

# Fluorogenic Aptamers and Fluorescent Nucleoside Analogs as Probes for RNA Structure and Function

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Julius-Maximilians-  
Universität Würzburg

vorgelegt von

Christian Steinmetzger

aus Bad Gandersheim

Würzburg 2020





Diese Arbeit wurde im Zeitraum von 01.01.2016–30.06.2017 am Institut für Organische und Biomolekulare Chemie der Georg-August-Universität Göttingen sowie nachfolgend vom 01.07.2017–14.04.2020 am Institut für Organische Chemie der Julius-Maximilians-Universität Würzburg unter der Anleitung von Prof. Dr. Claudia Höbartner angefertigt.

Eingereicht bei der Fakultät für Chemie und Pharmazie am:

14.04.2020

Gutachter der schriftlichen Arbeit:

1. Gutachter: Prof. Dr. Claudia Höbartner

2. Gutachter: Prof. Dr. Christoph Lambert

Prüfer des öffentlichen Promotionskolloquiums:

1. Prüfer: Prof. Dr. Claudia Höbartner

2. Prüfer: Prof. Dr. Christoph Lambert

3. Prüfer: Prof. Dr. Matthias Lehmann

4. Prüfer: Prof. Dr. Roland Mitric

5. Prüfer Prof. Dr. Jürgen Seibel

Datum des öffentlichen Promotionskolloquiums:

03.07.2020

Doktorurkunde ausgehändigt am:

\_\_\_\_\_



# Acknowledgements

First and foremost, I want to express my deep and sincere gratitude to my thesis supervisor and mentor, Prof. Dr. Claudia Höbartner. I have benefited immensely from your continuous guidance and support from the moment I started working in your group. Thank you for giving me the opportunity to work in such an interesting field of research and for always being available to discuss new ideas, problems and results.

I would also like to thank Prof. Dr. Christoph Lambert for agreeing to review my thesis and all members of the examination board for accepting the invitation to my defense.

Thank you to all group members, present and past, for making my time spent in the lab and the office over the past few years enjoyable and interesting. I would like to mention our former technician Jan, who helped me find my way around the lab when I first joined the group for a short practical course, as well as Laura, Steffi, Ann-Kathrin and Sebastian for their tremendously helpful work in the bio and chemistry labs. Thank you to the many practical course students who have been part of the Chili project, namely Philipp, Julia, Florian, Thorge and Claudia, as well as to my bachelor students Tim and Carmen—this has been a team effort and working with you has been great fun! Thank you, Irene, for always sharing and discussing the latest NMR news about the Chili aptamer with me, and thank you, Caro, for always sharing your good mood!

I am grateful to Julia, Irene and Caro for taking the time to proofread my thesis. I appreciate your feedback and your help in improving this thesis!

Many thanks to the analytical departments at the Universities of Göttingen and Würzburg for providing NMR and MS measurements and getting good results even when I once again had only the tiniest amount of hardly soluble substance to spare.

Finally, I want to thank my friends and family, who have always been there for me these past few years. Thank you for your unwavering support and patience!



Parts of this thesis have been published:

- Steinmetzger, C.; Palanisamy, N.; Gore, K. R.; Höbartner, C. A Multicolor Large Stokes Shift Fluorogen-Activating RNA Aptamer with Cationic Chromophores. *Chem. Eur. J.* **2019**, *25*, 1931–1935.
- Steinmetzger, C.; Bessi, I.; Lenz, A.-K.; Höbartner, C. Structure-fluorescence activation relationships of a large Stokes shift fluorogenic RNA aptamer. *Nucleic Acids Res.* **2019**, *47*, 11538–11550.
- Steinmetzger, C.; Bäuerlein, C.; Höbartner, C. Supramolecular fluorescence resonance energy transfer in nucleobase-modified fluorogenic RNA aptamers. *Angew. Chem. Int. Ed.* **2020**, *59*, 6760–6764.

*Science is fun. Science is curiosity. We all have natural curiosity. Science is a process of investigating. It's posing questions and coming up with a method. It's delving in.*

*Sally Ride*





## TABLE OF CONTENTS

ABSTRACT.....	V
ZUSAMMENFASSUNG.....	VII
<b>1 INTRODUCTION .....</b>	<b>1</b>
1.1 STRUCTURE AND FUNCTION OF RNA.....	1
1.1.1 Beyond the central dogma.....	1
1.1.2 Fluorescence in RNA .....	3
1.1.3 Small molecule recognition by RNA aptamers.....	8
1.2 FLUOROGENIC RNA APTAMERS .....	13
1.2.1 The serendipitous discovery of RNA-based fluorescence activation.....	13
1.2.2 In vitro selection against protein-like fluorogens .....	15
1.2.3 Recent developments .....	21
1.3 RESEARCH OBJECTIVES AND THESIS OUTLINE.....	27
<b>2 RESULTS AND DISCUSSION.....</b>	<b>29</b>
2.1 DESIGN OF THE CHILI APTAMER.....	29
2.1.1 A DMHBI-binding aptamer.....	29
2.1.2 Engineering Chili from 13-2min .....	29
2.2 FUNCTIONAL HALLMARKS OF CHILI .....	32
2.2.1 Spectral characteristics .....	32
2.2.2 Folding state.....	34
2.3 OPTIMIZATION OF THE CHILI SYSTEM.....	38
2.3.1 Screening of HBI ligands.....	39
2.3.2 Systematic mutagenesis of Chili.....	55
2.4 DETAILED CHARACTERIZATION OF SELECTED CHILI–HBI COMPLEXES .....	59
2.4.1 Binding specificity and metal ion dependence .....	59
2.4.2 Relative fluorescence quantum yields .....	62
2.4.3 Binding analysis by fluorescence spectroscopy .....	64
2.4.4 Calorimetric analysis of binding thermodynamics.....	67
2.4.5 Fluorescence activation kinetics .....	70
2.4.6 Ligand-induced stabilization of the binding site .....	71
2.5 AN APTAMER-TAG FOR GEL-BASED FLUORESCENCE QUANTIFICATION .....	74
2.5.1 In-gel visualization of the Chili aptamer .....	74
2.5.2 A two-color FRET reporter for monitoring DNAzyme activity.....	75
2.6 DEVELOPMENT OF A SUPRAMOLECULAR FRET PLATFORM.....	79
2.6.1 FRET studies of RNA structure and function.....	79
2.6.2 Choice of the FRET donor.....	80
2.6.3 Synthesis of 4-cyanoindole phosphoramidite building block .....	83

2.7	CHARACTERIZATION OF 4-CYANOINDOLE RIBONUCLEOSIDE .....	92
2.7.1	Impact of 4-cyanoindole incorporation on oligonucleotide structure.....	92
2.7.2	Spectroscopic properties of 4-cyanoindole .....	95
2.8	FÖRSTER RESONANCE ENERGY TRANSFER WITHIN THE CHILI APTAMER .....	100
2.8.1	Synthesis of 4-cyanoindole-modified Chili aptamer constructs.....	100
2.8.2	Observation of intrasupramolecular FRET .....	102
3	CONCLUSION AND OUTLOOK .....	111
4	EXPERIMENTAL PART.....	113
4.1	MATERIALS AND METHODS.....	113
4.1.1	General remarks.....	113
4.1.2	In vitro transcription .....	113
4.1.3	Solid-phase synthesis .....	114
4.1.4	Enzymatic ligation .....	117
4.1.5	Polyacrylamide gel electrophoresis (PAGE) .....	119
4.1.6	Oligonucleotide analysis .....	120
4.1.7	UV/Vis spectroscopy .....	121
4.1.8	pK <sub>a</sub> determination .....	121
4.1.9	Extinction coefficient determination .....	123
4.1.10	Fluorescence spectroscopy.....	124
4.1.11	Fluorescence lifetime analysis .....	127
4.1.12	Quantum yield determination .....	130
4.1.13	Screening assay.....	134
4.1.14	Equilibrium binding titrations .....	136
4.1.15	Ligand competition assay .....	139
4.1.16	Isothermal titration calorimetry .....	140
4.1.17	Binding kinetics .....	143
4.1.18	Thermal denaturation curves of double-stranded oligonucleotides .....	145
4.1.19	Thermal stability of the Chili aptamer .....	148
4.1.20	Stern-Volmer titrations .....	150
4.1.21	Förster resonance energy transfer .....	154
4.2	CHEMICAL SYNTHESIS.....	159
4.2.1	General considerations .....	159
4.2.2	Preparation of HBI derivatives .....	159
4.2.3	Preparation of 4-Cyanoindole phosphoramidite building blocks .....	192
4.2.4	Additional compounds and reagents .....	203
4.3	OLIGONUCLEOTIDE SYNTHESIS .....	206
4.3.1	In vitro transcription .....	206
4.3.2	Solid-phase synthesis .....	209
4.3.3	Enzymatic ligation .....	209
4.4	COMPUTATIONAL METHODS .....	210
4.4.1	Geometry optimization .....	210
4.4.2	Electrostatic potential maps .....	210
4.4.3	Sample input file .....	210

5	APPENDIX.....	210
5.1	SUPPLEMENTARY FIGURES.....	210
5.2	SUPPLEMENTARY TABLES.....	226
5.3	CARTESIAN COORDINATES .....	229
5.3.1	HBI derivatives .....	229
5.3.2	Methylated nucleobases.....	235
5.4	BUFFER SOLUTIONS AND REAGENTS.....	237
5.5	ABBREVIATIONS.....	239
	REFERENCES.....	243



## ABSTRACT

RNA plays a key role in numerous cellular processes beyond the central dogma of molecular biology. Observing and understanding this wealth of functions, discovering new ones and engineering them into purpose-built tools requires a sensitive means of observation. Over the past decade, fluorogenic aptamers have emerged to fill this niche. These short oligonucleotides are generated by *in vitro* selection to specifically interact with small organic fluorophores and can be utilized as genetically encoded tags for RNAs of interest.

The most versatile class of fluorogenic aptamers is based on derivatives of hydroxybenzylidene imidazolone (HBI), a conditional fluorophore mimicking the chromophore structure found in green and red fluorescent proteins. The respective aptamers are well-known by the “vegetable” nomenclature, including Spinach, Broccoli and Corn, and have found numerous applications for studying RNA function *in vitro* and in cells.

Their success, however, is somewhat overshadowed by individual shortcomings such as a propensity for misfolding, dependence on unphysiologically high concentrations of magnesium ions or, in the case of Corn, dimerization that might affect the function of the tagged RNA. Moreover, most fluorogenic aptamers exhibit limited ligand promiscuity by design, thereby restricting their potential for spectral tuning to a narrow window of wavelengths.

This thesis details the characterization of a new fluorogenic aptamer system nicknamed Chili. Chili is derived from an aptamer that was originally selected to bind 4-hydroxy-3,5-dimethoxyhydroxybenzylidene imidazolone (DMHBI), resulting in a green fluorescent complex. Unlike other aptamers of its kind, Chili engages in a proton transfer cycle with the bound ligand, resulting in a remarkably large Stokes shift of more than 130 nm.

By means of an empirical ligand optimization approach, several new DMHBI derivatives were found that bind to Chili with high affinity, furnishing complexes up to 7.5 times brighter compared to the parent ligand. In addition, Chili binds to  $\pi$ -extended DMHBI derivatives that confer fluorescence in the yellow–red region of the visible spectrum. The highest affinity and degree of fluorescence turn-on for both green and red fluorogenic ligands were achieved by the incorporation of a unique, positively charged substituent into the HBI scaffold.

Supplemented by NMR spectroscopy, kinetic and thermodynamic studies showed that the binding site of Chili is loosely preorganized in the absence of ligand and likely forms a G-quadruplex upon ligand binding.

To showcase future applications, Chili was incorporated into a FRET sensor for monitoring the cleavage of an RNA substrate by a 10-23 DNAzyme.

Besides aptamers as macromolecular fluorescent complexes, fluorescent nucleobase analogs are powerful small isomeric components of RNA suitable for studying structure and folding. Here, the highly emissive nucleobase analog 4-cyanoindole (4CI) was developed into a ribonucleoside (r4CI) for this purpose. A new phosphoramidite building block was synthesized to enable site-specific incorporation of 4CI into RNA.

Thermal denaturation experiments confirmed that 4CI behaves as a universal nucleobase, i.e. without bias towards any particular hybridization partner. Photophysical characterization established r4CI as a generally useful fluorescent ribonucleoside analog. In this work, it was employed to gain further insight into the structure of the Chili aptamer. Using several 4CI-modified

Chili–HBI complexes, a novel base–ligand FRET assay was established to obtain a set of combined distance and orientation restraints for the tertiary structure of the aptamer.

In addition to their utility for interrogating structure and binding, supramolecular FRET pairs comprising a fluorescent nucleobase analog donor and an innately fluorogenic acceptor hold great promise for the construction of color-switchable RNA aptamer sensor devices.

## ZUSAMMENFASSUNG

Weit über das zentrale Dogma der Molekularbiologie hinaus ist RNA an einer Vielzahl zellulärer Prozesse beteiligt. Um diese Prozesse aufzuklären, sie umfassend zu verstehen und sich zunutze zu machen bedarf es geeigneter Detektionsmethoden für RNA. Innerhalb des letzten Jahrzehnts wurden fluorogene Aptamere als ideales Werkzeug für diesen Zweck erkannt. Dabei handelt es sich um vergleichsweise kurze Oligonukleotide, die mittels *in vitro*-Selektion zur spezifischen Bindung bestimmter organischer Fluorophore erzeugt werden. Analog zu fluoreszierenden Proteinen können sie zur Fluoreszenzmarkierung von RNA eingesetzt werden.

Die wichtigste Klasse fluorogener Aptamere bindet und aktiviert Derivate des latenten Fluorophors 4-Hydroxybenzylidenimidazol (HBI), welcher ursprünglich im Kern fluoreszierender Proteine autokatalytisch aus einem Tripeptid-Fragment entsteht und deren spektrale Eigenschaften bestimmt. Vertreter dieser Klasse, namentlich Spinach, Broccoli und Corn, haben sich als alltägliches Werkzeug zur Fluoreszenzmarkierung von RNA etabliert.

Diesem Erfolg gegenüber stehen Unzulänglichkeiten, die das Potential einzelner Aptamere begrenzen. Beispielsweise kann es zur Ausbildung inaktiver Faltungszustände der RNA kommen oder die Fluoreszenzaktivierung erfordert eine hohe Magnesiumkonzentration, welche in Zellen nicht frei verfügbar ist. Im Fall des Corn-Aptamers bildet sich ein Homodimer, was unter Umständen die zu untersuchende RNA beeinträchtigen kann. Darüber hinaus ist, aufgrund der spezifischen Fluorophorbindung, jeweils nur geringes Potenzial zur gezielten Beeinflussung spektraler Eigenschaften vorhanden.

Kern dieser Arbeit ist die umfassende Charakterisierung des neuen Chili-Systems. Chili ist die optimierte Version eines Aptamers, welches einen grün fluoreszierenden Komplex mit 4-Hydroxy-3,5-dimethoxybenzylidenimidazol (DMHBI) ausbildet. Im Gegensatz zu anderen HBI-bindenden Aptameren vollzieht Chili einen Protonenaustausch mit seinem Liganden, woraus Fluoreszenz-emission mit einer ungewöhnlich hohen Stokes-Verschiebung von über 130 nm resultiert.

Die Struktur des ursprünglichen Liganden wurde im Hinblick auf höhere Affinität und stärkere Fluoreszenzemission optimiert, wobei ein bis zu 7.5-facher Gewinn an Helligkeit erzielt wurde. Als besonders vorteilhaft hat sich dafür die Einführung eines positiv geladenen Substituenten herausgestellt, der in dieser Form ein Alleinstellungsmerkmal von Chili ist. Auch stark modifizierte DMHBI-Derivate, die ein größeres konjugiertes System besitzen, werden von Chili gebunden und fluoreszieren daraufhin im gelben bis roten Bereich des sichtbaren Spektrums.

Studien zur Ligandenbindungskinetik und thermischen Denaturierung des Aptamers legen nahe, dass die zunächst strukturarme Bindungstasche durch die Aufnahme des Liganden einen G-Quadruplex ausbildet, was ebenfalls durch NMR-spektroskopische Daten bestätigt wird.

Als Beispiel für mögliche Anwendungen wurde das Chili-Aptamer eingesetzt, um die Spaltung eines RNA-Substrats durch ein 10-23 DNA-Enzym zu beobachten, wobei FRET zwischen dem Aptamer und einem Fluoreszenzmarker am Substrat als Reporter ausgenutzt wurde.

Neben makromolekularen Aptamer-Komplexen können fluoreszierende Nukleobasenanaloga als isomorphe Einheiten in RNA integriert werden, um deren Faltungszustand zu untersuchen. In dieser Arbeit wurde das fluoreszierende Nukleobasenanalogon 4-Cyanodinol (4CI) in das entsprechende Ribonukleosid (r4CI) umgewandelt und daraus ein neuer Phosphoramiditbaustein zum Einbau des fluoreszierenden von 4CI in RNA synthetisiert.

Anhand thermischer Denaturierungsexperimente wurde gezeigt, dass es sich bei 4Cl um eine universelle Base handelt, die ungeachtet des Hybridisierungskontexts toleriert wird. Die photophysikalische Charakterisierung von r4Cl zeigte, dass das fluoreszierendes Ribonukleosid-Analogon seine nützlichen Eigenschaften nach dem Einbau in Oligonukleotide beibehält, sodass es zur Strukturanalyse des Chili-Aptamers verwendet werden konnte. Mithilfe 4Cl-modifizierter Chili-HBI-Komplexe wurden erstmals intramolekulare FRET-Paare dieser Art erzeugt und zur Bestimmung kombinierter Abstands- und Orientierungsparameter genutzt.

Über ihre Verwendung für Struktur- und Bindungsstudien hinaus stellen supramolekulare FRET-Paare aus fluoreszierenden Nucleobasen-Analoga als Donoren und intrinsisch fluorogenen Akzeptoren eine Möglichkeit dar, neue schaltbare Aptamer-basierte Sensoren zu entwickeln, welche auf die Erkennung ihrer Zielspezies mit einem Wechsel der Fluoreszenzemissionswellenlänge reagieren.

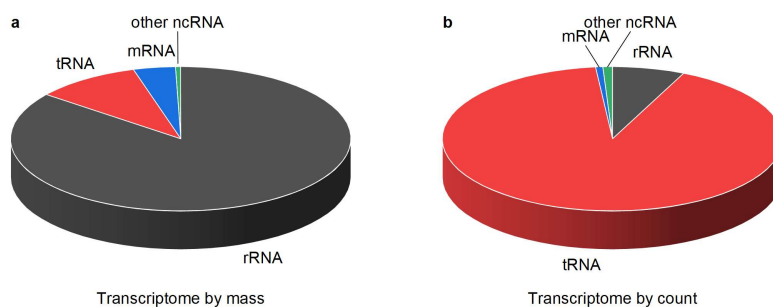


# 1 INTRODUCTION

## 1.1 STRUCTURE AND FUNCTION OF RNA

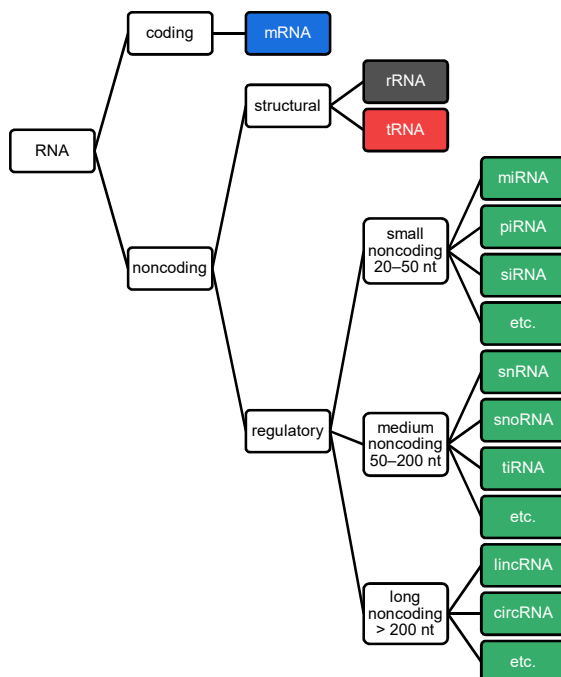
### 1.1.1 Beyond the central dogma

RNA has long been considered a mediator molecule between genetic information, stored in the DNA of an organism, and its functional realization in the form of proteins. Sequence information is transcribed into messenger RNA (mRNA), which is used by ribosomal RNA (rRNA) as an essential part of the ribosome to assemble the corresponding polypeptide chain with the help of transfer RNA (tRNA). Even though three types of RNA represent the majority of the transcriptome both by mass and number (Figure 1), this central dogma of molecular biology only captures a small fraction of the numerous roles RNA plays in cellular processes.



**Figure 1.** Estimated composition of the mammalian somatic transcriptome. Messenger, transfer and ribosomal RNA account for the majority both by mass (a) and number of individual molecules (b). Less abundant, noncoding transcripts (ncRNA) perform a multitude of functions, e.g. RNA splicing and gene regulation. This figure was adopted from a compilation of transcriptome composition data.<sup>1</sup>

For example, while more than 80% of the human genome is transcribed into RNA,<sup>2</sup> less than 2% of the genetic information content is eventually translated into proteins.<sup>3</sup> Other noncoding RNAs (ncRNA) in addition to tRNA and rRNA account for the majority of transcribed information (Figure 2). While numerous types of ncRNAs have been identified and classified according to their size and function, much remains to be learned and it is estimated that, even just in the human genome, upwards of 10000 long noncoding RNAs (lncRNA) have not yet been characterized.<sup>4,5</sup>



**Figure 2.** Simplified classification scheme for the diverse types of RNA. The vast majority of genetic information is transcribed into numerous types of noncoding RNAs. Although insight into their roles in biological processes is advancing steadily, many of their functions are yet to be discovered. RNAs between 20 and 200 nt in length are sometimes grouped together as small or short noncoding RNAs.<sup>6-9</sup>

Abbreviations: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), microRNA (miRNA), piwi-interacting RNA (piRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), tRNA-derived stress-induced RNA (tiRNA), long intergenic noncoding RNA (lincRNA), circular RNA (circRNA).

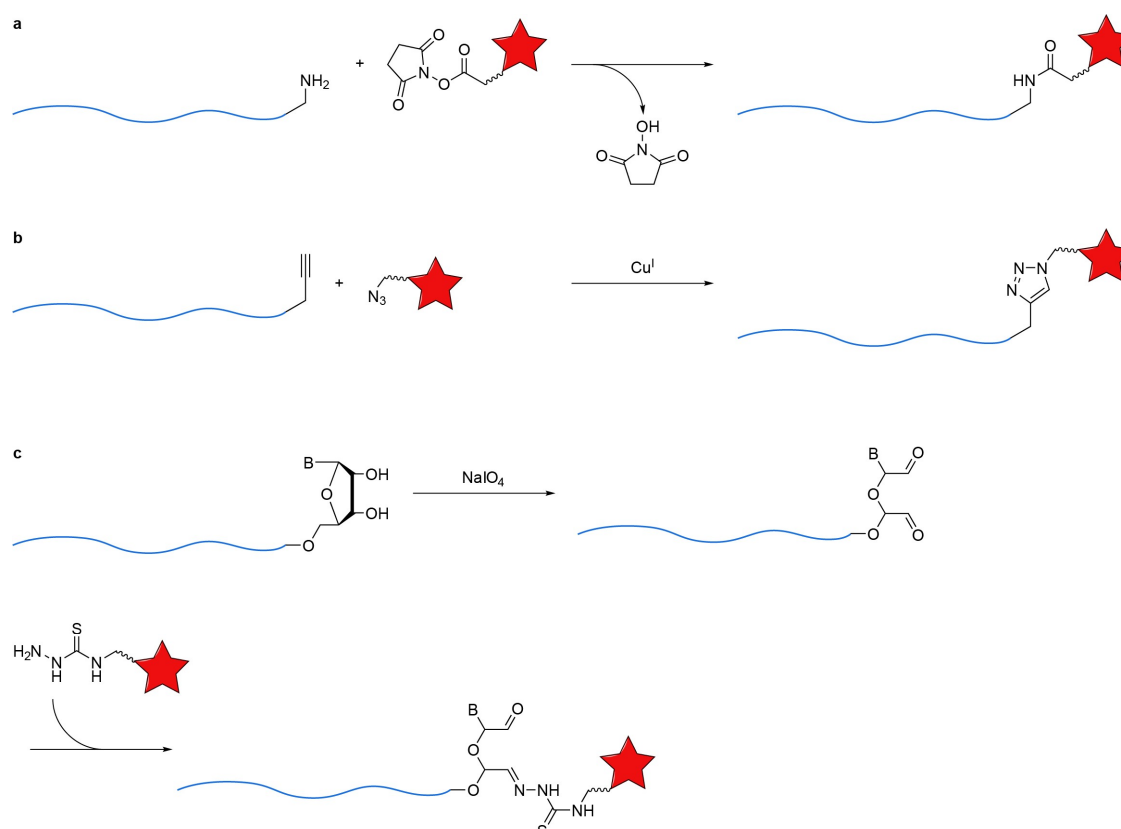
Nowadays, it is well established that RNA can catalytically process other nucleic acid strands in the form of ribozymes, e.g. by inducing sequence-specific phosphodiester bond cleavage or ligation, and can also act as a sensor for environmental parameters such as temperature<sup>10</sup> and the presence of both biomacromolecules<sup>11</sup> and small molecule metabolites.<sup>12,13</sup> Such sensors, called riboswitches, undergo structural changes upon stimulus recognition and can thereby function as regulatory elements for the control of gene expression.<sup>14</sup> Elucidating and understanding these numerous, multifaceted roles requires a means of observation that can shed light on the life cycle and localization of specific RNA sequences.

### 1.1.2 Fluorescence in RNA

One of the best-characterized and most prominent tools for imaging of biomacromolecules is the green fluorescent protein (GFP), which was discovered in the jellyfish species *Aequorea victoria*. This protein and a host of related, either naturally occurring or engineered ones can be fused to proteins of interest and expressed in live cells, thereby providing detailed visual information about their function in a native environment. The actual fluorophore in this case is a modified tripeptide within the larger protein, which forms autocatalytically during maturation.<sup>15</sup> In analogy, lighting up RNA with RNA would therefore be highly desirable for the reasons stated above. Unfortunately, unmodified nucleic acids do not exhibit useful fluorescence on their own and can therefore not be used as genetically encoded tags. This has led to the development of numerous approaches for covalent and noncovalent labeling of nucleic acids with emissive reporters, which will be introduced here briefly.

#### Covalent labeling of RNA

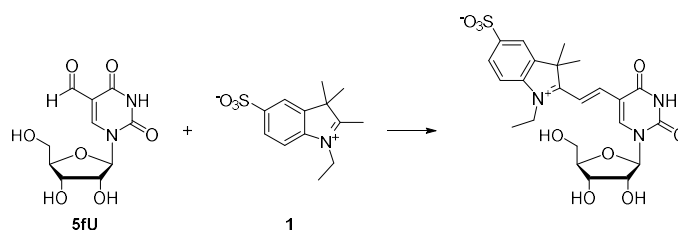
Covalent labeling is the method of choice for tagging small RNAs and, in its simplest form, requires a reactive handle for fluorophore attachment. Reactive groups are commonly incorporated at the 3'-end during solid-phase synthesis (e.g. amino, alkyne) or can be furnished post-synthetically in enzymatically synthesized RNA and are then treated with appropriately derivatized fluorophores. (Scheme 1).<sup>16,17</sup>



**Scheme 1.** Illustration of three commonly used strategies for covalent labeling of nucleic acids equipped with reactive handles. <sup>16,17</sup> a) An amino-modified oligonucleotide is reacted with an *N*-hydroxysuccinimide (NHS) ester derivative of the fluorophore. b) An alkyne-modified oligonucleotide is reacted with an azide derivative of the fluorophore via copper-catalyzed azide-alkyne cycloaddition (CuAAC). c) The 3'-terminal ribonucleotide is oxidized with  $\text{NaIO}_4$  to produce a dialdehyde, which is then reacted with a thiosemicarbazide derivative of the fluorophore.

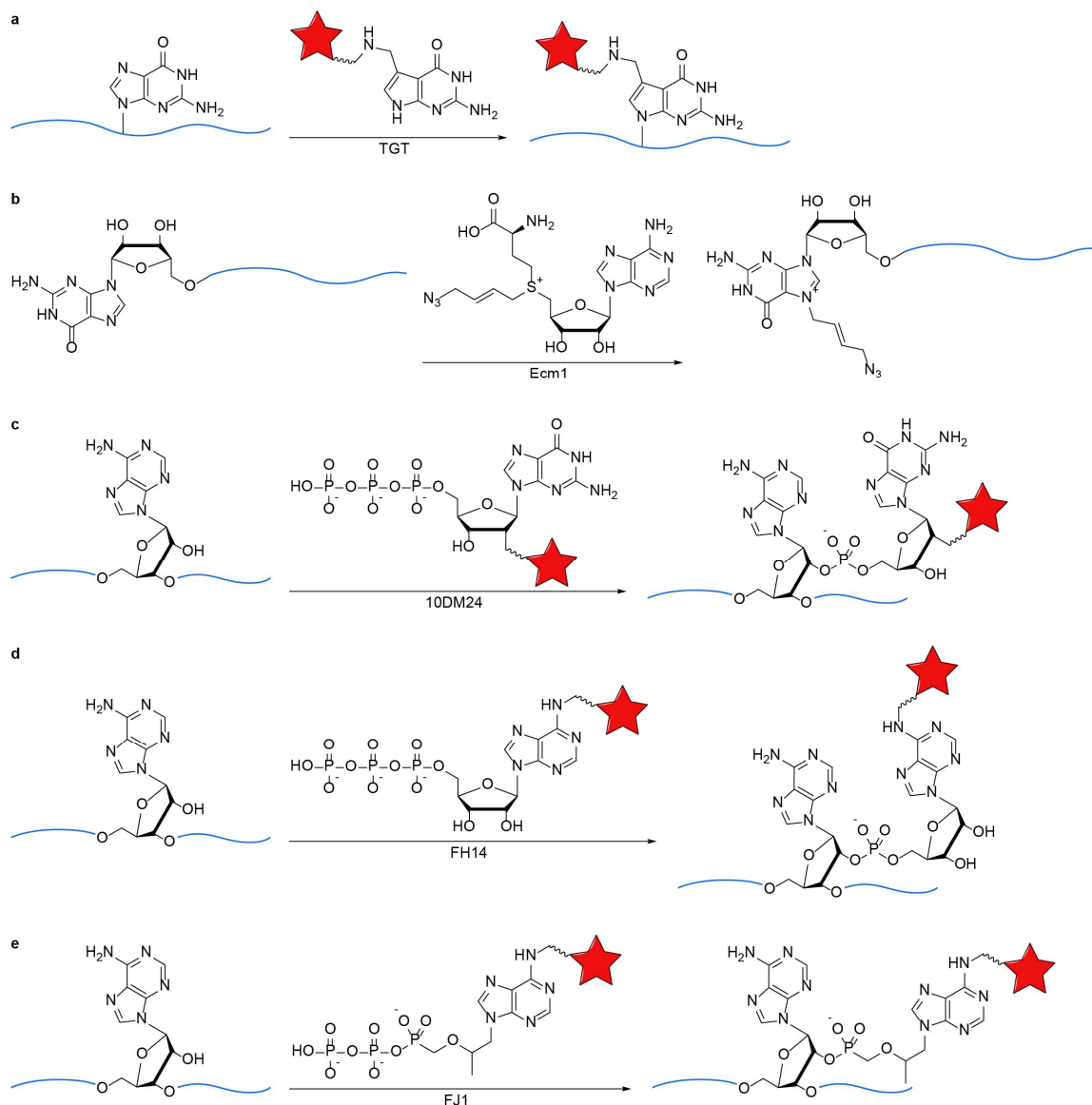
Certain naturally occurring RNA modifications can also be addressed directly in this manner. Samanta et al. demonstrated that *N*-ethyl-2,3,3-trimethylindolenium-5-sulfonate (**1**) can undergo

Aldol condensation with 5-formylpyrimidine sites in short DNA and RNA oligonucleotides to generate fluorescent hemicyanine dyes in situ (Scheme 2).<sup>18</sup>



**Scheme 2.** Example of a covalent labeling strategy that exploits naturally occurring reactive modifications in RNA. The indolenium reagent **1** undergoes Aldol condensation with 5-formyluridine (5fU) to form a fluorescent hemicyanine dye.<sup>18</sup>

Several RNA labeling methods have been developed that forego the need for reactive groups and can instead address native nucleic acids directly in a site-specific manner. For example, transglycosylase and methyltransferase enzymes can be repurposed as they are able to process not only their natural cofactors but also derivatives thereof. This has been exploited for transferring fluorescent dyes (Scheme 3a) or for the installation of functional groups in preparation for biorthogonal labeling reactions (Scheme 3b).<sup>19,20</sup> As an alternative to protein enzymes, deoxyribozymes and ribozymes have been evolved for this task using in vitro selection methods (see section 1.1.3). These catalyze the branching ligation between adenine residues in a defined sequence context and fluorescently labeled GTP or ATP derivatives to form 2',5'-phosphodiester bond (Scheme 3c, d).<sup>21,22</sup> These linkages are, however, susceptible to cleavage by debranching enzymes and the use of nucleoside triphosphates as cosubstrates might interfere with cellular function. Therefore, a new generation of ribozymes has been evolved, which utilize derivatives of the antiviral drug Tenofovir instead. These compounds exhibit less background reactivity and, because they are ligated to the RNA of interest by means of a phosphonate ester linkage, are more resistant to enzymatic degradation (Scheme 3e).<sup>23</sup>

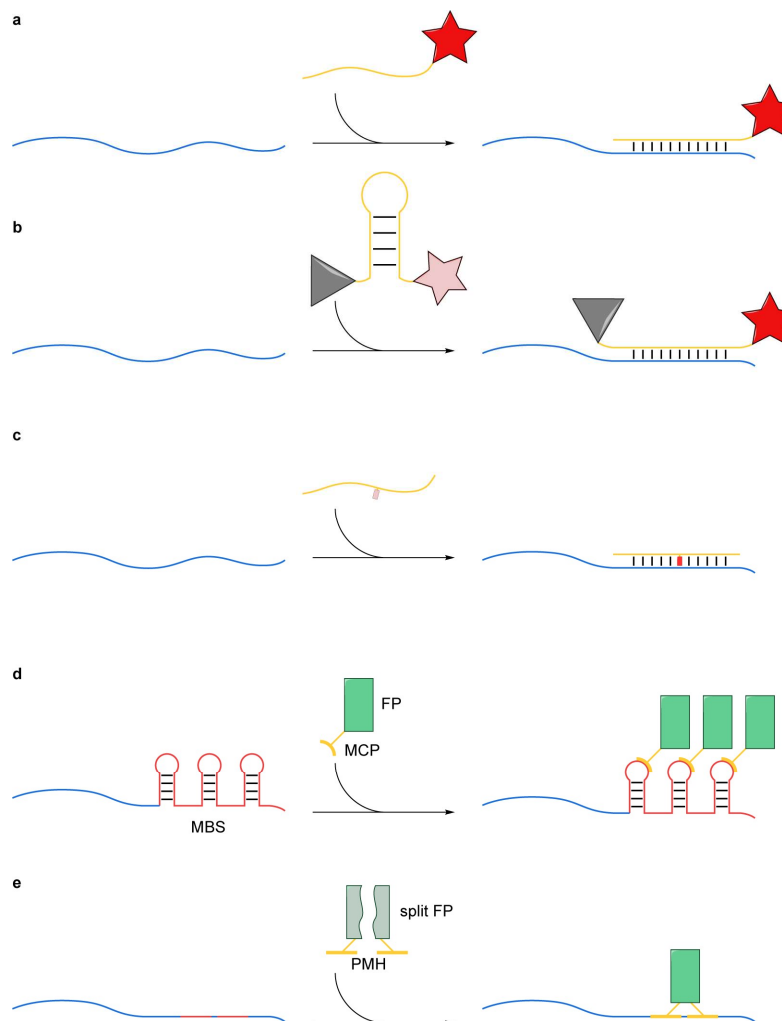


**Scheme 3.** Illustration of five commonly used strategies for covalent labeling of native nucleic acids.<sup>21–24</sup> a) Bacterial tRNA guanine transglycosylase (TGT) catalyzes the site-specific exchange of guanine with the modified nucleobase PreQ<sub>1</sub> (7-aminomethyl-7-deazaguanine). The enzyme also accepts PreQ<sub>1</sub> derivatives that are functionalized at the C7-aminomethyl group and can therefore be repurposed for transferring fluorophores to an RNA of interest. b) The RNA methyltransferase Ecm1 uses S-adenosyl methionine (SAM) to install a methyl group at the N7 position of guanine in the 5' cap region of eukaryotic mRNA. By using modified SAM analogs, reactive groups can be transferred and used in biorthogonal labeling reactions. c, d) The deoxyribozyme 10DM24 and the ribozyme FH14 catalyze the site-specific formation of branched 2',5'-phosphodiester bonds between adenine residues and various fluorophore-labeled GTP or ATP derivatives, respectively. e) The FJ1 ribozyme catalyzes a similar ligation reaction with derivatives of the antiviral drug Tenofovir as its cosubstrate. This results in the formation of a branched phosphonate ester linkage, which is more stable towards degradation compared to a phosphodiester.

### Noncovalent labeling of RNA

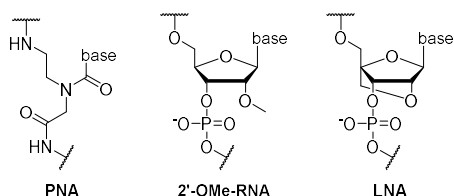
Fluorescently labeled RNA oligonucleotides themselves serve as probes for the noncovalent detection of larger RNAs of interest by fluorescence in situ hybridization (FISH). For this technique, the probe oligonucleotide is designed to be complementary to a sequence stretch in the RNA of interest that should ideally be free of secondary structure (Scheme 4a). FISH is well suited for RNA imaging in fixed cells, but drawbacks include the need to introduce an exogenous RNA into the cellular environment and high background fluorescence due to unbound probe molecules that can freely diffuse through the cell.<sup>25,26</sup> Molecular beacons are a variant of FISH probes that were developed to diminish or even eliminate background fluorescence contamination.<sup>27,28</sup> They are

labeled with a fluorophore and a suitable quenching agent at either end and form a stable hairpin in the absence of a complementary target RNA. Therefore, the fluorophore remains dark until hybridization to the RNA of interest occurs (Scheme 4b). An alternative approach for fluorescence activation utilizes forced intercalation (FIT) of a fluorophore into double-stranded RNA. An environment-sensitive fluorophore is incorporated as a nucleobase surrogate at an internal position of the probe. Upon hybridization with the target, the fluorophore intercalates between the newly formed base pairs, which results in fluorescence enhancement (Scheme 4c). While DNA oligonucleotides can be employed to construct FIT probes, a peptide nucleic acid (PNA) backbone is commonly used as it increases resistance towards degradation in live cells and confers favorable properties such as the ability to discriminate mismatches due to single base mutations in the target.<sup>29–31</sup>



**Scheme 4.** Illustration of five commonly used strategies for noncovalent labeling of nucleic acids.<sup>26</sup> a) For fluorescence in situ hybridization (FISH), a fluorescently labeled probe oligonucleotide binds the target due to sequence complementarity. b) To decrease background fluorescence, the probe can be replaced by a molecular beacon, a hairpin-forming oligonucleotide bearing a quencher (grey) in addition to the fluorophore. Hybridization to the target separates these two components and results in fluorescence enhancement. c) Alternatively, a DNA or, more commonly, PNA (peptide nucleic acid) probe can be modified with an environment-sensitive fluorophore at an internal position. Hybridization with the target leads to forced intercalation (FIT) of the fluorophore, which results in fluorescence enhancement. d) The RNA of interest can be extended by an MS2 binding site (MBS) hairpin, which recruits a coat protein (MCP) fused to a fluorescent reporter protein (FP, green). The fluorescence signal of the target can be increased by appending multiple copies of the MBS. e) RNA-binding proteins of the Pumilio homology domain (PMH) can be engineered to recognize specific 8 nt regions on the RNA of interest (red). Two adjacent regions can then be utilized to reconstitute a split fluorescent protein, thereby activating its fluorescence.

Besides PNA, other types of modified backbones that have been used to construct hybridization probes include 2'-*O*-methyl-modified RNA and locked nucleic acid (LNA, Chart 1). In addition to protection from enzymatic degradation, these modifications reduce nonspecific activation of the probes by DNA- or RNA-binding proteins and substantially increase their binding affinity to the target, which allows more effective binding of structured RNAs.<sup>32,33</sup>

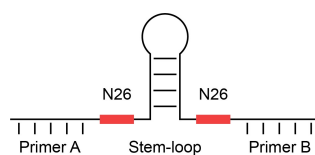


**Chart 1.** Instead of natural nucleic acid backbones, modified polymers have been used for generating hybridization probes that are protected from degradation in live cells and offer stronger, more specific binding to their targets. Examples of artificial backbone types include peptide nucleic acid (PNA), 2'-*O*-methyl-modified RNA and locked nucleic acid (LNA).<sup>34</sup>

For live cell and in vivo imaging, an endogenous, i.e. genetically encoded reporter is often desirable. Such a system has been constructed from a coat protein (MCP) that recognizes and binds a specific RNA hairpin (MS2 binding site, MBS) in the genomic RNA of the bacteriophage MS2.<sup>35</sup> The MBS hairpin is attached to the RNA of interest and recruits a construct comprising MCP fused to a fluorescent reporter protein such as GFP. Similar to the problems encountered with FISH, the steady expression of a fluorescent protein results in background fluorescence that hampers specific detection of the target. Therefore, multiple copies of the MBS hairpin are commonly used in series to increase the local brightness at the target RNA above the nonspecific background (Scheme 4d). However, this approach introduces a new problem: By attaching multiple hairpins to the RNA, each of which recruits a large fusion protein, the size of the probe easily dwarves the RNA of interest. This might interfere with the function and subcellular localization of the transcript.<sup>26</sup> To mitigate the need for multiple copies of the MS2 system, split fluorescent protein complementation strategies have been developed. Therein, the reporter protein is expressed as two fragments that are inactive and need to be reconstituted to become fluorescent. Each fragment is fused to an RNA-binding protein, e.g. the Pumilio homology domain (PUM-HD) of the human Puf protein.<sup>36</sup> PUM-HD recognizes a stretch of eight nucleotides via an equal number of subdomains, which can be engineered to target a sequence of choice present in the RNA of interest. Two such PUM-HD fusion proteins can therefore be used to colocalize the reporter protein fragments on the target for fluorescence activation (Scheme 4f).<sup>37</sup> At this point, the probe is genetically encoded, binds to a specific part of the target without needing an additional recruiter sequence and displays fluorescence turn-on only after binding. However, the issue of large probe size with respect to a typical RNA of interest remains. This is finally addressed by the introduction of aptamers, short RNA oligonucleotides that interact with small molecule ligand in a highly specific manner. In the following sections, the process of generating RNA aptamers for a given ligand and common features of RNA–ligand interactions will be discussed, before giving a detailed overview of their application as fluorescent tags.

### 1.1.3 Small molecule recognition by RNA aptamers

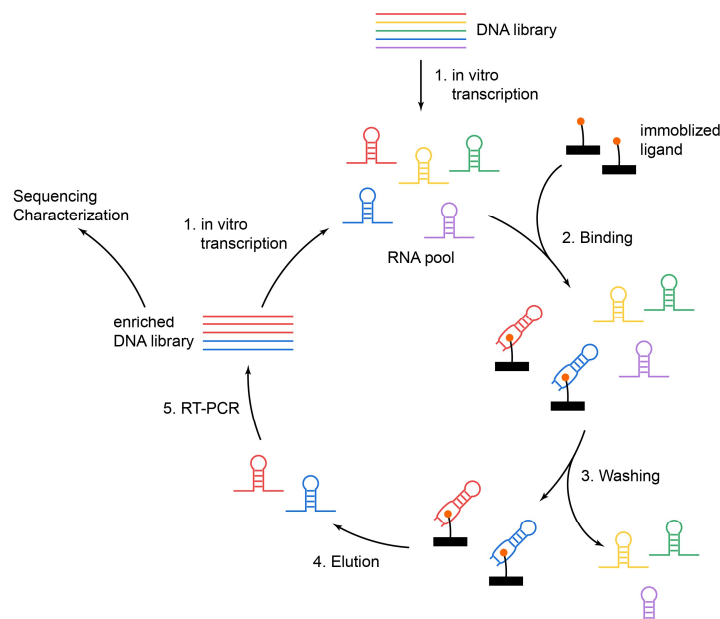
RNA molecules can adopt diverse tertiary structure motifs beyond the simple A form duplex, which is key to their functional versatility. A salient feature of RNA is its ability to interact as a riboswitch with small organic compounds in a highly specific manner. Metabolite-activated riboswitches, which are typically located in the 5'-untranslated regions (UTR) of mRNA, comprise a ligand-binding aptamer domain and an expression platform as the actual regulatory structural element. Binding of the target molecule to the aptamer domain induces a conformational change that sequesters or releases regulatory elements acting in cis on the mRNA, which modulates the transcription of the nascent RNA molecule or its translation into a protein.<sup>38</sup> While naturally occurring riboswitches were first described only in 2002,<sup>39-41</sup> artificial standalone aptamers capable of binding to organic dyes had already been developed more than 10 years earlier by Ellington and Szostak using an in vitro selection approach.<sup>42</sup> At the same time, similar but independent experiments for generating RNAs capable of binding to T4 DNA polymerase were carried out by Tuerk and Gold, who termed their procedure SELEX (systematic evolution of ligands by exponential enrichment).<sup>43</sup> Generally speaking, in vitro selection experiments start from a library of random DNA or RNA sequences. Figure 3 shows the key features of a library design that has been used for generating several classes of aptamers with high binding affinity:<sup>44</sup> Two random regions of 26 nt each are connected by a constant 12 nt stem-loop and flanked by constant primer binding sites required for PCR amplification. The stem-loop is supposed to facilitate folding of the structure motifs that arise in the random region. In theory, such a library could contain  $4^{52} \approx 10^{31}$  different sequences. In practice, a typical library size is approximately  $10^{15}$  molecules after initial PCR amplification, i.e. at best one out of every  $10^{16}$  possible sequences is sampled.



**Figure 3.** Example of a standard partially structured library design for in vitro selection of RNA aptamers. Two random regions of typically 26 nt are joined by a 12 nt stem-loop.<sup>44</sup>

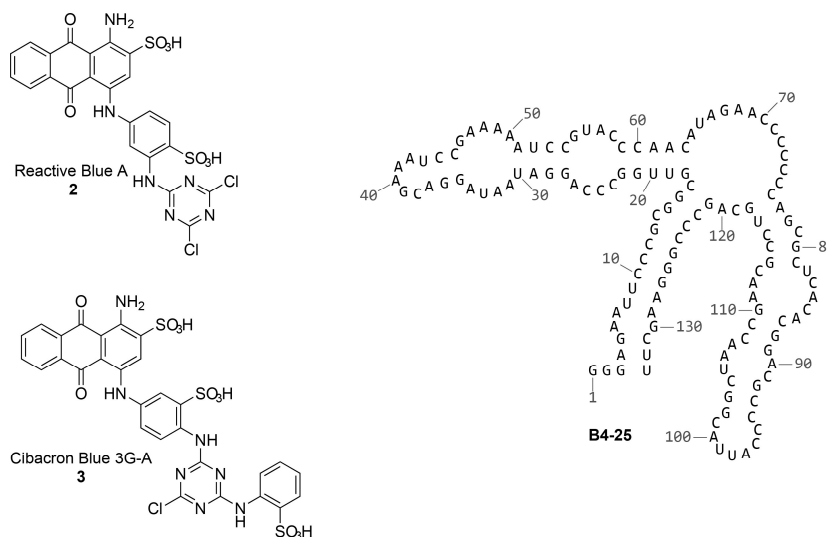
The PCR-amplified DNA library is transcribed into an RNA pool and incubated with the target molecule, which has been immobilized on a solid support either by forming a covalent bond or by using an affinity tag. Individual pool members that are capable of binding to the target are captured on the solid support, while non-binders can be removed by washing with a suitable buffer. Next, those RNAs that remained bound are eluted from the solid support by treatment with the free target compound. These sequences are converted into a new DNA pool by RT-PCR, at which point the next selection cycle begins. Over the course of several rounds, binding-competent sequences become enriched; the selection can be guided to a certain extent e.g. by lowering the incubation time, increasing the stringency of the washing step, including competitor molecules or varying temperature and buffer composition. Finally, the DNA library is sequenced and the resulting aptamers or aptamer families are screened for their individual binding affinity towards the target (Figure 4).<sup>45</sup>





**Figure 4.** General overview of a typical in vitro selection experiment for finding RNA aptamers. An initial DNA library is transcribed into RNA (1), which is exposed to the target ligand immobilized on a suitable matrix (2). Binders and non-binders are separated by suitable washing and elution steps (3,4). RT-PCT is used to generate a DNA library in which the binding RNA sequences have been enriched (5). From there on, a new selection round can begin or the library can be sequenced to identify and analyze its individual members.<sup>45</sup>

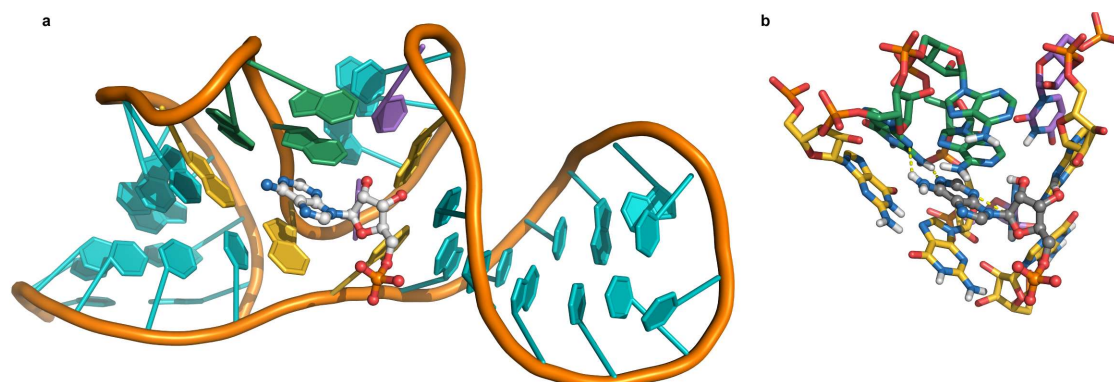
Already during the first studies, it became evident that aptamers are easily capable of discriminating between their targets and closely related compounds in a highly specific manner. For example, B4-25, one of the first dye-binding aptamer, recognizes Reactive Blue A (**2**), but rejects Cibacron Blue 3G-A (**3**) despite both sharing the same anthraquinone core (Chart 2). The precise structural basis for ligand discrimination is known only for a small set of aptamers that were thoroughly characterized by X-ray crystallography and NMR spectroscopy.



**Chart 2.** Predicted secondary structure of the RNA aptamer B4-25, which selectively binds the anthraquinone dye Reactive Blue A (**2**), but not the related Cibacron Blue 3G-A (**3**).<sup>42</sup>

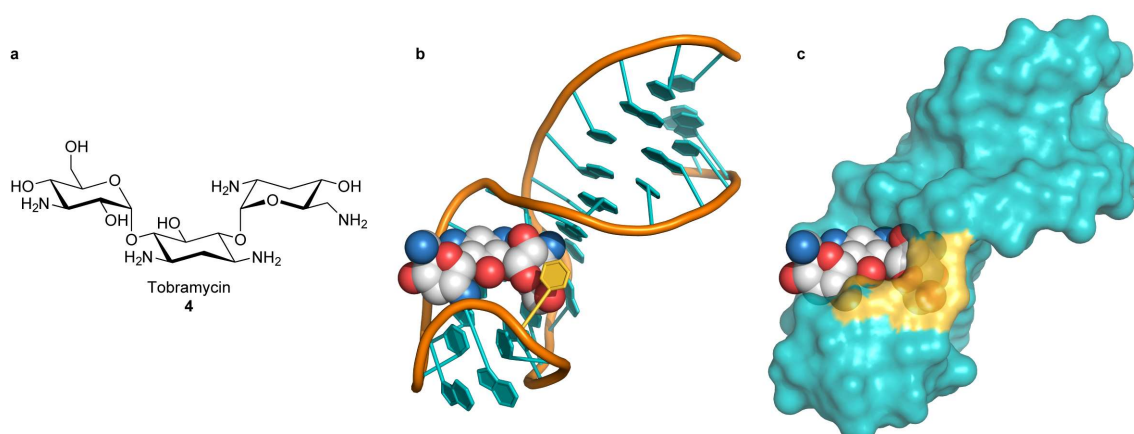
NMR-based structure elucidation of two ATP aptamer variants that are able to also recognize AMP revealed four important types of RNA–ligand interactions: AMP is part of a GNRA tetraloop-like motif,<sup>46</sup> taking the place of the fourth residue. Here, its Watson-Crick faces engages in hydrogen bonding with a mismatched guanine in close proximity, whereas N3 is hydrogen-bonded to the amino group of an adjacent adenine. The ligand is additionally held in place by  $\pi$  stacking

interactions with adenine and guanine residues located directly above and below its aromatic ring plane. Even the 2'- and 3'-hydroxy groups of AMP are involved in polar and Van-der-Waals contacts with the aptamer, meaning that specific recognition is achieved by recruiting a majority of the polar moieties available in the ligand (Figure 5).<sup>47,48</sup>



**Figure 5.** a) NMR structure of an ATP aptamer in complex with AMP (PDB: 1RAW). Nucleotides not engaged in ligand binding are colored in cyan. b) The perimeter of the binding site is formed by an internal four nucleotide loop (yellow). The ligand is part of a GNRA-like motif (green) and is engaged in hydrogen bonding with nearby residues (violet).<sup>47</sup>

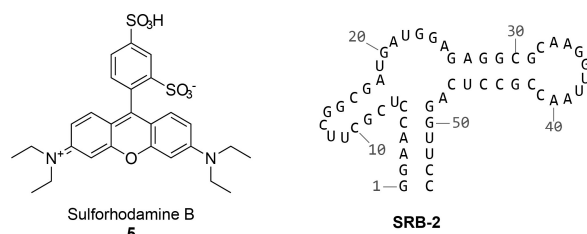
This situation is certainly unique due to the ligand being a nucleoside monophosphate, and altogether different binding modes are observed for other systems.<sup>49</sup> The aminoglycoside antibiotic tobramycin (**4**), for example, binds to a well-defined position in the major groove of two cognate aptamers in a sort of shape-adapted manner (Figure 6). Residues that are looped out of the RNA duplex serve as flaps that shield part of the ligand from the surrounding medium and increase its buried surface area. Because the amino groups of tobramycin are protonated at neutral pH, they can potentially engage in favorable electrostatic interactions with the negatively charged phosphate backbone of the RNA, which has been suggested as an explanation for the high binding affinity of the complex.<sup>50,51</sup>



**Figure 6.** a) Tobramycin (**4**) specifically binds to a section of the major groove in a cognate aptamer (PDB: 1TOB). b) The amino groups of the ligands are protonated at neutral pH and undergo electrostatic interactions with the phosphate backbone of the aptamer. c) A surface representation of the aptamer shows that a single cytosine (yellow) acts as a flap and essentially buries the ligand within the major groove.<sup>51</sup>

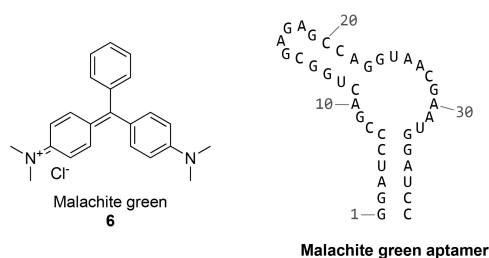
Aptamers such as the ones discussed above, which are capable of binding a cognate ligand with high specificity, naturally lend themselves to the construction of molecular switches for the control of gene expression<sup>52</sup> and can be employed on a more direct level as sensor devices for in vitro applications if conformational changes upon complex formation lead to the modulation of an observable signal such as the fluorescence of an attached reporter group.<sup>53</sup>

An alternative approach to aptamer utilization is exploitation of a specific property inherent to the ligand, which can be localized to an RNA of interest when an aptamer–ligand complex is formed. This is exemplified in an early work by Szostak and coworkers, who selected an aptamer capable of binding to the fluorescent dye sulforhodamine B (**5**, Chart 3). A 54 nt minimal active variant, SRB-2, binds to its cognate ligand with a dissociation constant ( $K_d$ ) of  $3.1 \pm 0.6 \times 10^{-7} \text{ mol L}^{-1}$  and minimal change in fluorescence intensity. Screening against a small library of similar fluorophores and ligand fragments revealed that the oxygen atom of the xanthene moiety is important for binding and suggested that the presence of an overall negative charge in the ligand is beneficial.<sup>54</sup> Later studies by Sunbul and Jäschke, however, demonstrated that some neutral and positively charged xanthene, acridine and oxazine derivatives can bind to SRB-2 even stronger than sulforhodamine B.<sup>55</sup> Functional SRB-2 was incorporated into elongated RNA constructs, e.g. as part of the 3'-UTR of a >2000 nt mRNA coding for enhanced green fluorescent protein (EGFP) and could also be expressed in COS-7 cells. Live cell fluorescent tagging of a target RNA, however, was not practical due to nonspecific binding of the ligand to basic amino acids.<sup>56</sup>



**Chart 3.** Structures of sulforhodamine B (**5**) and its cognate aptamer.<sup>54</sup>

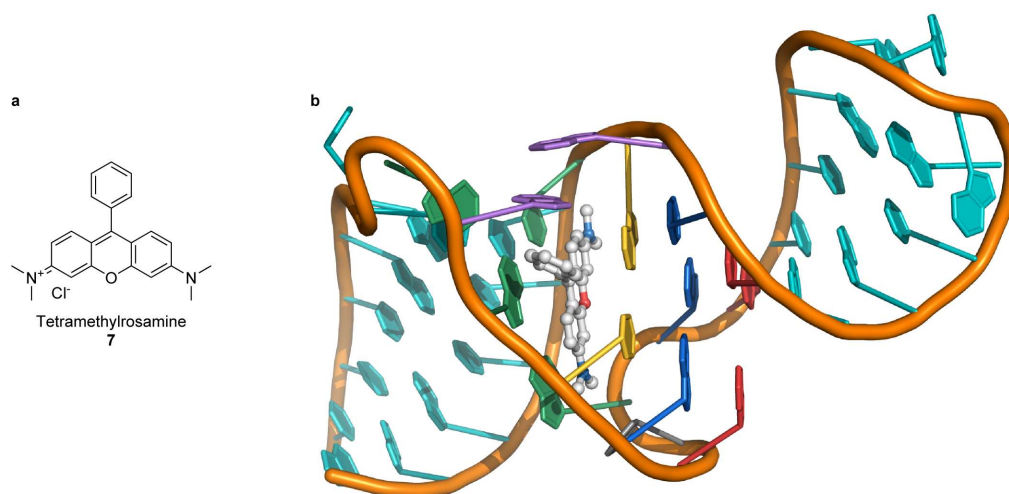
Grate and Wilson sought to identify an RNA aptamer that can bind the triphenylmethane dye malachite green (**6**).<sup>57</sup> Previously, antibodies covalently modified with malachite green had been used for chromophore-assisted laser inactivation (CALI) of proteins; after binding of such an antibody to its target protein, laser irradiation leads to the formation of highly reactive hydroxyl radicals that cause localized damage. This enables deactivation of proteins with high spatiotemporal control to study their function in a live cell environment.<sup>58,59</sup> An analogous RNA aptamer could therefore be used to apply CALI on the transcript level.



**Chart 4.** Structures of the triphenylmethane dye malachite green (**6**) and its cognate aptamer.<sup>57</sup>

SELEX with agarose-bound malachite green yielded a 112 nt aptamer, from which a 38 nt minimal binding-competent variant was generated after structural probing by alkaline hydrolysis (Chart 4). Binding to malachite green occurred with a  $K_d$  of approximately  $1 \times 10^{-6} \text{ mol L}^{-1}$  as measured by a shift in the absorption maximum of the ligand. In vitro experiments showed that the formed complex experiences highly site-specific backbone cleavage under 630 nm laser irradiation.<sup>57</sup> X-ray-crystallographic analysis of the malachite green aptamer in complex with the structurally similar rhodamine dye tetramethylrosamine (**7**) revealed an unusual binding site architecture (Figure 7). The xanthene moiety of the ligand was found to be stacked between a G–C Watson-Crick base pair on one side and a base quadruple formed around a central guanine on the other

side. At the same time, two adenine residues undergo stacking with the phenyl group of **7**. Two consecutive base triples complete the extensive set of stacking interactions in the binding site.<sup>60</sup>

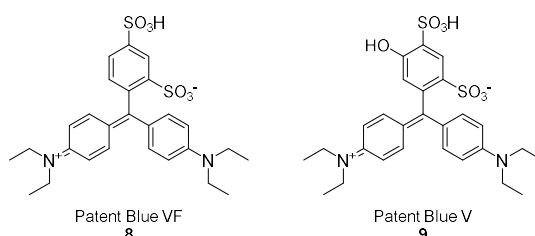


**Figure 7.** a) Tetramethylrosamine (**7**) can bind to the malachite green aptamer instead of its original ligand. b) Crystal structure of the malachite green aptamer in complex with **7** (PDB: 1F1T). The xanthene core of the ligand is stacked between a base quadruple (green) and a base pair (yellow), which is stabilized by interaction with additional base triples (blue and red). Additional stacking interactions occur between the phenyl side chain of the ligand and two adenine residues (violet).<sup>60</sup>

## 1.2 FLUOROGENIC RNA APTAMERS

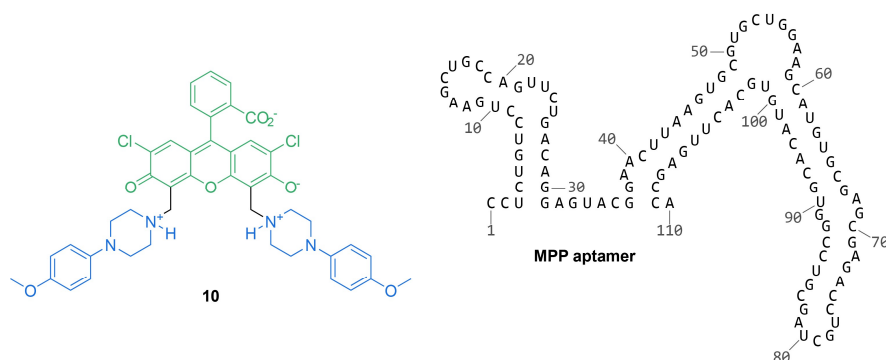
### 1.2.1 The serendipitous discovery of RNA-based fluorescence activation

Under normal circumstances, triphenylmethane dyes such as malachite green (**6**) exhibit extremely low fluorescence quantum yields due to efficient non-radiative deactivation caused by vibrational and torsional motion of the scaffold.<sup>61</sup> Fluorescence enhancement occurs if these pathways are discouraged, e.g. in a high viscosity medium or at low temperature. Based on the just discussed crystal structure, Tsien and coworkers recognized that the malachite green aptamer might constitute an environment that is capable of rigidifying bound malachite green and thereby enhancing its fluorescence quantum yield. They found that formation of the complex resulted in a more than 2300-fold increase in fluorescence intensity relative to the unbound ligand. This activation was specific to the cognate aptamer, as the aforementioned SRB-2 aptamer was not capable of inducing malachite green fluorescence. However, when SRB-2 was incubated with either Patent Blue VF (**8**) or Patent Blue V (**9**, Chart 5), two triphenylmethane analogs of sulforhodamine B, approximately 100-fold fluorescence enhancement was observed.



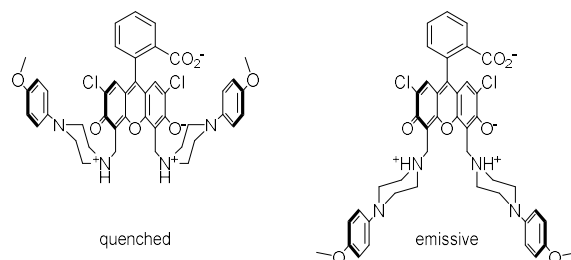
**Chart 5.** Structures of triphenylmethane dyes that show fluorescence activation by the SRB-2 aptamer.

The discovery of RNA-dependent fluorogenic behavior was a significant leap forward in the development of fluorescence-based RNA detection and imaging technology as it addresses many of the design challenges posed by earlier systems: An added aptamer tag is less likely to interfere with the folding and proper function than hybridization probes and molecular beacons. Also, a short oligonucleotide of only a few 10–100 nt is considerably smaller than the probes used in protein-based methods (MS2 system, RNA-binding proteins), which weigh in at several 100 kDa and therefore can sometimes dwarf the RNA of interest, possibly affecting cellular trafficking and interactions with other macromolecules.<sup>62</sup> In the case of malachite green as a fluorogen, however, concerns have been raised with respect to potential cytotoxic effects in light of its propensity to generate reactive oxygen species upon illumination.<sup>63,64</sup> It was shown only recently that live cell imaging of *E. coli* expressing the malachite green aptamer is possible when supplementing the growth medium with 5  $\mu\text{M}$  of the ligand.<sup>65</sup> In the meantime, the high potential of fluorogenic aptamers had spurred the development of several new fluoromodules (i.e. combinations of aptamers and ligands). For example, a 110 nt RNA aptamer was selected to activate the 2,7-dichlorofluorescein derivative **10**, which was originally designed as a reversibly quenched chemosensor (Chart 6).<sup>66</sup>



**Chart 6.** The engineered 2,7-dichlorofluorescein derivative **10** comprises fluorophore (green) and MPP quencher (blue) moieties, the latter of which can be sequestered by an RNA aptamer.<sup>67</sup>

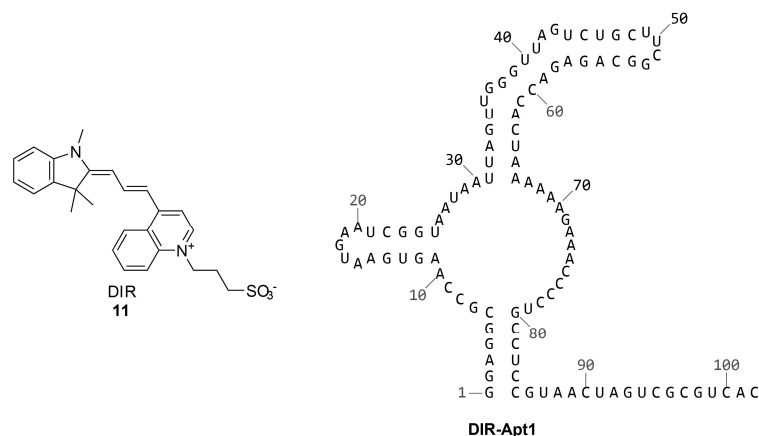
In compound **10**, the fluorophore is rendered non-emissive by the presence of *N*-(4-methoxyphenyl)piperazinyl (MPP) side chains at positions 4 and 5 of the xanthene core, which serve as donors for photoinduced electron transfer (PET). According to NOESY NMR, the molecule preferentially assumes a closed conformation that is efficiently quenched by PET. However, sequestration of the side chains by  $\beta$ -cyclodextrin leads to an open conformation, which interrupts PET and thereby restores fluorescence (Chart 7).<sup>67</sup> For in vitro selection, only the MPP moiety was immobilized on a solid matrix in order to generate aptamers that specifically interact with the quencher group. Compound **10** itself was deemed unsuitable as a selection substrate because the resulting aptamers would likely only bind to the non-emissive conformer without inducing fluorescence activation. Although binding of the MPP aptamer to compound **10** was found to be weak, the resulting 13-fold fluorescence turn-on of bound vs unbound ligand demonstrated the viability of the concept, which was later picked up and polished by the Jäschke group (see section 1.2.3).<sup>66</sup>



**Chart 7.** In solution, the 2,7-dichlorofluorescein derivative **10** adopts a closed conformation (left). The proximity between the fluorophore moiety and the amine side chains results in fluorescence quenching by PET. Binding of the side chain to  $\beta$ -cyclodextrin or the MPP aptamer results in an open conformation, which is no longer quenched by PET.<sup>67</sup>

Armitage and coworkers pursued a rational design approach for the synthesis of a new fluorogenic cyanine dye, dimethylindole red (DIR, **11**). Unlike their symmetrical counterparts, unsymmetrical cyanines are typically quenched due to torsional motions of the central methine bridge, which lead to efficient non-radiative deactivation of the excited state.<sup>68</sup> Conformational restriction, e.g. due to DNA or RNA intercalation, results in fluorescence enhancement, which could potentially be exploited in a targeted manner by an aptamer. It was reasoned that the bulky dimethylindole moiety should prevent intercalation into non-cognate nucleic acids; additionally, the negatively charged sulfonate side chain could prevent interactions with the phosphate backbone unless specifically tolerated by an aptamer. After in vitro selection, 32 DIR-binding aptamers were identified, three of which were capable of strongly activating DIR fluorescence. The most promising candidate, DIR-Apt1 (Chart 8), binds with a  $K_d$  of  $8.6 \pm 1.7 \times 10^{-8}$  mol L<sup>-1</sup> and leads to a

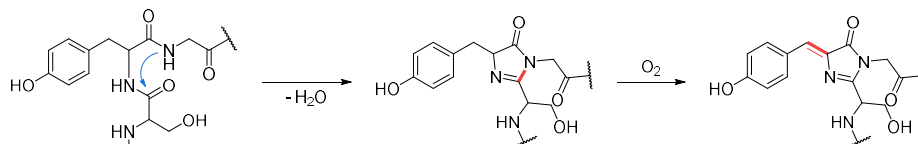
more than 40-fold fluorescence enhancement. Control experiments with calf thymus DNA and the malachite green aptamer confirmed specific activation of DIR.<sup>69</sup>



**Chart 8.** Predicted secondary structure of the DIR-Apt1 aptamer and its cognate ligand dimethylindole red (DIR, **11**).<sup>69</sup>

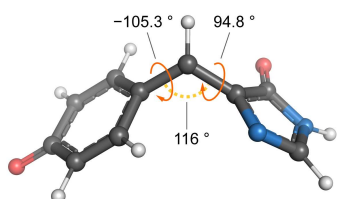
### 1.2.2 In vitro selection against protein-like fluorogens

While the development of new aptamer fluoromodules was flourishing, their application to live cell imaging had not yet progressed far because of concerns about the potential toxicity of exogenous probes, limited cell permeability of some fluorogens and often low contrast between the RNA of interest and nonspecific background fluorescence.<sup>70</sup> A major breakthrough, however, was achieved in 2011 by Jaffrey and coworkers, who proposed the use of biomimetic fluorogens as a new ligand class.<sup>71</sup> They turned to the hydroxybenzylidene imidazolone (HBI) chromophore scaffold, which is formed during the maturation of GFP by an autocatalytic cyclization of a serine-tyrosine-glycine tripeptide (Scheme 5).



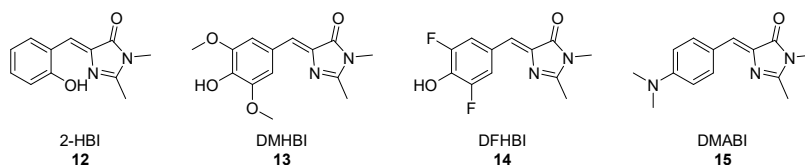
**Scheme 5.** Autocatalytic cyclization of a serine-tyrosine-glycine tripeptide followed by oxidation leads to the formation of an HBI chromophore embedded in the β-barrel of GFP.<sup>72</sup>

On its own, HBI is not fluorescent, as it undergoes efficient non-radiative deactivation after photoexcitation. Due to the tremendous importance of fluorescent proteins for live cell imaging, the photophysics of HBI have been studied in great detail; it has been suggested that geometry relaxation in the first excited state ( $S_1$ ) is required to generate an emissive species, which is susceptible to torsional motions of the benzylidene and imidazolone moieties ("hula twist"). Within the β-barrel environment of fluorescent proteins, the conformational freedom of HBI is limited, resulting in a high quantum yield of fluorescence. Once this restriction is lifted, however, a conical intersection between  $S_1$  and the ground state ( $S_0$ ) becomes accessible, opening up a highly efficient pathway for radiationless deactivation (Figure 8).<sup>73</sup> It was reasoned that the structural distortion responsible for fluorescence quenching is subtle enough to prevent unspecific activation of HBI by cellular components. The binding pocket of an RNA aptamer on the other hand, might be able to impose sufficient conformational restraints to induce fluorescence.



**Figure 8.** Gas-phase geometry of the HBI anion at the conical intersection between  $S_1$  and  $S_0$  (CASSCF/6-31G\*).<sup>73</sup>

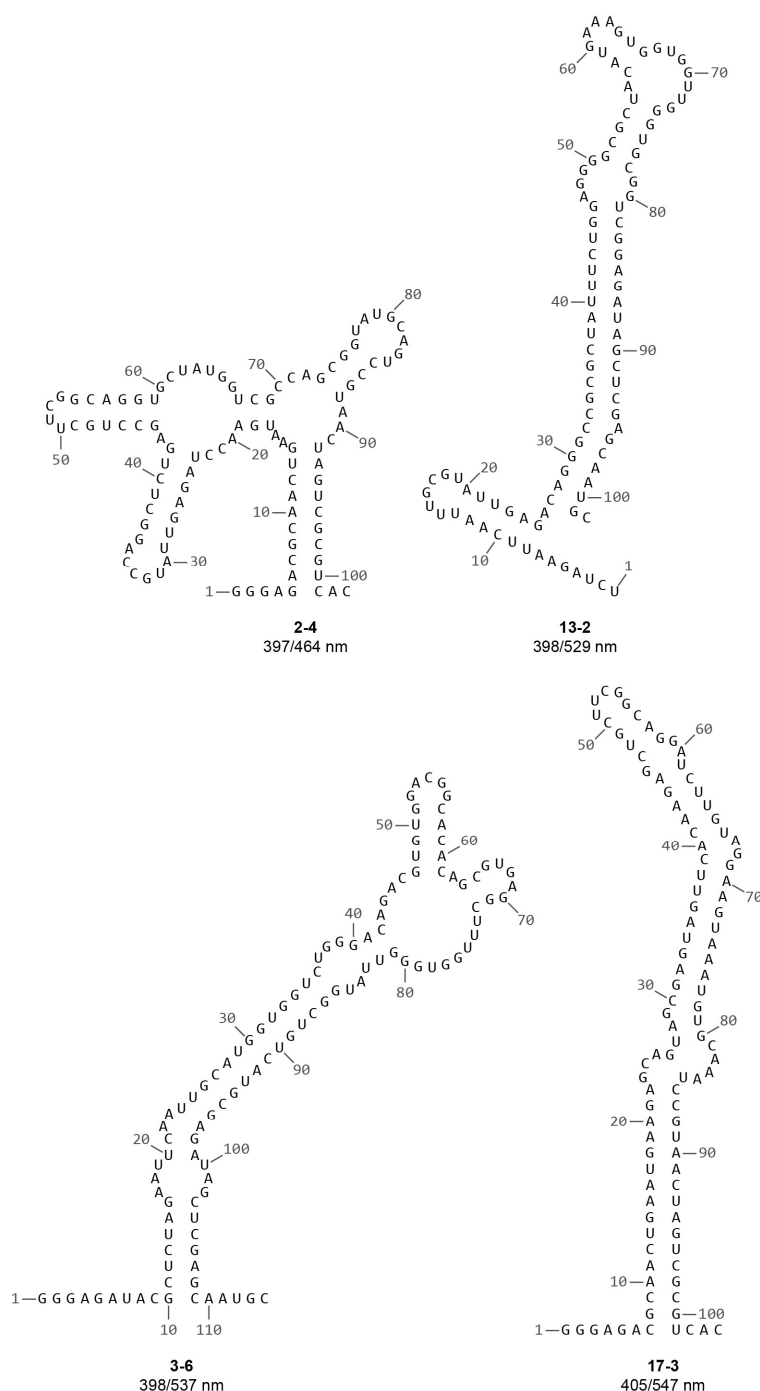
Four new HBI derivatives (2-HBI **12**, DMHBI **13**, DFHBI **14** and DMABI **15**, Chart 9) were synthesized for the selection of suitable aptamers. Unlike malachite green (**6**), these did not show unspecific fluorescence activation when exposed to HEK293T cells and were non-cytotoxic when illuminated according to an MTT assay.<sup>74</sup>



**Chart 9.** Four HBI derivatives were used as substrates for the *in vitro* selection of RNA mimics of fluorescent proteins.<sup>71</sup>

After ten rounds of SELEX against immobilized DMHBI, several aptamers were identified, among which a 102 nt clone called 13-2 (Chart 10) exhibited the highest degree of fluorescence turn-on, resulting in a brightness (extinction coefficient  $\times$  quantum yield) of 12% relative to GFP; based on truncation and mutagenesis studies, a 60 nt minimal domain (13-2min, see Chart 19 in section 2.1.1) was obtained. The 13-2–DMHBI complex was formed with a  $K_d$  of  $4.64 \times 10^{-7}$  mol L<sup>-1</sup> and displayed an emission maximum of 529 nm when excited at 398 nm. Interestingly, subsequent selection against new starting libraries yielded three additional aptamers with excitation/emission maxima between 397/464 and 405/547 nm, respectively, indicating that DMHBI experiences spectral tuning by the RNA (Chart 10). The available color palette was further extended by selecting aptamers against HBI analogs **12** and **15**.

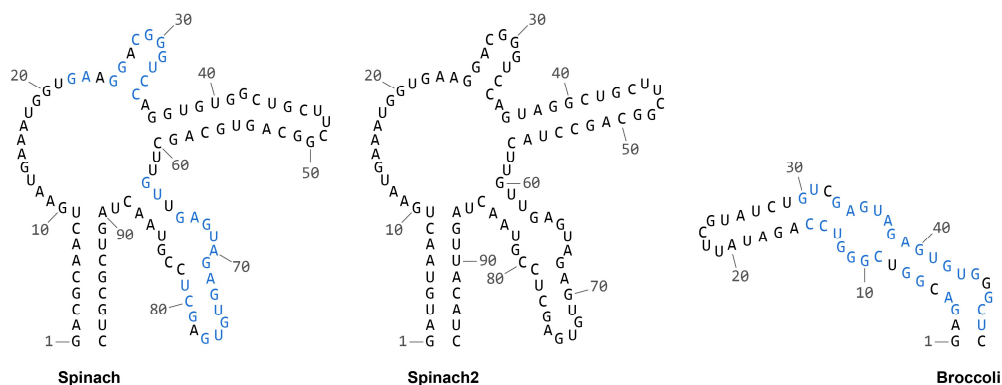




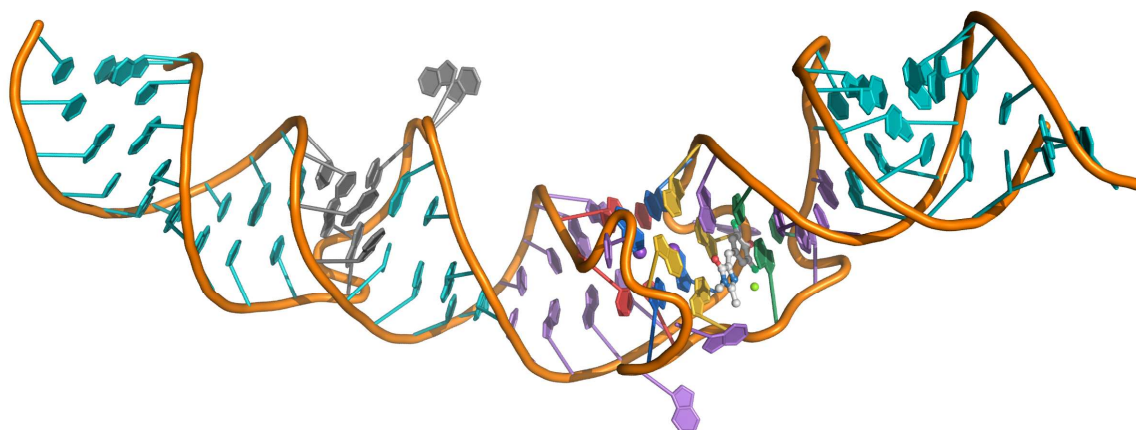
**Chart 10.** Predicted secondary structures of four DMHBI-binding aptamers and excitation/emission maxima of the resulting fluorescent complexes.<sup>71</sup>

Excitation spectra of the four aptamer–DMHBI complexes revealed that the ligand is predominantly bound in its neutral phenol form. It is well known that the fluorescence of enhanced GFP (eGFP), a serine–threonine mutant containing HBI in its phenolate form, is substantially brighter than that of wt-GFP because of the higher extinction coefficient of the anionic chromophore.<sup>75</sup> Therefore, a new aptamer, 24-2 (Chart 11), was selected against the DFHBI (**14**), which is completely deprotonated at neutral pH ( $pK_a = 5.5$ ) due to the electron-withdrawing effect of the fluorine substituents *ortho* to the hydroxy group. The 24-2–DFHBI complex formed with a  $K_d$  of  $5.37 \times 10^{-7}$  mol L<sup>-1</sup> and its green fluorescence ( $\lambda_{ex/em} = 469/501$  nm) mimics that of eGFP at approximately 53% of the latter’s brightness. In analogy to the widespread “fruit” nomenclature of fluorescent proteins,<sup>76</sup> this new 95 nt RNA aptamer was nicknamed

Spinach. After transforming *E. coli* cells with a plasmid encoding Spinach inserted into a tRNA scaffold that facilitates folding and lessens degradation by cellular nucleases,<sup>77</sup> incubation with DFHBI quickly lead to fluorescence activation. In HEK293T cells, a 5S rRNA tagged with Spinach-tRNA was used to probe e.g. the effect of an RNA polymerase III inhibitor on transcript abundance and the formation of stress granules.<sup>71</sup>



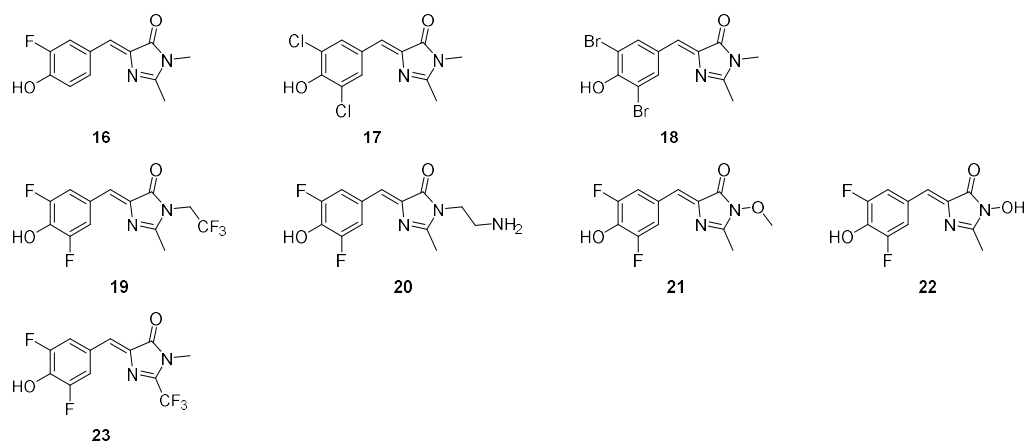
**Chart 11.** Predicted secondary structures of the Spinach, Spinach2 and Broccoli aptamers.<sup>78,79</sup> The sequence homology between Spinach and Broccoli is indicated by the blue-colored nucleosides.<sup>80</sup> Note that the actual tertiary structure of Spinach, and by extension Spinach2 and Broccoli, differs significantly from this early prediction.<sup>81,82</sup>



**Figure 9.** Crystal structure of the Spinach aptamer in complex with DFHBI (PDB: 4TS2). The ligand stacks on top of a two-tiered G-quadruplex (yellow, blue) that is supported by a mixed tetrad (red). Additional stacking interactions occur with a base triple (green) at the junction to the apical stem. Ancillary residues (violet) fence off the binding site. The basal stem is split by a short junction (grey).  $K^+$  ions intercalating between the quartets are shown as purple spheres.<sup>82</sup>

In 2014, two groups independently reported the crystal structures of apo (i.e. unliganded) Spinach and Spinach–DFHBI (Figure 9).<sup>81,82</sup> The binding site was found to contain a two-tiered G-quadruplex of unprecedented topology, which supports the ligand in a planar conformation. The ligand is further enclosed by stacking with a base triple that forms the junction between binding site and apical stem. The quadruplex transitions into the basal stem by means of an additional mixed tetrad.<sup>82</sup> In the absence of ligand, the tertiary structure of the aptamer remains largely intact, but the base triple collapses onto the proximal quartet of the G-quadruplex.<sup>81</sup> Knowledge of the structural basis for fluorescence activation allowed guided truncation of the RNA down to a minimal binding-competent domain to yield the 51 nt Baby Spinach aptamer, which exhibited 95% of the wt-Spinach brightness in complex with DFHBI.

Not long after the introduction of Spinach, shortcomings of this new fluoromodule began to surface. For example, Spinach-tagged (CGG)<sub>60</sub> repeat RNA in COS-7 cells did not show detectable fluorescence after incubation with the ligand despite stable expression. It was found that Spinach was susceptible to thermal denaturation with a melting temperature of 35 °C and was therefore mostly unfolded under live cell imaging conditions. Using a rational design approach, mismatched base pairs were corrected and superfluous unpaired bulges were removed to increase thermal stability; systematic mutagenesis studies were performed, which identified six brightness-enhancing point mutations. The resulting improved aptamer, Spinach2 (Chart 11), was up to 20 times brighter than its predecessor under otherwise identical conditions while maintaining similar excitation and emission maxima.<sup>78</sup> These wavelengths, however, are offset from the spectral windows of commercially available filter cubes for fluorescence microscopy with green fluorophores, which are often optimized for GFP or fluorescein isothiocyanate (FITC). Jaffrey and coworkers therefore synthesized new derivatives of DFHBI to modify the spectral characteristics of the ligand (Chart 12). Removing one of the fluorine atoms on the benzylidene moiety or replacing both by either chlorine or bromine (compounds **16–18**) did not significantly alter brightness or peak wavelengths. Modification of the imidazolone-N3 substituent, on the other hand, resulted in substantially redshifted excitation maxima (compounds **19–22**) but also lowered the binding affinity severely except in case of the trifluoroethyl derivative **19** (DFHBI-1T,  $\lambda_{ex/em} = 482/505$  nm in complex with Spinach2,  $K_d = 5.6 \times 10^{-7}$  mol L<sup>-1</sup>). Installation of a trifluoromethyl group at imidazolone-C2 (**23**) lead to a greater redshift that also extended to the emission maximum of the Spinach2 complex but was detrimental to both affinity and brightness.<sup>83</sup>

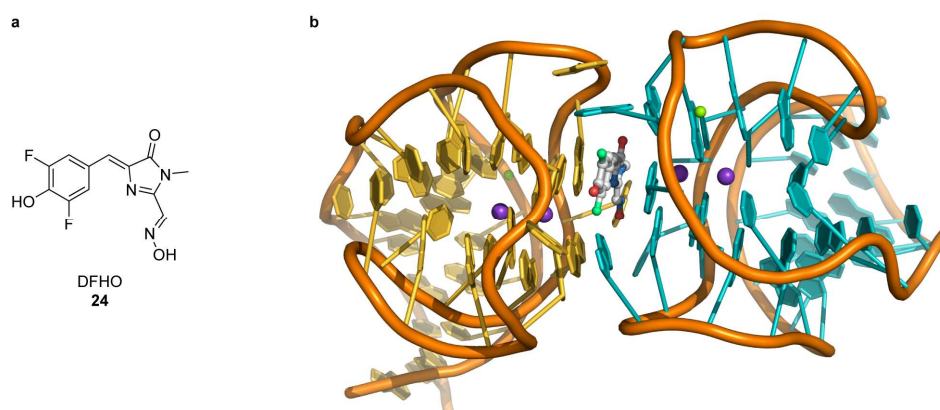


**Chart 12.** Structures of new DFHBI derivatives that are activated by the Spinach2 aptamer.<sup>83</sup>

A general challenge of generating bright RNA fluoromodules by in vitro selection lies in the procedure itself, because in a typical SELEX experiment enrichment is governed by the binding affinity between library members and the ligand. However, strongly binding aptamers do not necessarily induce ligand fluorescence, whereas RNAs with strong fluorogenic potential might be rejected due to lower affinity. In order to meet the demand for both affinity and brightness, the next iteration of HBI-activating aptamers was generated using a combined in vitro selection/live cell screening approach. As usual, SELEX of a starting library was performed against immobilized DFHBI. After six rounds, when pool diversity was still high but fluorogenic sequences became more abundant, the library was cloned into *E. coli* and screened by fluorescence-activated cell sorting (FACS). Several aptamer clones were identified which strongly activated DFHBI fluorescence under conditions that make proper RNA folding challenging, i.e. at 37 °C and with limited availability of free Mg<sup>2+</sup>. The clone with the brightest *E. coli* colony fluorescence, 29-1, also activated the improved fluorophore DFHBI-1T and was used for truncation experiments and as the input for

reselection with a doped library, ultimately yielding the 49 nt Broccoli aptamer (Chart 11). The DFHBI-1T complexes of Broccoli and Spinach2 showed virtually identical fluorescence brightness under optimal conditions, but the melting temperature of the new aptamer was 11 °C higher. Moreover, in  $Mg^{2+}$ -free buffer, the Broccoli activated DFHBI-1T fluorescence more than 5 times stronger than Spinach2. These improved in vitro characteristics also translated into the behavior of the fluoromodule in live cells. Imaging of HEK293T 5S rRNA was possible in the absence of supplemental  $Mg^{2+}$  and, unlike Spinach2, the aptamer folded correctly even when not placed in a stabilizing tRNA scaffold.<sup>79</sup> Surprisingly, despite their phenotypical differences the Spinach family and Broccoli share a highly conserved core, which underlines the functional importance of subtle sequence variations.<sup>84</sup>

Imaging studies with RNA mimics of GFP had mostly been of qualitative nature up to this point, as exact quantification of live cell fluorescence was hampered by the limited long term photostability of these systems.<sup>85,86</sup> Therefore, Jaffrey and coworkers sought to identify new aptamers that would bind to derivatives of DFHBI containing additional chemical handles for RNA–ligand interactions, which might improve the photostability of the fluorescent complex. Based on their previous efforts for the synthesis of improved ligands,<sup>83</sup> the choice fell on a modification of imidazolone-C2 with polar, conjugated side chains to achieve redshifted fluorescence. Among several tested candidates, all except the oxime-based DFHO (**24**, Figure 10a) showed nonspecific fluorescence activation by cellular constituents.<sup>87</sup> After eight rounds of conventional SELEX, a 76 nt aptamer that activated DFHO fluorescence was isolated and used as the starting point for the FACS-based selection of mutants that are functional under low- $Mg^{2+}$  conditions. Extensive truncation of the brightest clone resulted in a minimal 28 nt aptamer that was nicknamed Corn because of its yellow fluorescence in complex with DFHO ( $\lambda_{ex/em} = 505/545$  nm). Corn was moderately dependent on  $Mg^{2+}$  for binding of DFHO to form a complex with high affinity ( $K_d = 7 \times 10^{-8}$  mol L<sup>-1</sup>). As anticipated, this new fluoromodule exhibited increased photostability compared to Spinach- and Broccoli-based systems. Reselection of the 29-1 aptamer with DFHO yielded two additional fluorogenic RNAs (orange Broccoli,  $\lambda_{ex/em} = 513/562$  nm and red Broccoli,  $\lambda_{ex/em} = 518/582$  nm), but unlike Corn, their DFHO complexes were highly susceptible to photobleaching.<sup>88</sup> Their binding sites differ from that of original Broccoli and Spinach by the identity of a single nucleotide, which presumably controls spectral tuning of the ligand by interaction with its oxime moiety.<sup>80</sup>



**Figure 10.** a) DFHO (**24**) is an oxime-based ligand for the Corn aptamer. b) The active species of Corn is a homodimer (PDB: 5BJO). Two protomers (yellow and cyan) associate at their exposed G-quartet faces, between which the ligand can stack.  $K^+$  ions intercalating between the quartets are shown as purple spheres.<sup>89</sup>

Corn differs from other aptamers in that its functional form is a homodimer. The 28 nt RNA folds into a two-tiered G-quadruplex supported by two mixed sequence tetrads and a short base-paired

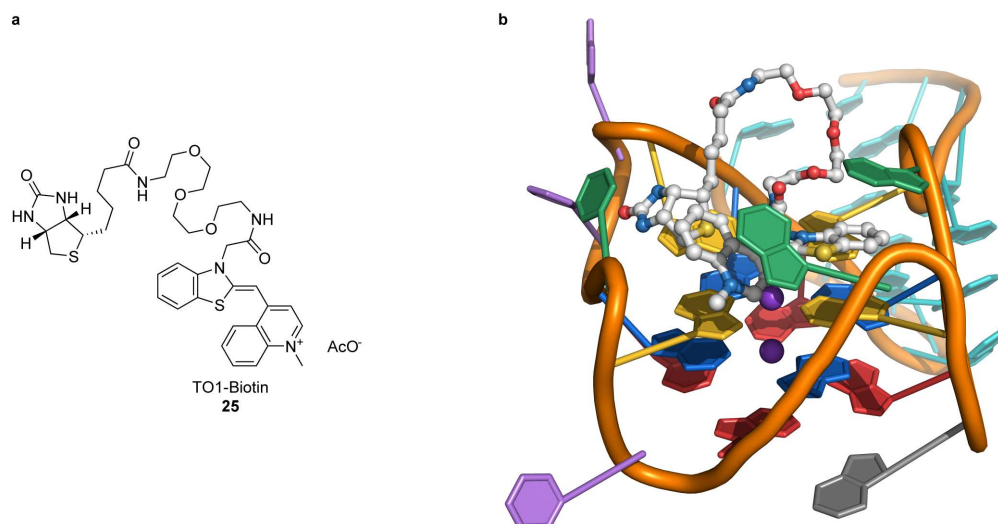
stem. Two of these protomers associate at their exposed G-quartet faces without requiring intermolecular base pairing interactions to form a dimer that is the actual fluorogenic species (Figure 10b).<sup>89</sup> RNA polymerase III activity under the control of different promoters and in response to the administration of a transcription inhibitor could readily be quantified in HEK293T cells using Corn as a reporter. It was however, deemed unsuitable as a tag for mRNA imaging because dimerization would likely interfere with the function of its target.<sup>88</sup> A follow-up study demonstrated that Corn could be fused to an RNA aptamer that binds *S*-adenosyl methionine (SAM) to generate an artificial riboswitch-like metabolite sensor. In the absence of SAM, the Corn domain of the construct was unable to dimerize and therefore did not induce DFHO fluorescence. Expression in HEK293T cells allowed monitoring of fluorescence intensity in response to treatment with the reversible SAM biosynthesis inhibitor cycloleucine.<sup>90</sup>

### 1.2.3 Recent developments

While Spinach and Broccoli variants are the most frequently used aptamer tags for RNA imaging at the moment, the development of new fluoromodules with interesting properties has steadily continued.<sup>45,91–94</sup> In the following section, some key examples will be discussed to provide an overview of the current state of the field.

#### **Mango aptamers**

Unrau and coworkers proposed fluorescent efficiency, which is defined as fluorescence enhancement divided by the dissociation constant, as an optimization goal for the selection of new fluorogenic aptamers. They reported relative efficiency values of 1, 62 and 11–20 for the GFP-MS2 system, the malachite green aptamer and Spinach, respectively, and reasoned that a fluoromodule with a low unbinding rate and high extinction coefficient might help to bridge the gap between Spinach and the malachite green aptamer. Due to its small size and low toxicity, the unsymmetrical cyanine dye thiazole orange (TO) was chosen as a selection target. To discourage nonspecific activation by cellular DNA or RNA, the benzothiazole moiety was derivatized with an acetate group. The selection substrate TO1-Biotin (**25**, Figure 11a) was prepared by linking this carboxylate group to biotin via a PEG chain, which allowed immobilization on streptavidin-coated magnetic beads. During *in vitro* selection, short incubation times were used to preferentially enrich RNAs that bind TO1-Biotin quickly, followed by thorough washing with ligand-containing buffer to remove RNAs that form labile complexes. After twelve rounds of selection, characterization of the isolated sequence families revealed an aptamer that bound TO1-Biotin with a  $K_d$  of  $3.2 \pm 0.7 \times 10^{-9} \text{ mol L}^{-1}$  and enhanced ligand fluorescence by a factor of 1100 ( $\lambda_{\text{ex/em}} = 510/535 \text{ nm}$ ), resulting in a relative efficiency of 970.<sup>95</sup> This new aptamer, which was named Mango, possesses a fluorogenic core of only 29 nt and folds into a three-tiered G-quadruplex that stacks with the thiazole orange moiety of the ligand and is stabilized by a short stem. Unexpectedly, the biotin side chain was found to be essential for ligand binding, despite being intended solely as an affinity tag for the *in vitro* selection. The crystal structure of the ligand complex shows that biotin directly interacts with nucleotides of the binding site by Van-der-Waals contacts and hydrogen bonding.<sup>96</sup> Despite its high fluorescence efficiency, the brightness of Mango–TO1-biotin was only half that of GFP.<sup>95</sup>



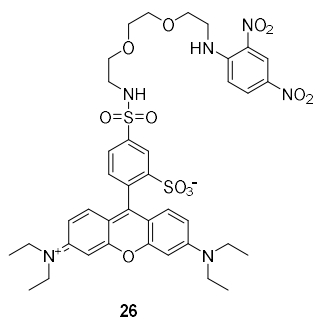
**Figure 11.** a) The fluorophore thiazole orange was attached to biotin via a PEG linker to furnish the aptamer selection substrate TO1-Biotin (**25**).<sup>95</sup> b) Crystal structure of the Mango aptamer in complex with TO1-Biotin (PDB: 5VF3). In addition to the thiazole orange moiety, biotin also interacts with nucleotides in the binding site (green). The ligand stacks on top of a three-tiered G-quadruplex (yellow, blue, red). The adenine residue connecting the two outer quartets (grey) augments one of them into a pentad. Three ancillary nucleotides (violet) connect the ligand-facing quartet to the other two layers of the G-quadruplex.  $K^+$  ions intercalating between the quartets are shown as purple spheres.<sup>96</sup>

Interestingly, the Spinach aptamer also binds TO1-Biotin with low affinity ( $K_d = 8.3 \pm 1.3 \times 10^{-7}$  mol L<sup>-1</sup>) but 4.5 times higher fluorescence brightness.<sup>97</sup> This suggests that the selection methodology for Mango might have been biased for strong binders at the expense of optimal fluorescence activation. The DNA pool obtained after the final round of the Mango selection was therefore subjected to a reselection by microfluidic-assisted in vitro compartmentalization ( $\mu$ IVC),<sup>98</sup> which produced three aptamer variants (Mango II–IV) that were found to be up to four times brighter than the original Mango I in complex with TO1-Biotin, while retaining high binding affinity.<sup>99</sup> The tertiary structure of the Mango II complex is largely similar to that of Mango I, but exhibits some key differences: in Mango II, the residues forming the G-tetrad that directly stacks with the ligand show less deviation from a planar arrangement than in Mango I, whereas the distal tetrad is stabilized further by interaction with two coplanar adenine residues, augmenting it into a hexad.<sup>100</sup> Mango III, on the other hand, contains a two-tiered G-quadruplex that narrows down to a canonical duplex via a base triple and is connected to a highly unusual, non-canonical parallel 3 bp duplex. Structure-guided mutagenesis and reselection with a partially randomized binding site produced two Mango III derivatives that displayed slightly higher brightness and minor spectral tuning compared to the wt aptamer.<sup>101</sup> Very recently, the favorable properties of Mango II–TO1-Biotin were exploited for single-molecule mRNA imaging in both fixed and live COS-7 cells. While that study employed an array of 24 aptamer copies to further increase fluorescence intensity, smaller arrays are likely still bright enough for this task. Despite its low dissociation constant, Mango II–TO1-Biotin gradually undergoes ligand recycling with the surrounding medium and therefore allows image acquisition over an extended time, which enabled super-resolution imaging of Mango II-labeled  $\beta$ -actin mRNA in fixed HEK293T cells via structured illumination microscopy (SIM).<sup>102</sup>

### Rainbow aptamer

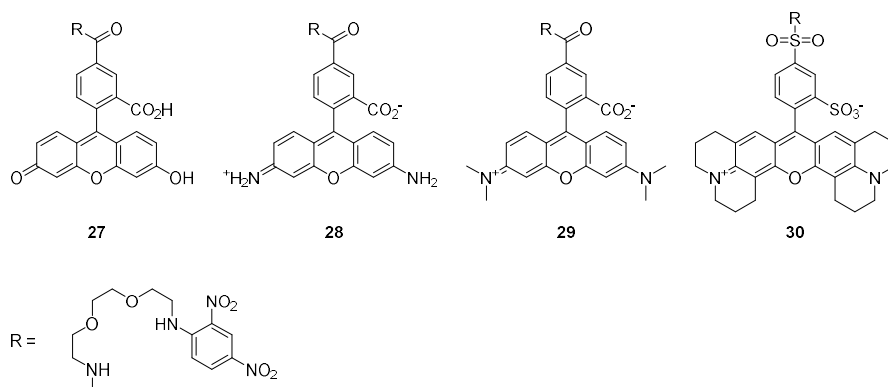
Sunbul and Jäschke expanded on the previously introduced concept of combining fluorophores with quencher groups to convert the SRB-2 aptamer (Chart 3) from a regular fluorescent to a fluorogenic tag. To this end, they linked the cognate ligand sulforhodamine B (**5**) to five different

contact quenchers via a short PEG chain. A derivative bearing a 2,4-dinitroaniline moiety (**26**, Chart 13) displayed 105-fold fluorescence enhancement after binding to SRB, but the dissociation constant increased by a factor of approximately five compared to unmodified **5**. This likely reflects the interaction energy between the fluorophore and the contact quencher, the separation of which was not accounted for when selecting SRB-2. Nonetheless, the binding affinity was still high enough to monitor the *in vitro* transcription of an SRB-2-tagged RNA in real time and to observe *E. coli* cells expressing the aptamer in a tRNA scaffold via fluorescence microscopy.<sup>103</sup>



**Chart 13.** Sulforhodamine B (**5**) was converted to the fluorogenic derivative **26** by attachment of 2,4-dinitroaniline as a quencher.

This system was later complemented by *in vitro* selection of the 75 nt DNB aptamer, which binds to the quencher 2,4-dinitroaniline instead of a specific fluorophore. The quencher was then conjugated to different rhodamine dyes (**27–30**, Chart 14), thereby converting them to conditional fluorophores. Quantum yields of these derivatives were in the range of 0.027–0.086, i.e. the quenching effect was moderate. Binding to the DNB aptamer occurred with micromolar dissociation constants and resulted in 5–73-fold fluorescence enhancement. *E. coli* cells expressing both the DNB and SRB-2 aptamer in individual tRNA scaffolds could be double-stained with **28** as a green fluorophore and a sulforhodamine B derivative quenched by 4-nitrobenzylamine as a yellow fluorophore.<sup>104</sup>



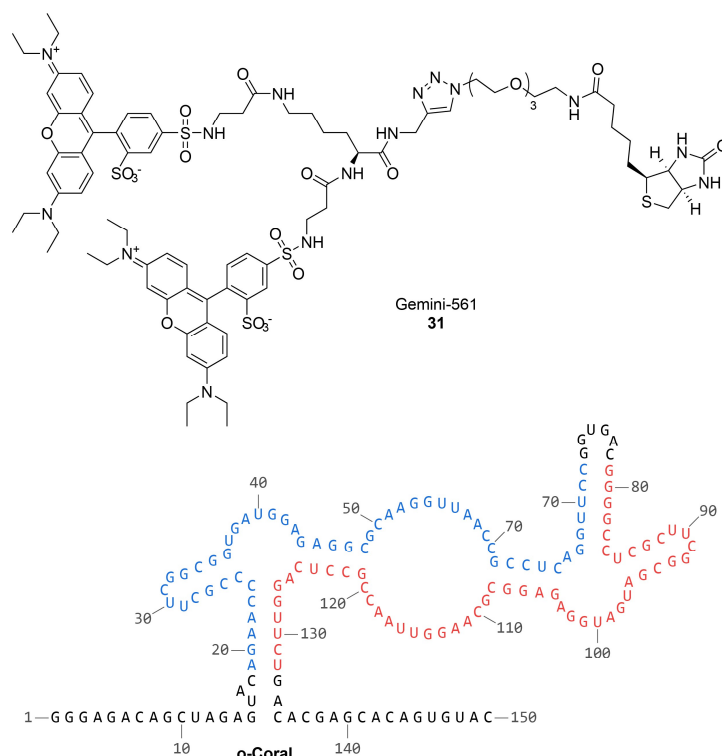
**Chart 14.** Different rhodamine dyes were linked to 2,4-dinitroaniline as a quencher to generate fluorogenic ligands for the DNB aptamer.

An elegant example for a variation of this concept is found in the riboglow platform, which comprises a naturally occurring cobalamin riboswitch<sup>105</sup> as a tag for the RNA of interest and rhodamine or cyanine fluorophores that are quenched by covalently attached cobalamin.<sup>106</sup>

### o-Coral aptamer

The SRB-2 and DNB systems suffer from the same fundamental design flaw: Both aptamers were selected to bind either the fluorophore or the quencher moiety of their respective ligands alone and are therefore not adapted for separating the two interacting components, which decreases binding affinity. A clever solution for this problem has been described by Ryckelynck and coworkers, who have designed the conditional fluorophore Gemini-561 (**31**). This compound is a

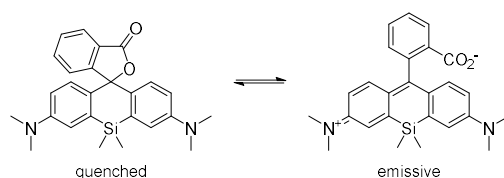
sulforhodamine B dimer that is quenched to formation of an intramolecular H aggregate. A cognate aptamer was isolated using a combined *in vitro* selection and  $\mu$ IVC screening strategy starting from an SRB-2 aptamer-derived mutant library. During selection, spontaneous dimerization resulted in the formation of what is essentially a dimeric SRB-2 aptamer. Secondary structure probing by nuclease digestion and systematic mutagenesis experiments suggest intertwined folding of the two SRB-2 subdomains (Chart 15). Truncation of nonessential linker nucleotides between the domains yielded the 150 nt o-Coral aptamer that binds to Gemini-561 with a  $K_d$  of  $7.3 \pm 0.2 \times 10^{-8} \text{ mol L}^{-1}$ . The extinction coefficient of the resulting complex is  $140000 \text{ L mol}^{-1} \text{ cm}^{-1}$ , which reflects the presence of two chromophores and results in a fluorescence brightness three times greater than that of eGFP.<sup>107</sup>



**Chart 15.** The dimeric sulforhodamine B-derivative Gemini-561 (**31**) binds to the o-Coral aptamer. A secondary structure model of o-Coral was established by mutagenesis and enzymatic probing and experiments. SRB-2 aptamer-like subdomains are colored blue and red.<sup>107</sup>

### Silicon rhodamine aptamer (SiRA)

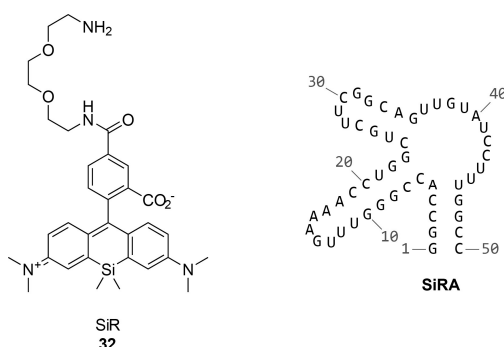
The spectral range of rhodamine-based fluoromolecules was recently expanded by Jäschke and coworkers, who selected a new aptamer for silicon rhodamine (SiR), a near-infrared fluorophore that has been popularized for super-resolution live cell imaging.<sup>108</sup> SiR dyes containing a carboxyl group at position 3 of the phenyl ring exist in an environment-dependent equilibrium between a dark spirolactone and a fluorescent zwitterionic form (Scheme 6). Binding to an RNA aptamer might influence the equilibrium such as to produce detectable fluorescence enhancement without the need for a covalently attached quencher.



**Scheme 6.** 3-Carboxy silicon rhodamine exists in equilibrium between a closed and an open form, which modulates fluorescence emission.



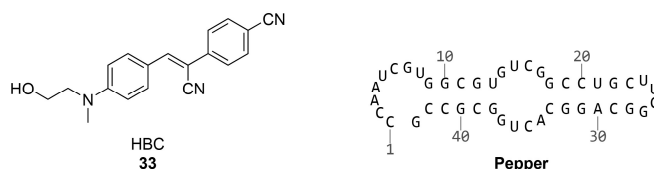
For *in vitro* selection, the target was attached to a solid support via an additional PEG amide linker connected to position 5 of the phenyl ring (compound **32**). Sequencing of the selection pool after 14 rounds of SELEX yielded 30 distinct aptamers, the majority of which shared a conserved tetranucleotide and short stem-loop in a three-way junction. A 50 nt minimal aptamer based on this motif, SiRA (Chart 16) bound to **32** with a  $K_d$  of  $4.3 \pm 0.7 \times 10^{-7}$  mol L<sup>-1</sup>. The resulting complex fluoresced seven times brighter than free **32** due to a combined increase of the extinction coefficient and fluorescence quantum yield, which reflects stabilization of the zwitterionic form of the bound ligand. Similar to Mango-TO1-biotin, the PEG chain seems to be involved in ligand binding, as its removal leads to a sixfold increase of the  $K_d$ . The fluorescence turn-on effect was enhanced by attaching a 2,4-dinitroaniline quencher to the PEG chain, but this also incurred a decrease of the binding affinity. Despite these seemingly unfavorable properties, imaging of a GFP mRNA tagged with five copies of SiRA embedded into a tRNA scaffold in live *E. coli* cells was readily possible. Moreover, the subcellular location of this RNA construct was imaged by stimulated emission depletion (STED) microscopy, which constituted the first application of fluorogenic aptamer-enabled super-resolution microscopy to live cells.<sup>109</sup>



**Chart 16.** Predicted secondary structure of the silicon rhodamine aptamer (SiRA). The cognate ligand **32** requires the presence of a PEG chain on the phenyl ring for optimal binding affinity.<sup>109</sup>

### Pepper aptamer

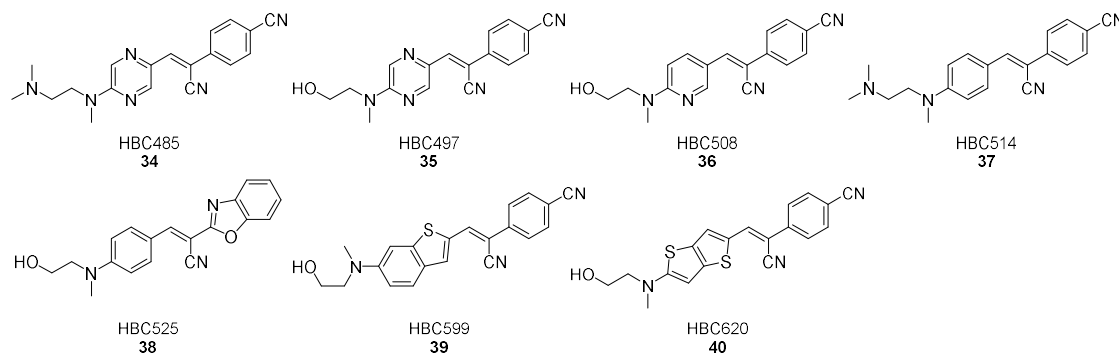
A novel class of donor-acceptor systems derived from (*Z*)- $\alpha$ -4-((2-hydroxyethyl)(methyl)amino)-benzylidene-4-cyanophenylacetonitrile (HBC, **33**) is the most recent addition to the growing palette of fluorogenic dyes. Like HBI, HBC derivatives are quenched in solution but become fluorescent when embedded into a confining matrix and therefore are a target for activation by an RNA aptamer. Conventional SELEX followed by truncation and systematic mutagenesis yielded such a 43 nt aptamer, Pepper (Chart 17), which enhanced HBC fluorescence 3000-fold ( $\lambda_{\text{ex/em}} = 485/530$  nm) upon binding with a  $K_d$  of approximately  $3.5 \times 10^{-9}$  mol L<sup>-1</sup>.



**Chart 17.** Predicted secondary structure of the Pepper aptamer. Pepper activates the fluorescence of HBC (**33**) and its derivatives **34–40**.<sup>110</sup>

HEK293T cells expressing Pepper fluoresced 8.5 and 11 times brighter than those expressing Broccoli or Corn, respectively, in the presence of their cognate ligands without requiring incorporation into a stabilizing scaffold such as tRNA. Due to its donor-acceptor nature, spectral tuning of HBC over sizeable part of the visible spectrum was easily achieved by heteroatom substitution and minor expansion of the  $\pi$  electron system. The resulting derivatives **34–40** (Chart

18) were named for their emission maximum in complex with Pepper. Because these structural modifications were tolerated well in terms of complex stability and brightness, Pepper can be classified as a promiscuous aptamer.



**Chart 18.** Modifications to the HBC core were used for spectral tuning of fluorescent Pepper complexes.

To demonstrate the utility of such an aptamer, Pepper was used for sequential CRISPR/Cas-based multicolor imaging of mammalian cells. HeLa cells were engineered to express GFP-labeled dCas9 bearing a nuclear localization signal and a Pepper-tagged single guide RNA (sgRNA) targeting human centromeric satellite DNA. The modified dCas9 was imported into the nucleus, where it colocalized with the sgRNA construct on genomic loci. Green Pepper fluorescence was enabled by addition of HBC to the cell culture medium. Subsequent washing and restaining with six of the HBC derivatives shown in Chart 18 allowed deliberate changing of the fluorescence emission color, which would not be possible for direct CRISPR/Cas-based imaging with fluorescent proteins.<sup>110</sup>

Overall, the examples presented in this introduction show how fluorogenic aptamers have shaped the field of RNA detection and imaging just as fluorescent proteins have done for their domain of biomacromolecules. Their nature as smaller, less structurally invasive probes compared to earlier fluorescent reporters such as molecular beacons and the MS2 system make them an invaluable addition to the RNA imaging toolbox in their own right. Structure elucidation of aptamer–ligand complexes has also shed light onto the various kinds of interactions that allow RNA to confer functionality to small organic molecules. Nonetheless, the continued development of new fluoromodules is highly relevant for consolidating desirable properties that are still unique to certain aptamer families and for transferring other beneficial characteristics of fluorescent proteins, e.g. long Stokes shift emission, to fluorogenic aptamer complexes.

### 1.3 RESEARCH OBJECTIVES AND THESIS OUTLINE

The research presented herein comprises two overarching projects that aim at the development, characterization and implementation of novel tools for fluorescent labeling and imaging of RNA. These projects deal with the large Stokes shift fluorogenic aptamer Chili (described in sections 2.1–2.5) and the fluorescent nucleoside analog 4-cyanoindole riboside (described in sections 2.6–2.8).

The first part begins with an overview of how an optimized RNA aptamer, Chili, was previously engineered from the DMHBI-binding 13-2 aptamer in this laboratory. This is followed by an investigation into methods for the preparation of suitable ligands to improve on the previously used azlactone methodology. Characterization of new Chili–ligand complexes commenced with a screening assay to identify compounds with redshifted emission maxima or a higher fluorogenic potential than the parent ligand DMHBI. Promising derivatives were chosen for in-depth biophysical characterization of their binding behavior. Binding affinities were determined by fluorescence titration as well as isothermal titration calorimetry in order to extract additional thermodynamic parameters. Dynamic changes occurring during the binding event were elucidated using time-dependent changes in fluorescence emission. The underlying relationship between ligand binding and RNA folding, in particular a reorganization of the binding site, was observed by temperature-dependent UV/Vis spectroscopy. Mutagenesis experiments gave additional insight into the structure of the binding site itself and the influence of auxiliary base pairs on reorganization kinetics. Finally, a Chili-based sensor device is described, which was used to visualize the cleavage activity of a 10-23 DNA enzyme.

The second part describes the incorporation the universal fluorescent nucleobase analog 4-cyanoindole into RNA oligonucleotides. An unconventional synthetic pathway was devised to prepare the required phosphoramidite building block for solid-phase synthesis. Modified oligonucleotides were characterized with respect to their structural and photophysical properties. In the end, 4-cyanoindole was incorporated site-specifically at various positions throughout the Chili aptamer in order to enable Förster resonance energy transfer to bound ligands. Distance-orientation relationships between the modification site and the binding site, which will be a useful supplement to traditional structure elucidation methods, were generated based on steady-state and time-resolved FRET datasets.



## 2 RESULTS AND DISCUSSION

[PARTS OF THIS CHAPTER HAVE BEEN PUBLISHED]<sup>111–113</sup>

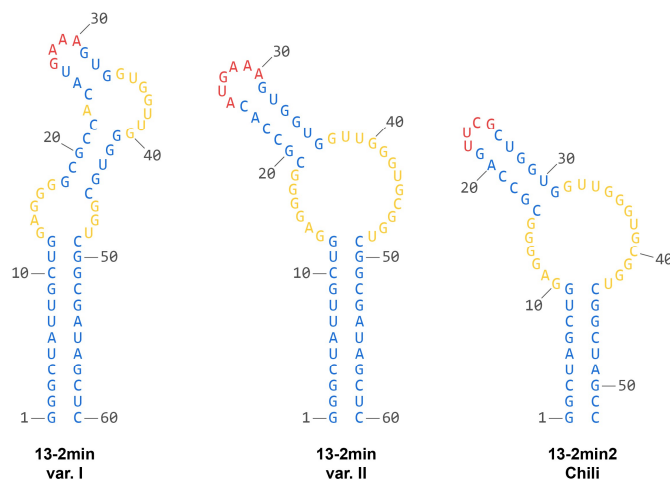
### 2.1 DESIGN OF THE CHILI APTAMER

#### 2.1.1 A DMHBI-binding aptamer

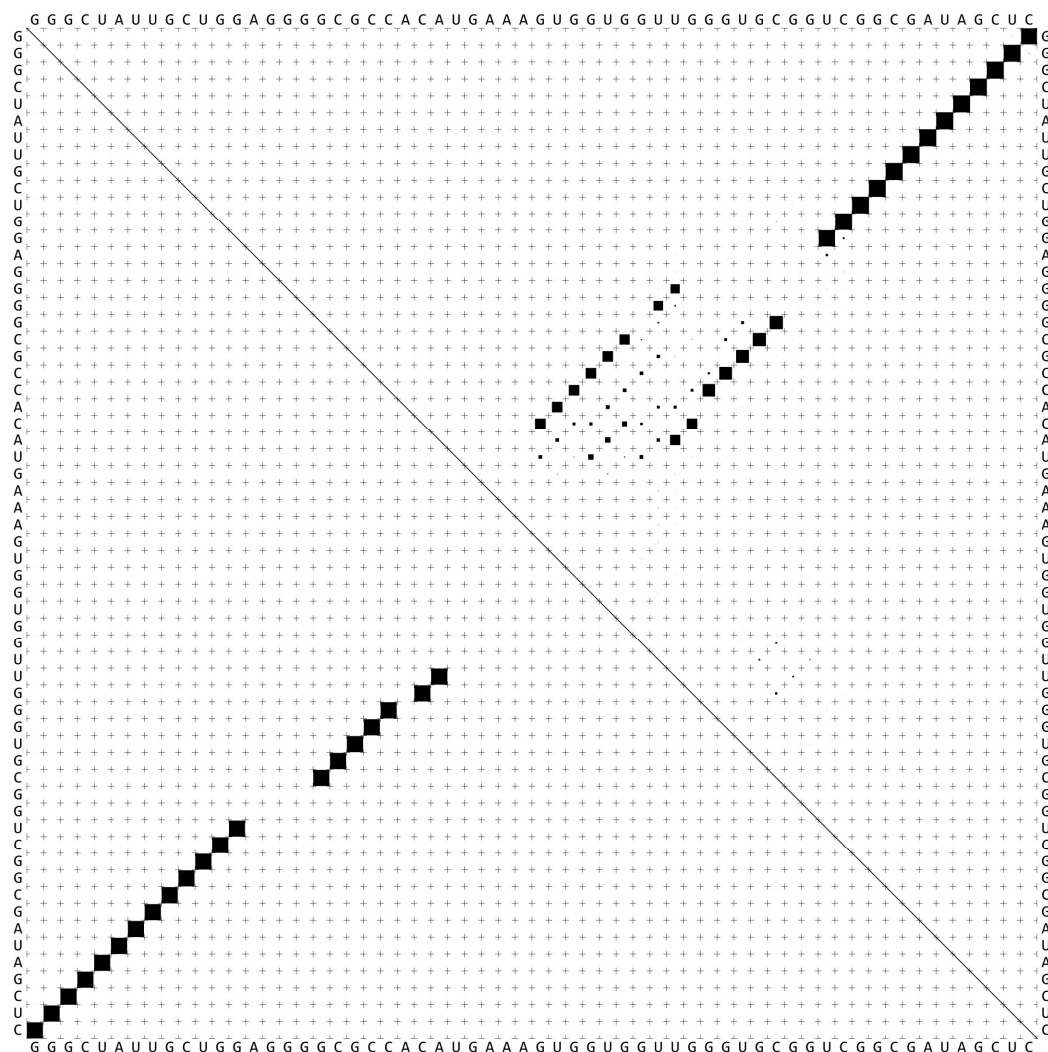
In their seminal 2011 publication, Jaffrey and coworkers selected various aptamers against the fluorescent-protein analog fluorogen DMHBI (**13**). Unlike the Spinach–DFHBI system, DMHBI complexes were not investigated in detail because their binding affinity and brightness were not considered promising enough for future applications. Intriguingly, however, various DMHBI complexes showed an indication of spectral tuning behavior, i.e. a dependence of the emission wavelength on the nature of the RNA aptamer, something the Spinach–DFHBI system was lacking at that time.<sup>71</sup> A well-functioning DMHBI-activating RNA aptamer therefore would be an important addition to the growing toolbox of genetically encodable fluorescent tags.

#### 2.1.2 Engineering Chili from 13-2min

13-2min, a 60 nt aptamer with the highest degree of fluorescence turn-on for DMHBI, was reported to preferentially adopt a global hairpin secondary structure containing two bulges (Chart 19, var. I).<sup>71</sup> However, secondary structure prediction (RNAfold web server, ViennaRNA web services<sup>114</sup>) suggested the folding landscape of 13-2min to be more complex, allowing for several alternative folding states (Figure 12). While the reported secondary structure occurs with low probability in this ensemble, other secondary structures featuring e.g. a basal stem and an apical stem-loop that flank a larger central bulge (Chart 19, var. II) are more abundant.



As discussed previously (see section 1.1.3), ligand recognition by RNA aptamers is based on highly specific intramolecular interactions, which require folding into a well-defined conformation. The conformational heterogeneity of 13-2min therefore could hamper ligand uptake because only a fraction of the RNA likely is available in a binding-competent conformation, whereas the remaining species are inactive.



**Figure 12.** Dot plot for the 13-2min aptamer displaying the secondary structure of the RNA as a  $60 \times 60$  matrix. Squares indicate the presence of a base pair between the respective nucleotides. The lower left section of the plot corresponds to the lowest-energy folding state, while the upper right section shows the ensemble of predicted folding states. The probability of base pair formation is proportional to the area of the square at that position. The reported structure of 13-2min (var. I in Chart 19) represents only a small fraction of the ensemble and several alternative folding patterns occur with a higher probability. The plot was generated using the RNAfold web server, ViennaRNA web services.<sup>114</sup>

In an effort to enhance the performance of 13-2min, our working group used a rational design approach to engineer an aptamer variant called 13-2min2 with the goal of reducing conformational heterogeneity and increasing the folding capacity of the RNA. Individual G–U wobble base pairs in the bottom stem were removed and the original loop region of 13-2min was replaced by a highly stable UUCG tetraloop (Table 1).

**Table 1.** Sequence alignment of 13-2min and 13-2min2 showing consensus regions (green), stems (blue), loops (red) and bulges (yellow).

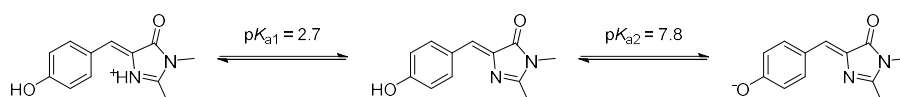
RNA	5'-sequence-3'					
	10	20	30	40	50	60
13-2min	GGGCUAUU	GCUGGAGGGGGCGCCA	CAUGAAAAG	JGGUGGUJUGGGUGCGGUCGGC	GAUAGCUC	
13-2min2	G-GCUA--	GCUGGAGGGGGCGCCA	G-UUCG-C	JGGUGGUJUGGGUGCGGUCGGC	--UAGC-C	
	[Blue]		[Yellow]	[Blue]	[Red]	[Blue]

Chemical probing of 13-2min2 by digestion with  $Tb^{3+}$  was performed before and after addition of DMHBI. The hydrolysis pattern confirmed the presence of a stem-loop element next to a large bulge (Chart 19). Nucleotides 34–38 within the bulge became more susceptible to hydrolytic cleavage upon ligand addition which implicates them in a binding-induced refolding of the RNA.<sup>111</sup> 13-2min2, nicknamed the Chili aptamer, displayed a substantial increase in fluorescence enhancement of DMHBI over 13-2min in vitro according to initial measurements and was the focus of the synthetic and analytic work presented within this thesis.

## 2.2 FUNCTIONAL HALLMARKS OF CHILI

### 2.2.1 Spectral characteristics

In aqueous environment, HBI derivatives can exist in a cationic, neutral and anionic protonation state (Scheme 7). The respective acid dissociation constants ( $pK_a$ ) of these forms are influenced by the nature of substituents attached to the benzylidene moiety, which can be exploited for generating a specific protonation state at a given pH. For example, while the parent compound HBI has a  $pK_a$  of 7.8, introduction of two electron-withdrawing fluorine substituents in DFHBI lowers the  $pK_a$  to 5.4, resulting in complete deprotonation at neutral pH.<sup>115</sup>



**Scheme 7.** Protonation states of HBI. Tautomeric forms and protonation sites correspond to experimentally observed species.<sup>115,116</sup>

DMHBI, bearing electron-donating methoxy substituents, has  $pK_a$  values of  $2.34 \pm 0.03$  and  $7.96 \pm 0.01$ , respectively, for the two protonation steps as determined by spectrophotometric titration in a universal buffer (Figure 13). Therefore, the ratio of phenol to phenolate species is approximately 3:1 under typical RNA aptamer binding conditions at pH 7.5 (Figure 14a). Excitation of the phenol at 400 nm results in very weak fluorescence emission ( $\phi = 6.6 \pm 0.1 \times 10^{-4}$ ) peaking at 485 nm with a shoulder at 537 nm, which arises due to increased acidity in the excited state (Figure 14b, Table 2, 3).<sup>117</sup> In the presence of excess Chili RNA, the absorption peak of the phenolate species is not observed, which indicates that the RNA exclusively binds the neutral ligand (Figure 14a). This characteristic sets Chili apart from the Spinach aptamer, which was designed to bind an anionic ligand and thereby mimic the characteristics of enhanced GFP. Excitation of Chili–DMHBI at 400 nm results in 100-fold fluorescence enhancement ( $\phi = 7.6 \pm 0.1 \times 10^{-2}$ ) over DMHBI alone with a single emission maximum at 537 nm.

**Table 2.** Spectral characteristics of DMHBI alone and in complex with the Chili aptamer.

Fluorophore	$\lambda_{ex}$ nm	$\lambda_{em}$ nm	Stokes shift nm	$\epsilon^{[a]}$ $L mol^{-1} cm^{-1}$	$\phi_{rel}^{[b]}$	Brightness <sup>[c]</sup>
DMHBI <sup>[d]</sup>	378	485	107	19500	$6.6 \pm 0.1 \times 10^{-4}$	—
	479	537	58	28600		
Chili–DMHBI	400	537	137	18000	$7.6 \pm 0.1 \times 10^{-2}$	1400

[a] Extinction coefficients measured from absorption spectra of a dilution series (ligand only) or those used for quantum yield determination (RNA–ligand complex). Details are given in section 4.1.9.

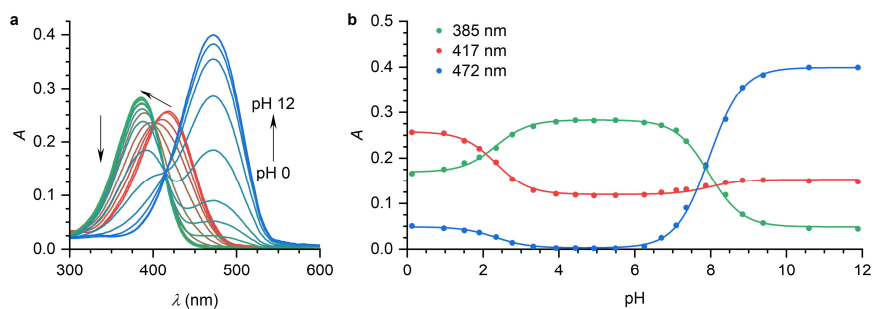
[b] Relative quantum yield measured using coumarin 153 as a reference compound. Details are given in section 4.1.12.

[c] Brightness =  $\phi \times \epsilon \times L^{-1} mol$  cm.

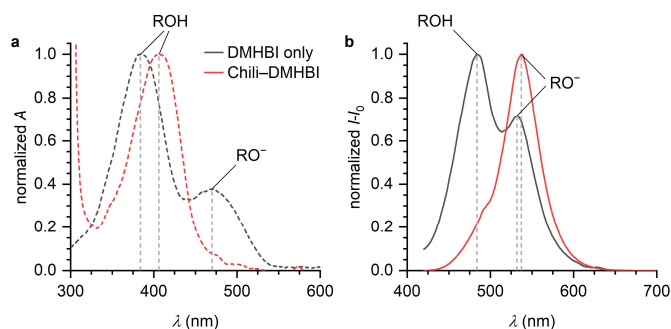
[d] Values are reported for the neutral (top, at pH 4.9) and deprotonated (bottom, at pH 11.9) species.

The large pseudo-Stokes shift of more than 130 nm arises from virtually complete deprotonation of the ligand prior to fluorescence emission (Figure 14b). As this degree of proton loss is not observed for the unbound ligand, one or more nucleotides in the binding site must be involved in the deprotonation. This RNA-mediated excited state proton transfer (Scheme 8) distinguishes Chili from other fluorogenic aptamer systems, which typically exhibit Stokes shifts well below 80 nm.<sup>45,91–94</sup>





**Figure 13.** a) UV/Vis spectra of DMHBI in aqueous buffer. Spectra were measured at 21 pH levels between 0 and 12. b) Titration curves of DMHBI at the absorption maxima of the cationic (red), neutral (green) and anionic (blue) species.  $pK_a$  values were obtained by global fitting of the absorbance with the Henderson-Hasselbalch equation. Details are given in section 4.1.12. Conditions: 11  $\mu\text{M}$  ligand, 30 mM Britton-Robinson buffer.



**Figure 14.** a) Normalized absorption spectra of DMHBI alone and in complex with the Chili aptamer. The RNA exclusively binds to neutral DMHBI. b) Normalized emission spectra of DMHBI alone and in complex with the Chili aptamer. Excited state proton abstraction from the phenol moiety is significantly enhanced in the RNA-bound state. Conditions: 17  $\mu\text{M}$  RNA, 10  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Fluorescence spectra were measured with  $\lambda_{\text{ex}} = 400$  nm.

**Table 3.** Absorption/emission maxima and ground state/excited state  $pK_a$  for three HBI derivatives. Values for HBI and DFHBI were taken from the literature.<sup>115</sup>

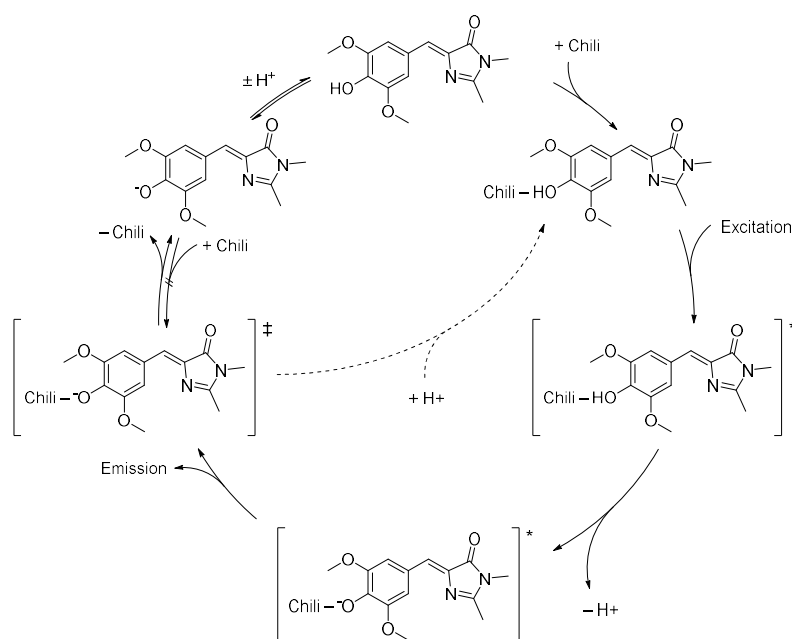
Compound	$\lambda_a^{[a]}$ nm	$\lambda_{\text{em}}^{[a]}$ nm	$pK_a^{[b]}$	$pK_a^{*[c]}$
HBI	370	448	7.8	-0.6
	425	494		
DFHBI	363	445	5.4	-4.0
	418	500		
DMHBI	385	485	$7.96 \pm 0.01$	-0.3
	470	537		

[a] Values for the phenol (top) and phenolate (bottom) species.

[b] Only the  $pK_a$  for the phenol-phenolate equilibrium is listed.

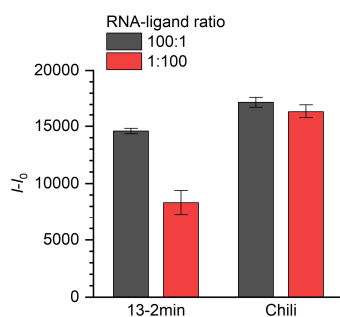
[c] Calculated using the Förster cycle:<sup>117</sup>

$$pK_a^* = pK_a - \frac{hc}{k_B T \ln 10} \frac{(\tilde{\nu}_{\text{abs, neutral}} - \tilde{\nu}_{\text{fl, neutral}}) + (\tilde{\nu}_{\text{abs, anion}} - \tilde{\nu}_{\text{fl, anion}})}{2}$$



### 2.2.2 Folding state

Chili was engineered to promote folding into a single conformer that can sequester the ligand more readily than the 13-2min aptamer. The structural homogeneity of both RNAs was assessed by means of a folding assay, in which RNA and DMHBI were first incubated in a 100:1 ratio. Even if the RNA is not fully folded, adding it in large excess ensures formation of 1 equivalent of the respective fluorescent complex. Next, RNA and DMHBI were incubated in a 1:100 ratio. If any inactive RNA species were present, less than 1 equivalent of the complex would be formed, resulting in a diminished fluorescence brightness. The proportion of correctly folded RNA is then given by the ratio of both brightness values.

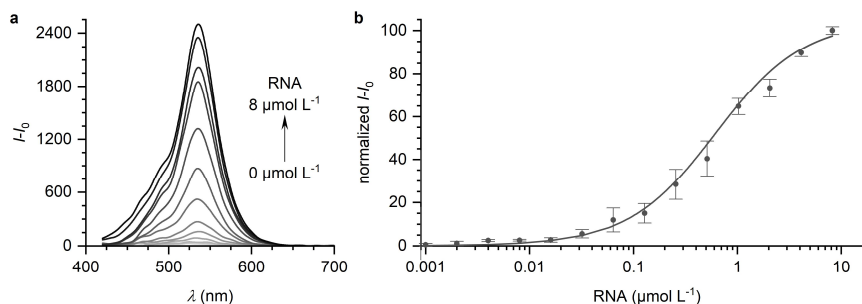


**Figure 15.** Folding assay for 13-2min and the Chili aptamer using DMHBI. Chili is less prone to misfolding than 13-2min and displays a higher intrinsic fluorescence enhancement. Conditions: 10  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand (black) or 0.1  $\mu\text{M}$  RNA, 10  $\mu\text{M}$  ligand (red), 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ) were integrated after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

57% of the 13-2min aptamer were found to adopt an active conformation. Therefore, 13-2min behaves similar to the improved Spinach2 aptamer.<sup>78</sup> For Chili, however, the same degree of fluorescence enhancement was observed irrespective of the RNA/ligand ratio, i.e. this aptamer shows vastly improved folding behavior (Figure 15, Table 4). While folding is the principal contributor to the increased fluorescence enhancement with Chili, this assay also shows that the intrinsic brightness of Chili–DMHBI is 18% higher than that of the 13-2min complex.

The binding affinity of the aptamer towards DMHBI remained largely unaffected by the improved folding behavior. The dissociation constant ( $K_d$ ) was determined by titrating RNA into solutions containing a fixed concentration of the ligand (Figure 16a). Data analysis was performed by fitting to a single-site binding model (Figure 16b). This approach was preferred to the commonly used method of titrating ligand into RNA and fitting with a simple hyperbola, as it does not result in

unspecific fluorescence enhancement at high ligand concentrations and more accurately captures the concentration changes that occur during the titration (see section 4.1.14). The resulting  $K_d$  of  $5.7 \pm 1.2 \times 10^{-7} \text{ mol L}^{-1}$  is similar to that of 13-2min–DMHBI ( $K_d = 4.6 \times 10^{-7} \text{ mol L}^{-1}$ ) and is on the same order of magnitude as that of DFHBI binding to Spinach or the folding-optimized variant Spinach2 (Table 4).



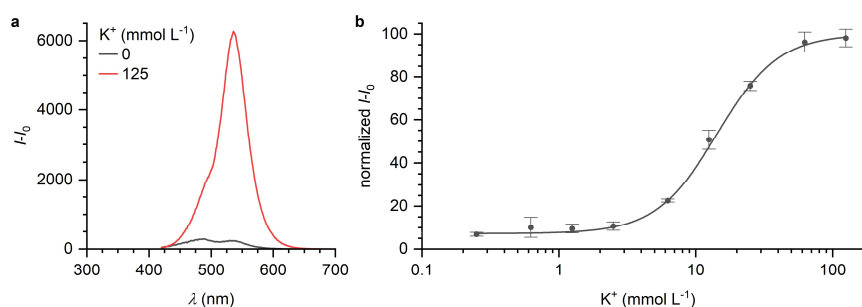
**Figure 16.** a) Addition of the Chili aptamer to DMHBI results in fluorescence enhancement. b) Titration of DMHBI with Chili. Data points were fitted with a single-site binding model.  $K_d = 5.7 \pm 1.2 \times 10^{-7} \text{ mol L}^{-1}$ . Conditions: 0–8.2  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

**Table 4.** Results of the folding assay and binding affinity measurement for the 13-2min and Chili aptamers in complex with DMHBI. For comparison, published values for the Spinach and Spinach2 aptamers are reported as well.<sup>78</sup>

Ligand	RNA	Degree of folding	$K_d$ $\text{mol L}^{-1}$
DMHBI	13-2min	$57 \pm 7\%$	$4.6 \times 10^{-7}$ <sup>[a]</sup>
	Chili	$95 \pm 4\%$	$5.7 \pm 1.2 \times 10^{-7}$
DFHBI	Spinach	$32 \pm 4\%$	$4.5 \pm 0.1 \times 10^{-7}$ <sup>[a]</sup>
	Spinach2	$58 \pm 5\%$	$4.3 \pm 0.2 \times 10^{-7}$ <sup>[a]</sup>

[a] Reported for titrating ligand into RNA.<sup>71</sup>

Chili contains a high number of consecutive guanines in its bulge region, which suggests that a G-quadruplex might be involved in its function. G-quadruplexes have emerged as a privileged tertiary structure motifs in fluorogenic aptamers and were identified as part of the binding sites of Spinach aptamers,<sup>81,82,118</sup> Corn,<sup>89</sup> and several Mango variants<sup>96,100,101</sup> by X-ray crystallography. G-quadruplexes comprise multiple layers of cyclic tetrads formed by Hoogsteen-bonded guanine residues. Metal cations intercalate within or between tetrads and are required to properly stabilize the quadruplex. While several group I and II cations are capable of inducing quadruplex formation,  $\text{Na}^+$  and  $\text{K}^+$  are not only the most physiologically relevant but also the most effective ones.<sup>119</sup> In the absence of  $\text{K}^+$ , the Chili aptamer does not induce fluorescence in DMHBI (Figure 17a).



**Figure 17.** a)  $\text{K}^+$  ions are required to observe fluorescence of the Chili–DMHBI complex. b) Titration of Chili–DMHBI with KCl. Data points were fitted with a Hill model.  $K_d = 1.2 \pm 0.1 \times 10^{-2} \text{ mol L}^{-1}$ ,  $n = 2.5 \pm 0.4$ . Conditions: 0.5  $\mu\text{M}$  RNA,

0.5  $\mu\text{M}$  ligand, 0–125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Titration of Chili–DMHBI with KCl results in a steady increase of fluorescence enhancement up to a concentration of 100 mM  $\text{K}^+$  with an apparent  $K_d$  of  $1.2 \pm 0.1 \times 10^{-2} \text{ mol L}^{-1}$  and a Hill coefficient of  $2.5 \pm 0.4$  (Figure 17b). Though no indication for the actual number of bound metal ions, this magnitude of cooperativity has been associated with the formation of two-tiered G-quadruplexes.<sup>120</sup> The apparent  $\text{K}^+$  affinity is similar to that of Spinach–DFHBI.<sup>121,122</sup> Detailed studies of G-quadruplex formation are described in section 2.4.6.

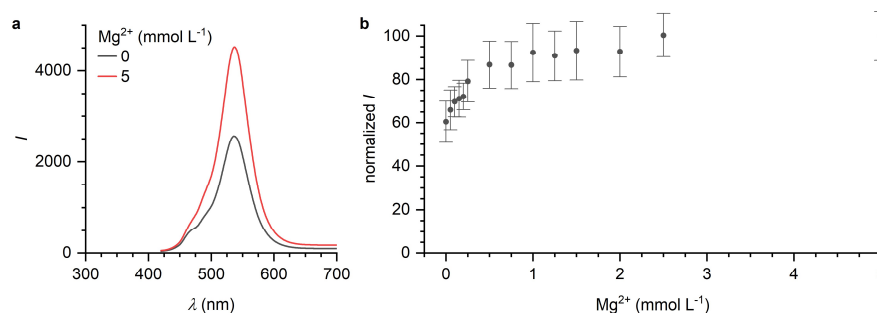
**Table 5.** Apparent  $\text{K}^+$  affinities of Chili, Spinach and Broccoli complexes obtained from fitting titration data with a Hill model.

Ligand	RNA	$K_d$ $\text{mol L}^{-1}$	Hill coefficient
DMHBI	Chili	$1.2 \pm 0.1 \times 10^{-2}$	$2.5 \pm 0.4$
DFHBI <sup>[a]</sup>	Spinach	$8.5 \pm 0.4 \times 10^{-3}$	$1.6 \pm 0.1$
	Broccoli	$1.4 \pm 0.1 \times 10^{-3}$	$1.7 \pm 0.2$

[a] Values reported by Shinde and coworkers.<sup>121,122</sup>

In cellular environments, the global tertiary structure of RNA molecules is stabilized by  $\text{Mg}^{2+}$  ions, which are partially chelated by the negatively charged phosphate backbone, resulting in decreased net charge. They are also present throughout the environment, which leads to a dielectric screening effect.<sup>123–125</sup> They are freely available at a concentration of less than 1  $\text{mmol L}^{-1}$ .<sup>126</sup> For prospective live cell applications, it is therefore highly desirable to utilize fluorogenic aptamers that do not require high concentrations of  $\text{Mg}^{2+}$  for proper function.

Even though the 13-2 aptamer was selected in the presence of 5  $\text{mmol L}^{-1}$   $\text{Mg}^{2+}$ , this does not reflect the requirements of Chili in complex with DMHBI, which already exhibits half-maximal fluorescence enhancement in  $\text{Mg}^{2+}$ -free buffer (Figure 18a). Maximal fluorescence is reached at a concentration of approximately 0.5–1  $\text{mmol L}^{-1}$   $\text{Mg}^{2+}$  (Figure 18b). In comparison, the Broccoli aptamer, which was selected under low- $\text{Mg}^{2+}$  conditions, performs optimally at approximately 0.3  $\text{mmol L}^{-1}$   $\text{Mg}^{2+}$  (Table 6).<sup>79</sup>



**Figure 18.** a) In the absence of  $\text{Mg}^{2+}$ , Chili–DMHBI exhibits approximately half of its maximal fluorescence enhancement. b) Titration of Chili–DMHBI with  $\text{MgCl}_2$ . 1  $\text{mM}$   $\text{Mg}^{2+}$  is sufficient to induce optimal fluorescence enhancement. Conditions: 0.5  $\mu\text{M}$  RNA, 5  $\mu\text{M}$  DMHBI, 125 mM KCl, 0–5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ) were integrated and normalized. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

**Table 6.** Magnesium requirements of the Chili–DMHBI complex in comparison with the Spinach2 and Broccoli complexes of DFHBI-1T.<sup>79</sup> The Broccoli aptamer has been optimized to work at low-magnesium concentrations.

Ligand	RNA	Mg <sup>2+</sup> for $I_{\max}$ mmol L <sup>-1</sup>	residual fluorescence w/o Mg <sup>2+</sup>
DMHBI	Chili	0.5	50%
DFHBI-1T	Spinach2	1.0	11%
	Broccoli	0.3	61%

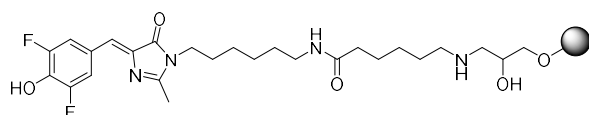
Overall, the combination of large Stokes shift, well defined folding behavior and low Mg<sup>2+</sup> requirement makes Chili a prime candidate for the development of a broadly useful fluorogenic aptamer platform. The following chapter details the steps taken to enhance ligand affinity and brightness of the Chili system.

### 2.3 OPTIMIZATION OF THE CHILI SYSTEM

A well-performing fluorogenic aptamer system is characterized by three key features: 1) Proper folding of the RNA to reliably provide a binding site for the fluorophore under various conditions. This is particularly important with respect to live cell imaging studies of low abundance transcripts, where brightness is strongly limited by the low copy number of aptamers, or for the construction of biosensors consisting of a receptor module fused to an aptamer at a defined entry point. While structural stability can be further enhanced by incorporating the aptamer in a larger scaffold such as a tRNA or three-way-junction, predictable folding *in vitro* is a prerequisite for designing such constructs. 2) Intrinsically bright emission due to strong absorption of excitation light in combination with a large fluorescence quantum yield. Even though the brightness of a target RNA can be increased by simply utilizing multiple copies of the aptamer, an increased probe size incurs the risk of impeding the natural function and interactions of the transcript. Additionally, with the ongoing development of aptamers for super-resolution imaging, a small footprint of the probe is required to achieve high spatial resolution. 3) Tight association between the fluorogen and the RNA to yield the highest possible level of fluorescent species. While complex formation can be forced by increasing the fluorophore concentration, this might result in higher background fluorescence due to non-specific activation.

While Chili satisfies the first of these requirements well, the brightness and binding strength of its DMHBI complex are modest. Systematic mutagenesis studies are routinely used to improve and miniaturize fluorogenic aptamers not only after *in vitro* selection of an initial clone, but also to generate variants that are optimized towards a specific goal such as the superfolding Spinach2 aptamer.<sup>78</sup>

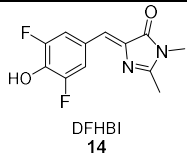
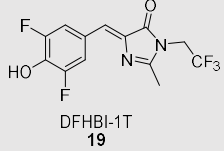
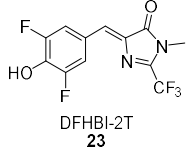
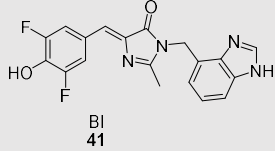
In the past, modification of the ligand has been exploited far less systematically than RNA mutagenesis for the optimization of aptamer systems. This stems from the fact that each aptamer is selected to bind a well-defined target molecule with high selectivity. A certain degree of structural flexibility, however, is present by design, as during the usual SELEX procedure the prospective ligand is covalently attached to an affinity matrix via a flexible linker. In case of HBI derivatives, the linker is most conveniently introduced at N3 of the imidazolone ring (Figure 19).



**Figure 19.** Structure of DFHBI attached to a sepharose matrix via an aminohexyl linker on imidazolone N3. The attachment site can be used as an entry point for chemical modifications.

Jaffrey and coworkers showed that different imidazolone N3 substituents not only influence the binding affinity of DFHBI (**14**) towards Spinach variants and the Broccoli aptamer, but also give rise to spectral shifts and brightness alterations for the respective ligand. Moving the modification site to the adjacent C2, however, was detrimental to both affinity and brightness. Initial studies resulted in two promising derivatives, DFHBI-1T (**19**) and DFHBI-2T (**23**).<sup>83</sup>

**Table 7.** Spectral properties of Spinach2 and Broccoli in complex with DFHBI derivatives. Reported values are sometimes inconsistent between different sources and were therefore taken from the initial publication describing each system.

Ligand	RNA	$\lambda_{\text{ex}}$ nm	$\lambda_{\text{em}}$ nm	$K_{\text{d}}$ nmol L <sup>-1</sup>	Brightness <sup>[a]</sup>
 DFHBI 14	Spinach2	447	501	530	15840
	Broccoli	— <sup>[b]</sup>	—	—	—
 DFHBI-1T 19	Spinach2	482	505	560	29140
	Broccoli	470	505	305	27824
 DFHBI-2T 23	Spinach2	500	523	1300	3480
	Broccoli	—	—	—	—
 BI 41	Spinach2	—	—	—	—
	Broccoli	470	505	51	22478

[a] Calculated as  $\phi \times \epsilon$  from reported values.<sup>79,83,127</sup>

[b] Value not reported.

Recently, an *in silico* screening approach was implemented based on molecular docking of DFHBI derivatives to the binding pocket of Spinach. The highest scoring candidate, BI (**41**), binds to the closely related Broccoli aptamer with high affinity to form a brightly fluorescent complex. In addition, it confers increased photobleaching resistance to the complex, which is attributed to a low unbinding rate after excitation (Table 7).<sup>127</sup>

A major focus of this thesis was the design and characterization of novel ligands for Chili using an empirical optimization approach to thoroughly explore the chemical space of the fluorophore binding site (section 2.3.1). Aspects under consideration were the influence of steric bulk imposed by the imidazolone-N3 substituent, tuning of  $\pi$ - $\pi$  stacking interactions by incorporation of aromatic side chains at that position, spectral shifts caused by substituents on the benzylidene moiety and tolerance for red fluorescent protein-derived fluorophore comprising a larger conjugated system.

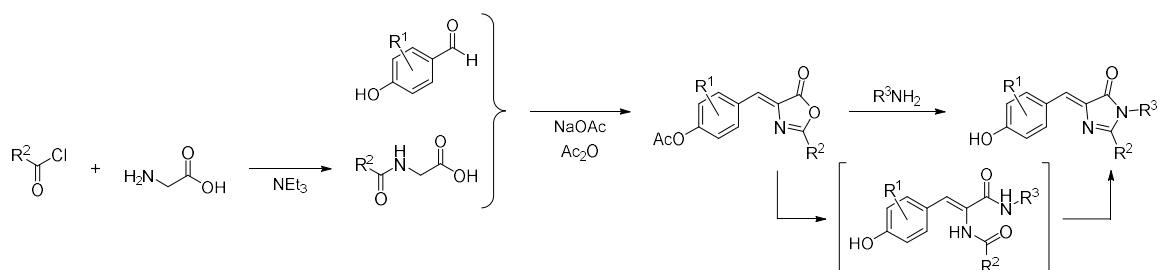
Since Chili was created by rational engineering of 13-2min, which itself was obtained by truncation and mutation of the 13-2 aptamer, systematic mutagenesis studies were focused to the ligand binding site; additional truncation experiments were performed as well (section 2.3.2).

### 2.3.1 Screening of HBI ligands

#### 2.3.1.1 Synthetic pathways

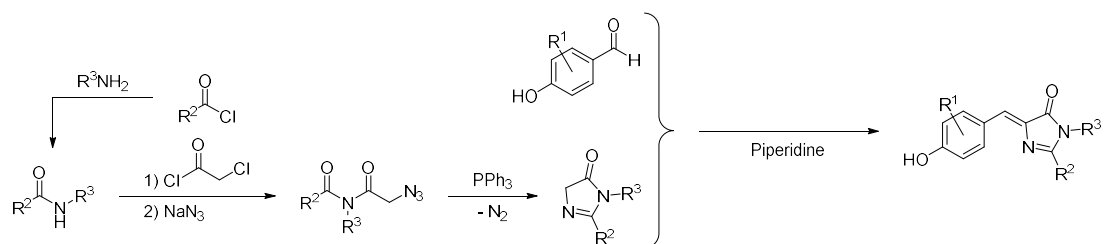
HBI derivatives are commonly prepared by aminolysis of an oxazolone (azlactone) intermediate, which is obtained via Erlenmeyer azlactone synthesis. This reaction is conceptually related to the Perkin reaction; N-acylation of glycine is followed by base-catalyzed condensation with an aromatic aldehyde. Most often, *N*-acetyl- or *N*-benzoylglycine are employed, which introduce a methyl or phenyl group at C2 of the five-membered ring, respectively (Scheme 9). A major drawback of this approach lies in the conversion from oxazolone to imidazolone, as the bisamide

intermediate does not readily cyclize if the amine bears an electron-deficient substituent such as an aryl group.



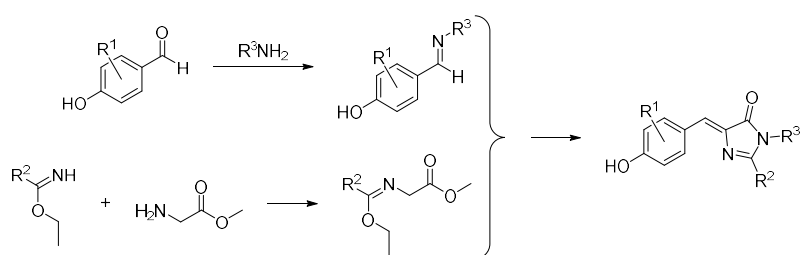
**Scheme 9.** General reaction scheme for the preparation of HBI derivatives by aminolysis of oxazolones. The aminolysis proceeds in separate ring-opening and -closing steps.

Alternatively, Knoevenagel condensation between an aromatic aldehyde and a preformed imidazolone, introduced by Burgess and coworkers in 2008,<sup>128</sup> directly yields HBI derivatives. C2/N3-substituted imidazolones are obtained in four steps by acylation of the respective amine and introduction of an azidoacetyl moiety followed by a tandem Staudinger reaction/imine formation (Scheme 10). While even aromatic amines are easily acylated in the first step, this route is overall not well-suited for the synthesis of a diverse compound library as both the C2- and N3-substituent are already introduced in the first step.



**Scheme 10.** General reaction scheme for the preparation of HBI derivatives by Knoevenagel condensation.

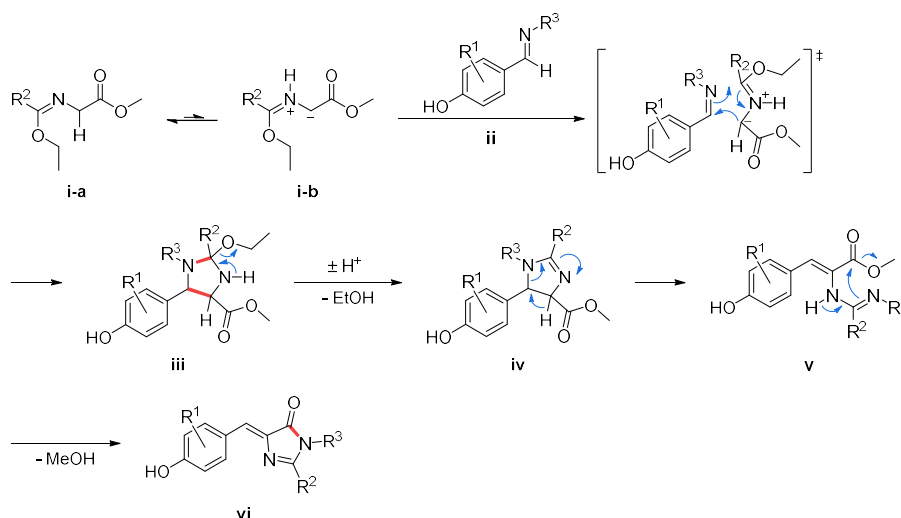
In 1993, Lerestif et al. first described an 1,3-dipolar cycloaddition between imines and imidates derived from  $\alpha$ -amino acid esters, which cleanly yielded a range of imidazolone derivatives (Scheme 11).<sup>129</sup> Later studies expanded the scope of this reaction and investigated the influence of microwave irradiation and acid catalysis.<sup>130,131</sup>



**Scheme 11.** General reaction scheme for the preparation of HBI derivatives by cycloaddition.

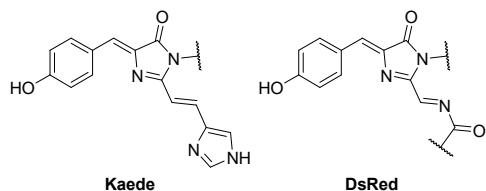
The imidate ester transiently forms an azomethine ylide, which behaves as a 1,3-dipole and reacts with the imine to form an imidazoline intermediate. Loss of ethanol initiates a rearrangement that results in the formation of the respective imidazolone. Formation of the new exocyclic double bond occurs with excellent *Z* selectivity (Scheme 12).





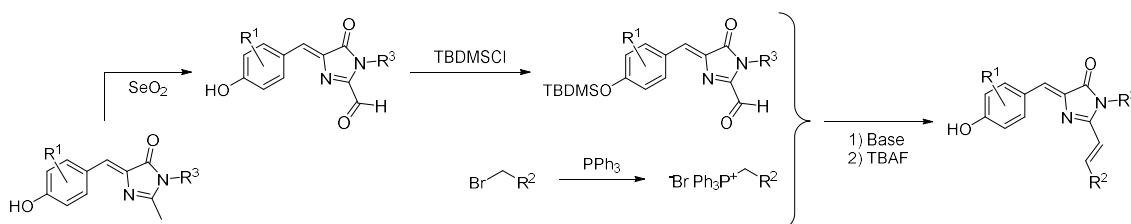
**Scheme 12.** Reaction mechanism for the cycloaddition between an imidate **i-a** and an imine **ii**. The 1,3-dipolar azomethine ylide tautomer **i-b** reacts with **ii** to form an imidazolone **iii**. Loss of EtOH leads to the unsaturated intermediate **iv**, which rearranges via **v** to form an HBI derivative **vi**.

The general utility of the cycloaddition was impressively demonstrated when Baldrige et al. reported a combinatorial library of 57 arylidene imidazolone derivatives prepared by this method.<sup>132</sup> Several follow-up studies by the same group also made extensive use of this methodology, including the direct installation of a limited number of different alkylidene and arylidene groups at imidazolone-C2.<sup>133–135</sup> Substituents at imidazolone-N3, however, were limited to simple aliphatic groups for all reported compounds. Nevertheless, due to its broad scope the cycloaddition reaction was used extensively during this thesis to prepare the majority of imidazolone-C2-alkyl HBI ligands.



**Chart 20.** The chromophores of the red fluorescent proteins Kaede and DsRed differ by the nature of their exocyclic double bond.

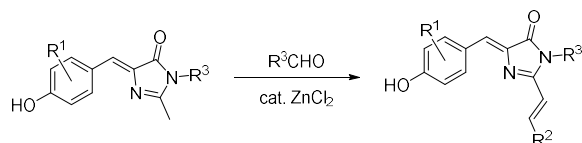
Based on the parent red fluorescent protein chromophores,  $\pi$ -extended HBI derivatives are categorized as either Kaede-like or DsRed-like, the difference being the presence of an additional C–C or C–N double bond, respectively (Chart 20).<sup>136</sup> Initial reports of synthetic Kaede-type chromophores were based around the azlactone strategy.<sup>137,138</sup> A novel approach was published in 2008 by Lukyanov and coworkers, who subjected C2-methyl imidazolones to a Riley oxidation followed by a Wittig reaction. Formation of the new double bond proceeded with high *E* selectivity despite using semi-stable phosphonium ylides in the absence of  $\text{Li}^+$  (Scheme 13).<sup>139</sup>



**Scheme 13.** General reaction scheme for the preparation of  $\pi$ -extended Kaede-like HBI derivatives by Wittig reaction.

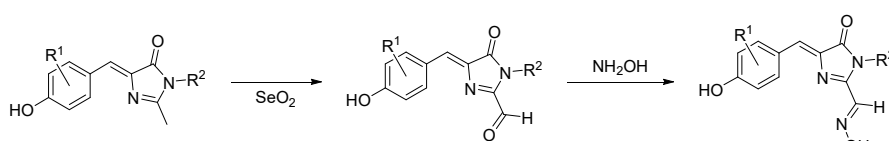
Shortly thereafter, it was demonstrated that C2-methyl imidazolones can directly undergo aldol condensation with a variety of aromatic and heteroaromatic aldehydes. Among five different

Lewis acids,  $\text{ZnCl}_2$  was found to be the most effective catalyst for this transformation (Scheme 14).<sup>140</sup> This methodology was later extended to borylated HBI derivatives, which readily underwent aldol condensation in the presence of piperidine as a catalytic base.<sup>141</sup>



**Scheme 14.** General reaction scheme for the preparation of  $\pi$ -extended Kaede-like HBI derivatives by Aldol condensation.

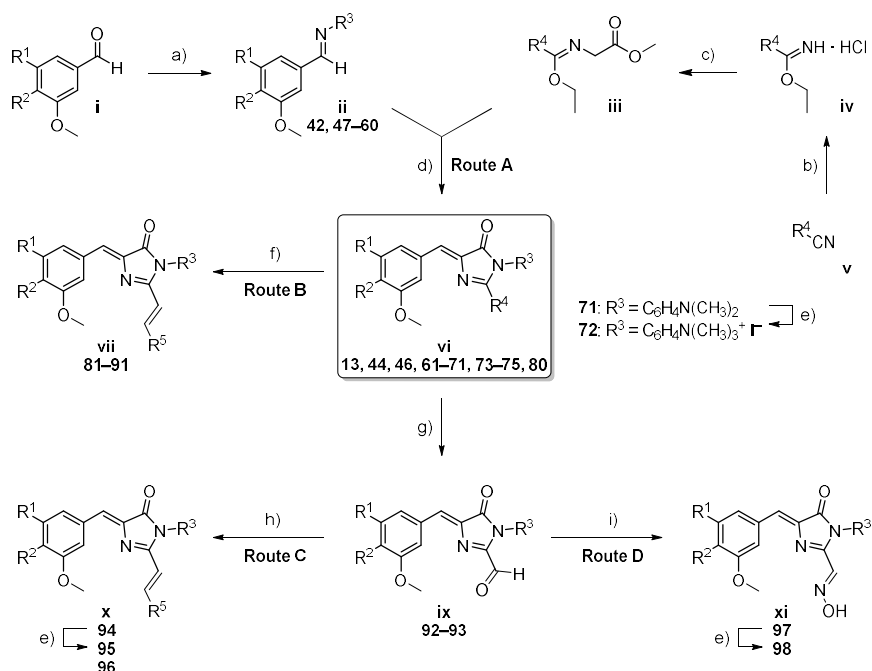
To date, there are no examples of actual synthetic DsRed-type chromophores due to the high sensitivity of acylimines towards hydrolytic degradation. By using a significantly more stable but structurally similar oxime moiety instead, Jaffrey and coworkers were able to synthesize DFHO as a ligand for the Corn aptamer (Scheme 15).<sup>142</sup> In accordance with their different substrate scopes, all three of the strategies outlined above were used in this thesis to generate a variety of Kaede- and DsRed-type HBI derivatives.



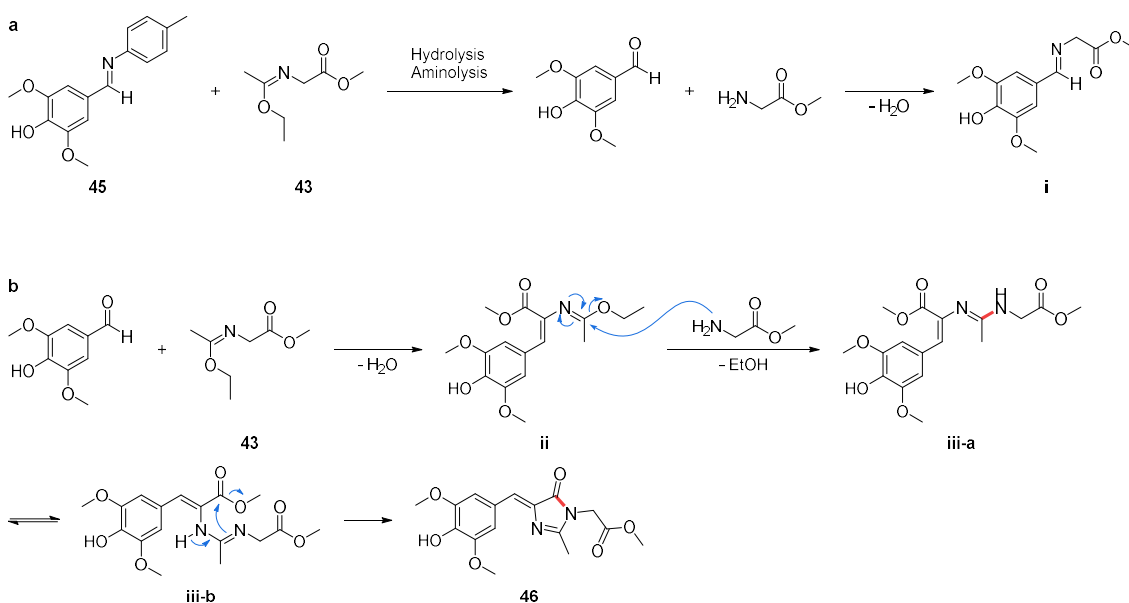
**Scheme 15.** General reaction scheme for the preparation of  $\pi$ -extended DsRed-analogous HBI oximes.

### 2.3.1.2 Preparation of a modified ligand library

In this thesis, the cycloaddition approach (Scheme 16, Route A) was first utilized for the synthesis of DMHBI (**13**) from the precursor imine (**42**) and imidate ester (**43**). Compound (**42**) was readily prepared from 4-hydroxy-3,5-dimethoxybenzaldehyde and a tenfold excess of methylamine. The imine was isolated in pure form by filtration and removal of the volatiles (Table 8). Following a reported procedure that improves on the original protocol, **43** was obtained from glycine methyl ester hydrochloride and ethyl acetimidate hydrochloride with a yield of 85%. The subsequent cycloaddition proceeded cleanly at ambient temperature using ethanol as a solvent. In contrast to the azlactone approach, DMHBI was obtained with a yield of 82% without requiring additional purification of the crude material. As a first modification to the core structure of DMHBI, a phenyl spacer was inserted into the imidazolone-N3-methyl moiety to generate DMHBTI (**44**). To synthesize the corresponding imine **45**, a stoichiometric mixture of the parent aldehyde and *para*-toluidine was heated to reflux in toluene using a Dean-Stark trap, thereby removing the water produced during the reaction. Evaporation of the solvent then yielded the imine, which was sufficiently pure for conversion to the HBI derivative (Table 8). In this case, the cycloaddition proceeded slower than with imine **42**. After 84 h at ambient temperature, the target compound was isolated from the reaction mixture with a yield of 54% alongside a major side product (DMHBTI-spdt, **46**), which was separated by column chromatography and characterized by NMR spectroscopy as well as ESI MS. The formation of such a side product had already been observed during the original development of the imine-imidate cycloaddition and was attributed either to transimination between the reactants or hydrolysis of the imine by residual water (Scheme 17).<sup>130</sup>



**Scheme 16.** Synthetic pathways towards modified HBI ligands (Route A) and derivatives with extended conjugated system (Routes B–D). Exemplary reagents and conditions: a)  $\text{R}^3\text{NH}_2$ ,  $\text{MgSO}_4$ , r.t., 24 h; or  $\text{R}^3\text{NH}_2$ , toluene, reflux, 16 h. b)  $\text{AcCl}$ , EtOH,  $0^\circ\text{C}$ , 7 h. c) Glycine methyl ester hydrochloride,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ , r.t., 3 h. d) EtOH, r.t., 16 h; or toluene,  $120^\circ\text{C}$ , 16 h. e) MeI, DMF, r.t., 24 h. f)  $\text{R}^5\text{CHO}$ , cat.  $\text{Sc}(\text{OTf})_3$ , dioxane,  $110^\circ\text{C}$ , 24 h. g)  $\text{SeO}_2$ , dioxane, reflux, 2 h. h)  $\text{R}^5\text{CH}_2\text{PPh}_3\text{Br}$ ,  $n\text{BuLi}$ ,  $0^\circ\text{C}$ , 30 min to r.t., 16 h. i)  $\text{NH}_2\text{OH} \cdot \text{HCl}$ ,  $\text{K}_2\text{CO}_3$ , MeOH, r.t., 24 h.

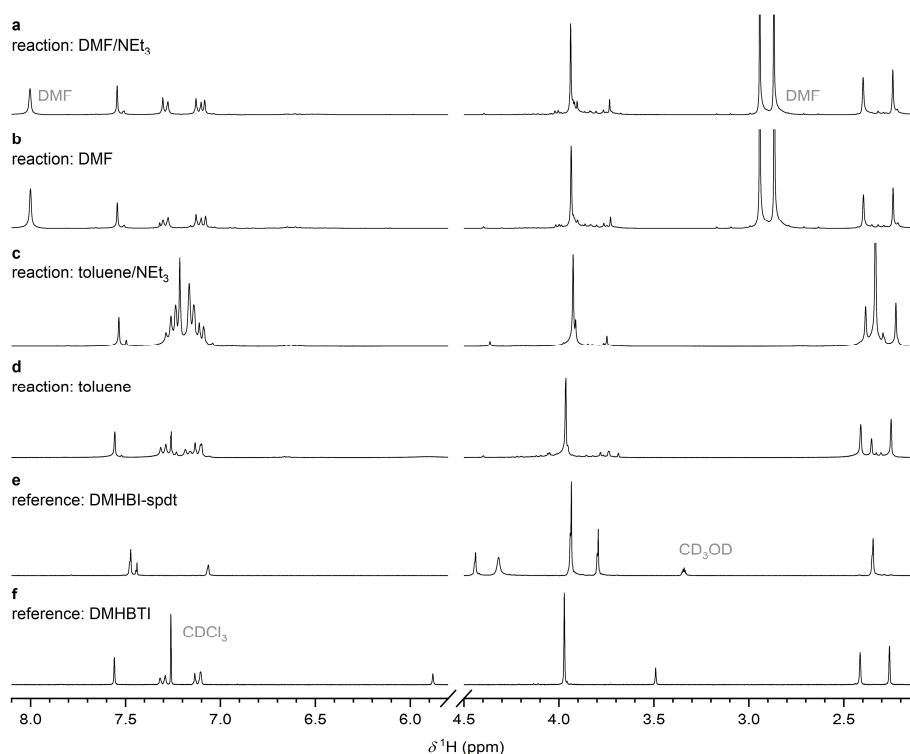


**Scheme 17.** a) Hydrolysis of imine **45** and imidate **43** by residual water would generate the respective aldehyde and amine. These could then react to form a new imine (**i**), which would then undergo cycloaddition as usual, resulting in the formation of the side product **46**. The presence of free amines in the reaction mixture would further facilitate decomposition of the starting materials. b) A second conceivable pathway following decomposition is a condensation between the aldehyde and **43** to form intermediate **ii**. Aminolysis of the imidate group by glycine methyl ester produces the amidine **iii**, which would undergo ring closure to form the side product **46**. Note that this reaction is identical to the final step depicted in scheme Scheme 12.

To improve the rate and selectivity of the cycloaddition, two additional solvents, toluene and DMF, were screened at elevated temperature. Triethylamine was included as an additive for base catalysis.

Solvent	Temperature °C	Additive	Time <sup>[a]</sup> h
Toluene	120	—	16
		5% NEt <sub>3</sub>	< 6
DMF	150	—	< 6
		5% NEt <sub>3</sub>	< 6

Test reactions were run until TLC indicated complete consumption of the starting material. Afterwards, the solvent was removed and the crude product was analyzed by <sup>1</sup>H NMR. In all cases, the side reaction did not take place and only DMHBTI was formed. The addition of triethylamine substantially accelerated the reaction in toluene but had no discernable effect otherwise. The average yield of the four test reactions was 83%. For subsequent syntheses of DMHBTI derivatives, toluene was preferred over DMF due to its higher volatility.



**Figure 20.** <sup>1</sup>H NMR spectra of the crude reaction mixtures described above (a–d) and the pure reference compounds (e, f) in CDCl<sub>3</sub>, 300 MHz. The DMHBTI-spdt sample contains CD<sub>3</sub>OD to increase solubility. Formation of the side product was negligible in all four trial reactions, whereas the reaction in EtOH yielded product and side product in a ratio of 2.3 to 1.

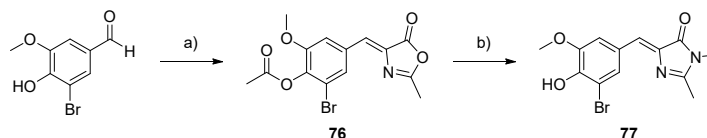
With an optimized protocol in hand, synthesis of a ligand library commenced with the preparation of 14 new imines in addition to the two mentioned above. When volatile amines, i.e. ones with a boiling point of less than 100 °C were used as a starting material, they were employed in large excess over the aldehyde in the presence of a drying agent. Otherwise, equal amounts of both starting materials were employed and water was removed by means of a Dean-Stark trap. Yields were above 80% in most cases and the crude products were sufficiently pure for the subsequent cycloaddition step without further purification (Table 8).

**Table 8.** A series of imines was synthesized as starting materials for cycloaddition reactions with imidate esters.

	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	Yield %
<b>42</b>	OMe	OH	Me	83
<b>47</b>	OMe	OH	Et	90
<b>48</b>	OMe	OH	<i>i</i> Pr	96
<b>49</b>	OMe	OH	<i>t</i> Bu	— <sup>[a]</sup>
<b>50</b>	OMe	OH	4-MeCy	> 99
<b>51</b>	OMe	OH	PhCH <sub>2</sub>	> 99
<b>52</b>	OMe	OH	4-MeOPhCH <sub>2</sub>	> 99
<b>53</b>	OMe	OH	Ph	97
<b>45</b>	OMe	OH	4-MePh	> 99
<b>54</b>	OMe	OH	4-MeOPh	94
<b>55</b>	OMe	OH	4-CF <sub>3</sub> Ph	98
<b>56</b>	OMe	OH	4-CF <sub>3</sub> OPh	97
<b>57</b>	OMe	OH	4-Me <sub>2</sub> NPh	68
<b>58</b>	OMe	OH	4- <i>t</i> BuPh	> 99
<b>59</b>	H	OH	4-MeOPh	93
<b>60</b>	OMe	H	4-MeOPh	> 99

[a] Not characterized due to low stability.

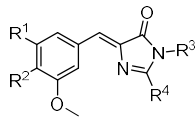
Next, these imines were employed in cycloaddition reactions with an imidate ester to generate modified HBI scaffolds (Table 9). The brominated HBI analog **77** was instead prepared independently in two steps via azlactone intermediate **76** (Scheme 18).



**Scheme 18.** Preparation of **77** via Erlenmeyer azlactone synthesis. Reagents and conditions: a) *N*-acetylglycine, NaOAc, Ac<sub>2</sub>O, 90 °C, 2 h, 68%. b) MeNH<sub>2</sub>, K<sub>2</sub>CO<sub>3</sub>, EtOH, reflux, 4 h, 21%.

Compound **71**, which contains a 4-(dimethylamino)phenyl substituent at N3, was treated with an excess of MeI in DMF to afford the cationic HBI derivative DMHBI<sup>+</sup> (**72**). While the cycloaddition gave a wide range of yields depending on the nature of the imine that were uncorrelated with the steric and electronic properties of R<sup>3</sup>, it is evidently superior to the azlactone approach when taking ease of purification and total yield over all reaction steps (azlactone formation and hydrolysis vs imine formation and cycloaddition) into consideration.

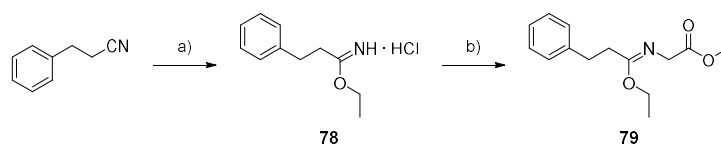
**Table 9.** Overview of HBI derivatives with modified benzylidene moieties (R1, R2) and side chains (R3, R4). Compounds were prepared according to Route A in Scheme 16 unless noted otherwise.

		R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	R <sup>4</sup>	Yield %	
	DMHBI	<b>13</b>	OMe	OH	Me	Me	82
	DMHBI-Et	<b>61</b>	OMe	OH	Et	Me	35
	DMHBI- <i>i</i> Pr	<b>62</b>	OMe	OH	<i>i</i> Pr	Me	93
	DMHBI- <i>t</i> Bu	<b>63</b>	OMe	OH	<i>t</i> Bu	Me	18
	DMHBI-MeCy	<b>64</b>	OMe	OH	4-MeCy	Me	40
	DMHBI-Bn	<b>65</b>	OMe	OH	PhCH <sub>2</sub>	Me	77
	DMHBI-PMBn	<b>66</b>	OMe	OH	4-MeOPhCH <sub>2</sub>	Me	67
	DMHBPI	<b>67</b>	OMe	OH	Ph	Me	93
	DMHBTI	<b>44</b>	OMe	OH	4-MePh	Me	54 <sup>[a]</sup>
	DMHBI-spdt	<b>46</b>	OMe	OH	MeO(CO)CH <sub>2</sub>	Me	24
	DMHBAI	<b>68</b>	OMe	OH	4-MeOPh	Me	30
	DMHBTI <sup>F</sup>	<b>69</b>	OMe	OH	4-CF <sub>3</sub> Ph	Me	27
	DMHBAI <sup>F</sup>	<b>70</b>	OMe	OH	4-CF <sub>3</sub> OPh	Me	45
	DMHBI-DMA	<b>71</b>	OMe	OH	4-Me <sub>2</sub> NPh	Me	79
	DMHBI <sup>+</sup>	<b>72</b>	OMe	OH	4-Me <sub>3</sub> N <sup>+</sup> Ph	Me	> 99
	DMHBI <sup>C</sup>	<b>73</b>	OMe	OH	4- <i>t</i> BuPh	Me	76
	MHBAI	<b>74</b>	H	OH	4-MeOPh	Me	25
	DMBAI	<b>75</b>	OMe	H	4-MeOPh	Me	69
	BMHBI	<b>77</b>	Br	OH	Me	Me	21 <sup>[b]</sup>
	DMHBI-PhEt	<b>80</b>	OMe	OH	Me	Ph(CH <sub>2</sub> ) <sub>2</sub>	73

[a] An average yield of 83% was obtained when scouting for optimized reaction conditions.

[b] Synthesized via the Erlenmeyer azlactone route.

To diversify the ligand library, additional HBI derivatives with modified substituents at position C2 of the imidazolone ring were prepared. In the simplest case, this was achieved by replacing the methyl imidate ester **43** in the cycloaddition reaction with the phenylethyl derivative **79**. The precursor imidate hydrochloride **78** was prepared from 3-phenylpropionaldehyde via Pinner reaction with a yield of 75%. The required anhydrous HCl was generated in situ from acetyl chloride and the solvent EtOH according to a reported protocol (Scheme 19).<sup>143</sup> After conversion to the imidate ester, cycloaddition with imine **42** proceeded readily in EtOH at ambient temperature to afford DMHBI-PhEt (**80**).

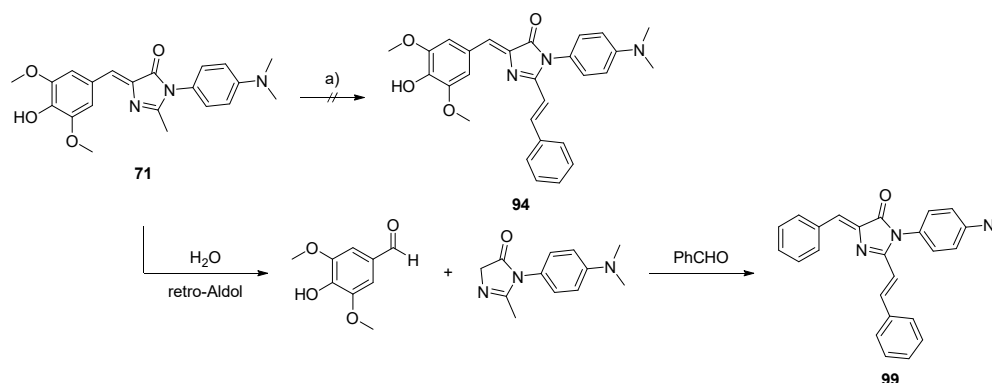


**Scheme 19.** Synthesis of a modified imidate via Pinner reaction. Reagents and conditions: a) AcCl, EtOH, 4 °C, 7 h, 75%. b) Glycine methyl ester hydrochloride, NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, r.t., 6 h, 72%.

As discussed above, a major point of interest was the synthesis of Kaede- and DsRed-type HBI derivatives (Table 10) following three different synthetic approaches, namely aldol condensation with aromatic aldehydes (Scheme 16, Route B), Wittig reaction with the respective phosphonium ylides (Route C) or reaction with hydroxylamine hydrochloride (Route D). The latter two approaches required oxidation of the C2-methyl group of the HBI core, which was conveniently achieved via a Riley-type oxidation with SeO<sub>2</sub> in dioxane. Even though Yampolsky et al. reported that the oxidation product thus obtained from unsubstituted HBI underwent complete decomposition upon chromatographic purification and therefore required transient protection of

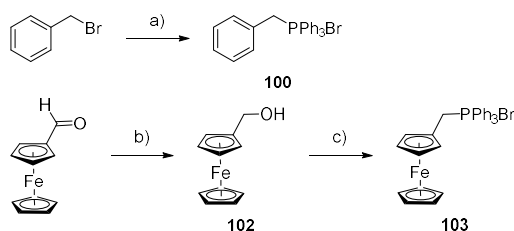
the phenolic OH group,<sup>139</sup> no such difficulties were encountered in this work. Specifically, compounds **13** and **71** were converted to the formylated derivatives **92** and **93**, respectively, in moderate to good yields.

For most reactant combinations, the Lewis acid-catalyzed reaction between an HBI derivative and an aromatic aldehyde proceeded reasonably well, albeit with the formation of numerous side products. While initially  $\text{ZnCl}_2$  was used as the Lewis acid, it was found that by using  $\text{Sc}(\text{OTf})_3$  instead the formation of side products could be reduced, simplifying the isolation of the target compounds. One of the side products from the reaction of **71** with benzaldehyde was isolated in pure form and was identified as the benzylidene imidazolone derivative **99** by NMR and HR-MS, indicating that a retro-Aldol reaction of the 4-hydroxy-3,5-dimethoxybenzylidene imidazolone moiety had taken place (Scheme 20). The target compound **94**, however, could not be isolated from the reaction mixture.



**Scheme 20.** Reaction of **71** with benzaldehyde did not lead to the desired product **94**. Instead, a retro-Aldol reaction took place which eventually resulted in the formation of **99**. Reagents and conditions: PhCHO, cat.  $\text{Sc}(\text{OTf})_3$ , dioxane, 110 °C, 48 h, 4%.

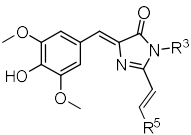
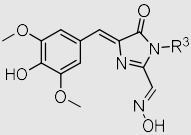
Further evidence implicating retro-Aldol reaction as a competing pathway was obtained from the reaction of **13** with ferrocene carbaldehyde, which only yielded trace amounts of the target compound **96** in addition to 4-hydroxy-3,5-dimethoxybenzaldehyde. Compounds **94** and **96** were therefore synthesized by Wittig reaction of the respective formylated HBI derivatives with either the benzylphosphonium ylide **100** or the ferrocenylphosphonium ylide **103**, both of which were prepared according to Scheme 21 following standard methods.<sup>144–147</sup>



**Scheme 21.** Synthesis of phosphonium ylides as starting materials for Wittig reactions with formylated HBI derivatives. Reagents and conditions: a)  $\text{PPh}_3$ , toluene, 90 °C, 15 h, 90%. b)  $\text{NaBH}_4$ , THF/MeOH, r.t., 1 h, 97%. c)  $\text{HPPH}_3\text{Br}$  (**101**), toluene, reflux, 2 h, 71%.

Finally, **93** was reacted with hydroxylamine hydrochloride under basic conditions to obtain the intermediate **97**, which was methylated along with **94** using an excess of MeI in DMF to obtain the cationic DsRed analogs **95** and **98**, respectively.

**Table 10.** Overview of Kaede- and DsRed-type DMHBI derivatives with modified side chains (R<sup>5</sup>). Compounds were prepared according to Routes B, C or D in Scheme 16.

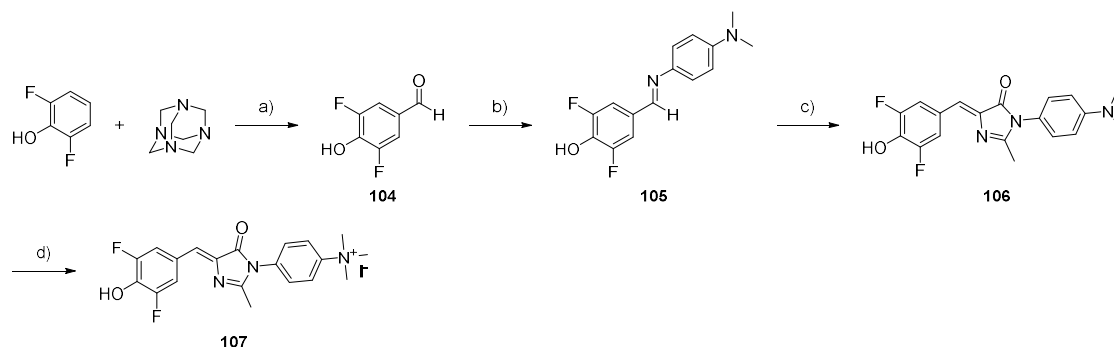
		R <sup>3</sup>	R <sup>5</sup>	Route	Yield %		
	DMHBI-Styr	<b>81</b>	Me	Ph	B	48	
	DMHBI-2Py	<b>82</b>	Me	2-Py	B	15	
	DMHBI-3Py	<b>83</b>	Me	3-Py	B	35	
	DMHBI-4Py	<b>84</b>	Me	4-Py	B	67	
	DMHBI-Imi	<b>85</b>	Me	4-Imi	B	63	
	DMHBI-Ind	<b>86</b>	Me	3-Ind	B	30	
	DMHBTI-2Py	<b>87</b>	4-MePh	2-Py	B	91	
	DMHBTI-3Py	<b>88</b>	4-MePh	3-Py	B	43	
	DMHBTI-4Py	<b>89</b>	4-MePh	4-Py	B	38	
	DMHBTI-Imi	<b>90</b>	4-MePh	4-Imi	B	25 <sup>[a]</sup>	
	DMHBTI-Ind	<b>91</b>	4-MePh	3-Ind	B	28	
	DMHBI-Styr-DMA	<b>94</b>	4-Me <sub>2</sub> NPh	Ph	C	37	
	DMHBI-Styr <sup>+</sup>	<b>95</b>	4-Me <sub>3</sub> N <sup>+</sup> Ph	Ph	C	> 99	
	DMHBI-Fc	<b>96</b>	Me	Fc	C	12	
		DMHBO-DMA	<b>97</b>	4-Me <sub>2</sub> NPh	—	D	37
		DMHBO <sup>+</sup>	<b>98</b>	4-Me <sub>3</sub> N <sup>+</sup> Ph	—	D	> 99

[a] 10:1 mixture of *E* and *Z* isomers.

All synthesized ligands were fully and unambiguously characterized by NMR spectroscopy and HR-MS. Kaede-type derivatives were obtained as *E* isomers at the newly formed exocyclic double bond as proven by the respective <sup>3</sup>J<sub>HH</sub> coupling constants, which were always on the order of 16 Hz. The sole exception was compound **90**, which was isolated as a 10:1 mixture of *E* and *Z* isomers. Additionally, the configuration of the central double bond was exemplarily determined for compound **81** using high-resolution <sup>1</sup>H-<sup>13</sup>C HMBC data; the observed <sup>3</sup>J<sub>CH</sub> coupling constant of 4.5 Hz for the α,β-unsaturated carbonyl moiety is highly characteristic for a *Z*-configured double bond.<sup>148</sup>

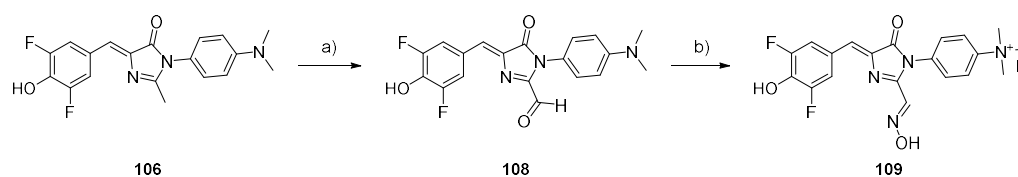
Beyond the ligands described so far, a new set of DFHBI derivatives bearing the cationic trimethylammoniumphenyl side chain at N3 were synthesized to test them for binding to the Spinach and Broccoli aptamers. To this end, 2,6-difluorophenol was formylated under Duff conditions following a reported procedure<sup>149</sup> (77%) and the resulting aldehyde **104** converted to imine **105**. Intermediate **106** was then synthesized in 36% yield via cycloaddition with **43** and methylated quantitatively to obtain DFHBI<sup>+</sup> (**107**, Scheme 22).





**Scheme 22.** Synthesis of DFHBI<sup>+</sup> (**107**). Reagents and conditions: a) TFA, reflux, 26 h, 77%. b) 4-(Dimethylamino)aniline, toluene, reflux, 16 h, >99%. c) **43**, toluene, 120 °C, 16 h, 36%. d) MeI, DMF, r.t., 24 h, >99%.

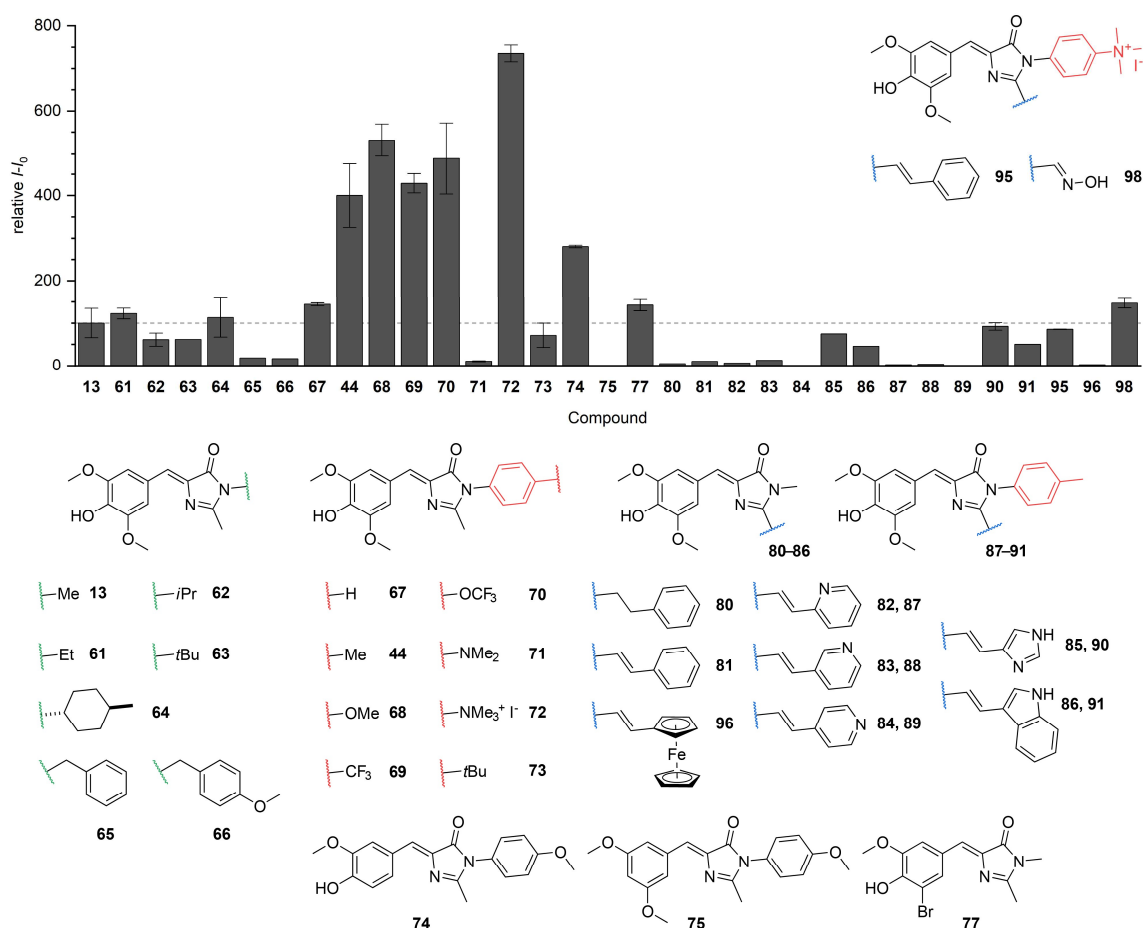
Additionally, **106** was oxidized to the aldehyde **108** (41%) and then reacted with  $\text{NH}_2\text{OH} \cdot \text{HCl}$  followed by MeI to obtain the cationic oxime DFHBO<sup>+</sup> (**109**, Scheme 23).



**Scheme 23.** Synthesis of DFHBO<sup>+</sup> (**109**). Reagents and conditions: a)  $\text{SeO}_2$ , dioxane, reflux, 1.5 h, 41%. b) 1.  $\text{NH}_2\text{OH} \cdot \text{HCl}$ ,  $\text{K}_2\text{CO}_3$ , MeOH, r.t., 41 h; 2. MeI, DMF, r.t., 26 h.

### 2.3.1.3 Structure–fluorescence activation relationships

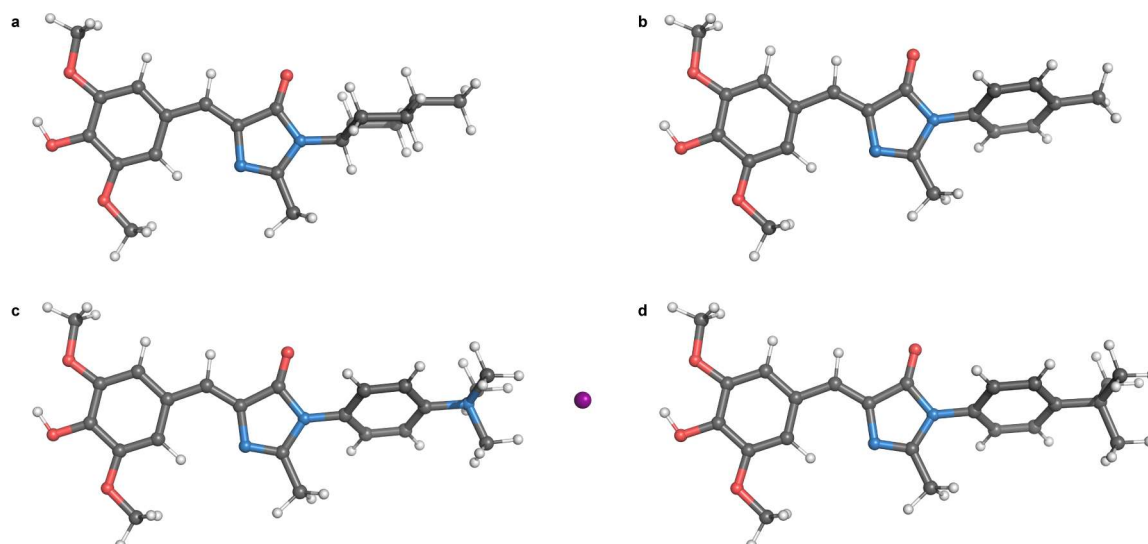
A screening assay was implemented to compare the fluorogenic potential of novel HBI derivatives to that of the parent ligand DMHBI in conjunction with the Chili aptamer (Figure 21). Briefly, the RNA aptamer was annealed by heating to 95 °C for 3 min and then cooled to ambient temperature. It was then incubated with one equivalent of ligand for 24 h to ensure proper equilibration. The binding buffer contained both  $\text{K}^+$  (125 mM) and  $\text{Mg}^{2+}$  (5 mM), mimicking the conditions present during the selection of the 13-2 aptamer. Fluorescence spectra were collected at the respective excitation and emission maxima, keeping all other measurement parameters, such as detector sensitivity and integration time, constant. RNA and ligand were employed at a concentration of 0.5  $\mu\text{M}$  each, which is on the order of the dissociation constant for the 13-2–DMHBI complex. As a result, a high screening brightness requires novel ligands to display both high affinity and a large fluorescence-turn-on. In addition to the overall brightness, special attention was paid to possible spectral shifts, as this might indicate a change in the ligand binding mode.



**Figure 21.** Fluorescence screening results for activation of new HBI ligands of the GFP-, Kaede- DsRed-type by the Chili aptamer. Conditions: 0.5  $\mu\text{M}$  RNA, 0.5  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 80 mM HEPES pH 7.5. Fluorescence spectra were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ) relative to Chili-DMHBI. Spectral data and intensity values are listed in Table 11, the respective fluorescence and excitation spectra are shown in Supplementary Figure 1 and 2.

The first set of ligands under investigation comprised those bearing increased steric bulk at position N3 of the imidazolone ring. Increasing the substituent size from methyl (**13**) to ethyl (**61**) had no discernible effect, but branching out to isopropyl (**62**) and *tert*-butyl (**63**) resulted in a 40% loss of fluorescence enhancement. However, actual steric bulk seems to be a secondary factor in this case, as is shown by the *trans*-4-methylcyclohexyl-substituted derivative **64**, which is as bright as the parent ligand DMHBI. Likely, the inclusion of branched substituents increases the rotational degrees of freedom of the ligand to such a degree that non-radiative deactivation via internal conversion is enhanced. This assumption is supported by the finding that not only branched alkyl, but also benzyl substituents at N3 (**65**, **66**), have a detrimental impact on fluorescence intensity.<sup>150</sup>

A second set of ligands, consisting of compounds **67–73**, was used to study the effect of  $\pi$ - $\pi$  interactions between an aromatic ligand side chain and the binding site of the aptamer. Their electronic properties were controlled by the nature of their *para*-substituents. In particular, ligands bearing  $\text{CH}_3$  groups were directly contrasted with those bearing  $\text{CF}_3$  groups instead. Not only has the trifluoromethyl group a strong electron-withdrawing effect, but it is also known to confer lipophilic properties and therefore might have an impact on RNA binding.<sup>151</sup> Geometry optimizations (DFT: B3LYP-D3/def2-TZVP) were performed for selected ligands and show that their global structure is largely unaffected by the nature of the side chain, i.e. the main HBI scaffold is planar in all cases. The aromatic N3-substituents are, however, twisted out of that plane by 50–59° (Figure 22, Supplementary Figure 4).



**Figure 22.** DFT-optimized geometries of a) DMHBI-MeCy (**64**), b) DMHBTI (**44**) c) DMHBI<sup>+</sup> (**72**) and d) DMHBI<sup>c</sup> (**73**) in the gas phase (B3LYP-D3/def2-TZVP).

A 40% enhancement of fluorescence activation over the parent ligand DMHBI was observed for the phenyl-substituted DMHBPI (**67**). Moreover, the excitation spectrum of the respective RNA–ligand complex was redshifted by 10 nm. A similar redshift occurs with the other HBI derivatives in this series and must stem from a specific interaction with the RNA, as it is not found in the unbound ligands (Supplementary Figure 1h–o). Spectral tuning of the emission maxima, however, was very limited (< 3 nm). Similar trends were observed for the Spinach–DFHBI-1T complex, albeit with a greater magnitude.<sup>83</sup> There, the imidazolone N3-substituent is located in close proximity to a nucleotide that has been shown to be important for spectral tuning in a series of Spinach-related Broccoli aptamer variants.<sup>80</sup>

DMHBTI (**44**) presents a particularly interesting case, showing a fourfold increase of fluorescence activation over DMHBI and also over the virtually isosteric cycloalkyl derivative DMHBI-MeCy (**64**, Figure 22a, b). Compared to its constitutional isomer DMHBI-Bn (**65**), fluorescence intensity of the Chili complex was even increased by a factor of 20. Identical behavior was found for N3-aryl ligands with electron-donating (**44**, **68**) and electron-withdrawing (**69**, **70**) *para*-substituents. Incorporation of a strongly electron-donating dimethylamino group in DMHBI-DMA (**71**) led to almost total loss of fluorescence activation. This was to be expected however, as the dimethylaniline (DMA) moiety has been used as an electron donor for fluorescence quenching by photoinduced electron transfer (PET) in several instances.<sup>152</sup> Later experiments (see section 2.8.2.3) demonstrated that the ligand is indeed quenched but still capable of binding to the Chili aptamer. *N*-methylation of **71** interrupted the PET pathway and led to complete recovery of the fluorescence activation; in fact, the resulting cationic ligand DMHBI<sup>+</sup> (**72**) fluoresced more than sevenfold brighter (Figure 23) than DMHBI in complex with the Chili aptamer. To better understand the influence of the trimethylammonium moiety, the isomorphous neutral analog DMHBI<sup>c</sup> (**73**), which bears a *tert*-butyl group instead, was synthesized. Despite their gas phase geometries being completely superimposable (Figure 22c, d, Supplementary Figure 5), DMHBI<sup>+</sup> is tenfold more emissive than **73**. It is worth mentioning that the calculated gas phase dipole moment of **73** (1.9 D) is smaller than those of compounds **68–70** (up to 3.2 D) and especially that of DMHBI<sup>+</sup> (16.7 D), which seems to indicate that a larger ligand dipole moment is beneficial for the Chili system. Due to its bright fluorescence, DMHBI<sup>+</sup> was included as a model ligand in the studies described in section 2.4.

**Table 11.** Substitution patterns and spectral data for the screened ligands shown in **Figure 21**. Fluorescence intensity is given relative to Chili-DMHBI.

Ligand		R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	R <sup>4</sup>	R <sup>5</sup>	$\lambda_{ex/em}$ nm	rel. fluores- cence int.
DMHBI	<b>13</b>	OMe	OH	Me	Me	—	400/537	1.00
DMHBI-Et	<b>61</b>	OMe	OH	Et	Me	—	400/537	1.22
DMHBI- <i>i</i> Pr	<b>62</b>	OMe	OH	<i>i</i> Pr	Me	—	400/537	0.61
DMHBI- <i>t</i> Bu	<b>63</b>	OMe	OH	<i>t</i> Bu	Me	—	400/534	0.61
DMHBI-MeCy	<b>64</b>	OMe	OH	4-MeCy	Me	—	400/537	1.13
DMHBI-Bn	<b>65</b>	OMe	OH	PhCH <sub>2</sub>	Me	—	400/537	0.18
DMHBI-PMBn	<b>66</b>	OMe	OH	4-MeOPhCH <sub>2</sub>	Me	—	400/535	0.16
DMHBPI	<b>67</b>	OMe	OH	Ph	Me	—	410/539	1.44
DMHBTI	<b>44</b>	OMe	OH	4-MePh	Me	—	410/539	4.00
DMHBAI	<b>68</b>	OMe	OH	4-MeOPh	Me	—	410/538	5.31
DMHBTI <sup>F</sup>	<b>69</b>	OMe	OH	4-CF <sub>3</sub> Ph	Me	—	413/540	4.29
DMHBAI <sup>F</sup>	<b>70</b>	OMe	OH	4-CF <sub>3</sub> OPh	Me	—	413/540	4.87
DMHBI-DMA	<b>71</b>	OMe	OH	4-Me <sub>2</sub> NPh	Me	—	413/540	0.10
DMHBI <sup>+</sup>	<b>72</b>	OMe	OH	4-Me <sub>3</sub> N <sup>+</sup> Ph	Me	—	413/542	7.35
DMHBI <sup>C</sup>	<b>73</b>	OMe	OH	4- <i>t</i> BuPh	Me	—	410/539	0.71
MHBAI	<b>74</b>	H	OH	4-MeOPh	Me	—	397/513	2.80
DMBAI	<b>75</b>	OMe	H	4-MeOPh	Me	—	n.d. <sup>[a]</sup>	n.d.
BMHBI	<b>77</b>	Br	OH	Me	Me	—	386/520	1.42
DMHBI-PhEt	<b>80</b>	OMe	OH	Me	Ph(CH <sub>2</sub> ) <sub>2</sub>	—	400/539	0.04
DMHBI-Styr	<b>81</b>	OMe	OH	Me	—	Ph	462/601	0.10
DMHBI-2Py	<b>82</b>	OMe	OH	Me	—	2-Py	467/616	0.06
DMHBI-3Py	<b>83</b>	OMe	OH	Me	—	3-Py	465/611	0.12
DMHBI-4Py	<b>84</b>	OMe	OH	Me	—	4-Py	n.d.	n.d.
DMHBI-Imi	<b>85</b>	OMe	OH	Me	—	4-Imi	463/594 <sup>[b]</sup>	0.74
DMHBI-Ind	<b>86</b>	OMe	OH	Me	—	3-Ind	469/539	0.46
DMHBTI-2Py	<b>87</b>	OMe	OH	4-MePh	—	2-Py	464/618	0.02
DMHBTI-3Py	<b>88</b>	OMe	OH	4-MePh	—	3-Py	467/613	0.03
DMHBTI-4Py	<b>89</b>	OMe	OH	4-MePh	—	4-Py	n.d.	n.d.
DMHBTI-Imi	<b>90</b>	OMe	OH	4-MePh	—	4-Imi	461/592 <sup>[c]</sup>	0.92
DMHBTI-Ind	<b>91</b>	OMe	OH	4-MePh	—	3-Ind	478/539	0.50
DMHBI-Styr <sup>+</sup>	<b>95</b>	OMe	OH	4-Me <sub>3</sub> N <sup>+</sup> Ph	—	Ph	465/603	0.85
DMHBI-Fc	<b>96</b>	OMe	OH	Me	—	Fc	460/573	0.02
DMHBO <sup>+</sup>	<b>98</b>	OMe	OH	4-Me <sub>3</sub> N <sup>+</sup> Ph	—	—	456/592	1.47

[a] No signal detected.

[b] Emission shoulder at 542 nm.

[c] Dual emission profile with additional excitation and emission peaks at 416 and 542 nm, respectively.

Given the fact that structurally and electronically diverse modifications to the imidazolone-N3 substituent were tolerated by the Chili aptamer, its ligand recognition mechanism must be dominated by interactions with the benzylidene moiety. When a methoxy group of DMHBI is replaced by bromine, the resulting ligand BMHBI (**77**) shows a moderate blueshift of the excitation and emission maxima (386 and 520 nm, respectively) without loss of fluorescence enhancement. Complete removal of a methoxy group, however, leads to an approximately 50% decline in fluorescence intensity when comparing MHBAI (**74**) and DMHBAI (**68**) and is accompanied by a more pronounced blueshift of the emission maximum (513 nm). This suggests that both methoxy groups contribute to an optimal occupation of the binding site due to steric effects. Moreover, the presence of the para-hydroxy group is essential for fluorescence activation; in DMBAI (**75**) removal

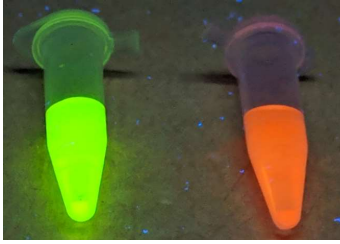
of this group prevents binding to the Chili aptamer (as will be discussed in section 2.4.1.2) and thereby abolishes fluorescence activation.

At the outset of screening Kaede-type HBI derivatives, two isosteric ligands, in which the imidazolone C2-methyl group has been replaced by an ethylphenyl or **(80)** styryl **(81)** moiety, respectively, were compared. As expected, spectral differences between DMHBI and the saturated compound **80** were negligible. Fluorescence intensity, however, dropped by a factor of 25 for the latter. DMHBI-Styr **(81)** was only slightly more emissive in complex with Chili, suggesting that steric bulk at C2 is detrimental to ligand binding or activation. Due to extended conjugation, the excitation and emission maxima are redshifted by over 60 nm, resulting in red fluorescence emission at 601 nm. The fluorescence intensity could be restored back to the level of Chili–DMHBI by replacing the N3-methyl group of **81** with a trimethylammoniumphenyl substituent to generate DMHBI-Styr<sup>+</sup> **(95)**. Spectral shifts between **81** and **95** were minimal ( $\leq 3$  nm difference). Inclusion of the cationic “plus” substituent is therefore a viable strategy for brightening Chili aptamer ligands.

Additional hydrogen bond acceptor sites were then introduced by replacing the styryl group of **81** with different vinylpyridyl side chains in an attempt to potentially create more contact points between the ligand and the binding site. DMHBI-2Py **(82)** and DMHBI-3Py **(83)** were minimally activated by the Chili aptamer with emission maxima of 615 and 610 nm, respectively, whereas DMHBI-4Py **(84)** was completely non-fluorogenic. The latter has recently also been studied as potential ligand for the fluorogen-activating FAST protein, but displayed no activity.<sup>153</sup> Interestingly, these three compounds showed a noticeably higher level of background fluorescence than other HBI derivatives discussed up to this point (Supplementary Figure 2b–d).

DMHBI-Imi **(85)**, which more closely mimics the Kaede chromophore by virtue of its C2-vinylimidazolyl side chain, was activated significantly stronger by the aptamer and reached a brightness similar to that of Chili–DMHBI. Remarkably, though, when the complex was excited at 463 nm, its emission spectrum was bimodal. Fitting the spectrum with a sum of Gaussian function revealed two emission bands peaking at 542 nm (38% relative area) and 592 nm (62% relative area), suggesting the presence of two photophysically distinct species (Supplementary Figure 2o). In particular, the higher energy band must arise from a fluorophore population in which the extended conjugation has been lost. A possible mechanism lies in proton transfer from the imidazole to a nearby acceptor site upon excitation of the fluorophore. This hypothesis is corroborated by the largely similar behavior of the indole analog DMHBI-Ind **(86)** in complex with Chili, which shows a single prominent emission maximum at 537 nm in addition to a less pronounced low-energy shoulder. With indole being a well-characterized photoacid,<sup>154</sup> proton loss from the NH group after excitation is likely and could result in loss of the extended conjugation. These experiments were repeated with similar results when screening a series of DMHBTI-based ligands **(87–91)**. Among these, DMHBTI-Imi **(90)** was obtained as a 10:1 mixture of *E* and *Z* isomers at the vinylic double bond. While it still showed the same dual emission as its simpler homolog DMHBI-Imi, the relative intensity of the bands was now dependent on the excitation wavelength, likely due to the presence of the two electronically distinct geometric isomers.

DMHBO<sup>+</sup> **(98)** was investigated as hydrolytically stable DsRed chromophore analog. Despite its small size, with a  $\pi$  extension comprising only a single double bond in the oxime moiety, excitation and emission maxima of 456 and 592 nm, respectively, were observed in complex with Chili. Moreover, it is the strongest emitting ligand for Chili found in this spectral range (Figure 23) and was therefore selected for further in depth characterization.



**Figure 23.** The Chili aptamer forms highly fluorescent complexes with the cationic ligands DMHBI<sup>+</sup> (left) and DMHBO<sup>+</sup> (right). Conditions: 10–15  $\mu\text{M}$  RNA, 1.6–2.3 eq. ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5,  $\lambda_{\text{ex}} \approx 365$  nm (handheld UV lamp).

## 2.3.2 Systematic mutagenesis of Chili

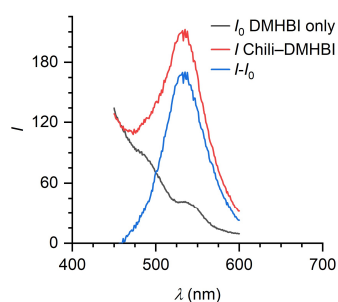
### 2.3.2.1 Binding site analysis

Mutagenesis of the Chili aptamer was performed for all 22 nucleotides within the ligand binding site. For most mutants, single-nucleotide transitions were introduced to conserve the steric footprint of the respective nucleobase; at position 9, an A–U transversion was introduced instead to avoid a stretch of seven consecutive guanine residues (Table 12).

**Table 12.** Mutagenesis pattern of the Chili ligand binding pocket. Blue and yellow nucleotides indicate transition and transversion mutants, respectively. Fluorescence enhancement of DMHBI at 540 nm is given relative to the wt aptamer.

Chili variant	5'-sequence-3'	rel. fluorescence intensity
	10    15                    35    40	
wt	...GGAGGGGC...GGUUGGGUGCGGUC...	1.00
G9A	... <b>A</b> GAGGGGC...GGUUGGGUGCGGUC...	0.04
G10A	...G <b>A</b> AGGGGC...GGUUGGGUGCGGUC...	0.06
A11U	...GG <b>U</b> GGGGC...GGUUGGGUGCGGUC...	0.05
G12A	...GGA <b>A</b> GGGC...GGUUGGGUGCGGUC...	0.08
G13A	...GGAG <b>A</b> GGC...GGUUGGGUGCGGUC...	0.03
G14A	...GGAGG <b>A</b> GC...GGUUGGGUGCGGUC...	0.20
G15A	...GGAGGG <b>A</b> C...GGUUGGGUGCGGUC...	0.09
C16U	...GGAGGGG <b>U</b> ...GGUUGGGUGCGGUC...	0.00
G31A	...GGAGGGGC... <b>A</b> GUUGGGUGCGGUC...	0.07
G32A	...GGAGGGGC...G <b>A</b> UUGGGUGCGGUC...	0.10
U33C	...GGAGGGGC...GG <b>C</b> UGGGUGCGGUC...	0.09
U34C	...GGAGGGGC...GGU <b>C</b> GGUGCGGUC...	1.02
G35A	...GGAGGGGC...GGUU <b>A</b> GGUGCGGUC...	0.63
G36A	...GGAGGGGC...GGUUG <b>A</b> GUGCGGUC...	0.96
G37A	...GGAGGGGC...GGUUGG <b>A</b> UGCGGUC...	0.13
U38C	...GGAGGGGC...GGUUGGG <b>C</b> GCGGUC...	0.58
G39A	...GGAGGGGC...GGUUGGGU <b>A</b> CGGUC...	0.07
C40U	...GGAGGGGC...GGUUGGGUG <b>U</b> GGUC...	0.74
G41A	...GGAGGGGC...GGUUGGGUG <b>A</b> GUC...	0.07
G42A	...GGAGGGGC...GGUUGGGUGCG <b>A</b> UC...	0.05
U43C	...GGAGGGGC...GGUUGGGUGCGG <b>C</b> C...	0.13
C44U	...GGAGGGGC...GGUUGGGUGCGGU <b>U</b> ...	0.54

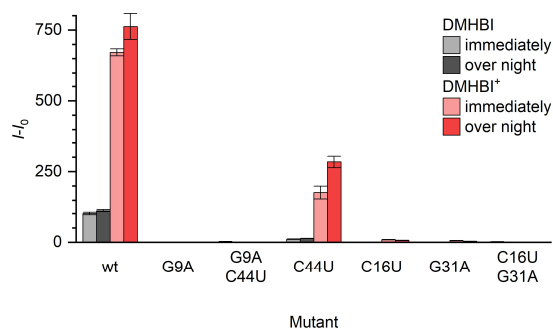
The fluorescence enhancement of DMHBI was measured for each mutant relative to the wild type by means of a high-throughput assay on a 96-well plate (Figure 24, see section 4.1.13 for method details). Using an RNA concentration on the same order of magnitude as the dissociation constants of wt-Chili–DMHBI ensured high sensitivity towards changes in affinity, which directly contribute to the observed fluorescence enhancement.



**Figure 24.** Fluorescence spectra of the RNA complexes were measured in a 96-well plate-based assay and background-corrected by subtraction of a ligand-only spectrum to remove scattering and residual fluorescence. Conditions: 0.5  $\mu\text{M}$  RNA, 2  $\mu\text{M}$  DMHBI, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5,  $\lambda_{\text{ex}} = 405$  nm.

No mutations were tolerated in the regions defined by nucleotides 9 to 16, 31 to 33 and 41 to 43. In fact, mutations at any of these positions resulted in a more than 80% reduction of fluorescence enhancement at any position, which indicates that these parts of the bulge are highly important for ligand binding or activation. In the region defined by nucleotides 34 to 40, however, only two guanine residues were found to be essential (Table 12). It is likely that conformational flexibility, e.g. in form of non-interacting nucleobases, is required to shift G37 and G39 into a specific arrangement for ligand binding. This is in agreement with the  $Tb^{3+}$  digestion data (section 2.1.2), which showed that this region of the bulge becomes more accessible to the surrounding solvent upon binding of DMHBI.

Interestingly, simultaneous mutation of G9 and C44, located at the junction between the bulge and basal stem, to an A and C, respectively, resulted in complete loss of function, whereas introduction of a transition mutation at position 44 only led to a 56% decrease in fluorescence enhancement. In contrast, mutation of the C16–G31 base pair at the junction between the bulge and the apical stem-loop was not tolerated at all. Therefore, the double mutants G9A/C44U and C16U/G31A were analyzed in more detail with a conventional cuvette-based assay using the brighter ligand DMHBI<sup>+</sup> in addition to DMHBI. Neither of the double mutants was active, showing that both G9 and G31 are essential. Overnight incubation with DMHBI<sup>+</sup> resulted in a more pronounced increase of fluorescence intensity of Chili-C44U compared to wt-Chili, suggesting that U at this position hampers folding of the aptamer (Figure 25).



**Figure 25.** Fluorescence emission of wt-Chili and single/double mutants in the presence of DMHBI or DMHBI<sup>+</sup> measured after incubation for 3 min and overnight. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 80 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{ex}$  = 400 nm for DMHBI, 413 nm for DMHBI<sup>+</sup>) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. (n = 3).

### 2.3.2.2 Analysis of stem mutants

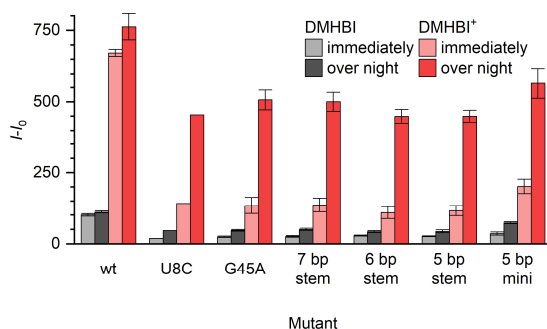
During the development of Chili from the 13-2min aptamer, three base pairs were removed from the basal stem without negative impact on the fluorescence activation of DMHBI; in fact, folding of the aptamer into a binding-competent conformation was strongly enhanced (see section 2.1.2). Still, a G–U wobble pair at the junction between stem and binding site remained in place; while they are almost as stable as canonical Watson-Crick base pairs, G–U wobble pairs deviate in their glycosidic bond angles, which leads to a perturbation of the local secondary structure in A-form RNA.<sup>155</sup> It was therefore hypothesized that the presence of a wobble pair at the stem–bulge junction might serve a functional purpose. Indeed, replacement by either a C8–G45 or U8–A45 Watson-Crick base pair resulted in diminished ligand activation and prolonged the maturation time of the aptamer, which is evidenced by a substantial rise of the fluorescence intensity after overnight incubation (Figure 26).



**Table 13.** Mutagenesis pattern of the Chili stem regions. Blue and yellow nucleotides indicate transition and transversion mutants, respectively.

Chili variant	5'-sequence-3'					Length nt
	5	20	30	45	50	
wt	G	GCCAGUUCGCU	G	G	G	52
U8C	G	GCCAGUUCGCU	G	G	G	52
G45A	G	GCCAGUUCGCU	G	A	G	52
7 bp stem	G	GCCAGUUCGCU	G	A	G	50
6 bp stem	G	GCCAGUUCGCU	G	A	G	48
5 bp stem	G	GCCAGUUCGCU	G	A	G	46
5 bp mini	G	GCC	GGA	G	G	42

Therefore, this base pair seems to isolate stem and ligand binding site from each other. Truncation of the stem by 1–3 bp (resulting in 7, 6 and 5 bp stem mutants, respectively; Table 13) did not significantly change the fluorescence intensity of the resulting aptamer complexes. A miniaturized Chili variant was then designed by combining the 5 bp basal stem with a shorter apical stem containing a GGAA tetraloop, leaving a 42 nt construct that retained approximately 75% of the wt-Chili fluorescence intensity with DMHBI and DMHBI<sup>+</sup> after overnight incubation. Both the 5 bp stem variant and the miniaturized aptamer (5 bp mini) were chosen for further characterization of their ligand binding affinities.

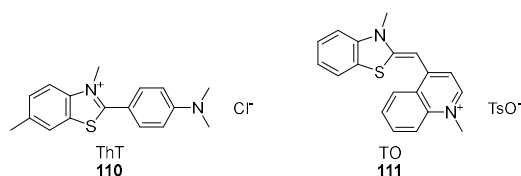


**Figure 26.** Fluorescence emission of wt-Chili and truncation mutants in the presence of DMHBI or DMHBI<sup>+</sup> measured after incubation for 3 min and overnight. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 80 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}}$  = 400 nm for DMHBI, 413 nm for DMHBI<sup>+</sup>) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. (n = 3).

Overall, systematic mutagenesis showed that wt-Chili already displays optimal fluorescence enhancement of HBI ligands, but tolerates further minimization very well.

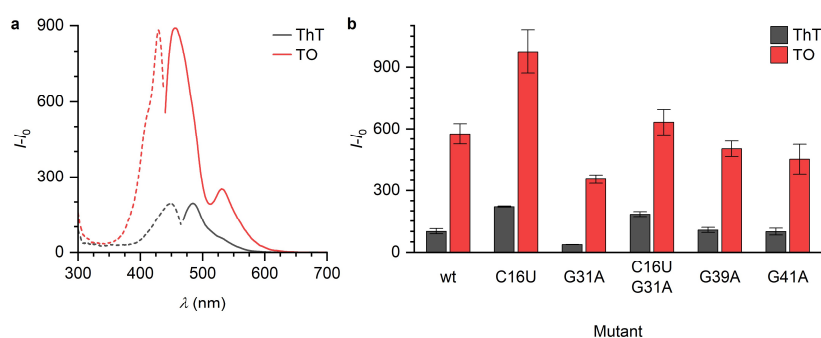
### 2.3.2.3 Activation of non-HBI quadruplex-binding fluorogenic ligands

The high guanine content of the Chili aptamer in conjunction with its dependence on K<sup>+</sup> ions for fluorescence activation suggests that a G-quadruplex is involved in the function of the RNA (see section 2.2.2). Therefore, the wt aptamer as well as several of the G–A transition mutants described above were examined for fluorescence activation of thioflavin T (ThT, **110**) and thiazole orange (TO, **111**, Chart 21), two fluorogenic probes that bind to quadruplex DNA and RNA with moderate selectivity.<sup>156–161</sup> Notably, the quadruplex-forming Spinach aptamer strongly activates ThT fluorescence.<sup>97</sup> Both ThT and TO are activated by the Corn aptamer to an even higher degree than its cognate ligand DFHO at the same concentration and X-ray co-crystal structures show binding of these fluorogens at the exposed face of the homodimeric quadruplexes.<sup>162</sup>



**Chart 21.** The fluorogenic dyes ThT and TO can bind to DNA- and RNA quadruplexes.

Chili activates ThT and TO fluorescence with an intensity of 0.25 and 1.46 times that of DMHBI, respectively (Figure 27). However, a direct comparison between the fluorescence intensity of these ligands and DMHBI (or DFHO in case of the Corn aptamer) is hardly meaningful due to their substantially different photophysical properties. Mutants C16U, G31A and C16U/G31A, all of which did not enhance fluorescence of cognate ligands, induced ThT and TO fluorescence. Interestingly, the double mutant showed the same degree of activation as wt-Chili, whereas the individual mutants were either more or less active. The compensatory nature of the double mutation suggests that C16 and G31 form a base pair, as was surmised from secondary structure analysis (see section 2.1.2).



**Figure 27.** a) Fluorescence excitation (dashed) and emission (solid) spectra of wt-Chili in complex with thioflavin T (ThT) or thiazole orange (TO). b) Fluorescence emission of wt-Chili and single/double mutants in the presence of ThT or TO measured after incubation for 3 min. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 80 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{ex}$  = 449 nm for ThT, 430 nm for TO) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. (n = 3).

While mutants G39A and G41A did not activate DMHBI fluorescence, their complexes with ThT and TO are as bright as those of wt-Chili. If either of the two nucleotides takes part in G-quadruplex formation, it is conceivable that a mixed-sequence quartet resulting from a mutation would disturb the topology sufficiently to prevent specific interaction with the cognate ligand, but still allow for non-specific activation of the other fluorogens.

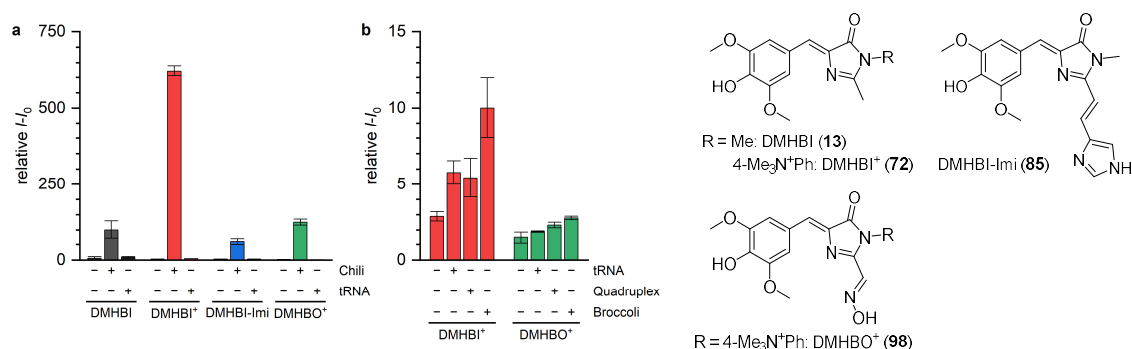
## 2.4 DETAILED CHARACTERIZATION OF SELECTED CHILI–HBI COMPLEXES

Several strongly fluorogenic ligands for the Chili aptamer as well as promising aptamer mutants with reduced size were identified from the comprehensive screening results discussed in the preceding section. To gain insight into structural and spectroscopic properties of the free ligands and their respective Chili complexes in greater detail, a wide array of state-of-the-art characterization techniques was employed. After establishing specific fluorescence activation, the relationship between ligand protonation state and binding was investigated. Relative quantum yields of Chili complexes were determined and binding affinities were measured by fluorescence titration as well as isothermal titration calorimetry. The resulting static picture of the underlying RNA–ligand interactions was then expanded by considering dynamic processes that accompany the complex formation, which yielded strong direct experimental evidence for the formation of a G-quadruplex. Special focus was placed on DMHBI<sup>+</sup> (**72**), DMHBI-Imi (**85**) and DMHBO<sup>+</sup> (**98**) as the most promising members of their structural families (GFP-, Kaede- and DsRed-like, respectively). Other ligands were included in the characterization whenever required to address specific questions.

### 2.4.1 Binding specificity and metal ion dependence

#### 2.4.1.1 Fluorescence activation by non-cognate nucleic acids

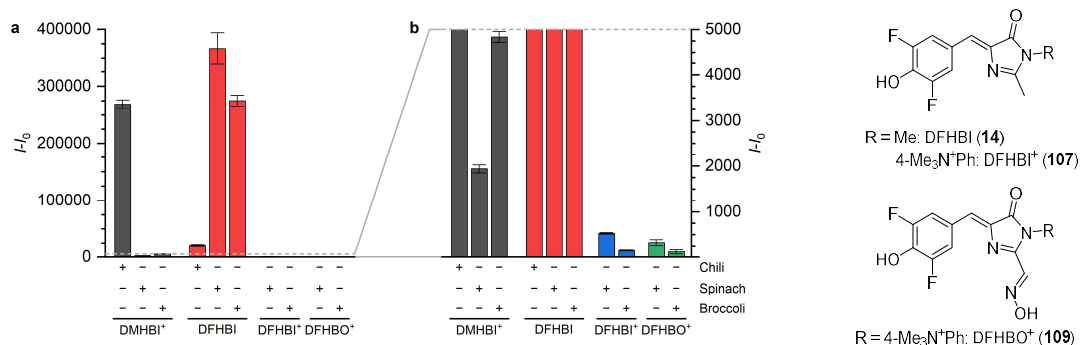
The first step of a more detailed characterization of GFP-, Kaede- and DsRed-like ligands was an examination of their specificity towards the Chili aptamer, i.e. whether they would also be activated by non-cognate nucleic acids. This was especially important for the two HBI derivatives bearing a permanent positive charge, because such compounds can be highly susceptible to non-specific interactions with nucleic acids by groove binding or intercalation and can also take part in stacking with G-quadruplex domains.<sup>163–166</sup> Activation of the parent ligand DMHBI, cationic DMHBI<sup>+</sup>, the Kaede-mimic DMHBI-Imi and the cationic DsRed-type chromophore DMHBO<sup>+</sup> was therefore tested with tRNA from *E. coli*, the 17 nt quadruplex-forming DNA oligonucleotide G<sub>3</sub>T<sup>167</sup> as well as the DFHBI-binding RNA aptamer Broccoli. In these control experiments, off-target activation was found to be minimal for all four compounds, which underlines the high specificity of the Chili–DMHBI interaction (Figure 28).



**Figure 28.** a) Integrated emission intensities of HBI ligands in buffer and in the presence of either the Chili aptamer or *E. coli* tRNA after blank correction (buffer) relative to Chili–DMHBI. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 80 mM HEPES pH 7.5. b) Integrated emission intensities of DMHBI<sup>+</sup> and DMHBO<sup>+</sup> in buffer and in the presence of a tenfold excess of the indicated RNA (tRNA, Broccoli) or DNA (quadruplex) after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Additionally, in order to investigate whether strong fluorescence activation of cationic HBI derivatives is unique to the Chili aptamer, the fluorinated analogs DFHBI<sup>+</sup> (**107**) and DFHBO<sup>+</sup> (**109**) were examined for their behavior in the presence of the Spinach or Broccoli aptamer. The crystal structure of the Spinach–DFHBI complex suggests that replacement by either compound should

be possible on steric grounds. Moreover, members of the Broccoli family were shown to elicit fluorescence emission from the uncharged DsRed-like oxime DFHO.<sup>80</sup>

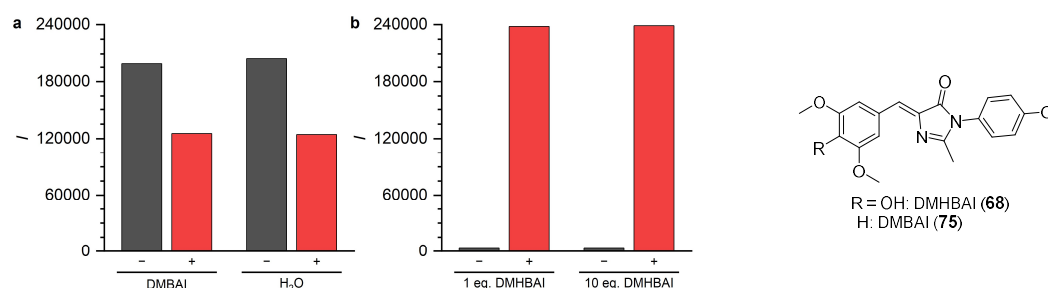


**Figure 29.** a) Integrated emission intensities of HBI ligands in buffer and in the presence of the Chili, Spinach or Broccoli aptamer after blank correction (buffer). b) Magnified view of the data in panel a using the same y-axis units. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand, 125 mM KCl, 5 mm MgCl<sub>2</sub>, 80 mM HEPES pH 7.5. Values are given as mean  $\pm$  s.d. (n = 3).

However, neither aptamer induced fluorescence above background levels in combination with the “plus” variants of DFHBI and DFHO, suggesting that this kind of compound is rejected entirely (Figure 29). Spinach, and by extension Broccoli, were selected specifically with the goal in mind of binding ligands that are negatively charged. As such, it is reasonable to assume that the net neutral charge of DFHBI<sup>+</sup> and DFHBO<sup>+</sup> at neutral pH prevents binding to their cognate aptamers. Therefore, the use of HBI derivatives with a permanent positive charge is a further exclusive feature of Chili in addition to highly Stokes-shifted fluorescence emission due to protons transfer.

#### 2.4.1.2 Recognition of the phenolic hydroxy group

Ligand screening revealed that DMBAI (75), in which the phenolic hydroxy group is absent, does not become activated upon incubation with Chili. A competition experiment between DMBAI and the hydroxylated analog DMHBAI (68), which is strongly fluorogenic, was performed to investigate whether the hydroxy group is required for binding of the ligand or only for its subsequent activation. Briefly, treatment of preformed Chili–DMHBAI complex with equal volumes of either buffer or an excess of the competitor DMBAI resulted in the same decrease of fluorescence intensity due to dilution of the Chili complex. This indicated that DMBAI is not capable of displacing DMHBAI from the ligand binding site of the aptamer (Figure 30a). Similarly, treatment of a sample containing Chili and an excess of DMBAI with 1 or 10 eq. of DMHBAI resulted in the same degree of fluorescence enhancement (Figure 30a). Therefore, DMBAI could not have occupied the ligand binding site in the first place.

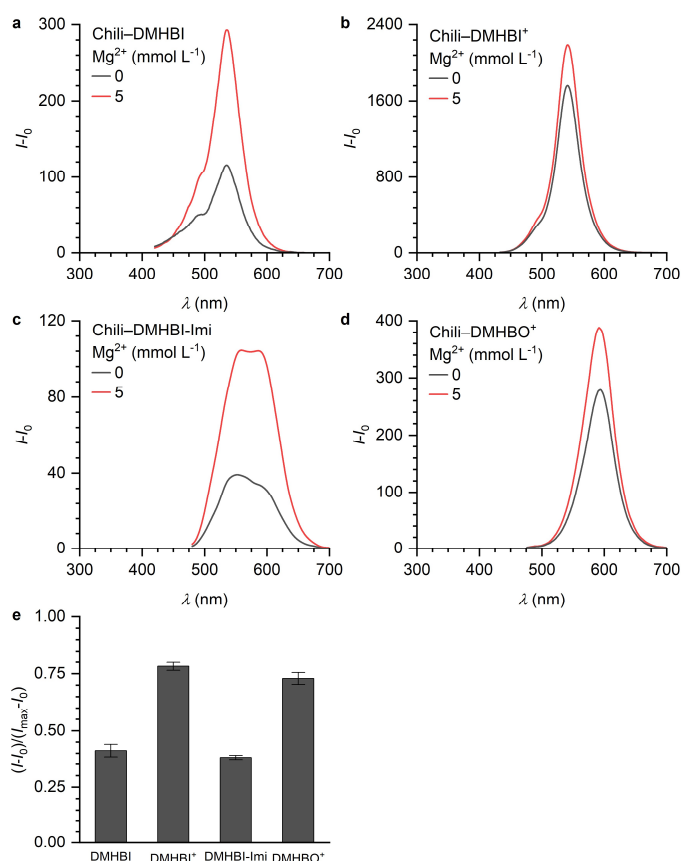


**Figure 30.** a) Integrated emission intensities for Chili–DMHBAI before and after adding an excess of DMBAI or an equivalent volume of water. Conditions: 0.5  $\mu$ M RNA, 0.5  $\mu$ M ligand (DMHBAI), 125 mM KCl, 5 mm MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. b) Integrated emission intensities for Chili in the presence of DMBAI before and after adding 1 or 10 eq. of DMHBAI. 0.5  $\mu$ M RNA, 5  $\mu$ M ligand (DMBAI), 125 mM KCl, 5 mm MgCl<sub>2</sub>, 40 mM HEPES pH 7.5.

Taken together, these findings demonstrate that the Chili aptamer specifically recognizes the phenolic hydroxy group of its ligands.

#### 2.4.1.3 Metal ion dependence

As shown earlier, ligand activation by the Chili aptamer requires both  $K^+$  and  $Mg^{2+}$  ions. In combination with DMHBI, approximately  $0.5\text{--}1\text{ mmol L}^{-1}$   $Mg^{2+}$  are required to achieve optimal brightness, whereas half-maximal fluorescence activation is already observed in the absence of added  $Mg^{2+}$  (see section 2.2.2). To investigate whether the new family of “plus” ligands, so far showing promising properties with regard to binding and brightness under standard conditions, have an influence on the  $Mg^{2+}$ -dependence of the Chili system, their brightness in  $Mg^{2+}$ -free buffer was examined. Indeed, both  $DMHBI^+$  and  $DMHBO^+$  retain approximately 80% of the fluorescence intensity observed with  $5\text{ mM}$   $Mg^{2+}$ . As a control, the same experiment was performed with the neutral Kaede-analog DMHBI-Imi, which was found to behave identically to DMHBI, suggesting that while  $Mg^{2+}$  is directly involved in ligand uptake or activation, a positively charged side chain can serve as a functional surrogate (Figure 31).



**Figure 31.** Fluorescence spectra of Chili in complex with a) DMHBI ( $\lambda_{ex}$  = 400 nm), b) DMHBI<sup>+</sup> ( $\lambda_{ex}$  = 413 nm), c) DMHBI-Imi ( $\lambda_{ex}$  = 460 nm) or d) DMHBO<sup>+</sup> ( $\lambda_{ex}$  = 456 nm) in the absence (black) or presence (red) of  $Mg^{2+}$ . e) Integrated emission intensities of  $Mg^{2+}$ -free samples after blank correction (buffer). The intensities of each respective  $Mg^{2+}$ -containing sample was set to 1. Conditions:  $0.5\ \mu\text{M}$  RNA,  $0.5\ \mu\text{M}$  ligand,  $125\text{ mM}$  KCl,  $40\text{ mM}$  HEPES pH 7.5,  $+ 5\text{ mM}$   $MgCl_2$ . Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

The lower  $Mg^{2+}$  requirement of the cationic ligands is accompanied by an increased acidity of their phenolic hydroxy groups as determined by spectrophotometric titration (Supplementary Figure 8, Supplementary Table 1); notably, the  $pK_a$  of DMHBO<sup>+</sup> is  $6.89 \pm 0.02$ , more than one log unit lower than that of DMHBI (Table 14, see section 2.2.1). It is conceivable that  $Mg^{2+}$  coordinates to the hydroxy group of a Chili-bound ligand, which would facilitate the proton abstraction that generates the actual emissive species (see Scheme 8 on page 34). In cationic ligands, the positively

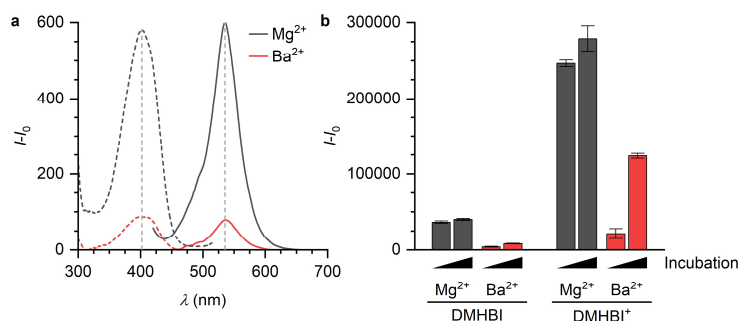
charged side chain increases the acidity of the hydroxy group, potentially rendering  $\text{Mg}^{2+}$  functionally obsolete.

**Table 14.**  $\text{pK}_a$  values for the phenol/phenolate equilibrium of four HBI derivatives. Values for the other protonation states of these and more ligands are given in Supplementary Table 1.

Ligand	$\text{pK}_a$
DMHBI <sup>[a]</sup>	$7.96 \pm 0.01$
DMHBI <sup>+</sup>	$7.63 \pm 0.02$
DMHBI-Imi	$7.87 \pm 0.02$
DMHBO <sup>+</sup>	$6.89 \pm 0.02$

[a] As reported in section 2.2.1.

In the crystal structure of the Spinach–DFHBI complex, a direct interaction between the benzylidene moiety of the ligand and a  $\text{Mg}^{2+}$  ion close to the RNA backbone has been observed. Upon substituting  $\text{Ba}^{2+}$  for  $\text{Mg}^{2+}$ , an 80% drop in fluorescence intensity occurred and was accompanied by a 27 nm blueshift of the ligand absorption.<sup>82</sup> QM/MM simulations of the complex have shown that the absorption maximum of Spinach-bound DFHBI is governed by the local electrostatic field, which is influenced by surrounding metal ions.<sup>168,169</sup> No such spectral tuning was observed when fluorescence excitation and emission spectra of Chili–DMHBI were measured in binding buffer containing 5 mM  $\text{Ba}^{2+}$ . However, the brightness of the complex decreased tenfold compared to the same concentration of  $\text{Mg}^{2+}$  (Figure 32a). It recovered to a limited degree over the course of 24 h and then remained at a stable level. When the experiment was repeated with DMHBI<sup>+</sup>, the recovery effect was more pronounced, but surprisingly, the maximal fluorescence enhancement in the presence of  $\text{Ba}^{2+}$  was lower than in the complete absence of divalent ions (Figure 32b).



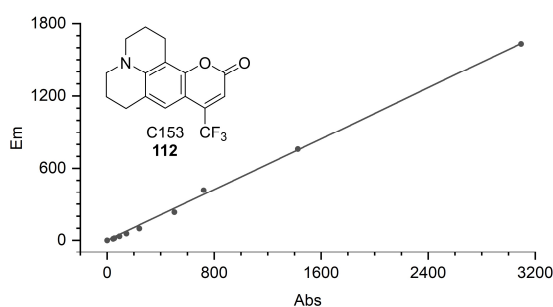
**Figure 32.** a) Fluorescence excitation (dashed) and emission (solid) spectra of Chili–DMHBI after 3 min incubation in the presence of  $\text{Mg}^{2+}$  or  $\text{Ba}^{2+}$  ( $\lambda_{\text{ex/em}} = 400/537$  nm). b) Fluorescence intensities of Chili complexes with DMHBI ( $\lambda_{\text{ex}} = 400$  nm) and DMHBI<sup>+</sup> ( $\lambda_{\text{ex}} = 413$  nm) in the presence of  $\text{Mg}^{2+}$  or  $\text{Ba}^{2+}$ . Emission spectra were integrated after blank correction (buffer). Conditions: 0.5  $\mu\text{M}$  RNA, 0.5  $\mu\text{M}$  ligand, 125 mM KCl, 40 mM HEPES pH 7.5, + 5 mM  $\text{MgCl}_2$  or  $\text{BaCl}_2$ . Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

The absence of spectral tuning indicates that DMHBI derivatives in complex with Chili are less susceptible to electrostatic field effects than Spinach-bound DFHBI. The loss of fluorescence enhancement instead suggests that divalent metal ions primarily affect the global stability of the active RNA conformation. Overall, the results presented in this section underscore the high potential of the Chili system and its “plus” ligands for in-cell applications, where high affinity and brightness as well as tolerance to limited  $\text{Mg}^{2+}$  availability are especially important.

#### 2.4.2 Relative fluorescence quantum yields

Fluorescence quantum yields of DMHBI derivatives and their Chili complexes were determined by the relative method, i.e. by comparing their absorbance and fluorescence intensity to that of a

reference fluorophore with a precisely known quantum yield (see section 4.1.12 for method details). Absorption and emission maxima of the reference fluorophore and of the sample under study should ideally be as similar as possible in order to minimize the effect of wavelength-dependent spectrofluorometer characteristics (lamp intensity, detector sensitivity).<sup>170</sup> For the Chili system, the most suitable quantum yield standard was coumarin 153 (C153, **112**) with  $\lambda_{\text{abs}} = 423 \text{ nm}$  and  $\lambda_{\text{em}} = 531 \text{ nm}$ . A quantum yield of  $0.53 \pm 0.01$  was determined by absolute measurement using an integrating sphere setup, which compares well with previously reported values (Figure 33).<sup>171</sup>



**Figure 33.** Plot of the integrated emission vs absorbed light for samples of C153 in EtOH ( $\lambda_{\text{ex}} = 418 \text{ nm}$ , for further details see section 4.1.12.1). The absolute quantum yield was determined from a linear fit to the data.  $\phi_{\text{abs}} = 0.53 \pm 0.01$ . Conditions: 0–50  $\mu\text{M}$  C153, EtOH.

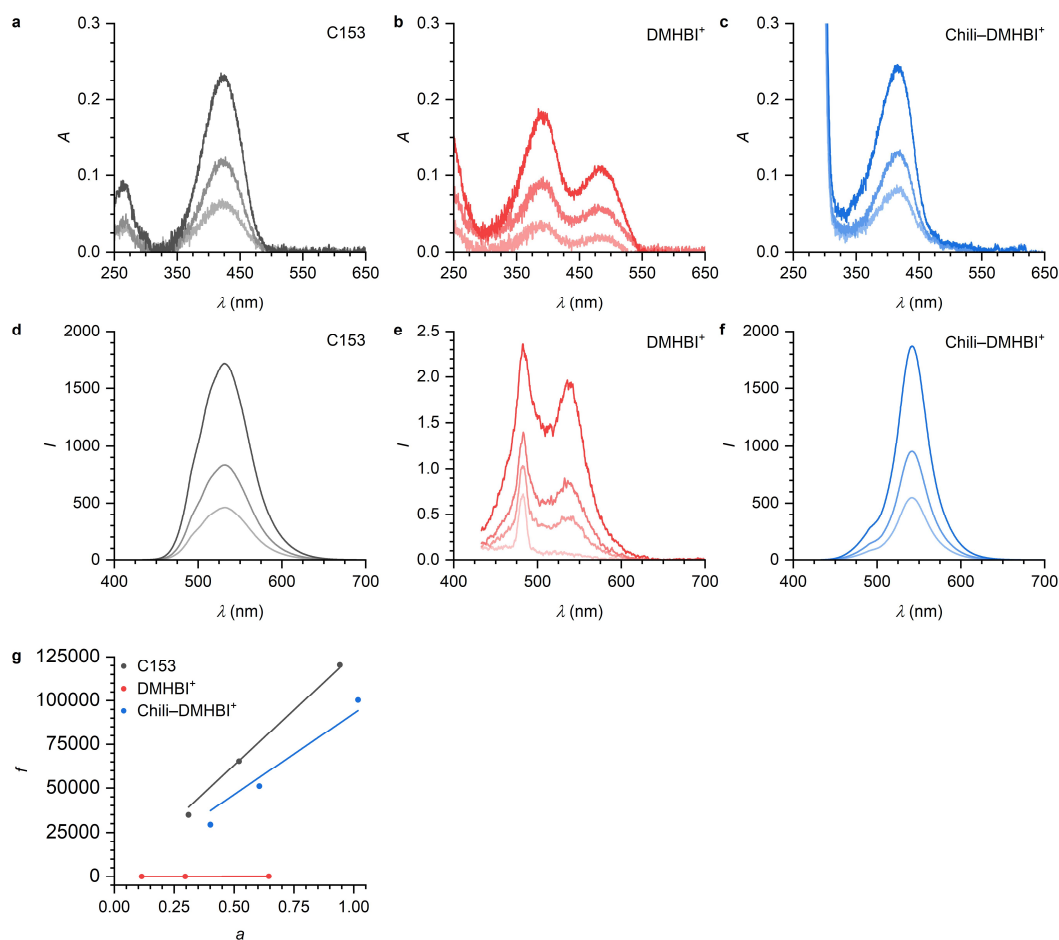
Relative quantum yields were then obtained from absorbance and fluorescence spectra of each ligand alone and in presence of the Chili aptamer (Figure 34, Supplementary Figure 3). The RNA was used in excess over the ligand to ensure complete binding, as the presence of residual unbound ligand would erroneously decrease the measured quantum yield. Extinction coefficients of the respective Chili complexes were estimated from the absorbance spectra and used to calculate their fluorescence brightness (Table 15).

**Table 15.** Fluorescence quantum yields of HBI derivatives before and after activation by the Chili aptamer were measured using the relative method. The resulting brightness values of the complexes were calculated by multiplying quantum yield and extinction coefficient.

Ligand	$\phi_{\text{rel}}$		$\epsilon$	Brightness <sup>[a]</sup>
	free	bound	L mol <sup>-1</sup> cm <sup>-1</sup>	bound
DMHBI <sup>[b]</sup> <b>13</b>	$6.6 \pm 0.1 \times 10^{-4}$	$7.6 \pm 0.1 \times 10^{-2}$	18000	1400
DMHBI <sup>+</sup> <b>72</b>	$1.1 \pm 0.1 \times 10^{-3}$	$4.0 \pm 0.1 \times 10^{-1}$	21000	8400
DMHBI-Imi <b>85</b>	$3.3 \pm 0.4 \times 10^{-3}$	$7.5 \pm 0.2 \times 10^{-2}$	20000	1500
DMHBO <sup>+</sup> <b>98</b>	$1.0 \pm 0.1 \times 10^{-3}$	$9.4 \pm 0.4 \times 10^{-2}$	22000	2100

[a] Brightness =  $\phi \times \epsilon \times L^{-1} \text{ mol cm}$ .

[b] As reported in section 2.2.1.



**Figure 34.** a–c) UV/Vis spectra of the quantum yield standard C153 in EtOH, DMHBI<sup>+</sup> and Chili-DMHBI<sup>+</sup> in buffer at concentrations of 2.5, 5 and 10  $\mu\text{M}$ . d–f) Fluorescence emission spectra ( $\lambda_{\text{ex}} = 413 \text{ nm}$ ) of the same samples in addition to blank samples (EtOH and buffer, respectively). g) Plot of integrated fluorescence intensities ( $f$ ) vs absorption factors ( $a$ ). Solid lines represent linear fits. For details about the calculation of  $a$ ,  $f$  and relative quantum yields see section 4.1.12.

Overall, the increase in quantum yield and resulting brightness closely follow the emission intensities observed during the screening assay (2.3.1.3). The fluorescence quantum yield of Chili-DMHBI<sup>+</sup> ( $\phi = 0.40$ ) is comparable to that of other RNA fluoromolecules such as Red Broccoli-DFHBO ( $\phi = 0.34$ ) and Mango IV-TO1-biotin ( $\phi = 0.41$ ).<sup>93</sup>

### 2.4.3 Binding analysis by fluorescence spectroscopy

#### 2.4.3.1 Equilibrium binding titrations with modified ligands

Fluorescence titrations were performed with a total of eight new ligands (Table 16) in combination with the Chili aptamer following the same protocol described earlier, i.e. titrating the ligand at a fixed concentrations with increasing amounts of RNA (see sections 2.2.2 and 4.1.14). For initial experiments, a ligand concentration of 0.1  $\mu\text{M}$  was used; if the resulting titration curve did not display well-behaved saturation behavior, i.e. was either too steep or too shallow, the experiment was repeated with an adjusted ligand concentration for better sampling of the binding isotherm (Supplementary Figure 6).<sup>172,173</sup>

DMHBTI (**44**) and DMHBAl (**68**) were found to bind more strongly to the Chili aptamer than the parent ligand DMHBI. Surprisingly, their fluorinated analogs DMHBTI<sup>F</sup> (**69**) and DMHBAl<sup>F</sup> (**70**), which are equally as emissive according to screening data, did not show saturation in their binding isotherms even after increasing the ligand concentration. While dissociation constants could be obtained by fitting the data with a Hill model, these values cannot be taken as reliable indicators



of the binding affinity due to the lack of saturation. An uncorrelated relationship between brightness and affinity is well documented in cases where a single ligand has been studied in combination with different RNA aptamers, e.g. for DFHBI complexes of Spinach variants or TO1-biotin complexes of various Mango derivatives.<sup>94,99</sup> Examples of the reverse case, i.e. a single aptamer binding to several different ligands, are naturally more sparse due to the specificity of each pair, but likewise show no clear trend.<sup>93</sup> In the present case, the low affinity of the fluorinated ligands with respect to their nonfluorinated analogs can possibly be attributed to the highly lipophilic nature of their side chains,<sup>151</sup> which might interact unfavorably with the polar RNA environment.

**Table 16.** Dissociation constants for the binding of various bright ligands to Chili. Titration curves (Figure 36, Supplementary Figure 6) were fitted with a single-site-binding model or alternatively the Hill equation.

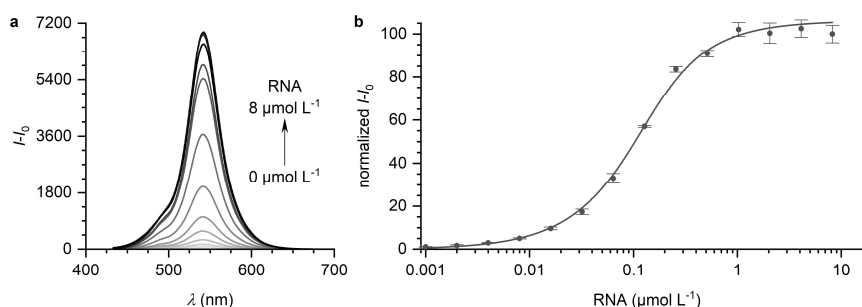
Ligand		rel. $I-I_0$ <sup>[a]</sup>	$K_d$ mol L <sup>-1</sup>
DMHBI	<b>13</b>	1.00	$5.7 \pm 1.2 \times 10^{-7}$ <sup>[b]</sup>
DMHBTI	<b>44</b>	4.00	$1.5 \pm 0.2 \times 10^{-7}$
DMHBAl	<b>68</b>	5.31	$6.5 \pm 0.7 \times 10^{-8}$
DMHBTI <sup>F</sup>	<b>69</b>	4.29	$1.4 \pm 0.1 \times 10^{-7}$ <sup>[c]</sup>
DMHBAl <sup>F</sup>	<b>70</b>	4.87	$1.5 \pm 0.2 \times 10^{-6}$ <sup>[c]</sup>
DMHBI <sup>+</sup>	<b>72</b>	7.35	$6.3 \pm 0.6 \times 10^{-8}$
DMHBI <sup>C</sup>	<b>73</b>	0.71	$7.4 \pm 0.7 \times 10^{-7}$ <sup>[c]</sup>
DMHBI-Imi	<b>85</b>	0.74	$7.1 \pm 0.2 \times 10^{-8}$
DMHBO <sup>+</sup>	<b>98</b>	1.47	$1.2 \pm 0.2 \times 10^{-8}$

[a] Screening result, see Table 11.

[b] As reported in section 2.2.2.

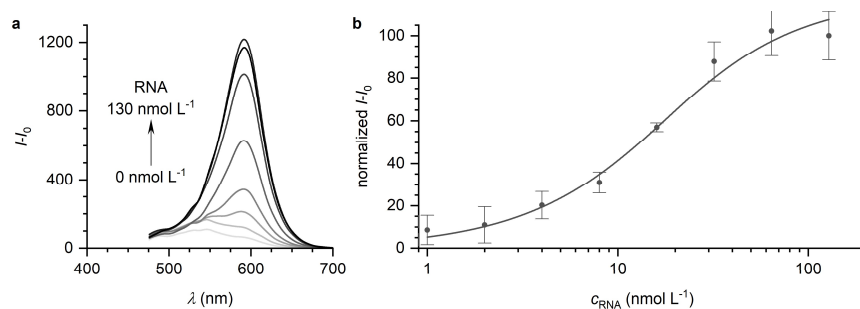
[c] Fitted with a Hill model.

The cationic ligand DMHBI<sup>+</sup> (**72**) binds with a dissociation constant of approximately 63 nM, similar to DMHBAl (Figure 35). While certainly not a bright fluorophore, DMHBI<sup>C</sup> (**73**), the neutral isomorphous analog of DMHBI<sup>+</sup>, was investigated as well in order to uncover further differences imparted by the positive charge of the “plus” substituent. Again, the titration curve did not reach proper saturation, suggesting weak binding to the Chili aptamer. After fitting the data with a Hill model, its affinity was estimated to be ten times smaller than that of DMHBI<sup>+</sup>. Therefore, the positive charge of DMHBI<sup>+</sup> seems to increase affinity, but only to a certain level. In light of these results, DMHBI<sup>+</sup> has emerged as the best-performing ligand among those exhibiting fluorescence in the green region of the visible spectrum by combining tight binding and bright emission.



**Figure 35.** a) Fluorescence enhancement of DMHBI<sup>+</sup> by Chili. b) Titration of DMHBI<sup>+</sup> with Chili. Data points were fitted with a single-site binding model.  $K_d = 6.3 \pm 0.6 \times 10^{-8}$  mol L<sup>-1</sup>. Conditions: 0–8.2  $\mu$ M RNA, 0.1  $\mu$ M ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 413$  nm) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. (n = 3).

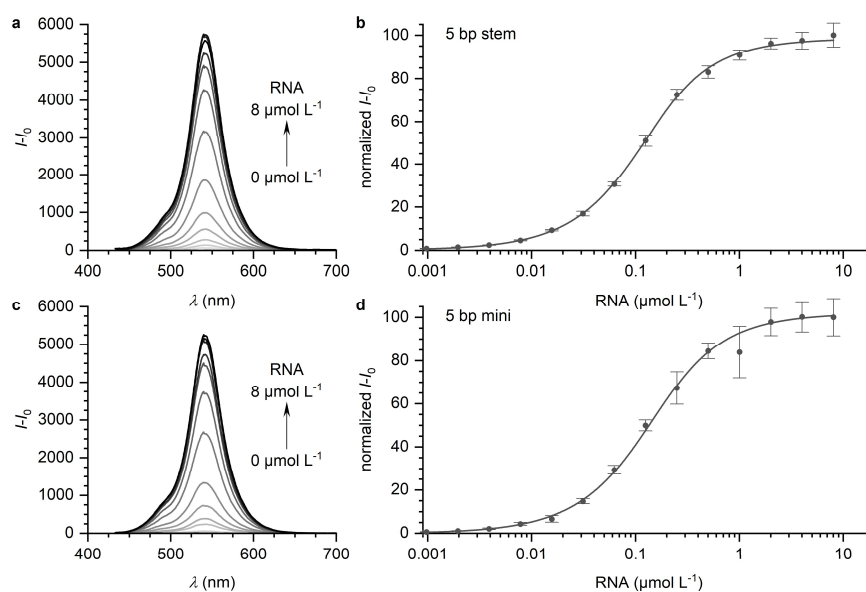
A similarly high affinity of approximately 71 nM was measured for the Kaede-type chromophore DMHBI-lmi (**85**). The presence of a bulky substituent at imidazolone-C2 is apparently tolerated well by the ligand binding site. By comparison, the binding site of the Spinach aptamer would not be able to accommodate such an extended ligand, as it would clash with the backbone between nucleotides A64 and G65 when inserted into a model of the RNA (Supplementary Figure 7).<sup>82</sup> Therefore, the binding site topology of Chili and Spinach must be different. Notably, the cationic DsRed mimic DMHBO<sup>+</sup> (**98**) binds to Chili with a  $K_d$  of 12 nM (Figure 36). The oxime moiety of this ligand might form an additional hydrogen bond with a nucleotide in the binding site, raising the affinity substantially above that of DMHBI<sup>+</sup>. In fact, while low-nanomolar dissociation constants have been reported for complexes of Mango-family aptamers<sup>95,99,101</sup> and the RiboGlow<sup>106</sup> platform, such a high affinity is unprecedented and currently unsurpassed (see Table 7 on page 39) for RNA mimics of fluorescent proteins.



**Figure 36.** a) Fluorescence enhancement of DMHBO<sup>+</sup> by Chili. b) Titration of DMHBO<sup>+</sup> with Chili. Data points were fitted with a single-site binding model.  $K_d = 1.2 \pm 0.2 \times 10^{-8} \text{ mol L}^{-1}$ . Conditions: 0–128 nM RNA, 10 nM ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 456 \text{ nm}$ ) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

### 2.4.3.2 Equilibrium binding titrations with truncated RNAs

Two short Chili constructs, the 5 bp stem truncation mutant and the miniaturized aptamer (5 bp mini, see Table 13) were also investigated for their affinity towards DMHBI<sup>+</sup> (Figure 37). The resulting dissociation constants were on the same order of magnitude as that of wt-Chili, making these short RNAs interesting candidates for future NMR studies, which require large amounts of sample (Table 17). In the course of this work, however, further studies were still conducted with the full-length Chili aptamer due to its faster maturation compared to the truncation mutants.



**Figure 37.** Fluorescence spectra and resulting titration curves upon treating DMHBI<sup>+</sup> with either a 5 bp stem (a, b) or miniaturized (c, d) variant of the Chili aptamer (see Table 13). Data points were fitted with a single-site binding model. Conditions: 0–8.2  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Fluorescence spectra ( $\lambda_{\text{ex}} = 413$  nm) were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

**Table 17.** Dissociation constants for the binding of DMHBI<sup>+</sup> to Chili mutants. Titration curves (Figure 37) were fitted with a single-site-binding model.

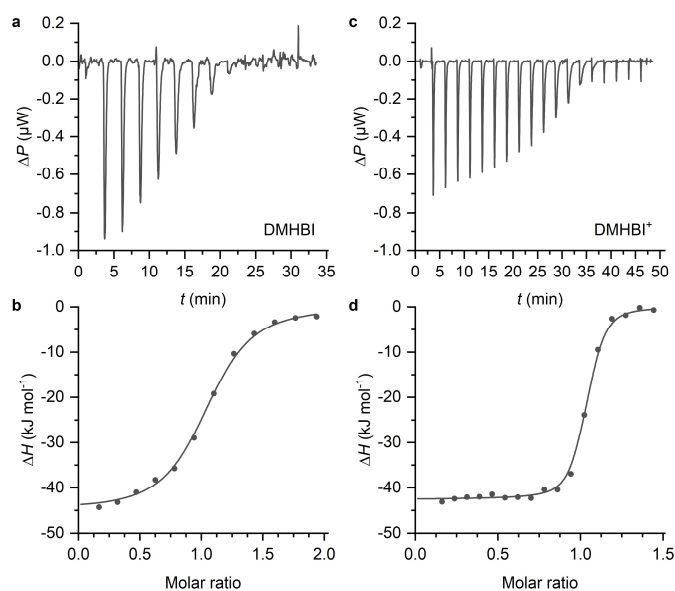
RNA	Length bp	$K_d$ $\text{mol L}^{-1}$
5 bp stem	46	$6.8 \pm 0.3 \times 10^{-8}$
5 bp mini	42	$8.9 \pm 0.7 \times 10^{-8}$

### 2.4.4 Calorimetric analysis of binding thermodynamics

In case of fluorogenic aptamers, affinity measurement by fluorescence titration is the usual and obvious approach, generating a reliable signal even at low analyte concentrations. However, determining the thermodynamic parameters ( $\Delta H^0$ ,  $\Delta S^0$ ,  $\Delta G^0$ ) that govern the binding equilibrium, as described by the dissociation constant ( $K_d$ ), would require a series of titrations to be performed at different temperatures; a van't Hoff plot of  $\ln K_d$  vs  $T^{-1}$  then allows calculation of  $\Delta H^0$ ,  $\Delta S^0$  from the slope and y-axis intercept, respectively, of a linear fit.<sup>172</sup> More directly, these parameters can be obtained, in addition to the dissociation constant and stoichiometry, by measuring the heat change associated with the complex formation via isothermal titration calorimetry (ITC, see section 4.1.16). Knowledge about the thermodynamic parameters is useful for gaining insight into

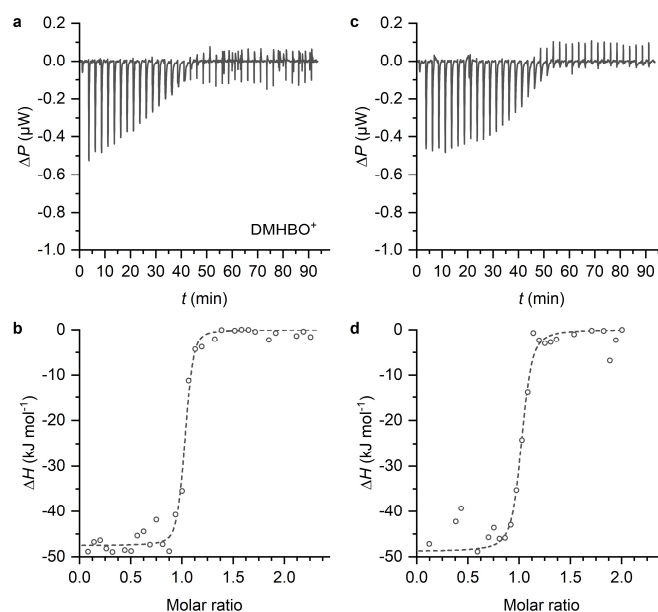
the binding mechanism of the system under investigation and can potentially help to identify distinct modes of interaction for different ligands.<sup>174</sup> While ITC has already been used extensively to study interactions between riboswitches or aptamers and small molecule ligands in the past,<sup>175,176</sup> according to a literature survey the results obtained in the course of this thesis constituted the first application of the method to a fluorogenic aptamer. Only recently, Sjekloća and Ferré-D'Amaré used ITC to show that the dimeric Corn aptamer binds the non-cognate fluorogenic ligand ThT with a 2:2 stoichiometry.<sup>162</sup>

The primary goal of using ITC for studying the Chili system was to elucidate mechanistic differences of binding to neutral or cationic ligands. Titration experiments were therefore performed with DMHBI (**13**), DMHBI<sup>+</sup> (**72**), DMHBI<sup>C</sup> (**73**) and DMHBO<sup>+</sup> (**98**). Suitable experimental parameters were determined by simulating idealized data based on  $K_a$  values from fluorescence titrations and an estimated binding enthalpy of approximately  $-50 \text{ kJ mol}^{-1}$  (typical value for RNA–ligand interactions<sup>177</sup>) using the MicroCal PEAQ-ITC software package. For typical experiments, 13 to 37 injections of ligand (100–150  $\mu\text{M}$ ) into an RNA sample (10–15  $\mu\text{M}$ ) were performed at 25 °C. The resulting thermograms were then processed and fitted to a single-site-binding model as described in section 4.1.16.1 (Figure 38).



**Figure 38.** Baseline-corrected differential power thermograms and integrated heat binding isotherms for the titration of Chili with DMHBI (a, b) or DMHBI<sup>+</sup> (c, d) at 25 °C. Data points were fitted with a single-site binding model. Conditions: 10 or 15  $\mu\text{M}$  RNA, 100 or 150  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Background correction data was obtained by titrating the ligand into RNA-free buffer. A single representative run of three is shown.

The high ligand concentration required for these experiments lead to rapid precipitation of DMHBI<sup>C</sup> from its stock solution, which precluded measurements. The same problem was encountered with DMHBO<sup>+</sup>, albeit to a lesser extent. Still, because the ligand concentration could not be kept constant during the course of the titration, thermodynamic parameters could not be determined accurately for this ligand (Figure 39).



**Figure 39.** Baseline-corrected differential power thermograms and integrated heat binding isotherms for the titration of Chili with DMHBO<sup>+</sup> at 25 °C (two independent runs). Fits with a single-site binding model (dashed lines) are affected by noisy data. Conditions: 15 μM RNA, 150 μM ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Background correction data was obtained by titrating the ligand into RNA-free buffer.

The calorimetric dissociation constants determined for DMHBI ( $3.7 \pm 0.6 \times 10^{-7} \text{ mol L}^{-1}$ ) and DMHBI<sup>+</sup> ( $2.4 \pm 0.6 \times 10^{-8} \text{ mol L}^{-1}$ ) are in good agreement with those obtained by fluorescence titration (section 2.4.3.1). Surprisingly, the difference in binding enthalpy was minimal (−44 and −42 kJ mol<sup>−1</sup>, respectively). However, while the entropic contribution is unfavorable for DMHBI (−276 J mol<sup>−1</sup> K<sup>−1</sup>), it is slightly favorable for DMHBI<sup>+</sup> (+78 J mol<sup>−1</sup> K<sup>−1</sup>), which leads to a more negative free energy change at 25 °C for the latter (Table 18). Factors governing the change in entropy that accompanies an association reaction are, among others, a decrease in the number of unbound molecules in solution, reduced conformational freedom especially for the ligand but possibly also for the macromolecule and also a change in hydration, e.g. release of water due to a smaller solvent-accessible surface area in the complex. Such a desolvation would commonly manifest itself in form of a strongly positive contribution to the binding entropy.<sup>178</sup> In the present case, this likely reflects an extended hydration shell surrounding the cationic side chain of DMHBI<sup>+</sup>, which is shed upon migration of the ligand into the size-constrained binding cavity of the Chili aptamer. The low affinities of DMHBI<sup>c</sup> as well as of the trifluoromethylated ligands **69** and **70** might therefore be caused by unfavorable binding entropies due to the hydrophobic character of their side chains.

**Table 18.** Thermodynamic data for the formation of Chili complexes from ITC experiments. In case of DMHBO<sup>+</sup>, only the enthalpy was estimated due to data quality.

Ligand		$\Delta H^0$ kJ mol <sup>−1</sup>	$\Delta S^0$ J mol <sup>−1</sup> K <sup>−1</sup>	$\Delta G_{298 \text{ K}}^0$ <sup>[a]</sup> kJ mol <sup>−1</sup>	$K_d$ mol L <sup>−1</sup>
DMHBI	<b>13</b>	−44	−276	−37	$3.7 \pm 0.6 \times 10^{-7}$
DMHBI <sup>+</sup>	<b>72</b>	−42	+75	−44	$2.4 \pm 0.6 \times 10^{-8}$
DMHBO <sup>+</sup>	<b>98</b>	−49	—	—	—

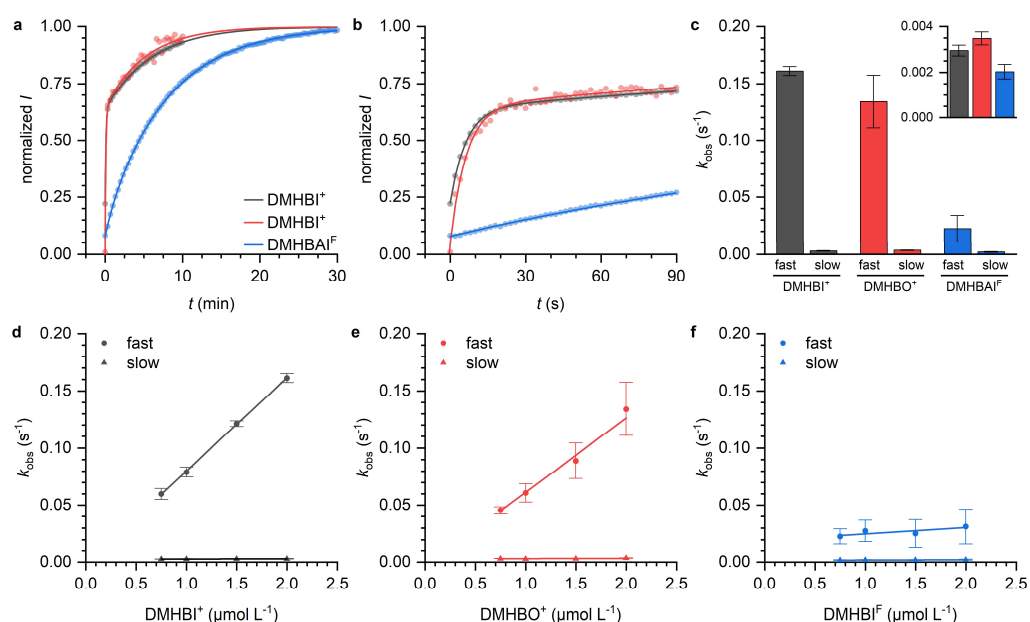
[a]  $\Delta G^0 = \Delta H^0 - T\Delta S^0$ .

For DMHBO<sup>+</sup>, an approximate binding enthalpy of −49 kJ mol<sup>−1</sup> was estimated from the initial parts of two binding isotherms (Figure 39b, d). Compared with the binding enthalpy measured for DMHBI<sup>+</sup>, this finding is consistent with the earlier interpretation that the oxime moiety can engage in an additional noncovalent interaction.

### 2.4.5 Fluorescence activation kinetics

Binding and unbinding kinetics of the fluorogenic aptamers Spinach and Corn with their cognate ligands have previously been used to substantiate affinities calculated from equilibrium binding assays. These data also showed that the respective complexes remain intact for prolonged periods after photoexcitation.<sup>86,88</sup>

In order to gain insight into the dynamics of the Chili system, the aptamer was incubated with DMHBI<sup>+</sup> under pseudo-first order conditions, i.e. using a large excess of ligand over RNA, which resulted in a gradual fluorescence increase (Figure 40a, b). Even though the equilibrium binding assays discussed above showed the formation of a simple 1:1 complex between the two components, the time course of fluorescence activation was unexpectedly complex, requiring a biexponential model for satisfactory fitting of the data. The same behavior was observed when using ligands with a higher (DMHBO<sup>+</sup>) or lower (DMHBAI<sup>F</sup>) affinity than DMHBI<sup>+</sup>, demonstrating that this is a general property of the aptamer. In each case, one of the apparent rate constants ( $k_{\text{obs}}$ ) corresponded to a fast process and was dependent on ligand concentration, whereas the other one was associated with a slow, ligand-independent process (Figure 40c).



**Figure 40.** a) Fluorescence activation kinetics of Chili with three different ligands under pseudo-first order conditions (25 nM RNA, 2  $\mu\text{M}$  ligand). Data points were collected in 2 s intervals for up to 30 min, every 10<sup>th</sup> point is plotted. Solid lines represent biexponential fits. b) First 90 s of the data in panel a without decimation. c)  $k_{\text{obs}}$  values obtained from the biexponential fits in panel a. The inset shows only  $k_{\text{obs}}$  for the slower process. d–f) Plot of  $k_{\text{obs}}$  vs ligand concentration. Solid lines represent linear fits, the slope of which is  $k_{\text{on}}$ . Conditions: 0.025 nM RNA, ligand as indicated, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Bimolecular association rate constants ( $k_{\text{on}}$ ) for the fast process were extracted from the concentration dependence of  $k_{\text{obs}}$  by linear fitting (Figure 40d–f). The corresponding dissociation rate constants ( $k_{\text{off}}$ ) were then calculated according to  $k_{\text{off}} = k_{\text{on}} \times K_{\text{d}}$  using the dissociation constants listed in Table 16. Even though the rate constants were affected by error propagation from the determination of  $k_{\text{obs}}$  by multiexponential fitting, it becomes clear that both DMHBI<sup>+</sup> and DMHBO<sup>+</sup> bind with comparable rates on the order of  $10^4 \text{ L mol}^{-1} \text{ s}^{-1}$  (Table 19), which is typical for the interaction of small molecules with RNA aptamers and riboswitches<sup>179</sup> and mirrors the binding rates of Spinach–DFHBI and Corn–DFHO.<sup>86,88</sup> In agreement with its considerably lower affinity, DMHBAI<sup>F</sup> binds slower than the other two ligands by an order of magnitude, which is consistent with the earlier interpretation that the lipophilic character of its side chain might hinder uptake

into the binding pocket of the aptamer. While  $k_{\text{off}}$  is similar for DMHBI<sup>+</sup> and DMHBAl<sup>F</sup>, the oxime ligand DMHBO<sup>+</sup> dissociates approximately ten times slower from the binding site with a residence time ( $k_{\text{off}}^{-1}$ ) upwards of 20 min. This presumably reflects the presence of an additional noncovalent interaction between the oxime moiety and the ligand binding site as discussed earlier (see section 2.4.4).

**Table 19.** Binding and unbinding rate constants for the formation of Chili aptamer complexes.

Ligand		$k_{\text{on}}$ L mol <sup>-1</sup> s <sup>-1</sup>	$k_{\text{off}}^{[a]}$ s <sup>-1</sup>
DMHBAl <sup>F</sup>	<b>70</b>	$5.5 \pm 2.9 \times 10^3$	$8.3 \pm 5.4 \times 10^{-3}$
DMHBI <sup>+</sup>	<b>72</b>	$8.2 \pm 0.1 \times 10^4$	$5.1 \pm 0.6 \times 10^{-3}$
DMHBO <sup>+</sup>	<b>98</b>	$6.5 \pm 0.5 \times 10^4$	$7.8 \pm 1.9 \times 10^{-4}$

[a]  $k_{\text{off}} = k_{\text{on}} \times K_{\text{d}}$ .

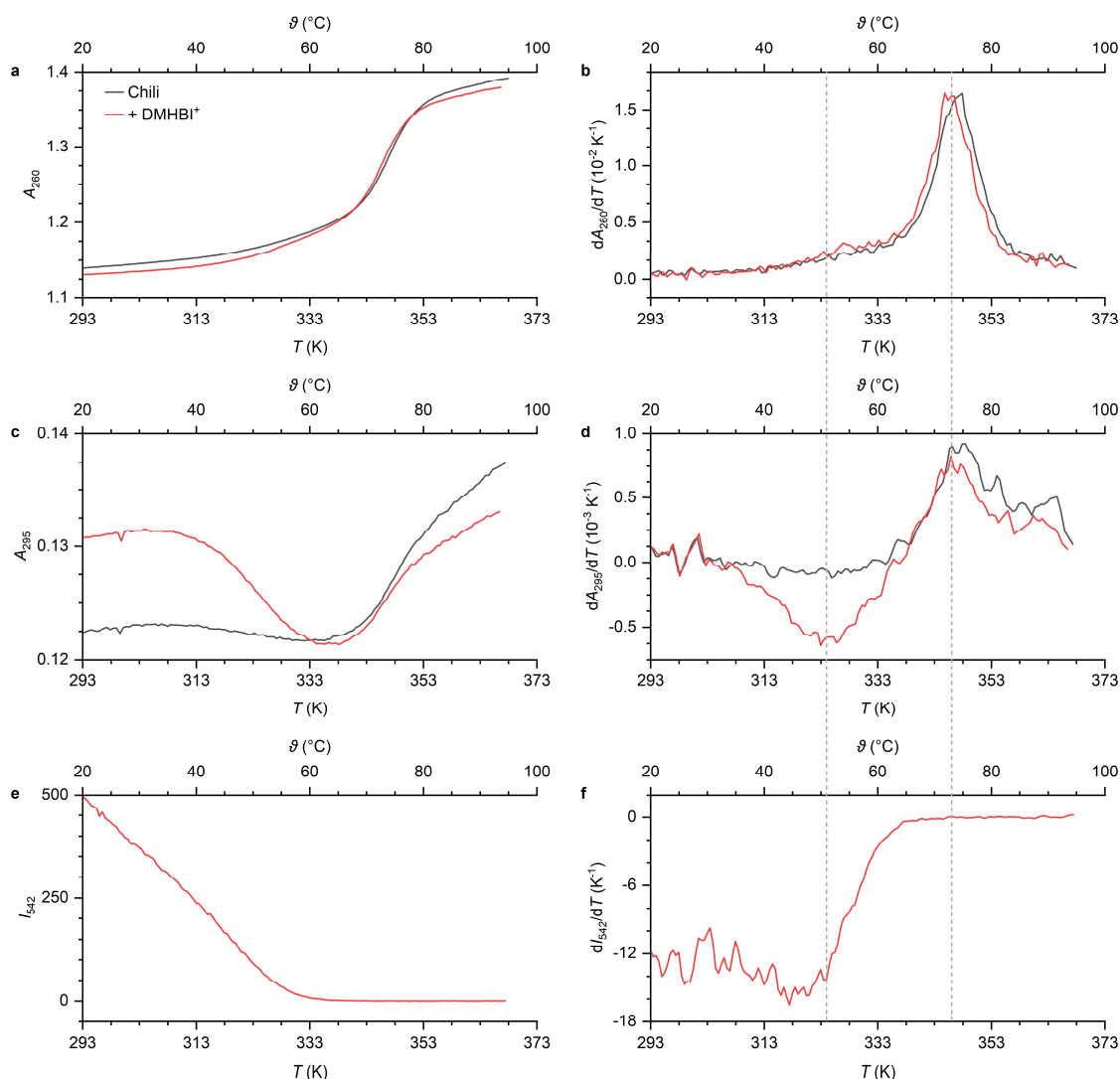
While the fast process obviously corresponds to the direct activation of the ligand into the binding site of the aptamer, the precise nature of the slow process is more elusive. Jaffrey and coworkers recently demonstrated that by switching the ligand from DFHBI to BI (see Table 7 on page 39), the fluorescence activation kinetics of the Broccoli aptamer change from monophasic to biphasic. It was speculated that BI is able to convert a non-fluorogenic conformation of the aptamer to an active species after binding.<sup>127</sup> As an alternative to such an induced fit mechanism, slow fluorescence enhancement could also occur by conformational selection, i.e. sequestration of an active RNA conformation from two or more slowly interconverting folding states due to trapping by ligand binding. Because no ligand concentration dependence of the associated  $k_{\text{obs}}$  was observed, these mechanisms could not be distinguished experimentally in the present case.<sup>180,181</sup> However, in light of the well-defined folding behavior displayed by the Chili aptamer (see section 2.2.2), induced fit appears to be the more probable mechanism and was corroborated by thermal denaturation studies described in the following section.

#### 2.4.6 Ligand-induced stabilization of the binding site

One of the factors governing aptamer applicability for cellular studies is the brightness of the fluorescent complex at the elevated temperatures required for cell culture, which encompasses thermal denaturation of the RNA and temperature-dependent deactivation of the fluorogenic ligand. This thermal stability is routinely evaluated by measuring the fluorescence intensity of an aptamer–ligand complex over a range of temperatures and can be compared in terms of an apparent melting temperature, at which the intensity drops to half of the value measured at 20 °C. For example, DFHBI complexes of the Spinach aptamer and its superfolder variant Spinach2 show apparent melting temperatures of 35 and 38 °C, respectively.<sup>78</sup> Recently, it was shown that the nature of the ligand can have a major influence in this regard; by changing the ligand of the Broccoli aptamer from DFHBI-1T (**19**, see Table 7) to the optimized derivative BI, the apparent melting temperature rose from 30 °C to 42 °C.<sup>127</sup> By considering changes not only to fluorescence emission, but also to the UV absorbance of the Spinach and Mango aptamers themselves and of the respective complexes with TO1-biotin, Unrau and coworkers showed that Spinach folds into its preferred conformation even in the absence of ligand, whereas the structure of Mango changes substantially with ligand uptake.<sup>97</sup>

Thermal denaturation experiments with the Chili aptamer were performed using the RNA alone and in complex with DMHBI<sup>+</sup>. The cationic ligand was chosen because of its bright fluorescence even in the absence of Mg<sup>2+</sup> ions; the Lewis-acidic character of Mg<sup>2+</sup> could otherwise lead to backbone scission at elevated temperatures.<sup>182</sup> In addition to fluorescence emission, UV

absorbance was recorded at 260 nm to monitor the folding state of canonical structure elements (i.e. duplex and hairpin domains) as well as at 295 nm, which would reveal changes to a putative G-quadruplex (see sections 4.1.18 and 4.1.19 for details).<sup>183</sup>

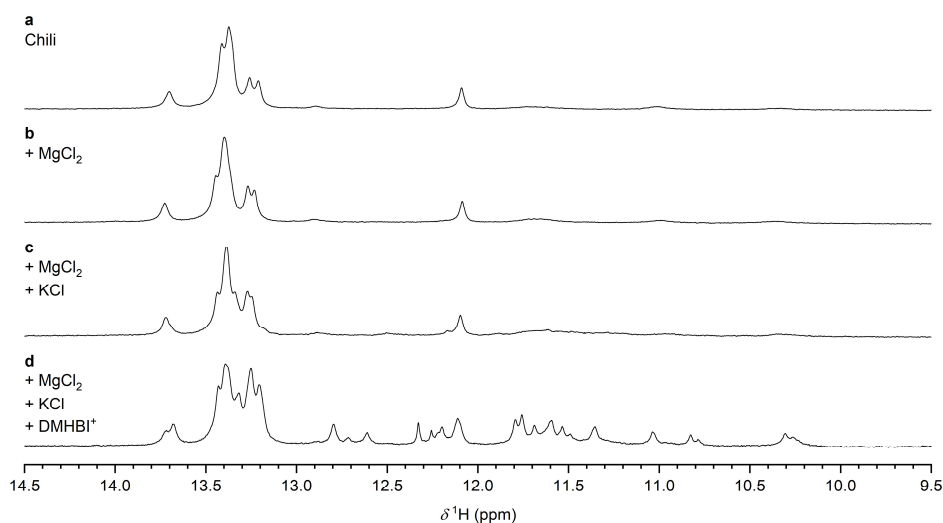


**Figure 41.** Thermal melting profiles of the Chili aptamer alone (black) and in complex with DMHBI<sup>+</sup> (red). a) Absorbance and b) first derivative of the absorbance at 260 nm showing a hyperchromic transition at 73 °C. c) Absorbance and d) and first derivative of the absorbance at 295 nm. The RNA–ligand complex shows a prominent hypochromic transition at 51 °C. e) Fluorescence emission and f) first derivative of the fluorescence emission ( $\lambda_{\text{ex/em}} = 413/542$  nm). Loss of fluorescence coincides with the transition occurring at 51 °C. Conditions: 2  $\mu\text{M}$  RNA, 125 mM KCl, 40 mM HEPES pH 7.5, + 2  $\mu\text{M}$  ligand. Curves are shown for the second of four reversible temperature ramps.

The UV absorbance at 260 nm shows a sharp hyperchromic transition at 73 °C (Figure 41a, b) for both Chili and Chili–DMHBI<sup>+</sup>, which is assigned to the dissociation of the basal and apical stems (Chart 19 on page 29). At 295 nm, an additional hypochromic transition is observed at 51 °C only in the presence of DMHBI<sup>+</sup> (Figure 41c, d). This highly characteristic decrease in absorbance indicates the ligand-induced formation of a G-quadruplex and is directly correlated with the degree of fluorescence enhancement (Figure 41e, f). No fluorescence is observed above 63 °C, at which point the quadruplex, i.e. the core of the ligand binding site, is fully denatured; because dissociation of the complex precedes melting of Watson-Crick secondary structure motifs, the absorbance profile at 260 nm is unaffected by the presence of DMHBI<sup>+</sup>.



These findings were supported by  $^1\text{H}$  NMR spectroscopy of the Chili aptamer (experiments performed by Irene Bessi in this working group), which showed that the ligand binding site is not preorganized in the absence of the ligand. Watson-Crick base pairs give rise to imino proton resonances in the region of 12–15 ppm, whereas imino protons in noncanonical structure motifs, such as G-quartets, are typically more shielded and yield signals between 10 and 12 ppm.<sup>184,185</sup> Prior to ligand addition, imino resonances were observed at > 12 ppm, which indicates base pair formation in the basal and apical helices of Chili (Figure 42a). Addition of  $\text{Mg}^{2+}$  (1 mM) and  $\text{K}^+$  (50 mM) did not have a notable effect on these signals, suggesting that these metal ions by themselves are not sufficient for inducing a structural change of the binding site (Figure 42b, c). After  $\text{DMHBI}^+$  was added, several new imino resonances appeared, which is consistent with the ligand-dependent formation of a G-quadruplex (Figure 42d).<sup>112</sup> In contrast, the Spinach and Corn aptamers were shown to retain G-quadruplex domains even in their apo state.<sup>81,162</sup>



**Figure 42.** Imino  $^1\text{H}$  NMR spectra of the Chili aptamer a) before and b–d) after sequential addition of  $\text{MgCl}_2$ ,  $\text{KCl}$  and  $\text{DMHBI}^+$ . Conditions: 150  $\mu\text{M}$  RNA, 25 mM TRIS pH 7.4, 10%  $\text{D}_2\text{O}/90\%$   $\text{H}_2\text{O}$ , + 5 mM  $\text{MgCl}_2$ , + 50 mM  $\text{KCl}$ , + 150  $\mu\text{M}$  ligand. Metal ions alone were not sufficient to induce notable changes, whereas addition of the ligand resulted in the appearance of several new imino resonances.<sup>112</sup> NMR spectra were recorded by Irene Bessi.

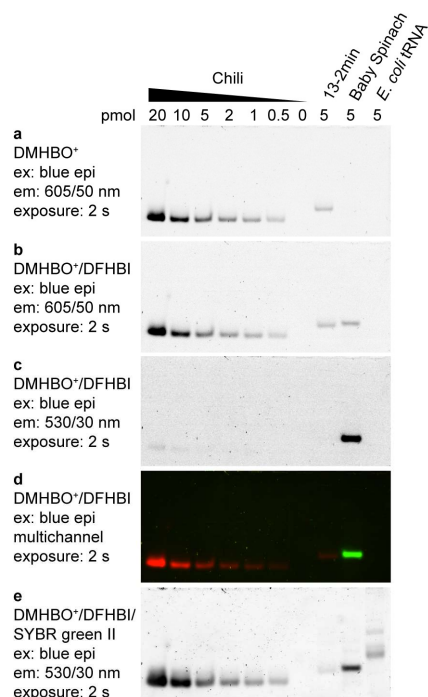
Overall, the lack of hypochromicity at 295 nm prior to addition of  $\text{DMHBI}^+$  in conjunction with the biphasic time course of fluorescence activation (section 2.4.5) and the structural changes observed by  $^1\text{H}$  NMR spectroscopy strongly supports binding of the ligand by an induced fit mechanism.

## 2.5 AN APTAMER-TAG FOR GEL-BASED FLUORESCENCE QUANTIFICATION

### 2.5.1 In-gel visualization of the Chili aptamer

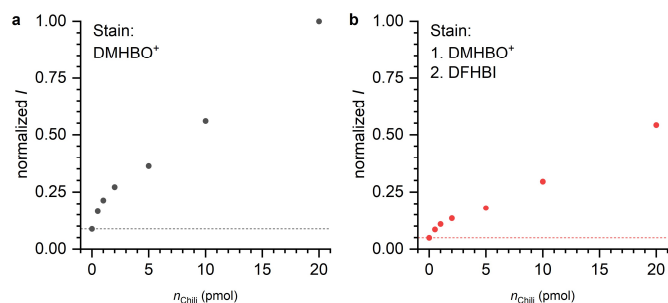
In-gel fluorescent staining of aptamer-tagged RNAs of interest has been proposed as a simple and time-efficient alternative to northern blotting for the detection of a specific RNA. Jaffrey and coworkers were able to demonstrate that RNA amounts greater than approximately 100 pg can be visualized on a denaturing polyacrylamide gel by using Broccoli–DFHBI-1T as a fluorescent label.<sup>186</sup> The same group later showed that Broccoli variants and the Corn aptamer are also amenable to in-gel staining with the ligand DFHO.<sup>88</sup> Inspired by the high affinity of the Chili aptamer for the DsRed-type ligand DMHBO<sup>+</sup> combined with its unique spectral properties, the utility of this system for in-gel fluorescence imaging of RNA was to be investigated.

As a first step, varying amounts of Chili, its unoptimized precursor 13-2min, the DFHBI-binding 51 nt Baby Spinach aptamer<sup>82</sup> and tRNA from *E. coli* as a negative control were run on a 10% native polyacrylamide gel, which was then stained with a 1  $\mu$ M solution of DMHBO<sup>+</sup> for 15 min at ambient temperature without prior washing. Blue epi illumination (broadband, centered at 475 nm) in combination with a 605/50 nm emission filter (red channel) allowed facile detection of Chili RNA (0.5–20 pmol, 6.9–274 ng) without contamination from nonspecific background fluorescence (Figure 43a). Fluorescence from 13-2min was relatively weak due to its poor folding behavior (see section 2.2.2). As expected, neither Baby Spinach nor control tRNA activated DMHBO<sup>+</sup> fluorescence (see also section 2.4.1.1). Next, the gel was counterstained using 1  $\mu$ M DFHBI for 15 min at ambient temperature. While the total intensity of Chili fluorescence was slightly reduced due to elution into the staining medium, the overall quantification remained unaffected (Figure 43b, 44). In the green channel (530/30 nm emission filter), strong fluorescence of Baby Spinach–DFHBI over a weak background signal was observed (Figure 43c). Crosstalk between the green and red channel was minor and allowed for orthogonal visualization of Chili and Baby Spinach fluorescence (Figure 43d). Finally, the gel was stained a third time with SYBR Green II for 15 min at ambient temperature, which revealed the presence of control tRNA (Figure 43e).



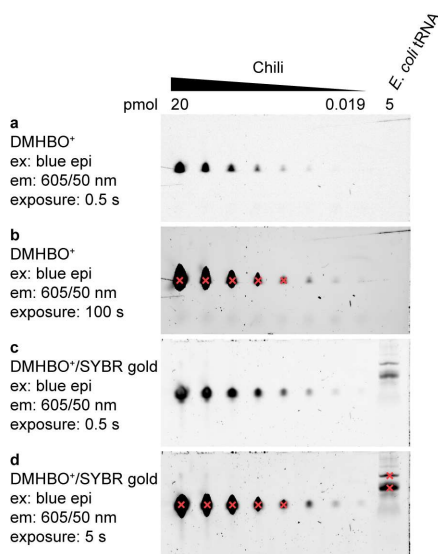
**Figure 43.** The indicated amounts of the Chili, 13-2min and Baby Spinach aptamers as well as *E. coli* tRNA were run on a 10% native polyacrylamide gel at a constant voltage of 200 V for 20 min in TBE. The gel was subsequently stained with 1  $\mu$ M DMHBO<sup>+</sup> in binding buffer, 1  $\mu$ M DFHBI in binding buffer and 1x SYBR green II in TBE for 15 min each without intermediate washing steps and imaged in between. The binding buffer contained 125 mM KCl, 5 mM MgCl<sub>2</sub> and 40 mM HEPES pH 7.5.

- After staining with DMHBO<sup>+</sup>, blue epi illumination, 605/50 nm emission filter.
- After counterstaining with DFHBI, blue epi illumination, 605/50 nm emission filter.
- Same as the previous image, 530/30 nm emission filter.
- Combined false color representation of panels c (red channel) and d (green channel).
- After staining with SYBR green II, blue epi illumination, 530/30 nm emission filter. The tRNA had not been visible before.



**Figure 44.** Quantification of Chili–DMHBO<sup>+</sup> fluorescence from panels a and b in Figure 43 a) The band intensity rises in a linear fashion for RNA amounts  $\geq 2$  pmol. b) Counter-staining with DFHBI leads to an overall lower band intensity due to elution from the gel, but the signal remains linear. Dashed lines denote the level of background fluorescence.

In-gel staining of Chili with DMHBO<sup>+</sup> was also possible under denaturing conditions. Using an appropriate exposure time, 19 fmol of the aptamer (260 pg) were detected (Figure 45). Encouraged by these results, the next step was to tag and visualize an RNA of interest with the Chili aptamer.

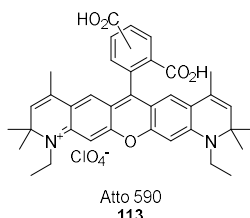


**Figure 45.** A concentration series of the Chili aptamer (starting with 20 pmol, eight steps, 2.7-fold dilution) and *E. coli* tRNA (5 pmol) were run on a 10% denaturing polyacrylamide gel at a constant voltage of 200 V for 20 min in TBE. The gel was subsequently stained with 1  $\mu$ M DMHBO<sup>+</sup> in binding buffer and 1x SYBR gold in TBE for 15 min each without intermediate washing steps and imaged in between. The binding buffer contained 125 mM KCl, 5 mM MgCl<sub>2</sub> and 40 mM HEPES pH 7.5.

- After staining with DMHBO<sup>+</sup>, blue epi illumination, 605/50 nm emission filter.
- Same as the previous image with longer exposure.
- After staining with SYBR gold, blue epi illumination, 605/50 nm emission filter. The tRNA had not been visible before.
- Same as the previous image with longer exposure. Overexposed bands are marked in red.

### 2.5.2 A two-color FRET reporter for monitoring DNazyme activity.

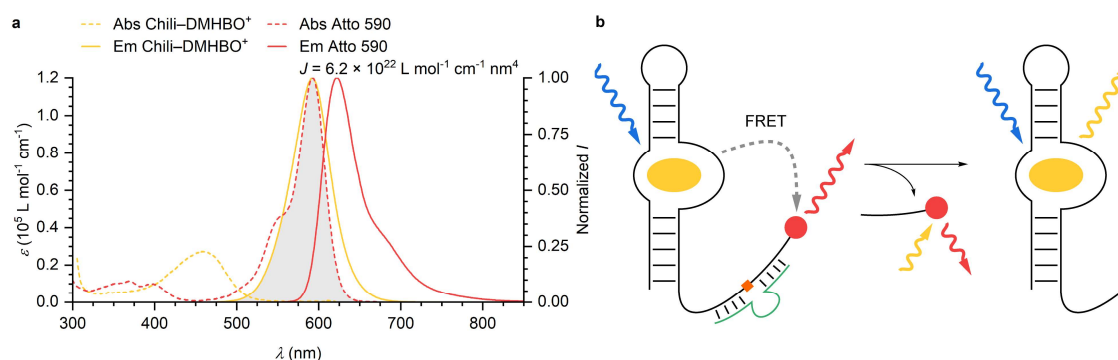
For investigating the utility of Chili as a fluorescent reporter, a model system was designed based on a 20 nt fragment of a 25S rRNA that is part of the large ribosomal subunit in *S. cerevisiae*.<sup>187</sup> A 10-23 DNazyme with appropriate complementary binding arms accepts this oligonucleotide as its substrate and catalyzes strand cleavage between G9 and U10 in the presence of Mg<sup>2+</sup>. The RNA substrate is 3'-labeled with the fluorescent dye Atto 590 (**113**, Chart 22) and tagged with Chili at its 5'-end.



**Chart 22.** Structure of the fluorescent dye Atto 590.

Due to their large spectral overlap, Chili–DMHBO<sup>+</sup> and Atto 590 form an excellent FRET pair (Figure 46a, see section 4.1.21.1 for background information on FRET). Before cleavage, excitation of the Chili complex at 456 nm would therefore result in donor-sensitized emission from Atto 590

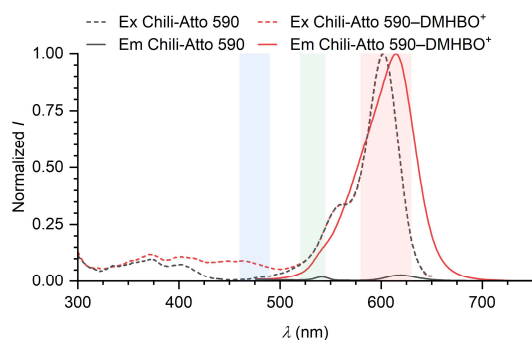
at 622 nm, whereas the cleavage product would only display direct fluorescence emission from DMHBO<sup>+</sup> at 592 nm (Figure 46b).



**Figure 46.** a) Absorption and fluorescence emission spectra of Chili-DMHBO<sup>+</sup> and Atto 590. The spectral overlap between donor emission and acceptor absorption is shaded. The overlap integral was calculated as described in section 4.1.21.1. b) Schematic representation of the FRET reporter construct. Excitation of Chili-bound DMHBO<sup>+</sup> (yellow) leads to donor-sensitized emission from 3'-Atto 590 (red). The RNA substrate is cleaved site-specifically (orange) by a 10-23 DNAzyme, thereby removing the Atto 590 label. Afterwards, direct fluorescence from the bound ligand can be observed.

The 10-23 DNAzyme and 3'-amino-functionalized RNA substrate (**RS-NH<sub>2</sub>**) were prepared by automated solid-phase synthesis using appropriate CPG supports and purified according to standard methods. Next, T4 polynucleotide kinase was used to 5'-phosphorylate **RS-NH<sub>2</sub>**. The product was isolated by ethanol precipitation and coupled with Atto 590-NHS ester in DMF at 37 °C for 3 h. After ethanol precipitation, the crude p-**RS**-Atto 590 was ligated to the Chili aptamer using T4 RNA ligase with the help of a 24 nt DNA splint oligonucleotide to generate Chili-**RS**-Atto 590, which was isolated and purified by PAGE.<sup>111</sup>

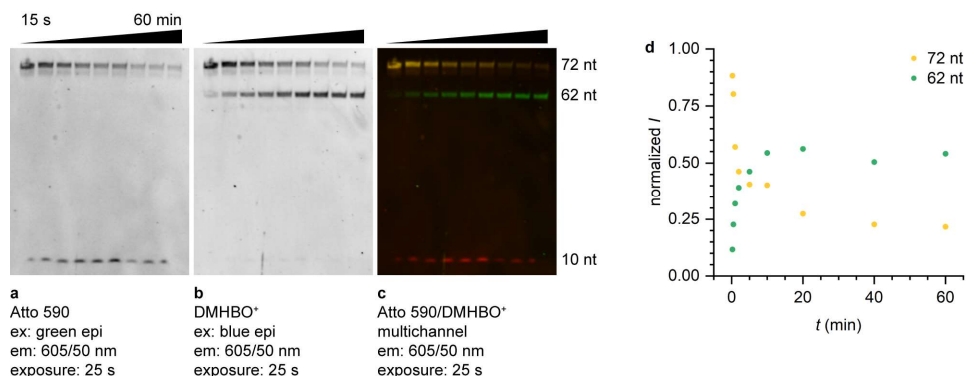
Fluorescence excitation and emission spectra of Chili-**RS**-Atto 590 were measured before and after addition of DMHBO<sup>+</sup> and demonstrate that the FRET pair is working as intended (Figure 47). The pronounced short-wavelength tail of the emission band arises from incomplete energy transfer. For a gel-based assay, blue and green epi illumination (broadband, centered at 475 and 530 nm, respectively) in combination with a 605/50 nm emission filter will allow orthogonal detection of donor-sensitized emission from Atto 590 before enzymatic cleavage, direct Chili-DMHBO<sup>+</sup> fluorescence thereafter and direct Atto 590 fluorescence at all times.



**Figure 47.** Normalized fluorescence excitation and emission spectra ( $\lambda_{ex/em} = 456/670$  nm) of Chili-Atto 590 before (black) and after (red) addition of DMHBO<sup>+</sup>. Shaded areas correspond to the excitation and emission windows used for image acquisition in Figure 48. Blue epi illumination generates direct DMHBO<sup>+</sup> fluorescence and FRET emission from Atto 590, whereas green epi illumination is selective for direct Atto 590 fluorescence. Conditions: 0.5  $\mu$ M RNA, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5, + 0.5  $\mu$ M ligand.

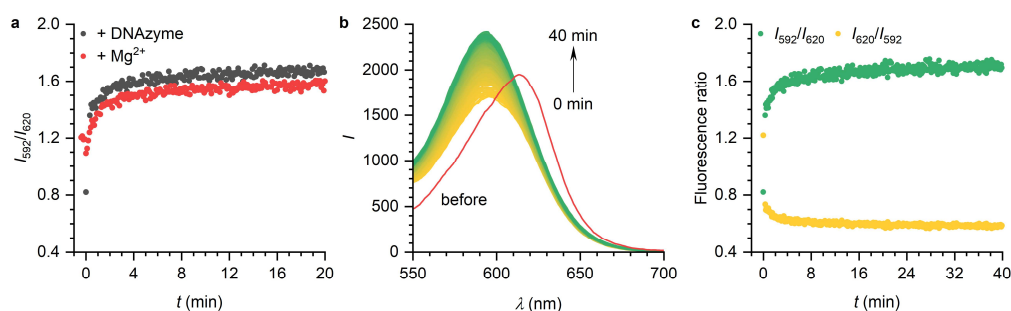
To monitor the time course of DNA-catalyzed substrate cleavage under single-turnover conditions, Chili-**RS**-Atto 590 was first annealed to the 10-23 DNAzyme in Mg<sup>2+</sup>-free buffer. The reaction was initiated by addition of 5 mM MgCl<sub>2</sub> at 25 °C; aliquots of the reaction mixture were removed at nine time points between 15 s and 60 min, quenched by addition of stop solution (formamide, TRIS, EDTA, pH 8.3) and stored at 0 °C. For analysis, the samples were resolved by denaturing PAGE (15%) and the gel was imaged to visualize the 10 nt Atto 590-labeled cleavage product. Next, the

gel was stained by treatment with a 1  $\mu\text{M}$  solution of DMHBO<sup>+</sup> for 5 min at ambient temperature and imaged again to visualize full-length Chili-RS-Atto 590 as well as the 62 nt fragment resulting from cleavage (Figure 48).



**Figure 48.** Samples from the 10-23 DNAzyme-catalyzed cleavage of Chili-RS-Atto 590 (time points taken after 15 s, 30 s, 1 min, 2 min, 5 min, 10 min, 20 min, 40 min and 60 min) were resolved on a 15% denaturing polyacrylamide gel at a constant power of 20 W for 45 min in TBE. The gel was imaged before and after staining with 1  $\mu\text{M}$  DMHBO<sup>+</sup> in binding buffer containing 125 mM KCl, 5 mM MgCl<sub>2</sub> and 40 mM HEPES pH 7.5 for 5 min. a) Before staining, green epi illumination, 605/50 nm emission filter. b) After staining with DMHBO<sup>+</sup>, blue epi illumination, 605/50 nm emission. c) Combined false color representation of panels a (red channel, direct Atto 590 fluorescence) and b (green channel, direct DMHBO<sup>+</sup> fluorescence and Atto 590 FRET). Therefore, the uncleaved 72 nt construct appears yellow. d) Quantification of the 72 and 62 nt RNA bands from panel b. The sum of fractional intensities at the first time point was normalized to 1.

RNA cleavage by the DNAzyme was also measured in a cuvette-based assay. First, it was confirmed that the reaction can be initiated equally well by addition of either Mg<sup>2+</sup> or the DNAzyme to Chili-RS-Atto 590–DMHBO<sup>+</sup> (Figure 49a). In this case, the use of DMHBO<sup>+</sup> was especially advantageous due to its low Mg<sup>2+</sup> requirement for fluorescence activation (see section 2.4.1.3). Fluorescence spectra ( $\lambda_{\text{ex}} = 456 \text{ nm}$ ) were recorded in 6 s intervals after addition of the DNAzyme (Figure 49b) and the ratio of fluorescence emission at 592 and 620 nm, which is a measure of regular fluorescence vs FRET, was monitored over the course of 40 min (Figure 49c). Here, it is evident that binding of the DNAzyme to the Chili construct immediately leads to a change in FRET efficiency (red spectrum in Figure 49b) followed by gradual cleavage of the substrate with concomitant loss of FRET.



**Figure 49.** a) Time course of regular vs FRET fluorescence of Chili-RS-Atto 590–DMHBO<sup>+</sup> cleavage by the 10-23 DNAzyme after initiation by addition of either Mg<sup>2+</sup> or the catalytic DNA. b) Emission spectra of Chili-RS-Atto 590–DMHBO<sup>+</sup> prior to cleavage by the 10-23 DNAzyme (red) and at 6 s intervals during the reaction (yellow–green). The initial shift of the spectral profile is caused by the formation of the enzyme–substrate complex. c) The ratio of fluorescence intensity at 592 and 620 nm in the spectra from panel b shows depletion of the FRET-active species (yellow) and simultaneous accumulation of the cleavage product (green). Conditions: 0.5  $\mu\text{M}$  RNA, 0.5  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5, + 5  $\mu\text{M}$  DNAzyme,  $\lambda_{\text{ex}} = 456 \text{ nm}$ .

Overall, this proof-of-concept study demonstrates that Chili is viable as a fluorescent marker for studying the fate of an RNA of interest. For assays performed purely in vitro, the capability of FRET

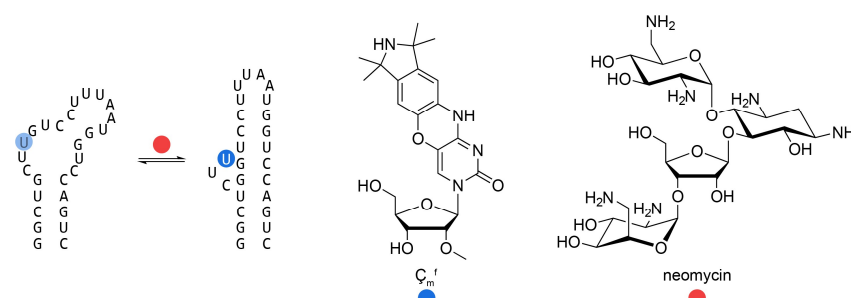
within a chemically labeled construct allows for highly specific quantification, whereas the in-gel results show great promise for future applications as a genetically encodable tag.

## 2.6 DEVELOPMENT OF A SUPRAMOLECULAR FRET PLATFORM

### 2.6.1 FRET studies of RNA structure and function

#### 2.6.1.1 Motivation

So far, the relationships between structure and fluorescence activation at work in Chili–HBI complexes have been observed by exploiting a self-reporting system, i.e. by monitoring differences in fluorescence enhancement upon ligand binding. Further insight could be gained from using an additional independent means of signal readout. Interactions between structured RNAs and their cognate ligands have previously been studied using a variety of fluorescence-based techniques. A notable example is the 2APfold approach, which involves incorporation of the fluorescent nucleobase analog 2-aminopurine (2AP) into the aptamer domains of riboswitches.<sup>188</sup> Suitable modification sites are determined based on crystallographic data or from chemical probing, e.g. by SHAPE.<sup>189</sup> 2AP fluorescence is determined by the sequence context, hybridization and solvent accessibility of the local environment. Conformational changes, which are induced by metal ion or ligand binding, modulate the 2AP emission. The resulting time-dependent increases or decreases of the observed brightness yield both apparent binding affinities and also timescales of structural rearrangement. 2APfold has been applied to the study of riboswitches for, among others, adenine<sup>190</sup>, preQ<sub>1</sub><sup>191</sup> and thiamine pyrophosphate<sup>192</sup> and gave valuable insights into their regulatory mechanisms, interconversion of different folding states and domain-dependent reorganization dynamics. This concept has recently been expanded to the use of other fluorescent nucleosides such as the extended cytosine analog  $\zeta_m^f$  (**114**), which was site-specifically incorporated into the neomycin aptamer (Figure 50). Stopped-flow fluorescence kinetics confirmed that binding between the aptamer and its ligand (**115**) proceeds via a conformational selection mechanism.<sup>193</sup>

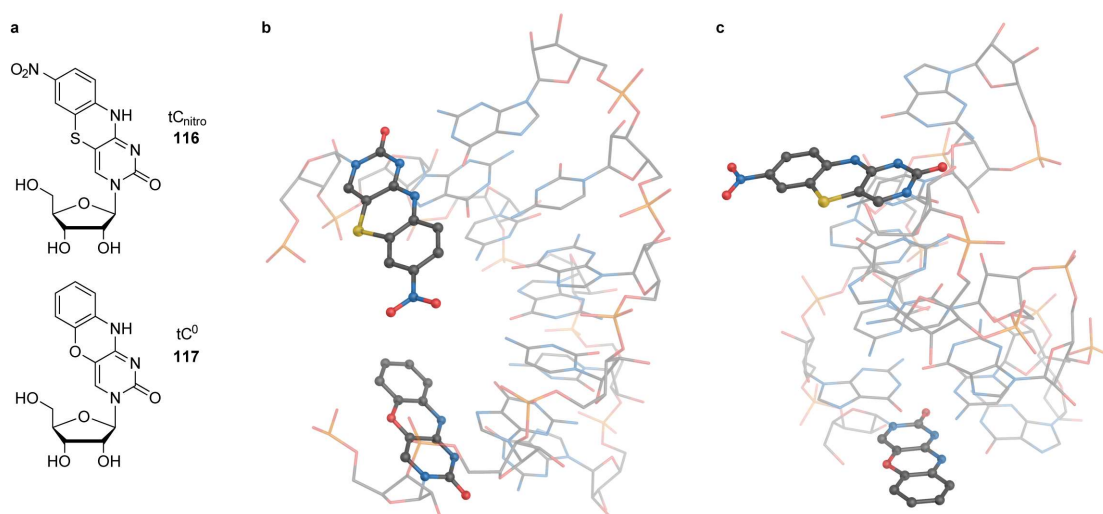


**Figure 50.** Conformational changes upon binding between the neomycin aptamer and its cognate ligand (**115**, red) were monitored by fluorescence enhancement of  $\zeta_m^f$  (**114**, blue), which was incorporated into a region of the aptamer that experiences a major structural rearrangement upon complex formation.<sup>193</sup>

#### 2.6.1.2 Base–base FRET reveals global structural information

Wilhelmsson and coworkers have demonstrated that FRET between two modified nucleobase analogs can also be used to directly obtain structural information about an oligonucleotide. By incorporating both an emissive donor and a non-emissive acceptor into DNA, position-dependent quenching of the donor fluorescence was observed and used to quantitatively model the distance and mutual orientation of the FRET pair within a B-form duplex.<sup>194,195</sup> Even small structural changes to a duplex, such as a decrease in helical twist and rise caused by interaction with a minor groove binder, were modelled by this approach.<sup>196</sup> By using an emissive FRET acceptor, Sugiyama and coworkers were able to visually monitor a colorimetric change stemming from a B–Z-form DNA transition that resulted in a decrease of the FRET efficiency.<sup>197</sup> Until very recently, base–base FRET pairs have only been constructed in DNA. The Wilhelmsson group reported the first application of this concept to the study of RNA using two isomeric tricyclic C analogs, the emissive donor tC<sup>0</sup> (**116**) and the non-emissive acceptor tC<sub>nitro</sub> (**117**). When incorporated into an

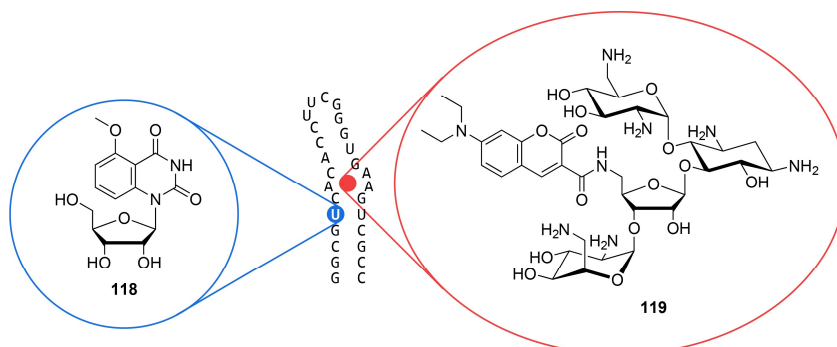
RNA duplex rich in GC repeats, a  $\text{NaClO}_4$ -induced transition from the regular A-form to the rare Z-form manifested itself in systematic decreases of the transfer efficiencies at various donor–acceptor distance. The observed structural changes are consistent with the few X-ray crystallography- and NMR-based studies of Z-form RNA (Figure 51).<sup>198</sup>



**Figure 51.** Models of  $(XGCGCG)_2$  RNA duplexes with  $X = tC^0$  or  $tC_{\text{nitro}}$ , respectively. This sequence motif has been employed by Wilhelmsson and coworkers to study an A–Z RNA transition.<sup>198</sup> a) Structures of the modified nucleosides. b) A-form duplex adopted from a canonical structure built using the W3DNA 2.0 server.<sup>199</sup> c) Z-Form duplex adapted from a reported NMR structure (PDB: 1T4X). Separation and orientation changes between the two modified bases brought about by an A–Z transition result in a decreased FRET efficiency.

### 2.6.1.3 Base–ligand FRET reports on binding events

Inside a fluorogenic aptamer such as Chili, a fluorescent nucleobase analog could potentially form a FRET pair with the bound ligand. One variation of this base–ligand FRET concept has been described by Tor and coworkers, who employed the isomorphous emissive U analog **118** to screen several aminoglycoside antibiotic candidates targeting the bacterial rRNA A-site. The aminoglycosides were derivatized with 7-*N,N*-diethylaminocoumarin (**119**), which forms a FRET pair with **118** upon binding to an A-site model oligonucleotide (Figure 52).



**Figure 52.** A 27 nt A-site construct was modified with an emissive U analog (blue). Binding to a fluorescently labeled neomycin derivative (red) resulted in base–ligand FRET.<sup>193</sup>

Titration experiments were used to determine  $EC_{50}$  values from the decrease of donor emission and concomitant rise in donor-sensitized acceptor emission. Preformed FRET-active complexes allowed observing the competitive displacement of bound aminoglycosides.<sup>200</sup>

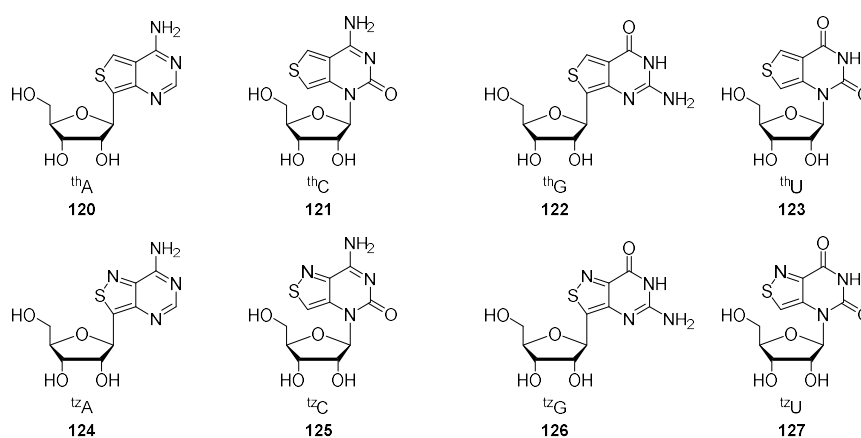
### 2.6.2 Choice of the FRET donor

The examples presented in the preceding section highlight the utility of developing a base–ligand FRET pair for the Chili aptamer. Due to their strong affinity for Chili,  $\text{DMHBI}^+$  and  $\text{DMHBO}^+$  are



obvious choices for the role of the energy transfer acceptor. A suitable donor fluorophore is then characterized by a high quantum yield and large spectral overlap with either ligand, which is quantified in terms of the overlap integral  $J$ . In practice, the aforementioned FRET pair  $tC^0$ – $tC_{\text{nitro}}$ , which has an overlap integral of  $1.7 \times 10^{14} \text{ nm}^4 \text{ L mol}^{-1} \text{ cm}^{-1}$  with a mean donor quantum yield of approximately 0.22 in double-stranded RNA, displayed observable energy transfer for distances up to 13 bp, in A-form RNA.<sup>198,201</sup> Considering the dimensions of the 52 nt Chili aptamer and the location of the binding site within the predicted secondary structure (Chart 19), a comparable combination of quantum yield and spectral overlap was a reasonable starting point for the selection of a prospective FRET donor.

At first glance, the emissive RNA alphabet developed by Tor and coworkers, comprising isomorphous thieno[3,4-*d*]pyrimidine analogs (**120–123**) of each of the four canonical nucleosides, satisfies these conditions (Chart 23). Emission maxima are in the range of 409–453 nm with quantum yields between 0.21 and 0.46 for the monomeric nucleosides (Table 20).<sup>202</sup>



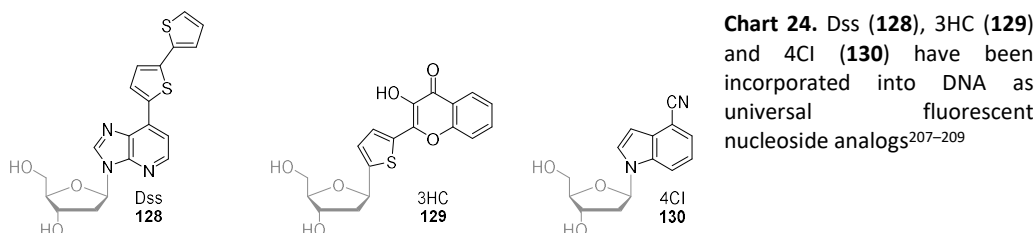
**Chart 23.** An emissive RNA alphabet (top row) was reported by Tor and coworkers.<sup>202</sup> The missing, functionally important purine-N7-like nitrogen of  $^{th}A$  (**120**) and  $^{th}G$  (**122**) was later restored in a second series of isofunctional analogs (bottom row).<sup>203</sup>

Compared to native adenosine and guanosine, analogs **120** and **122** were missing the functionally important N7 of the purine scaffold. This issue was overcome in a second-generation alphabet consisting of the isothiazolo[4,3-*d*]pyrimidine derivatives **124–127**. Here, the Hoogsteen face of the adenosine and guanosine analogs is restored, which allows them to mimic their natural counterparts more faithfully. Based on their emission maxima (392–459 nm), these compounds would also be suitable as FRET donors for Chili complexes. However, because their fluorescence quantum yields are substantially lower than those of the first-generation alphabet, the effective energy transfer distance would be decreased (Table 20).<sup>203</sup>

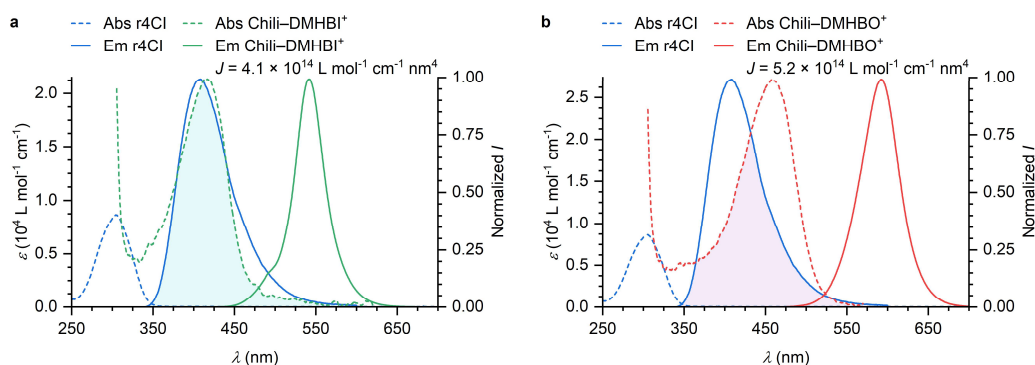
**Table 20.** Spectral properties of emissive RNA alphabets.<sup>202,203</sup>

Fluorophore	$\lambda_{\text{ex}}$ (nm)	$\lambda_{\text{em}}$ (nm)	$\phi$
<sup>th</sup> A <b>120</b>	341	420	0.21
<sup>th</sup> C <b>121</b>	320	429	0.41
<sup>th</sup> G <b>122</b>	321	453	0.46
<sup>th</sup> U <b>123</b>	304	409	0.41
<sup>tz</sup> A <b>124</b>	338	410	0.05
<sup>tz</sup> C <b>125</b>	325	411	0.05
<sup>tz</sup> G <b>126</b>	333	459	0.25
<sup>tz</sup> U <b>127</b>	312	392	0.01

To date, only <sup>th</sup>G and <sup>th</sup>U have been converted to phosphoramidite building blocks for solid-phase synthesis and studied after incorporation into RNA.<sup>202,204</sup> The other analogs would require additional characterization to establish their properties in the context of RNA oligonucleotides prior to site-specific incorporation into the Chili aptamer. Neither the nucleoside analogs themselves nor the respective phosphoramidites are commercially available at the moment and would have to be prepared in house by multi-step synthetic routes. In light of this, a reasonable alternative was to search for a universal nucleoside analog, i.e. one that is not biased towards a particular sequence or hybridization context. Universal bases often are hydrophobic aromatic systems containing fewer, if any, hydrogen bonding sites than their canonical counterparts. Instead, they non-specifically interact with surrounding bases by  $\pi$  stacking.<sup>205,206</sup> There are only few prominent examples of such universal bases displaying fluorescence and these have been restricted to the study of DNA (Chart 24).

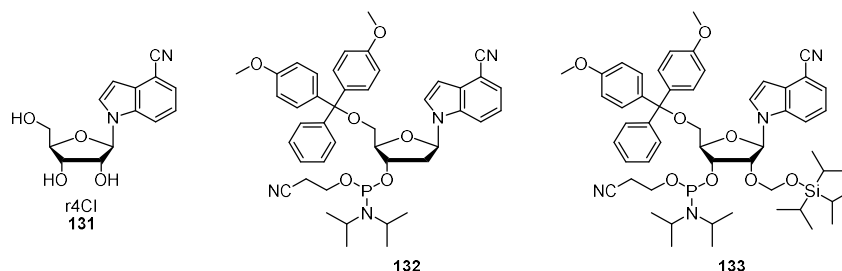


Dss (**128**) has originally been designed as part of an expanded genetic alphabet. Nucleoside triphosphates of Dss were enzymatically incorporated into DNA and RNA opposite a shape-complementary formylpyrrole base surrogate. Its high replication and transcription specificity notwithstanding, pairing of Dss with either the formylpyrrole or any of the four canonical bases did not affect the stability of modified DNA duplexes.<sup>207</sup> With an emission maximum of 442 nm, Dss could potentially act as a FRET donor for Chili–DMHBO<sup>+</sup>, but less so for the DMHBI<sup>+</sup> complex. In addition, the extended size of the modified nucleobase might have a negative impact on the folding of a highly structured RNA aptamer. The 3-hydroxychromenone derivative 3HC (**129**) exhibits dual emission due to excited state intramolecular proton transfer (ESIPT) with emission maxima in the region of 435 and 530 nm, respectively, for the individual tautomers. In most contexts, the long-wavelength band is dominant, which, in combination with variable quantum yields and severe quenching by adjacent guanines, makes 3HC unsuitable as a FRET donor.<sup>208,210</sup> 4-Cyanoindole (4CI, **130**), on the other hand, was shown to be highly emissive in a variety of sequence contexts and its sharp emission maximum at 412 nm is ideally matched to the absorption of Chili in complex with DMHBI<sup>+</sup> or DMHBO<sup>+</sup> (Figure 53).<sup>209,211</sup>



**Figure 53.** Absorption and fluorescence emission spectra of r4Cl, Chili-DMHBI<sup>+</sup> (a) and Chili-DMHBO<sup>+</sup> (b). The spectral overlap between donor emission and acceptor absorption is shaded. Overlap integrals were calculated as described in section 4.1.21.1.

In fact, despite the spectral differences between these two ligands, their overlap integral with 4CI is very similar due to the higher extinction coefficient of Chili-DMHBO<sup>+</sup>. To characterize the properties of 4CI in RNA and compare it with the DNA analog, it was therefore planned to synthesize the corresponding ribonucleoside r4Cl (**131**), the known DNA phosphoramidite **132** and also the previously unreported RNA phosphoramidite building block **133** (Chart 25).

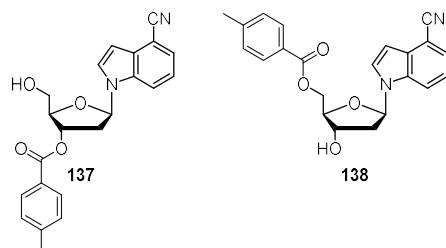


**Chart 25.** r4Cl (**131**) is the first example of a universal fluorescent ribonucleoside. While 4-cyanoindole has been incorporated into DNA by solid-phase synthesis using **132**,<sup>209</sup> the RNA phosphoramidite building block **133** has not been described previously.

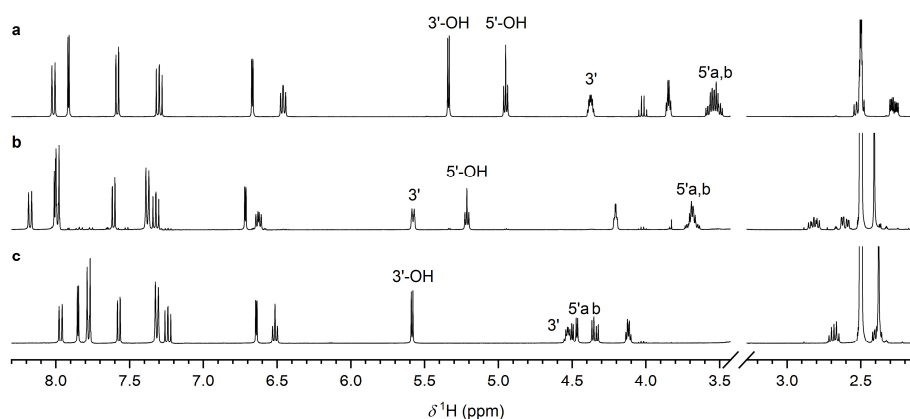
### 2.6.3 Synthesis of 4-cyanoindole phosphoramidite building block

#### 2.6.3.1 Synthesis of the DNA phosphoramidite

Following reported procedures,<sup>209</sup> **130** was deprotonated using NaH in MeCN and subsequently reacted with Hoffer's chlorosugar (**134**) for 6 h at ambient temperature, to yield the protected intermediate **135** (98%). Removal of the *p*-toluoyl groups was effected by treatment with NaOH in MeOH for 30 min at ambient temperature to afford the deoxyribonucleoside **136** (d4Cl) in 96% yield (Scheme 24). Purification of this compound by column chromatography was time-consuming as it tended to precipitate during the process. By using NaOMe instead of NaOH it was anticipated that **136** could be isolated by simple extraction rather than chromatography. However, reaction with NaOMe did not proceed to completion and the 3'-*O*- and 5'-*O*-protected isomers **137** and **138** (Chart 26, Figure 54) had to be removed by column chromatography.

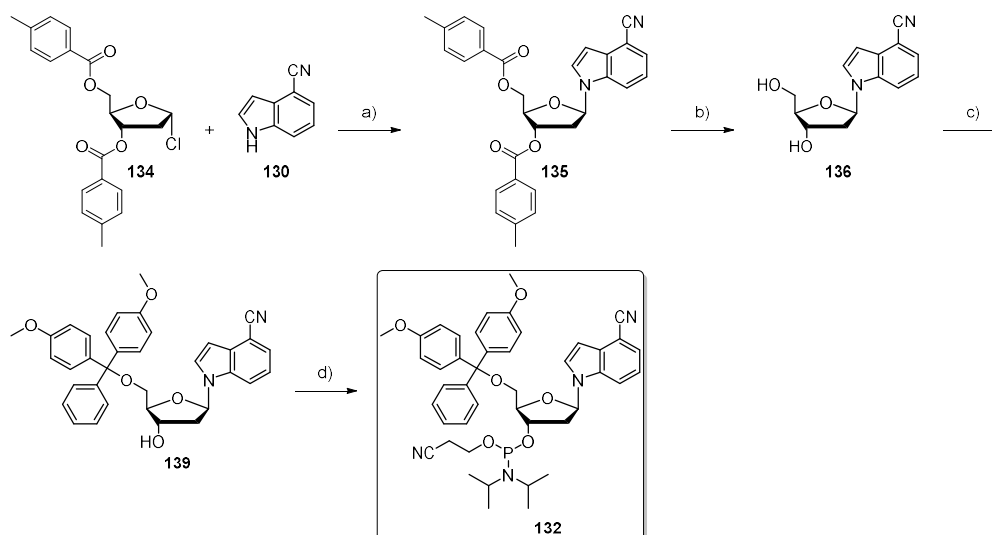


**Chart 26.** Treatment of **135** with NaOMe in MeOH did not only produce the fully deprotected deoxyribonucleoside **136**, but also the partially protected compounds **137** and **138**.



**Figure 54.**  $^1\text{H}$  NMR spectra of a) **136**, b) **137** and c) **138** in  $\text{DMSO-}d_6$ , 400 MHz. The water resonance at 3.33 ppm has been omitted for clarity. The identity of compounds **137** and **138** was verified by additional 2D NMR experiments and HR-MS.

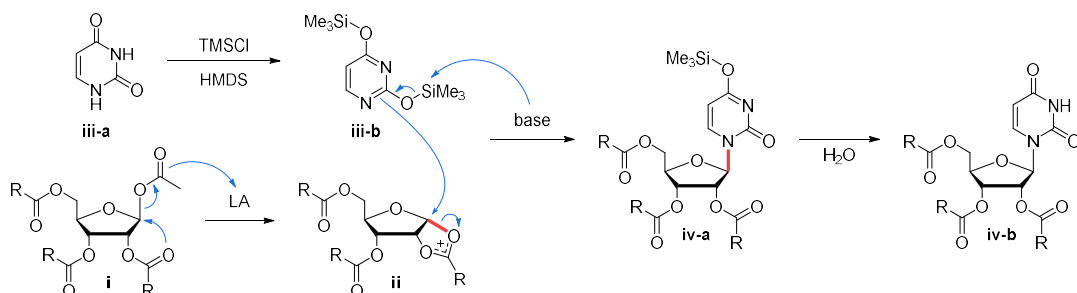
Reaction of **136** with DMTCl and  $\text{NEt}_3$  in pyridine cleanly yielded the 5'-*O*-DMT-protected intermediate **139** (65%). Phosphitylation with CEPCI and  $\text{Me}_2\text{NEt}$  in  $\text{CH}_2\text{Cl}_2$  at ambient temperature for 2.5 h afforded the DNA phosphoramidite **132** with a yield of 81% (Scheme 24).



**Scheme 24.** Synthesis of the 4-cyanoindole deoxyribonucleoside phosphoramidite building block **132** according to a published procedure.<sup>209</sup> Reagents and conditions: a) NaH, MeCN, r.t., 6 h, 98%. b) NaOH, MeOH, r.t., 30 min, 96%. c) DMTCl,  $\text{NEt}_3$ , pyridine, r.t., 4 h, 65%. d) CEPCI,  $\text{Me}_2\text{NEt}$ ,  $\text{CH}_2\text{Cl}_2$ , r.t., 2.5 h, 81%.

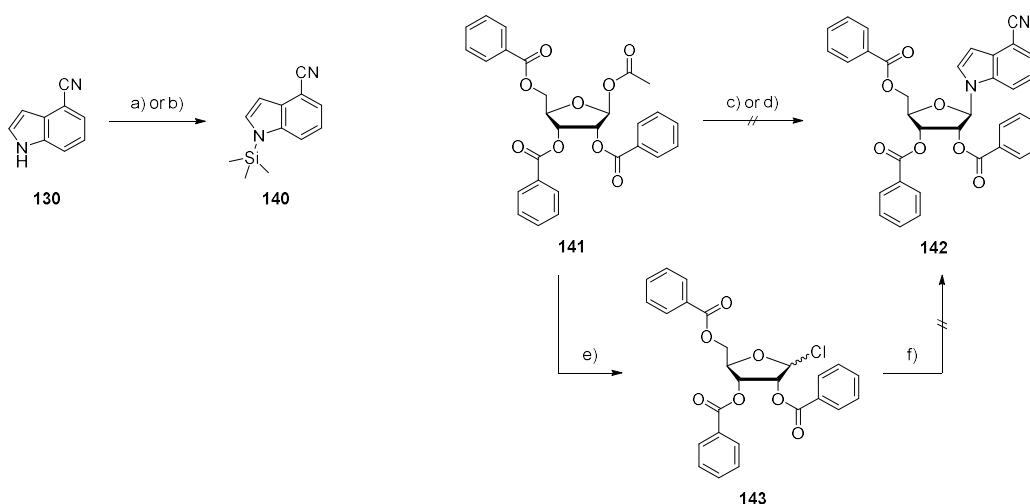
### 2.6.3.2 Synthesis of the RNA phosphoramidite

Initial efforts to synthesize a 4-cyanoindole RNA phosphoramidite building block were centered towards the nucleosidation of **130** with ribose derivatives. In the Vorbrüggen reaction, the 1'-*O*-acetyl group of the fully protected ribose **i** is first activated using a strong Lewis acid such as  $\text{TMSOTf}$  or  $\text{SnCl}_4$ . Loss of acetate results in the formation of acyloxonium intermediate **ii**, which is stabilized by the neighboring 2'-*O*-benzoyl group. Internal delivery of the benzoyl group shields the  $\alpha$  face of the furanose ring from nucleophilic attack. Consequently, reaction with the silylated heterocycle **iii** results in preferential formation of a  $\beta$ -glycosidic bond in the nucleoside **iv** (Scheme 25).<sup>212,213</sup>



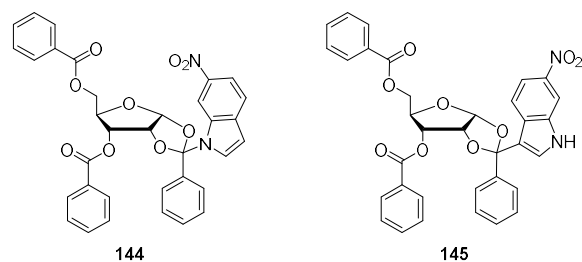
**Scheme 25.** Reaction mechanism for the Vorbrüggen nucleosidation. The protected ribose derivative **i** ( $R = \text{Ph}$  or  $\text{CH}_3$ ) is activated by a suitable Lewis acid (LA) to form acyloxonium intermediate **ii**. A nucleobase (e.g. uracil, **iii-a**) is silylated in situ to generate **iii-b** and thereby increase its N-nucleophilicity. Formation of the  $\beta$ -glycosidic bond leads to ribonucleoside **iv-a**, which is converted to **iv-b** during workup.<sup>212,213</sup>

Vorbrüggen reaction between **130** and **141** using either  $\text{SnCl}_4$  in DCE or  $\text{TMSOTf}$  in MeCN, in analogy to procedures describing the glycosylation of various kinds of heterocyclic nucleobase analogs, did not yield the protected nucleoside **142**, but instead led to decomposition of the starting materials.<sup>214–216</sup> Only minor amounts of **130** were recovered in either case (Scheme 26).



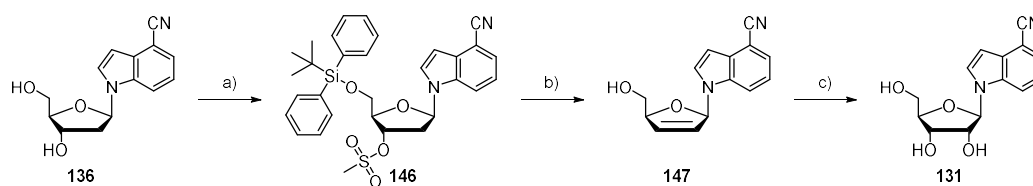
**Scheme 26.** Reaction of 4-cyanoindole (**130**) with the protected ribose derivative **141** under various conditions did not result in product formation. Reagents and conditions: a) TMSCl, HMDS, reflux, 2 h, product was used directly for the next reaction. b) BSA, MeCN, r.t., 1 h, product was used directly for the next reaction. c) **140** prepared according to a,  $\text{SnCl}_4$ , DCE, r.t., 4 h. d) **140** prepared according to b,  $\text{TMSOTf}$ , MeCN, r.t., 4 h. e) 1.  $\text{TiCl}_4$ ,  $\text{CH}_2\text{Cl}_2$ , r.t., 2 h; 2.  $\text{H}_2\text{O}$ , product was used directly for the next reaction. f) 1. **130**, NaH, MeCN,  $0^\circ\text{C}$ , 5 min; 2. **143**, MeCN,  $0^\circ\text{C}$  to r.t., 20 h.

Alternatively, the protected ribose **141** was reacted with  $\text{TiCl}_4$  to form the activated 1'-chloro derivative **143** according to a published procedure for the synthesis of 5-nitroindole ribonucleoside.<sup>217</sup> After treatment of **143** with the in situ-generated sodium salt of **130**, no **142** could be isolated from the reaction mixture. Again, only a small amount of **130** was recovered, suggesting that the compound degrades under the reaction conditions (Scheme 26). Direct ribosylation of indole derivatives has previously been reported to yield unexpected products such as **144** and **145** (Chart 27)<sup>218</sup>.



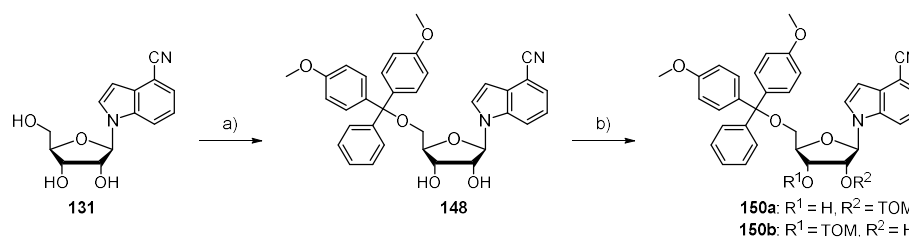
**Chart 27.** Side products reported from the reaction between 6-nitroindole and 1-bromo-2,3,5-tri-*O*-benzoyl-D-ribofuranose.<sup>218</sup>

In a more roundabout approach, indolic ribonucleoside have been prepared from the respective deoxyribonucleosides.<sup>219–221</sup> This methodology was adopted for the synthesis of r4Cl (**131**) as d4Cl was already available.

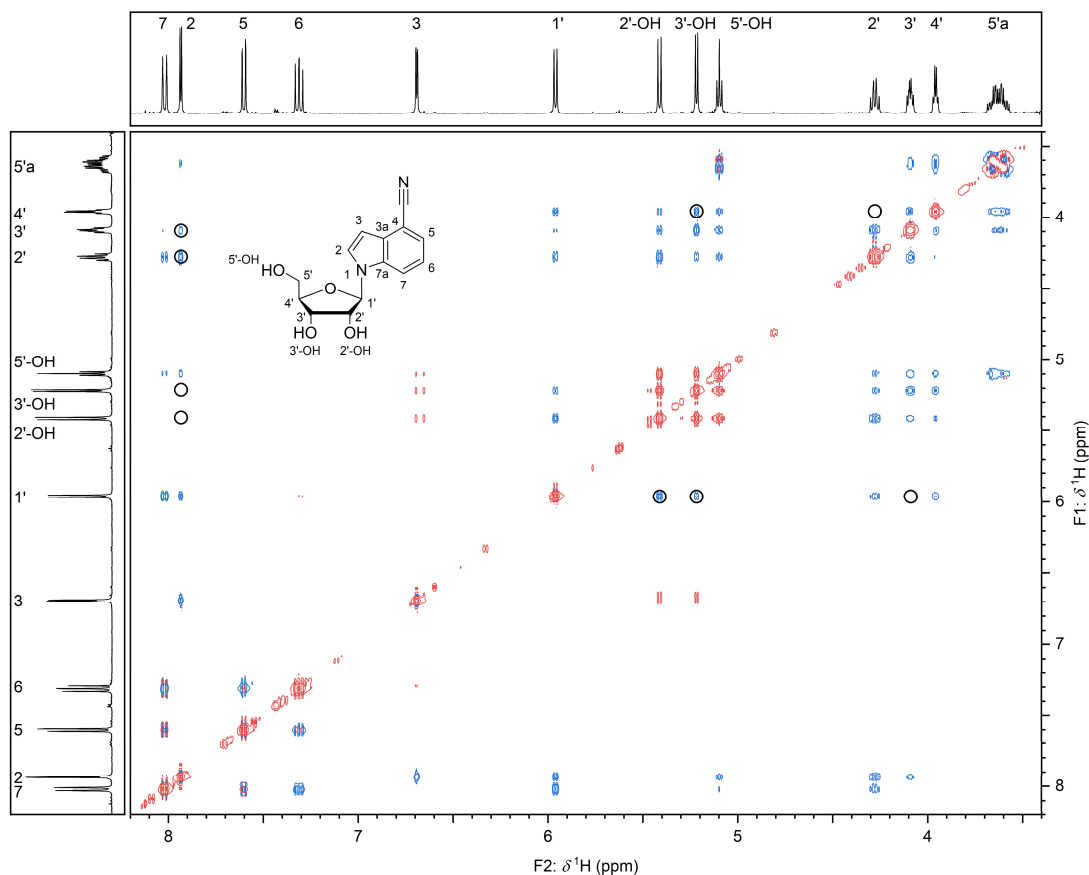


**Scheme 27.** Synthesis of 4-cyanoindole ribonucleoside (**131**, r4Cl). Reagents and conditions: a) 1. TBDPSCI, pyridine, r.t., 16 h; 2. MsCl, pyridine, r.t., 2 h, 72%. b) KOtBu, DMSO/H<sub>2</sub>O, r.t., 15 min, 82%. c) NMO, cat. OsO<sub>4</sub>, acetone/H<sub>2</sub>O, r.t., 18 h, 75%.

Silyl protection of **136** with TBDPSCI in pyridine proceeded selectively at the primary 5'-hydroxy group due to steric demand. Subsequently, the 3'-hydroxy group was activated with MsCl in a one-pot reaction to yield **146** (72%). Treatment with KOtBu in DMSO/H<sub>2</sub>O for 15 min at ambient temperature resulted in base-induced elimination of the 3'-OMs group and concomitant oxidative cleavage of the silyl ether to afford the didehydro-dideoxyribonucleoside **147** with a yield of 82%. Upjohn dihydroxylation with OsO<sub>4</sub> and *N*-methylmorpholine-*N*-oxide (NMO) was used to generate the ribonucleoside **131** (75%) with excellent diastereoselectivity (Scheme 27). <sup>1</sup>H NOESY NMR was used to ascertain the relative configuration of the dihydroxylation product. Correlations were observed between 1'-H/2'-OH, 1'-H/3'-OH and 3'-OH/4'-H and also from indole-2-H to 2'-H and 3'-H, but not to 2'-OH or 3'-OH. Intense correlation peaks between 1'-H/3'-H and 2'-H/4'-H, which would arise in a lyxofuranose isomer, were absent (Figure 55). After having obtained the ribonucleoside, the synthesis of the phosphoramidite building block was resumed. Tritylation of the 5'-hydroxy group was performed using DMTCl and NEt<sub>3</sub> in pyridine for 16 h at ambient temperature to obtain the protected nucleoside **148** with a yield of 89% (Scheme 28).

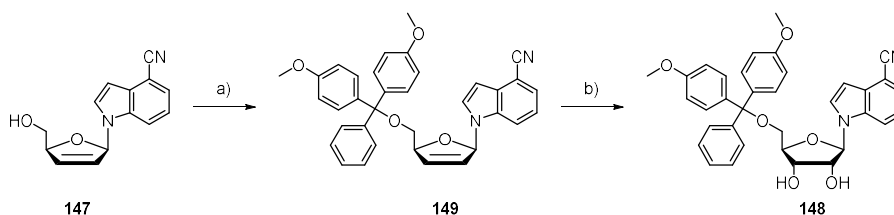


**Scheme 28.** Installation of DMT- and TOM-protecting groups on the ribonucleoside **131** Reagents and conditions: a) DMTCl, NEt<sub>3</sub>, pyridine, r.t., 16 h, 89%. b) 1. (tBu)<sub>2</sub>SnCl<sub>2</sub>, (*i*Pr)<sub>2</sub>NEt, DCE, 70 °C, 45 min; 2. TOMCl, DCE, r.t., 30 min, 77%.



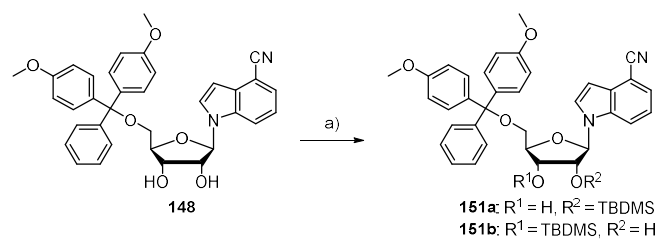
**Figure 55.**  $^1\text{H}$  NOESY NMR of **131** in  $\text{DMSO}-d_6$ , 400 MHz. A mixing time of 1.000 s was used. Encircled correlations distinguish the ribo configuration of the furanose from the alternative lyxo configuration.

For subsequent runs, the order of dihydroxylation and tritylation was swapped in order to avoid purification of the highly polar unprotected ribonucleoside by column chromatography. Tritylation of **147** was complete in 4 h and gave the protected intermediate **149** (62%), which was subjected to dihydroxylation as described above (yield > 99%). The diastereoselectivity of the dihydroxylation was not affected by this change in strategy (Scheme 29).



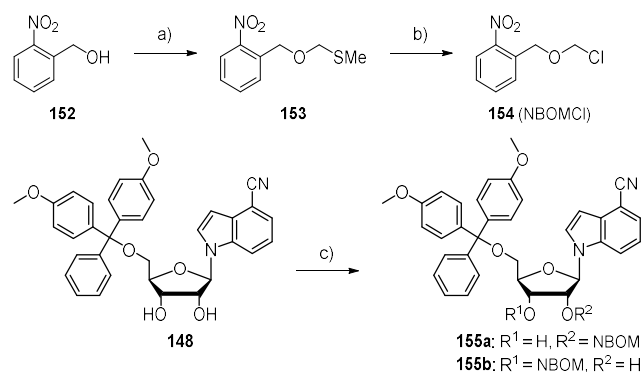
**Scheme 29.** Tritylation of **147** prior to dihydroxylation was preferred in order to avoid purification of the highly polar ribonucleoside **131**. Reagents and conditions: a) DMTCl,  $\text{NEt}_3$ , pyridine, r.t., 4 h, 62%. b) NMO, cat.  $\text{OsO}_4$ , acetone/ $\text{H}_2\text{O}$ , r.t., 18 h, > 99%.

Ribonucleoside **148** was converted in situ to the 2',3'-*O*-di-*tert*-butylstannylene acetal by treatment with  $(t\text{Bu})_2\text{SnCl}_2$  and  $(i\text{Pr})_2\text{NET}$  in DCE at 70 °C for 2 h. Addition of TOMCl gave a mixture of the corresponding 2'- and 3'-*O*-TOM-protected isomers **150a** and **150b**, which could not be separated completely by column chromatography regardless of the chosen stationary and mobile phases (Scheme 28). Therefore, different sets of protecting groups for the 2'-hydroxy group were evaluated. Silylation of **148** with TBDMSCl and  $\text{AgNO}_3$  proceeded only sluggishly and with low conversion. The material obtained after column chromatography of the reaction mixture still contained impurities, but was not purified further because  $^1\text{H}$  NMR indicated the presence of both isomers, **151a** and **151b**, in approximately equal amounts (Scheme 30).



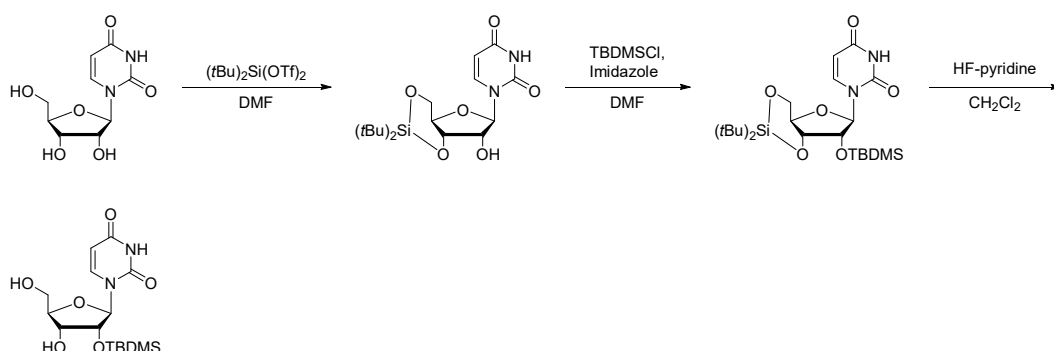
**Scheme 30.** Installation of TBDMS-protecting groups on the DMT-protected ribonucleoside **148**. Reagents and conditions: a) TBDMSCl, AgNO<sub>3</sub>, THF/pyridine, r.t., 20 h, product was not isolated in pure form.

As an alternative to conventional fluoride-labile protecting groups, RNA phosphoramidite building blocks have previously also been synthesized bearing photolabile moieties at the 2'-position such as a 2-nitrobenzyloxymethyl ether (NBOM). The required alkylation reagent NBOMCl (**154**) was prepared in two steps from 2-nitrobenzyl alcohol according to a published procedure.<sup>222</sup> Alkylation of **148** with NBOMCl was performed in a similar fashion to the reaction with TOMCl, but using (*n*Bu)<sub>2</sub>SnCl<sub>2</sub> instead of the *tert*-butyl derivative in the activation step. The resulting regioisomers **155a** and **155b**, however, could not be separated (Scheme 31).



**Scheme 31.** Preparation of NBOMCl for the installation of a photolabile protecting group on the DMT-protected ribonucleoside **148**. Reagents and conditions: a) DMSO, Ac<sub>2</sub>O, AcOH, r.t., 24 h, 63%. b) SO<sub>2</sub>Cl<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, r.t., 30 min, > 99%. c) 1. (*n*Bu)<sub>2</sub>SnCl<sub>2</sub>, (*i*Pr)<sub>2</sub>NEt, DCE, r.t., 1 h; 2. NBOMCl, DCE, 70 °C, 30 min, 27%.

Serebryany and Beigelman have described a strategy to allow for regioselective 2'-*O*-silylation of ribonucleosides by prior installation of a bridging 3',5'-*O*-di-*tert*-butylsilylene acetal. Selective cleavage of the acetal with HF-pyridine would then result in a 2'-*O*-protected ribonucleoside (Scheme 32).<sup>223</sup>

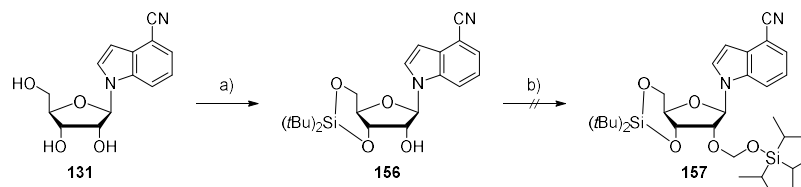


**Scheme 32.** General reaction scheme for the selective 2'-*O*-silyl protection of nucleosides (e.g. uridine) via a bridging 3',5'-*O*-silylene acetal.

In the present case, use of the TOM group was strongly preferred over a silyl ether in order to prevent 2'/3'-migration of the protecting group during subsequent reaction steps. While activation of the nucleoside by formation of a stannylene acetal prior to alkylation with TOMCl is no longer possible, Serebryany and Beigelmann have reported suitable alternative conditions,

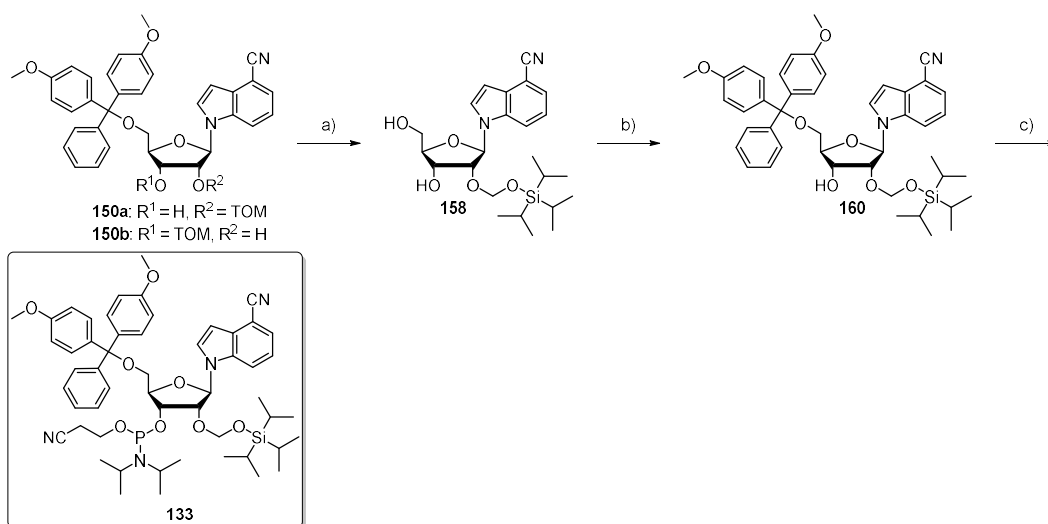


namely forgoing the activation step and using DBU as a base.<sup>224</sup> Conversion of **131** to the bridged silylene acetal **156** was achieved by treatment with  $(t\text{Bu})_2\text{Si}(\text{OTf})_2$  and  $\text{NEt}_3$  in DMF with an isolated yield of 30%. Alkylation with TOMCl and DBU in THF, however, did not result in detectable conversion to the desired compound **157** and the starting material was recovered completely (Scheme 33).

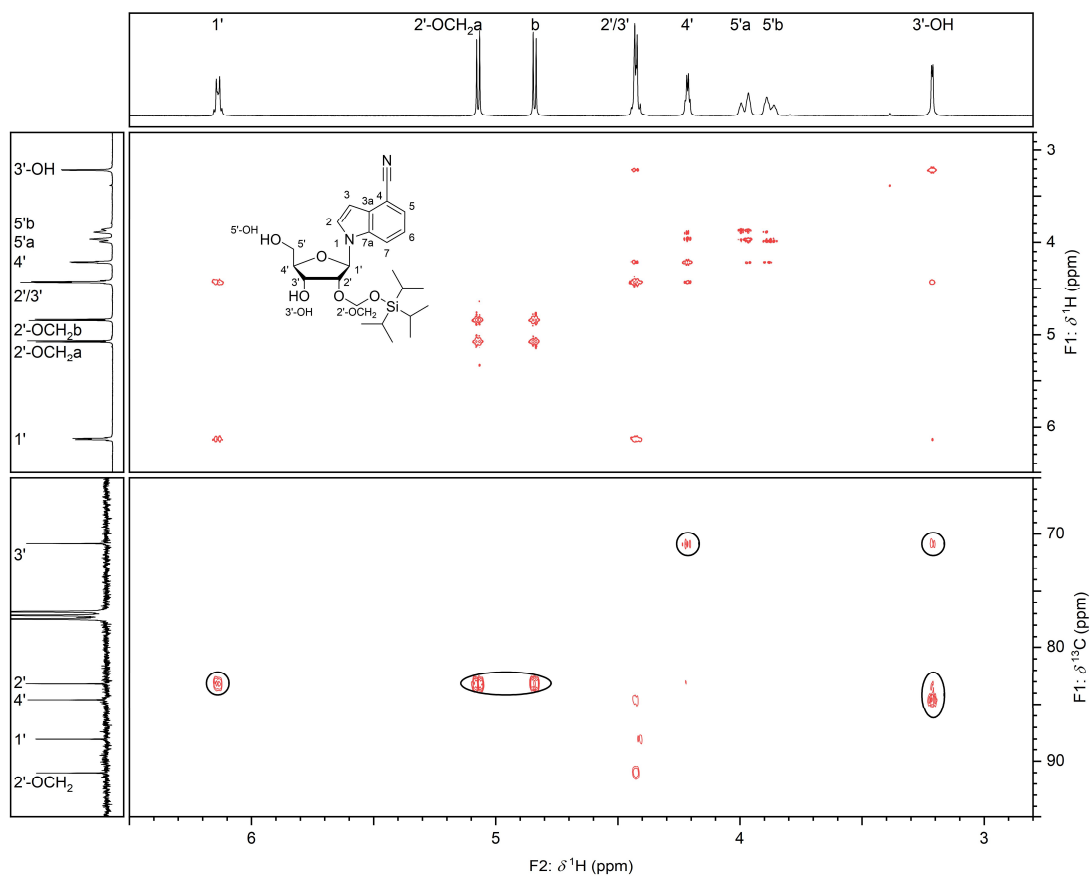


**Scheme 33.** Attempted regioselective 2'-*O*-protection of **131** via the Beigelman approach. Reagents and conditions: a)  $(t\text{Bu})_2\text{Si}(\text{OTf})_2$ , DMF, 0 °C to r.t., 45 min, 30%. b) TOMCl, DBU, THF, 0 °C, 1 h to r.t., 6 h.

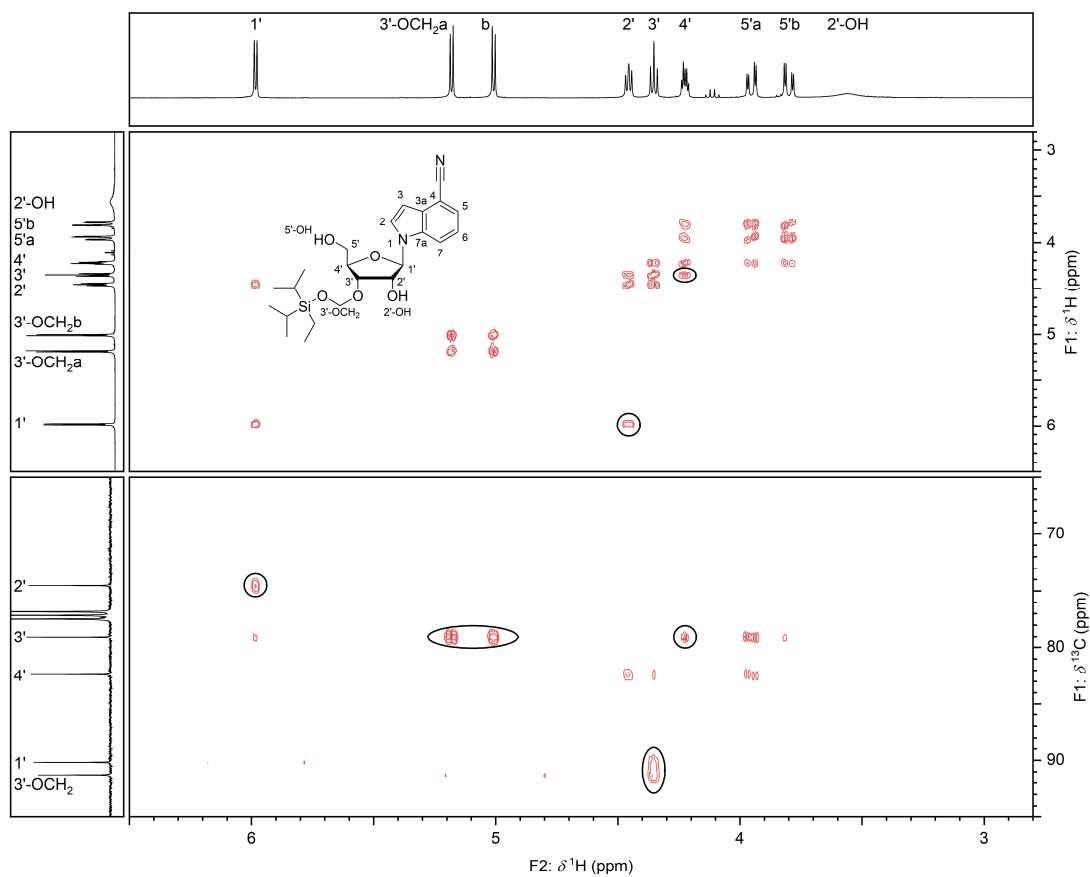
It was reasoned that separation of the 2'- and 3'-*O*-protected isomers of **150**, **151** or **155** was not feasible due to the low polarity of the indole moiety compared to regular nucleobases and their analogs. Therefore, the 5'-*O*-DMT group of **150** was removed by treatment with 1% TFA in DCM for 15 min at ambient temperature, which not only resulted in a substantial increase in polarity for both isomers as determined by TLC, but also made them readily separable by column chromatography (Scheme 34). Compounds **158** and **159** were unambiguously identified by  $^1\text{H}$  COSY and  $^1\text{H}$ - $^{13}\text{C}$  HMBC NMR. While 2'- and 3'-H are virtually isochronous and therefore indistinguishable in the 2'-protected isomer **158**, two-bond correlations between 1'-H/C2' and 4'-H/C3' could be used to assign the positions of the  $(i\text{Pr})_3\text{SiOCH}_2$  and OH groups, respectively. Additionally, the prominent cross peak between the OH group and C4' could only have arisen from a 3'-OH group. (Figure 56). For the 3'-protected isomer **159**, additional three-bond correlations between 1'-H/2'-H and 4'-H/3'-H facilitated assignment of the individual resonances of the ribose moiety (Figure 57).



**Scheme 34.** Synthesis of the 4-cyanoindole ribonucleoside phosphoramidite building block **133**. Reagents and conditions: a) TFA,  $\text{CH}_2\text{Cl}_2$ , r.t., 15 min, 39%. b) DMTCl,  $\text{NEt}_3$ , pyridine, r.t., 4 h, 65%. c) CEPCI,  $\text{Me}_2\text{NEt}$ ,  $\text{CH}_2\text{Cl}_2$ , r.t., 2.5 h, 79%.



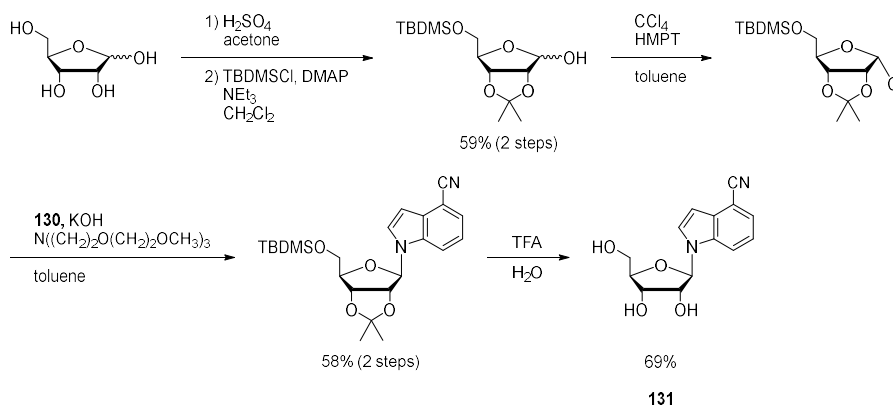
**Figure 56.**  $^1\text{H}$  COSY (top) and  $^1\text{H}$ - $^{13}\text{C}$  HMBC (bottom) NMR of **158** in  $\text{CDCl}_3$ , 400 MHz. Encircled correlations were used to assign the position of the TOM protecting group. The COSY alone was not sufficient because the signals for 2'- and 3'-H are overlapping.



**Figure 57.**  $^1\text{H}$  COSY (top) and  $^1\text{H}$ - $^{13}\text{C}$  HMBC (bottom) NMR of **159** in  $\text{CDCl}_3$ , 400 MHz. Encircled correlations were used to assign the position of the TOM protecting group. No peaks were observed for the 2'-OH group due to chemical exchange.

The 2'-*O*-TOM-protected nucleoside **158** was isolated with a yield of 39%. Tritylation was performed under the same conditions as described above and afforded compound **160** (54%), which was finally phosphitylated with CEPCI and Me<sub>2</sub>NEt in CH<sub>2</sub>Cl<sub>2</sub> at ambient temperature for 2.5 h. The RNA phosphoramidite building block **133** was obtained with a yield of 79% (Scheme 34).

Recently, after this synthetic pathway had been submitted for publication, Harki and coworkers described a different route towards the ribonucleoside **131**, which entailed 1'-chlorination of a protected ribose with CCl<sub>4</sub> and HMPT followed by nucleosidation with 4-cyanoindole (**130**) using KOH in toluene. Final deprotection with TFA afforded compound **131** over six steps with a combined yield of 24% (Scheme 35).<sup>225</sup> In comparison, the route presented in Scheme 27 above gave a yield of 43% over five steps.



**Scheme 35.** Alternative synthesis of r4Cl (**131**) described by Harki and coworkers.<sup>225</sup> The step economy and overall yield of the DNA–RNA conversion strategy detailed in this thesis are superior to the direct nucleosidation with a protected ribose.

## 2.7 CHARACTERIZATION OF 4-CYANOINDOLE RIBONUCLEOSIDE

### 2.7.1 Impact of 4-cyanoindole incorporation on oligonucleotide structure

#### 2.7.1.1 Incorporation of 4CI into RNA and DNA oligonucleotides

To answer the question of whether 4CI functions as a universal nucleobase surrogate in RNA the same way it does in DNA, its influence on strand hybridization thermodynamics needed to be considered. To this end, 16 bp RNA strands, containing either of the four canonical Watson-Crick base pairs or a base pair with 4CI, were designed to resemble the bottom and top stems of the Chili aptamer (Chart 28).

#### RNA

5' -GGCUAGCUG<sup>A</sup>CGCCAG-3'

3' -CCGAUCGAC<sup>N</sup>GCGGUC-5'

N = U (**R1<sub>AU</sub>**), r4CI (**R1<sub>AX</sub>**)

5' -GGCUAGCUG<sup>C</sup>CGCCAG-3'

3' -CCGAUCGAC<sup>N</sup>GCGGUC-5'

N = G (**R2<sub>CG</sub>**), r4CI (**R2<sub>CX</sub>**)

5' -GGCUAGCUG<sup>G</sup>CGCCAG-3'

3' -CCGAUCGAC<sup>N</sup>GCGGUC-5'

N = C (**R3<sub>Gc</sub>**), r4CI (**R3<sub>Gx</sub>**)

5' -GGCUAGCUG<sup>U</sup>CGCCAG-3'

3' -CCGAUCGAC<sup>N</sup>GCGGUC-5'

N = A (**R4<sub>UA</sub>**), r4CI, (**R4<sub>UX</sub>**), d4CI (**R4<sub>UY</sub>**)

#### DNA

5' -GGCTAGCTG<sup>T</sup>CGCCAG-3'

3' -CCGATCGAC<sup>N</sup>GCGGTC-5'

N = dA (**D1<sub>TA</sub>**), d4CI (**D1<sub>Ty</sub>**)

**Chart 28.** Sequences of 16 bp duplexes (RNA, DNA or mixed) for studying the influence of 4CI incorporation on oligonucleotide structure.

RNA oligonucleotides were prepared by solid-phase synthesis with 2'-O-TOM-protected phosphoramidites of acetylated nucleobases and the 4CI phosphoramidite building block **133** using standard methods with final detritylation. Cleavage from the solid support, deprotection and purification were performed as described in section 4.1.3.3. Yields were in the range of 18–136 nmol, the purity of the synthesis products and integrity of the 4CI modification were confirmed by HPLC and HR-MS (Table 21).

**Table 21.** Description of native and modified 16 nt oligonucleotides for thermodynamic analyses obtained by solid-phase synthesis. See also section 4.3.2.

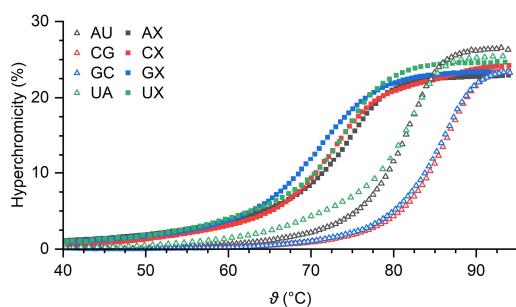
Number <sup>[a]</sup>	Description	Yield <sup>[b]</sup> nmol	<i>m/z</i> calc.	<i>m/z</i> obs.	Species
<b>R1a</b>	native top strands of <b>R1–R4</b> , <b>D1</b>	41	5132.743	5132.778	[M–H] <sup>–</sup>
<b>R2a</b>		79	5108.732	5108.802	[M–H] <sup>–</sup>
<b>R3a</b>		38	5148.738	5148.792	[M–H] <sup>–</sup>
<b>R4a</b>		90	5109.716	5109.755	[M–H] <sup>–</sup>
<b>D1a</b>		60	4895.844	4895.905	[M–H] <sup>–</sup>
<b>R1b</b>	native bottom strands of <b>R1–R4</b> , <b>D1</b>	51	5069.710	5069.779	[M–H] <sup>–</sup>
<b>R2b</b>		111	5108.732	5108.803	[M–H] <sup>–</sup>
<b>R3b</b>		114	5068.726	5068.793	[M–H] <sup>–</sup>
<b>R4b</b>		120	5092.737	5092.807	[M–H] <sup>–</sup>
<b>D1b</b>		136	4864.850	4864.909	[M–H] <sup>–</sup>
<b>RXb</b>	modified bottom strands of <b>R1–R4</b> , <b>D1</b>	51	5099.736	5099.814	[M–H] <sup>–</sup>
<b>RYb</b>		18	5083.741	5083.808	[M–H] <sup>–</sup>
<b>DYb</b>		21	4871.849	4871.903	[M–H] <sup>–</sup>

[a] X = r4Cl, Y = d4Cl.

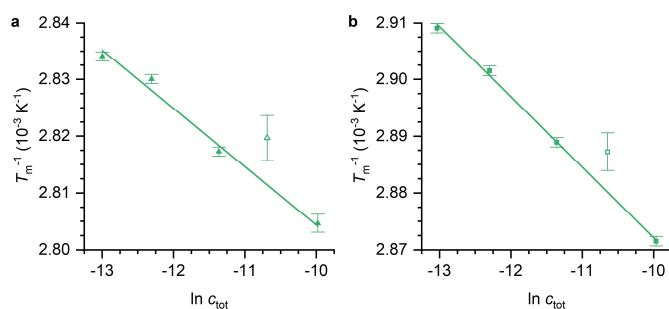
[b] 600 nmol synthesis scale.

### 2.7.1.2 Thermal stability analysis

Thermal denaturation curves of the RNA duplexes were measured at five different concentrations (1, 2, 5, 10 and 20  $\mu\text{M}$  per single strand) in sodium phosphate buffer (PBS). An example is shown in Figure 58. Each measurement comprised an initial annealing ramp followed by two cooling/heating cycles. Melting temperatures ( $T_m$ ) and thermodynamic parameters of the duplex formation ( $\Delta H^0$ ,  $\Delta S^0$ ,  $\Delta G_{298\text{K}}^0$ ) were extracted from the melting curves using a nonlinear least squares fitting algorithm, the derivation of which is given in section 4.1.18.1. The curves from both cycles were fitted individually and the results were averaged.  $\Delta H^0$ ,  $\Delta S^0$  and  $\Delta G_{298\text{K}}^0$  were further averaged between all five concentrations (Table 22, Supplementary Table 2). Fit errors for  $T_m$  were below 0.5  $^\circ\text{C}$ , which is considered to be the limiting accuracy of this method.<sup>226</sup> For comparison, the concentration dependence of  $T_m$  was also used to calculate thermodynamic parameters from a Van't Hoff analysis (see section 4.1.18.1 for details), shown exemplary for **R4** in Figure 59. The melting temperatures of **R1<sub>AU</sub>/R4<sub>UA</sub>** and **R2<sub>CG</sub>/R3<sub>GC</sub>** were consistent within each pair and in agreement with values predicted based on a nearest-neighbor model; thermodynamic parameters obtained via the Van't Hoff method, however, appeared to be systematically overestimated.<sup>227,228</sup> Therefore, only the directly fitted parameters will be considered in the following discussion.



**Figure 58.** Thermal denaturation curves of native and 4Cl-modified 16 bp RNA duplexes **R1–R4**. The change in absorbance at 260 nm was converted to percentage hyperchromicity. Conditions: 10  $\mu\text{M}$  per single-stranded oligonucleotide, PBS pH 7.4. Data is shown for the third of five ramps. Fitted curves and results for other concentrations are depicted in Supplementary Figure 9 on page 219.



**Figure 59.** Van't Hoff plots for a) **R4<sub>UA</sub>** and b) **R4<sub>UX</sub>**.  $c_{\text{tot}}$  denotes the added concentrations of both single-stranded oligonucleotides in each sample as determined by UV absorbance at 260 nm. Data points were fitted with a linear model, excluding those indicated by open symbols due to systematic deviations. Results for other duplexes are depicted in Supplementary Figure 9 on page 219.

**Table 22.** Thermodynamic data of 16 bp RNA, DNA and hybrid duplexes<sup>[a]</sup> at a concentration of 10  $\mu\text{M}$  per single strand in PBS (pH 7.4) obtained from direct fitting of the thermal denaturation curves or by Van't Hoff analysis. Melting temperatures for other concentrations are reported in Supplementary Table 2 on page 227.

	$T_m^{[b]}$ °C	$\Delta T_m^{[c]}$ °C	$\Delta H^0$ kJ mol <sup>-1</sup>	$\Delta S^0$ J mol <sup>-1</sup> K <sup>-1</sup>	$\Delta G_{298\text{K}}^0$ kJ mol <sup>-1</sup>	$\Delta H^0$ kJ mol <sup>-1</sup>	$\Delta S^0$ J mol <sup>-1</sup> K <sup>-1</sup>	$\Delta G_{298\text{K}}^0$ kJ mol <sup>-1</sup>
	Nonlinear fit				Van't Hoff analysis			
<b>R1<sub>AU</sub></b>	81.2 ± 0.4		-717	-1919	-145	-889	-2412	-169
<b>R1<sub>AX</sub></b>	74.3 ± 0.2	-6.9	-580	-1575	-111	-756	-2102	-129
<b>R2<sub>CG</sub></b>	85.7 ± 0.4		-717	-1892	-153	-858	-2280	-172
<b>R2<sub>CX</sub></b>	71.8 ± 0.2	-13.9	-557	-1505	-109	-594	-1635	-107
<b>R3<sub>GC</sub></b>	85.6 ± 0.3		-707	-1863	-152	-1026	-2765	-202
<b>R3<sub>GX</sub></b>	69.4 ± 0.2	-16.2	-455	-1214	-92	-675	-1882	-114
<b>R4<sub>UA</sub></b>	81.5 ± 0.5		-757	-2027	-153	-810	-2188	-158
<b>R4<sub>UX</sub></b>	73.2 ± 0.4	-8.3	-629	-1710	-119	-670	-1840	-121
<b>R4<sub>UY</sub></b>	71.6 ± 0.3	-9.9	-594	-1621	-110	-559	-1526	-103
<b>D1<sub>TA</sub></b>	69.2 ± 0.2		-493	-1339	-94	-672	-1874	-113
<b>D1<sub>TY</sub></b>	58.1 ± 0.2	-11.1	-418	-1162	-71	-403	-1129	-66

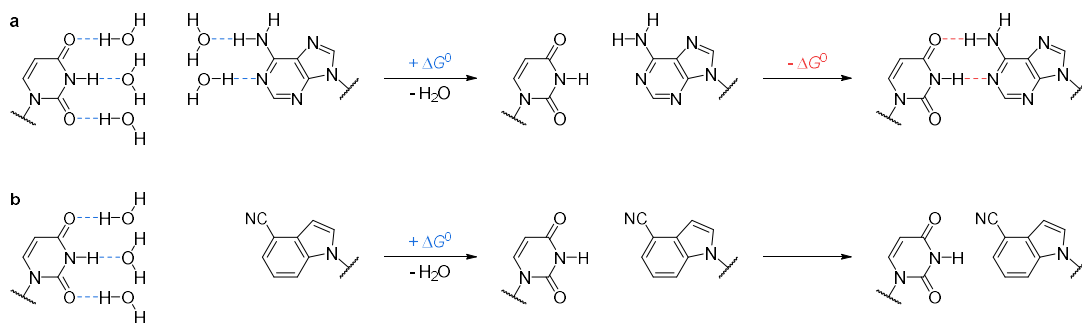
[a] X = r4Cl, Y = d4Cl.

[b] Values are given as mean ± s.d. for two cooling/heating cycles.

[c] Difference of means relative to the unmodified duplex.

Incorporation of 4Cl reduced the thermal stability of all four duplexes. However, the resulting  $T_m$  values were within a narrow range also observed for other universal RNA nucleoside analogs.<sup>206,229</sup> Differences between pairing with either of the canonical purine or pyrimidine bases were minor, indicating the expected lack of discriminatory hydrogen bonding.

In **R1** and **R4**, the change in  $\Delta G_{298\text{K}}^0$  upon incorporation of 4Cl is +34 kJ mol<sup>-1</sup>. This value clearly exceeds the hydrogen bond strength of a single A–U base pair, which is suggested to be on the order of 8 kJ mol<sup>-1</sup>.<sup>230</sup> Such a discrepancy has previously been rationalized on grounds of a desolvation penalty,<sup>231</sup> the argument being as follows: Unhybridized nucleobases engage in hydrogen bonding with surrounding water molecules, but need to shed this solvation shell in order to form a base pair. During this process, hydrogen bonds are both broken and formed, resulting in an overcompensation of the desolvation energy. By replacing one of the canonical bases with 4Cl, the opposite base still loses its solvation shell due to steric constraints, but the system does not gain stabilization from new interbase hydrogen bonds (Scheme 36). The comparatively larger destabilization of **R2** and **R3** is consistent with stronger hydrogen bonding in a G–C base pair.



**Scheme 36.** a) During the formation of a canonical Watson-Crick base pair, e.g. between A and U, the free energy cost of desolvation is compensated by newly formed hydrogen bonds. b) A universal base analog such as 4CI does not engage in hydrogen bonding. Therefore, hybridization carries an energetic penalty.

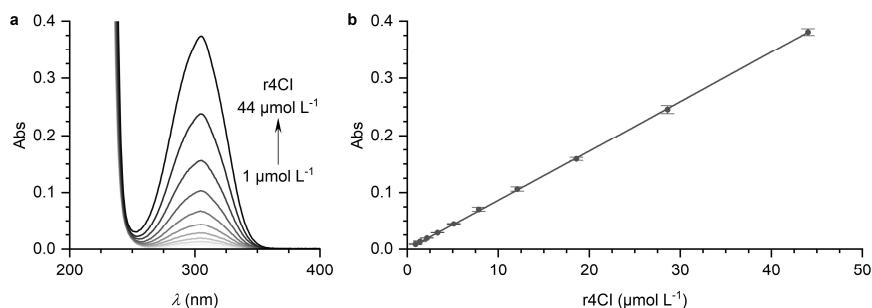
All four RNA duplexes display the same pattern of change in  $\Delta H^0$  and  $\Delta S^0$ : Compared with the formation of Watson-Crick base pairs, hybridization with 4CI carries a less favorable enthalpic component, which is partially compensated by a more favorable entropic contribution, suggesting an overall increase in conformational freedom around the modification site. In fact, molecular dynamics simulations of a base pair between A and a universal 2,4-difluorophenyl nucleobase analog, which is well accommodated in double-stranded RNA, indicate rapid exchange between completely intrahelical and partially extrahelical base pair geometries in the absence of hydrogen bonding.<sup>232</sup>

The melting temperature of hybrid duplex **R4<sub>UV</sub>**, which contains the 4CI deoxyribonucleoside **136** opposite U, is slightly reduced by an additional 1.6 °C (at 10  $\mu\text{M}$ ) with respect to the otherwise identical **R4<sub>UX</sub>**. While not investigated in further detail, this result suggests that d4CI can be used as a surrogate for r4CI in preliminary or exploratory studies with simple RNAs, which might be advantageous due to the less laborious preparation of the DNA phosphoramidite **132**. The effect of d4CI-incorporation on a larger, more complex RNA such as the Chili aptamer will be examined in section 2.8.2. For reference, in the 16 bp all-DNA duplex **D1**,  $T_m$  decreased by 11.1 °C (at 10  $\mu\text{M}$ ) upon exchanging the central T–A base pair for T–4CI. Overall, these results show that 4CI functions well as a universal nucleobase analog not only in DNA, but in RNA to the same degree.

## 2.7.2 Spectroscopic properties of 4-cyanoindole

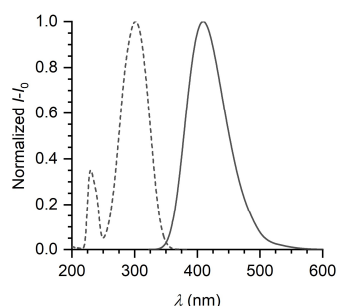
### 2.7.2.1 Characterization of the free nucleoside

Figure 60a shows the UV/Vis spectrum of r4CI (**131**) in aqueous buffer. The extinction coefficient of r4CI at the absorption maximum (305 nm) was determined by serial dilution to be  $8640 \pm 20 \text{ L mol}^{-1} \text{ cm}^{-1}$  (Figure 60b, Table 23), which is close to the value reported for d4CI.<sup>209</sup>



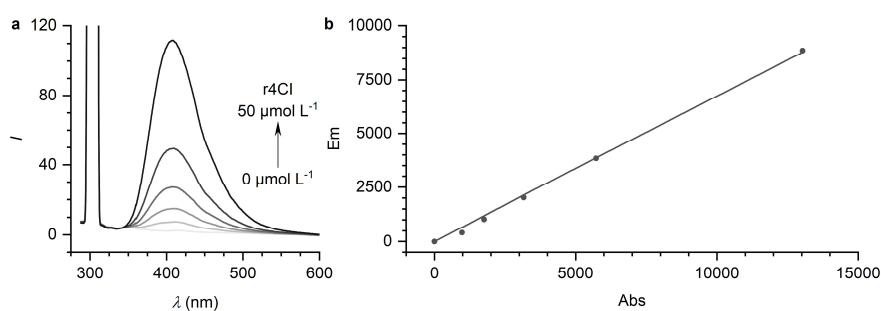
**Figure 60.** a) UV/Vis spectra of r4CI at different concentrations of the nucleoside. b) Plot of the absorbance at 305 nm vs concentration of r4CI. The extinction coefficient was determined from a linear fit to the data.  $\epsilon = 8640 \pm 20 \text{ L mol}^{-1} \text{ cm}^{-1}$ . Conditions: 0.9–44  $\mu\text{M}$  r4CI, PBS pH 7.4. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Excitation at 303 nm resulted in intense blue fluorescence emission at 412 nm, i.e. with a Stokes shift of 109 nm. In addition to the main band, the excitation spectrum of r4Cl shows a second peak at 230 nm, which originates from the formation of a  $\pi\sigma^*$  state (Figure 61).<sup>233</sup>



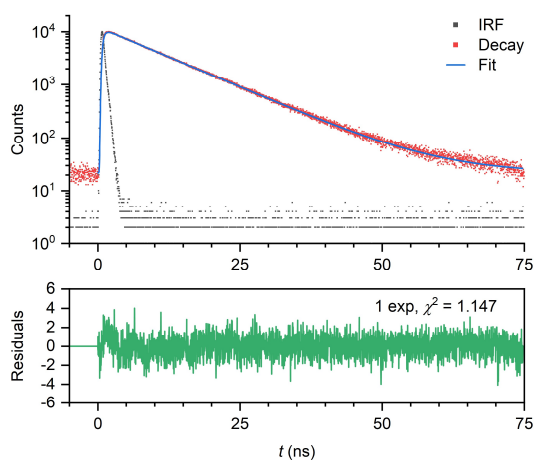
**Figure 61.** Fluorescence excitation and emission spectra of r4Cl ( $\lambda_{\text{ex/em}} = 303/410$  nm). Spectra were normalized after blank correction (buffer). Conditions:  $1 \mu\text{M}$  r4Cl in PBS pH 7.4.

Using an integrating sphere, the absolute fluorescence quantum yield ( $\phi_{\text{abs}}$ ) of r4Cl was measured to be  $0.67 \pm 0.01$  (Figure 62). For comparison, a relative quantum yield ( $\phi_{\text{rel}}$ ) of 0.78 has been reported for 4Cl itself against tryptophan as a standard.<sup>234</sup> Harki and coworkers give a value of 0.90 for  $\phi_{\text{rel}}$  of d4Cl relative to quinine sulfate, which appears unusually large in this context, but can be reconciled considering the differences in absorption and emission maxima between quinine sulfate and 4Cl.<sup>209</sup>



**Figure 62.** a) Fluorescence spectra of r4Cl ( $\lambda_{\text{ex}} = 303$  nm) at different concentrations of the nucleoside. b) Plot of the integrated emission vs absorbed light (for further details see section 4.1.12.1). The absolute quantum yield was determined from a linear fit to the data.  $\phi_{\text{abs}} = 0.67 \pm 0.01$ . Conditions: 0–50  $\mu\text{M}$  r4Cl, PBS pH 7.4.

The fluorescence lifetime of r4Cl is 9.48 ns as determined by TCSPC measurement; the decay kinetics are monoexponential, indicating that a single chemical species is present (Figure 63).



**Figure 63.** Top panel: Fluorescence lifetime measurement of r4Cl with monoexponential fit ( $\lambda_{\text{ex/em}} = 303/412$  nm). Bottom panel: Weighted residuals and  $\chi^2$  for the fit. Conditions:  $1 \mu\text{M}$  r4Cl in PBS pH 7.4.

### 2.7.2.2 Behavior of r4Cl in RNA

Incorporation of r4Cl into an oligonucleotide (**R4**) had limited effect on the position of excitation and emission maxima. In particular, the emission maximum in single- resp. double-stranded RNA



shows a hypsochromic shift of 3–5 nm with respect to the free nucleoside (Table 23). This slight difference would not have any significant effect on the spectral overlap with Chili–HBI complexes. Assuming that the extinction coefficient of 4CI remained unchanged, the relative quantum yield was simply determined from the ratio of fluorescence emission relative to the free nucleoside.  $\phi_{\text{rel}}$  was 0.17 for single-stranded RNA in a 5'-C/3'-G context.

**Table 23.** Spectroscopic properties of r4CI as a nucleoside and in single-/double-stranded RNA.

	$\lambda_a$ nm	$\epsilon$ L mol <sup>-1</sup> cm <sup>-1</sup>	$\lambda_{\text{ex}}$ nm	$\lambda_{\text{em}}$ nm	$\phi$	$\tau^{[f]}$ ns
r4CI	305	8640 ± 20	303	412	0.67 ± 0.01 <sup>[d]</sup>	9.48 <sup>[g]</sup>
ssRNA <sup>[a]</sup>	— <sup>[c]</sup>	—	303	409	0.17 <sup>[e]</sup>	7.24 <sup>[h]</sup>
dsRNA <sup>[b]</sup>	—	—	304	407	0.05 <sup>[e]</sup>	1.12 <sup>[h]</sup>

[a] **RXb**. [b] **R4<sub>UX</sub>**.

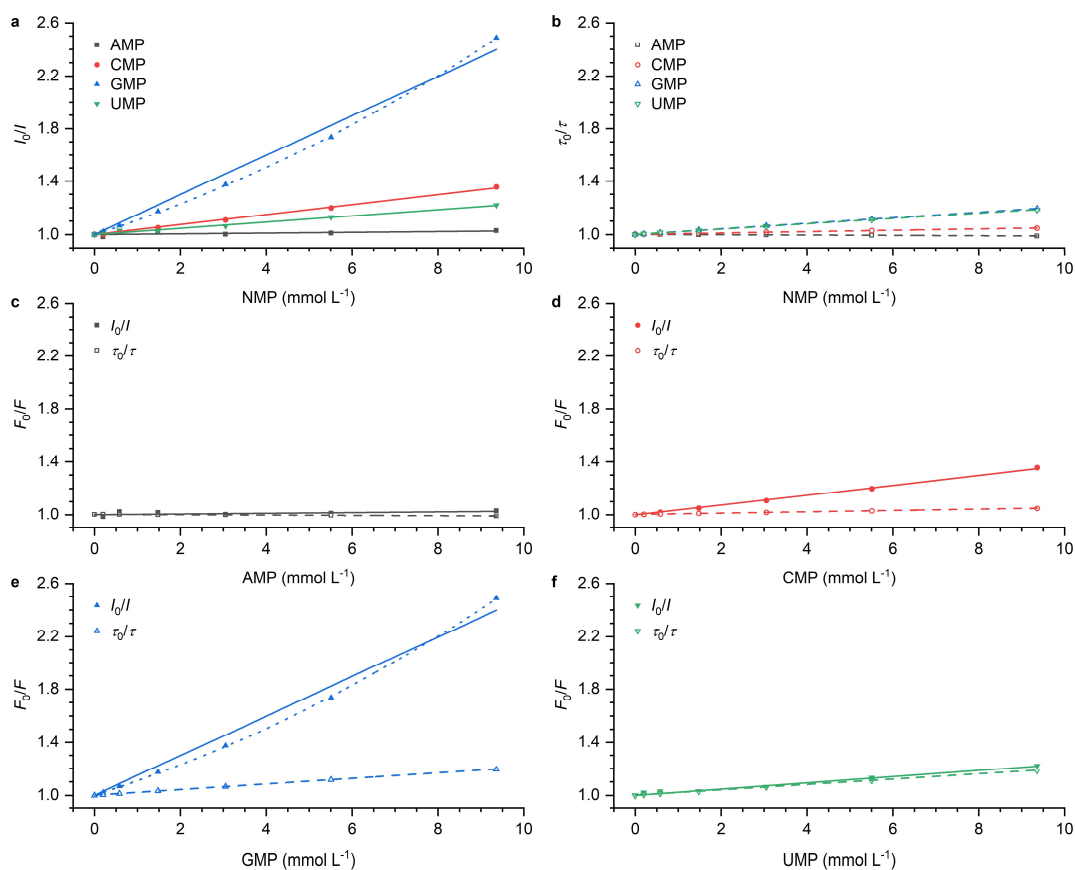
[c] Not measured.

[d] Absolute quantum yield. [e] Estimated relative to r4CI.

[f]  $\lambda_{\text{ex/em}} = 318/410$  nm

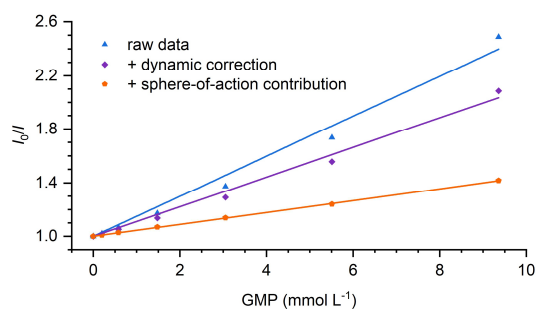
[g] Monoexponential fit. [h] Amplitude average for triexponential fit (Supplementary Figure 10).

Quenching of fluorescent nucleobase analogs in oligonucleotides is a common phenomenon and has been observed in several cases, including tC<sup>0,235</sup>, 2AP<sup>236</sup> and also d4CI.<sup>209</sup> Neighboring guanines have been found to be particularly effective quenchers operating by photoinduced electron transfer (PET) from the nucleobase to the excited fluorophore.<sup>237,238</sup> Stern-Volmer titrations of r4CI with each of the four canonical ribonucleotides showed decreasing quenching efficiency in the series G >> C ≈ U > A (Figure 64a, b). AMP has virtually no effect on 4CI fluorescence (Figure 64c), whereas CMP appears to predominantly be a static quencher as evidenced by the negligible slope of the fluorescence lifetime ( $\tau_0/\tau$ ) plot (Figure 64d). UMP, on the other hand, acts as a purely dynamic quencher (Figure 64f). From the upward curvature of the steady state ( $I_0/I$ ) plot it is clear that GMP quenches r4CI not only dynamically, but additionally with an apparent static contribution. The dynamic quenching constant ( $K_D$ ) for GMP was estimated to be 20.6 ± 0.1 L mol<sup>-1</sup> according to the  $\tau_0/\tau$  data (Figure 64e, see section 4.1.20.1 for further details).



**Figure 64.** Stern-Volmer plots for the quenching of r4Cl by the four canonical ribonucleotides as determined by a) steady-state intensity ( $\lambda_{\text{ex/em}} = 303/410$  nm) and b) fluorescence lifetime ( $\lambda_{\text{ex/em}} = 318/410$  nm). c-f) Combined steady-state and lifetime data for each ribonucleotide. Solid and dashed lines denote linear fits to the data points, dotted lines denote a quadratic fit. All y-axis intercepts were set to 1. Conditions: 5  $\mu\text{M}$  r4Cl, 0–9.4 mM NMP, PBS pH 7.4.

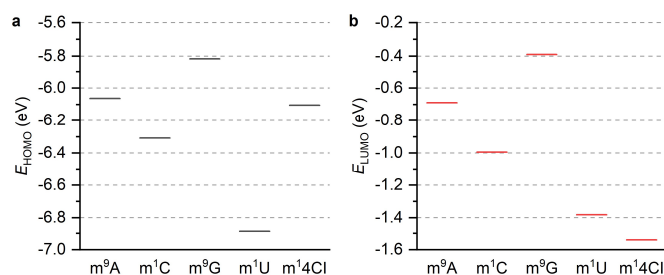
While a simple combination of dynamic and static quenching apparently fits the data well, this alone did not allow for straightforward extraction of the static quenching constant ( $K_S$ ) as can be seen from a plot that includes a dynamic correction.<sup>239</sup> Inclusion of an exponential term describing a sphere-of-action contribution fully explains the observed quenching behavior with a  $K_S$  of  $44.4 \pm 3.6$  L mol<sup>-1</sup> and an interaction radius of  $2.5 \pm 0.1$  nm. Considering that electron transfer efficiency obeys an exponential distance dependence on a sub-nanometer length scale, the interaction radius is too large to directly describe PET quenching between guanine and 4Cl.<sup>240,241</sup> This suggests that, at short distances, intermolecular interactions accelerate the formation of a stacked, statically quenched complex between the two molecules.<sup>239</sup>



**Figure 65.** The  $I_0/I$  values for r4Cl quenching by GMP (blue) were divided by the  $\tau_0/\tau$  data to correct for dynamic quenching. The remaining nonlinearity (violet) is due to an additional quenching factor. Division by an exponential term describing a sphere-of-action model reveals the linear static quenching contribution (orange).

Nonetheless, PET quenching remains a valid pathway at short distances. A frontier molecular orbital analysis at the B3LYP-D3/def2-TZVP level of theory, as previously conducted by Sugiyama

and coworkers for a fluorescent thymine analog,<sup>242</sup> reveals that guanine, but not any of the other nucleobases, would be capable of donating an electron to photoexcited 4Cl (Figure 66).



**Figure 66.** a) Highest occupied and b) lowest unoccupied molecular orbital energies (DFT: B3LYP-D3/def2-TZVP) of A, C, G, U and 4Cl, each bearing a methyl group at the glycosidic nitrogen to reduce computational effort. A qualitative analysis suggests that PET quenching of 4Cl fluorescence by guanine is possible.<sup>241</sup>

Further quenching of 4Cl fluorescence occurred in double-stranded RNA ( $\phi_{rel} = 0.05$ ) and is likely caused by the presence of an additional guanine in the opposite strand 5' of the modification site (Table 23). The fluorescence decay of r4Cl becomes multiexponential upon incorporation into the oligonucleotide (Table 24, Supplementary Figure 10). A slow decay component, with a lifetime similar to free r4Cl, arises from an unstacked and therefore weakly interacting subpopulation of the fluorophore. Two faster components correspond to different levels of interaction with the surrounding oligonucleotide environment.<sup>243,244</sup> Complex decay kinetics are a common occurrence with fluorescent nucleotides; for example, 2AP fluorescence decays in double-stranded DNA are generally described by four exponential components.<sup>236</sup> Interestingly, the amplitude average lifetime of r4Cl in single-stranded **R4** seems unusually long in relation to its diminished quantum yield, which suggests the formation of spectroscopically silent species, as has been proposed for tC<sup>0</sup>-modified RNA.<sup>198</sup> Such species would exhibit lifetime components shorter than the detection limit of the TCSPC setup and originate from static or very fast dynamic quenching effects.<sup>245</sup>

**Table 24.** Individual fluorescence lifetime components of 4Cl in single- and double-stranded RNA. Values were obtained from a triexponential fit.

	$\tau_1$	$\alpha_1$	$\tau_2$	$\alpha_2$	$\tau_3$	$\alpha_3$
	ns		ns		ns	
ssRNA <sup>[a]</sup>	9.58	0.64	5.34	0.18	1.03	0.19
dsRNA <sup>[b]</sup>	9.35	0.10	1.78	0.04	0.11	0.85

[a] **RXb**. [b] **R4<sub>ux</sub>**.

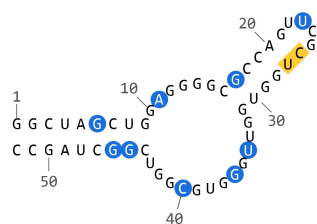
While the Chili aptamer is rich in guanines that can quench 4Cl fluorescence, this has been taken into consideration when designing the RNA sequences **R1–R4**. Accordingly, the findings presented in this section show that r4Cl will remain suitably emissive upon incorporation into the aptamer.

## 2.8 FÖRSTER RESONANCE ENERGY TRANSFER WITHIN THE CHILI APTAMER

### 2.8.1 Synthesis of 4-cyanoindole-modified Chili aptamer constructs

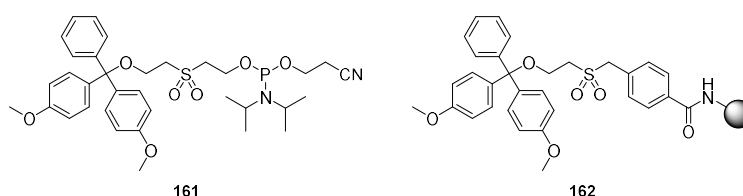
#### 2.8.1.1 Preparation of modified aptamer fragments

The choice of r4Cl incorporation sites and preparative approach towards fluorescent Chili aptamer constructs was inspired by previously established methodologies for the synthesis of site-specifically modified riboswitch domains.<sup>246,247</sup> Instead of SHAPE probing, which is the preferred tool to identify replacement sites for 2APfold analyses,<sup>188</sup> existing Tb<sup>3+</sup> digestion data<sup>111</sup> (section 2.1.2) and systematic mutagenesis results (section 2.3.2) were used to select a total of nine suitable positions for the incorporation of r4Cl. Four of these are located in the bottom and top stems of the aptamer, one in the loop, while the remainder is situated directly within the central bulge (Figure 67).



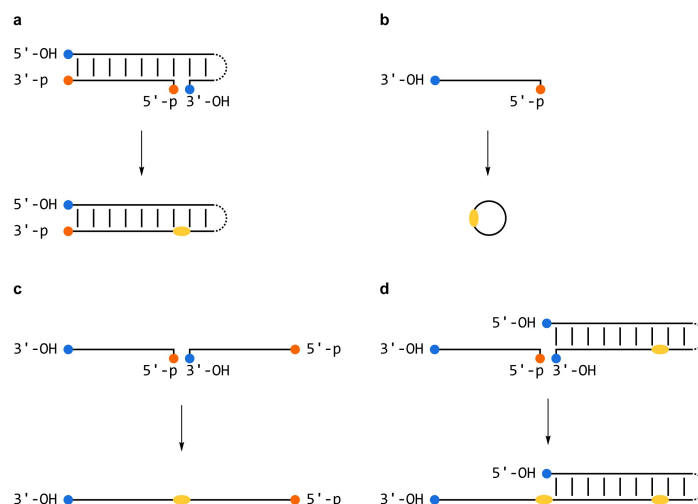
**Figure 67.** Nine positions of the Chili aptamer were modified by incorporation of r4Cl (blue). Full-length aptamers were prepared from 26 nt fragments by ligation with T4 RNA ligase. The ligation site is highlighted in yellow.

Although a modified 52 nt RNA is within reach of solid-phase synthesis using 2'-*O*-TOM-protected monomers,<sup>248</sup> it was deemed advantageous to synthesize shorter fragments instead, which would then be assembled into full-length aptamers by enzymatic ligation with T4 RNA ligase, which catalyzes the formation of a new phosphodiester bond between a 5'-phosphorylated donor fragment and an acceptor fragment bearing a 3'-hydroxy group. Nucleotides C26 and U27 were chosen as the ligation junction because they are prearranged in close proximity within the apical stem of the aptamer (Figure 67) and reasonably accommodate the preferences of T4 RNA ligase with regard to the nature of the acceptor ( $A > C \geq G > U$ ) and donor ( $pC > pU \approx pA > pG$ ) nucleotides.<sup>249,250</sup>



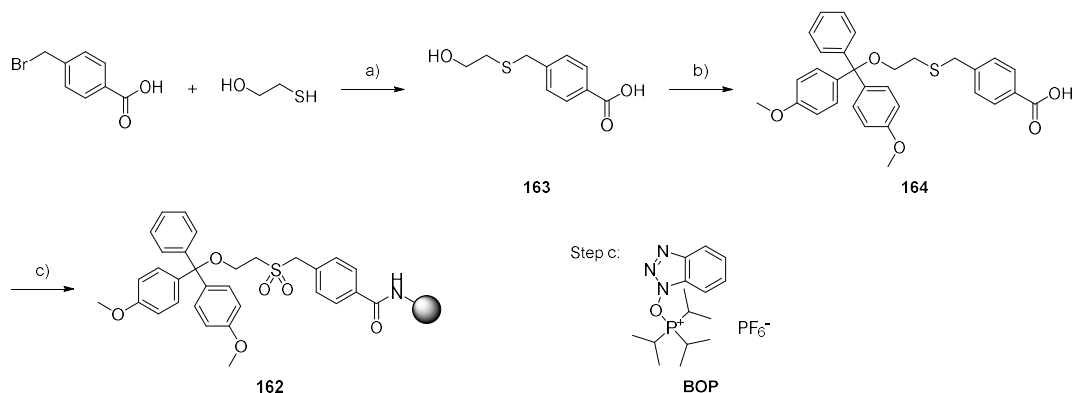
**Chart 29.** Amidite **161** was used for 5'-phosphorylation of the oligonucleotide fragments. The 3'-phosphate group was introduced using the modified CPG support **162**.

Donor oligonucleotides were 5'-phosphorylated directly during solid-phase synthesis by using the chemical phosphorylation amidite **161**, which has been prepared in this group previously (Chart 29). Additionally, donor oligonucleotides were 3'-phosphorylated by means of a functionalized CPG support (**162**) to prevent the formation of byproducts resulting from circularization of the donor, ligation with a second donor fragment or with the full-length product (Scheme 37, see section 4.1.4 for further details).



**Scheme 37.** a) Ligation between a 3',5'-bisphosphorylated donor oligonucleotide and an acceptor strand bearing a 3'-hydroxy group. The newly formed phosphodiester bond is indicated in yellow. If the donor strand is not phosphorylated at the 3'-end, side reactions such as b) circularization, c) dimerization or d) ligation with a full-length RNA might occur.

The CPG support was prepared starting from 4-bromomethylbenzoic acid, which was reacted with 2-mercaptoethanol to form compound **163** with 98% yield. After DMT-protection, the resulting intermediate **164** was loaded onto an amino-functionalized CPG support using standard amide coupling chemistry. Unreacted amino groups were capped with acetic anhydride and finally the thioether moiety was oxidized to a sulfone with *m*CPBA (Scheme 38). The loading density of **162** was determined from the absorbance of the trityl cation released upon treatment with dichloroacetic acid.



**Scheme 38.** Synthesis of the modified CPG support **162** for introduction of 3'-phosphate groups during RNA solid-phase synthesis. Reagents and conditions: a)  $\text{NET}_3$ , MeOH, r.t., 19 h, 98%. b) DMTCl, pyridine, r.t., 15 h, 72%. c) 1. Native amino CPG, BOP reagent,  $(i\text{Pr})_2\text{NET}$ , DMF, 40 °C, 2 h, 72%; 2.  $\text{Ac}_2\text{O}$ , 2,6-lutidine, *N*-methylimidazole, THF, r.t., 10 min; 3. *m*CPBA,  $\text{CH}_2\text{Cl}_2$ , r.t., 30 min.

Two native (**F1**, **F2**) and nine r4Cl-modified (**F3–F11**) 26 nt fragments of the Chili aptamer were prepared by solid-phase synthesis on a 600 nmol scale (section 4.1.3). An additional fragment (**F12**) containing d4Cl instead of r4Cl was synthesized as well in order to test the resulting ligation products for differences in function. Synthesis yields ranged between 36 and 119 nmol and all products were analyzed by HPLC and HR-MS (Table 25). Fragment **F10** was not phosphorylated completely during the synthesis and was therefore used in excess during the subsequent ligation step.

**Table 25.** Description of modified 26 nt Chili aptamer fragments obtained by solid-phase synthesis. See also section 4.3.2.

Number	Description <sup>[a]</sup>	Yield <sup>[b]</sup> nmol	<i>m/z</i> calc.	<i>m/z</i> obs.	Species
<b>F1</b>	Chili 5' fragment	36	8425.161	8425.315	[M-H] <sup>-</sup>
<b>F2</b>	Chili 3' fragment	105	8580.030	8580.140	[M-H] <sup>-</sup>
<b>F3</b>	Chili 5' fragment G6X	78	8416.166	8416.131	[M-H] <sup>-</sup>
<b>F4</b>	Chili 5' fragment A11X	38	8432.160	8432.295	[M-H] <sup>-</sup>
<b>F5</b>	Chili 5' fragment G17X	55	8416.166	8416.140	[M-H] <sup>-</sup>
<b>F6</b>	Chili 5' fragment U23X	69	8455.188	8455.142	[M-H] <sup>-</sup>
<b>F7</b>	Chili 3' fragment U34X	84	8610.056	8610.052	[M-H] <sup>-</sup>
<b>F8</b>	Chili 3' fragment G36X	64	8571.034	8571.144	[M-H] <sup>-</sup>
<b>F9</b>	Chili 3' fragment C40X	119	8611.040	8610.998	[M-H] <sup>-</sup>
<b>F10</b>	Chili 3' fragment G45X	88 <sup>[c]</sup>	8571.034	8571.098	[M-H] <sup>-</sup>
<b>F11</b>	Chili 3' fragment G46X	59	8571.034	8571.071	[M-H] <sup>-</sup>
<b>F12</b>	Chili 3' fragment G36Y	56	8555.039	8555.145	[M-H] <sup>-</sup>

[a] X = r4Cl, Y = d4Cl.

[b] 600 nmol synthesis scale.

[c] Incompletely 5'-phosphorylated.

### 2.8.1.2 Preparation of full-length aptamers by enzymatic ligation

Enzymatic ligations were performed on a 5 nmol scale using a 10% excess of the respective unmodified fragments **F1** or **F2** (section 4.1.4). Ligation products were purified by PAGE and characterized by HPLC and HR-MS. The observed isotope patterns of the charged species were in agreement with those calculated from their molecular formulas (Table 26).

**Table 26.** Description of modified full-length Chili aptamer constructs obtained by ligation of 26 nt fragments with T4 RNA ligase. See also section 4.3.3.

Number	Fragments <sup>[a]</sup>	Description <sup>[b]</sup>	Yield <sup>[c]</sup> nmol	<i>m/z</i> calc.	<i>m/z</i> obs.	Species
<b>L1</b>	<b>F3/F2</b>	Chili-G6X	1.92	771.047	771.105	[M-22H] <sup>22-</sup>
<b>L2</b>	<b>F4/F2</b>	Chili-A11X	12.5	808.574	808.576	[M-21H] <sup>21-</sup>
<b>L3</b>	<b>F5/F2</b>	Chili-G17X	2.87	771.047	771.103	[M-22H] <sup>22-</sup>
<b>L4</b>	<b>F6/F2</b>	Chili-U23X	2.52	809.670	809.729	[M-21H] <sup>21-</sup>
<b>L5</b>	<b>F1/F7</b>	Chili-U34X	2.36	809.670	809.336	[M-21H] <sup>21-</sup>
<b>L6</b>	<b>F1/F8</b>	Chili-G36X	0.60	807.812	807.917	[M-21H] <sup>21-</sup>
<b>L7</b>	<b>F1/F9</b>	Chili-C40X	1.92	772.866	772.925	[M-22H] <sup>22-</sup>
<b>L8</b>	<b>F1/F10</b>	Chili-G45X	1.85	771.047	771.102	[M-22H] <sup>22-</sup>
<b>L9</b>	<b>F1/F11</b>	Chili-G46X	2.73	771.047	771.061	[M-22H] <sup>22-</sup>
<b>L10</b>	<b>F1/F12</b>	Chili-G36Y	1.00	— <sup>[d]</sup>	—	—

[a] 5'-Acceptor fragment/3'-donor fragment used in the ligation reaction.

[b] X = r4Cl, Y = d4Cl.

[c] 5 nmol reaction scale except for **L2** (28 nmol).

[d] Not measured.

## 2.8.2 Observation of intrasupramolecular FRET

### 2.8.2.1 Spectroscopic and functional properties of modified Chili aptamer constructs

Fluorescence spectra of the synthesized Chili aptamer constructs in the absence of ligand show that the modified nucleobase remained emissive in all present cases (Figure 68). Its emission maximum varied between 393 and 404 nm (compared to 412 nm for the free nucleoside), indicating different levels of interaction with surrounding nucleotides. This is further exemplified

by the respective fluorescence decays, which required between two and four exponential components for proper fitting, yielding amplitude average lifetimes between 0.96 and 6.96 ns (Supplementary Figure 11). The relative quantum yield of 4CI was estimated for each construct from the ratio of the measured lifetime to that of the free nucleoside (Table 27). While this approach neglects the conceivable formation of dark states (see section 2.7.2), it was deemed appropriate for a qualitative assessment of the photophysical properties to the degree required for this investigation.

**Table 27.** Steady-state and fluorescence lifetime parameters of modified Chili aptamer constructs in the absence of ligand. Conditions: 0.5  $\mu\text{M}$  RNA, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Number	Description <sup>[a]</sup>	$I-I_0$ <sup>[b]</sup>	$\langle\tau\rangle$ <sup>[c]</sup> ns	$\lambda_{\text{em, max}}$ nm	$\phi_{\text{rel}}$ <sup>[d]</sup>
<b>L1</b>	Chili-G6X	1392 $\pm$ 102	6.80 $\pm$ 0.12	397	0.48
<b>L2</b>	Chili-A11X	613 $\pm$ 42	3.60 $\pm$ 0.28	404	0.25
<b>L3</b>	Chili-G17X	161 $\pm$ 6	0.96 $\pm$ 0.03	393	0.07
<b>L4</b>	Chili-U23X	2793 $\pm$ 94	6.96 $\pm$ 0.05	404	0.49
<b>L5</b>	Chili-U34X	591 $\pm$ 9	2.14 $\pm$ 0.16	401	0.15
<b>L6</b>	Chili-G36X	944 $\pm$ 25	4.00 $\pm$ 0.20	404	0.28
<b>L7</b>	Chili-C40X	762 $\pm$ 66	4.05 $\pm$ 0.12	404	0.29
<b>L8</b>	Chili-G45X	297 $\pm$ 26	1.41 $\pm$ 0.13	401	0.10
<b>L9</b>	Chili-G46X	1179 $\pm$ 38	3.94 $\pm$ 0.40	404	0.28
<b>L10</b>	Chili-G36Y	462 $\pm$ 5	3.13 <sup>[e]</sup>	405	— <sup>[f]</sup>

[a] X = r4Cl, Y = d4Cl.

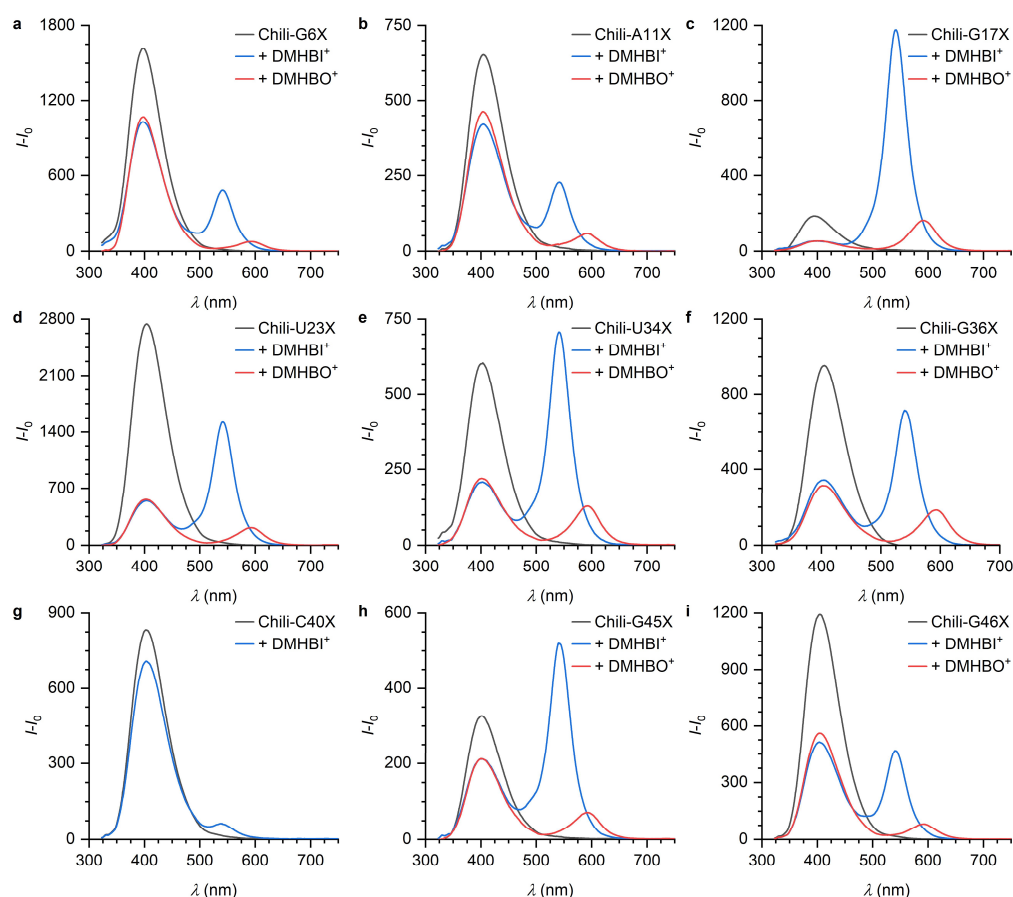
[b] Steady-state intensities ( $\lambda_{\text{ex/em}} = 303/410$  nm) after blank correction (buffer).

[c] Amplitude average fluorescence lifetime ( $\lambda_{\text{ex/em}} = 318/410$  nm) for multiexponential fits (Supplementary Figure 11).

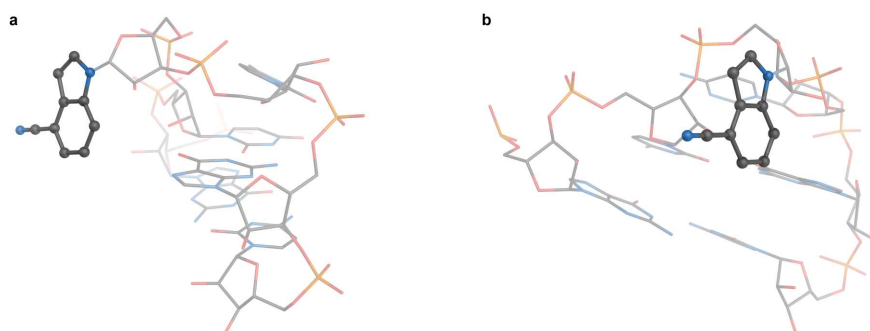
[d] Estimated from the amplitude average lifetime relative to that of r4Cl ( $\tau = 9.48$  ns,  $\phi_{\text{abs}} = 0.67$ ).

[e] Single experiment. [f]  $\phi_{\text{abs}}$  of d4Cl was not measured independently.

Both quantum yield and lifetime were highest for Chili-U23X (**L4**,  $\phi_{\text{rel}} \approx 0.49$ ,  $\langle\tau\rangle = 6.96$  ns), which also displays the least complex decay kinetics. Here, r4Cl replaces the second nucleotide in the apical UUCG tetraloop of the aptamer (Figure 67 on page 100). This second nucleotide of UUCG tetraloops, which generally are well conserved structure motifs, is conformationally flexible and bulged out of the base stack into the solvent environment (Figure 69).<sup>251,252</sup> Interestingly, the quantum yield and lifetime were very similar when the modification replaced G6 in the bottom stem of the Chili aptamer (**L1**,  $\phi_{\text{rel}} \approx 0.48$ ,  $\langle\tau\rangle = 6.80$  ns), indicating that this region is relatively flexible. Indeed, <sup>1</sup>H NMR data of the native Chili aptamer corroborate that the stem is incompletely formed in the absence of ligand.<sup>112</sup> Incorporation of d4Cl instead of r4Cl at position 36 in the central bulge resulted in a construct (**L10**) that displayed weaker fluorescence. Presumably, different conformations of the sugar unit lead to shifted positions of the modified nucleobase with respect to its environment, which incidentally increases quenching in **L10**.



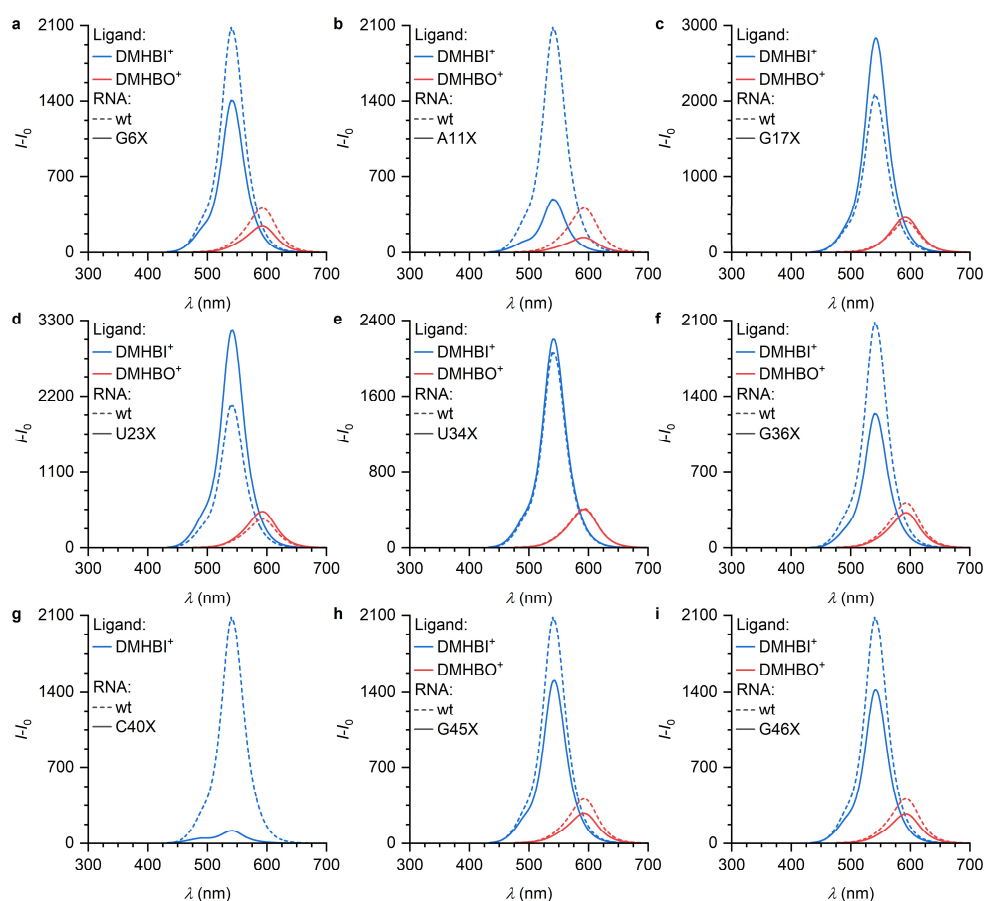
**Figure 68.** Fluorescence emission spectra ( $\lambda_{\text{ex}} = 303 \text{ nm}$ ) of modified Chili aptamer constructs alone and in the presence of ligand. Data for Chili-G36Y (**L10**) is shown in Supplementary Figure 14. Conditions:  $0.5 \mu\text{M}$  RNA,  $125 \text{ mM}$  KCl,  $5 \text{ mM}$   $\text{MgCl}_2$ ,  $40 \text{ mM}$  HEPES pH 7.5, +  $1 \mu\text{M}$  ligand.



**Figure 69.** a) Side-on and b) frontal view of a proposed structural model for the gUNCGc tetraloop (N = 4-cyanoindole) as it is found in Chili-U23X. The modified base protrudes out of the base stack, exposing it to the solvent environment. The model was adapted from the reported NMR structure of a UUCG tetraloop (PDB: 4E8Q) by replacing the second uridine residue with 4CI (DFT-optimized structure, B3LYP-D3/def2-TZVP). Carbon atoms 2 and 7a of 4CI were matched to carbon atoms 6 and 2 of U, respectively.

Next, it was tested whether the modified Chili aptamer constructs retained the ability to activate the fluorescence of  $\text{DMHBI}^+$  and  $\text{DMHBO}^+$ . After overnight incubation with a twofold excess of either ligand, all constructs except Chili-C40X (**L7**, not investigated with  $\text{DMHBO}^+$ ) showed proper fluorescence emission when the bound ligand was excited directly at  $413 \text{ nm}$  ( $\text{DMHBI}^+$ ) or  $456 \text{ nm}$  ( $\text{DMHBO}^+$ ), respectively. Steady-state intensities varied between the constructs and were usually different to that of the unmodified aptamer (Figure 70). The behavior of Chili-C40X will be discussed in more detail later in this section.



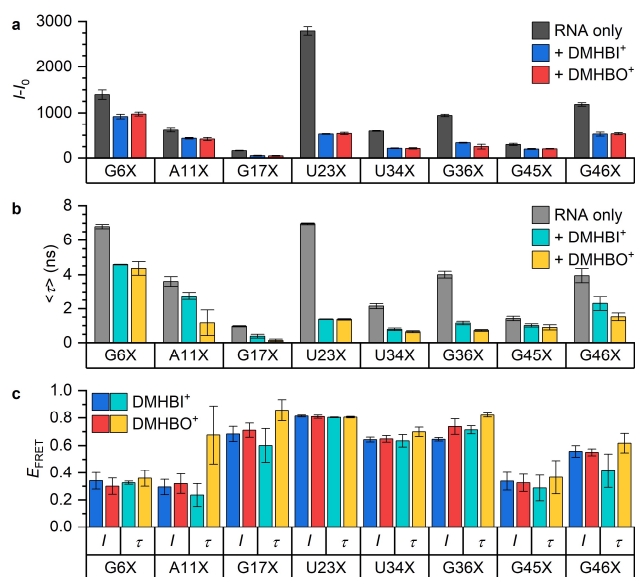


**Figure 70.** Fluorescence emission spectra ( $\lambda_{\text{ex}} = 413 \text{ nm}$  for DMHBI<sup>+</sup>,  $456 \text{ nm}$  for DMHBO<sup>+</sup>) of modified Chili aptamer constructs in the presence of ligand. Data for Chili-G36Y (**L10**) is shown in Supplementary Figure 14. Conditions:  $0.5 \mu\text{M}$  RNA,  $1 \mu\text{M}$  ligand,  $125 \text{ mM}$  KCl,  $5 \text{ mM}$  MgCl<sub>2</sub>,  $40 \text{ mM}$  HEPES pH 7.5.

When the modified Chili–ligand complexes were excited at  $303 \text{ nm}$ , donor-sensitized emission from the respective ligand was observed. At the same time, the steady-state fluorescence intensity (Figure 68) and lifetime (Supplementary Figure 12, 5) of the donor 4CI decreased as a result of energy transfer to the ligand. The emission maximum of 4CI did not change for a given construct upon ligand complexation, suggesting that differences with respect to the free nucleoside are due to the immediate sequence context rather than more remote tertiary interactions.

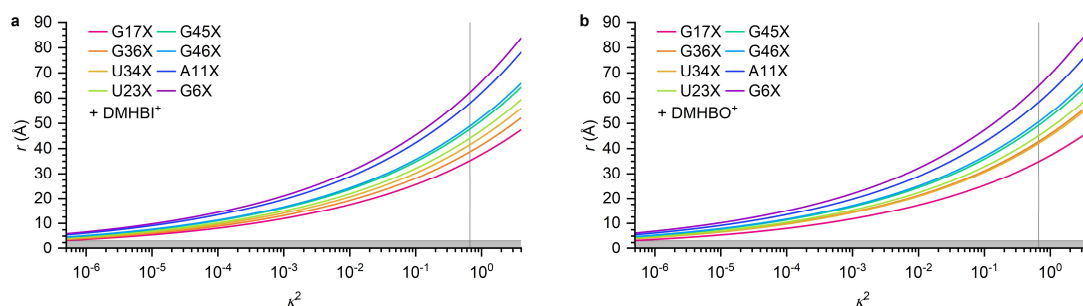
### 2.8.2.2 Quantification of intrasupramolecular FRET

FRET efficiencies were quantified according to the formula  $E_{\text{FRET}} = 1 - (F_{\text{bound}}/F_{\text{free}})$ , where  $F$  represents either fluorescence intensity ( $I$ ) or lifetime ( $\tau$ ). Generally, the results obtained from both methods ( $I$  and  $\tau$ ) were in excellent agreement not only with each other, but also across both ligands for a given Chili construct (Figure 71, Table 28, 16 on page 107). This can only be the case if the distance and orientation between the FRET donor and either acceptor are comparable, i.e. if both ligands occupy the binding site of the aptamer in a similar fashion. Interestingly, **L10**, which contains d4CI at position 36, displays notably smaller FRET efficiencies than the all-RNA construct **L6**, implying that these two modified aptamers are not entirely structurally homologous.



**Figure 71.** a) Steady-state fluorescence intensity ( $\lambda_{\text{ex/em}} = 303/410$  nm) and b) amplitude average fluorescence lifetimes ( $\lambda_{\text{ex/em}} = 303/410$  nm) of modified Chili aptamer constructs before and after addition of the indicated ligand. c) Calculated FRET efficiencies for each modification site. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

The spatial information encoded in FRET efficiencies can in principle be used as a spectroscopic ruler of sorts to determine the distance between the donor and acceptor. This requires knowledge about the mutual orientation of the transition dipole moments of the FRET partners, which is quantified in terms of the orientation factor  $\kappa^2$  (see section 4.1.21.1 for details). In the dynamic isotropic regime, the FRET partners are able to freely reorient themselves within the donor excited state lifetime, resulting in  $\kappa^2 = 2/3$ . However, the existence of a well-defined ligand binding site implies that the acceptor fluorophore is restricted to one specific orientation within the Chili aptamer reference frame. Moreover, donor fluorophores embedded in base-paired sections of the aptamer are very likely biased towards intrahelically stacked geometries (see section 2.7.2). In this case, a structural model of the 4CI labeling sites would be necessary to triangulate the ligand position.<sup>253,254</sup> Further analysis of the collected FRET data was therefore limited to the evaluation of a combined distance/orientation parameter. Spectral overlap integrals ( $J$ ) for each FRET pair were calculated under the assumption that the estimated fluorescence quantum yield of 4CI in a given Chili aptamer construct was not altered by ligand uptake, which is based on the observation that the emission maximum of the fluorophore did not change throughout this process. As a superficial guide, isotropic Förster radii ( $R_0$  for  $\kappa^2 = 2/3$ ) were then calculated as described in section 4.1.21.1. Their values are in agreement between both ligands (Table 30). For the combined data analysis, the donor–acceptor separation was plotted as a function of  $\kappa^2$  with the experimental FRET efficiencies as an adjustable parameter (Figure 72).



**Figure 72.** Relationship between the orientation factor  $\kappa^2$  and possible donor–acceptor distances in the modified Chili aptamer constructs. The curves were constructed from equation 111 in section 4.1.21.1 based on the data in Table 30. Vertical lines indicate  $\kappa^2 = 2/3$ , which would be valid in the dynamic isotropic regime. Distances below 3.2 Å, which corresponds to the orthogonal separation of the ligand and a G quartet in the crystal structure of the Spinach aptamer,<sup>82</sup> are shaded in grey.

The strongest coupling between donor and acceptor, which in this context means a favorable combination of distance and mutual transition dipole moment orientation, occurs when the 4Cl modification site is located at position 17 (**L3**) in the apical stem. In contrast, placement at position 6 (**L1**) in the basal stem results in the weakest coupling. Given that this position and also position 11 in the binding site are predicted to be in fairly close proximity to the ligand (Figure 67) and considering the accessible size of a 52 nt RNA aptamer (shorter than an extended 26 bp duplex with a length of approximately 70 Å<sup>255</sup>), this suggests donor–acceptor orientations resulting in a  $\kappa^2$  significantly smaller than 2/3 for these two positions.

**Table 28.** Steady-state and fluorescence lifetime parameters of modified Chili aptamer constructs in the presence of DMHBI<sup>+</sup>. Conditions: 0.5 μM RNA, 1 μM ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Values are given as mean ± s.d. (n = 3).

Number	Description <sup>[a]</sup>	$I-I_0$ <sup>[b]</sup>	$\langle\tau\rangle$ <sup>[c]</sup> ns	$E_f$	$E_r$
<b>L1</b>	Chili-G6X	918 ± 51	4.56 ± 0.02	0.34 ± 0.06	0.32 ± 0.01
<b>L2</b>	Chili-A11X	432 ± 18	2.74 ± 0.22	0.29 ± 0.06	0.24 ± 0.03
<b>L3</b>	Chili-G17X	51 ± 9	0.38 ± 0.11	0.68 ± 0.05	0.60 ± 0.12
<b>L4</b>	Chili-U23X	524 ± 10	1.38 ± 0.20	0.81 ± 0.01	0.80 ± 0.01
<b>L5</b>	Chili-U34X	212 ± 10	0.78 ± 0.08	0.64 ± 0.02	0.63 ± 0.05
<b>L6</b>	Chili-G36X	335 ± 8	1.15 ± 0.10	0.65 ± 0.01	0.71 ± 0.02
<b>L8</b>	Chili-G45X	197 ± 9	1.01 ± 0.10	0.34 ± 0.07	0.29 ± 0.09
<b>L9</b>	Chili-G46X	523 ± 46	2.30 ± 0.42	0.56 ± 0.04	0.41 ± 0.12
<b>L10</b>	Chili-G36Y	243 ± 13	1.23 <sup>[d]</sup>	0.47 ± 0.03	0.61 <sup>[d]</sup>

[a] X = r4Cl, Y = d4Cl.

[b] Steady-state intensities ( $\lambda_{\text{ex/em}} = 303/410$  nm) after blank correction (buffer).

[c] Amplitude average fluorescence lifetime ( $\lambda_{\text{ex/em}} = 318/410$  nm) for multiexponential fits (Supplementary Figure 12).

[d] Single experiment.

**Table 29.** Steady-state and fluorescence lifetime parameters of modified Chili aptamer constructs in the presence of DMHBO<sup>+</sup>. Conditions: 0.5 μM RNA, 1 μM ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Values are given as mean ± s.d. (n = 3).

Number	Description <sup>[a]</sup>	$I-I_0$ <sup>[b]</sup>	$\langle\tau\rangle$ <sup>[c]</sup> ns	$E_f$	$E_r$
<b>L1</b>	Chili-G6X	973 ± 42	4.36 ± 0.39	0.30 ± 0.06	0.36 ± 0.06
<b>L2</b>	Chili-A11X	416 ± 34	1.17 ± 0.75	0.32 ± 0.07	0.67 ± 0.21
<b>L3</b>	Chili-G17X	47 ± 8	0.14 ± 0.07	0.71 ± 0.05	0.86 ± 0.08
<b>L4</b>	Chili-U23X	537 ± 27	1.36 ± 0.04	0.81 ± 0.01	0.80 ± 0.01
<b>L5</b>	Chili-U34X	209 ± 14	0.65 ± 0.06	0.65 ± 0.02	0.70 ± 0.03
<b>L6</b>	Chili-G36X	249 ± 53	0.72 ± 0.05	0.74 ± 0.06	0.82 ± 0.02
<b>L8</b>	Chili-G45X	200 ± 7	0.89 ± 0.15	0.33 ± 0.06	0.37 ± 0.12
<b>L9</b>	Chili-G46X	531 ± 24	1.51 ± 0.23	0.55 ± 0.02	0.62 ± 0.07

[a] X = r4Cl.

[b] Steady-state intensities ( $\lambda_{\text{ex/em}} = 303/410$  nm) after blank correction (buffer).

[c] Amplitude average fluorescence lifetime ( $\lambda_{\text{ex/em}} = 318/410$  nm) for multiexponential fits (Supplementary Figure 13).

**Table 30.** Calculated spectral overlap integrals ( $J$ ) and isotropic Förster radii ( $R_0$ ) for the FRET pairs formed from 4CI and Chili-bound DMHBI<sup>+</sup> and DMHBO<sup>+</sup>, respectively.

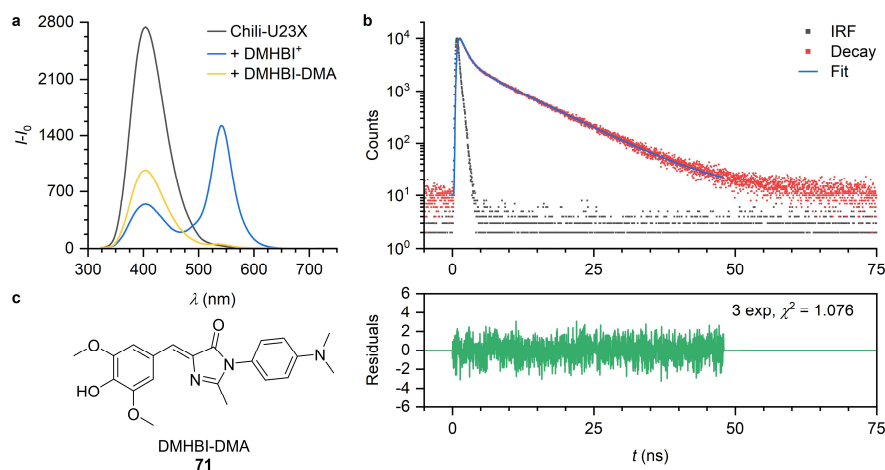
Number	Description <sup>[a]</sup>	$J$		$R_0$ <sup>[b]</sup>	
		L mol <sup>-1</sup> cm <sup>-1</sup> nm <sup>4</sup>	Å	L mol <sup>-1</sup> cm <sup>-1</sup> nm <sup>4</sup>	Å
		+ DMHBI <sup>+</sup>		+ DMHBO <sup>+</sup>	
<b>L1</b>	Chili-G6X	$4.1 \times 10^{14}$	56	$4.4 \times 10^{14}$	56
<b>L2</b>	Chili-A11X	$4.1 \times 10^{14}$	50	$4.9 \times 10^{14}$	52
<b>L3</b>	Chili-G17X	$3.8 \times 10^{14}$	40	$4.1 \times 10^{14}$	40
<b>L4</b>	Chili-U23X	$4.2 \times 10^{14}$	56	$4.8 \times 10^{14}$	57
<b>L5</b>	Chili-U34X	$4.0 \times 10^{14}$	46	$4.5 \times 10^{14}$	47
<b>L6</b>	Chili-G36X	$4.1 \times 10^{14}$	51	$4.4 \times 10^{14}$	52
<b>L8</b>	Chili-G45X	$4.0 \times 10^{14}$	43	$4.6 \times 10^{14}$	44
<b>L9</b>	Chili-G46X	$4.1 \times 10^{14}$	51	$4.8 \times 10^{14}$	52

[a] X = r4Cl.

[b] Isotropic Förster radius assuming an orientation factor  $\kappa^2$  of 2/3.

### 2.8.2.3 Interaction of Chili with a non-fluorogenic ligand

In addition to mapping the ligand binding site as discussed up to this point, the presence of a FRET reporter system also gives direct insight into the binding of non-fluorogenic ligands, which are spectroscopically silent under the conditions of standard fluorescence screening assays. The high FRET efficiency and initial brightness observed for Chili-U23X (**L4**), combined with the remote location of the modification site in the apical tetraloop, make this construct a suitable platform for such investigations. Indeed, incubation with DMHBI-DMA (**71**), the uncharged precursor of DMHBI<sup>+</sup>, resulted in a significant decrease of the 4CI steady-state emission and fluorescence lifetime due to energy transfer, albeit to a lesser degree than observed with DMHBI<sup>+</sup> or DMHBO<sup>+</sup> (Figure 73, Table 31).



**Figure 73.** The non-fluorogenic ligand DMHBI-DMA (**71**) binds to the Chili aptamer, which is observed by FRET quenching of 4CI incorporated at position 23 (**L4**). a) Steady-state emission spectra ( $\lambda_{\text{ex}} = 303$  nm) before and after ligand addition. In contrast to DMHBI<sup>+</sup>, no donor-sensitized emission is observed. b) Top panel: Fluorescence lifetime measurement with triexponential fit ( $\lambda_{\text{ex/em}} = 303/412$  nm). Bottom panel: Weighted residuals and  $\chi^2$  for the fit. Conditions: 0.5  $\mu\text{M}$  RNA, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5, + 1  $\mu\text{M}$  ligand.

**Table 31.** Steady-state and fluorescence lifetime parameters of Chili-U23X in the presence of DMHBI-DMA. Conditions: 0.5  $\mu\text{M}$  RNA, 1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5. Values are given as mean  $\pm$  s.d. ( $n = 3$ ).

Number	Description <sup>[a]</sup>	$I-I_0$ <sup>[b]</sup>	$\langle\tau\rangle$ <sup>[c]</sup>	$E_f$	$E_\tau$
			ns		
<b>L4</b>	Chili-U23X	$961 \pm 15$	$2.25 \pm 0.01$	$0.66 \pm 0.01$	$0.68 \pm 0.01$

[a] X = r4Cl.

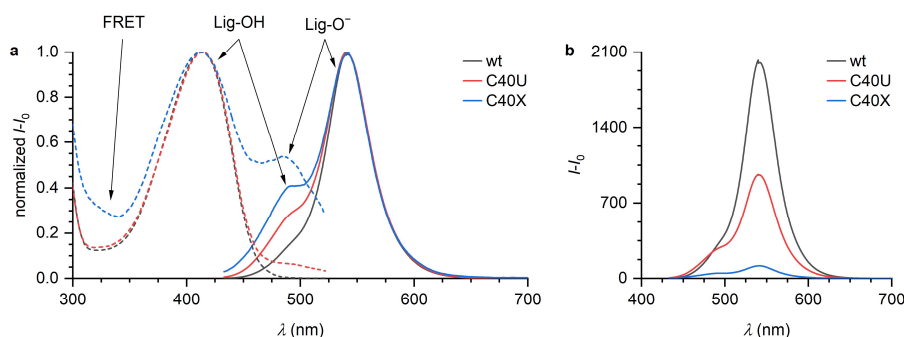
[b] Steady-state intensities ( $\lambda_{\text{ex/em}} = 303/410$  nm) after blank correction (buffer).

[c] Amplitude average fluorescence lifetime ( $\lambda_{\text{ex/em}} = 318/410$  nm) for triexponential fit.

This result unequivocally confirms association between Chili and DMHBI-DMA, which could not be concluded based on earlier screening data alone (see section 2.3.1.3). The lower FRET efficiency relative to both of the other ligands indicates that either DMHBI-DMA occupies the binding site in a slightly different orientation, possibly due to weaker electrostatic interactions, or binds with a smaller affinity, as was also observed for other neutral ligands with aromatic N3 substituents (see section 2.4.3.1).

#### 2.8.2.4 Evidence for RNA/ligand proton exchange

As mentioned earlier, Chili-C40X (**L7**) activated the fluorescence of DMHBI<sup>+</sup> the weakest among all modified constructs (Figure 70). Upon closer inspection, it became obvious that the emission spectrum ( $\lambda_{\text{em}} = 413$  nm) of Chili-C40X–DMHBI displayed a pronounced shoulder at approximately 485 nm, which was also observed for sufficiently concentrated samples of unbound DMHBI<sup>+</sup> (see section 2.4.1.3) and originated from the neutral phenol form of the ligand. Moreover, the excitation spectrum ( $\lambda_{\text{em}} = 542$  nm) of the complex contained an additional band at 485 nm, indicating that a fraction of RNA molecules bind the ligand in its anionic phenolate form (Figure 74a).



**Figure 74.** a) Normalized excitation and emission spectra ( $\lambda_{\text{ex/em}} = 413/542$  nm) of DMHBI<sup>+</sup> in complex with wt Chili and its variants containing r4Cl or U instead of C at position 40. The origin of each band is labeled. b) The fluorescence enhancement of the ligand is diminished by replacing C40. Conditions: 0.5  $\mu\text{M}$  RNA, 1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM  $\text{MgCl}_2$ , 40 mM HEPES pH 7.5.

This finding prompted a new examination of Chili mutant C40U, which had previously been used for systematic mutagenesis studies of the ligand binding site (see section 2.3.2.1). Indeed, Chili-C40U showed, to a lesser extent, the same spectral characteristics as the 4Cl-modified construct. Fluorescence enhancement of DMHBI<sup>+</sup> was at an intermediate level with respect to wt Chili and Chili-C40X (Figure 74b). Overall, these observations firmly establish nucleotide 40 as part of the proton transfer network (see Scheme 8 in section 2.2.1) that is responsible for the large Stokes shift characteristic of the Chili aptamer.



### 3 CONCLUSION AND OUTLOOK

In the course of this thesis, a fluorogenic aptamer system based on the 52 nt DMHBI-binding Chili RNA has been characterized, optimized and applied as a fluorescent reporter for gel-based visualization of DNA enzyme activity. A library of chemically modified HBI derivatives was synthesized and has been used to explore structure–fluorescence activation relationships. Chili differs from other HBI-activating RNA aptamers as it induces fluorescence emission with Stokes shifts upwards of 120 nm by altering the protonation state of its ligands. Moreover, it was found that the binding affinity and fluorescence intensity of Chili complexes were greatly augmented by including a quaternary ammonium ion in the side chain of its ligands. This is a specific feature of the Chili aptamer and it was explored to enhance both the binding and fluorescence of various ligands, furnishing complexes with emission wavelengths that cover a broad range of the visible spectrum. It was also shown in this work that the related Spinach and Broccoli aptamers reject their cognate fluorinated ligands when these were equipped with a permanent positive charge. The dissociation constants of Chili–ligand complexes are highly sensitive to the precise structure of the ligand, reaching low nanomolar values in the best case, which is the highest affinity yet reported for an HBI-binding aptamer. Systematic mutagenesis studies and temperature-dependent optical spectroscopy strongly suggest that the aptamer is loosely preorganized in the apo state and that interaction with the ligand stabilizes a G-quadruplex within the binding site. G-quadruplexes are a common structural motif in ligand binding sites among different classes of fluorogenic aptamers such as Spinach, Corn and Mango variants. The quadruplexes in these aptamers are, however, already formed prior to ligand binding so that only limited structural reorganization takes place upon complex formation. The ligand binding dynamics of Chili are therefore highly unusual for a fluorogenic aptamer, warranting further studies of the underlying mechanism.

Structure elucidation efforts have continued with the design of an intrasupramolecular FRET pair comprising a fluorescent nucleoside as the donor and Chili-bound ligand as the acceptor. To this end, a phosphoramidite building block for the site-specific incorporation of 4-cyanoindole ribonucleoside into RNA has been synthesized. 4-Cyanoindole is a universal nucleobase analog, i.e. it is not biased towards hybridization with a specific canonical nucleobase. While several universal fluorescent deoxyribonucleosides have been described before, the 4-cyanoindole derivative reported here is the first example of such a ribonucleoside. It has been characterized in detail and its properties suggest that it will be widely applicable for studying other functional RNA structures. In this work, it was shown that Chili aptamer variants modified with 4-cyanoindole remained binding-competent and displayed position-specific FRET efficiencies after ligand binding. Distance-orientation parameters were calculated from these results and will help to support ongoing crystallographic and NMR-based investigations into the tertiary structure of Chili. A major open question concerns the precise mechanism of ligand binding and fluorescence activation responsible for the observed large Stokes shift. This work provided a first insight into active site nucleotides that may be involved in binding the protonated ligand in its ground state or in the deprotonation of the excited state. Joint efforts with other members of the research group are expected to give more detailed insights into the three-dimensional structure of the Chili aptamer in its ligand-bound state in the near future.

Further directions of research concern the use of Chili as a fluorescent reporter for in vitro assays and its performance in live cell imaging applications. It will be of particular interest to see how the unique properties of Chili, such as large Stokes shift emission, complement existing fluoromolecules. One project that aims to establish Chili as a reporter module in a drug screening

assay is currently in its initial phase. The long-term goal is to identify small molecules with drug-like properties that interfere with essential regulatory RNA elements in the life cycle of pathogenic bacteria. For example, the T-box riboswitch of *S. aureus*, which regulates the de novo biosynthesis of methionine, has recently been proposed as a drug target for combating this prime example of a multiresistant gram-positive pathogen. Here, the challenge lies in monitoring the influence of potential drug candidates at the RNA level. It is proposed that Chili could potentially be used to directly monitor conformational changes of the RNA target in response to interaction with drug candidates or it could be established as a fluorescence reporter that is transcribed under the control of the riboswitch. This is just one example, but many other prospective directions of implementing Chili in biosensor or other analytical devices are conceivable, both for RNA research in particular and for chemical biology in general.



## 4 EXPERIMENTAL PART

### 4.1 MATERIALS AND METHODS

#### 4.1.1 General remarks

Section 4.1 describes the biochemical and spectroscopic techniques employed to obtain the results presented in this thesis. Each subsection contains a brief outline of the materials and instrumentation used for the experiments. General protocols are provided that detail sample preparation, data acquisition and analysis. For some of the more specialized techniques, a brief introduction is included in which the fundamentals of the method are discussed and key equations are derived.

Synthetic methods for the preparation of small molecules and oligonucleotides is given in sections 4.2 and 4.3. Computational methods are outlined in section 4.4.

The preparation of commonly used buffer solution and reagents, referenced throughout the procedures, is described in the appendix (section 0) or otherwise indicated in the materials paragraph of the respective procedure.

#### 4.1.2 In vitro transcription

##### 4.1.2.1 Overview

Enzymatic synthesis by in vitro transcription is the method of choice for the preparation of unmodified RNA oligonucleotides. A synthetic DNA template codes for the target RNA sequence downstream of a promotor binding region. After annealing of this promotor to the template, transcription is initiated in 5'-to-3' direction. For T7 RNA polymerase, the structural requirements of the initiation site necessitate the presence of two consecutive guanines at the 5'-end of the transcript.<sup>256,257</sup>

##### 4.1.2.2 Materials

DNA templates for in vitro transcription were purchased from Microsynth and purified by denaturing PAGE (15–20%). Ribonucleoside triphosphates (NTPs) were purchased from Jena Bioscience.

T7 RNA polymerase was prepared in house (by Ann-Kathrin Lenz) following a published procedure with minor modifications.<sup>258</sup>

##### 10x annealing buffer

TRIS (40 mM, pH 8.0), EDTA (1 mM), NaCl (150 mM)

##### 10x reaction buffer

TRIS (400 mM, pH 8.0), MgCl<sub>2</sub> (300 mM), DTT (100 mM), spermidine (20 mM)

##### 4.1.2.3 General protocol

Reactions were typically performed in a volume of 100  $\mu$ L.

The DNA template (2  $\mu$ M) and T7 promotor D008 (2  $\mu$ M) were dissolved in 1x annealing buffer and annealed by heating to 95 °C for 5 min in a volume of 50  $\mu$ L. After cooling to ambient temperature, the NTPs (4 mM each), 10x reaction buffer and T7 RNA polymerase (4  $\mu$ L, stock concentration 4 mg  $\mu$ L<sup>-1</sup>) were added. The mixture was brought up to 100  $\mu$ L with H<sub>2</sub>O and incubated at 37 °C for 4 h, by which time it had become turbid due to precipitation of magnesium pyrophosphate. To stop the reaction and dissolve the precipitate, 0.5 M EDTA pH 8.0 (20  $\mu$ L) was added. Finally,

loading buffer (80  $\mu$ L) was added and the transcript was purified by denaturing PAGE (10%). The isolated product was analyzed by HPLC.

Typical yields were 2–4 nmol RNA from 100  $\mu$ L transcription reactions as determined by UV absorbance.

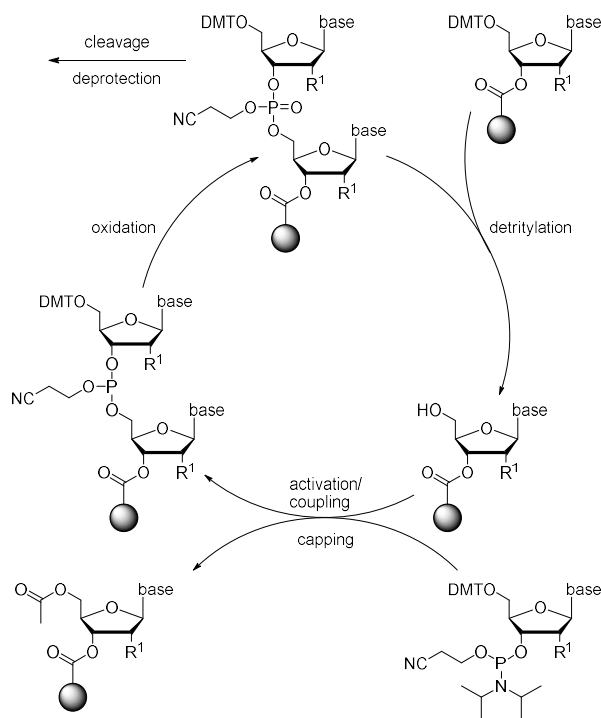
### 4.1.3 Solid-phase synthesis

#### 4.1.3.1 Overview

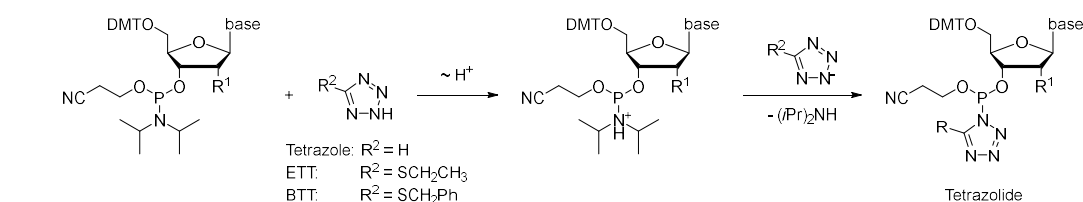
Modified DNA and RNA oligonucleotides are commonly prepared by automated solid-phase synthesis using phosphoramidite building blocks. To achieve selective coupling of these monomers, reactive groups need to be masked using a suitable set of orthogonal protecting groups that can be removed at appropriate times. Commonly, 5' hydroxy groups are protected with the acid-labile 4,4'-dimethoxytrityl (DMT) group, which is also used to monitor the progress of the synthesis. Amino groups on the nucleobases are protected using base-labile acyl groups that are removed during cleavage of the oligonucleotide from the solid support.<sup>259</sup>

Importantly, during RNA solid-phase synthesis, an additional 2' hydroxy protecting group is required in order to prevent branching of the nascent oligonucleotide. In this laboratory, the acetal-based [(triisopropylsilyl)oxy]methyl (TOM) protecting group is preferred due to higher coupling yields and stability towards migration compared to silyl ethers such as *tert*-butyldimethylsilyl (TBDMS) or triisopropylsilyl (TIPS).<sup>260–262</sup>

Each reaction cycle for solid-phase synthesis (Scheme 39) comprises four steps. First, the terminal DMT group of the (oligo)nucleotide is removed by treatment with a moderately strong organic acid such as di- or trichloroacetic acid (DCA or TCA, respectively). Next, the incoming phosphoramidite building block is activated by treatment with a tetrazole-based reagent (Scheme 40) and can react with the liberated 5' hydroxyl group to form a phosphite triester. In the following capping step, unreacted 5' hydroxy groups are reacted with a mixture of acetic anhydride and suitable base as well as a catalyst. Without capping, deletion sites would accumulate and could potentially interfere with the isolation of the desired full-length oligonucleotide. Finally, oxidation with iodine converts the acid-labile phosphite triester to a phosphate triester linkage, which prevents degradation during the detritylation step of the next cycle.



**Scheme 39.** Reaction cycle for the automated solid-phase synthesis of DNA and RNA. The 3' terminal nucleotide is attached to a solid support. Removal of the DMT protecting group enables coupling with the tetrazolid derivative (Scheme 40) of the next phosphoramidite building block in the sequence. Unreacted 5' hydroxy functions are blocked with acetyl groups to prevent further elongation in subsequent cycles. Oxidation converts the acids sensitive P<sup>III</sup> linkage to P<sup>V</sup> in preparation for the next detritylation step. After all nucleotides have been incorporated, the oligonucleotide is cleaved from the solid support and all protecting groups are removed.<sup>259</sup>



**Scheme 40.** During the activation step, the diisopropylamino group of the incoming phosphoramidite is protonated by the activator, which facilitates nucleophilic substitution. Especially for RNA solid-phase synthesis, ETT and BTT are preferred over tetrazole itself due to their higher reactivity.<sup>263</sup>

Treatment with ammonia or amine bases leads to cleavage of the synthesis product from the solid support and will also remove base-labile protecting groups. For RNA, incubation with a fluoride source such as tetrabutylammonium fluoride (TBAF) is required to remove the 2' hydroxy protecting groups. In this case, the byproducts of the deprotection are removed by desalting. The full-length oligonucleotide is finally isolated and purified by gel electrophoresis (see section 4.1.5).

#### 4.1.3.2 Materials and instrumentation

DNA and RNA oligonucleotides were prepared on an ABI 392 DNA/RNA Synthesizer or Pharmacia LKB Gene Assembler Plus.

Standard 5'-O-DMT- or 5'-O-DMT/2'-O-TOM-protected 3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidites were purchased from SAFC or ChemGenes. 2-[2-(4,4'-Dimethoxytrityloxy)ethyl-sulfonyl]ethyl-[(2-cyanoethyl)-(N,N-diisopropyl)]phosphoramidite for chemical 5'-phosphorylation has previously been prepared in this group following a published procedure.<sup>264</sup> Standard 5'-O-DMT- or 5'-O-DMT/2'-O-Ac-protected CPG supports (1000 Å, 25–35 μmol g<sup>-1</sup> for DNA, 20–30 μmol g<sup>-1</sup> for RNA) were purchased from Sigma Aldrich. Native amino CPG support (1000 Å, 74 μmol g<sup>-1</sup>) was purchased from ChemGenes. 3'-Phosphate CPG support was prepared as described below (see sections 4.2.4.1–4.2.4.5).

### Molecular sieves 3 Å

A batch of molecular sieves was thoroughly rinsed with water to remove any dust particles and then pre-dried at 120 °C for several days. The molecular sieves were activated by heating under reduced pressure for at least 1 d.

### MeCN

MeCN (DNA synthesis grade) was dried over freshly activated molecular sieves (3 Å) for at least 2 d prior to use.

### Phosphoramidites

Phosphoramidites were employed as 70 or 100 mM solutions in anhydrous MeCN. CH<sub>2</sub>Cl<sub>2</sub> was used as the solvent for the r4Cl phosphoramidite **133** instead due to its limited solubility in MeCN. The resulting solutions were dried over freshly activated molecular sieves (3 Å, few grains per vial) for 16 h at 4 °C prior to use.

### Detritylation ABI

Dichloroacetic acid (3 mL) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (97 mL, peptide synthesis grade).

### Detritylation Pharmacia

Dichloroacetic acid (3 mL) was dissolved in DCE (97 mL).

### Activator

ETT (0.82 g) was dissolved in MeCN (25 mL) to make a 0.25 M solution (two batches). The mixture was centrifuged at 4000 rpm for 10 min and the supernatant was transferred to a glass bottle. The reagent was dried over freshly activated molecular sieves (3 Å) for 1 d prior to use.

### Cap A ABI

*N*-Methylimidazole (10 g) was dissolved in MeCN (50 mL)

### Cap A Pharmacia

DMAP (3.0 g) was dissolved in MeCN (50 mL).

### Cap B ABI/Pharmacia

*sym*-Collidine (15 mL) and Ac<sub>2</sub>O (10 mL) were dissolved in MeCN (25 mL).

### Oxidizer ABI

50 mM I<sub>2</sub> in pyridine/H<sub>2</sub>O (commercially available)

### Oxidizer Pharmacia

I<sub>2</sub> (0.40 g) was dissolved in MeCN (100 mL). Afterwards, H<sub>2</sub>O (46 mL) and *sym*-collidine (9.2 mL) were added.

## **4.1.3.3 General protocol**

### Method A, ABI

The CPG support corresponding to the 3' end of the respective oligonucleotide (600 nmol, 20–30 mg) was packed into a column and sealed off with two filter discs and the required reagents

were prepared as described above. The synthesizer was set up according to standard protocols. All oligonucleotides were synthesized with final detritylation using the following preprogrammed routines:

DNA: "0.6uM DNA lv1cap Test Jan", 2 min coupling time

RNA: "0.6uM RNA Test Jan", 4 min coupling time

Stepwise coupling yields were assessed from the absorbance of the trityl cation released during the detritylation step. After the synthesis had finished, the solid support was subjected to workup without removal from the column.

#### Method B, Pharmacia

The CPG support corresponding the 3' end of the respective oligonucleotide (600 nmol, 20–30 mg) was packed into a column and sealed off with two filter discs and the required reagents were prepared as described above. The synthesizer was set up according to standard protocols. All oligonucleotides were synthesized with final detritylation using the following preprogrammed routines:

Unmodified phosphoramidites: "07CH-N4", 4 min coupling time

Modified phosphoramidites: "07CH-X12", 12 min coupling time

Stepwise coupling yields were assessed from the absorbance of the trityl cation released during the detritylation step. After the synthesis had finished, the solid support was dried under reduced pressure and transferred into a screw cap vial for further workup.

#### Workup and purification

Cleavage from the solid support and removal of base-labile protecting groups was achieved by treatment with 25% NH<sub>3</sub> at 25 °C for 30 min followed by 25% NH<sub>3</sub>/aq. MeNH<sub>2</sub> 1:1 at 25 °C for 5 h.

2'-O-Silyl-deprotection of the RNA was performed by treatment with tetrabutylammonium fluoride (TBAF, 1 M in THF) at 25 °C for 16 h followed by addition of Tris-HCl (1 M, pH 8.0) and desalting.

RNA was desalted using an ÄKTA start with three HiTRAP Desalting columns (5 mL each) from GE Healthcare (eluent: H<sub>2</sub>O; flow rate: 2 mL min<sup>-1</sup>; temperature: 25 °C).

The crude solid-phase synthesis products were purified by denaturing PAGE (15–20%) and analyzed by HPLC and ESI MS.

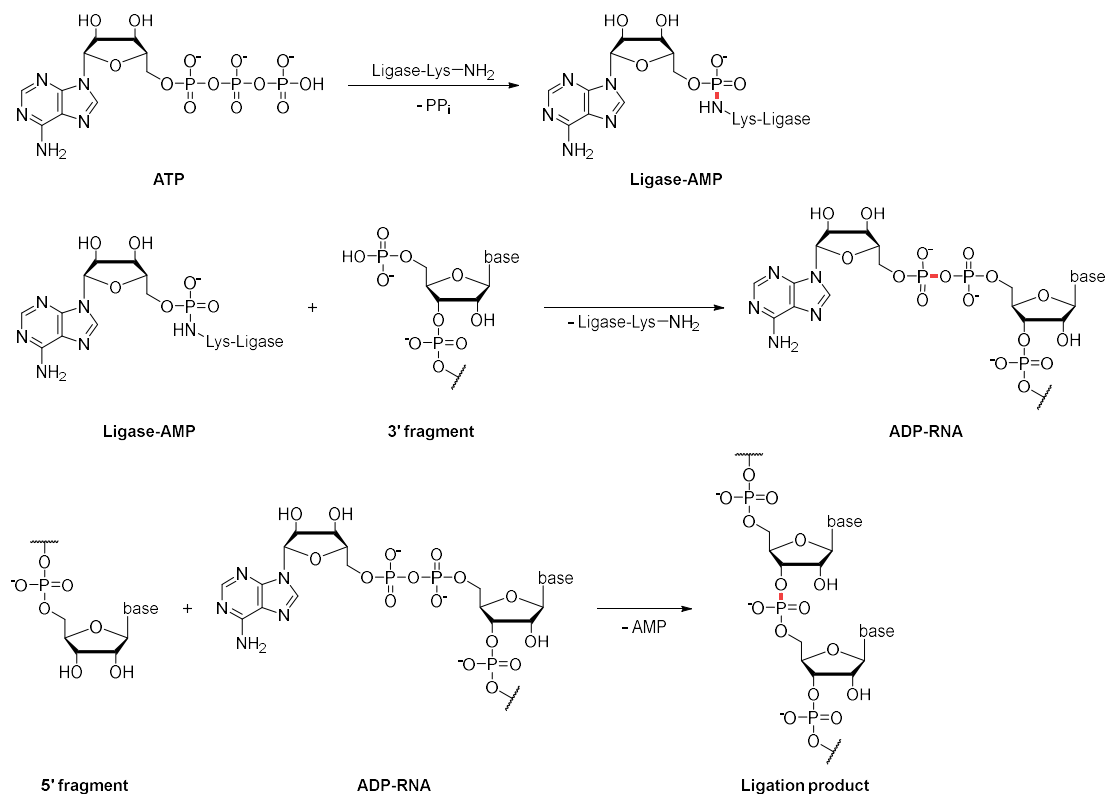
Typical yields were 60–140 nmol DNA/RNA from 600 nmol of CPG support as determined by UV absorbance.

### 4.1.4 Enzymatic ligation

#### **4.1.4.1 Overview**

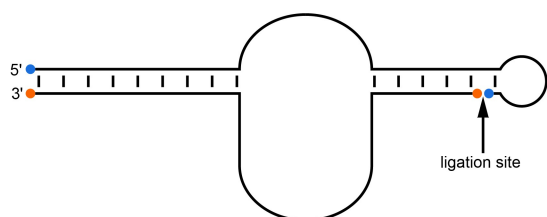
While in vitro transcription is routinely used for the synthesis of long unmodified RNAs, its use for the incorporation of modified nucleotides is limited to cases in which the corresponding triphosphates can be processed by an appropriate RNA polymerase. Additionally, site-specific incorporation is only possible under specific circumstances. Solid-phase synthesis addresses these issues, but is limited by the length of the target oligonucleotide, around 50–100 nt using standard methods. For increased efficiency, 4-cyanoindole-modified variants of the Chili aptamer were

therefore prepared by enzymatic ligation of shorter fragments with T4 RNA ligase 1. This enzyme catalyzes the ATP-dependent formation of a new phosphate linkage between the 5'-phosphorylated terminus of a donor RNA and the 3'-hydroxy group of a corresponding acceptor RNA (Scheme 41).<sup>265,266</sup> The ligation site was chosen such that the two 26 nt fragments can easily anneal without requiring an additional splint oligonucleotide (Figure 75).<sup>267</sup>



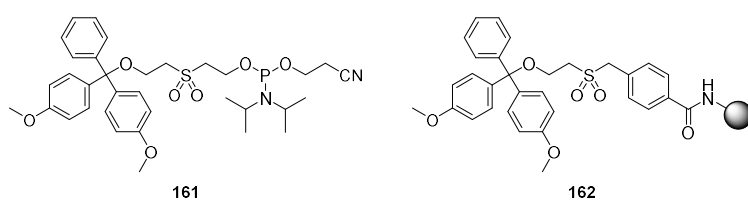
**Scheme 41.** The ligation of two RNA fragments by T4 RNA ligase 1 proceeds via three discrete steps. After adenylation of an amino group in the active site, a phosphodiester is formed with the 3'-end of the donor strand. Nucleophilic attack by the 5'-hydroxy group of the acceptor strand is accompanied by the release of ATP and results in the formation of a new phosphate linkage.<sup>266</sup>

Donor fragments were synthesized bearing a 3'-phosphate group in order to prevent side reactions such as circularization or oligomerization.



**Figure 75.** Two 26 nt fragments were ligated to generate full-length modified Chili aptamer constructs. The ligation site is located within the hairpin to ensure proper annealing without addition of a DNA splint. Blue: Hydroxy groups, orange: phosphate groups.

Both 5'- and the 3'-phosphate groups were installed during solid-phase synthesis using a chemical phosphorylation amidite (**161**) or a functionalized CPG support (**162**), respectively (Chart 30)



**Chart 30.** Chemical phosphorylation amidite (**161**) and functionalized CPG support (**162**) for the installation of 5'- and 3'-phosphate groups, respectively.

Ligation reactions were performed overnight at 37 °C using an excess of the unmodified RNA strand. Full-length products were isolated and purified by PAGE (see section 4.1.5). The quality and identity of the ligation products was assessed by HPLC and HR-MS.

Enzymatic ligation was also used to attach Chili to the 5'-end of a 3'-Atto 590-labeled model RNA substrate for cleavage by a 10-23 DNAzyme. In this case, it was necessary to use a 24 nt DNA splint oligonucleotide which could hybridize to both the aptamer and the substrate in order to furnish a ligation site.<sup>111</sup>

#### 4.1.4.2 Materials

T4 RNA ligase 1 and the corresponding 10x reaction buffer were purchased from New England Biolabs.

##### 10x reaction buffer

TRIS (500 mM, pH 7.5), MgCl<sub>2</sub> (100 mM), DTT (10 mM)

#### 4.1.4.3 General protocol

Reactions were typically performed in a volume of 50 µL.

The respective RNA fragments (100 µM 4Cl-modified RNA, 110 µM unmodified RNA) were dissolved in H<sub>2</sub>O (35 µL) and the RNA was annealed by heating to 95 °C for 3 min. After cooling to ambient temperature, 10x reaction buffer, 10 mM ATP (5 µL) and T4 RNA ligase 1 (50 units) were added and the reaction mixture was incubated at 37 °C for 12 h. Next, it was diluted with loading buffer (50 µL) and purified by denaturing PAGE (10%). The isolated product was analyzed by HPLC and ESI MS.

Typical yields were 2–3 nmol RNA from 50 µL ligation reactions as determined by UV absorbance.

#### 4.1.5 Polyacrylamide gel electrophoresis (PAGE)

Oligonucleotides were separated and purified according to their charge state by means of denaturing polyacrylamide gel electrophoresis (PAGE). Gels were prepared from acrylamide stock solutions containing 5% bis-acrylamide as a crosslinker and 7 M urea as a denaturing agent. The concentration of acrylamide was chosen based on the length of the respective oligonucleotide:

Oligonucleotide length	Acrylamide concentration
20 nt	20%
20–50 nt	15%
> 50 nt	10%

Polymerization was initiated by the addition of 25% ammonium peroxodisulfate (APS) and tetramethylethylenediamine (TMEDA or TEMED) in the following proportions:

Component	Amount preparative	Amount analytical
gel stock	45 mL	10 mL
APS	125 µL	25 µL
TMEDA	30 µL	10 µL

Preparative gels (0.7×200×300 mm) with four sample pockets were cured for 30–40 min at ambient temperature and pre-run for 30 min at a constant power of 35 W in 1x TBE. Next, the pockets were thoroughly rinsed and the samples (up to 200 µL containing 50 µL loading buffer per

pocket) were added. The actual run time of the samples at 35 W was between 2 h and 2.5 h depending on the composition of the gel and the length of the oligonucleotide.

Bands were visualized and documented by UV shadowing on a TLC plate, cut out of the gel, crushed and eluted by soaking in 1x TEN buffer (300  $\mu$ L per band). Two rounds of elution (1.5 h at 37 °C and overnight at 4 °C) were performed. In between rounds, the buffer was removed and stored at 4 °C.

To precipitate the oligonucleotides, the buffer was diluted 1:3 with abs. EtOH, frozen in liquid nitrogen and centrifuged for 30 min at 15000 rpm and 4 °C. The supernatants were carefully removed, the remaining pellet was covered with 70% EtOH (75  $\mu$ L) and centrifuged as before. After removal of the supernatants, the pellets were dried under reduced pressure and dissolved in H<sub>2</sub>O (150  $\mu$ L).

Sample concentration and quality were assayed by UV absorbance (see section 4.1.6.3) and HPLC analysis (see section 4.1.6.1), respectively. Additionally, the identity of solid-phase synthesis and ligation products was confirmed by mass spectrometry (see section 4.1.6.2)

For specific experiments, analytical gels (1.0×80×100 mm) with ten sample pockets were used. These experiments were performed under native conditions, i.e. no urea was present in the gel stock. Gels were cured for 30 min at ambient temperature. Samples (5  $\mu$ L + 1  $\mu$ L loading buffer per pocket) were directly added to the thoroughly rinsed pockets without any additional pre-run time. These gels were run in 1x TBE for 20 min at a constant voltage of 200 V.

Bands were visualized and imaged by staining with appropriate dyes for 15 min as indicated in the respective figures.

#### 4.1.6 Oligonucleotide analysis

##### 4.1.6.1 HPLC

Anion-exchange HPLC analyses were performed using an ÄKTApurifier from GE Healthcare with a DNAPac PA200 column (2×250 mm) from Thermo Scientific (eluent: Dionex buffers A and B; gradient: 0–48% B over 12 CV; flow rate 0.5 mL min<sup>-1</sup>; temperature: 60–80 °C).

Typical sample concentrations were in the range of 50–200 pmol oligonucleotide in 15  $\mu$ L of H<sub>2</sub>O.

##### 4.1.6.2 Mass spectrometry

High-resolution ESI mass spectra in negative ion mode were acquired on a Bruker micrOTOF-Q III instrument.

Monoisotopic masses of oligonucleotides were obtained by charge deconvolution of the raw spectra.

For DNA and RNA samples, a concentration of 200 pmol oligonucleotide in 8  $\mu$ L of H<sub>2</sub>O was used. Samples were diluted with hexafluoroisopropanol (HFIP) in MeCN and directly injected into the ESI source.

##### 4.1.6.3 Quantification

An Implen NanoPhotometer P 360 was used for DNA and RNA quantification. Concentrations of DNA and RNA samples were determined from their UV absorbance at 260 nm based on tabulated extinction coefficients of the individual nucleosides:<sup>268,269</sup>



Base	$\epsilon_{260 \text{ nm}}$ $\text{L mol}^{-1} \text{ cm}^{-1}$
A	15400
C	7300
G	11700
T	8800
U	9900

Any nearest-neighbor interactions were neglected. In case of 4-cyanoindole-containing oligonucleotides, the modified base was omitted from the calculation due to its low extinction coefficient at 260 nm.

#### 4.1.7 UV/Vis spectroscopy

##### 4.1.7.1 Materials and instrumentation

UV/Vis data were collected with a JASCO V-770 spectrophotometer equipped with a PAC-743 cell changer (regular absorption spectra). Alternatively, a VARIAN CARY 100 Bio spectrophotometer equipped with a 6x6 Multicell Block Peltier Series II cell changer and a VARIAN CARY Temperature Controller (thermal denaturation curves) was used.

Samples were measured in VARIAN semi-micro quartz cuvettes or in Roth semi-micro polystyrene cuvettes (10 mm path length) unless noted otherwise for specific methods. Spectral bandwidths were adjusted according to the manufacturer's instructions for each cuvette type. Scan speed and integration time were balanced to record a stable signal within a reasonable timeframe.

All measurements were conducted at 25 °C unless noted otherwise.

##### 4.1.7.2 Data analysis

All spectra were automatically subjected to a blank correction obtained from pure solvent or buffer in the same type of cuvette.

If necessary, residual constant baseline offsets, which occurred occasionally when using cuvettes with small window sizes in the JASCO spectrophotometer, were removed by shifting the entire spectrum up- or downward.

#### 4.1.8 $pK_a$ determination

##### 4.1.8.1 Overview

For a weak acid HA in equilibrium with its corresponding base  $A^-$ , the concentrations of both species are related to the pH of the system by the Henderson-Hasselbalch equation:

$$\text{HA} \xrightleftharpoons{K_a} \text{H}^+ + \text{A}^-$$

$$\text{pH} = \text{p}K_a - \log \frac{[\text{HA}]}{[\text{A}^-]} \quad (1)$$

The acid dissociation constant can be determined by measuring any property that varies proportionally with the composition of the system. In spectrophotometric titrations, this is the absorbance at a given wavelength ( $A_\lambda$ ), which comprises the absorbance values of the reaction partners weighted by their respective mole fractions  $\chi$ :

$$A_\lambda = A_{\text{HA}}\chi_{\text{HA}} + A_{\text{A}^-}\chi_{\text{A}^-} \quad (2)$$

One of the mole fractions can be eliminated from this equation:

$$\chi_{\text{HA}} = \frac{[\text{HA}]}{[\text{HA}] + [\text{A}^-]} \quad \chi_{\text{A}^-} = \frac{[\text{A}^-]}{[\text{HA}] + [\text{A}^-]} \quad (3)$$

$$\begin{aligned} A_{\lambda} &= A_{\text{HA}}\chi_{\text{HA}} + A_{\text{A}^-}(1 - \chi_{\text{HA}}) \\ &= A_{\text{HA}}\chi_{\text{HA}} + A_{\text{A}^-} - A_{\text{A}^-}\chi_{\text{HA}} \\ &= (A_{\text{HA}} - A_{\text{A}^-})\chi_{\text{HA}} + A_{\text{A}^-} \end{aligned} \quad (4)$$

After rearrangement, this yields

$$\chi_{\text{HA}} = \frac{A_{\lambda} - A_{\text{A}^-}}{A_{\text{HA}} - A_{\text{A}^-}} \quad (5)$$

An analogous expression is obtained for the deprotonated species:

$$\chi_{\text{A}^-} = \frac{A_{\text{HA}} - A_{\lambda}}{A_{\text{HA}} - A_{\text{A}^-}} \quad (6)$$

Using equation 3, these mole fractions are inserted into the Henderson-Hasselbalch equation:

$$\text{pH} = \text{p}K_{\text{a}} - \log \frac{\chi_{\text{HA}}}{\chi_{\text{A}^-}} = \text{p}K_{\text{a}} - \log \frac{(A_{\lambda} - A_{\text{A}^-})(A_{\text{HA}} - A_{\text{A}^-})}{(A_{\text{HA}} - A_{\text{A}^-})(A_{\text{HA}} - A_{\lambda})} = \text{p}K_{\text{a}} - \log \frac{A_{\lambda} - A_{\text{A}^-}}{A_{\text{HA}} - A_{\lambda}} \quad (7)$$

This expression is recast to isolate the experimentally measured  $A_{\lambda}$ :

$$10^{\text{p}K_{\text{a}} - \text{pH}} = \frac{A_{\lambda} - A_{\text{A}^-}}{A_{\text{HA}} - A_{\lambda}} \quad (8)$$

$$1 + 10^{\text{p}K_{\text{a}} - \text{pH}} = \frac{A_{\text{HA}} - A_{\lambda} + A_{\lambda} - A_{\text{A}^-}}{A_{\text{HA}} - A_{\lambda}} = \frac{A_{\text{HA}} - A_{\text{A}^-}}{A_{\text{HA}} - A_{\lambda}} \quad (9)$$

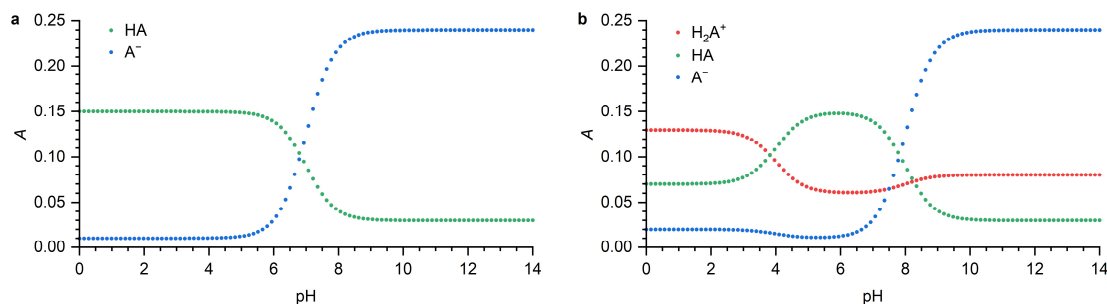
$$A_{\text{HA}} - A_{\lambda} = \frac{A_{\text{HA}} - A_{\text{A}^-}}{1 + 10^{\text{p}K_{\text{a}} - \text{pH}}} \quad (10)$$

$$A_{\lambda} = A_{\text{HA}} + \frac{A_{\text{A}^-} - A_{\text{HA}}}{1 + 10^{\text{p}K_{\text{a}} - \text{pH}}} \quad (11)$$

Titration curves described by equation 11 are simulated in Figure 76a. If multiple protonation steps occur over the course of the titration (Figure 76b), this equation can simply be expanded:

$$A = a + \frac{b - a}{1 + 10^{\text{p}K_{\text{a}1} - \text{pH}}} + \frac{c - b}{1 + 10^{\text{p}K_{\text{a}2} - \text{pH}}} \quad (12)$$

The individual absorbance parameters are not physically meaningful when overlapping bands occur in the spectra and have been replaced by the abstract parameters  $a$ ,  $b$  and  $c$ .



**Figure 76.** Simulated titration curves for a) a single protonation step with  $\text{p}K_{\text{a}} = 7$  and b) two consecutive protonation steps with  $\text{p}K_{\text{a}1} = 4$  and  $\text{p}K_{\text{a}2} = 8$ . Data points were calculated with equation 12.

#### 4.1.8.2 Materials

HBI derivatives were used as stock solutions (1 mM) in DMSO.

#### 4.1.8.3 General protocol

The following samples were prepared

Ligand pH 1–12:		
	Added volume	Final concentration
Britton-Robinson buffer	680 $\mu\text{L}$	
1 mM ligand in DMSO	10 $\mu\text{L}$	11 $\mu\text{M}$
H <sub>2</sub> O	210 $\mu\text{L}$	
Final volume	900 $\mu\text{L}$	

Ligand pH 0.4		
	Added volume	Final concentration
1 M HCl	340	0.38 M
1 mM ligand in DMSO	10 $\mu\text{L}$	11 $\mu\text{M}$
H <sub>2</sub> O	550	
Final volume	900 $\mu\text{L}$	

Ligand pH 0.1		
	Added volume	Final concentration
1 M HCl	680	0.76 M
1 mM ligand in DMSO	10 $\mu\text{L}$	11 $\mu\text{M}$
H <sub>2</sub> O	210	
Final volume	900 $\mu\text{L}$	

UV/Vis spectra of each sample were measured using the following parameter settings:

- Spectral range: 300–750 nm
- Spectral bandwidth: M2.0 nm
- Response: 0.06 s
- Data interval: 0.5 nm
- Scan speed: 1000 nm/min

Blank correction data was obtained from each sample prior to addition of the ligand.

The absorbance  $A$  at the absorption maxima of each observed species was fitted to a suitable form of equation 12 (page 122) accounting for up to three protonation steps.

#### 4.1.9 Extinction coefficient determination

##### 4.1.9.1 Materials

HBI derivatives and 4-cyanoindole ribonucleoside were used as stock solutions (HBI: 2 mM, r4Cl: 4 mM) in DMSO.

#### 4.1.9.2 General protocol

A stock solution of the analyte was serially diluted nine times with DMSO in ratio of 65:35 to prepare a concentration series. Aliquots of each dilution step were then mixed with buffer as follows:

HBI samples:		
	Added volume	Final concentration
Diluted analyte	10 $\mu\text{L}$	22–0.5 $\mu\text{M}$
Britton-Robinson buffer	680 $\mu\text{L}$	
H <sub>2</sub> O	210 $\mu\text{L}$	
Final volume	900 $\mu\text{L}$	

r4Cl samples:		
	Added volume	Final concentration
Diluted analyte	10 $\mu\text{L}$	44–1 $\mu\text{M}$
5x PBS pH 7.4	180 $\mu\text{L}$	
H <sub>2</sub> O	710 $\mu\text{L}$	
Final volume	900 $\mu\text{L}$	

For HBI derivatives, the pH of the Britton-Robinson buffer was chosen such as to ensure the presence of only the phenolic or phenolate species in a given experiment.

UV/Vis spectra of each sample were measured using the following parameter settings:

- Spectral range: 300–750 nm (HBI) or 200–750 nm (r4Cl)
- Spectral bandwidth: M2.0 nm
- Response: 0.06 s
- Data interval: 0.5 nm
- Scan speed: 1000 nm/min

Blank correction data was obtained from each sample prior to addition of the analyte.

Extinction coefficients  $\epsilon$  were calculated by fitting plots of the absorbance  $A$  at the absorption maximum vs analyte concentration to the Beer-Lambert law:

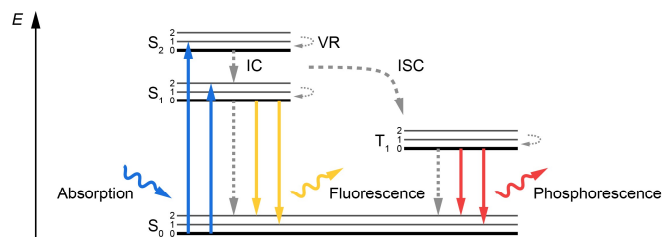
$$A = \epsilon cd \quad (13)$$

The path length  $d$  was set to 10 mm.

#### 4.1.10 Fluorescence spectroscopy

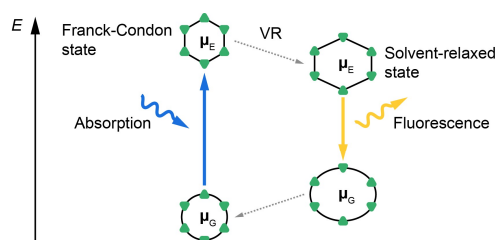
##### 4.1.10.1 Overview

The Jablonski diagram (Figure 77) summarizes the pathways responsible for the excitation and deactivation of a fluorophore.<sup>241,270</sup> Initially in its electronic ground state ( $S_0$ ), absorption of a photon with suitable energy leads to a transition into the vibrational levels of a higher-lying excited state ( $S_n$ ). The timescale of absorption is on the order of  $10^{-15}$  s, significantly faster than the following processes and also near instantaneous with respect to any displacement of the atomic nuclei (Franck-Condon principle).



**Figure 77.** Simplified Jablonski diagram for transitions between the electronic ground state ( $S_0$ ), the first two excited singlet states ( $S_1$ ,  $S_2$ ) and a triplet state ( $T_1$ ) for an organic fluorophore. For each state, three vibrational energy levels are depicted (energies not to scale). Photon absorption populates excited states that partially dissipate energy into their surroundings (vibrational relaxation, VR). Overlapping vibrational levels enable non-radiative transitions to lower-energy states of the same or different multiplicity by internal conversion (IC) or intersystem crossing (ISC) respectively. Fluorescence typically occurs from the lowest vibrational level of  $S_1$  due to its long lifetime with respect to other processes. Emission of a phosphorescence photon from  $T_1$  is slow due to the required change in multiplicity when returning to  $S_0$ .<sup>241,270</sup>

Excess energy vibrational is dissipated within  $10^{-12}$ – $10^{-10}$  s by redistribution between vibrational modes of the same electronic state and by collisions with surrounding molecules in a process termed vibrational relaxation (VR). Coupling between the vibrational energy levels of different electronic states with the same multiplicity leads to radiationless internal conversion (IC) on a timescale of  $10^{-11}$ – $10^{-9}$  s. By the same mechanism, intersystem crossing (ISC) populates an excited state of different multiplicity, albeit slower ( $10^{-10}$ – $10^{-8}$  s). Because the energy difference between  $S_0$  and  $S_1$  is typically significantly greater than that between higher excited states,  $S_1$  acts as a temporary population sink with a lifetime of  $10^{-10}$ – $10^{-7}$  s.



**Figure 78.** Initially, the fluorophore is in equilibrium with the surrounding solvent (green) at its ground state geometry. Excitation leads to a virtually instantaneous change of the electronic structure, altering e.g. the dipole moment ( $\mu$ ), but leaving the geometry unaffected (Franck-Condon principle). As the excited fluorophore reaches a new equilibrium with its surroundings, the energy of the system decreases. After emission of a fluorescence photon, the excited state geometry relaxes back to the initial state.<sup>241</sup>

During this time, the solvent surrounding the fluorophore relaxes to accommodate its new electronic structure, which is different from that in the  $S_0$  state. The concomitant stabilization of the fluorophore, in combination with energy losses from VR and IC, finally results in emission of a fluorescence photon with a lower energy than that of the initially absorbed one (Figure 78). This energy difference is known as the Stokes shift. Due to the speed of non-radiative deactivation processes relative to the lifetime of  $S_1$ , this state is usually the origin of fluorescence emission even if a higher excited state had been populated initially (Kasha's rule). Therefore, the emission wavelength is independent of the excitation wavelength.

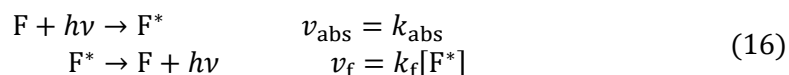
Under constant illumination ( $k_{\text{abs}}$ ), the excited state population ( $F^*$ ) reaches a steady state described by the following differential equation:

$$\frac{d[F^*]}{dt} = k_{\text{abs}} - (k_f + k_{\text{IC}} + k_{\text{ISC}})[F^*] = k_{\text{abs}} - (k_f + k_{\text{nr}})[F^*] = 0 \quad (14)$$

The rate constants for internal conversion ( $k_{ic}$ ), intersystem crossing ( $k_{isc}$ ) and other conceivable non-radiative deactivation pathways have been combined into  $k_{nr}$  on the right-hand side. From the steady-state condition, it follows that

$$k_{abs} = (k_f + k_{nr})[F^*] \quad (15)$$

For any photoinduced process, a primary quantum yield  $\phi$  can be defined that relates the number of events to the number of photons absorbed. Division by the time interval of the process converts the numbers to rates ( $\nu$ ).



Therefore, the quantum yield of fluorescence is given by

$$\phi_f = \frac{\nu_f}{\nu_{abs}} = \frac{k_f[F^*]}{(k_f + k_{nr})[F^*]} = \frac{k_f}{k_f + k_{nr}} \quad (17)$$

Experimentally, the observed quantity is fluorescence intensity  $I$ , which is proportional to the quantum yield by instrument-dependent parameters. The quantum yield can be measured directly by employing an appropriate set of spectral corrections (absolute quantum yield) or by comparison with a suitable reference substance (relative quantum yield). Both methods are described in section 4.1.12.

#### 4.1.10.2 Materials and instrumentation

Steady-state fluorescence data were collected with a JASCO FP-8300 spectrofluorometer equipped with an FCT-817S cell changer (regular excitation and emission spectra, association kinetics) or with an ILF-835 integrating sphere (absolute quantum yields). Alternative, a VARIAN CARY Eclipse spectrofluorometer equipped with either a Peltier Multicell Holder cell changer and a VARIAN Cary Temperature Controller (thermal denaturation curves) or an Agilent Microplate Reader Accessory (plate-based screening assays) was used.

Samples were measured in Hellma ultra-micro quartz cuvettes (10×2 mm, 3×3 mm or 1.5×1.5 mm) unless noted otherwise for specific methods. Excitation and emission monochromator wavelength were chosen as appropriate for the respective measurement. Spectral bandwidths and photomultiplier tube (PMT) voltage were adjusted to achieve the highest possible signal intensity without compromising the spectral resolution or introducing excessive noise. Scan speed and integration time were balanced to achieve a stable signal while limiting exposure of the sample to the excitation light source.

All measurements were conducted at 25 °C unless noted otherwise.

#### 4.1.10.3 Data analysis

Where indicated, excitation and emission spectra were manually corrected for background fluorescence and Raman scattering from the solvent by subtracting blank spectra obtained from pure solvent, buffer or an appropriate sample.

Spectral intensities were either reported at the indicated wavelengths or also, for emission spectra, as integrated peak areas. The latter was preferred when changes in the peak emission wavelength could potentially affect quantification results due to wavelength-dependent differences in the detector sensitivity or when a high dynamic range between samples resulted in spectra with markedly different signal-to-noise ratios.

#### 4.1.11 Fluorescence lifetime analysis

##### 4.1.11.1 Overview

After irradiation of a fluorescent sample by a short pulse of light, idealized as a  $\delta$  function, an excited state population  $F^*$  is generated. As discussed in section 4.1.10.1, this population decays by a combination of radiative and non-radiative processes with the respective rate constants  $k_f$  and  $k_{nr}$  according to a first order rate law:

$$\frac{d[F^*]}{dt} = -(k_f + k_{nr})[F^*] \quad (18)$$

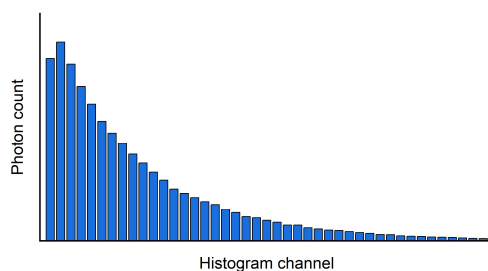
Integration of the differential equation leads to a simple exponential decay. Considering that the experimentally observed fluorescence intensity is proportional to the fluorophore concentration,

$$I(t) = I_0 \exp(-(k_f + k_{nr})t) = I_0 \exp\left(-\frac{t}{\tau}\right) \quad (19)$$

describes the attenuation of emission intensity immediately after excitation ( $I_0$ ) with a characteristic lifetime  $\tau$ , after which 63% percent of the initial population have returned to the ground state. It follows that  $\tau$  is related to the quantum yield of fluorescence ( $\phi$ ) by

$$\tau = \frac{1}{k_f + k_{nr}} = \frac{\phi}{k_f} \quad (20)$$

There are two approaches to measuring fluorescence lifetimes, the time domain method using pulsed excitation and the frequency domain method relying on intensity-modulated continuous excitation.<sup>241</sup> In this thesis, time domain measurements were performed using a time-correlated single photon counting (TCSPC) setup. During a TCSPC measurement, the sample is periodically irradiated with a pulsed light source. Photon detection events are limited to below 2% of the excitation rate to ensure the detection of individual photons, which are then binned into discrete channels according to their arrival time at the detector (Figure 79).



**Figure 79.** During a TCSPC experiment, photon arrival times are sorted into discrete time channels according to their arrival time. The photon count per channel follows a Poisson distribution. For the data presented in this thesis, a channel width of  $2.57 \times 10^{-11}$  s was used.

The excitation pulse, also called instrument response function (IRF), has a finite width, resulting in excitation of fluorophores over an extended period of time. Therefore, the experimentally observed photon count  $N(t)$  is a convolution of the actual fluorescence decay  $I(t)$  and the instrument response function  $R(t)$ :

$$N(t) = (I * R)(t) = \int_0^t I(t')R(t - t')dt' \quad (21)$$

To recover  $I(t)$ , a model function  $M(t)$  is set up, its convolution integral with the separately measured  $R(t)$  is calculated and fitted to  $N(t)$  by a nonlinear least squares algorithm. Convolution and fitting are iterated until an optimized set of parameters is obtained.<sup>271</sup> In this work, the model function was a sum of exponential decays:

$$M(t) = \sum_n \alpha_n \exp\left(-\frac{t}{\tau_n}\right) \quad \sum_n \alpha_n = 1 \quad (22)$$

The nonlinear least squares algorithm minimizes the goodness-of-fit parameter  $\chi^2$ , which is defined as follows:

$$\chi^2 = \sum_{i=1}^{n_c} \left( \frac{N_{\text{obs}}(i) - N_{\text{fit}}(i)}{\sigma(i)} \right)^2 = \sum_{i=1}^{n_c} \left( \frac{N_{\text{obs}}(i) - N_{\text{fit}}(i)}{\sqrt{N_{\text{obs}}(i)}} \right)^2 \quad (23)$$

Here,  $n_c$  is the number of channels of the photon histogram and  $\sigma(i)$  is the standard deviation in the  $i^{\text{th}}$  channel. As photon counting obeys Poisson statistics, the standard deviation is simply the square root of the counts in that channel. Usually, it is not  $\chi^2$  itself, which is optimized, but its reduced form  $\chi_R^2$ , which is corrected for the number of fit parameters ( $n_p$ ) in the model:

$$\chi_R^2 = \frac{\chi^2}{n_c - n_p - 1} \quad (24)$$

A fitted model that fully describes the observed decay trace would result in a  $\chi_R^2$  of 1. In practice, up to four exponential terms were included in the model until  $\chi_R^2$  was close to 1 and did not improve significantly anymore. The fit quality was assessed by examining the properly weighted residuals  $r_{\text{fit}}(i)$ , which should be randomly distributed about 0:

$$r_{\text{fit}}(i) = \frac{N_{\text{obs}}(i) - N_{\text{fit}}(i)}{\sqrt{N_{\text{obs}}(i)}} \quad (25)$$

Additionally, the autocorrelation ( $a$ ) of the residuals was inspected for systematic deviations.<sup>272</sup> It was conveniently calculated by taking the Fourier transform ( $\mathcal{F}$ ) of the normalized residuals, squaring the result (\* denotes complex conjugation) and applying an inverse Fourier transform:

$$\begin{aligned} r_{\text{norm}} &= \frac{r_{\text{fit}}(i)}{\sum r_{\text{fit}}^2(i)} \\ R &= \mathcal{F}(r_{\text{norm}}) \\ a(l) &= \mathcal{F}^{-1}(RR^*) \end{aligned} \quad (26)$$

For multiexponential models, an average lifetime was calculated. There are two concurrent definitions employing different weighting schemes of the individual amplitudes and lifetimes, an intensity average lifetime given by

$$\langle \tau \rangle_f = \frac{\sum_n \alpha_n \tau_n^2}{\sum_n \alpha_n \tau_n} \quad (27)$$

and the amplitude average lifetime

$$\langle \tau \rangle_a = \frac{\sum_n \alpha_n \tau_n}{\sum_n \alpha_n} = \sum_n \alpha_n \tau_n \quad (28)$$

The latter definition was found to be more appropriate in the majority of analyses and is used throughout this work.<sup>273</sup>

#### 4.1.11.2 Materials and instrumentation

Fluorescence lifetime data was collected using a Horiba DeltaFlex system with a DeltaDiode excitation source (DD-320, DD-405L or DD-485L) and polarization filters. The excitation polarizer



was oriented in the vertical position and the emission polarizer was oriented at 54.7° from the vertical (magic angle conditions) to eliminate anisotropic effects.

Samples were measured in Hellma ultra-micro quartz cuvettes (10×2 mm, 3×3 mm or 1.5×1.5 mm). The excitation source and emission monochromator wavelength were generally chosen to be as close as possible to the steady-state excitation and emission maxima of the fluorophore. A stop rate of close to but not more than 2% was achieved by suitable adjustment of the Em bandwidth. For typical experiments, the excitation pulse repetition rate was set to 10 MHz, which corresponds to a measurement interval of 105 ns and is sufficient for samples with decay time of up to 20 ns.<sup>241</sup> The photon arrival time was binned into 4096 channels, resulting in a width of 0.026 ns/channel. Data was collected until 10000 counts had accumulated in the peak channel.

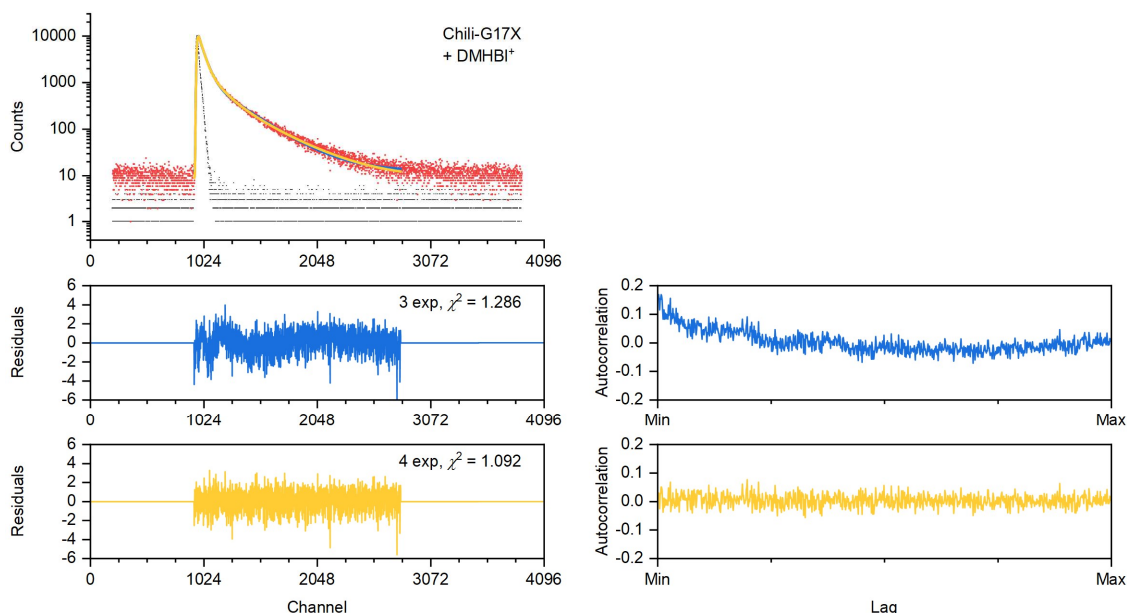
The instrument response function was recorded from a water or buffer sample by setting the emission monochromator to the actual peak wavelength of the excitation source. All other parameters were left unchanged. If the stop rate exceeded 2%, the excitation intensity was attenuated using neutral density filters.

All measurements were conducted at 25 °C.

#### 4.1.11.3 Data analysis

Fluorescence decay curves were analyzed by iterative reconvolution fitting with an appropriate instrument response function using a proprietary algorithm implemented in Horiba DAS6 (version 6.8). Fitting was restricted to portions of the data with photon counts above the noise level of the respective curve.

Starting from a monoexponential model, up to three additional exponential terms were included until the fit statistics ( $\chi^2$ ) did not show significant changes; generally,  $\chi^2 < 1.2$  was considered sufficient if the autocorrelation of the properly weighted fit residuals did not show systematic variations. An example is shown in Figure 80.

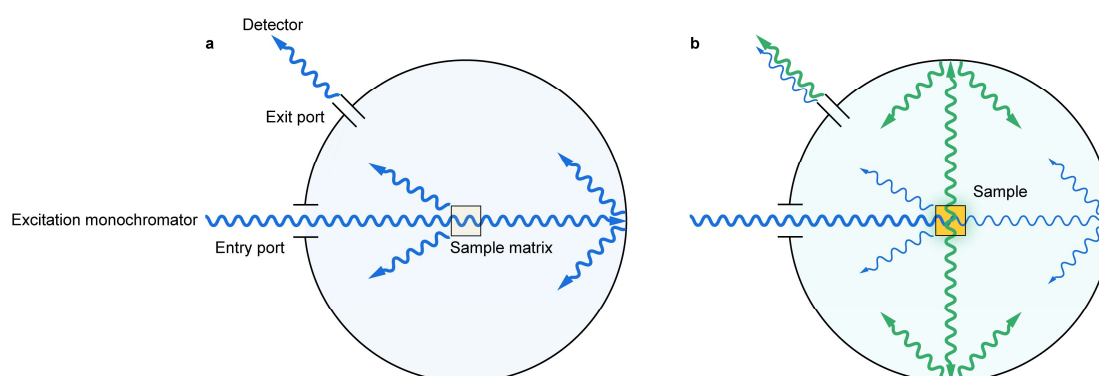


**Figure 80.** Comparison of multiexponential models for analyzing the fluorescence decay of Chili-G17X–DMHBI<sup>+</sup>. Top panel: Fits with three (blue) and four (yellow) exponential terms are visually indistinguishable. Middle panels: The Properly weighted residuals and their autocorrelation clearly show systematic deviations. A sum of three exponentials is not appropriate to describe the observed decay. Bottom panels: Addition of a fourth exponential term gives a sufficient, albeit complex, model.

## 4.1.12 Quantum yield determination

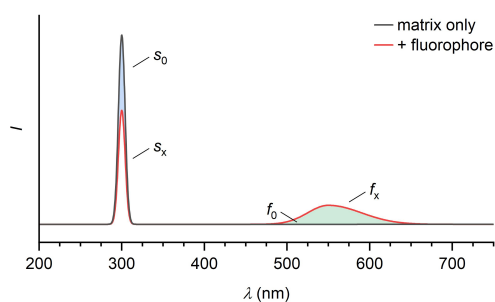
### 4.1.12.1 Overview

In addition to extinction coefficient  $\epsilon$  and excited state lifetime  $\tau$ , the quantum yield  $\phi$  is the third and final important quantity necessary for a comprehensive description of a fluorophore. Recalling the discussion in section 4.1.10.1,  $\phi$  is proportional to experimentally observable steady-state fluorescence intensities by instrument-dependent factors and therefore cannot be easily obtained from such results. The direct fluorometric determination of quantum yields has been established only recently<sup>274</sup> and is based on the use of an integrating sphere, a sample compartment with a highly reflective surface coating (Figure 81).



**Figure 81.** Schematic operating principle of an integrating sphere. a) Prior to addition of the sample, incident excitation light (blue) is scattered of the matrix and reflected inside the sphere several times before reaching the exit port towards the detector. The scattering intensity  $s_0$  is a measure for the baseline photon flux inside the sphere. Background emission and instrumental noise ( $f_0$ ) will be registered at longer wavelengths. b) After addition of the sample, part of the excitation light is absorbed and does not contribute to scattering. The difference between  $s_0$  and the residual scattering intensity  $s_x$  quantifies the decrease in photon flux due to absorption by the sample. At the same time, the fluorescence emission  $f_x$  (green) is reflected multiple times inside the sphere before reaching the detector.<sup>274</sup>

With an appropriate calibration, obtained either from a standardized light source or a well-characterized fluorophore, an integrating sphere setup allows measurement of the photon flux emitted and absorbed by a sample, the ratio of which is the quantum yield. For this, two separate measurements are necessary. In the first one, an emission spectrum of the sample matrix, i.e. a cuvette containing buffer, is recorded, taking care to include the scattered excitation light. The area of this scattering peak ( $s_0$ ) is proportional to the flux of the incident light reduced by absorption from non-fluorophore contents of the sphere. Detector noise might lead to spurious signal intensity at longer wavelengths ( $f_0$ ). Next, the experiment is repeated with a solution of the fluorophore in buffer under otherwise identical conditions. The scatter is now reduced additionally due to light absorption by the fluorophore ( $s_x$ ). At the same time, a fluorescence peak occurs with the area  $f_x$  (Figure 82).



**Figure 82.** General principle of absolute quantum yield calculations. The intensity of scattered ( $s$ ) and emitted ( $f$ ) light are determined in the presence and absence of the fluorophore using a calibrated integrating sphere. The difference in intensities between both experiments is proportional to the flux of photons absorbed (blue) or emitted (green) by the fluorophore. The intensities depicted here are not to scale.

The absolute quantum yield of the fluorophore is then simply calculated according to

$$\phi_{\text{abs}} = \frac{f_x - f_0}{s_0 - s_x} \quad (29)$$

Here, the denominator is proportional to the flux of photons absorbed by the fluorophore and the nominator is proportional to the emitted photon flux. For increased confidence in the result, the measurement can be repeated at different fluorophore concentrations to check for inner filter and aggregation effects. The quantum yield is then obtained by fitting a plot of  $f_x - f_0$  vs  $s_0 - s_x$  with the above equation. Generally, higher accuracy is achieved with more strongly absorbing samples.<sup>275</sup>

Due to requiring relatively large amounts (approximately 220  $\mu\text{L}$ ) of highly concentrated samples, absolute quantum yield measurements are prohibitively expensive for the RNA–ligand complexes investigated in this thesis. Their quantum yields were therefore determined by the relative method.<sup>171,276</sup> In this approach, the UV/Vis absorbance ( $A_x$ ) and fluorescence emission intensity ( $I_x$ ) of an analyte are measured using standard methods. Next, this measurement is repeated with a suitable reference compound to obtain  $A_{\text{std}}$  and  $I_{\text{std}}$  using the same instrument parameters; in addition to being well characterized in terms of its quantum yield ( $\phi_{\text{std}}$ ), the reference substance must absorb and emit light in the same wavelength regions as the analyte and should be employed in such a concentration that its absorbance roughly matches that of the analyte at the fluorescence excitation wavelength ( $\lambda_{\text{ex}}$ ). The measured absorbance ( $A$ ) is then converted to an absorption factor ( $a$ ), which is proportional to the fraction of light absorbed. The excitation monochromator of the spectrofluorometer has a finite bandwidth ( $\Delta\lambda_{\text{ex}}$ ), which is taken into account by integration:

$$a = \int_{\lambda_{\text{ex}} - \Delta\lambda_{\text{ex}}/2}^{\lambda_{\text{ex}} + \Delta\lambda_{\text{ex}}/2} 1 - 10^{-A} d\lambda \quad (30)$$

The emission spectra are background-corrected by subtraction of a solvent/buffer-only spectrum and integrated over the same range for both analyte ( $f_x$ ) and reference ( $f_{\text{std}}$ ), covering their emission bands entirely. In case different solvents were used for the analyte and reference, the respective refractive indices ( $\eta$ ) are included in the calculation of the relative quantum yield:

$$\phi_{\text{rel}} = \phi_{\text{std}} \cdot \frac{f_x}{a_x} \cdot \frac{a_{\text{std}}}{f_{\text{std}}} \cdot \frac{\eta_x^2}{\eta_{\text{std}}^2} \quad (31)$$

Similar to the absolute quantum yield measurement, the experiment can be repeated for a number of different analyte and reference concentrations. A linear model with zero intercept and slope parameter  $m$  is then used to fit plots of  $f$  vs  $a$ , and calculate  $\phi_{\text{rel}}$ :

$$\phi_{\text{rel}} = \phi_{\text{std}} \cdot \frac{m_x}{m_{\text{std}}} \cdot \frac{\eta_x^2}{\eta_{\text{std}}^2} \quad (32)$$

#### 4.1.12.2 Materials

Absolute fluorescence quantum yield samples were measured in a Jasco 6808-H150A liquid cell (1×10×250 mm side length). UV/Vis spectra for relative quantum yield determinations were recorded in a Starna 16.10 sub-micro cuvette (10 mm path length).

HBI derivatives and 4-cyanoindole ribonucleoside were used as stock solutions (100  $\mu\text{M}$  and 10 mM) in DMSO. Coumarin 153 was used as a stock solution (100  $\mu\text{M}$  and 10 mM) in EtOH. Rhodamine B was used as a stock solution (11.5 mM) in ethylene glycol.

#### 4.1.12.3 General protocol

##### Calibration of the integrating sphere

For the range of 350–750 nm, excitation and emission correction files were generated using a Jasco ESC-842 tungsten halogen lamp during the initial setup of the system.

A solution of Rhodamine B in ethylene glycol was used to obtain correction files for the range of 250–600 nm with the help of the spectral correction utility of the instrument.

##### Absolute quantum yield determination

Absolute quantum yield measurements were performed for coumarin 153 and 4-cyanoindole ribonucleoside. The following solutions were prepared:

C153 sample:

	Added volume	Final concentration
10 mM fluorophore in EtOH	2.5 $\mu\text{L}$	50 $\mu\text{M}$
EtOH	497.5 $\mu\text{L}$	
Final volume	500 $\mu\text{L}$	

r4CI sample:

	Added volume	Final concentration
100 mM fluorophore in DMSO	2.5 $\mu\text{L}$	50 $\mu\text{M}$
5x PBS pH 7.4	100 $\mu\text{L}$	1x
H <sub>2</sub> O	397.5 $\mu\text{L}$	
Final volume	500 $\mu\text{L}$	

An aliquot of each sample was serially diluted to prepare a concentration series (50–0.1  $\mu\text{M}$ , 9 steps for C153; 50–2  $\mu\text{M}$ , 5 steps for r4CI). Fluorescence spectra of the samples and of solvent or buffer were recorded in an integrating sphere using these parameters:

	C153/EtOH	r4CI/1x PBS
• Ex wavelength:	418 nm	303 nm
• Em range:	403–750 nm	288–600 nm
• Ex bandwidth:	2.5 nm	5 nm
• Em bandwidth:	2.5 nm	5 nm
• Response:	1 s	1 s
• PMT voltage:	450 V	350 V
• Data interval:	0.2 nm	0.2 nm
• Scan speed:	100 nm/min	100 nm/min

The spectra were integrated over  $\text{Ex} \pm 15$  nm to obtain the scattering values as well as between  $\text{Ex}+15$  nm and the end of the Em range to obtain the emission values. Absolute quantum yields were obtained from linear fits of the data with equation 29.

### Relative quantum yield determination

Relative quantum yield measurements were performed for a number of HBI derivatives and their Chili complexes using coumarin 153 as the reference substance. The following solutions were prepared:

RNA–ligand sample:

	Added volume	Final concentration
Chili RNA aptamer	as required	10/12.3/17 $\mu\text{M}$
1 M KCl	31.25 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	20 $\mu\text{L}$	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM $\text{MgCl}_2$	12.5 $\mu\text{L}$	5 mM
100 $\mu\text{M}$ ligand in DMSO	6.25/12.5/25 $\mu\text{L}$	2.5/5/10 $\mu\text{M}$
$\text{H}_2\text{O}$	as required	
Final volume	250 $\mu\text{L}$	

HBI sample:

	Added volume	Final concentration
100 $\mu\text{M}$ fluorophore in DMSO	6.25/12.5/25 $\mu\text{L}$	2.5/5/10 $\mu\text{M}$
1 M KCl	31.25 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	20 $\mu\text{L}$	40 mM
100 mM $\text{MgCl}_2$	12.5 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	as required	
Final volume	250 $\mu\text{L}$	

C153 sample:

	Added volume	Final concentration
100 $\mu\text{M}$ fluorophore in EtOH	6.25/12.5/25 $\mu\text{L}$	2.5/5/10 $\mu\text{M}$
EtOH	as required	
Final volume	250 $\mu\text{L}$	

Fluorescence spectra of the RNA–ligand, ligand-only and C153 sample were measured using the following parameters:

- Ex wavelength: maximum of the RNA–ligand complex
- Em range: Ex+20–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 2.5 nm
- Response: 1 s
- PMT voltage: adjusted for optimal signal intensity at the concentration of C153
- Data interval: 0.2 nm
- Scan speed: 100 nm/min

Background-corrected spectra (buffer or EtOH, respectively) were integrated to obtain values for  $f_x$  and  $f_{\text{std}}$ . Next, the samples were transferred to a UV/Vis cuvette and absorption spectra were measured with these settings:

- Spectral range: 220–750 nm
- Spectral bandwidth: M1.0 nm
- Response: 0.06 s
- Data interval: 0.2 nm
- Scan speed: 400 nm/min

The absorbance around the excitation wavelength was converted to absorption factors  $a_x$  and  $a_{std}$  according to equation 30. Relative quantum yields were obtained by linear fitting with equation 31.

#### 4.1.13 Screening assay

##### 4.1.13.1 Materials

Microplate-based assays were performed in black Corning 96 Well Half Area plates with flat bottom.

Ligands were used as stock solutions (100  $\mu\text{M}$ ) in DMSO.

##### 4.1.13.2 General protocol

###### Cuvette-based assay

Aptamer, ligand and binding buffer solutions were prepared as follows:

2x aptamer solution:

	Added volume	Final concentration
RNA aptamer	as required	1 $\mu\text{M}$
1 M KCl	12.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	16 $\mu\text{L}$	80 mM
Anneal 3 min at 95 $^{\circ}\text{C}$		
Incubate 20 min at 25 $^{\circ}\text{C}$		
100 mM $\text{MgCl}_2$	5 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	as required	
Final volume	100 $\mu\text{L}$	

2x ligand solution:

	Added volume	Final concentration
100 $\mu\text{M}$ ligand in DMSO	1 $\mu\text{L}$	1 $\mu\text{M}$
1 M KCl	12.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	16 $\mu\text{L}$	80 mM
100 mM $\text{MgCl}_2$	5 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	65.5 $\mu\text{L}$	
Final volume	100 $\mu\text{L}$	

1x buffer solution:

	Added volume	Final concentration
1 M KCl	12.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	16 $\mu\text{L}$	80 mM
100 mM $\text{MgCl}_2$	5 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	66.5 $\mu\text{L}$	
Final volume	100 $\mu\text{L}$	

Samples were prepared by mixing aliquots of the aptamer and ligand solutions (7.5  $\mu\text{L}$  each) and incubating for 3 min at ambient temperature. For time-dependent assays, the samples were measured again after overnight incubation at 4  $^{\circ}\text{C}$ . For background correction, solutions of the ligand and buffer (7.5  $\mu\text{L}$  each) were mixed.

Fluorescence spectra were measured with the following settings:

- Ex wavelength: maximum of the RNA–ligand complex
- Em range: Ex+20–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 5 nm
- Response: 50 ms
- PMT voltage: 680 V
- Data interval: 0.2 nm
- Scan speed: 500 nm/min

Background-corrected spectra were integrated.

#### Microplate-based assay

Sample solutions were prepared as follows:

Aptamer–ligand solution:		
	Added volume	Final concentration
RNA aptamer	as required	0.5 $\mu\text{M}$
1 M KCl	12.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	8 $\mu\text{L}$	40 mM
Anneal 3 min at 95 $^{\circ}\text{C}$		
Incubate 20 min at 25 $^{\circ}\text{C}$		
100 mM $\text{MgCl}_2$	5 $\mu\text{L}$	5 mM
100 $\mu\text{M}$ ligand in DMSO	2 $\mu\text{L}$	2 $\mu\text{M}$
$\text{H}_2\text{O}$	as required	
Final volume	100 $\mu\text{L}$	

Ligand solution:		
	Added volume	Final concentration
100 $\mu\text{M}$ ligand in DMSO	2 $\mu\text{L}$	2 $\mu\text{M}$
1 M KCl	12.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	16 $\mu\text{L}$	80 mM
100 mM $\text{MgCl}_2$	5 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	64.5 $\mu\text{L}$	
Final volume	100 $\mu\text{L}$	

Samples were transferred to a 96 well plate for measurement (95  $\mu\text{L}$  per well). Background correction data was obtained from the ligand sample.

Fluorescence spectra were measured with the following settings:

- Ex wavelength: 405 nm
- Em range: 450–600 nm
- Ex bandwidth: 10 nm

- Em bandwidth: 20 nm
- PMT voltage: high
- Data interval: 1.0 nm
- Scan speed: 600 nm/min

Fluorescence intensities were reported at 540 nm.

#### 4.1.14 Equilibrium binding titrations

##### 4.1.14.1 Overview

Formation of a 1:1 complex (C) between a host or receptor (R) and a guest or ligand (L) leads to an equilibrium that is described by the dissociation constant  $K_d$ .



Each receptor possesses a single binding site, the fractional occupation ( $\theta$ ) of which is given by

$$\theta = \frac{[C]}{[R] + [C]} \quad 1 - \theta = \frac{[R]}{[R] + [C]} \quad (34)$$

The mass action law can therefore be restated as

$$K_d = \frac{(1 - \theta)[L]}{\theta} \quad (35)$$

from which  $\theta$  can be isolated:

$$K_d = \frac{[L] - \theta[L]}{\theta} \quad (36)$$

$$K_d\theta = [L] - \theta[L] \quad (37)$$

$$\theta(K_d + [L]) = [L] \quad (38)$$

$$\theta = \frac{[L]}{K_d + [L]} \quad (39)$$

Considering an arbitrary number ( $n$ ) instead of a single binding site yields the well-known Hill equation:

$$\theta = \frac{[L]^n}{K_d + [L]^n} \quad (40)$$

The equilibrium concentrations of unbound R and L are usually not known during the course of a titration experiments. They can, however, be inferred from their concentrations prior to reaching equilibrium:

$$[R] = [R_0] - [C] \quad [L] = [L_0] - [C] \quad (41)$$

Therefore, the fraction of occupied sites can be expressed in terms of the initial concentration of R:

$$\theta = \frac{[C]}{[R_0]} \quad (42)$$



Consequently, the concentration of C is given by

$$[C] = [R_0] \frac{[L]^n}{K_d + [L]^n} \quad (43)$$

This equation describes a binding hyperbola and is commonly used for fitting titration experiments with the tacit assumption of  $[L] = [L_0]$ , which is only valid if L is in large excess over R at all points during the titration. For a more rigorous treatment, the mass action law is rewritten once again with the help of equation 41:

$$K_d = \frac{([R_0] - [C])([L_0] - [C])}{[C]} = \frac{[C]^2 - ([R_0] + [L_0])[C] + [R_0][L_0]}{[C]} \quad (44)$$

This results in a quadratic equation relating the concentration of C to the dissociation constant:

$$[C]^2 - ([R_0] + [L_0] + K_d)[C] + [R_0][L_0] = 0 \quad (45)$$

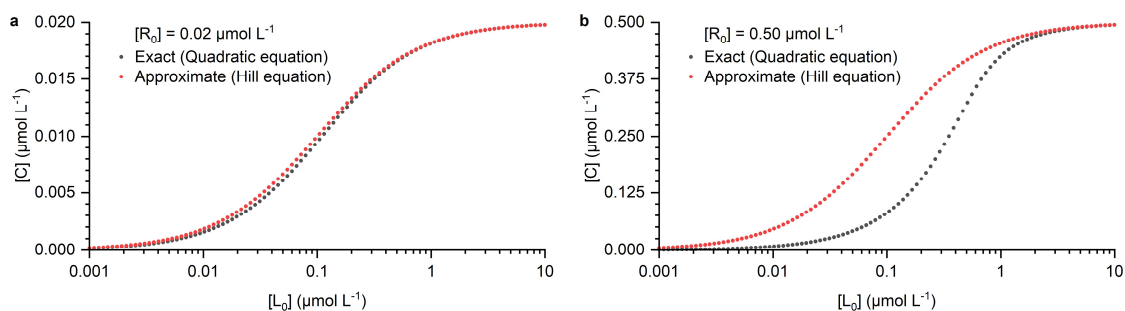
The valid solution of this equation is

$$[C] = \frac{1}{2} \left( [R_0] + [L_0] + K_d - \sqrt{([R_0] + [L_0] + K_d)^2 - 4[R_0][L_0]} \right) \quad (46)$$

The concentration of the complex is proportional to the intensity ( $I$ ) observed in a fluorescence titration experiment by a factor  $b$ . If required, an offset ( $a$ ) can be included to account for background signal at the start of the experiment:

$$I = b[C] \quad I = a + (b - a)[C] \quad (47)$$

Depending on the circumstances,  $[C]$  is calculated either exactly from equation 46 or, if justified, approximately from equation 43. Deviations between both formalisms become readily apparent when a high initial concentration of R leads to significant depletion of L in the early stages of the titration (Figure 83).



**Figure 83.** Simulated titration curves for the 1:1 association between a receptor (R) and its ligand (L) with a  $K_d$  of  $0.1 \mu\text{mol L}^{-1}$ . a) When the initial concentration of R is low, the concentration of L does not change significantly from its initial value to the equilibrium value and both can be used interchangeably. b) A high initial concentration of R leads to pronounced depletion of L and results in deviations that are not taken into account by the Hill equation. Data points were calculated using equations 43 or 46, respectively.

In practice, it was preferred to titrate varying concentrations of RNA into a fixed concentration of ligand to avoid various issues such as limited solubility of the ligand and increasing background fluorescence due to aggregation. Because equation 46 is symmetric with respect to the use of R and L, both approaches are mathematically equivalent. The Hill equation fundamentally describes a binding hyperbola, and so equation 43 is still valid when the identity of R and L are swapped, however the exponential coefficient  $n$  is then a purely empirical cooperativity parameter and should be 1 for single-site binding.

#### 4.1.14.2 Materials

Ligands were used as stock solutions (100  $\mu\text{M}$ ) in DMSO.

#### 4.1.14.3 General protocol

Typical aptamer, ligand and binding buffer solutions were prepared as follows:

2x aptamer solution:

	Added volume	Final concentration
62.5 $\mu\text{M}$ Chili RNA aptamer	11.5 $\mu\text{L}$	16 $\mu\text{M}$
H <sub>2</sub> O	33.5 $\mu\text{L}$	
Final volume	45 $\mu\text{L}$	

4x ligand solution:

	Added volume	Final concentration
100 $\mu\text{M}$ ligand in DMSO	2 $\mu\text{L}$	0.4 $\mu\text{M}$
100 mM MgCl <sub>2</sub>	100 $\mu\text{L}$	20 mM
H <sub>2</sub> O	398 $\mu\text{L}$	
Final volume	500 $\mu\text{L}$	

4x buffer solution:

	Added volume	Final concentration
1 M KCl	250 $\mu\text{L}$	500 mM
500 mM HEPES pH 7.5	160 $\mu\text{L}$	160 mM
H <sub>2</sub> O	90 $\mu\text{L}$	
Final volume	500 $\mu\text{L}$	

The 2x aptamer solution was serially diluted 1:1 with H<sub>2</sub>O to make a 15-step dilution series (7.5  $\mu\text{L}$  sample volume per step). Next, 4x buffer (3.75  $\mu\text{L}$  each) was added, the samples were annealed at 95 °C for 3 min and then kept at ambient temperature for 20 min. Finally, 4x ligand (3.75  $\mu\text{L}$  each) was added to bring the sample volume to a total of 15  $\mu\text{L}$ . All samples were incubated at 4 °C for 16 h. For background correction, a solution of 1x ligand and 1x buffer in H<sub>2</sub>O was prepared.

The above protocol results in a titration series containing 0.0005–8  $\mu\text{M}$  RNA and 0.1  $\mu\text{M}$  ligand. Samples with different RNA and ligand conditions (as indicated in the respective figures) were prepared analogously.

The parameters for fluorescence measurements were adjusted accordingly:

- Ex wavelength: maximum of the RNA–ligand complex
- Em range: Ex+20–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 5 nm
- Response: 1 s
- PMT voltage: adjusted for optimal signal intensity at the highest RNA concentration
- Data interval: 0.2 nm
- Scan speed: 500 nm/min

Background-corrected spectra were integrated and the resulting intensities were fitted to equation 47, including a baseline offset if required. If the data quality did not warrant fitting with

the exact model (indicated by slow convergence of the nonlinear least squares), the Hill equation was used instead in the calculation. In such cases, Hill coefficients ( $n$ ) were very close to 1 as expected for single-site binding.

#### 4.1.15 Ligand competition assay

##### 4.1.15.1 Materials

Ligand solutions in H<sub>2</sub>O (10  $\mu$ M) were freshly prepared from stock solutions in DMSO (10 mM).

##### 4.1.15.2 General protocol

Two samples containing Chili and either DMHBAI (**68**) or DMBAI (**77**) were prepared as follows:

Chili–DMHBAI sample (A):

	Added volume	Final concentration
Chili RNA aptamer	As required	0.5 $\mu$ M
1 M KCl	3.75 $\mu$ L	125 mM
500 mM HEPES pH 7.5	2.4 $\mu$ L	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM MgCl <sub>2</sub>	1.5 $\mu$ L	5 mM
10 $\mu$ M DMHBAI in H <sub>2</sub> O	1.5 $\mu$ L	0.5 $\mu$ M
H <sub>2</sub> O	As required	
Final volume	30 $\mu$ L	

Chili–DMBAI sample (B):

	Added volume	Final concentration
Chili RNA aptamer	As required	0.5 $\mu$ M
1 M KCl (1 M)	3.75 $\mu$ L	125 mM
500 mM HEPES pH 7.5	2.4 $\mu$ L	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM MgCl <sub>2</sub>	1.5 $\mu$ L	5 mM
10 $\mu$ M DMBAI in H <sub>2</sub> O	15 $\mu$ L	5 $\mu$ M
H <sub>2</sub> O	As required	
Final volume	30 $\mu$ L	

They were split into two aliquots (15  $\mu$ L) each, incubated for 3 min at ambient temperature and fluorescence spectra were recorded with the parameters stated below. Afterwards, the first aliquot of sample A was mixed with DMBAI solution (10  $\mu$ M, 15  $\mu$ L), while the second aliquot was mixed with an equal volume of H<sub>2</sub>O. Likewise, the first aliquot of sample B was mixed with DMHBAI solution (10  $\mu$ M, 15  $\mu$ L), while the second aliquot was mixed with a more dilute solution of DMHBAI in H<sub>2</sub>O (1  $\mu$ M, 15  $\mu$ L). After 3 min incubation at ambient temperature, spectra were measured again using the same parameters as before:

- Ex wavelength: 410 nm
- Em range: 430–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 5 nm
- Response: 50 ms
- PMT voltage: 680 V

- Data interval: 0.2 nm
- Scan speed: 500 nm/min

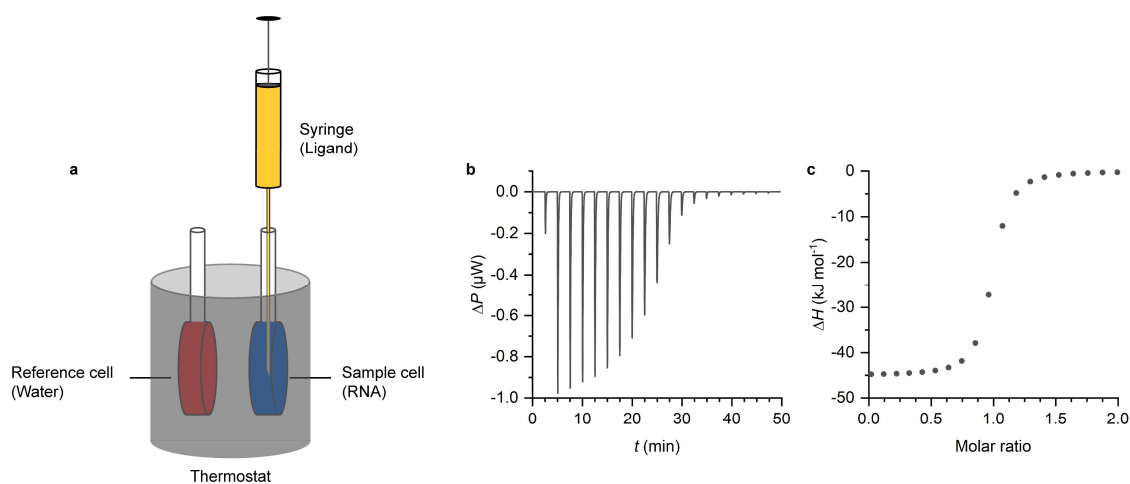
The resulting emission spectra were integrated.

#### 4.1.16 Isothermal titration calorimetry

##### 4.1.16.1 Overview

Any interaction between two reaction partners is accompanied by a change in energy of the system, which comprises enthalpic effects arising from the formation and breaking of covalent or noncovalent bonds as well as entropic effects related to the number of degrees of freedom the system can access. Together, these thermodynamic parameters determine the equilibrium constant of the reaction, which can be determined by any method capable of monitoring, either directly or indirectly, the composition of the reaction mixture. While this is sufficient for the characterization of a noncovalent association, detailed information about the associated enthalpy and entropy changes can give additional insight into the underlying binding mechanism.

By monitoring the heat released or consumed during a reaction, isothermal titration calorimetry (ITC) provides detailed information about equilibrium constants, stoichiometry and reaction enthalpy. During an ITC experiment, one of the reactants is placed inside a sample cell that is maintained at the same temperature as an adjacent reference cell. Aliquots of the second reagent are titrated into the sample cell, allowing sufficient time for reequilibration after each addition. To compensate for the heat exchange between the reaction mixture and its surroundings, differential power ( $\Delta P$ ) is applied to the sample and reference cells. Integration of  $\Delta P$  for each injection step results in a titration curve of molar enthalpy vs molar ratio of the reactants. The difference between lower and upper asymptote of this binding isotherm corresponds the reaction enthalpy  $\Delta H$ . The molar ratio at the inflection point quantifies the reaction stoichiometry  $N$ ; the slope at the inflection point is determined by the equilibrium constant  $K$ .



**Figure 84.** a) Schematic illustration of an isothermal titration calorimeter. The macromolecular sample and a reference substance, e.g. water, are loaded into two identical thermostatted cells which are kept at the same temperature by heating elements. Aliquots of the small molecule ligand are titrated into the sample, which either releases or consumes heat. b) The differential power required to keep the sample and reference at the same temperature is recorded for each injection of the ligand. c) The enthalpy change associated with each titration step is obtained by integration of the differential power and plotted against the molar ratio of the reactants. The bottom asymptote of the curve reflects the binding enthalpy and the inflection point is located at the molar ratio corresponding to the stoichiometry of the supramolecular complex. Fitting with a suitable model yield the dissociation constant of the reaction.

Binding isotherms were fitted with a model for a single set of identical binding sites. A receptor (R) and ligand (L) react to form a complex (C). The ligand occupies a certain fraction  $\theta$  of the total number of available binding sites:

$$\begin{aligned} R + L &\stackrel{K}{\rightleftharpoons} C \\ K &= \frac{[C]}{[R][L]} = \frac{\theta}{(1-\theta)[L]} \end{aligned} \quad (48)$$

At any point during the titration, the total concentration of ligand,  $[L_0]$ , is partitioned between unbound ligand and ligand occupying a fraction of  $N$  binding sites:

$$[L_0] = [L] + N\theta[R_0] \quad (49)$$

Combining these two expressions results in a quadratic equation for the fractional occupation:

$$\begin{aligned} K &= \frac{\theta}{(1-\theta)([L_0] - N\theta[R_0])} = \frac{\theta}{[L_0] - N\theta[R_0] - \theta[L_0] + N\theta^2[R_0]} \\ N\theta^2[R_0] - N\theta[R_0] - \theta[L_0] - \frac{\theta}{K} + [L_0] &= 0 \\ \theta^2 - \theta \left( 1 + \frac{[L_0]}{N[R_0]} + \frac{1}{N[R_0]K} \right) + \frac{[L_0]}{N[R_0]} &= 0 \end{aligned} \quad (50)$$

The valid solution of this equation is

$$\theta = \frac{1}{2} \left( 1 + \frac{[L_0]}{N[R_0]} + \frac{1}{N[R_0]K} - \sqrt{\left( 1 + \frac{[L_0]}{N[R_0]} + \frac{1}{N[R_0]K} \right)^2 - 4 \frac{[L_0]}{N[R_0]}} \right) \quad (51)$$

The heat content  $Q$  of the sample contained in a volume  $V_0$  at the start of the titration is defined as 0. Therefore, after reaching a certain fractional occupation in the  $i^{\text{th}}$  titration step, the current heat content depends on the change in enthalpy  $\Delta H$  relative to the initial state:

$$\begin{aligned} Q(i) &= N\theta[R_0] \cdot \Delta H V_0 \\ &= \frac{N[R_0] \cdot \Delta H V_0}{2} \left( 1 + \frac{[L_0]}{N[R_0]} + \frac{1}{N[R_0]K} - \sqrt{\left( 1 + \frac{[L_0]}{N[R_0]} + \frac{1}{N[R_0]K} \right)^2 - 4 \frac{[L_0]}{N[R_0]}} \right) \end{aligned} \quad (52)$$

The observable quantity in ITC is the heat released between individual titration steps:

$$\Delta Q(i)_{\text{raw}} = Q(i) - Q(i-1) \quad (53)$$

A correction term is then introduced to account for the increase in volume ( $\Delta V_i$ ) due to addition of the titrant:

$$\boxed{\Delta Q(i)_{\text{corr}} = Q(i) - Q(i-1) + \frac{\Delta V_i}{V_0} \left( \frac{Q(i) + Q(i-1)}{2} \right)} \quad (54)$$

This equation is fitted to the ITC thermogram using a proprietary Levenberg-Marquardt algorithm implemented in the MicroCal PEAQ-ITC software package.

#### 4.1.16.2 Materials

A stock solution of the Chili RNA aptamer (150  $\mu\text{M}$ ) was dialyzed against ultrapure  $\text{H}_2\text{O}$  using a Slide-A-Lyzer MINI device (3.5K MWCO, 0.5 mL, ThermoFisher Scientific) according to the manufacturer's instructions.

Ligand stock solutions (10 mM) were prepared in DMSO.

#### 4.1.16.3 General protocol

Typical samples were prepared as follows:

RNA sample:		
	Added volume	Final concentration
150 $\mu\text{M}$ Chili RNA aptamer, dialyzed	150 $\mu\text{L}$	15 $\mu\text{M}$
1 M KCl	187.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	120 $\mu\text{L}$	40 mM
DMSO	22.5 $\mu\text{L}$	1.5%
Anneal 3 min at 95 $^{\circ}\text{C}$		
Incubate 20 min at 25 $^{\circ}\text{C}$		
100 mM $\text{MgCl}_2$	75 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	945 $\mu\text{L}$	
Final volume	1500 $\mu\text{L}$	

Ligand sample:		
	Added volume	Final concentration
$\text{H}_2\text{O}$	150 $\mu\text{L}$	
1 M KCl	187.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	120 $\mu\text{L}$	40 mM
10 mM ligand in DMSO	22.5 $\mu\text{L}$	150 $\mu\text{M}$
Heat 3 min to 95 $^{\circ}\text{C}$		
Incubate 20 min at 25 $^{\circ}\text{C}$		
100 mM $\text{MgCl}_2$	75 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	945 $\mu\text{L}$	
Final volume	1500 $\mu\text{L}$	

Buffer sample:		
	Added volume	Final concentration
$\text{H}_2\text{O}$	150 $\mu\text{L}$	
1 M KCl	187.5 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	120 $\mu\text{L}$	40 mM
DMSO	22.5 $\mu\text{L}$	1.5%
Heat 3 min to 95 $^{\circ}\text{C}$		
Incubate 20 min at 25 $^{\circ}\text{C}$		
100 mM $\text{MgCl}_2$	75 $\mu\text{L}$	5 mM
$\text{H}_2\text{O}$	945 $\mu\text{L}$	
Final volume	1500 $\mu\text{L}$	

For DMHBI<sup>+</sup>, the final concentrations of RNA and dye were reduced to 10 and 100  $\mu\text{M}$ , respectively. Measurements were performed under the following conditions:

- Volume of RNA sample in the cell: 280  $\mu\text{L}$

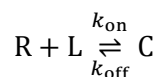
- Volume of ligand sample in the syringe: 40  $\mu\text{L}$
- Temperature: 25  $^{\circ}\text{C}$
- Reference power: 41.9  $\mu\text{W}$
- Feedback: High
- Stir speed: 750 rpm
- Initial delay: 60 s
- First injection: 0.4  $\mu\text{L}$  over 0.8 s
- Other injections: 12 $\times$ 3.0  $\mu\text{L}$  over 6.0 s (DMHBI) or 18 $\times$ 2.0  $\mu\text{L}$  over 4.0 s (DMHBI $^+$ )
- Spacing: 150 s

A baseline correction was performed by subtracting the mean injection heat of ligand into buffer from the titration data. The data points were fitted with the model described in section 4.1.16.1. For an initial fit, the number of binding sites was constrained to 1 and the active concentration of RNA in the cell was varied. Refinement of the fit was performed by varying the number of binding sites and fixing the RNA concentration.

#### 4.1.17 Binding kinetics

##### 4.1.17.1 Overview

The dynamics of 1:1 complex (C) formation between a receptor (R) and ligand (L) are characterized by the rate constants of the forward and reverse reactions:



Here,  $k_1$  is the second order rate constant for the association and  $k_{-1}$  is the first order rate constant of dissociation. The differential rate law for the formation of C is

$$\frac{d[\text{C}]}{dt} = -\frac{d[\text{R}]}{dt} = -\frac{d[\text{L}]}{dt} = k_{\text{on}}[\text{R}][\text{L}] - k_{\text{off}}[\text{C}] \quad (55)$$

An equilibrium with the dissociation constant  $K_d$  is reached when both processes occur at the same rate:

$$\left. \frac{d[\text{C}]}{dt} \right|_{\text{eq}} = 0 \quad (56)$$

$$k_{\text{on}}[\text{R}][\text{L}] = k_{\text{off}}[\text{C}] \quad (57)$$

$$\frac{[\text{R}][\text{L}]}{[\text{C}]} = \frac{k_{\text{off}}}{k_{\text{on}}} = K_d \quad (58)$$

The differential rate law can be rearranged by expressing the concentration of C in terms of reacted R. If L is present in large excess over R at the start of the complex formation, its consumption can be neglected and its concentration is included in a new rate constant  $k'$ .

$$[\text{R}] = [\text{R}_0] - [\text{C}] \quad [\text{L}] = [\text{L}_0] - [\text{C}] \approx [\text{L}_0] \quad (59)$$

$$\begin{aligned} \frac{d[\text{C}]}{dt} &= k_{\text{on}}[\text{R}][\text{L}_0] - k_{\text{off}}[\text{C}] = k'[\text{R}] - k_{\text{off}}([\text{R}_0] - [\text{R}]) \\ &= (k' + k_{\text{off}})[\text{R}] - k_{\text{off}}[\text{R}_0] \end{aligned} \quad (60)$$

Integration results in an exponential time course for the formation of C:

$$[C] = \frac{k'[R_0]}{k' + k_{\text{off}}} \{1 - \exp[-(k' + k_{\text{off}})t]\} \quad (61)$$

Experimentally, the concentration of C is proportional to the observed fluorescence intensity enhancement ( $I$ ) over the background ( $I_0$ ) and increases with an apparent rate constant of  $k_{\text{obs}}$ :

$$k_{\text{obs}} = k_{\text{on}}[L_0] + k_{\text{off}} \quad (62)$$

$$I = I_0 + A \exp(-k_{\text{obs}}t) \quad (63)$$

By repeating the experiment for a number of ligand concentrations and plotting  $k_{\text{obs}}$  vs  $[L_0]$ ,  $k_{\text{on}}$  is obtained from the slope of a linear fit to the data. In principle,  $k_{\text{off}}$  is given by the y-axis intercept of such a fit. However, the result is significant only when  $[L_0]$  is on the order of  $K_d$ .<sup>277,278</sup>

For the systems investigated in this thesis, biphasic fluorescence time courses were observed due to separate binding and refolding processes, both of which resulted in fluorescence enhancement. The fit function therefore included a second exponential term with a separate  $k_{\text{obs}}$ .<sup>279</sup>

$$I = I_0 + A_1 \exp(-k_{\text{obs}1}t) + A_2 \exp(-k_{\text{obs}2}t) \quad (64)$$

#### 4.1.17.2 Materials

Ligands were used as stock solutions (350  $\mu\text{M}$ ) in DMSO.

Samples were measured in a Jasco FMM-200 micro quartz cuvette (5×5 mm) containing a PTFE stir bar.

#### 4.1.17.3 General protocol

The following solutions were prepared:

RNA sample:		
	Added volume	Final concentration
Chili RNA aptamer	as required	26.25 nM
1 M KCl	525 $\mu\text{L}$	131.25 mM
500 mM HEPES pH 7.5	336 $\mu\text{L}$	42 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM MgCl <sub>2</sub>	220 $\mu\text{L}$	5.5 mM
H <sub>2</sub> O	as required	
Final volume	4000 $\mu\text{L}$	
Ligand samples:		
	Added volume	Final concentration
350 $\mu\text{M}$ ligand in DMSO	1.8/2.4/3.6/4.8/6 $\mu\text{L}$	15.75/21/31.5/42/52.5 $\mu\text{M}$
H <sub>2</sub> O	as required	
Final volume	40 $\mu\text{L}$	

For each concentration, an aliquot the respective ligand sample (20  $\mu\text{L}$ ) was quickly injected into an aliquot of the RNA sample (400  $\mu\text{L}$ ) while the fluorescence intensity was monitored for up to 30 min under continuous stirring. Measurements were repeated after incubating the RNA sample for 72 h at 4 °C. The following parameter settings were used:

- Ex wavelength: maximum of the RNA–ligand complex



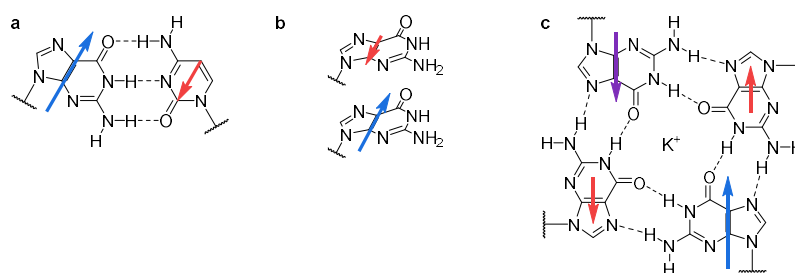
- Em wavelength: maximum of the RNA–ligand complex
- Ex bandwidth: 1 nm
- Em bandwidth: 20 nm
- Response: 50 ms
- PMT voltage: adjusted for optimal signal intensity
- Data interval: 2 s

Curves were fitted with a biexponential association model (equation 64) to obtain apparent rate constants  $k_{\text{obs}}$ . Plots of  $k_{\text{obs}}$  vs ligand concentration were then fitted with a linear equation, the slope of which is the association rate  $k_{\text{on}}$ . Dissociation rates were then calculated from experimentally determined dissociation constants according to  $k_{\text{off}} = K_{\text{d}} k_{\text{on}}$ .

#### 4.1.18 Thermal denaturation curves of double-stranded oligonucleotides

##### 4.1.18.1 Overview

In general, the absorbance of individual nucleobases does not simply behave additively when present in an oligonucleotide. For canonical structure elements such as helices formed from stacked Watson-Crick base pairs, interactions between the transition dipole moments of the bases lead to a decrease in total absorbance; a double strand will have a lower absorbance than its two constituent single strands (Figure 85a, b).<sup>280</sup>



**Figure 85.** Schematic depiction of transition dipole moment interactions between nucleobases in a) a Watson-Crick base pair, b) a  $\pi$ -stacked arrangement and c) a G-tetrad. In the first two cases the transition dipole moment of the first chromophore (blue) induces an opposite transition dipole moment in the second one (red), which results in a net decrease of the absorbance. In a G-tetrad, however, the specific arrangement of guanines leads to a synergistic enhancement of the transition dipole moment in the second-to-next chromophore (violet).

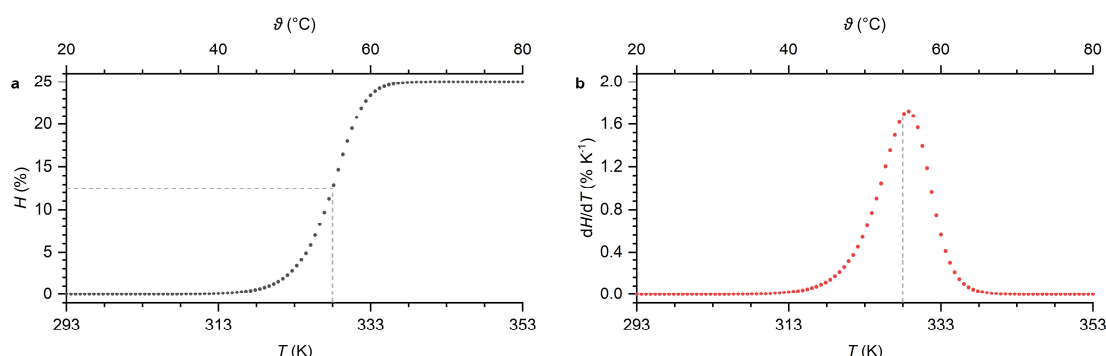
Thermal denaturation of a double-stranded oligonucleotide therefore results in a characteristic sigmoidal increase in absorbance, which is usually expressed in terms of the percentage hyperchromicity  $H$ :

$$H = 100 \frac{A - A_{\text{min}}}{A_{\text{min}}} \quad (65)$$

Such hyperchromic transitions are easily observed in the region between 250 and 280 nm. Denaturation of non-canonical structure elements, e.g. G-quartets or i-motifs, has little impact in this spectral region. Instead, a characteristic decrease in absorbance, i.e. a hypochromic effect, can be observed at 295 nm (Figure 85c).<sup>183</sup>

The denaturation of an oligonucleotide is characterized by a melting temperature  $T_m$ , which is defined as the temperature  $T$  at which half of the single strands are engaged in duplex formation. It can be estimated from a plot of  $H$  vs  $T$  by finding the midpoint between the bottom and top baselines of the curve (Figure 86a). Inaccuracies occur for highly sloping baseline and melting transitions that occur over a broad temperature range. In well-behaved cases, however, an accuracy of 0.5 °C is feasible.<sup>226</sup> In principle,  $T_m$  should coincide with the inflection point of the thermal denaturation curve, i.e. the maximum of  $dH/dT$ . For bimolecular systems, however, this

relationship does not strictly hold true. Instead, the maximum of  $dH/dT^{-1}$  occurs at a folded fraction of  $2^{1/2} - 1 \approx 0.414$ , meaning  $T_m$  will be overestimated. While the error introduced by deriving with respect to  $T$  instead of  $T^{-1}$  is small, the overall deviation from the actual  $T_m$  can be significant (Figure 86b).



**Figure 86.** a) Simulated thermal denaturation curve for a non-self-complementary duplex with the following parameters:  $T_m = 55\text{ }^\circ\text{C}$ ,  $\Delta H_m = -600\text{ kJ mol}^{-1}$ ,  $c_{\text{tot}} = 10\text{ }\mu\text{mol L}^{-1}$ ,  $A_{\text{min}} = 1.00$ ,  $A_{\text{max}} = 1.25$ . The data points are calculated using equation 83. b) Numerical first derivative of the thermal denaturation curve.  $T_m$  is indicated with a dashed line and lies  $0.8\text{ }^\circ\text{C}$  below the maximum of the derivative curve.

A more robust protocol for the determination of  $T_m$  relies on direct non-linear fitting of the thermal denaturation curve.

The derivation of the underlying fit equation starts from the mass action law for a bimolecular association reaction between two single-stranded oligonucleotides (S) to form a duplex (D) with an equilibrium constant  $K$ :



A sum of concentrations is introduced, which remains constant throughout the reaction by accounting for the dimerization:

$$c_{\text{tot}} = [S^1] + [S^2] + 2[D] \quad (67)$$

Under the usual experimental conditions, equal concentrations of both single-stranded oligonucleotides are employed:

$$[S^1] = [S^2] = [S] \quad (68)$$

The fractional degree of hybridization ( $\alpha$ ) is then given by

$$\alpha = \frac{2[D]}{c_{\text{tot}}} \quad (69)$$

$$(1 - \alpha) = \frac{[S^1] + [S^2]}{c_{\text{tot}}} = \frac{2[S]}{c_{\text{tot}}} \quad (70)$$

After solving for the respective concentrations, the mass action law is expressed in terms of  $\alpha$ :

$$[D] = \frac{\alpha c_{\text{tot}}}{2} \quad (71)$$

$$[S] = \frac{(1 - \alpha)c_{\text{tot}}}{2} \quad (72)$$

$$K = \frac{4\alpha c_{\text{tot}}}{2(1 - \alpha)^2 c_{\text{tot}}^2} = \frac{2\alpha}{(1 - \alpha)^2 c_{\text{tot}}} \quad (73)$$

Changes in the degree of hybridization are thus related to the temperature dependence of the equilibrium constant:

$$K = \exp\left(-\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R}\right) \quad (74)$$

In this expression, the enthalpy ( $\Delta H^0$ ) and entropy ( $\Delta S^0$ ) of the reaction jointly describe the melting point and the local slope of the thermal denaturation curve. During non-linear regression, this would result in a parameter dependency. In contrast to  $\Delta S^0$ ,  $T_m$  is directly observable and would be a more suitable regression parameter.

When a dimerization equilibrium has been reached at a given temperature, the Gibbs energy of the reaction is ( $\Delta G$ ) is 0:

$$\Delta G = \Delta G^0 + RT \ln K = 0 \quad (75)$$

$$\Delta G^0 = -RT \ln K \quad (76)$$

At  $T_m$  ( $\alpha = 0.5$ ), the expression of the equilibrium constant reduces to

$$K_{T_m} = \frac{4}{c_{\text{tot}}} \quad (77)$$

The definition of  $\Delta G^0$  is then used to isolate  $\Delta S^0$ :

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad (78)$$

$$-RT_m \ln \frac{4}{c_{\text{tot}}} = \Delta H^0 - T_m \Delta S^0 \quad (79)$$

$$\Delta S^0 = \frac{\Delta H^0}{T_m} + R \ln \frac{4}{c_{\text{tot}}} \quad (80)$$

Substitution into equation 74 results in

$$K = \exp\left(-\frac{\Delta H^0}{RT} + \frac{\Delta H^0}{RT_m} + \ln \frac{4}{c_{\text{tot}}}\right) \quad (81)$$

Rearranging equation 73,  $\alpha$  can be calculated as function of temperature using the parameters  $\Delta H^0$  and  $T_m$ :

$$\alpha = \frac{Kc_{\text{tot}} - \sqrt{2Kc_{\text{tot}} + 1} + 1}{Kc_{\text{tot}}} \quad (82)$$

A full description of the thermal denaturation curve relates  $\alpha$  to the observed change in absorbance going from the lower to the upper baseline ( $\text{base}_D$  and  $\text{base}_S$ , respectively):

$$A = \alpha \cdot \text{base}_D + (1 - \alpha) \cdot \text{base}_S \quad (83)$$

$$\text{base}_D = m_D T + b_D \quad (84)$$

$$\text{base}_S = m_S T + b_S \quad (85)$$

#### 4.1.18.2 Materials

Samples with a concentration of 1–5  $\mu\text{M}$  per individual oligonucleotide were measured in VARIAN semi-micro cuvettes (10 mm path length). For higher concentrated samples, Hellma macro cuvettes with a reduced sample volume (1 mm path length) were used.

#### 4.1.18.3 General protocol

To assess the concentration dependence of  $T_m$ , samples with concentrations between 1 and 20  $\mu\text{M}$  of each single strand were prepared:

1/2/5  $\mu\text{M}$  oligonucleotide sample:

	Added volume	Final concentration
DNA/RNA	As required	1/2/5 $\mu\text{M}$
Complementary DNA/RNA	As required	1/2/5 $\mu\text{M}$
5x PBS pH 7.4	100 $\mu\text{L}$	1x
H <sub>2</sub> O	As required	
Final volume	500 $\mu\text{L}$	

10/20  $\mu\text{M}$  oligonucleotide sample:

	Added volume	Final concentration
DNA/RNA	As required	10/20 $\mu\text{M}$
Complementary DNA/RNA	As required	10/20 $\mu\text{M}$
5x PBS pH 7.4	60 $\mu\text{L}$	1x
H <sub>2</sub> O	As required	
Final volume	300 $\mu\text{L}$	

Inside the cuvettes, samples were layered with 0.5 cm of silicon oil to minimize evaporation during the measurement. Five temperature ramps between 10 and 95  $^{\circ}\text{C}$  were collected with the following settings:

- Wavelengths: 260/270 nm
- Spectral bandwidth: 1 nm
- Averaging time: 2 s
- Heating rate: 0.5  $^{\circ}\text{C}/\text{min}$

Melting temperatures and reaction enthalpies were determined by fitting the obtained curves (at 260 nm, the first ramp was not included in the analysis) to the combined equations 81–85. An initial value for  $T_m$  was estimated from the inflection point of the melting curve or by fitting the whole curve. Then, the data range for fitting was decreased to  $T_m \pm 10 \text{ K}$  to obtain a reliable value for  $\Delta H^{\circ}$ .

#### 4.1.19 Thermal stability of the Chili aptamer

##### 4.1.19.1 Materials

For UV/Vis measurements VARIAN semi-micro cuvettes (10 mm path length) were used. Samples were then transferred to Starna sub-micro cuvettes (5 $\times$ 2 mm) for fluorescence measurements.

DMHBI<sup>+</sup> was used as a stock solution (100  $\mu\text{M}$ ) in DMSO.

#### 4.1.19.2 General protocol

The thermal stability of the Chili aptamer alone and in complex with DMHBI<sup>+</sup> was determined semi-quantitatively by UV/Vis and fluorescence spectroscopy. Addition of Mg<sup>2+</sup> was omitted to prevent cleavage of the RNA at elevated temperatures. The following samples were prepared:

RNA sample:		
	Added volume	Final concentration
Chili RNA aptamer	As required	2 $\mu$ M
1 M KCl	62.5 $\mu$ L	125 mM
500 mM HEPES pH 7.5	40 $\mu$ L	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
H <sub>2</sub> O	As required	
Final volume	500 $\mu$ L	

RNA–ligand sample:		
	Added volume	Final concentration
Chili RNA aptamer	As required	2 $\mu$ M
1 M KCl	62.5 $\mu$ L	125 mM
500 mM HEPES pH 7.5	40 $\mu$ L	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 $\mu$ M DMHBI <sup>+</sup> in DMSO	10 $\mu$ L	2 $\mu$ M
H <sub>2</sub> O	As required	
Final volume	500 $\mu$ L	

Inside the semi-micro cuvettes, samples were layered with 0.5 cm of silicon oil to minimize evaporation during the measurement. Five UV/Vis temperature ramps between 10 and 95 °C were collected with the following settings:

- Wavelengths: 260/295 nm
- Spectral bandwidth: 1 nm
- Averaging time: 2 s
- Heating rate: 0.5 °C/min

Aliquots of the samples (100  $\mu$ L) were then removed, transferred to sub-micro cuvettes and again layered with silicon oil to collect identical fluorescence temperature ramps:

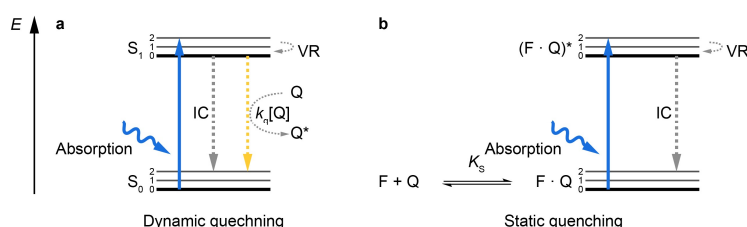
- Ex wavelength: 413 nm
- Em wavelength: 542 nm
- Ex bandwidth: 5 nm
- Em bandwidth: 5 nm
- Response: 0.1 s
- PMT voltage: 800 V
- Heating rate: 0.5 °C/min

For visual analysis, numerical first derivatives of the resulting absorbance and fluorescence curves were calculated.

## 4.1.20 Stern-Volmer titrations

## 4.1.20.1 Overview

Molecular contact between a fluorophore and quencher reduces the fluorescence intensity of said fluorophore without affecting its chemical makeup. In collisional (dynamic) quenching, this contact occurs transiently when the quencher encounters the fluorophore by diffusion during its excited state lifetime and induces non-radiative return to the ground state (Figure 87a). Static quenching, on the other hand, describes the permanent formation of a non-fluorescent complex between the fluorophore and quencher (Figure 87b).



**Figure 87.** a) Collision of an excited fluorophore with a quencher Q deactivates the excited state with a rate constant  $k_q$ . b) Complex formation between a fluorophore and a quencher with the association constant  $K_s$  generates a nonfluorescent species.

Recalling the differential equation that describes the population of an excited state under steady-state conditions,

$$\frac{d[F^*]}{dt} = k_{\text{abs}} - (k_f + k_{\text{IC}} + k_{\text{ISC}})[F^*] = 0 \quad (14)$$

the presence of a collisional quencher (Q) introduces an additional contribution with the rate constant  $k_q$  to radiationless decay:

$$\frac{d[F^*]}{dt} = k_{\text{abs}} - (k_f + k_{\text{IC}} + k_{\text{ISC}} + k_q[Q])[F^*] = 0 \quad (86)$$

This, in turn, leads to a decreased fluorescence quantum yield:

$$\phi = \frac{k_f}{k_f + k_{\text{IC}} + k_{\text{ISC}} + k_q[Q]} \quad (87)$$

Therefore, the ratio of quantum yields in the absence and presence of the quencher is given by

$$\begin{aligned} \frac{\phi_0}{\phi} &= \left( \frac{k_f}{k_f + k_{\text{IC}} + k_{\text{ISC}}} \right) \left( \frac{k_f + k_{\text{IC}} + k_{\text{ISC}} + k_q[Q]}{k_f} \right) = \frac{k_f + k_{\text{IC}} + k_{\text{ISC}} + k_q[Q]}{k_f + k_{\text{IC}} + k_{\text{ISC}}} \\ &= 1 + \frac{k_q}{k_f + k_{\text{IC}} + k_{\text{ISC}}}[Q] \end{aligned} \quad (88)$$

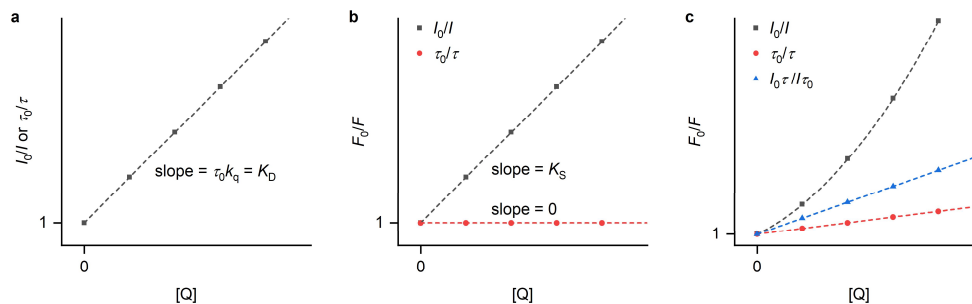
Using the definition of the fluorescence lifetime (see section 4.1.11.1), this expression simplifies to the Stern-Volmer equation:

$$\frac{\phi_0}{\phi} = 1 + \tau_0 k_q [Q] \quad (89)$$

Since both intensity and lifetime are proportional to the fluorescence quantum yield, it follows that either parameter obeys the same linear trend with respect to the quencher concentration (Figure 88a):

$$\boxed{\frac{I_0}{I} = \frac{\tau_0}{\tau} = 1 + K_D[Q]} \quad (90)$$

Here, the lifetime in the absence of quencher and the quenching rate constant have been combined into the Stern-Volmer constant for dynamic quenching ( $K_{SV} = K_D$ ).



**Figure 88.** Simulated Stern-Volmer plots of steady-state intensity ( $I$ ) and fluorescence lifetime ( $\tau$ ) in the case of a purely dynamic quenching, b) purely static quenching and c) a mixture thereof. A dynamic quencher affects both parameters to the same degree, whereas a static quencher has no discernable influence on the excited state lifetime. This results in a quadratic relationship between quencher concentration and intensity, but not lifetime, in the last case.

In case of static quenching, the fluorophore and quencher form a dark complex (C) with the association constant  $K_S$ :



The initial fluorophore concentration  $[F_0]$  is conserved according to

$$[F_0] = [F] + [C] \quad (92)$$

Substitution into the mass action law results in

$$K_S = \frac{[F_0] - [F]}{[F][Q]} \quad (93)$$

which rearranges into

$$\frac{[F_0]}{[F]} = 1 + K_S[Q] \quad (94)$$

This equation can also be expressed in terms of the observed fluorescence intensity:

$$\boxed{\frac{I_0}{I} = 1 + K_S[Q]} \quad (95)$$

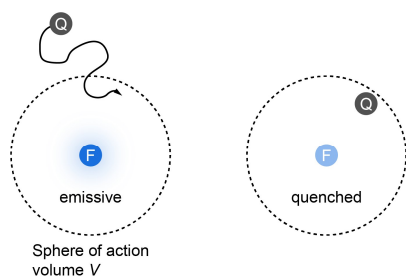
Under common experimental conditions,  $[Q]$  is much larger than  $[F_0]$ , meaning that depletion of the quencher due to complexation can be neglected. As only the fluorescence of unbound fluorophore is observed in this case, the lifetime is not affected by the presence of a static quencher (Figure 88b):

$$\frac{\tau_0}{\tau} = 1 \quad (96)$$

When dynamic and static quenching occur simultaneously, a plot of  $I_0/I$  vs quencher concentration shows a characteristic upward sloping deviation from linearity. The corresponding modified Stern-Volmer equation is quadratic in the quencher concentration:

$$\boxed{\frac{I_0}{I} = (1 + K_D[Q])(1 + K_S[Q])} \quad (97)$$

As before, the lifetime remains unaffected by static quenching and can therefore be used to separate the contributions of both quenching pathways. Formally, this can also be treated as a dynamic correction by taking the ratio of  $I_0/I$  and  $\tau_0/\tau$  (Figure 88c). Under specific circumstances, however, a combination of collisional quenching and the formation of a dark complex is not sufficient to fully explain nonlinear variations of  $I_0/I$  with respect to the quencher concentration. In these cases, a sphere-of-action model, which comes in the form of an exponential factor in the Stern-Volmer equation, can be invoked.<sup>239,270,281</sup> The sphere of action describes a volume within which the presence of a quencher leads to instantaneous deactivation of the fluorophore at the moment of excitation (Figure 89).



**Figure 89.** In the sphere-of-action model, the emission of a fluorophore is deactivated once the quencher comes within a certain distance.

In a given volume  $V$ , the mean number of quencher molecules  $\lambda$  is calculated from the molar concentration:

$$\lambda = VN_A[Q] \quad (98)$$

The probability of finding  $n$  quenchers within the sphere of action follows a Poisson distribution:

$$P(n) = \frac{\lambda^n}{n!} \exp(-\lambda) \quad (99)$$

Fluorescence occurs, when no quencher resides within the sphere, i.e. with a probability of

$$P(0) = \exp(-\lambda) \quad (100)$$

As this is phenomenologically similar to the formation of a dark complex, the contribution to the observed fluorescence intensity is

$$\frac{I_0}{I} = \exp(-VN_A[Q]) \quad (101)$$

The full Stern-Volmer equation for this case is therefore given by

$$\boxed{\frac{I_0}{I} = (1 + K_D[Q])(1 + K_S[Q]) \exp(-VN_A[Q])} \quad (102)$$

In practice,  $K_D$  will be calculated from lifetime data, which have been unaffected by static and sphere-of-action quenching and used as a fixed parameter when fitting intensity data to this equation.



#### 4.1.20.2 Materials

Ribonucleoside monophosphates (NMPs) were purchased from Jena Bioscience or Sigma Aldrich.

Stock solutions of the NMPs (100 mM) were prepared in H<sub>2</sub>O immediately prior to use. Their exact concentration was determined by UV absorbance.

4-Cyanoindole ribonucleoside was used as a stock solution (100 μM) in DMSO.

#### 4.1.20.3 General protocol

Four identical samples of the fluorophore r4Cl in PBS were prepared:

Fluorophore sample:		
	Added volume	Final concentration
100 μM fluorophore in DMSO	5 μL	5 μM
5x PBS pH 7.4	20 μL	1x
H <sub>2</sub> O	75 μL	
Final volume	100 μL	

Separately, samples containing both quencher and fluorophore were prepared:

Fluorophore/quencher sample:		
	Added volume	Final concentration
100 μM fluorophore in DMSO	5 μL	5 μM
100 mM NMP in H <sub>2</sub> O	20 μL	20 mM
5x PBS pH 7.4	20 μL	1x
H <sub>2</sub> O	55 μL	
Final volume	100 μL	

Steady-state emission spectra and fluorescence lifetime data were collected for each of the fluorophore samples. Next, aliquots of the respective quencher sample were added to gradually increase the quencher concentrations and measurements were repeated:

Added volume	Final concentration
μL	mM
0	0
1	0.20
2	0.58
5	1.48
10	3.05
20	5.51
50	9.36

Steady-state measurements were performed using these parameters:

- Ex wavelength: 303 nm
- Em range: 323–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 2.5 nm
- Response: 2 s
- PMT voltage: 480 V

- Data interval: 0.2 nm
- Scan speed: 500 nm/min

Values for  $F_0$  and  $F$  were obtained from the emission intensity at 410 nm.

For lifetime measurements, the following parameters were used:

- Ex wavelength: 318 nm
- Em wavelength: 410 nm
- Em bandwidth: 16 nm
- Time calibration: 0.026 ns/channel
- Number of channels: 4096
- Peak count: 10000

A water sample was used to measure the instrument response function (Em wavelength = 318 nm).

Decay curves were fitted to a monoexponential model as described in section 4.1.11.3

#### 4.1.21 Förster resonance energy transfer

##### 4.1.21.1 Overview

As discussed in section 4.1.10.1, an excited fluorophore can return to its electronic ground state either by emission of a photon (fluorescence) or by a number of non-radiative pathways. The deexcitation of a fluorophore, the donor (D), can be coupled to the excitation of a different one, the acceptor (A), i.e. two fluorophores can be energy transfer partners. The trivial case of energy transfer is the absorption of a donor fluorescence photon by the acceptor, which can occur when there is spectral overlap between the two fluorophores and in essence is a form of static quenching (see section 4.1.20.1). In contrast to the radiative case, non-radiative energy transfer proceeds via distance-dependent resonant coulombic coupling of donor and acceptor transition dipole moments. As such, Förster or fluorescence energy transfer (FRET) is intrinsically linked to the spatial arrangement of the donor–acceptor pair.<sup>282</sup>

According to Förster theory, the rate of energy transfer is given by

$$k_{\text{ET}} = \frac{1}{\tau_{\text{D}}} \left( \frac{R_0^6}{r_{\text{DA}}^6} \right) \quad (103)$$

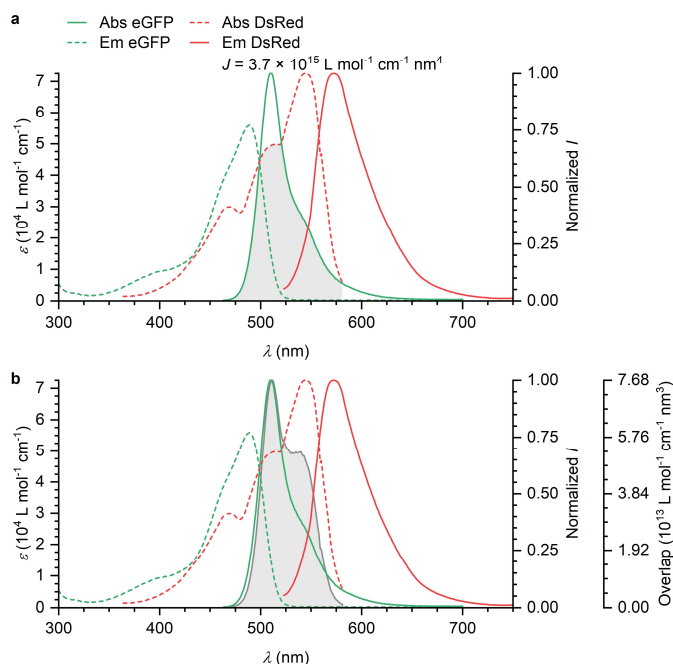
where  $\tau_{\text{D}}$  is the fluorescence lifetime of the donor,  $r_{\text{DA}}$  is the separation between donor acceptor and  $R_0$  is the Förster radius characterizing the DA pair:

$$R_0^6 = \frac{9 \ln 10 \kappa^2 \phi_{\text{D}} J_{\text{DA}}}{128 \pi^5 N_{\text{A}} \eta^4} \quad (104)$$

Here,  $\phi_{\text{D}}$  is the donor fluorescence quantum yield,  $\eta$  is the refractive index of the surrounding medium,  $\kappa^2$  is an orientation factor for the FRET partners and  $J$  is the spectral overlap integral of the DA pair:

$$J_{\text{DA}} = \int_0^{\infty} I_{\text{D}}(\lambda) \varepsilon_{\text{A}}(\lambda) \lambda^4 d\lambda \quad (105)$$

The overlap integral is calculated from the donor emission spectrum ( $I_D$ , area-normalized to 1) and the acceptor absorption spectrum ( $\epsilon_A$ , in extinction coefficient units). Importantly, due to the  $\lambda^4$  term in the above equation, the actual overlap integral is completely different from the simple spectral overlap that is commonly shown for illustrative purposes (Figure 90).



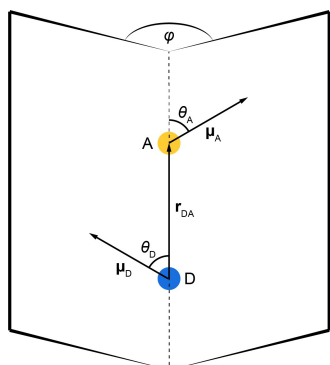
**Figure 90.** a) The emission spectrum of the fluorescent protein eGFP overlaps with the absorption spectrum of DsRed as shown by the shaded area.<sup>76</sup> This is quantified in terms of the spectral overlap integral  $J$ , which is calculated from equation 105 and takes on a value of  $3.7 \times 10^{15} \text{ L mol}^{-1} \text{ cm}^{-1} \text{ nm}^4$  for this particular FRET pair. b) In this graph, the shaded area represents the actual integrand of equation 105, which is skewed to longer wavelengths due to its  $\lambda^4$  dependence. During the calculation of the overlap integral, the donor emission spectrum is area-normalized to 1, i.e. its y-axis unit is formally  $\text{nm}^{-1}$ . Therefore, the integrand has the unit  $\text{L mol}^{-1} \text{ cm}^{-1} \text{ nm}^3$ .

Information about the spatial arrangement of the donor and acceptor transition dipole moments ( $\mu_D$  and  $\mu_A$ , respectively) is encoded in the orientation factor  $\kappa^2$ , which is defined as

$$\kappa^2 = [\mu_A \cdot \mu_D - 3(\mu_A \cdot \mathbf{r}_{DA})(\mu_D \cdot \mathbf{r}_{DA})]^2 \quad (106)$$

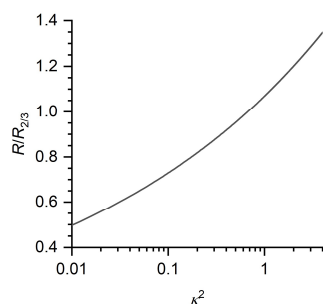
in vector form. For a more visually intuitive understanding (Figure 91), it is instructive to consider this equation in terms of the angles defined by the transition dipole moments and the DA separation vector:

$$\kappa^2 = (\sin \theta_D \sin \theta_A \cos \varphi - 2 \cos \theta_D \cos \theta_A)^2 \quad (107)$$



**Figure 91.** Illustration of the vectors and angles used in the definition of the orientation factor  $\kappa^2$  for a FRET pair.  $\kappa^2$  is calculated according to equations 106 or 107.

It is evident that  $\kappa^2$  will be maximal with a value of 4 when the transition dipole moments are collinear but will be 0 for perpendicular orientations. If the FRET partners can isotropically reorient themselves at their location within the lifetime of the donor excited state, i.e. fast with respect to the timescale of donor emission and energy transfer, a dynamically averaged  $\kappa^2$  of 2/3 results (Figure 92). This is often tacitly assumed when dealing with fluorophores not rigidly embedded into a larger system, e.g. if they are conjugated to a biomacromolecule via a flexible linker, but is not generally valid for more rigid constructs such as the ones described in this thesis.<sup>283</sup>



**Figure 92.** The orientation factor  $\kappa^2$  between the donor and acceptor dipole moments can take on values between 0 and 4; in the dynamic isotropic averaging regime, its value is 2/3. The Förster radius  $R$  of the FRET pair varies with the inverse sixth power of the orientation factor according to equation 104.

Experimentally, in a FRET assay one observes the steady-state emission intensity of fluorescence lifetime of the donor alone and in the presence of the acceptor. From these values, a FRET efficiency ( $E$ ) is calculated:

$$E = 1 - \frac{I_{DA}}{I_D} = 1 - \frac{\tau_{DA}}{\tau_D} \quad (108)$$

Formally, the FRET efficiency is the quantum yield of the energy transfer process. Recalling the discussion in sections 4.1.10.1, it follows that

$$E = \frac{k_{ET}}{k_f + k_{nr} + k_{ET}} = \frac{k_{ET}}{\tau_D^{-1} + k_{ET}} \quad (109)$$

Here, the inverse donor lifetime has been substituted for  $k_f$  and  $k_{nr}$ . This expression can be simplified by inserting equation 103:

$$E = \frac{R_0^6}{R_0^6 + r_{DA}^6} \quad (110)$$

Experimental FRET efficiencies can therefore be translated into the donor–acceptor separation if information about their mutual orientation is available. Lacking such information, a combined analysis is possible by examining the separation as a function of the orientation factor:

$$r_{AD} = \left( \frac{1 - E}{E} \cdot \frac{9 \ln 10 \kappa^2 \phi_D J_{DA}}{128 \pi^5 N_A \eta^4} \right)^{1/6} \quad (111)$$

#### 4.1.21.2 Materials

Ligands were used as stock solutions (20  $\mu\text{M}$ ) in DMSO.

#### 4.1.21.3 General protocol

Samples were prepared as follows:

## Aptamer-only solution:

	Added volume	Final concentration
RNA aptamer	as required	0.5 $\mu\text{M}$
1 M KCl	6.25 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	4 $\mu\text{L}$	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM MgCl <sub>2</sub>	2.5 $\mu\text{L}$	5 mM
DMSO	2.5 $\mu\text{L}$	5%
H <sub>2</sub> O	as required	
Final volume	50 $\mu\text{L}$	

## Aptamer–ligand solution:

	Added volume	Final concentration
RNA aptamer	as required	0.5 $\mu\text{M}$
1 M KCl	6.25 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	4 $\mu\text{L}$	40 mM
Anneal 3 min at 95 °C		
Incubate 20 min at 25 °C		
100 mM MgCl <sub>2</sub>	2.5 $\mu\text{L}$	5 mM
20 $\mu\text{M}$ ligand in DMSO	2.5 $\mu\text{L}$	1 $\mu\text{M}$
H <sub>2</sub> O	as required	
Final volume	50 $\mu\text{L}$	

## Buffer solution:

	Added volume	Final concentration
1 M KCl	6.25 $\mu\text{L}$	125 mM
500 mM HEPES pH 7.5	4 $\mu\text{L}$	40 mM
100 mM MgCl <sub>2</sub>	2.5 $\mu\text{L}$	5 mM
DMSO	2.5 $\mu\text{L}$	5%
H <sub>2</sub> O	34.75	
Final volume	50 $\mu\text{L}$	

Steady-state emission spectra and fluorescence lifetime data were collected for each of the samples. The buffer sample was used for background correction.

Steady-state measurements were performed using these parameters:

- Ex wavelength: 303/413/456 nm
- Em range: 323/433/476–750 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 2.5 nm
- Response: 2 s
- PMT voltage: 600 V
- Data interval: 0.2 nm
- Scan speed: 500 nm/min

For lifetime measurements, the following parameters were used:

- Ex wavelength: 318/408/477 nm
- Em wavelength: 410/542/592 nm
- Em bandwidth: 32/12/6 nm (for the respective wavelength pairs above)
- Time calibration: 0.026 ns/channel
- Number of channels: 4096
- Peak count: 10000

The buffer sample was used to measure the instrument response function (Em wavelength = 318/408/477 nm).

Decay curves were fitted to multiexponential models as described in section 4.1.11.3. FRET efficiencies were calculated from the blank-corrected steady-state intensities at 410 nm ( $I_D$  and  $I_{DA}$ , respectively) or from the amplitude average fluorescence lifetimes ( $\tau_D$  and  $\tau_{DA}$ , respectively) according to equation 108.

## 4.2 CHEMICAL SYNTHESIS

### 4.2.1 General considerations

#### 4.2.1.1 Starting materials, solvents and reagents

All standard chemicals and solvents were purchased from commercial suppliers. Anhydrous solvents were prepared by drying with activated molecular sieves (3 Å) for > 2 d or obtained from a solvent purification system. Organic solvents for optical spectroscopy were purchased from Acros Organics.

Aluminum-backed plates coated with silica gel and a fluorescent indicator were used for thin layer chromatography (TLC). The plates were visualized with UV light. In case of working with HBI derivatives, exposing the plates to ammonia vapor rendered these compounds orange or violet. Compounds containing 4-cyanoindole were additionally stained with Ehrlich's reagent (1% 4-(*N,N*-dimethylamino)benzaldehyde and 5% HCl in EtOH, violet color upon heating), which also rendered DMT-protected nucleosides red. Silica gel 60, 0.032–0.063 mm (230–450 mesh) was used for column chromatography.

#### 4.2.1.2 NMR spectroscopy

NMR spectra were acquired on Bruker Avance III and Avance III HD spectrometers between 300 and 600 MHz as well as Varian Mercury Plus and Inova spectrometers between 300 and 600 MHz.

Chemical shifts ( $\delta$ ) in ppm are referenced to the solvent residual signals ( $^1\text{H}$  and  $^{13}\text{C}$ ) or on the unified scale ( $^{19}\text{F}$  and  $^{31}\text{P}$ ).<sup>284</sup> Coupling constants ( $J$ ) are reported in Hz with the following multiplet designations: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad).

All spectral assignments were verified by additional 2D experiments.

Typical sample concentrations were in the range of 2–5  $\mu\text{mol}$  of analyte in 0.5 mL of solvent.

#### 4.2.1.3 Mass spectrometry

High-resolution ESI mass spectra in positive or negative ion mode were acquired on Bruker micrOTOF, micrOTOF-Q III and maXis instruments.

Monoisotopic masses of oligonucleotides were obtained by charge deconvolution of the raw spectra.

Typical sample concentrations for small molecules were in the range of 1–2  $\mu\text{mol}$  of analyte in 1.0 mL of solvent.

### 4.2.2 Preparation of HBI derivatives

#### 4.2.2.1 General procedure A, imine synthesis with volatile imines

A suspension of the aldehyde (25.0 mmol, 1.00 eq.) and  $\text{MgSO}_4$  (30.0 mmol, 1.20 eq.) in the amine (250 mmol, 10.0 eq.) was stirred ( $\text{CH}_2\text{Cl}_2$  was added if necessary) at ambient temperature for 24 h. Afterwards, the solution was filtered over a Celite plug. The solids were rinsed with  $\text{CH}_2\text{Cl}_2$  (3 $\times$ 20 mL) and the filtrate was evaporated under reduced pressure. The resulting product was usually sufficiently pure for all further reactions.

#### 4.2.2.2 General procedure B, imine synthesis with non-volatile imines

A solution of the aldehyde (20.0 mmol, 1.00 eq.) and the amine (20.0 mmol, 1.00 eq.) in toluene (80 mL) was heated to reflux with a Dean-Stark trap for 16 h. Afterwards, the solvent was completely removed under reduced pressure. The resulting product was usually sufficiently pure for all further reactions.

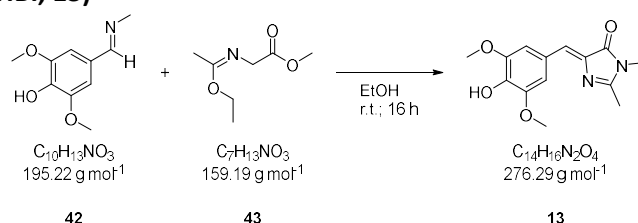
#### 4.2.2.3 General procedure C, cycloaddition reaction

A mixture of the imine (2.00 mmol, 1.00 eq.) and imidate (2.40 mmol, 1.20 eq) in either EtOH or toluene (2 mL) was stirred at ambient temperature or at 120 °C, respectively, until TLC showed completion (usually 24 h, up to 5 d for some compounds). In case the product had precipitated, the solids were collected by filtration and washed with Et<sub>2</sub>O (50 mL). Otherwise, the reaction mixture was evaporated to dryness, and the crude product was purified by column chromatography.

#### 4.2.2.4 General procedure D, Aldol condensation

The HBI derivative (200 μmol, 1.00 eq.), aldehyde (250 μmol, 1.25 eq.) and scandium triflate (30.0 μmol, 15.0 mol%) were dissolved in anhydrous dioxane (1 mL) in a closed vial. The mixture was stirred at 110 °C (oil bath temperature) until TLC showed completion (up to 48 h). Afterwards, the solvent was removed under reduced pressure. Purification of the residue by washing with MeOH or by column chromatography afforded the product.

#### 4.2.2.5 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2,3-dimethyl-3,5-dihydro-4H-imidazol-4-one (DMHBI, **13**)



The title compound **13** was synthesized according to General procedure C on a 5.00 mmol scale. Yellow solid (1.13 g, 4.09 mmol, 82%).

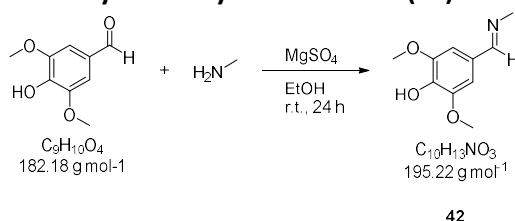
<sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 9.02 (s<sub>br</sub>, 1 H, OH), 7.62 (s, 2 H, Ph-2,6-H), 6.89 (s, 1 H, benzylidene-H), 3.80 (s, 6 H, OCH<sub>3</sub>), 3.09 (s, 3 H, NCH<sub>3</sub>), 2.34 (s, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 169.4 (Imi-C4), 162.2 (Imi-C2), 147.6 (Ph-C3,5), 138.4 (Ph-C4), 136.2 (Imi-C5), 125.7 (benzylidene-C), 124.3 (Ph-C1), 110.1 (Ph-C2,6), 56.0 (OCH<sub>3</sub>), 26.1 (NCH<sub>3</sub>), 15.4 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>14</sub>H<sub>16</sub>N<sub>2</sub>NaO<sub>4</sub>, [M+Na]<sup>+</sup>): 299.1002, found: 299.1004;

TLC (toluene/acetone 1:1): *R<sub>f</sub>* = 0.51.

#### 4.2.2.6 4-Hydroxy-3,5-dimethoxy-*N*-methylbenzaldimine (**42**)



The title compound **42** was synthesized according to General procedure A on a 25.0 mmol scale using a 33% solution of MeNH<sub>2</sub> in EtOH. Pale yellow solid (4.05 g, 20.8 mmol, 83%) that was sufficiently pure for all further reactions. For analysis, the crude product was recrystallized from EtOH.

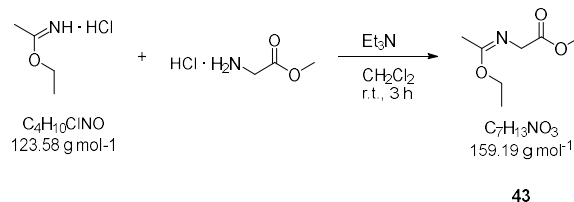
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.16 (q, *J* = 1.6 Hz, 1 H, CHN), 6.97 (s, 2 H, Ph-2,6-H), 3.92 (s, 6 H, OCH<sub>3</sub>), 3.49 (d, *J* = 1.6 Hz, 3 H, NCH<sub>3</sub>);



$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 162.2 (CHN), 147.6 (Ph-C3,5), 137.9 (Ph-C4), 127.5 (Ph-C1), 104.9 (Ph-C2,6), 56.5 ( $\text{OCH}_3$ ), 48.1 ( $\text{NCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{10}\text{H}_{14}\text{NO}_3$ ,  $[\text{M}+\text{H}]^+$ ): 196.0968, found: 196.0973.

#### 4.2.2.7 Methyl (Z)-2-((1-ethoxyethylidene)amino)acetate (**43**)

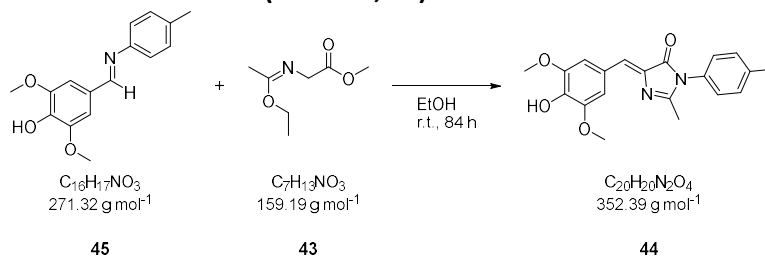


To a suspension of ethyl acetimidate hydrochloride (4.63 g, 37.5 mmol, 1.00 eq.) and methyl glycinate hydrochloride (4.71 g, 37.5 mmol, 1.00 eq.) in dry  $\text{CH}_2\text{Cl}_2$  (150 mL) was added  $\text{Et}_3\text{N}$  (5.2 mL, 37.5 mmol, 1.00 eq.). The resulting mixture was stirred at ambient temperature for 3 h. Afterwards, it was washed with  $\text{H}_2\text{O}$  (2 $\times$ 150 mL) and brine (150 mL) and the organic phase was dried over  $\text{MgSO}_4$ . Evaporation of the solvent under reduced pressure afforded the title compound **43** as a colorless liquid (5.08 g, 31.9 mmol, 85%). The product can be stored under an inert atmosphere at  $-20^\circ\text{C}$  for several weeks without decomposition.

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 4.10 (q,  $J$  = 7.1 Hz, 2 H,  $\text{OCH}_2\text{CH}_3$ ), 4.05 (s, 2 H,  $\text{NCH}_2$ ), 3.73 (s, 3 H,  $\text{OCH}_3$ ), 1.87 (s, 3 H,  $\text{CCH}_3$ ), 1.26 (t,  $J$  = 7.1 Hz, 3 H,  $\text{OCH}_2\text{CH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_7\text{H}_{13}\text{NNaO}_3$ ,  $[\text{M}+\text{Na}]^+$ ): 182.0788, found: 182.0788.

#### 4.2.2.8 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(4-methylphenyl)-3,5-dihydro-4H-imidazol-4-one (**44**)



The title compound **44** was synthesized according to General procedure C on a 1.25 mmol scale. After purification by column chromatography (Hex/EtOAc 70:30–25:75 + 1% AcOH) it was obtained as a yellow solid (237 mg, 0.67 mmol, 54%). Analytical data for a side product that was isolated during column chromatography are given below (see section 4.2.2.10).

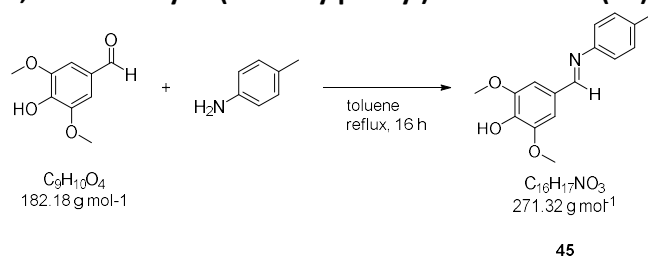
$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.56 (s, 2 H, Ph-2,6-H), 7.33–7.29 (m, 2 H, NAr-3,5-H), 7.14–7.11 (m, 2 H, NAr-2,6-H), 7.10 (q,  $J$  = 0.6 Hz, 1 H, benzylidene-H), 5.88 (s, 1 H, OH), 3.97 (s, 6 H,  $\text{OCH}_3$ ), 2.41 (s, 3 H, NAr- $\text{CH}_3$ ), 2.25 (d,  $J$  = 0.6 Hz, 3 H,  $\text{CCH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 170.2 (Imi-C4), 161.0 (Imi-C2), 147.2 (Ph-C3,5), 139.0 (NAr-C1), 137.5 (Ph-C4), 136.8 (Imi-C5), 131.1 (NAr-C4), 130.5 (NAr-C3,5), 128.5 (benzylidene-C), 127.3 (NAr-C2,6), 126.0 (Ph-C1), 109.5 (Ph-C2,6), 56.5 ( $\text{OCH}_3$ ), 21.4 (NAr- $\text{CH}_3$ ), 16.7 ( $\text{CCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{20}\text{H}_{21}\text{N}_2\text{O}_4$ ,  $[\text{M}+\text{H}]^+$ ): 353.1496, found: 353.1494;

TLC (Hex/EtOAc 60:40 + 1% AcOH):  $R_f$  = 0.30.

#### 4.2.2.9 4-Hydroxy-3,5-dimethoxy-*N*-(4-methylphenyl)benzaldimine (45)



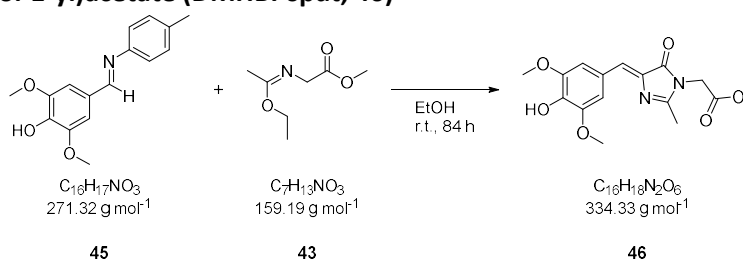
The title compound **45** was synthesized according to General procedure B on a 25.0 mmol scale. Yellow solid (6.78 g, 25.0 mmol, > 99%).

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.34 (s, 1 H, CHN), 7.21–7.17 (m, 2 H, NAr-3,5-H), 7.17 (s, 2 H, Ph-2,6-H), 7.15–7.10 (m, 2 H, NAr-2,6-H), 5.90 (s<sub>br</sub>, 1 H, OH), 3.97 (s, 6 H, OCH<sub>3</sub>), 2.37 (s, 3 H, NAr-CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>): δ (ppm) = 159.4 (CHN), 149.6 (NAr-C1), 147.4 (Ph-C3,5), 138.0 (Ph-C4), 135.6 (NAr-C4), 129.9 (NAr-C3,5), 128.2 (Ph-C1), 120.9 (NAr-C2,6), 105.7 (Ph-C2,6), 56.6 (OCH<sub>3</sub>), 21.1 (NAr-CH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>16</sub>H<sub>18</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 272.1281, found: 272.1283.

#### 4.2.2.10 Methyl (Z)-4-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-5-oxo-4,5-dihydro-1*H*-imidazol-1-yl)acetate (DMHBI-spdt, 46)



The title compound **46** was obtained as a side product during the synthesis of DMHBTI (**44**, see section 4.2.2.8). Yellow solid (102 mg, 0.31 mmol, 24%).

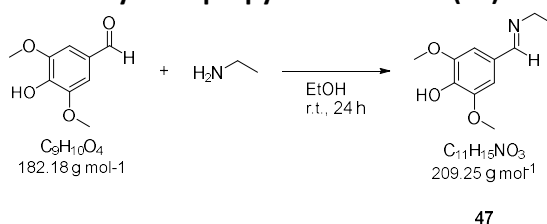
<sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD): δ (ppm) = 7.53 (s, 2 H, Ph-2,6-H), 7.01 (s, 1 H, benzylidene-H), 4.51 (s, 2 H, NCH<sub>2</sub>), 3.90 (s, 6 H, Ph-3,5-OCH<sub>3</sub>), 3.79 (s, 3 H, COCH<sub>3</sub>), 2.34 (s, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CD<sub>3</sub>OD): δ (ppm) = 171.7 (Imi-C5), 170.1 (COCH<sub>3</sub>), 162.5 (Imi-C2), 149.2 (Ph-C3,5), 140.3 (Ph-C4), 136.9 (Imi-C4), 130.2 (benzylidene-C), 126.2 (Ph-C1), 111.3 (Ph-C2,6), 56.8 (Ph-3,5-OCH<sub>3</sub>), 53.2 (COCH<sub>3</sub>), 42.2 (NCH<sub>2</sub>), 15.3 (CCH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>16</sub>H<sub>18</sub>N<sub>2</sub>NaO<sub>6</sub>, [M+Na]<sup>+</sup>): 357.1057, found: 357.1057;

TLC (Hex/EtOAc 60:40 + 1% AcOH): *R<sub>f</sub>* = 0.09.

#### 4.2.2.11 4-Hydroxy-3,5-dimethoxy-*N*-isopropylbenzaldimine (47)



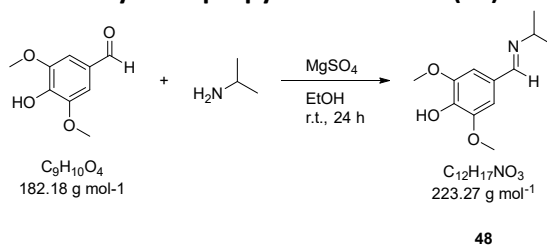
The title compound **47** was synthesized according to General procedure A on a 5.00 mmol scale using a 70% solution of EtNH<sub>2</sub> in H<sub>2</sub>O and without the addition of MgSO<sub>4</sub>. Pale yellow solid (816 mg, 4.50 mmol, 90%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.15 (t, *J* = 1.3 Hz, 1 H, CHN), 6.99 (s, 2 H, Ph-2,6-H), 5.83 (s<sub>br</sub>, 1 H, OH), 3.92 (s, 6 H, OCH<sub>3</sub>), 3.62 (qd, *J* = 7.3, 1.3 Hz, 2 H, CH<sub>2</sub>CH<sub>3</sub>), 1.29 (t, *J* = 7.3 Hz, 3 H, CH<sub>2</sub>CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 160.2 (CHN), 147.2 (Ph-C3,5), 137.4 (Ph-C4), 127.7 (Ph-C1), 104.9 (Ph-C2,6), 56.4 (OCH<sub>3</sub>), 55.5 (CH<sub>2</sub>CH<sub>3</sub>), 16.4 (CH<sub>2</sub>CH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>11</sub>H<sub>16</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 210.1125, found: 210.1122.

#### 4.2.2.12 4-Hydroxy-3,5-dimethoxy-*N*-isopropylbenzalimine (**48**)



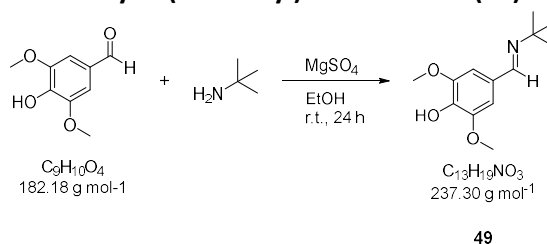
The title compound **48** was synthesized according to General procedure A on a 5.00 mmol scale. Orange foam (1.07 g, 4.79 mmol, 96%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.17 (s, 1 H, CHN), 6.98 (s, 2 H, Ph-2,6-H), 3.92 (s, 6 H, OCH<sub>3</sub>), 3.51 (hept, *J* = 6.4 Hz, 1 H, CH(CH<sub>3</sub>)<sub>2</sub>), 1.25 (d, *J* = 6.4 Hz, 6 H, CH(CH<sub>3</sub>)<sub>2</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 158.2 (CHN), 147.4 (Ph-C3,5), 137.4 (Ph-C4), 128.0 (Ph-C1), 105.0 (Ph-C2,6), 61.6 (CH(CH<sub>3</sub>)<sub>2</sub>), 56.5 (OCH<sub>3</sub>), 24.3 (CH(CH<sub>3</sub>)<sub>2</sub>);

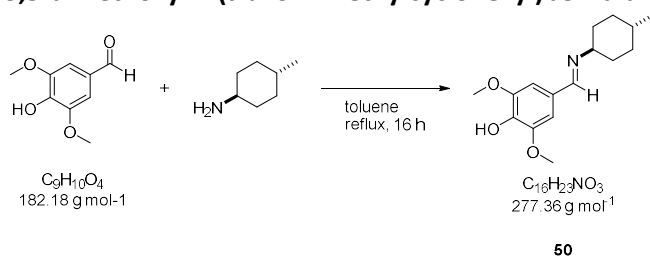
HR-MS (ESI+): *m/z* calc. (C<sub>12</sub>H<sub>18</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 224.1281, found: 224.1282.

#### 4.2.2.13 4-Hydroxy-3,5-dimethoxy-*N*-(*tert*-butyl)benzalimine (**49**)



The title compound **49** was synthesized according to General procedure A on a 12.5 mmol scale. Due to its low stability the crude product was used in the next step without further characterization.

#### 4.2.2.14 4-Hydroxy-3,5-dimethoxy-*N*-(*trans*-4-methylcyclohexyl)benzalimine (**50**)



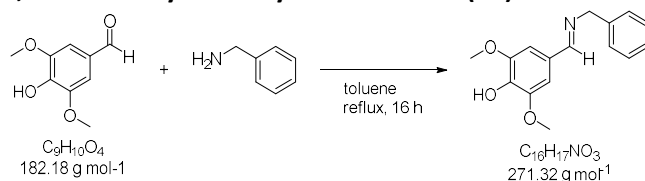
The title compound **50** was synthesized according to General procedure B on a 5.00 mmol scale. Yellow foam (1.39 g, 5.00 mmol, > 99%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.19 (s, 1 H, CHN), 6.98 (s, 2 H, Ph-2,6-H), 3.93 (s, 6 H, OCH<sub>3</sub>), 3.20–3.03 (m, 1 H, Cy-1-H), 1.83–1.70 (m, 4 H, Cy-2,6-H, Cy-3,5-H), 1.70–1.54 (m, 2 H, Cy-2,6-H), 1.52–1.33 (m, 1 H, Cy-4-H), 1.07 (td,  $J$  = 12.3, 3.6 Hz, 2 H, Cy-3,5-H), 0.92 (d,  $J$  = 6.5 Hz, 3 H, Cy-CH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (75 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 158.6 (CHN), 147.4 (Ph-C3,5), 128.4 (Ph-C1), 105.1 (Ph-C2,6), 70.0 (Cy-C1), 56.6 (OCH<sub>3</sub>), 34.4 (Cy-C2,6), 33.9 (Cy-C2,6, Cy-C3,5), 32.1 (Cy-C4), 22.6 (Cy-CH<sub>3</sub>);

**HR-MS** (ESI+):  $m/z$  calc. (C<sub>16</sub>H<sub>24</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 278.1754, found: 278.1787.

#### 4.2.2.15 4-Hydroxy-3,5-dimethoxy-*N*-benzylbenzaldimine (**51**)



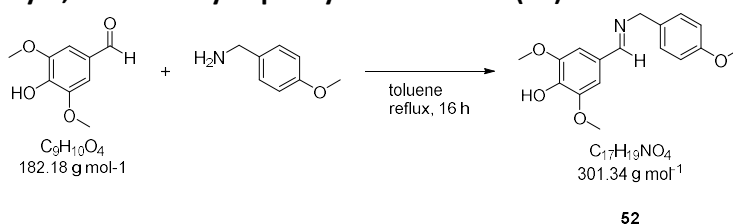
The title compound **51** was synthesized according to General procedure B on a 25.0 mmol scale. A first batch of the product crystallized from the reaction mixture after cooling to ambient temperature. It was isolated by filtration; the remainder was obtained by evaporation of the filtrate. Pale yellow solid (6.78 g, 25.0 mmol, > 99%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.26 (t,  $J$  = 1.4 Hz, 1 H, CHN), 7.42–7.20 (m, 5 H, Bn-2,6H, Bn-3,5-H, Bn-4-H), 7.05 (s, 2 H, Ph-2,6-H), 5.84 (s<sub>br</sub>, 1 H, OH), 4.81 (d,  $J$  = 1.3 Hz, 2 H, CH<sub>2</sub>), 3.93 (s, 6 H, OCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 161.7 (CHN), 147.2 (Ph-C3,5), 139.4 (Bn-C1), 137.5 (Ph-C4), 128.6 (Bn-C3,5), 128.1 (Bn-C2,6), 127.8 (Ph-C1), 127.1 (Bn-C4), 105.3 (Ph-C2,6), 65.0 (CH<sub>2</sub>), 56.6 (OCH<sub>3</sub>);

**HR-MS** (ESI+):  $m/z$  calc. (C<sub>16</sub>H<sub>18</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 272.1281, found: 272.1281.

#### 4.2.2.16 4-Hydroxy-3,5-dimethoxy-*N*-phenylbenzaldimine (**52**)



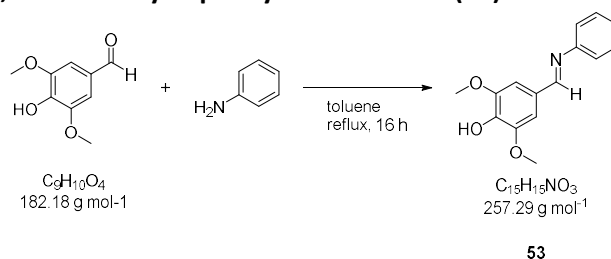
The title compound **52** was synthesized according to General procedure B on a 12.5 mmol scale. Pale yellow solid (3.77 g, 12.5 mmol, > 99%).

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.22 (t,  $J$  = 1.4 Hz, 1 H, CHN), 7.26–7.22 (m, 2 H, NAr-2,6-H), 7.03 (s, 2 H, Ph-2,6-H), 6.90–6.87 (m, 2 H, NAr-3,5-H), 4.74 (d,  $J$  = 1.3 Hz, 2 H, CH<sub>2</sub>), 3.89 (s, 6 H, Ph-OCH<sub>3</sub>), 3.79 (s, 3 H, NAr-OCH<sub>3</sub>);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 161.3 (CHN), 158.7 (NAr-C4), 147.3 (Ph-C3,5), 137.6 (Ph-C4), 131.4 (NAr-C1), 129.3 (NAr-C2,6), 127.7 (Ph-C1), 114.0 (NAr-C3,5), 105.2 (Ph-C2,6), 64.3 ( $\text{CH}_2$ ), 56.5 (Ph- $\text{OCH}_3$ ), 55.4 (NAr- $\text{OCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{17}\text{H}_{20}\text{NO}_4$ ,  $[\text{M}+\text{H}]^+$ ): 302.1387, found: 302.1392.

#### 4.2.2.17 4-Hydroxy-3,5-dimethoxy-*N*-phenylbenzalimine (53)



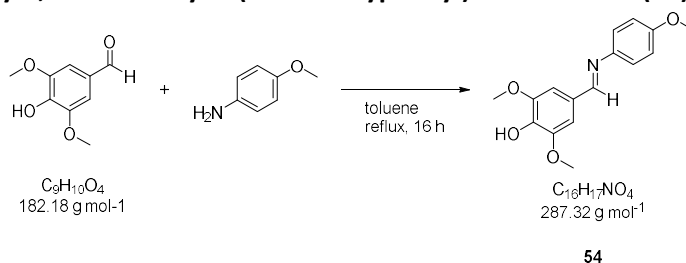
The title compound **53** was synthesized according to General procedure B on a 12.5 mmol scale. Dark yellow solid (3.11 g, 12.1 mmol, 97%).

$^1\text{H}$  NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 9.11 (s, 1 H, OH), 8.44 (s, 1 H, CHN), 7.43–7.36 (m, 2 H, NAr-3,5-H), 7.23 (s, 2 H, Ph-2,6-H), 7.23–7.17 (m, 3H, NAr-2,6-H, NAr-4-H), 3.84 (s, 6 H,  $\text{OCH}_3$ ).

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 160.5 (CHN), 151.9 (NAr-C1), 148.0 (Ph-C3,5), 139.2 (Ph-C4), 129.2 (NAr-C3,5), 126.6 (Ph-C1), 125.4 (NAr-C4), 120.9 (NAr-C2,6), 106.2 (Ph-C2,6), 56.0 ( $\text{OCH}_3$ ), 9, 10.;

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{15}\text{H}_{16}\text{NO}_3$ ,  $[\text{M}+\text{H}]^+$ ): 258.11247, found: 258.11292.

#### 4.2.2.18 4-Hydroxy-3,5-dimethoxy-*N*-(4-methoxyphenyl)benzalimine (54)



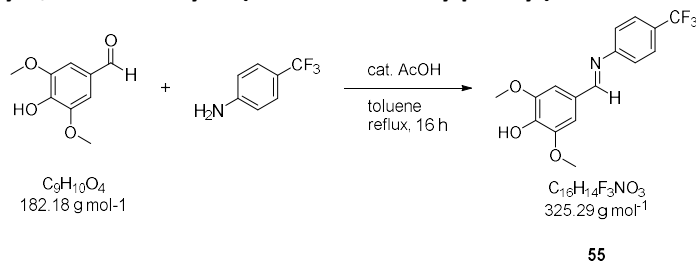
The title compound **54** was synthesized according to General procedure B on a 12.5 mmol scale. Off-white solid (3.36 g, 11.7 mmol, 94%).

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.35 (s, 1 H, CHN), 7.25–7.18 (m, 2 H, NAr-2,6-H), 7.16 (s, 2 H, Ph-2,6-H), 6.96–6.89 (m, 2 H, NAr-3,5-H), 5.93 ( $s_{br}$ , 1 H, OH), 3.97 (s, 6 H, Ph- $\text{OCH}_3$ ), 3.83 (s, 3 H, NAr- $\text{OCH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 158.2 (CHN), 158.1 (NAr-C4), 147.3 (Ph-C3,5), 145.1 (NAr-C1), 137.9 (Ph-C4), 128.2 (Ph-C1), 122.1 (NAr-C2,6), 114.5 (NAr-C3,5), 105.6 (Ph-C2,6), 56.6 (Ph- $\text{OCH}_3$ ), 55.7 (NAr- $\text{OCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{16}\text{H}_{18}\text{NO}_4$ ,  $[\text{M}+\text{H}]^+$ ): 288.1230, found: 288.1237.

#### 4.2.2.19 4-Hydroxy-3,5-dimethoxy-*N*-(4-trifluoromethylphenyl)benzalimine (55)



A solution of 4-hydroxy-3,5-dimethoxybenzaldehyde (911 mg, 5.00 mmol, 1.00 eq.) and 4-trifluoromethylaniline (806 mg, 5.00 mmol, 1.00 eq.) in toluene (20 mL) was treated with AcOH (4 drops) and heated to reflux for 16 h. After cooling to ambient temperature, residual solids were removed by filtration over Celite and washed successively with toluene, CHCl<sub>3</sub> and MeOH (20 mL each). The filtrate was evaporated to afford the crude product **55** (1.60 g, 4.92 mmol, 98%) as an off-white solid sufficiently pure for further reactions. An analytically pure sample was prepared by recrystallization from heptane/toluene (5:1, 1.25 g in 30 mL).

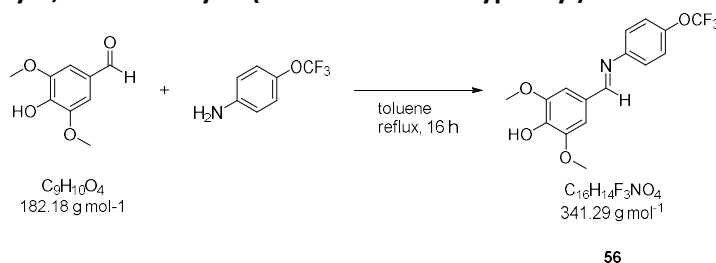
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.29 (s, 1 H, CHN), 7.63 (m, 2 H, NAr-3,5-H), 7.24 (m, 2 H, NAr-2,6-H), 7.18 (s, 2 H, Ph-2,6-H), 6.14 (s<sub>br</sub>, 1 H, OH), 3.95 (s, 6 H, OCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 161.4 (CHN), 155.3 (NAr-C1), 147.4 (Ph-C3,5), 138.7 (Ph-C4), 127.5 (q, *J* = 32.5 Hz, NAr-C4), 127.5 (Ph-C1), 126.4 (q, *J* = 3.8 Hz, NAr-C3,5), 124.4 (q, *J* = 271.3 Hz, CF<sub>3</sub>), 121.1 (NAr-C2,6), 106.1 (Ph-C2,6), 56.7 (OCH<sub>3</sub>);

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>): δ (ppm) = -62.0 (CF<sub>3</sub>);

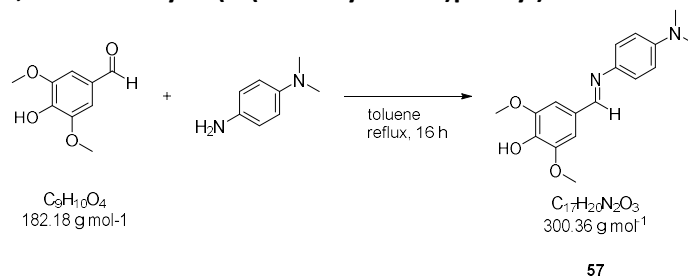
HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>16</sub>H<sub>15</sub>F<sub>3</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 326.0999, found: 326.1013.

#### 4.2.2.20 4-Hydroxy-3,5-dimethoxy-*N*-(4-trifluoromethoxyphenyl)benzalimine (56)



The title compound **56** was synthesized according to General procedure B on a 5.00 mmol scale. Brown solid (1.66 g, 4.86 mmol, 97%). Due to its low stability the crude product was used in the next step without further characterization.

#### 4.2.2.21 Hydroxy-3,5-dimethoxy-*N*-(4-(dimethylamino)phenyl)benzalimine (57)



The title compound **57** was synthesized according to General procedure B on a 10.0 mmol scale. For purification, the crude product was taken up in hot toluene (30 mL). Addition of pentane

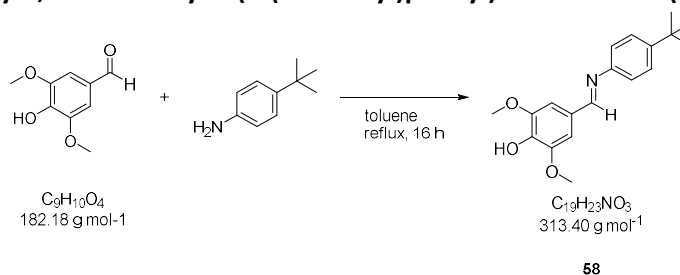
(30 mL) resulted in the formation of a yellow suspension from which a dark solid separated. The solid was removed by decantation and the supernatant was treated again with pentane (60 mL). Decantation and addition of pentane (60 mL) were repeated once more, resulting in the formation of yellow needles, which were collected by filtration and washed with a small amount of pentane to obtain compound **57** in a pure form (2.03 g, 6.77 mmol, 68%). A second batch of the product was obtained from the dark solid by dissolving it in hot toluene (30 mL), cooling to ambient temperature and slowly adding the solution to hexane (150 mL) while stirring. This resulted in a dark yellow precipitate that was collected and washed with a small amount of hexane.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.38 (s, 1 H, CHN), 7.30–7.11 (m, 4 H, NAr-2,6-H, Ph-2,6-H), 6.81–6.71 (m, 2 H, NAr-3,5-H), 5.80 ( $s_{\text{br}}$ , 1 H, OH), 3.98 (s, 6 H,  $\text{OCH}_3$ ), 2.98 (s, 6 H,  $\text{NCH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 155.6 (CHN), 149.6\* (NAr-C4), 147.4 (Ph-C3,5), 141.2\* (NAr-C1), 138.3\* (Ph-C4), 128.6\* (Ph-C1), 122.3 (NAr-C2,6), 113.1 (NAr-C3,5), 105.4 (Ph-C2,6), 56.6 ( $\text{OCH}_3$ ), 40.9 ( $\text{NCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{17}\text{H}_{21}\text{N}_2\text{O}_3$ ,  $[\text{M}+\text{H}]^+$ ): 301.1547, found: 301.1551.

#### 4.2.2.22 4-Hydroxy-3,5-dimethoxy-*N*-(4-(*tert*-butyl)phenyl)benzaldimine (**58**)



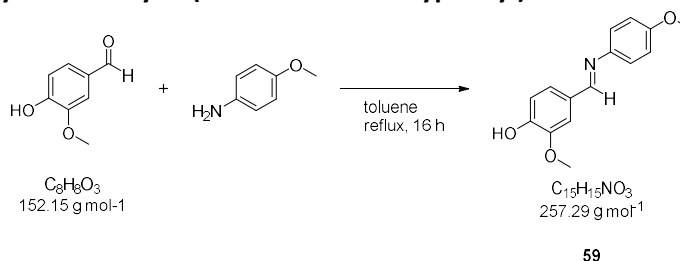
The title compound **58** was synthesized according to General procedure B on a 7.50 mmol scale. Off-white solid (2.35 g, 7.50 mmol, > 99%).

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.35 (s, 1H, CHN), 7.47–7.35 (m, 2 H, NAr-2,6-H), 7.21–7.11 (m, 4 H, Ph-2,6-H, NAr-3,5-H), 5.92 ( $s_{\text{br}}$ , 1 H, OH), 3.97 (s, 6 H,  $\text{OCH}_3$ ), 1.35 (s, 9 H,  $\text{C}(\text{CH}_3)_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 159.4 (CHN), 149.5 (NAr-C4), 149.0 (NAr-C1), 147.4 (Ph-C3,5), 138.1 (Ph-C4), 128.2 (Ph-C1), 126.1 (NAr-C2,6), 120.6 (NAr-C3,5), 105.8 (Ph-C2,6), 56.6 ( $\text{OCH}_3$ ), 34.6 ( $\text{C}(\text{CH}_3)_3$ ), 31.6 ( $\text{C}(\text{CH}_3)_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{19}\text{H}_{24}\text{NO}_3$ ,  $[\text{M}+\text{H}]^+$ ): 314.1751, found: 314.1761.

#### 4.2.2.23 4-Hydroxy-3-methoxy-*N*-(4-trifluoromethoxyphenyl)benzaldimine (**59**)



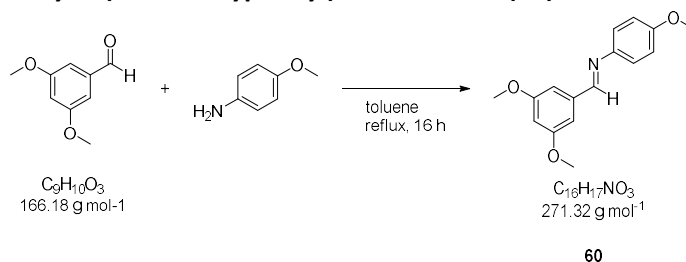
The title compound **59** was synthesized according to General procedure B on a 12.5 mmol scale. It crystallized from the reaction mixture after cooling to ambient temperature. Off-white solid (2.99 g, 11.6 mmol, 93%).

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.66 (s<sub>br</sub>, 1 H, OH), 8.45 (s, 1 H, CHN), 7.50 (d, *J* = 1.9 Hz, 1 H, Ph-2-H), 7.30 (dd, *J* = 8.2, 1.9 Hz, 1 H, Ph-6-H), 7.24–7.20 (m, 2 H, NAr-3,5-H), 6.98–6.92 (m, 2 H, NAr-2,6-H), 6.88 (d, *J* = 8.2 Hz, 1 H, Ph-5-H), 3.84 (s, 3 H, Ph-OCH<sub>3</sub>), 3.76 (s, 3 H, NAr-OCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 158.1 (CHN), 157.4 (NAr-C1), 149.8 (Ph-C4), 147.9 (Ph-C3), 144.7 (NAr-C4), 128.1 (Ph-C1), 123.7 (Ph-C6), 122.1 (NAr-C3,5), 115.3 (Ph-C5), 114.4 (NAr-C2,6), 110.2 (Ph-C2), 55.5 (Ph-OCH<sub>3</sub>), 55.3, (NAr-OCH<sub>3</sub>);

**HR-MS** (ESI+): *m/z* calc. (C<sub>15</sub>H<sub>16</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 258.11247, found: 258.11302.

#### 4.2.2.24 3,5-dimethoxy-*N*-(4-methoxyphenyl)benzalimine (**60**)



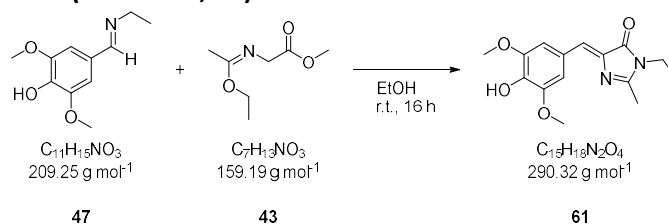
The title compound **60** was synthesized according to General procedure B on a 5.00 mmol scale. Light brown solid (1.36 g, 5.00 mmol, > 99%).

**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.39 (d, *J* = 0.4 Hz, 1 H, CHN), 7.30–7.18 (m, 2 H, NAr-2,6-H), 7.06 (dd, *J* = 2.3, 0.4 Hz, 2 H, Ph-2,6-H), 6.97–6.90 (m, 2 H, NAr-3,5-H), 6.57 (t, *J* = 2.3 Hz, 1 H, Ph-4-H), 3.86 (s, 6 H, Ph-OCH<sub>3</sub>), 3.83 (s, 3 H, NAr-OCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (75 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 161.2 (Ph-C3,5), 158.5 (NAr-C4), 158.4 (CHN), 144.8 (NAr-C1), 138.6 (Ph-C1), 122.4 (NAr-C2,6), 114.5 (NAr-C3,5), 106.4 (Ph-C2,6), 104.1 (Ph-C4), 55.7 (Ph-OCH<sub>3</sub>), 55.6 (NAr-OCH<sub>3</sub>);

**HR-MS** (ESI+): *m/z* calc. (C<sub>16</sub>H<sub>18</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 272.1281, found: 272.1284.

#### 4.2.2.25 (*Z*)-3-Ethyl-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4*H*-imidazol-4-one (**61**)



The title compound **61** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–96:4 + 1% AcOH) it was obtained as a yellow solid (202 mg, 697  $\mu$ mol, 35%).

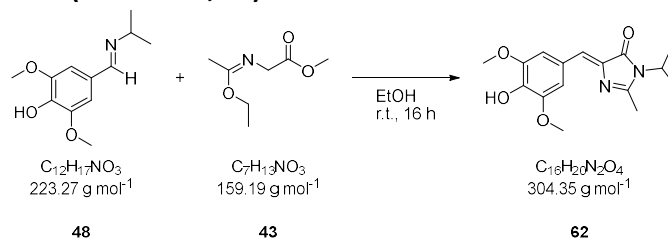
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.50 (s, 2 H, Ph-2,6-H), 7.01 (q, *J* = 0.6 Hz, 1H, benzylidene-H), 5.91 (s<sub>br</sub>, 1 H, OH), 3.95 (s, 6 H, OCH<sub>3</sub>), 3.67 (q, *J* = 7.3 Hz, 2 H, CH<sub>2</sub>CH<sub>3</sub>), 2.40 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>), 1.25 (t, *J* = 7.3 Hz, 3 H, CH<sub>2</sub>CH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 170.4 (Imi-C4), 161.4 (Imi-C2), 147.2 (Ph-C3,5), 137.5 (Ph-C4), 136.7 (Imi-C5), 128.0 (benzylidene-C), 125.9 (Ph-C1), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 35.5 (CH<sub>2</sub>CH<sub>3</sub>), 15.8 (CCH<sub>3</sub>), 14.8 (CH<sub>2</sub>CH<sub>3</sub>);



HR-MS (ESI+):  $m/z$  calc. ( $C_{15}H_{19}N_2O_4$ ,  $[M+H]^+$ ): 291.13353, found: 291.13393.

**4.2.2.26 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-isopropyl-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-*i*Pr, 62)**



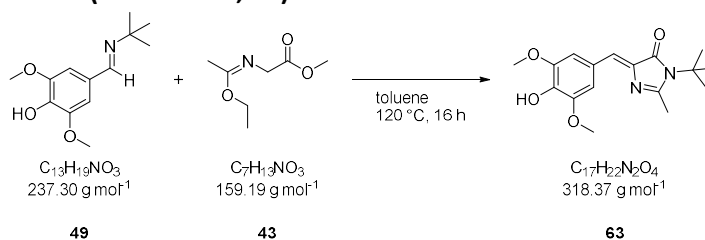
The title compound **62** was synthesized according to General procedure C on a 2.00 mmol scale. Brown solid (563 mg, 1.86 mmol, 93%).

$^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.50–7.47 (m, 2 H, Ph-2,6-H), 6.95 (q,  $J$  = 0.6 Hz, 1 H, benzylidene-H), 5.94 ( $s_{\text{br}}$ , 1H, OH), 4.26 (hept,  $J$  = 6.9 Hz, 1 H,  $\text{CH}(\text{CH}_3)_2$ ), 3.94 (s, 6 H,  $\text{OCH}_3$ ), 2.43 (d,  $J$  = 0.6 Hz, 3 H,  $\text{CCH}_3$ ), 1.47 (d,  $J$  = 7.0 Hz, 6 H,  $\text{CH}(\text{CH}_3)_2$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 170.7 (Imi-C4), 161.7 (Imi-C2), 147.2 (Ph-C3,5), 137.4 (Ph-C4), 127.5 (benzylidene-C), 126.0 (Ph-C1), 109.4 (Ph-C2,6), 56.5 ( $\text{OCH}_3$ ), 45.5 ( $\text{CH}(\text{CH}_3)_2$ ), 20.7 ( $\text{CH}(\text{CH}_3)_2$ ), 17.2 ( $\text{CCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $C_{16}H_{21}N_2O_4$ ,  $[M+H]^+$ ): 305.1496, found: 305.1492.

**4.2.2.27 (Z)-3-(*tert*-Butyl)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-*t*Bu, 63)**



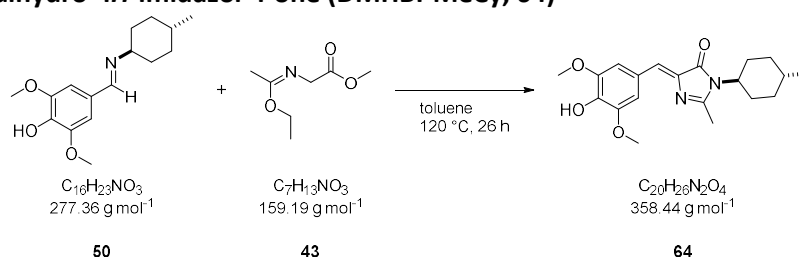
The title compound **63** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (Hex/EtOAc 100:0–0:100 + 1% AcOH) the crude product was obtained as a yellow solid. Residual impurities were removed by sublimation (120 °C, 0.001 mbar). Yellow foam (118 mg, 0.37 mmol, 18%).

$^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.49 (s, 2 H, Ph-2,6-H), 6.89 (s, 1 H, benzylidene-H), 5.85 (s, 1 H, OH), 3.94 (s, 6 H,  $\text{OCH}_3$ ), 2.55 (s, 3 H,  $\text{CCH}_3$ ), 1.63 (s, 9 H,  $\text{C}(\text{CH}_3)_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 171.9 (Imi-C4), 162.4 (Imi-C2), 147.1 (Ph-C3,5), 137.2 (Ph-C4, Imi-C5), 126.8 (benzylidene-C), 126.2 (Ph-C1), 109.3 (Ph-C2,6), 57.7 ( $\text{C}(\text{CH}_3)_3$ ), 56.5 ( $\text{OCH}_3$ ), 29.9 ( $\text{C}(\text{CH}_3)_3$ ), 22.2 ( $\text{CCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $C_{17}H_{23}N_2O_4$ ,  $[M+H]^+$ ): 319.1652, found: 319.1654.

**4.2.2.28 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(trans-4-methylcyclohexyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-MeCy, 64)**



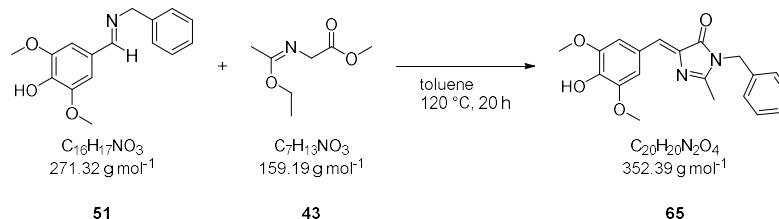
The title compound **64** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 99:1–96:4 + 1% AcOH) it was obtained as a light brown solid (286 mg, 799 μmol, 40%).

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.49 (s, 2 H, Ph-2,6-H), 6.94 (s, 1 H, benzylidene-H), 5.86 (s, 1 H, OH), 3.95 (s, 6 H, OCH<sub>3</sub>), 3.76 (tt, *J* = 12.4, 4.0 Hz, 1 H, Cy-1-H), 2.43 (s, 3 H, CCH<sub>3</sub>), 2.18 (qd, *J* = 12.9, 3.6 Hz, 2 H, Cy-2,6-H<sup>ax</sup>), 1.86–1.80 (m, 2 H, Cy-3,5-H<sup>eq</sup>), 1.73 (dd, *J* = 13.4, 3.7 Hz, 2 H, Cy-2,6-H<sup>eq</sup>), 1.54–1.41 (m, 1 H, Cy-4-H), 1.06 (qd, *J* = 13.2, 3.5 Hz, 2 H, Cy-3,5-H<sup>ax</sup>), 0.93 (d, *J* = 6.5 Hz, 3 H, Cy-CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (150 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.8 (Imi-C4), 161.9 (Imi-C2), 147.1 (Ph-C3,5), 137.3 (Ph-C4), 127.4 (Ph-C1, benzylidene-C), 126.0 (Imi-C5), 109.4 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 53.9 (Cy-C1), 34.7 (Cy-C3,5), 31.7 (Cy-C4), 30.2 (Cy-C2,6)', 22.3 (Cy-CH<sub>3</sub>), 17.4 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>20</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 359.1965, found: 359.1964.

**4.2.2.29 (Z)-3-Benzyl-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-Bn, 65)**



The title compound **65** was synthesized according to General procedure C on a 2.50 mmol scale. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–96:4 + 1% AcOH) it was obtained as a yellow solid (678 mg, 1.92 mmol, 77%).

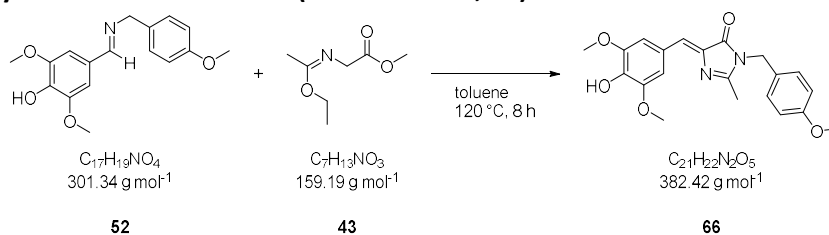
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.52 (s, 2 H, Ph-2,6-H), 7.36–7.30 (m, 2 H, Bn-3,5-H), 7.30–7.26 (m, 1 H, Bn-4-H), 7.24–7.20 (m, 2 H, Bn-2,6-H), 7.09 (q, *J* = 0.6 Hz, 1 H, benzylidene-H), 5.91 (s, 1 H, OH), 4.83 (s, 2 H, CH<sub>2</sub>), 3.95 (s, 6 H, OCH<sub>3</sub>), 2.25 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.7 (Imi-C4), 161.4 (Imi-C2), 147.2 (Ph-C3,5), 137.5 (Ph-C4), 136.9 (Imi-C5), 136.3 (Bn-C1), 129.1 (Bn-C3,5), 128.7 (benzylidene-C), 128.0 (Bn-C4), 127.1 (Bn-C2,6), 125.9 (Ph-C1), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 44.0 (CH<sub>2</sub>), 16.3 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>21</sub>H<sub>21</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 353.1496, found: 353.1508;

TLC (CHCl<sub>3</sub>/EtOH 96:4 + 1% AcOH): *R<sub>f</sub>* = 0.56.

#### 4.2.2.30 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methoxybenzyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-PMBn, 66)



The title compound **66** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–96:4 + 1% AcOH) it was obtained as a yellow foam (509 mg, 1.33 mmol, 67%).

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.50 (s, 2 H, Ph-2,6-H), 7.18–7.13 (m, 2 H, Bn-2,6-H), 7.06 (q, *J* = 0.6 Hz, 1 H, benzylidene-H), 6.88–6.82 (m, 2 H, Bn-3,5-H), 6.00 (s<sub>br</sub>, 1 H, OH), 4.75 (s, 2 H, CH<sub>2</sub>), 3.93 (s, 6 H, Ph-OCH<sub>3</sub>), 3.78 (s, 3 H, Bn-OCH<sub>3</sub>), 2.25 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>);

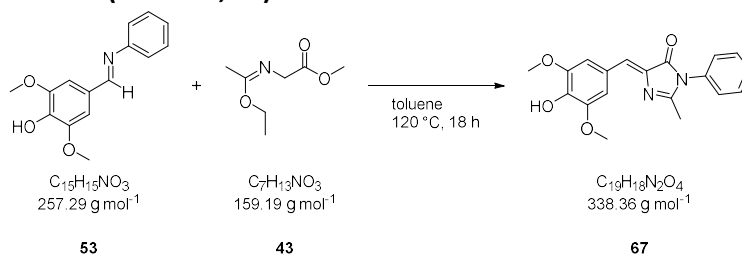
<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.5 (Imi-C4), 161.4 (Imi-C2), 159.3 (Bn-C4), 147.1 (Ph-C3,5), 137.5 (Ph-C4), 136.7 (Imi-C5), 128.5 (Bn-C2,6), 128.3 (Bn-C1), 128.3 (benzylidene-C), 125.8 (Ph-C1), 114.4 (Bn-C3,5), 109.6 (Ph-C2,6), 56.5 (Ph-OCH<sub>3</sub>), 55.5 (Bn-OCH<sub>3</sub>), 43.5 (CH<sub>2</sub>), 16.4 (CCH<sub>3</sub>).

170.7 (Imi-C4), 161.4 (Imi-C2), 147.2 (Ph-C3,5), 137.5 (Ph-C4), 136.9 (Imi-C5), 136.3 (Bn-C1), 129.1 (Bn-C3,5), 128.7 (benzylidene-C), 128.0 (Bn-C4), 127.1 (Bn-C2,6), 125.9 (Ph-C1), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 44.0 (CH<sub>2</sub>), 16.3 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>21</sub>H<sub>23</sub>N<sub>2</sub>O<sub>5</sub>, [M+H]<sup>+</sup>): 383.1601, found: 383.1612;

TLC (CHCl<sub>3</sub>/EtOH 96:4 + 1% AcOH): *R<sub>f</sub>* = 0.63.

#### 4.2.2.31 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-phenyl-3,5-dihydro-4H-imidazol-4-one (DMHBPI, 67)



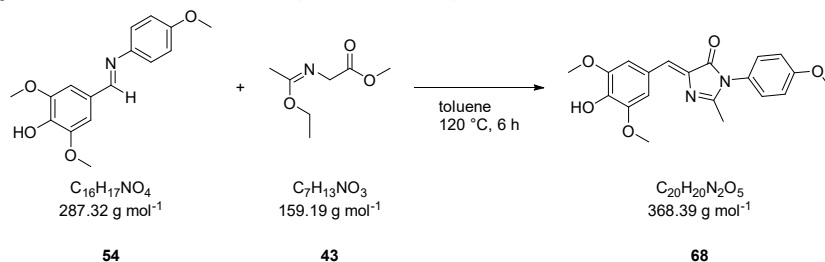
The title compound **67** was synthesized according to General procedure C on a 1.25 mmol scale. After purification by column chromatography (Hex/EtOAc 50:50 + 1% AcOH) it was obtained as a yellow solid (100 mg, 0.30 mmol, 93%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.56 (s, 2 H, Ph-2,6-H), 7.55–7.47 (m, 2 H, NAr-3,5-H), 7.47–7.39 (m, 1 H, NAr-4-H), 7.28–7.21 (m, 2 H, NAr-2,6-H), 7.11 (q, *J* = 0.6 Hz, 1 H, benzylidene-H), 5.95 (s, 1 H, OH), 3.96 (s, 6 H, OCH<sub>3</sub>), 2.27 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (1 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.0 (Imi-C4), 160.6 (Imi-C2), 147.2 (Ph-C3,5), 137.6 (Ph-C4), 136.6 (Imi-C5), 133.8 (NAr-C1), 129.8 (NAr-C3,5), 128.9 (NAr-C4), 128.7 (benzylidene-C), 127.4 (NAr-C2,6), 125.9 (Ph-C1), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 16.7 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>19</sub>H<sub>18</sub>N<sub>2</sub>NaO<sub>4</sub>, [M+Na]<sup>+</sup>): 361.11588, found: 361.11490.

#### 4.2.2.32 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methoxyphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBAl, 68)



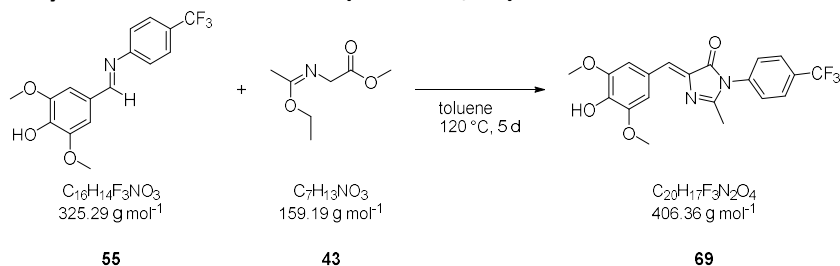
The title compound **68** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (Hex/EtOAc 50:50–25:75 + 1% AcOH) it was obtained as a yellow solid (223 mg, 0.61 mmol, 30%).

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.55 (s, 2 H, Ph-2,6-H), 7.18–7.12 (m, 2 H, NAr-2,6-H), 7.09 (s, 1 H, benzylidene-C), 7.03–6.97 (m, 2 H, NAr-3,5-H), 5.91 (s<sub>br</sub>, 1 H, OH), 3.96 (s, 7 H, Ph-OCH<sub>3</sub>), 3.85 (s, 3 H, NAr-OCH<sub>3</sub>), 2.24 (s, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.2 (Imi-C4), 161.0 (Imi-C2), 159.8 (NAr-C4), 147.1 (Ph-C3,5), 137.5 (Ph-C4), 136.7 (Imi-C5), 128.6 (NAr-C2,6), 128.4 (benzylidene-C), 126.3 (NAr-C1), 126.0 (Ph-C1), 115.1 (NAr-C3,5), 109.6 (Ph-C2,6), 56.5 (Ph-OCH<sub>3</sub>), 55.7 (NAr-OCH<sub>3</sub>), 16.7 (CCH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>20</sub>H<sub>21</sub>N<sub>2</sub>O<sub>5</sub>, [M+H]<sup>+</sup>): 369.1445, found: 369.1453.

#### 4.2.2.33 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-trifluoromethylphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBTF, 69)



The title compound **69** was synthesized according to General procedure C on a 1.00 mmol scale. After purification by column chromatography (Hex/EtOAc 80:20–30:70 + 1% AcOH) it was obtained as a yellow solid (108 mg, 0.27 mmol, 27%).

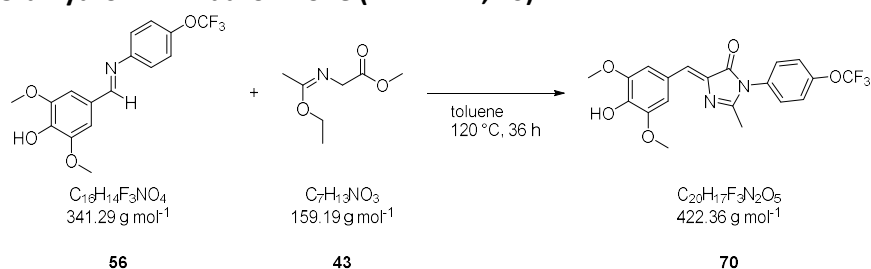
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.81–7.76 (m, 2 H, NAr-3,5-H), 7.56 (s, 2H, Ph-2,6-H), 7.43–7.39 (m, 2H, NAr-2,6-H), 7.14 (d, *J* = 0.7 Hz, 1 H, benzylidene-H), 5.92 (s, 1 H, OH), 3.97 (s, 6 H, OCH<sub>3</sub>), 2.32 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 169.5 (Imi-C4), 159.2 (Imi-C2), 147.2 (Ph-C3,5), 137.9 (Ph-C4), 137.0 (NAr-C1), 136.1 (Imi-C5), 130.9 (q, *J* = 33.2 Hz, NAr-C4), 129.6 (benzylidene-C), 127.7 (NAr-C2,6), 127.0 (q, *J* = 3.7 Hz, NAr-C3,5), 125.7 (Ph-C1), 123.8 (q, *J* = 272.3 Hz, CF<sub>3</sub>), 109.7 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 16.8 (CCH<sub>3</sub>);

<sup>19</sup>F NMR (470 MHz, CDCl<sub>3</sub>): δ (ppm) = –62.7 (CF<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>20</sub>H<sub>18</sub>F<sub>3</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 407.1213, found: 407.1213.

**4.2.2.34 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-trifluoromethoxyphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBAI<sup>F</sup>, 70)**



The title compound **70** was synthesized according to General procedure C on a 2.00 mmol scale. Precipitation of the product was completed by adding Et<sub>2</sub>O (20 mL) to the reaction mixture. Orange crystalline solid (378 mg, 0.89 mmol, 45%).

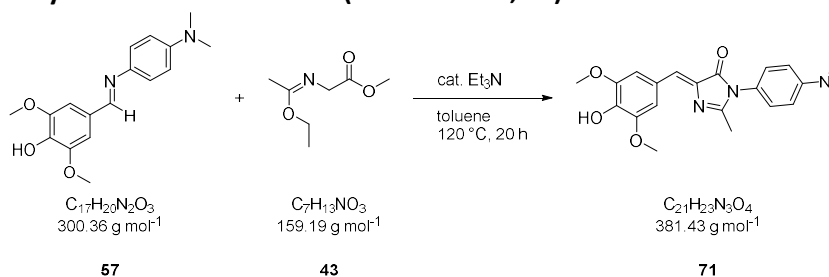
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.56 (s, 2 H, Ph-2,6-H), 7.38–7.34 (m, 2 H, NAr-3,5-H), 7.32–7.28 (m, 2 H, NAr-2,6-H), 7.12 (q, *J* = 0.7 Hz, 1H, benzylidene-H), 5.94 (s, 1 H, OH), 3.97 (s, 6 H, OCH<sub>3</sub>), 2.29 (d, *J* = 0.7 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 169.7 (Imi-C4), 159.8 (Imi-C2), 149.2 (d, *J* = 2.0 Hz, NAr-C4), 147.2 (Ph-C3,5), 137.9 (Ph-C4), 136.1 (Imi-C5), 132.2 (NAr-C1), 129.3 (benzylidene-C), 128.9 (NAr-C2,6), 125.8 (Ph-C1), 122.3 (NAr-C3,5), 109.7 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 16.7 (CCH<sub>3</sub>);

<sup>19</sup>F NMR (470 MHz, CDCl<sub>3</sub>): δ (ppm) = -57.9 (OCF<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>20</sub>H<sub>18</sub>F<sub>3</sub>N<sub>2</sub>O<sub>5</sub>, [M+H]<sup>+</sup>): 423.1162, found: 423.1169.

**4.2.2.35 (Z)-3-(4-(Dimethylamino)phenyl)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-DMA, 71)**



A mixture of imine **57** (601 mg, 2.00 mmol, 1.00 eq.), imidate **43** (382 mg, 2.40 mmol, 1.20 eq.) and Et<sub>3</sub>N (4 drops) in toluene (2 mL) was stirred at 120 °C for 20 h. Afterwards, the solvent was evaporated under reduced pressure. The crude product was dissolved in hot CHCl<sub>3</sub> (50 mL) and precipitated by the addition of hexane (100 mL). The precipitate was washed with hexane (3×15 mL) to afford compound **71** as a yellow solid (599 mg, 1.57 mmol, 79%).

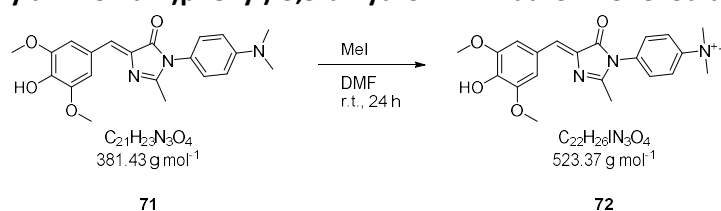
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.55 (s, 2 H, Ph-2,6-H), 7.13–7.01 (m, 3 H, benzylidene-H and NAr-2,6-H), 6.82–6.72 (m, 2 H, NAr-3,5-H), 5.90 (s<sub>br</sub>, 1 H, OH), 3.96 (s, 6 H, OCH<sub>3</sub>), 3.00 (s, 6 H, NCH<sub>3</sub>), 2.24 (s, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.5 (Imi-C4), 161.9 (Imi-C2), 150.6 (NAr-C4), 147.1 (Ph-C3,5), 137.4 (Ph-C4), 137.0 (Imi-C5), 128.1 (NAr-C2,6), 127.9 (Ph-C1), 126.1 (benzylidene-C), 122.0 (NAr-C1), 112.8 (NAr-C3,5), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 40.7 (NCH<sub>3</sub>), 16.7 (CCH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>21</sub>H<sub>24</sub>N<sub>3</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 382.1761.1002, found: 382.1772;

TLC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 98:2 + 1% AcOH): *R<sub>f</sub>* = 0.83.

**4.2.2.36 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(4-(trimethylammonium)phenyl)-3,5-dihydro-4H-imidazol-4-one iodide (DMHBI<sup>+</sup>, 72)**



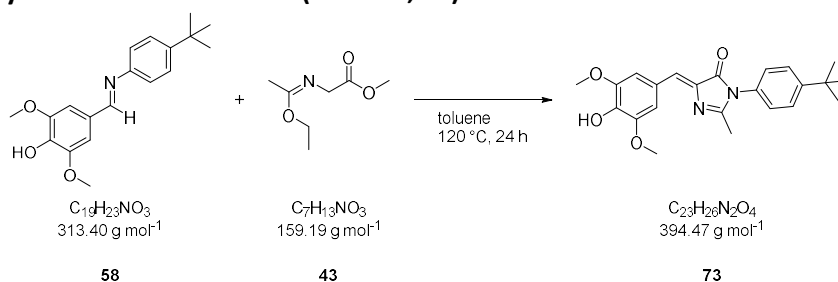
Compound **71** (66.8 mg, 175  $\mu\text{mol}$ , 1.00 eq.) and methyl iodide (109  $\mu\text{L}$ , 1.75 mmol, 10.0 eq.) were dissolved in DMF (3.5 mL) and stirred at ambient temperature for 24 h. Removal of the solvent under reduced pressure afforded the product **72** as a dark solid (92.0 mg, 175  $\mu\text{mol}$ , > 99%).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.28 (s<sub>br</sub>, 1 H, OH), 8.18–8.11 (m 2 H, NAr-3,5-H), 7.76–7.69 (m, 2 H, NAr-2,6-H), 7.71 (s, 2 H, Ph-2,6-H), 7.06 (s, 1 H, benzylidene-H), 3.81 (s, 6 H, OCH<sub>3</sub>), 3.66 (s, 9 H, NCH<sub>3</sub>), 2.26 (s, 3 H, CCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 168.8 (Imi-C4), 159.8 (Imi-C2), 147.9 (Ph-C3,5), 146.5 (NAr-C4), 139.0 (Ph-C4), 135.4 (Imi-C4), 135.0 (NAr-C1), 128.8 (NAr-C2,6), 127.6 (benzylidene-C), 124.3 (Ph-C1), 121.9 (NAr-C3,5), 110.3 (Ph-C2,6), 56.5 (NCH<sub>3</sub>), 56.1 (OCH<sub>3</sub>), 16.5 (CCH<sub>3</sub>);

**HR-MS** (ESI<sup>+</sup>): *m/z* calc. (C<sub>22</sub>H<sub>26</sub>N<sub>3</sub>O<sub>4</sub>, [M-I]<sup>+</sup>): 396.1918, found: 396.1926.

**4.2.2.37 (Z)-3-(4-*tert*-Butylphenyl)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI<sup>c</sup>, 73)**



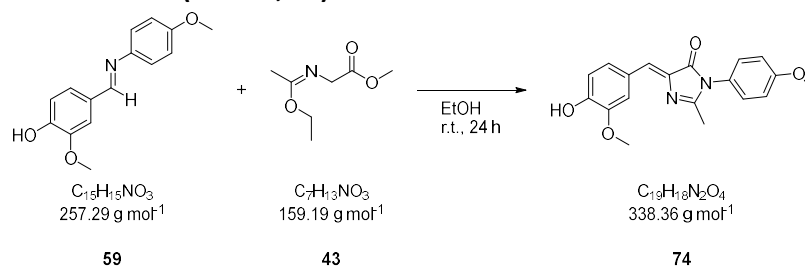
The title compound **73** was synthesized according to General procedure C on a 2.00 mmol scale. Precipitation of the product was completed by adding Et<sub>2</sub>O (20 mL) to the reaction mixture. Dark yellow solid (604 mg, 1.53 mmol, 76%).

**<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.56 (s, 2 H, Ph-2,6-H), 7.53–7.48 (m, 2 H, NAr-3,5-H), 7.19–7.13 (m, 2 H, NAr-2,6-H), 7.13–7.08 (m, 1 H, benzylidene-H), 5.92 (s<sub>br</sub>, 1 H, OH), 3.96 (s, 6 H, OCH<sub>3</sub>), 2.29 – 2.26 (m, 3 H, Imi-2-CH<sub>3</sub>), 1.35 (s, 9 H, C(CH<sub>3</sub>)<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 170.2 (Imi-C4), 161.1 (Imi-C2), 152.0 (NAr-C4), 147.2 (Ph-C3,5), 137.5 (Ph-C4), 136.7 (Imi-C5), 131.0 (NAr-C1), 128.5 (benzylidene-C), 126.9 (NAr-C2,6), 126.8 (NAr-C3,5), 126.0 (Ph-C1), 109.5 (Ph-C2,6), 56.5 (OCH<sub>3</sub>), 34.9 (C(CH<sub>3</sub>)<sub>3</sub>), 31.4 (C(CH<sub>3</sub>)<sub>3</sub>), 16.8 (Imi-2-CH<sub>3</sub>);

**HR-MS** (ESI<sup>+</sup>): *m/z* calc. (C<sub>23</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 395.1965, found: 395.1965.

#### 4.2.2.38 (Z)-5-(4-Hydroxy-3-methoxybenzylidene)-2-methyl-3-(4-methoxyphenyl)-3,5-dihydro-4H-imidazol-4-one (MHBAI, 74)



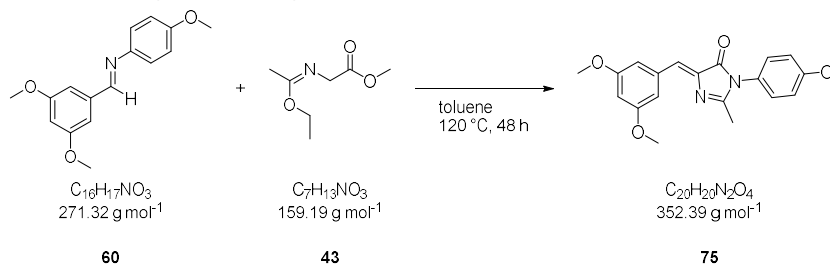
The title compound **74** was synthesized according to General procedure C on a 1.25 mmol scale. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–94:6 + 1% AcOH) it was obtained as a dark yellow solid (107 mg, 316 μmol, 25%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.04 (d, *J* = 1.9 Hz, 1H, Ph-C-2-H), 7.57 (ddd, *J* = 8.3, 1.9, 0.6 Hz, 1 H, Ph-6-H), 7.18–7.13 (m, 2 H, NAr-3,5-H), 7.13–7.12 (m, 1 H, benzylidene-H), 7.03–6.98 (m, 2 H, NAr-2,6-H), 6.96 (d, *J* = 8.3 Hz, 1 H, Ph-5-H), 6.05 (s<sub>br</sub>, 1 H, OH), 3.98 (s, 3 H, Ph-OCH<sub>3</sub>), 3.85 (s, 3 H, NAr-OCH<sub>3</sub>), 2.24 (d, *J* = 0.6 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.4 (Imi-C4), 160.9 (Imi-C2), 159.9 (NAr-C1), 148.3 (Ph-C4), 146.8 (Ph-C3), 136.4 (Imi-C5), 128.7 (NAr-C3,5), 128.5 (benzylidene-C), 127.7 (Ph-C6), 127.2 (Ph-C1), 126.4 (NAr-C4), 115.1 (NAr-C2,6), 114.8 (Ph-C5), 113.9 (Ph-C2), 56.1 (Ph-OCH<sub>3</sub>), 55.7 (NAr-OCH<sub>3</sub>), 16.6 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>19</sub>H<sub>18</sub>N<sub>2</sub>NaO<sub>4</sub>, [M+Na]<sup>+</sup>): 361.11483, found: 361.11588.

#### 4.2.2.39 (Z)-5-(3,5-Dimethoxybenzylidene)-3-(4-methoxyphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMBAI, 75)



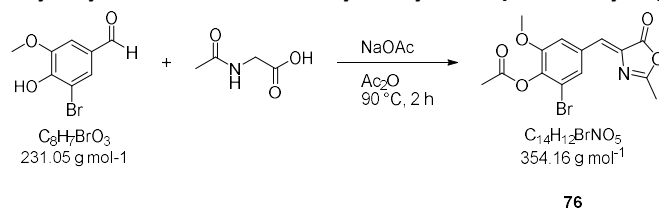
The title compound **75** was synthesized according to General procedure C on a 2.00 mmol scale. Precipitation of the product was completed by adding pentane (20 mL) to the reaction mixture. Orange crystalline solid (484 mg, 1.37 mmol, 69%).

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.40 (dd, *J* = 2.3, 0.5 Hz, 2 H, Ph-2,6-H), 7.19–7.12 (m, 2 H, NAr-3,5-H), 7.11–7.08 (m, 1 H, benzylidene-H), 7.05–6.97 (m, 2 H, NAr-2,6-H), 6.53 (t, *J* = 2.3 Hz, 1 H, Ph-4-H), 3.85 (s, 3 H, NAr-OCH<sub>3</sub>), 3.85 (s, 6H, Ph-OCH<sub>3</sub>), 2.24 (d, *J* = 0.7 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.40 (Imi-C4), 162.5 (Imi-C2), 160.9 (Ph-C3,5), 160.0 (NAr-C1), 138.8 (Imi-C5), 136.0 (Ph-C1), 128.7 (NAr-C3,5), 127.9 (benzylidene-C), 126.2 (NAr-C4), 115.1 (NAr-C2,6), 110.1 (Ph-C2,6), 103.3 (Ph-C4), 55.7 (NAr-OCH<sub>3</sub>), 55.6 (Ph-OCH<sub>3</sub>), 16.6 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>20</sub>H<sub>21</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 353.1496, found: 353.1495.

#### 4.2.2.40 (Z)-4-(4-Acetyloxy-3-bromo-5-methoxybenzylidene)-2-methyl-5(4H)-oxazolone (76)



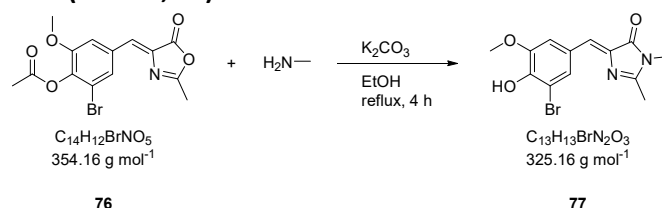
3-Bromo-4-hydroxy-5-methoxybenzaldehyde (1.00 g, 4.33 mmol, 1.00 eq.), *N*-acetylglycine (507 mg, 4.33 mmol, 1.00 eq.) and NaOAc (355 mg, 4.33 mmol, 1.00 eq.) were suspended in Ac<sub>2</sub>O (3 mL). The mixture was stirred at 90 °C for 2 h and then cooled to ambient temperature. It was diluted with EtOH (3.5 mL) and kept at 0 °C for 16 h. The precipitate was filtered off and rinsed with cold EtOH (2 mL), hot H<sub>2</sub>O (5 mL) and hexane (2×5 mL) to afford the title compound **76** as a yellow solid (1.05 g, 2.96 mmol, 68%).

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.84 (d, *J* = 1.8 Hz, 1 H, Ph-2-H), 7.80 (d, *J* = 1.8 Hz, 1 H, Ph-6-H), 6.98 (s, 1 H, benzylidene-H), 3.89 (s, 3 H, OCH<sub>3</sub>), 2.41 (s, 3 H, CCH<sub>3</sub>), 2.38 (s, 3 H, OC(O)CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 167.7 (OC(O)CH<sub>3</sub>), 167.4 (Oxa-C5), 167.1 (Oxa-C2), 152.7 (Ph-C5), 140.1 (Ph-C4), 133.6 (Oxa-C4), 132.7 (benzylidene-C), 129.0 (Ph-C1), 128.7 (Ph-C2), 117.7 (Ph-C3), 114.4 (Ph-C6), 56.5 (OCH<sub>3</sub>), 20.6 (OC(O)CH<sub>3</sub>), 15.9 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>15</sub>H<sub>15</sub>BrNO<sub>5</sub>, [M+MeOH-H]<sup>-</sup>): 384.0088, found: 384.0079.

#### 4.2.2.41 (Z)-5-(3-Bromo-4-hydroxy-5-methoxybenzylidene)-2,3-dimethyl-3,5-dihydro-4H-imidazol-4-one (BMHBI, 77)



Oxazolone **76** (300 mg, 847 μmol, 1.00 eq.), a 40% solution of MeNH<sub>2</sub> in H<sub>2</sub>O (0.23 mL, 2.67 mmol 3.15 eq.) and K<sub>2</sub>CO<sub>3</sub> (162 mg, 1.17 mmol, 1.38 eq.) were suspended in EtOH (4 mL) and heated to reflux for 4 h. After cooling to ambient temperature, the precipitate was filtered off and dissolved in aqueous acetate buffer (50 mL, pH 3.5). The solution was extracted with EtOAc (1×50 mL, 2×20 mL) and the combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>. After removal of the solvent, the residue was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 97:3) to afford the title compound **77** as a dark yellow solid (58 mg, 174 μmol, 21%).

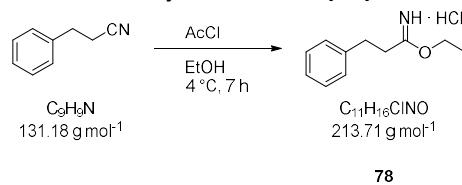
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.89 (d, *J* = 1.8 Hz, 1 H, Ph-6-H), 7.80 (d, *J* = 1.8 Hz, 1 H, Ph-2-H), 6.94 (q, *J* = 0.7 Hz, 1 H, benzylidene-H), 6.29 (s<sub>br</sub>, 1 H, OH), 3.96 (s, 3 H, OCH<sub>3</sub>), 3.18 (s, 3 H, NCH<sub>3</sub>), 2.38 (d, *J* = 0.7 Hz, 3 H, CCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 170.7 (Imi-C4), 162.2 (Imi-C2), 147.2 (Ph-C5), 145.1 (Ph-C4), 137.9 (Imi-C5), 129.8 (Ph-C2), 127.8 (Ph-C1), 126.1 (benzylidene-C), 113.2 (Ph-C6), 108.4 (Ph-C3), 56.6 (OCH<sub>3</sub>), 26.8 (NCH<sub>3</sub>), 15.9 (CCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>13</sub>H<sub>13</sub>BrKN<sub>2</sub>O<sub>3</sub>, [M+K]<sup>+</sup>): 362.9741, found: 362.9750.



#### 4.2.2.42 Ethyl 3-phenylpropionimidate hydrochloride (78)



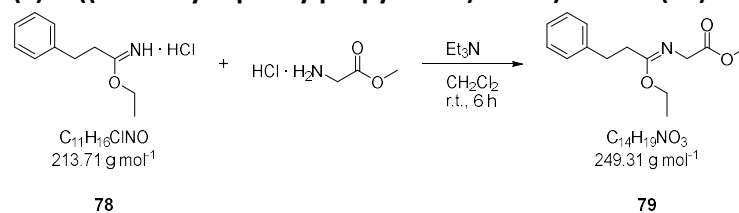
The title compound **78** was prepared according to a previously reported procedure:<sup>143</sup> 3-Phenylpropionitrile (787 mg, 6.00 mmol, 1.00 eq.) was dissolved in EtOH (4.2 mL, 72.0 mmol, 12.0 eq.) and cooled to 0 °C. AcCl (3.41 mL, 48.0 mmol, 8.00 eq.) was added dropwise and the resulting yellow solution was stirred at 4 °C for 7 h. The product was precipitated by addition of Et<sub>2</sub>O (20 mL), collected by filtration and washed with Et<sub>2</sub>O (30 mL), giving a white, crystalline solid (963 mg, 4.51 mmol, 75%).

<sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 11.95 (s<sub>br</sub>, 1 H, NH · HCl), 11.08 (s<sub>br</sub>, 1 H, NH · HCl), 7.34–7.29 (m, 2 H, Ph-3,5-H), 7.28–7.20 (m, 3 H, Ph-2,6-H, Ph-4-H), 4.37 (q, *J* = 7.0 Hz, 2 H, OCH<sub>2</sub>CH<sub>3</sub>), 2.94 (s, 4 H, PhCH<sub>2</sub>CH<sub>2</sub>), 1.30 (t, *J* = 7.0 Hz, 3 H, OCH<sub>2</sub>CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 178.3 (CNO), 138.9 (Ph-C1), 128.6 (Ph-C3,5), 128.3 (Ph-C2,6), 126.6 (Ph-C4), 69.0 (OCH<sub>2</sub>CH<sub>3</sub>), 34.1 (PhCH<sub>2</sub>CH<sub>2</sub>), 30.5 (PhCH<sub>2</sub>CH<sub>2</sub>), 13.3 (OCH<sub>2</sub>CH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>11</sub>H<sub>16</sub>NO, [M–Cl]<sup>+</sup>): 178.1226, found: 178.1230.

#### 4.2.2.43 Methyl (Z)-2-((1-ethoxy-3-phenylpropylidene)amino)acetate (79)



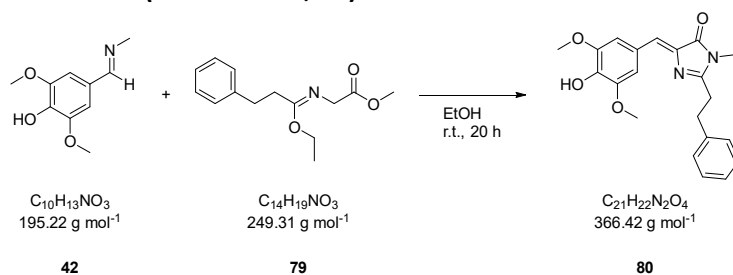
To a suspension of ethyl 3-phenylpropionimidate hydrochloride (**78**, 748 mg, 3.50 mmol, 1.00 eq.) and methyl glycinate hydrochloride (439 mg, 3.50 mmol, 1.00 eq.) in dry CH<sub>2</sub>Cl<sub>2</sub> (14 mL) was added Et<sub>3</sub>N (0.49 mL, 3.50 mmol, 1.00 eq.). The resulting mixture was stirred at ambient temperature for 6 h. Afterwards, it was washed with H<sub>2</sub>O (2×14 mL) and brine (14 mL) and the organic phase was dried over MgSO<sub>4</sub>. Evaporation of the solvent under reduced pressure afforded the title compound **79** as a pale yellow oil (631 mg, 2.53 mmol, 72%).

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.35–7.25 (m, 2 H, Ph-3,5-H), 7.25–7.12 (m, 4 H, Ph-2,6-H, Ph-4-H), 4.12 (q, *J* = 7.1 Hz, 2 H, OCH<sub>2</sub>CH<sub>3</sub>), 3.92 (s, 2 H, NCH<sub>2</sub>), 3.70 (s, 3 H, OCH<sub>3</sub>), 2.87 (dd, *J* = 9.0, 6.7 Hz, 2 H, PhCH<sub>2</sub>CH<sub>2</sub>), 2.52–2.45 (m, 2 H, PhCH<sub>2</sub>CH<sub>2</sub>), 1.27 (t, *J* = 7.1 Hz, 3 H, OCH<sub>2</sub>CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 171.6 (COOCH<sub>3</sub>), 165.8 (CNO), 140.8 (Ph-C1), 128.6 (Ph-C3,5), 128.4 (Ph-C2,6), 126.4 (Ph-C4), 61.1 (OCH<sub>2</sub>CH<sub>3</sub>), 52.1 (OCH<sub>3</sub>), 50.7 (NCH<sub>2</sub>), 32.3 (PhCH<sub>2</sub>CH<sub>2</sub>), 31.0 (PhCH<sub>2</sub>CH<sub>2</sub>), 14.4 (OCH<sub>2</sub>CH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>14</sub>H<sub>20</sub>NO<sub>3</sub>, [M+H]<sup>+</sup>): 250.1438, found: 250.1488.

**4.2.2.44 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-methyl-2-phenylethyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-PhEt, 80)**



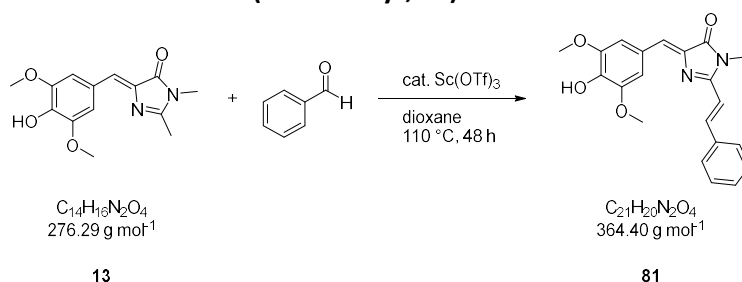
The title compound **80** was synthesized according to General procedure C on a 1.50 mmol scale. Part of the product precipitated from the reaction mixture and was filtered off. A second batch was obtained by evaporation of the filtrate and purified by column chromatography (CHCl<sub>3</sub>/EtOH 98:2–90:10 + 1% AcOH). Yellow foam (399 mg, 1.09 mmol, 73%).

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.58 (s, 2 H, Ph-2,6-H), 7.34–7.28 (m, 2 H, alkyl-Ar-3,5-H), 7.30–7.27 (m, 2 H, alkyl-Ar-2,6-H), 7.24 (m, 1 H, alkyl-Ar-4-H), 7.04 (s, 1 H, benzylidene-H), 6.00 (s<sub>br</sub>, 1 H, OH), 3.92 (s, 6 H, OCH<sub>3</sub>), 3.20 (dd,  $J$  = 8.5, 7.2 Hz, 2 H, CCH<sub>2</sub>CH<sub>2</sub>Ar), 3.08 (s, 3 H, NCH<sub>3</sub>), 2.88 (dd,  $J$  = 8.5, 7.2 Hz, 2 H, CCH<sub>2</sub>CH<sub>2</sub>Ar);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 170.8 (Imi-C4), 163.4 (Imi-C2), 147.0 (Ph-C3,5), 140.5 (alkyl-Ar-C1), 137.4 (Ph-C4), 137.0 (Imi-C5), 128.7 (alkyl-Ar-C3,5), 128.4 (alkyl-Ar-C2,6), 128.0 (benzylidene-C), 126.6 (alkyl-Ar-C4), 125.9 (Ph-C1), 109.6 (Ph-C2,6), 56.4 (OCH<sub>3</sub>), 31.2 (CCH<sub>2</sub>CH<sub>2</sub>Ar), 30.8 (CCH<sub>2</sub>CH<sub>2</sub>Ar), 26.5 (NCH<sub>3</sub>);

**HR-MS** (ESI+):  $m/z$  calc. (C<sub>21</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 367.1652, found: 367.1654.

**4.2.2.45 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-methyl-2-((E)-2-phenylvinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-Styr, 81)**



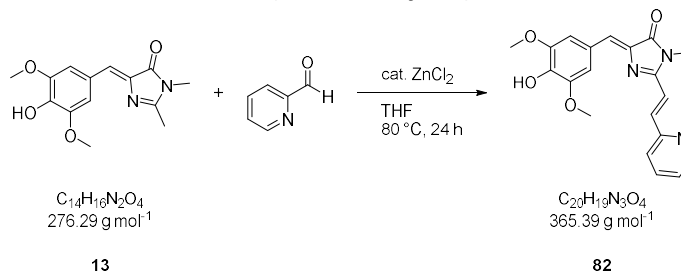
The title compound **81** was synthesized according to General procedure D on a 1.00 mmol scale. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–96:4 + 1% AcOH) it was obtained as an orange-red solid (176 mg, 0.48 mmol, 48%).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.23 (s<sub>br</sub>, 1 H, OH), 7.96 (d,  $J$  = 15.9 Hz, 1 H, CCHCHAr), 7.84–7.79 (m, 2 H, vinyl-Ar-2,6-H), 7.76 (s, 2 H, Ph-2,6-h), 7.51–7.40 (m, 3 H, vinyl-Ar-3,5-H, vinyl-Ar-4-H), 7.25 (d,  $J$  = 15.9 Hz, 1 H, CCHCHAr), 6.99 (s, 1 H, benzylidene-H), 3.86 (s, 6 H, OCH<sub>3</sub>), 3.28 (s, 3 H, NCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 169.9 (Imi-C4), 158.7 (Imi-C2), 147.9 (Ph-C3,5), 139.5 (CCHCHAr), 138.7 (Ph-C4), 137.2 (Imi-C5), 135.2 (vinyl-Ar-C1), 130.1 (vinyl-Ar-C4), 129.0 (vinyl-Ar-C3,5), 128.3 (vinyl-Ar-C2,6), 126.4 (benzylidene-C), 125.0 (Ph-C1), 114.1 (CCHCHAr), 110.1 (Ph-C2,6), 55.9 (OCH<sub>3</sub>), 26.4 (NCH<sub>3</sub>);

HR-MS (ESI+):  $m/z$  calc. ( $C_{21}H_{19}N_2O_4$ ,  $[M-H]^-$ ): 363.1350, found: 363.1333.

#### 4.2.2.46 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-methyl-2-((E)-2-(pyridin-2-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-2Py, **82**)



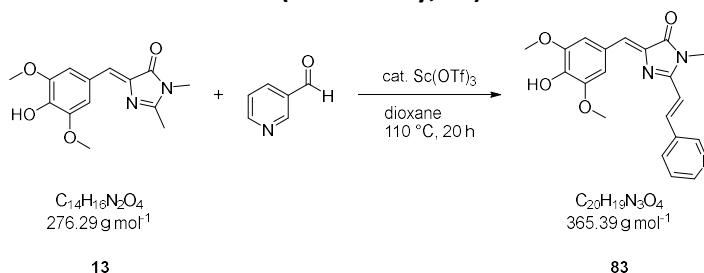
DMHBI (**13**, 207 mg, 750  $\mu$ mol, 1.00 eq.), pyridine-2-carbaldehyde (90.1 mg, 841  $\mu$ mol, 1.01 eq.) and anhydrous  $ZnCl_2$  (10.2 mg, 75.0  $\mu$ mol, 10mol%) were dissolved in THF (1.5 mL) and heated to 80 °C in a sealed tube for 16 h. Then, a second portion of the aldehyde (36.3 mg, 339  $\mu$ mol, 0.45 eq.) was added and the reaction was continued under the same conditions for 8 h. The solvent was removed under reduced pressure and the residue was washed with MeOH (50 mL) to afford the title compound **82** as an orange solid (41.4 mg, 113  $\mu$ mol, 15%).

<sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.23 (s<sub>br</sub>, 1 H, OH), 8.68 (ddd,  $J$  = 4.8, 1.2, 0.9 Hz, 1 H, vinyl-Ar-6-H), 7.95 (d,  $J$  = 15.6 Hz, 1 H, CCHCHAR), 7.89 (ddd,  $J$  = 7.7, 7.6, 1.7 Hz, 1 H, vinyl-Ar-4-H), 7.81 (ddd,  $J$  = 7.7, 1.2, 0.9 Hz, 1 H, vinyl-Ar-3-H), 7.76 (s, 2 H, Ph-2,6-H), 7.54 (d,  $J$  = 15.6 Hz, 1 H, CCHCHAR), 7.41 (ddd,  $J$  = 7.6, 4.8, 1.2 Hz, 1 H, vinyl-Ar-5-H), 7.04 (s, 1 H, benzylidene-H), 3.87 (s, 6 H, OCH<sub>3</sub>), 3.28 (s, 3 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 169.7 (Imi-C4), 158.2 (Imi-C2), 152.9 (vinyl-Ar-C2), 150.0 (vinyl-Ar-C6), 147.9 (Ph-C3,5), 139.0 (Ph-C4), 138.4 (CCHCHAR), 137.1 (vinyl-Ar-C4), 137.1 (Imi-C5), 127.3 (benzylidene-C), 124.8 (Ph-C1), 124.5 (vinyl-Ar-C3), 124.2 (vinyl-Ar-C5), 117.4 (CCHCHAR), 110.4 (Ph-C2,6), 56.0 (OCH<sub>3</sub>), 26.3 (NCH<sub>3</sub>);

HR-MS (ESI+):  $m/z$  calc. ( $C_{20}H_{20}N_3O_4$ ,  $[M+H]^+$ ): 366.1448, found: 366.1446.

#### 4.2.2.47 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-methyl-2-((E)-2-(pyridin-3-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-3Py, **83**)



The title compound **83** was synthesized according to General procedure D on a 200  $\mu$ mol scale. Brown solid (25.3 mg, 69.2  $\mu$ mol, 35%).

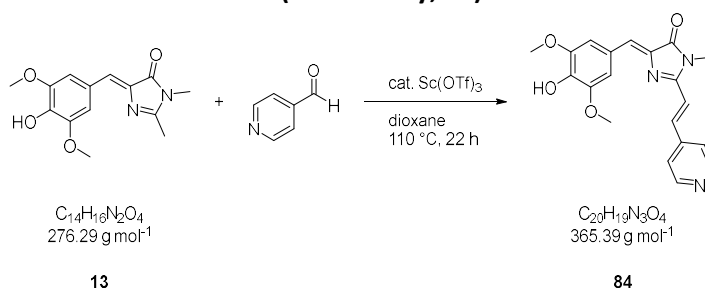
<sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.27 (s<sub>br</sub>, 1 H, OH), 8.97 (d,  $J$  = 2.2 Hz, 1 H, vinyl-Ar-2-H), 8.59 (dd,  $J$  = 4.8, 1.9 Hz, 1 H, vinyl-Ar-6-H), 8.30 (ddd,  $J$  = 8.0, 2.2, 1.9 Hz, 1 H, vinyl-Ar-4-H), 7.96 (d,  $J$  = 15.9 Hz, 1 H, CCHCHAR), 7.76 (s, 2 H, Ph-2,6-H), 7.50 (dd,  $J$  = 8.0, 4.8 Hz, 1 H, vinyl-5-H), 7.40 (d,  $J$  = 15.9 Hz, 1 H, CCHCHAR), 7.02 (s, 1 H, benzylidene-C), 3.86 (s, 6 H, OCH<sub>3</sub>), 3.29 (s, 3 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 169.8 (Imi-C4), 158.3 (Imi-C2), 150.4 (vinyl-Ar-C6), 149.8 (vinyl-Ar-C2), 147.9 (Ph-C3,5), 138.9 (Ph-C4), 137.0 (Imi-C5), 135.8 (CCHCHAR), 134.4 (vinyl-

Ar-C4), 130.9 (vinyl-Ar-C3), 126.9 (benzylidene-C), 124.8 (Ph-C1), 123.8 (vinyl-Ar-C5), 116.1 (CCHCHAR), 110.3 (Ph-C2,6), 55.9 (OCH<sub>3</sub>), 26.3 (NCH<sub>3</sub>);

**HR-MS** (ESI<sup>+</sup>): *m/z* calc. (C<sub>20</sub>H<sub>20</sub>N<sub>3</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 366.1448, found: 366.1441.

#### 4.2.2.48 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-methyl-2-((E)-2-(pyridin-4-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-4Py, **84**)



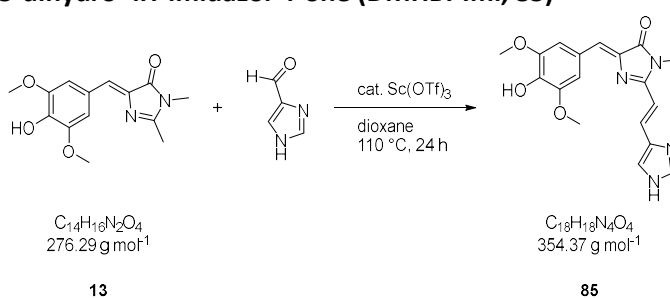
The title compound **84** was synthesized according to General procedure D on a 250 μmol scale. After purification by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 96:4–80:20) it was obtained as a dark brown solid (60.9 mg, 167 μmol, 67%).

**<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 9.22 (s<sub>br</sub>, 1 H, OH), 8.66 (dd, *J* = 4.2, 1.6 Hz, 2 H, vinyl-Ar-2,6-H), 7.88 (d, *J* = 15.8 Hz, 1 H, CCHCHAR), 7.76 (dd, *J* = 4.2, 1.6 Hz, 2 H, vinyl-Ar-3,5-H), 7.75 (s, 3 H, Ph-2,6-H), 7.48 (d, *J* = 15.8 Hz, 1 H, CCHCHAR), 7.05 (s, 1 H, benzylidene-H), 3.87 (s, 6 H, OCH<sub>3</sub>), 3.29 (s, 3 H, NCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 169.5 (Imi-C4), 157.8 (Imi-C2), 150.1 (vinyl-Ar-C2,6), 147.7 (Ph-C3,5), 141.9 (vinyl-Ar-C4), 139.0 (Ph-C4), 136.8 (Imi-C5), 136.2 (CCHCHAR), 127.5 (benzylidene-C), 124.6 (Ph-C1), 121.8 (vinyl-Ar-C3,5), 118.7 (CCHCHAR), 110.4 (Ph-C2,6), 55.9 (OCH<sub>3</sub>), 26.4 (NCH<sub>3</sub>);

**HR-MS** (ESI<sup>+</sup>): *m/z* calc. (C<sub>20</sub>H<sub>20</sub>N<sub>3</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 366.1448, found: 366.1447.

#### 4.2.2.49 2-((E)-2-(1H-Imidazol-4-yl)vinyl)-((Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene))-3-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-Imi, **85**)



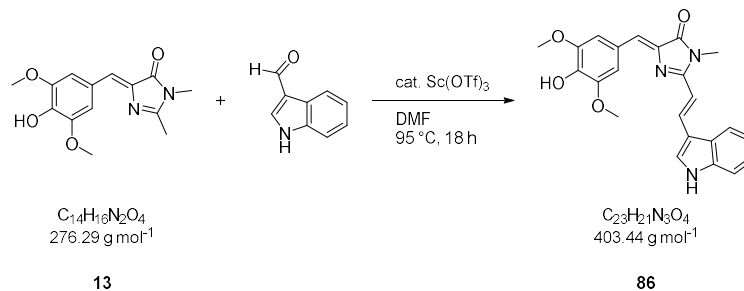
The title compound **85** was synthesized according to General procedure D on a 200 μmol scale. After purification by column chromatography (toluene/DMF 4:1–1:9) it was obtained as a brown solid (44.7 mg, 126 μmol, 63%).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 12.41 (s<sub>br</sub>, 1 H, NH), 9.11 (s<sub>br</sub>, 1 H Ph-4-OH), 7.85 (d, *J* = 15.4 Hz, 1 H, vinyl-2-H), 7.82 (s, 1 H, vinyl-Imi-2-H), 7.71 (s, 2 H, Ph-2,6-H), 7.62 (s, 1 H, vinyl-Imi-5-H), 6.95 (d, *J* = 15.4 Hz, 1 H, vinyl-1-H), 6.90 (s, 1 H, benzylidene-H), 3.85 (s, 6 H, OCH<sub>3</sub>), 3.22 (s, 3 H, NCH<sub>3</sub>);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) = 169.9 (Imi-C4), 159.1 (Imi-C2), 147.8 (Ph-C3,5), 138.5 (Ph-C4), 137.9\* (vinyl-Imi-C4), 137.4 (Imi-C5), 132.9\* (vinyl-C2), 125.1 (Ph-C1), 124.8 (benzylidene-C), 110.1 (Ph-C2,6), 109.5 (vinyl-C1), 56.0 (OCH<sub>3</sub>), 26.2 (NCH<sub>3</sub>);

HR-MS (ESI+):  $m/z$  calc. (C<sub>18</sub>H<sub>19</sub>N<sub>4</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 355.1401, found: 355.1399.

#### 4.2.2.50 2-((E)-2-(1H-Indol-3-yl)vinyl)-5-((Z)-4-hydroxy-3,5-dimethoxybenzylidene)-3-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-Ind, 86)



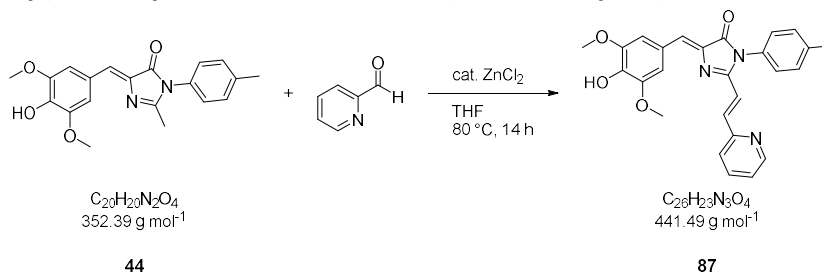
The title compound **86** was synthesized according to General procedure D on a 200  $\mu\text{mol}$  scale using DMF at 95 °C as the solvent instead of dioxane. After purification by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–75:25) it was obtained as an orange solid (24.4 mg, 60.5  $\mu\text{mol}$ , 30%).

$^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) = 8.28–8.22 (m, 1 H, CCHCHAR), 8.04 (d,  $J$  = 4.5 Hz, 1 H, vinyl-Ar-2-H), 7.99 (d,  $J$  = 7.5 Hz, 1 H, vinyl-Ar-4-H), 7.75 (s, 2 H, Ph-2,6-H), 7.53–7.47 (m, 1 H, vinyl-Ar-7-H), 7.28–7.24 (m, 1 H, vinyl-Ar-6-H), 7.24–7.18 (m, 1 H, vinyl-Ar-5-H), 6.92–6.86 (m, 1 H, CCHCHAR), 6.85 (s, 1 H, benzylidene-H), 3.88 (s, 6 H, OCH<sub>3</sub>), 3.30 (s, 3 H, NCH<sub>3</sub>);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) = 170.0 (Imi-C4), 159.6 (Imi-C2)', 147.9 (Ph-C3,5), 138.9 (Ph-C4), 137.4 (vinyl-Ar-C7a), 137.3 (Imi-C5), 134.1 (CCHCHAR), 131.5 (vinyl-Ar-C2), 125.0 (vinyl-Ar-C3a), 124.9 (Ph-C1), 123.5 (benzylidene-C), 122.6 (vinyl-Ar-C6), 120.9 (vinyl-Ar-C5), 119.8 (vinyl-Ar-C4), 113.2 (vinyl-Ar-C3), 112.4 (vinyl-Ar-C7), 109.9 (Ph-C2,6), 106.7 (CCHCHAR), 55.9 (OCH<sub>3</sub>), 26.1 (NCH<sub>3</sub>);

HR-MS (ESI+):  $m/z$  calc. (C<sub>23</sub>H<sub>22</sub>N<sub>3</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 404.1605, found: 404.1596.

#### 4.2.2.51 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-2-((E)-2-(pyridin-2-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI-2Py, 87)



DMHBTI (**44**, 118 mg, 335  $\mu\text{mol}$ , 1.00 eq.), pyridine-2-carbaldehyde (56.2 mg, 525  $\mu\text{mol}$ , 1.57 eq.) and anhydrous ZnCl<sub>2</sub> (4.8 mg, 35.0  $\mu\text{mol}$ , 10mol%) were dissolved in THF (1.5 mL) and heated to 80 °C in a sealed tube for 14 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (Hex/EtOAc 50:50–20:80 + 1% AcOH) to afford the title compound **87** as an orange-red solid (135 mg, 309  $\mu\text{mol}$ , 91%).

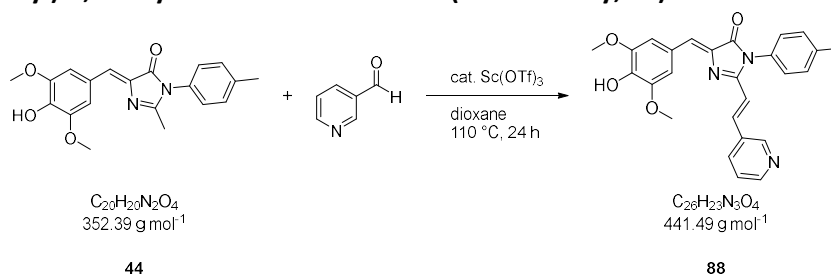
$^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.59 (ddd,  $J$  = 4.8, 1.9, 0.8 Hz, 1 H), 7.90 (d,  $J$  = 15.5 Hz, 1 H), 7.69 (s, 1 H), 7.71–7.65 (m, 2 H), 7.36–7.30 (m, 2 H), 7.32 (dt,  $J$  = 7.8, 1.0 Hz, 1 H), 7.25–7.18 (m, 5 H), 6.00 (s, 1 H), 4.02 (s, 6 H), 2.43 (s, 3 H);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 170.1, 157.7, 153.4, 150.2, 147.2, 138.8, 137.9, 137.5, 136.8, 130.7, 130.4, 129.2, 127.4, 126.5, 124.4, 123.9, 118.3, 110.0, 56.6, 21.5;

(complete spectral assignment was not possible due to strongly overlapping  $^1\text{H}$  resonances)

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{26}\text{H}_{24}\text{N}_3\text{O}_4$ ,  $[\text{M}+\text{H}]^+$ ): 442.1761, found: 442.1757.

#### 4.2.2.52 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-2-((E)-2-(pyridin-3-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI-3Py, 88)



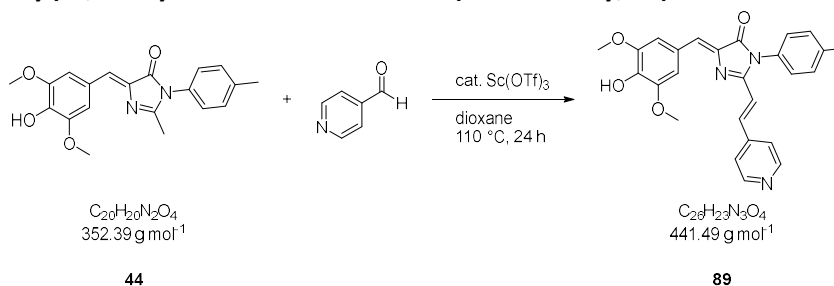
The title compound **88** was synthesized according to General procedure D on a 200  $\mu\text{mol}$  scale. After purification by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  99:1–94:6 + 5% AcOH) it was obtained as an orange-brown solid (37.7 mg, 85.3  $\mu\text{mol}$ , 43%).

$^1\text{H NMR}$  (500 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 9.28 (sbr, 1 H, OH), 8.76 (d,  $J$  = 2.1 Hz, 1 H, vinyl-Ar-2-H), 8.55 (dd,  $J$  = 4.7, 1.8 Hz, 1 H, vinyl-Ar-6-H), 8.02 (ddd,  $J$  = 8.1, 2.1, 1.8 Hz, 1 H, vinyl-Ar-4-H), 7.90 (d,  $J$  = 16.0 Hz, 1H, CCHCHAR), 7.81 (s, 2 H, Ph-2,6-H), 7.41 (dd,  $J$  = 8.1, 4.7 Hz, 1 H, vinyl-Ar-5-H), 7.38 (d,  $J$  = 8.1 Hz, 2 H, NAr-3,5-H), 7.31–7.26 (m, 2 H, NAr-2,6-H), 7.12 (s, 1 H, benzylidene-H), 6.77 (d,  $J$  = 16.0 Hz, 1 H, CCHCHAR), 3.88 (s, 6 H,  $\text{OCH}_3$ ), 2.41 (s, 3 H, NAr- $\text{CH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 168.9 (Imi-C4), 156.6 (Imi-C2), 150.5 (vinyl-Ar-C6), 149.6 (vinyl-Ar-C2), 147.9 (Ph-C3,5), 139.5 (Ph-C4), 138.1 (NAr-C4), 136.0 (Imi-C5), 135.9 (CCHCHAR), 134.0 (vinyl-Ar-C4), 130.7 (vinyl-Ar-C3), 130.5 (NAr-C1), 130.0 (NAr-C3,5), 128.0 (benzylidene-C), 127.4 (NAr-C2,6), 124.6 (Ph-C1), 123.9 (vinyl-Ar-C5), 116.0 (CCHCHAR), 110.5 (Ph-C2,6), 56.0 ( $\text{OCH}_3$ ), 20.7 (NAr- $\text{CH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{26}\text{H}_{24}\text{N}_3\text{O}_4$ ,  $[\text{M}+\text{H}]^+$ ): 442.1761, found: 442.1758.

#### 4.2.2.53 5-((Z)-4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-2-((E)-2-(pyridin-4-yl)vinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI-4Py, 89)



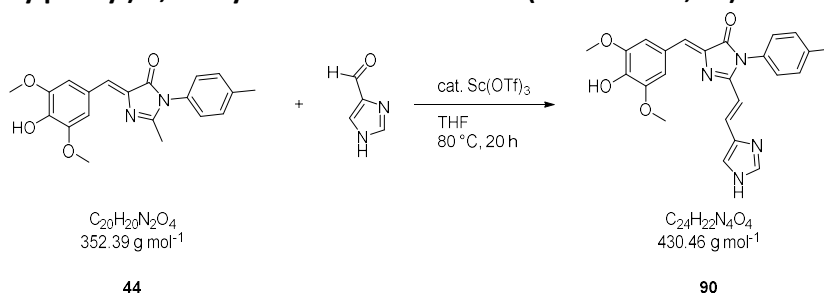
The title compound **89** was synthesized according to General procedure D on a 200  $\mu\text{mol}$  scale. After purification by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  98:2 + 5% AcOH) it was obtained as a red solid (33.7 mg, 76.3  $\mu\text{mol}$ , 38%).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 8.56 (d, *J* = 5.1 Hz, 2 H, vinyl-Ar-2,6-H), 7.76 (s, 2 H, Ph-2,6-H), 7.71 (d, *J* = 16.0 Hz, 1 H, CCHCHAR), 7.49 (d, *J* = 5.1 Hz, 2 H, vinyl-Ar-3,5-H), 7.38 (m, 2 H, NAr-3,5-H), 7.28 (m, 2 H, NAr-2,6-H), 7.11 (s, 1 H, benzylidene-H), 6.86 (d, *J* = 16.0 Hz, 1 H, CCHCHAR), 3.85 (s, 6 H, OCH<sub>3</sub>), 2.41 (s, 3 H, NAr-CH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 168.4 (Imi-C4), 154.0 (Imi-C2), 150.3 (vinyl-Ar-C2,6), 148.6 (Ph-C3,5), 142.1 (vinyl-Ar-C4), 137.9 (NAr-C4), 135.0 (CCHCHAR), 130.8 (NAr-C1), 129.9 (NAr-C3,5), 129.2 (benzylidene-C), 128.8, 128.1, 127.4 (NAr-C2,6), 125.3, 121.5 (vinyl-Ar-C3,5), 118.7 (CCHCHAR), 111.1 (Ph-C2,6), 55.8 (OCH<sub>3</sub>), 20.7 (NAr-CH<sub>3</sub>);

**HR-MS** (ESI+): *m/z* calc. (C<sub>26</sub>H<sub>24</sub>N<sub>3</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 442.1761, found: 442.1757.

#### 4.2.2.54 2-(2-(1H-Imidazol-4-yl)vinyl)-5-((Z)-4-hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI-Imi, 90)



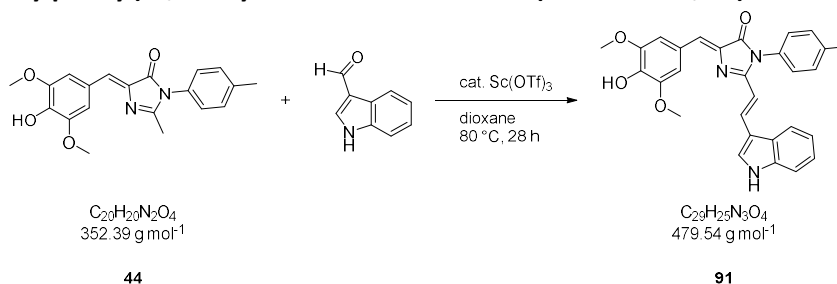
The title compound **90** was synthesized according to General procedure D on a 150  $\mu$ mol scale using THF at 80 °C as the solvent instead of dioxane. After purification by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 93:7) it was obtained as an orange solid (16.9 mg, 39.2  $\mu$ mol, 25%, 10:1 mixture of *E/Z* isomers at the newly formed C–C double bond).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 8.28–8.22 (m, 1 H, CCHCHAR), 8.04 (d, *J* = 4.5 Hz, 1 H, vinyl-Ar-2-H), 7.99 (d, *J* = 7.5 Hz, 1 H, vinyl-Ar-4-H), 7.75 (s, 2 H, Ph-2,6-H), 7.53–7.47 (m, 1 H, vinyl-Ar-7-H), 7.28–7.24 (m, 1 H, vinyl-Ar-6-H), 7.24–7.18 (m, 1 H, vinyl-Ar-5-H), 6.92–6.86 (m, 1 H, CCHCHAR), 6.85 (s, 1 H, benzylidene-H), 3.88 (s, 6 H, OCH<sub>3</sub>), 3.30 (s, 3 H, NCH<sub>3</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (125 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 170.0 (Imi-C4), 159.6 (Imi-C2)', 147.9 (Ph-C3,5), 138.9 (Ph-C4), 137.4 (vinyl-Ar-C7a), 137.3 (Imi-C5), 134.1 (CCHCHAR), 131.5 (vinyl-Ar-C2), 125.0 (vinyl-Ar-C3a), 124.9 (Ph-C1), 123.5 (benzylidene-C), 122.6 (vinyl-Ar-C6), 120.9 (vinyl-Ar-C5), 119.8 (vinyl-Ar-C4), 113.2 (vinyl-Ar-C3), 112.4 (vinyl-Ar-C7), 109.9 (Ph-C2,6), 106.7 (CCHCHAR), 55.9 (OCH<sub>3</sub>), 26.1 (NCH<sub>3</sub>);

**HR-MS** (ESI+): *m/z* calc. (C<sub>24</sub>H<sub>23</sub>N<sub>4</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 431.1714, found: 431.1712.

#### 4.2.2.55 2-((E)-2-(1H-Indol-3-yl)vinyl)-5-((Z)-4-hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI-Ind, 91)



The title compound **91** was synthesized according to General procedure D on a 150  $\mu\text{mol}$  scale at 80 °C. After purification by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{AcOH}$  100:1– $\text{CH}_2\text{Cl}_2/\text{MeOH}/\text{AcOH}$  10:1:1– $\text{MeOH}/\text{AcOH}$  10:1) it was obtained as a brownish solid (20.3 mg, 42.3  $\mu\text{mol}$ , 28%).

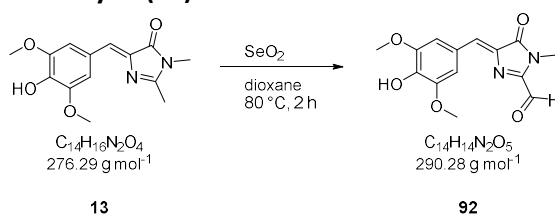
$^1\text{H}$  NMR (500 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 8.14 (d,  $J$  = 15.7 Hz, 1 H, CCHCHAr), 7.94 (s, 1 H, vinyl-Ar-2-H), 7.77 (s, 2 H, Ph-2,6-H), 7.46 (d,  $J$  = 8.1 Hz, 1 H, vinyl-Ar-4-H), 7.42 (m, 2 H, NAr-3,5-H), 7.38 (d,  $J$  = 8.1 Hz, 1 H, vinyl-Ar-7-H), 7.30 (m, 2 H, NAr-2,6-H), 7.19 (m, 1 H, vinyl-Ar-6-H), 7.09 (m, 1 H, vinyl-Ar-5-H), 6.93 (s, 1 H, benzylidene-H), 6.42 (d,  $J$  = 15.7 Hz, 1 H, CCHCHAr), 3.87 (s, 6 H,  $\text{OCH}_3$ ), 2.43 (s, 3 H, NAr- $\text{CH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) = 169.4 (Imi-C4), 157.7 (Imi-C2), 148.4 (Ph-C3,5), 138.3 (NAr-C4), 137.6 (vinyl-Ar-C7a), 134.0 (CCHCHAr), 131.8 (vinyl-Ar-C2), 131.2 (NAr-C1), 130.2 (NAr-C3,5), 127.8 (NAr-C2,6), 125.2 (benzylidene-C), 125.0 (vinyl-Ar-C3a), 122.9 (vinyl-Ar-C6), 121.2 (vinyl-Ar-C5), 119.1 (vinyl-Ar-C7), 113.2 (vinyl-Ar-C3), 112.9 (vinyl-Ar-C4), 110.2 (Ph-C2,6), 107.3 (CCHCHAr), 56.1 ( $\text{OCH}_3$ ), 21.0 (NAr- $\text{CH}_3$ );

(the  $^{13}\text{C}$  resonances of Ph-C1, Ph-C4 and Imi-C5 were not observed)

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{29}\text{H}_{26}\text{N}_3\text{O}_4$ ,  $[\text{M}+\text{H}]^+$ ): 480.1918, found: 480.1918.

#### 4.2.2.56 (Z)-4-(4-Hydroxy-3,5-dimethoxybenzylidene)-1-methyl-5-oxo-4,5-dihydro-1H-imidazole-2-carbaldehyde (**92**)



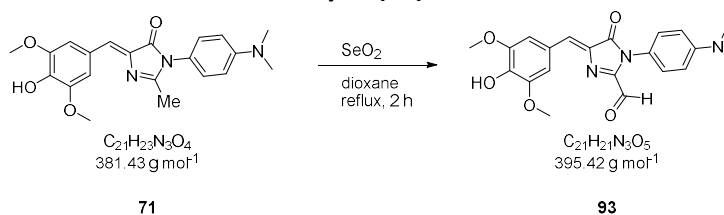
DMHBI (**13**, 414 mg, 1.50 mmol, 1.00 eq.) and  $\text{SeO}_2$  (200 mg, 1.80 mmol, 1.20 eq.) were suspended in dioxane (25 mL) and heated to reflux for 2 h. While still hot, the supernatant was decanted off from the deposited solids and the solvent was removed under reduced pressure. After purification by column chromatography ( $\text{CHCl}_3/\text{EtOH}$  98:2 + 1% AcOH) the title compound **92** was obtained as a red solid (884 mg, 3.05 mmol, 74%).

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 9.75 (s, 1 H, CHO), 7.61 (s, 2 H, Ph-2,6-H), 7.45 (s, 1 H, benzylidene-H), 6.11 (s, 1 H, OH), 3.98 (s, 6 H,  $\text{OCH}_3$ ), 3.49 (s, 3 H,  $\text{NCH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 185.4 (CHO), 170.3 (Imi-C4), 153.7 (Imi-C2), 147.5 (Ph-C3,5), 139.9 (Ph-C4), 137.7 (benzylidene-C), 137.3 (Imi-C5), 125.5 (Ph-C1), 111.2 (Ph-C2,6), 56.6 ( $\text{OCH}_3$ ), 28.1 ( $\text{NCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{15}\text{H}_{19}\text{N}_2\text{O}_6$ ,  $[\text{M}+\text{MeOH}+\text{H}]^+$ ): 323.1238, found: 323.1233.

#### 4.2.2.57 (Z)-1-(4-(Dimethylamino)phenyl)-4-(4-hydroxy-3,5-dimethoxybenzylidene)-5-oxo-4,5-dihydro-1H-imidazole-2-carbaldehyde (**93**)





Compound **71** (572 mg, 1.50 mmol, 1.00 eq.) and SeO<sub>2</sub> (200 mg, 1.80 mmol, 1.20 eq.) were suspended in dioxane (25 mL) and heated to reflux for 2 h. While still hot, the supernatant was decanted off from the deposited solids and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (CHCl<sub>3</sub>/EtOH 99:1–9:1 + 1% AcOH) afforded compound **93** as a black solid (413 mg, 1.04 mmol, 70%).

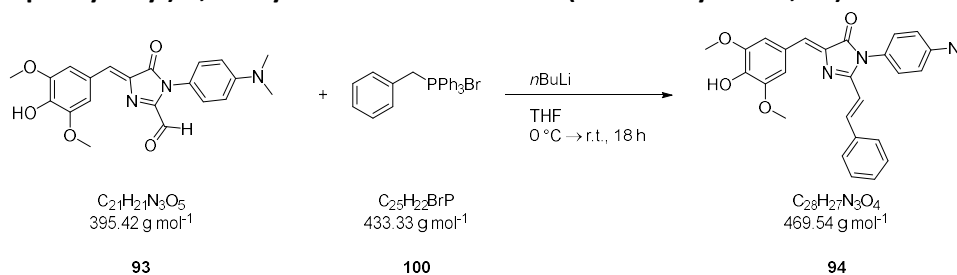
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 9.74 (s, 1 H, CHO), 7.66 (s, 2 H, Ph-2,6-H), 7.51 (s, 1 H, benzylidene-H), 7.17–7.07 (m, 2 H, NAr-H), 6.92–6.79 (m, 2H, NAr-H) 6.15 (s<sub>br</sub>, 1 H, OH), 3.99 (s, 6 H, OCH<sub>3</sub>), 3.02 (s, 6 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 183.5 (CHO), 170.0\* (Imi-C4), 152.9\* (Imi-C2), 149.9\* (NAr-C4), 147.5 (Ph-C3,5), 139.9 (Ph-C4), 137.9 (benzylidene-C), 136.8, 127.8, 125.6 (Ph-C1), 111.2 (Ph-C2,6), 56.6 (OCH<sub>3</sub>), 41.1\* (NCH<sub>3</sub>), two resonances were not observed;

HR-MS (ESI<sup>-</sup>): *m/z* calc. (C<sub>21</sub>H<sub>20</sub>N<sub>3</sub>O<sub>5</sub>, [M-H]<sup>-</sup>): 394.14084, found: 394.13946;

TLC (Hex/EtOAc 1:1): *R<sub>f</sub>* = 0.21.

#### 4.2.2.58 (Z)-3-(4-(Dimethylamino)phenyl)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-((E)-2-phenylvinyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-Styr-DMA, **94**)



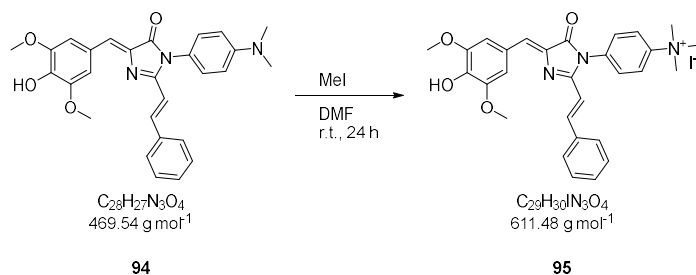
A suspension of phosphonium salt **100** (347 mg, 800 μmol, 1.00 eq.) in THF (5.4 mL) was cooled to 0 °C. *n*BuLi (2.5 M in hexane, 1.05 mL, 1.68 mmol, 2.10 eq.) was added dropwise and the resulting mixture was stirred for 30 min at the same temperature. Afterwards, HBI derivative **93** (316 mg, 800 μmol, 1.00 eq.) was added as a solid in four portions over the course of 30 min. The reaction was stirred at ambient temperature until TLC showed no further changes (18 h) and then quenched by addition of sat. aq. NH<sub>4</sub>Cl (10 mL) and H<sub>2</sub>O (10 mL). The mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub> (4×20 mL) and the combined organic phases were dried over MgSO<sub>4</sub>. After removal of the solvent under reduced pressure the residue was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/acetone 95:5–90:10 + 1% AcOH) to afford the title compound as a brown solid (139 mg, 295 μmol, 37%).

<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 9.27 (s<sub>br</sub>, 1 H, OH), 7.89 (d, *J* = 16.0 Hz, 1H, CCHCHAR), 7.81 (s, 2 H, Ph-2,6-H), 7.58–7.52 (m, 2 H, vinyl-Ar-2,6-H), 7.43–7.37 (m, 3 H, vinyl-Ar-3,5-H, vinyl-Ar-4-H), 7.22–7.14 (m, 2 H, NAr-2,6-H), 7.07 (s, 1 H, benzylidene-H), 6.88–6.82 (m, 2 H, NAr-3,5-H), 6.61 (d, *J* = 16.0 Hz, 1 H, CCHCHAR), 3.88 (s, 6 H, OCH<sub>3</sub>), 2.98 (s, 6 H, N(CH<sub>3</sub>)<sub>2</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 169.6 (Imi-C4), 157.9 (Imi-C2), 150.2 (NAr-C4), 147.9 (Ph-C3,5), 139.4 (CCHCHAR), 138.8 (Ph-C4), 136.6 (Imi-C5), 134.8 (vinyl-Ar-C1), 130.1 (vinyl-Ar-C4), 129.1 (vinyl-Ar-C3,5), 128.4 (NAr-C2,6), 127.9 (vinyl-Ar-C2,6), 127.0 (benzylidene-C), 125.0 (Ph-C1), 121.2 (NAr-C1), 114.1 (CCHCHAR), 112.4 (NAr-C3,5), 110.2 (Ph-C2,6), 55.9 (OCH<sub>3</sub>), 40.1 (N(CH<sub>3</sub>)<sub>2</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>28</sub>H<sub>27</sub>N<sub>3</sub>NaO<sub>4</sub>, [M+Na]<sup>+</sup>): 492.18938, found: 492.18866.

**4.2.2.59 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-((E)-2-phenylvinyl)-3-(4-(trimethylammonium)phenyl)-3,5-dihydro-4H-imidazol-4-one iodide (DMHBI-Styr<sup>+</sup>, 95)**



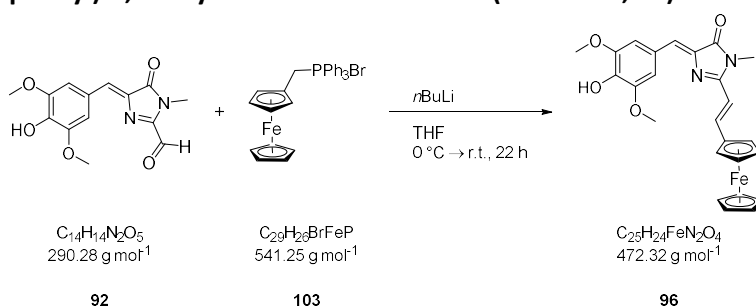
HBI derivative **94** (70.4 mg, 150  $\mu\text{mol}$ , 1.00 eq.) and methyl iodide (0.1 mL, 150 mmol, 1.00 eq.) were dissolved in DMF (3 mL) and stirred at ambient temperature for 24 h. Removal of the solvent under reduced pressure afforded the pure product **95** as a brown solid (91.7 mg, 150  $\mu\text{mol}$ , > 99%).

<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 9.39 (s, 1 H, OH), 8.22–8.14 (m, 2 H, NAr-3,5-H), 7.94 (d,  $J$  = 15.8 Hz, 1 H, CCHCHAR), 7.84 (s, 2 H, Ph-2,6-H), 7.78–7.73 (m, 2 H, NAr-2,6-H), 7.63–7.58 (m, 2 H, vinyl-Ar-2,6-H), 7.46–7.39 (m, 3 H, vinyl-Ar-3,5-H, vinyl-Ar-4-H), 7.16 (s, 1 H, benzylidene-H), 6.69 (d,  $J$  = 15.8 Hz, 1 H, CCHCHAR), 3.89 (s, 6 H, OCH<sub>3</sub>), 3.69 (s, 9 H, N(CH<sub>3</sub>)<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) = 168.7 (Imi-C4), 156.3 (Imi-C2), 147.9 (Ph-C3,5), 146.6 (NAr-C4), 139.8 (CCHCHAR), 139.2 (Ph-C4), 135.9 (Imi-C5), 134.8 (vinyl-Ar-C1), 134.6 (NAr-C1), 130.3 (vinyl-Ar-C4), 129.1 (vinyl-Ar-C3,5), 129.0 (NAr-C2,6), 128.1 (vinyl-Ar-C2,6), 128.0 (benzylidene-C), 124.8 (Ph-C1), 122.0 (NAr-C3,5), 113.8 (CCHCHAR), 110.4 (Ph-C2,6), 56.6 (N(CH<sub>3</sub>)<sub>3</sub>), 56.0 (OCH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>):  $m/z$  calc. (C<sub>29</sub>H<sub>30</sub>N<sub>3</sub>O<sub>4</sub>, [M-I]<sup>+</sup>): 484.22308, found: 484.22400.

**4.2.2.60 2-((E)-2-(Ferrocenyl)vinyl)-5-((Z)-4-hydroxy-3,5-dimethoxybenzylidene)-3-(4-methylphenyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-Fc, 96)**



A suspension of phosphonium salt **103** (217 mg, 400  $\mu\text{mol}$ , 1.00 eq.) in THF (3 mL) was cooled to 0 °C. *n*BuLi (2.5 M in hexane, 0.34 mL, 840  $\mu\text{mol}$ , 2.10 eq.) was added dropwise and the resulting mixture was stirred for 30 min at the same temperature. Afterwards, HBI derivative **92** (116 mg, 400  $\mu\text{mol}$ , 1.00 eq.) was added as a solid in three portions over the course of 30 min. The reaction was stirred at ambient temperature until TLC showed no further changes (22 h) and then quenched by addition of sat. aq. NH<sub>4</sub>Cl (3 mL) and H<sub>2</sub>O (10 mL). The mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub> (4×30 mL) and the combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>. After removal of the solvent under reduced pressure the residue was purified by column chromatography (CHCl<sub>3</sub>/EtOH 98:2–94:6 + 1% AcOH) and then filtered over silica to remove residual Ph<sub>3</sub>PO (eluting with

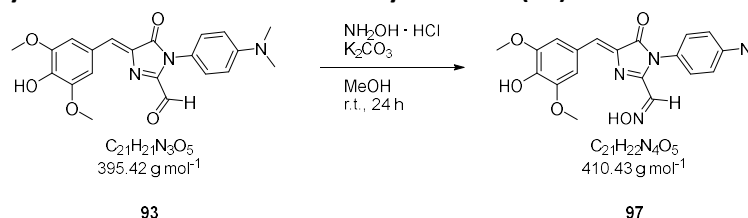
pentane/Et<sub>2</sub>O followed by CHCl<sub>3</sub>). The title compound was obtained as a red-brown solid (23.6 mg, 50.0 μmol, 12%).

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.89 (d, J = 15.5 Hz, 1 H, CCHCHFc), 7.64 (s, 2 H, Ph-2,6-H), 7.05 (s, 1 H, benzylidene-H), 6.39 (d, J = 15.5 Hz, 1 H, CCHCHFc), 5.89 (s, 1 H, OH), 4.57 (t, J = 1.9 Hz, 2 H, vinyl-Fc-2,5-H), 4.50 (t, J = 1.9 Hz, 2 H, vinyl-Fc-3,4-H), 4.20 (s, 5 H, Fc-H), 4.00 (s, 6 H, OCH<sub>3</sub>), 3.28 (s, 3 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (125 MHz, CDCl<sub>3</sub>): δ (ppm) = 171.0 (Imi-C4), 159.1 (Imi-C2), 147.2 (Ph-C3,5), 142.2 (CCHCHFc), 138.5 (Imi-C5), 137.2 (Ph-C4), 126.9 (Ph-C1), 126.1 (benzylidene-C), 109.7 (CCHCHFc), 109.5 (Ph-C2,6), 80.2 (vinyl-Fc-C1), 71.4 (vinyl-Fc-C3,4), 69.9 (Fc-C), 68.6 (vinyl-Fc-C2,5), 56.4 (OCH<sub>3</sub>), 26.8 (NCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>25</sub>H<sub>25</sub>FeN<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>): 473.1159, found: 473.1147.

#### 4.2.2.61 (Z)-1-(4-(Dimethylamino)phenyl)-4-(4-hydroxy-3,5-dimethoxybenzylidene)-5-oxo-4,5-dihydro-1H-imidazole-2-carbaldehyde oxime (97)



Aldehyde **93** (198 mg, 500 μmol, 1.00 eq.), hydroxylamine hydrochloride (41.7 mg, 600 μmol, 1.20 eq.) and K<sub>2</sub>CO<sub>3</sub> (38.0 mg, 275 μmol, 0.55 eq.) were suspended in MeOH (1.25 mL) and stirred at ambient temperature for 24 h. Afterwards, the solvent was removed under reduced pressure. Purification of the residue by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>-CH<sub>2</sub>Cl<sub>2</sub>/MeOH 98:2 + 1% AcOH) afforded the oxime **97** (75.1 mg, 183 μmol, 37%).

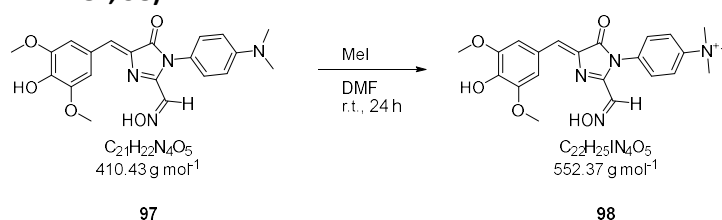
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 12.38 (s<sub>br</sub>, 1 H, NOH), 9.42 (s<sub>br</sub>, 1 H Ph-4-OH), 7.79 (s, 1 H, CHN), 7.74 (s, 2 H, Ph-2,6-H), 7.17 (s, 1 H, benzylidene-H), 7.13–7.07 (m, 2 H, NAr-3,5-H), 6.77–6.73 (m 2 H, NAr-2,6-H), 3.82 (s, 6 H, OCH<sub>3</sub>), 2.94 (s, 6 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 169.6 (Imi-C5), 154.0 (Imi-C2), 150.0 (NAr-C4), 147.9 (Ph-C3,5), 140.4 (CHN), 139.6 (Ph-C4), 136.0 (Imi-C4), 129.8 (benzylidene-C), 128.4 (NAr-C2,6), 124.5 (Ph-C1), 122.5 (NAr-C1), 112.1 (NAr-C3,5), 110.7 (Ph-C2,6), 56.0 (OCH<sub>3</sub>), 40.1 (NCH<sub>3</sub>);

HR-MS (ESI+): *m/z* calc. (C<sub>21</sub>H<sub>22</sub>N<sub>4</sub>NaO<sub>5</sub>, [M+H]<sup>+</sup>): 433.14879, found: 433.14870;

TLC (Toluene/Acetone 7:3 + 1% AcOH): *R<sub>f</sub>* = 0.28.

#### 4.2.2.62 (Z)-4-(4-Hydroxy-3,5-dimethoxybenzylidene)-5-oxo-1-(4-(trimethylammonium)phenyl)-4,5-dihydro-1H-imidazole-2-carbaldehyde oxime iodide (DMHBO<sup>+</sup>, 98)



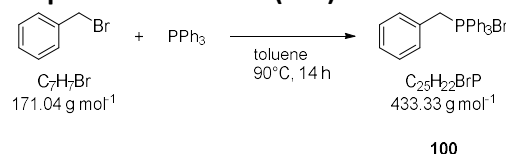
Compound **97** (51.3 mg, 125  $\mu\text{mol}$ , 1.00 eq.) and methyl iodide (78  $\mu\text{L}$ , 1.25 mmol, 10.0 eq.) were dissolved in DMF (2.5 mL) and stirred at ambient temperature for 24 h. Removal of the solvent under reduced pressure afforded the product **98** (70.0 mg, 125  $\mu\text{mol}$ , > 99%).

$^1\text{H NMR}$  (400 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  (ppm) = 12.35 (s, 1 H, NOH), 9.50 ( $s_{\text{br}}$ , 1 H Ph-4-OH), 8.12–8.07 (m, 2 H, NAr-2,6-H), 7.91 (s, 1 H, CHN), 7.77 (s, 2 H, Ph-2,6-H), 7.71–7.65 (m, 2 H, NAr-3,5-H), 7.25 (s, 1 H, benzylidene-H), 3.82 (s, 6 H,  $\text{OCH}_3$ ), 3.65 (s, 9 H,  $\text{NCH}_3$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  (ppm) = 168.7 (Imi-C5), 152.5 (Imi-C2), 148.0 (Ph-C3,5), 146.4 (NAr-C4), 140.5 (CHN), 139.9 (Ph-C4), 135.8 (NAr-C1), 135.4 (Imi-C4), 130.8 (benzylidene-C), 129.3 (NAr-C2,6), 124.3 (Ph-C1), 121.3 (NAr-C3,5), 110.9 (Ph-C2,6), 56.6 ( $\text{NCH}_3$ ), 56.1 ( $\text{OCH}_3$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{22}\text{H}_{25}\text{N}_4\text{O}_5$ ,  $[\text{M-I}]^+$ ): 425.18195, found: 425.18280.

#### 4.2.2.63 Benzyltriphenylphosphonium bromide (**100**)



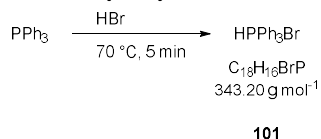
Benzyl bromide (5.94 mL, 50.0 mmol, 1.00 eq.) and triphenylphosphine (13.1 g, 50.0 mmol, 1.00 eq.) were suspended in toluene (125 mL) and stirred at 90 °C for 15 h. After cooling to ambient temperature, the precipitate was filtered off and washed with  $\text{Et}_2\text{O}$  (3×20 mL) to afford the title compound **100** as a white solid (19.5 g, 45.1 mmol, 90%). Spectral data matched those reported previously.<sup>147</sup>

$^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) =  $\delta$  7.80–7.70 (m, 3 H, Ph-4-H), 7.76–7.65 (m, 6 H, Ph-2,6-H), 7.66–7.56 (m, 6 H, Ph-3,5-H), 7.25–7.15 (m, 1 H, Bn-4-H), 7.14–7.06 (m, 2 H, Bn-3,5-H), 7.11–7.03 (m, 2 H, Bn-2,6\*H), 5.35 (d,  $J$  = 14.4 Hz, 2 H,  $\text{CH}_2$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 135.1 (d,  $J$  = 3.1 Hz, Ph-C4), 134.5 (d,  $J$  = 9.8 Hz, Ph-C2,6), 131.6 (d,  $J$  = 5.5 Hz, Bn-C2,6), 130.3 (d,  $J$  = 12.6 Hz, Ph-C3,5), 128.9 (d,  $J$  = 3.4 Hz, Bn-C3,5), 128.5 (d,  $J$  = 3.9 Hz, Bn-C4), 127.2 (d,  $J$  = 8.6 Hz, Bn-C1), 117.9 (d,  $J$  = 85.7 Hz, Ph-C1), 30.9 (d,  $J$  = 47.0 Hz,  $\text{CH}_2$ );

$^{31}\text{P}\{^1\text{H}\}$  NMR (162 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 23.14 (P);

#### 4.2.2.64 Triphenylphosphonium bromide (**101**)



Triphenylphosphine (13.1 g, 50.0 mmol, 1.00 eq.) was suspended in 48% aq. HBr (35 mL) and stirred at 70 °C for 5 min. The resulting clear solution was cooled to ambient temperature and extracted with  $\text{CHCl}_3$  (3×15 mL). The combined organic phases were dried over  $\text{Na}_2\text{SO}_4$ . Evaporation of the solvent under reduced pressure afforded the title compound **101** as a white solid (15.6 g, 45.5 mmol, 91%). Spectral data matched those reported previously.<sup>145</sup>

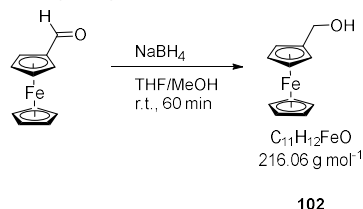
$^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 12.15 ( $s_{\text{br}}$ , 1 H, PH), 7.74–7.66 (m, 6 H, Ar-H), 7.67–7.54 (m, 3 H, Ar-4-H), 7.57–7.47 (m, 6 H, Ar-H);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 134.1 (d,  $J$  = 13.6 Hz, Ph-C4), 133.0 (Ph-C2,6), 129.8 (d,  $J$  = 11.0 Hz, Ph-C3,5), 123.7 (Ph-C1);

$^{31}\text{P}$  NMR (203 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = -9.07 (PH);

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{18}\text{H}_{16}\text{P}$ ,  $[\text{M}-\text{Br}]^+$ ): 263.0984, found: 263.0987.

#### 4.2.2.65 (Hydroxymethyl)ferrocene (**102**)



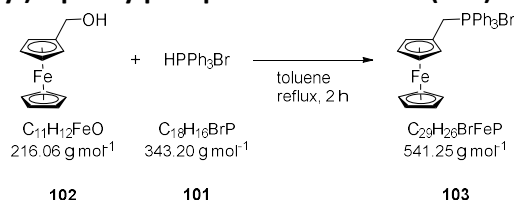
Ferrocene carbaldehyde (562 mg, 2.63 mmol, 1.00 eq.) was dissolved in a mixture of THF (25 mL) and MeOH (5 mL) at ambient temperature.  $\text{NaBH}_4$  (99.9 mg, 2.64 mmol, 1.01 eq.) was added in 5 portions over the course of 30 min; stirring was continued for 30 min. After removal of the solvent, the residue was taken up in EtOAc (25 mL), washed with  $\text{H}_2\text{O}$  ( $2 \times 10$  mL) and brine (10 mL). The organic phase was dried over  $\text{Na}_2\text{SO}_4$  and solvent was evaporated under reduced pressure to afford the title compound as a yellow, crystalline solid (554 mg, 2.56 mmol, 97%). Spectral data matched those reported previously.<sup>144</sup>

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 4.33 (s, 2 H,  $\text{CH}_2$ ), 4.24 (t,  $J$  = 1.9 Hz, 2 H, Fc-H), 4.18 (s, 5 H, Fc-H), 4.18 (t,  $J$  = 1.9 Hz, 2 H, Fc-H), 1.58 (s, 1 H, OH);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 88.6 (Fc-C1), 68.4 (Fc-C), 68.4 (Fc-C), 68.0 (Fc-C), 60.9 ( $\text{CH}_2$ );

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{11}\text{H}_{10}\text{Fe}$ ,  $[\text{M}-\text{H}_2\text{O}]^+$ ): 199.0205, found: 199.0211.

#### 4.2.2.66 (Ferrocenylmethyl)triphenylphosphonium bromide (**103**)



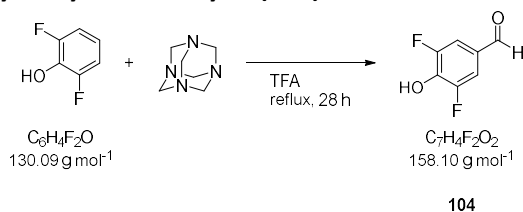
A suspension of (hydroxymethyl)ferrocene (**102**, 350 mg, 1.62 mmol, 1.00 eq.) and triphenylphosphonium bromide (**101**, 556 mg, 1.62 mmol, 1.00 eq.) in toluene (100 mL) was heated to reflux with a Dean-Stark trap for 2 h. After cooling to ambient temperature, the precipitate was filtered off and washed with cold  $\text{Et}_2\text{O}$  (10 mL) to afford the title compound **103** as a yellow solid (620 mg, 1.15 mmol, 71%). Spectral data matched those reported previously.<sup>146</sup>

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.73 (m, 9 H), Ph-H, 7.64 (m, 6 H, Ph-H), 5.08 (s, 2 H), 4.37 (s, 5 H, Fc-H), 4.04 (s, 2 H), 3.97 (s, 2 H);

$^{13}\text{C}\{^1\text{H}\}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 135.0 (d,  $J$  = 2.9 Hz, Ph-C4), 134.8 (d,  $J$  = 9.8 Hz, Ph-C2,6), 130.4 (d,  $J$  = 12.4 Hz, Ph-C3,5), 118.3 (d,  $J$  = 84.8 Hz, Ph-C1), 73.8 (Fc-C1), 71.1 (Fc-C), 70.4 (Fc-C), 68.9 (Fc-C), 29.3 (d,  $J$  = 42.5 Hz,  $\text{CH}_2$ );

$^{31}\text{P}\{^1\text{H}\}$  NMR (203 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 19.31 (P);

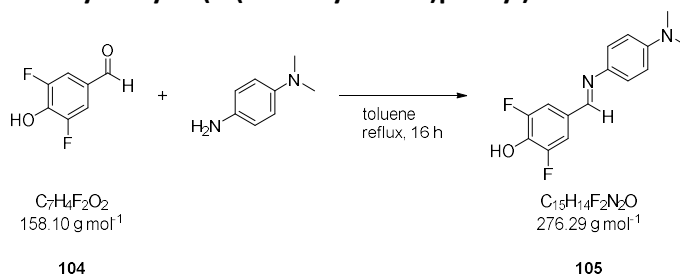
HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{29}\text{H}_{26}\text{FeP}$ ,  $[\text{M}-\text{Br}]^+$ ): 461.1116, found: 461.1118.

4.2.2.67 3,5-Difluoro-4-hydroxybenzaldehyde (**104**)

A solution of 2,6-difluorophenol (4.65 g, 35.7 mmol, 1.00 eq.) and hexamethylene tetramine (5.01 g, 35.7 mmol, 1.00 eq.) in trifluoroacetic acid (36 mL) was heated to reflux for 28 h. Afterwards, the volatiles were removed under reduced pressure, the residue was taken up in CH<sub>2</sub>Cl<sub>2</sub> (75 mL) and washed with sat. aq. NaHCO<sub>3</sub> (2×50 mL). The aqueous phase was first neutralized with 1 M NaOH, then brought to pH 1 with conc. HCl and back-extracted with CH<sub>2</sub>Cl<sub>2</sub> (3×50 mL). The combined organic phases were dried over MgSO<sub>4</sub> and the solvent was evaporated under reduced pressure to afford the product **104** as a pale yellow solid (4.37 g, 27.6 mmol, 77%). Spectral data matched those reported previously.<sup>149</sup>

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 9.81 (t, *J* = 1.9 Hz, 1 H, CHO), 7.54–7.44 (m, 2 H, Ph-2,6-H);

<sup>19</sup>F{<sup>1</sup>H} NMR (376 MHz, CDCl<sub>3</sub>): δ (ppm) = –133.0 (Ph-3,5-F).

4.2.2.68 3,5-Difluoro-4-hydroxy-*N*-(4-(dimethylamino)phenyl)benzaldimine (**105**)

The title compound **105** was synthesized according to General procedure B on a 20.0 mmol scale. Dark brown solid (5.48 g, 19.8 mmol, > 99%) that was sufficiently pure for all further reactions. For analysis, the crude product was recrystallized from toluene. The product is prone to decomposition upon exposure to moisture.

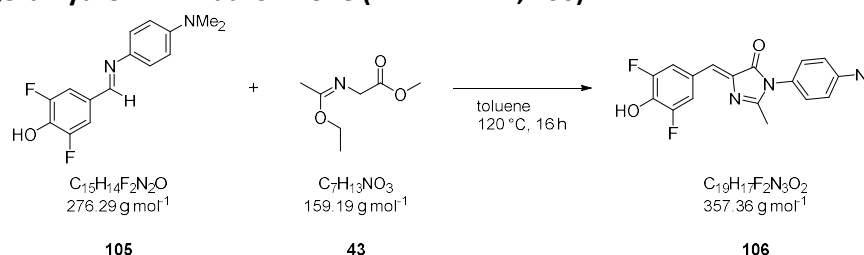
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 10.59 (s<sub>br</sub>, 1 H, OH), 8.50 (s, 1 H, CHN), 7.59–7.47 (m, 2 H, Ph-2,6-H), 7.28–7.19 (m, 2 H, NAr-2,6-H), 6.79–6.70 (m, 2 H, NAr-3,5-H), 2.92 (s, 6 H, NCH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 152.9 (CHN), 152.3 (d, *J* = 242.9 Hz, Ph-C3,5), 149.4 (NAr-C4), 139.3 (NAr-C1), 135.9 (Ph-C4), 127.8 (Ph-C1), 122.4 (NAr-C2,6), 112.6 (NAr-C3,5), 111.1 (d, *J* = 18.0 Hz, Ph-C2,6), 40.2 (NCH<sub>3</sub>);

<sup>19</sup>F{<sup>1</sup>H} NMR (376 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = –131.9 (Ph-3,5-F);

HR-MS (ESI+): *m/z* calc. (C<sub>15</sub>H<sub>15</sub>F<sub>2</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>): 277.11524, found: 277.11503.

#### 4.2.2.69 (Z)-3-(4-(Dimethylamino)phenyl)-5-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DFHBI-DMA, **106**)



The title compound **106** was synthesized according to General procedure C on a 2.00 mmol scale. After purification by column chromatography (hexane/EtOAc 6:4–4:6) it was obtained as an solid (108 mg, 0.27 mmol, 27%).

A mixture of imine **12** (553 mg, 2.00 mmol, 1.00 eq.) and imidate **3** (382 mg, 2.40 mmol, 1.20 eq.) in toluene (2 mL) was stirred at  $120^\circ\text{C}$  for 16 h. Afterwards, the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (hexane/EtOAc 6:4–4:6) affording compound **13** as an orange crystalline solid (256 mg, 0.72 mmol, 36%).

$^1\text{H NMR}$  (400 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  (ppm) = 10.98 (s<sub>br</sub>, 1 H, OH), 8.08–7.96 (m, 2 H, Ph-2,6-H), 7.18–7.11 (m, 2 H, NAr-2,6-H), 6.97 (s, 1 H, benzylidene-H), 6.83–6.75 (m, 2 H, NAr-3,5-H), 3.95 (s, 6 H, NCH<sub>3</sub>), 2.17 (s, 3 H, CCH<sub>3</sub>);

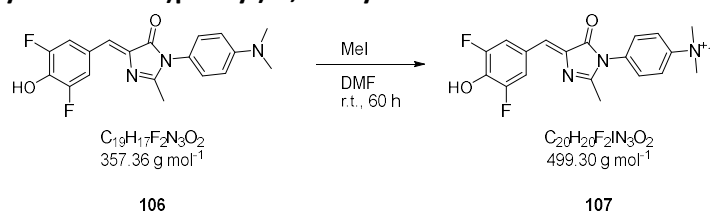
$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  (ppm) = 169.4 (Imi-C4), 163.8 (Imi-C2), 151.8 (d,  $J = 241.8$  Hz, Ph-C3,5), 150.3 (NAr-C4), 138.0 (Imi-C5), 135.8 (Ph-C4), 128.2 (NAr-C2,6), 124.7 (Ph-C1), 123.3 (benzylidene-C), 121.4 (NAr-C1), 115.4–115.0 (m, Ph-C2,6), 112.3 (NAr-C3,5), 40.1 (NCH<sub>3</sub>), 16.2 (CCH<sub>3</sub>);

$^{19}\text{F}\{^1\text{H}\}$  NMR (376 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  (ppm) =  $-132.3$  (Ph-3,5-F);

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{19}\text{H}_{18}\text{F}_2\text{N}_3\text{O}_2$ ,  $[\text{M}+\text{H}]^+$ ): 358.13616, found: 358.13516;

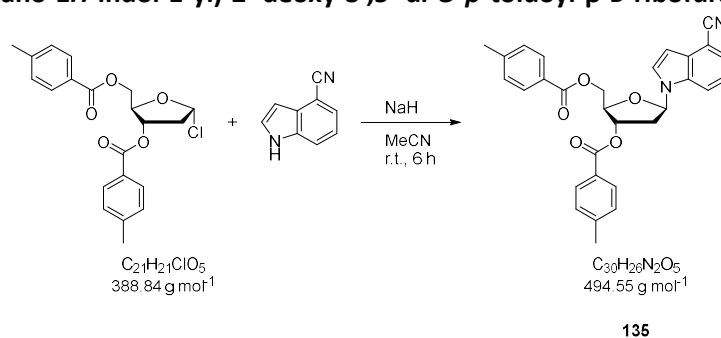
TLC (hexane/EtOAc 4:6):  $R_f = 0.49$ .

#### 4.2.2.70 (Z)-5-(3,5-Difluoro-4-hydroxybenzylidene)-2-methyl-3-(4-(trimethylammonium)phenyl)-3,5-dihydro-4H-imidazol-4-one iodide (DFHBI<sup>+</sup>, **107**)



Starting from compound **106**, DFHBI<sup>+</sup> **107** was synthesized analogously to the dimethoxy derivative **72** (see section 4.2.2.36).

## 4.2.3 Preparation of 4-Cyanoindole phosphoramidite building blocks

4.2.3.1 1'-(4-Cyano-1*H*-indol-1-yl)-2'-deoxy-3',5'-di-*O*-*p*-toluoyl-β-*D*-ribofuranose (135)

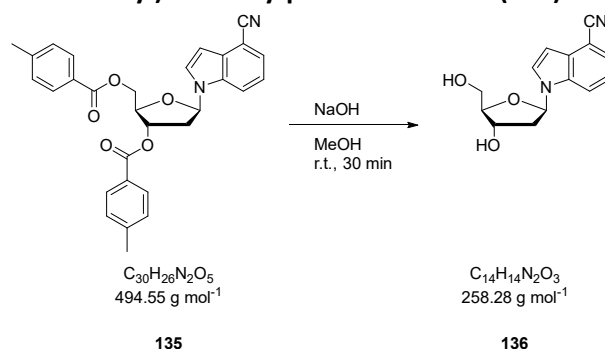
NaH (111 mg, 2.78 mmol, 1.20 eq., 60% dispersion in mineral oil) was added to a solution of 4-cyanoindole (330 mg, 2.32 mmol, 1.00 eq.) in dry MeCN (50 mL). The mixture was stirred at ambient temperature for 30 min, then 1-chloro-2-deoxy-3',5'-di-*O*-*p*-toluoyl-α-*D*-ribofuranose (1.08 g, 2.78 mmol, 1.20 eq.) was added in three portions over the course of 30 min. After 5 h, the mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> (20 mL), filtered over Celite and evaporated to dryness. Purification of the residue by column chromatography (hexane/EtOAc 90:10–70:30) afforded the tolyl-protected deoxyribonucleoside **135** as a white foam (1.13 g, 2.28 mmol, 98%). Analytical data matched those reported previously.<sup>209</sup>

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.02–7.93 (m, 2 H, 3'-*O*-toluoyl-2,6-H), 7.93–7.85 (m, 2 H, 5'-*O*-toluoyl-2,6-H), 7.75 (ddd, *J* = 8.5, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.47 (dd, *J* = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.45 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.33–7.27 (m, 2 H, 3'-*O*-toluoyl-3,5-H), 7.27–7.21 (m, 2 H, 5'-*O*-toluoyl-3,5-H), 7.16 (dd, *J* = 8.5, 7.4 Hz, 1 H, Ind-6-H), 6.73 (dd, *J* = 3.4, 0.9 Hz, 1 H, Ind-3-H), 6.46 (dd, *J* = 8.3, 5.6 Hz, 1 H, 1'-H), 5.74 (ddd, *J* = 6.5, 2.8, 2.6 Hz, 1 H, 3'-H), 4.70 (dd, *J* = 11.9, 3.6 Hz, 1 H, 5'-H<sub>a</sub>), 4.63 (dd, *J* = 11.9, 3.6 Hz, 1 H, 5'-H<sub>b</sub>), 4.60 (ddd, *J* = 3.6, 3.6, 2.8 Hz, 1 H, 4'-H), 2.87 (ddd, *J* = 14.1, 8.3, 6.5 Hz, 1 H, 2'-H<sub>eq</sub>), 2.70 (ddd, *J* = 14.1, 5.6, 2.6 Hz, 1 H, 2'-H<sub>ax</sub>), 2.45 (s, 3 H, 3'-*O*-toluoyl-4-CH<sub>3</sub>), 2.44 (s, 3 H, 5'-*O*-toluoyl-4-CH<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 166.3 (5'-*O*-CO), 166.1 (3'-*O*-CO), 144.7 (3'-*O*-toluoyl-C4), 144.4 (5'-*O*-toluoyl-C4), 135.6 (Ind-C7a), 130.7 (Ind-C3a), 129.9 (3'-*O*-toluoyl-C2,6), 129.8 (5'-*O*-toluoyl-C2,6), 129.5 (3'-*O*-toluoyl-C3,5), 129.4 (5'-*O*-toluoyl-C3,5), 126.84 (Ind-C2), 126.79 (5'-*O*-toluoyl-C1), 126.5 (3'-*O*-toluoyl-C1), 125.8 (Ind-C5), 121.9 (Ind-C6), 118.6 (CN), 115.0 (Ind-C7), 103.7 (Ind-C4), 102.5 (Ind-C3), 85.9 (C1'), 82.2 (C4'), 74.9 (C3'), 64.1 (C5'), 38.2 (C2'), 21.90 (3'-*O*-toluoyl-4-CH<sub>3</sub>), 21.86 (5'-*O*-toluoyl-4-CH<sub>3</sub>);

HR-MS (ESI<sup>+</sup>): *m/z* calc. (C<sub>30</sub>H<sub>26</sub>N<sub>2</sub>NaO<sub>5</sub>, [M+Na]<sup>+</sup>): 517.17339, found: 517.17514;

TLC (toluene/acetone 95:5): *R*<sub>f</sub> = 0.74.

4.2.3.2 1'-(4-Cyano-1*H*-indol-1-yl)-2'-deoxy-β-*D*-ribofuranose (136)



Compound **135** (1.00 g, 2.02 mmol, 1.00 eq.) was dissolved in 1% methanolic NaOH (60 mL). After 30 min the solution was evaporated to dryness. The residue was adsorbed onto silica gel and purified by column chromatography (hexane/EtOAc 60:40–EtOAc–EtOAc/MeOH 95:5) to obtain the unprotected deoxyribonucleoside **136** as an off-white foam (504 mg, 1.94 mmol, 96%). Analytical data matched those reported previously.<sup>209</sup>

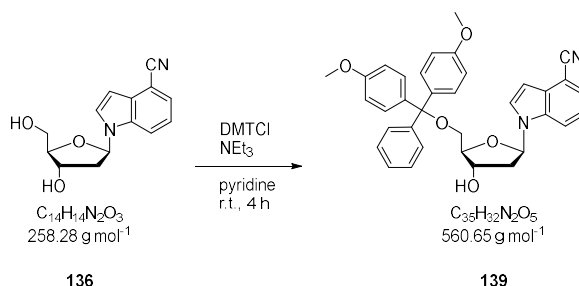
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 8.02 (ddd, *J* = 8.5, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.92 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.59 (dd, *J* = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.30 (dd, *J* = 8.5, 7.4 Hz, 1 H, Ind-6-H), 6.67 (dd, *J* = 3.4, 0.9 Hz, 1 H, Ind-C3), 6.46 (dd, *J* = 7.6, 6.0 Hz, 1 H, 1'-H), 5.33 (d, *J* = 4.2 Hz, 1 H, 3'-OH), 4.94 (dd, *J* = 5.6, 5.2 Hz, 1 H, 5'-OH), 4.42–4.32 (m, 1 H, 3'-H), 3.84 (ddd, *J* = 4.5, 4.5, 3.0 Hz, 1 H, 4'-H), 3.56 (ddd, *J* = 11.6, 5.6, 4.5 Hz, 1 H, 5'-H<sub>a</sub>), 3.50 (ddd, *J* = 11.6, 5.2, 4.5 Hz, 1 H, 5'-H<sub>b</sub>), 2.57–2.45 (m, 1 H, 2'-H<sub>eq</sub>), 2.27 (ddd, *J* = 13.2, 6.0, 3.2 Hz, 1 H, 2'-H<sub>ax</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 135.3 (Ind-C7a), 129.4 (Ind-C3a), 129.0 (Ind-C2), 125.3 (Ind-C5), 121.6 (Ind-C6), 118.5 (CN), 116.0 (Ind-C7), 101.6 (Ind-C4), 100.4 (Ind-C3), 87.3 (C4'), 84.8 (C1'), 70.6 (C3'), 61.7 (C5'), 39.9 (C2');

HR-MS (ESI+): *m/z* calc. (C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>NaO<sub>3</sub>, [M+Na]<sup>+</sup>): 281.08966, found: 281.08884;

TLC (CH<sub>2</sub>Cl<sub>2</sub>): *R*<sub>f</sub> = 0.10.

#### 4.2.3.3 1'-(4-Cyano-1H-indol-1-yl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxy-β-d-ribofuranose (**139**)



Compound **136** (111 mg, 431 μmol, 1.00 eq.) was coevaporated with pyridine (3×5 mL) and dissolved in anhydrous pyridine (12 mL) containing NEt<sub>3</sub> (65.4 mg, 646 μmol, 1.50 eq.). 4,4'-Dimethoxytrityl chloride (219 mg, 646 μmol, 1.50 eq.) was added in three portions over the course of 1 h and stirring was continued at ambient temperature for 3 h. The reaction was quenched by addition of MeOH (1 mL) and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 90:10–50:50 + 1% NEt<sub>3</sub>) afforded the DMT-protected deoxyribonucleoside **139** as a pale yellow foam (158 mg, 281 μmol, 65%). Analytical data match those reported previously.<sup>209</sup>

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.76 (ddd, *J* = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.47 (dd, *J* = 7.4, 0.8 Hz, 1 H, Ind-5-H), 7.40 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.40–7.37 (m, 2 H, Ph-2'',6''-H), 7.30–7.26 (m, 4 H, Ph-2,2',6,6'-H), 7.26–7.17 (m, 2 H, Ph-3'',5''-H), 7.17 (dd, *J* = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.81–6.76 (m, 4 H, Ph-3,3',5,5'-H), 6.70 (dd, *J* = 3.4, 0.8 Hz, 1 H, Ind-3-H), 6.39 (dd, *J* = 7.3, 6.0 Hz, 1 H, 1'-H), 4.64 (ddd, *J* = 6.5, 3.6, 3.5 Hz, 1 H, 3'-H), 4.09 (ddd, *J* = 4.6, 4.2, 3.5 Hz, 1 H, 4'-H), 3.78 (2 s, 6 H, OCH<sub>3</sub>), 3.38 (dd, *J* = 10.2, 4.2 Hz, 1 H, 5'-H<sub>a</sub>), 3.32 (dd, *J* = 10.2, 4.6 Hz, 1 H, 5'-H<sub>b</sub>), 2.64 (ddd, *J* = 13.6, 7.3, 6.5 Hz, 1 H, 2'-H<sub>eq</sub>), 2.44 (ddd, *J* = 13.6, 6.0, 3.6 Hz, 1 H, 2'-H<sub>ax</sub>);

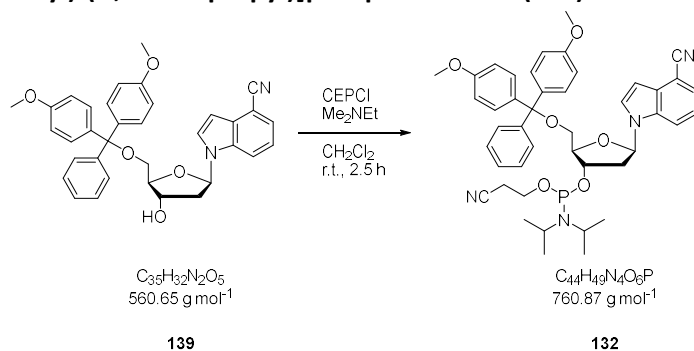
<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 158.7 (Ph-C4,4'), 144.6 (Ph-C1''), 135.8 (Ind-C7a), 135.7 (Ph-C1,1'), 130.6 (Ind-C3a), 130.2 (Ph-C2,2',6,6'), 128.3 (Ph-C2'',6''), 128.1 (Ph-C3'',5''), 127.2 (Ph-C4''), 127.1 (Ind-C2), 125.7 (Ind-C5), 121.8 (Ind-C6), 118.7 (CN), 115.2 (Ind-C7), 113.3 (Ph-

C3,3',5,5'), 103.4 (Ind-C4), 102.0 (Ind-C3), 86.8 (5'-OC), 85.5 (C4'), 85.3 (C1'), 72.8 (C3'), 63.9 (C5'), 55.4 (OCH<sub>3</sub>), 40.4 (C2');

**HR-MS** (ESI+): *m/z* calc. (C<sub>35</sub>H<sub>32</sub>N<sub>2</sub>NaO<sub>5</sub>, [M+Na]<sup>+</sup>): 583.22034, found: 583.22158;

**TLC** (hexane/EtOAc 80:20 + 1% NEt<sub>3</sub>): *R<sub>f</sub>* = 0.20.

#### 4.2.3.4 1'-(4-Cyano-1*H*-indol-1-yl)-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy-β-*D*-ribofuranose 3'-[(2-cyanoethyl)-(N,N-diisopropyl)]phosphoramidite (**132**)



To a solution of compound **139** (95.0 mg, 169 μmol, 1.00 eq.) and Me<sub>2</sub>NEt (60.2 mg, 508 μmol, 3.00 eq.) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (1 mL) was added 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (60.2 mg, 254 μmol, 1.50 eq.). The mixture was stirred at ambient temperature for 2.5 h. MeOH (1 mL) was added and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 95:5–70:30 + 1% NEt<sub>3</sub>) afforded the phosphoramidite **132** as a white foam (105 mg, 138 μmol, 81%, 1:1 mixture of diastereomers). Analytical data match those reported previously.<sup>209</sup>

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.82 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.81 (dt, *J* = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.48–7.46 (m, 2 H, Ind-5-H), 7.48–7.41 (m, 2 H, Ind-2-H), 7.42–7.37 (m, 4 H, Ph-2'',6''-H), 7.33–7.25 (m, 8 H, Ph-2,2'-6,6'-H), 7.25–7.15 (m, 6 H, Ph-3'',5''-H, Ph-4''-H), 7.15 (dd, *J* = 7.2, 1.1 Hz, 1 H, Ind-6-H), 7.13 (dd, *J* = 7.2, 1.2 Hz, 1 H, Ind-6-H), 6.82–6.72 (m, 8 H, Ph-3,3'-5-5'-H), 6.72–6.67 (m, 2 H, Ind-3-H), 6.43–6.34 (m, 2 H, 1'-H), 4.79–4.69 (m, 2 H, 3'-H), 4.26 (pseudo-q, *J* = 3.4 Hz, 1 H, 4'-H), 4.23 (pseudo-q, *J* = 3.8 Hz, 1 H, 4'-H), 3.78 (4 s, 12 H, OCH<sub>3</sub>), 3.91–3.64 (m, 4 H, OPOCH<sub>2</sub>), 3.64–3.56 (m, 4 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 3.42–3.23 (m, 4 H, 5'-H<sub>2</sub>), 2.75–2.65 (m, 2 H, 2'-H<sub>a</sub>), 2.62 (pseudo-t, *J* = 6.2 Hz, 2 H, NCCH<sub>a</sub>), 2.64–2.44 (m, 2 H, 2'-H<sub>a</sub>), 2.47 (pseudo-t, *J* = 6.2 Hz, 2 H, NCCH<sub>b</sub>), 1.20 (d, *J* = 6.6 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.18 (d, *J* = 6.7 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.18 (d, *J* = 6.8 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.10 (d, *J* = 6.8 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>);

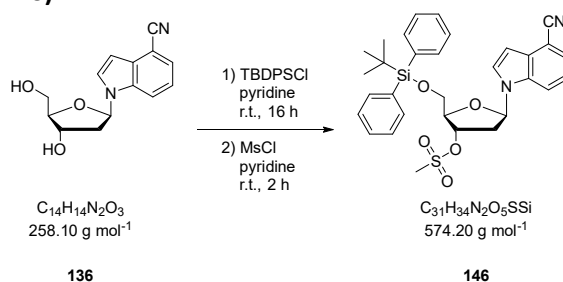
**<sup>13</sup>C{<sup>1</sup>H} NMR** (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 158.68 (Ph-C4,4'), 158.65 (Ph-C4,4'), 144.6 (Ph-C1''), 135.8 (Ind-C7a), 135.7 (Ph-C1,1'), 130.7 (Ind-C3a), 130.30 (Ph-C2,2',6,6'), 130.25 (Ph-C2,2',6,6'), 128.4 (Ph-C2'',6''), 128.3 (Ph-C2'',6''), 127.99 (Ph-C3'',5''), 127.98 (Ph-C3'',5''), 127.4 (Ph-C4''), 127.3 (Ph-C4''), 127.1 (Ind-C2), 127.0 (Ind-C2), 125.7 (Ind-C5), 121.72 (Ind-C6), 121.70 (Ind-C6), 118.79 (Ind-CN), 118.76 (Ind-CN), 117.6 (NCCH<sub>2</sub>), 115.5 (Ind-C7), 115.4 (Ind-C7), 113.2 (Ph-C3,3',5,5'), 103.41 (Ind-C4), 103.36 (Ind-C4), 102.0 (Ind-C3), 101.9 (Ind-C3), 86.62 (5'-OC), 86.61 (5'-OC), 85.7 (C1'), 85.2 (C4'), 73.8 (C3'), 63.6 (C5'), 63.5 (C5'), 58.6–58.2 (m, POCH<sub>2</sub>), 55.38 (OCH<sub>3</sub>), 55.36 (OCH<sub>3</sub>), 43.5–43.3 (m, NCH(CH<sub>3</sub>)<sub>2</sub>), 39.9 (C2'), 39.8 (C2'), 24.7 (NCH(CH<sub>3</sub>)<sub>2</sub>);

**<sup>31</sup>P{<sup>1</sup>H} NMR** (162 MHz, CDCl<sub>3</sub>): δ (ppm) = 148.8, 148.7;

**HR-MS** (ESI+): *m/z* calc. (C<sub>44</sub>H<sub>49</sub>N<sub>4</sub>NaO<sub>6</sub>P, [M+Na]<sup>+</sup>): 783.32874, found: 783.32581;

**TLC** (hexane/EtOAc 60:40 + 1% NEt<sub>3</sub>): *R<sub>f</sub>* = 0.50.

#### 4.2.3.5 1'-(4-Cyano-1*H*-indol-1-yl)-2'-deoxy-3'-*O*-mesyl-5'-*O*-(*tert*-butyldiphenylsilyl)- $\beta$ -D-ribofuranose (**146**)



**136** (505 mg, 1.95 mmol, 1.00 eq.) was dissolved in anhydrous pyridine (15 mL) and treated with *tert*-butyldiphenylsilyl chloride (642 mg, 2.34 mmol, 1.20 eq.) The solution was stirred at ambient temperature for 16 h, at which point TLC indicated complete consumption of the unprotected deoxyribonucleoside. Next, methanesulfonyl chloride (5.58 g, 48.7 mmol, 25.0 eq.) was added and stirring was continued for 2 h. The reaction was quenched by careful addition of MeOH (10 mL, strongly exothermic reaction) and the volatiles were removed under reduced pressure. The residue was taken up in H<sub>2</sub>O (100 mL) and extracted with EtOAc (100 mL). The organic phase was washed with 10% aq. citric acid solution (50 mL), brine (50 mL) and dried over MgSO<sub>4</sub>. After removal of the solvent, purification of the residue by column chromatography (hexane/EtOAc 70:30) afforded the title compound **146** as a pale yellow foam (807 mg, 1.40 mmol, 72%).

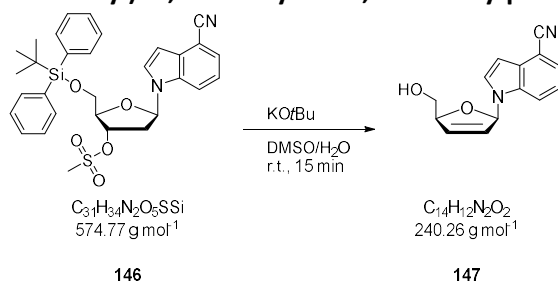
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.76 (ddd,  $J$  = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.68–7.57 (m, 4 H, Ph-2,6-H), 7.49 (dd,  $J$  = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.48–7.39 (m, 2 H, Ph-4-H), 7.37 (d,  $J$  = 3.5 Hz, 1 H, Ind-2-H), 7.37–7.29 (m, 4 H, Ph-3,5-H), 7.20 (dd,  $J$  = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.70 (dd,  $J$  = 3.5, 0.9 Hz, 1 H, Ind-3-H), 6.38 (dd,  $J$  = 8.7, 5.7 Hz, 1 H, 1'-H), 5.50 (pseudo-dt,  $J$  = 5.5, 2.2, 1 H, 3'-H), 4.34 (td,  $J$  = 3.2, 1.9 Hz, 1 H, 4'-H), 3.92 (dd,  $J$  = 11.5, 3.0 Hz, 1 H, 5'-H<sub>a</sub>), 3.85 (dd,  $J$  = 11.5, 3.7 Hz, 1 H, 5'-H<sub>b</sub>), 3.07 (s, 3 H, SCH<sub>3</sub>), 2.76 (ddd,  $J$  = 14.2, 8.7, 5.5 Hz, 1 H, 2'-H<sub>eq</sub>), 2.71 (ddd,  $J$  = 14.2, 5.7, 2.2 Hz, 1 H, 2'-H<sub>ax</sub>), 1.11 (s, 9 H, C(CH<sub>3</sub>)<sub>3</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 135.81 (Ind-C7a), 135.77 (Ph-C2,6), 135.6 (Ph-C2,6), 132.7 (Ph-C1), 132.3 (Ph-C1), 130.7 (Ind-C3a), 130.3 (Ph-C4), 130.2 (Ph-C4), 128.1 (Ph-C3,5), 128.1 (Ph-C3,5), 126.9 (Ind-C2), 125.9 (Ind-C5), 122.0 (Ind-C6), 118.6 (CN), 115.0 (Ind-C7), 103.6 (Ind-C4), 102.5 (Ind-C3), 85.3 (C1'), 84.4 (C4'), 80.3 (C3'), 63.5 (C5'), 38.72 (SCH<sub>3</sub>), 38.66 (C2'), 27.1 (C(CH<sub>3</sub>)<sub>3</sub>), 19.5 (C(CH<sub>3</sub>)<sub>3</sub>);

HR-MS (ESI<sup>+</sup>):  $m/z$  calc. (C<sub>31</sub>H<sub>34</sub>N<sub>2</sub>NaO<sub>5</sub>SSi, [M+Na]<sup>+</sup>): 597.18499, found: 597.18416;

TLC (CH<sub>2</sub>Cl<sub>2</sub>):  $R_f$  = 0.50.

#### 4.2.3.6 1'-(4-Cyano-1*H*-indol-1-yl)-2',3'-didehydro-2',3'-dideoxy- $\beta$ -D-ribofuranose (**147**)



Compound **146** (807 mg, 1.40 mmol, 1.00 eq.) and KOtBu (819 mg, 7.30 mmol, 5.20 eq.) were dissolved in a mixture of DMSO (14 mL) and H<sub>2</sub>O (0.14 mL) and stirred at ambient temperature for

15 min. Afterwards, the reaction mixture was diluted with H<sub>2</sub>O (200 mL) and extracted with EtOAc (3×150 mL). The combined organic layers were dried over MgSO<sub>4</sub>, evaporated and the oily residue was purified by column chromatography (hexane/EtOAc 50:50–40:60) to afford the title compound **147** as a colorless oil (277 mg, 1.15 mmol, 82%).

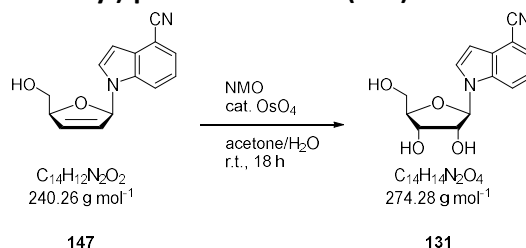
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.80 (ddd, *J* = 8.3, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.49 (dd, *J* = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.43 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.27 (dd, *J* = 8.3, 7.4 Hz, 1H, Ind-6-H), 6.99 (pseudo-dt, *J* = 3.2, 1.7 Hz, 1 H, 1'-H), 6.74 (ddd, *J* = 3.4, 0.9, 0.5 Hz, 1 H, Ind-3-H), 6.41 (ddd, *J* = 6.0, 2.0, 1.5 Hz, 1 H, 3'-H), 6.14–6.07 (m, 1 H, 2'-H), 5.02–4.93 (m, 1 H, 4'-H), 3.75 (ddd, *J* = 12.2, 5.8, 3.3 Hz, 1 H, 5'-H<sub>a</sub>), 3.64 (ddd, *J* = 12.2, 7.3, 4.2 Hz, 1 H, 5'-H<sub>b</sub>), 1.69 (dd, *J* = 7.3, 5.8 Hz, 1 H, 5'-OH);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 136.2 (Ind-C7a), 134.4 (C3'), 130.7 (Ind-C3a), 127.7 (Ind-C2), 126.1 (C2'), 125.9 (Ind-C5), 122.0 (Ind-C6), 118.6 (CN), 114.6 (Ind-C7), 103.5 (Ind-C4), 102.4 (Ind-C3), 90.2 (C1'), 87.4 (C4'), 64.2 (C5');

**HR-MS** (ESI+): *m/z* calc. (C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>NaO<sub>2</sub>, [M+Na]<sup>+</sup>): 263.07910, found: 263.07816;

**TLC** (CH<sub>2</sub>Cl<sub>2</sub>): *R<sub>f</sub>* = 0.30.

#### 4.2.3.7 1'-(4-Cyano-1*H*-indol-1-yl)-β-*D*-ribofuranose (**131**)



Compound **147** (253 mg, 1.05 mmol, 1.00 eq.) was dissolved in a mixture of acetone (10 mL) and H<sub>2</sub>O (1.2 mL). Afterwards, *N*-methylmorpholine *N*-oxide (320 mg, 2.73 mmol, 2.60 eq.) and OsO<sub>4</sub> (1.07 mL, 105 μmol, 10 mol%, 2.5% solution in *t*BuOH) were added and the mixture was stirred at ambient temperature for 2 h. Then, a second portion of NMO (320 mg, 2.73 mmol, 2.60 eq.) was added and stirring was continued for 16 h. The reaction mixture was concentrated to a volume of 2 mL under reduced pressure. Subsequently, 5% aq. Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (60 mL) was added to reduce the OsO<sub>4</sub> and the solution was extracted with EtOAc (2×60 mL). The combined organic phases were dried over MgSO<sub>4</sub>. After removal of the solvent, purification of the residue by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5–93:7) afforded the ribonucleoside **131** as a white foam (216 mg, 788 μmol, 75%).

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 8.01 (ddd, *J* = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.93 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.59 (dd, *J* = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.30 (dd, *J* = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.68 (dd, *J* = 3.4, 0.9 Hz, 1 H, Ind-3-H), 5.95 (d, *J* = 6.3 Hz, 1 H, 1'-H), 5.40 (d, *J* = 6.6 Hz, 1 H, 2'-OH), 5.21 (d, *J* = 4.9 Hz, 1 H, 3'-OH), 5.09 (pseudo-t, *J* = 5.2 Hz, 1 H, 5'-OH), 4.27 (ddd, *J* = 6.6, 6.3, 5.2 Hz, 1 H, 2'-H), 4.08 (ddd, *J* = 5.2, 4.9, 3.2 Hz, 1 H, 3'-H), 3.95 (ddd, *J* = 3.8, 3.7, 3.2 Hz, 1 H, 4'-H), 3.65 (ddd, *J* = 11.9, 5.2, 3.7 Hz, 1 H, 5'-H<sub>a</sub>), 3.59 (ddd, *J* = 11.9, 5.2, 3.8 Hz, 1 H, 5'-H<sub>b</sub>);

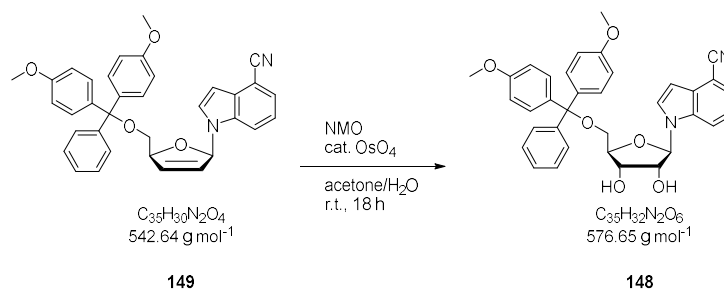
**<sup>13</sup>C{<sup>1</sup>H} NMR** (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) = 135.5 (Ind-C7a), 129.5 (Ind-C3a), 129.3 (Ind-C2), 125.3 (Ind-C5), 121.5 (Ind-C6), 118.5 (CN), 116.1 (Ind-C7), 101.6 (Ind-C4), 100.3 (Ind-C3), 89.0 (C1'), 85.3 (C4'), 74.4 (C2'), 70.3 (C3'), 61.4 (C5');

**HR-MS** (ESI+): *m/z* calc. (C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>NaO<sub>4</sub>, [M+Na]<sup>+</sup>): 297.08458, found: 297.08381;

TLC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5): *R<sub>f</sub>* = 0.07.

#### 4.2.3.8 1'-(4-Cyano-1*H*-indol-1-yl)-5'-*O*-(4,4'-dimethoxytrityl)-β-D-ribofuranose (**148**)

##### Method A



Compound **149** (119 mg, 220 μmol, 1.00 eq.) was dissolved in a mixture of acetone (5 mL) and H<sub>2</sub>O (0.6 mL). Afterwards, *N*-methylmorpholine *N*-oxide (66.9 mg, 571 μmol, 2.60 eq.) and OsO<sub>4</sub> (0.5 mL, 22.0 μmol, 10 mol%, 2.5% solution in *t*BuOH) were added and the mixture was stirred at ambient temperature for 2 h. Then, a second portion of NMO (66.9 mg, 571 μmol, 2.60 eq.) was added and stirring was continued for 16 h. The reaction mixture was concentrated to a volume of 2 mL under reduced pressure. Subsequently, 5% aq. Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (60 mL) was added to reduce the OsO<sub>4</sub> and the solution was extracted with EtOAc (2×60 mL). The combined organic phases were dried over MgSO<sub>4</sub>. After removal of the solvent, purification of the residue by column chromatography (CH<sub>2</sub>Cl<sub>2</sub> + 1% NEt<sub>3</sub>) afforded the DMT-protected ribonucleoside **148** as a pale yellow foam (127 mg, 220 μmol, > 99%).

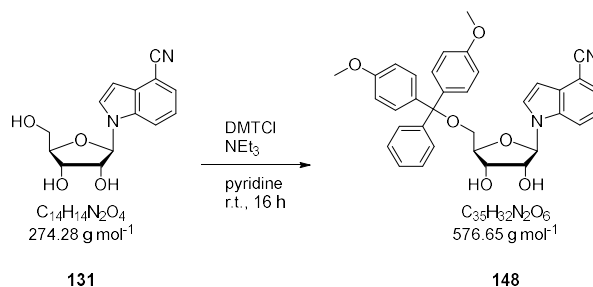
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 7.83 (ddd, *J* = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.54 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.46–7.40 (m, 3 H, Ind-5-H, Ph-2'',6''-H), 7.37–7.29 (m, 4 H, Ph-2,2',6,6'-H), 7.31–7.18 (m, 3 H, Ph-3'',5''-H, Ph-4''-H), 7.08 (dd, *J* = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.83–6.78 (m, 4 H, Ph-3,3',5,5'-H), 6.70 (dd, *J* = 3.4, 0.9 Hz, 1 H, Ind-3-H), 5.97 (d, *J* = 5.5 Hz, 1 H, 1'-H), 4.51 (dd, *J* = 5.6, 5.5 Hz, 1 H, 2'-H), 4.41 (dd, *J* = 5.6, 3.8 Hz, 1 H, 3'-H), 4.20 (ddd, *J* = 3.8, 3.5, 3.4 Hz, 1 H, 4'-H), 3.78 (2 s, 6 H, OCH<sub>3</sub>), 3.53 (dd, *J* = 10.6, 3.4 Hz, 1 H, H-5'<sub>a</sub>), 3.40 (dd, *J* = 10.6, 3.5 Hz, 1 H, H-5'<sub>b</sub>);

<sup>13</sup>C{<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 158.7 (Ph-C4,4'), 144.6 (Ph-C1''), 135.7 (Ind-C7a), 135.6 (Ph-C1,1''), 130.7 (Ind-C3a), 130.3 (Ph-C2,2',6,6'), 128.3 (Ph-C2'',6''), 128.1 (Ph-C3'',5''), 127.8 (Ind-C2), 127.2 (Ph-C4''), 125.7 (Ind-C5), 121.7 (Ind-C6), 118.7 (CN), 115.6 (Ind-C7), 113.4 (Ph-C3,3',5,5'), 103.3 (Ind-C4), 102.0 (Ind-C3), 90.4 (C1'), 86.9 (5'-OC), 83.5 (C4'), 74.7 (C2'), 71.1 (C3'), 63.4 (C5'), 55.4 (OCH<sub>3</sub>)

HR-MS (ESI+): *m/z* calc. (C<sub>35</sub>H<sub>32</sub>N<sub>2</sub>NaO<sub>6</sub>, [M+Na]<sup>+</sup>): 599.21526, found: 599.21636;

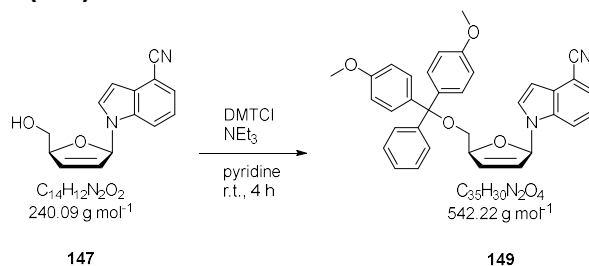
TLC (CH<sub>2</sub>Cl<sub>2</sub>): *R<sub>f</sub>* = 0.17.

##### Method B



**131** (120 mg, 438  $\mu\text{mol}$ , 1.00 eq.) was coevaporated with pyridine (3 $\times$ 5 mL) and dissolved in anhydrous pyridine (15 mL) containing  $\text{NEt}_3$  (66.5 mg, 657  $\mu\text{mol}$ , 1.50 eq.). 4,4'-Dimethoxytrityl chloride (178 mg, 526  $\mu\text{mol}$ , 1.20 eq.) was added in three portions over the course of 1 h and stirring was continued at ambient temperature for 15 h. The reaction was quenched by addition of MeOH (1.5 mL) and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 80:20– $\text{CH}_2\text{Cl}_2$ – $\text{CH}_2\text{Cl}_2$ /MeOH 98:2, the column was packed using 1%  $\text{NEt}_3$  in the eluent) afforded the DMT-protected ribonucleoside **148** as an off-white foam (226 mg, 392  $\mu\text{mol}$ , 89%). Analytical data matched those reported above.

#### 4.2.3.9 1'-(4-Cyano-1*H*-indol-1-yl)-2',3'-didehydro-2',3'-dideoxy-5'-*O*-(4,4'-dimethoxytrityl)- $\beta$ -D-ribofuranose (**149**)



Compound **147** (312 mg, 1.30 mmol, 1.00 eq.) was coevaporated with pyridine (3 $\times$ 5 mL) and dissolved in anhydrous pyridine (30 mL) containing  $\text{NEt}_3$  (0.22 mL, 1.56 mmol, 1.50 eq.). 4,4'-Dimethoxytrityl chloride (528 mg, 1.56 mmol, 1.50 eq.) was added in three portions over the course of 1 h and stirring was continued at ambient temperature for 3 h. The reaction was quenched by addition of MeOH (1 mL) and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 80:20– $\text{CH}_2\text{Cl}_2$ – $\text{CH}_2\text{Cl}_2$ /MeOH 99:1 + 1%  $\text{NEt}_3$ ) afforded the title compound **149** as a pale yellow foam (437 mg, 1.29 mmol, 62%).

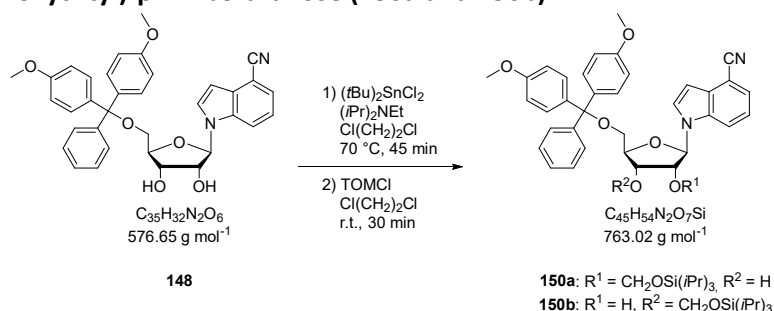
**$^1\text{H}$  NMR** (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.81 (ddd,  $J$  = 8.3, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.51 (dd,  $J$  = 7.5, 0.9 Hz, 1 H, Ind-5-H), 7.33–7.29 (m, 2 H, Ph-2'',6''-H), 7.27 (d,  $J$  = 3.4 Hz, 1 H, Ind-2-H), 7.25 (dd,  $J$  = 8.3, 7.5 Hz, 1 H, Ind-6-H), 7.22–7.17 (m, 4 H, Ph-2,2',6,6'-H), 7.17–7.11 (m, 3 H, Ph-3'',5''-H, Ph-4''-H), 6.99 (pseudo-dt,  $J$  = 3.2, 1.6 Hz, 1 H, 1'-H), 6.70–6.62 (m, 4 H, Ph-3,3',5,5'-H), 6.58 (dd,  $J$  = 3.3, 0.9 Hz, 1 H, Ind-3-H), 6.42 (pseudo-dt,  $J$  = 6.0, 1.7 Hz, 1 H, 3'-H), 6.07 (ddd,  $J$  = 6.0, 2.3, 1.4 Hz, 1 H, 2'-H), 5.09–5.01 (m, 1 H, 4'-H), 3.76 (2 s, 6 H,  $\text{OCH}_3$ ), 3.26 (dd,  $J$  = 9.9, 5.9 Hz, 1 H, 5'- $\text{H}_a$ ), 3.13 (dd,  $J$  = 9.9, 4.2 Hz, 1 H, 5'- $\text{H}_b$ );

**$^{13}\text{C}\{^1\text{H}\}$  NMR** (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 158.5 (Ph-C4,4'), 144.7 (Ph-C1''), 136.1 (Ind-C7a), 135.8 (Ph-C1,1'), 135.7 (Ph-C1,1'), 134.9 (C3'), 130.8 (Ind-C3a), 130.2 (Ph-C2,2',6,6'), 130.1 (Ph-C2,2',6,6'), 128.3 (Ph-C2'',6''), 128.0 (Ind-C2), 127.8 (Ph-C3'',5''), 126.9 (Ph-C4''), 125.7 (Ind-C5), 125.5 (C2'), 121.8 (Ind-C6), 118.8 (CN), 114.8 (Ind-C7), 113.1 (Ph-C3,3',5,5'), 103.3 (Ind-C4), 101.9 (Ind-C3), 90.3 (C1'), 86.3 (5'-OC), 86.1 (C4'), 65.9 (C5'), 55.3 ( $\text{OCH}_3$ ).

**HR-MS** (ESI+):  $m/z$  calc. ( $\text{C}_{35}\text{H}_{30}\text{N}_2\text{NaO}_4$ ,  $[\text{M}+\text{Na}]^+$ ): 565.20978, found: 565.21053;

**TLC** ( $\text{CH}_2\text{Cl}_2$ ):  $R_f$  = 0.70.

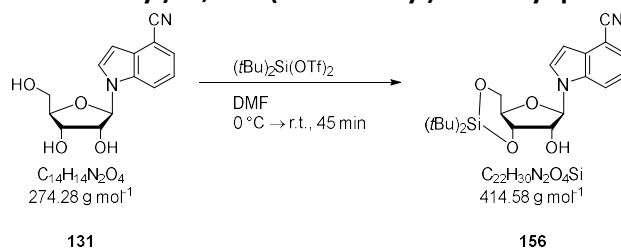
#### 4.2.3.10 1'-(4-Cyano-1*H*-indol-1-yl)-2'/3'-*O*-(triisopropylsilyloxy)methyl-5'-*O*-(4,4'-dimethoxytrityl)- $\beta$ -*D*-ribofuranose (**150a** and **150b**)



A solution of compound **148** (317 mg, 550  $\mu$ mol, 1.00 eq.) in dichloroethane (3 mL) was treated successively with (*i*Pr)<sub>2</sub>NEt (0.3 mL, 2.23 mmol, 4.05 eq.) and (tBu)<sub>2</sub>SnCl<sub>2</sub> (200 mg, 660  $\mu$ mol, 1.20 eq.) and heated to 70 °C for 45 min. After cooling to ambient temperature, (triisopropylsilyloxy)methyl chloride (151 mg, 676  $\mu$ mol, 1.23 eq.) was added and stirring was continued at that temperature for 30 min. The reaction was quenched by addition of MeOH (0.1 mL), diluted with CH<sub>2</sub>Cl<sub>2</sub> (30 mL) and washed with sat. aq. NaHCO<sub>3</sub> (3×50 mL). The organic phase was dried over MgSO<sub>4</sub>, evaporated and the residue was purified by column chromatography (hexane/EtOAc 80:20–70:30 + 1% NEt<sub>3</sub>) to afford a mixture of the 2'- and 3'-*O*-TOM-protected isomers **150a** and **150b** (322 mg, 422  $\mu$ mol, 77%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.86 (pseudo-dt, *J* = 8.4, 0.9 Hz), 7.81 (pseudo-dt, *J* = 8.3, 0.9 Hz), 7.58 (d, *J* = 3.4 Hz), 7.52 (d, *J* = 3.4 Hz), 7.49–7.40 (m), 7.37–7.18 (m), 7.13 (dd, *J* = 8.4, 7.4 Hz), 7.05 (dd, *J* = 8.3, 7.4 Hz), 6.86–6.75 (m), 6.73 (dd, *J* = 3.4, 0.8 Hz), 6.70 (dd, *J* = 3.4, 0.8 Hz), 6.11 (d, *J* = 6.0 Hz), 5.95 (d, *J* = 5.3 Hz), 5.10 (d, *J* = 4.9 Hz), 5.09 (d, *J* = 4.9 Hz), 4.89 (d, *J* = 4.9 Hz), 4.85 (d, *J* = 4.9 Hz), 4.57–4.48 (m), 4.38–4.29 (m), 4.27 (pseudo-q, *J* = 2.9 Hz), 3.78 (4 s), 3.57 (d, *J* = 5.4 Hz), 3.52 (ddd, *J* = 12.6, 10.6, 2.9 Hz), 3.41 (dd, *J* = 10.6, 3.3 Hz), 3.37–3.31 (m), 3.07 (pseudo-t, *J* = 2.0 Hz), 1.16–0.93 (m).

#### 4.2.3.11 1'-(4-Cyano-1*H*-indol-1-yl)-3',5'-*O*-(di-*tert*-butyl)silanediy- $\beta$ -*D*-ribofuranose (**156**)

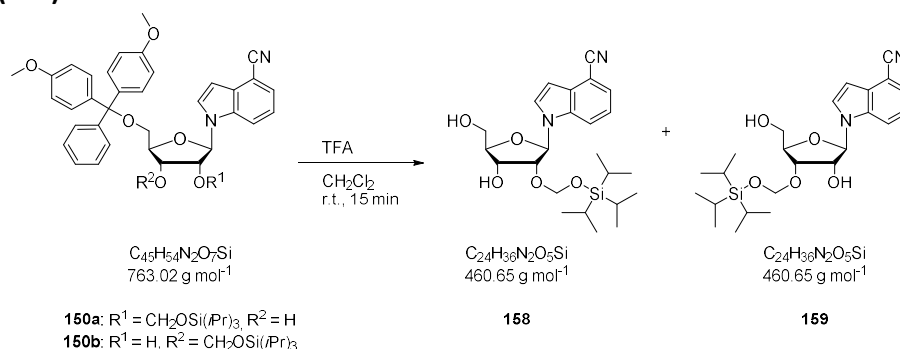


A solution of compound **131** (41.1 mg, 150  $\mu$ mol, 1.00 eq.) in DMF (1.5 mL) was cooled to 0 °C and slowly treated with (tBu)<sub>2</sub>Si(OTf)<sub>2</sub> (72.7 mg, 165  $\mu$ mol, 1.10 eq.). The mixture was warmed to ambient temperature over the course of 30 min and stirring was continued for 15 min. NEt<sub>3</sub> (62  $\mu$ L) was added, stirring was continued for 15 min and the volatiles were removed under reduced pressure. The residue was purified by column chromatography to obtain the silylene acetal **156** as a white foam (18.7 mg, 45.0  $\mu$ mol, 30%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.68 (pseudo-dt, *J* = 8.4, 0.9 Hz, 1 H, Ind-7-H), 7.52 (dd, *J* = 7.4, 0.9 Hz, 1 H, Ind-5-H), 7.30 (d, *J* = 3.4 Hz, 1 H, Ind-2-H), 7.30 (dd, *J* = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.81 (dd, *J* = 3.4, 0.9 Hz, 1 H, Ind-3-H), 6.11–6.07 (m, 1 H, 1'), 4.56–4.48 (m, 1 H, 5'-H<sub>a</sub>), 4.37–4.31 (m, 1 H, 2'-H), 4.22–4.15 (m, 1 H, 3'-H), 4.15–4.10 (m, 1 H, 4'-H), 4.10–4.03 (m, 1 H, 5'-H<sub>b</sub>), 2.70 (d, *J* = 0.8 Hz, 1 H, 2'-OH), 1.10 (s, 9 H, C(CH<sub>3</sub>)<sub>3</sub>), 1.06 (s, 9 H, C(CH<sub>3</sub>)<sub>3</sub>);

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 135.4 (Ind-C7a), 130.4 (Ind-C3a), 126.7 (Ind-C2), 126.0 (Ind-C5), 122.2 (Ind-C6), 118.5 (CN), 114.5 (Ind-C7), 103.8 (Ind-C4), 102.9 (Ind-C3), 91.7 (C1'), 76.8 (C3'), 74.7 (C2'), 74.5 (C4'), 67.7 (C5'), 27.5 ( $\text{C}(\text{CH}_3)_3$ ), 27.4 ( $\text{C}(\text{CH}_3)_3$ ), 22.9 ( $\text{C}(\text{CH}_3)_3$ ), 20.6 ( $\text{C}(\text{CH}_3)_3$ ).

**4.2.3.12 1'-(4-Cyano-1*H*-indol-1-yl)-2'-*O*-(triisopropylsilyloxy)methyl- $\beta$ -d-ribofuranose (158) and 1'-(4-Cyano-1*H*-indol-1-yl)-3'-*O*-(triisopropylsilyloxy)methyl- $\beta$ -d-ribofuranose (159)**



A 1:1 mixture of compounds **150a** and **150b** (470 mg, 616  $\mu\text{mol}$ , 1.00 eq.) was dissolved in  $\text{CH}_2\text{Cl}_2$  (50 mL) containing TFA (1 mL) and stirred at ambient temperature for 15 min. Then, MeOH (1 mL) was added and the volatiles were removed under reduced pressure. The addition and removal of MeOH were repeated until only a faint red color persisted. Purification of the residue by column chromatography (hexane/EtOAc 80:20–50:50) afforded both the 2'- and 3'-*O*-TOM-protected isomers. The desired 2'-*O*-TOM-protected compound **158** eluted second from the column and was obtained as a colorless oil (110 mg, 239  $\mu\text{mol}$ , 39%).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.77 (d,  $J = 8.4$  Hz, 1 H, Ind-7-H), 7.58 (d,  $J = 3.4$  Hz, 1 H, Ind-2-H), 7.49 (dd,  $J = 7.5, 0.8$  Hz, 1 H, Ind-5-H), 7.25 (dd,  $J = 8.4, 7.5$  Hz, 1 H, Ind-6-H), 6.79 (d,  $J = 3.4$  Hz, 1 H, Ind-3-H), 6.18–6.10 (m, 1 H, 1'-H), 5.07 (d,  $J = 4.9$  Hz, 1 H, 2'- $\text{OCH}_a$ ), 4.84 (d,  $J = 4.9$  Hz, 1 H, 2'- $\text{OCH}_b$ ), 4.47–4.39 (m, 2 H, 2'-H, 3'-H), 4.21 (pseudo-q,  $J = 2.8$  Hz, 1 H, 4'-H), 3.98 (pseudo-dt,  $J = 12.1, 2.9$  Hz, 1 H, 5'- $\text{H}_a$ ), 3.88 (pseudo-dt,  $J = 12.1, 4.0$  Hz, 1 H, 5'- $\text{H}_b$ ), 3.21 (d,  $J = 2.3$  Hz, 1 H, 3'-OH), 1.85 (pseudo-t,  $J = 5.3$  Hz, 1 H, 5'-OH), 1.16–0.97 (m, 21 H,  $\text{CH}(\text{CH}_3)_2$ );

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 136.1 (Ind-C7a), 130.6 (Ind-C3a), 127.3 (Ind-C2), 125.8 (Ind-C5), 122.0 (Ind-C6), 118.7 (CN), 115.0 (Ind-C7), 103.5 (Ind-C4), 102.6 (Ind-C3), 91.0 (2'- $\text{OCH}_2$ ), 88.0 (C1'), 84.6 (C4'), 83.2 (C2'), 70.8 (C3'), 62.8 (C5'), 17.86 ( $\text{CH}(\text{CH}_3)_2$ ), 17.85 ( $\text{CH}(\text{CH}_3)_2$ ), 11.9 ( $\text{CH}(\text{CH}_3)_2$ );

**HR-MS** (ESI+):  $m/z$  calc. ( $\text{C}_{24}\text{H}_{36}\text{N}_2\text{NaO}_5\text{Si}$ ,  $[\text{M}+\text{Na}]^+$ ): 483.22857, found: 483.22986;

**TLC** (hexane/EtOAc 60:40):  $R_f = 0.25$ .

Analytical data for the 3'-*O*-TOM-protected isomer **159**:

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.72 (d,  $J = 8.4$  Hz, 1 H, Ind-7-H), 7.51 (d,  $J = 3.4$  Hz, 1 H, Ind-2-H), 7.48 (dd,  $J = 7.4, 0.7$  Hz, 1 H, Ind-5-H), 7.25 (dd,  $J = 8.4, 7.4$  Hz, 1 H, Ind-6-H), 6.77 (dd,  $J = 3.4, 0.7$  Hz, 1 H, Ind-3-H), 5.98 (d,  $J = 4.4$  Hz, 1 H, 1'-H), 5.18 (d,  $J = 4.8$  Hz, 1 H, 3'- $\text{OCH}_{2a}$ ), 5.01 (d,  $J = 4.8$  Hz, 1 H, 3'- $\text{OCH}_{2b}$ ), 4.46 (dd,  $J = 5.6, 4.4$  Hz, 1 H, 2'-H), 4.35 (dd,  $J = 5.6, 5.6$  Hz, 1 H, 3'-H), 4.22 (ddd,  $J = 5.6, 3.0, 2.8$  Hz, 1 H, 4'-H), 3.95 (dd,  $J = 12.3, 2.8$  Hz, 1 H, 5'- $\text{H}_a$ ), 3.80 (dd,  $J = 12.3, 3.0$  Hz, 1 H, 5'- $\text{H}_b$ ), 3.49 ( $s_{br}$ , 1 H, OH), 1.19–1.01 (m, 21 H,  $\text{CH}(\text{CH}_3)_2$ );

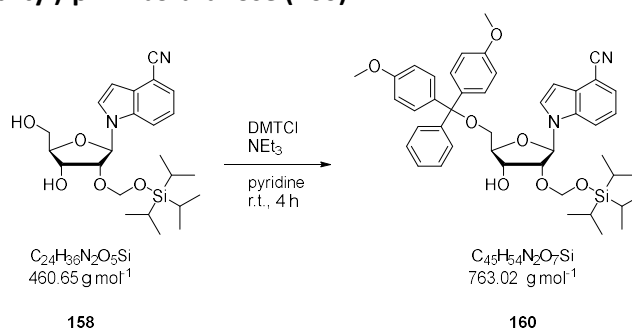


$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 135.9 (Ind-C7a), 130.5 (Ind-C3a), 127.2 (Ind-C2), 125.8 (Ind-C5), 121.9 (Ind-C6), 118.7 (CN), 114.9 (Ind-C7), 103.4 (Ind-C4), 102.5 (Ind-C3), 91.3 (3'-OCH<sub>2</sub>), 90.2 (C1'), 82.4 (C4'), 79.0 (C3'), 74.5 (C2'), 61.6 (C5'), 17.9 (CH(CH<sub>3</sub>)<sub>2</sub>), 11.9 (CH(CH<sub>3</sub>)<sub>2</sub>);

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{24}\text{H}_{36}\text{N}_2\text{NaO}_5\text{Si}$ ,  $[\text{M}+\text{Na}]^+$ ): 483.22857, found: 483.22953;

TLC (hexane/EtOAc 60:40):  $R_f$  = 0.30.

#### 4.2.3.13 1'-(4-Cyano-1H-indol-1-yl)-2'-O-(triisopropylsilyloxy)methyl-5'-O-(4,4'-dimethoxytrityl)- $\beta$ -D-ribofuranose (160)



Compound **158** (99.0 mg, 215  $\mu\text{mol}$ , 1.00 eq.) was coevaporated with pyridine (3 $\times$ 5 mL) and dissolved in anhydrous pyridine (12 mL) containing NEt<sub>3</sub> (32.6 mg, 323  $\mu\text{mol}$ , 1.50 eq.). 4,4'-Dimethoxytrityl chloride (87.4 mg, 256  $\mu\text{mol}$ , 1.50 eq.) was added in three portions over the course of 1 h and stirring was continued at ambient temperature for 3 h. The reaction was quenched by addition of MeOH (1 mL) and the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 80:20–70:30 + 1% NEt<sub>3</sub>) afforded the DMT- and TOM-protected ribonucleoside **160** as an off-white foam (88.6 mg, 116  $\mu\text{mol}$ , 54%).

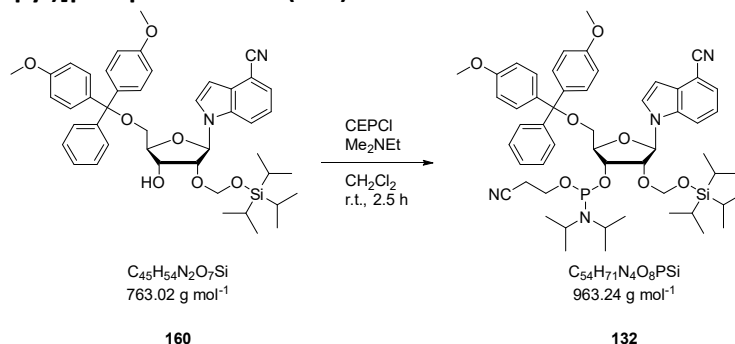
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.86 (ddd,  $J$  = 8.4, 0.9, 0.9 Hz, 1 H, Ind-7-H), 7.52 (d,  $J$  = 3.4 Hz, 1 H, Ind-2-H), 7.48–7.45 (m, 1 H, Ind-5-H), 7.45–7.41 (m, 2 H, Ph-2'',6''-H), 7.35–7.30 (m, 4 H, Ph-2,2',6,6'-H), 7.30–7.18 (m, 4 H, Ph-3'',5''-H, Ph-4''-H), 7.05 (dd,  $J$  = 8.4, 7.4 Hz, 1 H, Ind-6-H), 6.84–6.77 (m, 4 H, Ph-3,3',5,5'-H), 6.70 (dd,  $J$  = 3.4, 0.8 Hz, 1 H, Ind-3-H), 6.11 (d,  $J$  = 5.9 Hz, 1 H, 1'-H), 5.10 (d,  $J$  = 4.9 Hz, 1 H, 2'-OCH<sub>a</sub>), 4.85 (d,  $J$  = 4.9 Hz, 1 H, 2'-OCH<sub>b</sub>), 4.54 (dd,  $J$  = 5.1, 2.3 Hz, 1 H, 3'-H), 4.55–4.47 (m, 1 H, 2'-H), 4.27 (ddd,  $J$  = 3.3, 2.7, 2.3 Hz, 1 H, 4'-H), 3.79 (2 s, 6 H, OCH<sub>3</sub>), 3.50 (dd,  $J$  = 10.6, 2.7 Hz, 1 H, 5'-H<sub>a</sub>), 3.41 (dd,  $J$  = 10.6, 3.3 Hz, 1 H, 5'-H<sub>b</sub>), 3.07 (s<sub>br</sub>, 1 H, 3'-OH), 1.09–1.00 (m, 21 H, CH(CH<sub>3</sub>)<sub>2</sub>);

$^{13}\text{C}\{^1\text{H}\}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 158.7 (Ph-C4,4'), 144.6 (Ph-C1''), 135.7 (Ind-C7a), 135.6 (Ph-C1,1'), 130.7 (Ind-C3a), 130.3 (Ph-C2,2',6,6'), 128.3 (Ph-C2'',6''), 128.1 (Ph-C3'',5''), 127.9 (Ind-C2), 127.1 (Ph-C4''), 125.7 (Ind-C5), 121.7 (Ind-C6), 118.8 (CN), 115.6 (Ind-C7), 113.3 (Ph-C3,3',5-5'), 103.4 (Ind-C4), 102.1 (Ind-C3), 91.2 (2'-OCH<sub>2</sub>), 88.1 (C1'), 86.9 (5'-OC), 83.9 (C4'), 83.3 (C2'), 71.2 (C3'), 63.7 (C5'), 55.4 (OCH<sub>3</sub>), 17.9 (CH(CH<sub>3</sub>)<sub>2</sub>), 11.9 (CH(CH<sub>3</sub>)<sub>2</sub>);

HR-MS (ESI+):  $m/z$  calc. ( $\text{C}_{45}\text{H}_{55}\text{N}_2\text{O}_7\text{Si}$ ,  $[\text{M}+\text{H}]^+$ ): 764.38066, found: 764.54634;

TLC (hexane/EtOAc 70:30):  $R_f$  = 0.30.

**4.2.3.14 1'-(4-Cyano-1*H*-indol-1-yl)-2'-*O*-(triisopropylsilyloxy)methyl-5'-*O*-(4,4'-dimethoxytrityl)- $\beta$ -*D*-ribofuranose 3'-[[2-cyanoethyl-(*N,N*-diisopropyl)]phosphoramidite (**133**)**



To a solution of compound **160** (80.0 mg, 105  $\mu$ mol, 1.00 eq.) and Me<sub>2</sub>NEt (80.3 mg, 1.10 mmol, 10.5 eq.) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (1 mL) was added 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (37.2 mg, 157  $\mu$ mol, 1.50 eq.). The mixture was stirred at ambient temperature for 2.5 h. Afterwards, the solvent was removed under reduced pressure. Purification of the residue by column chromatography (hexane/EtOAc 95:5–70:30 + 1% NEt<sub>3</sub>) afforded the phosphoramidite **133** as an off-white foam (79.9 mg, 83.0  $\mu$ mol, 79%, 1:1 mixture of diastereomers).

**<sup>1</sup>H NMR** (600 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.85 (dd,  $J$  = 8.4, 1.0 Hz, 1 H, Ind-7-H), 7.84 (dd,  $J$  = 8.5, 1.0 Hz, 1 H, Ind-7-H), 7.62 (d,  $J$  = 3.4 Hz, 1 H, Ind-2-H), 7.59 (d,  $J$  = 3.4 Hz, 1 H, Ind-2-H), 7.48–7.41 (m, 6 H, Ind-5-H, Ph-H), 7.39–7.18 (m, 13 H, Ph-H), 7.09 (pseudo-t,  $J$  = 7.3 Hz, 1 H, Ind-6-H), 7.07 (pseudo-t,  $J$  = 7.3 Hz, 1 H, Ind-6-H), 6.84–6.77 (m, 8 H, Ph-H), 6.70–6.66 (m, 2 H, Ind-3-H), 6.15 (d,  $J$  = 6.3 Hz, 1 H, 1'-H), 6.11 (d,  $J$  = 6.0 Hz, 1 H, 1'-H), 4.93–4.85 (m, 4 H, 2'-OCH<sub>2</sub>), 4.70 (dd,  $J$  = 6.3, 5.0 Hz, 1 H, 2'-H), 4.62 (pseudo-t,  $J$  = 5.6 Hz, 1 H, 2'-H), 4.58–4.50 (m, 2 H, 3'-H), 4.36 (pseudo-q,  $J$  = 3.4 Hz, 1 H, 4'-H), 4.30 (pseudo-q,  $J$  = 3.1 Hz, 1 H, 4'-H), 3.95–3.88 (m, 1 H, POCH<sub>a</sub>), 3.88–3.81 (m, 1 H, POCH<sub>b</sub>), 3.79 (s, 3 H, OCH<sub>3</sub>), 3.79 (s, 3 H, OCH<sub>3</sub>), 3.78 (s, 3 H, OCH<sub>3</sub>), 3.78 (s, 3 H, OCH<sub>3</sub>), 3.65–3.55 (m, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>, POCH<sub>b</sub>), 3.55 (dd,  $J$  = 4.3, 3.1 Hz, 1 H, 5'-H<sub>a</sub>), 3.50 (dd,  $J$  = 10.6, 3.0 Hz, 1 H, 5'-H<sub>a</sub>), 3.39–3.34 (m, 1 H, 5'-H<sub>b</sub>), 3.34–3.29 (m, 1 H, 5'-H<sub>b</sub>), 2.70–2.64 (m, 1 H, NCCH<sub>a</sub>), 2.64–2.58 (m, 1 H, NCCH<sub>b</sub>), 2.37–2.33 (m, 1 H, NCCH<sub>a</sub>), 2.33–2.30 (m, 1 H, NCCH<sub>b</sub>), 1.18 (d,  $J$  = 6.8 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.17 (d,  $J$  = 6.7 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.17 (d,  $J$  = 6.8 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 1.03 (d,  $J$  = 6.8 Hz, 6 H, NCH(CH<sub>3</sub>)<sub>2</sub>), 0.92–0.83 (m, 42 H, SiCH(CH<sub>3</sub>)<sub>2</sub>);

**<sup>13</sup>C{<sup>1</sup>H} NMR** (150 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 158.75 (Ph-C4,4'), 158.73 (Ph-C4,4'), 144.6 (Ph-C1''), 144.5 (Ph-C1''), 135.9 (Ph-C1,1'), 135.77 (Ind-C7a), 135.75 (Ph-C1,1'), 135.68 (Ph-C1,1'), 135.65 (Ph-C1,1'), 135.62 (Ind-C7a), 130.83 (Ind-C3a), 130.77 (Ind-C3a), 130.4 (Ph-C2,2',6,6'), 130.3 (Ph-C2,2',6,6'), 128.5 (Ph-C2'',6''), 128.4 (Ph-C2'',6''), 128.1 (Ph-C3'',5''), 128.0 (Ind-C2), 127.8 (Ind-C2), 127.2 (Ph-C4''), 127.1 (Ph-C4''), 125.64 (Ind-C5), 125.59 (Ind-C5), 121.7 (Ind-C6), 121.6 (Ind-C6), 118.90 (Ind-CN), 118.87 (Ind-CN), 172.7 (NCCH<sub>2</sub>), 172.5 (NCCH<sub>2</sub>), 115.72 (Ind-C7), 115.65 (Ind-C7), 113.3 (Ph-C3,3',5,5'), 103.23 (Ind-C4), 103.21 (Ind-C4), 102.0 (Ind-C3), 101.8 (Ind-C3), 89.7 (2'-OCH<sub>a</sub>), 89.6 (2'-OCH<sub>a</sub>), 89.29 (2'-OCH<sub>b</sub>), 89.26 (2'-OCH<sub>b</sub>), 88.5 (C1'), 88.3 (C1'), 86.94 (5'-OC), 86.86 (5'-OC), 83.8 (C4'), 83.6 (C4'), 78.7 (C2'), 78.1 (C2'), 71.7 (C3'), 71.1 (C3'), 63.3 (C5'), 63.0 (C5'), 59.0 (POCH<sub>2</sub>), 58.0 (POCH<sub>2</sub>), 55.39 (OCH<sub>3</sub>), 55.36 (OCH<sub>3</sub>), 43.6 (NCH(CH<sub>3</sub>)<sub>2</sub>), 43.5 (NCH(CH<sub>3</sub>)<sub>2</sub>), 43.33 (NCH(CH<sub>3</sub>)<sub>2</sub>), 43.25 (NCH(CH<sub>3</sub>)<sub>2</sub>), 24.7 (NCH(CH<sub>3</sub>)<sub>2</sub>), 20.5 (NCCH<sub>2</sub>), 20.2 (NCCH<sub>2</sub>), 17.84 (SiCH(CH<sub>3</sub>)<sub>2</sub>), 17.83 (SiCH(CH<sub>3</sub>)<sub>2</sub>), 17.80 (SiCH(CH<sub>3</sub>)<sub>2</sub>), 17.78 (SiCH(CH<sub>3</sub>)<sub>2</sub>), 11.94 (SiCH(CH<sub>3</sub>)<sub>2</sub>), 11.92 (SiCH(CH<sub>3</sub>)<sub>2</sub>);

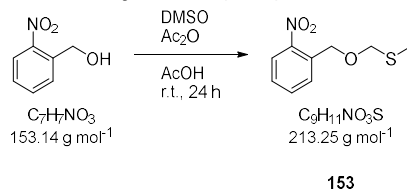
**<sup>31</sup>P{<sup>1</sup>H} NMR** (243 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 150.8, 150.2;

**HR-MS** (ESI+):  $m/z$  calc. ( $C_{54}H_{71}N_4NaO_8PSi$ ,  $[M+Na]^+$ ): 985.46765, found: 985.46781;

**TLC** (toluene/acetone 95:5):  $R_f = 0.70$ .

#### 4.2.4 Additional compounds and reagents

##### 4.2.4.1 2-Nitrobenzyl methylthiomethyl ether (153)

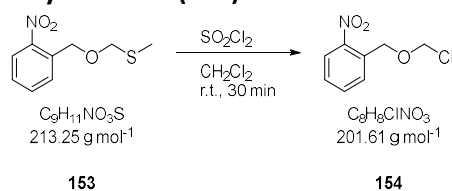


A solution of 2-nitrobenzylalcohol (323 mg, 2.17 mmol, 1.00 eq.),  $Ac_2O$  (15 mL) and AcOH (10 mL) in DMSO (20 mL) was stirred at ambient temperature for 24 h. Afterwards, the mixture was added dropwise to sat. aq.  $NaHCO_3$  (50 mL), stirred for 2 h and extracted with a 1:1 mixture of hexane and EtOAc (4×100 mL). The combined organic phases were dried over  $MgSO_4$ , evaporated and the residue was purified by column chromatography (hexane/EtOAc 9:1–7:3) to afford the title compound **153** as a dark oil (289 mg, 1.36 mmol, 63%). Spectral data matched those reported previously.<sup>222</sup>

**$^1H$  NMR** (400 MHz,  $CDCl_3$ ):  $\delta$  (ppm) = 8.02 (dd,  $J = 8.2, 1.3$  Hz, 1 H), 7.80–7.73 (m, 1 H), 7.69–7.60 (m, 1 H), 7.50–7.41 (m, 1 H), 4.99 (d,  $J = 0.8$  Hz, 2 H), 4.78 (s, 2 H), 2.20 (s, 3 H);

**TLC** (hexane/EtOAc 7:3):  $R_f = 0.58$ .

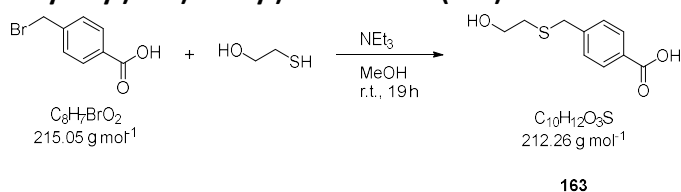
##### 4.2.4.2 2-Nitrobenzyloxymethyl chloride (154)



$SO_2Cl_2$  (207 mg, 1.54 mmol, 1.13 eq.) was added dropwise to a solution of compound **153** (289 mg, 1.36 mmol, 1.00 eq.) in  $CH_2Cl_2$  (10 mL) and stirred at ambient temperature for 30 min. Afterwards, the volatiles were removed under reduced pressure to afford the crude chloromethyl ether **154** as a dark oil (275 mg, 1.36 mg, > 99%) sufficiently pure for further reactions. Spectral data matches those reported previously.<sup>222</sup>

**$^1H$  NMR** (400 MHz,  $CDCl_3$ ):  $\delta$  (ppm) = 8.13 (dd,  $J = 8.2, 1.3$  Hz, 1 H), 7.79–7.72 (m, 1 H), 7.72–7.63 (m, 1 H), 7.54–7.45 (m, 1 H), 5.64 (s, 2 H), 5.16 (d,  $J = 1.0$  Hz, 2 H).

##### 4.2.4.3 4-(((2-Hydroxyethyl)thio)methyl)benzoic acid (163)



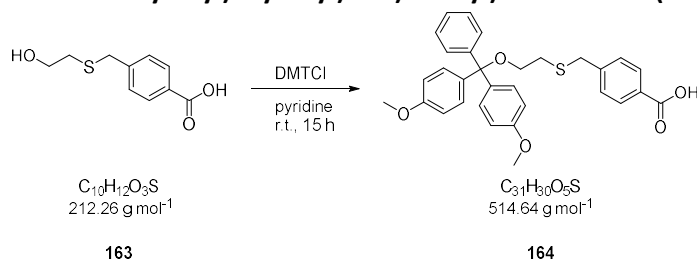
A solution of 4-bromomethylbenzoic acid (300 mg, 1.48 mmol, 1.00 eq.), 2-mercaptoethanol (126 mg, 1.62 mmol, 1.16 eq.) and  $NEt_3$  (280 mg, 2.76 mmol, 1.98 eq.) in MeOH (6 mL) was stirred at ambient temperature for 19 h. Afterwards, the volatiles were removed under reduced pressure and the residue was taken up in  $H_2O$  (12 mL). The solution was acidified by addition of HCl (1 M)

until precipitation of a white solid ceased. The resulting suspension was extracted with EtOAc (3×20 mL) and the combined organic phases were dried over Na<sub>2</sub>SO<sub>4</sub>. After evaporation of the solvent, the crude product was obtained as a white solid (289 mg, 1.36 mmol, 98 %) that was sufficiently pure for the next reaction.

<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>DN): δ (ppm) = 7.98–7.91 (m, 2 H), 7.49–7.40 (m, 2 H), 3.80 (s, 2 H), 3.59 (t, *J* = 6.6 Hz, 2 H), 2.53 (t, *J* = 6.6 Hz, 3 H);

TLC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5 + 1% AcOH): *R<sub>f</sub>* = 0.76.

#### 4.2.4.4 4-(((2-(4,4'-Dimethoxytrityl)oxyethyl)thio)methyl)benzoic acid (**164**)

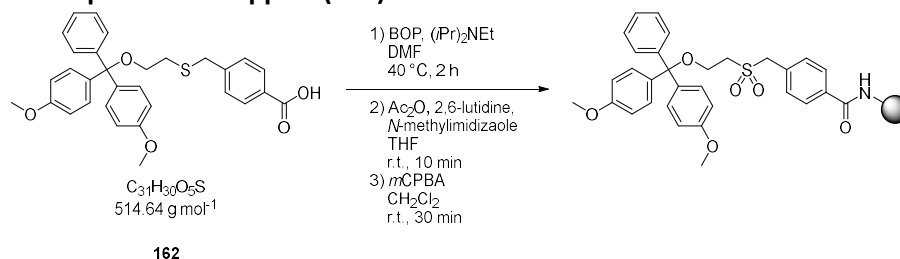


Compound **163** (284 mg, 1.34 mmol, 1.00 eq.) was coevaporated with pyridine (3×5 mL) and dissolved in anhydrous pyridine (7 mL). 4,4'-Dimethoxytrityl chloride (680 mg, 2.01 μmol, 1.50 eq.) was added in three portions over the course of 1 h and stirring was continued at ambient temperature for 14 h. The solvent was removed under reduced pressure. After coevaporation with toluene (3×5 mL), the residue was taken up in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and washed with citric acid (5% aq. Solution, 3×20 mL). The organic phase was dried over Na<sub>2</sub>SO<sub>4</sub>, evaporated and the residue was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1–90:10, the column was packed using 1% NEt<sub>3</sub> in the eluent) to afford an off-white foam (495 mg, 0.96 mmol, 72%). The product contained residual trimethylamine that did not impede the next reaction.

<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>DN): δ (ppm) = 7.90–7.83 (m, 2 H), 7.44–7.37 (m, 2 H), 7.32–7.18 (m, 9 H), 6.88–6.82 (m, 4 H), 3.75 (s, 6 H), 3.67 (s, 2 H), 3.09 (t, *J* = 6.7 Hz, 2 H), 2.59 (t, *J* = 6.7 Hz, 2 H);

TLC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5): *R<sub>f</sub>* = 0.13.

#### 4.2.4.5 3'-Phosphate CPG support (**162**)



In the following procedure, washing implies brief agitation with the respective solvent followed by centrifuging at 4200 rpm for 10 min and removal of the supernatant.

Native amino CPG support (250 mg, 18.5 μmol, 1.00 eq.) was washed with DMF (1.5 mL) and treated with a solution of compound **164** (25.0 mg, 48.7 μmol, 2.63 eq.), BOP reagent (22.0 mg, 49.8 μmol, 2.69 eq.) and (*i*Pr)<sub>2</sub>NEt (12 μL, 67.7 μmol, 3.66 eq.) in anhydrous DMF (1.5 mL). The reaction mixture was shaken at 40 °C for 2 h under exclusion of light. Afterwards, it was centrifuged at 4200 rpm for 5 min, the supernatant was removed and the solids were washed with anhydrous DMF (5 mL) followed by anhydrous MeCN (5 mL). A mixture of capping solutions A and

B (configured for ABI, 2 mL each) was added and the mixture was shaken at ambient temperature for 10 min. The supernatant was removed as described above and the solids were successively washed with MeCN, MeOH and CH<sub>2</sub>Cl<sub>2</sub> (5 mL each). Then, a solution of *m*CPBA (75% by weight, 173 mg, 750 μmol, 40.5 eq.) in CH<sub>2</sub>Cl<sub>2</sub> (12 mL) was added and the mixture was shaken at ambient temperature for 30 min. The supernatant was removed as described above, the solids were successively washed with MeCN and CH<sub>2</sub>Cl<sub>2</sub> (5 mL each) and the CPG support was carefully dried under reduced pressure. Two batches of **162** with loading densities of 51 and 56 μmol g<sup>-1</sup> were obtained.

#### Capping assay

Approximately 0.5 mg of **162** were suspended in ninhydrin solution (0.5 mL, 0.025% in EtOH) and heated to 65 °C for 2 min. No blue coloration due to uncapped amino groups was observed.

#### Loading density assay

Approximately 3.0 mg of **162** were suspended in dichloroacetic acid (10 mL, 3% solution in dichloroethane). An aliquot was removed and its absorbance spectrum between 200 and 800 nm was measured in a 1 cm cuvette. From the absorbance (*A*) of the DMT cation at 498 nm ( $\epsilon = 70000 \text{ L mol}^{-1} \text{ cm}^{-1}$ ) and the exact mass (*m*), the loading density was calculated as follows:

$$\frac{n}{m} = \frac{A}{\epsilon \cdot 1 \text{ cm}} \cdot \frac{10 \text{ mL}}{m}$$

### 4.3 OLIGONUCLEOTIDE SYNTHESIS

#### 4.3.1 In vitro transcription

Transcription reactions were performed on a 100  $\mu$ L scale as described in section 4.1.2.3. The required DNA templates, T7 promotor and transcript sequences are listed below.

#	5'-sequence-3' <sup>[a]</sup>	Length nt	Description
D008	CTGTAATACGACTCACTATA	20	T7 promotor
D551	GAGCTATCGCCGACCGCACCCAACCACCACTTTCATGTGGCGC CCCTCCAGCAATAGCCC TATAGTGAGTCGTATTACAG	80	Template for Tr044
D1082	GGTGCTCACTCTACTCAACAGTAGCGAACTACTGGACCCGT CCTTCACCTATAGTGAGTCGTATTACAG	71	Template for Tr104
D1086	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr108
D1132	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCTCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr123
D1133	GGCTAGCCGACCGCACCCAACTACCAGCGAACTGGCACCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr124
D1134	GGCTAGCCGACCGCACTCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr125
D1135	GGCTAGCCGACCGTACCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr126
D1136	GGCTAGCCGACTGCACCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr127
D1138	GGCTAGCTGACCGCACCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr129
D1187	GAGCCCACACTCTACTCGACAGATACGAATATCTGGACCCGAC CGTCTCTATAGTGAGTCGTATTACAG	69	Template for Tr188
D1411	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCCCTC TAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr189
D1412	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCCCTT CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr190
D1413	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCCTTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr191
D1414	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCTCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr192
D1415	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGTCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr193
D1417	GGCTAGCCGACCGCACCCAACTACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr195
D1418	GGCTAGCCGACCGCACCCAATCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr196
D1419	GGCTAGCCGACCGCACCTAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr197
D1420	GGCTAGCCGACCGCATCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr198
D1421	GGCTAGCCGATCGCACCCAACCACCAAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr199
D1422	GGCTAGCCGACCGCACCCAACCACCAAGCGAACTGGCGCCCCTC CGGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr200

D1492	GGCTAGCCGACCGCACCCAACCACCAGCGAACTGGCGCCCCAC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr214
D1493	GGCTAGCCGACCGCACCCAACCACCAGCGAACTGGCACCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr215
D1495	GGCTAGCCGACCGCACCCAGCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr217
D1496	GGCTAGCCGACCGCACCCGACCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr218
D1497	GGCTAGCCGACCGCGCCCAACCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr219
D1498	GGCTAGCCGACCACACCCAACCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr220
D1499	GGCTAGCCGGCCGCACCCAACCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr221
D1500	GGCTAGCCAACCGCACCCAACCACCAGCGAACTGGCGCCCCTC CAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr222
D1639	GGCAGCTGACCGCACCCAACCACCAGCGAACTGGCGCCCCTCC AGCTGCCATAGTGAGTCGTATTACAG	70	Template for Tr303
D1640	GGAGCTGACCGCACCCAACCACCAGCGAACTGGCGCCCCTCCA GCTCCTATAGTGAGTCGTATTACAG	68	Template for Tr304
D1641	GGACTGACCGCACCCAACCACCAGCGAACTGGCGCCCCTCCAG TCCTATAGTGAGTCGTATTACAG	66	Template for Tr305
D1939	GGCTAGCCAACCGCACCCAACCACCAGCGAACTGGCGCCCCTC TAGCTAGCCTATAGTGAGTCGTATTACAG	72	Template for Tr392
D1942	GGACCGACCGCACCCAACCACCTTCCGGCGCCCCTCCAGTCCT ATAGTGAGTCGTATTACAG	64	Template for Tr395

[a] For visual orientation, the T7 promotor and corresponding annealing sites are highlighted in green.

#	5'-sequence-3' [a]	Length nt	Description	Promo- tor
Tr044	GGGCUAUUGCUGGAGGGGCGCCACAUGAAAGUGGU GGUUGGGUGCGGUCGGCGAUAGCUC	60	13-2 aptamer	D551
Tr104	GGUGAAGGACGGGUCCAGUAGUUCGCUACUGUUGA GUAGAGUGUGAGCACC	51	Baby Spinach aptamer	D1082
Tr108	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili aptamer wt	D1086
Tr123	GGCUAGCUGGAGAGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G13A	D1132
Tr124	GGCUAGCUGGAGGGGUGCCAGUUCGCUGGUAGUUG GGUGCGGUCGGCUAGCC	52	Chili C16U/G31A	D1133
Tr125	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG AGUGCGGUCGGCUAGCC	52	Chili G36A	D1134
Tr126	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUACGGUCGGCUAGCC	52	Chili G39A	D1135
Tr127	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCAGUCGGCUAGCC	52	Chili G41A	D1136
Tr129	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCAGCUAGCC	52	Chili G45A	D1138
Tr188	GAGACGGUCGGGUCCAGAUUUCGUAUCUGUCGAG UAGAGUGUGGGCUC	49	Broccoli aptamer	D1187

Tr189	GGCUAGCUAGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G9A	D1411
Tr190	GGCUAGCUGAAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G10A	D1412
Tr191	GGCUAGCUGGAAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G12A	D1413
Tr192	GGCUAGCUGGAGGAGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G14A	D1414
Tr193	GGCUAGCUGGAGGGACGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili G15A	D1415
Tr195	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUAGUUG GGUGCGGUCGGCUAGCC	52	Chili G31A	D1417
Tr196	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGAUUG GGUGCGGUCGGCUAGCC	52	Chili G32A	D1418
Tr197	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUA GGUGCGGUCGGCUAGCC	52	Chili G35A	D1419
Tr198	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GAUGCGGUCGGCUAGCC	52	Chili G37A	D1420
Tr199	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGAUCGGCUAGCC	52	Chili G42A	D1421
Tr200	GGCUAGCCGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili U8C	D1422
Tr214	GGCUAGCUGGUGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili A11A	D1492
Tr215	GGCUAGCUGGAGGGGUGCCAGUUCGCUGGUGGUUG GGUGCGGUCGGCUAGCC	52	Chili C16U	D1493
Tr217	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGCUG GGUGCGGUCGGCUAGCC	52	Chili U33C	D1495
Tr218	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUCG GGUGCGGUCGGCUAGCC	52	Chili U34C	D1496
Tr219	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGCGCGGUCGGCUAGCC	52	Chili U38C	D1497
Tr220	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGUGGUCGGCUAGCC	52	Chili C40U	D1498
Tr221	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGCCGGCUAGCC	52	Chili U43C	D1499
Tr222	GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUUGGCUAGCC	52	Chili C44U	D1500
Tr303	GGCAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUGG GUGCGGUCAGCUGCC	50	Chili G45A 7 bp stem	D1639
Tr304	GGAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUGGG UGC GGUCAGCUCC	48	Chili G45A 6 bp stem	D1640
Tr305	GGACUGGAGGGGCGCCAGUUCGCUGGUGGUUGGGU GCGGUCAGUCC	46	Chili G45A 5 bp stem	D1641
Tr392	GGCUAGCUAGAGGGGCGCCAGUUCGCUGGUGGUUG GGUGCGGUUGGCUAGCC	52	Chili G9A/C44U	D1939
Tr395	GGACUGGAGGGGCGCCGGAAGGUGGUUGGGUGCGG UCGGUCC	42	Chili 5 bp mini	D1942

[a] For visual orientation, the T7 promotor and corresponding annealing sites are highlighted in green. Loop structures in the transcripts are highlighted in blue.



### 4.3.2 Solid-phase synthesis

RNA solid-phase syntheses were performed on a 0.6  $\mu\text{mol}$  scale as described in section 4.1.3.3. The synthesis products and corresponding HR-MS results (see section 4.1.6.2) are listed below.

#	5'-sequence-3' [a, b, c]	Length nt	Description
R610	GGCUAGCUGGAGGGGCGCCAGUUCGC	26	Chili 5' fragment
R611	pUGGUGGUUUGGGUGCGGUCGGCUAGCCp	26	Chili 3' fragment
R681	GGCUAGCUGACGCCAG	16	$T_m$ sample A
R682	GGCUAGCUGCCGCCAG	16	$T_m$ sample C
R683	GGCUAGCUGGGGCCAG	16	$T_m$ sample G
R684	GGCUAGCUGUCGCCAG	16	$T_m$ sample U
R685	CUGGCGUCAGCUAGCC	16	$T_m$ complement U
R686	CUGGCGGCAGCUAGCC	16	$T_m$ complement G
R687	CUGGCGCCAGCUAGCC	16	$T_m$ complement C
R688	CUGGCGACAGCUAGCC	16	$T_m$ complement A
R689	CUGGCGXCAGCUAGCC	16	$T_m$ complement X
R693	GGCUAGCUGGXGGGGCGCCAGUUCGC	26	Chili 5' fragment A11X
R735	pUGGUGGUUUGYUGCGGUCGGCUAGCCp	26	Chili 3' fragment G36Y
R737	pUGGUGGUUUGXUGCGGUCGGCUAGCCp	26	Chili 3' fragment G36X
R750	GGCUAXCUGGAGGGGCGCCAGUUCGC	26	Chili 5' fragment G6X
R751	GGCUAGCUGGAGGGGCGCCAGUUCGC	26	Chili 5' fragment G17X
R752	GGCUAGCUGGAGGGGCGCCAGUXCGC	26	Chili 5' fragment U23X
R753	pUGGUGGUUGGGUGCGGUCGGCUAGCCp	26	Chili 3' fragment U34X
R754	pUGGUGGUUUGGGUGXGGUCGGCUAGCCp	26	Chili 3' fragment C40X
R755	pUGGUGGUUUGGGUGCGGUCXGCUAGCCp	26	Chili 3' fragment G45X
R756	pUGGUGGUUUGGGUGCGGUCGXCUAGCCp	26	Chili 3' fragment G46X

[a] 5'-p indicates use of the chemical phosphorylation reagent **161** for solid-phase synthesis.

[b] 3'-p indicates use of the 3'-phosphate CPG support **162** for solid-phase synthesis.

[c] X: r4Cl, Y: d4Cl.

### 4.3.3 Enzymatic ligation

Ligation reactions were performed on a 50  $\mu\text{L}$  scale as described in section 4.1.4.3. The ligation products are listed below.

5'-sequence-3' [a]	Length nt	Description
GGCUAXCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCGGCUAGCCp	52	Chili G6X
GGCUAGCUGGXGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCGGCUAGCCp	52	Chili A11X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCGGCUAGCCp	52	Chili G17X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCGGCUAGCCp	52	Chili U23X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUGGGUGCGGUCGGCUAGCCp	52	Chili U34X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUGXUGCGGUCGGCUAGCCp	52	Chili G36X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGYUGCGGUCGGCUAGCCp	52	Chili G36Y
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCXGCUAGCCp	52	Chili G45X
GGCUAGCUGGAGGGGCGCCAGUUCGCUGGUGGUUUGGGUGCGGUCGXCUAGCCp	52	Chili G46X

[a] X: r4Cl, Y: d4Cl

## 4.4 COMPUTATIONAL METHODS

### 4.4.1 Geometry optimization

DFT-optimized geometries were calculated with the software package ORCA version 4.0.1.2<sup>285,286</sup> using the B3LYP functional with D3BJ dispersion correction,<sup>287,288</sup> a def2-TZVP basis set<sup>289,290</sup> and the corresponding auxiliary basis set for the RIJCOSX approximation<sup>291</sup> on all atoms. Stationary points were characterized as minima on the potential energy surface by analytical frequency calculations. Tight convergence criteria were used throughout.

### 4.4.2 Electrostatic potential maps

Electrostatic potential maps were generated from a cube file of the electron density with a resolution of 80×80×80 grid points using a python script that is available online.<sup>292</sup>

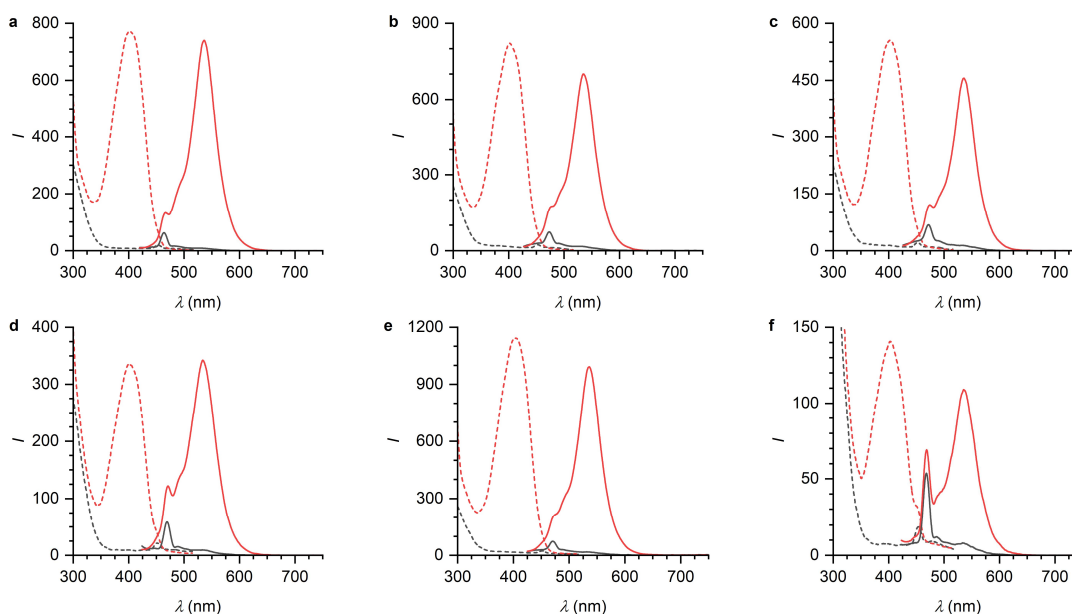
### 4.4.3 Sample input file

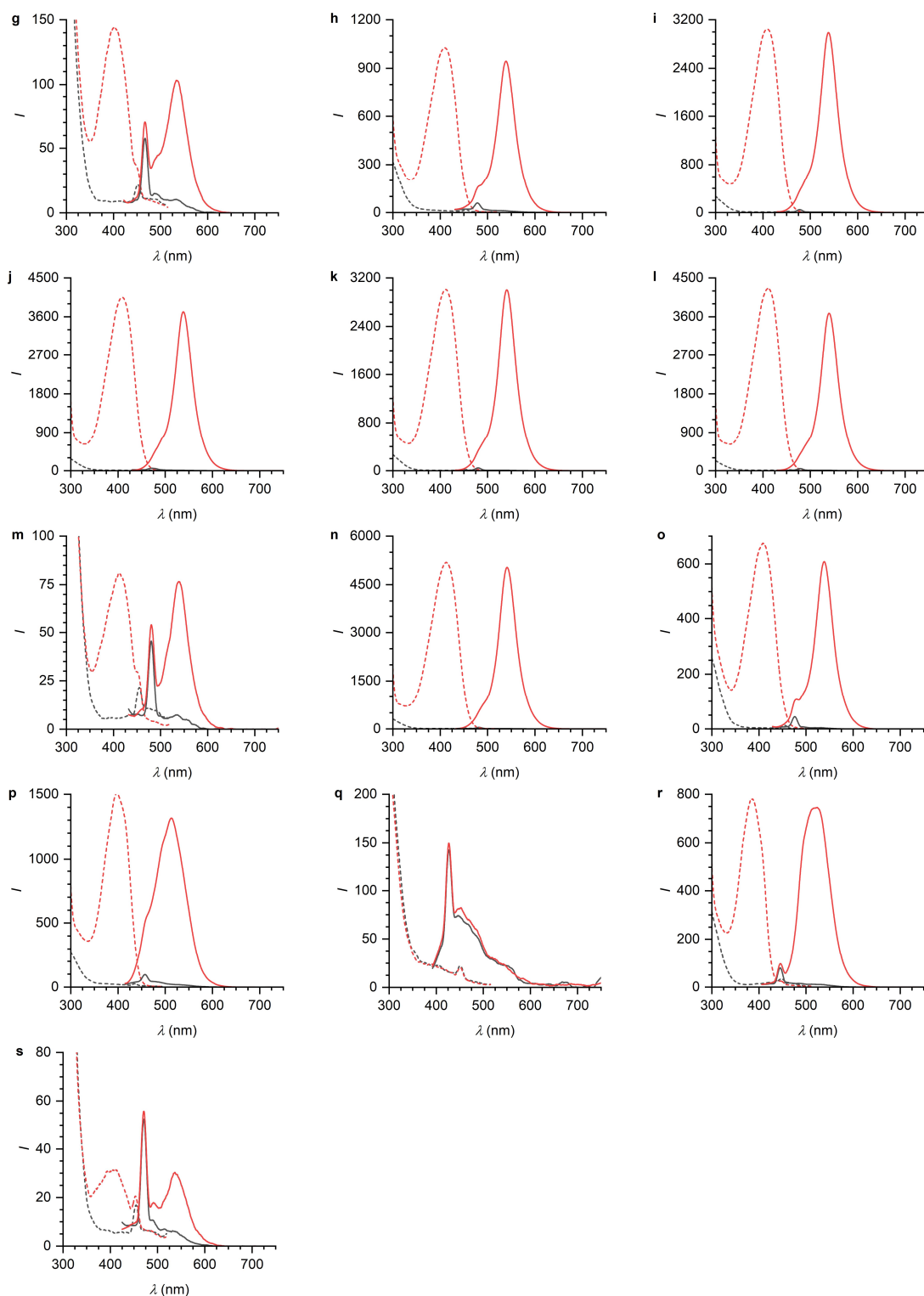
Starting from an xyz file which contains the initial model coordinates, the following script requests a geometry optimization followed by an analytical frequency calculation. An augmented basis set is assigned to the heavy atom iodine. At the end of the optimization, an electron density file is made available for further calculations.

```
# !B3LYP D3BJ def2-TZVP RIJCOSX def2/J Grid5 FinalGrid6 GridX6 TightSCF TightOpt Freq KeepDens
#
# %basis
#      newgto I "ma-def2-TZVP" end
# end
#
# *xyzfile 0 1 dmhb_p-trimethylammoniumphenyl_i_iodide_start.xyz
```

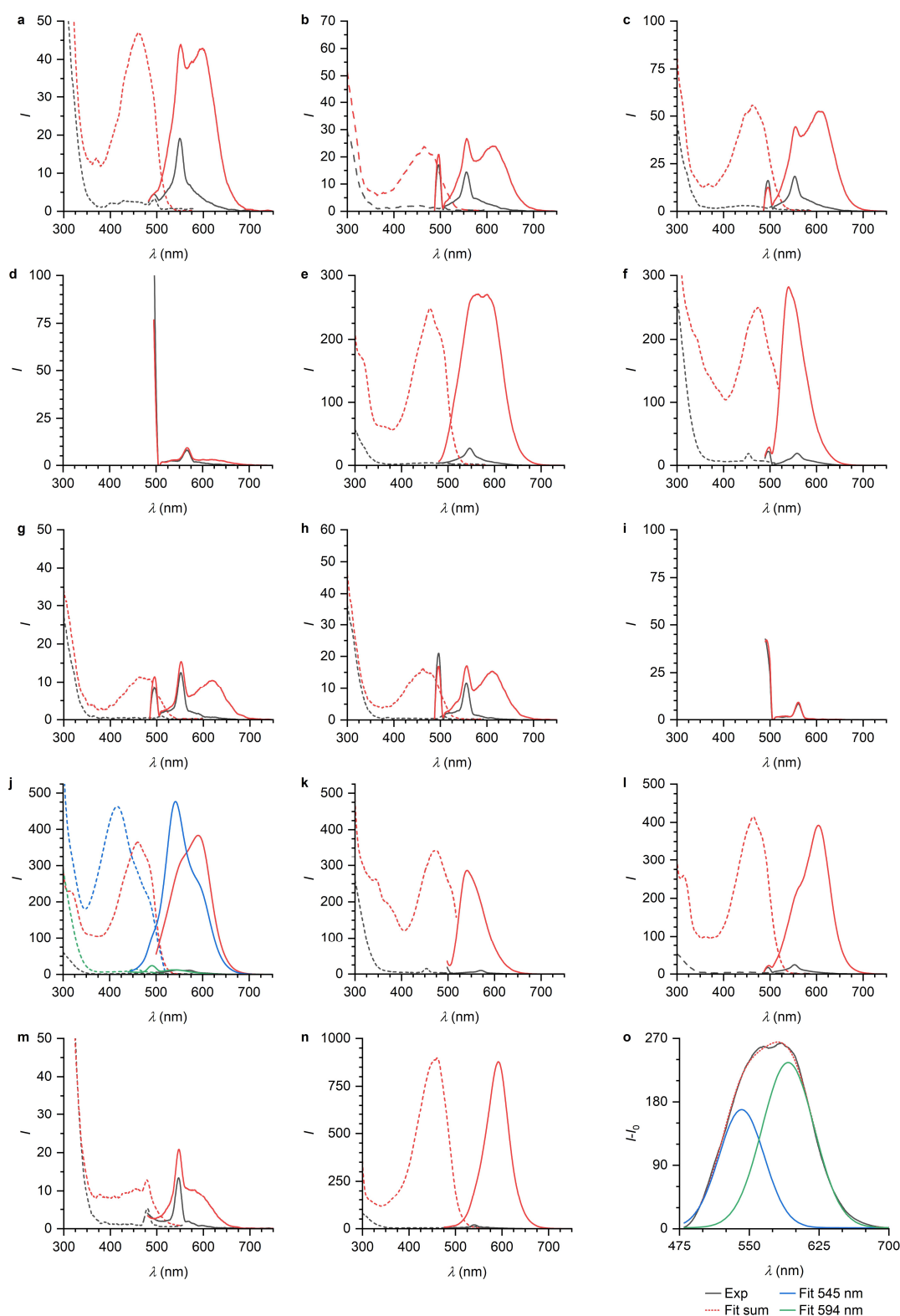
## 5 APPENDIX

### 5.1 SUPPLEMENTARY FIGURES

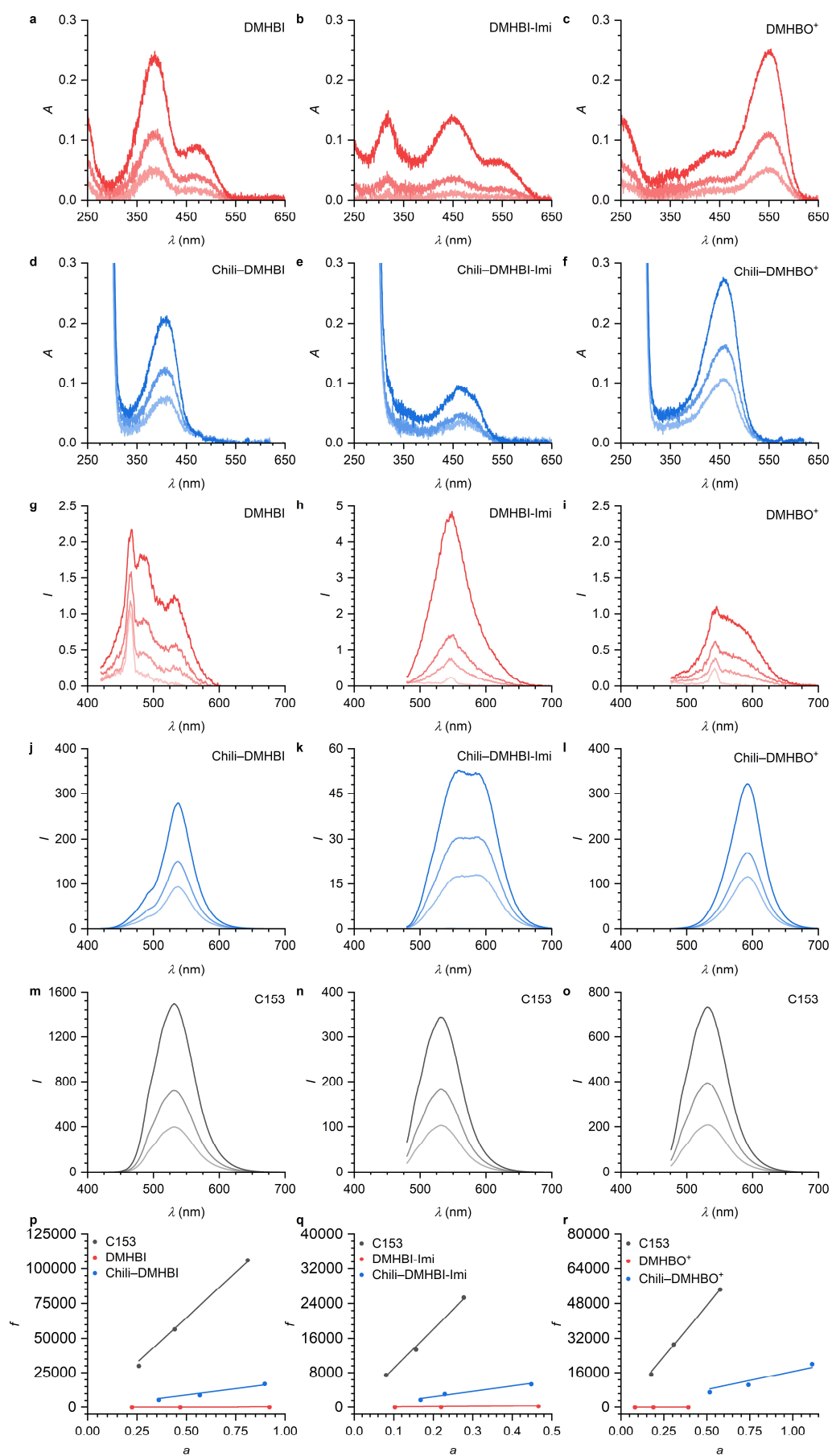




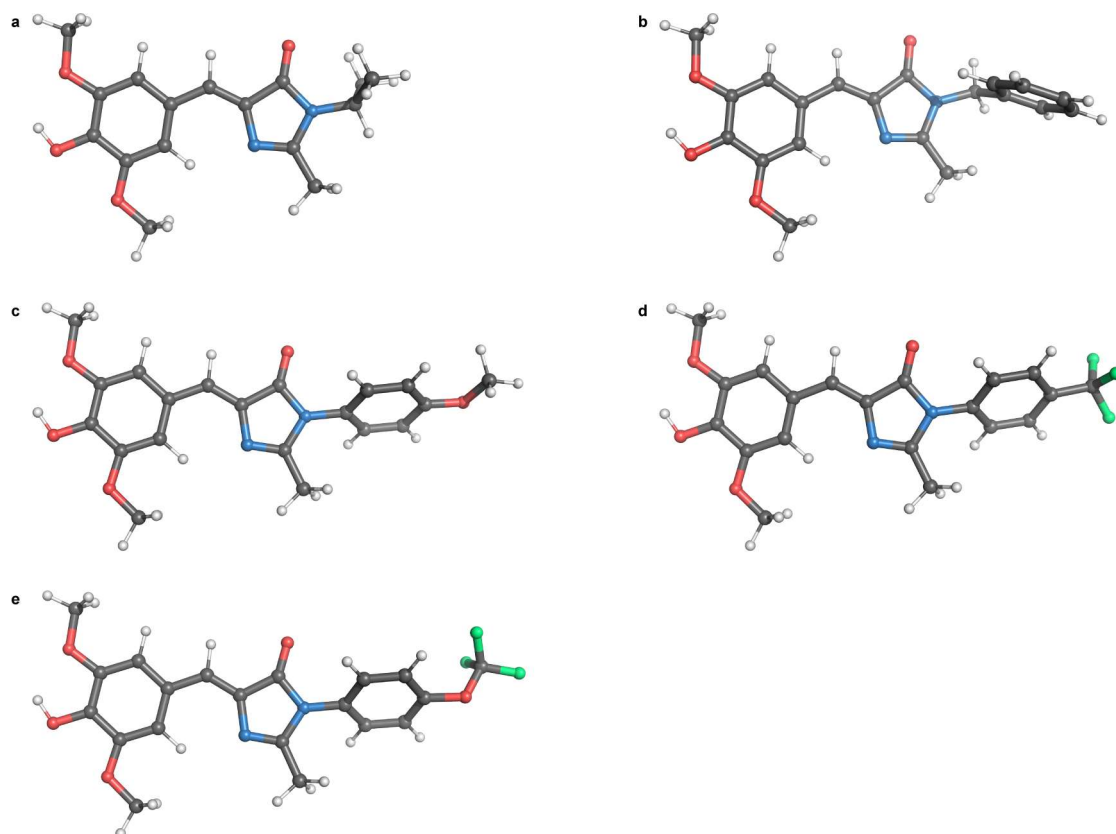
**Supplementary Figure 1.** Uncorrected fluorescence emission (solid) and excitation (dashed) spectra of Chili-HBI complexes with green fluorescence (red) and of the respective HBI ligands alone (black). The excitation and emission wavelengths used to obtain the spectra are given in parentheses. a) DMHBI (**13**, 400/537 nm), b) DMHBI-Et (**61**, 400/537 nm), c) DMHBI-*i*Pr (**62**, 400/537 nm), d) DMHBI-*t*Bu (**63**, 400/534 nm), e) DMHBI-MeCy (**64**, 400/537 nm), f) DMHBI-Bn (**65**, 400/537 nm), g) DMHBI-PMBn (**66**, 400/535 nm), h) DMHBPI (**67**, 410/539 nm), i) DMHBTI (**44**, 410/539 nm), j) DMHBAl (**68**, 410/538 nm), k) DMHBTI<sup>F</sup> (**69**, 413/540 nm), l) DMHBAl<sup>F</sup> (**70**, 413/540 nm), m) DMHBI-DMA (**71**, 413/540 nm), n) DMHBI<sup>+</sup> (**72**, 413/542 nm), o) DMHBI<sup>C</sup> (**73**, 410/539 nm), p) MHBAI (**74**, 395/513 nm), q) DMBAI (**75**, 372/535 nm, no signal was obtained at these or any other wavelengths), r) BMHBI (**77**, 386/520 nm), s) DMHBI-PhEt (**80**, 400/539 nm).



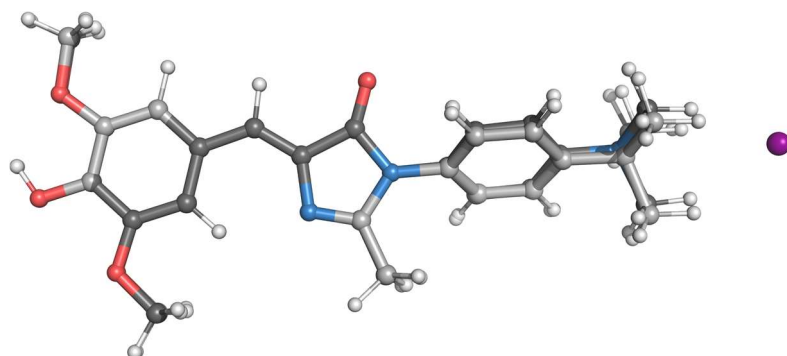
**Supplementary Figure 2.** Uncorrected fluorescence emission (solid) and excitation (dashed) spectra of the Chili-HBI complexes with  $\pi$ -conjugated C2 substituents, i.e. with red fluorescence (red), and of the respective HBI ligands alone (black). The respective excitation and emission wavelengths used to obtain the spectra are given in parentheses. a) DMHBI-Styr (**81**, 462/601 nm), b) DMHBI-2Py (**82**, 467/616 nm), c) DMHBI-3Py (**83**, 465/611 nm), d) DMHBI-4Py (**84**, 475/–, no signal was obtained at this or any other wavelength), e) DMHBI-Imi, (**85**, 463/594 nm), f) DMHBI-Ind (**86**, 469/539 nm), g) DMHBI-2Py (**87** 464/618 nm), h) DMHBI-3Py (**88**, 467/613 nm), i) DMHBI-4Py (**89**, 470/–, no signal was obtained at this or any other wavelength), j) DMHBTI-Imi (**90**, black/red: 480/598 nm, green/blue: 420/541 nm), k) DMHBTI-Ind (**91**, 478/539 nm), l) DMHBI-Styr<sup>+</sup> (**95**, 465/603 nm), m) DMHBI-Fc (**96**, 460/573 nm), n) DMHBO<sup>+</sup> (**98**, 456/592 nm). o) The blank-corrected emission spectrum of Chili in complex with DMHBI-Imi (**85**) was deconvoluted with two Gaussian peaks that are centered at 545 and 594 nm, respectively.



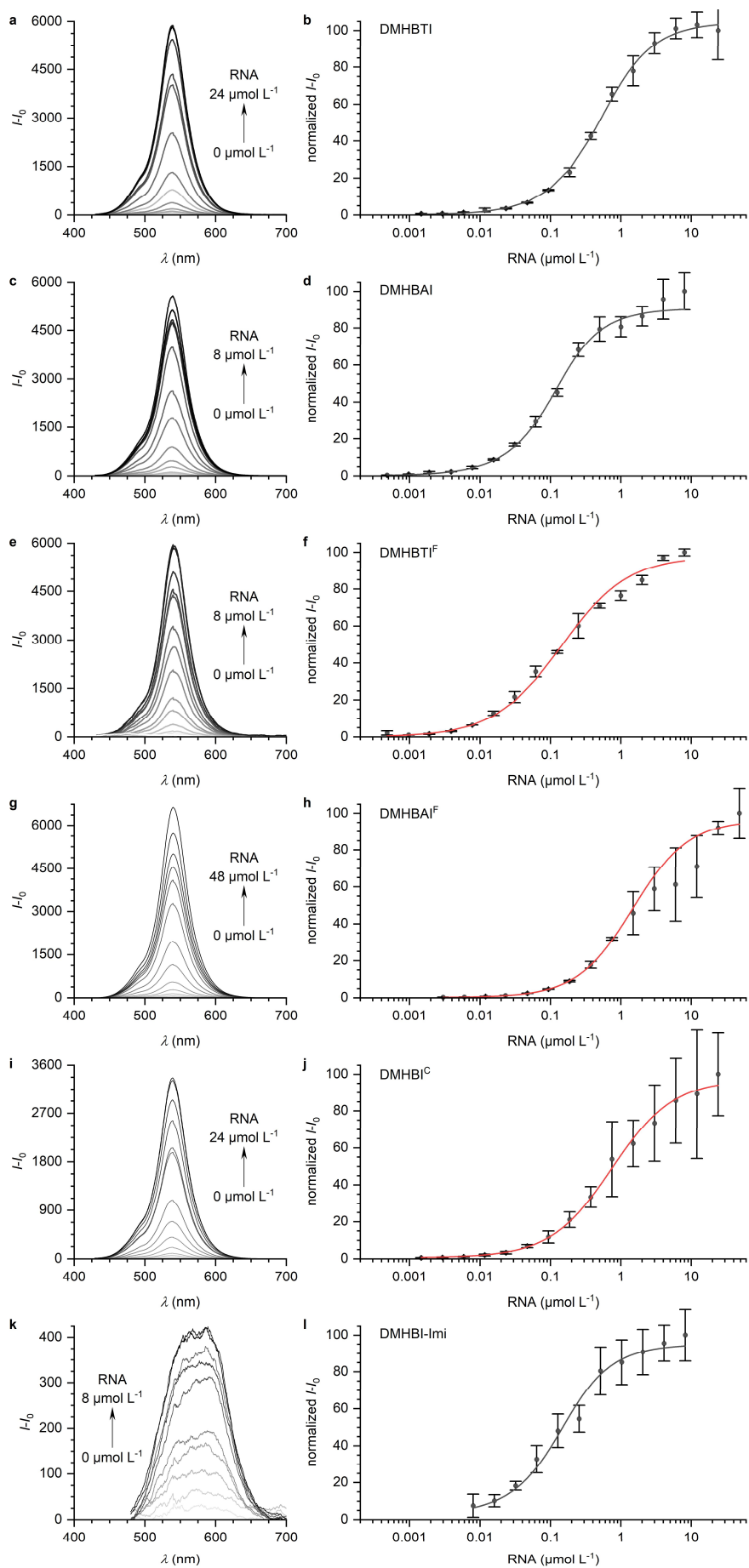
**Supplementary Figure 3.** a–g) UV/Vis spectra of DMHBI (**13**), DMHBI-Imi (**85**) and DMHBO<sup>+</sup> (**98**) alone or in complex with Chili at concentrations of 2.5, 5 and 10  $\mu\text{M}$ . g–l) Fluorescence emission spectra ( $\lambda_{\text{ex}} = 400, 460$  or 456 nm for DMHBI, DMHBI-Imi and DMHBO<sup>+</sup>, respectively) of the same samples in addition to blank samples (buffer). m–o) Fluorescence emission spectra ( $\lambda_{\text{ex}} = 400, 460$  or 456 nm) of Coumarin 153 in EtOH at concentrations of 2.5, 5 and 10  $\mu\text{M}$ . p–r) Plot of integrated fluorescence intensities ( $f$ ) vs absorption factors ( $a$ ) for each row of spectra. The corresponding absorption spectra of Coumarin 153 are shown in Figure 34a. Solid lines represent linear fits.



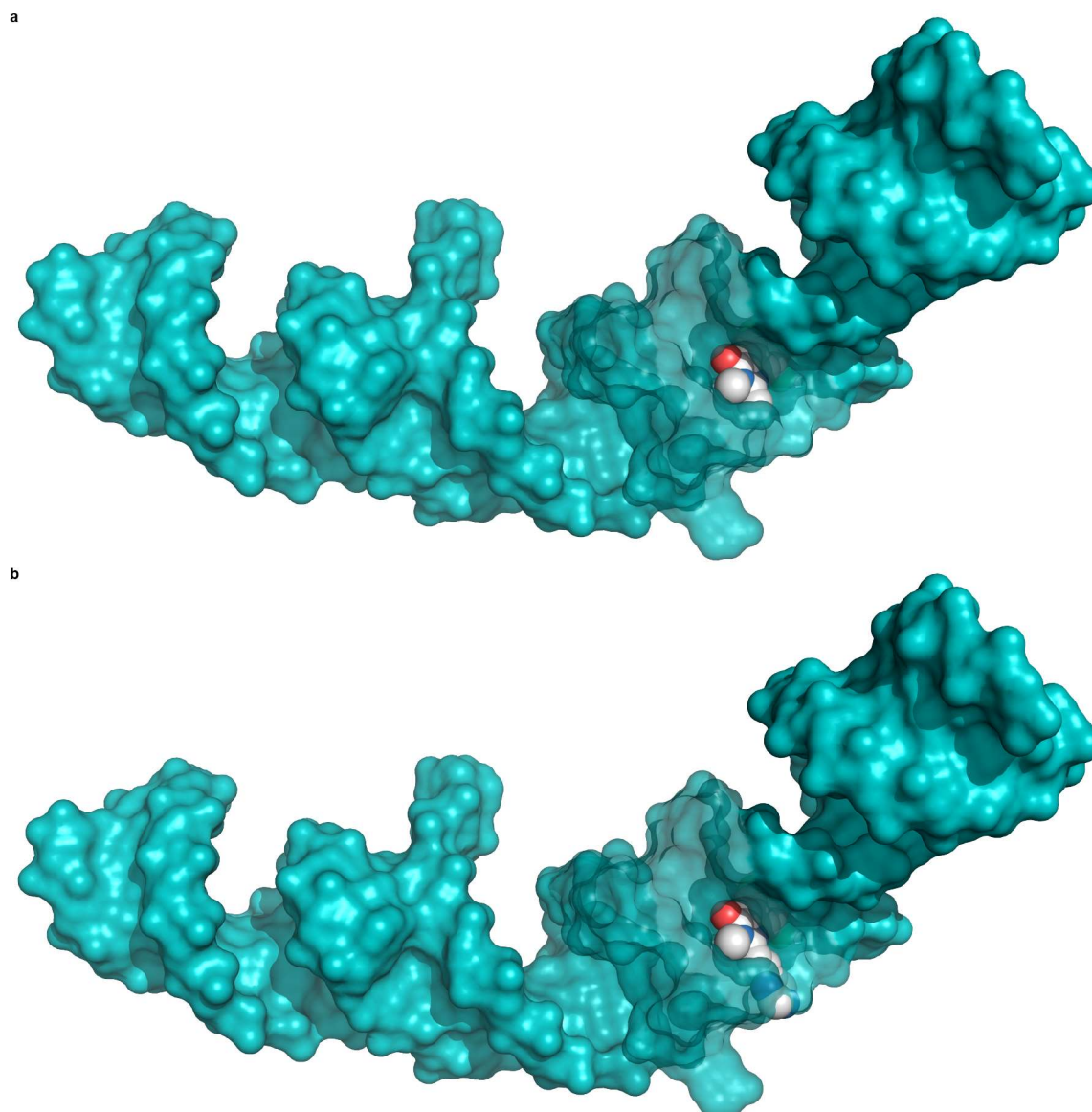
**Supplementary Figure 4.** DFT-optimized geometries of a) DMHBI-*i*Pr (**62**), b) DMHBI-Bn (**65**) c) DMHBAl (**68**) d) DMHBI<sup>F</sup> (**69**) and d) DMHBAl<sup>F</sup> (**70**) in the gas phase (B3LYP-D3/def2-TZVP).



**Supplementary Figure 5.** Structural alignment of DMHBI<sup>+</sup> (**72**, carbon atoms in dark grey) and DMHBI<sup>c</sup> (**73**, carbon atoms in light grey). There is minimal deviation between the two molecules in the gas phase (DFT: B3LYP-D3/def2-TZVP).

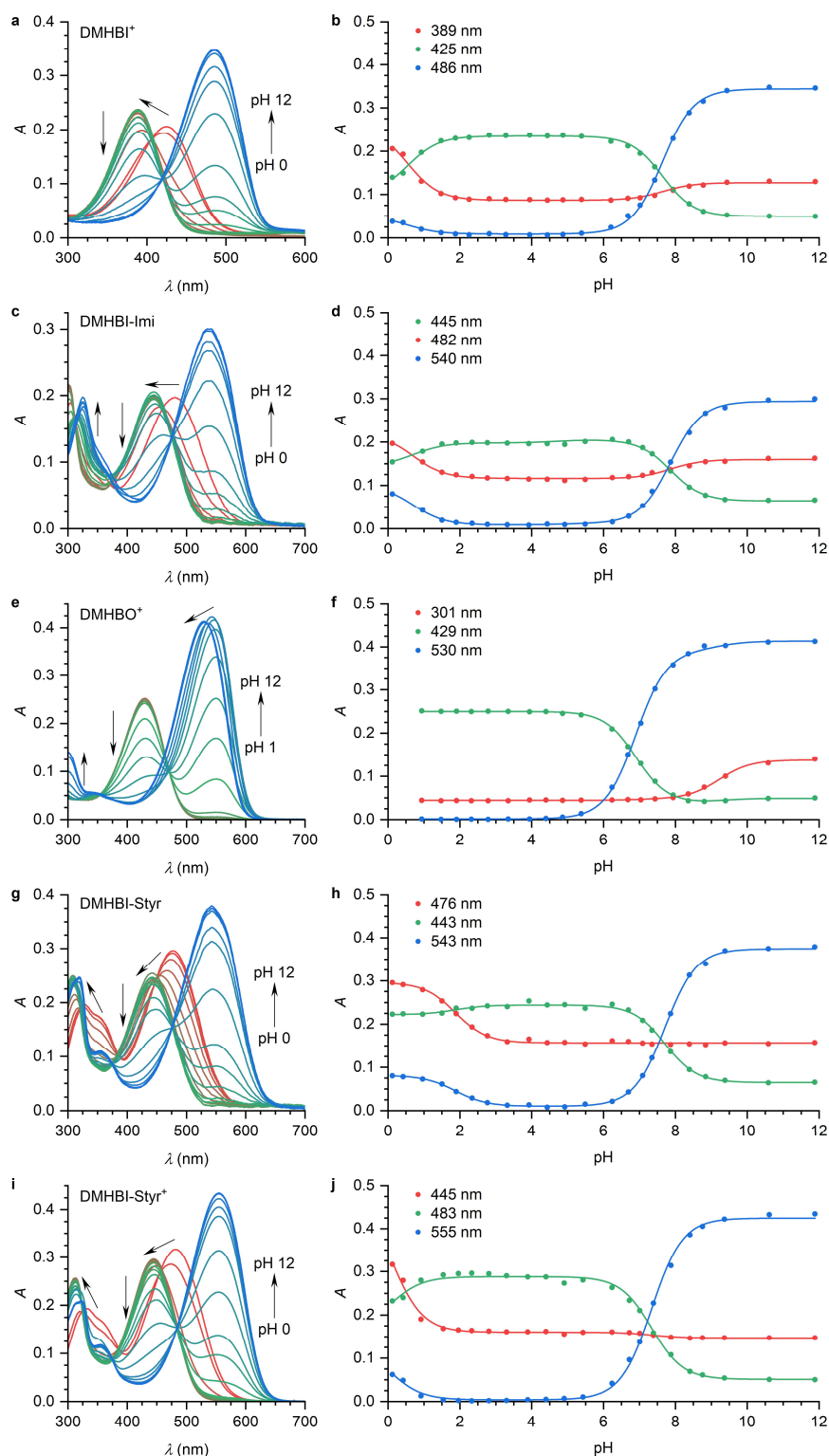


**Supplementary Figure 6.** Fluorescence enhancement (left) and titration curves (right) of various HBI derivatives with the Chili aptamer. a, b) DMHBTI (**44**, 0–24  $\mu\text{M}$  RNA, 0.3  $\mu\text{M}$  ligand,  $K_d = 1.5 \pm 0.2 \times 10^{-7} \text{ mol L}^{-1}$ ), c, d) DMHBAl (**68**, 0–8  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand,  $K_d = 6.5 \pm 0.7 \times 10^{-8} \text{ mol L}^{-1}$ ), e, f) DMHBTI<sup>F</sup> (**69**, 0–8  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand,  $K_{\text{Hill}} = 1.4 \pm 0.1 \times 10^{-7} \text{ mol L}^{-1}$ ), g, h) DMHBAl<sup>F</sup> (**70**, 0–48  $\mu\text{M}$  RNA, 0.5  $\mu\text{M}$  ligand,  $K_{\text{Hill}} = 1.5 \pm 0.2 \times 10^{-6} \text{ mol L}^{-1}$ ), i, j) DMHBI<sup>C</sup> (**73**, 0–24  $\mu\text{M}$  RNA, 0.5  $\mu\text{M}$  ligand,  $K_{\text{Hill}} = 7.4 \pm 0.7 \times 10^{-7} \text{ mol L}^{-1}$ ) and k, l) DMHBI-Imi (**85**, 0–8  $\mu\text{M}$  RNA, 0.1  $\mu\text{M}$  ligand,  $K_d = 7.1 \pm 0.2 \times 10^{-8} \text{ mol L}^{-1}$ ). Data points were fitted with a single-site binding model (black) or with the Hill equation (red). Conditions: 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5. Fluorescence spectra were integrated and normalized after blank correction (buffer). Values are given as mean  $\pm$  s.d. (n = 3).

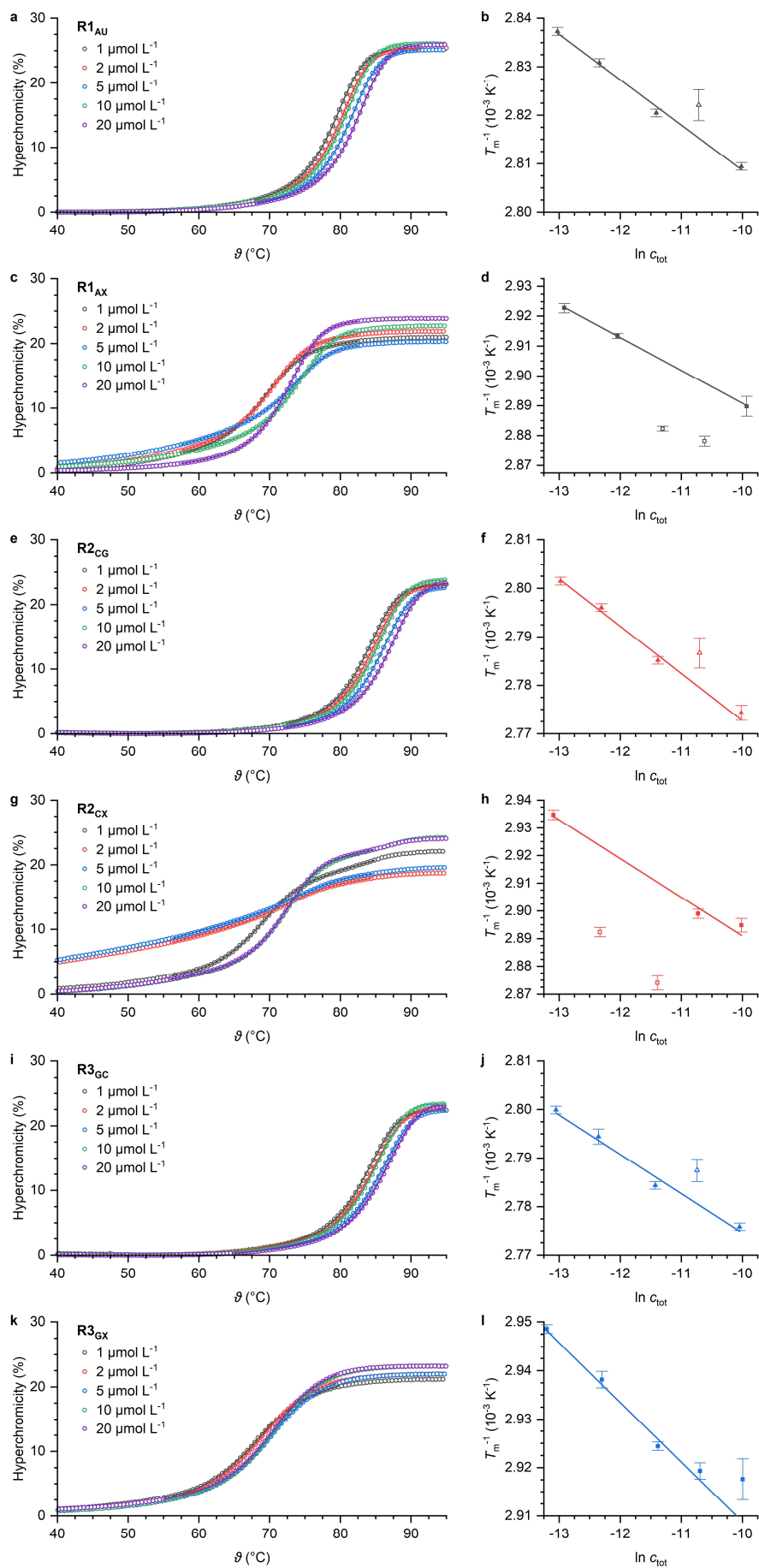


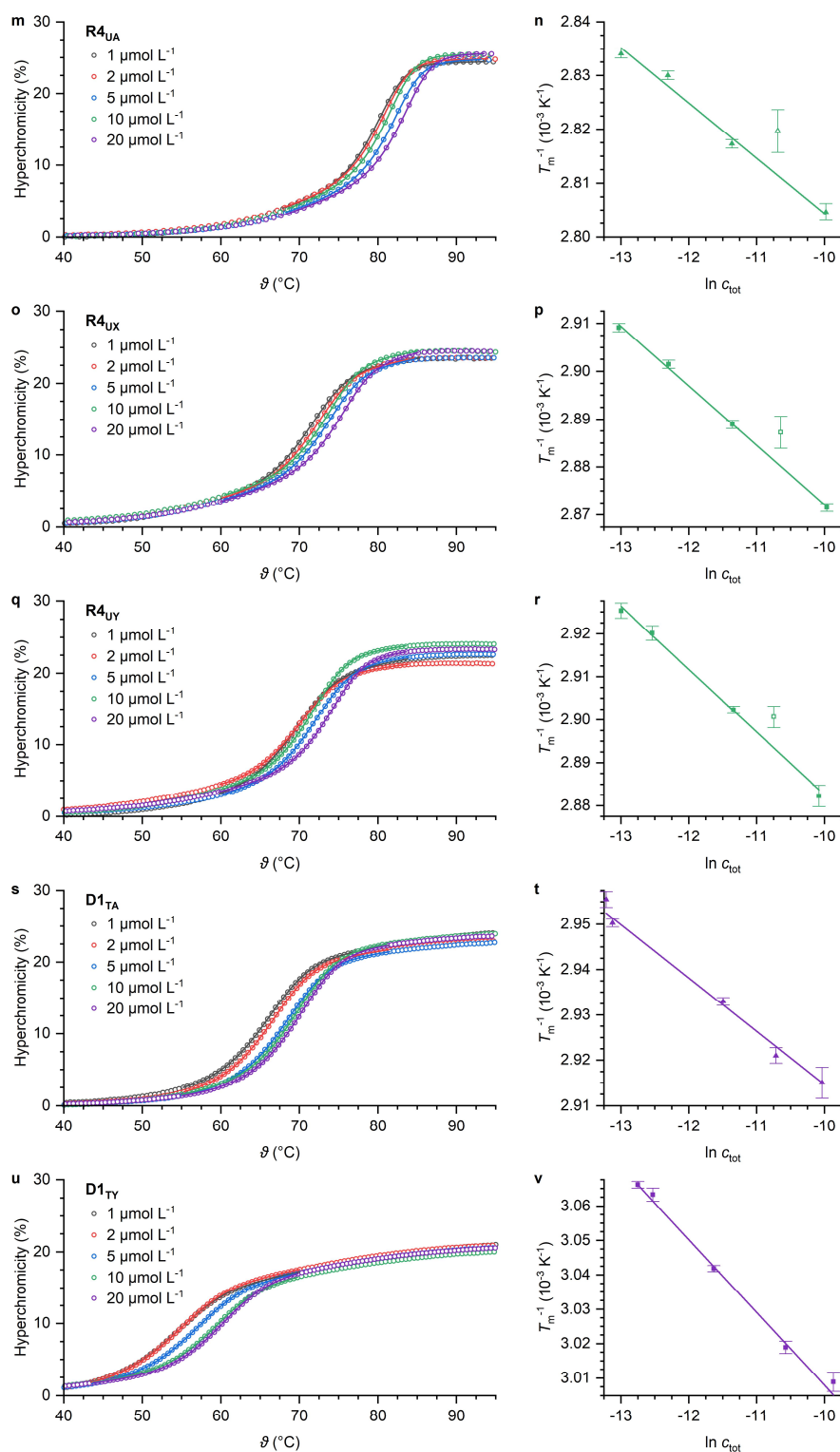
**Supplementary Figure 7.** Surface representation of the Spinach aptamer (PDB: 4TS2) in complex with a) DFHBI (**14**) and b) a hypothetical DFHBI derivative bearing a vinyl-4-imazolyl side chain at imidazolone-C2 (analogous to DMHBI-Imi, **85**). The extended side chain clashes with the backbone at the binding site.



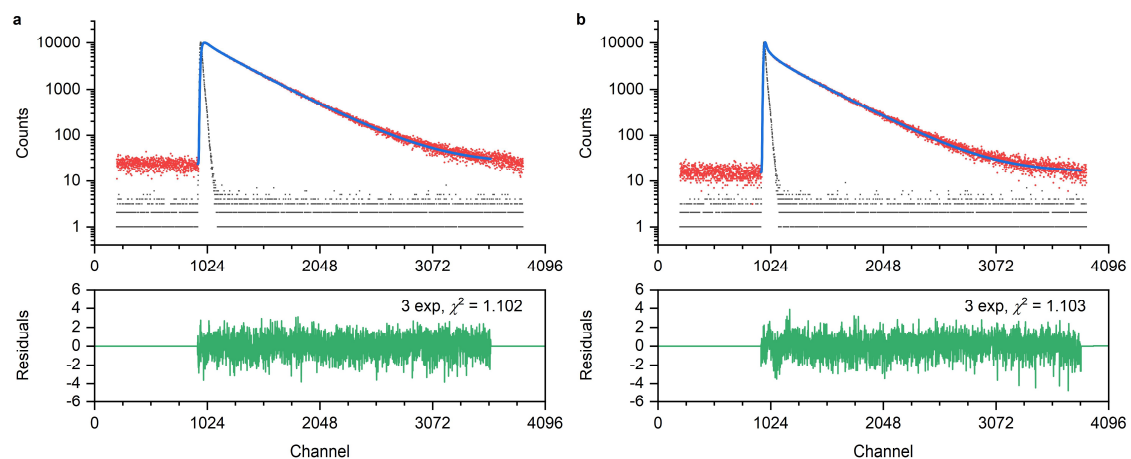


**Supplementary Figure 8.** UV/Vis spectra (left) and titration curves (right) at the absorption maxima of the cationic (red), neutral (green) and anionic (blue) species of various HBI derivatives in aqueous buffer. Spectra were measured at 20–22 pH levels between 0 and 12.  $pK_a$  values were obtained by global fitting of the absorbance with the Henderson-Hasselbalch equation and are reported in Supplementary Table 1. Conditions: 11  $\mu\text{M}$  ligand, 30 mM Britton-Robinson buffer.

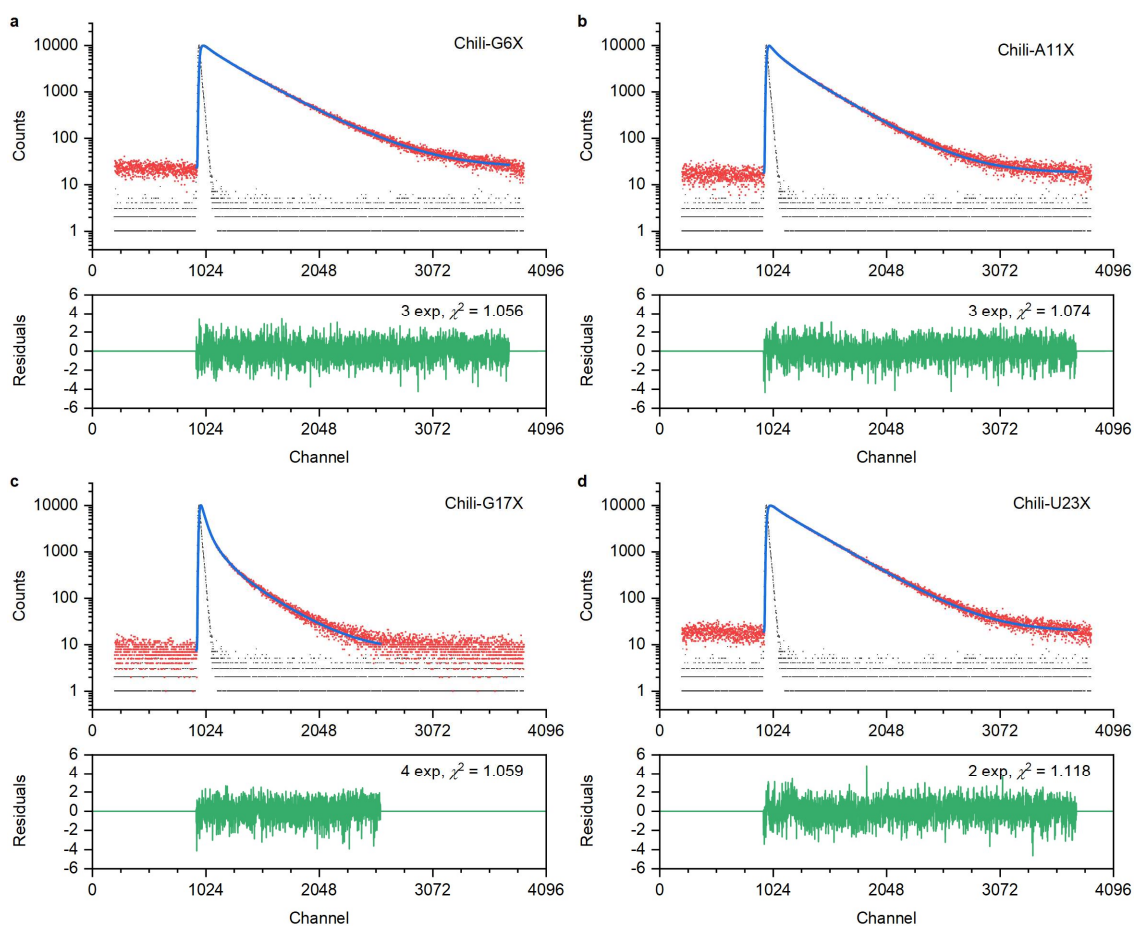


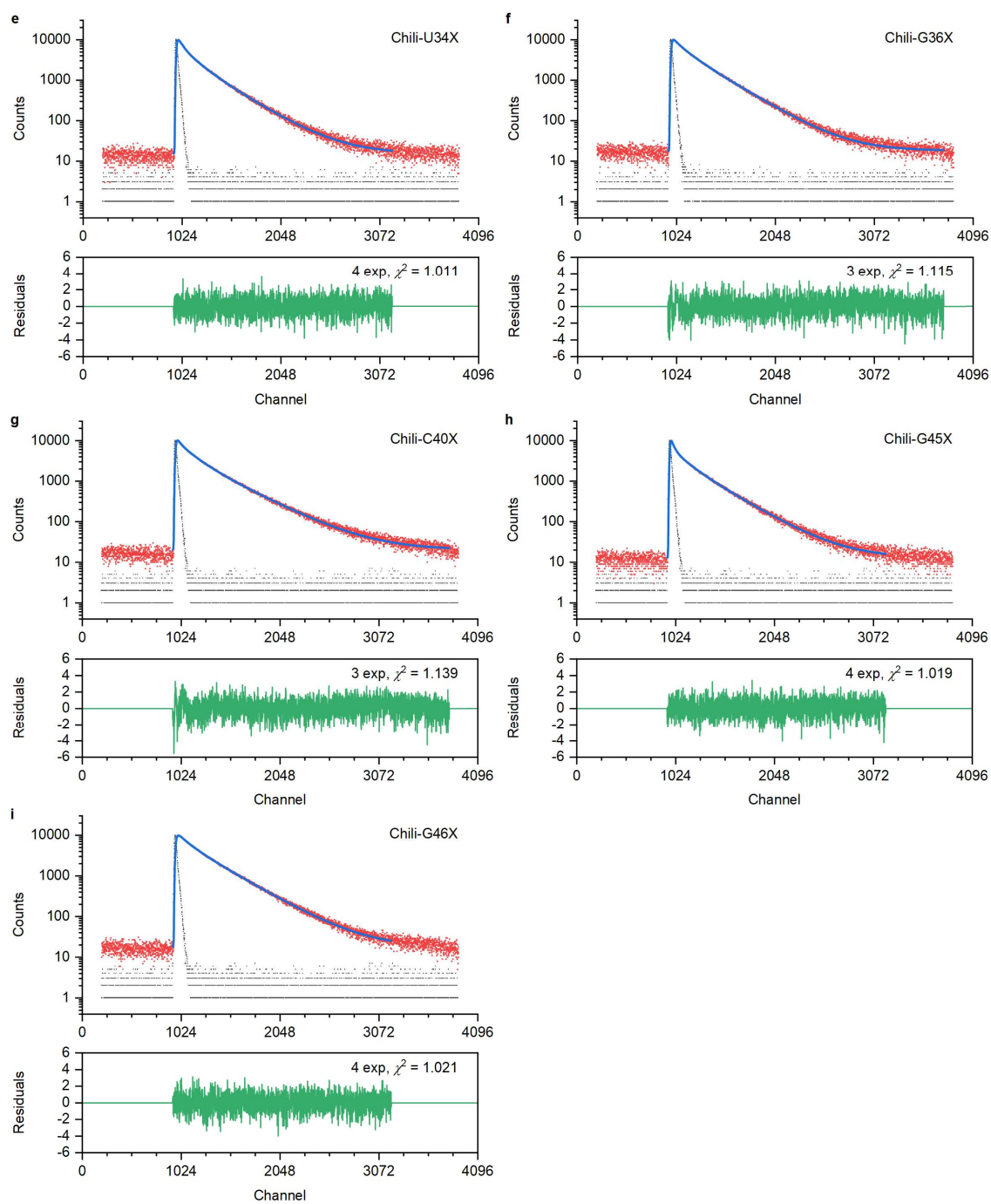


**Supplementary Figure 9.** Left: Thermal denaturation curves of native and 4Cl-modified 16 bp duplexes **R1–R4** and **D1** in PBS (pH 7.4). The change in absorbance at 260 nm was converted to percentage hyperchromicity to make facilitate comparison of the curves for different concentrations and fitted (solid lines) as described in section 4.1.18.1. Data is shown for the third of five ramps. Thermodynamic parameters reported in Supplementary Table 2 were obtained by direct fitting of the absorbance curves in order to also recover the actual sample concentrations, but differences to the fit results shown in this figure are negligible. Right: Van't Hoff plots using fitted  $T_m$  values averaged between two cooling/heating cycles. Open data points were excluded from the fits due to systematic deviations. As discussed in the main text, thermodynamic parameters obtained by this method were not analyzed further.

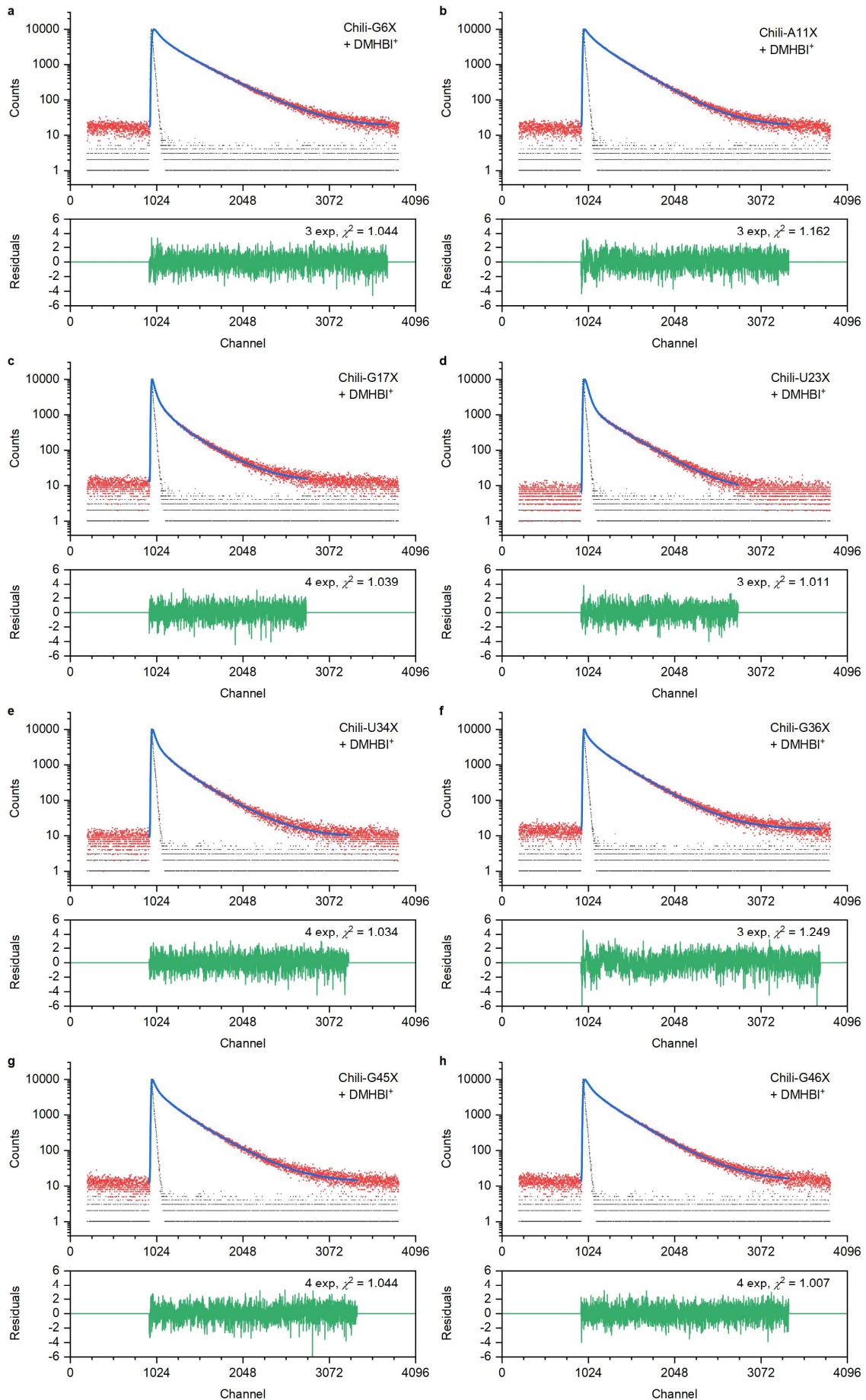


**Supplementary Figure 10.** Fluorescence lifetime measurement ( $\lambda_{\text{ex/em}} = 318/410$  nm) of various a) single-stranded **RXb** and b) double-stranded **R4**. Top panel: Sample (red), instrument response function (black), multiexponential fit (blue). Bottom panel: Weighted residuals and reduced  $\chi^2$  for the indicated fit. Conditions:  $1 \mu\text{M}$  RNA, PBS pH 7.4.

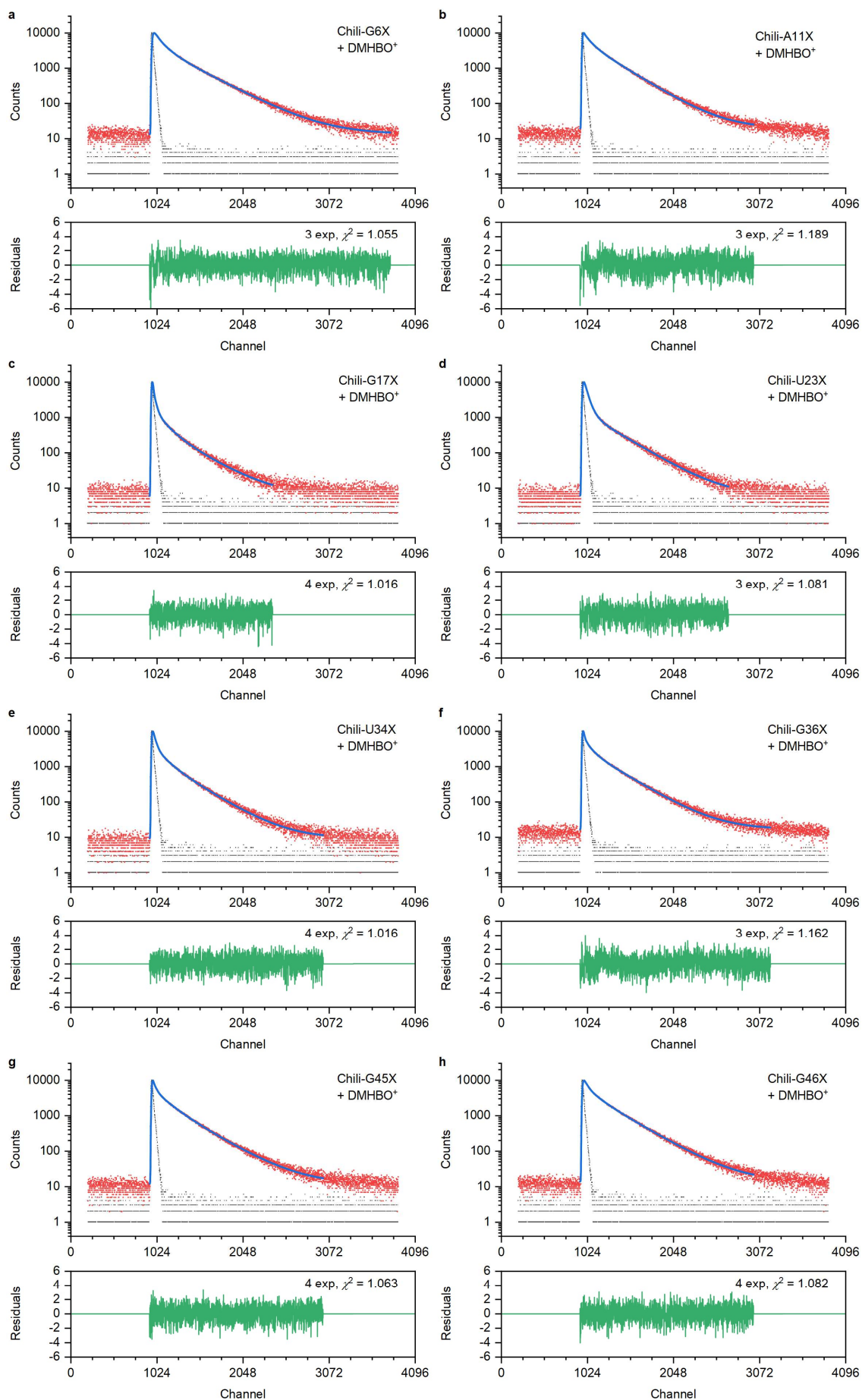




**Supplementary Figure 11.** Fluorescence lifetime measurement ( $\lambda_{ex/em} = 318/410$  nm) of various modified Chili aptamer constructs. Top panel: Sample (red), instrument response function (black), multiexponential fit (blue). Bottom panel: Weighted residuals and reduced  $\chi^2$  for the indicated fit. Conditions:  $0.5 \mu\text{M}$  RNA,  $125 \text{ mM}$  KCl,  $5 \text{ mM}$   $\text{MgCl}_2$ ,  $40 \text{ mM}$  HEPES pH 7.5.

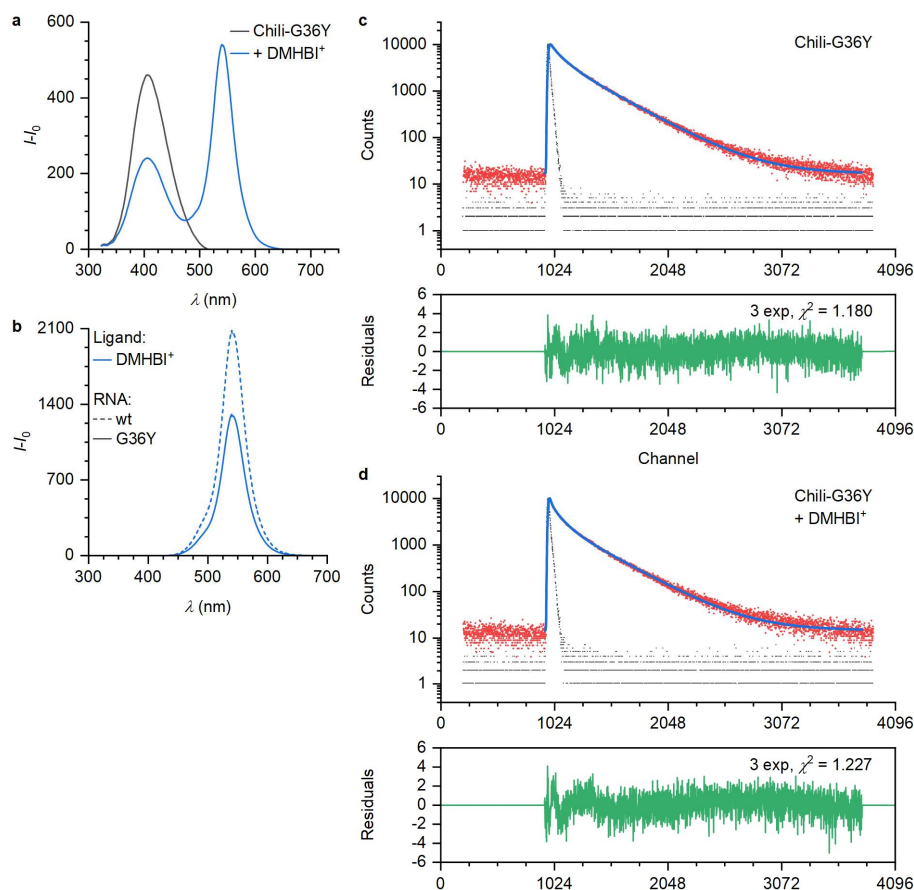


**Supplementary Figure 12.** Fluorescence lifetime measurement ( $\lambda_{\text{ex/em}} = 318/410 \text{ nm}$ ) of various modified Chili aptamer constructs in complex with DMHBI<sup>+</sup>. Top panel: Sample (red), instrument response function (black), multiexponential fit (blue). Bottom panel: Weighted residuals and reduced  $\chi^2$  for the indicated fit. Conditions: 0.5  $\mu\text{M}$  RNA, 1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5.





**Supplementary Figure 13.** Fluorescence lifetime measurement ( $\lambda_{\text{ex/em}} = 318/410$  nm) of various modified Chili aptamer constructs in complex with DMHBO<sup>+</sup>. Top panel: Sample (red), instrument response function (black), multiexponential fit (blue). Bottom panel: Weighted residuals and reduced  $\chi^2$  for the indicated fit. Conditions: 0.5  $\mu\text{M}$  RNA, 1  $\mu\text{M}$  ligand, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5.



**Supplementary Figure 14.** a) Fluorescence emission spectra ( $\lambda_{\text{ex}} = 303$  nm) of Chili-G36Y (**L10**) alone and in complex with DMHBI<sup>+</sup>. b) Fluorescence emission spectra ( $\lambda_{\text{ex}} = 413$  nm). c, d) Fluorescence lifetime measurement ( $\lambda_{\text{ex/em}} = 318/410$  nm). Top panel: Sample (red), instrument response function (black), multiexponential fit (blue). Bottom panel: Weighted residuals and reduced  $\chi^2$  for the indicated fit. Conditions: 0.5  $\mu\text{M}$  RNA, 125 mM KCl, 5 mM MgCl<sub>2</sub>, 40 mM HEPES pH 7.5, +1  $\mu\text{M}$  ligand.

## 5.2 SUPPLEMENTARY TABLES

**Supplementary Table 1.**  $pK_a$  values for the protonation equilibria of various HBI derivatives. Values were obtained by fitting of the titration curves shown in Figure 13 and Supplementary Figure 8 with the Henderson-Hasselbalch equation.

Ligand		$pK_a$ (NH <sup>+</sup> /N)	$pK_a$ (OH/O <sup>-</sup> )	$pK_a$ (other)
DMHBI	<b>13</b>	$2.34 \pm 0.03$	$7.96 \pm 0.01$	—
DMHBI <sup>+</sup>	<b>72</b>	$0.55 \pm 0.01$	$7.63 \pm 0.02$	—
DMHBI-Styr	<b>81</b>	$1.87 \pm 0.01$	$7.71 \pm 0.01$	—
DMHBI-Imi	<b>85</b>	$0.70 \pm 0.05$	$7.87 \pm 0.02$	$4.79 \pm 0.06^{[a]}$
DMHBI-Styr <sup>+</sup>	<b>95</b>	$0.09 \pm 0.12$	$7.34 \pm 0.02$	—
DMHBO <sup>+</sup>	<b>96</b>	—	$6.89 \pm 0.02$	$9.23 \pm 0.01^{[b]}$

[a] Imidazole, NH<sup>+</sup>/N.<sup>293</sup>

[b] Oxime, NOH/NO<sup>-</sup>.

**Supplementary Table 2.** Thermodynamic data of 16 bp RNA, DNA and hybrid duplexes<sup>[a]</sup> in PBS (pH 7.4) obtained from direct fitting of the thermal denaturation curves.

	$c^{[b]}$ $\mu\text{mol L}^{-1}$	$T_m^{[c]}$ $^{\circ}\text{C}$	$\Delta T_m^{[d]}$ $^{\circ}\text{C}$	$\Delta H^0$ $\text{kJ mol}^{-1}$	$\Delta S^0$ $\text{J mol}^{-1} \text{K}^{-1}$	$\Delta G_{298 \text{ K}}^0$ $\text{kJ mol}^{-1}$
<b>R1<sub>AU</sub></b>	1	79.3 ± 0.1				
	2	80.1 ± 0.1				
	5	81.4 ± 0.1		-717	-1919	-145
	10	81.2 ± 0.4				
	20	82.8 ± 0.1				
<b>R1<sub>AX</sub></b>	1	69.0 ± 0.2	-10.3			
	2	70.1 ± 0.1	-10.0			
	5	73.8 ± 0.1	-7.6	-580	-1575	-111
	10	74.3 ± 0.2	-6.9			
	20	72.9 ± 0.4	-9.9			
<b>R2<sub>CG</sub></b>	1	83.8 ± 0.1				
	2	84.5 ± 0.1				
	5	85.9 ± 0.1		-717	-1892	-153
	10	85.7 ± 0.4				
	20	87.8 ± 0.2				
<b>R2<sub>CX</sub></b>	1	67.6 ± 0.2	-16.2			
	2	72.6 ± 0.2	-11.9			
	5	74.8 ± 0.3	-11.1	-557	-1505	-109
	10	71.8 ± 0.2	-13.9			
	20	72.3 ± 0.3	-15.0			
<b>R3<sub>GC</sub></b>	1	84.0 ± 0.1				
	2	84.7 ± 0.2				
	5	86.0 ± 0.1		-707	-1863	-152
	10	85.6 ± 0.3				
	20	87.1 ± 0.1				
<b>R3<sub>GX</sub></b>	1	66.0 ± 0.1	-18.0			
	2	67.2 ± 0.2	-17.5			
	5	68.8 ± 0.1	-17.2	-455	-1214	-92
	10	69.4 ± 0.2	-16.2			
	20	69.6 ± 0.5	-17.5			
<b>R4<sub>UA</sub></b>	1	79.7 ± 0.1				
	2	80.2 ± 0.1				
	5	81.8 ± 0.1		-757	-2027	-153
	10	81.5 ± 0.5				
	20	83.4 ± 0.2				
<b>R4<sub>UX</sub></b>	1	70.6 ± 0.1	-9.1			
	2	71.5 ± 0.1	-8.7			
	5	73.0 ± 0.1	-8.8	-629	-1710	-119
	10	73.2 ± 0.4	-8.3			
	20	75.1 ± 0.1	-8.3			
<b>R4<sub>UY</sub></b>	1	68.7 ± 0.2	-11.0			
	2	69.3 ± 0.2	-10.9			
	5	71.4 ± 0.1	-10.4	-594	-1621	-110
	10	71.6 ± 0.3	-9.9			
	20	73.8 ± 0.3	-9.6			

Supplementary Table 2 cont.

	$c^{[b]}$ $\mu\text{mol L}^{-1}$	$T_m^{[b]}$ $^{\circ}\text{C}$	$\Delta T_m^{[c]}$ $^{\circ}\text{C}$	$\Delta H^0$ $\text{kJ mol}^{-1}$	$\Delta S^0$ $\text{J mol}^{-1} \text{K}^{-1}$	$\Delta G_{298\text{K}}^0$ $\text{kJ mol}^{-1}$
<b>D1<sub>TA</sub></b>	1	65.2 ± 0.2				
	2	65.8 ± 0.1				
	5	67.8 ± 0.1		-493	-1339	-94
	10	69.2 ± 0.2				
	20	69.9 ± 0.4				
<b>D1<sub>TY</sub></b>	1	53.0 ± 0.1	-12.2			
	2	53.3 ± 0.2	-12.5			
	5	55.6 ± 0.1	-12.2	-418	-1162	-71
	10	58.1 ± 0.2	-11.1			
	20	59.2 ± 0.3	-10.7			

[a] X = r4Cl, Y = d4Cl.

[b] Concentration of each individual strand.

[c] Values are given as mean ± s.d. for two cooling/heating cycles.

[d] Difference of means relative to the unmodified duplex.

## 5.3 CARTESIAN COORDINATES

## 5.3.1 HBI derivatives

5.3.1.1 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-isopropyl-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-*i*Pr, 62)

H	-0.448689	-2.367131	1.709123
C	-0.397167	-1.333419	2.008131
C	-1.532747	-0.755308	2.597711
C	-1.493672	0.592515	2.992410
H	-2.367475	1.035651	3.446754
C	-0.343127	1.334716	2.793171
O	-0.175083	2.654198	3.125841
C	0.790118	0.758359	2.209874
C	0.755938	-0.587742	1.816879
O	1.899068	-1.062418	1.261264
O	1.916006	1.485181	2.022422
C	-1.279767	3.348851	3.682209
C	1.925240	-2.426437	0.872802
H	1.738661	2.378410	2.350859
H	-0.944193	4.370220	3.844518
H	-1.582690	2.907788	4.635779
H	-2.130079	3.347777	2.994789
H	2.921441	-2.602199	0.473908
H	1.178292	-2.635805	0.101750
H	1.749903	-3.086167	1.727402
C	-3.061413	-2.769379	2.502569
C	-4.382369	-3.366097	2.821099
N	-2.276346	-3.728227	1.850250
N	-4.270266	-4.669167	2.314619
C	-3.003369	-4.798570	1.760301
O	-5.356871	-2.897925	3.379429
C	-2.750953	-1.492687	2.821450
H	-3.552286	-0.954607	3.318465
C	-5.322517	-5.687817	2.384953
C	-5.645060	-6.035061	3.837968
C	-2.541549	-6.058262	1.121351
H	-1.528290	-5.908580	0.757277
H	-3.183817	-6.337701	0.282380
H	-2.545883	-6.890074	1.830298
H	-4.907156	-6.574867	1.905597
H	-6.046619	-5.166090	4.356985
H	-4.749000	-6.370851	4.362609
C	-6.558132	-5.249808	1.598698
H	-7.303597	-6.046875	1.604981
H	-6.298892	-5.030110	0.561628
H	-6.996355	-4.357481	2.043024
H	-6.384500	-6.836954	3.873386

5.3.1.2 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(*trans*-4-methylcyclohexyl)-3,5-dihydro-4H-imidazol-4-one (DMHBI-MeCy, 64)

H	-0.434333	-2.339495	1.672485
C	-0.391253	-1.316625	2.007506
C	-1.536030	-0.765096	2.604718
C	-1.507654	0.568497	3.046132
H	-2.387935	0.991875	3.506776
C	-0.358575	1.322351	2.884515
O	-0.200287	2.630633	3.263347
C	0.784135	0.771514	2.295031
C	0.760587	-0.560130	1.855462
O	1.912596	-1.011369	1.298482
O	1.909315	1.508688	2.147098
C	-1.309415	3.297634	3.843661
C	1.944922	-2.357363	0.851988
H	1.724170	2.389371	2.503815
H	-0.978842	4.313231	4.047394
H	-1.612731	2.817562	4.778082
H	-2.158009	3.320295	3.154473
H	2.944971	-2.514166	0.454756
H	1.205087	-2.534368	0.066093
H	1.763475	-3.053730	1.675704
C	-3.048953	-2.784094	2.421502
C	-4.356442	-3.416783	2.728968

N	-2.255093	-3.701252	1.721754
N	-4.228841	-4.694795	2.164067
C	-2.964786	-4.779263	1.594881
O	-5.329527	-2.991486	3.322989
C	-2.749953	-1.519571	2.795709
H	-3.553689	-1.010894	3.319041
C	-5.237316	-5.750395	2.241919
C	-6.544607	-5.339923	1.558023
C	-5.475564	-6.187062	3.690704
C	-7.579228	-6.461540	1.646719
C	-7.837341	-6.903614	3.091304
C	-6.517611	-7.302273	3.760234
C	-8.864821	-8.029557	3.159372
C	-2.489399	-6.002228	0.897502
H	-1.486243	-5.816792	0.522027
H	-3.142033	-6.260705	0.059916
H	-2.464241	-6.861823	1.571996
H	-9.060656	-8.325601	4.192445
H	-8.508167	-8.912842	2.621659
H	-9.813952	-7.727891	2.710941
H	-6.924268	-4.438937	2.042125
H	-6.342303	-5.085708	0.514326
H	-7.234062	-7.329522	1.070370
H	-8.514554	-6.135852	1.184170
H	-8.235983	-6.038654	3.636041
H	-6.129227	-8.201756	3.265512
H	-6.695585	-7.576340	4.803406
H	-4.826770	-6.603566	1.696979
H	-5.812845	-5.321210	4.263371
H	-4.529591	-6.516469	4.128400

**5.3.1.3 (Z)-3-Benzyl-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBI-Bn, 65)**

H	-2.381237	-2.503561	-1.500380
C	-1.396059	-2.073025	-1.431822
C	-0.708441	-2.176356	-0.211872
C	0.579251	-1.625997	-0.097511
H	1.106289	-1.706798	0.841610
C	1.155911	-0.991726	-1.183273
O	2.401476	-0.419721	-1.205336
C	0.470645	-0.884945	-2.398193
C	-0.815529	-1.433227	-2.516440
O	-1.403430	-1.283623	-3.729446
O	1.034881	-0.259854	-3.457333
C	3.172540	-0.452503	-0.014662
C	-2.700783	-1.830783	-3.904431
H	1.910083	0.045411	-3.178041
H	4.106720	0.053856	-0.244634
H	3.380185	-1.481666	0.290692
H	2.662395	0.072019	0.797815
H	-2.980256	-1.608437	-4.931481
H	-3.421288	-1.373720	-3.220292
H	-2.699369	-2.913473	-3.749484
C	-2.460949	-3.446493	1.102785
C	-2.861355	-4.073788	2.386898
N	-3.491648	-3.625361	0.168749
N	-4.137324	-4.578241	2.099811
C	-4.430168	-4.283262	0.774213
O	-2.288575	-4.176610	3.453696
C	-1.267987	-2.830226	0.944307
H	-0.646179	-2.840302	1.834216
C	-4.911062	-5.374193	3.028385
C	-5.714259	-4.695203	0.155610
H	-5.732413	-4.360593	-0.878462
H	-6.562388	-4.257475	0.689123
H	-5.832963	-5.780395	0.190674
H	-4.421279	-5.244482	3.996479
H	-5.921352	-4.969286	3.114075
C	-4.970660	-6.842223	2.666822
C	-6.133467	-7.573090	2.895520
C	-3.862212	-7.489471	2.123837
C	-6.190469	-8.928703	2.591548
C	-3.918207	-8.842241	1.814631
C	-5.082313	-9.566829	2.047907

H	-7.004012	-7.077656	3.310215
H	-2.952052	-6.932734	1.941311
H	-7.102398	-9.483358	2.772762
H	-3.049999	-9.332125	1.392357
H	-5.125071	-10.620781	1.805064

**5.3.1.4 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(4-methylphenyl)-3,5-dihydro-4H-imidazol-4-one (DMHBTI, 44)**

H	-0.506210	-2.330653	1.784354
C	-0.427988	-1.287219	2.039794
C	-1.563205	-0.641569	2.555447
C	-1.486123	0.719142	2.896729
H	-2.360153	1.214545	3.293476
C	-0.298298	1.406347	2.722359
O	-0.090092	2.729069	3.015454
C	0.832662	0.764404	2.207046
C	0.759839	-0.594095	1.864918
O	1.903173	-1.133851	1.373039
O	1.993449	1.438631	2.036453
C	-1.174262	3.475652	3.545508
C	1.880237	-2.505620	1.011700
H	1.843894	2.347840	2.333601
H	-0.798692	4.483712	3.703102
H	-1.510778	3.058069	4.497987
H	-2.012769	3.503505	2.843890
H	2.880070	-2.733722	0.649877
H	1.150201	-2.695734	0.219577
H	1.646507	-3.139897	1.871412
C	-3.158092	-2.606840	2.521653
C	-4.520957	-3.126126	2.793726
N	-2.364303	-3.644612	2.017755
N	-4.407031	-4.491241	2.442049
C	-3.106850	-4.704871	1.985499
O	-5.520871	-2.582874	3.208824
C	-2.820957	-1.317788	2.752414
H	-3.631677	-0.713496	3.147771
C	-5.461539	-5.434883	2.527828
C	-6.663445	-5.188652	1.871183
C	-5.319429	-6.589631	3.289522
C	-7.701209	-6.103151	1.965494
C	-7.566763	-7.280997	2.702155
C	-6.359884	-7.505808	3.362765
H	-4.402824	-6.760678	3.837839
H	-6.235778	-8.402752	3.957989
H	-6.782291	-4.275456	1.305288
H	-8.635201	-5.897663	1.456168
C	-8.687037	-8.282451	2.767568
C	-2.655528	-6.027123	1.481922
H	-1.694692	-5.900112	0.989183
H	-3.380254	-6.446875	0.781444
H	-2.541421	-6.745418	2.296608
H	-8.613256	-8.903524	3.661067
H	-8.667046	-8.949552	1.900394
H	-9.659880	-7.788863	2.775123

**5.3.1.5 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-methoxyphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBAl, 68)**

H	-0.558660	-2.340095	1.766989
C	-0.458043	-1.297350	2.018263
C	-1.559475	-0.642405	2.592387
C	-1.454181	0.718065	2.927699
H	-2.301685	1.220339	3.370276
C	-0.272329	1.396390	2.688321
O	-0.038095	2.718395	2.966265
C	0.825384	0.744727	2.116433
C	0.724577	-0.613689	1.780994
O	1.836554	-1.162792	1.231236
O	1.979906	1.410809	1.884329
C	-1.095131	3.481892	3.525035
C	1.797870	-2.543177	0.906016
H	1.849603	2.322910	2.181507
H	-0.707398	4.490193	3.648222
H	-1.395999	3.084416	4.498221

H	-1.960575	3.501882	2.857099
H	2.780884	-2.781338	0.506836
H	1.034097	-2.753889	0.152006
H	1.602511	-3.154301	1.791756
C	-3.171119	-2.593532	2.634029
C	-4.526409	-3.099396	2.965386
N	-2.414215	-3.636273	2.084812
N	-4.447395	-4.460256	2.592826
C	-3.173318	-4.685501	2.074315
O	-5.499209	-2.547474	3.430928
C	-2.810909	-1.308881	2.852971
H	-3.595411	-0.698989	3.290067
C	-5.512343	-5.390854	2.704020
C	-6.708359	-5.156085	2.041302
C	-5.383374	-6.528976	3.499524
C	-7.765787	-6.052232	2.148677
C	-7.622599	-7.203735	2.921892
C	-6.422771	-7.434864	3.599270
H	-4.466831	-6.696294	4.049411
H	-6.332485	-8.323807	4.208796
H	-6.818774	-4.260072	1.446625
H	-8.687473	-5.842532	1.627360
O	-8.587801	-8.147685	3.080960
C	-9.816767	-7.974949	2.394441
H	-9.664628	-7.937840	1.311671
H	-10.330658	-7.065558	2.719511
H	-10.424170	-8.841325	2.644611
C	-2.767040	-6.010067	1.540712
H	-1.806852	-5.904509	1.041803
H	-3.510062	-6.390654	0.836934
H	-2.672653	-6.749509	2.338714

**5.3.1.6 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-trifluoromethylphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBTI<sup>F</sup>, 69)**

H	-0.473878	-2.336531	1.761483
C	-0.415882	-1.293484	2.024396
C	-1.598361	-0.631901	2.395715
C	-1.547387	0.728911	2.744095
H	-2.457482	1.235746	3.028884
C	-0.337227	1.398532	2.730007
O	-0.147249	2.714610	3.058950
C	0.841760	0.739164	2.364623
C	0.794149	-0.616787	2.005424
O	1.984631	-1.167984	1.663287
O	2.025020	1.392116	2.358037
C	-1.274031	3.470120	3.476345
C	1.999488	-2.542644	1.312082
H	1.853061	2.303427	2.636569
H	-0.902525	4.465853	3.705371
H	-1.729174	3.034707	4.370020
H	-2.021327	3.532959	2.680688
H	3.036757	-2.780126	1.089048
H	1.380239	-2.734489	0.431236
H	1.649038	-3.167301	2.138606
C	-3.197159	-2.589523	2.239639
C	-4.572069	-3.100588	2.433558
N	-2.363929	-3.646298	1.855364
N	-4.422655	-4.492651	2.190806
C	-3.086900	-4.718721	1.842599
O	-5.604829	-2.545059	2.734828
C	-2.875648	-1.293745	2.460610
H	-3.718242	-0.675660	2.754536
C	-5.455240	-5.438008	2.371220
C	-6.711028	-5.201663	1.814409
C	-5.244344	-6.578933	3.143823
C	-7.742217	-6.103281	2.026578
C	-7.524820	-7.249853	2.784138
C	-6.272277	-7.486810	3.340401
H	-4.284321	-6.744250	3.610664
H	-6.103347	-8.371588	3.938132
H	-6.876639	-4.306876	1.233878
H	-8.715968	-5.917355	1.595635
C	-8.656453	-8.196648	3.066503
C	-2.580746	-6.056273	1.438629



H	-1.630045	-5.921829	0.928146
H	-3.286230	-6.562790	0.778128
H	-2.418760	-6.704292	2.302311
F	-8.228972	-9.468245	3.218005
F	-9.580010	-8.198998	2.084185
F	-9.307027	-7.870848	4.208706

**5.3.1.7 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-3-(4-trifluoromethoxyphenyl)-2-methyl-3,5-dihydro-4H-imidazol-4-one (DMHBAI<sup>F</sup>, 70)**

H	-0.640484	-2.087243	2.645256
C	-0.652919	-1.014190	2.554148
C	-1.896176	-0.368775	2.452960
C	-1.938233	1.030368	2.328339
H	-2.894636	1.525501	2.248740
C	-0.760036	1.754730	2.304788
O	-0.661823	3.115559	2.181952
C	0.479091	1.112484	2.408553
C	0.525326	-0.284445	2.533592
O	1.767190	-0.821156	2.624269
O	1.629735	1.821891	2.387057
C	-1.860079	3.862820	2.040311
C	1.868478	-2.231980	2.738376
H	1.391291	2.755295	2.290142
H	-1.558534	4.902794	1.943307
H	-2.500862	3.748613	2.918764
H	-2.410067	3.556547	1.146379
H	2.932574	-2.449622	2.787749
H	1.430264	-2.732957	1.870481
H	1.376752	-2.593682	3.645946
C	-3.381713	-2.414228	2.544703
C	-4.750836	-2.977797	2.529106
N	-2.455642	-3.460718	2.630158
N	-4.513452	-4.371069	2.642009
C	-3.129857	-4.563159	2.682619
O	-5.839475	-2.454208	2.441941
C	-3.146325	-1.084213	2.464694
H	-4.044948	-0.479419	2.394261
C	-5.528222	-5.355702	2.647366
C	-6.494095	-5.350072	1.643008
C	-5.579493	-6.316733	3.653922
C	-7.488993	-6.316056	1.627823
C	-7.509520	-7.278984	2.624452
C	-6.566455	-7.291829	3.637490
H	-4.856205	-6.295392	4.456642
H	-6.616624	-8.050443	4.405911
H	-6.462904	-4.588057	0.878416
H	-8.236062	-6.329015	0.847522
O	-8.450529	-8.319657	2.584392
C	-9.709615	-8.055603	2.994667
C	-2.514211	-5.913624	2.738293
H	-1.459427	-5.819676	2.491765
H	-2.998394	-6.597807	2.039499
H	-2.599730	-6.351902	3.735023
F	-9.752463	-7.549339	4.239609
F	-10.390250	-9.199379	2.978058
F	-10.343146	-7.176733	2.194054

**5.3.1.8 (Z)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3-(4-(trimethylammonium)phenyl)-3,5-dihydro-4H-imidazol-4-one iodide (DMHBI<sup>+</sup>, 72)**

H	-0.13268	-0.90426	-10.07804
C	-0.33735	-0.12469	-10.79201
C	-0.69361	1.14606	-10.30983
C	-0.96319	2.18119	-11.2226
H	-1.23771	3.1569	-10.85
C	-0.86658	1.94175	-12.58103
O	-1.09102	2.85776	-13.57263
C	-0.51251	0.6749	-13.06183
C	-0.24942	-0.36325	-12.15378
O	0.08275	-1.55316	-12.70958
O	-0.42173	0.44008	-14.3875
C	-1.42991	4.18427	-13.19653
C	0.35399	-2.6372	-11.83464
H	-0.63101	1.26784	-14.84469

H	-1.54711	4.74017	-14.12321
H	-0.63558	4.63615	-12.59655
H	-2.36837	4.20574	-12.63609
H	0.59184	-3.48269	-12.47526
H	-0.51707	-2.87643	-11.21806
H	1.2049	-2.41785	-11.18338
C	-0.53216	0.65826	-7.83302
C	-0.68407	1.14125	-6.44606
N	-0.0892	-0.66977	-7.81219
N	-0.25599	0.02104	-5.68
C	0.06606	-1.00691	-6.57541
O	-1.0636	2.19687	-5.98911
C	-0.79355	1.44264	-8.90628
H	-1.11833	2.44655	-8.65121
C	-0.20726	1.46E-4	-4.27378
C	-1.30923	0.42093	-3.53742
C	0.94575	-0.40321	-3.60145
C	-1.27206	0.41396	-2.14899
C	-0.13264	-0.0187	-1.48721
C	0.97858	-0.42544	-2.21814
H	1.8261	-0.68561	-4.16026
H	1.88321	-0.7484	-1.72478
H	-2.19551	0.76243	-4.05044
H	-2.14512	0.74384	-1.61058
N	-0.06381	-0.06775	0.00112
C	-1.31294	0.43688	0.6725
H	-2.15093	-0.18676	0.3749
H	-1.47523	1.47155	0.38384
H	-1.14278	0.35854	1.7507
C	0.14174	-1.49602	0.45789
H	-0.69229	-2.08614	0.08663
H	0.17727	-1.48693	1.55228
H	1.07572	-1.86482	0.04663
C	0.48918	-2.35802	-6.1253
H	0.42464	-3.03576	-6.97296
H	-0.14529	-2.72319	-5.31551
H	1.51903	-2.35913	-5.76168
C	1.08201	0.78341	0.50333
H	0.92364	1.79819	0.14738
H	2.01381	0.38862	0.11277
H	1.07246	0.72862	1.59636
I	0.25814	-0.31276	3.99704

**5.3.1.9 (Z)-3-(4-*tert*-Butylphenyl)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-methyl-3,5-dihydro-4*H*-imidazol-4-one (DMHBI<sup>c</sup>, 73)**

H	-0.565099	-2.329345	1.741030
C	-0.460734	-1.292754	2.014554
C	-1.545866	-0.657948	2.640223
C	-1.436087	0.695281	3.002754
H	-2.270499	1.182889	3.484847
C	-0.267224	1.386270	2.736521
O	-0.031745	2.704148	3.031896
C	0.814805	0.753625	2.115105
C	0.710282	-0.597844	1.754401
O	1.806406	-1.129425	1.157418
O	1.957399	1.431310	1.859260
C	-1.076056	3.451832	3.634197
C	1.762686	-2.502953	0.804850
H	1.831641	2.336790	2.177869
H	-0.691518	4.461050	3.759806
H	-1.343678	3.038454	4.610432
H	-1.961951	3.474989	2.993727
H	2.732912	-2.728228	0.368628
H	0.975241	-2.701593	0.072251
H	1.599215	-3.133296	1.683598
C	-3.150039	-2.613658	2.660684
C	-4.487147	-3.146563	3.021855
N	-2.414560	-3.618073	2.019920
N	-4.419482	-4.484437	2.567094
C	-3.167830	-4.670101	1.980575
O	-5.438700	-2.629084	3.563850
C	-2.784023	-1.339222	2.926245
H	-3.549576	-0.752887	3.424773
C	-5.474151	-5.427988	2.656574

C	-6.717239	-5.136305	2.112144
C	-5.282820	-6.647894	3.299379
C	-7.745706	-6.068586	2.184598
C	-7.565677	-7.311657	2.790739
C	-6.311356	-7.574182	3.351458
H	-4.331210	-6.865930	3.765377
H	-6.130756	-8.516697	3.851011
H	-6.879914	-4.178617	1.637968
H	-8.699216	-5.808764	1.749363
C	-8.670554	-8.368331	2.869315
C	-9.955544	-7.922548	2.160961
H	-9.785852	-7.730322	1.099902
H	-10.375301	-7.021040	2.610595
H	-10.706575	-8.710371	2.239308
C	-8.182761	-9.668878	2.202264
H	-7.928996	-9.495827	1.154711
H	-8.965534	-10.429549	2.242879
H	-7.299960	-10.072490	2.698853
C	-2.777086	-5.952392	1.342037
H	-1.852784	-5.795838	0.791337
H	-3.557239	-6.304269	0.664321
H	-2.618000	-6.737150	2.084155
C	-9.003658	-8.645505	4.347386
H	-9.355212	-7.737974	4.841384
H	-8.133007	-9.009043	4.894311
H	-9.787714	-9.402406	4.422506

### 5.3.2 Methylated nucleobases

#### 5.3.2.1 m<sup>a</sup>A

N	-3.7018836213	-0.5144343935	0.3664408933
C	-2.3886410613	-0.2422802258	0.0634284355
C	-2.2935773031	1.1497825651	0.0619038894
N	-3.5133594209	1.7268089964	0.3589011498
C	-4.3157919413	0.7080213809	0.5310100162
C	-1.0247536232	1.6818192466	-0.2237594962
N	-0.0138942066	0.8361690599	-0.4628517572
C	-0.2516519792	-0.4806327586	-0.4270687174
N	-1.3996743988	-1.1060882099	-0.1752189366
H	0.6046053825	-1.1150295805	-0.6289697725
N	-0.7951777509	3.0116869065	-0.2806899695
C	-4.2955055474	-1.8311361956	0.4757531671
H	-5.3641725465	0.7741257587	0.7797807740
H	-1.5146269329	3.6516540539	0.0043235753
H	0.1516020730	3.3321894200	-0.3814162507
H	-5.3446113213	-1.7225036881	0.7439415724
H	-3.7844846882	-2.4138985393	1.2418184519
H	-4.2183611124	-2.3620437969	-0.4731370250

#### 5.3.2.2 m<sup>1</sup>C

C	-0.7409710867	0.8810612678	0.0627360277
C	0.5094680578	1.4030148632	0.0858989443
N	1.6024914165	0.6169216166	0.0078046875
C	1.4841732422	-0.8062602036	-0.1021953866
N	0.2232920367	-1.3242564718	-0.1171063428
C	-0.8290251948	-0.5419451198	-0.0432712345
O	2.5082273480	-1.4593272964	-0.1700944581
N	-2.0491872193	-1.1352585759	-0.0921882680
H	-2.0757037978	-2.1402403457	-0.0574733100
H	-2.8814193867	-0.6206529824	0.1283626794
H	0.6874936588	2.4673948406	0.1667024946
H	-1.6131724693	1.5140878282	0.1169690396
C	2.9503422122	1.1671573670	0.0307372092
H	3.5074303366	0.7577997768	0.8724539046
H	2.8877567903	2.2504513106	0.1175587843
H	3.4802210554	0.8971791247	-0.8817297712

#### 5.3.2.3 m<sup>9</sup>G

N	-1.9227216919	1.5315504324	-0.7474006332
C	-1.0857156477	2.5276825427	-0.8345003217
N	0.2088539784	2.1795962402	-0.4931063561
C	0.1693818995	0.8530625240	-0.1662200219
C	-1.1575768922	0.4647363399	-0.3290538856
N	1.2316806649	0.1060795834	0.2197845966
C	0.9148288381	-1.1333061116	0.4819580521

C	-1.5188081723	-0.8940547522	-0.0459913806
N	-0.3531195641	-1.6248188738	0.3626492800
O	-2.5869876059	-1.4649764741	-0.0926763091
N	1.8745839358	-1.9959294030	0.9486547787
H	-1.3311266841	3.5340628742	-1.1367559676
C	1.3810954177	3.0297980561	-0.4709782539
H	-0.5511268527	-2.5777416297	0.6327283588
H	2.8129928968	-1.6384051604	0.8656289071
H	1.7753384697	-2.9744358789	0.7325560281
H	1.0928804373	4.0304197039	-0.7866333870
H	2.1438639337	2.6450838417	-1.1479462069
H	1.7994626390	3.0767861452	0.5345927222

**5.3.2.4 m<sup>1</sup>U**

C	-1.1977325788	0.6781828248	-0.3097902338
C	-1.0744953459	-0.6528632279	-0.1556532143
N	0.1128433620	-1.2716409498	0.1367760437
C	1.2886001198	-0.5413685213	0.2941052325
N	1.1362540235	0.8187200453	0.1319303143
C	-0.0356009335	1.5338038663	-0.1658708464
O	2.3500663812	-1.0693284535	0.5519401347
O	-0.0078203954	2.7427901753	-0.2772310589
H	-1.9215969526	-1.3175828786	-0.2559854337
C	0.2117061460	-2.7187950512	0.2979325217
H	1.9797442728	1.3643431104	0.2432685308
H	-2.1464927789	1.1349257412	-0.5396176143
H	-0.7747350416	-3.1510944643	0.1494202947
H	0.5735186565	-2.9640125475	1.2954423796
H	0.9081010650	-3.1332296691	-0.4296570506

**5.3.2.5 m<sup>1</sup>4Cl**

N	-2.1687120000	-0.0573310000	0.0195840000
C	-0.9305660000	0.5432060000	0.0221820000
C	0.0448150000	-0.4894370000	0.0363460000
C	-0.6573820000	-1.7308650000	0.0421370000
C	-1.9912120000	-1.4246350000	0.0316180000
C	1.4049940000	-0.1267640000	0.0416030000
C	1.7482080000	1.2236400000	0.0325200000
C	0.7624940000	2.2153120000	0.0186120000
C	-0.5867750000	1.8914690000	0.0133500000
H	1.0617350000	3.2546920000	0.0117490000
C	2.4135720000	-1.1347640000	0.0561350000
C	-3.4348390000	0.6395800000	0.0058950000
H	-2.8478130000	-2.0786330000	0.0316280000
H	-0.2274480000	-2.7177200000	0.0526400000
H	-1.3413740000	2.6669570000	0.0027280000
H	2.7929580000	1.5016290000	0.0365120000
N	3.2172200000	-1.9618500000	0.0684590000
H	-4.2409990000	-0.0906210000	0.0074030000
H	-3.5282940000	1.2602170000	-0.8879700000
H	-3.5389700000	1.2756490000	0.8876470000

## 5.4 BUFFER SOLUTIONS AND REAGENTS

### 1 M TRIS pH 8.0

TRIS (60.57 g) was dissolved in H<sub>2</sub>O (350 mL) and the pH was adjusted to 8.0 by addition of conc. HCl. Afterwards, H<sub>2</sub>O was added for a final volume of 500 mL.

### 0.25 M TRIS pH 8.0

TRIS (30.29 g) was dissolved in H<sub>2</sub>O (700 mL) and the pH was adjusted to 8.0 by addition of conc. HCl. Afterwards, H<sub>2</sub>O was added for a final volume of 1000 mL.

### 0.5 M EDTA pH 8.0

Na<sub>2</sub>EDTA (93.06 g) were dissolved in H<sub>2</sub>O (350 mL) and the pH was adjusted to 8.0 by addition of 5 M NaOH. Afterwards, H<sub>2</sub>O was added for a final volume of 500 mL.

### 1 M NaCl

NaCl (29.22 g) was dissolved in H<sub>2</sub>O (500 mL).

### 3 M NaCl

NaCl (87.66 g) was dissolved in H<sub>2</sub>O (500 mL).

### 10x TBE buffer

TRIS (108 g) and boric acid (55 g) were mixed with 0.5 M EDTA pH 8.0 (40 mL) and the mixture was topped up to 900 mL with H<sub>2</sub>O. After everything had dissolved, H<sub>2</sub>O was added for a final volume of 1000 mL

### 1x TEN buffer

To a mixture of 1 M TRIS pH 8.0 (5 mL), 0.5 M EDTA pH 8.0 (1 mL) and 3 M NaCl (50 mL), H<sub>2</sub>O was added for a final volume of 500 mL.

### 10/15/20% gel stock

Urea (420 g) was suspended in 10x TBE (100 mL). 40% Rotiphorese gel acrylamide (250/375/500 mL) was added and the mixture was topped up to 900 mL with H<sub>2</sub>O. After everything had dissolved, H<sub>2</sub>O was added for a final volume of 1000 mL.

### 25% APS

Ammonium peroxodisulfate (3.75 g) was dissolved in H<sub>2</sub>O (15 mL).

### Loading buffer

Formamide (80 mL) was mixed with 10x TBE (10 mL) and 0.5 M EDTA pH 8.0 (10 mL). Afterwards, bromophenol blue (25 mg) and xylene cyanol (25 mg) were added.

### Dionex buffer A

0.25 M TRIS pH 8.0 (100 mL) were diluted with H<sub>2</sub>O (400 mL). Urea (360 g) was added in small batches. After everything had dissolved, H<sub>2</sub>O was added for a final volume of 1000 mL and the solution was passed through a 0.2 µm cellulose acetate filter.

### Dionex buffer B

0.25 M TRIS pH 8.0 (100 mL) were diluted with H<sub>2</sub>O (400 mL). Urea (360 g) was added in small batches followed by NaClO<sub>4</sub> · H<sub>2</sub>O (70.23 g). After everything had dissolved, H<sub>2</sub>O was added for a final volume of 1000 mL and the solution was passed through a 0.2 μm cellulose acetate filter.

0.2 M NaH<sub>2</sub>PO<sub>4</sub>

0.2 M Na<sub>2</sub>HPO<sub>4</sub>

5x PBS buffer pH 7.4

0.2 M NaH<sub>2</sub>PO<sub>4</sub> (14.0 mL), 0.2 M Na<sub>2</sub>HPO<sub>4</sub> (48.5 mL) and 1 M NaCl (187.5 mL) were mixed. The ratio of H<sub>2</sub>PO<sub>4</sub><sup>-</sup> to HPO<sub>4</sub><sup>2-</sup> was calculated according to the Henderson-Hasselbalch equation and the pH of the buffer solution was checked by pH meter.

Britton-Robinson buffer<sup>294</sup>

AcOH (2.30 mL), H<sub>3</sub>PO<sub>4</sub> (2.72 mL) and H<sub>3</sub>BO<sub>3</sub> (2.48 g) were dissolved in H<sub>2</sub>O for a final volume of 1000 mL. Aliquots (40 mL) of this stock solution were titrated with 1 M HCl or NaOH to make 20 separate buffers with a pH between 1 and 12, which were analyzed by pH meter.

## 5.5 ABBREVIATIONS

μIVC	Microfluidic-assisted in vitro compartmentalization
2AP	2-Aminopurine
4CI	4-Cyanoindole
AMP	Adenosine monophosphate
APS	Ammonium peroxydisulfate
ATP	Adenosine triphosphate
BOP	(Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
BTT	5-Benzylthio-1 <i>H</i> -tetrazole
C153	Coumarin 153
CALI	Chromophore-assisted laser inactivation
CASSCF	Complete active space self consistent field
CEPCI	2-Cyanoethyl <i>N,N</i> -diisopropylchlorophosphoramidite
CMP	Cytidine monophosphate
COSY	Correlation spectroscopy
CPG	Controlled pore glass
CRISPR	Clustered regularly interspaced short palindromic repeats
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
d4CI	4-Cyanoindole deoxyribonucleoside
DBU	1,8-Diazabicyclo(5.4.0)undec-7-ene
DCA	Dichloroacetic acid
DCM	Dichloromethane
DFT	Density functional theory
DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulfoxide
DMTCl	4,4'-Dimethoxytrityl chloride
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EC <sub>50</sub>	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced GFP
ESI	Electrospray ionization
ESIPT	Excited state intramolecular proton transfer
ETT	5-Ethylthio-1 <i>H</i> -tetrazole
FAST	Fluorescence-activating and absorption-shifting tag
FISH	Fluorescence in situ hybridization
FIT	Forced intercalation
FITC	Fluorescein isothiocyanate
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
GMP	Guanosine triphosphate
GTP	Guanosine triphosphate
HBI	Hydroxybenzylidene imidazolone
HEPES	2-(4-(2-Hydroxyethyl)piperazin-1-yl)ethanesulfonic acid
HFIP	Hexafluoroisopropanol
HMBC	Heteronuclear multiple bond correlation
HMPT	Hexamethylphosphoric triamide
HPLC	High-performance liquid chromatography
IC	Internal conversion

ISC .....	Intersystem crossing
LNA .....	Locked nucleic acid
lncRNA .....	Long noncoding RNA
MBS .....	MS2 binding site
MCP .....	MS2 coat protein
<i>m</i> CPBA .....	<i>meta</i> -Chloroperoxybenzoic acid
mRNA .....	Messenger RNA
MS .....	Mass spectrometry
MWCO .....	Molecular weight cut-off
NBOMCl .....	2-Nitrobenzyloxymethyl chloride
ncRNA .....	Noncoding RNA
NHS .....	<i>N</i> -Hydroxysuccinimide
NMO .....	<i>N</i> -Methylmorpholine- <i>N</i> -oxide
NMP .....	Nucleoside monophosphate
NMR .....	Nuclear magnetic resonance
NOESY .....	Nuclear Overhauser enhancement spectroscopy
NTP .....	Nucleoside triphosphate
PAGE .....	Polyacrylamide gel electrophoresis
PBS .....	Phosphate-buffered saline
PCR .....	Polymerase chain reaction
PEG .....	Polyethylene glycol
PET .....	Photoinduced electron transfer
PMH .....	Pumilio homology domain
PNA .....	Peptide nucleic acid
QM/MM .....	Quantum mechanics/molecular mechanics
r4Cl .....	4-Cyanoindole ribonucleoside
RNA .....	Ribonucleic acid
rRNA .....	Ribosomal RNA
RT .....	Reverse transcription
SAM .....	<i>S</i> -adenosyl methionine
SELEX .....	Systematic evolution of ligands by exponential enrichment
sgRNA .....	Single guide RNA
SHAPE .....	Selective 2'-hydroxyl acylation analyzed by primer extension
SIM .....	Structured illumination microscopy
SiR .....	Silicon rhodamine
SiRA .....	Silicon rhodamine aptamer
STED .....	Stimulated emission depletion
TBAF .....	Tetra- <i>n</i> -butylammonium fluoride
TBE .....	TRIS/Borate/EDTA
TCA .....	Trichloroacetic acid
TCSPC .....	Time-correlated single photon counting
TEMED .....	Tetramethylethylenediamine
TEN .....	TRIS/EDTA/NaCl
TFA .....	Trifluoroacetic acid
TGT .....	tRNA guanine transglycosylase
THF .....	Tetrahydrofurane
ThT .....	Thioflavin T
TLC .....	Thin layer chromatography
TMEDA .....	Tetramethylethylenediamine
TO .....	Thiazole orange
TOMCl .....	(Triisopropylsilyloxy)methyl chloride



TRIS.....	Tris(hydroxymethyl)aminomethane
tRNA.....	Transfer RNA
UMP.....	Uridine monophosphate
UTR.....	Untranslated region
VR.....	Vibrational relaxation



## References

- (1) Palazzo, A. F.; Lee, E. S. Non-coding RNA: what is functional and what is junk? *Front. Genet.* **2015**, *6*, 2.
- (2) Hangauer, M. J.; Vaughn, I. W.; McManus, M. T. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.* **2013**, *9*, e1003569.
- (3) Di Iulio, J.; Bartha, I.; Wong, E. H. M.; Yu, H.-C.; Lavrenko, V.; Yang, D.; Jung, I.; Hicks, M. A.; Shah, N.; Kirkness, E. F.; *et al.* The human noncoding genome defined by genetic diversity. *Nat. Genet.* **2018**, *50*, 333–337.
- (4) Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874.
- (5) Cech, T. R.; Steitz, J. A. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* **2014**, *157*, 77–94.
- (6) Diamantopoulos, M. A.; Tsiakanikas, P.; Scorilas, A. Non-coding RNAs: the riddle of the transcriptome and their perspectives in cancer. *Ann. Transl. Med.* **2018**, *6*, 241.
- (7) Amin, N.; McGrath, A.; Chen, Y.-P. P. Evaluation of deep learning in non-coding RNA classification. *Nat. Mach. Intell.* **2019**, *1*, 246–256.
- (8) Boivin, V.; Faucher-Giguère, L.; Scott, M.; Abou-Elela, S. The cellular landscape of mid-size noncoding RNA. *Wiley Interdiscip. Rev. RNA* **2019**, *10*, e1530.
- (9) Dahariya, S.; Paddibhatla, I.; Kumar, S.; Raghuvanshi, S.; Pallepati, A.; Gutti, R. K. Long non-coding RNA: Classification, biogenesis and functions in blood cells. *Mol. Immunol.* **2019**, *112*, 82–92.
- (10) Johansson, J.; Mandin, P.; Renzoni, A.; Chiaruttini, C.; Springer, M.; Cossart, P. An RNA Thermosensor Controls Expression of Virulence Genes in *Listeria monocytogenes*. *Cell* **2002**, *110*, 551–561.
- (11) Kreuzer, K. D.; Henkin, T. M. The T-Box Riboswitch: tRNA as an Effector to Modulate Gene Regulation. *Microbiol. Spectr.* **2018**, *6*.
- (12) Mandal, M.; Breaker, R. R. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451–463.
- (13) Winkler, W. C.; Breaker, R. R. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **2005**, *59*, 487–517.
- (14) Serganov, A.; Patel, D. J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* **2007**, *8*, 776–790.
- (15) Day, R. N.; Davidson, M. W. The fluorescent protein palette: Tools for cellular imaging. *Chem. Soc. Rev.* **2009**, *38*, 2887–2921.
- (16) Qin, P. Z.; Pyle, A. M. Site-specific labeling of RNA with fluorophores and other structural probes. *Methods* **1999**, *18*, 60–70.
- (17) Paredes, E.; Das, S. R. Click chemistry for rapid labeling and ligation of RNA. *ChemBioChem* **2011**, *12*, 125–131.
- (18) Samanta, B.; Seikowski, J.; Höbartner, C. Fluorogenic Labeling of 5-Formylpyrimidine Nucleotides in DNA and RNA. *Angew. Chem. Int. Ed.* **2016**, *55*, 1912–1916.
- (19) Alexander, S. C.; Busby, K. N.; Cole, C. M.; Zhou, C. Y.; Devaraj, N. K. Site-Specific Covalent Labeling of RNA by Enzymatic Transglycosylation. *J. Am. Chem. Soc.* **2015**, *137*, 12756–12759.
- (20) Holstein, J. M.; Anhäuser, L.; Rentmeister, A. Modifying the 5'-Cap for Click Reactions of Eukaryotic mRNA and To Tune Translation Efficiency in Living Cells. *Angew. Chem. Int. Ed.* **2016**, *55*, 10899–10903.

- (21) Büttner, L.; Javadi-Zarnaghi, F.; Höbartner, C. Site-specific labeling of RNA at internal ribose hydroxyl groups: terbium-assisted deoxyribozymes at work. *J. Am. Chem. Soc.* **2014**, *136*, 8131–8137.
- (22) Ghaem Maghami, M.; Scheitl, C. P. M.; Höbartner, C. Direct in Vitro Selection of Trans-Acting Ribozymes for Posttranscriptional, Site-Specific, and Covalent Fluorescent Labeling of RNA. *J. Am. Chem. Soc.* **2019**, *141*, 19546–19549.
- (23) Ghaem Maghami, M.; Dey, S.; Lenz, A.-K.; Höbartner, C. Repurposing Antiviral Drugs for Orthogonal RNA-Catalyzed Labeling of RNA. *Angew. Chem. Int. Ed.* [Online early access]. DOI: 10.1002/anie.202001300.
- (24) Muthmann, N.; Hartstock, K.; Rentmeister, A. Chemo-enzymatic treatment of RNA to facilitate analyses. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, e1561.
- (25) Tyagi, S. Imaging intracellular RNA distribution and dynamics in living cells. *Nat. Methods* **2009**, *6*, 331–338.
- (26) Mannack, L. V. J. C.; Eising, S.; Rentmeister, A. Current techniques for visualizing RNA in cells. *F1000Research* **2016**, *5*.
- (27) Tyagi, S.; Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **1996**, *14*, 303–308.
- (28) Silverman, A. P.; Kool, E. T. Quenched probes for highly specific detection of cellular RNAs. *Trends Biotechnol.* **2005**, *23*, 225–230.
- (29) Köhler, O.; Jarikote, D. V.; Seitz, O. Forced intercalation probes (FIT Probes): thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide-polymorphism detection. *ChemBioChem* **2005**, *6*, 69–77.
- (30) Hövelmann, F.; Bethge, L.; Seitz, O. Single labeled DNA FIT probes for avoiding false-positive signaling in the detection of DNA/RNA in qPCR or cell media. *ChemBioChem* **2012**, *13*, 2072–2081.
- (31) Hövelmann, F.; Seitz, O. DNA Stains as Surrogate Nucleobases in Fluorogenic Hybridization Probes. *Acc. Chem. Res.* **2016**, *49*, 714–723.
- (32) Tsourkas, A.; Behlke, M. A.; Bao, G. Hybridization of 2'-O-methyl and 2'-deoxy molecular beacons to RNA and DNA targets. *Nucleic Acids Res.* **2002**, *30*, 5168–5174.
- (33) Wang, L.; Yang, C. J.; Medley, C. D.; Benner, S. A.; Tan, W. Locked nucleic acid molecular beacons. *J. Am. Chem. Soc.* **2005**, *127*, 15664–15665.
- (34) Silverman, A. P.; Kool, E. T. Oligonucleotide Probes for RNA-Targeted Fluorescence In Situ Hybridization. *Adv. Clin. Chem.* **2007**, *43*, 79–115.
- (35) Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. H.; Long, R. M. Localization of ASH1 mRNA Particles in Living Yeast. *Mol. Cell* **1998**, *2*, 437–445.
- (36) Wang, X.; McLachlan, J.; Zamore, P. D.; Hall, T. M.T. Modular Recognition of RNA by a Human Pumilio-Homology Domain. *Cell* **2002**, *110*, 501–512.
- (37) Ozawa, T.; Natori, Y.; Sato, M.; Umezawa, Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat. Methods* **2007**, *4*, 413–419.
- (38) Breaker, R. R. Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*.
- (39) Mironov, A. S.; Gusarov, I.; Rafikov, R.; Lopez, L. E.; Shatalin, K.; Kreneva, R. A.; Perumov, D. A.; Nudler, E. Sensing Small Molecules by Nascent RNA. *Cell* **2002**, *111*, 747–756.
- (40) Nahvi, A.; Sudarsan, N.; Ebert, M. S.; Zou, X.; Brown, K. L.; Breaker, R. R. Genetic Control by a Metabolite Binding mRNA. *Chem. Biol.* **2002**, *9*, 1043–1049.
- (41) Winkler, W.; Nahvi, A.; Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002**, *419*, 952–956.
- (42) Ellington, A. D.; Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818–822.

- (43) Tuerk, C.; Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **1990**, *249*, 505–510.
- (44) Davis, J. H.; Szostak, J. W. Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11616–11621.
- (45) Bouhedda, F.; Autour, A.; Ryckelynck, M. Light-Up RNA Aptamers and Their Cognate Fluorogens: From Their Development to Their Applications. *Int. J. Mol. Sci.* **2017**, *19*.
- (46) Heus, H. A.; Pardi, A. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **1991**, *253*, 191–194.
- (47) Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* **1996**, *2*, 628–640.
- (48) Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature* **1996**, *382*, 183–186.
- (49) Patel, D. J.; Suri, A. K.; Jiang, F.; Jiang, L.; Fan, P.; Kumar, R. A.; Nonin, S. Structure, recognition and adaptive binding in RNA aptamer complexes. *J. Mol. Biol.* **1997**, *272*, 645–664.
- (50) Jiang, L.; Suri, A. K.; Fiala, R.; Patel, D. J. Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex. *Chem. Biol.* **1997**, *4*, 35–50.
- (51) Jiang, L.; Patel, D. J. Solution structure of the tobramycin-RNA aptamer complex. *Nat. Struct. Mol. Biol.* **1998**, *5*, 769–774.
- (52) Spöring, M.; Finke, M.; Hartig, J. S. Aptamers in RNA-based switches of gene expression. *Curr. Opin. Biotechnol.* **2019**, *63*, 34–40.
- (53) Ilgu, M.; Nilsen-Hamilton, M. Aptamers in analytics. *Analyst* **2016**, *141*, 1551–1568.
- (54) Holeman, L. A.; Robinson, S. L.; Szostak, J. W.; Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Fold. Des.* **1998**, *3*, 423–431.
- (55) Sunbul, M.; Jäschke, A. SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging. *Nucleic Acids Res.* **2018**, *46*, e110.
- (56) Eydeler, K.; Magbanua, E.; Werner, A.; Ziegelmüller, P.; Hahn, U. Fluorophore binding aptamers as a tool for RNA visualization. *Biophys. J.* **2009**, *96*, 3703–3707.
- (57) Grate, D.; Wilson, C. Laser-mediated, site-specific inactivation of RNA transcripts. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6131–6136.
- (58) Jay, D. G.; Keshishian, H. Laser inactivation of fasciclin I disrupts axon adhesion of grasshopper pioneer neurons. *Nature* **1990**, *348*, 548–550.
- (59) Liao, J. C.; Roider, J.; Jay, D. G. Chromophore-assisted laser inactivation of proteins is mediated by the photogeneration of free radicals. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2659–2663.
- (60) Baugh, C.; Grate, D.; Wilson, C. 2.8 Å crystal structure of the malachite green aptamer. *J. Mol. Biol.* **2000**, *301*, 117–128.
- (61) Oster, G.; Nishijima, Y. Fluorescence and Internal Rotation: Their Dependence on Viscosity of the Medium 1. *J. Am. Chem. Soc.* **1956**, *78*, 1581–1584.
- (62) Bao, G.; Rhee, W. J.; Tsourkas, A. Fluorescent probes for live-cell RNA detection. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 25–47.
- (63) Kraus, G. A.; Jeon, I.; Nilsen-Hamilton, M.; Awad, A. M.; Banerjee, J.; Parvin, B. Fluorinated analogs of malachite green: synthesis and toxicity. *Molecules* **2008**, *13*, 986–994.
- (64) Lux, J.; Peña, E. J.; Bolze, F.; Heinlein, M.; Nicoud, J.-F. Malachite green derivatives for two-photon RNA detection. *ChemBioChem* **2012**, *13*, 1206–1213.
- (65) Yerramilli, V. S.; Kim, K. H. Labeling RNAs in Live Cells Using Malachite Green Aptamer Scaffolds as Fluorescent Probes. *ACS Synth. Biol.* **2018**, *7*, 758–766.
- (66) Sparano, B. A.; Koide, K. A strategy for the development of small-molecule-based sensors that strongly fluoresce when bound to a specific RNA. *J. Am. Chem. Soc.* **2005**, *127*, 14954–14955.

- (67) Sparano, B. A.; Shahi, S. P.; Koide, K. Effect of binding and conformation on fluorescence quenching in new 2',7'-dichlorofluorescein derivatives. *Org. Lett.* **2004**, *6*, 1947–1949.
- (68) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. Experimental and computational investigation of unsymmetrical cyanine dyes: understanding torsionally responsive fluorogenic dyes. *J. Am. Chem. Soc.* **2007**, *129*, 5710–5718.
- (69) Constantin, T. P.; Silva, G. L.; Robertson, K. L.; Hamilton, T. P.; Fague, K.; Waggoner, A. S.; Armitage, B. A. Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromolecules. *Org. Lett.* **2008**, *10*, 1561–1564.
- (70) Schifferer, M.; Griesbeck, O. Application of aptamers and autofluorescent proteins for RNA visualization. *Integr. Biol.* **2009**, *1*, 499–505.
- (71) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **2011**, *333*, 642–646.
- (72) Schultz, C. Fluorescent revelations. *Chem. Biol.* **2009**, *16*, 107–111.
- (73) Martin, M. E.; Negri, F.; Olivucci, M. Origin, nature, and fate of the fluorescent state of the green fluorescent protein chromophore at the CASPT2//CASCF resolution. *J. Am. Chem. Soc.* **2004**, *126*, 5452–5464.
- (74) Tolosa, L.; Donato, M. T.; Gómez-Lechón, M. J. General Cytotoxicity Assessment by Means of the MTT Assay. *Methods Mol. Biol.* **2015**, *1250*, 333–348.
- (75) Cormack, B. P.; Valdivia, R. H.; Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **1996**, *173*, 33–38.
- (76) Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nat. Methods* **2019**, *16*, 277–278.
- (77) Ponchon, L.; Dardel, F. Recombinant RNA technology: the tRNA scaffold. *Nat. Methods* **2007**, *4*, 571–576.
- (78) Strack, R. L.; Disney, M. D.; Jaffrey, S. R. A superfolder Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat. Methods* **2013**, *10*, 1219–1224.
- (79) Filonov, G. S.; Moon, J. D.; Svendsen, N.; Jaffrey, S. R. Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **2014**, *136*, 16299–16308.
- (80) Filonov, G. S.; Song, W.; Jaffrey, S. R. Spectral Tuning by a Single Nucleotide Controls the Fluorescence Properties of a Fluorogenic Aptamer. *Biochemistry* **2019**, *58*, 1560–1564.
- (81) Huang, H.; Suslov, N. B.; Li, N.-S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat. Chem. Biol.* **2014**, *10*, 686–691.
- (82) Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat. Struct. Mol. Biol.* **2014**, *21*, 658–663.
- (83) Song, W.; Strack, R. L.; Svendsen, N.; Jaffrey, S. R. Plug-and-play fluorophores extend the spectral properties of Spinach. *J. Am. Chem. Soc.* **2014**, *136*, 1198–1201.
- (84) Ageely, E. A.; Kartje, Z. J.; Rohilla, K. J.; Barkau, C. L.; Gagnon, K. T. Quadruplex-Flanking Stem Structures Modulate the Stability and Metal Ion Preferences of RNA Mimics of GFP. *ACS Chem. Biol.* **2016**, *11*, 2398–2406.
- (85) Wang, P.; Querard, J.; Maurin, S.; Nath, S. S.; Le Saux, T.; Gautier, A.; Jullien, L. Photochemical properties of Spinach and its use in selective imaging. *Chem. Sci.* **2013**, *4*, 2865–2873.
- (86) Han, K. Y.; Leslie, B. J.; Fei, J.; Zhang, J.; Ha, T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J. Am. Chem. Soc.* **2013**, *135*, 19033–19038.

- (87) Jaffey, S. R. Weill Medical College of Cornell University, New York, NY. Personal communication, 2017.
- (88) Song, W.; Filonov, G. S.; Kim, H.; Hirsch, M.; Li, X.; Moon, J. D.; Jaffrey, S. R. Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat. Chem. Biol.* **2017**, *13*, 1187–1194.
- (89) Warner, K. D.; Sjekloća, L.; Song, W.; Filonov, G. S.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat. Chem. Biol.* **2017**, *13*, 1195–1201.
- (90) Kim, H.; Jaffrey, S. R. A Fluorogenic RNA-Based Sensor Activated by Metabolite-Induced RNA Dimerization. *Cell Chem. Biol.* **2019**, *26*, 1725–1731.e6.
- (91) Ouellet, J. RNA Fluorescence with Light-Up Aptamers. *Front. Chem.* **2016**, *4*, 29.
- (92) Truong, L.; Ferré-D'Amaré, A. R. From fluorescent proteins to fluorogenic RNAs: Tools for imaging cellular macromolecules. *Protein. Sci.* **2019**, *28*, 1374–1386.
- (93) Trachman, R. J.; Ferré-D'Amaré, A. R. Tracking RNA with light: selection, structure, and design of fluorescence turn-on RNA aptamers. *Q. Rev. Biophys.* **2019**, *52*, e8.
- (94) Neubacher, S.; Hennig, S. RNA Structure and Cellular Applications of Fluorescent Light-Up Aptamers. *Angew. Chem. Int. Ed.* **2019**, *58*, 1266–1279.
- (95) Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocar, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J. RNA mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking. *ACS Chem. Biol.* **2014**, *9*, 2412–2420.
- (96) Trachman, R. J.; Demeshkina, N. A.; Lau, M. W. L.; Panchapakesan, S. S. S.; Jeng, S. C. Y.; Unrau, P. J.; Ferré-D'Amaré, A. R. Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat. Chem. Biol.* **2017**, *13*, 807–813.
- (97) Jeng, S. C. Y.; Chan, H. H. Y.; Booy, E. P.; McKenna, S. A.; Unrau, P. J. Fluorophore ligand binding and complex stabilization of the RNA Mango and RNA Spinach aptamers. *RNA* **2016**, *22*, 1884–1892.
- (98) Ryckelynck, M.; Baudrey, S.; Rick, C.; Marin, A.; Coldren, F.; Westhof, E.; Griffiths, A. D. Using droplet-based microfluidics to improve the catalytic properties of RNA under multiple-turnover conditions. *RNA* **2015**, *21*, 458–469.
- (99) Autour, A.; C Y Jeng, S.; D Cawte, A.; Abdolazadeh, A.; Galli, A.; Panchapakesan, S. S. S.; Rueda, D.; Ryckelynck, M.; Unrau, P. J. Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat. Commun.* **2018**, *9*, 656.
- (100) Trachman, R. J.; Abdolazadeh, A.; Andreoni, A.; Cojocar, R.; Knutson, J. R.; Ryckelynck, M.; Unrau, P. J.; Ferré-D'Amaré, A. R. Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness. *Biochemistry* **2018**, *57*, 3544–3548.
- (101) Trachman, R. J.; Autour, A.; Jeng, S. C. Y.; Abdolazadeh, A.; Andreoni, A.; Cojocar, R.; Garipov, R.; Dolgosheina, E. V.; Knutson, J. R.; Ryckelynck, M.; *et al.* Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nat. Chem. Biol.* **2019**, *15*, 472–479.
- (102) Cawte, A. D.; Unrau, P. J.; Rueda, D. S. Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat. Commun.* **2020**, *11*, 1283.
- (103) Sunbul, M.; Jäschke, A. Contact-mediated quenching for RNA imaging in bacteria with a fluorophore-binding aptamer. *Angew. Chem. Int. Ed.* **2013**, *52*, 13401–13404.
- (104) Arora, A.; Sunbul, M.; Jäschke, A. Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers. *Nucleic Acids Res.* **2015**, *43*, e144.
- (105) Johnson, J. E.; Reyes, F. E.; Polaski, J. T.; Batey, R. T. B12 cofactors directly stabilize an mRNA regulatory switch. *Nature* **2012**, *492*, 133–137.

- (106) Braselmann, E.; Wierzba, A. J.; Polaski, J. T.; Chromiński, M.; Holmes, Z. E.; Hung, S.-T.; Batan, D.; Wheeler, J. R.; Parker, R.; Jimenez, R.; *et al.* A multicolor riboswitch-based platform for imaging of RNA in live mammalian cells. *Nat. Chem. Biol.* **2018**, *14*, 964–971.
- (107) Bouhedda, F.; Fam, K. T.; Collot, M.; Autour, A.; Marzi, S.; Klymchenko, A.; Ryckelynck, M. A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nat. Chem. Biol.* [Online early access]. DOI: 10.1038/s41589-019-0381-8.
- (108) Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigsmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa, I. R.; Luo, Z.-G.; *et al.* A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* **2013**, *5*, 132–139.
- (109) Wirth, R.; Gao, P.; Nienhaus, G. U.; Sunbul, M.; Jäschke, A. SiRA: A Silicon Rhodamine-Binding Aptamer for Live-Cell Super-Resolution RNA Imaging. *J. Am. Chem. Soc.* **2019**, *141*, 7562–7571.
- (110) Chen, X.; Zhang, D.; Su, N.; Bao, B.; Xie, X.; Zuo, F.; Yang, L.; Wang, H.; Jiang, L.; Lin, Q.; *et al.* Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat. Biotechnol.* [Online early access]. DOI: 10.1038/s41587-019-0249-1.
- (111) Steinmetzger, C.; Palanisamy, N.; Gore, K. R.; Höbartner, C. A Multicolor Large Stokes Shift Fluorogen-Activating RNA Aptamer with Cationic Chromophores. *Chem. Eur. J.* **2019**, *25*, 1931–1935.
- (112) Steinmetzger, C.; Bessi, I.; Lenz, A.-K.; Höbartner, C. Structure-fluorescence activation relationships of a large Stokes shift fluorogenic RNA aptamer. *Nucleic Acids Res.* **2019**, *47*, 11538–11550.
- (113) Steinmetzger, C.; Bäuerlein, C.; Höbartner, C. Supramolecular Fluorescence Resonance Energy Transfer in Nucleobase-Modified Fluorogenic RNA Aptamers. *Angew. Chem. Int. Ed.* **2020**, *59*, 6760–6764.
- (114) Gruber, A. R.; Lorenz, R.; Bernhart, S. H.; Neuböck, R.; Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res.* **2008**, *36*, W70-4.
- (115) Liptonok, S. P.; Conyard, J.; Page, P. C. B.; Chan, Y.; You, M.; Jaffrey, S. R.; Meech, S. R. Photoacid Behaviour in a Fluorinated Green Fluorescent Protein Chromophore: Ultrafast Formation of Anion and Zwitterion States.†. *Chem. Sci.* **2016**, *7*, 5747–5752.
- (116) Bell, A. F.; He, X.; Wachter, R. M.; Tonge, P. J. Probing the ground state structure of the green fluorescent protein chromophore using Raman spectroscopy. *Biochemistry* **2000**, *39*, 4423–4431.
- (117) Finkler, B.; Spies, C.; Vester, M.; Walte, F.; Omlor, K.; Riemann, I.; Zimmer, M.; Stracke, F.; Gerhards, M.; Jung, G. Highly photostable "super"-photoacids for ultrasensitive fluorescence spectroscopy. *Photochem. Photobiol. Sci.* **2014**, *13*, 548–562.
- (118) Fernandez-Millan, P.; Autour, A.; Ennifar, E.; Westhof, E.; Ryckelynck, M. Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI. *RNA* **2017**, *23*, 1788–1795.
- (119) Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4*, 38.
- (120) Mullen, M. A.; Assmann, S. M.; Bevilacqua, P. C. Toward a digital gene response: RNA G-quadruplexes with fewer quartets fold with higher cooperativity. *J. Am. Chem. Soc.* **2012**, *134*, 812–815.
- (121) Savage, J. C.; Davare, M. A.; Shinde, U. P. Subtle sequence variations alter tripartite complex kinetics and G-Quadruplex dynamics in RNA aptamer Broccoli. *Chem. Commun.* [Online early access]. DOI: 10.1039/C9CC09375C.
- (122) Savage, J. C.; Shinde, U. P. Oregon Health & Science University, Portland, OR. Personal communication, 2020.



- (123) Brion, P.; Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113–137.
- (124) Serra, M. J.; Baird, J. D.; Dale, T.; Fey, B. L.; Retatagos, K.; Westhof, E. Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA* **2002**, *8*, 307–323.
- (125) Draper, D. E. RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.* **2008**, *95*, 5489–5495.
- (126) Grubbs, R. D. Intracellular magnesium and magnesium buffering. *Biometals* **2002**, *15*, 251–259.
- (127) Li, X.; Kim, H.; Litke, J. L.; Wu, J.; Jaffrey, S. R. Fluorophore-Promoted RNA Folding and Photostability Enables Imaging of Single Broccoli-Tagged mRNAs in Live Mammalian Cells. *Angew. Chem. Int. Ed.* **2020**, *59*, 4511–4518.
- (128) Wu, L.; Burgess, K. Syntheses of highly fluorescent GFP-chromophore analogues. *J. Am. Chem. Soc.* **2008**, *130*, 4089–4096.
- (129) Lerestif, J. M.; Bazureau, J. P.; Hamelin, J. Cycloaddition with stabilized imidates as potential azomethines ylides : A new route to 2-imidazoline and 4-yliden-s-imidazolinone. *Tetrahedron Lett.* **1993**, *34*, 4639–4642.
- (130) Lerestif, J. M.; Perrocheau, J.; Tonnard, F.; Bazureau, J. P.; Hamelin, J. 1,3-Dipolar cycloaddition of imidate ylides on imino-alcohols: Synthesis of new imidazolones using solvent free conditions. *Tetrahedron* **1995**, *51*, 6757–6774.
- (131) Kerneur, G.; Lerestif, J. M.; Bazureau, J. P.; Hamelin, J. Convenient Preparation of 4-Alkylidene-1H-imidazol-5(4H)-one Derivatives from Imidate and Aldehydes by a Solvent-Free Cycloaddition under Microwaves. *Synthesis* **1997**, *1997*, 287–289.
- (132) Baldridge, A.; Kowalik, J.; Tolbert, L. Efficient Synthesis of New 4-Arylideneimidazol-5-ones Related to the GFP Chromophore by 2+3 Cyclocondensation of Arylideneimines with Imidate Ylides. *Synthesis* **2010**, *2010*, 2424–2436.
- (133) Conyard, J.; Kondo, M.; Heisler, I. A.; Jones, G.; Baldridge, A.; Tolbert, L. M.; Solntsev, K. M.; Meech, S. R. Chemically modulating the photophysics of the GFP chromophore. *J. Phys. Chem. B* **2011**, *115*, 1571–1577.
- (134) Baldridge, A.; Feng, S.; Chang, Y.-T.; Tolbert, L. M. Recapture of GFP chromophore fluorescence in a protein host. *ACS. Comb. Sci.* **2011**, *13*, 214–217.
- (135) Samanta, S. R.; Da Silva, J. P.; Baldridge, A.; Tolbert, L. M.; Ramamurthy, V. A latent reaction in a model GFP chromophore revealed upon confinement: photohydroxylation of ortho-halo benzylidene-3-methylimidazolidiones via an electrocyclization process. *Org. Lett.* **2014**, *16*, 3304–3307.
- (136) Miyawaki, A.; Shcherbakova, D. M.; Verkhusha, V. V. Red fluorescent proteins: chromophore formation and cellular applications. *Curr. Opin. Struct. Biol.* **2012**, *22*, 679–688.
- (137) McCapra, F.; Razavi, Z.; Neary, A. P. The fluorescence of the chromophore of the green fluorescent protein of *Aequorea* and *Renilla*. *J. Chem. Soc., Chem. Commun.* **1988**, 790.
- (138) He, X.; Bell, A. F.; Tonge, P. J. Synthesis and spectroscopic studies of model red fluorescent protein chromophores. *Org. Lett.* **2002**, *4*, 1523–1526.
- (139) Yampolsky, I. V.; Kislukhin, A. A.; Amatov, T. T.; Shcherbo, D.; Potapov, V. K.; Lukyanov, S.; Lukyanov, K. A. Synthesis and properties of the red chromophore of the green-to-red photoconvertible fluorescent protein Kaede and its analogs. *Bioorg. Chem.* **2008**, *36*, 96–104.
- (140) Chuang, W.-T.; Chen, B.-S.; Chen, K.-Y.; Hsieh, C.-C.; Chou, P.-T. Fluorescent protein red Kaede chromophore; one-step, high-yield synthesis and potential application for solar cells. *Chem. Commun.* **2009**, 6982–6984.

- (141) Baleeva, N. S.; Myannik, K. A.; Yampolsky, I. V.; Baranov, M. S. Bioinspired Fluorescent Dyes Based on a Conformationally Locked Chromophore of the Fluorescent Protein Kaede. *Eur. J. Org. Chem.* **2015**, *2015*, 5716–5721.
- (142) Jaffrey, S. R.; Paige, J. S. Coupled recognition/detection system for in vivo and in vitro use. US2010024622, Feb 18, 2010.
- (143) Yadav, V.; Babu, K.G. A Remarkably Efficient Markovnikov Hydrochlorination of Olefins and Transformation of Nitriles into Imidates by Use of AcCl and an Alcohol. *Eur. J. Org. Chem.* **2005**, *2005*, 452–456.
- (144) D'Antona, N.; Morrone, R.; Gambera, G.; Pedotti, S. Enantiorecognition of planar "metallocenic" chirality by a nitrile hydratase/amidase bienzymatic system. *Org. Biomol. Chem.* **2016**, *14*, 4393–4399.
- (145) Hercouet, A.; Le Corre, M. Triphenylphosphonium Bromide: A Convenient and Quantitative Source of Gaseous Hydrogen Bromide. *Synthesis* **1988**, *1988*, 157–158.
- (146) Kalita, D.; Morisue, M.; Kobuke, Y. Synthesis and electrochemical properties of slipped-cofacial porphyrin dimers of ferrocene-functionalized Zn-imidazolyl-porphyrins as potential terminal electron donors in photosynthetic models. *New J. Chem.* **2006**, *30*, 77–92.
- (147) Manoni, F.; Connon, S. J. Catalytic asymmetric Tamura cycloadditions. *Angew. Chem. Int. Ed.* **2014**, *53*, 2628–2632.
- (148) Couto, U. R.; Navarro-Vázquez, A.; Tormena, C. F. Unexpected behavior of the 3 JCH coupling constant in unsaturated compounds. *Magn. Reson. Chem.* **2019**, *57*, 939–945.
- (149) Ieda, N.; Nakagawa, H.; Horinouchi, T.; Peng, T.; Yang, D.; Tsumoto, H.; Suzuki, T.; Fukuhara, K.; Miyata, N. Peroxynitrite generation from a NO-releasing nitrobenzene derivative in response to photoirradiation. *Chem. Commun.* **2011**, *47*, 6449–6451.
- (150) Gepshtein, R.; Huppert, D.; Agmon, N. Deactivation mechanism of the green fluorescent chromophore. *J. Phys. Chem. B* **2006**, *110*, 4434–4442.
- (151) Leroux, F. R.; Manteau, B.; Vors, J.-P.; Pazenok, S. Trifluoromethyl ethers--synthesis and properties of an unusual substituent. *Beilstein J. Org. Chem.* **2008**, *4*, 13.
- (152) Sessler, J. L.; Sathiosatham, M.; Brown, C. T.; Rhodes, T. A.; Wiederrecht, G. Hydrogen-bond-mediated photoinduced electron-transfer: novel dimethylaniline-anthracene ensembles formed via Watson-Crick base-pairing. *J. Am. Chem. Soc.* **2001**, *123*, 3655–3660.
- (153) Povarova, N. V.; Zaitseva, S. O.; Baleeva, N. S.; Smirnov, A. Y.; Myasnyanko, I. N.; Zagudaylova, M. B.; Bozhanova, N. G.; Gorbachev, D. A.; Malyshevskaya, K. K.; Gavrikov, A. S.; *et al.* Red-Shifted Substrates for FAST Fluorogen-Activating Protein Based on the GFP-Like Chromophores. *Chem. Eur. J.* **2019**, *25*, 9592–9596.
- (154) Chen, Y.; Rich, R. L.; Gai, F.; Petrich, J. W. Fluorescent species of 7-azaindole and 7-azatryptophan in water. *J. Phys. Chem.* **1993**, *97*, 1770–1780.
- (155) Varani, G.; McClain, W. H. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO reports* **2000**, *1*, 18–23.
- (156) Gabelica, V.; Maeda, R.; Fujimoto, T.; Yaku, H.; Murashima, T.; Sugimoto, N.; Miyoshi, D. Multiple and cooperative binding of fluorescence light-up probe thioflavin T with human telomere DNA G-quadruplex. *Biochemistry* **2013**, *52*, 5620–5628.
- (157) Xu, S.; Li, Q.; Xiang, J.; Yang, Q.; Sun, H.; Guan, A.; Wang, L.; Liu, Y.; Yu, L.; Shi, Y.; *et al.* Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes. *Sci. Rep.* **2016**, *6*, 24793.
- (158) Li, Y.; Xu, S.; Wu, X.; Xu, Q.; Zhao, Y.; Lou, X.; Yang, X. Thioflavin T as a fluorescence light-up probe for both parallel and antiparallel G-quadruplexes of 29-mer thrombin binding aptamer. *Anal. Bioanal. Chem.* **2016**, *408*, 8025–8036.

- (159) Nygren, J.; Svanvik, N.; Kubista, M. The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **1998**, *46*, 39–51.
- (160) Lubitz, I.; Zikich, D.; Kotlyar, A. Specific high-affinity binding of thiazole orange to triplex and G-quadruplex DNA. *Biochemistry* **2010**, *49*, 3567–3574.
- (161) La Renaud de Faverie, A.; Guédin, A.; Bedrat, A.; Yatsunyk, L. A.; Mergny, J.-L. Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res.* **2014**, *42*, e65.
- (162) Sjekloča, L.; Ferré-D'Amaré, A. R. Binding between G Quadruplexes at the Homodimer Interface of the Corn RNA Aptamer Strongly Activates Thioflavin T Fluorescence. *Cell Chem. Biol.* **2019**, *26*, 1159-1168.e4.
- (163) Monchaud, D.; Teulade-Fichou, M.-P. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* **2008**, *6*, 627–636.
- (164) Granzhan, A.; Ihmels, H.; Jäger, K. Diazonium- and tetraazoniapolycyclic cations as motif for quadruplex-DNA ligands. *Chem. Commun.* **2009**, 1249–1251.
- (165) Del Castillo, P.; Horobin, R. W.; Blázquez-Castro, A.; Stockert, J. C. Binding of cationic dyes to DNA: distinguishing intercalation and groove binding mechanisms using simple experimental and numerical models. *Biotech. Histochem.* **2010**, *85*, 247–256.
- (166) Kotar, A.; Kocman, V.; Plavec, J. Intercalation of a Heterocyclic Ligand between Quartets in a G-Rich Tetrahelical Structure. *Chem. Eur. J.* **2020**, *26*, 814–817.
- (167) Rachwal, P. A.; Findlow, I. S.; Werner, J. M.; Brown, T.; Fox, K. R. Intramolecular DNA quadruplexes with different arrangements of short and long loops. *Nucleic Acids Res.* **2007**, *35*, 4214–4222.
- (168) Li, X.; Chung, L. W.; Li, G. Multiscale Simulations on Spectral Tuning and the Photoisomerization Mechanism in Fluorescent RNA Spinach. *J. Chem. Theory Comput.* **2016**, *12*, 5453–5464.
- (169) Bose, S.; Chakrabarty, S.; Ghosh, D. Electrostatic Origin of the Red Solvatochromic Shift of DFHBDI in RNA Spinach. *J. Phys. Chem. B* **2017**, *121*, 4790–4798.
- (170) Resch-Genger, U.; Rurack, K. Determination of the photoluminescence quantum yield of dilute dye solutions (IUPAC Technical Report). *Pure Appl. Chem.* **2013**, *85*, 2005–2026.
- (171) Würth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-Genger, U. Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nat. Protoc.* **2013**, *8*, 1535–1550.
- (172) Hirose, K. A Practical Guide for the Determination of Binding Constants. *J. Incl. Phenom. Macrocycl. Chem.* **2001**, *39*, 193–209.
- (173) Thordarson, P. Determining association constants from titration experiments in supramolecular chemistry. *Chem. Soc. Rev.* **2011**, *40*, 1305–1323.
- (174) Doyle, M. L. Characterization of binding interactions by isothermal titration calorimetry. *Curr. Opin. Biotechnol.* **1997**, *8*, 31–35.
- (175) Feig, A. L. Applications of isothermal titration calorimetry in RNA biochemistry and biophysics. *Biopolymers* **2007**, *87*, 293–301.
- (176) Cai, S.; Yan, J.; Xiong, H.; Liu, Y.; Peng, D.; Liu, Z. Investigations on the interface of nucleic acid aptamers and binding targets. *Analyst* **2018**, *143*, 5317–5338.
- (177) Reinstein, O.; Yoo, M.; Han, C.; Palmo, T.; Beckham, S. A.; Wilce, M. C. J.; Johnson, P. E. Quinine binding by the cocaine-binding aptamer. Thermodynamic and hydrodynamic analysis of high-affinity binding of an off-target ligand. *Biochemistry* **2013**, *52*, 8652–8662.
- (178) Jelesarov, I.; Bosshard, H. R. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.* **1999**, *12*, 3–18.

- (179) Wickiser, J. K.; Cheah, M. T.; Breaker, R. R.; Crothers, D. M. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **2005**, *44*, 13404–13414.
- (180) Vogt, A. D.; Di Cera, E. Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochemistry* **2012**, *51*, 5894–5902.
- (181) Paul, F.; Weikl, T. R. How to Distinguish Conformational Selection and Induced Fit Based on Chemical Relaxation Rates. *PLOS Comput. Biol.* **2016**, *12*, e1005067.
- (182) AbouHaidar, M. G.; Ivanov, I. G. Non-enzymatic RNA hydrolysis promoted by the combined catalytic activity of buffers and magnesium ions. *Z. Naturforsch. C* **1999**, *54*, 542–548.
- (183) Mergny, J.-L.; Phan, A.-T.; Lacroix, L. Following G-quartet formation by UV-spectroscopy. *FEBS Letters* **1998**, *435*, 74–78.
- (184) Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. NMR spectroscopy of RNA. *ChemBioChem* **2003**, *4*, 936–962.
- (185) Adrian, M.; Heddi, B.; Phan, A. T. NMR spectroscopy of G-quadruplexes. *Methods* **2012**, *57*, 11–24.
- (186) Filonov, G. S.; Kam, C. W.; Song, W.; Jaffrey, S. R. In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. *Chem. Biol.* **2015**, *22*, 649–660.
- (187) Johnston, M.; Hillier, L.; Riles, L.; Albermann, K.; André, B.; Ansorge, W.; Benes, V.; Brückner, M.; Delius, H.; Dubois, E.; *et al.* The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII. *Nature* **1997**, *387*, 87–90.
- (188) Haller, A.; Soulière, M. F.; Micura, R. The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Acc. Chem. Res.* **2011**, *44*, 1339–1348.
- (189) Soulière, M. F.; Haller, A.; Rieder, R.; Micura, R. A powerful approach for the selection of 2-aminopurine substitution sites to investigate RNA folding. *J. Am. Chem. Soc.* **2011**, *133*, 16161–16167.
- (190) Rieder, R.; Lang, K.; Graber, D.; Micura, R. Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. *ChemBioChem* **2007**, *8*, 896–902.
- (191) Haller, A.; Rieder, U.; Aigner, M.; Blanchard, S. C.; Micura, R. Conformational capture of the SAM-II riboswitch. *Nat. Chem. Biol.* **2011**, *7*, 393–400.
- (192) Lang, K.; Rieder, R.; Micura, R. Ligand-induced folding of the thiM TPP riboswitch investigated by a structure-based fluorescence spectroscopic approach. *Nucleic Acids Res.* **2007**, *35*, 5370–5378.
- (193) Gustmann, H.; Segler, A.-L. J.; Gophane, D. B.; Reuss, A. J.; Grünwald, C.; Braun, M.; Weigand, J. E.; Sigurdsson, S. T.; Wachtveitl, J. Structure guided fluorescence labeling reveals a two-step binding mechanism of neomycin to its RNA aptamer. *Nucleic Acids Res.* **2019**, *47*, 15–28.
- (194) Börjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. Nucleic acid base analog FRET-pair facilitating detailed structural measurements in nucleic acid containing systems. *J. Am. Chem. Soc.* **2009**, *131*, 4288–4293.
- (195) Bood, M.; Füchtbauer, A. F.; Wranne, M. S.; Ro, J. J.; Sarangamath, S.; El-Sagheer, A. H.; Rupert, D. L. M.; Fisher, R. S.; Magennis, S. W.; Jones, A. C.; *et al.* Pentacyclic adenine: a versatile and exceptionally bright fluorescent DNA base analogue. *Chem. Sci.* **2018**, *9*, 3494–3502.
- (196) Wranne, M. S.; Füchtbauer, A. F.; Dumat, B.; Bood, M.; El-Sagheer, A. H.; Brown, T.; Gradén, H.; Grøtli, M.; Wilhelmsson, L. M. Toward Complete Sequence Flexibility of Nucleic Acid Base Analogue FRET. *J. Am. Chem. Soc.* **2017**, *139*, 9271–9280.
- (197) Han, J. H.; Yamamoto, S.; Park, S.; Sugiyama, H. Development of a Vivid FRET System Based on a Highly Emissive dG-dC Analogue Pair. *Chem. Eur. J.* **2017**, *23*, 7607–7613.

- (198) Füchtbauer, A. F.; Wranne, M. S.; Bood, M.; Weis, E.; Pfeiffer, P.; Nilsson, J. R.; Dahlén, A.; Grøtli, M.; Wilhelmsson, L. M. Interbase FRET in RNA: from A to Z. *Nucleic Acids Res.* **2019**, *47*, 9990–9997.
- (199) Li, S.; Olson, W. K.; Lu, X.-J. Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* **2019**, *47*, W26–W34.
- (200) Xie, Y.; Dix, A. V.; Tor, Y. FRET enabled real time detection of RNA-small molecule binding. *J. Am. Chem. Soc.* **2009**, *131*, 17605–17614.
- (201) Füchtbauer, A. F.; Preus, S.; Börjesson, K.; McPhee, S. A.; Lilley, D. M. J.; Wilhelmsson, L. M. Fluorescent RNA cytosine analogue - an internal probe for detailed structure and dynamics investigations. *Sci. Rep.* **2017**, *7*, 2393.
- (202) Shin, D.; Sinkeldam, R. W.; Tor, Y. Emissive RNA alphabet. *J. Am. Chem. Soc.* **2011**, *133*, 14912–14915.
- (203) Rovira, A. R.; Fin, A.; Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015**, *137*, 14602–14605.
- (204) Srivatsan, S. G.; Greco, N. J.; Tor, Y. A highly emissive fluorescent nucleoside that signals the activity of toxic ribosome-inactivating proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 6661–6665.
- (205) Loakes, D. Survey and summary: The applications of universal DNA base analogues. *Nucleic Acids Res.* **2001**, *29*, 2437–2447.
- (206) Koller, A. N.; Bozilovic, J.; Engels, J. W.; Gohlke, H. Aromatic N versus aromatic F: bioisosterism discovered in RNA base pairing interactions leads to a novel class of universal base analogs. *Nucleic Acids Res.* **2010**, *38*, 3133–3146.
- (207) Kimoto, M.; Mitsui, T.; Yokoyama, S.; Hirao, I. A unique fluorescent base analogue for the expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2010**, *132*, 4988–4989.
- (208) Spadafora, M.; Postupalenko, V. Y.; Shvadchak, V. V.; Klymchenko, A. S.; Mély, Y.; Burger, A.; Benhida, R. Efficient Synthesis of Ratiometric Fluorescent Nucleosides Featuring 3-Hydroxychromone Nucleobases. *Tetrahedron* **2009**, *65*, 7809–7816.
- (209) Passow, K. T.; Harki, D. A. 4-Cyanoindole-2'-deoxyribonucleoside (4CIN): A Universal Fluorescent Nucleoside Analogue. *Org. Lett.* **2018**, *20*, 4310–4313.
- (210) Dziuba, D.; Postupalenko, V. Y.; Spadafora, M.; Klymchenko, A. S.; Guérineau, V.; Mély, Y.; Benhida, R.; Burger, A. A universal nucleoside with strong two-band switchable fluorescence and sensitivity to the environment for investigating DNA interactions. *J. Am. Chem. Soc.* **2012**, *134*, 10209–10213.
- (211) Ahmed, I. A.; Acharyya, A.; Eng, C. M.; Rodgers, J. M.; DeGrado, W. F.; Jo, H.; Gai, F. 4-Cyanoindole-2'-deoxyribonucleoside as a Dual Fluorescence and Infrared Probe of DNA Structure and Dynamics. *Molecules* **2019**, *24*.
- (212) Vorbrüggen, H.; Höfle, G. On the Mechanism of Nucleoside Synthesis. *Chem. Ber.* **1981**, *114*, 1256–1268.
- (213) Chapter 3. Nucleosides and Nucleotides. In *Nucleic Acids in Chemistry and Biology*, 3rd ed.; Blackburn, G. M., Gait, M. J., Loakes, D., Williams, D. M., Eds.; Royal Society of Chemistry: Cambridge, 2007; pp 77–142.
- (214) Liang, Y.; Hnatiuk, N.; Rowley, J. M.; Whiting, B. T.; Coates, G. W.; Rablen, P. R.; Morton, M.; Howell, A. R. Access to oxetane-containing psico-nucleosides from 2-methyleneoxetanes: a role for neighboring group participation? *J. Org. Chem.* **2011**, *76*, 9962–9974.
- (215) Vorbrüggen, H.; Ruh-Pohlentz, C. Synthesis Of Nucleosides. *Organic Reactions*; Wiley: Hoboken, N.J., 2004; pp 1–630.
- (216) Vorbrüggen, H.; Lagoja, I. M.; Herdewijn, P. Synthesis of Ribonucleosides by Condensation Using Trimethylsilyl Triflate. *Curr. Protoc. Nucleic Acid. Chem.* **2006**, *27*, 1.13.1–1.13.16.

- (217) Harki, D. A.; Graci, J. D.; Edathil, J. P.; Castro, C.; Cameron, C. E.; Peterson, B. R. Synthesis of a Universal 5-Nitroindole Ribonucleotide and Incorporation into RNA by a Viral RNA-Dependent RNA Polymerase. *ChemBioChem* **2007**, *8*, 1359–1362.
- (218) Sokolova, T. N.; Shevchenko, V. E.; Preobrazhenskaya, M. N. Interaction of indoles with glycosyl halides in the presence of silver oxide. *Carbohydr. Res.* **1980**, *83*, 249–261.
- (219) Chen, J. J.; Wei, Y.; Drach, J. C.; Townsend, L. B. Synthesis and antiviral evaluation of trisubstituted indole N-nucleosides as analogues of 2,5,6-trichloro-1-(beta-D-ribofuranosyl)benzimidazole (TCRB). *J. Med. Chem.* **2000**, *43*, 2449–2456.
- (220) Williams, J. D.; Chen, J. J.; Drach, J. C.; Townsend, L. B. Design, synthesis, and antiviral activity of certain 3-substituted 2,5,6-trichloroindole nucleosides. *J. Med. Chem.* **2004**, *47*, 5753–5765.
- (221) Božilović, J.; Bats, J. W.; Engels, J. W. Synthesis and structure of fluoroindole nucleosides. *Can. J. Chem.* **2007**, *85*, 283–292.
- (222) Yong, P. K.; Banerjee, A. Photochemistry of 2-nitrobenzyl enol ethers: oxidative C=C bond scission. *Org. Lett.* **2005**, *7*, 2485–2487.
- (223) Serebryany, V.; Beigelman, L. An efficient preparation of protected ribonucleosides for phosphoramidite RNA synthesis. *Tetrahedron Lett.* **2002**, *43*, 1983–1985.
- (224) Serebryany, V.; Beigelman, L. Synthesis of 2'-O-substituted ribonucleosides. *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1007–1009.
- (225) Passow, K. T.; Antczak, N. M.; Sturla, S. J.; Harki, D. A. Synthesis of 4-Cyanoindole Nucleosides, 4-Cyanoindole-2'-Deoxyribonucleoside-5'-Triphosphate (4CIN-TP), and Enzymatic Incorporation of 4CIN-TP into DNA. *Curr. Protoc. Nucleic Acid. Chem.* **2020**, *80*, e101.
- (226) Mergny, J.-L.; Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **2003**, *13*, 515–537.
- (227) Xia, T.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719–14735.
- (228) Kibbe, W. A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* **2007**, *35*, W43–6.
- (229) Parsch, J.; Engels, J. W. C-F...H-C hydrogen bonds in ribonucleic acids. *J. Am. Chem. Soc.* **2002**, *124*, 5664–5672.
- (230) Petersheim, M.; Turner, D. H. Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* **2002**, *22*, 256–263.
- (231) Schweitzer, B. A.; Kool, E. T. Hydrophobic, Non-Hydrogen-Bonding Bases and Base Pairs in DNA. *J. Am. Chem. Soc.* **1995**, *117*, 1863–1872.
- (232) Zacharias, M.; Engels, J. W. Influence of a fluorobenzene nucleobase analogue on the conformational flexibility of RNA studied by molecular dynamics simulations. *Nucleic Acids Res.* **2004**, *32*, 6304–6311.
- (233) Hebestreit, M.-L.; Schneider, M.; Lartian, H.; Betz, V.; Heinrich, M.; Lindic, M.; Choi, M. Y.; Schmitt, M. Structures, dipole moments and excited state lifetime of isolated 4-cyanoindole in its ground and lowest electronically excited singlet states. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14766–14774.
- (234) Hilaire, M. R.; Ahmed, I. A.; Lin, C.-W.; Jo, H.; DeGrado, W. F.; Gai, F. Blue fluorescent amino acid for biological spectroscopy and microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 6005–6009.
- (235) Sandin, P.; Börjesson, K.; Li, H.; Mårtensson, J.; Brown, T.; Wilhelmsson, L. M.; Albinsson, B. Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res.* **2008**, *36*, 157–167.

- (236) Jones, A. C.; Neely, R. K. 2-Aminopurine as a fluorescent probe of DNA conformation and the DNA-enzyme interface. *Q. Rev. Biophys.* **2015**, *48*, 244–279.
- (237) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One-Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies. *J. Phys. Chem.* **1996**, *100*, 5541–5553.
- (238) Heinlein, T.; Knemeyer, J.-P.; Piestert, O.; Sauer, M. Photoinduced Electron Transfer between Fluorescent Dyes and Guanosine Residues in DNA-Hairpins. *J. Phys. Chem. B* **2003**, *107*, 7957–7964.
- (239) Doose, S.; Neuweiler, H.; Sauer, M. A close look at fluorescence quenching of organic dyes by tryptophan. *ChemPhysChem* **2005**, *6*, 2277–2285.
- (240) Adams, D. M.; Brus, L.; Chidsey, C. E. D.; Creager, S.; Creutz, C.; Kagan, C. R.; Kamat, P. V.; Lieberman, M.; Lindsay, S.; Marcus, R. A.; *et al.* Charge Transfer on the Nanoscale: Current Status. *J. Phys. Chem. B* **2003**, *107*, 6668–6697.
- (241) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3. ed. (corr. at 4. print.); Springer: New York, NY, 2010.
- (242) Hirashima, S.; Han, J. H.; Tsuno, H.; Tanigaki, Y.; Park, S.; Sugiyama, H. New Size-Expanded Fluorescent Thymine Analogue: Synthesis, Characterization, and Application. *Chem. Eur. J.* [Online early access]. DOI: 10.1002/chem.201900843.
- (243) Remington, J. M.; Philip, A. M.; Hariharan, M.; Kohler, B. On the origin of multiexponential fluorescence decays from 2-aminopurine-labeled dinucleotides. *J. Chem. Phys.* **2016**, *145*, 155101.
- (244) Remington, J. M.; McCullagh, M.; Kohler, B. Molecular Dynamics Simulations of 2-Aminopurine-Labeled Dinucleoside Monophosphates Reveal Multiscale Stacking Kinetics. *J. Phys. Chem. B* **2019**, *123*, 2291–2304.
- (245) Sholokh, M.; Sharma, R.; Shin, D.; Das, R.; Zaporozhets, O. A.; Tor, Y.; Mély, Y. Conquering 2-aminopurine's deficiencies: highly emissive isomorphous guanosine surrogate faithfully monitors guanosine conformation and dynamics in DNA. *J. Am. Chem. Soc.* **2015**, *137*, 3185–3188.
- (246) Höbartner, C.; Rieder, R.; Kreutz, C.; Puffer, B.; Lang, K.; Polonskaia, A.; Serganov, A.; Micura, R. Syntheses of RNAs with up to 100 nucleotides containing site-specific 2<sup>′</sup>-methylseleno labels for use in X-ray crystallography. *J. Am. Chem. Soc.* **2005**, *127*, 12035–12045.
- (247) Lang, K.; Micura, R. The preparation of site-specifically modified riboswitch domains as an example for enzymatic ligation of chemically synthesized RNA fragments. *Nat. Protoc.* **2008**, *3*, 1457–1466.
- (248) Micura, R. Small Interfering RNAs and Their Chemical Synthesis. *Angew. Chem. Int. Ed.* **2002**, *41*, 2265.
- (249) England, T. E.; Uhlenbeck, O. C. Enzymatic oligoribonucleotide synthesis with T4 RNA ligase. *Biochemistry* **1978**, *17*, 2069–2076.
- (250) Romaniuk, E.; McLaughlin, L. W.; Neilson, T.; Romaniuk, P. J. The effect of acceptor oligoribonucleotide sequence on the T4 RNA ligase reaction. *Eur. J. Biochem.* **1982**, *125*, 639–643.
- (251) Ennifar, E.; Nikulin, A.; Tishchenko, S.; Serganov, A.; Nevskaya, N.; Garber, M.; Ehresmann, B.; Ehresmann, C.; Nikonov, S.; Dumas, P. The crystal structure of UUCG tetraloop. *J. Mol. Biol.* **2000**, *304*, 35–42.
- (252) Villa, A.; Widjajakusuma, E.; Stock, G. Molecular dynamics simulation of the structure, dynamics, and thermostability of the RNA hairpins uCACGg and cUUCGg. *J. Phys. Chem. B* **2008**, *112*, 134–142.
- (253) Muschielok, A.; Andrecka, J.; Jawhari, A.; Brückner, F.; Cramer, P.; Michaelis, J. A nano-positioning system for macromolecular structural analysis. *Nat. Methods* **2008**, *5*, 965–971.

(254) Preus, S.; Kilså, K.; Miannay, F.-A.; Albinsson, B.; Wilhelmsson, L. M. FRETmatrix: A general methodology for the simulation and analysis of FRET in nucleic acids. *Nucleic Acids Res.* **2013**, *41*, e18.

(255) Cantor, C. R.; Saenger, W. *Principles of Nucleic Acid Structure*; Springer Advanced Texts in Chemistry; Springer New York: New York, NY, 1984.

(256) Gallo, S.; Furler, M.; Sigel, R. K.O. *In vitro* Transcription and Purification of RNAs of Different Size. *Chimia* **2005**, *59*, 812–816.

(257) Gößringer, M.; Helmecke, D.; Köhler, K.; Schön, A.; Kirsebom, L. A.; Bindereif, A.; Hartmann, R. K. Enzymatic RNA Synthesis Using Bacteriophage T7 RNA Polymerase. In *Handbook of RNA Biochemistry*, Second, completely revised and enlarged edition; Hartmann, R. K., Bindereif, A., Schön, A., Westhof, E., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2014; pp 1–28.

(258) Rio, D. C. Expression and purification of active recombinant T7 RNA polymerase from *E. coli*. *Cold Spring Harb. Protoc.* [Online early access]. DOI: 10.1101/pdb.prot078527.

(259) Höbartner, C.; Wachowius, F. Chemical Synthesis of Modified RNA. In *The Chemical Biology of Nucleic Acids*; Mayer, G., Ed.; John Wiley & Sons, Ltd: Chichester, UK, 2010; pp 1–37.

(260) Wu, X.; Pitsch, S. Synthesis and pairing properties of oligoribonucleotide analogues containing a metal-binding site attached to beta-D-allofuranosyl cytosine. *Nucleic Acids Res.* **1998**, *26*, 4315–4323.

(261) Pitsch, S.; Weiss, P. A.; Jenny, L.; Stutz, A.; Wu, X. Reliable Chemical Synthesis of Oligoribonucleotides (RNA) with 2'-O-[(Triisopropylsilyl)oxy]methyl(2'-O-tom)-Protected Phosphoramidites. *Helv. Chim. Acta* **2001**, *84*, 3773–3795.

(262) Höbartner, C.; Kreutz, C.; Flecker, E.; Ottenschläger, E.; Pils, W.; Grubmayr, K.; Micura, R. The Synthesis of 2'-O-[(Triisopropylsilyl)oxy] methyl (TOM) Phosphoramidites of Methylated Ribonucleosides (m<sup>1</sup>G, m<sup>2</sup>G, m<sup>2</sup><sub>2</sub>G, m<sup>1</sup>I, m<sup>3</sup>U, m<sup>4</sup>C, m<sup>6</sup>A, m<sup>6</sup><sub>2</sub>A) for Use in Automated RNA Solid-Phase Synthesis. *Monatsh. Chem.* **2003**, *134*, 851–873.

(263) Wei, X. Coupling activators for the oligonucleotide synthesis via phosphoramidite approach. *Tetrahedron* **2013**, *69*, 3615–3637.

(264) Horn, T.; Urdea, M. S. A chemical 5'-phosphorylation of oligodeoxyribonucleotides that can be monitored by trityl cation release. *Tetrahedron Lett.* **1986**, *27*, 4705–4708.

(265) Paredes, E.; Evans, M.; Das, S. R. RNA labeling, conjugation and ligation. *Methods* **2011**, *54*, 251–259.

(266) Turunen, J. J.; Pavlova, L. V.; Hengesbach, M.; Helm, M.; Müller, S.; Hartmann, R. K.; Frilander, M. J. RNA Ligation. In *Handbook of RNA Biochemistry*, Second, completely revised and enlarged edition; Hartmann, R. K., Bindereif, A., Schön, A., Westhof, E., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2014; pp 45–88.

(267) Moore, M. J.; Query, C. C. Joining of RNAs by splinted ligation. *Methods Enzymol.* **2000**, *317*, 109–123.

(268) Seydel, J. K.; Garret, E. R. Spectrophotometric Analyses of 2-Deoxy-D-ribose in Presence of Nucleosides. *Anal. Chem.* **1965**, *37*, 271–274.

(269) Battersby, T. R.; Albalos, M.; Friesenhahn, M. J. An unusual mode of DNA duplex association: Watson-Crick interaction of all-purine deoxyribonucleic acids. *Chem. Biol.* **2007**, *14*, 525–531.

(270) Valeur, B.; Berberan-Santos, M. N. *Molecular Fluorescence: Principles and Applications*, 2nd ed.; Wiley-VCH: Weinheim, 2013.

(271) Smith, D. A.; McKenzie, G.; Jones, A. C.; Smith, T. A. Analysis of time-correlated single photon counting data: a comparative evaluation of deterministic and probabilistic approaches. *Methods Appl. Fluoresc.* **2017**, *5*, 42001.



- (272) Grinvald, A.; Steinberg, I. Z. On the analysis of fluorescence decay kinetics by the method of least-squares. *Anal. Biochem.* **1974**, *59*, 583–598.
- (273) Sillen, A.; Engelborghs, Y. The Correct Use of "Average" Fluorescence Parameters. *Photochem. Photobiol.* **1998**, *67*, 475–486.
- (274) Porrès, L.; Holland, A.; Pålsson, L.-O.; Monkman, A. P.; Kemp, C.; Beeby, A. Absolute measurements of photoluminescence quantum yields of solutions using an integrating sphere. *J. Fluoresc.* **2006**, *16*, 267–272.
- (275) Leyre, S.; Coutino-Gonzalez, E.; Joos, J. J.; Ryckaert, J.; Meuret, Y.; Poelman, D.; Smet, P. F.; Durinck, G.; Hofkens, J.; Deconinck, G.; *et al.* Absolute determination of photoluminescence quantum efficiency using an integrating sphere setup. *Rev. Sci. Instrum.* **2014**, *85*, 123115.
- (276) Würth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-Genger, U. Comparison of methods and achievable uncertainties for the relative and absolute measurement of photoluminescence quantum yields. *Anal. Chem.* **2011**, *83*, 3431–3439.
- (277) Bagshaw, C. R. *Biomolecular Kinetics: A Step-by-Step Guide*, 1st ed.; Foundations of Biochemistry and Biophysics; CRC Press: Boca Raton, FL, 2017.
- (278) Stroberg, W.; Schnell, S. On the validity and errors of the pseudo-first-order kinetics in ligand-receptor binding. *Math. Biosci.* **2017**, *287*, 3–11.
- (279) Heredia, V. V.; Thomson, J.; Nettleton, D.; Sun, S. Glucose-induced conformational changes in glucokinase mediate allosteric regulation: transient kinetic analysis. *Biochemistry* **2006**, *45*, 7553–7562.
- (280) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I.; Hearst, J. E. *Nucleic acids: Structures, properties, and functions*; University Science Books: Sausalito (CA), 2000.
- (281) Eftink, M. R.; Ghiron, C. A. Fluorescence quenching of indole and model micelle systems. *J. Phys. Chem.* **1976**, *80*, 486–493.
- (282) Balzani, V.; Juris, A.; Ceroni, P. *Photochemistry and photophysics: Concepts, research, applications*; Wiley-VCH: Weinheim, 2014.
- (283) Dale, R. E.; Eisinger, J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. *Biopolymers* **1974**, *13*, 1573–1605.
- (284) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29*, 2176–2179.
- (285) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (286) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1327.
- (287) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (288) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (289) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (290) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (291) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

- (292) Retegan, M. <https://gist.github.com/mretegan/5501553>.
- (293) Berezin, M. Y.; Kao, J.; Achilefu, S. pH-dependent optical properties of synthetic fluorescent imidazoles. *Chem. Eur. J.* **2009**, *15*, 3560–3566.
- (294) Britton, H. T. S.; Robinson, R. A. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* **1931**, 1456–1462.