THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Enlightening RNA biology – Insights from Fluorescent Nucleobase Analogues

from fluorophore characterization to applications in live-cell imaging

PAULINE PFEIFFER

Department of Chemistry and Chemical Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2025

Enlightening RNA biology – Insights from Fluorescent Nucleobase Analogues
From fluorophore characterization to applications in live-cell imaging

Pauline Pfeiffer
ISBN 978-91-8103-269-7

Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5727
ISSN 0346-718X

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

This work was made possible by the funding from the Chalmers **Nano Area of Advance** (formerly Excellence Initiative Nano) PhD program.

Front cover:
Artistic outline of the thesis content with a fluorescent base analogue (FBA, yellow glow) in the middle. Left: cellular uptake of fluorescent nucleosides and live-cell images using confocal microscopy (green), two-photon microscopy (magenta), and FLIM (jet colour scale). Top right: FBA incorporated into an oligonucleotide and its photophysical characterization by circular dichroism, absorption melting, fluorescence lifetime determination, and UV-Vis absorption spectroscopy. Bottom right: labelling of long RNA *in vitro* and the subsequent live-cell imaging of the labelled RNA and the resulting fluorescent protein. All images and spectra are based on data presented in this thesis.

Back cover:
Photo by Jesper Nilsson with active support from Stefano Racioppi.

Printed by Chalmers digitaltryck
Gothenburg, Sweden 2025

Enlightening RNA biology – Insights from Fluorescent Nucleobase Analogues
From fluorophore characterization to applications in live-cell imaging

PAULINE PFEIFFER
Department of Chemistry and Chemical Engineering
Chalmers University of Technology

ABSTRACT

Well-characterized, native-like probes to study RNA and its nucleotides are essential for advancing both fundamental research and the development of RNA-based therapeutics. Fluorescent nucleobase analogues (FBAs) offer unique opportunities to probe nucleic acids under physiological conditions while preserving native interaction patterns. This thesis explores the use of fluorescent nucleosides to investigate RNA and nucleotide metabolism, with a particular focus on live-cell imaging applications. It includes a comprehensive photophysical characterization of FBAs, emphasizing their suitability for various microscopy techniques. These insights are applied to monitor the spontaneous cellular uptake of FBAs using fluorescence microscopy, including confocal imaging, two-photon excitation microscopy, and fluorescence lifetime imaging. One adenine analogue, 2CNqA, emerged as a key compound due to its ability to enter cells and localize to the nucleus, *i.e.* the site of RNA synthesis. In addition to imaging, techniques such as flow cytometry and spectroscopy were employed to elucidate the uptake pathways and the metabolic processing of 2CNqA, including its incorporation into RNA. This work also demonstrates the incorporation of 2CNqA into long RNA via *in vitro* transcription, including a thorough characterization of 2CNqA incorporation degree and its photophysical properties in an RNA environment. The subsequent cellular delivery of this FBA-labelled mRNA and the expression of the encoded fluorescent protein are investigated for applicability for RNA studies in living cells.

Together, the thorough characterization of FBAs, the observed spontaneous uptake and subsequent RNA labelling, and the *in vitro* transcription-based labelling strategy constitute a versatile toolbox for studying dynamic aspects of RNA biology and RNA-based drug delivery in living cells. This thesis also highlights the potential of diverse readout methods for advancing RNA imaging and analysis.

Keywords: Fluorescent base analogues, RNA, nucleotides, nucleosides, RNA building blocks, fluorescence, spectroscopy, fluorescence microscopy, live cell imaging, *in vitro* transcription

SAMMANFATTNING

Välkarakteriserade, naturliknande prober för att studera RNA och dess nukleotider är avgörande för att främja både grundforskning och utvecklingen av RNA-baserade läkemedel. Fluorescerande nukleobasanaloger (FBAs) erbjuder unika möjligheter att undersöka nukleinsyror under fysiologiska förhållanden samtidigt som de bevarar naturliga interaktionsmönster.

Denna avhandling utforskar användningen av fluorescerande nukleosider för att undersöka RNA och nukleotidmetabolism, med särskilt fokus på applikationer inom avbildning av levande celler. Den innehåller även en omfattande fotofysikalisk karakterisering av FBAs, särskilt med avseende på deras lämplighet för olika mikroskopitekniker. Dessa insikter används sedan för att följa det spontana cellulära upptaget av FBAs med hjälp av fluorescensmikroskopi, inklusive konfokalmikroskopi, tvåfotonexcitationsmikroskpi och fluorescenslivslängdsavbildning. En adeninanalog, 2CNqA, framträdde som en nyckelmolekyl tack vare sin förmåga att ta sig in i celler och lokalisera till cellkärnan – celldelen för RNA-syntes. Förutom avbildning användes tekniker som flödescytometri och spektroskopi för att utreda upptagsvägar och metabol omsättning av 2CNqA, inklusive dess inkorporering i RNA. Avhandlingen visar också hur 2CNqA kan inkorporeras i långt RNA via *in vitro*-transkription, följt av en noggrann karakterisering av inkorporeringsgraden och dess fotofysiska egenskaper i RNA-miljön. Den cellulära leveransen av detta FBA-märkta mRNA och uttrycket av det kodade fluorescerande proteinet undersöks för tillämpbarhet för RNA-studier i levande celler.

Tillsammans representerar den noggranna karakteriseringen, det spontana upptaget av FBAs och den *in vitro*-transkriptionsbaserade märkningsstrategin en mångsidig verktygslåda för att studera dynamiska aspekter av RNA-biologi och RNA-baserade läkemedels leverans i levande celler. Avhandlingen lyfter också fram potentialen hos olika avläsningsmetoder för att främja RNA-avbildning och analys.

Nyckelord: Fluorescerande basanalog, RNA, nukleotider, nukleosider, RNA-byggstenar, fluorescens, spektroskopi, fluorescensmikroskopi, avbildning av levande celler, *in vitro*-transkription

LIST OF APPENDED PAPERS

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text. Shared first authorship indicated by ₸.

Paper I **Multiphoton characterization and live cell imaging using fluorescent adenine analogue 2CNqA.** Jesper R. Nilsson, Carlos Benitez-Martin, Henry Samson, Pauline Pfeiffer, Tom Baladi, Anders Dahlén, Steven Magennis, L. Marcus Wilhelmsson. *Phys. Chem. Chem. Phys.*, **2023**, 25, 20218-20224.

Paper II **Metabolic RNA labeling in non-engineered cells following spontaneous uptake of fluorescent nucleoside phosphate analogues.** Pauline Pfeiffer, Jesper R. Nilsson, Audrey Gallud, Tom Baladi, Hoang-Ngoan Le, Mattias Bood, Malin Lemurell, Anders Dahlén, Morten Grøtli, Elin K. Esbjörner, and L. Marcus Wilhelmsson *Nucleic Acids Research,* **2024**, 52, 10102-10118.

Paper III **Expanding fluorescent base analogue labelling of long RNA by *in vitro* transcription.** Pauline Pfeiffer₸, Alma F. E. Karlsson₸, Jesper R. Nilsson, L. Marcus Wilhelmsson
Manuscript submitted to *Journal of Biological Chemistry*

Paper IV **Increased Brightness of Fluorescent Uridine, qU, inside Single- and Double-Stranded RNA.** Alma F. E. Karlsson₸, Pauline Pfeiffer₸, Hoang-Ngoan Le, Tom Baladi, Anders Dahlén, L. Marcus Wilhelmsson
Manuscript in preparation for *Scientific Reports*

Paper V **Monitoring Nucleoside Metabolism in Living Cells with a Nucleobase Analogue *via* Fluorescence Lifetime Imaging.** Pauline Pfeiffer, Niusha Bagheri, Chen Qian, Jerker Widengren, L. Marcus Wilhelmsson
Accepted for publication in *Chemical Communication*

NOT INCLUDED IN THIS THESIS

Interbase FRET in RNA: from A to Z. Anders F. Füchtbauer, Moa S. Wranne, Mattias Bood, Erik Weis, Pauline Pfeiffer, Jesper R. Nilsson, Anders Dahlén, Morten Grøtli, L. Marcus Wilhelmsson. *Nucleic Acids Research*, **2019**, 47, 9990–9997.

Getting DNA and RNA out of the dark with 2CNqA: a bright adenine analogue and interbase FRET donor. Anna Wypijewska del Nogal, Anders F. Füchtbauer, Mattias Bood, Jesper R. Nilsson, Moa S. Wranne, Sangamesh Sarangamath, Pauline Pfeiffer, Vinoth Sundar Rajan, Afaf H. El-Sagheer, Anders Dahlen, Tom Brown, Morten Grøtli, L. Marcus Wilhelmsson. *Nucleic Acids Research*, **2020**, 48,7640–7652.

Fluorescent Nucleoside Phosphates. Patent WO2022218943A1, submitted to PCT by Stealth Labels Biotech AB with Elin Esbjörner Winters, Marcus Wilhelmsson, Pauline Pfeiffer, Audrey Gallud, Jesper Nilsson. **2021.** *Patent pending.*

CONTRIBUTION REPORT

Paper I Carried out the cell culture work and the single-photon characterization of 2CNqATP. Contributed to preparation of the microscopy figures.

Paper II Performed a majority of the experiments, processed the experimental data, performed the analysis, designed the figures, and drafted the manuscript.

Paper III Helped to outline the study, designed the template sequence with J.R.N., performed the capping and tailing reactions of the 2CNqA-RNA, and the lipofection together with the flow cytometry and microscopy readout, analysed the data, and designed the corresponding figures. Drafted and wrote the manuscript together with A.F.E.K..

Paper IV Performed the circular dichroism measurements and determined the melting temperatures, analysed and interpreted the data, and prepared the corresponding figures. Drafted and wrote the manuscript together with A.F.E.K..

Paper V Conceived the project together with L.M.W., performed FLIM experiments together with N.B., analysed the FLIM data together with the other authors, performed and analysed TCSPC measurements in solution, designed the figures, drafted and wrote the manuscript.

Note: No contribution to the organic synthesis work was made by the author of this thesis.

“Doch Wissenschaft kann niemals absolut sein. Sie ist die Kunst der Annäherung.”

“But science can never be absolute. It is the art of approaching.”

Frank Schätzing, “Nachrichten aus einem unbekannten Universum”

CONTENT

1. INTRODUCTION

Ribonucleic acid (RNA) gained unprecedented global attention during the COVID-19 pandemic, as the first mRNA-based vaccines became widely known to the public. However, within the scientific community, RNA has long been a subject of intense research. The rapid development of the COVID-19 mRNA vaccines was the result of decades of foundational work which highlights the role of RNA and its building blocks in modern medicine. Beyond vaccines, nucleoside analogues have also emerged as powerful drugs in the treatment of viral infections and cancer — though these advances remain less visible to the general public.

As the fight against diseases continues, the development of new therapeutic strategies and diagnostic methods is essential. Therefore, research tools are needed to study RNA and its building blocks to accomplish further design and development of drug candidates. These tools must provide reliable readouts without interfering with the system of interest, *e.g.* natural cellular pathways — unlike drug candidates themselves, which are specifically designed to interfere with the targeted pathways. Continuous innovation in research methodologies is vital to advancing both fundamental understanding and therapeutic applications.

Among many techniques used in molecular biology, fluorescence-based methods are very valuable and widely used. They complement structural techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), and cryo-electron microscopy (cryo-EM), while offering advantages like high sensitivity and specificity, compatibility with live-cell experiments, and cost-effective instrumentation. These features would make fluorescence also a well-suited technique for studying nucleic acids, but the natural nucleobases are virtually non-fluorescent. Therefore, a fluorescent moiety is required for readout. Common fluorophores, that are often bulky dyes, can be attached externally. The alternative approach — central to this thesis — is to use modified bases that are intrinsically fluorescent.

Significant progress has been made in the design and characterization of fluorescent base analogues (FBAs), including prior work by the Wilhelmsson group. Building on this foundation, this thesis presents the photophysical characterization of FBAs and explores their application as tools for studying nucleotide metabolism. Unlike studies

that focus on incorporating FBAs into oligonucleotides via solid-phase synthesis, this work emphasizes the use of fluorescent nucleosides as RNA building blocks for enzymatic incorporation.

The thesis is based on two published articles and three manuscripts. It starts with a detailed background on RNA as molecule and as a therapeutic agent, as well as an overview of RNA labelling strategies. The following methodology part outlines the experimental techniques employed throughout the research this work is based on. The core of the thesis text presents the original research, organized by methodology rather than publication chronology, reflecting the interconnected nature of the research projects. It starts with the photophysical characterization of FBAs, followed by imaging-based investigations, and concludes with applications of FBAs as research tools to study cellular processes. It is important to note that the design and synthesis of the FBAs and oligonucleotides were carried out prior to this work and that the author did not contribute to these aspects.

The final section of this text presents additional projects that were carried out as part of this work but have not yet been published. These projects are included as an outlook, highlighting the potential of FBAs in future research contexts.

Overall, this thesis provides a comprehensive study on how fluorescent nucleosides can be utilized in RNA research, particularly in live-cell studies with focus on imaging. It contributes to the groundwork for future research of nucleic acid biology and the investigations of novel therapeutic strategies and drug candidates.

2. BACKGROUND

2.1 Biochemistry

In this chapter, the background of this thesis is described, including the current research landscape and state-of-the-art of the field.

2.1.1 Ribonucleic Acid

Nucleic acids are, alongside proteins, lipids, and carbohydrates, one of the major macromolecule classes in cells. The two naturally occurring nucleic acids are ribonucleic acid (RNA) and deoxyribonucleic acids (DNA). They are both polymers of building blocks called nucleo*s*ides which are connected by phosphodiester bonds (Fig. 1A). RNA nucleo*t*ides consist of a sugar moiety (β-D-ribofuranose), phosphate group(s) that are attached at the 5' carbon of the sugar, and a nucleobase attached via a N-glycosidic bond at the 1' position (Fig. 1B).

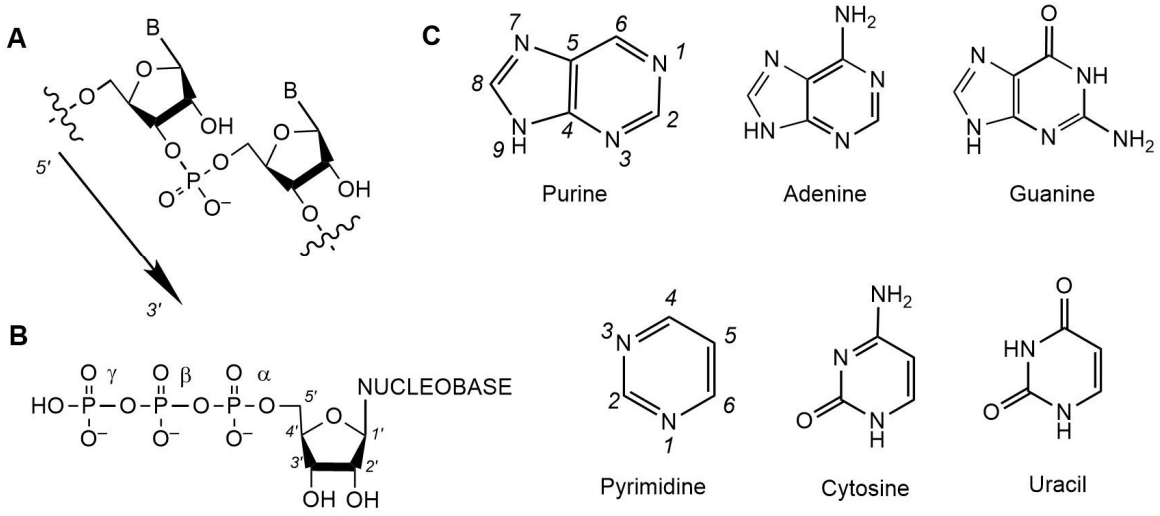

Figure 1. Molecular structure of RNA. A) Shown is a dinucleotide connected via a phosphodiester bond. The arrow indicates the typical reading direction of RNA. B) Building block of RNA: nucleoside triphosphate, which is threefold negatively charged at physiological pH. Indicated in italics are the annotations of sugar atoms and phosphate groups. C) The most common (canonical) nucleobases of RNA, which are derivatives of purine or pyrimidine, with their annotation of atoms indicated in italics.

The alternating sugar and phosphate groups form the "backbone" of the RNA polymer, while the nucleobases define its sequence and are crucial for the functionality of RNA by hydrogen bonding, called base pairing. There are four main nucleobases found in RNA, the two purine

bases adenine and guanine (bound to the sugar at position 9) and the two pyrimidine bases cytosine and uracil (bound to the sugar at position 1) (Fig. 1C). Other (non-canonical or modified) bases can be found in RNA as well, often methylated purines and pyrimidines, like pseudouridine or hypoxanthine, that play a major role in specific cellular processes.[1] RNA exhibits diverse secondary and tertiary structures that can vary significantly depending on sequence, environment, and interactions with other molecules. A prevalent structural motif is the inter- or intrastrand A-form helix, which arises from base pairing between nucleobases.[2,3]

The central dogma of molecular biology, proposed by Francis Crick in 1957, describes the flow from hereditary information (encoded in the DNA), through messenger molecules (mRNA) to the synthesis of polypeptides (proteins) that determine the phenotype.[4] As research continued it became evident that RNA molecules are very diverse and also act as catalysts for biochemical reactions (ribozymes[5]), regulate gene expression (*e.g.* siRNA and miRNA[6]), modify other RNAs (*e.g.* snoRNA[7]), contribute to chromatin remodelling and genome stability (*e.g.* lncRNA and piRNA[8,9]), and are involved in cellular signalling[10]. The characteristics of RNA, including their reaction mechanisms, interactions, nucleotide structures, and macromolecular diversity, have therefore gained interest for biochemical method development and therapeutic applications.[11,12] Although several methods and technologies to detect and study RNA in biological systems already exist, there is still a need for more.

2.1.2 RNA Therapeutics

Building blocks, oligomers, and polymers of RNA have gained significant attention and have become increasingly important in human medicine. Due to the described roles of RNA in cell biology (Ch. 2.1.1), RNA therapeutics offer precise and targeted treatments for a variety of diseases. This chapter will give a short overview of therapeutic approaches based on nucleic acids, with a focus on RNA derivatives. It should be noted that the chemical definition of RNA is not as clear as described above (Ch. 2.1.1) due to, sometimes heavy, molecular modifications on the sugar, the phosphate-sugar backbone, and/or the nucleobases.

RNA therapeutics defines a disease treatment where RNA is a direct target, or where it constitutes the drug itself.[13] During this thesis work, the COVID-19 pandemic hit and gave RNA research great public attention due to the approval of the first mRNA-based vaccines.[14] The RNA in these vaccines is an mRNA sequence coding for a modified version of the spike protein of the virus. Once injected, the mRNA is taken up by cells, particularly antigen-

presenting cells like dendritic cells and monocytes. Inside their cytoplasm, the mRNA is thereafter translated by the ribosomes into a harmless mimic of the viral spike protein, which is transported to the cell membrane and presented on the cell surface. Recognition of the artificial viral motif by the immune system then leads to an immune response, which prepares the body for potential infection by the actual virus.[15] The first mRNA flu vaccine was tested in mice in 1990[16], but difficulties with rapid degradation prevented it from advancing to clinical trials in humans. The necessary breakthrough in mRNA vaccine technology came with the development of advanced delivery methods.[17] Usually, the body quickly degrades foreign RNA as a defence mechanism, making it difficult for vaccine RNA to "deliver the message".[18] Furthermore, long mRNA molecules cannot cross cell membranes due to their significant size, charge, and hydrophilicity.[19] The game-changing development was lipid nanoparticles (LNPs) that encapsulate the mRNA, protecting it from degradation and facilitating cellular uptake via the endocytic pathway.[20,21] The progress of LNP-mRNA formulations for cancer immunotherapy began in the early 2010s, and the first LNP-encapsulated siRNA drug was approved for polyneuropathy in 2018.[20,22] Hence, when the COVID-19 pandemic hit, much of the required knowledge and methods were in place to rapidly develop effective mRNA vaccines.

Following the pandemic, the 2023 Nobel Prize in Physiology or Medicine was awarded to Katalin Karikó and Drew Weissman "for their discoveries concerning nucleoside base modifications that enabled the development of effective mRNA vaccines against COVID-19".[23] Their work is a major contribution to the progress on the mRNA side, demonstrating how chemical modifications can reduce the immune response triggered by RNA and enhance protein production.[24–26] Research on chemical modifications to find new nucleobase analogues to further improve the therapeutic effect of mRNA is very much an active research field.

In addition to the long mRNAs (thousands of nucleotides) used in vaccines and for protein replacement therapy, RNA therapeutics also include single- or double-stranded oligonucleotides in the range of 10–100 nucleotides. These oligonucleotide-based therapeutics display a large diversity in terms of which biochemical process they target, with examples including aptamers, small interfering RNAs (siRNAs), micro RNAs (miRNAs), or antisense oligonucleotides (ASOs).[27,28] Gapmers are one particular example of an ASO design that consist of a DNA segment ("gap") flanked by modified nucleotides, and are usually in the range of 15–20 nucleotides.[29] Modifications are, for example, phosphorothioate linkages or 2'-4'-ethylenoxy-bridged ribose (cEt) on the flanking sides to improve nuclease stability, RNA

affinity, and bioavailability (example in Fig. 13).[30] Gapmers are taken up via the endosomal pathway[31], base-pair with target mRNAs inside cells and the resulting double strand recruits RNase H which degrades the RNA part of the hybrid duplex in a catalytic manner, *i.e.* the gapmer remains intact and can form hybrids in many cycles. This exemplifies how an efficient natural mechanism to regulate gene expression[32] at the RNA level can be exploited for drug development, evidenced by the fact that several gapmer drugs have received regulatory approval[33].

Synthetic compounds based on monomers of ribonucleic acids, also play a major role in human medicine, for example as antiviral or anticancer drugs. Chemical modifications can be made to the phosphate group(s), the sugar moiety, or the nucleobase (compare Fig. 1) with almost infinite possibilities. For therapeutic purposes, a common approach is to create a molecule that interferes with the natural pathways of nucleic acid polymerization to downregulate or modify DNA or RNA synthesis.[34] One example is Remdesivir, a pre-existing drug candidate that was evaluated as an antiviral drug during the 2014 Ebola outbreak.[35] It was later also explored as an inhibitor of the replication of the SARS-CoV-2 and was eventually approved by the European Medicines Agency for treatment of severe COVID-19 during the pandemic.[36] Such nucleoside analogues usually mimic the natural building blocks, *i.e.* nucleoside triphosphates. However, the polar or charged nature of nucleotides prevents them from diffusing across the hydrophobic lipid bilayer of cell membranes. Instead, transporters that allow natural and modified nucleosides (*i.e.* without the phosphate groups) to pass play a major role and are in focus of research.[37] Another approach to deliver nucleoside analogues is to administer them as prodrugs, which are therapeutically inactive compounds that are metabolized to active drugs inside the body. This means that, in addition to chemical modifications for the action as a drug inside cells, further modifications are introduced to mask the charge and allow diffusion over the cell membrane. Hence, the prodrugs for nucleotide-based drugs are often designed to have increased lipophilicity.[38] A commonly used motif is the ProTide technology that allows the delivery of nucleoside analogues as monophosphates.[39] With these modifications, the bioavailability of the drug increases[40], but the required processing inside the cells adds many steps to the mechanism of action of the drug, which in turn lowers its efficiency.[41] The most crucial step inside the cell in the process of becoming the active drug is the addition of the phosphate groups by intracellular or viral kinases[42], where the first re-phosphorylation step has been reported to be rate-limiting[43]. With the phosphate groups attached, the drug can then interact with the cellular transcription machinery. To develop, design, and predict the efficiency

of nucleic acids-based drug candidates, it is important to have techniques and methods to investigate the natural nucleic acids and their building blocks inside the cells, and to study the drug candidates regarding their cellular uptake and intracellular processing.

2.1.3 Metabolic RNA labelling

The understanding of RNA turnover, *i.e.* the relative rates of transcription and degradation, as well as their regulation, is a key area of basic research in various fields plays an important role in advancing the understanding of diseases. The study of all ribonucleic acids inside cells is known as transcriptomics and today the standard application to analyse this transcriptome is RNA-sequencing.[44] This allows quantification of transcripts, but since RNA has to be extracted from cells it gives no information about RNA dynamics inside cells. Spatial transcriptomics localizes transcriptional activity and monitors RNA dynamics within intact tissue and is typically done by *in situ* sequencing or *in situ* hybridization (ISH).[45] However, this approach is limited by a fixation step, meaning that it cannot capture real-time changes inside living cells. Approaches focusing more specifically on individual transcripts inside cells are among others fluorescent RNA-aptamers or RNA fluorescent *in situ* hybridization (RNA-FISH). Aptamer-readout, on one hand, is live-cell compatible but requires genetic perturbation (sequence insertion).[46] FISH, on the other hand, does not interfere with the genomic material but relies on cell fixation, which, again, limits its temporal resolution.[47] A more straightforward method is to incorporate artificial functionalities into RNA by metabolic processes which allow to study biological processes in a native environment not necessarily focusing on individual transcripts. The combination of introduced modifications with *e.g.* RNA-sequencing, proteomics, or techniques like flow cytometry and cellular imaging is a powerful strategy to study nucleic acids and their building blocks on several levels. Furthermore, the study of nucleosides and nucleotides in metabolism, *e.g.* the quantification of influx and efflux via membrane transporters[48], the specificity for certain polymerases, and mechanistic insights of the catalysis[49], is a fundamental part of RNA turnover research. Hence, labelling the monomers of RNA offers great opportunities.

To label RNA metabolically, ribonucleoside triphosphate derivatives are needed as substrates for RNA polymerases. However, as described previously (Ch. 2.1.2), these negatively charged molecules are not easily internalized by cells. Hence, delivery methods like transfection, microinjection, or chemical masking are used, but come with problems like toxicity, low efficiency, and the requirement of intracellular conversion, respectively. The active uptake of

nucleoside triphosphates by nucleotide transporters has only been reported in bacteria.[50] In eukaryotic cells, only nucleosides can be internalized via nucleoside transporters[51] which requires subsequent re-phosphorylation of the compounds (native or modified) inside the cells — a bottleneck for labelling of RNA.[38,40] Therefore, approaches to enhance the uptake and facilitate the delivery of modified nucleotides are highly sought after. One way of achieving this is to modify cells genomically, for example to overexpress enzymes to enhance the uptake of modified nucleosides. Wang *et al.* showed that by engineering cells to overexpress uridine-cytidine kinase 2 to enhance phosphorylation, the labelling of RNA with different fluorescent pyrimidine analogues is possible.[52,53] Another approach is to design the nucleosides or nucleotides in a way so that they can be internalized by wildtype cells. Chemical designs like the previously described prodrugs (Ch. 2.1.2) are investigated for metabolic labelling applications. These two approaches, *i.e.* engineering of cells and chemical design of the labels, can be combined to develop cell (enzyme)-nucleoside pairs that make metabolic RNA labelling possible for specific purposes. For example, in 2022 Singha *et al.* published a vinyluracil–uracil phosphoribosyltransferase pair that allowed cell-specific RNA labeling.[54]

Besides the chemical design for (enhanced) uptake, the label for the readout must be considered carefully. Different nucleoside/nucleotide labels can be used, depending on the desired redout method, availability of compounds, and experimental setup. Metabolic labelling of RNA has been performed using radioactive or isotope labels[48], but for *e.g.* microscopy applications fluorescence readout is the better choice. The introduction of the fluorescent attribute is often achieved by metabolic incorporation of functionalized nucleosides into RNA and subsequent fluorophore attachment. One approach is the detection and affinity purification using an antibody against the modified nucleoside 5-Bromouridine (BrU).[55] This method is possible to combine with various readout methods, but the dependency on antibody detection limits the applicability.[44] An often-used method for functionalization (including labelling) of nucleic acids as well as their building blocks is bioorthogonal chemistry.[56] For such click-reactions, alkyne- or azide-containing nucleosides are used where the fluorophore of choice can be clicked on post-synthetically by Cu-catalyzed azide–alkyne cycloaddition (CuACC), or Cu-free strain-promoted azide–alkyne cycyloaddition (SPAAC).[57] The latter approach is particularly important for labelling of long RNA as it is compatible with biological systems, *e.g.* labelling inside living cells.[44] Different analogues are available and allow great flexibility in the experimental setup, *e.g.* 4-thiouridine (4sU)[58] or 5-ethynyluridine (EU)[59]. Recently, using SPAAC chemistry, Wang *et al.* presented cell- and polymerase-specific labelling of RNA by

feeding azidocytidine to cells that are overexpressing deoxycytidine kinase.[60] However, all of these methods always require a second step where the actual label, *i.e.* the fluorophore, is attached. Furthermore, the cells usually need to be fixated, or the reactions proceed with slow kinetics, limiting the applicability in living systems. To overcome these limitations the use of building blocks that are intrinsically fluorescent and label RNA directly upon incorporation is emerging.

2.1.4 Fluorescent Dyes for Nucleic Acids

Fluorescence-based techniques are very popular and widely applied due to their ability to study biomolecules in native conditions and even inside living cells. Since natural nucleic acids are virtually non-emissive, fluorophore labelling is necessary to allow their study using fluorescence-based methods. The fluorophore can be attached non-covalently (*e.g.* the intercalating dye ethidium bromide[61]) or, for more specific and controllable labelling, covalently. The covalent modifications can be divided into external and internal modifications. External modifications are fluorophores attached by linkers to the nucleic acid, without being involved in the actual base stack.[62] These fluorophores can be incorporated synthetically by solid-phase polymerization, or they can be attached to nucleic acids after introducing a building block that contains a reactive "handle," onto which the fluorophore is subsequently coupled (Ch. 2.1.3). Almost any commercially available dye can be used, for example fluorescein, cyanine-, or Atto-dyes. The photophysical characteristics, *i.e.* the brightness, of these dyes often allows for single-molecule detection, but their size and hydrophobic linkers can perturb the structure and interaction of nucleic acids with other (bio-)molecules.[63] Furthermore, the dye is positioned outside the base stack which limits the analysis of local structural information. In this thesis the focus lies on internal modifications, where a fluorophore replaces a natural nucleobase and therefore gives an "on-site" readout from a small modification. Several planar aromatic fluorophores have been developed and incorporated into nucleic acids, *e.g.* pyrene or phenanthrene.[64] However, their inability to form hydrogen bonds with the pairing base on the opposite strand reduces the utility to study native structure and behaviour of nucleic acids.[62] Therefore, the research field of fluorescent base analogues (FBA) has emerged. The approach is to create molecules that are bright, photostable, absorb light outside the nucleobase absorption band (> 300 nm), and have a large Stokes' shift (difference in energy between absorption and emission maxima). Together with the demands of resembling the natural bases, *i.e.* form Watson-Crick base pairs, and to not perturb the natural structure and interaction of the nucleic acids, the design, characterization, and eventual application is a complex process.[62,65]

Many FBAs have been developed and applied to study, for example, nucleic acid interactions and structures to address questions in biology, biochemistry, and biophysics of nucleic acids.[65–67] The adenine analogue 2-aminopurine (2-AP) was one of the first FBAs, published in 1969, and is today widely used for labelling of nucleic acids (Fig. 2).[68] The fluorescence properties of many FBAs, including 2-AP, are sensitive to their microenvironment. This can be used to monitor changes in the local environment of the fluorophore, which in turn can be applied as a tool to, for example, study DNA-protein interactions.[69] However, for other applications, *e.g.* in Förster resonance energy transfer (FRET) or fluorescence anisotropy measurements, environment insensitive probes are preferred to more easily obtain quantitative results. A representative FBA of this group is the tricyclic cytosine analogue tC (Fig. 2). This molecule was developed in 1995 and in 2001 Wilhelmsson *et al.* also found that it is highly fluorescent inside nucleic acids while being virtually micro-environment insensitive.[70–72] From this starting point, new FBAs have been developed and characterized and the herein presented work is based on previously designed, characterized, and applied FBAs from the Wilhelmsson laboratory. The tricyclic cytosine analogue tC^O (1,3-diaza-2-oxophenoxazine, Fig. 2)) is the oxo-homologue of the mentioned tC, that was developed in 1995, and characterized as an FBA by Sandin *et al.* in 2008.[70,73] It has been and still is applied in structural studies of nucleic acids[74–76] (paper not included in thesis), to investigate interactions *e.g.* between DNA and proteins[77] or small molecule binding to nucleic acids[78], DNA dynamics[79], and to label DNA enzymatically[80]. Furthermore, two analogues from the adenine family play a major role in this thesis: the cyano-modified quadracyclic adenine 2CNqA and the pentacyclic adenine pA (Fig. 2). 2CNqA is, compared to its mother compound qA[81], significantly less sensitive to its microenvironment.[82] It is so far the brightest qA derivative and has been characterized also in an oligonucleotide context where it was found to be one of the brightest reported FBAs in duplex DNA.[83] Furthermore, it was successfully used as a donor for FRET between FBAs to study the structure of both DNA and RNA, thereby constituting a promising tool for analysing conformations of nucleic acids[83] (paper not included in thesis). pA is an extension of the qA scaffold, afforded by incorporating an additional benzene ring.[82,84] It has been thoroughly characterized as an FBA and its versatility was demonstrated by its use as a donor for interbase FRET in oligonucleotides, as a successful two-photon fluorophore, and as a label in microscopy.[84,85] All three FBAs — 2CNqA, pA, and tC^O — have been shown to be useful fluorescent labels for gapmers for live-cell imaging and flow cytometry readout.[86] A key development for this thesis was published in 2021, where tC^O was synthesized as its nucleoside triphosphate derivative for the first time and used to enzymatically label mRNA.[87] With that, the synthesis strategy for

FBA-triphosphates and investigations of FBAs for enzymatic and metabolic labelling built the foundation for future biochemical applications.

Besides research on already established FBAs from our group, the herein presented work also includes the characterization of a new base analogue inside RNA oligomers. The synthesis and photophysical characterization of the quadracyclic uridine monomer (qU, Fig. 2), *i.e.* as a nucleobase and a ribonucleoside, was published 2024.[88] As a monomer, qU has high brightness and absorbs in the visible region, which makes it an interesting candidate for biological applications and readouts. Furthermore, it shows an equilibrium between two spectrally separated tautomeric forms and is pH sensitive, making it an interesting probe for live-cell imaging in various environments inside cells.

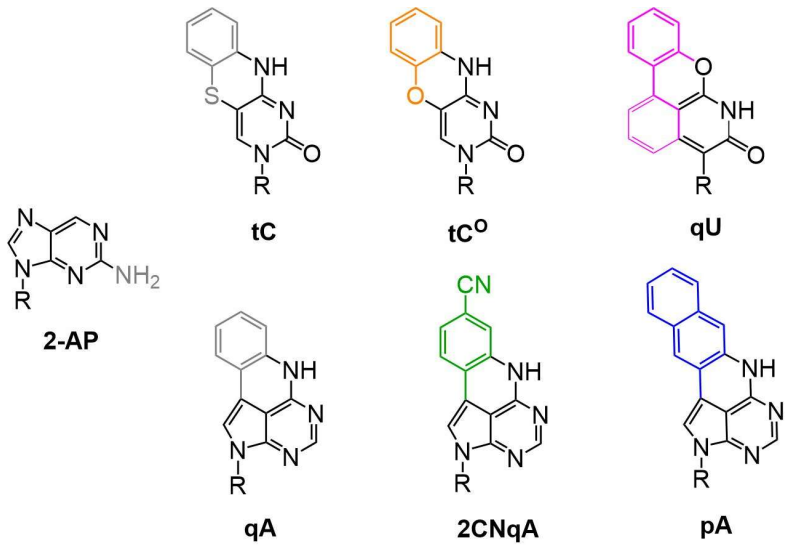

Figure 2. Structures of FBAs that are described in the text. The natural scaffold is drawn in black, while the FBAs important for this thesis are highlighted in bright colours. R = attachment site for the sugar moiety of the corresponding nucleoside.

The overall aim of FBA development is to enlarge the toolbox for nucleic acid labelling for various, but specific, needs and, thus, complementing the variety of techniques already available to study nucleic acids.

2.2 Methods

This thesis explores chromophores/fluorophores in biological environments and reactions and therefore, a solid understanding of the methods used is required. In this chapter, the fundamental principles of the main techniques are introduced and further explained in the light of the research presented.

2.2.1 Light Absorption by Nucleic Acids

When atoms form molecules, their atomic orbitals combine to create molecular orbitals. Each molecular orbital can hold up to two electrons with opposite spins. In the ground state of a molecule, these orbitals are filled with electrons, starting from the orbital with the lowest energy. If all occupied molecular orbitals contain paired electrons, the molecule is said to be in a singlet state. However, if there is one unpaired electron, the molecule is in a doublet state, as seen in radicals. The electronic and vibrational states of a molecule can be illustrated in a Jabłoński diagram where the transitions between them become easily understandable (Fig. 4). Overall, the total energy of a molecule consists of translational, rotational, vibrational, and electronic energy, where each electronic state has several vibrational levels, which in turn have rotational levels. In the Jabłoński diagram, electronic states with a few of their vibrational levels are depicted. Light can be absorbed by a molecule, causing it to transition from a low-energy to a high-energy state, provided that the energy of the light matches the energy gap between the molecule's electronic levels and that the molecule is oriented correctly. This molecule is thereafter in an excited state. The molar absorptivity (ε in $M^{-1} \times cm^{-1}$) describes how much light a molecule can absorb at a specific wavelength by expressing the probability of excitation from the ground state to an excited state (or more general from a lower to a higher energy state).[89] Light with energy in the UV and visible region (vis) is generally needed for transitions between electronic states. In nucleic acids, the chromophores that absorb in the UV region are the nucleobases, while the backbone, consisting of sugar and phosphate groups, absorbs at even higher energy (< 200 nm[90]). Each base has an extensive conjugated π electron system which gives rise to intense π-π* transitions with absorption peaks in the range of 150 to 300 nm.[91] The absorption spectra of DNA or RNA polymers present an average spectrum with respect to the base composition and degenerate interactions, with the most characteristic peak maximum appearing at around 260 nm.

Non-degenerate interactions, like hybridization or separation of nucleic acid strands, can change the intensity of the band at 260 nm. As two strands dissociate and lose interactions the bases become less stacked and the intensity at 260 nm increases, known as hyperchromicity (with the reverse phenomena known as hypochromicity). This is useful for example in determining the melting temperature, *i.e.* the separation of double strands or other secondary structures that are stemming from base-pairing. The melting temperature T_m is defined as the temperature where half of the double strands are denatured, *i.e.* dissociated into two single strands. This is dependent on the type of nucleic acid, its sequence (*e.g.* the GC-content), the

solvent, the salt concentration, the pH, and possible interactions with other molecules.[92] In the presented work, the melting temperature is an important parameter for investigating the effect of fluorescent base analogues on base-pairing and -stacking, compared to the corresponding natural base (Ch. 3.1.1 *vide infra*).

2.2.1.1 Circular Dichroism

Another technique that is important for this thesis work, and that is based on the absorption of light by molecules, is circular dichroism (CD). In dichroism, the absorption of light by molecules differs depending on the polarization direction of the incoming light beam. In circularly polarized light the magnitude of the electric field vector is constant, and the direction is modulated, leading to a helix shape along the direction of propagation. This means that two orientations are possible: a right-handed and a left-handed helix. In measuring CD, the difference in absorption between these two polarizations is investigated and allows conclusions about the 2D- and 3D-conformation of (bio-)molecules.[91] To generate a CD signal, the solution must contain optically active molecules, that are either inherently chiral or rendered asymmetric by their surrounding environment. Like most biomolecules, nucleic acids are asymmetric due to the chiral sugar in the backbone, leading to a weak CD signal in the π- π* region of the nucleobases. Nucleic acids form helical structures, which furthermore have a super asymmetry. This leads to electronic interactions between the chromophores (*i.e.* nucleobases) resulting in a strong CD signal that is characteristic for each conformation. Hence, the recording of CD spectra allows to draw conclusions about the secondary structures of nucleic acids.[93] In the presented work, this was particularly important for the investigation of the effect on the helical structure when FBAs are incorporated into oligomers (Ch. 3.1.1 *vide infra*). Since both right- and left-handed circularly polarized light follow Beer-Lambert Law, CD is defined as the difference in their extinction coefficients $\Delta\varepsilon$ (Eq. 1 with A absorption, l light pathlength, c concentration of the solution).

$$\Delta A_{(\lambda)} = A_{(\lambda),Left} - A_{(\lambda),Right} = \left[\varepsilon_{(\lambda),Left} - \varepsilon_{(\lambda),Right}\right] \times l \times c = \Delta\varepsilon \times l \times c \qquad \text{[Equation 1]}$$

The light leaving the sample is hence elliptically polarized, due to the superimposition of right- and left-handed circularly polarized light that was absorbed differently. In this work, CD is expressed as ellipticity θ (unit: mdeg), which describes the angle between the maximum and minimum of the magnitude of the electric field vector (E) of the elliptically polarized light that leaves the sample (optical rotatory dispersion, Fig. 3)).[94]

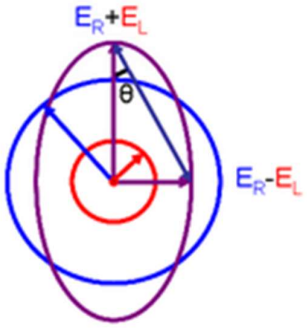

Figure 3. Constitution of the ellipticity of the superimposed light in CD. Circles show the electric field vectors of the left (red) and right (blue) circularly polarized light when looking along the axis of propagation. The resulting superimposed light (purple), that leaves the sample is elliptically polarized and is described by the angle θ as shown.

2.2.1.2 Multiphoton Absorption

An additional light-matter interaction is multiphoton absorption, which is based on the *simultaneous* absorption of multiple photons by a molecule without an intermediate excited state. This can also be depicted in a Jabłoński diagram (Fig. 4). Neither photon is at resonance with the energy states of the system, but the combined frequency matches the resonance of an electronic energy state of the system. Since it requires spatial and temporal overlap of two or more photons, the occurrence of this process is highly unlikely in nature, but it can be achieved in a research environment. For two-photon absorption (2PA, TPA) the probability of photons absorbed per time unit depends on the square of the intensity of the incoming light, making the process non-linear. Therefore, the cross section is not intuitively understandable (as in one-photon absorption by cm^2) and was first described by Maria Göppert-Mayer[95], in whose honour the unit of the two-photon cross section (GM) was named, with $1\ GM = 10^{-50}\ cm^4 s/_{photon}$.

To facilitate the simultaneous absorption of two photons, a pulsed laser system delivering a high photon flux is needed to provide the required laser power in short "packages" with high peak power, while the average laser power is comparably low. The 2PA in this thesis is important for microscopy purposes, where the focal point (determined by the objective lens) is the point where the photons are "crowded" enough (Ch. 2.2.3.2 *vide infra*).[96] We characterized one FBA as two-photon probe in solution (Ch. 3.1.3 *vide infra*) and subsequently applied it in two-photon fluorescence imaging of living cells (Ch. 3.2.3 *vide infra*).

2.2.2 Fluorescence

After a molecule has absorbed light and has been excited, it will eventually return to its ground state. Most often this happens by transferring the energy as heat to the surroundings — referred to as non-radiative decay. In some cases, other processes where photons are emitted, can take place, namely fluorescence and phosphorescence. Since this work is based on fluorescent base analogues the first process is of interest, but both processes are displayed in Fig. 4. In short, phosphorescence includes a step of intersystem crossing (ISC) to an excited triplet state (two orbitals with unpaired electrons, T_1) from where the molecule can relax by emitting a photon. In fluorescence, the emission generally comes from relaxation from the lowest vibrational level of the first excited singlet state (S_1). Notably, irrespective of whether the excitation has taken place with one or several photons, the de-excitation pathways are the same processes.

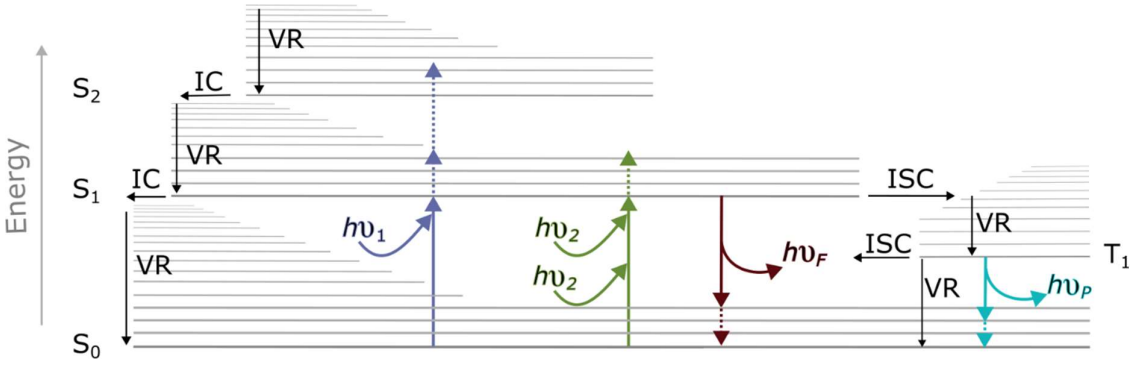

Figure 4. Jabłoński diagram showing the photophysical processes described in this thesis. Non-radiative processes are indicated in black arrows including vibrational relaxation (VR), internal conversion (IC), and intersystem crossing (ISC). One-photon absorption is indicated in purple, 2PA in green, fluorescence emission in dark red, and phosphorescence emission in cyan. S indicates singlet state and T triplet state. Dotted arrows indicate the possibility of processes going to different vibrational levels.

2.2.2.1 Steady-state Fluorescence

To describe fluorescence, it is helpful to look at the rates of the different processes involved, *i.e.* the radiative decay k_r (emission of a photon) and the non-radiative processes k_{nr} which involve internal conversion (IC) and the following vibrational relaxation (VR) as well as ISC (Fig. 4). Other processes that decrease the fluorescence intensity are static and dynamic quenching (k_q), which depend on the concentration of a quenching species [Q]. Altogether, the efficiency of the fluorescence process is measured as the fluorescence quantum yield (ϕ_f),

which is the radiative rate constant divided by the sum of all rates, *i.e.* describing the number of absorbed photons in relation to the number of emitted photons (Eq. 2).

$$\phi_f = \frac{photons\ emitted}{photons\ absorbed} = \frac{k_r}{k_r + k_{nr} + k_q[Q]} \quad \text{[Equation 2]}$$

In this work the quantum yields are determined in relation to a reference compound with known quantum yield[97] by Eq. 3, which takes into account the measured integrated emission I, the absorption A at the excitation wavelength, and the refractive indices η of the solutions. S and R refer to sample and reference, respectively.

$$\phi_{f,S} = \frac{I_S/A_S \eta_S^2}{I_R/A_R \eta_R^2} \phi_{f,R} \quad \text{[Equation 3]}$$

The quantum yield multiplied by the molar absorptivity results in a value on the brightness (with the unit $M^{-1} \times cm^{-1}$) of a compound, which is an important parameter for detectability, and hence of interest for *e.g.* microscopy or flow cytometry applications. To measure steady-state fluorescence, meaning collecting the average emission of randomly oriented molecules in solution, a fluorimeter is used. An excitation monochromator selects the wavelength of the excitation light that hits the sample. Emission is typically detected at a right angle (to avoid detecting excitation light) after passing the emission monochromator, which selects the detection wavelength. To record an emission spectrum the excitation monochromator is fixed, and the emission monochromator scans the emission wavelength range. This gives information on the relative intensity of the emission at the scanned wavelength. In contrast, if the emission monochromator is fixed and the excitation monochromator is scanning, an excitation spectrum is recorded. This provides information regarding the absorption of the emitting species in the solution. All three spectra — absorption, excitation, and emission — give useful information in the characterization of fluorophores.[89]

2.2.2.2 Time-resolved Fluorescence

The average time a molecule spends in the excited state is defined as the inverse sum of rates of the processes that depopulate the excited state (Eq. 4) and is called fluorescence lifetime (τ).

$$\tau = \frac{1}{k_r + k_{nr} + k_q[Q]} \quad \text{[Equation 4]}$$

The intrinsic lifetime of a molecule is the time the molecule would spend in the excited state in the absence of non-radiative processes. Emission is a random process where each excited fluorophore has the same probability to emit a photon at a given timepoint t, meaning that some molecules emit at $t = \tau$, while others send out a photon at $t > \tau$ or $t < \tau$. This leads to an exponential decay of the emitting population (Eq. 5), with I_0 being the intensity at $t = 0$.

$$I(t) = I_0 e^{-t/\tau} \quad \text{[Equation 5]}$$

Time-correlated single photon counting (TCSPC) is a commonly used technique for determining fluorescence lifetimes. A pulsed laser source excites the sample, which emits photons that are detected at a right angle in a time-dependent manner. The heart of a TCSPC instrument is the time-to-amplitude converter (TAC), recording the time between the excitation pulse and one emitted photon arriving on the detector. After many signal periods (excitation pulses and corresponding arriving photons on the detector) a large number of photons are collected in time bins (channels) displaying a distribution over the detected time resulting in a histogram representing the decay of the sample (Fig. 5).[98]

Since the detector registers only the first photon arriving, the decay is biased toward shorter lifetimes. To avoid this effect, no more than 1 photon per 100 laser pulses (*i.e.* detection rate of 1%) is measured.[89] Because the excitation pulse is not indefinitely sharp and no detector is perfect, the instrument pulse (or instrument response function, IRF), must be monitored and accounted for during analysis. This is because the measured decay presents the convolution of the IRF and the actual fluorescence decay of the sample.

For analysis, a mathematical model which is convolved with the IRF and compared to the actual decay is used. By iterating this re-convolution process, the parameters that describe the measured decay best are found. As described above, for fluorescence lifetime an exponential decay model is applied. Multiexponential decays are described by the sum of several single exponential decays (Eq. 6), where I is the intensity at time t, α is the pre-exponential factor describing the amplitude of the different components at $t = 0$ (sum normalized to unity), τ is the lifetime, and n is the number of components.

$$I(t) = \sum_{i=1}^{n} \alpha_i {}^{-t/\tau} \quad \text{[Equation 6]}$$

Since almost any decay can be fitted to this model by using a sufficient number of exponentials, the aim when analysing TCSPC data is to use as few exponentials as possible to describe the decay best, in accordance with the proposed mechanism or research hypothesis.[89] In this thesis,

the amplitude-weighted average lifetime (< τ >) for fluorophores with multiexponential decay (Eq. 7 as an example of a bi-exponential decay) is in the foreground, since it is proportional to the steady-state intensity.

$$< \tau >= \int_0^\infty I(t)dt = \alpha_1\tau_1 + \alpha_2\tau_2 \quad \text{[Equation 7]}$$

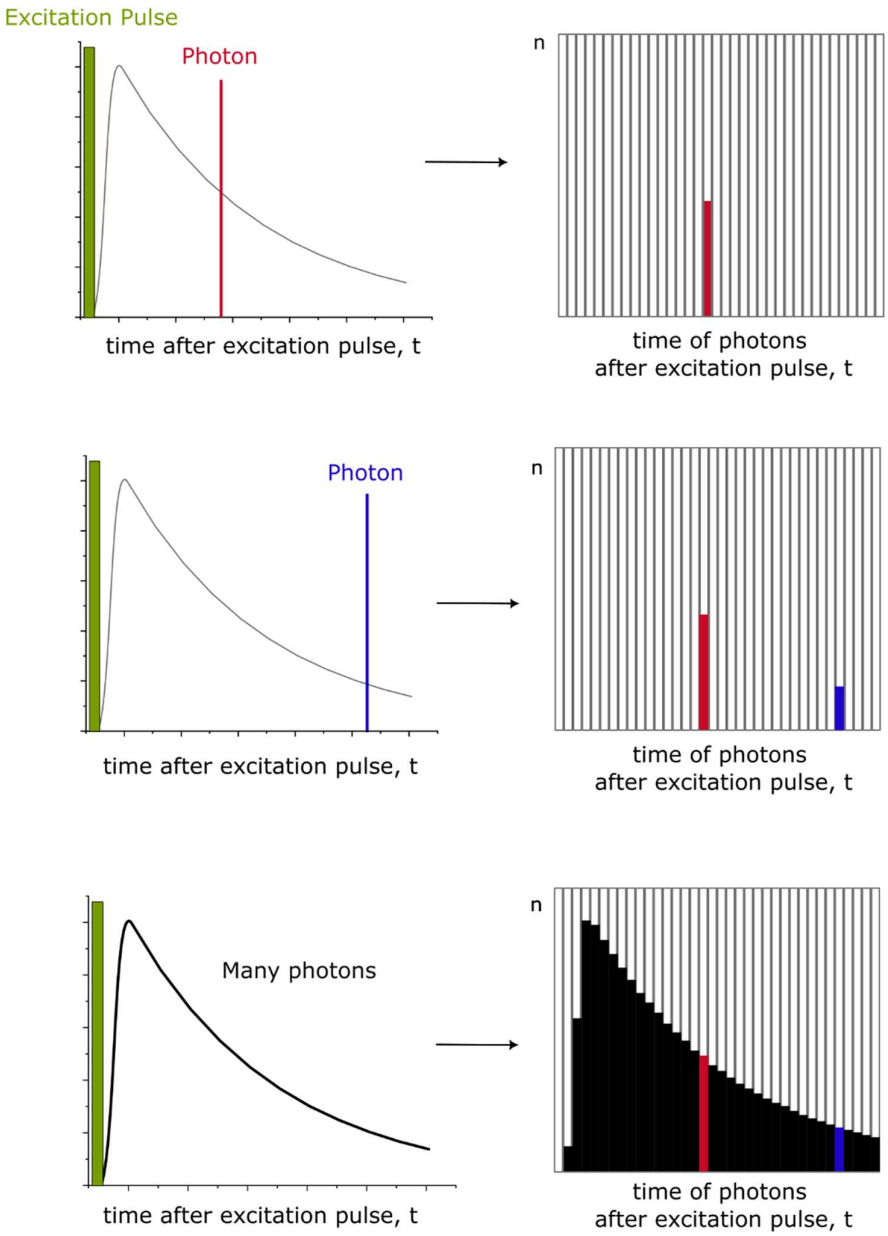

Figure 5. The recording principle of TCSPC. The first photon (red) is stored in a channel assigned to the arrival time at the detector. In the second signal period the process is repeated (blue photon). After collecting many photons (black), the histogram representing the decay is created. Figure modified after Becker 2017.[98]

2.2.3 Light Microscopy

The motivation to enlarge the small world in order to understand processes that are invisible to the naked human eye is very old and dates back at least to the 13th century.[99] Microscopy in biology gave us the insight to learn about bacteria, eukaryotic cells, cellular structures and is still making breakthrough at all scales of biology. With more techniques and methods at hand, microscopy is an important tool to watch biochemistry in action and is expanding into the world of nanoscopy.

Apart from magnifying glasses, the simplest light microscopes are compound microscopes. These use a primary lens (objective) to produce an enlarged real image of the object, which is then further magnified by a second lens (ocular) to create a virtual image that can be observed on the retina of the eye. These microscopes are called bright-field microscopes, since visible light is passing through the sample giving contrast by attenuation of transmitted light due to differences in the density in the sample. The resulting image is seen as a dark sample on a white background, hence the name. To enhance detectability or visibility, staining samples or parts of the samples is a common method. However, depending on the research question, this might not be sufficient, for example when examining behaviour of specific molecules inside cells. In such cases, introducing fluorescent labels to the sample or to particular molecules, and using fluorescence microscopy, can be a more suitable approach. Fluorescence microscopy provides high contrast, high specificity, and the possibility of quantification of even live-cell images. The use of fluorophores requires an excitation light source and a detection system for the emitted light. In fluorescence microscopes this is often provided by filters to select the desired wavelengths. The simplest fluorescence microscopes are therefore imaging in epi-fluorescence mode, where the excitation light is illuminating the sample through the objective and the emitted light is focused through the same objective onto the detector. Main limitations are out-of-focus light that is detected from all planes of the sample, and damage to the whole sample by the excitation light (*i.e.* bleaching), which is especially important to consider in a live-cell context (*i.e.* phototoxicity).

2.2.3.1 Confocal Microscopy

A more advanced technique circumventing the limitations of imaging in epi-fluorescence mode is confocal laser scanning microscopy (CLSM).[100] The key components in these microscopes

are the insertion of a pinhole at the focal point of the objective lens, which leads to efficient blocking of out-of-focus light, and the use of lasers as illumination source, *i.e.* creating point illumination. Only light emitted from fluorophores close to the focal plane is detected, enhancing the optical resolution. The use of lasers requires a scanning mode, where the laser is moved across the sample by scanning mirrors. Because the emitted light from the sample travels an identical path "back" through the scanning system, the beam focussed on the detector aperture is inherently descanned by the mirrors. This descanned signal is detected by photomultipliers (instead of cameras) and the actual image is reconstructed in the connected software. This approach comes with high flexibility in the setup of the image, for example the choice of the pixel size or zoom, which can be of importance for downstream image analysis. Because only a thin plane is imaged, optical sectioning of thick samples and computational reconstruction of 3D images is possible.[101] Drawbacks of CLSM are low light efficiency, since much is blocked by the pinhole, and slow imaging, due to the scanning mode. More advanced technical developments and methods like spinning disk confocal microscopy solve parts of these drawbacks.[102]

2.2.3.2 Two-photon Microscopy

The use of fluorophores in microscopy comes with more advantages, based merely on molecular photophysics. In the context of this thesis, two previously discussed parameters are of importance: multiphoton excitation and fluorescence lifetime. Two-photon imaging uses fluorophores with an appropriate two-photon cross section to excite them by two simultaneously absorbed photons (Ch. 2.2.1.2). In the context of imaging, this excitation mode comes with several advantages: biological samples, like tissues, scatter light with a higher scattering coefficient in the shorter wavelength range. Since two-photon excitation uses wavelengths in the near-IR region, the penetration depth is greater. Furthermore, major intracellular absorbers (*e.g.* water, melanin) absorb very little light in the near-infrared range.[103] Due to pulsed excitation, the photons are "crowded" enough in only a very small focal volume (femtoliter-range), leading to emission from a very small part of the sample, efficiently reducing out-of-focus light. A pinhole is therefore unnecessary and can even be detrimental, since the emitted light is of shorter wavelength than the excitation beam and emitted photons are more prone to scattering and should not be lost due to a pinhole. Instead, non-descanned detection is preferred.[104] Since only a very small volume is exposed to the excitation light of long wavelengths, the risk of photobleaching and phototoxic effects is greatly reduced, making multiphoton imaging very useful for biological applications.[105]

2.2.3.3 Fluorescence Lifetime Imaging

So far, only steady-state emission has been described for imaging purposes. However, the information regarding the lifetime of fluorophores in samples can be monitored in a spatial manner using fluorescence lifetime imaging microscopy (FLIM). FLIM is based on TCSPC measurements, as previously described (Ch. 2.2.2.2), whereby the TCSPC module in a FLIM microscope also receives information from the scan controller. This means that every detected photon comes with the time information in the decay, but also information on pixel, line, and frame.[98] This efficiently results in an image containing a fluorescence decay curve in each pixel. Now not only can the steady-state intensity be displayed with spatial resolution, but also the lifetime, which can be very useful to study different fractions of the same fluorophore in different microenvironments. Since the fluorescence lifetime depends on the molecular state of interaction, but not on the fluorophore concentration, molecular effects can be studied in samples with varying fluorophore concentrations.[106] The analysis of lifetime images may be done by two approaches. Since each pixel contains a fluorescence decay an exponential fit can be applied (Ch. 2.2.2.2) and the images can be displayed in colour code so that the average lifetime of each pixel is visualized. Another way of analysis is the phasor approach which operates in the frequency domain. Instead of fitting the decay, a Fourier transformation is applied to the decay data in each pixel.[107] In the phasor plot, for each pixel a pointer ("phasor") is defined, where the transformation expresses the decay as two parameters: the phase shift (φ) and the modulation (m), which represent the phase and amplitude of the Fourier components, respectively (Fig. 6).[98]

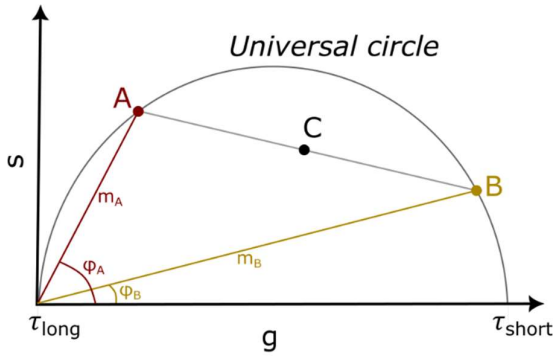

*Figure 6. Phasor plot. The phasor coordinates are the real (g) and imaginary (s) parts of the Fourier transformation. Mono-exponential decays fall on the universal circle (A, B), while multiexponential decays (C) are located inside. Phasors of combinations of several components are linear combinations of the phasors of these components (*e.g. *C as combination of A and B). Figure modified after Becker 2017.*[98]

Each fluorescence decay corresponds to a unique position in the phasor plot, which depends on its average lifetime or lifetime distribution. One key feature of the phasor plot is its reciprocity, meaning that it is possible to map back from a position in the phasor plot to the corresponding pixel in the original image. Instead of determining fluorescence lifetimes and fitting decay components, the phasor approach aims to separate and identify fluorophores (or fluorophore fractions) based on their distinct lifetime signatures.

The herein described microscopy techniques are just scratching the surface of the microscopy world. A wide variety of optical imaging methods are available, many of which can be combined, *e.g.* multiphoton FLIM. This diversity of techniques makes imaging a vital part of scientific research.

2.2.4 Flow Cytometry

Flow cytometry is a laser-based technique that is used to analyse cells or particles in suspension. Multiple parameters can be measured at once, as single cells pass through a laser array in a fluidics system. The main readout parameters are relative cell size, internal cellular complexity and fluorescence intensity. This allows for the evaluation of cell distributions, cellular subtypes, and populations. With that, flow cytometry finds wide applications, ranging from cell biology research to diagnostics in clinics.[108]

The cell suspension is injected into a microcapillary flow system by a needle and is then hydrodynamically focused, allowing every cell to pass by a laser array one at a time. The system detects the forward scattered light (at minimum of 2˚ angle to avoid direct laser light detection[109], FSC), the sideward scattered light (at 90˚ angle[109], SSC) and light emitted from excited fluorescent species. The FSC signal provides information about the cell size, while the SSC signal gives information about the internal complexity or granularity of the cells allowing assessment of cellular structural features.[110] When a cell passes by the laser it results in a signal (SSC or FSC) over the time passing, which is detected and defined by its peak height, width, and area. These parameters can be plotted in dot plots, where every detected cell is represented, and then used for gating of cell populations. In Fig. 7 a typical analysis is shown with the aim to analyse fluorescence signals of single living cells (hepatocytes Huh-7, Ch. 2.2.6 *vide infra*), meaning excluding dead cells, cell debris, or cell clusters. Hence, the population of single living cells (Fig. 7, blue) can be selected and analysed for fluorescence signals using different laser lines, for example from a fluorescence protein or a fluorescence nucleobase analogue.

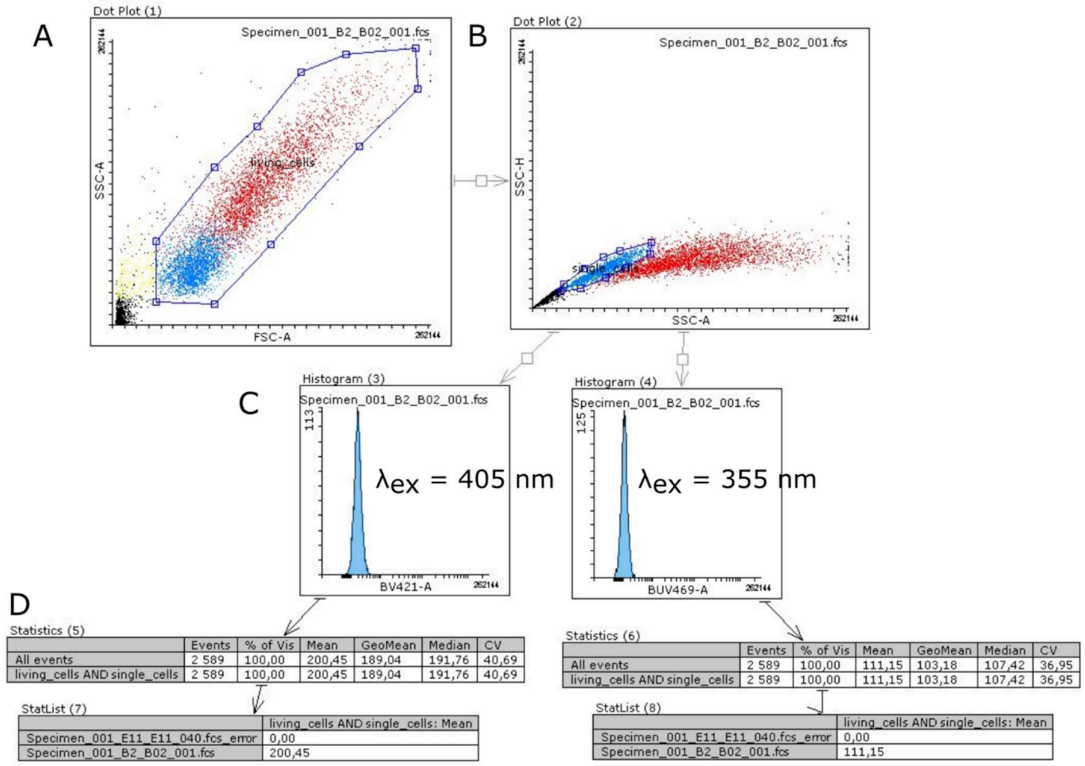

Figure 7. Example of a typical workflow in Flowing Software for flow cytometry data analysis. A) Dot plot of the peak area (-A on axis) of the detected side scatter signal (SSC) and forward scatter signal (FSC) allowing for gating of living cells population (aspect ratio 1, red and blue population). B) Dot plot of height and area (-H and -A on axis, respectively) of the SSC signal allowing for subsequent gating of single and living cells (aspect ratio 1, blue population) for further analysis. C) Histograms of the fluorescence signal from two laser lines (405 nm excitation and 355 nm excitation) of the gated cell population (blue population). D) Statistical analysis of the histograms, including for example total event counts, mean fluorescence intensity, and median fluorescence intensity. Note that these are random data and therefore the numbers do not have scientific significance.

Flow cytometers provide a wide combination of excitation wavelengths and emission filters, fitting most common dyes. However, detecting novel fluorophores can require adjustment of the instrument setup and consideration of customized features. In general, the selection of fluorophores is based on the biological question, *e.g.* which biomolecule to label, what fluorophores are available for a specific biomolecule and how the labelling is done, as well as considerations of spectral features of the fluorophores, especially when used in combination.

The development of flow cytometry instrumentation and the proof of utility of the fluidics system has allowed for the analysis of small particles in the range of 100–1000 nm, like extracellular vesicles, or virus particles (flow virometry).[111]

2.2.5 Long RNA

2.2.5.1 Synthesis and Purification

Solid-phase synthesis of nucleic acids is a widely used method to produce DNA and RNA. However, this synthetic cycle, where nucleoside phosphoramidites (natural as well as modified) are added stepwise according to the desired sequence, is limited to around 100 nucleotides.[112] For long RNA molecules, enzymatic *in vitro* transcription (IVT) is the method of choice. A DNA-dependent RNA polymerase is the enzyme that assembles the nucleotides following a DNA template sequence. RNA polymerases used in the laboratory often stem from bacteriophages, with the most common one, and used in this thesis work, being the T7-polymerase, which consists of a single subunit of about 100 kDa.[113] The DNA template contains several important elements: a promoter sequence, which serves as the transcription initiation signal recognized by the polymerase; a sequence of interest, such as one encoding a reporter protein; and additional regulatory elements, including the Kozak or Shine-Dalgarno sequences that facilitate translation initiation, as well as untranslated regions (UTRs). The resulting RNA transcripts are used in analytical techniques, structural studies, biochemical and genetic studies, or as functional molecules[114], *e.g.* as messenger RNA (mRNA). For RNA to become mRNA two more enzymatic reactions are needed: 5'-capping and 3'-tailing. In the capping reaction, an enzyme adds a 7-methylguanosine to the 5'-end of the transcribed RNA via a 5'-5' triphosphate linkage (known as Cap-0 structure).[115] In eukaryotic systems this cap plays a major role in translation initiation, prevents the degradation of the RNA by exonucleases, and has other regulatory functions. Other capping groups are found in nature and artificially synthesized groups are explored to enhance downstream protein synthesis, increase stability, and reduce immune responses.[116] The tailing step of RNA is the addition of a polyadenine (poly(A)) tail to the 3'-end. This tail, together with the cap-structure, is a crucial part for translation initiation in eukaryotes (ribosome assembly)[117] and regulation of mRNA decay[118]. Since non-template dependent addition of poly(A) results in heterogenous length of RNA, a poly(T)-tail is often encoded in the DNA template for research and medical use.[116] After the production of (m)RNA, thorough purification is needed to ensure clean RNA of the desired length. When choosing RNA purification methods, several parameters need to be considered

— such as the type of RNA to be purified, the amount of RNA, its origin (*e.g.* produced in an *in vitro* reaction or extracted from cells), the degree of protocol control or modification, and intended downstream applications. On a research scale, purification using standard extraction (phenol:chloroform:isoamyl alcohol) and precipitation (ethanol or LiCl) methods, polyacrylamide gel extraction, or silica-based spin columns are most common. When working with larger scale, purification using solid-phase extraction or liquid chromatography are widely used.[119]

2.2.5.2 Analysis and Quality Control

To analyse the final RNA product, two main methods were used in the herein presented work. One is spectral analysis, based on the absorption of light by the nucleobases, with the characteristic peak centred around 260 nm (Ch. 2.2.1). A clean peak allows to draw conclusions about the purity of the sample and guideline values for purity based on absorption ratios where contaminants absorb are available ($^{A_{260}}/_{A_{280}} \approx 2$ for protein contamination and $^{A_{260}}/_{A_{230}} \approx 2 - 2.4$ for organic compound contamination[118]). By knowing the molar absorptivity of the RNA strand at 260 nm (or for mixed sequences an average value), the concentration of a nucleic acid solution can be determined by just measuring the absorption at 260 nm and applying Beer-Lambert Law (Eq. 8, with I light intensity, I_0 the intensity of the incident light, *i.e.* of the reference beam, ε molar extinction coefficient, l light pathlength, A absorption).

$$log_{10}\left(^{I_0}/_{I}\right) = A = \varepsilon \times c \times l \quad \text{[Equation 8]}$$

However, absorption does not provide conclusions on the integrity of the RNA polymer, since it is based on the per-base-absorption. Electrophoretic analysis can be employed to assess the size distribution and homogeneity of RNA samples and was herein used as second method of analysis. Gel electrophoresis is a chromatography technique were charged molecules (like nucleic acids due to their negative phosphate backbone) are separated through a cross-linked polymer gel matrix when placed in an electric field. By comparing to a standard ("ladder"), the molecular weight or length of the RNA molecules in a sample can be estimated.[120] In this work agarose gels were used, but polyacrylamide gels are also very common in the field.

To achieve high-quality and pure RNA for downstream applications, many parameters must be considered and tested. This includes all steps — from the design of the DNA template to the optimization of the *in vitro* transcription reaction, the post-transcriptional modifications, and purification.

2.2.6 Cell Culture

Cell culture refers to cells that are grown under controlled conditions outside their natural environment and hence provides an important tool as model system to study biochemistry and biology. Effects of toxic compounds or drugs can be investigated, as well as mutagenesis or carcinogenesis. Drug screening and development are based on cell cultures, since they provide consistency and reproducibility in combination with easy accessibility, low costs, relatively easy handling and fast turnaround compared to *in vivo* experiments.[121] The importance and ethics of cell lines are highlighted in the story of Henrietta Lacks[122], from whom the first cells were cultured and which still play a major role in research. In this thesis, immortal human cell lines, *i.e.* tumorous cells that do not stop dividing, are used.

In vertebrates the liver is the central organ of metabolism, as it acts in the breakdown and excretion of metabolites and exogenous substances, including therapeutic and toxins. Therefore, the major cell line used in this thesis work is a well differentiated hepatic carcinoma cell line with an epithelial-like morphology, named Huh-7, that partially retained the hepatic functions.[123] These cells were taken from the liver tumour of a Japanese male in 1982.[124]

Besides the molecular study of the fluorescent base analogues in a cellular context, cell culture allows to easily investigate the toxicity of these compounds by standardized assays probing for example for metabolic activity (herein used AlamarBlue assay[125]) or membrane integrity (lactate dehydrogenase leakage (LDH) assay[126]).

3. MAIN WORK

This chapter portrays the work of the scientific articles that this thesis is based on. **Paper I** presents the thorough characterization of our adenine analogue 2CNqA as a single- and multiphoton probe and the subsequent use of this molecule when incorporated into antisense oligonucleotides (ASOs) in two-photon live cell imaging. 2CNqA also played a major role in **Paper II**, where we observed the uptake of nucleoside phosphates by human cells in culture. In this work, we furthermore investigated the uptake mechanisms and explored the possibility of metabolic labelling of RNA by natural pathways. In **Paper V** we applied these observations in fluorescence lifetime imaging to investigate 2CNqA processing inside living cells. We also successfully used 2CNqA in *in vitro* reactions to intrinsically label mRNA which we subsequently delivered to cultured human cells to investigate its translatability. These results are presented in **Paper III**. **Paper IV** focusses on the characterization of the new analogue qU when incorporated into oligomers. This is an important stage in the evaluation of novel fluorescent base analogues for their possible applicability in biological assays and highlights an early step of FBA development.

Since these projects developed and grew together, they are herein described in intertwined chapters and not in chronological order. The first chapter covers the photophysical characterization of FBAs that was performed in different projects. The second part focusses on the imaging of FBAs by different fluorescence microscopy techniques. The last section covers the applicability of FBAs as tools to study cellular processes by different fluorescence readout methods.

In addition to the published results, this chapter also presents other, yet related, interesting findings and observations — well in line with Schätzing's quote that opens this thesis, which states that science is the art of approaching[127], *i.e.* it is a process with strong potential to reveal insights beyond the original scope.

3.1 Photophysical Characterization of FBAs

3.1.1 A new Uridine Analogue Incorporated into RNA

Once a design for a base analogue is completed and the compound is synthesized, the next step is to thoroughly characterize its photophysical properties in solution, usually first in a monomeric form, *i.e.* as a nucleobase or a nucleoside. In case of our new quadracyclic uridine

analogue qU (Fig. 2) this initial characterization was done and published prior to the herein presented work.[88] The qU ribonucleoside shows quite complex photophysics, but generally it absorbs light in the visible region (300–500 nm), a range that is well suitable for excitation using a 405 nm laser, which is common in confocal microscopes. Furthermore, the qU monomer exhibits a high quantum yield of 27% and lifetime of 2.44 ns at pH 7.[88] It displays dual-band emission characteristics, with lifetimes that depend on both pH and excitation wavelength, due to the presence of different ground- and excited-state species.[88] At physiological pH, qU is present in two forms that are in equilibrium, with one form being the iminol, *i.e.* protonated on the oxygen at position 2 of the pyrimidine scaffold (compare Fig. 1), and the other one being the amide *i.e.* protonated on the nitrogen at position 3 (as shown in Fig. 2). It was concluded from the absorption spectra and quantum chemical calculations that the iminol, which is the tautomer that would not form Watson-Crick hydrogen bonds, is dominant in the monomeric form in solution.[88]

With this knowledge, we took the next step to characterize qU inside a base stack and focussed on RNA oligomers. These oligonucleotides were synthesized by solid-phase synthesis with defined sequences that allow investigations on the effect of neighbouring bases and base-pairing (Tab. 1). By designing sequences where qU is flanked by different neighbouring bases, we were able to study how its electronic microenvironment influences its fluorescent properties in single- and double-stranded RNA oligonucleotides (ssRNA and dsRNA). Sequences where one qU-modified sequence was hybridized with strands with varying opposite bases to qU allowed further investigations on mismatch/base-pairing of the fluorescent nucleobase. Besides recording absorption and emission spectra and determining the fluorescence quantum yield and lifetime of qU in ssRNA and dsRNA, we also investigated the thermal stability and secondary structures of the modified dsRNAs by comparing to the corresponding unmodified sequences. This comprehensive characterization is presented in **Paper IV.**

A first and important observation was that the fluorescence quantum yield and lifetime inside RNA are both considerably increased compared to the monomer in a pH 7 solution ($\phi_{monomer}$ = 27%, $\tau_{avg,monomer}$= 2.44 ns). An increase upon incorporation is less common than a decrease and it is advantageous for possible applications of base analogues. On average, the fluorescence lifetime of qU in ssRNA is 7.1 ns and just slightly lower in dsRNA (6.5 ns). As expected, the same trend was found for the fluorescence quantum yield which is 52% in ssRNA and 43% in

dsRNA. However, comparing the quantum yield and lifetime values of the different sequences, our results show that the local environment of qU matters significantly, *i.e.* the direct neighbours affect the fluorescence quantum yield and lifetime (Tab. 1).

*Table 1. Denotation of the samples and the corresponding sequences in the qU RNA-oligomer study and the results for quantum yield (ϕ), fluorescence lifetime (τ), and melting experiments (ΔT_m, the difference between melting temperature of modified and unmodified samples). Top: sequences for the qU-modified samples, where the name indicates the direct neighbours of qU (X). Bottom: mismatch (-mis) sequences with the base opposite to qU indicated in the name and bold in the sequence. Mismatches are complementary to the UXC sequence, except for the opposite base of qU. Not written out are the sequences that are unmodified (U instead of qU) and that are fully complementary to the modified sequences (*i.e. *A-match). Values presented as mean ± standard deviation.*

		Φ (%)		**τ (ns)**		
Name	**Sequence 5'→3'**	**ssRNA**	**dsRNA**	**ssRNA**	**dsRNA**	**ΔT_m (°C)**
AXA	CGC AA**qU** AUC G	66 ± 2	51 ± 1	9.3 ± 0.2	9.8 ± 0.0	–10.8 ± 1.4
AXC	CGC AA**qU** CUC G	60 ± 1	45 ± 2	8.3 ± 0.2	6.0 ± 0.0	–14.0 ± 1.6
AXU	CGC AA**qU** UUC G	34 ± 1	40 ± 0	4.8 ± 0.1	9.4 ± 0.0	–11.8 ± 1.9
CXA	CGC AC**qU** AUC G	68 ± 1	47 ± 3	9.5 ± 0.3	3.5 ± 0.4	–6.0 ± 2.2
CXC	CGC AC**qU** CUC G	54 ± 1	42 ± 2	7.7 ± 0.1	6.4 ± 0.0	–6.8 ± 4.9
GXG	CGC AG**qU** GUC G	51 ± 1	51 ± 5	7.0 ± 0.1	7.8 ± 0.1	–9.3 ± 2.0
GXU	CGC AG**qU** UUC G	43 ± 1	49 ± 0	5.7 ± 0.0	7.3 ± 0.1	–13.8 ±1.8
UXA	CGC AU**qU** AUC G	62 ± 3	49 ± 3	9.1 ± 0.1	5.4 ± 0.2	–5.8 ± 2.4
UXG	CGC AU**qU** GUC G	42 ± 1	30 ± 2	5.5 ± 0.1	5.6 ± 0.0	–5.8 ± 1.7
UXU	CGC AU**qU** UUC G	33 ± 1	32 ± 0	4.5 ± 0.1	4.3 ± 0.1	–7.0 ± 1.6
UXC	CGC AU**qU** CUC G	55 ± 2	37 ± 1	7.1 ± 0.1	6.0 ± 0.0	–9.3 ± 2.0
G-mis	CGA G**G**A UGC G	-	57 ± 0	-	6.5 ± 0.1	–15.3 ± 1.6
U-mis	CGA G**U**A UGC G	-	72 ± 3	-	8.1 ± 0.1	+6.3 ± 1.8
C-mis	CGA G**C**A UGC G	-	69 ± 2	-	7.7 ± 0.0	+6.3 ± 1.2

We observed that not only the identity of the surrounding bases matters, but also their position, *i.e.* whether they are on the 5′ or 3′ side relative to qU. For single strands where qU has an adenine as 3’ neighbour the fluorescence quantum yield and lifetime are the highest, whereas a uridine at this position gives overall the lowest values. Cytidine as the 3’ neighbour gives the three next highest quantum yields and longest average lifetimes following the top adenine 3’-neighbour sequences. Conversely, this sequence dependency is much less pronounced when comparing different 5'-neighbours. After hybridization, *i.e.* for the dsRNA sequences, such trends are less clear, but it can be noted that double purine surroundings (AXA, GXG) result in the

highest quantum yields. This change from the single-stranded case indicates that there are additional and more complex effects on quantum yield and lifetime when qU is in an environment that forces it into a more stacked and helical geometry with the surrounding bases.

We also examined the absorption and emission spectra of these sequences (Fig. 8) in detail and compared them to the monomeric form (from data published by Le *et al.*[88]). For the ssRNA sequences (Fig. 8A), the local environment (*i.e.* the direct neighbours) appears to have minimal influence on the spectra, meaning the overall spectral shape is preserved. However, compared to the qU monomer, the absorption spectra of the ssRNAs exhibit a slight red shift, with the peak centred at 385 nm (versus 380 nm for the monomer) and a changed ratio of the two main peaks (at 330 nm and 385 nm). For the dsRNA (Fig. 8B), the spectral features are preserved throughout the sequences as well, and in this case, the ratio between the two main peaks is similar to that of the monomer. Some sequences also display a distinct shoulder around 435 nm, a wavelength range where the monomer absorbs as well.

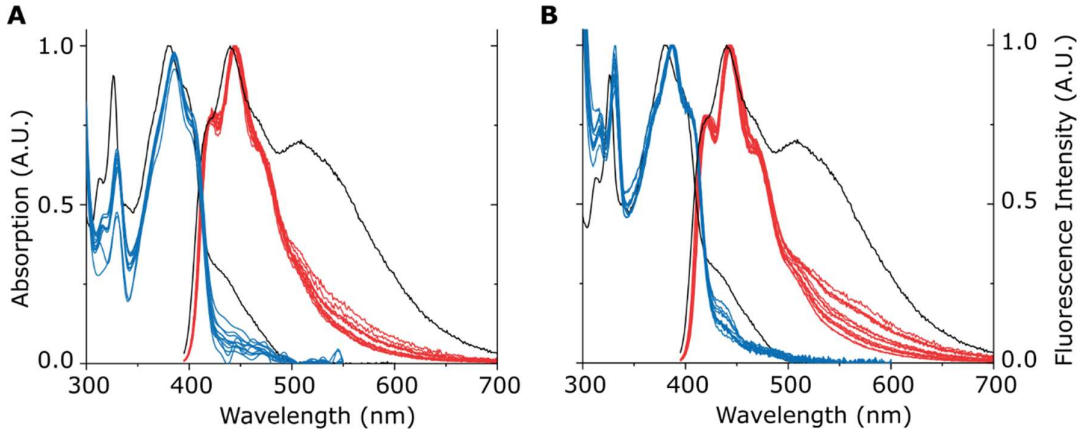

Figure 8. Normalized absorption (blue) and emission (red) spectra of qU inside oligomers. Per-spectrum sequence details (as in Tab. 1) are omitted for clarity. Black spectra show absorption and emission of qU monomer at pH 7 for comparison (taken from Le et al.[88]). Left axis absorption, right axis emission. A) single stranded RNA oligomers. Absorption spectra are smoothed using 5 pt FFT filter. B) double stranded RNA oligomers.

Overall, the spectral shape of the absorption spectra, with a significant reduction of the long-band absorption peak (450 nm), suggests that the iminol form of qU is even more dominant inside the base stack and that no shift towards an increased amount of the Watson-Crick-pairing amide form is observed. The emission spectra (Fig. 8 red) reveal that the fluorescence peak centred at 440 nm is preserved across all sequences, which is consistent with the monomer

data. However, the second emission peak around 530 nm, as observed for the monomer, is essentially absent in the oligomers. At neutral pH, this emission band was ascribed to the amide form of the monomer.[88] This again suggests that the equilibrium between the tautomers of qU is shifted even more towards the iminol form inside oligomers. Notably, fluorescence in this longer wavelength region (> 500 nm) shows a slight sensitivity to sequence variation, indicating that subtle electronical differences do influence the excited-state behaviour, yet not considerably.

We further investigated the impact of qU incorporation on the base-pairing and overall RNA structure. To this end, we recorded UV melting curves to calculate the melting temperature of the dsRNAs, and recorded CD spectra (Ch. 2.2.1.1), respectively. First, we recorded CD spectra of qU modified dsRNA oligomers and compared them to their unmodified counterparts. The CD spectra of both modified and unmodified sequences display the reported CD fingerprint of the canonical A-form double helix of RNA[128], indicating that qU incorporation does not affect the overall A-form RNA structure. However, deviations between the CD spectra of unmodified and modified duplexes are observed (Fig. 9).

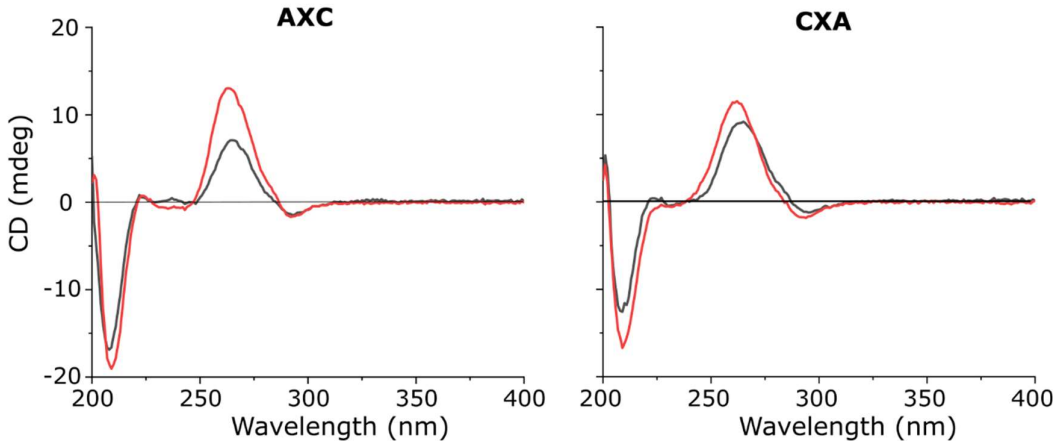

Figure 9. CD spectra of two example RNA duplexes (AXC and CXA), highlighting the sequence dependency of the nearest neighbours of qU (X). Compare Tab. 1 for sequence details. Y-axis has the same scale in both graphs. Black lines: qU-modified sequences, red lines: unmodified sequences, i.e. *U instead of qU.*

Since the absorption features, *i.e.* position of peaks and molar absorptivity, of qU are different to U, CD-signal changes are expected upon their exchange, but these effects cannot explain all changes. For example, comparing the sequences AXC and CXA we observed differences compared to their unmodified counterpart (Fig. 9, comparing black with red lines), but also when comparing these two modified duplexes (Fig. 9, comparing black lines), with a shift of

the 260 nm peak. In these strands qU has the same direct neighbours (A and C), but the order is inversed, meaning that the sequence of the base stack affects the overall A-form helix structure.

The intensity differences in the CD spectra between modified and unmodified strands show that almost all qU modified strands give less CD signal than the unmodified ones, even though qU has the higher molar absorptivity[88]. This may suggest that the overall concentration of A-form helices, especially looking at the extent of perfect A-form for the whole 10-mer duplex, is less in samples where qU is incorporated. This supports our hypothesis that qU is the iminol tautomer inside oligomers, *i.e.* the non-base pairing and, hence, helix disturbing form.

To further investigate the base-pairing characteristics of qU with the canonical adenine in the complementary strand, we recorded absorption melting curves, *i.e.* monitoring the hypochromicity at 260 nm over a certain temperature range (Ch. 2.2.1), of the sequences containing different neighbours of qU. As all melting curves show a sigmoidal shape (Fig. 10), this data suggest that also qU-modified sequences form A-form RNA duplexes with its complementary sequence.

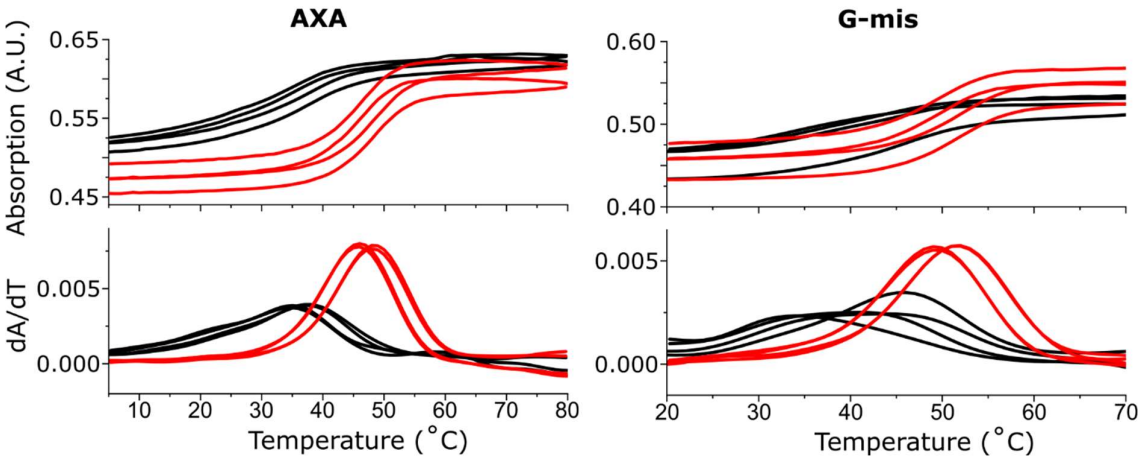

Figure 10. Four melting ramps (top) of one representative nearest neighbour sample (AXA) and one mismatch (G-mis) sequence and their 1st derivative (bottom). Unmodified sequences (U instead of qU) are shown in red, while the qU-modified sequences are displayed in black. Curves are smoothened using a 5 pt FFT-filter. Compare Tab. 1 for sequence details.

However, the melting ranges are very different between the modified and unmodified sequences (Fig. 10, compare black and red lines). It is also worth noting that the hypochromicity is smaller for the qU-modified duplexes indicating that the duplex in this case is less perfect. Calculating the melting temperatures (T_m) and the differences between the

unmodified and modified sequences (ΔT_m), we observed that qU destabilizes the duplexes by 9.1 ± 3.1°C with a ΔT_m range from –5.8°C to –14.0°C (Tab. 1). This, again, supports the hypothesis that qU inside oligomers is in its iminol form and does not base-pair properly with adenine. Furthermore, as for the CD measurements, we observed a sequence dependency of the destabilization effect. For example, with a purine base on the 5' side of qU the duplexes are more destabilized ($\Delta T_m < -9°C$) than with 5' pyrimidine ($\Delta T_m \geq -7°C$).

Taken together, qU is destabilizing the A-form RNA structure by highly compromised base-pairing interactions compared to natural U. Clear patterns of how the photophysics of qU are affected by the structure and binding of the different sequences, *i.e.* comparing the fluorescence quantum yield and lifetime with trends in the melting temperatures and CD data, were difficult to identify. However, we hypothesised that qU exists mainly as the iminol tautomer inside the base stack, which leads to destabilized duplexes. The destabilization by qU is in contrast to other fluorescent base analogues that we have designed and investigated previously. The adenine analogue qA[81] with its derivatives 2CNqA[83], qAN1[129], qAN4[130] and pA[84] on average increase duplex stability of DNA or RNA. The cytosine analogue tC [131] shows similar duplex stability as natural DNA double strands, whereas its homologue tC^{O} [73] displays slightly increased stability of double stranded DNA and RNA.

To investigate the base-pairing capabilities of qU with the other canonical bases, we also studied non-Watson-Crick base interactions of qU by hybridizing one sequence with varying bases opposite to qU in the complementary strands. As qU is designed to be a uridine analogue, the complementary base we used for the sequence-dependency study described above was adenine. When adenine (A-match) is replaced by cytosine (C-mismatch), guanine (G-mismatch), or uridine (U-mismatch) (Tab. 1), we observed small differences between the qU-modified mismatch strands compared to the unmodified mismatch strands in the CD spectra, indicating that the overall A-form RNA helix is again preserved, but altered, as for the A-match sequences. Recording melting curves and calculating the resulting melting temperatures (Tab. 1) we observed that duplexes where qU has a pyrimidine as complementary base are stabilized by 6.3°C, while the G-mismatch gives the most destabilized duplex ($\Delta T_m = -15.3°C$). Therefore, we hypothesized that, because qU is an enlarged nucleobase with enhanced hydrophobicity, it localizes more easily in the hydrophobic environment inside the base stack where the opposite base is smaller, *i.e.* a pyrimidine. An interesting continuation of this investigation would be to look at the effects of different nearest neighbours of qU in the mismatch samples, *i.e.* a combination of the two parts of the herein presented characterization.

Future investigations could also focus on the observed tautomer equilibrium and include experiments on the pH dependency of qU when incorporated into oligomers. Le *et al.*[88] observed that qU is highly sensitive to pH where the deprotonation of the hydroxy group of the iminol tautomer results in interesting shifts towards longer wavelengths. Since the iminol form is stabilized in oligomers, this would make qU a promising tool for spectroscopy and microscopy applications, particularly in RNA and antisense oligonucleotide (ASO) therapeutics. Our characterization of qU in an RNA context is therefore the next step after the previous characterization as monomer, towards applications of a new base analogue in RNA, that extends the FBA-RNA alphabet with a fluorescent uracil.

3.1.2 Characterization of FBA-Nucleotides

In addition to qU discussed above (Ch. 3.1.1), several other fluorescent base analogues (FBAs) developed in the Wilhelmsson group had already been thoroughly characterized — both as free nucleobases and when incorporated into DNA or RNA strands — and applied in various studies (Ch. 2.1.4). During the curse of this thesis, the focus shifted toward the use of FBAs for fluorescence microscopy of live cells. Following the publication of the synthesis and applicability of tC^O as triphosphate[87], the synthesis of other FBA-triphosphates (FBA-TPs) was established and opened the door for a broad range of investigations.[132] However, before applying the triphosphates of these analogues in *in vitro* or *in cellulo* experiments, it is essential to establish a comprehensive understanding of their photophysical properties in aqueous solution, in order to avoid misinterpretation of experimental results. Three of our FBAs, namely tC^O, 2CNqA, and pA (Ch. 2.1.4, Fig. 2), have been shown to be suitable base analogues with useful photophysical properties as nucleobases and inside oligomers[73,74,83,84,86,129,133] and, hence, were chosen for investigations as ribonucleotides for cell studies. Therefore, in **Paper I** and **Paper II** we continued the single-photon characterization of the nucleotide triphosphates of tC^O and 2CNqA and presented the synthesis and characterization of the nucleoside triphosphate of pA, and the monophosphate version of 2CNqA. We recorded absorption, emission, and excitation spectra, calculated the fluorescence quantum yields, and recorded the fluorescence lifetimes (Fig. 11, Tab. 2). The absorption in the near-UV region is a common feature of FBAs, and the tail of the lowest energy band allows excitation at 405 nm, which is a common wavelength in conventional microscopes and flow cytometers.

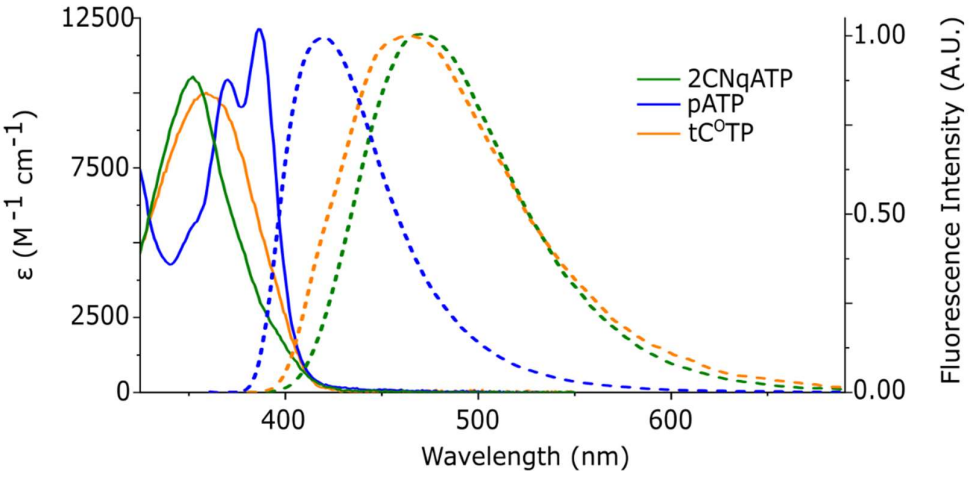

Figure 11. UV–vis absorption (molar absorptivity ε, solid line) and normalized emission (dotted line) spectra of the FBA-TPs reported in ***Paper I*** *and* ***Paper II****. Chemical structures in Fig. 2 display same colour code. 2CNqAMP spectra are omitted for clarity, they overlap with the spectra of 2CNqATP.*

Table 2. Photophysical characteristics of the FBA-TPs (chemical structures Fig. 2). ε_{260}*: molar absorptivity at 260 nm, reported in given references;* ϕ*: fluorescence quantum yield;* $\lambda_{abs,max}$*: wavelength of absorption maximum;* $\lambda_{em,max}$*: wavelength of emission maximum;* $\varepsilon_{max,\,FBA} \times \phi$*: brightness at the absorption maximum;* $\varepsilon_{405,\,FBA} \times \phi$*: brightness at 405 nm.*

FBA-TP	ε_{260} **(M^{-1} cm^{-1})**	ϕ **(%)**	$\lambda_{abs,\,max}$ **(nm)**	$\lambda_{em,\,max}$ **(nm)**	$\varepsilon_{max,\,FBA} \times \phi$ **(M^{-1} cm^{-1})**	$\varepsilon_{405,\,FBA} \times \phi$ **(M^{-1} cm^{-1})**	τ **(ns)**
tC^{O}TP	11000[73]	27	360	453	2700	417	3.4
2CNqATP	14600[83]	48	352	471	4900	504	9.9
pATP	22300[84]	49	386	420	6000	909	6.0

The brightness values range from 2700 to 6000 $M^{-1}cm^{-1}$, which are exceptionally high for FBAs[62,82,83,130] (compare 2-aminopurine with a brightness of 5300 $M^{-1}cm^{-1}$, and drops ~100-fold upon incorporation into nucleic acids[68], Ch. 2.1.4). Together with the first studies successfully using these analogues for imaging[84,86,87], this characterization motivates further imaging applications. For 2CNqATP we observed a lifetime of 9.9 ns with a mono-exponential decay, which is in contrast to incorporated 2CNqA, where we have observed bi-exponential decays resulting in an average lifetime of around 6 ns[83]. This difference opens the door for imaging applications using the fluorescence lifetime, instead of intensity-based techniques.

We also investigated the pH-dependency of the fluorescence quantum yield and lifetime of 2CNqATP and pATP (data unpublished). The fluorescence quantum yield of both compounds decreases with decreasing pH, where pATP is more affected (Fig. 12).

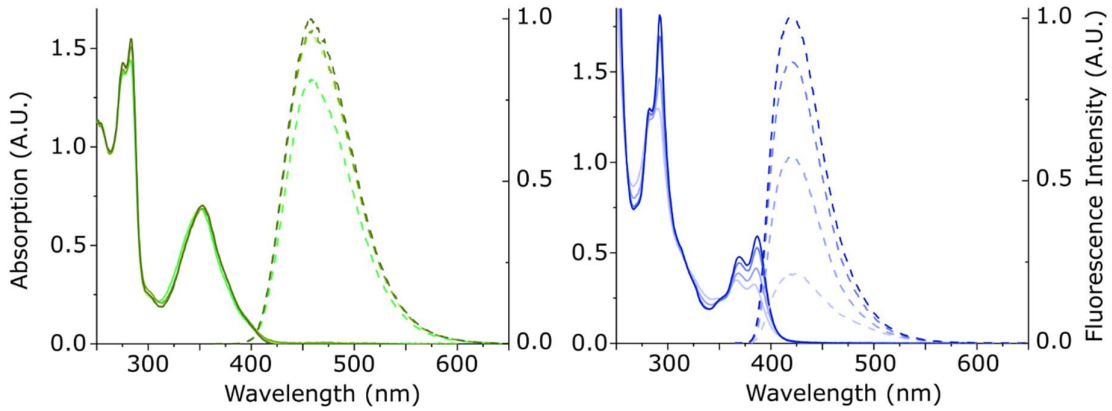

Figure 12. Absorption spectra normalized to 260 nm (solid lines, left axis) and emission spectra normalized to unity at the peak maximum (dotted lines, right axis) of 2CNqATP (green) and pATP (blue) at pH 4.5 (lightest colour), pH 5.5, pH 6.5, and pH 7.5 (darkest colour).

However, the fluorescence lifetime is stable with values of 9.9 ns for 2CNqATP and 6 ns for pATP. Interestingly, in an earlier study in the group, we observed that the lifetime of 2CNqA decreases with increasing H^+ concentration when it is in an ASO environment (data by Dr. J.R. Nilsson, unpublished). This highlights that 2CNqA is slightly sensitive to its microenvironment, which is important to know when studying 2CNqA in complex systems like cells. This observation might even be useful to monitor 2CNqA-labelled ASOs during cellular uptake via endocytosis, a cellular ingestion where the ASO-containing vesicles become more acidic during the uptake process. Future investigations of pH effects on 2CNqA inside long RNAs might reveal possible applications in LNP-based delivery studies.

3.1.3 Multiphoton Characterization

FBA imaging requires excitation using near-UV light, *e.g.* the common 405 nm laser line in a confocal microscope (Fig. 11, Tab. 2). However, when imaging in living systems excitation using short wavelengths can lead to phototoxic effects.[103] Furthermore, penetration into thicker biological samples, like tissues, becomes less at shorter wavelengths, meaning that deep tissue detection of fluorophores excited in or near the UV region is a problem.[105] Multiphoton excitation of fluorophores mitigates these shortcomings (Ch. 2.2.1.2). To explore the use of

fluorescent base analogues as imaging probes for two-photon imaging, we thoroughly characterized the adenine analogue 2CNqA as a multiphoton probe (**Paper I**). Two-photon absorption properties and detectability inside living cells was investigated for 2CNqA as ribonucleoside triphosphate and incorporated at one or two positions inside a 16-mer ASO in aqueous solution (Fig. 13).

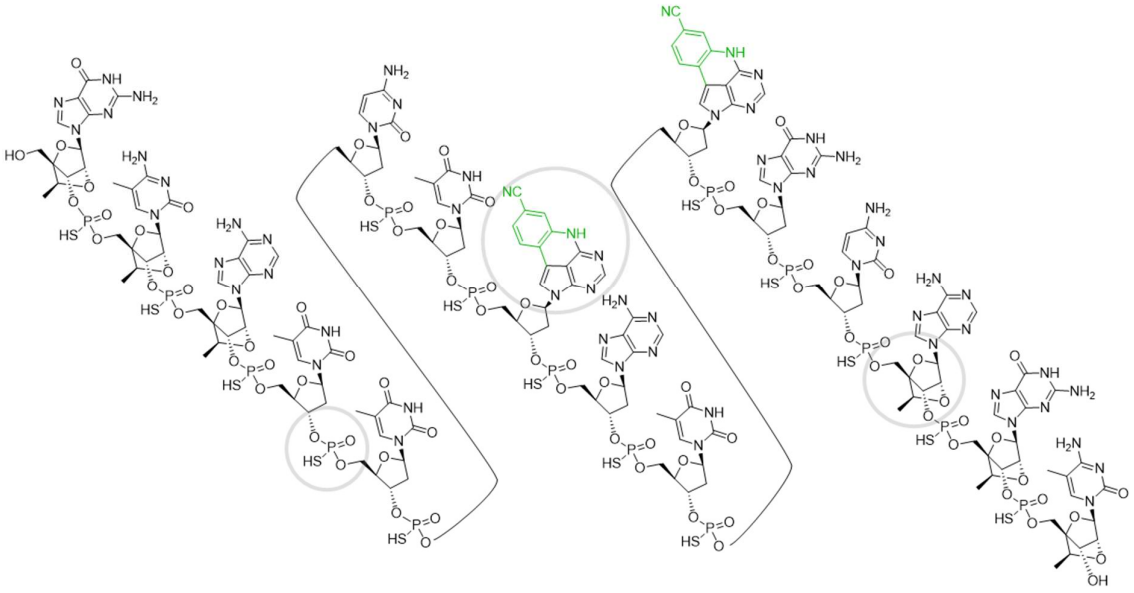

*Figure 13. Molecular structure of the ASO sequence used in **Paper I**. The overall sequence is 5'–**GCA**TTCT<u>A</u>AT<u>A</u>GC**AGC**–3 with the underlined As being replaced by 2CNqA at position 11 for single labelling, or positions 8 and 11 for double labelling (modification in green in the figure). The sugar moieties of the three bases at the ends of the ASO are cEts, and all nucleosides are linked via phosphorothioates (Ch. 2.1.2). Grey rings highlight examples of these modifications in the structure. The sequence is shown in three blocks only to fit the page; this has no scientific meaning.*

We determined the two-photon cross section and emission spectra and found that the two-photon excitation maximum is at 700 nm in all three cases (Fig. 14A). 2CNqA inside ASOs shows a pronounced shoulder at around 780 nm, again highlighting that the spectral characteristics of 2CNqA change slightly depending on its microenvironment (Ch. 3.1.2). To confirm that the absorption is indeed a two-photon process we created a multiplicity plot. The excitation intensity at 700 nm was modulated and the corresponding emission intensities were recorded and integrated. By graphing the logarithm of the laser power (*i.e.* excitation intensity) to the logarithm of the integrated emission intensity, a linear plot was achieved (Fig. 14B). Linear regression resulted in a slope of ~2, indicating a two-photon process.

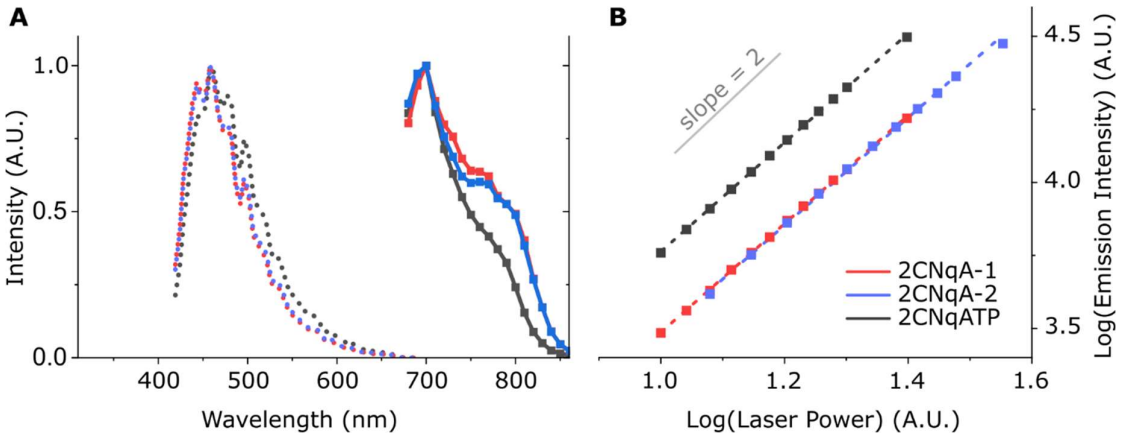

Figure 14. Two-photon characterization of 2CNqA inside an ASO environment (red and blue) and as 2CNqATP (black). A) Two-photon emission spectra (dotted lines) and two-photon cross section spectra (boxed, solid lines). Normalized to unity at peak maximum. B) Absorption multiplicity plot. Integrated emission intensities were determined at various excitation intensities at 700 nm. Linear regression of the log-log plot resulted in a slope of 1.9 for all three samples. Slope of two (grey) shown for comparison.

The cross section of the triphosphate (5.8 GM) is close to the single-labelled ASO (2CNqA-1: 6.9 GM), and the double-labelled ASO shows a doubled cross section (2CNqA-2: 13 GM) which is well in agreement with the number of incorporated 2CNqAs. These cross sections at 700 nm are exceptionally high for FBAs, which was confirmed by high brightness values (as count rate per molecule) from fluorescence correlation spectroscopy (FCS) measurements. Our adenine analogue pA was previously characterized as two-photon probe, but shows only about half of the brightness of 2CNqA.[85] Therefore 2CNqA is a great FBA candidate, also for two-photon imaging.

3.2. Live-cell Imaging using FBAs

3.2.1 Confocal Microscopy

In 2021 Baladi *et al.* published that tC^O incorporated into long RNA can be used to track translatable RNA inside cells.[87] Together with the knowledge of the photophysical properties of our FBA-TPs and their compatibility with a standard 405 nm laser line (Ch. 3.1.2), this finding lead to further investigations for imaging in cellular environment. Having tC^O, pA and 2CNqA (Fig. 2) as ribonucleoside triphosphates (-TP) at hand, we added these monomers to the cell medium of living cells without formulation or transfection and followed the emission signal using a laser scanning confocal microscope. This procedure was against rational experimental design and curiosity driven, and because charged molecules (triphosphates have

a net charge at physiological pH of –3.3[134]) do not easily cross hydrophobic bilayers like cell membranes (Ch. 2.1.2 and Ch. 2.1.3), we did not expect much intercellular signal. Surprisingly, cells exposed to 2CNqATP and pATP showed strong fluorescence, while tC^{O}TP-exposed cells displayed signal only on the autofluorescence level (Fig. 15).

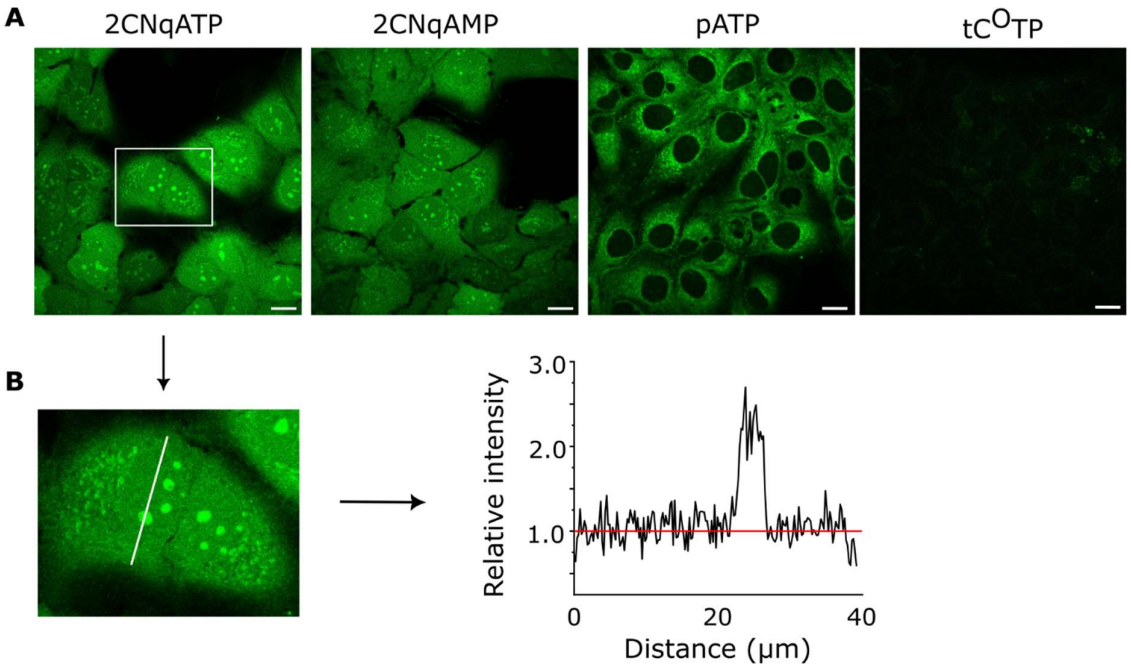

Figure 15. Spontaneous uptake of fluorescent nucleoside phosphates. A) Confocal fluorescence images of Huh-7 cells exposed to 2.5 µM of the indicated compound for 24 h. B) Representative fluorescence intensity line profile analysis of a cell exposed to 2CNqATP.

Because the photophysical properties of these compounds were well characterized (Ch. 3.1.2), we could exclude quenching effects on tC^{O} and instead stated that the adenine analogues enter the cells, while the cytosine analogue does not. This serendipitous finding led to the main project of this thesis work, **Paper II**. We observed that the cellular localization differed strongly between 2CNqA and pA: pA localizes in cytosolic structures, whereas 2CNqA is more evenly distributed and reaches the nuclei and accumulates inside the nucleoli (Fig. 15A). Simple line profile analysis showed that the intensity in these nuclear sub-structures is 2.3 (± 0.1) times higher than in the nuclei and cytoplasm (Fig. 15B). Because of this nuclear localization, *i.e.* inside the cellular compartments where RNA synthesis happens, 2CNqA was in focus for our further studies. Therefore, we also studied the uptake of its monophosphate (-MP) version and observed the same localization as for the triphosphate (Fig. 15A). Interestingly, when recording the uptake as timelapse we observed that the internalization of the monophosphate derivative

occurs much faster (Ch. 3.3.1, *vide infra*). To the date this thesis was written, the uptake of 2CNqATP or -MP was tested in adherent cultures of at least five different cell lines where the same uptake and localization pattern was observed.

Imaging cells exposed to 2CNqATP or 2CNqAMP was hereafter the basis for other investigations on imaging and applicability studies. Because the uptake of this compound was unexpected, and the pathways it takes unknown, we observed interesting effects, based on confocal imaging. For example, co-exposure with the nuclear dye DRAQ5 lead to an effect on the localization of 2CNqA, where the nucleolar signal changed to a more punctuated pattern inside the nuclei (Fig. 16, data unpublished). To the date this thesis was written we cannot explain this observation. However, further investigations can lead to better understanding of 2CNqA as adenine analogue inside cells and could provide insights into cellular processes during nuclei staining.

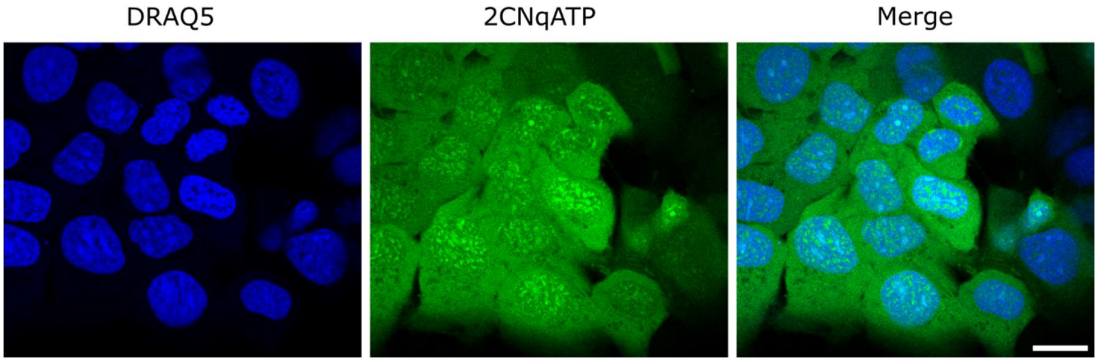

Figure 16. Confocal fluorescence images of Huh-7 cells co-exposed to DRAQ5 (λ_{ex} = 640 nm, displayed in blue) and 2CNqATP (λ_{ex} = 405 nm, displayed in green). Scale bar = 20 µm. Compare intracellular localization of 2CNqA to Fig 15.

3.2.2 Fluorescence Lifetime Imaging

Not only did we investigate intensity-based imaging, but we also tested the applicability of 2CNqA in fluorescence lifetime imaging (FLIM), which resulted in a submitted communication (**Paper V**). From our thorough photophysical characterization (Ch. 3.1.2), we knew that the 2CNqA triphosphate displays a mono-exponential decay, with a resulting lifetime of 9.9 ns (**Paper I**), while the decays of metabolically incorporated 2CNqA need to be fitted to a biexponential decay function with one resulting lifetime being 9.2 ns and the second one 1.4 ns (**Paper V**). While the longer lifetime of the incorporated derivative is very close to the one of the monomer (*i.e.* triphosphate, Tab.

2) — both being around 9 ns — the shorter lifetime will be distinguishable from the longer-lived species by FLIM. Thus, changes in the amplitude contribution of these lifetime species indicate a change in the microenvironment of the 2CNqA-base, which hence could be due to metabolic RNA incorporation. Therefore, we repeated the exposure of cultured cells to 2CNqATP and monitored the cellular uptake using FLIM. As the setup did not allow a timelapse recording, we studied two exposure timepoints, 4 h and 24 h, that were based on the observations in **Paper II**, to get an insight in the cellular processing of 2CNqA over time. The FLIM images were analysed in the time domain by fitting the fluorescent decays and in the frequency domain by phasor plots.

By fitting the image decays to a bi-exponential decay model we observed one lifetime of 10 ns and a second one of 3 ns. We then compared the two exposure times (4 h and 24 h) by looking at the amplitudes of these two lifetime species (Tab. 3).

Table 3. Fitted amplitudes (α) of the exponential decay components normalized to unity. The amplitudes were fitted with τ_1 fixed to 10 ns, and τ_2 to 3 ns. Samples are numbered, with the time cells were exposed to 2CNqATP indicated.

Sample	α_1	α_2
#1; 4 h	0.52	0.48
#2; 4 h	0.47	0.53
#3; 24 h	0.40	0.60
#4; 24 h	0.40	0.60

After four hours exposure both fluorescent species contribute to a similar extent. However, the contribution of the shorter-lived species increases over time (24 h), which results in a shorter average lifetime. As the shorter lifetime species only stems from non-monomeric 2CNqA, this result shows a change in microenvironment that indicates that 2CNqA is processed by the cells. In **Paper II** we presented the metabolic incorporation of 2CNqA into RNA (Ch. 3.3.2, *vide infra*), which supports our interpretation that we observe RNA-incorporation in the FLIM images.

Because measurements in aqueous solution do not straightforwardly represent the complex cellular environment that 2CNqA experiences and that we detect in the FLIM images, we also analysed the FLIM data in a model-independent manner by separating fluorophore populations based on their overall decay features by the phasor approach (Ch. 2.2.3.3). We observed non-mono-exponential decays, as most phasors did not locate on the semi-circle of the plot. The photon-weighted centre of

mass distribution of the images corresponds well to the average lifetimes of the decay fits, with shorter exposure time (4 h) resulting in longer average lifetimes. For comparison, we drew an intensity fraction line through the phasor centres to display the images with relative lifetimes along this line. We observed clearly that cells exposed for four hours display a larger contribution of the longer lifetime species compared to cells exposed for 24 h.

Looking at all resulting images (from the decay fitting and phasor analysis), we observed that the longer lifetimes are present in the cytosol, while the nuclei display shorter lifetimes (Fig. 17).

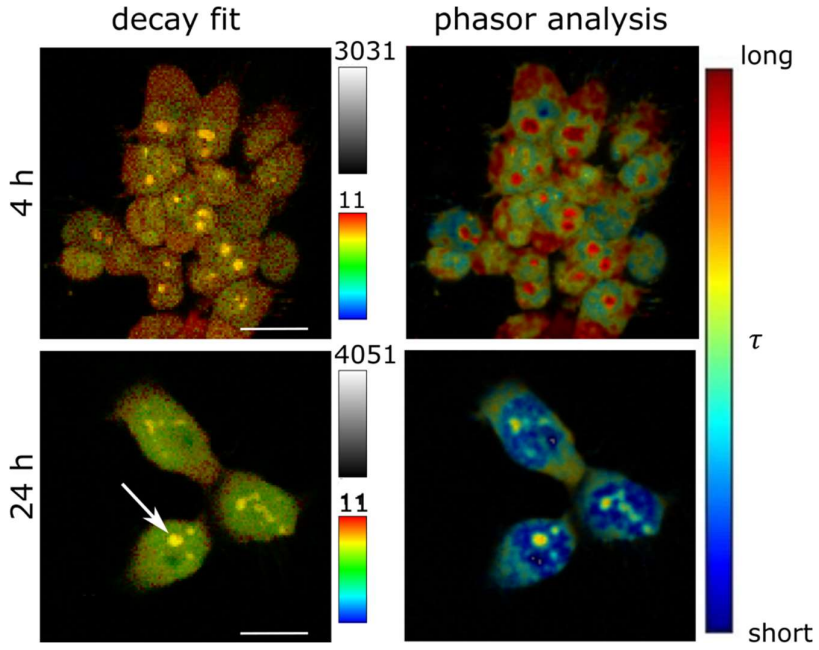

Figure 17. FLIM images of HEK293T cells exposed to 2CNqATP for 4 h or for 24 h (samples #1 and #3 in Tab. 3). Left: Resulting images from the analysis of the fluorescence decays with colourized lifetime scale from 0 ns (blue) to 11 ns (red) and intensity scale, i.e. *photon counting events from zero (dark) to indicated count number (bright). Right: Images from the phasor analysis where a fraction line was drawn through the phasor centres of the recorded images. Colouring is based on this fraction line to display the lifetime distribution. Scale bars = 30 µm. Arrow highlights one nucleolus as example.*

The nuclei are the cellular compartments where RNA synthesis takes place, hence, having an increased contribution of the short-lived species in these cellular structures supports the hypothesis that this is an effect of RNA incorporation (Ch. 3.3.2, *vide infra*). Most strikingly, the nucleoli (nuclear sub-structures, Fig 17 arrow) display a

higher fraction of the long-lived species. We knew from **Paper II** that 2CNqA accumulates inside the nucleoli, since they appear as bright spots in the intensity-based images (Fig. 15). Using FLIM, we obtained additional information on the fluorescent species of 2CNqA in the nucleoli, *i.e.* the long-lived species of 2CNqA is predominant in these nuclear structures.

Taken together, **Paper V** presents the first study to use FLIM measurements to monitor FBAs in live cells. We were able to detect different fluorescent species of 2CNqA inside living cells, which is a powerful approach for further investigations on its metabolization, as well as a basis for studies using RNA labelling and nucleoside metabolism.

3.2.3 Two-photon Imaging

Because we found 2CNqA suitable as a two-photon probe in **Paper I**, we also investigated our FBAs in two-photon imaging. In **Paper I**, we show the applicability of 2CNqA inside ASOs for two-photon live-cell imaging. Since we used a microscope setup capable of confocal (*i.e.* one-photon excitation) and two-photon imaging we were able to acquire images of the same region of interest of our live-cell samples in both modes.

The labelled ASOs appear as bright spots within the cells, consistent with their localization in vesicles formed during endosomal uptake, which is the reported internalization pathway for ASOs[31]. These spots overlap very well in the images of both excitation modes, highlighting that it is possible to image 2CNqA inside ASOs using both single- and multiphoton excitation (Fig. 18). Additionally, the emission from control cells is low and the overlap between one- and two-photon signals is poor, which confirms the detection of 2CNqA in the ASO exposed cells.

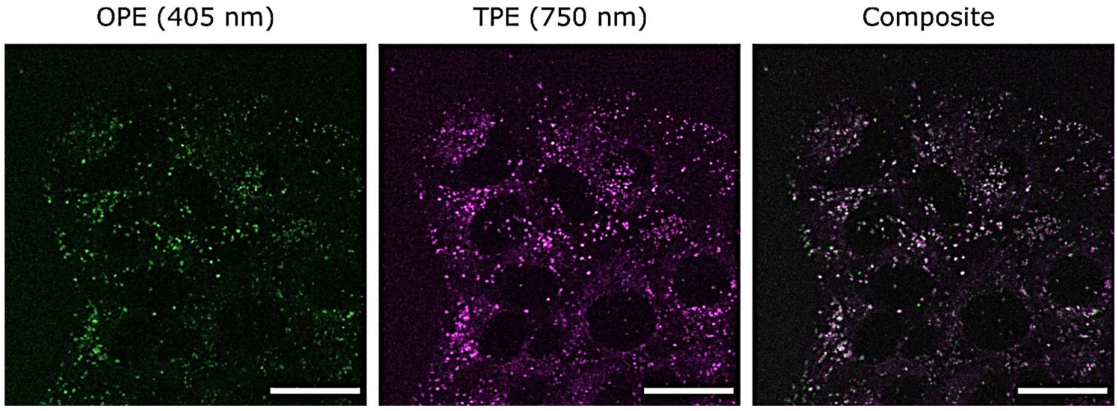

Figure 18. Examples of fluorescence microscopy images of Huh-7 cells exposed to ASO with a single incorporation of 2CNqA. 2CNqA was monitored using one- (OPE) and two-photon (TPE) excitation through an emission filter (410–680 nm) using identical detector gain. Scale bars = 30 µm.

In the following work, we monitored the cellular uptake of nucleosides of 2CNqA and pA into cells using two-photon imaging. Both FBAs were promising probes based on our findings in **Paper I** and **Paper II**, as well as previous work in the group[85]. Building on the results from **Paper II** we included experiments where we exposed live cells to 2CNqATP, 2CNqAMP, and pATP and imaged in one- and two-photon mode (Fig. 19, data unpublished).

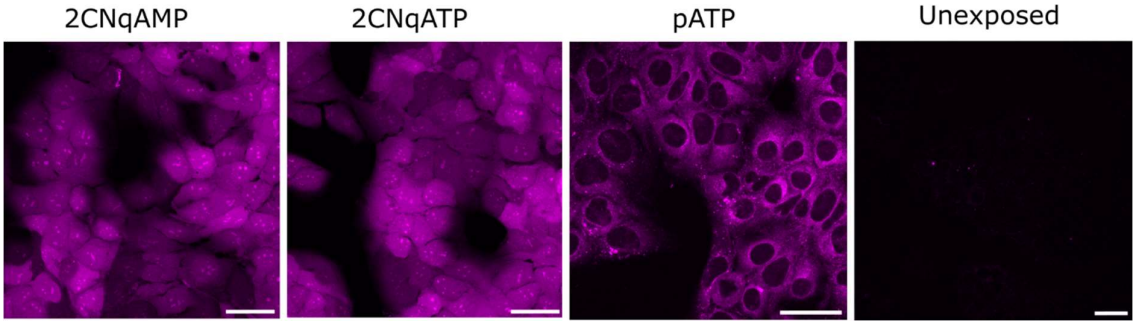

Figure 19. Two-photon images of living cells exposed to 2CNqAMP, 2CNqATP, or pATP. For autofluorescence level images of unexposed cells were recorded as well. Excitation wavelength was set to 750 nm and emission was recorded between 410 nm and 680 nm. Scale bars = 50 µm.

The uptake of these FBAs was readily detectable in both imaging modes, and their intracellular localization appeared consistent with observations reported in **Paper II**. Imaging in two-photon mode is beneficial for FBAs, for two reasons: I) the most common UV laser line for confocal imaging is 405 nm, not 360 nm, but FBA excitation using the 405 nm laser line is not optimal, because it does not match the absorption maximum of the FBAs. Using 750 nm, the FBAs are excited more efficiently; II) as

described previously (Ch. 2.2.3.2), using longer wavelength penetration depth is better and allows deeper imaging in biological samples. Hence, showing that imaging 2CNqA and pA in two-photon mode is feasible we open the door for applications like organoid or tissue imaging.

Having demonstrated the applicability of 2CNqA in two-photon imaging and FLIM we have presented a strong imaging probe. In two-photon FLIM the advantages of both imaging techniques — looking at fluorescence lifetime at high spatial resolution with increased penetration depth and minimum photobleaching — are combined, which makes it a valuable technique in biomedical research[135]. Hence, 2CNqA could be a probe applicable for nucleotide and RNA studies in this field.

3.3 FBAs as Research Tools

3.3.1 Investigation of Cellular Processes

Following the observation that 2CNqA and pA are internalized by eukaryotic cells, and that this can be visualized by live-cell imaging (Ch. 3.2), we aimed for a more detailed analysis of the cellular uptake processes (**Paper II**). First, we tested for cytotoxicity of our compounds where we did not observe any effects on cell morphology, cell membrane integrity (tested using LDH leakage assay), nor metabolic activity (tested using AlamarBlue assay) in a range of concentrations at 24 h exposure time. We then repeated the uptake experiment at 4˚C where no intracellular signal was observed (Fig. 20A). Because active processes are disfavoured at lower temperatures, we concluded that the internalization of 2CNqA and pA happens via energy-dependent cellular processes, *i.e.* active uptake via enzymes, such as transporters. When performing experiments where washing of cells and media exchange after exposure was needed, we observed that pA is washed out, while 2CNqA-exposed cells show stable fluorescence intensity (Fig. 20B). Hence, an equilibrium between cellular import and export determined pA concentration, whereas 2CNqA stably accumulates inside cells. This is interesting, because it shows that the two A-analogues are sensed differently by the cells after internalization. Investigating the uptake and excretion of pA, its localization inside cells, and its intracellular processing might be a future research project that can shine light on the structure-activity relationship of base modifications and their uptake, as well as intracellular localization and reactions — especially when comparing pA to the structurally similar 2CNqA (Fig. 2). To understand

the structure-activity relationship of modified nucleosides will enable the rational design of new and more efficient analogues for a variety of applications.

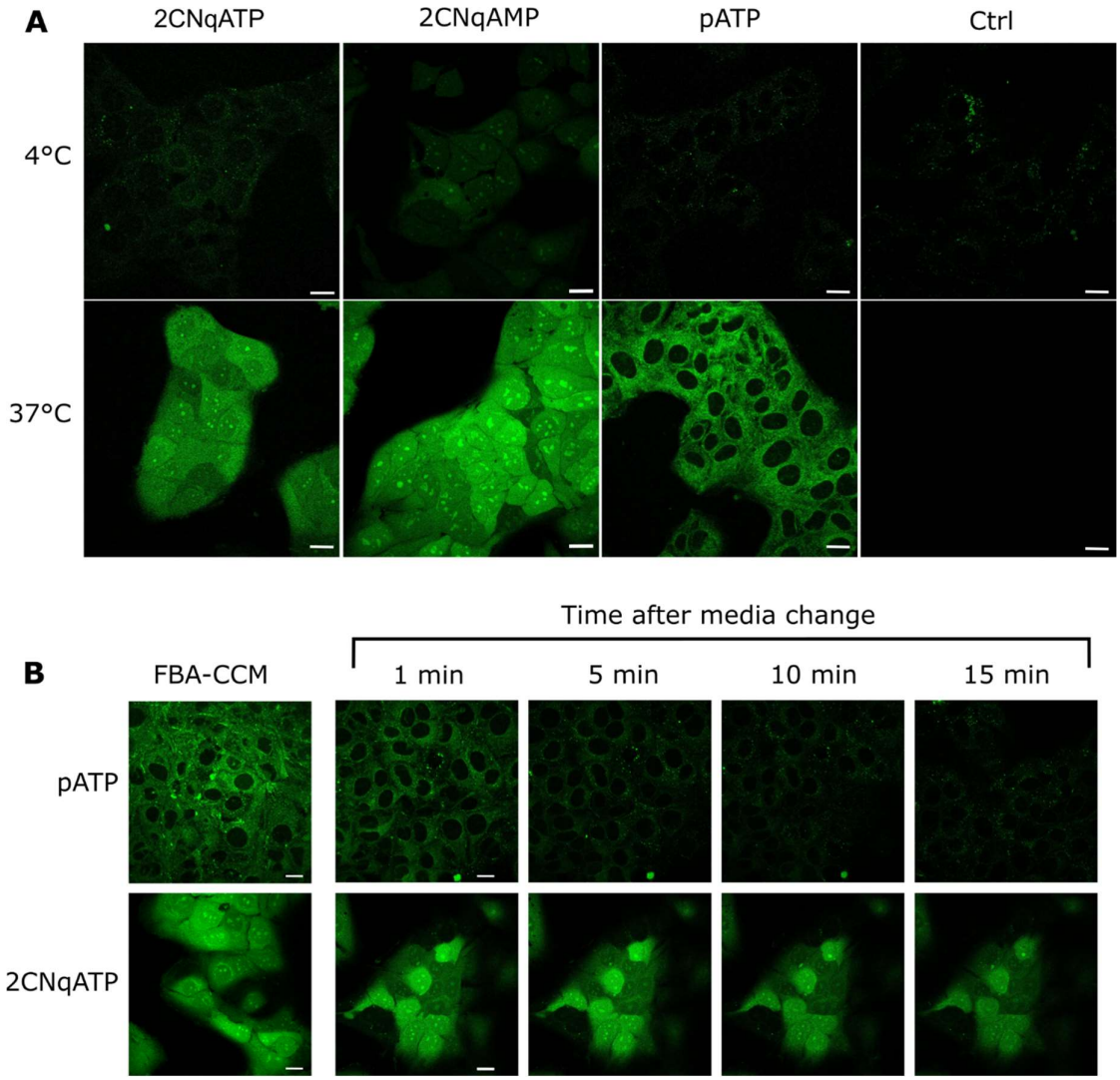

Figure 20. A) Confocal microscopy images of the uptake of 2CNqATP and pATP after 1.5 h at 4°C compared to the uptake at 37°C. Control (Ctrl) are cells exposed to buffer in cell culture medium. B) Confocal microscopy images of pATP and 2CNqATP uptake before and after removal of the treatment solution (FBA-CCM) and adding fresh cell culture medium (CCM). Scale bars = 20 µm

For the continuation of the herein presented project (*i.e.* **Paper II**) the clearance of pA meant that it could not be included for flow cytometry readout and other experiments where detaching and washing was required. Furthermore, since 2CNqA shows stable accumulation and nuclear localization, *i.e.* where RNA synthesis takes place, the focus was put on 2CNqA in the continued work of this thesis. To investigate its uptake closer, we also included the monophosphate version of the nucleoside. As for 2CNqATP

exposure, we observed uptake via energy-dependent processes and the same intracellular localization for 2CNqAMP.

For investigations on dose-dependent effects of the uptake we added different concentrations of 2CNqAMP and 2CNqATP to cells and observed a saturation effect above 2.5 µM FBA in the cell media for both compounds (Fig. 21A).

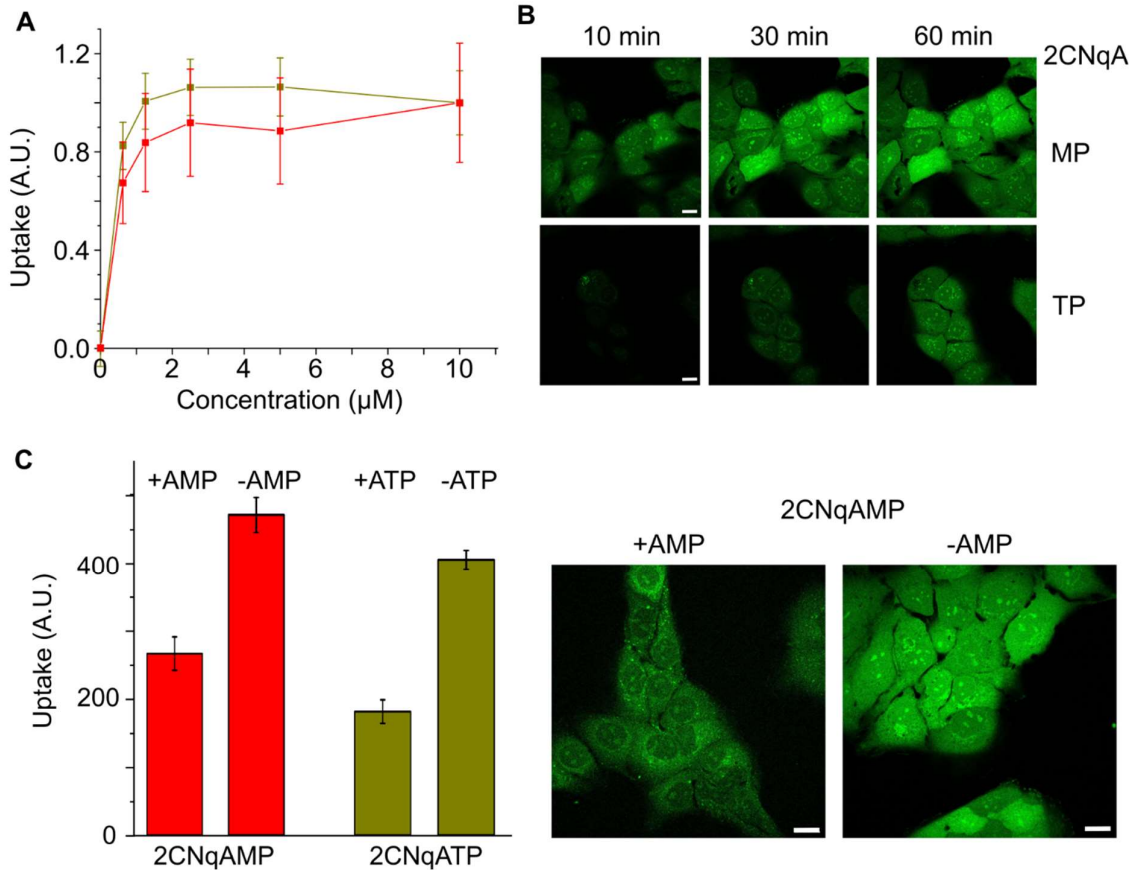

Figure 21. A) Uptake of 2CNqATP (yellow) and 2CNqAMP (red) at 37°C after 24 h as function of extracellular FBA concentration. Readout is based on the fluorescence signal from cells recorded using a plate reader or a flow cytometer. B) Confocal images show the uptake of 2CNqAMP (top) and 2CNqATP (bottom) at the indicated time points after exposure at 37°C. C) Uptake of 2CNqATP (yellow) and 2CNqAMP (red) after 3 h at 37°C in the presence (+AXP) and absence (-AXP) of a 250-fold excess of canonical ATP (for 2CNqATP) or canonical AMP (for 2CNqAMP) measured using flow cytometry. Right: representative confocal microscopy images of the 2CNqAMP uptake. Same observation was made for 2CNqATP in competition with ATP. Scale bars = 20 µm.

However, performing experiments to study the uptake kinetics (based on microscopy and flow cytometry), we observed that the monophosphate is internalized twice as

fast, indicating that the phosphate groups play a role in the uptake process (Fig. 21B). Hence, to test if our analogues enter via natural ATP pathways and adenine transporters, we competed the uptake with an excess of ATP or AMP, correspondingly (Fig. 21C). This resulted in a 2-fold decrease of the intracellular 2CNqA signal, suggesting that competition with the natural counterpart occurs. Furthermore, the localization pattern of 2CNqA was altered upon the presence of ATP or AMP, indicating that competition also occurs inside the cells, *i.e.* after internalization. We also tested pATP uptake in competition with ATP under the microscope and made the same observation. This means that both analogues likely use the same pathway as the natural adenosine derivates to enter cells. In summary, we observed an active uptake of the A-analogues that is in competition with the natural counterparts, with the phosphate groups playing a role in the internalization process. By studying literature on ATP pathways, we hypothesized that our analogues enter the natural adenine salvage mechanism, which is part of the purinergic signalling pathway (Fig. 22). In this process, the phosphate groups of extracellular ATP are cleaved off consecutively, and each cleavage step leads to a signal inside the cells. In the last step, adenine gets internalized by nucleoside transporters to be re-phosphorylated inside the cells.

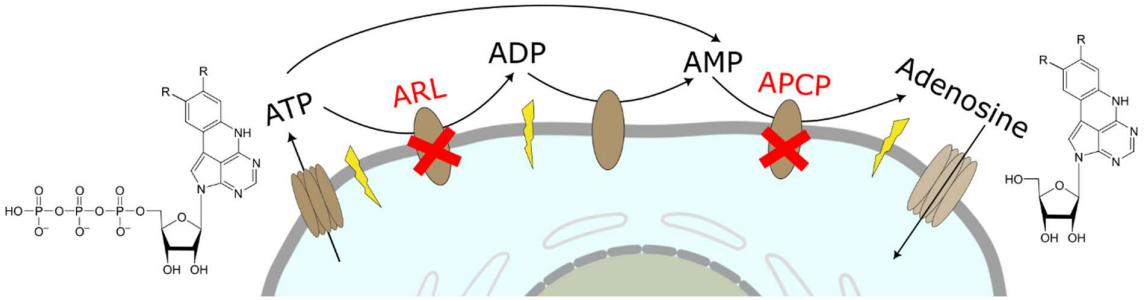

Figure 22. Illustration of the purinergic signalling pathway. ATP is released from the cells (or added to the cell medium) and the phosphate groups are cleaved consecutively where ATP, ADP, and adenosine activate receptors that trigger intracellular signalling (yellow flashes). Adenosine is then internalized via nucleoside transporters. In red are indicated the inhibitors that were tested during this thesis work. Molecular structure of adenine analogues is shown to highlight that they probably enter this pathway.

Following this hypothesis, we started experiments to investigate 2CNqA uptake while inhibiting enzymes involved in the salvage pathway (data unpublished). One compound is the commercially available ARL-67156, which is an antagonist of the ectoATPase that converts ATP to adenosine diphosphate (ADP) on the outside of the cell membrane (Fig. 22).[136] Another

investigation was done by inhibiting a ecto-5'-nucleotidase, an enzyme converting adenosine monophosphate (AMP) to adenosine using APCP (Adenosine-5'-(α,β-methylene)diphosphate) (Fig. 22).[137] The outcome of the inhibition experiments performed during the thesis work was inconclusive. Challenges were toxicity and the verification that these compounds work in the tested cell system (Huh-7 cell line). The targets of these inhibitors are enzyme families that are expressed differently in certain cells or tissues. Furthermore, not all enzymes of one family can be equally efficiently inhibited.[138–140] Also, control experiments are needed to exclude effects due to co-exposures of the inhibitors and the FBA that might result in misleading interpretation. Therefore, an elaborate screening of conditions to reduce cytotoxic effects and optimize the inhibition conditions using natural ATP or AMP is needed prior to further experiments including 2CNqATP/MP uptake. Likewise, inhibition of the nucleoside transporters might be a way forward to study the uptake pathway of 2CNqA and pA as well. Overall, these initial inhibition experiments highlight the complexity of enzymatic pathways and the challenges in studying them in living systems.

As a first step towards using FBAs as tools, we present experiments to establish 2CNqA for better understanding of intercellular processes in **Paper II**. We were able to monitor the effects of Actinomycin D (ActD) and $NaAsO_2$ on the nucleoli by imaging 2CNqA (Fig. 23). It was reported previously that both compounds inhibit the ribosome synthesis that takes place in the nucleoli, but that they activate two different stress regulation pathways.[141] This results in shrinking of the nucleoli under ActD exposure, but not under $NaAsO_2$. Using 2CNqA as a nucleoli dye we were able to monitor these differences by live-cell microscopy where previously antibody-labelling of nucleolar sub-compartments was used[141]. Analysis of images of 2CNqA exposed cells has shown that the nucleoli area decreases by a factor of 2–3 under ActD exposure.

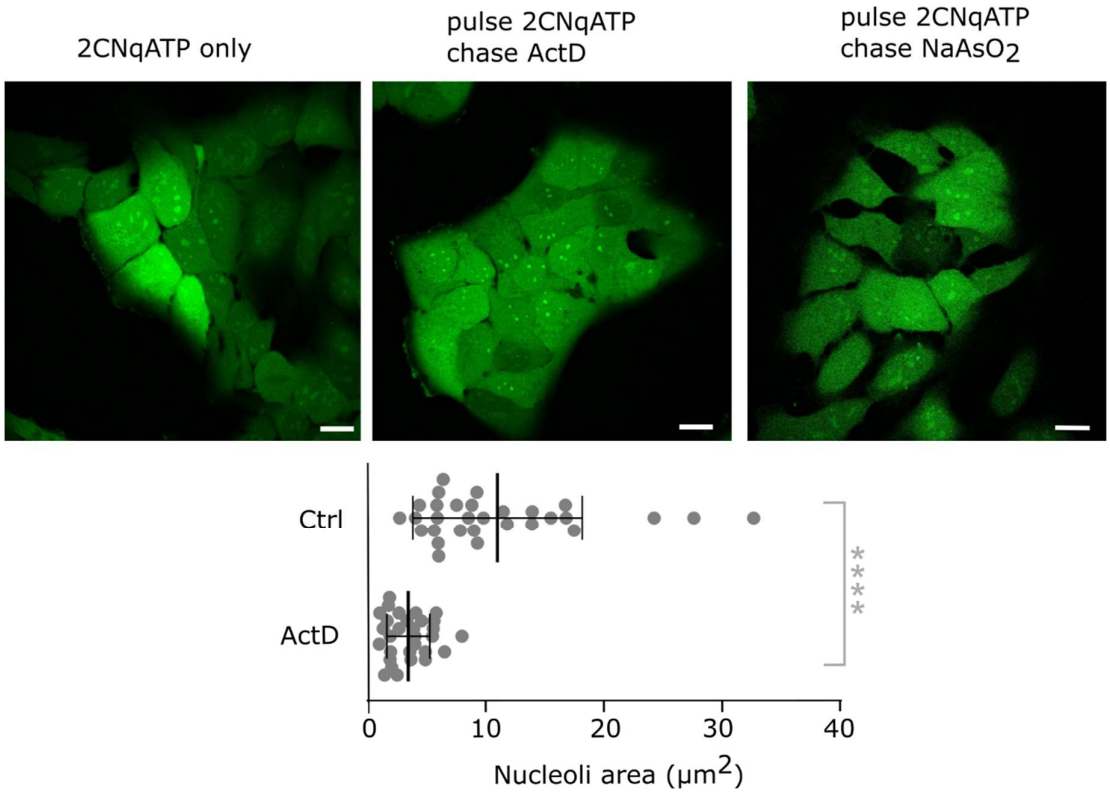

*Figure 23. The effect of Actinomycin D (ActD) and $NaAsO_2$ treatment visualized using 2CNqATP. Scale bars = 20 µm. To analyse the effect of ActD quantitatively, thirty nucleoli in three different images were measured, plotted and are presented with the mean and the standard deviation. The P-value was calculated using a two-tailed heteroscedastic t test; **** $P < 0.0001$.*

As the study of **Paper II** was very explorative and based on a serendipitous finding, we approached the investigations from several angles and with many ideas. The observed uptake, the start of its characterization, and the processing inside cells is a whole new field of study and opened the door for many explorations for the use of FBAs as research tools.

3.3.2 Metabolic RNA-labelling

2CNqA enters the cells and localizes in the cytosol, but also inside the nuclei and nucleoli, which are the sites of RNA synthesis (Ch. 3.2 and Ch. 3.3.1, **Paper II**). Hence, 2CNqA was a good candidate to investigate metabolic RNA labelling using FBAs. Therefore, total cellular RNA from cells that were exposed to 2CNqAMP and 2CNqATP over various time spans was extracted and afterwards investigated spectroscopically for 2CNqA signature (Fig. 24). Control experiments where the compounds were added to cell lysate, *i.e.* incorporation by living cells is not possible, were included.

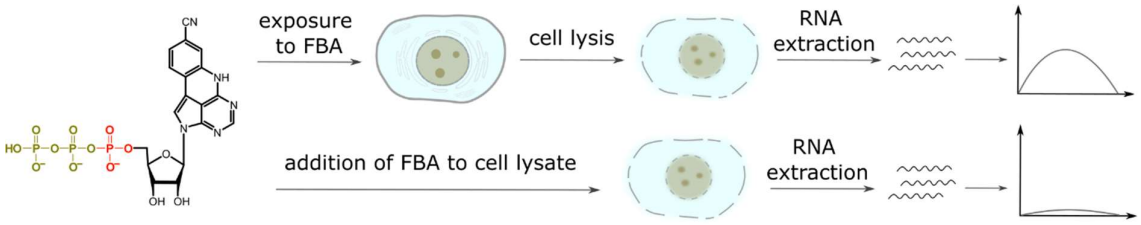

Figure 24. Experimental layout for cell treatment and RNA extraction. Top: cell exposure to 2CNqATP (green) or 2CNqAMP (red) for a certain time with subsequent cell lysis, RNA extraction, and spectroscopic readout. Bottom: control experiment where the compound was added to cell lysate prior to RNA extraction.

By recording emission spectra of the RNA extracts, we observed the FBA emission band around 460 nm (Fig. 25). This band increased with cell exposure time, *i.e.* the RNA became more fluorescent, while the emission from the control was very low. This was a first indication that 2CNqA gets incorporated into cellular RNA.

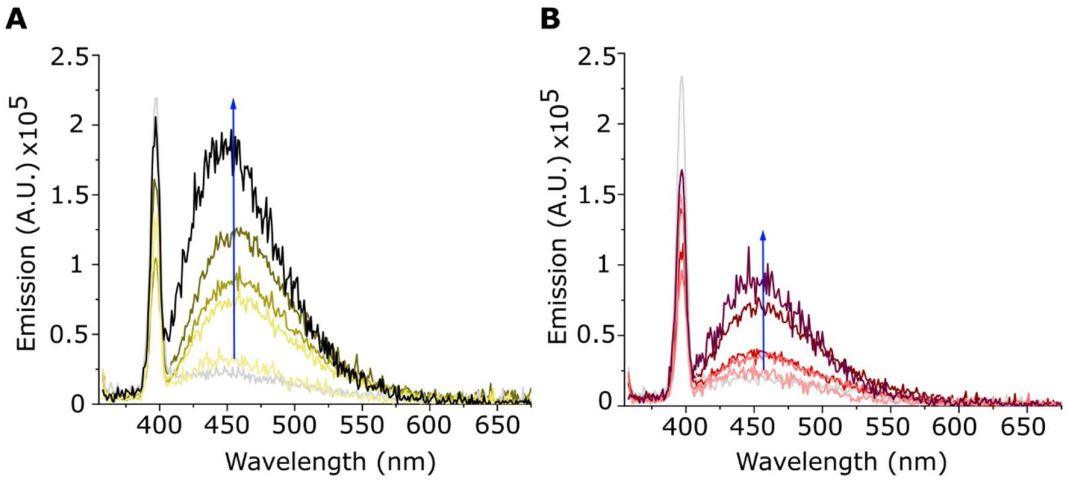

Figure 25. Emission spectra of purified RNA from cells exposed to A) 2CNqATP and B) to 2CNqAMP for 2, 8, 12, 24, 48 h (light to dark colour, blue arrows). Spectra are normalized to RNA absorption at 260 nm to account for RNA concentration differences. Grey spectra are of control experiment.

Upon further analysis, we observed a red shift of *ca.* 8 nm in the excitation spectra of extracted RNA compared to 2CNqATP monomer absorption (Fig. 26). To examine if this is an effect of incorporation, we were overlaying these spectra with absorption spectra of single- and double-stranded RNA oligonucleotides containing a single 2CNqA-base from a previous study. We observed an overlap of the 2CNqA-incorporated samples, and hence, the same shift compared to the triphosphate version. This highlights that the shift in absorption is a fingerprint of the slight sensitivity

of 2CNqA to its microenvironment and indicates that 2CNqA is indeed incorporated in the total cellular RNA extracts.

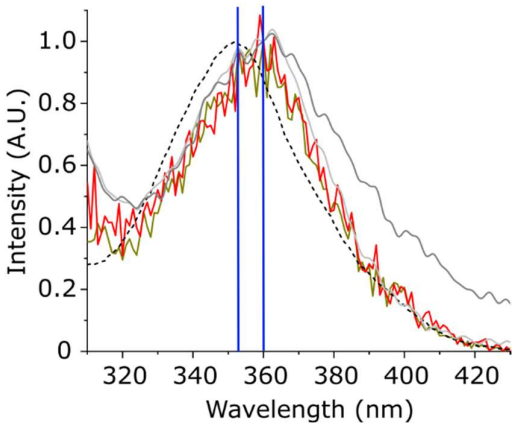

Figure 26. Normalized excitation spectra of the extracted RNA from cells exposed to 2CNqATP (yellow) and 2CNqAMP (red) at time point 24 h. For comparison, normalized absorption spectra of 2CNqATP (dashed black line), 2CNqA-ssRNA (dark grey line), and 2CNqA-dsRNA (light grey line) are added. Blue lines at peak maxima.

Having established this, we looked closer at the data at hand and found another interesting observation when comparing the emission spectra of 2CNqATP- and 2CNqAMP-exposed cells: the emission signal is lower from extracts of 2CNqAMP-exposed cells at all timepoints (Fig. 25), suggesting a lower degree of incorporation, even though the cellular uptake of this compound was more efficient, in the sense that it appeared faster. Preliminary experiments using FLIM to compare the lifetime species (Ch. 3.2.2) in images of 2CNqAMP-treated cells to compare to 2CNqATP-treatment have shown the same trend (data unpublished). Assuming that both compounds are converted to the 2CNqA nucleoside in the uptake process, both derivatives experience the same need for intracellular re-phosphorylation and transport. This surprising observation of faster apparent uptake kinetics (Ch.3.3.1), but lower efficiency of RNA incorporation, suggests a more complex cellular process, and at the point this thesis was written we cannot fully explain this discrepancy. A potential direction for future studies addressing this could involve investigations on the effects of elevated triphosphate or monophosphate concentrations in the extracellular environment. During purinergic signalling, the cleavage of phosphate groups outside the cell membrane can trigger intracellular signalling pathways (Fig. 22), which may, in turn, stimulate RNA synthesis[142] and facilitate the incorporation of 2CNqA. A more detailed analysis of our co-exposure experiments with ATP and AMP could provide initial insights into this mechanism.

In a next step, we used the spectroscopic data from our extracted RNA samples to determine the incorporation degree of 2CNqA. By constructing a 2CNqATP standard curve we were able to estimate the efficiency of 2CNqA incorporation into cellular RNA. We calculated that one of every 10^4–10^5 bases is an 2CNqA-base after 24 h of exposure to 2CNqATP. This number only includes 2CNqA nucleotides incorporated into all newly synthesized RNA molecules and is hence affected by long-lived RNA that was present prior to 2CNqA exposure. Furthermore, this value is not exact, as accurate quantification poses a challenge: many parameters for the standard curve need to be considered, and thorough control experiments for exact sample readout need to be performed.

To investigate if 2CNqA ribonucleotides are converted into DNA building blocks inside cells, and therefore are no longer available for RNA synthesis, we co-exposed cells to hydroxyurea, which inhibits the ribonucleotide reductase, *i.e.* the conversion of ribonucleotide diphosphates to their corresponding 2′-deoxy derivatives[143]. We did not observe any effects on the cellular localization of 2CNqA, nor on the emission profile of the RNA extracts, suggesting that 2CNqA remains as ribonucleotide inside cells. However, as described in several parts of this thesis, such co-exposures require a detailed screening, and hydroxyurea is known to have other effects on cells as well[144] that might have to be considered for further conclusions. To assess possible effects of 2CNqA internalization and RNA incorporation on different RNA types that are synthesized inside cells, we also performed agarose gel electrophoresis. This revealed that the RNA size distribution is unaltered compared to extracts from unexposed cells. The 2CNqA-RNA extracts display the typical pattern of total RNA extracts from human cells, showing clear bands for 28S rRNA, 18S rRNA, and smaller RNAs. Overall, 2CNqA does not seem to alter intercellular processes and is apparently accepted as RNA building block inside cells.

In summary, the results presented in **Paper II** highlight the use of 2CNqA as a label for RNA and to the best of our knowledge, this was the first time metabolic labelling was successful in wildtype, *i.e.* non-engineered, cells without assisted uptake of nucleosides. It is particularly stunning, because 2CNqA upon cellular uptake must be processed by several cellular enzymes (*e.g.* kinases for phosphorylation, polymerases) to get incorporated into RNA by the cellular machinery. Our readout methods involved standard techniques and are, hence, available in most biochemistry laboratories, which is an important advantage for future applications. This also means that other advanced methods can be applied to investigate our findings further and use them to monitor RNA metabolism, as for example RNA sequencing methods, or

— as we have started to do — using FLIM as imaging approach (**Paper V**, Ch. 3.2.2). To investigate the distribution of 2CNqA in different kinds of RNA (tRNA, rRNA, mRNA, *etc*), more specific RNA analysis, like capillary electrophoresis with fluorescence detection, liquid chromatography, or hybridization capture technologies, can be performed in future studies.

3.3.3 Enzymatic *in vitro* Labelling

Labelling of RNA inside cells by fluorescent building blocks (Ch.3.3.2) results in fluorescent RNA of all kinds. To label RNA more specifically, RNA synthesis including the labelling is done *ex cellulo*. As described earlier (Ch. 2.2.5.1), shorter oligonucleotides can be synthesized and labelled via solid phase synthesis, which allows the study of *e.g.* labelled ASOs as presented by Nilsson *et al.*[86] and as we presented in **Paper I**, or structural studies of new analogues inside oligomers as presented in **Paper IV**. For synthesis of long RNA, enzymatic *in vitro* reactions are used where a DNA template dictates the resulting RNA sequence (Ch. 2.2.5.1). In 2021 Baladi *et al.* presented the use of $tC^{O}TP$ to label RNA by *in vitro* transcription (IVT).[87] Having 2CNqATP presented as analogue for RNA labelling inside cells (**Paper II**), we tested it for *in vitro* labelling, which resulted in **Paper III**. We used a well-established IVT system that is based on the T7-phage RNA polymerase, and that is available as a kit containing all reagents, except the DNA template. Our in-house designed DNA template codes for the fluorescent mCherry protein and contains the alpha globulin 3'- and 5' UTRs for increased translation, which are based on the UTRs of mRNA vaccines[145]. We also included the Kozak sequence for translation initiation in eukaryotic systems, as well as the Shine-Dalgarno sequence for expression in prokaryotic systems.

By adding different fractions of 2CNqATP to the ATP pool of the IVT, we observed the possibility to label RNA fluorescently using our adenine analogue. We analysed the synthesized long RNA spectroscopically and by gel electrophoresis. The resulting RNA has the expected length of 884 nucleotides at all tested 2CNqATP:ATP ratios, which indicates that the T7-polymerase can synthesize full-length RNA in the presence of 2CNqA (Fig. 27A). Spectroscopic analysis of the RNA has shown that 2CNqA gets incorporated, since we observed an absorption band around 360 nm stemming from 2CNqA absorption (Fig. 27B). This peak displayed a red shift of 10 nm compared to 2CNqATP, as we observed also in **Paper II** (Ch. 3.3.2, Fig. 26), which is a fingerprint for 2CNqA incorporation. Using the absorption values at 260 nm (where all canonical

nucleobases and 2CNqA absorb) and at 360 nm (2CNqA absorption only), and the corresponding molar absorptivities, we calculated the fraction of 2CNqA-bases inside the long RNA. This fraction rises with increasing 2CNqA content in the IVT reaction, which we could fit to a linear equation with a slope of 0.39 (Fig. 27C).

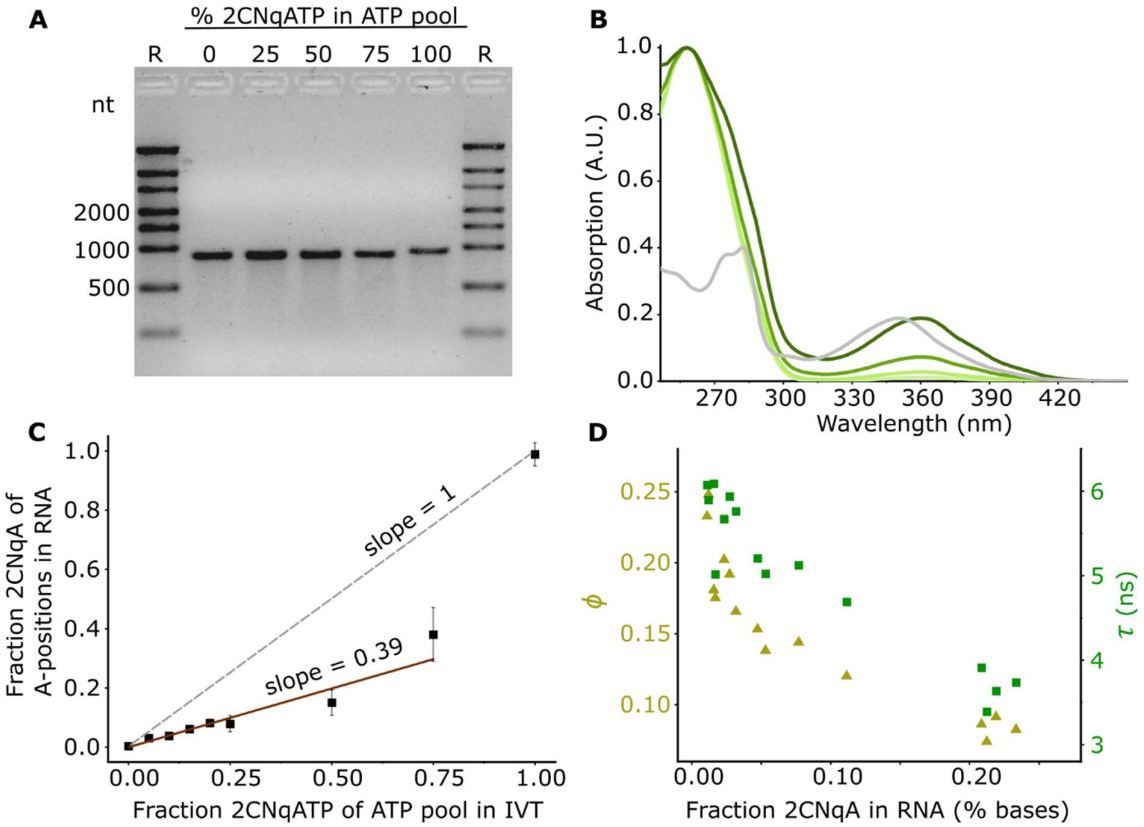

Figure 27. A) Agarose gel electrophoresis of resulting RNA from one representative IVT reaction set. R: RiboRuler, nt: nucleoitdes B) UV-vis spectra normalised to the absorption at 260 nm. 25%, 50%, 75%, and 100% 2CNqATP of the ATP pool in the reactions (light to dark colour). In grey, absorption spectrum of the 2CNqATP monomer normalised to the 2CNqA peak at 360 nm in the 100% 2CNqA-RNA spectrum for comparison. C) 2CNqA incorporation degree as mean value of several reactions, expressed as percentage of adenine positions in the resulting RNA sequence, plotted against the fraction of 2CNqATP of the ATP pool in the IVT reaction. Red line presents the linear fit, excluding the reaction where all ATP was replaced by 2CNqATP in the IVT. The dashed line depicting a 1:1 reactivity is included for comparison. D) Fluorescence quantum yields (ϕ, yellow), and amplitude-weighted fluorescence lifetimes (τ, green) of individual 2CNqA-labelled RNA samples plotted against the 2CNqA incorporation degree.

This represents an apparent reactivity of 1:2.6 between the modified and natural ATP and indicates that the T7-polymerase has a slight preference for the natural ATP, which is

furthermore supported by a lower amount of resulting RNA at higher 2CNqATP:ATP ratios. Previously in our group, we observed that tC^O is equally well incorporated as natural C[87], which is in contrast to the A-analogue. A comparison to the commercially available Cy5-CTP revealed that the size of the modified building blocks, as well as pi-stacking and hydrophobicity, play a role for the incorporation efficiency (**Paper III**). For all RNA samples where 2CNqA was present in the IVT, we observed an emission signal with a maximum around 460 nm, which is in accordance with the reported emission of 2CNqA[82,83] (compare previous chapters). Because quantum yield and lifetime are important parameters for fluorescence-based detection methods, like fluorescence microscopy, we determined both parameters for all labelling degrees (Fig. 27D). We observed that they depend on the incorporation degree and display a decrease with increasing 2CNqA incorporation, which we ascribe to dynamic quenching. This leads to a non-linearity between incorporation degree and brightness, which is important to consider for applications where overall brightness is crucial.

Next, we were interested in whether the 2CNqA-labelled RNA can be translated into a functional protein inside cells. Therefore, we capped and tailed the RNA to achieve translation competent mRNA (Ch. 2.2.5.1) and transfected cultured cells using lipofectamine. 2CNqA-labelled RNA with an incorporation degree of 1%–2% was detectable under the confocal microscope after 30 min exposure time using 405 nm excitation (Fig. 28A). After five hours, diffuse mCherry fluorescence was detectable across the cytosol and nuclei, demonstrating that our labelled RNA can be translated into a fluorescent protein. However, we observed weak mCherry signal from the labelled RNA compared to unlabelled RNA and therefore, quantified the translatability using flow cytometry (Fig. 28B, red). This revealed that the mCherry intensity drops by a factor of ten for the lowest investigated incorporation degree (0.8%), compared to mCherry signal from unlabelled RNA. Interestingly, the number of cells displaying mCherry signal is affected comparably little (Fig. 28B, blue). We concluded that 2CNqA affects, yet does not completely inhibit, the translation of the mRNA. To exclude that a non-negligible fraction of unlabelled RNA is present in samples from 2CNqA-IVTs, which could contribute to the mCherry signal, we calculated the probability of unlabelled strands at different labelling degrees using binomial statistics. At the lowest incorporation degree of 0.8% this probability is 0.08%, meaning that the observed mCherry fluorescence most likely originates from the translation of labelled RNA.

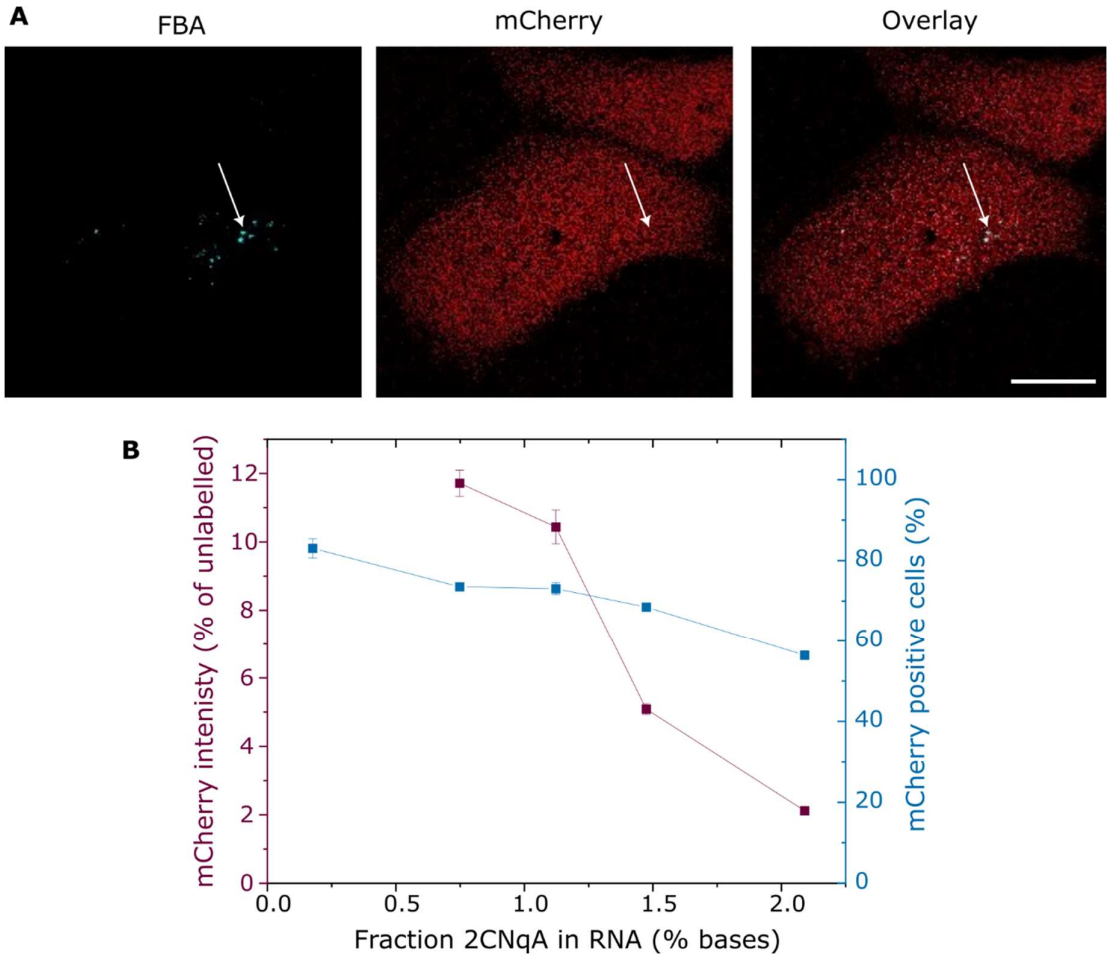

Figure 28. A) Live-cell confocal fluorescence microscopy of Huh-7 cells exposed to 2CNqA-labelled mRNA five hours after exposure to 1.1%-labelled 2CNqA-RNA. 2CNqA displayed in cyan (arrow) and mCherry in red. Scale bar = 20 µm. B) Flow cytometry analysis of Huh-7 cells exposed to mRNA labelled with 2CNqA at different incorporation degrees. Cells were exposed to lipofectamine-formulated mRNA for 24 h and all events were gated for single living cell population for analysis.

Testing the translation of mRNA labelled with Cy5 at C-positions, we observed that the translation efficiency is negatively affected to the same extent as it was for 2CNqA. Fluorophore size leading to steric hindrance, hydrophobicity, or stabilization/destabilization of RNA structures can be effects that result in perturbed translation. However, more detailed structural studies are needed to investigate the molecular basis of these results.

Taken together, we were able to label long RNA using our adenine analogue 2CNqA by a standard IVT reaction and we could tune the labelling degree by adjusting the 2CNqATP:ATP ratio in the reaction. With the previously reported tC^O as a label for C-positions in the

sequence[87], we present with this work increased sequence flexibility for RNA labelling. As shown in other projects of this thesis, 2CNqA is a suitable multiphoton probe (**Paper I**) and can be used for FLIM (**Paper V**). Hence, having 2CNqA-labelled RNA at hand, studies based on these two imaging techniques can be carried out.

As described above, placing 2CNqA in the coding part of mRNA had a negative effect on the translatability of the RNA. Therefore, in continuation of the work presented in **Paper III**, we test in ongoing work different enzymes to label RNA in non-coding regions. Our hypothesis is that FBAs incorporated into the UTRs or the poly(A) tail do not disturb translation as much as we observed when FBAs are present in the open reading frame of the RNA. Therefore, we started testing different enzymes and various conditions, based on extensive literature research. tC^O was successfully added onto the 3'-ends of short RNA oligonucleotides using the enzyme terminal deoxynucleotidyl transferase (TdT)[87], which is therefore a good candidate for 2CNqA-labelling as well. Another enzyme we found promising was the poly(U) polymerase from *Schizosaccharomyces pombe*, which was shown to add modified nucleotides to the 3'-end of RNA.[146] Furthermore, as we use an *Escherichia coli* poly(A) polymerase for adding the 3'-poly(A) tail to RNA, and have 2CNqA as an adenine analogue at hand, adding 2CNqA to the tailing reaction is a promising labelling strategy. We also include a poly(A) polymerase form S*accharomyces cervisiae* to test an eukaryotic enzyme. The resulting RNA is analysed by gel electrophoresis, UV-vis absorption spectroscopy, and steady-state emission readout. At the point the thesis was written, we observe only low labelling degrees in all cases. By gel electrophoresis we see that the increase in length is minimal, meaning that only a few nucleotides are added to the long strands. We are, however, able to detect emission signals, and a small peak outside the RNA absorption bands corresponding to the low-energy absorption bands of the FBAs. This suggests that it is possible to add FBAs onto long RNA using the above-mentioned enzymes. To optimise incorporation, reactions are currently being carried out in which different cofactors, amounts of triphosphate and FBAs, as well as enzyme units are varied. Furthermore, it is nowadays standard in many laboratories to encode the poly(A) tail in the DNA template as poly(T) (Ch. 2.2.5.1). Adding 2CNqA to IVT reactions using such a template could lead to labelled RNA where a large fraction of fluorescent labels is located in the tail. This would render the RNA fluorescent while the coding region displays only a low labelling degree, which should leave the mRNA translation more unperturbed. In summary, there are ways forward and many experiments are left to perform to overcome the reduced translatability we observed for 2CNqA incorporated into the coding region.

4. ADDITIONAL WORK AND OUTLOOK

Many possible future investigations following the original work of this thesis were described in chapter 3 and in the attached publications. During the time that the experimental part of this thesis was performed, ideas to explore the applications of fluorescent base analogues in various biological systems emerged and initial investigations into some of them were started. This chapter gives an overview of three side projects that highlight the manifold options for future research.

Following the spontaneous uptake of FBA nucleotides (**Paper II**), we investigated if it is possible to detect 2CNqA-labelled bioparticles that originate from cells that have internalized 2CNqATP. This laboratory work was done as collaborations with groups from the Department of Life Sciences at Chalmers and the Department of Microbiology and Immunology at the University of Gothenburg. We infected cells with virus particles after the cells had internalized 2CNqATP and then harvested the forming virus particles and tried to detect 2CNqA fluorescence inside them. We also extracted RNA from the particles to see if we can detect metabolically labelled RNA. In a similar approach, cells were exposed to 2CNqATP and after 24 h incubation time extracellular vesicles (EVs) were harvested. The work with such EVs and virus particles is challenging for two main reasons: i) these particles are very small (nm range) and in combination with the relatively low brightness of our FBAs (compared to large conventional fluorophores, such as GFP, that are often used) the readout becomes very difficult and ii) in the case of virus particles, biohazard regulations had to be considered and required many safety considerations. Therefore, the work performed during this thesis work focused on setting up the detection methods using inactivated viruses and EVs. We realized early on that it is of outermost importance to include conclusive negative and positive controls at all stages (*e.g.* particles labelled with another dye than the FBAs, or controls for autofluorescence level). Readout methods we started with included small particle analysis using a flow cytometer, fluorescence microscopy, nanoparticle tracking analysis, and spectroscopic analyses. The results of our preliminary experiments are promising, but more work is needed to set up innovative detection methods to prove that FBA derivatives are present inside such particles. This could include the immobilization of the particles on surfaces or using a flow device for advanced microscopy readouts. Once these hurdles are overcome, the approach of using 2CNqA-exposed cells for labelling small particles presents a direct

non-invasive readout, that complement protein engineering approaches and coupled-reaction assays.

Another idea that emerged from the spontaneous uptake in human (eukaryotic) cells (**Paper II**) was to test if the spontaneous internalization also occurs in bacteria (prokaryotic cells). For this, I collaborated with Dr. Ann-Britt Schäfer at the Department of Life Sciences at Chalmers. We started using the gram-positive bacterium *Bacillus subtilis*, since it is the established model organism in the laboratory. We added 2CNqATP to bacteria in regular growth medium, as well as to bacteria in starvation medium to enhance possible uptake. Image analysis revealed a significant increase in fluorescence signal in bacteria exposed to 2CNqATP under starvation conditions (Fig. 29).

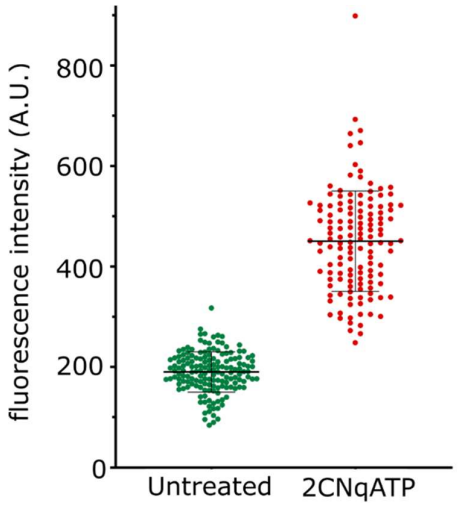

Figure 29. Bacterial uptake of 2CNqATP. Bacillus subtilis *in starvation medium was exposed to 2.5 µM for 1 h and then imaged under an epifluorescence microscope. Three images per condition were analysed using MicrobeJ where debris, doublets and out of focus cells were excluded.*

An advanced screening for future investigations should consider different concentrations, time points, different media, and complementary analysis techniques. Furthermore, because bacteria are very diverse, testing the uptake in different bacteria or strains would be needed to assess the potential of our observation. Many research questions of such a project would be similar to the uptake in human cells (*e.g.* which uptake pathways are involved, dose response to treatments and equilibria, possibility of metabolic labelling of RNA) and could put FBAs in the light of tools for research in the field of, for example, antibiotic resistance[147]. Another interesting approach would be to test the uptake in plant cells in order to obtain a more

comprehensive overview on how chemical modifications of base analogues affect the interaction with various cells. If the spontaneous uptake happens there as well, FBAs could also be an interesting tool in plant RNA research, for example how plants respond to environmental changes[148], or to develop RNA vaccines for plants[149].

Based on our *in vitro* experiments using enzymes to label RNA (**Paper III,** Ch. 3.3.3) grew the idea of using enzymatic reactions to modify the base structure of a nucleobase analogue to create a new and possibly fluorescent base. In a first try, I used the non-fluorescent analogues tC_{nitro} and qA_{nitro} [129,150] and the enzyme nitroreductase (from *Escherichia coli*) to reduce the nitro groups of these molecules to amino groups. Following the reaction spectroscopically, I observed an absorption band appearing around 560 nm during the reaction, which suggests a newly formed compound in the reaction mix. The main challenge of the spectroscopy-based readout is that the enzyme needs $NADH+H^+$ for the reaction, which is a dinucleotide that absorbs and fluoresces in the same wavelength range as most FBAs. However, I observed oxidation of $NADH+H^+$ to the non-fluorescent NAD^+ during the reaction, which was a good indicator that the reaction is working. Further analysis of the resulting molecules (foremost "tC_{amino}") using a suited purification protocol (*e.g.* column chromatography, HPLC) and analysis (*e.g.* NMR) are needed. Other enzymes and analogues can be tested as well, or possibly a coupled enzymatic reaction, like the nitroreductase with a deaminase to create hydroxy groups. A computational approach (*e.g.* density-function calculations[82]) can help to predict and compare the resulting spectra of new molecules from these reactions. This approach, together with the observed cellular uptake of base analogues, could be a way to develop assays to monitor specific reactions inside cells.

5. CONCLUDING REMARKS

In this thesis, I present research on fluorescence base analogues (FBA) for studying nucleotide and RNA metabolism. This work stretches from fluorophore characterization to methodological investigations using different microscopy techniques, and eventually to biological studies inside living cells. In total, this thesis is based on five interconnected articles.

Spectroscopic characterizations were performed for **Paper I**, **Paper II**, and **Paper IV**. **Paper IV** presents the characterization of the new FBA qU inside RNA oligomers where we show that it displays complex photophysics inside oligonucleotides as it does as a monomer. Its iminol tautomer is pronounced inside oligomers which leads to destabilization of duplexes. However, continued characterization based on this thesis work (*e.g.* pH dependency experiments) may make this analogue an interesting candidate for future applications. **Paper I** and **Paper II** present the continued characterization of the triphosphates pATP and tC^{O}TP, but especially of our analogue 2CNqA as monomer, *i.e.* as nucleoside triphosphate (2CNqATP) and monophosphate (2CNqAMP), as well as inside ASOs, including single- and two-photon spectroscopy.

Paper II was the central focus of my work, where we present, besides the photophysical characterization, the spontaneous cellular uptake of our two adenine analogues: pA and 2CNqA. In addition, we demonstrate the versatility of 2CNqA for studies that are based on imaging, flow cytometry readout, and spectroscopic techniques. We explored the potential for metabolic labelling of RNA following the uptake of 2CNqATP and 2CNqAMP by cells. Beyond the publication, we also filed a patent application on fluorescent adenosine triphosphates to study biological mechanisms, where my results contributed. The startup LanteRNA (Stealth Labels Biotech AB) is currently commercializing these findings.

The project that **Paper II** is based on encompassed a broad range of experiments and conceptual approaches that aimed at deepening the understanding of our observations and exploring potential applications, and not all results are included in the article. One of the follow-up projects led to a communication article, **Paper V,** where we demonstrate the applicability of 2CNqA in fluorescence lifetime imaging of living cells. This study provided more detailed insights into the fluorescent species of 2CNqA present in different cellular compartments and marks the first use of an FBA

for FLIM in living cells. As emphasized in the thesis, there remains significant potential to combine the findings of **Paper I** (two-photon imaging), **Paper II** (observed uptake), and **Paper V** (FLIM) to pave the way for applications of 2CNqA in cellular studies.

The work of **Paper III** presents a different application of 2CNqA, focusing not on cellular uptake, but on its use in *in vitro* reactions. We demonstrate the labelling of RNA via *in vitro* transcription using 2CNqATP and compare its performance to our previously published tC^{O} [87] and the widely used Cy5 dye. The labelled RNA was delivered to cells, where we were able to detect it using confocal microscopy. We also monitored subsequent protein synthesis and found that, similar to Cy5, 2CNqA exerts an inhibitory effect on protein production. In the thesis, I also present an outlook with ideas on how to overcome this limitation, which is ongoing work in the Wilhelmsson group and at LanteRNA. Together with the findings in **Paper I** (two-photon imaging) and **Paper V** (FLIM) 2CNqA shows promise as a versatile tool to study long RNAs.

In summary, chapters 3 and 4 of this thesis highlight the opportunities and the remaining challenges that result from my work. Especially the focus on studies in living cells using FBAs presents pioneer work that is ongoing in our group and in the FBA field. Many of our observations were both surprising and intriguing, and while not all can be fully explained at the current stage, we present promising initial results for future investigations which may inspire further research across various fields. Besides the FBAs that are presented and applied in this thesis, my work also contributes to a broader understanding on how base analogues — and potentially other small molecule derivatives or drugs — can be designed for intended applications. Therefore, the herein presented results bridge the gap between molecular design, synthesis, and initial characterization of FBAs (compare doctoral thesis by Dr. Moa S. Wranne[151]), and their functional evaluation in complex biological systems, such as living cells. Looking ahead, one promising direction to complement this work is the use of computational design to predict new molecules based on the insights gained in this thesis. This iterative refinement, from design to application and back to structural optimization, holds strong potential for establishing a robust structure–activity relationship framework for future FBA development.

ACKNOWLEDGEMENTS

I would like to thank everyone who has contributed to my education and has shaped me as a scientist and person. You are so many — it is just impossible to name everyone here.

The person who stands out the most is my fantastic supervisor, **Marcus Wilhelmsson**, who welcomed me as bachelor-level Erasmus trainee, guided me through my master thesis, and provided me strong supervision throughout my PhD studies. I could not have completed this work without your endless encouragement and mentorship.

I am grateful to all the people who supervised and guided me on the path of becoming a doctor. Special thanks to my co-supervisor **Elin Esbjörner,** who accepted me as master thesis student, and with that opened the door to FBA research in living cells for me, and to **Fredrik Westerlund,** for being a fantastic examiner and for the thorough reading of the thesis draft.

Jesper, for being my partner in crime in nearly all my scientific projects — as creative mind, colleague, and supervisor. You deserve a medal for performing lab work with the person having world's worst patience. Many thanks for being such a supportive friend through all these years.

All the past and present group members of the Wilhelmsson group. Especially **Moa** and **Anders**, who introduced me as trainee to the world of FBAs, and **Tom,** for the feedback on the chemical structures in my figures and patiently explaining organic chemistry concepts of FBAs.

Audrey, for being an amazing master thesis supervisor and for teaching me so many of the methods and techniques that became central to this thesis.

All my office mates over the years, especially **Wera** and **Hanna,** for the many fruitful scientific chats and discussions and for the support during the most intense times of being a PhD student.

Felix, **Jesper, Sofi**, and **Aleks** at LanteRNA, for always seeing me as a part of your little gang.

All the wonderful people in the chemistry building whom I have had the pleasure to meet and work with over the past eight years. You are too many to name here, but I am grateful to each and every one of you!

My flat mates, for bearing with me and my moods at home — special thanks to **Sandra**, for keeping the music in my life. Also, **Stenkolsateljén** on Ringön, for the creative space that became my second home and office.

UNDER Lea's TRUST, for making this world a better place and for reminding me to never take my good education for granted. I am grateful for being part of your team which gives me the opportunity to grow and gain skills beyond academia.

All my friends in Sweden, Germany, and wherever you are right now, for always being there for me, even across long distances.

And last, but by no means least, my family — who *always* believes in me. Special thanks to the scientist **Papa Tim**, **Mama Katy,** and my sister **Lisa-Lou**, who so often gave me new perspectives on science and continuously nurtured my scientific fascination, like pointing out the belemnite fossils in the stairs at the chemistry entrance at Chalmers.

BIBLIOGRAPHY

1. Chawla, M., *et al.* An atlas of RNA base pairs involving modified nucleobases with optimal geometries and accurate energies. *Nucleic Acids Research* **43**, 6714–6729 (2015).
2. Hintz, H., *et al.* Architecture of RNA. in *Fundamentals of RNA Structure and Function* (ed. Grover, N.) 21–48 (Springer International Publishing, Cham, 2022).
3. Vicens, Q. & Kieft, J. S. Thoughts on how to think (and talk) about RNA structure. *Proceedings of the National Academy of Sciences* **119**, e2112677119 (2022).
4. Crick, F. H. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).
5. Doherty, E. A. & Doudna, J. A. Ribozyme Structures and Mechanisms. *Annual Review of Biophysics and Biomolecular Structure* **30**, 457–475 (2001).
6. Cai, Y., *et al.* A Brief Review on the Mechanisms of miRNA Regulation. *Genomics, Proteomics & Bioinformatics* **7**, 147–154 (2009).
7. Xiao, L., *et al.* Disorders and roles of tsRNA, snoRNA, snRNA and piRNA in cancer. *Journal of Medical Genetics* **59**, 623–631 (2022).
8. Fernandes, J. C. R., *et al.* Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. *Noncoding RNA* **5**, 17 (2019).
9. Chen, S. *et al.* The biogenesis and biological function of PIWI-interacting RNA in cancer. *J Hematol Oncol* **14**, 93 (2021).
10. Nelson, J. W. & Breaker, R. R. The lost language of the RNA World. *Science Signaling* **10**, eaam8812 (2017).
11. Wang, F., *et al.* RNA therapeutics on the rise. *Nature Reviews Drug Discovery* **19**, 441–442 (2020).
12. Shi, Y. *et al.* Chemically Modified Platforms for Better RNA Therapeutics. *Chem. Rev.* **124**, 929–1033 (2024).
13. Hastings, M. L. & Krainer, A. R. RNA therapeutics. *RNA* **29**, 393–395 (2023).
14. Commissioner, O. of the. FDA Takes Key Action in Fight Against COVID-19 By Issuing Emergency Use Authorization for First COVID-19 Vaccine. *FDA* https://www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19 (2020).
15. Verbeke, R., *et al.* Innate immune mechanisms of mRNA vaccines. *Immunity* **55**, 1993–2005 (2022).
16. Wolff, J. A. *et al.* Direct gene transfer into mouse muscle in vivo. *Science* **247**, 1465–1468 (1990).
17. Al Fayez, N. *et al.* Recent Advancement in mRNA Vaccine Development and Applications. *Pharmaceutics* **15**, 1972 (2023).
18. Bitounis, D., *et al.* Strategies to reduce the risks of mRNA drug and vaccine toxicity. *Nat Rev Drug Discov* **23**, 281–300 (2024).
19. Dowdy, S. F. Overcoming cellular barriers for RNA therapeutics. *Nat Biotechnol* **35**, 222–229 (2017).
20. Hou, X., *et al.* Lipid nanoparticles for mRNA delivery. *Nat Rev Mater* **6**, 1078–1094 (2021).
21. Samaridou, E., *et al.* Lipid nanoparticles for nucleic acid delivery: Current perspectives. *Advanced Drug Delivery Reviews* **154–155**, 37–63 (2020).

22. Mehta, M. *et al.* Lipid-Based Nanoparticles for Drug/Gene Delivery: An Overview of the Production Techniques and Difficulties Encountered in Their Industrial Development. *ACS Mater Au* **3**, 600–619 (2023).
23. Karlsson Hedestam, G. & Sandberg, R. Discoveries concerning nucleoside base modifications that enabled the development of effective mRNA vaccines against COVID-19. https://www.nobelprize.org/prizes/medicine/2023/advanced-information/ (2023).
24. Karikó, K., *et al.* Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23**, 165–175 (2005).
25. Karikó, K. *et al.* Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol Ther* **16**, 1833–1840 (2008).
26. Karikó, K., *et al.* Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Research* **39**, e142 (2011).
27. Roberts, T. C., *et al.* Advances in oligonucleotide drug delivery. *Nat Rev Drug Discov* **19**, 673–694 (2020).
28. Kulkarni, J. A. *et al.* The current landscape of nucleic acid therapeutics. *Nat. Nanotechnol.* **16**, 630–643 (2021).
29. Crooke, S. T., *et al.* Mechanisms of Antisense Drug Action, an Introduction. in *Antisense Drug Technology* (CRC Press, 2007).
30. Geary, R. S., *et al.* Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv Drug Deliv Rev* **87**, 46–51 (2015).
31. Juliano, R. L., *et al.* Cellular uptake and intracellular trafficking of antisense and siRNA oligonucleotides. *Bioconjug Chem* **23**, 147–157 (2012).
32. Lim, K. R. Q. & Yokota, T. Invention and Early History of Gapmers. in *Gapmers: Methods and Protocols* (eds. Yokota, T. & Maruyama, R.) 3–19 (Springer US, New York, NY, 2020).
33. Vinjamuri, B. P., *et al.* Review on Commercial Oligonucleotide Drug Products. *Journal of Pharmaceutical Sciences* **113**, 1749–1768 (2024).
34. Kamzeeva, P. N., *et al.* Recent Advances in Molecular Mechanisms of Nucleoside Antivirals. *Curr Issues Mol Biol* **45**, 6851–6879 (2023).
35. Eastman, R. T. *et al.* Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19. *ACS Cent Sci* **6**, 672–683 (2020).
36. Grein, J. *et al.* Compassionate Use of Remdesivir for Patients with Severe Covid-19. *N Engl J Med* **382**, 2327–2336 (2020).
37. Koczor, C. A., *et al.* The Role of Transporters in the Toxicity of Nucleoside and Nucleotide Analogs. *Expert Opin Drug Metab Toxicol* **8**, 665–676 (2012).
38. Jones, R. J. & Bischofberger, N. Minireview: nucleotide prodrugs. *Antiviral Research* **27**, 1–17 (1995).
39. Mehellou, Y., *et al.* The ProTide Prodrug Technology: From the Concept to the Clinic. *J. Med. Chem.* **61**, 2211–2226 (2018).

40. Rautio, J., *et al.* The expanding role of prodrugs in contemporary drug design and development. *Nat Rev Drug Discov* **17**, 559–587 (2018).
41. Kell, D. B. & Dobson, P. D. The Cellular Uptake of Pharmaceutical Drugs is Mainly Carrier-mediated and is thus an Issue not so Much of Biophysics but of Systems Biology. *Beilstein-Institut* 149–168 (2009).
42. Zhang, Y., *et al.* Current prodrug strategies for improving oral absorption of nucleoside analogues. *Asian Journal of Pharmaceutical Sciences* **9**, 65–74 (2014).
43. Meier, C. Nucleoside diphosphate and triphosphate prodrugs – An unsolvable task? *Antivir Chem Chemother* **25**, 69–82 (2017).
44. E. Kleiner, R. Interrogating the transcriptome with metabolically incorporated ribonucleosides. *Molecular Omics* **17**, 833–841 (2021).
45. Williams, C. G., *et al.* An introduction to spatial transcriptomics for biomedical research. *Genome Medicine* **14**, 68 (2022).
46. Paige, J. S., *et al.* RNA Mimics of Green Fluorescent Protein. *Science* **333**, 642–646 (2011).
47. Femino, A. M., *et al.* Visualization of Single RNA Transcripts in Situ. *Science* **280**, 585–590 (1998).
48. Choi, J.-S. & Berdis, A. J. Visualizing nucleic acid metabolism using non-natural nucleosides and nucleotide analogs. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1864**, 165–176 (2016).
49. Kuchta, R. D. Nucleotide Analogues as Probes for DNA and RNA Polymerases. *Curr Protoc Chem Biol* **2**, 111–124 (2010).
50. Feldman, A. W. *et al.* A Tool for the Import of Natural and Unnatural Nucleoside Triphosphates into Bacteria. *J. Am. Chem. Soc.* **140**, 1447–1454 (2018).
51. King, A. E., *et al.* Nucleoside transporters: from scavengers to novel therapeutic targets. *Trends in Pharmacological Sciences* **27**, 416–425 (2006).
52. Wang, D., *et al.* Live-cell RNA imaging with metabolically incorporated fluorescent nucleosides. *J. Am. Chem. Soc.* **144**, 14647–14656 (2022).
53. Zhang, Y. & Kleiner, R. E. A Metabolic Engineering Approach to Incorporate Modified Pyrimidine Nucleosides into Cellular RNA. *J. Am. Chem. Soc.* **141**, 3347–3351 (2019).
54. Singha, M. K. *et al.* An Optimized Enzyme-Nucleobase Pair Enables In Vivo RNA Metabolic Labeling with Improved Cell-Specificity. *Biochemistry* **61**, 2638–2642 (2022).
55. Kofoed, R. H., *et al.* Investigation of RNA Synthesis Using 5-Bromouridine Labelling and Immunoprecipitation. *J Vis Exp* **135**, 57056 (2018).
56. George, J. T. & Srivatsan, S. G. Bioorthogonal chemistry-based RNA labeling technologies: evolution and current state. *Chem Commun (Camb)* **56**, 12307–12318 (2020).
57. Perrone, D., *et al.* Modified Nucleosides, Nucleotides and Nucleic Acids via Click Azide-Alkyne Cycloaddition for Pharmacological Applications. *Molecules* **26**, 3100 (2021).
58. Moreno, S. *et al.* Synthesis of 4-thiouridines with prodrug functionalization for RNA metabolic labeling. *RSC Chem. Biol.* **3**, 447–455 (2022).
59. Jao, C. Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc Natl Acad Sci U S A* **105**, 15779–15784 (2008).

60. Wang, D., Zhang, Y. & Kleiner, R. E. Cell- and Polymerase-Selective Metabolic Labeling of Cellular RNA with 2′-Azidocytidine. *J. Am. Chem. Soc.* **142**, 14417–14421 (2020).
61. Reinhardt, C. G. & Krugh, T. R. A comparative study of ethidium bromide complexes with dinucleotides and DNA: direct evidence for intercalation and nucleic acid sequence preferences. *Biochemistry* **17**, 4845–4854 (1978).
62. Wilhelmsson, L. M. Fluorescent nucleic acid base analogues. *Quarterly Reviews of Biophysics* **43**, 159–183 (2010).
63. Michel, B. Y., *et al.* Probing of Nucleic Acid Structures, Dynamics, and Interactions With Environment-Sensitive Fluorescent Labels. *Front. Chem.* **8**, (2020).
64. Ren, R.X.F., *et al.* Naphthalene, phenanthrene and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA. *J. Am. Chem. Soc.* **118**, 7671–7678 (1996).
65. Wilhelmsson, L. M. & Yitzhak, T. *Fluorescent Analogs of Biomolecular Building Blocks: Design and Applications*. (Wiley, Hoboken, New Jersey, 2016).
66. Xu, W., Chan, K. M. & Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nature Chem* **9**, 1043–1055 (2017).
67. Tanpure, A. A., *et al.* Fluorescent Nucleoside Analogs: Probes for Investigating Nucleic Acid Structure and Function. *Israel Journal of Chemistry* **53**, 366–378 (2013).
68. Ward, D. C., *et al.* Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* **244**, 1228–1237 (1969).
69. Allan, B. W. & Reich, N. O. Targeted Base Stacking Disruption by the *Eco* RI DNA Methyltransferase. *Biochemistry* **35**, 14757–14762 (1996).
70. Lin, K.-Y. & Matteucci, M. D. A Cytosine Analogue Capable of Clamp-Like Binding to a Guanine in Helical Nucleic Acids. *J. Am. Chem. Soc.* **120**, 8531–8532 (1998).
71. Wilhelmsson, L. M., *et al.* Highly Fluorescent DNA Base Analogue that Forms Watson−Crick Base Pairs with Guanine. *J. Am. Chem. Soc.* **123**, 2434–2435 (2001).
72. Sandin, P. Fluorescent properties of DNA base analogue tC upon incorporation into DNA -- negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Research* **33**, 5019–5025 (2005).
73. Sandin, P. *et al.* Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Research* **36**, 157–167 (2008).
74. Börjesson, K. *et al.* Nucleic Acid Base Analog FRET-Pair Facilitating Detailed Structural Measurements in Nucleic Acid Containing Systems. *Journal of the American Chemical Society* **131**, 4288–4293 (2009).
75. Füchtbauer, A. F. *et al.* Interbase FRET in RNA: from A to Z. *Nucleic Acids Research* **47**, 9990–9997 (2019).
76. Dumat, B., *et al.* Studying Z-DNA and B- to Z-DNA transitions using a cytosine analogue FRET-pair. *Nucleic Acids Research* **44**, e101–e101 (2016).
77. Kuznetsova, A. A. *et al.* New Environment-Sensitive Multichannel DNA Fluorescent Label for Investigation of the Protein-DNA Interactions. *PLoS ONE* **9**, e100007 (2014).
78. Bood, M. *et al.* Interbase-FRET binding assay for pre-microRNAs. *Sci Rep* **11**, 9396 (2021).

79. Baral, S. *et al.* Evidence for intrinsic DNA dynamics and deformability in damage sensing by the Rad4/XPC nucleotide excision repair complex. *Nucleic Acids Research* **53**, gkae1290 (2025).
80. Turner, M. B. & Purse, B. W. Fluorescent Tricyclic Cytidine Analogues as Substrates for Retroviral Reverse Transcriptases. *ChemPlusChem* **85**, 855–865 (2020).
81. Dierckx, A. *et al.* Quadracyclic Adenine: A Non-Perturbing Fluorescent Adenine Analogue. *Chemistry - A European Journal* **18**, 5987–5997 (2012).
82. Foller Larsen, A. *et al.* Development of bright fluorescent quadracyclic adenine analogues: TDDFT-calculation supported rational design. *Scientific Reports* **5**, (2015).
83. Wypijewska Del Nogal, A. *et al.* Getting DNA and RNA out of the dark with 2CNqA: a bright adenine analogue and interbase FRET donor. *Nucleic Acids Res* **48**, 7640–7652 (2020).
84. Bood, M. *et al.* Pentacyclic adenine: a versatile and exceptionally bright fluorescent DNA base analogue. *Chemical Science* **9**, 3494–3502 (2018).
85. Fisher, R. S. *et al.* Pulse-shaped two-photon excitation of a fluorescent base analogue approaches single-molecule sensitivity. *Phys. Chem. Chem. Phys.* **20**, 28487–28498 (2018).
86. Nilsson, J. R. *et al.* Fluorescent base analogues in gapmers enable stealth labeling of antisense oligonucleotide therapeutics. *Sci Rep* **11**, 11365 (2021).
87. Baladi, T. *et al.* Stealth Fluorescence Labeling for Live Microscopy Imaging of mRNA Delivery. *J. Am. Chem. Soc.* **143**, 5413–5424 (2021).
88. Le, H.-N. *et al.* Synthesis and photophysical characterization of a pH-sensitive quadracyclic uridine (qU) analogue. *Chem. Eur. J.**30,*** e202303539.
89. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Springer US, Boston, MA, 1999).
90. Ito, A. & Ito, T. ABSORPTION SPECTRA OF DEOXYRIBOSE, RIBOSEPHOSPHATE, ATP AND DNA BY DIRECT TRANSMISSION MEASUREMENTS IN THE VACUUM-UV (150—190 nm) AND FAR-UV (190—260 nm) REGIONS USING SYNCHROTRON RADIATION AS A LIGHT SOURCE. *Photochemistry and Photobiology* **44**, 355–358 (1986).
91. Van Holde, *et al. Principles of Physical Biochemistry*. (Prentice Hall, Upper Saddle River, NJ, 1998).
92. Wilson, K. *Principles and Techniques of Practical Biochemistry*. (Cambridge Univ. Press, Cambridge, 1997).
93. Fasman, G. D. *Circular Dichroism and the Conformational Analysis of Biomolecules*. (Plenum Press, New York, 1996).
94. Nakanishi, K., *et al. Circular Dichroism: Principles and Applications*. (J. Wiley & sons, New York Chichester Weinheim, 2000).
95. Göppert-Mayer, M. Über Elementarakte mit zwei Quantensprüngen. *Annalen der Physik* **401**, 273–294 (1931).
96. Denk, W., *et al.* Multi-Photon Molecular Excitation in Laser-Scanning Microscopy. in *Handbook Of Biological Confocal Microscopy* (ed. Pawley, J. B.) 535–549 (Springer US, Boston, MA, 2006).

97. Eaton, D. F. Reference materials for fluorescence measurement. *Pure and Applied Chemistry* **60**, 1107–1114 (1988).
98. Becker, W. *The Bh TCSPC Handbook*. (Becker&Hickl GmbH, Berlin, Germany, 2017).
99. *Atti Della Fondazione Giorgio Ronchi E Contributi Dell'Istituto Nazionale Di Ottica*. (La Fondazione, 1975).
100. Elliott, A. D. Confocal Microscopy: Principles and Modern Practices. *Curr Protoc Cytom* **92**, e68 (2020).
101. Hibbs, A. R. *Confocal Microscopy for Biologists*. (Springer US, Boston, MA, 2004).
102. Jonkman, J. & Brown, C. M. Any Way You Slice It—A Comparison of Confocal Microscopy Techniques. *J Biomol Tech* **26**, 54–65 (2015).
103. König, K. Multiphoton microscopy in life sciences. *Journal of Microscopy* **200**, 83–104 (2000).
104. Non-Descanned FLIM Detection in Multiphoton Microscopes. *Becker & Hickl GmbH* https://www.becker-hickl.com/literature/application-notes/non-descanned-flim-detection-in-multiphoton-microscopes/.
105. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat Methods* **2**, 932–940 (2005).
106. Becker, W. Fluorescence lifetime imaging – techniques and applications. *Journal of Microscopy* **247**, 119–136 (2012).
107. Digman, M. A., Caiolfa, V. R., Zamai, M. & Gratton, E. The Phasor Approach to Fluorescence Lifetime Imaging Analysis. *Biophys J* **94**, L14–L16 (2008).
108. Mandy, F. F., *et al.* Principles of flow cytometry. *Transfusion Science* **16**, 303–314 (1995).
109. Bakke, A. C. The Principles of Flow Cytometry. *Laboratory Medicine* **32**, 207–211 (2001).
110. Goetz, C., *et al. Flow Cytometry Basics for the Non-Expert*. (Springer International Publishing, Cham, 2018).
111. Zamora, J. L. R. & Aguilar, H. C. Flow virometry as a tool to study viruses. *Methods* **134–135**, 87–97 (2018).
112. Bartosik, K., *et al.* Synthesis of Nucleobase-Modified RNA Oligonucleotides by Post-Synthetic Approach. *Molecules* **25**, 3344 (2020).
113. Ebright, R. H. RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II. *Journal of Molecular Biology* **304**, 687–698 (2000).
114. Nielsen, H. *RNA: Methods and Protocols*. (Humana Press, Totowa, NJ, 2011).
115. Banerjee, A. K. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiological Reviews* **44**, 175–205 (1980).
116. Perenkov, A. D., *et al.* In Vitro Transcribed RNA-Based Platform Vaccines: Past, Present, and Future. *Vaccines (Basel)* **11**, 1600 (2023).
117. Hinnebusch, A. G. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu. Rev. Biochem.* **83**, 779–812 (2014).
118. Farrell, R. E. *RNA Methodologies: A Laboratory Guide for Isolation and Characterization*. (Elsevier, Academic, Amsterdam, 2010).

119. Gjerde, D. T. *RNA Purification and Analysis: Sample Preparation, Extraction, Chromatography*. (John Wiley & Sons, Incorporated, Hoboken, 2009).
120. Bloomfield, V. A., *et al. Nucleic Acids: Structures, Properties, and Functions*. (University Science Books, Sausalito, Calif, 2000).
121. Zhao, C. Cell culture: in vitro model system and a promising path to in vivo applications. *Journal of Histotechnology* **46**, 1–4 (2023).
122. Skloot, R. *The Immortal Life of Henrietta Lacks*. (Crown Publishers, New York, 2010).
123. Nakabayashi, H. *et al.* PHENOTYPICAL STABILITY OF A HUMAN HEPATOMA CELL LINE, HuH-7, IN LONG-TERM CULTURE WITH CHEMICALLY DEFINED MEDIUM. *The Japanese Cancer Association* (1984).
124. Nakabayashi, H., *et al.* Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res.* **42**, 3858–3863 (1982).
125. Fields, R. D. & Lancaster, M. V. Dual-attribute continuous monitoring of cell proliferation/cytotoxicity. *Am Biotechnol Lab* **11**, 48–50 (1993).
126. Cory, A. H., *et al.* Use of an Aqueous Soluble Tetrazolium/Formazan Assay for Cell Growth Assays in Culture. *Cancer Communications* **3**, 207–212 (1991).
127. Schätzing, F. *Nachrichten aus einem unbekannten Universum: eine Zeitreise durch die Meere*. (Kiepenheuer & Witsch, Köln, 2006).
128. Riazance, J. H. *et al.* Evidence for Z-form RNA by vacuum UV circular dichroism. *Nucleic Acids Res* **13**, 4983–4989 (1985).
129. Wranne, M. S. *et al.* Toward Complete Sequence Flexibility of Nucleic Acid Base Analogue FRET. *Journal of the American Chemical Society* **139**, 9271–9280 (2017).
130. Füchtbauer, A. F. *et al.* Lighting Up DNA with the Environment-Sensitive Bright Adenine Analogue qAN4. *ChemPlusChem* **85**, 319–326 (2020).
131. Engman, K. C. *et al.* DNA adopts normal B-form upon incorporation of highly fluorescent DNA base analogue tC: NMR structure and UV-Vis spectroscopy characterization. *Nucleic Acids Research* **32**, 5087–5095 (2004).
132. Nilsson, J. R. *et al.* Multiphoton characterization and live cell imaging using fluorescent adenine analogue 2CNqA. *Phys. Chem. Chem. Phys.* **25**, 20218–20224 (2023).
133. Bood, M., *et al.* Fluorescent nucleobase analogues for base–base FRET in nucleic acids: synthesis, photophysics and applications. *Beilstein Journal of Organic Chemistry* **14**, 114–129 (2018).
134. Bruice, P. Y. *Organic Chemistry*. (Fourth edition. Upper Saddle River, NJ : Pearson/Prentice Hall, [2004] ©2004, 2004).
135. Ranawat, H., *et al.* Recent trends in two-photon auto-fluorescence lifetime imaging (2P-FLIM) and its biomedical applications. *Biomed Eng Lett* **9**, 293–310 (2019).
136. Schäkel, L. *et al.* Nucleotide Analog ARL67156 as a Lead Structure for the Development of CD39 and Dual CD39/CD73 Ectonucleotidase Inhibitors. *Front Pharmacol* **11**, 1294 (2020).
137. Caiazzo, E., *et al.* The Ecto-5'-Nucleotidase/CD73 Inhibitor, α,β-Methylene Adenosine 5'-Diphosphate, Exacerbates Carrageenan-Induced Pleurisy in Rat. *Frontiers in Pharmacology* **10**, 775 (2019).

138. Wood, E. *et al.* Cell-type specificity of ectonucleotidase expression and upregulation by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Archives of Biochemistry and Biophysics* **407**, 49–62 (2002).
139. Lévesque, S. A., *et al.* Specificity of the ecto-ATPase inhibitor ARL 67156 on human and mouse ectonucleotidases. *British Journal of Pharmacology* **152**, 141–150 (2007).
140. Zimmermann, H. Extracellular metabolism of ATP and other nucleotides. *Naunyn-Schmied Arch Pharmacol* **362**, 299–309 (2000).
141. Szaflarski, W. *et al.* Early rRNA processing is a stress-dependent regulatory event whose inhibition maintains nucleolar integrity. *Nucleic Acids Res* **50**, 1033–1051 (2021).
142. Beldi, G. *et al.* The role of purinergic signaling in the liver and in transplantation: effects of extracellular nucleotides on hepatic graft vascular injury, rejection and metabolism. *Front Biosci* **13**, 2588–2603 (2008).
143. Yarbro, J. W. Mechanism of action of hydroxyurea. *Semin Oncol* **19**, 1–10 (1992).
144. Singh, A. & Xu, Y.-J. The Cell Killing Mechanisms of Hydroxyurea. *Genes (Basel)* **7**, 99 (2016).
145. Ma, Q. *et al.* Optimization of the 5′ untranslated region of mRNA vaccines. *Sci Rep* **14**, 19845 (2024).
146. Vo, J. M. *et al.* Synthesis of modified nucleotide polymers by the poly(U) polymerase Cid1: application to direct RNA sequencing on nanopores. *RNA* **27**, 1497–1511 (2021).
147. Meng, L. *et al.* Metabolic RNA labeling for probing RNA dynamics in bacteria. *Nucleic Acids Research* **48**, 12566–12576 (2020).
148. Marquardt, S. & Manavella, P. A. A ribose world: current status and future challenges of plant RNA biology. *J Exp Bot* **74**, 2203–2207 (2023).
149. Voloudakis, A. E., *et al.* RNA-Based Vaccination of Plants for Control of Viruses. *Annu Rev Virol* **9**, 521–548 (2022).
150. Preus, S., *et al.* Characterization of Nucleobase Analogue FRET Acceptor tCnitro. *J. Phys. Chem. B* **114**, 1050–1056 (2010).
151. Wranne, M. *DNA and RNA Base Analogue Fret: From Fluorophore Design to Biochemical Applications*. (Chalmers University of Technology, Göteborg, 2018).

QR Codes

Links to the published articles, including supplementary data.

Paper I

Multiphoton characterization and live cell imaging using fluorescent adenine analogue 2CNqA (DOI: 10.1093/nar/gkae722)

Paper II

Metabolic RNA labeling in non-engineered cells following spontaneous uptake of fluorescent nucleoside phosphate analogues (DOI: 10.1038/s41598-021-90629-1)