Applications of Optical Control of Oligonucleotide and Protein Function

by

James Bride Hemphill

B.S. Molecular Genetics, University of Vermont, 2007

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

Dietrich School of Arts and Sciences

This dissertation was presented

by

James Bride Hemphill

It was defended on

July 29th 2015

and approved by

W. Seth Horne, Assistant Professor, Department of Chemistry

Kabirul Islam, Assistant Professor, Department of Chemistry

Michael Tsang, Associate Professor, Department of Developmental Biology

Dissertation Advisor: Alexander Deiters, Professor, Department of Chemistry

Applications of Optical Control of Oligonucleotide and Protein Function

James Bride Hemphill, PhD

University of Pittsburgh, 2015

Optical regulation using light as an external trigger was applied to the control of biological processes with high spatio-temporal resolution. Photoremovable caging groups were site-specifically incorporated onto oligonucleotides and proteins to optically regulate their function in biological environments, typically for the photochemical control of gene expression. These caging group modifications enabled both OFF → ON and ON → OFF optochemical switches for important chemical biology tools. Oligonucleotides containing caging group modifications were synthesized to regulate nucleic acid function with light. Specifically, photocaged triplex-forming oligonucleotides were developed to optochemically control transcription in cell culture. Light-activated antagomirs were designed for the optical inhibition of miR-21 and miR-122 function in the regulation of endogenous microRNA activity. This technology was then applied to the study of miR-22 and miR-124 function in cortical neuron migration during cerebral corticogenesis. Splice-switching oligonucleotides were engineered to optically control mRNA splicing pathways in both human cells and zebrafish. The optical control of plasmid-based gene expression was demonstrated with a caged promoter, and applied to the photochemical activation of transcription in a live animal model. The caging of oligonucleotides was also applied to DNA computation in the production of optically controlled logic gates and amplification cycles, providing spatio-temporal control over hybridization cascades to add new functionality to DNA computation modules. These studies in DNA computation led to the development of novel biosensors for

logic gate-based detection of specific micro RNA signatures in live cells. In addition, proteins were optically controlled through the site-specific installation of caging groups on amino acid side chains that are essential for protein function using unnatural amino acid mutagenesis in mammalian cells with an expanded genetic code. A caged lysine analogue was incorporated into T7 RNA polymerase to photochemically regulate transcription in the development of a light-activated synthetic gene network and light-triggered RNA interference. A light-activated Cas9 endonuclease was engineered through the installation of a caged lysine analogue to optically control CRISPR/Cas9 editing of both exogenous and endogenous genes. Lastly, a system for the incorporation of unnatural amino acids in zebrafish was studied in efforts to produce the first vertebrate species with an expanded genetic code.

Table of Contents

List of Tables

List of Figures

xiii

1.0   Introduction to Caged Oligonucleotides and Proteins

Gene expression is tightly controlled in single cells and multicellular organisms, and the precise activation and deactivation of genes in high spatio-temporal resolution is required for many biological processes.[1] In order to dissect or reprogram genetic circuits with the same spatio-temporal resolution observed in nature, light can be used as an external control element. Light activation provides advantages over other methods of external triggers, such as specific timing or location of irradiation, controlled amplitude, and wavelength specificity that is fully orthogonal in most biological systems.[2] The use of light-removable protecting groups to control biological events was first applied to cyclic adenosine monophasphate[3] (cAMP) and adenosine triphosphate[4] (ATP). While initially developed by organic chemists for use as protecting groups for multi-step syntheses,[5] these "caging" groups found many new roles in the regulation of biochemical processes.[6] Photoactivated nucleotides were initially used to study the kinetics of ion channels,[7] muscle fibers,[8] and triphosphate-based reactions,[9] providing insight into dynamic biological activities, such as sensory neuronal function.[10] These findings were some of the earliest applications of caged small molecules to study biological processes with light, allowing for rapid activation and analysis of function in ways that were previously not possible. Building on that research, glutamate, γ-aminobutyric acid, inositol trisphosphate, and calcium have all been rendered light-activatable for the optical interrogation of neurotransmitters and subsequent photostimulation of neural circuits (Figure 1.1).[11]

1
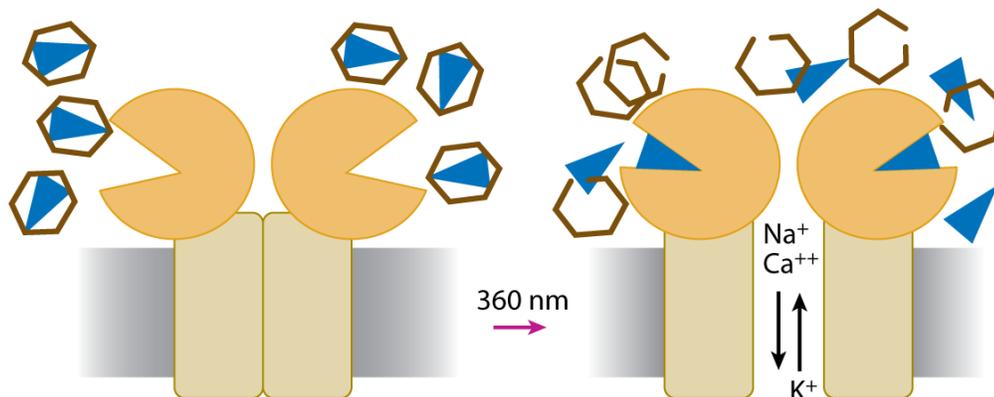
Figure 1.1: Photostimulation of caged neurotransmitters.
Caged neurotransmitters (brown hexagons) can be activated with light, allowing the decaged compounds (blue triangles) to interact with neuronal targets, for example in the activation of ion channels. Image adapted with permission from Szobota and Isacoff, *Annu. Rev. Biophys.* **2010**, 39:329-48.

The installation of caging groups on larger biological macromolecules is a versatile approach to the optical regulation of DNA, RNA, and protein activity.[12] Several photoremovable protecting groups are available for researchers to interrogate biological systems using light, each with different synthetic requirements for installation and unique parameters for the release of the active biomolecule.[13] Biomolecules are caged with these photolabile groups in a manner that specifically inhibits function, which is then restored through light irradiation induced "decaging". The central focus of this thesis is the development of chemical biology tools based on caged oligonucleotides and caged proteins that enable control of important biological processes in a precise manner that other research tools cannot provide. The two main photoremovable protecting groups that were utilized in this thesis are derivatives of the *ortho*-nitrobenzyl alcohol (ONB) and coumarin-4-ylmethanol molecules.[13]

Caging groups can be installed onto nucleic acid bases to optically activate oligonucleotide duplex formation (Figure 1.2A), which is critical for the transfer of genetic

information and gene expression. Additionally, caged oligonucleotides have applications in the field of DNA computation as logic gate components. Proteins can also be optically controlled through the site-specific installation of caging groups on amino acid residues at active sites or in close proximity to residues required for biological activity (Figure 1.2B). The photochemical control of biological processes with caged biomolecules should fulfill the following requirements: A) caging group linkage that is stable under physiological conditions, B) complete inhibition of the biological function of interest, C) decaging with non-damaging light,[14] D) rapid and efficient decaging, as well as E) generation of nontoxic byproducts.[15]
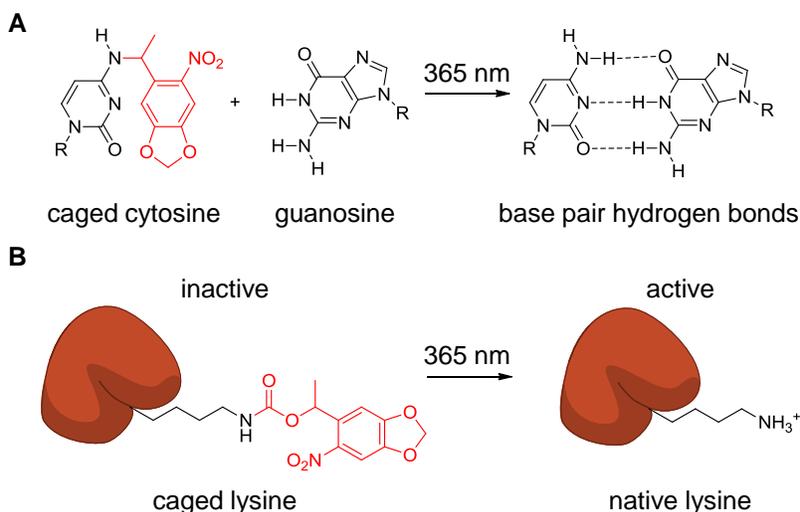
Figure 1.2: Photocaging of biological molecules.
A) DNA nucleobase caging prevents base pair hybridization until decaged through UV irradiation. R groups represent oligonucleotide chains. The *ortho*-nitrophenylethyl (NPE) caging group is shown in red. B) Site-specific amino acid caging inactivates protein function until decaged with UV irradiation. The 6-nitropiperonyloxycarbonyl (NPOC) caging group is shown in red.

1.1   Optical Control of Oligonucleotide Function

The term oligonucleotide describes a broad class of nucleic acid polymers, including deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). These organic polymers consist of a sugar-phosphate backbone that is connected at the 3' and 5' positions, with heterocyclic nitrogenous bases that are linked to the sugars at the 4' position (Figure 1.3).[16] There are structural differences between DNA and RNA that are biologically relevant for specific function. The 2' position of the RNA sugar ring has a hydroxyl group, in comparison to the 2' hydrogen on DNA. There is a modification between the DNA base thymine and the RNA base uracil, in which the DNA base form is methylated. DNA exists as an anti-parallel double-stranded helix, where as RNA is observed as single stranded helical polymers that can form complex secondary structures, such as stem-loops and bulges.
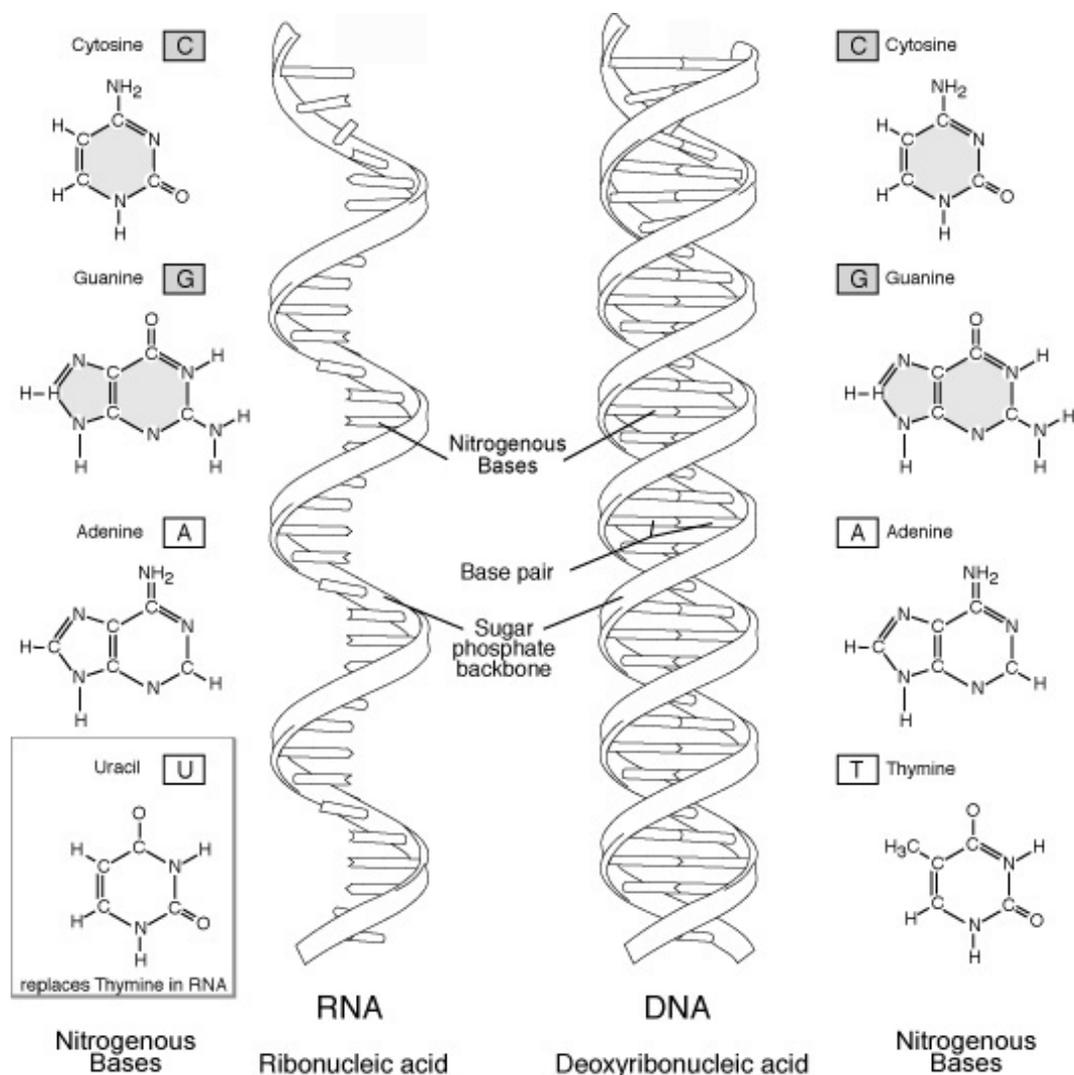
Figure 1.3: Comparison of RNA and DNA helices as well as nucleobase structures. Image adapted from the National Human Genome Research Institute.

There are many known and well defined roles of oligonucleotides in biology, specifically in genetic information transfer and in the control of gene expression. The central dogma describes the flow of genetic information,[17] in which DNA is transcribed into RNA and RNA is translated into protein. Within this genetic flow, there are many different applications of synthetic oligonucleotides for the regulation of gene function. The design parameters for synthetic oligonucleotides can alter gene expression by interacting with specific substrates within

the central dogma; for example, binding to DNA with triplex-forming oligonucleotides (TFOs) can inhibit transcription into RNA,[18] mRNA templates can be inhibited with antisense agents,[19] microRNA (miRNA) and RNA interference (RNAi) pathways can be controlled with antagomirs or synthetic small-interfering RNA (siRNA) duplexes,[20-22] splice-switching oligonucleotides (SSOs) can control RNA splicing pathways of translation into functional protein products,[23] protein function can be inhibited by aptamers,[24] etc. These many classes of oligonucleotides have broad structural modification requirements and utilize different mechanism for controlling gene expression.

Synthetic oligonucleotides have been manufactured with a variety of modified functional groups. The modifications can be introduced at different domains of the oligonucleotide. Common regions for modification include the phosphate backbone,[25] sugar ring,[26] nucleobase,[27] and chain terminus.[28] These modifications have broad and unique characteristics, such as fluorophores for oligonucleotide labeling or sulfur containing bases capable of forming disulfide bonds in bioconjugation reactions. A large focus of nucleic acid modification is to provide important functionality for biological applications, including increased stability, tighter binding affinity, and enhanced *in vivo* efficacy.[29] However, the application of caging groups provides the opportunity to photochemically control oligonucleotide function with high levels of spatial and temporal resolution. These caging groups are typically introduced at the common oligonucleotide modification sites described above. The function of caged oligonucleotides can generally be described by two classes: those containing photoremovable protecting groups on structural components, and those containing photocleavable linkers between structural components (Figure 1.4).[30] The application of these optically controlled oligonucleotides, for either *in vivo* or *in vitro*

studies, can also generally be described in two classes: those that are optically activated, or those that are optically deactivated.
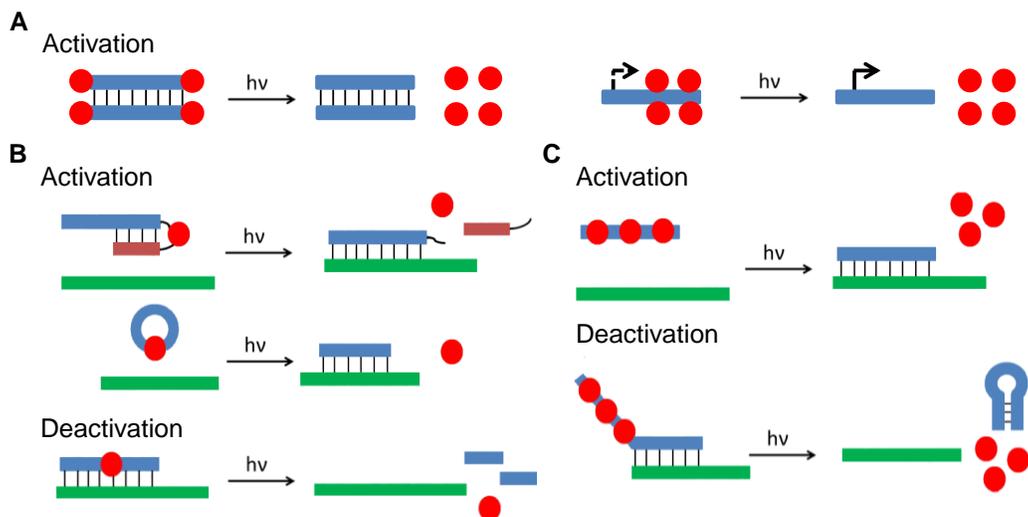
Figure 1.4: Caging approaches to regulate oligonucleotide binding with light.
A) Photocaged phosphate backbones. B) Photocleavable linkers. C) Photocaged nucleobases. Modified oligonucleotides are indicated in blue, with the photoresponsive moieties indicated in red. Target strands are indicated in green. Arrow above DNA represents transcription. Image adapted with permission from Liu and Deiters, *Acc. Chem. Res.* **2014**, 47(1):45-55. Copyright 2013 American Chemical Society.

Oligonucleotides caged in either of these manners have been used for the *in vitro* photochemical control of probes for abasic sites in RNA,[31] recombinant DNA manipulation,[32] DNA aptamer activation,[33] and ribozyme function.[34, 35] There are many examples of optically controlled gene regulation using caged oligonucleotides *in vivo* with high spatial and temporal control, including caged plasmid DNA and mRNA,[36] DNA decoys,[37] antisense technology,[38, 39] and RNA interference.[40, 41]

Methods for the generation of optically controlled oligonucleotides involve the modification of nucleic acid structures, either through post-synthesis chemical reactions or during solid-phase synthesis of caged oligonucleotides. A phosphate backbone caged plasmid

was generated through chemical modification of a native nucleic acid structure (Figure 1.4A), but lacked sequence-specificity and the ability to control the number of caging groups attached.[42] While backbone caging inactivated transcription of the plasmid, full restoration was not achieved with UV irradiation due to incomplete removal of the numerous caging groups. Similar approaches to apply caging groups nonspecifically to the backbone of plasmid DNA and mRNA structures for the control of gene expression were also met with the same limitations (i.e., only partial recovery of function after long irradiation times).[36, 43] This approach was then applied to the control of small interfering RNA (siRNA) oligonucleotides, but again clean OFF → ON light switching could not be realized.[44] However, significant progress has been made over the years in the selective installation of phosphate caging groups and control over the number of caging group additions. Through these improvements, the potential of light activated RNA interference (LARI) with terminal phosphate caged siRNA duplexes was realized for gene regulation in cell culture (Figure 1.5).[45]
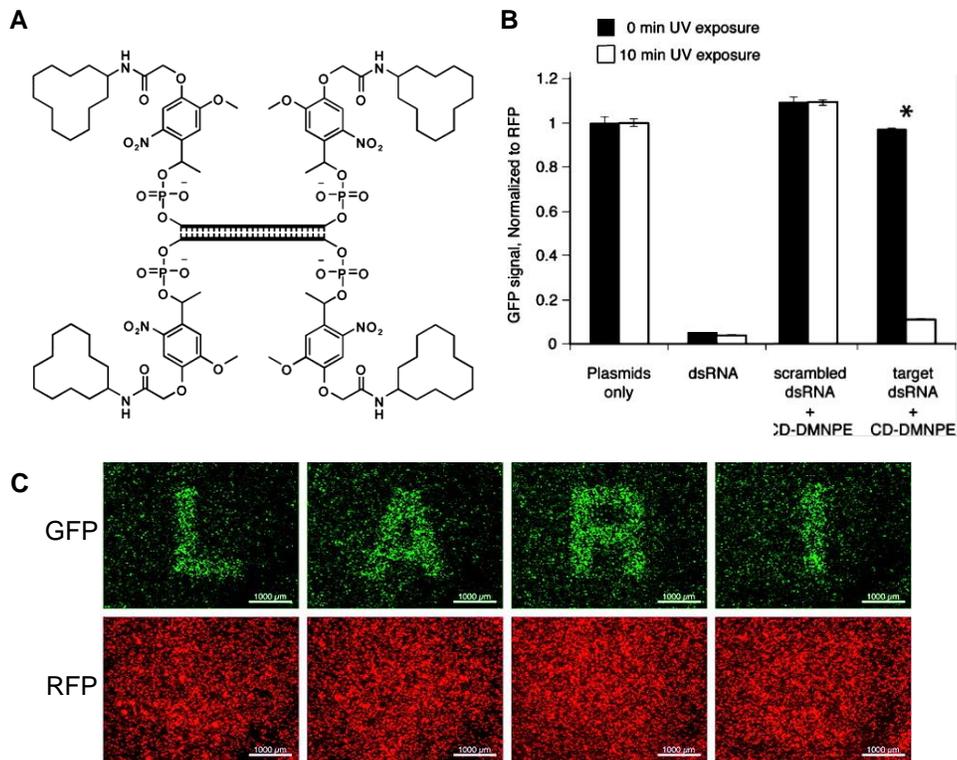
Figure 1.5: Patterned gene expression using light-activated RNAi.
A) Caging groups were installed at the 5' and 3' terminal phosphates on a dsRNA duplex targeting GFP. B) The siRNA duplexes were transfected into cells with GFP and RFP expression plasmids, then irradiated with UV light for 10 min. C) UV irradiation through spatially restricted masks enables spatial control of RNAi. Image adapted with permission from Jain et al, *J. Am. Chem. Soc*. **2011**, 133(3):440-446. Copyright 2011 American Chemical Society.

In this study, the siRNA targeting GFP was chemically modified in the duplex form, demonstrating post-synthesis addition of caging groups onto oligonucleotides. A total of four caging groups were incorporated per duplex, as modification to each terminal phosphate was achieved (Figure 1.5A). To test the optical activation of RNAi, the siRNA duplexes were transfected into cells with expression vectors for two reporter genes, green fluorescent protein (GFP; RNAi target) and red fluorescent protein (RFP; transfection control). After transfection, the cells were UV irradiated and expression of the fluorescent reporters was analyzed (Figure 1.5B). The noncaged siRNA duplex showed significant reduction of GFP expression, while a negative control scrambled siRNA sequence containing the caging group modifications showed

9

no activity. The caged siRNA duplex targeting GFP showed no inhibition of GFP expression in the absence of UV irradiation, and photochemical activation with UV irradiation reduced GFP expression to similar levels as the noncaged siRNA duplex. This caged siRNA methodology was then applied to the spatial control of RNAi, by irradiating cells transfected with the caged oligonucleotide through spatially restricted masks (Figure 1.5C). Spatial control of the caged siRNA was demonstrated with inhibition of GFP expression, as the optical activation was restricted to the irradiated areas outside of the "LARI" masks.

Working towards the site-specific incorporation of internal photoreactive moieties into oligonucleotides, photocleavable linkers have been introduced to link multiple oligonucleotide segments, or to link the terminal ends of a single oligonucleotide (i.e. circularization). These reactions are also performed post-synthesis, to modify nucleic acids through chemical ligation, and enable light-induced degradation (Figure 1.4B). The use of photocleavable linkers for photochemical regulation of oligonucleotide function in biological environments has been well demonstrated,[46] especially with morpholinos (MOs), which are chemically altered oligonucleotides that contain a morpholine ring and dimethyl amine-modified phosphate backbone.[47] These modifications enhance the MO stability and efficacy for *in vivo* targeted gene knockdown commonly applied in zebrafish studies.[48] While the composition of MOs differs from other nucleic acids, the use of photochemical caging groups can still be applied for the optical control of oligonucleotide activity. For example, a MO antisense agent containing a photocleavable linker was applied to the optochemical regulation of endogenous gene inhibition in zebrafish (Figure 1.6).[38]
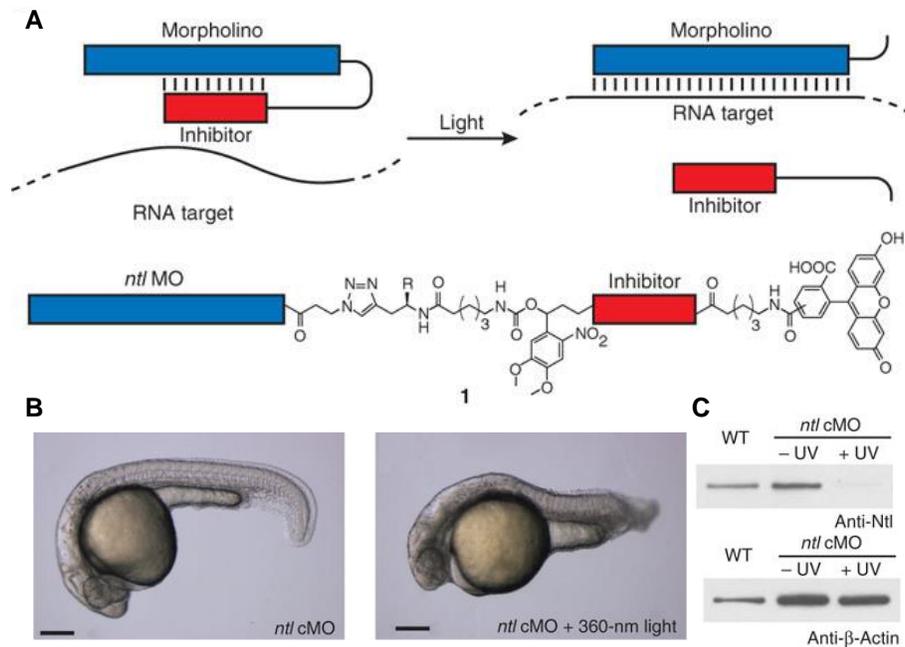
Figure 1.6: Light-controlled gene silencing in zebrafish embryos.
A) The MO morpholino (blue) is connected to a complementary inhibitor strand (red) through a photocleavable linker, enabling light-triggered hybridization to RNA targets. The structure of the linker for the caged *ntl* morpholino (cMO) is shown. B) Embryos were injected with the *ntl* cMO and exposed to UV light or kept in the dark. C) Western blot of *ntl* expression, with a β-actin control. Image adapted with permission from MacMillan Publishers Ltd: *Nature Chem. Bio*. 133(3):440-446, copyright 2011.

Here, a MO targeting *ntl*, a gene responsible for formation of the tail and notochord, was bound through a photocleavable linker to an inhibitor strand containing a complementary sequence (Figure 1.6A). Thus, the binding of the caged MO (cMO) to its target mRNA was prohibited until the inhibitor strand was removed through photochemical cleavage of the linker. When injected into zebrafish embryos, activation of cMO function through light irradiation resulted in the loss of posterior structures and tail deformation (Figure 1.6B). Additionally, the loss of *ntl* was observed by Western blot (Figure 1.6C). This study demonstrated that photoresponsive nucleic acids could be applied to live animals to achieve spatio-temporal regulation of gene expression.

Alternative methods for the synthesis of caged oligonucleotides were developed to further control the specific number and location of the caging groups. To produce oligonucleotides containing small, site-specific photoreactive moieties, ONB photocleavable backbones were incorporated at the interbase linkages of oligonucleotide chains to perform light-induced degradation and phototriggered hybridization at near-UV wavelengths.[49] This method utilized solid-phase synthesis with modified phosphoramidites to determine the specific location of the caging group.[50] A similar strategy with a phosphoramidite containing an ONB-ether derivative demonstrated enhanced utility as a protecting group at the 2' hydroxyl group during RNA synthesis, with increased stability during post-synthesis deprotection, while still maintaining the photoresponsive properties of the caged RNA oligonucleotide.[51] Moreover, photocleavable linkers within the sequence of a single stranded oligonucleotide have been applied to light-induced degradation for OFF → ON photoswitching in live animals.[52] The site-specific incorporation of an ONB caging group into oligonucleotides using this method was also successfully demonstrated with caging of the 2' sugar position in RNA, and utilized to photochemically control catalytic activity of the hammerhead ribozyme (Figure 1.7).[34]
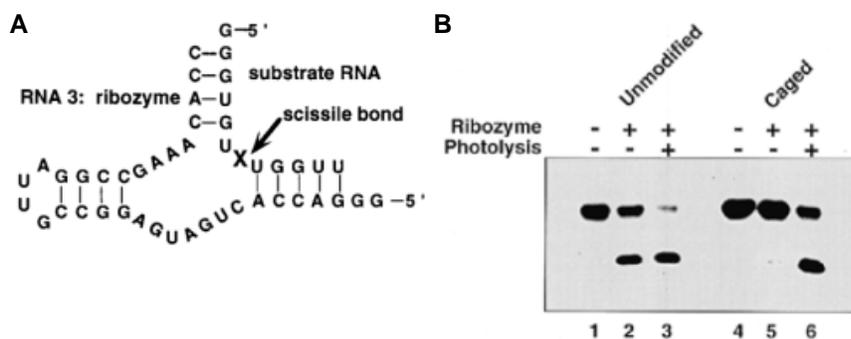
Figure 1.7: Photocontrol of a ribozyme reaction with caged RNA.
A) Canonical secondary structure of the hammerhead ribozyme substrate complex. The caging group was synthetically installed onto the 2' sugar position at location X. B) Denaturing PAGE analysis of ^{32}P-end-labeled unmodified and caged RNA substrates. Image adapted from Chaulk

and MacMillan, *Nucl. Acids Res.* **1998***, 26(13): 3173-3178 by permission of Oxford University Press.

In this study, the photocaged modification was incorporated site-specifically onto a hydroxyl group at the 2' sugar position of a RNA residue responsible for ribozyme cleavage (Figure 1.7A). Through *in vitro* analysis, the caged RNA substrate only showed cleavage after UV photolysis of the caging group (Figure 1.7B). These results demonstrated that the reactive 2' hydroxyl functionality in a RNA ribozyme could be caged through sugar modification to initiate catalytic cleavage upon photolysis. However, the methods described thus far all lack the ability to disrupt the base pair hybridization at nucleobase sites, which are most often associated with the oligonucleotide "information storage" that ultimately confers activity in gene regulation.

The most widely discussed modifications in this thesis are nucleobase caging groups (Figure 1.4C),[30] in which the sites responsible for hydrogen bonding to other bases are blocked to sterically restrict duplex formation and hybridization of the oligonucleotide complement.[53] Previous reports demonstrated that nucleobases could be modified with ONB derivatives using modified phosphoramidites for nucleobase protection and photolysis deprotection during solid-phase synthesis.[54] However, the utility of these photoremovable protecting groups in biological applications was not immediately realized. The first biological application for the incorporation of a caged nucleoside was achieved with a modified phosphoramidite in the synthesis of a single stranded oligonucleotide containing thiolether linkages at an adenosine residue, for the photochemical activation of DNA enzymatic function, although this caging site was not utilized for the inhibition of base pairing.[55] This caging group unfortunately required activation wavelengths between 250-280 nm, thus representing limited utility in the interrogation of true biological systems, which rely on nondamaging UV-A (or longer) wavelengths for light activation. Therefore, several groups expanded upon the approach, applying nucleobase caging

13

technology to a number of oligonucleotide structures for photochemical activation at near-UV wavelengths. The incorporation of nucleobase caging group on a thymine residue produced a temporary mismatch that enabled the photoactivation of transcription using a caged promoter region in a DNA template.[56] A similar caged thymidine residue was utilized in the optochemcial regulation of a DNA-based aptamer targeting thrombin (Figure 1.8).
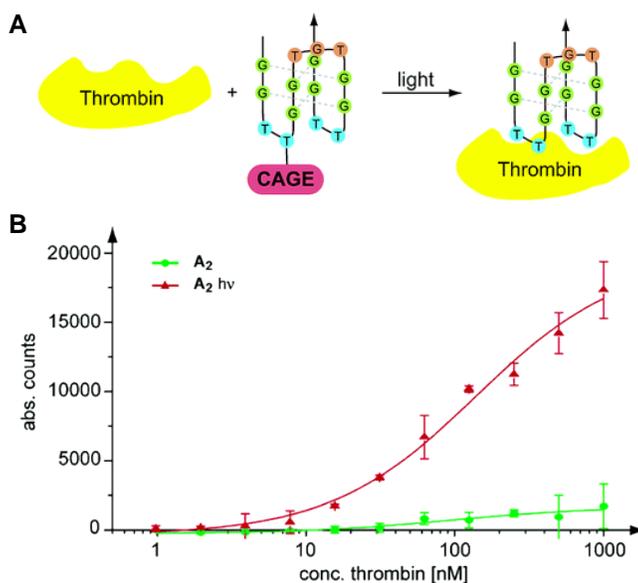
Figure 1.8: Light regulation of aptamer activity.
A) The G-qaudruplex DNA-based thrombin aptamer was caged at a critical thymidine residue to inhibit binding to thrombin in the absence of UV light. B) Binding assay of the caged aptamer (A_2) to thrombin, with and without irradiation. Image adapted with permission from Heckel and Mayer, *J. Am. Chem. Soc.* **2005**, 127(3):822-823. Copyright 2005 American Chemical Society.

The thrombin aptamer was modified with a synthetically installed caged thymidine residue, which did not interrupt the G-quadruplex secondary structure formation (Figure 1.8A). However, the site-specific caging group was placed in a location that interrupted binding of the aptamer to the thrombin protein substrate. In the absence of UV irradiation, no binding between the aptamer and thrombin could be detected (Figure 1.8B). However, light irradiation restored aptamer recognition of the protein, demonstrating that aptamer function could be

photochemically activated with caged nucleobases. Additionally, the functional inhibition of thrombin-dependent blood clotting was demonstrated with a human plasma clotting assay, showing that the optical control extended beyond simple binding assays.

In conjunction with the caged thymine, a nucleobase caged guanine residue was used to control G-quadruplex formation with light, expanding the number of reported caged DNA residues.[57] Similarly, nucleobase caging of guanine within an RNA oligonucleotide was used to study folding kinetic of RNA-RNA interactions.[58] In a rather expansive effort, all four naturally occurring RNA nucleobases were photoprotected and applied to the light regulation of RNA folding and tertiary structure formation.[59] In order to add an additional layer of functionality into the nucleobase caging approach, a photocleavable quencher moiety enabled not only disruption of duplex formation, but also the ability to monitor decaging through an adjacent fluorophore.[60] While these initial nucleobase caging approaches had some drawbacks, including specialized synthesis condition requirements or relatively poor decaging kinetics, they paved the way for more advanced technologies to be developed. For example, the use of nucleobase caging was subsequently applied to siRNA molecules for optical regulation of RNA interference (Figure 1.9).[40]
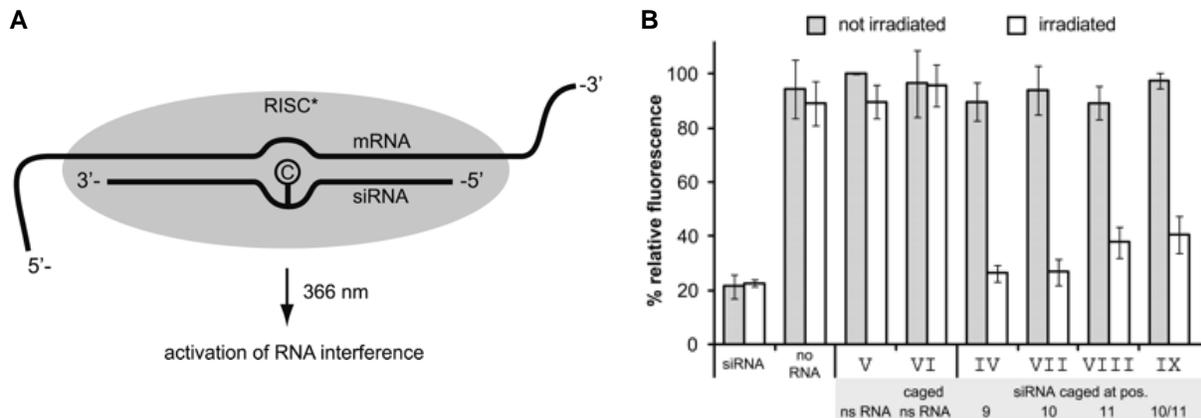
Figure 1.9: Light-activated RNA interference with nucleobase-caged siRNAs.
A) Nucleobase caging groups (©) were incorporated into an siRNA strand targeting GFP. B) The relative GFP expression was determined for a series of modified siRNA oligonucleotides. Image adapted with permission from Mikat and Heckel, *RNA*. **2007**, 13:2341-2347. Copyright 2007 RNA Society

Here, the nucleobase caged guanine and thymidine residues were synthetically installed onto siRNA strands targeting GFP. These caging groups did not inhibit complete hybridization between the siRNA and mRNA duplex, but rather introduced temporary mismatches to create a "bulge" in the duplex recognition site (Figure 1.9A). These siRNA oligonucleotides were transfected into cells expressing the GFP and RFP reporter genes, and UV irradiated or kept in the dark. The GFP expression levels were then normalized to the RFP expression (Figure 1.9B). The nucleobase caged siRNA reagents were not active in the absence of irradiation, and no inhibition of GFP expression was observed. However, after UV irradiation the siRNAs containing caging groups at the mRNA cleavage site were activated, and reduced GFP expression was observed. These results demonstrated that the nucleobase caging approach could be applied to optical control of RNAi with site-specific modified residues.

Since those initial developments, many new nucleobase caging groups incorporated into oligonucleotides through synthetic means (Figure 1.10). This general approach allowed the photosensitive modification to be placed at any base residue of interest through the incorporation

of synthetic modified phosphoramidites during solid-phase oligonucleotide synthesis.[61] Since the nucleobase caging group can be applied to virtually any DNA or RNA base at defined residues, the study of very specific base-base interactions was made possible.
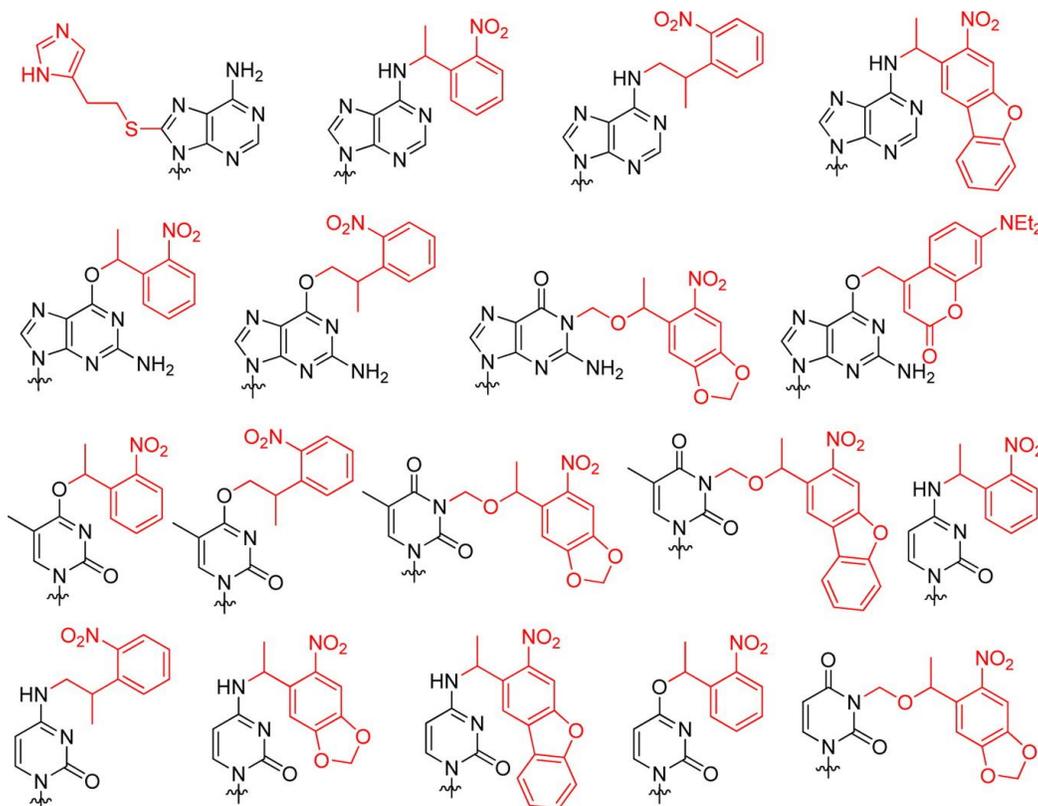
Figure 1.10: Chemical structures of caged nucleobases incorporated into oligonucleotides. The light-removable caging groups are shown in red. Image adapted with permission from Liu and Deiters, *Acc. Chem. Res.* **2014**, 47(1):45-55. Copyright 2013 American Chemical Society.

Oligonucleotides containing caged nucleobases have been synthesized as standard DNA, phosphorothiotate (PS) backbone modified, and 2'-O-methyl (2'OMe) nucleobase modified compositions. The application of photochemically modified oligonucleotides will be further described in Chapter 2.0 for the regulation of gene function and in Chapter 3.0 for the regulation of DNA computation.

1.1.1 Synthesis of Modified Oligonucleotides.

Oligonucleotide synthesis was performed using standard β-cyanoethyl phosphoramidite chemistry on solid-phase supports.[50] A refurbished Applied Biosystems Model 394 automated DNA/RNA synthesizer (Life Technologies) and a MerMade4 synthesizer (Bioautomation) were used for all oligonucleotide manufacturing. All oligonucleotide synthesis reagents were obtained from Glen Research. The synthesis cycle (Figure 1.11) is based on a reactive 3' phosphorous coupled with a 5' hydroxide to polymerize phosphoramidites into nucleotide chains. Synthesis was performed in the 3' → 5' direction with the first base immobilized on CPG resin column at the 3' sugar position. The initiating synthesis step was the removal of a 5' DMTr protecting group and formation of a reactive 5' OH group through treatment with TCA. This was performed in several successive reactions since the DMTr cation is highly reactive and can re-tritylate any reactive nucleophiles. The removal of the DMTr cation can be observed and quantified by the absorbance at 498 nm, a process known as trityl monitoring. The incoming phosphoramidite was prepared for coupling through activation with tetrazole, forming a tetrazoyl phosphoramidite. This reactive species coupled with the 5' OH of the previous base and formed a 5' → 3' phosphorous linkage. A capping step was performed to acetylate any remaining 5' OH groups in order to truncate any failed coupling reactions. Acetic anhydride and N-methyl imidazole were combined to form a strong acetylating agent that capped failed sequences, preventing further coupling with activated phosphoramidites and terminating the oligonucleotide chain elongation. The phosphite triester linkage of successful coupling reactions was then oxidized through treatment with an I_2 / H_2O / THF / pyridine mixture to form a more stable phosphate triester. This final step of a single base addition cycle leaves a 5' DMTr protecting group on the newly

18

added nucleotide, which can be coupled with incoming activated phosphoramidites through additional cycles to complete the oligonucleotide chain. Once all nucleotides have been added and the final 5' DMTr group has been removed through TCA treatment, the oligonucleotide was cleaved from the resin with concentrated ammonium hydroxide/methyl amine (AMA, 1:1) leaving a 3' hydroxide group. Deprotection of the cyanoethyl phosphate protecting groups and exocyclic amines on certain bases (A, C, and G) was also performed with the AMA treatment at 65 °C for 2 hr, forming phosphodiester linkages and nucleobase structures that are naturally seen in native DNA. Subsequent purification was performed with sephadex columns utilizing size exclusion.[62] UV absorption at 260 nm was used to calculate molar concentrations from the oligonucleotide extinction coefficient (IDT OligoAnalyzer 3.1). Polyacrylamide gel (PAGE) analysis was used to confirm the size and purity of the final product. In the cases where multiple bands were observed, truncated oligonucleotides were removed from the final product through band excision and elution of the full-length oligonucleotide. Further analysis of the oligonucleotide can be conducted with HPLC and mass spectrometry methods.[63]
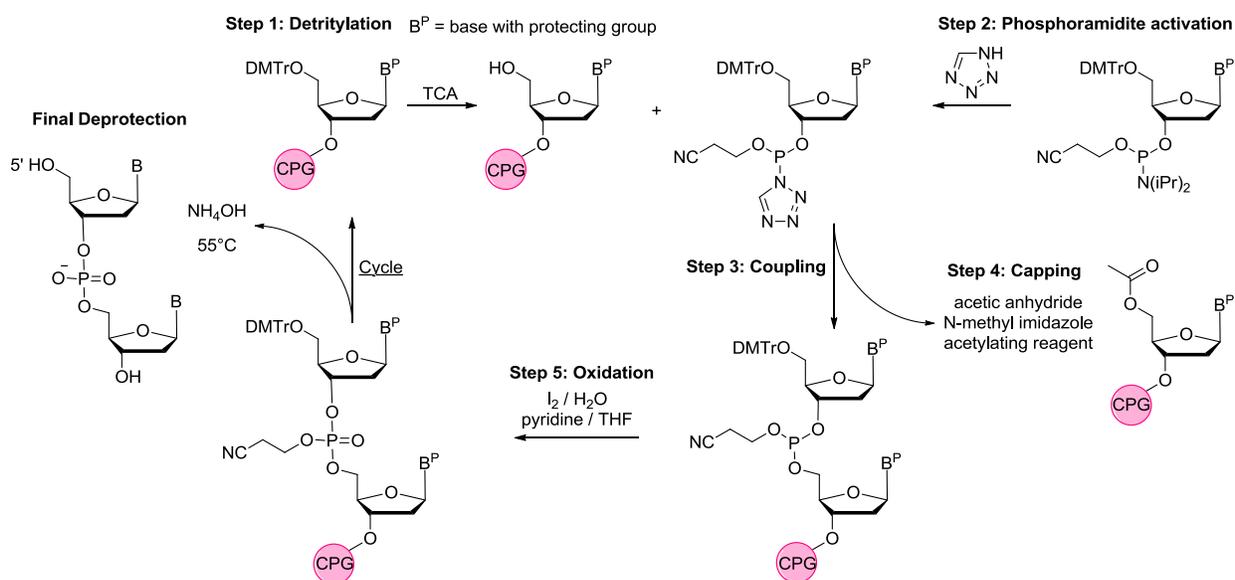
Figure 1.11: DNA synthesis cycle with phosphoramidites.

A number of caged DNA phosphoramidites have been synthesized in the Deiters lab and are incorporated into oligonucleotides through the same synthesis cycle as described above (Figure 1.12). These monomers contain the ONB derivatives *ortho*-nitrophenylethyl (NPE) or 6-nitropiperonyloxymethyl (NPOM), and were designed to remain stable under standard oligonucleotide synthesis conditions, while exhibiting efficient decaging at UV wavelengths. When placed on exocyclic amines, the caging modifications also act as protecting groups during oligonucleotide synthesis, with functional groups that are orthogonal to the synthesis conditions.
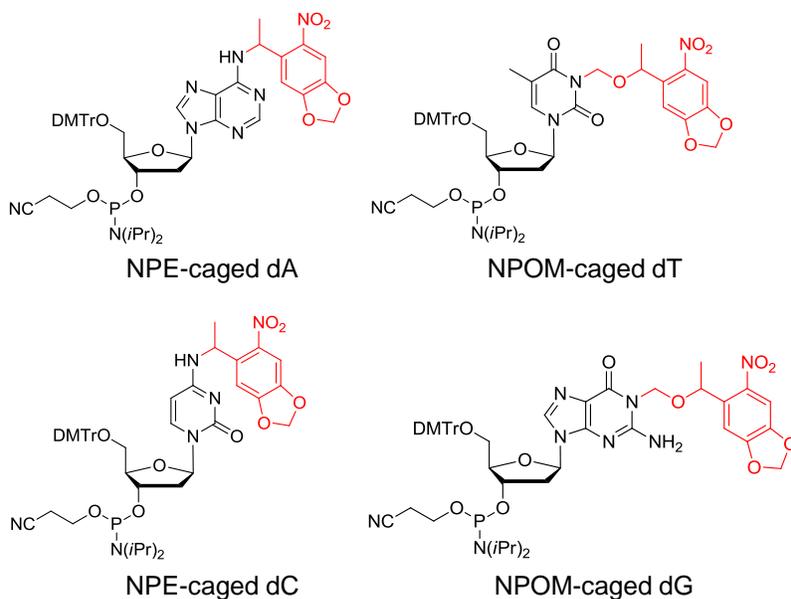
Figure 1.12: Caged DNA phosphoramidites synthesized in the Deiters lab. Caging groups have been applied to phosphoramidites for all four native DNA bases (A, T, C, and G). NPE and NPOM caging groups are indicated in red.

Modifications such as 2'OMe sugar rings have been shown to increase affinity for the DNA or RNA targets, as well as cellular stability of the oligonucleotide.[64] 2'OMe caged phosphoramidites have been synthesized in the Deiters lab, as well as phosphoramidites that employ light-cleavable groups at positions that are not involved in base hybridization (Figure 1.13). Phosphorothioate (PS) linkages were also of interest due to the increased stability and applications in molecular biology, such as siRNA or antisense oligonucleotides.[65] The sulfurization of nucleotide backbone linkages was performed using a standard ABI synthesis cycle and the Beaucage sulfurizing reagent.[66]
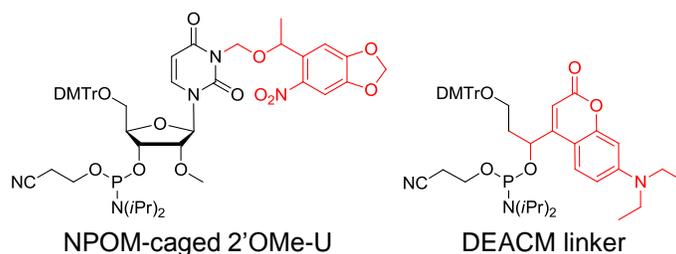
Figure 1.13: Modified caged phosphoramidites synthesized in the Deiters lab. Caging groups and photocleavable linkers have been installed on specialized phosphoramidites. Photolabile groups are indicated in red.

Table 5.1 includes the sequence information on all ~110 oligonucleotides that have been successfully synthesized along with the corresponding application. More detailed information regarding synthesis conditions can be found in Section 5.2.

1.2   Optical Control of Protein Function

Several methods for the generation of light-activated proteins exist.[67] Initially, optical induction of protein function and gene expression was demonstrated with caged small molecule effectors that interact with receptors to control activation of protein function with light.[68] These systems showed good OFF → ON switching behavior, with low activity in the absence of light and high levels of gene expression after irradiation, although diffusion of the activated small molecule effector limited the timeframe in which the function could be observed. Similar methods were applied to induction of the Tet-ON system with a caged doxycycline molecule, which showed some delay in light activation response and also experienced issues with diffusion of the activated compound.[69] However, the issue of diffusion was subsequently overcome by linking the ligand and receptor with a photoswitchable domain, which can be turned ON or OFF with

22

specific wavelengths of light without complete release of the ligand substrate.[70] Caged small molecules were also developed for the inhibition of gene expression using a photochemically caged ribosomal inhibitor to turn OFF translation with light.[71] While these methods all involve the caging of biologically active small molecules that target protein function, the installation of photochemically responsive groups onto the protein of interest is an alternative method that allows for more efficient photoregulation (Figure 1.14). Light-activation induces changes in structural perturbations, conformations, protein-protein interactions, or active site accessibility.
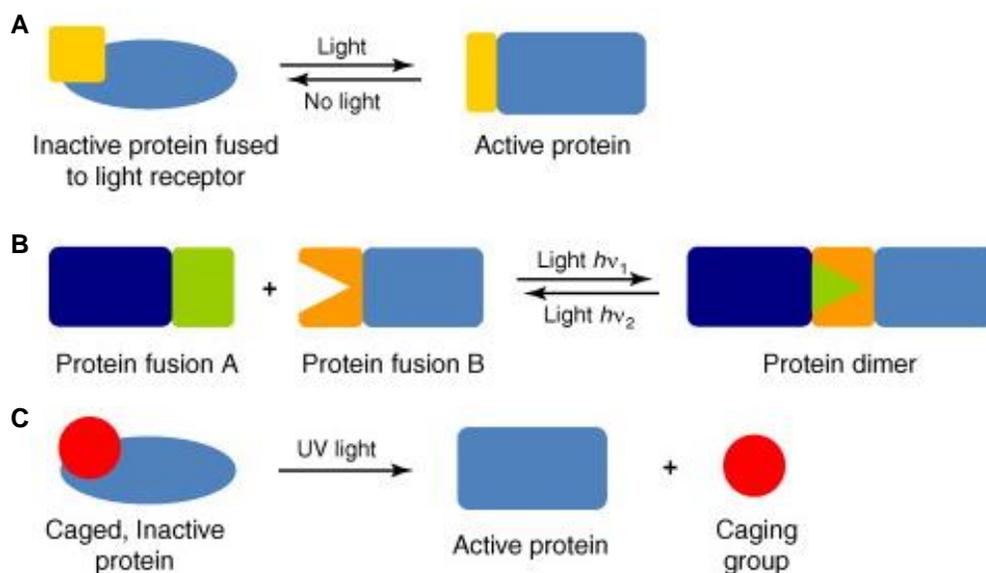
Figure 1.14: Select approaches to directly control protein function with light.
A) Protein activity can be reversibly regulated by light irradiation when fused to a light receptor (e.g., light-oxygen-voltage domain). B) Two proteins can be dimerized by light irradiation when fused to natural photoreceptors (e.g., phytochromes). D) Proteins containing caged amino acids can be activated via irradiation. Proteins are indicated in blue, caging groups are indicated in red, and light-responsive fusion domains are indicated in orange/green. Adapted from *Trends Biotechnol.*, 28(9), Riggsbee and Deiters, **Recent advances in the photochemical control of protein function**, 468-475, Copyright (2010) with permission from Elsevier.

Among the methods for direct protein modification, light-oxygen-voltage (LOV) domains,[72] phytochromes (Phy),[73] and cryptochromes (Cry)[74] have all been utilized to efficiently control proteins with light.[75] These optogenetic approaches rely on the engineering of constructs

with a naturally occurring light-sensitive protein from plants or microbial organisms fused to the

target protein of interest in order to render it light-responsive.[76] The fusion proteins can then be

photoswitched to control protein dimerization and protein conformation, enabling optical control

of protein function (Figure 1.15).[77]
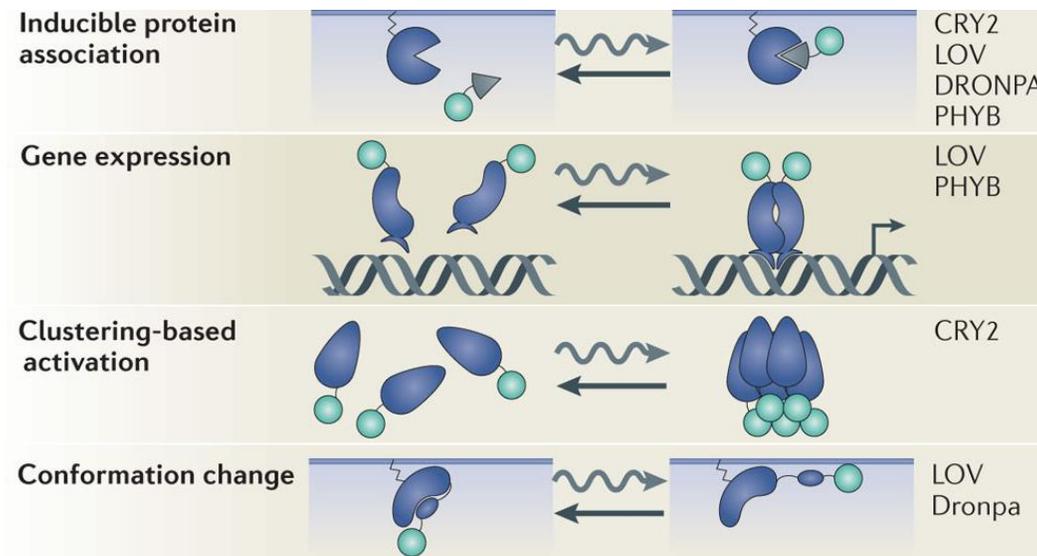
Figure 1.15: Controlling protein function with optogenetic tools.
The photosensitive proteins that have been used for different regulatory approaches are listed to the right. LOV, light-oxygen-voltage; PHYB, phytochrome B; CRY2, cryptochrome 2; DRONPA, photoswitchable fluorescent protein. The light-stimulated protein reverts back to its original state either in the dark or through a second light exposure. Photosensitive proteins are indicated in navy blue, binding partners are indicated as gray triangles, and protein regulation targets are shown as teal circles. Curvy arrows indicate the response of the system to light, whereas straight arrows indicate dark- or light-stimulated reversion. Image adapted with permission from MacMillan Publishers Ltd: *Nature Rev. Mol. Cell Bio.,* 15:551-558, copyright 2014.

As shown in the optogenetics illustration above, heterodimerization can be used to recruit

a protein domain to a substrate for inducible protein association. To initiate gene expression,

heterodimerization or homodimerization techniques can be used to recruit transcriptional

activators or other DNA-modifying proteins. Fused protein domains can be used to form natural

clusters that are dependent on domain density to activate function with light. Lastly,

conformational changes in the photosensitive protein can be used to expose concealed functional

domains or activate protein function. Adaptations of the yeast two-hybrid system have been modified with these light-sensitive fusion proteins to reversibly control gene expression with light in eukaryotes.[78]

As an example for the utilization of conformational changes induced through photostimulation of LOV domains, fusion constructs were developed for the manipulation of kinase activity with a genetically encoded photoactivatable phosphorylation inhibitor (Figure 1.16).[79] Protein phosphorylation by kinases is an essential mechanism of cell signaling and cellular regulation of a wide range of processes and plays very important roles in human pathology.[80] Optical interrogation of kinase function and cell signaling has advantages over the use of gene manipulation or small molecule inhibitors, which often provide poor spatio-temporal resolution and limited specificity for targets that utilize common cofactors. In this study, a highly selective naturally occurring inhibitory peptide (PKI)[81] of cyclic-AMP dependent kinase (PKA) was fused with a LOV domain to produce a photoactivated PKI (PA-PKI). This construct was inactive until irradiated with blue light, which enabled the PKI peptide to bind to PKA, inhibiting downstream phosphorylation events (Figure 1.16A). The LOV domain was linked to a carboxy-terminal helical extension (Jα)[82] that is unwound through conformational changes after irradiation, resulting in increased flexibility and accessibility of the PKI peptide (Figure 1.16B). In order to test the response of PA–PKI to light, the PKA-dependent phosphorylation of the transcription factor cAMP response element-binding protein (CREB)[83] was analyzed (Figure 1.16C). A reduction of CREB phosphorylation was observed through Western blot analysis when the PA-PKI was exposed to blue light, in contrast to no irradiation (Figure 1.16D). This study demonstrated that coupling genetically encoded photosensitive domains to inhibitory peptides enables photoswitchable endogenous kinase signaling, in a versatile approach to control

phosphorylation activity in living cells. The minimally invasive approach may allow for spatio-temporal precision in the study of kinase pathways that can be inhibited by peptides.
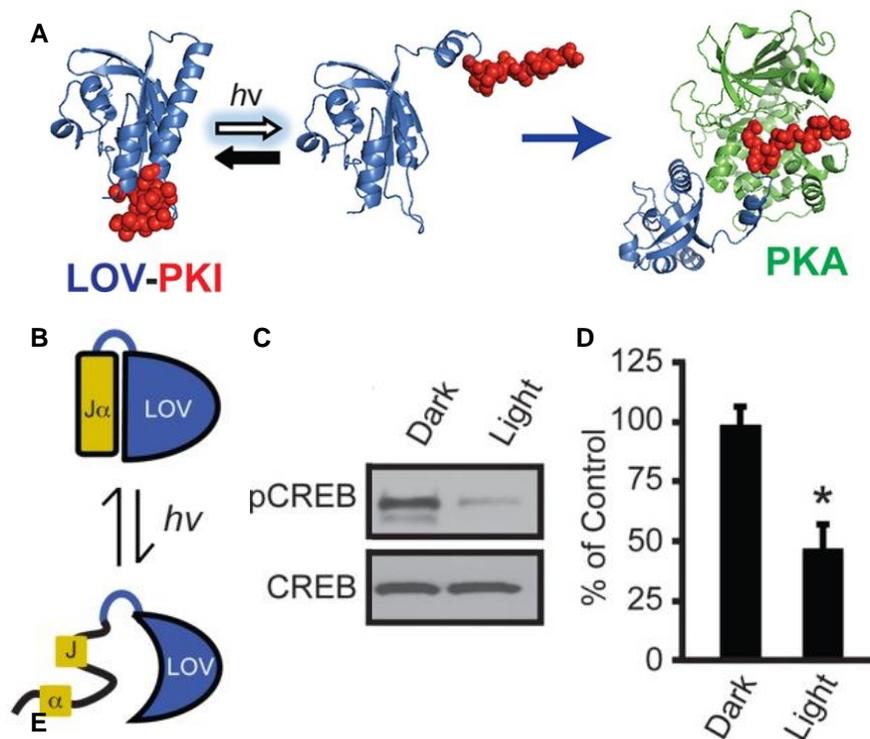
Figure 1.16: Optical control of kinase activity with photoswitchable inhibitory peptides. A) Schematic of the LOV fusion PA–PKI and light-induced binding to PKA. B) Schematic showing the light response of LOV. Blue represents the LOV globular domain, yellow represents the Jα helix. Upon irradiation with blue light, a conformational change in LOV causes the unwinding of Jα. C) Western blot of CREB phosphorylation (pCREB) with PA-PKI expression, both with and without blue light (460 nm). D) Quantification of C. Image adapted with permission from Yi et al, *ACS Synth Biol.* **2014**, 3(11):788-795. Copyright 2014 American Chemical Society.

As an alternative approach, light-induced protein dimerization has been used to control cell signaling. For example, the Phy/PIF dimerization can be adapted for conditional control of Rac1 function through translocation of an activator protein to the cell membrane (Figure 1.17).[84] The Rac1 protein is a GTPase that regulates actin cytoskeletal dynamics during cell migration.[85] The use of phytochromes for gene regulation relies on several factors to detect red or near-infrared light and produce photoisomerization: a chromophore (phycocyanobilin, PCB), a

phytochrome (PhyB), and a phytochrome interaction factor (PIF3).[86] In this study, Rac1 was activated by the catalytic Tiam-DHPH domain module after membrane recruitment[87] and production of lamellipodial cell protrusions (Figure 1.17A). The PhyB and PIF3 photosensitive proteins were fused to constructs expressing the reporter proteins mCherry and YFP for cellular tracking, as well as the Tiam-DPDH (Figure 1.17B). Photochemical recruitment of the PIF-Tiam-DHPH chimera, in conjunction with PhyB and PCB, caused a pronounced lamellipodial phenotype within 20 min in ~80% of cells (Figure 1.17B). The spatially localized light activation of Rac1 cell protrusion was then tested. Red laser stimulation was used for localized recruitment of the PIF-Tiam-DHPH domain in serum-depleted cells, producing a localized lamellipodial 'bloom' within 10 min (Figure 1.17C). By slowly extending the point of activating light away from the cell, the ability to draw out cellular morphology up to 30 μm from the main body of the cell was demonstrated. The high spatial and temporal resolution of light control allows this module to function as an analytical tool to pattern cellular behavior. In fact, the optical regulation of Rac1 was used to demonstrate that GTPase activation occurs within seconds, suggesting that downstream signaling steps were responsible for the 5-10 min delay in lamellipodial protrusion formation. This study showcases the potential for programming specific cell geometries and intercellular connections with light using the Phy/PIF system to sculpt cell shape in an unprecedented manner. An advantage of the Phy/PIF system is the ability to reversibly photoswitch between active (red) and inactive (far-red) states using light.

27
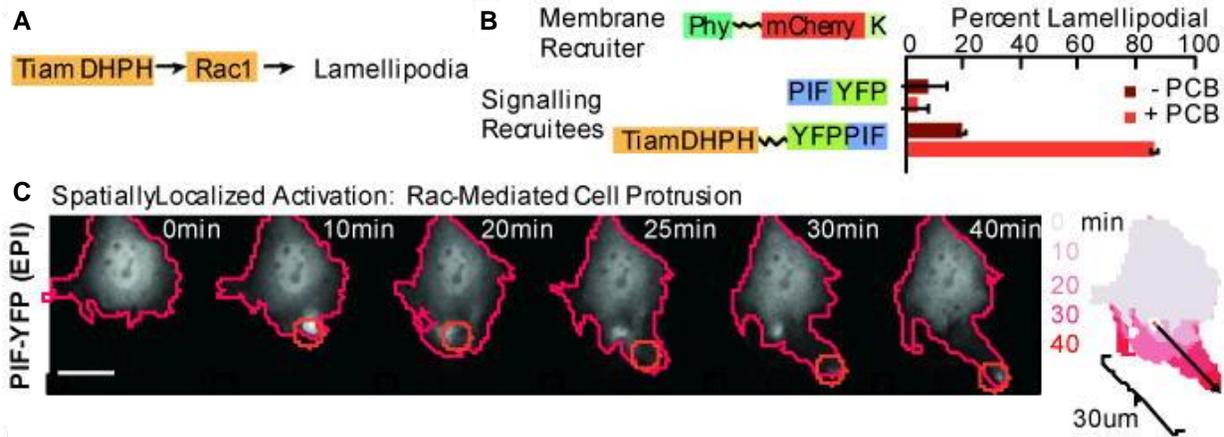
Figure 1.17: Optical control of cell signaling using a light-switchable protein dimerization. A) The catalytic Tiam-DHPH domain activates the respective G-protein Rac1, which in turn acts through effector proteins to modify the actin cytoskeleton. B) Constructs with the PIF-Tiam-DHPH and Phy domains were assayed for their ability to induce lamellipodia in serum-depleted cells with red light irradiation (650 nm). The percentage of cells that produced lamellipodia within 20 min was determined with live microscopy. C) Local induction and extrusion of lamellipodia in live cells was demonstrated by focusing red light (650 nm) onto a small portion of the cell and slowly extending the light away from the cell body. Superimposed outlines of the cell show directed extension 30 μm along the line of light movement. Image adapted with permission from MacMillan Publishers Ltd: *Nature,* 461: 991-1001, copyright 2014.

A significant drawback of the Phy/PIF system is the requirement to supplement the cells with PCB. The Cry domains represent an alternative approach to the optical control of protein dimerization that relies on an endogenous chromophore cofactor (flavin adenine dinucleotide, FAD), which does not require additional supplementation. For example, cryptochrome (Cry) induced dimerization of a split Cre recombinase and optical control of DNA recombination was demonstrated (Figure 1.18).[88] Cre recombinase is an extensively used enzyme in genetic engineering that recombines DNA at specific loxP sites.[89] The protein can be split into two nonfunctional domains to activate DNA recombination upon dimerization and reconstitution.[90] The Cry2 domain binds to the cryptochrome-interacting basic-helix-loop-helix (CIBN) domain when photoexcited with blue light.[91] In this study, the split Cre domains were fused to the Cry2 and CIBN domains to optically control Cre recombination (Figure 1.18A). Additionally, a plasmid containing a transcriptional stop sequence flanked by loxP sites preceding EGFP was

used as a reporter for Cre recombinase activity (Figure 1.18A). This reporter expresses EGFP when the Cre recombinase is dimerized and activated. The Cre expression and reporter constructs were transfected into cells, and irradiated with pulsed blue light (450 nm), ranging up to 24 h (Figure 1.18B). In the absence of blue light activation, no recombination resulting in the expression of EGFP was observed. Cells containing both Cry2-CreN and CIBN-CreC that were exposed to pulsed blue light showed increased EGFP expression, with the most significant increase in recombinase activity was observed with 24 h irradiation (Figure 1.18C). Since this system is entirely genetically encoded and can be activated using common light sources, these modules are useful for optically controlling a broad range of biological phenomena.[92]
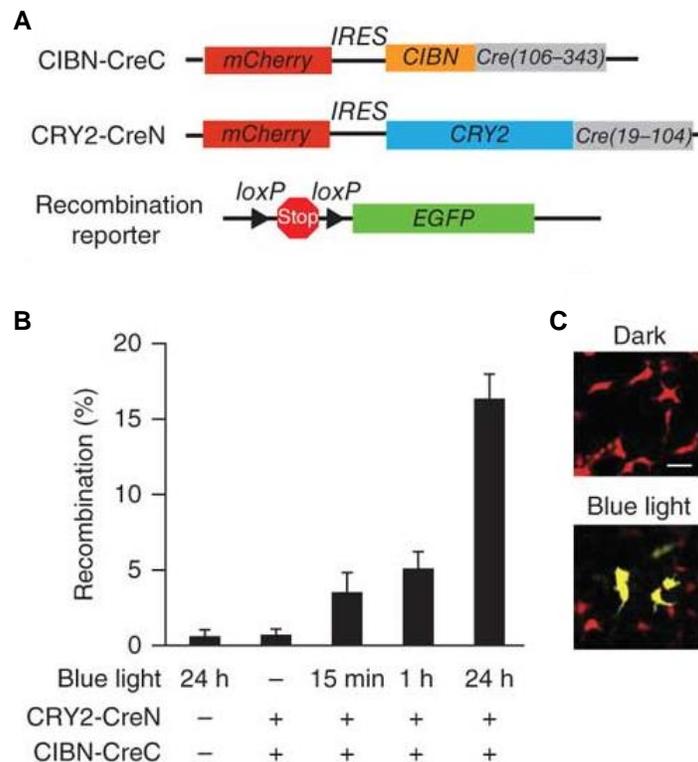
Figure 1.18: Optical control of DNA recombination using a blue-light protein dimerization.
A) Schematic showing the two split Cre recombinase constructs (CIBN-CreC and CRY2-CreN) and the EGFP reporter construct. B) Cre reporter recombination was measured 48 h after cellular transfection with the Cre reporter and CRY2/CIBN constructs. Cells were exposed to blue light pulses (450 nm) for the indicated durations or kept in the dark. C) Fluorescent images from cells

However, the engineering of light-responsive fusion constructs with the protein of interest requires significant trial-and-error on a case-by-case basis, since the target protein must be split into two nonfunctional fragments and then specifically fused to the photoresponsive domains to optically control dimerization. These fusion domains can also interfere with the native protein activity, as large modifications can encumber protein function in some biological applications. Therefore, the use of smaller site-specific caging modifications was developed through incorporation of caged amino acids into proteins. Through this method, the optical regulation achieved with a single caging group modification does not encumber the native protein with fusion constructs to light-responsive protein domains that are several orders of magnitude larger (from 20 to >120 kDa) than a caging group (0.24 kDa).

The addition of modified amino acid residues (e.g., photocaged) for the incorporation of new functionality into proteins has been accomplished a number of ways: chemical modification of amino acids, solid-phase synthesis of proteins containing modified amino acids, and *in vivo* incorporation of modified amino acids (Figure 1.19). These methodologies have been applied to the modification of proteins with amino acids containing site-specific caging groups.[93] An advantage of using synthetic caging groups is the ability to tune the photochemical properties of caged proteins, such that they respond to UV irradiation, avoid interference by ambient light, and allow for the simultaneous application of the full palette of fluorescent protein reporters, while providing clean OFF → ON light switching.
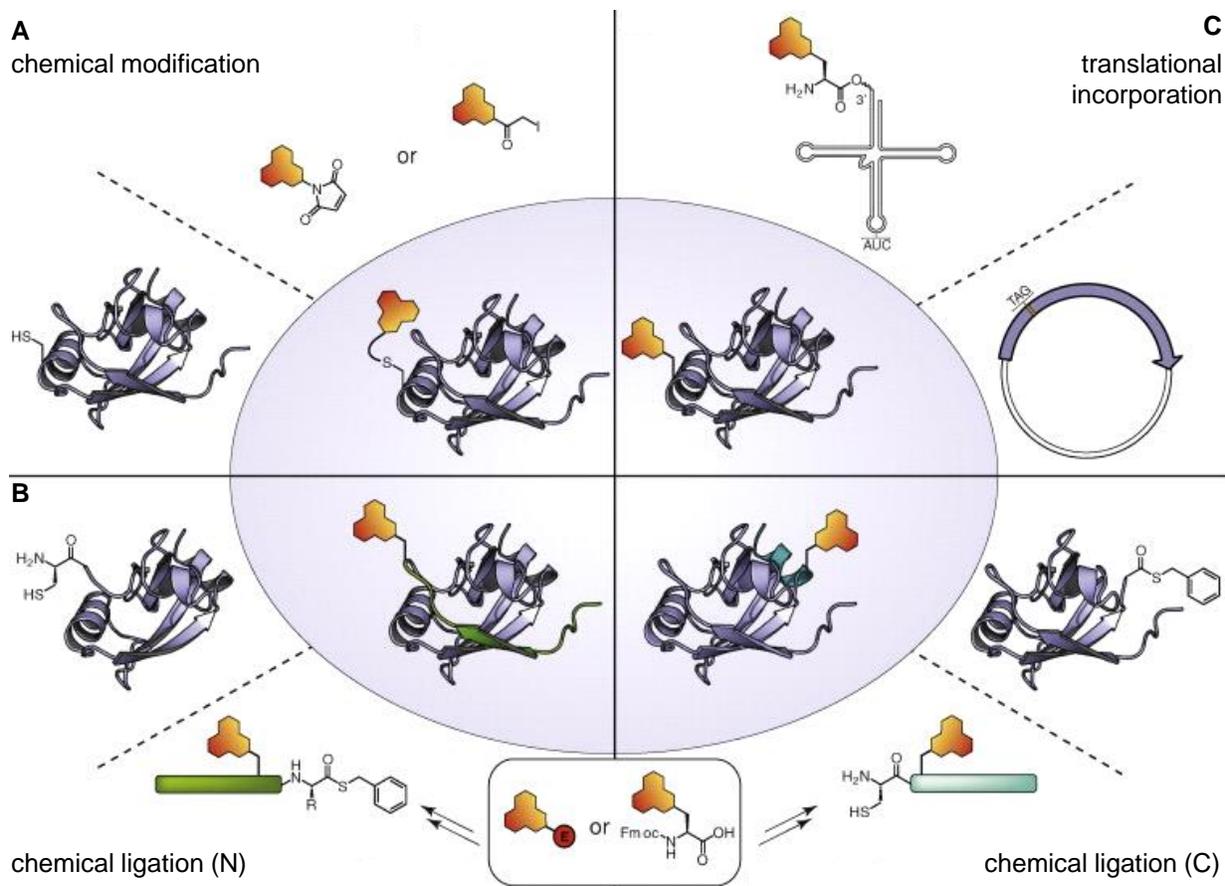
Figure 1.19: Methods for incorporation of modified amino acids into proteins.
A) Chemical modification of a solvent exposed amino acid (e.g., cysteine) with a thiol-reactive agent. B) Chemical ligation between an expressed protein and a synthetic peptide, which is amended to the N-terminus (left) or the C-terminus (right) of the protein. C) Incorporation of a modified amino acid through suppression of an amber stop codon during translation. A tRNA molecule designed to recognize and read through the amber codon is charged with the modified amino acid, either synthetically or endogenously with a mutant tRNA synthetase. Adapted from *Trends Biotechnol*., 28(2), Loving et al, **Monitoring protein interactions and dynamics with solvatochromic fluorophores**, 73-83, Copyright (2010) with permission from Elsevier.

The generation of caged peptides or proteins can be achieved through synthetic chemical methodologies and introduced into biological systems for the optical control of biological processes.[94] Traditionally, these protein-caging modifications have been introduced in a nonspecific manner, through *in vitro* modification and downstream purification of the photoresponsive protein. General chemical modifications were performed using functional groups of the protein backbone or amino acid side chains for reactions to introduce

photoresponsive moieties onto a protein of interest (Figure 1.19A).[95] For example, the caging of

the cofilin protein was achieved through chemical modification of an amino acid side chain, in

which a cysteine residue was altered with an ONB derivative (Figure 1.20A).[96] The cofilin

protein is responsible for modulating actin dynamics through the depolymerization of F-actin,[97]

and is regulated by LIM kinase phosphorylation of a serine residue that inactivates cofilin

function.[98] However, a constitutively active cofilin containing a cysteine mutation maintains a

high rate of F-actin cleavage and cannot be phosphorylated by LIM kinase.[99] It was shown that

chemical conversion of the cysteine on the constitutively active cofilin produced a caged cofilin,

which was unable to bind to F-actin in the absence of photolysis. Light irradiation restored

cofilin activity and F-actin cleavage. In this study, the chemical modification of a protein was

used to optically control the function of a kinase target. While these methods were relatively

simple to employ, they were somewhat limited in the ability to control the location specificity of

the modification, required that the amino acids be surface exposed (i.e., internal sites buried

within a protein structure are not accessible for chemical modification), and could only be

applied to specific amino acid residues (e.g., Cys and Lys). To overcome challenges regarding

nonspecific protein modification through chemical means, the use of solid-phase peptide

synthesis was enabled for the manufacturing of caged peptides containing site-specific caging

group modifications.[100] However, these approaches for the production of caged proteins were

limited by the constraints of peptide synthesis, allowing only peptides of a specific size and with

a limited number of caged amino acid incorporations to be generated. Larger peptide chains have

since been produced through chemical ligation methods, enabling the photochemical caging of

full proteins through synthetic means (Figure 1.19B).[101] For example, the caging of Smad2 was

achieved through chemical ligation, in which a thioester modification on an expressed protein

was linked to a synthetic peptide containing caged phosphate groups (Figure 1.20B).[102] Smad2 is a kinase that accumulates in the nucleus to regulate transcription after phosphorylation, which releases Smad2 from the Smad anchor for receptor activation (SARA), and subsequently forms a Smad2 homotrimer mediated by the MH2 domain.[103] The caging strategy for Smad2 involved the chemical ligation of an expressed Smad2-MH2 domain modified with a thioester[104] and complexed to the SARA binding domain (SARA-SBD) with a synthetic caged phosphopeptide[105] through expressed protein ligation.[106] The resulting caged Smad2 complex contained two caged phosphate groups, which could be removed with UV irradiation to activate the Smad2 homotrimerization and downstream biological activity (Figure 1.20B). The optically controlled behavior of caged Smad2-MH2 was determined with a nuclear import assay to study the kinetics of biological signaling and transport processes in live cells. Labeled Smad2 complexes were delivered to cells, and nuclear accumulation was analyzed through cellular imaging (Figure 1.20C). In the presence of the SARA-SBD, nonphosphorylated Smad2-MH2 (OP) was excluded from the nucleus, whereas diphosphorylated Smad2-MH2 (2P) accumulated in the nucleus. The caged Smad2-MH2 was excluded from the nucleus in the absence of irradiation, whereas decaging of the protein with light led to near complete nuclear accumulation. In this study, a caged Smad2-MH2 protein was prepared by a semisynthetic route utilizing chemical ligation in the production of a phosphate caged protein and optical control of phosphorylation-dependent function.
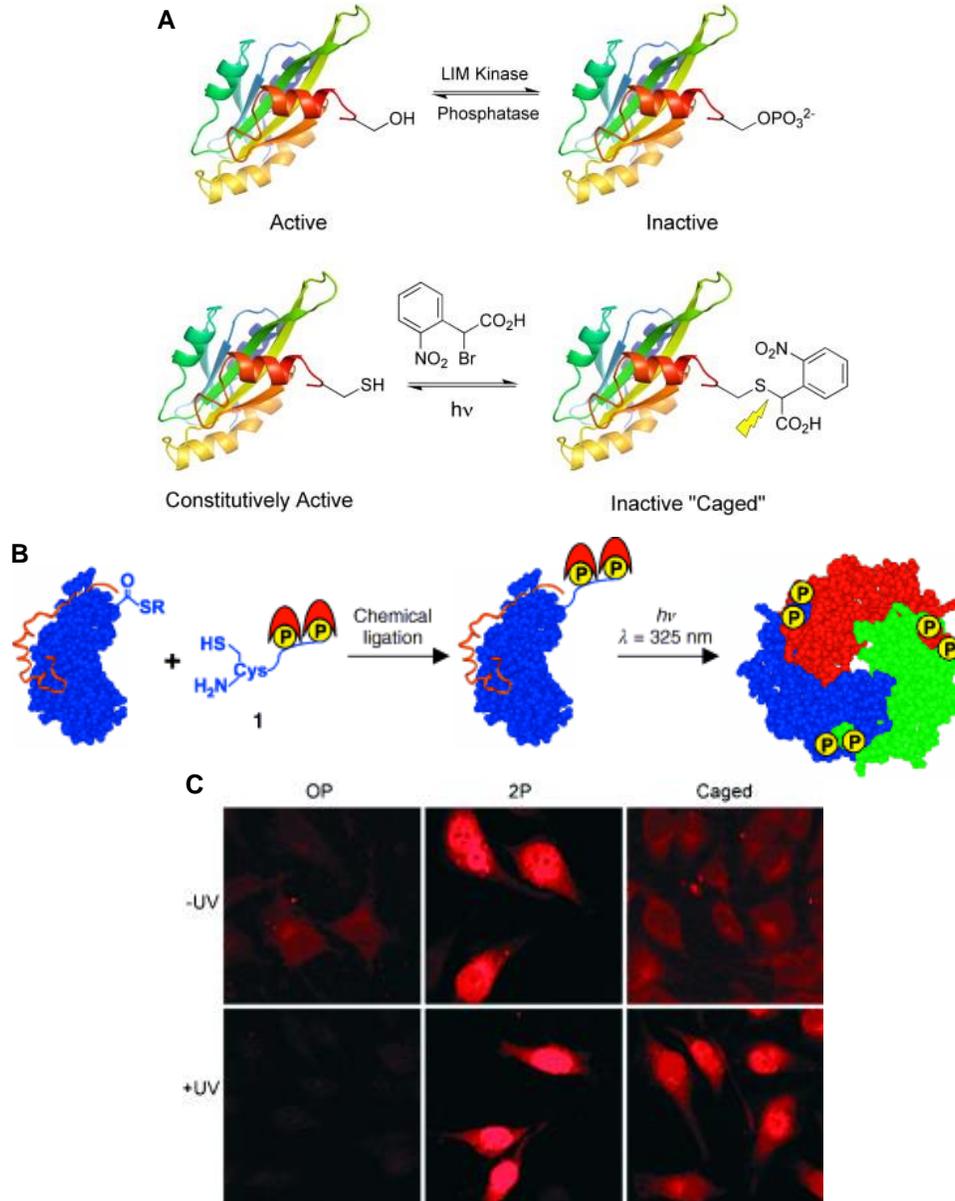
Figure 1.20: Chemical modification and ligation strategies for photocaging proteins.
A) The design and synthesis of caged cofilin. LIM kinase phosphorylates a serine residue on native cofilin. However, when a constitutively active cysteine mutant is chemically modified with an ONB derivative, the caged cofilin is inactive until UV irradiated. Adapted from *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1804(3), Priestman and Lawrence, **Light-mediated remote control of signaling pathways**, 547-558, Copyright (2010) with permission from Elsevier. B) Semisynthesis of caged Smad2-MH2. Expressed protein ligation was used to ligate a recombinant Smad2-MH2-α-thioester/SARA-SBD protein complex to a doubly caged phosphopeptide to produce the caged Smad2-MH2/SARA-SBD heterodimer. Caged Smad2-MH2 is activated by exposure to UV light and subsequently releases SARA-SBD to form a homotrimer. Smad2-MH2 is shown in globular form, SARA-SBD is shown in orange, phosphorylated residues are symbolized by yellow circles, and caging groups are symbolized by red crescents. C) Nuclear import assay of Smad2-MH2 variants. Nonphosphorylated (OP), doubly phosphorylated (2P), and caged Smad2-MH2 are shown before (top) and after (bottom) UV laser irradiation (20 min) in HeLa cells. Image adapted from **Photocontrol of Smad2 a Multiphosphorylated Cell-Signaling**

34

The methods described above for *in vitro* chemical production of caged proteins were all generally limited for biological applications due to purification requirements, low quantities generated, restricted bioorthogonal reactions, and the need for delivery (typically injection) into biological systems. Therefore, the use of genetically expressed modified proteins (i.e., caged proteins that are produced within the cell or organism of interest) provides additional value for biological applications in the field of optochemical control of protein function (Figure 1.19C).[107]

Cells can be equipped with an expanded genetic code that allows for the site-specific *in vivo* incorporation of a modified amino acid.[108, 109] The growing field of unnatural amino acid (UAA) mutagenesis has led to new approaches in developing methods to control protein activity with photochemical modifications in recognition of the amber codon TAG,[110] which is a low frequency stop codon. Mutations can be made in a gene of interest to introduce the amber codon for site-specific incorporation of the UAA. Initially, the chemical synthesis of misacylated tRNAs in conjunction with amber codon mutagenesis provided a method to site-specifically install caged amino acids onto proteins and photochemically control protein function,[111] although the chemical process of misacylation proved to be experimentally challenging. Therefore, orthogonal synthetase/tRNA pairs have been engineered to incorporate specific UAAs through cellular generation of amber-suppressor tRNAs that are selectively charged with a UAA by a corresponding tRNA synthetase (Figure 1.21).[112]
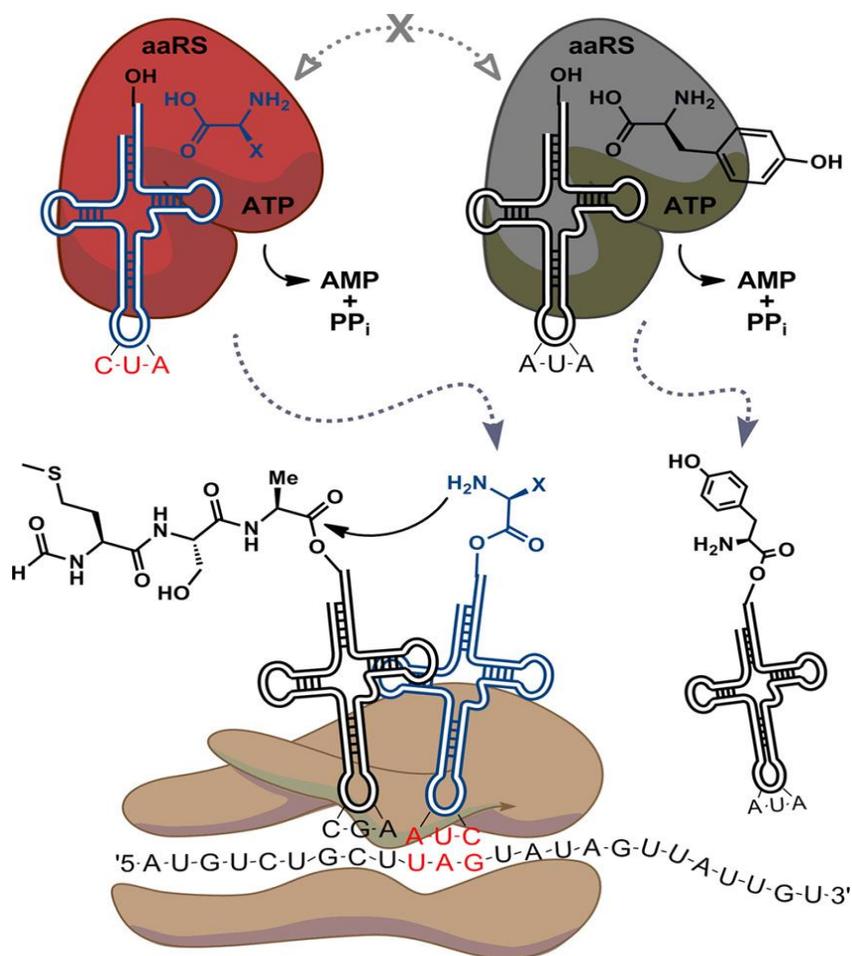
Figure 1.21: Unnatural amino acid (UAA) mutagenesis in live cells.
The mutant aaRS (red) aminoacylates the orthogonal tRNA$_{CUA}$ (blue) which leads to incorporation of the UAA (blue X) in response to the amber codon (red) by the ribosome. The endogenous synthetases (grey) and tRNAs (black) only incorporate native amino acids (black). Image adapted with permission from Young and Schultz, *J. Biol. Chem.* **2010**, 285(15):11039-44.

Caged proteins have been developed with a variety of amino acid analogues through UAA incorporation in bacterial, yeast, and mammalian cells. Several light-responsive UAAs have been synthesized and successfully incorporated in response to an amber codon, such as ONB,[113] 6-nitropiperonyloxycarbonyl (NPOC),[114] and coumarin-based[115] derivatives of Cys,[116, 117] Lys,[118] Ser,[119] and Tyr[120] (Figure 1.22). Advantages of this methodology over other approaches for light-activated proteins are that the location specificity of the caged amino acid is determined through mutagenesis of the protein encoding gene, photochemical removal of the

caging group restores the native amino acid residue, and the light-activated proteins are produced in live cells.
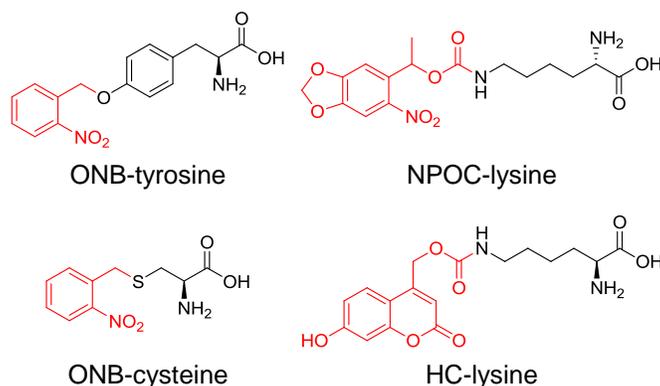
Figure 1.22: Examples of caged amino acids.
Caging groups have been synthetically installed onto the side chains of several amino acids. Photolabile caging groups are indicated in red.

As an example for the specific application of UAA mutagenesis in the incorporation of caged amino acids into proteins, a photocaged lysine derivative (NPOC-lysine, Figure 1.22) was incorporated into MEK1 kinase to optically control kinase signaling networks in live cells (Figure 1.23).[121] In this study, the caged amino acid was placed into the ATP binding pocket of the MEK1 catalytic domain, perturbing its function until light induced decaging generates the native protein (Figure 1.23A). MEK1 is part of the Raf/MEK/ERK signaling pathway, and when activated by phosphorylation it in turn phosphorylates the protein kinases ERK1 and ERK2.[122] Once phosphorylated, ERK1/2 detach from MEK1 and translocate into the nucleus to regulate gene expression,[123] while dephosphorylation by nuclear MAPK phosphatases (MKP) results in ERK1/2 returning to the cytoplasm, producing a feedback loop within this dynamic phosphorylation pathway (Figure 1.23B).[124] Stimulation of this pathway can be achieved in cells expressing MEK1 and ERK2 with epidermal growth factor (EGF), which activates the entire

MAP kinase pathway, resulting in ERK2 translocation from the cytoplasm to the nucleus. Tracking of the pathway activation and subsequent nuclear translocation can be observed through imaging of ERK2 fused to an enhanced green fluorescent protein (EGFP-ERK2). To demonstrate that caged MEK1 (C-MEK1) could be photochemically activated, the caged protein was expressed in live cells containing the EGFP-ERK2 reporter. Real-time measurements of EGFP-ERK2 nuclear translocation showed that the C-MEK1 rapidly induced nuclear translocation of the EGFP-ERK2 reporter through photochemical activation of the phosphorylation pathway (Figure 1.23C). The anchoring of ERK2 in the nucleus was retained, in contrast to cells stimulated with EGF that showed temporally limited nuclear translocation. These findings demonstrated that the photoactivated C-MEK1 acts as a stationary stimulus and optical activity of the subnetwork is not subjected to negative feedback and compensatory effects that were observed when stimulating the EGF receptor on the cell surface. The UAA mutagenesis strategy for creating site-specifically caged kinases that are expressed in living mammalian cells and can be activated with light irradiation enabled the study of single step kinetics in signaling cascades. This method is generally applicable to creating photoactivated kinases for other signaling pathways to provide precise quantitative insights into the kinetics of kinase networks and to dissect temporal regulation.
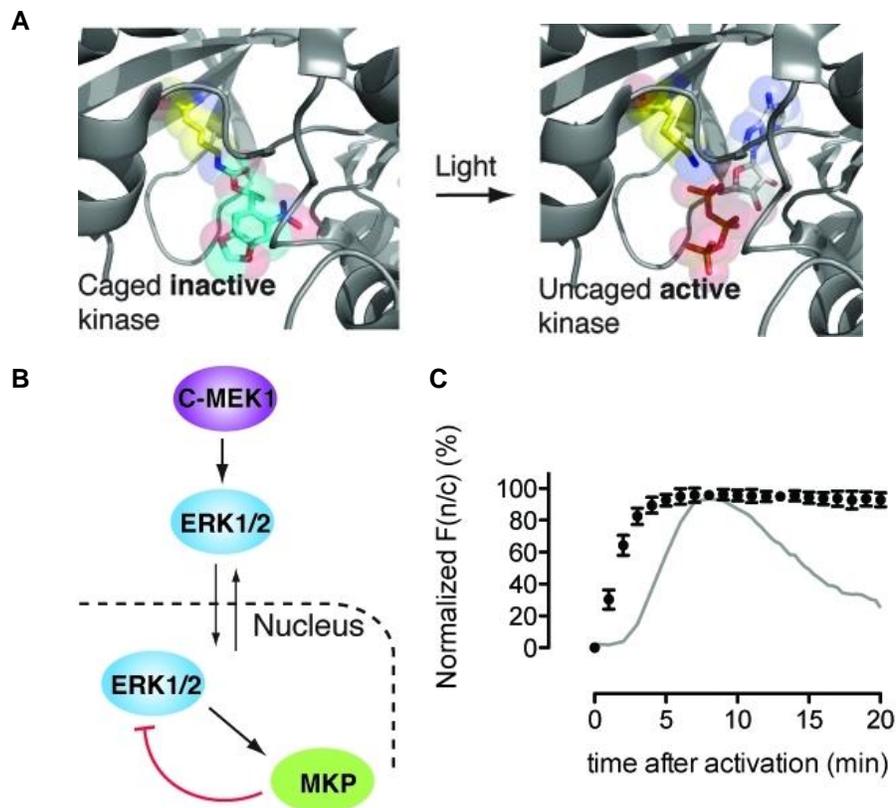
Figure 1.23: Genetically encoded caged MEK1 using UAA mutagenesis.
A) MEK1 was modified with a genetically encoded photocaged lysine at the K97 residue, sterically blocking the ATP binding pocket, which renders the kinase inactive until light irradiation removes the caging group. B) Scheme of the potential negative feedback within the photoactivatable subnetwork. C) Normalized F(n/c) (ratio of nuclear to cytoplasmic EGFP fluorescence intensities) as a function of time after. The gray line shows the normalized F(n/c) observed when cells are stimulated with EGF. Image adapted with permission from Gautier et al, *J. Am. Chem. Soc.* **2011**, 133(7):2124-2127. Copyright 2011 American Chemical Society.

The site-specific incorporation of caged unnatural amino acids has been applied to a number of proteins for photochemical control of activity including Cre recombinase,[125] DNA polymerase,[126] RNA polymerase,[127] zinc finger nuclease,[128] kinsases,[114, 121] transcription factors,[129] ion channels, [130] intein,[131] and reporter genes such as luciferase or EGFP.[115, 117] The use of UAA mutagenesis to express caged proteins for the optical regulation of gene function will be further described in Chapter 4.0 , focusing primarily on the NPOC-lysine, shown in Figure 1.22, and its application in mammalian cell systems. These proteins are caged in a manner

that deactivates function, typically through inhibition of active-site residues or steric blocking of conformational changes required for protein activity. Light irradiation then activates the caged protein to initiate a biological process of interest, which is monitored to demonstrate optical regulation of gene expression or protein function.

2.0   Application of Caged Oligonucleotides in the Regulation of Gene Function

This chapter will describe the application of nucleobase caging groups and photocleavable linkers to several biologically relevant oligonucleotides (as described in Section 1.1) The use of caged oligonucleotides for the optical control of gene function was further expanded to include new classes of regulatory nucleic acids. Importantly, the use of photocaged oligonucleotides for the regulation of gene function was demonstrated in live animals, allowing for optically controlled *in vivo* studies.

2.1   Optical Regulation of Transcription through Caged Triplex-Forming Oligonucleotides

This material was reprinted, in part, with permission from Govan, J. M.; Uprety, R.; **Hemphill, J.**; Lively, M. O.; Deiters, A. *ACS Chem. Biol.* **2012**, 20 (7), 1247-56. Oligonucleotide syntheses were performed, in part, by the author of this thesis. All biological experiments were conducted by Jeane Govan in the Deiters lab.

Triplex-forming oligonucleotides (TFOs) bind target duplex DNA in a sequence-specific manner and can block transcription factors to suppress gene expression.[132] These single-stranded oligonucleotides recognize double-stranded DNA templates and bind in the major groove through Hoogsteen hydrogen bonds to form T:A:T and G:G:C triplex structures (Figure 2.1A and

B).[133] TFOs have been applied to the inhibition of protein–DNA binding,[134] DNA replication,[135] and gene expression,[136] but generally lack precise spatial and temporal control which limits potential applications. Caged TFOs were developed for the photochemical regulation of Hoogsteen base-pairing, allowing activation or deactivation of TFO function and gene transcription in mammalian cells (Figure 2.1C).
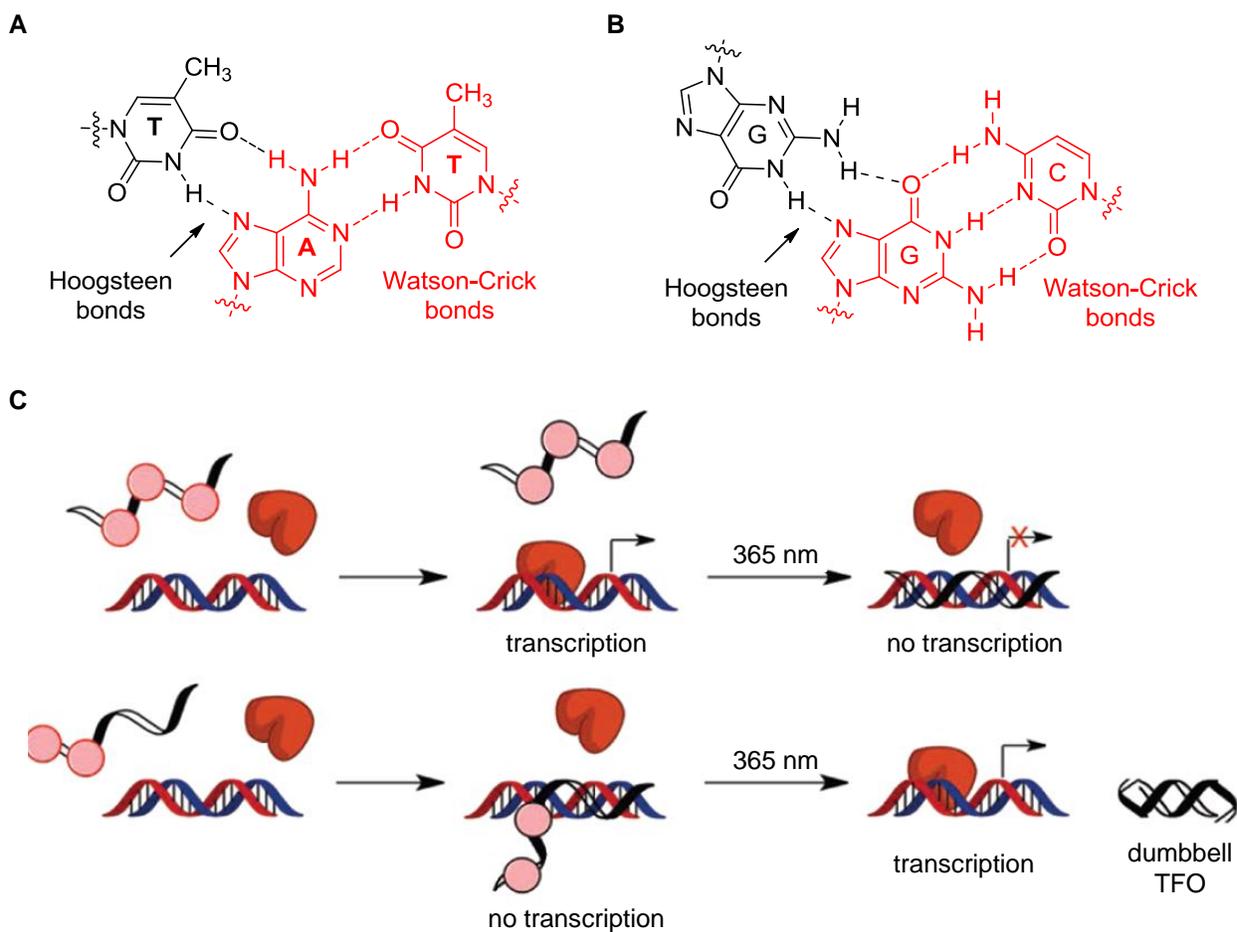
Figure 2.1: Schematic of light-controlled triplex forming oligonucleotides (TFOs). A) Hydrogen bond formation between the A:T pair in duplex DNA (red) and a T in the TFO. B) Hydrogen bond formation between the G:C pair in duplex DNA (red) and a G in the TFO. Watson-Crick (red) and Hoogsten (black) hydrogen bonds are indicated. C) Schematic of a light-activated hairpin TFO (top) and a light-deactivated dumbbell TFO (bottom). Image adapted with permission from Govan et al, *ACS Chem Bio*. **2012**, 7(7):1247-56.

The cyclin D1 promoter was selected as a target duplex for the design of light-controlled TFOs that utilize hairpin and dumbbell secondary structures (Table 2.1). Hairpin loop structures on the 5′ and 3′ termini of antisense agents have been shown to stabilize oligonucleotides in tissue culture while maintaining their antisense properties,[137] a methodology that has previously not been applied to TFOs. In the caged form, the hairpin TFO will not bind the DNA target, while after irradiation the caging groups are cleaved and the TFO is activated, thereby photochemically inhibiting transcription. In contrast, a caged dumbbell TFO was designed to achieve photochemical inhibition of TFO activity and light-activation of transcription. In the caged form, the dumbbell TFO will bind the DNA target, while after irradiation the TFO is deactivated through an internal dumbbell structure formation that will no longer bind the DNA target, thereby photochemically activating transcription.

Table 2.1: Sequences of the light-controlled triplex-forming oligonucleotides.
Hairpin (HP) and dumbbell (DB) TFOs are shown. Caged base residues are underlined, bolded, and highlighted red ("**T**" represents an NPOM-caged thymidine residue, "**C**" represents an NPE-caged cytidine residue). Caged DNA oligonucleotides were synthesized according to the general protocol 5.2**Error! Reference source not found.**. Oligonucleotide residues synthesized by the Lively Lab (Center for Structural Biology, Wake Forest University School of Medicine) are indicated with a caret (^).

Strand	Sequence (5' → 3')
HP-TFO-1^ hairpin	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTAT CGCGCGCAAAGCGCGCG
CHP-TFO-1^ 4T caged hairpin	GCGCGCGAAACGCGCGCTACG**T**GGG**T**GGGGG**T**GGGGGG**T**AT CGCGCGCAAAGCGCGCG
DB-TFO-1 dumbbell	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTAT CGCGCGCAAAGCGCGCGATACCCCCCACCCCCACCCACGTA
CDB-TFO-1 4C caged dumbbell	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTAT CGCGCGCAAAGCGCGCGATA**C**CCCC**C**ACCC**C**CAC**C**CACGTA

TFO inhibition of gene expression was first investigated with a luciferase reporter containing an upstream cyclin D1 promoter (pCyclin D1 Δ-944).[138] The caged hairpin TFO

(CHP-TFO-1) showed UV dependent inhibition of gene expression for the luciferase reporter construct (Figure 2.2A). The caged dumbbell TFO (CDB-TFO-1) was active in the absence of UV irradiation, though exposure to UV light removes the caging groups which deactivates the TFO through dumbbell formation and recovers gene expression (Figure 2.2A). These findings were validated with photochemical regulation of an endogenous cyclin D1 gene using qRT-PCR to quantify cyclin D1 mRNA (Figure 2.2B). The caged TFOs recovered relative cyclin D1 expression to similar levels observed with the corresponding noncaged TFOs when irradiated, indicating full UV decaging.
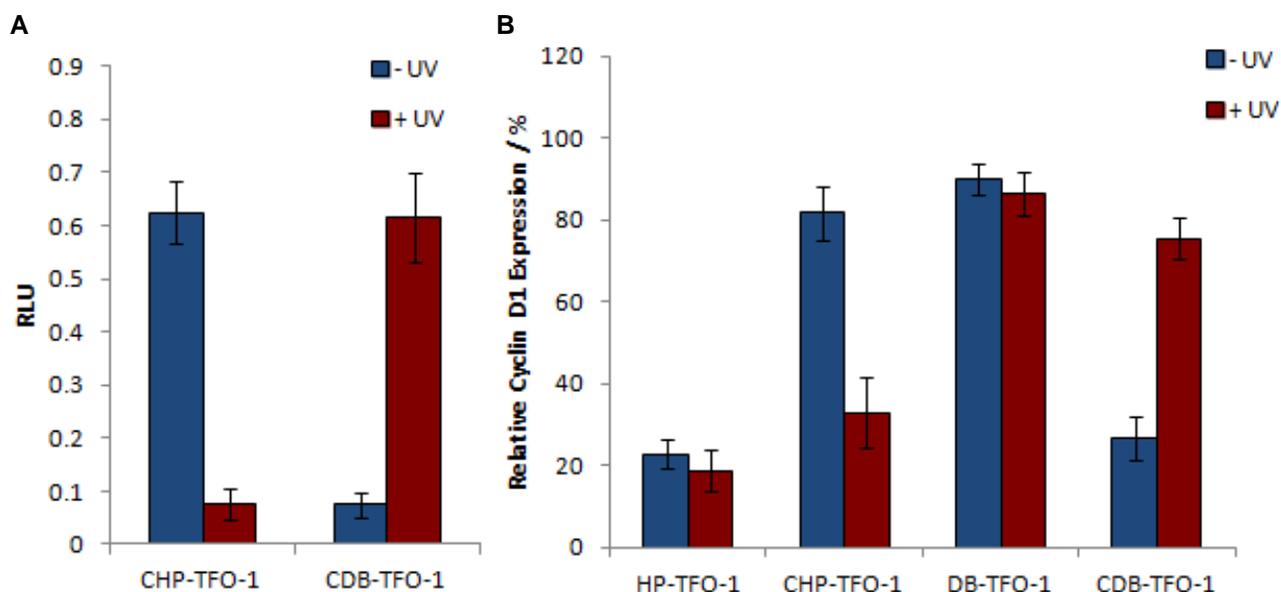
Figure 2.2: Photochemical activation of gene transcription in mammalian cells.
A) Normalized relative luciferase units (RLU) for a luciferase reporter gene downstream of the cyclin D1 promoter. B) Relative cyclin D1 expression in MBA-MD-231 cells. Error bars represent standard deviations from three independent experiments. Image adapted with permission from Govan et al, *ACS Chem Bio*. **2012**, 7(7):1247-56. These experiments were conducted by Jeane Govan.

Caged triplex-forming oligonucleotides were synthesized and successfully applied to the photochemical regulation of transcription in mammalian cells, enabling light-activation and

light-deactivation of gene transcription. Through the direct incorporation of caging groups on nucleobases within the TFO molecules, light-induced inhibition of a specific promoter (cyclin D1) and thus inhibition of gene transcription was achieved. In addition, a caged TFO design was developed that enables the light-activation of gene expression. This was achieved through the synthesis of a new caged deoxycytidine phosphoramidite and its incorporation into dumbbell-forming TFOs. These methodologies were validated in human cells by photochemically controlling the transcription of a transiently transfected reporter gene (luciferase) and an endogenous gene (cyclin D1). Since TFOs are versatile inhibitors of transcription, and have implications in the precise regulation of gene promoter activity in tissue culture or multicellular organisms, these caged transcriptional regulators can be applied as photochemical ON/OFF switches to study gene expression. Potential applications of caged TFOs range from basic biological studies of gene expression to new gene therapeutic approaches with high spatial and temporal resolution.

2.2   Optical Control of MicroRNA Function using Caged Antagomirs

This material was reprinted, in part, with permission by The Royal Society of Chemistry from Connelly, C.M.; Uprety, R.; **Hemphill, J.;** Deiters, A. *Mol Biosyst*. **2012,** 8(11), 2987-93. Oligonucleotide syntheses were performed by the author of this thesis. All biological experiments were conducted by Colleen Connelly in the Deiters lab.

MicroRNAs (miRNAs) are small noncoding RNAs that down regulate gene expression in a sequence specific fashion by binding the 3′ untranslated regions of target mRNAs.[139] It has been estimated more than 30% of all gene expression is regulated by miRNA activity[140] and the

expression of certain miRNAs has been linked to a wide range of human diseases.[141, 142] Gene regulation through miRNA pathways has become an increasingly studied field due to the prevalence of misregulation linked to disease states, highlighting the important role of understanding miRNA regulation and biogenesis.[143] The most common tools to study and control miRNA function are chemically modified complementary oligonucleotides termed antagomirs.[144, 145] Here, the use of caged antagomirs was demonstrated for the photochemical control of miRNA function and inhibition of gene expression (Figure 2.3).
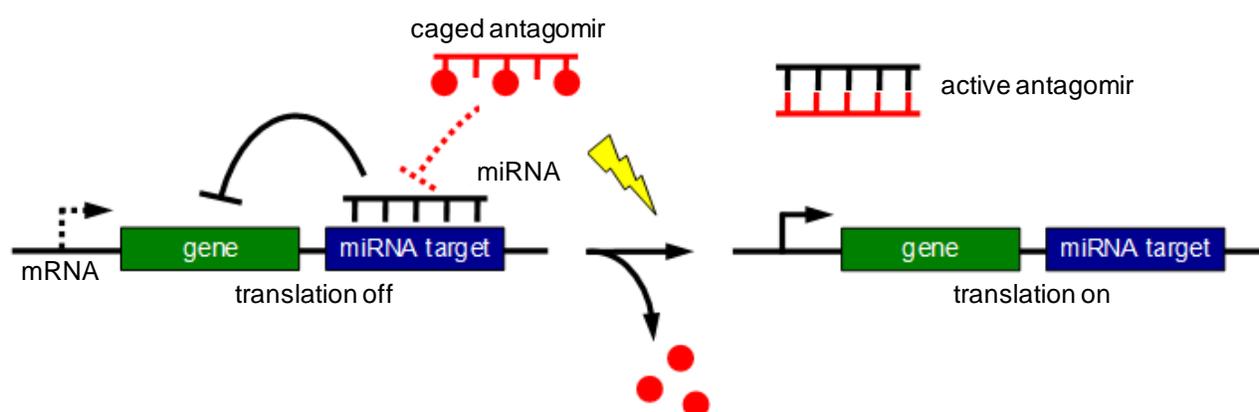
Figure 2.3: Regulation of gene expression with nucleobase caged antagomirs.
Image adapted with permission from Connelly et al, *Mol Biosyst.* **2012**, 8(11):2987-93.

MicroRNAs miR-21 and miR-122 were selected as targets to demonstrate light-activated inhibition of miRNA function. Over expression of miR-21 is observed in many cancer types,[146] while miR-122 is a liver specific miRNA[147] that has roles in hepatitis C virus replication and infection.[148] Caged antagomirs were synthesized with 2′OMe modified nucleotides and phosphorothioate backbones, containing three or four caging groups depending on the antagomir sequence (Table 2.2). These nucleobase-caged antagomirs have no effect on miRNA-mediated gene silencing until activation through photochemical removal of the caging groups, which

46

allows the antagomirs to bind to the miRNAs and blocks their function, leading to the light-activation of gene expression.

Table 2.2: Sequences of the noncaged and caged antagomirs.
All oligonucleotides contain phosphorothioate backbone linkages and 2'OMe modified nucleobases. Antagomirs for miR-21 and miR-122 are shown. Caged 2'OMe NPOM-uridine residues are underlined, bolded, and highlighted red ("**U**"). Caged oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
miR-21 antagomir	AUCAACAUCAGUCUGAUAAGCUA
miR-21 4U caged antagomir	A**U**CAACA**U**CAGUCUGA**U**AAGC**U**A
miR-122 antagomir	ACAAACACCAUUGUCACACUCCA
miR-122 3U caged antagomir	ACAAACACCA**U**UG**U**CACAC**U**CCA

The photochemical control of antagomir activity was investigated in mammalian cell culture using the psiCHECK-2 (Promega) reporter plasmid containing downstream miRNA target sequences.[149, 150] The presence of endogenous miRNAs decreased luciferase signal through inhibition of gene expression, and antagomir binding to the miRNA recovered luciferase expression. The caged miR-21 antagomir showed no inhibition of miRNA function in the absence of UV irradiation, but exhibited an increase in gene expression upon decaging similar to a noncaged antagomir (Figure 2.4A). The caged miR-122 antagomir also showed no gene expression in the absence of irradiation and UV dependent miRNA inhibition to recover luciferase activity (Figure 2.4B). These results demonstrate that caged antagomir activity is fully restored upon UV irradiation with similar levels of gene expression observed as the noncaged control antagomirs.
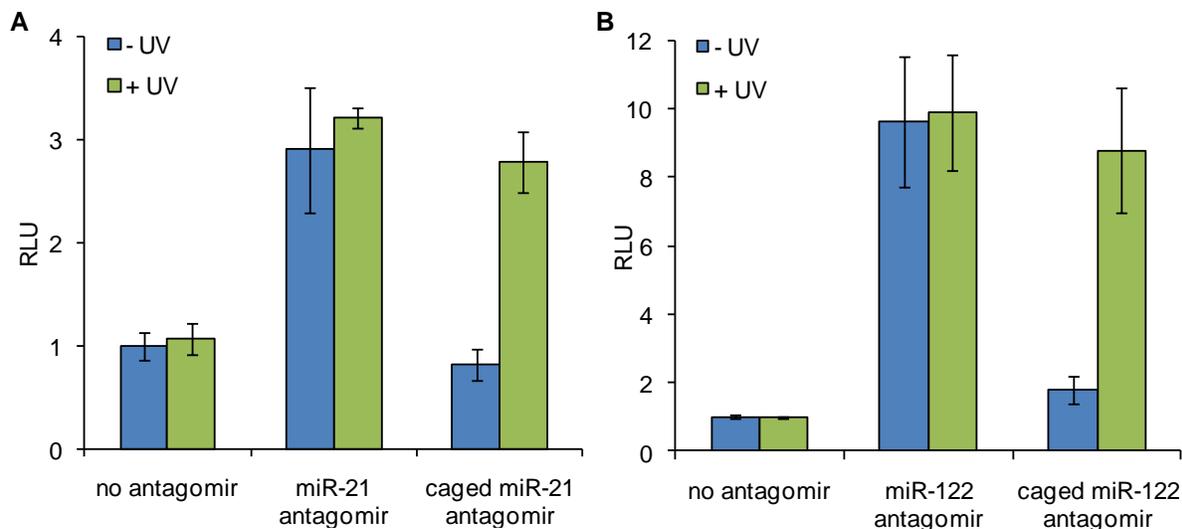
Figure 2.4: Light-activated inhibition of miRNA function.
A) Photochemical inhibition of miR-21 function. B) Photochemical inhibition of miR-122 function. Huh7 cells expressing the psi-CHECK-2 reporters were transfected with the caged and noncaged antagomirs then UV irradiated (365 nm, 25 W) for 5 min. Dual luciferase assays were performed at 48 h, and *Renilla* luciferase was normalized to the control firefly luciferase. The error bars represent standard deviations from three independent experiments. Image adapted with permission from Connelly et al, *Mol Biosyst.* **2012** , 8(11):2987-93. These experiments were conducted by Colleen Connelly.

The miR-21 target sequence was then placed downstream of an EGFP reporter gene to achieve spatial control for the photochemical regulation of miRNA activity in mammalian cells. Irradiations were performed with a fiber optic probe in a defined region (Figure 2.5). EGFP expression was only observed within the irradiated area, demonstrating spatial control over antagomir activity and photochemically activated inhibition of miR-21 function.
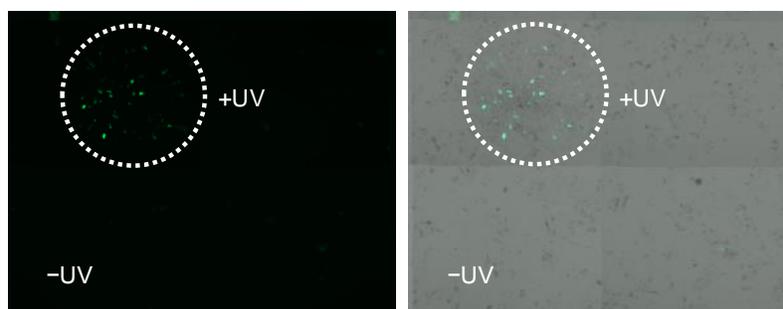
Figure 2.5: Spatial activation of EGFP expression with a caged miR-21 antagomir. Huh7 cells were co-transfected with an EGFP sensor for miR-21 function and the caged miR-21 antagomir (100 pmol). Cells were irradiated at 365 nm using a LED fiber optics probe (Prizmatix) and were imaged after 48 h. The EGFP channel (left) and corresponding brightfield merged images (right) are shown. Image adapted with permission from Connelly et al, *Mol Biosyst.* **2012** , 8(11):2987-93. These experiments were conducted by Colleen Connelly.

Light-activated antagomirs were developed for mature miR-21 and miR-122 through the site-specific installation of caging groups that photochemically regulate antagomir-miRNA hybridization. These synthetic antagomirs were applied to the photochemical regulation of miRNA function in mammalian cells using both a luciferase reporter and an EGFP reporter. Both temporal and spatial control over miRNA function was achieved using a caged miR-21 antagomir. Photocaging enables the precise activation of miRNA antagomirs and deactivation of miRNA function using UV irradiation, with sequence specificity that allows for antagomirs to be designed to target any miRNA of interest. These caged antagomirs can be used to further explore the roles of miRNA in signaling networks or human diseases with high spatial and temporal resolution.

2.3 Optical Control of microRNAs Controlling Polarization of Migrating Cortical Neurons

In addition to the initial studies of photochemically caged antagomirs in cell culture, the technology was later applied to the study of neuronal migration in mouse brain slices. The optical regulation of miRNA function through caged antagomirs was used to determine miRNA-based roles of radial migration during cerebral corticogenesis, in collaboration with the Nguyen lab (University of Liège, GIGA-Neurosciences, Developmental Neurobiology Unit, Belgium).

This material was reprinted, in part, with permission from Volvert, M.L.; Prévot, P.P.; Close, P.; Laguesse, S.; Pirotte, S.; **Hemphill, J.**; Rogister, F.; Kruzy, N.; Sacheli, R.; Moonen, G.; Deiters, A.; Merkenschlager, M.; Chariot, A.; Malgrange, B.; Godin, J.D.; Nguyen, L. *Cell Reports.* **2014**, 7(4), 1168-1183. Oligonucleotide syntheses were performed by the author of this thesis. All biological experiments were conducted in the Nguyen lab.

The cerebral cortex comprises six layers of neurons born in the progenitor zones of the forebrain. Dorsal cortical progenitors generate temporal cohorts of neurons that undergo active migration to reach their final positions in successive cortical layers where they extend neurites to finalize contacts with target cells.[151] Cell migration and branching require dynamic cell shape remodeling orchestrated both by extracellular and intracellular cues that ultimately converge on the cytoskeleton.[152] Untangling the mechanisms that drive neuron migration and integration to appropriate neuronal networks is thus critical for understanding the biological basis of these disorders as well as the emergence of cortical architecture, connectivity and functions during development. Although most projection neurons undergo somal translocation at early phases of corticogenesis, they combine different migration modes to reach their final position at later stages.[153] Bipolar progenitors leave the ventricular zone (VZ) and start radial migration. When

they reach the intermediate zone (IZ), they become multipolar.[154] This morphological conversion is a critical regulation step as mutations in genes that control the multipolar stage often lead to radial migration defects. Over the past few years, several studies started to define the mechanisms underlying radial migration, and most identified regulators were actin or microtubule (MT) cytoskeleton-associated proteins.[153] In spite of these advances, little is known about the molecular basis and more particularly the epigenetic control of cell shape conversion during the successive steps of radial migration in the cortex.

Epigenetics has recently been extended to the study of miRNAs that regulate signaling pathways to control neurogenesis[155] including those that orchestrate successive steps of corticogenesis.[156] Several miRNAs are abundant in the developing cerebral cortex, among which some show dynamic expression that correlates with developmental milestones of the cortex. Critical roles have been revealed for Dicer in cortical neurogenesis,[157] and although cortical phenotypes resulted from loss of mature miRNAs, functional connections to individual miRNA have been mostly correlative. Therefore, there is currently no evidence that miRNAs directly control the migration of postmitotic projection neurons.[156]

Initial research findings indicate that conditional removal of Dicer in postmitotic projection neurons impedes radial migration, and Dicer is required for proper polarization of neurons during migration. Additionally, overexpression of the corepressor silencing transcription factor CoREST[158] contributes to radial migration defects of Dicer knockout neurons. It was shown that the CoREST/REST transcriptional repressor complex contains predicted conserved miRNA target site sequences in the 3'UTR (Figure 2.6A). The role of miRNA regulation of CoREST in radial migration was confirmed, showing that miR-22 and miR-124 promote radial

migration of projection neurons by targeting CoREST, and were further identified to control

neuronal migration in the developing cortex.
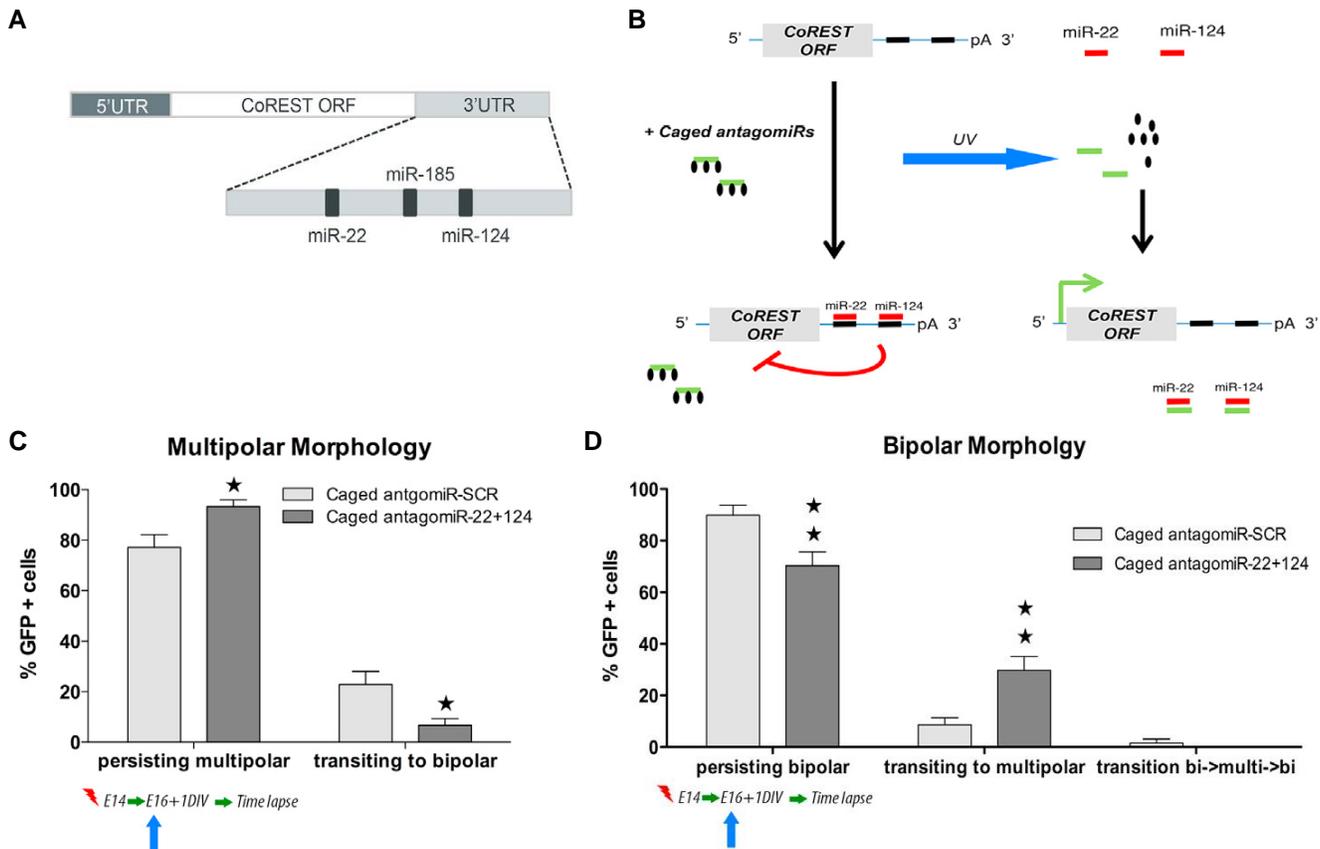
Figure 2.6: Optochemically activated antagomir inhibition of CoREST.
A) miRNA target sites in the 3'UTR of the CoREST ORF. B) Photochemical activation of
nucleobase-caged antagomirs for postmitotic neuronal migration in cultured brain slices. UV light
(blue arrow) activates the caged antagomirs, which bind to endogenous miR-22 or miR-124 and
block their function in cortical neurons, leading to CoREST upregulation. C and D) Histograms
showing the percentage of neurons with light-activated (blue arrow) caged antagomirs that
undergo multipolar to bipolarconversion (C) or that maintained a stable bipolar morphology
during migration (D) in 1 day cultured slices from corresponding brains. Image adapted with
permission from Volvert et al, *Cell Rep.* **2014**, 7(4):1168-83. These experiments were conducted
by the Nguyen lab.

Optical tools were then exploited to induce spatial and temporal changes of endogenous

miR-22 and miR-124 activity (Figure 2.6B). Brains were sliced 2 days after electorporation of

the caged antagomirs (Table 2.3) and UV illuminated, followed by real-time imaging 24 h later.

52

It was shown that the light-activated antagomirs reduced bipolar conversion (Figure 2.6C) and impaired bipolar stability (Figure 2.6D) of neurons navigating in the IZ. These results demonstrate that endogenous miR-22 and miR-124 are both required for radial migration by regulating the dynamic morphological remodeling of projection neurons. These results also suggest that the migration phenotype resulting from the conditional removal of Dicer in postmitotic neurons mostly arise as a consequence of the lack of miR-22 and miR-124 maturation. The expression of CoREST in the developing cortical wall results from the interplay between transcriptional and translational mechanisms, including miR-22 and miR-124.

Table 2.3: Sequences of the noncaged and caged antagomirs.
All oligonucleotides contain phosphorothioate backbone linkages and 2'OMe modified nucleobases. Antagomirs for miR-22 and miR-124 are shown. Caged 2'OMe NPOM-uridine residues are underlined, bolded, and highlighted red ("**U**"). Caged oligonucleotides were synthesized according to the general protocol 5.2**Error! Reference source not found.**.

Strand	Sequence (5' → 3')
miR-22 antagomir	AACAGUUCUUCAACUGGCAGCUU
miR-22 3U caged antagomir	AACAGU**U**CUUCAAC**U**GGCAGC**U**U
miR-124 antagomir	AGGCAUUCACCGCGUGCCUUA
miR-124 3U caged antagomir	AGGCAU**U**CACCGCG**U**GCCU**U**A

A tight regulation of projection neuron migration is fundamental for the establishment of functional connectivity in the developing neocortex. Although several genes and molecular pathways have been associated with neuronal migration in physiological or pathological conditions,[152] the epigenetic control of this process remains poorly investigated. Dicer was conditionally removed in postmitotic neurons, then combined with global gene expression profiling and miRNA expression pattern analysis to identify the miRNAs and corresponding gene targets involved in radial migration regulation. miR-22 and miR-124 were discovered to be enriched in the cortical wall where they target CoREST to fine-tune expression of Doublecortin

(Dcx), thereby promoting migration to the appropriate layer in the cortical plate. CoREST was the only gene from a short list of regulators of neuronal polarization whose expression was upregulated in Dicer-depleted projection neurons. This correlates with reduction of transcriptional Dcx of as well as downregulation of its expression in cortical neurons depleted for Dicer. At the functional level, lack of Dicer disrupted radial migration at distinct steps, including multipolar-bipolar cell transition in the IZ and bipolar stability during locomotion. This phenotype was rescued after targeting CoREST or expressing Dcx in Dicer-depleted neurons. Although the fine-tuning of additional targets by other miRNAs during the migration of projection neurons cannot be excluded, the regulation of CoREST expression by miR-124 and miR-22 is a critical step for proper polarization of projection neurons during their migration in the developing cerebral cortex. The dynamic expression of Dcx is controlled by activators and repressors, the latter including the CoREST complex, which is targeted by miR-22 and miR-124. The use photochemically caged antagomirs assisted in the discovery of this miRNA-based role in CoREST regulation and neuronal migration pathways.

2.4   Optical Control of Gene Expression through Site-Specific Promoter Caging

This material was reprinted in its entirety with permission from **Hemphill, J.**; Govan, J. M.; Uprety, R.; Tsang, M.; Deiters, A. *J. Am. Chem. Soc.* **2014**, 136(19), 7152-7158.

In cell and molecular biology, double-stranded circular DNA constructs, known as plasmids, are extensively used to express a gene of interest. These gene expression systems rely on a specific promoter region to drive the transcription of genes either constitutively (i.e., in a continually "ON" state) or conditionally (i.e., in response to a specific transcription initiator).

However, controlling plasmid-based expression with high spatial and temporal resolution in cellular environments and in multicellular organisms remains challenging. To overcome this limitation, nucleobase-caging groups were site-specifically installed within a plasmid promoter region to enable optochemical control of transcription and, thus, gene expression, via photolysis of the caging groups. Through the light-responsive modification of plasmid-based gene expression systems, optochemical activation of an exogenous fluorescent reporter gene was demonstrated in both tissue culture and a live animal model, as well as light-induced overexpression of an endogenous signaling protein.

For transcriptional activation in eukaryotic systems, a gene expression plasmid contains a promoter sequence upstream of the gene of interest, such as the commonly used cytomegalovirus (CMV) promoter,[159] which further includes a specific transcription initiator sequence called the "TATA box".[160] Transcription is initiated when a subunit of transcription factor IID (TFIID), referred to as the TATA box binding protein (TBP), binds to the TATA box sequence and recruits additional components of the transcriptional machinery such as RNA polymerase II.[161] Since the binding of TBP to the TATA box is the driving force to activate transcription, the TATA box is an optimal site to optochemically control the transcription of a plasmid by applying nucleobase caging technologies. In order to develop a generally applicable system for the optical activation of transcription in vivo, modified short oligomer fragments were inserted into plasmids in a site-specific manner[162] and introduced caged nucleobases at 1−3 defined sites into the TATA-box promoter region of large (>4 kB) gene expression plasmids. The newly developed caged plasmid system was successfully applied to in vivo optochemical control of gene expression in mammalian cells and in zebrafish.

The optochemical regulation of an enhanced green fluorescent protein (EGFP) reporter gene was selected as a proof-of-concept model for the application of a caged promoter region. An EGFP plasmid was modified to contain two nicking sites in the CMV promoter flanking the TATA box, in order to remove a short DNA fragment and replace it with a caged DNA insert (Figure 2.7). The site-specific installation of caged nucleobases within the TATA box will inhibit initiation of EGFP transcription by TBP. Upon UV irradiation, the caging groups would be cleaved and transcription would be activated. To test this hypothesis, 1–3 NPOM-caged thymidine residues[163] were site-specifically incorporated into oligonucleotides containing the TATA box sequence via automated solid-phase DNA synthesis (Table 2.6). A mutant promoter region that contains three T → C base substitutions (TATATAA → CACACAA) was designed as a negative control, based on the ability to replace the same thymidine residues with caged Ts and on a previous analysis that showed less than 4% transcriptional efficiency with two or more mutations in the TATA box region.[164] This T^{mut} negative control contains thymidine mutations at the same sites in the TATA box as the designed NPOM-caged promoter constructs. The installation of the caged thymidine nucleotides was predicted to disrupt hydrogen-bonding interactions between base-pairs within the TATA box region; however, nucleotides outside of this region will still be able to hybridize to their complement sequences as needed for ligation of the inserts into the plasmid.[162]
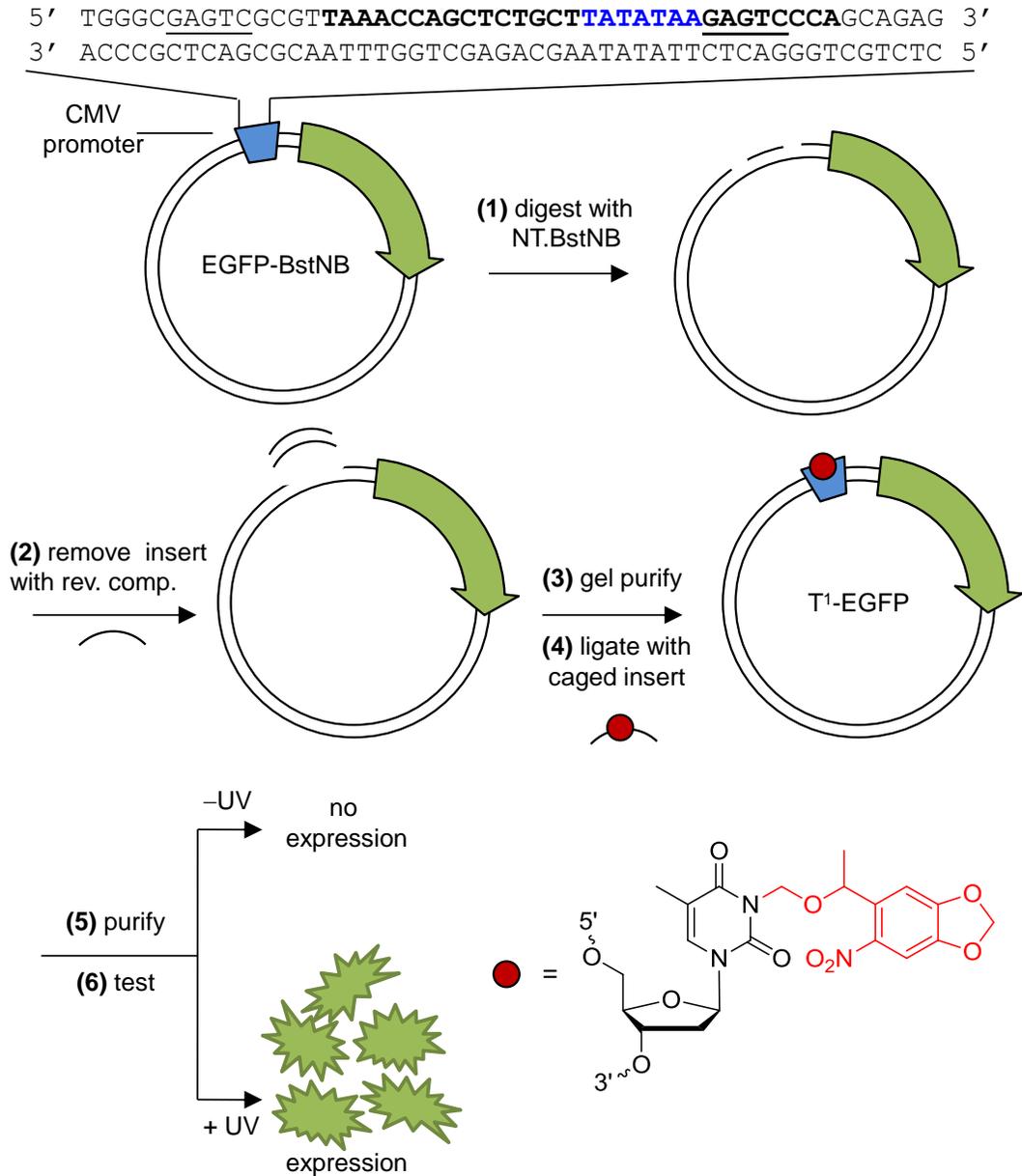
Figure 2.7: Construction of the site-specifically caged plasmids.
(1) pEGFP-Bst is digested with Nt.BstNB and (2) annealed with the reverse complement to remove the TATA box region. (3) The digested plasmid is gel purified and (4) ligated with a phosphorylated caged TATA box insert. (5) The caged plasmid is then column-purified and (6) applied in mammalian cells or live animal models that were either kept in the dark (no expression of EGFP) or irradiated with UV light (EGFP expression). The restriction sites are underlined and the TATA box recognition sequence is shown in blue. The NPOM-caged thymidine is represented by a red circle and the NPOM modification is indicated in red within the nucleotide structure.

To ensure that the caged DNA does not completely inhibit hybridization to its complement sequence, melting temperatures (T_M) were determined (Table 2.4). With each

addition of a caged nucleotide, the T_M slightly decreased, which is reflective of the caging groups interfering with the hydrogen bonding interactions of the nucleobases.[165] As expected, removal of the caging groups through UV exposure completely restored binding of the caged TATA box sequences to their complement sequence similar to the noncaged T_M. Importantly, hybridization for all caged primers was still detected at the ligation temperature of 4 °C, a prerequisite for construction of the promoter-caged plasmid.

Table 2.4: Melting temperatures (T_Ms) of the caged TATA box oligonucleotides.
The T_Ms were experimentally determined with the pEGFP-BstNB plasmid TATA sequence complement. Standard deviations were derived from three separate T_M measurements.

DNA	−UV	+UV
T^0	73.5 ± 0.1	-
T^1	73.5 ± 0.1	74.0 ± 0.2
T^2	72.5 ± 0.2	73.0 ± 0.2
T^3	70.2 ± 0.3	73.5 ± 0.3

The caged plasmids were then assembled and verified by agarose gel electrophoresis (Figure 2.8) prior to subsequently assessment for function in mammalian cell culture.
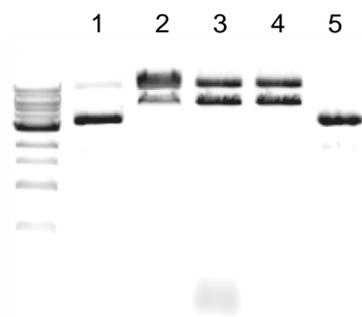
Figure 2.8: Gel analysis of caged plasmid construction.
pEGFP-Bst (lane 1) was digested with Nt.BstNB (lane 2). The reverse complement to the nicked DNA fragment was added and annealed (lane 3). The digested plasmid was gel purified (bottom band, lane 4) and the phosphorylated caged DNA oligo was ligated into the purified plasmid (lane 5). Aliquots of each step were collected and analyzed on a 1% agarose gel. These experiments were conducted by Jeane Govan.

A cell viability assay (CellTiter-Glo, Bio-Rad) was performed to analyze the effects of cellular UV-A application with increasing intervals of 365 nm irradiations, demonstrating that UV-A light does not reduce cell viability for exposures of up to 20 min in the irradiation setup (Figure 2.9).
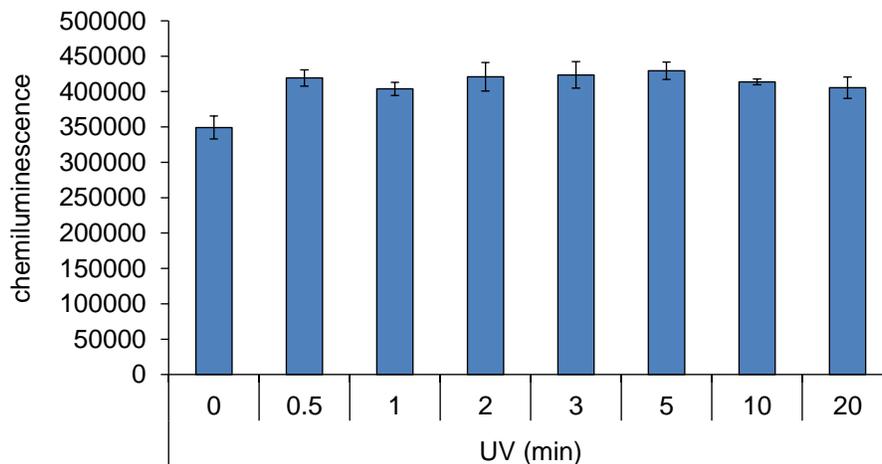
Figure 2.9: Effect of UV exposure on cell viability.
A CellTiter-Glo viability assay (Bio-Rad) was performed in HEK293T cells with increasing intervals of UV irradiation (365 nm, 25 W).

The caged EGFP plasmids were then cotransfected with a DsRed expression vector as a control, and the transfected HEK293T cells were either irradiated for 5 min (365 nm, 25 W) or kept in the dark. After 48 h incubation the cells were imaged for EGFP and DsRed expression (Figure 2.10). As expected, T^{mut} achieved the greatest reduction in EGFP expression from a series of mutants analyzed (data not shown), and the thymidine caging groups were capable of inhibiting EGFP transcription to low basal levels similar to those of the mutated nonfunctional TATA box sequence. After UV irradiation, the caged plasmids regained full functionality, showing expression levels virtually identical to the noncaged EGFP expression plasmid (T^0), which validates that the incorporation of NPOM-caged thymidine nucleotides within the TATA

box sequence can be applied to the optochemical regulation of a plasmid expression vector. The high levels of EGFP expression observed for the T^1–T^3-caged plasmids after UV irradiation suggest that optical control for each of the constructs results in similar levels of gene activation.
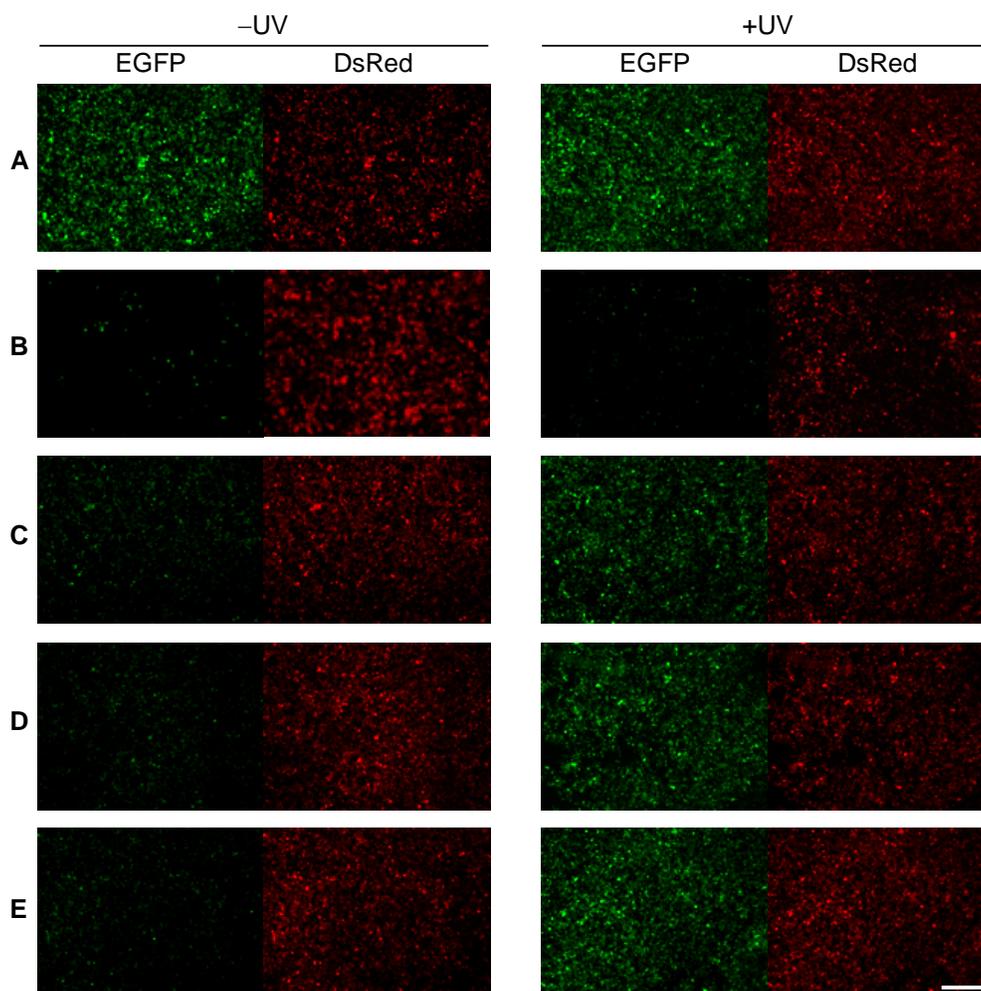
Figure 2.10: Light-activation of EGFP expression with caged plasmids.
A) T^0 noncaged EGFP. B) T^{mut} negative control. C) T^1 caged EGFP. D) T^2 caged EGFP. E) T^3 caged EGFP. HEK293T cells were transfected with noncaged and caged EGFP plasmids and pDsRed-N1. The cells were irradiated for 5 min with UV light (365 nm, 25 W) or kept in the dark. Cells were imaged on 5X magnification after 48 h. EGFP channel is shown to the left and the DsRed channel is shown on the right. Scale bar indicates 200 μm. These experiments were conducted by Jeane Govan.

The light-activation of EGFP expression results were also quantified by flow cytometry (Figure 2.11 and Figure 2.12A), verifying the initial micrograph findings. With each additional caging group a slight reduction in the EGFP background expression was observed. The T^3-caged plasmid was then applied to all subsequent experiments, since it showed the lowest background EGFP expression before light exposure. One of the main advantages of using nucleobase caging technology to photoregulate gene expression is the ability to perform localized and temporal control over biological activity. To this end, HEK293T cells were cotransfected with the T^3-caged EGFP plasmid and the control DsRed expression vector. Following transfection, only a small subset of cells was irradiated with UV light, followed by imaging after 48 h incubation. As shown in Figure 2.12B, EGFP expression was localized to the irradiated area, while DsRed expression was observed in all cells. This demonstrates that the developed TATA box-caging methodology can be applied to optochemically regulate gene expression with spatio-temporal control.

A
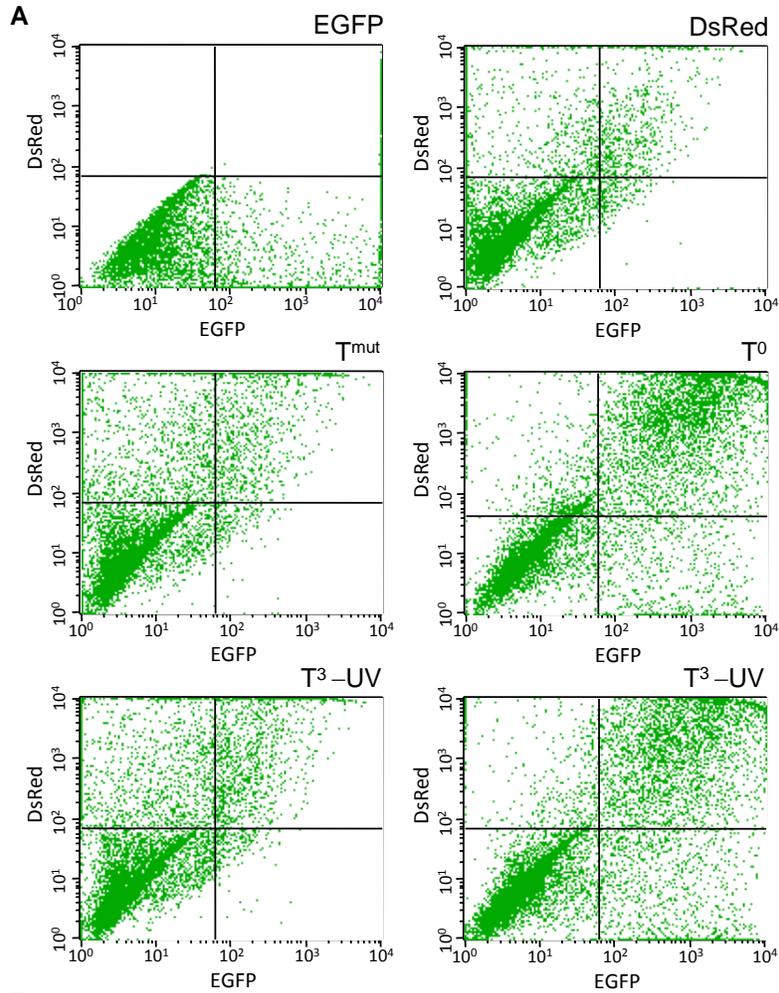
B

	% normalized EGFP expressing cells			
	−UV		+ UV	
	Average	Std Dev	Average	Std Dev
T^{mut}	15.25	6.75	21.42	4.16
T^0	71.39	3.56	81.69	6.41
T^1	22.84	7.72	81.06	8.24
T^2	16.50	6.18	84.10	11.82
T^3	13.69	2.29	76.55	9.64

Figure 2.11: Fluorescent cell counts.
A) Representative EGFP and DsRed fluorescent cell count dot plots for the following conditions: EGFP plasmid alone, DsRed plasmid alone, T^{mut} negative control, T^0 noncaged positive control, T^3 caged plasmid −UV, and T^3 caged plasmid +UV. B) Raw values for the cells expressing both EGFP and DsRed normalized to the number of cells expressing only DsRed. These experiments were conducted in part by Jeane Govan.
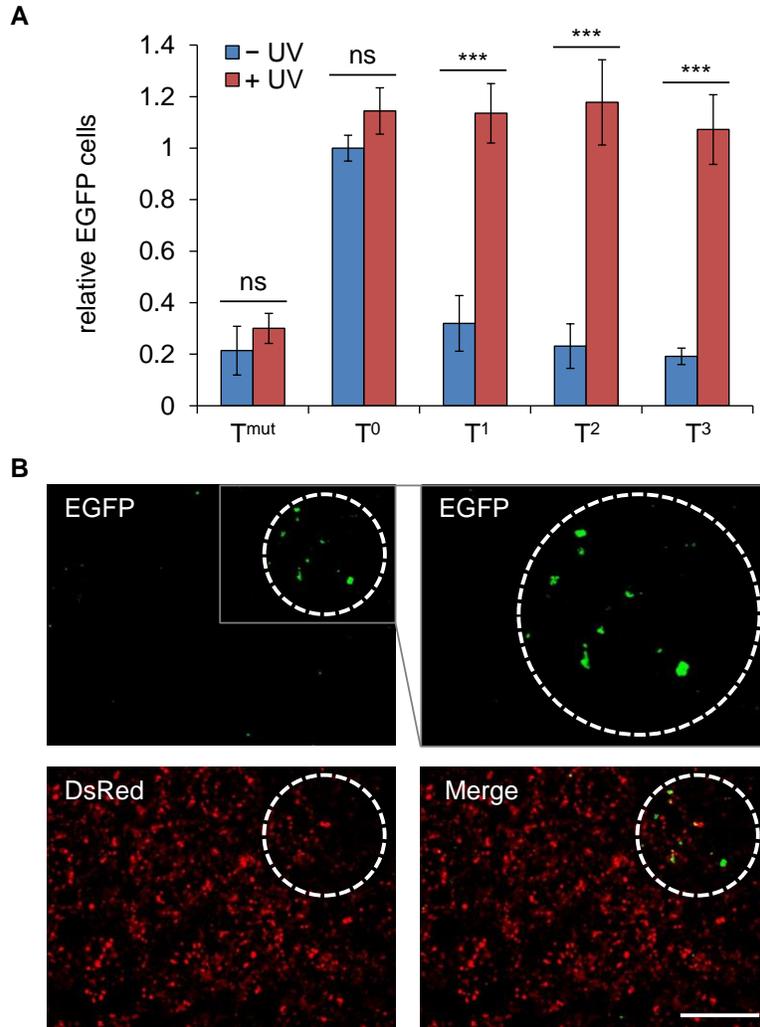
Figure 2.12: Light-activation of EGFP expression.
A) Quantification of light-activated EGFP expression. HEK293T cells were transfected with noncaged and caged EGFP plasmids and a DsRed control plasmid. The cells were irradiated for 5 min (365 nm, 25 W) or kept in the dark. After 48 h incubation, the cells were analyzed by flow cytometry. The number of cells expressing both EGFP and DsRed was normalized to the number of cells expressing only DsRed and set relative to the noncaged plasmid. Standard deviations were calculated form three individual experiments. ns = not significant ($P > 0.05$), *** = highly significant ($P < 0.001$). B) Spatial activation of EGFP expression. HEK293T cells were transfected with T^3-caged EGFP and DsRed plasmids. Cells within the white dashed circle were irradiated through a microscope filter cube (DAPI, BP377/28, 40X) for 30 s and were imaged (5X magnification) after 48 h incubation. An enlarged region of the EGFP channel is shown in the gray box. Scale bar indicates 200 μm.

In order to demonstrate the general applicability of the developed methodology, the optochemical overexpression of an endogenous gene was investigated. Pololike kinase 3 (Plk3) is a serine/threonine kinase that is essential for cells entering into mitosis, spindle formation,

segregation of the chromosomes, and cytokinesis.[166] Plk3 is also a tumor suppressor that when overexpressed can induce cell cycle arrest, chromatin condensation, and apoptosis.[167] Since the ectopic expression of Plk3 leads to disruption of microtubule integrity, a change in cell morphology occurs, namely cytokinesis defects and formation of binucleated and polynucleated cells.[168] Due to its involvement in the cell cycle and the phenotypic change when Plk3 is overexpressed, the optical activation of Plk3 expression was investigated through the engineered caged plasmids. To this end, Plk3 was fused to the C-terminus of the EGFP expression vector based on a previously reported plasmid.[168] The EGFP-Plk3 plasmid was then modified with the caged TATA box DNA sequences as previous stated (see Figure 2.7). HeLa cells were transfected with the modified EGFP-Plk3 plasmids, and the cells were either irradiated for 5 min (365 nm, 25 W) or kept in the dark, followed by incubation for 48 h. This cell line was used to analyze previously reported phenotypic changes[168] and to improve single cell imaging capabilities, since HeLa cells have distinct morphology and form cellular monolayers, in contrast to the previously used HEK293T cell line. The cells were fixed and stained to identify actin filaments as well as nuclei, in addition to EGFP expression. As expected, the positive control T^0-noncaged EGFP-Plk3 plasmid showed transcription of both genes and a significant change in phenotype, specifically the formation of binucleated cells and a loss in cellular structure, while the negative control T^{mut} showed normal cellular morphology (Figure 2.13).
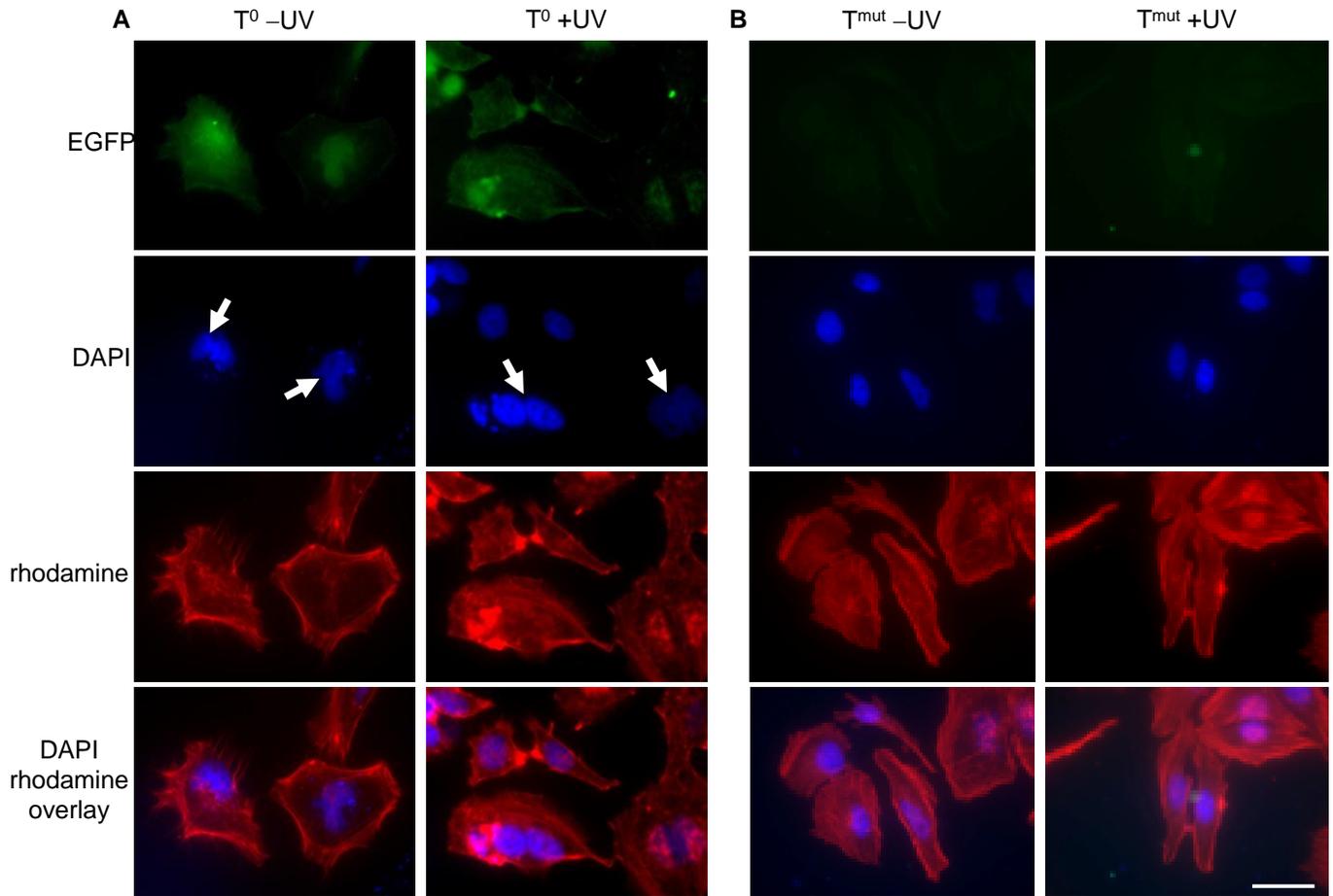
Figure 2.13: Controls for the light-induced over-expression of Plk3.
HeLa cells were transfected with T^0 noncaged EGFP-Plk3 plasmid (A) and the T^{mut} negative control EGFP-Plk3 plasmid (B). The cells were irradiated with UV light (365 nm, 5 min, 25 W), or kept in the dark, and incubated for 48 h. The cells were fixed and stained with DAPI (blue) and Rhodamine-phalloidin (red) and imaged on 63X magnification. White arrows indicate binucleated cells, and scale bar indicates 50 μm. These experiments were conducted by Jeane Govan.

The T^3-caged plasmid showed low EGFP expression as well as little morphological change when the cells were kept in the dark. After UV irradiation, the caging groups were removed and transcription of EGFP-Plk3 was activated, as shown by the increase in EGFP expression and the observation of binucleated cells (Figure 2.14A). Additionally, the overexpression of Plk3 can lead to apoptosis by activating caspase-3.[169] Thus, the downstream activation of caspase-3 activity was measured in response to the optochemically driven overexpression of Plk3. HeLa cells were transfected with the noncaged or caged EGFP-Plk3

plasmids and either kept in the dark or irradiated (365 nm, 5 min) and caspase-3 activity was measured after 48 h (Ac-DEVD-AFC substrate, Calbiochem). The T^0-noncaged TATA box demonstrated a 5-fold increase in caspase-3 activity over nontreated cells. In contrast, the T^{mut} plasmid only led to basal levels of caspase-3 activity, confirming it as a negative control for Plk3-driven caspase-3 activation (Figure 2.14B). In the absence of UV light, the T^3-caged plasmid was inactive as only basal caspase-3 activity similar to nontreated cell was observed, as expected. However, after light-induced activation an increase in caspase-3 activity was detected, indicative of Plk3-driven downstream pathway regulation. These results demonstrate that the overexpression of an endogenous gene can be optochemically activated through the site-specific incorporation of NPOM-caged thymidine nucleotides within the TATA box transcription regulatory region. Here, this was applied to the induction of a phenotypic change and activation of a downstream signaling pathway; however, broad applicability of the developed methodology to the temporal activation of gene function is conceivable.
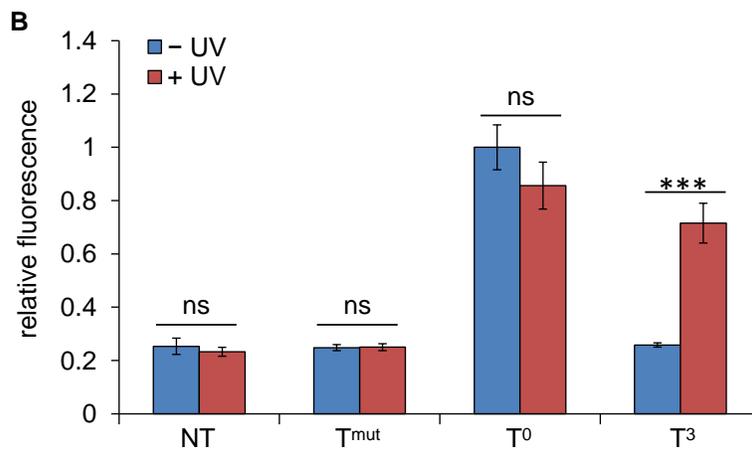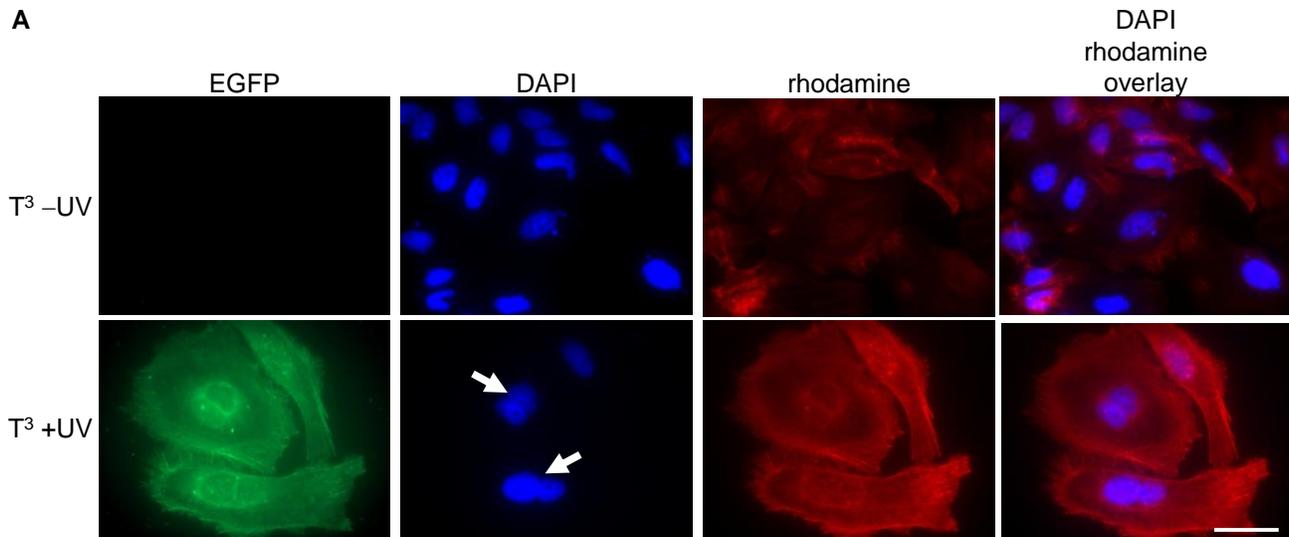
Figure 2.14: Caged plasmid-based light activation of endogenous signaling pathways.
A) Light-induced expression of Plk3. HeLa cells were transfected with the T^3-caged EGFP-Plk3 plasmid followed by irradiation of the caged construct (365 nm, 5 min, 25 W) and incubated for 48 h. The cells were then fixed and stained with DAPI (nuclei) and Rhodamine-phalloidin (actin filaments) prior to imaging (63X magnification). White arrows indicate binucleated cells, and scale bar indicates 50 μm. B) Light-induced activation of caspase-3. HeLa cells were transfected with the T^{mut} negative control, T^0-noncaged, and T^3-caged EGFP-Plk3 plasmids. The cells were either irradiated (365 nm, 5 min, 25 W) or kept in the dark and lysed after 48 h. The lysate was assayed with a fluorogenic caspase-3 substrate (Calbiochem). Fluorescence units were normalized to the noncaged control, and standard deviations were calculated from three individual experiments. ns = not significant ($P > 0.05$), *** = highly significant ($P < 0.001$). These experiments were conducted in part by Jeane Govan.

In order to demonstrate the applicability of the caged TATA box construct to the

optochemical control of gene function in an animal, the expression of a fluorescent reporter was

tested in zebrafish embryos. The zebrafish was selected because it is a common model organism

for developmental studies, its transparency facilitates irradiation, and plasmid-driven gene expression has been well documented via direct microinjection into fertilized eggs.[170] Microinjections with modified EGFP plasmids were performed at the 1-cell stage, and embryos were either irradiated (365 nm, 2 min) or kept in the dark. After 24 h incubation the embryos were dechorionated and imaged for EGFP expression directly and 24 h later. The T^0-noncaged plasmid was used as a positive control, exhibiting mosaic EGFP expression patterns commonly observed with DNA injection due to differential plasmid content in each cell (Figure 2.15A). When the TATA box region of the vector was caged, transcription was deactivated and only a minimal level of EGFP expression was observed. However, after UV exposure, the embryos injected with the T^3-caged plasmid showed activation of EGFP expression, conferring the optical control observed in a single cell environment to gene expression in a multicellular animal (Figure 2.15B). More than 45% of live embryos injected with the caged plasmid construct expressed EGFP after UV irradiation, with levels of transcriptional activation similar to the noncaged control (Figure 2.15C). Although a small population of embryos exhibited low levels of EGFP expression in the absence of UV exposure, a 5-fold increase in EGFP expressing embryos was observed.
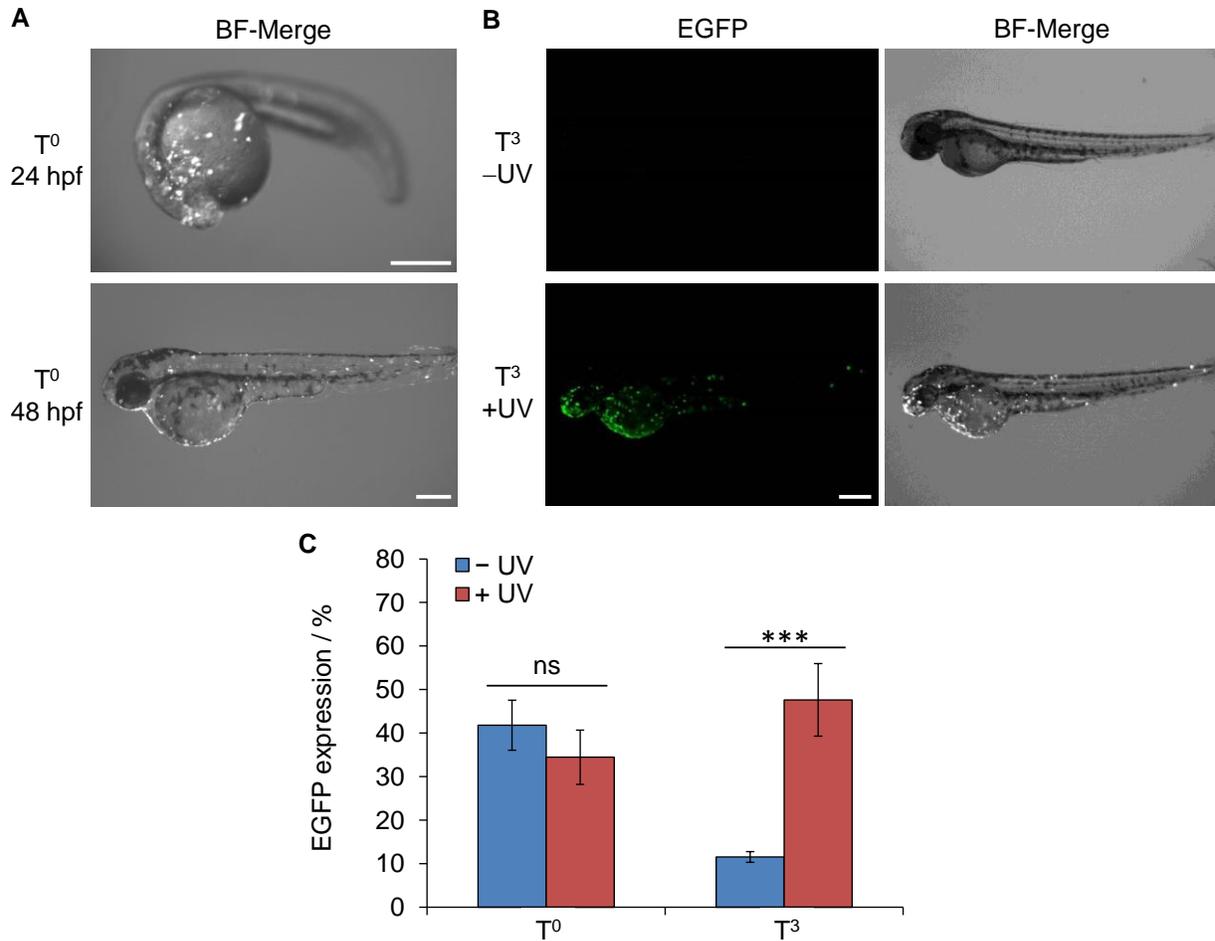
Figure 2.15: Light-induced EGFP expression in zebrafish.
A) Zebrafish embryos were microinjected with the control T^0 noncaged EGFP plasmid (50 pg) at the 1-cell stage and were incubated at 28 °C. Imaging was then performed with dechorionated embryos at 24 hpf and 48 hpf for EGFP expression. B) Embryos were microinjected at the 1-cell stage with the T^3-caged EGFP expression plasmid or the noncaged T^0 plasmid (50 pg). Embryos were then irradiated (2 min, 365 nm) or kept in the dark and incubated at 28 °C for 24 hpf, followed by dechorionation. Imaging was performed at 48 hpf. C) Frequency of the EGFP phenotype for each condition. Error bars represent standard deviations from three (T^0) or four (T^3) independent experiments. N = 9–24. ns = not significant ($P > 0.05$), *** = highly significant ($P < 0.001$). Scale bars indicate 250 μm.

Additionally, a UV irradiation time course was performed, indicating that longer exposures do not significantly enhance the frequency of EGFP expressing embryos (Figure 2.16A). Late-stage irradiations at 8 hpf (75% epiboly stage) were also performed to examine caged plasmid activation during gastrulation, and EGFP expression was observed at 24 hpf

69

(Figure 2.16B). Although the total number of EGFP positive embryos was slightly lower compared to irradiations at earlier stages, the presence of similar mosaic expression patterns as observed in the case of 1 hpf irradiation shows that caged plasmids can provide a means to activate gene expression through UV irradiation later in development (Figure 2.16C). These results demonstrate that the caged promoter sequence allows for the construction of plasmid-based optochemical gene expression that can be readily applied to live aquatic embryos for the regulation of gene function.
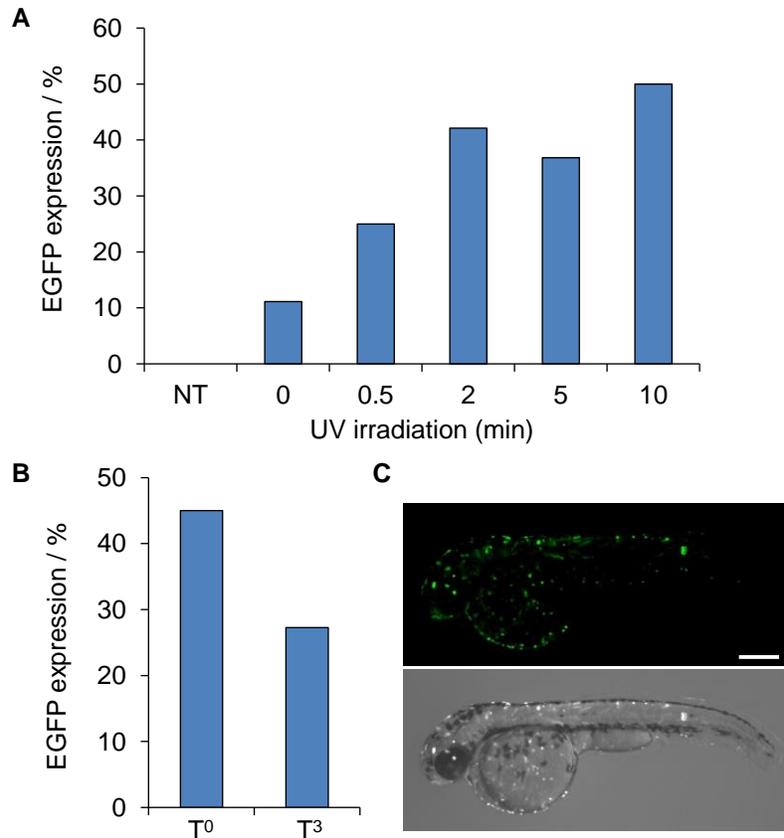
Figure 2.16: Optimization of UV irradiation time.
A) Zebrafish embryos were microinjected with the T^3 caged EGFP plasmid (50 pg) and UV irradiated (365 nm) across a range of times. At 24 hpf the frequency of EGFP expression was determined through embryo counting. N = 6–20. B) Zebrafish embryos were microinjected with the T^0 noncaged and T^3 caged EGFP plasmids (50 pg) at the one cell stage and were incubated at 28 °C. Late-stage 2 min UV irradiations (365 nm) were then performed at 8 hpf (75% epiboly) and the embryos were incubated at 28 °C. At 24 hpf the frequency of EGFP expression was determined through embryo counting. N = 11–20. C) Zebrafish injected with the T^3 caged EGFP plasmid that were late-stage UV irradiated at 8 hpf were dechorionated and imaged at 48 hpf (top: EGFP; bottom: brightfield/EGFP merge). Scale bar indicates 250 μm.

In summary, a system was engineered in which plasmid function can be optochemically regulated with high spatial and temporal resolution. A site-specifically caged promoter region was inserted into expression plasmids via ligation with synthetic nucleobase-caged DNA strands. By installing NPOM-caged thymidine nucleotides within the TATA box promoter region, transcription was inhibited and activity was not observed until the caging groups were removed through a brief exposure to UV light. The optical OFF → ON switching of plasmid function was

assessed using a fluorescent reporter gene in live cells, and spatial control of activation was demonstrated for TATA box-driven gene expression in human tissue culture. Additionally, it was shown that the engineered system could be used to regulate cellular signaling cascades by optochemically triggering overexpression of an endogenous gene, polo-like kinase 3 (Plk3). The effect of Plk3 overexpression was observed by a phenotype change leading to binucleation only in irradiated cells and through upregulation of caspase-3 activity after light-induced Plk3 activation. Lastly, the caged vector methodology was applied to the optochemical triggering of gene expression in live animals. Specifically, light-activated gene expression was achieved in the zebrafish embryo, a multicellular model organism that is extensively used for genetic studies. In contrast to caging of the oligonucleotide phosphate backbone,[171] this approach is completely site-specific, generally applicable, and does not require auxiliary proteins, as only 1–3 NPOM-caging groups are synthetically incorporated onto nucleobases in the TATA-box region of the expression plasmids. Thus, only a few photolysis reactions are required to optically activate gene expression from an otherwise inactive expression vector. This method adds a new and precise synthetic biology tool to the light-regulation of gene function in cells and organism and has broad applicability in the regulation of plasmid-encoded protein expression, as demonstrated in mammalian cell culture and zebrafish embryos.

2.4.1   Methods and Materials

Construction of the Caged Plasmids. See Section 5.3 for specific information on molecular biology techniques The Nt.BstNB restriction sites were cloned into the pEGFP-N1 19 bases immediately upstream and downstream of the TATA box through PCR amplification with sequence specific primers (Table 2.5) following protocol 5.3.2. The pEGFP-BstNB plasmid was

72

digested (40 μg) with Nt.BstNB at 55 °C for 2 h (500 μL) following protocol 5.3.3. The reverse

complement to the 34 bp DNA insert fragment was added to the digestion reaction (25 μL of a

100 μM solution) and annealed in TAE/Mg^{2+} buffer (0.04 M tris-acetate, 1 mM

ethylenediaminetetraacetic acid (EDTA), 12.5 mM magnesium acetate, pH ~7.4) by cooling the

solution from 95 to 12 °C over 10 min. The complete plasmid digest reaction then was gel

purified with the E.Z.N.A. Gel Extraction (Omega) kit. The noncaged and caged TATA box

sequences (T^{mut} and T^0–T^3, Table 2.6) were 5′ phosphorylated with T4 polynucleotide kinase at

50 μM (50 μL) and were ligated (10 μL insert, 60 μL reaction) into the purified plasmid

following protocol 5.3.4. The ligated product was column purified with the E.Z.N.A. Cycle Pure

(Omega) kit and quantified with a Nanodrop spectrometer

Table 2.5: Primer sequences used in the cloning of the pEGFP-BstNB plasmid.
Restriction sites are underlined.

Strand	Sequence (5'− 3')
Nt.BstNB forward	TATATAAGACC<u>GAGTC</u>CCGTCGTCAGATCCGC
Nt.BstNB reverse	AGCAGAGCTGGTTTAACGC<u>GACTC</u>GCCCAACCGC
TATA insert rev. comp.	ACGGTTGGAGGTCTGTGTGAGCAGAGCTGGTTTA

Table 2.6: Sequences of the synthetic caged TATA box oligonucleotides.
Sequences were designed for insertion into the TATA promoter region of the pEGFP-BstNB plasmid. NPOM-caged thymidine residues are underlined, bolded, and highlighted red ("**T**"). Caged DNA oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5'− 3')
T^{mut}	TAAACCAGCTCTGCTCACACAGACCTCCAACCGT
T^0	TAAACCAGCTCTGCTTATATAGACCTCCAACCGT
T^1	TAAACCAGCTCTGCTTA**T**ATAGACCTCCAACCGT
T^2	TAAACCAGCTCTGCT**T**ATA**T**AGACCTCCAACCGT
T^3	TAAACCAGCTCTGCT**T**A**T**A**T**AGACCTCCAACCGTU

Melting Temperatures. The T_M of each TATA box duplex was measured using a CFX96 Touch Real Time PCR Detection System (Bio-Rad). TATA box DNA duplexes (20 µL, 1 µM) were incubated in TAE/Mg^{2+} buffer (0.04 M tris-acetate, 1 mM EDTA, and 12.5 mM magnesium acetate) and annealed over a temperature gradient from 95 to 4 °C over 10 min. The samples were then heated in the presence of SYBR green (1 µL of 20X SsoFast EvaGreen Supermix, Bio-Rad) from 0 to 100 °C at a rate of 0.5 °C/min with a dwell time of 10 sec, and the fluorescence was measured every 0.5 °C. The T_M was determined by the maximum of the first derivative of the fluorescence vs temperature plot. Standard deviations were calculated from three individual experiments.

Analysis of the Effect of UV-A Exposure on Cell Viability. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into a 96-well plate and grown to ~70% confluence within 24 h following protocol 5.4.1. Cells were then irradiated for 0–20 min (UV transilluminator, 365 nm, 25 W) following protocol 5.4.3, then incubated for 24 h. After the overnight incubation, 150 µL of the cellular media was removed and 50 µL CellTiter-Glo (Bio-Rad) reagent was added. Chemiluminescence was measured on a BioTek Synergy 4 plate reader after 10 min. Standard deviations were calculated from three individual experiments.

Light Activation of EGFP Expression. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 96-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. The cells were transfected with pEGFP-BstNB (T^{mut}, T^0–T^3, 150 ng/well) and pDsRed-N1 (300 ng/well) plasmids using bPEI following protocol 5.4.2. The following day cells were either irradiated for 5 min (UV transilluminator, 365

nm, 25 W) or kept in the dark following protocol 5.4.3, then incubated for 48 h. The cells were imaged on a Zeiss Observer Z1 microscope (5X magnification) following protocol 5.4.5.

Flow Cytometry Analysis. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 24-well and grown to ∼70% confluence within 24 h following protocol 5.4.1. The cells were transfected with pEGFP-BstNB (T^0–T^3, 750 ng/well) and pDsRed-N1 (1500 ng/well) plasmids using bPEI following protocol 5.4.2. The following day cells were either irradiated for 5 min (UV transilluminator, 365 nm, 25 W) or kept in the dark following protocol 5.4.3, then incubated for 48 h. The cells were trypsinized and resuspended in DMEM media. Flow cytometry was performed on a FACSCalibur (Becton-Dickinson) instrument (EGFP; 488 nm argon laser, 530/50 nm BPF. DsRed; 633 nm excitation argon laser, 671/50 nm BPF). Cells were gated for EGFP and DsRed fluorescence (above $10^{2.5}$ RFUs) then analyzed with Cellquest Pro Software until 20,000 cells had been counted for each condition tested. For each of the triplicates, the data were averaged, normalized to the T^0-noncaged control, and standard deviations were calculated. p values were calculated from unpaired t tests.

Spatial Activation of Gene Expression. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 96-well plates and grown to ∼70% confluence within 24 h following protocol 5.4.1. Cells were transfected with the T^3-caged EGFP (50 ng) and pDsRed-N1 (300 ng) plasmids using lipofectamine transfection reagent for 4 h following protocol 5.4.2. Localized irradiation was performed with a Zeiss Observer Z1 microscope for 30 sec following protocol 5.4.3. The cells were then incubated for 48 h and imaged on a Zeiss Observer Z1 microscope (5X magnification) following protocol 5.4.5.

Plk3 Phenotypic Cell Assay. See Section 5.3 for specific information on molecular biology techniques and Section 5.4 for specific information on cell culture techniques. The Plk3 gene was fused to the C terminus of EGFP in pEGFP-Bst to form the pEGFP-Bst-Plk3 plasmid. The pEGFP-Bst-Plk3 plasmid was constructed by amplifying Plk3 from Drosophila Plk3 cDNA (ATCC) with specific primers, and the pEGFP-BstNB plasmid was PCR amplified for Plk3 insertion (Table 2.7) following protocol 5.3.1. Both PCR products were digested with *SpeI* and *MfeI* and ligated together following protocol 5.3.3. The caged pEGFP-Bst-Plk3 plasmid was constructed using the same protocol as described above. HeLa cells were passaged into 4 well chamber slides and grown to 70% confluency within 24 h following protocol 5.4.1. The cells were transfected with noncaged and caged pEGFP-Bst-Plk3 plasmids (150 ng) using lPEI following protocol 5.4.2. The cells were irradiated with a UV transilluminator (365 nm, 5 min, 25 W) following protocol 5.4.3, and were incubated for 48 h. The cells were fixed and stained with DAPI (blue) and Rhodamine-phalloidin (red) fluorescent dyes following protocol 5.4.4, then imaged on a Zeiss Z1 Observer microscope (63X magnification) following protocol 5.4.5.

Table 2.7: Primer sequences for Plk3 amplification and insert into pEGFP-BstNB.

Primer	Sequence (5'− 3')
Plk3 forward	CGTAAGCAATTGGACTTCTTTACC
Plk3 reverse	CCTACGACTAGTCTAGGCTGGGCT
pEGFP-BstNB forward	GGAACTAGTCAGCGGCCGCGACTCT
pEGFP-BstNB reverse	CCTACGCAATTGCTTGTACAGCTCGTC

Plk3 Caspase 3 Activity Assay. See Section 5.4 for specific information on cell culture techniques. HeLa cells were passaged into 24-well plates and grown to ~70% confluency within 24 h following protocol 5.4.1. The cells were transfected with noncaged and caged pEGFP-Bst-Plk3 plasmids (150 ng) using lPEI following protocol 5.4.2. The cells were irradiated with UV

light (365 nm, 5 min, 25 W UV transilluminator) following protocol 5.4.3, and were incubated for 48 h. The cells were lysed and total protein was quantified with a Nanodrop spectrometer following protocol 5.4.8. HeLa cell protein extract (100 μg) was incubated with 50 μM Caspase-3 substrate (Ac-DEVD-AFC, Calbiochem) in activity buffer (50 mM HEPES, 150 mM NaCl, 50 mM $MgCl_2$, 250 μM EDTA, 10% sucrose, 0.1% CHAPS, pH 7.2) at 37 °C for 20 h. The fluorescence was measured on a BioTek Synergy 4 plate reader (400/505 nm). For each of the triplicates, the data were averaged, normalized to the T^0-noncaged control, and standard deviations were calculated. *p* values were calculated from unpaired *t* tests.

Zebrafish Injections. See Section 5.5 for specific information on zebrafish techniques. Embryos were microinjected with 50 pg of the plasmid in 1 nL following protocol 5.5.2. Embryos were then irradiated following injection (typically at the 4- or 8-cell stage) for 2 min with a 365 nm UV transilluminator. Late-stage irradiation experiments were performed at 8 hpf. Imaging was performed at 24 and 48 h following protocol 5.5.3. EGFP scores were calculated with embryo counts of [(EGFP positive/alive)·100]. For each of the replicates, the data were averaged, and standard deviations were calculated. *p* values were calculated from unpaired *t* tests.

2.5   Optical Control of Alternative Splicing through Caged Splice-Switching Oligonucleotides

The spliceosome machinery is composed of several proteins and multiple small RNA molecules that are involved in gene regulation through the removal of introns from pre-mRNAs in order to assemble exon-based mRNA containing protein-coding sequences. Splice-switching oligonucleotides (SSOs) are genetic control elements that can be used to specifically control the expression of genes through correction of aberrant splicing pathways. A current limitation with SSO methodologies is the inability to achieve conditional control of their function paired with high spatial and temporal resolution. This limitation was addressed through site-specific installation of light-removable nucleobase-caging groups as well as photocleavable backbone linkers into synthetic SSOs. This enables optochemical OFF → ON and ON → OFF switching of their activity and thus precise control of alternative splicing. The use of light as a regulatory element allows for tight spatial and temporal control of splice switching in mammalian cells and animals.

Alternative splicing (AS) plays a large role in the regulation of gene expression,[172] as the spliceosome is responsible for processing of pre-mRNA into coding sequence by eliminating introns and assembling the correct exons.[173] The splicing of pre-mRNA in humans was documented in the mid 1980s,[174] and the field has dramatically expanded since. It has been estimated that 95% of all human genes show AS pathways, highlighting the importance of spliceosome activity in RNA regulation.[175] Additionally, up to 50% of mutations associated with genetic disorders result in altered pre-mRNA splicing pathways,[176] linking aberrant mRNA splicing to a wide variety of disease states, including cystic fibrosis and Duchenne muscular dystrophy.[177] Splice-switching oligonucleotides (SSOs) act by hybridizing to the pre-mRNA and blocking splice sites in a sequence-specific manner, which prevents interaction with components of the spliceosome such that splicing pathways are altered. SSOs are commonly used tools to

control gene expression via alternative mRNA splicing pathways through the removal of aberrant introns or exons as well as exon retention or skipping.[23] For example, the coding region of a protein can be interrupted via a mutant intron gene (Figure 2.17A). Through AS pathways, the aberrant sequences can be blocked with SSOs, allowing for corrected exon splicing and resulting in the expression of a functional gene product (Figure 2.17B). Besides applications as research tools, SSOs are proposed for use as therapeutic agents to correct splice mutations,[178] and have been applied to genetic disorders for the initiation of several human clinical trials[179] because these oligonucleotides are able to restore proper gene function in disease states.
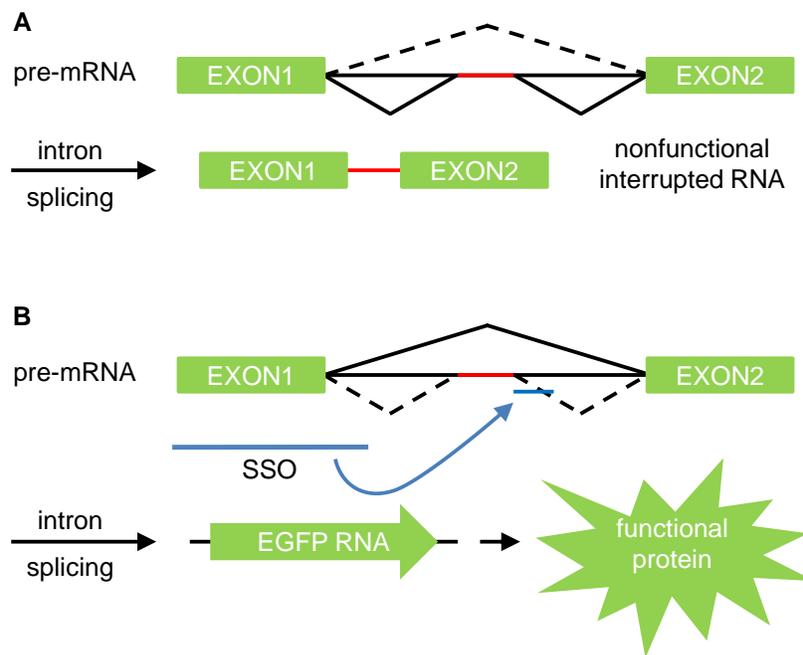
Figure 2.17: Schematic of splice-switching oligonucleotide function.
Pre-mRNA processing and regulation of gene expression through alternative splicing (AS) are shown. Active splicing pathways are indicated with solid black lines and AS pathways are indicated with dashed lines. The mutant intron is shown in red, and the activated SSO is shown in blue. A) In the absence of the SSO, a mutant intron from the pre-mRNA is present in the mRNA sequence and disrupts formation of a functional gene product. B) The SSO binds to the target site, blocks the spliceosome from interacting with the target site, and creates an AS pathway that will remove the mutant intron from the mRNA allowing for the expression of functional protein (e.g., EGFP).

However, no conditional control of SSO activity has been reported, and neither temporal nor spatial regulation of SSO function and AS has been achieved. The specific timing and location of AS play very important roles in many regulatory pathways in cells across the animal kingdom, for example during neuronal cell differentiation in the human brain.[180] Similar spatial and temporal activity of AS has been observed in fruit flies,[181] mice,[182] zebrafish,[183] and nematodes,[184] demonstrating that distinct AS patterns are responsible for many essential biological functions. Methods to accurately perturb the spatial and temporal patterns of AS are critical for understanding the complex mechanisms that underlie gene regulatory control by the spliceosome.[185] However, traditional SSOs are constitutively in an ON state, i.e., the splicing pathways are immediately altered and cannot be controlled with any spatial and temporal resolution, precluding precise investigation of AS in cells and multicellular organisms. In order to achieve conditional control of alternative splicing, optochemical tools were developed based on the introduction of caged nucleobases and photocleavable linkers into SSOs. This approach enabled both optochemical activation and deactivation of splicing pathways, leading to efficient two-directional genetic control through OFF → ON as well as ON → OFF light switches. These optochemical regulation tools provide conditional control of SSO activity with high resolution in cellular and embryonic environments, providing spatial and temporal capabilities for dissection of AS pathways as well as the general control of gene function.

Before applying the light-controlled SSO approach to a complex biological system, such as a developing zebrafish embryo, the β-globin intron 1 was selected as a proof-of-principle target, since it contains an aberrant splice site that contributes to the genetic blood disorder β-thalassaemia and that can be corrected with SSOs.[186, 187] HeLa cell lines stably expressing enhanced green fluorescent protein (EGFP) and luciferase (Luc) genes interrupted with the

mutant β-globin intron (HeLa:EGFP654, HeLa:Luc705)[188] were used to analyze the optochemical regulation of SSOs and the ability to control AS pathways with light. The reporter genes contain an altered coding region with an aberrant splice site to impair functional fluorescent protein expression until it is blocked by a sequence-specific SSO. The activity of control SSOs was analyzed using TAMRA-labeled oligonucleotides to track transfection, for both a noncaged SSO and a five-base-mismatched oligo (Table 2.8). TAMRA fluorescence showed successful transfection of the labeled oligos, and EGFP expression was observed in the HeLa:EGFP654 cell line only in the presence of the positive control SSO (Figure 2.18A). No detrimental effect on the SSO-controlled expression of EGFP in the presence of UV light was detected, ensuring that the reporter construct itself was not UV responsive (Figure 2.18B).
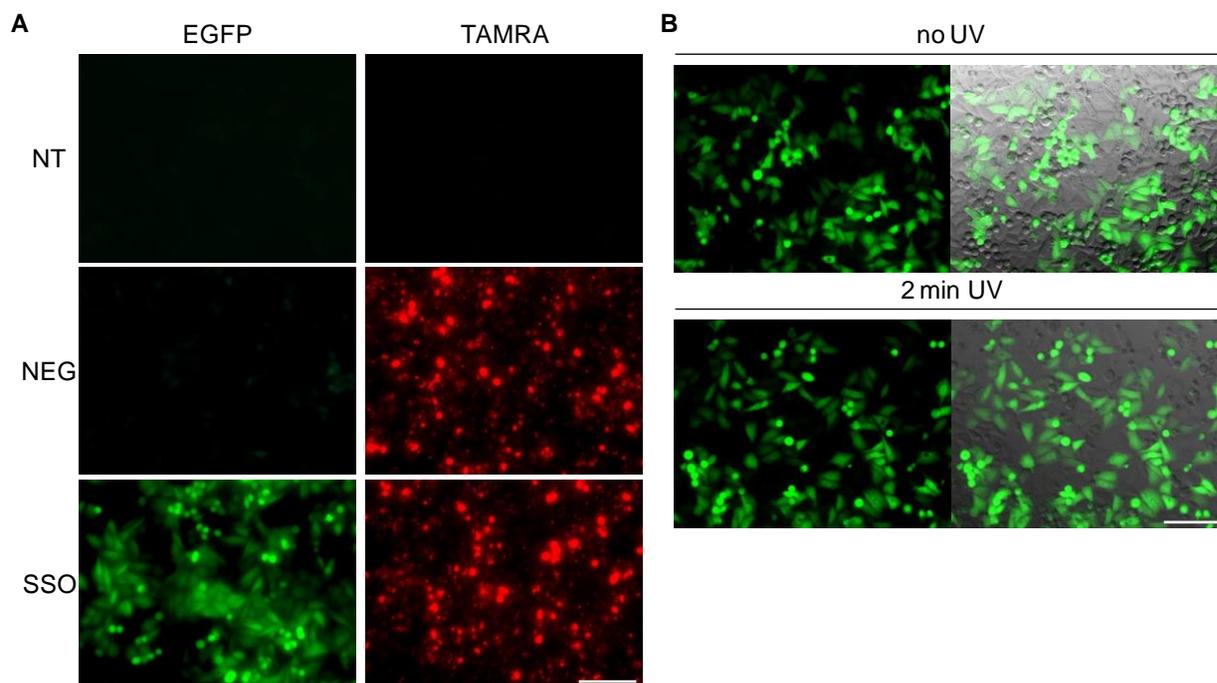
Figure 2.18: Confirmation of HeLa:EGFP654 reporter function.
A) HeLa:EGFP654 cells were passaged into 96-well plates and transfected with noncaged TAMRA labeled SSOs (50 nM) using X-tremeGENE siRNA reagent (Roche). The EGFP and TAMRA fluorescence were observed at 24 h on 20X magnification. NT = no oligo treatment, NEG = mismatch negative control. B) HeLa:EGFP654 cells were transfected with the noncaged SSO (50 nM) using X-tremeGENE siRNA reagent (Roche), then UV irradiated on a

transilluminator (365 nm, 6.3 mW/cm2), and imaged for EGFP fluorescence on 20X magnification 24 h after UV exposure. Scale bars indicate 0.2 mm.

While the HeLa:Luc705 cell line showed a significant 20-fold increase in luciferase when transfected with the noncaged SSO, a decrease of luciferase expression was observed for cells transfected and exposed to short UV exposures (Figure 2.19). Therefore, the HeLa:EGFP654 cell line was utilized for all subsequent analysis of the optical regulation of splice switching in mammalian cells. In order to optochemically regulate SSO activity, oligonucleotides with light-responsive modifications were synthesized containing 2′OMe nucleotides and phosphorothioate linkages, which are highly effective oligonucleotide chemistries for cellular applications of SSOs.[189]
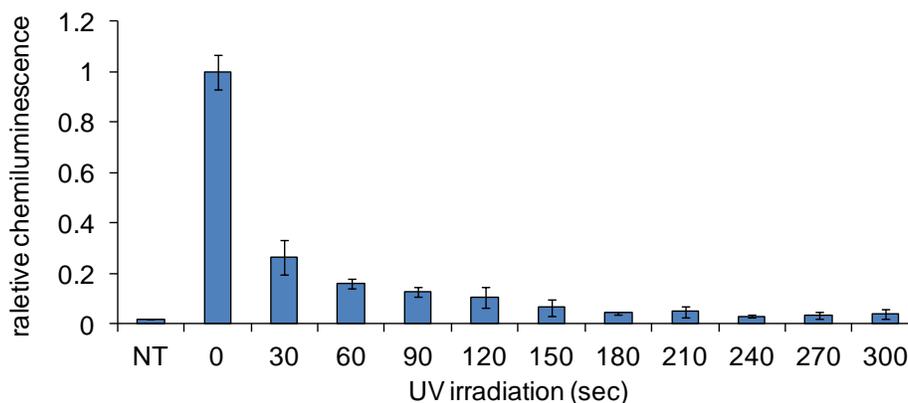
Figure 2.19: Splice-switch driven luciferase expression with UV irradiation.
HeLa:Luc705 cells were transfected with the noncaged SSO (50 nM) using X-tremeGENE siRNA reagent (Roche), then UV irradiated on a transilluminator (365 nm, 6.3 mW/cm2) for 30 sec bursts up to 5 min. A luciferase bright glow assay was performed 24 h after UV exposure. Chemiluminescence was normalized to a no UV control and error bars represent the standard deviation of experimental triplicates.

First, a nucleobase-caged SSO was engineered that contained light-removable protecting groups that inhibit hybridization to the complementary pre-mRNA target. A 2′OMe-NPOM-caged uridine phosphoramidite was synthesized (Figure 2.20A) and incorporated into the SSO at

four positions throughout the sequence, as indicated in Table 2.8, to achieve full inhibition of duplex formation. The light-activated SSO (LASSO) was synthesized using phosphoramidite chemistry, gel extracted to obtain a high-purity full-length oligomer (Figure 2.20B and Figure 2.21), and subsequently transfected into the HeLa reporter cell line.
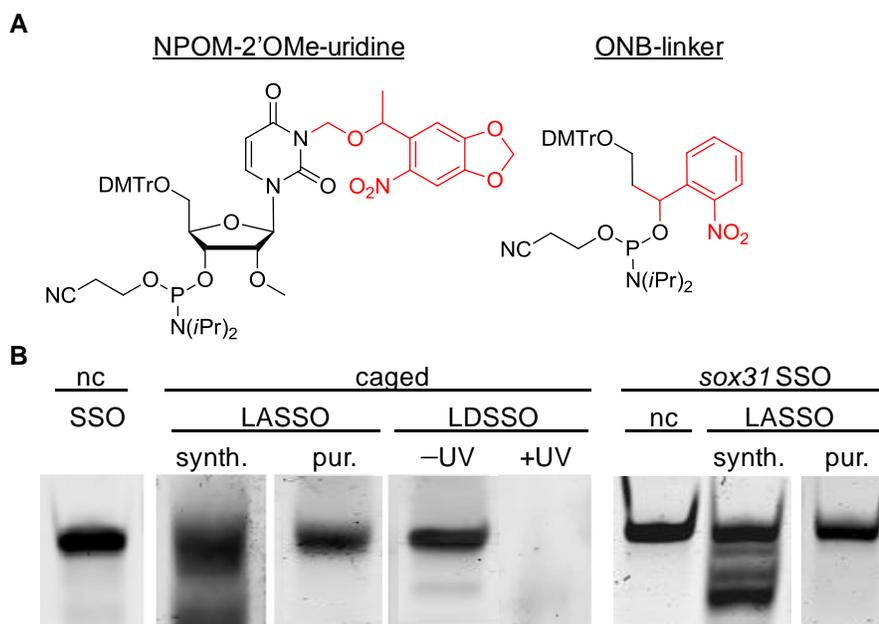
Figure 2.20: Splice-switching oligonucleotide synthesis.
A) Chemical structures of the two specialized phosphoramidites used for the synthesis of photochemically regulated oligonucleotides. B) Gel analysis of 10 pmol of each oligonucleotide synthesis. The abbreviation "nc" stands for noncaged. The abbreviations "synth." and "pur." represent the crude synthesis and purification products, respectively.
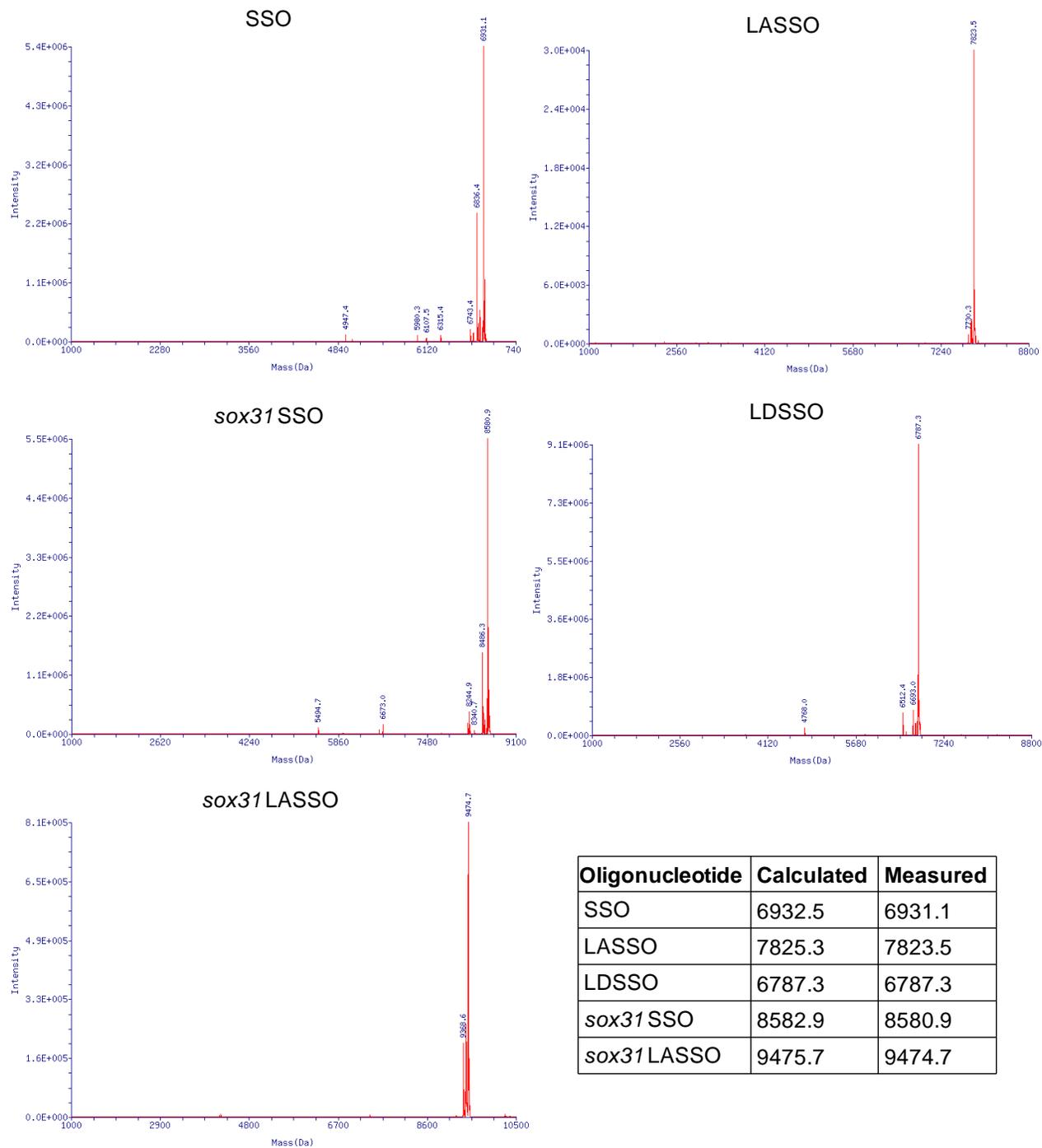
Figure 2.21: Analysis of splice-switching oligonucleotides.
ESI mass spectra of synthesized oligonucleotides and tabulated results of calculated and measured masses (in Da). No photolysed products (loss of 1-4 NPOM groups or linker cleavage) were detected. Electrospray ionization mass spectrometry (ESI-MS) was performed by Novatia (Newtown, PA).

Oligonucleotide	Calculated	Measured
SSO	6932.5	6931.1
LASSO	7825.3	7823.5
LDSSO	6787.3	6787.3
sox31 SSO	8582.9	8580.9
sox31 LASSO	9475.7	9474.7

The LASSO was designed to be inactive when transfected into cells until briefly exposed to UV light, thus enabling efficient OFF → ON photoswitching for aberrant splice correction (Figure 2.22A and B). In the absence of UV irradiation, no EGFP expression was observed, but after a short UV exposure the AS pathway was activated through SSO decaging, leading to removal of the mutant intron and recovery of EGFP reporter gene expression (Figure 2.22D). Only cells that were UV-irradiated exhibit splice-switching activity and EGFP expression.
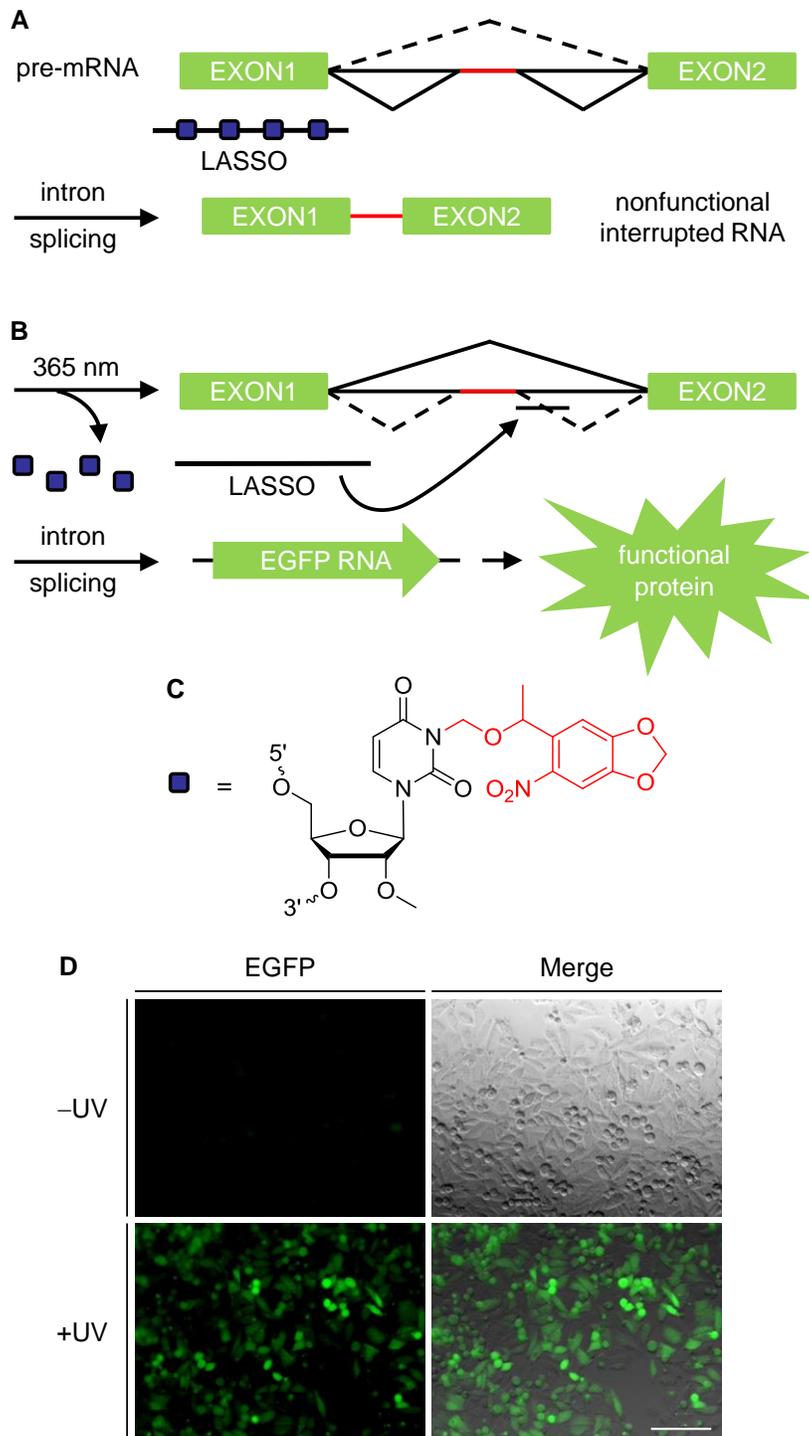
Figure 2.22: Optochemical regulation of splice switching with the LASSO (OFF → ON). Schematics for the two RNA splicing pathways (black) and the mutant intron (red) are shown. A) In the absence of irradiation, the LASSO does not bind to the target site due to inhibition of base hybridization from the nucleobase caging groups. B) UV decaging of the LASSO enables conditional splice correction of the aberrant mutant intron and activation of gene expression. C) Dark blue boxes represent 2′OMe-NPOM-caged uridine residues, with the light-removable group indicated in red. D) HeLa:EGFP654 cells were transfected with the LASSO, irradiated for 2 min

or kept in the dark, and imaged for EGFP expression after 24 h. Images of the EGFP channel (left) and the EGFP channel merged with a brightfield image (right) are shown. Scale bar indicates 0.2 mm.

In order to maximize light activation, the parameters for UV irradiation and LASSO concentration were further analyzed, revealing a 200 nM concentration and a 2 min irradiation time as being optimal for light-activated splice correction (Figure 2.23). Importantly, no background leakiness of the LASSO before irradiation was observed, confirming excellent OFF to ON switching behavior.
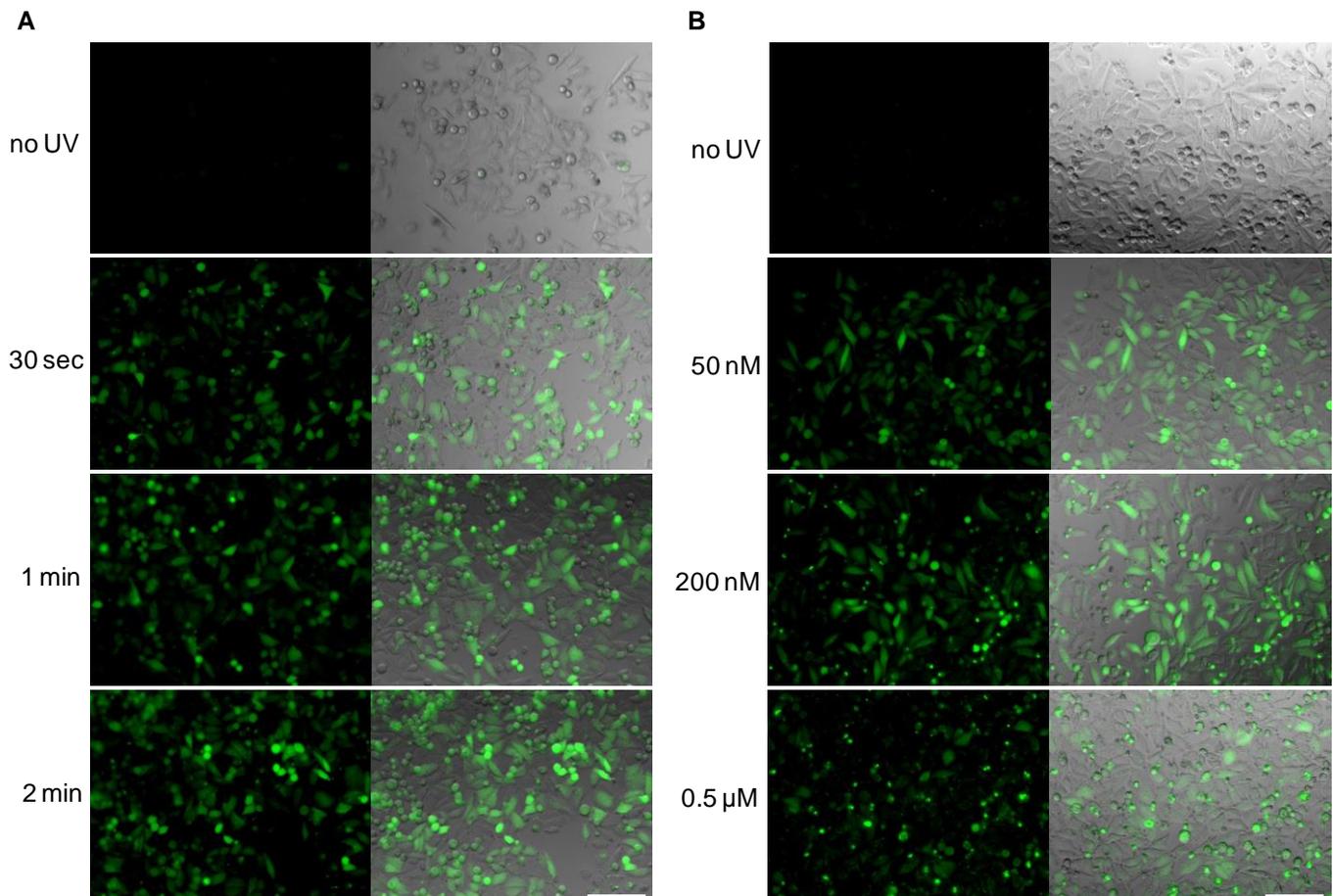
Figure 2.23: Optimization of photochemical LASSO activation in HeLa:EGFP654 cells.
A) Cells were transfected of the LASSO (50 nM) using X-tremeGENE siRNA reagent (Roche), followed by different irradiation times (UV transilluminator, 365 nm, 6.3 mW/cm2) and imaging after 24 h (20X magnification). B) The 2 min exposure time was then used across a range of SSO transfection concentrations. Scale bars indicate 0.2 mm.

As a second approach, a light-cleavable oligonucleotide was synthesized to enable ON →

OFF photoswitching of gene expression using an *ortho*-nitrobenzyl (ONB) linker (Figure

2.20A). The ONB group was incorporated at two locations internally within the SSO sequence,

as indicated in Table 2.8, both to maintain binding with the target site and to ensure complete

cleavage of the full-length oligonucleotide through a brief UV exposure. This light-

deactivated SSO (LDSSO) was designed to be active in the absence of UV irradiation, but it can

be deactivated with UV light, inducing oligonucleotide cleavage (Figure 2.24A and B). The

successful synthesis and UV-dependent cleavage of the LDSSO was confirmed by gel analysis

prior to cellular applications (Figure 2.20B and Figure 2.21). Transfection of the LDSSO

confirmed that the installation of ONB linkers did not interfere with the SSO activity and that

splice-switching correction for the expression of EGFP occurred in the absence of UV exposure

(Figure 2.24D). In contrast, cells that were UV-irradiated showed no EGFP expression due to

deactivation of splice-switching oligonucleotide as a result of light-induced oligonucleotide

fragmentation. Importantly, splice correction with the LDSSO was completely deactivated

through UV irradiation, demonstrating an excellent ON → OFF switching behavior of the SSO

pathway. Thus, in conjunction with the LASSO described above, both optical activation and

deactivation of splice-switching pathways are possible.
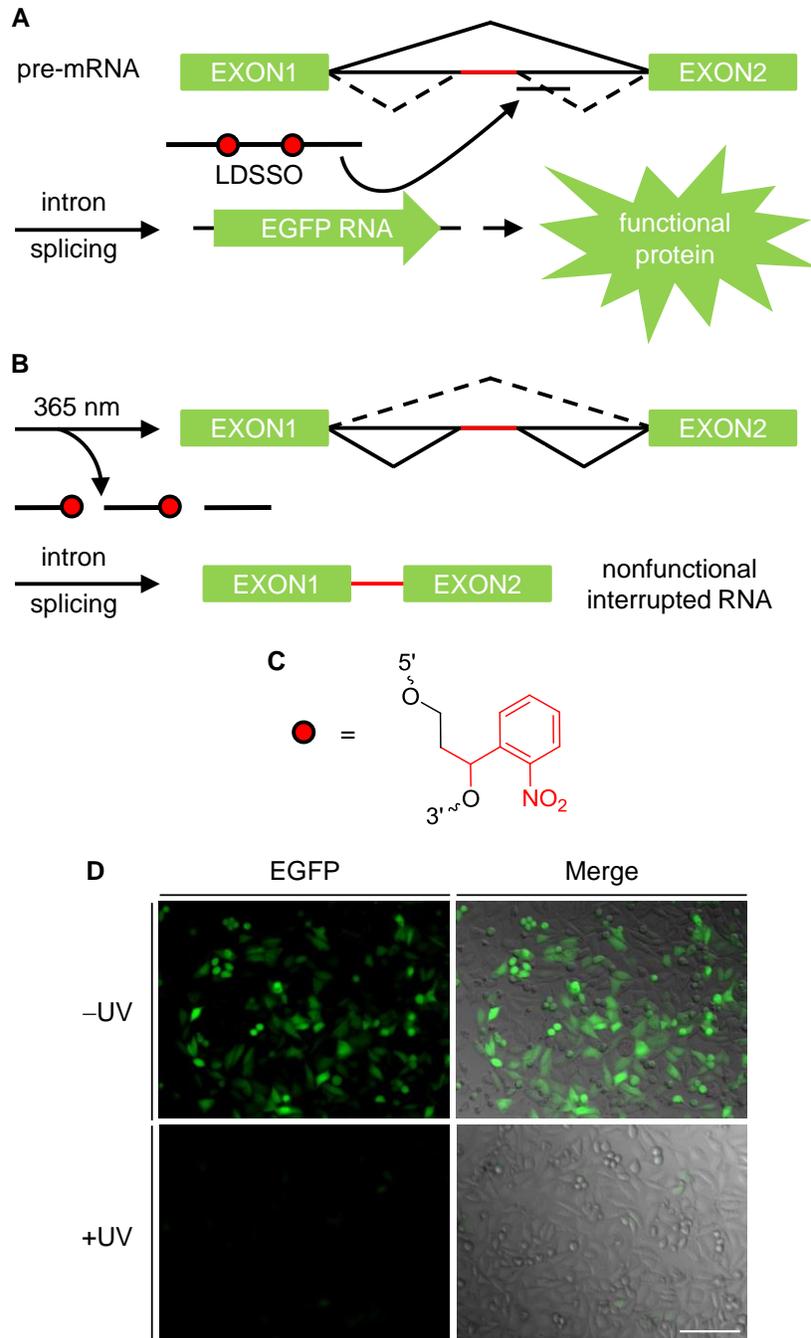
Figure 2.24: Optochemical regulation of splice switching with the LDSSO (ON → OFF). Schematics for the two RNA splicing pathways (black) and the mutant intron (red) are shown. A) In the absence of irradiation, the LDSSO binds to the target site for aberrant splice correction and activation of gene expression. B) UV cleavage of the LDSSO deactivates splice correction, and gene expression is interrupted by the mutant intron. C) Red circles represent ONB linker residues, with the photocleavable group indicated in red. D) HeLa:EGFP654 cells were transfected with the LDSSO, irradiated for 2 min or kept in the dark, and imaged for EGFP expression after 24 h. Images of the EGFP channel (left) and the EGFP channel merged with a brightfield image (right) are shown. Scale bar indicates 0.2 mm.

The effect of optochemical regulation of SSO activity and aberrant intron splicing on EGFP expression was then quantified through flow cytometry. Conditions for cell transfection and fluorescent cell counting were optimized both for the LASSO and LDSSO (Figure 2.25), followed by analysis of the optical OFF → ON and ON → OFF switching with the light-regulated SSOs and comparison to a noncaged positive control (Figure 2.26). The EGFP reporter expression for the LASSO was fully restored, showing nearly identical expression levels as that of the control after a 2 min UV irradiation for light-activation of SSO function. The light-deactivated SSO showed similarly high levels of EGFP expression as that of the noncaged SSO in the absence of UV irradiation but greatly reduced expression levels after UV exposure. Both SSO light switches exhibited >10-fold changes in EGFP expression between the OFF and ON states. These findings are in agreement with the observations from cellular micrographs presented in Figure 2.22 and Figure 2.24, validating the application of photocaged bases and photocleavable linkers for optochemical activation and deactivation of aberrant splicing pathways.
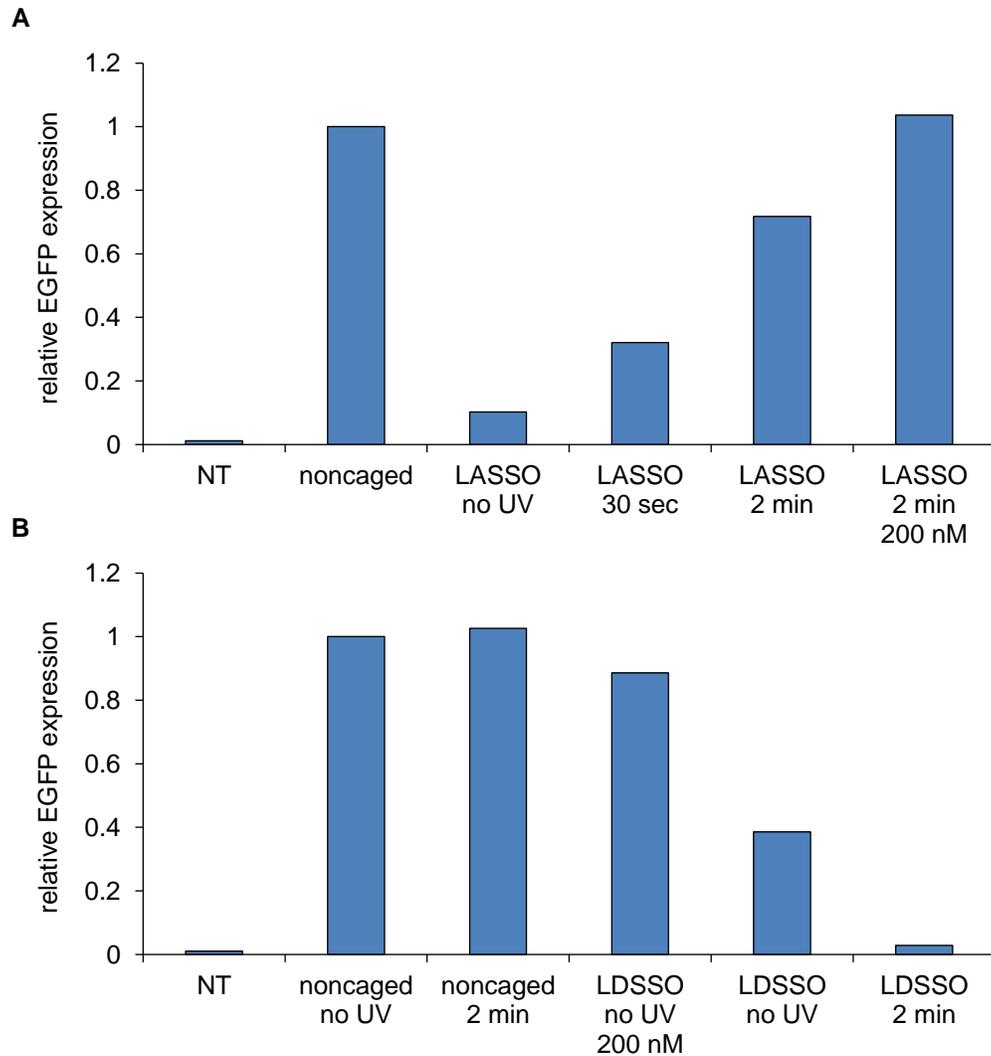
Figure 2.25: Optimization for flow cytometry quantification of splice-switching. HeLa:EGFP654 cells were transfected in 6-well plates with the LASSO (A) and LDSSO (B) oligonucleotides (50 nM unless otherwise indicated), irradiated with a UV transilluminator (365 nm, 6.3 mW/cm2), and analyzed for EGFP expression after 24 h using a FACSCalibur (Becton-Dickinson) flow cytometer (20,000 gating events). The gated EGFP positive cells were normalized relative to the noncaged control.
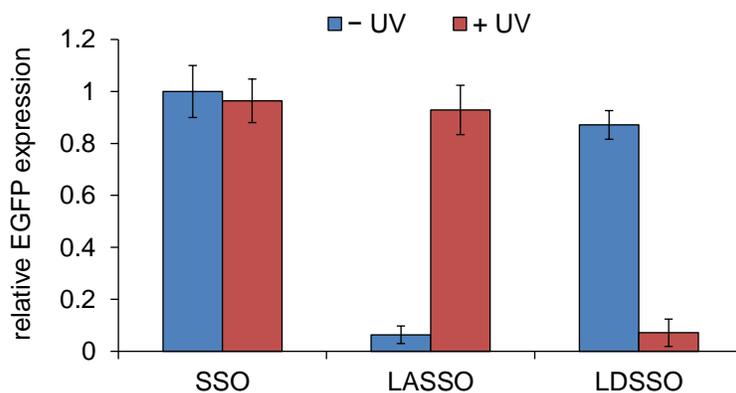
Figure 2.26: Flow cytometry quantification of optically regulated splice-switching. HeLa:EGFP654 cells were transfected with the SSOs, irradiated for 2 min or kept in the dark, and analyzed for EGFP expression after 24 h. The gated EGFP positive cells were normalized to the noncaged control. Error bars represent standard deviations from three independent experiments through counts of 20,000 cells each.

The ability to specifically control the location of SSO optochemical regulation through localized UV illumination was then investigated. Irradiations for the LASSO were performed using irradiation masks as well as microscope optics. When UV irradiation was applied in a spatially defined area in conjunction with the light-activated SSO, only the cells that were exposed to 365 nm light exhibited EGFP expression through an OFF → ON light switch of SSO activity (Figure 2.27). The time of irradiation was decreased from 2 min to 30 s for localized irradiations because these were performed through focusing microscope optics using a Xe/Hg lamp. The tight spatial control conferred with this method can potentially be used to analyze splice switching and aberrant intron correction in small subsets of cellular populations as well as within defined locations in multicellular organisms.
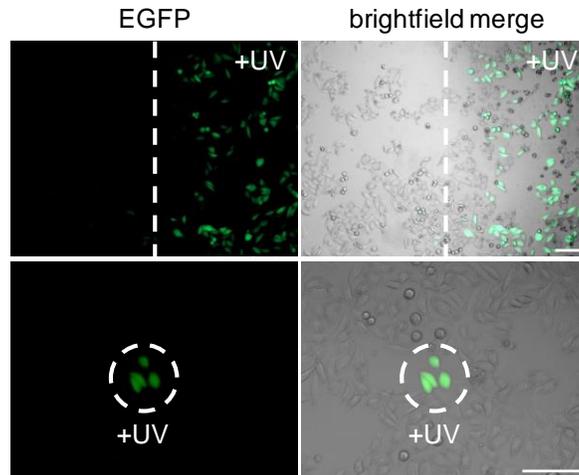
Figure 2.27: Spatial control of EGFP expression using the LASSO.
Localized irradiations were performed using a vertical mask (top) or a partially closed microscope shutter (bottom). Cells were imaged after 24 h. Images of the EGFP channel (left) and the EGFP channel merged with a brightfield image (right) are shown. Scale bars indicate 0.2 mm.

In order to translate the methodology for the optical regulation of splice-switching from cell culture experiments to multicellular organisms, an endogenous gene in zebrafish was targeted to demonstrate conditional control of RNA splicing. The zebrafish embryo was selected as a model system because it has been shown that UV-A exposure does not affect zebrafish embryonic development, hatch rate, mortality, or global gene expression,[190] making this translucent and *ex utero* developing animal highly suited for photoactivation studies with microinjected caged oligonucleotides.[170] A light-activated splice-switching oligonucleotide was designed targeting *sox31* (also known as *sox19b*), a member of the B1 Sox gene family that is responsible for many critical processes during zebrafish development, both before and after the midblastula transition (MBT).[191] During the blastula stage, *sox31* is expressed throughout the blastoderm, where it then assists in epiboly during the gastrulation stage.[192] Splice variants of *sox31* have been previously generated with morpholino oligonucleotides (MO) targeting cryptic splice sites in *sox31*.[193] Alternative splicing-based inhibition of *sox31* resulted in normal development during the first 3 hpf, followed by arrested embryonic development and loss of

93

epiboly for zebrafish that failed to undergo gastrulation. Although these findings were made with MOs, there is precedence for the use of 2′OMe-modified oligonucleotides as antisense agents in zebrafish.[194] As such, a 2′OMe *sox31* LASSO was synthesized with NPOM-caged uridine residues incorporated at four positions throughout the sequence, as indicated in Table 2.8, and purified for injection into zebrafish embryos (Figure 2.20B). The *sox31* LASSO was designed to be inactive in the absence of UV exposure, allowing for correct RNA splicing and functional protein production until a brief UV exposure removes the nucleobase caging groups to activate the LASSO, which will then bind the RNA target site to initiate AS for the production of a nonfunctional interrupted RNA (see Figure 2.22). To this end, embryos were injected with the oligonucleotides (5 ng) and either irradiated (365 nm, 2 min) or kept in the dark. At 8 h postfertilization (hpf), the embryos were imaged during gastrulation at ~75% epiboly (Figure 2.28A), and the frequency of embryos exhibiting epiboly defects was determined (Figure 2.28B). The noninjected embryos were not affected by UV exposure, as no significant changes in development were observed. Injection with the noncaged positive control *sox31* SSO showed developmental arrest and high frequency of the no epiboly phenotype. A negative control scrambled 2′OMe oligonucleotide shows low levels of developmental arrest, slightly above those observed for noninjected embryos. The *sox31* LASSO showed normal epiboly formation and development through gastrulation for the injected embryos in the absence of UV irradiation, similar to the negative control. However, activation of the LASSO with UV irradiation inhibited *sox31* expression and induced epiboly failure to the same high levels as that observed for injection of the noncaged control. Thus, the conditional control of splice switching of an endogenous gene was successfully demonstrated in a living animal.
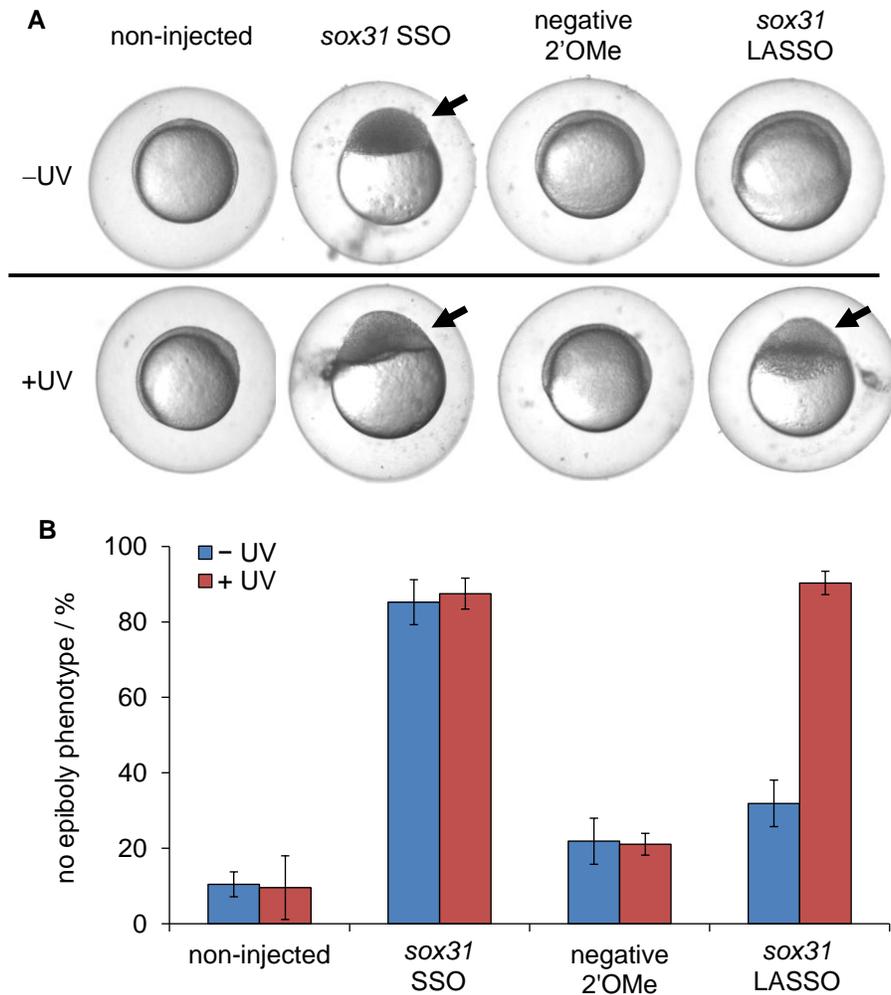
Figure 2.28: Optical regulation of splice switching in living zebrafish embryos.
For a schematic of RNA splicing and LASSO activity refer to Figure 2.22. A) Zebrafish embryos
were injected with the SSOs (5 ng) and imaged at 8 hpf. The major phenotypes observed from
treatment with the light-activated and control SSOs, both without (top) and with (bottom) UV
irradiation, are shown. Arrows indicate the absence of epiboly formation. B) Zebrafish embryo
scores were determined for the frequency of no epiboly formation with and without UV exposure
(365 nm, 2 min). Error bars represent standard deviations from three independent experiments.
$N = 18$–44.

Additionally, UV exposures at different time points from 1 to 7 hpf were performed to
determine temporal requirements for the activity of *sox31* during early embryonic development.
Only embryos irradiated after 4 hpf (after the MBT) showed a distinct decrease in epiboly
defects (Figure 2.29), supporting functional *sox31* as a requirement for epiboly.[193] Thus,

the *sox31* LASSO enabled OFF → ON photoswitching of a dominant negative splice variant, thereby inhibiting *sox31* and other SoxB1 genes in the developing zebrafish embryo.
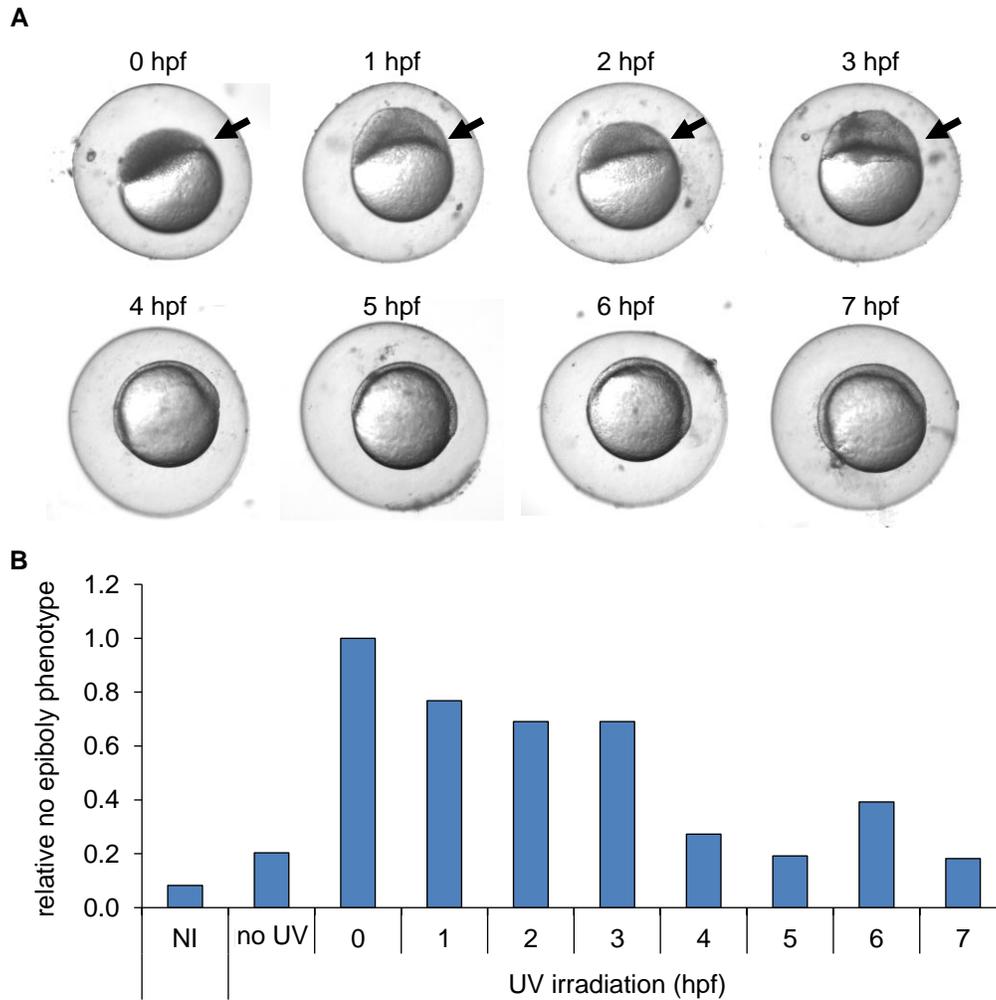
Figure 2.29: *Sox31* LASSO light-activation time points.
Zebrafish embryos were injected with the *sox31* LASSO (5 ng) and UV irradiated (365 nm, 2 min) at 1-7 hpf. A) At 8 hpf, the embryos were imaged to demonstrate the majority phenotype. B) The frequency of embryos exhibiting a no epiboly phenotype was determined. $N = 17$–22.

Alternative splicing (AS) is an important factor in the regulation of gene expression, responsible for the processing of pre-mRNA in many organisms. Misregulation of the spliceosome resulting in aberrant splicing is associated with a wide range of human genetic

disorders, and distinct changes in endogenous alternative splicing can lead to gene expression patterns that are tightly regulated in many biological pathways. Splice-switching oligos (SSOs) provide a method to investigate AS and control splicing pathways through hybridization with the pre-mRNA, which blocks splice sites and prevents interaction with components of the spliceosome, thus modulating gene expression and protein function. In order to address the lack of tools to conditionally regulate SSO function with spatial and temporal resolution, a method to regulate AS pathways with light as an external control element was developed. The installation of caged nucleobases in synthetic SSOs enabled optochemical activation of splice switching and alternative expression pathways in live cells and animals. Light-triggered deactivation of splice switching was demonstrated through the installation of photocleavable linkers into the oligonucleotide backbone. The NPOM-caged nucleobases and ONB-linker groups site-specifically introduced into SSOs were used to develop both OFF → ON and ON → OFF light switches for gene splicing, with excellent switching behavior and no detectable background activity before light exposure. These findings are supported by cellular micrographs as well as florescent cell quantification using an EGFP reporter system in human cells. In addition, spatial control of SSO function in a cellular monolayer was achieved through localized UV exposure, showing distinct splice-switch regulation and activation of the caged SSO in a subset of cells. The conditional control of splice switching was demonstrated with the correction of a mutant intron, but it could be readily applied to other mechanisms of AS, such as mutant exon removal or processing of genes that contain multiple splice variants. To this end, the developed methodology was applied to the regulation of mRNA processing for an endogenous gene in a living organism, by targeting *sox31* splicing in zebrafish embryos. The induction of developmental arrest through the inhibition of epiboly during zebrafish gastrulation was

conditionally controlled with injection of a *sox31* light-activated SSO, successfully demonstrating temporal regulation of AS pathways in a complex animal model. By conveying optochemical regulation to SSOs, a new level of precision was added for conditional control over these gene regulatory tools and potential therapeutic reagents. The developed methodology will aid in the investigation of spatial and temporal mechanisms underlying spliceosome correction, aberrant splice switching, and their function in cells as well as multicellular organisms.

2.5.1 Methods and Materials

Oligonucleotides. The noncaged and negative control SSOs were received as a gift from Rudolph Juliano (UNC). The negative control scramble oligonucleotide was purchased from Ambion. Electrospray ionization mass spectrometry (ESI-MS) was performed by Novatia (Newtown, PA). All oligonucleotide sequences are shown in Table 2.8.

Table 2.8: Oligonucleotides used for photochemically regulated splice-switching.
All oligonucleotides contain phosphorothioate backbone linkages and 2'OMe modified nucleobases. Mismatched base pairs are indicated with an asterisk (*). NPOM-caged 2'OMe uridine residues are underlined, bolded, and highlighted red ("**U**"). Light-cleavable linker substitutions are indicated with a dash (-). Caged oligonucleotides were synthesized according to the general protocol 5.2**Error! Reference source not found.**.

Strand	Sequence (5' → 3')
SSO	GUUAUUCUUUAGAAUGGUGC
NEG	GUA*AUUA*UUUAU*AAUC*GUC*C
LASSO	G**U**UAU**U**CUU**U**AGAA**U**GGUGC
LDSSO	GUUAU-CUUUAGA-UGGUGC
sox31 SSO	AGCCCUUUUCUCAAAACAAACCUGU
sox31 LASSO	AGCCC**U**UUU**U**C**U**CAAAACAAACC**U**GU

98

Cellular Transfection of SSOs. See Section 5.4 for specific information on cell culture techniques. The reporter cell lines (HeLa:EGFP654 and HeLa:Luc705) were passaged into 96-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 50–200 nM of each SSO using 1 µL of X-tremeGENE siRNA reagent for 4 h following protocol 5.4.2.

Fluorescence Imaging of EGFP. See Section 5.4 for specific information on cell culture techniques. Fluorescent imaging was performed after 24 h incubation using a Zeiss Observer Z1 microscope (20X magnification) following protocol 5.4.5. The EGFP signal was normalized to a common setting for fluorescent intensity across all experiments (black = 300; white = 3000; gamma = 0.8) in Zen Pro 2011 imaging software. Fluorescent and brightfield merged images are shown with scale bars.

Luciferase assay. See Section 5.4 for specific information on cell culture techniques. A bright glow assay was performed after 24 h incubation to quantify luciferase activity. Luminescence was read from a white 96-well plate on a BioTek Synergy 4-plate reader, with the gain set at 150 and 1 sec integration per well. The luminescence was normalized to the maximum activation observed, and error bars represent the standard deviation of experimental triplicates.

Optochemical Regulation of SSOs. See Section 5.4 for specific information on cell culture techniques. Cellular irradiations were performed post-transfection with a 365 nm UV transilluminator (6.3 mW/cm^2) for 30 sec to 2 min (EGFP reporter) or 5 min (luciferase reporter) following protocol 5.4.3. Spatially distinct UV irradiations were performed with a transilluminator through precut vertical slits in tinfoil for 2 min, and localized irradiations were performed with a Zeiss Observer Z1 microscope for 30 sec following protocol 5.4.3.

Fluorescent Cell Counting. See Section 5.4 for specific information on cell culture techniques. HeLa:EGFP64 cells were passaged into 6-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 50–200 nM of each SSO using 5 μL of X-tremeGENE siRNA reagent for 4 h following protocol 5.4.2. After transfection, UV irradiations were performed as described above followed by 24 h incubation. Cells were then trypsinized and resuspended in PBS buffer for fluorescent analysis. Flow cytometry was performed on a FACSCalibur (Becton-Dickinson) instrument (488 nm argon laser, 530/50 nm BPF) and analyzed using Cellquest Pro software. Cells were gated for EGFP fluorescence (above $10^{2.5}$ RFUs) and analyzed until 20000 cells had been counted for each condition tested. The frequency of EGFP positive cells (gated/total) was normalized to the noncaged control SSO. Error bars represent the standard deviation of experimental triplicates.

Zebrafish Maintenance and Injections. See Section 5.5 for specific information on zebrafish techniques. Embryos were microinjected at the 1-2 cell stage with 5 ng of the oligonucleotides in 5 nL following protocol 5.5.2. Embryos were then irradiated immediately following injection for 2 min with a 365 nm UV transilluminator and incubated for 8 h. UV exposures were performed at different timepoints between 1-7 hpf. Imaging was performed at 8 hpf following protocol 5.5.3. Phenotype scores were calculated with embryo counts of [(no epiboly/alive) × 100]. For each of the replicates, the data were averaged, and standard deviations were calculated.

3.0   DNA Computation: Photochemical Regulation and Application in Live Cells

In addition to applying photocaging of oligonucleotides to the control of gene expression (as described in Chapter 2.0 the use of caged oligonucleotides in optical control of DNA computation events was analyzed (Sections 3.1 and 3.4). These studies were transitioned from photochemical activation to cellular applications of DNA computation (Sections 3.2, 3.3, and 3.5). The goals of these projects were to construct nucleic acid networks that utilize light as an input and to implement logic gates that perform computation functions in living cells. In the latter example, the DNA logic gates acted as biosensors for the detection of endogenous nucleic acids within cellular environments.

Since a molecular computer encoded a solution to the Hamiltonian path problem in 1994 through DNA hybridization,[195] demonstrating that DNA algorithms can be used to perform computation operations, there have been many developments in the field of DNA computation.[196] These DNA-based chemical circuits rely on nucleic acid base pairing interactions for the assembly of structures that are both precise in their activation and highly controllable. Researchers in the field construct these DNA computers to mimic the computation power of electronic systems (i.e., silicon-based) with devices composed solely of biological components. In a silicon-based computing device, the inputs consist of and are processed into electrical signals. For example, binary code utilizes digits of 0 (NO) and 1 (YES) to indicate ON/OFF states. Circuits then utilize these binary digits as inputs in the construction of logic

gates for the execution of Boolean logic functions and generation of electrical output signals. Logic gates are a basic computation device commonly implemented with diodes or transistors that function as electronic switches. However, in the examples presented below the electrical voltages associated with the input and output of a logic gate have been replaced with nucleic acid strands (e.g., DNA or RNA). Though there are many hurdles in producing DNA computation devices that would rival silicon-based computation, such as scaling limitations and the low computation speed of DNA-based devices,[197] there are distinct advantages in developing biologically relevant computation systems. DNA computation devices have the ability to interact with biological and chemical environments, which is an important step towards developing *in vivo* cellular computation, and can potentially be interfaced with genetic regulatory molecules for diagnostic and therapeutic applications.[198, 199] Additionally, DNA computation modules have potential as nanodevices for molecular bioassembly of complex DNA structures or in the development of smart materials.[200] It has been estimated that DNA computation can operate at over 100 teraflops, as well as store a single bit of information in one cubic nanometer, with storage capacities greatly exceeding silicon computation devices.[197] DNA computation also provides increased storage density for information compared to electrical systems.[201] The focus of this chapter is on the construction of novel nucleic acid-based DNA computation devices for both cell-free and cell-based systems (Figure 3.1).[199]
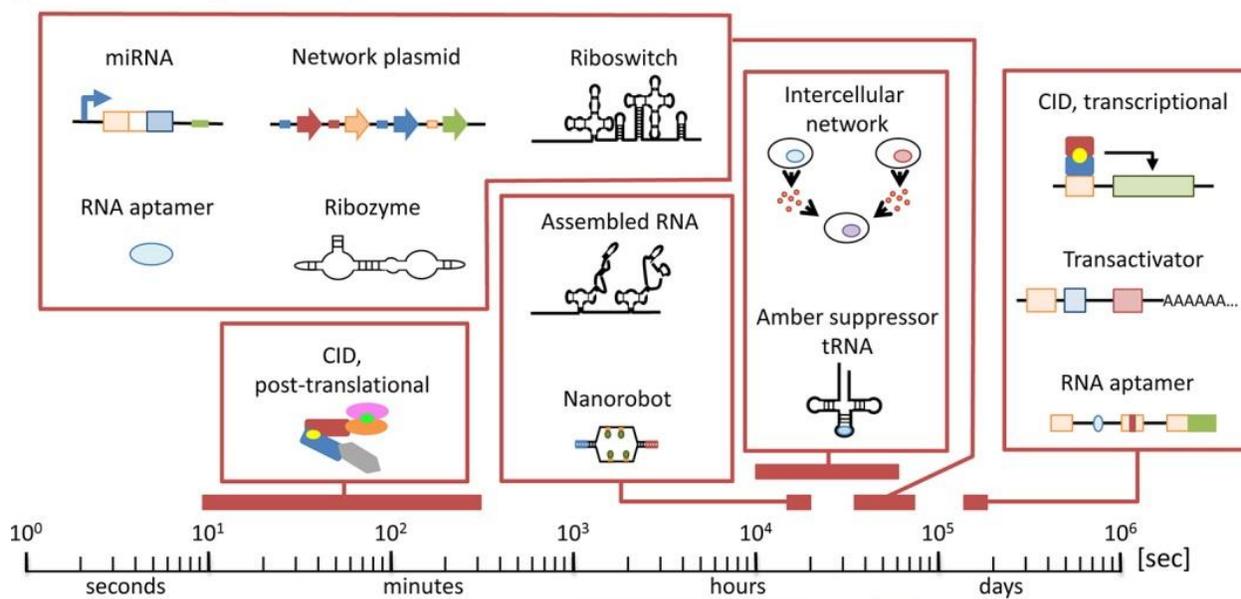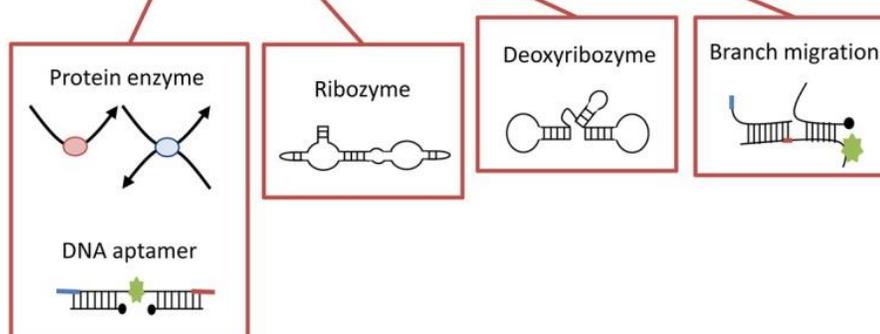
Figure 3.1: Cell-based and cell-free nucleic acid logic devices.
Representative time-scales for the activation of biomolecule-based logic devices are shown. Image adapted with permission from Miyamoto et al, *ACS Synth. Biol.*, **2013**, 2(2):72-82. Copyright 2013 American Chemical Society.

A variety of DNA devices have been engineered that recognize nucleic acid strands or ligand inputs through sequence-specific interactions in the generation of output signals, which can be used for the construction of synthetic circuits or as molecular computers with unique biological functions.[202] The devices that will be discussed in this chapter are limited to logic gates or signal amplification cascades with Boolean function, which, by definition, respond to one or more input to produce a single defined output. The function of these DNA logic gates is

103

based on strand displacement and branch migration, which occur on the order of a few hours as shown in Figure 3.1. The method for activation of the logic gates discussed in this chapter (e.g, OR gates, AND gates, signal amplification networks) utilizes toe-hold mediated strand displacement in the generation of a defined output. A toe-hold is a short single stranded DNA sequence (e.g., 6 bases) extended from a DNA duplex that binds to a complementary sequence on another strand to facilitate DNA hybridization, branch migration, and strand displacement (Figure 3.2).[203] The toe-holds are designed to interact with specific nucleic acid sequences as biological inputs, which displace bound oligonucleotides from DNA duplex structures. In the absence of a toe-hold, strand displacement is kinetically slow (Figure 3.2A). In contrast, if there are two toe-holds present on each termini of the duplex, strand displacement is kinetically fast and considered reversible, a process known as toe-hold exchange (Figure 3.2B). However, in the presence of a single toe-hold, strand displacement is kinetically fast in a single direction, and thus considered irreversible (Figure 3.2C). The first step of toe-hold mediated strand displacement (hybridization) involves complementary binding of the incoming oligo to the toe-hold sequence, which is rapid and reversible ($k_1 = 10^6$ $M^{-1}s^{-1}$).[204] The second step (branch migration) involves removing the bound strand from the toe-hold containing duplex ($k_2 = 1$ s^{-1}), and although the process is slower than the first step, it generally proceeds in a single direction, which makes the reaction irreversible.[204] These calculated rate constants are dependent on the DNA concentration, sequence, and the length of a toe-hold domain, which can alter the rate of binding or displacement.[203] For example, reducing the toe-hold length by a single base (e.g., from 6 to 5) will reduce the rate of the first step by a factor of 10. Due to the important role of a toe-hold in strand displacement reactions, certain strands can be rendered nonreactive through toe-hold sequestering, such as binding a complementary domain to make the toe-hold

inaccessible.[205] Unique toe-holds for each nucleic acid component are utilized to initiate strand displacement reactions with sequence-specific inputs.
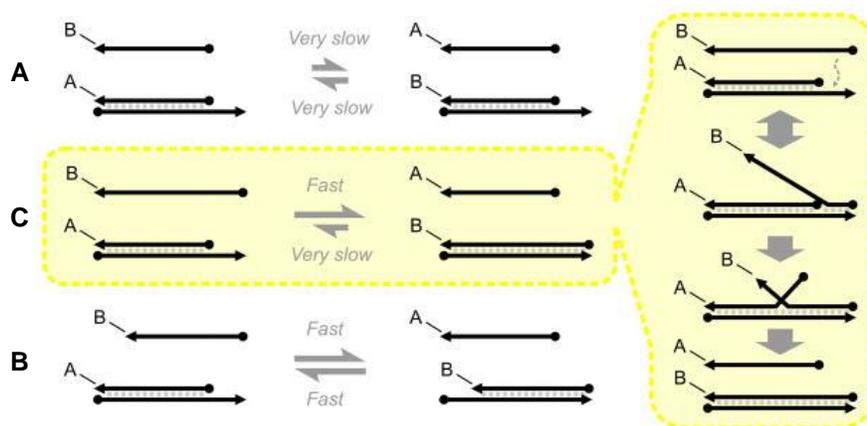
Figure 3.2: Schematic of toe-hold mediated strand displacement.
A) Slow displacement without toe-holds. B) Fast reversible strand displacement with two toe-holds. C) Fast irreversible strand displacement with a single toe-hold (yellow box). A reaction pathway for toe-hold mediated irreversible strand displacement is shown to the right. Round and triangular ends of the lines represent 5′ and 3′ termini, respectively. Dashed line with arrow indicates toe-hold binding. 392-400. Adapted from *Curr. Opin. Biotechnol.*, 21(15), Chen and Ellington, **Shaping up nucleic acid computation**, 392-400, Copyright (2010) with permission from Elsevier

DNA logic gates are a powerful computational device because the outputs are chemically equivalent to the inputs, and the output of one gate can act as the input for a following gate – similar to electronic gates. The advances in DNA logic gate engineering have enabled serial connections of gates, thus generating signaling cascades that can be assembled into complex molecular circuits based on nucleic acid hybridization.[206, 207] For example, a complex circuit containing translator gates was used to combine multiple AND gates with OR gates, in conjunction with signal amplification and restoration (Figure 3.3).[208]
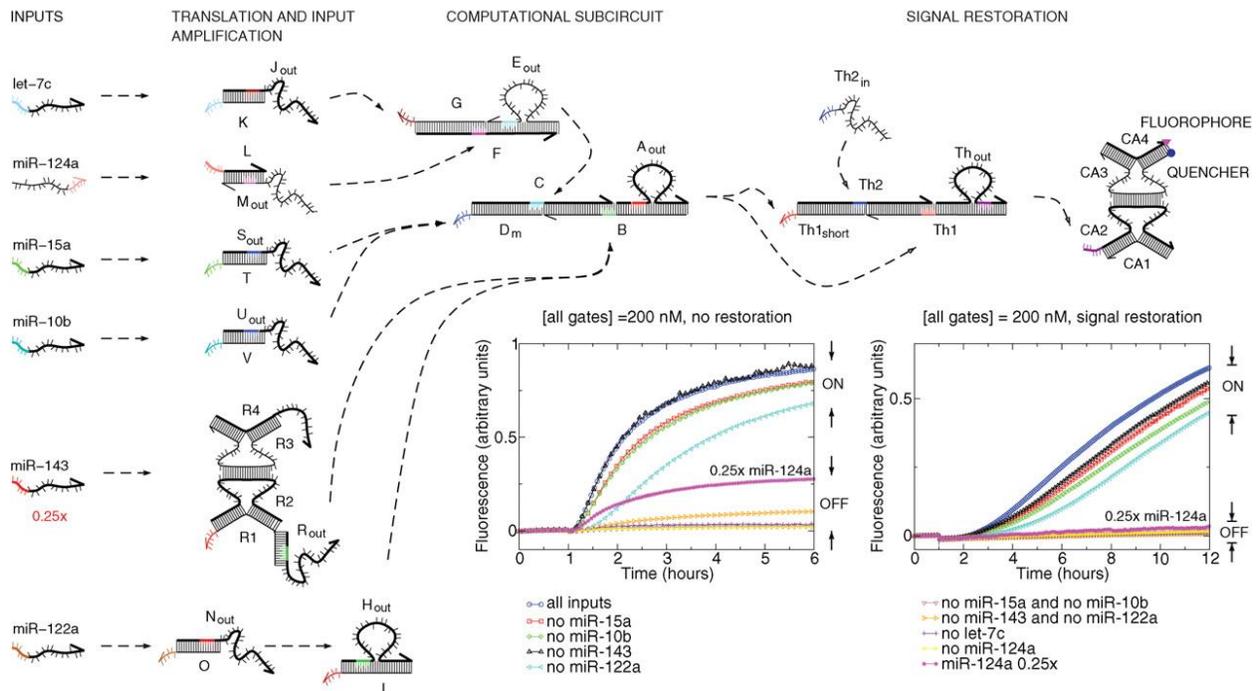
Figure 3.3: Complex circuit with connected logic gates.
Circuit diagram for let-7c AND miR-124a AND (miR-15a OR miR-10b) AND (miR-143 OR miR-122a). Signal propagation through the circuit containing AND gates, OR gates, sequence translators, and a signal restoration module is shown. The five-layer circuit consists of a total of 11 gates and accepts six inputs. Fluorescence traces of circuit operation without and with the signal restoration module (threshold plus amplifier) are shown. Adapted from Seelig et al, *Science*. **2006**, 314(5805):1585-1588. Reprinted with permission from AAAS.

Here, the serial connection of AND gates with OR gates allowed the specific analysis of patterns of six miRNA inputs. Each miRNA input interacted with a translator gate, which produced an output strand that acts as an input for a downstream gate. The final interaction of the completed circuit was the displacement of a fluorophore:quencher containing duplex in a reporter gate to generate an output signal (i.e., fluorescence). This circuit utilized eleven total gate structures, which all interacted autonomously in solution. As the size of a complex circuit increased, the speed of activation decreased (from completion in 5-6 h to requiring >12 h for completion), as shown in the gate kinetic diagrams. However, in both cases the OFF and ON states can be clearly distinguished.

DNA logic gates have been designed as synthetic chemical circuits with functions in memory simulation,[209] deoxyribozyme function,[210] G-quadruplex formation,[211] aptamer ligand binding,[212] activation of protein translation,[213] molecular beacon probes,[214] edge detection,[215] and game simulation.[216] Since enzyme-free DNA logic gates recognize nucleic acid inputs through strand hybridization to activate computation cascades, the oligonucleotide outputs generated can be interfaced with cellular environments. The ability to interface with biological systems is a strong driving force to further develop DNA computing devices that recognize specific biological changes, which is paramount towards the eventual goal of developing modular cellular circuitry and molecular computation devices. Cells have been programmed to recognize DNA hybridization for cellular self-assembly pathways,[205] showing the potential for linking DNA computation with biological systems. Applications of nucleic acid-based logic gates in biological systems include targeted cellular delivery,[217] mRNA detection with amplification,[218] and targeted therapy of cancerous cell markers.[219] For example, programmable DNAzymes libraries have been constructed to replicate a variety of Boolean logic gates with miRNA or mRNA inputs, such as the miR-21 AND miR-125b logic gate that was applied in live cells (Figure 3.4).[220]
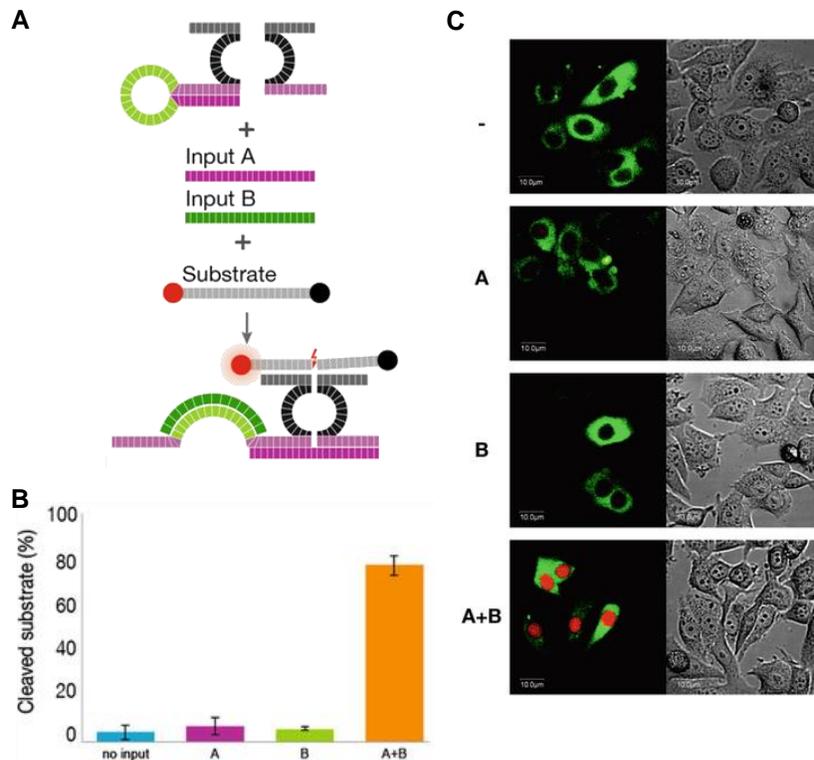
Figure 3.4: Development of miR-21 AND miR-125b DNAzyme logic gate in live cells. A) Illustration of the DNAzyme Dz13 cleaving a fluorescently labeled RNA substrate. Input A = miR-21, Input B = miR-125b. B) Quantitative substrate cleavage results. Error bars represent standard deviation errors from three independent experiments. C) Cells were microinjected with each input combination. Fluorophore-labeled dextran was used to normalize fluorescence and mark injected cells. Representative injected cells were imaged 5 minutes after injection. Fluorescent (left) and phase (right) images are shown. Image adapted with permission from Kahan-Hanum et al, *Sci. Rep.*, **2013**, 3:1535. Copyright 2013 Nature Publishing Group.

In this study, the Dz13 DNAzyme[221] was reprogrammed into a library of Boolean logic gates, where the Dz13 catalytic activity was controlled with specific input conditions. To form an AND gate the DNAzyme's catalytic core was split into two parts, and a blocking arm containing a stem-loop was loop was added to one half of the DNAzyme core (Figure 3.4A). The toe-holds and complementary sequences were designed using miR-21 (Input A) and miR-122 (Input B) sequences. The functional DNAzyme complex will only be formed when both inputs are present, which remove the blocking arm and bring the split catalytic core halves together. The activity of the AND gate was analyzed with a fluorescently labeled RNA substrate that produced an output

signal upon DNAzyme cleavage only when both inputs were present, confirming the AND gate function *in vitro* (Figure 3.4B). Subsequently, the *in vivo* activity of the AND gate was tested. The AND gate, fluorescent substrate, and exogenous inputs were microinjected into cells followed by microscopy imaging (Figure 3.4C). The AND gate properly operated within living cells and the computation output was observed in less than 5 min. The DNAzyme substrate cleavage only occurred when both inputs were injected, and output signal was localized in the cell nucleus. While this report demonstrated activation of DNA computation in a biological environment, the use of exogenous inputs limited the detection of endogenous miRNA profiles in specific cell lines. Moreover, gate activation may have occurred already during injection, before inputs and gate were released into the cell. In another interesting biological application of a DNA nanodevice, pH was sensed in *Caenorhabditis elegans* (Figure 3.5).[222]
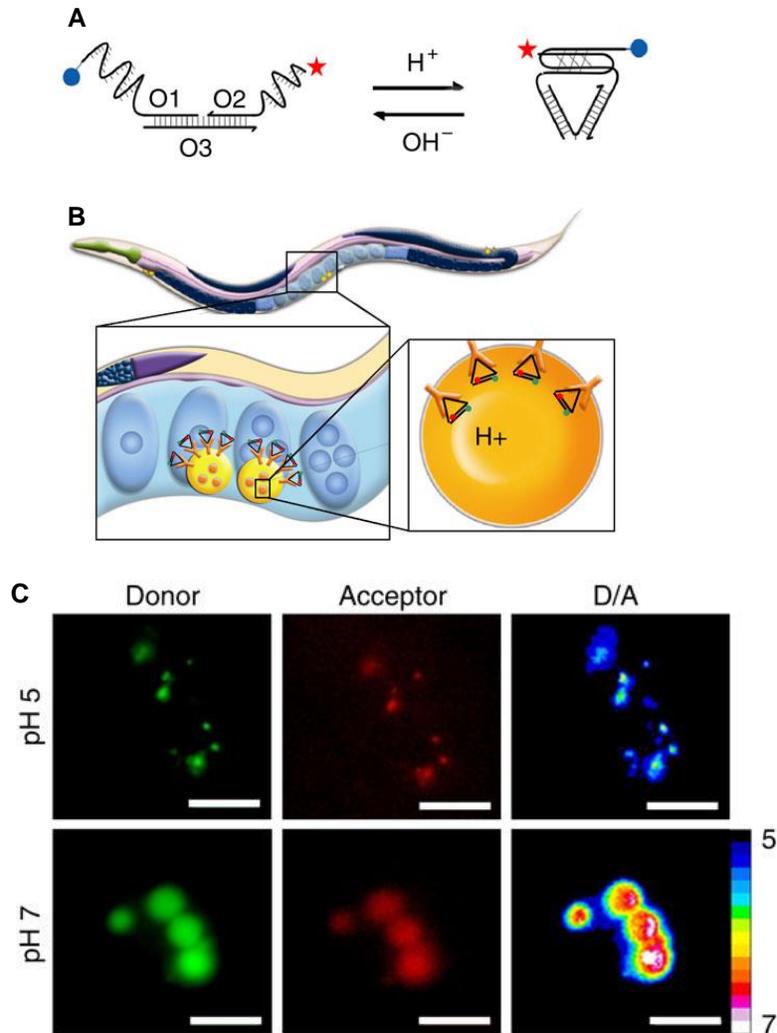
Figure 3.5: DNA-based sensing of pH changes in a living organism.
A) Schematic showing the structure and function of the I-switch nanodevice. B) I-switch uptake in coelomocytes postinjection in *C. elegans*.[223] C) Donor channel, acceptor channel, and respective pseudocolour D/A images of wild-type coelomocytes labelled with the I-switch and clamped at pH 5 or 7. Scale bar indicates 10 μm. Image adapted with permission from MacMillan Publishers Ltd: *Nat. Comm.,* 2:340, copyright 2011.

The DNA nanomachine was triggered within a live animal in response to pH changes associated with endocytosis. A previously described autonomous DNA nanomachine that undergoes a conformational change in response to pH changes, known as the I-switch,[224] was utilized for the detection of acidic compartments within *C. elegans*. The I-switch contains two DNA duplexes connected by a flexible hinge, and cytosine-rich single stranded overhangs at the

termini, which can be protonated under acidic conditions to form an I-motif (Figure 3.5A).[225] In this state, two fluorohores are brought into close proximity, allowing for fluorescence resonance energy transfer (FRET)-based pH sensing. *C. elegans* have proton-rich large scavenger cells known as coelomocytes, which are involved in endocyctosis of macromolecules from the body cavity, and can be used to study endocytosis pathways (Figure 3.5B).[226] The function of the I-switch was determined *in vivo* through worm injections at two pHs and subsequent imaging of the FRET pair (Figure 3.5C). The D/A ratios observed for endosomes at pH 5 are lower than those observed for endosomes at pH 7, demonstrating that the injected I-switch retains pH sensing function within endosomes. The DNA nanomachine allowed for the mapping of spatio-temporal pH changes that occur during endosome maturation. This report demonstrated that rationally designed DNA-based molecular devices can interrogate complex biological phenomena in live organisms. In addition, there are several examples of oligonucleotide computation concepts that utilize protein components and existing networks in live cells, including RNA gates targeting mRNA,[227] transcriptional regulation with CRISPR/Cas9 gene circuits,[228] and RNAi logic gates.[229, 230] Directly relevant to the research discussed in Section 3.5, amplification circuits have been engineered for the detection of RNAs in live cells.[218, 231]

Current limitations of the DNA computation technology that will be addressed in this chapter include spatio-temporal control of logic gate activation (Section 3.1), activation of logic gates in cellular environments for the detection of endogenous nucleic acids (Sections 3.2 and 3.3), and amplification of the input signal (Sections 3.4 and 3.5).

3.1   DNA Computation: A Photochemically Controlled AND Gate

This material was reprinted, in part, with permission from Prokup, A.; **Hemphill, J.**; Deiters, A. *J. Am. Chem. Soc.* **2012**, 134(8), 3810-3815. Oligonucleotide syntheses and assistance in logic gate design were performed by the author of this thesis. All experiments were conducted by Alexander Prokup in the Deiters lab.

DNA-based logic gates have previously been operated through purely chemical means, controlling logic operations through DNA strands or other biomolecules. Although gates can operate through this manner, it limits temporal and spatial control of DNA-based logic operations. Light represents a powerful input with a wide range of advantages over chemical or biological inputs. Photochemical inputs close the gap between DNA computation and silicon-based electrical circuitry, since electromagnetic waves can be directly converted into electrical output signals and vice versa. This connection is important for the further development of an interface between DNA logic gates and electronic devices. The use of chemical inputs introduces variables, such as cellular uptake, processing, and diffusion that reduce the reliability of a logic gate to be controlled in a biological environment. A system in which the logic gate machinery is pre-assembled and activated with light provides enhanced control and specificity. Photochemically caged nucleic acids allow for light-activation of DNA hybridization in a precise manner that other research tools cannot accommodate. However, photocaged nucleic acids have not been used in cellular computation or the development of DNA-based logic gates. Here, photochemical control of logic gate function was demonstrated by employing caging groups on DNA strands responsible for toe-hold displacement. The photochemical triggering of a functional logic gate allows for spatial and temporal activation, which can be used to enhance control over signaling cascades of complex DNA computation circuits.[209]

A light-triggered AND gate was designed based on the concept of toe-hold mediated strand displacement, which forms the basis for many DNA computational elements.[207] As shown in Figure 3.6, the light-triggered AND gate will only deliver an output signal if both photochemical input signals of different wavelengths (I_1 and I_2) are present. The gate complex is composed of three ssDNA oligomers: a fluorophore strand G_F, a quencher strand G_Q and a toe-hold containing strand G_T. The fluorophore and quencher moieties are in close proximity preventing fluorescence. In order to activate the gate, A_4 and B_0 need to induce a toe-hold displacement cascade resulting in the removal of G_Q from G_F. The A_4 strand binds to the toe-hold of G_T separating the gate complex, allowing B_0 to bind to the toe-hold exposed on G_F. This event releases G_Q, permitting emission of the excited fluorophore. It was hypothesized that caging groups installed on select thymidine bases of the A_4 strand will prevent hybridization and thus prevent strand exchange. Therefore, without the proper light inputs for decaging (input I_1 = 365 nm) and excitation (input I_2 = 532 nm) no output signal will be observed. Thus, step 1 involves UV irradiation at 365 nm for decaging of the nucleotides. After caging group removal, complementary regions are exposed, enabling DNA:DNA hybridization. In step 2, A_4 will dislodge G_T via a toe-hold mediated strand displacement mechanism.[207] Following step 2, the gate complex consists of only fluorophore and quencher strands. Step 3 occurs spontaneously because a second toe-hold region is exposed on the gate complex after the G_T strand was expelled by the A_4 strand. During step 4, quencher and fluorophore strands are separated by a second toe-hold mediated exchange with the strand B_0. In step 5, irradiation at 532 nm now leads to excitation of the fluorophore and observation of fluorescence emission as the output signal.
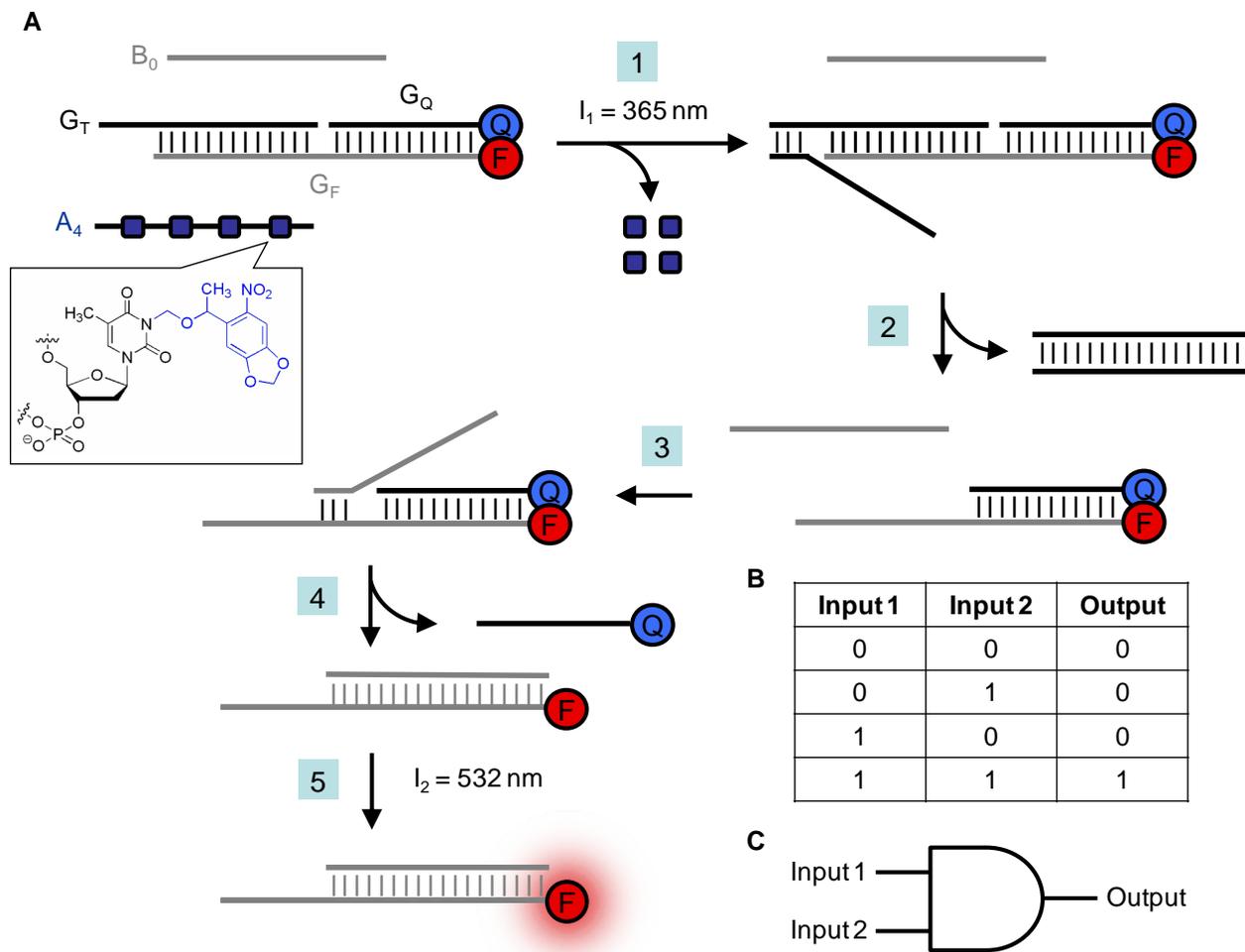
Figure 3.6: A photochemically control DNA logic gate.
A) Light-triggered DNA-based AND gate using irradiation at 365 and 532 nm as input signals I_1 and I_2, respectively, and fluorescence as the output signal. The NPOM (6-nitropiperonyloxymethylene) caging group installed on thymidine nucleotides is represented by a blue square. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). B) AND gate truth table. C) AND gate circuit diagram. Image adapted with permission from Prokup et al, *J. Am. Chem Soc.* **2012**, 134(8):3810-5.

In order to determine the effect of caging groups for the photochemical control of an AND gate, caging groups were initially added to the A strand (for a detailed discussion of nucleobase caging, see Section 1.1). A set of four oligonucleotides bearing 1–4 caging groups was synthesized (A_1–A_4, Table 3.1) and individually tested for function to study the design requirements for suppression of strand displacement.

Table 3.1: Inputs for the photochemically activated AND gate. NPOM-caged thymidine residues are underlined, bolded, and highlighted red ("**T**"). Caged DNA oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
A_0	TATGGTTGTTTATGTGTTCCCTGATCTTTAGCCTTA
A_1	TA**T**GGTTGTTTATGTGTTCCCTGATCTTTAGCCTTA
A_2	TA**T**GG**T**TGTTTATGTGTTCCCTGATCTTTAGCCTTA
A_3	TA**T**GG**T**TGTTTATG**T**GTTCCCTGATCTTTAGCCTTA
A_4	TATGG**T**TGTTTATG**T**GTTCCC**T**GATCTT**T**AGCCTTA

As the number of caging groups was increased, the fluorescence output of the gate linearly decreased in the absence of UV irradiation with $I_1 = 365$ nm (Figure 3.7A). Optimal suppression of the output signal was observed with A_4, which contained four NPOM caging groups evenly distributed throughout the DNA strand and displayed no activity. An optimization of the UV irradiation time for decaging was then conducted, and a time course was performed with the A_4 strand. Maximum fluorescence was observed after 15 min of UV irradiation at $I_1 = 365$ nm followed by a brief excitation at $I_2 = 532$ nm (Figure 3.7B). Longer I_1 irradiation times lead to a decrease in fluorescence, possibly due to photobleaching of the fluorophore.[232] Activation of logic gates using noninvasive UV irradiation as an input signal shows that a DNA-based light switch can be generated, which holds promise for developing new applications of externally regulated DNA computation devices.
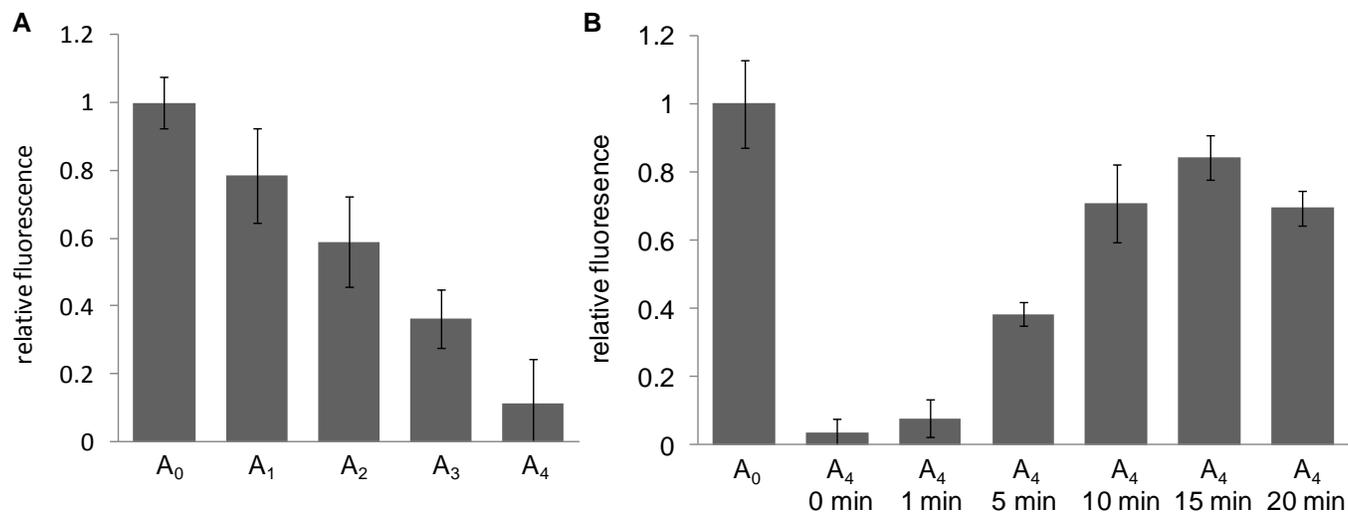
Figure 3.7: Optimization of caging group number and UV irradiation time for the A strand. A) The logic gate was not irradiated with 365 nm light in order to keep all caging groups in place, but only with 532 nm light (I_2). B) The logic gate was irradiated with a time course of 365 nm light (I_1) and activated was detected with 532 nm light (I_2). Averages of three independent experiments are shown, and error bars represent standard deviations. Image adapted with permission from Prokup et al, *J. Am. Chem Soc.* **2012**, 134(8):3810-5. These experiments were conducted by Alexander Prokup.

To investigate whether the light-triggered AND gate could be controlled with temporal and spatial resolution using UV light as an input, several experiments were conducted with the caged A_4 strand. First, three separate sets of logic gates were irradiated at different time points, and a fluorescent signal was only observed after UV irradiation (Figure 3.8A). Thus, temporal control over the light-triggered AND gate was achieved. Second, a step response of the gate was elicited through subsequent UV irradiations in two intervals (Figure 3.8B). The output signal of the caged AND gate is dependent upon the irradiation interval and increases with additional UV exposure. The tunable nature of the step response displays a unique feature to control output intensity of a DNA-based AND gate using subsequent input stimuli I_1. Achieving a tunable step-response allows light-triggered DNA logic gates to be used as molecular controllers that can be adapted to enhance circuit cascades. These factors demonstrate the improvements upon existing DNA logic gates through temporal activation with light input signals. Third, to demonstrate

spatial control of DNA computation via locally restricted light irradiation, the AND gate complex and the caged strand A_4 were embedded into a low-melt agarose gel. The gel was either kept in the dark or irradiated in a patterned fashion with I_1 = 365 nm UV light, followed by imaging of the gel via excitation at I_2 = 532 nm. A distinct pattern was obtained and no fluorescence was observed in the absence of UV irradiation, demonstrating the ability to apply the developed light-triggered AND gate in spatially controlled DNA computation (Figure 3.8C). This demonstrates that logic gate operations can be performed in semi-solid structures and are not limited to solution-based applications. Since electronic systems depend on solid structures and spatially separated devices, identification and recognition of spatially separated signals potentially allows for organization of circuits to create an important link between non-electronic and electronic computational systems.
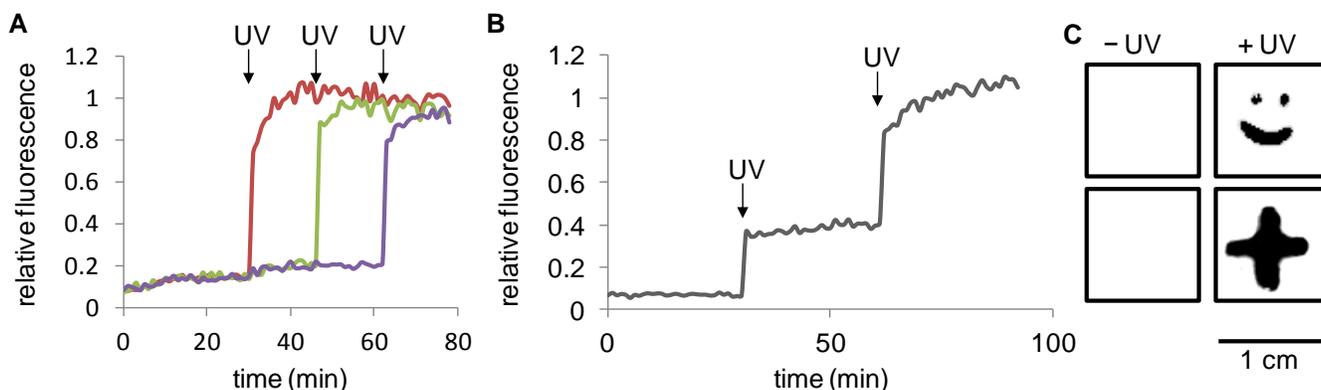
Figure 3.8: UV irradiation at different time points to demonstrate temporal control.
A) Baseline fluorescence was measured for 30 min, and three individual gates containing caged A_4 at were irradiated with I_1 = 365 nm light at 30 (red), 45 (green), and 60 min (purple). B) A single logic gate containing caged A_4 was irradiated for two intervals resulting in a steplike response. Graphs represent an average of three independent experiments. C) Spatially restricted activation of the DNA logic gate through patterned UV irradiation using masks followed by gel imaging with I_2 = 532 nm. Image adapted with permission from Prokup et al, *J. Am. Chem Soc.* **2012**, 134(8):3810-5. These experiments were conducted by Alexander Prokup.

In conclusion, a photochemically controlled AND gate was developed through the incorporation of caged thymidine nucleotides in a DNA-based logic gate. Strands of DNA were synthesized using specialized phosphoramidites, which enabled the use of specific wavelengths of light as inputs for a DNA-based AND gate. Temporal control over DNA computation was achieved through introducing four caging groups and activating separate gate complexes at different time points, displaying fundamental properties of a light-switch for molecular circuits. When a single gate complex was irradiated at two intervals, a step response in the output signal was observed, suggesting that the phototriggered AND gate can act as a tunable DNA-based circuit. Integration of a light-activated AND gate for purposes of a step response could allow the gate to function as a manual feedback controller. Within a cascade of gates, the light-triggered AND gate can operate as a switch or controller and will allow for more complex and better controlled circuit designs. Moreover, photochemical activation enabled DNA-based logic operations in a spatially localized fashion. This was demonstrated by light-triggered pattern formation in a semisolid substrate, where DNA computation events were only observed in areas that received irradiation with both input wavelengths. Design rules were established that enabled light-activation of the gate and will be applicable to further developments, e.g., the generation of other light-triggered logic gates. The use of light to control a DNA-based logic gate creates a new paradigm of inputs that will be beneficial when used in a biological context. Light allows for spatial and temporal control with high specificity, while overcoming the downfalls of chemical-based inputs such as diffusion and delivery kinetics. Photochemical inputs also shorten the gap between DNA computation and silicon-based electrical circuitry, since light waves can be directly converted into electrical output signals and vice versa. This connection is supremely important for further developing the interface of DNA logic gates and electronic devices and,

thus, the interface of biological systems with electrical circuits. Thus, the photochemical control demonstrated here lays the foundation for the programming of complex, DNA-based computation devices with unprecedented spatial and temporal resolution.

3.2 DNA Computation in Mammalian Cells: AND Gates

The next step taken for the design of new paradigms in DNA computation was the application of DNA-based logic gates in a biological system. The use of DNA logic gate operations in live cells requires biomolecular assembly of DNA structures, which has been well documented.[233] Here, the first *in vivo* photochemical activation of a light-triggered AND gate in mammalian cells, and subsequent application for the detection of endogenous nucleic acids with DNA logic gates targeting microRNAs are described. This material was reprinted, in part, with permission from **Hemphill, J.**; Deiters, A. *J. Am. Chem. Soc.* **2013**, 135(28), 10512-10518. All experiments were conducted by the author of this thesis.

Oligonucleotide AND gates were engineered to respond to specific microRNA (miRNA) inputs in live mammalian cells. Both single and dual-sensing miRNA-based computation devices were synthesized for the cell-specific identification of endogenous miR-21 and miR-122. A logic gate response was observed with miRNA expression regulators, exhibiting molecular recognition of miRNA profile changes. Nucleic acid logic gates that are functional in a cellular environment and recognize endogenous inputs significantly expand the potential of DNA computation to monitor, image, and respond to cell-specific markers.

At the time of this study, synthetic DNA-based logic gates had not yet been reported in live mammalian cells and therefore held great promise for the computation of endogenous

biological inputs independent of protein components and gene expression systems. MicroRNAs are small noncoding ssRNAs that down-regulate gene expression in a sequence-specific fashion by binding the 3′ untranslated regions of target mRNAs,[139] and it has been estimated up to 30% of all genes are regulated by miRNAs.[140] The expression and misregulation of certain miRNAs has been linked to a wide range of human diseases,[142] highlighting the importance in understanding miRNA regulation and function.[141, 143]

It was demonstrated that logic gates can be used to detect miRNAs in vivo through DNA computation and will enable cell-specific gate activation based on endogenous miRNA expression patterns. Although DNA-hydridization probes have been used to detect cellular miRNAs,[234] for example, through miRNA-specific molecular beacons,[235] DNA computation in live cells allows for Boolean logic gate operations with miRNA inputs, and the generated oligonucleotide outputs enable applications beyond miRNA pattern detection. While a previous example of miRNA detection by logic gate operations in live cells requires plasmid constructs for the expression of protein gate components,[229] the use of DNA logic gates to identify endogenous miRNA patterns as described here is independent of the cellular transcription and translation machinery and does not require the transfer of genetic information, allowing the circuits to be orders of magnitude smaller in size based on DNA sequence (e.g., 268 bp versus >7000 bp for a dual-miRNA input gate).

MicroRNAs miR-21 and miR-122 were selected as inputs for miRNA-specific intracellular logic gate activation. Overexpression of miR-21 is observed in many cancer types,[146, 236] while miR-122 is involved in hepatocellular carcinoma[237] and hepatitis C virus replication.[148] The DNA-based AND gate is derived from a design by Seelig et al.[208] and computes two input strands that initiate toe-hold mediated strand exchange, displacing a

120

quencher and fluorophore duplex, leading to a fluorescent output. The use of mature miRNAs as endogenous AND gate inputs is both sequence- and cell-specific. Additionally, the fluorescent output can be observed without the need to perform cell lysis or RNA purification, which can introduce variability between experiments.[238] Here, the developed methodology was applied to the detection of cellular miRNA expression patterns, utilizing DNA computation as a method for the development of synthetic circuits that are functional in live cells.

In order to visualize DNA computation in a cellular system, a photochemically activated logic gate was utilized to trigger toe-hold mediated strand exchange and gate output at a defined time point. First, the light-triggered DNA logic gate (oligonucleotide sequences can be found in Table 3.2) was tested for activity in mammalian cells through transfection of the AND gate components into HEK293T cells. The noncaged A and B strands were used to determine if gate components would remain stable within live cells, and if activated gate fluorescence could be observed in vivo (Figure 3.9). Logic gate activation was observed 4 h after transfection and was dependent on the presence of both input strands, confirming cellular stability of the AND gate duplex and fluorescent imaging of DNA computation through standard microscopy techniques. Imaging cytometry was unfortunately unsuccessful for the quantification of logic gate activation in cells, since the punctuate fluorescence patterns did not the satisfy requirements of identifiers used in common cell counting software, such as size cutoffs and shape recognition.
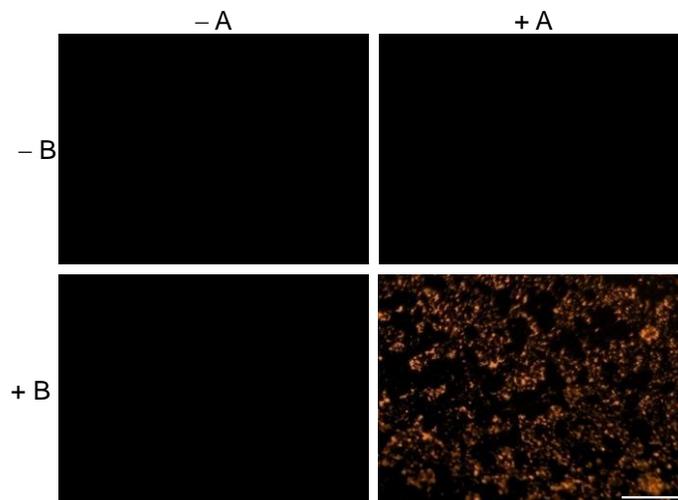
Figure 3.9: Application of logic gate for DNA computation in HEK293T cells.
Transfections were performed with the AND gate and A/B strands then imaged after 4 h. Scale bar
indicates 200 μm.

To ensure that the logic gate could be triggered inside a cellular environment and to
demonstrate intracellular toe-hold mediate strand exchange without activation in preceding
mixing steps, photochemical control of DNA computation was used (for a detailed discussion of
nucleobase caging, see Section 1.1). The AND gate was cotransfected with the caged A_{4T} input,
and irradiation (5 min, 365 nm) was performed after gate transfection was completed, followed
by cellular imaging after 1 h (Figure 3.10). Successful cellular activation of the AND gate
through UV decaging of the input strand with minimal background was observed. This temporal
control over oligonucleotide hybridization confirms that DNA computation based on toe-hold
mediated strand exchange can be conducted in live cells, since only logic gates that were
successfully transfected into the cells and also exposed to UV irradiation were activated.
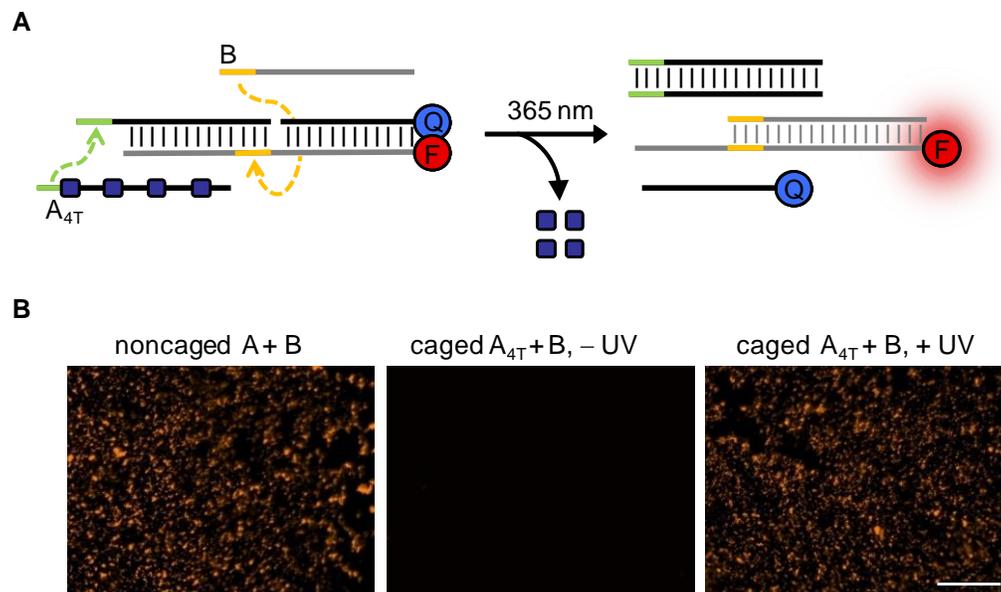
Figure 3.10: Photochemically controlled DNA-based AND gate activation in HEK293T cells. A) Simplified schematic of the light-activated AND gate with a caged A_{4T} strand. NPOM-caged thymidine nucleotides are indicated as dark blue boxes. Toe-hold regions are shown for the A (green) and B (orange) input activation cascades, along with corresponding arrows representing hybridization steps. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). B) Cells were transfected with the AND gate and A/B inputs then imaged for TAMRA fluorescent output after 4 h. Cells were also transfected with the AND gate and caged A_{4T}/B strands for 4 h then UV irradiated and imaged 1 h after exposure. Scale bar indicates 200 μm.

In order to develop a DNA logic gate that can respond to an endogenous cellular activator, a miR-21-based AND gate was engineered that replaces the A input with mature miR-21. The gate sequences were amended to add miR-21 relevant toe-hold regions on G_T and recognition sites within the gate complex, which required an 8 base truncation from the AND gate duplex to accommodate the smaller miR-21 input (Table 3.3). The B strand was altered to have a toe-hold recognition sequence that is complementary to the 3′ sequence of miR-21. Additionally, the fluorophore and quencher moieties were switched to the opposite strands to increase DNA synthesis yields. To verify that toe-hold mediated strand exchange was not inhibited by the alterations in gate sequence and introduction of ssRNA as an input, the miR-21 AND gate duplex was purified and incubated with synthetic miR-21 and B_{21} strands (Figure 3.11A). The miR-21 AND gate is only activated when both the miR-21 and B_{21} inputs are

123

present, providing a miR-21 responsive DNA computation device. The miR-21-based AND gate was then tested for function in HEK293T cells, which display low endogenous expression levels of miR-21.[239] As expected, the logic gate was confirmed to produce an output signal dependent upon transfection of both synthetic miR-21 RNA and B_{21} stands (data not shown). Importantly, no background fluorescence was observed with the AND gate alone or with only a single input strand, confirming the gate duplex stability in vivo. Next, two human cell lines which express high levels of miR-21 (Hela and Huh7, >10-fold increase in relative miR-21 concentration compared to the control HEK293T cells)[239] were transfected with the AND gate components in order to test detection of endogenous miR-21 through DNA computation. Experiments with these cell lines were performed in the absence of synthetic miR-21, and the AND gate was cotransfected with only the B_{21} input strand, followed by fluorescence imaging after 4 h (Figure 3.11B). The miR-21 AND gate duplex was not activated in the absence of B_{21}, and activation in the presence of B_{21} indicates that toe-hold mediated strand exchange was initiated with endogenous miR-21 as the input strand. The presence of miR-21 with the engineered AND gate was observed in both cell lines, validating the application of AND gates for the detection of miRNA in mammalian cells. The logic gates show high transfection and activation efficiency in a cellular monolayer, as >90% of the cells show a fluorescent output from endogenous gate/miRNA interactions.
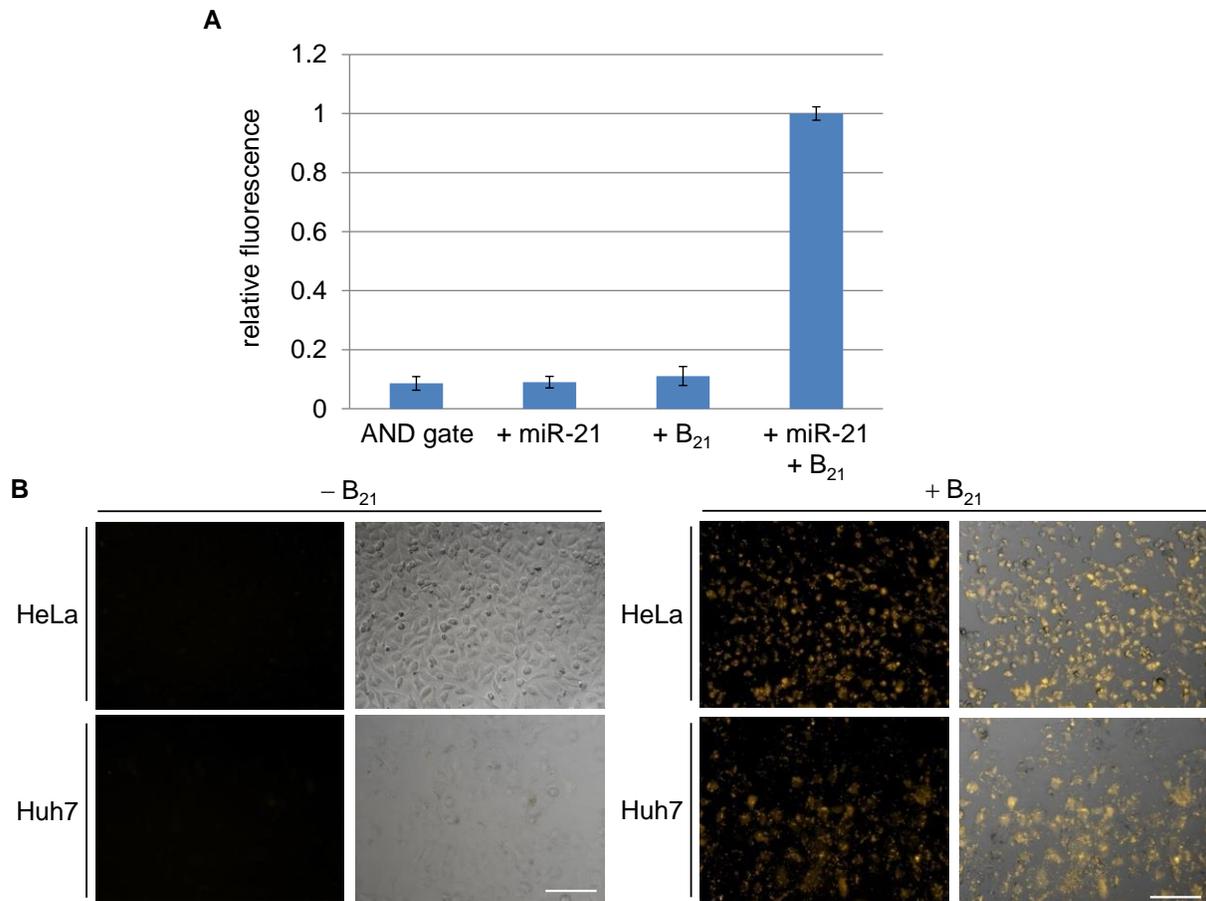
Figure 3.11: Engineered miR-21-based AND gate for detection of endogenous miR-21. A) The logic gate was tested for activation with miR-21 and B_{21} strands. TAMRA fluorescence was observed at 4 h and normalized to the activated AND gate. An average of three independent experiments is shown and error bars represent standard deviations. The miR-21 AND gate requires both inputs for generation of fluorescent output. B) HeLa (top) and Huh7 (bottom) cells were transfected with the miR-21 gate and B_{21} strand then imaged after 4 h. Scale bars indicate 200 μm.

A transfection time course was performed in order to determine the minimum amount of time required for AND gate activation in vivo, and early detection of miR-21 can be observed in >50% of the cells after only 1.5 h, suggesting that a short cellular incubation is sufficient for computation of miRNA expression patterns (Figure 3.12).
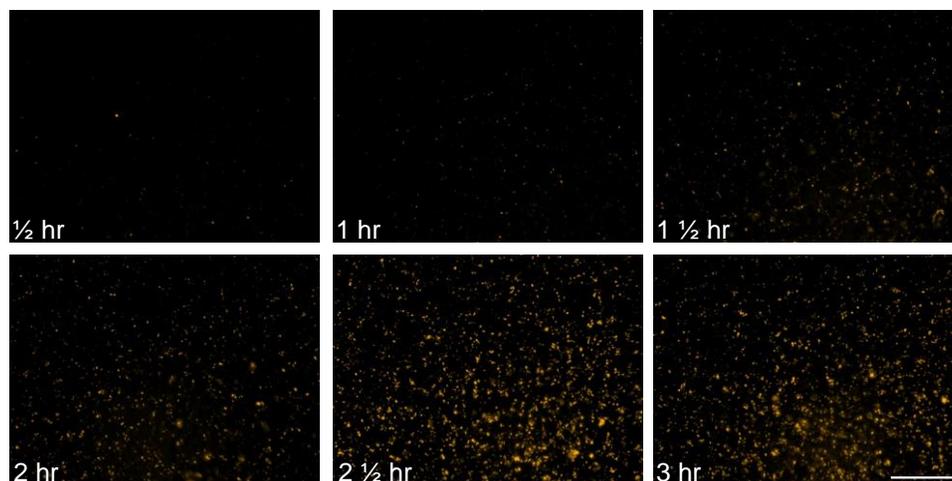
Figure 3.12: miR-21 AND gate time course in HeLa cells.
Cells were transfected with the miR-21 gate and B_{21} then imaged at specific time points. Scale bar indicates 200 μm.

Fluorescence was also imaged after an extended cellular incubation time of 24 h (Figure 3.13). In this experiment, endogenous miRNA-driven logic gate activation was still observed, and no fluorescence was observed in the absence of B_{21}. This data suggests that the nonactivated gate duplex was stable in HeLa cells for up to 24 h, and duplex degradation does not occur within an overnight time frame.
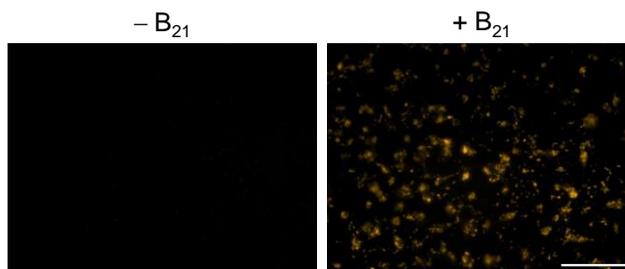
Figure 3.13: Long-term exposure of the miR-21 AND gate in HeLa cells.
Cells were transfected with the miR-21 gate and B_{21} then imaged after 24 h incubation. Scale bar indicates 200 μm.

126

Additional imaging experiments were performed, including nuclear staining, to identify the location of the DNA computation event (Figure 3.14). Apparent localization of activated miR-21 logic gates could be observed around the nuclear periphery – potentially in endosomes, suggesting that DNA logic gates can potentially be applied to the spatial identification of miRNA-rich and/or logic gate-rich areas.
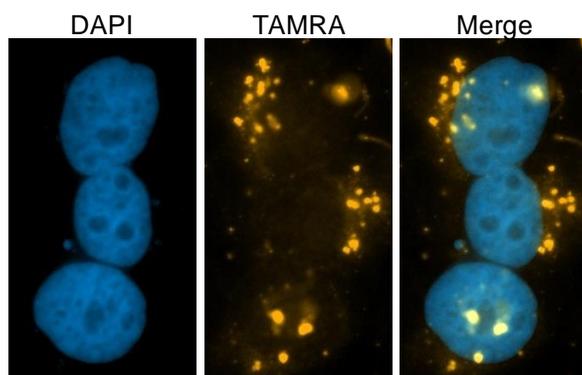
Figure 3.14: miR-21 AND gate cellular localization in HeLa cells.
Cells were transfected with the miR-21 gate and B_{21} then fixed and nuclear counter stained with DAPI and imaged to observe gate activation with sub-cellular resolution.

The miR-21 logic gate was then altered to be photochemically controlled through UV activation of a caged input strand. Light-activation allows for controlled cellular interaction of the logic gate output and enables temporally distinct detection of endogenous nucleic acids. A caged B_{21-4T} strand gate was synthesized with four light cleavable NPOM-dT bases and tested for the photochemical activation of the miR-21 AND gate (Figure 3.15B). An increase in gate fluorescent output was observed after a 5 min UV irradiation when the AND gate was analyzed in solution. The caged miR-21 AND gate was then transfected into HeLa cells for photochemically activated miRNA detection (Figure 3.15C). As expected, in conjunction with the caged B_{21-4T} strand, the miR-21 AND gate only detects endogenous miR-21 after UV

exposure. The combination of photocaging and miRNA detection allows for the use of light as a second input when analyzing endogenous miRNA expression in biological environments.
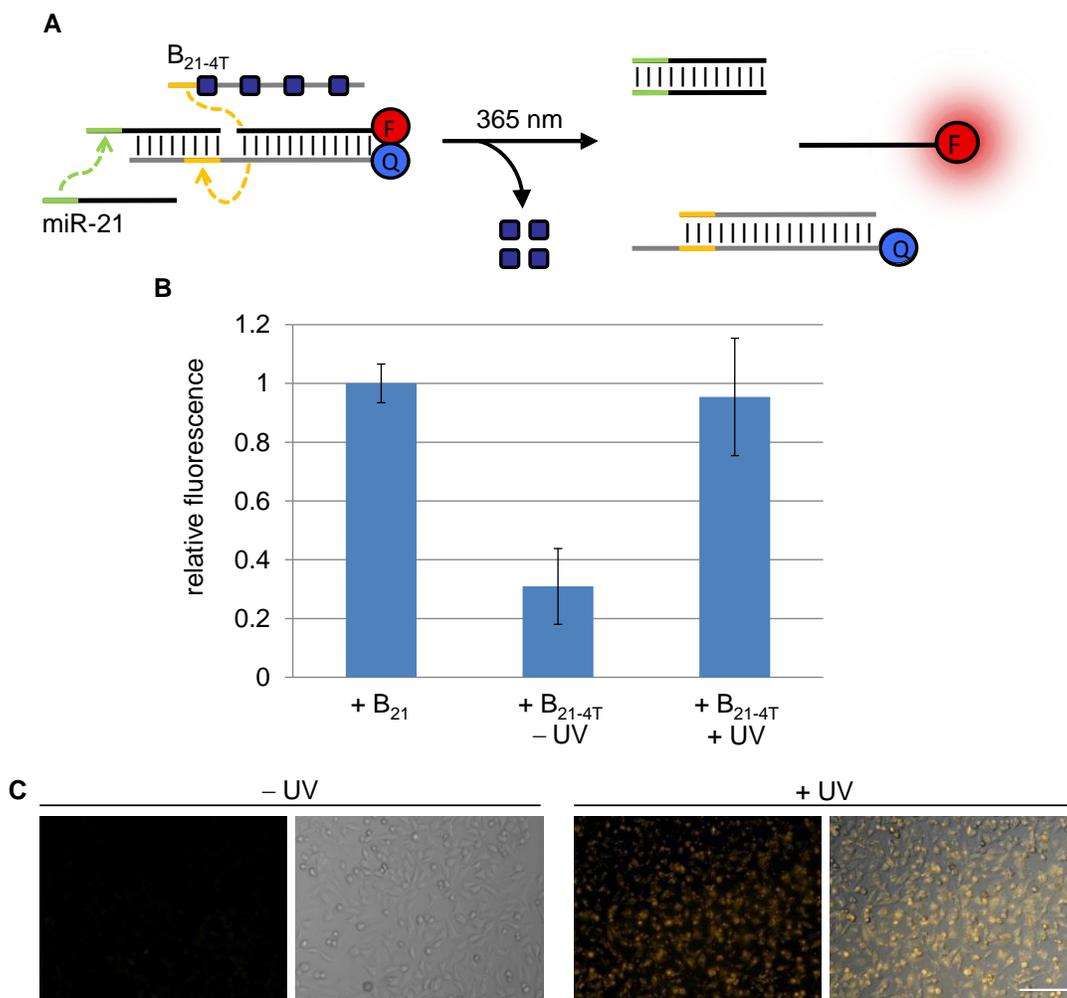
Figure 3.15: A photochemically triggered miR-21 AND gate.
A) Simplified schematic of the AND gate with a caged B_{21-4T} strand. NPOM-caged thymidine nucleotides are indicated as dark blue boxes. Toe-hold regions are shown for the miR-21 (green) and B_{21-4T} (orange) input activation cascades, along with corresponding arrows representing hybridization steps. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). B) The AND gate was tested for photochemical (365 nm, 5 min) activation of caged B_{21-4T}. TAMRA fluorescence was observed at 4 h and normalized to the noncaged activated AND gate. An average of three independent experiments is shown and error bars represent standard deviations. C) HeLa cells were cotransfected with the miR-21 gate and caged B_{21-4T} for 4 h and then UV irradiated and imaged 1 h after exposure. Scale bars indicate 200 μm.

Environmental changes can greatly affect miRNA expression patterns and quickly monitoring these effects with DNA computation devices has potential in diagnostic application and therapeutic responses. Here, the AND gate was applied to the detection of changes in miR-21 levels in HeLa cells through both small molecule treatment with a miR-21 inhibiting azobenzene[149] and antagomir treatment with a reverse complement of the mature miR-21 sequence,[21] perturbing different stages of the miRNA pathway.[145] Cells were treated with either the miR-21 antagomir (100 pmol) or the small molecule (10 μM) for 48 h, followed by transfection of the miR-21 AND gate and subsequent fluorescence imaging (Figure 3.16). The logic gate output is deactivated through treatment with either the small molecule or the miRNA antagomir, demonstrating successful DNA computation of miRNA responses to environmental changes.
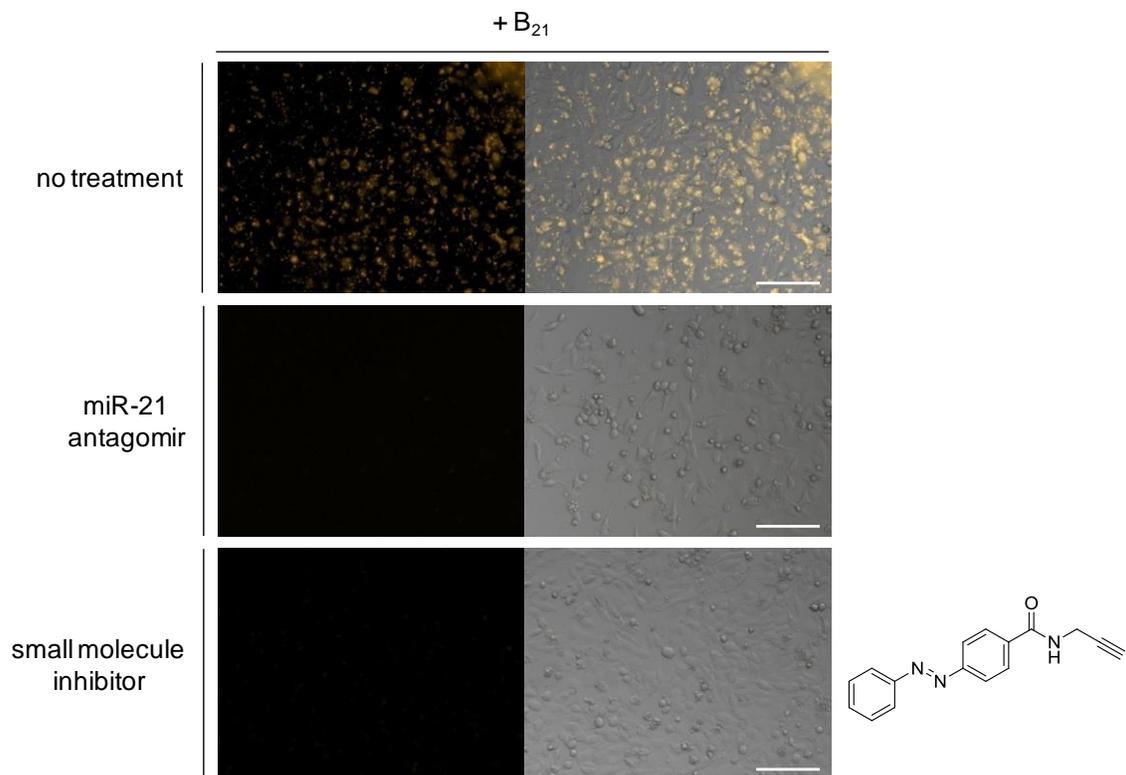
Figure 3.16: miR-21 AND gate detection of functionally depleted miR-21 in HeLa cells. Cells were treated for 48 h with a miR-21 antagomir (100 pmol) or a miR-21 small molecule inhibitor (10 μM) followed by cotransfection with the miR-21 gate and B_{21} as previously described. Scale bars indicate 200 μm. The azobenzene derivative shown is a known miR-21 inhibitor.

A dose–response experiment was performed to observe the effects of decreased small molecule concentrations on detection of miR-21 inhibition (Figure 3.17). Inhibition of miR-21 expression from treatment of HeLa cells with the small molecule can be detected at 250 nM through deactivation of a miR-21 AND gate. Thus, the detection of miRNA with logic gates has been shown to be sensitive to miRNA inhibitors at submicromolar concentrations, comparable to genetically encoded luciferase reporters.[240]
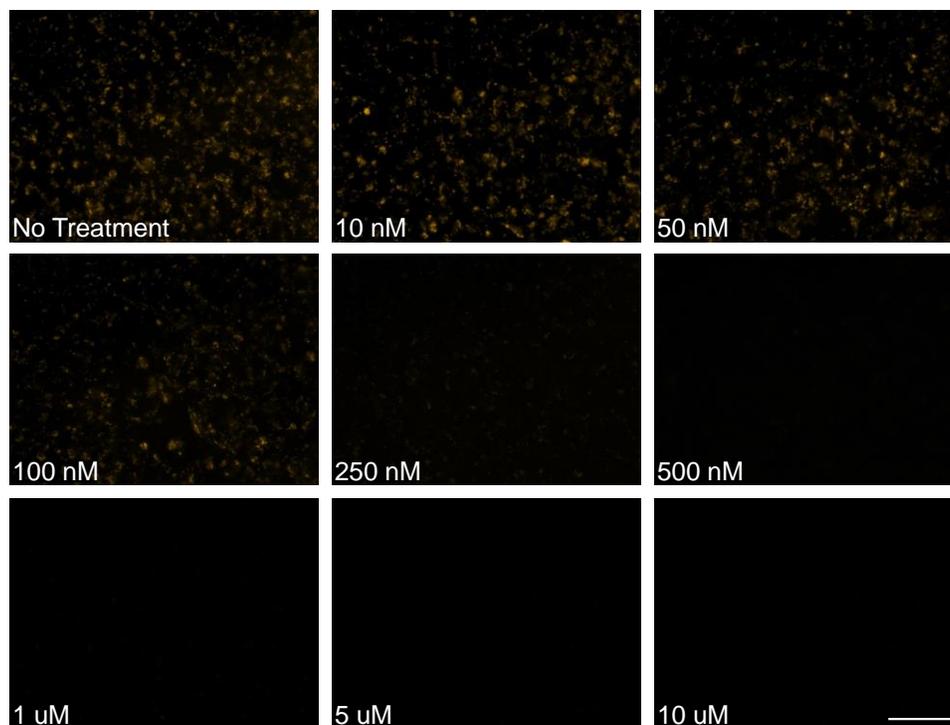
Figure 3.17: Dose response for miR-21 AND gate with small molecule treatment. HeLa cells were treated for 48 h with a small molecule inhibitor (10 nM to 10 μM) then transfected with the miR-21 gate and the B_{21} strand. Scale bar indicates 200 μm.

In order to demonstrate the general applicability of sequence-specific DNA logic gates for the detection of various miRNAs in different cellular environments, a miR-122-based AND gate was engineered. The miR-122 AND gate was designed through reconfiguration of the two input logic gate system (Table 3.4) and confirmed to be activated in the presence of both miR-122 RNA and B_{122} inputs (Figure 3.18A). No background fluorescence was observed in HEK293T or HeLa cells, which do not express miR-122,[239] and logic gate function was specifically activated through the addition of synthetic miR-122 (data not shown). To validate endogenous cellular activation, the miR-122 AND gate was applied to Huh7 cells, which express high levels of miR-122 (>10-fold increase in relative miR-122 concentrations compared to the control HEK293T and HeLa cells).[239] Efficient logic gate activation was observed in the presence of cellular miR-122 and the addition of the second B_{122}input strand (Figure 3.18B).

131

These experiments confirm that the AND gate design can be readily altered to recognize different miRNAs, in different cell types, and can be applied to the imaging of cell-specific miRNA markers.
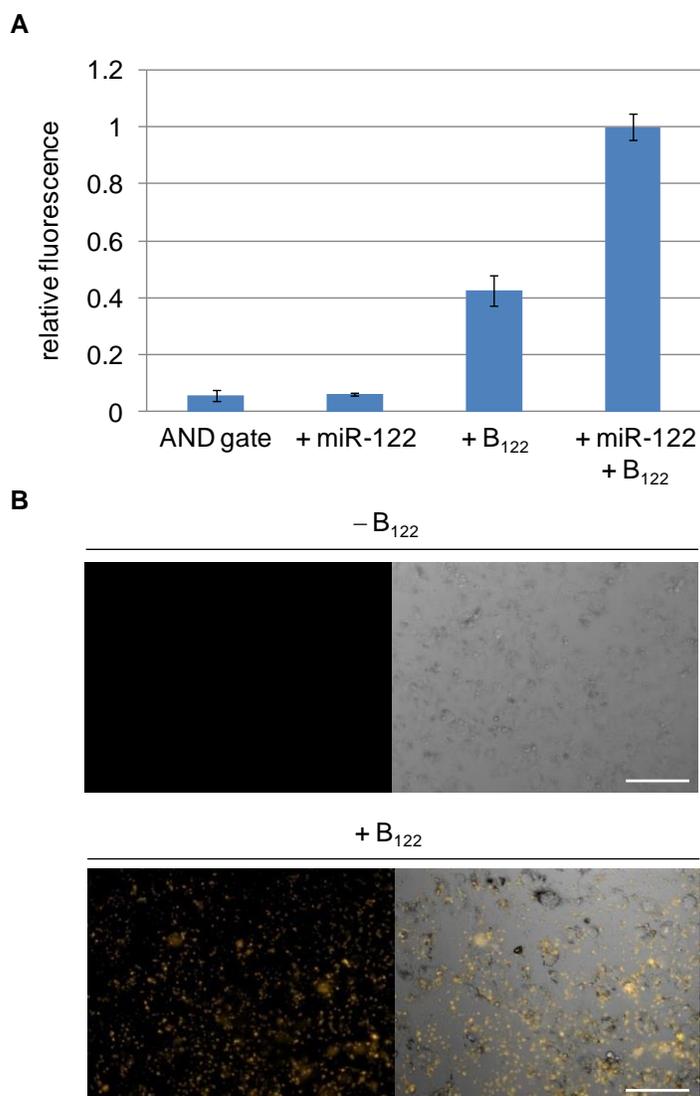
Figure 3.18: Engineered miR-122-based AND gate.
A) The logic gate was tested for activation with miR-122 and B_{122} strands. TAMRA fluorescence was observed at 4 h and normalized to the activated AND gate. An average of three independent experiments is shown and error bars represent standard deviations. B) Huh7 cells were cotransfected with the miR-122 gate and B_{122} strand and imaged after 4 h. Scale bars indicate 200 μm.

A distinct application of DNA logic gates that are functional in cellular computation events is the ability to trigger an output signal dependent upon the recognition of multiple endogenous inputs. Expression patterns of miRNAs associated with cancer and other disease states often involve the simultaneous up-and-down regulation of several miRNAs,[145] highlighting the importance of simultaneous detection for the computation of diagnostic and therapeutic outputs based on miRNA expression patterns. In order to develop a dual-miRNA logic gate, the design of translator gates[208] was adapted to interact with miR-21 and miR-122 inputs (Table 3.5 and Figure 3.19), converting them into DNA output strands that subsequently trigger AND gate activation and fluorescent output. This modular design will allow for the general application of the cellular DNA logic gate system for dual-miRNA activation through simple tailoring of the connected translator subcircuits to any two miRNA inputs.
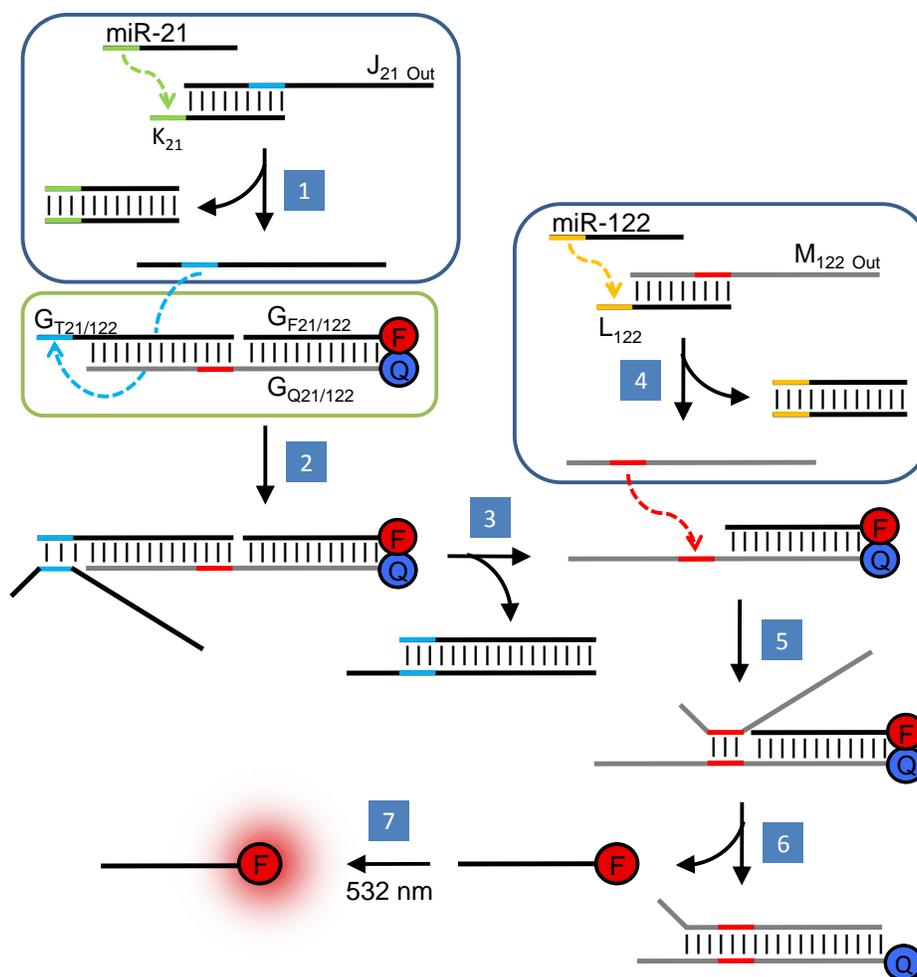
Figure 3.19: The miR-21/122 translator-coupled AND gate activation cascade. Translator gates (blue boxes) interact with two miRNA inputs that displace the $J_{21\ Out}$ and $M_{122\ Out}$ strands. These translator gate outputs activate the miR-21/122 AND gate (green box) through toe-hold mediated strand exchange in sequential steps producing a fluorescent output (ex: 532 nm). Toe- hold regions are shown for the miR-21 (green/blue) and miR-122 (orange/red) activation cascades, along with corresponding arrows representing hybridization steps. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA).

The miR-21/122 AND gate system was first tested with synthetic RNA molecules and was only activated in the presence of both miRNAs (Figure 3.20B). The translator-coupled AND gate device was successfully applied in Huh7 cells, which express high functional levels of both miRNAs,[239, 241] thus displaying a rapid Boolean operation with endogenous miR-21 and miR-122 as inputs (Figure 3.20C). There was no background with a single translator gate, and both translator gate interactions were required to trigger the AND gate, further verifying complete

specificity for the two cellular miRNA inputs. These experiments validated that the translator-coupled AND gate circuit maintains a low fluorescence background, remains stable intracellularly, and is activated only in the presence of both miRNAs.
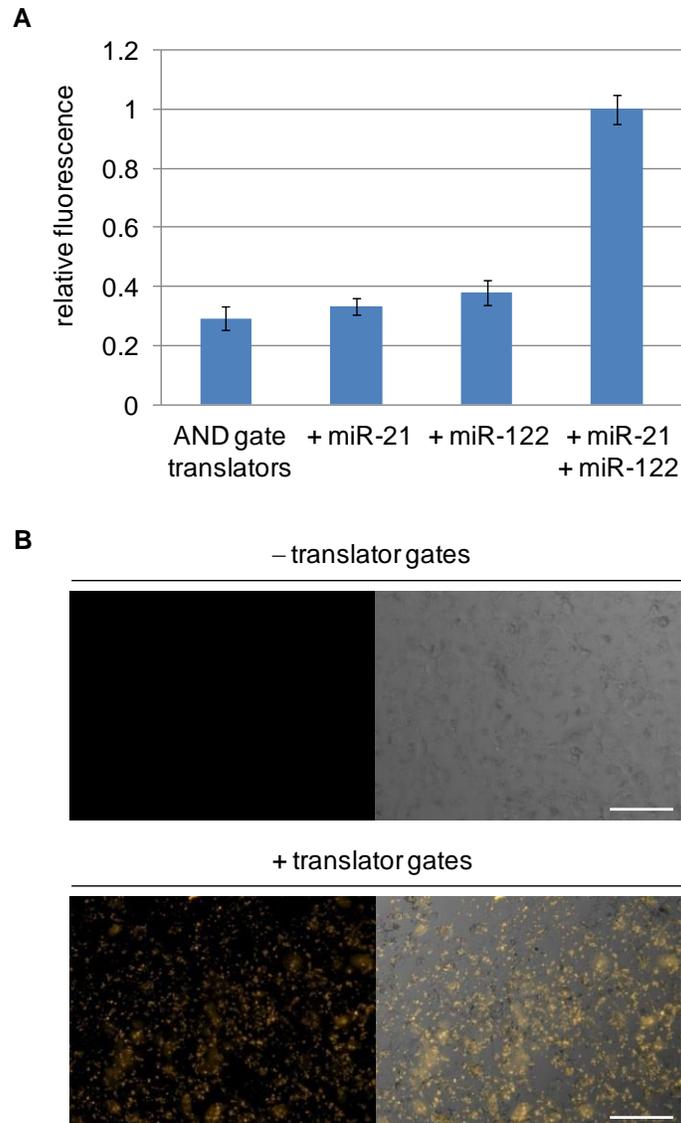
Figure 3.20: Engineered miR-21/122 translator-coupled AND gate.
A) The logic gate was tested for activation with the miR-21 and miR-122 RNA strands. TAMRA fluorescence was observed at 4 h and normalized to the activated AND gate. An average of three independent experiments is shown and error bars represent standard deviations. B) Huh7 cells were cotransfected with the miR-21/122 AND gate and translator gates, followed by imaging after 4 h. Scale bars indicate 200 μm.

Moreover, in HeLa cells that endogenously express miR-21 but not miR-122,[239] the addition of synthetic miR-122 was sufficient for gate activation, demonstrating that dual-sensing miRNA gates can be activated with a mix of endogenous and synthetic miRNA inputs (Figure 3.21).
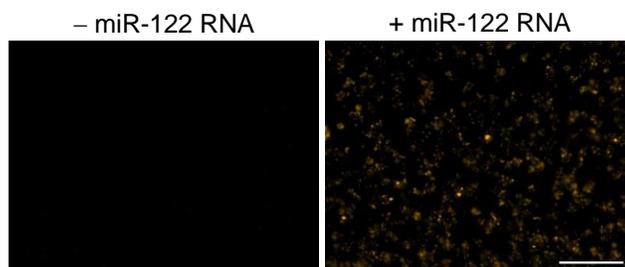
Figure 3.21: miR-21/122 translator-coupled AND gate application in HeLas.
Cells were transfected with the miR-21/122 gate and both translator gates in the absence and presence of synthetic miR-122 RNA then imaged after 4 h. Scale bar indicates 200 μm.

The successful activation of DNA logic gates in live cells lays the groundwork to pursue the development of biological applications for complex molecular computation. Light-directed DNA assembly within mammalian cells was demonstrated as a stepping stone toward cellular activation of synthetic DNA computation circuits. Logic gates were developed and tested for in vivo activation of a fluorescent output dependent upon the presence of endogenous miRNAs. The cell-specific miRNA detection was achieved through DNA computation with AND gates that were engineered to recognize, with high specificity, either miR-21 or miR-122 or both miRNAs. The response of an endogenously activated logic gate was demonstrated to identify reduced miRNA function due to environmental changes, for example, exposure to miRNA small molecule inhibitors. The ability to compute the presence (or absence through the application of other Boolean logic gates) of multiple miRNAs in live cells has implications in the diagnosis and treatment (e.g., through the release of therapeutic agents) of human disease, especially cancer.

Moreover, this modular computation approach can be easily reconfigured for the detection of other endogenous nucleic acid inputs. The DNA devices developed have potential application in the detection of miRNA markers for cellular expression profiling and can be expanded to the in vivo detection of other miRNAs, as well as interfaced with additional gates to develop complex DNA-based circuits that can recognize and respond to multi-miRNA patterns.

3.2.1 Methods and Materials

Logic Gate Duplex Purification. See Section 5.6 for specific information on DNA computation techniques. Logic gate duplexes were purified following protocol 5.6.1. All oligonucleotide sequences are shown in the following tables: a letter prior to the base indicates sugar modification (r = RNA, m = 2'OMe), an asterisk (*) indicates phosphorothioate linkages between bases, toe-hold regions are underlined, NPOM-caged thymidines are highlighted in red and bolded ("**T**"), Q = BHQ2 quencher, F = Tetramethylrhodamine (TAMRA) fluorophore:

Table 3.2: Photochemically activated AND gate.
Caged DNA oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
A	TATGGTTGTTTATGTGTTCCCTGATCTTTAGCCTTA
A_{4T}	TATGG**T**TGTTTATG**T**GTTCCC**T**GATCTT**T**AGCCTTA
B	AGATGTTAGTTTCACGAAGACAATGATTAAGGC
G_F	TGTTTATGTGTTCCCTGATCTTTAGCCTTAATCATTGTCTTCGTGAAACTAACATCTAAC-F
G_Q	Q-GTTAGATGTTAGTTTCACGAAGACAATGAT
G_T	TAAGGCTAAAGATCAGGGAACACATAAACAACCATA

137

Table 3.3: miR-21 AND gate.

Caged DNA oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
miR-21	<u>rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA</u>
B$_{21}$	GTTAGATGTTAGTTTCACGAAGACAATGAT<u>TCAACA</u>
B$_{21-4T}$	GT**T**AGATGT**T**AGTT**T**CACGAAGACAATGAT**T**CAACA
G$_{FM}$	F-GTTAGATGTTAGTTTCACGAAGACAATGAT
G$_{Q21}$	ATCAGACTGA<u>TGTTGA</u>ATCATTGTCTTCGTGAAACTAACATCTAAC-Q
G$_{T21}$	TCAACATCAGTCTGAT<u>AAGCTA</u>
antagomir	mA*mU*mC*mA*mA*mC*mA*mU*mC*mA*mG*mU*mC*mU*mG*mA* mU*mA*mA*mG* mC*mU*mA

Table 3.4: miR-122 AND gate.

Strand	Sequence (5' → 3')
miR-122	<u>rUrGrGrArGrUrGrUrGrArCrArArUrGrGrUrGrUrUrG</u>
B$_{122}$	GTTAGATGTTAGTTTCACGAAGACAATGAT<u>CAAACA</u>
G$_{FM}$	F-GTTAGATGTTAGTTTCACGAAGACAATGAT
G$_{Q122}$	GTGACAATGG<u>TGTTTG</u>ATCATTGTCTTCGTGAAACTAACATCTAAC-Q
G$_{T122}$	CAAACACCATTGTCAC<u>ACTCCA</u>

Table 3.5: miR-21 / miR-122 translator-coupled AND gate.

Strand	Sequence (5' → 3')
G$_{F21/122}$	F-GTTAGATGTTAGTTTCACGAAGACAATGAT
G$_{T21/122}$	TGGAGTTAAAGATCAGGGAACACATAAACA<u>TCAACA</u>
G$_{Q21/122}$	TGTTTATGTGTTCCCTGATCTTTA<u>ACTCCA</u>ATCATTGTCTTCGTGAAACTAACATCTAAC-Q
J$_{21\ Out}$	ATCAGACTGA<u>TGTTGA</u>TGTTTATGTGTTCCCTGATCTTTAACTCCA
K$_{21}$	ACATCAACATCAGTCTGAT<u>AAGCTA</u>
L$_{122}$	CAAACACCATTGTCACACT<u>CCAATC</u>
M$_{122\ Out}$	GTTAGATGTTAGTTTCACGAAGACAATGAT<u>TGGAGT</u>GTGACAATGG

Gate Functional Examination. See Section 5.6 for specific information on DNA computation techniques. Gate fluorescent output signals were determined following protocol 5.6.2.

Cellular Logic Gate Transfection. See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. Cells were passaged into a 96-well plate and grown to ~70% confluence within 24 h following protocol 5.4.1. The cells were transfected with logic gates using X-tremeGENE siRNA transfection reagent following protocol 5.6.3.

Live Cell Imaging of TAMRA Fluorescence. See Section 5.4 for specific information on cell culture techniques. Transfected cells were imaged on a Zeiss Observer Z1 microscope (20X magnification) following protocol 5.4.5.

Imaging of Subcellular TAMRA Fluorescence Localization. See Section 5.4 for specific information on cell culture techniques. HeLa cells were seeded into 4-well chamber slides following protocol 5.4.1, and transfected with the miR-21 AND gate at 50 nM with 200 nM B_{21} input strand using 5 µL of X-tremeGENE (Roche) in 1 mL of Opti-Mem (Invitrogen) at 37 °C for 4 h. After 4 h, the cells were fixed then nuclear counter stained with DAPI (Invitrogen) following protocol 5.4.4, and imaged with a Zeiss Observer Z1 (63X magnification) following protocol 5.4.5.

Photochemical Activation in Mammalian Cells. Cells were incubated for 1 h post-transfection and UV irradiated for 5 min on a 365 nm transilluminator following protocol 5.4.3. Cellular images were taken 1 h post UV irradiation following protocol 5.4.5.

Small Molecule and Antagomir Treatment of HeLa Cells. HeLa cells were treated with (*E*)-4-(phenyldiazenyl)-*N*-(prop-2-ynyl)benzamide (10 µM) in 1% DMSO/DMEM growth media for 48 h. The dose–response treatments were performed in 1% DMSO/DMEM at decreasing concentrations of the small molecule (as low as 10 nM). Alternatively, cells were transfected with 100 pmol of the miR-21 antagomir using X-tremeGENE following protocol

5.4.2. Logic gate transfections were performed after 48 h incubations using X-tremeGENE following protocol 5.6.3.

3.3   DNA Computation in Mammalian Cells: OR Gate, OR/AND Circuit, and Release of a Splice-Switching Oligonucleotide

The next step taken to expand the DNA logic gate-based detection of endogenous miRNAs was the incorporation of new Boolean logic functions into miRNA-based computation devices. These devices will enable logic operations beyond the AND gate function described in Sections 3.1 and 3.2.

As the first example, an OR gate was designed for the detection of a single miRNA from a pair of inputs. The concept of a duplex structure containing two overhangs for OR gate functionality was previously described for an aptamer-based OR gate for ligand binding,[212] and has been applied to toe-hold strand displacement for the development of functional protein outputs,[242] but has not been utilized for endogenous nucleic acid detection. Other methods for DNA-based OR gates with nucleic acid inputs have been demonstrated, including stem-loop hairpin molecular beacon structures[243] and translator gates.[208] This OR logic gate utilizes an internal fluorophore:quencher duplex that contains two toe-hold specific for miRNA sequences, one on either side. The toe-holds for miR-21 and miR-122 were amended to the fluorophore and quencher strands, such that when either miRNA is present the duplex will dissociate through toe-hold mediated strand exchange, and a fluorescent output will be observed (Table 3.6 and Figure 3.22A). Therefore, there are 3 conditions that will produce the output signal for the miR-21 OR miR-122 logic gate: miR-21 present, miR-122 present, or both miRNAs present (Figure 3.22B).
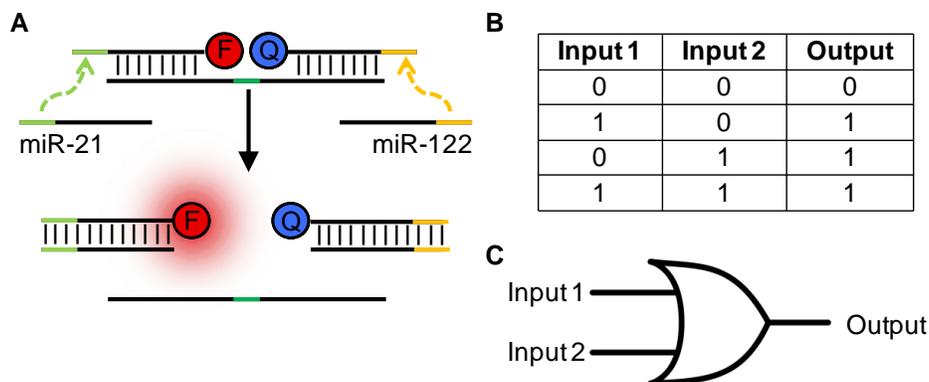
Figure 3.22: miR-21 OR miR-122 logic gate.
A) The nucleic acid-based Boolean OR gate activation cascade that responds to single stranded miR-21 and miR-122 inputs. Toe-hold regions are shown for the miR-21 (green) and miR-122 (orange), along with corresponding arrows representing hybridization steps. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). B) OR gate truth table. C) OR circuit diagram.

The linker length between the internal fluorophore and quencher strands was investigated to determine the appropriate spacing for fluorophore quenching and proper duplex formation, since the dye modifications could introduce steric hindrance of base hybridization due to the internal placement within the duplex. Therefore, linker lengths of 1-3 cytosine nucleotides were examined in solution for different sized duplexes, and the corresponding response to synthetic miRNA inputs was observed (Figure 3.23A). This analysis demonstrated that the quenching of the TAMRA signal and the activation by miRNA inputs was not dependent on linker length. The OR gate containing a single C linker was utilized for all subsequent experiments. It is worth noting that the fluorescent activation of the OR logic gate with miR-122 is less than that observed for the miR-21 input, and that the output is slightly increased with both miRNAs compared to either individually (Figure 3.23B).
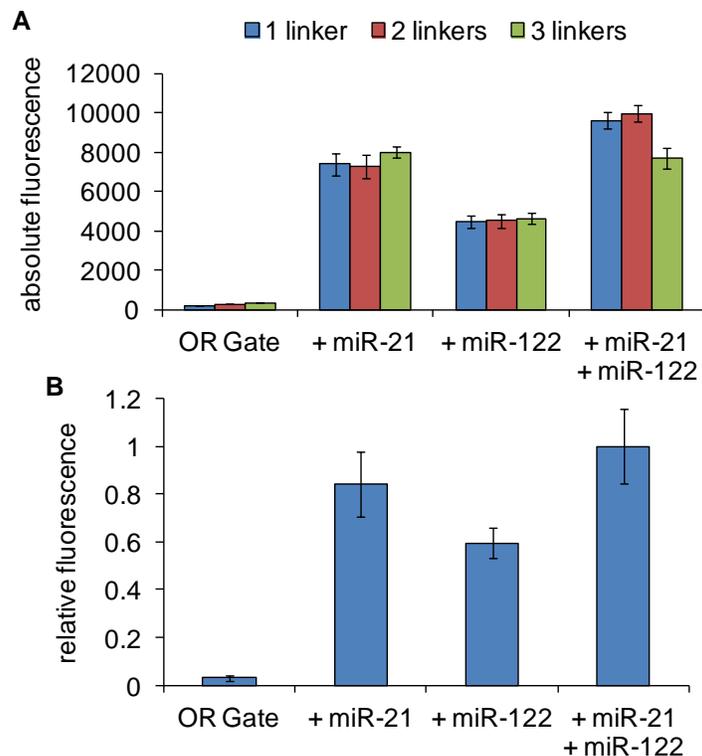
Figure 3.23: Analysis of miR-21 OR miR-122 logic gate activation.
A) The OR gates (200 nM) were incubated with the miRNA inputs (800 nM) and TAMRA fluorescence was measured at 4 h. B) Relative fluorescence of the OR gate containing a single C linker, normalized to the signal generated from the presence of both miRNA inputs. An average of three independent experiments is shown and error bars represent standard deviations.

The miR-21 OR miR-122 logic gate was then applied to the detection of endogenous miRNAs in mammalian cells, to confirm biological function of the new duplex structure. The logic gate was transfected into five cell lines (Figure 3.24), with varied miRNA expression profiles as determined through miRNA-based database searches.[244] One of the cell lines does not overexpress either miRNA of interest (HEK293T), while the four others show mixed profiles for the overexpression of either or both miRNAs (HeLa, HepG2, Huh7, PC3). Unfortunately, some background fluorescent was observed in the HEK293T cell line that does not overexpress either miRNA. However, there is an increase in fluorescence observed in cell lines that overexpress the miRNAs, as observed in the four positive cell lines. Variability in the total fluorescent output

was observed for different cell lines, independent of miRNA signatures. For example, PC3s showed the greatest logic gate activation while only overexpressing miR-21, compared to cell lines that overexpressed both miRNAs (HepG2, Huh7). Since this finding does not support the increased activation of the OR gate with both miRNAs, as shown in Figure 3.22B, this phenomenon is likely attributable to either differences in relative miRNA expression levels, miRNA copy numbers, or transfection efficiency between the cells lines, i.e., cell lines that transfect more readily will have an increased uptake of the logic gate and show enhanced activation over cell lines that transfect poorly.
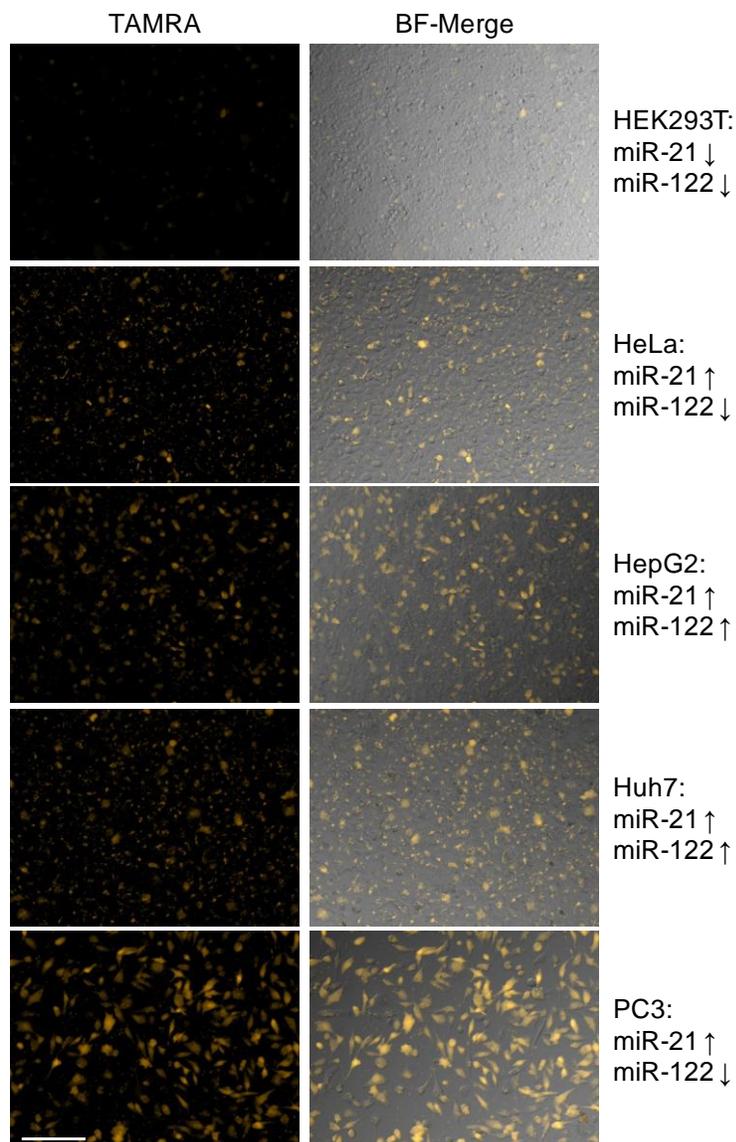
Figure 3.24: miR-21 OR miR-122 logic gate activation in cells.
The OR gate was transfected (200 nM) into 5 cell lines (HEK293T, HeLa, HepG2, Huh7, and PC3), with each relevant miRNA profile indicated on the right (↓ = low levels of miRNA expression, ↑ = high levels of miRNA expression), then imaged after 4 h. TAMRA and brightfield-merged channels are shown. Scale bar indicates 200 μm.

The development of connected miRNA-based logic gate circuits was then examined through the coupling an OR gate and an AND gate. This DNA computation module utilizes translator gates to recognize either miR-21 or miR-122 to generate the same output, which then interacts with a fluorophore:quencher duplex that contains an additional miRNA target site for

144

miR-125b (Table 3.7 and Figure 3.25). Since the function of this device is based on logic gate circuit connectivity by linking an OR gate to the activation cascade of an AND gate, the activation of a fluorescent output signal is only produced if the miR-21 OR miR-122 condition is met in conjunction with miR-125b. Therefore, there are 3 conditions that will generate the output signal for the (miR-21 OR miR-122) AND miR-125b circuit: miR-21 and miR-125b present, miR-122 and miR-125b present, or all three miRNAs present (Figure 3.25B).
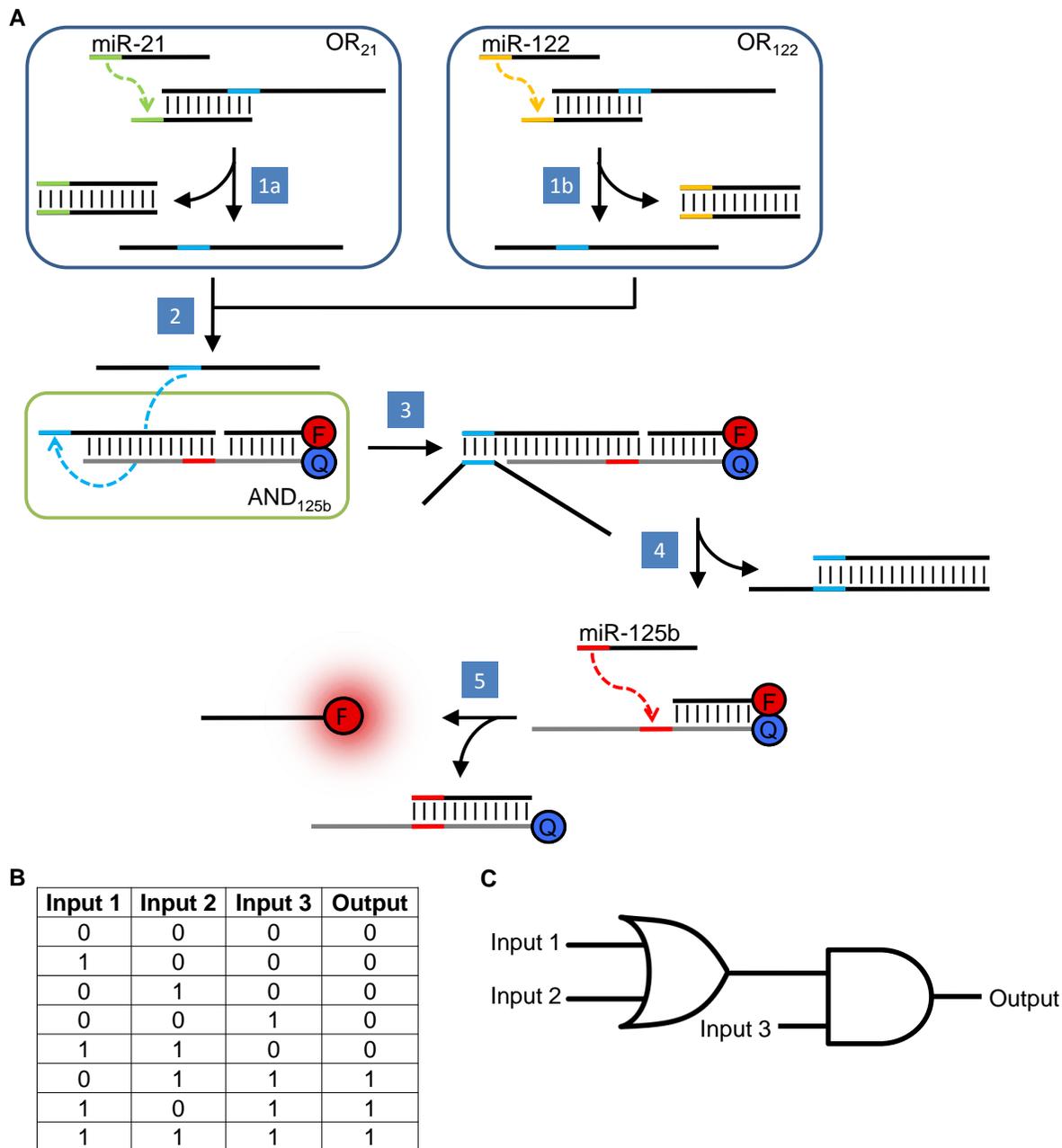
Figure 3.25: (miR-21 OR miR-122) AND miR-125b logic circuit.
A) The nucleic acid-based (miR-21 OR miR-122) AND miR-125b logic circuit is activated with either of two OR inputs only in conjunction with the AND input. The OR/AND logic circuit activation cascade is shown, with single stranded inputs for miR-21, miR-122, and miR-125b. OR translator gates (blue boxes) interact with two miRNA inputs that displace a activator strands with the same toe-hold recognition sequences, which interact with the AND miR-125b gate (green box). Toe-hold regions are shown for the miR-21 (green), miR-122 (orange), and miR-125b (red), along with corresponding arrows representing hybridization steps. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). B) The OR/AND truth table. C) The OR/AND circuit diagram.

146

The (miR-21 OR miR-122) AND miR-125b logic circuit was first analyzed in a test tube in the presence of each activating translator gate with the corresponding miRNA and miR-125b (Figure 3.26). Proper circuit function was observed, as the fluorophore gate duplex had minimal background with the translator gates alone, and each condition from the truth table that should produce a fluorescent output was confirmed ([1,0,1] [0,1,1] and [1,1,1]). However, there was some leakiness observed in the presence of miR-125b alone, suggesting the circuit design may limit the detection of endogenous miRNA patterns through an increase in miR-125b sensitivity.
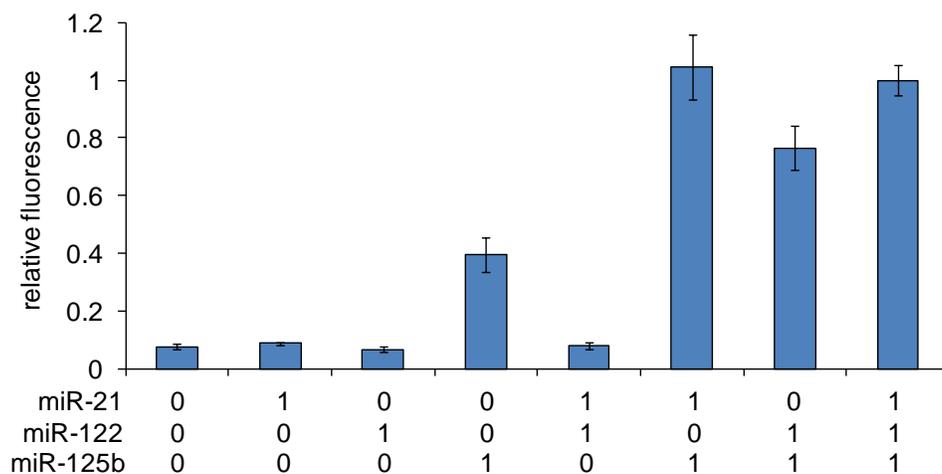
Figure 3.26: Analysis of the OR/AND logic circuit activation.
The fluorophore:quencher AND gate duplex (200 nM) was incubated with the OR translator gates (200 nM) and corresponding miRNA inputs (800 nM). TAMRA fluorescence was observed at 4 h, and set relative to the [1,1,1] input activation with all three miRNAs. Error bars represent standard deviations from an average of three independent experiments

The (miR-21 OR miR-122) AND miR-125b logic circuit was then applied to the endogenous detection of complex miRNA profiles in mammalian cells. The five cell lines described above for the OR gate were also utilized in the analysis of the (miR-21 OR miR-122) AND miR-125b logic circuit cellular activation (Figure 3.27), with varied miRNA expression profiles as determined through miRNA-based database searches.[244] Two positive cell lines that

contain either miR-21 or miR-122 in the presence of miR-125b (HeLa and PC3), which should active the logic and generate a fluorescent output based on the truth table. Low background was observed in the three cell lines that do not show overexpression of the required miRNAs for gate activation (HEK293T, Huh7, HepG2). The successful detection of multiple miRNAs with a connected DNA computation module was demonstrated, as fluorescent activation of the (miR-21 OR miR-122) AND miR-125b logic circuit was observed selectively in HeLa and PC3 cell lines. However, the logic circuit activation was not robust, and high variability existed between experiments. Logic gate transfections were repeated with each cell line 5-10 times, and both background fluorescence in the negative control HEK293T cell line as well as lack of activation in the positive HeLa cell line have occasionally been observed. Thus, additional optimization of the computation system may be required for the application of multi-layer DNA logic gate circuits to the robust detection of complex miRNA signatures in live cells.
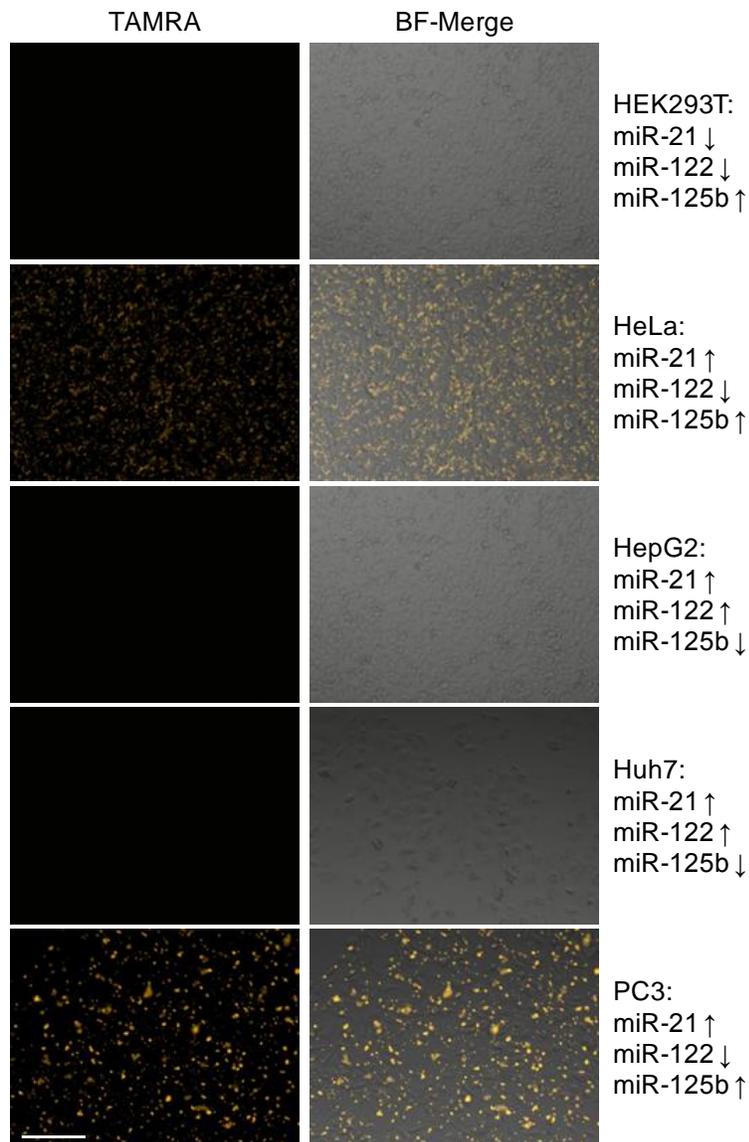
Figure 3.27: (miR-21 OR miR-122) AND miR-125b logic circuit in cells.
The OR/AND circuit fluorophore:quencher duplex was transfected (50 nM) with the translator gates (200 nM) into 5 cell lines (HEK293T, HeLa, HepG2, Huh7, and PC3), with each relevant miRNA profile indicated on the right ($\downarrow$ = low levels of miRNA expression, $\uparrow$ = high levels of miRNA expression), then imaged after 4 h. TAMRA and brightfield-merged channels are shown. Scale bar indicates 200 μm.

In order to analyze the location of miRNA detection within live cells and potentially identify miRNA-rich regions or cellular compartments, subcellular imaging with nuclear staining was performed for the HeLa and PC3 cellular transfections of the (miR-21 OR miR-122) AND

miR-125b logic circuit (Figure 3.28). Indeed, localization of the activated logic gates appears similar to Figure 3.14 (miR-21 AND gate), in a phenomenon that needs to be further examined.
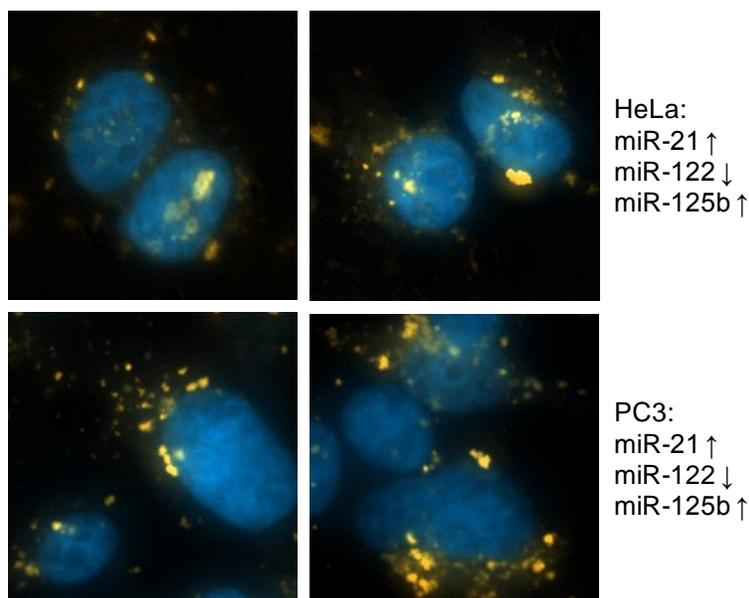
Figure 3.28: Cellular localization of the (miR-21 OR miR-122) AND miR-125b logic circuit. Cells were transfected with the (miR-21 OR miR-122) AND miR-125b circuit fluorophore:quencher duplex (50 nM) and translator gates (200 nM) for 4 h, then fixed and nuclear counter stained with DAPI. Imaging was performed to observe gate activation with sub-cellular resolution.

It has been shown that miRNAs are often associated with endosomal compartments, in addition to cytoplasmic foci P-bodies and GW-bodies.[245] Anti-miRNA oligonucleotides are also associated with endosomal compartments, suggesting that the targeting of miRNAs occurs within the same endosomal compartments.[246] Therefore, it is a reasonable assumption that the DNA logic gates are localized in endosomes and are activated by miRNA inputs within these compartments.[247] Initial application of an endosomal cellular pH indicator (pHrodo, Life Technologies) has been met with limited success (data not shown), due to spectral overlap between the pHrodo indicator and the TAMRA fluorophore utilized for the DNA logic gates, which will need to be further investigated before any conclusions can be made. A full

150

understanding of the subcellular trafficking of DNA logic gates and the corresponding target miRNA will require additional investigation to distinguish the importance of the localized fluorescence observed with activated miRNA-based logic gates in relation to miRNA targeting within the cell.

Since DNA computation provides a unique approach to the detection of miRNA patterns with oligonucleotides, the generated nucleic acid outputs can then be directly interfaced with cellular systems for regulation of gene expression. To this end, the release of a biologically relevant splice-switching oligonucleotide output (SSO, described in Section 2.5) in response to a DNA computation event was investigated. A DNA logic gate was designed to release an SSO in response to miR-21 (Table 3.8). This miR-21 SSO AND gate (Figure 3.29A) was applied to the regulation of alternative splicing (AS) pathways for the correction of an aberrant mutant intron. A luciferase reporter gene was used that contains an altered coding region with an aberrant splice site to impair functional protein expression until it is blocked by the sequence-specific SSO output (Figure 3.29B and C).

151
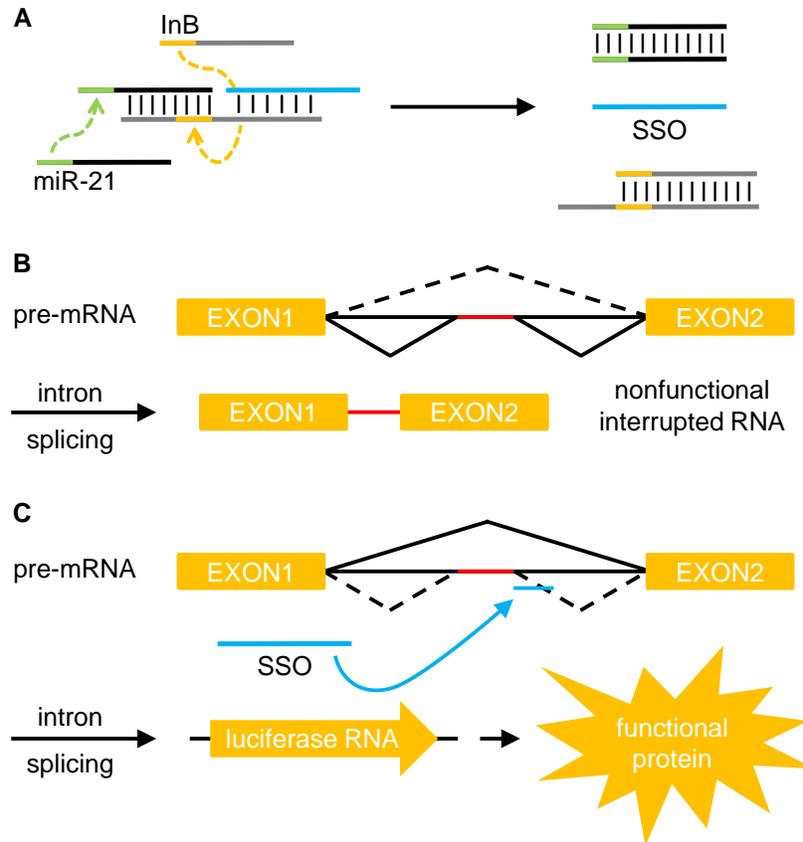
Figure 3.29: The miR-21 SSO AND gate.
A) Simplified schematic of the AND gate. Toe-hold regions are shown for the miR-21 (green) and InB (orange) input activation cascades, along with corresponding arrows representing hybridization steps. The SSO output is shown in light-blue. B) In the absence of the SSO, a mutant intron from the pre-mRNA is present in the mRNA sequence and disrupts formation of a functional gene product. C) The SSO binds to the target site, blocks the spliceosome from interacting with the target site, and creates an AS pathway that will remove the mutant intron from the mRNA allowing for the expression of functional protein (e.g., luciferase). Active splicing pathways are indicated with solid black lines, AS pathways are indicated with dashed lines, and the mutant intron is shown in red.

Initial experiments were performed with the pLuc705 plasmid[248] through transient transfection with the plasmid, SSO, and miR-21 SSO AND gate in either HEK293T or HeLa cells. However, transfection with the different oligonucleotide constructs proved to be challenging for reliable SSO activation under various conditions (data not shown). Therefore, a HeLa cell line stably expressing the interrupted luciferase gene containing a mutant β-globin intron (HeLa:Luc705, described in Section 2.5)[188] was used to analyze the miRNA-based logic

gate activity. Since the HeLa cell line shows high levels of miR-21 overexpression,[239] transfection of the AND gate with the InB input should result in the release of a regulatory SSOs to control AS pathways. To this end, transfections were performed with the SSO control and a range of conditions for the miR-21 SSO AND gate (Figure 3.30). The cell line showed low basal expression of luciferase and activation of AS pathways with SSO transfection, confirming previous results described in Section 2.5. However, transfection with the miR-21 SSO AND gate alone showed an increase in luciferase expression compared to nontreated cells at both 50 and 200 nM, suggesting that either degradation of the gate duplex or interaction of the 3' SSO overhang with the pre-mRNA were resulting in the premature release of the SSO. Additionally, the logic gate was not activated with either the InB input or both synthetic inputs. Therefore, the miR-21 SSO AND gate did not function as intended, as luciferase expression was observed independent of the required input strands, and the oligonucleotide transfection experiments were discontinued.
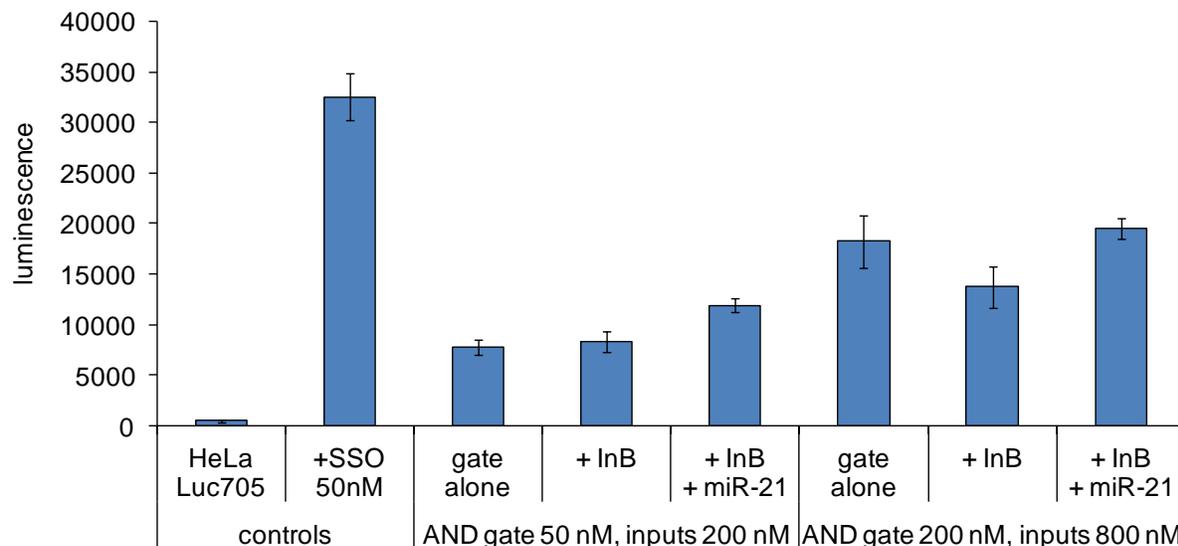
Figure 3.30: Analysis of miR-21 SSO AND gate driven luciferase expression.
HeLa:Luc705 cells were transfected with the SSO (50 nM) or the miR-21 SSO AND gate (50 and
200 nM) with synthetic input strands (200 or 800 nM). A luciferase assay was performed after 24
h incubation. Error bars represent the standard deviation of experimental triplicates

Since cellular transfection techniques proved ineffective at properly delivering the SSO

for activation of AS pathways in response to specific miRNA signatures, the use of "naked"

oligonucleotide delivery was subsequently investigated. A small molecule that increases

intracellular trafficking of oligonucleotides for enhanced pharmacological SSO activity[187] was

applied to the miR-21 SSO AND gate for delivery of nucleic acids into HeLa:Luc705 cells. The

gate duplex and input strands were added to the media for oligonucleotide delivery in the

absence of transfection reagent. The cells were soaked with the oligonucleotides overnight, then

treated with Retro-1 (Figure 3.31A) for 4 h. A luciferase assay was performed after an additional

4 h incubation (Figure 3.31B). The use of naked oligos significantly decreased the efficiency of

SSO activation for AS pathways and the expression of luciferase compared to the use of

transfection reagents as shown in Figure 3.30. There was no observable difference between the

luciferase expressed with the miR-21 SSO AND gate alone or with the corresponding inputs at

either concentration. Unfortunately, the Retro-1 treatment did not show any enhancement of the

154

SSO control or the logic gate activation, as no increases in luciferase expression from splice-switching was observed. Therefore, the use of Retro-1 for the delivery of SSO-based logic gates and intracellular trafficking of oligonucleotides was discontinued. The further understanding of DNA logic gate cellular localization and trafficking will be required to make progress on the development of a computation-based release for biologically active agents driven by the detection of certain miRNA expression patterns and potential therapeutic application.
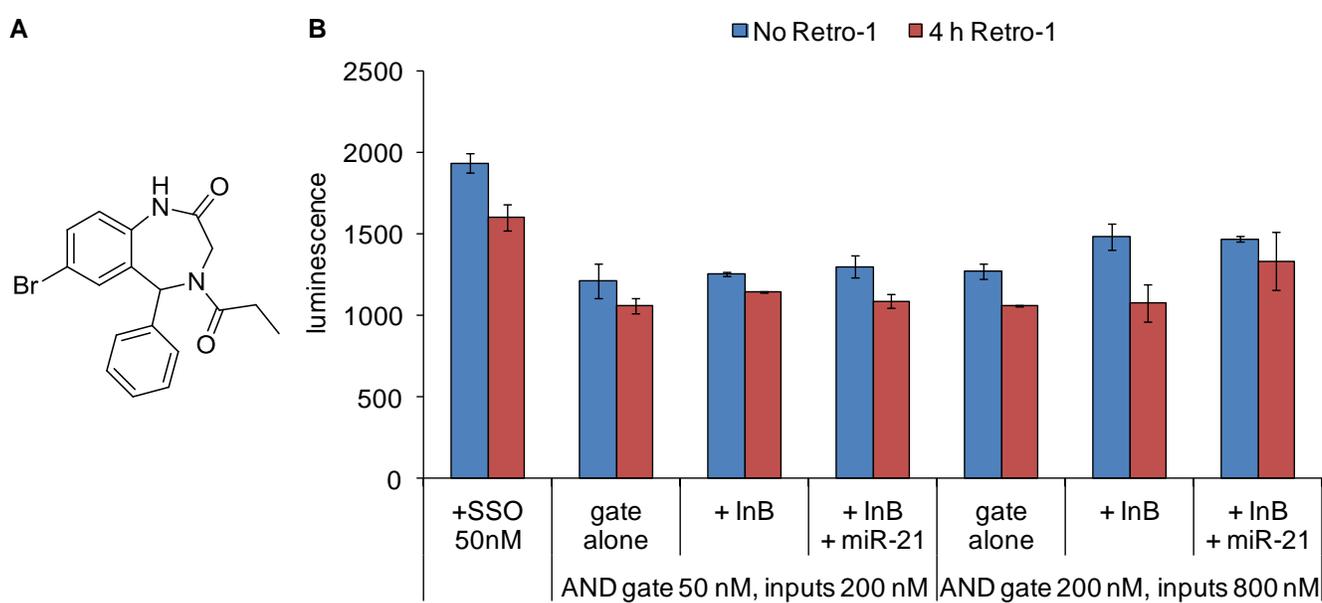
Figure 3.31: Analysis of miR-21 SSO AND cellular delivery technique.
A) Structure of Retro-1. B) HeLa:Luc705 cells were soaked overnight with the SSO (50 nM) or the miR-21 SSO AND gate (50 and 200 nM) with synthetic input strands (200 or 800 nM). Cells were then treated with Retro-1 (100 μM) for 4 h. A luciferase assay was performed 4 h post Retro-1 treatment. Error bars represent the standard deviation of experimental triplicates

In summary, three Boolean logic gates were analyzed for miRNA detection in mammalian cells: an OR gate, an OR/AND circuit, and an SSO-releasing AND gate. The OR gate was designed with an internal fluorophore:quencher pair and was confirmed in test tube experiments with synthetic miRNAs. Additionally, OR gate application in mammalian cells was successfully utilized for the detection of miR-21 or miR-122. However, compared to the AND

gate discussed in Section 3.2, increased background fluorescence was observed in cells lacking either miRNA. The OR/AND logic circuit was designed with a duplex containing a terminal fluorophore:quencher pair that interacts with two translator gates, and was also confirmed in solution with synthetic miRNAs. The OR/AND logic circuit successfully detected complex miRNA signatures in two positive control cell lines. However, reproducibility remains a hurdle in cellular applications, as varied levels of background fluorescence have been observed. Lastly, a miR-21 SSO AND gate was designed, but the release of a regulatory oligonucleotide could not be demonstrated in a cellular environment through transfections or intracellular trafficking driven by Retro-1 treatment. Future studies for these projects will focus on reducing background of logic gate fluorescence, increasing reproducibility, and confirming localization in subcellular compartments. To that end, a number of transfection techniques are available that will deliver the DNA logic gates to cells, such as cationic polymers, viral vectors, DNA nanoparticles, electroporation, or peptide conjugation.[249] These methods may introduce an alternative method to liposomal transfection (e.g., via X-tremeGENE siRNA reagent), which will allow the DNA logic gates to penetrate different subcellular compartments that were previously inaccessible. Alternative delivery methods may also generate less background or increased reproducibility for DNA logic gates applications, through protection of duplex degradation or specific subcellular localization.

3.3.1   Methods and Materials

Logic Gate Duplex Purification.   See Section 5.6 for specific information on DNA computation techniques. Logic gate duplexes were purified following protocol 5.6.1. All oligonucleotide sequences are shown in the following tables: a letter prior to the base indicates

sugar modification (r = RNA, m = 2'OMe), an asterisk (*) indicates phosphorothioate linkages between bases, toe-hold regions are underlined or bolded/italicized, linker regions highlighted green, Q = BHQ2 quencher, F = Tetramethylrhodamine (TAMRA) fluorophore:

Table 3.6: miR-21 OR miR-122 logic gate.

Strand	Sequence (5' → 3')
miR-21	rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA
miR-122	rUrGrGrArGrUrGrUrGrArCrArArUrGrGrUrGrUrUrUrG
OR gate 1C	ATCAGACTGATGTTGACTGGAGTGTGACAATGG
OR gate 2C	ATCAGACTGATGTTGACCTGGAGTGTGACAATGG
OR gate 3C	ATCAGACTGATGTTGACCCTGGAGTGTGACAATGG
F-OR21/122	F-TCAACATCAGTCTGATAAGCTA
Q-OR21/122	CAAACACCATTGTCACACTCCA-Q

Table 3.7: (miR-21 OR miR-122) AND miR-125b logic circuit.

Strand	Sequence (5' → 3')
miR-21	rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA
miR-122	rUrGrGrArGrUrGrUrGrArCrArArUrGrGrUrGrUrUrUrG
miR-125b	rUrCrCrCrUrGrArGrArCrCrCrUrArArCrUrUrGrUrGrA
21-Comp	*AGTAGT*TCAACATCAGTCTGATAAGCTA
122-Comp	*AGTAGT*CAAACACCATTGTCACACTCCA
21-Act	ATCAGACTGATGTTGA*ACTACT*CGTAGGTGTAGGAAAG*TCACAA*
122-Act	GTGACAATGGTGTTTG*ACTACT*CGTAGGTGTAGGAAAG*TCACAA*
Toe-hold	TTGTGACTTTCCTACACCTACG*AGTAGT*
F-OR/AND	F-TCCCTGAGACCCTAAC
Q-OR/AND	CGTAGGTGTAGGAAAGTCACAAGTTAGGGTCTCAGGGA -Q

Table 3.8: miR-21 SSO AND gate.

Strand	Sequence (5' → 3')
miR-21	rUrArGrCrUrUrArUrCrArGrArCrUrGrA<u>rUrGrUrUrGrA</u>
SSO	mG*mU*mU*mA*mU*mU*mC*mU*mU*mU*mA*mG*mA*mA*mU*mG*mG*mU*mG*mC
Gate$_{SSO}$	TTCTAAAGAATAAC*TAGCTT*ATCAGACTGA
InB	*AAGCTA*GTTATTCTTTAGAA
Toe$_{SSO}$	<u>TCAACA</u>TCAGTCTGATAAGCTA

Gate Functional Examination. See Section 5.6 for specific information on DNA computation techniques. Gate fluorescent output signals were determined following protocol 5.6.2.

Cellular Logic Gate Transfection. See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. Cells were passaged into a 96-well plate and grown to ~70% confluence within 24 h following protocol 5.4.1. The cells were transfected with logic gates using X-tremeGENE following protocol 5.6.3.

Live Cell Imaging of TAMRA Fluorescence. See Section 5.4 for specific information on cell culture techniques. Transfected cells were imaged on a Zeiss Observer Z1 microscope (20X magnification) following protocol 5.4.5.

Imaging of Subcellular TAMRA Fluorescence Localization. See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. HeLa and PC3 cells were seeded into 4-well chamber slides following protocol 5.4.1, and transfected with the (miR-21 OR miR-122) AND miR-125b circuit at 50 nM with 200 nM translator gates using 5 μL X-tremeGENE (Roche) in 1 mL of Opti-Mem (Invitrogen) at 37 °C for 4 h. After 4 h, the cells were fixed then nuclear counter stained with

DAPI (Invitrogen) following protocol 5.4.4, and imaged with a Zeiss Observer Z1 (63X magnification) following protocol 5.4.5.

Luciferase assay: See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. HeLa:Luc705 cells were passaged into a 96-well plate and grown to ~70% confluence within 24 h following protocol 5.4.1. The cells were transfected with the SSO control (50 nM), the miR-21 SSO AND gate (50 or 200 nM), and the miR-21/InB inputs (200 or 800 nM) using X-tremeGENE following protocol 5.6.3. Additionally, cells were soaked with the oligonucleotide reagents (same concentrations as above) in DMEM overnight at 37 °C. Retro-1 treatment was performed at 100 μM in DMEM with dilution from a 20 mM DMSO stock solution for 4 h. A Bright-Glo assay (Promega) was performed 24 h post transfection or 4 h post Retro-1 treatment by removing 150 μL of cell media, adding 50 μL of the kit reagent, and shaking for 20 min prior to reading chemiluminescence on a Tecan M1000.

3.4   Optically Controlled Signal Amplification for DNA Computation

This material was reprinted, in part, with permission from Prokup, A.; **Hemphill, J.**; Liu, Q.; Deiters, A. *ACS Synth. Bio*. **2015**, available online. Oligonucleotide syntheses and assistance in fuel-catalyst cycle design were performed by the author of this thesis. All experiments were conducted by Alexander Prokup in the Deiters lab.

In order to operate, DNA computation devices require an exchange of DNA strands between logic gates, which can become inefficient in large circuits, leading to a dampening in signal.[208] Options are available to amplify a low DNA output signal: for example, the use of

159

branched DNA to accumulate labeled DNA strands on an output strand,[250] a kinetically trapped metastable DNA fuel,[251] the hybridization chain reaction (HCR),[252, 253] and an entropy-driven fuel-catalyst cycle.[206, 254] However, the aforementioned amplification methods are limited in their ability to be externally controlled, particularly in a temporal and spatial fashion. While these signal amplifiers function efficiently, they cannot be switched ON or OFF quickly and noninvasively. Therefore, light-regulated variants of HCR and the fuel-catalyst cycle amplification were developed for DNA computation. This design relies on introducing nucleobase-caging groups into DNA strands to sterically block DNA/DNA hybridization until irradiation with UV light induces caging group removal and DNA duplex formation. Thus, a simple chemical modification to an existing structure enables the DNA amplification devices to be either turned ON or OFF through application of photochemical triggers in a spatial and temporal manner. To move the capabilities of these devices beyond solution-based operations, the components were embedded in agarose gels. Irradiation with customizable light patterns and at different time points demonstrated both spatial and temporal control. Overall, the addition of small light-cleavable photocaging groups to DNA signal amplification circuits enabled conditional control as well as fast photocontrol of signal amplification.

HCR allows for the detection of small concentrations of nucleic acids by generating an amplified signal through the opening of metastable hairpins to form a long nicked duplex, even in complex biological environments.[255] Three components are required for the reaction: two hairpins and an initiator strand. In the absence of the initiator, the hairpins will not cross react, as there are no complementary sequences exposed. However, when the initiator is added, it will hybridize to the toe-hold of hairpin 1 (H1) and expose a toe-hold for hairpin 2 (H2). After H2 binds to H1, a new toe-hold will be revealed, allowing the concatemer to continue growing. Overall, the presence of the input signal (i.e., the initiator strand N) is amplified through the

production of high molecular weight duplexes. To obtain optical control over HCR, the initiator

strand was blocked with four nucleobase caging groups to prevent binding to H1 (Table 3.9).

Table 3.9: Initiator strands for photochemically activated HCR.
NPOM-caged thymidine residues are underlined, bolded, and highlighted red ("**T**"). Caged DNA
oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
Initiator N_0	AGTCTAGGATTCGGCGTGTATATA
Initiator N_4	AGTC**T**AGGA**T**TCGGCG**T**GTATA**T**A

This design utilizes the known hairpin and initiator sequences published by the Pierce

group,[252] with modified toe-hold regions to increase the number of thymidines. These hairpin

sequences have been designed to prevent premature signal generation in the absence of initiator,

and the extra thymidine residues also allowed for increased flexibility in selecting nucleobases

caging sites. Activation of the initiator strand was achieved by irradiation with UV light to

photochemically remove the caging groups. Consequently, UV irradiation can act as a switch to

turn on HCR (Figure 3.32A). No background was observed in the absence of UV light.

Irradiation of N_4 with UV light triggered the formation of higher molecular weight products

similar to the HCR products generated by N_0. In the absence of UV light, no HCR occurred and

no higher molecular weight products were formed (Figure 3.32B). Removal of the caging groups

through UV exposure restored activity to the initiator, which was evident by the formation of the

same HCR products as were generated by the noncaged initiator. Thus, photochemical control of

HCR was achieved through a synthetic modification of a single DNA component.
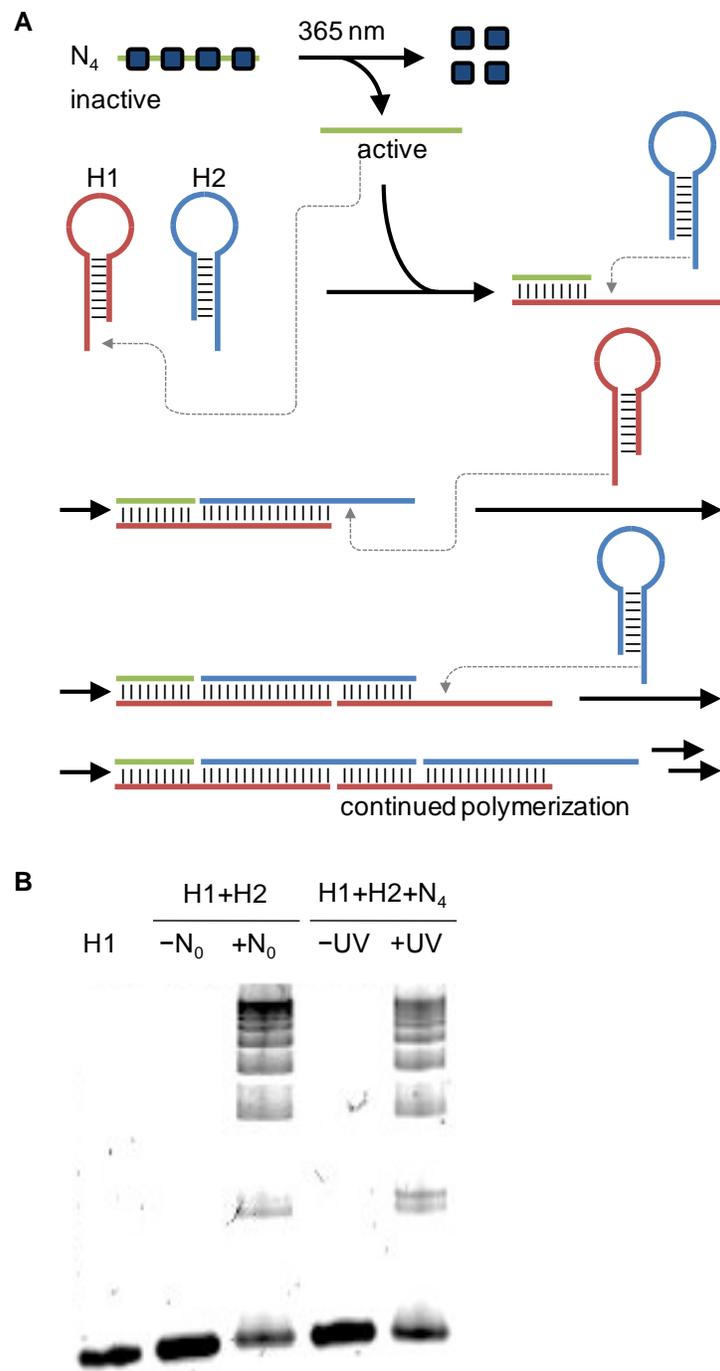
Figure 3.32: Photochemical control of HCR.
A) Schematic of the hybridization chain reaction with the caged initiator strand N_4. Oligonucleotides are shown as colored lines, and NPOM caging groups are represented by blue boxes. Two hairpins (H1 and H2) are metastable until light-triggering of the initiator strand causes the formation of higher molecular weight product strands. B) PAGE analysis of HCR reactions with noncaged (N_0) and caged initiator (N_4) strands. Image adapted with permission from Prokup et al, *ACS Synth. Bio.* **2015**. These experiments were conducted by Alexander Prokup.

A second DNA-based device to achieve signal amplification involves a fuel-catalyst cycle.[206] The cycle begins with binding of the catalyst strand to the substrate complex (duplex containing the substrate, signal, and output strands). After a toe-hold mediated strand exchange, the catalyst displaces the signal strand revealing a toe-hold for the fuel strand. Binding of the fuel strand completely removes the output strand and catalyst, which creates a waste duplex. Displacement of the catalyst strand allows the cycle to continue, and the output strand is then able to interact with the reporter gate, releasing the quencher strand from the fluorophore strand. Thus, an increase in fluorescence corresponds to an active cycle. The fuel-catalyst cycle amplification arises from the release of a surplus of signal and output strands from a limited supply of catalyst. If the catalyst strand is synthetically caged (Table 3.10), the caging groups will prevent activation of the cycle by blocking hybridization of the catalyst strand to the substrate complex. To turn the amplification cycle OFF, a caged inhibitor strand was conceived (Table 3.10), which is completely complementary to the catalyst thereby blocking its function.

Table 3.10: Strands for photochemically controlled fuel-catalyst cycle.
NPOM-caged thymidine residues are underlined, bolded, and highlighted red ("**T**"). Caged DNA oligonucleotides were synthesized according to the general protocol 5.2.

Strand	Sequence (5' → 3')
Catalyst C_0	CATTCAATACCCTACGTCTCCA
Catalyst C_4	C**A**TTC**A**AT**A**CCCT**A**CGTCTCCA
Inhibitor I_0	TGGAGACGTAGGGTATTGAATG
Inhibitor I_4	TGG**A**GACGT**A**GGGT**A**TTGA**A**TG

In the absence of caging groups, adding a catalyst strand to the substrate complex and fuel strands will release the signal and output strands. Removal of the catalyst caging groups with 365 nm light will initiate the cycle (Figure 3.33A). When the inhibitor strand is caged, hybridization to the catalyst is prevented and the cycle operates normally. After irradiation, the

inhibitor will bind to the catalyst prohibiting continuation of the cycle (Figure 3.33A). In order to generate a fluorescent output signal, a reporter gate can be added (Figure 3.33B). The gate will interact with the output strand, releasing the fluorophore strand. The free fluorophore is then able to emit a fluorescent signal. Activation or deactivation of the cycle through photochemical means enables reliable regulation of amplification by optically switching the cycle from either OFF → ON or ON → OFF.
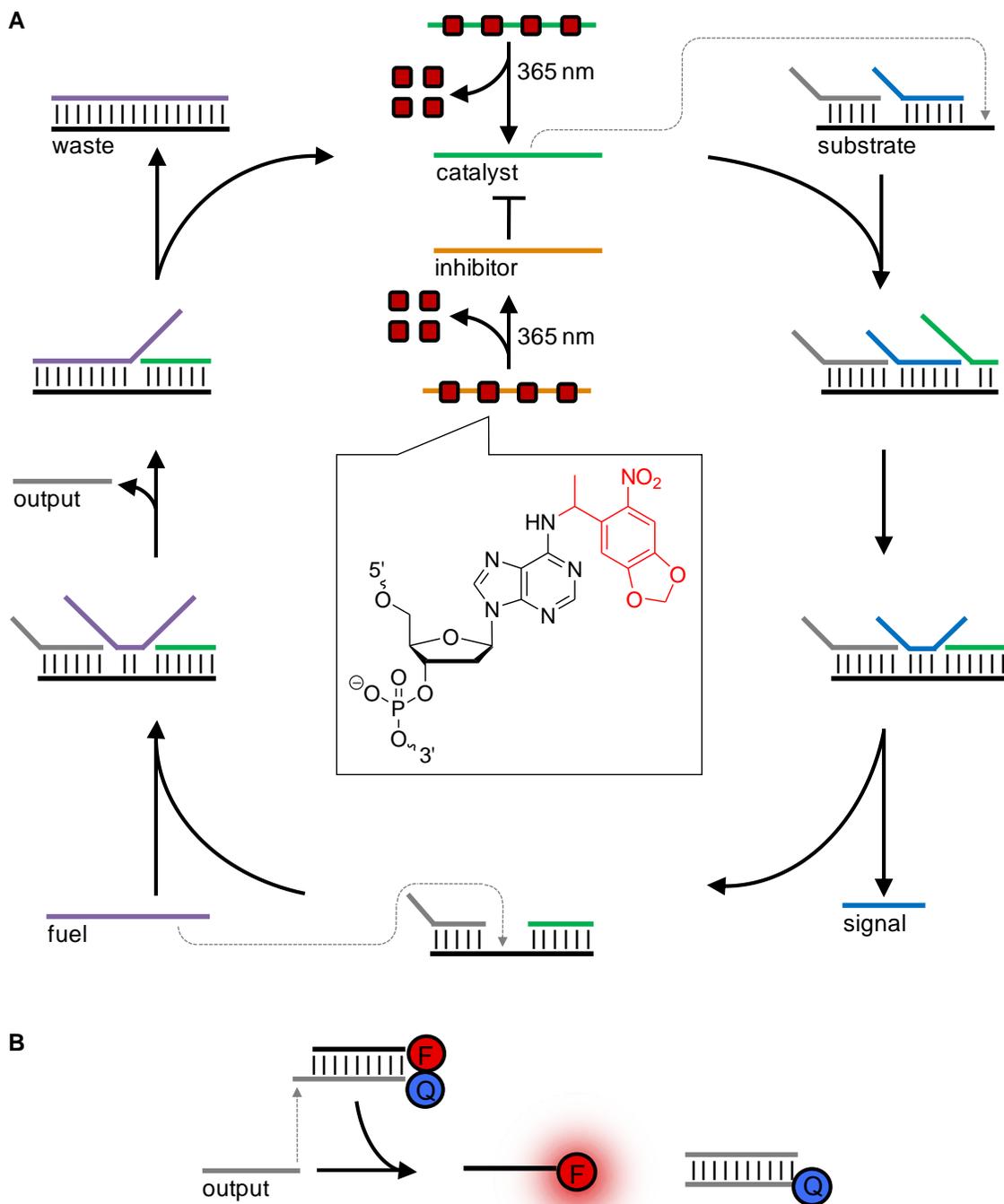
Figure 3.33: Photochemcial control of the fuel-catalyst cycle.
A) Schematic of the fuel-catalyst cycle with caged inhibitor and caged catalyst strands. B) Schematic of the fluorescent reporter gate. Colored lines represent DNA oligomers and red boxes indicate caging groups. The output strand is able to interact with the reporter gate, releasing a fluorophore strand. Quencher Q = Iowa Black RQ. Fluorophore F = tetramethylrhodamine (TAMRA). Image adapted with permission from Prokup et al, *ACS Synth. Bio.* **2015**.

Oligonucleotide strands and preformed gate structures were mixed in solution, and a fluorescence reporter gate was used to measure the activity of the fuel-catalyst cycle. After replacing C_0 with the caged C_4, no signal was produced, effectively turning the amplification cycle OFF (Figure 3.34A). Removal of the caging groups with UV light restored catalyst activity, generating a signal. Conversely, introduction of I_4 to the substrate complex, fuel strand, and C_0 did not affect normal operation of the cycle, and strand exchange cascades continued to produce signal. However, irradiating the caged inhibitor strand prevented amplification by sequestration of the catalyst strand, which switched the cycle OFF (Figure 3.34B). Thus, the activity of the fuel-catalyst cycle could be photochemically controlled through the use of the caged inhibitor or caged catalyst strand. More than 5-fold changes in fluorescence signal were observed upon light-activation of the caged catalyst or caged inhibitor strand. This corresponds to efficient photoswitching of the amplification cycle, since a similar change was observed for the addition of noncaged catalyst and inhibitor strands.
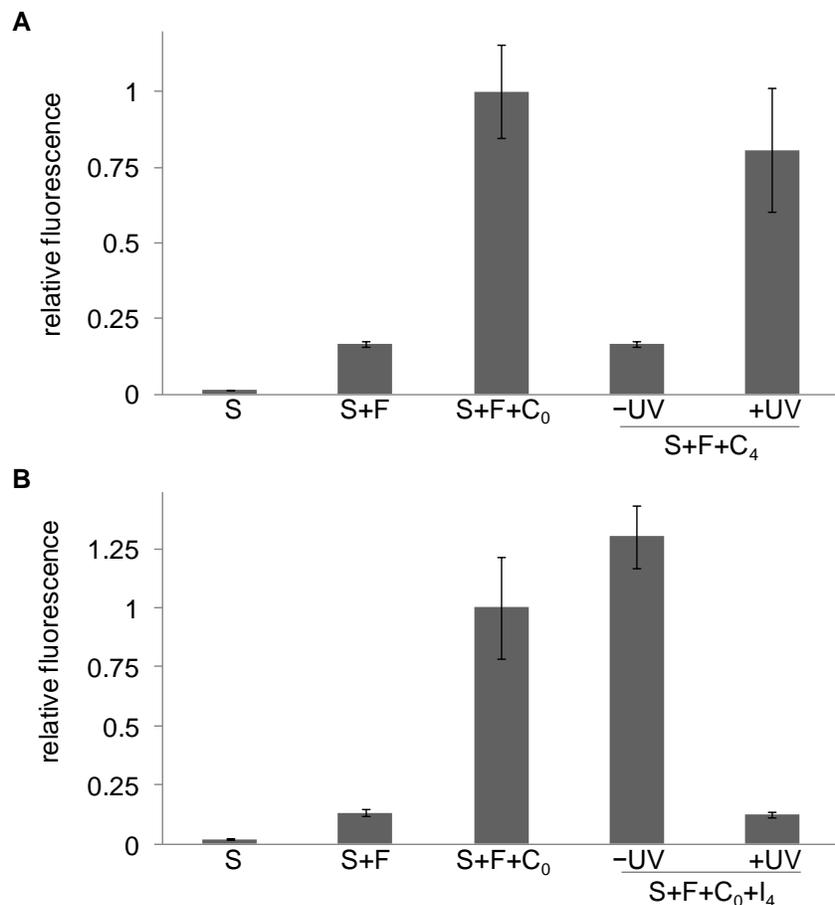
Figure 3.34: Optical OFF → ON and ON → OFF switching of the fuel-catalyst cycle.
A) Fuel-cycle with caged catalyst C_4. B) Fuel cycle with caged inhibitor I_4. Samples labeled +UV were irradiated with 365 nm light for 10 min before addition. Additional single letter abbreviations are used for substrate (S) and fuel (F). Excitation and emission wavelengths were 545 and 585 nm, respectively. Error bars represent standard deviations from three independent experiments. Image adapted with permission from Prokup et al, *ACS Synth. Bio.* **2015**. These experiments were conducted by Alexander Prokup.

With successful demonstration of optical control in a solution-based amplification cycle, this system was then transitioned semisolid media. Amplification of a signal inside a semisolid can expand the applications of DNA computation systems beyond solution-based devices, since it greatly constrains diffusion thereby enabling spatial control. Additionally, a solid structure creates a modular unit that could facilitate the physical separation of components in a DNA cascade. Much like electric components, embedded DNA computation devices can act as stand-alone elements of a larger circuit. To demonstrate spatial control of the fuel-catalyst cycle using

167

the caged catalyst C$_4$, oligonucleotide components were embedded into low-melt agarose. Spatially restricted illumination of the gel with a fiber optics probe (Figure 3.35A) or through a mask (Figure 3.35B) enabled the amplification cycle to be activated in specific and independent regions. Although masks allowed for customizable shapes, the edges were not as well-defined as those produced by an LED fiber optic light source. To create a gradient effect, the gel was irradiated for different time intervals (Figure 3.35C). The gradient demonstrated how signal intensity could be tuned by varying the applied UV irradiation. Variable light intensities will create diverse populations of signal intensity, adding depth to the recognition of an OFF → ON transition. The ability to create signals in any desired pattern using optical regulation could allow for better control in investigating or modeling biological events. Thus, spatial and temporal control of the gel-based fuel-catalyst cycle offers enhanced flexibility in controlling signal amplification.
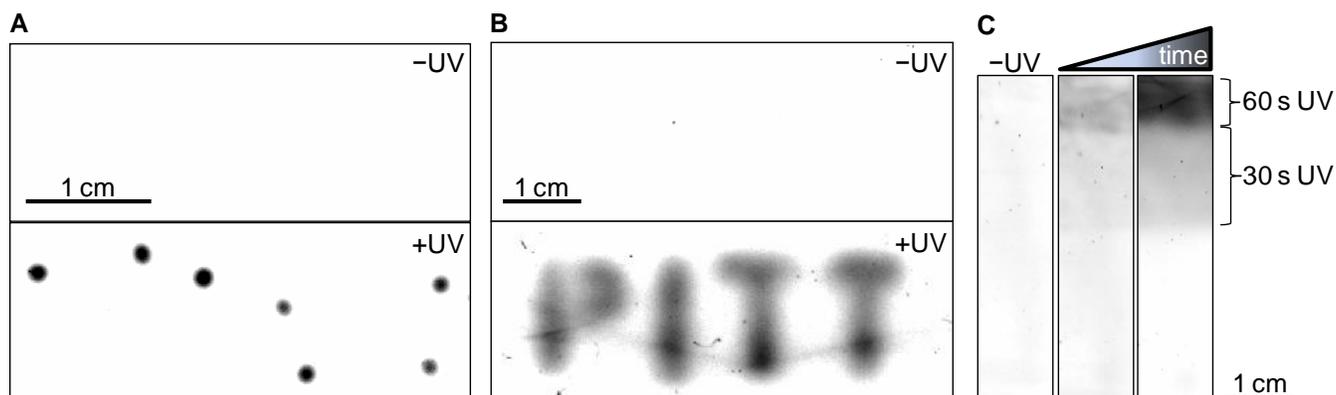
Figure 3.35: Spatial control of the fuel-catalyst cycle by using low-melt agarose gels. 1–2% gels were embedded with the substrate complex, fuel strand, and caged catalyst C$_4$. A) A fiber optics probe was used to irradiate (1 min, 365 nm) spatially independent areas in the shape of the big dipper constellation. B) Illumination through a mask was used to pattern "PITT". C) Multiple irradiation time intervals were used to create a gradient effect, tuning activation of the catalyst cycle. All fluorescent imaging was performed using green LED excitation, and emitted light was detected at 580–630 nm. Image adapted with permission from Prokup et al, *ACS Synth. Bio.* **2015**. These experiments were conducted by Alexander Prokup.

In conclusion, modification of oligonucleotides with nucleobase-caging groups enabled optical control over HCR and a fuel-catalyst cycle, DNA devices that allow for DNA signal amplification. Crucial DNA strands were modified with photocleavable caging groups to optically control the individual reaction circuits. For HCR, a caged initiator strand was synthesized. Upon decaging, the initiator strand was able to interact with the hairpins, causing amplification through DNA strand polymerization. Only a low concentration of initiator is necessary to start the HCR, which can be achieved through minimal light exposure. A fuel-catalyst cycle was also successfully optically switched from OFF $\rightarrow$ ON or ON $\rightarrow$ OFF by using either a caged catalyst or a caged inhibitor, respectively. To prevent DNA/DNA hybridization in the absence of illumination, and thus to control the cycle with light, four evenly spaced caging groups were added to the DNA strands. Quantification of the output was made possible with a reporter gate generating a fluorescent signal. Conducting light-activation in a semisolid containing the DNA circuits led to remarkable spatial and temporal control. Localized illumination of the gel embedded with the DNA circuits enabled triggering of signal amplification in customizable patterns as well as tunable gradients. For each cycle, light acts as a dependable switch for triggering computational events as it is tunable and noninvasive. Caged oligonucleotides represent a modular framework that can be easily fitted to existing DNA-based architectures.

The methodology development reported here may find application in more complex DNA computation circuits that contain output amplification. Temporal control enables precise sequencing of gate and subnetwork functions and, in the case of temporal control of an amplification cycle, allows for upstream circuit completion before output amplification, thereby potentially reducing the background signal by preventing premature activation. In addition,

temporal control of DNA circuits enables modification of any system (e.g., drug treatment of cells) that is interfaced with a DNA computation network before circuit activation. Light is an excellent external control element that can be used as a switch with very high temporal and spatial resolution without the need for other physical or chemical alterations (e.g., injections). The application of light-activated DNA circuits and amplification devices is especially advantageous in systems that do not allow for later addition of oligonucleotide triggers, for example, in semisolid media as shown here or in enclosed biological environments, such as organisms.

3.5 DNA Computation in Mammalian Cells: Signal Amplification of miRNA

In addition to developing a miR-21 AND gate (as described in Section 3.2), a miR-21 DNA fuel-catalyst cycle (as described in Section 3.4) was also tested to enable signal amplification of miRNA sensing in mammalian cells. The catalytic cycle has been described by Winfree to increase the fluorescent output signal through amplification of a single oligonculeotide,[206] and may decrease the detection limit requirements of miRNA logic gates for cellular applications. Oligonucleotide hybridization reactions have shown utility in the amplification and detection of nucleic acids in live cells. The hairpin DNA cascade reaction was recently applied to the amplification of low abundance mRNA in cells and subsequent detection using a fluorescent reporter gate.[218]Additionally, the cascade hybridization reaction (CHR) was applied to the amplification and detection of low abundance miRNAs in cells through the activation of FRET pairs (Figure 3.36).[231]
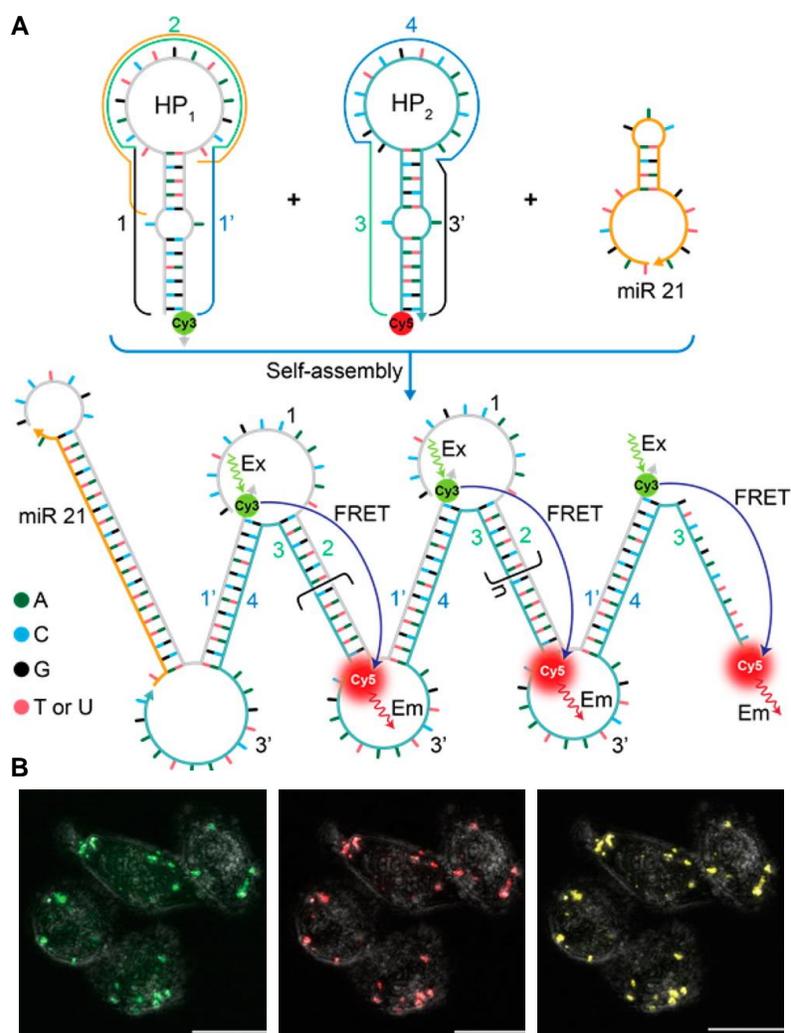
Figure 3.36: Cellular miRNA imaging using a DNA cascade hybridization reaction (CHR). A) Scheme illustrating the activation of a miRNA CHR. Molecular beacon hairpins containing a Cy3 and Cy5 fluorophore interact with the miR-21 target, forming linked hairpin chains. The fluorophores are then brought in close proximity to activate FRET donor emission. B) FRET analysis of miR-21 CHR in HeLa cells. Cy3 green channel (left), Cy5 red channel (middle) and FRET donor channel (right) are shown. Confocal mages were taken 4 h after transfection. Scale bars indicate 25 μm. Image adapted from Cheglakov et al, *J. Am. Chem. Soc.*, **2015**, 137(19):6116-6119. Copyright 2015 American Chemical Society.

Two programmable molecular beacon hairpins were modified to recognize miR-21 and produce a FRET emission signal through CHR (Figure 3.36A). A single input sequence is amplified from the continuously increasing FRET signal generated as the hairpin cascade reaction occurs, polymerizing long chains of connected hairpins. The miR-21 CHR device was evaluated through transfection into the HeLa cell line that overexpress miR-21, followed by

171

imaging of the fluorophores (Figure 3.36B). When both the hairpins were transfected into cells, the FRET donor emission could be observed, confirming the detection of miR-21. Therefore, it was determined that a similar approach utilizing the fuel-catalyst cycle would be tested for amplified miRNA detection in live cells.

In order to develop biologically relevant fuel-catalyst cycles, the toe-hold mediated displacement cascade was redesigned for the recognition of two miRNAs (miR-21 and miR-122). The substrate gate duplex contains a recognition domain that is sequence-specific for a 22 mer oligonucleotide catalyst, and the mature miRNA sequences of 22 bases allows for facile reconstruction of the catalytic domains without any changes to the recognition domains. Therefore, the substrate gate and fuel strands were altered for a sequence-specific strand exchange cascade in response to miR-21 or miR-122 catalysts (Table 3.11 and Table 3.12). After a completed cycle, the substrate:fuel duplex becomes a waste product, the catalyst strand is recycled to initiate further catalytic cycles, and an output strand is released that interacts with a reporter gate to activate a fluorescent output signal (see Figure 3.33). It was proposed that the catalytic cycle could be used to recognize miRNAs as catalysts and activate the reporter gate for *in vivo* miRNA detection in mammalian cells with a highly sensitivity DNA computation modules.

The miR-21 and miR-122 catalytic cycles were first tested for activation of the reporter gate. The reporter gate is not activated by the substrate gates and fuel strands alone, but the addition of 20 nM miRNAs enhanced the fluorescent signal (Figure 3.37). Both the miR-21 and miR-122 cycles showed similar levels of activation with low background. Higher concentrations of the substrates and fuel strands produced an increase in reporter gate activation with the same amount of miRNA catalyst. Therefore, the reporter gate activation demonstrates the miRNA-

based catalytic cycles enabled amplified fluorescent detection of miR-21 or miR-122, dependent upon the substrate gate and fuel strand concentrations. Although activation of the catalytic cycle was demonstrated, these results (4-fold increase in fluorescence) are significantly lower than the amplification observed in the original fuel-catalyst cycle publication (over 100-fold increase in fluorescence),[206] suggesting further optimization of the component concentrations should be examined to increase the overall amplification of miRNA catalysts.
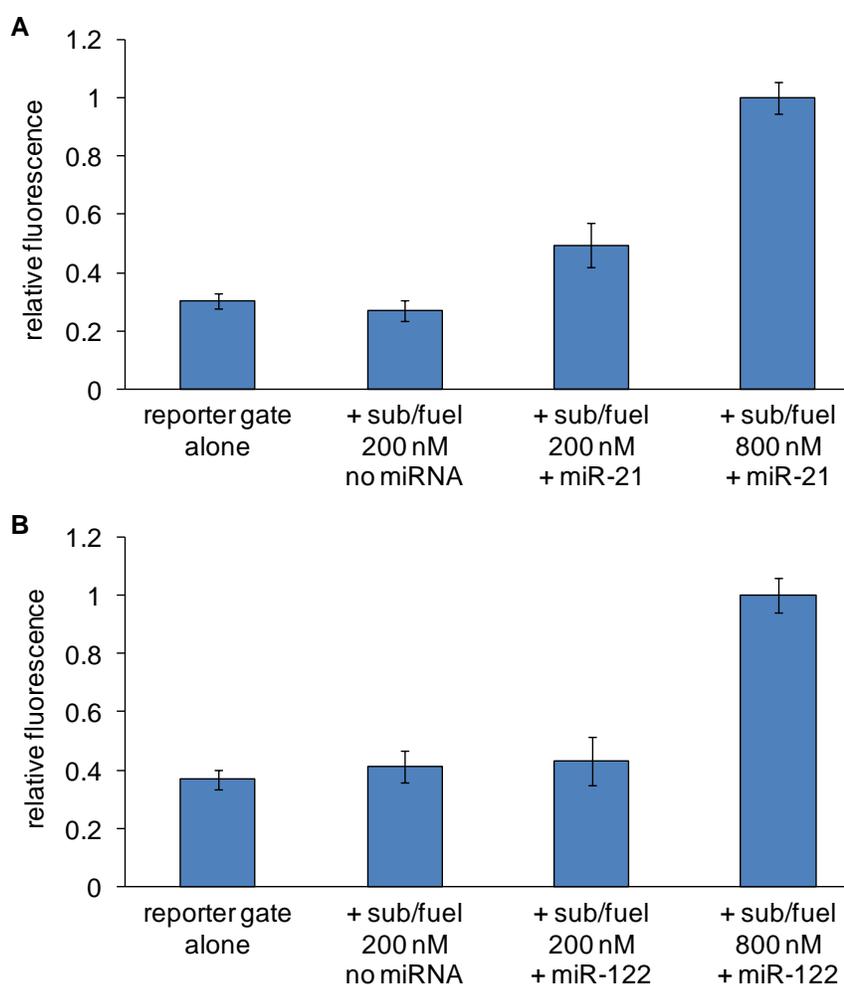
Figure 3.37: Validation of miRNA fuel-catalyst cycles.
A) The miRNA-based catalytic fuel cycle for miR-21. B) The miRNA-based catalytic fuel cycle for miR-122. The reporter gate (50 nM) was combined in solution with the substrate gates and fuel strands (200 or 800 nM), then the amplification cycles were initiated with miRNA catalysts (20 nM). TAMRA fluorescence was observed at 4 h (non-optimized) and normalized to the highest

activated condition. An average of three independent experiments is shown and error bars represent standard deviations.

Further analysis was then performed to determine the minimal concentration requirement for miRNA detection with the optimized catalytic cycle conditions, using the miR-21 catalytic cycle. For these experiments, the 800 nM concentration of substrate gate and fuel strand were used with a dilution series of miR-21, and subsequent signal amplification was determined with the reporter gate (Figure 3.38). While the 20 nM concentration of miR-21 signal matched the complete activation of the reporter gate with an output strand control at 200 nM, there are still significant increases in fluorescence at reduced concentrations down to 1 nM, which shows a 2-fold increase over the fuel-catalyst cycle in the absence of miR-21. The 5 nM and 10 nM concentrations of miR-21 show 3- and 5-fold increases, respectively. This data suggests that while the miRNA catalytic cycle can confidently detect 20 nM catalysts in solution, there is the potential to reduce the detection limit even further and get signal amplification of low abundance miRNAs.
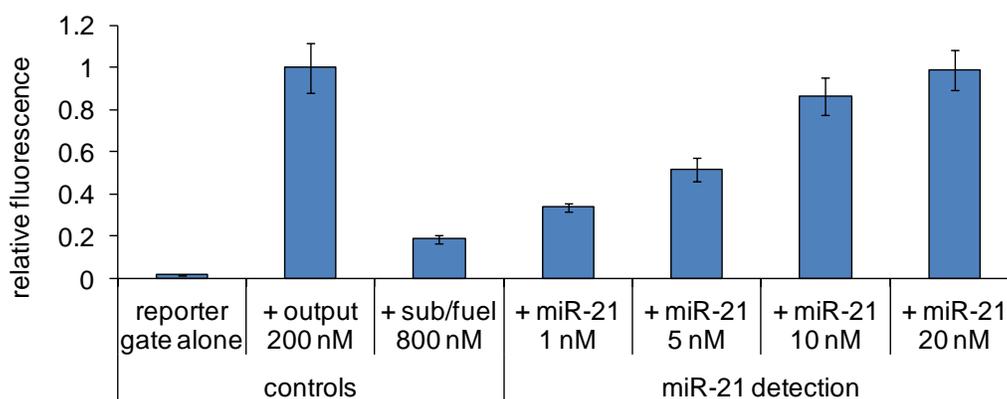
Figure 3.38: Analysis of miR-21 detection limit.
The miR-21 catalytic fuel cycle was performed with the reporter gate (50 nM) along with the corresponding substrate gate and fuel strand (800 nM). The miR-21 catalyst was tested from 1-20 nM, and compared to complete activation of the reporter gate with the output strand (200 nM). TAMRA fluorescence was observed at 4 h and normalized to the activated reporter gate. An average of three independent experiments is shown and error bars represent standard deviations.

174

In order to apply the miRNA-based catalytic cycles for the detection of endogenous miRNAs in live cells, the reporter gate duplex activation in the absence of the fuel-catalyst components was first analyzed *in vivo*. HEK293T cells were transfected with the reporter gate duplex in the absence and presence of the output strand and imaged after 4 h. The reporter gate showed no background fluorescence and is activated only in the presence of the output strand, validating its activity in a cellular environment (Figure 3.39).
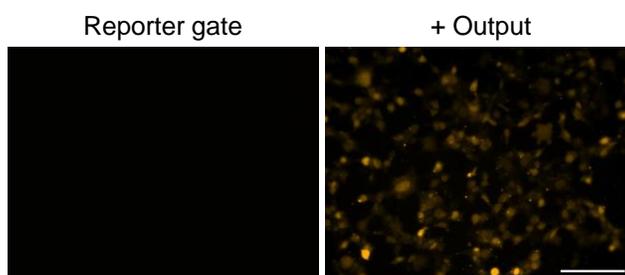
Figure 3.39: Fuel-catalytic cycle reporter gate analysis in HEK293T cells.
The reporter gate was transfected into HEK293T cells alone at 50 nM (left) and with 200 nM output stand (right). TAMRA fluorescence was observed at 4 h. Scale bar indicates 200 μm.

Co-transfection of the reporter gate with the catalytic cycles was then performed into different cell lines in order to determine the feasibility of detecting miRNAs *in vivo* with signal amplification. Initial experiments were performed with the miR-21 catalytic cycle in HEK293T (↓), HeLa (↑), and Huh7 (↑) cell lines. When all of the catalytic cycle components were transfected, very few cells remained for analysis, and no fluorescence was observed. Unfortunately, these experiments showed minimal success, as decreased cell viability and cell adhesion inhibited the ability to use the catalytic cycle *in vivo*. Further investigation of the singular components of the catalytic cycle did not indicate any trends for which gate or strand was causing the loss of cells (data not shown). Additionally, efforts to increase cell viability, such as replacing the gate elution buffers with water or PBS, and modifying the transfection

concentrations, were unsuccessful. Reporter gate activation with the miRNA-based catalytic cycles *in vivo* was not demonstrated due to these limitations.

In summary, catalytic cycles have been developed that recognize miR-21 or miR-122 and produces a fluorescent output signal. Substrate gates with miRNA-based toe-hold regions were used to initiate the catalytic cycles with low concentrations of synthetic miRNA. However, the concentrations of each component of the miRNA-based catalytic cycles need to be further optimized to enhance the amplification of miRNA catalysts. Since the goal of this project was not to perform amplification and detection of miRNA in solution, but rather in cellular environments, the transfection experiments were more critical to gauge the ability of the catalytic cycles for endogenous miRNA detection. To that end, the reporter gate activation with the output strand has been demonstrated in mammalian cells. Current research is focused towards activating the reporter gate with the miRNA-based catalytic cycles *in vivo*. For amplification of endogenous miRNA sensing, the *in vivo* link between the reporter gate and catalytic cycle must be established. The many strand exchanges have been difficult to optimize within a cellular environment, since this system is more complex than the reported AND gates, which involve 2-4 strand dissociations compared to 5 dissociations required for the reporter gate activation with the catalytic cycle. Similar reports for detection of RNAs in live cells with signal amplification have been demonstrated through hybridization cascades using FRET pair assembly[231] and hairpin amplifiers,[218] suggesting that different designs may be optimal for signal amplification of nucleic acid signatures in live cells. The use of DNA-based catalytic cycles and signal amplification system will presumably enable a lower detection limit for the endogenous sensing of miRNAs than the AND gates demonstrated in Sections 3.2 and 3.3. One benefit of this system is that the reporter gate has an interchangeable design, in which the substrate gates can be altered to

respond to a new catalyst, and the catalytic cycle can be developed to recognize alternate miRNAs or other endogenous oligonucleotides of biological interest. Once the *in vivo* link between a miRNA-based catalytic cycle and the output reporter gate has been successfully completed, the next step will be light-activation of the catalytic cycle with a caged fuel strand to enable spatial and temporal sensing of low abundance miRNAs in live cells.

3.5.1   Methods and Materials

Fuel-catalyst Duplex Purification. See Section 5.6 for specific information on DNA computation techniques. Duplexes were purified following protocol 5.6.1. All oligonucleotide sequences are shown in the following tables; a letter prior to the base indicates sugar modification (r = RNA), miRNA recognition sequences are underlined, Q = BHQ2 quencher, F = Tetramethylrhodamine (TAMRA) fluorophore:

Table 3.11: miR-21 fuel-catalyst cycle.

Strand	Sequence (5' → 3')
Catalyst miR-21	rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrG
Substrate$_{21}$	<u>TCAACATCAGTCTGATAAGCTA</u>AGGGCCGTAAGTTAGTTGGAGACGTAGG
Fuel$_{21}$	CCTACGTCTCCAACTAACTTACGGCCCT<u>TAGCTTATCAGACTGA</u>
Signal strand$_{21}$	CCACATACATCATATTCCCT<u>TAGCTTATCAGACTGA</u>
Output strand	CTTTCCTACACCTACGTCTCCAACTAACTTACG
Reporter gate F	F-CTTTCCTACACCTACG
Reporter gate Q	TGGAGACGTAGGTGTAGGAAAG-Q

Table 3.12: miR-122 fuel-catalyst cycle.

Strand	Sequence (5' → 3')
Catalyst miR-122	rUrGrGrArGrUrGrUrGrArCrArArUrGrGrUrGrUrUrUrG
Substrate$_{122}$	<u>CAAACACCATTGTCACACTCCA</u>AGGGCCGTAAGTTAGTTGGAGACGTAGG
Fuel$_{122}$	CCTACGTCTCCAACTAACTTACGGCCCT<u>TGGAGTGTGACAATGG</u>
Signal strand$_{122}$	CCACATACATCATATTCCCT<u>TGGAGTGTGACAATGG</u>
Output strand	CTTTCCTACACCTACGTCTCCAACTAACTTACG
Reporter gate F	F-CTTTCCTACACCTACG
Reporter gate Q	TGGAGACGTAGGTGTAGGAAAG-Q

Fuel-catalyst Cycle Examination. See Section 5.6 for specific information on DNA computation techniques. Fluorescence was measured following protocol 5.6.2.

Fuel-catalyst Cellular Transfections. See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. Cells were passaged into a 96-well plate and grown to ~70% confluence within 24 h following protocol 5.4.1. The cells were transfected with the fuel-catalyst cycle using X-tremeGENE following protocol 5.6.3.

Live Cell Imaging of TAMRA Fluorescence. See Section 5.4 for specific information on cell culture techniques and Section 5.6 for specific information on DNA computation techniques. Transfected cells were imaged on a Zeiss Observer Z1 microscope (20X magnification) following protocol 5.4.5.

178

4.0   Application of Caged Proteins in the Regulation of Gene Function

As discussed in Section 1.2, proteins can be caged with light-responsive amino acids through UAA mutagenesis. This chapter will discuss the application of NPOC-caged lysine (Figure 1.22, further abbreviated as PCK) for the optical control of RNA polymerization and gene editing in mammalian cells. Additionally, a system for the incorporation of unnatural amino acids in live animals was developed in efforts to produce the first vertebrate species with an expanded genetic code.

4.1   Genetically Encoded Light-Activated Transcription for Spatiotemporal Control of Gene Expression and Gene Silencing in Mammalian Cells

This material was reprinted, in part, with permission from **Hemphill, J.**; Chou, C.; Chin, J.W.; Deiters, A. *J. Am. Chem. Soc.* **2013**, 135(36), 13433-13439.

A caged T7 RNA polymerase was expressed in cells with an expanded genetic code and used in the photochemical activation of genes under control of an orthogonal T7 promoter, demonstrating tight spatial and temporal control. The synthetic gene expression system was validated with two reporter genes (luciferase and EGFP) and applied to the light-triggered transcription of short hairpin RNA constructs for the induction of RNA interference.

The addition of nonmammalian RNA polymerases to the genetic circuitry of cells enables the development of orthogonal genetic expression platforms that can be manipulated to activate specific genes that will not be expressed by the endogenous cellular genetic machinery. Caged polymerases are versatile building blocks for the construction of synthetic gene networks to control gene expression with light as an input.[256] Protein expression under photochemical control can be used to build synthetic biological circuits that respond to light with specific gene function outputs, thus conveying precise spatial and temporal control over these circuits. A genetically encoded light-activated T7 RNA polymerase was developed that enables the promoter-specific photochemical regulation of transcription for both coding and noncoding RNAs in a mammalian cell system.

T7 RNA polymerase (T7RNAp) is a bacteriophage RNA polymerase that is related to the PolI family of DNA polymerases,[257] which is especially useful for exogenous gene control in the reprogramming of genetic and network relationships, as the polymerase and its corresponding promoter are completely orthogonal to all endogenous polymerases in mammalian cells.[258] Therefore, genes of interest can be conditionally expressed through the light activation of a T7RNAp temporarily rendered inactive through the incorporation of a caged amino acid in its active site. Importantly, photochemically activated transcription can be used to not only control gene expression but also to express noncoding RNA sequences of other biological function. A light-activated T7RNAp has previously been generated in through the site-specific incorporation of an *o*-nitrobenzyl caged tyrosine in *E. coli* cells with an expanded genetic code.[127] However, its application for the photochemical control of transcription in mammalian cells was hampered since the *M. janaschii* tRNA synthetase/tRNA pair is not orthogonal in mammalian cells and transfection of the caged protein was required. To overcome this limitation, site-specific

incorporation of a genetically encoded PCK into T7RNAp through an engineered, fully orthogonal pyrrolysine synthetase/tRNA pair was used to photochemically control RNA polymerization in mammalian cells with (Figure 4.1). The photocaged lysine exhibits increased solubility in mammalian growth media, more rapid decaging due to better leaving group qualities of the released carbamate, a bathochromic shift in absorbance maximum and thus decaging wavelength, and produces a less reactive byproduct upon photolysis than the previously employed caged tyrosine.[114]
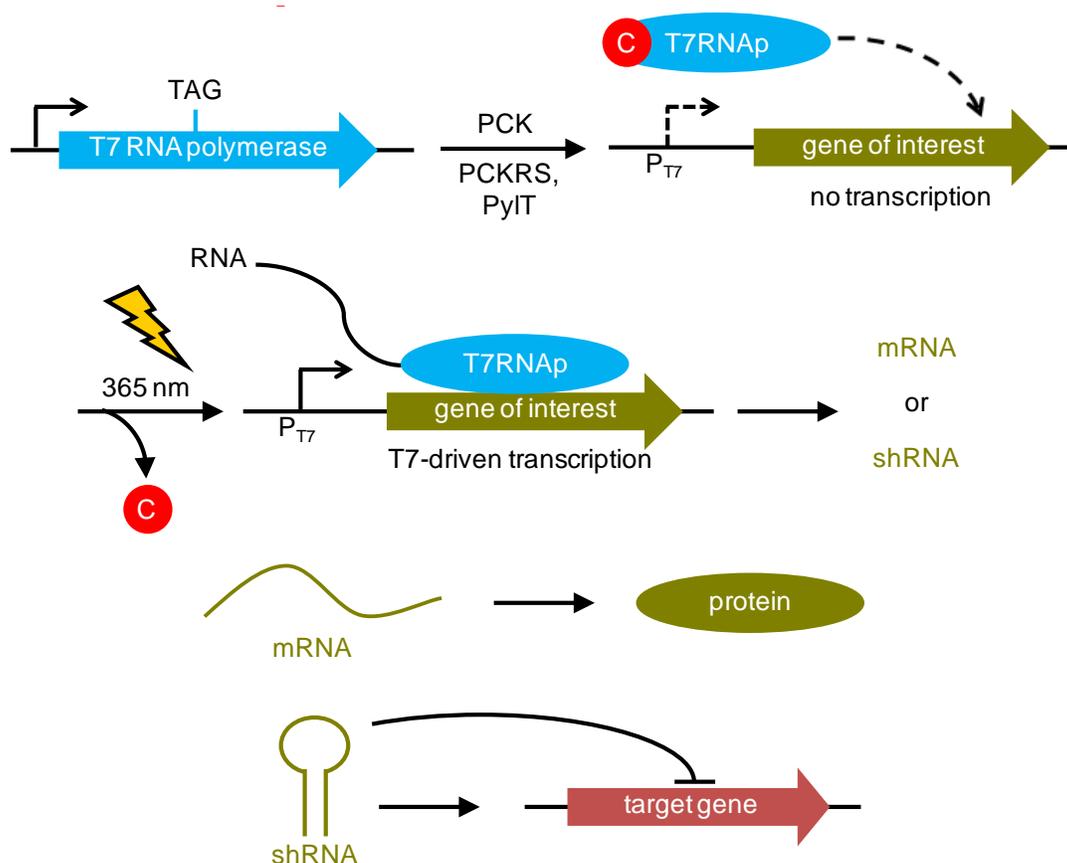
Figure 4.1: Schematic for expression and function of caged T7RNAp.
Light-activated T7RNAp is expressed in mammalian cells through site-specific incorporation of PCK (red circle) into its active site in response to an amber stop codon (TAG) via an engineered tRNA synthetase (PCKRS) that misacylates an amber-suppressor tRNA (PylT) with PCK. The caged T7RNAp is completely inactive until irradiation with 365 nm UV light induces decaging and activation of T7-driven transcription. Depending on the function of the transcribed RNA,

light-induced protein expression from mRNA or light-induced gene silencing via RNA interference from shRNA is achieved.

Transcription catalyzed by T7RNAp involves a number of protein conformational changes, and the active site of T7RNAp, which is conserved among a number of polymerases found in nature, has been studied in detail.[259] These structural and kinetic studies indicate that the active site lysine 631 (K631) is critical for T7RNAp function, as it interacts with phosphate groups of the incoming nucleotide triphosphates (NTPs) at the interface of the DNA template and RNA product strands, stabilizing the NTPs through hydrogen bonding with the 1α-phosphate residue.[260, 261] Mutations at the K631 position have been shown to inhibit transcription.[262] Thus, it was hypothesized that a sterically demanding caging group installed on K631 will block incoming NTPs from the active site and inhibit T7RNAp activity. In addition, the carbamate linker changes the electronic nature of the ε-amino group, effectively preventing protonation and removing the positive charge. The caged T7RNAp is expected to be completely inactive until a brief UV irradiation at 365 nm removes the caging group from K631 and enables transcription (Figure 4.2).

caged, **inactive** T7RNAp   decaged, **active** T7RNAp
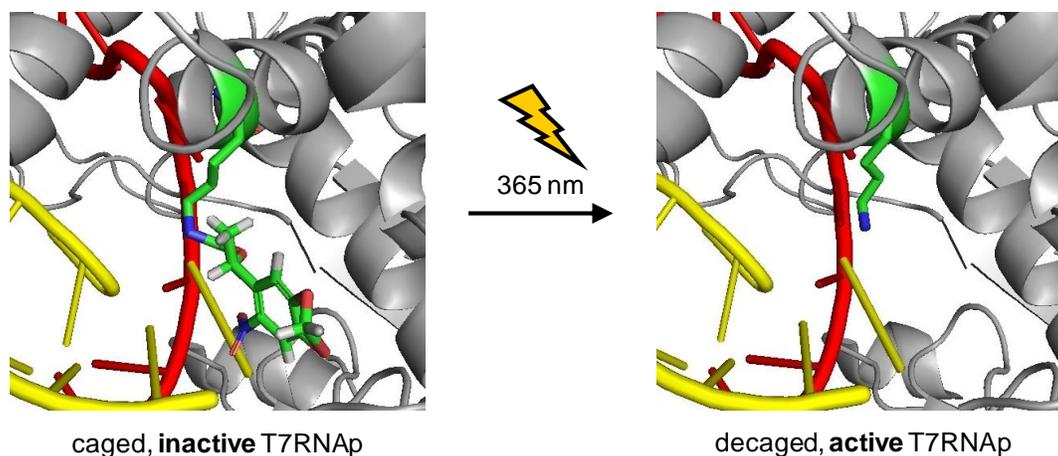
Figure 4.2: Structural representation of caged T7RNAp active site residue. K631-caged (left) and wild-type (right) T7RNAp are depicted. The K631 residue is indicated in green, and the DNA template and RNA transcript are indicated in red and yellow, respectively. Images were generated in PyMol from PDB 1S76.[261]

An amber stop codon TAG was introduced at K631 of T7RNAp (Table 4.1), and the gene was cloned with a 6-His tag into a plasmid (pAG31) containing the MbPylRS synthetase mutant for PCK incorporation (PCKRS: M241F, A267S, Y271C, and L274M), derived from the *M. barkeri* pyrrolysine tRNA synthetase.[114]   Previous studies for the incorporation of UAAs have shown that there are no off-target effects on endogenous gene expression through the repurposing of the low-frequency TAG stop codon in mammalian cells.[108, 263] As a positive control, wild-type (WT) T7RNAp was also cloned into the same vector. For the expression of caged T7RNAp, cotransfections were performed with a plasmid (pAG38) containing four copies of the pyrrolysine tRNA$_{CUA}$ (PylT). Western blot analysis of HEK293T cells that were transfected with the T7RNAp expression constructs and incubated for 24 h in the absence or presence of PCK revealed the expression of hexahistidine-tagged T7RNAp only in the presence of PCK, indicative of the high fidelity of the engineered synthetase (Figure 4.3).
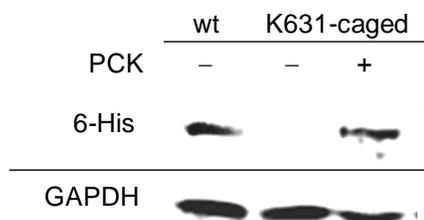
Figure 4.3: Expression of caged T7RNAp.
HEK293T cells were transfected with pAG31:T7RNAp WT, or pAG31:T7RNApK631TAG and pAG38 expression plasmids then incubated in the presence of PCK (2 mM) for 24 h. Western blot analysis was performed using 6-His and GAPDH primary antibodies with a FITC-labeled secondary antibody for detection.

The activity and photochemical control of K631-caged T7RNAp was investigated using a firefly luciferase reporter gene under control of the T7 promoter.[127] First, the effects of UV exposure on the luciferase assay were studied, revealing that an irradiation times up to 20 min had only minor detrimental effects on T7-driven luciferase activity in cells expressing the wild-type polymerase, demonstrating a low UV sensitivity for the luciferase assay (Figure 4.4A). Cells were then transfected with the T7-driven luciferase reporter and the K631-caged T7RNAp expression plasmids, incubated in PCK for 24 h, and irradiated with 365 nm light at increasing intervals up to 20 min (Figure 4.4B). Light-activated luciferase expression was observed in cells expressing the caged T7RNAp, with a tunable linear response of gene expression to an increasing irradiation time. A 3 min irradiation exhibited a 9-fold increase compared to the nonirradiated control, and luciferase activity did not significantly increase with additional UV exposure of up to 20 min. Importantly, the presence of the caging group at K631 fully blocked transcriptional activity of the enzyme, since the absence of polymerase (−T7), caged T7RNAp (−PCK), or light exposure (0 min UV) all showed identical, very low levels of background luminescence.
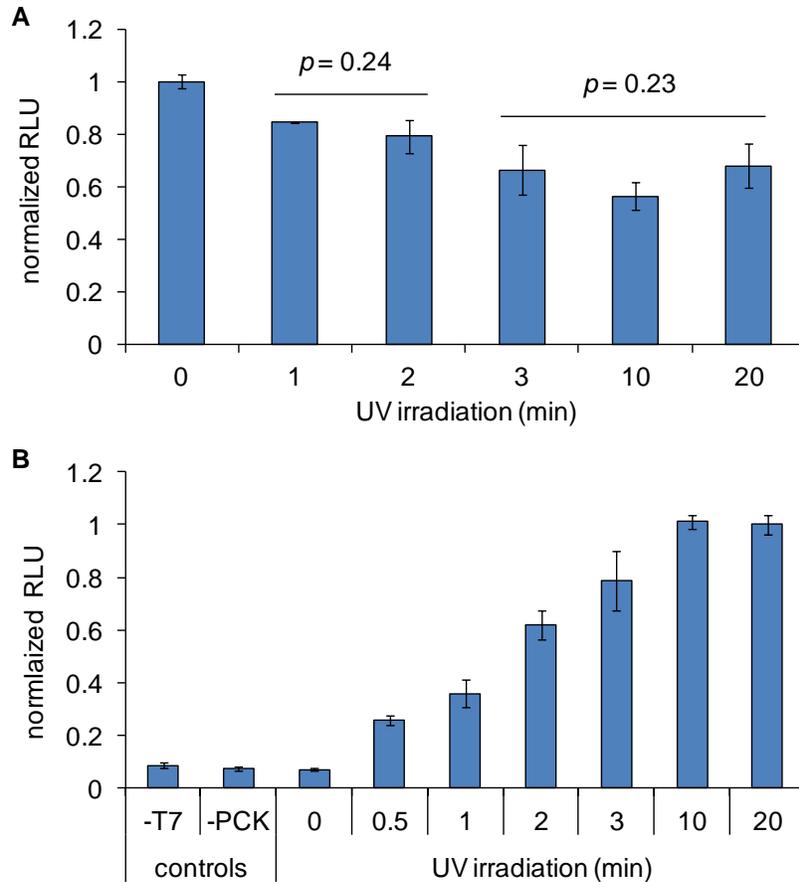
Figure 4.4: Luciferase reporter UV irradiation and photochemical activation.
A) HEK293T cells were transfected with the pAG31:T7RNAp WT and T7-IRES-Luc plasmids. Luminescence units were normalized to a no UV control. An analysis of variance (ANOVA) was conducted. Luciferase activities after a 1 and 2 min irradiation were statistically identical (p = 0.24). A similar statistical insignificance in variance was observed between 3, 10, and 20 min irradiations (p = 0.23). B) HEK293T cells were transfected with the pAG31:T7RNApK631TAG, pAG38, and T7-IRES-Luc plasmids then incubated with 2 mM PCK overnight. Luminescence units were normalized to the activation observed after a 20 min irradiation. UV irradiations were performed on a 365 nm transilluminator, and luciferase bright glow assays were performed 24 h after UV exposure. Error bars represent standard deviations of experimental triplicates.

A T7-IRES-EGFP reporter plasmid was also constructed for fluorescent imaging of the photochemical gene expression system. The T7-driven EGFP reporter construct was validated in cells expressing the wild-type polymerase and transfection conditions were optimized for enhanced EGFP expression (Figure 4.5). The 48 h expression transfected with 2-fold increase of lipofectamine reagent showed the highest EGFP levels.
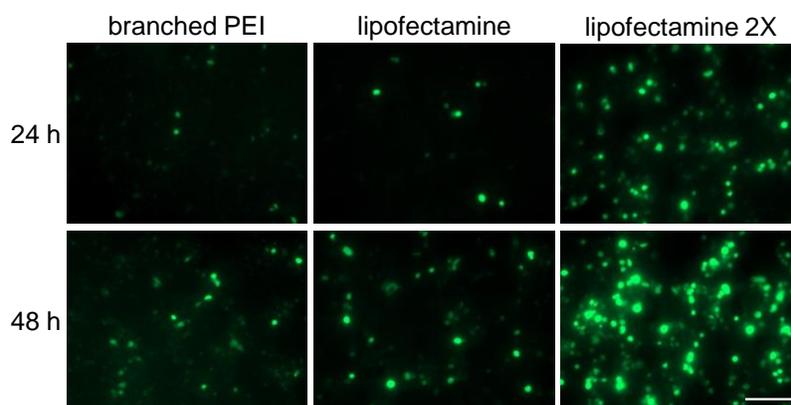
Figure 4.5: Wild-type T7RNAp expression of an EGFP reporter.
HEK293T were transfected with the pAG31:T7RNAp WT and T7-IRES-EGFP plasmids under various conditions. Fluorescence was observed at 24 and 48 h. Scale bar indicates 200 μm.

The photochemical activation of caged T7RNAp for the expression of T7-driven EGFP was then confirmed through cotransfection with the PCKRS/PylT expression machinery, and conditions were optimized for optimal OFF → ON switching of EGFP expression (Figure 4.6A). The 48 h incubation increased total EGFP expression, and no enhancement was observed through longer UV irradiation times. No background EGFP fluorescence was observed in cells containing the caged T7RNAp construct in the absence of PCK or UV exposure (Figure 4.6B). Photochemical control over EGFP expression was only achieved through treatment of cells with both PCK and UV irradiation. Fluorescent cell counting was performed to quantify the photochemical activation of T7RNAp-dependent expression of EGFP (Figure 4.6C). A 5-fold increase in the number of fluorescent cells was observed after both 2 and 10 min UV irradiations, while the absence of either PCK or UV exposure showed only basal levels of fluorescence, indicating that the caged polymerase is completely inactive and displays no background expression.

186
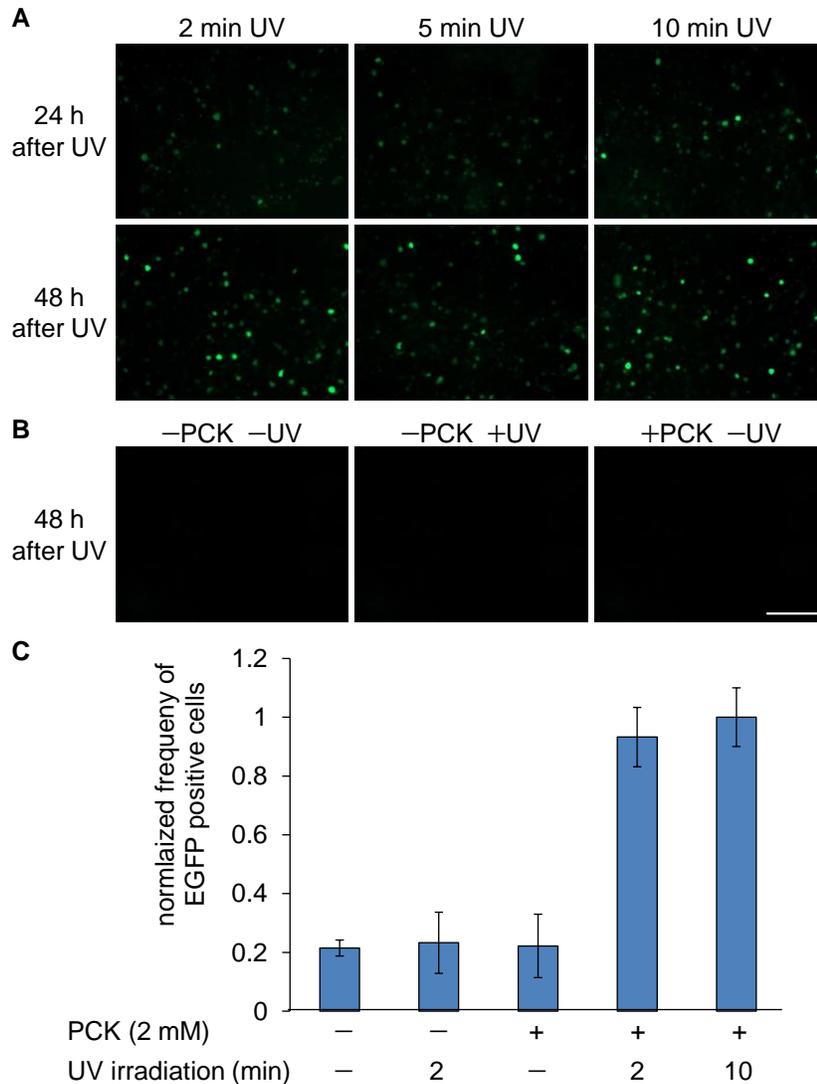
Figure 4.6: Light activation of EGFP transcription in HEK293T cells.
A) UV irradiation time course for the activation of K631-caged T7RNAp. Cells were transfected with the pAG31:T7RNApK631TAG, pAG38, and T7-IRES-EGFP plasmids then incubated with 2 mM PCK overnight. UV irradiations were performed on a 365 nm transilluminator for 2, 5, and 10 min. The expression of EGFP was observed at 24 and 48 h time points after UV irradiation. B) K631-caged T7RNAp expression of EGFP in the absence of PCK or UV irradiation at 48 h. Scale bar indicates 200 μm. C) Fluorescent cell counting was performed 48 h after UV exposure. The frequencies of EGFP-positive cells (gated/total) were normalized to the 10 min UV irradiation, and error bars represent the standard deviation obtained from three experimental replicates.

In order to demonstrate spatial control of RNA expression in mammalian tissue culture, the photochemical activation of the T7RNAp-driven EGFP reporter was performed as in Figure 4.6A; however, spatially restricted UV irradiations were conducted through specifically shaped

187

masks (pinhole and horizontal slit). Only areas exposed to the patterned UV light were EGFP-positive due to localized activation of RNA polymerization (Figure 4.7). These experiments successfully demonstrated spatial and temporal control of gene function using a site-specifically caged T7RNAp encoded in mammalian cells.
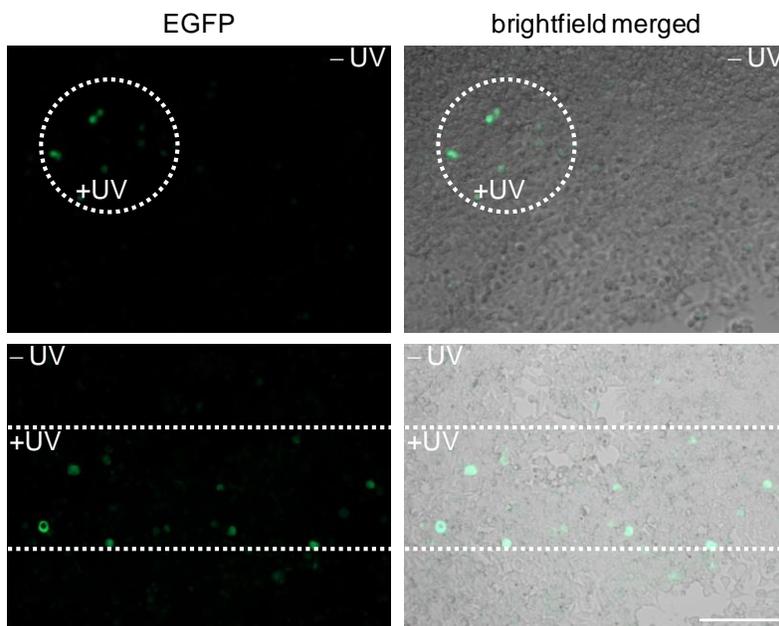
Figure 4.7: Spatial control of photochemically activated gene expression.
Cells were cotransfected with the K631-caged T7RNAp polymerase expression system and EGFP reporter as previously described, then incubated with 2 mM PCK for 24 h. Masks were used to irradiate small populations of cells at 365 nm for 2 min, and EGFP expression was observed at 48 h. Irradiations were performed through a small pinhole (top) and a horizontal slit (bottom). Scale bar indicates 200 μm.

To further demonstrate the applicability of the light-activated T7RNAp/promoter system, it was subsequently applied to the photochemical activation of RNA interference through T7-driven expression of short hairpin RNAs (shRNAs).[264] These shRNAs are processed into siRNAs exhibiting prolonged activity for gene knockdown.[22] The shRNAs can be designed to exhibit promoter-specific expression,[265] and can be genetically integrated to generate stable RNAi regulators.[266] The use of plasmid-based shRNA for gene suppression has been successfully

applied in a variety of mammalian cell systems,[267] and T7RNAp-driven shRNAs have been reported for the regulation of gene expression in live mammalian cells[268] as well as zebrafish embryos.[269] Here, the developed photochemical gene regulation system was coupled with an Eg5 shRNA component as a proof of principle study in the photochemical regulation of RNA interference. The Eg5 gene encodes a motor protein involved in mitosis,[270] and inhibition of Eg5 with siRNA oligonucleotides has been shown to produce binucleated cells,[271, 272] providing a distinct phenotypic readout for RNAi activity. The Eg5 shRNA expression cassette was cloned into the T7-IRES-EGFP reporter plasmid with built-in T7RNAp expression components including the T7 RNA promoter and terminator sequences (Table 4.2). This construct was designed to fluorescently track cells that were successfully transformed, as well as exhibit expression and photochemical activation of caged T7RNAp.

First, shRNA expression was tested for activity in the presence of a WT T7RNAp gene in HeLa cells. The resulting binucleated cellular phenotype was observed, in agreement with the phenotype observed through transfection of an siRNA oligonucleotide positive control, as expected (Figure 4.8). A binucleated phenotype was observed with both the Eg5 siRNA and plasmid-based WT T7-expressed shRNA controls, while EGFP expression successfully tracked cells for the identification of T7RNAp-driven shRNA activity. The K631-caged T7RNAp expression system was then used to achieve UV activation of Eg5 shRNA expression through cotransfection of the required plasmids, followed by incubation in the presence of PCK and UV irradiation (Figure 4.8). The EGFP expression and cellular binucleation were fully dependent upon UV irradiation and activation of the caged T7RNAp. The shRNA construct has been validated to enable light-activated RNA interference.
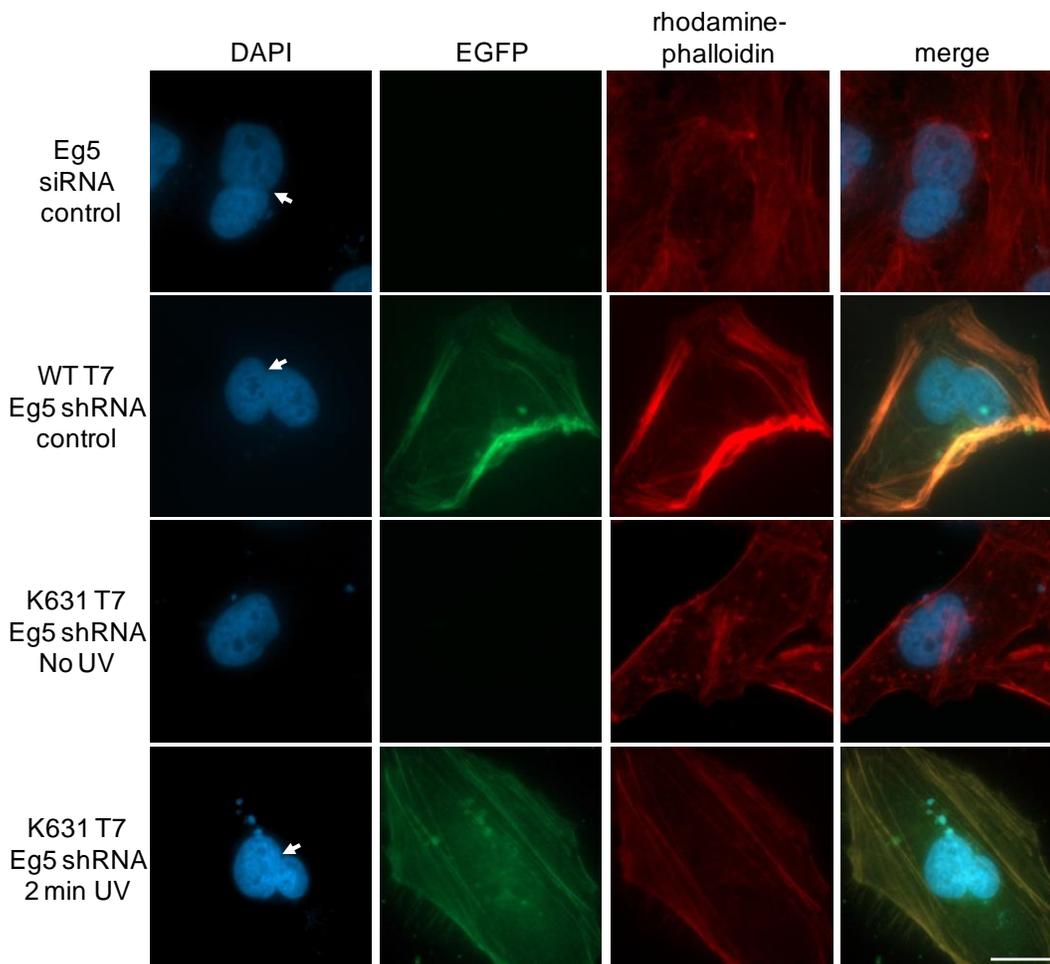
Figure 4.8: Light activation of Eg5 shRNA using the caged T7RNAp system.
HeLa ells were transfected with an Eg5 siRNA (50 pmol), as well as WT T7RNAp or the K631-caged T7RNAp expression system with the T7-IRES-EGFP-Eg5 shRNA construct. The caged T7RNAp expressions were incubated in 2 mM PCK for 24 hr, and irradiated for 2 min at 365 nm. At 48 h post treatment, cells were fixing and staining with DAPI (nucleus) and rhodamine-phalloidin (actin) prior to imaging. Scale bar indicates 20 μm, and binucleated cells are labeled with a white arrow.

Light activation of shRNA expression was verified through the observation of a binucleated cell phenotype only after UV irradiation, which was then quantified through cell counting (Figure 4.9). The relative frequency of binucleated cells observed indicated that the photochemically regulated expression of Eg5 shRNA with K631-caged T7RNAp is activated to nearly identical levels as wild-type T7RNAp. The frequency of binucleated cells observed for the Eg5 siRNA was also similar to the frequency observed for the T7RNAp plasmid-driven Eg5

190

shRNA. In the absence of UV irradiation the relative frequency of cellular binucleation was

similar to nontreated cells. The photochemical activation of K631‑caged T7RNAp produced a

similar binucleation frequency as the Eg5 siRNA and WT T7RNAp.
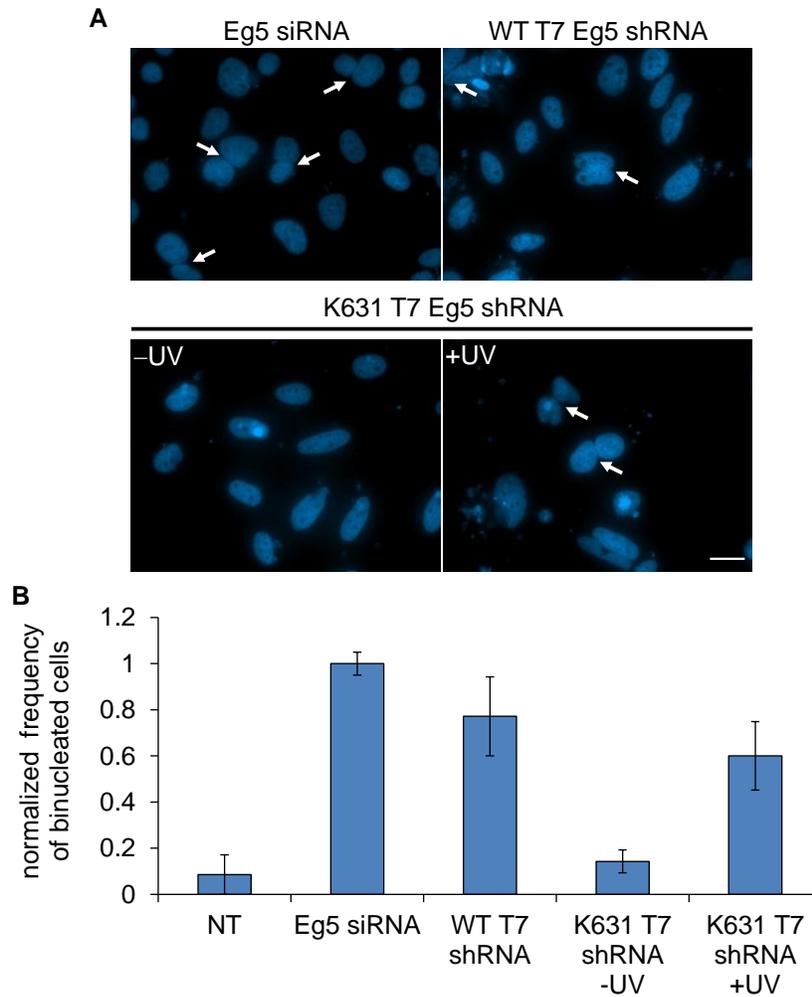
Figure 4.9: Relative frequency of binucleated phenotype in HeLa cells.
Transfections were performed with an Eg5 siRNA control and the T7RNAp plasmid-driven Eg5 shRNA as previously described. A) The cells were fixed 48 h post treatment and stained with DAPI (nucleus) prior to imaging. Scale bar indicates 20 μm. Binucleated cells are labeled with a white arrow. B) Counts of 50 cells were performed for each condition to determine the relative frequency of binucleated cells, normalized to the Eg5 siRNA positive control. Error bars represent standard deviations from triplicates.

Additionally, qRT-PCR was performed on cellular RNA extracts to quantify the relative Eg5 mRNA expression levels and to further validate the photochemical gene silencing (Figure 4.10). A 55% reduction of Eg5 mRNA was observed in the case of both wild-type T7RNAp-driven shRNA expression and light-activated T7RNAp, in accordance with the Eg5 siRNA reagent as a positive control. In the absence of UV irradiation, the caged T7RNAp expression system showed similar levels of Eg5 mRNA as nontreated cells. These findings demonstrate that the caged T7RNAp-controlled genetic circuit can be applied to the photochemical regulation of gene expression and the light induction of noncoding RNAs to achieve external control over RNA interference.
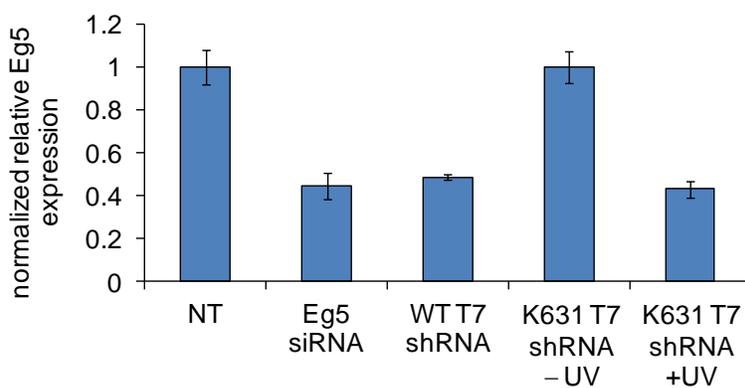
Figure 4.10: Quantification of Eg5 mRNA in HEK293T cells.
Transfections were performed with an Eg5 siRNA control and compared to T7RNAp plasmid-driven Eg5 shRNA as previously described. Total RNA was isolated from cells 48 h after the corresponding treatment and qRT-PCR was performed with Eg5 specific primers. The Eg5 mRNA was normalized to GADPH mRNA as a control and set relative to nontreated cells (NT), with error bars representing standard deviations from triplicates.

In summary, a light-activated T7RNAp gene expression system was successfully developed in mammalian cells engineered with additional protein biosynthetic machinery for genetically encoded protein modification with a caged lysine amino acid (PCK). Site-specific incorporation of PCK at the K631 active site residue in T7RNAp led to full inactivation of RNA

polymerization. Light irradiation at 365 nm removed the caging group from the active site and enabled the expression of transcripts that were placed under control of the T7 promoter. This was successfully demonstrated for two reporter genes, luciferase and EGFP. Temporal control of UV exposure allowed for finely tuned gene expression that gradually increased with longer durations of irradiation. In addition, precise spatial control of T7RNAp activity through localized UV irradiation was achieved, as shown by the patterned expression of an EGFP reporter in a monolayer of mammalian cells. Furthermore, photochemical regulation of T7RNAp activity was used to achieve precise control over the expression of noncoding RNAs, as exemplified by the light induction of shRNA transcription and thus RNA interference. Here, a unique combination of individual parts, including an orthogonal tRNA/tRNA synthetase pair, a photocaged lysine amino acid, a caged T7RNAp, and T7 promoter-driven shRNA expression cassette, were interfaced with the RNA interference pathway to construct the circuitry for light-triggered, sequence-specific gene silencing. The ability to photochemically control the expression of shRNAs allows for spatial and temporal dissection of the RNAi pathway. Overall, a genetically encoded expression platform has been developed that can photochemically regulate specific genes under control of the T7 promoter in mammalian cells. In theory, any gene under control of a T7 promoter can now be photochemically regulated in live cells. The engineered spatial and temporal control can be applied to the precise triggering of genetic networks allowing for a wide range of cellular applications.

4.1.1   Methods and Materials

Plasmid Construction. See Section 5.3 for specific information on molecular biology techniques. An amber stop was introduced into T7 RNA polymerase through mutagenesis of the

pBH161 plasmid.[273] A single nucleotide was changed for the K631 AAG → TAG mutation using primers shown in Table 4.1 following protocol 5.3.2. The T7RNAp gene was then PCR amplified from the pBH161 plasmid with primers to introduce *NheI* and *Eco*RI restriction sites following protocol 5.3.1. The ~2.7 kB T7RNAp gene was cloned between the *NheI* and *MfeI* restriction sites of the pAG31 plasmid[114] following protocols 5.3.3 and 5.3.4. The Eg5 shRNA insert was prepared with the T7 RNA promoter, Eg5 shRNA sense and antisense strands connected with a loop sequence, then followed by a T7 RNA terminator region (Table 4.2). These elements were purchased as single-stranded DNA oligos (IDT) and ligated into the T7-IRES-EGFP plasmid at *BglII and SacI* restriction sites following protocols 5.3.3 and 5.3.4. Plasmid maps are shown in Figure 4.11.

Table 4.1: Primer sequences used in the development of K631TAG T7RNAp.
The T7RNAp gene PCR products were into pAG31. Mutations introduced are underlined and restriction sites are bolded.

Primer	Sequence (5' → 3')
K631TAG Forward	GTTACTCGCAGTGTGACTT̲AGCGTTCAGTCATGACGC
K631TAG Reverse	GCGTCATGACTGAACGCTA̲AGTCACACTGCGAGTAAC
Forward *Nhe1*	TTAA**GCTAGC**ACCATGGGCAGCAGCCATC
Reverse *EcoR1*	CGGT**GAATTC**TTACGCGAACGCGAAGTC

Table 4.2: Sequences of the Eg5 shRNA construct and the siRNA oligonucleotides.
Restriction enzyme sites indicated in bold, T7-promoter and terminator underlined, and ribonucleic acid bases indicated with "r".

Strand	Sequence (5' → 3')
Eg5 shRNA insert	**GATC**<u>TAATACGACTCACTATAGGGAGA</u>ATAGACTTCATCCTTGTTGT TCAGAGCTAACAACAAGGATGAAGTCTAT<u>ATCTGTTTTTTT</u>**AGCT**
Eg5 shRNA insert	AAAAAAACAGATATAGACTTCATCCTTGTTGTTAGCTCTGAACAAC AAGGATGAAGTCTATTCTCCCTATAGTGAGTCGTATTA
siRNA sense	rCrArArCrArArGrGrArUrGrArArGrUrCrUrArUTT
siRNA antisense	rArUrArGrArCrUrUrCrArUrCrCrUrUrGrUrUrGTT
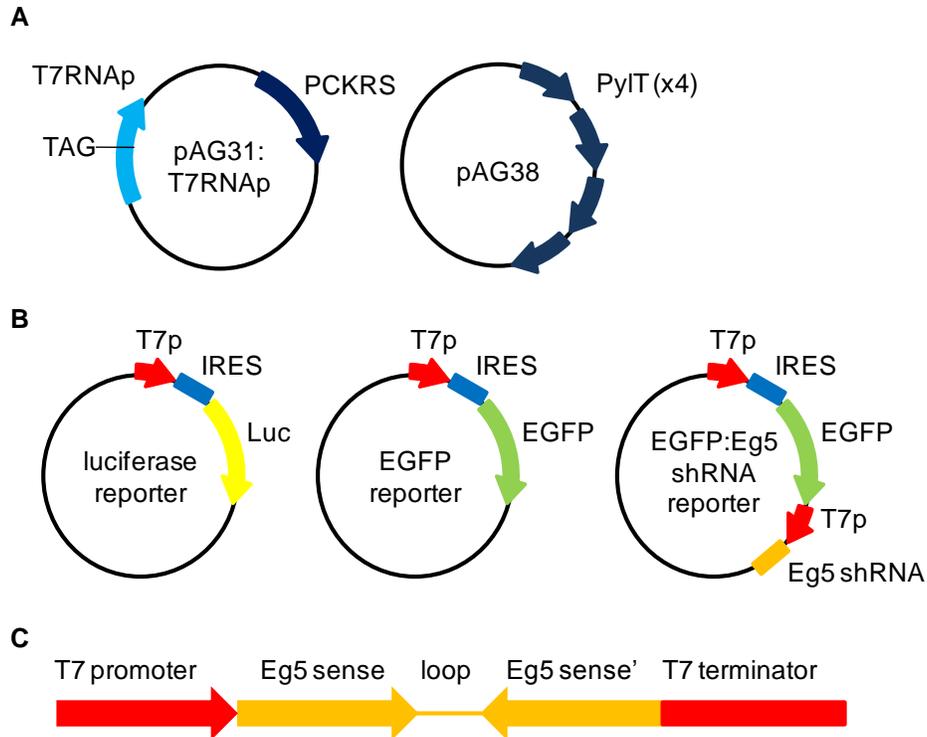
Figure 4.11: Plasmid constructs for T7RNAp experiments.
A) Plasmids used for the expression of caged T7RNAp. B) Plasmids used for the analysis of activity with specific reporter genes. C) The Eg5 shRNA construct segments.

Western Blot of Caged T7RNAp. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 6-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 2 μg of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. Protein isolations were performed following protocol 5.4.8. Western blots were performed following protocol 5.4.9 using a mouse anti-6-His primary antibody and a mouse anti-GADPH control (Santa Cruz Biotechnology). The primary antibodies were detected with a goat anti-mouse-FITC fluorescent secondary antibody.

Photochemical Regulation of Reporter Genes. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 96-well plates and grown to

195

~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 200 ng of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. The PCK-containing media was removed after the overnight incubation followed by exposure at 365 nm using a UV transilluminator (25 W) following protocol 5.4.3. For the luciferase reporter, a Bright-Glo assay (Promega) was performed 24 h post UV irradiation to quantify luciferase activity by removing 150 μL of cell media, adding 50 μL of the kit reagent, and shaking for 20 min. Chemiluminescence was measured on a BioTek Synergy 4 plate reader. For the EGFP reporter, fluorescent imaging was performed using a Zeiss Observer Z1 microscope (20X magnification) following protocol 5.4.5. Spatially distinct UV irradiations were performed through precut vertical slits in tinfoil, and irradiated at 365 nm with a transilluminator (25 W) for 2 min following protocol 5.4.3.

Fluorescent Cell Sorting. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 96-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 200 ng of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. The PCK-containing media was removed after the overnight incubation followed by exposure at 365 nm using a UV transilluminator (25 W) following protocol 5.4.3. Following a 48 h incubation cells were trypsinized, pooled, washed, and resuspended in PBS. Flow cytometry was performed on a FACSCalibur (Becton-Dickinson) instrument (488 nm argon laser, 530/50 nm BPF). Cells were gated for EGFP fluorescence (above $10^{2.5}$ RFUs) then analyzed with Cellquest Pro Software until 20,000 cells were counted for each condition tested. The frequency of EGFP positive cells (gated/total) was averaged from three replicates and normalized to the UV-activated EGFP expression.

Eg5 shRNA Expression and Imaging. See Section 5.4 for specific information on cell culture techniques. HeLa cells were passaged into 4-well chamber slides following protocol 5.4.1, and transfected with 1 µg of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The Eg5 siRNA duplex was prepared by annealing the oligonucleotides (Table 4.2)[274] at 10 µM in TAE/Mg^{2+} buffer (0.04 M tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), 12.5 mM magnesium acetate, pH ~7.4) from 65 to 12 °C over 10 min, and transfected at 50 pmol as a positive control with X-tremeGENE siRNA for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. The PCK-containing media was removed after the overnight incubation followed by exposure at 365 nm for 2 min using a UV transilluminator (25 W) following protocol 5.4.3. After 48 h incubation cells were fixed with 3.75% formaldehyde, followed by DAPI (Invitrogen) and rhodamine-phalloidin (Invitrogen) counter stains following protocol 5.4.4. Imaging was performed with a Zeiss Observer Z1 microscope (63X magnification) following protocol 5.4.5. Relative frequencies of binucleated cells were determined through three counts of 50 cells for each condition and were normalized to the Eg5 siRNA positive control.

Quantification of Eg5 mRNA by qRT-PCR. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 6-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 2 µg of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The Eg5 siRNA oligonucleotide was annealed as described and transfected at 200 pmol with X-tremeGENE siRNA for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. The PCK-containing media was removed after the overnight incubation followed by exposure at 365 nm for 2 min using a UV transilluminator (25 W) following protocol 5.4.3. After 48 h

incubation, total RNA was isolated from cells following protocol 5.4.6. Quantitative RT-PCR was performed with a GAPDH primer set[275] and an Eg5 primer set[271] (Table 4.3) following protocol 5.4.7. The threshold cycles (Ct) of each sample were normalized to the GAPDH control gene, and the inhibition of Eg5 expression is represented relative to nontreated cells.

Table 4.3: Sequences of qRT-PCR primers used.

Primer	Sequence (5' → 3')
GAPDH forward	TGCACCACCAACTGCTTAGC
GAPDH reverse	GGCATGGACTGTGGTCATGAG
Eg5 forward	CAGCTGAAAAGGAAACAGCC
Eg5 reverse	GGCATGGACTGTGGTCATGAG

4.2   Optical Control of CRISPR/Cas9 Gene Editing

This material was reprinted, in part, with permission from **Hemphill, J.**; Borchardt, E.K.; Brown, K.A.; Asokan, A.; Deiters, A. *J. Am. Chem. Soc.* **2015**, 137(17), 5642-5645.

The CRISPR/Cas9 system has emerged as an important tool in biomedical research for a wide range of applications, with significant potential for genome engineering and gene therapy. In order to achieve conditional control of the CRISPR/Cas9 system, a genetically encoded light-activated Cas9 was engineered through the site-specific installation of a caged lysine amino acid. Several potential lysine residues were identified as viable caging sites that can be modified to optically control Cas9 function, as demonstrated through optical activation and deactivation of both exogenous and endogenous gene function.

Many bacteria and archaea utilize an adaptive immune defense based on a system of clustered regularly interspaced short palindromic repeats (CRISPR) that target invasive nucleic acids through the interaction of CRISPR-associated (Cas) proteins and CRISPR arrays, which are transcribed and processed into short CRISPR RNAs (crRNAs).[276] The crRNA guides the Cas proteins to sequence-specific duplex cleavage. Type II CRISPR systems employ an additional RNA, the trans-activating crRNA (tracrRNA), which hybridizes with the crRNA.[277] These two RNAs can be combined to allow Cas9 targeting with a single guide RNA (gRNA).[278] The Cas9 enzyme has been optimized for site-specific DNA cleavage and nicking followed by non-homologous end-joining (NHEJ) or homology-directed repair (HR), enabling gene editing, gene deletion, and gene mutation[279] in human cells[280] and animal models.[281] The ease of customized gRNA design allows for sequence-specific and highly efficient gene targeting without the need for protein engineering.[282] In addition, a catalytically inactive Cas9 has been engineered into a transcriptional activator and repressor, expanding the utility of Cas9 as a gene regulatory tool.[283]

Optically regulating Cas9 function enables precise spatial and temporal control of gene editing. In order to develop a system for optochemical control of CRISPR/Cas9 gene editing (Figure 4.12), genetic code expansion was used by adding an engineered pyrrolysyl tRNA (PylT)/tRNA synthetase (PCKRS) pair to the translational machinery of human cells to enable the site-specific incorporation of PCK into proteins (as discussed in Section 1.2).[114] Recently, the use of Cry domains was applied to the light-induced recruitment of a transcriptional activator for a DNA-bound enzymatically inactive Cas9 variant (dCas9).[284] Additionally, a fully functional Cas9 was modified with Cry domains in the development of light-activated CRISPR/Cas9 gene editing.[285] These optogenetic systems utilized blue light optical dimerization of Cry2 and CIBN (discussed in Section 1.2) domains fused to split protein fragments to control transcription or

genomic editing for multiple targets in cellular systems, and successfully demonstrated both spatial control as well as light switching reversibility of Cas9 activity. In contrast, our approach for Cas9 modification with a site-specific caged lysine analogue enables the generation of wild-type Cas9 through light-induced decaging, which is not encumbered by photosensitive fusion domains.
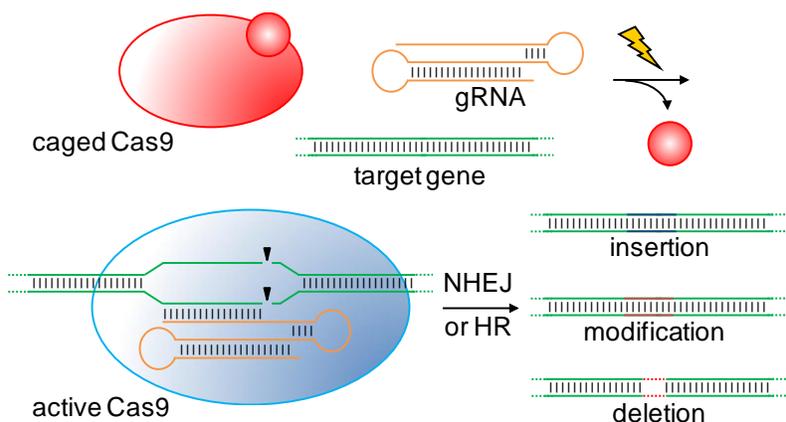
Figure 4.12: Light-activation of caged Cas9 enables optochemical control of gene editing. The caged Cas9 protein contains a site-specifically incorporated PCK, rendering it inactive until the caging group is removed through light exposure. This generates wild-type Cas9, which induces sequence-specific DNA cleavage followed by subsequent non-homologous end-joining (NHEJ) or homology-directed repair (HR).

Multiple lysines of interest were identified as potential caging sites for the inhibition of CRISPR/Cas9 function (Figure 4.13). K76, K163, K510, and K742 are in close proximity to the gRNA nucleic acid binding sites based on recent crystal structures,[286] and thus may be essential for Cas9-gRNA interaction. K866 undergoes a significant conformational change upon binding of the gRNA, orienting the lysine to become surface exposed, which may be necessary to properly position the target DNA strand for cleavage. However, the exact role of this residue has not been determined.
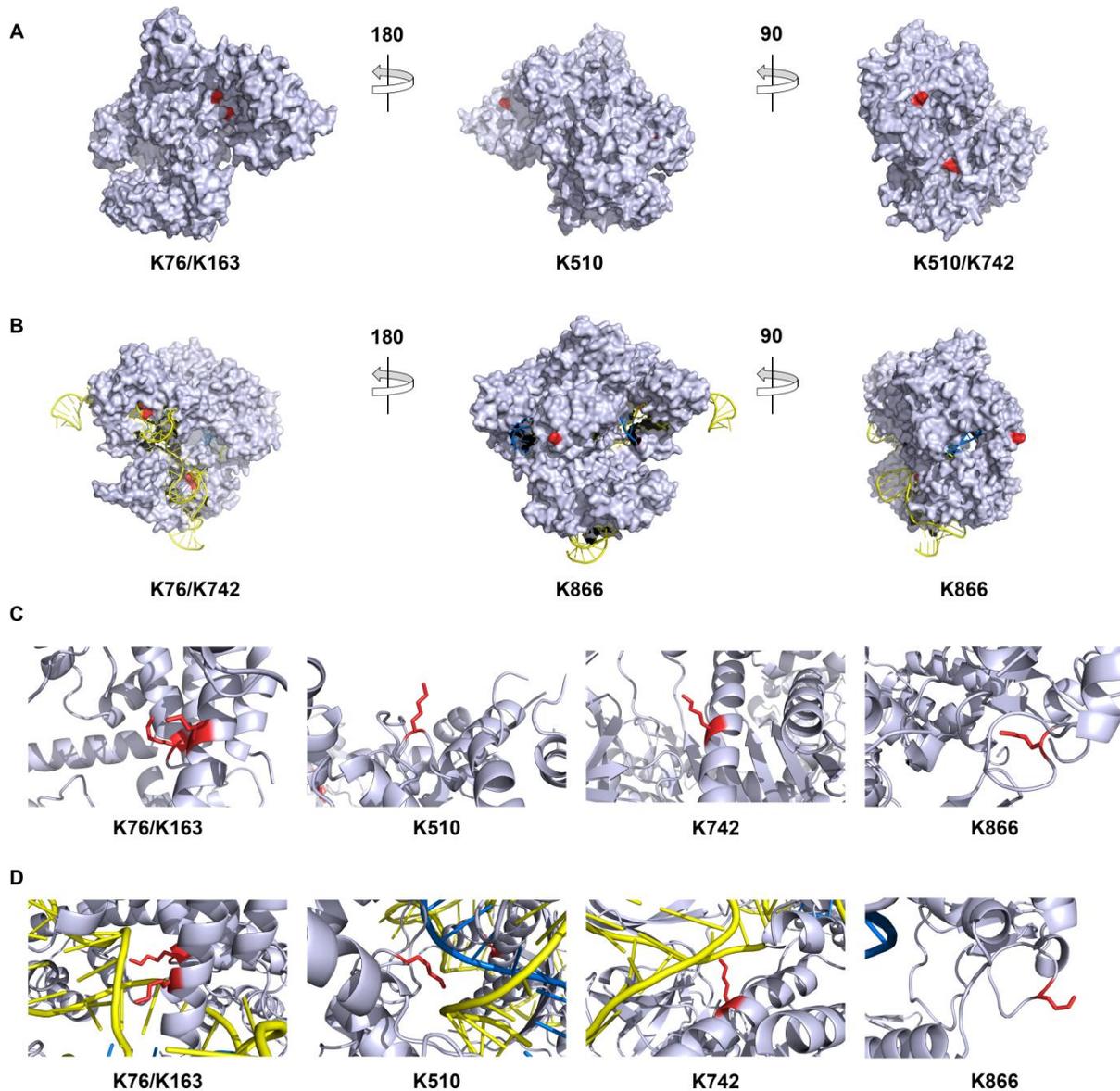
Figure 4.13: Structural annotation of critical lysines on Cas9.
A) Lysines of interest (red) depicted on a surface model of unbound Cas9 (PDB: 4CMP). B) Lysines of interest depicted on a surface model of bound Cas9 (PDB: 4OO8), with gRNA (yellow) and target DNA (blue) shown. C) Detailed view of each lysine of interest in the unbound Cas9 structure. D) Detailed view of each lysine of interest in the bound Cas9 structure.

The Cas9 gene was first cloned into the PCKRS expression plasmid (pAG31, described in Section 4.1)[114] and an HA-tag was added. Alanine mutations were introduced at the five lysine sites to assay the lysine residues of interest for Cas9 activity. Western blots confirmed all Cas9

alanine mutants expressed well in HEK293T cells (Figure 4.14A). Amber stop codon (TAG) mutations were introduced at all sites of interest since the K→PCK mutation may induce an additional level of perturbation compared to a K→A mutation. For the expression of caged Cas9, cotransfections were performed with the PylT expression plasmid (pAG38) in HEK293T cells and the cells were incubated for 24 h in the absence or presence of PCK. Western blots confirmed PCK-dependent expression of the caged Cas9 mutants (Figure 4.14B).
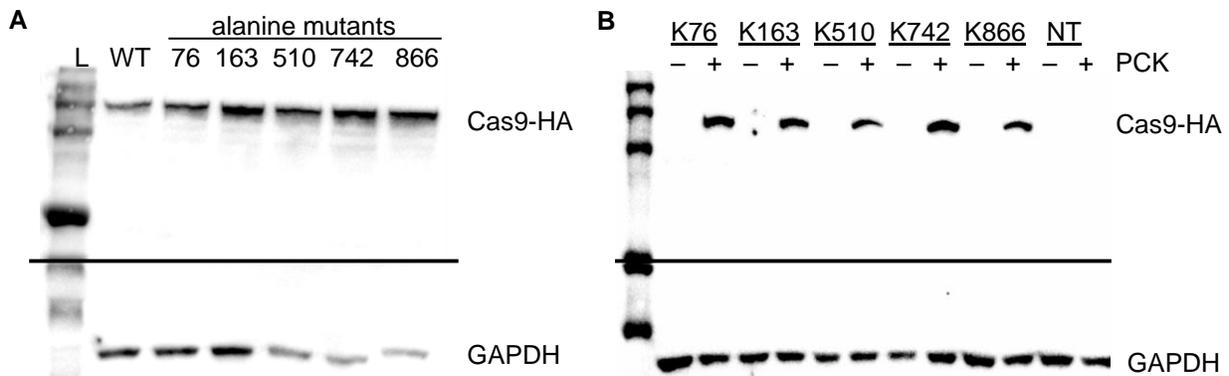
Figure 4.14: Western blot analysis of Cas9 expressions.
HEK293T cells were transfected with the plasmid expression systems and protein was purified for chemiluminescent detection of the C-terminal HA tag. A) Cas9 expressions with the pAG31:Cas9 K→A mutant plasmids. B) Caged Cas9 expressions with the pAG31:Cas9 K→TAG mutant and pAG38 plasmids, incubated in the absence or presence of PCK (2 mM) for 24 h. Western blot analysis was performed using HA and GAPDH primary antibodies with a HRP secondary antibody for chemiluminescent detection. The Cas9-HA and GAPDH control bands are annotated, and a horizontal line indicates a cut site on the transfer membrane for antibody staining. These experiments were conducted by Kalyn Brown.

A dual reporter assay based on pIRG[287] was developed by the Asokan lab, which switches from expressing DsRed to expressing EGFP in the presence of functional Cas9 and matching gRNAs (Figure 4.15A and B). Two gRNAs (Table 4.7) were designed to target sequences upstream and downstream of the DsRed-terminator cassette. Upon cotransfection with active Cas9, these gRNAs direct the excision of DsRed, and the reporter plasmid is repaired to activate EGFP expression, as successfully demonstrated in HEK293T cells with wild-type (WT)

Cas9 (Figure 4.15C and D). It was also verified that the reporter plasmid assay is not responsive

to UV exposure in the absence of Cas9 (Figure 4.15C).
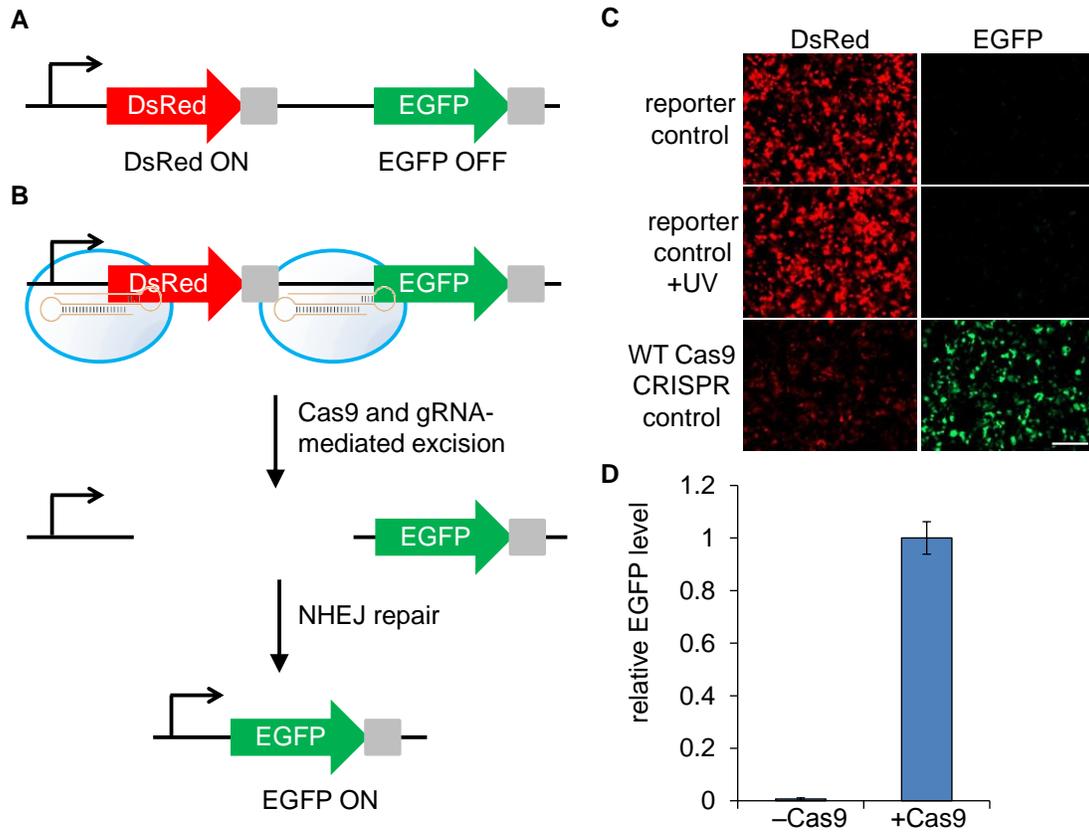
Figure 4.15: Dual reporter CRISPR/Cas9 activity assay.
A) Depiction of the dual reporter locus. A DsRed gene (red arrow) and an EGFP gene (green arrow) are separated by a transcription termination sequence (grey boxes). In the absence of Cas9, transcription terminates immediately following DsRed, allowing only DsRed expression. B) When functional Cas9 (blue) and gRNAs (orange) are present, the complex mediates excision of the DsRed-terminator cassette and NHEJ repair allows expression of EGFP. C) HEK293T cells were transfected with the pRG dual reporter, pAG31:Cas9 WT, and gRNA2 expression plasmids. DsRed and EGFP fluorescence were imaged (10X magnification) at 48 h. Scale bar indicates 200 μm. D) Analysis of EGFP expression by imaging cytometry. Error bars represent standard deviations from three replicates.

This assay was used in an initial alanine scan of residues K76, K163, K540, K742, and

K866, showing that four of the Cas9 alanine mutants were still active (Figure 4.16). However,

K866 was identified as being essential for activity, suggesting it was a potential target for the introduction of PCK.
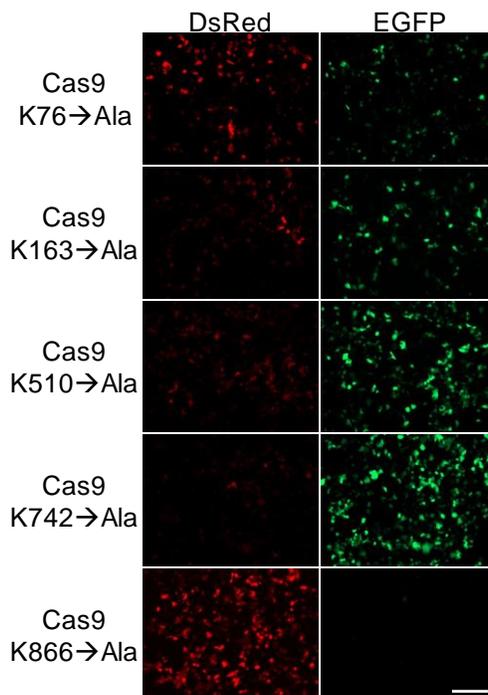
Figure 4.16: Cas9 alanine mutant activity scanning.
The pAG31:Cas9 K→A mutant expression plasmids were transfected into HEK293T cells together with the dual reporter system (pRG and gRNA2 plasmids). DsRed and EGFP fluorescence were imaged (10X magnification) at 48 h. Scale bar indicates 200 μm.

The function of the caged Cas9 mutants was then tested in the presence and absence of UV exposure (365 nm, 2 min) using the dual reporter assay (Figure 4.17). The incorporation of PCK at K76, K163, and K866 showed full inhibition of Cas9 activity in the absence of UV exposure, while the K742PCK mutant was still functional, similar to wild-type. Additionally, the K510PCK mutant showed a low level of undesired background activity in the absence of UV exposure. After light-activation, the K163PCK, K510PCK, and K866PCK mutants showed successful light-activation of Cas9 as observed through the expression of EGFP, while the K76PCK mutant was surprisingly not activated. In contrast to wild-type Cas9, all cells

expressing light-activated Cas9 mutants still showed DsRed fluorescence, since the caged Cas9 activation occurred 24 h after transfection, while the wild-type Cas9 was immediately active once expressed. Thus, in the case of light-activated Cas9, DsRed protein that has already been expressed persists, with a half-life greater than 4 days.[288]
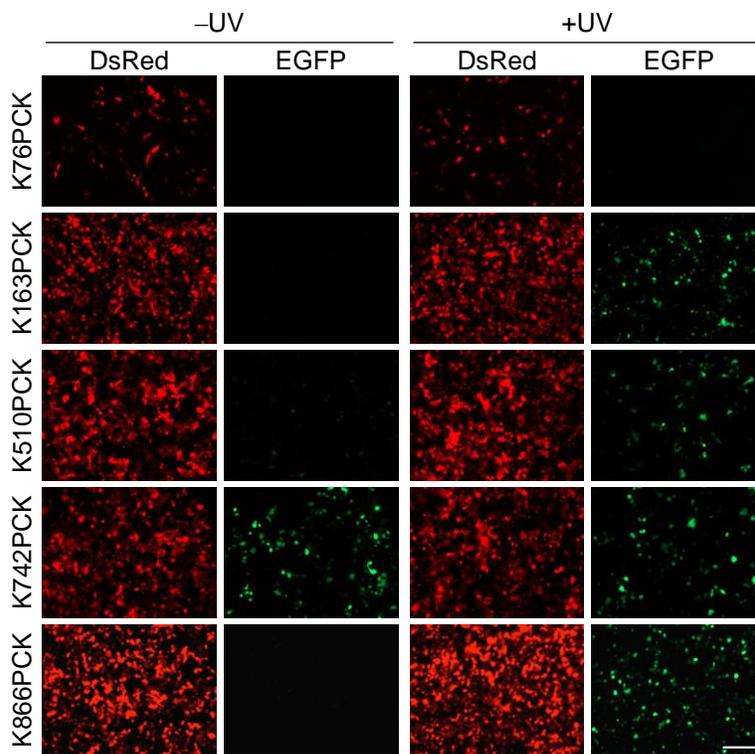
Figure 4.17: Cas9 PCK mutant activity scanning.
HEK293T cells were transfected with the caged Cas9 expression system (pAG31:Cas9 K→TAG and pAG38) as well as the dual reporter constructs (pRG and gRNAs), then incubated in the presence of PCK (2 mM) for 24 h. The cells were kept in the dark or UV irradiated for 2 min and imaged (10X magnification) after 48 h. Scale bar indicates 200 μm.

Two lysine sites were successfully identified that are amenable to optically control of Cas9 function, presenting two different pathways for the light-activation of CRISPR/Cas9 gene editing: (1) K163, which might interact with the gRNA, and (2) K866, which may play a role in positioning the nontarget DNA strand (Figure 4.13). The Cas9 K866PCK mutant showed

minimal background activity before irradiation and high activity after light exposure for 2 min, reaching wild-type Cas9 levels (Figure 4.18A). In order to show spatial control of CRISPR/Cas9 gene editing in mammalian cells, the activation of K866-caged Cas9 was performed through patterned UV irradiation of a layer of HEK293T cells (Figure 4.18B).
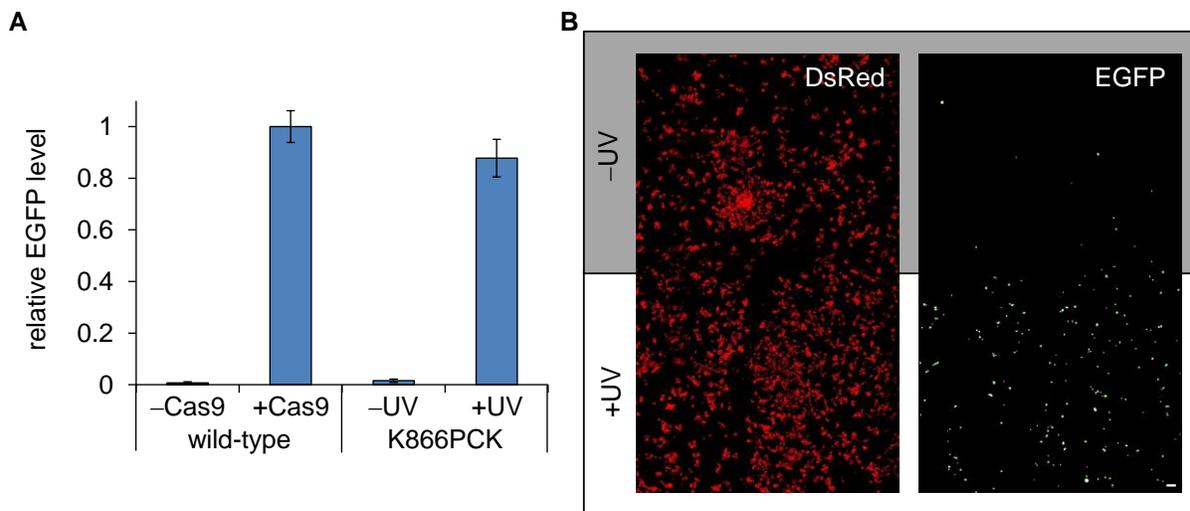
Figure 4.18: Analysis of K866-caged Cas9 light activation.
A) HEK293T cells were transfected with WT Cas9 or the K866-caged Cas9 expression system and dual reporter constructs as previously described, then incubated in the presence of PCK (2 mM) for 24 h. UV irradiations were performed on a 365 nm transilluminator for 2 min. Imaging was performed 48 h after UV exposure (10X magnification) and EGFP expression was quantified by imaging cytometry. Error bars represent standard deviations from three replicates. B) HEK293T cells were transfected with the K866-caged Cas9 expression system and dual reporter construct then incubated in the presence of PCK (2 mM) for 24 h. Spatial control of CRISPR/Cas9 gene editing was performed with UV irradiation of a subset of cells through a defined mask. Imaging was performed (20X magnification) with tile stitching (2X3) after 48 h. Scale bar indicates 200 μm.

A UV irradiation time course was also performed with the WT and K866-caged Cas9 (Figure 4.19), demonstrating that exposure of >2 min did not further enhance activation. Taken together, these experiments successfully demonstrate spatial and temporal control of gene function using a site-specifically caged Cas9 that is genetically encoded in mammalian cells, allowing for conditional light-activation of CRISPR/Cas9 gene editing.
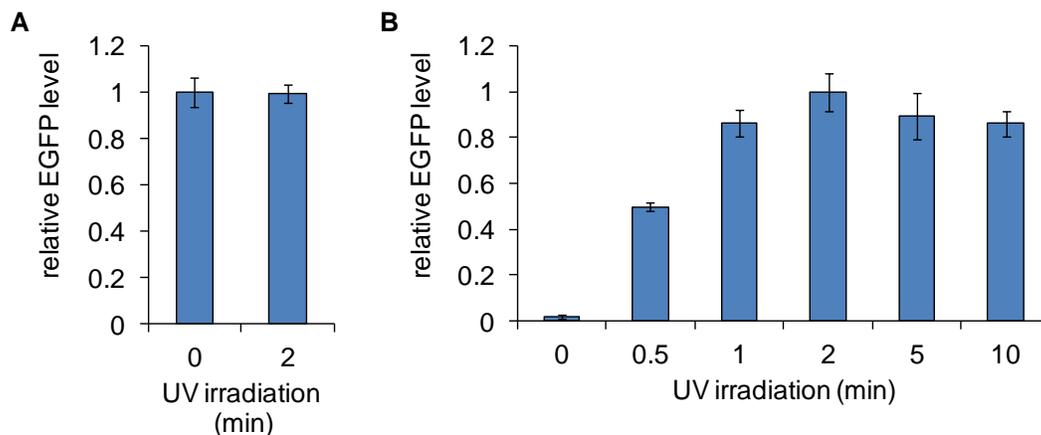
Figure 4.19: Cas9 UV irradiation optimization.
A) WT Cas9. B) K866-caged Cas9. Expression plasmids were transfected into HEK293T cells with the dual reporter system and PCK incorporation constructs, then incubated in the presence of PCK (2 mM) for 24 h. The cells were then UV irradiated for 0-10 min and imaged after 48 h incubation. The images were counted for EGFP expressing cells. Error bars represent the standard deviations of three replicates.

The mechanism of deactivation by the K163PCK and K866PCK mutations was further analyzed through plasmid cleavage assays with purified Cas9 (Figure 4.20), showing no DNA cleavage or nicking activity in the absence of UV irradiation. The absence of any catalytic activity of the caged enzyme suggests that interaction with the gRNA and/or target DNA may be inhibited through introduction of PCK, thereby deactivating Cas9 function.
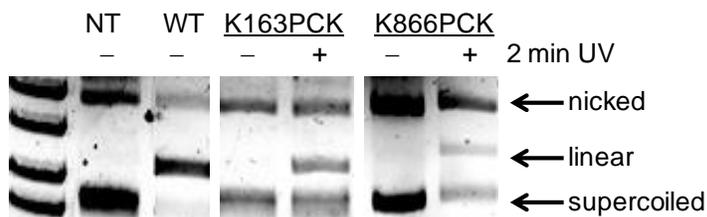
Figure 4.20: Cas9 DNA cleavage assays.
The Cas9 proteins were expressed in HEK293T cells and purified from lysate using the HA-tag. Nontreated (NT) and wild-type (WT) Cas9 purifications are included. The PCK-caged Cas9 proteins were UV irradiated for 2 min prior to purification. The Cas9 purifications were then incubated with the dual reporter plasmid and synthetically transcribed EGFP gRNA overnight at

37 °C. The cleavage assay products were analyzed on an agarose gel. Nicked, linear, and supercoiled plasmid populations are indicated.

Silencing of an endogenous gene was then demonstrated through light-activated gene editing using the optically activated Cas9. The transmembrane transferrin receptor CD71 (also known as TfR1), associated with leukemia and lymphoma,[289] was selected as a target. Multiple gRNAs for the 5′UTR[283] and exons within the coding region of CD71 were designed in order to inhibit protein function via disruption of upstream regulatory elements or disruption of the amino acid sequence, and subsequently cloned into the PylT expression plasmid. First, CD71 knockdown by wild-type CRISPR/Cas9 was confirmed through quantification via both qRT-PCR and phenotype analysis with fluorescent antibody staining of HeLa cells (Figure 4.21). Indeed, repression of CD71 mRNA with the 5′UTR targeting gRNA was observed (~70%) while the exon-targeting gRNAs showed no effect on mRNA levels. In addition, reduced levels of CD71 protein (~75%) were observed for all gRNAs, relative to nontreated cells.
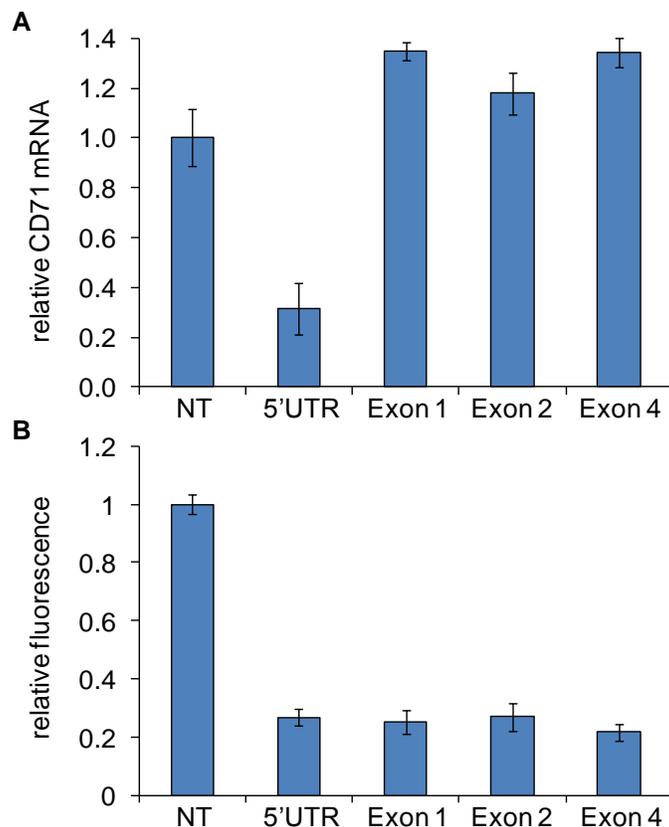
Figure 4.21: Silencing of CD71 expression with wild-type Cas9.
HeLa cells were transfected with the pAG31:Cas9 WT and pAG38:CD71 gRNA expression constructs then incubated for 48 h. A) Quantitative real-time PCR analysis of CD71 mRNA, normalized to the GAPDH control gene. B) Fluorescence detection of cell-surface CD71 protein. Data are shown relative to nontreated (NT) cells, and error bars represent the standard deviations of three replicates.

Next, CD71 suppression was performed with the K866-caged Cas9 in the absence and presence of light-activation. Quantification of CD71 mRNA revealed a reduction by approximately 60% only in the case of the gRNA targeting the CD71 5′UTR (Figure 4.22A). This is not surprising, as the introduction of mutations into the coding region of the CD71 gene should not affect transcription. Light-activation of Cas9 function reduced the presence of CD71 protein on the cell surface by approximately 50% (Figure 4.22B), demonstrating the ability to optically control the silencing of endogenous genes with a caged CRISPR/Cas9 system. The reduced repression observed for the light-activated K866-caged Cas9 compared to wild-type

209

Cas9 (10–25%) may require further optimization for endogenous gene editing applications, although the frequency of indels has not been determined, limiting quantitative assessment on the genomic level. Thus, the optically activated CRISPR/Cas9 system can be applied to the control of endogenous genes on the genomic level.
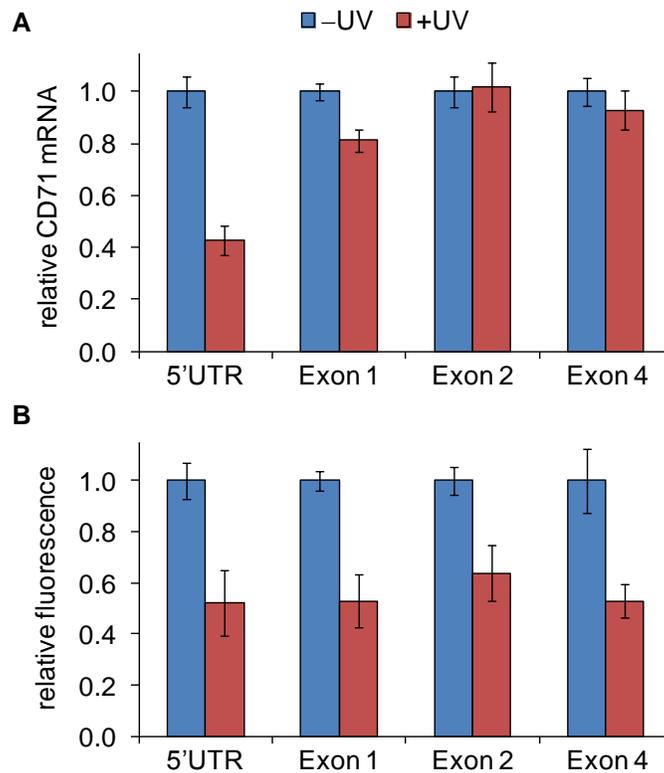
Figure 4.22: Light-activated CRISPR/Cas9 silencing of CD71 in HeLa cells.
HeLa cells were transfected with the pAG31:Cas9K866TAG and pAG38:CD71 gRNA expression constructs then incubated in the presence of PCK (2 mM) for 24 h. The cells were then UV irradiated for 2 min and analyzed after 48 h. A) Quantitative real-time PCR analysis of CD71 mRNA, normalized to the GAPDH control gene. B) Fluorescence detection of cell-surface CD71 protein. Data are shown relative to nonirradiated cells for each condition, and error bars represent standard deviations from three replicates.

In summary, a genetically encoded light-activated CRISPR/Cas9 system was successfully developed for conditional control of gene editing and gene expression. Through both alanine and UAA scanning, lysine residues that are important for Cas9 function were identified. Of these

sites, K866 was shown to be essential for activity and was utilized in the generation of a light-activated CRISPR/Cas9 system through incorporation of a photocaged lysine at that position. It was shown that the genetically encoded caged Cas9 could be applied to gene editing—for activation and deactivation—of both fluorescent reporters and endogenous genes in human cells. Importantly, the caged enzyme was completely inactive before UV illumination, and its activity could be restored up to wild-type levels through a brief 120 sec exposure to 365 nm light. Many cell types and model organisms have been modified by Cas9-mediated gene editing,[290] demonstrating the broad applicability of the CRISPR/Cas9 system. UAA mutagenesis based on the pyrrolysine system is expanding into model organisms, such as *C. elegans* and *D. melanogaster*,[291] providing future opportunities for optical control of CRISPR/Cas9 function beyond cell culture. Light-activation of CRISPR/Cas9 may allow for the study of gene function with high precision, and may reduce toxicity from off-target mutations[292] by restricting the function of Cas9 to certain locations or time points.

4.2.1 Methods and Materials

Plasmid Constructs. See Section 5.3 for specific information on molecular biology techniques. The CMV-driven Cas9 gene was PCR amplified from the hCas9 expression vector[280] (Addgene 41815) with primers shown in Table 4.4 to introduce both *NheI* and *MfeI* restriction sites as well as an HA tag on the C-terminus following protocol 5.3.1. The ~4.9 kB Cas9 gene insert was cloned between the *NheI* and *MfeI* restriction sites of the pAG31 PCKRS expression plasmid[114] following protocols 5.3.3 and 5.3.4. Alanine mutations and amber stop codons were introduced into Cas9 at five sites through site-directed mutagenesis with primers shown in Table 4.5 and Table 4.6 following protocol 5.3.2. The U6-driven CD71 gRNAs (Table 4.7) were also

constructed from gBlocks through PCR amplification with primers to introduce *Bsu36I* and *PacI* restriction sites following protocol 5.3.1, and then cloned between the *Bsu36I* and *PacI* restriction sites of the pAG38 PylT expression plasmid[114] following protocols 5.3.3 and 5.3.4. Exon-based gRNA sequences were identified from the human CD71 mRNA coding regions (NIH Gene ID: 7037) for target sites containing downstream PAM sequences and minimal predicted off-target effects against the human genome (hg19) using the Optimized CRISPR Design Tool from the Zhang lab, with a minimum quality score cutoff of 92. Plasmid maps are shown in Figure 4.23. The pRG dual reporter and gRNA2 plasmids were generated by Erin Borchardt in the Asokan lab.

Table 4.4: Sequences for gene insertion of Cas9 into pAG31.
Restriction sites bolded and HA tag underlined.

Strand	Sequence (5' → 3')
hCas9 *Nhe1* Forward	TTAA**GCTAGC**ACCATGGACAAGAAGT
hCas9-HA *Mfe1* Reverse	CGGT**GAATTC**TTA<u>AGCGTAATCTGGAAC ATCGTATGGGTA</u>CACCTTCCTCTTCTTC

Table 4.5: Sequences of primers used in the development of K→Ala Cas9 mutants. Mutations introduced capitalized and bolded**.**

Primer	Sequence (5' → 3')
K76AlaForward	gcaga**GC**gaatcggatctgctacctgcaggagatc
K76ala Reverse	cgattc**GC**tctgcgggtatatctgcgccgtgctgt
K163Ala Forward	tatgatc**GC**atttcggggacacttcctcatcgaggggg
K163Ala Reverse	cccgaaat**GC**gatcatatgcgccagcgcgagatagat
K510Ala Forward	cttcct**GC**acactctctgctgtacgagtacttcacagtttataacgagctcaccaa
K510Ala Reverse	agagtgt**GC**aggaagcacctttttcgttaggcagattttttatcaaagttagtcatcc
K742Ala Forward	accgtt**GC**ggtcgtggatgaactcgtcaaagtaa
K742Ala Reverse	cgacc**GC**aacggtctgcagtattccctttttgat
K866Ala Forward	agaggg**GC**gagtgataacgtcccctcagaag
K866Ala Reverse	tcactc**GC**ccctctattttttatcggatcttgtcaacac

Table 4.6: Sequences of primers used in the development of K→TAG Cas9 mutants. Mutations introduced capitalized and bolded.

Primer	Sequence (5' → 3')
K76TAG Forward	cgcaga**T**agaatcggatctgctacctgcaggagatc
K76TAG Reverse	ccgattct**A**tctgcgggtatatctgcgccgtgctgt
K163TAG	tatgatc**T**a**G**tttcggggacacttcctcatcgaggggg
K163TAG Reverse	cccgaaa**C**t**A**gatcatatgcgccagcgcgagatagat
K510TAG	cttcct**T**a**G**cactctctgctgtacgagtacttcacagtttataacgagctcaccaa
K510TAG Reverse	agagtg**C**t**A**aggaagcacctttttcgttaggcagattttttatcaaagttagtcatcc
K742TAG	accgtt**T**aggtcgtggatgaactcgtcaaagtaa
K742TAG Reverse	cgacct**A**aacggtctgcagtattccctttttgat
K866TAG	agaggg**T**agagtgataacgtcccctcagaag
K866TAG Reverse	tcactct**A**ccctctattttttatcggatcttgtcaacac

Table 4.7: Sequences of gRNA constructs.
Primer binding sites capitalized, restriction sites bolded, poly-T transcription terminator sequences underlined, and gRNAs in blue (guide target sequences capitalized and underlined).

Strand	Sequence (5' → 3')
DsRed gRNA	U6:<u>TCGACTCTAGAGGATCCAC</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>
EGFP gRNA	U6:<u>TAGCTAGTCTAGGTCGATGC</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>
Bsu361 Forward	atat**cctaagg**ACTGGTCAACTGGCTAATCG
Pac1 Reverse	atat**ttaattaa**CGATACGATGAGCTAGGCAT
CD71 gRNA 5'UTR	U6:<u>GGACGCGCTAGTGTGAGTGC</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>ATGCCTAGCTCATCGTATCG
CD71 gRNA Exon 1	U6:<u>GTCATATACCCGGTTCAGCC</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>ATGCCTAGCTCATCGTATCG
CD71 gRNA Exon 2	U6:<u>CTGCAGCACGTCGCTTATAT</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>ATGCCTAGCTCATCGTATCG
CD71 gRNA Exon 4	U6:<u>GGGTTATGTGGCGTATAGTA</u>gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt<u>tttttttt</u>ATGCCTAGCTCATCGTATCG
U6-promoter	ACTGGTCAACTGGCTAATCGtgtacaaaaaagcaggctttaaaggaaccaattcagtcgactggatccggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgtttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccg
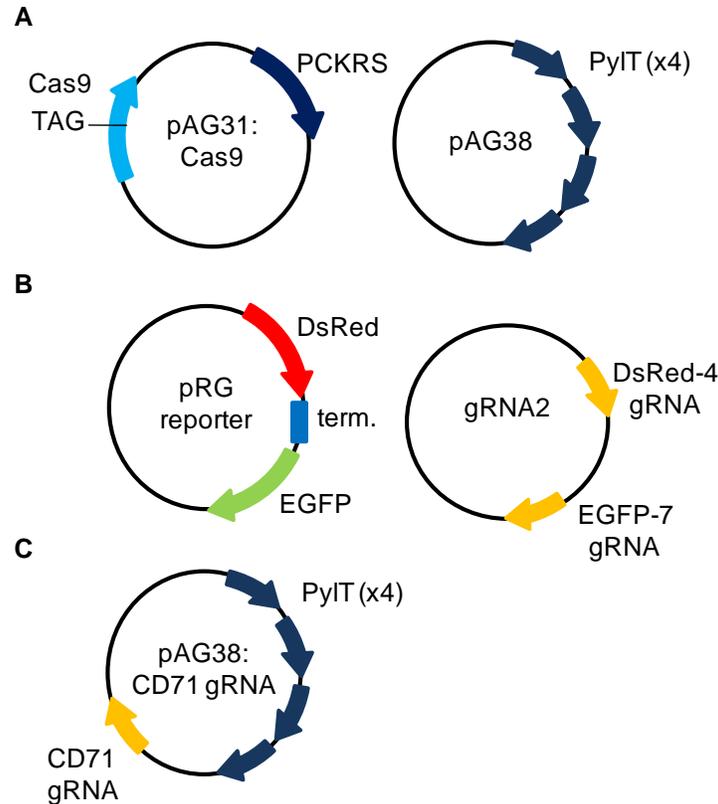
Figure 4.23: Plasmid constructs for light-activated CRISPR/Cas9.
A) Expression of caged Cas9. B) Analysis of activity with specific reporter genes. C) CD71 gRNA construct.

Western Blots. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 6-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 2 μg of each plasmid using lipofectamine for 4 h following protocol 5.4.2. For the expression of caged Cas9, the cellular growth media was supplemented with PCK (2 mM) for 48 h. Protein isolations were performed following protocol 5.4.8. Western blots were performed following protocol 5.4.9 using a mouse-anti-HA primary antibody and a mouse-anti-GADPH control (Santa Cruz Biotechnology). The primary antibodies were detected with a goat-anti-mouse-HRP secondary antibody for chemiluminescent analysis with the VisiGlo kit (Amresco).

Optical Activation of Reporter Gene Editing. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 96-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 200 ng of each plasmid using bPEI overnight following protocol 5.4.2. The transfection mixtures were supplemented with PCK (2 mM) and removed after overnight incubations, followed by exposure to 365 nm using a UV transilluminator (25 W) following protocol 5.4.3. Fluorescent imaging of the dual reporter was performed after 48 h incubations on a Zeiss Observer Z1 microscope (10X magnification) following protocol 5.4.5. Fluorescent cell counting was performed in ImageJ software (NIH - settings: threshold 5-10%, size >200 pixels^2, and circularity 0-1). Error bars represent the standard deviations of three replicates. For the spatial control experiments, UV irradiations were performed through a tin foil mask to only expose a subset of cells to 365 nm light following protocol 5.4.3. Microscopy imaging was then performed on a Nikon A1 confocal microscope (20X magnification) in a tiled grid (10x10) and stitched using Elements software.

DNA Cleavage Assays. See Section 5.4 for specific information on cell culture techniques. HEK293T cells were passaged into 6-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 2 µg of each plasmid using lipofectamine for 4 h following protocol 5.4.2. For the expression of caged Cas9, the cellular growth media was supplemented with PCK (2 mM) for 48 h. Protein isolations and HA-tag immunoprecipitation were performed on total protein lysate (~200 µL) following protocol 5.4.8. Synthetic EGFP gRNA was produced from a PCR template for *in vitro* T7RNAp transcription. The single stranded template (Table 4.8) was PCR amplified with T7 forward and reverse primers and used in 20 µL transcription reactions with the MEGAscript T7 transcription

216

kit (Life Technologies) using 500-1000 ng of PCR template according to the manufacturer's protocol. The DNA cleavage assays were performed by incubating 40 µL of the Cas9 HA-bead immobilized protein purifications, 2 µL of synthetic EGFP gRNA purification (~6000 ng total, pre-annealed by cooling from 65 to 12 °C over 10 min in TAE/Mg^{2+} buffer [0.04 M tris-acetate, 1 mM EDTA, and 12.5 mM magnesium acetate]), and 200 ng of the dual reporter plasmid in Cas9 activity buffer[278] [20mM HEPES, 150 mM KCl, 0.5 mM DTT, 0.1 mM EDTA, 10 mM MgCl$_2$, pH 7.4] shaking overnight at 37 °C. The reaction products were then denatured at 72 °C for 20 min, ran on a 0.8% agarose gel (200V, 1 h), and stained with ethidium bromide.

Table 4.8: Sequences of templates and primers used for synthetic gRNA transcription. T7RNAp promoter sequence capitalized.

Strand	Sequence (5' → 3')
EGFP gRNA template	TAATACGACTCACTATAGGGAGAtagctagtctaggtcgatgcgttttagagct agaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctt
T7 Forward	TAATACGACTCACTATAGGG
T7 Reverse	aaagcaccgactcggtgcca

Photochemical Regulation of Endogenous CD71. See Section 5.4 for specific information on cell culture techniques. HeLa cells were passaged into 96-well plates and grown to ~70% confluence within 24 h following protocol 5.4.1. Transfections were performed with 200 ng of each plasmid using lipofectamine for 4 h following protocol 5.4.2. The cellular growth media was supplemented with PCK (2 mM) for 24 h. The PCK containing media was then removed after the overnight incubation, followed by exposure to 365 nm using a UV transilluminator (25 W) following protocol 5.4.3. After 48 h incubation, both quantification of CD71 mRNA and fluorescent antibody detection of CD71 protein were performed. Quantification of CD71 mRNA was performed by qRT-PCR, in which total RNA was isolated

from cells following protocol 5.4.6. Reverse transcription and qRT-PCR was performed with the primer sets shown in Table 4.9 following protocol 5.4.7. The threshold cycles (Ct) of each sample were normalized to the GAPDH control gene, and set relative to either nontreated or nonirradiated cells. Protein quantification of cell-surface expressed CD71 was performed with anti-human CD71 APC fluorescent antibody (eBioscience) after the media was removed and a 200 μL 37 °C PBS wash was performed. The antibody was added (5 μL of 0.06 μg antibody in 50 μL of PBS) for 1 h at 37 °C. The cells were then washed three times in 200 μL 37 °C PBS and analyzed on a Tecan M1000 plate reader (ex: 635/5; em: 660/10).

Table 4.9: Sequences of qRT-PCR primers used.

Primer	Sequence (5' → 3')
GAPDH forward	TGCACCACCAACTGCTTAGC
GAPDH reverse	GGCATGGACTGTGGTCATGAG
CD71 forward	AAAATCCGGTGTAGGCACAG
CD71 reverse	GCACTCCAACTGGCAAAGAT

4.3   Expression of Caged Proteins in Live Zebrafish

In order to express caged proteins in a live animal, initial attempts were made to transition the technology for the pyrrolysyl tRNA/tRNA synthetase pair (PylT/PylRS) system of UAA incorporation to zebrafish embryos. Previously, UAA mutagenesis technology has been demonstrated in both *C. elegans* and *D. melanogaster* through utilization of orthogonal synthetase/tRNA pairs.[291, 293] However, this will be the first example of an expanded genetic code in a live vertebrate animal, which is a major advancement in the field of UAA

218

incorporation. The initial goals of this project focused on the incorporation of light-activated amino acids to enable precise optical control over protein function with unprecedented spatial and temporal resolution in zebrafish.

To rapidly assay for incorporation of the UAAs into an exogenous gene, the mCherry-TAG-EGFP fusion reporter containing a TAG amber codon inserted between the two fluorescent proteins[294] was used for microinjection into zebrafish embryos (Figure 4.24). The reporter mRNA was injected into 1-cell stage embryos together with mRNA encoding the synthetase (PylRS) and its cognate tRNA (PylT). In the absence of the UAA incorporation machinery, only expression of mCherry should be detected. However, in the presence of the PylRS/PylT pair and the amino acid of interest, the TAG codon will be recognized for site-specific incorporation allowing for EGFP to be translated.
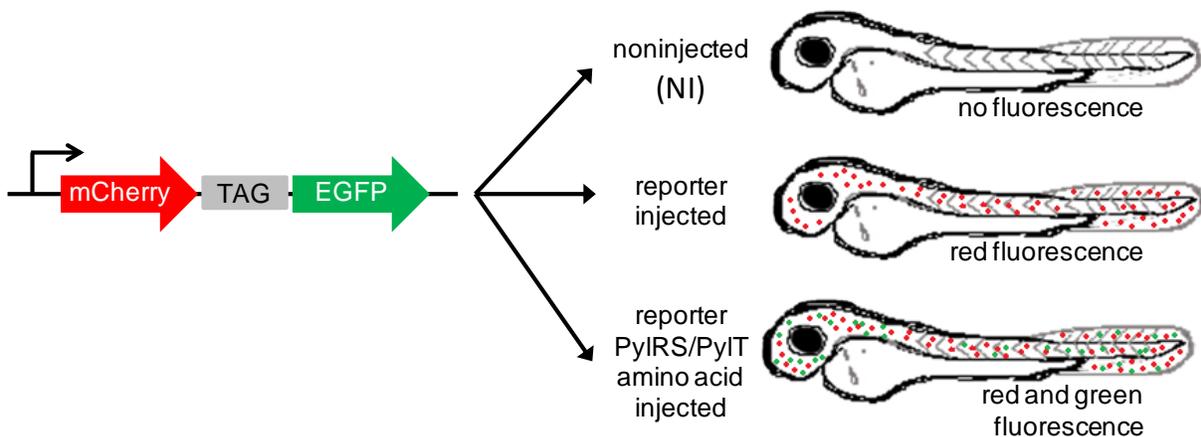
Figure 4.24: Schematic for analysis of UAA incorporation in zebrafish.
The mCherry-TAG-EGFP reporter (left) contains a TAG codon between mCherry and EGFP gene constructs. The mRNA is microinjected into zebrafish embryos at the 1-cell stage, and the developed embryos are imaged at 24 h for fluorescence (right). In the absence of the reporter mRNA, there is no fluorescence observed (top). With the reporter mRNA alone, only mCherry is expressed (middle). When the reporter mRNA, synthetase mRNA, and PylT are injected, both mCherry and EGFP are expressed with the amino acid present (bottom).

The coding sequences of both the fluorescent reporter and PylRS mutants of interest were cloned into vectors for *in vitro* transcription of the mRNAs (generated by Michael Tsang). Each component of the incorporation machinery was then tested individually for the effects of microinjection on mortality rates. The reporter mRNA and synthetase mRNA injections were tolerated at 100 pg, showing only a ~2.5-fold increases in mortality (Figure 4.25).
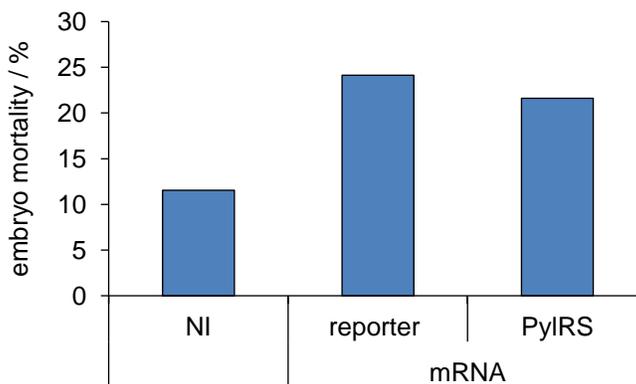
Figure 4.25: mRNA injection toxicity in zebrafish embryos.
Embryos were injected with the mCherry-TAG-EGFP reporter and PylRS mRNAs (100 pg) at the
1-cell stage. Mortality rates were calculated at 24 h. N = 26–29.

When the PylT was expressed from a U6 promoter via plasmid injection, near total toxicity was observed at 100 pg (Figure 4.26). However, when the PylT was injected as a synthetic oligonucleotide, injections up to 5 ng were tolerated. The PylT sequence contains a 3' CAA overhang, which is the site of amino acid attachment that is normally added by endogenous RNA polymerases in cells,[295] and typically is not included in the tRNA sequence on plasmid expression vectors.[296] Since the synthetic PylT will not be transcribed, the addition of a synthetic 3' CCA will skip processing from endogenous RNA polymerases in the zebrafish and enable a more direct path to aminoacylation by the synthetase. All subsequent experiments were performed with the synthetic PylT.
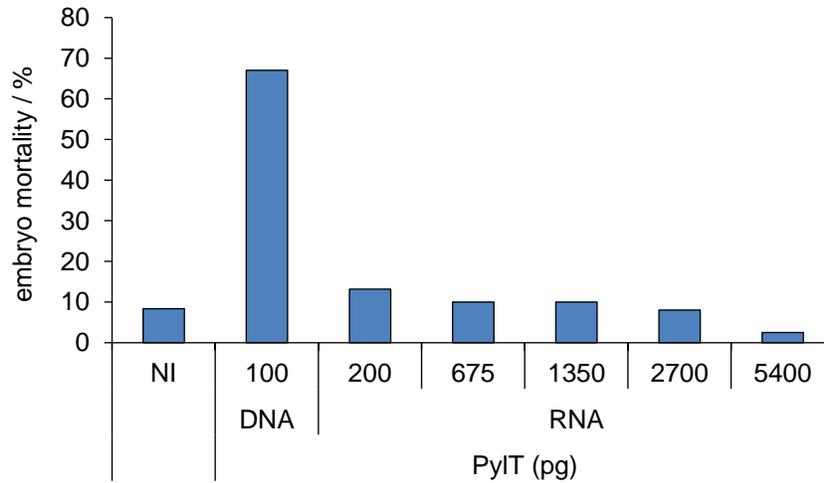
Figure 4.26: PylT injection toxicity in zebrafish embryos.
Embryos were injected with the PylT as either a DNA plasmid expression construct (100 pg) or synthetic oligonucleotide (200-5400 pg) at the 1-cell stage. Mortality rates were calculated at 24 h. $N = 38$-70.

The function of the reporter mRNA was then validated through injection and imaging of mCherry in zebrafish embryos after 24 h. The mCherry expression was observed in ~80% of embryos injected with the reporter mRNA, and an increased amount of mRNA did not increase the number of mCherry positive embryos (Figure 4.27A). The mCherry expression was observed at high intensity throughout the embryo (Figure 4.27B). Although yolk sac auto fluorescence was observed, no distinct EGFP expression was observed in any of the injected embryos since the TAG codon was read as a stop codon.
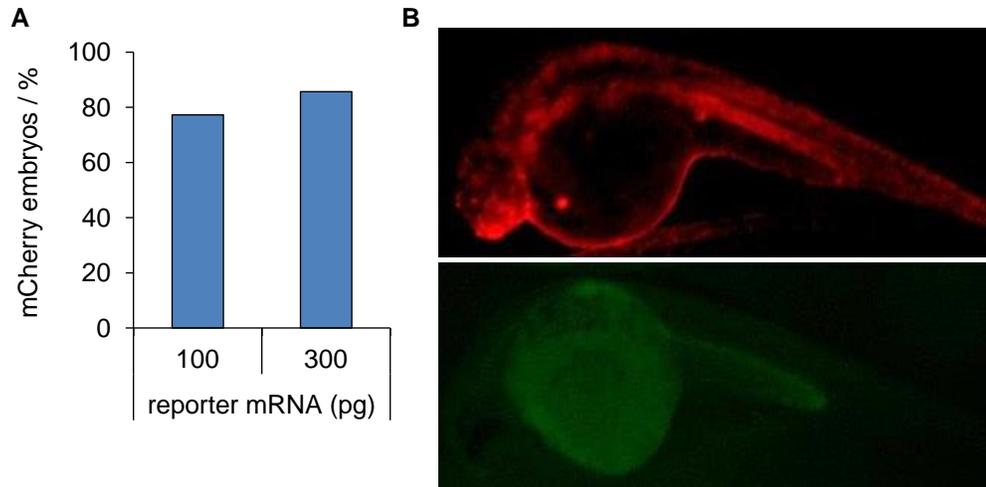
Figure 4.27: mCherry expression from reporter mRNA.
Embryos were injected with the reporter mRNA (100 or 300 pg) at the 1-cell stage. A) Frequency of mCherry positive embryos. B) Embryo micrographs (30X) of mCherry (top) and EGFP (bottom) fluorescence at 24 h from injection of mCherry-TAG-EGFP mRNA. N = 25-29.

To test the incorporation of amino acids with the PylRS/PylT pair, a "sloppy" synthetase mutant (generated by Jihe Liu) that incorporates phenylalanine (Phe) at the TAG codon was used for microinjection (PheRS, containing N311A and C313K mutations).[297] Mortality studies of Phe supplementation (up to 2 mM) through injection into the yolk were performed, and only minor increases in mortality were observed (Figure 4.28A). Previous experiments without Phe injection did not yield robust results. The incorporation of Phe was then tested with injection of the reporter mRNA, PheRS mRNA, PylT synthetic RNA, and yolk injection of the amino acid. Since the optimal amount of each component for successful amino acid incorporation was not known, the embryos were injected with the RNA mix at a 1:2:20 ratio (reporter:PheRS:PylT) based on the toxicity studies, which was followed by Phe yolk injection of 1 mM. After 24 h, the fish were imaged and counted for the expression of the fluorescent reporter genes (Figure 4.28B). The reporter mRNA produced mCherry expression in ~70% of the embryos analyzed, and EGFP expression was observed in ~40% of the mCherry positive embryos. These findings confirmed the site-specific incorporation of Phe at the TAG codon using the synthetase/tRNA pair through

EGFP expression (Figure 4.28C). Since the reporter alone did not exhibit any EGFP fluorescence, these early findings indicate that the PylRS/PylT system can be applied to live zebrafish embryos to potentially expand the genetic code and incorporate new amino acids in response to the TAG amber codon. However, continued optimization of site-specific amino acid incorporation needs to be performed to find optimal conditions (i.e., increase the amount of EGFP expressed) that enhance the efficiency of the system.
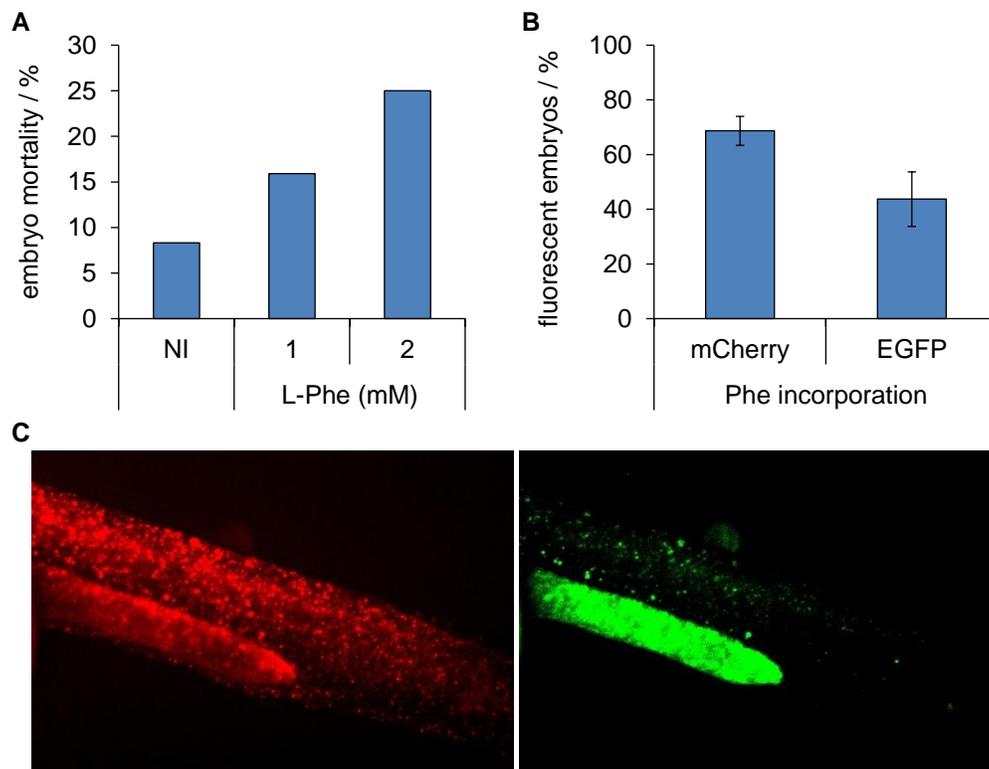
Figure 4.28: Incorporation of Phe in response to the TAG codon.
A) Phe was direct injected into the yolk of zebrafish embryos at 1-2 mM. Mortality rates were calculated at 24 h. B) Embryos were injected with the reporter mRNA, PheRS mRNA, and PylT synthetic RNA in the blastomrere at the 1-cell stage. Yolk injections of Phe (1 mM) were then performed. Fluorescent counts were determined at 24 h, and error bars represent standard deviations from three replicates. C) Embryo micrographs (60X) of mCherry (left) and EGFP (right) fluorescence at 24 h. $N = 36$-55.

While the Phe incorporation results were very encouraging, the system is based on a common amino acid found endogenously in the organism, and the incorporation of a truly UAA still needs to be demonstrated. Therefore, the injection of PCK into embryonic yolks was tested. Initial examination of PCK showed that yolk injections of the amino acid produced undesirably high rates of mortality, even at concentrations as low as 0.25 mM (Figure 4.29A). Efforts to reduce the toxicity of the synthetic amino acid solution, including pH balancing in Tris buffer, which increased the pH of the PCK solution from ~3 (acidic) to ~7 (neutral), proved to be unsuccessful. While the project was initially designed for the incorporation of PCK, the toxicity of the amino acid in this particular instance was determined to be too high to move forward. Thus, the toxicity of a coumarin-based caged analogue of lysine (HCK)[115] was tested, which showed greatly reduced toxicity compared to PCK (Figure 4.29B). These results demonstrated that the amino acid sensitivity of zebrafish embryos should always be examined prior to developing amino acid-specific incorporation systems.
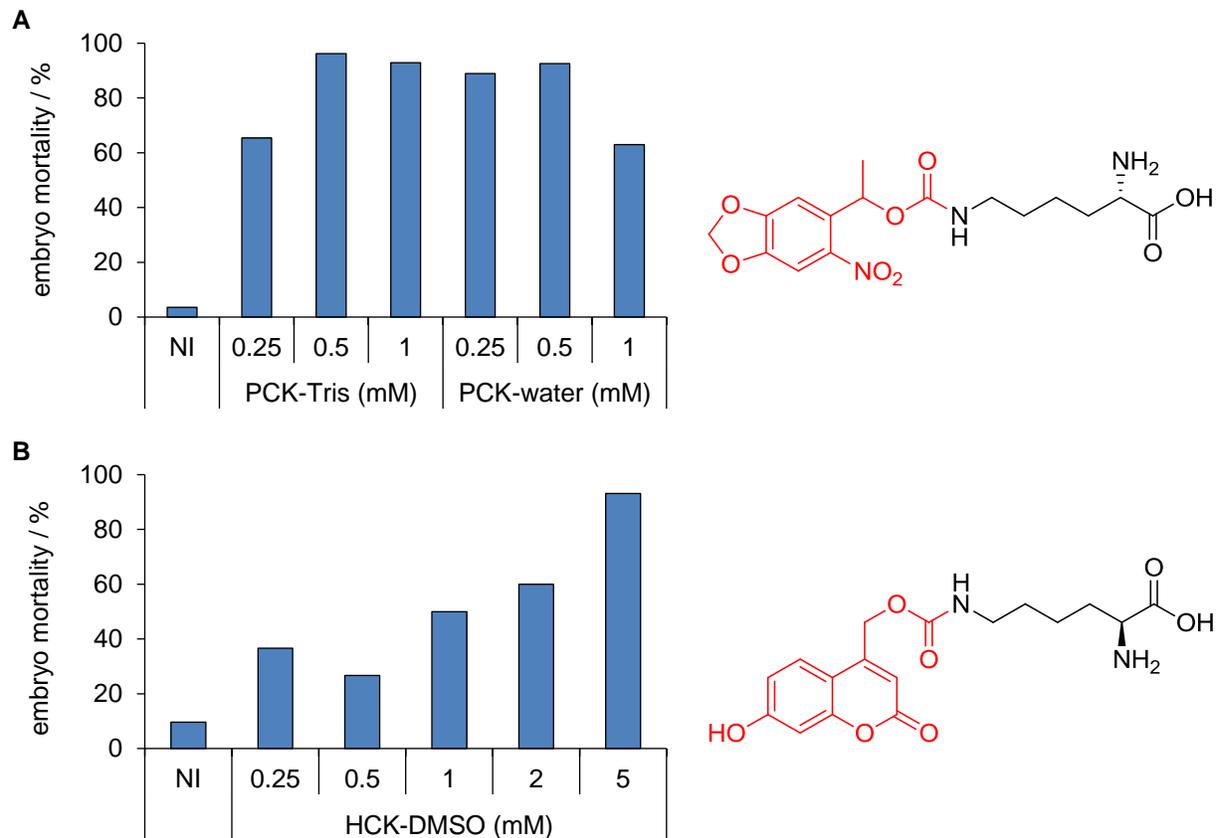
Figure 4.29: Toxicity of photocaged amino acids.
Caged lysine analogues were direct injected into the yolk of zebrafish embryos at the 1-cell stage, and mortality rates were calculated at 24 h. A) PCK (TFA salt) was diluted in Tris buffer pH 7.4 and water. B) HCK (TFA salt) was diluted into DMSO. The chemical structures of the amino acids are shown to the right, with the photolabile moieties indicated in red. $N = 26\text{-}32$.

With the initial studies on the PylT and amino acid toxicity completed, as well as successful demonstrations of Phe incorporation, future work includes the encoding of HCK and downstream expression of caged proteins. Once the activity of the PylRS/PylT pair for UAA incorporation in zebrafish has been confirmed through RNA injection, transgenic lines will be generated for caged amino acid incorporation and the expression of caged proteins. This will be the first vertebrate species with an expanded genetic code, and the precise spatiotemporal control of protein function will have a transformative effect on the entire zebrafish field.

4.3.1 Methods and Materials

Plasmid Constructs and RNA Preparation. The mCherry-TAG-EGFP reporter and pyrrolysyl synthetase (wild-type and mutant versions) were cloned into the pCS2+ plasmid for *in vitro* transcription of the mRNA by the Michael Tsang. The synthetic PylT was ordered from IDT with a 3' CCA overhang: rGrGrArArArCrCrUrGrArUrCrArUrGrUrArGrArUrCrGrArArC rGrGrArCrUrCrUrArArArArUrCrCrGrUrUrCrArGrCrCrGrGrGrUrUrArGrArUrUrCrCrCrGrGrGr GrUrUrUrCrCrGrCrCrA. The PylT oligo was resuspended to 100 μM in DEPC-treated water.

Zebrafish Maintenance and Injections. See Section 5.5 for specific information on zebrafish techniques. Embryos were microinjected into the blastomere at the 1-cell stage with the reporter mRNA, synthetase mRNAs, PylT expression vector (pAG38), and PylT synthetic RNA diluted 1:1 in phenol red following protocol 5.5.2. For Phe incorporation, a 1:2:20 ratio of RNAs (100 pg reporter mRNA, 200 pg synthetase mRNA (PheRS), and 2000 pg tRNA (PylT)) was injected in 2.2 nL. Yolk injections of amino acid solutions were performed with the final concentrations based on an approximate yolk volume of 110 nL.[298] Embryos were incubated in the dark at 28 °C for 24 h, then counted and imaged following protocol 5.5.3. mCherry scores were calculated with embryo counts of [(mCherry positive/alive)·100%]. EGFP scores were calculated with embryo counts of [(EGFP/mCherry)·100%]. For the fluorescent counts, the data from three independent experiments were averaged and standard deviations were calculated.

5.0   Expanded Methods and Materials

5.1   Sequences of Synthesized Oligonucleotides

Table 5.1: All oligonucleotides synthesized.
Nucleobase sequences are noted with capital letters (A = adenine, C = cytosine, G = guanine, T = thymine, and U = uracil). A letter prior to the base indicates sugar modification (m = 2'OMe). An asterisk (*) indicates phosphorothioate linkages between bases. Caged bases have been underlined and colored red, while cleavable linkers are shown with a triangle or box (DEACM = ▽, ONB = □). Terminal modifications are indicated (AmMO = amino modified, Uaq = 2′-anthraquinoylamido-2′-deoxyuridine cap).

Strand	Sequence (5' → 3')	Application
NF-kB HP decoy	TGGGGACTTTCCAGTTTCTGGAAAGTCCCCA	Hairpin decoy
DsRed HP decoy	GCGCGCGCAAACGCGCGCTACAACTCGGTGATGACGTTCTCGGAGGAGCGGCGCGCGCAAAGCGCGCG	Hairpin decoy
Eg5 HP decoy	GCGCGCGAAACGCGCGCTACCCGAGCTCTCTTATCAACAGCGGCGCGCGCAAAGCGCGCG	Hairpin decoy
CDTFO uncaged	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTATCGCGCGCAAAGCGCGCGATACCCCCCACCCCCACCCACGTA	Deactivated Triplex forming oligo
CDTFO 2C caged	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTATCGCGCGCAAAGCGCGCGATACC**C**CCCACCCC**C**CACCCACGTA	Deactivated Triplex forming oligo, caged
CDTFO 4C caged	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTATCGCGCGCAAAGCGCGCGATA**C**CCCC**C**ACCC**C**CAC**C**CACGTA	Deactivated Triplex forming oligo, caged
CDTFO 5C caged	GCGCGCGAAACGCGCGCTACGTGGGTGGGGGTGGGGGGTATCGCGCGCAAAGCGCGCGATA**C**CCC**C**CAC**C**CC**C**AC**C**CACGTA	Deactivated Triplex forming oligo, caged
miR-21 antagomir	mA*mU*mC*mA*mA*mC*mA*mU*mC*mA*mG*mU*mC*mU*mG*mA*mU*mA*mA*mG*mC*mU*mA	PS-2'OMe miRNA antagomir
miR-21 2U caged antagomir	mA*mU*mC*mA*mA*mC*mA*mU*mC*mA*mG*mU*mC*mU*mG*mA***mU***mA*mA*mG*mC***mU***mA	PS-2'OMe miRNA antagomir, caged

227

Strand	Sequence (5' → 3')	Application
miR-21 4U caged antagomir	mA***mU***mC*mA*mA*mC*mA***mU***mC*mA*mG*mU*mC*mU*mG*mA***mU***mA*mA*mG*mC***mU***mA	PS-2'OMe miRNA antagomir, caged
miR-22 antagomir	mA*mA*mC*mA*mG*mU*mU*mC*mU*mU*mC*mA*mA*mC*mU*mG*mG*mC*mA*mG*mC*mU*mU	PS-2'OMe miRNA antagomir (Nguyen collaboration)
miR-22 3U caged antagomir	mA*mA*mC*mA*mG*mU***mU***mC*mU*mU*mC*mA*mA*mC***mU***mG*mG*mC*mA*mG*mC***mU***mU	PS-2'OMe miRNA antagomir, caged (Nguyen collaboration)
miR-122 antagomir	mA*mC*mA*mA*mA*mC*mA*mC*mC*mA*mU*mU*mG*mU*mC*mA*mC*mA*mC*mU*mC*mC*mA	PS-2'OMe miRNA antagomir
miR-122 3dT antagomir	mA*mC*mA*mA*mA*mC*mA*mC*mC*mA*dTmU*mG*dT*mC*mA*mC*mA*mC*dT*mC*mC*mA	PS-2'OMe miRNA antagomir mixmer
miR-122 3U caged antagomir	mA*mC*mA*mA*mA*mC*mA*mC*mC*mA***mU***mU*mG***mU***mC*mA*mC*mA*mC***mU***mC*mC*mA	PS-2'OMe miRNA antagomir, caged
miR-124 antagomir	mA*mG*mG*mC*mA*mU*mU*mC*mA*mC*mC*mG*mC*mG*mU*mG*mC*mC*mU*mU*mA	PS-2'OMe miRNA antagomir (Nguyen collaboration)
miR-124 3U caged antagomir	mA*mG*mG*mC*mA*mU***mU***mC*mA*mC*mC*mG*mC*mG***mU***mG*mC*mC*mU***mU***mA	PS-2'OMe miRNA antagomir, caged (Nguyen collaboration)
miR-134 antagomir	mC*mC*mC*mC*mU*mC*mU*mG*mG*mU*mC*mA*mA*mC*mC*mA*mG*mU*mC*mA*mC*mA	PS-2'OMe miRNA antagomir (Schratt collaboration)
miR-134 3U caged antagomir	mC*mC*mC*mC***mU***mC*mU*mG*mG***mU***mC*mA*mA*mC*mC*mA*mG***mU***mC*mA*mC*mA	PS-2'OMe miRNA antagomir, caged (Shratt collaboration)
miR-221 antagomir (full)	mG*mA*mA*mA*mC*mC*mC*mA*mG*mC*mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*mU	PS-2'OMe miRNA antagomir
miR-221 antagomir (18 mer)	mC*mC*mA*mG*mC*mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*mU	5' truncated miR-221 antagomir
miR-221 antagomir (13 mer)	mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*mU	5' truncated miR-221 antagomir
miR-221 antagomir (full) 3'Uaq	mG*mA*mA*mA*mC*mC*mC*mA*mG*mC*mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*Uaq	PS-2'OMe miRNA antagomir, 3'Uaq
miR-221 antagomir (18 mer) 3'Uaq	mC*mC*mA*mG*mC*mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*Uaq	5' truncated miR-221 antagomir, 3'Uaq
miR-221 antagomir (13 mer) 3'Uaq	mA*mG*mA*mC*mA*mA*mU*mG*mU*mA*mG*mC*Uaq	5' truncated miR-221 antagomir, 3'Uaq

Strand	Sequence (5' → 3')	Application
Chordin antisense (PO)	mAmUmCmCmAmCmAmGmCmAmGmCmCmCmCmUmCmCmAm UmCmAmUmCmC	2'OMe zebrafish antisense agent (Yoder collaboration)
Chordin antisense (PS)	mA*mU*mC*mC*mA*mC*mA*mG*mC*mA*mG*mC*mC*mC*mC *mU*mC*mC*mA*mU*mC*mA*mU*mC*mC	PS-2'OMe zebrafish antisense agent (Yoder collaboration)
Chordin antisense (PS)	mA***mU***mC*mC*mA*mC*mA*mG*mC*mA*mG*mC*mC*mC*m C***mU***mC*mC*mA***mU***mC*mA***mU***mC*mC	PS-2'OMe zebrafish antisense agent, caged (Yoder collaboration)
Pitx2 antisense	mGmGmUmAmCmAmGmUmAmCmAmGmUmAmGmGmCmUmCm AmCmAmGmAmCmA	2'OMe *Xenopus* antisense agent (Yoder collaboration)
Jnk1 antisense	mUmGmCmUmGmUmCmAmCmGmCmUmUmGmCmUmUmCmGm GmCmUmCmAmUmA	2'OMe *Xenopus* antisense agent (Yoder collaboration)
Wnt11 antisense	mCmCmAmGmUmGmAmCmGmGmGmUmCmGmGmAmGmCmCm AmUmUmGmGmUmA	2'OMe *Xenopus* antisense agent (Yoder collaboration)
Renilla luc antisense 5'Amino mod	AmMOmC*mG*mU*mU*mU*mC*mC*mU*mU*mU*mG*mU*mU* mC*mU*mG*mG*mA*	PS-2'OMe antisense, 5'Amino (Chakrabarti collaboration)
Firefly luc antisense 5 'amino mod	AmMOmU*mU*mC*mU*mU*mU mA*mU*mG*mU*mU*mU*mU*mU*mG*mG*mC*mG*	PS-2'OMe antisense, 5'Amino (Chakrabarti collaboration)
GFP antisense 5' amino mod	AmMOmU*mA*mG*mU*mU*mG*mA*mA*mC*mG*mC*mU*mU* mC*mC*mA*mU*mC*	PS-2'OMe antisense, 5'Amino (Chakrabarti collaboration)
Renilla luc antisense 3' amino mod	mC*mG*mU*mU*mU*mC*mC*mU*mU*mU*mG*mU*mU*mC*m U*mG*mG*mA*AmMO	PS-2'OMe antisense, 3'Amino (Chakrabarti collaboration)
Firefly luc antisense 3' amino mod	mU*mU*mC*mU*mU*mU*mA*mU*mG*mU*mU*mU*mU*mU*m G*mG*mC*mG*AmMO	PS-2'OMe antisense, 3'Amino (Chakrabarti collaboration)
GFP antisense 3' amino mod	mU*mA*mG*mU*mU*mG*mA*mA*mC*mG*mC*mU*mU*mC*m C*mA*mU*mC*AmMO	PS-2'OMe antisense, 3'Amino (Chakrabarti collaboration)
TATA box T[0] noncaged	TAAACCAGCTCTGCTTATATAGACCTCCAACCGT	TATA box
TATA box T[1] caged	TAAACCAGCTCTGCTTA**T**ATAGACCTCCAACCGT	TATA box, caged
TATA box T[2] caged	TAAACCAGCTCTGCT**T**ATA**T**AGACCTCCAACCGT	TATA box, caged
TATA box T[3] caged	TAAACCAGCTCTGCT**TATAT**AGACCTCCAACCGT	TATA box, caged
Dro TATA	CGCGTTTCCAAAATGTATAAAGAGTCACCG	Drosophila TATA box (Reeves collaboration)

Strand	Sequence (5' → 3')	Application
Dro TATA 2T caged	CGCGTTTCCAAAATG**TAT**AAAGAGTCACCG	Drosophila TATA box, caged (Reeves collaboration)
Dro TATA 3T caged	CGCGTTTCCAAAA**TGTAT**AAAGAGTCACCG	Drosophila TATA box, caged (Reeves collaboration)
DsRed antisense addition A1	ATGAAGGTGCCGT▽CCTGCAGGGAGGAG	DsRed antisense agent, cleavable
DsRed antisense deletion D1	ATGAAGGTGCCGT▽CTGCAGGGAGGAG	DsRed antisense agent, cleavable
DsRed antisense deletion D2	ATGAAGGTGCCG▽CTGCAGGGAGGAG	DsRed antisense agent, cleavable
DsRed antisense deletion D3	ATGAAGGT▽CCGTCCTG▽AGGGAGGAG	DsRed antisense agent, cleavable
EGFP antisense (PO)	GAGCTGCACGCTGCCGTC	EGFP antisense agent
EGFP antisense (PS)	G*A*G*C*T*G*C*A*C*G*C*T*G*C*C*G*T*C*	EGFP antisense agent
EGFP antisense addition A1	GAGCTGCAC▽GCTGCCGTC	EGFP antisense agent, cleavable
EGFP antisense deletion D1	GAGCTGCAC▽CTGCCGTC	EGFP antisense agent, cleavable
EGFP antisense deletion D2	GAGCTGCA▽CTGCCGTC	EGFP antisense agent, cleavable
EGFP antisense (PO) deletion D3	GAGCTG▽ACGCT▽CCGTC	EGFP antisense agent, cleavable
EGFP antisense (PS) deletion D3	G*A*G*C*T*G▽A*C*G*C*T▽C*C*G*T*C	EGFP antisense agent, cleavable
Catalyst	CATTCAATACCCTACGTCTCCA	Catalytic fuel cycle
Catalyst 3A caged	C**A**TTCA**A**TACCCT**A**CGTCTCCA	Catalytic fuel cycle, caged
Catalyst 4A caged	C**A**TTC**A**AT**A**CCCT**A**CGTCTCCA	Catalytic fuel cycle, caged
Catalyst C1 cleavable	CATTCAATAC□CTACGTCTCCA	Catalytic fuel cycle, cleavable

Strand	Sequence (5' → 3')	Application
Catalyst C2 cleavable	CATTCAA☐ACCCTA☐GTCTCCA	Catalytic fuel cycle, cleavable
Inhibitor	TGGAGACGTAGGGTATTGAATG	Catalytic fuel cycle
Inhibitor (10)	TGGAGACGTA	Catalytic fuel cycle
Inhibitor 3G caged	TG**G**A**G**AC**G**TA	Catalytic fuel cycle, caged
Inhibitor 4G caged	T**GG**A**G**AC**G**TA	Catalytic fuel cycle, caged
Inhibitor 4A caged	TGG**A**GACGT**A**GGGT**A**TTGA**A**TG	Catalytic fuel cycle, caged
Output 4T caged	CTTTCC**T**ACACCTACG**T**CTCCAAC**T**AACT**T**ACGG	Catalytic fuel cycle, caged
A input 1T caged	TA**T**GGTTGTTTATGTGTTCCCTGATCTTTAGCCTTA	AND gate caged input, caged
A input 2T caged	TA**T**GG**T**TGTTTATGTGTTCCCTGATCTTTAGCCTTA	AND gate caged input, caged
A input 3T caged	TA**T**GG**T**TGTTTATG**T**GTTCCCTGATCTTTAGCCTTA	AND gate caged input, caged
A input 4Tcaged	TATGG**T**TGTTTATG**T**GTTCCC**T**GATCTT**T**AGCCTTA	AND gate caged input, caged
G$_T$ 1T caged	GGCTAAAGATCAGGGAACACATAAACAACCA**T**A	AND gate caged gate toehold, caged
B input	GTTAGATGTTAGTTTCACGAAGACAATGATTAAGGC	AND gate input
B input 4T caged	GTTAGA**T**GTTAGTT**T**CACGAAGACAA**T**GAT**T**AAGGC	AND gate caged input, caged
Toehold (6)	TATATA	AND gate
B$_{21}$ input	GTTAGATGTAGTTCACGAAGACAATGATCAACA	miR-21 AND gate input
B$_{21}$ input 4T caged	GT**T**AGATG**T**AGT**T**CACGAAGACAATGA**T**CAACA	miR-21 AND gate input, caged
B$_{21}$ input 4G caged	GTTAGAT**G**TA**G**TTCAC**G**AAGACAAT**G**ATCAACA	miR-21 AND gate input, caged
HCR Initiator	AGTCTAGGATTCGGCGTGTATATA	Hybridization chain reaction

Strand	Sequence (5' → 3')	Application
HCR Initiator 3T caged	AGTCTAGGATTCGGCGTG**TATAT**A	HCR toehold caged initiator
HCR Initiator 4T caged	AGTC**T**AGGA**T**TCGGCG**T**GTATA**T**A	HCR evenly caged initiator
Strand 1	ACATTCCTAAGTCTGAACATTACAGCTTGCTACACAGAAGAGCCGCCATAGTA	DNA tetrahedron assembly
Strand 2	TATCACCAGGCAGTTGAAGTGTAGCAAGCTGTAATGATGCGAGGGTCCAATAC	DNA tetrahedron assembly
Strand 3	TCAACTGCCTGGTGATAAACGACACTACGTGGGAACTACTATGGCGGCTCTTC	DNA tetrahedron assembly
Strand 4	TTCAGACTTAGGAATGTCTTCCCACGTAGTGTCGTTGTATTGGACCCTCGCAT	DNA tetrahedron assembly
Strand 1 2-cleavable	ACATTCCTAAGTCTGAA☐CATTACAGCTTGCTACAC☐AGAAGAGCCGCCATAGTA	DNA tetrahedron assembly, cleavable
Strand 2 2-cleavable	TATCACCAGGCAGTTGA☐AGTGTAGCAAGCTGTAAT☐GATGCGAGGGTCCAATAC	DNA tetrahedron assembly, cleavable
Strand 3 2-cleavable	TCAACTGCCTGGTGATA☐AACGACACTACGTGGGAA☐CTACTATGGCGGCTCTTC	DNA tetrahedron assembly, cleavable
Strand 4 2-cleavable	TTCAGACTTAGGAATGT☐CTTCCCACGTAGTGTCGT☐TGTATTGGACCCTCGCAT	DNA tetrahedron assembly, cleavable
SSO #623	mG*mU*mU*mA*mU*mU*mC*mU*mU*mU*mA*mG*mA*mA*mU*mG*mG*mU*mG*mC	PS-2'OMe splice-switching
SSO 4dT	mG*dT*mU*mA*mU*dT*mC*mU*mU*dT*mA*mG*mA*mA*dT*mG*mG*mU*mG*mC	PS-2'OMe splice-switching mixmer
SSO 2U caged	mG*mU*mU*mA*mU*mU*mC***mU***mU*mU*mA*mG*mA*mA***mU***mG*mG*mU*mG*mC	PS-2'OMe splice-switching, caged
SSO 4U caged	mG***mU***mU*mA*mU***mU***mC*mU*mU***mU***mA*mG*mA*mA***mU***mG*mG*mU*mG*mC	PS-2'OMe splice-switching, caged
SSO 1-cleavable	mG*mU*mU*mA*mU*mU*mC*mU*mU*☐mA*mG*mA*mA*mU*mG*mG*mU*mG*mC	PS-2'OMe splice-switching, cleavable
SSO 2-cleavable	mG*mU*mU*mA*mU*☐mC*mU*mU*mU*mA*mG*mA*☐mU*mG*mG*mU*mG*mC	PS-2'OMe splice-switching, cleavable
Sox31 SSO	mA*mG*mC*mC*mC*mU*mU*mU*mU*mC*mU*mC*mA*mA*mA*mA*mC*mA*mA*mA*mC*mC*mU*mG*mU	PS-2'OMe splice-switching zebrafish

Strand	Sequence (5' → 3')	Application
Sox31 SSO 4U caged	mA*mG*mC*mC*mC***mU***mU*mU***mU***mC***mU***mC*mA*mA*mA*mA*mC*mA*mA*mA*mC*mC***mU***mG*mU	PS-2'OMe splice-switching zebrafish, caged
SSO gate	mU*mU*mC*mU*mA*mA*mA*mG*mA*mA*mU*mA*mA*mC*mU*mA*mG*mC*mU*mU*mA*mU*mC*mA*mG*mA*mC*mU*mG*mA	PS-2'OMe miR-21 SSO AND gate
SSO input B	mA*mA*mG*mC*mU*mA*mG*mU*mU*mA*mU*mU*mC*mU*mU*mU*mA*mG*mA*mA	PS-2'OMe miR-21 SSO AND gate
SSO toehold	mU*mC*mA*mA*mC*mA*mU*mC*mA*mG*mU*mC*mU*mG*mA*mU*mA*mA*mG*mC*mU*mA	PS-2'OMe miR-21 SSO AND gate
ZFN Comp 3T caged	CAC**T**GCGGC**T**CCGGCCCCG**T**C	Zinc finger nuclease binding domains, caged
ZFN HP 4T caged	GACGGGGCCGGAGCCGCAGTGGATGTAGGGAAAAGCCCGGCCCTTTTGGGCCGGGCT**T**TTCCC**T**ACATCCAC**T**GCGGC**T**CCGGCCCCGTC	Zinc finger nuclease binding domains, caged
ZFN PHP 3T caged	GACGGGGCCGGAGCCGCAGTGGATGTAGGGAAAAGCCCGGGGGTTTTCCCCCGGGCT**T**TTCCC**T**ACA**T**C	Zinc finger nuclease binding domains, caged
T7 template	GGGAGAATTCAACTGCCATCTAGGCACAGGTAAATTAATTAGAGATGAAATAGAGTGGAGTGAGGGATTAGTTTTGGAGTACTCGCT	T7RNAp PCR template

5.2   Oligonucleotide Synthesis

Standard synthesis cycles were used for all oligonucleotide production on either a refurbished Applied Biosystems Model 394 automated DNA/RNA synthesizer (Life Technologies) or a new MerMade4 synthesizer (Bioautomation) using β-cyanoethyl phosphoramidite chemistry. The synthesizers required an argon tank (4.8, 300 SZ, Matheson, PARN30048) at >500 PSI to perform a complete synthesis cycle. A high-quality (>99.9%), low-water (<50 ppm) source of anhydrous acetonitrile was also required (VWR, EM-AX0152-1). When the synthesizer had not been used in several weeks, a startup cleaning cycle performed block and line flushes with acetonitrile then argon, followed by filling the lines with fresh reagents. The startup cycle on the

ABI Model 394 was performed automatically with the "ABI Begin" function, while the MerMade4 startup cycle was performed manually. All reagents for oligonucleotide synthesis were purchased from Glen Research (Table 5.2).

Table 5.2: Glen Research reagents for oligonucleotide synthesis.

Reagent	Catalog #	Amount
Activator	30-3100-52	200 mL
Cap Mix A	40-4110-52	200 mL
Cap Mix B	40-4220-52	200 mL
Oxidizing Solution	40-4330-52	200 mL
Deblocking Mix	40-4140-57	450 mL
Diluent	40-4050-50	100 mL
dA-CE phosphoramidite	10-1000-10	1g
dC-CE phosphoramidite	10-1010-10	1g
dG-CE phosphoramidite	10-1020-10	1g
dT-CE phosphoramidite	10-1030-10	1g
DNA 1000Å 40 nM supports (ex. A)	20-2101-45	4 pack
2'OMe-A-CE phosphoramidite	10-3100-10	1 g
2'OMe-C-CE phosphoramidite	10-3110-10	1 g
2'OMe-G-CE phosphoramidite	10-3121-10	1 g
2'OMe-U-CE phosphoramidite	10-3130-10	1 g
2'OMe 1000Å 0.2 μM supports (ex. A)	20-3700-42	4 pack
Beaucage sulfurizing reagent	40-4036-10	1 g
5'Amino C6 TFA phosphoramidite	10-1916-02	0.25 g
3'Amino PT C6 CPG support	20-2956-42	pack of 4

Reagent bottles were connected to the synthesizer in a hood using the bottle change function. Each reagent has a specific shelf date and an estimated usage volume provided by Glen Research, as predicted by the synthesis scale. The amidites were manually dissolved in anhydrous acetonitrile to 0.1 M according to the manufacturer's protocol. Specialized caged

234

phosphoramidites were dissolved in anhydrous acetonitrile to a final concentration of 0.05 M to reduce reagent usage. DNA oligonucleotides were synthesized at the 40 nM scale with 1000Å solid-phase CPG support columns. 2'OMe phosphorothioate oligonucleotides were synthesized at the 0.2 µM scale with solid-phase CPG support columns. All solid-phase supports contained the 3' terminal base. Synthesis cycles were provided by Applied Biosystems and Bioautomation. Coupling times of 25 sec were used for DNA monomers. Synthesis of modified oligonucleotides was also performed using β-cyanoethyl phosphoramidite chemistry but with increased coupling times according to the manufacturer's instructions, such as 6 min coupling times for 2'OMe monomers. The amino modifications (3' and 5' AmMo) did not require any alterations the standard synthesis conditions. The coupling time required for caged phosphoramidites was typically 10 min. The sulfurization steps were performed by replacing the oxidizing reagent with the Beaucage sulfurizing reagent (3*H*-1,2-benzodithiole-3-one-1,1-dioxide) at 0.05 M in acetonitrile for 15 min. All cycle alterations were made following the user manual instructions. Each synthesis cycle was monitored visually by observing the release of DMTr cations after the first- and last-two deprotection steps, which produced a bright orange solution from the trityl collection outlet. The end procedure included removal of the last DMTr protecting group (DMT OFF) and automated column elution (CE). At the completion of a synthesis cycle, the oligonucleotides were automatically eluted from the solid-phase supports with ~1 mL ammonium hydroxide methylamine (AMA, 1:1) and then deprotected at 65 °C for 2 h in an oil bath. The full-length oligonucleotides were then purified with GE Healthcare illustra Nap-10 Sepahdex™ G-25 resin columns (VWR, 95017-011). The columns were rinsed 3 times with 1 mL of water, the entire deprotection solution was applied to the column, and a 1 mL water elution was collected. Oligonucleotides are all stored in solution at −20 °C. UV absorbance was

used to confirm the presence of nucleic acids through the absorption at 260 nm (NanoDrop) and calculate molar concentrations with the oligonucleotide extinction coefficient (IDT OligoAnalyzer 3.1). Gel analysis provided a method to confirm the size as well as to calculate the accurate concentration for the full-length oligonucleotide through a dilution series and comparison to a control standard. Typically, 10 pmol (10 μL of 1 μM) of each oligonucleotide was loaded onto a 20% TBE-PAGE gel (45 min, 200 V). Run times and gel percentages may vary based on the oligonucleotide size. Staining was then performed with Sybr Gold nucleic acid stain (Life Technologies, 16500100). The gel was soaked for 15 min in 1:10,000 Sybr Gold diluted in 50 mL TBE, then rinsed with water and imaged on a fluorescent gel imager (Bio-Rad ChemiDoc Sybr Gold settings; standard filter 1, UV transillumination). Purity of the final product was determined by comparison of the full-length product in relation to the total yield through normalization of a fluorescently stained gel. In cases where the purity was <90%, typically through contaminating truncation products, the oligonucleotides were further purified with a 20% TBE-PAGE gel through UV shadowing using a handheld UV lamp against a TLC plate while shielding the caged oligos, band excision of the full-length product, and elution (overnight at 37 °C or 2 h at 65 °C) in 500 μL water. Multiple elution steps have been shown to increase the total yield by 50%, since passive elution is dependent on a concentration equilibrium between the gel slice and the buffer system. The eluted oligonucleotides solutions were spun down to remove any debris, such as contaminating gel pieces, and were analyzed through UV absorption followed by additional PAGE analysis. Oligonucleotides were concentrated through water evaporation (SpeedVac) or ethanol precipitation (separation of nucleic acids from solution, resuspension in decreased volume) when necessary.

5.3   General Molecular Biology Techniques

5.3.1   Polymerase Chain Reaction (PCR)

Several kits were used for different PCR requirements (all purchased from New England Biolabs). *Taq* DNA polymerase was used for general PCR where a low mutation rate is not required, such as PCR screening or amplification of a short insert (<100 bp). However, for gene amplification (e.g., from plasmids), a high fidelity Phusion polymerase was used to limit the introduction of mutations. The size and reaction inputs for PCRs varied based on the application requirements (Table 5.3). Reactions were performed in commercial polymerase buffers with 200 μM dNTPs and ~2 μL enzyme per 200 μL reaction. Primers were added at 0.5-1 mM.

Table 5.3: PCR reaction scales and input amounts based on application.

Application	Input	Scale
Colony Screen	colony	15 μL
Gene amplification	Plasmid (5 ng)	200 μL
T7RNAp template	ssDNA (20 pmol)	1 mL

The cycle times and temperature for each PCR step (denature, anneal, extend) were performed according to the manufacturer's provided protocols and the melting temperatures of each oligonucleotide primer (IDT). Anneal times varied depending on the base content and size of the primers for extension, while extension times varied depending on the total size of the PCR product (~30 sec per kilobase). PCR products were column-purified with PCR cleanup kits, such as the QIAQuick (Qiagen) or E.Z.N.A. Cycle Pure (Omega) kits. The PCR product was typically eluted in 50 μL of 65 °C water, but smaller volumes were sometimes used to increase the concentration.

5.3.2 Site-Directed Mutagenesis (SDM)

SDM was first performed with the QuikChange Lightning kit (Agilent) following the manufacturer's protocol. Primers for the QuikChange mutagenesis reactions were designed to be fully complementary, with the desired mutations placed approximately in the middle on both primer strands, and ~15-20 bases that bind to the plasmid on either side, for a total size of ~35-45 bases as well as $T_MS > 78$ °C. Alternatively, an overlap extension SDM method was utilized,[299] in which the extension was performed with high fidelity Phusion polymerase (New England Biolabs). Primers for the overlap SDM reactions were designed to be partially complimentary for ~15 bases, again with the desired mutations placed approximately in the middle on both primer strands, but in this case there were ~20 bp on either end of the primers that are only complementary to the plasmid, and the T_M of the duplex region should be at least 10 °C less than the full-length primer T_MS.

For both methods, polymerization reactions were performed in 25 μL reactions with 1 μM primers and 5-10 ng of plasmid DNA. The annealing and extension temperatures were altered depending on the primer T_MS and the size of the gene product or plasmid. After the reactions were completed, the SDM reactions were treated with 0.4 μL *Dpn1* (New England Biolabs) at 37 °C for 1 h to digest the methylated parent plasmid isolated from *E. coli*. The SDM reactions were then transformed into chemically competent bacteria (see protocol 5.3.5) and analyzed for the proper base mutations through sequencing.

5.3.3   Restriction Enzyme Digestion

Sequence-specific restriction enzymes (REs) were ordered from New England Biolabs and buffer components were provided by the manufacturer. RE reactions were typically performed for 2 h at 37 °C, unless the enzyme requires a different temperature for activity, and the enzyme was then heat inactivated at 80 °C for 20 min. The reaction volume varied based on the application requirements. For plasmid mapping, 10 μL reactions with 200 ng of the plasmid were performed with 0.2 μL of each RE. The entire reaction mixture was subsequently analyzed by agarose-gel electrophoresis. For gene insert or plasmid backbone digestion, 200 μL reactions with 8 μg of the plasmid were performed with 4 μL of each RE. Plasmid backbones that were utilized as vectors for ligation were then treated with 4 μL of alkaline phosphatase (New England Biolabs) for 2 h at 37 °C to remove the phosphate groups and reduce self-circularization during ligation reactions (see protocol 5.3.4). The resulting digested products (inserts or backbones) were gel purified on 0.8% agarose gels through band excision and isolation with the QIAQuick (Qiagen) or E.Z.N.A. (Omega) gel extraction kits, followed by elution with 30 μL of 65 °C water. The DNA concentrations (ng / μL) were determined with a NanoDrop spectrophotometer.

5.3.4   Plasmid Ligation.

The backbone and insert gene products were produced via PCR, followed by RE digest and gel purification, as described above (protocols 5.3.1 and 5.3.3). When synthetic oligonucleotides were used as an insert and were not previously digested by an RE, an intermediate step was required prior to ligation to add a 5′ phosphate group with T4 polynucleotide kinase (PNK, New England Biolabs). The T4-PNK reaction was performed with 10 μM DNA in 20 μL with 1 μL

T4-PNK for 1 h at 37 °C in commercial T4-PNK buffer containing ATP, according to the manufacturer's protocol. The T4-PNK treated oligonucleotides were then combined (10 μM in 40 μL) and annealed (95 °C to 12 °C over 10 min) in TAE/Mg^{2+} buffer (0.04 M tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), and 12.5 mM magnesium acetate). Ligation reactions were then performed at a 10 μL scale, most commonly with a 1:6 vector:insert ratio, using 100 pg of the vector backbone and either Quick ligase or T4 DNA ligase (New England Biolabs). The quick ligations were performed for 15 min at room temperature, while the T4 DNA ligations were performed overnight at 4 °C. Negative control reactions were performed with the appropriate amount of water without any insert. Additionally, quick ligations were subsequently spiked with T4 DNA ligase and stored overnight at 4 °C to potentially increase the efficiency of the ligation protocol; both ligations were transformed into bacterial cells (see protocol 5.3.5).

5.3.5   Plasmid *E. coli* Transformation

Competent cells for plasmid transformation are either included in commercial DNA manipulation kits or are generated separately. For all cloning in this thesis the *E. coli* strains Top10, BL21, and NovaBlue were used. The generation of chemically competent cells was performed with streaking out cell stocks on an LB-agar plates and incubating overnight at 37 °C. A single colony was then used to inoculate 5 mL of LB-broth which was incubated shaking overnight at 37 °C. The overnight culture (1 mL) was then transferred into 50 mL of LB-broth and incubated shaking at 37 °C until the OD$_{600}$ reached 0.3-0.5, usually within 2-3 h. The culture was then chilled on ice for 20 min and cells were pelleted by centrifugation (16,000 g, 5 min, 4 °C). The pellet was resuspended in an ice-cold transformation solution (85% LB-broth, 10%

PEG wt/vol, 5% DMSO vol/vol, 50 mM $MgCl_2$, pH 6.5) with a calculated volume of OD_{600} (determined above from the 50 mL culture) X 10 mL. The suspension was then aliquoted into 1.5 mL tubes cooled to -80 ^{o}C (50-100 µL) and placed in a dry ice ethanol bath before long-term storage at -80 ^{o}C for. The competency of the cells was calculated through transformation of a pUC19 control plasmid DNA (colony count X (transformation volume µL / plating volume µL) / plasmid DNA µg), and should be greater than 10^6 colonies per µg of plasmid DNA. The transformations were typically performed with 50 µL competent cell stocks, which were briefly thawed on ice. The DNA of interest was then added to the competent cells, mixed gently through pipetting up and down once or twice, and stored on ice for 20 min. For plasmids, 1-5 µL of ~100 ng total DNA was added. For SDM and ligation reactions, a range of 1-10 µL was added. When high transformation efficiency was expected (plasmid transformations), the amount of competent cells could be reduced to 25 µL per reaction. After the 20 min on ice, the cell-DNA mixture was heat shocked in a 42 ^{o}C water bath for 30 sec, and placed back on ice for an additional 2 min. SOC media was then added (100-450 µL) and grown in the absence of antibiotic for 1 hr shaking at 37 ^{o}C. The cultures were then plated onto LB-agar plates that were prepared with antibiotic, depending upon the resistance for the plasmid of interest. To prepare the culture plates, LB-agar was heated in the microwave for 2 min and stock antibiotic solutions were diluted into an appropriate volume. Frequently used antibiotics include ampicillin (50 mg/mL stock in water, 50 µg/mL final, 1:1000), kanamycin (50 mg/mL stock in water, 50 µg/mL final, 1:1000), and tetracycline (5 mg/mL stock in ethanol, 25 µg/mL final, 1:200). For co-transformation of two plasmids, multiple antibiotics can be used on a single plate. The LB-agar was poured into 10 cm plates (~10 mL) and cooled on the bench until solidified. Plating of different amounts was performed until the entire SOC culture was plated, e.g., from a 500 µL culture, 100 and 400 µL

241

would be spread onto two separate agar plates. The cultures were spread, typically with a glass rod that was dipped in ethanol and flame sterilized, while spinning the plate. LB-agar plate cultures were incubated overnight at 37 ^{o}C, at which time individual colonies were picked and inoculated into 5 mL cultures. All further cultures were grown in the presence of the corresponding antibiotic. Glycerol stocks of the transformed cells for all plasmid bacterial cultures were produced by mixing saturated LB-broth culture with glycerol (1:1 to a final volume of 1 mL) and storing at −80 ^{o}C.

5.3.6   Plasmid Purification

Plasmids were purified from bacterial cultures grown to saturation in LB-broth containing the appropriate antibiotics. Cultures of 5 mL and 100 mL were used for plasmid production. Smaller cultures were purified using spin column miniprep kits, such as QIAprep (Qiagen) or E.Z.N.A. Plasmid (Omega) kits. These purifications were eluted with 50 µL of 65 ^{o}C water. Larger cultures were purified using midiprep kits (same manufacturers), and eluted with 500 µL of 65 ^{o}C water. The plasmid concentrations (ng/µL) were determined with a NanoDrop spectrophotometer.

5.3.7   Plasmid Analysis

Plasmids were analyzed through a number of different methods. PCR-based screens were directly performed on bacterial colonies with 15 µL reactions (see protocol 5.3.1). The supercoiled plasmids were also analyzed on a 0.8% agarose gel with 200 ng of the plasmid and ethidium bromide DNA staining. RE digests of the plasmids (see protocol 5.3.3) were used to

confirm insert sizes with 10 μL reactions of 200 ng plasmid and 0.2 μL of each RE. These reactions were incubated at 37 °C for 2 hr and analyzed on a 0.8% agarose gel. All plasmid constructs were confirmed through sequencing (Genewiz or University of Pittsburgh Genomics and Proteomics Core Laboratories) with 15 μL solutions of 1 μg DNA and 20 pmol of primer for the region of interest.

5.4   General Cell Culture Techniques

5.4.1   Cell Growth and Maintenance

All cell stocks were maintained at 37 °C and 5% CO_2 in 10 cm culture plates, and manipulations were performed in a biosafety cabinet. The majority of the cell lines (including HEK293T, HeLa, HeLa:EGFP654, HeLa:Luc705, HepG2, Huh7) used in this research were grown in Dulbecco's modified Eagle's medium (DMEM, Hyclone), supplemented with 10% fetal bovine serum (FBS, Hyclone) and 2% penicillin/streptomycin (PS, MP Biomedicals). The DMEM solution was made by dissolving 5.89 g DMEM powder, 1.63 g sodium bicarbonate, and 0.145 g L-glutamine in 445 mL of Milli-Q purified water (EDM Millipore) that was autoclaved. The pH was then adjusted to ~7.4 with 65% hydrochloric acid (~10-15 drops). Subsequently, 50 mL of FBS and 5 mL of a 100X PS solution were added and the media was filter sterilized. The PC-3 cell line was grown in F-12K medium (Hyclone), which was prepared from a commercial reagent according to the manufacturer's protocol. A fresh media exchange (10 mL) was performed every 2-3 days for each plate (10 cm). The cell lines were passaged when confluence reached >90% by removing the media and adding 1 mL of TrypLE Express reagent (Invitrogen), then shaking at room

temperature for ~5 min. Cells were lifted from the plate through pipetting and 9 mL of growth media was added. The passages were then performed in a 1:10 dilution by adding 1 mL of the cell resuspension to 9 mL of media in a new 10 cm plate. For passages into multi-well plates, the following volumes were used: 200 μL per well in a 96-well plate (~10^4 cells), 500 μL per well (~2.5×10^4 cells) in a 8-well chamber slide, 1 mL per well (~5×10^4 cells) in a 4-well chamber slide, and 2 mL per well (~10^5 cells) in a 6-well plate. The specific cell dilution and number varied based on the cell line or confluency at passage, and cells were grown for 1-2 days prior to transfection. White plates were used for chemiluminescence experiments, while black plates were used for fluorescence experiments. Plates were purchased from BD-Falcon, Greiner, or CoStar (Fisher, VWR). Passages of cell stocks were tracked by date, cell line, media requirements, and passage number. Cells were discarded once the passage number went above 25, and new cells were thawed. To thaw cells, the frozen cell vials were placed in a 37 °C water bath for 5 min, and then added to plates containing growth media. The media was changed after 1 day to remove residual DMSO. Frozen stocks of cell lines were made by passaging ~5 mL of cell resuspension into a 25 cm plate (25 mL total), growing the cells to confluency (5-7 days), lifting the cells with 5 mL trypsin, pelleting the cells at 3,000 g for 10 min, and resuspending the cells in 5 mL growth media containing 5% DMSO. The cell suspension was then aliquoted into 1 mL cryogenic preservation vials (~10^6 cells) and placed into a –80 °C overnight freezer before transfer to a liquid nitrogen dewar.

5.4.2   Cell Transfection

Cell transfections were performed to introduce foreign DNA into the cells, such as an expression plasmid or synthetic oligonucleotide. The cellular growth media was removed and replaced with

75% total volume of Opti-MEM media (Invitrogen) per well (150 µL for 96-well, 750 µL for 6-well). Transfection mixes were then prepared by mixing the nucleic acids in Opti-MEM media (50 µL per well for 96-well plate, 250 µL per well for 6-well plate) and incubating with the appropriate transfection reagent. Plasmid transfections were typically performed with 200 ng each plasmid in a 96-well plate (chemiluminescence or fluorescence readouts) and 2,000 ng of each plasmid in a 6-well plate (protein or RNA isolations). These transfections were performed with either branched polyethylene imine (bPEI), linear polyethylene imine (lPEI), or Lipofectamine 2000 (Invitrogen) at 1-2 µL per well for 96-well plates and 5-10 µL per well for 6-well plates, according to the manufacturers' protocols. The PEI reagents were used for overnight transfections, while the lipofectamine reagent was used for 4 h transfections (preferred). The PEI transfections could also be performed with PS-free DMEM in replacement of Opti-MEM. Oligonucleotide transfection concentrations varied based on the application, and are noted in each experimental section. Typically, 20-200 pmol were used for genetic regulators (TFOs, antagomirs, siRNA duplexes) and 50-200 mM were used for logic gates. Oligonucleotide transfections were performed with the X-tremeGENE siRNA reagent (Roche) for 4 h with 1-2 µL reagent per well for 96-well plates and 5-10 µL reagent per well for 6-well plates. The transfection reagents were mixed with the plasmid(s) or oligonucleotide(s), then incubated for 20 min at room temperature, and added to the cells for the appropriate amount of time. The media was removed at the completion of the transfection and replaced with fresh growth media. At this step, the photocaged lysine amino acid (PCK) was added at 1-2 mM concentration for expression of caged proteins. The protocols presented above should be used as points of reference, with the understanding that they can be modulated based on the experiment of interest. Every new experiment requires some degrees of optimization, including analysis of transfection

reagents/amounts, alterations in nucleic acid concentrations, applications in different cell lines, and different cell culture plate sizes.

5.4.3 Cell Irradiations

The photochemical control of proteins and oligonucleotides was commonly performed in cell culture with UV exposure. Whole well irradiations were performed directly on top of a UV transilluminator (25 W, 6.3 mW/cm^2) for various exposure times, typically 2 min were found as optimal unless otherwise noted. No temperature control was performed. Spatially distinct UV irradiations were performed with the UV transilluminator through precut designs in tinfoil sheets, which were taped onto the underside of the cell culture plates and left on during incubation periods to locate the area of interest. Additionally, localized irradiations were performed with a Zeiss Observer Z1 microscope (40X objective, NA 0.75 plan-apochromat; Zeiss) and a DAPI filter set (68 HE ex. BP377/28) with a partially closed aperture to irradiate only a specific subset of cells, with the x:y location noted to locate the area of interest.

5.4.4 Cell Fixing

Fixed cells were prepared in Nunc Lab-Tek II 4-well chamber slides (Thermo Scientific), which can be stained and stored for imaging. Cells were transfected and incubated for the required period of time, then washed twice with 500 μL PBS pH 7.4 at 4 °C. The cells were then fixed on ice with 500 μL of 3.75% formaldehyde in PBS for 15 min. After the fixing step, the cells were washed with PBS three times, and permeabilized with 200 μL of 0.5% TritonX100 for 30 sec at room temperature. The cells were again washed and stained with various reagents, including

fluorescent protein-specific antibodies (specific manufacturer's protocol), rhodamine-phalloidin actin staining (7 μL per mL of PBS containing 1% of BSA, 200 μL per well, 20 min room temperature, Life Technologies), and DAPI nuclear staining (15 μL of a 14.3 mM solution per mL PBS, 200 μL per well, 2 min room temperature, Invitrogen). After staining procedures, the cells were washed three times with 200 μL of PBS, the chamber slide top was removed, and the mount was air-dried. Prolong Gold antifade reagent (Life Technologies) was then gently added dropwise, and a coverslip was placed on the top to dry overnight.

5.4.5 Cell Imaging

Before imaging, growth media was replaced with 37 °C phenol red-free DMEM-high modified growth media (Thermo Scientific). The following filter cubes were used on a Zeiss Observer Z1 microscope: DAPI (filter set 68 HE; ex. BP377/28; em. BP464/100), GFP/EGFP/FITC (filter set 38 HE; ex. BP470/40; em. BP525/50), and DsRed/mCherry/Rhodamine/TAMRA (filter set 43 HE; ex. BP575/25; em. BP605/70). The following Zeiss objectives were used: 5X (NA 0.16 plan-neofluora), 10X (NA 0.25 plan-apochromat), 20X (NA 0.8 plan-apochromat) and 63X oil (NA 1.4 plan-apochromat). Cellular images were taken in brightfield as well as fluorescent channels, and then processed in Zen Pro 2011 imaging software. Additional images were taken on a Nikon A1 confocal microscope with the assistance of Nikon representatives. Fluorescent cell counting was performed with ImageJ software (settings: threshold 5-10%, size >200 pixels^2, and circularity 0-1).

5.4.6 RNA Isolation

Total RNA isolations from $\sim\!10^6$ cells were performed using Trizol reagent (Invitrogen) in 6-well plates following the manufacturer's protocol. The cellular media was removed, cells were rinsed with 4 °C PBS pH 7.4, and 1 mL of Trizol reagent was added. The cells were then lifted via pipetting, incubated for 5 min at room temperature, and collected with centrifugation at 16,000 g for 10 min at 4 °C. Cholorform was added (200 μL), mixed by shaking, and incubated for 5 min at room temperature. The reaction was then phase separated by centrifugation at 16,000 g for 10 min at 4 °C, and the upper aqueous layer containing the RNA (~50% volume) was transferred to a new tube. The RNA was precipitated from the reaction through addition of 600 μL of isopropanol, stored at –20 °C for 30 min, and centrifuged at 16,000 g for 20 min at 4 °C. The RNA pellets were then dried with open caps for 30 min, resuspended in ~20 μL nuclease free water, and the final concentrations (ng/μL) were determined with a NanoDrop spectrophotometer.

5.4.7 RNA Analysis by Quantitative Real-Time PCR (qRT-PCR)

The first step of qRT-PCR was the reverse transcription of the RNA into DNA with the iScript cDNA Synthesis kit (Bio-Rad), using 20 ng of the RNA isolations (2 μL of 10 ng/μL dilutions) in 40 μL reactions with 1 μL of the reverse transcriptase enzyme, according to the manufacturer's protocol. The RT reactions were heated to 25 °C for 5 min, then 42 °C for 45 min, and finally 85 °C for 5 min before being cooled on ice. Quantitative PCR was then performed with the SsoFast Evagreen Supermix (Bio-Rad) in 20 μL reactions using gene-specific primers, according to the manufacturer's protocol. The cDNA product from the RT

reaction (5 μL) was added to the qPCR reactions containing 10 μL of the 2X SSoFast reagent and 10 μM primers, one set for the target gene and one set for the control gene (e.g., GAPDH) in separate reactions (each in triplicate). Primer sequences are shown in each project-specific experimental section. The qPCRs were heated at 95 °C for 3 min then cycled 40 times from 95 °C for 10 sec to 55 °C for 30 sec. Amplification of the PCR product was monitored with Sybr Green fluorescence using a Bio-Rad CFX96 Real-Time System. The Ct values and relative gene expressions $(\Delta(\Delta Ct))$[300] were determined with the Bio-Rad CFX Manager 3.1 software. Technical triplicates of the qPCRs were used to determine normalized gene expression relative to the GAPDH controls.

5.4.8 Protein Isolation from Mammalian Cells

Total protein isolation from ~10^6 cells in a single well of a 6-well plate was used to perform western blot analysis of caged protein expression or purification of protein from mammalian cell expressions. The cells were rinsed with 1 mL 4 °C PBS pH 7.4, 200 μL of GE Healthcare mammalian cell lysis buffer (Sigma-Aldrich) was added, and cells were incubated while shaking on ice for 10 min. The cells were then lifted via pipetting and centrifuged at 16,000 g for 20 min at 4 °C. The lysis solution was removed from the cell debris. HA-tag purification was performed on total protein lysate with 20 μL of mouse-anti-HA antibody (Santa Cruz Biotechnology) and 100 μL of Protein A Sepharose 4B suspension (Life Technologies) overnight at 4 °C in a refrigerator (preferably shaking or rocking). The sepharose beads were then washed three times with 300 μL 4 °C PBS by pelleting the beads through centrifugation (3,000 g for 2 min at 4 °C) and resuspended to 100 μL total volume with PBS.

5.4.9   Protein Analysis

Standard western blot techniques were used for analysis of proteins. Total protein extract (40 µL) was denatured with protein loading dye at 75 °C for 20 min. The denatured protein solution was then size separated on a 1.5 mM 10% SDS-PAGE gel, which was subsequently soaked in transfer solution (25 mM Tris-base, 192 mM glycine, 20% methanol, 0.025% wt/vol SDS). The protein gel was then transferred to a GE Healthcare Hybond-LFP PVDF membrane (Fisher) at 80 V for 90 min on ice in a Bio-Rad Mini Trans-Blot module. Blocking of the membrane was performed with 5% milk in TBST buffer (10 mM Tris-base, 150 mM NaCl, pH 7.6, 10% wt/vol Tween 20) at room temperature for 1 h. Primary antibody (1:1000 in 15 mL TBST) binding was performed at 4 °C overnight, followed by several rinses and secondary antibody (1:1000 in 15 mL TBST, depending on the manufacturer's suggestions) binding for 1 h at room temperature. All specific antibodies are indicated in the methods and materials sections. Chemiluminescent detection was performed with the VisiGlo kit (Amresco) and imaged on a ChemiDoc MP (Bio-Rad), according to the manufacturer's protocol.

5.5   General Zebrafish Techniques

5.5.1   Zebrafish Mating

All zebrafish experiments were performed with the University of Pittsburgh's Institutional Animal Care and Use Committee approval. The Oregon AB* strain was maintained under standard conditions at the University of Pittsburgh School of Medicine in accordance with

Institutional and Federal guidelines. All feeding schedules, temperature control, water quality monitoring, and light cycles were controlled by the zebrafish facility staff. Natural matings were set up the afternoon before planned injections, typically at 3 pm. Male and female fish (5-20 each) were placed in a breeding tank with a separating divider and placed on the breeding shelves. It is preferred that these fish have not been breed for about two weeks, and that they range in age from 6 months to 18 months old, based on the hatch date written on the tank. The following day, the divider was removed to obtain embryos from natural matings, typically at 9 am. Once the fish have breed, the fertilized embryos were collected and the breeding tank replaced with fresh water. Breeding would typically occur over the course of 1-2 h, and several hundred embryos could be gathered from a single successful breeding set. Once breeding is complete, the male and female fish were separated and then placed back onto the maintenance shelves with the corresponding breeding results indicated on the tank by date.

5.5.2 Zebrafish Injection

Borosilicate glass capillary tubes needles (World Precision Instruments, 1.0 mm; 4 in) were heated and pulled to form micropipette needles. The needles were then marked every 1 mm, representing a volume of ~30 nL between the hash marks. Injections solutions were diluted 1:1 in phenol red dye (1% phenol red in water). The injection solutions were then transferred to small petri dishes and covered with mineral oil. The needles was clipped to form an opening at the tip, and then loaded with ~2 μL of the injection solution. Microinjects were performed with using a World Precision Instruments Pneumatic PicoPump injector. Three injections of 30 nL (volume between two hash mark) were performed and timed with a stopwatch. The times were averaged and divided by 30 to determine the gating time required for injection of a single nL.

Gating times were then multiplied by the volume necessary to deliver the proper amount of reagent. The embryos were aligned in a 3% agarose injection tray containing divots to hold the embryos in place. The loaded needle was then lowered to slowly penetrate the chorion. Plasmid constructs and oligonucleotides were injected directly into the blastomere at the 1-cell stage for maximum efficiency. Blastomere injections were performed in volumes of 1-5 nL for plasmids (50-200 ng), SSOs (1-5 ng), mRNAs (100-300 pg), and tRNAs (0.2-5.4 ng). Amino acids (0-2 mM) were injected into the yolk in volumes of 1-10 nL to allow for diffusion into each cell. Once injections were completed, the embryos were rinsed out of the agarose tray with E3 zebrafish medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.1% methylene blue, pH 7.8) and placed into clean 10 cm dishes. UV irradiations (365 nm transilluminator) were then performed if necessary. Embryos were incubated in the dark at 28 °C, followed by imaging analysis.

5.5.3   Zebrafish Imaging

Early-stage zebrafish embryos (<8 hpf) were directly imaged. Late-stage zebrafish (>16 hpf) were treated with Tricaine (Sigma, MS-222) using a 20X stock solution diluted in E3 zebrafish medium for 5 min and manually dechorionated with forceps. All imaging was performed on a Leica MZ16FA stereo fluorescence microscope with a 1X objective (0.14 NA) at zooms ranging from 30X to 60X, depending on the number and size of the fish. EGFP (filter 41017), mCherry (filter 41043), and brightfield (no filter) images were collected with a QImaging Retiga-EXi Fast 1394 digital camera.

5.6   General DNA Computation Techniques

5.6.1   Duplex Gate Purification

Oligonucleotides were ordered from IDT (nonmodified) and Alpha DNA (5′ TAMRA and 3′ BHQ2 modifications). Gate complexes were assembled at 20 µM in 200 µL of TAE/Mg^{2+} buffer (0.04 M tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), 12.5 mM magnesium acetate, pH ~7.4) and annealed by cooling the solution from 95 to 12 °C over 10 min. The gates were then purified on 1.5 mm 20% native TBE-PAGE gels (200 V, 40 min). The full size duplex bands were identified through UV back shadowing on a TLC plate, excised, cut into small pieces of ~1 mm^3, and eluted by shaking overnight in 500 µL TAE/Mg^{2+} buffer. The supernatant was removed and a second elution was performed for the gel slice through shaking overnight in 500 µL TAE/Mg^{2+} buffer to increase the total yield of purified gate. All elutions were subsequently centrifuged at 3,000 g for 5 min to remove any contaminating agarose debris from the solution. Elutions were also performed in water and PBS buffer pH 7.4 for the fuel-catalysts cycle substrate gate. Gate concentrations (typically 5-10 µM) were determined by UV absorption at 260 nm (Nanodrop) and calculated with the duplex extinction coefficient (IDT OligoAnalyzer 3.1). All oligonucleotide sequences are shown in tables for each experimental section.

5.6.2   Fluorescence Analysis

Synthetic miRNA strands were ordered from Sigma. Each reaction was set up in 50 µL TAE/Mg^{2+} buffer in triplicate wells and incubated at 37 °C for 4 h. The AND gates were used at 200 nM with 200 nM translator gates and 800 nM input strands. The OR gate was used at 200

nM with 800 nM input strands. The (miR-21 OR miR-122) AND miR-125b gate circuit was used at 200 nM with 200 nM translator gates and 800 nM input strands. The fuel-catalyst cycle reporter gate was used at 50 nM with 200 nM or 800 nM substrate gate and fuel strand, and a range from 1-20 nM of the synthetic miRNA. TAMRA fluorescence was measured on BioTek Synergy 4 (ex. 532; em. 576; reading from bottom) and Tecan M1000 (ex. 545; em. 585; reading from bottom) plate readers in black 96-well plates (BD Falcon) after 4 h reaction time and normalized to the positive control for the activated logic gate or fuel-catalyst cycle.

5.6.3 Logic Gate and Catalytic Cycle Cellular Transfections

Transfections were performed in 96-well plates using 1-2 µL of X-tremeGENE siRNA transfection reagent (Roche) per well in 100 µL of Opti-MEM. The light-triggered AND gate was transfected at 200 nM with 800 nM input strands. The miRNA-based AND gates were transfected at 50 nM with 50 nM translator gates and 200 nM input strands. The OR gate was transfected at 200 nM. The (miR-21 OR miR-122) AND miR-125b gate was transfected at 50 nM with 200 nM translator gates. The catalytic cycle reporter gate was transfected at 50 nM with 200 nM of the output activating strand. The substrate gates and fuel strands were transfected at 200 nM or 800 nM. For 4-well chamber slides, 4 µL of the X-tremeGENE reagent in 200 µL of Opti-MEM was used per well, with the same concentrations as indicated above. After 4 h at 37 °C, the Opti-MEM transfection mixtures were removed from the cells and replaced with clear growth media for imaging following protocol 5.4.5. The TAMRA signal was normalized to a standard setting for fluorescence intensity (black = 300; white = 2000; gamma = 0.6) in Zen Pro 2011. Fluorescence merged with brightfield images are shown.

Appendix A

List of Abbreviations

2'OMe   2'-O-methyl modification

5′UTR   5' untranslated region

A   adenine (oligonucelotides)

A   alanine (proteins)

ABI   Applied Biosystems

AMA   ammonium hydroxide/methyl amine

AmMo   amino modified

Amp   ampicillin

AS   alternative splicing

ATP   adenosine triphosphate

bp   base pair

bPEI   branched polyethylene imine

C   cytosine

cAMP   cyclic adenosine monophosphate

Cas9   CRISPR-associated endonuclease

CD71	transmembrane transferrin receptor (TfR1)
CDB-TFO	caged dumbbell TFO
CHP-TFO	caged hairpin TFO
CHR	cascade hybridization reaction
CIBN	cryptochrome-interacting basic-helix-loop-helix
cMO	caged MO
CMV	cytomegalovirus promoter
CoREST	corepressor silencing transcription factor
Cre	tyrosine recombinase enzyme
CREB	cAMP response element-binding protein
CRISPR	clustered regularly interspaced short palindromic repeats
Cry	cryptochrome domain
Cys	cysteine
DAPI	4',6-diamidino-2-phenylindole fluorophore, nuclear stain
dCas9	catalytically inactive Cas9 variant
DEACM	diethyl amino coumarin
DEPC	diethylpyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DMTr	dimethoxytrityl-[bis-(4-methoxyphenyl)phenylmethyl]
DNA	deoxyribonucleic acid
DsRed	*Discosoma sp.* red fluorescent protein
EDTA	ethylenediaminetetraacetic acid

Eg5	kinesin-5 motor protein
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
ERK	extracellular signal-regulated kinases
FACS	fluorescence-activated cell sorting
FAD	flavin adenine dinucleotide
FRET	fluorescence resonance energy transfer
G	guanine
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein
gRNA	guide RNA
H_2O	water
HA	human influenza hemagglutinin
HCR	hybridization chain reaction
HCK	coumarin-based caged analogue of lysine
hpf	hours post fertilization
HR	homology-directed repair
IRES	internal ribosome entry site
IZ	intermediate zone
K	lysine
Kan	kanamycin
kb	kilobases
LARI	light activated RNA interference

LASSO	light activated SSO
LDSSO	light deactivated SSO
LED	light-emitting diode
LOV	light-oxygen-voltage domain
lPEI	linear polyethylene imine
Luc	luciferase
Lys	lysine
MAP	mitogen-activated protein kinase
MBT	midblastula transition
mCherry	*Discosoma sp.* red fluorescent protein
MEK	mitogen-activated protein kinase kinase
MH2	functional domain for homotrimerization of Sma2
MO	morpholino
mRNA	messenger RNA
miRNA	micro RNA
NHEJ	non-homologous end-joining
NPE	*ortho*-nitrophenylethyl caging group
NPOM	6-nitropiperonyloxymethyl caging group
NPOC	6-nitropiperonyloxycarbonyl caging group
NTP	nucleoside triphosphate
ONB	*ortho*-nitrobenzyl caging group
PAGE	polyacrylamide gel
PBS	phosphate-buffered saline

PCB	phycocyanobilin chromophore
PCK	photocaged lysine (NPOC)
PCKRS	PCK-tRNA synthetase mutant
PCR	polymerase chain reaction
Phe	phenylalanine
PheRS	Phe-tRNA synthetase mutant
Phy	phytochrome domain
PKA	cyclic-AMP dependent kinase
PKI	PKA inhibitory peptide
PIF	phytochrome interaction factor
Plk3	pololike kinase 3
PS	phosphorothiotate modification
PylRS	pyrrolysyl-tRNA synthetase
PylT	pyrrolysyl tRNA$_{CUA}$
qRT-PCR	quantitative real-time PCR
Rac1	Ras-related C3 botulinum toxin substrate 1
RISC	RNA inducing silencing complex
RLU	relative luminescence unit
RFP	*Discosoma sp.* red fluorescent protein
RFU	relative fluorescence unit
RNA	ribonucleic acid
RNAi	RNA interference
SARA	Smad anchor for receptor activation

SARA-SBD	SARA binding domain
SDM	site-directed mutagenesis
shRNA	short hairpin RNA
siRNA	small interfering RNA
sox31	member of the B1 Sox gene family
SSO	splice-switching oligonucleotide
T	thymine
TAG	amber codon
T7RNAp	T7 RNA polymerase
TAE/Mg^{2+}	tris base, acetic acid, EDTA, magnesium buffer
TAMRA	tetramethylrhodamine
TATA box	Goldberg-Hogness transcription initiator sequence
TBE	tris base, borate, EDTA buffer
TBP	TATA box binding protein
TCA	trichloroacetic acid
Tet	tetracycline
TFIID	transcription factor IID
T_M	melting temperature
TFA	trifluoroacetic acid
TFO	triplex-forming oligonucleotide
THF	tetrahydrofuran
tRNA	transfer RNA
Tyr	tyrosine

U	uracil
U6	Pol III promoter
UAA	unnatural amino acid
UV	ultraviolet light
UV-A	wavelength of ultraviolet light $320 - 400$ nm
VZ	ventricular zone
WT	wild-type
YFP	yellow fluorescent protein

Bibliography

1. Brevini, T. A.; Cillo, F.; Antonini, S.; Tosetti, V.; Gandolfi, F., Temporal and spatial control of gene expression in early embryos of farm animals. *Reprod Fertil Dev* **2007,** *19* (1), 35-42; Li, E.; Davidson, E. H., Building developmental gene regulatory networks. *Birth Defects Res C Embryo Today* **2009,** *87* (2), 123-30; Chan, T. M.; Longabaugh, W.; Bolouri, H.; Chen, H. L.; Tseng, W. F.; Chao, C. H.; Jang, T. H.; Lin, Y. I.; Hung, S. C.; Wang, H. D.; Yuh, C. H., Developmental gene regulatory networks in the zebrafish embryo. *Biochim Biophys Acta* **2009,** *1789* (4), 279-98.

2. Mayer, G.; Heckel, A., Biologically active molecules with a "light switch". *Angew Chem Int Ed Engl* **2006,** *45* (30), 4900-21; Young, D. D.; Deiters, A., Photochemical control of biological processes. *Org Biomol Chem* **2007,** *5* (7), 999-1005.

3. Engels, J.; Schlaeger, E. J., Synthesis, structure, and reactivity of adenosine cyclic 3',5'-phosphate benzyl triesters. *J Med Chem* **1977,** *20* (7), 907-11.

4. Kaplan, J. H.; Forbush, B.; Hoffman, J. F., Rapid photolytic release of adenosine 5'-triphosphate from a protected analogue: utilization by the Na:K pump of human red blood cell ghosts. *Biochemistry* **1978,** *17* (10), 1929-35.

5. Pelliccioli, A. P.; Wirz, J., Photoremovable protecting groups: reaction mechanisms and applications. *Photochem Photobiol Sci* **2002,** *1* (7), 441-58.

6. McCray, J. A.; Trentham, D. R., Properties and uses of photoreactive caged compounds. *Annu Rev Biophys Biophys Chem* **1989,** *18*, 239-70; Dormán, G.; Prestwich, G. D., Using photolabile ligands in drug discovery and development. *Trends Biotechnol* **2000,** *18* (2), 64-77; Ellis-Davies, G. C., Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nat Methods* **2007,** *4* (8), 619-28; Specht, A.; Bolze, F.; Omran, Z.; Nicoud, J. F.; Goeldner, M., Photochemical tools to study dynamic biological processes. *HFSP J* **2009,** *3* (4), 255-64.

7. Pierce, D. H.; Scarpa, A.; Trentham, D. R.; Topp, M. R.; Blasie, J. K., Comparison of the kinetics of calcium transport in vesicular dispersions and oriented multilayers of isolated sarcoplasmic reticulum membranes. *Biophys J* **1983,** *44* (3), 365-73; Dolphin, A. C.; Wootton, J. F.; Scott, R. H.; Trentham, D. R., Photoactivation of intracellular guanosine triphosphate analogues reduces the amplitude and slows the kinetics of voltage-activated calcium channel currents in sensory neurones. *Pflugers Arch* **1988,** *411* (6), 628-36; Blatt, M. R.; Thiel, G.; Trentham, D. R., Reversible inactivation of K+ channels of Vicia stomatal guard cells following the photolysis of caged inositol 1,4,5-trisphosphate. *Nature* **1990,** *346* (6286), 766-9.

8. Goldman, Y. E.; Hibberd, M. G.; McCray, J. A.; Trentham, D. R., Relaxation of muscle fibres by photolysis of caged ATP. *Nature* **1982,** *300* (5894), 701-5; Walker, J. W.; Somlyo, A. V.; Goldman, Y. E.; Somlyo, A. P.; Trentham, D. R., Kinetics of smooth and

skeletal muscle activation by laser pulse photolysis of caged inositol 1,4,5-trisphosphate. *Nature* **1987,** *327* (6119), 249-52; Dantzig, J. A.; Walker, J. W.; Trentham, D. R.; Goldman, Y. E., Relaxation of muscle fibers with adenosine 5'-[gamma-thio]triphosphate (ATP[gamma S]) and by laser photolysis of caged ATP[gamma S]: evidence for Ca2+-dependent affinity of rapidly detaching zero-force cross-bridges. *Proc Natl Acad Sci U S A* **1988,** *85* (18), 6716-20.

9.   McCray, J. A.; Herbette, L.; Kihara, T.; Trentham, D. R., A new approach to time-resolved studies of ATP-requiring biological systems; laser flash photolysis of caged ATP. *Proc Natl Acad Sci U S A* **1980,** *77* (12), 7237-41.

10. Walker, J. W.; Reid, G. P.; Trentham, D. R., Synthesis and properties of caged nucleotides. *Methods Enzymol* **1989,** *172*, 288-301.

11. Callaway, E. M., Caged neurotransmitters. Shedding light on neural circuits. *Curr Biol* **1994,** *4* (11), 1010-2; Nerbonne, J. M., Caged compounds: tools for illuminating neuronal responses and connections. *Curr Opin Neurobiol* **1996,** *6* (3), 379-86; Kramer, R. H.; Fortin, D. L.; Trauner, D., New photochemical tools for controlling neuronal activity. *Curr Opin Neurobiol* **2009,** *19* (5), 544-52; Szobota, S.; Isacoff, E. Y., Optical control of neuronal activity. *Annu Rev Biophys* **2010,** *39*, 329-48; Hess, G. P.; Lewis, R. W.; Chen, Y., Caged neurotransmitters and other caged compounds: design and application. *Cold Spring Harb Protoc* **2014,** *2014* (10), pdb.top084152; Amatrudo, J. M.; Olson, J. P.; Agarwal, H. K.; Ellis-Davies, G. C., Caged compounds for multichromic optical interrogation of neural systems. *Eur J Neurosci* **2015,** *41* (1), 5-16.

12. Lee, H. M.; Larson, D. R.; Lawrence, D. S., Illuminating the chemistry of life: design, synthesis, and applications of "caged" and related photoresponsive compounds. *ACS Chem Biol* **2009,** *4* (6), 409-27; Deiters, A., Principles and applications of the photochemical control of cellular processes. *Chembiochem* **2010,** *11* (1), 47-53; Brieke, C.; Rohrbach, F.; Gottschalk, A.; Mayer, G.; Heckel, A., Light-controlled tools. *Angew Chem Int Ed Engl* **2012,** *51* (34), 8446-76; Herrmann, A., Using photolabile protecting groups for the controlled release of bioactive volatiles. *Photochem Photobiol Sci* **2012,** *11* (3), 446-59; Ahmed, I.; Fruk, L., The power of light: photosensitive tools for chemical biology. *Mol Biosyst* **2013,** *9* (4), 565-70.

13. Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J., Photoremovable protecting groups in chemistry and biology: reaction mechanisms and efficacy. *Chem Rev* **2013,** *113* (1), 119-91.

14. Meunier, J. R.; Sarasin, A.; Marrot, L., Photogenotoxicity of mammalian cells: a review of the different assays for in vitro testing. *Photochem Photobiol* **2002,** *75* (5), 437-47; Forman, J.; Dietrich, M.; Monroe, W. T., Photobiological and thermal effects of photoactivating UVA light doses on cell cultures. *Photochem Photobiol Sci* **2007,** *6* (6), 649-58; Dong, Q.; Svoboda, K.; Tiersch, T.; Monroe, W., Photobiological effects of UVA and UVB light in zebrafish embryos: Evidence for a competent photorepair system. *Journal of Photochemistry and Photobiology B-Biology* **2007,** *88* (2-3), 137-146.

15. Adams, S. R.; Tsien, R. Y., CONTROLLING CELL CHEMISTRY WITH CAGED COMPOUNDS. *Annual Review of Physiology* **1993,** *55*, 755-784; Humphrey, D.; Rajfur, Z.; Imperiali, B.; Marriott, G.; Roy, P.; Jacobson, K., Design, synthesis, and characterization of caged compounds. *CSH Protoc* **2007,** *2007*, pdb.ip25.

16. WATSON, J. D.; CRICK, F. H., The structure of DNA. *Cold Spring Harb Symp Quant Biol* **1953,** *18*, 123-31.

17.   Crick, F., Central dogma of molecular biology. *Nature* **1970,** *227* (5258), 561-3.
18.   Kim, H. G.; Miller, D. M., A novel triplex-forming oligonucleotide targeted to human cyclin D1 (bcl-1, proto-oncogene) promoter inhibits transcription in HeLa cells. *Biochemistry* **1998,** *37* (8), 2666-72.
19.   Eckstein, F.; Krieg, A. M.; Stein, C. A.; Agrawal, S.; Beaucage, S.; Cook, P. D.; Crooke, S.; Gait, M. J.; Gewirtz, A.; Hélène, C.; Miller, P.; Narayanan, R.; Nicolin, A.; Nielsen, P.; Ohtsuka, E.; Seliger, H.; Stec, W.; Tidd, D.; Wagner, R.; Zon, J., On the quality control of antisense oligonucleotides. *Antisense Nucleic Acid Drug Dev* **1996,** *6* (3), 149.
20.   Meister, G.; Landthaler, M.; Dorsett, Y.; Tuschl, T., Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. *RNA* **2004,** *10* (3), 544-50.
21.   Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M., Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **2005,** *438* (7068), 685-9.
22.   Amarzguioui, M.; Rossi, J. J.; Kim, D., Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Lett* **2005,** *579* (26), 5974-81.
23.   Kole, R.; Krainer, A. R.; Altman, S., RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* **2012,** *11* (2), 125-40.
24.   Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **1992,** *355* (6360), 564-6.
25.   Micklefield, J., Backbone modification of nucleic acids: synthesis, structure and therapeutic applications. *Curr Med Chem* **2001,** *8* (10), 1157-79.
26.   Faria, M.; Ulrich, H., Sugar boost: when ribose modifications improve oligonucleotide performance. *Curr Opin Mol Ther* **2008,** *10* (2), 168-75.
27.   Wilson, J. N.; Kool, E. T., Fluorescent DNA base replacements: Reporters and sensors for biological systems. *Org Biomol Chem* **2006,** *4* (23), 4265-74.
28.   Thuong, N. T.; Asseline, U., Modification of the 5' terminus of oligonucleotides for attachment of reporter and conjugate groups. *Curr Protoc Nucleic Acid Chem* **2001,** *Chapter 4*, Unit 4.2.
29.   Venkatesan, N.; Kim, S. J.; Kim, B. H., Novel phosphoramidite building blocks in synthesis and applications toward modified oligonucleotides. *Curr Med Chem* **2003,** *10* (19), 1973-91; Bell, N. M.; Micklefield, J., Chemical modification of oligonucleotides for therapeutic, bioanalytical and other applications. *Chembiochem* **2009,** *10* (17), 2691-703.
30.   Liu, Q.; Deiters, A., Optochemical Control of Deoxyoligonucleotide Function via a Nucleobase-Caging Approach. *Acc Chem Res* **2014,** *47* (1), 45-55.
31.   Lenox, H. J.; McCoy, C. P.; Sheppard, T. L., Site-specific generation of deoxyribonolactone lesions in DNA oligonucleotides. *Org Lett* **2001,** *3* (15), 2415-8; Zheng, Y.; Sheppard, T. L., Half-life and DNA strand scission products of 2-deoxyribonolactone oxidative DNA damage lesions. *Chem Res Toxicol* **2004,** *17* (2), 197-207; Trzupek, J. D.; Sheppard, T. L., Photochemical generation of ribose abasic sites in RNA oligonucleotides. *Org Lett* **2005,** *7* (8), 1493-6; Horn, T.; Downing, K.; Gee, Y.; Urdea, M. S., Controlled chemical cleavage of synthetic DNA at specific sites. *Nucleosides & Nucleotides* **1991,** *10* (1-3), 299-302.
32.   Yamana, K.; Yoshikawa, A.; Nakano, H., Synthesis of a new photoisomerizable linker for connecting two oligonucleotide segments. *Tetrahedron letters* **1996,** *37* (5), 637-640; Young, D. D.; Edwards, W. F.; Lusic, H.; Lively, M. O.; Deiters, A., Light-triggered

polymerase chain reaction. *Chem Commun (Camb)* **2008,** (4), 462-4; Young, D. D.; Lusic, H.; Lively, M. O.; Deiters, A., Restriction enzyme-free mutagenesis via the light regulation of DNA polymerization. *Nucleic Acids Res* **2009,** *37* (8), e58; Young, D. D.; Govan, J. M.; Lively, M. O.; Deiters, A., Photochemical regulation of restriction endonuclease activity. *Chembiochem* **2009,** *10* (10), 1612-6.

33.   Heckel, A.; Mayer, G., Light regulation of aptamer activity: an anti-thrombin aptamer with caged thymidine nucleobases. *J Am Chem Soc* **2005,** *127* (3), 822-3; Mayer, G.; Müller, J.; Mack, T.; Freitag, D. F.; Höver, T.; Pötzsch, B.; Heckel, A., Differential regulation of protein subdomain activity with caged bivalent ligands. *Chembiochem* **2009,** *10* (4), 654-7.

34.   Chaulk, S. G.; MacMillan, A. M., Caged RNA: photo-control of a ribozyme reaction. *Nucleic Acids Res* **1998,** *26* (13), 3173-8.

35.   Silverman, S. K., Catalytic DNA (deoxyribozymes) for synthetic applications-current abilities and future prospects. *Chem Commun (Camb)* **2008,** (30), 3467-85; Lusic, H.; Lively, M. O.; Deiters, A., Light-activated deoxyguanosine: photochemical regulation of peroxidase activity. *Mol Biosyst* **2008,** *4* (6), 508-11; Young, D. D.; Garner, R. A.; Yoder, J. A.; Deiters, A., Light-activation of gene function in mammalian cells via ribozymes. *Chem Commun (Camb)* **2009,** (5), 568-70; Richards, J. L.; Seward, G. K.; Wang, Y. H.; Dmochowski, I. J., Turning the 10-23 DNAzyme on and off with light. *Chembiochem* **2010,** *11* (3), 320-4.

36.   Ando, H.; Furuta, T.; Tsien, R. Y.; Okamoto, H., Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos. *Nat Genet* **2001,** *28* (4), 317-25.

37.   Govan, J. M.; Lively, M. O.; Deiters, A., Photochemical control of DNA decoy function enables precise regulation of nuclear factor κB activity. *J Am Chem Soc* **2011,** *133* (33), 13176-82.

38.   Shestopalov, I. A.; Sinha, S.; Chen, J. K., Light-controlled gene silencing in zebrafish embryos. *Nat Chem Biol* **2007,** *3* (10), 650-1.

39.   Tang, X.; Swaminathan, J.; Gewirtz, A. M.; Dmochowski, I. J., Regulating gene expression in human leukemia cells using light-activated oligodeoxynucleotides. *Nucleic Acids Res* **2008,** *36* (2), 559-69; Young, D. D.; Lusic, H.; Lively, M. O.; Yoder, J. A.; Deiters, A., Gene silencing in mammalian cells with light-activated antisense agents. *Chembiochem* **2008,** *9* (18), 2937-40; Deiters, A.; Garner, R. A.; Lusic, H.; Govan, J. M.; Dush, M.; Nascone-Yoder, N. M.; Yoder, J. A., Photocaged morpholino oligomers for the light-regulation of gene function in zebrafish and Xenopus embryos. *J Am Chem Soc* **2010,** *132* (44), 15644-50; Young, D. D.; Lively, M. O.; Deiters, A., Activation and deactivation of DNAzyme and antisense function with light for the photochemical regulation of gene expression in mammalian cells. *J Am Chem Soc* **2010,** *132* (17), 6183-93.

40.   Mikat, V.; Heckel, A., Light-dependent RNA interference with nucleobase-caged siRNAs. *RNA* **2007,** *13* (12), 2341-7.

41.   Blidner, R. A.; Svoboda, K. R.; Hammer, R. P.; Monroe, W. T., Photoinduced RNA interference using DMNPE-caged 2'-deoxy-2'-fluoro substituted nucleic acids in vitro and in vivo. *Mol Biosyst* **2008,** *4* (5), 431-40; Shah, S.; Jain, P. K.; Kala, A.; Karunakaran, D.; Friedman, S. H., Light-activated RNA interference using double-stranded siRNA precursors modified using a remarkable regiospecificity of diazo-based photolabile groups. *Nucleic Acids Res* **2009,** *37* (13), 4508-17; Casey, J. P.; Blidner, R. A.; Monroe,

W. T., Caged siRNAs for spatiotemporal control of gene silencing. *Mol Pharm* **2009,** *6* (3), 669-85; Govan, J. M.; Young, D. D.; Lusic, H.; Liu, Q.; Lively, M. O.; Deiters, A., Optochemical control of RNA interference in mammalian cells. *Nucleic Acids Res* **2013,** *41* (22), 10518-28.

42. Monroe, W. T.; McQuain, M. M.; Chang, M. S.; Alexander, J. S.; Haselton, F. R., Targeting expression with light using caged DNA. *J Biol Chem* **1999,** *274* (30), 20895-900.

43. Ando, H.; Furuta, T.; Okamoto, H., Photo-mediated gene activation by using caged mRNA in zebrafish embryos. *Methods Cell Biol* **2004,** *77*, 159-71; Okamoto, H.; Hirate, Y.; Ando, H., Systematic identification of factors in zebrafish regulating the early midbrain and cerebellar development by odered differential display and caged mRNA technology. *Frontiers in Bioscience* **2004,** *9*, 93-99.

44. Shah, S.; Rangarajan, S.; Friedman, S. H., Light-Activated RNA Interference. *Angewandte Chemie* **2005,** *117* (9), 1352-1356.

45. Jain, P. K.; Shah, S.; Friedman, S. H., Patterning of gene expression using new photolabile groups applied to light activated RNAi. *Journal of the American Chemical Society* **2010,** *133* (3), 440-446.

46. Tang, X.; Maegawa, S.; Weinberg, E. S.; Dmochowski, I. J., Regulating gene expression in zebrafish embryos using light-activated, negatively charged peptide nucleic acids. *J Am Chem Soc* **2007,** *129* (36), 11000-1; Ouyang, X.; Shestopalov, I. A.; Sinha, S.; Zheng, G.; Pitt, C. L.; Li, W. H.; Olson, A. J.; Chen, J. K., Versatile synthesis and rational design of caged morpholinos. *J Am Chem Soc* **2009,** *131* (37), 13255-69; Wang, Y.; Wu, L.; Wang, P.; Lv, C.; Yang, Z.; Tang, X., Manipulation of gene expression in zebrafish using caged circular morpholino oligomers. *Nucleic Acids Res* **2012,** *40* (21), 11155-62; Yamazoe, S.; Shestopalov, I. A.; Provost, E.; Leach, S. D.; Chen, J. K., Cyclic caged morpholinos: conformationally gated probes of embryonic gene function. *Angew Chem Int Ed Engl* **2012,** *51* (28), 6908-11.

47. Summerton, J.; Weller, D., Morpholino antisense oligomers: design, preparation, and properties. *Antisense and Nucleic Acid Drug Development* **1997,** *7* (3), 187-195.

48. Nasevicius, A.; Ekker, S. C., Effective targeted gene 'knockdown' in zebrafish. *Nature genetics* **2000,** *26* (2), 216-220.

49. Ordoukhanian, P.; Taylor, J.-S., Design and Synthesis of a Versatile Photocleavable DNA Building Block. Application to Phototriggered Hybridization. *Journal of the American Chemical Society* **1995,** *117* (37), 9570-9571; Ordoukhanian, P.; Taylor, J. S., Caged single and double strand breaks. *Bioconjug Chem* **2000,** *11* (1), 94-103.

50. Caruthers, M. H.; Beaucage, S. L.; Efcavitch, J. W.; Fisher, E. F.; Matteucci, M. D.; Stabinsky, Y., New chemical methods for synthesizing polynucleotides. *Nucleic Acids Symp Ser* **1980,** (7), 215-23; Caruthers, M. H.; Beaucage, S. L.; Becker, C.; Efcavitch, J. W.; Fisher, E. F.; Galluppi, G.; Goldman, R.; deHaseth, P.; Matteucci, M.; McBride, L., Deoxyoligonucleotide synthesis via the phosphoramidite method. *Gene Amplif Anal* **1983,** *3*, 1-26.

51. Pitsch, S.; Weiss, P. A.; Wu, X.; Ackermann, D.; Honegger, T., Fast and Reliable Automated Synthesis of RNA and Partially 2′-O-Protected Precursors (Caged RNA') Based on Two Novel, Orthogonal 2′-O-Protecting Groups, Preliminary Communication. *Helvetica chimica acta* **1999,** *82* (10), 1753-1761.

52.	Tomasini, A. J.; Schuler, A. D.; Zebala, J. A.; Mayer, A. N., PhotoMorphs: a novel light-activated reagent for controlling gene expression in zebrafish. *Genesis* **2009,** *47* (11), 736-43; Tallafuss, A.; Gibson, D.; Morcos, P.; Li, Y.; Seredick, S.; Eisen, J.; Washbourne, P., Turning gene function ON and OFF using sense and antisense photo-morpholinos in zebrafish. *Development* **2012,** *139* (9), 1691-9.

53.	Rodrigues-Correia, A.; Koeppel, M. B.; Schäfer, F.; Joshi, K. B.; Mack, T.; Heckel, A., Comparison of the duplex-destabilizing effects of nucleobase-caged oligonucleotides. *Anal Bioanal Chem* **2011,** *399* (1), 441-7.

54.	Alvarez, K.; Vasseur, J.-J.; Beltran, T.; Imbach, J.-L., Photocleavable Protecting Groups as Nucleobase Protections Allowed the Solid-Phase Synthesis of Base-Sensitive SATE-Prooligonucleotides. *The Journal of Organic Chemistry* **1999,** *64* (17), 6319-6328.

55.	Ting, R.; Lermer, L.; Perrin, D. M., Triggering DNAzymes with Light: A Photoactive C8 Thioether-Linked Adenosine. *Journal of the American Chemical Society* **2004,** *126* (40), 12720-12721.

56.	Kröck, L.; Heckel, A., Photoinduced transcription by using temporarily mismatched caged oligonucleotides. *Angew Chem Int Ed Engl* **2005,** *44* (3), 471-3.

57.	Mayer, G.; Kröck, L.; Mikat, V.; Engeser, M.; Heckel, A., Light-induced formation of G-quadruplex DNA secondary structures. *Chembiochem* **2005,** *6* (11), 1966-70.

58.	Wenter, P.; Fürtig, B.; Hainard, A.; Schwalbe, H.; Pitsch, S., Kinetics of Photoinduced RNA Refolding by Real-Time NMR Spectroscopy. *Angewandte Chemie International Edition* **2005,** *44* (17), 2600-2603.

59.	Höbartner, C.; Silverman, S. K., Modulation of RNA tertiary folding by incorporation of caged nucleotides. *Angew Chem Int Ed Engl* **2005,** *44* (44), 7305-9.

60.	Tang, X.; Dmochowski, I. J., Phototriggering of caged fluorescent oligodeoxynucleotides. *Org Lett* **2005,** *7* (2), 279-82.

61.	Heckel, A., Nucleobase-caged phosphoramidites for oligonucleotide synthesis. *Curr Protoc Nucleic Acid Chem* **2007,** *Chapter 1*, Unit 1.17.

62.	Cashion, P. J.; Notman, H. J.; Cadger, T.; Sathe, G. M., Purification of oligodeoxynucleotides on Sephadex G-10 columns. *J Chromatogr* **1977,** *135* (1), 189-92.

63.	Crain, P. F.; McCloskey, J. A., Applications of mass spectrometry to the characterization of oligonucleotides and nucleic acids. *Curr Opin Biotechnol* **1998,** *9* (1), 25-34; van Dongen, W. D.; Niessen, W. M., Bioanalytical LC-MS of therapeutic oligonucleotides. *Bioanalysis* **2011,** *3* (5), 541-64; Basiri, B.; Bartlett, M. G., LC-MS of oligonucleotides: applications in biomedical research. *Bioanalysis* **2014,** *6* (11), 1525-42.

64.	Nawrot, B.; Sipa, K., Chemical and structural diversity of siRNA molecules. *Curr Top Med Chem* **2006,** *6* (9), 913-25; Zhang, H. Y.; Du, Q.; Wahlestedt, C.; Liang, Z., RNA Interference with chemically modified siRNA. *Curr Top Med Chem* **2006,** *6* (9), 893-900; Prakash, T. P.; Bhat, B., 2'-Modified oligonucleotides for antisense therapeutics. *Curr Top Med Chem* **2007,** *7* (7), 641-9; Lennox, K. A.; Behlke, M. A., A direct comparison of anti-microRNA oligonucleotide potency. *Pharm Res* **2010,** *27* (9), 1788-99.

65.	Eckstein, F.; Gish, G., Phosphorothioates in molecular biology. *Trends Biochem Sci* **1989,** *14* (3), 97-100.

66.	Beaucage, S. L., Oligodeoxyribonucleotides synthesis. Phosphoramidite approach. *Methods Mol Biol* **1993,** *20*, 33-61.

67.   Riggsbee, C. W.; Deiters, A., Recent advances in the photochemical control of protein function. *Trends Biotechnol* **2010,** *28* (9), 468-75; Baker, A. S.; Deiters, A., Optical control of protein function through unnatural amino acid mutagenesis and other optogenetic approaches. *ACS Chem Biol* **2014,** *9* (7), 1398-407.

68.   ; Lin, W.; Lawrence, D. S., A strategy for the construction of caged diols using a photolabile protecting group. *The Journal of organic chemistry* **2002,** *67* (8), 2723-2726; Hayashi, K.-i.; Hashimoto, K.; Kusaka, N.; Yamazoe, A.; Fukaki, H.; Tasaka, M.; Nozaki, H., Caged gene-inducer spatially and temporally controls gene expression and plant development in transgenic Arabidopsis plant. *Bioorganic & medicinal chemistry letters* **2006,** *16* (9), 2470-2474.

69.   Cambridge, S. B.; Geissler, D.; Keller, S.; Cürten, B., A caged doxycycline analogue for photoactivated gene expression. *Angewandte Chemie International Edition* **2006,** *45* (14), 2229-2231.

70.   Link, K. H.; Shi, Y.; Koh, J. T., Light activated recombination. *Journal of the American Chemical Society* **2005,** *127* (38), 13088-13089.

71.   Goard, M.; Aakalu, G.; Fedoryak, O. D.; Quinonez, C.; Julien, J. S.; Poteet, S. J.; Schuman, E. M.; Dore, T. M., Light-mediated inhibition of protein synthesis. *Chemistry & biology* **2005,** *12* (6), 685-693.

72.   Herrou, J.; Crosson, S., Function, structure and mechanism of bacterial photosensory LOV proteins. *Nature reviews microbiology* **2011,** *9* (10), 713-723.

73.   Rockwell, N. C.; Su, Y.-S.; Lagarias, J. C., Phytochome structure and signaling mechanisms. *Annual review of plant biology* **2006,** *57*, 837.

74.   Lin, C.; Shalitin, D., Cryptochrome structure and signal transduction. *Annual Review of Plant Biology* **2003,** *54* (1), 469-496.

75.   Casal, J. J., Phytochromes, cryptochromes, phototropin: photoreceptor interactions in plants. *Photochemistry and Photobiology* **2000,** *71* (1), 1-11; Möglich, A.; Yang, X.; Ayers, R. A.; Moffat, K., Structure and function of plant photoreceptors. *Annual review of plant biology* **2010,** *61*, 21-47.

76.   Möglich, A.; Moffat, K., Engineered photoreceptors as novel optogenetic tools. *Photochemical & Photobiological Sciences* **2010,** *9* (10), 1286-1300; Toettcher, J. E.; Voigt, C. A.; Weiner, O. D.; Lim, W. A., The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. *Nat Methods* **2011,** *8* (1), 35-8; Miesenböck, G., Optogenetic control of cells and circuits. *Annu Rev Cell Dev Biol* **2011,** *27*, 731-58; Fenno, L.; Yizhar, O.; Deisseroth, K., The development and application of optogenetics. *Annu Rev Neurosci* **2011,** *34*, 389-412; Konermann, S.; Brigham, M. D.; Trevino, A.; Hsu, P. D.; Heidenreich, M.; Le Cong; Platt, R. J.; Scott, D. A.; Church, G. M.; Zhang, F., Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **2013**.

77.   Shimizu-Sato, S.; Huq, E.; Tepperman, J. M.; Quail, P. H., A light-switchable gene promoter system. *Nature biotechnology* **2002,** *20* (10), 1041-1044; Levskaya, A.; Chevalier, A. A.; Tabor, J. J.; Simpson, Z. B.; Lavery, L. A.; Levy, M.; Davidson, E. A.; Scouras, A.; Ellington, A. D.; Marcotte, E. M., Synthetic biology: engineering Escherichia coli to see light. *Nature* **2005,** *438* (7067), 441-442; Tyszkiewicz, A. B.; Muir, T. W., Activation of protein splicing with light in yeast. *Nature Methods* **2008,** *5* (4), 303-305; Yazawa, M.; Sadaghiani, A. M.; Hsueh, B.; Dolmetsch, R. E., Induction of protein-protein interactions in live cells using light. *Nat Biotechnol* **2009,** *27* (10), 941-5;

Losi, A.; Gärtner, W., The evolution of flavin-binding photoreceptors: an ancient chromophore serving trendy blue-light sensors. *Annu Rev Plant Biol* **2012,** *63*, 49-72; Tischer, D.; Weiner, O. D., Illuminating cell signalling with optogenetic tools. *Nature Reviews Molecular Cell Biology* **2014,** *15* (8), 551-558.

78.   Drepper, T.; Krauss, U.; Zu Berstenhorst, S. M.; Pietruszka, J.; Jaeger, K.-E., Lights on and action! Controlling microbial gene expression by light. *Applied microbiology and biotechnology* **2011,** *90* (1), 23-40.

79.   Yi, J. J.; Wang, H.; Vilela, M.; Danuser, G.; Hahn, K. M., Manipulation of endogenous kinase activity in living cells using photoswitchable inhibitory peptides. *ACS synthetic biology* **2014,** *3* (11), 788-795.

80.   Lahiry, P.; Torkamani, A.; Schork, N. J.; Hegele, R. A., Kinase mutations in human disease: interpreting genotype–phenotype relationships. *Nature Reviews Genetics* **2010,** *11* (1), 60-74.

81.   Walsh, D. A.; Ashby, C. D.; Gonzalez, C.; Calkins, D.; Fischer, E. H.; Krebs, E. G., Purification and characterization of a protein inhibitor of adenosine 3', 5'-monophosphate-dependent protein kinases. *Journal of Biological Chemistry* **1971,** *246* (7), 1977-1985.

82.   Harper, S. M.; Neil, L. C.; Gardner, K. H., Structural basis of a phototropin light switch. *Science* **2003,** *301* (5639), 1541-1544.

83.   Montminy, M., Transcriptional regulation by cyclic AMP. *Annual review of biochemistry* **1997,** *66* (1), 807-822.

84.   Levskaya, A.; Weiner, O.; Lim, W.; Voigt, C., Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **2009,** *461* (7266), 997-1001.

85.   Ridley, A. J.; Schwartz, M. A.; Burridge, K.; Firtel, R. A.; Ginsberg, M. H.; Borisy, G.; Parsons, J. T.; Horwitz, A. R., Cell migration: integrating signals from front to back. *Science* **2003,** *302* (5651), 1704-1709; Raftopoulou, M.; Hall, A., Cell migration: Rho GTPases lead the way. *Developmental biology* **2004,** *265* (1), 23-32.

86.   Quail, P. H., Phytochrome photosensory signalling networks. *Nature Reviews Molecular Cell Biology* **2002,** *3* (2), 85-93.

87.   Michiels, F.; Stam, J. C.; Hordijk, P. L.; Van Der Kammen, R. A.; Ruuls-Van Stalle, L.; Feltkamp, C. A.; Collard, J. G., Regulated membrane localization of Tiam1, mediated by the NH2-terminal pleckstrin homology domain, is required for Rac-dependent membrane ruffling and C-Jun NH2-terminal kinase activation. *The Journal of cell biology* **1997,** *137* (2), 387-398; Inoue, T.; Do Heo, W.; Grimley, J. S.; Wandless, T. J.; Meyer, T., An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nature methods* **2005,** *2* (6), 415-418.

88.   Kennedy, M. J.; Hughes, R. M.; Peteya, L. A.; Schwartz, J. W.; Ehlers, M. D.; Tucker, C. L., Rapid blue-light-mediated induction of protein interactions in living cells. *Nat Methods* **2010,** *7* (12), 973-5.

89.   Nagy, A., Cre recombinase: the universal reagent for genome tailoring. *genesis* **2000,** *26* (2), 99-109.

90.   Jullien, N.; Sampieri, F.; Enjalbert, A.; Herman, J. P., Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. *Nucleic acids research* **2003,** *31* (21), e131-e131.

91.   Liu, H.; Yu, X.; Li, K.; Klejnot, J.; Yang, H.; Lisiero, D.; Lin, C., Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science* **2008,** *322* (5907), 1535-1539.

92.   Tucker, C. L., Manipulating cellular processes using optical control of protein-protein interactions. *Prog Brain Res* **2012,** *196*, 95-117.

93.   Curley, K.; Lawrence, D. S., Light-activated proteins. *Curr Opin Chem Biol* **1999,** *3* (1), 84-8; Shigeri, Y.; Tatsu, Y.; Yumoto, N., Synthesis and application of caged peptides and proteins. *Pharmacol Ther* **2001,** *91* (2), 85-92; Lawrence, D. S., The preparation and in vivo applications of caged peptides and proteins. *Curr Opin Chem Biol* **2005,** *9* (6), 570-5.

94.   Zou, K.; Miller, W. T.; Givens, R. S.; Bayley, H., Caged thiophosphotyrosine peptides. *Angewandte Chemie International Edition* **2001,** *40* (16), 3049-3051; Veldhuyzen, W. F.; Nguyen, Q.; McMaster, G.; Lawrence, D. S., A light-activated probe of intracellular protein kinase activity. *Journal of the American Chemical Society* **2003,** *125* (44), 13358-13359; Vázquez, M. E.; Nitz, M.; Stehn, J.; Yaffe, M. B.; Imperiali, B., Fluorescent caged phosphoserine peptides as probes to investigate phosphorylation-dependent protein associations. *Journal of the American Chemical Society* **2003,** *125* (34), 10150-10151.

95.   Turner, A. D.; Pizzo, S. V.; Rozakis, G.; Porter, N. A., Photoreactivation of irreversibly inhibited serine proteinases. *Journal of the American Chemical Society* **1988,** *110* (1), 244-250; Stoddard, B. L.; Koenigs, P.; Porter, N.; Petratos, K.; Petsko, G. A.; Ringe, D., Observation of the light-triggered binding of pyrone to chymotrypsin by Laue x-ray crystallography. *Proceedings of the National Academy of Sciences* **1991,** *88* (13), 5503-5507; Marriott, G., Caged protein conjugates and light-directed generation of protein activity: preparation, photoactivation, and spectroscopic characterization of caged G-actin conjugates. *Biochemistry* **1994,** *33* (31), 9092-7; Pan, P.; Bayley, H., Caged cysteine and thiophosphoryl peptides. *FEBS Lett* **1997,** *405* (1), 81-5.

96.   Ghosh, M.; Ichetovkin, I.; Song, X.; Condeelis, J. S.; Lawrence, D. S., A new strategy for caging proteins regulated by kinases. *Journal of the American Chemical Society* **2002,** *124* (11), 2440-2441.

97.   Condeelis, J., Life at the leading edge: the formation of cell protrusions. *Annual review of cell biology* **1993,** *9* (1), 411-444.

98.   Arber, S.; Barbayannis, F. A.; Hanser, H.; Schneider, C.; Stanyon, C. A.; Bernard, O.; Caroni, P., Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase. *Nature* **1998,** *393* (6687), 805-809.

99.   Bamburg, J. R.; McGough, A.; Ono, S., Putting a new twist on actin: ADF/cofilins modulate actin dynamics. *Trends in cell biology* **1999,** *9* (9), 364-370.

100.   Tatsu, Y.; Shigeri, Y.; Sogabe, S.; Yumoto, N.; Yoshikawa, S., Solid-phase synthesis of caged peptides using tyrosine modified with a photocleavable protecting group: application to the synthesis of caged neuropeptide Y. *Biochem Biophys Res Commun* **1996,** *227* (3), 688-93; Sreekumar, R.; Ikebe, M.; Fay, F. S.; Walker, J. W., Biologically active peptides caged on tyrosine. *Methods Enzymol* **1998,** *291*, 78-94; Walker, J. W.; Gilbert, S. H.; Drummond, R. M.; Yamada, M.; Sreekumar, R.; Carraway, R. E.; Ikebe, M.; Fay, F. S., Signaling pathways underlying eosinophil cell motility revealed by using caged peptides. *Proc Natl Acad Sci U S A* **1998,** *95* (4), 1568-73; Curley, K.; Lawrence, D. S., Photoactivation of a signal transduction pathway in living cells. *Journal of the American Chemical Society* **1998,** *120* (33), 8573-8574; Tatsu, Y.; Shigeri, Y.; Ishida, A.;

Kameshita, I.; Fujisawa, H.; Yumoto, N., Synthesis of caged peptides using caged lysine: application to the synthesis of caged AIP, a highly specific inhibitor of calmodulin-dependent protein kinase II. *Bioorg Med Chem Lett* **1999,** *9* (8), 1093-6; Marriott, G.; Walker, J. W., Caged peptides and proteins: new probes to study polypeptide function in complex biological systems. *Trends Plant Sci* **1999,** *4* (8), 330-334; Humphrey, D.; Rajfur, Z.; Vazquez, M. E.; Scheswohl, D.; Schaller, M. D.; Jacobson, K.; Imperiali, B., In situ photoactivation of a caged phosphotyrosine peptide derived from focal adhesion kinase temporarily halts lamellar extension of single migrating tumor cells. *Journal of Biological Chemistry* **2005,** *280* (23), 22091-22101.

101.   Dawson, P. E.; Kent, S. B. H., Synthesis of native proteins by chemical ligation 1. *Annual review of biochemistry* **2000,** *69* (1), 923-960; Hahn, M. E.; Muir, T. W., Manipulating proteins with chemistry: a cross-section of chemical biology. *Trends in biochemical sciences* **2005,** *30* (1), 26-34; Pellois, J. P.; Muir, T. W., A ligation and photorelease strategy for the temporal and spatial control of protein function in living cells. *Angewandte Chemie* **2005,** *117* (35), 5859-5863.

102.   Hahn, M. E.; Muir, T. W., Photocontrol of Smad2, a Multiphosphorylated Cell-Signaling Protein, through Caging of Activating Phosphoserines. *Angewandte Chemie International Edition* **2004,** *43* (43), 5800-5803.

103.   Hunter, T., Signaling—2000 and beyond. *Cell* **2000,** *100* (1), 113-127; Siegel, P. M.; Massagué, J., Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. *Nature Reviews Cancer* **2003,** *3* (11), 807-820.

104.   Wu, J.-W.; Hu, M.; Chai, J.; Seoane, J.; Huse, M.; Li, C.; Rigotti, D. J.; Kyin, S.; Muir, T. W.; Fairman, R., Crystal structure of a phosphorylated Smad2: Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-β signaling. *Molecular cell* **2001,** *8* (6), 1277-1289.

105.   Rothman, D. M.; Vázquez, M. E.; Vogel, E. M.; Imperiali, B., General method for the synthesis of caged phosphopeptides: tools for the exploration of signal transduction pathways. *Organic letters* **2002,** *4* (17), 2865-2868.

106.   Muir, T. W., Semisynthesis of proteins by expressed protein ligation. *Annual review of biochemistry* **2003,** *72* (1), 249-289.

107.   Gautier, A.; Gauron, C.; Volovitch, M.; Bensimon, D.; Jullien, L.; Vriz, S., How to control proteins with light in living systems. *Nat Chem Biol* **2014,** *10* (7), 533-41; Priestman, M. A.; Lawrence, D. S., Light-mediated remote control of signaling pathways. *Biochim Biophys Acta* **2010,** *1804* (3), 547-58.

108.   Wang, L.; Xie, J.; Schultz, P. G., Expanding the genetic code. *Annu Rev Biophys Biomol Struct* **2006,** *35*, 225-49.

109.   Wang, Q.; Parrish, A. R.; Wang, L., Expanding the genetic code for biological studies. *Chem Biol* **2009,** *16* (3), 323-36; Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010,** *79*, 413-44.

110.   Young, T. S.; Schultz, P. G., Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J Biol Chem* **2010,** *285* (15), 11039-44.

111.   Mendel, D.; Ellman, J. A.; Schultz, P. G., Construction of a light-activated protein by unnatural amino acid mutagenesis. *Journal of the American Chemical Society* **1991,** *113* (7), 2758-2760; England, P. M.; Lester, H. A.; Davidson, N.; Dougherty, D. A., Site-specific, photochemical proteolysis applied to ion channels in vivo. *Proceedings of the National Academy of Sciences* **1997,** *94* (20), 11025-11030; Philipson, K. D.; Gallivan, J.

P.; Brandt, G. S.; Dougherty, D. A.; Lester, H. A., Incorporation of caged cysteine and caged tyrosine into a transmembrane segment of the nicotinic ACh receptor. *Am J Physiol Cell Physiol* **2001,** *281* (1), C195-206.

112.   Cropp, T. A.; Schultz, P. G., An expanding genetic code. *TRENDS in Genetics* **2004,** *20* (12), 625-630; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W., Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol* **2008,** *4* (4), 232-4; Li, W. T.; Mahapatra, A.; Longstaff, D. G.; Bechtel, J.; Zhao, G.; Kang, P. T.; Chan, M. K.; Krzycki, J. A., Specificity of pyrrolysyl-tRNA synthetase for pyrrolysine and pyrrolysine analogs. *J Mol Biol* **2009,** *385* (4), 1156-64.

113.   Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G., A genetically encoded photocaged tyrosine. *Angew Chem Int Ed Engl* **2006,** *45* (17), 2728-31.

114.   Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W., Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc* **2010,** *132* (12), 4086-8.

115.   Luo, J.; Uprety, R.; Naro, Y.; Chou, C.; Nguyen, D. P.; Chin, J. W.; Deiters, A., Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. *J Am Chem Soc* **2014,** *136* (44), 15551-8.

116.   Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G., A genetically encoded photocaged amino acid. *J Am Chem Soc* **2004,** *126* (44), 14306-7.

117.   Uprety, R.; Luo, J.; Liu, J.; Naro, Y.; Samanta, S.; Deiters, A., Genetic encoding of caged cysteine and caged homocysteine in bacterial and mammalian cells. *Chembiochem* **2014,** *15* (12), 1793-9.

118.   Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G., A facile system for encoding unnatural amino acids in mammalian cells. *Angew Chem Int Ed Engl* **2009,** *48* (22), 4052-5; Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W., Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J Am Chem Soc* **2010,** *132* (42), 14819-24.

119.   Lemke, E. A.; Summerer, D.; Geierstanger, B. H.; Brittain, S. M.; Schultz, P. G., Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat Chem Biol* **2007,** *3* (12), 769-72.

120.   Tong, Y.; Brandt, G. S.; Li, M.; Shapovalov, G.; Slimko, E.; Karschin, A.; Dougherty, D. A.; Lester, H. A., Tyrosine decaging leads to substantial membrane trafficking during modulation of an inward rectifier potassium channel. *J Gen Physiol* **2001,** *117* (2), 103-18.

121.   Gautier, A.; Deiters, A.; Chin, J. W., Light-activated kinases enable temporal dissection of signaling networks in living cells. *J Am Chem Soc* **2011,** *133* (7), 2124-7.

122.   Shaul, Y. D.; Seger, R., The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2007,** *1773* (8), 1213-1226.

123.   Khokhlatchev, A. V.; Canagarajah, B.; Wilsbacher, J.; Robinson, M.; Atkinson, M.; Goldsmith, E.; Cobb, M. H., Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation. *Cell* **1998,** *93* (4), 605-615; Rubinfeld, H.; Hanoch, T.; Seger, R., Identification of a cytoplasmic-retention sequence in ERK2. *Journal of Biological Chemistry* **1999,** *274* (43), 30349-30352.

124. Volmat, V.; Camps, M.; Arkinstall, S.; Pouysségur, J.; Lenormand, P., The nucleus, a site for signal termination by sequestration and inactivation of p42/p44 MAP kinases. *Journal of Cell Science* **2001,** *114* (19), 3433-3443; Costa, M.; Marchi, M.; Cardarelli, F.; Roy, A.; Beltram, F.; Maffei, L.; Ratto, G. M., Dynamic regulation of ERK2 nuclear translocation and mobility in living cells. *Journal of cell science* **2006,** *119* (23), 4952-4963.

125. Edwards, W. F.; Young, D. D.; Deiters, A., Light-Activated Cre Recombinase as a Tool for the Spatial and Temporal Control of Gene Function in Mammalian Cells. *Acs Chemical Biology* **2009,** *4* (6), 441-445.

126. Chou, C.; Young, D. D.; Deiters, A., A light-activated DNA polymerase. *Angew Chem Int Ed Engl* **2009,** *48* (32), 5950-3.

127. Chou, C. J.; Young, D. D.; Deiters, A., Photocaged T7 RNA Polymerase for the Light Activation of Transcription and Gene Function in Pro- and Eukaryotic Cells. *Chembiochem* **2010,** *11* (7), 972-977.

128. Chou, C.; Deiters, A., Light-activated gene editing with a photocaged zinc-finger nuclease. *Angew Chem Int Ed Engl* **2011,** *50* (30), 6839-42.

129. Arbely, E.; Torres-Kolbus, J.; Deiters, A.; Chin, J. W., Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *J Am Chem Soc* **2012,** *134* (29), 11912-5.

130. Kang, J. Y.; Kawaguchi, D.; Coin, I.; Xiang, Z.; O'Leary, D. D.; Slesinger, P. A.; Wang, L., In vivo expression of a light-activatable potassium channel using unnatural amino acids. *Neuron* **2013,** *80* (2), 358-70.

131. Ren, W.; Ji, A.; Ai, H. W., Light activation of protein splicing with a photocaged fast intein. *J Am Chem Soc* **2015,** *137* (6), 2155-8.

132. Seidman, M. M.; Glazer, P. M., The potential for gene repair via triple helix formation. *J Clin Invest* **2003,** *112* (4), 487-94.

133. Vasquez, K. M.; Glazer, P. M., Triplex-forming oligonucleotides: principles and applications. *Q Rev Biophys* **2002,** *35* (1), 89-107; Simon, P.; Cannata, F.; Concordet, J. P.; Giovannangeli, C., Targeting DNA with triplex-forming oligonucleotides to modify gene sequence. *Biochimie* **2008,** *90* (8), 1109-16; Jain, A.; Wang, G.; Vasquez, K. M., DNA triple helices: biological consequences and therapeutic potential. *Biochimie* **2008,** *90* (8), 1117-30.

134. Jain, A.; Magistri, M.; Napoli, S.; Carbone, G. M.; Catapano, C. V., Mechanisms of triplex DNA-mediated inhibition of transcription initiation in cells. *Biochimie* **2010,** *92* (3), 317-20.

135. Diviacco, S.; Rapozzi, V.; Xodo, L.; Helene, C.; Quadrifoglio, F.; Giovannangeli, C., Site-directed inhibition of DNA replication by triple helix formation. *FASEB J* **2001,** *15* (14), 2660-8.

136. Guntaka, R. V.; Varma, B. R.; Weber, K. T., Triplex-forming oligonucleotides as modulators of gene expression. *Int J Biochem Cell Biol* **2003,** *35* (1), 22-31.

137. Abdelgany, A.; Wood, M.; Beeson, D., Hairpin DNAzymes: a new tool for efficient cellular gene silencing. *J Gene Med* **2007,** *9* (8), 727-38.

138. Herber, B.; Truss, M.; Beato, M.; Müller, R., Inducible regulatory elements in the human cyclin D1 promoter. *Oncogene* **1994,** *9* (7), 2105-7.

139. Carthew, R. W., Gene regulation by microRNAs. *Curr Opin Genet Dev* **2006,** *16* (2), 203-8.

140. Lewis, B. P.; Burge, C. B.; Bartel, D. P., Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005,** *120* (1), 15-20.

141. Sevignani, C.; Calin, G. A.; Siracusa, L. D.; Croce, C. M., Mammalian microRNAs: a small world for fine-tuning gene expression. *Mamm Genome* **2006,** *17* (3), 189-202.

142. Tong, A. W.; Nemunaitis, J., Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? *Cancer Gene Ther* **2008,** *15* (6), 341-55; Lindsay, M. A., microRNAs and the immune response. *Trends Immunol* **2008,** *29* (7), 343-51; Port, J. D.; Sucharov, C., Role of microRNAs in cardiovascular disease: therapeutic challenges and potentials. *J Cardiovasc Pharmacol* **2010,** *56* (5), 444-53; Cullen, B. R., Viruses and microRNAs: RISCy interactions with serious consequences. *Genes Dev* **2011,** *25* (18), 1881-94.

143. Sayed, D.; Abdellatif, M., MicroRNAs in development and disease. *Physiol Rev* **2011,** *91* (3), 827-87.

144. Esau, C. C., Inhibition of microRNA with antisense oligonucleotides. *Methods* **2008,** *44* (1), 55-60.

145. Liu, Z.; Sall, A.; Yang, D., MicroRNA: An emerging therapeutic target and intervention tool. *Int J Mol Sci* **2008,** *9* (6), 978-99.

146. Schetter, A. J.; Leung, S. Y.; Sohn, J. J.; Zanetti, K. A.; Bowman, E. D.; Yanaihara, N.; Yuen, S. T.; Chan, T. L.; Kwong, D. L.; Au, G. K.; Liu, C. G.; Calin, G. A.; Croce, C. M.; Harris, C. C., MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* **2008,** *299* (4), 425-36; Medina, P. P.; Nolde, M.; Slack, F. J., OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature* **2010,** *467* (7311), 86-90.

147.   .

148. Jopling, C. L.; Yi, M.; Lancaster, A. M.; Lemon, S. M.; Sarnow, P., Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **2005,** *309* (5740), 1577-81.

149. Gumireddy, K.; Young, D. D.; Xiong, X.; Hogenesch, J. B.; Huang, Q.; Deiters, A., Small-molecule inhibitors of microrna miR-21 function. *Angew Chem Int Ed Engl* **2008,** *47* (39), 7482-4.

150. Young, D. D.; Connelly, C. M.; Grohmann, C.; Deiters, A., Small molecule modifiers of microRNA miR-122 function for the treatment of hepatitis C virus infection and hepatocellular carcinoma. *J Am Chem Soc* **2010,** *132* (23), 7976-81.

151. Marín, O.; Rubenstein, J. L., Cell migration in the forebrain. *Annu Rev Neurosci* **2003,** *26*, 441-83; Bielas, S.; Higginbotham, H.; Koizumi, H.; Tanaka, T.; Gleeson, J. G., Cortical neuronal migration mutants suggest separate but intersecting pathways. *Annu Rev Cell Dev Biol* **2004,** *20*, 593-618; Casanova, M. F.; Trippe, J., Regulatory mechanisms of cortical laminar development. *Brain Res Rev* **2006,** *51* (1), 72-84; Mandel, G.; Fiondella, C. G.; Covey, M. V.; Lu, D. D.; Loturco, J. J.; Ballas, N., Repressor element 1 silencing transcription factor (REST) controls radial migration and temporal neuronal specification during neocortical development. *Proc Natl Acad Sci U S A* **2011,** *108* (40), 16789-94.

152. Heng, J. I.; Chariot, A.; Nguyen, L., Molecular layers underlying cytoskeletal remodelling during cortical development. *Trends Neurosci* **2010,** *33* (1), 38-47.

153. Gupta, A.; Tsai, L. H.; Wynshaw-Boris, A., Life is a journey: a genetic look at neocortical development. *Nat Rev Genet* **2002,** *3* (5), 342-55.

154. Noctor, S. C.; Martínez-Cerdeño, V.; Ivic, L.; Kriegstein, A. R., Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci* **2004,** *7* (2), 136-44.

155. Shi, Y.; Zhao, X.; Hsieh, J.; Wichterle, H.; Impey, S.; Banerjee, S.; Neveu, P.; Kosik, K. S., MicroRNA regulation of neural stem cells and neurogenesis. *J Neurosci* **2010,** *30* (45), 14931-6; Lang, M. F.; Shi, Y., Dynamic Roles of microRNAs in Neurogenesis. *Front Neurosci* **2012,** *6*, 71; Kawahara, H.; Imai, T.; Okano, H., MicroRNAs in Neural Stem Cells and Neurogenesis. *Front Neurosci* **2012,** *6*, 30.

156. Volvert, M. L.; Rogister, F.; Moonen, G.; Malgrange, B.; Nguyen, L., MicroRNAs tune cerebral cortical neurogenesis. *Cell Death Differ* **2012,** *19* (10), 1573-81.

157. Davis, T. H.; Cuellar, T. L.; Koch, S. M.; Barker, A. J.; Harfe, B. D.; McManus, M. T.; Ullian, E. M., Conditional loss of Dicer disrupts cellular and tissue morphogenesis in the cortex and hippocampus. *J Neurosci* **2008,** *28* (17), 4322-30; Kawase-Koga, Y.; Otaegi, G.; Sun, T., Different timings of Dicer deletion affect neurogenesis and gliogenesis in the developing mouse central nervous system. *Dev Dyn* **2009,** *238* (11), 2800-12; Nowakowski, T. J.; Mysiak, K. S.; Pratt, T.; Price, D. J., Functional dicer is necessary for appropriate specification of radial glia during early development of mouse telencephalon. *PLoS One* **2011,** *6* (8), e23013.

158. Andrés, M. E.; Burger, C.; Peral-Rubio, M. J.; Battaglioli, E.; Anderson, M. E.; Grimes, J.; Dallman, J.; Ballas, N.; Mandel, G., CoREST: a functional corepressor required for regulation of neural-specific gene expression. *Proc Natl Acad Sci U S A* **1999,** *96* (17), 9873-8.

159. Kamensek, U.; Sersa, G.; Vidic, S.; Tevz, G.; Kranjc, S.; Cemazar, M., Irradiation, cisplatin, and 5-azacytidine upregulate cytomegalovirus promoter in tumors and muscles: implementation of non-invasive fluorescence imaging. *Mol Imaging Biol* **2011,** *13* (1), 43-52.

160. Isomura, H.; Stinski, M. F.; Kudoh, A.; Nakayama, S.; Murata, T.; Sato, Y.; Iwahori, S.; Tsurumi, T., A cis element between the TATA Box and the transcription start site of the major immediate-early promoter of human cytomegalovirus determines efficiency of viral replication. *J Virol* **2008,** *82* (2), 849-58.

161. Alberini, C. M., Transcription factors in long-term memory and synaptic plasticity. *Physiol Rev* **2009,** *89* (1), 121-45.

162. Yuan, B.; O'Connor, T. R.; Wang, Y., 6-Thioguanine and S^6-methylthioguanine are mutagenic in human cells. *ACS Chem Biol* **2010,** *5* (11), 1021-7.

163. Lusic, H.; Young, D. D.; Lively, M. O.; Deiters, A., Photochemical DNA activation. *Org Lett* **2007,** *9* (10), 1903-6.

164. Reiter, W. D.; Hüdepohl, U.; Zillig, W., Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. *Proc Natl Acad Sci U S A* **1990,** *87* (24), 9509-13.

165. Young, D. D.; Edwards, W. F.; Lusic, H.; Lively, M. O.; Deiters, A., Light-triggered polymerase chain reaction. *Chem Commun (Camb)* **2008,** (4), 462-4.

166. Barr, F. A.; Silljé, H. H.; Nigg, E. A., Polo-like kinases and the orchestration of cell division. *Nat Rev Mol Cell Biol* **2004,** *5* (6), 429-40.

167. Xu, D.; Wang, Q.; Jiang, Y.; Zhang, Y.; Vega-Saenzdemiera, E.; Osman, I.; Dai, W., Roles of Polo-like kinase 3 in suppressing tumor angiogenesis. *Exp Hematol Oncol* **2012,** *1* (1), 5.

168. Jiang, N.; Wang, X.; Jhanwar-Uniyal, M.; Darzynkiewicz, Z.; Dai, W., Polo box domain of Plk3 functions as a centrosome localization signal, overexpression of which causes mitotic arrest, cytokinesis defects, and apoptosis. *J Biol Chem* **2006,** *281* (15), 10577-82.

169. Wang, Q.; Xie, S.; Chen, J.; Fukasawa, K.; Naik, U.; Traganos, F.; Darzynkiewicz, Z.; Jhanwar-Uniyal, M.; Dai, W., Cell cycle arrest and apoptosis induced by human Polo-like kinase 3 is mediated through perturbation of microtubule integrity. *Mol Cell Biol* **2002,** *22* (10), 3450-9; Wang, L.; Gao, J.; Dai, W.; Lu, L., Activation of Polo-like kinase 3 by hypoxic stresses. *J Biol Chem* **2008,** *283* (38), 25928-35.

170. Holder, N.; Xu, Q., Microinjection of DNA, RNA, and protein into the fertilized zebrafish egg for analysis of gene function. *Methods Mol Biol* **1999,** *97*, 487-90; Xu, Q., Microinjection into zebrafish embryos. *Methods Mol Biol* **1999,** *127*, 125-32.

171. Yamaguchi, S.; Chen, Y.; Nakajima, S.; Furuta, T.; Nagamune, T., Light-activated gene expression from site-specific caged DNA with a biotinylated photolabile protection group. *Chem Commun (Camb)* **2010,** *46* (13), 2244-6.

172. Scherrer, K.; Imaizumi-Scherrer, M. T.; Reynaud, C. A.; Therwath, A., On pre-messenger RNA and transcriptions. A review. *Mol Biol Rep* **1979,** *5* (1-2), 5-28.

173. Lamond, A. I., The spliceosome. *Bioessays* **1993,** *15* (9), 595-603.

174. Leff, S. E.; Rosenfeld, M. G.; Evans, R. M., Complex transcriptional units: diversity in gene expression by alternative RNA processing. *Annu Rev Biochem* **1986,** *55*, 1091-117.

175. Pan, Q.; Shai, O.; Lee, L. J.; Frey, B. J.; Blencowe, B. J., Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **2008,** *40* (12), 1413-5; Wang, E. T.; Sandberg, R.; Luo, S.; Khrebtukova, I.; Zhang, L.; Mayr, C.; Kingsmore, S. F.; Schroth, G. P.; Burge, C. B., Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008,** *456* (7221), 470-6; Ozsolak, F.; Milos, P. M., RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* **2011,** *12* (2), 87-98.

176. Cartegni, L.; Chew, S. L.; Krainer, A. R., Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **2002,** *3* (4), 285-98.

177. Tazi, J.; Durand, S.; Jeanteur, P., The spliceosome: a novel multi-faceted target for therapy. *Trends Biochem Sci* **2005,** *30* (8), 469-78.

178. Sazani, P.; Kole, R., Modulation of alternative splicing by antisense oligonucleotides. *Prog Mol Subcell Biol* **2003,** *31*, 217-39; Sazani, P.; Kole, R., Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *J Clin Invest* **2003,** *112* (4), 481-6.

179. van Deutekom, J. C.; Janson, A. A.; Ginjaar, I. B.; Frankhuizen, W. S.; Aartsma-Rus, A.; Bremmer-Bout, M.; den Dunnen, J. T.; Koop, K.; van der Kooi, A. J.; Goemans, N. M.; de Kimpe, S. J.; Ekhart, P. F.; Venneker, E. H.; Platenburg, G. J.; Verschuuren, J. J.; van Ommen, G. J., Local dystrophin restoration with antisense oligonucleotide PRO051. *N Engl J Med* **2007,** *357* (26), 2677-86; Kinali, M.; Arechavala-Gomeza, V.; Feng, L.; Cirak, S.; Hunt, D.; Adkin, C.; Guglieri, M.; Ashton, E.; Abbs, S.; Nihoyannopoulos, P.; Garralda, M. E.; Rutherford, M.; McCulley, C.; Popplewell, L.; Graham, I. R.; Dickson, G.; Wood, M. J.; Wells, D. J.; Wilton, S. D.; Kole, R.; Straub, V.; Bushby, K.; Sewry, C.; Morgan, J. E.; Muntoni, F., Local restoration of dystrophin expression with the

morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *Lancet Neurol* **2009,** *8* (10), 918-28; Goemans, N. M.; Tulinius, M.; van den Akker, J. T.; Burm, B. E.; Ekhart, P. F.; Heuvelmans, N.; Holling, T.; Janson, A. A.; Platenburg, G. J.; Sipkens, J. A.; Sitsen, J. M.; Aartsma-Rus, A.; van Ommen, G. J.; Buyse, G.; Darin, N.; Verschuuren, J. J.; Campion, G. V.; de Kimpe, S. J.; van Deutekom, J. C., Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N Engl J Med* **2011,** *364* (16), 1513-22; Cirak, S.; Arechavala-Gomeza, V.; Guglieri, M.; Feng, L.; Torelli, S.; Anthony, K.; Abbs, S.; Garralda, M. E.; Bourke, J.; Wells, D. J.; Dickson, G.; Wood, M. J.; Wilton, S. D.; Straub, V.; Kole, R.; Shrewsbury, S. B.; Sewry, C.; Morgan, J. E.; Bushby, K.; Muntoni, F., Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet* **2011,** *378* (9791), 595-605; Mendell, J. R.; Rodino-Klapac, L. R.; Sahenk, Z.; Roush, K.; Bird, L.; Lowes, L. P.; Alfano, L.; Gomez, A. M.; Lewis, S.; Kota, J.; Malik, V.; Shontz, K.; Walker, C. M.; Flanigan, K. M.; Corridore, M.; Kean, J. R.; Allen, H. D.; Shilling, C.; Melia, K. R.; Sazani, P.; Saoud, J. B.; Kaye, E. M.; Group, E. S., Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* **2013,** *74* (5), 637-47.

180. Grabowski, P., Alternative splicing takes shape during neuronal development. *Curr Opin Genet Dev* **2011,** *21* (4), 388-94; Grabowski, P. J.; Black, D. L., Alternative RNA splicing in the nervous system. *Prog Neurobiol* **2001,** *65* (3), 289-308.

181. Goodwin, S. F.; Taylor, B. J.; Villella, A.; Foss, M.; Ryner, L. C.; Baker, B. S.; Hall, J. C., Aberrant splicing and altered spatial expression patterns in fruitless mutants of Drosophila melanogaster. *Genetics* **2000,** *154* (2), 725-45.

182. Strehler, E. E.; Zacharias, D. A., Role of alternative splicing in generating isoform diversity among plasma membrane calcium pumps. *Physiol Rev* **2001,** *81* (1), 21-50.

183. Ott, E. B.; Te Velthuis, A. J.; Bagowski, C. P., Comparative analysis of splice form-specific expression of LIM Kinases during zebrafish development. *Gene Expr Patterns* **2007,** *7* (5), 620-9; Rissone, A.; Sangiorgio, L.; Monopoli, M.; Beltrame, M.; Zucchi, I.; Bussolino, F.; Arese, M.; Cotelli, F., Characterization of the neuroligin gene family expression and evolution in zebrafish. *Dev Dyn* **2010,** *239* (2), 688-702.

184. Mullen, G. P.; Rogalski, T. M.; Bush, J. A.; Gorji, P. R.; Moerman, D. G., Complex patterns of alternative splicing mediate the spatial and temporal distribution of perlecan/UNC-52 in Caenorhabditis elegans. *Mol Biol Cell* **1999,** *10* (10), 3205-21.

185. Nilsen, T. W.; Graveley, B. R., Expansion of the eukaryotic proteome by alternative splicing. *Nature* **2010,** *463* (7280), 457-63.

186. Dominski, Z.; Kole, R., Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc Natl Acad Sci U S A* **1993,** *90* (18), 8673-7; Ming, X.; Alam, M. R.; Fisher, M.; Yan, Y.; Chen, X.; Juliano, R. L., Intracellular delivery of an antisense oligonucleotide via endocytosis of a G protein-coupled receptor. *Nucleic Acids Res* **2010,** *38* (19), 6567-76; Nakagawa, O.; Ming, X.; Huang, L.; Juliano, R. L., Targeted intracellular delivery of antisense oligonucleotides via conjugation with small-molecule ligands. *J Am Chem Soc* **2010,** *132* (26), 8848-9; Kotula, J. W.; Pratico, E. D.; Ming, X.; Nakagawa, O.; Juliano, R. L.; Sullenger, B. A., Aptamer-mediated delivery of splice-switching oligonucleotides to the nuclei of cancer cells. *Nucleic Acid Ther* **2012,** *22* (3), 187-95.

187. Ming, X.; Carver, K.; Fisher, M.; Noel, R.; Cintrat, J. C.; Gillet, D.; Barbier, J.; Cao, C.; Bauman, J.; Juliano, R. L., The small molecule Retro-1 enhances the pharmacological actions of antisense and splice switching oligonucleotides. *Nucleic Acids Res* **2013,** *41* (6), 3673-87.

188. Sazani, P.; Kang, S. H.; Maier, M. A.; Wei, C.; Dillman, J.; Summerton, J.; Manoharan, M.; Kole, R., Nuclear antisense effects of neutral, anionic and cationic oligonucleotide analogs. *Nucleic Acids Res* **2001,** *29* (19), 3965-74.

189. Resina, S.; Kole, R.; Travo, A.; Lebleu, B.; Thierry, A. R., Switching on transgene expression by correcting aberrant splicing using multi-targeting steric-blocking oligonucleotides. *J Gene Med* **2007,** *9* (6), 498-510.

190. Dong, Q.; Svoboda, K.; Tiersch, T.; Monroe, W., Photobiological effects of UVA and UVB light in zebrafish embryos: Evidence for a competent photorepair system. *J Photochem Photobio B* **2007,** *88* (2-3), 137-146.

191. Okuda, Y.; Yoda, H.; Uchikawa, M.; Furutani-Seiki, M.; Takeda, H.; Kondoh, H.; Kamachi, Y., Comparative genomic and expression analysis of group B1 sox genes in zebrafish indicates their diversification during vertebrate evolution. *Dev Dyn* **2006,** *235* (3), 811-25; Okuda, Y.; Ogura, E.; Kondoh, H.; Kamachi, Y., B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. *PLoS Genet* **2010,** *6* (5), e1000936; Hu, S.; Wu, Z.; Yan, Y.; Li, Y., Sox31 is involved in central nervous system anteroposterior regionalization through regulating the organizer activity in zebrafish. *Acta Biochim Biophys Sin (Shanghai)* **2011,** *43* (5), 387-99.

192. Girard, F.; Crémazy, F.; Berta, P.; Renucci, A., Expression pattern of the Sox31 gene during zebrafish embryonic development. *Mech Dev* **2001,** *100* (1), 71-3.

193. Hu, S. N.; Yu, H.; Zhang, Y. B.; Wu, Z. L.; Yan, Y. C.; Li, Y. X.; Li, Y. Y.; Li, Y. P., Splice blocking of zygotic sox31 leads to developmental arrest shortly after Mid-Blastula Transition and induces apoptosis in zebrafish. *FEBS Lett* **2012,** *586* (3), 222-8.

194. Schneider, P. N.; Olthoff, J. T.; Matthews, A. J.; Houston, D. W., Use of fully modified 2'-O-methyl antisense oligos for loss-of-function studies in vertebrate embryos. *Genesis* **2011,** *49* (3), 117-23.

195. Adleman, L. M., Molecular computation of solutions to combinatorial problems. *Science* **1994,** *266* (5187), 1021-4.

196. Lai, Y. H.; Sun, S. C.; Chuang, M. C., Biosensors with built-in biomolecular logic gates for practical applications. *Biosensors (Basel)* **2014,** *4* (3), 273-300; Han, D.; Kang, H.; Zhang, T.; Wu, C.; Zhou, C.; You, M.; Chen, Z.; Zhang, X.; Tan, W., Nucleic acid based logical systems. *Chemistry* **2014,** *20* (20), 5866-73.

197. Parker, J., Computing with DNA. *EMBO Rep* **2003,** *4* (1), 7-10.

198. Benenson, Y., Biomolecular computing systems: principles, progress and potential. *Nat Rev Genet* **2012,** *13* (7), 455-68.

199. Miyamoto, T.; Razavi, S.; DeRose, R.; Inoue, T., Synthesizing Biomolecuale-Based Boolean Logic Gates. *ACS Synthetic Biology* **2012,** *Online*.

200. Ma, D.-L.; He, H.-Z.; Chan, D. S.-H.; Leung, C.-H., Simple DNA-based logic gates responding to biomolecules and metal ions. *Chemical Science* **2013,** *4* (9), 3366-3380; Pu, F.; Ren, J.; Qu, X., Nucleic acids and smart materials: advanced building blocks for logic systems. *Adv Mater* **2014,** *26* (33), 5742-57.

201. Church, G. M.; Gao, Y.; Kosuri, S., Next-generation digital information storage in DNA. *Science* **2012,** *337* (6102), 1628-1628.

202. Chen, X.; Ellington, A. D., Shaping up nucleic acid computation. *Curr Opin Biotechnol* **2010,** *21* (4), 392-400; Krishnan, Y.; Simmel, F. C., Nucleic acid based molecular devices. *Angew Chem Int Ed Engl* **2011,** *50* (14), 3124-56; Jung, C.; Ellington, A. D., Diagnostic applications of nucleic acid circuits. *Acc Chem Res* **2014,** *47* (6), 1825-35.

203. Zhang, D. Y.; Winfree, E., Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society* **2009,** *131* (47), 17303-17314.

204. Srinivas, N.; Ouldridge, T. E.; Šulc, P.; Schaeffer, J. M.; Yurke, B.; Louis, A. A.; Doye, J. P. K.; Winfree, E., On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic acids research* **2013,** *41* (22), 10641-10658.

205. Yin, P.; Choi, H. M.; Calvert, C. R.; Pierce, N. A., Programming biomolecular self-assembly pathways. *Nature* **2008,** *451* (7176), 318-22.

206. Zhang, D. Y.; Turberfield, A. J.; Yurke, B.; Winfree, E., Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **2007,** *318* (5853), 1121-5.

207. Qian, L.; Winfree, E., Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **2011,** *332* (6034), 1196-201.

208. Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E., Enzyme-free nucleic acid logic circuits. *Science* **2006,** *314* (5805), 1585-8.

209. Qian, L.; Winfree, E.; Bruck, J., Neural network computation with DNA strand displacement cascades. *Nature* **2011,** *475* (7356), 368-72.

210. Lederman, H.; Macdonald, J.; Stefanovic, D.; Stojanovic, M. N., Deoxyribozyme-based three-input logic gates and construction of a molecular full adder. *Biochemistry* **2006,** *45* (4), 1194-1199.

211. Zhu, J.; Zhang, L.; Li, T.; Dong, S.; Wang, E., Enzyme-Free Unlabeled DNA Logic Circuits Based on Toehold-Mediated Strand Displacement and Split G-Quadruplex Enhanced Fluorescence. *Advanced Materials* **2013,** *25* (17), 2440-2444.

212. Yoshida, W.; Yokobayashi, Y., Photonic Boolean logic gates based on DNA aptamers. *Chem Commun (Camb)* **2007,** (2), 195-7.

213. Anderson, J. C.; Voigt, C. A.; Arkin, A. P., Environmental signal integration by a modular AND gate. *Mol Syst Biol* **2007,** *3*, 133.

214. Lake, A.; Shang, S.; Kolpashchikov, D. M., Molecular logic gates connected through DNA four-way junctions. *Angew Chem Int Ed Engl* **2010,** *49* (26), 4459-62.

215. Tabor, J. J.; Salis, H. M.; Simpson, Z. B.; Chevalier, A. A.; Levskaya, A.; Marcotte, E. M.; Voigt, C. A.; Ellington, A. D., A synthetic genetic edge detection program. *Cell* **2009,** *137* (7), 1272-81.

216. Stojanovic, M. N.; Stefanovic, D., A deoxyribozyme-based molecular automaton. *Nat Biotechnol* **2003,** *21* (9), 1069-74.

217. Douglas, S. M.; Bachelet, I.; Church, G. M., A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **2012,** *335* (6070), 831-4.

218. Wu, C.; Cansiz, S.; Zhang, L.; Teng, I. T.; Qiu, L.; Li, J.; Liu, Y.; Zhou, C.; Hu, R.; Zhang, T.; Cui, C.; Cui, L.; Tan, W., A Nonenzymatic Hairpin DNA Cascade Reaction Provides High Signal Gain of mRNA Imaging inside Live Cells. *J Am Chem Soc* **2015,** *137* (15), 4900-3.

219. You, M.; Zhu, G.; Chen, T.; Donovan, M. J.; Tan, W., Programmable and multiparameter DNA-based logic platform for cancer recognition and targeted therapy. *J Am Chem Soc* **2015,** *137* (2), 667-74.

220. Kahan-Hanum, M.; Douek, Y.; Adar, R.; Shapiro, E., A library of programmable DNAzymes that operate in a cellular environment. *Sci. Rep.* **2013,** *3*.

221. Dass, C. R.; Choong, P. F. M., C-jun: pharmaceutical target for DNAzyme therapy of multiple pathologies. *Die Pharmazie-An International Journal of Pharmaceutical Sciences* **2008,** *63* (6), 411-414.

222. Surana, S.; Bhat, J. M.; Koushika, S. P.; Krishnan, Y., An autonomous DNA nanomachine maps spatiotemporal pH changes in a multicellular living organism. *Nat Commun* **2011,** *2*, 340.

223. Altun, Z. F.; Hall, D., Handbook of C. elegans anatomy, 2008. *WormAtlas*.

224. Modi, S.; Swetha, M. G.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y., A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nature Nanotechnology* **2009,** *4* (5), 325-330.

225. Guéron, M.; Leroy, J.-L., The i-motif in nucleic acids. *Current opinion in structural biology* **2000,** *10* (3), 326-331.

226. Fares, H.; Greenwald, I., Genetic analysis of endocytosis in Caenorhabditis elegans: coelomocyte uptake defective mutants. *Genetics* **2001,** *159* (1), 133-145.

227. Benenson, Y., RNA-based computation in live cells. *Curr Opin Biotechnol* **2009,** *20* (4), 471-8.

228. Nielsen, A. A.; Voigt, C. A., Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol* **2014,** *10*, 763.

229. Rinaudo, K.; Bleris, L.; Maddamsetti, R.; Subramanian, S.; Weiss, R.; Benenson, Y., A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol* **2007,** *25* (7), 795-801; Xie, Z.; Wroblewska, L.; Prochazka, L.; Weiss, R.; Benenson, Y., Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* **2011,** *333* (6047), 1307-11.

230. Leisner, M.; Bleris, L.; Lohmueller, J.; Xie, Z.; Benenson, Y., MicroRNA circuits for transcriptional logic. *Methods Mol Biol* **2012,** *813*, 169-86.

231. Cheglakov, Z.; Cronin, T. M.; He, C.; Weizmann, Y., Live Cell MicroRNA Imaging Using Cascade Hybridization Reaction. *J Am Chem Soc* **2015,** *137* (19), 6116-9.

232. Cunningham, C. W.; Mukhopadhyay, A.; Lushington, G. H.; Blagg, B. S.; Prisinzano, T. E.; Krise, J. P., Uptake, distribution and diffusivity of reactive fluorophores in cells: implications toward target identification. *Mol Pharm* **2010,** *7* (4), 1301-10.

233. Shih, W., Biomolecular assembly: dynamic DNA. *Nat Mater* **2008,** *7* (2), 98-100.

234. Wang, F.; Niu, G.; Chen, X.; Cao, F., Molecular imaging of microRNAs. *Eur J Nucl Med Mol Imaging* **2011,** *38* (8), 1572-9.

235. Kang, W. J.; Cho, Y. L.; Chae, J. R.; Lee, J. D.; Choi, K. J.; Kim, S., Molecular beacon-based bioimaging of multiple microRNAs during myogenesis. *Biomaterials* **2011,** *32* (7), 1915-22; Yao, Q.; Zhang, A. M.; Ma, H.; Lin, S.; Wang, X. X.; Sun, J. G.; Chen, Z. T., Novel molecular beacons to monitor microRNAs in non-small-cell lung cancer. *Mol Cell Probes* **2012,** *26* (5), 182-7.

236. Zhu, S.; Wu, H.; Wu, F.; Nie, D.; Sheng, S.; Mo, Y. Y., MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. *Cell Res* **2008,** *18* (3), 350-9.

237. Negrini, M.; Gramantieri, L.; Sabbioni, S.; Croce, C. M., microRNA involvement in hepatocellular carcinoma. *Anticancer Agents Med Chem* **2011,** *11* (6), 500-21.

238. Bustin, S. A., Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* **2002,** *29* (1), 23-39; Bustin, S. A.; Nolan,

T., Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* **2004,** *15* (3), 155-66.

239. Tian, W.; Dong, X.; Liu, X.; Wang, G.; Dong, Z.; Shen, W.; Zheng, G.; Lu, J.; Chen, J.; Wang, Y.; Wu, Z.; Wu, X., High-throughput functional microRNAs profiling by recombinant AAV-based microRNA sensor arrays. *PLoS One* **2012,** *7* (1), e29551.

240. Connelly, C. M.; Thomas, M.; Deiters, A., High-throughput luciferase reporter assay for small-molecule inhibitors of microRNA function. *J Biomol Screen* **2012,** *17* (6), 822-8.

241. Chen, H. L.; Huang, J. Y.; Chen, C. M.; Chu, T. H.; Shih, C., MicroRNA-22 can reduce parathymosin expression in transdifferentiated hepatocytes. *PLoS One* **2012,** *7* (4), e34116.

242. Prokup, A.; Deiters, A., Interfacing Synthetic DNA Logic Operations with Protein Outputs. *Angewandte Chemie* **2014,** *126* (48), 13408-13411.

243. Genot, A. J.; Bath, J.; Turberfield, A. J., Reversible logic circuits made of DNA. *J Am Chem Soc* **2011,** *133* (50), 20080-3.

244. CellMiner. http://discover.nci.nih.gov/cellminer/;
microRNA.org. http://www.microrna.org/microrna/home.do;
microRNA Cancer Association Database. http://mircancer.ecu.edu/;
miRNAMap. http://mirnamap.mbc.nctu.edu.tw/.

245. Gibbings, D. J.; Ciaudo, C.; Erhardt, M.; Voinnet, O., Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. *Nat Cell Biol* **2009,** *11* (9), 1143-9; Kulkarni, M.; Ozgur, S.; Stoecklin, G., On track with P-bodies. *Biochem Soc Trans* **2010,** *38* (Pt 1), 242-51; Gibbings, D.; Voinnet, O., Control of RNA silencing and localization by endolysosomes. *Trends Cell Biol* **2010,** *20* (8), 491-501.

246. Torres, A. G.; Fabani, M. M.; Vigorito, E.; Williams, D.; Al-Obaidi, N.; Wojciechowski, F.; Hudson, R. H.; Seitz, O.; Gait, M. J., Chemical structure requirements and cellular targeting of microRNA-122 by peptide nucleic acids anti-miRs. *Nucleic Acids Res* **2012,** *40* (5), 2152-67.

247. Hoekstra, D.; Rejman, J.; Wasungu, L.; Shi, F.; Zuhorn, I., Gene delivery by cationic lipids: in and out of an endosome. *Biochem Soc Trans* **2007,** *35* (Pt 1), 68-71; Zuhorn, I. S.; Engberts, J. B.; Hoekstra, D., Gene delivery by cationic lipid vectors: overcoming cellular barriers. *Eur Biophys J* **2007,** *36* (4-5), 349-62.

248. Kang, S. H.; Cho, M. J.; Kole, R., Up-regulation of luciferase gene expression with antisense oligonucleotides: implications and applications in functional assay development. *Biochemistry* **1998,** *37* (18), 6235-9.

249. Colosimo, A.; Goncz, K. K.; Holmes, A. R.; Kunzelmann, K.; Novelli, G.; Malone, R. W.; Bennett, M. J.; Gruenert, D. C., Transfer and expression of foreign genes in mammalian cells. *Biotechniques* **2000,** *29* (2), 314-8, 320-2, 324 passim; Pichon, C.; Billiet, L.; Midoux, P., Chemical vectors for gene delivery: uptake and intracellular trafficking. *Curr Opin Biotechnol* **2010,** *21* (5), 640-5.

250. Collins, M. L.; Irvine, B.; Tyner, D.; Fine, E.; Zayati, C.; Chang, C.; Horn, T.; Ahle, D.; Detmer, J.; Shen, L. P.; Kolberg, J.; Bushnell, S.; Urdea, M. S.; Ho, D. D., A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *Nucleic Acids Res* **1997,** *25* (15), 2979-84.

251. Seelig, G.; Yurke, B.; Winfree, E., Catalyzed relaxation of a metastable DNA fuel. *J Am Chem Soc* **2006,** *128* (37), 12211-20.

252. Dirks, R. M.; Pierce, N. A., Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci U S A* **2004,** *101* (43), 15275-8.

253. Choi, H. M.; Beck, V. A.; Pierce, N. A., Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* **2014,** *8* (5), 4284-94.

254. Zhang, D. Y.; Winfree, E., Dynamic allosteric control of noncovalent DNA catalysis reactions. *J Am Chem Soc* **2008,** *130* (42), 13921-6.

255. Choi, H. M.; Chang, J. Y.; Trinh, l. A.; Padilla, J. E.; Fraser, S. E.; Pierce, N. A., Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol* **2010,** *28* (11), 1208-12.

256. Gardner, L.; Deiters, A., Light-controlled synthetic gene circuits. *Curr Opin Chem Biol* **2012,** *16* (3-4), 292-9.

257. Cramer, P., Common structural features of nucleic acid polymerases. *Bioessays* **2002,** *24* (8), 724-9.

258. Moffatt, B. A.; Dunn, J. J.; Studier, F. W., Nucleotide sequence of the gene for bacteriophage T7 RNA polymerase. *J Mol Biol* **1984,** *173* (2), 265-9.

259. Steitz, T. A., Visualizing polynucleotide polymerase machines at work. *EMBO J* **2006,** *25* (15), 3458-68.

260. Maksimova, T. G.; Mustayev, A. A.; Zaychikov, E. F.; Lyakhov, D. L.; Tunitskaya, V. L.; Akbarov, A. K. h.; Luchin, S. V.; Rechinsky, V. O.; Chernov, B. K.; Kochetkov, S. N., Lys631 residue in the active site of the bacteriophage T7 RNA polymerase. Affinity labeling and site-directed mutagenesis. *Eur J Biochem* **1991,** *195* (3), 841-7; Woody, A. Y.; Osumi-Davis, P. A.; Hiremath, M. M.; Woody, R. W., Pre-steady-state and steady-state kinetic studies on transcription initiation catalyzed by T7 RNA polymerase and its active-site mutants K631R and Y639F. *Biochemistry* **1998,** *37* (45), 15958-64.

261. Yin, Y. W.; Steitz, T. A., The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **2004,** *116* (3), 393-404.

262. Osumi-Davis, P. A.; de Aguilera, M. C.; Woody, R. W.; Woody, A. Y., Asp537, Asp812 are essential and Lys631, His811 are catalytically significant in bacteriophage T7 RNA polymerase activity. *J Mol Biol* **1992,** *226* (1), 37-45.

263. Xie, J.; Schultz, P. G., An expanding genetic code. *Methods* **2005,** *36* (3), 227-38.

264. Sharp, P. A.; Zamore, P. D., Molecular biology. RNA interference. *Science* **2000,** *287* (5462), 2431-3.

265. Harborth, J.; Elbashir, S. M.; Vandenburgh, K.; Manninga, H.; Scaringe, S. A.; Weber, K.; Tuschl, T., Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev* **2003,** *13* (2), 83-105.

266. Lambeth, L. S.; Smith, C. A., Short hairpin RNA-mediated gene silencing. *Methods Mol Biol* **2013,** *942*, 205-32.

267. Sui, G.; Soohoo, C.; Affar, e. B.; Gay, F.; Shi, Y.; Forrester, W. C., A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc Natl Acad Sci U S A* **2002,** *99* (8), 5515-20; Paddison, P. J.; Caudy, A. A.; Bernstein, E.; Hannon, G. J.; Conklin, D. S., Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* **2002,** *16* (8), 948-58; Paddison, P. J.; Caudy, A. A.; Hannon, G. J., Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* **2002,** *99* (3), 1443-8; Brummelkamp, T. R.; Bernards, R.; Agami,

R., A system for stable expression of short interfering RNAs in mammalian cells. *Science* **2002,** *296* (5567), 550-3.

268. Yu, J. Y.; DeRuiter, S. L.; Turner, D. L., RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc Natl Acad Sci U S A* **2002,** *99* (9), 6047-52; Peng, Y.; Lu, J. X.; Shen, X. F., shRNA driven by Pol II/T7 dual-promoter system effectively induce cell-specific RNA interference in mammalian cells. *Biochem Biophys Res Commun* **2007,** *360* (2), 496-500.

269. Wang, N.; Sun, Y. H.; Liu, J.; Wu, G.; Su, J. G.; Wang, Y. P.; Zhu, Z. Y., Knock down of gfp and no tail expression in zebrafish embryo by in vivo-transcribed short hairpin RNA with T7 plasmid system. *J Biomed Sci* **2007,** *14* (6), 767-76.

270. Ferenz, N. P.; Gable, A.; Wadsworth, P., Mitotic functions of kinesin-5. *Semin Cell Dev Biol* **2010,** *21* (3), 255-9.

271. Zhu, C.; Zhao, J.; Bibikova, M.; Leverson, J. D.; Bossy-Wetzel, E.; Fan, J. B.; Abraham, R. T.; Jiang, W., Functional analysis of human microtubule-based motor proteins, the kinesins and dyneins, in mitosis/cytokinesis using RNA interference. *Mol Biol Cell* **2005,** *16* (7), 3187-99.

272. Ma, N.; Tulu, U. S.; Ferenz, N. P.; Fagerstrom, C.; Wilde, A.; Wadsworth, P., Poleward transport of TPX2 in the mammalian mitotic spindle requires dynein, Eg5, and microtubule flux. *Mol Biol Cell* **2010,** *21* (6), 979-88.

273. He, B.; Rong, M.; Lyakhov, D.; Gartenstein, H.; Diaz, G.; Castagna, R.; McAllister, W. T.; Durbin, R. K., Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expr Purif* **1997,** *9* (1), 142-51.

274. Harborth, J.; Elbashir, S. M.; Bechert, K.; Tuschl, T.; Weber, K., Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J Cell Sci* **2001,** *114* (Pt 24), 4557-65.

275. Nielsen, R.; Courtoy, P. J.; Jacobsen, C.; Dom, G.; Lima, W. R.; Jadot, M.; Willnow, T. E.; Devuyst, O.; Christensen, E. I., Endocytosis provides a major alternative pathway for lysosomal biogenesis in kidney proximal tubular cells. *Proc Natl Acad Sci U S A* **2007,** *104* (13), 5407-12.

276. Wiedenheft, B.; Sternberg, S. H.; Doudna, J. A., RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012,** *482* (7385), 331-8.

277. Garneau, J. E.; Dupuis, M.; Villion, M.; Romero, D. A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadán, A. H.; Moineau, S., The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010,** *468* (7320), 67-71; Deltcheva, E.; Chylinski, K.; Sharma, C. M.; Gonzales, K.; Chao, Y.; Pirzada, Z. A.; Eckert, M. R.; Vogel, J.; Charpentier, E., CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011,** *471* (7340), 602-7.

278. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012,** *337* (6096), 816-21.

279. Hsu, P. D.; Scott, D. A.; Weinstein, J. A.; Ran, F. A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E. J.; Wu, X.; Shalem, O.; Cradick, T. J.; Marraffini, L. A.; Bao, G.; Zhang, F., DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **2013,** *31* (9), 827-32; Ran, F. A.; Hsu, P. D.; Lin, C. Y.; Gootenberg, J. S.; Konermann, S.; Trevino, A. E.; Scott, D. A.; Inoue, A.; Matoba, S.; Zhang, Y.; Zhang, F., Double nicking by RNA-

guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **2013,** *154* (6), 1380-9.

280. Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E.; Norville, J. E.; Church, G. M., RNA-guided human genome engineering via Cas9. *Science* **2013,** *339* (6121), 823-6.

281. Gratz, S. J.; Cummings, A. M.; Nguyen, J. N.; Hamm, D. C.; Donohue, L. K.; Harrison, M. M.; Wildonger, J.; O'Connor-Giles, K. M., Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **2013,** *194* (4), 1029-35; Wang, H.; Yang, H.; Shivalila, C. S.; Dawlaty, M. M.; Cheng, A. W.; Zhang, F.; Jaenisch, R., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **2013,** *153* (4), 910-8.

282. Strauβ, A.; Lahaye, T., Zinc fingers, TAL effectors, or Cas9-based DNA binding proteins: what's best for targeting desired genome loci? *Mol Plant* **2013,** *6* (5), 1384-7; Niu, J.; Zhang, B.; Chen, H., Applications of TALENs and CRISPR/Cas9 in Human Cells and Their Potentials for Gene Therapy. *Mol Biotechnol* **2014,** *56* (8), 681-8.

283. Gilbert, L. A.; Larson, M. H.; Morsut, L.; Liu, Z.; Brar, G. A.; Torres, S. E.; Stern-Ginossar, N.; Brandman, O.; Whitehead, E. H.; Doudna, J. A.; Lim, W. A.; Weissman, J. S.; Qi, L. S., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **2013,** *154* (2), 442-51.

284. Nihongaki, Y.; Yamamoto, S.; Kawano, F.; Suzuki, H.; Sato, M., CRISPR-Cas9-based Photoactivatable Transcription System. *Chem Biol* **2015,** *22* (2), 169-74; .

285. Nihongaki, Y.; Kawano, F.; Nakajima, T.; Sato, M., Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nature biotechnology* **2015**.

286. Nishimasu, H.; Ran, F. A.; Hsu, P. D.; Konermann, S.; Shehata, S. I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O., Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **2014,** *156* (5), 935-49; Jinek, M.; Jiang, F.; Taylor, D. W.; Sternberg, S. H.; Kaya, E.; Ma, E.; Anders, C.; Hauer, M.; Zhou, K.; Lin, S.; Kaplan, M.; Iavarone, A. T.; Charpentier, E.; Nogales, E.; Doudna, J. A., Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **2014,** *343* (6176), 1247997.

287. De Gasperi, R.; Rocher, A. B.; Sosa, M. A.; Wearne, S. L.; Perez, G. M.; Friedrich, V. L.; Hof, P. R.; Elder, G. A., The IRG mouse: a two-color fluorescent reporter for assessing Cre-mediated recombination and imaging complex cellular relationships in situ. *Genesis* **2008,** *46* (6), 308-17.

288. Verkhusha, V. V.; Kuznetsova, I. M.; Stepanenko, O. V.; Zaraisky, A. G.; Shavlovsky, M. M.; Turoverov, K. K.; Uversky, V. N., High stability of Discosoma DsRed as compared to Aequorea EGFP. *Biochemistry* **2003,** *42* (26), 7879-84.

289. Parenti, R.; Salvatorelli, L.; Magro, G., Anaplastic Thyroid Carcinoma: Current Treatments and Potential New Therapeutic Options with Emphasis on TfR1/CD71. *Int J Endocrinol* **2014,** *2014*, 685396.

290. Sander, J. D.; Joung, J. K., CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **2014,** *32* (4), 347-55.

291. Greiss, S.; Chin, J. W., Expanding the genetic code of an animal. *J Am Chem Soc* **2011,** *133* (36), 14196-9; Bianco, A.; Townsley, F. M.; Greiss, S.; Lang, K.; Chin, J. W., Expanding the genetic code of Drosophila melanogaster. *Nat Chem Biol* **2012,** *8* (9), 748-50.

292. Fu, Y.; Foden, J. A.; Khayter, C.; Maeder, M. L.; Reyon, D.; Joung, J. K.; Sander, J. D., High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **2013,** *31* (9), 822-6; Lin, Y.; Cradick, T. J.; Brown, M. T.; Deshmukh, H.; Ranjan, P.; Sarode, N.; Wile, B. M.; Vertino, P. M.; Stewart, F. J.; Bao, G., CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res* **2014,** *42* (11), 7473-85.

293. Parrish, A. R.; She, X.; Xiang, Z.; Coin, I.; Shen, Z.; Briggs, S. P.; Dillin, A.; Wang, L., Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs. *ACS Chem Biol* **2012,** *7* (7), 1292-302; Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W., Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotechnol* **2014,** *32* (5), 465-72.

294. Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W., Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *J Am Chem Soc* **2009,** *131* (25), 8720-1.

295. Hou, Y. M., CCA addition to tRNA: implications for tRNA quality control. *IUBMB Life* **2010,** *62* (4), 251-60.

296. Wang, W.; Takimoto, J. K.; Louie, G. V.; Baiga, T. J.; Noel, J. P.; Lee, K. F.; Slesinger, P. A.; Wang, L., Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nat Neurosci* **2007,** *10* (8), 1063-72.

297. Wang, Y. S.; Russell, W. K.; Wang, Z.; Wan, W.; Dodd, L. E.; Pai, P. J.; Russell, D. H.; Liu, W. R., The de novo engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives. *Mol Biosyst* **2011,** *7* (3), 714-7.

298. Jardine, D.; Litvak, M. K., Direct yolk sac volume manipulation of zebrafish embryos and the relationship between offspring size and yolk sac volume. Blackwell Science Ltd: Journal of Fish Biology, 2003; Vol. 63, pp 388-397.

299. Ling, M. M.; Robinson, B. H., Approaches to DNA mutagenesis: an overview. *Anal Biochem* **1997,** *254* (2), 157-78; Zheng, L.; Baumann, U.; Reymond, J.-L., An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic acids research* **2004,** *32* (14), e115-e115.

300. Schmittgen, T. D.; Livak, K. J., Analyzing real-time PCR data by the comparative CT method. *Nature protocols* **2008,** *3* (6), 1101-1108.