Developing light-activated nucleic acids for gene knockdown in cell-free and living systems



Giacomo Mazzotti

Oriel College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2023

Abstract

Nucleic acids encode the information of all biological functions, which are regulated through the expression of genes. Today, various deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) structures can be reliably synthesized and manipulated to recreate naturally-occurring phenomena or introduce new functions in cell-free and living systems. Controlling the expression of a gene to protein has become a fundamental tool for therapeutic and synthetic biology research. One way of controlling gene expression is by using specific wavelengths of light. Light is an ideal stimulus as it can be externally applied to a system with high spatiotemporal precision, and it is bioorthogonal to most biological processes.

The synthesis of a protein can be inhibited by degrading or deactivating the corresponding gene, a process called gene knockdown. The two nucleic acids mainly used for gene knockdown are antisense oligonucleotides (ASOs), short single stranded DNA sequences, and small interfering RNAs (siRNAs), short double stranded RNA sequences. ASOs and siRNAs can be synthesized and modified to be activated in response to light irradiation.

This work presents the development of an orthogonal two-wavelength control of cell-free gene expression. Light-activated (LA)-ASOs and LA-siRNAs were generated from commercially available amino-modified sequences. The amine groups were functionalized with UV or blue light-activatable photocages, which inactivated the nucleic acids by steric

hinderance. Upon irradiation with a specific wavelength of light, the knockdown activity was restored, as the photocages were cleaved off, uncaging to the initial amino-modified nucleic acid structure. Cell-free protein synthesis (CFPS) was controlled with UV light-activated ASOs, by triggering RNase H-mediated mRNA degradation on demand. Additionally, by combining UV LA-ASOs to a previously developed blue LA-DNA template, transcription was activated with UV and translation halted orthogonally with blue light. Finally, ASOs targeting two different genes were modified with orthogonal photocages for precise and selective two-wavelength gene knockdown in CFPS.

Progress towards light-activated gene knockdown in mammalian cells using siRNA will also be presented. Two different strategies for generating LA-siRNA will be discussed. The siRNA photocages were decorated with a cell-penetrating peptide and a fluorophore, for self-delivery to cells and fluorescent tracing of the nucleic acid. Images demonstrating cellular uptake without the use of transfection reagents will be finally shown.

Acknowledgements

I wish to express my profound gratitude to everyone that has supported me over the past four and a half years. First and foremost, I would like to extend my appreciation to Dr. Michael Booth for believing in my abilities and giving me the opportunity to be one of the first few people to establish an entirely new research group. You have been an exceptional supervisor, and I am looking forward to working on new, interesting projects with you at UCL. Secondly, I want to express my gratitude to Prof. Hagan Bayley for welcoming us into his "extended" research group. In the lab and outside the lab, you truly made us feel like a big family.

To the Booth group, the best team one could ask for. I will miss working with you. Special thanks go to Dr. Jefferson Smith for being an absolute superstar, an inspiration for us all, and, more importantly, a great friend. To Razia Chowdhry, for our late nights in the lab talking absolute non-sense for hours while performing questionable experiments. We made it to the end, and I am proud of you. To Dr. Assala Al Samad, a true point of reference for everyone. I have learned a great deal from you, as we all did. Our coffee breaks, lunch, and conversations have been invaluable, and I look forward to continuing them in London. To Khoa Chung, for all the pleasant pints, long conversations, and great fights over any board game. To Denis Hartmann for your help with the second half of the ASO project. I

would also like to extend my gratitude to the Booth group at UCL, which supported me during the final months of writing this very text.

To the Bayley group, you have made the lab a great workplace. A special mention goes to Pablo Martín Baniandrés for being my hermano at 360°, no matter what. To Dr. Matthew Cornall, always lifting everyone's spirit with jokes and banter, even if your choice of lab tunes was often questionable. Jorin Riexinger for being a great scientist and constant source of positivity. The pub trips and the adventures in Crete have been memorable.

To the Oriel College crew, Lauren, Mike, Josh, Darius, and everyone in the MCR. You have been my family away from home. Special mentions go to the lejendary Eoin O'Sullivan and Lukas Boehnert. The stories, the memories, the trips together. Eoin, sheer luck led me to you on my first night in Oxford, and I have followed your Irish buttocks ever since. I am immensely proud of your rugby ascend to the top, and I look forward to seeing you and your dad at my wedding. Lukas, you have not been the first person I met in Oxford, and that's ok. Your pasta still gives me nightmares, but I will miss your dad jokes. Potato potato. One day, I'll teach you how to make chickpeas and balsamic.

To my family, for their patience and unwavering support. To Federica Fabbri, for standing by me throughout the years. I know it has not been the easiest journey. Thank you for enduring all the late flights and public transports, all the "hikes" on flat grounds, the rainy days, covid at my mom's, and *that* day with the beagles. Plus, everything else.

I also want to express my gratitude to the Engineering and Physical Sciences Research Council and UK Research and Innovation for giving me the opportunity to undertake this research.

Published work

The following paper has been published based on work outlined in Chapter 2 of this thesis:

- G. Mazzotti, D. Hartmann, M. J. Booth. Precise, Orthogonal Remote-Control of Cell-Free Systems Using Photocaged Nucleic Acids. *J. Am. Chem. Soc.* **2023**, 145, 17, 9481-9487. doi.org/10.1021/jacs.3c01238.

The following review paper has been referenced throughout this thesis:

- D. Hartmann, J. M. Smith, G. Mazzotti, R. Chowdhry, M. J. Booth. Controlling gene expression with light: a multidisciplinary endeavour. *Biochem. Soc. Trans.* **2020**, 48 (4), 1645-1659. doi.org/10.1042/BST20200014.

Abbreviations

2’-F = 2’-fluoro

2’-MOE = 2’-methoxyethyl

2’-OMe = 2’-O-methyl

AA = Aminoallyl

AAV = Adeno-associated vectors

Ad = Adenoviral

ASCVD = Atherosclerotic cardiovascular disease

ASGPR = Asialoglycoprotein receptor

ASO = Antisense oligonucleotide

AuNPs = Gold nanoparticles

bLA = blue light-activated

Cas 9 = CRISPR-associated protein 9

CFE = Cell-free expression

CFPS = Cell-free protein synthesis

CPP = Cell-penetrating peptide

CRISPR = Clustered regularly interspaced palindromic repeats

Cy5 = Cyanine-5

ds = double stranded

DSPC = 1,2-diastearoyl-*sn*-glycero-3-phosphocholine

E. coli = Escherichia coli

GalNAc = N-acetyl galactosamine

GFP = Green fluorescent protein

IPTG = Isopropyl ß-*D*-1-thiogalactopyranoside

ITR = Inverted terminal repeats

IVT = in vitro transcription

IVTT = in vitro transcription and translation

LA = Light-activated

LCA = Leber congenital amaurosis

LDL = Low density lipoprotein

LNA = Locked nucleic acid

LNP = Lipid nanoparticle

mAb = Monoclonal antibody

miRNA = micro RNA

mRNA = Messenger RNA

mC = mCherry

mV = mVenus

ncRNA = Non-coding RNA

mSA = Monovalent streptavidin

NHS = N-hydroxysuccinimide

NIR = Near-infrared

nt = nucleotide

PAGE = Polyacrylamide gel electrophoresis

PC = Photocleavable

PCB = Photocleavable biotin

PCR = Polymerase chain rection

PEG = Polyethylene glycol

PMO = Phosphorodiamidate morpholino oligonucleotide

PNA = Peptide nucleic acid

PS = Phosphorothioate

PURE = Protein synthesis Using Recombinant Elements

RBS = Ribosome binding site

RNAi = RNA interference

RNAP = RNA polymerase

RuBEP = $Ru(bipyridine)_2(3\text{-}ethynylpyridine)_2$

siRNA = Small interfering RNA

SA = Streptavidin

shRNA = Small hairpin RNA

SNA = Spherical nucleic acid

SNP = Single-nucleotide polymorphism

ss = single stranded

tetSA = Tetravalent streptavidin

tRNA = Transfer RNA

TTR = Transthyretin-mediated amyloidosis

UCNP = Upconverting nanoparticle

UTP = uridine triphosphate

UTR = Untraslated region

uvLA = UV light-activated

WCF = Watson-Crick-Franklin

Contents

Introduction 15

1.1 Early years of nucleic acid research 15

1.1.1 Nucleic acids structure and central dogma 16

1.2 Nucleic acids as therapeutics 18

1.2.1 Current medicine vs. future medicine 18

1.2.2 Small molecule therapeutics 20

1.2.3 Protein therapeutics 21

1.2.4 Nucleic acid therapeutics 22

1.3 Nucleic acids in synthetic biology 35

1.3.1 Cell-free expression systems 37

1.4 Remote control of gene expression 38

1.4.1 Control of gene expression using small molecules 39

1.4.2 Control of gene expression with light 40

1.3 Summary of thesis work 48

Orthogonal on/off control of cell-free expression using photocaged nucleic acids 50

2.1 Historical remarks 50

2.2 RNase H-mediated mRNA degradation 52

2.3 Other ASO mechanisms 53

2.4 Antisense oligonucleotides in synthetic biology 54

2.5 Light-activated antisense oligonucleotides 55

2.6 Orthogonal systems in synthetic biology 57

2.7 Chapter overview 59

2.8 Preparation of modified ASOs 61

2.9 RNase H-mediated mRNA degradation tests 65

2.10 Optimization of modified-ASO sequence 70

2.11 Control of in vitro transcription and translation with UV 74

2.12 Two-wavelength control of cell-free protein synthesis 78

2.13 Two-wavelength control of gene knockdown 82

2.14 Conclusions 88

Development of light-activated and self-delivered siRNAs for the control of gene knockdown in mammalian cells 90

3.1 Historical remarks 90

3.2 RNAi mechanism 92

3.3 siRNA therapeutics 94

3.4 Controlling siRNA with light 96

3.5 Cell-penetrating peptides 97

3.6 Chapter overview 98

3.7 Preparation of modified siRNAs 100

3.8 Aminoallyl siRNAs 106

3.9 C6-amino linker 112

3.10 Self-delivery to mammalian cells 122

3.3 Conclusions 127

Conclusion and future directions 129

Methods 134

5.1 Materials 134

5.2 General methods 135

5.2.1 Agarose gel electrophoresis 135

5.2.2 Polyacrylamide gel electrophoresis 135

5.2.3 Oligonucleotide mass spectrometry 136

5.3 Methods Chapter 2 136

5.3.1 Formation of biotinylated UV-photocleavable antisense oligonucleotides 137

5.3.2 First HPLC purification method of modified antisense oligonucleotides 137

5.3.3 Transcription of mVenus and mCherry mRNA 137
5.3.4 RNase H-mediated mRNA degradation 138
5.3.5 Binding of streptavidin 138
5.3.6 UV irradiation conditions 138
5.3.7 Second HPLC purification of modified antisense oligonucleotides 139
5.3.8 Preparation of blue light-activatable antisense oligonucleotides 139
5.3.9 Preparation of linear template DNA 139
5.3.10 Blue light irradiation conditions 140
5.3.11 In vitro transcription control using uvLA-V3 140
5.3.12 Two-wavelength control of in vitro transcription 141
5.3.13 Light-controlled gene knockdown in cell-free protein synthesis with uvLA-V3 142
5.3.14 Control of transcription and gene knockdown with two wavelengths in cell-free protein synthesis 142
5.3.15 Two-wavelength control of RNase H-mediated mRNA degradation 143
5.3.16 Two wavelength-controlled gene knockdown in cell-free protein synthesis 143
5.4 Methods Chapter 3 144
5.4.1 Preparation of modified siRNA 144
5.4.2 Binding of streptavidin 145
5.4.3 HEK 293 T transfection 145
5.4.4 3T3-GFP transfection 146

5.4.5 Epifluorescence microscopy 146

5.4.6 UV irradiation conditions 146

Appendix A 147

A.1 DNA sequences 147

A.2 mRNA sequences 148

A.3 Sequences of oligonucleotides used in PCR 149

A.4 Antisense oligonucleotide sequences 149

Appendix B 150

B.1 Modified antisense oligonucleotides LC-MS spectra 150

B.2 LC-MS spectra uvLA-V1 biotin 150

B.3 LC-MS spectra uvLA-V2 biotin 151

B4. LC-MS spectra uvLA-V3 biotin 152

B5. LC-MS spectra bLA-C1 biotin 154

Bibliography 155

1

Introduction

1.1 Early years of nucleic acid research

The term nucleic acid was first coined by Richard Altmann in 1889 [1]. The name derived from “nuclein”, a phosphorous-rich substance isolated from cell nuclei by his mentor Friedrich Miescher, almost two decades earlier [2]. Over the following years, important structural information on nucleic acids was elucidated by several scientists – most notoriously Albrecht Kossel and Phoebus Levene – and the term evolved into the more detailed designations deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). In 1928, Frederick Griffith demonstrated how injecting mice with either living virulent

pneumococcus or a mixture of avirulent pneumococcus and heat-killed virulent pneumococcus led to the death of the mice [3]. This was the first showcase of bacterial transformation [4]. Years later, in 1944, after understanding the potential of radioactive isotopes, Oswald Avery, Colin MacLeod, and Maclyn McCarty determined that nucleic acids were responsible for bacterial transformation and, therefore, for transferring genetic information. In their experiment, *Escherichia coli* was infected with a virus containing ^{32}P-labelled DNA and ^{35}S-labelled protein coating. As the virus replicated within the cells, only the ^{32}P-labelled material was recovered [5]. Finally in 1953, the DNA double helix structure was elucidated by Watson, Crick, Franklin, and Wilkins, providing a clear understanding of how genetic information is copied. This revelation marked the beginning of a new era for medicine, biology, and chemistry.

1.1.1 Nucleic acids structure and central dogma

Nucleic acids are composed of nucleotides, each containing a purine or pyrimidine nitrogenous base, 5-membered ring sugar, and a phosphate group. The sugar is deoxyribose for DNA and ribose for RNA, while the bases are adenine (A), guanine (G), cytosine (C), and thymidine (T) in DNA or uracil (U) in RNA.

After the DNA structure, Crick envisioned the central dogma of molecular biology, which states how information is transferred sequentially residue-by-residue from DNA to RNA to protein and can only be transferred from nucleic acid to nucleic acid or nucleic acid to protein, but not from protein to either protein or nucleic acid (Figure 1.1) [6, 7]. These findings opened new fields of research involving nucleic acids, giving a direct link between

their chemical composition and biological function. In the following decades, great effort was put into finding ways to synthesize nucleic acids and sequencing them.

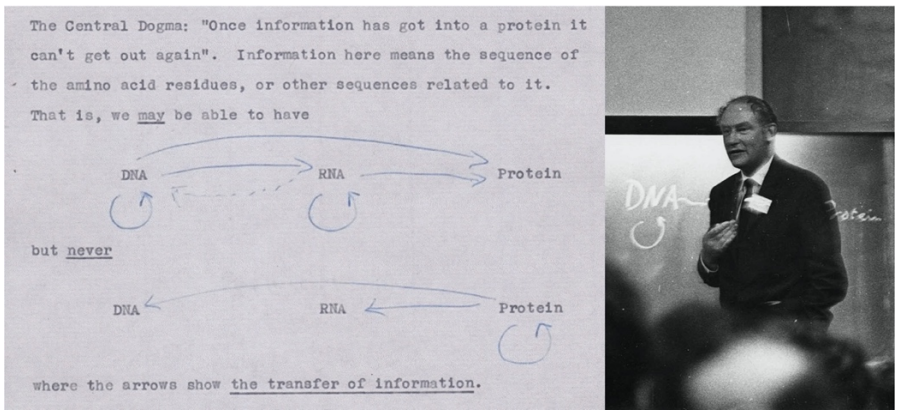


Figure 1.1: Francis Crick and the central dogma

Francis Crick in 1957 during a renowned lecture "On protein synthesis" at University College London said to have "permanently altered the logic of biology". At the time, experimental data was not conclusive in excluding the direct production of proteins from DNA.

During the 1980s, the phosphoramidite approach for oligonucleotide synthesis was introduced and established as a robust and relatively simple method to build nucleic acids one nucleotide at the time [8]. Subsequently, the method was adapted for solid phase synthesis and automated synthesizers, with its core chemistry remaining largely unchanged after forty years. Another groundbreaking scientific achievement of the same period is the polymerase chain reaction (PCR) method [9], which allows for rapid DNA amplification and has become an indispensable tool in laboratories across a broad range of disciplines.

1.2 Nucleic acids as therapeutics

The first time that the therapeutic potential of nucleic acids was demonstrated was in 1972, when Theodore Friedmann and Richard Roblin introduced the concept of gene therapy, explaining how a dysfunctional gene could potentially be replaced with a functional gene copy [10]. At the time, the role of RNA in the central dogma was that of an "unstable intermediate" carrier of information, later called mRNA. It was Paul Zamecnik who discovered that RNA could base-pair with more RNA or DNA forming a double helix, and he had the intuition of using this finding for therapeutic purposes. Zamecnik designed the first synthetic antisense oligonucleotide, a short DNA sequence that could bind and degrade Rous Sarcoma Virus mRNA, inhibiting the virus replication and cell transformation [4, 5]. This laid the foundations for the development of antisense therapeutics, which represent almost the entirety of nucleic acid drugs available today. With the development of new delivery technologies and large scale production and testing, nucleic acid therapeutics are now at the forefront of medical research globally, with outstanding results such as the COVID-19 mRNA vaccine [13].

1.2.1 Current medicine vs. future medicine

With the recent advances in next-generation sequencing technologies and their reduction in cost, it is now possible to sequence an entire genome in less than a minute and for less than one hundred pounds [14]. This increased accessibility to genetic information is driving substantial and rapid progress in precision medicine and personalized medicine. Precision

medicine is a term used to describe the precise intervention on a patient at a molecular level, whereas personalized medicine refers to the tailoring of a treatment for each individual case [15]. The two terms are often used interchangeably in life sciences, and they refer to the ability to prevent/treat/monitor a disease with methods tailored to an individual's unique genetic, biochemical, and physiological profile (Figure 1.2). This would be impossible using only conventional small molecule therapeutics, as small molecules almost uniquely aim to alter protein functions in the cell, but only less than 1.5% of the human genome is currently recognized to encode functional proteins [16], and only the 10-14% of proteins possess a binding site to make them a druggable target [17].

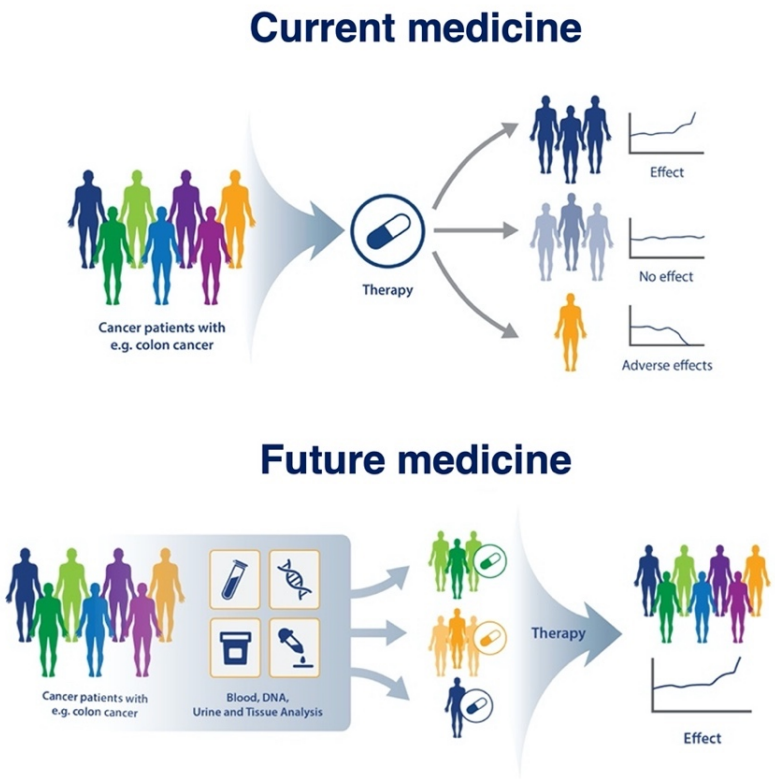


Figure 1.2: Current medicine vs. future medicine
In current medicine one treatment is designed for a multitude of patients that might react differently. With future medicine, personalized therapies are tailored to each individual.

1.2.2 Small molecule therapeutics

Small molecule drugs constitute over 90% of the currently marketed therapeutics, due to their low cost of production in industrial scale synthesis, their excellent stability *in vivo*, and the possibility to easily design and screen different analogs to find what can reach the highest possible number of pharmacologic targets. Despite their efficacy, small molecule therapeutics exhibit several limitations that preclude their use for certain diseases.

To be clinically relevant, small molecules must establish a strong interaction with the target protein, binding deep grooves or cavities with high affinity [18]. This constraint restricts their applicability to specific protein categories, primarily receptors, ion channels, or transporters [19]. Small molecules also lack selectivity, affecting the targeted protein in every tissue, which can lead to side effects. Moreover, small molecules retain the same function independently of feedback from the organism. This can pose a limitation in certain instances. One example is with beta-blocker ligands for heart conditions, which act by blocking beta-adrenergic receptors, irrespective of slow heart rate at rest or faster heart rate during physical activity. During exercise, when there is a higher demand for oxygen and energy levels, the heart of an individual assuming beta-blockers will work more than normal to increase the required blood flow.

These limitations must be complemented by protein and nucleic acid therapeutic approaches to yield a broader and more comprehensive spectrum of treatment options.

1.2.3 Protein therapeutics

Proteins carry out highly specific functions in the cell. Recombinant production systems have facilitated their large-scale production across several organisms, with the flexibility to easily introduce sequence mutations or modifications, to improve the protein activity or facilitate its purification. The most approved protein therapeutics are monoclonal antibodies (~50%), generally simple antibodies able to bind a target antigen and guide the innate immune system to it. Other types of engineered protein therapeutics are used to substitute missing or mutated proteins, such as coagulation factors, replacement enzymes, or growth factors [20]. Protein therapeutics are also being researched as regulators for signaling protein-protein interactions. These interactions do not usually have a site easily bound by small molecules and the best solution would be to deliver a protein with modified signaling properties that can affect the cellular behavior [19].

However, protein therapeutics present some limitations, such as their size that can be over 100 kD, complicating their delivery. Some proteins can be difficult to produce recombinantly as they can be toxic for the host organism or require post-translational modifications to function, which can be challenging to introduce. Furthermore, proteins must be folded with defined secondary and tertiary structures to function efficiently, which can be difficult to maintain if the protein contains sensitive functional groups (e.g., cysteine residues). Finally, during their production and purification, pH, detergents, or reactions within the host organism can alter important protein properties [16]. As the field of protein engineering is making great advances, proteins are the best therapeutic option for several applications. However, nucleic acids present unique characteristics needed to expand on proteins and small molecule therapeutic applications.

1.2.4 Nucleic acid therapeutics

Conventional small molecules and protein treatments target protein-based processes. In contrast, nucleic acid therapeutics act at the genetic level, bypassing a layer of biological complexity. Targeting genes not only provides an alternative, but often represents the sole option to treating diseases caused by "undruggable" proteins. The mechanism of action of nucleic acid therapeutics can proceed in two ways: by delivering an exogenous gene (e.g., by implementing a missing gene or replacing a faulty one), or by silencing the expression of an existing gene [13]. Once nucleic acids reach the target cell, the endogenous cellular machinery performs the task encrypted in the nucleic acid sequence.

Based on their mechanism and delivery platform, nucleic acid therapeutics can be classified as following (Figure 1.3):

Antisense therapeutics include a variety of relatively short nucleic acids, amongst which the most effective are antisense oligonucleotides (ASOs) and small interfering RNAs (siRNAs). ASOs are single stranded DNA sequences, typically 16-24 nucleotide long, whereas siRNAs are 21-23-nucleotide double stranded RNAs. The name antisense derives from their ability to form Watson-Crick-Franklin (WCF) base pairs with RNA in the cell, causing the knockdown of a target gene. ASOs act by RNase H-mediated RNA degradation or by steric block, whereas siRNAs activate the RNA interference (RNAi) mechanism. A more detailed description of ASOs and siRNAs will be discussed in Chapter 2 and Chapter 3 of this work, respectively. This group includes other types of nucleic acids acting similarly, such as micro RNAs (miRNAs).

mRNA vaccines encode the gene of a desired antigen in the form of messanger (m)RNA, so to be recognized and translated by the host cell. Some of the mRNA features typically needed for its expression in mammalian cells are untranslated regions (UTRs) flanking the coding

sequence, a 5'-end cap and a poly(A) tail. Once the mRNA is delivered to the cell the antigen is expressed and an immune response triggered. The mRNA is usually encapsulated in lipid nanoparticles (LNPs) for biodistribution and to prevent mRNA degradation while in circulation.

Adeno-associated vectors (AAV) are virus-derived capsids containing short and modified single stranded DNA viral genomes [21]. The DNA sequence contains viral genes needed for encapsulation into new viral vectors and integration in the human genome, such as T-shaped inverted terminal repeats (ITRs). Most of the other viral genes are substituted by the desired therapeutic genes.

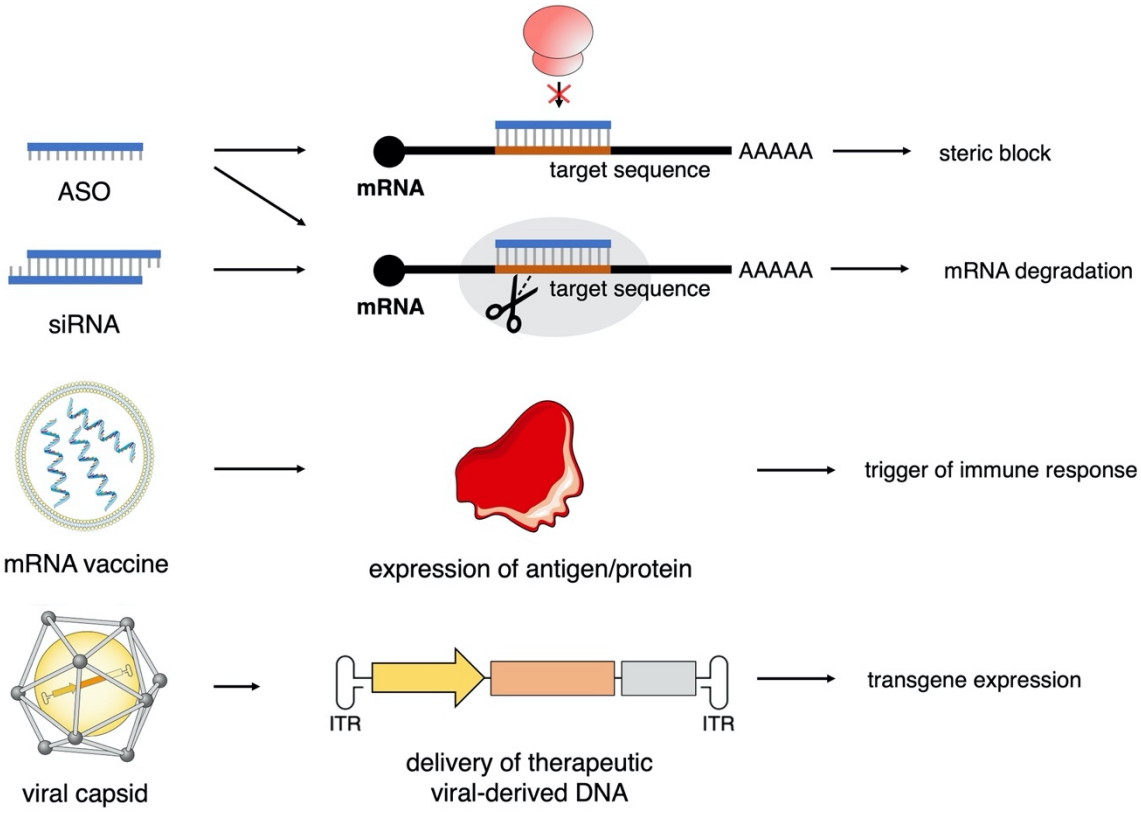


Figure 1.3: Types of nucleic acid therapeutics approved
Antisense therapeutics (ASOs and siRNAs) act by inhibiting mRNA translation. mRNA vaccines are delivered via lipid nanoparticles to express an antigen and trigger an immune response. Therapeutic genes are delivered via viral-derived DNA.

Adenoviral (Ad) vector vaccines are also encapsulated into viral capsids, but they contain a much larger double stranded DNA viral genome. Viral replication genes are replaced with the target antigen gene, which triggers the desired immune response once expressed. Contrary to AAV, Ad-dsDNA has a low capacity of integrating in the host genome [22].

1.2.4.1 Therapeutic advantages of nucleic acids

Given the specificity of nucleic acids and the tools available for their production and modification, a large number of sequences can be quickly designed and screened. Companies are developing platforms to produce and deliver nucleic acid therapeutics for different applications. Ideally, once a production workflow is established, different diseases can be targeted simply by changing the nucleic acid sequence. A remarkable example is Alnylam Pharmaceuticals, a biopharmaceutical company that produces four clinically-approved antisense therapies. These therapies share a similar structure and delivery system, while targeting different genes. Following the approval of the first therapy, the subsequent three quickly gained approval in a four-year span, demonstrating how a reliable platform can lead to rapid drug development. This standardized production approach, could also contribute in lowering the overall cost of the process – currently very high – similarly to what occurred with monoclonal antibodies throughout the years [23].

Antisense therapeutics – which are the most approved category of nucleic acid-based drugs – inhibit the synthesis of a target protein. This can be exploited for diseases caused by malfunctioning proteins that cannot be strongly bound by a small molecule drug, or that cannot be replaced by an engineered protein. Additionally, preventing the synthesis of a target protein can offer advantages over targeting the function of an altered protein already formed and circulating in the organism.

Nucleic acid therapeutics often represent the only possible treatment for rare or orphan genetic disease [13]. This includes diseases that require targeting of a single-nucleotide polymorphism (SNP) to selectively knockdown a specific mutant allele [24]. Moreover, the larger part of the genome does not encode proteins, but instead produces noncoding RNA (ncRNA) transcripts, that help regulating gene expression and protein functions in the cell [25]. The down regulation of ncRNAs has been linked to most cancer hallmarks and several biological processes, including immune disorders, and neurological diseases. A direct way to target ncRNAs is by using antisense nucleic acids, whereas using small molecule drugs or protein therapeutics would be overly complicated.

Finally, gene editing tools such as clustered regularly interspaced palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas 9), that would allow the permanent and precise modification of the genome are also in development for therapeutic purposes [26]. These technologies allow the insertion, deletion, or change of DNA bases to correct base mutations or mutate disease-related genes for therapeutic effects.

Despite the number of advantages offered by nucleic acids for therapeutic purposes, of all approved drugs between 2008 and 2022, only ~3% correspond to nucleic acid therapeutics: fifteen antisense therapeutics, two mRNA vaccines for COVID-19, two AAV, and two Ad vector COVID-19 vaccines for emergency use only. This corresponds to only 1-2 nucleic acid approved every year.

1.2.4.2 Nucleic acid limitations

Nucleic acids are relatively large and negatively charged due to the phosphate groups on the backbone. As an innate defense mechanism, the organism can recognize foreign nucleic acids

based on localization, local concentration, and structure [27]. Depending on the nucleic acid function, the therapeutic effect occurs inside the cytosol or nucleus of a cell [28].

The main extracellular obstacles encountered are the following: *circulation* through circulatory and lymphatic systems can result in a quick excretion, as most nucleic acids are small enough to be cleared from the organism by dedicated organs such as spleen, kidney, or liver. The method of *administration* plays a key role in designing the nucleic acid. All the approved therapeutics to date are administered via injection, often *in situ* to facilitate the uptake (Figure 1.4). While circulating, nucleic acids are subjected to *nuclease degradation* and triggering of the *immune response*. For this reason, nucleic acid treatments are often coupled to immunosuppressive injections. Finally, if targeting the nervous system, the blood-brain-barrier represents an additional obstacle [29].

Once the target tissue is reached, more barriers are encountered for the cellular uptake: first, the right cell type has to be targeted for *specificity*. Most methods of delivery involve *endocytosis*. Through this pathway, nucleic acids are encapsulated in the endosome – an internal vesicle formed by the cell – from which they have to escape to reach the cytosol, in a process called *endosomal release*. At this time, the nucleic acid could be degraded by several enzymes activated upon acidification or expelled back outside the cell via exocytosis. Physical methods such as electroporation can be used for cell delivery, but they are usually not feasible for most organs *in vivo*. Finally, an extra barrier is present for delivery to the nucleus. Small, modified nucleic acids can often enter the nucleus by signal-independent diffusion, whereas longer sequences have to contain a nuclear localization signal and transportation by dedicated proteins [22, 23].

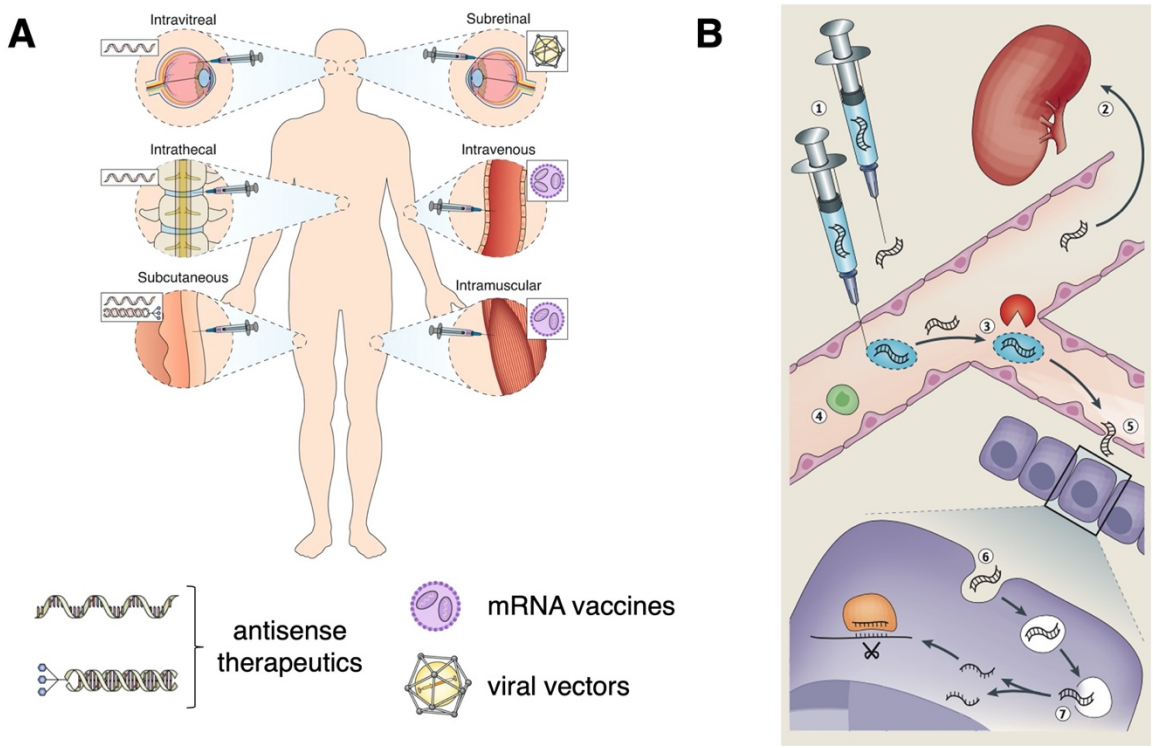


Figure 1.4: Barriers encountered by nucleic acid therapeutics

A) Depicts different ways to inject nucleic acid therapeutic formulations to different parts of the human body. B) Several barriers encountered by nucleic acid therapeutics, before reaching the cellular target. 1 – injection method, 2 – excretion from the organism, 3 – nuclease degradation, 4 – triggering of immune response, 5 – specific cell-targeting, 6 – cellular uptake, 7 – endosomal escape.

Even considering the tremendous potential of nucleic acid therapeutics, these limitations are reflected in the resources and time needed for manufacturing and clinical testing. The resulting cost is very high, especially for large scale manufacturing, with consequent high cost of treatment. This has led to the discontinuation of some approved nucleic acid therapeutics, because of other cheaper ways of treatment. By developing new technologies and optimizing the present ones, the costs will decrease in the future and more drugs will be commercialized.

1.2.4.3 Nucleic acid modification methods

The limitations described are mainly overcome by chemical modification of the nucleic acid. A nucleic acid can be modified at different positions and each modification offers a set of advantages.

The non-bridging oxygen of the *phosphate backbone* can be substituted with a sulfur atom to substantially decrease nuclease degradation, a modification called *phosphorothioate* (PS). The work on PS is originally from 1966, when Frick Eckstein reported the synthesis of nucleosides containing a PS group and the resistance to phosphodiesterases of a PS-dinucleotide [24, 25]. Phosphorothioates are now a fundamental component of several drug development strategies involving antisense gene silencing, with a large portion or entire sequences completely modified with PS. However, each PS substitution adds a chiral center to the nucleic acid sequence. Therefore, oligonucleotides of standard length for therapeutic purpose can reach high degrees of chirality. For example, a fully PS-substituted 19-mer results in more than half a million stereoisomers [34]. Even if isomer mixtures result in an alteration of the physical and biological property of the nucleic acid, such as the potency, this is a well-tolerated occurrence in the field, as it usually does not result in toxic effects, although it can increase the immune response [35]. The sulfur in phosphorothioates have the advantage of allowing selective alkylation to introduce further modifications at specific locations [36]. This can be exploited for different purposes, such as neutralizing the overall charge of the nucleic acid groups to improve its delivery. Some of the properties affected by the introduction of PS are the reduced affinity to complementary sequences (T_m is decreased by ~0.8 °C for each nucleotide linkage) [37], and the non-specific binding to certain proteins [30, 31]. Moreover, substitution of an oxygen with a sulfur maintains the overall negative charge, which can be good for the solubility of the nucleic acid, but still requires the same delivery methods as unmodified nucleic acids.

The *carbohydrate* portion of each nucleotide is also commonly modified, in particular by alkylating the 2'-OH group of the ribose. This is typically either simply methylated (2'-OMe) or methoxyethylated (2'-MOE) to increase resistance to nuclease degradation and improve the stability in vivo. Moreover, for DNA bases, the 2'-OH can be substituted with an electronegative substituent, typically 2'-F. This helps the sugar adopting a C3'-*endo* conformation, which is the conformation favored by RNA bases. For DNA antisense oligonucleotides, this results in a more stable duplex formation with the target mRNA [40]. This kind of modification offers a valid alternative to partially substitute the DNA sequence with RNA bases to mimic the sugar conformation, which would instead result in degradation. Other 2'-modifications, such as 2'-MOE adopt a similar C3'-*endo* conformation, with the MOE group projecting into the minor groove after the duplex is formed [41]. Phosphorothioates and 2'-modifications have been combined successfully in several FDA-approved drugs available today.

The sugar can also be modified to constrain the base configuration. A methylene bridge is installed between the 2'-oxygen and the 4'-carbon, locking the sugar into a N-type pucker to again improve the duplex stability with the target RNA and resistance to nucleases [42]. These are called locked nucleic acids (LNAs).

The five *nucleobases* can also be modified, however very few modifications have found a successful application. One of the concerns regarding modified nucleobases is that, once metabolized, they could compete with other natural nucleotides for incorporation into the genome [43]. Nonetheless, therapeutic nucleic acids containing C5-methyl cytidine and C5-methyl uridine are present in the literature. The C5-methyl groups can stack between the planar nucleobases, contributing in excluding water molecules from the duplex formed, therefore increasing its thermal stability [44]. Modifications with groups larger than a methyl have been shown to interfere with base pairing [45].

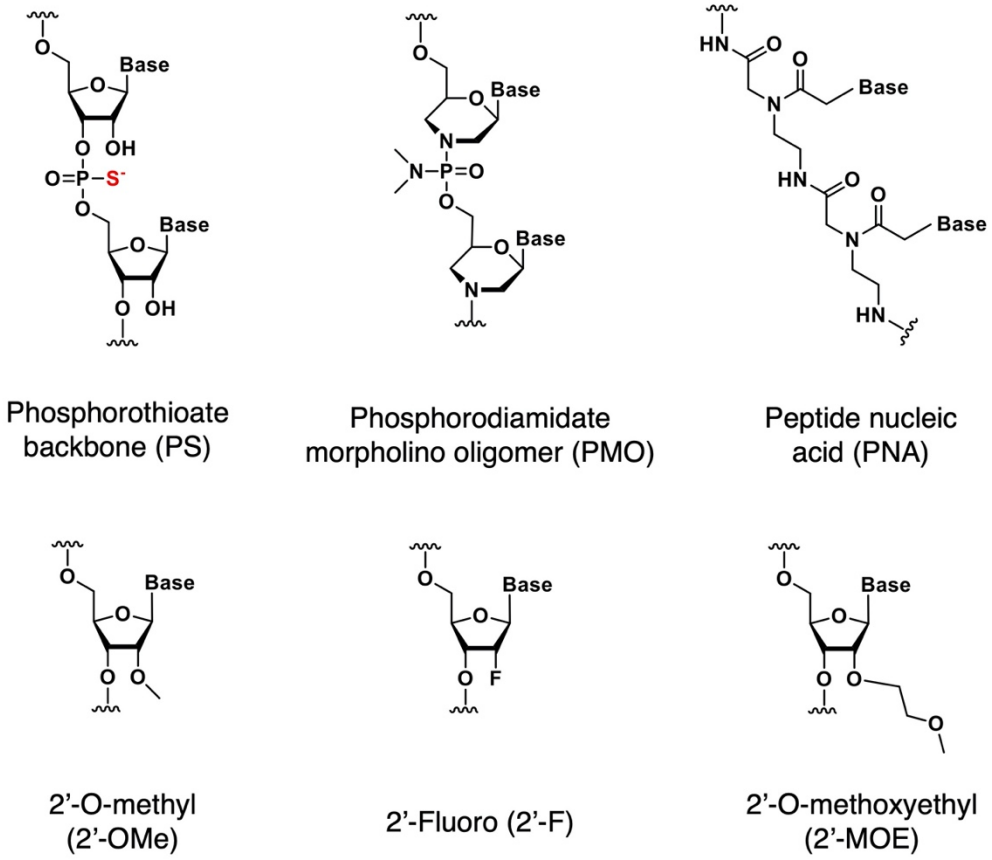


Figure 1.5: Most common nucleic acid modifications

Most approved nucleic acid therapeutics are composed of a combination of these nucleotide modifications.

Finally, phosphorodiamidate morpholino oligomers (PMOs) and peptide nucleic acids (PNA) are very effective DNA/RNA analogues. Their polymer chain is not charged, and this results in the formation of stronger PMO:DNA/RNA or PNA:DNA/RNA duplexes, compared to DNA:DNA/RNA because of the absence of electrostatic repulsions. This feature has been exploited especially in antisense therapeutics that do not require RNA degradation and it will be discussed further in Chapter 2. PMOs and PNAs are composed of natural DNA nucleobases attached to a synthetic backbone. In PMOs, the sugar/phosphate backbone is replaced by

methylenemorpholine rings linked by phosphorodiamidate groups, while in PNAs, the backbone consists of N-2-aminoethylglycine repeating units [46].

1.2.4.4 Nucleic acid delivery methods

While nucleic acid modifications help solving the stability issues and can improve the drug potency, another major barrier to overcome is cellular delivery. In the recent years, a few delivery technologies have been successful and have truly pushed the field of nucleic acid therapeutics forward. To date, most nucleic acid therapeutics have been focused on the delivery to the liver or to favorable areas, such as eyes and spinal cord [47]. The liver is a good target, as it possesses a discontinuous sinusoidal endothelium, meaning that it favors the uptake of larger nanoparticles before renal clearance, compared to other organs. Moreover, the liver is a highly vascularized tissue as the kidneys and spleen are, again favoring the accumulation of larger particles. Diversely, eyes possess the advantage of immune privilege, exhibiting a limited inflammatory response to safeguard vision from potential damage. Moreover, even though the nervous system presents the additional blood-brain barrier, nucleic acids injected in the spinal cord are taken up by neurons and glial cells and quickly distribute throughout the central nervous system [29].

The main technologies developed for nucleic acid cell delivery are the following:

Lipid nanoparticles (LNPs), spherical vesicles made of lipids that can encapsulate nucleic acids, and have been one of the most successful means of delivery . Originally, they were only formed by neutral lipids, but that resulted in poor nucleic acid encapsulation [48]. To improve on that, cationic lipids were used to form stable complexes with the nucleic acid negative charge called lipoplexes. However, lipoplexes tend to trigger a strong immune response and get rapidly eliminated by organ clearance. The third version of LNPs and most widely used

today is the pH titratable cationic ionizable lipids. These are composed of aminolipids, which are positively charged at acidic pH, and almost completely uncharged at physiological pH, for improved pharmacokinetics. Usually, the lipid composition also includes other PEGylated lipids to help vial stability before administration, then pharmacokinetics and biodistribution once in circulation. Phospholipids are also added, as they are well-known to increase overall LNP stability, together with other important helper lipids, such as cholesterol, which increases the lipid membrane rigidity, reducing drug leakage and non-specific protein binding. As stated above, due to the fenestrated capillaries that favor the uptake of larger objects, LNPs naturally accumulate in the liver. Here, they interact with apolipoprotein E, which facilitates their hepatocyte uptake via the low-density lipoprotein (LDL) receptors. This is a robust and efficient mechanism for nucleic acid delivery, and one of the reasons why delivery to the liver has been the main focus for nucleic acid therapeutics [48].

Bioconjugation to different chemical moieties can promote the targeting of specific cells and cellular uptake through cellular receptors [47]. These ligands include peptides, antibodies, and sugars, and they are usually covalently attached to the nucleic acid. If the ligand of choice interferes with the activity of the nucleic acid, this can be attached through labile functional groups that are removed after cell delivery, such as pH-sensitive or photolabile groups. In clinical trials, the most successful bioconjugation used to date has been triantennary N-acetylgalactosamine (tri-GalNAc), a trimeric carbohydrate group for the targeting of asialoglycoprotein receptor (ASGPR). This group has had the most success when covalently attached to one of the siRNA strands. to improve targeting and cellular uptake of hepatocytes, where ASGPR is over expressed [49].

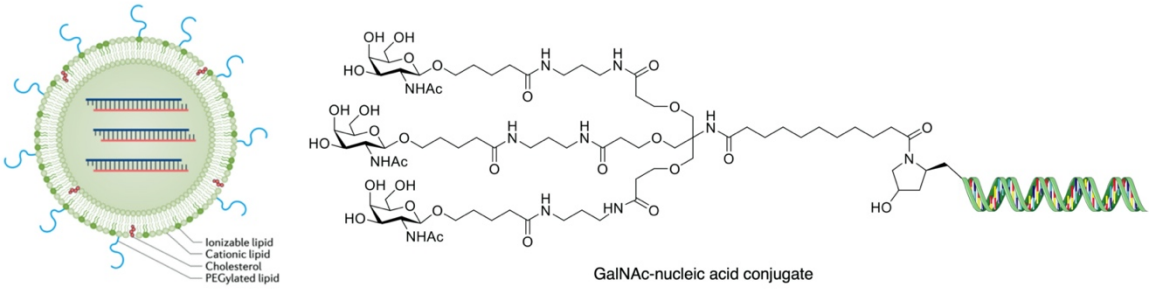


Figure 1.6: Approved nucleic acid delivery methods
Lipid nanoparticles with the most common lipid formulation on the left. Triantennary-GalNAc conjugated to a nucleic acid therapeutic on the right.

Other forms of bioconjugation being studied include aptamers and antibodies for receptor-specific targeting, and cell-penetrating peptides (CPPs) [50]. CPPs are short peptide sequences either synthetic or derived from protein translocation motifs, which can help internalizing a cargo into the cell through different mechanisms, some of them still not completely understood. Lastly, nanocarriers of different types can be conjugated to nucleic acids. One of the most promising structures are spherical nucleic acids (SNAs), hydrophobic gold or silica nanoparticles densely loaded with hydrophilic nucleic acids via thiol linkages. Due to the high density, the nucleic acids are less susceptible to nuclease degradation and have shown to deliver to cells while causing very minimum immune response [51].

1.2.4.5 Therapeutic examples

As mentioned above, lipid nanoparticles have been at the forefront of nucleic acid therapeutic delivery. Two successful clinical applications that use LNPs are described here.

Patisiran was the first RNA interference (RNAi)-based drug approved. The therapeutic siRNA was designed for the treatment of hereditary transthyretin-mediated amyloidosis (hTTR), a rare genetic disease caused by a mutation in the gene encoding the protein TTR. Due to the mutation, TTR cannot assume the correct tetrameric folding, and this results in the formation of misfolded monomeric units that accumulate in various body tissues overtime, including heart and nervous system [52]. As TTR is produced by the liver, siRNA is encapsulated in LNPs and administered via IV, spontaneously accumulating in the liver where the drug is delivered. The sequence is a twenty-one RNA double strand, modified with eleven 2'-OMe groups, and containing terminal 3'-deoxythymidines.

mRNA COVID-19 vaccine is another notable example. The mRNA encodes the spike protein gene. The LNP formulation is similar to the one used for patisiran, with ionizable lipids being the key for successful delivery and pharmacokinetics. The lipid nanoparticles are formed by the components discussed above as following: ALC-3015, an ionizable aminolipid that is protonated in the acidic endosome, favoring mRNA delivery; the other lipids used to improve biodistribution and confer the nanoparticle a stable structure are the PEGylated lipid ALC-0159; the phospholipid 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC), and cholesterol.

Other interesting clinical examples, not using lipid nanoparticles are summarized here. *Inclisiran* is another siRNA-based drug, it is designed to knockdown mutations of the PCKS9 gene to reduce LDL levels and prevent atherosclerotic cardiovascular disease (ASCVD). ASCVD causes more than 19 million deaths annually [53], and the major treatment before inclisiran was the monoclonal antibody (mAb) evolocumab, or a combination of evolocumab and another mAb. These antibody treatments require an average of one injection every two weeks, for a total of ~3640 mg of mAb per year. Contrarily, inclisiran only requires two injections a year for a total of ~600 mg of nucleic acid a year [54]. Inclisiran acts in the liver

by targeted hepatocyte delivery through tri-GalNAc groups, and it is modified with 2'-OMe, 2'-F, and terminal phosphorothioates.

Tivanisiran is a completely naked, unmodified siRNA, which is delivered via intravitreal injection, without the aid of any nanocarrier. Because administered directly to the eye, rapid degradation and excretion are avoided [55].

Voretigene neparvovec-rzyl is an AAV drug commercialized as Luxturna. It delivers the missing wild-type gene for the treatment of Leber congenital amaurosis (LCA) type 2, a congenital retinal dystrophy that results in severe vision loss at an early age and had no treatment before Luxturna. The drug is delivered by subretinal injection to the back of the eye, where it is internalized by the cell via membrane receptor interactions with the AAV capsid. A copy of the functioning gene is ultimately delivered to the nucleus [21].

In conclusion, several nucleic acid therapeutics have been approved for clinical use for the treatment of several diseases, including rare and orphan diseases, or as a competitive alternative to other existing drugs. The COVID-19 mRNA vaccine has shown how effective nucleic acid therapeutics can be. However, the limitations in drug administration and delivery and the high cost of production have negatively influenced the number of drugs approved, which corresponds to only a few every year.

1.3 Nucleic acids in synthetic biology

Beyond therapeutics, nucleic acids represent a key aspect of modern biology, especially in the growing area of synthetic biology. Synthetic biology is a discipline composed of two main

subfields, one is to reproduce natural biology processes artificially, the other to use natural biology components to create unnatural systems [56]. Synthetic biology applications are increasing every year in a variety of different fields, from engineered plant-based hamburgers, to new bio-sourced materials, to biological bacteria-based fertilizers [57].

The field of synthetic biology originated in 1961, with Francis Jacob and Jacques Monrod studies on the *E. coli* lac operon and their first understanding of regulatory circuits [58]. In the 1980s and 1990s with the exponential improvement in genetic manipulation and DNA sequencing, researchers started to reverse engineer cellular networks, artificially manipulate biological systems, envisioning biotechnological and health applications. The first real genetic circuit considered the birth of synthetic biology was built in the year 2000 by James Collins et al. [59]. It was a genetic toggle switch in *E. coli* composed by two promoters that would induce the expression of reciprocal repressors. Two promoters part of the same plasmid could be induced by two different stimuli: isopropyl ß-*D*-1-thiogalactopyranoside (IPTG) and heat. The result was to obtain the desired behavior from E. coli in response to external stimuli, without using traditional genetic engineering of proteins or other regulatory elements.

Today, with the advances in phosphoramidite chemistry for rapid and inexpensive nucleic acid synthesis, plasmids and gene circuits can be easily generated. Moreover, higher level of gene circuit complexity can be reached with gene editing, gene silencing, and other techniques, for a variety of novel applications.

1.3.1 Cell-free expression systems

Cell-free protein synthesis (CFPS) systems are a fundamental tool in synthetic biology. They allow the performing of biological reactions with defined components and without the complexity of a cellular environment [60]. Moreover, they are highly versatile and achieve rapid RNA/protein synthesis starting from natural or synthetic DNA. CFPS systems can be derived from crude cellular extracts of bacteria, insects, wheat germs, rabbit reticulocytes, or they can be reconstructed in vitro. This is done by isolating the fundamental components of the transcription and translation machineries. Transcription is reproduced by using an RNA polymerase (RNAP) and nucleotide triphosphates in an appropriate buffer. Translation is more complicated, as it requires a complex mixture of recombinant proteins, ribosomes, transfer RNAs (tRNAs), and amino acids. However, different CFPS systems are easily accessible, as they have been commercialized by different companies (e.g., the Protein synthesis Using Recombinant Elements (PURE) system).

Cell-free systems avoid all cell-based synthetic biology limitations, such as laborious genetic encoding, gene delivery, and slow design-build-test cycles. They are ideal for the expression and purification of proteins otherwise challenging to express, such as toxic proteins or proteins containing non-canonical amino acids [61]. Non-canonical amino acids can be added to a protein with CFPS by implementing orthogonal tRNAs and aminoacyl tRNA synthase in the cell-free mixture. CFPS have also been used for high-throughput drug screening, as DNA constructs can directly be added into a cell-free reaction mixture producing the enzyme that is being tested, without the need of cellular transformation or protein purification [62]. Other CFPS applications are the study of biological processes [63], and the engineering of gene circuits [64]. CFPS systems are very portable, they can be freeze-dried, applied to a solid support, or used in small aliquots. As a result of the several applications,

CFPS technologies have gained great interest from the biomanufacturing field and industry [65]. Finally, CFPS systems can be encapsulated in lipid bilayers to form synthetic cells, which have a great potential for controlled and targeted drug delivery or the study of cellular communication.

On the other hand, CFPS are not yet suited for the synthesis of proteins that require post-translational modifications to function, and cannot yet reproduce DNA self-replication mechanisms present in living cells. Moreover, even though great efforts are put into the development of large scale production CFPS technologies, the high cost to standardize and automate these systems are a substantial limitation for industry [59, 65].

1.4 Remote control of gene expression

Nucleic acids have applications in existing therapeutics and possess great potential for the development of new drugs. However, several limitations, such as stability in vivo, targeting of specific organs, and non-specificity have to be overcome for further development.

One of the most effective ways to tackle these limitations is the ability to control the nucleic acid activity remotely. By controlling a nucleic acid, the processes of transcription and translation can be controlled, therefore controlling the expression of a target gene. A nucleic acid can be completely inactivated while in circulation, to then be activated on demand when the desired cellular target is reached. Synthetic biology, and in particular cell-free expression systems would also expand their potential with the possibility of controlling gene expression externally. In the future, synthetic cells could be directed toward the target tissue, the drug

produced *in situ* and released by applying an external stimulus. More complex gene circuits could also be designed and controlled remotely by responding to different inputs. This can be done with different external inputs, such as chemical or molecular triggers, or with physical stimuli, such as temperature or light.

1.4.1 Control of gene expression using small molecules

One way of controlling gene expression is by using small molecules, as small molecules can easily cross the cellular membrane to reach their target. This have revolutionized the biotechnology field, starting from inducible protein expression with IPTG in *E. coli* [67]. Genetic circuits of this kind are usually engineered with components derived from bacteria and it therefore controlled with bacterial transcription factors. The lac (LacI), arabinose (AraC), and tetracycline (TetR) repressors are the most well-known examples, as they reliably inhibit transcription. In the absence of their ligand, the repressor protein binds to the operator sequence, which is next to the promoter sequence, preventing RNAP binding by steric hinderance. Upon addition of the controlling small molecule, the repressor is removed and RNAP can bind and start the transcription process [68]. Similar systems with different ligands have been extensively used to control the expression of several genes in gene circuits in living cells and cell-free systems [61, 62].

In cell-free systems, other small molecule inputs have been used to create RNA switches and transcription factor-based biosensors, where the analyte is the trigger for activation of expression. RNA toehold switches can be engineered for a variety of different applications. The RNA switch consists in a toehold sequence containing the gene promoter for the activation of a reporter protein. The promoter is repressed by a sequence loop

containing the ribosome-binding site (RBS). Once the target RNA sequence is present, it binds to part of the repressing loop, freeing the RBS, activating the promoter and thus the expression of the target gene [71]. This kind of biosensor can be produced low-cost, freeze-dried, and applied to a portable solid support, and it has been used for the detection of Zika virus. On the other hand, transcription factor-based biosensors are composed of modules of several plasmids encoding the following proteins: a metabolic enzyme that can convert an undetectable molecule into a ligand for transcription-factor detection, the transcription factor, and a reporter module for final visualization [72]. The modules can be designed to detect a wide range of analytes in complex medium by engineering the structure of the first metabolic enzyme. Although the systems described work well as biosensors, the addition of a small molecule to a system has several limitations. For example, for *in vivo* applications, the molecule added externally could be difficult to deliver to the desired tissue, it could be toxic to the organism, or be excreted before reaching the target. In cell-free systems, adding an activator will alter the concentration of the internal components. This also applies to enclosed systems, such as synthesis cells, where, additionally, the membrane is not permeable to many activators.

1.4.2 Control of gene expression with light

Physical stimuli such as light, temperature, and magnetism have the major advantage over small molecules of being applied externally to a closed system. For therapeutic purposes, this represents the least invasive method for the remote control of nucleic acids. Above all, light is an ideal stimulus as it is bioorthogonal to most cellular processes and has low toxicity [73]. Moreover, light can be applied and adjusted quite easily, favoring the rapid screening of a large number of reactions, an important quality for biomanufacturing and industry

applications. Unlike other remote stimuli, multiple orthogonal wavelengths of light can be applied to control different biological processes in the same system. With recently developed microirradiation laser technologies, light can be applied with extremely high spatial and temporal resolution, down to sub-micron regions of cellular compartments [74]. With all these advantages, light could be used for therapeutic applications to precisely activate nucleic acids in specific tissues, or trigger drug release from synthetic cells in particular locations. In synthetic biology, specific parts of synthetic tissues or gene circuits could be activated with light *in vitro*.

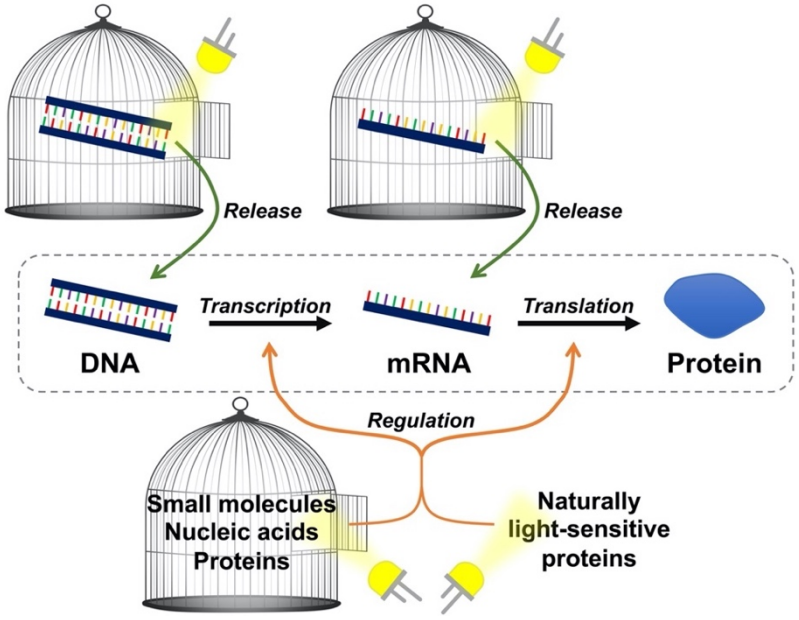


Figure 1.7: Different ways of controlling the central dogma via chemical or biological means

Transcription and translation can be controlled by chemically modifying nucleic acids with photolabile groups, by activating light-sensitive small molecule promoter/repressor triggers, or by exploiting light-triggered change in conformation of naturally light-sensitive proteins.

Gene expression can be controlled with light via chemical or biological means (Figure 1.7). Chemical methods generally consist in generating gene switches activated by light-sensitive small molecules, or in attaching light-activatable modifications to nucleic acids. On the other hand, biological methods exploit the conformational changes of light-sensitive proteins [73].

1.4.2.1 Light-activation via chemical means

Light-sensitive molecules are called chemical photo*cages*, as they *lock* the nucleic acid in an inactive conformation – the "off state" – preventing its function. Once the light stimulus is applied, a structural change in the molecule is triggered, usually the rearrangement of a ring system that leads to a bond cleavage. This trigger causes the release of the photocage from the nucleic acid, which is therefore converted to its active form – the "on state". The most common chemical photocages are 2-nitrobenzyl and coumarin derivatives, which are photoactivated at 365 and 365-400 nm respectively. These functionalities have the advantage of being very stable under natural light and their photoactivation mechanism is well-understood, as they have been used extensively as photocleavable protecting groups [69, 70]. Their synthesis and functionalization are also simple and inexpensive, which makes them ideal for their coupling to nucleic acids and other bioactive molecules. However, depending on the power of the irradiating source and the wavelength applied, short wavelengths of light can have low tissue penetration and can potentially damage living tissues, such as with UV-B light. As a result, several of these technologies have mainly been studied in zebrafish, as they are transparent and favour light tissue penetration.

As a consequence, efforts are being put towards developing new photocages that are activated at longer wavelengths, have lower toxicity, and higher tissue penetration, such as

near-infrared (NIR) light. Quinolines, Ru(bipyridine)$_2$(3-ethynylpyridine)$_2$ (RuBEP), and cyanines are cleaved at higher wavelengths of light. Cyanines in particular can be modified with more functionalities to extend their wavelength activation up to the near-infrared (NIR) spectrum. However, these photocages are not as easily prepared by synthesis or easily purchased as 2-nitrobenzyls and coumarins are. Some of the main functional groups with their correspondent wavelength of activation are shown below (Figure 1.8).

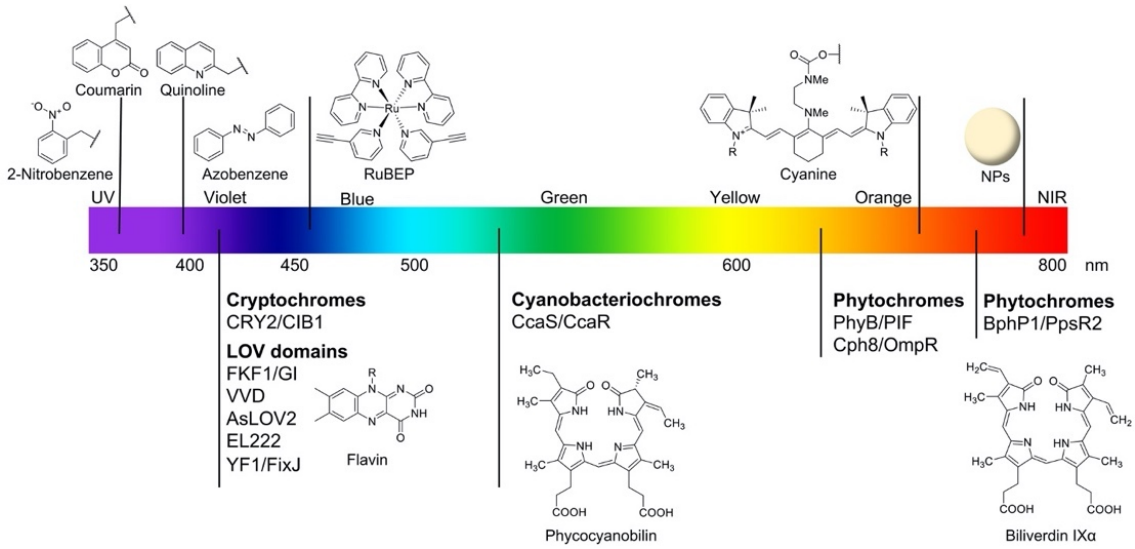


Figure 1.8: Functional groups for the activation of photocages at different wavelengths of light

The synthesis and functionalization of photocages responsive at wavelengths < 400 nm is more well-studied and more inexpensive. However, these photocages can result in cell toxicity or poor tissue penetration. Researchers are working on the synthesis of photocages responsive to higher wavelengths of light. Some functionalities, such as cyanines can be modified to cleave at a broad range of wavelengths, up to near-infrared, which makes them a very viable option for nucleic acid applications.

Upconverting nanoparticles (UCNPs) have also been used to cage, then release nucleic acids. UCNPs are lanthanide-doped NPs that can absorb NIR irradiation and release UV-vis radiation. Therefore, nucleic acids can be attached to UCNPs via 2-nitrobenzyl linkages, which are then cleaved upon application of NIR light [77].

Finally, azobenzenes have also been used to control gene expression. These structures undergo reversible conformational change once irradiated with two different wavelengths of light, typically ~400 nm and ~450 nm, depending on the substituents appended. If intercalated between DNA base pairing on a promoter, they can destabilize the DNA duplex in the *cis*-form with one wavelength, preventing gene expression. When irradiation is switched to the other wavelength, the azobenzene isomerizes to the *trans*-form, enabling DNA expression [78].

1.4.2.2 Light-activation via biological means

Control via biological means is mainly carried out by using either natural or engineered light-sensitive proteins. Commonly used photosensitive proteins contain a photosensitive tryptophan conformation, or a bound photoactive cofactor, such as flavin-based or tetrapyrrole-based chromophores [79]. A common way to control gene expression is by splitting the protein into two inactive domains. Upon light irradiation, a conformational change is triggered, bringing the two domains together and activating protein activity. Proteins can also be engineered to introduce unnatural amino acids in their sequence via orthogonal ribosomes and tRNA/tRNA synthase pairs. These amino acids can contain a photocage in the active site of the protein, making it inactive until irradiation with light.

1.4.2.3 Light-activated nucleic acid examples

Small molecule ligands for promoters/repressors can be modified with photocages. Several of these technologies have been showed to work from cell-free expression systems to *in vivo*,

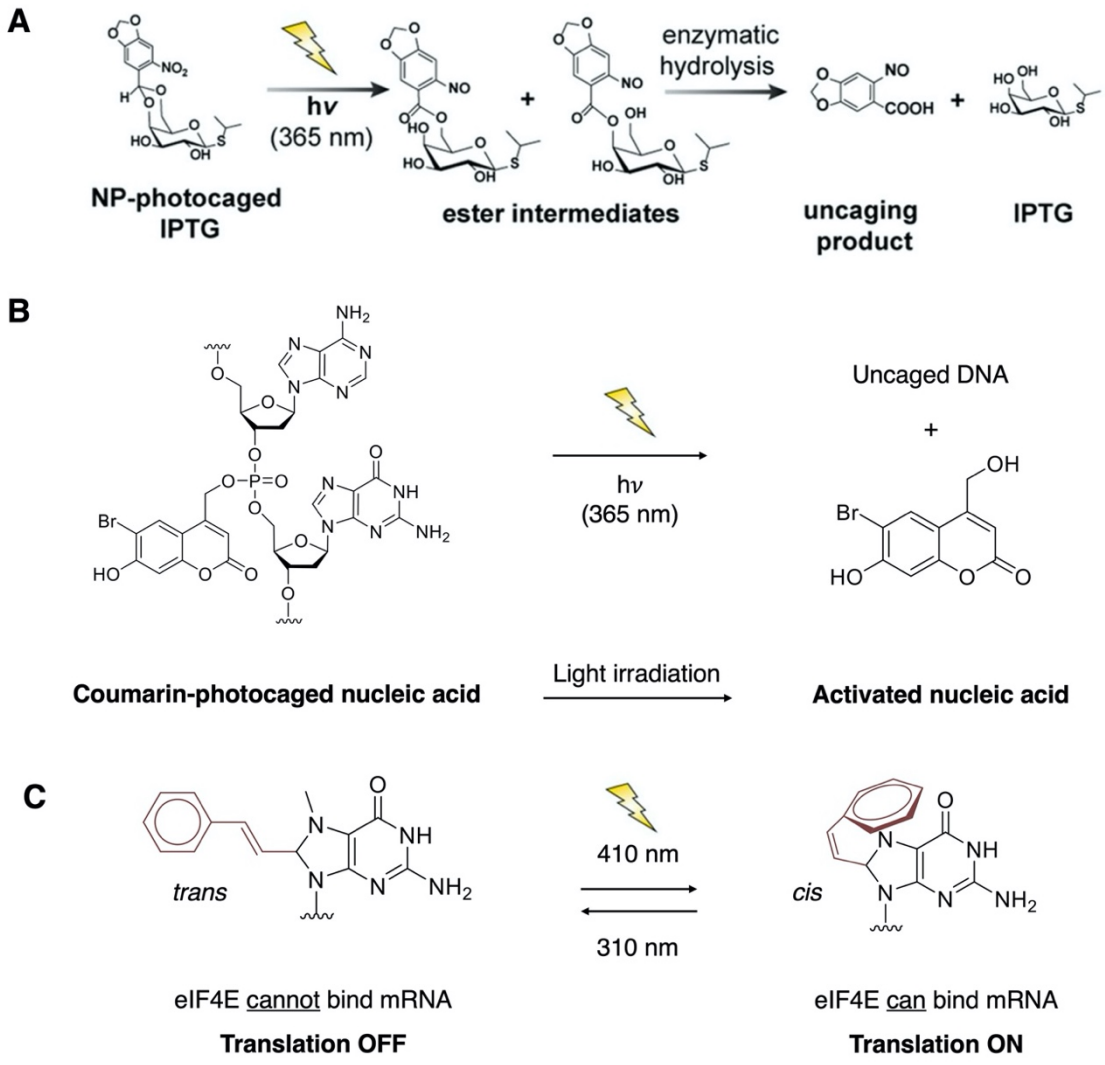


Figure 1.9: Methods to control gene expression with light

Some of the methods to control gene expression with light discussed in the text. A) in a bacteria culture, IPTG is caged with a nitrophenyl (NP) derivative. Once photoactivated, the NP group is photocleaved and enzymes present complete the release of free IPTG, which activates gene expression. B) Coumarin derivatives are present along the phosphate backbone of nucleic acids, making them inactive. The nucleic acids are activated by UV irradiation. C) mRNA translation is blocked by a 5'-cap in the *trans* configuration. Light triggers a photoisomerization to *cis*, allowing the eukaryotic initiation factor 4E to bind mRNA, hence activating translation.

especially in zebrafish as they are transparent and favour light tissue penetration. For example, 2-nitrobenzyls have been used for to control tetracycline- or IPTG-induced systems [72, 73]. IPTG or tetracycline are modified with 2-nitrobenzyl cages and are added to the system in their photocaged, inactive form (Figure 1.9A). Upon UV irradiation, IPTG or tetracycline are uncaged and free to bind their respective transcription repressor, activating gene expression. With a similar mechanism, coumarin photocages have been attached to tamoxifen derivatives – an estradiol analog – to control the activation of estrogen receptors [82].

Nucleic acids can also be photocaged by directly modifying their structure. Photocages can be attached to different parts of the nucleic acid structure, as discussed for other types of nucleic acid modifications in Section 1.2.4.3. Different research groups have worked on controlling gene expression by attaching 2-nitrobenzyl or coumarin derivatives to the phosphate backbone (Figure 1.9B) [75, 76]. However, even though the off state inhibited the nucleic acid activity well, the activity could be recovered only partially after light irradiation. The 5'-cap of mRNA can also be modified with light activated moieties, as it is necessary for translation (Figure 1.9C) [84, 85]. Nucleobases can be photocaged by attaching 2-nitrobenzyl moieties to 5-hydroxy-methylcytosine or 5-hydroxy-methyluracil [87]. This type of photocaging does not affect base pairing; instead, the caging groups block RNA polymerase access to the major groove.

The strategy of modification of the nucleobases has also been exploited by our group to control cell-free expression in synthetic tissues and synthetic cells [81, 82]. Amino-modified deoxythymidine bases part of a DNA promoter sequence were reacted with 2-nitrobenzyl photocages, containing a terminal biotin (Figure 1.10). This light-activated primer was used to amplify a gene of interest by PCR. The resulting DNA template was then bound to streptavidin, which turns off gene expression by steric hinderance. Upon UV irradiation, the photocages containing the streptavidin are cleaved off, and gene expression is activated

(Figure 1.9A/C). With this method, any gene activated by this modified promoter could potentially be controlled with light, by repeating the process described. In the examples mentioned, seven streptavidin photocages were added on a DNA template with a T7 promoter for the best off-state results.

A coumarin, blue light-activated photocage was also developed by our group [90]. With the same method of modifying a primer containing amines and integrate it into a DNA template by PCR, two DNA templates were generated for the cell-free synthesis of the α- and ω-subunit of a split ß-galactosidase. One of the DNA sequences was modified with UV

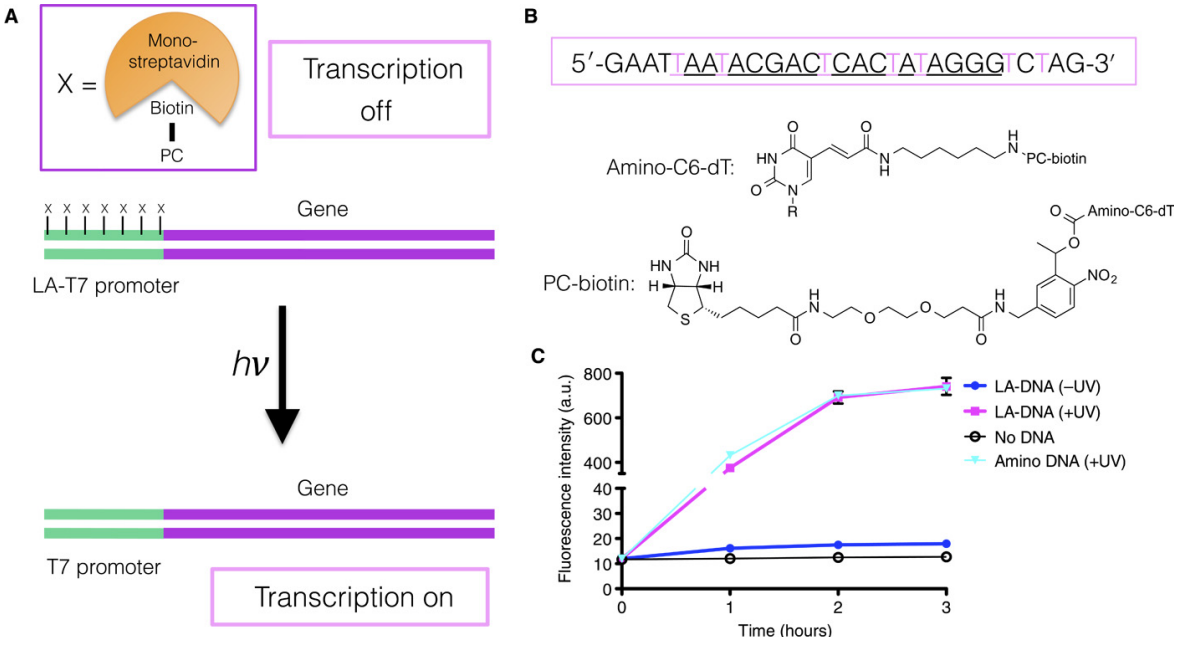


Figure 1.10: Light-activated transcription in a cell-free expression system
A) The T7 promoter of a linear DNA template is modified with photocleavable (PC) streptavidin groups, forming a light-activated (LA)-DNA. B) To form LA-DNA, a PCR primer is modified with C6-amino-deoxythimidines, then reacted with a PC-biotin linker, before binding streptavidin. C) The cell-free production of a fluorescent protein is measured in various conditions.

photocages, while the other with blue light photocages. Only when both UV and blue light were irradiated, and both of the protein subunits expressed could the ß-galactosidase correctly assemble and convert its substrate. The two photocages used were also orthogonal, as the irradiation with only one of the two wavelengths would not cause the substrate to react. This generated a dual-wavelength controlled cell-free DNA-based AND-gate.

1.3 Summary of thesis work

The overall aim of this work was to generate light-activated (LA)-antisense oligonucleotides (ASOs) and LA-small interfering RNAs (siRNAs) to control gene knockdown respectively in cell-free and living systems using light. By modifying ASOs with photocages responsive to orthogonal wavelengths, one of the goals of this work was to control the knockdown of two different genes in cell-free systems using orthogonal inputs. The other goal was to combine LA-ASOs with a LA-DNA template to control the activation and deactivation of protein synthesis at different expression levels, also by using different wavelengths of light. Additionally, the goal of generating light-activated small interfering RNAs was to control gene silencing in mammalian cells with UV. Here, a strategy for siRNA self-delivery was also included. This work will be partitioned into the following chapters:

Chapter 2 describes a method for the modification and purification of ASOs with UV- and blue light-active photocages. The ASOs activity and photoactivation efficiency are then evaluated using different cell-free systems. By using a UV-ASO coupled to a blue-activated

DNA template, the activation of gene expression with one wavelength, followed by its deactivation with a different wavelength are shown. Finally, two distinct ASOs targeting two different mRNAs are shown to control the knockdown of two different genes with orthogonal light irradiation in cell-free protein synthesis systems.

Chapter 3 describes the development of a strategy to generate siRNAs that are both light-controllable and can be delivered to different mammalian cell lines without the use of transfection reagents. A nucleic acid-protein construct is generated, to which cell-penetrating peptides are attached for self-delivery. The effect on the gene knockdown efficiency of modifications at different positions along the siRNA sequence is also evaluated. By binding a fluorophore to the light-activated siRNA, the construct is traceable by microscopy. Images showing the successful delivery of the nucleic acid are presented.

2

Orthogonal on/off control of cell-free expression using photocaged nucleic acids

2.1 Historical remarks

The inhibition of translation in a cell-free system was reported for the first time in 1977 and named “hybrid-arrested translation”, as it was caused by the hybridization of single-stranded DNA and mRNA [91]. However, the discovery of antisense oligonucleotides (ASOs) as therapeutics is assigned to Paul Zamecnik for his work in 1978, when the first 13-nucleotide antisense oligonucleotide (ASO) was synthesized and used in tissue culture to inhibit the

translation of Rous sarcoma RNA virus, which is crucial for the virus replication. The 5'- and 3'-terminal hydroxyl groups were also modified to increase the ASO stability in the cell, a technique still in use today for nucleic acid therapeutics [2, 3].

After two decades studying the mechanism of action of ASOs and optimizing cell delivery and modifications needed for the ASO stability, fomivirsen, the first ASO drug was approved in 1998 to prevent the replication of human cytomegalovirus (CMV) in patients with AIDS and other immunodeficiencies [92]. The virus causes cytomegalovirus retinitis, a sight-threatening disease caused by the progressive destruction of retinal cells. Fomivirsen is delivered on site by intravitreal injection, which also conveniently allows the use of only few ASO modifications for stability. Precisely, the backbone of the 21-nucleotide ASO is fully modified with phosphorothioates (PS), but there is no 2'-OH modification, which would be needed to prevent degradation in blood vessels but is not for eye delivery. Fomivirsen was discontinued soon after its approval due to its high cost. Instead, more comprehensive and more highly active antiretroviral therapies are in use today, usually containing multiple drugs acting at different stages of the HIV life cycle [47]. After fomivirsen, several other ASO-based therapeutics have been approved by the EMA and FDA for the treatment of different diseases based on their mode of action.

ASOs are short oligodeoxynucleotides (~18-30 nts) that can be classified into three generation categories based on the chemical modifications applied [93], all of which are currently in use. Fomivirsen is part of the first generation ASOs modified on the phosphate backbone, usually with one of the non-bridging oxygens substituted by a sulphur to form phosphorothioates. The second generation ASOs are substituted at the 2'-OH, while the third generation includes all the more structural modifications, such as LNAs, PMOs and PNAs, often used for alternative splicing mechanism. Gapmers are another interesting alternative structure. They are composed of unmodified or fully PS-modified DNA sequences flanked

by RNA sequences. This structure increases base pairing affinity increasing the ASO potency, while also reducing nuclease degradation.

2.2 RNase H-mediated mRNA degradation

The most common mechanism for ASO gene knockdown is the RNase H-mediated mRNA degradation. RNases H are non-sequence specific endonucleases endogenous to nearly all organisms, with a highly conserved structure, and they are divided in three families called HI, HII, and HIII. This enzyme is present in both the nucleus and cytoplasm, and it can also be found in subcellular compartments, such as mitochondria [94]. The RNase H endogenous cellular function is to remove RNA primers from Okazaki fragments and contribute to regulating other mechanisms for DNA replication. In retroviruses, RNase H is encoded as part of the reverse transcriptase [9, 10].

When a synthetic ASO is delivered to the cell, it generally distributes in the cytoplasm to target mRNA and in the nucleus to target ncRNAs and pre-mRNA. RNase H is a divalent Mg^{2+} or Mn^{2+} cation-dependent endonuclease, it recognizes the DNA:RNA hybrid duplex and cleaves the phosphodiester bond of the RNA strand, leaving 3'-hydroxyl and 5'-phosphate groups. This way, the mRNA is degraded and the synthesis of the corresponding protein inhibited. This is the mechanism exploited by fomivirsen and other commercialized drugs such as mipomersen. Mipomersen is a gapmer that was introduced in 2013 and it is formed by ten DNA central bases flanked by five RNA bases on each side, with its backbone fully PS-modified and the flanking RNA bases containing 2'-MOE and 5-methyl pyrimidine. It is administered subcutaneously, and it lowers the level of total

cholesterol by targeting the mRNA encoding apolipoprotein B-100 in the liver, which is the main component of low-density lipoprotein (LDL). Mipomersen has had a great impact as it offers a valid alternative to people intolerant to statin, which is the most efficient and almost only other therapy prescribed for this purpose [97].

2.3 Other ASO mechanisms

Antisense oligonucleotides can be designed to bind to an RNA target with high affinity without recruiting RNase H, but instead sterically block interactions between the target mRNA and other proteins. This can be used to simply block access to the ribosome binding site, or to regulate pre-mRNA splicing in the nucleus. Pre-mRNA splicing is a post-transcriptional process that occurs for more than 90% of human genes, before the mature mRNA can exit the nucleus to be translated in the cytoplasm [98]. Through splicing, eukaryotic organisms can expand the variety of proteins encoded in a single gene, by generating multiple transcript isoforms from that same gene. The pre-mRNA is formed by a series of sequences that ultimately form the mature mRNA called exons, and spaced from one another by sequences of variable size called introns. During the splicing process, RNA-binding proteins called splicing factors remove introns to chain exons together forming mRNA. ASOs can be used to bind the pre-mRNA and block the binding of splicing factors, therefore changing the abundance of a particular gene isoform and the expression level of the corresponding protein. For beneficial therapeutic effects a specific splice variant can be upregulated or downregulated through this mechanism [99]. Great examples of this type of therapy are eteplirsen and golodirsen, which are two phosphorodiamidate morpholino

oligonucleotides (PMOs) targeting exon 51 and 53 respectively of the pre-mRNA of dystrophin, a structural protein present mainly in skeletal and cardiac muscles and encoded by the largest human gene. Exon deletions in patients result in the aberrant expression of dystrophin, leading to Duchenne muscular dystrophy in 70% of the cases. Eteplirsen and golodirsen trigger a mechanism called exon skipping. By binding to the target exon containing a premature stop codon, they block the exon from being spliced. The splicing factors skip that exon and proceed to the next one and this results in the expression of a shortened but functional version of dystrophin [14, 15]. The advantage of using PMOs is that they have a neutral backbone, which helps their delivery, they are resistant to nucleases, and water soluble. Once hybridized to RNA they are not recognized by RNase H. The disadvantage of this type of therapy is that the active ASOs have to be determined empirically, and the process can be long and costly.

2.4 Antisense oligonucleotides in synthetic biology

Despite the potential of antisense oligonucleotides as tools in synthetic biology and synthetic cells, the literature is lacking ASOs applications in this field. ASOs are very short, stable, and inexpensive and only small amounts are required to achieve mRNA degradation. RNase H is also a very accessible, stable, and cheap enzyme. Therefore, screening of vast libraries of ASOs in cell-free expression (CFE) systems with RNase H could be achievable. This could be exploited to optimize an ASOs sequences, modifications, and mode of action for later use as therapeutics. Moreover, ASOs could be used in the design of logic gates to add a higher level of complexity in regulating gene expression in gene circuits.

One example of ASOs used in cell-free systems is the transfection of ASOs in synthetic cells using lipofectamine [102]. Even though a very high concentration of ASO was needed to cause gene knockdown (10 μM) in synthetic cells, this was carried out as proof of concept to simulate transfection in tissue culture and to screen the activity of several modified ASOs.

Another interesting study aimed at evaluating the efficacy and turnover capability of different LNA-ASOs in cell-free expression. Once the best sequences and modifications were determined cell-free, the ASO best candidates were used for gene knockdown in mice [103]. The *in vitro* results corresponded to the *in vivo* knockdown, demonstrating the validity of this approach.

Even considering these several useful applications of ASOs in cell-free systems, the literature is lacking some more substantial work in this field.

2.5 Light-activated antisense oligonucleotides

Antisense oligonucleotides can be caged with photocleavable groups and activated with specific wavelengths of light as other nucleic acids. For ASOs targeting a certain mRNA expressing a protein, the OFF state is represented by no change in mRNA and protein expression levels, whereas the ON state corresponds to the ASO activation and therefore protein knockdown with lower level of mRNA and protein expressed.

One of the strategies used to generate UV-photoactivated ASOs has been the development of circular ASOs. One of these ASOs has been designed by connecting the active DNA strand to a partially complementary and inactive strand through a

photocleavable (PC) linker (Figure 2.1A). Once UV light is shone breaking the PC link, the target mRNA present in the mixture and with higher affinity for the ASO can anneal with it and get degraded by RNase H [104]. Other examples are when the PC linker connects the 5'- and 3'-end of the ASO, locking it into an inactive circular configuration (Figure 2.1B) [19, 20]. In this example, the photorelease was measured using FRET by attaching two different fluorophores to the two opposite ends of the ASO.

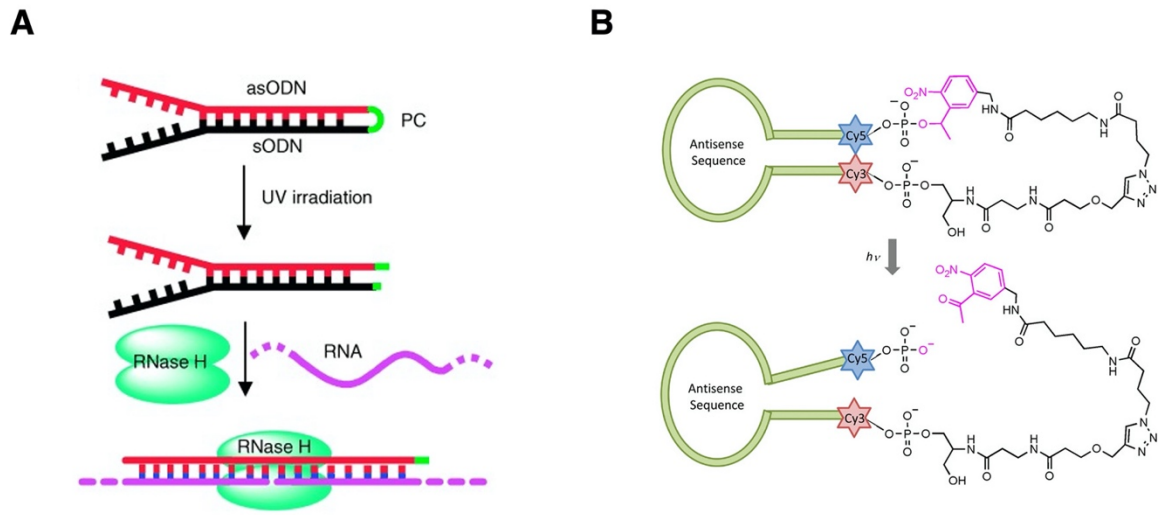


Figure 2.1: UV activation of circular photocleavable ASOs
A) The antisense oligodeoxynucleotide (asODN) is annealed to an inactive and not fully complementary single DNA strand via a photocleavable link. Once UV is applied, the antisense ODN forms a duplex with the mRNA present instead, causing its degradation via the RNase H mechanism. B) The 5'- and 3'-ends of an ASO are circularized via a photocleavable linker, making the ASO inactive. Once the PC is cleaved, the ASO becomes linear and free to bind mRNA.

These light-activated ASOs have shown some success. The issue with these structures derives from the fact that even when circularized or partially annealed to another strand, at 37 °C the ASO has enough freedom to find partial base complementarity with the target mRNA, and the really active RNase H can quickly find and cleave the mRNA.

Moreover, these structures are quite difficult to synthesize and require modified phosphoramidites, which can increase their cost.

Our research group developed a photocaged ASO with a different strategy. It was modified with photocaged phosphorothioates to be used in a cell-free system, similar to the one used in this work. The presence of neutral photocages on the ASO backbone decreases the overall ASO negative charge to potentially help with cell delivery, and allows control of cell-free protein expression with light [36].

2.6 Orthogonal systems in synthetic biology

The concept of orthogonality is at the core of synthetic biology. A protein function can be controlled by using analogues of the protein endogenous ligand, but this can lead to interference with other cellular processes. A higher level of control is achieved by engineering orthogonal enzymes or receptors that only respond to the exogenous ligand [107]. Orthogonal systems are also engineered around the central dogma. For example, by engineering translation systems with orthogonal aminoacyl-tRNA synthetases and tRNAs, unnatural amino acids can be inserted in proteins [108]. In principle, there could be two distinct genetic codes running in the same cell.

Orthogonal gene switches responding to light have been constructed in mammalian cells for the control of three protein dimers responding to UV/blue/NIR light [109], and in zebrafish to control the formation of two different phenotypes by using two coumarin gene switches responsive to two slightly different wavelengths of light [110]. Orthogonal light inputs could be used in cell-free systems to achieve activation and deactivation with different

wavelengths of light. This could create a new level of complexity for gene circuits, or it could be used in synthetic cells to time the production of a beneficial gene within a therapeutic window, then cease it when necessary. Azobenzene-modified bases have been attached to DNA for this purpose, with the azobenzene moiety reversibly intercalating between the DNA strands and blocking transcription when irradiated at 400 nm, or coming off and allowing transcription with 450 nm (Figure 2.2). However, this system does not have a very tight off state and the RNA synthesized cannot be reversibly degraded [78].

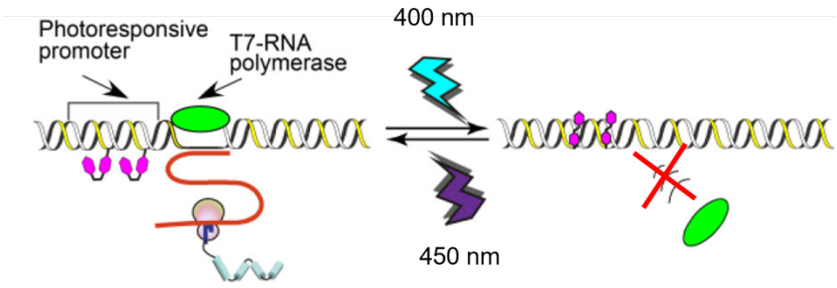


Figure 2.2: Reversible activation of transcription using azobenzenes
The DNA T7 promoter is modified with azobenzene moieties, which can intercalate in the double strand reversibly. Two different and orthogonal wavelengths can then be used to control T7 RNA polymerase binding.

Recently, our group worked on the formation of a blue light-activated photocage. The previously developed UV and the blue light photocages were attached to the T7 promoter of two DNA templates forming two complementary halves of ß-galactosidase, which would fully form and function only when both wavelengths were activated, creating an AND-gate [90].

2.7 Chapter overview

In this chapter, light-activated (LA)-antisense oligonucleotides were prepared by attaching photocages to commercially available amino-modified ASOs. The photocages (PC) were a photolabile moiety, either a 2-nitrobenzyl group responding to UV-A light (365 nm), or a coumarin-derivative responding to blue light (455 nm), with a terminal biotin/streptavidin. The bulky streptavidin stops any interaction between the ASO and RNase H, hence preventing mRNA degradation and allowing DNA expression. Several 16-22 nucleotide long amino-ASOs were screened, containing 3-4 amine modifications distributed at different positions in the sequence. Initially, UV light-activated (uvLA)-ASOs were used to turn off the expression of mVenus, first by degrading the mRNA product formed by *in vitro* transcription (IVT) reactions, then by controlling the level of protein expression with UV irradiation at different times (Figure 2.3). Secondly, in combination with a blue-light activated (bLA)-DNA template prepared by Denis Hartmann from our group, mVenus expression could be turned on with 455 nm light, and orthogonally turned off with 365 nm. Lastly, by modifying two different ASOs with UV and blue light photocages, the expression of mVenus and mCherry could be orthogonally and precisely deactivated, both in IVT and in vitro transcription and translation (IVTT) systems.

By modifying the DNA sequences used, the methods described could potentially be adapted to any mRNA/protein target. With this method, modified-ASOs can be generated quickly and inexpensively, their sequence easily designed and modified, favouring rapid screening and optimization. Moreover, phosphorothioate or 2'-OH modifications could be inserted and the ASOs tested as a screening method for gene knockdown in mammalian cells.

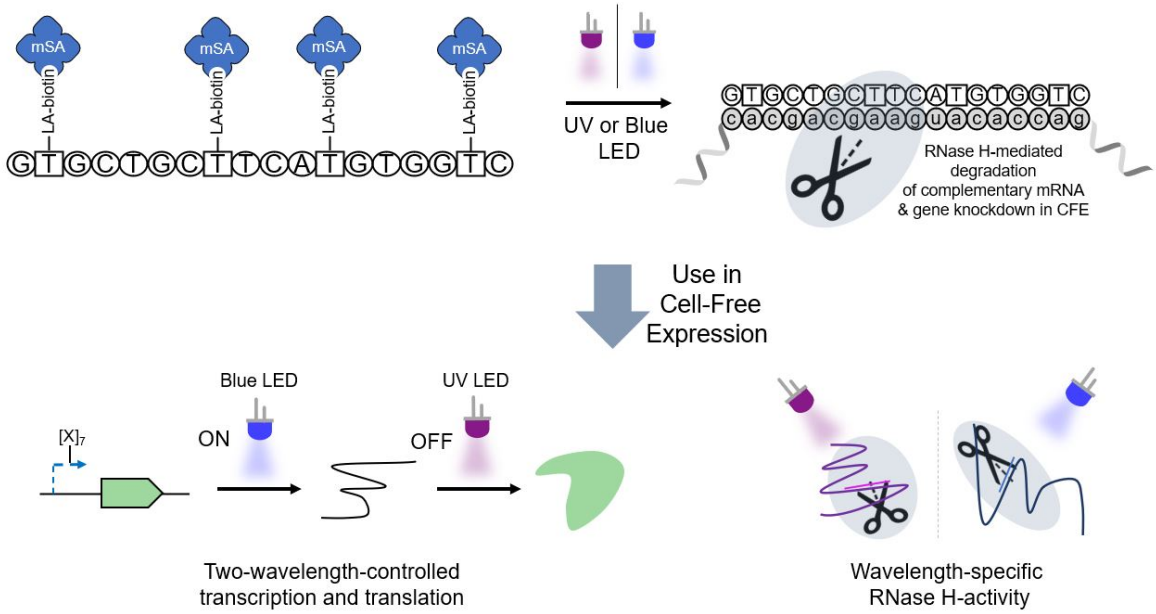


Figure 2.3. Orthogonal light-activated ASOs for the control of cell-free expression
General scheme of the project. ASOs are modified with UV or blue light-activated photocages and used in combination with a photoactive DNA to control gene expression orthogonally in cell-free systems.

This work was carried out in collaboration with Denis Hartmann, who contributed by running some of the HPLC purifications and LC-MS analyses showed, carrying out some of the final IVTT experiments showed, and designing the figure schematics. Denis also synthesized the coumarin derivative for blue light-activation and generated the bLA-DNA template.

2.8 Preparation of modified ASOs

As the goal was to control cell-free protein synthesis (CFPS) in *in vitro* transcription and translation (IVTT) systems, the protein mVenus was chosen as model protein, since the amounts of mVenus produced could easily be determined with common plate readers. This protein derives from *Aequorea victoria* green fluorescent protein (GFP) and it is spontaneously fluorescent. It is also very bright compared to other fluorescent proteins, with good sensitivity even at low levels of protein [111]. Moreover, mVenus is relatively short and does not require particular conditions to fold and fluoresce correctly, making it a good target for our purposes overall.

Plasmids containing the mVenus gene under a T7 promoter were already available through other group members, and a linear T7 mVenus-DNA template was easily obtained by PCR using standard procedures. This was then transcribed into the ~1 kb long mVenus-mRNA using a commercial T7 RNA polymerase and triphosphates. The first ASO screening to target mVenus was done previously in our group [36], using the following criteria: the ASO had to be 16-24 nucleotides long and it had to bind the initial part of the mRNA sequence to avoid expression of longer, truncated, partially fluorescent proteins. Furthermore, since each amine modification would later be introduced as a modified-deoxythymidine, the sequence had to contain as many dTs as possible. Ideally, the thymidine bases would be spread throughout the sequence to test ASOs containing modifications in different number and positions. The screening was carried out in standard RNase H-mediated mRNA degradation conditions, using a ratio of mVenus-mRNA:ASO of 1:4. The conditions were equivalent to the ones used in this work and can be found in the Methods chapter (Chapter 5). After this preliminary screening, the 20-mer **mV-220** resulted in the

best candidate, the name indicating the binding starting position (220) on mVenus-mRNA, with position 1 being the A of the AUG start codon:

mV-220

TTC ATG TGG TCG GGG TAG CG

Three different versions of this oligonucleotide were purchased, each containing aminohexyl modifications in different positions. Terminal modifications were attached to the 5'-end phosphate group, while internal modifications to the deoxythymidine methyl group (Figure 2.4). The modified sequences used **V1**, **V1b** and **V1c** are shown below:

Entry	Target	Nucleotide position on mRNA sequence	Length in number of bases	Number of amines	Sequence (5' → 3')	
1	mV	220	20	0	TTCATGTGGTCGGGGTAGCG	
2	mV	220	20	3	TXCATGTGGXCGGGGXAGCG	→ **V1**
3	mV	220	20	3	2XCAXGTGGTCGGGGTAGCG	→ **V1b**
4	mV	220	20	4	2XCAXGTGGXCGGGGTAGCG	→ **V1c**

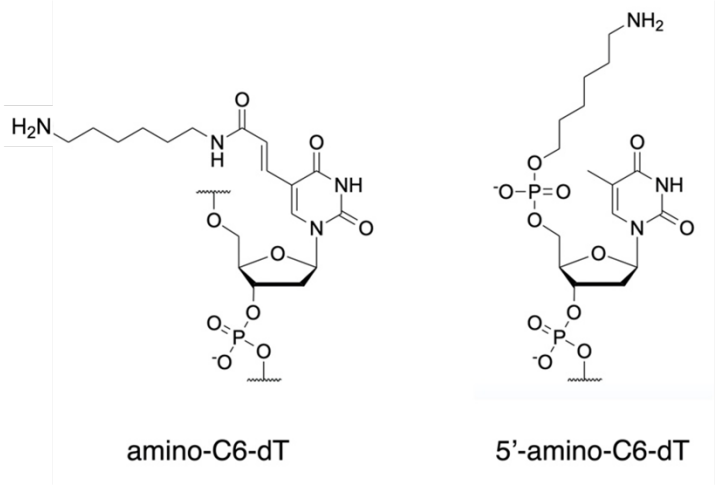


Figure 2.4. Amino C6-dT and 5'-amino-C6-dT structures
Structure of the amino-modified bases used in this work.

The amino-ASOs were then reacted with 2-nitrobenzyl N-hydroxysuccinimide (NHS) ester groups containing a terminal biotin, named photocleavable biotin (PCB) groups. The products were analyzed, and purified by HPLC (Figure 2.5). During HPLC purification,

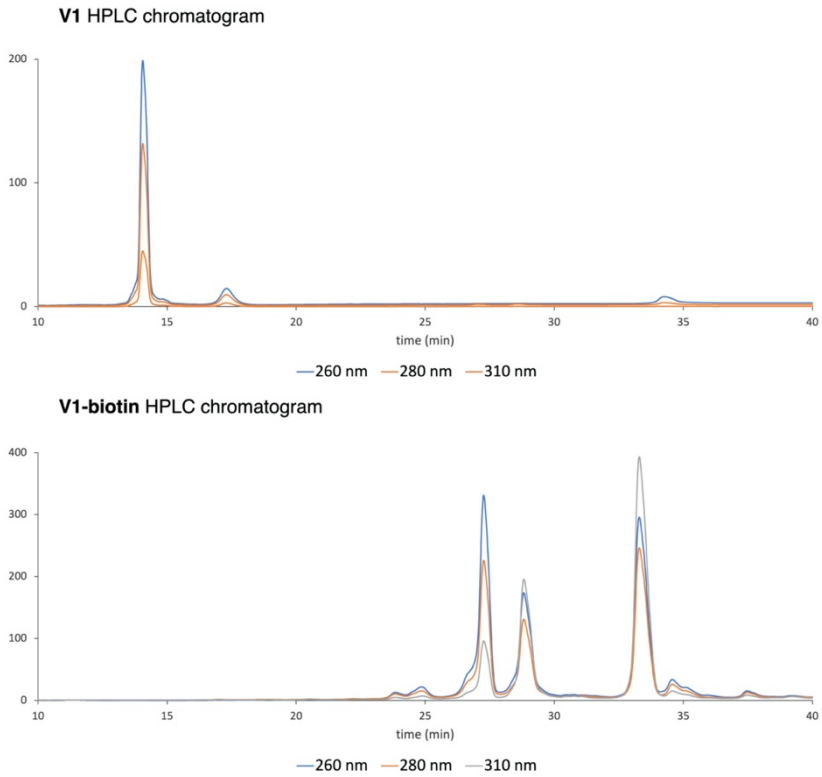


Figure 2.5. HPLC chromatogram of V1-biotin crude

HPLC chromatogram of the starting amino-ASO V1, then reacted with PCB-NHS ester to yield V1-biotin. The reaction crude shows no starting material left, and the product at ~27 min was later purified. Since V1 contains three reactive amine sites, by analyzing the peak shape, the product contained a shoulder corresponding to partial reaction, where only 1/3 or 2/3 amines reacted.

the DNA peak was easily identified because it showed a characteristic 260/280/310 nm absorbance ratio. The other peaks were by-products of the PCB-NHS ester. Figure 2.5 shows the HPLC chromatogram of the amino-ASO **V1** before NHS reaction, then the crude mixture of fully reacted product **V1-biotin** at ~27 min time. After HPLC purification, lyophilization and resuspension in buffer, the product was analyzed by polyacrylamide gel electrophoresis (PAGE) (Figure 2.6A). The same procedure was repeated for **V1b** and **V1c** (Figure 2.6B). The fully modified **V1-biotin** product corresponds to all three amines of **V1** modified with PCB. However, the HPLC chromatogram shows a small, but visible shoulder before the

main peak at ~27 min. The shoulder corresponds to the product with only 1/3 or 2/3 modified amines, and although most of this minor product could be separated during the HPLC purification, a small percentage of the impurity remained, slightly affecting some of the results. The HPLC method used was a gradient of 5-40% water/acetonitrile (ACN) in ammonium acetate (NH_4OAc) over the course of 40 min at room temperature. This method was later improved by heating the HPLC column during purification, changing the buffer to triethylammonium bicarbonate (TEAB) and by decreasing the gradient curve.

In the gel, the reaction of the amino-ASO with the PCB-NHS ester corresponds to a band shift for the increased molecular weight. After UV irradiation, the PCB groups is cleaved off and the resulting band indicates the restoring of the amino-ASO.

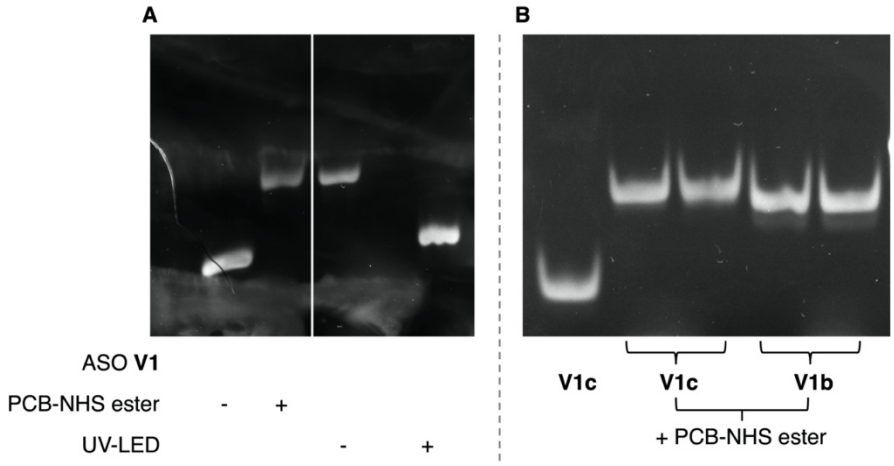


Figure 2.6. PAGE analysis of modified V1, V1b, and V1c
PAGE analysis showing different modified ASOs before and after reaction with the PCB-NHS ester. The products were loaded after HPLC purification.

2.9 RNase H-mediated mRNA degradation tests

To test the efficiency of the modified ASOs, RNase H-mediated mRNA degradation reactions were carried out. Initially, **V1** was tested against mVenus-mRNA, then for preliminary streptavidin binding studies and response to UV light. The first RNase H-mediated mRNA degradation reaction showed how the presence of three amines on the ASO is well tolerated by the enzyme, as shown in figure 2.7. As the mRNA is around 1 kb long and the ASO binds starting from position 220, the gel bands of the resulting fragments after incubation with RNase H correspond to ~250 and ~750 nucleotides and yield the specific ASO band signature on agarose gel. The reaction tested was carried out with 5 units of RNase H (unit defined as the amount of enzyme required to produce 1 nmol of cleaved product in 20 min, at 30 °C and pH 7.7, using a model oligonucleotide as substrate) and a ratio of 4:1 mRNA/ASO. As the conditions lead to a good, but incomplete degradation of mRNA, an increasing number of RNase H units was used for the next testing with streptavidin.

Streptavidin is a tetrameric protein, and monovalent streptavidin was used to bind the biotinylated ASO. This is a variant of the protein in which three of the four monomers composing the tetramer are substituted by inactive or "dead" monomers. These are still able to form the same tetrameric structure, but only one monomer has an active site that can bind biotin. It is often used to avoid formation of higher structures formed by the same streptavidin binding to multiple biotinylated nucleic acids. Monovalent streptavidin was kindly provided by Dr. Mark Howarth lab, Department of Biochemistry, Oxford.

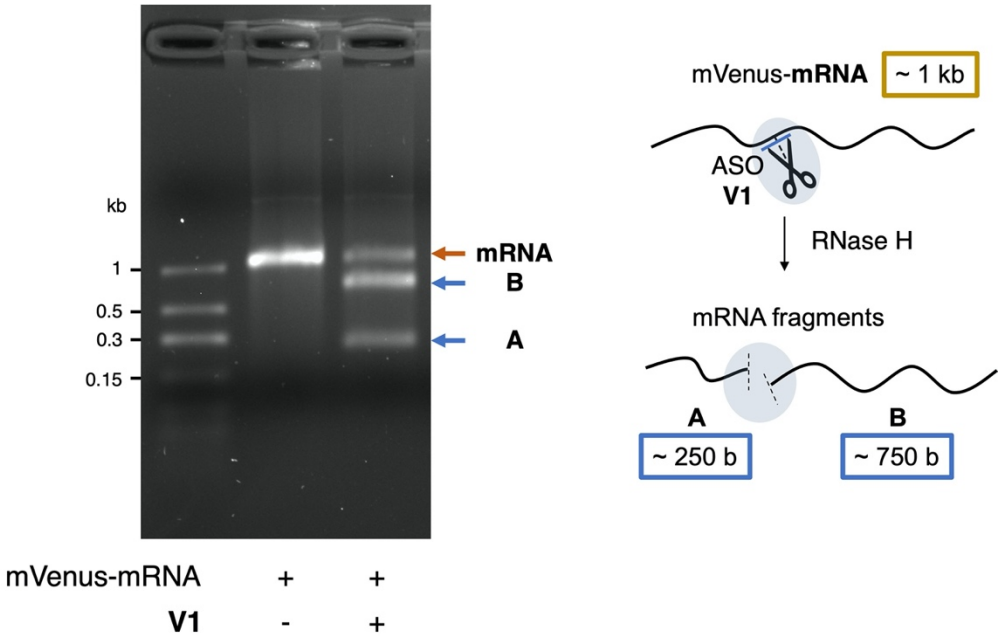


Figure 2.7. RNase H-mediated mRNA degradation

On the left, agarose gel to analyze the mVenus degradation reaction. Purified mRNA is shown intact in the absence of ASO. Once the ASO is added, a signature degradation pattern is observed. A molecular weight marker is shown on the first lane on the left. On the right, a scheme representing the reason of the signature degradation pattern. The ASO binds at around position 220 of the mVenus-mRNA sequence; here, RNase H binds and the degradation occurs. The remaining fragments size are indicated in the figure and labeled as A and B.

The binding of streptavidin to the PCB-ASO had the desired effect of blocking any interaction with RNase H, inhibiting mRNA degradation (Figure 2.8). The degradation was restored once UV irradiation is applied. The ASO knockdown efficiency after applying UV was almost equivalent to the one of the amino-ASO. However, adding ten to forty units of enzyme caused the mRNA to degrade excessively and with no specificity. This is clearly visible with twenty and forty RNase H units, where most of the degradation product corresponds to the ~250 bp band or further degradation resulting in a smear. For this reason, the following experiments were carried out with six units of RNase H.

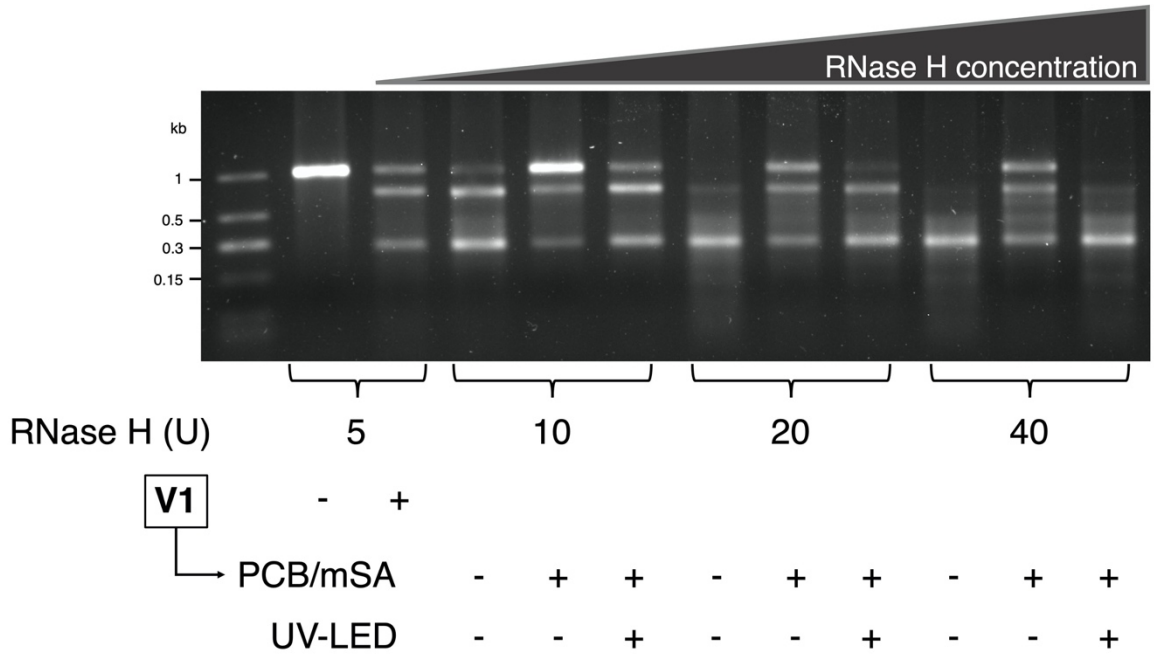


Figure 2.8. ON and OFF states at increasing RNase H concentrations
Increasing amounts of RNase H were tested in combination with the application of UV. The streptavidin-modified ASOs were compared to amino-ASOs (positive control). The addition of streptavidin clearly inhibits mRNA degradation, while applying UV restores the degradation at levels comparable to the positive control of amino-ASOs. The amount of RNase H added was reduced for the following experiments.

Finally, the three different PCB-ASOs **V1**, **V1b**, and **V1c** prepared and purified in the previous section were compared. After HPLC purification, they were incubated with either monovalent or tetravalent/wild-type streptavidin (mSA or tetSA, respectively). The now UV light-activated ASOs were named uvLA-**V1**, uvLA-**V1b** and uvLA-**V1c**, incubated with RNase H, and the mixture irradiated (or not) with UV light. The agarose gel analysis showed a tight OFF state for **V1** and **V1c**, with no mRNA degradation. Upon UV illumination, the ON state was better for **V1** than **V1c** most likely because of **V1** only having three groups coming off versus the four of **V1c**. Lastly, for all three ASOs, LA-groups containing tetSA cleaved off less efficiently than the ones with mSA (Figure 2.9).

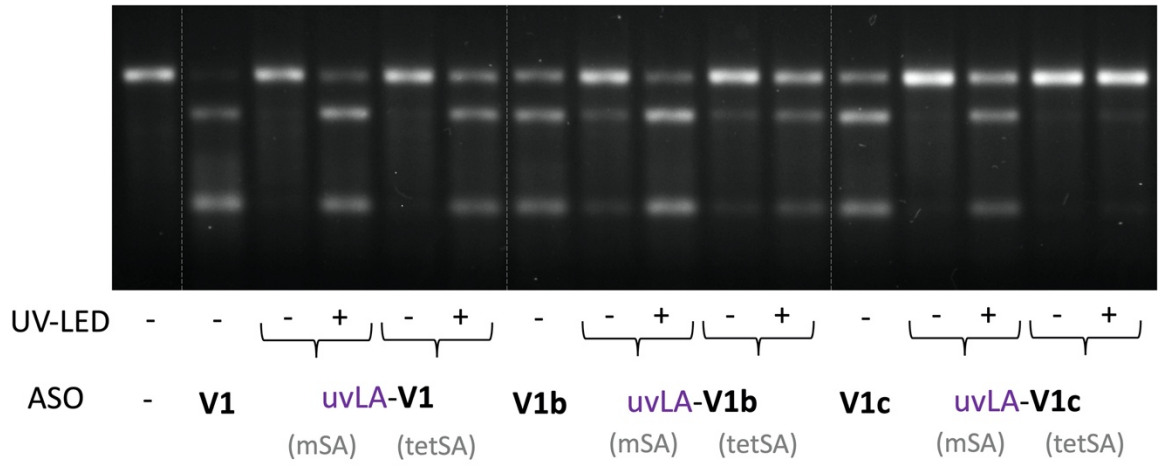


Figure 2.9. RNase H-mediated mRNA degradation to compare different uvLA-ASOs

Comparison of efficiency in mRNA degradation of different ASOs sequences, modified with either monovalent or tetravalent streptavidin. Purified mVenus-mRNA is shown on the first lane from the left. The gel is then divided in sections of 5 lanes, each comparing the unmodified amino-ASO, then addition of mSA +/- UV, then addition of tetSA +/- UV.

To achieve a better ON/OFF ratio, one further dT in **V1** was substituted with an amino-C6-dT, for a total of four modifications. Moreover, a new set of antisense oligonucleotides targeting the same mVenus-mRNA region was designed and screened, with the purpose of improving the overall ASO potency. These ASOs partially included **V1** sequence or were a shortened version of it and they are shown in the table in figure 2.11. The sequences were initially tested in a standard RNase H-mediated degradation reaction, and they all resulted active, without being able to rank them based on their potency (Figure 2.11). Since the ASO would finally be used in an *in vitro* transcription and translation (IVTT) system, the sequences were also tested using the commercial PURExpress mixture. The DNA used was a linear PCR template containing a T7 promoter and the mVenus gene. RNase H and ASOs were added to the reaction mixture. Two different concentrations of RNase H were used to test this reaction and, not surprisingly, adding more RNase H resulted in a decrease of total protein expression (Figure 2.11).

Entry	Target	Nucleotide position on mRNA sequence	Length in number of bases	Number of amines	Sequence (5' → 3')	
5	mV	220	20	4	TXCATGXGGXCGGGGXAGCG	→ V2
6	mV	229	18	0	GTGCTGCTTCATGTGGTC	
7	mV	229	18	4	GXGCTGCXTCAXGTGGXC	→ V3
8	mV	224	16	0	TTCATGTGGTCGGGGT	
9	mV	229	17	0	TGCTGCTTCATGTGGTC	
10	mV	230	16	0	TGCTGCTTCATGTGGT	

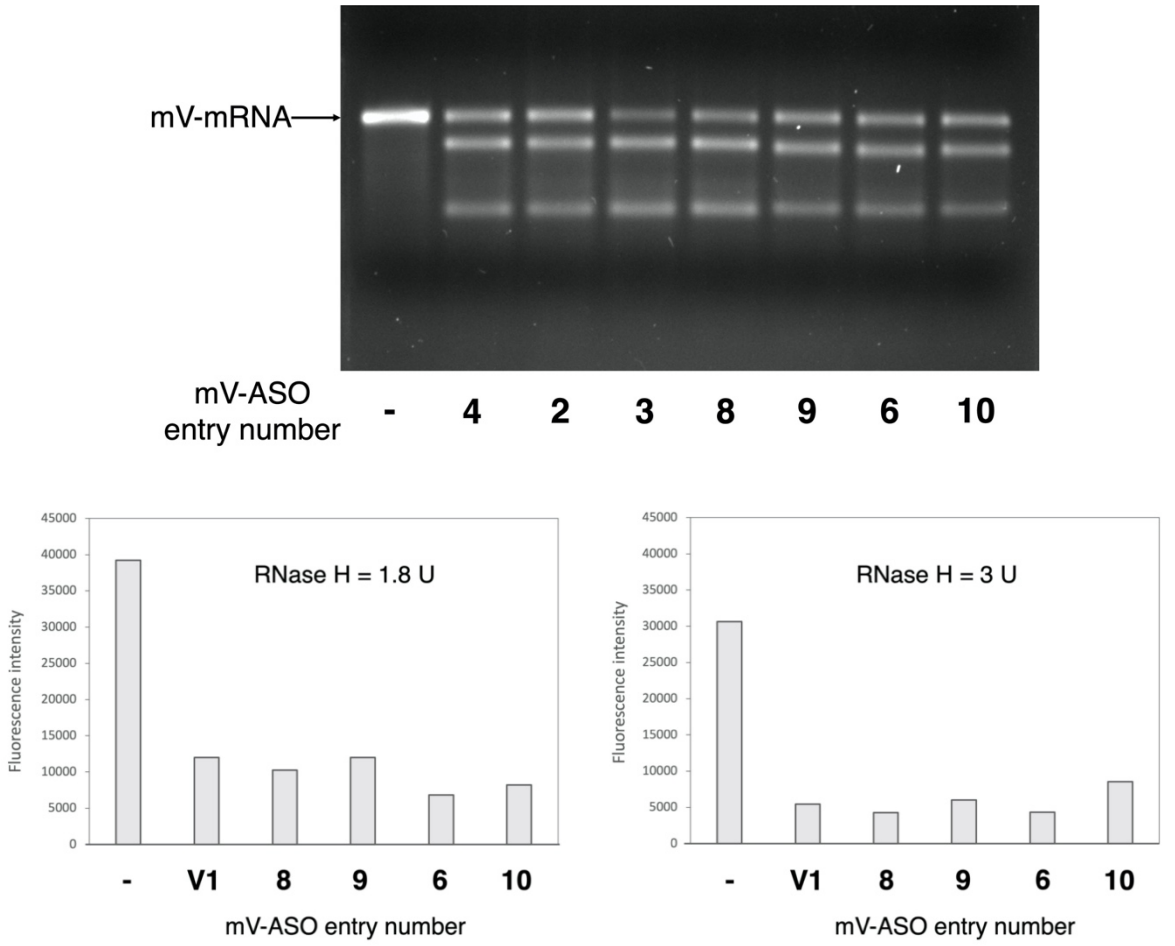


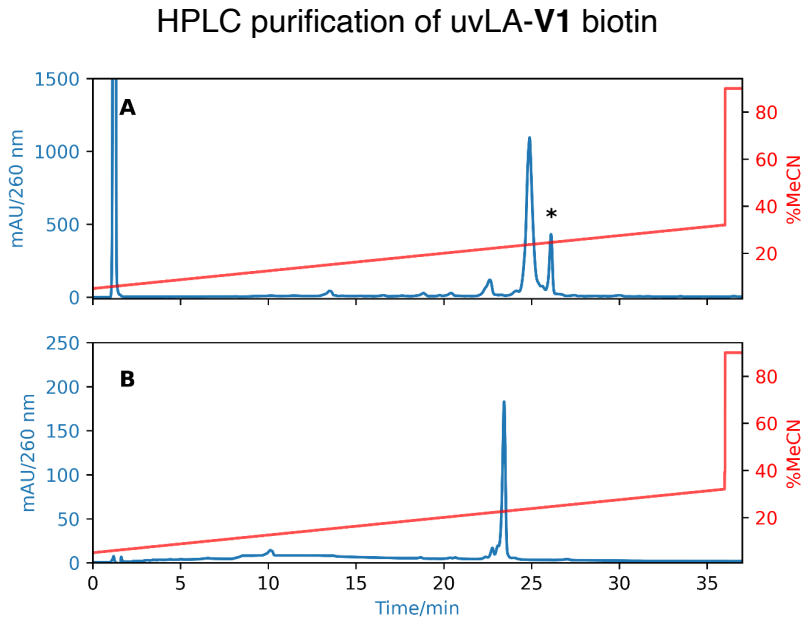
Fig. 2.11. RNase H-mediated mRNA degradation and IVTT reaction supplemented with RNase H to evaluate the potency of different ASOs

From top to bottom: sequences designed for higher potency and for the insertion of a fourth modification. Agarose gel of a RNase H-mediated mRNA degradation reaction comparing the different amino-ASOs. IVTT reaction comparing the different ASO sequences. The last graph does not contain error bars as only a single experiment per condition was carried out. As the results showed consistency across different RNase H concentrations, the experiment was not repeated further.

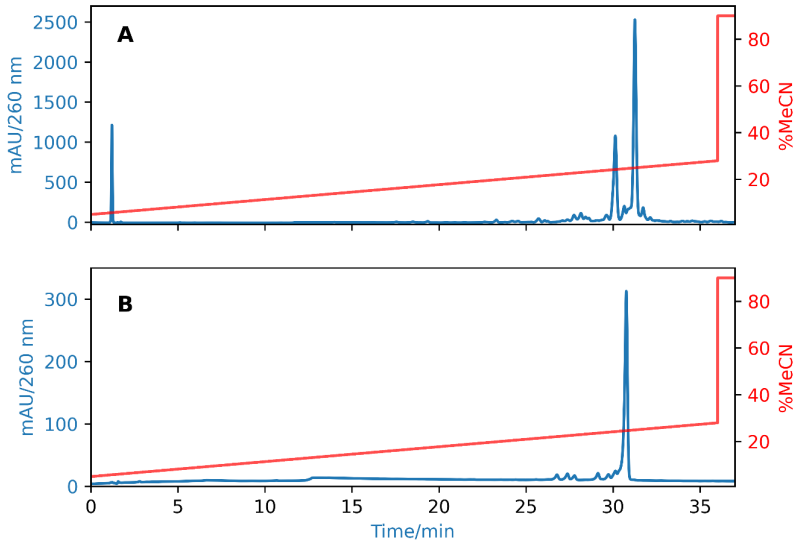
2.10 Optimization of modified-ASO sequence

Based on the results, the ASO in entry number 6 was selected and purchased with four amine modifications (**V3**). **V1**, **V2** and **V3** were modified with PCB-NHS ester and purified by HPLC using the improved method discussed above. The method uses TEAB buffer, a slower gradient, and column heating to 50 °C. The products were analyzed by mass spectrometry and PAGE before binding to streptavidin. Details on the method can be found in Chapter 5.

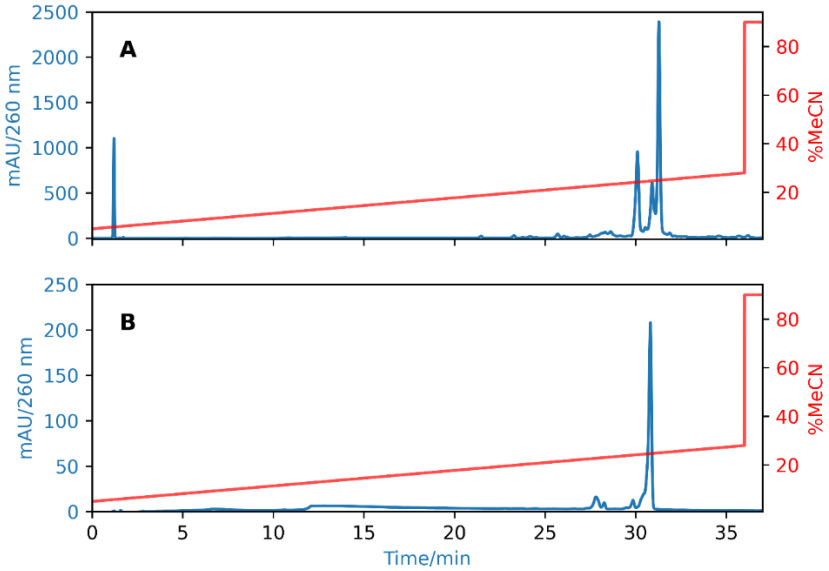
The chromatograms showed how the ASOs purification was successful, and the purified products only contained very minor impurities. The LC-MS spectra indicated in the table of figure 2.12 are shown in Appendix B.



HPLC purification of uvLA-**V2** biotin



HPLC purification of uvLA-**V3** biotin



Oligonucleotide	Number of PC-biotin groups	Purified product	Expected mass	Mass after deconvolution
V1	3	uvLA-**V1** biotin	9149	9148
V2	4	uvLA-**V2** biotin	9740	9740
V3	4	uvLA-**V3** biotin	9032	9032

Fig. 2.12. HPLC purification of uvLA-ASOs containing three or four modifications

Each HPLC chromatogram shows the crude (top) and purified product (bottom). The table corresponds to the mass spectrometry results for the purified products. The LC-MS spectra are shown in appendix B.

V2 and **V3** were also analyzed on PAGE after HPLC purification, agarose gel, binding to streptavidin, and UV irradiation. The UV irradiation used for the ASOs loaded on this gel was not powerful enough to completely cleave off all four photocleavable groups. This is visible in figure 2.13, where UV irradiation before and after streptavidin binding resulted in the product still containing a band corresponding to incomplete cleavage. This was improved in later experiments by increasing the UV power applied or the irradiation time. Binding to streptavidin was complete as no starting material band was visible on the gel. However, a band of higher size was present on the gel, due to unspecific binding. This band does not respond to UV irradiation, and it was ignored for future experiments.

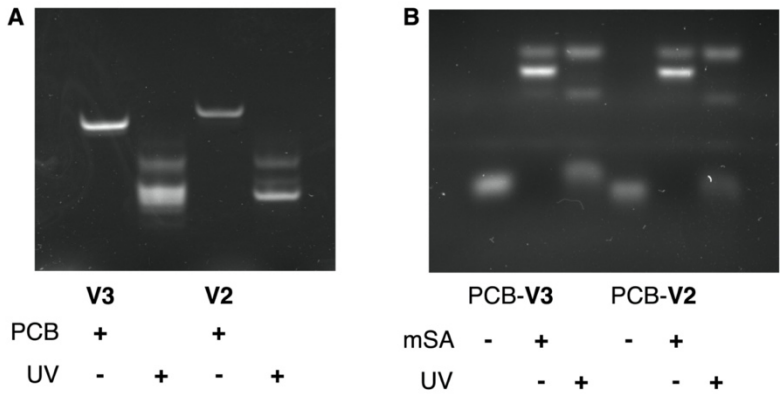


Figure 2.13. Analysis of HPLC purified uvLA-ASOs, binding to streptavidin, and photocleavage

On the left, a PAGE analysis of the HPLC purified ASOs containing four PCB modifications. After UV, the bands recovered contain fully cleaved ASOs and ASOs with only 3/4 PCB cleaved off. On the right, binding to monovalent streptavidin with subsequent UV irradiation. The band at the very top corresponds to unspecific binding of streptavidin. The band does not get photocleaved and it was ignored for following applications.

The unmodified sequences, the amino-ASOs **V1**, **V2** and **V3**, and the uvLA-ASOs were finally compared in a RNase H-mediated degradation reaction (Figure 2.14). The sequence of ASO mV229 is the unmodified sequence of **V3**. Same for unmodified mV220, and amino-modified **V1** and **V2**, which respectively contain three and four amines. In the gel, both mV229 and **V3** were more active than the mV220 sequence. Moreover, having four amines in **V2** and **V3** versus three amines in **V1** did not affect the knockdown efficiency, but improved the off state. Interestingly, the degradation products had a slightly different pattern, as mV229 cleaves slightly after mV220 along the mRNA sequence, making one of the resulting fragments smaller and the other larger.

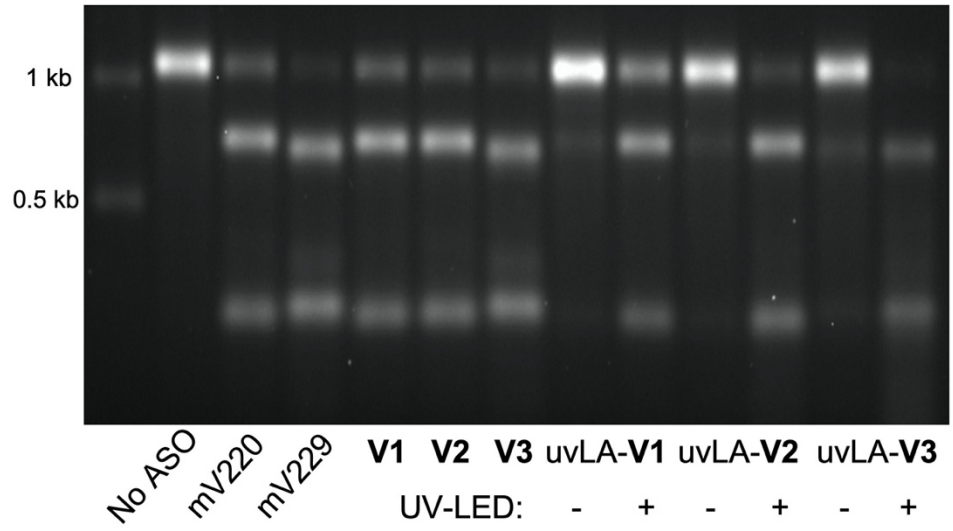


Figure 2.14. RNase H-mediated mRNA degradation assay to compare different photocaged ASO sequences

From left to right: a molecular weight marker, purified mVenus-mRNA, the unmodified ASOs mV220 and mV229, the amino-ASOs V1, V2, V3, then all the light-activated ASOs showing no mRNA degradation before UV irradiation, then full degradation when UV is applied. The gel showed how the most active sequence was V3.

2.11 Control of in vitro transcription and translation with UV

The next goal was to reverse mRNA transcription by degrading the mRNA formed in an *in vitro* transcription (IVT) reaction. This was done by incubating a T7 DNA template expressing mVenus with T7 RNA Polymerase, triphosphates, and RNase H. The buffer used for T7 RNAP was compatible with RNase H. Initially, the condition for the plateauing of mVenus-mRNA expression were determined (Figure 2.15). After three hours, the amount of mRNA produced was almost saturated. At this point, the amino-ASO **V3** was added to completely degrade the mRNA formed.

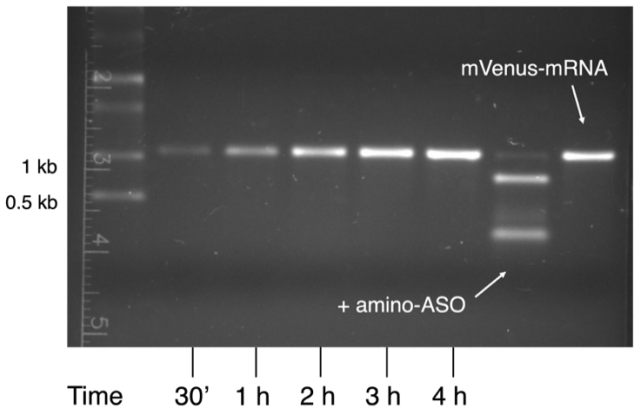


Figure 2.15. In vitro transcription reaction showing mRNA formation and following degradation upon addition of amino-ASO

The agarose gel shows the formation and degradation of mVenus-mRNA. Upon addition of the DNA template in the IVT mixture, mRNA formation is measured by analyzing an aliquot at 30 min, 1 h, 2 h, 3 h, and 4 h. The difference in the amount of mRNA between 3 and 4 hrs is negligible. The active amino-ASO V3 is also added at 3 h and incubated for 1 h. The result is complete mRNA degradation. The last lane on the right corresponds to the purified mVenus-mRNA and it was used as molecular weight reference.

Since the exact amount of mRNA produced then degraded was practically unknown, a screening of RNase H-mediated degradation reaction conditions was carried out to compare the off state of **V2** and **V3** at different concentrations at a fixed amount of RNase H (Figure 2.16).

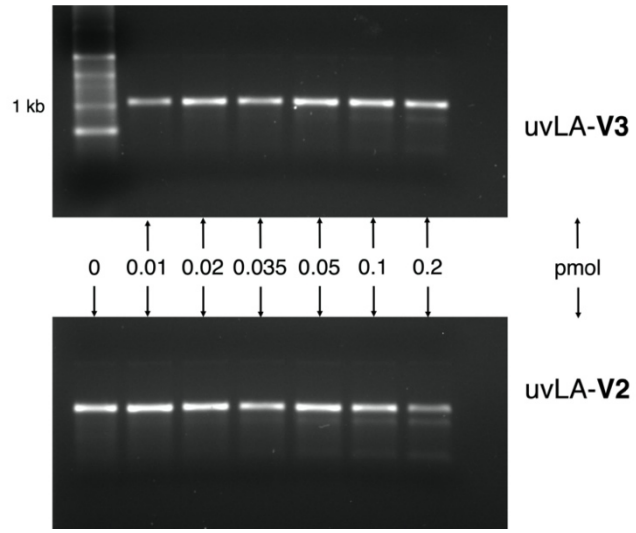


Figure 2.16. Comparison of off states between uvLA-V2 and uvLA-V3 at different concentrations

Different concentrations of modified-ASO were added to a RNase H-mediated mRNA degradation assay to determine the maximum amount of LA-ASO that could possibly be added without causing any mRNA degradation. The experiment was used to test the off state of both uvLA-V3 and uvLA-V2.

As 0.035 pmol of uvLA-**V3** resulted in the highest amount we could possibly add maintaining the off state, the same reaction conditions were repeated to evaluate the on state efficiency. In an IVT reaction in the presence of RNase H and 0.035 pmol of uvLA-**V3**, when no UV was applied for three hours, the mRNA remained intact, as expected. At the three-hour time point, UV was applied to the reaction tube, and after letting the reaction incubate for another hour, it was analyzed by agarose gel. As a result, the ASO did activate and degrade mVenus-mRNA at levels comparable to the positive control **V3** at the same concentration (Figure 2.17).

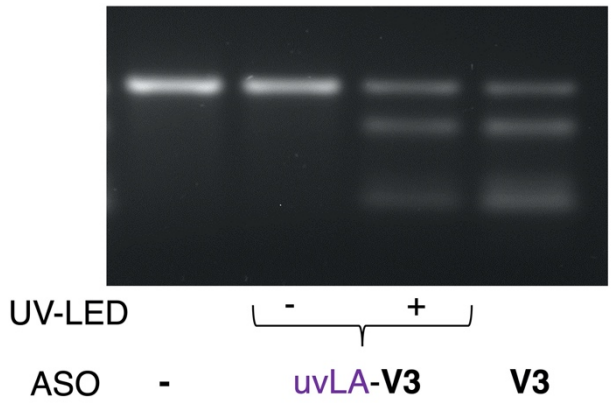


Figure 2.17. Control of in vitro transcription with UV

An IVT reaction in the presence of RNase H and uvLA-V3 was incubated for three hours. No mRNA degradation was observed. After three hours, UV was applied and the reaction incubated for another hour. At the same time, in a separate reaction tube not containing any ASO, the amino-ASO V3 (positive control) was added. The agarose gel result showed the same level of mRNA degradation caused by the activated uvLA-V3 and the amino-ASO.

Finally, to test the system with *in vitro* transcription and translation systems, **V3** or uvLA-**V3**, and the mV DNA template were added to the CFPS system (supplemented with RNase H). The amount of intact mVenus expressed after four hours was measured by fluorescence (Figure 2.18). In the absence of ASOs, UV irradiation only caused a non-significant decrease (p-value = 0.23) in mVenus expression, confirming little-to-no UV-damage. The amino-ASO **V3** decreased mV synthesis by 89% compared to the no ASO control, with no significant change after UV irradiation (p-value = 0.265). Finally, the off state of uvLA-**V3** did not affect protein synthesis significantly (~15% decrease, p-value = 0.084), whereas it recovered 84% of **V3** knockdown efficiency after five minutes UV irradiation, reducing protein synthesis by ~75% when compared to the no ASO control. This

demonstrated that it is possible to control gene knockdown in IVTT using light-activated ASOs.

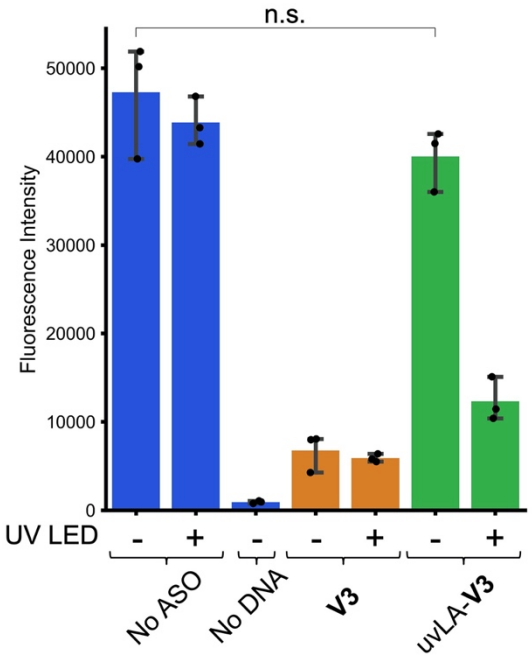


Fig. 2.18. Control of in vitro transcription and translation with UV

Fluorescence output of mVenus, corresponding to its expression levels. RNase H was added to the IVTT reaction mixture. On the far left, with no ASO present, the UV irradiation caused a very slight decrease in mVenus expression. The same when the amino-ASO V3 was present as positive control. In the presence of uvLA-V3, the off state was comparable to no ASO present, while the on state was comparable to the positive control.

2.12 Two-wavelength control of cell-free protein synthesis

Our group recently reported the synthesis of a blue light-activated photocage, to be attached to the T7 promoter of a DNA template to control its expression with 455 nm wavelength

[112]. The molecule was completely orthogonal to the UV-activated 2-nitrobenzyl photocage used in this work. This allowed for the orthogonal activation of DNA transcription and following mRNA degradation. The same molecule was used in this work to modify a mVenus DNA template with a T7 promoter and an antisense oligonucleotide. The structure of the molecule and overall scheme of the dual-wavelength work is shown in figure 2.19.

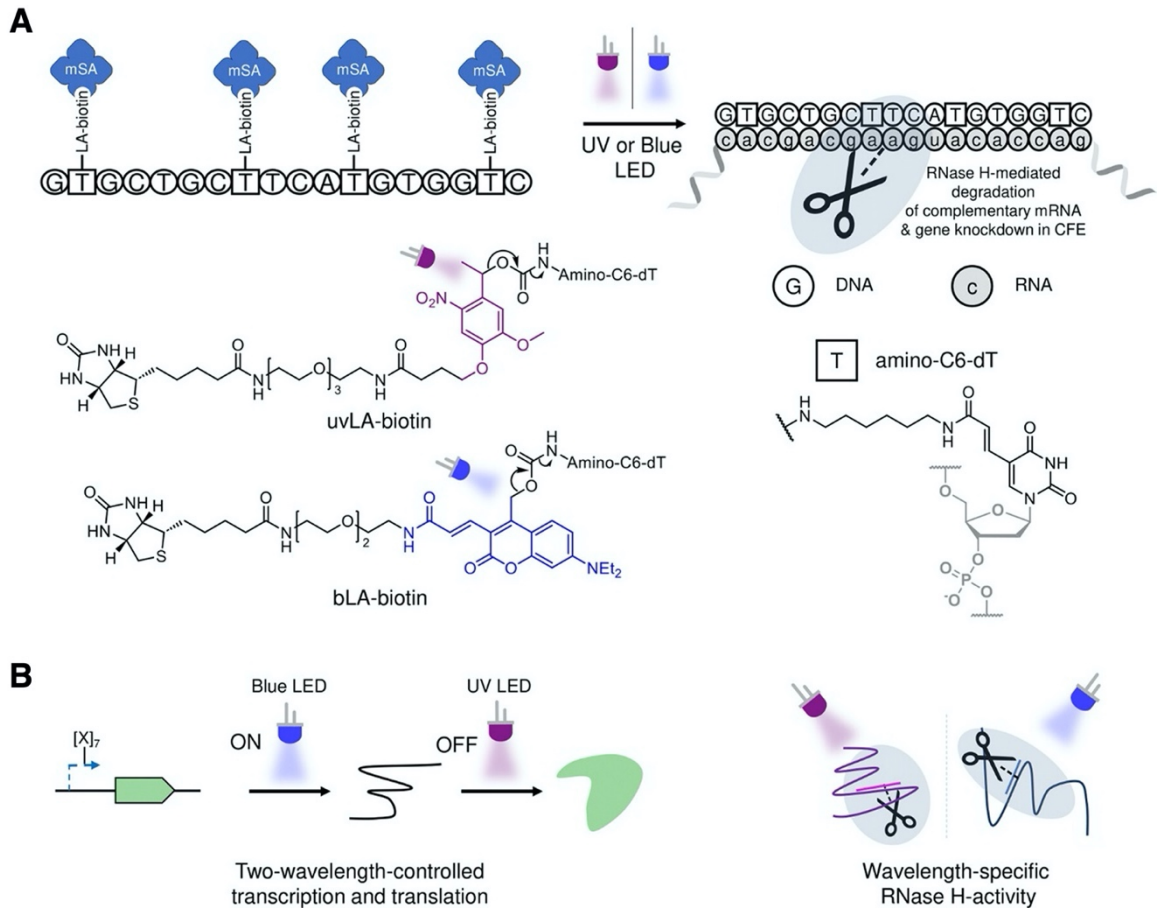


Figure 2.19. UV and blue light photocage structures and general dual-wavelength activation scheme

A) The structure for the UV (2-nitrobenzyl) and blue (coumarin) light-activated photocages, and the general scheme for the modification and activation of V3 are shown. B) General scheme of the dual-wavelength activation work. A DNA template was modified with blue light photocage and illuminated to turn ON transcription. An ASO was modified with the UV light photocage and irradiated to turn OFF translation.

To achieve control of the activation and deactivation of *in vitro* transcription, the blue LA (bLA)-DNA template was prepared and used in combination with uvLA-**V3**. In the first testing experiments, two different amounts of uvLA-**V3** were added: 0.035 pmol to achieve a perfect off state and good but not complete ON, or 0.05 pmol to have a great on state, but a slightly leaky OFF. The analysis corresponded exactly to the controls without

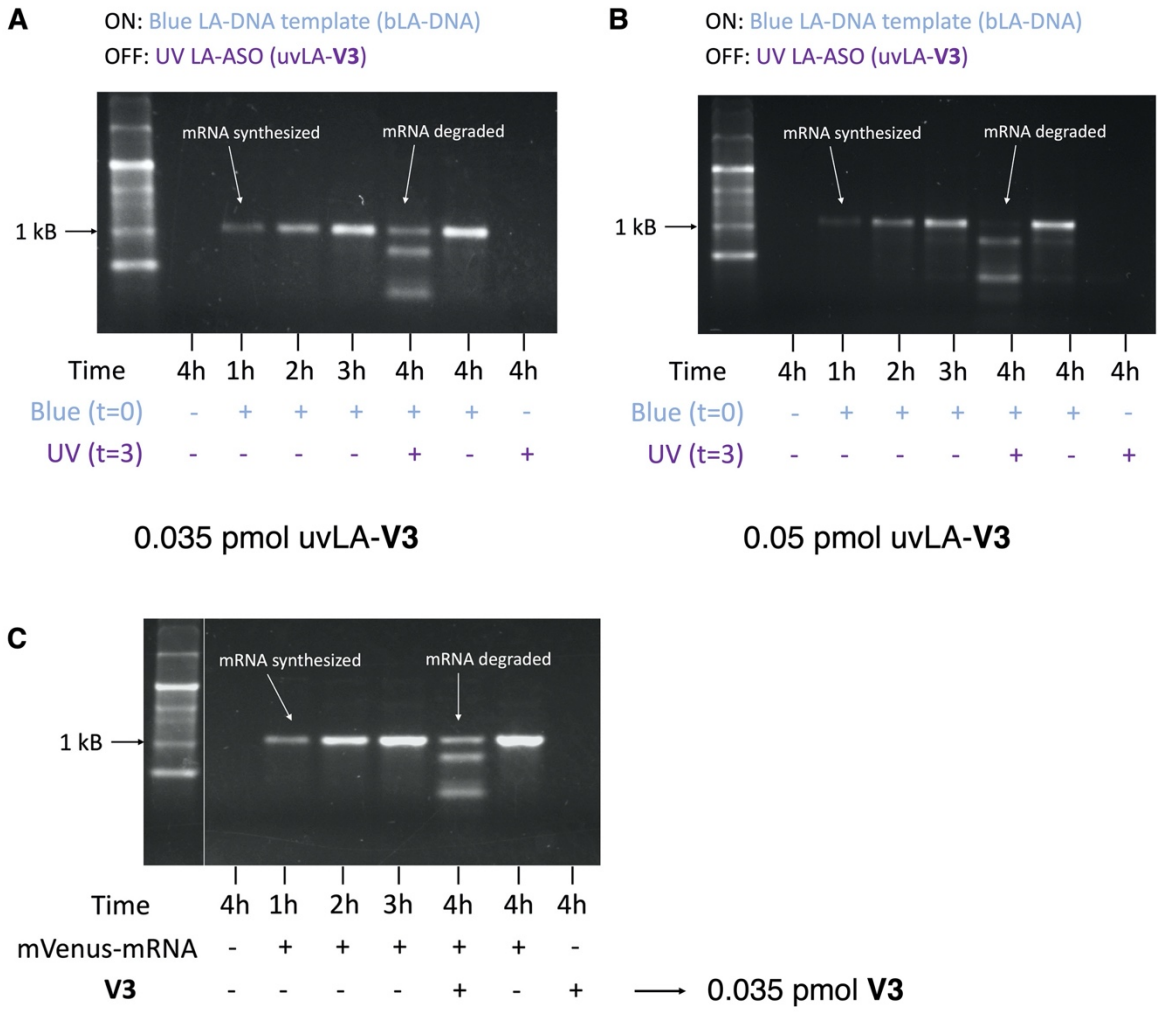


Fig. 2.20. Orthogonal control of IVT activation and mRNA degradation
Gene expression was initiated by irradiation with blue light. After three hours UV light was shone and the mRNA product degraded. A) and B) compare two different ASO concentrations. 0.035 pmol of ASO result in a very tight off state, whereas 0.05 pmol result in a very active on state, but slightly leaky off state. C) shows the control experiment run in parallel. The transcription was initiated by adding mRNA to the tube, and degradation by adding the amino-ASO control.

photocages, where T7 mVenus (mV) DNA template was added at the starting of the reaction, and after three hours the amino-ASO **V3**, was added to achieve a similar mRNA degradation result (Figure 2.20).

As the system worked very well in IVT, it was also tested in CFPS. The bLA-mV DNA template was incubated for four hours at 37 °C in the presence of uvLA-**V3** and RNase H under different illumination conditions. Blue light alone resulted in an equal protein output compared to the control with no ASO (p-value = 0.196). This was due to the orthogonality of the two photocages, with uvLA-**V3** not activating under blue light. Irradiation with UV light alone did not activate the bLA-mV DNA template, therefore no protein was synthesized, as only the ASO was activated. Finally, blue light was applied first to activate transcription and translation, then UV light was applied at different timepoints to halt protein synthesis at different expression levels upon demand, by degrading the mRNA already generated. When the reaction was irradiated with both UV and blue light before incubation at 37 °C, a high degree of gene knockdown (74%) was observed after four hours, as any mRNA formed would be quickly degraded by RNase H using the uncaged ASO. Secondly, when blue light was applied first, then UV after thirty minutes incubation, protein synthesis had already initiated but was halted at 51%, compared to when the ASO was absent. This was in line with previous data, showing how CFPS systems produce a substantial amount of mRNA and protein in the first stages of the reaction. Lastly, after activating protein synthesis with blue light and incubating for three hours before shining UV, the mVenus protein synthesis output had virtually no reduction (p-value = 0.519), as its production already reached a plateau. This was expected, as the CFPS reagents used usually resulted in a plateau around three to four hours. This demonstrated a two wavelength ON and OFF switch for CFPS, with the ability to temporally activate and halt protein synthesis (Figure 2.21).

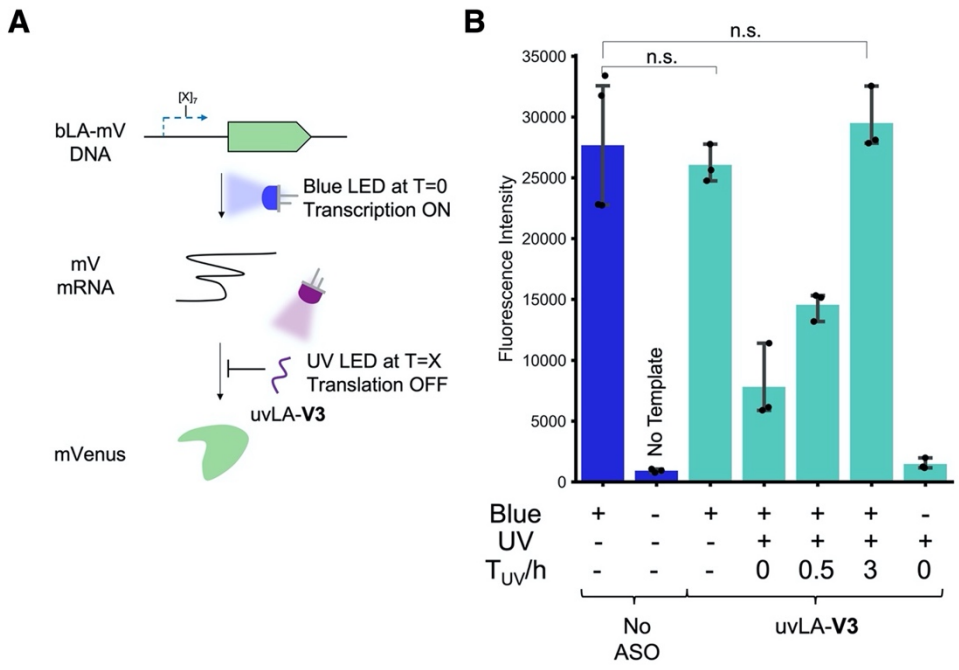


Figure 2.21. Orthogonal control of IVTT activation and halting of protein synthesis at different expression levels

A) General scheme showing activation of transcription with blue light first, by activating the coumarin-modified bLA-mV DNA template. Then the orthogonal deactivation of translation at different time points by uncaging the uvLA-V3 ASO with UV. B) Cell-free protein expression halted at time = 0, 0.5 h, 3 h. The orthogonality of the system was shown first when irradiated with blue light only, showing the protein expression level equal to when the ASO was absent. Moreover, when the system was irradiated with UV only, transcription was not activated.

2.13 Two-wavelength control of gene knockdown

With two orthogonal photocages that could both be attached to amino-C6-dT modifications, the next goal was to control the transcription and translation of two different mRNAs/proteins with sequence specific ASOs activated by orthogonal wavelengths. The second target selected was mCherry, a red fluorescent protein derived by a different organism than mVenus (dsRed vs. avGFP), and with a very different gene sequence. These

two proteins would allow for simultaneous fluorescent measurements, as they have different excitation/emission spectra.

The aim was initially to identify a new ASO sequence that could target the mRNA encoding mCherry (mC), without binding mV-mRNA. The ASO screening followed the same criteria as for mVenus, although ultimately the mC-ASO would have to be screened against mVenus-mRNA to confirm it would not cause any degradation, and vice versa mV-ASO would have to be screened against mCherry-mRNA. All the mC-ASOs screened are shown in the table below (Figure 2.22).

Entry	Target	Nucleotide position on mRNA sequence	Length in number of bases	Number of amines	Sequence (5' → 3')	
11	mC	27	19	0	CTCCTTGATGATGGCCATG	
12	mC	81	19	0	CTCGAACTCGTGGCCGTTC	
13	mC	139	18	0	CTTCAGCTTGGCGGTCTG	
14	mC	147	19	0	CTTGGTCACCTTCAGCTTG	
15	mC	222	16	0	CTTCACGTAGGCCTTG	
16	mC	303	16	0	GTCCTCGAAGTTCATC	
17	mC	141	16	0	CTTCAGCTTGGCGGTC	
18	mC	139	16	0	TCAGCTTGGCGGTCTG	
19	mC	139	16	4	XCAGCTXGGCGGXCXG	→ C1
20	mC	140	15	0	TCAGCTTGGCGGTCT	
21	mC	135	22	0	CTTCAGCTTGGCGGTCTGGGTG	
22	mC	139	23	0	GTCACCTTCAGCTTGGCGGTCTG	
23	mC	141	21	0	GTCACCTTCAGCTTGGCGGTC	
24	mC	142	20	0	GTCACCTTCAGCTTGGCGGT	

Figure 2.22. Table showing all the mC-ASO sequences screened

The table shows all the sequences screened. Entries 11-16 were tested in a first batch, then entries 17-24 following initial results. As for mV-ASOs, the nucleotide position indicated refers to the binding of the first base of the ASO from its 3'-end on the mRNA starting from position 1 of the 5'-end of mRNA.

The screening was done by RNase H-mediated mRNA degradation, and it identified a good target area along the mC-mRNA sequence, with ASO sequences containing multiple Ts across the length of the ASO (Figure 2.23). Amongst different sequences tested, the only one showing promising result was ASO entry **13**. By changing the sequence length and

binding position around that mC-mRNA target area, the sequences **17**-**24** were designed and tested against both mC- and mV-mRNA (Figure 2.24).

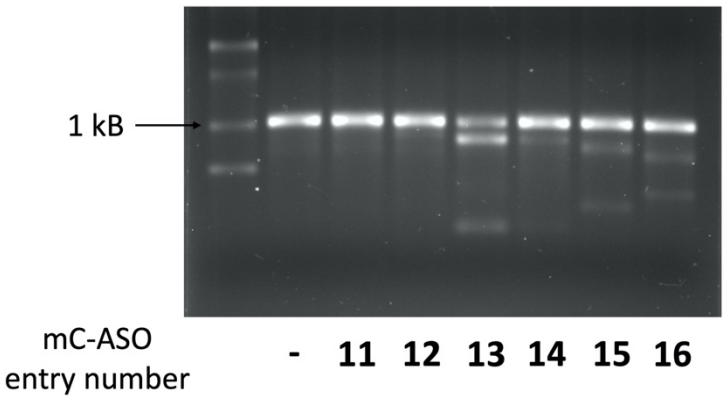


Figure 2.23. First screening of mCherry ASO
The agarose gel shows an RNase H-mediated mC-mRNA degradation reaction, with the first lane from the left being the positive control of intact mRNA. A molecular weight marker is also shown. The most active ASO causing the most degradation was entry 13.

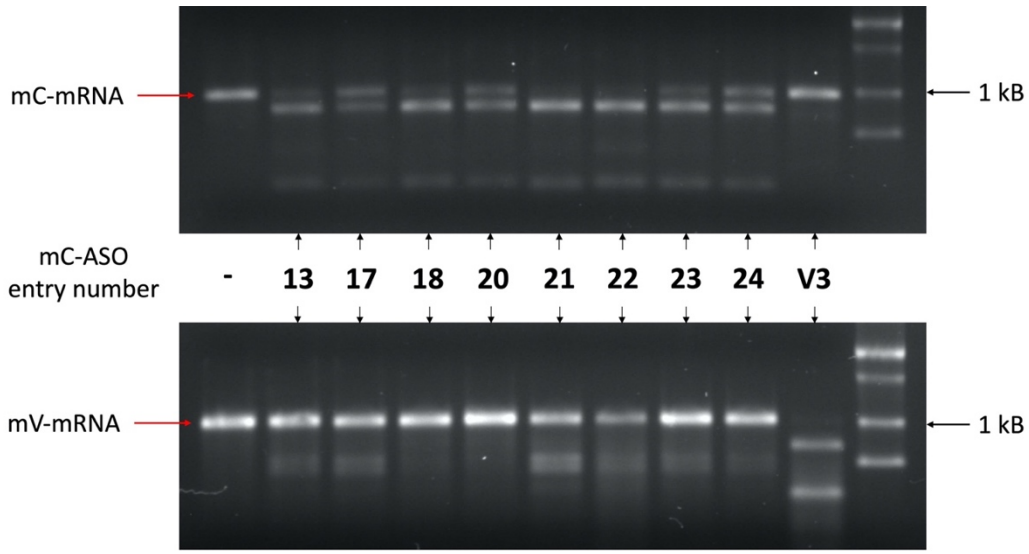


Figure 2.24. Second screening of mCherry ASO for orthogonality with mVenus
The agarose gel shows an RNase H-mediated mRNA degradation reaction with ASO designed to target mCherry against mC-mRNA and mV-mRNA to test their orthogonality. On the left, the purified mRNAs (negative control), on the right a molecular weight ladder. V3 was used as positive control for mV degradation.

The sequence causing the highest mC-mRNA degradation and lowest mV-mRNA degradation was ASO entry **18**, which was finally purchased with four amino-C6-dT bases (**C1**). Other experiments showed how **V3** was causing more crosstalk, degrading more mC-mRNA than **V2**, which would have been an issue when moving to IVTT reactions. Therefore, for this work **V2** was the ASO of choice for targeting mVenus. The orthogonality between the amino-ASOs **V2** and **C1** was tested (Figure 2.25) and with this, **C1** was modified with the blue light photocage, purified by HPLC, and analyzed by mass spectrometry (data shown in Appendix B). The orthogonality between the two wavelengths of activation was validated with the UV-vis spectrum in figure 2.26.

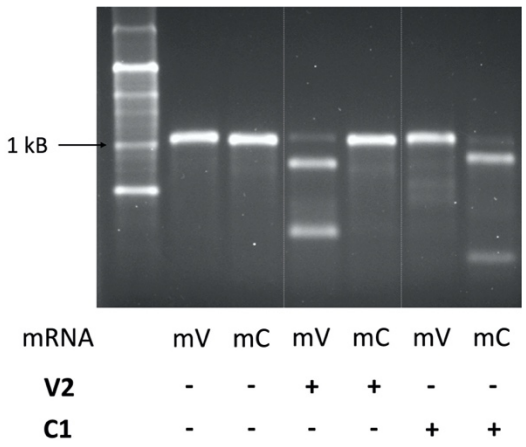


Figure 2.25. Orthogonality between V2 and C1 in RNase H-mediated mRNA degradation

The amino-ASOs V2 and C1 target mVenus and mCherry respectively. In this agarose gel showing a RNase H-mediated mRNA degradation reaction, both ASOs were tested against both mV- and mC-mRNA to evaluate the orthogonality of the system. From left to right: a molecular weight marker; mV- and mC-mRNA without any ASO present; both mRNAs with the addition of only V2, showing degradation of mV-mRNA and mC-mRNA remaining intact; both mRNAs with the addition of only C1, showing degradation of mC-mRNA and mV-mRNA remaining intact.

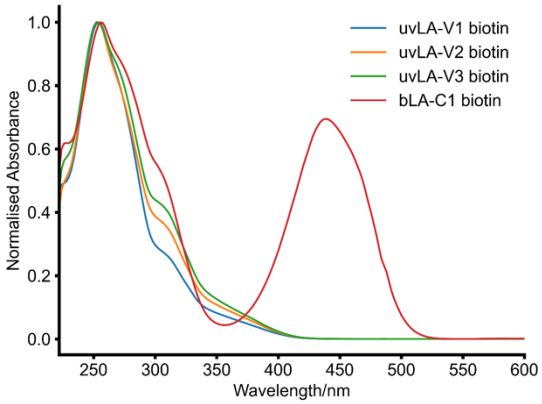


Figure 2.26. Orthogonality between V2 and C1 in UV-vis spectrum
A UV-vis spectrum ranging from 230 to 600 nm of all the LA-ASOs prepared (before binding to streptavidin). The UV responsive ASOs and the blue light responsive ASO show no overlap and therefore perfect orthogonality.

Monovalent streptavidin was finally bound to both ASOs forming uvLA-**V2** and bLA-**C1**. The two photocaged ASOs were incubated with mV-mRNA, mC-mRNA, and RNase H to test light-controlled mRNA degradation. In the absence of irradiation, the two ASOs maintained their off state and both mRNAs remained intact. Irradiation with blue light alone resulted in only mC-mRNA degrading. Vice versa, when UV light was applied, mV-mRNA was the only mRNA to be degraded by RNase H (Figure 2.27).

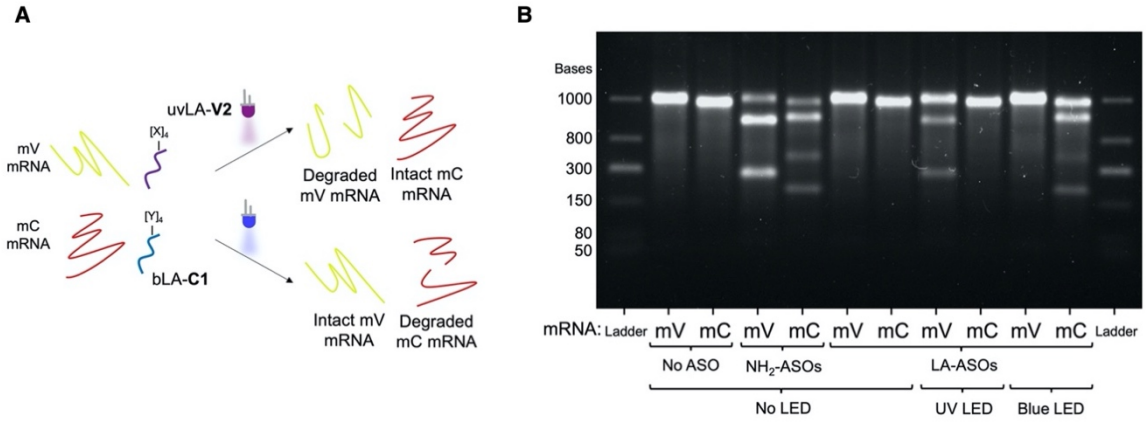


Figure 2.27. Orthogonality between light-activated ASOs in RNase H-mediated mRNA degradation reaction

A) General scheme showing the selective degradation of mV-mRNA with UV and mC-mRNA with blue light. B) RNase H-mediated mRNA degradation reaction showing the orthogonality between the UV- and blue-activated ASOs.

Finally, the orthogonal LA-ASOs were used to control the cell-free protein synthesis of the two different proteins in IVTT. The DNA templates of mV and mC, bLA-**C1**, uvLA-**V2**, were added to a CFPS system supplemented with RNase H. The mixtures were irradiated (or not irradiated) with UV or blue light, before being incubated for four hours. Their final fluorescence was then measured. Without any illumination, the expression level of both proteins corresponded to the one in the absence of ASOs (17% increase for mC and 23% decrease for mV, both non-significant). Upon UV irradiation, uvLA-**V2** uncaged and selectively knocked down mV fluorescence by 56%, only causing 16% of unwanted mC degradation (non-significant, p-value = 0.129). Contrarily, blue light uncaged bLA-**C1**, which selectively knocked down mC by 62%, with an unwanted 30% reduction observed for mV fluorescence (significant against positive control, p-value = 0.032). This demonstrated we were able to control gene knockdown in CFPS using two orthogonally light-controlled ASOs (Figure 2.28).

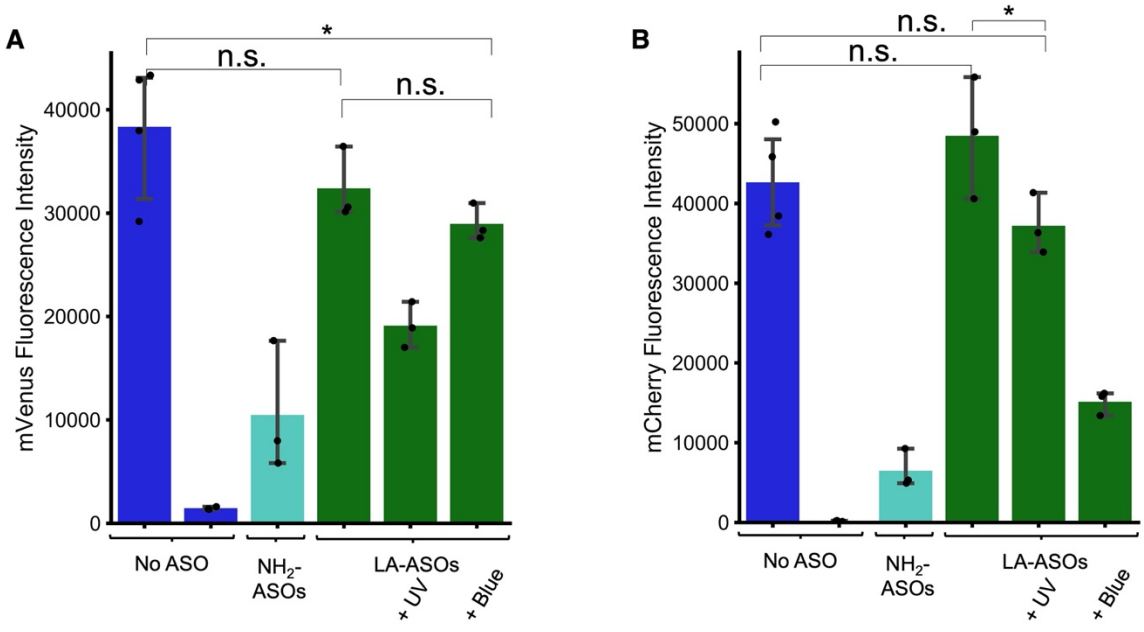


Figure 2.28. Dual-light activated orthogonal control of gene knockdown in CFPS
The two graphs represent measurements taken simultaneously at two different excitation and emission wavelengths. A) mVenus is knocked down by UV activation and not with blue light. B) mCherry is knockdown with blue light and not with UV.

2.14 Conclusions

The role of antisense oligonucleotides in cell-free protein synthesis systems is underexplored, despite their potential in synthetic biology. In cell-free mixtures, ASOs can degrade a target mRNA relatively quickly and specifically in the presence of RNase H, preventing the expression of the corresponding protein. This process can be controlled with light in a spatiotemporal manner, by modifying the ASOs with photocages. This represents one of the only available tools to control gene knockdown remotely in a cell-free system.

Several ASO sequences were tested for the knockdown of fluorescent protein genes mVenus and mCherry. The most active sequences were selected and modified with UV or blue light sensitive photocages. This was done by reacting NHS esters containing the photocleavable group and a terminal biotin with commercially available amino-modified ASOs. Streptavidin was then bound to the biotinylated ASOs to form the photocage.

The amino-modifications did not cause any decrease in ASO activity, while the photocages blocked RNase H from approaching the ASO by steric hinderance, thus preventing the mRNA from being degraded and allowing protein synthesis to occur normally. Once irradiated with light, the RNase H activity was restored, the mRNA was degraded, and protein synthesis was halted.

As the UV and blue light photocages used were completely orthogonal to each other, they were combined in two different cell-free setups. The first, by using a previously developed blue light-activated DNA template to activate transcription with one wavelength, then halt translation with the orthogonal wavelength using a UV-activated ASO. The second, by generating two orthogonally active ASOs targeting two different mRNAs, for selective two-wavelength gene knockdown in cell-free protein synthesis.

As all the previously reported work on controlling cell-free protein synthesis only focuses on the remote activation of gene expression, with this approach, a method to also turn gene expression OFF with light on demand was developed.

3

Development of light-activated and self-delivered siRNAs for the control of gene knockdown in mammalian cells

3.1 Historical remarks

Due to the importance of their findings, after only about eight years from their first publication on RNA interference, Andrew Fire and Craig Mello were awarded the Nobel prize in Physiology or Medicine in 2006. Their discovery clarified many contradictory

experimental observations of the time and revealed a mechanism for controlling gene expression. The first visualization of this phenomenon occurred years before, in 1990, when by inserting transgenes to overexpress chalcone synthase, the enzyme responsible for flower pigmentation, instead of obtaining a brighter coloured petunia, Richard Jorgensen et al. obtained interesting decoloured patterns on the petals (Figure 3.1) [113].

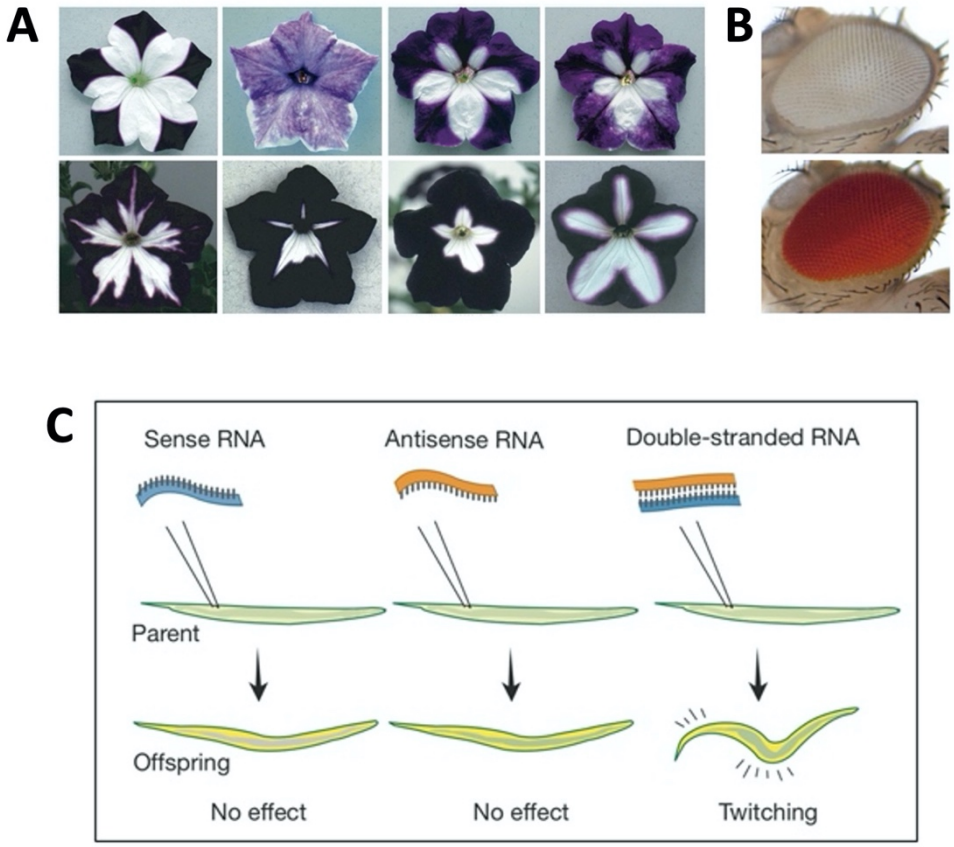


Figure 3.1: RNA interference in plants, insects, and worms
Interesting phenotypes in different organisms resulting from RNA interference. A) Decolored patterns in petunia (ref. [69]). B) Wild-type *Drosophila* expresses the red pigmentation gene, which loses pigmentation once the gene is silenced. C) Nobel prize RNAi work on *Caenorhabditis elegans*.

Later, David Baulcombe et al. described RNA as the trigger of this effect, after observing similar patterns in *Nicotiana benthamiana* as a response to plant viruses [114]. Then finally, Andrew Fire and Craig Mello assigned this gene silencing effect to dsRNA, after being able to show control over *Caenorhabditis elegans* behaviour by targeting different genes with different dsRNA sequences [115]. The dsRNA had to be complementary to the mature mRNA sequence (intron or promoter sequences in pre-mRNA did not trigger the response), and sense or antisense single strands alone did not affect expression. The term RNA interference (RNAi) was coined and the mechanism of gene silencing elucidated soon after. Although some exceptions in some fungi and protists exist [116], this mechanism can be found in most eukaryotic organisms and used for therapeutic purposes, which is why it gained interest so quickly and became a fundamental tool in medicine and research.

3.2 RNAi mechanism

Interestingly, the first paper showing RNAi was most likely published in 1928, when the upper leaves of a tobacco plant were observed to have become immune to a virus that was infecting the rest of the plant [117]. In fact, RNAi is most likely an ancient defence mechanism against genetic material from viruses, as there are a few different RNAi mechanisms all conserved in animals, plants, and fungi, suggesting a common ancestor. Prokaryotes also have an RNAi-like defence system, which is functionally analogous but not homologous to eukaryotic RNAi [118].

RNA interference consists of three major pathways mediated by three different forms of RNA: small interfering (si)RNAs, which are short dsRNAs and will be the main focus of this work; microRNAs (miRNAs), also small, non-protein-coding RNAs; Piwi-interacting (pi)RNAs, longer RNAs that do not require the proteins required by siRNAs and miRNAs to be produced. In the endogenous eukaryotic and most relevant pathway, the viral information is transcribed by RNA Polymerase II into a primary micro RNA (pri-miRNA) transcript. Pri-miRNA is then processed further into a shorter small hairpin RNA (shRNA) of ~30 bp by the ribonuclease Drosha, then exported to the cytoplasm by the karyopherin Exportin-5 [36, 37].

In the cytoplasm, the pre-miRNAs or other shRNAs formed are recognized by the endoribonuclease Dicer due to their loops, and Dicer cleaves them into shorter sequences. Dicer can also bind exogenous synthetic small interfering RNAs, which are dsRNA sequences typically ~19-21 nucleotide long and containing two-nucleotide 3' overhangs on each strand (Figure 3.2). Dicer induces the formation of a large RNA-induced silencing complex (RISC), which binds the short dsRNA sequences, loads the antisense strand onto an Argonaute protein (Ago), while discarding the sense strand. Because of this, the antisense strand is called the "guide" strand, while the sense strand is called the "passenger" strand. Ago is responsible for recruiting mRNA, finding the target sequence complementary to the guide strand, and promoting mRNA degradation, thus inhibiting the translation of mRNA [38, 39]. Ago activity ATP-independent and catalytic, since one RNA-Ago complex can degrade multiple mRNAs [123]. The endogenous post-transcriptional control can also follow other minor mechanisms without Dicer [124].

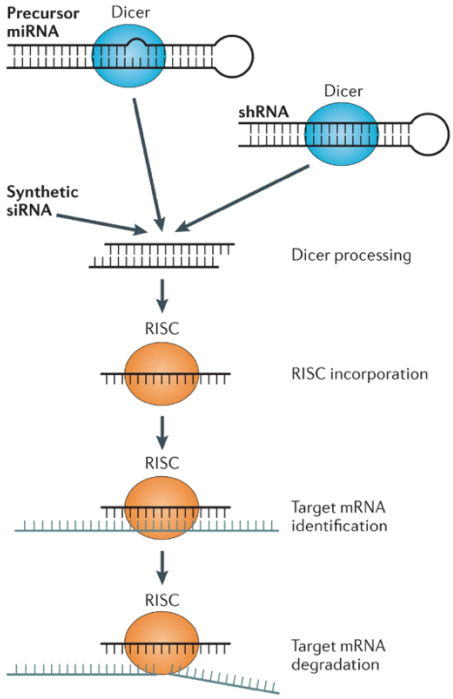


Figure 3.2: Simplified RNA interference mechanism
A simplified version of the RNAi mechanism, which can occur from naturally occurring RNAs or synthetic siRNAs (from ref. [21]).

3.3 siRNA therapeutics

Since the RNAi mechanism is present in human cells, it presents great potential for therapeutic purposes. The first demonstration of this potential was shown with *Fas* siRNA injected into mice to protect them from autoimmune hepatitis [125]. Fas is a receptor which mediates cell apoptosis and it is implicated in a number of liver diseases. By injecting *Fas* siRNA into a vein in the mice tail, the expression of Fas was reduced and the mice protected from liver failure.

Generally, siRNA-based therapeutics have the same stability, circulation, specificity, and delivery issues, which are overcome with technologies such as lipid nanoparticle carriers or common chemical modifications [28]. RNAi takes place in the cytoplasm; therefore, siRNAs do not have to be enter the nucleus, which facilitates their delivery. RNAi is also catalytic, which results in high knockdown efficiency with low amounts of siRNA. Moreover, with siRNAs being transported to the target site by the RISC, a lower concentration of siRNA is needed compared to ASOs, which instead have to find base complementarity with the target mRNA before RNase H can intervene. However, ASOs are easier and less expensive to synthesize than siRNAs, as RNA possesses the extra 2'-OH which requires extra protection and deprotection steps during solid phase synthesis, reducing coupling efficiency and leading to side products. Being double stranded, siRNAs also keep all the aromatic nucleobases on the inside, leaving only the hydrated surface on the outside, which does not interact with cell surfaces and is rapidly excreted, whereas ASOs can interact with serum proteins and cell surfaces, lasting longer in the organism [126].

As for ASOs therapeutics, the siRNA-based drugs approved have been few and all approximately in the last five years. Some examples are siRNAs fully alkylated at the 2'-OH position – such as the first siRNA-based therapeutic approved patisiran discussed in Chapter 1 – or siRNAs covalently bound to tri-GalNAc asialoreceptor binding motifs – such as lumasiran and givosiran – [127]. To date, the last siRNA drug accepted was vutrisiran in June 2022. Vutisiran has a similar therapeutic effect to patisiran and is modified with tri-GalNAc [128].

Not a therapeutic, but a last interesting application is the engineering of certain plants to produce siRNA for self-protection against insect pests [129].

3.4 Controlling siRNA with light

RNAi has been controlled in mammalian cells by using light-activated siRNAs. One of the methods to modify siRNA has been to add photocages to both the 5'- and 3'-terminal phosphates on both strands [130]. The bulkiness of the photocages seems to have an important effect on the off state and the authors were able to show patterning with mammalian tissue culture, by placing a photomask over the cells and irradiating from top. This resulted in a knockdown in gene expression only where the UV light could pass through.

Another interesting application has been the modification of the passenger strand with an azobenzene moiety [131]. Here, the azobenzene was covalently attached to two parts of the strand, bridging them together. This had a negative impact on base paring, but it achieved reversible control of gene knockdown.

All these applications have a similar mechanism and they almost uniquely use small molecule photocages to control siRNA, with the advantages and disadvantages discussed in the previous chapters. However, all these technologies require traditional methods for transfection, such as lipid formulations. This hinders the use of light-activated (LA)-siRNAs in tissue culture and in *in vivo* settings, and novel systems where light activation is coupled with the ability of self-delivery or targeted cell uptake could advance the field.

3.5 Cell-penetrating peptides

A way to aid cell delivery, while avoiding the use of often toxic lipid vehicles, nanoparticles, injections, or viral-based vectors is the use of cell-penetrating peptides (CPPs). These are typically short peptides (usually shorter than thirty amino acids) that have been studied for the past thirty years for their ability of translocating themselves and cargos of different kinds across the cell membrane. The first CPP to be discovered in 1988 was the short positively charged domain of the HIV-transactivator of transcription (TAT) protein. This cationic peptide of eleven amino acids sequence allows the whole protein of ~100 amino acids to cross mammalian cell membranes and enter the cell nucleus. Hundreds of other CPPs, mainly cationic, amphipathic, or hydrophobic have been described over the years. While for most of them the mechanism of cellular uptake is not yet entirely understood, it generally follows common endocytosis pathways or a direct translocation pathway, depending on the CPP sequence, cell type, and conditions.

One of these CPPs was used to show successful transport of streptavidin across the cell membrane [132]. In this work, the CPPs used were polydisulfides covalently bound to part of the lysine residues of each streptavidin monomer. The protein biotin-binding pockets were used to visualize internalization in the cell, by the use of biotinylated fluorophores or by detection of binding to mitochondria through protein labelling tools. In the paper, delivery of a peptide nucleic acid (PNA) was also shown, although the release was triggered by treating the cells with an excess of free biotin and the PNA had no function in the cells other than being fluorescently labelled. Using the same technology, a synthetic metal catalyst was delivered to cells and the conversion of a substrate using that catalyst was shown [133]. Streptavidin has also been shown to internalize in 3T3 fibroblast and Jurkat T

cells using TAT peptide [134]. However, in this paper, endosomal release was not spontaneous, but was instead stimulated with poly(propylacrylic acid), which lowers the pH.

Gold nanoparticles (AuNPs) have also been used in combination with siRNA, streptavidin, and TAT for controlled gene knockdown. In this case, the siRNA was attached to the AuNPs through 5'-thiols and released from the AuNPs by NIR irradiation. To aid cell delivery, the 3'-end was biotinylated and bound to streptavidin-TAT. However, the protein was not released from the siRNA once in the cell and this significantly lowered the knockdown efficiency [130, 131].

In this work, biotin-TAT was purchased and bound to streptavidin to help cell delivery. Moreover, to help visualization, a biotinylated-fluorophore was also bound to streptavidin. Specifically, cyanine-5 (Cy5), with an excitation wavelength of 651 nm and emission of 670 nm.

3.6 Chapter overview

The goal of this project was to control gene expression in mammalian cells. Following a similar route to the one showed for the light-activated antisense oligonucleotides, amino-siRNA sequences were modified with biotin/streptavidin UV photocages. Once in the cell, the siRNA would be inactive until irradiated with UV light. UV would then trigger RNAi activation and therefore the knockdown of a target gene. Since streptavidin has a tetrameric structure and only one of the four monomers would be attached to siRNA while the other three would have free biotin-binding sites, fluorophores and cell-penetrating peptides would be attached to the protein. These molecules could be easily interchanged with other

biotinylated ligands, depending on the application of the light-activated siRNA. The fluorophore attached would help visualizing the siRNA inside the cell, while the cell-penetrating peptide would allow for self-delivery of the siRNA/streptavidin construct to mammalian cells, without the aid of any transfecting agents. Therefore, the final construct would be light-activated, self-delivered, and traceable by fluorescence (Figure 3.3). While the amino-ASOs of Chapter 2 were purchased, two strategies were exploited to generate siRNAs containing modified bases: one was to synthesize both strands using an enzymatic method, introducing amino-modified bases in the form of triphosphates compatible with the polymerase, while the other was to synthesize one strand containing the amino group and produce the other via transcription.

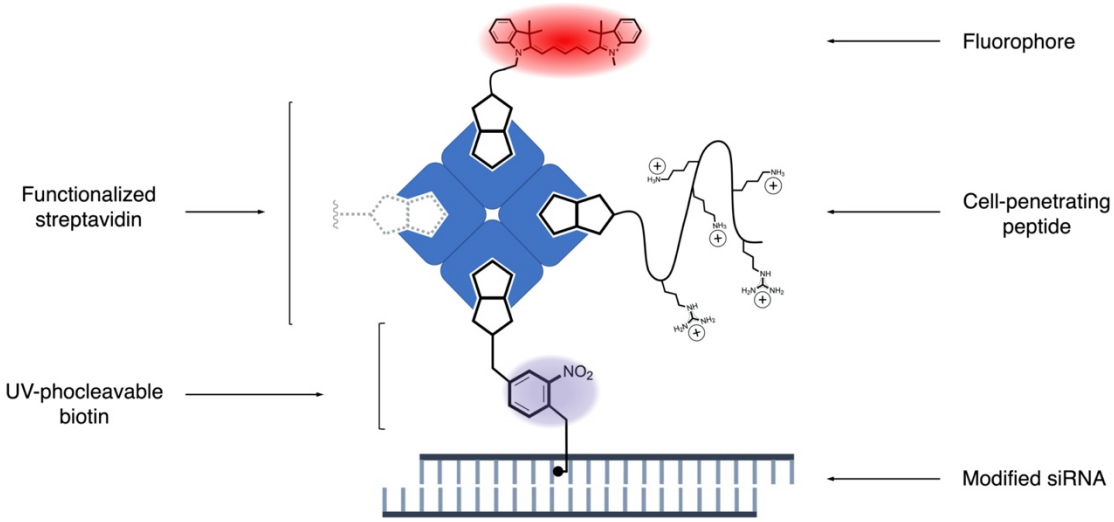


Figure 3.3: Light-activated siRNA with multi-functionalized streptavidin
A streptavidin can be attached to the siRNA via photocleavable groups attached to modified bases, then decorated with different functionalities. In this work, the photocleavable group used was a 2-nitrobenzyl derivative and the functionalities were the fluorophore cyanine-5, and the cell-penetrating peptide TAT.

3.7 Preparation of modified siRNAs

Synthetic small interfering RNAs usually contain 2 nt 3'-overhangs on both strands to be recognized correctly by RISC and maximize gene knockdown [137]. Most commonly, the 2 nts used as overhangs are deoxythymidines, as having terminal deoxynucleotides can reduce nuclease degradation but does not affect the knockdown efficiency. However, in a thorough study comparing several bases at different positions [138], 3'-UU overhangs on the antisense strand were shown to be among the most effective pairs. Overhangs should also be identical on both strands because the passenger strand can in part contribute to the RNAi mechanism. Moreover, terminal A or G should be avoided because unpaired purines have a higher chance of being degraded by common RNases.

The first target selected for mammalian cells knockdown was green fluorescent protein GFP. The siRNA sequence used was adopted from an siRNA designed by Simon Friedman et al. and containing dT overhangs [139]. In their work, the sequence showed remarkable GFP knockdown. The siRNA was named GFP-siRNA (Figure 3.4).

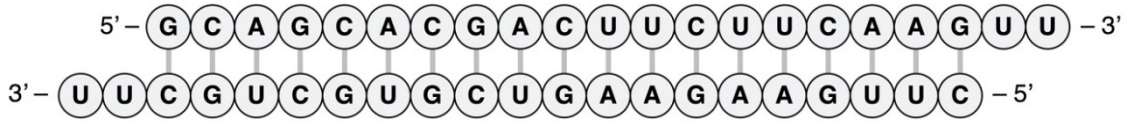


Figure 3.4: GFP-siRNA sequence
The sequence targets GFP-mRNA. Both strands contain 3'-UU overhangs.

The first challenge was to introduce amino-modified bases in the sequence. The work previously done in our group showed how seven modifications on a DNA template containing a T7 promoter were needed to completely prevent RNA polymerase from approaching, hence preventing transcription [88]. Not knowing the number of modifications needed to block RNAi, a quick and affordable method to introduce commercially available amino-modified nucleotide triphosphates at potentially any position was adopted from the work of Xudong Zhu et al. [140], and it is outlined in figure 3.5:

1. Two DNA oligonucleotides containing a T7 promoter sequence, followed by a 10-nucleotide sequence called leader sequence, followed by either the target or its complementary sequence were purchased. Both DNA strands contained TT at the 3'- end.
2. DNA oligonucleotides complementary to the two just described were also purchased and annealed to their complementary, forming two dsDNA sequences:

   T7 promoter + leader sequence + target sequence + TT.

   T7 promoter + leader sequence + complementary sequence + TT.
3. Transcription of the top strand of the two DNA templates was carried out using standard T7 RNA polymerase (RNAP) and triphosphates and yielding ssRNAs containing the leader sequence, followed by the target or complementary sequence and terminal 3'-UU. At this stage, any modified triphosphate compatible to the RNA polymerase could be added to one or both strands. For this work, an aminoallyl (AA)-modified uridine triphosphate was added, therefore introducing an amino-modification at every uridine position.

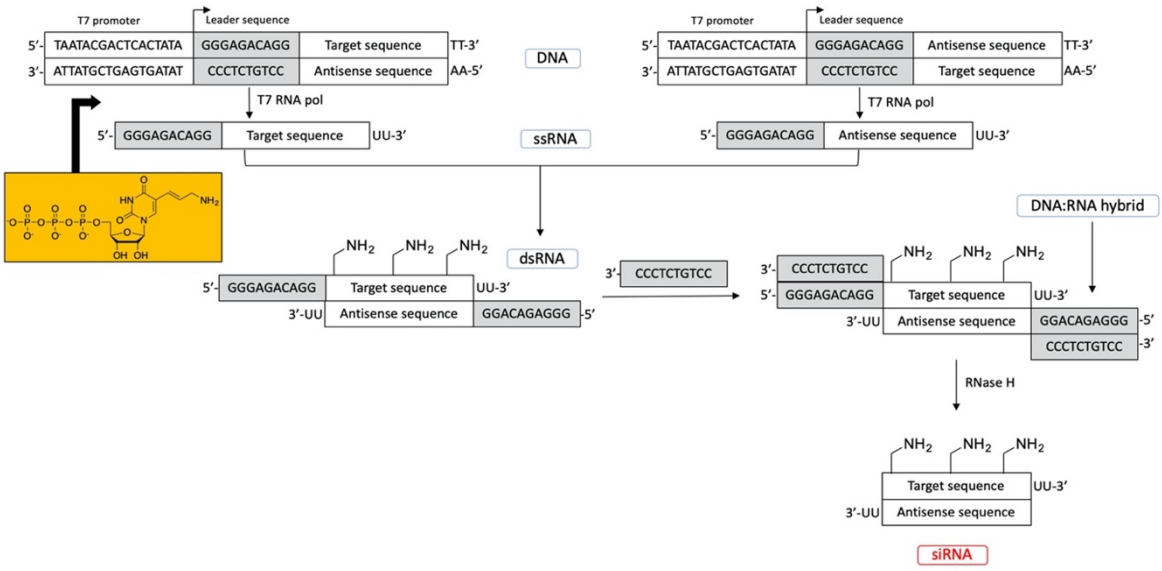


Figure 3.5: Method followed for the generation of amino-modified siRNAs
The diagram shows the steps followed to generate amino-modified siRNA starting from commercially available DNA oligonucleotides.

4. The two partially complementary ssRNAs were mixed to form a dsRNA, where the only parts that could anneal were the target and complementary sequence, since the two leader sequences and the 3'-UU overhangs were the same exact sequence and could not base pair:

   Top strand: 5' – leader sequence + target sequence + UU – 3'.

   Bottom strand: 5' – leader sequence + complementary sequence + UU – 3'.

5. A 10-nucleotide DNA sequence called UHO and complementary to the leader sequence was added to the dsRNA. Here, RNase H was added to selectively degrade the DNA:RNA duplex and yield the final siRNA, containing amino-modified nucleotides and 3'-UU overhangs.

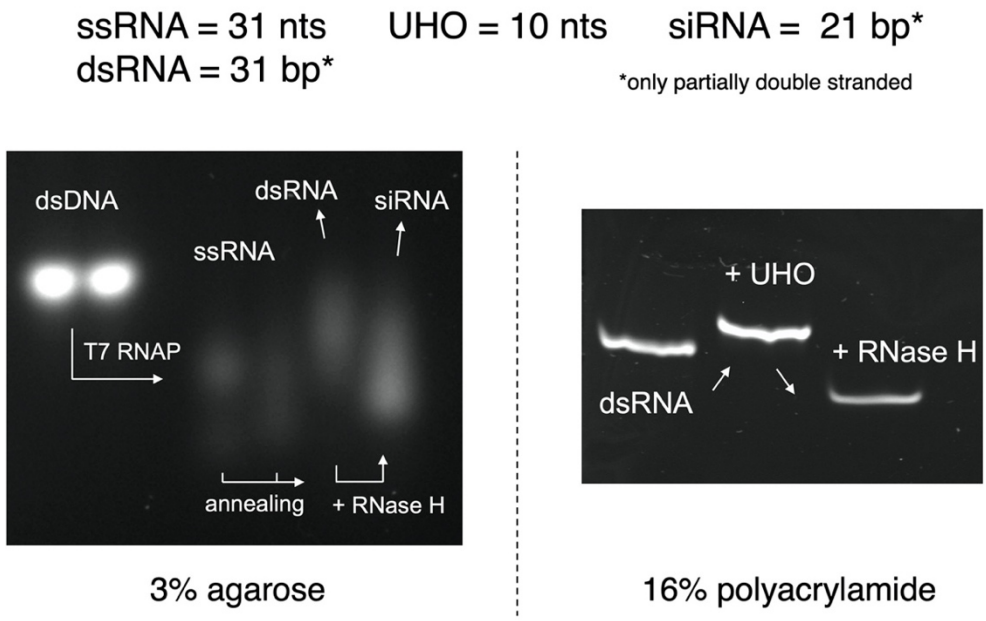


Figure 3.6: Comparison between agarose and polyacrylamide gel electrophoresis analysis

The main steps described in the text are analysed by gel electrophoresis. The use of agarose gel electrophoresis to analyze siRNAs is not as accurate as electrophoresis with high percentage polyacrylamide gels (16-20%). However, the steps used to generate siRNAs can still be followed on agarose gel.

Although every step could be followed by agarose gel electrophoresis, a more precise analysis was obtained using polyacrylamide gel electrophoresis (PAGE), as the size of the nucleic acids produced was too small even at high agarose percentage (Figure 3.6).

The T7 RNA polymerase transcription in this procedure yields very high amounts of RNA from a low amount of inexpensive DNA, therefore this protocol could be easily scaled up to produce large quantities of siRNA. However, HPLC purification of the unmodified antisense strand and following LC-MS analysis showed the presence of an extra base inserted in ~25% of the product. This is a common effect observed with the T7 RNAP used

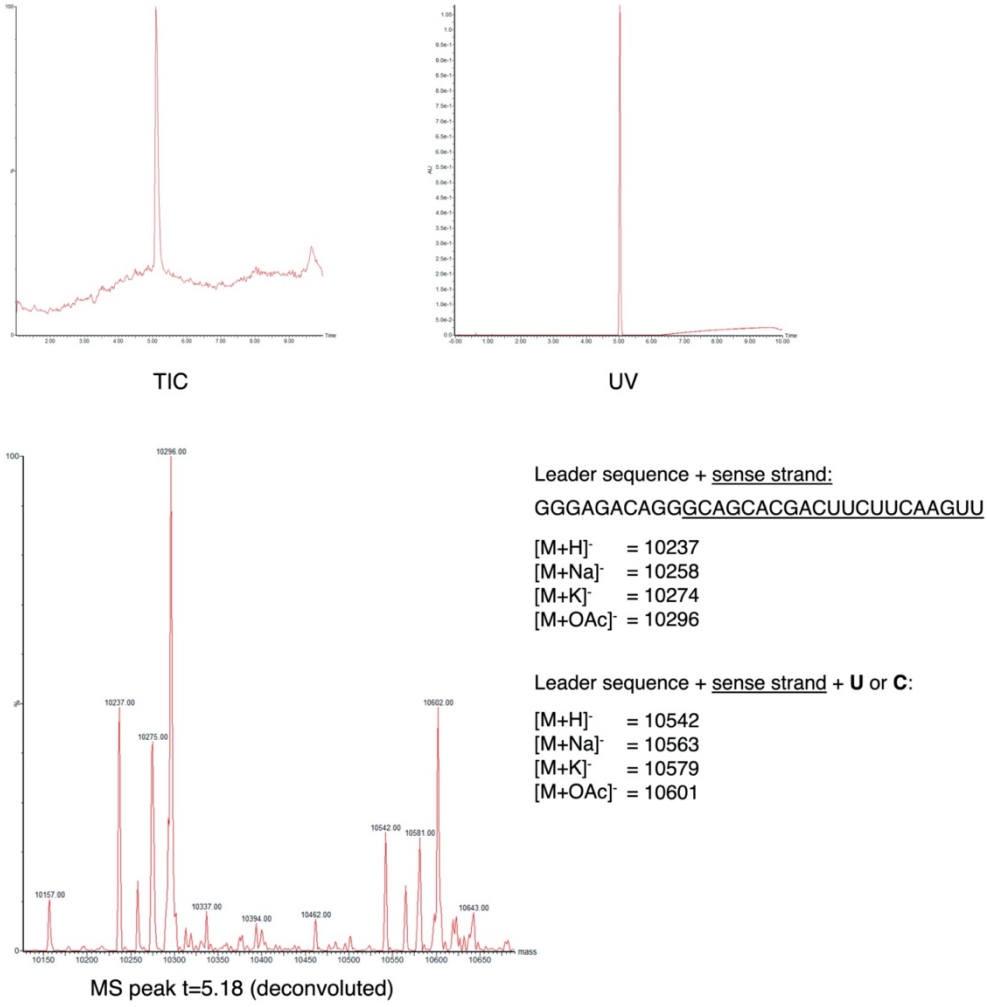


Figure 3.7: Mass spectrometry spectra after ssRNA transcription

The mass peaks correspond to the ion adducts of two separate products with a difference of one nucleotide. The ion adduct $[M+OAc]^-$ is derived by the NH_4OAc buffer used during HPLC purification.

in the transcription reaction of short transcripts, especially when the reaction is carried out with high concentrations of DNA template like in this case [52, 53]. Even if sequence dependent, the extra base is usually inserted at the 3'-end through different termination pathways. This can be observed in the mass spectrometry spectrum (Figure 3.7), where the secondary and inseparable product containing an extra cytidine or uridine (the difference between the two nucleotides was not detectable in this MS analysis) is detected after deconvolution, while the chromatogram and gel electrophoresis show a single clean product. The $[M+OAc]^-$ ion adduct derives from the HPLC purification carried out with an NH_4OAc

buffer. As the last step of the procedure used is an RNase H degradation, the extra base inserted could also be derived from incomplete RNase digestion.

The presence of an extra base partially influenced the knockdown efficiency of the siRNA since the polymerase error described occurred for both sense and antisense strand. The sequence was tested in 3T3 fibroblasts stably expressing GFP, and the fluorescence intensity measured by epifluorescence microscopy (Figure 3.8). After some preliminary studies to optimize the concentration, the siRNA was transfected at 20 nM, using standard lipofectamine transfection procedures. A scrambled sequence (SCR) was also designed, by keeping the same exact base content as GFP-siRNA, but changing the base order, so that it would not target any gene present in the cell line. As shown in the graph below, lipofectamine alone and lipofectamine transfection of the SCR-siRNA did not affect GFP fluorescence (respectively 96% and 94% GFP expression). Transfection with GFP-siRNA resulted in 64% of GFP expression showing gene knockdown, even if affected by the sequence errors described.

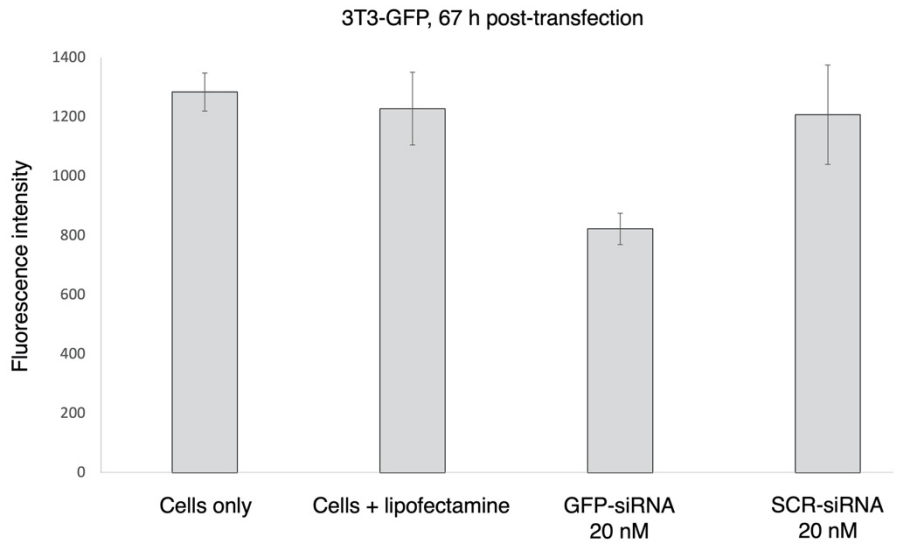


Figure 3.8: Gene knockdown of GFP-siRNA compared to a scrambled control
Knockdown efficiency of GFP-siRNA transfected into 3T3 fibroblasts stably expressing GFP. The fluorescence was measured by epifluorescence microscopy.

3.8 Aminoallyl siRNAs

With the same procedure described in Section 3.7, aminoallyl-uridine triphosphate (AA-UTP) was added while generating the sense strand or the antisense strand of the same sequence, which were then annealed together or with the complementary unmodified strand in different combinations. The following dsRNA sequences were generated:

- amino-modified sense strand + unmodified antisense strand: 6AA-S-siRNA;
- unmodified sense strand + amino-modified antisense strand: 7AA-AS-siRNA;
- both strands containing amine modifications: 13AA-SAS-siRNA.

The modified sequences were reacted with NHS ester photocleavable biotin groups, the same UV photocleavable group showed in Chapter 2. As the reaction would most likely yield a complex mixture of products, only approximate purification and desalting were carried out. However, due to the fact that the double stranded reaction products would be incompatible to analyze by HPLC, as the organic solvents would break the double stranded duplex, the NHS ester reaction was initially carried out on the aminoallyl sense and antisense separately. Therefore, the sense ssRNA containing six amine modifications 6AA-S-ssRNA was initially treated with NHS ester-PCB and analyzed by HPLC and PAGE (Figure 3.9). The ssRNA fully reacted with the NHS ester, as the starting material completely disappeared. However, the gel resulted in a ladder of products, indicating how not every amine reacted with the NHS ester. Instead, the five visible bands were increasingly larger in size, as more and more amines reacted. This result was similar to what was observed with the C6-amino linker on the ASOs, resulting in an HPLC peak with a small shoulder preceding the main peak corresponding to the not fully reacted product. Though, in that case most of the reaction went to completion. Here, the reaction was not complete as the aminoallyl reactivity is probably much lower than the one of longer C6-alkyl amine groups,

due to the lower exposure in solution. Especially with some of the modified bases being next to each other in the sequence.

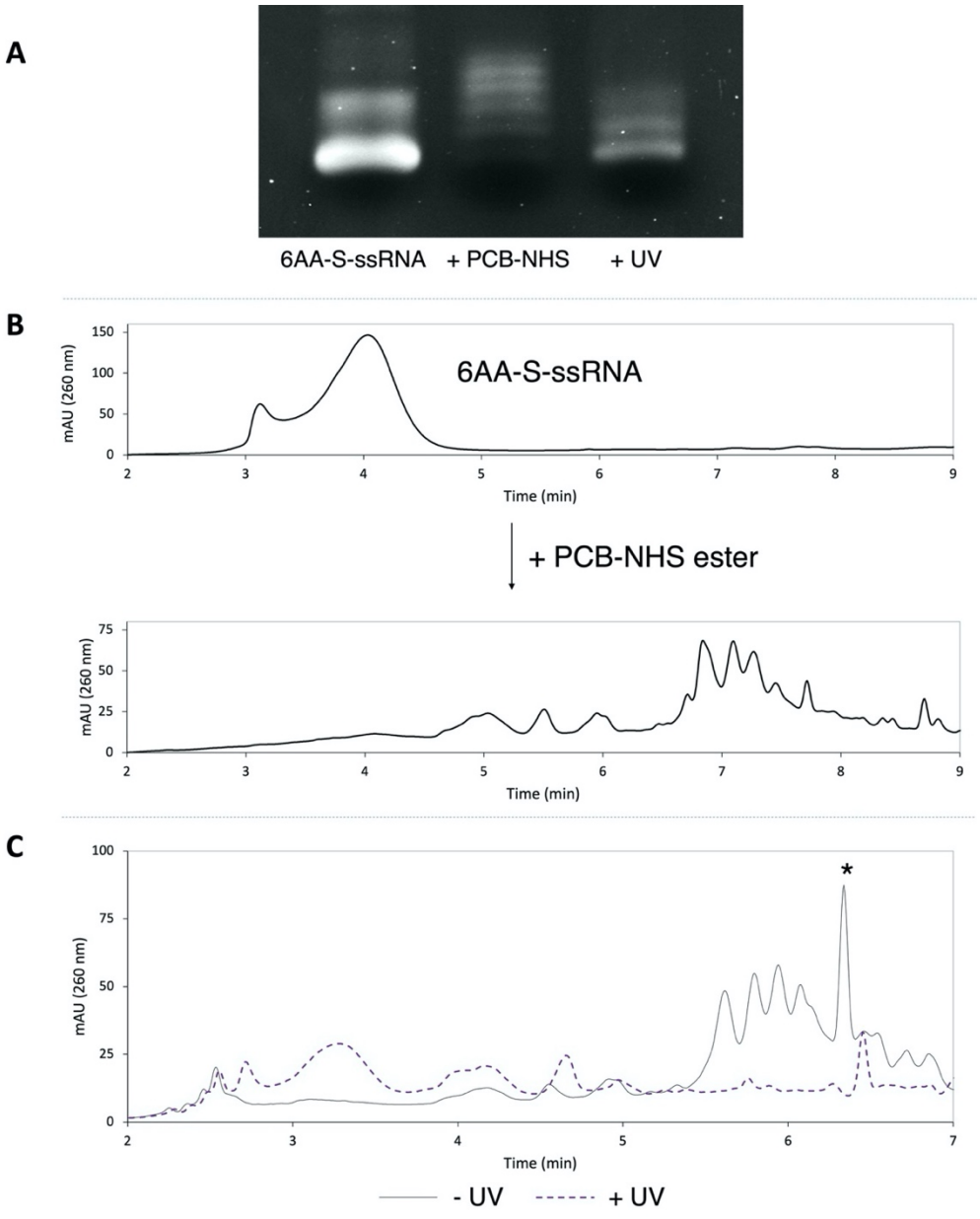


Figure 3.9: Reaction between PCB-NHS ester and AA-siRNA
As shown in figure A, the reaction resulted in a multitude of products, but the starting material completely reacted. B) HPLC chromatogram showing what the gel suggested. C) Partial recovery of starting material after UV irradiation.

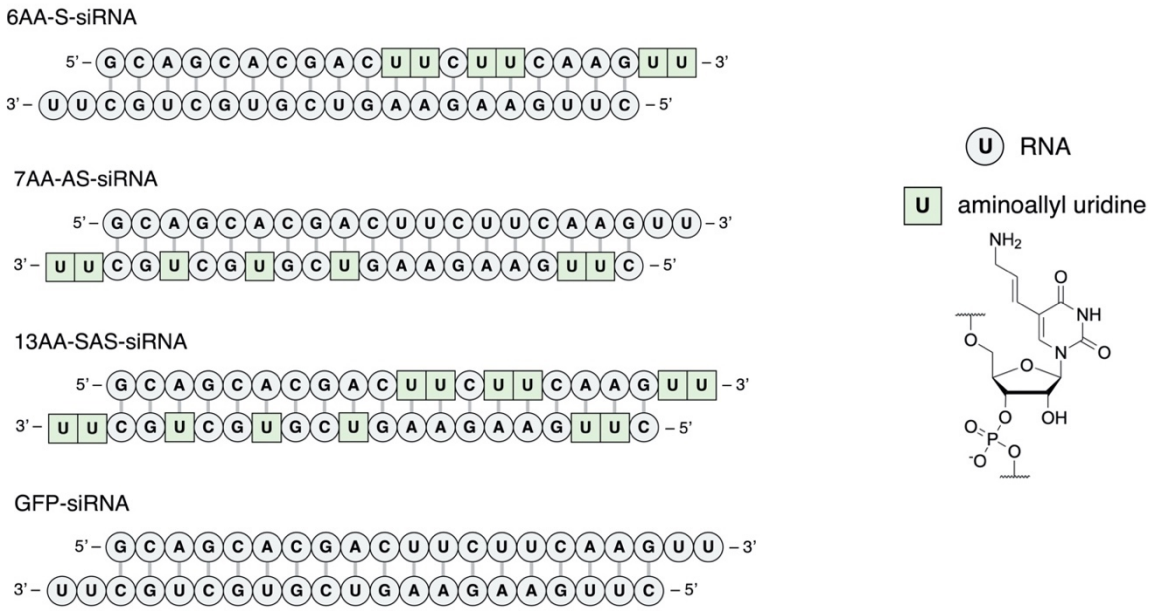


Figure 3.10: AA-siRNAs generated and directly reacted to PCB-NHS ester
Aminoally-modified siRNA sequences showed were reacted with the PCB NHS ester.

PAGE analysis showed how UV irradiation yielded the fully recovered aminoallyl-ssRNA as the major product, but some ssRNAs with one or two photocages on were left. The HPLC spectrum showed a similar result. The aminoallyl-ssRNA was a broad peak due to the several amines present, fully converted into a complicated mixture of inseparable products after addition of PCB-NHS ester, each corresponding to a different number of PCB groups bound. Upon UV irradiation, the aminoallyl-ssRNA peak was partially restored, but with some of the PCB groups not fully uncaging. HPLC purification of a single PCB product would have been practically impossible, therefore, for the following experients, the PCB-NHS ester was directly reacted to the aminoallyl-siRNAs sequences showed in figure 3.10. After the reaction following the same procedure used for single stranded nucleic acids, the products were extensively washed with water using centrifugal spin filters to remove the dimethylformamide (DMF) used in the reaction, the excess reagents present, and some of

the salt. The AA-siRNAs were analyzed before and after UV irradiation (Figure 3.11). PAGE analysis resulted in all modified siRNAs forming multiple products once reacted with the NHS ester, then most of the product recovering after UV irradiation.

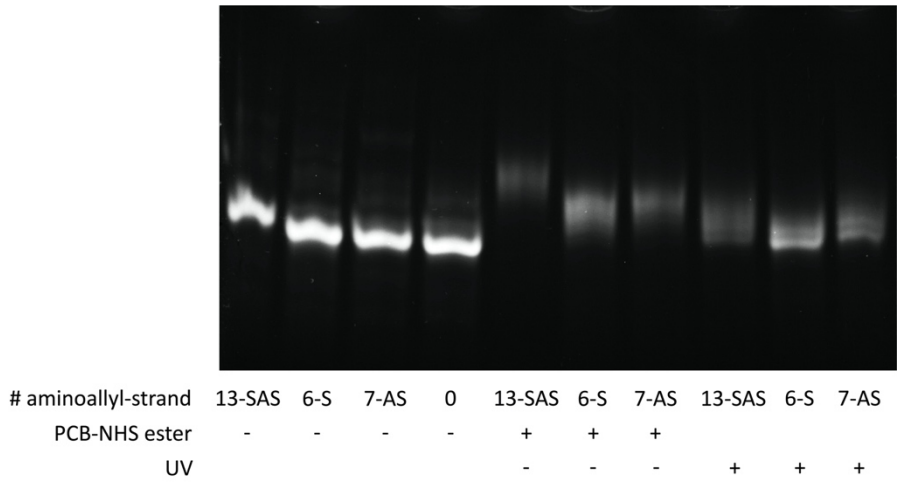


Figure 3.11: Reaction between PCB-NHS ester and AA-siRNAs
Aminoallyl-siRNAs before and after reaction with PCB-NHS ester and after UV.

The two siRNAs containing amines on only one of two strands were then tested in 3T3-GFP cells to evaluate the effect on RNAi (Figure 3.12). As expected, modifying the sense strand had little to no effect on gene silencing, while amines on the antisense strand almost completely inhibited siRNA activity. This can be explained by the fact that the antisense guide strand is the one selected by RISC to find base complementarity with mRNA before degrading it. Even if amine groups on a single nucleobase have little-to-no effect on single base pairing, six or seven amines caused this effect to be much greater. As modification of the sense strand resulted in a negligible effect on siRNA potency, it confirmed how having amines in the sequence would not affect gene knockdown after uncaging.

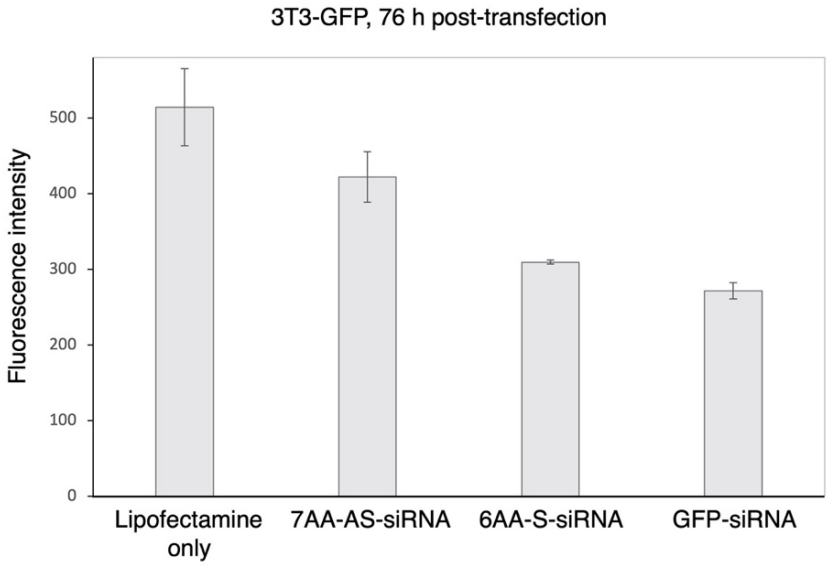


Figure 3.12: Effect of different strand modifications on gene knockdown
Amino-allyl modifications on either strand of the siRNA were compared to each other and compared to GFP-siRNA. The results shown are gene knockdown measured on 3T3 fibroblasts stably expressing GFP and have been measured by epifluorescence microscopy. Modification of the antisense strand (7AA-AS-siRNA) has a much higher negative effect on the siRNA knockdown efficiency.

To lower the number of aminoally modifications on the antisense strand and hopefully lower their negative effect on siRNA activity, different amounts of aminoallyl-UTP were added to the antisense strand transcription reaction mixture, 25% and 50% AA-UTP (the remaining % being unmodified UTP). The siRNAs were named 25p-AS-siRNA and 50p-AS-siRNA. The positioning of the amine groups throughout the sequence would be a mixture of all the possibilities and a distribution of number of modifications. Nonetheless, this strategy was chosen because after adding streptavidin, blocking the antisense strand might have had a greater impact on turning off gene silencing. After the first transcription with different percentages of AA-UTP, the two ssRNAs were analyzed by HPLC before and after addition of the PCB-NHS ester. The chromatogram showed how not all the starting material was converted to product, while in a parallel reaction with fully aminoallyl-

modified sense strand it was indeed completely converted (Figure 3.13A). This indicates

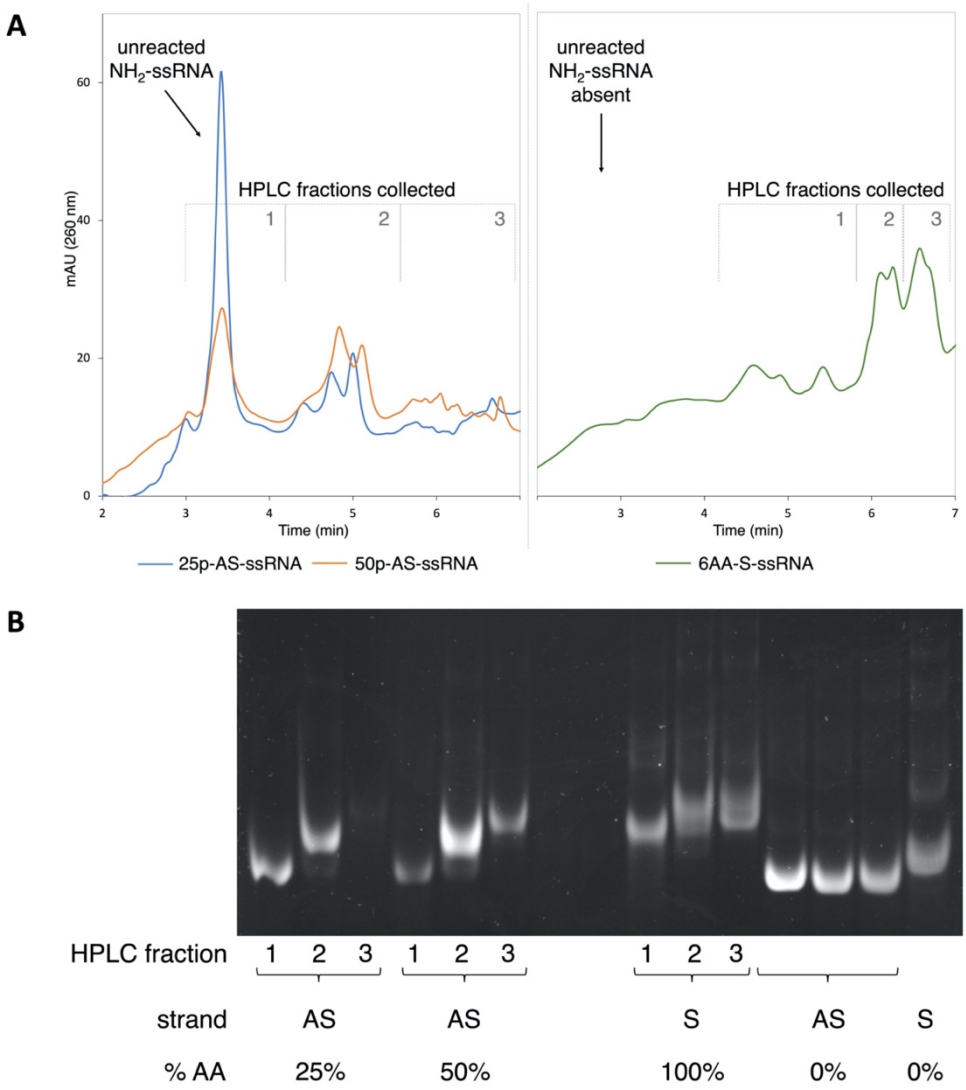


Figure 3.13: Purification of partially modified ssRNA

A) HPLC chromatograms of RNA with different amounts of aminoallyl-UTP added during transcription. The presence of unreacted amino-siRNA for transcriptions with 25% and 50% AA-UTP added (left) shows how T7 RNAP prefers to insert an unmodified base compared to an amino-modified one. B) The HPLC fractions were analysed by PAGE.

how the T7 RNAP preferably inserts an unmodified UTP when available in solution instead of an aminoallyl-UTP. This happens more often when more unmodified UTP is in solution, which is why the unreacted amino-ssRNA peak is larger when adding 25% AA-UTP compared to 50% AA-UTP. For further analysis, the HPLC chromatograms were divided in three sections, and each section purified and analysed by PAGE. The more apolar fractions corresponded to larger products, indicating an increasing number of PCB modifications inserted and a lower number of amines left. The gel confirms full reaction of the sense strand, while for the antisense strand 25% and 50% seemed to have the same mass (Figure 3.13B). This was due to the fact that both 25% and 50% yielded the same unmodified or partially modified transcription products, just at different ratios. This strategy was therefore abandoned moving forward.

3.9 C6-amino linker

A better option to improve the reactivity of the amino groups was to follow the work done for the ASOs by using C6-amino modifications. For this purpose, a sense strand containing two amino-C6-uridine modified bases was kindly synthesized by Dr. Afaf El-Sagheer and named 2xC6-NH_2-ssRNA (Figure 3.14). The modified bases were inserted during RNA oligonucleotide synthesis by adding commercially available C6-amino modified uridine phosphoramidites to a standard RNA oligonucleotide phosphoramidite synthesis procedure. The amine position was chosen following the criteria of placing the photocages in the middle of the sequence and not on adjacent bases.

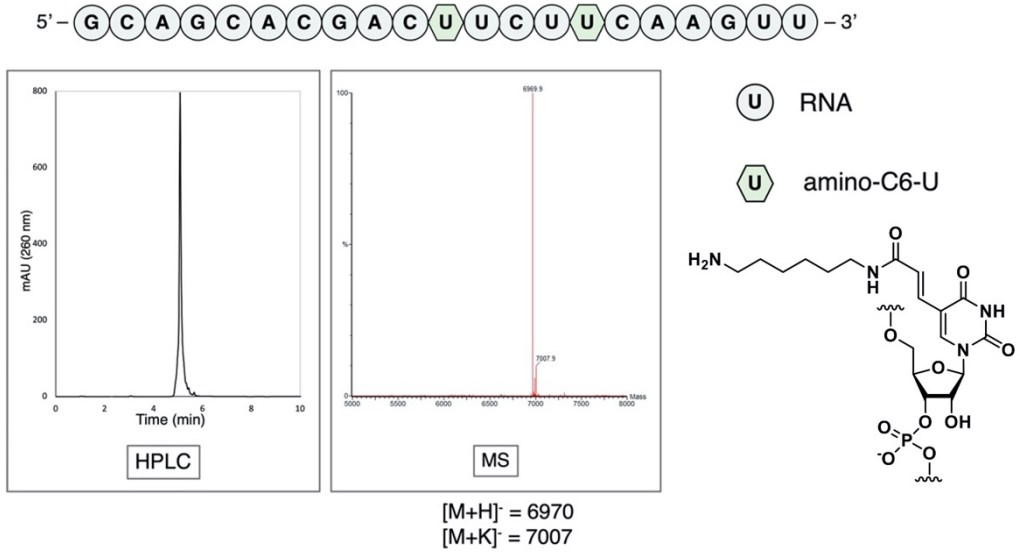


Figure 3.14: Structure, HPLC, and MS spectra of synthetic 2xC6-NH_2-ssRNA
The purity and molecular weight of the modified ssRNA sequence was confirmed by HPLC and mass spectrometry. The structure of the C6-amino linker modified bases is shown.

On the other hand, the antisense strand was generated by T7 RNAP transcription. The siRNA was finally made by following the same method used for the aminoallyl-siRNA. 2xC6-NH_2-ssRNA was first modified with the same NHS-ester 2-nitrobenzyl PCB groups used in the rest of this work to obtain 2xPCB-ssRNA. The modified ssRNA was HPLC purified, then analyzed by LC-MS (Figure 3.15).

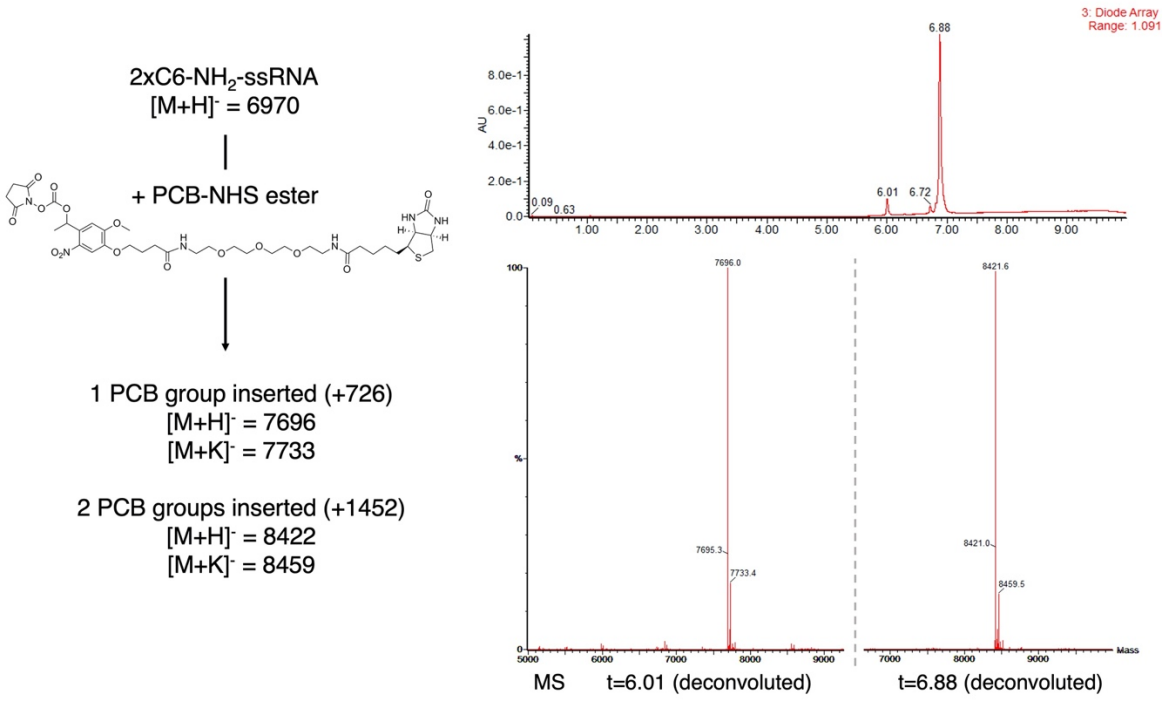


Figure 3.15: Reaction outcome of 2xC6-NH_2-ssRNA and PCB-NHS ester

The amino-modified siRNA containing two C6-amino linker was modified with PCB-NHS ester group, the analysed by LC-MS. In the chromatogram, the major product corresponds to both amines fully reacted with the NHS ester (peak at t=6.88), while a minor product (t=6.01) corresponds to only one of the two amines reacted.

The HPLC purification, which followed the same procedure used for ASOs, resulted in a major product containing both PCB groups, plus a small impurity containing only one modification, which was inseparable using the HPLC method followed. PAGE and HPLC after purification showed UV-photocleavage of the product with recovery of the amino-ssRNA (Figure 3.16).

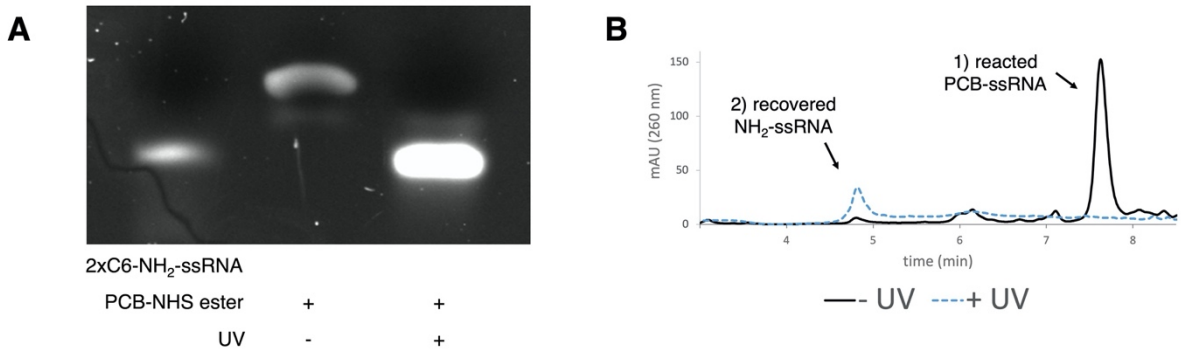


Figure 3.16: Light-responsive 2xC6-PCB-ssRNA

The ssRNA containing two photocleavable modifications was tested after HPLC purification. On both PAGE (A) and HPLC (B) analyses, the starting material fully reacted with the NHS ester containing the photocleavable group. Upon UV irradiation, the photocleavable group cleaved off and the amino-ssRNA was recovered.

2xPCB-ssRNA was then annealed to the unmodified antisense transcription product still containing the leader sequence, and the resulting dsRNA treated with UHO and RNase H to obtain 2xPCB-siRNA. The siRNA was analysed by native PAGE (Figure 3.17), and resulted in a single clean band. This was a major improvement compared to the aminoallyl-siRNA described earlier in this chapter. The siRNA now containing two biotinylated photocleavable (PCB) groups was also analysed after binding to monovalent streptavidin (mSA) to form the photocages (Figure 3.17). As the polyacrylamide gel used was of high percentage, streptavidin could not travel through. Therefore, siRNA fully bound to mSA remained in the well. Upon UV irradiation, the siRNA was almost completely recovered. A faint band was still visible on top of the gel, but this could be due to the UV conditions applied.

2xPCB-siRNA

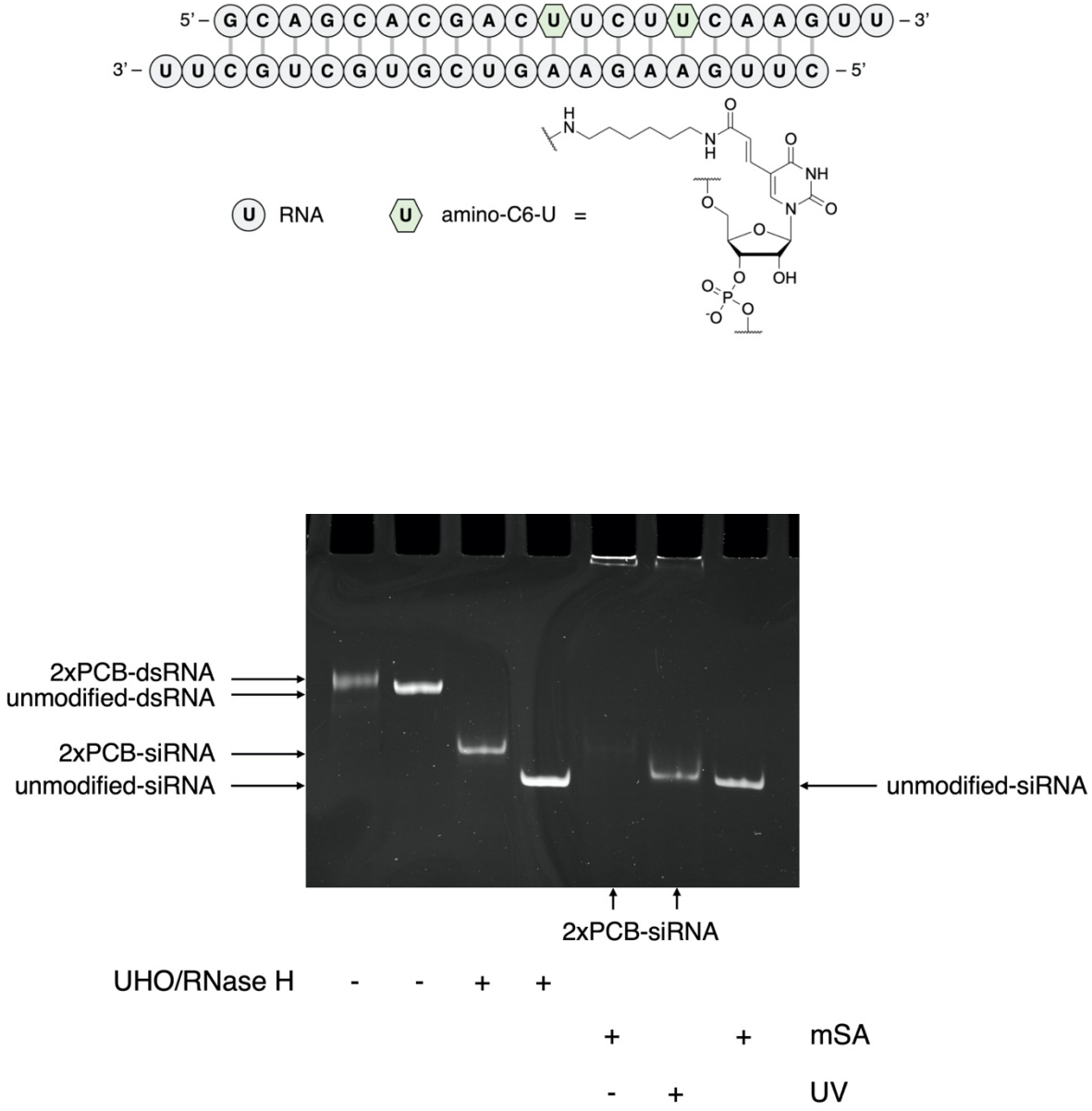


Figure 3.17: Formation of 2xPCB-siRNA

Top: the siRNA structure containing two PCB-modifications. Bottom: the 2xPCB-modified ssRNA sense strand was annealed to the antisense strand produced by transcription. Then, UHO/RNase H processing yielded the PCB-siRNA. The biotins of the PCB group bound to streptavidin, which was too large to travel through a high percentage polyacrylamide gel. UV irradiation cleaved off the PCB/streptavidin groups and recovered the amino-siRNA, which could travel through the gel. The identity of the uncaged siRNA was confirmed by comparing the molecular weight to an unmodified-siRNA.

To better resolve the protein-siRNA conjugate, an agarose gel analysis was carried out (Figure 3.18A). On agarose, the siRNA bound to streptavidin could easily travel through the gel. Different UV conditions were also tested. With agarose, it was clearly visible how the PCB-siRNA starting material was fully bound to streptavidin, forming a brighter band at higher molecular weight, and a fainter band right below. The two bands corresponded to one (the fainter) and two (the brighter) PCB groups bound to streptavidin, indicating how most of the product contained two mSA. This was even more evident when the mixture was irradiated with UV, as most of the brighter band was either cleaved off to recover the starting siRNA band or the fainter band, corresponding to only one of the two mSA cleaving off. Moreover, from this experiment, UV conditions of 2 mW for 5 minutes were selected for future experiments, as the irradiation time showed a good on state and it would be not too long or powerful to cause cellular damage. A faint band at the top of both siRNA-mSA bands was also visible. This was assigned to unspecific binding of mSA to the siRNA, as it did not decrease in intensity with UV irradiation, but still stained with the nucleic acid stain used.

The same streptavidin binding experiment was carried out by including three biotinylated cell-penetrating peptide TAT, which were incubated with streptavidin prior siRNA binding. As TAT is highly positively charged, the siRNA-streptavidin-TAT construct could not move from the well during electrophoresis, and the size difference appeared larger than the one expected (Figure 3.18B). The siRNA band was recovered after irradiation at 365 nm.

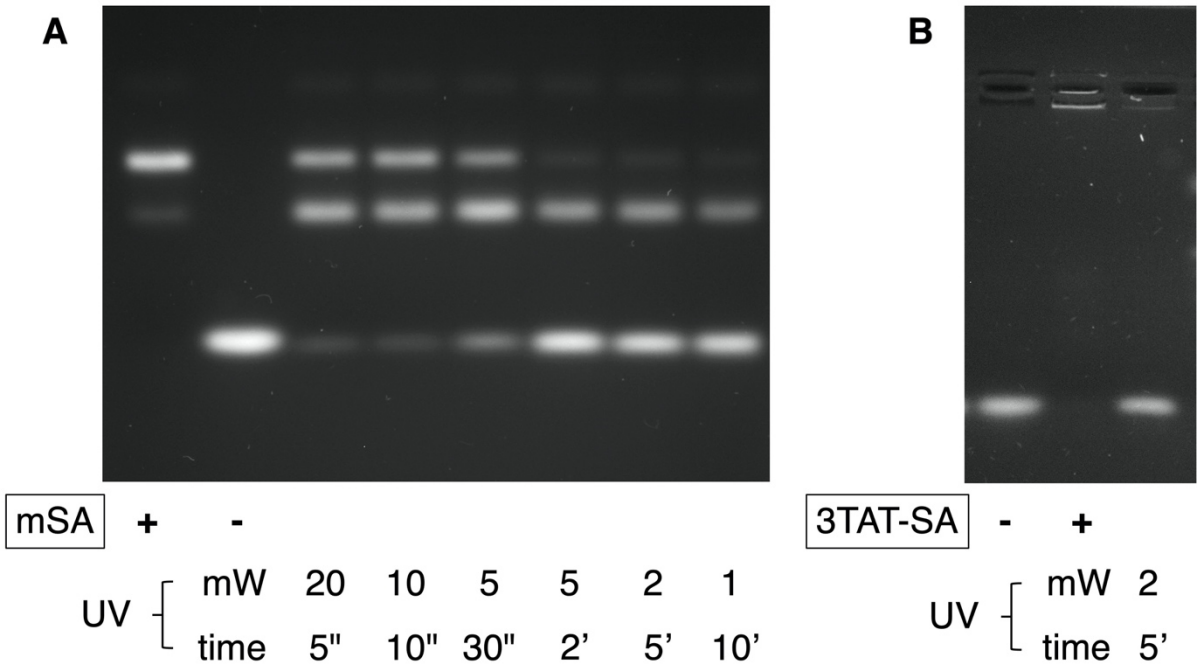


Figure 3.18: 2xPCB-siRNA binding to streptavidin

A) After complete binding to monovalent streptavidin, different UV intensities were tested. B) When the siRNA was bound to streptavidin containing highly positively charged peptides, it could not travel across the agarose gel.

Before binding to PCB/streptavidin, a siRNA containing two amines was generated and tested in mammalian cells. The siRNA was prepared following the same general method and by using 2xC6-NH_2-ssRNA as sense strand. The test in mammalian cells was done to confirm that amines on the sequence would not affect the gene knockdown efficiency. As shown in figure 3.19, not only the siRNA caused gene silencing of GFP, but it appeared to have higher activity than the unmodified siRNA prepared previously. This was hypothesized to be caused by the insertion of extra bases observed while generating GFP-siRNA via transcription method (explained in Section 3.7) Interestingly, PCB modifications alone did not seem to interfere with gene silencing.

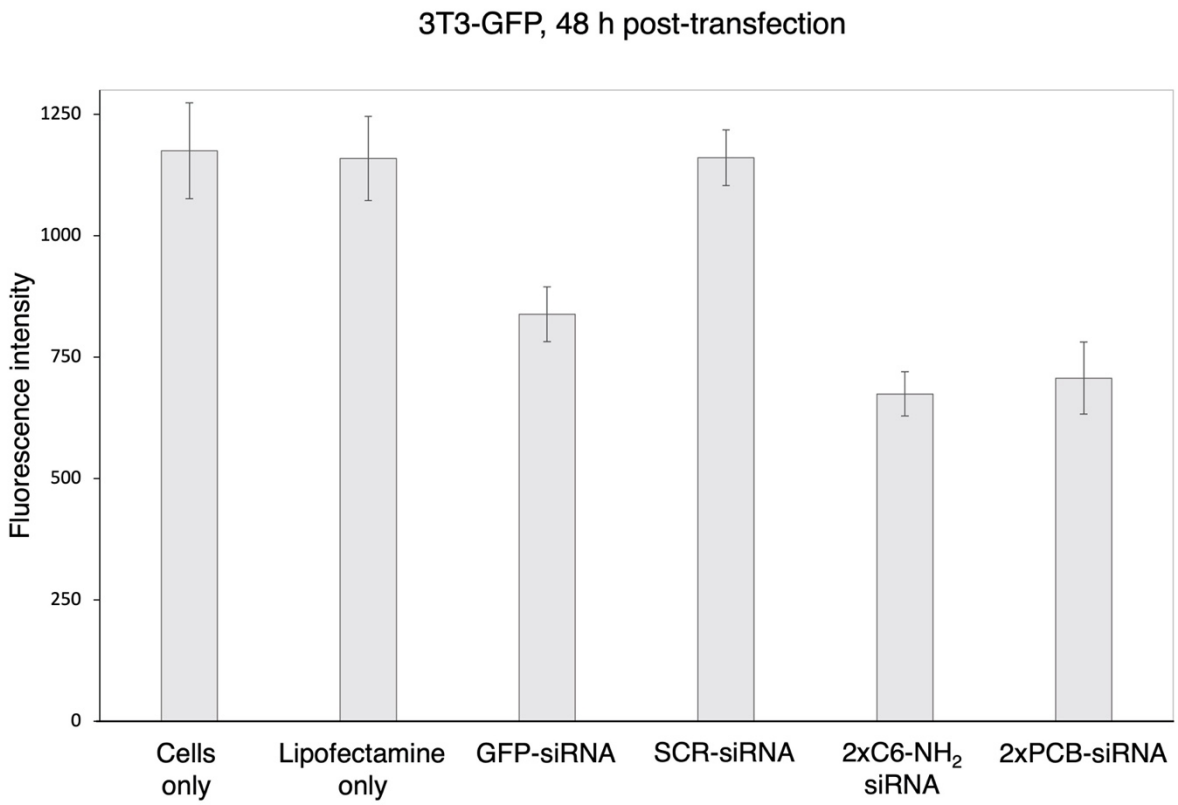


Figure 3.19: Testing of different siRNAs for knockdown in mammalian cells
Different siRNAs were transfected in 3T3 fibroblasts stably expressing GFP and their knockdown efficiency was compared. The cells were analysed by epifluorescence microscopy and by plate reader.

A better knockdown efficiency was obtained with the same experiment and cells, by using a cell tracking dye (Deep Red Cell Tracker), then using a software to carry out cell segmentation using Cell Profiler and analyze knockdown at single cell resolution. After these measurements, a box and whisker plot was constructed, and the results normalized for percentage GFP expression (Figure 3.20). The box and whisker plot shows several high expression outliers, which are caused by the fact that often many cells had off-scale fluorescence for the detection limit of the epifluorescence microscope, due to GFP abnormal overexpression. This can be caused by cellular stress or other unknown and uncontrollable

variables. The broad error bars observed in the plot were mostly caused by this effect. Nonetheless, the resulting analysis showed great knockdown and represented a very reliable technique for knockdown analysis. However, the process was very time consuming, as it was only partially automated. Therefore, for routine experiments it would have been almost impossible to proceed with this for each knockdown analysis. Therefore, the experiments shown in the next figures were analyzed by simple fluorescence measuring by epifluorescence microscopy or by plate reader.

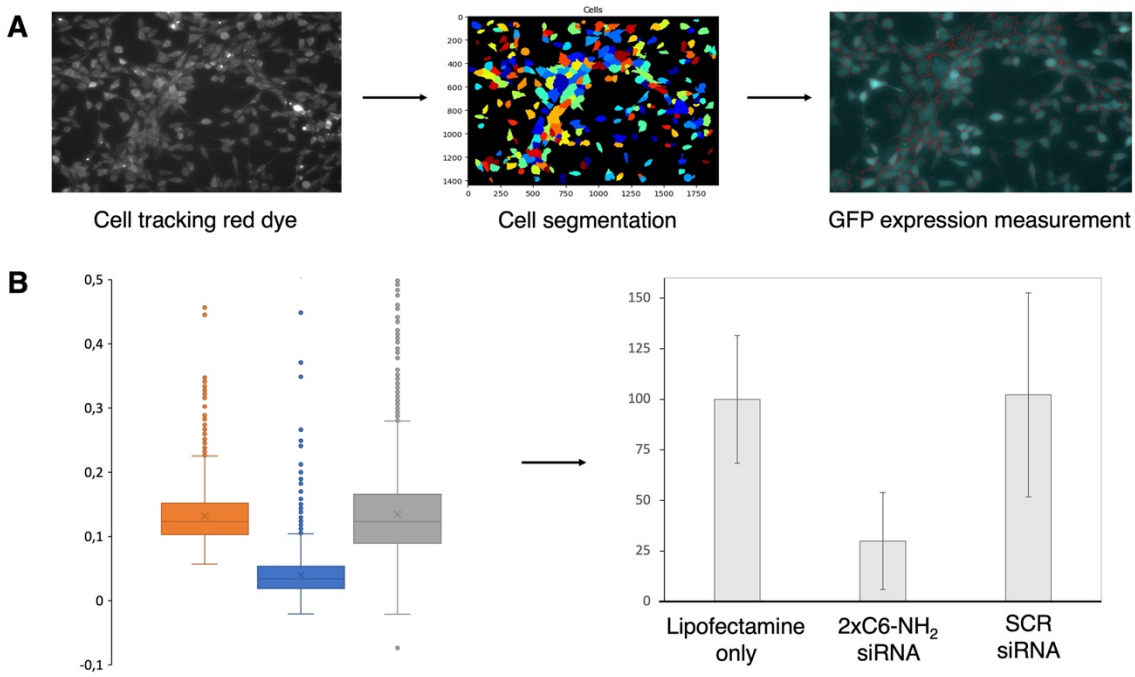


Figure 3.20: Single-cell analysis of gene knockdown

A) Cell segmentation process carried out using Cell Profiler. The mask was created by tracking individual cells using a cell tracking dye. Single cell GFP fluorescence was then measured based on the segmentation profile created. B) A box whisker plot (left) was generated with the cell fluorescent measurements. The graph on the right was derived from the box whisker plot for easier visualization. The fluorescence was normalized against the lipofectamine only fluorescence.

Other than 3T3, siRNA knockdown was tested in another cell line, HEK 293T. Here, the wild-type cell line was used, not stably expressing any fluorescent protein. Therefore, a plasmid expressing mVenus with a mammalian cell system had to be co-transfected with the siRNA for fluorescent measurements. Since mVenus is derived from GFP, the siRNA sequence used previously aligned perfectly with the mVenus-mRNA target sequence and, once tested, it indeed knocked down mVenus expression, with great efficacy (Figure 3.21).

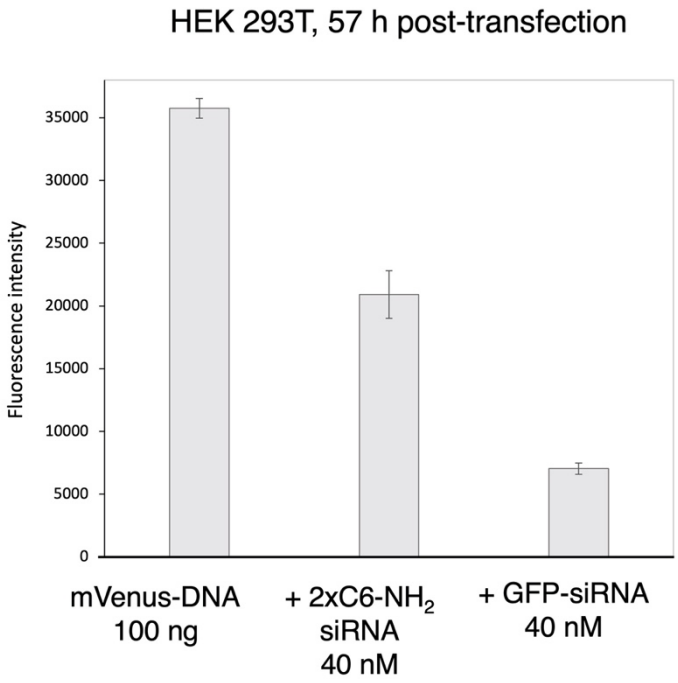


Figure 3.21: Gene knockdown after co-transfection of siRNA and mVenus plasmid
HEK 293T cells were co-transfected using lipofectamine with a DNA plasmid expressing mVenus and siRNA. The siRNAs used were the amino-siRNA and the unmodified siRNA used previously to target GFP.

3.10 Self-delivery to mammalian cells

Finally, self-delivery experiments were carried out. First, the self-delivery of a streptavidin (SA) bound to one equivalent of biotinylated Cy5 (1 Cy5) and two equivalents of biotinylated TAT (2 TAT) was tested (SA [2 TAT – 1 Cy5]) (Figure 3.22), without binding to any siRNA. The same test was repeated for streptavidin only bound to Cy5, with no TAT (SA [0 TAT – 1 Cy5]). The protein was simply added to the cell wells during the

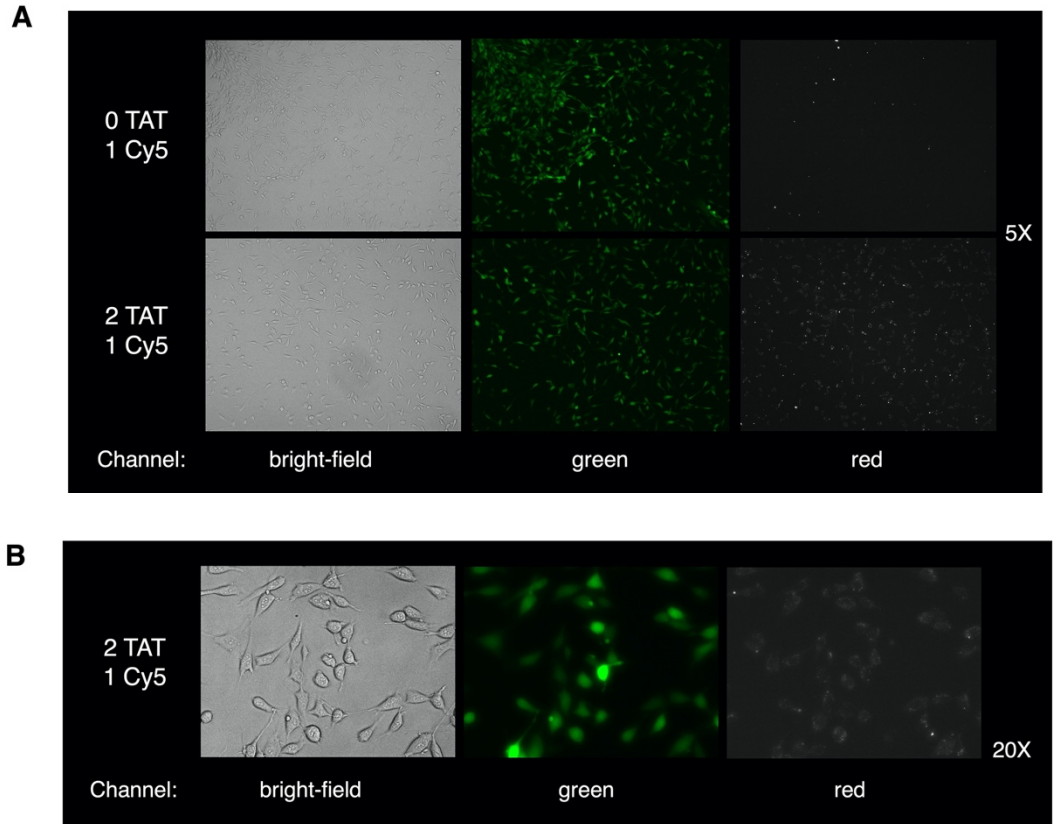


Figure 3.22: Delivery of SA-TAT-Cy5 to mammalian cells

The green channel corresponds to GFP fluorescence in the cells. The red channel corresponds to the Cy5 bound to streptavidin being internalized in the cells or not.
A) Streptavidin bound to cell-penetrating peptide TAT shows self-delivery, while it does not deliver when TAT is absent. B) A greater magnification of some of the cells showing successful siRNA self-delivery.

"transfection" procedure, but without the use of any transfection reagents. Cells were incubated for 6-12 hours after SA [2 TAT - 1 Cy5] was added. The media was then changed to media not containing any colored pH indicator that might interfere with Cy5 red fluorescence. The cells were analyzed by microscopy at different magnitudes. When TAT was not attached to streptavidin, while Cy5 was [0 TAT – 1 Cy5], the cells did not internalized any red fluorescence, while they were still healthy and expressing GFP (Figure 3.22A). When two equivalents of TAT were added to streptavidin, every cell showed fluorescence in the red channel, while also maintaining normal GFP expression levels. Red fluorescence was homogenously distributed throughout the whole cell, suggesting successful deliver to the cytoplasm (Figure 3.22B).

To follow, streptavidin with Cy5 and 2 TAT was finally attached to siRNA and delivered to cells without the use of lipofectamine (Figure 3.23). As for just the protein, the siRNA-SA construct without any TAT did not deliver the siRNA, while the one with TAT bound did. A composite image of the green and red channel in figure 3.23B shows how the red fluorescence nicely overlaps with the cells of the green channel, indicating how these were not just streptavidin aggregates. However, higher magnification microscopy showed how the construct did not distribute as homogeneously throughout the whole cell as just the protein did, indicating formation of the endosome (Figure 3.23C). The endosomes formed resulted in very bright circular spots due to the high local concentration of Cy5. Even though this was a completely expected observation, as most delivery mechanisms go through endosome formation and endosomal escape, such a protein-siRNA construct could have difficulties being released in the cytoplasm.

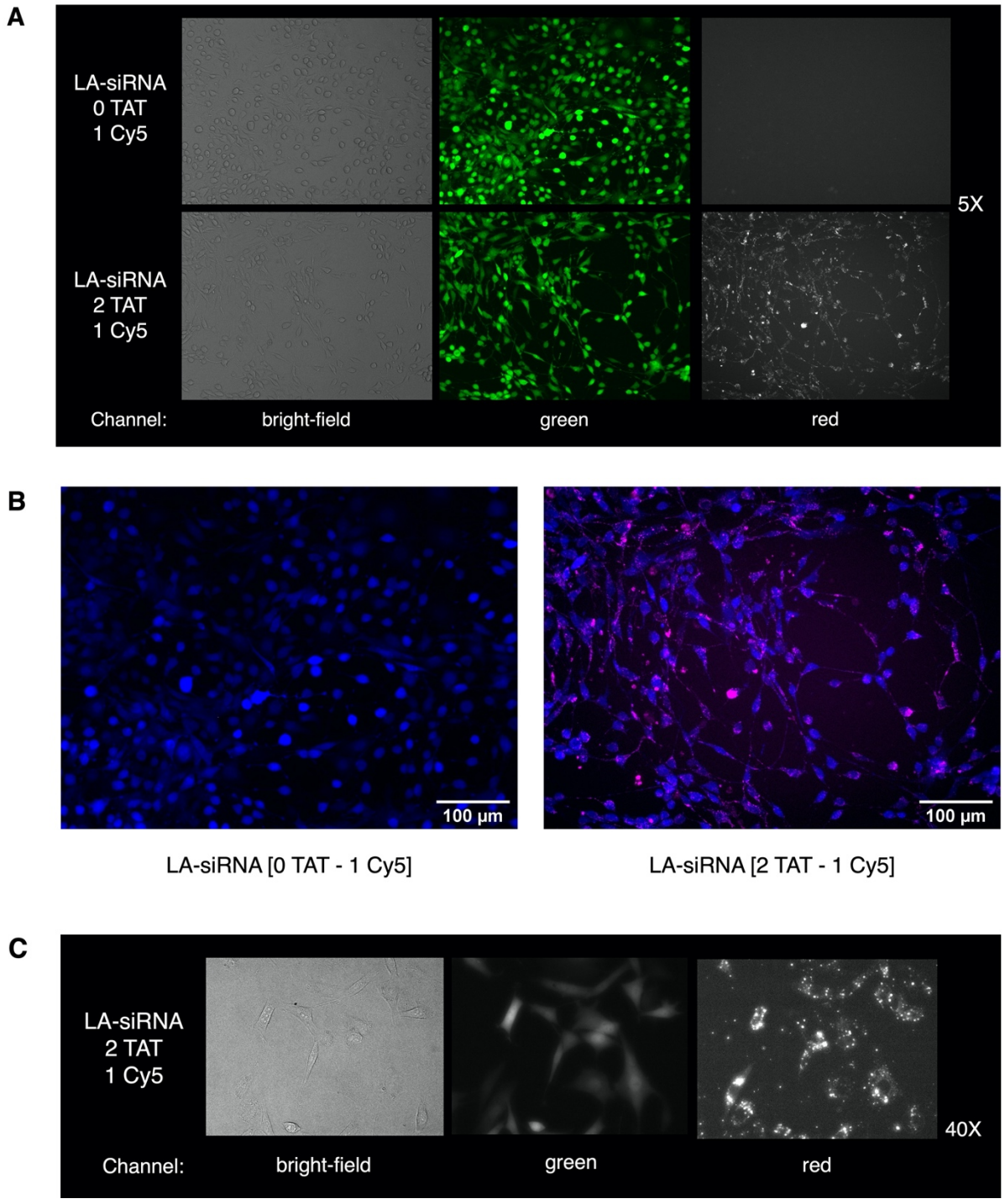


Figure 3.23: Delivery of siRNA bound to SA-TAT-Cy5 to mammalian cells

A) Without cell-penetrating peptide the construct cannot enter the cell, while it can when TAT is present. B) Composite image with clear colors, showing overlapping between cells and delivered siRNA. C) Higher magnification shows endosomal formation.

Next, before trying the self-deliver of modified siRNA then activate it by irradiating with UV, UV damage to the cells was tested (Figure 3.24). Here, after irradiating with the same conditions that would be used for the siRNA, only negligible change in protein expression was observed, sixteen hours after UV was applied.

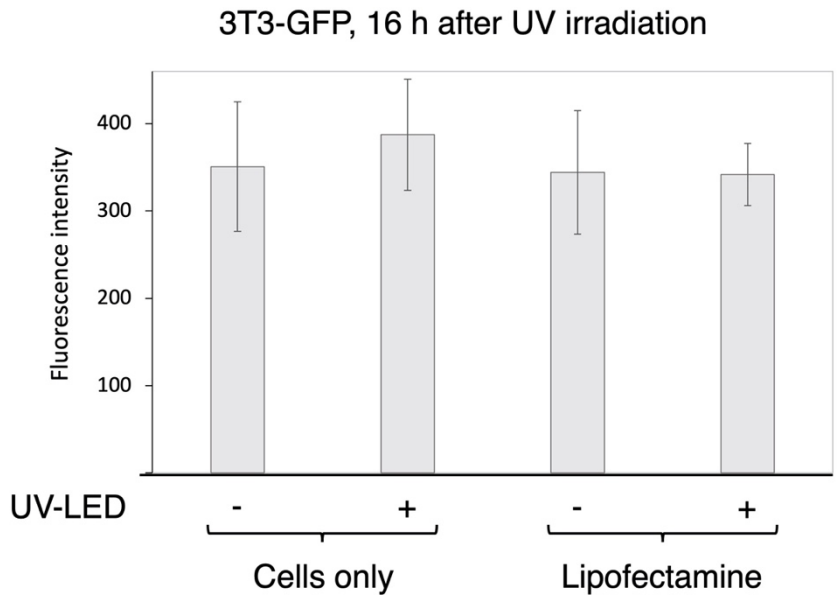


Figure 3.24: Testing of UV damage to cells
UV irradiation did not cause any UV damage to the cells after normal irradiation conditions.

Finally, the self-delivery and light activation of gene knockdown was tested in HEK cells. Figure 3.25 shows the green and red fluorescence channels, to the left and right respectively. The two channels corresponded to mVenus expression and Cy5 fluorescence inside the cell. Cy5 fluorescence seemed to indicate how the light-activated siRNA was indeed delivered to cell, as the increase in fluorescence was comparable to the delivery of

just SA-TAT-Cy5. However, this experiment could not clarify if the protein-siRNA construct was delivered to cytoplasm, or it was just trapped into the endosome and later degraded by the cell. If delivered, it indeed showed an off state compared to the amino-siRNA. However, upon UV irradiation, the gene knockdown did not turn on. This could be due the fact that the siRNA was degraded in the endosome. Otherwise, when compared to previous gels analyses, the UV irradiation only partially cleaved off the streptavidin groups, and that could also be a reason for very minimal gene knockdown. Moreover, some of the UV irradiation is partially blocked by the cell tissue, so it might have reached the siRNA with much lower intensity than the one needed to uncage the siRNA. This might have only released a very small percentage of the uncaged siRNA, that combined to a low percentage of endosomal release might be the reason for such minimal gene knockdown.

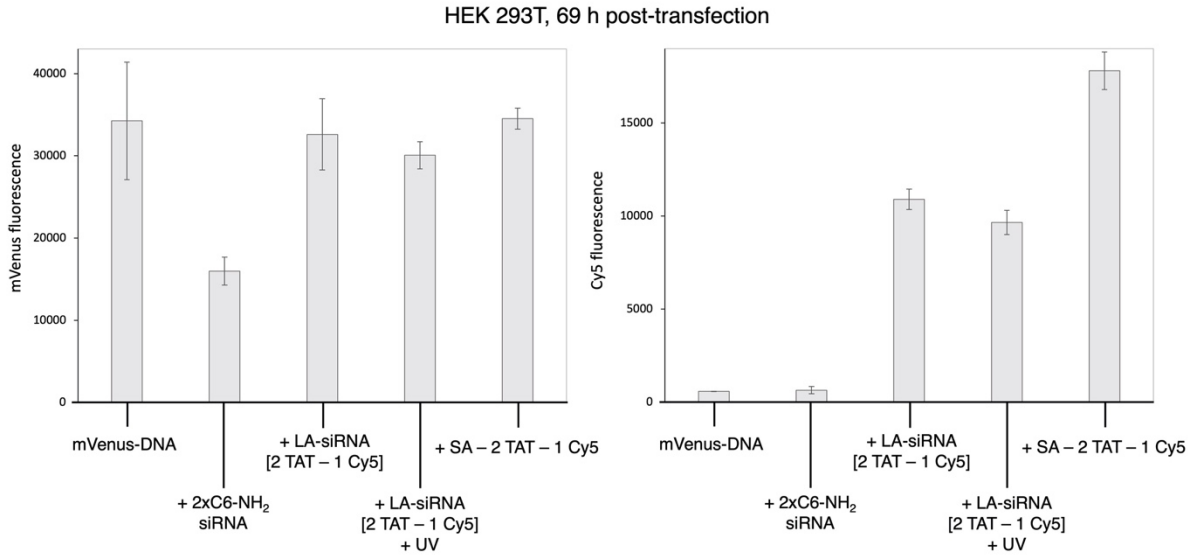


Figure 3.25: Delivery of LA-siRNA with off state

The two graphs represent fluorescence reading taken at two different fluorescent channels at the same time. On the left, mVenus knockdown is shown. On the right, cell internalization of the streptavidin-siRNA complex is shown. The streptavidin-siRNA complex did indeed internalize into the cell (Cy5 fluorescence levels are higher than the controls), but only caused minimal knockdown after UV irradiation, compared to the amino-siRNA control.

3.3 Conclusions

Small interfering RNAs are the most commonly used RNA interference therapeutics. Once in the cytoplasm, they activate the cell machinery to cause gene silencing by causing degradation of a target mRNA. New technologies are needed to expand siRNA applications. Amongst them, the ability of controlling siRNA in spatiotemporal manner would be a powerful tool for laboratory and clinical research. The delivery of siRNAs to mammalian cells is also confined to the use of transfecting reagents, often toxic, incompatible with the cellular environment, and that cannot easily be transferred to experiments in vivo.

The formation of UV light-activated siRNAs was developed in this work. Amino-siRNAs were modified with photocleavable–biotin groups via NHS ester, then bound to streptavidin moieties. Wild-type streptavidin is a tetramer and can therefore bind three more biotinylated products other than siRNA. This allows the inclusion of small molecules or other compounds with a specific function, which can then be easily removed by UV irradiation with the rest of the photocage. The siRNA sequence was modified with two different types of amines. Starting from commercial DNA oligonucleotides, aminoallyl modifications were included in the transcription reaction to produce RNA using aminoallyl-modified triphosphates compatible with common RNA polymerases. These triphosphates allow rapid and inexpensive insertion of a high number of modifications. Modifications of either strand of the siRNA were tested in mammalian cells. Modifications of the siRNA antisense strand had a negative impact on gene silencing efficiency, while modifications on the sense strand did not affect RNAi. However, with this method it was impossible to purify a single fully modified photocleavable-siRNA by HPLC, because of the poor reactivity of the aminoallyl groups.

A different siRNA was therefore produced using a similar method but different modifications. This time, the amino-modified sense strand was made synthetically by a collaborator and contained two C6-amino linkers. The modification with photocleavable moieties and subsequent purification of the modified-ssRNA and then siRNA were successful.

To allow self-delivery of the nucleic acid, cell-penetrating peptides were attached to wild-type streptavidin, then the protein bound to the biotinylated-siRNA to form UV-active photocages. The construct was characterized before and after UV irradiation *in vitro*. To trace cell delivery by fluorescence, biotinylated cyanine-5 was also added to streptavidin prior siRNA binding. Once added to tissue cultures, the siRNA-streptavidin construct was successfully delivered to multiple cell lines and visualized by epifluorescence microscopy.

4

Conclusion and future directions

Nucleic acid technologies have important applications as therapeutics and as synthetic biology tools. In medicine, they present a list of advantages compared to traditional small molecule drugs and protein therapeutics. As only a fraction of the human genome can be currently treated with these two more traditional methods, nucleic acids offer an alternative to potentially target any gene, with base-to-base precision. Ultimately, the goal would be to standardize a platform to produce and consistently deliver nucleic acids to any target organ, by just modifying their sequence. Any advance in this direction is leading towards more precise personalized medicine, in which a patient is treated based on a thorough profile evaluation, resulting in a personalized therapy.

On the other hand, synthetic biology is leading the technological progress of many fields, by implementing the design of newly engineered genetic circuits, organisms, and biomaterials [143]. At the core of this progress is the use of cell-free protein synthesis systems, which have become readily accessible. These systems have been used for various applications, among which the engineering of biosensors, genetic circuits, and drug screening. CFPS can be encapsulated into lipid bilayers to form synthetic cells, which in the future will constitute new methods for controlled drug delivery and the construction of artificial life [144].

In both medicine and synthetic biology, the ability to control gene expression externally represents an essential tool. This can be done by molecular stimuli, by using small activating ligands, or by physical stimuli, such as light, temperature, or magnetism. Even though the use of small molecules has had great applications in different fields, it requires the intervention on an open system, which can disrupt other reactions by altering the parameters, or cause toxicity. Physical stimuli on the other hand can be applied remotely to closed systems. Above all, light has been one of the most successfully used methods, as it has low toxicity, is bioorthogonal to most cellular processes, and can be applied with very high spatial and temporal precision.

The modification of nucleic acids with light-sensitive photocages allows for the control of gene expression with specific wavelengths. One way of altering gene expression is by using antisense oligonucleotides and small interfering RNAs for gene knockdown. In the work presented, these nucleic acids have been modified by reacting photocages to amino-modified sequences. The photocages are formed by a photocleavable functional group, either a 2-nitrobenzyl for UV light activation, or a coumarin for blue light activation. Then, as these functional groups contain a terminal biotin, streptavidin is attached to the other end. The streptavidin photocages create high steric hinderance and prevent the approach of

proteins to the nucleic acid. As UV and blue light have low tissue penetration and if applied at high intensity can cause cellular damage, the development of photocages activated at longer wavelengths is desirable. The advantage of the methods described to photocage nucleic acids is that any photocleavable scaffold can be synthesized with a functional group to react to modified DNA or RNA (e.g., NHS ester) and with a terminal biotin, both of which are standard features. Some photocages reactive to higher wavelengths are present in the literature. Amongst them, the green light-activated BODIPY derivatives responding to ~520 nm [145], and near-infrared cyanine photocages responding to ~690 nm (Figure 4.1) [146].

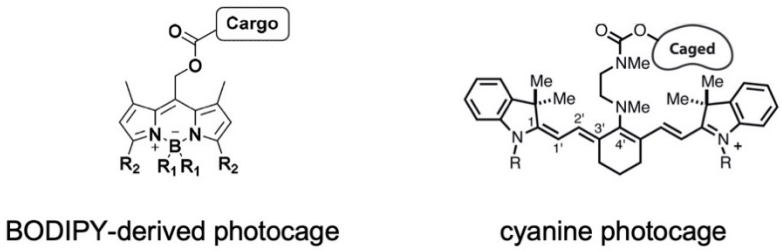


Figure 4.1: BODIPY-and cyanine-derived photocages
These structures can be used as photocages. They can photocleave at wavelengths higher than UV/blue, therefore penetrating cellular tissue further and causing lower toxicity.

After sequence optimization, antisense oligonucleotides responding to either UV or blue light have been generated with the photocaged streptavidin strategy described and have been used to control RNase H-mediated mRNA degradation in cell-free systems. CFPS are usually activated on demand with light and other methods, but strategies to turn off gene expression are lacking in the literature. Moreover, ASOs have been understudied in the field

of synthetic biology, even though they could have interesting applications. As the two UV and blue photocages used have demonstrated to be completely orthogonal [90], two different setups for dual-wavelength control of gene expression were designed. The first one used the uvLA-ASO and a previously generated bLA-DNA template, where a T7 promoter was caged with seven coumarin/biotin/streptavidin photocages. Here, DNA transcription was activated selectively and protein translation halted at different expression levels with UV. The other was to selectively target the mRNA of two different proteins with orthogonal ASOs. This resulted in the orthogonal control of gene knockdown of two genes in the same reaction mixture.

With this work in cell-free expression systems, the engineering of complex logic gates could be achieved. Moreover, the CFPS mixtures could be encapsulated into lipid bilayers to form synthetic cells. By setting up light irradiation at different orientations, the expression in synthetic cells could be patterned for higher spatiotemporal control. This could be coupled to mammalian cells or bacteria cultures to induce a specific response only in a determined area, elevating further the reachable level of complexity. A future application of such systems that could be envisioned is the light activated production of a nucleic acid drug *in situ* with one wavelength, then the deactivation with the other, after a certain feedback from the organism. Finally, this technology could also be applied to tissue cultures and *in vivo*. For this, ASO sequences would be modified with phosphorothioates, 2'-modifications, or other methods to increase the stability of the ASO.

The photocaged streptavidin strategy was also implemented in siRNAs. Two methods producing different types of uvLA-siRNA modifications have been shown, with observations on the sequence, the photocleavage, HPLC purification, and gene knockdown efficiency. The tetrameric nature of wild-type streptavidin was exploited to decorate the photocage with cell-penetrating peptides and a fluorophore for tracing. This has shown

successful delivery of the siRNA to the cell, without the aid of any transfection reagents. However, UV irradiation did not result in an activation of gene knockdown. By increasing the UV irradiation, the desired gene knockdown might be achieved, although a precise optimization of UV conditions would be required to avoid damaging the cells. Another way of achieving light-activation would be to substitute the UV photocage with a blue light one, which was already used to control ASOs and would be easily adapted to siRNA.

Nonetheless, with this system, the delivery of a nucleic acid-protein conjugate was carried out. It was impossible to understand if the conjugate was delivered to the cytoplasm or not, but this could be established with further engineering. One way to analyze delivery to the cytoplasm would be to substitute the photocleavable groups with different labile groups, such as disulfides, which could be cleaved by glutathione in the cellular environment. If successful, this method could become a great control for delivery to run in parallel to other nucleic acids/proteins. Streptavidin could also be bound to different and more effective cell-penetrating peptides [127, 128]. Moreover, by observing formation of the endosome and endosomal escape, small molecule sensors could be added to better understand the phenomenon [147].

As changing the photocages could expand the applications of this technology, so it would be changing the streptavidin moiety. Surface receptors unique to a cell-type could be selectively targeted by different molecules, such as antibodies, or peptides, which could attached to a nucleic acid by substituting the streptavidin unit. These could be attached to a photocage to aid cellular delivery, then cleaved off by light irradiation, without interfering with the nucleic acid activity.

5

Methods

5.1 Materials

All solvents and reagents were purchased from Sigma/Merck unless stated otherwise. NHS-PC-biotin was purchased from Click Chemistry Tools and stored in dry DMF in aliquots at -80 °C. PURExpress was purchased from New England Biolabs (NEB). All enzymes were purchased from NEB, unless stated otherwise. Standard DNA oligonucleotides were synthesized by Integrated DNA Technologies (IDT). Amino-modified DNA oligonucleotides were synthesized by ATDBio. Monovalent streptavidin was a gift from

Prof. Mark Howarth (University of Oxford). 2xC6-NH_2-ssRNA was kindly synthesized and gifted by Dr. Afaf El-Sagheer (University of Oxford).

5.2 General methods

The following methods were used consistently throughout each of the chapters. They are written in full here and then referred in each respective chapter when used. Gel electrophoresis data was analyzed using ImageLab software.

5.2.1 Agarose gel electrophoresis

Agarose gels were prepared at 1.5% agarose in 1X TBE buffer and 1X Gel-Red® nucleic acid stain (Biotium), then run at 100 V in 1X TBE buffer. Samples were prepared using RNA loading dye (NEB, B0363S), heated to 70 °C for 10 minutes then cooled on ice before loading to denature the RNA. Samples were run against a low-range ssRNA ladder (NEB, N0364S) or ssRNA ladder (NEB, N0362S).

5.2.2 Polyacrylamide gel electrophoresis

Gels were cast by hand with the BioRad Mini-PROTEAN® handcasting equipment at a thickness of 0.75 mm. Short single stranded nucleic acids were run in denaturing polyacrylamide gels. Denaturing gels were made with 16-20% PAA/BisAA (19:1), 7 M Urea, 1X TBE buffer, polymerised using 0.8% APS and 0.05% TEMED, then run at 250 V

in 1X TBE buffer in a cold environment. Samples were loaded with 95% formamide, 0.015% SDS, 5 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol. Gels were stained with Gel-Red® nucleic acid stain (Biotium) following the manufacturer's protocol. Short double stranded nucleic acids were run in native polyacrylamide gels. Native gels were made with 16% PAA/BisAA (19:1), 1X TBE buffer, polymerised using 0.8% APS and 0.05% TEMED, then run at 150 V in 1X TBE buffer. Samples were loaded with 10% glycerol, 5 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol loading buffers. Gels were finally stained with Gel-Red® nucleic acid stain (Biotium) following the manufacturer's protocol.

5.2.3 Oligonucleotide mass spectrometry

Oligonucleotide mass spectra were recorded on a Waters Xevo G2 QTOF ESI- UPLC-MS system. A gradient of MeOH in Et_3N and hexafluoroisopropanol (HFIP) was used (buffer A, 8.6 mM Et_3N, 200 mM HFIP in 5% MeOH/H_2O (v/v); buffer B, 20% buffer A in MeOH). Samples were analysed using a gradient of 0 to 70% buffer B over 8 minutes and 0% B to 10 minutes and data was then deconvoluted using MassLynx v4.1.

5.3 Methods Chapter 2

Experiments containing blue light-activatable biotin groups were performed under reduced laboratory lighting conditions (direct overhead lights turned off).

5.3.1 Formation of biotinylated UV-photocleavable antisense oligonucleotides

Amino-modified oligonucleotides (20 μM) were added to a 0.5 mL DNA LoBind tube (Eppendorf) and reacted with PC Biotin-NHS ester (5 mM) with 100 mM NaHCO3, in a total volume of 50 μL (50% water 50% DMF). Reactions were incubated in the dark at room temperature for 1 hour, with frequent mixing and centrifugation, then left at 4 °C overnight. Reactions were quenched with 350 μL of 20 mM tris pH 7. The mixtures were then partially desalted by centrifugation with Amicon Ultra-0.5 mL centrifugal filters, according to the manufacturer's protocol, then purified further by HPLC.

5.3.2 First HPLC purification method of modified antisense oligonucleotides

The modified oligonucleotides were purified by HPLC on a semi-preparative C-18 column (Vydac, 25 cm x 10 mm), at room temperature, using a gradient of 5-35% MeCN over 40 minutes with 10 mM NH_4OAc pH 7 throughout. The resulting oligonucleotides were lyophilised, resuspended in H_2O and analysed by LCMS for purity (details on LCMS below).

5.3.3 Transcription of mVenus and mCherry mRNA

In vitro transcription was performed using the linear DNA templates prepared, with the HiScribe™ T7 High Yield RNA Synthesis Kit (NEB), typically at 37 °C for 4 hours, following the manufacturer's protocol. The reactions were then diluted by 10-fold with water and incubated in the presence of DNase I (ThermoFisher) for 15 minutes at 37 °C, to remove the DNA template. The resulting RNA was then purified using the GeneJet RNA

Clean-up and Concentration Kit (ThermoFisher) following the manufacturer's protocol and eluted in H_2O. Prior to use, mRNA was tested for size and purity against a commercial ssRNA size marker.

5.3.4 RNase H-mediated mRNA degradation

1 pmol of mVenus mRNA was incubated with 6 U of RNase H (recombinant E. coli, Takara) and 0.2 pmol of antisense oligonucleotide (ASO) in a buffer system containing 30 mM HEPES pH 7, 100 mM KCl, 20 mM MgCl2 and 2 mM DTT. The samples were incubated at 37 °C for 1 hour. After incubation, RNA loading dye (NEB, B0363S) was added, the samples heated to 70 °C for 10 minutes and analysed by agarose gel electrophoresis.

5.3.5 Binding of streptavidin

The biotinylated antisense oligonucleotides were incubated with monovalent streptavidin or tetravalent streptavidin (NEB) in a 4-fold excess to the number of photocleavable biotin groups. The reactions were incubated in the dark for 2-3 hours at room temperature, then left at 4 °C overnight.

5.3.6 UV irradiation conditions

Samples in an open 200 μL PCR tube were held in a PCR tube rack (StarLabs) over aluminium foil and irradiated top-down. Irradiation was performed with a ThorLabs 365 nm LED (M365L3) equipped with a collimator (COP5-A) from a distance of 34 cm at an irradiance of 2.12 mW·cm^{-2}, controlled by a ThorLabs driver (LEDD1B) set at 1 A maximum drive current. The power was measured at the position of the sample using the ThorLabs

analog handheld laser power meter console PM100A, using a photodiode power sensor (Thorlabs S120VC), which has a measuring surface area of 0.7088 cm^2 (Ø of aperture: 0.95 cm).

5.3.7 Second HPLC purification of modified antisense oligonucleotides

The modified oligonucleotides were purified by HPLC on an Agilent Polaris C-18 column (150 x 4.5 mm), heated to 50 °C, using a gradient of 5-28% MeCN over 36 minutes with 10 mM TEAB pH 8.5 throughout. The resulting oligonucleotides were lyophilised, resuspended in H_2O and analysed by LCMS for purity.

5.3.8 Preparation of blue light-activatable antisense oligonucleotides

The amine-modified oligonucleotide (Entry 19, 23.5 μM) was added to a 1.5 mL DNA LoBind Tube (Eppendorf) and reacted with bLA-Biotin PFP Carbonate (as prepared previously [90], 17.5 mM) with MOPS pH 8.5 (119 mM) in 59% DMF in a total reaction volume of 34 μL. The reaction was incubated in a Thermomixer (Eppendorf) at 37 °C and 800 RPM overnight. The solution was diluted to 50 μL with H_2O, washed with $CHCl_3$ (3 x 200 μL), 50 μL DMSO added, transferred to a HPLC injection vial and purified by HPLC.

5.3.9 Preparation of linear template DNA

PCR reactions were carried out using DreamTaq DNA polymerase MasterMix (2X, ThermoFisher), forward and reverse primers at 0.25 μM concentration and 0.04 ng/μL of the Hind III-digested plasmid as template, in a total reaction volume of 25 μL. The PCR was carried out according to the manufacturer's protocol for 35 cycles with an annealing

temperature of 52 °C for 30 seconds, an extension time of 72 °C for 1 minute 15 seconds/kbp and a final extension at 72 °C for 10 minutes. The resulting DNA was then purified using the GeneJet PCR purification columns (ThermoFisher) following the manufacturer's protocol.

5.3.10 Blue light irradiation conditions

Samples in an open 200 μL PCR tube were held in a PCR tube rack (StarLabs) over aluminium foil and irradiated top-down. Irradiation was performed with a ThorLabs 455 nm LED (M455L4) equipped with a collimator (COP4-A) from a distance of 30 cm at an irradiance of 64 $mW \cdot cm^{-2}$ for 1 minute. The power was measured at the position of the sample using the ThorLabs analog handheld laser power meter console PM100A, using a photodiode power sensor (Thorlabs S120VC), which has a measuring surface area of 0.7088 cm^2 (Ø of aperture: 0.95 cm).

5.3.11 In vitro transcription control using uvLA-V3

Reagents from the HiScribe™ T7 High Yield RNA Synthesis kit (NEB) were combined in a PCR tube as a master mix considering 3 µL final volume per reaction condition. Each 3 µL reaction contained the following: 0.05 µL T7 RNA Polymerase, 1X T7 RNA Polymerase buffer, and 10 mM NTPs, to which 0.001 pmol of mVenus linear DNA template, 3 U of RNase H (recombinant *E. coli*, Takara) and 0.05 pmol of caged ASO uvLA-**V3** were also added. The reactions were incubated in the dark at 37 °C for 3 hours. At this point, a 3 µL aliquot was transferred to a different tube and irradiated with UV light (conditions above) for 5 minutes, while 0.05 pmol of amino-ASO **V3** were added to a separate 3 µL aliquot as control, before all tubes were put back at 37 °C and incubated for 1 more hour. Each reaction

was finally quenched with 0.5 μL of 100 mM EDTA, RNA loading dye (NEB, B0363S) was added, the samples heated to 70 °C for 10 minutes and analysed by agarose gel electrophoresis.

5.3.12 Two-wavelength control of in vitro transcription

Reagents from the HiScribe™ T7 High Yield RNA Synthesis kit (NEB) were combined in a PCR tube as a master mix considering 3 μL final volume per reaction condition. Each 3 μL reaction contained the following: 0.05 μL T7 RNA Polymerase, 1X T7 RNA Polymerase buffer, 10 mM NTPs, to which 0.001 pmol of bLA-mVenus DNA, 3 U of RNase H (recombinant *E. Coli*, Takara) and 0.05 pmol of caged ASO uvLA-**V3** were also added. The reaction was started by irradiating the tube with blue light (conditions above) for 1 minute, followed by incubation in the dark at 37 °C. Every hour 3 μL were removed from the master mix reaction tube and quenched with 0.5 μL of 100 mM EDTA. After 3 hours, one 3 μL was separately irradiated with UV light (conditions above) for 5 minutes, then put back at 37 °C. All reactions were quenched with 0.5 μL of 100 mM EDTA after 1 more hour (4 hours total). The reaction controls included one reaction where no light was applied, and one reaction where only UV light was applied at the start. In parallel, an equivalent set of reactions were run with non-caged nucleic acids as control. A linear mV DNA template was added to start the reaction and the amino-ASO **V3** was added to initiate RNase H-mediated degradation. All reactions were finally mixed with RNA loading dye (NEB, B0363S), heated to 70 °C for 10 minutes and analysed by agarose gel electrophoresis.

5.3.13 Light-controlled gene knockdown in cell-free protein synthesis with uvLA-V3

In a 200 μL PCR tube, 5 ng/μl linear mV template DNA was added to PURExpress® (NEB, E6800) with 0.67 U/μL RNase H (Takara) and 0.2 ng/μL NH_2-**V3** or uvLA-**V3** ASO. The resulting solutions were kept at room temperature in the dark and illuminated with UV as required (5 minutes), before placing them in a thermocycler and incubated at 37 °C for 4 hours. 2 μL of each solution was then placed into 39 μL of H_2O and mixed by pipetting. 40 μL of the resulting solutions were then transferred into a 384 well plate and placed into a plate reader (Tecan Infinity M1000) and fluorescence measurements were taken ($\lambda_{Ex/Em}$: 515/527 nm, Gain 173).

5.3.14 Control of transcription and gene knockdown with two wavelengths in cell-free protein synthesis

In a 200 μL PCR tube, 5 ng/ul of blue light-activatable mV template DNA (as prepared previously)[90] were added to PURExpress® (NEB, E6800) with 0.67 U/μL RNase H (Takara) and 0.33 ng/μL uvLA-**V3** and 0.2 ng/μL NH_2-**V3** or uvLA-**V3** ASO. The resulting solutions were kept at room temperature in the dark and illuminated as required (1 minute with blue and 3 mins with UV), before placing them in a thermocycler and incubated at 37 °C. To illuminate at different timepoints, tubes were cooled to room temperature in the dark and illuminated as required, before placing them back into the thermocycler for a total of 4 hours incubation time. 2 μL of each solution were then placed into 39 μL of H_2O and mixed by pipetting. 40 μL of the resulting solutions were then transferred into a 384 well plate and

placed into a plate reader (Tecan Infinity M1000) and fluorescence measurements were taken ($\lambda_{Ex/Em}$: 515/527 nm, Gain 173).

5.3.15 Two-wavelength control of RNase H-mediated mRNA degradation

1 pmol of either mV or mC mRNA was incubated with 6 U of RNase H (recombinant E. coli, Takara) and 1.2 ng of each bLA-C1 and uvLA-V3 in a buffer system containing 30 mM HEPES pH 7, 100 mM KCl, 20 mM MgCl2 and 2 mM DTT. The samples were illuminated as required (5 minutes for UV irradiation or 1 minute for blue irradiation), incubated at 37 °C for 1 hour. After incubation, RNA loading dye (NEB, B0363S) was added, the samples heated to 70 °C for 10 minutes and analysed by agarose gel electrophoresis.

5.3.16 Two wavelength-controlled gene knockdown in cell-free protein synthesis

In a 200 μL PCR tube, 5 ng/μl each of linear mV and mC DNA (as prepared were added to PURExpress® (NEB, E6800) with 0.67 U/μL RNase H (Takara), 0.5 ng/μL bLA-C1 ASO and 0.33 ng/μL uvLA-V2 ASO. The resulting solutions were kept at room temperature in the dark and illuminated as required (3 minutes for UV irradiation or 1 minute for blue irradiation), before placing them in a thermocycler at 37 °C for 4 hours. 2 μL of each solution were then placed into 39 μL of H2O and mixed by pipetting. 40 μL of the resulting solutions were then transferred into a 384 well plate and placed into a plate reader (Tecan Infinity

M1000) and fluorescence measurements were taken (λEx/Em: 515/527 nm for mV, gain 183, and λEx/Em: 587/610 nm for mCherry, gain 234).

5.4 Methods Chapter 3

All mammalian cells were cultured at 37 °C in a 5% CO_2 atmosphere in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum, 1X GlutaMAX and 1X MEM non-essential amino acid solution. If otherwise, it will be stated in the text.

5.4.1 Preparation of modified siRNA

The two DNA oligonucleotides were mixed in equal molar amount, denatured for 1 minute at 90 °C, then slowly cooled to room temperature over 1 hour. In vitro transcription was performed using the linear DNA templates prepared with the HiScribe™ T7 High Yield RNA Synthesis Kit (NEB), typically at 37 °C for 16 hours, following the manufacturer's protocol for short transcripts. When aminoallyl-UTP was needed, it was added at 1.5X concentration of the other triphosphates. The resulting RNA was purified by ethanol precipitation at -20 °C overnight. The two purified ssRNA strands were mixed in equal molar amount and annealed for 1 hour RNase H buffer (30 mM HEPES pH 7, 2 mM $MgCl_2$ 0.1 M KCl, 2 mM DTT). The UHO oligonucleotide was added in 1.5X molar excess of the dsRNA, RNase H was also added and the mixture incubated for 1 hour 30 minutes at 37 °C. The reactions were incubated with DNase I (ThermoFisher) for 15 minutes at 37 °C, to

remove the DNA template, then purified with ethanol precipitation. The resulting siRNA was analyzed by native polyacrylamide electrophoresis.

5.4.2 Binding of streptavidin

The biotinylated siRNA was incubated with monovalent streptavidin or tetravalent streptavidin (NEB) in a 4-fold excess to the number of photocleavable biotin groups. The reactions were incubated in the dark for 2-3 hours at room temperature, then left at 4 °C overnight. When biotin-TAT (AnaSpec) or biotin-Cy5 (AAT) were added, streptavidin was incubated with them for 1 hour at room temperature before being added to siRNA.

5.4.3 HEK 293 T transfection

HEK cells were culture with the conditions described, but with FluoroBrite™ DMEM (Thermo) for lower epifluorescence microscope background. When 80-90% confluent they were seeded at a density of 25,000 cells per well in a 96-well plate and grown overnight prior transfection. 100 ng of mVenus-N1 plasmid, different concentration of siRNA (typically 20 nM as final concentration in the well), 0.3 μL of Lipofectamine 3000 (Thermo) and 0.2 μL enhancer lipid P3000 were added to serum-free medium OptiMEM and incubated at room temperature for 15 minutes, as for manufacturer protocol. The mixture was then simply transferred to the 96-well plate. The cells were analyzed by epifluorescence microscopy and plate reader, as described below.

5.4.4 3T3-GFP transfection

Different amounts of siRNA (typically 20 nM as final concentration in the well) and 0.2 μL of Lipofectamine RNAiMAX (Thermo) were mixed in serum-free medium OptiMEM and incubated at room temperature for 15 minutes, as for manufacturer protocol. The mixture was added to a 96-well plate. 3T3 cells were seeded in the plate at 15,000 cells per well and the volume diluted with OptiMEM. The cells were placed in the incubator at 37 °C for 6 hours. At this point, the transfecting medium was replaced with full DMEM medium and the cells were placed back in the incubator. After 48-72 hours the cells were analyzed by epifluorescent microscopy or plate reader, as described below.

5.4.5 Epifluorescence microscopy

Samples held inside imaging chambers were imaged with a Lecia DMi8 inverted epifluorescence microscope. Brightfield, TXR, and GFP filter settings were used.

5.4.6 UV irradiation conditions

Samples in a 96-well plate with opened lid were illuminated over aluminium foil and top-down. Irradiation was performed with a ThorLabs 365 nm LED (M365L3) equipped with a collimator (COP5-A) from a distance of 34 cm at an irradiance of 1.75 mW·cm^{-2}, controlled by a ThorLabs driver (LEDD1B) set at 1 A maximum drive current. The power was measured at the position of the sample using the ThorLabs analog handheld laser power meter console PM100A, using a photodiode power sensor (Thorlabs S120VC), which has a measuring surface area of 0.7088 cm^2 (Ø of aperture: 0.95 cm).

Appendix A

A.1 DNA sequences

Gene	Sequence (5' → 3')
mVenus	GAAATTAATACGACTCACTATAGGGTCTAGAAATAATTTTGTTTAACTTTAAGAA GGAGGTATACAT**ATG**GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTCA TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCT CGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCA CGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATC TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGG GCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGT CTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATC CGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAG AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGA GCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGT CCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCT GTACAAG**TAA**TGAGGATCCCGGGAATTCTCGAGTAAGGTTAACCTGCAGGAG GCCTTTAATTAAGGTGGTGCGGCCGCGCTAGCGGTCCCGGGGGATCGATCC GGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAG CAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTT GCTGAAAGGAGGAACTATATC
mCherry	GAAATTAATACGACTCACTATAGGGTCTAGAAATAATTTTGTTTAACTTTAAGAA GGAGGTATACAT**ATG**GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCA AGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGA GTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGA CCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACA TCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGC CGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAG CGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCC TCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACT TCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCT CCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGC AGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCA CCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACAT CAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTAC GAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAG **TAA**TGAGGATCCCGGGAATTCTCGAGTAAGGTTAACCTGCAGGAGGCCTTTA ATTAAGGTGGTGCGGCCGCGCTAGCGGTCCCGGGGGATCGATCCGGCTGCT AACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAAC TAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAA GGAGGAACTATATC

A.2 mRNA sequences

Gene	Sequence (5' → 3')
mVenus	GGGUCUAGAAAUAAUUUUGUUUAACUUUAAGAAGGAGGUAUACAU**AUG**GUG AGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCU GGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGCGAGG GCGAUGCCACCUACGGCAAGCUGACCCUGAAGCUCAUCUGCACCACCGGCA AGCUGCCCGUGCCCUGGCCCACCCUCGUGACCACCCUCGGCUACGGCCUG CAGUGCUUCGCCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUCAAG UCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUCAAGGA CGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCC UGGUGAACCGCAUCGAGCUGAAGGGCAUCGACUUCAAGGAGGACGGCAAC AUCCUGGGGCACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUC ACCGCCGACAAGCAGAAGAACGGCAUCAAGGCCAACUUCAAGAUCCGCCAC AACAUCGAGGACGGCGGCGUGCAGCUCGCCGACCACUACCAGCAGAACACC CCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCUA CCAGUCCAAGCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCU GCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUGU ACAAGUAAUGAGGAUCCCGGGAAUUCUCGAGUAAGGUUAACCUGCAGGAGG CCUUUAAUUAAGGUGGUGCGGCCGCGCUAGCGGUCCCGGGGGAUCGAUCC GGCUGCUAACAAAGCCCGAAAGGAAGCUGAGUUGGCUGCUGCCACCGCUG AGCAAUAACUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGU UUUUUG
mCherry	GGGUCUAGAAAUAAUUUUGUUUAACUUUAAGAAGGAGGUAUACAU**AUG**GUG AGCAAGGGCGAGGAGGAUAACAUGGCCAUCAUCAAGGAGUUCAUGCGCUU CAAGGUGCACAUGGAGGGCUCCGUGAACGGCCACGAGUUCGAGAUCGAGG GCGAGGGCGAGGGCCGCCCCUACGAGGGCACCCAGACCGCCAAGCUGAAG GUGACCAAGGGUGGCCCCCUGCCCUUCGCCUGGGACAUCCUGUCCCCUCA GUUCAUGUACGGCUCCAAGGCCUACGUGAAGCACCCCGCCGACAUCCCCG ACUACUUGAAGCUGUCCUUCCCCGAGGGCUUCAAGUGGGAGCGCGUGAUG AACUUCGAGGACGGCGGCGUGGUGACCGUGACCCAGGACUCCUCCCUGCA GGACGGCGAGUUCAUCUACAAGGUGAAGCUGCGCGGCACCAACUUCCCCU CCGACGGCCCCGUAAUGCAGAAGAAGACCAUGGGCUGGGAGGCCUCCUCC GAGCGGAUGUACCCCGAGGACGGCGCCCUGAAGGGCGAGAUCAAGCAGAG GCUGAAGCUGAAGGACGGCGGCCACUACGACGCUGAGGUCAAGACCACCU ACAAGGCCAAGAAGCCCGUGCAGCUGCCCGGCGCCUACAACGUCAACAUCA AGUUGGACAUCACCUCCCACAACGAGGACUACACCAUCGUGGAACAGUACG AACGCGCCGAGGGCCGCCACUCCACCGGCGGCAUGGACGAGCUGUACAAG UAAUGAGGAUCCCGGGAAUUCUCGAGUAAGGUUAACCUGCAGGAGGCCUU UAAUUAAGGUGGUGCGGCCGCGCUAGCGGUCCCGGGGGAUCGAUCCGGCU GCUAACAAAGCCCGAAAGGAAGCUGAGUUGGCUGCUGCCACCGCUGAGCAA UAACUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGUUUUUU G

A.3 Sequences of oligonucleotides used in PCR

Name	Sequence (5' → 3')
T7 Forward primer	GAAATTAATACGACTCACTATAGGGTCTAG
Reverse primer	GATATAGTTCCTCCTTTCAG

A.4 Antisense oligonucleotide sequences

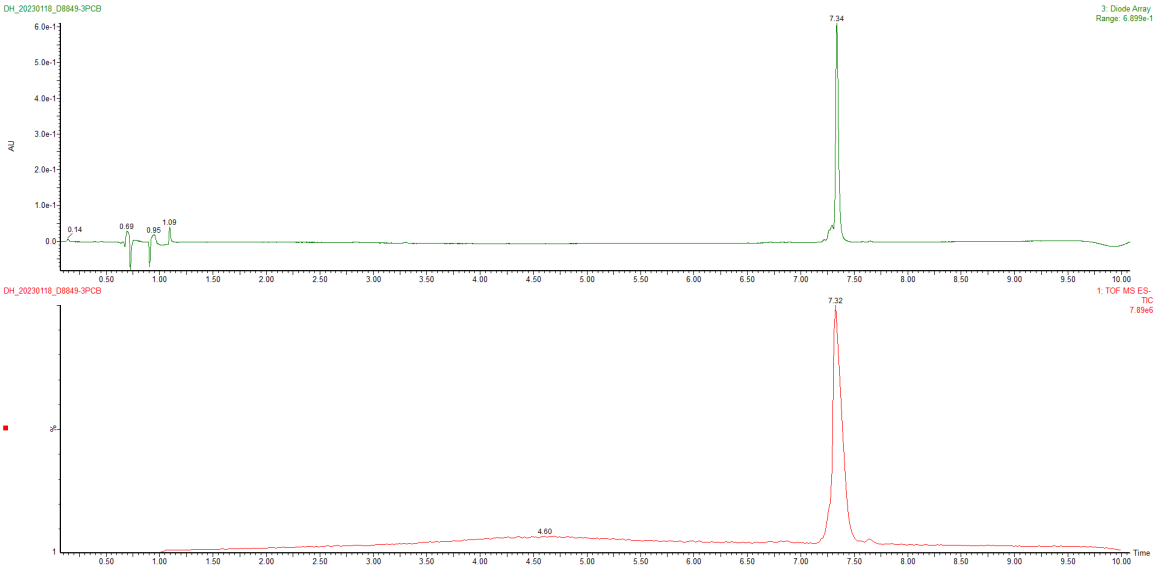
Entry	Target	Nucleotide position on mRNA sequence	Length in number of bases	Number of amines	Sequence (5' → 3')	
1	mV	220	20	0	TTCATGTGGTCGGGGTAGCG	→ V1
2	mV	220	20	3	TXCATGTGGXCGGGGXAGCG	→ V1b
3	mV	220	20	3	2XCAXGTGGTCGGGGTAGCG	→ V1c
4	mV	220	20	3	2XCAXGTGGXCGGGGTAGCG	→ V2
5	mV	220	20	4	TXCATGXGGXCGGGGXAGCG	
6	mV	229	18	0	GTGCTGCTTCATGTGGTC	→ V3
7	mV	229	18	4	GXGCTGCXTCAXGTGGXC	
8	mV	224	16	0	TTCATGTGGTCGGGGT	
9	mV	229	17	0	TGCTGCTTCATGTGGTC	
10	mV	230	16	0	TGCTGCTTCATGTGGT	
11	mC	27	19	0	CTCCTTGATGATGGCCATG	
12	mC	81	19	0	CTCGAACTCGTGGCCGTTC	
13	mC	139	18	0	CTTCAGCTTGGCGGTCTG	
14	mC	147	19	0	CTTGGTCACCTTCAGCTTG	
15	mC	222	16	0	CTTCACGTAGGCCTTG	
16	mC	303	16	0	GTCCTCGAAGTTCATC	
17	mC	141	16	0	CTTCAGCTTGGCGGTC	
18	mC	139	16	0	TCAGCTTGGCGGTCTG	→ C1
19	mC	139	16	4	XCAGCTXGGCGGXCXG	
20	mC	140	15	0	TCAGCTTGGCGGTCT	
21	mC	135	22	0	CTTCAGCTTGGCGGTCTGGGTG	
22	mC	139	23	0	GTCACCTTCAGCTTGGCGGTCTG	
23	mC	141	21	0	GTCACCTTCAGCTTGGCGGTC	
24	mC	142	20	0	GTCACCTTCAGCTTGGCGGT	

Appendix B

B.1 Modified antisense oligonucleotides LC-MS spectra

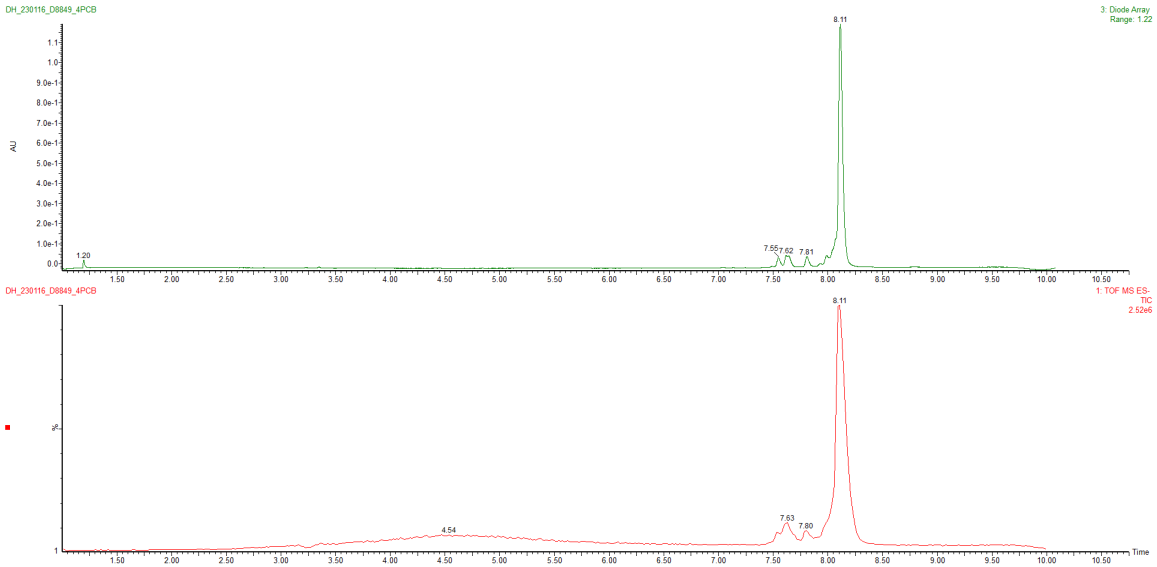
Oligonucleotide	Number of PC-biotin groups	Purified product	Expected mass	Mass after deconvolution
V1	3	uvLA-**V1** biotin	9149	9148
V2	4	uvLA-**V2** biotin	9740	9740
V3	4	uvLA-**V3** biotin	9032	9032
C1	4	bLA-**C1** biotin	8320	8320

B.2 LC-MS spectra uvLA-V1 biotin



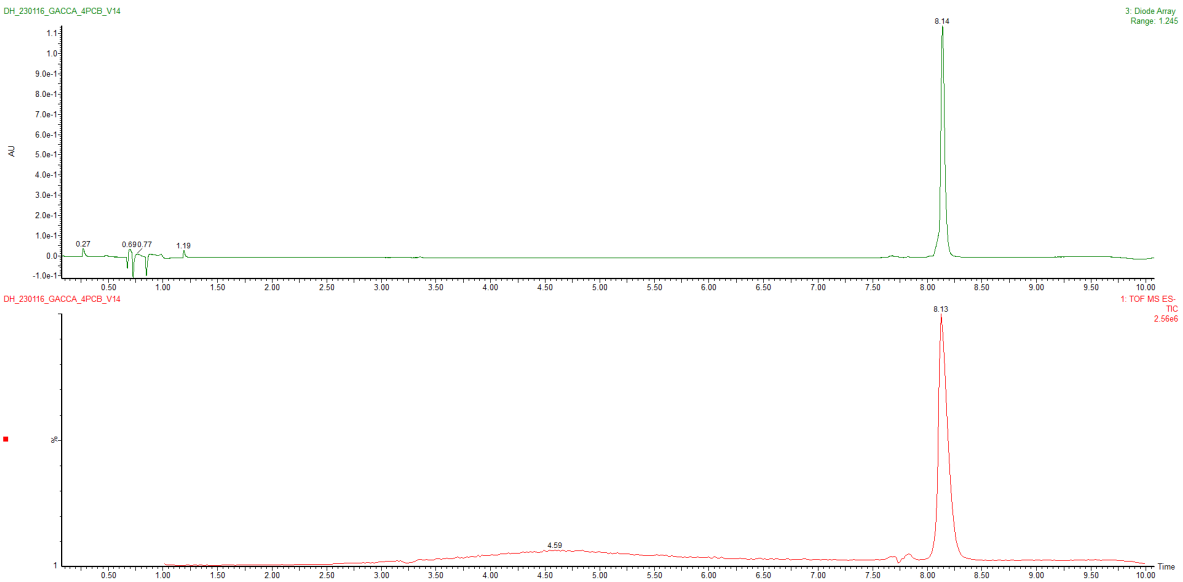


B.3 LC-MS spectra uvLA-V2 biotin



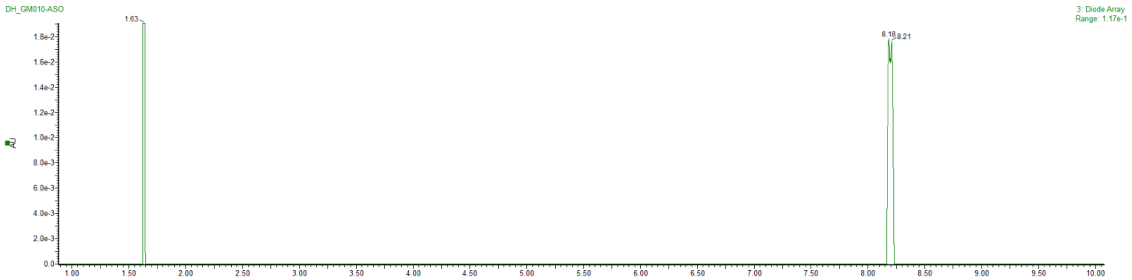


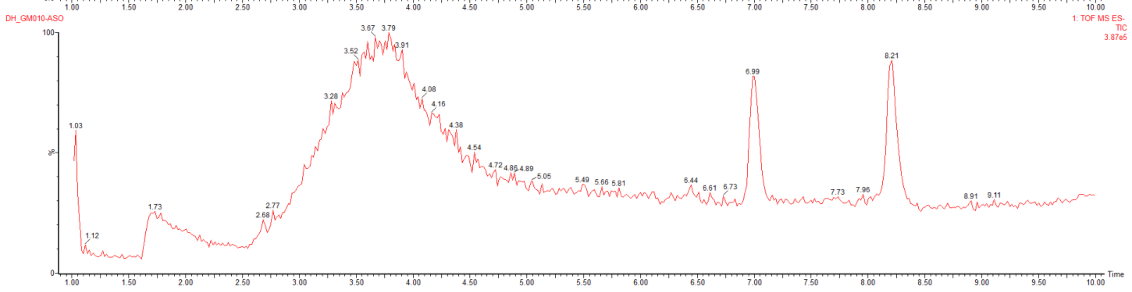
B4. LC-MS spectra uvLA-V3 biotin

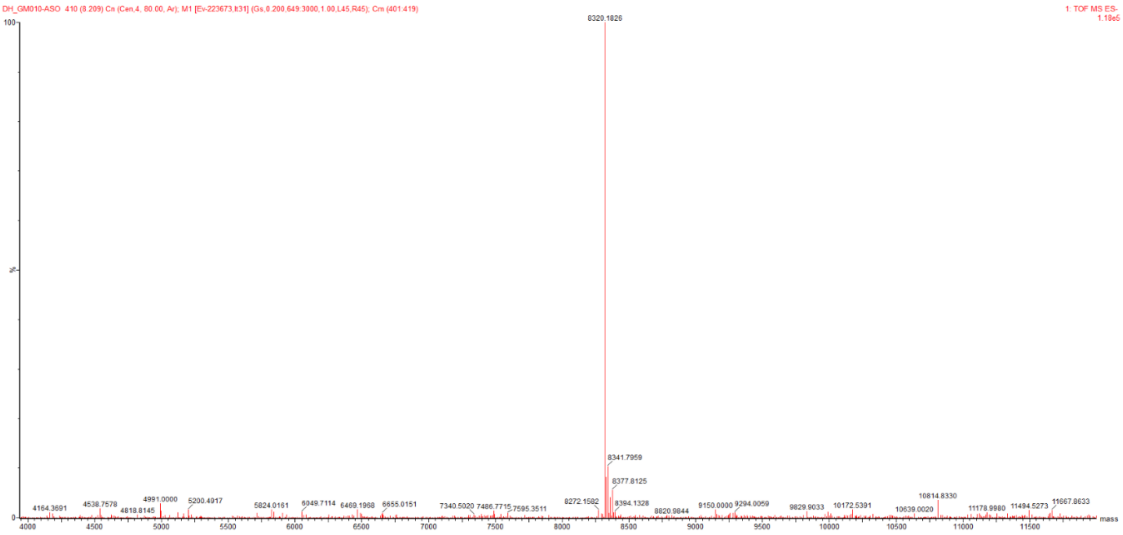



DH_230116_GACCA_4PCB_V14 405 (8.125) M1 [Ev+237936,tt42] (Gs,0.330,772.3000,1.00,LS,RS); Cm (400:417)
1: TOF MS ES-
6.56e5
100
%
0
9032.0000
9088.0000
6600
6800
7000
7200
7400
7600
7800
8000
8200
8400
8600
8800
9000
9200
9400
9600
9800
10000
10200
10400
10600
10800
11000
11200
11400
11600
11800
mass

B5. LC-MS spectra bLA-C1 biotin







Bibliography

[1] R. Altmann. “Ueber Nucleinsäuren.”. *Arch. f. Anat. u. Physiol* **1889**, 524–536.

[2] R. Dahm. “Friedrich Miescher and the discovery of DNA”. *Dev. Biol.* **2005**, *278* (2), 274–288. doi: 10.1016/j.ydbio.2004.11.028.

[3] B. Y. F. Griffith. “The Significance of pneumococcal types. Occurrence of a Variety of Serological Types in the Sputum from an individual case of pneumonia”. *J. Hyg. (Lond).* **1928**, *27* (2), 113–159.

[4] B. H. Judd. “Nucleic Acids as Genetic Material”. *eLS* **2001**. doi: 10.1038/npg.els.0000807.

[5] O. Avery, C. MacLeod, and M. Mccarty. “Studies on the chemical nature of the substance inducing transformation of Pneumococcal types”. *J. Exp. Med.* **1943**, *79* (6).

[6] M. Cobb. “60 years ago, Francis Crick changed the logic of biology”. *PLoS Biol.* **2017**, *15* (9). doi: 10.1371/journal.pbio.2003243.

[7] M. Cobb. “The lecture that changed biology,” *BBC News*, 2017. https://www.bbc.co.uk/news/science-environment-41311087

[8] R. Hogrefe. “A Short History of Oligonucleotide Synthesis.” https://www.trilinkbiotech.com/a-short-history-of-oligonucleotide-synthesis

[9] R. K. Saiki *et al*. “Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia”. *Science* **1985**, *230*1350–1354. doi: 10.1126/science.2999980.

[10] T. Friedmann and R. Roblin. “Gene Therapy for Human”. *Science* **1972**, *175* (4025), 949–955.

[11] M. L. Stephenson and P. C. Zamecnik. “Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide”. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75* (1), 285–288. doi: 10.1073/pnas.75.1.285.

[12] P. C. Zamecnik and M. L. Stephenson. “Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide”. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75* (1), 280–284. doi: 10.1073/pnas.75.1.280.

[13] J. A. Kulkarni *et al*. “The current landscape of nucleic acid therapeutics”. *Nat. Nanotechnol.* **2021**, *16* (6), 630–643. doi: 10.1038/s41565-021-00898-0.

[14] E. Pennisi. “Upstart DNA sequencers could be a ‘game changer’”. *Science* **2022**, *376* (6599), 1257–1258. doi: 10.1126/science.add4867.

[15] L. H. Goetz and N. J. Schork. “Personalized medicine: motivation, challenges, and progress”. *Fertil. Steril.* **2018**, *109* (6), 952–963. doi: 10.1016/j.fertnstert.2018.05.006.

[16] Z. E. Sauna *et al*. “Recent advances in (therapeutic protein) drug development”. *F1000Research* **2017**, *6*. doi: 10.12688/f1000research.9970.1.

[17] T. R. Damase, R. Sukhovershin, C. Boada, F. Taraballi, R. I. Pettigrew, and J. P. Cooke. "The Limitless Future of RNA Therapeutics". *Front. Bioeng. Biotechnol.* **2021**, *9* (March), 1–24. doi: 10.3389/fbioe.2021.628137.

[18] A. L. Hopkins and C. R. Groom. "The druggable genome". *Nat. Rev. Drug Discov.* **2002**, *1* (9), 727–730. doi: 10.1038/nrd892.

[19] E. Gurevich and V. Gurevich. "Therapeutic Potential of Small Molecules and Engineered Proteins," in *Arrestins - Pharmacology and Therapeutic Potential*, 2014, 1–14.

[20] B. Leader, Q. J. Baca, and D. E. Golan. "Protein therapeutics: a summary and pharmacological classification". *Nat. Rev. Drug Discov.* **2008**, *7*21–39.

[21] D. Wang, P. W. L. Tai, and G. Gao. "Adeno-associated virus vector as a platform for gene therapy delivery". *Nat. Rev. Drug Discov.* **2019**, *18* (5), 358–378. doi: 10.1038/s41573-019-0012-9.

[22] S. A. Mendonça, R. Lorincz, P. Boucher, and D. T. Curiel. "Adenoviral vector vaccine platforms in the SARS-CoV-2 pandemic". *npj Vaccines* **2021**, *6* (97). doi: 10.1038/s41541-021-00356-x.

[23] B. Kelley. "Industrialization of mAb production technology: The bioprocessing industry at a crossroads". *MAbs* **2009**, *1* (5), 443–452. doi: 10.4161/mabs.1.5.9448.

[24] F. Conroy *et al*. "Chemical engineering of therapeutic siRNAs for allele-specific gene silencing in Huntington's disease models". *Nat. Commun.* **2022**, *13*. doi: 10.1038/s41467-022-33061-x.

[25] M. Winkle, S. M. El-Daly, M. Fabbri, and G. A. Calin. "Noncoding RNA therapeutics — challenges and potential solutions". *Nat. Rev. Drug Discov.* **2021**, *20* (8), 629–651. doi: 10.1038/s41573-021-00219-z.

[26] T. Li *et al*. "CRISPR/Cas9 therapeutics: progress and prospects". *Signal Transduct. Target. Ther*. **2023**, *8* (36). doi: 10.1038/s41392-023-01309-7.

[27] M. Schlee and G. Hartmann. "Discriminating self from non-self in nucleic acid sensing". *Nat. Rev. Immunol*. **2016**, *16* (9), 566–580. doi: 10.1038/nri.2016.78.

[28] A. Wittrup and J. Lieberman. "Knocking down disease: A progress report on siRNA therapeutics". *Nat. Rev. Genet*. **2015**, *16* (9), 543–552. doi: 10.1038/nrg3978.

[29] C. Rinaldi and M. J. A. Wood. "Antisense oligonucleotides: The next frontier for treatment of neurological disorders". *Nat. Rev. Neurol*. **2018**, *149*–22. doi: 10.1038/nrneurol.2017.148.

[30] R. L. Juliano. "The delivery of therapeutic oligonucleotides". *Nucleic Acids Res*. **2016**, *44* (14), 6518–6548. doi: 10.1093/nar/gkw236.

[31] J. D. Torres-Vanegas, J. C. Cruz, and L. H. Reyes. "Delivery systems for nucleic acids and proteins: Barriers, cell capture pathways and nanocarriers". *Pharmaceutics* **2021**, *13* (3). doi: 10.3390/pharmaceutics13030428.

[32] F. Eckstein. "Nucleoside Phosphorothioates". *J. Am. Chem. Soc*. **1966**, *88* (18), 4292–4294.

[33] F. Eckstein. "A Dinucleoside Phosphorothioate". *Tetrahedron Lett*. **1967**, (13), 1157–1160.

[34] N. Iwamoto *et al*. "Control of phosphorothioate stereochemistry substantially increases the efficacy of antisense oligonucleotides". *Nat. Biotechnol*. **2017**, *35* (9), 845–851. doi: 10.1038/nbt.3948.

[35] H. Jahns *et al*. "Chirality matters: Stereo-defined phosphorothioate linkages at the termini of small interfering RNAs improve pharmacology in vivo". *Nucleic Acids Res*. **2022**, *50* (3), 1221–1240. doi: 10.1093/nar/gkab544.

[36] D. Hartmann and M. J. Booth. “Light-controlled cell-free protein synthesis using phosphorothioate-caged antisense oligonucleotides”. *Commun. Chem.* **2023**, *6* (59). doi: https://doi.org/10.1038/s42004-023-00860-2.

[37] M. Manoharan. “2’-Carbohydrate modifications in antisense oligonucleotide therapy: Importance of conformation, configuration and conjugation”. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1999**, *1489*117–130. doi: 10.1016/S0167-4781(99)00138-4.

[38] A. Grünweller, E. Wyszko, B. Bieber, R. Jahnel, V. A. Erdmann, and J. Kurreck. “Comparison of different antisense strategies in mammalian cells using locked nucleic acids, 2′-O-methyl RNA, phosphorothioates and small interfering RNA”. *Nucleic Acids Res.* **2003**, *31* (12), 3185–3193. doi: 10.1093/nar/gkg409.

[39] D. A. Brown *et al.* “Effect of phosphorothioate modification of oligodeoxynucleotides on specific protein binding”. *J. Biol. Chem.* **1994**, *269* (43), 26801–26805. doi: 10.1016/s0021-9258(18)47090-1.

[40] M. Egli and M. Manoharan. “Chemistry, structure and function of approved oligonucleotide therapeutics”. *Nucleic Acids Res.* **2023**, *51* (6), 2529–2573. doi: 10.1093/nar/gkad067.

[41] P. S. Pallan *et al.* “Structure and nuclease resistance of 2′,4′-constrained 2′-O-methoxyethyl (cMOE) and 2′-O-ethyl (cEt) modified DNAs”. *Chem. Commun.* **2012**, *48* (66), 8195–8197. doi: 10.1039/c2cc32286b.

[42] S. Ochoa and V. T. Milam. “Modified nucleic acids: Expanding the capabilities of functional oligonucleotides”. *Molecules* **2020**, *25* (20). doi: 10.3390/molecules25204659.

[43] W. Brad Wan and P. P. Seth. “The Medicinal Chemistry of Therapeutic Oligonucleotides”. *J. Med. Chem.* **2016**, *59* (21), 9645–9667. doi: 10.1021/acs.jmedchem.6b00551.

[44] Biosynthesis. "5-Methylcytosine" oligonucleotide modification – https://www.biosyn.com/oligonucleotideproduct/5-methylcytidine,%205-me-c%20oligonucleotide%20modification.aspx

[45] M. Hawner and C. Ducho. "Cellular Targeting of Oligonucleotides by Conjugation with Small Molecules". *Molecules* **2020**, *25*5963. doi: 10.3390/MOLECULES25245963.

[46] A. Gupta, A. Mishra, and N. Puri. "Peptide nucleic acids: Advanced tools for biomedical applications". *J. Biotechnol.* **2017**, *259*148–159. doi: 10.1016/j.jbiotec.2017.07.026.

[47] T. C. Roberts, R. Langer, and M. J. A. Wood. "Advances in oligonucleotide drug delivery". *Nat. Rev. Drug Discov.* **2020**, *19* (10), 673–694. doi: 10.1038/s41573-020-0075-7.

[48] E. Samaridou, J. Heyes, and P. Lutwyche. "Lipid nanoparticles for nucleic acid delivery: Current perspectives". *Adv. Drug Deliv. Rev.* **2020**, *154–155*37–63. doi: 10.1016/j.addr.2020.06.002.

[49] Y. Huang. "Preclinical and Clinical Advances of GalNAc-Decorated Nucleic Acid Therapeutics". *Mol. Ther. - Nucleic Acids* **2017**, *6* (March), 116–132. doi: 10.1016/j.omtn.2016.12.003.

[50] S. Benizri, A. Gissot, A. Martin, B. Vialet, M. W. Grinstaff, and P. Barthélémy. "Bioconjugated Oligonucleotides: Recent Developments and Therapeutic Applications". *Bioconjug. Chem.* **2019**, *30* (2), 366–383. doi: 10.1021/acs.bioconjchem.8b00761.

[51] C. H. Kapadia, J. R. Melamed, and E. S. Day. "Spherical Nucleic Acid Nanoparticles: Therapeutic Potential". *BioDrugs* **2018**, *32* (4), 297–309. doi: 10.1007/s40259-018-0290-5.

[52] I. Urits *et al*. "A Review of Patisiran (ONPATTRO®) for the Treatment of Polyneuropathy in People with Hereditary Transthyretin Amyloidosis". *Neurol. Ther*. **2020**, *9* (2), 301–315. doi: 10.1007/s40120-020-00208-1.

[53] R. S. Vasan, D. M. Enserro, V. Xanthakis, A. S. Beiser, and S. Seshadri. "Temporal Trends in the Remaining Lifetime Risk of Cardiovascular Disease Among Middle-Aged Adults Across 6 Decades: The Framingham Study". *Circulation* **2022**, *145* (17), 1324–1338. doi: 10.1161/CIRCULATIONAHA.121.057889.

[54] K. K. Ray *et al*. "Long-term efficacy and safety of inclisiran in patients with high cardiovascular risk and elevated LDL cholesterol (ORION-3): results from the 4-year open-label extension of the ORION-1 trial". *Lancet Diabetes Endocrinol*. **2023**, *11* (2), 109–119. doi: 10.1016/S2213-8587(22)00353-9.

[55] J. Moreno-Montañés, A. M. Bleau, and A. I. Jimenez. "Tivanisiran, a novel siRNA for the treatment of dry eye disease". *Expert Opin. Investig. Drugs* **2018**, *27* (4), 421–426. doi: 10.1080/13543784.2018.1457647.

[56] S. A. Benner and A. M. Sismour. "Synthetic biology". *Nat. Rev. Genet*. **2005**, *6* (7), 533–543. doi: 10.1038/nrg1637.

[57] C. A. Voigt. "Synthetic biology 2020–2030: six commercially-available products that are changing our world". *Nat. Commun*. **2020**, *11*10–15. doi: 10.1038/s41467-020-20122-2.

[58] D. E. Cameron, C. J. Bashor, and J. J. Collins. "A brief history of synthetic biology". *Nat. Rev. Microbiol*. **2014**, *12* (5), 381–390. doi: 10.1038/nrmicro3239.

[59] T. S. Gardner, C. R. Cantor, and J. J. Collins. "Construction of a genetic toggle switch in Escherichia coli". *Nature* **2000**, *403* (6767), 339–342. doi: 10.1038/35002131.

[60] K. Khambhati, G. Bhattacharjee, N. Gohil, D. Braddick, V. Kulkarni, and V. Singh. "Exploring the Potential of Cell-Free Protein Synthesis for Extending the Abilities of Biological Systems". *Front. Bioeng. Biotechnol.* **2019**, *7* (October). doi: 10.3389/fbioe.2019.00248.

[61] R. W. Martin *et al*. "Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids". *Nat. Commun.* **2018**, *9* (1203). doi: 10.1038/s41467-018-03469-5.

[62] L. E. Contreras-Llano and C. Tan. "High-throughput screening of biomolecules using cell-free gene expression systems". *Synth. Biol.* **2018**, *3*. doi: 10.1093/synbio/ysy012.

[63] K. G. Wandera, S. P. Collins, F. Wimmer, R. Marshall, V. Noireaux, and C. L. Beisel. "An enhanced assay to characterize anti-CRISPR proteins using a cell-free transcription-translation system". *Methods* **2020**, *17*242–50. doi: 10.1016/j.ymeth.2019.05.014.

[64] D. Siegal-Gaskins, Z. A. Tuza, J. Kim, V. Noireaux, and R. M. Murray. "Gene circuit performance characterization and resource usage in a cell-free 'breadboard'". *ACS Synth. Biol.* **2014**, *3* (6), 416–425. doi: 10.1021/sb400203p.

[65] D. Garenne, M. C. Haines, E. F. Romantseva, P. Freemont, E. A. Strychalski, and V. Noireaux. "Cell-free gene expression". *Nat. Rev. Methods Prim.* **2021**, *1* (49). doi: 10.1038/s43586-021-00046-x.

[66] B. Melinek *et al*. "Toward a Roadmap for Cell-Free Synthesis in Bioprocessing". *Bioprocess Int.* **2020**.

[67] H. A. de Boer, L. J. Comstock, and M. Vasser. "The tac promoter: a functional hybrid derived from the trp and lac promoters.". *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*21–25. doi: 10.1073/pnas.80.1.21.

[68] R. Lutz and H. Bujard. “Independent and tight regulation of transcriptional units in escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements”. *Nucleic Acids Res*. **1997**, *25* (6), 1203–1210. doi: 10.1093/nar/25.6.1203.

[69] C. L. Dent, G. Lau, E. A. Drake, A. Yoon, C. C. Case, and P. D. Gregory. “Regulation of endogenous gene expression using small molecule-controlled engineered zinc-finger protein transcription factors”. *Gene Ther*. **2007**, *14* (18), 1362–1369. doi: 10.1038/sj.gt.3302985.

[70] J. Liang, M. J. Mclachlan, and H. Zhao. “Orthogonal control of endogenous gene expression in mammalian cells using synthetic ligands”. *Biotechnol. Bioeng*. **2013**, *110* (5), 1419–1429. doi: 10.1002/bit.24807.

[71] K. Pardee *et al*. “Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components”. *Cell* **2016**, *165* (5), 1255–1266. doi: 10.1016/j.cell.2016.04.059.

[72] P. L. Voyvodic *et al*. “Plug-and-play metabolic transducers expand the chemical detection space of cell-free biosensors”. *Nat. Commun*. **2019**, *10*1697. doi: 10.1038/s41467-019-09722-9.

[73] D. Hartmann, J. M. Smith, G. Mazzotti, R. Chowdhry, and M. J. Booth. “Controlling gene expression with light: a multidisciplinary endeavour”. *Biochem. Soc. Trans*. **2020**, *48* (4), 1645–1659. doi: 10.1042/BST20200014.

[74] X. Kong, N. M. Wakida, and K. Yokomori. “Application of Laser Microirradiation in the Investigations of Cellular Responses to DNA Damage”. *Front. Phys*. **2021**, *8*. doi: 10.3389/fphy.2020.597866.

[75] Y. V. Il’ichev, M. A. Schwörer, and J. Wirz. “Photochemical Reaction Mechanisms of 2-Nitrobenzyl Compounds: Methyl Ethers and Caged ATP”. *J. Am. Chem. Soc*. **2004**, *126* (14), 4581–4595. doi: 10.1021/ja039071z.

[76] C. Hamerla *et al*. "Photochemical mechanism of DEACM uncaging: A combined time-resolved spectroscopic and computational study". *Phys. Chem. Chem. Phys.* **2020**, *22* (24), 13418–13430. doi: 10.1039/c9cp07032j.

[77] S. Wen, J. Zhou, K. Zheng, A. Bednarkiewicz, X. Liu, and D. Jin. "Advances in highly doped upconversion nanoparticles". *Nat. Commun.* **2018**, *9*. doi: 10.1038/s41467-018-04813-5.

[78] Y. Kamiya, T. Takagi, H. Ooi, H. Ito, X. Liang, and H. Asanuma. "Synthetic gene involving azobenzene-tethered t7 promoter for the photocontrol of gene expression by visible light". *ACS Synth. Biol.* **2015**, *4* (4), 365–370. doi: 10.1021/sb5001092.

[79] A. Baumschlager. "Engineering Light-Control in Biology". *Front. Bioeng. Biotechnol.* **2022**, *10*. doi: 10.3389/fbioe.2022.901300.

[80] S. B. Cambridge *et al*. "Doxycycline-dependent photoactivated gene expression in eukaryotic systems". *Nat. Methods* **2009**, *6* (7), 527–531. doi: 10.1038/nmeth.1340.

[81] D. Binder, D. Kohlheyer, J. Pietruszka, J. Frunzke, K. Jaeger, and V. F. Wendisch. "Light-Controlled Cell Factories : Employing Photocaged Isopropyl- ␣ - D - Thiogalactopyranoside for Light-Mediated Optimization of lac Promoter-Based Gene Expression and (␣) -Valencene Biosynthesis in". *Appl. Environ. Microbiol.* **2016**, *82* (20), 6141. doi: 10.1128/AEM.01457-16.Editor.

[82] P. T. Wong *et al*. "Control of an Unusual Photo-Claisen Rearrangement in Coumarin Caged Tamoxifen through an Extended Spacer". *ACS Chem. Biol.* **2017**, *12* (4), 1001–1010. doi: 10.1021/acschembio.6b00999.

[83] H. Ando, T. Furuta, R. Y. Tsien, and H. Okamoto. "Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos". *Nat. Genet.* **2001**, *28* (4), 317–325. doi: 10.1038/ng583.

[84] A. Schroeder *et al*. "Remotely activated protein-producing nanoparticles". *Nano Lett*. **2012**, *12* (6), 2685–2689. doi: 10.1021/nl2036047.

[85] S. Ogasawara. "Control of cellular function by reversible photoregulation of translation". *ChemBioChem* **2014**, *15* (18), 2652–2655. doi: 10.1002/cbic.201402495.

[86] S. Ogasawara. "Transcription Driven by Reversible Photocontrol of Hyperstable G-Quadruplexes". *ACS Synth. Biol*. **2018**, *7* (11), 2507–2513. doi: 10.1021/acssynbio.8b00216.

[87] Z. Vaníková, M. Janoušková, M. Kambová, L. Krásný, and M. Hocek. "Switching transcription with bacterial RNA polymerase through photocaging, photorelease and phosphorylation reactions in the major groove of DNA". *Chem. Sci*. **2019**, *10* (14), 3937–3942. doi: 10.1039/c9sc00205g.

[88] M. J. Booth, V. R. Schild, A. D. Graham, S. N. Olof, and H. Bayley. "Light-activated communication in synthetic tissues". *Sci. Adv*. **2016**, *2* (4), e1600056. doi: 10.1126/sciadv.1600056.

[89] J. M. Smith, D. Hartmann, and M. J. Booth. "Engineering cellular communication between light-activated synthetic cells and bacteria". *Nat. Chem. Biol*. **2023**, *19,* 1138–1146. doi: 10.1101/2022.07.22.500923.

[90] D. Hartmann, R. Chowdhry, J. M. Smith, and M. J. Booth. "Orthogonal Light-Activated DNA for Patterned Biocomputing within Synthetic Cells". *JACS* **2023**, *145* (17), 9471–9480. doi: 10.26434/chemrxiv-2022-p8xgb.

[91] B. M. Paterson, B. E. Roberts, and E. L. Kuff. "Structural gene identification and mapping by DNA•mRNA hybrid-arrested cell-free translation". *Proc. Natl. Acad. Sci. U. S. A*. **1977**, *74* (10), 4370–4374. doi: 10.1073/pnas.74.10.4370.

[92] L. Grillone and R. Perry. “Fomivirsen”. *Drugs* **1999**, *57* (3), 375–380. doi: 10.1201/9781315152110.

[93] V. Barresi, C. Musmeci, A. Rinaldi, and D. F. Condorelli. “Transcript-Targeted Therapy Based on RNA Interference and Antisense Oligonucleotides: Current Applications and Novel Molecular Targets”. *Int. J. Mol. Sci.* **2022**, *23* (16). doi: 10.3390/ijms23168875.

[94] X. H. Liang, H. Sun, J. G. Nichols, and S. T. Crooke. “RNase H1-Dependent Antisense Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the Nucleus”. *Mol. Ther.* **2017**, *25* (9), 2075–2092. doi: 10.1016/j.ymthe.2017.06.002.

[95] M. Hyjek, M. Figiel, and M. Nowotny. “RNases H: Structure and mechanism”. *DNA Repair (Amst).* **2019**, *84* (July), 102672. doi: 10.1016/j.dnarep.2019.102672.

[96] M. Nowotny, S. A. Gaidamakov, R. J. Crouch, and W. Yang. “Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis”. *Cell* **2005**, *121* (7), 1005–1016. doi: 10.1016/j.cell.2005.04.024.

[97] M. E. Visser *et al*. “Mipomersen, an apolipoprotein B synthesis inhibitor, lowers low-density lipoprotein cholesterol in high-risk statin-intolerant patients: A randomized, double-blind, placebo-controlled trial”. *Eur. Heart J.* **2012**, *33* (9), 1142–1149. doi: 10.1093/eurheartj/ehs023.

[98] C. Di *et al*. “Function, clinical application, and strategies of Pre-mRNA splicing in cancer”. *Cell Death Differ.* **2019**, *26* (7), 1181–1194. doi: 10.1038/s41418-018-0231-3.

[99] X. Shen and D. R. Corey. “Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs”. *Nucleic Acids Res.* **2018**, *46* (4), 1584–1600. doi: 10.1093/nar/gkx1239.

[100] Saifullah, N. Motohashi, T. Tsukahara, and Y. Aoki. “Development of Therapeutic RNA Manipulation for Muscular Dystrophy”. *Front. Genome Ed.* **2022**, *4*. doi: 10.3389/fgeed.2022.863651.

[101] K. R. Q. Lim, R. Maruyama, and T. Yokota. “Eteplirsen in the treatment of Duchenne muscular dystrophy”. *Drug Des. Devel. Ther.* **2017**, *11*533–545. doi: 10.2147/DDDT.S97635.

[102] W. Sato *et al*. “A gene expression control technology for cell-free systems and synthetic cells via targeted gene silencing and transfection”. *bioRxiv* **2022**. doi: 10.1101/2022.07.28.501919.

[103] T. Yamamoto *et al*. “Evaluation of multiple-turnover capability of locked nucleic acid antisense oligonucleotides in Cell-Free RNase H-mediated antisense reaction and in mice”. *Nucleic Acid Ther*. **2014**, *24* (4), 283–290. doi: 10.1089/nat.2013.0470.

[104] X. Tang and I. J. Dmochowski. “Controlling RNA digestion by RNase H with a light-activated DNA hairpin”. *Angew. Chemie - Int. Ed.* **2006**, *45* (21), 3523–3526. doi: 10.1002/anie.200600954.

[105] L. Wu, Y. Wang, J. Wu, C. Lv, J. Wang, and X. Tang. “Caged circular antisense oligonucleotides for photomodulation of RNA digestion and gene expression in cells”. *Nucleic Acids Res*. **2013**, *41*677–686. doi: 10.1093/nar/gks996.

[106] L. Yang, H. B. Kim, J. Y. Sul, S. B. Yeldell, J. H. Eberwine, and I. J. Dmochowski. “Efficient Synthesis of Light-Triggered Circular Antisense Oligonucleotides Targeting Cellular Protein Expression”. *ChemBioChem* **2018**, *19* (12), 1250–1254. doi: 10.1002/cbic.201800012.

[107] A. R. Buskirk and D. R. Liu. “Creating small-molecule-dependent switches to modulate biological functions”. *Chem. Biol.* **2005**, *12* (2), 151–161. doi: 10.1016/j.chembiol.2004.11.012.

[108] C. C. Liu, M. C. Jewett, J. W. Chin, and C. A. Voigt. "Toward an orthogonal central dogma". *Nat. Chem. Biol.* **2018**, *14* (2), 103–106. doi: 10.1038/nchembio.2554.

[109] K. Müller, R. Engesser, J. Timmer, M. D. Zurbriggen, and W. Weber. "Orthogonal optogenetic triple-gene control in mammalian cells". *ACS Synth. Biol.* **2014**, *3* (11), 796–801. doi: 10.1021/sb500305v.

[110] L. Fournier *et al*. "A blue-absorbing photolabile protecting group for in vivo chromatically orthogonal photoactivation". *ACS Chem. Biol.* **2013**, *8* (7), 1528–1536. doi: 10.1021/cb400178m.

[111] G. J. Kremers, J. Goedhart, E. B. Van Munster, and T. W. J. Gadella. "Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET förster radius". *Biochemistry* **2006**, *45* (21), 6570–6580. doi: 10.1021/bi0516273.

[112] D. Hartmann, R. Chowdhry, J. M. Smith, and M. J. Booth. "Orthogonal Light-Activated DNA for Patterned Biocomputing within Synthetic Cells". *J. Am. Chem. Soc.* **2023**, *145* (17), 9471–9480. doi: 10.1021/jacs.3c02350.

[113] C. Napoli, C. Lemieux, and R. Jorgensen. "Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans". *Plant Cell* **1990**, *2* (4), 279–289. doi: 10.2307/3869076.

[114] F. Ratcliff, B. D. Harrison, and D. C. Baulcombe. "A similarity between viral defense and gene silencing in plants". *Science (80-.).* **1997**, *276* (5318), 1558–1560. doi: 10.1126/science.276.5318.1558.

[115] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. "Potente and specific genetic interference by double-stranded RNA in Caenorhabditis elegans". *Nature* **1998**, *391* (February), 806–811.

[116] P. Svoboda. “Key Mechanistic Principles and Considerations Concerning RNA Interference”. *Front. Plant Sci.* **2020**, *11*1237. doi: 10.3389/fpls.2020.01237.

[117] D. C. Baulcombe. “RNAi silencing in plants”. *Nature* **2004**, *431*356–363.

[118] S. A. Shabalina and E. V. Koonin. “Origins and evolution of eukaryotic RNA interference”. *Trends Ecol. Evol.* **2008**, *23* (10), 578–587. doi: 10.1016/j.tree.2008.06.005.

[119] R. Mishra, A. Kumar, H. Ingle, and H. Kumar. “The Interplay Between Viral-Derived miRNAs and Host Immunity During Infection”. *Front. Immunol.* **2020**, *10*. doi: 10.3389/fimmu.2019.03079.

[120] R. Yi, Y. Qin, I. G. Macara, and B. R. Cullen. “Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs”. *Genes Dev.* **2003**, *17* (24), 3011–3016. doi: 10.1101/gad.1158803.

[121] R. L. Setten, J. J. Rossi, and S. ping Han. “The current state and future directions of RNAi-based therapeutics”. *Nat. Rev. Drug Discov.* **2019**, *18* (6), 421–446. doi: 10.1038/s41573-019-0017-4.

[122] A. J. Varley and J. P. Desaulniers. “Chemical strategies for strand selection in short-interfering RNAs”. *RSC Adv.* **2021**, *11* (4), 2415–2426. doi: 10.1039/d0ra07747j.

[123] B. Haley and P. D. Zamore. “Kinetic analysis of the RNAi enzyme complex”. *Nat. Struct. Mol. Biol.* **2004**, *11* (7), 599–606. doi: 10.1038/nsmb780.

[124] V. N. Kim, J. Han, and M. C. Siomi. “Biogenesis of small RNAs in animals”. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (2), 126–139. doi: 10.1038/nrm2632.

[125] E. Song *et al.* “RNA interference targeting Fas protects mice from fulminant hepatitis”. *Nat. Med.* **2003**, *9* (3), 347–351. doi: 10.1038/nm828.

[126] J. K. Watts and D. R. Corey. “Silencing disease genes in the laboratory and the clinic”. *J. Pathol.* **2012**, *226* (2), 365–379. doi: 10.1002/path.2993.

[127] M. M. Zhang, R. Bahal, T. P. Rasmussen, J. E. Manautou, and X. bo Zhong. "The growth of siRNA-based therapeutics: Updated clinical studies". *Biochem. Pharmacol.* **2021**, *189*. doi: 10.1016/j.bcp.2021.114432.

[128] S. J. Keam. "Vutrisiran: First Approval". *Drugs* **2022**, *82* (13), 1419–1425. doi: 10.1007/s40265-022-01765-5.

[129] K. Y. Zhu and S. R. Palli. "Mechanisms, applications, and challenges of insect RNA interference". *Annu. Rev. Entomol.* **2020**, *65*293–311. doi: 10.1146/annurev-ento-011019-025224.

[130] P. K. Jain, S. Shah, and S. H. Friedman. "Patterning of gene expression using new photolabile groups applied to light activated RNAi". *J. Am. Chem. Soc.* **2011**, *133* (3), 440–446. doi: 10.1021/ja107226e.

[131] M. L. Hammill, G. Islam, and J. P. Desaulniers. "Synthesis, Derivatization and Photochemical Control of ortho-Functionalized Tetrachlorinated Azobenzene-Modified siRNAs". *ChemBioChem* **2020**, *21* (16), 2367–2372. doi: 10.1002/cbic.202000188.

[132] J. López-Andarias *et al.* "Cell-Penetrating Streptavidin: A General Tool for Bifunctional Delivery with Spatiotemporal Control, Mediated by Transport Systems Such as Adaptive Benzopolysulfane Networks". *J. Am. Chem. Soc.* **2020**, *142* (10), 4784–4792. doi: 10.1021/jacs.9b13621.

[133] Y. Okamoto *et al.* "A cell-penetrating artificial metalloenzyme regulates a gene switch in a designer mammalian cell". *Nat. Commun.* **2018**, *9*. doi: 10.1038/s41467-018-04440-0.

[134] B. Albarran, R. To, and P. S. Stayton. "A TAT-streptavidin fusion protein directs uptake of biotinylated cargo into mammalian cells". *Protein Eng. Des. Sel.* **2005**, *18* (3), 147–152. doi: 10.1093/protein/gzi014.

[135] X. Huang *et al*. "Light-activated RNA interference in human embryonic stem cells". *Biomaterials* **2015**, *63*70–79. doi: 10.1016/j.biomaterials.2015.06.006.

[136] X. Huang, Y. Lai, G. B. Braun, and N. O. Reich. "Modularized Gold Nanocarriers for TAT-Mediated Delivery of siRNA". *Small* **2017**, *13* (8). doi: 10.1002/smll.201602473.

[137] S. M. Elbashir, J. Martinez, A. Patkaniowska, W. Lendeckel, and T. Tuschl. "Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate". *EMBO J*. **2001**, *20* (23), 6877–6888. doi: 10.1093/emboj/20.23.6877.

[138] J. Xie, Z. Chen, X. Zhang, H. Chen, and W. Guan. "Identification of an RNase that preferentially cleaves A/G nucleotides". *Sci*. *Rep*. **2017**, *7*. doi: 10.1038/srep45207.

[139] A. Kala and S. H. Friedman. "Enhanced light-activated RNA interference using phosphorothioate-based dsRNA precursors of siRNA". *Pharm*. *Res*. **2011**, *28* (12), 3050–3057. doi: 10.1007/s11095-011-0529-z.

[140] X. Zhu, T. Li, Y. Dang, Y. Feng, and P. Huang. "A novel in vitro transcription method for producing siRNAs without specific sequence requirements". *Mol*. *Biotechnol*. **2005**, *31* (3), 187–192. doi: 10.1385/MB:31:3:187.

[141] Y. Gholamalipour, A. Karunanayake Mudiyanselage, and C. T. Martin. "NAR breakthrough article 3 end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character—RNA-Seq analyses". *Nucleic Acids Res*. **2018**, *46* (18), 9253–9263. doi: 10.1093/nar/gky796.

[142] J. F. Milligan, D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck. "Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates". *Nucleic Acids Res*. **1987**, *15* (21), 8783–8798. doi: 10.1093/nar/15.21.8783.

[143] A. P. Liu *et al*. "The living interface between synthetic biology and biomaterial design". *Nat. Mater*. **2022**, *21* (4), 390–397. doi: 10.1038/s41563-022-01231-3.

[144] J. F. Pelletier *et al*. "Genetic requirements for cell division in a genomically minimal cell". *Cell* **2021**, *184* (9), 2430-2440.e16. doi: 10.1016/j.cell.2021.03.008.

[145] J. A. Peterson *et al*. "Family of BODIPY Photocages Cleaved by Single Photons of Visible/Near-Infrared Light". *J. Am. Chem. Soc*. **2018**, *140* (23), 7343–7346. doi: 10.1021/jacs.8b04040.

[146] A. P. Gorka, R. R. Nani, J. Zhu, S. Mackem, and M. J. Schnermann. "A near-IR uncaging strategy based on cyanine photochemistry". *J. Am. Chem. Soc*. **2014**, *136* (40), 14153–14159. doi: 10.1021/ja5065203.

[147] H. Du Rietz, H. Hedlund, S. Wilhelmson, P. Nordenfelt, and A. Wittrup. "Imaging small molecule-induced endosomal escape of siRNA". *Nat. Commun*. **2020**, *11*. doi: 10.1038/s41467-020-15300-1.