Development of Fluorescent Nucleobase Analogues

- Intrinsically labelled nucleic acids for molecular binding investigations

MATTIAS BOOD

UNIVERSITY OF GOTHENBURG

Department of Chemistry and Molecular Biology

University of Gothenburg

2019

DOCTORAL THESIS

Submitted for fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

Development of Fluorescent Nucleobase Analogues
- Intrinsically labelled nucleic acids for molecular binding investigations

ISBN 978-91-7833-502-2 (PRINT)
ISBN 978-91-7833-503-9 (PDF)

Printed in Gothenburg, Sweden 2019
Printed by Chalmers Reproservice

“This too shall pass”

Persian poets

Abstract

This thesis focuses on the design, synthesis and utilization of fluorescent nucleobase analogues (FBAs). FBAs are an important class of compounds, used in the research of nucleic acids. The class of canonical FBAs, *i.e.* like the natural nucleobases, are of special interest as they can replace the natural nucleobases without significantly perturbing the overall structure and biological function of the nucleic acid. The overarching goal of the project was to establish a molecular binding interaction assay based on novel FBAs, to study ligand binding to oligonucleotides.

This thesis starts with explaining the design rationale behind the class of quadra- and penta-cyclic adenine analogues, followed by the developed synthetic methods to such constructs. The developed synthetic scheme was used to prepare a library of over 50 novel multicyclic adenine analogues.

One of the brightest molecules, pA, was incorporated and characterized inside DNA and was found to not perturb the overall structure of duplex DNA significantly. Moreover, pA was characterized as one of the brightest adenine analogues in DNA and RNA at the time of publishing. Follow-up studies revealed that pA can be detected *via* two-photon spectroscopy at a ratio of signal to background as low as five to one, meaning that our developed FBAs are approaching super resolution imaging applications.

Another remarkable compound that was identified from the early screening study was 2CNqA, which just recently turned out to be the brightest FBA in DNA and RNA to date. The interbase FRET (Förster resonance energy transfer) properties were studied of 2CNqA in both DNA and RNA, and the probe accurately reports FRET of at least 1.5 turns of DNA, making it suitable to study changes over short DNA and RNA.

The thesis is concluded with the synthesis, incorporation and characterization of the FRET pair tC^{O}-tC_{nitro} in RNA where they were used to monitor changes from A- to Z-form RNA. Furthermore, the FRET pair was then used to study the antibiotic class of aminoglycosides binding to RNA, faithfully reporting on their relative binding affinity of a pre-microRNA construct.

Keywords: Fluorescent nucleobase analogue, FRET, surface plasmon resonance, isothermal titration calorimetry, DNA, RNA, pre-microRNA, aminoglycoside, binding interaction.

List of publications

I. Second-Generation Fluorescent Quadracyclic Adenine Analogues: Environment-Responsive Probes with Enhanced Brightness
Dumat, B.†, Bood, M.†, Wranne, M. S., Lawson, C. P., Larsen, A. F., Preus, S., Streling, J., Gradén, H., Wellner, Erik., Grøtli, M., Wilhelmsson, L. M.
Chemistry – A European Journal, **2015**, *21*, 4039-4048.

II. Pentacyclic adenine: a versatile and exceptionally bright fluorescent DNA base analogue
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., Höök, F., Brown, T., Kim, B. H., Dahlén, A., Wilhelmsson, L. M., Grøtli, M.
Chemical Science, **2018**, *9*, 3494-3502.

III. Adenine analogue 2CNqA – the brightest fluorescent base analogue inside DNA and RNA
Wypijewska, A., Füchtbauer, A. F., Bood, M., Nilsson, J. R., Wranne, M. S., Pfeiffer, P., Sarangamath, S., Rajan, E.J.S., V., El-Sagheer, A. H., Dahlén, A., Brown, T., Grøtli, M., Wilhelmsson, L. M.
Manuscript in preparation

IV. Interbase FRET in RNA – From A to Z
Füchtbauer, A. F.†, Wranne, M. S.†, Bood, M., Weis, E., Pfeiffer, P., Nilsson, J. R., Dahlén, A., Grøtli, M., Wilhelmsson, L. M.
Manuscript submitted to Nucleic Acids Research, under revision.

V. RNA Interbase FRET Binding Interaction Assay
Bood, M., Wypijewska, A., Nilsson, J., Edfeldt, F., Dahlén, A., Wilhelmsson, L. M., Grøtli, M.
Manuscript in preparation

† Equally contributing authors

Publications not included in the thesis

- Development of bright fluorescent quadracyclic adenine analogues: TDDFT-calculation supported rational design
  Foller Larsen, A., Dumat, B., Wranne, M. S., Lawson, C. P., Preus, S., Bood, M., Gradén, H., Grøtli, M., Wilhelmsson, L. M.
  Scientific Reports, **2015**, *5*, 12653.

- Toward Complete Sequence Flexibility of Nucleic Acid Base Analogue FRET
  Wranne, M. S., Füchtbauer, A. F., Dumat, B., Bood, M., El-Sagheer, A. H., Brown, T., Gradén, H., Grøtli, M., Wilhelmsson, L. M.
  Journal of the American Chemical Society, **2017**, *139*, 9271-9280.

- Fluorescent nucleobase analogues for base–base FRET in nucleic acids: synthesis, photophysics and applications
  Bood, M.†, Sarangamath, S.†, Wranne, M. S., Grøtli, M., Wilhelmsson, L. M.
  Beilstein Journal of Organic Chemistry, **2018**, *14*, 114-129.

- Pulse-shaped two-photon excitation of a fluorescent base analogue approaches single-molecule sensitivity
  Fisher R. S., Nobis, D., Füchtbauer, A. F., Bood, M., Grøtli, M., Wilhelmsson, L. M., Jones, A. C., Magennis, S. W.
  Physical Chemistry Chemical Physics, **2018**, *20*, 28487-28498.

Contribution report

Paper I. Planned and performed the synthesis. Wrote the experimental section of the synthesis. Wrote the synthetic section of the article together with A.F.L. Proofread the manuscript.

Paper II. Planned and performed the synthesis together with A.F.F. Synthesised and purified the oligonucleotides together with A.F.F. Performed fluorescence measurements together with A.F.F., M.S.W, J. J.R. and S.S. Wrote the manuscript.

Paper III. Contributed to the synthesis of the DNA building blocks together with A.F.F. Planned and performed synthesis of the RNA building blocks. Synthesised and purified the oligonucleotides. Performed fluorescent measurements together with A.W.d.N., A.F.F., M.S.W., P.F., J.N., V.E.J.S.R, and S.S. Proofread the manuscript.

Paper IV. Contributed to the synthesis together with A.F.F. and supervised synthesis performed by E.W. Synthesised and purified the oligonucleotides. Proofread the manuscript.

Paper V. Designed, synthesised and purified the oligonucleotides. Performed SPR measurements together with F.E. Performed ITC measurements. Interpreted ITC data together with J.N. Performed FRET measurements. Interpreted FRET data together with L.M.W. Wrote the manuscript.

Abbreviations

2-AP	2-aminopurine
2CNqA	2-cyano quadracyclic adenine
A	adenosine
ABI	applied biosystems
AcOH	acetic acid
ASO	antisense oligonucleotide
Boc	*tert*-butyloxycarbonyl
bp	base pair
C	cytosine
CE	2-cyanothyl
CEP-Cl	chloro-(2-cyanoethoxy)diisopropylaminophosphine
DABCO	1,4-diazabicyclo[2.2.2]octane
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DEA	diethylamine
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DMSO	dimethyl sulfoxide
DMTr	dimethoxytrityl
dsDNA	double-stranded DNA
EDTA	ethylenediaminetetraacetic acid
EtI	ethyl iodide
EtOAc	ethyl acetate
EtOH	ethanol
FBA	fluorescent nucleobase analogue
FID	fluorescent indicator displacement

FRET	förster resonance energy transfer
$FRET_{eff}$	förster resonance energy transfer efficiency
G	guanosine
h	hour
HBPin	4,4,5,5-tetramethyl-1,3,2-dioxaborolane
HMDS	hexamethyldisilazane
IR	infrared
ITC	isothermal titration calorimetry
LC-MS	liquid chromatography–mass spectrometry
LC-TOF-MS	liquid chromatography–time-of-flight–mass spectrometry
LiHMDS	lithium bis(trimethylsilyl)amide
MB	molecular beacon
MeCN	acetonitrile
MeOH	methanol
min	minute
miR	microRNA
mRNA	messenger RNA
MST	microscale thermophoresis
MW	microwave
NaHMDS	sodium bis(trimethylsilyl)amide
nm	nanometre
NMR	nuclear magnetic resonance
nt	nucleotide
OP10	ÄKTA OligoPilot 10
pA	pentacyclic adenine
pre-FBA	preliminary fluorescent nucleobase analogue
pre-miR	precursor-microRNA

pri-miR	primary-microRNA
qA	quadracyclic adenine
RP-HPLC	reverse phase high performance liquid chromatography
RT	room temperature
s	second
SAR	structure activity relationship
SPR	surface plasmon resonance
SPS	solid-phase synthesis
ssDNA	single-stranded DNA
T	thymidine
TBAF	tetra-butylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl ether
TBDMS-Cl	*tert*-butyldimethylsilyl chloride
TBDMSOM	*tert*-butyldimethylsilyloxymethyl
TBDMSOTf	*tert*-butyldimethylsilyl trifluoromethanesulfonate
TBDPS-Cl	*tert*-butyl(chloro)diphenylsilane
tC	tricyclic cytosine
TCSPC	time-correlated single photon counting
TEAA	triethylammonium acetate
TEAB	triethylammonium bicarbonate
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TMS	trimethylsilyl
TMS-Cl	chlorotrimethylsilane
TMS-OTf	trimethylsilyl trifluoromethanesulfonate
U	uracil

Contents

1. General introduction and aims of the thesis

During the past two decades a new class of regulators, microRNA (miR), have been identified to play a fundamental role in the regulation of cell development and function.[1] Since the discovery of miRs in *C. elegans* in 1993,[2,3] more than 1900 genes coding for over 2600 miRs have been identified in humans.[4] Their biogenesis is well characterized and the canonical pathway (Figure 1) can be briefly described as; genes coding for primary-miR (pri-miR, over 1 kb) are expressed and then processed inside the nucleus to precursor-miR (pre-miR, 70–90 nucleotides, nt), exported to the cytoplasm and further processed to mature miR (20–25 nt). The mature miR can in turn bind to and silence messenger RNA (mRNA), resulting in lowered protein expression. One third of the entire proteome is estimated to be regulated by these types of processes.[5]

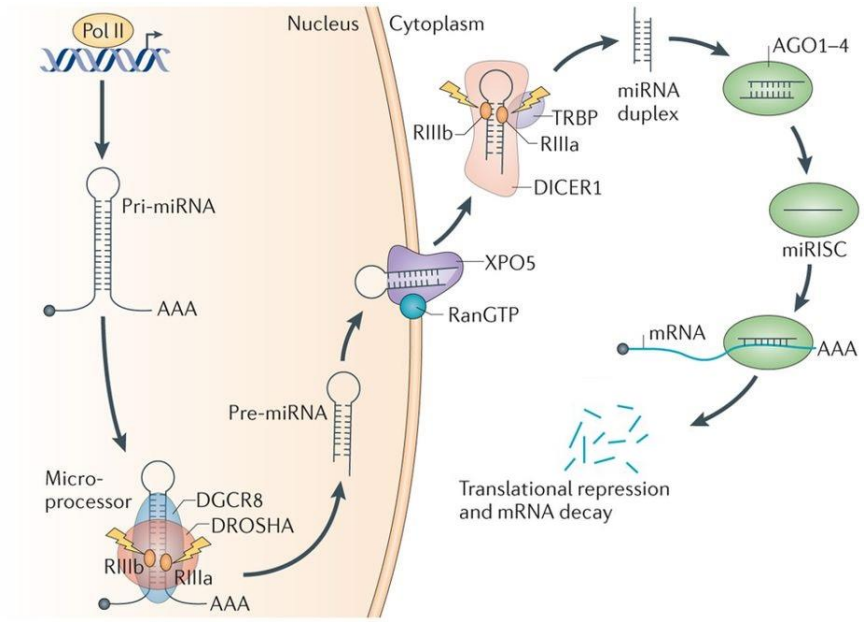


Figure 1. Simplified scheme of miR biogenesis and function. Adapted with permission.[6]

As with all cellular processes, the biogenesis of miRs is not without fault and the dysregulation of miR levels has been associated with a number of disease states,[7] including all forms of cancer.[8,9] While miR are now well validated therapeutic targets, it is difficult to rationally design selective small molecule miR modulators, as structural motifs are shared amongst

several miRs.[10] Antisense oligonucleotides (ASOs) that target miRs, have been explored therapeutically with miravirsen in early clinical trials.[11] Nevertheless, ASOs that target mRNA in general has been in development for over three decades and even as several concerns regarding stability, uptake and delivery has been solved, issues regarding toxicity still exists.[12]

In the drug discovery process, assays are employed to discover lead molecules and to build structure-activity relationships (SAR) that can guide lead optimization and identify potential toxic or off-target effects.[13] With the identification of new therapeutic targets, the development of novel assays are also required. During the past decade we have seen the development of several new assays for the identification of small molecule miR binders, unfortunately with limited success.[14,15]

The overall goal of my PhD project was to develop an *in vitro* assay suitable for monitoring small molecule binders to pre-miRNAs. This was to be achieved through the use of internally placed fluorescent RNA base analogues as Förster resonance energy transfer (FRET) pairs. In order to realize this goal, the following four milestones were defined:

- Develop new fluorescent nucleobase analogues (FBAs) with desirable photophysical properties.

- Synthesise phosphoramidite building blocks of FBAs amenable for solid-phase synthesis (SPS) of DNA and characterize the FBAs in an oligonucleotide context.

- Develop the synthesis of phosphoramidite building blocks of FBAs for incorporation into RNA.

- Develop a novel pre-miR binding interaction assay based on suitable RNA FRET FBA pairs.

2. Background

The primary objective of this chapter is to introduce the theoretical background of this thesis and to provide a broad overview of nucleic acids, spectroscopy and binding interaction assays.

2.1 Nucleic acids

In eukaryotic life, DNA carries the genetic information required to produce the entire organism. DNA can be transcribed to RNA which acts as the transcript from which proteins are translated. Lately, RNA has turned out to be more complex than previously thought, being both functional and taking part in several other important regulatory pathways.[16]

2.1.1 Structure and composition of oligonucleotides

The monomeric units of DNA and RNA consists of a heterocyclic nucleobase linked *via* a C-N glycosidic bond to a pentose monosaccharide equipped with a 5'-OH phosphate group. If the pentose monosaccharide is ribose then the oligomer formed from linked monomers is defined as RNA, but if the pentose monosaccharide is deoxyribose then the oligomer is defined as DNA. Nucleosides are divided into two main categories: purines; constituted by adenosine (A) and guanosine (G) which can base-pair to the pyrimidines; thymidine (T) or uracil (U) and cytidine (C) respectively (Figure 2a). Thymine occurs in DNA and uracil in RNA.

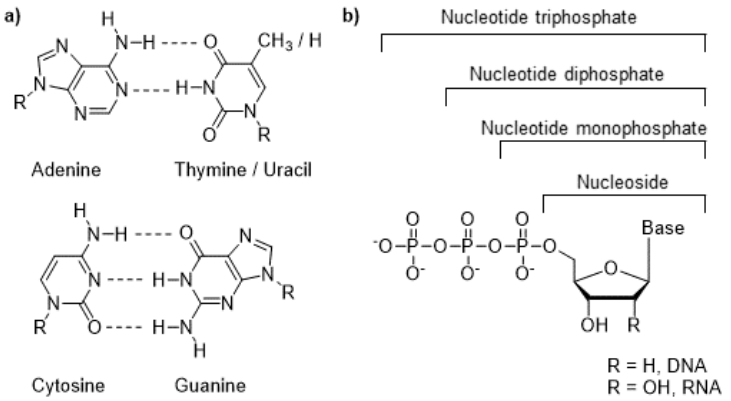


Figure 2. a) The four nucleobases of DNA and uracil of RNA. R = deoxyribose or ribose. b) The nucleobase is constituted of either pyrimidines or purines; a nucleoside is the nucleobase with an attached (deoxy)ribose and a nucleotide furthermore carries from one, up to three 5'-OH phosphate groups.

If the monomeric nucleoside contains a 5'-OH phosphate group, it is termed nucleotide. Linked nucleotides *via* 5' to 3' phosphate bonds create oligonucleotides. DNA is commonly found in the B-form right-handed double helical structure and contains approximately 10 base pairs (bp) per turn of the double helix. While several other forms of double helical DNA exist,[17,18] the A- and Z-form DNA are also considered biologically active.[19–21]

The extra 2'-OH of ribonucleosides makes RNA more susceptible to hydrolysis than DNA, as the 2'-OH can attack the 3'-OH phosphate group.[22] The extra hydroxyl on the pentose ring shifts the pentose ring conformation from a C2'-endo found in B-DNA to C3'-endo found in right-handed double helical A-RNA and A-DNA.[22]

More noteworthy, RNA does not normally carry a complementary strand. This characteristic of RNA allows it to bend and self-base-pair in a unique fashion, creating a completely different set of secondary and tertiary structures, some of which are highlighted in Figure 3.[23] In this sense, RNA is much more dynamic than DNA and a single type of RNA ensemble may sometimes be composed of several RNA conformations.

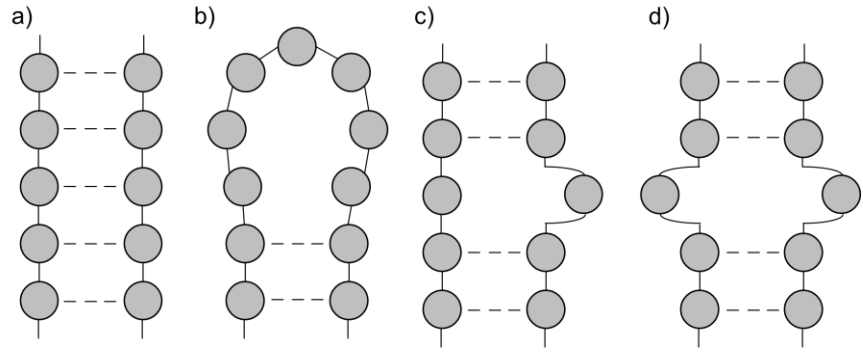


Figure 3. Representation of various structural motifs in RNA. a) Stem or double helix. b) Hairpin loop. c) Bulge. d) Internal loop or mismatch.

2.1.2 Targeting RNA with antisense oligonucleotides

While the majority of commercial pharmaceuticals target proteins, other modalities are now being explored for pharmaceutical opportunities. Of the ~20,000 protein coding genes, approximately 3,000 are considered to be disease-related, where only ~700 of these have been therapeutically accessed.[24,25] The main reason why more proteins have not been therapeutically accessed is partly due to the fact that they are difficult to

target, with some even labelled "undruggable".[24] However, in theory, it should be possible to target proteins deemed undruggable by interfering with their biogenesis.

ASOs can interfere with the biogenesis of undruggable proteins by influencing the translation of RNA to proteins. ASOs binds to mRNA and inhibits the translation process to proteins.[26] Although the development of oligonucleotide-based drugs has been ongoing for more than three decades, only eight drugs have entered the market as of mid 2019.[27–29] The low number of drugs that have made it to the market are due to issues regarding poor *in vivo* biological activity, toxic off-target effects as well as poor absorption and distribution.[30]

2.1.2 Targeting RNA with small molecules

To date, the antibiotic linezolid, is the only synthetic small molecule drug on the market that specifically targets RNA (Linezolid, Figure 4).[31] Other small RNA binding molecules discovered includes the antibiotic family of; aminoglycosides (Neomycin, Figure 4), macrolides (Erythromycin, Figure 4), tetracyclines (Tetracycline, Figure 4) and oxazolidinones (Pleuromutilin, Figure 4).[32]

The majority of RNA binding compounds described to date are either intercalating, highly stacking/lipophilic and/or highly basic species resulting in positively charged compounds in physiological conditions. Such characteristics often leads to undesirable properties such as not being orally bioavailable or non-specific interactions between the small molecule and the RNA of interest.[33] There currently seems to be a consensus in the field that RNA should be druggable using small molecules, as long as careful consideration is taken with regards to the choice of target, screening techniques to identify hits and identification of RNA motifs.[34]

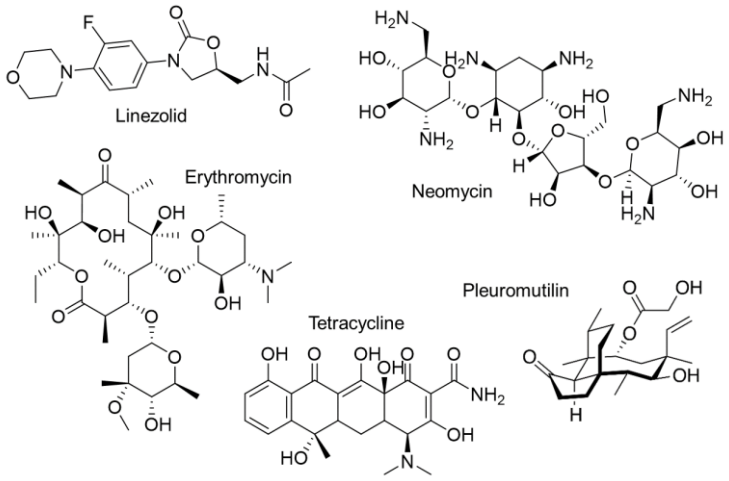


Figure 4. The only designed small molecule targeting RNA (Linezolid), and four other discovered RNA binding small molecules, all of which are antibiotics.

2.2 Spectroscopy

Spectroscopy can be used in a multitude of experiments to measure various optical parameters. The focus in this thesis is mainly with the absorption and emission of light from small molecular labels/probes and oligonucleotides.

2.2.1 Absorption and emission of light

The absorption of light by molecules can be measured with a spectrophotometer. The instrument records how much of the incident light (I_O) is passed through the sample (I) at each wavelength. The absorption is calculated for each wavelength using equation 1.

$$A = log\left(\frac{I_0}{I}\right) \qquad [1]$$

The absorption can be related to the concentration of the molecule under study using the Beer-Lambert law, where c is concentration (mol dm^{-3}), ε is molar absorptivity (absorbance) and l is the path length (cm, Equation 2).

$$A = c \cdot \varepsilon \cdot l \qquad [2]$$

Some molecules, when absorbing light, become excited from the ground state, S_0, to the first excited state, S_1, or higher states (Figure 5). In most cases the molecule relaxes to the ground state *via* non-radiative transitions

(rate determined by k_{nr}; Figure 5) such as internal conversion followed by vibrational relaxation. However, some molecules can relax to the S_0 ground state from the excited S_1 state, *via* the emission of a photon (rate determined by k_r; Figure 5). As some energy is lost to the surroundings mainly *via* vibrational relaxation in this process, the emitted light is always of longer wavelength than the absorbed light.

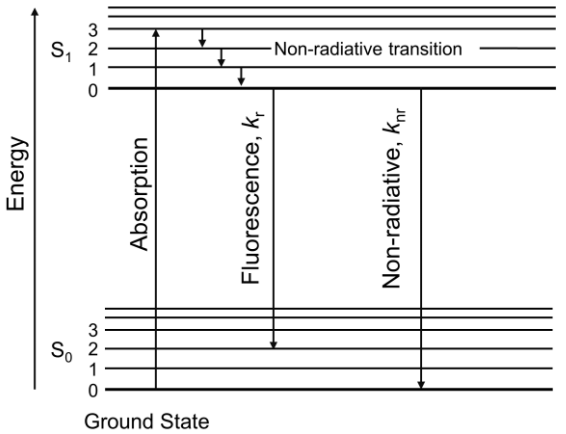


Figure 5. Jablonski diagram showing the S_0 ground state and the S_1 excited state and main photophysical processes involved in absorption and emission.

The fluorescent lifetime (τ) describes the average time a fluorophore spends in the excited states. It is defined as the inverse sum of all processes that decreases the excited state, where k_r is the rate at which the molecule relaxes to the ground state while emitting a photon and k_{nr} is the rate at which the molecules relaxes to the ground state without emitting a photon (Equation 3). The fluorescence lifetime can be measured *via* time-correlated single photon counting (TCSPC).

$$\tau = \frac{1}{k_r + k_{nr}} \qquad [3]$$

The fluorescence quantum yield (Φ_F) is defined as the ratio between the number of photons a molecule emits and the number of photons it absorbs, which is equal to the ratio between the rate at which relaxation of the molecule relaxes *via* emission of a photon and the total amount of rates that depopulate the excited state (Equation 4).

$$\Phi_F = \frac{photons\ emitted}{photons\ absorbed} = \frac{k_r}{k_r + k_{nr}} \qquad [4]$$

2.2.2 Förster resonance energy transfer

FRET is a process where molecules in an excited state can donate their energy to a proximal molecule in the ground state that acts as an acceptor. The following criteria must be met in order for FRET to occur: (I) the emission from the donor needs to overlap with the absorption band of the acceptor (Figure 6a); (II) the distance between the donor and acceptor needs to be close enough in space (Figure 6b-c); and (III) the orientation of the transition dipole moments (molecular antennae) of the donor and acceptor must not be perpendicular to each other (Figure 6d).

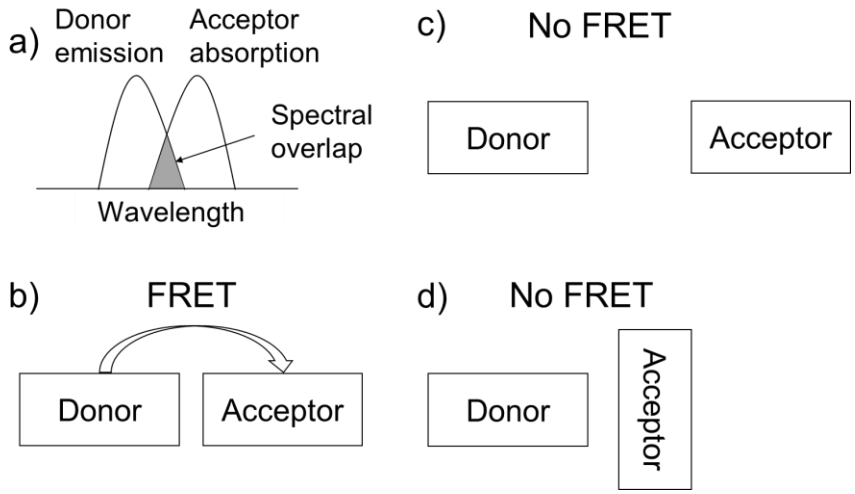


Figure 6. Distance and orientation dependence of FRET between a donor and acceptor molecule. a) Spectral overlap between donor and acceptor. b) FRET occurs due to correct distance between donor and acceptor. c) FRET cannot occur due to too long distance between donor and acceptor. d) FRET does not occur due to incorrect orientation of donor and acceptor.

The efficiency of the FRET process ($FRET_{eff}$) can be measured by both steady-state fluorescence and TCSPC (Equation 5).

$$\mathrm{FRET}_{eff} = 1 - \frac{I_{DA}}{I_D} = 1 - \frac{\tau_{DA}}{\tau_D} \qquad [5]$$

Here D is donor, A is acceptor, DA refers to donor and acceptor present, I is intensity and τ is for average lifetime.

2.3 Binding interaction assays

An assay can be defined as an investigative procedure for qualitatively or quantitatively measuring the presence, amount or functional activity of a target entity. This thesis focuses on assays that look at the molecular binding of small molecules to oligonucleotides, in particular that of RNA.[35]

2.3.1 Label-free assays

Label-free assays measure the interaction of molecules without the use of reporter tags such as radioisotopes or fluorescent dyes. A number of label-free binding interaction assays exist such as monitoring small molecules interactions with proteins *via* nuclear magnetic resonance (NMR) techniques,[36] affinity-chromatography coupled with mass detection,[37] and various spectroscopy-based methods.[38] Below follows a brief description of the assays used in this thesis, which will be described more thoroughly in section 3.4.1 and 3.4.2.

2.3.1.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC), is a technique where a change in temperature is measured when a ligand is added to a sample of interest. ITC has found extensive use in the study of ligand-protein and ligand-DNA interactions.[39–41] Only a few studies have employed ITC for studying ligand-RNA interactions.[42] Notable examples include tetracycline binding to riboswitches,[43] aminoglycosides binding to 16 S ribosomal RNA,[44] and aminoglycosides binding to the HIV-1 RNA dimerization initiation site.[45]

2.3.1.2 Surface plasmon resonance

Surface plasmon resonance (SPR) is a physical occurrence that happens when polarized light hits a thin metal film at the interface of media with different refractive indices and is the basis for several optical biosensors.[46] SPR applied to the measurement of small molecules interacting with macromolecules can be considered a label-free technique. SPR techniques have been used to study ligand-RNA interactions.[47] Early studies focused on aminoglycosides binding to various RNA targets.[48] SPR was also used to guide the design of bi-functional ligands for targeting the Rev response element, a highly structured 350 nt segment of HIV-1 mRNA.[49] More recently SPR was used to validate hits generated from a fluorescent indicator displacement (FID) assay for ligands binding pre-miR-29a.[50]

2.3.2 Labelled assays

Ideally, an assay should interfere as little as possible with the system under investigation. As previously mentioned, label-free assays are designed so that a physical parameter is measured, for example conductivity, mass or thermodynamic changes, they interfere minimally with the system investigated. However, in many cases label-free assays cannot be performed under biologically relevant conditions or lack the required throughput to allow screening of large libraries. A combination of orthogonal assay techniques is therefore required to validate both ligand-target interaction and ensure biological relevance.[13] In efforts to provide an overview of the assay which we set out to develop in my PhD project, the following sections describes a few different techniques.

2.3.2.1 Microscale thermophoresis

Microscale thermophoresis (MST), is a technique used to measure binding interactions in which it is required that the sample is fluorescently labelled. By either intrinsically labelling, or covalently attaching a fluorophore to a sample of interest, the mobility of the sample can be measured in a capillary tube with an induced temperature gradient (Figure 9).[51] In comparison to SPR, MST requires no immobilization, uses a small sample size and can be used with cell lysate as the sample media, providing biologically relevant conditions.[51] MST mainly provides binding constant data of ligand-complex interactions, but other parameters such as binding stoichiometry, thermodynamic parameters, such as enthalpy of the process (ΔH) and binding kinetics can be derived. The main drawback with MST is the limitation of which fluorescent reporter label can be used. Each instrument is built with a specific excitation laser that is combined with the infrared (IR) laser generating the temperature gradient, therefore making the change of the excitation wavelength rather laborious. The MST instruments are built either to excite tryptophan residues in proteins (excitation wavelength of 280 nm) or with an excitation source in the visible region ranging from 480 nm to 720 nm, therefore excluding those fluorescent dyes that exhibit an excitation maximum in between this range. MST has recently been used to study neomycin binding to the Rev responsive element of HIV-1 mRNA.[52]

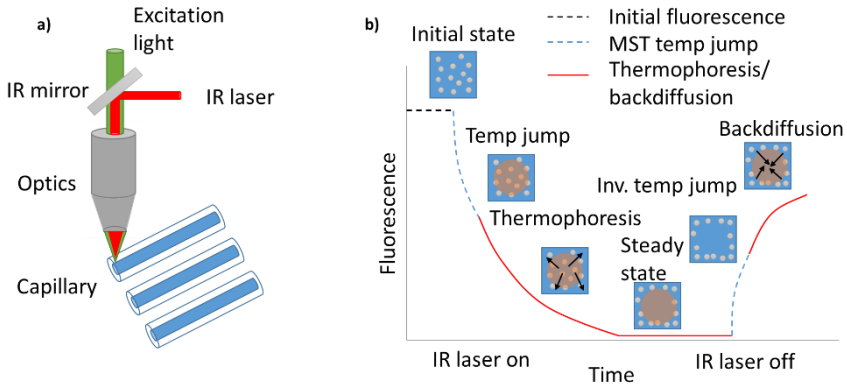


Figure 7. The basics of a microscale thermophoresis (MST) experiment.

2.3.2.2 Fluorescent indicator displacement

FID assays work by first binding a fluorescent substance of medium affinity to the desired sample (Figure 10).[53] Then the ligand under analysis is added. If the fluorescent substance is displaced by the ligand, change in emission is observed as the microenvironment around the probe, especially the polarity, changes. This signal change can be used to provide binding affinity information relative to the displaced probe. This type of system has been used in a high throughput format to study the interaction of small molecules with proteins,[54] DNA,[55] and RNA.[56]

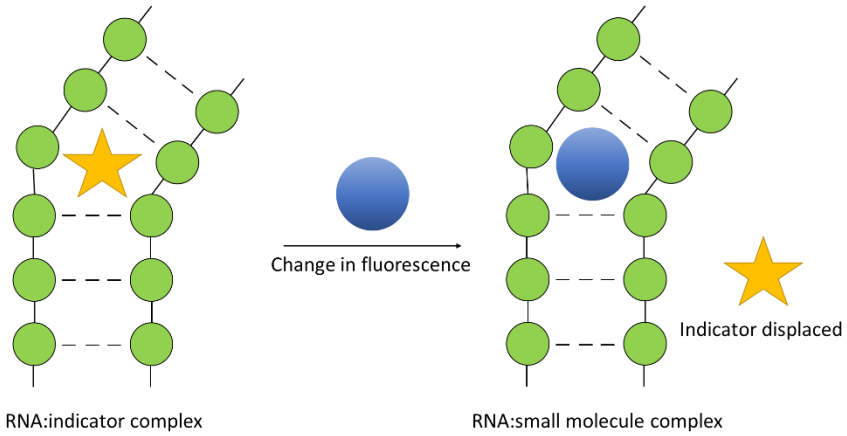


Figure 8. The principle of fluorescent indicator displacement (FID) assays.

2.3.2.3 Molecular beacon

Molecular beacon (MB) fluorescence assay is a versatile assay where a short fluorescently labelled oligonucleotide (~25 nt) reports on the specific target through hybridization. The 5'- and 3'-ends of the oligonucleotide are labelled with a FRET pair, which, in an unbound state are in close proximity resulting is high $FRET_{eff}$ between the probes. As the oligonucleotide binds to its target through hybridization, the distance between the FRET pair increases, causing the $FRET_{eff}$ to be lower as a result.[57] Recently, molecular beacons have been used to image RNA in live cells,[58] and in screening the inhibition of miR maturation by small molecules.[59]

2.3.3 Internucleobase labelled assays

New modalities in drug discovery refers to the next generation of peptides, peptidomimetics, oligonucleotide-based molecules and novel hit finding technologies.[60,61] One such example of a modality is the modulation of miR levels for therapeutic use. Even though great progress has been achieved in the development of miR modulators, small molecules that modulate miR are yet to reach clinical trials. This could be due to the fact that RNA itself is difficult to target,[34] but possibly also due to the methods used to identify and evaluate the current small molecules as modulators of miR maturation. It is possible that since small molecule libraries are developed around proteins, no good RNA binding compounds are included.[14] Indeed, the identified small molecule miR binders are all either highly lipophilic and/or highly charged molecules, and the few that are not, are potentially interfering with the maturation process by other means such as Dicer inhibition.[15] There currently is a lack of understanding regarding which structural elements in RNA can be selectively targeted and by what type of compounds,[62] something that interbase labelled assays could potentially shed light on.

The concept of using FBAs to probe small molecule interactions to oligonucleotides has previously been exemplified with 2-aminopurine (2-AP), which was used to probe small molecule binding to HIV-1 TAR RNA.[63] Recently, a spin-labelled fluorescent probe was used in a structure guided fluorescence labelling approach, which revealed a two-step binding mechanism of neomycin to its RNA aptamer.[64] The fluorescent probe, however, altered the binding affinity of neomycin significantly, thus better probes are required.

2.4 Fluorescent nucleobase analogues

FBAs are fluorescent molecules that can be divided into two categories: (I) canonical, meaning that they are of similar size, shape and hydrogen bonding properties to mimic the native nucleobases; and (II) non-canonical, meaning that no limit is put on the design of the molecule other than the function and photophysical properties of the probe.[65] The central theme of this thesis is canonical FBAs and section 2.4.1 will briefly introduce these fluorescent entities. For a more comprehensive overview of both canonical and non-canonical FBAs see Wilhelmsson and Tor,[66] and the recent review articles from Tanpure *et al*.,[67] and Xu *et al*.[65]

2.4.1 Overview of canonical FBAs

FBAs are powerful tools for studying structure and dynamics of nucleic acids as they can be placed close to the site of interest without perturbing the biological function of the nucleic acid. Depending on the intended application, FBAs, in general, should:

- Retain the hydrogen bonding properties of the native nucleobase they are replacing.
- Be small enough not to impact tertiary structure formation.
- Have a high brightness for detection.
- Be stable towards photodegradation.
- Absorb light outside the absorption band of the natural nucleobases that is preferably significantly red-shifted.

Fifty years ago, the foundation for fluorescent nucleobase analogues was laid when Stryer *et al*. published the fluorescence properties of three substances. One of which was 2-AP (Figure 9), considered the golden standard in the field of FBAs.[68] Since then a number of FBAs have been developed and some of the most notable examples include; 8-vdA,[69] the tricyclic cytosines tC and tC^{O},[70,71] the pteridines (3-MI, 6-MI, 6MAP and DMAP),[72,73] and the emissive isomorphic RNA alphabet (^{th}A, ^{th}C, ^{th}G, and ^{th}U).[74]

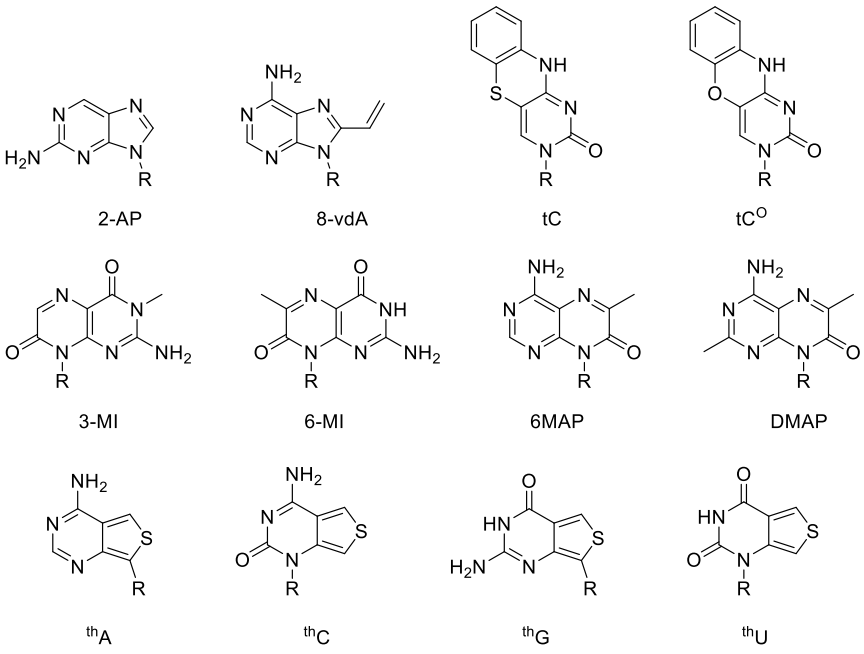


Figure 9. Overview of fluorescent nucleobase analogues. R = (deoxy)ribose.

FBAs have been used to study a number of processes where nucleic acids are involved. Recent examples include oxidative DNA alkylation repair,[75] the effect of mercury on DNA metabolism,[76] DNA duplex formation,[77] and the use of FBAs for ultra-sensitive oligonucleotide detection.[78]

The biggest challenges for the design and synthesis of FBAs includes the red shifting of the absorption for improved live cell imaging. Importantly also, making FBAs bright enough for single-molecule analyses and super-resolution imaging. All while still keeping the FBA small enough to not adversely affect the biological properties of the studied system.[65]

2.4.2 FRET FBA pairs

A great number of FBAs have been developed and used in numerous applications. However, in order to gain valuable structural information, such as distance and orientation, more than one label is required. FRET FBA pairs (henceforth FRET pairs) are an example of a spectroscopic ruler.[79]

Interbase FRET pairs are nucleoside pairs consisting of a fluorescent donor and a fluorescent acceptor. The donor absorbs incoming light,

becomes excited and can then, *via* FRET, donate its energy to a neighbouring acceptor molecule. The acceptor molecule either emits a photon or the energy is lost in a non-radiative pathway and the molecule returns to the ground state.

Before the commencement of this project, only one interbase FRET pair had been developed, namely the tC^{O}-tC_{nitro} as deoxyribonucleosides.[80] Since then, only one other interbase FRET pair has been developed by Sugiyama and co-workers in 2017, thdG-tC in DNA.[81] During this project our group has developed several new interbase FRET pairs including qAN1-qA_{nitro} in DNA,[82] pA-qA_{nitro} in DNA (**Paper II**), qAN4-qA_{nitro} in DNA (manuscript in preparation), 2-cyano-qA (2CNqA)-qA_{nitro} and 2CNqA-tC_{nitro} in RNA (**Paper III**), and tC^{O}-tC_{nitro} in RNA (**Paper IV**).

3. Methodology

This chapter gives a brief overview of the main methods in the synthesis of nucleosides, the standard approach to synthesise oligonucleotides and the techniques used in this thesis for measuring the binding interaction of small molecules and oligonucleotides.

3.1 Synthetic strategies

Several different strategies to synthesise FBAs can be employed. In this thesis, they are categorized into either divergent or convergent syntheses, depending on the linearity of the synthesis performed. A normal linear synthesis is done when the intermediates are moved towards the desired product one step at a time (Figure 10a).[83] A convergent synthetic approach focuses on synthesizing fragments of approximately the same size and complexity that are connected towards the end of the synthetic scheme. The convergent approach is more common for the synthesis of larger amounts of material, as the number of reduced linear steps leads to an overall higher yield (Figure 10b). In contrast, a divergent synthetic strategy aims at generating a common intermediate from which several different products can be obtained, which allows for library synthesis around a common scaffold (Figure 10c).[84]

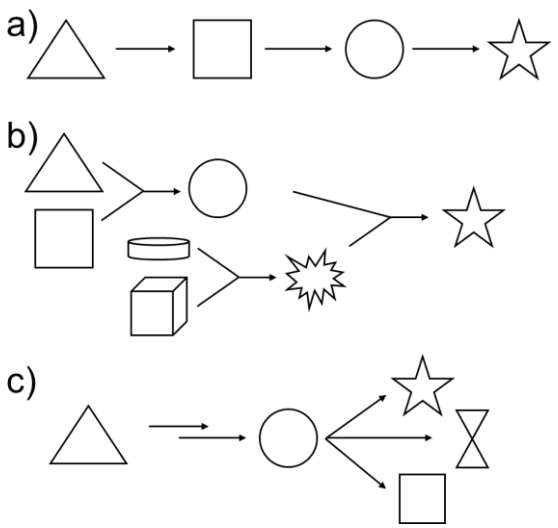
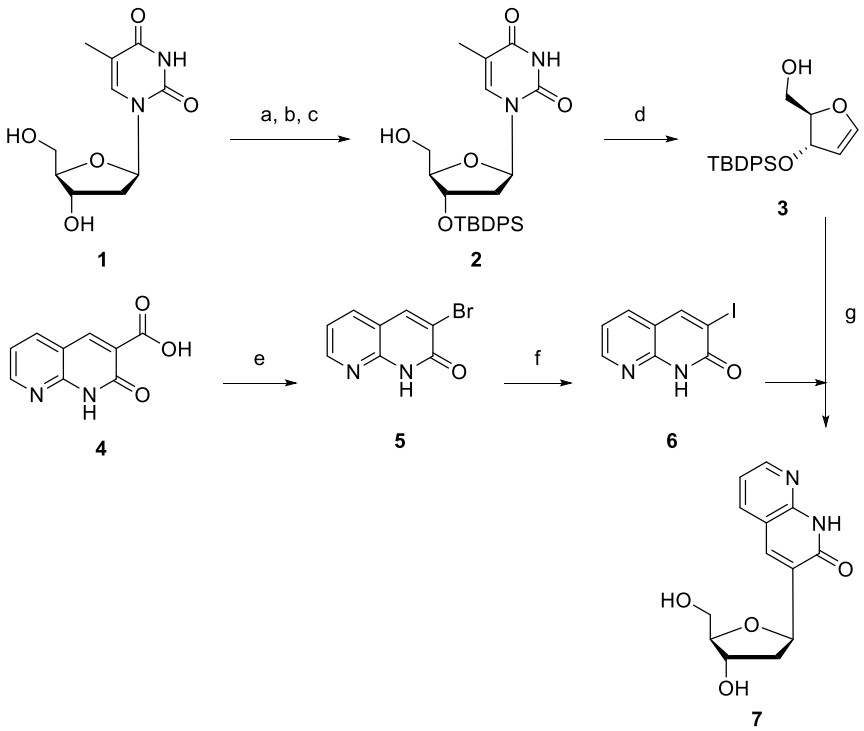


Figure 10. Different synthetic strategies. a) Linear synthesis, carrying the same building block continuously forward. b) Convergent synthesis, connecting similarly sized fragments together. c) Divergent synthesis of several products *via* a common intermediate.

3.1.1 Convergent synthesis

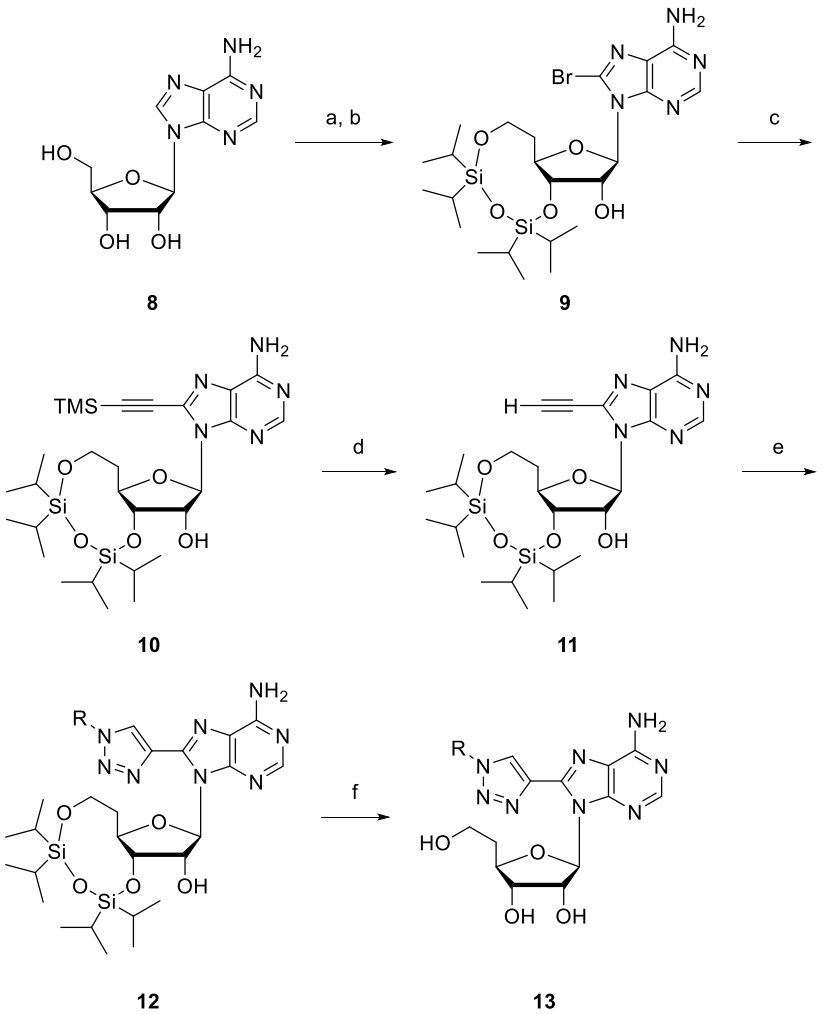
The convergent synthetic approach applied to the synthesis of nucleosides refers to a synthetic scheme where the nucleobase is fully constructed before performing a glycosylation with the desired sugar component. The convergent synthetic strategy was successfully applied in the synthesis of bicyclic thymine (bT, **7**, Scheme 1).[85] The desired sugar component (**3**) was first synthetically prepared starting from thymidine (**1**). The glycal was then coupled to the bicyclic core (**6**) *via* a Heck coupling. The convergent approach has been employed successfully by our group in the synthesis of the qAN1-qA$_{nitro}$ FRET pair.[82]



Scheme 1. Synthesis of the FBA bT. Reagents and conditions: (a) TBDMS-Cl, imidazole, DMF, rt, overnight, 83%; (b) TBDPS-Cl, imidazole, DMF, 60 °C, overnight, 100%; (c) TFA:H_2O 10:1, DCM, 0 °C, 4 h, 92%; (d) Ammonium sulphate, HMDS, 80 °C, then rt, TMS-Cl, reflux 4 h, 85%; (e) DMAP, pyridine, THF, Br_2, BBr_3, 85 °C, 2 h, 93%; (f) NaI, CuI, dioxane, *trans-N,N'*-dimethylcyclohexane-1,2-diamine, 110 °C, 12 h, 54%; (g) $Pd(OAc)_2$, $AsPh_3$, *t*$BuNH_2$, DMF 60 °C, 32 h, then TBAF, AcOH, 0 °C, 45 min, followed by ACN:AcOH 1:1, sodium triacetoxyborohydride, 0 °C, 1 h, 82%.

3.1.2 Divergent synthesis

The most common approach to synthesise modified nucleobase phosphoramidite building blocks is to work in a combined linear and divergent fashion.[86] The necessary synthetic handles are first installed on the desired nucleobase, such as chlorine or iodine for substitution and coupling reactions, respectively, which usually involves harsh conditions not suitable once the desired sugar component has been attached. These conditions may cause significant depurination. Glycosylation is then performed to achieve the (deoxy)ribonucleoside either *via* a substitution type reaction for deoxyribose or a Vorbrüggen reaction for ribose.[87,88] The glycosylated base can then be used for further functionalization on the previously installed chemical handles such as; substitution, coupling reactions and click chemistry. A valuable feature with this approach is that most of the variability is introduced towards the end of the synthetic scheme, which simplifies FBA library generation. Small changes can have a dramatic impact on photophysical properties such as quantum yield of fluorescence.[89] One successful example of applying the divergent strategy for generating a small FBA library was used by Dyrager *et al.* in the generation of C-8 triazole substituted adenines (Scheme 2).[90] The alkyne (**10**) required for click chemistry was introduced *via* a Sonogashira coupling of (**9**) and nine different azides were installed on the C-8 position of adenine (**12**).



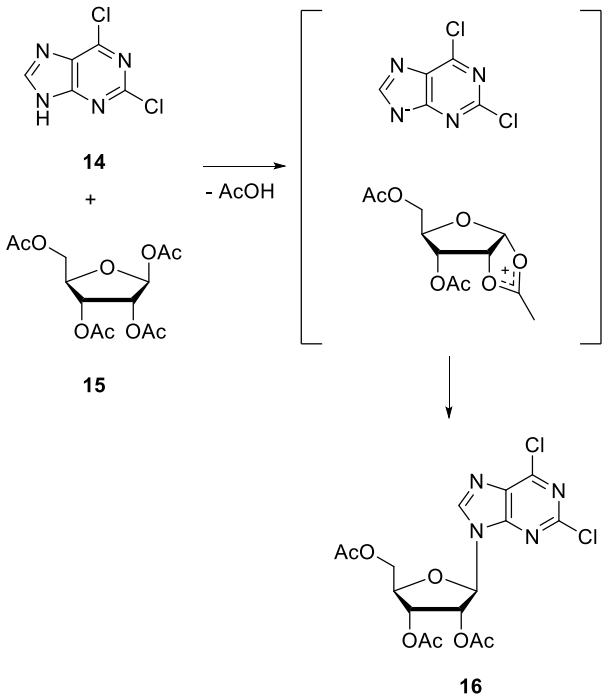
Scheme 2. Synthesis of C-8 substituted adenine *via* the divergent methodology. Reagents and conditions: (a) Br_2/NaOAc buffer, 83%; (b) tetraisopropyldisiloxane dichloride, pyridine, 65%; (c) TMS/acetylene, $Pd(PPh_3)_2Cl_2$, CuI, NEt_3, THF, 50 °C, 50 min, 80%; (d) NH_3 (aq. 25%)/EtOAc (1.5:1, v/v), rt, 14 h, 81%; (e) one of three protocols used: NaN_3 in DCM/H_2O, triflic anhydride, 0 °C, 2 h, then $NaHCO_3$, benzylamine derivative, $CuSO_4$-$5H_2O$, rt, 30 min, MeOH, followed by **11**, TBTA, L-ascorbic acid sodium salt, 60 °C, microwave (MW), 5 min; (f) TBAF (1 M in THF; 2 eq.), THF, rt, overnight, 99%.

3.2 Synthesis of nucleosides

FBAs are typically either chemically functionalized natural nucleosides or constructed from the beginning where a novel heterocycle is investigated. Due to the large number of FBAs in the literature and the diversity of such compounds, few general synthetic approaches to synthesise FBAs exist.[65, 86] In this section, the most common synthetic methods to synthesise nucleosides *via N*-glycosylation are briefly described.[91]

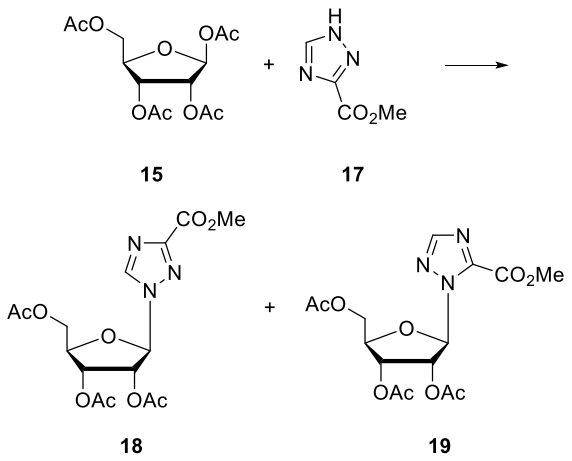
3.2.1 Fusion synthesis

Fusion synthesis, also known as melt condensation, employs a nucleobase which reacts with a C1'-acetoxysugar. The reaction is normally performed under Lewis acidic conditions with high temperatures (>150 °C) and under vacuum, all while releasing volatile acetic acid (Scheme 3).[92] The purine (**14**) was melted with the acetoxy-sugar (**15**), releasing AcOH yielding a bicyclic intermediate sugar that can accept a nucleophilic attack from the purine affording **16**.



Scheme 3. Early example of fusion synthesis to achieve *N*-glycosylation. Reagents and conditions: equimolar amounts of **14** and **15**, 115 °C, 15 min.

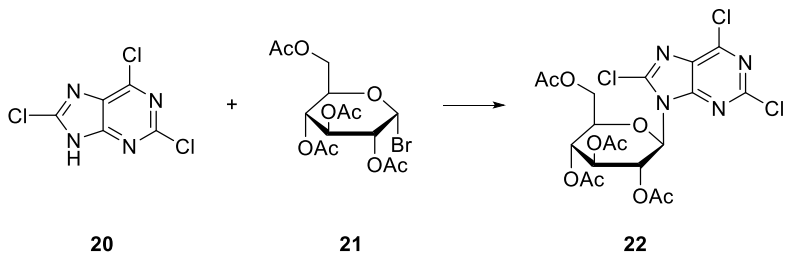
The fusion procedure was used in the initial synthesis of the antiviral ribavirin (Scheme 4).[93] Compound **15** was melted with the triazole (**17**) furnishing the isomers **18** and **19** in a 10:1 mixture.



Scheme 4. Fusion synthesis to produce the antiviral ribavirin. Reagents and conditions: ~0.1 mol% bis(*p*-nitrophenyl)phosphate, 165 °C, vacuum, 20min, 85% (10:1).

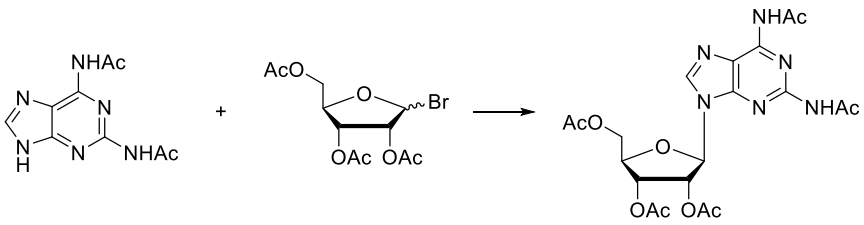
3.2.2 Metal salt method

The metal salt method was first used by Fischer *et al.* in the synthesis of the glucopyranoside (**22**, Scheme 5), where the silver salt of 2,6,8-trichloropurine (**20**) was coupled with the bromo-sugar (**21**).[94]



Scheme 5. Metal salt method used by Fischer *et. al.* to couple a purine with a pyranose. Reagents and conditions: Silver salt of **20** mixed with **21** in xylene.

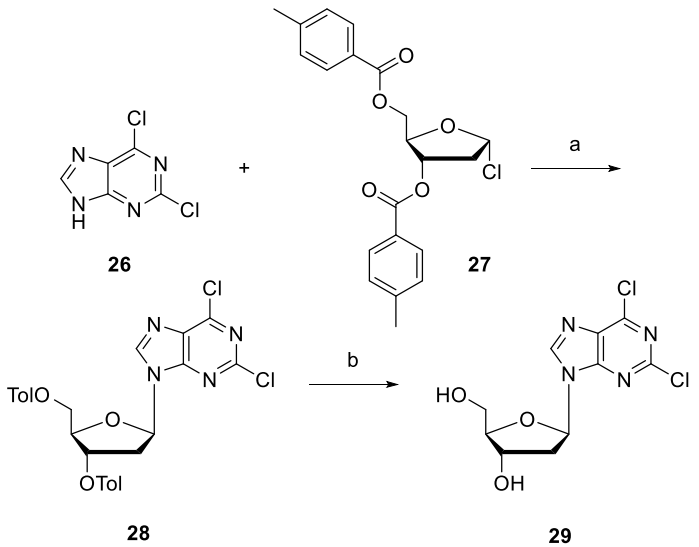
The metal salt method was later developed to employ mercury due to the increased reactivity and better solubility (Scheme 6).[95] The mercury salt of **23** was coupled to the bromo-pentose (**24**) to furnish **25** in a modest yield (57%).



23 **24** **25**

Scheme 6. Improved metal salt method utilizing mercury. Reagents and conditions: mercury salt of **23** mixed with **24** in xylene, reflux, 3 h, 57%.

More recently, the toxic heavy metals previously used in the metal salt procedure have been replaced by sodium (Scheme 7). The sodium salt of 2,6-dichloro-purine (26) was coupled to Hoffer's α-chloro sugar (27) providing 28 in good yield (82%). Deprotection of the toluoyl groups was achieved by heating 28 with methanolic ammonia up to 150 °C for 20 h furnishing 29 in a good yield (71%). Where the previous methods often provided anomeric mixtures, positional isomers and low yields, the sodium salt method provided a cleaner reaction with higher yields, all while avoiding the use of mercury.[87]

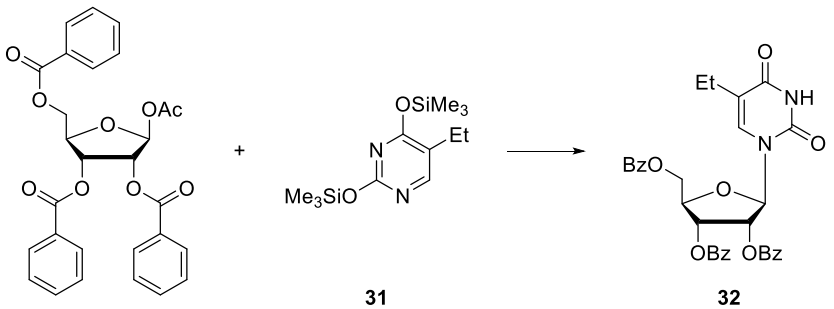


26 **27** **28** **29**

Scheme 7. Metal salt method employing sodium. Reagents and conditions: (a) 1 eq. NaH, acetonitrile, rt, 30 min, then 1 eq. of **27**, 50°C, 2 h, 82%. (b) NH_3/MeOH, 135-150 °C, 20h, 71%.

3.2.3 Vorbrüggen reaction

The Vorbrüggen reaction (also known as the silyl-Hilbert-Johnson reaction) is based on the seminal work from Hilbert and Johnson, in which pyrimidines were reacted with halo-hexose sugars to form the *N*-glycosidic bond.[96,97] The Vorbrüggen reaction avoids the halo-sugars employed in the Hilbert-Johnson reaction in favour of -OAc or -OR sugars that are easier to synthesise, modify, purify and store. The method is mild and performed at room temperature with Friedel-Craft catalysts such as $SnCl_4$, $ZnCl_2$ or TMSOTf (Scheme 8). The benzoyl protected acetoxy-sugar **30** was coupled to the pyrimidine **31** providing **32** in an excellent yield (95%).



30 **31** **32**

Scheme 8. Vorbrüggen reaction in the synthesis of nucleosides. Reagents and conditions: $SnCl_4$, 1,2-dichloroethane, rt, 48h, 95%.

3.3 Oligonucleotide chemistry

Linked nucleotides form an oligonucleotide. This section describes the most widely used methodologies in chemical synthesis, purification and analysis of short oligonucleotides (<100 nt). The standard practices described do not apply for longer oligonucleotides (>100 nt) and will therefore not be considered.

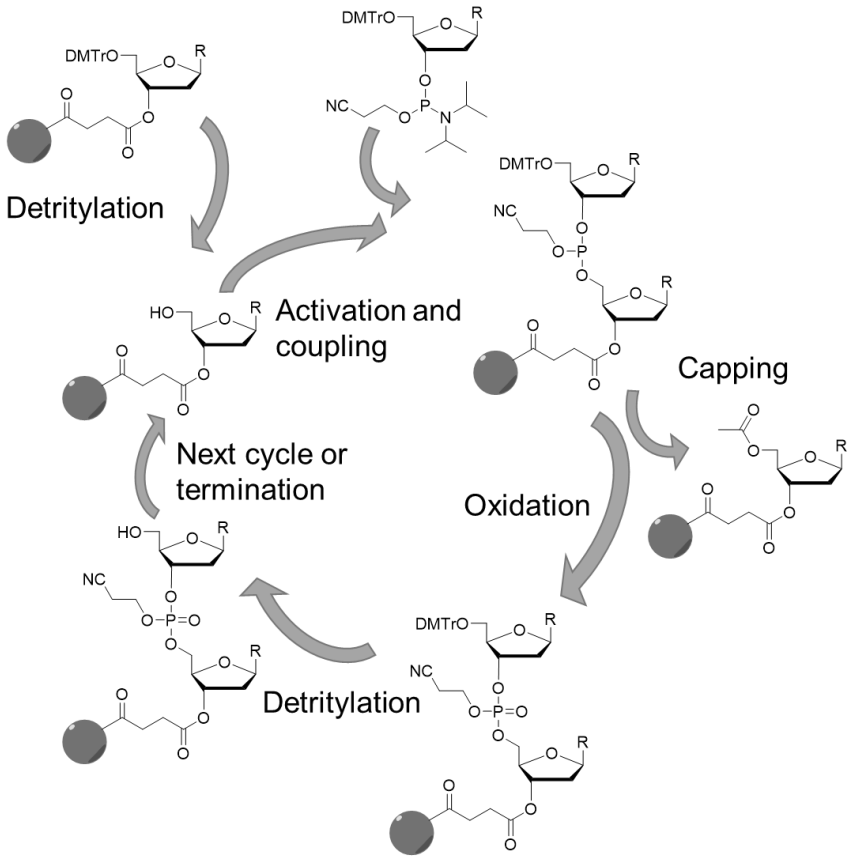
3.3.1 Oligonucleotide synthesis

Early oligonucleotide chemistry connected two synthetic nucleotides together via *H*-phosphonate chemistry.[98] The *H*-phosphonate chemistry was eventually adopted for solid-phase synthesis (SPS) of small oligonucleotides.[99] In parallel to this, phosphodi- and tri-ester chemistries were developed which increased selectivity,[100,101] that in turn opened up the use of more efficient coupling agents that dramatically reduced the

length of the synthesis and increased the yields significantly. In the 1970s, development of phosphite triester chemistry led Caruthers *et. al*. to pioneer the phosphoramidite chemistry that is being widely used today.[102] Most commonly oligonucleotide synthesis is performed *via* SPS by fully automated machines, where each nucleotide, as a protected phosphoramidite building block, is incorporated by four distinct steps in a growing chain (Scheme 9).[103] These steps consists of:

1. Detritylation: the trityl group on the solid support is cleaved using TCA in DCM.
2. Activation and coupling: the monomeric phosphoramidite building block are activated with 5-(benzylthio)-1*H*-tetrazole (BTT) and coupled to the nucleotide attached to the solid support.
3. Capping: the unreacted material attached to the solid support is capped with a mixture of acetic anhydride and 1-methylimidazole to prevent accumulating n-1 species.
4. Oxidation: the reactive phosphor(III) is oxidized to phosphor(V) by a mixture of I_2, water and pyridine.

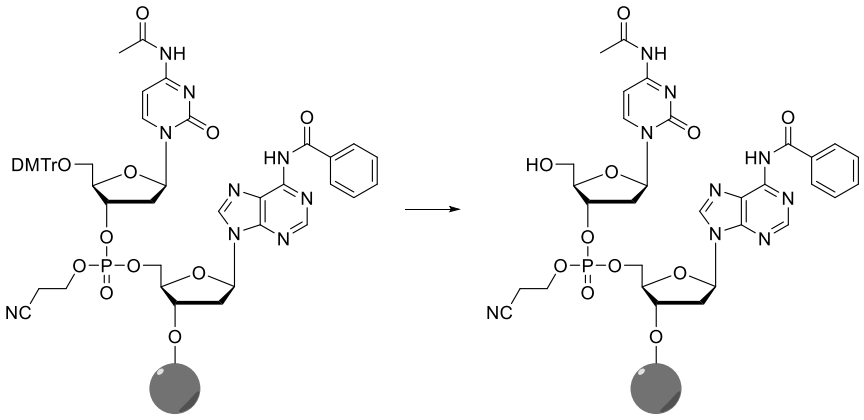
Principally, the synthesis of DNA and RNA are identical. However, due to the 2'-OH in RNA, an additional protection group is required. Most commonly, TBDMS is used as it strikes the right balance between effectively removing the reactivity and potential 2'-OH to 3'-OH phosphate migration during synthesis and being small enough as not to add too much steric bulk. There is no defined maximum length of an RNA oligonucleotide synthesized with TBDMS protecting groups, although many commercial suppliers only provide oligonucleotides up to 45 nt. For longer RNA constructs other 2'-OH protection groups are recommended. Nonetheless, where the coupling step in DNA synthesis usually takes less than 30 seconds, it can take at least 10 times longer for RNA depending on a number of factors such as the choice of solid support, length of the oligonucleotide, or use of modified bases amongst others.



Scheme 9. Oligonucleotide synthesis cycle, R = nucleobase.

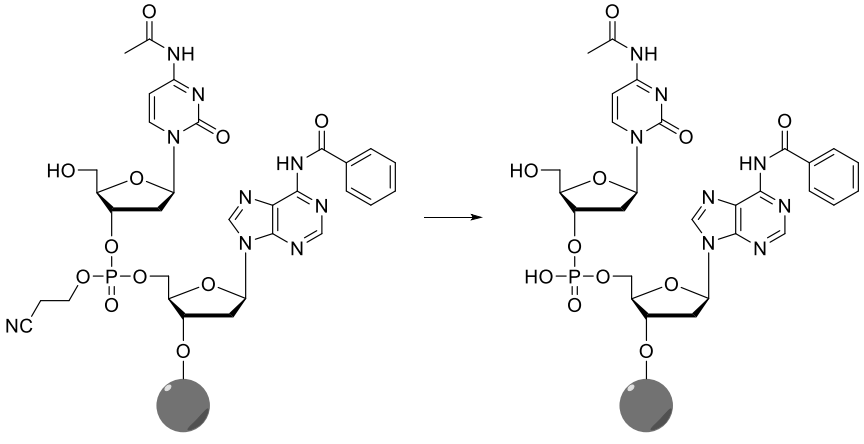
3.3.2 Oligonucleotide workup, purification and analysis

Upon synthetic completion of the oligonucleotide, some post-synthetic handling is required. The final 4,4-dimethoxytrityl (DMTr) protecting group is removed at the end of the synthesis using the same procedure as used during detritylation step of the oligonucleotide synthesis, *i.e.*, while the oligonucleotide is still on the solid support (Scheme 10). However, the DMTr protection group can also be kept on the oligonucleotide as a hydrophobic handle to simplify reverse phase high performance liquid chromatography (RP-HPLC) purification for difficult separations.



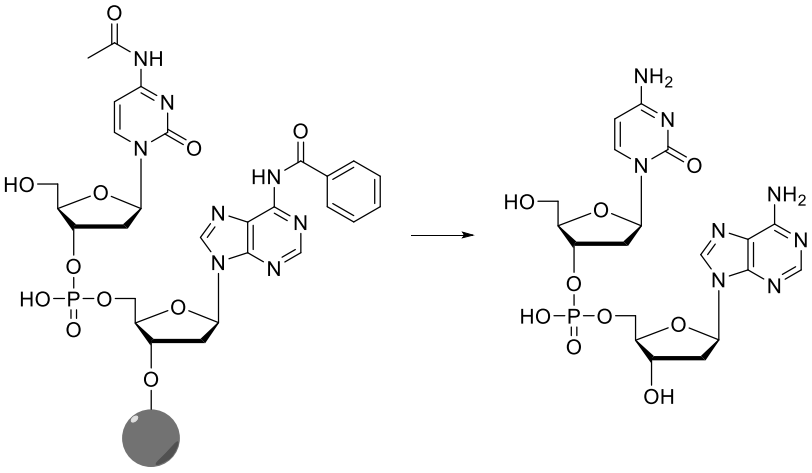
Scheme 10. Detritylation of final 5'-OH trityl protection group.

The 2-cyanoethyl (CE) phosphate protection groups are cleaved with a non-nucleophilic base such as DEA in acetonitrile *via* β-elimination (Scheme 11). This procedure is performed with the oligonucleotide still attached to the solid support and is done in a flow to rapidly remove any formed reactive acrylonitrile directly, which otherwise can undergo Michael addition to thymine or uracil.



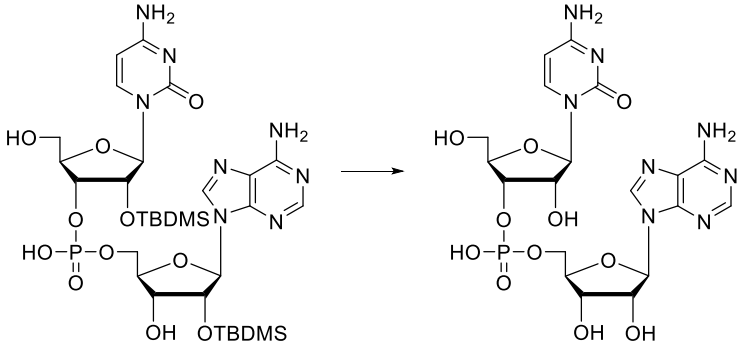
Scheme 11. Removal of the 2-cyanoethyl protection groups.

Cleavage from the resin is achieved by using a mixture of concentrated aqueous ammonia and ethanol, which also removes the nucleobase protection groups (Scheme 12).



Scheme 12. Deprotection of the nucleobase protection groups and cleavage from the solid support.

In the case of RNA, the 2'-OH protecting TBDMS groups needs to be cleaved, which is performed using a fluoride source, such as Et_3N*3HF (Scheme 13). As RNA handling requires a fluoride source, additional steps of removing excess fluoride need to be added to the procedure. Precipitating the oligonucleotide with *n*-butanol followed by centrifugation and removing the supernatant is adequate.



Scheme 13. Final deprotection step of RNA.

The crude oligonucleotide is finally purified by RP-HPLC using a mobile phase consisting of acetonitrile and a slightly basic buffered solution. Short DNA sequences (<30nt) can be purified using readily available C8 or C18 columns, however longer sequences (and especially RNA) sets high requirements on the stationary phase, where a small pore size is

critical to obtain high purity oligonucleotides. Ion-exchange chromatography can also be used, but this requires de-salting upon completion, whereas the buffer components in an RP-HPLC are usually volatile and depending on the desired counterion no extra salt swapping or desalting is required. Less practical, polyacrylamide gel electrophoresis can be employed to purify the oligonucleotide, generating high purity samples of the oligonucleotides, unfortunately in very small amounts.

The final analysis of the synthesised oligonucleotide is to identify that the correct oligonucleotide has been synthesised, and to check the impurity profile. Usually, a detailed liquid chromatography–mass spectrometry (LC-MS) can provide the required data, but greater detail can be obtained by running a combined LC-time-of-flight-MS (LC-TOF-MS).

3.4 Binding interaction measurements

In this section two of the most common and readily available techniques for measuring small molecule binding interaction with an oligonucleotide are presented, which has been used in **Paper V**. The chapter ends with an explanation of how binding interaction of small molecules and oligonucleotides can be achieved by steady-state fluorescence and FRET.

3.4.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is measured in a calorimeter with a sample and reference cell located in an adiabatic jacket (Figure 11a).[104] The raw data is displayed as a change in μcal/s against time and further processed to kcal/mol against the molar ratio of added ligand compared to sample (molar ratio). The data provided from analysis (Figure 11b) includes; binding affinity (K_a), enthalpy changes (ΔH) and stoichiometry of binding, usually referred to as the *n* value.

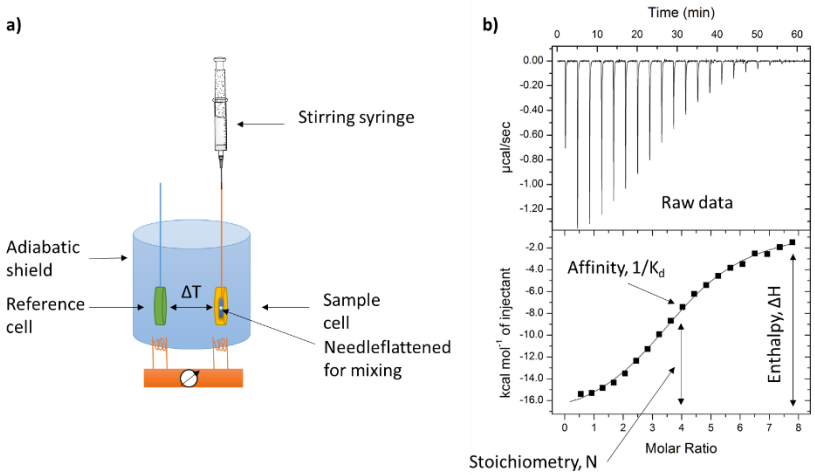


Figure 11. The basics of an isothermal titration calorimetry (ITC) experiment.

ITC is a general technique and is often employed to measure small molecules binding to larger macromolecules.[39] The two most common ITC machines are the full volume ITC (1000 uL cell volume) and the reduced volume ITC (200 uL cell volume). Standard conditions involve titrating a small amount of a ten times concentrated ligand to sample. Approximately twenty additions are performed, noting the enthalpic change in each instance over the course of one hour, giving ample time for equilibrium between most types of ligand and sample to form. The technique is laborious, requiring several reference runs to be performed. Collecting a complete binding interaction data set for one small molecule binding to one protein or oligonucleotide can take a full day.

A complete set of data from ITC consists of four experiments:

1. Sample of interest in cell to which the ligand of interest is titrated.
2. Sample of interest in cell to which the buffer used is titrated.
3. Buffer in cell to which the ligand of interest is titrated.
4. Buffer in cell to which the buffer used is titrated.

The results of experiments 2-4 are subtracted from experiment 1 which then shows the true enthalpic change contributed by adding the ligand of interest to the sample of interest. ITC is extremely sensitive and to obtain high-quality data the concentration and binding affinity of the sample and the ligand needs to match. Moreover, a careful selection of buffer conditions is required to avoid non-specific interactions between the

ligand and the buffer or the sample and the buffer. Another drawback of ITC is the relatively large amounts of sample and ligand that are required compared, for example, to techniques which employ fluorescence readout.

3.4.2 Surface plasmon resonance

Surface plasmon resonance (SPR) is measured on an SPR instrument. A large variety of instruments exists to suit the needs of the application, ranging from low-throughput systems where various parameters can be modified, to high-throughput systems, capable of screening 10,000 ligand interactions per day. The sample or the tested ligand is adhered to a metal surface (Figure 12). Normally, the biotin-streptavidin system is used, but covalent links have also been used.[105]

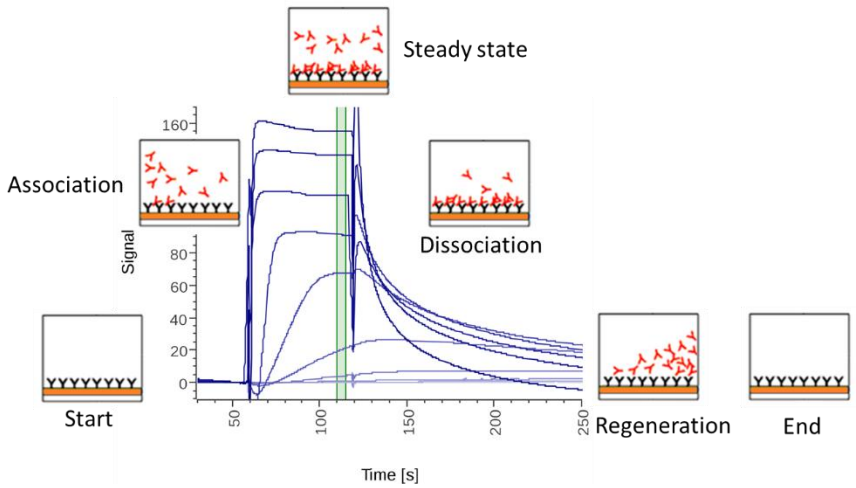


Figure 12. The main phases of a surface plasmon resonance (SPR) experiment where a sample is adhered to a surface, the ligand of interest is bound and then dissociated, followed by a regeneration of the sample chip. Green marked area indicates data acquisition for binding affinity.

The excess sample is washed away, followed by association of the ligand of interest. After a limited time, the ligand is allowed to dissociate which is followed by a sample regeneration by washing the chip. Data are recorded for several different concentrations of the sample being tested and are displayed in a histogram with time on the x-axis and response units (RU) on the y-axis. Marked in green (Figure 12) indicates steady-state and is where the response value is collected. An SPR experiment provides the binding constant (K_D) and stoichiometry of binding, and detailed binding kinetics such as k_{on} and k_{off}.

3.4.3 Steady-state emission spectroscopy

In steady-state emission spectroscopy the emission of photons from a molecule is measured using a spectrofluorometer. The sample is typically continuously illuminated by a white light lamp where one wavelength at the time is monitored and the emission of the sample is captured in a detector. Several different cuvettes exist, and the most common ones are standard 1.5-3 mL cuvettes, but sizes range down to 60 uL reduced volume cuvettes for precious samples. When observing biomolecules such as proteins and oligonucleotides one needs to be aware of any effect of surface adhesion that the plastic pipette tips and quartz surfaces might have on the sample of interest.[106]

The intensity of emitted light that is measured can be compared to that of a known compound providing the fluorescence quantum yield. By observing the change in quantum yield for a sample labelled only with a FRET donor compared to a sample with a paired FRET acceptor, we can determine whether the change in quantum yield originates from changes in the local microenvironment of the FRET donor, or if the change originates from a difference in distance or orientation of the FRET pair. If we instead use a FRET pair where the donor and acceptor are virtually unresponsive to changes in the local microenvironment, such as tC^{O}-tC_{nitro},[80] we would know that any change in the fluorescence quantum yield comes from a change in either relative distance or orientation between the probes directly.

To test if a ligand interacts with an oligonucleotide, the following experiment may be performed. The fluorescently labelled oligonucleotide sample is dissolved in the desired buffer and added to a cuvette. An emission spectrum is recorded and then the ligand of interest is added where upon a new emission spectrum is recorded. Any change in measured emission, accounting for dilution, originates from a binding event of the ligand of interest to the oligonucleotide.

3.4.4 Time-resolved emission spectroscopy

Time-resolved fluorescence lifetimes can be measured using a time-correlated single photon counter (TCSPC) setup. The sample is placed in a cuvette in which a pulsed laser can excite the sample. Each pulse excites a part of the sample and the aim is to measure the decay of the excited state as a function of time. Because emission is a random process, some molecules will emit a photon after a short period of time, where others will

emit after a longer period. For most small molecules, the fluorescence lifetimes are in the nanosecond regime and the special electronic setup makes it possible to measure the time between the excitation of the sample and the detection of a photon in the detector.

Both steady-state and time-resolved emission spectroscopy can be used to determine quantum yield, and, thereby, also $FRET_{eff}$. However, time-resolved emission spectroscopy is not concentration dependent and can provide additional insight into subtle changes in quantum yield where for instance a ligand is titrated into the sample, reducing the sample concentration.

4. Original work

This chapter summarises the work presented in the five papers that comprise this thesis. The first section explains the design and synthetic process undertaken to create fluorescent adenine analogues (**Paper I**), followed by the synthesis and incorporation of adenine FBAs into DNA (**Paper II**). The synthesis chapter is concluded by explaining the synthesis and incorporation of FBAs into RNA (**Papers III–IV**). The final section demonstrates how our developed probes can be used to study the binding interaction of RNA and small molecules (**Paper V**).

4.1 Design and synthesis of new FBAs

This chapter intends to present the key transformations and synthetic steps that yielded novel fluorescent nucleobase analogues (**Papers I–IV**). Figure 13 illustrates the A and C analogues that have been developed.

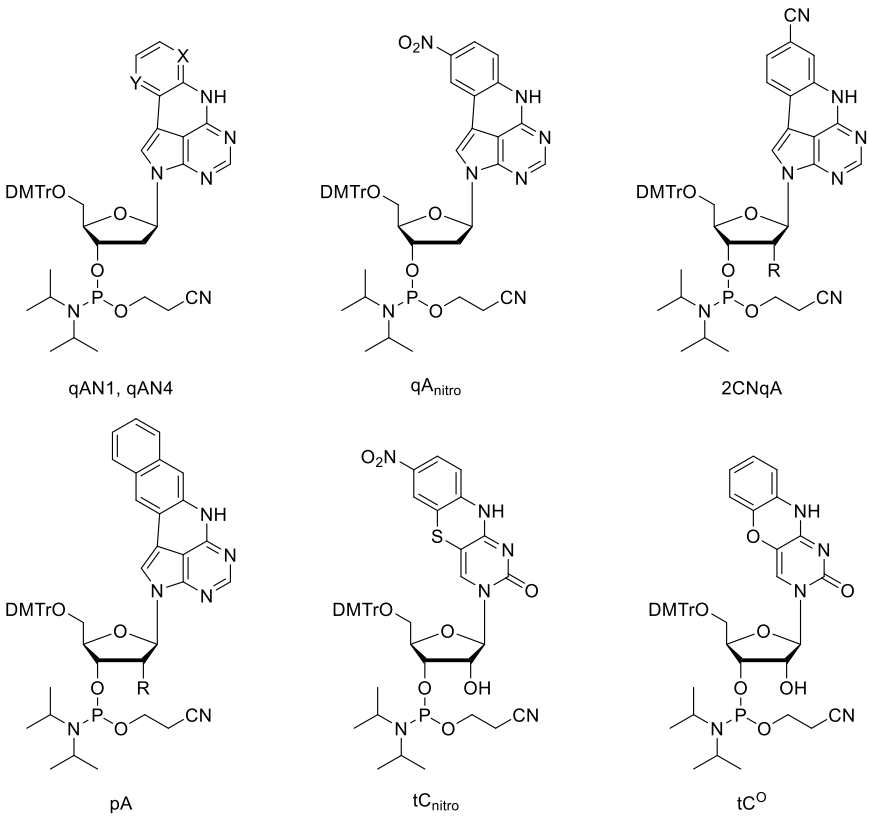


Figure 13. Structures of developed DNA and RNA phosphoramidites. For X = N, Y = H, qAN1, for X = H, Y = N, qAN4. R = deoxyribose or ribose.

4.1.1 Design of non-perturbing FBAs

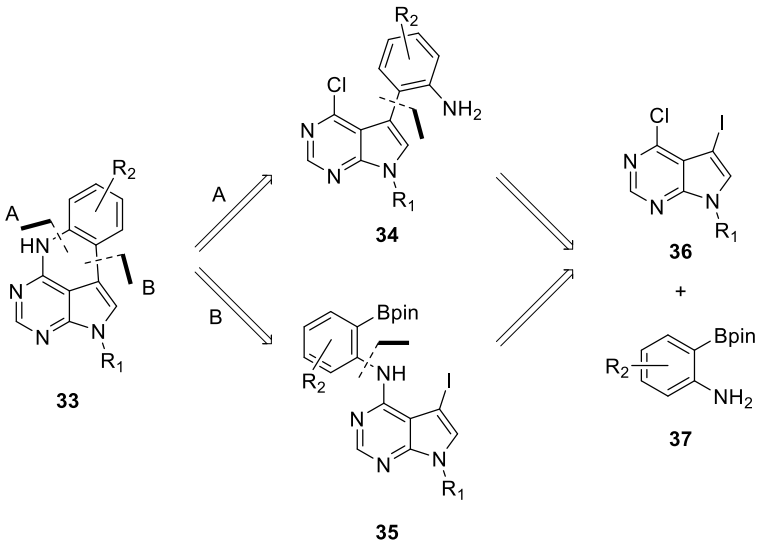
The development of bright non-perturbing FBAs remains a challenge. Small modifications to the native nucleobases can introduce fluorescence.[68,107] More dramatic changes, such as extending the conjugation *via* aromatic ring-fusion or introducing fluorescent labels such as pyrene conjugated to the nucleobase, can lead to greater brightness (molar absorptivity multiplied by quantum yield of fluorescence, $\varepsilon\Phi_F$) but is limited by issues such as interaction with the tertiary structure of the oligonucleotide.[108,109] Thus, a careful consideration of the geometrical constraints is required before attempting to construct a novel, bright, and most importantly, non-perturbing FBA. The adenine scaffold offers several sites for modifications: C2, C8, the C6 exocyclic amino functionality and the N7 to C7 substitution leading to 7-deazaadenines.[86] In general, the modifications encountered by the hydrogen bonding surface such as C2 or C6 exocyclic amino substitutions are problematic as they can potentially interfere with the hydrogen-bonding properties of the nucleobase. Substitutions on the C8 have proven to destabilise the double helix by pushing the glycosidic bond from anti to syn.[110–112] However, switching N7 to C7 and placing substitutions on the C7 position can be tolerated.[113]

To date, there are no accurate means to predict the fluorescent properties of small molecules and thus correctly guide their synthesis. In general, conjugated aromatic molecules increase molar absorptivity, and the introduction of heteroatoms red-shift the molecule's absorption. However, it is unclear as to how the quantum yield is affected by such modifications. In some cases, calculating the S_0–S_1 oscillator strength can be indicative of the relative quantum yield within a set of compounds.[89] Furthermore, the incorporation of FBAs into oligonucleotides generally quenches the emission, and the mechanisms are not completely characterised, further increasing the complexity of the FBA design.[82] To generate a building block required for DNA or RNA, SPS typically requires great effort in terms of protecting reactive fragments, and the entire synthesis of such a compound can easily become a tedious multi-step process. Therefore, it is necessary to know that the FBA that is being made will be valuable in terms of photophysical properties, and our approach has been to synthesise only the heterocyclic moiety of the FBA that is responsible for fluorescent properties. By shortening the number of steps and simplifying the previously developed chemistries we could synthesise focused libraries of advanced heterocyclic scaffolds.

4.1.2 Fluorescent multicyclic adenine analogues

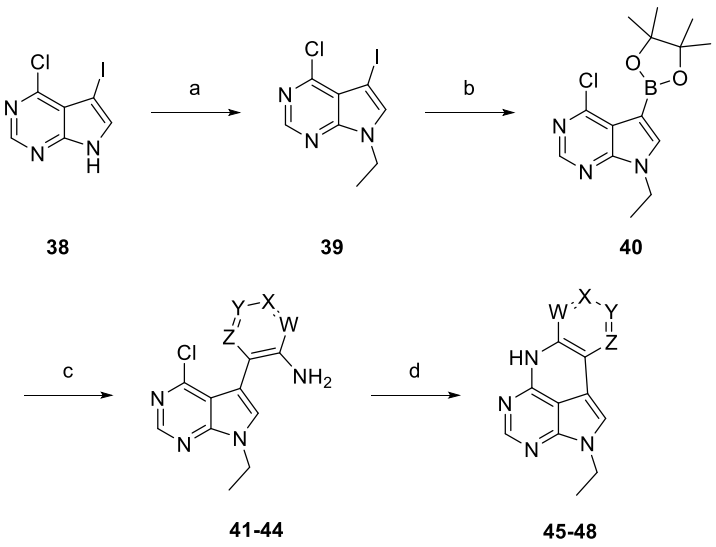
Our starting point was the quadracyclic adenine (qA) that was developed previously by our group.[114] We chose this scaffold, as we knew it was an excellent adenine analogue in terms of size, shape, and base pairing properties. Unfortunately, the photophysical properties were not great, with a modest quantum yield of 0.07 as a monomer and almost completely quenched with a quantum yield of 0.003 on average in double-stranded DNA (dsDNA).[114] Upon examining the structure, it was apparent that we could introduce modest modifications to ideally increase quantum yield as well as red-shift the absorption. However, by observing the linear 11-step synthetic scheme of qA with a total yield of 1.5%, it was apparent that synthesising several modified qA derivatives would be extremely laborious. Ideally, we would only synthesise the heterocyclic moiety as a preliminary-FBA (pre-FBA). The reasoning was that if the substituent on the N-9 of the purine scaffold was non-aromatic and non-conjugated to the aromatic system, it would adequately resemble the deoxyribose component to indicate whether the synthesised pre-FBA was a good candidate for DNA phosphoramidite synthesis.

By performing a retrosynthetic analysis of the qA scaffold we found two short and similar routes leading to commercially available reagents. By first disconnecting the secondary amine between the top and the bottom aromatic rings, we ended up with the possibility of performing an intramolecular cyclisation by substituting the chloride for an exocyclic amino group (**34**, route A, Scheme 14). The next disconnection was made at the C-7 deaza-position of adenine, furnishing a 6-chloro-7-iodo-deazapurine scaffold (**36**), which is commercially available lacking substitutions on N9 and an ortho-boronpinacolester-aniline species (**37**). In the second route (route B, Scheme 14), the first disconnection is made at the C-C bond joining the top and bottom aromatic rings, assumed to be possible to form from an intramolecular cross-coupling reaction such as Suzuki-Miyaura cross-coupling (**35**). The second disconnection performed at the secondary amine generates the same starting materials as for route A.



Scheme 14. Retrosynthetic analysis of **33**. R_1 = protecting group or alkyl chain. R_2 = freely selectable substituent.

We opted for route A and the synthesis commenced from commercially available 6-chloro-7-iodo-deazapurine (**38**, Scheme 15, **Paper I**). In order to not impact the photophysical properties, we determined that an ethyl group instead of the ribose on the N9 position would be adequate. The ethylated product (**39**) was obtained in 90% yield under anhydrous conditions with a slight excess of ethyl iodide and caesium carbonate in DMF and without the need for column chromatography. Miyaura borylation of **39** with pinacolborane and an excess of triethylamine in dioxane provided **40** in 86% yield after flash chromatography. A catalyst screening was carried out resulting in the conditions presented in Scheme 15 (for details see **Paper I**). A slight excess in equivalents of the ortho-iodo-anilines was coupled to **40** *via* a Suzuki-Miyaura cross-coupling utilising $PdCl_2(PPh_3)_4$ as the catalyst in good yield.[115] Finally, the cyclisation was performed by treating **41**–**44** with chlorotrimethylsilane (TMS-Cl) to increase the exocyclic nitrogen's nucleophilicity,[116] followed by the addition of excess lithium bis(trimethylsilyl)amide (LiHMDS) to yield **45**–**48** in moderate to good yield (55-71%).



Scheme 15. Synthesis of ethylated qAN1-4 FBAs. Compounds **41** and **45**, W = N and X, Y, Z = CH. Compounds **42** and **46**, X = N and W, Y, Z = CH. Compounds **43** and **47**, Y = N and X, W, Z = CH. Compounds **44** and **48**, Z = N, W, X, Y = CH. Reagents and reaction conditions: (a) EtI (1.2 eq.), Cs_2CO_3 (1.2 eq.), DMF, rt, 4 h, 90%. (b) HBpin (1.1 eq.), $Pd(PPh_3)_4$ (3 mol%), Et_3N (10 eq.), dioxane, 80 °C, 24 h, 86%. (c) Aniline (1.1 eq.), $PdCl_2(PPh_3)_4$ (3 mol%), K_3PO_4 (2.5 eq.), MeCN-H_2O 1:1, 80 °C, 2 h, 56-79%. (d) TMS-Cl (1.05 eq.), rt, 30 min, ii) LiHMDS (2.5 eq.), 100 °C, MW, 3 h, 55-71%.

Using this chemistry, we synthesised 12 additional qA-derivatives.[89] However, to how incorporation into DNA and/or RNA would affect photophysical properties was yet to be observed. In parallel to developing the first probes for DNA SPS, over 20 pre-FBAs were synthesised in the group, and their photophysical data were evaluated (**49**–**60**, data not published, Figure 14), However, only the most interesting species were progressed into phosphoramidites for oligonucleotide incorporation.

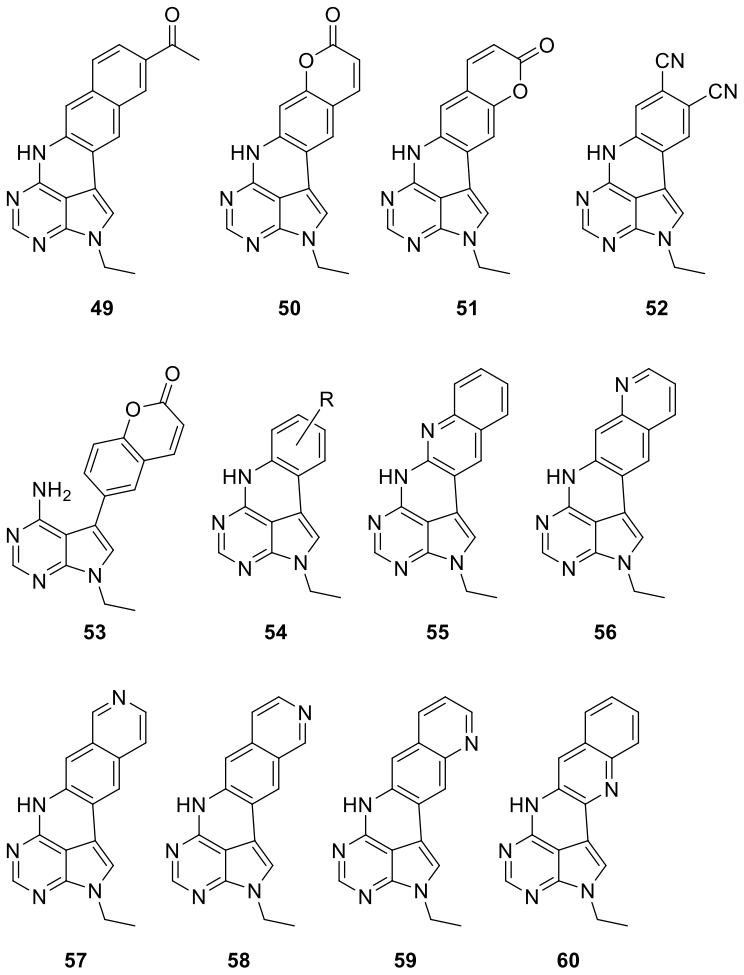


Figure 14. Various synthesised aromatic heterocycles (**49**–**60**), data not published.

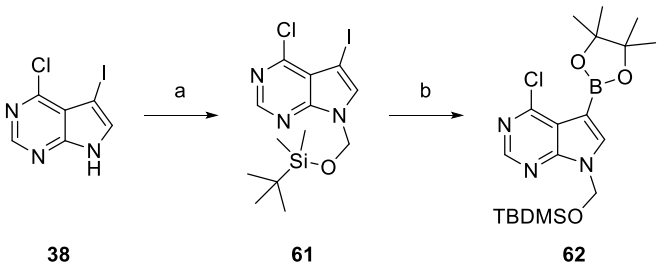
4.1.3 Synthesis of DNA phosphoramidites

Derived from the photophysical characterisations performed in **Paper I**, the first target FBA for incorporation was qAN1. The synthesis and incorporation of the fluorescent probe, FRET donor qAN1, and the FRET acceptor/quencher qA_{nitro} was achieved, thereby constituting the first adenine–adenine interbase FRET-pair.[82] During the development of the synthetic scheme for the qAN1 FBA, we continued to synthesise several other fluorescent adenine pre-FBAs.[89] The size-expanded quadracyclic

adenine scaffold, pentacyclic adenine (pA, **Paper II**, data not published for the pre-FBA) and the 2CNqA compound[89] were of particular interest.

Thus, we set out to establish a viable synthetic route built on a convergent design strategy, utilising a late-stage glycosylation. The idea was that the synthesised heterocyclic nucleobase scaffold can be functionalised freely in the N9 position, employing any glycosylation chemistry. DNA phosphoramidite of 2CNqA was made identically as the route of pA shown below.

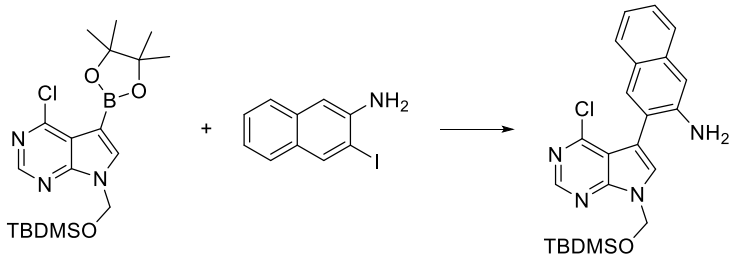
The multi-gram synthesis started from commercially available 6-chloro-7-deaza-7-iodo-purine (**38**), which was protected in a two-step procedure using formaldehyde under basic conditions followed by *tert*-butyldimethylsilyl trifluoromethanesulfonate (TBDMSOTf) in pyridine to yield the N-9 protected deaza-purine in 86% yield after filtration (**61**, Scheme 16).[117] The common intermediate used for both pA and 2CNqA was prepared through the borylation of **61**, under previously developed conditions (**Paper I**) in high yield (91%). This two-step protocol was used to prepare batches of >25 g of **62** (data not published).



38 **61** **62**

Scheme 16. Synthesis of common intermediate used to prepare various quadra- and penta-cyclic DNA phosphoramidites. Reagents and reaction conditions: (a) i) HCHO (2 eq.), NaOH (0.1 eq.), MeCN, 50 °C, 1 h. ii) TBDMS-OTf (1.2 eq.), pyridine, 0 °C, 30 min, 86%. (b) HBpin (1.2 eq.), $Pd(PPh_3)_4$ (2.5 mol%), Et_3N (1.5 eq.), dioxane, 90 °C, 4.5 h, 91%.

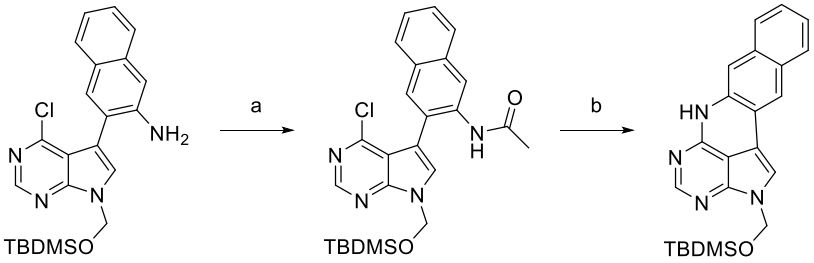
Utilising the previously mentioned Suzuki-Miyaura cross-coupling (**Paper I**),[82,89] the reaction between **62** and 3-amino-2-iodo-naphtalene (**63**) with $PdCl_2(PPh_3)_4$ furnished **64** in good yield (71%, Scheme 17).



62 **63** **64**

Scheme 17. Synthesis of **64** en route to pA DNA phosphoramidite. Reagents and reaction conditions: 3-amino-2-iodo-naphtalene (1 eq.), $PdCl_2(PPh_3)_4$ (4 mol%), K_2CO_3 (2.5 eq.), MeCN-H_2O 19:1, 80 °C, 1.5 h, 71%.

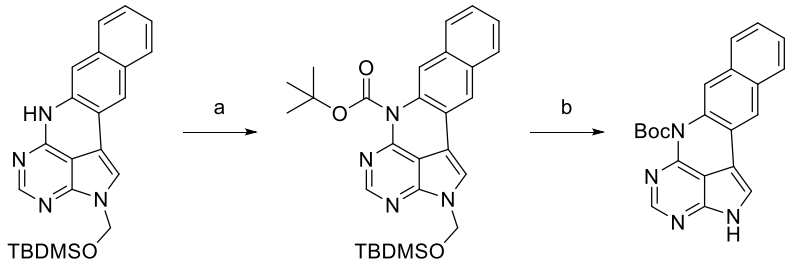
The synthetic route was continued with employing a two-step protocol to cyclise the scaffold (Scheme 18). The exocyclic amino group of **64** was activated for nucleophilic aromatic substitution by being reacted with acetyl chloride under basic conditions in DCM to furnish **65**. Then, the reaction mixture was evaporated until dry, followed by re-dissolving in THF to which LiHMDS in excess was added and the reaction mixture was heated in the microwave. This yielded the cyclised pentacyclic adenine scaffold (**66**) in good yield (73%, Scheme 18). We needed to use different activation procedures for the exocyclic amine than for the ethylated qAN1-4 synthesis (Scheme 15). As we scaled the reaction up from the synthesis of the pre-FBAs (**Paper I**),[89] the activation by silylation was difficult to control, with doubly silylated material generated as the main by-product.



64 **65** **66**

Scheme 18. Synthesis of the pentacyclic scaffold **66**. Reagents and reaction conditions: (a) AcCl (1.15 eq.), pyridine (1.25 eq.), DCM, rt, 1 h. (b) LiHMDS (1.8 eq.), THF, 100 °C, MW, 30 min, 73%.

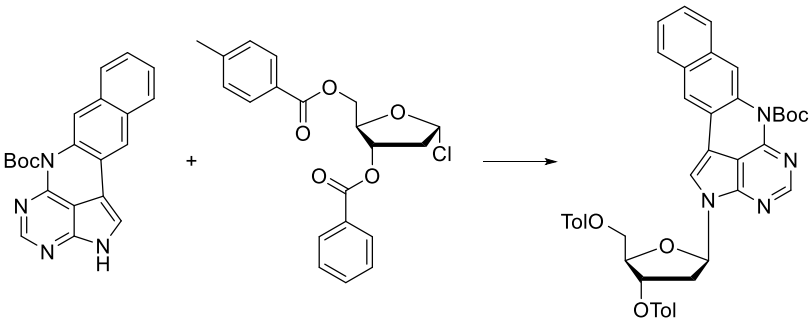
To prepare the nucleobase for glycosylation, a protection group was required at the N6 position and the N9 was required to be unprotected. First, the *tert*-butyloxycarbonyl (Boc) protection of **66** was achieved using Boc anhydride in excess with 4-dimethylaminopyridine (DMAP) in THF with a good yield (80%) of **67** (Scheme 19). Then, *tert*-butyldimethylsilyloxymethyl (TBDMSOM) removal was performed using tetra-butylammonium fluoride (TBAF) and ethylenediamine in high yield (92%) of **68**.



66 **67** **68**

Scheme 19. Synthesis of **68** ready for *N*-glycosylation. Reagents and reaction conditions: (a) Boc_2O (2.2 eq.), DMAP (2.5 eq.), THF, rt, 24 h, 80%. (b) TBAF (1 eq.), ethylenediamine (2 eq.), THF, 0 °C, 15 min, 92%.

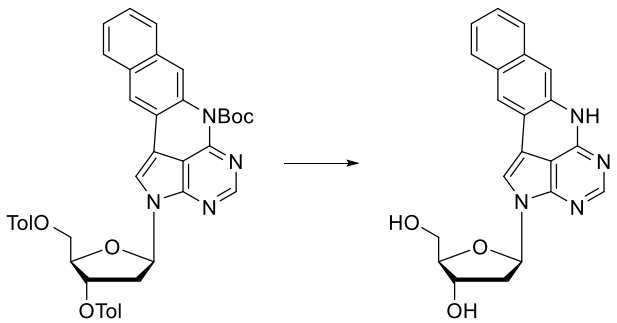
Glycosylation was then performed using the metal-salt method (Scheme 20),[87] in which sodium hydride was pre-stirred with the nucleobase (**68**), thus achieving the full deprotonation of the N9 and keeping the sodium as a counter-ion. Hoffer's α-chloro-sugar (**27**),[118] was then added in a slight excess which furnished **69** in a modest yield (57%).



68 **27** **69**

Scheme 20. Synthesis of protected pA nucleoside **69** by the metal salt method. Reagents and reaction conditions: i) NaH (1.35 eq.), MeCN, 0 °C, 3 h. ii) Hoffer's α-chloro-sugar (1.2 eq.), rt, 2 h, 57%.

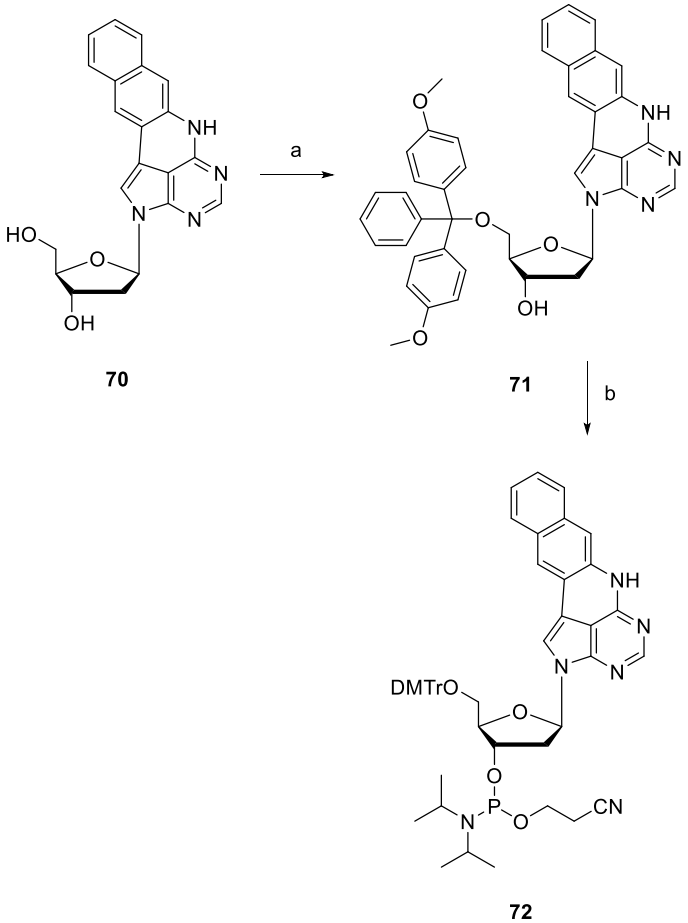
Previously, for deprotection during the synthesis of qA, a low-yielding two-step deprotection approach was employed, in which the Boc group was first cleaved under acidic conditions, and then the toluoyl groups were cleaved under basic conditions.[114] Increasing the basicity of the sodium methoxide by changing the solvent from methanol to acetonitrile, we achieved global deprotection without the need for chromatographical purification, thus turning the low yield into quantitative (Scheme 21).



69 **70**

Scheme 21. Synthesis of the unprotected pA nucleoside **70**. Reagents and reaction conditions: NaOMe (6 eq.), MeCN, 50 °C, 20 min, 99%.

The completion of the DNA phosphoramidite for SPS (**72**, Scheme 22) was achieved by DMTr-protection of the primary alcohol of **70** with a modest yield (59%), followed by the phosphitylation treatment of **71** with chloro-(2-cyanoethoxy)diisopropylaminophosphine (CEP-Cl) providing **72** in high yield (88%).



Scheme 22. Synthesis of pA DNA phosphoramidite that is ready for SPS. Reagents and reaction conditions: (a) DMTr-Cl (1.3 eq.), pyridine, rt, 1.5 h, 59%. (b) CEP-Cl (2 eq.), *N*-methylmorpholine (4 eq.), DCM, rt, 2 h, 88%.

The subsequent characterisation revealed several interesting properties, such as a high retained brightness upon incorporation into the double-stranded DNA and a remarkably high two-photon cross-section useful for imaging (**Paper II**).

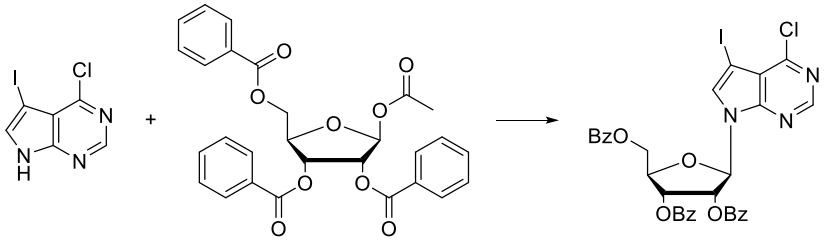
4.1.4 Synthesis of RNA phosphoramidites

The carefully developed synthetic route that promptly provided adenine analogue phosphoramidite building blocks for DNA SPS unfortunately did not translate well for the RNA chemistry. We established that the Boc protection group previously used was not stable at the N6 position under mild Vorbrüggen conditions. The Lewis acidic $SnCl_4$ or trimethylsilyl trifluoromethanesulfonate (TMS-OTf) mainly led to a complete Boc removal and a complex reaction mixture of various glycosylated heterocyclic species. Moreover, screening protecting groups did not improve the situation, as benzyl, albeit working for the glycosylation, proved too difficult to cleave, and other protection groups were generally not stable during the glycosylation.

Instead, we remodelled the entire synthetic strategy to a more traditional and linear one, where the established protocols for glycosylation of purines was successfully employed. Several attempts were made to cyclise the scaffold using our improved conditions with LiHMDS, but eventually only 1,4-diazabicyclo[2.2.2]octane (DABCO) in combination with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) was adequate as other reactions led to depurination, thus lowering the yield significantly.[114]

At first, 2-cyano quadracyclic adenine (2CNqA) (**Paper III**) was synthesised as an RNA phosphoramidite. However, following the protocol below, with a few minor adjustments, the pA RNA phosphoramidite was synthesised accordingly (manuscript in preparation).

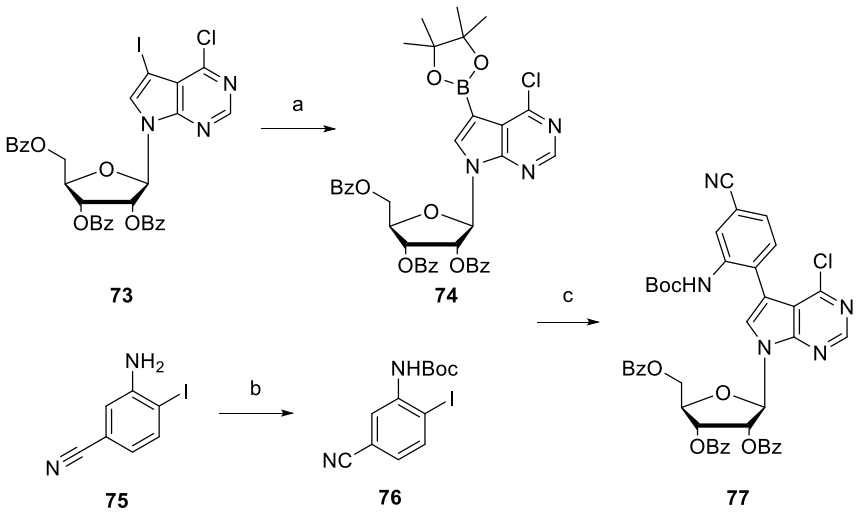
The linear synthesis began with a multi-gram Vorbrüggen *N*-glycosylation of 6-chloro-7-iodo-deazapurine (**38**) with the benzoyl protected ribose (**30**, Scheme 23) with a modest yield (60%).



38 **30** **73**

Scheme 23. Vorbrüggen reaction employed to yield **73**. Reagents and reaction conditions: i) *N*,*O*-bis(trimethylsilyl)acetamide (1.1 eq.), MeCN, rt, 20 min. ii) ribose (1.3 eq.), TMS-OTf (1.1 eq.), 80 °C, 2 h, 60%.

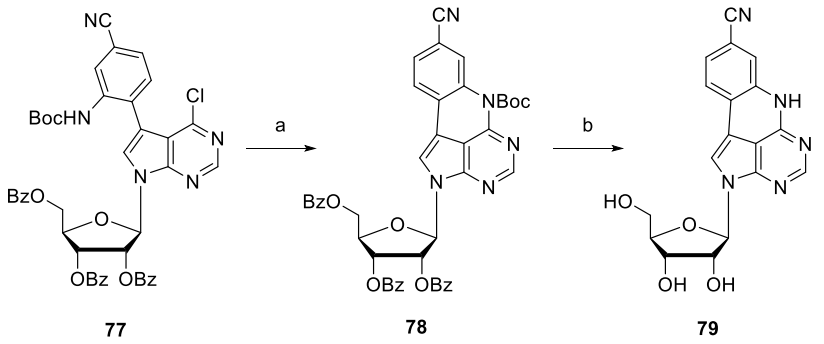
The borylation of **73** using the previously described method in **Papers I–II**, furnished **74** in good yield (76%, Scheme 24). The required 3-amino-4-iodobenonitrile (**75**) was Boc protected using sodium bis(trimethylsilyl)amide (NaHMDS) as a base, followed by the addition of a diluted Boc anhydride solution in THF at -78 °C to avoid forming doubly protected material, thus yielding **76**. Compounds **74** and **76** were coupled *via* the Suzuki-Miyaura cross-coupling reaction previously described; however, it was performed under strictly anhydrous conditions to mitigate Boc deprotection as observed from using a mixture of acetonitrile and water, in a good yield of **77** (83%).



73 **74** **75** **76** **77**

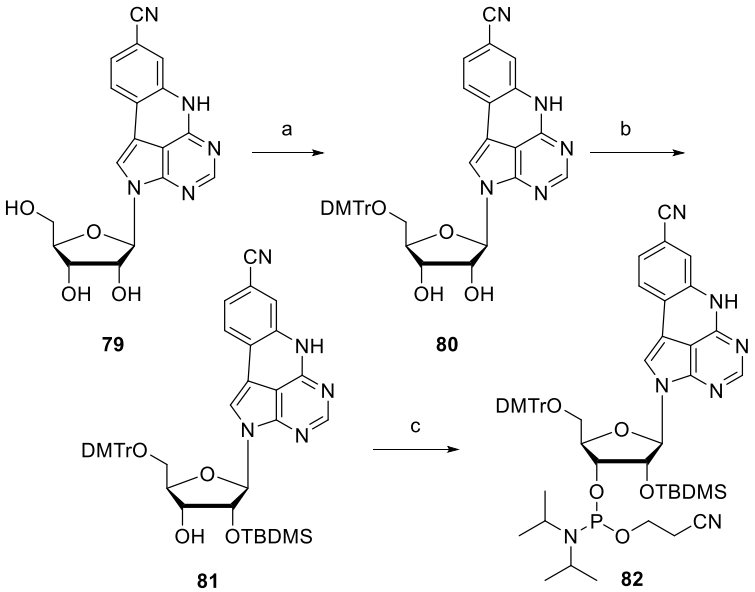
Scheme 24. Synthesis of the advanced intermediate **77** towards a 2CNqA RNA phosphoramidite. Reagents and reaction conditions: (a) $Pd(PPh_3)_4$ (2 mol%), HBPin (1.5 eq.), Et_3N (10 eq.), THF, 80 °C, 36 h, 76%. (b) Boc_2O (1.1 eq.), NaHMDS (2 eq.), THF, -78 °C, 1 h, 81%. (c) Compound **76** (1 eq.), **74** (1.5 eq.) $PdCl_2(PPh_3)_2$ (5 mol%), K_2CO_3 (2.5 eq.), DME, 80 °C, 55 h, 83%.

The material was now setup for intramolecular cyclisation according to previously reported conditions.[114] DABCO was added to transform the C6 chloride of **77** to a better leaving group and DBU was used as a sterically hindered base (Scheme 25). The reaction progressed sluggishly as previously reported, however the global deprotection was nearly quantitative, thus resulting in a modest yield (46%) over two steps. Compound **79** was isolated through precipitation and filtration, thereby avoiding the need for flash chromatography.



Scheme 25. Synthesis of the unprotected 2CNqA nucleoside. Reagents and reaction conditions: (a) DBU, DABCO, DMF, 70 °C, 12 h. (b) NaOMe, MeOH, rt, 1 h, 46% over two steps.

The synthesis of the 2CNqA RNA phosphoramidite was concluded by three routine steps (Scheme 26). First, tritylation of **79** using DMTr-Cl in pyridine afforded **80** in good yield (80%). Second, the protection of the 2'-OH in **80** was achieved using TBDMS-Cl, which resulted in a modest yield of **81** (67%). Finally, the phosphitylation of **81** with CEP-Cl provided the 2CNqA RNA phosphoramidite (**82**) in excellent yield (96%).



Scheme 26. Completion of the 2CNqA RNA phosphoramidite. (a) DMTr-Cl, pyridine, rt, 3 h, 80%. (b) TBDMS-Cl, $AgNO_3$, THF, pyridine, rt, 7 h, 67%. (c) CEP-Cl, *N*,*N*-diisopropylethylamine, THF, rt, 20 h, 96%.

4.2 Oligonucleotide chemistry

With the FBAs synthesised as the protected phosphoramidites, they were ready for incorporation into an oligonucleotide. During our synthesis of DNA oligonucleotides 10 nt to 33 nt long, a few issues presented themselves (**Papers II–III**).[82] The modified building blocks required a mixture of MeCN and toluene for solubilisation prior to synthesis, as opposed to the native nucleotides that were dissolved in pure MeCN. The total equivalent of phosphoramidite utilised in each coupling reaction was 25. The DNA oligonucleotides were synthesised according to the standard protocols of an Applied Biosystems (ABI) synthesiser using a coupling time of 60 s for the native nucleotides and an ample time for the modified building blocks (10 mins). The oligonucleotides were synthesised with the final trityl protecting group removed, cleaved from the solid support, and the nucleobases deprotected with concentrated aqueous ammonia at 55 °C for 4 h. The purification was readily achieved by RP-HPLC with a gradient of MeCN and triethylammonium bicarbonate (TEAB) usually obtaining a purity of >95%. All of the DNA oligonucleotides used in this project were synthesised on an ABI 394 automated oligo synthesiser on a 1 µmol scale, with solid support pre-loaded cartridges containing the first nucleotide of the sequence.

For the synthesis of the RNA oligonucleotides an OligoPilot ÄKTA 10 (OP10) was used (**Papers III-V**). Initially, we attempted to work at a similar scale (1 µmole) of synthesis as for the ABI. However, while small cartridges (1–3 µmole) can work with the OP10, we could not manage to design a protocol that lead to a successful and efficient synthesis. Eventually, we opted for a full 32 µmol scale of the desired RNA oligonucleotides. A minimum of three equivalents of FBA phosphoramidite was necessary for optimum coupling efficiency, leading to the expenditure of nearly 100 µmole of FBA phosphoramidite as compared to the DNA synthesis in which 25 µmole was adequate. The overall efficiency, *i.e.* amount of FBA phosphoramidite used to obtain a certain amount of oligonucleotide, was significantly higher for RNA synthesis on the OP10 compared to the DNA synthesis on the ABI. Overall, standard parameters were employed for the RNA oligonucleotide synthesis. The coupling time of the native nucleotides were set to 5 mins and the FBA phosphoramidites to 20 mins.

The RNA oligonucleotides were initially cleaved from the solid support with a prolonged version of the previously used protocol for the DNA

oligonucleotides (55 °C, 12 h). We noticed that prolonged exposure of our RNA oligonucleotides containing a biotin-C6 handle (**Paper V**) to concentrated ammonia led to the degradation of the handle, which forced us to reduce the reaction time to 5 h. Subsequently, the TBDMS protecting groups were cleaved with triethylamine trihydroflouride and the RNA oligonucleotides was precipitated using *n*-butanol.

The purification systems in place did not allow for a straightforward purification of 32 µmol of oligonucleotides, in fact, no system was prepared for RNA purification. Strict restrictions were placed on buffer preparation to avoid contamination by RNase that degrade RNA. A RP-HPLC system was created from unused components and a new C18 column with small particle and pore sizes (5 µm, 130 Å) was purchased. The price of a large C18 column that could purify a 32 µmole scale RNA oligonucleotide synthesis made it unfeasible. Each RNA oligonucleotide was split into 4 or 8 µmole portions and purified individually using a gradient comprising MeCN and triethylammonium acetate (TEAA), sometimes requiring several re-purifications to obtain the spectroscopical quality (>95 %), which was ultimately determined by a LC-TOF-MS instrument.

4.3 Photophysical properties of the FBAs

In this section the collected highlights of the photophysical characterisation of the FBAs from **Papers I–IV** is shown. The specialised applications of total internal reflection fluorescence microscopy and 2-photon excitation have been performed by our collaborators (pA, **Paper II**).

4.3.1 Paper I, qAN1-4

The synthesis of the qAN1-4 series yielded two bright probes, qAN1 and qAN4, with a quantum yield of 0.18 and 0.32 respectively, compared to the parent compound qA (0.07). For both qAN1 and qAN4 a red-shift in the absorption of ~30nm was observed and a blue-shift in the emission of 25 and 10nm respectively were observed as compared to qA. However, the absorption and emission of qAN1 and qAN4 remained well separated, minimising self-quenching. A slight degree of sensitivity to polarity was observed for the entire series; the deviating emission spectrum for qAN2 and qAN3 in water was due to poor solubility (Figure 15). The quantum yield varied from 0.13 in methanol to 0.38 in DMSO for qAN1 and 0.12 in DCM to 0.49 in DMSO for qAN4. The relatively high quantum yield and brightness of qAN1 recommended it as a good FBA which we characterised fully in 2017 as the first adenine–adenine DNA FRET pair.[82]

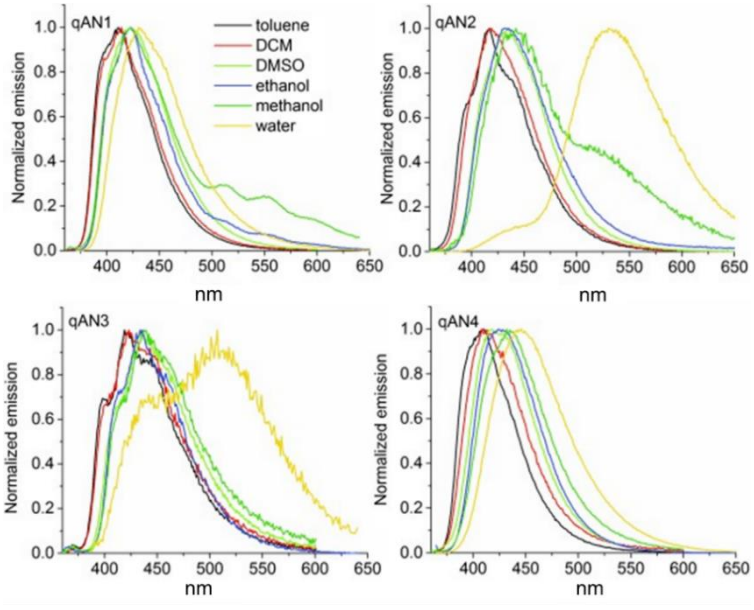


Figure 15. Normalized emission of ethylated qAN1-4 in various solvents.

4.3.2 Paper II, pA-qA$_{nitro}$

During the small library synthesis of the quadracyclic scaffold, we envisioned that an overall increase in molar absorptivity, leading to higher overall brightness, may arise from fusing yet another aromatic ring onto the qA scaffold (Figure 14). The developed synthetic scheme of qAN1 as a DNA phosphoramidite was adopted for pA (**Paper II**). As a monomer, pA had a high brightness (up to 15 200 M^{-1} cm^{-1}), retained high quantum yield across a variety of solvents, $0.64 \leq \Phi_F \leq 0.84$, and a clear absorption to emission band separation. The incorporation of pA into DNA proved it to be an excellent adenine FBA despite its increased size, mainly due to the outer ring positioned in the major groove. From the collected UV-melting experiments a slight increase in duplex stability was observed with a T_m increase of 1.1 °C on average. The brightness of pA within ssDNA of 2,130 M^{-1} cm^{-1} and dsDNA of 1,370 M^{-1} cm^{-1} was on average three-fold higher than for qAN1, which was reported to be one of the brightest FBAs in DNA.[82]

The pA FBA paired nicely with qA$_{nitro}$ in terms of making a new FRET-pair in DNA (Figure 16). For short distances (up to 4 bp separation), the $FRET_{eff}$ is nearly 1.0, *i.e.* when the relative angle between the bases approaches 90°, a local $FRET_{eff}$ minima is obtained at 7 bp separation ($FRET_{eff}$ ~0.15). The $FRET_{eff}$ increases as the bases becomes aligned (8–10 bp separation). Moreover, at longer distances the $FRET_{eff}$ falls close to 0.0 (> 12 bp separation). In conclusion, monitoring the changes in which distance or orientation has a big impact on $FRET_{eff}$, such as at 6 bp separation, would be highly sensitive.

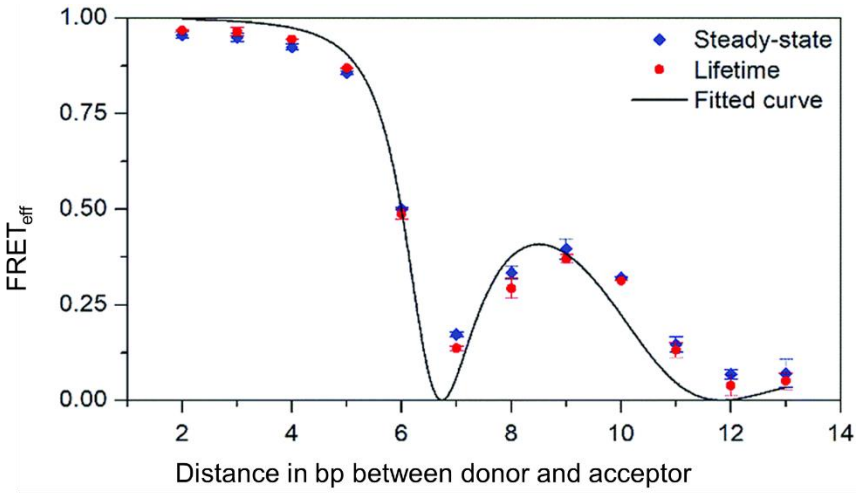


Figure 16. Distance and orientation influencing the $FRET_{eff}$ of the pA-qA_{nitro} FRET pair. Blue diamonds represent $FRET_{eff}$ from steady-state spectroscopy and red circles $FRET_{eff}$ from lifetime data. The black line is the theoretical, calculated $FRET_{eff}$.

4.3.3 Paper III, 2CNqA-qA_{nitro} and 2CNqA-tC_{nitro}

The initial photophysical characterisation of the heterocyclic 2CNqA indicated that it would be interesting to study in an oligonucleotide context due to its high fluorescent quantum yield.[89] We opted to install 2CNqA in both the DNA and RNA contexts to better understand how FRET is affected by their different duplex structures (**Paper III**). As an adenine mimic, 2CNqA stabilised the duplex DNA on average by 3.6 °C, which is significantly higher than pA (1.1 °C on average). Moreover, in one RNA strand, 2CNqA stabilised the duplex RNA by 1.5 °C. The observed circular dichroism spectra were largely unchanged compared to the unmodified oligonucleotide sequences, and thus, 2CNqA is a non-perturbing adenine analogue.

The fluorescent quantum yield of 2CNqA was measured in every possible combination of nearest nucleobase neighbour for ssDNA and dsDNA, and the quantum yield was 0.26 and 0.28 respectively on average. The values in ssDNA ranged from 0.10 to 0.44, while in dsDNA the range was 0.22 to 0.32. The low dependence on sequence surroundings was observed for tC^{70} and tC^{O},[71] and the quenching effect of neighbouring nucleobases when going from ssDNA to dsDNA was observed for qAN1.[82] The high quantum yield and molar absorptivity (~10,000 M^{-1} cm^{-1}) in DNA yielded

brightness values of 4,400 and 3,200 M^{-1} cm^{-1} for ssDNA and dsDNA respectively, producing the highest single-photon excitation brightness values ever reported for an FBA in an oligonucleotide context. In RNA oligonucleotides the 2CNqA FBA showed a somewhat more modest quantum yield, on average 0.12. The higher observed quenching in A-form RNA than B-form DNA is possibly due the lower distance between the nucleobases *i.e.*, higher pi stacking as compared to B-form DNA.

The FRET pairs studied were 2CNqA-qA$_{nitro}$ in DNA (Figure 17, left) and 2CNqA-tC$_{nitro}$ in RNA (Figure 17, right). For the FRET experiments, both the DNA and the RNA 2CNqA probes performed well due to the high overlap integral between donor and acceptor in both instances. The observed high overlap integral of 2CNqA-qA$_{nitro}$ in DNA, especially at long distances, indicated that the FRET pair can report on longer interbase FRET measurements.

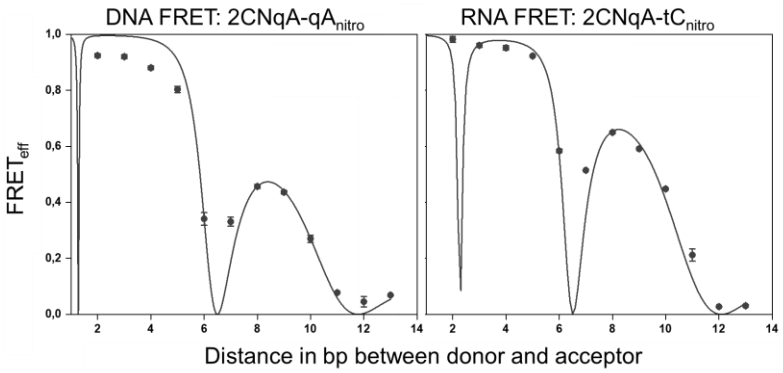


Figure 17. Distance and orientation influencing the FRET$_{eff}$ of the 2CNqA-qA$_{nitro}$ in DNA (left) and 2CNqA-tC$_{nitro}$ in RNA (right). Black dots represent experimentally determined FRET$_{eff}$. The black line is the theoretical, calculated FRET$_{eff}$.

4.3.4 Paper IV, tC^{O}-tC$_{nitro}$

The ribonucleoside tC^{O} was previously synthesised, incorporated into RNA and photophysically characterised.[119] In **Paper IV,** we reported the synthesis of the FRET acceptor ribo-tC$_{nitro}$ and the incorporation of the two probes in RNA, that constituted the first interbase FRET pair in RNA. Like tC^{O}, tC$_{nitro}$ has a stabilising effect on duplex RNA (ΔT_m +2.8 °C on average). The tC^{O}-tC$_{nitro}$ FRET pair perform well in RNA and the FRET$_{eff}$ trend is similar to that of 2CNqA-tC$_{nitro}$ in RNA.

In **Paper IV,** the transition from right–handed A-form RNA to left–handed Z-form RNA was studied using FRET (Figure 18). The Z-form RNA mainly occurs in GC-repeat sequences, providing the tC^{O}-tC_{nitro} FRET pair with an excellent opportunity to study the conformational change. The structure of the Z-form RNA in the literature was used to predict the apparent $FRET_{eff}$ at several virtual donor and acceptor separations (Figure 18, top). In general, our measured $FRET_{eff}$ corresponded well with the published structures of Z-form RNA, except for at 6–8 bp separations. The overall lower measured $FRET_{eff}$ may be because the Z-RNA structure differs, as the literature has used 6 bp GC-repeats, whereas we used 14 bp GC-repeats. The difference between A- and Z-form RNA was 25–87% change in $FRET_{eff}$ for seven out of the eight studied bp separations (Figure 18, bottom). The results clearly indicated that our FRET pair can be used to monitor structural changes in RNA with high sensitivity.

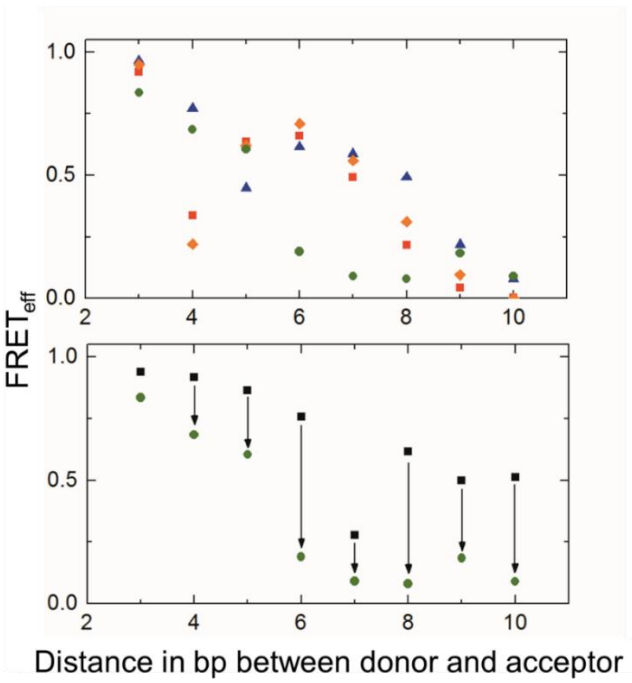


Figure 18. Representation of change in $FRET_{eff}$ from A- to Z-form RNA. Top, green circles indicates measured $FRET_{eff}$, the rest are calculated $FRET_{eff}$ values based on literature data. Bottom, change in measured $FRET_{eff}$ from A-form duplex RNA (black squares) to Z-form duplex RNA (green circles) under high salt conditions (8 M $NaClO_4$).

4.4 RNA interbase-FRET binding interaction assay

The foundation of this PhD project entailed synthesising novel adenine FBAs (**Paper I**). This was followed by the incorporation and characterisation of our probes in DNA, as well as development of new FRET pairs (**Paper II**). The FRET pairs were then compared between DNA and RNA (**Paper III**) and followed up by investigating structural changes in RNA (**Paper IV**). With three of the four milestones achieved, we focused on developing a small molecule binding interaction assay to pre-miR constructs using our non-perturbing RNA FRET pairs (**Paper V**). For achieving the last milestone, a proof-of-concept study was designed in which we decided to work with pre-miR-21, a known oncomiR.[120]

The native pre-miR-21 is a 72 nt long hairpin RNA with several structural features, although, most of the structural features are present in approximately half of the sequence. Furthermore, from our experience of RNA oligonucleotide synthesis combined with the need for several modifications, such as FBA replacement or biotin-labelling, we knew that the synthesis would be problematic. Instead, we simulated the folding of the 72 nt pre-miR-21 sequence and identified exactly where we could obtain the most interesting structural features to study using the FRET pair (Figure 19, left). This led to the design of a 43 nt, truncated pre-mir-21 construct, providing the same hairpin type fold as the RNA sequence (Figure 19, right).

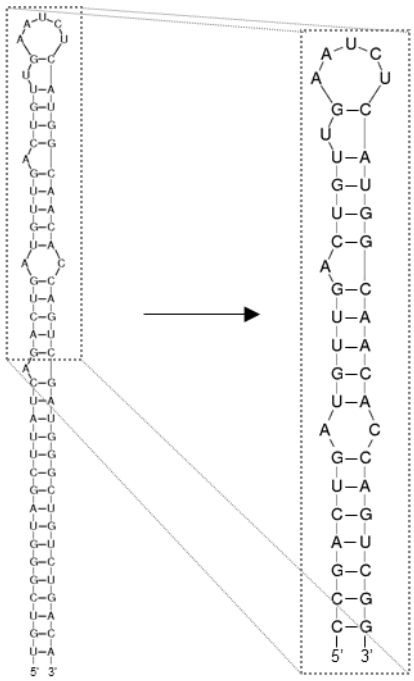


Figure 19. Representation of folded native pre-miR-21 (left) and folded truncated pre-miR-21 (right).

As the assay was intended to utilise FRET pairs, the sequence design and placement of the FRET pairs were of utmost importance. For instance, FRET in DNA between pA–qA$_{nitro}$ (**Paper II**) indicated that 6 bp separation between donor and acceptor would yield approximately 0.50 FRET$_{eff}$, thus creating a sensitive system for changes in the relative distance or orientation of the probes. However, for RNA, a 6 bp separation would result in a FRET$_{eff}$ of 0.75, whereas a 7 base-pair separation would result in FRET$_{eff}$ of 0.50 (0.15 for DNA) (**Papers III–IV**). A complicating factor of FRET pair placement and sequence design is that both the relative distance and orientation of the FRET pair are different in a sequence with a high structure content as compared to a regular duplex A-form RNA. For instance, an abasic site is significantly shorter than a full bp and that any contraction in length results in a twist of the helical structure thus potentially changing the probe orientation, which affects the FRET$_{eff}$. Nonetheless, we aimed for approximately 7 base-pair separations for the placement of our FRET pair in two variations; one with the FRET pair in

the loop region (Figure 20, left) and one with the FRET pair in the stem region around the A-abasic site and A-C mismatch (Figure 20, right).

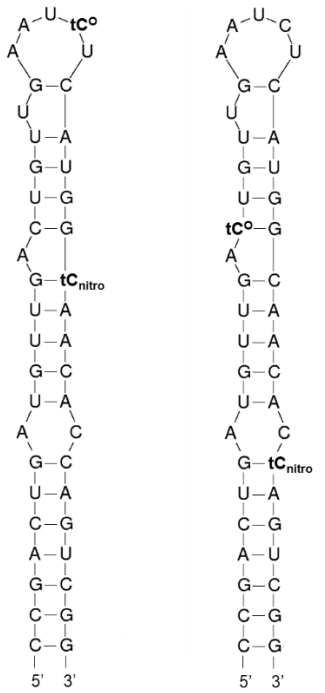


Figure 20. Sequence design of FRET labelled truncated pre-miR-21. a) FRET pair placed to monitor changes in the hairpin. b) FRET pair placed to monitor changes in the stem region.

Initially, we aimed to test a few known RNA binders to the non-labelled truncated pre-miR-21 sequence by ITC. Neomycin has been shown to bind RNA, especially pre-miR-21, making it a good starting point for our assay development.[121] Unfortunately, as the ITC was running, seemingly large enthalpic spikes were observed upon the injection of neomycin to the pre-miR-21 construct. By performing test runs of titrating neomycin to the utilised typical phosphate buffer containing ethylenediaminetetraacetic acid (EDTA), we found non-specific interactions. These non-specific interactions were potentially from a protonation event, due to the observed magnitude of the peaks, occurring between the buffer and neomycin. After screening several buffer conditions, sodium cacodylate 20mM and NaCl 80mM at pH 7.2 provided only small changes in the heat of dilution enthalpy as neomycin was added. The binding of neomycin to truncated pre-miR-21 construct yielded a K_d of 5.2 µM.

SPR was then used to validate our ITC data and neomycin tested at seven different concentrations to biotinylated truncated pre-miR-21 presented a K_d of 2.7 µM. Furthermore, a panel of 11 other aminoglycosides was run against both truncated native pre-miR-21 and truncated FRET labelled pre-miR-21 (Table 1). The obtained dissociation constants showed no major impact of our incorporated FRET pair on the binding affinity of aminoglycosides. The K_d of the twelve aminoglycosides ranged from 3 to 300 µM, which was also the metric used to rank them (SPR Rank, Table 1).

We then tested the binding of all 12 aminoglycosides using our FRET pair. The $FRET_{eff}$ of the RNA sequences before the addition of aminoglycoside was determined and then compared to the $FRET_{eff}$ after addition. In general, a reduction in emission was observed, which represented an increase in $FRET_{eff}$. The experiment was performed for both hairpin and stem FRET labelled RNA oligonucleotide in two different ways. First, aminoglycoside was added to reach 15 µM in the cuvette with 1 µM RNA oligonucleotide, resulting in various changes in emission depending on the aminoglycoside used, as they vary in binding affinity. Based on the dissociation constants from SPR for the 12 aminoglycosides, the required amount of aminoglycoside to reach 90% degree of complexation was calculated and added. The change in emission after adding 15 µM was divided as a fraction of the 90% degree of complexation emission, thus normalising the emission change (Δ-$FRET_{eff}$, Table 1).

Table 1. List of dissociation constants of twelve aminoglycosides based on SPR experiments and change in $FRET_{eff}$ of hairpin- and stem-labelled RNA oligonucleotides

Aminoglycoside	SPR, truncated pre-miR-21 K_d [µM]	Δ-$FRET_{eff}$	SPR Rank	FRET Rank
Neomycin	2.7	0.95	1	1
Sisomicin	3.9	0.69	2	4
Tobramycin	8.1	0.81	3	2
Amikacin	14.8	0.55	4	7
Gentamicin	16.2	0.64	5	6
Apramycin	22.4	0.66	6	5
Netilmicin	28.4	0.73	7	3
Kanamycin	44.2	0.49	8	8
Ribostamycin	90.2	0.32	9	9
Geneticin	101	0.40	10	10
Streptomycin	134	0.13	11	11
Hygromycin	291	-0.01	12	12

The ranking based on obtained dissociation constant values from SPR providing the SPR rank matches well with the ranking based on changes in $FRET_{eff}$, except for amikacin, which for an unknown reason did not impact the relative $FRET_{eff}$ as much as expected from its K_d. None of the aminoglycosides showed a selective binding to either the stem- or hairpin-labelled pre-miR-21 construct (Figure 20). Instead, only unspecific interactions were observed, supported by SPR data (binding stoichiometry of 2 to 4 aminoglycosides per RNA oligonucleotide).

In summary, a novel RNA small molecule binding interaction assay based on interbase FRET between tC^{O}–tC_{nitro} was developed. It was capable of distinguishing between good and poor binders of a small library of aminoglycosides. This final study fulfils the last milestone and the overall goal of the PhD project planned five years ago.

5. Concluding remarks

This thesis describes the development of novel and highly fluorescent nucleobase analogues, the construction of phosphoramidite building blocks on multi-gram scale for DNA and RNA solid-phase synthesis and a new powerful small molecule binding interaction assay to RNA.

A significant effort was performed early in this project work (**Paper I**) to develop the synthetic tools required for the construction of the FBAs described herein. By making an FBA model, significant synthetic effort was avoided, which greatly accelerated the process to synthesise multicyclic adenine compounds and identify the brightest candidates. Small modifications, such as changing a carbon atom to a nitrogen atom, had a massive impact on the fluorescence quantum yield in water moving from qA (0.07) to qAN1 (0.18), to 2CNqA (0.42) and finally to pA (0.64).

In **Paper II**, the synthetic scheme to access multicyclic adenine derivatives had been significantly improved upon from the previously described synthesis of a qA. The improved synthesis made multi-gram synthesis of phosphoramidite building blocks feasible. Once incorporated into DNA oligonucleotides, pA proved to maintain a relatively high brightness, and functioned well as a FRET donor with the qA_{nitro} as FRET acceptor. However, the most interesting properties were studied by our collaborators that identified pA as the brightest two-photon excitable FBA to date and together with its other properties, pA is a very promising molecule for super resolution experiments and imaging applications.

As my goal was to develop an RNA based molecular binding interaction assay, the chemistry to generate RNA phosphoramidite building blocks was pursued (**Papers III-IV**). The synthesis of these compounds proved difficult and, unfortunately, we could not adopt our high yielding synthetic scheme from the DNA synthesis to accommodate other glycosylation's except using deoxyribose derivatives and thus we had to develop a new synthetic pathway for the 2CNqA (**Paper III**).

Gratefully, I was able to follow through with the initial PhD project description and have constructed a small molecule RNA binding interaction assay based on our developed FRET pairs (**Paper V**). By placing FRET pairs in a structurally defined RNA sequence, we could accurately report on the relative binding affinities within a series of

aminoglycosides. This study marks the beginning for internucleobase FRET applications in RNA and holds great promise for future studies.

In conclusion, our group has done a tremendous research effort in bringing new fluorescent nucleobase analogues to the scientific community and the applications of them has only just started to emerge.

Acknowledgements

I would like to express my utmost gratitude to all who has made this thesis possible and that has supported me through the years. Special thanks go to,

My main supervisor **Morten Grøtli**, for taking me in to your research group six years ago and your mentorship from BSc to PhD, you have been absolutely fantastic.

My co-supervisor **Marcus Wilhelmsson**, for the monumental help in all aspects of physical chemistry.

My industrial co-supervisors **Malin Lemurell**, for making this project possible and **Anders Dahlén** for being incredibly helpful with the oligonucleotide parts.

My examiner **Kristina Luthman**, for helping me develop as a chemist.

Anders, **Christopher**, **Erik** and **Tom** for all the joint struggle in making these *nucleoside nightmares*.

Jesper and **Moa** for teaching me so much spectroscopy in so little time.

All former and current collaboration members of the FBA project from the **Grøtli** and **Wilhelmsson** groups that I have worked with, **Blaise**, **Sangamesh**, **Pauline**, **Tristan**, **Anna**, and **Vinoth**.

Cassandra for proof-reading this thesis.

Our international collaborators, especially **Jong Jin Ro**, **Byeang H. Kim**, **Rachel Fischer**, **Steven Magennis**, **Anita Jones**, **Afaf El-Sagheer** and **Tom Brown**.

My colleagues from the CVRM New Modalities team, KK4, now KK2, for welcoming me and teaching me the AZ way, and for being fabulous lab mates, **Gradén**, **Maria S.**, **Maria Ö.**, **Broddan**, **Jocke**, **Rouven**, **Peter**, **Martin**, **Johan**, **Laurent**, **Ahlke**, **Tom**, **Andrei** and **Bill**.

Everyone at **SSL** and **SAM** for teaching and assisting me in oligonucleotide purification and analysis, especially **Linda** and **Thomas**.

Everyone in the **AZ Youth Club** and all the people at **Floor 5** for fun lunches and good discussions.

My **family** and **friends**, for all the encouragement.

The incredible support from the love of my life, **Tove**, what a journey we have done!

Last and least, little baby **Hugo** who will not remember anything from this era but whom I love so.

References and notes

1. Robinson, V. L., Rethinking the central dogma: Noncoding RNAs are biologically relevant. *Urol. Oncol. Semin. Ori.* **2009,** *27* (3), 304-306.
2. Lee, R. C.; Feinbaum, R. L.; Ambros, V., The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **1993,** *75* (5), 843-854.
3. Wightman, B.; Ha, I.; Ruvkun, G., Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **1993,** *75* (5), 855-862.
4. Kozomara, A.; Griffiths-Jones, S., miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014,** *42* (Database issue), D68-73.
5. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C., Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol.* **2018,** *9*, 402-402.
6. Lin, S.; Gregory, R. I., MicroRNA biogenesis pathways in cancer. *Nat. Rev. Cancer* **2015,** *15*, 321-333.
7. Li, Y.; Kowdley, K. V., MicroRNAs in Common Human Diseases. *Genomics Proteomics Bioinformatics* **2012,** *10* (5), 246-253.
8. Croce, C. M., Causes and consequences of microRNA dysregulation in cancer. *Nat. Rev. Genet.* **2009,** *10* (10), 704-714.
9. Iorio, M. V.; Croce, C. M., MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol. Med.* **2012,** *4* (3), 143-159.
10. Matsui, M.; Corey, D. R., Non-coding RNAs as drug targets. *Nat. Rev. Drug Discov.* **2017,** *16* (3), 167-179.
11. Titze-de-Almeida, R.; David, C.; Titze-de-Almeida, S. S., The Race of 10 Synthetic RNAi-Based Drugs to the Pharmaceutical Market. *Pharm. Res.* **2017,** *34* (7), 1339-1363.
12. Frazier, K. S., Antisense oligonucleotide therapies: the promise and the challenges from a toxicologic pathologist's perspective. *Toxicol. Pathol.* **2015,** *43* (1), 78-89.
13. Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L., Principles of early drug discovery. *Br. J. Pharmacol.* **2011,** *162* (6), 1239-1249.
14. Disney, M. D., Targeting RNA with Small Molecules To Capture Opportunities at the Intersection of Chemistry, Biology, and Medicine. *J. Am. Chem. Soc.* **2019,** *141* (17), 6776-6790.
15. A. Lorenz, D.; L. Garner, A., Approaches for the Discovery of Small Molecule Ligands Targeting microRNAs. RNA Therapeutics, Topics in Medicinal Chemistry **2017**, *27*, 79-110.

16. Morris, K. V.; Mattick, J. S., The rise of regulatory RNA. *Nat. Rev. Genet.* **2014,** *15*, 423-437.
17. van Dam, L.; Levitt, M. H., BII Nucleotides in the B and C Forms of Natural-sequence Polymeric DNA: A New Model for the C Form of DNA. *J. Mol. Biol.* **2000,** *304* (4), 541-561.
18. Premilat, S.; Albiser, G., A new D-DNA form of poly(dA-dT).poly(dA-dT): an A-DNA type structure with reversed Hoogsteen pairing. *Eur. Biophys. J.* **2001,** *30* (6), 404-410.
19. Franklin, R. E.; Gosling, R. G., The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Cryst.* **1953,** *6* (8-9), 673-677.
20. Mitsui, Y.; Langridge, R.; Shortle, B. E.; Cantor, C. R.; Grant, R. C.; Kodama, M.; Wells, R. D., Physical and Enzymatic Studies on Poly d(I–C).Poly d(I–C), an Unusual Double-helical DNA. *Nature* **1970,** *228* (5277), 1166-1169.
21. Pohl, F. M.; Jovin, T. M., Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.* **1972,** *67* (3), 375-396.
22. Voet, D.; Voet, J. G., *Biochemistry*. John Wiley & Sons: Hoboken, NJ, **2011**.
23. Filipowicz, W.; Hohn, T., *Post-Transcriptional Control of Gene Expression in Plants*. Springer Netherlands: Dordrecht, **1996**.
24. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nat. Rev. Drug Discov.* **2002,** *1* (9), 727-730.
25. Ezkurdia, I.; Juan, D.; Rodriguez, J. M.; Frankish, A.; Diekhans, M.; Harrow, J.; Vazquez, J.; Valencia, A.; Tress, M. L., Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum. Mol. Genet.* **2014,** *23* (22), 5866-5878.
26. Crooke, S. T., Molecular Mechanisms of Antisense Oligonucleotides. *Nucleic acid Ther.* **2017,** *27* (2), 70-77.
27. Stein, C. A.; Castanotto, D., FDA-Approved Oligonucleotide Therapies in 2017. *Mol. Ther.* **2017,** *25* (5), 1069-1075.
28. Ottesen Eric, W., ISS-N1 makes the first FDA-approved drug for spinal muscular atrophy. In *Transl. Neurosci.* **2017**, *8*, 1-6.
29. Paik, J.; Duggan, S., Volanesorsen: First Global Approval. *Drugs* **2019,** *79* (12), 1349-1354.
30. Rinaldi, C.; Wood, M. J. A., Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* **2017**, *14*, 9-21.

31. Stevens, D. L.; Dotter, B.; Madaras-Kelly, K., A review of linezolid: the first oxazolidinone antibiotic. *Expert Rev. Anti Infect. Ther.* **2004,** *2* (1), 51-59.
32. Thomas, J. R.; Hergenrother, P. J., Targeting RNA with Small Molecules. *Chem. Rev.* **2008,** *108* (4), 1171-1224.
33. Guan, L.; Disney, M. D., Recent Advances in Developing Small Molecules Targeting RNA. *ACS Chem. Biol.* **2012,** *7* (1), 73-86.
34. Warner, K. D.; Hajdin, C. E.; Weeks, K. M., Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **2018,** *17*, 547-558.
35. Blakeley, B. D.; DePorter, S. M.; Mohan, U.; Burai, R.; Tolbert, B. S.; McNaughton, B. R., Methods for identifying and characterizing interactions involving RNA. *Tetrahedron* **2012,** *68* (43), 8837-8855.
36. Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C.; Meyer, B.; Oschkinat, H.; Peng, J.; Schwalbe, H.; Siegal, G., Perspectives on NMR in drug discovery: a technique comes of age. *Nat. Rev. Drug Discov.* **2008,** *7* (9), 738-745.
37. Deng, G.; Sanyal, G., Applications of mass spectrometry in early stages of target based drug discovery. *J. Pharm. Biomed. Anal.* **2006,** *40* (3), 528-538.
38. Cunningham, B. T.; Laing, L. G., Advantages and application of label-free detection assays in drug screening. *Expert Opin. Drug Dis.* **2008,** *3* (8), 891-901.
39. Freyer, M. W.; Lewis, E. A., Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. *Methods Cell Biol.* **2008**, *84*, 79-113.
40. Saboury, A. A., A review on the ligand binding studies by isothermal titration calorimetry. *J. Iran. Chem. Soc.* **2006,** *3* (1), 1-21.
41. Chaires, J. B., Energetics of drug-DNA interactions. *Biopolymers* **1997,** *44* (3), 201-15.
42. Feig, A. L., Applications of isothermal titration calorimetry in RNA biochemistry and biophysics. *Biopolymers* **2007,** *87* (5-6), 293-301.
43. Muller, M.; Weigand, J. E.; Weichenrieder, O.; Suess, B., Thermodynamic characterization of an engineered tetracycline-binding riboswitch. *Nucleic Acids Res.* **2006,** *34* (9), 2607-2617.
44. Kaul, M.; Pilch, D. S., Thermodynamics of aminoglycoside-rRNA recognition: the binding of neomycin-class

aminoglycosides to the A site of 16S rRNA. *Biochemistry* **2002,** *41* (24), 7695-7706.
45. Bernacchi, S.; Freisz, S.; Maechling, C.; Spiess, B.; Marquet, R.; Dumas, P.; Ennifar, E., Aminoglycoside binding to the HIV-1 RNA dimerization initiation site: thermodynamics and effect on the kissing-loop to duplex conversion. *Nucleic Acids Res.* **2007,** *35* (21), 7128-7139.
46. Nguyen, H. H.; Park, J.; Kang, S.; Kim, M., Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications. *Sensors* **2015,** *15* (5), 10481-10510.
47. Vo, T.; Paul, A.; Kumar, A.; Boykin, D. W.; Wilson, W. D., Biosensor-surface plasmon resonance: A strategy to help establish a new generation RNA-specific small molecules. *Methods* **2019**, in press, 10.1016/j.ymeth.2019.05.005.
48. Wong, C.-H.; Liang, F.-S., Surface Plasmon Resonance Study of RNA–Aminoglycoside Interactions. In *Methods Enzymol.* **2003**, *362*, 340-353.
49. Nakatani, K.; Horie, S.; Goto, Y.; Kobori, A.; Hagihara, S., Evaluation of mismatch-binding ligands as inhibitors for Rev-RRE interaction. *Bioorg. Med. Chem.* **2006,** *14* (15), 5384-5388.
50. Fukuzumi, T.; Murata, A.; Aikawa, H.; Harada, Y.; Nakatani, K., Exploratory Study on the RNA-Binding Structural Motifs by Library Screening Targeting pre-miRNA-29 a. *Chem. Eur. J.* **2015,** *21* (47), 16859-16867.
51. Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D., MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **2014,** *1077*, 101-113.
52. Moon, M. H.; Hilimire, T. A.; Sanders, A. M.; Schneekloth, J. S., Jr., Measuring RNA-Ligand Interactions with Microscale Thermophoresis. *Biochemistry* **2018,** *57* (31), 4638-4643.
53. Wicks, S. L.; Hargrove, A. E., Fluorescent indicator displacement assays to identify and characterize small molecule interactions with RNA. *Methods* **2019**, in press, 10.1016/j.ymeth.2019.04.018.
54. McCallum, M. M.; Pawlak, A. J.; Shadrick, W. R.; Simeonov, A.; Jadhav, A.; Yasgar, A.; Maloney, D. J.; Arnold, L. A., A fluorescence-based high throughput assay for the determination of small molecule-human serum albumin protein binding. *Anal. Bioanal. Chem.* **2014,** *406* (7), 1867-1875.
55. Tse, W. C.; Boger, D. L., A Fluorescent Intercalator Displacement Assay for Establishing DNA Binding Selectivity and Affinity. *Acc. Chem. Res.* **2004,** *37* (1), 61-69.

56. Zhang, J.; Umemoto, S.; Nakatani, K., Fluorescent Indicator Displacement Assay for Ligand−RNA Interactions. *J. Am. Chem. Soc.* **2010,** *132* (11), 3660-3661.
57. Zheng, J.; Yang, R.; Shi, M.; Wu, C.; Fang, X.; Li, Y.; Li, J.; Tan, W., Rationally designed molecular beacons for bioanalytical and biomedical applications. *Chem. Soc. Rev.* **2015,** *44* (10), 3036-3055.
58. Monroy-Contreras, R.; Vaca, L., Molecular beacons: powerful tools for imaging RNA in living cells. *J. Nucleic Acids* **2011**, ID: 741723.
59. Bose, D.; Jayaraj, G. G.; Kumar, S.; Maiti, S., A Molecular-Beacon-Based Screen for Small Molecule Inhibitors of miRNA Maturation. *ACS Chem. Biol.* **2013,** *8* (5), 930-938.
60. Valeur, E.; Guéret, S. M.; Adihou, H.; Gopalakrishnan, R.; Lemurell, M.; Waldmann, H.; Grossmann, T. N.; Plowright, A. T., New Modalities for Challenging Targets in Drug Discovery. *Angew. Chem. Int. Ed.* **2017,** *56* (35), 10294-10323.
61. Valeur, E.; Jimonet, P., New Modalities, Technologies, and Partnerships in Probe and Lead Generation: Enabling a Mode-of-Action Centric Paradigm. *J. Med. Chem.* **2018,** *61* (20), 9004-9029.
62. Disney, M. D.; Winkelsas, A. M.; Velagapudi, S. P.; Southern, M.; Fallahi, M.; Childs-Disney, J. L., Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules Targeting Structured RNAs. *ACS Chem. Biol.* **2016,** *11* (6), 1720-1728.
63. BRADRICK, T. D.; MARINO, J. P., Ligand-induced changes in 2-aminopurine fluorescence as a probe for small molecule binding to HIV-1 TAR RNA. *RNA* **2004,** *10* (9), 1459-1468.
64. Reuss, A. J.; Gustmann, H.; Braun, M.; Wachtveitl, J.; Segler, A.-L. J.; Gophane, D. B.; Sigurdsson, S. T.; Grünewald, C.; Weigand, J. E., Structure guided fluorescence labeling reveals a two-step binding mechanism of neomycin to its RNA aptamer. *Nucleic Acids Res.* **2018,** *47* (1), 15-28.
65. Xu, W.; Chan, K. M.; Kool, E. T., Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017,** *9* (11), 1043-1055.
66. Wilhelmsson, L. M., Fluorescent nucleic acid base analogues. *Q. Rev. Biophys.* **2010,** *43* (2), 159-183.
67. Tanpure, A. A.; Pawar, M. G.; Srivatsan, S. G., Fluorescent Nucleoside Analogs: Probes for Investigating Nucleic Acid Structure and Function. *Isr. J. Chem.* **2013,** *53* (6-7), 366-378.
68. Ward, D. C.; Reich, E.; Stryer, L., Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside,

2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* **1969,** *244* (5), 1228-1237.
69. Ben Gaied, N.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mely, Y., 8-vinyl-deoxyadenosine, an alternative fluorescent nucleoside analog to 2'-deoxyribosyl-2-aminopurine with improved properties. *Nucleic Acids Res.* **2005,** *33* (3), 1031-1039.
70. Sandin, P.; Wilhelmsson, L. M.; Lincoln, P.; Powers, V. E. C.; Brown, T.; Albinsson, B., Fluorescent properties of DNA base analogue tC upon incorporation into DNA — negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Res.* **2005,** *33* (16), 5019-5025.
71. 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* (1), 157-167.
72. Hawkins, M. E.; Pfleiderer, W.; Balis, F. M.; Porter, D.; Knutson, J. R., Fluorescence Properties of Pteridine Nucleoside Analogs as Monomers and Incorporated into Oligonucleotides. *Anal. Biochem.* **1997,** *244* (1), 86-95.
73. Hawkins, M. E.; Pfleiderer, W.; Jungmann, O.; Balis, F. M., Synthesis and fluorescence characterization of pteridine adenosine nucleoside analogs for DNA incorporation. *Anal. Biochem.* **2001,** *298* (2), 231-240.
74. Rovira, A. R.; Fin, A.; Tor, Y., Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015,** *137* (46), 14602-14605.
75. Beharry, A. A.; Lacoste, S.; O'Connor, T. R.; Kool, E. T., Fluorescence Monitoring of the Oxidative Repair of DNA Alkylation Damage by ALKBH3, a Prostate Cancer Marker. *J. Am. Chem. Soc.* **2016,** *138* (11), 3647-3650.
76. Schmidt, O. P.; Mata, G.; Luedtke, N. W., Fluorescent Base Analogue Reveals T-Hg(II)-T Base Pairs Have High Kinetic Stabilities That Perturb DNA Metabolism. *J. Am. Chem. Soc.* **2016,** *138* (44), 14733-14739.
77. Burns, D. D.; Teppang, K. L.; Lee, R. W.; Lokensgard, M. E.; Purse, B. W., Fluorescence Turn-On Sensing of DNA Duplex Formation by a Tricyclic Cytidine Analogue. *J. Am. Chem. Soc.* **2017,** *139* (4), 1372-1375.
78. Fisher, R. S.; Nobis, D.; Füchtbauer, A. F.; Bood, M.; Grøtli, M.; Wilhelmsson, L. M.; Jones, A. C.; Magennis, S. W., Pulse-

shaped two-photon excitation of a fluorescent base analogue approaches single-molecule sensitivity. *PCCP* **2018,** *20* (45), 28487-28498.
79. Bood, M.; Sarangamath, S.; Wranne, M. S.; Grøtli, M.; Wilhelmsson, L. M., Fluorescent nucleobase analogues for base-base FRET in nucleic acids: synthesis, photophysics and applications. *Beilstein J. Org. Chem.* **2018,** *14*, 114-129.
80. 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* (12), 4288-4293.
81. Yamamoto, S.; Han, J. H.; Park, S.; Sugiyama, H., Development of a vivid FRET system based on a highly emissive dG-dC analogue pair. *Chem. Eur. J.* **2017**, *23* (31), 7607-7613.
82. 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* (27), 9271-9280.
83. Smith, M. B. **2018**, *Organic Synthesis 4th Edition*.
84. Burke, M. D.; Schreiber, S. L., A Planning Strategy for Diversity-Oriented Synthesis. *Angew. Chem. Int. Ed.* **2004,** *43* (1), 46-58.
85. Lawson, C. P.; Füchtbauer, A. F.; Wranne, M. S.; Giraud, T.; Floyd, T.; Dumat, B.; Andersen, N. K.; H. El-Sagheer, A.; Brown, T.; Gradén, H.; Wilhelmsson, L. M.; Grøtli, M., Synthesis, oligonucleotide incorporation and fluorescence properties in DNA of a bicyclic thymine analogue. *Sci. Rep.* **2018,** *8* (1), ID: 13970.
86. Matarazzo, A.; Hudson, R. H. E., Fluorescent adenosine analogs: a comprehensive survey. *Tetrahedron* **2015,** *71* (11), 1627-1657.
87. Kazimierczuk, Z.; Cottam, H. B.; Revankar, G. R.; Robins, R. K., Synthesis of 2'-deoxytubercidin, 2'-deoxyadenosine, and related 2'-deoxynucleosides via a novel direct stereospecific sodium salt glycosylation procedure. *J. Am. Chem. Soc.* **1984,** *106* (21), 6379-6382.
88. Niedballa, U.; Vorbrueggen, H., Synthesis of nucleosides. 12. General synthesis of N-glycosides. Iv. Synthesis of nucleosides of hydroxy and mercapto nitrogen heterocycles. *J. Org. Chem.* **1974,** *39* (25), 3668-3671.
89. Foller Larsen, A.; Dumat, B.; Wranne, M. S.; Lawson, C. P.; Preus, S.; Bood, M.; Gradén, H.; Marcus Wilhelmsson, L.; Grøtli, M., Development of bright fluorescent quadracyclic adenine analogues: TDDFT-calculation supported rational design. *Sci. Rep.* **2015,** *5*, ID: 12653.

90. Dyrager, C.; Börjesson, K.; Dinér, P.; Elf, A.; Albinsson, B.; Wilhelmsson, L. M.; Grøtli, M., Synthesis and Photophysical Characterisation of Fluorescent 8-(1H-1,2,3-Triazol-4-yl)adenosine Derivatives. *Eur. J. Org. Chem.* **2009,** *10*, 1515-1521.
91. Vorbrüggen, H. **2004**, Synthesis Of Nucleosides. In *Organic Reactions*, pp 1-630.
92. Diekmann, E.; Friedrich, K.; Fritz, H.-G., Didesoxy-Ribonucleoside durch Schmelzkondensation. *J. Prakt. Chem./Chem. Ztg.* **1993,** *335* (5), 415-424.
93. Witkowski, J. T.; Robins, R. K.; Sidwell, R. W.; Simon, L. N., Design, synthesis, and broad spectrum antiviral activity of 1-.beta.-D-ribofuranosyl-1,2,4-triazole-3-carboxamide and related nucleosides. *J. Med. Chem.* **1972,** *15* (11), 1150-1154.
94. Fischer, E.; Helferich, B., Synthetische Glucoside der Purine. *Ber. Dtsch. Chem. Ges.* **1914,** *47* (1), 210-235.
95. Davoll, J.; Lowy, B. A., A New Synthesis of Purine Nucleosides. The Synthesis of Adenosine, Guanosine and 2,6-Diamino-9-β-D-ribofuranosylpurine1. *J. Am. Chem. Soc.* **1951,** *73* (4), 1650-1655.
96. Johnson, T. B.; Hilbert, G. E., THE SYNTHESIS OF PYRIMIDINE-NUCLEOSIDES. *Science* **1929,** *69* (1796), 579-580.
97. Niedballa, U.; Vorbrüggen, H., A General Synthesis of Pyrimidine Nucleosides. *Angew. Chem. Int. Edit.* **1970,** *9* (6), 461-462.
98. Michelson, A. M.; Todd, A. R., Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3′: 5′-internucleotidic linkage. *J. Chem. Soc.* **1955,** (0), 2632-2638.
99. Garegg, P. J.; Lindh, I.; Regberg, T.; Stawinski, J.; Strömberg, R.; Henrichson, C., Nucleoside H-phosphonates. III. Chemical synthesis of oligodeoxyribonucleotides by the hydrogenphosphonate approach. *Tetrahedron Lett.* **1986,** *27* (34), 4051-4054.
100. Gilham, P. T.; Khorana, H. G., Studies on Polynucleotides. I. A New and General Method for the Chemical Synthesis of the C5″-C3″ Internucleotidic Linkage. Syntheses of Deoxyribo-dinucleotides1. *J. Am. Chem. Soc.* **1958,** *80* (23), 6212-6222.
101. Letsinger, R. L.; Ogilvie, K. K., Nucleotide chemistry. XIII. Synthesis of oligothymidylates via phosphotriester intermediates. *J. Am. Chem. Soc.* **1969,** *91* (12), 3350-3355.
102. Beaucage, S. L.; Caruthers, M. H., Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981,** *22* (20), 1859-1862.

103. Reese, C. B., Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Organic & Biomolecular Chemistry* **2005,** *3* (21), 3851-3868.
104. Huddler, D., Zartler, R. E. **2017**, Thermodynamics in Drug Discovery. In *Applied Biophysics for Drug Discovery*, pp 7-28.
105. Drescher, D. G.; Ramakrishnan, N. A.; Drescher, M. J., Surface plasmon resonance (SPR) analysis of binding interactions of proteins in inner-ear sensory epithelia. *Methods in molecular biology (Clifton, N.J.)* **2009,** *493*, 323-343.
106. Goebel-Stengel, M.; Stengel, A.; Taché, Y.; Reeve, J. R., The importance of using the optimal plasticware and glassware in studies involving peptides. *Anal. Biochem.* **2011,** *414* (1), 38-46.
107. Shin, D.; Sinkeldam, R. W.; Tor, Y., Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2011,** *133* (38), 14912-14915.
108. Godde, F.; Toulme, J. J.; Moreau, S., Benzoquinazoline derivatives as substitutes for thymine in nucleic acid complexes. Use of fluorescence emission of benzo[g]quinazoline-2,4-(1H,3H)-dione in probing duplex and triplex formation. *Biochemistry* **1998,** *37* (39), 13765-13775.
109. Okamoto, A.; Kanatani, K.; Saito, I., Pyrene-labeled base-discriminating fluorescent DNA probes for homogeneous SNP typing. *J. Am. Chem. Soc.* **2004,** *126* (15), 4820-4827.
110. Eason, R. G.; Burkhardt, D. M.; Phillips, S. J.; Smith, D. P.; David, S. S., Synthesis and Characterization of 8-Methoxy-2′-Deoxyadenosine-Containing Oligonucleotides to Probe the Syn Glycosidic Conformation of 2′-deoxyadenosine Within DNA. *Nucleic Acids Res.* **1996,** *24* (5), 890-897.
111. Millen, A. L.; McLaughlin, C. K.; Sun, K. M.; Manderville, R. A.; Wetmore, S. D., Computational and Experimental Evidence for the Structural Preference of Phenolic C-8 Purine Adducts. *J. Phys. Chem.* **2008,** *112* (16), 3742-3753.
112. Dierckx, A.; Dinér, P.; El-Sagheer, A. H.; Kumar, J. D.; Brown, T.; Grøtli, M.; Wilhelmsson, L. M., Characterization of photophysical and base-mimicking properties of a novel fluorescent adenine analogue in DNA. *Nucleic Acids Res.* **2011,** *39* (10), 4513-4524.
113. Seela, F.; Thomas, H., Duplex Stabilization of DNA: Oligonucleotides containing 7-substituted 7-deazaadenines. *Helv. Chim. Acta* **1995,** *78* (1), 94-108.
114. Dierckx, A.; Miannay, F.-A.; Ben Gaied, N.; Preus, S.; Björck, M.; Brown, T.; Wilhelmsson, L. M., Quadracyclic Adenine: A Non-Perturbing Fluorescent Adenine Analogue. *Chem. Eur. J.* **2012,** *18* (19), 5987-5997.

115. Zhang, L.; Li, Z.; Wang, X.-j.; Yee, N.; Senanayake, C. H., Regioselective Synthesis of Polysubstituted N2-Alkyl/Aryl-1,2,3-Triazoles via 4-Bromo-5-iodo-1,2,3-triazole. *Synlett* **2012,** *23* (07), 1052-1056.
116. Rajeswari, S.; Jones, R. J.; Cava, M. P., A new synthesis of amides from acyl fluorides and N-silylamines. *Tetrahedron Lett.* **1987,** *28* (43), 5099-5102.
117. Magnin, G. C.; Dauvergne, J.; Burger, A.; Biellmann, J.-F., t-Butyldimethylsilyloxymethyl group, a versatile protecting group of adenine. *Tetrahedron Lett.* **1996,** *37* (43), 7833-7836.
118. Hoffer, M., α-Thymidin. *Chem. Ber.* **1960,** *93* (12), 2777-2781.
119. 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* (1), ID: 2393.
120. Feng, Y.-H.; Tsao, C.-J., Emerging role of microRNA-21 in cancer. *Biomed. Rep.* **2016,** *5* (4), 395-402.
121. Yan, H.; Bhattarai, U.; Guo, Z.-F.; Liang, F.-S., Regulating miRNA-21 Biogenesis By Bifunctional Small Molecules. *J. Am. Chem. Soc.* **2017,** *139* (14), 4987-4990.