not unique to the tubulin family and can arise from other mutations. For example, mutations in microtubule actin crosslinking factor 43 (*MACF43)*, a gene encoding for a protein involved in facilitating actin-microtubule interactions at the plus end of microtubules, can also give rise to similar brain malfunctions.[20] Because neuronal cells contain highly polarized cellular locations and display a greater/clearer impact on phenotypic expression with altered components, many mutations in cytoskeletal subunits, and thus RNA localization, have been linked to several neuronal diseases. However, the

cytoskeleton plays a crucial role in the direct transport of mRNA in all cells, therefore, future studies in different cell types may reveal a wide range of cellular diseases resulting from mutations in cytoskeletal subunits and thus, mutations in RNA localization.

Identifying RNAs that interact with the cytoskeleton

For our studies, we aim to investigate the role of the cytoskeleton in mediating RNA localization to address how mutations within this system lead to diseases like tubulinopathies. We seek to identify which RNA cytoskeleton interactions are needed for proper localization to further understand the biological significance of cytoskeletal structures in RNA localization. In order to do this, we have gathered possible RNAs that interact with one of the three cytoskeletal structures, microtubules, as a means for identifying a link between RNA mislocalization and tubulinopathy. We aim to understand the role of the cytoskeleton in RNA localization using two techniques: nucleobase oxidation and single molecule fluorescence *in situ* hybridization (smFISH).

Although a variety of techniques could alternatively be used, nucleobase oxidation and smFISH are high resolution techniques that can identify and visualize RNA-cytoskeleton interactions *in vivo.* Previously, cell fractionation, involving the fractionation, extraction, and sequencing of transcripts from a pool of cells, was the basis for many RNA localization studies. Although fractionation and similar techniques have furthered investigations, they were not capable of providing information at a high specificity, extracted large pools of RNA found in the vicinity, and explored results pertaining to an average count rather than on a single-cell basis.[21] Alternatively, an improved technique called nucleobase oxidation utilizes an excited fluorophore such as dibromofluorescein (DBF) to identify novel localized RNA within living cells.[22] When excited though blue light, fluorophores such as DBF are able to use the energy to produce highly reactive oxygen species that have a short half-life (~3 μs), and short diffusion distance (~268nm). Therefore, upon excitation, nearby molecules close to the generated reactive oxygen species will be modified by oxidation. More specifically, the excited fluorophore utilizes the energy to bring a ground state triplet oxygen to a singlet state radical, which will go on to tag the guanosine residue of nearby RNAs, producing 8-Oxoguanine (Fig. 1B).[22] Since 8-Oxoguanine (8-oxoG) is highly susceptible to nucleophilic attack, its chemistry can be exploited to purify RNAs that have been modified away from the rest of the RNA pool (Fig.1A).[22] This will allow us to gain more accurate results and identify possible novel RNA interactions.

The HaloTag protein, a bacterial derived peptide, has the ability to tightly bind a diverse range of fluorophores modified with a HaloTag ligand. When the HaloTag is fused to a protein with a known localization, the HaloTag is not only localized to the site of interest, but it also causes localization of the fluorophore modified with a HaloTag ligand. Therefore, nucleobase oxidation can be used in conjunction to the HaloTag interaction system to tag nearby and localized RNAs with 8-oxoG. DBF or analogs of DBF with improved properties such as rapid cell labeling or high specificity can be utilized to create Haloligand-fluorophore chimeric proteins.

After fluorophore-mediated in-cell labeling of localized RNAs, click chemistry can be used *in vitro* to attach biotin to the 8-Oxoguanosine residue (Fig. 1B), which can then be used to facilitate the purification of tagged RNA molecules from the rest of the lysate with the help of streptavidin (Fig. 1A). Using a Halo-Tag fusion to either Tubulin Alpha 1a or microtubule-actin crosslinking factor 43, will allow for the purification of RNA associated with the cytoskeleton. HeLa cell lines, stably expressing each of these fusion proteins, were treated with JF646, an analog of DBF, to initiate its localization by the Halo domain to either tubulin (TUBA1A) or the plus end of microtubules (MACF43). This facilitated the tagging and purification of RNA molecules near these cellular structures *in vivo*. We have successfully extracted RNA from the cell lysate as well as performed click chemistry to attach biotin to tagged RNA and await further purification of tagged RNA for RNA sequencing.

To validate our pending nucleobase oxidation results, we will use smFISH to view the localization of interesting RNAs identified. Single-molecule fluorescent *in Situ* hybridization (smFISH) allows for observing the localization of RNA within a single cell.[21,23] However, smFISH relies on cell fixation and can only analyze a few RNAs at a time, thus, limiting us to visualize a snapshot of the localization of a few RNAs. Furthermore, this technique involves the production of expensive fluorescently labeled oligonucleotide probes that are effective in detecting abundant mRNAs but have not been demonstrated to work with mRNAs that exist at lower quantities.[24] In order to use smFISH for subsequent validation of RNAs identified by nucleobase oxidation, we aim to increase the range of detection and reduce the cost to visualize many RNAs to gain a deeper understanding of the role of the cytoskeleton in RNA spatial arrangement.

To create a method for developing smFISH probes with a higher detection sensitivity, our study attempted to 3' end label oligonucleotide probes with dye-conjugated nucleotides using terminal deoxynucleotidyl transferase (TdT).[24] To increase the sensitivity of our smFISH probes, we used various conditions to increase the number of attached dyes to the 3' end (Fig. 4C). By attaching more fluorescent dyes to the probe, we assumed that this would yield more sensitivity than a probe with a single fluorophore label. We were successful in attaching more than one dye to our probes, however, further purifications from unincorporated dye are needed before these probes are used in smFISH.

By utilizing a combination of techniques to address how the cytoskeleton plays a role in RNA localization, a deeper understanding of underlying mechanisms can be obtained. Nucleobase oxidation by the HaloTag-TUBA1A or HaloTag-MACF43 allows for the identification RNAs whose localizations are dependent on TUBA1A or MACF43. By identifying this subset of RNAs, an analysis of sequence similarity may lead to the identification of potential zip-codes, giving rise to additional studies focusing on *trans*-acting factors involved in this process. Not only would these studies deepen our understanding of the biological significance of elements involved, but they may lead to the identification of problematic phenotypes resulting from mis-localization and may provide future targets for therapeutic use.

Results

Generating Halo tagged MACF43 and TUBA1A

To expand upon the biological role of the cytoskeleton in mediating RNA localization, we aimed to identify RNAs that interact with the cytoskeleton using spatially restricted nucleobase oxidation. In order to perform nucleobase oxidation in our experiments, a Halo-tagged protein of interest must be generated by fusing a HaloTag, which binds specifically to certain fluorophores, to the N-terminus side of a protein with a known localization (Fig. 1A). We chose to create chimeric proteins of alpha tubulin 1a and microtubule-actin cross-linking factor 43 to investigate RNA interactions within the microtubule components of the cytoskeleton. To tag and express these proteins *in vivo* for nucleobase oxidation, we utilized plasmids that express antibiotic resistance and the HaloTag gene. For proper protein folding, we decided to fuse *MACF43* and *TUBA1A* coding sequences at the 3' end of the Halo-tag gene.

After each construct was created, we sought to verify the proper ligation of *MACF43* and *TUBA1A* in each plasmid. We conducted a restriction digest on the purified plasmids with XhoI and pShaI, which was expected to produce four bands (345bp, 664bp, 722bp, 7.69kp) for the Halo-TUBA1A plasmid and four bands (32bp, 265bp, 345bp, and 7.69kp) for the Halo-MACF43 plasmid. Although a DNA agarose gel of the restriction digest shown in Figure 2 displayed the hypothesized number of bands, the expected 7.69kb sized bands for both digest shown at the top lanes had larger molecular weights and could not be identified correctly as the molecular ladder bands were difficult to distinguish at larger weights (* in Fig. 2). However, the other bands were of proper size and each digest produced the expected number of bands therefore

suggesting that proper ligation of *MACF43* and *TUBA1A* into the plasmid had occurred. As a further validation, sanger sequencing was performed and showed that plasmids were capable of expressing Halo tagged MACF43 and TUBA1A (not shown).

Cell imaging reveals proper cell transformation

Once the plasmids expressing either Halo-*MACF43* or Halo-*TUBA1A* under a Tet-on inducible system were validated, each plasmid was concentrated and transfected into HeLa cells. To maximize the yield of tagged mRNA, we grew cells to an 85-95% confluence. The fluorophore JF646, a DBF derivative that contains a portion of the HaloTag ligand that allows it to bind to the HaloTag, was added to cells for 30 minutes to allow for the creation of localized oxygen radicals in downstream steps (Fig. 1A). Upon excitation with green light (~560nm), JF646 dyes introduced into cells expressing Halo-MACF43 or Halo-TUBA1A were seen interspersed throughout the cytoplasmic regions (Fig. 3). This result was expected as both proteins are localized to cytoskeletal structures throughout the cytoplasm. To verify that cells could only retain dye in the presence of the HaloTag fusion protein, HeLa cells were grown in media lacking doxycycline (- dox), therefore, they did not express the HaloTag fusion proteins (not shown). When JF646 (dye) was introduced and cells imaged, it became evident that the dye was not retained through other means after the wash step as no excitation was obtained in the dye channel (~646nm). Furthermore, when negative controls, consisting of non-transfected HeLa cells incubated with the fluorophore, were subjected to the same excitation wavelength, no fluorescence was observed (not shown). This control demonstrated that only the cells that expressed Halo tagged proteins have the ability to retain the fluorophore after washing the cells, thus, the fluorescence observed in cells expressing the Halo tagged proteins was not a result of residual dye. The remaining portion of the cells were then briefly exposed to green light to generate reactive oxygen species that will react with nearby RNA to modify guanine residues into 8-Oxoguanine from within a limited range (~268nm). Treated cells were washed and total RNA isolated. Next, we performed click chemistry on our RNA samples to attach biotin. Had time permitted, the next step would been to purify the RNA samples by running modified RNAs on a streptavidin column for chromatography. Afterwards, we would have sequenced purified samples through RNA-seq to determine the identity of these RNA localized around cytoskeletal structures.

Denaturing Gel suggests inefficient attachment of more than one nucleotide to unlabeled smFISH probes

To obtain a means for validating RNAs possibly identified to interact with Halo tagged MACF43 and TUBA1A, we will need observe their localization *in vivo*. Several techniques, such as single-molecule Fluorescence *in situ* hybridization (smFISH), can be used to achieve this. To prepare to use smFISH to validate our results from nucleobase oxidation and generate probes to RNAs identified, we decided to optimize an in-house smFISH probe generation protocol by creating smFISH probes targeting the *luciferase* mRNA. Commercially available smFISH oligonucleotide probes are a set of short oligos designed to bind over a large section of a complementary sequence of a target RNA. Each oligo within this set is labelled with a single fluorophore at the 3' end, which limits the sensitivity of the probe. Secondly, commercially available probes are often expensive and not financially practical for usage to identify the localization of many different RNAs, a consideration for using smFISH to validate our nucleobase oxidation samples. To increase sensitivity and lower cost, I aimed to create a set of oligonucleotide probes containing more than one attached fluorescent dye by doing the following: attaching the dye to modified nucleotides and using terminal deoxynucleotidyl transferase (TdT) to catalyze the addition of more than one of these dye-conjugated nucleotides to the 3' end of oligonucleotide probes (Fig. 4B). In detail, Amino-11-deoxyUTP (dUTP) and

Amino-11-dideoxyUTP (ddUTP) were put through an NHS-ester reaction with a dye, Quasar 570 Succinimidyl Ester. These dye-conjugated nucleotides were then purified through ethanol precipitation and used as reagents for a Terminal Deoxynucleotidyl Transferase (TdT) reaction, which facilitated the addition of fluorescently labelled dUTP or ddUTP onto the 3' end of a set of unlabeled oligonucleotide probes that target the entire sequence of the *luciferase* mRNA. In this reaction, dUTP would theoretically be added continuously to the probe until a chain terminating ddUTP is added. ddUTP would terminate polymerization by TdT as it lacks the free 3' hydroxyl group necessary for facilitating elongation in this process. Altering the ratio of dUTP to ddUTP used should therefore change the number of dUTPs added by increasing or decreasing the likelihood of TdT attaching either dUTP or ddUTP. A ratio of 2:1 dye-dUTP to dye-ddUTP was first tried when labeling smFISH probes. This ratio was hypothesized to produce probes with dye attachments ranging from +1 to +3 dye-nucleotide and would statistically favor attachment of more than one given the larger concentration of non-terminating dye-nucleotides (Figure 4b). To visualize attachment, probes were purified away from unincorporated dye and run on a 15% Polyacrylamide-8M Urea denaturing gel, stained with Gel-green to visualize all nucleic acids, and imaged at wavelengths exciting Gel-green (~480nm, green) and Succinimidyl Ester (~570nm, red), the dye used to label oligos. Bands indicating attachment of >+1 dye-nucleotides to unlabeled probes were not seen as only the two green-fluorescent bands were observed as opposed to red bands. The higher molecular weight green-fluorescence band suggested that attachment of one nucleotide occurred (Fig. 5). Unexpectedly, faint intermediate bands were seen emanating from the brightest band, possibly indicating denaturation or improper creation of the oligo probe set. Overall, these findings suggest that dye-conjugate dUTP:ddUTP ratio of 2:1 was not effective for labelling oligos. It remained possible, however, that with a 2:1 ratio, attachment of chain terminating dye-ddUTP was favored, thus leading to the observed results.

To see if we could improve labelling, we tried adjusting reaction conditions as well as using different ratios of dye-conjugated dUTP to ddUTP. From all the attempted optimizations, only pH seemed to improve results by preventing the formation of scattered intermediates seen in Figure 5. With a higher dUTP ratio, however, we observed the similar inefficient labelling of the smFISH probes with one dye attachment (Fig. 6 lanes 2-4). For both 9:1 and 5:1 ratios, the +1 attachment was observed (low molecular weight red band) and there were faint red-fluorescent high molecular weight bands that suggested that more than one dye-nucleotide attachment was occurring (lane 4 and 5). However, the intensity of unlabeled probes (* in Fig. 6; lowest MW band) in samples that underwent the reactions for attachment were similar to the unlabeled probe control, which suggested inefficient addition of dye-conjugated nucleotides through the TdT reaction. Altogether, these results suggested inefficiencies in either the conjugation of the fluorophore to dUTP and ddUTP or attachment of these dye-nucleotides to unlabeled oligonucleotide probes.

Addition of dye after TdT reaction improved dye attachment efficiency

Based on our reactions above, we predicted that the dye was inhibiting dUTP attachment by TdT. This is further highlighted by Figure 4a, which shows that Quasar 570 (dye) is a bulky modification to add onto a nucleotide base and it creates a molecule with a long carbon chain consisting mainly of single bonds. These single bonds allow the carbon chain to freely rotate and may lead to steric hindrance, preventing its use as TdT substrate. To test if the dye was negatively affecting attachment, reactions with only Amino-11-dUTP or Amino-11-ddUTP were used as reagents for the TdT reaction. TdT reaction products were purified through ethanol precipitation and run on a 15% PA – 8M Urea denaturing gel. As seen in Figure 7, band intensity of unlabeled probe (lowest molecular weight green band) was reduced in both TdT reactions using only Amino-11-ddUTP (lane 2) and Amino-11-dUTP (lane 5) when compared to the unlabeled probe control (lane 1). This result suggests efficient attachment of nucleotides by

TdT, given that all three were loaded at the same concentration of DNA and the unlabeled probe bands in the TdT reactions were less abundant when compared to the control. Furthermore, products from the reaction involving only dUTP suggest a variability of +1 to +5 nucleotide attachments (lane 5).

Next, we conjugated the fluorophore (Quasar 570) to purified probes from each TdT reaction through an NHS-ester labelling reaction. As shown in Figure 7, a small amount of ddUTP-attached probe not conjugated to the dye was seen (lane 3 and 4, higher molecular weight green bands) and little to no dUTP-attached probe remained (lane 6 and 7, higher molecular weight green bands) when compared to controls (lane 2 and 5 respectively). This observation demonstrates that a significant portion of the probes labeled with amino-11-nucleotides had undergone attachment to the dye. Overall, adding amino-11-nucleotides to probes using TdT followed probe purification and fluorophore conjugation alleviated low dye attachment efficiencies. Additionally, using only either ddUTP or dUTP improved attachment efficiency. Our protocol efficiently generated dye-conjugated probes with either one attached dye or multi-dye attachments.

smFISH with homemade probes reveals excess dye present

Next, we tested our homemade probes against commercially available probes in HeLa cell lines expressing *luciferase* mRNA. In these cell lines, transfected constructs expressed both *luciferase* and *GFP* genes under the control of an inducible Tet-on expression system. We exposed transfected cells to doxycycline to induce expression of GFP and the *luciferase* mRNA from the transfected plasmid. GFP acted as a reporter gene to identify which cells harbored the transfected constructs and allowed us to identify cells where the *luciferase* mRNA was induced.

We first performed smFISH using the following probes against luciferase: Stellaris commercially available probes and dye-conjugated dUTP probes. When smFISH was conducted on cells with commercially available probes, fluorescence was observed but at a low intensity (Fig. 8, column 1). Defined puncta were visible in the red channel for commercial Stellaris probes, but these puncta were of less intensity and difficult to spot (Fig.8, Merge). As expected, GFP was not present in cells that were not exposed to doxycycline (Figure 8, column 2), which indicates no induction of the target RNA. However, when uninduced cells treated with smFISH homemade dye-conjugated dUTP probes were imaged for Quasar 570 (dye), we detected a large amount of fluorescence present within the cell (Fig. 8, Dye). This finding was unexpected as uninduced cells did not express luciferase mRNA, yet our results suggested otherwise. In cells treated with doxycycline, GFP was present indicating proper expression through the Tet-on system (Fig. 8, lane 3). However, a similar dye fluorescence intensity as seen for uninduced cells was observed (Fig. 8, Merge). Given that our reporter gene, *GFP*, was not induced in all cells that display strong red fluorescence, this fluorescence observed was likely due to nonspecific background noise.

The large background noise we observed with the homemade probes could be caused by free dye that was carried over from the ethanol precipitation used to purify probes after dye labelling. To test this, homemade probes were further purified through size-exclusion chromatography to remove free dye then used to conduct smFISH on transfected cells exposed to doxycycline. Though GFP was visible at a lower intensity when compared to the previous experiments in Figure 6, we observed less intense red fluorescence from the further purified smFISH probes, indicating that there was less unincorporated dye (Fig. 8). Although homemade probes displayed stronger fluorescent signal when compared to commercially available probes, the fluorescence did not seem to accurately represent RNA localization as well as the commercial probes as there were fewer discernable puncta. Unfortunately, untreated – DOX samples with smFISH probes that had extra purification were not generated for comparison. Had time permitted, the next step would have involved using greater amounts of washes and

purification methods on homemade probes in an attempt to remove as much unincorporated dye and unbound probe as possible.

Discussion

RNA localization remains a multilayered system involving many components to accomplish proper trafficking as a means for regulating RNA expression. Given the wide-ranging interactions involved, many mutations within the participating molecules may result in abnormal cellular functions. For example, several mutations in the building blocks of cytoskeleton have been implicated in altered RNA localization, thus, leading to disease. These malfunctions highlight the importance of the cytoskeletal components, such as microtubules or motor proteins, to mediate the direct transport of RNA to its specified cellular localization. Most strikingly, this is well-illustrated by prevalent mutations in *TUBA1A,* which results in *TUBA1A*-associated "tubulinopathy". Given the large presence of α-tubulin mRNA in cells, mutations in this gene can significantly alter numerous functions.[19] To further investigate what RNAs are localized to the microtubules and thus result in disease when lost, our group sought to explore RNA found in near proximity to tubulin structures. We used nucleobase oxidation with HaloTag fused to TUBA1A or MACF43 as a means for localizing the fluorophores needed to produce oxygen radicals. By spatially restricting fluorophores to the cytoskeleton, we were able to achieve tagging of nearby RNAs. Furthermore, we aimed to verify the localization of identified RNAs through smFISH. To address the likelihood of the identified RNAs being expressed at low concentrations and test for multiple RNAs, we developed a robust protocol to create smFISH oligoprobes with more than one attached dye for increased sensitivity. Although the methods explored through our study showed great promise, there are more optimizations necessary to ensure each assay functions accurately to examine RNAs localized to the cytoskeleton.

Spatially restricted nucleobase oxidation was a technique that allowed us to isolate RNAs localized to the cytoskeleton. Through this procedure, tagging of nearby transcripts may be achieved using proteins with known localizations. We looked at the RNA interactions with TUBA1A and MACF43 as these proteins are important in the formation of tubulin heterodimers and microtubule-actin interactions respectively. An inducible DNA construct that expressed the HaloTag properly fused to the 5' end of *TUBA1A* and *MACF43* gene allowed us to produce proteins with the ability to retain reactive fluorophores, thus, allowing us to tag nearby RNA molecules with 8-Oxoguanine. Though we were not able to isolate 8-oxoG tagged mRNA for further sequencing and verification, due to time constraints, we were able to extract RNA from cells after *in vivo* nucleobase oxidation.

More importantly, steps were taken to ensure minimal presence of unincorporated fluorophore for the tagging of 8-Oxoguanine by oxygen radicals and the correct localization of the dye. By utilizing multiple negative controls such as HeLa cells grown in media lacking doxycycline and therefore not expressing the HaloTag protein, we were able to verify that cells could only retain the fluorophore in the presence of the HaloTag fusion protein as no excitation was obtained in the dye channel (~646nm). In contrast, HeLa cells exposed to doxycycline and therefore expressing the Halotag protein were able to retain the fluorophore after wash steps, further suggesting that DNA constructs were correctly made (Fig. 3). When cells exposed to doxycycline were imaged, puncta were seen at large quantities around the cytoplasm, which was expected as microtubules are spread out throughout the cell. Altogether, this suggested that the transfected plasmid was only capable of producing enough HaloTag fusion protein for dye localization when activated through the Tet-on system.

After the proper precautions were taken to ensure that only nearby RNA to tubulin structures were being tagged with 8-Oxoguanine, nucleobase oxidation was conducted on each sample. We were able to extract RNA produced by the cells and attach biotin onto tagged RNAs through click-chemistry for feasible downstream extraction. Had time allowed, RNAs tagged with

biotin would have been purified through a streptavidin column followed by RNA sequencing to determine their identify. These following steps may lead to future studies involving investigations on the role/function of identified RNAs and any possible implications in disease. Additionally, mutant versions of *TUBA1A* could be fused to the HaloTag domain to investigate RNA interactions that are present or absent as a result of altered interactions, which may further identify what mRNAs are implicated in tubulinopathies through cross-comparison. Furthermore, identified RNA would allow for future studies involving verification of localization through smFISH and sequence similarity studies to pinpoint possible novel zip-code sequences.

Optimizing smFISH probe generation to increase the number of attached dyes

To investigate whether these identified RNAs were either involved in interactions with microtubules or if they were false positives due to the procedure, we wanted to utilize smFISH as a means for tracking localization of identified RNA near TUBA1A and MACF43 structures. However, to address the possibility of low detection for some identified RNAs with commercially available probes and to generate a procedure to make different sets of probes to validate multiple RNAs, we needed to do the following: 1) optimize oligo probe generation and 2) make probes containing more than one attached fluorophore. By increasing the fluorophores attached, we would theoretically obtain oligo probes with higher sensitivity and detection. In the first attempts to label smFISH probes targeting the *luciferase* mRNA, single dye-conjugated nucleotides were not the preferred substrate in the TdT reactions, thus, labelling of oligonucleotide occurred at a relatively small efficiency (Fig. 6). Originally, a higher ratio of labelled dUTP compared to labelled ddUTP (*i.e*. two-fold difference) was hypothesized to favor attachment of nonterminating dUTP twice as much as opposed to elongation terminator ddUTP. However, this was quickly disproven as a 2 to 1 (dUTP to ddUTP) ratio resulted in oligos with only one dye-conjugated nucleotide attachment (Fig. 5). Optimizations such as adjusting the pH of the reaction, increasing the reaction time, or increasing nucleotide ratios were undertaken, with only the adjustment of pH providing the highest impact by reducing abnormal intermediates (Fig. 6). Overall, these adjustments did little to improve dye attachment to our oligos and suggested that the TdT enzyme could not utilize multiple dye-conjugated nucleotides as substrates. Given that Quasar 570 (dye) is a bulky modification that creates a molecule with a long carbon chain consisting mainly of single bonds, there is a possibility of steric hindrance resulting from the freely rotating long carbon chain (Fig. 4A), therefore preventing its use as TdT substrate. For example, the long chain of the dye may interact with the TdT polymerase in a way that prematurely knocks the polymerase off before addition of the dye-conjugated dUTP and subsequent labelled nucleotides. However, it is also possible that the dye-conjugated nucleotide cannot fit into the NTP pore or DNA exit channel of TdT, thus, resulting in attachment of only one dye-nucleotide.

To counteract the inability of TdT to incorporate dye-conjugated nucleotides, the procedure was adjusted to use Amino-11-nucleotides as reagent for probe elongation. These Amino-11-nucleotides are modified nucleotides that contain an attached primary amine on a long linker that can undergo the NHS-ester reaction. By integrating amino-11-nucleotides, we can elongate probes with the modified nucleotides first and then conjugate the dye through an NHS-ester reaction after the TdT reaction is complete. This alteration in our procedure resulted in effective elongations. We observed that probes not only had addition of one amino-11-nucleotide but reactions with only amino-11-dUTP demonstrated multiple additions (up to +5). Our studies show that TdT end labelling using amino-11-nucleotides prior to dye addition is an effective method for generating labelled oligos for smFISH and a useful method to add multiple sites for dye conjugation.

smFISH probes with increased dye attachments display high background noise

After conjugation of the fluorophore to the elongated probes, smFISH was conducted on HeLa cells expressing the *luciferase* mRNA to see if elongated probes were viable for smFISH. Our commercially available Stellaris probes showed faint puncta, which suggested that *luciferase* mRNA was expressed but at a low concentration (Fig. 8). Although foci were clearly visible when conducting smFISH with our homemade 3'end labeled probes, there was a large amount of background noise that was not seen with the Stellaris probes. It became evident that the purification steps after the NHS ester reaction were not enough to remove excess dye from the labelled probes. To fix this issue, homemade probes were subjected to a size-exclusion column chromatography to remove any free dye. Though this helped to alleviate some of the background noise, it did not sufficiently reduce the background noise seen in smFISH. One possible reason why our purified probes still displayed high levels of background fluorescence was that during the purification, the weight of the dyes attached to the probes were not taken into consideration when choosing the resin for the size-exclusion chromatography. This mistake can result in copurification of some free-flowing dye and the labeled probes. Collectively, these extra purification steps resulted in visible foci during smFISH, which signify that labelling is properly working but that improved purification steps are required.

Overall, our lab was successful in attaching more than one dye to unlabeled oligonucleotide probes for smFISH. Creating our own probes with multiple dye attachments presents several potential benefits such as the ability to counteract the lesser detection/sensitivity seen from commercially available probes. Although further purifications steps are needed to reduce background noise, puncta of *luciferase* expressed at low levels were visible. This can help in visualizing localization of RNA identified from nucleobase oxidation that are expressed at low concentrations. Also, dyes that absorb at different wavelengths could be used, which would allow for investigating multiple mRNA simultaneously. This could prove beneficial in our mRNA localization studies as it would allow for us to investigate how multiple mRNA interact with the cytoskeleton. Furthermore, we could also gain a spatial understanding of how the presence of each RNA influences the interactions necessary for its localization. Although this technique may prove useful, especially if multiple targets can be examined at once with greater sensitivity, it has the downside of only working after the cells have been fixed, thus allowing for snapshot views rather than dynamic information about mRNA localization.

Use of nucleobase oxidation and smFISH for RNA localization

By utilizing both techniques described above, we can gain a better understanding of how the cytoskeleton interacts with RNA and therefore begin to fill in the gaps of the multilayered RNA localization system. Given its complex nature, addressing portions of the underlying mechanisms can provide clues to how the cytoskeleton, in our case microtubules, interact with RNA as well as how the lack of these interactions may lead to problematic phenotypes. Although improved purification steps will be required, these techniques can also be applied to other cytoskeletal components as a means for identifying localized RNA. From this, further studies can be conducted on the identified RNA to investigate their cellular functions and potential sequence similarities, which may lead to the identification of zip-codes for future verification. More importantly, this would allow for studies focusing on the significant implication of these processes in disease as well as the importance of *trans*-acting factors involved in the RNA localization. Though the viability of HeLa cells provide a faster mean for identifying RNA, interactions between RNA and cytoskeletal components must be further investigated in non-oncogenic cells to understand the role of the cytoskeleton in RNA mislocalization. Furthermore, our procedure has the ability to tag other molecular structures and does not have to be restricted solely to RNA.[22] Quasar 570 can also be replaced by several different dyes, which allows for studies involving multiple RNA simultaneously. This, paired with possible RNA

localization studies, can allow for future investigations on the connection between RNA localization, cytoskeletal components, and tubulinopathies resulting from mutations. Although further optimizations are required before usage, our lab was able to attach more than one dye to the unlabeled smFISH probes and created cell lines that successfully expressed HaloTag fused proteins, laying the groundwork for future optimizations.

Materials and Methods

TdT mediated attachment of Amino-11-d/ddUTP to unlabeled smFISH probes

The online Stellaris probe designer program was used to create a probe set for targeting the *luciferase* mRNA. Each unlabeled ssDNA oligo probe from the set was pooled together to create an equimolar mixture of a final concentration greater than or equal to 100µM. A mixture (1nmol oligonucleotide mixture, 1X Terminal deoxynucleotidyl transferase (TdT), 330µM Amino-11-dUTP or Amino-11-ddUTP, and 0.25µMCobalt Chloride) was briefly spun, and 0.8 units/µL TdT Enzyme was added. The TdT reaction mixture was incubated away from light in a PCR machine at 37°C with the hot lid set to the same temperature for 16-18 hours.

Purification of Amino-dUTP/ddUTP attached probes

A 200uL solution (15µL TdT reaction, 0.3M Sodium acetate (pH: 5.5), 1.5µg or greater of GlycoBlue coprecipitant) was added to 800µL of 100% ethanol prechilled at -20°C and then inverted 4 times to mix and incubated at 20°C for 20 minutes. Samples were centrifuged at 16,000 rcf for 20 minutes at 4°C, supernatant discarded, and 1mL of 80% ethanol prechilled at 4°C was added. The purification was vortexed to detach pellet, centrifuged at 16,000 rcf for 5 minutes at 4°C and supernatant discarded. Twice, 1mL 80% ethanol prechilled at 4°C was added and then vortexed until pellet completely detached, transferred into a new Eppendorf tube, centrifuged at 16,000 rcf for 5 minutes at 4°C, and supernatant discarded. Pellets were left to air dry on ice for 30 min to 1 hour and resuspended in 30-50µL nuclease free water. Using the ssDNA setting on Nanodrop 2.0, the concentration of the resuspended sample was obtained. Purified Amino-11-d/ddUTP probes were stored at -20°C until further use.

Denaturing Gel Analysis of Elongated probes

A 15% (v/v) polyacrylamide - 8M Urea stock (15% (v/v) Acrylamide/Bis solution, 1X TBE, 8M Urea) was heated until Urea fully dissolved. Stock was stored at room temperature away from light until further use. Gel casting components were washed with dH_20 and 70% ethanol. To cast gel, 10mL of the 15% PA-8M Urea stock, 1:200 fresh 10% (w/v) Ammonium persulfate, 1:1600 TEMED was quickly mixed and added to the dried and assembled gel cast. Gel was left to polymerize for 30 minutes to 1 hour. Samples in 3X loading dye (2.5% Ficoll-400, 10mM EDTA, 3.3 mM Tris-HCl, 0.08% SDS, 0.02% Dye 1 [pink/red], 0.001% Dye 2 [blue], pH 8) were prepared on ice. The polymerized gel was rinsed with dH_20, placed in and rinsed with 1X TBE. Gel was pre-ran for 30 minutes at 20mA (~200V) and wells rinsed again with the 1X TBE. Samples were loaded and gel run at 165V until the xylene cyanol (blue) and bromophenol blue (purple) markers reached 2/3 of the gel length and the bottom, respectively. Gel was imaged using UV302 autofocus setting on an Azure Biosystems C15 machine. Gel was also imaged for Gelgreen stain (~488nm) and Quasar 570 (~540nm) on a Sapphire Biomolecular Imager. Gel was stained with 1X Gelgreen in 1X TBE for 15 minutes and reimaged using the same channels as before.

Conjugation of NHS-ester Dye to Amino-11-d/ddUTP elongated probes

40mM dye (either Quasar 570 or Quasar 670) in anhydrous dimethyl sulfoxide (DMSO) was used to make a 2nmol/µL stock and was created along with a 1M Sodium bicarbonate stock

with a pH of 8.4 (adjusted using HCL). The average molecular weight of a probe 7,590 g/mol) was used to convert the grams of probe to the moles of probe in each mixture. From a denaturing gel, the average attachment of Amino-11-nucleotide to probes was visually obtained and used to calculate the mass of the average attached nucleotide ($moles\ x\ nucleotides\ attached$). From this, the amount of dye to add was calculated by using the obtained mass and multiplying it by the ratio of dye wanted (i.e. an excess of 5x the amount of probe). The NHS-ester reaction mixture was created by combining the following in a 30µL max volume: 1x mass elongated probes, 5x mass dye, and 0.1M Sodium bicarbonate (pH: 8.4). Reaction mixture was incubated away from light at room temperature (25°C) for 2-4 hours, purified through ethanol precipitation as described in "purification of Amino-11-d/ddUTP attached probes", and concentration measured. Stock solutions were stored at -20°C until further use.

Single molecule Fluorescence *in Situ* Hybridization

HeLa cells previously transfected with a construct containing a Tet-on expression system for the *luciferase* gene (pTL025) were grown to a 95% confluence. Glass coverslips were placed into wells of a 12-well plate and each added complete media (DMEM) containing 0.05 µg/mL Puromycin and 2µg/mL Doxycycline**.** Approximately 25,000 cells were plated into each well and left incubating at 37°C for 48 hours (5% CO_2). After incubation, wells had media removed and were washed with 1mL of phosphate buffer saline (PBS). 1mL of fixation buffer (3.7% formaldehyde in 1X PBS solution) was added and samples left incubating at room temperature for 10 minutes. Fixation buffer was removed, and samples were washed twice with 1mL PBS. 70% ethanol was added and left to incubate for one hour at room temperature. After incubation, ethanol was removed from samples and 1mL wash buffer (10% formamide and 2X saline-sodium citrate buffer containing 3.0M NaCl and 0.3M sodium citrate at pH 7.0) was added. Samples were left to incubate at room temperature for 2-5 minutes and a hybridization chamber was created. 125nM dye-conjugated probes were added into 200uL hybridization buffer (10% dextran sulfate, 2X saline-sodium citrate, 10% formamide) and the whole mixture was placed in the hybridization chamber. Coverslips were placed with cells facing towards the mixture and the hybridization chamber was closed, wrapped with aluminum foil and left to incubate at 37°C overnight (16 hours). After incubation, coverslips were gently transferred cell side up to a new 6-well plate, where 1mL wash buffer was added, and left to incubate in the dark at 37°C for 30 minutes. Wash buffer was discarded, 1mL of DAPI stain buffer (100ng/mL DAPI in 1X wash buffer) added and left to incubate for 5 minutes away from light at room temperature. 5uL Fluoromount G was added to each microscope slide and coverslips mounted with the cell side facing the droplet and left to dry overnight at 4°C away from light. DAPI, TRITC, and Cy5 channels were used in a wide-view microscope to visualize DAPI, GFP, and Dye excitation respectively.

Halo-Tag mediated tagging of Surrounding mRNA:

PCR Amplification of TUBA1A and Macf43 genes

A 25uL PCR reaction (1X Q5 high fidelity mastermix, 1.0 µM of each primer (primers 622 and 623 for amplification of *TUBA1A* and primers 624 and 625 for amplification of *Macf43*), 325ng of the plasmid containing both *TUBA1A* and *Macf43*) was created. Samples were placed in a PCR machine with the following settings: initial denaturing at 98°C for 30 seconds, 30 cycles- 98°C for 10 seconds, 66°C for *Macf43* primers and 65°C for *TUBA1A* specific primers for 30 seconds, 72°C for 30 seconds- and a final extension at 72°C for 2 minutes. Lid was set to 95°C and the hold temperature set to 4°C. Samples were stored at -20°C until further use.

Restriction Digest and DNA Gel Electrophoresis

A 50uL restriction digest (0.1X cutsmart, 0.4units/µL XhoI restriction enzyme, 3µg pTL032 plasmid) was created. Mixture was gently mixed, incubated at 37°C on a heatblock for 1-2hours, and either placed on ice for subsequent steps or stored at -20°C until further use. A 1% (w/v) agarose gel containing 1X Gelgreen and 1X TAE was prepared. Samples in1X DNA gel loading dye were loaded onto the gel and ran at 100V for 1-2 hours. Gel was imaged using the EpiBlue autofocus setting on Azure Biosystems C15 machine. A UVP Visi-Blue Transilluminator was used for visualization and excision of DNA bands. Extracted bands were purified following Zymoclean™ Gel DNA Recovery Kit procedure and concentration of each sample was obtained using nanodrop (2.0). Samples were stored at 20°C until further use.

Infusion cloning, competent cell transformation, and LB-media growth

The following components were mixed after being thawed on ice: 50ng of XhoI digest pTL032, amount of PCR product, and 1X mastermix from clontech infusion HD cloning kit. Mixture was incubated at 50°C for 15 minutes, incubated on ice for 2 minutes, and 1uL added to Stellar Competent Cells. Cells were incubated on ice for 30 minutes, at 42°C for 45 seconds, and on ice for 1 minute. After, 100uL SOC Medium was added and cells left to incubate at 37°C for 1 hour. 50uL of transformed cells were plated on 1X kanamycin agar plates and left to grow/incubate at 37°C for 16 hours. Single transformants from plates were transferred to LB-ampicillin media and left to incubate at 37°C for 16 hours.

ZR Plasmid Miniprep, Double Restriction Enzyme Digest, and DNA sequencing

Plasmids were isolated from transformed competent cells by following ZR Plasmid Miniprep™ - Classic Kit procedure and concentration (ng/uL) was measured using nanodrop (2.0). A 25uL double restriction enzyme digest (500ng plasmid, 0.2 units/µL PshAI, 0.4 units/µL XhoI, 0.1X Cutsmart buffer) was conducted at 37°C for two hours. Samples were analyzed on 1% (w/v) agarose gel. Gel was left running for ~2 hours and imaged using the EpiBlue autofocus setting on the Azure Biosystems C15. Isolates were verified for proper gene integration through Quintarabio overnight Seq-A-Strip DNA sequencing service.

HeLa Cell transfection with Plasmids

Frozen HeLa cell stocks were thawed on ice then transferred into a 6-well plate containing complete media (1X Fetal bovine serum (tetracycline free), 1X penicillin-streptomycin, and 1X DMEM and maintained until 80-90% confluent. A mixture (2µg *Macf43* or *TUBA1A* PCR product, 100ng pTL020, and 123.75uL of Lipofectamine LTX reagent) was added to cells and left to incubate at 37°C for 24 hours. Media was replaced with fresh complete media and cells allowed to recover at 37°C for 48 hours then media was replaced with fresh media containing 0.5µg/mL puromycin. Cells were left to grow until 95% confluent, trypsinized (37°C for 5 minutes) and transferred into a new 6-well plate. Complete media with 1µg/mL puromycin was added to cells. HeLa cells were grown to 95% confluence, trypsinized, split 3 times, and then frozen stocks created (-80°C overnight). Stocks were transferred into liquid nitrogen until further use.

Imaging Halo-Tag with Fluorophores

Transfected cells from T75 flasks, containing MACf43 or TUBA1, were transferred to 12-well plates containing coverslips and allowed to reach 90% confluence in complete media. Cells had complete media discarded and were washed with PBS as described in “HeLa Cell Transfection”. 500uL of 2% formaldehyde in PBS solution was added to each of the wells, cells allowed to incubate for 30 minutes at room temperature, and solution discarded. 500uL of 25nM JF646 dye

in PBS was added to each of the 12-plate wells. Solution was discarded from each well and cells washed with 1mL PBS along with 5 minutes of incubation at room temperature before repeating a total of three times. 500uL of 100ng/mL DAPI in PBS was added to cells and left to incubate at room temp. for 10 minutes. DAPI solution was discarded and wells washed once with PBS. Mounting slides were added 10uL of Fluoromount G and coverslips transferred over. Nail polish was added to the coverslips and allowed to dry for sealing. Samples were covered with aluminum wrap and left incubating at room temperature for 30 minutes. Slides were either visualized using a wide-view microscope or left to incubate in the dark at 4°C for approximately 52 hours.

Acknowledgements
I want to thank Dr. Jenn Garcia and Dr. Kristin Moore for substantial contributions in providing feedback through-out.

References

1. Bovaird, Samantha, et al. “Biological Functions, Regulatory Mechanisms, and Disease Relevance of RNA Localization Pathways.” *FEBS Letters*, vol. 592, no. 17, 5 Sept. 2018, pp. 2948–2972.,
2. Cody, Neal A.l., et al. “The Many Functions of MRNA Localization during Normal Development and Disease: from Pillar to Post.” *Wiley Interdisciplinary Reviews: Developmental Biology*, vol. 2, no. 6, 2013, pp. 781–796., doi:10.1002/wdev.113.
3. Chin A., Lécuyer E. (2017). RNA localization: Making its way to the center of the stage. *BBA General Subjects*, 2956-2970.
4. Schratt GM, Tuebing F., Nigh EA, Kane CG, Sabatini ME, Kiebler M and Greenberg ME (2006). A brainspecific microRNA regulates dendritic spine development. Nature 439, 283-289.
5. Taliaferro J., Wang E., Burge C. (2014). Genomic analysis of RNA localization. *RNA Biology*, 11(8), 1040-1050.
6. Jambhekar A., DeRisi J. (2007). *Cis-acting* determinants of asymmetric, cytoplasmic RNA transport. *RNA,* 13(5), 625-642.
7. Wang, Eric T., et al. “Dysregulation of MRNA Localization and Translation in Genetic Disease.” *The Journal of Neuroscience*, vol. 36, no. 45, Sept. 2016, pp. 11418–11426., doi:10.1523/jneurosci.2352-16.2016.
8. Haimovich, G. and Gerst, J. E. (2018). Single-molecule Fluorescence *in situ* Hybridization (smFISH) for RNA Detection in Adherent Animal Cells. *Bio-protocol* 8(21): e3070.
9. Wilkie G., Davis I. (2001). *Drosophila wingless* and Pair-Rule Transcripts Localize Apically by Dynein-Mediated Transport of RNA Particles. *Cell*, 105(2), 209-219
10. Tian, Li et al. “RNA-Binding Protein RBP-P Is Required for Glutelin and Prolamine mRNA Localization in Rice Endosperm Cells.” *The Plant cell* vol. 30,10 (2018): 2529-2552. doi:10.1105/tpc.18.00321
11. Xiao, Qingpin, et al. “Cytoskeleton Molecular Motors: Structures and Their Functions in Neuron.” *International Journal of Biological Sciences*, vol. 12, no. 9, 2016, pp. 1083–1092., doi:10.7150/ijbs.15633.
12. Kato, Yusuke, et al. “Overview of the Mechanism of Cytoskeletal Motors Based on Structure.” *Biophysical Reviews*, vol. 10, no. 2, Dec. 2017, pp. 571–581., doi:10.1007/s12551-017-0368-1.
13. Medioni, C., et al. “Principles and Roles of MRNA Localization in Animal Development.” *Development*, vol. 139, no. 18, 2012, pp. 3263–3276., doi:10.1242/dev.078626.
14. Mofatteh, Mohammad, and Simon L. Bullock. “SnapShot: Subcellular MRNA Localization.” *Cell*, vol. 169, no. 1, 23 Mar. 2017, doi:10.1016/j.cell.2017.03.004.
15. Blower, Michael D. “Molecular insights into intracellular RNA localization.” *International review of cell and molecular biology* vol. 302 (2013): 1-39. doi:10.1016/B978-0-12-407699-0.00001-7
16. Suter, B (2018). RNA localization and transport. *BBA, Gen Regulatory Mechanisms*, *1861(10),* 938-951.
17. Neelamraju, Yaseswini, et al. “Mutational Landscape of RNA-Binding Proteins in Human Cancers.” *RNA Biology*, vol. 15, no. 1, 2017, pp. 115–129., doi:10.1080/15476286.2017.1391436.
18. Hebebrand, M., Hüffmeier, U., Trollmann, R. *et al.* The mutational and phenotypic spectrum of *TUBA1A*-associated tubulinopathy. *Orphanet J Rare Dis* 14, 38 (2019).
19. Aiken, Jayne et al. “The *α*-Tubulin gene *TUBA1A* in Brain Development: A Key Ingredient in the Neuronal Isotype Blend.” *Journal of developmental biology* vol. 5,3 (2017): 8. doi:10.3390/jdb5030008
20. Erdogan, Burcu, et al. “The Microtubule plus-End-Tracking Protein TACC3 Promotes Persistent Axon Outgrowth and Mediates Responses to Axon Guidance Signals during Development.” *Neural Development*, vol. 12, no. 1, 2017, doi:10.1186/s13064-017-0080-7.
21. Kwon S. (2013). Single-molecule fluorescence in situ hybridization: quantitative imaging of single RNA molecules. *BMB reports*, *46*(2), 65–72.
22. Li Y., Aggarwal M., Nguyen K., et al. (2017). Assaying RNA localization in Situ with Spatially Restricted Nucleobase Oxidation. *ACS Chemical Biology*, 12(11), 2709-2714.
23. Tsanov N., Samacoits A., Chouaib R., et al. (2016). smFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Research, 44(22),* 1-11*.*
24. Gasper I., Wippich F., Ephrussi A., et al. (2017). Enzymatic production of single-molecule FISH and RNA capture probes. *RNA Journal, 23, 1582-1591.*
25. Fazal F., Han S., Parker K., et al (2019). Atlas of Subcellular RNA localization Revealed by APEX-Seq. *Cell, 178, 473-490.*

Figures

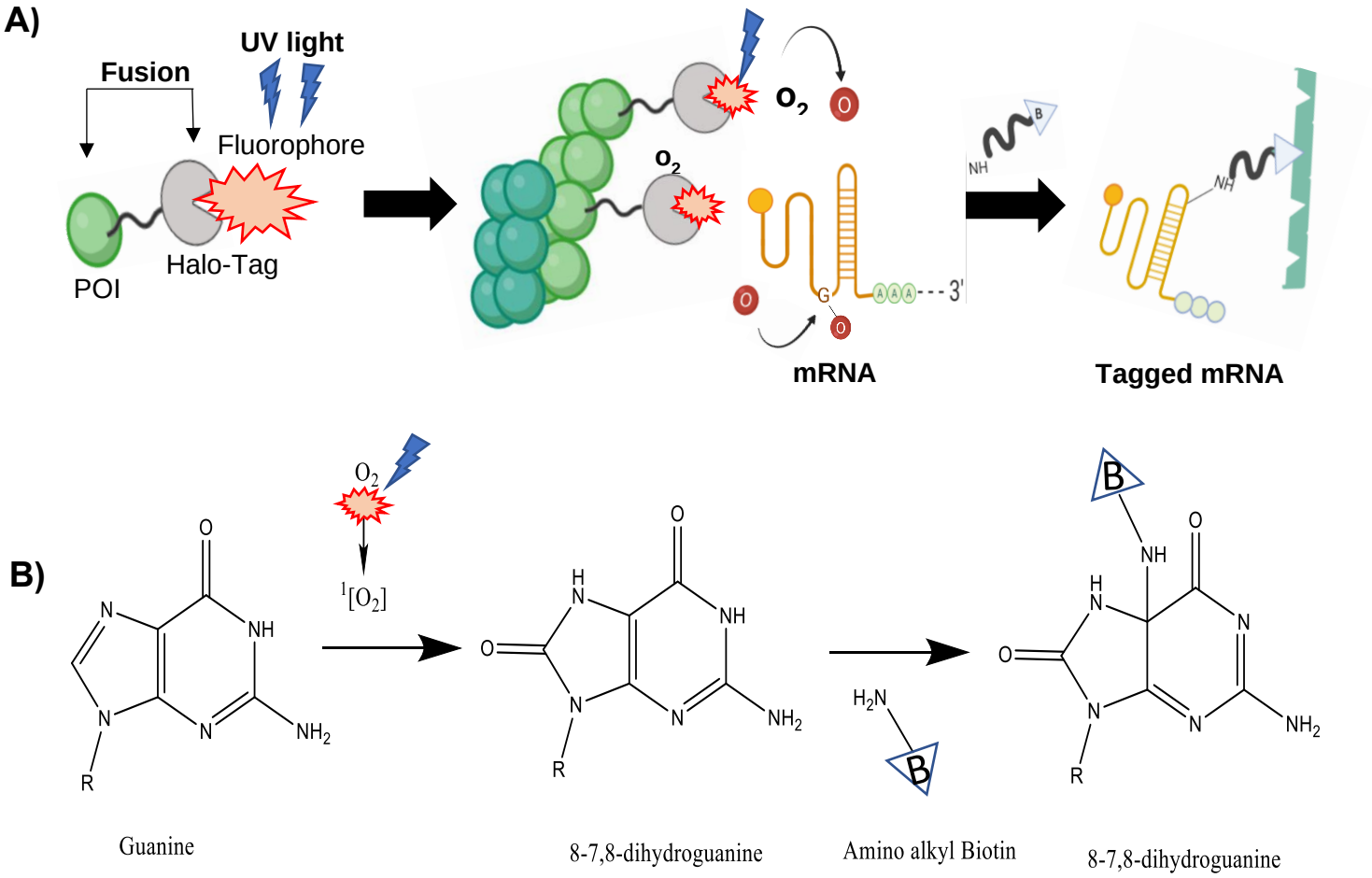

Figure 1. Nucleobase Oxidation coupled with HaloTag mediated localization allows for nearby mRNA tagging. (A) The schematic depicts a simplified procedure for localized nucleobase oxidation of RNA. The method begins with the expression of a Halo tagged protein with a known localization (*i.e*. Protein of interest, POI). A modified fluorophore containing the Haloligand is introduced and subsequently binds the HaloTag protein, which brings the fluorophore to a specific cellular location. The fluorophore uses free energy from blue light to excite triplet oxygen from a ground state to a singlet state. The oxygen radical interacts with guanine residues in nearby RNA to form 8-7,8-Oxoguanine or 8-Oxoguanine. Modified RNAs can be purified by adding amino alkyl biotin to result in addition of biotin to the RNA. This attached biotin is used to purify tagged RNA through a streptavidin column chromatography. **(B)** A schematic representing the chemistry of 8-Oxoguanine RNA labelling and attachment of biotin. A nucleophile, amino alkyl biotin in this case, attacks the electrophile, 8-oxoguanine to form a guanine residue containing biotin.

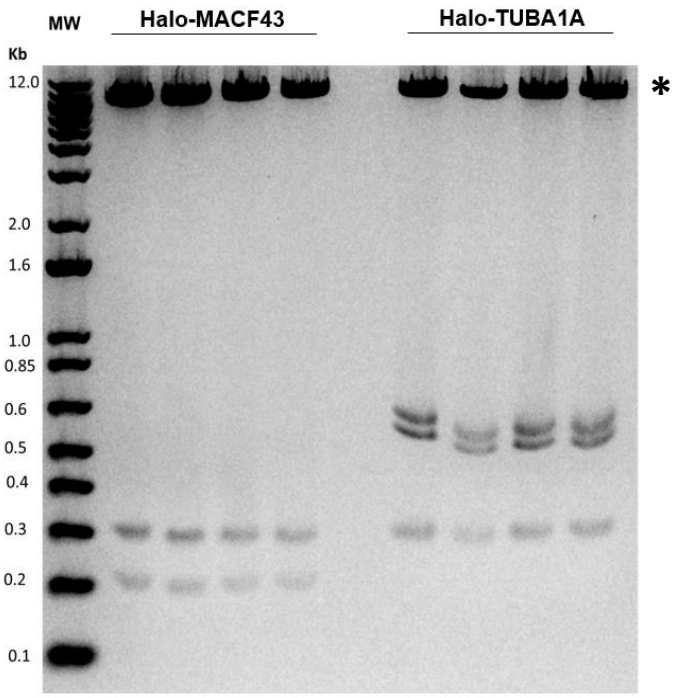

Figure 2. Double restriction digest of fused plasmids suggests correct integration of *TUBA1A* and *MACF43.* Plasmids were extracted using ZR Plasmid Miniprep-Classic. A double restriction digest was performed using XhoI and PshAI on each plasmid. Four bands were expected for digest of plasmids containing *TUBA1A* (lanes 7-10) and three were expected for *Macf43* (lanes 2-5). 1kb plus ladder was used for the molecular weight (lane 1). (*****) indicates unexpected 12kb bands

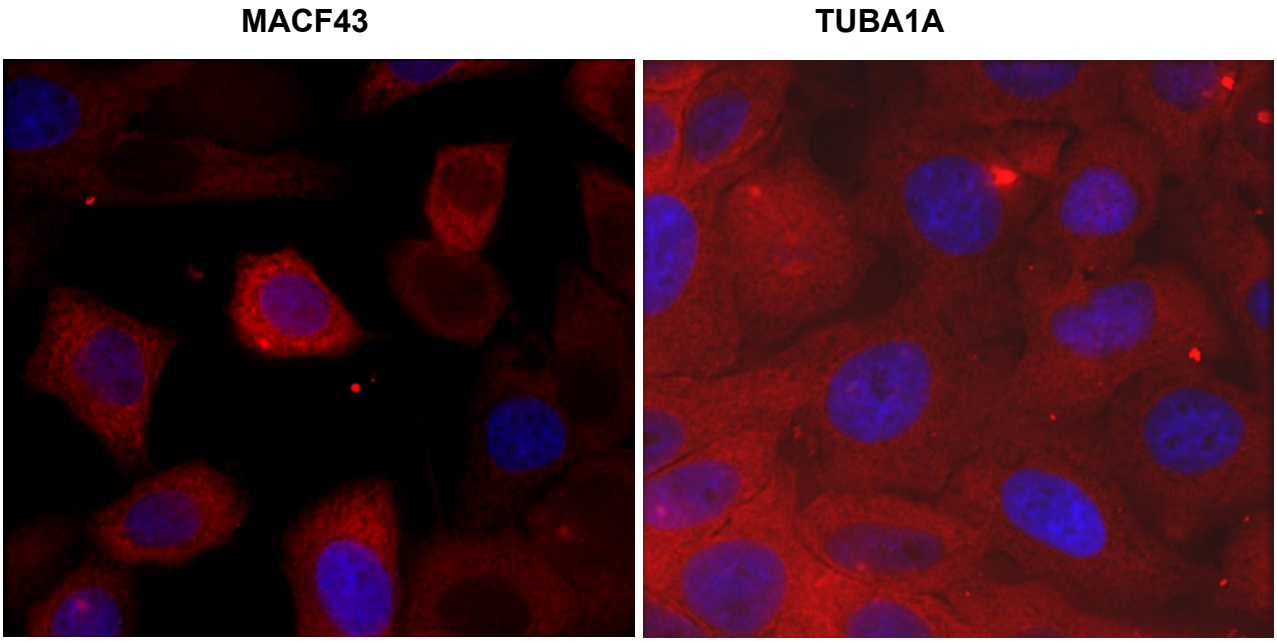

Figure 3. Microscopy of expressed HaloTag fusion protein shows intracellular binding of fluorophore. HeLa cells were transfected with a plasmid containing Halotag fusion to either *TUBA1A* or *Macf43.* Cells, grown to 95% confluence, were exposed to doxycycline for induction of the Halo-tagged protein. Cells were exposed to JF646 (dye) for fluorophore localization, DAPI used for nuclear staining, and cells washed to remove unbound dye. A wide-view microscope was used to image both the dye and DAPI stain. Control samples (WT HeLa cells) were not visible during dye excitation and are not shown. Dye (red) and DAPI (blue).

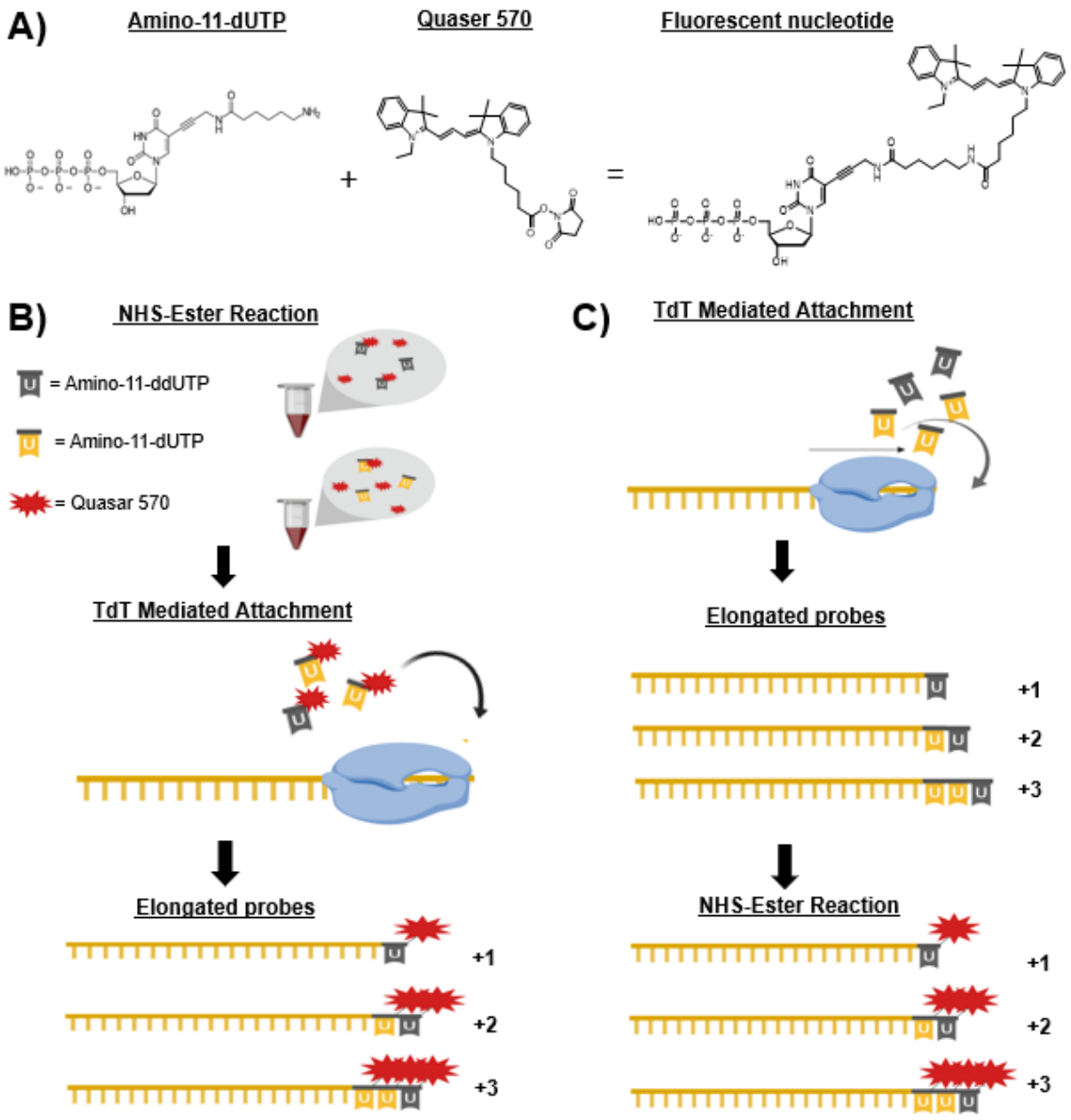

Figure 4. NHS-ester reaction and TdT mediated attachment utilized for unlabeled smFISH probe elongation. The following schematic shows the components involved in the NHS-ester reaction and the steps taken for addition of more than one dye to unlabeled smFISH luciferase probes. The theoretical results are shown for each. **(A)** NHS-ester reaction involves Amino-11-dUTP or Amino-11-ddUTP (not shown) and Quasar Succinimidyl NHS ester. (B) Method 1 involved conjugation of Quasar 570 Succinimidyl NHS-ester dye to Amino-nucleotides first followed by terminal deoxynucleotidyl transferase (TdT) mediated attachment of modified dye-nucleotides to unlabeled smFISH probes. **(B)** Method B first involves TdT mediated attachment of Amino-nucleotides not conjugated to the dye followed by NHS-ester reaction with Quasar 570 smFISH probes.

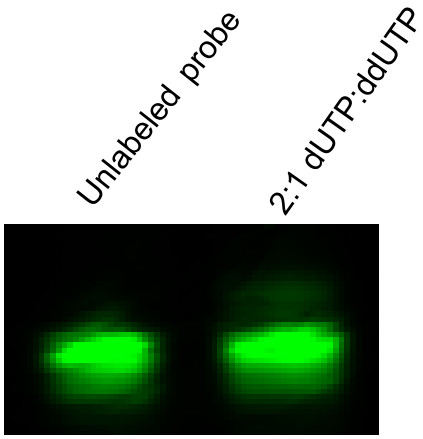

Figure 5. Denaturing gel suggest inadequate dye-Amino-11-nucleotide attachment to smFISH probes. Samples were run at 15mA in a 15% Polyacrylamide- 8M Urea denaturing gel and stained with Gelgreen (1X). Fluorescence setting on Sapphire Biomolecular Imager with excitation for Gelgreen (488nm) and dye (540nm) was used to obtain the image. Quasar 570 was conjugated to Amino-11-nucleotides through an NHS-ester reaction and used for elongation of the unlabeled luciferase 3' with TdT. Unlabeled *luciferase* probe (lane 1) and 2:1 ratio dUTP to ddUTP (lane 2).

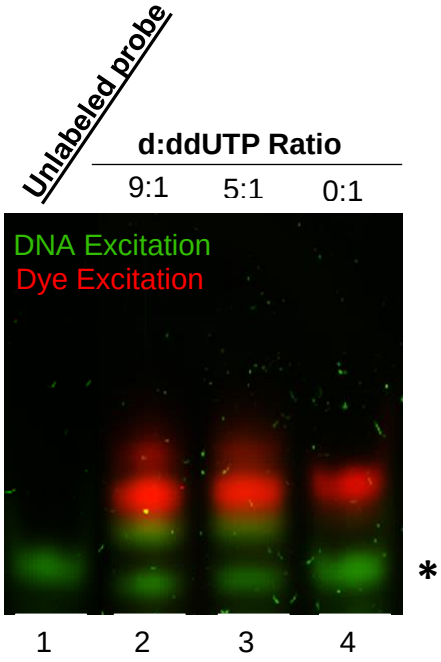

Figure 6. Denaturing gel shows smFISH probes with a single dye attachment. Quasar 570 (dye, red) was conjugated to Amino-11-nucleotides through an NHS-ester reaction and used as reagents for 3' labelling of the oligos with terminal deoxynucleotidyl transferase. TdT reactions were run at 15mA in a 15% Polyacrylamide- 8M Urea denaturing gel and stained with Gelgreen (1X). Fluorescence setting on Sapphire Biomolecular Imager with excitation for Gelgreen (488nm) and Quasar 570 dye (540nm) was used to obtain the image. (*****) refers to the unlabeled probe or the lowest molecular weight band.

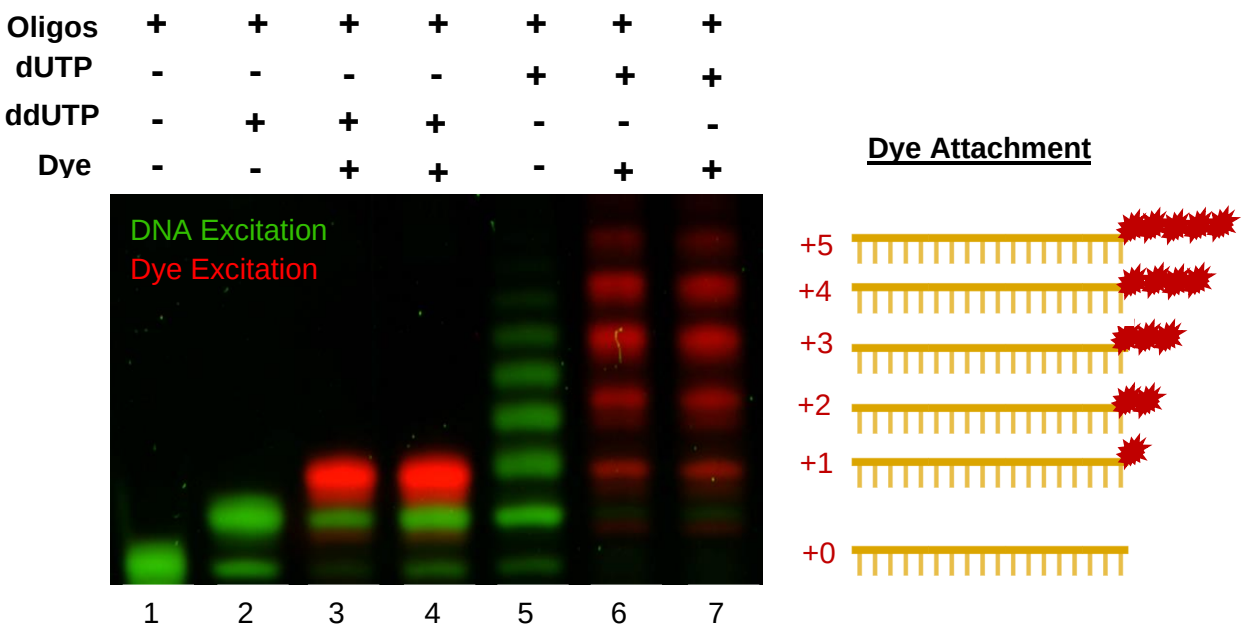

Figure 7. Denaturing gel shows smFISH probes with varying degrees of dye attachment. Amino-11-nucleotides were attached to unlabeled luciferase probes first and conjugated to dye last in reactions that contained only amino-11-dUTP (lanes 5-7) or amino-11-ddUTP (lanes 2-4). Quasar 570 (dye) was conjugated to Amino-11-nucleotides through an NHS-ester reaction (lanes 3, 4, 6, 7) as described in Figure 4. Oligos elongated with Amino-11-nucleotides and oligos with conjugated dyes were run at 15mA in a 15% PA- 8M Urea denaturing gel and stained with 1X Gelgreen. Simplified schematic of multiple dye attachments shown on the right.

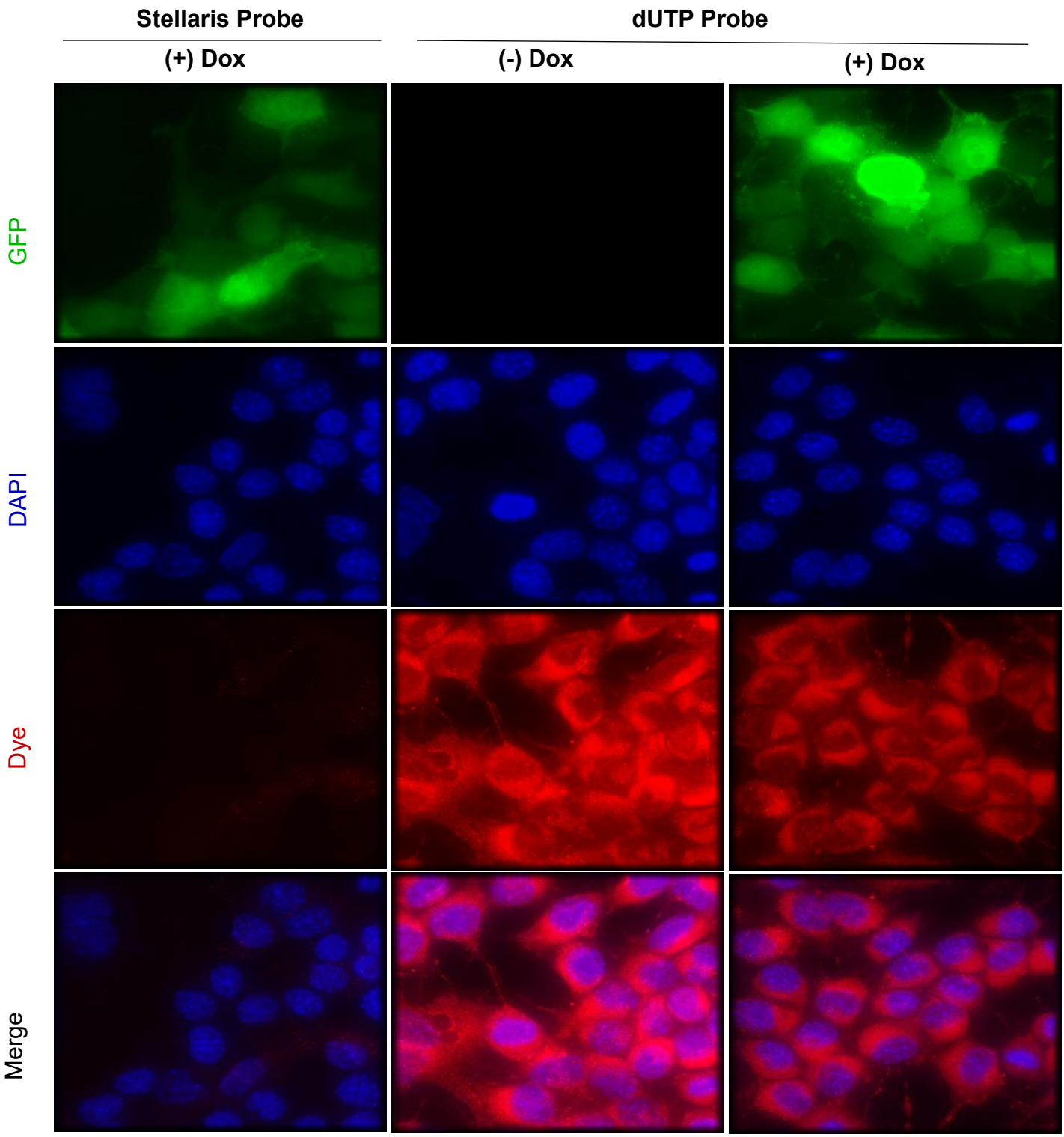

Figure 8. Microscopy of smFISH conducted with homemade probes reveal large background noise. HeLa cells with a Tet-on expression system construct for the *luciferase* and *GFP* genes were used. Commercially available probes (Stellaris) and homemade probes (dUTP) were used for smFISH. Images were obtained using a wide-view microscope. GFP is in green, DAPI staining for nuclear visualization is in blue, and smFISH probes are in red. HeLa cells were incubated in the presence of doxycycline and smFISH was performed using commercially available luciferase probes (column 1). smFISH performed with homemade probes on HeLa cells not introduced doxycycline (column 2). smFISH was performed with homemade probes on HeLa cells exposed to doxycycline (column 3). *Luciferase* was induced by incubating with Doxycycline for 48 hours.

Size Exclusion dUTP Probe

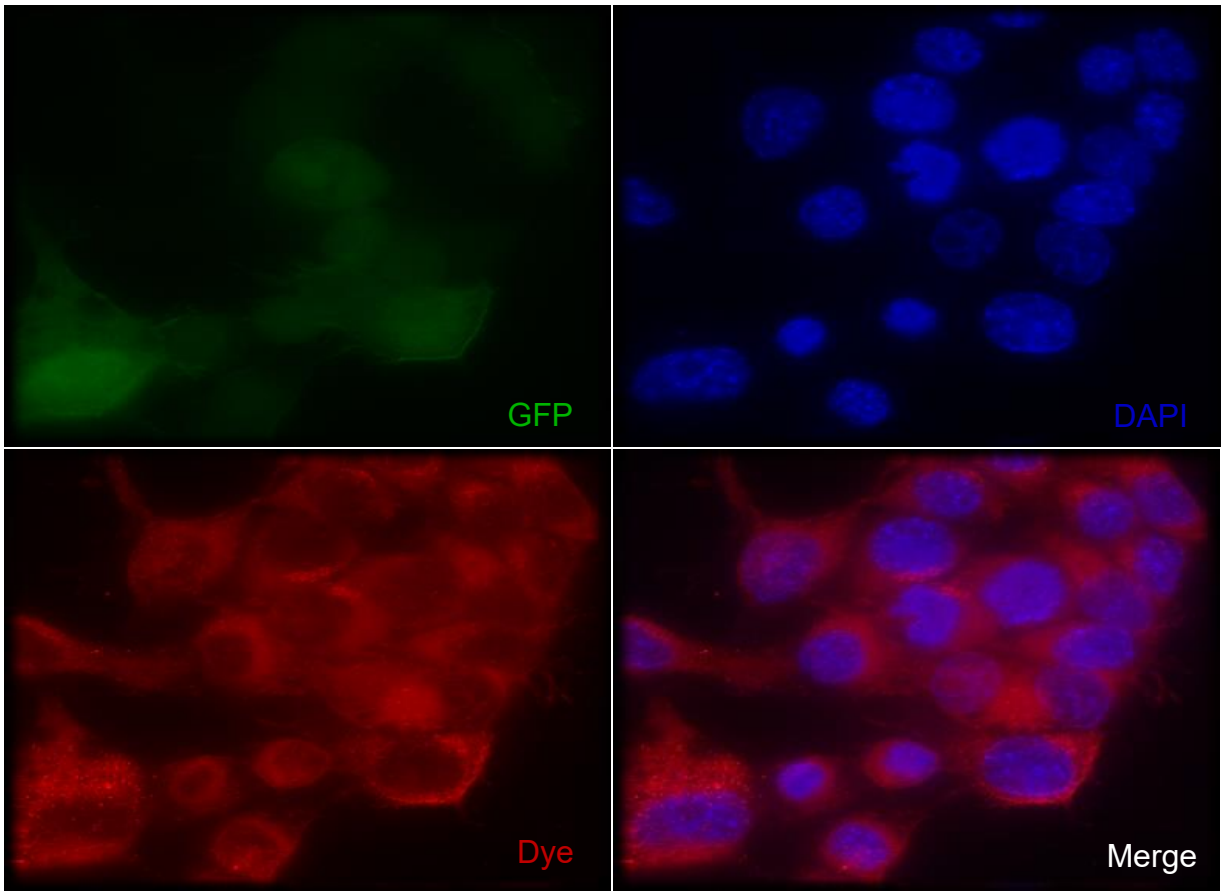

Figure 9. Microscopy of smFISH conducted with further purified homemade probes reveals less background noise. HeLa cells with a Tet-on expression system construct for the luciferase and GFP genes were used. Cells were exposed to doxycycline during growth. Homemade probes (dUTP) further purified through size exclusion chromatography were used for smFISH. Images were obtained using a wide-view microscope. GFP is in green, DAPI staining for nuclear visualization is in blue, and smFISH probes are in red.