Base-Functionalized Nucleoside Analog Probes: Design, Synthesis and Applications in Nucleic Acid Labeling and Diagnosis

A thesis

Submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

By

Arun A. Tanpure

Reg. No. 20103050

INDIAN INSTITUTE OF SCIENCE EDUCATION
AND RESEARCH PUNE

2015

This dissertation is dedicated to the memory of my grandparents

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE
(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)
Dr. Homi Bhabha Road, Pashan, Pune–411 008

Dr. Seergazhi G. Srivatsan
Associate Professor, Chemistry

Dr. Seergazhi G. Srivatsan
Associate Professor
Department of Chemistry
IISER, Pune

CERTIFICATE

Certified that the work incorporated in the thesis entitled ***"Base-Functionalized Nucleoside Analog Probes: Design, Synthesis and Applications in Nucleic Acid Labeling and Diagnosis"*** submitted by **Mr. Arun A. Tanpure** was carried out by the candidate under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or Institution.

S.G. Srivatsan

Date: September 7, 2015
Place: Pune

Dr. Seergazhi G. Srivatsan

Tel.: +91-20-25908086 Fax: +91-20-25908186 E-mail: srivatsan@iiserpune.ac.in

DECLARATION

I declare that this written submission represents my ideas in my own words and where others ideas have included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of Dr. S. G. Srivatsan.

Date: September 7, 2015

Place: Pune

Mr. Arun Ankush Tanpure

Reg. No. 20103050

Acknowledgements

It gives me immense pleasure to express my sincere gratitude to my research advisor Dr. S. G. Srivatsan for his persistent and generous support throughout my doctoral research. He has provided me an impeccably blended scientific environment of independence and assistance as well made it possible for me to grow as a researcher in a way I would have never imagined. Apart from healthy scientific discussions and debates, he has influenced my personal life in many ways. He is and will always be a very special mentor for me.

I am extremely grateful to Prof. K. N. Ganesh (Director, IISER Pune) for providing the state of art research facilities. Furthermore, I would like to acknowledge my doctoral research advisory committee members Dr. G. J. Sanjayan and Dr. Kundan Sengupta for their critical evaluation through various RAC meetings. I would like to specifically thank Prof. Sanjeev Galande for allowing me to carry out radiolabeling experiments in his lab.

I have been very fortunate to have cordial colleagues like Dr. Maroti, Anupam, Harita, Pooja, Sweta, Siddharth, Pramod, Ashok, Sudeshna, Pragya, Jerrin, Dr. Cornelia, Sangemesh and Manisha. I appreciate their support during various experiments and in maintaining a cheerful environment in the lab all along. I would like to extend my gratitude to all faculty members and research scholars of chemistry department for their candid support and encouragement especially at the beginning of my PhD.

It is my prime duty to acknowledge Council of Scientific and Industrial Research (CSIR) India for my graduate research fellowship. I am particularly grateful to National Institute of Genetics (NIG) Japan for inviting and funding my attendance at *The 8th Sokendai student seminar*, 2011 held in Tsumagoi, Japan. I would like to acknowledge the Department of Science and Technology (DST), Deutsche Forschungsgemeinschaft (DFG) and the Lindau council for selecting me as one of the participants and providing financial support for attending The *63rd Lindau Nobel Laureate Meeting* in 2013. I sincerely thank Royal Society of Chemistry (RSC) for inviting me to the headquarters at Cambridge, UK as an intern to participate in the science publishing program with entire financial support.

I am particularly indebted to my parents and my entire family, for their absolute love, blessings and sacrifices. The values and virtues they have instilled have always kept me in high spirit and made me capable to overcome various professional challenges.

Arun Ankush Tanpure

Chapter 1 is a reprint of: A. A. Tanpure, M. G. Pawar, S. G. Srivatsan, Fluorescent nucleoside analogs: probes for investigating nucleic acid structure and function. *Isr. J. Chem.* **2013**, *53*, 366–378.
The dissertation author is the main author and researcher for this work.

Chapter 2A is a reprint of: A. A. Tanpure, S. G. Srivatsan, A microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA abasic site. *Chem. Eur. J.* **2011**, *17*, 12820–12827.
The dissertation author is the main author and researcher for this work.

Chapter 2B is a reprint of: A. A. Tanpure, S. G. Srivatsan, Synthesis and photophysical characterization of a fluorescent nucleoside analogue that signals the presence of an abasic site in RNA. *ChemBioChem* **2012**, *13*, 2392–2399.
The dissertation author is the main author and researcher for this work.

Chapter 3A is a reprint of: A. A. Tanpure, S. G. Srivatsan, Conformation-sensitive nucleoside analogs as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, *doi: 10.1093/nar/gkv743.*
The dissertation author is the main author and researcher for this work.

Chapter 4 is a reprint of: A. A. Tanpure, S. G. Srivatsan, Synthesis, photophysical properties and incorporation of a highly emissive and environment-sensitive uridine analogue based on the lucifer chromophore. *ChemBioChem* **2014**, *15*, 1309–1316.
The dissertation author is the main author and researcher for this work.

Chapter 5 is a reprint of: A. A. Sawant, A. A. Tanpure, P. P. Mukherjee, S. Athavale, A. Kelkar, S. Galande, S. G. Srivatsan, A versatile toolbox for posttranscriptional chemical labeling and Imaging of RNA. *Nucleic Acid Res.* **2015** *(Just accepted)*
The dissertation author is one of the main authors and researcher for this work.

Table of content

Chapter 2: Detection of abasic sites in DNA and RNA oligonucleotide by using fluorescent nucleoside analogs

Chapter 2A: 5-Benzofuran-modified uridine as a probe for the fluorescence detection of DNA abasic site

Chapter 2B: 5-Benzofuran-modified 2'-deoxyuridine as a probe for the fluorescence detection of RNA abasic site

Chapter 3: Fluorescent nucleoside analogs as probes for study the DNA and RNA non-canonical structures

Chapter 3A: 5-Benzofuran-modified nucleoside analogs as topology-specific probes for human telomeric DNA and RNA G-quadruplex structures

Chapter 3B: Benzofuran-conjugated fluorescent uridine analog to probe the conformational switch in i-motif forming human telomeric DNA and RNA repeats

Abbreviations

ε	molar extinction coefficient
μL	microliter
μM	micro molar
2-AP	2-aminopurine
A	adenine
Abs	absorption
ACN	acetonitrile
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BODIPY	boron-dipyrromethene
BrU	5-Bromouridine
C	cytosine
CD	circular dichroism
CMP	cytidine monophosphate
CTP	cytidine monophosphate
cP	centipoise
Cu	copper
CuAAC	copper(I)-catalyzed alkyne-azide cycloaddition
DAPI	4', 6-diamidino-2-phenylindole
DEAE	diethylaminoethyl
DMAP	4-Dimethylaminopyridine
DMSO	N, N-dimethyl sulfoxide
DMF	Dimethylformamide
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
ds	double stranded
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
em	emission
EPR	electron paramagnetic resonance
EdU	ethynyl deoxyuridine
EU	ethynyl uridine
FI	fluorescence intensity
FRET	förster resonance energy transfer
G	guanine
GFP	green fluorescent protein
GMP	guanosine monophosphate
GQ	G-quadruplex
GTP	guanosine triphosphate
HPA	hydroxylpiccolinic acid
HPLC	high performance liquid chromatography
h-Telo	human telomeric
invDA	Inverse-electron demand Diels–Alder
IR	infrared
in vitro	outside living organism
in vivo	inside living organism
IU	5-Iodouridine
Kd	apparent binding constant
LED	light emitting diode
MAD	multiwavelength anomalous dispersion
MALDI-TOF	matrix assisted laser desorption ionisation-time of flight
max	maximum
MeOH	methanol
$(MeO)_3PO$	trimethyl phosphate

mg	milligram
MHz	megahertz
6-MI	6-methyl isoxanthopterin
mM	milimolar
nm	nanometer
nmol	nanomolar
NMP	nucleoside monophosphate
NMPs	nucleoside monophosphate
NMR	nuclear magnetic resonance
NTPs	nucleoside triphosphates
ON	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
PAH	polycyclic aromatic hydrocarbon
pC	pyrroloC
Pd	palladium
PDS	pyridostatin
$POCl_3$	phosphorus oxychloride
ppm	parts per million
rel	relative
R_f	retention factor
RIPs	ribosome inactivating proteins
RNA	ribonucleic acid
SAD	single-wavelength anomalous dispersion
SNP	single nucleotide polymorphism
SPAAC	strain-promoted alkyne-azide cycloadditions
ss	single stranded
T	thymine
TBDMS	tert-butyldimethylsilyl
TCSPC	time correlated single photon counting
TEAA	triethylammonium acetate
TEAB	triethylammonium bicarbonate
TEG	triethylene glycol
TERRA	telomeric repeat-containing RNA
THF	tetrahydrofuron
THPTA	tris-(3-hydroxypropyl triazolylmethyl)amine
TLC	thin layer chromatography
T_m	thermal melting
TOM	triisopropylsilyloxymethy
Tris	tris (hydroxymethyl) amino methane
U	uridine
UMP	uridine monophosphate
UTP	uridine triphosphate
UV	ultraviolet
WC	Watson-Crick

Synopsis

Base-Functionalized Nucleoside Analog Probes: Design, Synthesis and Applications in Nucleic Acid Labeling and Diagnosis

Background and Aim: Considering the vast therapeutic potential of nucleic acids, several biophysical methods and chemical probes have been and are being developed to advance the fundamental understanding of nucleic acids structure and function, and their interaction with other biomolecules.[1−4] In particular, base-modified fluorescent nucleoside analogs, which are environment-sensitive, are very useful in studying the structure and recognition properties of nucleic acids.[5−9] Although several structurally diverse nucleoside analogs exhibiting probe-like properties have been developed over the years, majority of these analogs have absorption maximum in the UV region and importantly, display drastic fluorescence quenching upon incorporation into nucleic acids. Due to these shortcomings many of these analogs cannot be easily implemented in nucleic acid analysis and in cell based studies.[10] Therefore, it is envisaged that the development of robust environment-sensitive nucleoside analogs with photophysical properties suitable for *in vitro* and in cell analysis will greatly advance the fundamental understanding of nucleic acid structure and function. Such nucleoside analogs will also be highly useful in devising new nucleic acid-based diagnostic tools.

Solid-phase oligonucleotide (ON) synthesis protocols and enzymatic methods using polymerases and ligases are now being routinely used for the synthesis of modified RNA ONs for a variety of applications. However, in several instances, the synthesis of modified phosphoramidite substrate involves several complicated chemical manipulations and or the modified substrate does not survive the chemical conditions used in the solid-phase ON synthesis protocol. On the other hand enzymatic incorporation of modified triphosphate substrate by *in vitro* transcription reaction works under mild conditions. However, many times the modified triphosphate substrate does not serve as a good substrate for the RNA polymerase enzyme. Alternatively, postsynthetic modification by using bioorthogonal chemical reactions such as azide-alkyne cycloaddition and azide-phosphine Staudinger ligation reactions have emerged as a valuable method to label nucleic acids.[11-13] However, RNA labeling using these reactions is not straightforward as methods that have been established for DNA often do not work for RNA due to its inherent instability.[14] Moreover, the azide group, which participates in a wider range of bioorthogonal reactions in comparison

to the alkyne functionality, cannot be easily incorporated into RNA as it undergoes Staudinger-type reaction with phosphoramidite substrates used in the solid-phase method.[15] Therefore, development of a milder method to incorporate azide functionality into RNA would provide easy access to azide-labeled RNA for further postsynthetic functionalization in a modular fashion by click and Staudinger reactions.

This thesis illustrates the design, synthesis and photophysical characterization of environment-sensitive fluorescent nucleoside analogs obtained by attaching heterocyclic rings onto the 5-position of a pyrimidine base. Some of these analogs are moderately emissive and are highly sensitive to solvent polarity and viscosity changes. Interestingly, these analogs when incorporated into DNA and RNA ONs were minimally perturbing and uniquely, retained appreciable fluorescence. Utilizing the probe-like properties of these analogs we have devised simple fluorescence-based assays to detect abasic sites in nucleic acids and discriminate different nucleic acid topologies (e.g., DNA and RNA G-quadruplexes (GQs) and i-motifs).

This thesis also illustrates the development of a simple and practical posttranscriptional chemical labeling method for RNA by using bioorthogonal reactions. In this strategy, we have used novel azide-modified uridine triphosphates to incorporate the azide functionality into RNA by transcription reaction. The azide-modified RNA transcripts are conveniently labeled with a variety of biophysical probes in a modular fashion by copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC), copper-free strain-promoted azide-alkyne cycloaddition (SPAAC) and azide-phosphine Staudinger ligation reaction. Taken together, straightforward synthesis, easy incorporation into nucleic acids and probe-like properties highlight the potential of these novel functionalized nucleoside analogs as efficient tools to study the structure and function of nucleic acids both *in vitro* and in cells.

The thesis is organized as follows:

Chapter 1: Base-modified nucleoside analogs in nucleic acid study

In this chapter, a concise literature background on the advancement of base-modified nucleoside analog probes for studying nucleic acid is provided. In particular, the development of structurally minimally perturbing base-modified fluorescent nucleoside probes and their incorporation into nucleic acids for various applications is discussed. In the second part of this chapter, an overview of recent development in the applications of bioorthogonal chemical reactions in labeling and imaging nucleic acids is presented. The

limitations of currently available nucleoside analog probes and methods to label nucleic acids, and the motivation for the present research work are also described.

Chapter 2: Detection of abasic sites in DNA and RNA oligonucleotides by using fluorescent nucleoside analogs

Chapter 2A: 5-Benzofuran-modified uridine as a probe for the fluorescence detection of DNA abasic site

We have synthesized a small series of fluorescent nucleoside analogs by conjugating heterobicycles such as indole, *N*-methyl indole and benzofuran at the 5-position of uridine (Figure 1A). Preliminary phtophysical studies indicated that indole- and *N*-methyl indole-conjugated uridine analogs (^{In}U and ^{MIn}U, respectively) were practically non-emmisive, whereas the benzofuran-conjugated uridine analog (^{Bf}U) was reasonably fluorescent. Detailed photophysical characterization revealed probe-like properties of ^{Bf}U, wherein it was found to be highly sensitive to microenvironment changes such as solvent polarity and viscosity (Figure 1B and 1C).

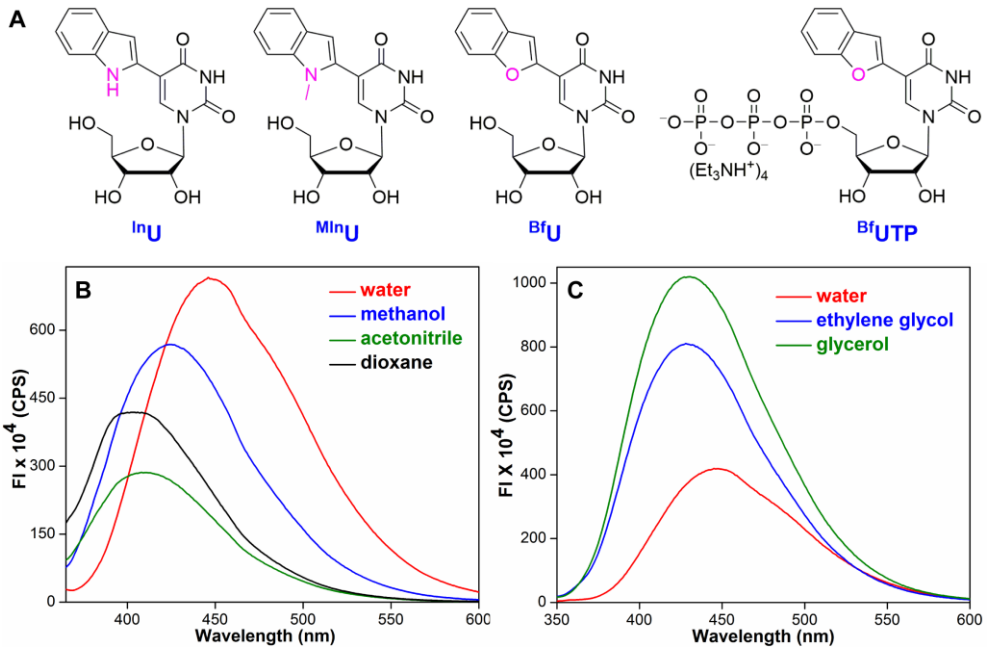


Figure 1. (**A**) Chemical structure of nucleoside analogs. Triphosphate ^{Bf}UTP suitable for enzymatic incorporation of ^{Bf}U into RNA ONs is also shown. (**B** and **C**) Emission studies of ^{Bf}U in solvents of different polarity and viscosity, respectively.

^{Bf}U was incorporated into RNA ONs by T7 RNA polymerase-catalysed *in vitro* transcription reaction using ^{Bf}UTP (Figure 2). Steady-state and time-resolved fluorescence spectroscopy studies of ^{Bf}U containing ONs validated the sensitivity of ^{Bf}U towards changes in its environment such as base-pair substitutions. In particular, ^{Bf}U containing RNA ON was found to report the presence of a DNA abasic site with an appreciable enhancement in its fluorescence intensity (Figure 2).[16]

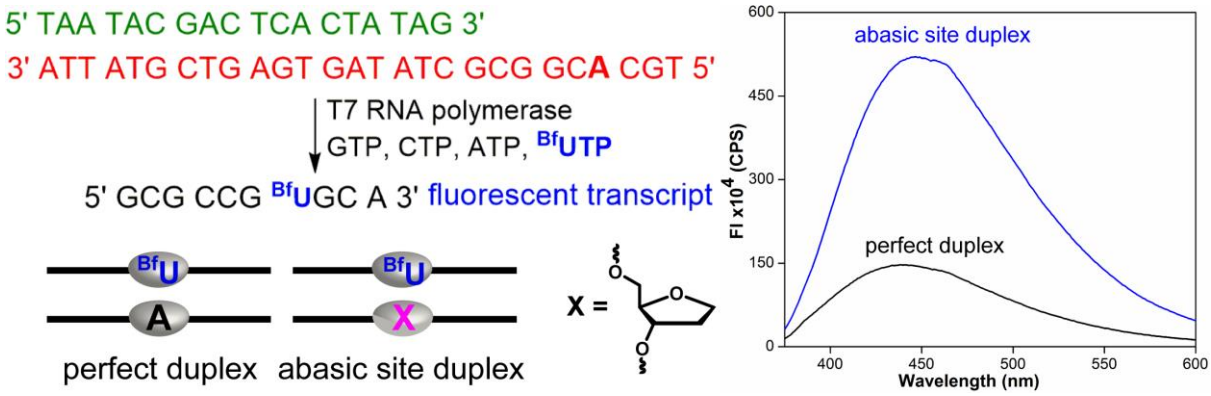


Figure 2. Schematic representation of enzymatic incorporation of ^{Bf}UTP by *in vitro* transcription reaction and detection of DNA abasic site using fluorescently labeled RNA ON.

Chapter 2B: 5-Benzofuran-modified 2'-deoxyuridine as a probe for the fluorescence detection of RNA abasic site

Responsiveness of ^{Bf}U ribonucleoside analog to neighbouring base environment prompted us to investigate the effect of placing benzofuran-modified nucleobase within DNA ONs. 5-Benzofuran-2'-deoxyuridine (^{Bfd}U), which essentially exhibits similar fluorescence properties as that of the corresponding ribonucleoside analog, was site-specifically incorporated into DNA ONs by solid-phase ON synthesis using phosphoramidite substrate. Upon incorporation into single stranded and double stranded ONs, the emissive nucleoside showed significantly enhanced emission intensity compared to the free nucleoside, a property rarely exhibited by the majority of fluorescent nucleoside analogs (Figure 3). The photophysical behaviour of ^{Bfd}U prompted us to assess the ability of the emissive nucleoside to detect the presence of an abasic site in a biologically relevant RNA motif.

The formation of abasic sites in RNA is rare and is almost exclusively associated with the depurination activity of protein toxins called ribosome inactivating proteins (RIPs). RIP toxins such as ricin and saporin catalyze the depurination of a specific adenosine residue, within the highly conserved sarcin–ricin loop region of eukaryotic 28S rRNA. As a result,

the depurinated ribosome has lower affinity for elongation factors essential for protein synthesis, which leads to the cell death. Current methods to monitor the specific depurination activity of RIP toxins are laborious or involve extensive radiolabeling procedures and hence are not amenable to fast-paced studies of RIPs and RNA abasic site. DNA ON complementary to the model RNA ON containing the conserved sarcin/ricin loop region of eukaryotic 28S rRNA, a commonly used substrate to study the depurination activity of ribosome inactivating proteins (RIPs), was appropriately labeled with ^{Bfd}U.[17] Hybridization of fluorescent DNA ON and RIP RNA substrate resulted in a strong emission band, while the duplex of the fluorescent DNA ON and depurinated product mimic exhibited a drastically quenched emission (Figure 3). This observation highlights the potential of ^{Bfd}U as a sensitive fluorescence hybridization probe for RNA abasic site.[18]

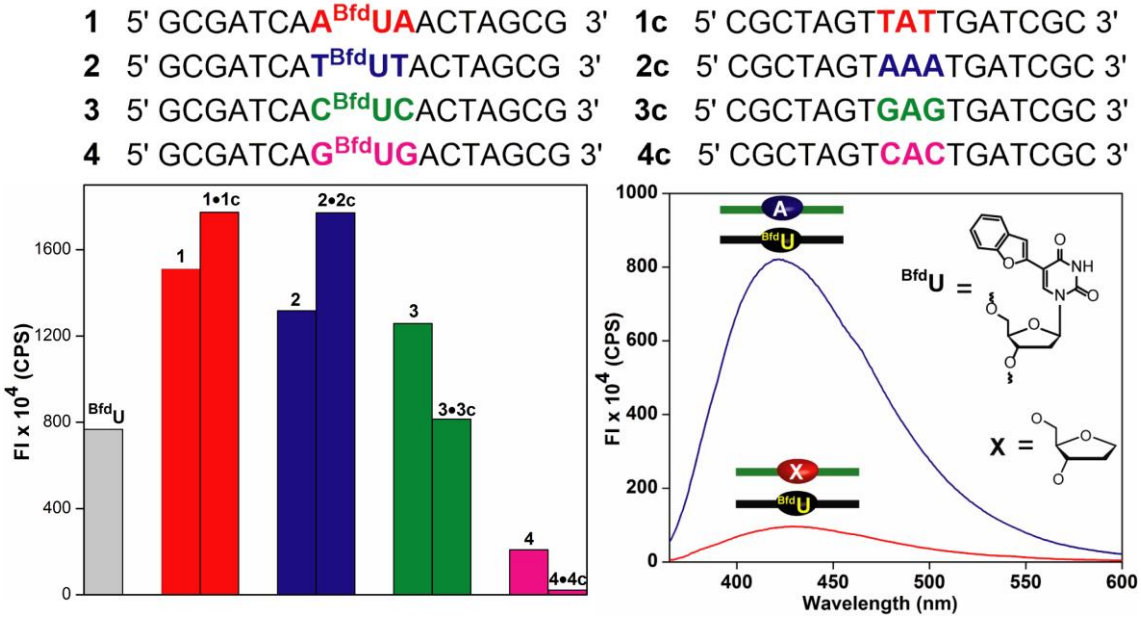


Figure 3. Fluorescence intensity of ^{Bfd}U, modified DNA ONs (**1–4**) and corresponding DNA duplexes, in which ^{Bfd}U is flanked between different nucleobases. Utility of ^{Bfd}U labeled DNA to detect the abasic site in therapeutically relevant RNA is also demonstrated.

In parallel, we have developed a label-free fluorescence method to detect the depurination activity of RIP toxin by using a fluorescent ligand, 2-amino-5,6,7-trimethyl-1,8-naphthyridine (ATMND), which is known to exhibits drastic quenching in fluorescence intensity upon pseudo-base pairing with a cytosine residue placed opposite an abasic site in an ON duplex. The potent depurination activity of saporin on the synthetic RNA substrate A generated the depurinated RNA product B, which upon hybridization to the complementary ON C placed a cytosine residue opposite the abasic site (Figure 4A). Addition of ATMND to

the above product duplex reported the formation of abasic site in RNA with appreciable quenching in fluorescence intensity (Figure 4B). This method offers a straightforward approach to study the kinetics of depurination reactions and could be implemented in high throughput screening (HTS) formats for the discovery of potential RIP inhibitors.[19]

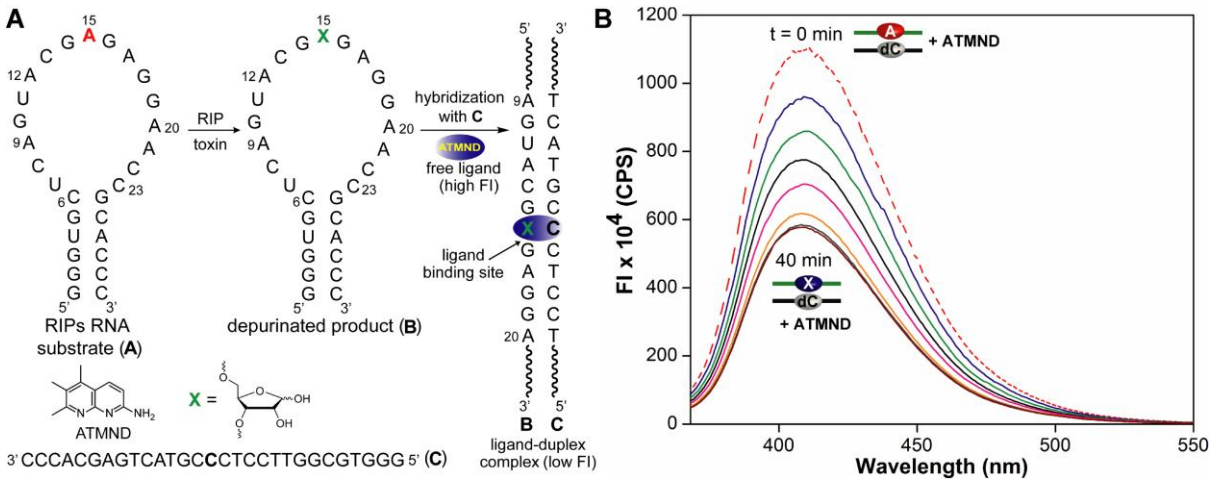


Figure 4. **(A)** Schematic illustration of label-free fluorescence detection of a RIP toxin-mediated depurination of a hairpin ON substrate A, depurinated RNA ON product B and complementary DNA ON C used in this study. While duplex A•C contains an A•dC mismatch at position 15, duplexe B•C contain an abasic site opposite dC **(B)** Saporin mediated depurination of RNA substrate A monitored by fluorescent ligand ATMND.

Chapter 3: Fluorescent nucleoside analogs as probes for study the DNA and RNA non-canonical structures

Chapter 3A: 5-Benzofuran-modified nucleoside analogs as topology-specific probes for human telomeric DNA and RNA G-quadruplex structures

In recent years, guanine-rich sequences with the potential to form non-canonical four-stranded nucleic acid structures, called G-quadruplexes (GQs), have received particular attention due to their interesting structural features and biological functions.[20–21] Owing to the structural diversity of GQs and their role in diseases, G-rich sequences are being rigorously evaluated as novel therapeutic targets for cancer chemotherapies.[21] However, paucity of efficient chemical probes that can detect different GQ topologies and quantitatively report ligand binding has been a major impediment in the advancement of GQ-directed therapeutic strategies. Therefore, therapeutic evaluation of GQs will greatly benefit from the development of fluorescence based biophysical tools that can distinguish different GQs based

topology and nucleic acid type, and are compatible to screening formats for the identification of small molecule ligands that selectively bind to different GQ structures is highly desired.

The microenvironment sensitivity of 5-benzofuran-uridine and the 3D structure of h-Telo DNA repeats in which the loop region (TTA) undergoes substantial conformational changes upon binding to a small molecule ligand inspired us to develop the benzofuran-modified uridine analog as a probe to detect GQ structures and GQ binders. Human telomeric DNA repeat $d[AG_3(T_2AG_3)_3]$ and telomeric repeat-containing RNA (TERRA) $r[U_2AG_3]_4$ were labeled with ^{Bfd}U and ^{Bf}U, respectively. Both steady-state fluorescence and lifetime analysis demonstrated the potential of ^{Bfd}U and ^{Bf}U to selectively detect GQ structures from corresponding duplexes, and distinguish between GQ topologies based on (i) ionic conditions (Na^+ vs. K^+) and (ii) nucleic acid type (DNA vs. RNA). Furthermore, the conformation-sensitivity of these probes was used in developing a fluorescence binding assays to quantitatively monitor the topology-specific binding of model ligands such as pyridostatin and BRACO19 to DNA and RNA GQs. It is expected that this unique class of GQ probes will provide an efficient platform to identify topology- and nucleic acid-specific GQ binders of therapeutic potential.[22]

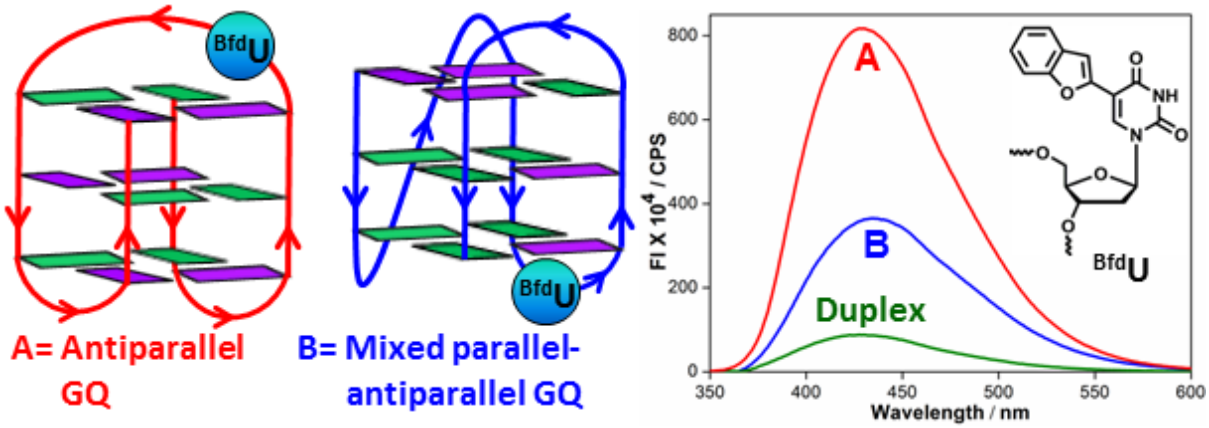


Figure 5. Incorporation of ^{Bfd}U into G-rich human telomeric DNA and detection of topology specific G-quadruplex structures.

Chapter 3B: Benzofuran-conjugated fluorescent uridine analog to probe the conformational switch in i-motif forming human telomeric DNA and RNA repeats

Cytidine-rich (C-rich) repeat oligomers in slightly acidic conditions form a tetraplex of hemiprotonated base pair of $C{\bullet}C^+$ and form stable self-assembled structure known as i-motif (Figure6). C-rich sequences often exist in promoter regions of the oncogene and in human

telomeric DNA, which can fold into an i-motif structure and controlling the oncogene expression at the transcriptional level.[23] Occurrence of i-motif structures in living cell has not been determined yet, and hence, less is known about these structures as potential therapeutic targets. Nevertheless, sequence specific molecular recognition properties of these pH sensitive self-assemblies have been programmed and utilized as building blocks of DNA nano-devices for various application and have received great attention.[24]

Superior probe-like properties of ^{Bf}U and ^{Bfd}U encouraged us to incorporate both the nucleosides into C-rich human telomeric DNA and RNA ONs. Using both steady-state fluorescence and lifetime analysis, we have monitored the pH dependent structural transition of these ONs from single strand to compact i-motif structure (Figure 6). Our results demonstrate the versatility of these probes in studying non-canonical structures in nucleic acids.

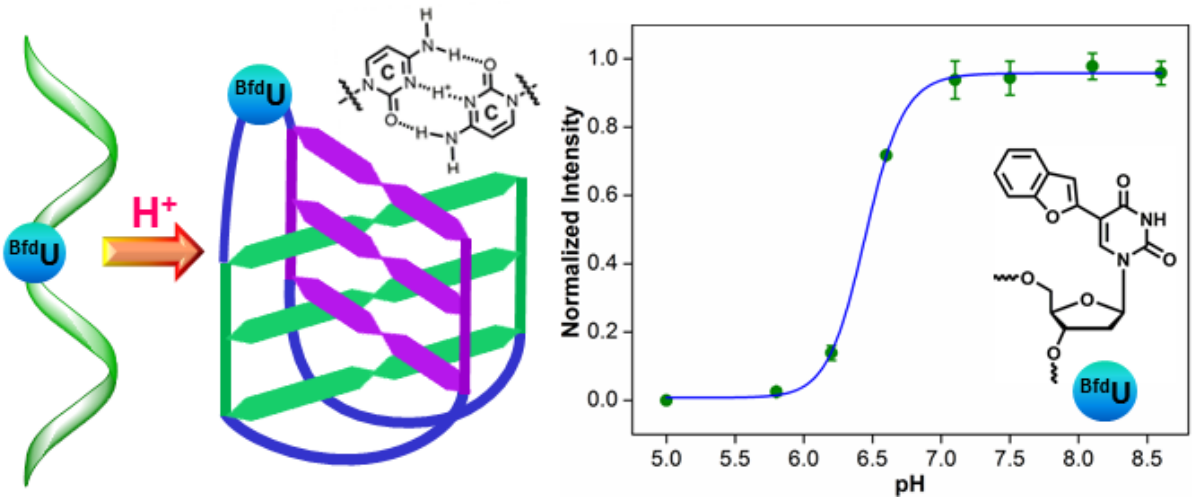


Figure 6. Site-specific labeling of ^{Bfd}U into C-rich human telomeric DNA ON and fluorescence based monitoring the pH dependent structural transition of i-motif structure.

Chapter 4: Synthesis, photophysical properties and incorporation of a highly emissive and environment-sensitive uridine analog based on the Lucifer chromophore

Although ^{Bf}U is highly suitable fluorescent nucleoside probe for *in vitro* applications, its absorption maxima in the UV region limits its application for in cells nucleic acid study. Therefore, we synthesized a bright and environment-sensitive uridine analog based on the Lucifer chromophore 1,8-naphthalimide core (^{Np}U). It was postulated that coupling 4-ethynyl-1,8-naphthalimide at the 5-position of uridine would impart the non-emissive nucleoside with superior probe-like properties akin to Lucifer dyes.[25] Notably, ^{Np}U displays

both excitation and emission maxima in the visible region, exhibits a high quantum yield and excellent solvatochromism (Figure 7).

The corresponding triphosphate and phosphoramidite derivatives of naphthalimide-modified nucleosides act as a good substrate for the synthesis of fluorescent RNA ONs by *in vitro* transcription reactions as well as by solid-phase ON synthesis protocols, respectively. Furthermore, the emissive nucleoside incorporated into ON duplexes exhibits appreciable fluorescence efficiency. Importantly, the fluorescence properties of the emissive nucleoside incorporated into ONs is sensitive to flanking bases and base pair substitutions. In particular, the emissive nucleoside signals the presence of purine repeats (dG and dA) with a significant increase in fluorescence intensity (Figure 7), a property seldom exhibited by most of the fluorescent nucleoside analogs.[26] Currently, the suitability of the fluorescent nucleoside analog in probing the cellular uptake and trafficking of RNA ONs by fluorescence microscopy is being evaluated.

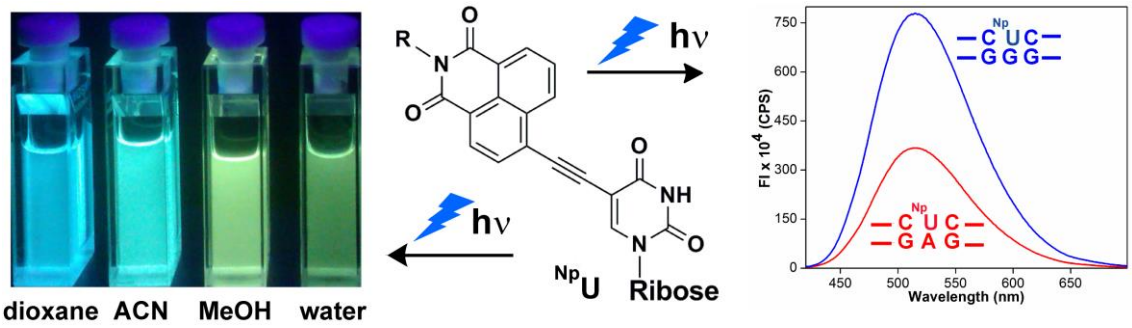
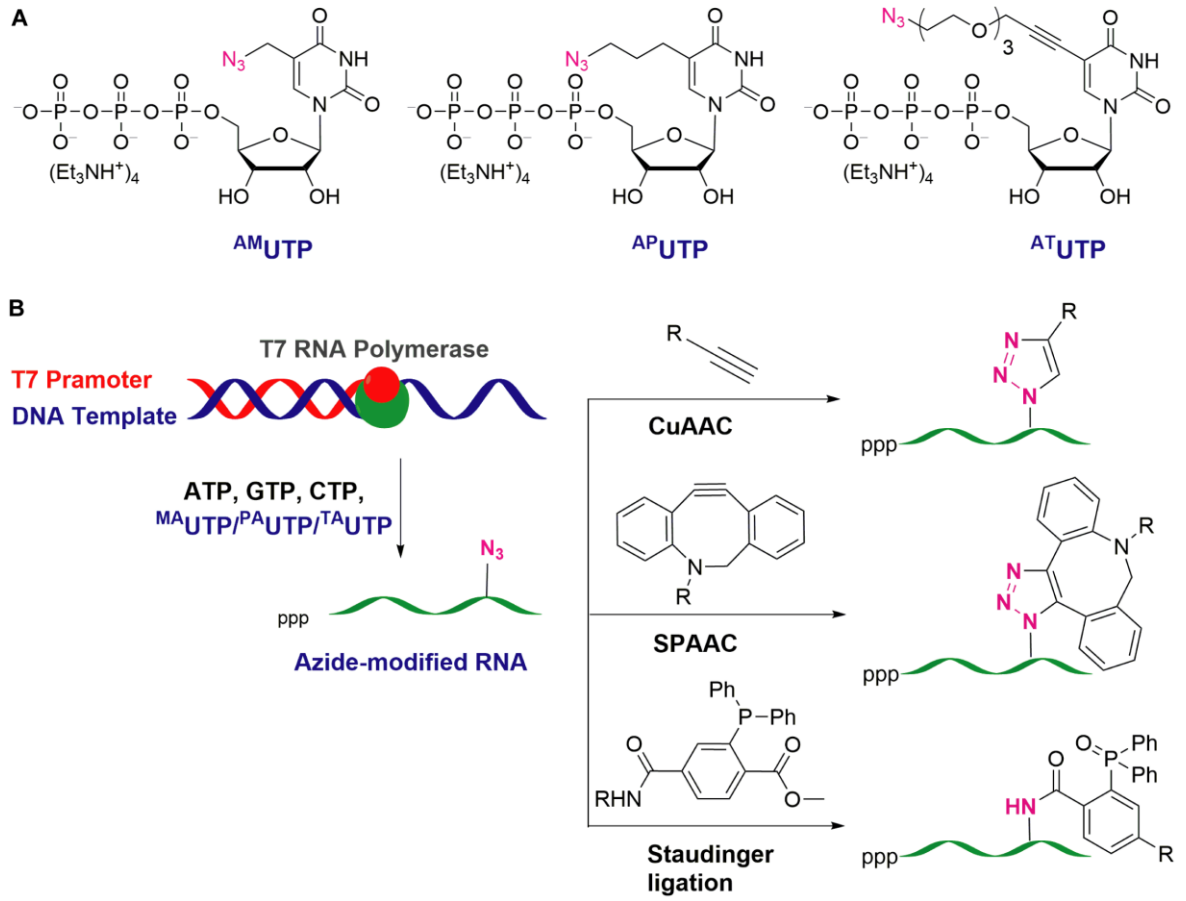


Figure 7. Structure, solvatochromism of ^{Np}U and fluorescence detection of purine repeats.

Chapter 5: Azide-Modified nucleotides a versatile toolbox for posttranscriptional functionalization of RNA by bioorthogonal chemical reactions

The azide group, which is the common reactive functionality in CuAAC, SPAAC and Staudinger ligation reactions, is not easily incorporable by conventional solid-phase ON synthesis protocols as it readily undergoes Staudinger-type reaction with phosphoramidite substrates. In this regard, we have synthesized new azide-modified ribonucleoside triphosphates ^{AM}UTP, ^{AP}UTP and ^{AT}UTP, and have effectively incorporated into oligoribonucleotides by *in vitro* transcription reactions catalyzed by T7 RNA polymerase (Scheme 1).[27–28] Reactions with various oligodeoxynucleotide templates exemplify the utility of transcription reaction in functionalizing oligoribonucleotides with more than one azide group. The azide-modified RNA ONs were functionalized *in vitro* with variety of

biophysical probes in a modular fashion by CuAAC, SPAAC and Staudinger ligation reactions (Scheme 1).[29] Notably, one of the analogs, 5-azidomethyl UTP (^{AM}UTP), was found to be incorporated into cellular RNA transcripts by endogenous RNA polymerases, which uniquely enabled the imaging of newly transcribing RNA by posttranscriptional click reactions (in-cell studies with AMUTP have performed by one of my laboratory colleagues and are described in his thesis). Taken together, our results demonstrate that this practical chemical labeling methodology will provide new opportunities to study RNA *in vitro* and in cells.



Scheme 1. (**A**) Structure of azide-modified uridine triphosphates substrate. (**B**) Enzymatic incorporation of azide-UTP into RNA ON and posttranscriptional chemical functionalization.

References:

[1] M. P. Latham, D. J. Brown, S. A. McCallum, A. Pardi, *ChemBioChem* **2005**, *6*, 1492–1505.

[2] O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007,** *40*, 1–53.

[3] L. Lina, S. Jia, H. Zhen, *Chem. Soc. Rev.* **2011**, *40*, 4591–4602.

[4] F. Wachowius, C. Höbartner, *ChemBioChem* **2010**, *11*, 469–480.

[5] A. Okamoto, Y. Saito, I. Saito, *J. Photochem. Photobiol. C* **2005**, *6*, 108–122.

[6] J. N. Wilson, E. T. Kool, *Org. Biomol. Chem.* **2006**, *4*, 4265–4274.

[7] L. M. Wilhelmsson, *Quart. Rev. Biophy*. **2010**, *43*, 159–183.

[8] A. A. Sawant, S. G. Srivatsan, *Pure Appl. Chem*. **2011**, *83*, 213–232.

[9] K. Phelps, A. Morris, P. A. Beal, *ACS Chem. Biol.* **2012**, *7*, 100–109.

[10] R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev*. **2010**, *110*, 2579–2619.

[11] S. H. Weisbrod, A. Marx, *Chem. Commun*. **2007**, 1828−1830.

[12] P. M. E. Gramlich, C. T. Wirges, A. Manetto, T. Carell, *Angew. Chem. Int. Ed*. **2008**, *47*, 8350−8358.

[13. A. H. El-Sagheer, T. Brown, *Chem. Soc. Rev.*, **2010**, *39*, 1388−1405.

[14] D. Schulz, A. Rentmeister, *ChemBioChem* **2014**, *15*, 2342–2347.

[15] T. Wada, A. Mochizuki, S. Higashiya, H. Tsuruoka, S. Kawahara, M. Ishikawa, M. Sekine, *Tet. Lett*. **2001**, *42*, 9215−9219.

[16] A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J*. **2011**, *17*, 12820–12827.

[17] S. G. Srivatsan, N. J. Greco, Y. Tor, *Angew. Chem. Int. Ed*. **2008**, *47*, 6661–6665.

[18] A. A. Tanpure, S. G. Srivatsan, *ChemBioChem*, **2012**, *13*, 2392–2399.

[19] A. A. Tanpure, P. Patheja, S. G. Srivatsan, *Chem. Commun.* **2012**, *48*, 501–503.

[20] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35*, 7429−7455.

[21] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261−275.

[22] A. A. Tanpure, S. G. Srivatsan, *Nucleic Acids Res.* **2015**, *doi: 10.1093/nar/gkv743.*

[23] M. Guéron, J.-L, Leroy, *Curr. Opin. Struct. Biol.* **2000**, *10*, 326–331.

[24] Y. Krishnan, F. C. Simmel, *Angew. Chem. Int. Ed*. **2011**, *50*, 3124–3156.

[25] R. M. Duke, E. B. Veale, F. M. Pfeffer, P. E. Kruger, T. Gunnlaugsson, *Chem. Soc. Rev.* **2010**, *39*, 3936–3953.

[26] A. A. Tanpure, S. G. Srivatsan, *ChemBioChem* **2014**, *15*, 1309–1316.

[27] H. Rao, A. A. Sawant, A. A. Tanpure, S. G. Srivatsan, *Chem. Commun*. **2012**, *48*, 498–500.

[28] H. Rao, A. A. Tanpure, A. A., Sawant, S. G. Srivatsan, *Nat. Protocols* **2012**, *7*, 1097–1112.

[29] A. A. Sawant, A. A. Tanpure, P. P. Mukherjee, S. Athavale, A. Kelkar, S. Galande, S. G. Srivatsan, *Nucleic Acids Res.* **2015**, (*Just accepted*).

List of Publications

[1] **A. A. Tanpure**, S. G. Srivatsan **(2011)** A microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogue: synthesis, enzymatic incorporation, and fluorescence detection of a DNA abasic site. *Chem. Eur. J. 17*, 12820–12827.

[2] H. Rao, A. A. Sawant, **A. A. Tanpure**, S. G. Srivatsan **(2012)** Posttranscriptional chemical functionalization of azide-modified oligoribonucleotides by bioorthogonal click and Staudinger reactions. *Chem. Commun. 48*, 498–500.

[3] **A. A. Tanpure**, P. Patheja, S. G. Srivatsan **(2012)** Label-free fluorescence detection of the depurination activity of ribosome inactivating protein toxins. *Chem. Commun. 48*, 501–503.

[4] H. Rao, **A. A. Tanpure**, A. A. Sawant, S. G. Srivatsan **(2012)** Enzymatic incorporation of an azide-modified UTP analog into oligoribonucleotides for post-transcriptional chemical functionalization. *Nature Protocols 7*, 1097–1112.

[5] **A. A. Tanpure**, S. G. Srivatsan **(2012)** Synthesis and photophysical characterisation of a fluorescent nucleoside analogue that signals the presence of an abasic site in RNA. *ChemBioChem 13*, 2392–2399.

[6] **A. A. Tanpure**, M. G. Pawar, S. G. Srivatsan **(2013)** Fluorescent nucleoside analogs: probes for investigating nucleic acid structure and function. *Isr. J. Chem. 53*, 366–378.

[7] **A. A. Tanpure**, S. G. Srivatsan **(2014)** Synthesis, photophysical properties and incorporation of a highly emissive and environment-sensitive uridine analogue based on the lucifer chromophore. *ChemBioChem 15*, 1309–1316.

[8] **A. A. Tanpure**, S. G. Srivatsan, (**2015**) Conformation-sensitive nucleoside analogs as topology-specific fluorescence turn-on probes for DNA and RNA G-quadruplexes. *Nucleic Acids Res. DOI: 10.1093/nar/gkv743.*

[9] A. A. Sawant, **A. A. Tanpure**, P. P. Mukherjee, S. Athavale, A. Kelkar, S. Galande, S. G. Srivatsan, **(2015)** A versatile toolbox for posttranscriptional chemical labeling and Imaging of RNA. *Nucleic Acid Res.* (*Just accepted*)

[10] **A. A. Tanpure**, S. G. Srivatsan, Benzofuran-conjugated fluorescent uridine analog to probe the conformational switch in i-motif forming human telomeric DNA and RNA repeats. (Manuscript under preparation)

Patent

[1] **A. A. Tanpure**, A. A. Sawant, S. Galande, S. G. Srivatsan **(2015)** Novel azide-modified UTP analogs for posttranscriptional chemical functionalization and imaging of RNA. *Indian Patent Application No.1555/MUM/2015*.

Chapter 1

Base-modified nucleoside analogs in nucleic acid study

1.1 Introduction

Nucleic acids perform a plethora of vital cellular functions such as storage and transfer of genetic information, catalysis and regulation of gene expression by interacting with nucleic acids, proteins and small molecule metabolites. Unlike proteins, nucleic acids are chemically less diverse. But amazingly they (especially RNA) expand their functional repertoire by using their inherent conformational dynamics and by adopting diverse 3-dimensional structures that rapidly interconvert between different functional states (Figure 1).[1,2] Several biophysical and theoretical tools have been developed to uncover the fundamentals of nucleic acid folding and recognition processes. The majority of biophysical investigations greatly rely on techniques, namely, fluorescence, electrophoresis, circular dichroism (CD), calorimetry, nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR), X-ray crystallography and microscopy.[3-5] Among these, fluorescence spectroscopy is by far the most attractive technique as it is easily accessible, versatile and provides information in real time with great sensitivity. Consequently, fluorescence spectroscopy has been extensively used to investigate pathways and kinetics of conformational transformations of nucleic acids, nucleic acid-protein and nucleic acid-small molecule complexes.[3] Importantly, advances in ultrafast and single-molecule fluorescence spectroscopy techniques have allowed the investigation of conformational dynamics and processing of nucleic acids *in vitro* as well as in cells in a wide range of time scales.[6,7]

Needless to say, fluorescent reporters have to be introduced into nucleic acids since natural nucleobases are practically nonemissive.[8,9] Common fluorophores such as ethidium bromide, DAPI, SYBR Green as well as Hoechst dyes, which bind to nucleic acids noncovalently by intercalation or along the grooves, are frequently used to visualize nucleic acids in gel electrophoresis and cell microscopy. Covalent attachment of commercially available fluorescent derivatives (fluorescein, rhodamine, Alexa and cyanine dyes, *etc*.) to the phosphate backbone, sugar or base is one of the most common methods to fluorescently modify nucleic acids.[10] Obvious advantageous of these probes are that they possess very high molar absorptivities and fluorescence quantum yields, and are significantly photostable. Consequently, custom synthesized oligonucleotides (ONs) labeled with these probes are easily accessible and also fluorescent precursors (e.g., dye-labeled phosphoramidites and triphosphates) that can be readily incorporated into ONs by either chemical or enzymatic methods are also commercially available.[10] Such probes incorporated into ONs have been widely used in FRET, fluorescence anisotropy and single-molecule fluorescence

experiments.[7,11] However, a main shortcoming of the majority of these probes is that upon binding or incorporation they significantly perturb the native structure of nucleic acids. Hence, in several instances the probe molecule is attached to the ON via a long linker to minimize structural perturbations.

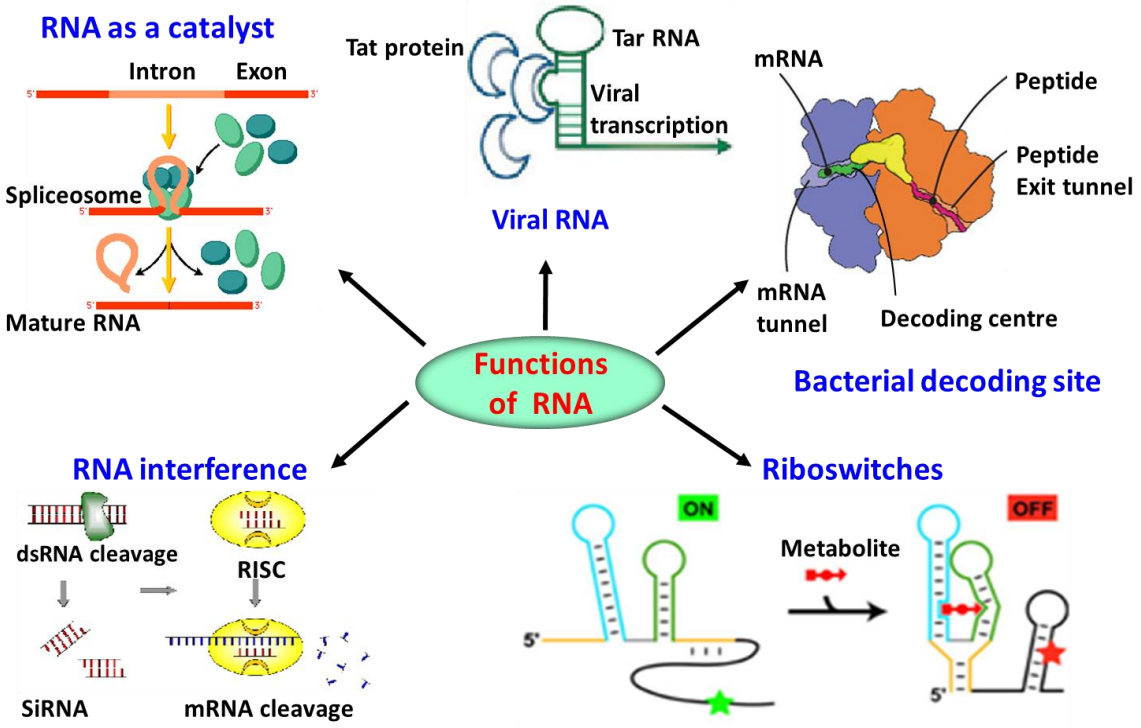


Figure 1. Pictorial representation of various functions of RNA in contemporary biology.

The conformational changes, which take place during nucleic acid folding or recognition event, are both at the global as well as nucleotide level.[1,2] Changes in the conformation of nucleotides (e.g., near the binding site) most often alter their surrounding physical properties and their interactions with neighbouring bases. Hence, introducing an environment-sensitive reporter molecule near the site of interest and surveying the local conformational changes by fluorescence spectroscopy has been shown to provide a better picture of the interaction under investigation.[12,13] An important consideration that has to be taken into account when incorporating such fluorescent probes is that they should minimally perturb the native structure and function of the target nucleic acid. In this regard, base-modified fluorescent nucleoside analogs that are capable of reporting changes in their conformation and surrounding environment in the form of changes in their fluorescence properties such as quantum yield, emission maximum, lifetime and anisotropy have found

wide applications in developing tools to investigate the dynamics, structure and function of nucleic acids.[12-16]

1.2 Biophysical techniques and chemical probes for studying structure and function of nucleic acids

1.2.1 X-ray crystallography

X-ray crystallography provides 3D structural insight into nucleic acid, which allows us better understanding of biological processes at the atomic level. X-ray crystallography has been very helpful in studying the mechanisms of replication initiation, transcription and translation regulation, catalytic RNAs, riboswitch function, and nucleic acid-protein interactions.[17] Heavy atom derivatization of nucleic acids for phase determination is essential in X-ray crystallographic analysis. Heavy atom labels are introduced into nucleic acid by soaking or co-crystallizing in salt solutions of heavy atoms or incorporation of bromine or iodine containing nucleoside into nucleic acid.[18] However, halogenated nucleosides are prone to undergo dehalogenation during X-ray crystallography analysis.[19] To overcome this limitation, anomalous scattering property of selenium atom has been extensively used in the multi-wavelength anomalous dispersion (MAD) phasing technique for X-ray crystallography study of nucleic acids.[20]

Egli, Huang and co-workers have successfully implemented 2'-methylselenouridine (**1**) for the derivatization of RNA and DNA for MAD phasing.[21] Thereafter, Micura and co-workers reported an advanced procedure for the site-specific labeling of standard nucleosides containing 2'-methylseleno groups in RNA.[22] Recently, Huang and co-workers have investigated the atom-specific replacement of the nucleobase oxygen with selenium. This research provides novel insights into the base pairing and stacking interactions of nucleic acids. The phosphoramidites building block of the 4-Se-T (**2**), 2-Se-T (**3**), 5-SeMe-T (**4**) and 6-Se-dG (**5**) for solid-phase synthesis and the 4-Se-T triphosphate for enzymatic synthesis have been developed (Figure 3).[20] Together, these efforts have proven the applicability of selenium atom as a MAD phasing label and paved the way for use of selenium-modified nucleic acids in X-ray crystallography-based structural analysis of DNA and RNA.

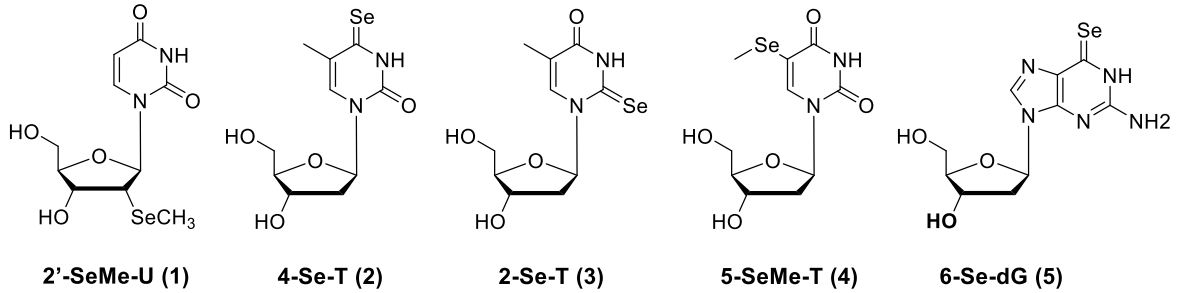


Figure 3. Representative examples of selenium containing nucleoside probe for X-ray diffraction analysis of nucleic acids.

1.2.2 Nuclear Magnetic Resonance (NMR) spectroscopy

NMR is a very useful technique to investigate nucleic acids in solution and to characterize their dynamic properties at atomic level. With isotope-labeled (^{13}C, ^{15}N isotopes) oligonucleotides, it is possible to explore base-pairing modes, local and global dynamics and interaction surfaces with proteins, small ligands and metal ions.[23] Currently, both biochemical and chemical methods for the synthesis of isotope labeled nucleic acids are available. Wenter and Pitsch reported the preparation of specifically labeled $^{15}N(3)$-pyrimidine and $^{15}N(1)$-purine for site-specific chemical labeling of RNA. Subsequently, they investigated folding kinetics of topologically favoured conformational exchange between different hairpin folds.[24] Williamson and co-workers developed a method for isotopic labeling of the purine (**6-9**) through biochemical pathway (Figure 4). Furthermore, they have illustrated the utility of this method to obtain NMR spectra of HIV-2 transactivation region of RNA.[25] Very recently, Mergny and co-workers demonstrated G-quadruplex DNA and ligand interaction in living cells using isotopically labeled DNA ON.[26]

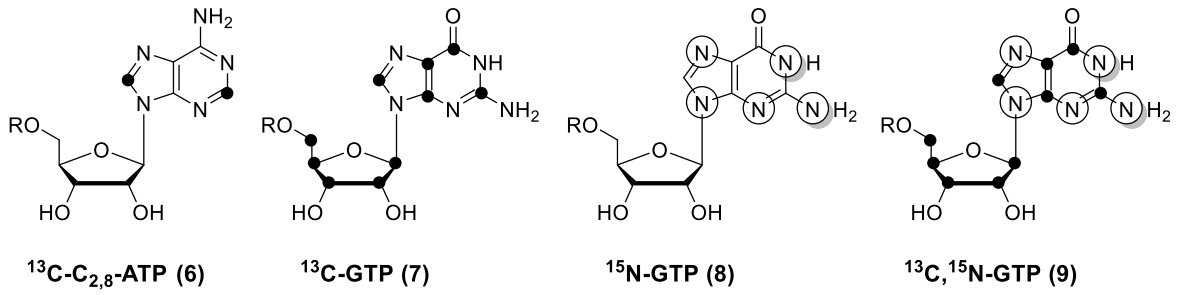


Figure 4. ^{13}C (•) and ^{15}N (○) labeled triphosphate substrates used in the synthesis of isotopically enriched RNA for NMR spectroscopy (R = $(PO_3)_3^{-4}$)

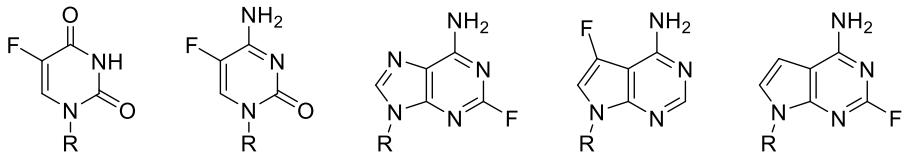


Figure 5. Fluorine-modified nucleoside probes to study various secondary structures of RNA. R = ribose

On the other hand, use of fluorine labels offers high intrinsic sensitivity as natural abundance of the ^{19}F nucleus is 100% which practically solves the problem of resonance degeneracy to study biomolecules.[27] Site-specific incorporation of fluorine atoms into nucleic acids was exploited to monitor conformational changes and folding of nucleic acids or to investigate binding processes in small molecule/nucleic acid or enzyme/nucleic acid complexes.[28] Micura group implemented 5-F pyrimidine labeling (Figure 5) and implemented to probe DNA and RNA secondary structures by 1D ^{19}F NMR.[29] Strobel and co-workers have developed a series of fluorine substituted adenosine analogs **10**-**14** (Figure 5), triphosphate substrate of each analogs is readily incorporated into RNA by T7 RNA polymerase. Further, they have conducted a nucleotide analog interference mapping (NAIM) study on a self-ligating construct of the Varkud Satellite (VS) ribozyme.[30]

1.2.3 Electron Paramagnetic Resonance (EPR) Spectroscopy

In recent years site-directed spin labeling (SDSL) of nucleic acids have received great attention in investigating nucleic acid–ligand interactions and identifying different metals binding sites in nucleic acids by EPR spectroscopy.[31] The rapid developments in EPR spectroscopy triggered the advancement of spin-labeling techniques for nucleic acids and hence, variety of synthetic methods have been developed to label nucleic acid with stable nitroxide as EPR probe. One of the first methods for spin labeling of nucleobase is the introduction of 4-thiouridine at specific sites of RNA by solid-phase ON synthesis and subsequent derivatization by disulfide (**15**) using appropriate reagent.[32] The limitation of this method is its restriction to uridine site only and further possible effects on base pairing and the overall stability of the RNA under investigation. In a similar approach Engels group introduced a new method by performing Pd-catalyzed Sonogashira cross-coupling reaction with TPA (2,2,5,5-tetramethyl-pyrrolin-1-oxyl-3-acetylene) and halogenated nucleosides. This strategy was applied in RNA for the derivatization of 5-iodouridine (**16**) and 5-

iodocytidine (**17**) with spin label, which facilitated the measurement of intramolecular distances in solvated RNA systems.[33] 5-TEMPA-modified deoxyuridine was introduced into the loop regions of G-quadruplex forming human telomeric DNA oligonucleotide and this EPR probe has used to investigate the polymorphic nature of DNA G-quadruplex *in vitro*.[34] Improved stability of the nitroxide radical probes and the sensitivity of the EPR have mutually utilised to determine the *in-cell* intramolecular diatance in nucleic acid molecules with high precision.[35] The accuracy of distance measurements and reliable determination of molecular motion is difficult with majority of spin-labeling probe within ONs as they contain flexible linkers. In this context, Sigurdsson and co-workers reported a selective reaction between aromatic isothiocyanates and 2'-amino groups in RNA. The resulting spin label (**18**) displayed limited mobility in RNA, making them promising candidates for precise distance measurements by pulsed EPR. Upon conjugation to RNA, a tetraethyl isoindoline derivative showed significant stability under reduction conditions.[36]

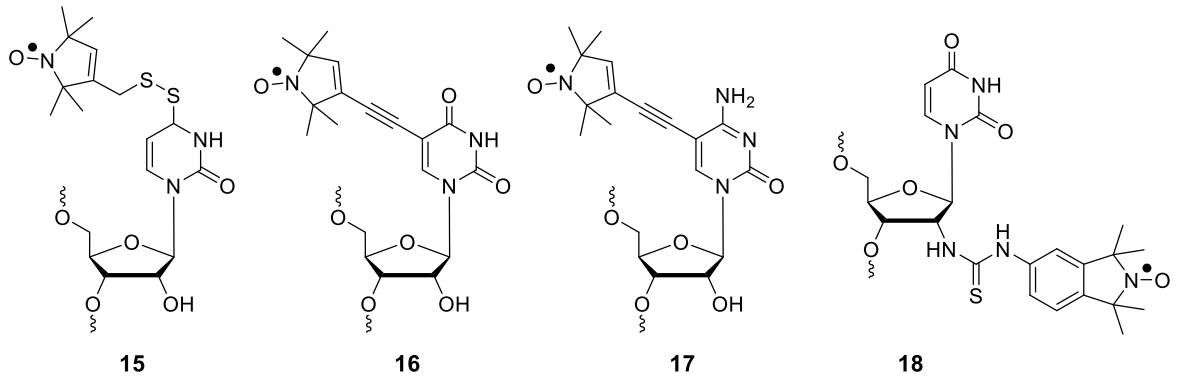


Figure 6. Selected examples of spin labeled nucleic acid probes for application in EPR spectroscopy.

1.2.4. Fluorescence spectroscopy

Fluorescence spectroscopy is one of the most informative and sensitive analytical techniques. Its compatibility in high throughput screening format, usefulness in studying the local and global conformational changes in nucleic acids during various recognition processes make it mostly used analytical tool in nucleic acid research. However, naturally occurring nucleosides are practically non-emissive hence chemist have synthesized a variety of base-modified fluorescent nucleoside analog probes with wide range of fluorescence properties.[12-15]

1.3 Base-Modified fluorescent nucleoside analogs: design and classification

Nucleoside analogs with useful photophysical properties have been developed by extending the π conjugation of nucleobases or by using known fluorophores as base surrogates. The various families of base-modified nucleoside analogs can be generally classified as size-expanded, extended, pteridine, polycyclic aromatic hydrocarbon (PAH) and minimally perturbing fluorescent nucleoside analogs.

1.3.1 Size-expanded fluorescent nucleobase analogs

Size-expanded nucleobase analogs are synthesized by either fusing or inserting aromatic or heterocylic rings onto pyrimidine and purine bases (Figure 7).[37] Many such modifications have afforded highly emissive nucleobases (e.g., **19**–**21**), which often retain the H-bonding complementarity as that of the native bases. Saito and co-workers have developed several such ring-expanded fluorescent nucleobase analogs (e.g., **20** and **21**), which photophysically distinguish the type of base opposite to the emissive base.[38] Saito named these analogs as base-discriminating fluorescent (BDF) nucleobases and effectively implemented them in the homogeneous detection of single nucleotide polymorphisms (SNPs).[39] Matteucci and co-workers synthesized ONs containing tricyclic cytidine nucleoside analogs, phenothiazine (tC, **22**) and phenoxazine (tC^o, **23**), with the intention to maximize the stacking interaction and hence, the stability of the duplex for antisense applications.[40] Later, Wilhelmsson and co-workers thoroughly characterized the photophysical properties of tC and tC^o in the free nucleoside form and also within ONs.[41] Unlike the majority of nucleoside probes, the quantum efficiencies of these fluorophores upon incorporation into ONs are not dramatically reduced. Utilizing this property they could assemble a useful FRET pair using (tC^o) as the donor and its nitro analog, 7-nitro-1,3-diaza-2-oxophenothiazine (tC_{nitro}), as the acceptor.[42] Subsequently, Sigurdsson's group performed a detailed study on the effects of flanking residues on the mismatch-detection abilities of a series of expanded cytidine analogs, tC^o, tC and $Ç^f$ (**24**).[43] They found $Ç^f$ to be superior in identifying all mismatches uniquely, suggesting that it could be utilized in the detection of SNPs. Recently, Wilhelmsson and co-workers have introduced a quadracyclic adenine **qA** (**25**) as an alternative to existing fluorescent adenine analogs. This adenine analog, when incorporated into ONs, maintains the duplex stability and also has a reasonable fluorescence quantum yield as compared to the most of the base-modified fluorescent analogs.[44]

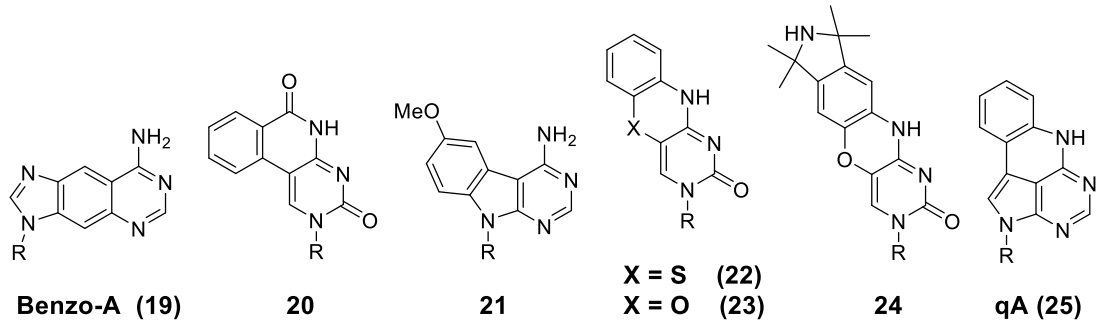


Figure 7. Selected examples of size-expanded fluorescent nucleoside analogs. R = ribose or 2'-deoxyribose.

1.3.2 Size extended fluorescent nucleobase analogs

Family of extended nucleobase analogs are typically synthesized by tethering known fluorophores via rigid or flexible linker to the nucleobases.[13,37] Several examples of extended analogs have been developed by attaching fluorescent aromatic systems such as pyrene (**26**), perylene (**27**), BODIPY (**28**), prodan (**29**), phenanthroline, anthracene, bipyridine, terpyridine and thiazole orange (Figure 8). ONs containing such fluorescent tags have been used by numerous groups for a variety applications including detection of SNPs, nucleic acid lesions, electron transfer process in nucleic acids, *etc*.[45] In an alternate strategy, extended analogs have also been prepared by tagging heterocyles to the nucleobases, which essentially extend the π conjugation thereby imparting favourable photophysical properties (Figure 8).[12-15,46] Many of these analogs exhibit sensitivity to changes in their environment and also form stable base pairs like natural bases when incorporated into ONs.

Engels's and Zhou's group have synthesized several 5-benzimidazolyl-2'-deoxyuridine derivatives, which show emission in the visible region and moderate antibacterial activity.[46g,46h] Recently, Hocek and co-workers constructed solvatochromic nucleoside analogs by conjugating aminophthalimide (**30**) and GFP-like (**31**) chromophores via an alkyne linker.[47] They enzymatically incorporated the corresponding triphosphates into DNA ON reporters and established fluorescence assays to detect the binding of DNA to p53, an important tumour suppressor protein. Kim and co-workers have incorporated piperazinephenyl- and pyrene-modified pyrimidine and purine nucleosides (**32**and **33**, respectively) into DNA ONs, and have studied their cellular uptake and i-motif formation by fluorescence spectroscopy.[48]

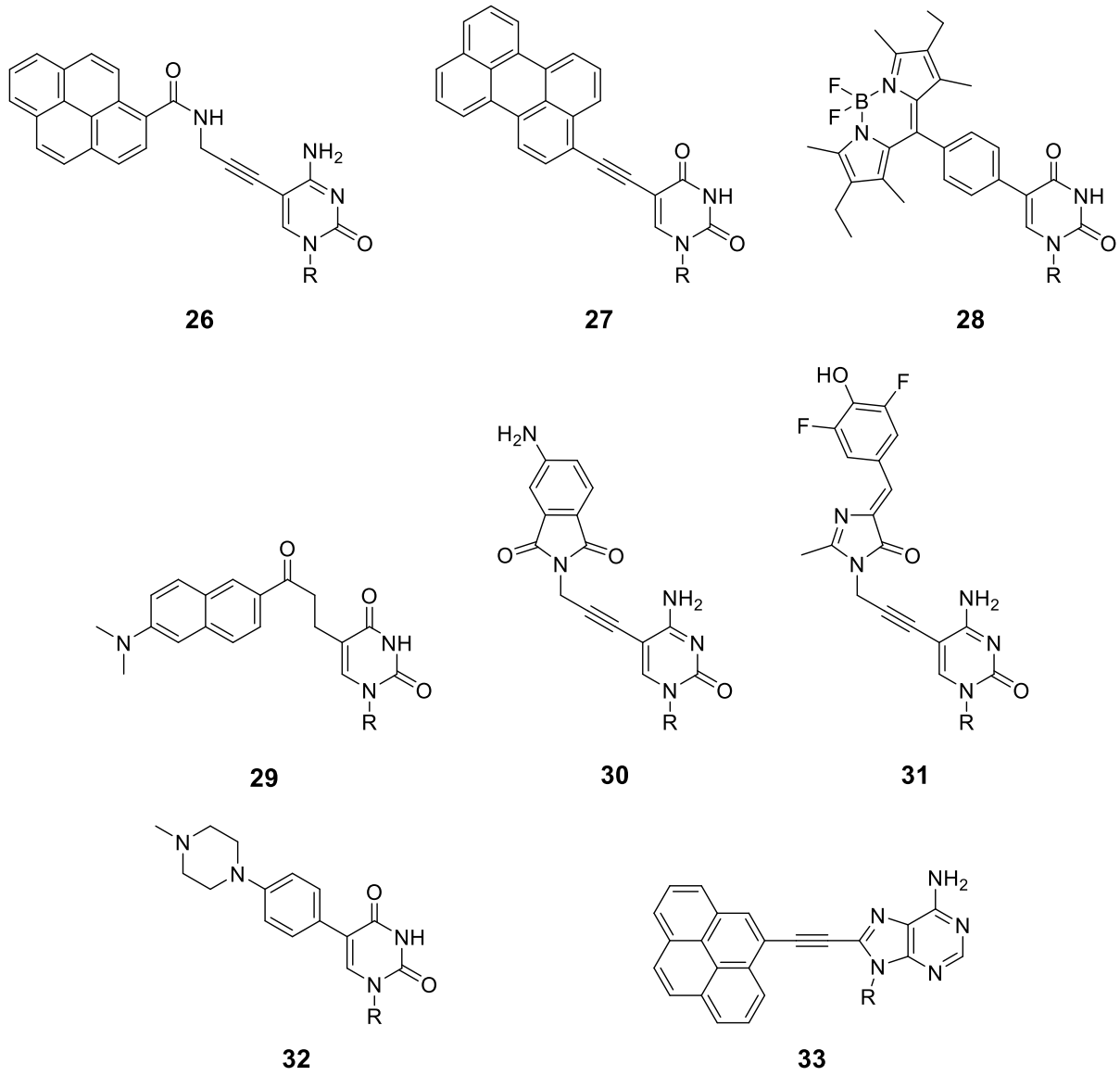


Figure 8. Selected examples of extended fluorescent nucleoside analogs. R = 2'-deoxyribose.

1.3.3 Fluorescent pteridine and polyaromatic hydrocarbon analogs

Pteridines are naturally occurring planar heterobicyclic aromatic compounds bearing structural resemblance to the natural purines.[49] Some of these analogs (e.g., 3MI (**34**) and 6MI (**35**)) are highly emissive and have been used as base surrogates to develop fluorescent nucleosides (Figure 9).[50,51] In a similar approach, natural bases have been replaced with fluorescent PAHs to generate PAH base analogs (**36, 37**) (Figure 9).[52,53] Recently, Kool and co-workers have used combinatorial approach to assemble a library of "oligodeoxyfluorosides" containing different PAH fluorophores in a DNA-like chain. These polychromophoric oligomers display large Stokes shifts and a broad array of quantum yields and emission wavelengths encompassing the complete visible region. These DNA

polyfluorophores have been elegantly implemented in diagnostic assays to detect toxic gases and metal ions, and in multiplexed imaging of live cells.[54]

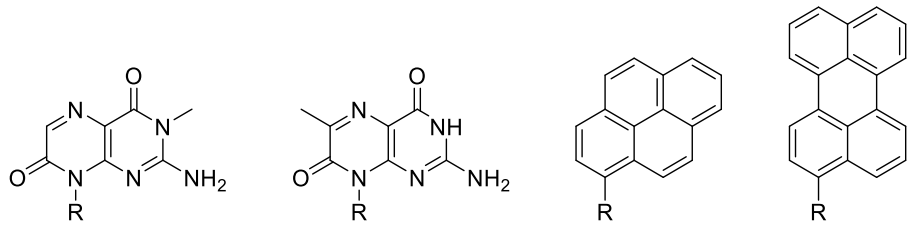


Figure 9. Representative examples of pteridine (**34** and **35**) and PAH (**36** and **37**) base analogs. R = ribose or 2'-deoxyribose.

While the majority of nucleobase analogs discussed above structurally deviate from their parent bases significantly, isomorphic base analogs closely resemble the structure and WC base pairing property of native bases.[12-15] A practical advantage of isomorphic nucleoside analog probes is that they can be specifically placed at a desired position in a target ON with minimum structural perturbations. Such probes are ideally suited for studying subtle alterations in conformation and dynamics that happen around the site of interest during a folding or recognition process. 2-Aminopurine (2-AP, **38**, Figure 10), a highly emissive and environment-sensitive isoster of adenine, was introduced by Stryer and co-workers more than four decades ago.[55] It base pairs strongly with thymine and uracil and serves as an adenine analog in several enzymatic reactions. Because of its nonperturbing nature and exquisite sensitivity to conformational changes, 2-AP by far has been the most widely utilized nucleoside analog in the study of structure, dynamics and binding properties of nucleic acids.[12-15]

1.3.4 Minimally perturbing fluorescent nucleoside analogs

Despite 2-AP's wide utility, excitation and emission maxima in the UV region and drastically reduced fluorescence efficiency within ONs significantly limits its applications, particularly in cell microscopy.[55,56] Therefore, much of the recent advances in the development of new generation fluorescent nucleoside analogs have been generally based on the following criteria. The nucleoside probe should (i) be structurally minimally invasive, (ii) have excitation and emission maxima in the visible region, (iii) have an appreciable quantum yield when incorporated into ONs and (iv) importantly, report changes in microenvironment via changes in the photophysical properties.

Tor's group initiated a program in this direction and synthesized a series of isomorphic nucleoside analogs by either conjugating or fusing aromatic five-membered heterocycles to pyrimidine and purine cores (Figure 10).[13] Preliminary photophysical characterization of furan-, thiophene-, oxazole- and thiazole-conjugated pyrimidines revealed that the furan-conjugated pyrimidines possessed probe-like properties in terms of emission maximum in the visible region with a reasonable quantum yield and sensitivity to polarity changes.[57] Furan-modified uridine (**39**) and cytidine (**40**) analogs were site-specifically incorporated into model DNA and RNA ON reporters without affecting the native structure. These fluorescent reporters were effectively employed in the detection of abasic sites and mutagenic oxidation product of guanosine (8-oxoG), and in the monitoring of RNA-aminoglycoside antibiotics and RNA-protein interactions.[57] Despite possessing high quantum yields, furan-conjugated purines were not evaluated further as they displayed emission in the UV region, and were found to be less sensitive to microenvironment changes. Tor and co-workers synthesized a fused thiophene derivative of uridine (**41**), which displayed high quantum yield and excellent fluorescence solvatochromism. It was incorporated into RNA ONs by both solid-phase chemical and enzymatic methods, and was utilized in devising fluorescence hybridization assays to detect mismatches in duplexes and monitor the formation of RNA abasic sites by the depurination activity of ribosome inactivating protein (RIP) toxins.[58]

Recently, Tor and co-workers assembled new FRET pairs using fluorescent uridine analogs based on 5-methoxyquinazoline-2,4(1*H*,3*H*)-dione (**42**) and 5-aminoquinazoline-2,4(1*H*,3*H*)-dione (**43**), which served as an effective donor for coumarin and acceptor for tryptophan, respectively.[59] Utilizing these FRET systems, fluorescence assays were established to investigate the binding of aminoglycoside antibiotics to a model bacterial ribosomal decoding site (A-site) RNA motif and HIV-1 Rev peptide to its cognate RNA target, Rev Responsive Element (RRE). Other recent examples of isomorphic nucleoside analogs from this group include a complete set of highly emissive RNA alphabet derived from thieno[3,4-*d*]-pyrimidine as the heterocyclic core, and pH and polarity-sensitive 6-aza-uridine derivatives.[60,61] The suitability of these analogs in biophysical probing of nucleic acid-related processes is yet to be explored.

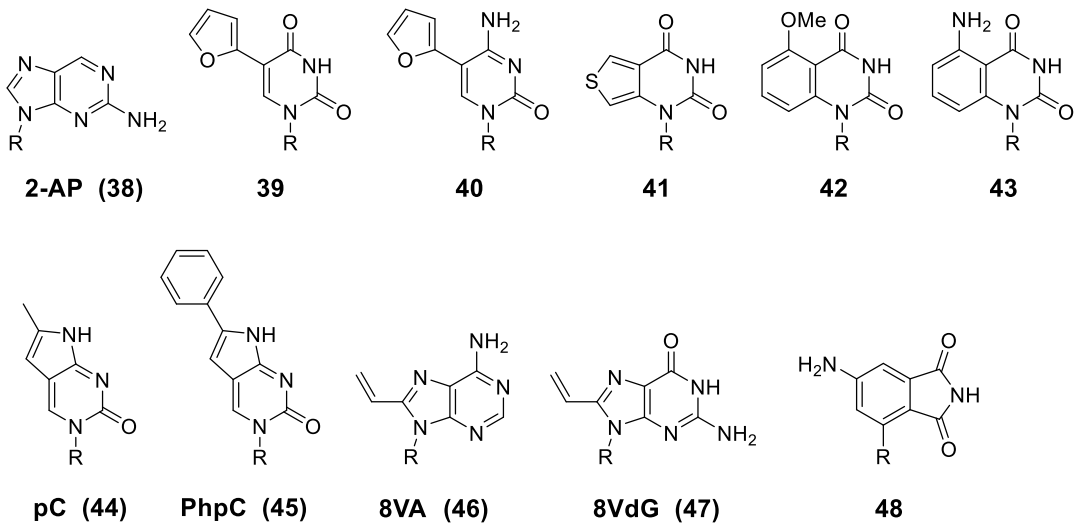


Figure 10. Examples of minimally perturbing fluorescent nucleoside analogs. R = ribose or 2'-deoxyribose.

Pyrrolocytidine (pC, **44**), a moderately emissive isoster of cytidine, displays emission maximum in the visible region.[63,63] Like many other fluorescent nucleoside analogs the quantum efficiency of pC drops drastically when incorporated into ON duplexes. Nevertheless, its isomorphicity and sensitivity to changes in its surrounding environment have been aptly utilized in studying the RNA secondary structure, conformational dynamics of tRNA 3'-end, thermodynamics of drug-DNA complexes and in the detection of 8-oxoG.[63-65] Recently, a few pC derivatives with improved fluorescence properties have been developed (Figure 10).[66-68] In particular, phenylpyrrolocytidine (PhpC, **45**) displays high quantum yield (0.31) and a red-shifted excitation and emission maximum (370 nm and 465 nm, respectively) compared to most other nucleoside analogs.[69] Hudson and co-workers incorporated PhpC into RNA ONs for studying the activity of HIV-1 RT ribonuclease H enzyme. A RNA-DNA hybrid containing a single PhpC served as an excellent substrate and reported the cleavage activity of the RNase H enzyme with a nearly 14-fold enhancement in fluorescence intensity.[69] In a similar approach, a siRNA labeled with multiple PhpC residues was used to monitor the trafficking and silencing activity of siRNA inside living cells using fluorescence microscopy.[70] Since PhpC also showed reduction in fluorescence upon incorporation into ONs, the authors had to use siRNA labeled with at least three PhpC residues for effective fluorescence monitoring inside the cells. However, this led to reduced silencing activity by the siRNA probably due to structural perturbation caused by multiple modifications.

Drawing inspiration from 8-vinyl-deoxyadenosine (**46**), which was introduced as an alternative fluorescent probe to 2-AP, Diederichsen and co-workers developed an isomorphic guanosine analog **47** by tethering a vinyl group at the 8 position of guanine.[71,72] The ability of this emissive and environment-sensitive nucleoside analog to adopt both *syn* and *anti* conformation was appropriately utilized for studying the different topologies of G-quadruplexes.[72] Recently, Wagenknecht, Ernsting and co-workers, synthesized a solvatochromic fluorescent C-nucleoside analog using 4-aminophthalimide (**48**), which when incorporated into a model double stranded DNA exhibited large Stokes shift and a reasonable quantum yield.[73]

1.4 Methods to incorporate modified-nucleoside analogs into nucleic acid

1.4.1 Solid-phase ON synthesis

Solid-phase ON synthesis is the method of choice for the incorporation of most of the modified nucleoside probes into DNA and RNA ONs. In this fully automated method the building block phosphoramidites are sequentially coupled to the growing chain of ON in the order required for the desired sequence (Figure 11).[74] The key step in the ONs synthesis is specific and sequential formation of 3'-5' internuceloside phosphodiester linkage. For the synthesis of nucleoside phosphoramidite substrate reactive hydroxy and exocyclic amino groups present in natural or synthetic nucleosides are needed to be protected appropriately. Generally, the 5'-hydroxyl group is protected by using an acid-labile DMT (4,4'-dimethoxytrityl) group, while 2'-hydroxy group present in ribonucleoside is protected with TBDMS (t-butyldimethylsilyl) group or with TOM (tri-iso-propylsilyloxymethyl) group, both being removable by treatment with fluoride ion.[75] The protection of the exocyclic amino groups has to be orthogonal to that of the 5'-hydroxy group because the latter is removed at the end of each synthetic cycle. In the standard scheme Bz (benzoyl) protection is used for A, dA, C, and dC, while G and dG are protected with isobutyryl group. More recently, Ac (acetyl) group is often used to protect C and dC (Figure 12).[76] Site-specific labeling is the main advantage of this method, however, this method suffer with some limitations. For example, this method involves rigours protection-deprotection steps and hence desired modifications sometime do not survive in these stringent conditions. Further, coupling efficiency is significantly lower for bulky substrate and yields are poor for longer ONs sequences.[74]

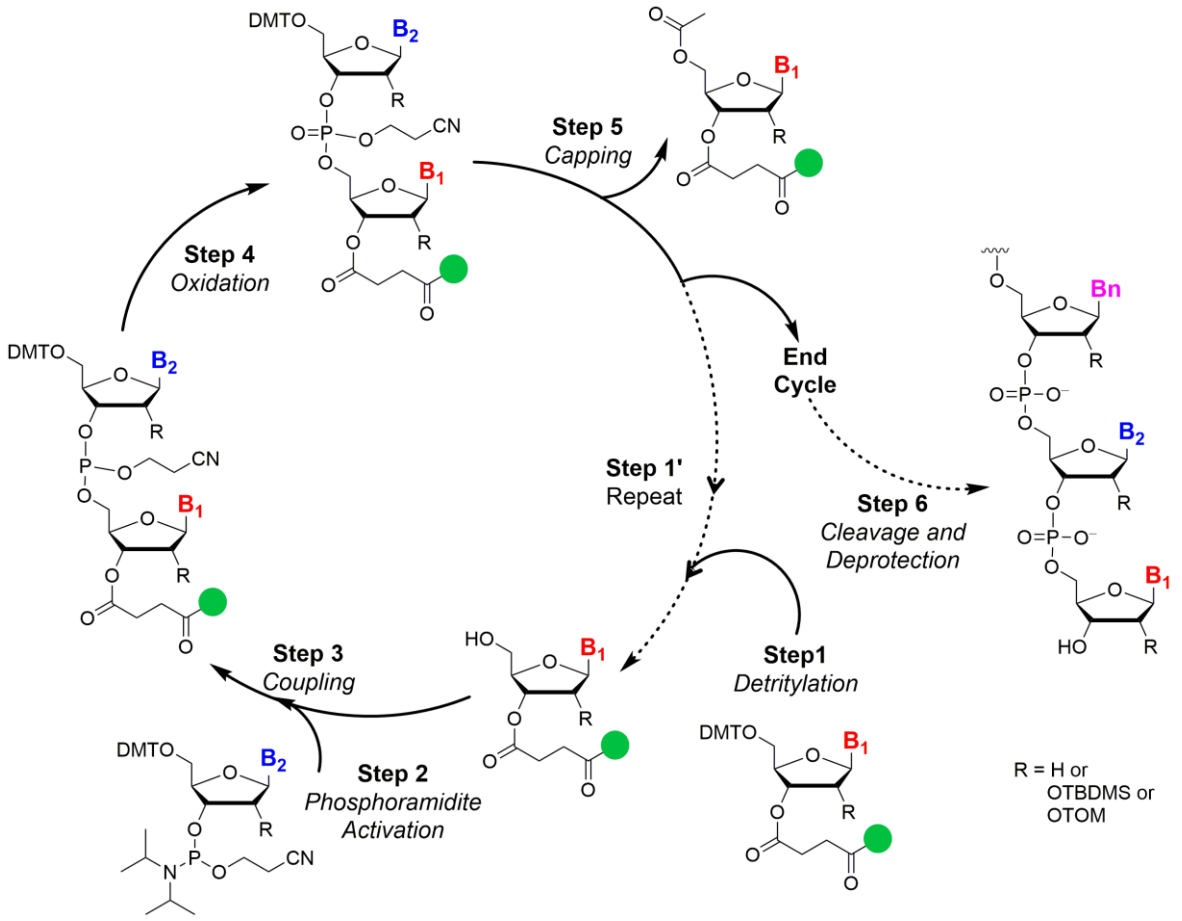


Figure 11. Standard cycle of nucleotide addition during solid-phase ON synthesis.[74]

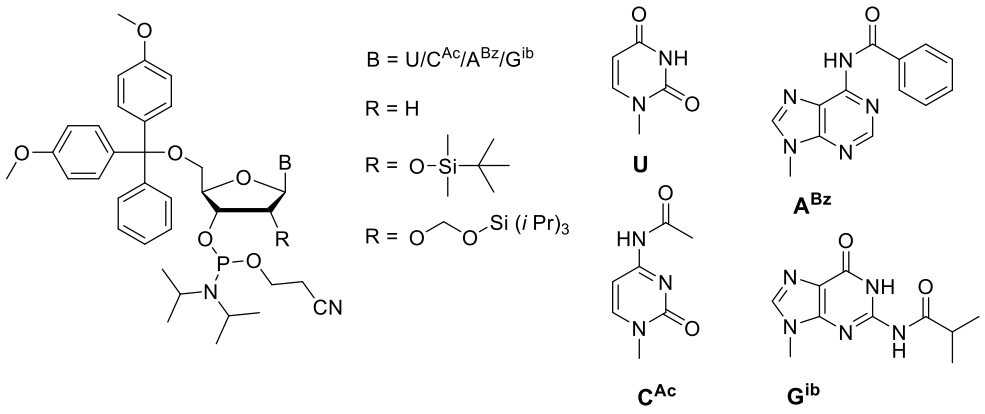


Figure 12. Different protecting sites in nucleoside to derive the desired phosporamedite substrate for solid-phase ON synthesis.

1.4.2 Enzymatic synthesis of ONs

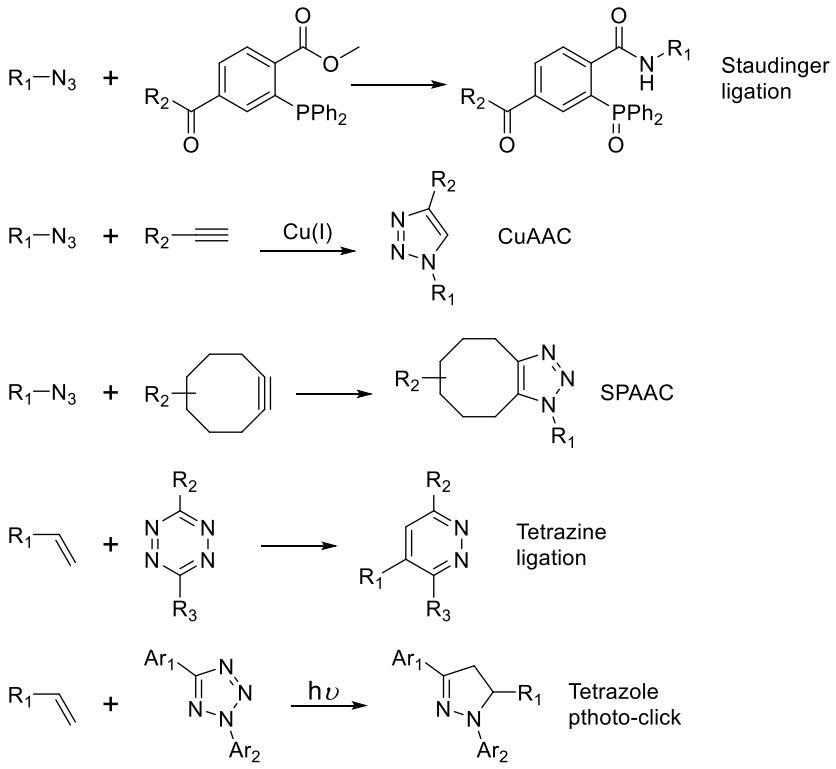
Modified nucleoside probes are incorporated into DNA ONs via DNA polymerase-catalysed primer extension and polymerase chain reaction.[47] T7 RNA polymerase catalysed *in vitro* transcription reaction is attractive alternative for the synthesis RNA ON containing modified

nucleoside probe.[57b] Unlike solid-phase chemical synthesis these methods operate under very mild conditions and hence the modified probe is unaffected during the course of incorporation into ONs. The modified triphosphate substrates necessary for these reactions are synthesized from corresponding nucleosides by performing a one-pot-two-step reaction in the presence of $POCl_3$ followed by a reaction with bistributylammonium pyrophosphate a method developed by Ludwig.[77] A large number of chemical strategies have been developed to effectively synthesize nucleoside triphosphates in the past decades.[77] However, many challenges remain in chemically selective triphosphate synthesis and purification due to the multiple functionalities of nucleosides, i.e., 5', 2', 3'-sugar hydroxyl groups and nucleobase amino groups. To address these issues Hung and co-workers have designed a mild and selective phosphorylating reagent. This elegant methodology provides high selectivity of the 5′-triphosphate synthesis which can be easily purified by RP-HPLC to offer high purity and yields.[77f, 77g] The success of these enzymatic methods greatly rely on the ability of the enzymes to accept and incorporate the modified triphosphates into growing chain of DNA and RNA ONs during the reaction. The incorporation efficiency of modified triphosphate is greatly relying on the size and functionality present in the modified-triphosphate. In particular, bulky modification on base and sugar-modified triphosphate perturb the native structure of the polymerase during various phases of transcription and resulted in to less yield or failure of the transcript. Enzymatic incorporation of modified nucleoside triphosphates, with T7 or a similar phage RNA polymerase, is particularly well suited for template-directed transcription reactions, provided that they are good substrates for the polymerase. This requirement limits the number of modified nucleotides that can be used in transcription reactions; thus only a few modifications can be enzymatically introduced into oligoribonucleotides. Nonetheless, transcription reaction can be used to produce reasonably large quantities of labeled RNAs by using small amounts of the triphosphate substrate a merit of this method over solid-phase chemical synthesis. Difficulties in site-specific incorporation is another disadvantages of this method but recently Hirao and co-workers have developed an elegant protocol for site-specific labeling of nucleic acids by *in vitro* replication and transcription by using unnatural base pair systems.[78]

1.4.3 Postsynthetic chemical functionalization

Recently, postsynthetic functionalization by using chemoselective reactions has emerged as a valuable tool to label glycans, proteins, lipids and nucleic acids for a variety of

applications.[79-81] In this methodology, a nucleoside containing an unnatural reactive group is incorporated into an ON sequence by chemical or enzymatic methods. Further functionalization is achieved postsynthetically by performing a bioorthogonal reaction between the labeled ON and a probe containing the cognate reactive functionality.[81] In the past decade, a number of reactions including Staudinger ligation, click reactions such as copper-catalysed azide-alkyne cycloaddition (CuAAC), copper free strain promoted azide-alkyne cycloaddition (SPAAC), photo-click reactions and tetrazine ligation have been well developed.[82] These reactions have been widely utilized for labeling biomolecules in the context of living cells and whole organisms (Scheme 1).[79a] Although these reactions are very well exploited to label the DNA in *in vitro* and in cell, postsynthetic RNA manipulations are less explored.



Scheme 1. Representative bioorthogonal chemical reactions used to label and image biomolecule in *in vitro* and in cells.

More than 80 years after the discovery of the Staudinger reaction, Saxon and Bertozzi modified this reaction such that it became a valuable method for metabolic engineering of cell surface glycans.[83] The Staudinger ligation was envisioned to meet most of the criteria necessary for a selective bioorthogonal ligation and since then this ligation method has been extensively used for various bioconjugations.[84] Weisbrod and Marx have successfully

shown the incorporation of azide-modified triphosphates **49** and **50** into DNA ON by primer extension and polymerase chain reactions (Figure 13). They have performed the Staudinger ligation reaction on DNA to label with affinity probe such as biotin.[85] Although the Staudinger ligation has received significant attention over the last years, limitations of this reaction arise with stability of the phosphine reagents in living system as well as slower kinetics of the reaction.[79a]

Copper-catalyzed 1,3-dipolar cycloaddition between azide and an alkyne has emerged as an excellent alternative to the Staudinger ligation. Currently, it is the most commonly used bioorthogonal chemical reaction for selective and controlled labeling of nucleic acids.[81] Carell and co-workers have initiated the incorporation of the two alkyne derivatives EdU (**51)** and **53** into a series of DNA ONs using corresponding phosphoramidite substrates. Furthermore, they have investigated multiple postsynthetic functionalization of alkyne-modified DNA ONs by using wide range of azide labels.[86] The Mitchison group developed a method for labeling DNA *in vivo* based on the incorporation of EdU into cellular DNA during DNA replication, and afterwards ethynyl functionality of EdU was used to image the DNA by performing CuAAC with fluorescent azides.[87] Subsequently, Salic group has applied this methodology to label RNA with EU (**52**) to image the cellular RNA by performing click reaction.[88]

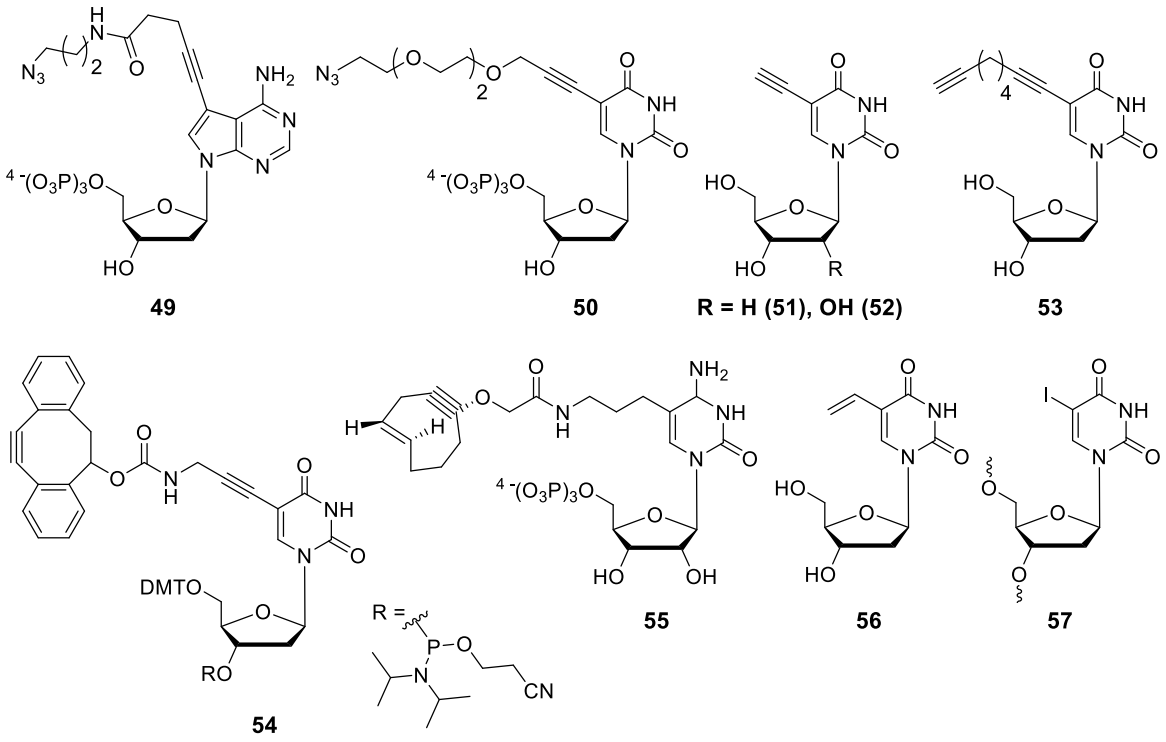


Figure 13. Various building blocks for postsynthetic chemical functionalization of nucleic acids.

In order to exclude the use cytotoxic copper, Bertozzi and co-workers have developed another biorthogonal chemical reaction called SPAAC reaction circumventing the need for copper. The incredible bioorthogonality and excellent kinetics of this reaction allows its application within cultured cells as well as in live zebrafish to label a glycans.[82b] Brown group has extended the utility of this reaction for internal DNA labeling by developing phosphoramidite monomer **54** which contains a strained cycloalkyne. The cyclooctyne ONs reacted efficiently with complementary azide probes under catalyst free reaction conditions.[89] This SPAAC chemistry may also be applicable for the synthesis of biologically important nucleic acid analogs that are inaccessible by conventional solid-phase synthesis. However, lack of efficient methods to metabolically incorporate the bulky cyclooctyne and azide functionality into nucleic acids for labeling and imaging in cells remains still a challenge.

Inverse-electron demand Diels–Alder (invDA) reactions between electron-deficient tetrazines and electron-rich dienophiles emerged as an attractive bioorthogonal chemical reaction as invDA reactions are compatible with cellular environment and importantly do not require a catalyst. Although, the tetrazine precursors are not easily available, invDA reactions have been widely used for labeling synthetic oligonucleotides *in vitro* as well as cellular and cell-surface proteins by using strained dienophiles such as norbornene, trans-cyclooctene, cyclopropene.[90] Royzen and co-workers reported the cytidine triphosphate analog **55** derivatized with a trans-cyclooctene group. This analog was efficiently incorporated into model RNA ON using *in vitro* transcription reactions. Furthermore, a reaction with fluorescein-labeled tetrazine was utilized to fluorescently tag the synthetic RNA without use of any catalyst.[91] Very recently, Rieder and Luedtke reported the metabolic incorporation of 5-vinyl-2'-deoxyuridine (**56**) into cellular DNA and visualized it by using a reactive fluorescent tetrazine.[92] By using the iodo-modified DNA ONs **57** Davis group described an efficient method for the introduction of a variety of sensitive and useful functional groups by Suzuki–Miyaura cross-coupling.[93] Together these approaches have provided an alternative method to conventional solid-phase and enzymatic methods for labeling nucleic acids for variety of applications.

1.5 Thesis outline

Although a significant number of minimally perturbing fluorescent nucleoside analogs with improved photophysical properties such as emission in visible region, high quantum yield and sensitivity to microenvironment have been developed, only a very few retain reasonable quantum yield when introduced into oligonucleotides. The progressive quenching exhibited by the majority of fluorophores upon incorporation into single stranded and double stranded ONs have essentially prohibited their implementation in certain fluorescence methods (e.g., anisotropy, single-molecule spectroscopy and cell microscopy). Further, there are few shortcomings in current chemical and enzymatic labeling methods as exemplified in the above sections. Therefore, there is a continuous demand for the development of robust nucleoside probes and labeling techniques, which can be used *in vitro* as well as in cell based analysis of nucleic acids.

We have been interested in the developing fluorescence nucleoside analogs which (i) are structurally minimally perturbing, (ii) are sensitive to their local environment, and (iii) retain appreciable fluorescence efficiency upon incorporation into DNA and RNA ONs. Towards this endeavour, we have developed environment-sensitive fluorescent nucleoside analogs by conjugating heterobicycles at the 5-position of uracil. The design strategy, synthesis and applications of heterobicycle-conjugated nucleosides are described in Chapters 2–4. In parallel, we have also developed a practical method to incorporate azide functionality into RNA ONs by transcription reaction using azide-modified UTP analogs. This is the first example of enzymatic incorporation of azide functionality into RNA, which also enabled the posttranscriptional chemical functionalization of RNA with a variety of biophysical probes in a modular fashion by using Staudinger ligation, CuAAC and SPAAC reactions. The design, synthesis and enzymatic incorporation of azide-modified UTPs into RNA, and development of posttranscriptional chemical functionalization protocols are described in Chapter 5 of this thesis.

1.6 References

[1] a) T. Hermann, D. Patel, *Structure* **2000**, *8*, R47–R54; b) N. Leulliot, G. Varani, *Biochemistry* **2001**, *40*, 7947–7956; c) M. Mandal, R. R. Breaker, *Nat. Rev. Mol. Cell Biol*. **2004**, *5*, 451–463; d) Z. Shajani, P. Deka, G. Varani, *Trends Biochem Sci*. **2006**, *31*, 421–424; e) D. M. J. Lilley, *Philosoph. Transac. Roy. Soc B*. **2011**, *366*, 2910–2917.

[2] a) K. B. Hall, *Curr. Opin. Chem. Biol*. **2008**, *12*, 612–618; b) H. M. Al-Hashimi, N. G. Walter, *Curr. Opin. Struct. Biol.* **2008**, *18*, 321–329; c) N. G. Walter, *Methods* **2009**, *49*, 85–86.

[3] a) U. Asseline, *Curr. Org. Chem*. **2006**, *10*, 491–518; b) S. G. Srivatsan, M. Famulok, *Comb. Chem. High Throughput Screening* **2007**, *10*, 698–705; c) X. Shi, D. Herschlag, *Methods Enzymol.* **2009**, *469*, 287–302; d) B. Juskowiak, *Anal. Bioanal. Chem*. **2011**, *399*, 3157–3176.

[4] a) V. J. DeRose, *Curr. Opin. Struct. Biol*. **2003**, *13*, 317–324; b) N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, J. W. Engels, *Nucleic Acids Res.* **2007**, *35*, 3128–3143; c) C. E. Aitken, A. Petrov, J. D. Puglisi, *Annu. Rev. Biophys*. **2010**, *39*, 491–513; d) M. F. Bardaro Jr, G. Varani, *WIREs RNA* **2012**, *3*, 122–132; e) P. Nguyen, P. Z. Qin, *WIREs RNA* **2012**, *3*, 62–72.

[5] a) J. M. Ogle, A. P. Carter, V. Ramakrishnan, *Trends Biochem. Sci*. **2003**, *28*, 259–266; b) A. Korostelev, S. Trakhanov, M. Laurberg, H. Noller, *Cell* **2006**, *126*, 1065–1077; c) S. R. Holbrook, *Annu. Rev. Biophys*. **2008**, *37*, 445–464; d) A. Serganov, D. J. Patel, *Curr. Opin. Struct. Biol*. **2012**, *22*, 279–286.

[6] T. Xia, *Curr. Opin. Chem. Biol*. **2008**, *12*, 604–611.

[7] a) T. Ha, *Biochemistry* **2004**, *43*, 4055–4063; b) R. Zhao, D. Rueda, *Methods* **2009**, *49*, 112–117.

[8] a) S. H. Weisbrod, A. Marx, *Chem. Commun*. **2008**, 5675–5685; b) F. Wachowius, C. Höbartner, *ChemBioChem* **2010**, *11*, 469–480; c) O. Khakshoor, E. T. Kool, *Chem. Commun*. **2011**, *47*, 7018–7024.

[9] a) J. Peon, A. H. Zewail, *Chem. Phys. Lett.* **2001**, *348*, 255–262; b) E. Nir, K. Kleinermanns, L. Grace, M. S. de Vries, *J. Phys. Chem. A* **2001**, *105*, 5106–5110.

[10] a) R. T. Ranasinghe, T. Brown, *Chem. Commun.* **2005**, 5487–5502; b) A. A. Martí, S. Jockusch, N. Stevens, J. Ju, N. J. Turro, *Acc. Chem. Res.* **2007**, *40*, 402–409.

[11] a) S. Tyagi, D. P. Bratu, F. R. Kramer, *Nat. Biotechnol*. **1998**, *16*, 49–53; b) A. Jenne, J. S. Hartig, N. Piganeau, A. Tauer, D. A. Samarsky, M. R. Green, J. Davies, M. Famulok, *Nat. Biotechnol.* **2001**, *19*, 56–61; c) M. Dorywalska1, S. C. Blanchard, R. L. Gonzalez Jr, H. D. Kim, S. Chu, J. D. Puglisi, *Nucleic Acids Res.* **2005**, *33*, 182–189; d) M. Hafner, A. Schmitz, I. Grüne, S. G. Srivatsan, B. Paul, W. Kolanus, T. Quast, E. Kremmer, I. Bauer, M. Famulok, *Nature* **2006**, *444*, 941–944; e) M. Stoop, C. J. Leumann, *Chem. Commun*. **2011**, *47*, 7494–7496.

[12] M. J. Rist, J. P. Marino, *Curr. Org. Chem.* **2002**, *6*, 775–793.

[13] R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev*. **2010**, *110*, 2579–2619.

[14] S. G. Srivatsan, A. A. Sawant, *Pure Appl. Chem*. **2011**, *83*, 213–232.

[15] K. Phelps, A. Morris, P. A. Beal, *ACS Chem. Biol*. **2012**, *7*, 100–109.

[16] K. B. Hall. *Methods Enzymol.* **2009**, *469*, 269–285.

[17] a) M. C. Wahl, M. Sundaralingam, *Curr. Opin. Chem. Biol.* **1995**, *5*, 282–295; b) M. Egli, *Curr. Opin. Chem. Biol.* **2004**, *8*, 580–91.

[18] B. H. Mooers, *Methods* **2009**, *47*, 168–176.

[19] E. Ennifar, P. Carpentier, J. L. Ferrer, P. Waltera, P. Dumas, *Acta Crystallogr. D Biol. Crystallogr.* **2002,** *58*, 1262-1268.

[20] L. Lin, J. Sheng, Z. Huang, *Chem. Soc. Rev.* **2011**, *40*, 4591–4602.

[21] Q. Du, N. Carrasco, M. Teplova, C. J. Wilds, M. Egli, Z. Huang, *J. Am. Chem. Soc.* **2002**, *124*, 24–25.

[22] C. Höbartner, R. Rieder, C. Kreutz, B. Puffer, K. Lang, A. Polonskaia, A. Serganov, R. Micura, *J. Am. Chem. Soc.* **2005**, *127*, 12035–12045.

[23] a) G. Varani , F. Aboul-ela , F. H.-T. Allain , *Progress in NMR spectroscopy* **1996**, *29*, 51–127; b) B. Fürtig, C. Richter, J. Wöhnert, H. Schwalbe, *ChemBioChem.* **2003**, *4*, 936–962; c) A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *J. Am. Chem. Soc.* **2007**, 129, 4386–4392. c) J. Rinnenthal, J. Buck, J. Ferner, A. Wacker, B. FÜrtig, H. Schwalbe, *Acc. Chem. Res.* **2011**, *44*, 292–1301; d) J. R. Bothe, E. N. Nikolova, C. D. Eichhorn, J. Chugh, A. L. Hansen, H. M. Al-Hashimi, *Nat. Methods.* **2011**, *8*, 919–931.

[24] a) P. Wenter, S. Pitsch, *Helv. Chim. Acta* **2003**, *86*, 3955–3974; b) P. Wenter, G. Bodenhausen, J. Dittmer, S. Pitsch, *J. Am. Chem. Soc.* **2006**, *128*, 7579– 7587.

[25] H. L. Schultheisz, B. R. Szymczyna, L. G. Scott, J. R. Williamson, *ACS Chem Biol.* **2008,** *3*, 499–511.

[26] G. F. Salgado, C. Cazenave, A. Kerkour, J. L. Mergny, *Chem. Sci.* **2015**, *6*, 3314–3320.

[27] S. L. Cobb, C. D. Murphy, *J. Fluorine Chem.* **2009**, **130**, 132–143.

[28] a) F. Rastinejad, C. Evilia, P. Lu, *Methods Enzymol.* **1995**, *261*, 560–275; b) C. Hammann, D. G. Norman, D. M. Lilley, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5503–5508 c) C. Kreutz, H. Kählig, R. Konrat, R. Micura, *J. Am. Chem. Soc.* **2005**, *127*, 11558–11559; d) N. Jain, Y. Li, L. Zhang, S. Meneni, B. Cho, *Biochemistry* **2007,** *46*, 13310–13321; e) N. B. Barhate, R. N. Barhate, P. Cekan, G. Drobny, S. T. Sigurdsson, *Org. Lett.* **2008,** *10*, 2745–2747; f) J. Riedl, R. Pohl, L. Rulíšek, M. Hocek *J. Org. Chem.* **2012**, *77*, 1026–1044; g) T. Lombés, R. Moumné, V. Larue, E. Prost, M. Catala,

T. Lecourt, F. Dardel, L. Micouin, C. Tisné, *Angew. Chem. Int. Ed.* **2012**, *51*, 9530–9534.

[29] B. Puffer, C. Kreutz, U. Rieder, M. O. Ebert, R. Konrat, R. Micura, *Nucleic Acids Res.* **2009,** *37*, 7728–7740.

[30] I. T. Suydam, S. A. Strobel, *J. Am. Chem. Soc.* **2008**, *130*, 13639–13648.

[31] V. J. DeRose, *Curr. Opin. Struct. Biol.* **2003**, *13*, 317–324.

[32] P. Z. Qin, K. Hideg, J. Feigon, W. L. Hubbell, *Biochemistry* **2003**, *42*, 6772–6783.

[33] N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, J. W. Engels, *Nucleic Acids Res.* **2007**, *35*, 3128–3143.

[34] V. Singh, M. Azarkh, T. E. Exner, J. S. Hartig, M. Drescher, *Angew. Chem. Int. Ed.* **2009**, *48*, 9728–9730.

[35] I. Krstić, R Hänsel, O. Romainczyk, J. W. Engels, V. Dötsch, T. F. Prisner, *Angew. Chem. Int. Ed.* **2011**, 50, 5070–5074.

[36] S. Saha, A. P. Jagtap, S. Th. Sigurdsson, *Chem. Commun*. **2015**, 51, 13142–13145.

[37] J. N. Wilson, E. T. Kool, *Org. Biomol. Chem.* **2006**, *4*, 4265–4274.

[38] A. Okamoto, Y. Saito, I. Saito, *J. Photochem. Photobiol. C* **2005**, *6*, 108–122.

[39] a) A. Okamoto, K. Tainaka, I. Saito, *J. Am. Chem. Soc.* **2003**, *125*, 4972–4973; b) A. Okamoto, K. Tainaka, T. Fukuta, I. Saito, *J. Am. Chem. Soc*. **2003**, *125*, 9296–9297.

[40] K.Y. Lin, R. J. Jones, M. Matteucci, *J. Am. Chem. Soc*. **1995**, *117*, 3873–3874.

[41] a) L. M. Wilhelmsson, A. Holmén, P. Lincoln, P. E. Nielsen, B. Nordén. *J. Am. Chem. Soc.* **2001**, *123*, 2434–2435; b) P. Sandin, K. Borjesson, H. Li, J. Martensson, T. Brown, L. M. Wilhelmsson, B. Albinsson, *Nucleic Acids Res.* **2008**, *36*, 157–167.

[42] K. Borjesson, S. Preus, A. H. El-Sagheer, T. Brown, B. Albinsson, L. M. Wilhelmsson, *J. Am. Chem. Soc.* **2009**, *131*, 4288–4293.

[43] a) P. Cekan, S. T. Sigurdsson, *Chem. Commun*. **2008**, 3393–3395; b) H. Gardarsson, A. S. Kale, S. T. Sigurdsson, *ChemBioChem* **2011**, *12*, 567–575.

[44] A. Dierckx, F. A. Miannay, N. B. Gaied, S. Preus, M. Björck, T. Brown, L. M. Wilhelmsson, *Chem. Eur. J*. **2012**, *18*, 5987–5997.

[45] a) K. Tainaka, K. Tanaka, S. Ikeda, K.-I Nishiza, T. Unzai, Y. Fujiwara, I. Saito, A. Okamoto, *J. Am. Chem. Soc.* **2007**, *129*, 4776–4784; b) Y. Saito, K. Motegi, S. S. Bag, I. Saito, *Bioorg. Med. Chem.* **2008**, *16*, 107–113; c) N. A. Grigorenko, C. J. Leumann, *Chem. Commun.* **2008**, 5417–5419; d) T. Ehrenschwender, H.-A, Wagenknecht, *J. Org. Chem.* **2011**, *76*, 2301–2304; e) S. Ikeda, T. Kubota, D. O. Wang, H. Yanagisawa, T. Umemoto, A. Okamoto, *ChemBioChem* **2011**, *12*, 2871–2880; f) A. Okamoto, K.

Sugizaki, M. Yuki, H. Yanagisawa, S. Ikeda, T. Sueoka, G. Hayashi, D. O. Wang, *Org. Biomol. Chem*. **2013**, *11*, 363–371.

[46] a) M. Kimoto, T. Mitsui, Y. Harada, A. Sato, S. Yokoyama, I. Hirao, *Nucleic Acids Res*. **2007**, *35*, 5360–5369; b) A. Dumas, N. W. Luedtke, *J. Am. Chem. Soc.* **2010**, *132*, 18004–18007; c) W. Hirose, K. Sato, A. Matsuda, *Angew. Chem. Int. Ed*. **2010**, *49*, 8392–8394; d) A. Dierckx, P. Dinér, A. H. El-Sagheer, D. K. Joshi, T. Brown, M. Grøtli, L. M. Wilhelmsson, *Nucleic Acids Res.* **2011**, *39*, 4513–4524; e) Y. Saito, S. Miyamoto, A. Suzuki, K. Matsumoto, T. Ishihara, I. Saito, *Bioorg. Med. Chem. Lett*. **2012**, *22*, 2753–2756, f) A. Suzuki, N. Takahashi, Y. Okada, I. Saito, N. Nemoto, Y. Saito, *Bioorg. Med. Chem. Lett*. **2013**, *23*, 886–892; g) J. Krim, C. Grünewald, M. Taourirte, J. C. Engels, *Bioorg. Med. Chem*. **2012**, *20*, 480–486; h) P. Guo, X. Xu, X. Qiu, Y. Zhou, S. Yan, C. Wang, C. Lu, W. Ma, X. Weng, X. Zhang, X. Zhou, *Org. Biomol. Chem*. **2013**, *11*, 1610–1613.

[47] a) J. Riedl, R. Pohl, N. P. Ernsting, P. Orság, M. Fojtac, M. Hocek, *Chem. Sci.* **2012**, *3*, 2797–2806; b) J. Riedl, P. Ménová, R. Pohl, P. Orság, M. Fojtac, M. Hocek, *J. Org. Chem.* **2012**, *77*, 8287−8293.

[48] a) N. Venkatesan, Y. J. Seo, E. K. Bang, S. M. Park, Y. S. Lee, B. H. Kim, *Bull. Korean Chem. Soc*. **2006**, *27*, 613–630; b) S. M. Park, S.-J. Nam, H. S. Jeong, W. J. Kim, B. H. Kim, *Chem. Asian J*. **2011**, *6*, 487–492; c) II J. Lee, B. H. Kim, *Chem. Commun.* **2012**, *48*, 2074–2076.

[49] M. E. Hawkins, *Methods Enzymol.* **2008**, *450*, 201–231.

[50] a) M. E. Hawkins, W. Pfleiderer, A. Mazumder, Y. G. Pommier, F. M. Balis, *Nucleic Acids Res*. **1995**, *23*, 2872–2880; b) M. E. Hawkins, W. Pfleiderer, F. M. Balis, D. Porter, J. R. Knutson, J. R. *Anal. Biochem*. **1997**, *244*, 86–95; c) M. E. Hawkins, W. Pfleiderer, O. Jungmann, F. M. Balis, *Anal. Biochem.* **2001**, *298*, 231–240.

[51] J. Parsons, T. Hermann. *Tetrahedron* **2007**, *63*, 3548–3552.

[52] Y. N. Teo, E. T. Kool, *Chem. Rev*. **2012**, *112*, 4221−4245.

[53] a) F. Wojciechowski, J. Lietard, C. J. Leumann, *Org. Lett*. **2012**, *14*, 5176–5179; b) T. Ono, S. Wang, C.-K. Koo, L. Engstrom, S. S. David, E. T. Kool, *Angew. Chem. Int. Ed.* **2012**, *51*, 1689–1692; c) S. S. Bag, S. Talukdar, K. Matsumoto, R. Kundu, *J. Org. Chem*. **2013**, *78*, 278–291.

[54] a) C.-K. Koo, F. Samain, N. Dai, E.T. Kool, *Chem. Sci.* **2011**, *2*, 1910–1917; b) S. S. Tan, S. J. Kim, E. T. Kool, *J. Am. Chem. Soc*. **2011**, *133*, 2664–2671; c) J. Guo, S. Wang, N. Dai, Y. N. Teo, E. T. Kool, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3493–

3498; d) S. Wang, J. Guo, T. Ono, E. T. Kool, *Angew. Chem. Int. Ed*. **2012**, *51*, 7176–7180.

[55] D. C. Ward, E. Reich, L. Stryer, *J. Biol. Chem*. **1969**, *244*, 1228–1237.

[56] a) M. Kawai, M. J. Lee, K. O. Evans, T. M. Nordlund, *J. Fluorescence* **2001**, *11*, 23–32; b) E. L. Rachofsky, R. Osman, J. B. A. Ross, *Biochemistry* **2001**, *40*, 946–956.

[57] a) N. J. Greco, Y. Tor, *J. Am. Chem. Soc*. **2005**, *127*, 10784–10785; b) S. G. Srivatsan, Y. Tor, *J. Am. Chem. Soc.* **2007**, *129*, 2044–2053; c) S. G. Srivatsan, Y. Tor, *Tetrahedron* **2007**, *63*, 3601–3607; d) N. J. Greco, R. W. Sinkeldam, Y. Tor, *Org. Lett.* **2009**, *11*, 1115–1118.

[58] a) S. G. Srivatsan, N. J. Greco, Y. Tor, *Angew. Chem. Int. Ed.* **2008**, *47*, 6661–6665; b) S. G. Srivatsan, H. Weizman, Y. Tor, *Org. Biomol. Chem.* **2008**, *6*, 1334–1338.

[59] a) Y. Xie, A. V. Dix, Y. Tor, *J. Am. Chem. Soc.* **2009**, *131*, 17605–17614; b) Y. Xie, A. V. Dix, Y. Tor, *Chem. Commun*. **2010**, *46*, 5542–5544; c)Y. Xie, T. Maxson, Y. Tor, *J. Am. Chem. Soc*. **2010**, *132*, 11896–11897.

[60] D. Shin, R.W. Sinkeldam, Y. Tor, *J. Am. Chem. Soc*. **2011**, *133*, 14912–14915.

[61] R. W. Sinkeldam, P. A. Hopkins, Y. Tor, *ChemPhysChem* **2012**, *13*, 3350–3356.

[62] D. A. Berry, K.-Y. Jung, D. S. Wise, A. D. Sercel, W. H. Pearson, H. Mackie, J. B. Randolph, R. L. Somers, *Tetrahedron Lett*. **2004**, *45*, 2457–2461.

[63] R. A. Tinsley, N. G. Walter, *RNA* **2006**, *12*, 522–529.

[64] C.-M. Zhang, C. Liu, T. Christian, H. Gamper, J. Rozenski, D. Pan, J. B. Randolph, E. Wickstrom, B. S. Cooperman, Y.-M. Hou, *RNA* **2008**, *14*, 2245–2253.

[65] a) X. Zhang, R. M. Wadkins, *Biophys. J.* **2009**, *96*, 1884–1891; b) Q. Zhang, Y. Wang, X. Meng, R. Dhar, H. Huang, *Anal. Chem*. **2013**, *85*, 201–207.

[66] R. H. E. Hudson, A. G. Choghamarani. *Nucleosides, Nucleotides and Nucleic Acids*, **2007**, *26*, 533–537.

[67] a) X. Ming, F. Seela, *Chem. Eur. J.* **2012**, *18*, 9590–9600; b) X. Ming, P. Ding, P. Leonard, S. Budow, F. Seela, *Org. Biomol. Chem*. **2012**, *10*, 1861–1869.

[68] M. S. Noé, A. C. Ríos, Y. Tor, *Org. Lett.* **2012**, *14*, 3150–3153.

[69] A. S. Wahba, A. Esmaeili, M. J. Damha, R. H. E. Hudson, *Nucleic Acids Res*. **2010**, *38*, 1048–1056.

[70] A. S. Wahba, F. Azizi, G. F. Deleavey, C. Brown, F. Robert, M. Carrier, A. Kalota, A. M. Gewirtz, J. Pelletier, R. H. E. Hudson, M. J. Damha, *ACS Chem. Biol*. **2011**, *6*, 912–919.

[71] N. B. Gaied, N. Glasser, N. Ramalanjaona, H. Beltz, P. Wolff, R. Marquet, A. Burger, Y. Mely, *Nucleic Acids Res.* **2005**, *33*, 1031–1039.

[72] A. Nadler, J. Strohmeier, U. Diederichsen, *Angew. Chem. Int. Ed*. **2011**, *50*, 5392–5396.

[73] M. Weinberger, F. Berndt, R. Mahrwald, N. P. Ernsting, H. A. Wagenknecht, *J. Org. Chem*. **2013**, *78*, 2589–2599.

[74] G. M. Blackburn, M. J. Gait, D. Loakes, D. M. Williams, *Nucleic Acids in Chemistry and Biology* **2006**, *3rd edition*, RSC publications.

[75] a) M. J. Gait, *Oligonucleotide Synthesis: A Practical Approach* **1984,** Oxford University Press; b) S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. L.Wu, *Helv. Chim. Acta* **2001**, *84*, 3773–3795; c) B. S. Sproat, *Methods Mol. Biol.* **2005**, *288*, 17–32.

[76] a) L. J. McBride, R. Kierzek, S. L. Beaucage, M. H. Caruthers, *J. Am. Chem. Soc.* **1986**, *108*, 2040–2048; b) M. P. Reddy, N. B. Hanna, F. Farooqui, *Nucleosides & Nucleotides* **1997**, *16*, 1589–1598.

[77] a) J. A. Ludwig, *Acta Biochim. Biophys. Acad. Sci. Hung.* **1981**, *16*, 131–13; b) K. Burgess, D. Cook, *Chem. Rev.* **2000**, *100*, 2047–2060; c) W. Wu, C. L. Freel Meyers, R. F. Borch, *Org. Lett.* **2004**, *6*, 2257–2260; d) A. T. Horhota, J. W. Szostak, L. W. McLaughlin, *Org. Lett.* **2006**, *8*, 5345–5347; e) Q. Sun, J. P. Edathil, R. Wu, E. D. Smidansky, C. E. Cameron, B. R. Peterson, *Org. Lett.* **2008**, *10*, 1703–1706, f) J. Caton-Williams, L. Lin, M. Smith, Z. Huang, *Chem. Commun.* **2011**, *47*, 8142–8144; g) J. Caton-Williams, M. Smith, N. Carrasco, Z. Huang, *Org. Lett.* **2011**, *13*, 4156–4159.

[78] M. Kimoto, Y. Hikida, and I. Hirao, *Israel J. Chem*. **2013**, *53*, 450–468.

[79] a) J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.* **2005**, *1*, 13−21; b) C. D. Spicer, B. G. Davis, *Nat. Commun.* **2014**, *5*:4740 doi: 10.1038/ncomms5740.

[80] a) A. G. Neef, C. Schultz, *Angew. Chem. Int. Ed.* **2009**, *48*, 1498−1500.

[81] a) P. M. E. Gramlich, C. T. Wirges, A. Manetto, T. Carell, *Angew. Chem. In. Ed.* **2008,** *47*, 8350−8358; b) A. H. El-Sagheer, T. Brown, *Chem. Soc. Rev.* **2010**, *39*, 1388−1405; c) D. Schulz, A. Rentmeister, *ChemBioChem* **2014**, *15*, 2342–2347.

[82] a) M. Köhn, R. Breinbauer, *Angew. Chem. Int. Ed.* **2004**, 43, 3106–3116; b) J. C. Jewetta, C. R. Bertozzi, *Chem. Soc. Rev.* **2010**, *39*, 1272–1279; c) J. E. Hein, V. V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1302–1315; d) N. K. Devaraj, R. Weissleder, *Acc. Chem. Res.* **2011**, *44*, 816–827.

[83] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.

[84] S. S. van Berkel, M. B. van Eldijk, J. C. M. van Hest, *Angew. Chem. Int. Ed.* **2011**, *50*, 8806−8827.

[85] S. H. Weisbrod, A. Baccaro, A. Marx, *Nucleic Acids Symposium Series* **2008**, *52* 383–384.

[86] J. Gierlich, G. A. Burley, P. M. E. Gramlich, D. M. Hammond, T. Carell, *Org. Lett.* **2006**, *8*, 3639–3642.

[87] A. Salic, T. J. Mitchison, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2415−2420.

[88] C.Y. Jao, A. Salic, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15779−15784.

[89] M. Shelbourne, T. Brown Jr, A. H. El-Sagheer, T. Brown, *Chem. Commun.* **2012**, *48*, 11184–11186.

[90] a) A.-C. Knall, C. Slugovc, *Chem. Soc. Rev.* **2013**, *42*, 5131−5142; b) J. Šečkutė, N. K. Devaraj, *Curr. Opin. Chem. Biol.* **2013**, *17*, 761−767.

[91] P. N. Asare-Okai, E. Agustin, D. Fabris, M. Royzen, *Chem. Commun*. **2014**, *50*, 7844–7847.

[92] U. Rieder, N. W. Luedtke, *Angew. Chem. Int. Ed.* **2014**, *53*, 1–6.

[93] L. Lercher, J. F. McGouran, B. M. Kessler, C. J. Schofield, B. G. Davis, *Angew. Chem. Int. Ed.* **2013**, *52*, 10553–10558.

Chapter 2

Chapter 2: Detection of abasic sites in DNA and RNA oligonucleotides by using fluorescent nucleoside analogs

Chapter 2A

5-Benzofuran-modified uridine as a probe for the fluorescence detection of DNA abasic site

2A.1 Introduction

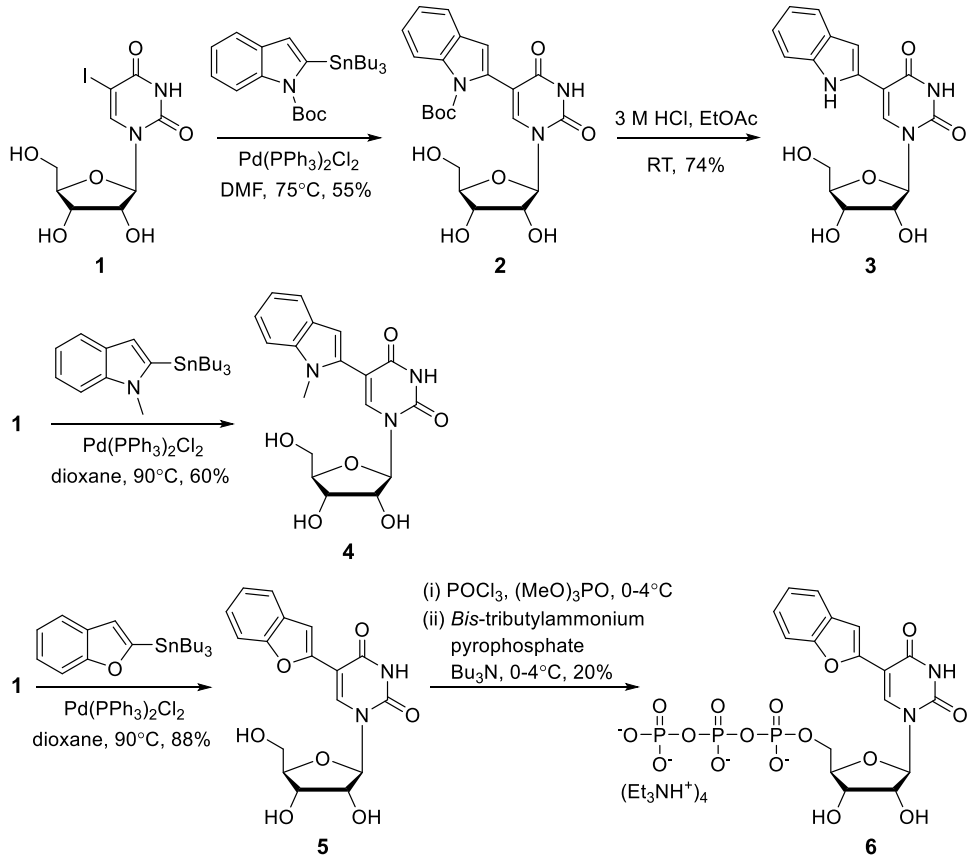
Fluorescent nucleoside analogs that photophysically report changes in nucleic acid conformation have become very important bioanalytical tools in advancing the understanding of nucleic acid structure and function.[1-7] Nucleic acids lack intrinsic fluorescence as natural nucleobases are practically nonemissive.[8] In order to impart better photophysical properties to otherwise nonemissive natural nucleobases, most design principles rely on naturally occurring fluorescent heterocycles, polycyclic aromatic hydrocarbons or on extending the π-conjugation by appending heterocycles to the bases.[6] Utilizing these design strategies, several classes of base-modified fluorescent nucleobase analogs such as size-expanded,[9] extended,[6,10] pteridine,[11] polycyclic aromatic hydrocarbon[9] and isomorphic,[6] have been developed. A significant number of these analogs that display useful photophysical properties have been implemented in several DNA-based, and to a relatively lesser extent in RNA-based biophysical assays.[12,13] Notably, the majority of these analogs, except isomorphic bases, substantially deviate from the native structure of the nucleobases. Isomorphic analogs are advantageous because they maintain the Watson-Crick (WC) base pairing, and minimally perturb the native structure and function of target oligonucleotides (ONs). In particular, a highly emissive and environment-sensitive adenosine analogue, 2-aminopurine (2AP), has been extensively utilized in designing several DNA- and RNA-based biophysical and discovery assays.[7,14,15] However, an excitation and emission maximum in the UV region and drastically low quantum yields displayed by 2AP, when incorporated into ONs, greatly limits its application to in vitro systems only.[15-17]

Numerous base-modified ribonucleoside analogs with emission in the visible region and a high quantum yield have been developed, but only a few analogs have been effectively implemented in exploring the functions of RNA molecules.[7] In particular, a very few structurally nonperturbing pyrimidine ribonucleoside analogs have shown promise as useful fluorescent probes.[17-19] In order to expand the repertoire of useful fluorescent pyrimidine ribonucleoside analogs, we sought to synthesize microenvironment-sensitive pyrimidine ribonucleoside analogs that are structurally nonperturbing, display emission in the visible region, and maintain a reasonable quantum yield upon incorporation into oligoribonucleotides (RNA ONs).

We have drawn inspiration from a naturally occurring fluorescent amino acid, tryptophan, to construct a series of emissive pyrimidine ribonucleoside analogs. Tryptophan, an indole derivative, is reasonably emissive, and its emission properties are highly sensitive to

its local environment.[20] Therefore, we hypothesize that appending bicyclic heterocycles such as an indole moiety to a nucleobase (e.g. 5-position of uracil) may enhance the π-conjugation and impart better photophysical properties to the nucleobase.[21,22] Here, we describe the synthesis and photophysical characterization of a focused set of new pyrimidine ribonucleoside analogs derived by tagging indole, *N*-methylindole and benzofuran at the 5-position of uracil. Notably, benzofuran-conjugated uridine analogue is appreciably emissive with an emission maximum in the visible region, and also exhibits excellent solvatochromism. We also report the synthesis of the corresponding ribonucleoside triphosphate substrate, and its effective incorporation into RNA ONs by transcription reactions catalyzed by T7 RNA polymerase. Furthermore, the emissive uridine analogue, when incorporated into oligoribonucleotide, positively signals the presence of a DNA abasic site.

2A.2 Results and Discussion



Scheme 1. Synthesis of modified ribonucleoside analogs **3**, **4** and **5**, and ribonucleotide analogue **6**.

2A.2.1 Synthesis of ribonucleoside analogs

Modified uridine analogs **3**, **4** and **5** have been synthesized according to the procedure illustrated in Scheme 1. In a two-step reaction, indole-conjugate uridine **3** has been synthesized by performing a palladium catalyzed cross-coupling reaction between 5-iodouridine **1** and tributylstannyl derivative of *N*-Boc-protected indole, followed by deprotection of the Boc-group in the presence of 3 M HCl. *N*-methylindole- (**4**) and benzofuran-conjugated uridine (**5**) have been synthesized by reacting 5-iodouridine with corresponding stannylated heterocycles in the presence of palladium catalyst.

2A.2.2 Photophysical characterization of ribonucleoside analogs

An important attribute exhibited by many conformation-sensitive fluorescent nucleoside analogue probes is that they photophysically respond to changes in solvent polarity.[6] Therefore, prior to incorporation into RNA ONs, photophysical properties have been examined in solvents of different polarity to evaluate the responsiveness of modified ribonucleosides to polarity changes. Increasing solvent polarity has marginal impact on the ground state electronic spectrum of modified ribonucleosides **3**–**5** (Figure 1–3 and Table 1-2). However, quantum yield, emission maximum and fluorescence lifetime are significantly affected by solvent polarity. Unexpectedly, indole-conjugated (**3**) and *N*-methylindole-conjugated (**4**) uridine analogs show weak fluorescence in nonpolar solvents, and practically no fluorescence in water (Figure 1, 2 and Table 1). Rewardingly, when excited at its lowest energy maximum (322 nm), the benzofuran-conjugated uridine **5** in water exhibits a very strong emission band (λ_{em}=447 nm) corresponding to a quantum yield of 0.212 ± 0.002 (Figure 3, Table 2). As the solvent polarity is sequentially decreased from water to dioxane, the ribonucleoside shows significant hypsochromic shift (447 to 404 nm) and a nearly 2-fold quenching in fluorescence intensity. Furthermore, relative quantum yield determined using 2AP as a standard in various solvents is in good agreement with the observed fluorescence intensity in respective solvents (Table 2). The responsiveness of ribonucleoside to changes in solvent polarity environment is also confirmed by plotting the Stokes shift in various solvents as a function of Reichardts's microscopic solvent polarity parameter, $E_T(30)$ (Figure 4).[23] A positive correlation between the Stokes shift and $E_T(30)$ further establishes the sensitiveness of the emissive ribonucleoside **5** to its surrounding environment.

Table 1. Photophysical properties of modified ribonucleosides **3** and **4** in various solvents.

Ribonucleoside	Solvent	λ_{max}[a] (nm)	λ_{em}[b] (nm)	Φ
3	Water	330	-	-
	Methanol	330	-	-
	Acetonitrile	340	458	$2.85 \times 10^{-2} \pm 3.0 \times 10^{-3}$
	Dioxane	340	448	$5.81 \times 10^{-2} \pm 3.9 \times 10^{-3}$
4	Water	286	-	-
	Methanol	286	-	-
	Acetonitrile	290	494	$2.09 \times 10^{-3} \pm 1.6 \times 10^{-4}$
	Dioxane	293	484	$6.72 \times 10^{-3} \pm 9.3 \times 10^{-4}$

[a] The lowest energy maximum is given. [b] In water and methanol, ribonucleosides **3** and **4** are practically nonfluorescent. See Figure 1 and 2 for absorption and emission spectra.

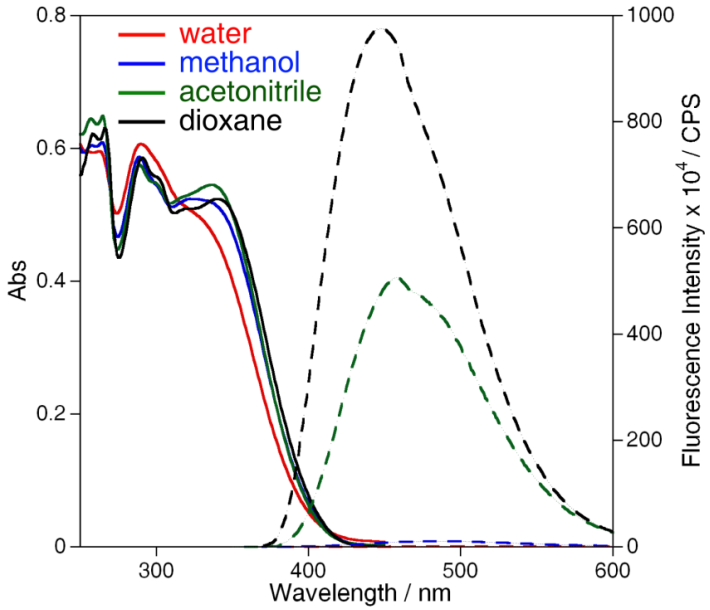


Figure 1. Absorption (50 µM, solid lines) and emission (50 µM, dashed lines) spectra of ribonucleoside **3** in various solvents. Samples were excited at respective lowest energy absorption maximum in various solvents (Table 2). Excitation and emission slit widths were maintained at 2 and 3 nm, respectively. All solutions for absorption and emission studies contained 5% DMSO. In water and methanol, ribonucleoside **3** is practically nonfluorescent

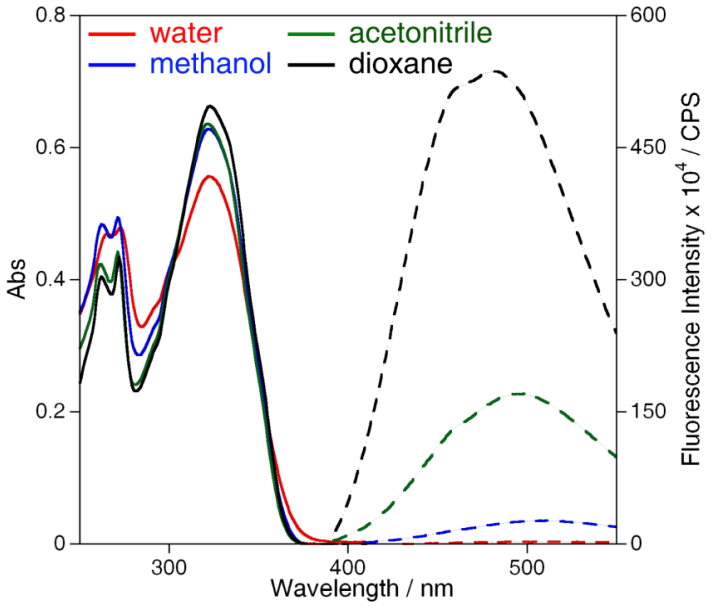


Figure 2. Absorption (50 µM, solid lines) and emission (50 µM, dashed lines) spectra of ribonucleoside **4** in various solvents. Samples were excited at respective lowest energy absorption maximum in various solvents (Table 2). Excitation and emission slit widths were maintained at 4 and 7 nm, respectively. All solutions for absorption and emission studies contained 5% DMSO. In water and methanol, ribonucleoside **4** is practically nonfluorescent.

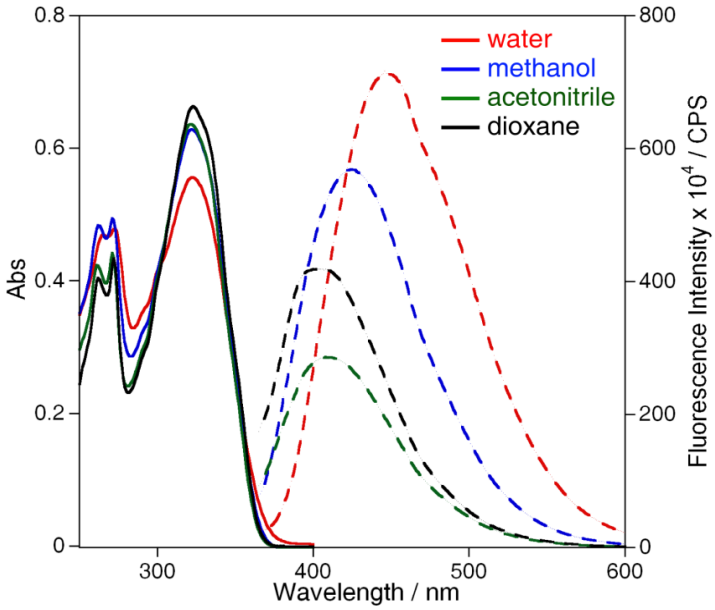


Figure 3. Absorption (25 µM, solid lines) and emission (5.0 µM, dashed lines) spectra of ribonucleoside 5 in various solvents. Samples were excited at 322 nm. Excitation and emission slit widths were maintained at 2 and 4 nm, respectively. All solutions for absorption and emission studies contained 2.5% and 0.5% DMSO, respectively.

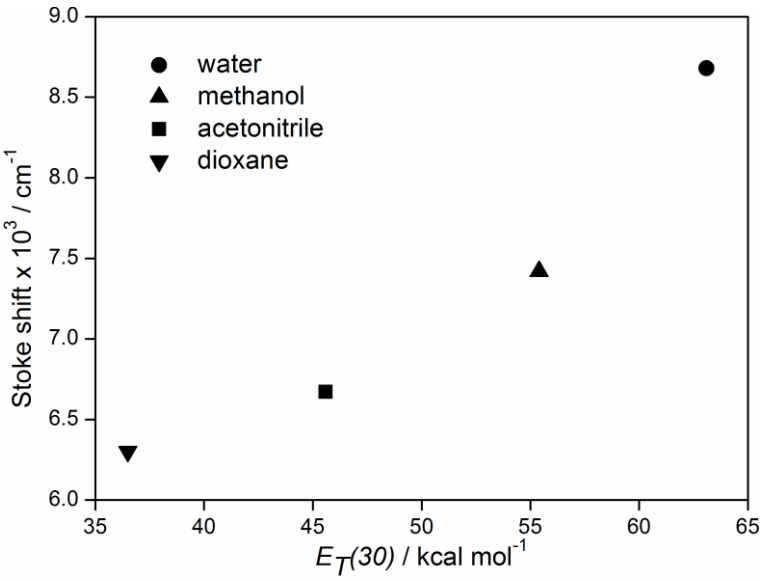


Figure 4. A plot of Stokes shift vs. $E_T(30)$ for ribonucleoside **5** in various solvents.

Time-resolved fluorescence spectroscopic measurements have also been performed to study the effect of solvent polarity on the excited state decay kinetics of ribonucleoside **5** (Table 2, Figure 5). Decay profile of the emissive ribonucleoside in solvents of different polarity reveals distinct decay kinetics. The ribonucleoside in water has the highest lifetime of 2.55 ns. As the solvent polarity is decreased from water to dioxane, a huge decrease in the lifetime (~6-fold) is observed in dioxane as compared to water. The decreasing trend in the lifetime is consistent with the intensity displayed by ribonucleoside in various solvents.

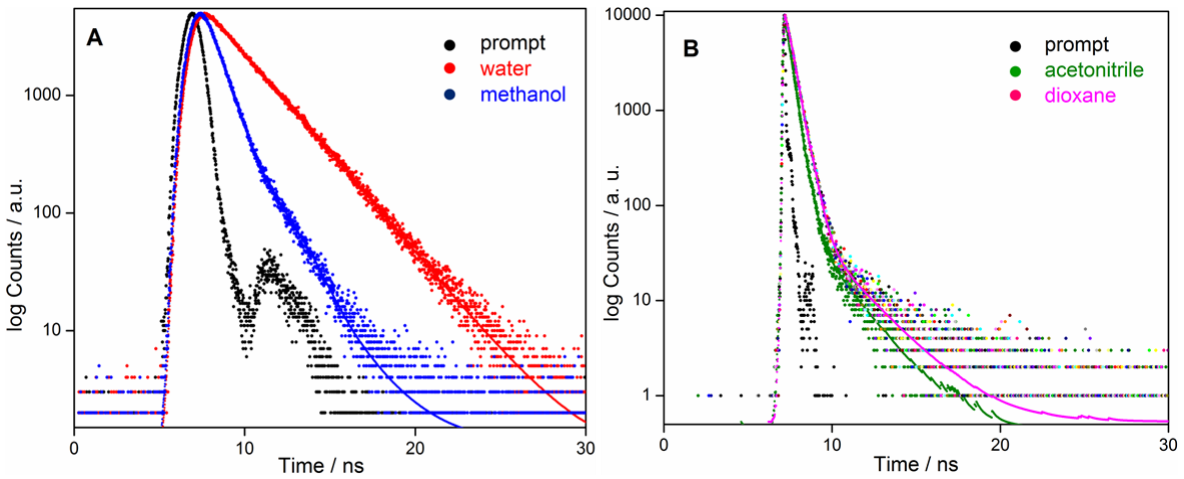


Figure 5. Excited state decay profile of ribonucleoside **5** in various solvents. Laser profile is shown in black (prompt). Curve fits are shown in solid lines.

Table 2. Photophysical properties of fluorescent ribonucleoside **5** in various solvents

Solvent	λ_{max}[a] [nm]	λ_{em} [nm]	I_{rel}[b]	Φ[c]	τ_{av}[c] [ns]	*r*
Water	322	447	1.0	0.212	2.55	0.016
Methanol	322	423	0.8	0.145	0.94	-
Acetonitrile	322	410	0.4	0.063	0.33	-
Dioxane	322	404	0.6	0.099	0.43	-
Ethylene glycol	322	428	1.9	0.377	3.13	0.125
Glycerol	322	427	2.4	0.528	3.35	0.315

[a] The lowest energy maximum is given. [b] Emission intensity relative to intensity in water. [c] Standard deviations for Φ and τ_{av} are $\leq$ 0.004 and 0.02 ns, respectively. k_r and k_{nr} are radiative and nonradiative decay rate constants, respectively. *r* = fluorescence anisotropy.

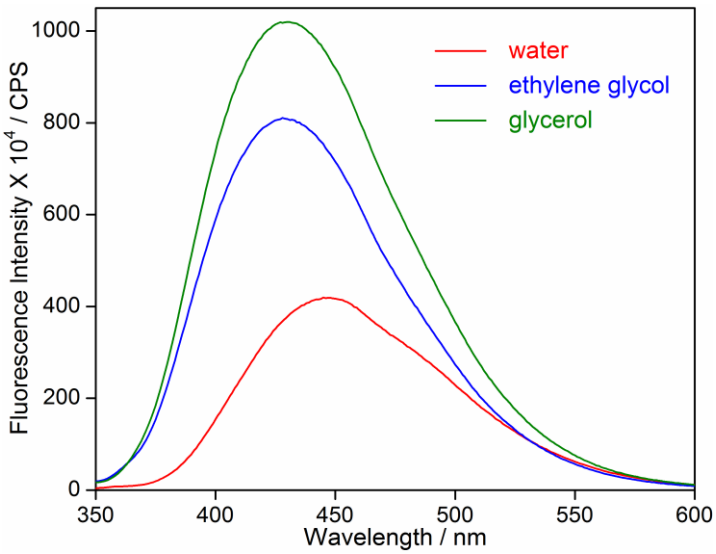


Figure 6. Emission (5.0 µM) spectra of ribonucleoside **5** in solvents of different viscosity. Samples were excited at 322 nm. Excitation and emission slit widths were maintained at 1 and 10 nm, respectively. All solutions contained 0.5% DMSO.

The aromatic heterobicyclic ring and the pyrimidine ring is separated by a rotatable aryl-aryl bond in fluorescent nucleoside analog **5**. The relative conformation of the two ring systems about the rotatable bond, which could be sensitive to molecular crowding effects and viscosity, could affect the conjugation and hence, the fluorescence properties of the nucleosides.[21] The conformation-sensitivity of nucleoside analog containing a molecular rotor element was evaluated by performing photophysical characterizations in solvents of similar polarity (to eliminate polarity effect) but different viscosity. In a viscous medium (glycerol) nucleoside **5** displayed a significant enhancement in fluorescence intensity with no apparent change in the emission maximum as compared to in a less viscous medium (ethylene

glycol, Figure 6 and Table 2). Fluorescence lifetime and anisotropy measurements also revealed longer lifetimes and higher anisotropy values in glycerol than in ethylene glycol because of the rigidification of the fluorophore in a more viscous medium (Figure 7, Table 2).

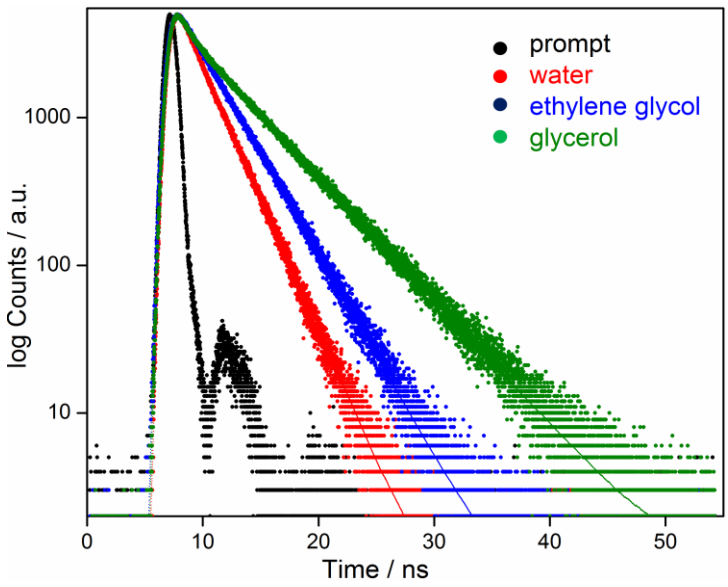


Figure 7. Excited state decay profile of ribonucleoside **5** in solvents of different viscosity. Laser profile is shown in black (prompt). Curve fits are shown in solid lines.

Properties such as emission in the visible region, a modest quantum yield, and responsiveness to polarity changes as illustrated by fluorescence measurements adequately confer probe-like attribute to the fluorescent ribonucleoside analogue **5**. Collectively, these photophysical characterizations indicated that the emissive nucleoside analog **5** is indeed environment- and conformation-sensitive, properties that are highly suitable for studying the nucleic acid function. Therefore, ribonucleoside **5** has been selected for incorporation into RNA ONs by transcription reactions. The corresponding ribonucleoside triphosphate substrate **6** necessary for in vitro transcription reactions has been synthesized by employing the method developed by Ludwig.[24] The modified triphosphate **6** has been synthesized in a one-pot two-step reaction by reacting ribonucleoside **5** with freshly distilled $POCl_3$, followed by a reaction with *bis*-tributylammonium pyrophosphate (Scheme 1).

2A.2.3 Enzymatic incorporation

Although, solid-phase chemical synthesis is the method of choice for synthesizing modified RNA ONs, enzymatic methods such as transcription and ligation reactions have also been effectively utilized in labeling RNAs.[25-27] In order to investigate the ability of T7 RNA

polymerase to incorporate the modified triphosphate **6** into RNA ONs, in vitro transcription reactions were conceived as explained in Figure 8. A series of promoter-template duplexes were assembled by annealing a DNA promoter strand to DNA templates, **D1**–**D5**. The templates were designed to contain one or two deoxyadenosine residues at different positions to direct single or double incorporations of the modified triphosphate **6**. Additionally, a deoxythymidine residue was placed at the 5'-end of each template strand to direct the addition of a single adenosine at the 3'-end of each transcript. Hence, in vitro transcription reactions in the presence of GTP, CTP, UTP/**6** and α-^{32}P ATP, if successful, would result in the formation of full-length oligoribonucleotide transcripts containing a ^{32}P-labeled adenosine at the 3'-end. Reaction products could then be resolved by analytical denaturing polyacrylamide gel electrophoresis and imaged. Failed transcription reactions resulting in transcripts shorter than the full-length products would not carry the ^{32}P-labeled adenosine, and hence, would remain undetected.

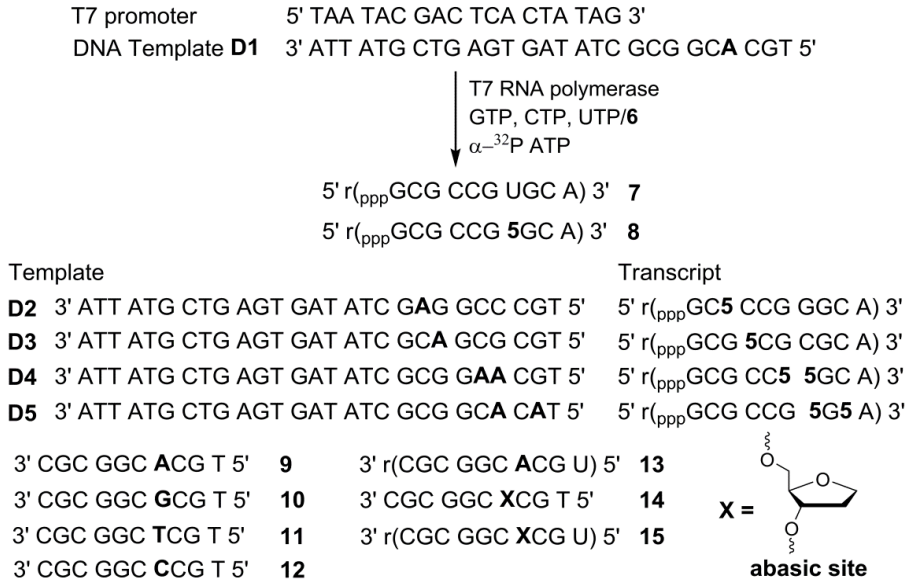


Figure 8. Incorporation of ribonucleoside triphosphate 6 by transcription reactions in the presence of DNA templates **D1**–**D5**. Synthetic ONs (**9**–**15**) used in this study. "r" preceding the parentheses with the sequence represents RNA oligonucleotides.

A transcription reaction in the presence of template **D1** results in the formation of the 10-mer modified full-length RNA transcript **8** containing the modified ribonucleoside at the +7 position (Figure 9, lane 2). Along with the full-length product, trace amounts of N+1 and N+2 products are also formed due to nontemplated random incorporation of ribonucleotides.[28] The triphosphate **6** is incorporated into the oligoribonucleotide with a moderate efficiency of 67 % relative to a reaction in the presence of natural UTP. Discernible

retardation in the mobility of modified transcript **8** with respect to unmodified transcript **7** clearly indicates the inclusion of a higher molecular weight emissive ribonucleoside during the transcription reaction (Figure 9, compare lanes 1 and 2). When both UTP and **6** are not included in a control transcription reaction, the full-length RNA product is not produced (Figure 9, lane 3). This result shows that the formation of full-length transcripts is not due to adventitious misincorporation. Interestingly, in a reaction containing UTP and **6** in equimolar concentrations, T7 RNA polymerase incorporates both natural and modified UTP into RNA transcripts (Figure 9, lane 4).

In order to test the ability of T7 RNA polymerase to incorporate the modification near the promoter region, transcription reactions were performed with templates **D2** and **D3**, which would direct the incorporation of **6** at +3 and +4 positions, respectively.[29] The incorporation efficiencies were found to be lower for reactions with templates **D2** and **D3** (Figure 9, lanes 6 and 8). Reactions in the presence of template **D4** and **D5** resulted in double incorporation of **6** in successive as well as in alternating sites (Figure 9, lanes 10 and 12). Albeit moderate yields, reactions with various templates clearly demonstrate the usefulness of in vitro transcription reactions in generating singly- and doubly-modified fluorescent RNA ONs.

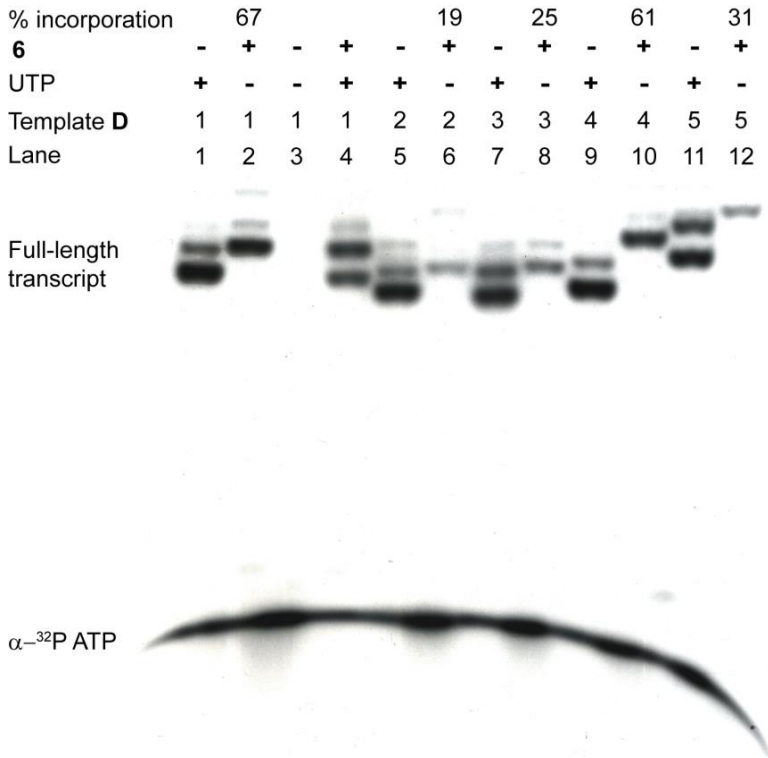


Figure 9. Polyacrylamide gel electrophoresis of transcripts obtained from in vitro transcription reactions with DNA templates **D1**−**D5** in the presence of UTP and modified UTP **6** under denaturing condition. Percentage incorporation of **6** is reported with respect to the amount of full-length product formed in the presence of UTP.

2A.2.4 Characterization of fluorescent oligoribonucleotide transcripts

Large-scale transcription reactions were performed in the presence of non-radiolabeled triphosphates and template **D1** to isolate transcript **8** for further chemical and photophysical characterizations. A typical reaction after gel electrophoretic purification gave nearly 14 nmol of the modified oligoribonucleotide product. The presence of benzofuran-conjugated uridine **5** in the oligoribonucleotide **8** was confirmed by MALDI-TOF mass analysis (Figure 10). Furthermore, reversed-phase HPLC analysis of ribonucleoside products obtained from an enzymatic digestion reaction of transcript **8** clearly revealed the presence of modified ribonucleoside **5** (Figure 11). The ribonucleoside composition determined from the area under the individual ribonucleoside peak also matched well with the sequence of the full-length oligoribonucleotide. Mass analysis of HPLC fractions corresponding to the retention time of each ribonucleoside unambiguously established the authenticity of the natural and modified ribonucleoside **5** present in transcript **8** (Table 3). Together, these results clearly reveal the incorporation of the emissive ribonucleotide into RNA ONs by RNA polymerase in transcription reactions.

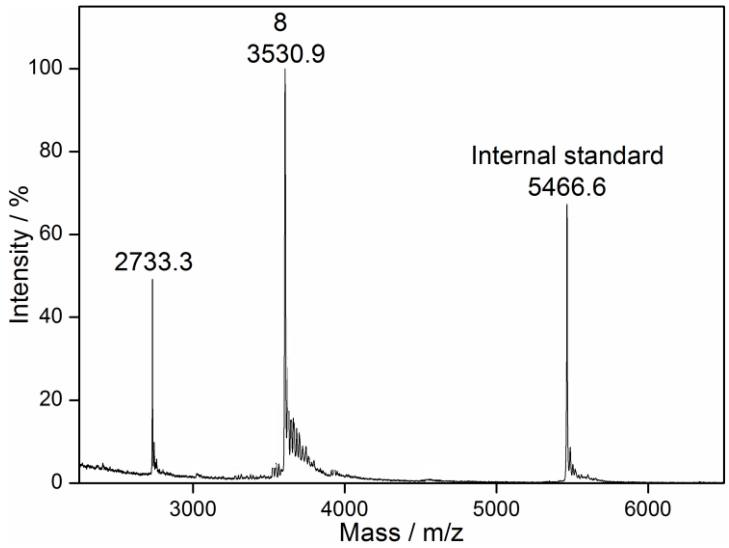


Figure 10. MALDI-TOF MS spectrum of the modified transcript **8** calibrated relative to the +1 ion of an internal 18-mer DNA ON standard (m/z: 5466.6). m/z 2732.6 is +2 ion of the internal DNA standard. Calcd. for **8**: 3531.0 [M]; found: 3530.9.

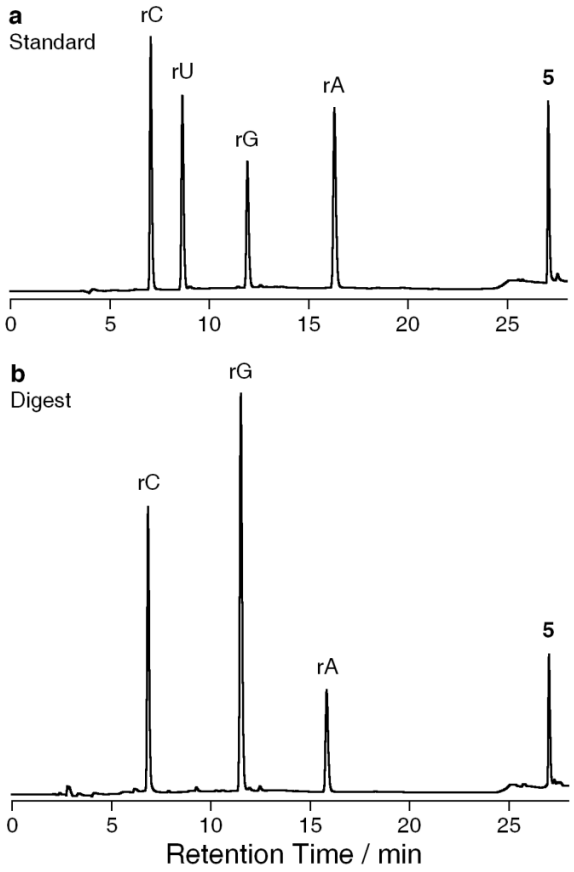


Figure 11. HPLC profile of ribonucleoside products obtained from an enzymatic digestion of transcript 8 at 260 nm. (a) Natural ribonucleosides and fluorescently modified ribonucleoside 5 mix. (b) Digested oligoribonucleotide **8**.

Table 3. MALDI-TOF mass analysis of HPLC fractions of oligoribonucleotide **8** digest.

HPLC fractions of the digest	Calcd. for	Found
rC	$C_9H_{13}N_3O_5K$: 282.0 [M+K^+]	282.2
rG	$C_{10}H_{13}N_5O_5K$: 322.1 [M+K^+]	322.2
rA	$C_{10}H_{13}N_5O_4$: 267.1 [M^+]	268.2
5	$C_{17}H_{16}N_2O_7$: 360.1 [M^+]	360.2

2A.2.5 Photophysical characterization of fluorescent RNA ONs

Fluorescence properties of emissive nucleoside analogs when incorporated into ONs are known to be influenced by a variety of mechanisms involving neighbouring bases.[15,21,30,31] Therefore, steady-state and time-resolved fluorescence spectroscopic measurements have been carried out with oligoribonucleotide **8** and duplexes of **8** to evaluate the influence of neighbouring bases on the fluorescence of ribonucleoside **5**. A series of duplexes have been

assembled by hybridizing oligoribonucleotide **8** to complementary DNA and RNA ONs in which the ribonucleoside **5** has been placed opposite to its complementary or mismatch bases (Figure 8).

The photophysical properties of oligoribonucleotide **8** reveal characteristic features that are different from those of the free ribonucleoside **5** in aqueous buffer. Upon excitation, the single stranded oligoribonucleotide **8** shows an emission band centered at 447 nm corresponding to a quantum yield of 0.040 ± 0.001, which is nearly 5-fold lower as compared to the free ribonucleoside (Figure 12, Table 4). Unlike free ribonucleoside, which exhibits monoexponential fluorescence intensity decay kinetics, oligoribonucleotide **8** displays biexponential decay kinetics (Table 4). The biexponential decay kinetics consists of a shorter and a longer lifetime component contributing equally to the overall lifetime (τ_{av}=2.21). Similar decay profiles have been reported for 2AP in which the free nucleoside exhibits decay kinetics corresponding to a single lifetime, but when incorporated into ONs shows decay kinetics corresponding to multiple lifetimes.[32] The quenching of ribonucleoside fluorescence in **8** can be possibly due to electron transfer process between the modified base and neighbouring bases, and or alterations in the conformation of the benzofuran moiety with respect to the nucleobase.[21,31-33] Several emissive nucleoside analogs containing conjugated and fused heterocycles, when incorporated into ONs, show similar fluorescence quenching.[6]

The fluorescence quenching is much more pronounced in RNA-DNA (**8•9**) and RNA-RNA (**8•13**) duplexes in which the emissive ribonucleoside **5** is placed opposite to its complementary bases, dA and A, respectively (Figure 12). Duplexes **8•9** and **8•13** show a slightly blue shifted emission maximum (~439 nm) corresponding to a quantum yield of 0.0085 ± 1.0 x 10^{-4} and 0.0082 ± 2.1 x 10^{-4}, which is nearly 4- and 5-fold lower than that of the single stranded oligoribonucleotide **8**, respectively (Table 4). The perfect duplexes also show biexponential decay kinetics similar to **8**. The drastic quenching of fluorescence intensity accompanied by a small spectral shift in perfect duplexes **8•9** and **8•13** can be possibly due to a combination of the following reasons: (a) electron transfer process between the base-paired fluorescent analogue and flanking bases, (b) desolvation effect, (c) and alterations in the conformation of the labeled uridine.[15,21,31]

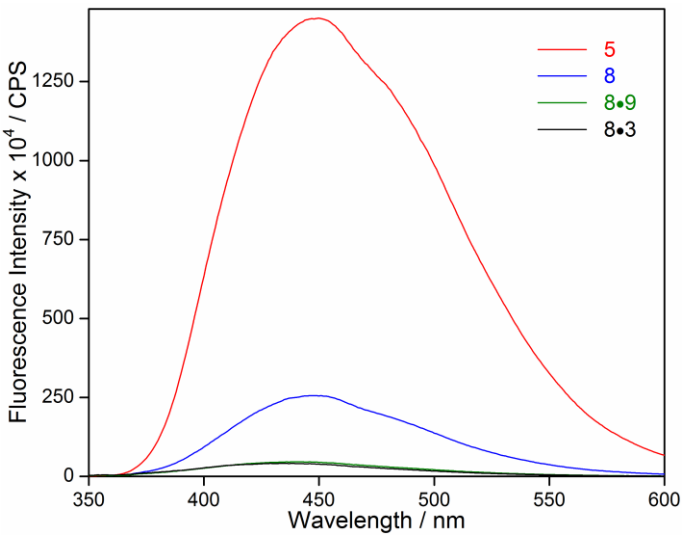


Figure 12. Emission spectra (1 µM) of ribonucleoside **5**, single stranded RNA transcript **8**, and duplexes **8•9** and **8•13** in 20 mM cacodylate buffer (pH 7.1). Samples were excited at the absorption maximum of the ribonucleoside (322 nm) with an excitation slit width of 4 nm and emission slit width of 7 nm.

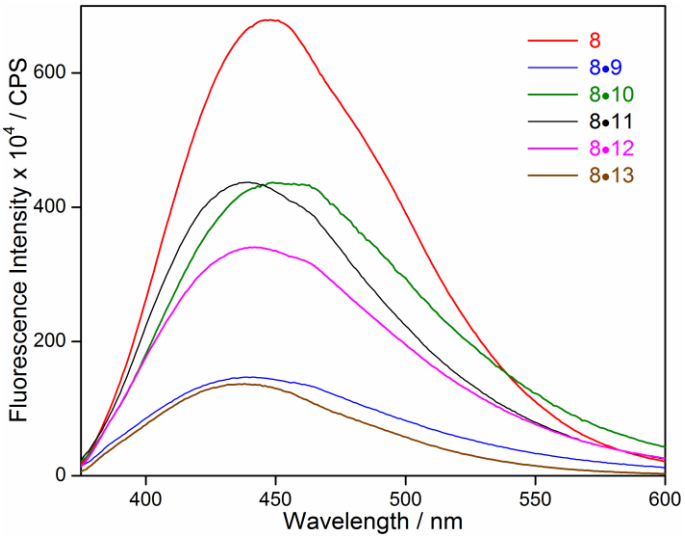


Figure 13. Emission spectra (1 µM) of oligoribonucleotide transcript **8** and duplexes containing emissive ribonucleoside **5** opposite to complementary base (**8•9** and **8•13)** and mismatched bases (**8•10**, **8•11**, and **8•12**) in 20 mM cacodylate buffer (pH 7.1). Samples were excited at 330 nm with an excitation slit width of 7 nm and emission slit width of 9 nm.

Interestingly, mismatched duplexes **8•10**, **8•11** and **8•12** containing **5** opposite to dG, dT and dC, respectively, exhibit 2 to 3-fold higher fluorescence intensity compared to the perfect duplexes (Figure 13). Similar to perfect duplexes, all mismatched duplex constructs display biexponential decay kinetics with varying shorter and longer lifetime components (Table 4). Collectively, these results demonstrate the ability of ribonucleoside **5** to photophysically distinguish between a single stranded ON and a perfect complementary

duplex, albeit with quenched emission. Nevertheless, this trait of ribonucleoside can be utilized in hybridization assays in which the kinetics of formation and dissociation of a perfect duplex can be studied by monitoring the changes in the photophysical properties of the fluorescently modified oligoribonucleotide.[18b,34]

Table 4. Photophysical properties of ON constructs containing ribonucleoside **5**.

Sample	λ_{em} (nm)	Φ[a]	τ_1 (ns)	τ_2 (ns)	τ_{av}[a] (ns)
5	447	0.203	2.55	-	2.55
8	447	0.040	0.92 (0.49)	3.47 (0.51)	2.21
8•9	439	0.009	0.32 (0.89)	3.64 (0.11)	0.70
8•10	449	ND	0.56 (0.65)	3.44 (0.35)	1.58
8•11	439	ND	0.35 (0.95)	3.14 (0.05)	0.50
8•12	443	ND	0.24 (0.94)	3.55 (0.06)	0.44
8•13	438	0.008	0.40 (0.88)	3.93 (0.12)	0.84
8•14	447	0.034	0.30 (0.78)	3.27 (0.22)	0.94
8•15	440	0.014	0.27 (0.89)	3.66 (0.11)	0.65

[a] Standard deviations for Φ and τ_{av} are $\leq$ 0.001 and 0.03 ns, respectively. Relative amplitude is given in parenthesis. k_r and k_{nr} are radiative and nonradiative decay rate constants, respectively. ND= not determined

2A.2.6 Fluorescence detection of a DNA abasic site

Abasic sites are frequently occurring DNA lesions that are vulnerable to strand breaks, and if unrepaired can be toxic as well as mutagenic.[35,36] Most methods that have been developed to detect abasic sites either use aldehyde-reactive probes, which irreversibly react with abasic sites of isolated DNA, or fluorescent probes, which show changes in fluorescence properties when placed opposite to abasic sites.[12d,37,38] Drastic quenching in fluorescence intensity upon placing the fluorescent probe **5** opposite to complementary bases prompted us to study the effect of abasic sites on the fluorescence properties of **5**. Abasic site containing duplexes have been constructed by hybridizing RNA transcript **8** to custom DNA and RNA ONs **14** and **15**, which contain a chemically stable abasic site surrogate, tetrahydrofuran (Figure 8). When excited at 330 nm, the abasic site-containing duplex **8•14** shows a strong emission band (λ_{em}=447 nm) corresponding to a quantum yield of 0.034 ± 0.001, which is nearly 4-fold higher than that of the perfect duplex **8•9** (Figure 14, Table 4).

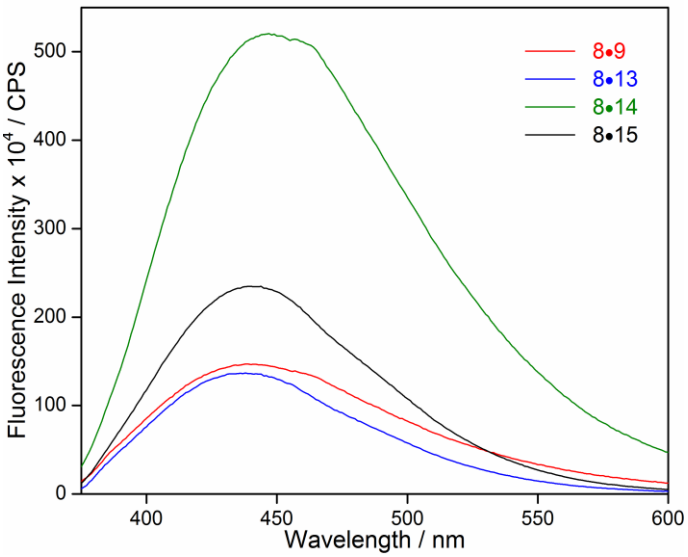


Figure 14. Emission spectra (1 μM) of duplexes containing the emissive ribonucleoside **5** opposite to complementary bases (**8•9** and **8•13)**, and abasic sites (**8•14** and **8•15)** in 20 mM cacodylate buffer (pH 7.1). Samples were excited at 330 nm with an excitation slit width of 7 nm and emission slit width of 9 nm.

Interestingly, an RNA-RNA duplex **8•15** possessing an abasic site opposite to **5** exhibits only a slight enhancement in fluorescence intensity as compared to its perfect duplex **8•13** (Figure 14, Table 4). Comparable emission maximum of duplex **8•14** and free ribonucleoside **5** in aqueous buffer indicates that the solvent polarity environment around the fluorophore in **8•14** and free nucleoside is similar (Table 4). In addition, a 4-fold higher k_r/k_{nr} ratio, calculated from Φ and τ, points out that the radiative pathway is appreciably favoured in the abasic site-containing duplex **8•14** over the perfect duplex (Table 4). Therefore, the emissive ribonucleoside, when incorporated into oligoribonucleotide, preferentially signals the presence of a DNA abasic site with a significant enhancement in fluorescence intensity.

2A.2.7 Stability of duplexes

The modified ribonucleoside upon incorporation can potentially perturb the native structure of RNA ONs, which can lead to ineffective hybridization. The observed fluorescence properties of duplexes will then be a consequence of a mixture of the more emissive single stranded oligoribonucleotide **8** and corresponding duplex. In order to investigate the influence of benzofuran modification on the duplex stablity, thermal denaturation and radioactive native gel mobility shift experiments were performed under the conditions employed in the fluorescence experiments. A slightly lower T_m values for modified duplexes

as compared to the corresponding control unmodified duplexes indicated a marginal destabilization due to benzofuran modification (Figure 15 and Table 5).

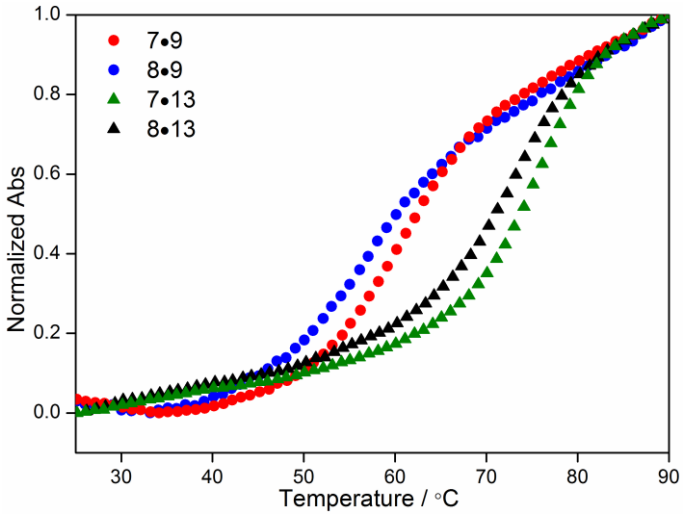


Figure 15. UV-thermal melting of control unmodified and fluorescently-modified duplexes (1 μM) in 20 mM cacodylate buffer (pH 7.1, 500 mM NaCl, 0.5 mM EDTA). Duplexes were hybridized by heating a 1:1 mixture of ONs at 90°C for 3 min, and cooling the solutions slowly to room temperature. Samples were placed in crushed ice for at least 30 min before analysis. T_m values are given below (Table 5).

Table 5. T_m values of control and modified duplexes

Control unmodified duplexes	T_m (°C)	Fluorescently modified duplexes	T_m (°C)
7•9	60.5 ± 0.6	**8•9**	58.3 ± 0.6
7•13	74.3 ± 0.7	**8•13**	73.5 ± 0.6
7•15	54.7 ± 1.0	**8•15**	54.5 ± 0.4

^{32}P-labeled transcript **8** synthesized by transcription reaction in the presence of template **D1** and α-^{32}P ATP was annealed to custom ONs **9**–**15**. The hybridized ONs were resolved on a non-denaturing polyacrylamide gel, and imaged. The gel-shift experiments revealed complete hybridization (Figure 16). Taken together, these results clearly indicate that the benzofuran moiety does not affect the hybridization efficiency. Hence, it can be concluded that the observed differences in the fluorescence properties of completely intact duplexes are exclusively due to the differences in the microenvironment of fluorescent ribonucleoside.

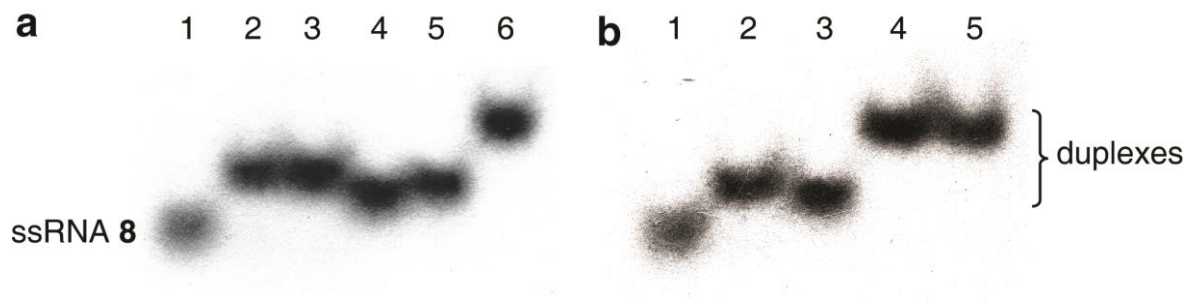


Figure 16. Gel mobility shift experiments to determine the hybridization efficiency of duplexes containing the emissive ribonucleoside 5 opposite to complementary bases, mismatched bases and abasic sites. (a) Lane 1, single stranded oligoribonucleotide 8. Lanes 2–6, duplexes 8•9, 8•11, 8•10, and 8•12, 8•13, respectively. (b) Lane 1, oligoribonucleotide 8. Lanes 2–5, duplexes 8•9, 8•14, 8•13, and 8•15, respectively.

2A.3 Conclusions

Preferably, incorporating the modified ribonucleoside reporter near the interaction site and closely maintaining the structural and functional integrity of the target oligoribonucleotide will provide a better understanding of the interaction under investigation. In this regard, microenvironment-sensitive emissive pyrimidine ribonucleoside analogue **5**, which is minimally perturbing and displaying emission in the visible region with a reasonable quantum yield, has the potential to be implemented in fluorescence-based assays to investigate nucleic acid structure, dynamics and recognition processes.

2A.4 Experimental Section

2A.4.1 Materials

Unless otherwise mentioned, materials obtained from commercial suppliers were used without any further purification. 5-iodouridine, *N*-methylindole, benzofuran, tributyltin chloride and *bis*(triphenylphosphine)-palladium(II) chloride were obtained from Sigma-Aldrich. $POCl_3$ was purchased from Acros Organics, and was distilled before use. Custom synthesized oligodeoxyribonucleotides were either purchased from Integrated DNA Technologies, Inc. or from Sigma-Aldrich. ONs were purified by polyacrylamide gel electrophoresis (PAGE) under denaturing condition, and desalted on Sep-Pak Classic C18 cartridges (Waters Corporation). Custom synthesized RNA ONs purchased from Dharmacon RNAi Technologies were deprotected according to the supplier's protocol, PAGE-purified, and desalted on Sep-Pak Classic C18 cartridges. T7 RNA polymerase, ribonuclease inhibitor

(RiboLock), RNase A, RNase T1 and NTPs were obtained from Fermentas Life Science. Calf intestinal alkaline phosphatase was procured from Invitrogen. Snake venom phosphodiesterase I was purchased from Sigma-Aldrich. Radiolabeled α-^{32}P ATP (2000 Ci/mmol) was purchased from the Board of Radiation and Isotope Technology, Government of India. Chemicals for preparing buffer solutions were purchased from Sigma-Aldrich (BioUltra grade). Autoclaved water was used in all biochemical reactions and fluorescence measurements.

2A.4.2 Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400. All MALDI-MS measurements were recorded on an Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer. Absorption spectra were recorded on a PerkinElmer, Lambda 45 UV-Vis spectrophotometer. UV-thermal melting studies on ONs were performed on a Cary 300Bio UV-Vis spectrophotometer. Steady State and time-resolved fluorescence experiments were carried out in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3). Reversed-phase flash chromatographic (C18 Redi*SepRf* column) purifications were carried out using Teledyne ISCO, Combi Flash *Rf*. HPLC analyses were performed using Dionex ICS 3000.

2A.4.3 Synthesis of ribonucleoside analogs

Synthesis of *N*-Boc-indole-conjugated uridine 2: To a suspension of 5-iodouridine (0.500 g, 1.35 mmol, 1 equiv) and *bis*(triphenylphosphine)-palladium(II) chloride (0.048 g, 0.07 mmol, 0.05 equiv) in degassed anhydrous DMF (7 ml) was added (*N*-Boc-indol-2-yl)tributylstannane[39] (1.024 g, 2.02 mmol, 1.5 equiv). The reaction mixture was heated at 75°C for 3 hr, cooled, and filtered through celite pad. Celite pad was washed with dioxane (3 x 10 ml). The solvent was evaporated, and the residue was purified by flash chromatography using a silica gel column (Silica Redi*SepRf* 120 g column, 0–25% methanol in chloroform for 25 min) to afford the product **2** as white foam (0.340 g, 55%). R_f=0.72 ($CHCl_3$/MeOH 8:2); ^{1}H NMR (400 MHz, [d_6]DMSO): δ=11.71 (br s, 1H), 8.24 (s, 1H), 8.07 (d, *J*=8 Hz, 1H), 7.59 (d, *J*=7.6 Hz, 1H), 7.32 (t, *J*=7.8 Hz, 1H), 7.24 (t, *J*=7.4 Hz, 1H), 6.64 (s, 1H), 5.85 (d, *J*=4.4 Hz, 1H), 5.46 (d, *J*=5.6 Hz, 1H), 5.15 (br, 2H), 4.09 (d, *J*=4.8 Hz, 1H), 4.01 (d, *J*=4.4 Hz, 1 H), 3.88 (br, 1 H), 3.66–3.53 (m, 2H), 1.47 (s, 9H) ppm; ^{13}C NMR (100 MHz, [d_6]DMSO): δ=162.3, 150.5, 149.3, 137.3, 136.4, 132.1, 128.3, 124.5, 122.8, 120.7, 114.6,

110.0, 109.7, 88.2, 84.8, 83.6, 74.1, 69.6, 60.4, 27.4 ppm; MALDI-TOF MS: *m/z* calcd. for $C_{22}H_{25}N_3O_8$+K^+: 498.1 [M+K^+]; found: 498.1.

Synthesis of indole-conjugated uridine 3: A solution of **2** (118 mg, 0.26 mmol) in HCl (3M in ethyl acetate, 3.5 ml) was stirred at RT for 4 hr. Solvent was evaporated, and the residue was dissolved in methanol. The pH was adjusted to ~7 by adding saturated sodium bicarbonate. Solvents were evaporated to dryness, and the residue was purified by flash chromatography using a silica gel column (Silica Redi*SepRf* 40 g column, 0–20% methanol in chloroform for 20 min) to afford the product **3** as a pale yellow powder (0.068 g, 74% yield). R_f=0.53 ($CHCl_3$/MeOH 8:2); ^{1}H NMR (400 MHz, [d_6]DMSO): δ=11.74 (br s, 1H), 11.08 (br s, 1H), 8.51 (s, 1H), 7.47 (d, *J*=2.8 Hz, 1H), 7.45 (d, *J*=3.2 Hz, 1H), 7.07–7.03 (m, 1H), 6.99–6.95 (m, 1H), 6.80 (d, *J*=1.6 Hz, 1H), 5.86 (d, *J*=4.8 Hz, 1H), 5.50 (d, *J*=5.2 Hz, 1H), 5.36 (t, *J*=5.0 Hz, 1H), 5.15 (d, *J*=5.2 Hz, 1H), 4.21 (dd, *J*=10.4, *J*=5.2 Hz, 1 H), 4.08 (dd, *J*=10.0, *J*=5.2 Hz, 1 H), 3.90 (dd, *J*=7.2, *J*=2.8 Hz, 1 H), 3.78–3.73 (m, 1H), 3.66–3.61 ppm (m, 1H); ^{13}C NMR (100 MHz, [d_6]DMSO): δ=161.8, 149.7, 136.9, 136.1, 131.1, 127.8, 121.1, 119.6, 119.2, 111.5, 106.4, 98.9, 88.5, 84.7, 73.7, 69.4, 60.5 ppm; HRMS: *m/z* calcd. for $C_{17}H_{17}N_3O_6$+Na^+: 382.1015 [M+Na^+]; found: 382.1017; λ_{max}(H_2O)=290 and 330 nm, ε_{290}=12093 $M^{-1}cm^{-1}$, ε_{330}=9826 $M^{-1}cm^{-1}$.

Synthesis of *N*-Methylindole-conjugated uridine 4: To a suspension of 5-iodouridine (0.252 g, 0.68 mmol, 1 equiv) and *bis*(triphenylphosphine)-palladium(II) chloride (0.025 g, 0.036 mmol, 0.05 equiv) in degassed anhydrous dioxane (6.2 ml) was added (*N*-methylindol-2-yl)tributylstannane[39] (0.567 g, 1.35 mmol, 2 equiv). The reaction mixture was heated at 90°C for 2 hr, cooled to RT, and filtered through celite pad. Celite pad was washed three times with dioxane (10 ml). The solvent was evaporated, and the residue was purified by flash chromatography using a silica gel column (Silica Redi*SepRf* 120 g column, 0–20% methanol in chloroform for 25 min) to afford the product **4** as a pale yellow solid (0.150 g, 60%). R_f=0.63 ($CHCl_3$/MeOH 8:2); ^{1}H NMR (400 MHz, [d_6]DMSO): δ=11.66 (br s, 1H), 8.22 (s, 1H), 7.53 (d, *J*=7.6 Hz, 1H), 7.43 (d, *J*=8.0 Hz, 1H), 7.18–7.14 (m, 1H), 7.05–7.02 (m, 1H), 6.42 (s, 1H), 5.86 (d, *J*=4.8 Hz, 1H), 5.47 (d, *J*=5.6 Hz, 1H), 5.12–5.09 (m, 2H), 4.14 (dd, *J*=10.4, *J*=5.2 Hz, 1H), 4.01 (dd, *J*=9.6, *J*=4.8 Hz, 1 H), 3.87 (dd, *J*=7.0, *J*=2.8 Hz, 1H), 3.64–3.61 (m, 1H), 3.60 (s, 3H), 3.55–3.50 ppm (m, 1H); ^{13}C NMR (100 MHz, [d_6]DMSO): δ=162.0, 150.5, 140.8, 137.5, 133.3, 127.0, 121.4, 120.1, 119.2, 109.8, 106.6,

102.3, 88.2, 84.8, 74.0, 69.7, 60.4, 30.8 ppm; HRMS: *m/z* Calcd. for $C_{18}H_{19}N_3O_6$+Na^+: 396.1172 [M+Na^+]; found: 396.1175; $\lambda_{max}(H_2O)$=273 and 286 nm, ε_{273}=11093 $M^{-1}cm^{-1}$, ε_{286}=10513 $M^{-1}cm^{-1}$.

Synthesis of benzofuran-conjugated uridine 5: To a suspension of 5-iodouridine (0.250 g, 0.68 mmol, 1 equiv) and *bis*(triphenylphosphine)-palladium(II) chloride (0.024 g, 0.034 mmol, 0.05 equiv) in degassed anhydrous dioxane (7.5 ml) was added 2-(tri-n-butylstannyl)benzofuran[40] (0.412 g, 1.01 mmol, 1.5 equiv). The reaction mixture was heated at 90°C for 2 hr, cooled to RT, and filtered through celite pad. The celite pad was washed with dioxane (3 x 10 ml), and the solvent was evaporated. The solid residue was dissolved in a minimum amount of hot 1:1 dioxane: methanol solution, and precipitated with hexane. The precipitate was allowed to stand at ~5°C for 2 hr, after which the precipitate was filtered and dried to afford the product **5** as off-white solid (0.210 g, 86%). R_f=0.51 ($CHCl_3$/MeOH 8:2); ^{1}H NMR (400 MHz, [d_6]DMSO): δ=11.79 (br s, 1H), 8.87 (s, 1H), 7.62 (d, *J*=7.6 Hz, 1H), 7.55 (d, *J*=8.4 Hz, 1H), 7.34 (s, 1H), 7.30–7.26 (m, 1H), 7.23–7.20 (m, 1H), 5.87 (d, *J*=4.4 Hz, 1H), 5.50 (d, *J*=5.2 Hz, 1H), 5.39 (t, *J*=4.0 Hz, 1H), 5.14 (d, *J*=5.6 Hz, 1H), 4.15 (dd, *J*=9.6, *J*=4.8 Hz, 1H), 4.09 (dd, *J*=10.2, *J*=5.0 Hz, 1 H), 3.96–3.95 (m, 1H), 3.82–3.78 (m, 1H), 3.69–3.65 (m, 1H) ppm; ^{13}C NMR (100 MHz, [d_6]DMSO): δ=160.3, 153.0, 149.6, 149.2, 137.3, 128.8, 124.3, 123.0, 121.0, 110.8, 104.7, 103.9, 88.8, 84.6, 74.4, 69.4, 60.1 ppm; HRMS: *m/z* Calcd. for $C_{17}H_{16}N_2O_7$+Na^+: 383.0855 [M+Na^+]; found: 383.0855; $\lambda_{max}(H_2O)$=272 and 322 nm, ε_{272}=19106 $M^{-1}cm^{-1}$, ε_{322}=22240 $M^{-1}cm^{-1}$, ε_{260}=17253 $M^{-1}cm^{-1}$.

Synthesis of benzofuran-conjugated uridine triphosphate 6: To an ice cold solution of ribonucleoside **5** (0.090 g, 0.25 mmol, 1 equiv) in trimethyl phosphate (1.2 ml) was added freshly distilled $POCl_3$ (60 μL, 0.64 mmol, 2.6 equiv). The solution was stirred for 42 h at ~4°C. TLC revealed only partial conversion of the ribonucleoside into the product. A solution of *bis*-tributylammonium pyrophosphate[41] (0.5 M in DMF, 2.6 ml, 5.2 equiv) and tributylamine (0.65 ml, 2.72 mmol, 11 equiv) was rapidly added under ice-cold condition. The reaction was quenched after 30 min with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, 15 ml), and was extracted with ethyl acetate (20 ml). The aqueous layer was evaporated and the residue was purified first on DEAE sephadex-A25 anion exchange column (10 mM–1M TEAB buffer, pH 7.5) followed by reversed-phase flash column chromatography (C18 Redi*SepRf*, 0–40% acetonitrile in 100 mM triethylammonium acetate

buffer, pH 7.2, 40 min). Appropriate fractions were lyophilized to afford the desired triphosphate product **6** as a tetratriethylammonium salt (51 mg, 20%). ^{1}H NMR (400 MHz, D_2O): δ=7.93 (s, 1H), 7.44 (d, *J*=8.0 Hz, 1H), 7.37 (d, *J*=7.6 Hz, 1H), 7.14 (t, *J*=7.6 Hz, 1H), 7.05 (t, *J*=7.4 Hz, 1H), 6.96 (s, 1H), 5.69 (d, *J*=4.8 Hz, 1H), 4.31–4.26 (m, 2H), 4.13 (br, 3H) ppm; ^{13}C NMR (100 MHz, D_2O): δ=161.4, 153.4, 150.3, 147.7, 136.7, 128.6, 125.0, 123.1, 121.1, 111.5, 106.4, 105.1, 89.1, 83.4 (d, *J*=7.1 Hz), 73.4, 69.9, 65.5 ppm; ^{31}P NMR (162 MHz, D_2O): δ=-5.85 (d, *J*=20.7 Hz, P_γ), -10.92 (d, *J*=20.6 Hz, P_α), -21.85 (t, *J*=19.7 Hz, P_β); MALDI-TOF MS (negative mode): *m/z* calcd. for $C_{17}H_{19}N_2O_{16}P_3$: 600.0 [M]; found: 599.1 $[M-H]^-$.

2A.4.4 Photophysical characterization of ribonucleoside analogs:

Steady-state fluorescence of ribonucleoside analogs in various solvents

Ribonucleoside analogs in water, methanol, acetonitrile and dioxane were excited at respective lowest energy absorption maximum with an appropriate excitation and emission slit widths, and emission profile in each solvent was recorded. Fluorescence experiments were performed in triplicate in a micro fluorescence cell (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3.

Time-resolved fluorescence measurements

Excited state lifetimes of benzofuran-conjugate uridine **5** in various solvents were determined using TCSPC fluorescence spectrophotometer (Horiba Jobin Yvon). While the concentration of ribonucleoside in water and methanol was 5 μM, concentration in acetonitrile and dioxane was 250 μM. Ribonucleoside **5** in water and methanol was excited using 339 nm LED source (IBH, UK, NanoLED-339L) with a band pass of 4 nm, and fluorescence signal at respective emission maximum was collected. Similarly, ribonucleoside in acetonitrile and dioxane was excited using 375 nm diode laser source (IBH, UK, NanoLED-375L) with a band pass of 12 nm. Lifetime measurements were performed in duplicate, and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay kinetics in various solvents was found to be either mono or biexponential with χ^2 (goodness of fit) values very close to unity.

2A.4.5 Quantum yield determination in various solvents

Quantum yield of ribonucleoside **5** in different solvents relative to 2-aminopurine standard was determined using the following equation.[42]

$$\Phi_{F(x)} = (A_s/A_x)\,(F_x/F_s)\,(n_x/n_s)^2\,\Phi_{F(s)}$$

Where s is the standard, x is the ribonucleoside, A is the absorbance at excitation wavelength, F is the area under the emission curve, n is the refractive index of the solvent, and Φ_F is the quantum yield.

2A.4.6 Enzymatic incorporation of ribonucleoside triphosphate 6

Transcription reactions with α-^{32}P ATP: Promoter-template duplexes were assembled by heating a 1:1 mixture (final 5 μM) of deoxyoligonucleotide templates (**D1**–**D5**) and an 18-mer T7 RNA polymerase consensus promoter deoxyoligonucleotide sequence in TE buffer (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.8) at 90°C for 3 min and cooling the solution slowly to room temperature. The duplexes were then placed on crushed ice for 20 min and stored at -40°C. Transcription reactions were performed in 40 mM Tris-HCl buffer (pH 7.9) containing 250 nM of annealed templates, 10 mM $MgCl_2$, 10 mM NaCl, 10 mM dithiothreitol (DTT), 2 mM spermidine, 1 U/μL RNase inhibitor (RiboLock), 1 mM GTP, 1 mM CTP, 1 mM UTP and or 1 mM modified UTP **6**, 20 μM ATP, 5 μCi α-^{32}P ATP and 3 U/μL T7 RNA polymerase in a total volume of 20 μL. After 3.5 h at 37°C, reactions were quenched by adding 20 μL of the loading buffer (7 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8). The samples were heated at 75°C for 3 min and cooled on an ice bath. The samples (4 μL) were loaded on an analytical denaturing polyacrylamide gel (18%) containing 7 M urea and run at a constant power (11 W) for nearly 4 h. The gel was exposed to X-ray film (1–2 h), and the exposed film was developed, fixed and dried. The bands corresponding to full-length transcripts were then quantified using a software (GeneTools from Syngene) to determine the transcription yield. The percentage incorporation of modified ribonucleoside triphosphate **6** has been reported with respect to transcription efficiency in the presence of natural NTPs. All reactions were performed in duplicate, and the errors in yields were found to be ≤ 4%.

Large-scale transcription reactions

Large-scale transcription reactions using template **D1** were performed in a 250 μL reaction volume under similar conditions to isolate RNA ONs **7** and **8** for further characterization and

fluorescence studies. The reaction contained 2 mM GTP, 2 mM CTP, 2 mM ATP, 2 mM UTP or 2 mM modified UTP **6**, 20 mM $MgCl_2$, 0.4 U/μL RNase inhibitor (RiboLock), 300 nM annealed template and 800 units T7 RNA polymerase. After incubating the reaction for 12 h at 37°C, the precipitated pyrophosphate was removed by centrifugation. The reaction volume was reduced approximately to 1/3 by Speed Vac, and 50 μL of loading buffer was added. The solution was heated at 75°C for 3 min and cooled on the ice bath. The sample was loaded onto a preparative 20% denaturing polyacrylamide gel (20 cm width, 45 cm height and 2 mm thickness) and was run at a constant power (25 W) for nearly 4.5 h. The gel was UV shadowed, and appropriate band was excised, extracted with 0.3 M sodium acetate and desalted using Sep-Pak classic C18 cartridge. Typical yields of transcripts were 14–15 nmol. Control transcript **7**, ε_{260}=9.10 x 10^4 $M^{-1}cm^{-1}$; modified transcript **8**, ε_{260}=9.86 x 10^4 $M^{-1}cm^{-1}$.

2A.4.7 Enzymatic digestion of transcript 8

Nearly 4 nmol of the modified oligoribonucleotide **8** from a large-scale transcription reaction was digested with snake venom phosphodiesterase I (0.01 U), calf intestine alkaline phosphatase (1 U/μL), and RNase A (0.25 μg) in a total volume of 100 μL in 50 mM Tris-HCl buffer (pH 8.5, 40 mM $MgCl_2$, 0.1 mM EDTA) for ~12 h at 37°C. After this period, RNase T1 (0.2 U/μL) was added, and the sample was incubated for another 4 h at 37°C. The ribonucleoside mixture obtained was analyzed by reversed-phase analytical HPLC using Phenomenex-Luna C18 column (250 x 4.6 mm, 5 micron) at 260 nm and 320 nm. Mobile phase A: 50 mM triethylammonium acetate buffer (pH 7.5), mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min.

2A.4.8 Photophysical characterization of fluorescently modified RNA ONs

Steady-state fluorescence: Oligoribonucleotide **8** (10 μM) was annealed to custom DNA and RNA ONs (**9**−**15**, 11 μM) by heating a 1:1.1 mixture of the ONs in 20 mM cacodylate buffer (pH 7.1, 500 mM NaCl, 0.5 mM EDTA) at 90°C for 3 min. Samples were then cooled slowly to RT, and were placed in crushed ice. Samples were diluted to give a final concentration of 1 μM (with respect to **8**) in cacodylate buffer. Fluorescently modified duplexes were excited at 330 nm (unless otherwise mentioned) with an excitation slit width of 7 nm and emission slit width of 9 nm. Fluorescence experiments were performed in triplicate in a micro

fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 at 18 ± 1°C.

Time-resolved fluorescence: Lifetimes of the duplexes (2 μM) were determined under the same conditions as mentioned above using TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3). Fluorescently modified RNA ONs were excited using 339 nm LED source (IBH, UK, NanoLED-339L), and fluorescence signal at respective emission maximum was collected. Experiments were performed in duplicate, and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay profiles for all ON constructs were found to be biexponential with χ^2 (goodness of fit) values very close to unity.

Quantum yield determination

Quantum yield of fluorescently modified oligoribonucleotide constructs was determined relative to the quantum yield of the ribonucleoside **5** in 20 mM cacodylate buffer (pH 7.1, 500 mM NaCl, 0.5 mM EDTA) using the equation given in section 2A.4.5.

2A.5 Refrences

[1] M. J. Rist, J. P. Marino, *Curr. Org. Chem.* **2002,** *6*, 775−793.

[2] A. Okamoto, Y. Saito, I. Saito, *J. Photochem. Photobiol. C: Photochem. Rev.* **2005,** *6*, 108−122.

[3] G. M. Wilson, *Rev. Fluoresc*. **2005,** 223−243.

[4] U. Asseline, *Curr. Org. Chem.* **2006,** *10*, 491−518.

[5] X. Shi, D. Herschlag, *Methods Enzymol.* **2009,** *469*, 288−302.

[6] R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev.* **2010,** *110*, 2579−2619.

[7] S. G. Srivatsan, A. A. Sawant, *Pure Appl. Chem.* **2011,** *83*, 213−232.

[8] a) M. Daniels, W. Hauswirth, *Science* **1971,** *171*, 675−677; b) J. M. L. Pecourt, J. Peon, B. Kohler, *J. Am. Chem. Soc.* **2000,** *122*, 9348−9349; c) E. Nir, K. Kleinermanns, L. Grace, M. S. de Vries, *J. Phys. Chem.* **2001,** *105*, 5106−5110.

[9] J. N. Wilson, E. T. Kool, *Org. Biomol. Chem.* **2006,** *4*, 4265−4274.

[10] R. T. Ranasinghe, T. Brown, *Chem. Commun*. **2005,** 5487−5502.

[11] M. E. Hawkins, in Topics in Fluorescence Spectroscopy, DNA Technology, Vol. 7 (Eds: J. R. Lakowicz), Kluwer Academic/Plenum Publishers, New York, 2003, pp. 151−175.

[12] Selected examples of DNA modification: a) L. H. Thoresen, G. S. Jiao, W. C. Haaland, M. L. Metzker, K. Burgess, *Chem. Eur. J.* **2003,** *9,* 4603−4610; b) R. S. Coleman, M.

A. Mortensen, *Tet. Lett.* **2003,** *44,* 1215−1219; c) S. T. Gaballah, J. D. Vaught, B. E. Eaton, T. L. Netzel, *J. Phys. Chem. B* **2005,** *109,* 5927−5934; d) N. J. Greco, Y. Tor, *J. Am. Chem. Soc.* **2005,** *127,* 10784−10785; e) F. Seela, E. Schweinberger, K. Y. Xu, V. R. Sirivolu, H. Rosemeyer, E. M. Becker, *Tetrahedron* **2007,** *63,* 3471−3482; f) M. Mizuta, K. Seio, K. Miyata, M. Sekine, *J. Org. Chem.* **2007,** *72,* 5046−5055; g) K. Tainaka, K. Tanaka, S. Ikeda, K. I. Nishiza, T. Unzai, Y. Fujiwara, I. Saito, A. Okamoto, *J. Am. Chem. Soc.* **2007,** *129,* 4776−4784; h) P. Sandin, K. Borjesson, H. Li, J. Martensson, T. Brown, L. M. Wilhelmsson, B. Albinsson, *Nucleic Acids Res.* **2008,** *36,* 157−167; i) N. A. Grigorenko, C. J. Leumann, *Chem. Commun.* **2008,** 5417−5419; j) Y. N. Teo, J. N. Wilson, E. T. Kool, *J. Am. Chem. Soc.* **2009,** *131,* 3923−3933; k) A. Dumas, N. W. Luedtke, *J. Am. Chem. Soc.* **2010,** *132,* 18004−18007; l) W. Hirose, K. Sato, A. Matsuda, *Angew. Chem. Int. Ed.* **2010,** *49,* 8392−8394; *Angew. Chem.* **2010,** *122,* 8570-8572; m) A. Omumi, D. G. Beach, M. Baker, W. Gabryelski, R. A. Manderville, *J. Am. Chem. Soc.* **2011,** *133,* 42−50; n) T. Ehrenschwender, H. A. Wagenknecht, *J. Org. Chem.* **2011,** *76,* 2301−2304; o) A. Nadler, J. Strohmeier, U. Diederichsen, *Angew. Chem. Int. Ed.* **2011,** 50, 5392−5396; *Angew. Chem.* **2011,** *123,* 5504−5508.

[13] Selected examples of RNA modification: a) A. Arzumanov, F. Godde, S. Moreau, J. J. Toulme, A. Weeds, M. J. Gait, *Helv. Chim. Acta* **2000,** *83,* 1424−1436; b) S. Shandrick, Q. Zhao, Q. Han, B. K. Ayida, M. Takahashi, G. C. Winters, K. B. Simonsen, D. Vourloumis, T. Hermann, *Angew. Chem. Int. Ed.* **2004,** *43,* 3177−3182; *Angew. Chem.* **2004,** 116, 3239−3244; c) M. Kaul, C. M. Barbieri, D. S. Plich, *J. Am. Chem. Soc.* **2006,** *128,* 1261−1271; d) K. Lang, R. Rieder, R. Micura, *Nucleic Acids Res.* **2007,** *35,* 5370−5378; e) C. Zhao, J. P. Marino, Tetrahedron **2007,** *63,* 3575−3584; f) S. G. Srivatsan, Y. Tor, *J. Am. Chem. Soc.* **2007,** *129,* 2044−2053; g) B. Heppell, D. A. Lafontaine, *Biochemistry* **2008,** *47,* 1490−1499; h) C. M. Zhang, C. Liu, T. Christian, H. Gamper, J. Rozenski, D. Pan, J. B. Randolph, E. Wickstrom, B. S. Cooperman, Y. M Hou, *RNA* **2008,** *14,* 2245−2253; i) S. Roday, M. B. Sturm, D. Blakaj, V. L. Schramm, *J. Biochem. Biophys. Methods* **2008,** *70,* 945−953; j) H. S. Jeong, S. Kang, J. Y. Lee, B. H. Kim, *Org. Biomol. Chem.* **2009,** *7,* 921−925; k) H. Peacock, O. Maydanovych, P. A. Beal, *Org. Lett.* **2010,** *12,* 1044−1047; l) U. Forster, K. Lommel, D. Sauter, C. Grünewald, J. W. Engels, J. Wachtveitl, *ChemBioChem* **2010,** *11,* 664−672; m) Y. Hikida, M. Kimoto, S. Yokoyama, I. Hirao, *Nat. Protoc.* **2010,** *5,* 1312−1323.

[14] a) T. Xia, *Curr. Opin. Chem. Biol.* **2008,** *12,* 604−611; b) K. B. Hall, *Methods Enzymol.* **2009,** *469,* 269−285.

[15] D. C. Ward, E. Reich, L. Stryer, *J. Biol. Chem.* **1969,** *244,* 1228−1237.

[16] M. Kawai, M. J. Lee, K. O. Evans, T. M. Nordlund, *J. Fluorescence* **2001,** *11,* 23−32.

[17] E. L. Rachofsky, R. Osman, J. B. A. Ross, *Biochemistry* **2001,** *40,* 946−956.

[18] a) D. A. Berry, K.Y. Jung, D. S. Wise, A. D. Sercel, W. H. Pearson, H. Mackie, J. B. Randolph, R. L. Somers, *Tet. Lett.* **2004,** *45,* 2457−2461; b) R. A. Tinsley, N. J. Walter, *RNA* **2006,** *12,* 522−529; c) A. S. Wahba, A. Esmaeili, M. J. Damha, R. H. E. Hudson, *Nucleic Acids Res.* **2010,** *38,* 1048−1056.

[19] a) S. G. Srivatsan, N. J. Greco, Y. Tor, *Angew. Chem., Int. Ed.* **2008,** *47,* 6661−6665; *Angew. Chem.* **2008,** *120,* 6763−6767; b) Y. Xie, A. V. Dix, Y. Tor. *J. Am. Chem. Soc.* **2009,** *131,* 17605−17614; c) Y. Xie, T. Maxson, Y. Tor, *J. Am. Chem. Soc.* **2010**, *132,* 11896−11897.

[20] J. R. Lakowicz, *in Principles of Fluorescence Spectroscopy, 3rd Ed.* Springer, New York, **2006**.

[21] R. W. Sinkeldam, A. J. Wheat, H. Boyaci, Y. Tor, *ChemPhysChem* **2011,** *12,* 567−570.

[22] M. G. Pawar, S. G. Srivatsan, *Org. Lett.* **2011,** *13,* 1114−1117.

[23] C. Reichardt, *Chem. Rev.* **1994,** *94,* 2319−2358.

[24] J. Ludwig, *Acta. Biochim. et Biophys. Acad. Sci. Hung.* **1981,** *16,* 131−133.

[25] F. Wachowius, C. Höbartner, *ChemBioChem* **2010,** *11,* 469−480.

[26] Enzymatic ligation has been used to synthesize labeled RNAs. a) C. Höbartner, R. Rieder, C. Kreutz, B. Puffer, K. Lang, A. Polonskaia, A. Serganov, R. Micura, *J. Am. Chem. Soc.* **2005,** 127, 12035−12045; b) R. Rieder, K. Lang, D. Graber, R. Micura, *ChemBioChem* **2007,** 8, 896−902.

[27] a) F. Huang, G. Wang, T. Coleman, N. Li, *RNA* **2003,** *9,* 1562−1570; b) R. Fiammengo, K. Musilek, A. Jaschke, *J. Am. Chem. Soc.* **2005,** *127,* 9271−9276; c) M. Kimoto, T. Mitsui, Y. Harada, A. Sato, S. Yokoyama, I. Hirao, *Nucleic Acids Res.* **2007,** *35,* 5360−5369; d) S. G. Srivatsan, Y. Tor, *Chem. Asian J.* **2009,** *4,* 419−427; e) G. Stengel, M. Urban, B. W. Purse, R. D. Kuchta, *Anal. Chem.* **2010,** *82,* 1082−1089.

[28] J. F. Milligan, O. C. Uhlenbeck, *Methods Enzymol.* **1989,** *180,* 51−62.

[29] J. F. Milligan, D. R. Groebe, G. W. Witherell, O. C. Uhlenbeck, *Nucleic Acids Res.* **1987,** *15,* 8783−8798.

[30] a) M. Torimura, S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa, R. Kurane, *Anal. Sci.* **2001,** *17,* 155−160; b) J. M. Jean, K. B. Hall, *Proc. Natl. Acad. Sci. USA* **2001,** *98,* 37−41.

[31] a) S. Doose, H. Neuweiler, M. Sauer, *ChemPhysChem* **2009,** *10,* 1389−1398; b) S. O. Kelly, J. K. Barton, *Science* **1999,** *283,* 375−381.

[32] T. M. Nordlund, S. Andersson, L. Nilsson, R. Rigler, A. Gräslund, L. W. McLaughlin, *Biochemistry* **1989,** *28,* 9095−9103.

[33] C. R. Guest, R. A. Hochstrasser, L. C. Sowers, D. P. Millar, *Biochemistry* **1991,** *30,* 3271−3279.

[34] a) M. E. Hawkins, F. M. Balis, *Nucleic Acids Res.* **2004,** *32,* e62; b) Y. J. Seo, S. Bhuniya, S. Tapadar, J. W. Yi, B. H. Kim, Bull. Korean *Chem. Soc.* **2007,** *28,* 1923−1924; c) S. Ikeda, T. Kubota, M. Yuki, H. Yanagisawa, S. Tsuruma, A. Okamoto, *Org. Biomol. Chem.* **2010,** *8,* 546−551.

[35] a) T. Lindahl, *Nature* **1993,** *362,* 709−715; b) C. D. Mol, S. S. Parikh, C. D. Putnam, T. P. Lo, J. A. Tainer, *Annu. Rev. Biophys. Biomol. Struct.* **1999,** *28,* 101−128.

[36] a) L. A. Loeb, B. D. Preston, *Annu. Rev. Genet.* **1986,** *20,* 201−230; b) W. Zhou, P. W. Doetsch, *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 6601−6605.

[37] a) H. Ide, K. Akamatsu, Y. Kimura, K. Michiue, K. Makino, A. Asaeda, Y. Takamori, K. Kubo, *Biochemistry* **1993,** *32,* 8276−8283; b) J. Lhomme, J. F. Constant, M. Demeunynck, *Biopolymers* **1999,** *52,* 65−83; c) H. Atamna, I. Cheung, B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **2000,** *97,* 686−691.

[38] a) E. L. Rachofsky, E. Seibert, J. T. Stivers, R. Osman, J. B. A. Ross, *Biochemistry* **2001,** 40, 957−967; b) T. Hirose, T. Ohtani, H. Muramatsu, A. Tanaka, *Photochem. Photobiol.* **2002,** *76,* 123−126; c) K. Sato, M. M. Greenberg, *J. Am. Chem. Soc.* **2005,** *127,* 2806−2807; d) L. Valis, N. Amann, H. A. Wagenknecht, *Org. Biomol. Chem.* **2005,** 3, 36−38; e) E. Fundador, J. Rusling, *Anal. Bioanal. Chem.* **2007,** *387,* 1883−1890; f) Z. Xu, Y. Sato, S. Nishizawa, N. Teramae, *Chem. Eur. J.* **2009,** *15,* 10375−10378.

[39] S. S. Labadie, E. Teng, J. *Org. Chem.* **1994,** *59,* 4250−4254.

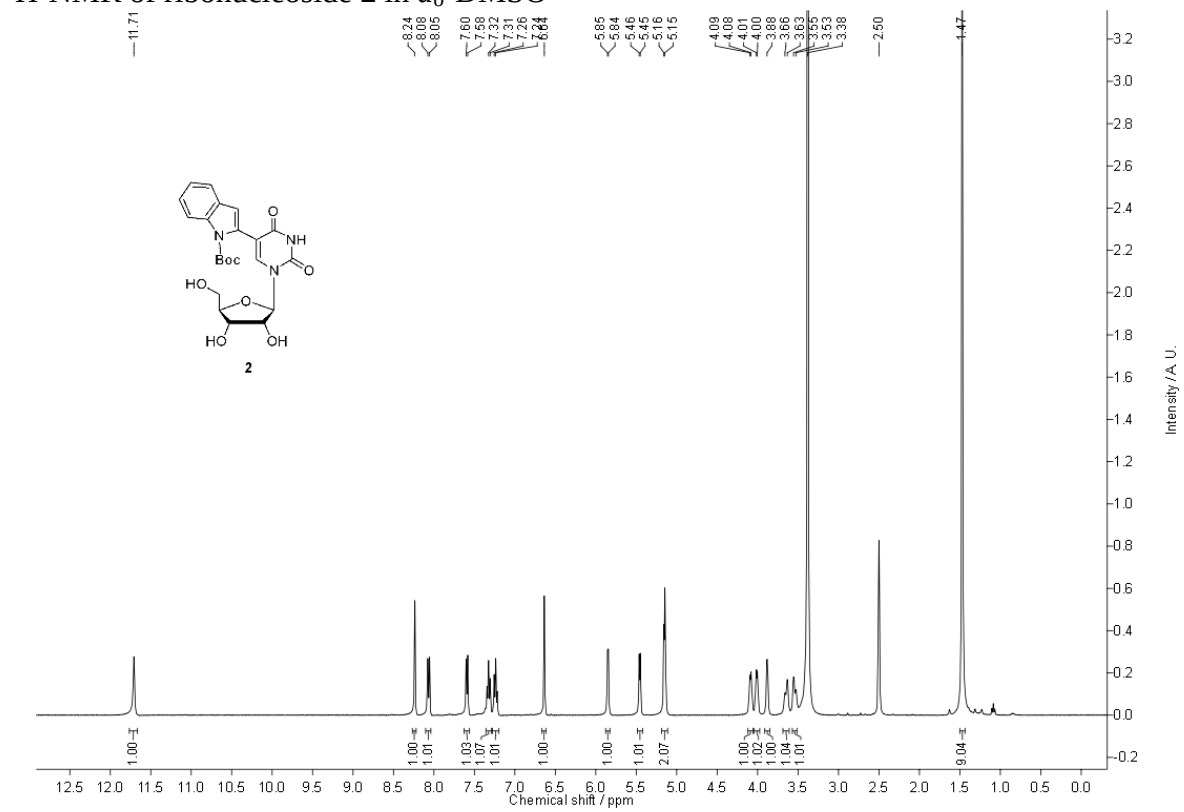
[40] L. S. Liebeskind, J. Wang, *J. Org. Chem.* **1993**, *58,* 3550−3556.

[41] J. G. Moffatt, *Can. J. Chem.* **1964**, *42,* 599−604.

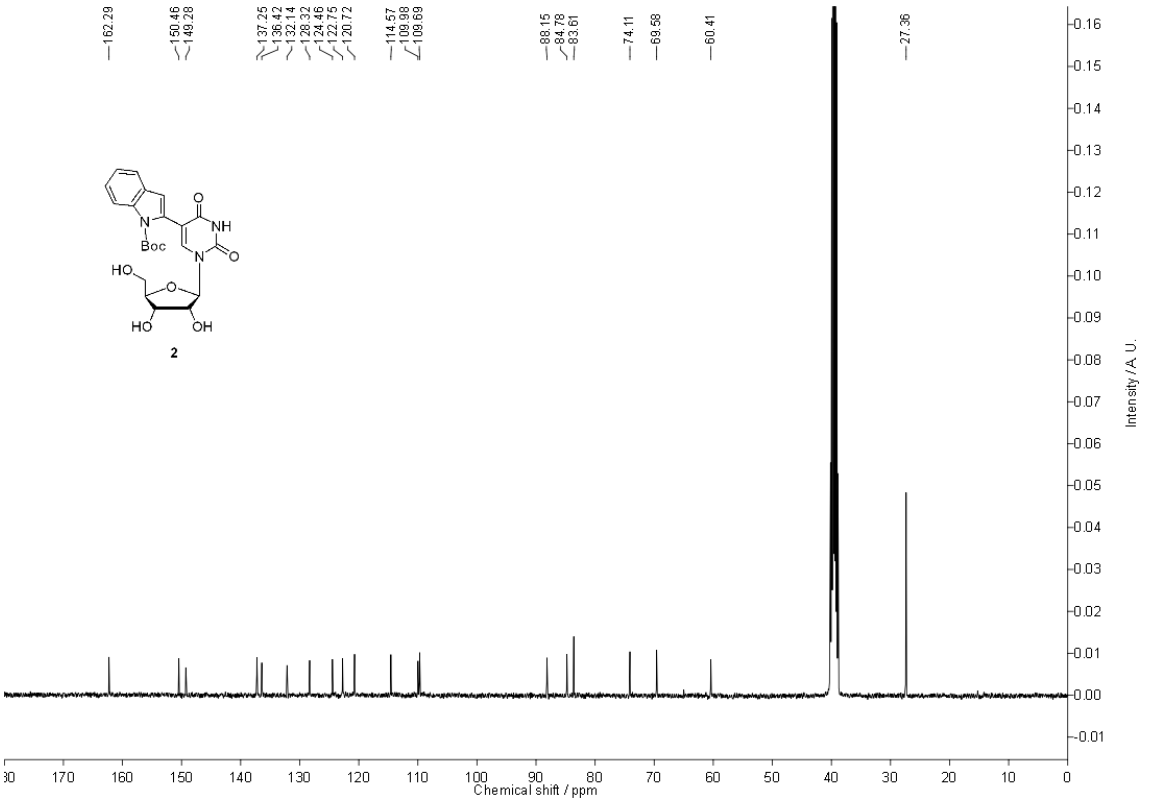
[42] D. Lavabre, S. Fery-Forgues, *J. Chem. Educ.* **1999**, *76,* 1260−1264.

2A.6 Appendix-I : Characterization data of synthesized compounds

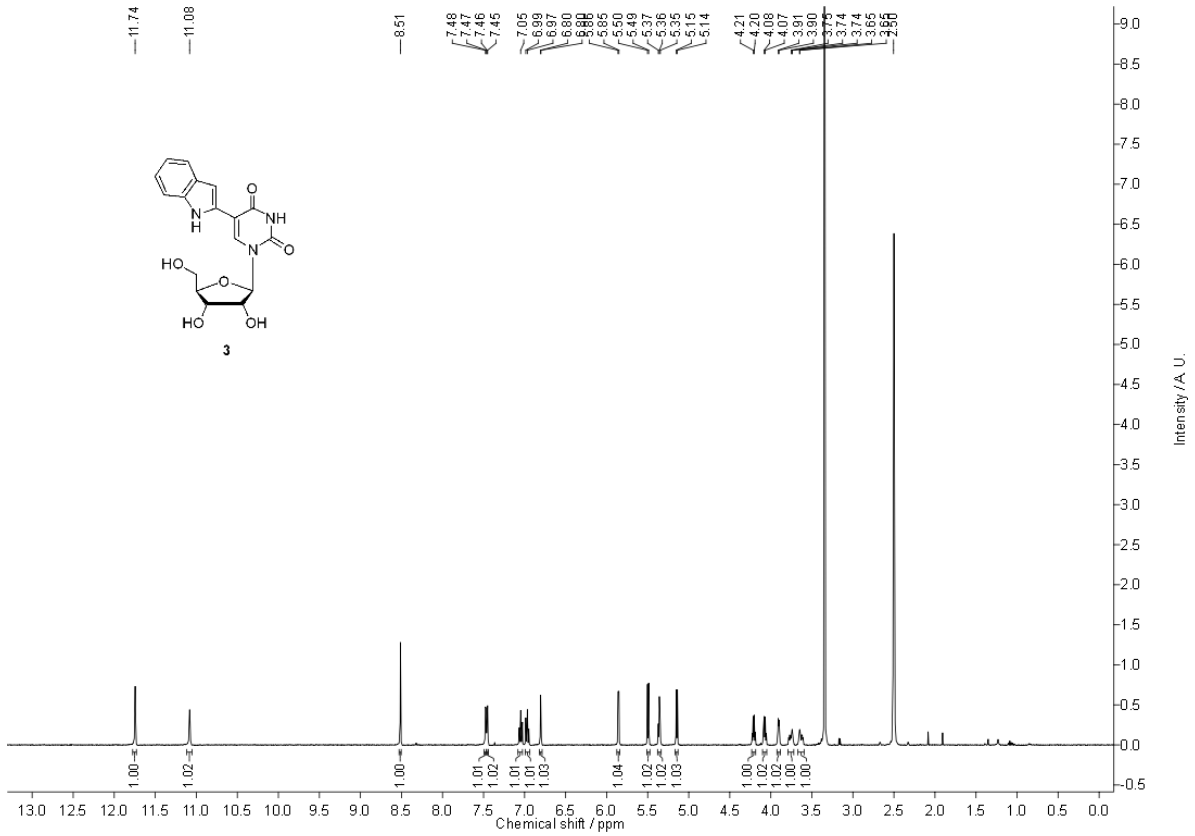
^{1}H-NMR of ribonucleoside **2** in d_6-DMSO



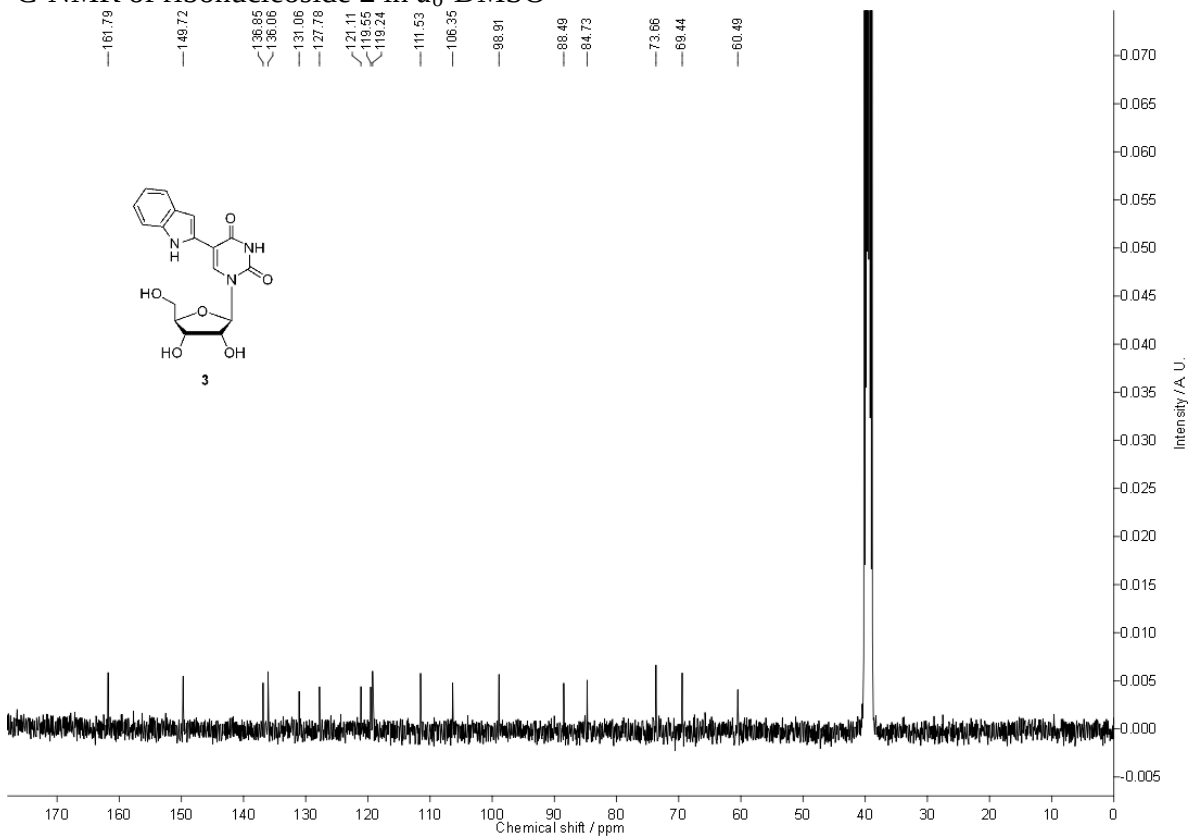
^{13}C-NMR of ribonucleoside **2** in d_6-DMSO



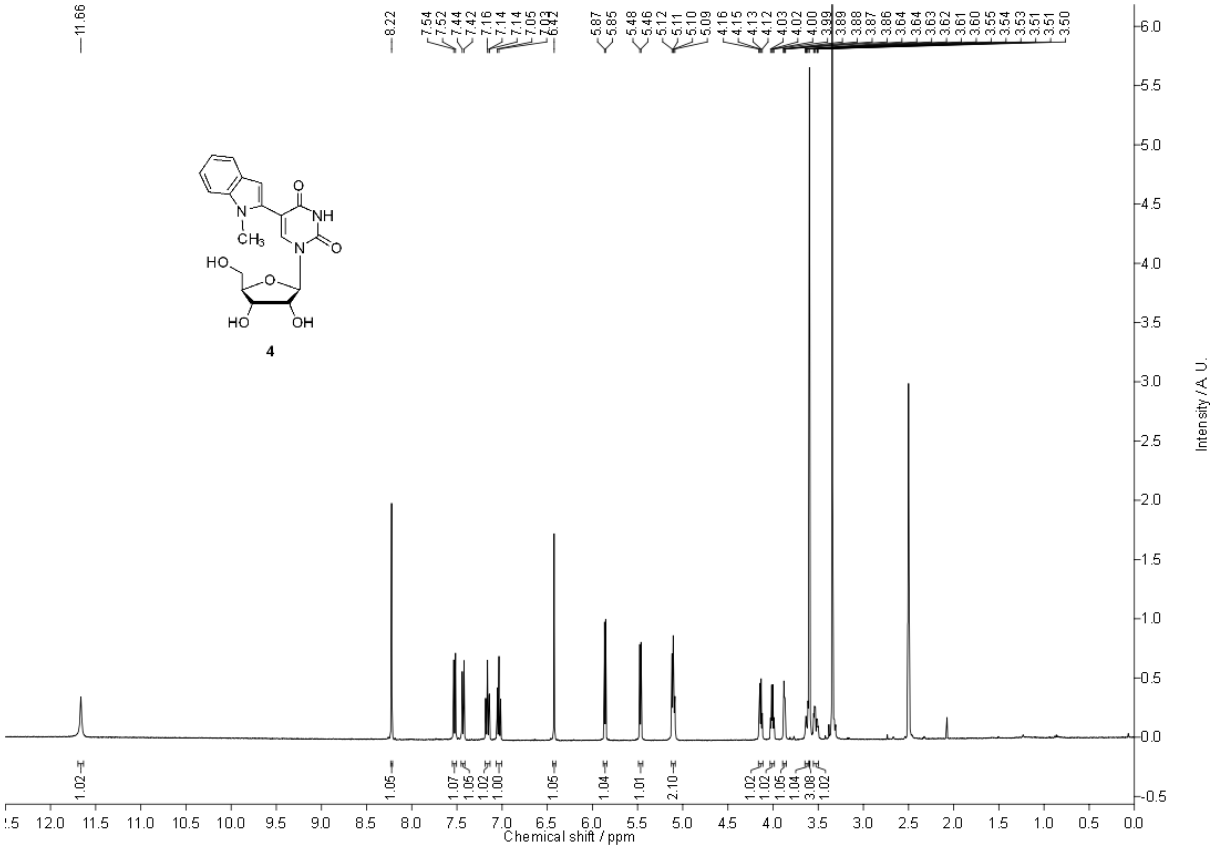
^{1}H-NMR of ribonucleoside **3** in d_6-DMSO



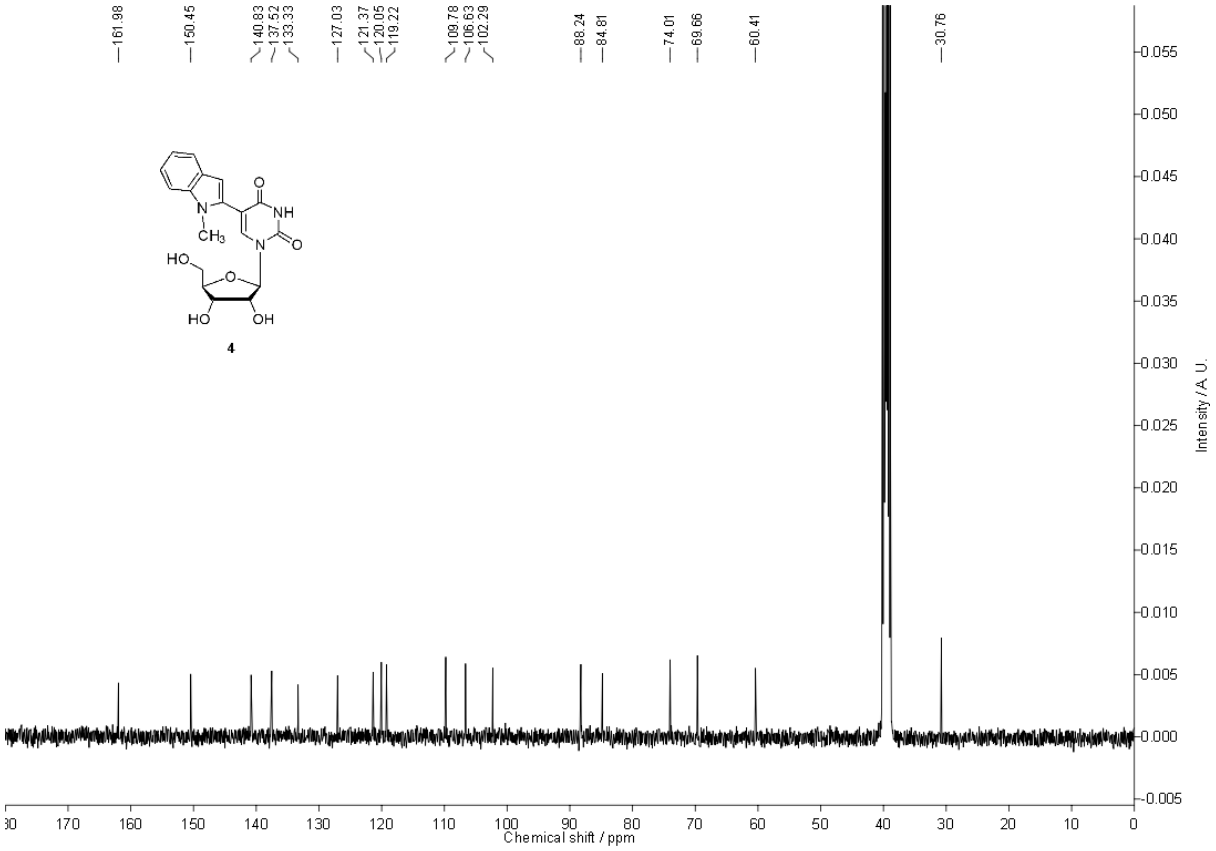
^{13}C-NMR of ribonucleoside **2** in d_6-DMSO



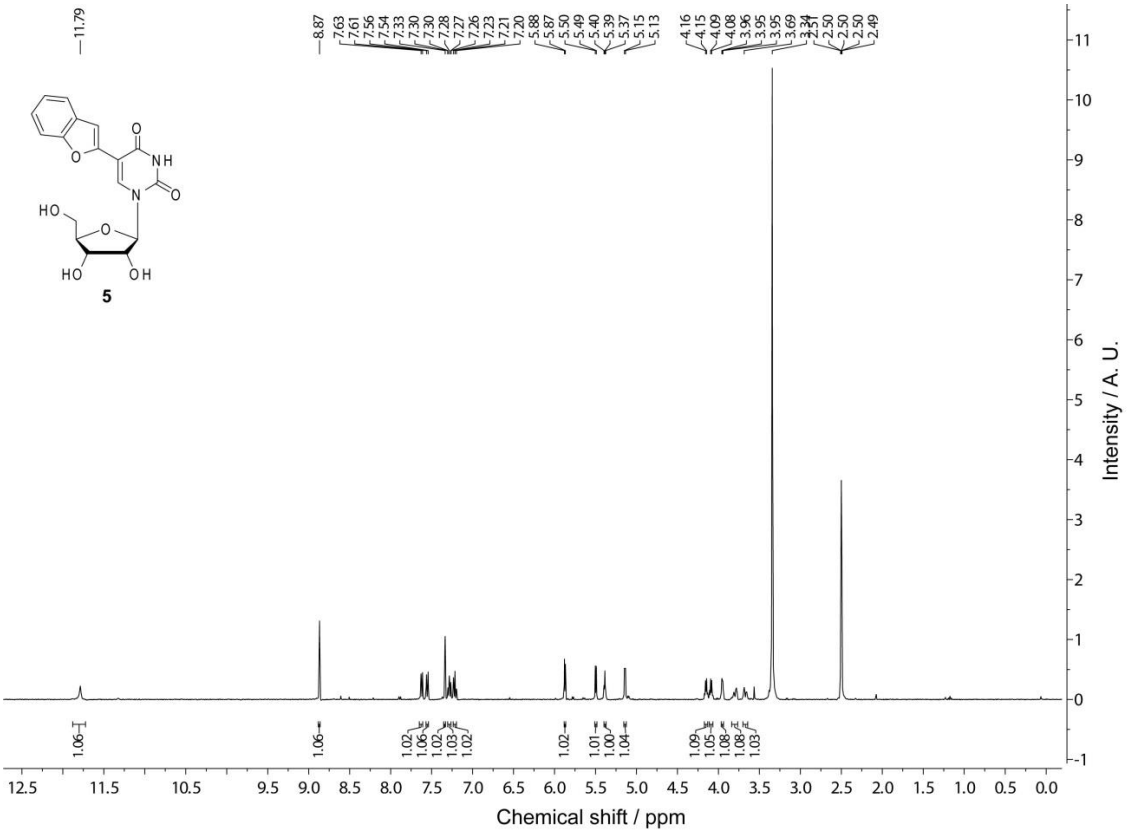
^{1}H-NMR of ribonucleoside **4** in d_6-DMSO



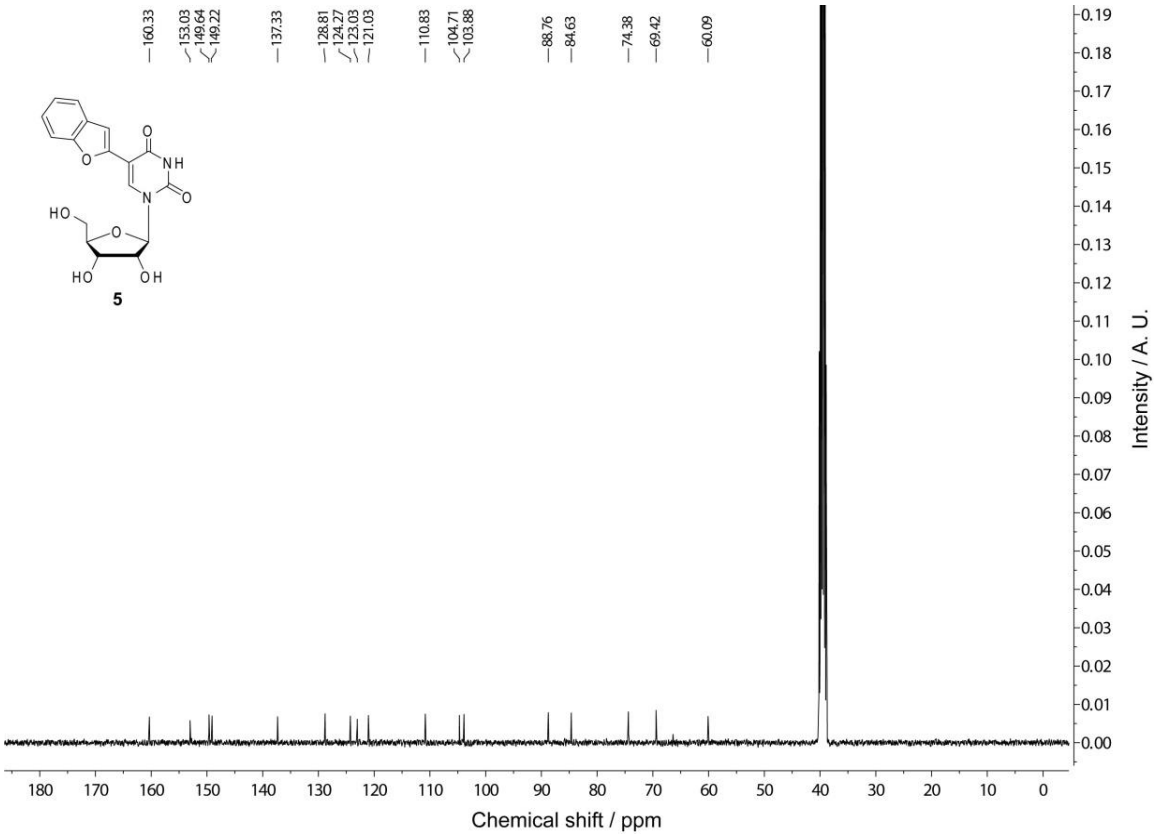
^{13}C-NMR of ribonucleoside **4** in d_6-DMSO



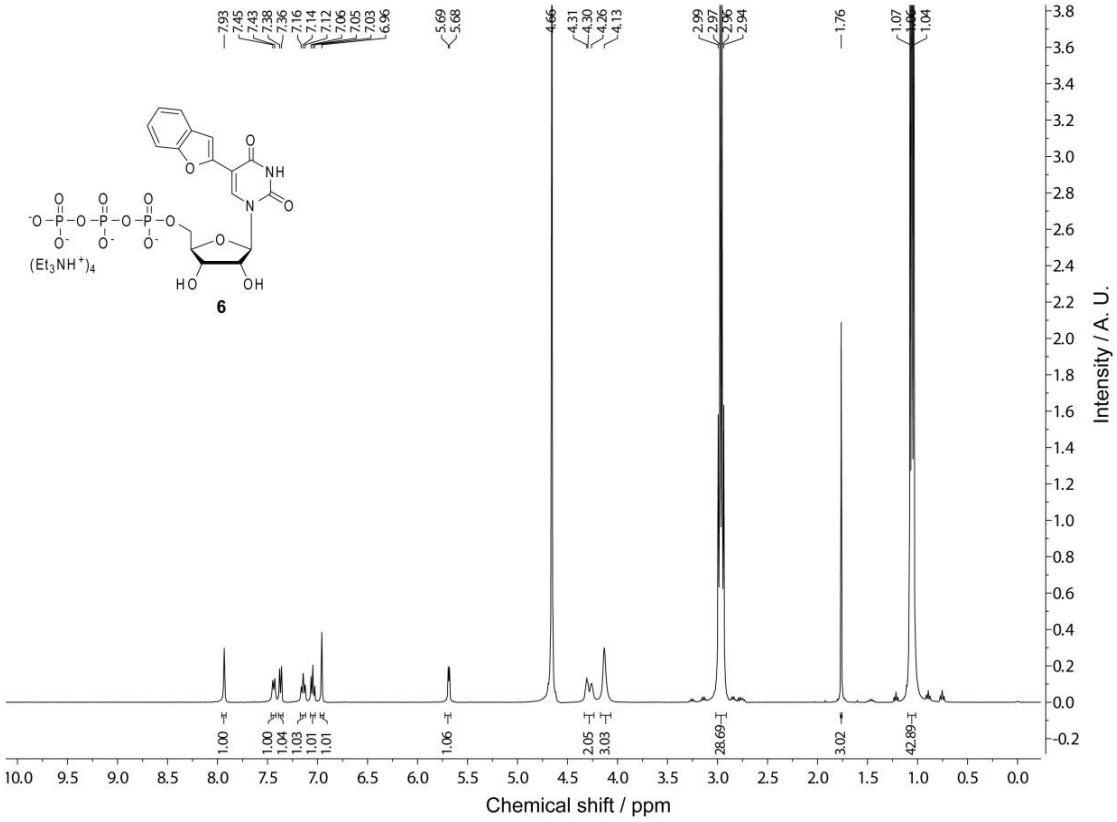
^{1}H-NMR of ribonucleoside **5** in d_6-DMSO



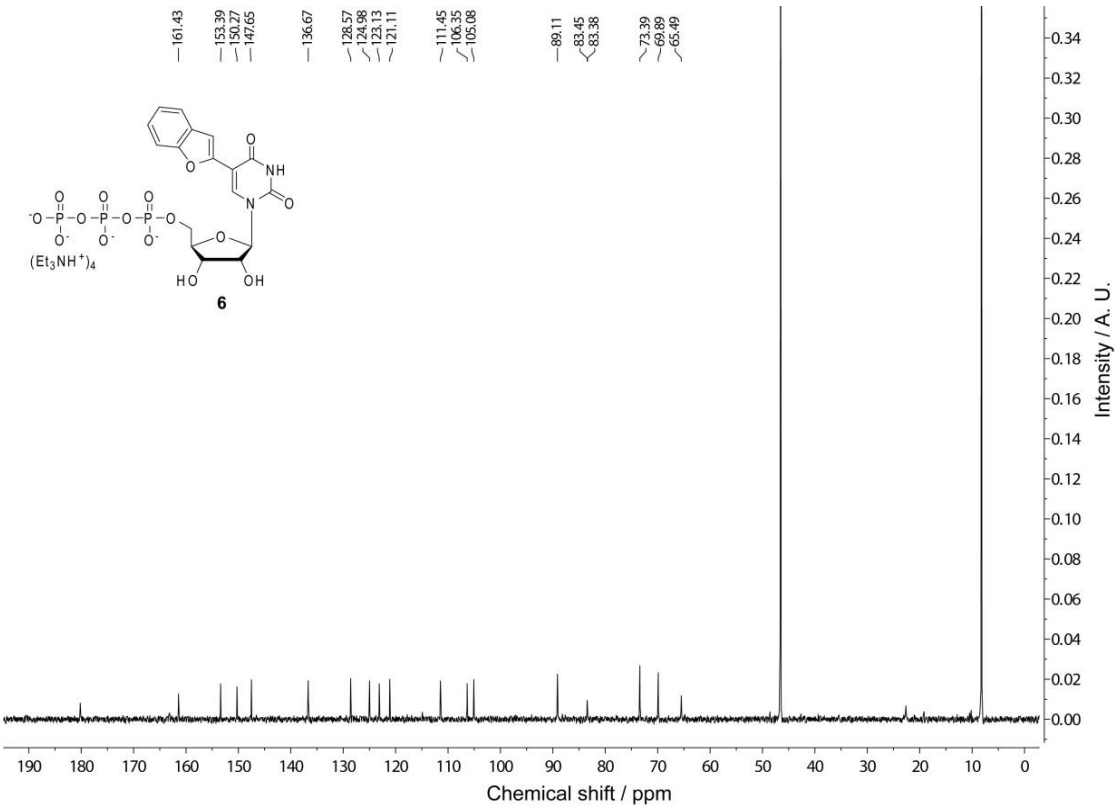
^{13}C-NMR of ribonucleoside **5** in d_6-DMSO



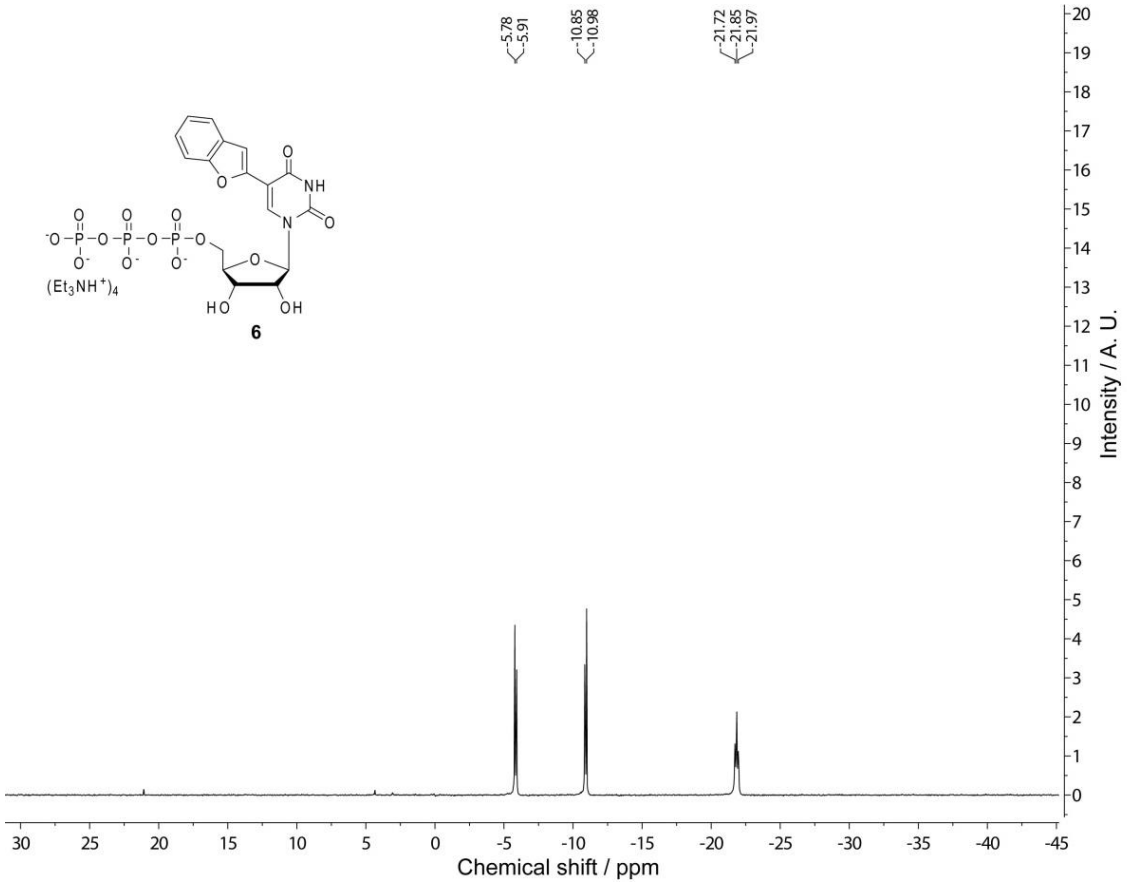
^{1}H-NMR of ribonucleoside triphosphate **6** in D_2O



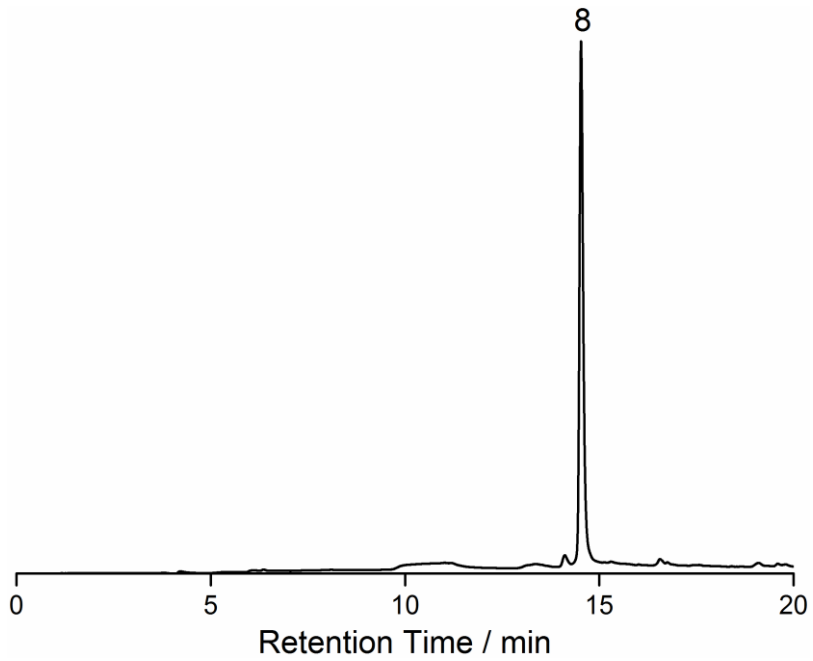
^{13}C-NMR of ribonucleoside triphosphate **6** in D_2O. Trace amounts of triethylammonium acetate buffer is also present.



^{31}P-NMR of ribonucleoside triphosphate **6** in D_2O



RP-HPLC chromatogram of modified RNA transcript **8** at 260 nm. Mobile phase A: 50 mM triethylammonium acetate buffer (pH 7.5), mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 10 min and 10−55% B in 10 min.



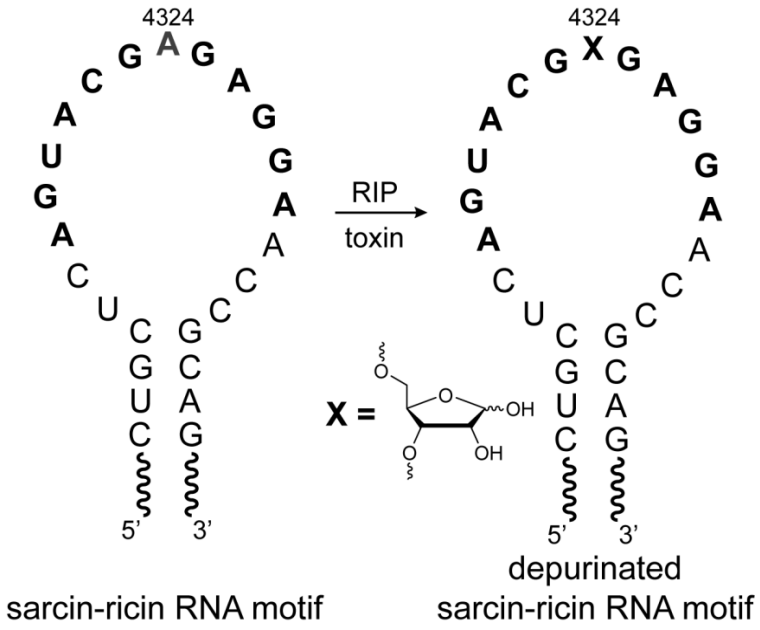
Chapter 2B

5-Benzofuran-modified 2'-deoxyuridine as a probe for the fluorescence detection of RNA abasic site

2B.1 Introduction

Unlike DNA abasic sites, which are common DNA lesions, the formation of abasic sites in RNA is rare and is almost uniquely associated with the depurination activity of ribosome inactivating protein (RIPs) toxins.[1] These toxins (eg., ricin and saporin) arrest the function of the highly conserved RNA motif of eukaryotic 28S rRNA called the sarcin–ricin loop, which interacts with elongation factors essential for the protein synthesis, by depurinating a specific adenosine residue (A_{4324}) on the loop (Scheme 1).[2] Consequently, the elongation factors necessary for the translation process no longer bind to the depurinated ribosome leading to cell death.[3] In particular, ricin has been considered as a potential bioterrorism agent because of its high toxicity, easy isolation procedures from natural source and lack of effective treatment against its exposure.[4] Traditionally, RIP toxins have been detected by using enzyme-linked immunosorbent assays (ELISA) and antibody-based immunoassays.[5] Methods have also been developed to monitor the specific depurination activity of RIP toxins by using radiolabeling, immunoaffinity chromatography, electrochemiluminescence and mass spectrometry techniques.[6] Usually these methods are laborious and involve elaborate assay setups and radiolabeling procedures. Alternatively, a few abasic-site-sensitive fluorescence probes have provided effective tools to directly detect the formation of abasic sites in RNA.[7,8] Nevertheless, abasic site detection assays that are compatible with screening formats are highly desired for the discovery of RIPs inhibitors.

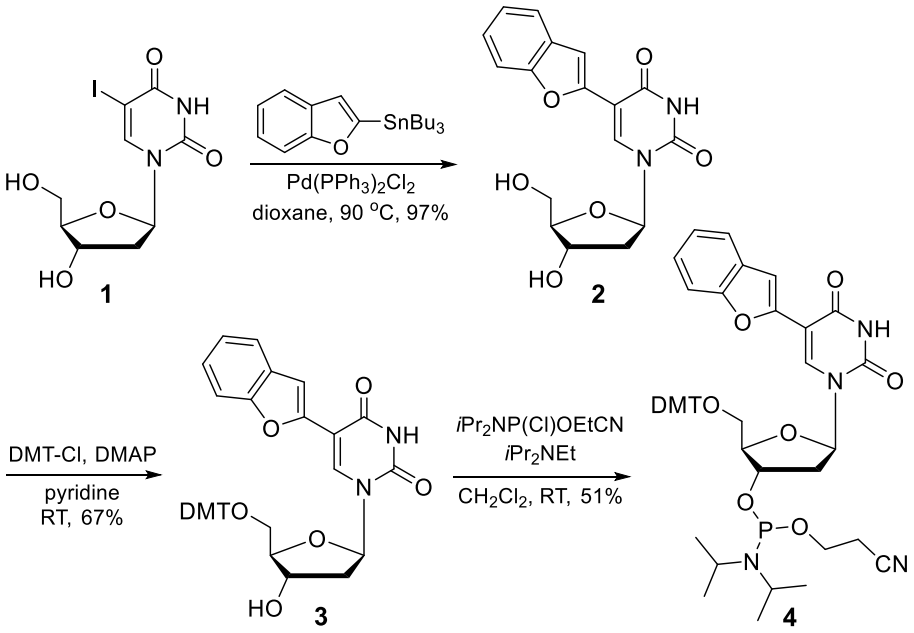
In the first part of this chapter we have described the microenvironment sensitive benzofuran-conjugated uridine as a probe to detect the DNA abasic site. Superior probe like properties of benzofuran-conjugated uridine motivated us to study the structure, dynamics and recognition properties of DNA. Here, we describe the synthesis, photophysical characterization and incorporation of benzofuran-conjugated 2'-deoxyuridine analogue **2** into DNA ONs. Remarkably, upon incorporation into single stranded and double stranded ONs the emissive nucleoside shows significantly enhanced emission intensity compared to the free nucleoside, a property, which is rarely displayed by the most of the fluorescent nucleoside analogs. Furthermore, using fluorescence spectroscopy we illustrate the photophysical behaviour of the emissive nucleoside incorporated into DNA ONs in different base environments. Finally, as a proof of responsiveness of the nucleoside to environmental changes we describe the ability of a DNA ON reporter labeled with modified nucleoside **2** in signalling the presence of an abasic site in a model depurinated sarcin–ricin RNA motif of the eukaryotic 28S rRNA.



Scheme 1. Schematic diagram showing the sequence of the sarcin-ricin hairpin motif of the ribosomal RNA. The conserved residues are highlighted in bold. RIP toxins catalytically depurinate the A_{4324} residue producing the depurinated RNA motif. The structure of the RNA abasic site X is also shown.

2B.2 Results and Discussion

2B.2.1 Synthesis of nucleoside 2



DMT-Cl = 4,4'-dimethoxytritylchloride; DMAP = 4-dimethylaminopyridine

Scheme 2. Synthesis of fluorescent nucleoside, 1-(2-deoxy-β-D-ribofuranosyl)-5-(benzofuran-2-yl)uracil **2** and corresponding phosphoramidite substrate **4**.

The fluorescent analogue, 5-(benzofuran-2-yl)-2'-deoxyuridine **2** has been synthesized under typical Stille cross-coupling reaction conditions by reacting 5-iodo-2'-deoxyuridine **1** with 2-(tri-n-butylstannyl)benzofuran in the presence of a palladium catalyst, $Pd(PPh_3)_2Cl_2$. The phosphoramidite substrate **4** necessary for the solid-phase ON synthesis was prepared first protecting the 5'-hydroxyl with dimethoxytrityl group followed by phosphitylation of the 3'-hydroxyl in the presence of 2-cyanoethyl diisopropylchlorophosphoramidite (Scheme 2).

2B.2.2 Photophysical characterization of nucleoside 2

Like previous study before incorporating into ONs, the photophysical properties of nucleoside **2** have been evaluated by performing UV absorption, steady-state, and time-resolved fluorescence spectroscopic measurements in solvents of different polarity and viscosity. The ground-state electronic spectrum of **2** in water reveals distinct absorption bands at 265 nm, 272 nm and 322 nm (Figure 1). When measured in different solvents, the absorption maxima of **2** are marginally affected by solvent polarity. However, the excited-state electronic properties are substantially influenced by changes in solvent polarity environment. An aqueous solution of the nucleoside upon excitation at its lowest energy maximum (322 nm) displays a strong emission band in the visible region (λ_{em} = 446 nm, Figure 1, Table 1). As the solvent polarity is gradually decreased from water to dioxane, the nucleoside exhibits a nearly 2.5-fold quenching in fluorescence intensity and a significant hypsochromic shift (446–406 nm). Furthermore, the quantum yields determined in various solvents follow a similar decreasing trend as that of the fluorescence intensity in respective solvents (Table 1). A positive correlation between the Stokes shift determined in various solvents and Reichardt's microscopic solvent polarity parameter, $E_T(30)$, also ascertains the sensitivity of the nucleoside to changes in solvent-polarity environment (Figure 2).[10] The effect of solvent polarity on the excited-state decay kinetics of nucleoside **2** has also been investigated by time-resolved fluorescence spectroscopy. The nucleoside in water has the highest lifetime of 2.38 ns, which decreases significantly to 0.33 ns in dioxane (Table 1, Figure 3).

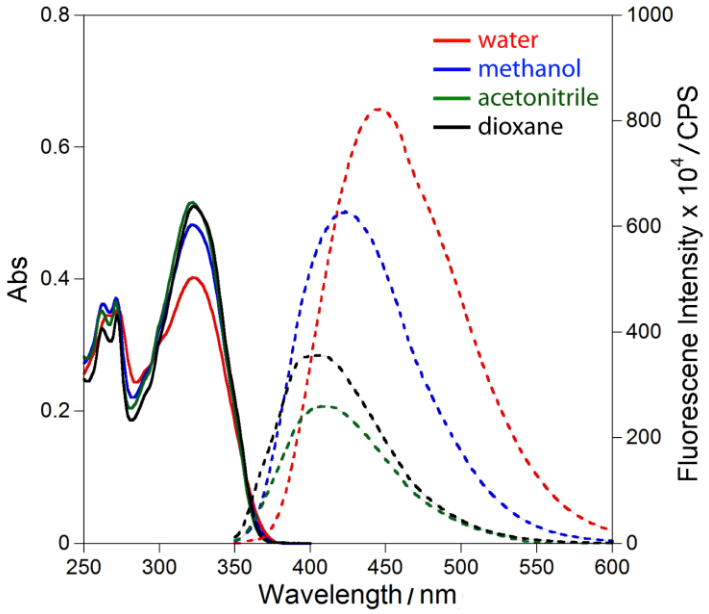


Figure 1. Absorption (25 µM, solid lines) and emission (5.0 µM, dashed lines) spectra of nucleoside **2** in different solvents. For fluorescence studies samples were excited at 322 nm, and excitation and emission slit widths were maintained at 3 nm and 5 nm, respectively.

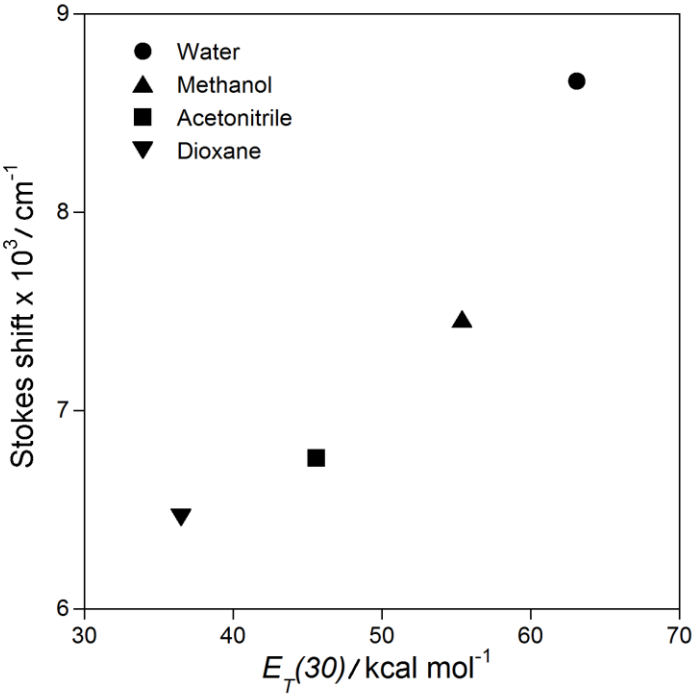


Figure 2. A plot of Stokes shift vs. $E_T(30)$, a microscopic solvent polarity parameter.

Table 1. Photophysical properties of **2** in different solvents

Solvent	λ_{max}[a] (nm)	λ_{em} (nm)	I_{rel}[b]	Φ[c]	τ_{ave}[c] (ns)
water	322	446	1.0	0.19	2.38
methanol	322	423	0.8	0.12	0.78
acetonitrile	322	411	0.3	0.04	0.27
dioxane	322	406	0.4	0.07	0.33
ethylene gycol	429	429	2.0	0.40	2.50
glycerol	427	427	2.7	0.54	3.56

[a] λ_{max} corresponding to the lowest energy maximum is given. [b] Relative emission intensity is given with respect to intensity in water. [c] Standard deviations for Φ and τ_{ave} are $\leq$ 0.002 and 0.02 ns, respectively.

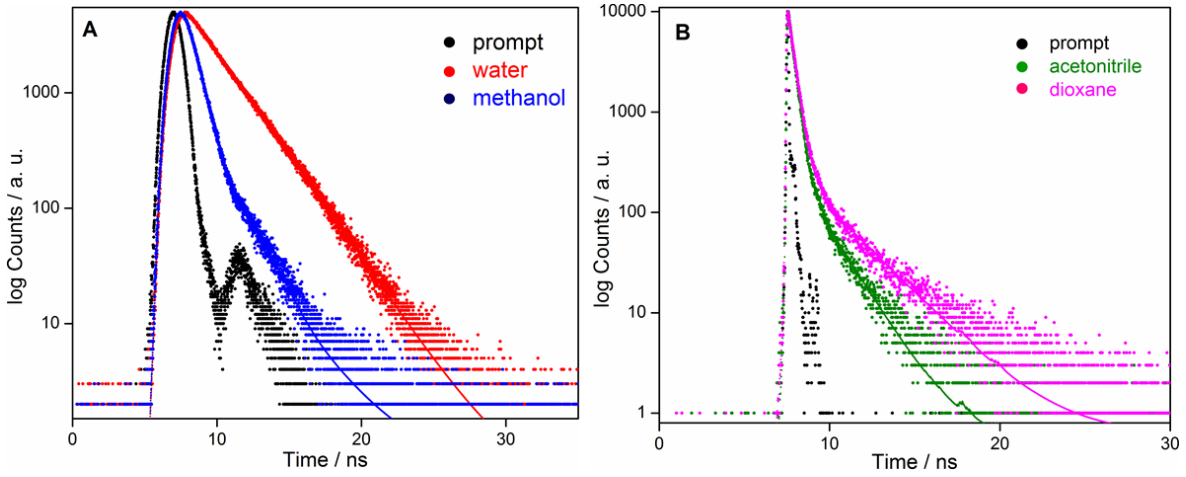


Figure 3. Excited state decay profile of nucleoside **2** in various solvents. Laser profile is shown in black (prompt). Curve fits are shown in solid lines. Fluorescence intensity decay kinetics in water was found to be monoexponential whereas in other solvents was found to be biexponential. For lifetimes see Table 1.

The relative conformation of the pyrimidine ring and conjugated aromatic ring separated by a rotatable aryl–aryl bond in modified nucleoside analogs, which has a direct impact on the conjugation and hence, the fluorescence properties, has been shown to be sensitive to molecular crowding effects and viscosity.[11] In comparison with low viscous media, viscous media restricts free rotation resulting in structural rigidification and usually enhanced fluorescence emission.[12] A conjuated nucleoside analogue incorpoarted into ONs can undergo structural rigidification-derigidification due to interactions with neighbouring bases during a folding or recognition process.[11] In order to assess the effects of changes in the conformation of the benzofuran moiety relative to the nucleobase, further fluorescence

characterization of nucleoside **2** has been performed in solvents of similar polarity but varying viscosity. As the solvent viscosity is increased from ethylene glycol ($\eta_{25°C}$ = 16.1 cP) to glycerol ($\eta_{25°C}$ = 934 cP), the nucleoside shows enhanced emission intensity (nearly 1.5-fold) in glycerol as compared to in ethylene glycol with no apparent change in emission maximum (Figure 4 and Table 1).

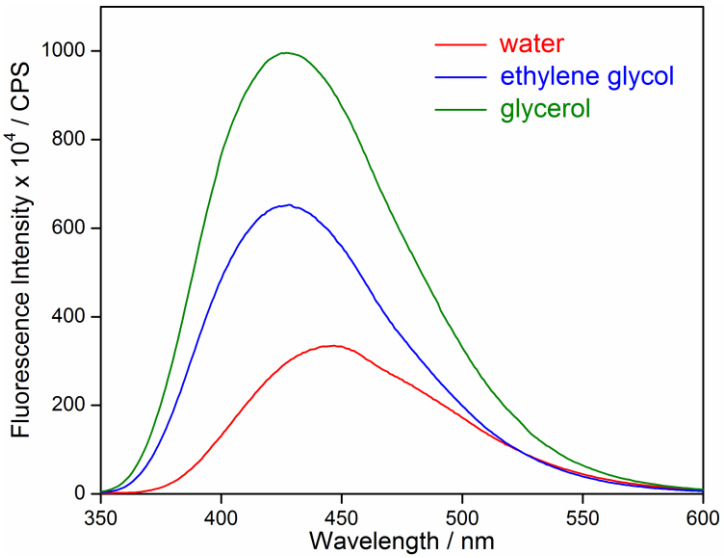


Figure 4. Emission spectra of nucleoside **2** (5.0 μM) in solvents of varying viscosity. Samples were excited at 322 nm, and excitation and emission slit widths were maintained at 1 nm and 10 nm, respectively.

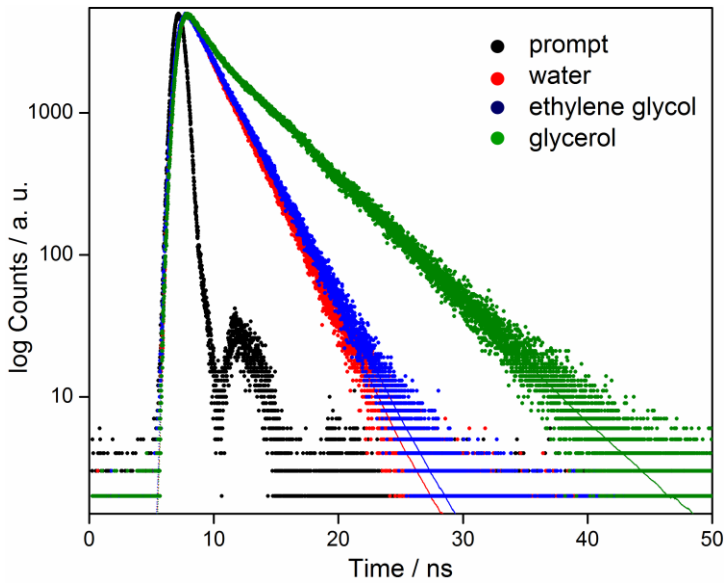


Figure 5. Excited state decay profile of nucleoside **2** in various solvents. Laser profile is shown in black (prompt). Curve fits are shown in solid lines. Fluorescence intensity decay kinetics in water and ethylene glycol was found to be monoexponential whereas in glycerol was found to be biexponential. For lifetimes see Table 1.

Life time measurements also reveal a longer decay time for the nucleoside in glycerol as compared to in ethylene glycol, which is consistent with the viscosity difference between the solvents (Figure 5 and Table 1). Taken together, emission in the visible region with a reasonable quantum yield, and the dual-sensitivity of the nucleoside fluorescence to solvent-polarity and viscosity changes prompted us to study the responsiveness of the emissive nucleoside within DNA ONs.

2B.2.3 Incorporation of nucleoside 2 into DNA ONs

To study the effect of flanking bases on the fluorescence of the modified nucleoside a series of DNA ONs **5**–**8** were synthesized in which the nucleoside **2** was placed in between different bases (Figure 6). The modified phosphoramidite was incorporated into 17-mer DNA ONs at a single position under standard solid-phase ON synthesis conditions. The deprotected ONs were then purified by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions. The integrity of full-length modified ONs was confirmed by mass analysis (Table 2).

5	5' GCGATCAA**2**AACTAGCG 3'	**5c**	5' CGCTAGTTATTGATCGC 3'
6	5' GCGATCAT**2**TACTAGCG 3'	**6c**	5' CGCTAGTAAATGATCGC 3'
7	5' GCGATCAC**2**CACTAGCG 3'	**7c**	5' CGCTAGTGAGTGATCGC 3'
8	5' GCGATCAG**2**GACTAGCG 3'	**8c**	5' CGCTAGTCACTGATCGC 3'
9	5' GCGATCAGTGACTAGCG 3'	**10**	5' CGCTAGTCTCTGATCGC 3'
11	5' CGCTAGTCGCTGATCGC 3'	**12**	5' CGCTAGTCCCTGATCGC 3'

Figure 6. Sequence of fluorescently modified (**5**–**8**) and custom synthesized (**9**–**12** and **5c**–**8c**) DNA ONs. **9** is a control unmodified DNA. While hybridization of **5**–**8** with **5c**–**8c**, respectively, will place the nucleoside **2** opposite to its complementary base, hybridization of **8** with **10**–**12** will place the nucleoside **2** opposite to mismatched bases.

Table 2. ε_{260} and mass data of modified ONs

Oligonucleotides	ε_{260} ($M^{-1}cm^{-1}$)	Calcd. mass	Observed mass
5	17.3×10^4	5305.5	5305.7
6	16.5×10^4	5287.5	5288.3
7	16.2×10^4	5257.5	5257.9
8	16.9×10^4	5337.5	5337.4
19	26.2×10^4	8963.8	8964.0

The presence of benzofuran moiety can potentially perturb the structure of the ONs and hence, the formation of a stable duplex with its complementary ONs. The stability of a duplex assembled by hybridizing one of the modified ONs (**8**) with its complementary ON (**8c**) was studied by UV-thermal denaturation experiment. Thermal denaturation analysis showed only a small difference in T_m between the modified duplex **8•8c** and control unmodified duplex **9•8c** indicating that the benzofuran modification has only marginal effect on the duplex stability (Figure 7, Table 3).

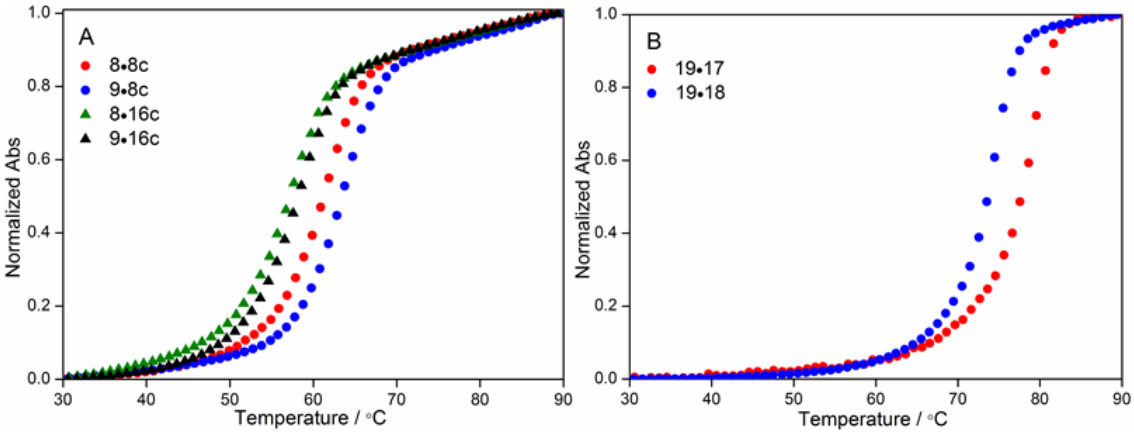


Figure 7. (**A**) Thermal melting of control unmodified and fluorescently-modified duplexes (1 μM). Small difference in T_m values between the modified duplexes and control unmodified duplexes indicates that the benzofuran moiety has minimum impact on the duplex stability. (**B**) Thermal melting of DNA probe-RIP RNA substrate duplex (**19•17**) and DNA probe-depurinated product mimic duplex (**19•18**). The UV-thermal melting profiles indicate that both RIP substrate **17** and product mimic **18** form stable duplexes with fluorescently modified DNA ON **19**.

Table 3. Tm values of control and modified duplexes

Duplexes	T_m (°C)	Duplexes	T_m (°C)
8•8c	61.5 ± 0.7	**8•11**	58.1 ± 0.6
9•8c (control)	63.0 ± 0.2	**8•12**	58.5 ± 0.9
8•16c	58.0 ± 0.8	**8•16a**	53.3 ± 0.6
9•16c (control)	60.7 ± 0.7	**19•17**	79.7 ± 0.8
8•10	56.6 ± 0.7	**19•18**	74.7 ± 0.3

2B.2.4 Photophysical characterization of nucleoside 2 in different base environment

Photophysical properties of fluorescent nucleosides incorporated into ONs can be altered by a variety of mechanisms such as stacking of the chromophore with flanking bases, collisional and hydrogen bonding interactions with neighbouring bases, solvation-desolvation effect,

rigidification–derigidification of the chromophore, and excited state processes involving neighbouring bases.[11,13,14] The influence of neighbouring bases on the fluorescence properties of nucleoside **2** has been studied by performing steady-state fluorescence measurements with DNA ONs **5**–**8** and duplexes constructed by hybridizing **5**–**8** with respective complementary ONs **5c**–**8c** (Figure 6). Single stranded ONs **5**–**7** in which the emissive nucleoside is placed between dA, dT and dC residues, respectively, exhibits significantly enhanced emission as compared to the free nucleoside (Figure 8, Table 4). Interestingly, nucleoside **2** incorporated into duplexes **5•5c** and **6•6c** show enhanced emission compared to respective single stranded ONs (Figure 6). This observation is particularly noteworthy because the majority of emissive nucleoside analogs (eg., 2-AP, pyrroloC) show progressive fluorescence quenching upon incorporation into single stranded and double stranded ONs, a major reason that has hampered the use of many such analogs in vivo assays.

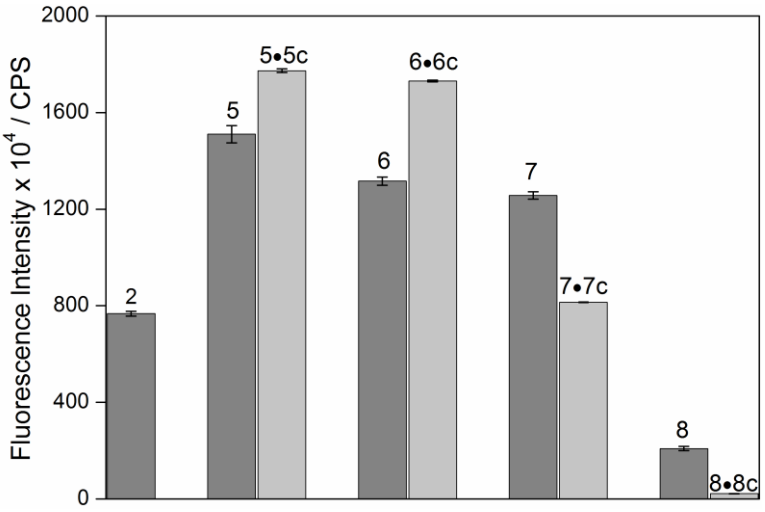


Figure 8. Relative fluorescence intensity of modified ONs (**5**–**8**, 1 μM) and corresponding duplexes (1 μM) in 20 mM cacodylate buffer (pH 7.0, 100 mM NaCl, 0.5 mM EDTA) at respective emission maximum (422–444 nm, see Table 4 for details). Samples were excited at 330 nm, and excitation and emission slit widths were maintained at 3 nm and 10 nm, respectively.

However, ON **8** in which the nucleoside **2** is flanked by guanosine residues displays a slightly blue shifted and markedly less intense (~3-fold) emission band compared to the free nucleoside **2** (Figure 6). The quenching effect is more pronounced (~10-fold) when the emissive nucleoside is placed opposite to the complementary base in a perfect duplex **8•8c**. 2-AP and several other fluorescent nucleoside analogs incorporated into ONs also exhibit similar fluorescence intensity quenching. Furthermore, duplexes (**8•10**, **8•11** and **8•12**) in which the emissive nucleoside is placed opposite to mismatched bases also show similar

fluorescence quenching as that of the perfect duplex **8•8c** (Figure 9). This fluorescence quenching along with a small spectral shift towards the blue region maybe due to a combination of the following reasons: an electron transfer process between the fluorescent nucleoside and adjacent guanosine residues,[13,14] desolvation effect and alterations in the conformation of the benzofuran moiety with respect to the nucleobase.[11] These results clearly reveal the influence of neighbouring bases on the fluorescence of nucleoside **2**. Such a property has been employed in devising nucleic acid-based molecular beacons and in the detection of a specific mismatch in DNA duplexes.[15]

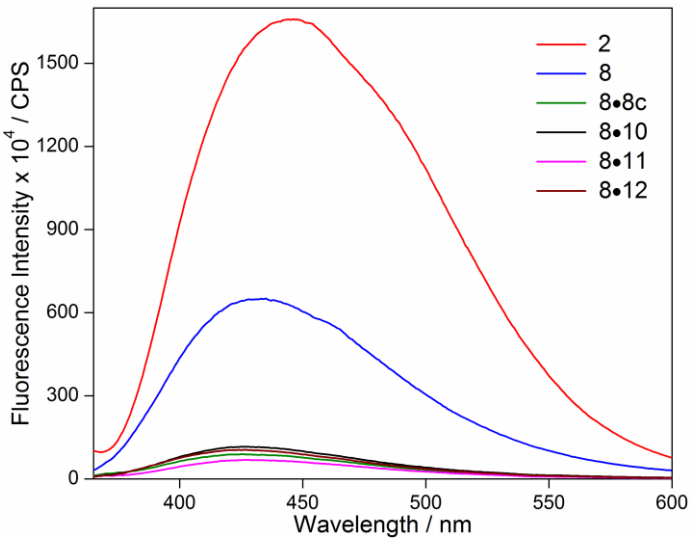


Figure 9. Emission spectra of **8** and duplexes assembled by hybridizing **8** with complementary (**8c**) and mismatched ONs (**10**–**12**). Samples were excited at 330 nm, and excitation and emission slit widths were maintained at 7 nm and 10 nm, respectively.

Table 4. Emission maximum of modified ONs (**5**–**8**) and duplexes made by hybridizing **5**–**8** with respective complementary ONs **5c**–**8c**.

Sample	λ_{em}	I_{rel}[a]	Sample	λ_{em}	I_{rel}[a]
2	444	1	**7**	433	1.64
5	435	1.97	**7•7C**	422	1.06
5•5C	428	2.31	**8**	432	0.27
6	434	1.72	**8•8C**	429	0.03
6•6C	422	2.26			

[a] I_{rel} Relative emission intensity is given with respect to intensity of the nucleoside **2**.

2B.2.5 Fluorescence detection of an RNA abasic site

Early methods to detect abasic sites were largely focused toward the detection of abasic sites in DNA as they are formed spontaneously or as intermediates in the base excision repair process of damaged nucleobases.[16,17] These methods relied on the specific and irreversible reaction between aldehyde-reactive probes and abasic sites of isolated DNA.[18] Later, Greenberg developed a more sensitive method by utilizing biotinylated cysteine to specifically detect the oxidized form of abasic lesion, 2-deoxyribonolactone.[19] Matray and Kool used pyrene modified deoxyribonucleotide triphosphate, which was preferentially incorporated opposite to an abasic site by DNA polymerase to identify the presence of abasic sites.[20] Alternatively, fluorescence-based methods using probes that show changes in their emission properties when placed adjacent or opposite to abasic sites were found to be more useful as they offered direct detection of abasic sites in DNA and RNA.[7,8,21,22] Shipova and Gates developed one of the first fluorimetric assays to monitor the time-depended generation of abasic sites by using 2-AP labeled DNA hairpin constructs.[22] In this study 2-AP was placed adjacent to a guanine residue, which was alkylated using leinamycin[23] and subsequently excised to produce an abasic site adjacent to the fluorescent probe. Upon treatment with leinamycin, 2-AP positively reported the formation of abasic sites with a significant enhancement in fluorescence intensity. Adopting a similar approach, pyrene- and thiophene-modified fluorescent nucleoside analogs and a fluorescent ligand that specifically binds to an abasic site was utilized in monitoring the formation of abasic sites in RNA ONs.[7,8]

Our previous study revealed that a short RNA ON reporter containing the benzofuran-conjugated ribonucleoside analogue could specifically signal the presence of a DNA abasic site in an RNA-DNA heteroduplex.[9] This observation encouraged us to study the impact of placing the deoxyribonucleoside **2** opposite an RNA abasic site. As before a series of DNA-RNA heteroduplexes was assembled by annealing modified DNA ONs **5**–**8** to respective complementary RNA ONs **13c**–**16c** and ONs **13a**–**16a** containing a chemically stable abasic site surrogate, tetrahydrofuran (Figure 10A). Nucleoside **2** flanked by dA, dT and dC residues when placed opposite to an abasic site in duplexes **5•13a**, **6•14a** and **7•15a** showed discernible quenching in fluorescence intensities as compared to corresponding prefect duplexes **5•13c**, **6•14c** and **7•15c** (Figure 10B). However, the perfect duplex **8•16c** and abasic site-containing duplex **8•16a** in which the nucleoside is flanked by dG residues exhibited very weak fluorescence.

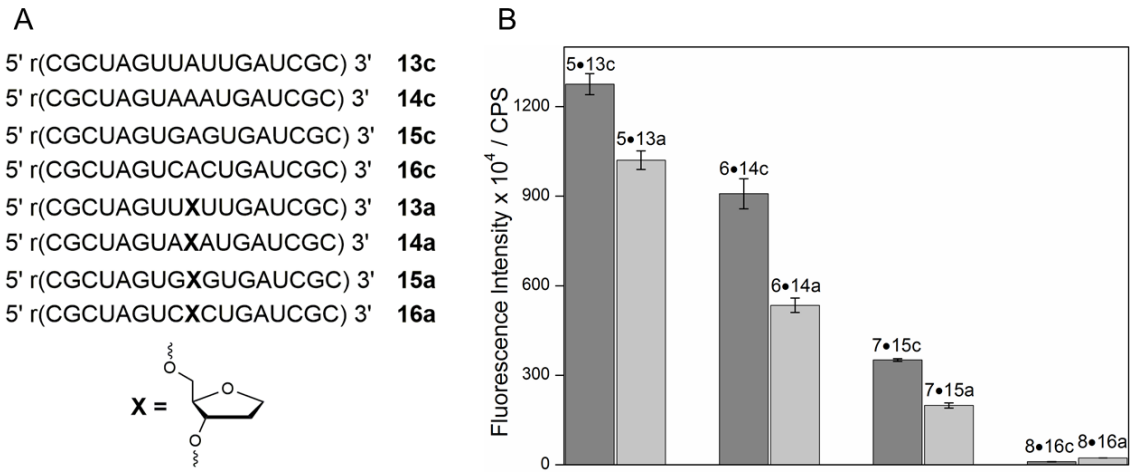


Figure 10. [**A**] Sequence of synthetic RNA ONs (**13c**–**16c** and **13a**–**16a**). While hybridization of **5**–**8** with **13c**–**16c**, respectively, will place the nucleoside **2** opposite to its complementary base, hybridization of **5**–**8** with **13a**–**16a** will place the nucleoside **2** opposite to a chemical stable abasic site surrogate **X**. [**B**] Fluorescence intensity (*FI*) of heteroduplexes (1 μM) in 20 mM cacodylate buffer (pH 7.0, 100 mM NaCl, 0.5 mM EDTA) at 430 nm. Samples were excited at 330 nm, and excitation and emission slit widths were maintained at 3 nm and 10 nm, respectively.

To further assess the ability of emissive nucleoside in detecting the presence of an abasic site in a biologically relevant RNA motif we decided to use model RNA ONs **17** and **18** (Figure 8A). ON **17** containing the conserved sarcin-ricin loop region of eukaryotic 28S rRNA is a commonly used substrate to study the depurination activity of RIPs.[6b,7,8] ON **18** containing a chemically stable abasic site substitute, tetrahydrofuran, acts as a model of the depurinated product that would be obtained upon depurination of substrate **17** by RIP toxins.[7b,8] A complementary DNA ON **19** labeled with the fluorescent analogue **2** was synthesized, which upon hybridization to a model sarcin–ricin RNA substrate **17** and depurinated product mimic **18** would place **2** opposite to a complementary base and an abasic site, respectively (Figure 11A). The UV-thermal melting profiles indicate that both RIP substrate **17** and product mimic **18** form stable duplexes with fluorescently modified DNA ON **19 (**Table 3**)**. A duplex of fluorescent ON and RIP RNA substrate (**19•17**) shows a strong emission band whose intensity is nearly 1.5-fold higher than the single stranded ON **19** (Figure 11B). Interestingly, a duplex of ON probe and depurinated product mimic **19•18** exhibits drastically quenched emission (~8-fold) as compared to the substrate duplex **19•17** (Figure 11B). Although the exact morphology of the fluorescent nucleoside **2** in the product duplex **19•18** is unknown, we believe that the observed quenching in fluorescence intensity can be possibly due to the following reasons. In the substrate duplex **19•17** the benzofuran moiety tagged at the 5-position of the base is extrahelical and projected towards the major

groove and hence, it is less stacked and located away from the guanosine residues of ON **17** (Figure 12). However, in product duplex **19•18** the benzofuran ring is presumably intrahelical because the nucleoside **2** opposite to an abasic site can potentially undergo anti-to-syn conformational change as it is not restricted by complementary base pairing.[11] In this situation the fluorophore is more stacked and closer to guanosine residues of ON **18** resulting in quenching of fluorescence intensity. Together, it can be inferred that the nucleoside **2** in product duplex **19•18** signals the presence of an abasic site in RNA, albeit with quenched emission. This attribute of the emissive nucleoside can be potentially implemented in a fluorescence hybridization assay to detect the depurination activity of highly toxic RIPs.[7,8]

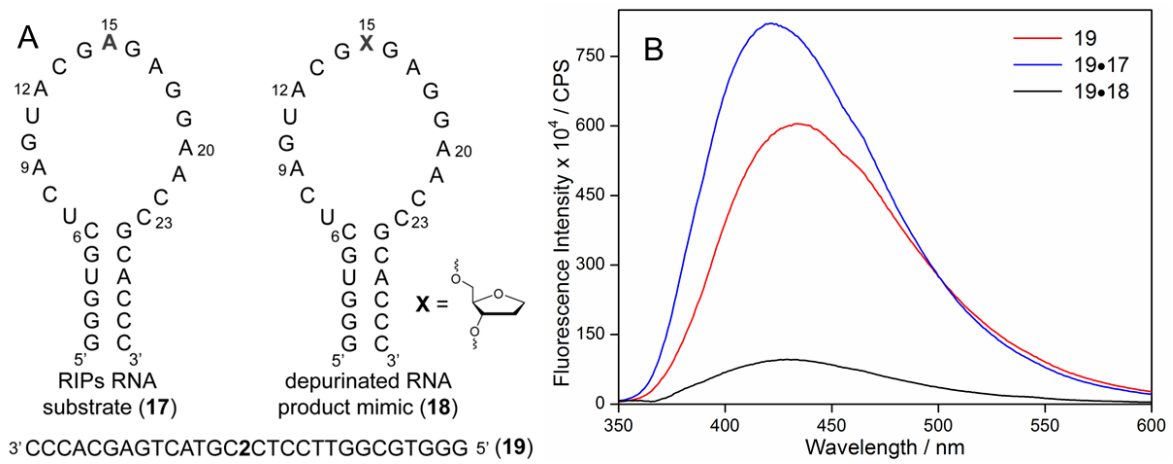


Figure 11. A) Sequence of synthetic RIPs RNA substrate (**17**) and depurinated product mimic (**18**) containing a chemically stable abasic site substitute **X**. Adenosine residue A_{15} corresponding to A_{4324} of rat 28S rRNA is specifically depurinated by RIP toxins to produce an abasic site. cDNA **19** containing the fluorescent nucleoside **2** is also shown. B) Emission spectra (1 μM) of substrate (**19•17**) and depurinated product duplexes (**19•18**). Samples were excited at 330 nm, and excitation and emission slit widths were maintained at 7 nm and 10 nm, respectively.

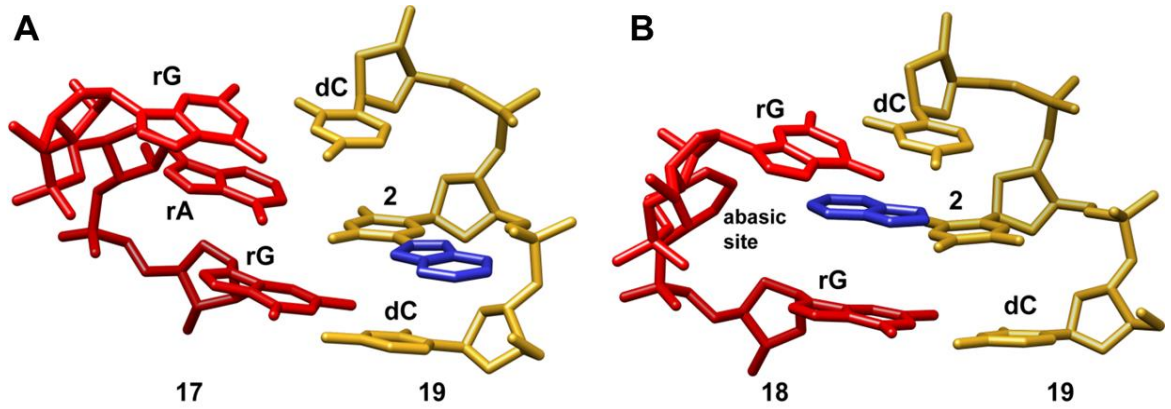


Figure 12. A schematic diagram showing the possible conformation of the emissive nucleoside **2** in (A) RIP substrate duplex (**19•17**) and (B) depurinated product duplex (**19•18**). For clarity only flanking bases are shown. The benzofuran moiety (blue) is persumably intrahelical in duplex **19•18** because the nucleoside **2** can potentially undergo anti-to-syn conformational change as it is not restricted by complementary base pairing. This figure has been generated using UCSF Chimera software.

2B.3 Conclusions

A microenvironment-sensitive fluorescent 2'-deoxyuridine analogue based on a 5-(benzofuran-2-yl)pyrimidine core, which displays emission in the visible region, has been incorporated into DNA ONs. Notably, upon incorporation into duplexes the nucleoside **2** remarkably maintains its fluorescence efficiency, a property, which is seldom exhibited by the majority of fluorescent nucleoside analogs. Furthermore, the ability of the fluorescent analogue incorporated into a DNA ON reporter to detect the presence of an abasic site in RNA highlights the potential of the nucleoside **2** as a fluorescent probe. Since the conformation of the emissive nucleoside, its surrounding environment and or its interaction with neighbouring bases are likely to be altered during a folding or recognition event, it is expected that this environment-sensitive nucleoside analogue will be a useful probe in studying the structure and function of nucleic acids.

2B. 4 Experimental Section

2B.4.1 Materials

5-iodo-2'-deoxyuridine, benzofuran, butyllithium, tributyltin chloride, *bis*(triphenylphosphine)-palladium(II) chloride, DMT-Cl, *N*,*N*-diisopropylethylamine were obtained from Sigma-Aldrich. 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite was purchased from Alfa Aesar. *N*-benzoyl-protected dA, dT, *N*,*N*-dimethylformamidine-

protected dG and *N*-acetyl-protected dC phosphoramidite substrates for solid phase ON synthesis were purchased from ChemGenes and Proligo. All other reagents and solid supports for the solid phase synthesis were obtained from ChemGenes. Synthetic DNA ONs were purchased from Integrated DNA Technologies, Inc. and purified by polyacrylamide gel electrophoresis (PAGE) under denaturing condition, and desalted on Sep-Pak Classic C18 cartridges (Waters Corporation). Custom synthesized oligoribonucleotides purchased from Dharmacon RNAi Technologies were deprotected according to the supplier's protocol, PAGE-purified, and desalted on Sep-Pak Classic C18 cartridges. Spectroscopy grade solvents and chemicals (BioUltra grade) for preparing buffer solutions were purchased from Sigma-Aldrich. Autoclaved water was used in all biochemical reactions and fluorescence analysis.

2B.4.2 Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400 and Bruker Ultrashield 500 MHz WB Plus spectrometers. Mass measurements were recorded on Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer, MicroMass ESI-TOF and Water Synapt G2 High Definition mass spectrometers. Modified DNA ONs were synthesized on an Applied Biosystems RNA/DNA synthesizer (ABI-394). Absorption spectra were recorded on a PerkinElmer, Lambda 45 UV-Vis spectrophotometer. UV-thermal melting studies of ONs were performed on a Cary 300Bio UV-Vis spectrophotometer. Steady State fluorescence and time-resolved experiments were carried out in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3).

2B.4.3 Synthesis

1-(2-deoxy-β-D-ribofuranosyl)-5-(benzofuran-2-yl)uracil 2: To a suspension of 5-iodo-2'-deoxyuridine (0.776 g, 2.18 mmol, 1 equiv) and *bis*(triphenylphosphine)-palladium(II) chloride (0.076 g, 0.11 mmol, 0.05 equiv) in degassed anhydrous dioxane (23 ml) was added 2-(tri-n-butylstannyl)benzofuran (1.33 g, 3.27 mmol, 1.5 equiv). The reaction mixture was heated at 90 °C for 2.5 h and filtered through celite pad. The celite pad was washed with hot dioxane (3 x 15 ml) and the filtrate was evaporated. The solid residue was washed extensively with hexane, dissolved in a minimum amount of hot 1:1 dioxane:methanol solution, and then the product was precipitated with hexane. The precipitate was stirred at ~4 °C for 2 h, filtered and dried to afford the product **2** as a white solid (0.758 g, 97%).

(CH_2Cl_2:MeOH = 9:1) R_f = 0.51; ^{1}H NMR (400 MHz, *d6*-DMSO): δ (ppm) 11.77 (s, 1H), 8.75 (s, 1H), 7.62 (d, *J* = 7.2 Hz, 1H), 7.55 (d, *J* = 8.0 Hz, 1H), 7.34 (s, 1H), 7.28 ($_{app}$t, *J* = 7.4 Hz, 1H), 7.22 ($_{app}$t, *J* = 7.4 Hz, 1H), 6.23 (t, *J* = 6.4 Hz, 1H), 5.32 (d, *J* = 4 Hz, 1H), 5.25 (t, *J* = 4.4 Hz, 1H), 4.37–4.30 (m, 1H), 3.89–3.87 (m, 1H), 3.74–3.65 (m, 2H), 2.29–2.19 (m, 2H); ^{13}C NMR (100 MHz**,** *d6*-DMSO**)**: δ (ppm) 160.3, 153.0, 149.4, 149.1, 137.1, 128.8, 124.3, 123.0, 121.0, 110.8, 104.7, 103.8, 87.6, 85.1, 70.1, 60.8, 40.5; HRMS: (m/z): Calculated for $C_{17}H_{16}N_2O_6Na$ $[M+Na]^+$ = 367.0906, found: 367.0909; $\lambda_{max}(H_2O)$ = 272 and 322 nm, ε_{272} = 14106 $M^{-1}cm^{-1}$, ε_{322} = 16093 $M^{-1}cm^{-1}$, ε_{260} = 12613 $M^{-1}cm^{-1}$.

DMT-protected 2'-deoxyuridine 3: A solution of **2** (0.702 g, 2.04 mmol, 1 equiv), DMT-Cl (0.87 g, 2.45 mmol, 1.2 equiv) and DMAP (62 mg, 0.51 mmol, 0.25 equiv) in anhydrous pyridine (7 ml) was stirred at room temperature under nitrogen atmosphere for 12 h. Pyridine was evaporated under reduced pressure. The residue was purified by silica gel column chromatography (2 % methanol in CH_2Cl_2 containing 1% triethylamine) to afford the product **3** as white foam (0.88 g, 67% yield). (CH_2Cl_2:MeOH = 92:8 + few drops of triethylamine) R_f = 0.55; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.45 (s, 1H), 7.51–7.46 (m, 3H), 7.44 (s, 1H), 7.37 (d, *J* = 8.8 Hz, 4H), 7.21 (t, *J* = 7.6 Hz, 2H), 7.13–7.07 (m, 2H), 6.96 (t, *J* = 7.8 Hz, 1H), 6.71 (dd, *J* = 1.4 Hz, 8.6 Hz, 4H), 6.43 (t, *J* = 6.6 Hz, 1H), 6.29 (d, *J* = 8.4 Hz, 1H), 4.50–4.47 (m, 1H), 4.12 (dd, *J* = 3.2 Hz, 6.4 Hz, 1H), 3.63 (s, 3H), 3.62 (s, 3H), 3.60 (d, *J* = 3.2, 1H), 3.36 (dd, *J* = 3.6 Hz, 10.6 Hz, 1H), 2.56–2.51 (m, 1H), 2.40–2.33 (m, 1H); ^{13}C NMR (100 MHz**,** $CDCl_3$ **)**: δ (ppm) 160.2, 158.6, 153.6, 149.4, 147.6, 144.6, 135.8, 135.7, 135.2, 130.2, 130.1, 129.0, 128.3, 128.1, 127.1, 124.3, 122.9, 121.0, 113.3, 110.9, 107.0, 105.9, 87.0, 86.5, 85.7, 72.3, 63.4, 55.3, 41.5; HRMS: (m/z): Calculated for $C_{38}H_{34}N_2O_8Na$ $[M+Na]^+$ = 669.2213, found: 669.2213.

Benzofuran modified 2'-deoxyuridine phosphoramidite 4: To a solution of **3** (0.538 g, 0.83 mmol, 1 equiv) in anhydrous dichloromethane (7.5 ml) was added DIPEA (0.73 ml, 4.16 mmol, 5 equiv) and stirred for 10 min. To this solution was slowly added 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.280 ml, 1.25 mmol, 1.5 equiv) and the stirring was continued for another 2 h. The reaction mixture was diluted with dichloromethane (15 ml) and was washed with 5% sodium bicarbonate solution (10 ml) followed by brine (10 ml). The organic extract was dried over sodium sulphate and evaporated, and the residue was purified by silica gel column chromatography (30–55% ethyl acetate in petroleum ether containing

1% triethylamine) to afford the product **4** as a white solid (0.358 g, 51 %). (EtOAc: petroleum ether = 7:3 + few drops of triethylamine) R_f = 0.85; ^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) 8.53 (s, 1H), 7.52 (d, *J* = 7.5 Hz, 2H), 7.47–7.37 (m, 6H), 7.21 (t, *J* = 7.8 Hz, 2H), 7.13–7.05 (m, 2H), 6.90 (t, *J* = 7.8 Hz, 1H), 6.72 (d, *J* = 8.0 Hz, 4H), 6.41 (t, *J* = 6.8 Hz, 1H), 6.08 (d, *J* = 8 Hz, 1H), 4.59–4.56 (m, 1H), 4.24 (br, 1H), 3.66–3.54 (m, 11H), 3.29 (dd, *J* = 2.5 Hz, 10.5 Hz, 1H), 2.59–2.55 (m,1H), 2.41 (t, *J* = 6 Hz, 2H), 2.39–2.34 (m, 1H), 1.20–1.11 (m, 12H); ^{13}C NMR (125 MHz**,** $CDCl_3$**)**: δ (ppm) 160.2, 158.6, 153.6, 149.3, 147.6, 144.6, 135.9, 135.7, 135.2, 130.3, 130.2, 129.0, 128.4, 128.0, 127.1, 124.2, 122.8, 120.9, 117.5, 113.3, 113.3, 110.9, 107.0, 105.9, 86.9, 86.2, 86.2, 85.7, 73.4, 73.4, 62.9, 58.4, 58.3, 55.3, 43.5, 43.4, 40.7, 24.8, 24.7, 24.7, 20.3; ^{31}P NMR (162 MHz, $CDCl_3$): δ (ppm) 149.8, 149.2; HRMS: (m/z): Calculated for $C_{47}H_{51}N_4O_9PNa$ $[M+Na]^+$ = 869.3291, found: 869.3293.

2B.4.4 Photophysical characterization of benzofuran-conjugated 2'-deoxyuridine 2

Steady-state fluorescence in various solvents: Modified nucleoside **2** (5 μM) in water, methanol, acetonitrile and dioxane was excited at lowest energy absorption maximum (322 nm) with excitation and emission slit widths of 3 nm and 5 nm, respectively. In case of ethylene glycol and glycerol the excitation and emission slit widths were maintained at 1 nm and 10 nm, respectively. All solutions contained 0.5% DMSO. Fluorescence experiments were performed in triplicate in a micro fluorescence cell (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 fluorescence spectrophotometer.

Time-resolved fluorescence measurements: Excited state lifetimes of **2** in various solvents were determined using TCSPC fluorescence spectrophotometer (Horiba Jobin Yvon). While the concentration of **2** in water, methanol, ethylene glycol and glycerol was 5 μM, concentration in acetonitrile and dioxane was 250 μM. Nucleoside **2** in water, methanol, ethylene glycol and glycerol was excited using 339 nm LED source (IBH, UK, NanoLED-339L) with a band pass of 4 nm, and fluorescence signal at respective emission maximum was collected. Similarly, **2** in acetonitrile and dioxane was excited using 375 nm diode laser source (IBH, UK, NanoLED-375L) with a band pass of 12 nm. Lifetime measurements were performed in duplicate, and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay kinetics in water and ethylene glycol were found to be monoexponential, whereas in other solvents were found to be biexponential with χ^2 (goodness of fit) values very close to unity.

2B.4.5 Synthesis and purification of modified ONs

Benzofuran modified DNA ONs **5**–**8** and **19** were synthesized on a 1.0 μmole scale (1000 Å CPG solid support). Phosphoramidite **4** was site-specifically incorporated into the ONs by standard DNA ON synthesis protocol with a final trityl-off step (coupling efficiencies were found to be 60–75%). The solid support was treated with 30% aqueous ammonium hydroxide (3 ml) for 15 h at ~50 °C. The aqueous ammonium hydroxide solution was evaporated to dryness on a Speed Vac, and deprotected ON products were purified by polyacrylamide gel electrophoresis using a 20% gel under denaturing conditions. Respective modified ON products were visualized by UV shadowing; product bands were excised from the gel and transferred to a poly-prep column. The gel pieces were crushed using a sterile glass rod and ONs were extracted with sodium acetate buffer (0.3 M, 3 ml) for 12 h. The resulting solutions were filtered and desalted using Sep-Pak classic C18 cartridges.

2B.4.6 MALDI-TOF mass measurements

Molecular weight of DNA ONs **5**–**8** and **19** were determined using Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer. 2 μL of (100 μM) each modified ON was mixed with 1 μL of 100 mM ammonium citrate buffer (pH 9), 2 μL of 100 μM DNA internal standard (10-mer for **5**–**8** and 18-mer for **19**) and 4 μL of saturated 3-hydroxypiccolinic acid (matrix) solution. The samples were desalted with an ion-exchange resin (Dowex 50W-X8, 100-200 mesh, ammonium form) and spotted on the MALDI plate, and were air dried. The resulting spectra were calibrated relative to the internal DNA standards. Sequence of 10-mer DNA internal standard: 5' TGCACGGCGC 3' (mass of +1 and +2 ions are 3029.0 and 1514.5, respectively). Sequence of 18-mer DNA internal standard: 5' TAATACGACTCACTATAG 3' (mass of +1 and +2 ions are 5466.6 and 2733.3, respectively).

2B.4.7 Thermal denaturation experiments

Duplexes were formed by heating a 1:1 mixture (10 μM) of the appropriate ONs in 20 mM cacodylate buffer (pH 7.0, 100 mM NaCl, 0.5 mM EDTA) at 90 °C for 3 min and cooling the solutions slowly to room temperature. Hybridized samples were diluted with cacodylate buffer to give a final concentration of 1 μM duplex in 20 mM cacodylate buffer (pH 7.0, 100 mM NaCl, 0.5 mM EDTA). Thermal denaturation was monitored at 260 nm.

2B.4.8 Steady-state fluorescence of modified DNA ONs

ONs **5**–**8** (10 μM) were annealed to respective complementary custom ONs (**5c**–**8c**, 11 μM) by heating the ONs in 20 mM cacodylate buffer (pH 7.0, 100 mM NaCl, 0.5 mM EDTA) at 90 °C for 3 min. Samples were then cooled slowly to RT, and placed on crushed ice for 2 h. Samples were diluted to give a final concentration of 1 μM (with respect to modified ON) in cacodylate buffer. Following the above procedure, DNA-RNA heteroduplexes (1 μM) were assembled by annealing modified DNA ONs **5**–**8** to respective complementary RNA ONs **13c**–**16c** and abasic-site containing ONs **13a**–**16a**. Fluorescently modified duplexes were excited at 330 nm with an excitation and emission slit width of 3 nm and 10 nm, respectively. Fluorescence experiments were performed in triplicate in a micro fluorescence cuvette at RT. Duplexes **19•17** and **19•18** were also prepared as mentioned above and fluorescence spectra were recorded by exciting the samples at 330 nm with an excitation and emission slit width of 7 nm and 10 nm, respectively.

2B.5 Refrences

[1] a) L. Barbieri, M. G. Battelli, F. Stirpe, *Biochim. Biophys. Acta* **1993**, *1154*, 237–282; b) L. Roncuzzi, A. Gasperi-Campani, *FEBS Lett.* **1996**, *392*, 16–20.

[2] a) Y. Endo, A. Glück, I. G. Wool, *J. Mol. Biol.* **1991**, *221*, 193–207; b) P. Ghosh, J. K. Batra, *Biochem. J.* **2006**, *400*, 99–104.

[3] a) D. Moazed, J. M. Robertson, H. F. Noller, *Nature* **1988**, *334*, 362–364; b) Y. Endo, Y. L. Chan, A. Lin, K. Tsurugi, I. G. Wool, *J. Biol. Chem.* **1988**, *263*, 7917–7920.

[4] a) J. Audi, M. Belson, M. Patel, J. Schier, J. Osterloh, *J. Am. Med. Ass.* **2005**, *294*, 2342–2351; b) Centers for Disease Control and Prevention, *Morb Mortal Wkly Rep*. **2003**, *52*, 1129–1131; c) M.-C. Ho, M. B. Sturm, S. C. Almo, V. L. Schramm, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20276–20281.

[5] a) M. A. Poli, V. R. Rivera, J. F. Hewetson, G. A. Merrill, *Toxicon* **1994**, *32*, 1371–1377; (b) C. Lubelli, A. Chatgilialoglu, A. Bolognesi, P. Strocchi, M. Colombatti, F. Stirpe, *Anal. Biochem*. **2006**, *355*, 102–109; c) R. E. Fulton, H. G. Thompson, *J. Immunoass. Immunochem*. **2007**, *28*, 227–241.

[6] a) M. Zamboni, M. Brigotti, F. Rambelli, L. Montanaro, S. Sperti, *Biochem. J.* **1989**, *259*, 639–643; b) S. Tang, R.-G. Hu, W.-Y. Liu, K.-C. Ruan, *Biol. Chem*. **2000**, *381*, 769–772; c) S.-A. Fredriksson, A. G. Hulst, E. Artursson, A. L. de Jong, C. Nilsson, B. L. M. van Baar, *Anal. Chem.* **2005**, *77*, 1545–1555; d) W. K. Keener, V. R. Rivera, C.

C. Young, M. A. Poli, *Anal. Biochem.* **2006**, *357*, 200–207; e) A. J. Haes, B. C. Giordano, G. E. Collins, *Anal. Chem*. **2006**, *78*, 3758–3764; f) F. Becher, E. Duriez, H. Volland, J. C. Tabet, E. Ezan, *Anal. Chem.* **2007**, *79*, 659–665.

[7] a) S. Roday, M. B. Sturm, D. Blakaj, V. L. Schramm, *J. Biochem. Biophys. Methods* **2008**, *70*, 945–953. b) S. G. Srivatsan, N. J. Greco, Y. Tor, *Angew. Chem. Int. Ed.* **2008**, *47*, 6661–6665; *Angew. Chem.* **2008**, *120*, 6763–6767; c) Q. -Q. Tan, D. -X. Dong, X. -W. Yin, J. Sun, H. -J. Ren, R. -X. Li, *J. Biotech*. **2009**, *139*, 156–162.

[8] A. A. Tanpure, P. Patheja, S. G. Srivatsan, *Chem. Comm.* **2012**, *48*, 501–503.

[9] A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J*. **2011**, *17*, 12820–12827.

[10] C. Reichardt, *Chem. Rev*. **1994**, *94*, 2319–2358.

[11] R. W. Sinkeldam, A. J. Wheat, H. Boyaci, Y. Tor, *ChemPhysChem* **2011**, *12*, 567–570.

[12] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy, 3rd ed.* Springer, New York.

[13] a) M. Kawai, M. J. Lee, K. O. Evans, T. M. Nordlund, *J. Fluoresc*. **2001**, *11*, 23–32; b) J. M. Jean, K. B. Hall, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 37–41.

[14] a) C. A. M. Seidel, A. Schulz, M. H. M. Sauer, *J. Phys. Chem.* **1996,** *100*, 5541–5553; b) S. O. Kelley, J. K. Barton, *Science* **1999**, *283*, 375–381; c) M. Torimura, S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa, R. Kurane, *Anal. Sci.* **2001**, *17*, 155–160.

[15] T. Heinlein, J.-P. Knemeyer, O. Piestert, M. Sauer, *J. Phys. Chem. B* **2003,** *107*, 7957–7964.

[16] a) T. Lindahl, *Nature* **1993**, *362*, 709–715; b) K. S. Gates, T. Nooner, S. Dutta, *Chem. Res. Toxicol.* **2004**, *17*, 839–856; c) H. A. Dahlmann, V. G. Vaidyanathan, S. J. Sturla. *Biochemistry* **2009**, *48*, 9347–9359.

[17] a) L. A. Loeb, B. D. Preston, *Annu. Rev. Genet*. **1986**, *20*, 201–230; b) W. Zhou, P. W. Doetsch, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6601–6605.

[18] a) M. Talpaert-Borl, M. Liuzzi, *Biochim. Biophys. Acta* **1983**, *740*, 410–416; b) H. Ide, K. Akamatsu, Y. Kimura, K. Michiue, K. Makino, A. Asaeda, Y. Takamori, K. Kubo, *Biochemistry* **1993**, *32*, 8276–8283; c) J. Lhomme, J. F. Constant, M. Demeunynck, *Biopolymers* **1999**, *52*, 65–83; d) D. Boturyn, J.-F. Constant, E. Defrancq, J. Lhomme, A. Barbin, C. P. Wild, *Chem. Res. Toxicol*. **1999**, *12*, 476–482. e) H. Atamna, I. Cheung, B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 686–691; f) D. Chakravarti, A. F. Badawi, D. Venugopal, J. L. Meza, L. Z. Crandall, E. G. Rogan, E.

L. Cavalieri, *Mutat. Res*. **2005**, *588*, 158–165; g) E. Fundador, J. Rusling, *Anal. Bioanal. Chem*. **2007**, *387*, 1883–1890.

[19] a) K. Sato, M. M. Greenberg, *J. Am. Chem. Soc.* **2005**, *127*, 2806–2807; b) S. Dhar, T. Kodama, M. M. Greenberg, *J. Am. Chem. Soc*. **2007**, *129*, 8702–8703.

[20] C. Brotschi, A. Häberli, C. J. Leumann, *Angew. Chem. Int. Ed.* **2001**, *40*, 3012–3014; *Angew. Chem.* **2001**, *113*, 3101–3103.

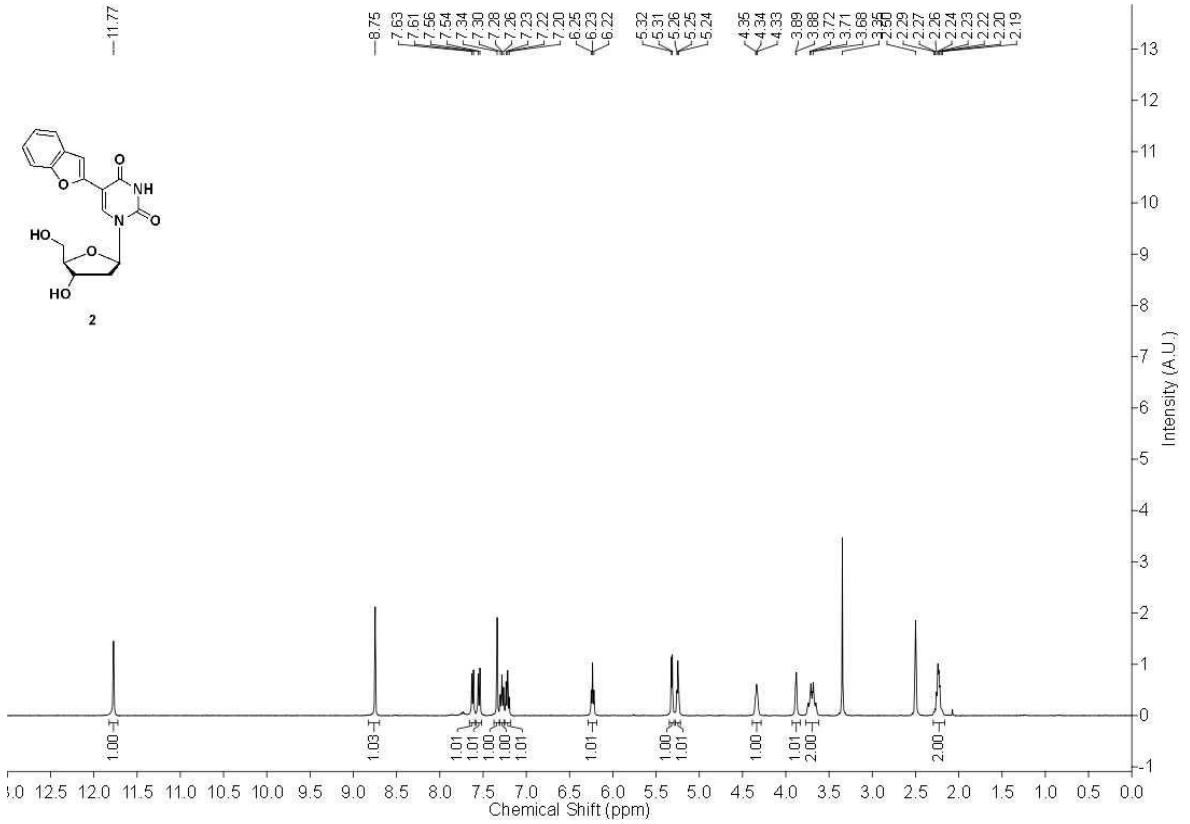
[21] Fluorescence detection of abasic sites in DNA: a) E. L. Rachofsky, E. Seibert, J. T. Stivers, R. Osman, J. B. A. Ross, *Biochemistry* **2001**, *40*, 957–967; b) T. Hirose, T. Ohtani, H. Muramatsu, A. Tanaka, *Photochem. Photobiol*. **2002**, *76*, 123–126; c) K. Yoshimoto, S. Nishizawa, M. Minagawa, N. Teramae, *J. Am. Chem. Soc.* **2003**, *125*, 8982–8983; d) L. Valis, N. Amann, H. A. Wagenknecht, *Org. Biomol. Chem.* **2005**, *3*, 36–38; e) N. J. Greco, Y. Tor, *J. Am. Chem. Soc*. **2005**, *127*, 10784–10785; f) Z. Xu, Y. Sato, S. Nishizawa, N.Teramae, *Chem. Eur. J*. **2009**, *15*, 10375–10378.

[22] E. Shipova, K. S. Gates, *Bioorg. Med. Chem. Lett*. **2005**, *15*, 2111–2113.

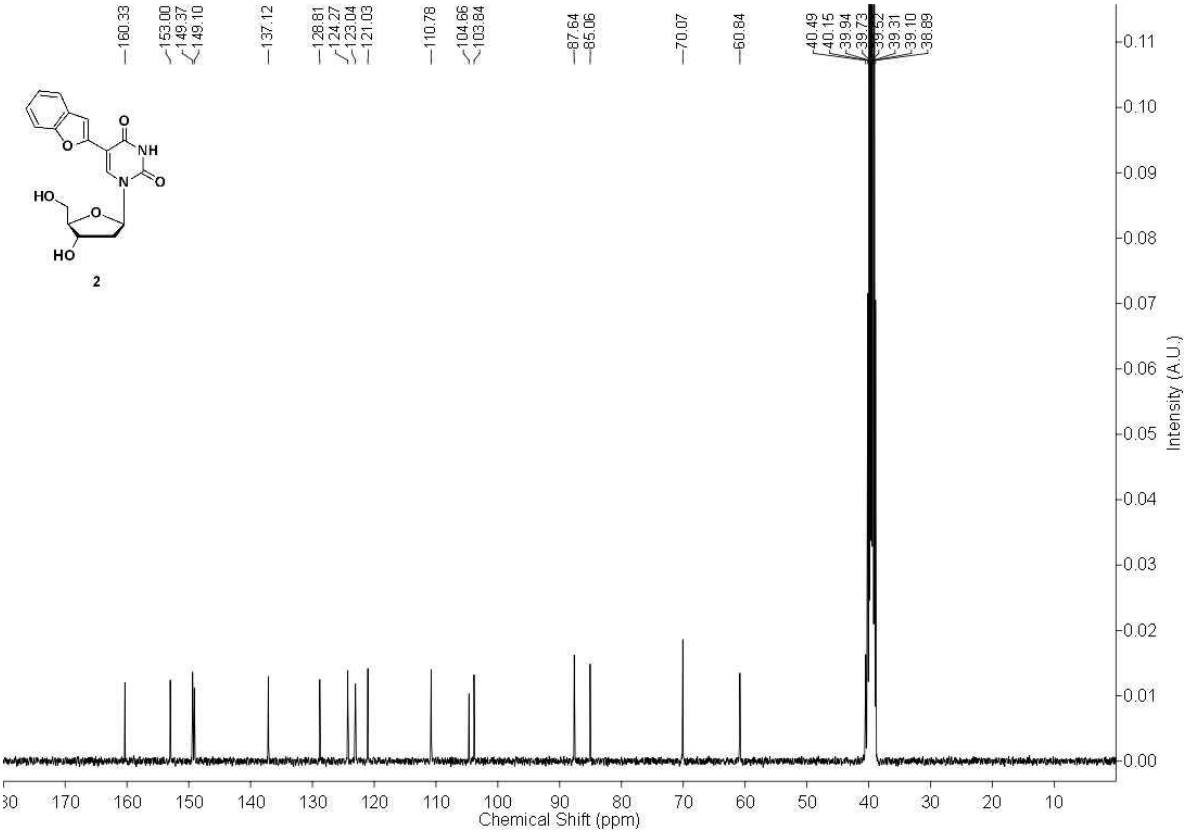
[23] T. Nooner, S. Dutta, K. S. Gates, *Chem. Res. Toxicol*. **2004**, *17*, 942–949.

2B.6 Appendix-II: Characterization data of synthesized compounds

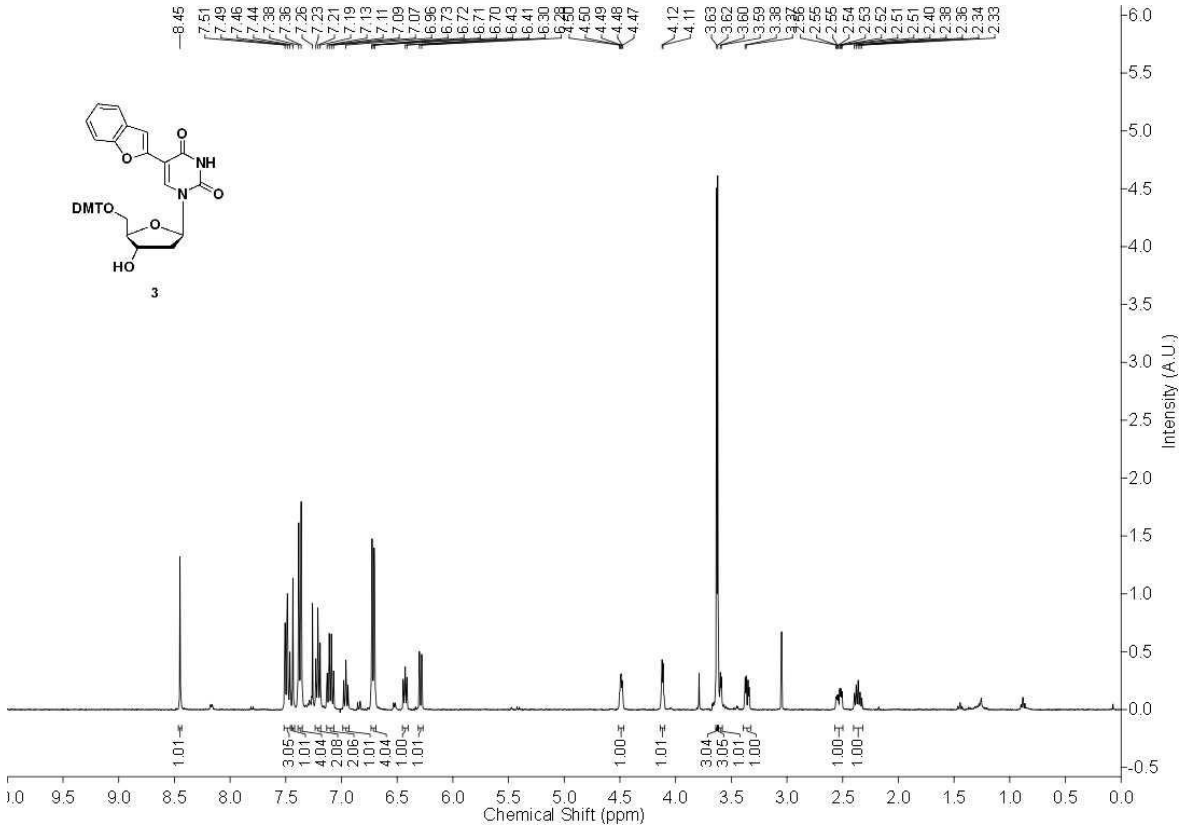
^{1}H NMR of **2** (400 MHz, *d6*-DMSO)



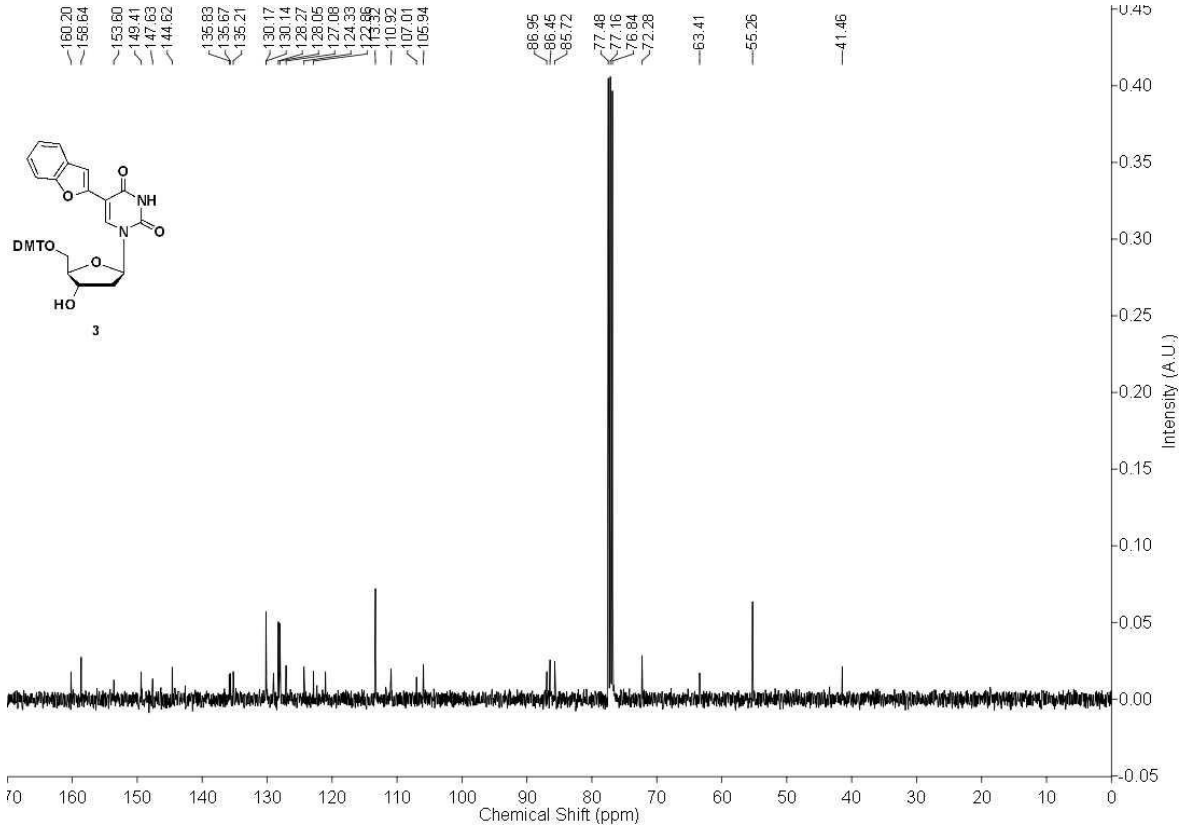
^{13}C NMR of **2** (100 MHz, *d6*-DMSO)



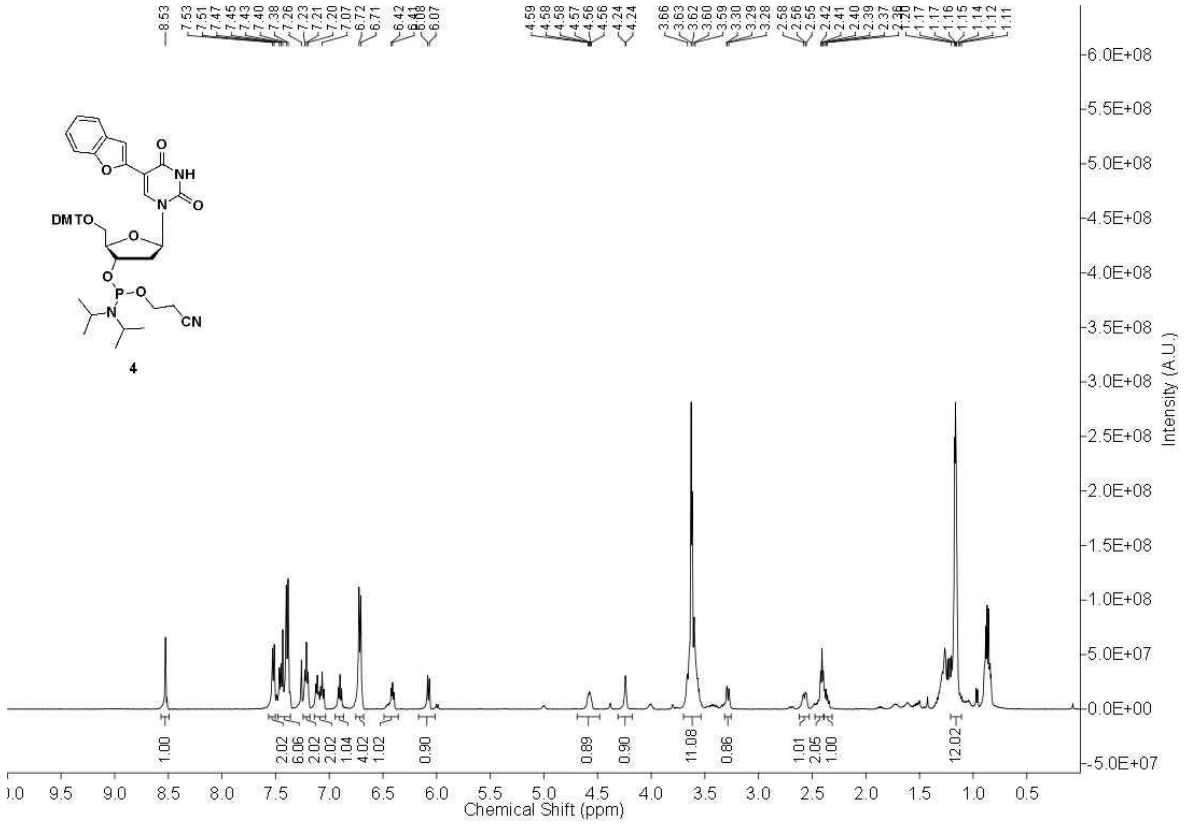
^{1}H NMR of **3** (400 MHz, $CDCl_3$)



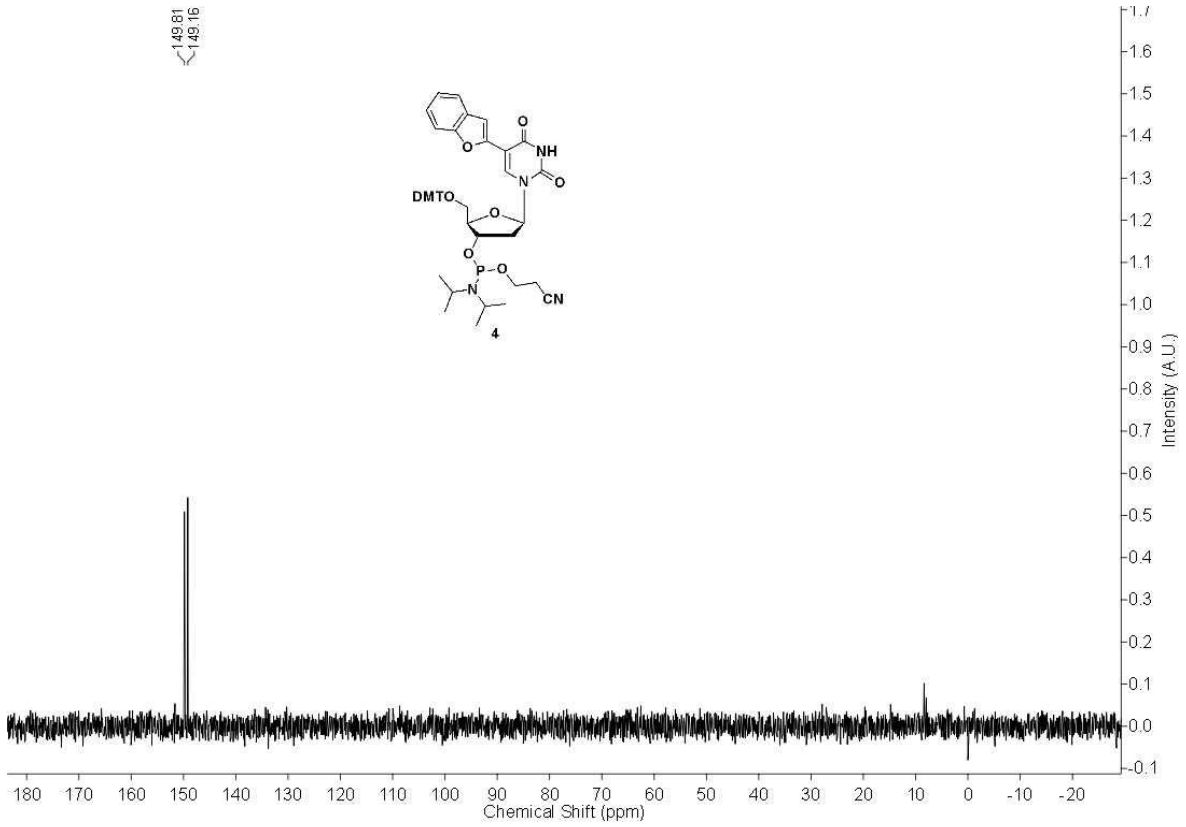
^{13}C NMR of **3** (100 MHz, $CDCl_3$)



^{1}H NMR of **4** (500 MHz, $CDCl_3$)



^{31}P NMR of **4** (162 MHz, $CDCl_3$)



Chapter 3

Fluorescent nucleoside analogs as probes for study the DNA and RNA non-canonical structures

Chapter 3A

5-Benzofuran-modified nucleoside analogs as topology-specific probes for human telomeric DNA and RNA G-quadruplex structures

3A.1 Introduction

Guanine-rich sequences with the potential to form non-canonical four-stranded nucleic acid structures called G-quadruplexes (GQs) have received particular attention in recent years due to their interesting structural features and biological functions.[1,2] These sequences are abundantly found in human genome, particularly in telomeric DNA repeats and certain oncogenic promoter regions, and in untranslated regions of mRNA and telomeric repeat-containing RNA (TERRA).[3] Compelling evidence suggests that these structures play important roles in chromosome maintenance and transcriptional- and translational-control of proliferation-associated genes (e.g., ***c-myc***, ***c-kit***, ***NRAS***, etc.).[4–6] In terms of structure, GQs exhibit a variety of folding topologies *in vitro*, which depend on the sequence and monovalent cations.[7–9] For example, human telomeric (h-Telo) DNA repeats $(TTAGGG)_n$ typically form antiparallel and mixed parallel-antiparallel stranded intramolecular GQ structures in the presence of Na^+ and K^+ ions, respectively.[10–13] However, equivalent RNA sequences form parallel GQ irrespective of Na^+ and K^+ ionic conditions.[14,15]

Owing to the structural diversity of GQs and their role in disease states, G-rich sequences are being rigorously evaluated as novel therapeutic targets for cancer chemotherapy.[1–6] This has led to a flurry of efforts in the development and therapeutic use of small molecule binders, which induce or stabilize GQ structures and modulate their biological function.[16–27] Recent visualization of DNA and RNA GQ structures in cells has further bolstered the interest in this direction.[28-34] Although many of these small molecules bind to GQs strongly, they still lack the required selectivity to differentiate different GQ topologies and nucleic acid type to progress for clinical trials. Furthermore, paucity of efficient chemical probes that can detect different GQ topologies and quantitatively report ligand binding has been a major impediment in the advancement of GQ-directed therapeutic strategies.[35–37] These shortcomings are also evident as no GQ-binding ligand, except for quarfloxin, has been tested in clinical trials.[38]

Fluorescence resonance energy transfer (FRET) approach has been widely applied to study the formation, stability and dynamics of various GQ structures.[39-41] In this approach, an appropriate FRET pair is covalently attached at the 5' and 3' ends of GQ forming oligonucleotide (ON) sequences and change in fluorescence upon folding in the presence of metal ions, complementary ONs or ligands is used as a measure to study various GQ structures.[39-43] Such fluorescently-tagged GQs and aptamers based on GQ structures (e.g., thrombin binding aptamer) have also been elegantly utilized in the fluorometric detection of

metal ions,[44-46] proteins [47-49] and in screening assays to identify quadruplex ligands.[50,51] In a different approach, the efficient excimer emission from pyrene-conjugated thrombin binding DNA aptamer enabled the sensitive detection of K^+ ions.[52] Interestingly, studies using differentially end-labeled GQ ONs, which exhibit quenching in fluorescence due to proximal ligand binding, revealed that GQ ligands bind to a GQ structure at distinct G-tetrads with varied binding affinities.[53] Such labeled GQ ONs could serve as probes to indentify G-tetrad-specific ligands as the physicochemical environment of G-tetrads of a GQ structure is different.

Ligands and metal complexes that show changes in their fluorescence upon binding to GQs have also been utilized as tools to detect and study the recognition properties of GQs.[35−37,54−57] Alternatively, fluorescent purine surrogates (e.g., 6-methylisoxanthopterin and 2-aminopurine) and base-modified 2'-deoxyguanosine analogs containing vinyl, styryl, aryl or heteroaryl group have been incorporated into ONs and utilized in the study of DNA GQs.[58−67] However, many of the fluorescent non-covalent binders and nucleoside analogs are structurally perturbing or poorly discriminate different GQ topologies or exhibit unfavourable photophysical properties (e.g., very low fluorescence due quenching by guanosine and emission in the UV region), which depend on the sequence and hence, hamper their implementation in discovery assays to indentify GQ binders.[35,68] For example, 2-aminopurine (2AP), a fluorescent adenine analogue, has been incorporated into the loop region (TTA) of h-Telo DNA repeat and used as a structure-selective probe for DNA GQs and ligand binding.[60,61] In an analogous study, 2AP and 6MI were incorporated into the loop positions of thrombin binding aptamer. Though these fluorescent analogs exhibited substantial enhancement in fluorescence intensity, which depended on the location of the analogs, they significantly destabilized the quadruplexes.[66] Moreover, their ability to discriminate between DNA and RNA GQs and quantitatively report topology-specific binding of ligands to DNA and RNA GQs has not been explored. Therefore, it is envisioned that the therapeutic evaluation of GQs will greatly benefit from the development of fluorescence-based biophysical tools that (i) can specifically sense the formation of GQs with enhancement in fluorescence intensity, (ii) can distinguish different GQs based on topology and nucleic acid type, and (iii) are compatible to screening formats for the identification of small molecules that selectively bind to DNA and RNA GQs.

Here, we describe examples of minimally perturbing environment-sensitive fluorescent pyrimidine nucleoside analogs, based on a 5-(benzofuran-2-yl)uracil core, which

when incorporated into one of the loop regions of human telomeric DNA and RNA ON sequences, clearly signal the formation of different DNA and RNA GQ structures with significant enhancement in fluorescence intensity. These emissive nucleosides selectively detect GQ structures from corresponding duplexes, and distinguish GQ topologies based on (i) ionic conditions (Na^+ vs. K^+) and (ii) nucleic acid type (DNA vs. RNA). Furthermore, we utilized the conformation-sensitivity of the probes in developing fluorescence binding assays to quantitatively determine topology-specific binding of ligands to DNA and RNA GQs. Our results demonstrate that these simple GQ sensors could provide an alternative platform for the development of screening assays to identify topology- and nucleic acid-specific GQ binders of therapeutic potential.

3A.2 Results and Discussion

3A.2.1 Fluorescence detection of DNA GQs

In previous chapter, we have introduced base-modified fluorescent pyrimidine analogs derived by attaching benzofuran moiety at the 5-position of uracil.[71−73] 5-Benzofuran-modified 2'-deoxyuridine (**1**) and uridine (**2**) analogs are reasonably emissive and their fluorescence properties are highly sensitive to changes in their microenvironment. Unlike the majority of fluorophores, these nucleoside**s**, when incorporated into ONs and hybridized to complementary ONs, retain appreciable fluorescence efficiency and report changes in flanking bases and base pair mismatches via changes in their fluorescence properties. These key observations and 3D structure of h-Telo DNA repeats in which the loop region (TTA) undergoes substantial conformational change upon binding to a small molecule ligand inspired us to develop the benzofuran-modified nucleoside analogue as a probe to detect GQ structures and GQ binders (Figure 1).

Human telomeric DNA repeats, which endcap eukaryotic chromosomes play an important role in chromosome maintenance. [74,75] While aberrant shortening of telomeric repeats during cell division can result in genomic instability, the preservation of telomere length by telomerase activity in tumour cells has been implicated in carcinogenesis.[76,77] Therefore, we selected h-Telo DNA repeat $d[AG_3(T_2AG_3)_3]$, which has received much of the attention, as the first test system. Phosphoramidite **3** was used in the synthesis of fluorescent h-Telo DNA ONs **5** and **6** in which the loop residues dT_{11} and dT_{12}, respectively, were replaced with 5-benzofuran-modified 2'-deoxyuridine **1** (Figure 2). We purposely chose to modify the loop dT residues as modifications on guanosine could affect the efficiency of

formation of GQ.[67,78] The integrity of full-length modified h-Telo DNA ONs **5** and **6** was confirmed by mass analysis (Table 1 and Figure 3).

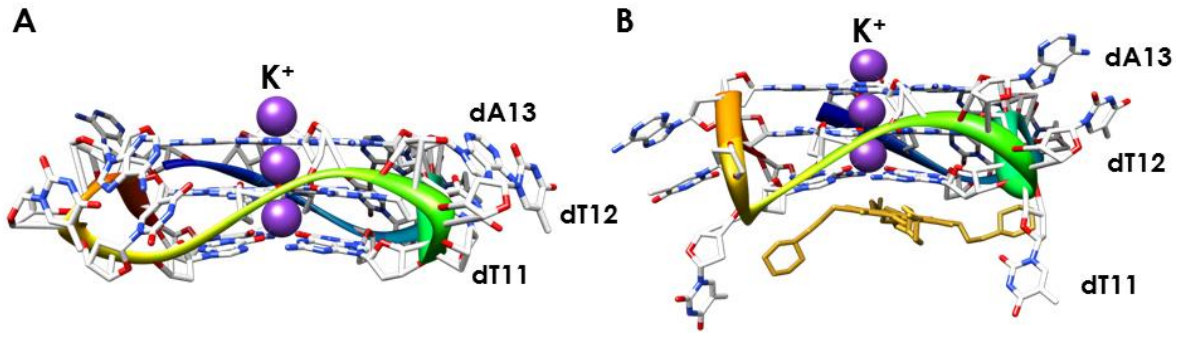
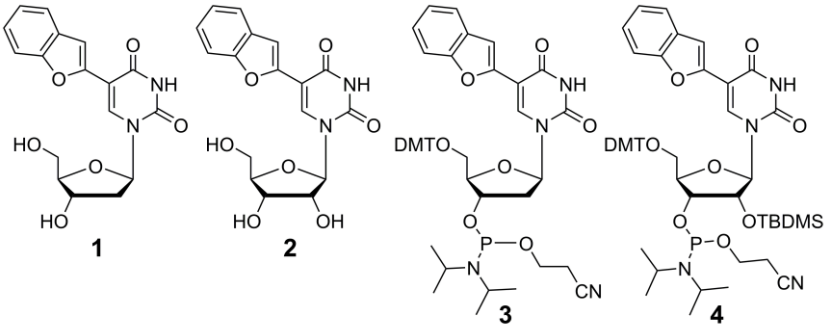


Figure 1. Crystal structure of h-Telo DNA repeat d[$AG_3(T_2AG_3)_3$] in the presence of K^+ ions. The conformation of loop residues dT_{11}, dT_{12} and dA_{13} in the absence (**A**) and presence (**B**) of a ligand (tetrasubstituted naphthalene diimide derivative) is shown. [11,23] The PDB accession numbers are 1KF1 and 4DA3, respectively. We envisaged that pyrimidine loop residues would be potential sites for introducing conformation-sensitive probes.



5 5' d(AGGGTTAGGG**1**TAGGGTTAGGG) 3'
6 5' d(AGGGTTAGGGT**1**AGGGTTAGGG) 3'
7 5' d(AGGGTTAGGGTTAGGGTTAGGG) 3'
8 5' d(CCCTAACCCTAACCCTAACCCT) 3'
9 5' r(UUAGGGUUAGGG**2**UAGGGUUAGGG) 3'
10 5' r(UUAGGGUUAGGGUUAGGGUUAGGG) 3'
11 5' d(AGGGTTAGGG**1**TAGGGTTAGGGTTAGGG**1**TAGGGTTAGGG) 3'

Figure 2. Chemical structure of 5-benzofuran-modified 2'-deoxyuridine (1) and uridine (2) analogs, and corresponding phosphoramidite substrates 3 and 4 used in the synthesis of GQ-forming DNA and RNA ONs (5, 6, 9 and 11). 5 and 6 are fluorescent h-Telo DNA ONs containing the modification at position 11 and 12, respectively. ON 7 is a control unmodified h-Telo DNA and 8 is complementary to h-Telo DNA 5–7. 9 and 10 are modified and control unmodified TERRA ONs, respectively. Sequence of doubly modified longer h-Telo DNA repeat 11 used in this study.

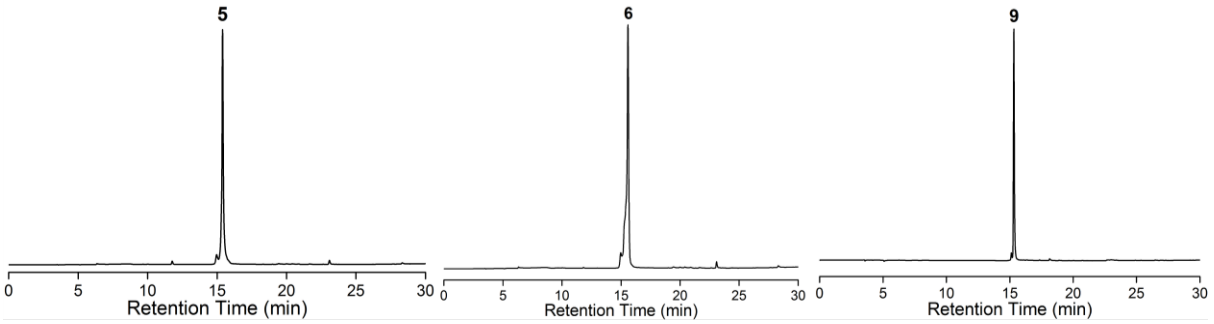


Figure 3. HPLC chromatograms of PAGE purified fluorescent ONs **5**, **6** and **9** at 260 nm. Mobile phase A = 100 mM triethylammonium acetate buffer (pH 7.5), mobile phase B = acetonitrile. Flow rate = 1 mL/min. Gradient = 0–10% B in 10 min and 10–100% B in 20 min. HPLC analysis was performed using Luna C18 column (250 x 4.6 mm, 5 micron).

Table1. ε_{260} and MALDI-TOF mass data of modified h-Telo DNA and TERRA ONs.

Modified ON	ε_{260} ($M^{-1}cm^{-1}$)	Calcd. mass	Observed mass
5	232713	7068.6	7068.5
6	232713	7068.6	7069.3
9	263253	7962.8	7963.1
11	420226	12898.3	12898.6

Steady-state fluorescence analysis of ONs **5** and **6** in a buffer solution containing KCl or NaCl (100 mM) was performed by exciting the samples at 330 nm. Telomeric DNA **5**, which could form hybrid-type of mixed parallel-antiparallel stranded GQ in K^+, displayed a significantly higher fluorescence intensity (~4 fold at λ_{em} = 435 nm) as compared to corresponding perfect duplex **5•8** (Figure 4A). Remarkably, **5** in the presence of Na^+ ions, which is known to induce antiparallel GQ structure exhibited a nearly 9-fold enhancement in fluorescence intensity as compared to duplex **5•8** with no apparent change in emission maximum. Further, time-resolved fluorescence measurements revealed distinct decay kinetics for different GQ topologies. The antiparallel GQ structure of **5** exhibited the highest lifetime of 3.00 ns followed by mixed-type GQ (1.40 ns) and duplex **5•8** (0.64 ns and 0.66 ns) in the presence of K^+ and Na^+ ions (Figure 4B, Table 2). This trend in lifetime was found to be consistent with the trend in emission intensity observed in steady-state experiments. The marked difference in emission intensity and lifetime exhibited by h-Telo DNA ON **5** in K^+ and Na^+ is likely due to distinctly different conformation of fluorescent nucleosides in these two different topologies.

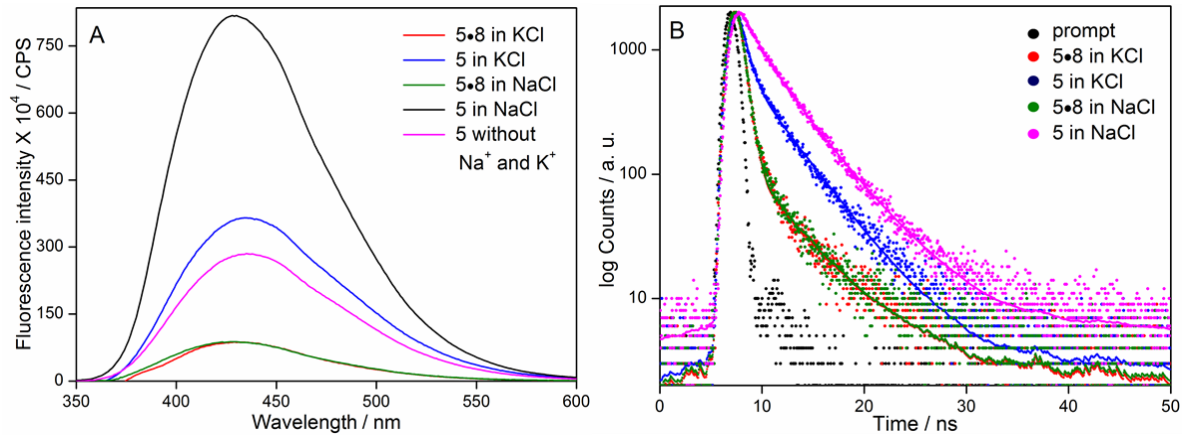


Figure 4. (**A**) Steady-state and (**B**) time-resolved fluorescence spectra of h-Telo DNA **5** and corresponding duplex **5•8** in the presence of K^+ or Na^+ ions. Laser profile is shown in gray (prompt). Fluorescence spectrum of ON **5** (magenta) in buffer without added Na^+ and K^+ ions is also shown. ONs (0.25 μM) were excited at 330 nm with excitation and emission slit widths of 8 and 10 nm, respectively.

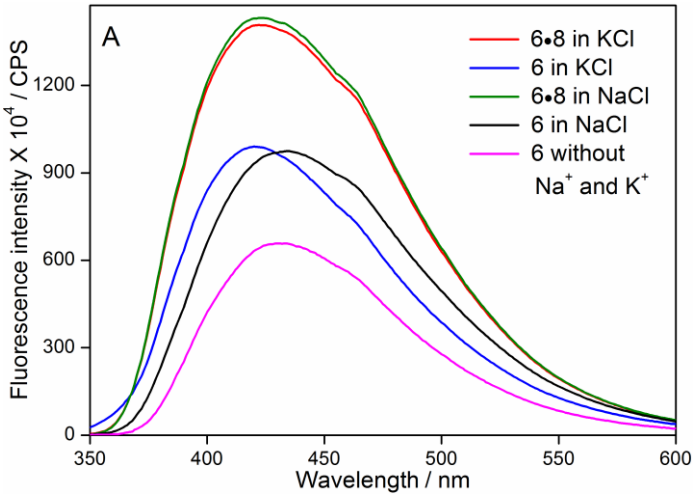


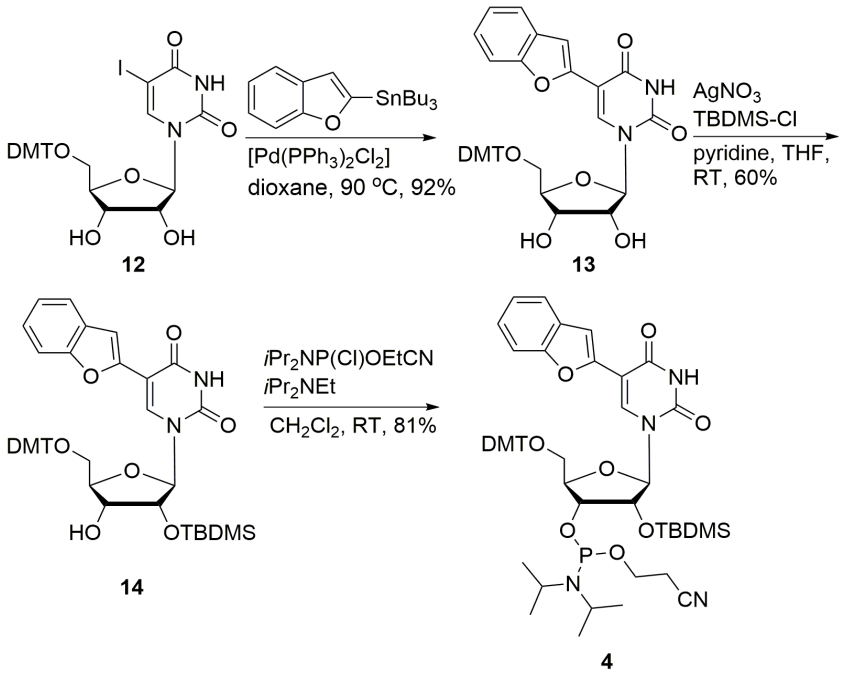
Figure 5. Fluorescence spectra of h-Telo DNA **6** and corresponding duplex **6•8** in the presence of K^+ or Na^+ ions. Fluorescence spectrum of ON **6** (magenta) in buffer without added Na^+ and K^+ ions is also shown. ONs (0.25 μM) were excited at 330 nm with excitation and emission slit widths of 8 and 10 nm, respectively.

Interestingly, the efficiency of sensing of different GQ topologies by emissive nucleoside was found to depend on the position of modification. ON 6, containing the emissive nucleoside 1 at position 12, in the presence of Na^+ or K^+ ions displayed reasonable quenching in fluorescence intensity as compared to duplex 6•8 (Figure 5). Albeit small difference in emission maximum, ON 6 failed to distinguish between different GQ structures in Na^+ and K^+ ionic conditions. Based on emission maximum, the fluorescent nucleoside 1 in h-Telo DNA ON 5 in the presence of K^+ and Na^+ ions (~435 nm) is slightly less solvent

exposed than the free nucleoside (447 nm, Table 1). While the emissive nucleoside in ON 6 in the presence of Na^+ ions (435 nm) is solvated similar to ON 5 in the presence of K^+ and Na^+ ions, it is even more less exposed to solvent in ON 6 in the presence of K^+ ions (420 nm).

3A.2.2 Fluorescence detection of RNA GQs

While most efforts in the study of GQ structures have been dedicated toward developing probes for DNA GQs, probe development for RNA GQs has received less attention. Therefore, we wanted to explore the efficacy of our emissive nucleoside in detecting the formation of RNA GQ structures. For this purpose we chose TERRA composed of extended tandem repeats of $(U_2AG_3)_n$, which has been considered to play important role in telomere structure maintenance and replication.[79,80] RNA GQs formed *in vitro* under near physiological conditions are thermodynamically more stable than their DNA counterparts, and have been shown to form parallel-stranded GQ structure in both Na^+ and K^+ ionic conditions.[14,15] The counterpart of h-Telo DNA **5**, TERRA ON **9** $(U_2AG_3)_4$, containing benzofuran-modified uridine analogue **2** at loop residue was synthesized by using phosphoramidite **4** (Scheme 1, Figure 2, Table 1).



Scheme 1. Synthesis of benzofuran-modified uridine phosphoramidite substrate **4** for the solid-phase synthesis of RNA ON **9**. DMT = 4,4'-dimethoxytrityl, TBDMS = 2'-*O*-*tert*-butyldimethylsilyl.

Upon excitation, RNA **9** in the presence of K^+ and Na^+ ions displayed a significant and comparable enhancement in fluorescence intensity (~5-fold) as compared to corresponding duplex **9•8** (Figure 6A). Excited-state decay kinetics analysis of **9** also revealed a similar lifetime of 2.80 ns and 2.90 ns in K^+ and Na^+ ions, respectively, which was nearly four times higher than that of duplex (Figure 6B, Table 2). These results clearly indicate that the conformation and microenvironment of the emissive nucleoside incorporated into **9** in K^+ and Na^+ is similar and consistent with the formation of parallel GQ structure expected for TERRA ON irrespective of ionic conditions.[14,15] It is worth mentioning here that a fluorescent adenine analogue, 2AP, when incorporated into the loop region of h-Telo DNA repeat exhibits emission maximum in the UV region (λ_{em} = 370 nm) and much lower quantum yield (0.06) as compared to benzofuran-modified h-Telo DNA and TERRA GQs (Table 2).[81] This is because, guanine is known to effectively quench the fluorescence of 2AP in ONs by electron transfer mechanism.[68]

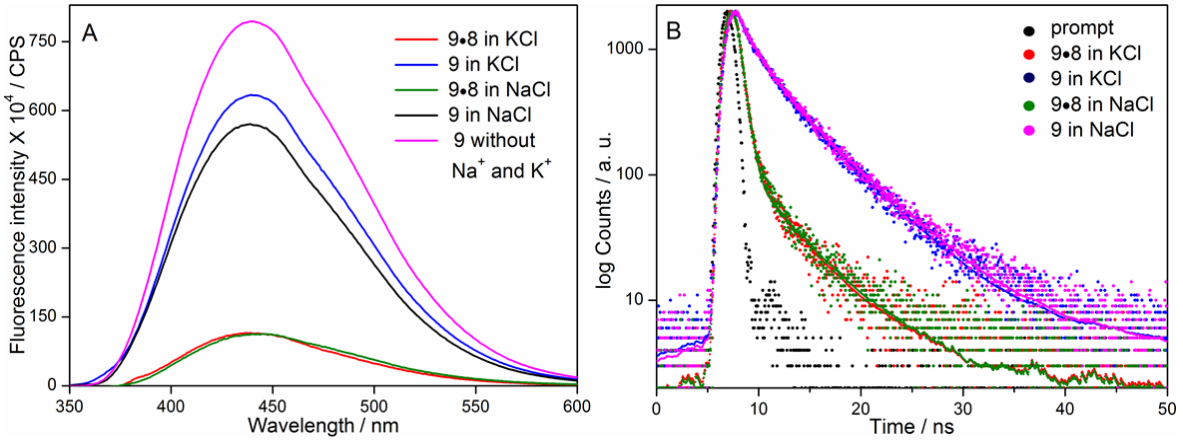


Figure 6. (**A**) Steady-state and (**B**) time-resolved fluorescence spectra of TERRA **9** and corresponding duplex **9•8** in the presence of K^+ or Na^+ ions. Laser profile is shown in gray (prompt). Fluorescence spectrum of ON **9** (magenta) in buffer without added Na^+ and K^+ ions is also shown. ONs (0.25 μM) were excited at 330 nm with excitation and emission slit widths of 8 and 10 nm, respectively.

We believe that distinct fluorescence properties such as emission maximum, quantum yield and lifetime exhibited by DNA and RNA GQ structures of **5** and **9**, and corresponding duplexes in different ionic conditions are due to distinct conformation and microenvironment of emissive nucleosides in these constructs. As a consequence, the observed enhancement in fluorescence intensity displayed by GQs as compared to duplexes can be possibly due to a combination of the following reasons: rigidification of the fluorophore, reduced electron transfer process and stacking interaction between the emissive nucleobase and neighbouring

bases in GQs as compared to base-paired emissive nucleobase in duplexes.[68,82] Taken together, these results clearly demonstrate the ability of benzofuran-modified nucleosides to distinguish different GQ topologies based on ionic conditions and nucleic acid type.

Table 2. Quantum yield and excited-state lifetime of modified h-Telo DNA **5** and TERRA **9** GQs and corresponding duplexes in different ionic conditions.

Sample	λ_{em} (nm)	Φ^a	$\tau_1{}^b$ (ns)	$\tau_2{}^b$ (ns)	$\tau_{Ave}{}^a$ (ns)
1 in water (53)	446	0.19	2.38	-	2.38
2 in water (52)	447	0.21	2.55	-	2.55
5 in KCl	435	0.11	0.74 (0.81)	4.03 (0.19)	1.40
5•8 in KCl	430	0.03	0.56 (0.98)	5.10 (0.02)	0.64
5 in NaCl	430	0.25	1.89 (0.58)	4.50 (0.42)	3.00
5•8 in NaCl	430	0.03	0.57 (0.98)	5.27 (0.02)	0.66
9 in KCl	440	0.19	1.53 (0.66)	5.37 (0.34)	2.80
9•8 in KCl	440	0.04	0.56 (0.98)	5.42 (0.02)	0.65
9 in NaCl	440	0.18	1.67 (0.67)	5.67 (0.33)	2.90
9•8 in NaCl	440	0.04	0.58 (0.98)	5.30 (0.02)	0.67

[a] Standard deviations for quantum yield (Φ) and average lifetime (τ_{av}) are ≤ 0.001 and ≤ 0.05 ns, respectively. [b] Relative amplitude is given in parenthesis.

3A.2.3 Circular dichroism and thermal melting studies

Benzofuran modification in h-Telo DNA **5** and TERRA **9** could potentially affect the formation and stability of GQ structures. Circular dichroism (CD) analysis of modified ONs **5** and **9**, and corresponding unmodified ONs **7** and **10**, respectively, in the presence of Na^+ or K^+ ions confirmed the formation of respective topologies thereby indicating that the modification did not hamper the formation of GQs (Figures 7). In the presence of K^+ ions h-Telo DNA ONs **5** and **7** exhibit a positive peak at ~290 nm, a smaller shoulder at ~270 nm and a smaller minimum at ~ 235 nm, which closely resemble the CD pattern of hybrid-type of mixed parallel-antiparallel stranded GQ structure. Such a CD profile has been observed for the same sequence in earlier reports.[12,83] In the presence of Na^+ ions h-Telo DNA ONs **5** and **7** exhibit CD pattern (positive peak at ~290 nm and strong negative peak at ~265 nm) characteristic of antiparallel stranded GQ structure.[10,84] On the other hand, irrespective of ionic conditions both modified and unmodified TERRA ONs show similar CD profiles characteristic of parallel stranded GQ structure.[14,84] Furthermore, UV-thermal melting

analysis of modified and unmodified GQ-forming ONs in different ionic conditions gave similar T_m values ($\Delta T_m \leq 2°C$) suggesting that the benzofuran modification had only a minor impact on GQ stability (Figure 8, Table 3)[85]

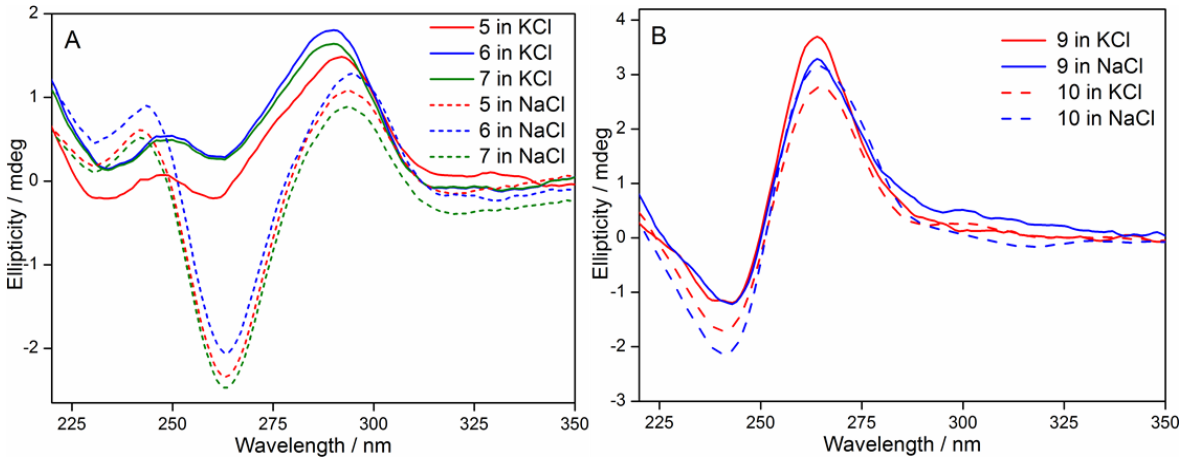


Figure 7. CD spectra of (A) fluorescently modified h-Telo DNA **5**, h-Telo DNA **6** and control unmodified h-Telo DNA **7** (B) fluorescently modified TERRA **9** and control unmodified TERRA **10** in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl. Both unmodified and modified ONs show similar CD profiles in respective ionic conditions.

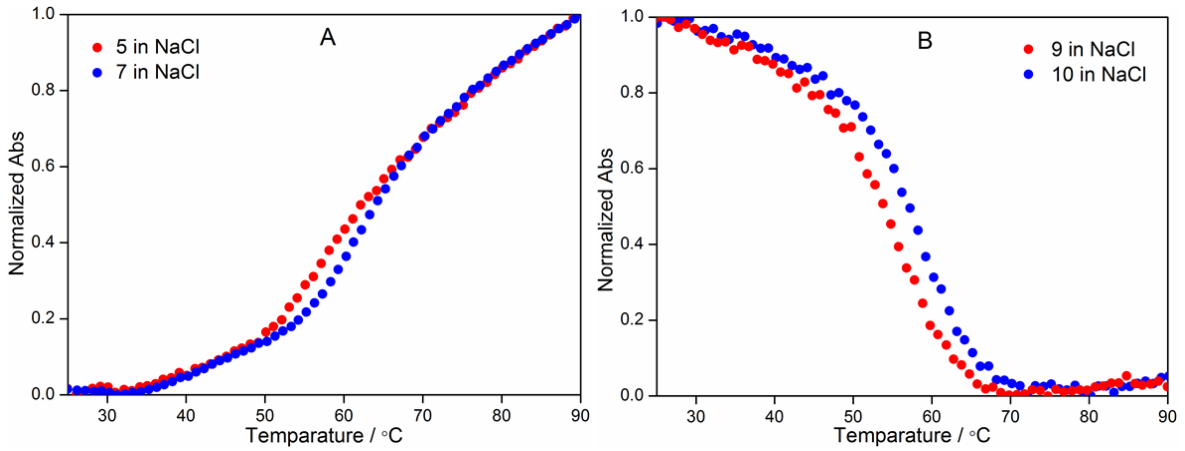


Figure 8. Representative UV-thermal melting profiles of fluorescently modified **(A)** h-Telo DNA **5** and control unmodified h-Telo DNA **7** in 10 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl at 260 nm. **(B)** TERRA **9** and control unmodified TERRA **10** in 10 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl at 295 nm.

Table 3. T_m values of fluorescently modified and control unmodified GQs in the presence of KCl and NaCl

Fluorescent GQ	T_m (°C)	Control unmodified GQ	T_m (°C)
5 in KCl	67 ± 0.5	**7** in KCl	67 ± 0.6
5 in NaCl	61 ± 0.4	**7** in NaCl	62 ± 0.7
9 in KCl	77 ± 0.9	**10** in KCl	78 ± 0.8
9 in NaCl	58 ± 0.9	**10** in NaCl	60 ± 0.8

3A.2.4 Topology-specific binding of small-molecules to DNA and RNA GQ structures

Next we sought to explore the compatibility of these GQ sensors in estimating topology-specific binding of small-molecules to different GQ structures. For this purpose, we chose two known GQ binders, pyridostatin (PDS) and BRACO19, which have been used as tools for biophysical and therapeutic analysis of GQ structures (Figure 9).[86,87] Fluorescent h-Telo DNA **5** was assembled into mixed-type and antiparallel GQs in buffers containing KCl and NaCl, respectively. GQs of **5** were excited at 330 nm and changes in fluorescence intensity upon addition of increasing concentrations of a ligand were monitored. Addition of PDS to mixed-type GQ resulted in a nearly 4-fold quenching in fluorescence intensity at the saturation concentration of PDS (2.5 μM, Figure 10A). The saturation binding isotherm yielded an apparent K_d of 919 ± 7 nM, which is comparable to literature report (Figure 10B).[88] Interestingly, binding of PDS to antiparallel GQ in Na^+ resulted in a concentration-dependent decrease in fluorescence intensity, which was significantly pronounced as compared to in the presence of K^+ (8-fold at PDS = 1.5 μM, Figure 10C, 10D). The K_d value (440 ± 21 nM) indicated that PDS has higher binding affinity for antiparallel h-Telo DNA GQ as compared to mixed-type h-Telo DNA GQ structure (Table 4). The fluorescent nucleoside also reported the binding of BRACO19 to both mixed-type and antiparallel GQs of **5** with significant quenching in fluorescence intensity (Figure 11). BRACO19 exhibited opposite binding affinities for GQs as compared to PDS (Table 4). While BRACO19 binding to mixed-type h-Telo DNA GQ was found to be stronger than PDS, binding of BRACO19 to antiparallel DNA GQ was weaker compared to PDS.

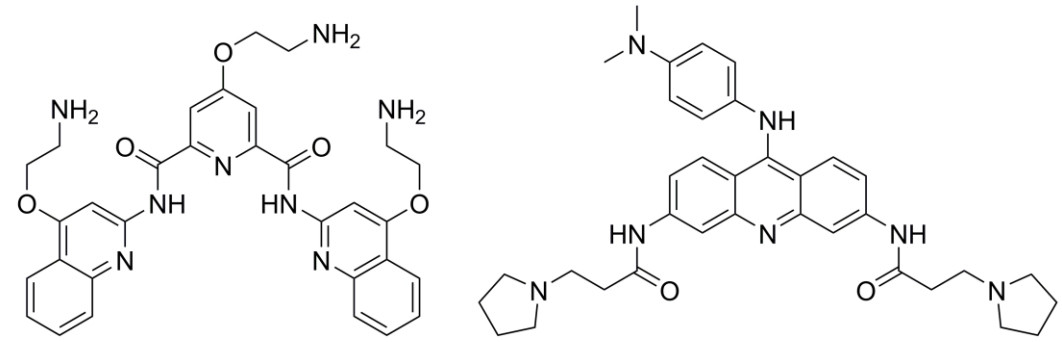


Figure 9. Chemical structure of GQ binders, pyridostatin (PDS) and BRACO19.

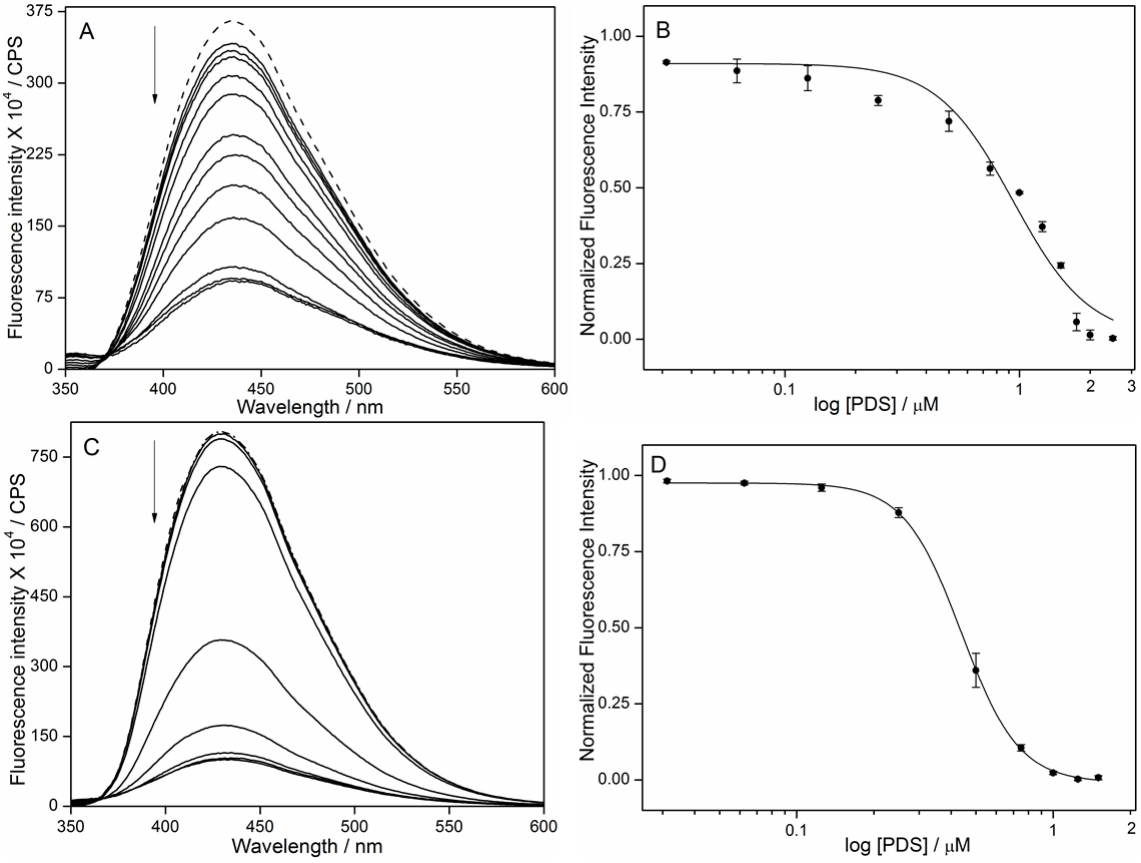


Figure 10. Emission spectra (solid lines) of mixed-type (**A**) and antiparallel (**C**) GQs of h-Telo DNA **5** (0.25 μM) in the presence of K^+ and Na^+ ions, respectively, with increasing concentrations of PDS. Dashed lines represent emission profile of GQs in the absence of PDS. Curve fit for the binding of PDS to h-Telo DNA **5** in the presence of (**B**) KCl and (**D**) NaCl. Normalized fluorescence intensity at λ_{em} = 430 nm is plotted against log [PDS].

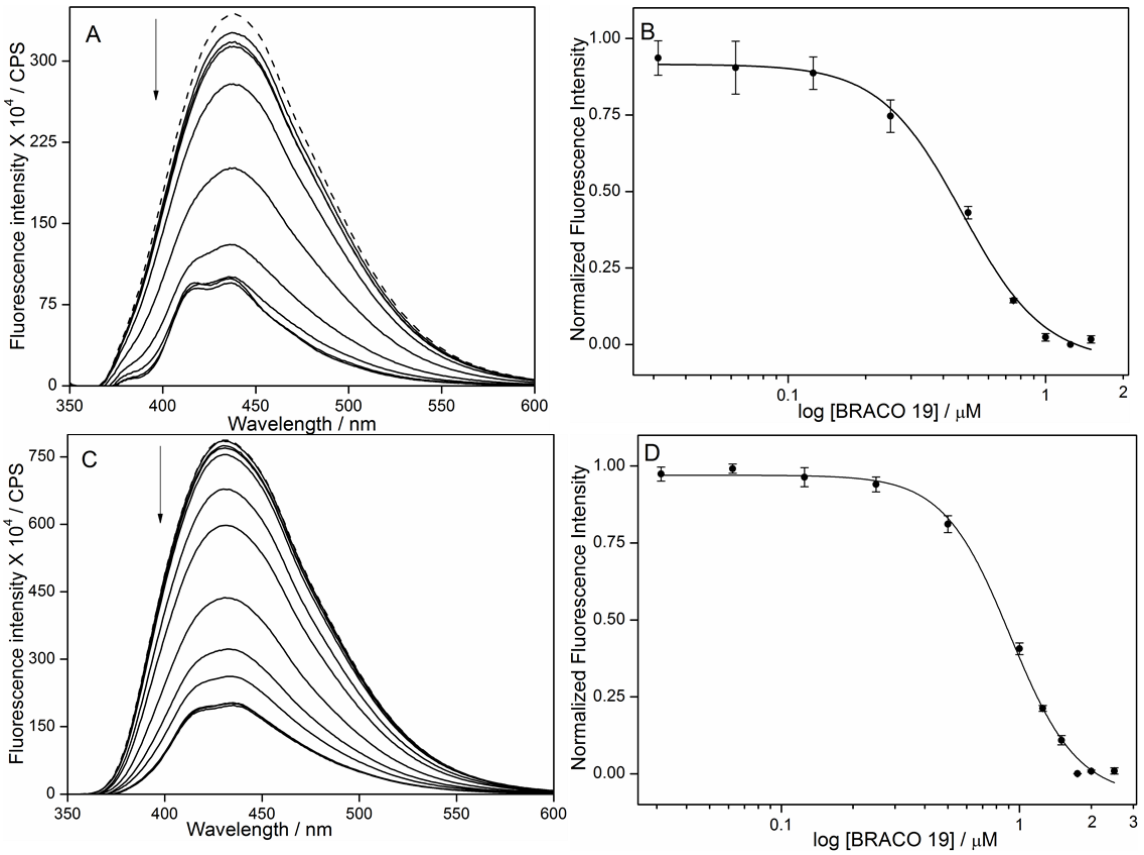


Figure 11. Emission spectra (solid lines) of mixed-type (**A**) and antiparallel (**C**) GQs of h-Telo DNA **5** (0.25 μM) in the presence of K^+ and Na^+ ions, respectively, with increasing concentrations of BRACO19. Dashed lines represent emission profile of GQs in the absence of BRACO19. Curve fit for the binding of BRACO19 to h-Telo DNA **5** in the presence of (**B**) KCl and (**D**) NaCl. Normalized fluorescence intensity at λ_{em} = 430 nm is plotted against log [BRACO19].

High-resolution structures of TERRA have revealed distinct features of RNA GQs involving 2'-hydroxyl group and loop residues that are not found in DNA GQs.[15] These studies suggest the possibility of developing small-molecule ligands that can selectively bind GQs based on nucleic acid type.[89] TERRA ON **9**, which forms parallel GQ structure in both K^+ and Na^+ ionic conditions, was subjected to binding studies. Concentration-dependent quenching in fluorescence intensity signalled the binding of PDS and BRACO19 to **9** (Figure 12, 13). The ligands exhibited significantly higher but similar binding affinities for TERRA as compared to h-Telo DNA in the presence of K^+ and Na^+ ions (Table 4). These results are consistent with the ability of TERRA ON **9** to form one type of stable GQ structure (i.e., parallel) as compared to corresponding DNA sequence, irrespective of ionic conditions. Importantly, control binding experiments with nucleoside **1** and duplexes of h-Telo DNA (**5•8**) and TERRA (**9•8**) resulted only in minor changes in fluorescence intensity indicating that the nucleoside incorporated into ONs **5** and **9** signalled only the specific binding of

ligands to GQ structures (Figure 14). In the absence of structural analysis, fluorescence quenching displayed by GQs of **5** and **9** is possibly due to the conformation attained by emissive nucleosides upon ligand binding, which is less rigid and appropriately positioned for nonradiative interactions with neighbouring bases and or ligands.[68,90,91] Further, differential binding affinities exhibited by PDS and BRACO19 for DNA and RNA GQs could be ascribed to the differences in chemical environment and conformation of G-tetrads of GQ assemblies.[53] Collectively, these observations highlight the potential of benzofuran-modified nucleoside analogs as probes for quantitative estimation of structure-specific binding of ligands to GQ structures.

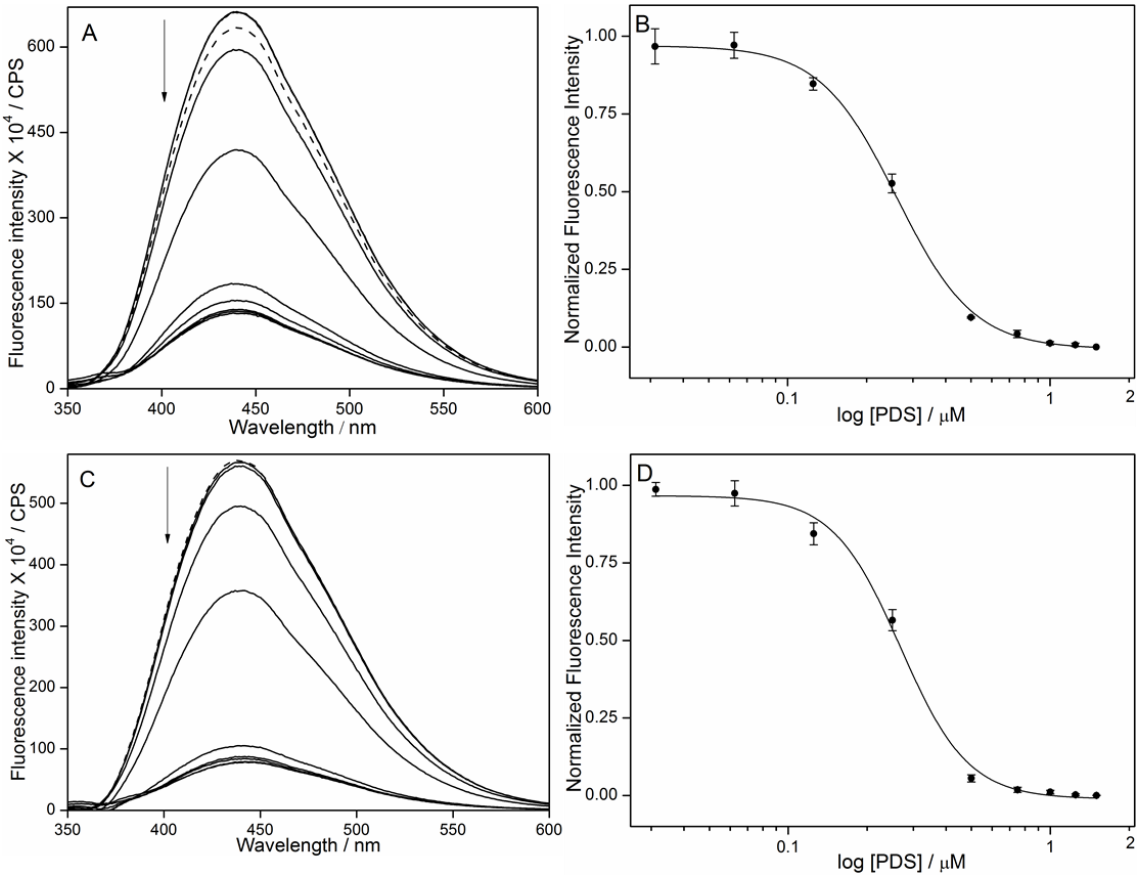


Figure 12. Emission spectra (solid lines) of parallel GQ of TERRA **9** (0.25 μM) in the presence of K^+ (**A**) and Na^+ (**C**) ions, respectively, with increasing concentrations of PDS. Dashed lines represent emission profile of GQs in the absence of PDS. Curve fit for the binding of PDS to TERRA **9** in the presence of (**B**) KCl and (**D**) NaCl. Normalized fluorescence intensity at λ_{em} = 430 nm is plotted against log [PDS].

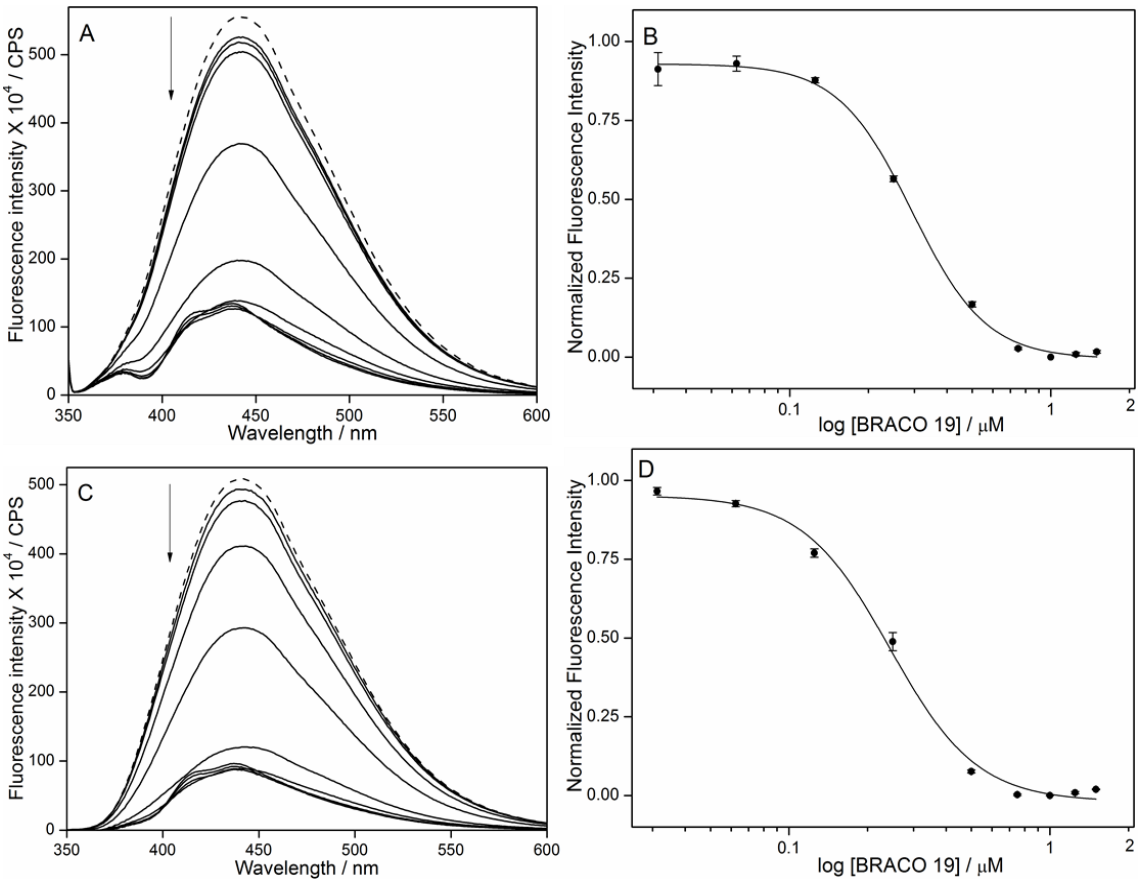


Figure 13. Emission spectra (solid lines) of parallel GQ of TERRA **9** (0.25 μM) in the presence of K^+ (**A**) and Na^+ (**C**) ions, respectively, with increasing concentrations of BRACO19. Dashed lines represent emission profile of GQs in the absence of BRACO19. Curve fit for the binding of BRACO19 to TERRA **9** in the presence of (**B**) KCl and (**D**) NaCl. Normalized fluorescence intensity at λ_{em} = 440 nm is plotted against log [BRACO19].

Table 4. Binding constant (K_d) for PDS and BRACO19 binding to h-Telo DNA **5** and TERRA **9**.

ON	PDS (nM)		BRACO19 (nM)	
	in KCl	in NaCl	in KCl	in NaCl
5	919 ± 07	440 ± 21	476 ± 25	938 ± 29
9	260 ± 10	268 ± 12	294 ± 04	244 ± 11
11	n.d.	539 ± 19	483 ± 23	462 ± 17

n.d. = not determined (although addition of PDS to **11** in the presence of KCl resulted in quenching, reliable binding constant could not be determined).

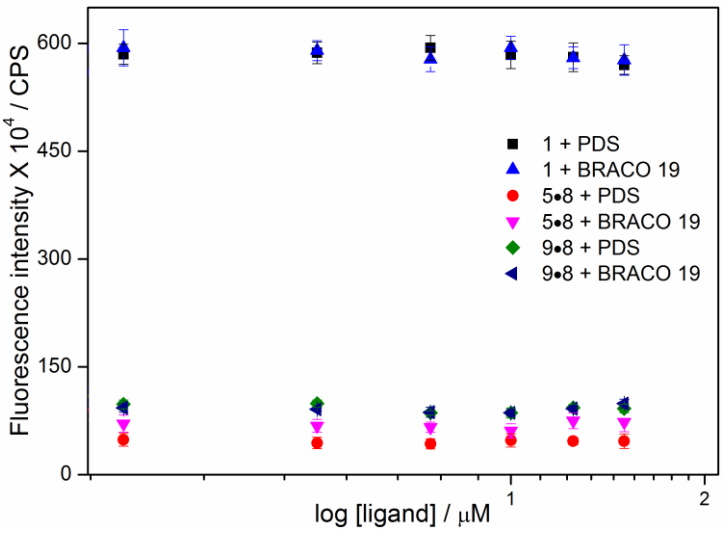


Figure 14. Control binding experiments. Fluorescence intensity of nucleoside **1** and duplexes of h-Telo DNA (**5•8**) and TERAA (**9•8**) (0.25 μM in 10 mM Tris-HCl buffer pH 7.5, 100 mM NaCl) as a function of increasing concentrations of PDS and BRACO19. Samples were excited at 330 nm with excitation and emission slit widths of 8 nm and 10 nm, respectively. Changes in FI of **1** and duplexes **5•8** and **9•8** are plotted at respective emission maximum (447 nm, 430 nm and 440 nm, respectively). Addition of increasing concentrations of ligands to nucleoside **1** and duplexes of h-Telo DNA (**5•8**) and TERAA (**9•8**) resulted only in minor changes in fluorescence intensity.

3A.2.5 Binding of ligands to higher-order h-Telo DNA GQ structures

The telomeric DNA, which terminates in a 100–200 bases 3' single stranded overhang of G-rich hexameric repeats, can form higher-order GQ structures (e.g., dimer or multimer). [92,93] Such consecutive GQ units have been also considered as a potential target as they play an important role in recognition and function of telomeric DNA. We synthesized longer h-Telo DNA repeat **11** containing two modifications, which could form two consecutive GQ units. [94] The CD and fluorescence profiles of ON **11** in the presence of KCl and NaCl revealed the formation of mixed-type and antiparallel GQ structures, respectively (Figure 15). Akin to ON **5**, addition of increasing concentrations of PDS and BRACO19 resulted in concentration-dependent decrease in fluorescence intensity, which however, yielded similar binding constants in both Na^+ and K^+ ionic conditions (Table 4). These results underscore the potential utility of emissive nucleoside in studying the structure as well as in identifying higher-order GQ binders.

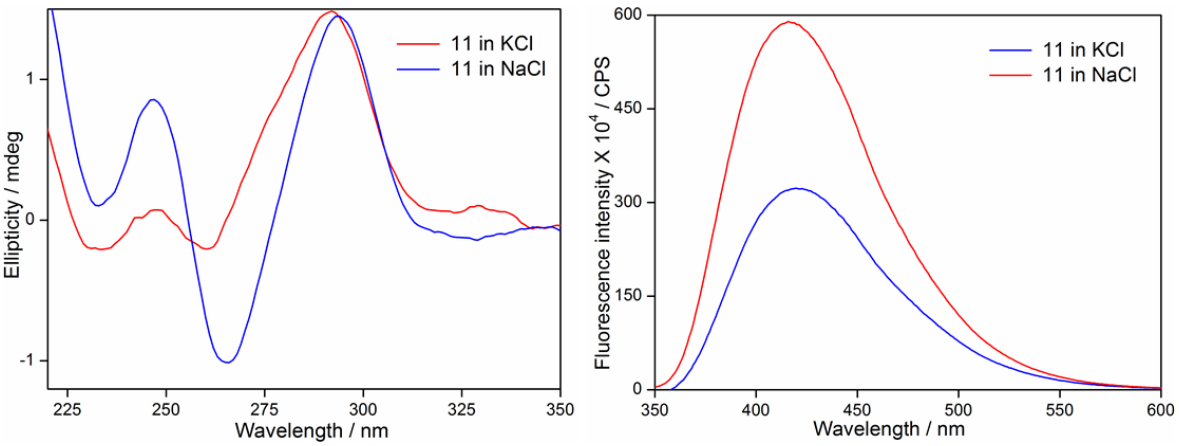


Figure 15. (**A**) CD and (**B**) fluorescence spectra of doubly modified longer h-Telo DNA repeat **11** in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl. CD and fluorescence profiles of ON **11** in the presence of KCl and NaCl were found to be similar to h-Telo DNA **5**.

3A.3 Conclusions

Environment-sensitive and minimally perturbing 5-benzofuran-modified 2'-deoxyuridine (**1**) and uridine (**2**) analogs, when incorporated into one of the loop residues of human telomeric DNA and RNA ONs, reported the formation of different DNA and RNA GQ structures with significant enhancement in fluorescence intensity. These GQ sensors enabled the development of a simple fluorescence method to detect and quantify topology- and nucleic acid-specific binding of ligands to GQ structures. Although many fluorescent purine nucleosides analogs have been used in the study of DNA GQs, their use in the discrimination of different GQ topologies based on nucleic acid type and in the quantitative estimation of ligands binding to DNA and RNA GQs has not been well explored.[35] Moreover, some of the easily accessible analogs (e.g., 2-AP, 6-MI, furyl-dG), though exhibit enhancement in fluorescence depending on the position of incorporation, significantly destabilize the GQ structures.[60,66,67] Hence, implementation of such probes in screening formats to identify efficient GQ binders may not be feasible. In this context, benzofuran-modified nucleoside analogs could serve as useful GQ probes and provide an efficient platform to identify topology- and nucleic acid-specific GQ binders of therapeutic potential.

3A.4 Experimental Section

3A.4.1 Materials

5-iodouridine, benzofuran, butyllithium, tributyltin chloride, *bis*(triphenylphosphine)-palladium(II) chloride, 4,4'-dimethoxytrityl chloride, Pyridostatin (PDS), BRACO19 were obtained from Sigma-Aldrich. 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite was purchased from Alfa Aesar. *N*-benzoyl-protected dA, dT, *N*,*N*-dimethylformamidine-protected dG and *N*-acetyl-protected dC phosphoramidite substrates for solid-phase DNA synthesis were purchased from ChemGenes Corporation. Solid supports for DNA synthesis were obtained from ChemGenes Corporation. TBDMS-protected ribonucleoside phosphoramidite substrates and solid supports required for RNA synthesis were purchased from Glen Research. All other reagents for solid-phase ON synthesis were obtained from ChemGenes Corporation. h-Telo DNA ONs **5**, **6** and **11** were synthesized by following our earlier report using phosphoramidite **3**.[72] Synthetic DNA ONs **7** and **8** were purchased from Integrated DNA Technologies, Inc. and purified by polyacrylamide gel electrophoresis (PAGE) under denaturing condition, and desalted on Sep-Pak Classic C18 cartridges (Waters Corporation). Custom synthesized oligoribonucleotide **10** was purchased from Dharmacon RNAi Technologies, deprotected according to the supplier's protocol, PAGE-purified and desalted using Sep-Pak Classic C18 cartridge. Chemicals (BioUltra grade) for preparing buffer solutions were purchased from Sigma-Aldrich. Autoclaved water was used for preparation of all buffer solutions and in biophysical analysis.

3A.4.2 Instrumentation

NMR spectra were recorded on 400 MHz Jeol ECS-400 spectrometer. Mass measurements were recorded on Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer and Water Synapt G2 High Definition mass spectrometers. Modified DNA and RNA ONs were synthesized on Applied Biosystems RNA/DNA synthesizer (ABI-394). Absorption spectra were recorded on Shimadzu UV-2600 spectrophotometer. HPLC analysis was performed using Agilent Technologies 1260 Infinity. UV-thermal melting studies of ONs were performed on Cary 300Bio UV-Vis spectrophotometer and CD analysis was performed on JASCO J-815 CD spectrometer. Steady-state and time-resolved fluorescence experiments were carried out in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3).

3A.4.3 Synthesis

5-Benzofuran-modified 5'-DMT-protected uridine 13: A mixture of 5'-DMT-protected 5-iodouridine **12** (0.750 g, 1.12 mmol, 1.0 equiv), 2-(tri-*n*-butylstannyl)benzofuran (0.681 g, 1.67 mmol, 1.5 equiv), and *bis*(triphenylphosphine)-palladium(II) chloride (0.047 g, 0.07 mmol, 0.06 equiv) was dissolved in anhydrous dioxane (15 ml). The reaction mixture was heated at 90 °C for 2 h and filtered through celite pad. The celite pad was washed with dichloromethane (2 × 20 ml), filtrate was evaporated and the resulting solid residue was purified by silica gel column chromatography to afford the compound **13** as a white foam (0.68 g, 92%). TLC (CH_2Cl_2:MeOH = 94:6 with few drops of Et_3N); R_f = 0.74; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.49 (s, 1H), 7.47 (d, *J* = 8.4 Hz, 2H), 7.36–7.33 (m, 6H), 7.18 (t, *J* = 7.8 Hz, 2H), 7.08–7.00 (m, 2H), 6.91 (t, *J* = 8.4 Hz, 1H), 6.69 (dd, *J* = 8.8, 2.4 Hz, 4H), 6.25 (d, *J* = 8.4 Hz, 1H), 6.04 (d, *J* = 5.2 Hz, 1H), 4.56 (br, 4H), 4.49 (t, *J* = 5.2 Hz, 1H), 4.30–4.26 (m, 2H), 3.59 (s, 3H), 3.58 (s, 3H); ^{13}C NMR (100 MHz**,** $CDCl_3$ **)**: δ (ppm) 161.6, 158.5, 153.5, 151.4, 147.6, 144.6, 135.9, 135.7, 135.3, 130.2, 130.1, 129.0, 128.3, 128.0, 127.0, 124.1, 122.7, 120.8, 113.3, 110.9, 107.0, 105.8, 90.5, 86.9, 75.8, 71.2, 63.2, 55.2, 53.1; HRMS: (m/z) Calculated for $C_{38}H_{34}N_2O_9K$ $[M+K]^+$ = 701.1901, found: 701.1908.

5-Benzofuran-modified 2'-*O*-TBDMS protected uridine 14: Compound **13** (0.600 g, 0.91 mmol, 1.0 equiv) and silver nitrate (0.346 g, 2.04 mmol, 2.2 equiv) were dissolved in anhydrous pyridine (1.8 ml). To the above solution was added anhydrous THF (9 ml). After stirring for 10 min *tert*-butyldimethylsilyl chloride (0.307 g, 2.04 mmol, 2.25 equiv) was added. Reaction mixture was stirred for additional 0.5 h and filtered through celite pad. Celite pad was washed with ethyl acetate (2 × 15 ml) and the combined organic solution was then washed with 5 % sodium bicarbonate (25 ml) followed by brine solution (25 ml). The organic extract was dried over sodium sulphate and evaporated. The resulting crude residue was purified by silica gel column chromatography to afford the product **14** as a white solid (0.425 g, 60%). TLC (petroleum ether:EtOAc = 60:40 with few drops of Et_3N); R_f = 0.6; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.55 (s, 1H), 7.55 (d, *J* = 7.6 Hz, 2H), 7.45–7.39 (m, 6H), 7.20 (t, *J* = 7.6 Hz, 2H), 7.10–7.02 (m, 2H), 6.82 (t, *J* = 7.8 Hz, 1H), 6.71 (dd, *J* = 8.8, 4.8 Hz, 4H), 6.19 (d, *J* = 6.8 Hz, 1H), 5.80 (d, *J* = 8.4 Hz, 1H), 4.58 (t, *J* = 5.4 Hz, 1H), 4.21 (dd, *J* = 7.8, 5.4 Hz, 1H), 4.13 (t, *J* = 4.0 Hz, 1H), 3.72 (dd, *J* = 11.0, 1.8 Hz, 1H), 3.63 (s, 3H), 3.60 (s, 3H), 3.25 (dd, *J* = 11.0, 2.6 Hz, 1H), 0.90 (s, 9H), 0.13 (s, 3H), 0.11 (s, 3H); ^{13}C NMR (100 MHz**,** $CDCl_3$ **)**: δ (ppm) 160.0, 158.7, 153.5, 149.5, 147.3, 144.7, 135.8, 135.5, 134.9,

130.2, 130.1, 128.8, 128.3, 128.0, 127.1, 124.3, 122.7, 120.8, 113.3, 111.0, 107.5, 106.2, 87.8, 87.1, 84.2, 76.0, 71.3, 63.3, 55.3, 31.1, 25.7, -4.6, -5.0; HRMS: (m/z): Calculated for $C_{44}H_{48}N_2O_9SiK$ $[M+K]^+$ = 815.2766, found: 815.2769.

5-Benzofuran-modified uridine phosphoramidite substrate 4: To a solution of compound **14** (0.250 g, 0.32 mmol, 1.0 equiv) in anhydrous dichloromethane (2.5 ml) was added *N*,*N*-diisopropylethylamine (0.140 ml, 0.80 mmol, 2.5 equiv) and stirred for 10 min. To this solution was slowly added 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.11 ml, 0.48 mmol, 1.5 equiv) and the reaction mixture was stirred for 12 h. The solvent was evaporated to dryness and the residue was redissolved in ethyl acetate (20 ml), which was washed with 5% sodium bicarbonate solution (15 ml) followed by brine solution (15 ml). The organic extract was dried over sodium sulphate, evaporated and the crude solid residue was purified by column chromatography to afford the product **4** as a white solid (0.254 g, 81%). TLC (petroleum ether:EtOAc = 60:40 with few drops of Et_3N); R_f = 0.70; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.61 (s, 1H), 7.58 (d, *J* = 7.6 Hz, 2H), 7.44 (d, *J* = 7.6 Hz, 6H), 7.21 (t, *J* = 7.6 Hz, 2H), 7.10 (t, *J* = 7.4 Hz, 1H), 7.01 (t, *J* = 7.6 Hz, 1H), 6.78–6.70 (m, 6H), 6.16 (d, *J* = 6.4 Hz, 1H), 5.61 (d, *J* = 8.4 Hz, 1H), 4.56 (dd, *J* = 6.4, 5.2 Hz, 1H), 4.35 (br, 1H), 4.17–4.11 (m, 1H), 3.80 (br, 1H), 3.62 (s, 3H), 3.60 (s, 3H), 3.54 (dd, *J* = 11.4, 6.2 Hz, 2H), 3.43–3.34 (m, 1H), 3.15 (dd, *J* = 11.0, 2.6 Hz, 1H), 2.28–2.21 (m, 2H), 1.16 (d, *J* = 6.8 Hz, 6H), 1.13 (d, *J* = 6.8 Hz, 6H), 0.86 (s, 9H), 0.06 (s, 6H); ^{13}C NMR (100 MHz**,** $CDCl_3$ **)**: δ (ppm) 160.0, 158.6, 153.5, 149.4, 147.5, 144.7, 136.0, 135.6, 135.1, 130.3, 128.9, 128.4, 127.1, 124.1, 122.7, 120.7, 117.4, 113.3, 111.0, 107.2, 105.9, 87.6, 87.1, 84.3, 75.2, 73.0, 62.9, 57.6, 55.3, 43.6, 25.8, 24.8, 22.8, 20.2, 18.1, -4.6, -4.7; ^{31}P NMR (162 MHz, CDCl3): δ (ppm) 151.6; HRMS: (m/z): Calculated for $C_{53}H_{65}N_4O_{10}PSiK$ $[M+K]^+$ = 1015.3845, found: 1015.3854.

3A.4.4 Solid-phase synthesis of modified DNA and RNA ONs

Benzofuran-modified h-Telo DNA ONs **5**, **6** and **11** were synthesized and purified according to our earlier report using phosphoramidite **3**.[53] Benzofuran-modified TERRA ON **8** was synthesized on a 1 μmole scale CPG solid support (1000 Å) by following standard solid-phase RNA ON synthesis protocol using 2'-*O*-TBDMS-protected phosphoramidite substrates. While incorporation of native 2'-*O*-TBDMS-protected phosphoramidites was performed with a coupling time of 10 min, incorporation of fluorescent 2'-*O*-TBDMS-protected

phosphoramidite substrate **4** was performed with a coupling time of 30 min. The coupling efficiency of **4** based on trityl monitor-assay was found to be ~35%. After deprotection of trityl group on the synthesizer, the solid support was treated with a 1:1 solution of 10 M methylamine in ethanol and water (1.5 mL) for 12 h. The mixture was centrifuged and the supernatant was evaporated to dryness on a Speed Vac. The residue was then dissolved in DMSO (100 μL) and was added TEA.3HF (150 μL). The resulting solution was heated at 65 °C for 2.5 h and was brought to RT. The completely deprotected ON solution was lyophilized and was purified by PAGE (20% gel) under denaturing conditions. The band corresponding to the full-length product was identified by UV shadowing. The ON was extracted with ammonium acetate buffer (0.5 M, 3 ml) and desalted using a Sep-Pak classic C18 cartridge.

3A.4.5 MALDI-TOF mass analysis of modified DNA and RNA ONs

2 μL of the modified ON (~200 μM) was combined with 1 μL of ammonium citrate buffer (100 mM, pH 9), 2 μL of a DNA internal standard (200 μM) and 4 μL of saturated 3-hydroxypicolinic acid solution. The sample was desalted using an ion-exchange resin (Dowex 50W-X8, 100-200 mesh, ammonium form), spotted on the MALDI plate and was air dried. The resulting spectrum was calibrated relative to an internal DNA ON standard.

3A.4.6 Photophysical analysis of fluorescently modified DNA and RNA GQs and corresponding duplexes

Steady-state fluorescence

GQ forming ONs **5**/**6**/**9**/**11** (0.25 μM) were heated at 90 °C for 3 min in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl. Samples were then cooled slowly to RT and were kept at 4 °C for 1 h. Duplexes (0.25 μM) of ONs **5/6/9** were assembled by heating a 1:1 mixture of the ON with complementary DNA ON **8** in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl at 90 °C for 3 min. Duplexes (**5•8**, **6•8** and **9•8)** were then cooled slowly to RT and stored at 4 °C for 1 h. Fluorescently modified GQs and duplexes were excited at 330 nm with excitation and emission slit widths of 8 and 10 nm, respectively. All fluorescence experiments were performed in triplicate in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 at 20 °C.

Quantum yield determination

Quantum yield of fluorescently modified DNA and RNA ONs were determined relative to the quantum yield of nucleoside **1** and **2** respectively, by using the following equation. [69]

$$\Phi_{F(x)} = (A_s/A_x)(F_x/F_s)(n_x/n_s)^2 \Phi_{F(s)}$$

Where s is respective fluorescent nucleoside, x is fluorescently modified DNA and RNA ON constructs, A is the absorbance at excitation wavelength, F is the area under the emission curve, n is the refractive index of the buffer, and Φ_F is the quantum yield. Quantum yield of nucleoside **1** and **2** is 0.19 and 0.21, respectively.

Time-resolved fluorescence

Excited-state lifetime of GQs and duplexes (1 μM), assembled as mentioned above, were determined using TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3). ONs were excited using 339 nm LED source (IBH, UK, NanoLED-339L) and fluorescence signal at respective emission maximum was collected (Table 1). All experiments were performed in duplicate and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay kinetics of all ON constructs were found to be biexponential with χ^2 (goodness of fit) values very close to unity.

3A.4.7 Circular dichroism analysis

h-Telo DNA ONs **5** and **7** and TERRA ONs **9** and **10** (4 μM) were assembled into GQs by heating the ONs at 90 °C for 3 min in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl. Samples were then cooled slowly to RT and kept at 4 °C for 1 h prior to CD analysis. CD spectra were collected from 350 to 220 nm on a Jasco J-815 CD spectrometer using 1 nm bandwidth at 20 °C. Experiments were performed in duplicate wherein each spectrum was an average of five scans. The spectrum of buffer was subtracted from all sample spectra.

3A.4.8 Thermal melting analysis

h-Telo DNA ONs **5** and **7** and TERRA ONs **9** and **10** (1 μM) were annealed by heating the ONs at 90 °C for 3 min in 10 mM Tris-HCl buffer (pH 7.5) containing either 100 mM KCl or 100 mM NaCl. Samples were then cooled slowly to RT and kept at 4 °C for 1 h prior to thermal melting analysis. UV-thermal melting analysis was performed in triplicate at 260 nm and 295 nm by using Cary 300Bio UV-Vis spectrophotometer.

3A.4.9 Fluorescence binding assay: Binding of PDS and BRACO19 to h-Telo DNA ONs 5, 11 and TERRA 9

A series of solution of DNA and RNA GQs (0.25 μM) of **5/11** and **9**, respectively, containing increasing concentrations of PDS/BRACO19 was prepared in Tris-HCl buffer (10 mM, pH 7.5) containing either 100 mM KCl or 100 mM NaCl. The concentration of ligand was increased from 13 nM to 2.5 μM. The samples were excited at 330 nm with excitation and emission slit widths of 8 and 10 nm, respectively. A spectral blank of the buffer in the absence of ON and ligand was subtracted from all measurements. All fluorescence experiments were performed in triplicate in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 at 20 °C. Control binding experiments with duplexes **5•8** and **9•8** were also performed under similar conditions. Normalized fluorescence intensity (F_N) versus log of ligand concentration plots were fitted using Hill equation (OriginPro 8.5.1) to determine the apparent binding constant K_d for the binding of ligands to GQs. [70]

$$F_N = \frac{F_i - F_s}{F_0 - F_s}$$

F_i is the fluorescence intensity at each titration point. F_0 and F_s are the fluorescence intensity in the absence of ligand (L) and at saturation, respectively. n is the Hill coefficient or degree of cooperativity associated with the binding.

$$F_N = F_0 + (F_s - F_0)\left(\frac{[L]^n}{[K_d]^n + [L]^n}\right)$$

The value of the Hill coefficient describes the cooperativity of ligand binding in the following way if $n>1$ is a positively cooperative binding means once one ligand molecule is bound to GQ, its affinity for other ligand molecules increases. While if $n<1$ is a negatively cooperative binding means once one ligand molecule is bound to GQ, its affinity for other ligand molecules decreases. We have observer a Hills coefficient of ~3 for each binding study.

3A.5 Refrences

[1] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35*, 7429−7455.

[2] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261−275.

[3] G. W. Collie, G. N. Parkinson, *Chem. Soc. Rev.* **2011**, *40*, 5867−5892.

[4] S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127*, 10584−10586.

[5] S. Kumari, A. Bugaut, J. L. Huppert, S. Balasubramanian, *Nat. Chem. Biol.* **2007**, *3*, 218−221.

[6] Y.-L. Zheng, N. Hu, Q. Sun, C. Wang, P. R. Taylor, *Cancer Res.* **2009**, *69*, 1604–1614.

[7] A. T. Phan, *FEBS J*. **2010**, *277*, 1107−1117.

[8] P. L. T. Tran, J.-L. Mergny, P. Alberti, *Nucleic Acids Res.* **2011**, *39*, 3282−3294.

[9] S. Zhang, Y. Wu, W. Zhang, *ChemMedChem* **2014**, *9*, 899−911.

[10] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263–282.

[11] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876–880.

[12] A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 2723–2735.

[13] F. M. Lannan, I. Mamajanov, N. V. Hud, *J. Am. Chem. Soc.* **2012**, *134*, 15324−15330.

[14] Y. Xu, K. Kaminaga, M. Komiyama, *J. Am. Chem. Soc.* **2008**, *130*, 11179–11184.

[15] H. Martadinata, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 2570–2578.

[16] A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou, J.-L. Mergny, *Biochimie*, **2008**, *90*, 131–155.

[17] D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem*. **2008**, *6*, 627−636.

[18] S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed.* **2010**, *49*, 4020–4034.

[19] Y. Xu, M. Komiyama, *Methods* **2012**, *57*, 100–105.

[20] P. V. L. Boddupally, S. Hahn, C. Beman, B. De, T. A. Brooks, V. Gokhale, L. H. Hurley, *J. Med. Chem.* **2012**, *55*, 6076–6086.

[21] J. M. Nicoludis, S. P. Barrett, J.-L. Mergny, L. A.Yatsunyk, *Nucleic Acids Res.* **2012**, *40*, 5432−5447.

[22] J. Mohanty, N. Barooah, V. Dhamodharan, S. Harikrishna, P. I. Pradeepkumar, A. C. Bhasikuttan, *J. Am. Chem. Soc.* **2013**, *135*, 367–376.

[23] M. Micco, G.W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka, S. Neidle, *J. Med. Chem.* **2013**, *56*, 2959–2974.

[24] T. Agarwal, M. K. Lalwani, S. Kumar, S. Roy, T. K. Chakraborty, S. Sivasubbu, S. Maiti, *Biochemistry* **2014**, *53*, 1117–1124.

[25] W. J. Chung, B. Heddi, F. Hamon, M.-P. Teulade-Fichou, A. T. Phan, *Angew. Chem. Int. Ed.* **2014**, *53*, 999–1002.

[26] A. Laguerre, N. Desbois, L. Stefan, P. Richard, C. P. Gros, D. Monchaud, *ChemMedChem*, **2014,** *9*, 2035−2039.

[27] Z. Yu, M. Han, J. A. Cowan, *Angew. Chem. Int. Ed.* **2015**, *54*, 1901–1905.

[28] K. Paeschke, S. Juranek, T. Simonsson, A. Hempel, D. Rhodes, H. J. Lipps, *Nat. Struct. Mol. Biol.* **2008**, *15*, 598–604.

[29] Y. Xu, Y. Suzuki, K. Ito, Komiyama, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14579–14584.

[30] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182–186.

[31] E. Y. N. Lam, D. Beraldi, D. Tannahill, S. Balasubramanian, *Nat. Commun.* **2013**, *4*, 1796–1804.

[32] A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh Jr., D. Sen, P. M. Lansdrop, *Nucleic Acids Res.* **2013,** *42*, 860−869.

[33] R.-Y Wu, K.-W. Zheng, J.-Y. Zhang, Y.-H. Hao, Z. Tan, *Angew. Chem. Int. Ed*. **2015**, *54*, 2447–2451.

[34] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J.-M. Perrier-Cornet, J. M. Y. Wong, D. Monchaud, *J. Am. Chem. Soc.* **2015**,*137*, 8521–8525.

[35] B. R. Vummidi, J. Alzeer, N. W. Luedtke, *ChemBioChem* **2013**, *14*, 540–558.

[36] E. Largy, A. Granzhan, F. Hamon, D. Verga, M. P. Teulade-Fichou, *Top. Curr. Chem.* **2013**, *330*,111–177.

[37] A. C. Bhasikuttan, J. Mohanty, *Chem. Commun.* **2015**, *51*, 7581−7597.

[38] D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C. B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. C. Lim, D. Von Hoff, K. Anderes, W. G. Rice, *Cancer Res.* **2009**, *69*, 7653–7661.

[39] J.-L. Mergny, J.-C. Maurizot, *ChemBioChem*, **2001**, *2*, 124−132.

[40] H. Ueyama, M. Takagi, S. Takenaka, *J. Am. Chem. Soc.* **2002,** *124*, 14286−14287.

[41] R. A. Darby, M. Sollogoub, C. McKeen, L. Brown, A. Risitano, N. Brown, C. Barton, T. Brown, K. R. Fox, *Nucleic Acids Res.* **2002**, *30*, e39.

[42] A. Risitano, K. R. Fox, *Biochemistry*, **2003**, *42*, 6507−6513.

[43] L. Ying, J. J. Green, H. Li, D. Klenerman, S. Balasubramanian, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14629−14634.

[44] S. Nagatoishi, T. Nojima, E. Galezowska, B. Juskowiak, S. Takenaka, *ChemBioChem* **2006**, ***7***, 1730−1737.

[45] S. Takenaka, B. Juskowiak, *Anal. Sci.* **2011**, *27*, 1167−1172.

[46] K. Ohtsuka, S. Sato, Y. Sato, K. Sota, S. Ohzawa, T. Matsuda, K. Takemoto, N. Takamune, B. Juskowiak, T. Nagai, S. Takenaka, *Chem. Commun.* **2012**, *48*, 4740−4742.

[47] H. Yang, J. Ji, Y. Liu, J. Kong, B. Liu, *Electrochem. Commun.* **2009,** *11*, 38−40.

[48] J. J. Li, X. Fang, W. Tan, *Biochem. Biophys. Res. Commun.* **2002**, *292*, 31−40.

[49] S. De Tito, F. Morvan, A. Meyer, J. J. Vasseur, A. Cummaro, L. Petraccone, B. Pagano, E. Novellino, A. Randazzo, C. Giancola, D. Montesarchio, *Bioconjug. Chem.* **2003,** *24*, 1917−1927.

[50] L. Lacroix, A. Séosse, J. L. Mergny, *Nucleic Acids Res.* **2011**, *39*, e21.

[51] A. Benz, V. Singh, T. U. Mayer, J. S. Hartig, *ChemBioChem*, **2011**, *12*, 1422−1426.

[52] S. Nagatoishi, T. Nojima, B. Juskowiak, S. Takenaka, *Angew. Chem. Int. Ed.* **2005**, *44*, 5067−5070.

[53] D. D. Le, M. D. Antonio, L. K. M. Chan, S. Balasubramanian, *Chem Commun.* **2015,** *51*, 8048−8050.

[54] M. Nikan, M. Di Antonio, K. Abecassis, K. McLuckie, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2013**, *52*, 1428–1431.

[55] K. Iida, T. Nakamura, W. Yoshida, M. Tera, K. Nakabayashi, K. Hata, K. Ikebukuro, K. Nagasawa, *Angew. Chem. Int. Ed.* **2013**, *52*, 12052–12055.

[56] A. Laguerre, L. Stefan, M. Larrouy, D. Genest, J. Novotna, M. Pirrotta, D. Monchaud, *J. Am. Chem. Soc.* **2014**, *136*, 12406–12414.

[57] D. Zhao, X. Dong, N. Jiang, D. Zhang, C. Liu, *Nucleic Acids Res.* **2014,** *42*, 11612−11621.

[58] J. C. Myers, S. A. Moore, Y. Shamoo, *J Biol Chem.* **2003**, *278*, 42300–42306.

[59] J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucleic Acids Res.* **2005,** *33*, 4649−4659.

[60] T. Kimura, K. Kawai, M. Fujitsuka T. Majima, *Chem. Commun.* **2006**, 401−402.

[61] T. Kimura, K. Kawai, M. Fujitsuka, T. Majima, *Tetrahedron* **2007**, *63*, 3585–3590.

[62] R. D. Gray, L. Petraccone, J. O. Trent, J. B. Chaires, *Biochemistry* **2010**, *49*, 179−194.

[63] A. Dumas, N. W. Luedtke, *Nucleic Acids Res.* **2011**, *39*, 6825−6834.

[64] M. Sproviero, K. L. Fadock, A. A. Witham, R. A. Manderville, P. Sharma, S. D. Wetmore, *Chem. Sci.* **2014,** *5*, 788−796.

[65] A. Nadler, J. Strohmeier, U. Diederichsen, *Angew. Chem. Int. Ed.* **2011**, *50*, 5392−5396.

[66] J. Johnson, R. Okyere, A. Joseph, K. Musier-Forsyth, B. Kankia, *Nucleic Acids Res.* **2013**, *41*, 220−228.

[67] M. Sproviero, K. L. Fadock, A. A. Witham, R. A. Manderville, *ACS Chem. Biol.* **2015,** *10*, 1311−1318.

[68] S. Doose, H. Neuweiler, M. Sauer, *ChemPhysChem* **2009**, *10*, 1389–1398.

[69] D. Lavabre, S. J. Fery-Forgues, *Chem. Educ.* **1999**, *76*, 1260−1264.

[70] V. K. Tam, D. Kwong, Y. Tor, *J. Am. Chem. Soc.* **2007,** *129*, 3257−3266.

[71] A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J.* **2011**, *17*, 12820–12827.

[72] A. A. Tanpure, S. G. Srivatsan, *ChemBioChem* **2012**, *13*, 2392–2399.

[73] P. M. Sabale, J. T. George, S. G. Srivatsan, *Nanoscale* **2014**, *6*, 10460–10469.

[74] E. H. Blackburn, *Nat. Struct. Biol*. **2000**, ***7***, 847−850.

[75] R. E. Verdun, J. Karlseder, *Nature* **2007**, *447*, 924−931.

[76] J. W. Shay, I. B. Roninson, *Oncogene* **2004**, *23*, 2919−2933.

[77] Y.-L. Zheng, N. Hu, Q. Sun, C. Wang, P. R. Taylor, *Cancer Res.* **2009**, *69*, 1604−1614.

[78] J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res.* **2007,** *35*, 3064−3075.

[79] C. M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto, J. Lingner, *Science* **2007**, *318*, 798–801.

[80] S. Schoeftner, M. A. Blasco, *Nat. Cell Biol.* **2008**, *10*, 228–236.

[81] T. Kimura, K. Kawai, M. Fujitsuka, T. Majima, *Chem Commun.* **2004,** 1438−1439.

[82] A. A. Tanpure, M. G. Pawar, S. G. Srivatsan, *Isr. J. Chem.* **2013**, *53*, 366–378.

[83] Y. Xue, Z-Y. Kan, Q. Wang, Y. Yao, J. Liu, Y-H. Hao, Z. Tan, *J. Am. Chem. Soc.* **2007,** *129*, 11185−11191.

[84] G. Biffi, D.Tannahill, S. Balasubramanian, *J. Am. Chem. Soc.* **2012,** *134*, 11974−11976.

[85] P. A. Rachwal, K. R. Fox, *Methods* **2007,** *43*, 291–301.

[86] M. J. B. Moore, C. M. Schultes, J. Cuesta, F. Cuenca, M. Gunaratnam, F. A. Tanious, W. D. Wilson, S. Neidle, *J. Med. Chem.* **2006**, **49**, 582−599.

[87] R. Rodriguez, S. Müller, J. A. Yeoman, C. Trentesaux, J.-F. Riou, S. Balasubramanian, *J. Am. Chem. Soc.* **2008**, *130*, 15758–15759.

[88] D. Koirala, S. Dhakal, B. Ashbridge, Y. Sannohe, R. Rodriguez, H. Sugiyama, S. Balasubramanian, H. Mao, *Nat Chem.* **2011,** *3*, 782–787.

[89] G. W. Collie, S. Sparapani, G. N. Parkinson, S. Neidle, *J. Am. Chem. Soc.* **2011**, *133*, 2721–2728.

[90] R. W. Sinkeldam, A. J. Wheat, H. Boyaci, Y. Tor, *ChemPhysChem* **2011**, *12*, 567–570.

[91] L. M. Wilhelmsson, *Quart. Rev. Biophy*. **2010**, *43*, 159−183.

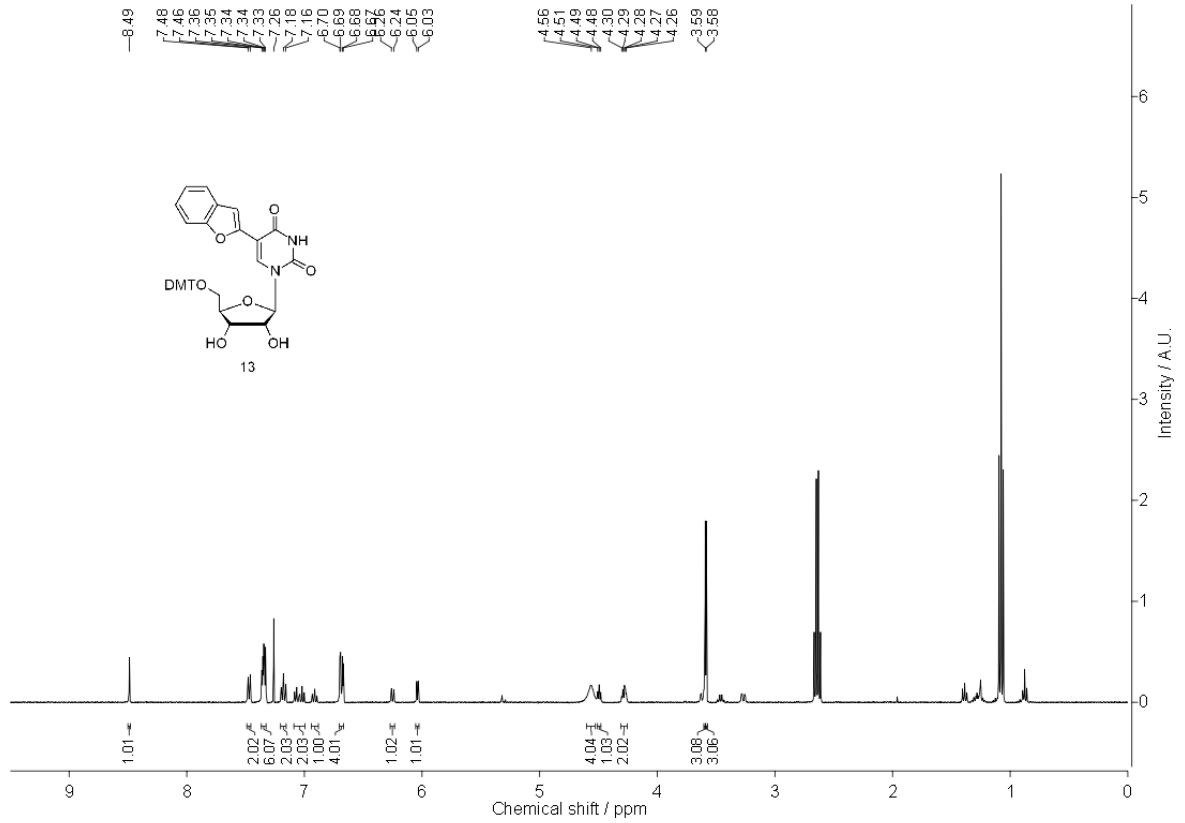
[92] Y. Xu, T. Ishizuka, K. Kurabayashi, M. Komiyama, *Angew. Chem. Int. Ed.* **2009**, *48*, 7833–7836.

[93] L. Petraccone, *Top Curr. Chem.* **2013**, *330*, 23–46.

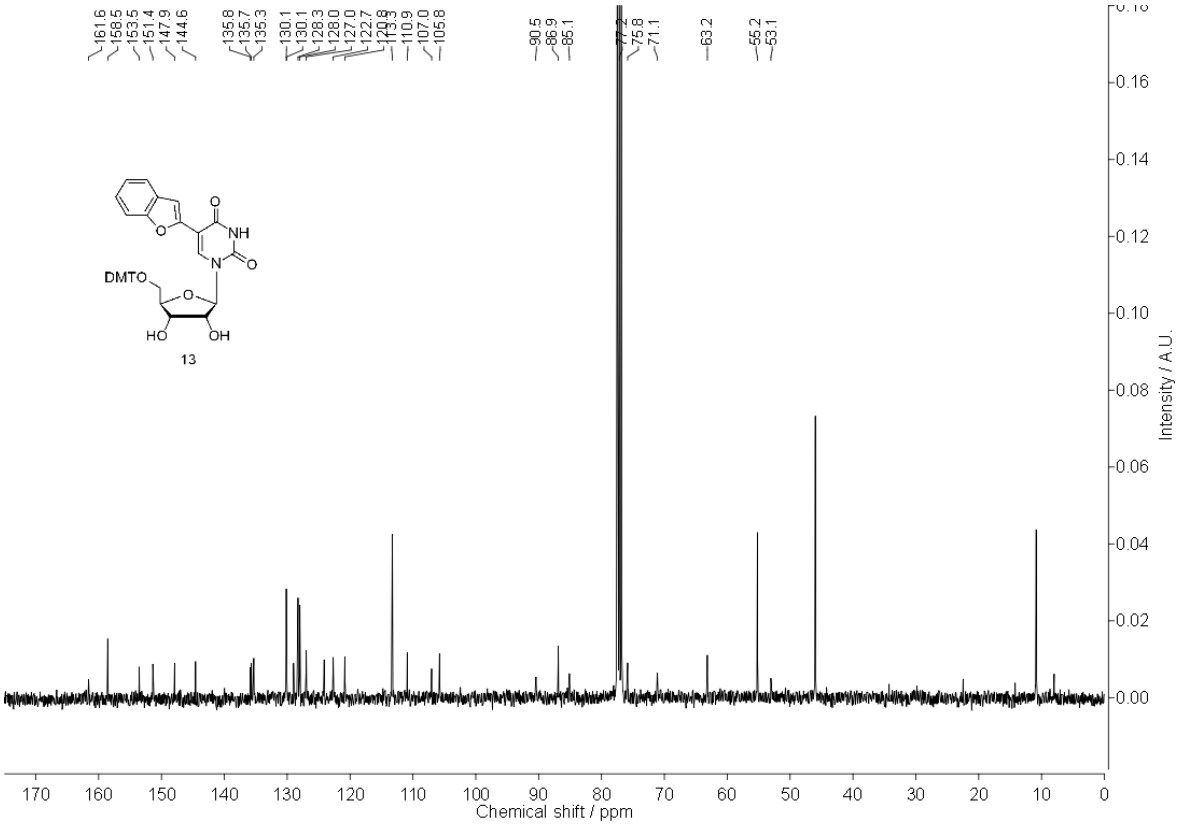
[94] C. Zhao, L. Wu, J. Ren, Y. Xu, X. Qu, *J. Am. Chem. Soc.* **2013**, *135*, 18786–18789.

3A.6 Appendix-III: Characterization data of synthesized compounds

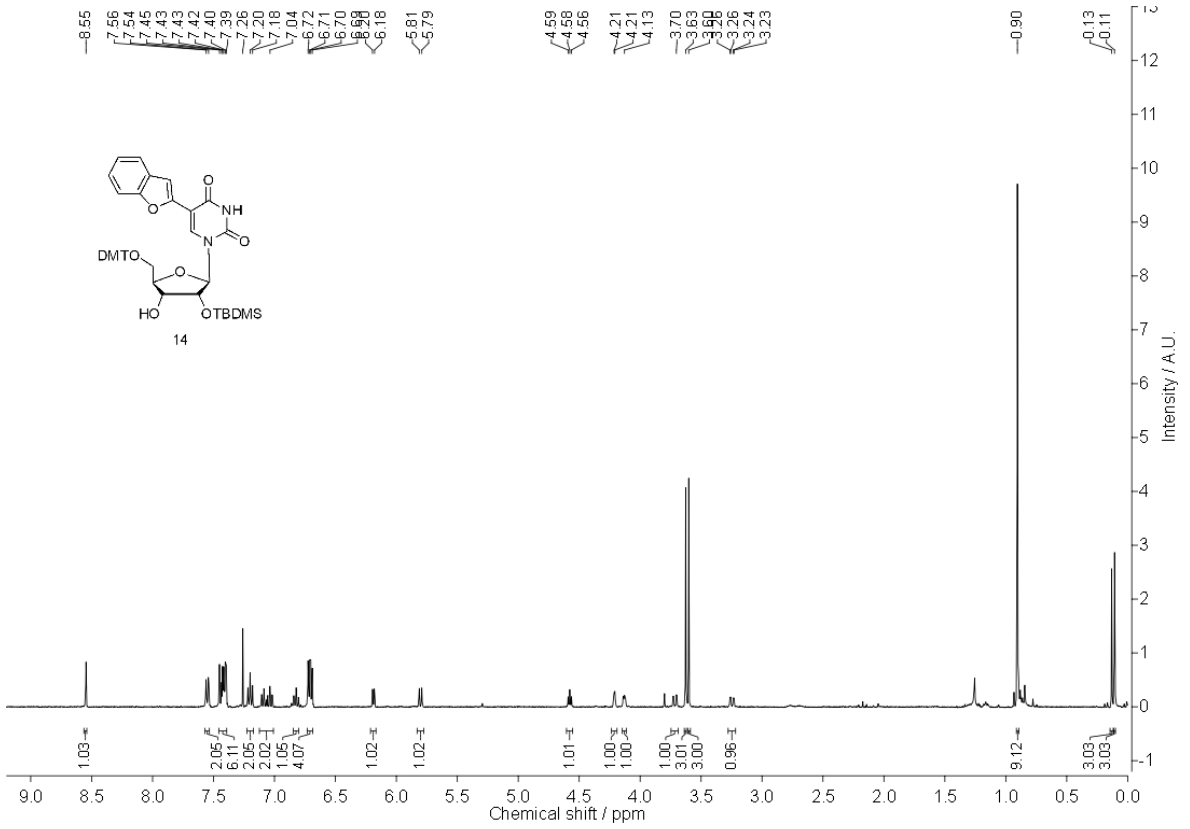
^{1}H NMR of 5-Benzofuran-modified 5'-DMT-protected uridine **13** in $CDCl_3$



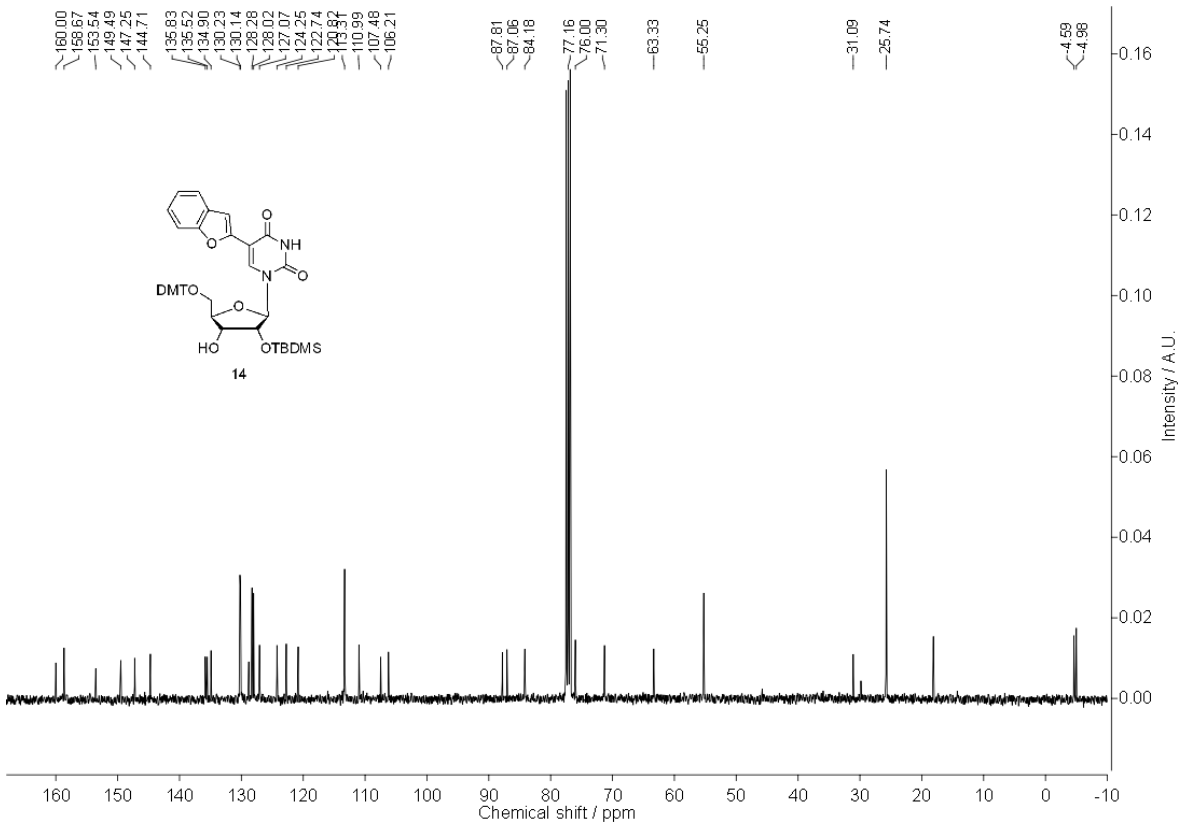
^{13}C NMR of 5-Benzofuran-modified 5'-DMT-protected uridine **13** in $CDCl_3$



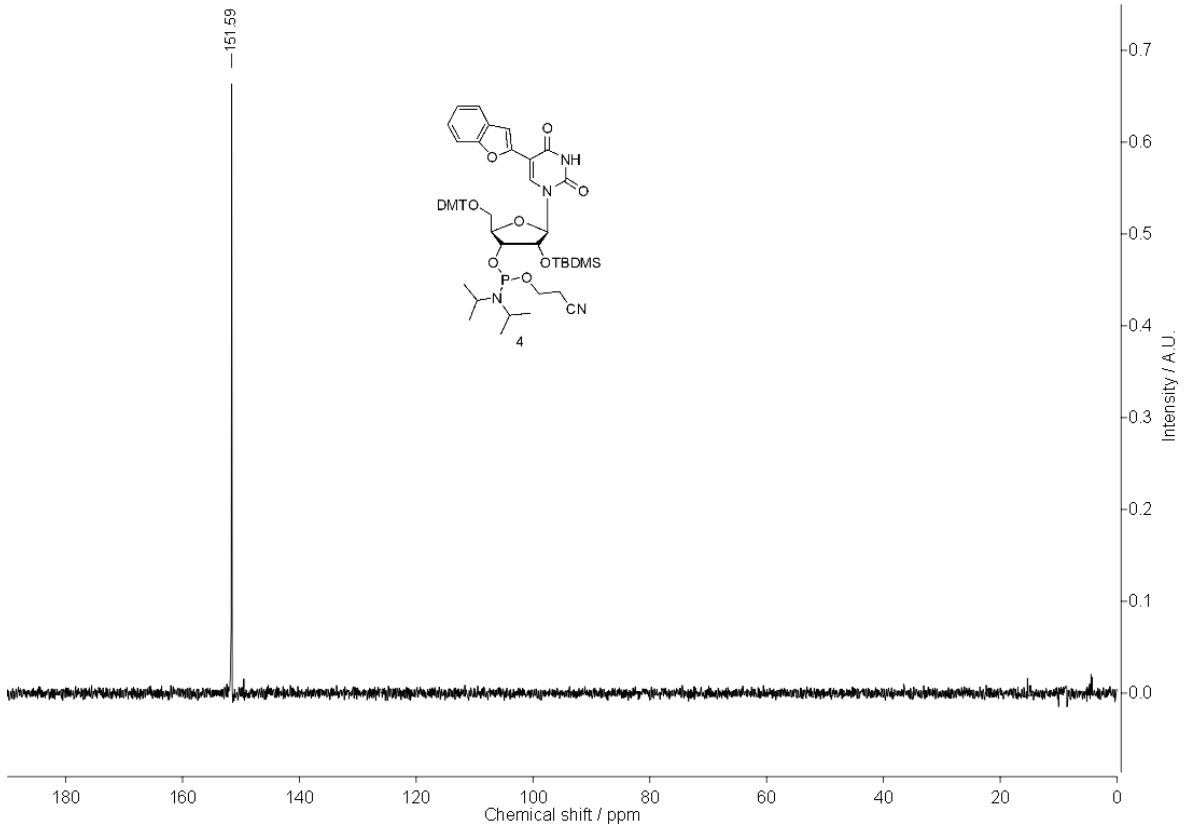
^{1}H NMR of 5-Benzofuran-modified 2'-*O*-TBDMS-protected uridine **14** in $CDCl_3$



^{13}C NMR of 5-Benzofuran-modified 2'-*O*-TBDMS-protected uridine **14** in $CDCl_3$



^{31}P NMR of phosphoramidite substrate **4** in $CDCl_3$



Chapter 3B

Benzofuran-conjugated fluorescent uridine analog to probe the conformational switch in i-motif forming human telomeric DNA and RNA repeats

3B.1 Introduction

Cytosine rich (C-rich) DNA oligonucleotides (ONs) in acidic condition form an intercalated tetraplex of hemiprotonated C•C^+ base pairs inducing a stable self-assembly called i- motif.[1] C-rich sequences abundantly present in human telomeric (h-telo) DNA and promoter regions of several oncogenes. Some of these sequences have been shown to form stable i-motif structure *in vitro* in slightly acidic pH.[2] Several proteins including transcription factors selectively bind to the C-rich DNA and induce i-motif structures. This protein DNA interaction leading to the control of oncogenic expression at the transcription level.[3,4] Although direct evidence for the existence of i-motif structure in cellular environment is yet to be proven, recent studies suggest the possibility of targeting these structures with specific ligands, which modulate the oncogenic expression at transcription level leading to potential therapeutics for genetic diseases.[5] Moreover, sequence specific molecular recognition properties of this pH sensitive self-assembly has been programmed and utilized as a building block for the design of several functional DNA nano-devices used for a variety of applications.[6]

Structurally, i-motif DNA exhibits a variety of folding topologies and wider range of stability *in vitro*, which greatly relies on the number of cytosine residues, the loop length and the analytical conditions.[2e,7] In addition, C-rich RNA sequences forming i-motif structures in solution have been identified.[8] Contrary to DNA i-motif conformation, the i-motif formed by RNA ON are much less stable possibly due to the hydrophilic 2' hydroxyl group, which prefer the hydration of RNA and thereby destabilizing the i-motif formation.[8] Sequence dependent folding and intercalation topologies of i-motif DNA structures have been investigated by X-ray crystallography and NMR at the atomic level.[9] Fluorescence resonance energy transfer (FRET) has been widely used to study the conformational dynamics and folding kinetics of DNA i-motif *in vitro* as well as *in vivo*.[10] Recently, base modified fluorescent nucleoside analogs have emerged as valuable probe to understand the folding transitions of i-motif DNA.[11] Kim and co-workers have introduced a pyrene-modified deoxyadenosine (^{Py}A) analog into c-rich DNA and characterized the i-motif formation by monitoring the exciplex emission of the fluorophore.[11b] Using a similar approach, Majima group has implemented a combination of 5-(pyrenylethynyl)-2'-deoxyuridine and 2'- anthraquinone-modified uridine analogs to study the electron transfer process during i-motif formation.[11c] Later, Mata and Luedtke used dimethylaniline (DMA)

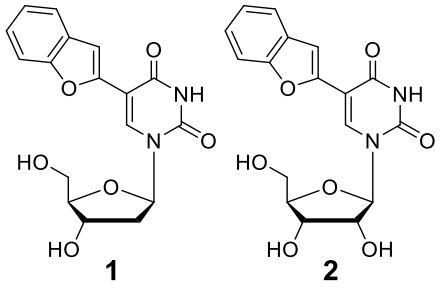
fused deoxycytidine (^{DMA}C) to compare the folding kinetics of i-motif with duplex.[11d] However, depending on the position within the DNA majority of these probes reasonably destabilize the i-motif structure, and hence, hamper the sensitivity of fluorescence based studies of i-motifs. In this context, development of minimally perturbing fluorescent probes, which faithfully report the structural transitions, will help in understanding these structures and further accelerate the therapeutic evaluation of this non-canonical structure of nucleic acids.

Here, we describe environment-sensitive and minimally perturbing fluorescent pyrimidine nucleoside analogs, based on a 5-(benzofuran-2-yl)uracil core, for the structural analysis of DNA and RNA i-motifs. Upon incorporating into one of the loop regions of C-rich h-telo DNA and RNA ON sequences, these analogs clearly signaled the transition from single stranded ONs into their respective i-motif structures. Importantly, steady-state as well as time-resolved fluorescence spectroscopy techniques indicated the potential of these probes in estimating the *ipH* of i-motif formation. Our results demonstrate that these i-motif sensors provide a simple platform to probe the structural polymorphism of biologically important C-rich DNA and RNA sequences.

3B.2 Results and Discussion

3B.2.1 Fluorescence detection of DNA i-motif formation

In the previous chapter, we have demonstrated the utility of 5-benzofuran-conjugated 2'-deoxyuridine (**1**) and uridine (**2**) in the detection of DNA and RNA G-quadruplex topologies. These results encouraged us to explore the utility of the fluorescent analogs as probes for i-motif structures in nucleic acids. Since the stability of i-motif DNA is significantly dependent on the number of cytosine residue in ON sequence, an increase in cytosine content induces i-motif formation even at neutral/slightly alkaline pH with higher stability.[7] Therefore, to assess the ability of fluorescent nucleoside **1** towards detection of i-motif formation, we have synthesized cytosine-dense DNA ON **3** by solid phase ON synthesis (Figure 1). We purposely chose to modify the dT residues as modifications on cytosine could alter the efficiency of i-motif formation. The purity and integrity of full-length modified DNA ON was confirmed by HPLC and mass analysis, respectively (Figure 2 and Table 1).



3 5' d(TCC CCC TCC CCC **1**CC CCC TCC CCC) 3'
4 5' d(CCC TAA CCC **1**AA CCC TAA CCC TAA) 3'
5 5' d(GCG ATC AC**1** CAC TAG CG) 3'
6 5' d(TCC CCC TCC CCC TCC CCC TCC CCC) 3'
7 5' d(CCC TAA CCC TAA CCC TAA CCC TAA) 3'
8 5' r(CCC UAA CCC **2**AA CCC UAA CCC UAA) 3'
9 5' r(CCC UAA CCC UAA CCC UAA CCC UAA) 3'

Figure 1. Chemical structure of 5-benzofuran-conjugated 2'-deoxyuridine (**1**) and uridine (**2**) used in the synthesis of i-motif-forming DNA and RNA ONs (**3**, **4** and **8**). **4** and **8** are fluorescent h-telo DNA and RNA ONs, respectively, containing the modification at one of the loop positions. **7** and **9** are control unmodified h-telo DNA and RNA ONs, respectively. **5** is control fluorescent DNA ON which does not forms i-motif even at acidic pH.

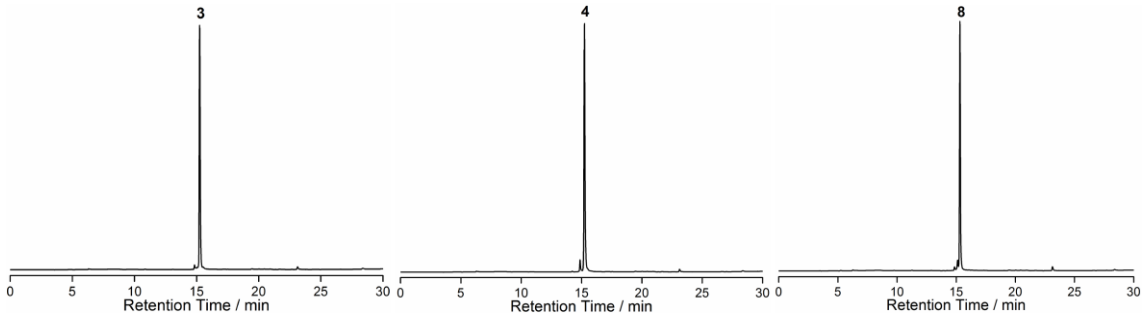


Figure 2. HPLC chromatograms of PAGE purified fluorescent ONs **3**, **4** and **8** at 260 nm. Mobile phase A = 100 mM triethylammonium acetate buffer (pH 7.6), mobile phase B = acetonitrile. Flow rate = 1 mL/min. Gradient = 0–10% B in 10 min and 10–100% B in 20 min.

Table 1. ε_{260} and mass data of modified DNA and RNA ONs.

Modified ON	ε_{260} ($M^{-1}cm^{-1}$)[a]	Calcd. mass	Observed mass
3	178813	7040.6	7041.5
4	224613	7232.8	7233.0
5	161613	5257.5	5257.9
8	230253	7571.1	7571.1

[a] The molar absorption coefficient of modified oligonucleotides were determined by using OligoAnalyzer 3.1. The extinction coefficient of nucleoside **1** (ε_{260} = 12613 $M^{-1}cm^{-1}$) and **2** (ε_{260} = 17253 $M^{-1}cm^{-1}$) were used in place of thymidine and uridine respectively.

Steady-state fluorescence analysis of DNA ON **3** was performed as function of pH in Tris-HCl buffer solution. When excited at 330 nm, DNA ON **3** at slightly basic pH (8.6)

displayed strong emission band centered at 440 nm (Figure 3A). As the pH was lowered from 8.6 to 5.0 DNA ON **3** showed a significant pH-dependent fluorescence quenching (~ 19 fold) without apparent changes in emission maximum (Figure 3A). Upon plotting the normalised fluorescence intensity at emission maximum against pH, a sigmoidal curve was observed, which yielded a *ipH* 7.2 (Figure 3B, Table 5).

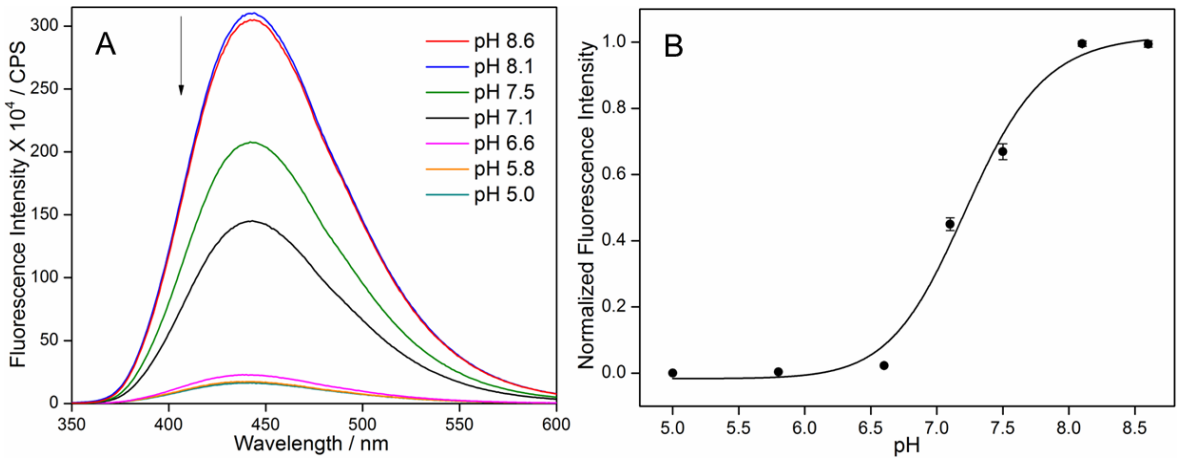


Figure 3. (A) Fluorescence spectra of DNA ON **3** (1 µM) in Tris-HCl buffer (100 mM) at different pH. (B) Curve fit obtained by plotting normalized fluorescence intensity at λ_{em} = 440 nm against pH was used to determine the transition point of i-motif formation.

Furthermore, time-resolved fluorescence analysis revealed distinct decay kinetics for DNA ON **3,** as we changed the pH of sample. DNA ON **3** at pH 8.6 was mostly in single stranded form and exhibited mono-exponential and highest lifetime (5.66 ns). While at pH 5.0 it formed stable i-motif assembly, which displayed a significantly lower and bi-exponential decay kinetics (2.11ns) (Figure 4A, Table 2). This trend in lifetime was found to be consistent with the trend in emission intensity observed in steady-state experiments. Importantly, a saturation lifetime isotherm yielded an *ipH* 7.0 (Figure 4B, Table 5), which was similar to the one obtained by steady-state analysis. These observations highlight the consistency of the fluorescence nucleoside analog **1** in reporting the pH dependent structural transition of DNA ON **3** from single strand to intercalated i-motif structure. The remarkable difference in the emission intensity and lifetime exhibited by DNA ON **3** at different pH is likely due to significant conformational change leading to effective stacking of the fluorophore **1** between adjacent nucleobases as a consequence of i-motif formation.

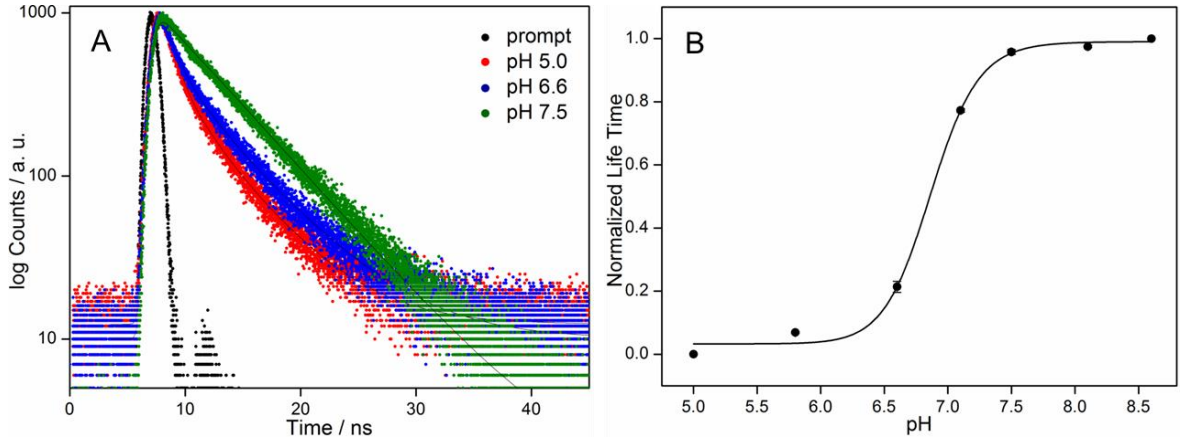


Figure 4. (A) Time-resolved fluorescence spectra of DNA ON **3** (1 µM) in Tris-HCl buffer (100 mM) at representative pH. (B) Curve fit obtained by plotting normalized lifetime against pH was used to determine the transition point of i-motif formation.

Table 2. Lifetime properties of DNA ON **3** in Tris-HCl buffer (100 mM) at different pH

ON 3 at	λ_{em} (nm)	τ_1[a] (ns)	τ_2[a] (ns)	τ_{Ave}[b] (ns)
pH 5.0	440	1.18 (0.74)	4.70 (0.26)	2.11
pH 5.8	440	1.24 (0.70)	4.98 (0.30)	2.35
pH 6.6	440	1.40 (0.64)	5.52 (0.36)	2.87
pH 7.1	440	1.94 (0.26)	5.88 (0.74)	4.85
pH 7.5	440	5.50 (1.00)	-	5.50
pH 8.1	440	5.58 (1.00)	-	5.58
pH 8.6	440	5.66 (1.00)	-	5.66

[a] Relative amplitude is given in parenthesis. [b] Standard deviations for τ_{ave} (lifetime) are ≤ 0.05 ns. Life time of DNA ON **3** at pH 5.0 to 7.1 was found bi-exponential while at pH 7.5 to 8.6 was mono-exponential.

Owing to biological relevance of C-rich h-telo DNA sequences $d(C_3TA_2)_n$ the folding kinetics and topologies have been well studied by NMR spectroscopy.[2a,2c,9b] It is also reported that while adopting intramolecular i-motif structure the loop region (TAA) of h-telo DNA undergoes substantial conformational change.[13] Hence, to further confirm the sequence independent conformational sensitivity of **1**, we have synthesized h-telo DNA ON **4** containing fluorescent uridine at one of the loop regions. When excited in Tris-HCl buffer at pH 8.6 h-telo DNA ON **4** exhibited a strong emission band at 440 nm. A progressive quenching in fluorescence (~ 5 fold) was observed as the pH was lowered (Figure 5A). In addition, lifetime analysis also revealed different decay kinetics for h-telo DNA **4** at different pH. It displayed longer mono-exponential life time (5.75 ns) at pH 8.6 and exhibited shorter

bi-exponential life time (4.20 ns) at pH 5.0 (Figure 6A, Table 3). Further, we have determined the *ipH* for h-telo DNA ON **4** by using steady-state and lifetime analysis data (Figure 5B, 6B respectively and Table 5). The obtained *ipH* values are comparable with h-telo DNA $(CCCTAA)_3CCC$, which was recently reported by using circular dichroisms (CD) spectroscopy.[7c] Appreciable difference in the emission intensity and lifetime exhibited by h-telo DNA ON **4** at different pH is likely due to the combinations of conformation change and effective stacking of fluorescent nucleosides with neighbouring bases as a consequence of i-motif formation at acidic pH.[2c] Importantly, control fluorescence experiments with nucleoside **1** and DNA ON **5** which does not fold into i-motif structure resulted in only marginal changes in fluorescence intensity (Figure 7). These results clearly indicate that the fluorescence properties of free nucleoside are not dependent on the pH of the solution. These results demonstrate that conformational-sensitive nucleoside analog **1** selectively detected the pH induced i-motif formation only upon incorporation into C-rich DNA ONs.

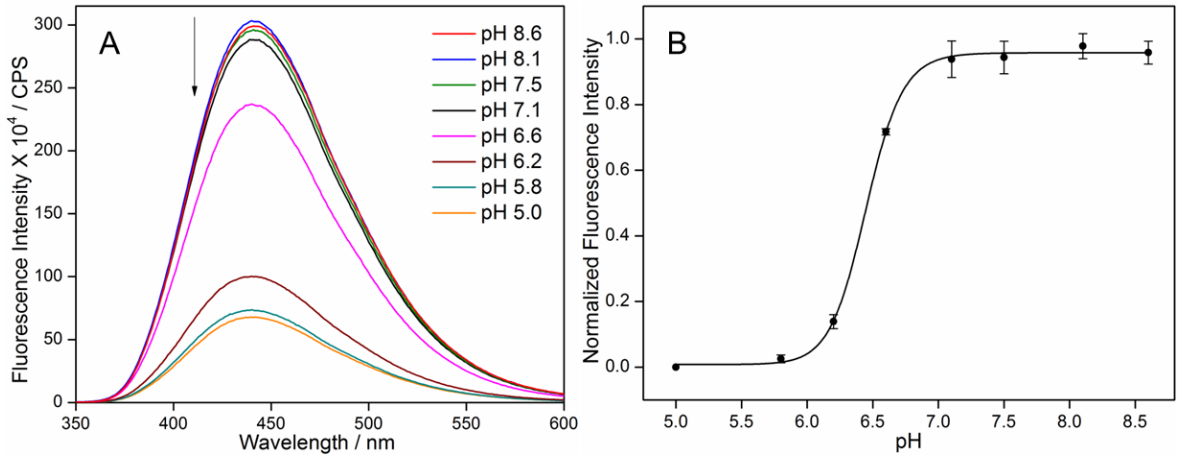


Figure 5. (A) Fluorescence spectra of h-telo DNA ON **4** (1 μM) in Tris-HCl buffer (100 mM) at different pH. (B) Curve fit obtained by plotting normalized fluorescence intensity at λ_{em} = 440 nm against pH was used to determine the transition point of i-motif formation.

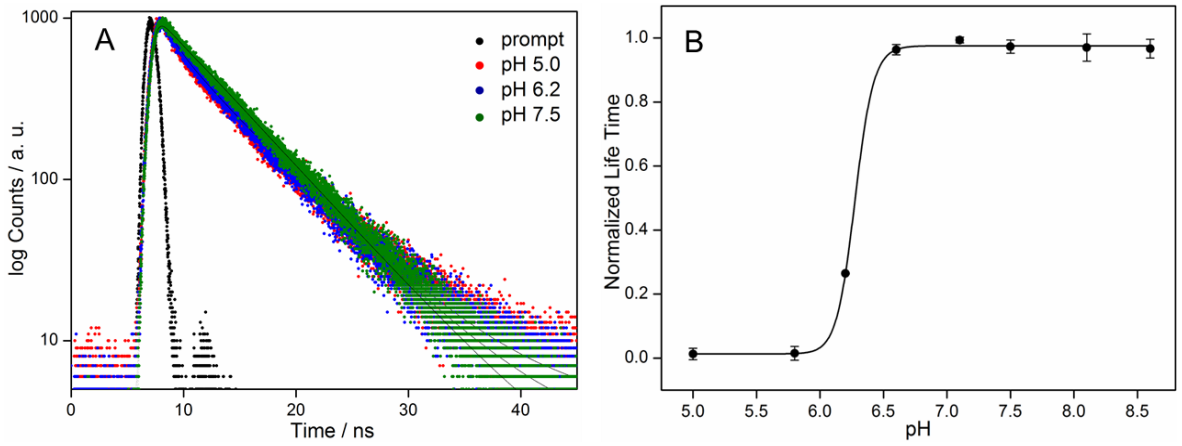


Figure 6. (A) Time-resolved fluorescence spectra of h-telo DNA ON **4** (1 μM) in Tris-HCl buffer (100 mM) at representative pH. (B) Curve fit obtained by plotting normalized lifetime against pH was used to determine the transition point of i-motif formation.

Table 3. Lifetime properties of h-telo DNA ON **4** in Tris buffer (100 mM with different pH)

Sample	λ_{em} (nm)	τ_1[a] (ns)	τ_2[a] (ns)	τ_{Ave}[b] (ns)
pH 5.0	440	1.51 (0.45)	6.41 (0.55)	4.20
pH 5.8	440	1.51 (0.44)	6.37 (0.56)	4.21
pH 6.2	440	1.76 (0.37)	6.26 (0.63)	4.61
pH 6.6	440	5.75	-	5.75
pH 7.1	440	5.80	-	5.80
pH 7.5	440	5.76	-	5.76
pH 8.1	440	5.76	-	5.76
pH 8.6	440	5.75	-	5.75

[a] Relative amplitude is given in parenthesis. [b] Standard deviations for τ_{ave} (lifetime) are ≤ 0.05 ns. Life time of h-telo DNA ON **4** in pH 5.0 to 6.2 was bi-exponential while, in pH 6.6 to 8.6 was mono exponential

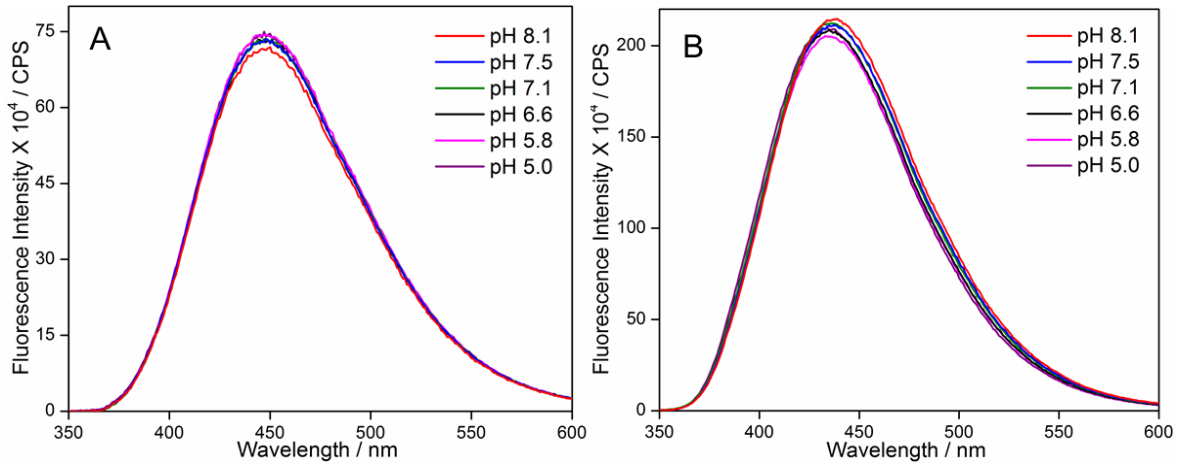


Figure 7. Fluorescence spectra of (A) Free nucleoside **1** and (B) DNA ON **5** (1 μM) in Tris-HCl buffer (100 mM) at different pH.

3B.2.2 Fluorescence detection of RNA i-motif formation

Unlike C-rich DNA, short RNA oligomers (~5 mer) do not form i-motif even in fairly acidic conditions.[14] However, longer RNA ON sequences such as $[r(UC_5)]_4$ form RNA i-motif, which is structurally similar to equivalent DNA but surprisingly less stable.[8] To study the i-motif formation in C-rich RNA, we synthesized h-telo RNA ON $r(C_3UA_2)_4$ **8** containing benzofuran-conjugated uridine **2** at one of the loop positions (Figure 1, 2 and Table 1). Like DNA ONs **3** and **4** h-telo RNA ON **8** was also subjected to steady-state fluorescence as well as lifetime analysis in Tris-HCl buffer solution as a function of pH.

When excited at its lowest energy maximum (330nm) h-telo RNA ON **8** displayed strong emission band at 440 nm at pH 7.5. Upon increasing the acidity of the sample from (pH 7.5 to 4.0) h-telo RNA ON **8** exhibited a significant pH-dependent fluorescence quenching (~ 6 fold) without apparent change in emission maximum (Figure 8A). The observed trend in fluorescence quenching of h-telo RNA ON **8** is similar to that of h-telo DNA ON **4**, which yielded a *ipH* 5.8 (Figure 8B, Table 5). This result implicates the tendency of RNA to form possible i-motif structure at relatively lower pH than the corresponding DNA ON.[8] Furthermore, time-resolved fluorescence analysis revealed distinct decay kinetics for h-telo RNA ON **8**, as we changed the pH of sample. ON **8** at pH 7.5 exhibited a highest lifetime of 6.35 ns. While at pH 4.0 it is expected to adopt i-motif form, which displayed a significantly lower lifetime (4.17 ns, Figure 9A, Table 4). This trend in lifetime was found to be consistent with the trend in emission intensity observed in steady-state experiments. Importantly, saturation lifetime profile yielded a *ipH* 5.6, which is similar to the *ipH* obtained by steady-state analysis (Figure 9B, Table 5). These observations highlight the potential of the fluorescent ribonucleoside analog **2** to probe the pH dependent conformational transition of C-rich RNA from single stranded state to possibly i-motif form. The significant difference in the emission intensity and lifetime exhibited by h-telo RNA ON **8** at different pH is probably due to the effective stacking interaction between the fluorophore **2** and adjacent nucleobases. On the basis of observed fluorescence and lifetime changes compared to the equivalent DNA ON we assume that C-rich RNA oligonucleotide also induce the pH dependent i-motif structure.

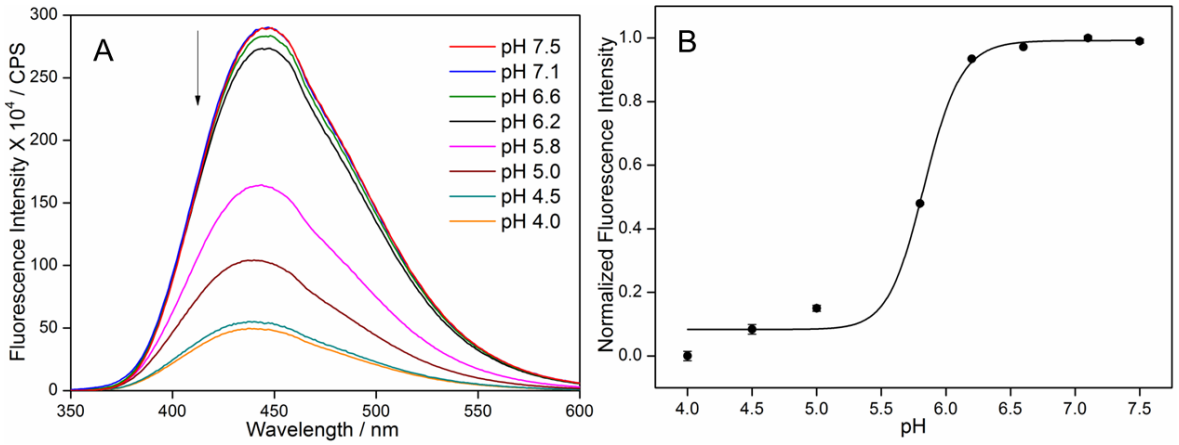


Figure 8. (A) Fluorescence spectra of h-telo RNA **8** (1 µM) in Tris-HCl buffer (100 mM) at different pH. (B) Curve fit obtained by plotting normalized fluorescence intensity at λ_{em} = 440 nm against pH was used to determine the transition point of i-motif formation.

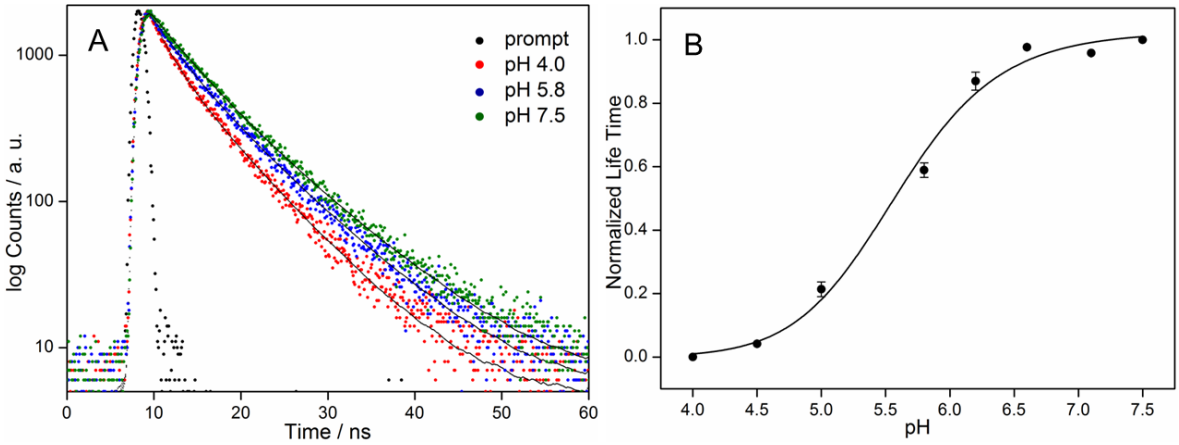


Figure 9. (A) Time-resolved fluorescence spectra of h-telo RNA ON **8** (1 µM) in Tris-HCl buffer (100 mM) at representative pH. (B) Curve fit obtained by plotting normalized lifetime against pH was used to determine the transition point of i-motif formation.

Table 5. Transitional pH of i-motif formation (*ipH*) obtained from steady-state fluorescence as well as life time analysis for DNA ONs **3**, **4** and h-telo RNA ON **8**.

ONs	*ipH* from steady-state fluorescence analysis	*ipH* from life time analysis
3	7.20 (±0.03)	7.00 (±0.01)
4	6.50 (± 0.01)	6.30 (±0.02)
8	5.80 (±0.01)	5.60 (±0.05)

3B.2.3 Circular dichroism and thermal melting studies

Benzofuran-modification in DNA and RNA could potentially affect the formation and overall stability of i-motif structures. CD analysis of modified DNA ON **3** and **4** and corresponding unmodified ON **6** and **7** at different pH confirmed the pH-dependent i-motif formation and also ascertained that the modification did not hamper i-motif structure (Figure 10). DNA ONs **3** and **6** exhibited a strong positive peak at 275 nm at pH 7.5 corresponding to a random structure. As the pH was lowered the ONs displayed noticeable changes in CD profile. At pH 5.0 these ONs exhibited a strong positive peak at 290 nm and negative peak at 265 nm characteristic of an i-motif structure.[7c] Similarly, h-telo DNA ON **4** and **7** displayed a positive peak at 270 nm and negative peak at 250 nm at pH 7.5. These ONs exhibited strong positive peak at 290 nm and negative peak at 255 nm at pH 5.0 characteristic of an i-motif structure.[7e,11,15] h-telo RNA ON **8** and **9** exhibited strong positive peak at 270 nm at pH 7.1 and displayed a noticeable pH-dependent red shifted hypochromic profile in CD. At pH 4.5 these RNA ONs displayed a positive peak at 275 nm, which could be characteristic of RNA i-motif (Figure 11). Furthermore, UV-thermal melting analysis of modified and unmodified DNA ONs at pH 5.0 resulted in similar T_m values ($\Delta T_m \leq 1°C$) suggesting that the benzofuran modification had only a minor impact on i-motif stability (Figure 12 and Table 6). DNA ONs **3** and **6** displayed remarkably higher T_m values as compared to h-telo DNA ONs **4** and **7**, which is consistent with the cytosine content. On the other hand modified h-telo RNA ON **8** and corresponding unmodified RNA ON **9** did not give reliable thermal melting profiles for the determination of T_m of respective i-motif structure.

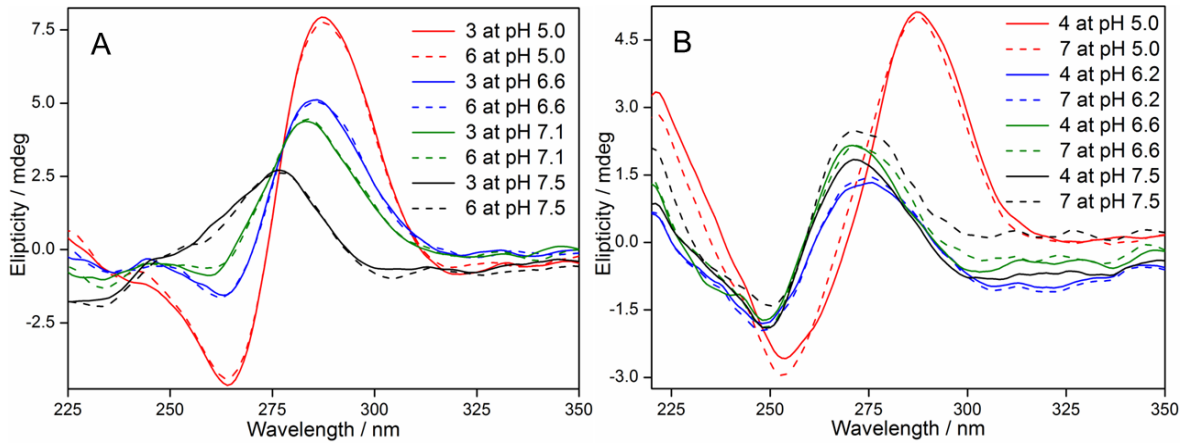


Figure 10. CD spectra of (A) fluorescently modified DNA ON **3** and control unmodified DNA ON **6** (B) fluorescently modified h-telo DNA ON **4** and control h-telo DNA ON **7** in 100 mM Tris-HCl-HCl buffer at different pH.

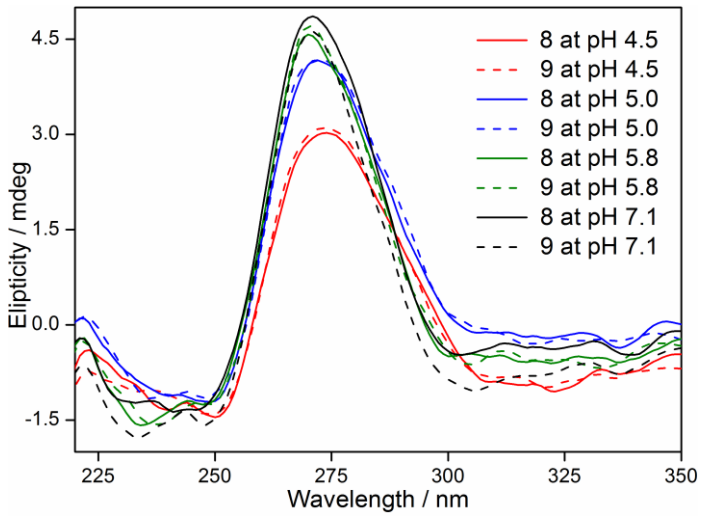


Figure 11. CD spectra of fluorescently modified h-telo RNA **8** and control h-telo DNA **9** in 100 mM Tris-HCl-HCl buffer at different pH.

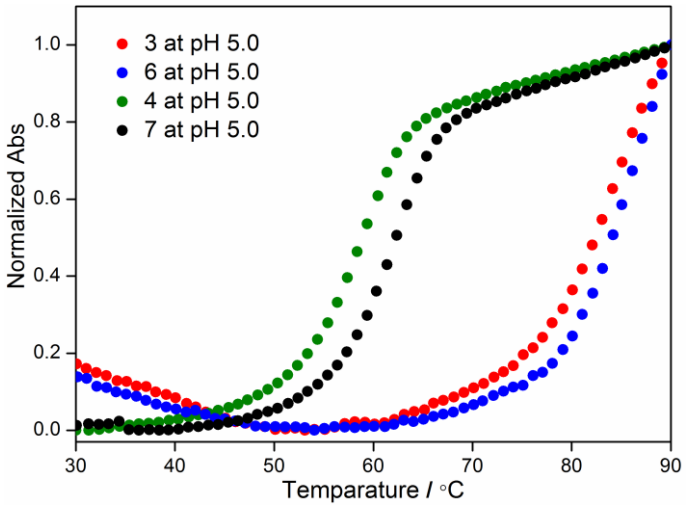


Figure 12. UV-thermal melting profiles of fluorescently modified DNA **3**, **4** and control unmodified DNA **6** and **7** in 100 mM Tris-HCl-HCl buffer at pH 5.0.

Table 6. T_m values of fluorescently modified and control unmodified DNA ONs in Tris-HCl buffer at pH 5.0.

Fluorescently-modified DNA ON	T_m (°C)	Control unmodified DNA ON	T_m (°C)
3	84 ± 0.7	**6**	85 ± 0.6
4	60 ± 0.4	**7**	61 ± 0.5

3B.3 Conclusions

5-benzofuran- conjugated 2'-deoxyuridine (**1**) and uridine (**2**) analogs upon incorporation into one of the loop residues faithfully reported the pH-dependent formation of DNA and RNA i-motifs by changes in fluorescent intensity and excited-state lifetime. The incorporation of modified nucleosides into i-motif forming ON sequences did not hamper the structure and stability of the i-motifs. Notably, these probes enabled the determination of i-motif transition pH (*ipH)* based on cytosine content in the ONs and nucleic acid type (DNA *vs.* RNA). Currently, we are exploring the possibility of using these fluorescent probes in studying protein/ligand binding to biologically relevant i-motif structures.

3B.4 Experimental section

3B.4.1 Materials

Benzofuran-modified nucleoside analogoues **1**, **2** and corresponding phosphoramedite substrate for solid phase oligonucleotide (ON) synthesis were synthesis as per our earliear reports [12]. *N*-benzoyl-protected dA, dT, *N*,*N*-dimethylformamidine-protected dG and *N*-acetyl-protected dC phosphoramidite substrates for solid phase DNA synthesis were purchased from ChemGenes. Solid supports for the solid phase DNA synthesis were obtained from ChemGenes. TBDMS-protected ribonucleoside phosphoramidite substrates and solid supports required for the solid-phase RNA synthesis were purchased from Glen Research. All other reagents for solid phase oligonucleotides synthesis were obtained from ChemGenes. Synthetic DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc. and purified by polyacrylamide gel electrophoresis (PAGE) under denaturing condition and desalted on Sep-Pak Classic C18 cartridges (Waters Corporation). Custom synthesized oligoribonucleotide was purchased from Dharmacon RNAi Technologies, deprotected according to the supplier's protocol, PAGE-purified and desalted on Sep-Pak Classic C18 cartridges. Chemicals (BioUltra grade) for preparing buffer solutions were purchased from Sigma-Aldrich. Autoclaved water was used for preparation of all buffer solution and fluorescence analysis.

3B.4.2 Instrumentation

Mass measurements were recorded on Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer. Modified DNA and RNA oligonucleotides were synthesized on Applied Biosystems RNA/DNA synthesizer (ABI-394). Absorption spectra were recorded on a PerkinElmer, Lambda 45 UV-Vis spectrophotometer. RP-HPLC analyses were performed using Agilent Technologies 1260 Infinity. UV-thermal melting studies of oligonucleotides were performed on a Cary 300Bio UV-Vis spectrophotometer and CD analysis was performed on JASCO J-815 CD spectrometer. Steady-state fluorescence and time-resolved experiments were carried out in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3).

3B.4.3 Solid-Phase synthesis and purification of modified DNA and RNA ONs

Benzofuran-modified DNA ONs **3, 4, 5** and RNA ON **8** were synthesized and purified by following our earlier reports [12]. See Table 1 for ε_{260} and MALDI-MS data of the modified oligonucleotides and Figure 2 for the HPLC chromatogram of ON **3, 4** and **8.**

3B.4.4 MALDI-TOF MS of DNA and RNA ONs

Molecular weight of benzofuran-modified DNA and RNA ONs were determined using Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer. 2 µL of the modified ON (200 µM) was combined with 1 µL of ammonium citrate buffer (100 mM, pH 9), 1.5 µL of a DNA standard (200 µM) and 4 µL of saturated 3-hydroxypicolinic acid solution. The sample was desalted with an ionexchange resin (*Dowex 50W-X8*, 100-200 mesh, ammonium form), spotted on the MALDI plate, and was air dried. The resulting spectrum was calibrated relative to an internal DNA ON standard.

3B.4.5 Photophysical analysis of fluorescently modified ONs

Steady-state fluorescence

Fluorescently modified DNA ONs **3**, **4** and RNA ON **8** (1.0 µM) were mixed in 100 mM Tris buffer containing different pH (4.0 to 8.6) and kept for 1 h at 20 °C. Similarly samples of DNA ON **5** which doesn't forms i-motif structure and nucleoside **1** were prepared in 100 mM Tris buffer containing different pH (4.0 to 8.6). Fluorescently modified ONs and nucleoside were excited at 330 nm with excitation and emission slit widths of 2 and 5 nm, respectively.

All fluorescence experiments were performed in triplicate in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 at 20 °C.

Time-resolved fluorescence

Fluorescently modified DNA ONs **3**, **4** and RNA ON **8** (1.0 μM) were mixed in 100 mM Tris buffer containing different pH (4.0 to 8.6) and lifetimes of the ONs were determined using TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3). Fluorescently modified ONs were excited using 339 nm LED source (IBH, UK, NanoLED-339L) and fluorescence signal at respective emission maximum was collected. All Experiments were performed in duplicate and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay kinetics for oligonucleotide constructs were found to be either mono-exponential or bi-exponential with χ^2 (goodness of fit) values very close to unity.

3B.4.6 Circular dichroism analysis

Fluorescently modified DNA ONs **3**, **4**, controlled unmodified DNA ONs **6**, **7** and Fluorescently modified RNA ON **8** (8.0 μM) were mixed in 100 mM Tris buffer containing different pH (4.0 to 8.6). CD spectra were collected from 350 to 225 nm on a Jasco J-815 CD spectrometer using 1 nm bandwidth at 20 °C. Experiments were performed in duplicate wherein each spectrum was an average of five scans. The spectrum of buffer was subtracted from all sample spectra.

3B.4.7 Thermal melting analysis

Fluorescently modified DNA ONs **3**, **4**, controlled unmodified DNA ONs **6**, **7** and Fluorescently modified RNA ON **8** (8.0 μM) were mixed in 100 mM Tris buffer containing different pH (4.0 to 8.6). UV-thermal melting analysis was performed in triplicate at 260 nm and 295 nm by using Cary 300Bio UV-Vis spectrophotometer.

3B.4.8 Determination of transitional pH

Transitional pH of i-motif formation (*ipH*) is considered as a pH at which 50% of the ON is in folded state (i-motif form). Normalized fluorescence intensity (F_N) versus pH plots were fitted using Hill equation (OriginPro 8.5.1) to determine the transitional pH.[16]

$$F_N = \frac{F_i - F_s}{F_0 - F_s}$$

F_i is the fluorescence intensity at each individual pH. F_0 and F_s are the fluorescence intensity in the single stranded state (s) and at completely folded state (i-motif), respectively. n is the Hill coefficient associated with the folding.

$$F_N = F_0 + (F_s - F_0)\left(\frac{[s]^n}{[ipH]^n + [s]^n}\right)$$

3B. 5 References

[1] K. Gehring, J.-L. Leroy, M. Guéron, *Nature* **1993**, *363*, 561−565.

[2] a) J.-L. Leroy, M. Guéron, J.-L. Mergny, C. Hélène, *Nucleic Acids Res.* **1994**, *22*, 1600−1606; b) G. Manzini, N. Yathindra, L. E. Xodo, *Nucleic Acids Res.* **1994**, *22*, 4634−4640; c) A. T. Phan, M. Guéron, J.-L. Leroy, *J. Mol. Biol.* **2000**, *299*, 123−144; d) T. Simonsson, M. Pribylova, M. Vorlickova, *Biochem. Biophys. Res. Commun.* **2000**, *278*, 158–166; e) S. Kendrick, Y. Akiyama, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2009**, *131*, 17667–17676.

[3] a) T. A. Brooks, L. H. Hurley, *Nat. Rev. Cancer* **2009**, *9*, 849–861; b) T. A. Brooks, S. Kendrick, L. Hurley, *FEBS J.* **2010**, *277*, 3459–3469.

[4] a) L. Lacroix, H. Lienard, E. Labourier, M. Djavaheri-Mergny, J. Lacoste, H. Leffers, J. Tazi, C. Hélène, J.-L. Mergny, *Nucleic Acids Res.* **2000**, *28*, 1564–1575; b) J. F. Cornuel, A. Moraillon, M. Guéron, *Biochimie* **2002**, *84*, 279–289; c) D. J. Uribe, K. Guo, Y. J. Shin, D. Sun, *Biochemistry* **2011**, *50*, 3796–3806; d) Y. M. K. Yoga, D. A. K. Traore, M. Sidiqi, C. Szeto, N. R. Pendini, A. Barker, P. J. Leedman, J. A. Wilce, M. C. J. Wilce, *Nucleic Acids Res.* **2012**, *40*, 5101–5114.

[5] a) Y. Chen, K. Qu, C. Zhao, L. Wu, J. Ren, J. Wang, X. Qu, *Nat. Commun.* **2012**, *3*, DOI:10.1038/ncomms2091; b) S. Kendrick, H.-J. Kang, M. Alam, M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4161–4171; c) H.-J. Kang, S. Kendrick, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4172–4185; d) Y. Cui, D. Koirala, H. Kang, S. Dhakal, P. Yangyuoru, L. H. Hurley, H. Mao, *Nucleic Acids Res.* **2014**, *42*, 5755–5764; e) H. A. Day, P. Pavlou, Z. A. E. Waller, *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418.

[6] a) D. Liu, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2003**, *42*, 5734–5736; b) Y. Peng, X. Wang, Y. Xiao, L. Feng, C. Zhao, J. Ren, X. Qu, *J. Am. Chem. Soc.* **2009**, *131*, 13813–13818; c) Y. Krishnan, F. C. Simmel, *Angew. Chem. Int. Ed*. **2011**, *50*, 3124–3156; d) C. Chen, F. Pu, Z. Huang, Z. Liu, J. Ren, X. Qu, *Nucleic Acids Res.* **2011**, *39*, 1638–1644; e)Y. Dong, Z. Yang, D. Liu, *Acc. Chem. Res.* **2014**, *47*, 1853–1860; f) S. Haldera, Y. Krishnan, *Nanoscale* **2015**, *7*, 10008–10012.

[7] (a) J.-L. Mergny, L. Lacroix, X. G. Han, J.-L. Leroy, C. Hélène, *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898; b) M. Gueron, J.-L. Leroy, *Curr. Opin. Struct. Biol.* **2000**, *10*, 326–331, c) J. Zhou, C. Wei, G. Jia, X. Wang, Z. Feng, C. Li, *Mol. BioSyst.* **2010**, *6*, 580–586; d) S. P. Gurung, C. Schwarz, J. P. Hall, C. J. Cardin, J. A. Brazier, *Chem. Commun.* **2015**, *51*, 5630–5632; e) L. Lannes, S. Halder, Y. Krishnan, H. Schwalbe, *ChemBioChem* **2015**, *16*, 1647 – 1656.

[8] a) D. Collin, K. Gehring, *J. Am. Chem. Soc.* **1998**, *120*, 4069–4072; b) K. Snoussi, S. Nonin-Lecomte, J.-L. Leroy, *J. Mol. Biol.* **2001**, *309*, 139–153.

[9] a) C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11636–11640; b) L. Cai, L. Chen, S. Raghavan, R. Ratliff, R. Moyzis, A. Rich, *Nucleic Acids Res.* **1998** , *26*, 4696–4705; c) A. L. Lieblein, J. Buck, K. Schlepckow, B. Furtig, H. Schwalbe, *Angew. Chem. Int. Ed.* **2012**, *51*, 250–253; d) S. Benabou, A. Aviñó, R. Eritja, C. González, R. Gargallo, *RSC Adv.* **2014**, *4*, 26956–26980.

[10] a) J. L. Mergny, *Biochemistry* **1999**, *38*, 1573–1581; b) S. Modi, M. G. Swetha, D. Goswami, G. D. Gupta, S. Mayor, Y. Krishnan, *Nat. Nanotechnol.* **2009**, *4*, 325–330; c) S. Surana, J. M. Bhat, S. P. Koushika, Y. Krishnan, *Nat. Commun.* **2011**, *2*, DOI: 10.1038/ncomms1340; d) J. Choi, S. Kim, T. Tachikawa, M. Fujitsuka, T. Majima, *J. Am. Chem. Soc.* **2011**, *133*, 16146–16153; e) J. Choi, T. Majima, *Photochem. Photobiol.* **2013**, *89*, 513–522.

[11] a) I. J. Lee, M. Park, T. Joo, B. H. Kim, *Mol. BioSyst.* **2012**, *8*, 486–490; b) I. J. Lee, B. H. Kim, *Chem. Commun.* **2012**, 48, 2074–2076; c) J. Choi, A. Tanaka, D. W. Cho, M. Fujitsuka, T. Majima, *Angew. Chem. Int. Ed.* **2013**, *52*, 12937–12941; d) M. Guillaume, N. W. Luedtke, *J. Am. Chem. Soc.* **2015,** *137*, 699–707.

[12] a) A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J.* **2011**, *17*, 12820–12827; b) A. A. Tanpure, S. G. Srivatsan, *ChemBioChem*, **2012**, *13*, 2392–2399; c) A. A. Tanpure, M.

G. Pawar, S. G. Srivatsan, *Isr. J. Chem.* **2013**, *53*, 366–378; d) A. A. Tanpure, S. G. Srivatsan, *Nucleic Acids Res.* **2015**, *doi: 10.1093/nar/gkv743.*

[13] A. T. Phan, J.-L. Leroy, *J. Biomol. Struct. Dyn.* **2000**, 17, 245–251.

[14] L. Lacroix, J.-L. Mergny, J.-L. Leroy, C. Hélène, *Biochemistry* **1996**, *35*, 8715–8722.

[15] C. Chen, M. Li, Y. Xing, Y. Li, C.-C. Joedecke, J. Jin, Z. Yang, D. Liu, *Langmuir* **2012**, *28*, 17743−17748.

[16] G. Vernier, A. Chenal, H. Vitrac, R. Barumandzadhe, C. Montagner, V. Forge, *Protein Science* **2007**, *16*, 391–400.

Chapter 4

Synthesis, photophysical properties and incorporation of a highly emissive and environment-sensitive uridine analogue based on the lucifer chromophore

4.1 Introduction

Base-modified fluorescent nucleoside probes that photophysically report subtle changes in their neighboring base environment have been very useful in designing bioanalytical assays to study the structure and function of nucleic acids.[1,2] Over the years several such responsive fluorescent nucleoside analogue probes have been appropriately implemented in assays to (i) detect base pair mismatches,[3] abasic sites,[4] mutagenic nucleobase modifications,[5] electron transfer process[6] and (ii) study nucleic acid topologies,[7] enzyme activities,[8] nucleic acid-small molecule and nucleic acid-protein complexes.[9] However, baring a very few examples, the majority of emissive nucleoside analogs, when incorporated into ON sequences, experience drastic fluorescence quenching due to interactions with neighbouring bases, which considerably limit their applications.[10] Notably, purines, especially guanine (G), quench the fluorescence of most of the fluorophores by photoinduced electron-transfer process, and hence, turn-on fluorescence detection of purine repeats and mismatches has been quite a challenge.[11-13] G-repeats (e.g., G-triad) are biologically important structural motifs and are frequently found in nucleic acids (e.g., telomeric and oncogene promoter regions). Hence, development of fluorescence probes that can signal the presence of G-repeats with an enhancement in fluorescence intensity is highly desired.[13] Therefore, much of the recent efforts in the development of new fluorescent nucleoside analogs are dedicated towards designing environment-sensitive analogs that have excitation and emission maximum in the visible region and high fluorescence efficiency within ONs, with the view of implementing them in both in vitro and in vivo assays.[14,15-18]

Kool and co-workers have assembled a library of DNA-like chain, called oligodeoxyfluorosides, containing different PAH fluorophores, which display large Stokes shifts and wide array of quantum yields and emission wavelengths.[16] These oligodeoxyfluorosides with tunable photophysical properties have been implemented in the detection of molecular species in solution and in the multiplexed imaging of cells. In a somewhat similar approach, siRNAs labeled with multiple phenylpyrrolocytidine residues have been used to monitor the trafficking and silencing activity of siRNA inside living cells using fluorescence microscopy.[17] More recently, a quadracyclic adenine analogue and 4-aminophthalimide C-nucleoside have been introduced as fluorescent nucleoside surrogates, which when incorporated into ONs form stable duplexes and retain a reasonable fluorescence efficiency as compared to most other fluorescent nucleoside analogs.[18] Predicting the photophysical behaviour of nucleoside analogs, free or incorporated into ONs, based on their

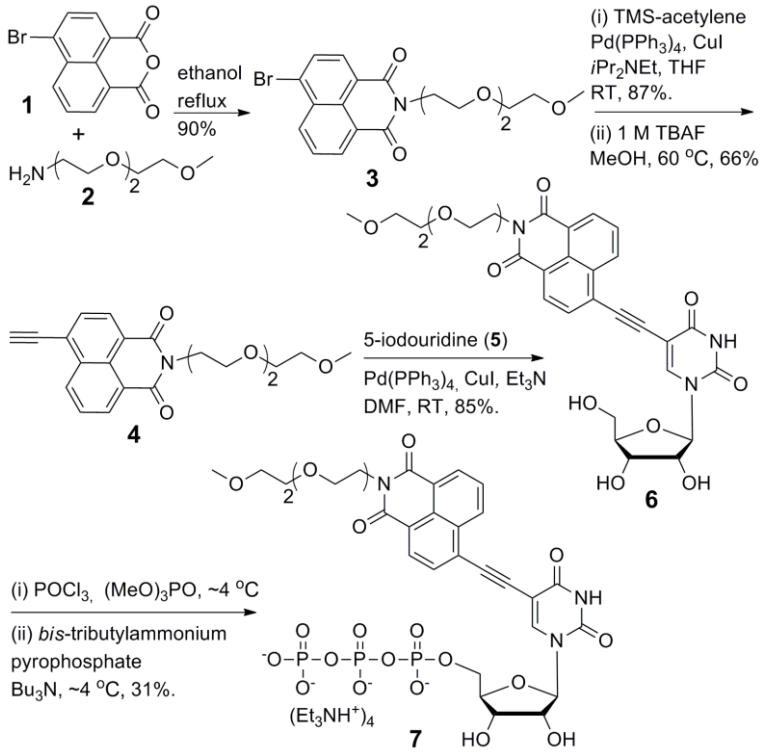
structure is not simple, and hence, probe design and implementation largely remain an empirical task. As a part of the continued efforts to develop new fluorescence probes with useful properties, here, we report the synthesis and photophysical properties of a highly emissive ribonucleoside analogue **6** obtained by conjugating the Lucifer chromophore, napthalimide core, at the 5-position of uridine via an ethynyl linker. The napthalimide-modified uridine analogue exhibits both excitation and emission maximum in the visible region and excellent fluorescence solvatochromism. Notably, the fluorescence properties of emissive nucleoside incorporated into ONs are sensitive to flanking bases and base pair substitutions. In particular, the emissive nucleoside signals the presence of purine repeats (dG and dA) with a significant increase in fluorescence intensity, a property seldom exhibited by most of the fluorescent nucleoside analogs.[11]

4.2 Results and Discussion

4.2.1 Synthesis and photophysical properties of naphthalimide-modified uridine 6

The design of fluorescent probe **6** is based on the 4-amino-1,8-naphthalimide core present in Lucifer dyes, which absorbs light in the visible region and exhibits large Stoke shifts and high quantum yields.[19,20] The photophysical properties of parent napthalimide, containing the electron-withdrawing imide moiety, largely depends on the substituent present at the 4-position of the dye. Typically, derivatives containing electron donating groups at this position exhibit very strong fluorescence due to the generation of a charge transfer excited state. [21] As the fluorescence properties of naphthalimide core can be tuned by rational modification, its derivatives have been extensively used as biological markers, sensors, electroluminescent materials, to name a few.[20] Encouraged by these reports, it was postulated that coupling 4-ethynyl-1,8-naphthalimide at the 5-position of uridine would impart superior probe-like properties akin to Lucifer dyes to the nucleobase. The naphthalimide-modified ribonucleoside **6** was synthesized according to the steps illustrated in Scheme 1. 4-ethynyl-1,8-naphthalimide derivative **4** was prepared by first reacting commercially available 4-bromo-1,8-naphthalic anhydride **1** with amine-modified triethylene glycol linker **2** followed by a Sonogashira cross coupling reaction with TMS-acetylene. The silyl protecting group was then deprotected in the presence of TBAF to afford the naphthalimide derivative **4** in moderate yields. Attachment of triethylene glycol linker at the imide position was chosen to enhance the solubility of the naphthalimide derivative in polar solvents for subsequent manipulations. The fluorescent uridine analogue **6** was then obtained by performing a

palladium-catalyzed cross coupling reaction between 4-alkyne-modified naphthalimide **4** and 5-iodouridine **5**.



Scheme 1. Synthesis of fluorescent uridine analogue **6** and its corresponding triphosphate **7**.

The ground state and excited state electronic properties of nucleoside **6** have been evaluated in solvents of different polarity to test the solvatochromic behaviour of the nucleoside. While ground state electronic spectrum is marginally affected by changes in solvent polarity, the excited state properties of **6** are significantly altered by solvent polarity changes (Figure 1, Table 1). In dioxane, the least polar solvent used in this study, the nucleoside exhibits a very strong emission band (λ_{em} = 480 nm) corresponding to a quantum yield of 87% (Figure 1, Table 1). As the solvent polarity is increased from dioxane to water the nucleoside shows a remarkable 48 nm red-shifted emission band (λ_{em} = 528) and a nearly 11-fold quenching in fluorescence intensity. Excited-state lifetime measurements reveal longer lifetimes in nonpolar solvents as compared to in polar solvents (Figure 2 and Table 1). A good positive correlation between Stokes shift in different solvents and $E_T(30)$ value, a microscopic solvent polarity parameter, further indicates the sensitivity of nucleoside to microenvironment changes (Figure 3). Taken together, high fluorescence efficiency,

absorption and emission profiles in the visible region and sensitivity to polarity changes prompted us to study the fluorescence properties of the emissive nucleoside within ONs.

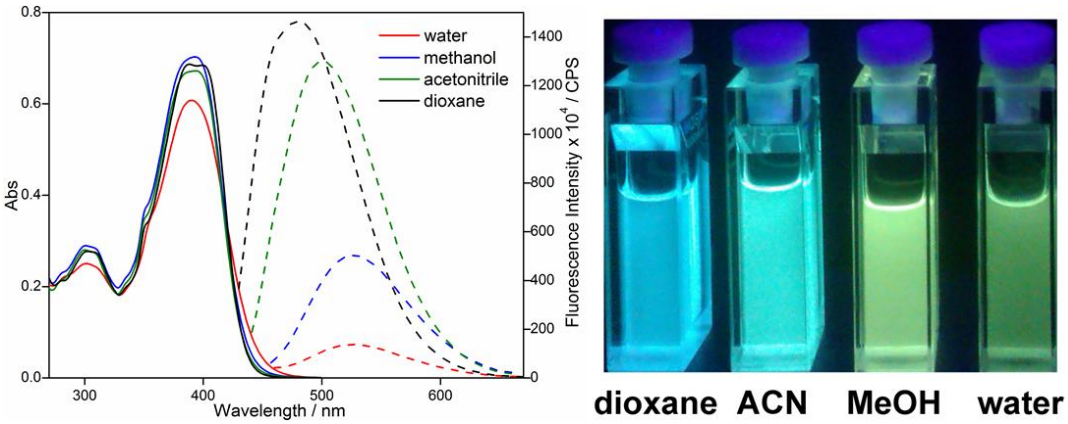


Figure 1. Absorption (25 µM, solid) and emission (5.0 µM, dash) spectra of nucleoside **6** in solvents of different polarity. All solutions for absorption and emission studies contained 2.5% and 0.5% DMSO, respectively.

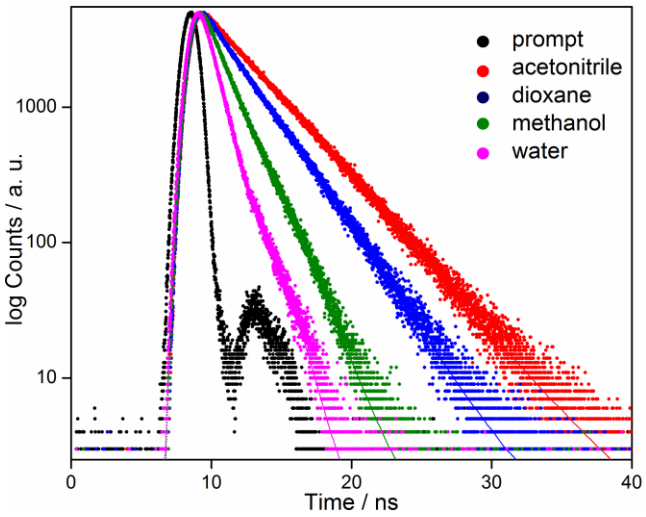


Figure 2. Excited state decay profile of ribonucleoside **6** in solvents of different polarity. Laser profile (prompt) is shown in black, curve fits are shown in solid lines.

Table 1. Photophysical properties of nucleoside **6**.

solvent	λ_{max}[a] (nm)	λ_{em} (nm)	I_{rel}[b]	Φ[c]	τ_{ave}[c] (ns)
water	391	528	1	0.11	1.03
methanol	393	527	3.7	0.24	1.66
acetonitrile	393	500	9.6	0.67	3.68
dioxane	393	480	10.8	0.87	2.78

[a] Lowest energy maximum. [b] Relative fluorescence intensity is given with respect to intensity in water. [c] Errors for Φ and τ_{ave} are ≤ 0.003 and 0.03 ns, respectively.

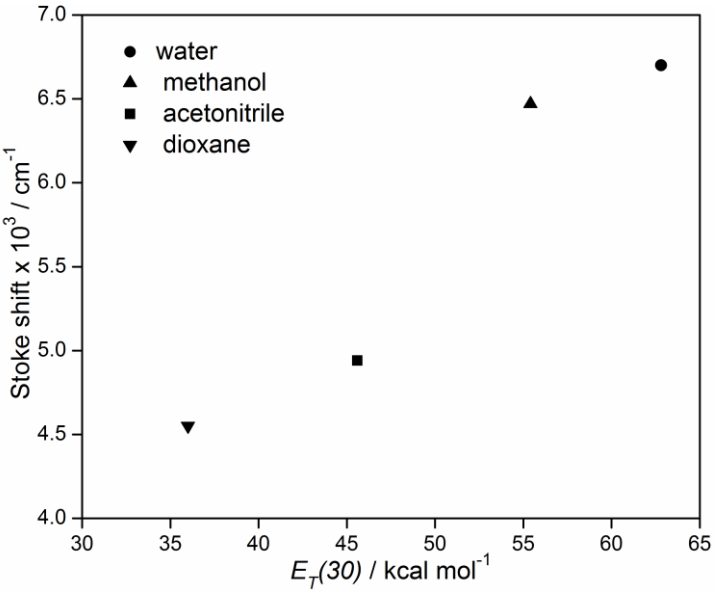


Figure 3. A plot of Stokes shift vs. $E_T(30)$ for ribonucleoside **6** in solvents of different polarity.

4.2.2 Enzymatic incorporation of nucleoside 6 into RNA ONs

Primer extension and in vitro transcription reactions in the presence of DNA and RNA polymerases, respectively, and ligation reactions have been effectively used in the synthesis of fluorescently modified DNA and RNA ONs.[9,21-24] Notably, Hirao and co-workers have elegantly used unnatural base pairs, which are orthogonal to A-T/U and G-C pairs, to site-specifically incorporate fluorescent analogs of unnatural bases into RNA ONs by in vitro transcription reactions.[9a,25] Therefore, to assess the photophysical behavior of the emissive nucleoside within ONs, we chose to incorporate the naphthalimide-modified uridine into RNA ONs by in vitro transcription reactions. The modified UTP **7** required for enzymatic incorporation was synthesized by treating nucleoside **6** with anhydrous $POCl_3$ and then with *bis*-tributylammonium pyrophosphate in ice cold conditions (Scheme 1).[26] The efficiency of T7 RNA polymerase to incorporate the modification into RNA ONs was first tested by radiolabeling experiments using a series of promoter-template DNA duplexes. The duplexes were formed by annealing a T7 RNA polymerase consensus promoter DNA ON with template ONs **T1**–**T5** (Figure 4).[27] The templates were predisposed to direct the incorporation of the monophosphate of **7** at one or two positions near the promoter region, away from the promoter region, and at the 3'-end of the transcript. In vitro transcription reactions were performed in the presence of GTP, CTP, α-^{32}P ATP and UTP/**7**, and the radiolabled transcription products were analyzed by polyacrylamide gel electrophoresis under denaturing conditions and phosphorimaged.

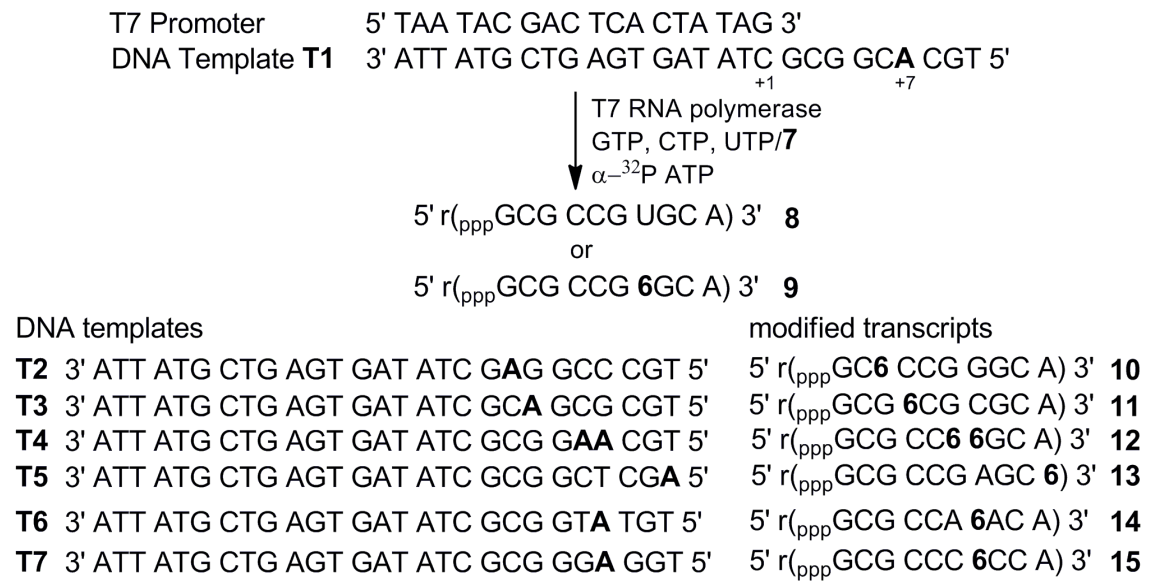


Figure 4. Incorporation of UTP **7** into RNA ONs by in vitro transcription reactions using templates **T1**–**T7**.

When a transcription reaction is performed in the presence of template **T1**, the RNA polymerase incorporates the monophosphate of **7** at the +7 position to afford the full-length transcript **9** in a good yield (Figure 5, lane 2).[28] The incorporation of a heavier ribonucleotide into the transcript is also evident from the retarded mobility exhibited by transcript **9** (Figure 5, compare lane 1 and 2). While a control reaction in the absence of UTP and **7** produces no full-length oligoribonucleotide product ruling out any misincorporation, a reaction in the presence of equal amounts of UTP and **7** indicates that the enzyme predominantly prefers natural UTP over modified UTP (Figure 5, lane 3 and 4, respectively). Attempts to introduce the modification near the promoter region (+3 and +4 positions) using templates **T2** and **T3** resulted in very low amounts of the full-length transcripts (Figure 5, lanes 6 and 8). This observation is reasonable because the enzyme often tends to be less tolerant to modifications during the initial phase of the polymerization process.[24b] Interestingly, a transcription reaction in the presence of template **T4** results in double incorporations in successive positions with a reasonable efficiency (Figure 5, lane 10). Also, when a reaction is carried out in the presence of template **T5** the RNA polymerase incorporates the modification at the 3′-end of the transcript (**13**) with an excellent efficiency (Figure 5, lane 12 and Table 2). Although, internally modified transcript **9** and 3′-end modified transcript **13** have same mass, their electrophoretic mobility is different (Figure 5, compare lane 2 and 12). Predicting the electrophoretic mobility of modified ONs based on charge and mass is not straight forward as hydrophobicity, hydration, conformation and

frictional properties of the appended group as well as its interactions with sieving medium could affect the mobility.[29] Therefore, the observed difference in mobility could be possibly due to the differences in the hydrophobicity, conformation and frictional properties of the internally- and 3′-end-attached fluorophore.

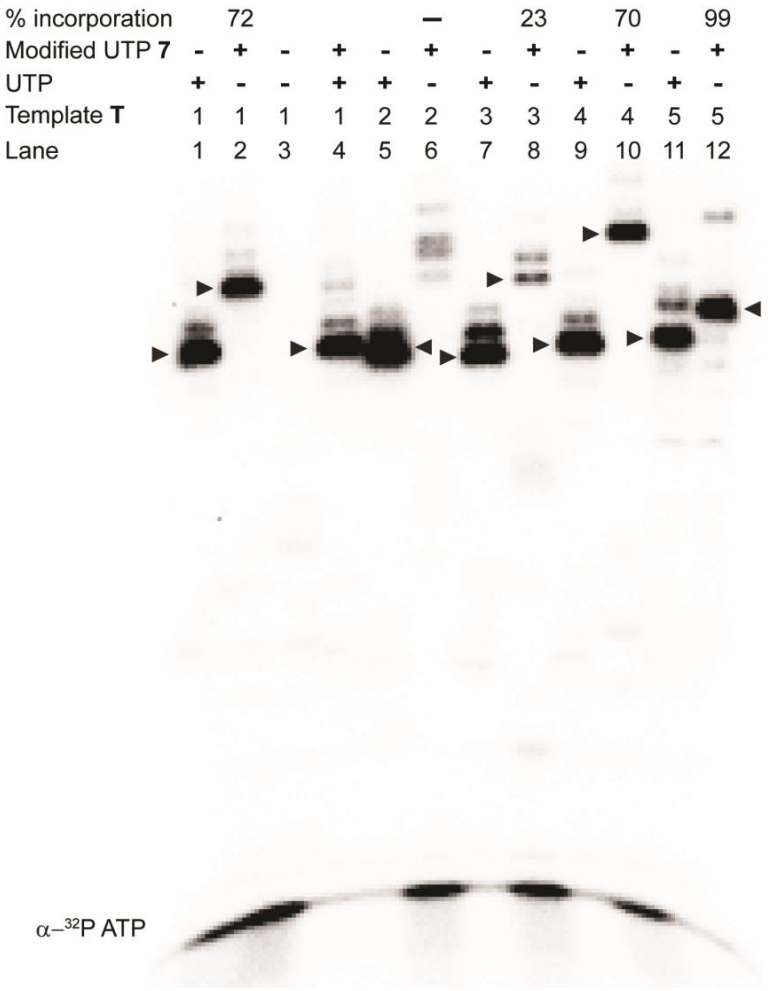


Figure 5. Phosphorimage of transcription products resolved by PAGE under denaturing conditions. Transcription reactions were performed with templates **T1**–**T5** in the presence of UTP and modified UTP **7**. % incorporation of **7** is given relative to a control reaction in the presence of natural NTPs. The modified full-length transcripts, denoted using arrow heads, were ascertained by mass analysis (Table 2). The incorporation of UTP **7** was very low in the presence of template **T2** and hence, the band corresponding to the full-length transcript **10** could not be assigned (lane 6).

RNA ONs were isolated from large-scale transcription reactions to further ascertain the incorporation of naphthalimide modification in the transcripts. Mass analysis confirmed the formation of modified full-length RNA ON products in the presence of fluorescent UTP by in vitro transcription reactions (Table 2). The presence of ribonucleoside **6** in the

transcription product was also established by an enzymatic digestion reaction. In this experiment the RNA ON **9** was subjected to digestion in the presence phosphodiesterase, alkaline phosphatase, RNase A and RNase T1. The resulting ribonucleoside products analyzed by HPLC revealed the presence of modified nucleoside in the transcript (Figure 6, Table 3).

Table 2. ε_{260} and mass data of modified RNA ONs.

RNA ON	ε_{260} ($M^{-1}cm^{-1}$)	Calcd. mass	Observed mass
9	91613	3780.3	3779.9
11	92413	3780.3	3780.8
12	90626	4106.6	4107.4
13	90313	3780.3	3780.9
14	98013	3748.3	3748.3
15	85713	3700.2	3700.7
20	91013	3462.3	3463.1

[a] The molar absorption coefficient of modified ONs were determined by using OligoAnalyzer 3.1. The extinction coefficient of nucleoside **6** (ε_{260} = 10,413 $M^{-1}cm^{-1}$) was used in place of natural uridine.

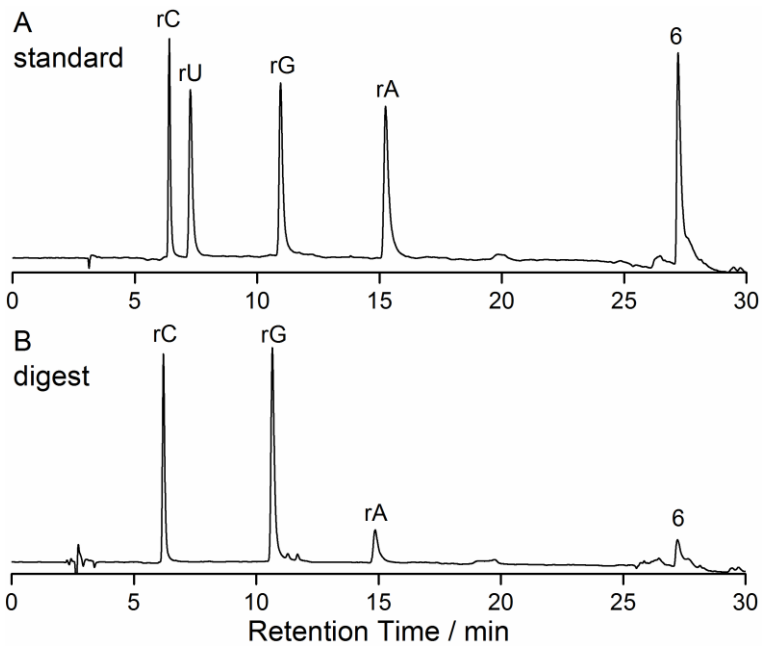


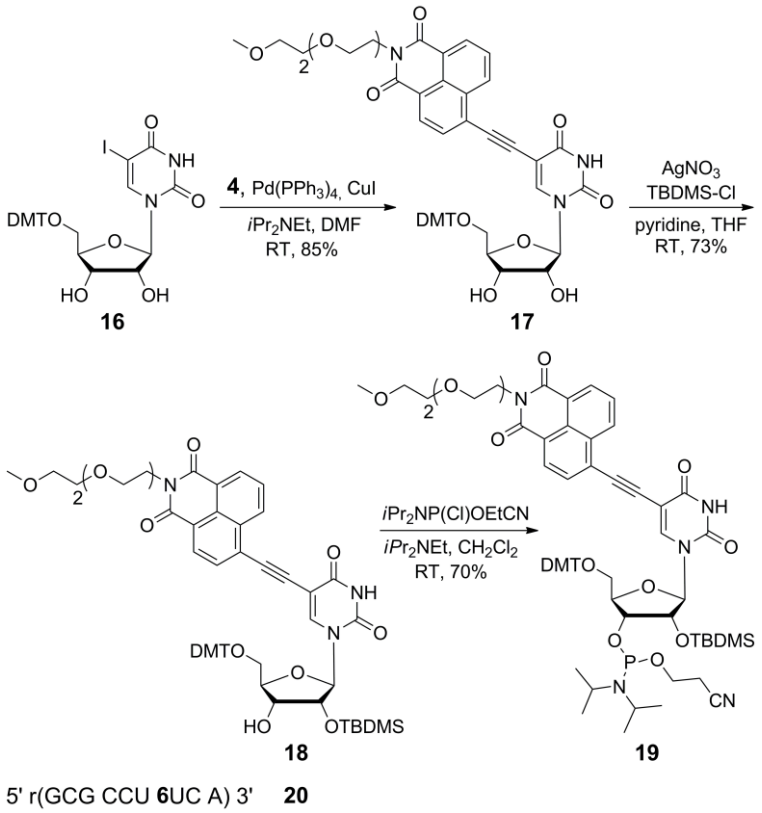
Figure 6. HPLC chromatogram of ribonucleoside products obtained from an enzymatic digestion of transcript **9** at 260 nm. a) Mixture of natural ribonucleosides and modified fluorescent ribonucleoside **6**. b) Digested ON **9**. The fractions corresponding to individual ribonucleoside fractions were further analyzed by mass spectroscopy (see Table 3).

Table 3. MALDI-TOF mass analysis of HPLC fractions of ON **9** digest.

HPLC fractions of the digest	Calcd. for	Found
rC	$C_9H_{13}N_3O_5K$ = 282.05 $[M+K]^+$	282.07
rG	$C_{10}H_{13}N_5O_5K$ = 306.08 $[M+K]^+$	306.30
rA	$C_{10}H_{13}N_5O_4$ = 290.09 $[M+Na]^+$	290.10
6	$C_{17}H_{16}N_2O_7$ = 648.16 $[M+K]^+$	648.20

4.2.3 Incorporation of nucleoside 6 into RNA ON by solid-phase synthesis.

Interactions with neighboring bases can affect the photophysical properties of an emissive nucleoside incorporated into ONs by various mechanisms.[30-32,11] Such fluorescent nucleoside analogs whose photophysical properties are modulated by subtle changes in their neighboring base environment are very useful in designing nucleic acid-based diagnostic tools.[1,2] Therefore, we decided to evaluate the effect of flanking bases and base pair substitutions on the fluorescence properties of emissive ribonucleoside **6**. RNA ONs **9**, **14**, and **15** containing **6** in between rG, rA and rC residues, respectively, were synthesized by large-scale transcription reactions (Figure 4, Table 2). However, enzymatic synthesis of an RNA ON containing modified uridine **6** in-between U residues was not attempted as it would be difficult to control the site of incorporation by in vitro transcription reactions. In addition, the RNA polymerase predominantly prefers natural UTP over modified UTP **7**, which would prevent the introduction of **6** in between U residues by transcription reaction (Figure 5, lane 4). Therefore, we synthesized the RNA ON **20**, containing **6** in between U residues, by solid-phase ON synthesis protocol. The 2'-*O*-(*tert*-butyldimethylsilyl)-protected phosphoramidite substrate **19** was synthesized by following the steps illustrated in scheme 2. The modified phosphoramidite substrate was then site-specifically incorporated into an RNA ON sequence **20** by using a standard RNA ON synthesis cycle. The ON was deprotected and purified, and the presence of fluorescent uridine analogue in the full-length product was confirmed by mass analysis (Table 2).



Scheme 2. Synthesis of naphthalimide-modified uridine phosphoramidite substrate **19** used in the solid-phase synthesis of RNA ON **20**. DMT = 4,4'-dimethoxytrityl, TBDMS = 2'-O-*tert*-butyldimethylsilyl.

4.2.4 Photophysical characterization of naphthalimide-modified RNA ONs

RNA ONs **9** and **14** in which **6** is flanked by purine residues, rG and rA, respectively, displayed a significantly quenched and slightly blue-shifted emission band as compared to the free nucleoside. While ON **20**, containing **6** in-between rU residues, exhibited discernible reduction in fluorescence intensity, ON **15**, containing **6** in between rC residues, displayed significant enhancement in fluorescence intensity as compared to the free nucleoside (Figure 7). The effect of neighbouring bases on fluorescence properties of nucleoside **6** was further evaluated by using duplexes made of ONs **9**, **14**, **15** and **20**. The duplexes were assembled by hybridizing fluorescent ONs with complementary DNA ONs in such a fashion that the emissive nucleoside was placed opposite to complementary and mismatched bases (Figure 8). The naphthalimide modification can potentially affect the hybridization efficiency of the RNA ONs, and hence, the observed fluorescence profile may be due to a combination of intact duplexes and unhybridized fluorescent ON. Therefore, prior to performing the fluorescence study, the effect of modification on the stability of duplexes was determined by thermal denaturation experiments. Although, the modification had slight destabilization

effect, the naphthalimide-modified ONs formed stable duplexes with T_m values well above room temperature (Figure 9, Table 4). Hence, this destabilization due to the presence of modification should not hamper the photophysical characterization of duplexes at room temperature.

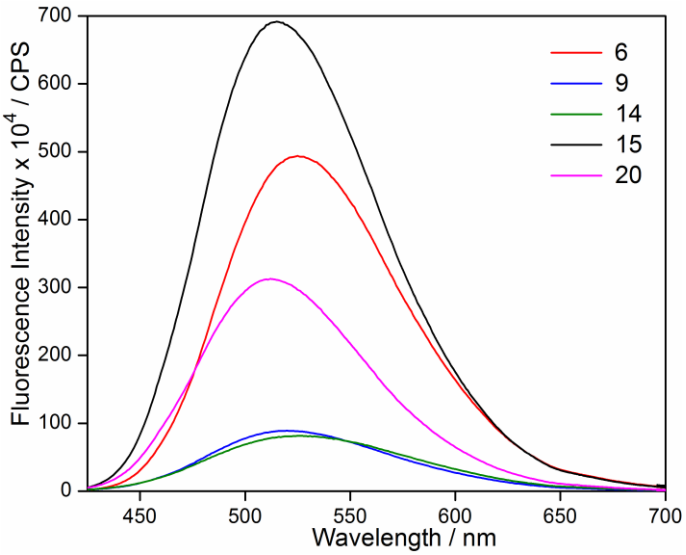


Figure 7. Emission spectra of nucleoside **6** (1 μM) and RNA ONs **9**, **14**, **15** and **20** (1 μM) containing **6** in between rG, rA, rC and rU residues, respectively. ONs were excited at 407 nm, and excitation and emission slit widths were kept at 6 and 8 nm respectively.

3' CGC GGC ACG T 5'	**9A**	3' CGC GGT ATG T 5'	**14A**
3' CGC GGC TCG T 5'	**9T**	3' CGC GGT TTG T 5'	**14T**
3' CGC GGC GCG T 5'	**9G**	3' CGC GGT GTG T 5'	**14G**
3' CGC GGC CCG T 5'	**9C**	3' CGC GGT CTG T 5'	**14C**
3' CGC GGG AGG T 5'	**15A**	3' CGC GGA AAG T 5'	**20A**
3' CGC GGG TGG T 5'	**15T**	3' CGC GGA TAG T 5'	**20T**
3' CGC GGG GGG T 5'	**15G**	3' CGC GGA GAG T 5'	**20G**
3' CGC GGG CGG T 5'	**15C**	3' CGC GGA CAG T 5'	**20C**

Figure 8. Sequence of custom DNA ONs used in this study. For example, hybridization of **9** with **9A**, **9T**, **9G** and **9C** would place the fluorescent nucleoside **6** opposite to complementary base dA and mismatched bases dT, dG and dC, respectively.

Table 4. T_m values of modified ONs duplexes[a]

Duplex	T_m [°C]	Duplex	T_m [°C]	Duplex	T_m [°C]	Duplex	T_m [°C]
9•9A	53.3	**14•14A**	43.7	**15•15A**	55.9	**20•20A**	43.1
9•9T	49.3	**14•14T**	42.2	**15•15T**	52.1	**20•20T**	41.6
9•9G	49.7	**14•14G**	40.3	**15•15G**	53.3	**20•20G**	39.1
9•9C	51.1	**14•14C**	40.9	**15•15C**	53.6	**20•20C**	39.7

[a] Errors for $T_m \leq 0.9$ °C.

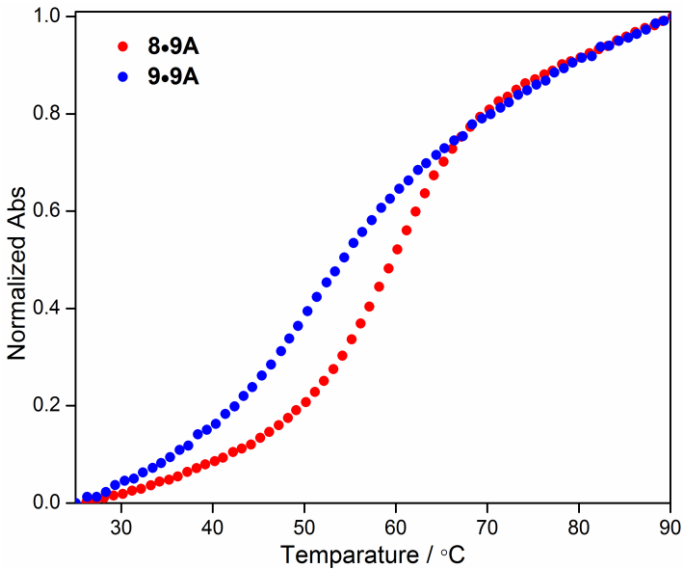


Figure 9. Representative UV-thermal melting profile of control unmodified (**8•9A**) and fluorescently modified (**9•9A**) duplexes (1 μM) at 260 nm. For ON sequence see Figure 4 and 8. T_m of duplex **8•9A** is 59.2 ± 0.6 °C and **9•9A** is 53.3 °C ± 0.7 °C. Although, the naphthalimide-modified duplex has lower T_m than corresponding unmodified duplex, the modified duplex forms a stable hybrid with T_m values well above room temperature. Hence, this destabilization due to the presence of modification should not hamper the photophysical characterization of duplexes at room temperature. See Table 4 for T_m values of other modified duplexes in which the emissive nucleoside was placed in-between different flanking bases and opposite to complementary and mismatched bases.

The nucleoside **6** flanked by rG residues, when placed opposite to complementary and mismatched bases in RNA-DNA heteroduplexes **9•9A**, **9•9T**, **9•9G** and **9•9C**, respectively, displayed further quenching in fluorescence intensity with no change in emission maximum as compared to single stranded RNA ON **9** (Figure 10A). Although, the duplexes made of ON **14** in which the emissive nucleoside is flanked by rA and placed opposite to dA, dT and dC bases showed enhancement in fluorescence intensity as compared to **14**, their overall fluorescence efficiency was weak (Figure 10B). The fluorescence quenching exhibited by nucleoside analogue **6** upon incorporation into ONs duplexes is a common feature, which is also shown by most of the other nucleoside analogs. This quenching in intensity could be due to electron transfer process between the modified base and neighbouring bases.[31] However, partial stacking interaction and relative conformation of the napthalimide core with respect to the uracil base in different base environment could also influence the fluorescence outcome.[30,32]

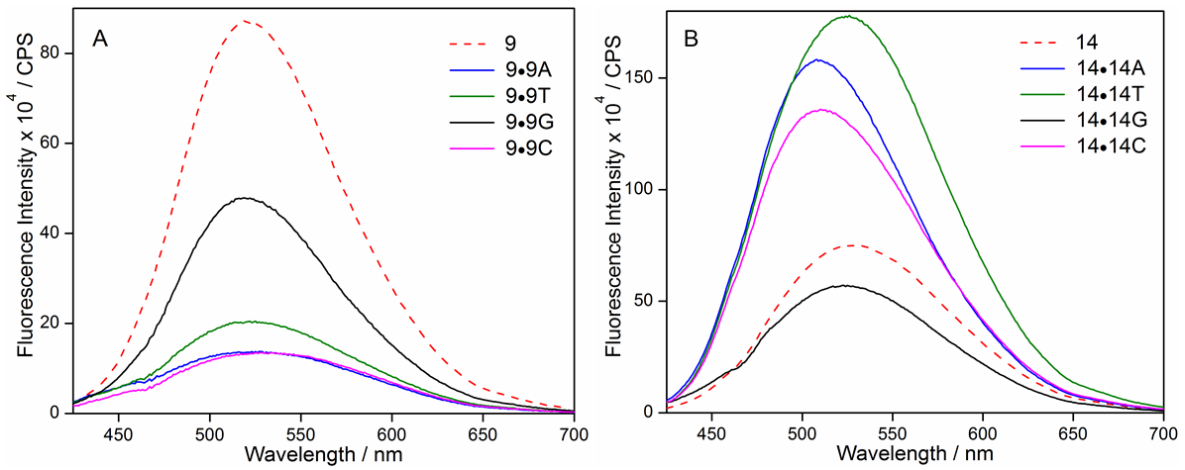


Figure 10. (A) Emisson spectra (1 μM) of ON **9** and duplexes made of **9**. (B) Emisson spectra (1 μM) of ON **14** and duplexes made of **14**. ONs were excited at 407 nm, and excitation and emission slit widths were kept at 6 and 8 nm, respectively.

The nucleoside analogue in-between rC residues in RNA ON **15** showed an intense band with emission maximum at 515 nm (Figure 11A). Surprisingly, the emissive uridine analogue placed opposite to dG residue in a duplex **15•15G** (containing consecutive dG residues) displayed more than 2-fold enhancement in fluorescence intensity as compared to when it was placed opposite to its complementary base (dA) in duplex **15•15A** (Figure 11A). However, the enhancement in fluorescence intensity was not very significant when the nucleoside analogue was placed opposite to dT and dC mismatches in duplexes **15•15T**, and **15•15C**, respectively. We were doubly surprised when the fluorescent uridine analogue incorporated in-between rU residues and placed opposite to complementary (**20•20A**) and mismatched bases (**20•20T**, **20•20G, 20•20C**) displayed a remarkable enhancement (> four-fold) in fluorescence intensity with no apparent change in emission maximum as compared to the single stranded ON **20** (Figure 11B). This observation is noteworthy because the majority of fluorescent nucleoside analogs including 2-aminopurine, upon incorporation into ONs, experience drastic quenching in fluorescence intensity due to stacking interactions and or electron transfer process.[2,11,30] In particular, this effect is more pronounced when the nucleoside analogue is placed in the vicinity of purines bases.[17]

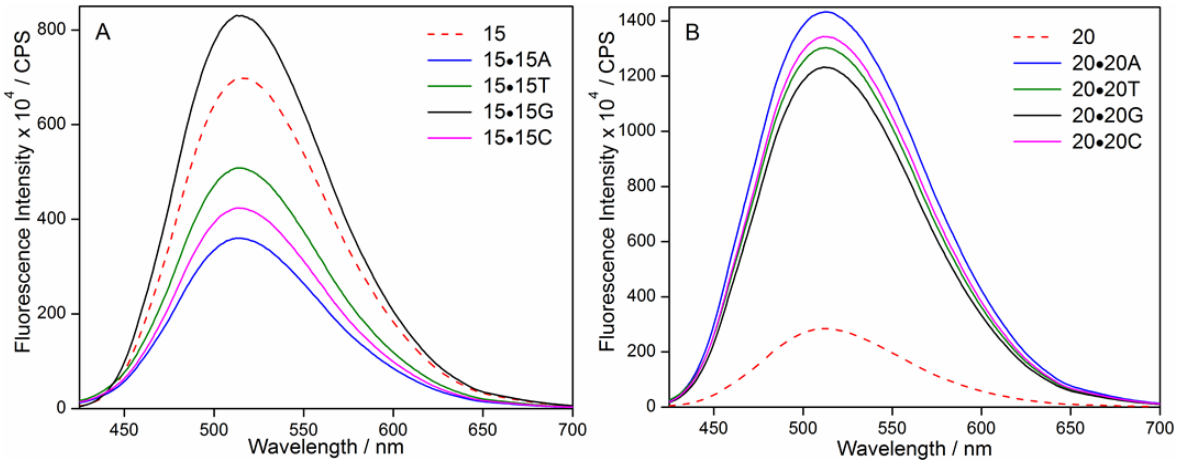


Figure 11. (A) Emisson spectra (1 μM) of ON **15** and duplexes made of **15**. (B) Emisson spectra (1 μM) of ON **20** and duplexes made of **20.** ONs were excited at 407 nm and excitation and emission slit widths were kept at 6 and 8 nm, respectively.

The fluorescence profile exhibited by nucleoside **6** in different nucleobase environment could be likely due to alterations in the relative conformation of the naphthalimide moiety with respect to the uracil base, electron transfer process between the emissive nucleobase and neighbouring bases and solvation-desolvation effects.[30-32,11] Together, these results clearly indicate that the fluorescence properties of emissive base are sensitive to flanking bases and base pair substitutions, which essentially arise due to changes in various interactions between the modified base and neighbouring bases. The fact that the fluorescence of the majority of fluorophores is quenched by purine residues,[11] the enhancement in fluorescence intensity exhibited by naphthalimide-modified nucleoside, when placed in a G-rich and A-rich environment, is rare and useful.[13] Such an analogue probe with an excitation and emission maximum in the visible region and ability to exhibit enhanced fluorescence efficiency in certain sequence context could be highly useful in devising hybridization assays to detect nucleobase repeats in nucleic acids (e.g., G-repeats and A-repeats).

4.3 Conclusions

Nucleic acids perform their function by either interacting with proteins, nucleic acids or small molecule metabolites, and during such recognition events they undergo conformational changes both at the global and nucleoside levels.[33] The changes in nucleoside conformation (e.g., near the recognition site) also alter its surrounding physical properties and interactions with adjacent nucleosides. Hence, environment-sensitive fluorescent nucleoside analogs with

properties such as excitation and emission maximum in the visible region and high quantum yields within ONs, which are suitable for both in vitro and cell-based applications, are highly desired. In this regards, fluorescent ribonucleoside **6**, based on the Lucifer chromophore, represents a new type of environment-sensitive nucleoside analogue with excitation and emission maximum in the visible region and a high quantum yield. The triphosphate and phosphoramidite derivative of naphthalimide-modified nucleoside acts as a good substrate for the synthesis of fluorescent RNA ONs by in vitro transcription reaction as well as by solid-phase ON synthesis protocol, respectively. Furthermore, the emissive nucleoside incorporated into ON duplexes exhibits appreciable fluorescence efficiency, and is also responsive to changes in its flanking bases and base pair substitutions. These favourable photophysical properties and amicability to enzymatic and chemical incorporation underscore the potential of naphthalimide-modified nucleoside analogue as a useful probe in the investigation of nucleic acids by fluorescence spectroscopy and microscopy.[17] We are currently evaluating the suitability of the fluorescent nucleoside analogue in probing the cellular uptake and trafficking of RNA ONs (e.g., siRNA) by fluorescence microscopy, and the results will be reported in due course.

4.4 Experimental Section

4.4.1 Materials

Except mentioned, materials obtained from commercial suppliers were used without any purification. 5-iodouridine, triethylene glycol monomethyl ether, 4-bromo-1,8-naphthalic anhydride, ethynyltrimethylsilane, tetrabutylammonium fluoride, tetrakis (triphenylphosphine) palladium(0), Copper(I) iodide, silver nitrate, *tert*-butyldimethylsilyl chloride and DEAE Sephadex A-25 resin were obtained from Sigma-Aldrich. $POCl_3$ was purchased from Across Organics, and was distilled prior to use. 2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite was purchased from Alfa Aesar. TBDMS-protected ribonucleoside phosphoramidite substrates and solid supports required for the solid-phase ON synthesis were purchased from Glen Research. Custom synthesized DNA ONs (ONs) were purchased from Integrated DNA Technologies, Inc. DNA ONs were purified by polyacrylamide gel electrophoresis (PAGE) under denaturing condition, followed by desalted on Sep-Pak Classic C18 cartridges (Waters Corporation). T7 RNA polymerase, ribonuclease inhibitor (RiboLock), RNase A, RNase T1 and NTPs were obtained from Fermentas Life Sciences. Calf intestinal alkaline phosphatase was procured from Invitrogen. Snake venom

phosphodiesterase I was purchased from Sigma-Aldrich. Radiolabeled α-^{32}P ATP (2000 Ci/mmol) was purchased from the Board of Radiation and Isotope Technology, Government of India. Chemicals for preparing buffer solutions were purchased from Sigma-Aldrich (BioUltra grade). Autoclaved water was used in all biochemical reactions and fluorescence analysis.

4.4.2 Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400. All MALDI-MS measurements were recorded on an Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer, MicroMass ESI-TOF and Water Synapt G2 High Defination mass spectrometer. Absorption spectra were recorded on a PerkinElmer, Lambda 45 UV-Vis spectrophotometer. UV-thermal melting studies of ONs were performed on a Cary 300Bio UV-Vis spectrophotometer. Steady State and time-resolved fluorescence experiments were carried out in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a TCSPC instrument (Horiba Jobin Yvon, Fluorolog-3). Reversed-phase flash chromatographic (C18 Redi*SepRf* column) purifications were carried out using Teledyne ISCO, Combi Flash *Rf*. HPLC analyses were performed using Agilent Technologies 1260 Infinity. Modified RNA ONs were synthesized on an Applied Biosystems oligonucleotide synthesizer (ABI-394).

4.4.3 Synthesis of ribonucleoside 6 and triphosphate substrate 7

4-Bromo-*N*-(triethyleneglycol monomethyl ether)-1,8-naphthalimide (3): Refluxed a mixture of 4-Bromo-1,8-napthalic anhydride **1** (1.4 g, 5.1 mmol, 1.0 equiv), 2-(2-(2-methoxyethoxy)ethoxy)ethanamine **2**[35] (1.08 g, 5.6 mmol, 1.1 equiv) and ethanol (14 ml) for 8 h. Evaporated the solvent completely, redissolved the product in ethyacetate and washed with water follwed by brine, an organic extract was dried over sodium sulphate and evaporated. The crude product obtained was purified by coloumn chromatograthy to give product **3** as pale yellow solid (2.92 g, 90%). TLC (Ethyl acetate) R_f = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.62 (dd, *J* = 7.2 Hz, *J* = 1.2 Hz, 1H), 8.54 (dd, *J* = 8.4 Hz, *J* = 1.2 Hz, 1H), 8.38 (d, *J* = 7.6 Hz, 1H), 8.02 (d, *J* = 8 Hz, 1H), 7.82 (dd, *J* = 8.4 Hz, *J* = 7.2 Hz, 1H), 4.42 (t, *J* = 6 Hz, 2H), 3.81 (t, *J* = 6.2 Hz, 2H), 3.71–3.69 (m, 2H), 3.62–3.60 (m, 2H), 3.58–3.56 (m, 2H), 3.44–3.42 (m, 2H), 3.30 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 163.7, 133.4, 132.2, 131.4, 131.2, 130.7, 130.4, 129.1, 128.2, 123.1, 122.3, 72.0, 70.7, 70.6,

70.3, 68.0, 59.1, 39.3. HRMS: *m/z* Calcd. for $C_{19}H_{20}BrNO_5Na$ $[M+Na]^+$ = 444.0423, found = 444.0420.

4-Ethnyl-*N*-(triethyleneglycol monomethyl ether)-1,8-naphthalimide (4): A mixture of 4-Bromo napthalimide derivative **3** (2.89 g, 6.8 mmol, 1.0 equiv), tetrakis(triphenylphosphine) palladium(0) (0.395 g, 0.34 mmol, 0.05 equiv), CuI (0.13 g, 0.68 mmol, 0.1 equiv) and *N*,*N*-diisopropylethylamine (4.2 mL, 24 mmol, 3.5 equiv) were dissolved in degassed anhydrous THF (57 ml). Trimethylsilylacetylene (3.9 mL, 27 mmol, 4 equiv) was added slowly to the above solution and reaction mixture was stirred at RT for 12 h in dark, solvent was evaporated completely and the residue was purified by silica gel column chromatography to afford the brown viscous oil (2.62 g, 87%). Into a solution of coupled product (2.5 g, 5.7 mmol, 1 equiv) and anhydrous methanol (50 ml) was added 1 M tetrabutylammonium fluoride solution in THF (22.8 mL, 22.8 mmol, 4.0 equiv). The reaction mixture was heated at 60 °C for 1 h, evaporated the solvent and dissolved the residue in ethyl acetate, washed with water and brine. Dry the organic extract over sodium sulphate and evaporated, the crude product was purified by column chromatography to give a product **4** as lemon yellow soft solid (1.37 g, 66%). TLC (Ethyl acetate) R_f = 0.63; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.63 (d, *J* = 8.4 Hz, 1H), 8.60 (d, *J* = 7.6 Hz, 1H), 8.50 (d, *J* = 7.6 Hz, 1H), 7.91 (d, *J* = 7.6 Hz, 1H), 7.80 ($_{app}$t, *J* = 7.8 Hz, 1H), 4.42 (t, *J* = 6 Hz, 2H), 3.82 (t, *J* = 6.0 Hz, 2H), 3.74 (s, 1H), 3.71–3.69 (m, 2H), 3.62–3.60 (m, 2H), 3.58–3.56 (m, 2H), 3.44–3.42 (m, 2H), 3.30 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 164.0, 163.8, 132.4, 132.0, 131.8, 131.7, 130.3, 128.0, 127.8, 126.4, 123.0, 122.8, 86.7, 80.4, 72.0, 70.7, 70.6, 70.3, 68.0, 59.1, 39.3. HRMS: *m/z* Calcd. for $C_{21}H_{21}NO_5Na$ $[M+Na]^+$ = 390.1317, found = 390.1321.

4-Ethynyl-*N*-(triethylene glycol monomethyl ether)-1,8-naphthalimide-conjugated uridine (6): Degassed a mixture of iodouridine **5** (0.1 g, 0.27 mmol, 1.0 equiv), 4-Ethynyl-*N*-(2-triethyleneglycol monomethyl ether)-1,8-naphthalimide **4** (0.140 g, 0.38 mmol, 1.25 equiv), CuI (0.005 g, 0.027 mmol, 0.1 equiv) and tetrakis(triphenylphosphine) palladium(0) (0.016 g, 0.014 mmol, 0.05 equiv) for 0.5 h and dissolved in degassed anhydrous DMF (4 ml), Et_3N (0.075 ml, 0.54 mmol, 2.0 equiv) was slowly added to the above solution. The reaction mixture was stirred in dark at RT for 4 h. Solvent was evaporated and the resulting solid was purified by reverse-phase flash column chromatography (C18 Redi*SepRf*, 0–60% acetonitrile in water for 40 min) to give the fluorescent ribinucleoside **6** as orange solid

(0.140 g, 85%). TLC (CH_2Cl_2:MeOH = 8:2); R_f = 0.72; ^{1}H NMR (400 MHz, d_6-DMSO): δ (ppm) 11.85 (br, 1H), 8.79 (s, 1H), 8.73 (d, *J* = 8.4 Hz, 1H), 8.54 (d, *J* = 6.8 Hz, 1H), 8.44 (d, *J* = 7.6 Hz, 1H), 7.99–7.92 (m, 2H), 5.78 (d, *J* = 3.2 Hz, 1H), 5.55 (d, *J* = 4 Hz, 1H), 5.45 (br, 1H), 5.13 (d, *J* = 4.8 Hz, 1H), 4.23 (t, *J* = 6 Hz, 2H), 4.13–4.07 (m, 2H), 3.92–3.79 (m, 2H), 3.67–3.64 (m, 3H), 3.55–3.53 (m, 2H), 3.47–3.40 (m, 4H), 3.30–3.27 (m, 2H), 3.14 (s, 3H); ^{13}C NMR (100 MHz, d_6-DMSO): δ (ppm) 163.2, 162.9, 161.5, 149.6, 145.4, 132.0, 131.4, 130.7, 130.2, 130.0, 128.3, 127.4, 126.5, 122.6, 121.5, 97.4, 92.9, 89.4, 89.1, 84.6, 74.2, 71.2, 69.7, 69.7, 69.6, 68.8, 66.9, 59.9, 58.0. HRMS: *m/z* Calcd. for $C_{30}H_{31}N_3O_{11}Na$ $[M+Na]^+$ = 632.1856, found = 632.1853. ε_{260} = 10,413 $M^{-1}cm^{-1}$, ε_{390} = 23,880 $M^{-1}cm^{-1}$.

4-Ethynyl-*N*-(triethyleneglycol monomethyl ether)-1,8-naphthalimide-conjugated uridine triphosphate (7): To an ice cold solution of ribonucleoside **6** (0.091 g, 0.15 mmol, 1.0 equiv) in trimethyl phosphate (1.2 ml) was added freshly distilled $POCl_3$ (35 μL, 0.37 mmol, 2.5 equiv). The solution was stirred for 24 h at ~4 °C. A solution of bis-tributylammonium pyrophosphate (0.5 M in DMF, 1.5 ml, 5.0 equiv) and tributylamine (0.35 ml, 1.5 mmol, 10.0 equiv) was rapidly added under ice-cold condition. The reaction was quenched after 30 min with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, 15 ml), and was washed with ethyl acetate (20 ml). The aqueous layer was evaporated and the residue was purified first on DEAE sephadex-A25 anion exchange column (10 mM-1M TEAB buffer, pH 7.5) followed by reverse-phase flash column chromatography (C18 Redi*SepRf*, 0–40% acetonitrile in 50 mM triethylammonium acetate buffer (TEAA), pH 7.2, 40 min). Appropriate fractions were lyophilized to give the desired triphosphate product **7** as a tetratriethylammonium salt (59 mg, 31%). 1H NMR (400 MHz, D_2O): δ (ppm) 7.97–7.87 (m, 4H), 7.42–7.38 (m, 2H), 5.72 (br, 1H), 4.30 (br, 2H), 4.16 (br, 3H), 3.99 (br, 2H), 3.58 (t, *J* = 4.4 Hz, 2H), 3.51 (br, 2H), 3.43 (br, 2H), 3.35–3.33 (m, 2H), 3.19 (t, *J* = 3.6 Hz, 2H), 3.03 (s, 3H); ^{13}C NMR (100 MHz, D_2O): δ (ppm) 164.6, 164.4, 162.8, 150.2, 144.7, 133.2, 131.8, 130.7, 130.3, 130.0, 128.0, 127.3, 126.5, 120.5, 119.8, 99.3, 91.2, 90.9, 89.5, 83.5, 73.6, 70.9, 69.6, 69.6, 69.5, 69.4, 67.3, 65.2, 57.9; ^{31}P NMR (162 MHz, D_2O): δ (ppm) -9.09 (br, P_γ), -11.32 (br, P_α), -22..51 (br, P_β) HRMS (negative mode): *m/z* Calcd. for $C_{30}H_{34}N_3O_{20}P_3$ [M] = 849.0949, found = 848.0955 $[M-H]^+$.

Naphthalimide-modified 5'-DMT-protected uridine (17): A mixture of 5'-DMT protected-iodouridine **16** (0.615 g, 0.92 mmol, 1.0 equiv),[36] 4-Ethynyl-*N*-(2-triethyleneglycol

monomethyl ether)-1,8-naphthalimide **4** (0.470 g, 1.28 mmol, 1.4 equiv), CuI (0.021 g, 0.113 mmol, 0.12 equiv) and tetrakis(triphenylphosphine) palladium(0) (0.065 g, 0.056 mmol, 0.06 equiv) was dissolved in degassed anhydrous DMF (12 ml). *N*,*N*-diisopropylethylamine (0.78 ml, 4.5 mmol, 4.9 equiv) was slowly added to the above solution. The reaction mixture was stirred in dark at RT for 5 h. Solvent was evaporated and the resulting solid was purified by silica gel column chromatography to afford the compound **17** as faint yellow foam (0.705 g, 85%). TLC (CH_2Cl_2:MeOH = 95:5 with few drops of Et_3N); R_f = 0.56; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.46 (S, 1H), 8.34 (d, *J* = 7.2 Hz, 1H), 8.26 (d, *J* = 7.2 Hz, 1H), 8.09 (d, *J* = 7.2 Hz, 1H), 7.47 (d, *J* = 7.2 Hz, 2H), 7.36 (d, *J* = 8 Hz, 4H), 7.23–7.18 (m, 3H), 7.02 (t, *J* = 7 Hz, 1H), 6.89 (d, *J* = 7.2 Hz, 1H), 6.76 ($_{app}$t, *J* = 5.6 Hz, 4H), 6.07 (br, 1H), 4.57 (br, 1H), 4.44 (br, 1H), 4.36 (br, 3H), 4.27 (br, 2H), 3.77 (t, *J* = 5.6 Hz, 2H), 3.68 (t, *J* = 5.6 Hz, 2H), 3.61 (t, *J* = 4 Hz, 2H), 3.57 (S, 9H), 3.45 (t, *J* = 6 Hz, 2H), 3.31–3.26 (m, 4H); ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 164.0, 163.8, 162.4, 158.7, 150.9, 144.7, 143.2, 135.9, 135.6, 132.7, 131.4, 131.0, 130.4, 130.0, 129.9, 128.3, 128.0, 127.6, 127.4, 127.1, 126.9, 122.3, 121.6, 113.5, 100.3, 91.3, 90.4, 90.1, 87.2, 84.9, 76.1, 72.0, 70.8, 70.7, 70.5, 70.2, 68.0, 63.0, 59.1, 55.2, 39.2; HRMS: *m/z* Calcd. for $C_{51}H_{50}N_3O_{13}$ $[M+H]^+$ = 912.3344, found = 912.3349.

Naphthalimide-modified 2'-*O*-TBDMS protected uridine (18): A mixture of compound **17** (0.840 g, 0.92 mmol, 1.0 equiv) and silver nitrate (0.350 g, 2.07 mmol, 2.25 equiv) was dissolved in anhydrous pyridine (2.5 ml). To the above solution was added anhydrous THF (12.6 ml) and stirred for 10 min. After this period, *tert*-butyldimethylsilyl chloride (TBDMS-Cl, 0.312 g, 2.07 mmol, 2.25 equiv) was added to the reaction mixture. The reaction mixture was stirred for additional 0.5 h and filtered through celite pad. Celite pad was washed with ethyl acetate (3 x 10 ml), and the filtrate was washed with 5 % sodium bicarbonate (25 ml) followed by brine (25 ml). The organic extract was dried over sodium sulphate and evaporated. The resulting crude residue was purified by silica gel column chromatography to afford the product **18** as pale yellow foam (0.690 g, 73%). TLC (EtOAc:petroleum ether = 60:40 with few drops of Et_3N); R_f = 0.50; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.57 (S, 1H), 8.50 (d, *J* = 7.2 Hz, 1H), 8.42 (d, *J* = 8.4 Hz, 1H), 8.18 (d, *J* = 8.0 Hz, 1H), 7.49 (d, *J* = 7.6 Hz, 2H), 7.44 (d, *J* = 8 Hz, 1H), 7.38 (dd, *J* = 8.6 Hz, *J* = 5.4 Hz, 5H), 7.23 (t, *J* = 7.6 Hz, 2H), 7.05 (t, *J* = 7.4 Hz, 1H), 6.75 (dd, *J* = 8.8 Hz, *J* = 2.8 Hz, 4H), 6.13 (d, *J* = 5.2 Hz, 1H), 4.60 (t, *J* = 5.0 Hz, 1H), 4.40 (t, *J* = 5.8 Hz, 1H), 4.25 (t, *J* = 4.0 Hz, 1H), 4.22 (br, 1H), 3.79 (t, *J* = 6.0 Hz, 2H), 3.70 (t, *J* = 4.8 Hz, 2H), 3.62 (t, *J* = 4.8 Hz, 2H), 3.59 (br, 4H), 3.57 (S,

6H), 3.45 (t, *J* = 4.8 Hz, 2H), 3.31 (S, 3H), 3.26 (br, 1H), 0.92 (S, 9H), 0.17 (S, 3H), 0.14 (S, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 164.2, 164.0, 162.2, 158.8, 150.1, 144.7, 143.0, 135.6, 135.3, 133.0, 131.6, 131.3, 130.5, 128.3, 128.1, 127.9, 127.8, 127.6, 127.3, 127.1, 122.5, 121.8, 113.5, 100.4, 91.3, 90.3, 88.6, 87.5, 84.4, 72.0, 71.5, 70.7, 70.6, 70.3, 68.0, 63.2, 59.1, 55.2, 53.0, 39.2, 29.8, 25.8, -4.6, -5.0; HRMS: *m/z* Calcd. for $C_{57}H_{64}N_3O_{13}Si$ $[M+H]^+$ =1026.4208, found = 1026.4224.

Naphthalimide-modified uridine phosphoramidite (19): To a solution of compound **18** (0.250 g, 0.24 mmol, 1.0 equiv) in anhydrous dichloromethane (2.5 ml) was added *N*,*N*-diisopropylethylamine (0.11 ml, 0.6 mmol, 2.5 equiv) and stirred for 10 min. To this solution was slowly added 2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.065 ml, 0.29 mmol, 1.2 equiv) and the reaction mixture was stirred for 12 h. The solvent was evaporated to dryness and was redissolved in ethyl acetate (20 ml), washed with 5% sodium bicarbonate solution (10 ml) followed by brine (10 ml). The organic extract was dried over sodium sulphate, evaporated and the residue was purified by column chromatography to afford the product **19** as pale yellow solid (0.210 g, 70 %). TLC (EtOAc:petroleum ether = 60:40 with few drops of Et_3N); R_f = 0.60; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.69 (S, 1H), 8.48 (d, *J* = 6.0 Hz, 1H), 8.42 (d, *J* = 8.4 Hz, 1H), 8.17 (d, *J* = 8.8 Hz, 1H), 7.52 (d, *J* = 7.6 Hz, 2H), 7.40–7.37 (m, 5H), 7.23 (d, *J* = 7.6 Hz, 2H), 7.00 (d, *J* = 7.2 Hz, 1H), 6.75 (dd, *J* = 7.0 Hz, *J* = 1.8 Hz, 5H), 6.20 (d, *J* = 6.4 Hz, 1H), 4.65 (t, *J* = 4.2 Hz, 1H), 4.40 (d, *J* = 6.0 Hz, 2H), 4.23 (d, *J* = 4.8 Hz, 1H), 3.78 (d, *J* = 6.0 Hz, 1H), 3.69 (t, *J* = 4.6 Hz, 2H), 3.65 (t, *J* = 3.2 Hz, 2H), 3.61 (t, *J* = 2.6 Hz, 2H), 3.59 ($_{app}$t, 2H, *J* = 1.8 Hz), 3.58 (s, 6H),3.52 (t, *J* = 2.8 Hz, 2H), 3.45 (t, *J* = 2.6 Hz, 2H) 3.3 (S, 3H), 2.70 (br, 2H), 2.69 (br, 2H), 2.67 (br, 2H), 1.14 (d, *J* = 7.2 Hz, 12H) 0.92 (S, 9H), 0.14 (S, 3H), 0.12 (S, 3H); ^{31}P NMR (162 MHz, $CDCl_3$): δ (ppm) 150.61, 149.50; HRMS: *m/z* Calcd. for $C_{66}H_{80}N_5O_{14}$ PSiNa $[M+Na]^+$ =1248.5106, found = 1248.5121.

4.4.4 Photophysical characterization of ribonucleoside analogue 6

Steady-state fluorescence of ribonucleoside analog in various solvents

Ribonucleoside analog **6** (5 μM) in water, methanol, acetonitrile, dioxane, was excited at respective lowest energy absorption maximum with an excitation and emission slit widths of 1 nm and 5 nm respectively. Fluorescence experiments were performed in triplicate in a micro fluorescence cell (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3.

Quantum yield determination in various solvents

Quantum yield of ribonucleoside **6** in different solvents relative to cumarin 153 as standard was determined using the following equation.[37]

$$\Phi_{F(x)} = (A_s/A_x)\,(F_x/F_s)\,(n_x/n_s)^2\,\Phi_{F(s)}$$

Where s is the standard, x is the ribonucleoside, A is the absorbance at excitation wavelength, F is the area under the emission curve, *n* is the refractive index of the solvent, and Φ_F is the quantum yield. Quantum yield of cumarin 153 in acetonitrile is 0.56.[38]

Time-resolved fluorescence measurements

Excited state lifetimes of ribonucleoside **6** in various solvents were determined using TCSPC fluorescence spectrophotometer (Horiba Jobin Yvon). Concentration of ribonucleoside **6** in all solvents was 5 μM, and was excited using 371 nm LED source (IBH, UK, NanoLED-371L) and fluorescence signal at respective emission maximum was collected. Lifetime measurements were performed in duplicate, and decay profiles were analyzed using IBH DAS6 analysis software. Fluorescence intensity decay kinetics in all solvents was found to be monoexponential with χ^2 (goodness of fit) values very close to unity.

4.4.5 Enzymatic incorporation of modified triphosphate 7 into RNA ONs

Transcription reactions in presence of radiolabeled α-^{32}P ATP: Promoter-template duplexes were constructed by heating a equimolar concentration (final 5 μM) of DNA templates (**T1–T5**) and 18-mer T7 RNA polymerase consensus promoter DNA sequence in TE buffer (10mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.8) at 90 °C for 3 min and cooling the solution slowly to room temperature. The duplexes were then placed in ice bath for 20 min and stored at -40 °C. Transcription reactions were performed in 40 mM Tris-HCl buffer (pH 7.9) containing 250 nM annealed templates, 10 mM $MgCl_2$, 10 mM NaCl, 10 mM dithiothreitol (DTT), 2 mM spermidine, 1 U/μL RNase inhibitor (RiboLock), 1 mM GTP, 1 mM CTP, 1 mM UTP and or 1 mM modified UTP **7**, 20 μM ATP, 5 μCi α^{32}P ATP and 3 U/μL T7 RNA polymerase in a total volume of 20 μL. reactions were quenched After 4 h at 37 °C, by adding 20 μL of the loading buffer (7 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8), heated to 75 °C for 3 min and cooled on an ice bath. The samples (5 μL) were loaded onto a sequencing 18% denaturing polyacrylamide gel. The products on the gel were exposed to X-ray film (10 min) and the exposed film was developed, fixed and dried. The bands were then quantified using the software GeneTools (from

Syngene), percentage incorporation of modified ribonucleoside triphosphate **7** is reported relative to transcription efficiency in the presence of natural NTPs. All reactions were performed in duplicate and the standard deviations were found to be ± 3%.

Large-scale transcription reactions: Large-scale transcription reactions using template **T1** and **T3–T7** were performed in a 250 μL reaction volume under similar conditions mentioned above to isolate ONs for further characterization and photophysical studies. Mixed a solutions of 2 mM GTP, 2 mM CTP, 2 mM ATP, 2 mM UTP or 2 mM modified UTP **7**, 20 mM $MgCl_2$, 0.4 U/μL RNase inhibitor (RiboLock), 300 nM annealed template and 800 units T7 RNA polymerase and incubation for 12 h at 37 °C. The reaction volume was reduced approximately to 1/3 by Speed Vac, and 30 μL of loading buffer was added. The samples were loaded onto a preparative 20% denaturing polyacrylamide gel and run at constant power of (24 W) for 5 h. The gel was UV shadowed, appropriate band was removed, extracted with 0.3 M sodium acetate and desalted using Sep-Pak classic C18 cartridge. Typical yields of the transcripts are 16–19 nmol.

4.4.6 Enzymatic digestion of transcript 9

4 nmol of the fluorescent modified ON **9** was digested with snake venom phosphodiesterase I (0.015 U), calf intestine alkaline phosphatase (1.5 U/μL), and RNase A (0.5 μg) in a total volume of 100 μL in 50 mM Tris-HCl buffer (pH 8.5, 40 mM $MgCl_2$, 0.1 mM EDTA) for 24 h at 37 °C. RNase T1 (0.6 U/μL) was then added, and the sample was incubated for another 12 h at 37 °C. The ribonucleoside mixture obtained was analyzed by reversed-phase analytical HPLC using Phenomenex, Luna C18 column (250 x 4.6 mm, 5 micron) at 260 nm and 390 nm. Mobile phase A: 50 mM triethylammonium acetate buffer (pH 7.1), mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min. The fractions corresponding to individual ribonucleoside fractions were further analyzed by mass spectroscopy.

4.4.7 Solid-phase synthesis of RNA ON 20

Napthalimide-modified RNA ON **20** was synthesized on a 1 μmole scale CPG solid support (1000 Å) by following standard solid-phase RNA ON synthesis protocol.[34] While incorporation of regular 2'-*O*-TBDMS-protected phosphoramidites were performed with a coupling time of 10 min, incorporation of fluorescent 2'-*O*-TBDMS-protected

phosphoramidite substrate **19** was performed with a coupling time of 30 min. The coupling efficiency in the presence of **19**, based on trityl monitor assay, was found to be 20%. The trityl protecting group was deprotected on the synthesizer. The solid support was treated with a 1:1 solution of 10 M methylamine in ethanol and water (1.5 mL) for 12 h. The mixture was centrifuged and the supernatant was evaporated to dryness on a Speed Vac. The residue was then dissolved in DMSO (100 μL) and treated with TEA.3HF (150 μL). The sample was heated at 65°C for 2.5 h and was brought to RM. The completely deprotected ON solution was lyophilized and was purified by PAGE using a 20% gel under denaturing conditions. The band corresponding to the full-length product was identified by UV shadowing. The ON was extracted with ammonium acetate buffer (0.5 M, 3 ml) and desalted using Sep-Pak classic C18 cartridges.

4.4.8 MALDI-TOF MS of ONs transcripts

Molecular weight of RNA transcripts were determined using Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer. 2 μL of the transcript (150 μM) was combined with 1 μL of ammonium citrate buffer (100 mM, pH 9), 2 μL of a DNA standard (200 μM, 18-mer) and 4 μL of saturated 3-hydroxypicolinic acid solution. The sample was desalted with an ionexchange resin (Dowex 50W-X8, 100-200 mesh, ammonium form), spotted on the MALDI plate, and was air dried. The resulting spectrum was calibrated relative to an internal 18-mer DNA ON standard.

4.4.9 Photophysical characterization of naphthalimide-modified ONs

ONs **9**, **14**, **15** and **20** (10 μM) were annealed to respective complementary custom DNA ONs by heating a 1:1.1 mixture of the ONs in 20 mM cacodylate buffer (pH 7.0, 500 mM NaCl, 0.5 mM EDTA) at 90 °C for 3 min. Samples were then cooled slowly to RT, placed in crushed ice for 2 h. Samples were diluted to give a final concentration of 1 μM (with respect to **9**, **14**, **15** and **20**) in cacodylate buffer. Fluorescently modified duplexes were excited at 407 nm with excitation and emission slit widths of 6 and 8 nm, respectively. Fluorescence experiments were performed in triplicate in a micro fluorescence cuvette (Hellma, path length 1.0 cm) on a Horiba Jobin Yvon, Fluorolog-3 at 20 °C.

References

[1] a) U. Asseline, *Curr. Org. Chem*. **2006**, *10*, 491–518; b) K. B. Hall, *Methods Enzymol.* **2009**, *469*, 269–285; c) X. Shi, D. Herschlag, *Methods Enzymol.* **2009**, *469*, 288–302; d) L. M. Wilhelmsson, *Quart. Rev. Biophy*. **2010**, *43*, 159–183; e) S. G. Srivatsan, A. A. Sawant, *Pure Appl. Chem*. **2011**, *83*, 213–232.

[2] R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev*. **2010**, *110*, 2579–2619.

[3] a) A. Okamoto, K. Tanaka, T. Fukuta, I. Saito, *J. Am. Chem. Soc*. **2003**, *125*, 9296–9297; b) A. Okamoto, Y. Saito, I. Saito, *J. Photochem. Photobiol. C*, **2005**, *6*, 108–122; c) E. Mayer-Enthart, H.-A. Wagenknecht, *Angew. Chem.* **2006**, *118*, 3451–3453; *Angew. Chem. Int. Ed*. **2006**, *45*, 3372–3375; d) K. Tainaka, K. Tanaka, S. Ikeda, K.-I Nishiza, T. Unzai, Y. Fujiwara, I. Saito, A. Okamoto, *J. Am. Chem. Soc.* **2007**, *129*, 4776–4784; e) F. Takei, H. Suda, M. Hagihara, J. Zhang, A. Kobori, K. Nakatani, *Chem. Eur. J.* **2007**, *13*, 4452–4457; f) Q. Xiao, R. T. Ranasinghe, A. M. P. Tang, T. Brown, *Tetrahedron*, **2007**, *63*, 3483–3490; g) S. G. Srivatsan, H. Weizman, Y. Tor, *Org. Biomol. Chem.* **2008**, *6*, 1334–1338; h) H. Gardarsson, A. S. Kale, S. T. Sigurdsson, *ChemBioChem*, **2011**, *12*, 567–575; i) X. Ming, F. Seela, *Chem. Eur. J*. **2012**, *18*, 9590–9600.

[4] a) T. J. Matray, E. T. Kool, *Nature* **1999**, *399*, 704–708; b) L. Valis, N. Amann, H.-A. Wagenknecht, *Org. Biomol. Chem*. **2005**, *3*, 36–38; c) E. Shipova, K. S. Gates, *Bioorg. Med. Chem. Lett*. **2005**, *15*, 2111– 2113; d) N. J. Greco, Y. Tor, *J. Am. Chem. Soc*. **2005**, *127*, 10784–10785; e) A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J*. **2011**, *17*, 12820–12827.

[5] a) N. J. Greco, R. W. Sinkeldam, Y. Tor, *Org. Lett.* **2009**, *11*, 1115–1118; b) W. Hirose, K. Sato, A. Matsuda, *Angew. Chem.* **2010**, *122*, 8570–8572; *Angew. Chem. Int. Ed*. **2010**, *49*, 8392–8394; c) Q. Zhang, Y. Wang, X. Meng, R. Dhar, H. Huang, *Anal. Chem*. **2013**, *85*, 201–207.

[6] a) S. O. Kelly, J. K. Barton, *Science* **1999**, *283*, 375–381; b) N. A. Grigorenko, C. J. Leumann, *Chem. Comm.* **2008**, 5417–5419; c) A. Zahn, C. J. Leumann, *Chem. Eur. J*. **2008**, *14*, 1087–1094.

[7] a) A. Nadler, J. Strohmeier, U. Diederichsen, *Angew. Chem*. **2011**, *123*, 5504 –5508; *Angew. Chem. Int. Ed*. **2011**, *50*, 5392–5396; b) S. Müllar, J. Strohmeier, U.

Diederichsen, *Org. Lett*. **2012**, *14*, 1382–1385; c) II J. Lee, B. H. Kim, *Chem. Commun.* **2012**, *48*, 2074–2076; d) K. T. Kim, B. H. Kim, *Chem. Commun.* **2013**, *49*, 1717–1719.

[8] a) S. G. Srivatsan, N. J. Greco, Y. Tor, *Angew. Chem*. **2008**, *120*, 6763–6767; *Angew. Chem. Int. Ed*. **2008**, *47*, 6661–6665; b) A. S. Wahba, A. Esmaeili, M. J. Damha, R. H. E. Hudson, *Nucleic Acids Res*. **2010**, *38*, 1048–1056; c) R. W. Sinkeldam, L. S. McCoy, D. Shin, Y. Tor, *Angew. Chem. Int. Ed.* **2013**, *52*, 14026–14030.

[9] a) R. Kawai, M. Kimoto, S. Ikeda, T. Mitsui, M. Endo, S. Yokoyama, I. Hirao, *J. Am. Chem. Soc.* **2005**, *127*, 17286–17295; b) R. Rieder, K. Lang, D. Graber, R. Micura, *ChemBioChem* **2007**, *8*, 896–902; c) S. G. Srivatsan, Y. Tor, *Tetrahedron* **2007**, *63*, 3601–3607; d) J. Parsons, M. P. Castaldi, S. Dutta, S. M. Dibrov, D. L. Wyles, T. Hermann. *Nat. Chem. Biol.* **2009**, *5*, 823–825; e) Y. Xie, A. V. Dix, Y. Tor, *J. Am. Chem. Soc.* **2009**, *131*, 17605–17614; f) J. Riedl, R. Pohl, N. P. Ernsting, P. Orság, M. Fojta, M. Hocek, *Chem. Sci*. **2012**, *3*, 2797–2806; g) J. Riedl, P. Ménová, R. Pohl, P. Orság, M. Fojta, M. Hocek, *J. Org. Chem.* **2012**, *77*, 8287−8293.

[10] a) P. Cekan, S. T. Sigurdsson, *Chem. Commun*. **2008**, 3393–3395; b) K. Borjesson, S. Preus, A. H. El-Sagheer, T. Brown, B. Albinsson, L. M. Wilhelmsson, *J. Am. Chem. Soc.* **2009**, *131*, 4288–4293; c) A. A. Tanpure, S. G. Srivatsan, *ChemBioChem* **2012**, *13*, 2392–2399.

[11] C. A. M. Seidel, A. Schulz, M. H. M. Sauer, *J. Phys. Chem.* **1996**, *100*, 5541–5553.

[12] J. Hou, X. Liu, J. Liu, *Biotechnol. Biotechnol. Eq.* **2012**, *26*, 3148–3154.

[13] a) B. R. Vummidi, J. Alzeer, N. W. Luedtke, *ChemBioChem* **2013**, *14*, 540–558. b) Y. Xu, Y. Suzuki, K. Ito, M. Komiyama, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14579–14584.

[14] Y. N. Teo, E. T. Kool, *Chem. Rev*. **2012**, *112*, 4221–4245.

[15] a) S. M. Park, S.-J. Nam, H. S. Jeong, W. J. Kim, B. H. Kim, *Chem. Asian J*. **2011**, *6*, 487–492; b) K. Phelps, A. Morris, P. A. Beal, *ACS Chem. Biol.* **2012**, *7*, 100–109; c) A. A. Tanpure, M. G. Pawar, S. G. Srivatsan, *Isr. J. Chem*. **2013**, *53*, 366–378.

[16] a) S. S. Tan, S. J. Kim, E. T. Kool, *J. Am. Chem. Soc*. **2011**, *133*, 2664–2671; b) J. Guo, S. Wang, N. Dai, Y. N. Teo, E. T. Kool, *Proc. Natl. Acad. Sci. USA*, **2011**, *108*, 3493–3498.

[17] A. S. Wahba, F. Azizi, G. F. Deleavey, C. Brown, F. Robert, M. Carrier, A. Kalota, A. M. Gewirtz, J. Pelletier, R. H. E. Hudson, M. J. Damha, *ACS Chem. Biol*. **2011**, *6*, 912–919.

[18] a) A. Dierckx, F.-A. Miannay, N. B. Gaied, S. Preus, M. Björck, T. Brown, L. M. Wilhelmsson, *Chem. Eur. J.* **2012**, *18*, 5987–5997; b) M. Weinberger, F. Berndt, R. Mahrwald, N. P. Ernsting, H.-A. Wagenknecht, *J. Org. Chem*, **2013,** *78*, 2589–2599.

[19] a) W. W. Stewart, *Nature* **1981**, *292*, 17–21; b) A. P. De Silva, H. Q. N. Gunaratne, J.-L. Habib-Jiwan, C. P. McCoy, T. E. Rice, J.-P. Soumillion, *Angew. Chem.* **1995**, *107*, 1889–1891; *Angew. Chem. Int. Ed.* **1995**, *34*, 1728–1731.

[20] S. Banerjee, E. B. Veale, C. M. Phelan, S. A. Murphy, G. M. Tocci, L. J. Gillespie, D. O. Frimannsson, J. M. Kelly, T. Gunnlaugsson, *Chem. Soc. Rev.* **2013**, *42*, 1601–1618.

[21] R. M. Duke, E. B. Veale, F. M. Pfeffer, P. E. Kruger, T. Gunnlaugsson, *Chem. Soc. Rev.* **2010**, *39*, 3936–3953.

[22] a) S. H. Weisbrod, A. Marx, *Chem. Commun*. **2008**, 5675–5685; b) F. Wachowius, C. Höbartner, *ChemBioChem* **2010**, *11*, 469–480.

[23] a) J. Riedl, R. Pohl, L. Rulíšek, M. Hocek, *J. Org. Chem*. **2012**, *77*, 1026–1044; b) B. Holzberger, J. Strohmeier, V. Siegmund, U. Diederichsen, A. Marx, *Bioorg. Med. Chem. Lett*. **2012**, *22*, 3136–3139.

[24] a) K. Lang, R. Rieder, R. Micura, *Nucl. Acids Res*. **2007**, *35*, 5370–5378; b) S. G. Srivatsan, Y. Tor, *Chem. Asian J.* **2009**, *4*, 419–427; c) G. Stengel, M. Urban, B. W. Purse , R. D. Kuchta, *Anal. Chem*. **2010**, *82*, 1082–1089; d) M. Kimoto, T. Mitsui, S. Yokoyama, I. Hirao, *J. Am. Chem. Soc*. **2010**, *132*, 4988–4989; e) Y. Hikida, M. Kimoto, S. Yokoyama, I. Hirao, *Nat. Protoc*. **2010**, *5*, 1312–1323; f) M. G. Pawar, S. G. Srivatsan, *Org. Lett.* **2011**, *13*, 1114–1117; g) M. G. Pawar, A. Nuthanakanti, S. G. Srivatsan, *Bioconjugate Chem.* **2013**, *24*, 1367−1377.

[25] M. Kimoto, T. Mitsui, Y. Harada, A. Sato, S. Yokoyama, I. Hirao, *Nucleic Acids Res*. **2007**, *35*, 5360–5369.

[26] J. Ludwig, *Acta. Biochim. et Biophys. Acad. Sci. Hung*. **1981**, *16*, 131–133.

[27] S. G. Srivatsan, Y. Tor, *J. Am. Chem. Soc.* **2007**, *129*, 2044–2053.

[28] Trace amounts of N+1, N+2 transcripts are also formed due to random noncode incorporations of nucleotides. J. F. Milligan, O. C. Uhlenbeck, *Methods Enzymol*. **1989**, *180*, 51–62.

[29] J. P. Peters, S. P. Yelgaonkar, S. G. Srivatsan, Y. Tor, L. J. Maher III, *Nucl. Acid. Res*. **2013**, *41*, 10593–10604.

[30] a) D. C. Ward, E. Reich, L. Stryer, *J. Biol. Chem*. **1969**, *244*, 1228–1237; b) E. L. Rachofsky, R. Osman, J. B. A. Ross, *Biochemistry*, **2001**, *40*, 946–956.

[31] S. Doose, H. Neuweiler, M. Sauer, *ChemPhysChem* **2009**, *10*, 1389–1398.

[32] R. W. Sinkeldam, A. J. Wheat, H. Boyaci, Y. Tor, *ChemPhysChem*, **2011**, *12*, 567–570.

[33] a) M. Mandal, R. R. Breaker, *Nat. Rev. Mol. Cell Biol*. **2004**, *5*, 451–463; b) Z. Shajani, P. Deka, G. Varani, *Trends Biochem Sci*. **2006**, *31*, 421–424; c) K. B. Hall, *Curr. Opin. Chem. Biol*. **2008**, *12*, 612–618; d) H. M. Al-Hashimi, N. G. Walter, *Curr. Opin. Struct. Biol.* **2008**, *18*, 321–329; e) D. M. J. Lilley, *Philosoph. Transac. Roy. Soc B*. **2011**, *366*, 2910–2917.

[34] B. S. Sproat, *Methods Mol. Biol*. **2005**, *288*, 17–32.

[35] J. Clerc, B. Schellenberg, M. Groll, A. S. Bachmann, R. Huber, R. Dudler, M. Kaiser, *Eur. J. Org. Chem.* **2010**, *21*, 3991–4003.

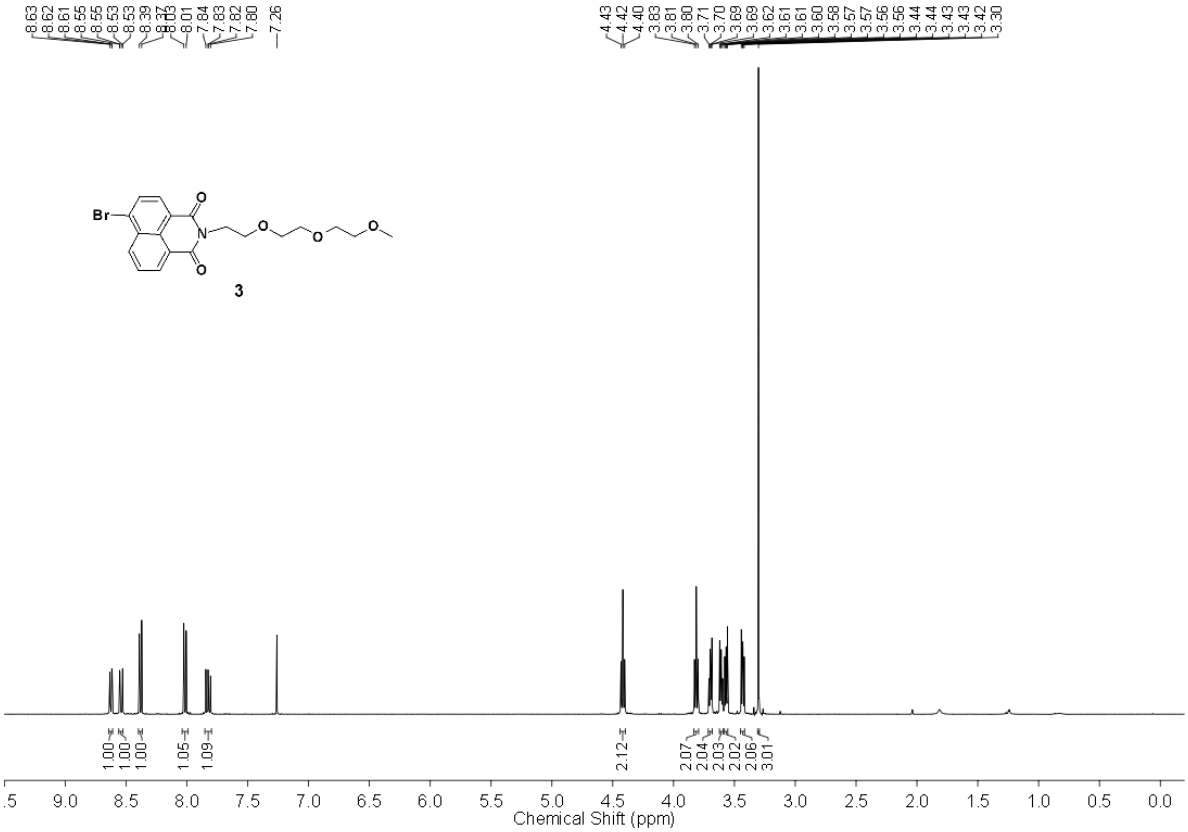
[36] K. Shah, H. Wu, T. M. Rana, *Bioconjugate Chem.* **1994**, *5*, 508–512.

[37] D. Lavabre, S. *J.* Fery-Forgues, *Chem. Educ.* **1999**, *76*, 1260–1264.

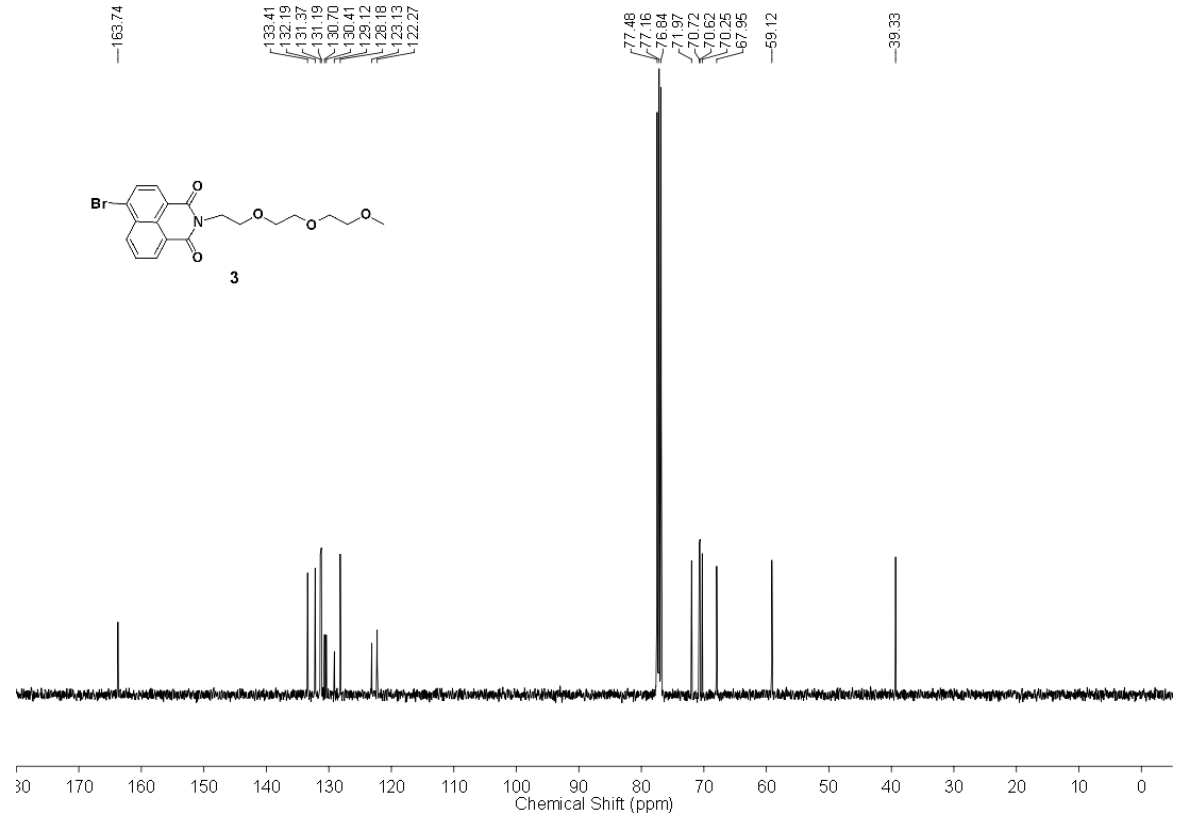
[38] G. Jones II, W. R. Jackson, C-Y. Choi, *J. Phys. Chem.* **1985**, *89*, 294–300.

4.6 Appendix-IV: Characterization data of synthesized compounds

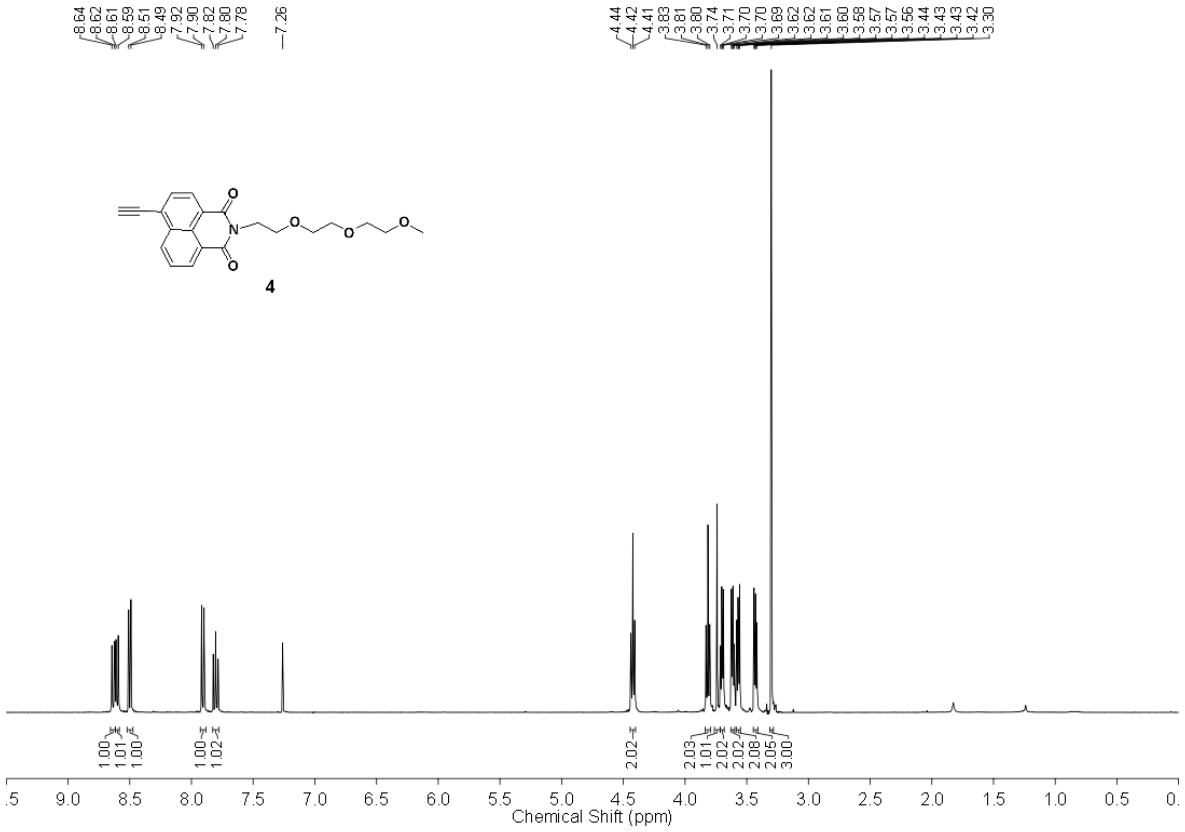
^{1}H-NMR of 4-Bromo napthalimide derivative **3** in $CDCl_3$



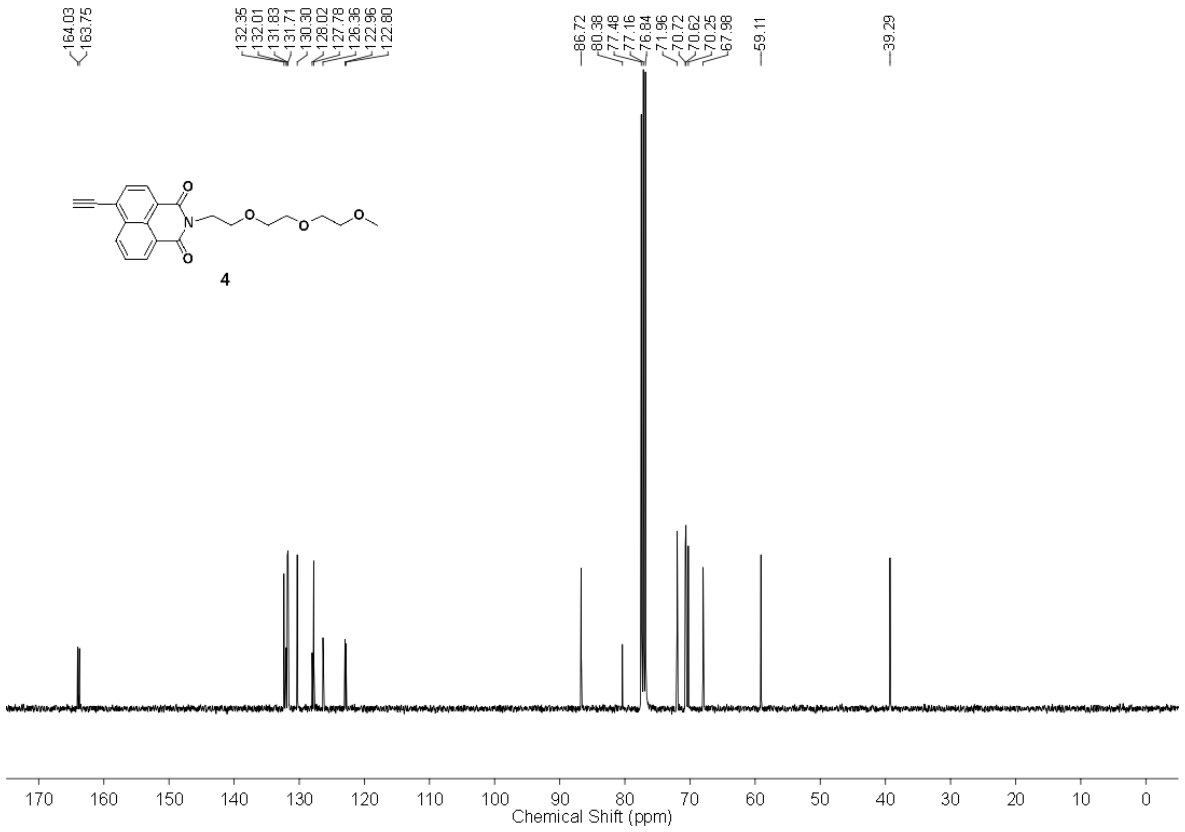
^{13}C-NMR of 4-Bromo napthalimide derivative **3** in $CDCl_3$



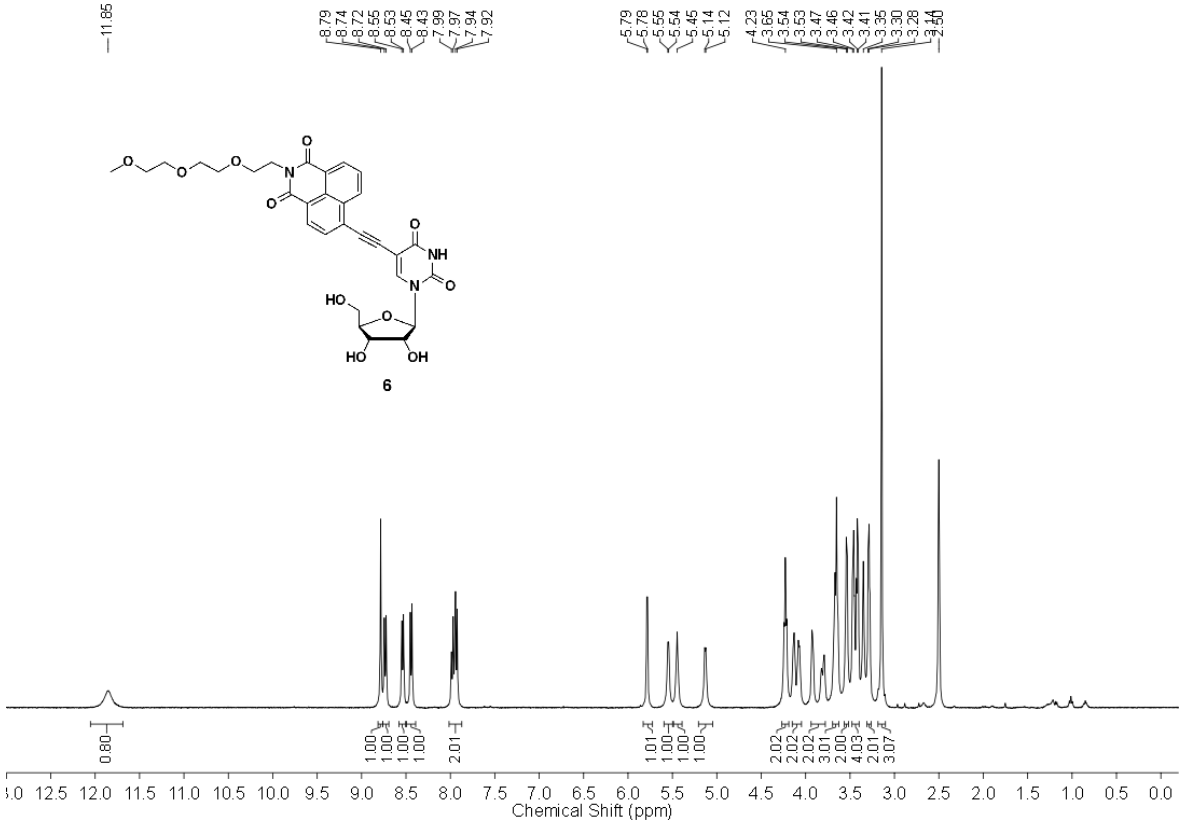
^{1}H-NMR of napthalimide alkyne **4** in $CDCl_3$



^{13}C-NMR of napthalimide alkyne **4** in $CDCl_3$

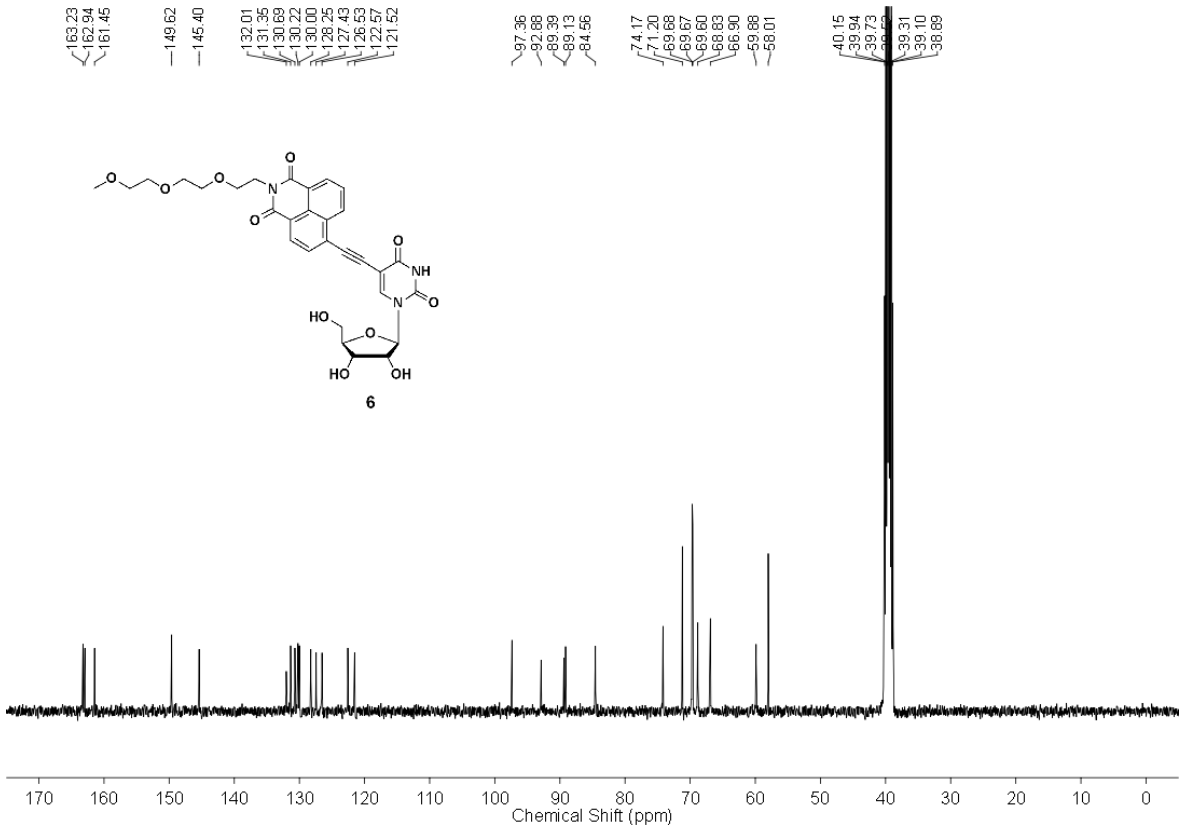


^{1}H-NMR of emissive ribonucleoside **6** in d_6-DMSO



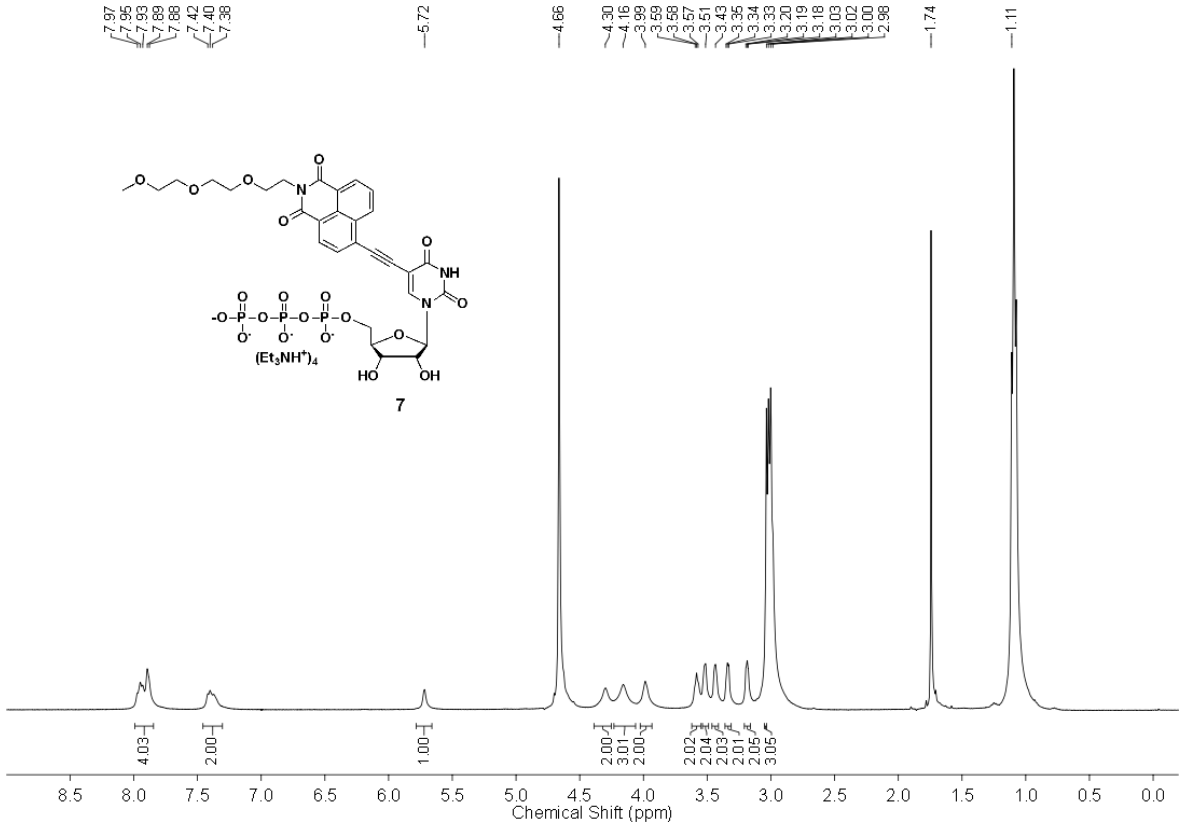
6

^{13}C-NMR of ribonucleoside **6** in d_6-DMSO

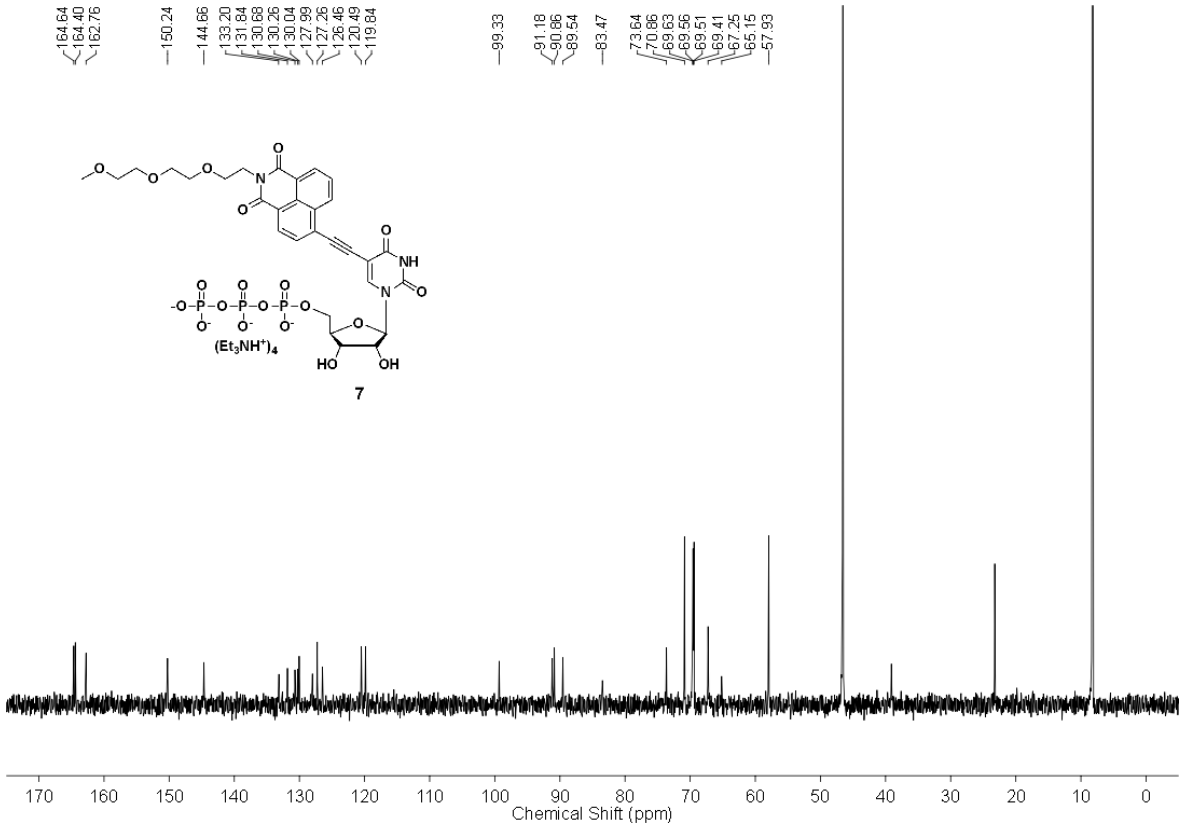


6

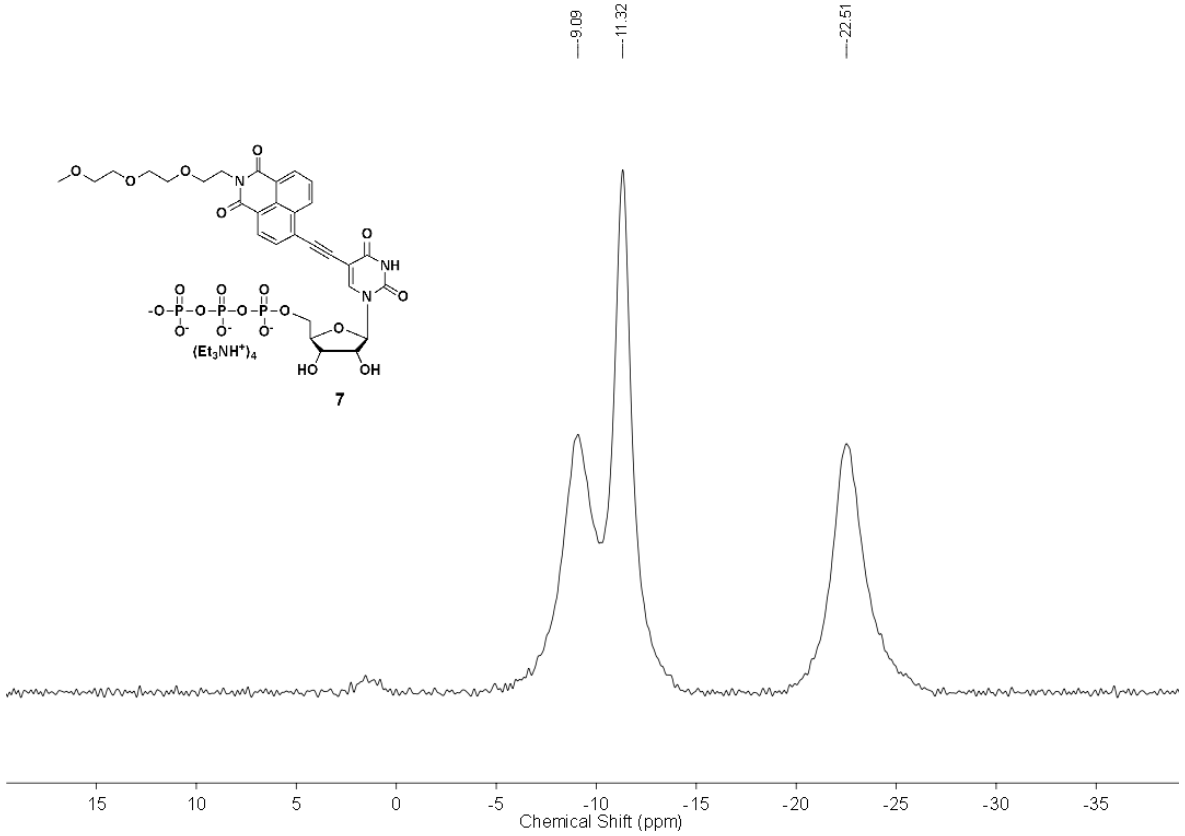
^{1}H-NMR of ribonucleotide **7** in D_2O (contains trace amount of TEAA buffer).



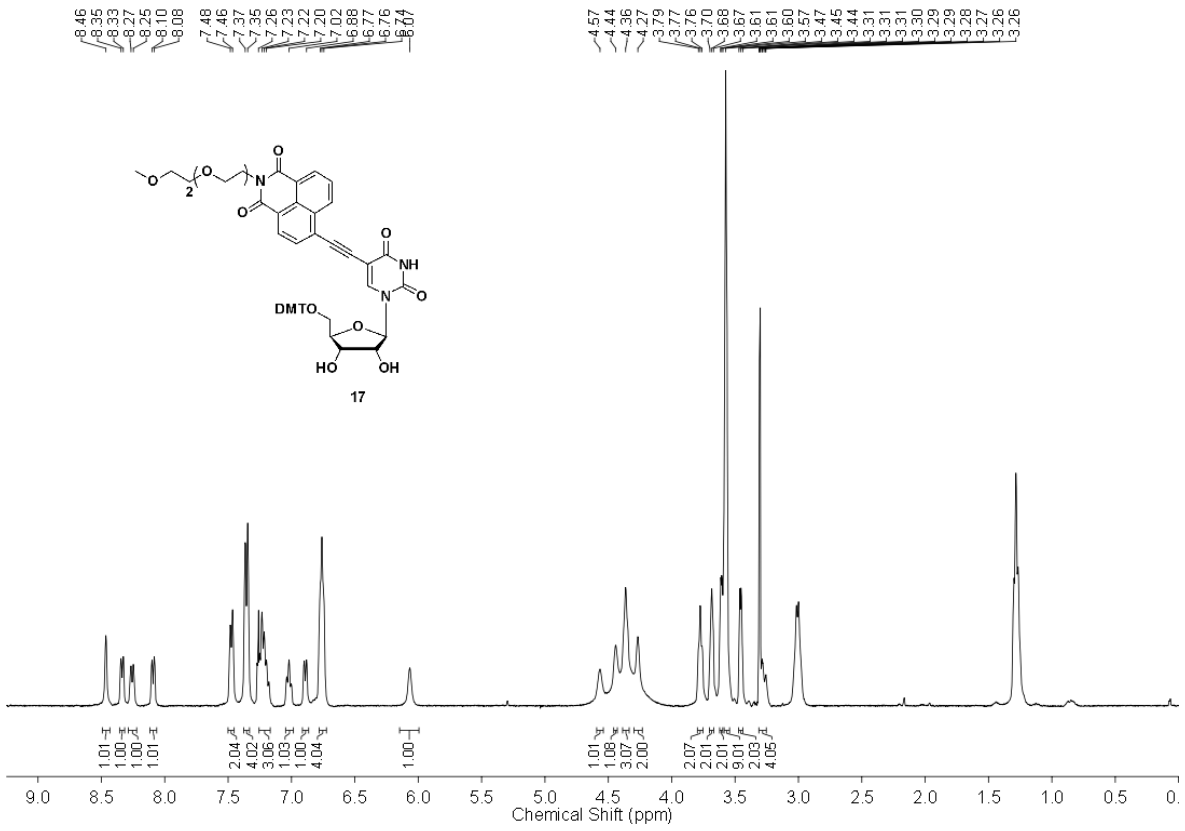
^{13}C-NMR of ribonucleotide **7** in D_2O (contains trace amount of TEAA buffer).



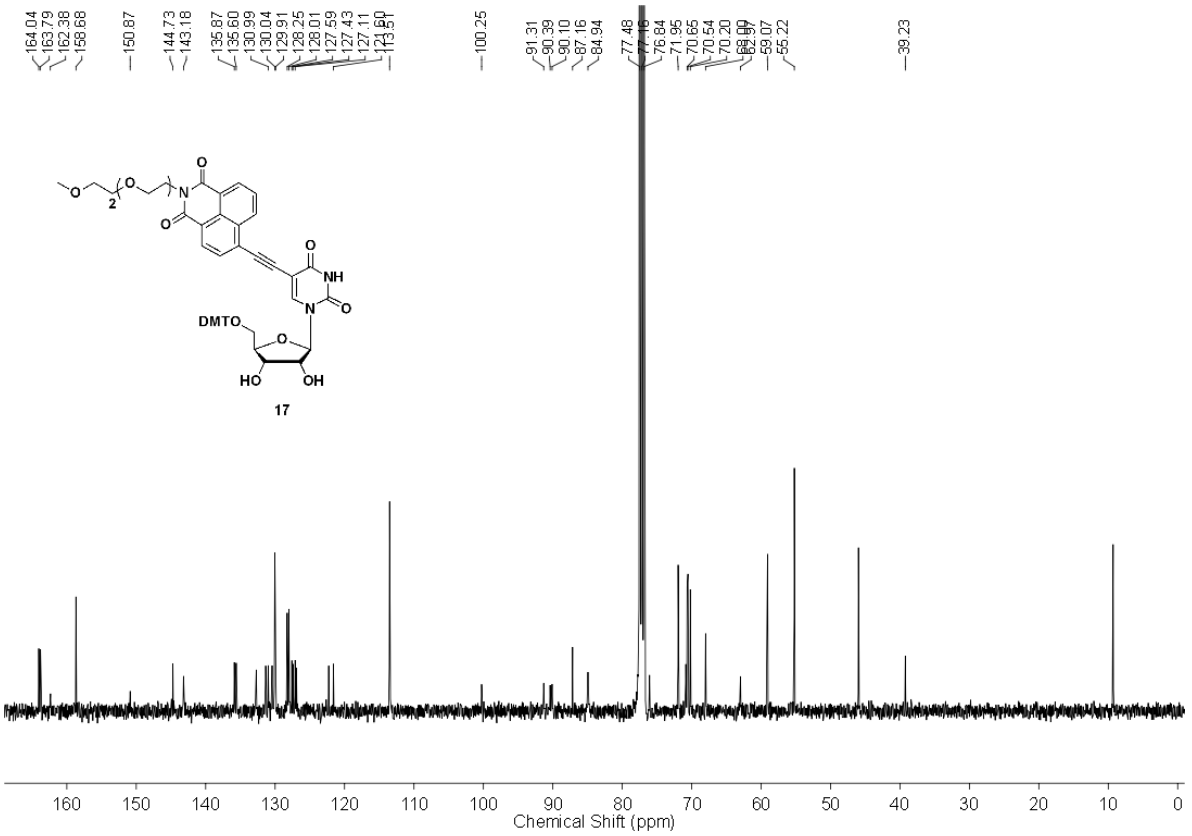
^{31}P-NMR of ribonucleotide **7** in D_2O.



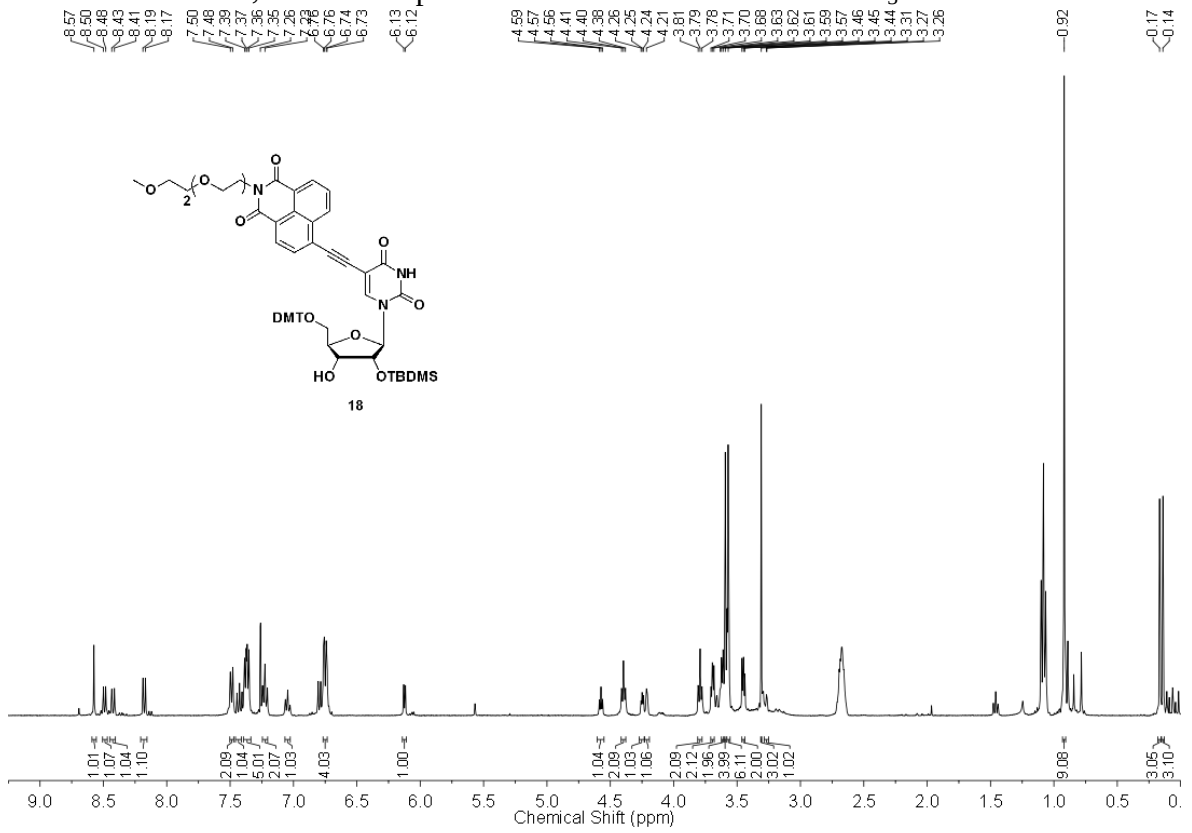
^{1}H-NMR of 5' DMT protected ribonucleoside **17** in $CDCl_3$



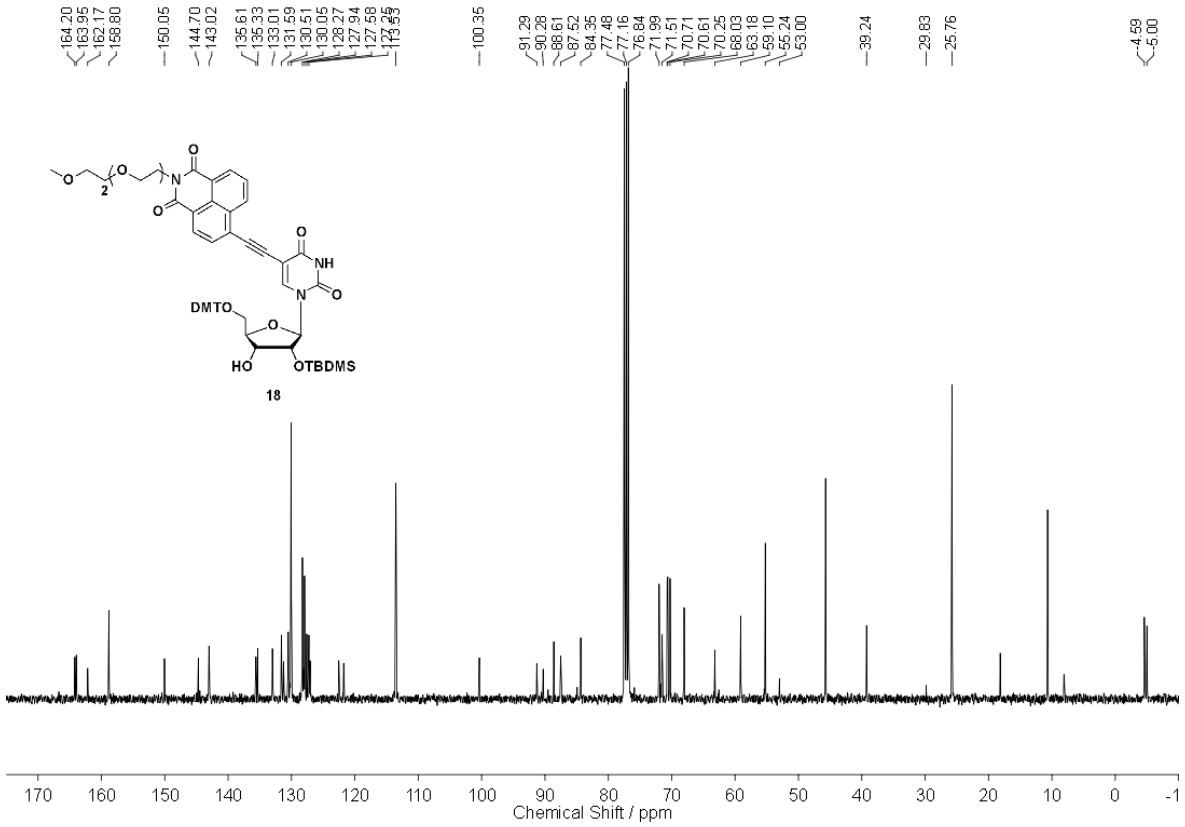
^{13}C-NMR of 5' DMT protected ribonucleoside **17** in $CDCl_3$



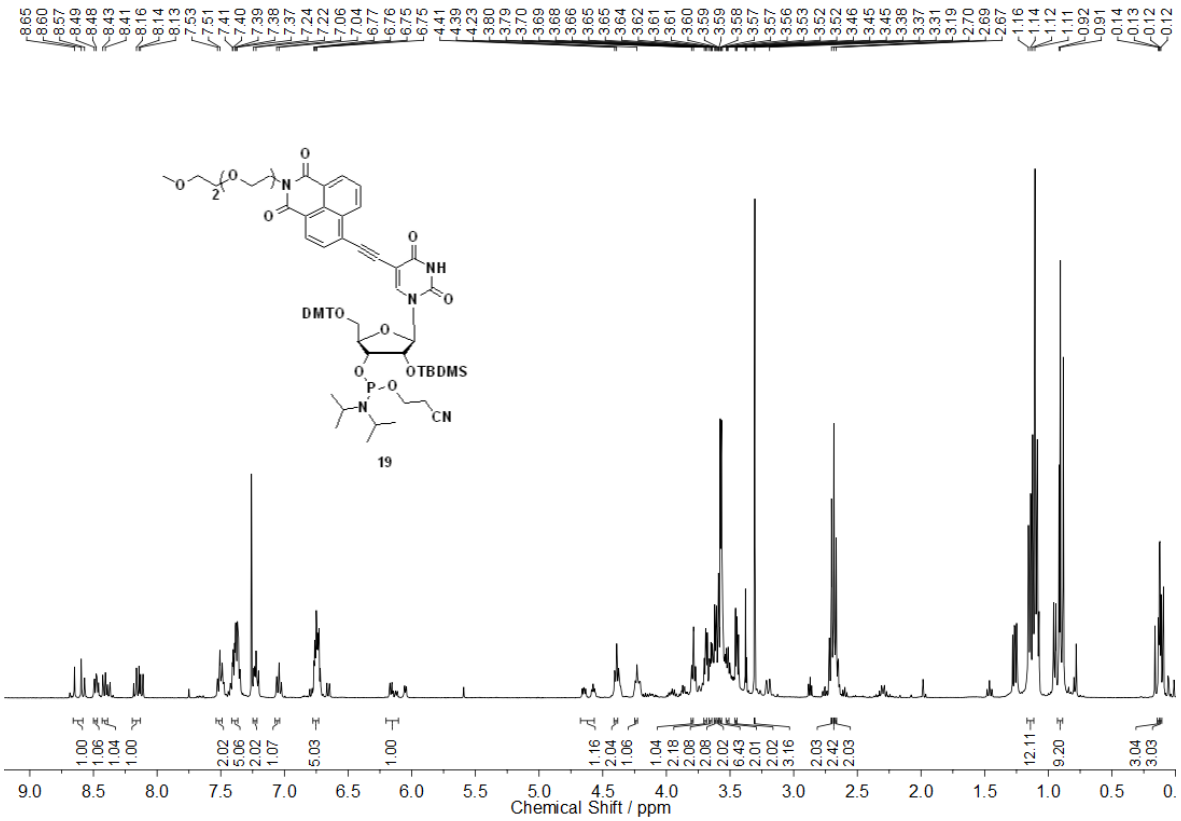
^{1}H-NMR of 5' DMT, 2' TBDMS protected ribonucleoside **18** in $CDCl_3$



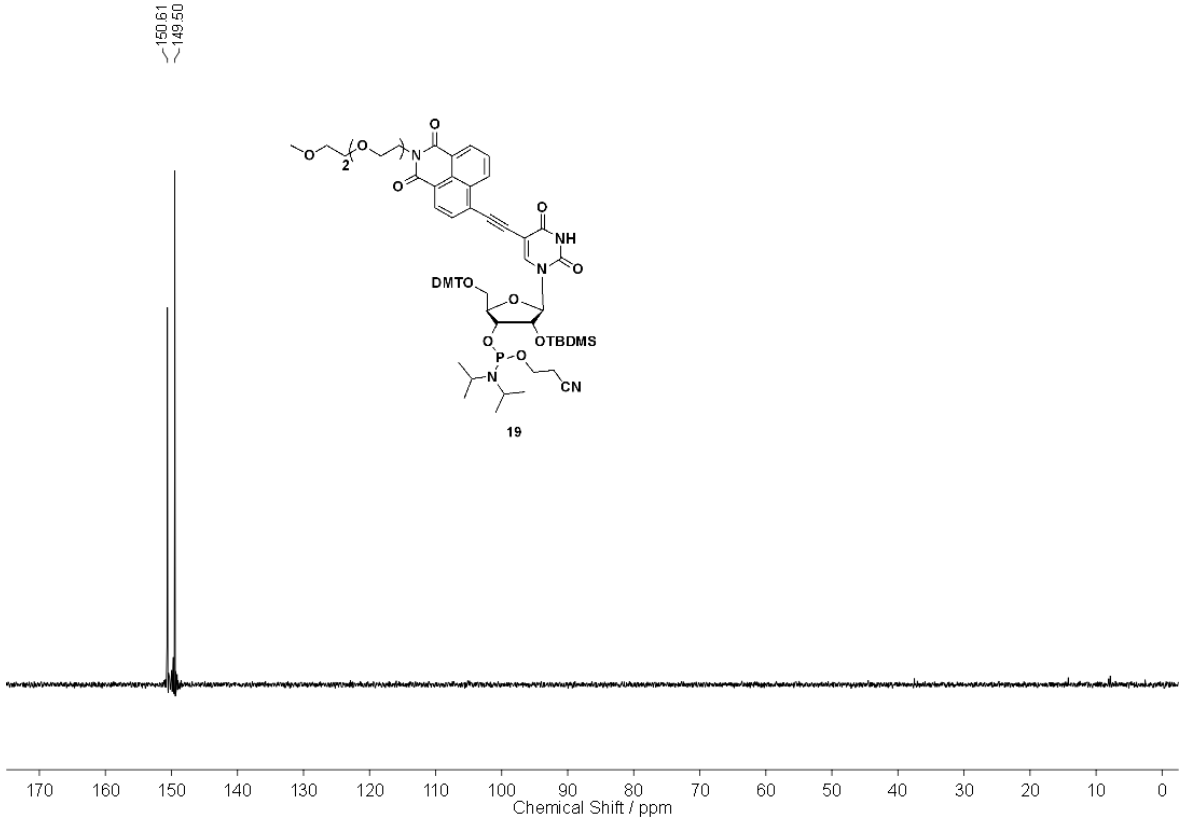
^{13}C-NMR of 5' DMT, 2' TBDMS protected ribonucleoside **18** in $CDCl_3$



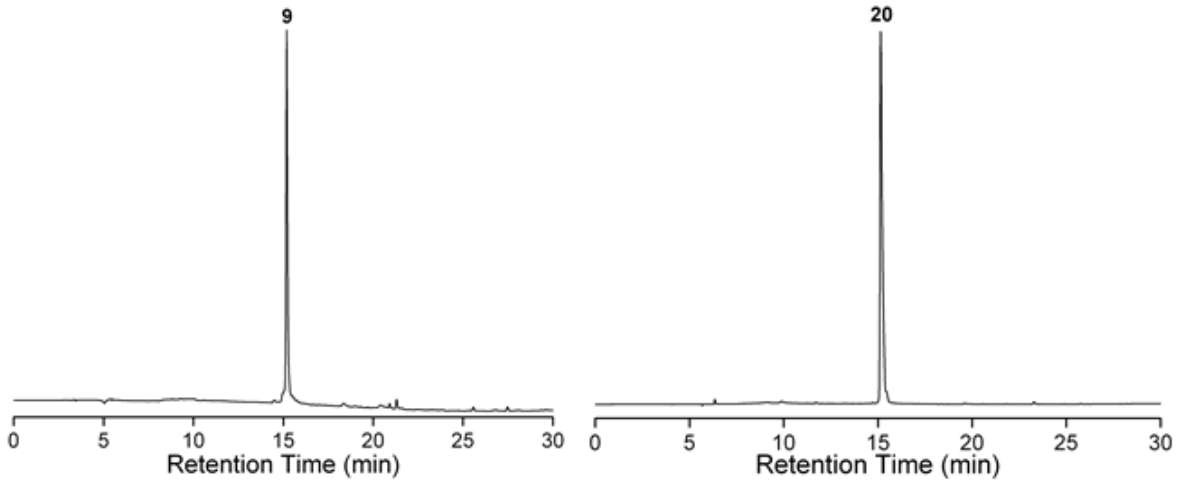
^{1}H-NMR of phosphoramidite **19** in $CDCl_3$ (Containing trace amount of triethylamine)



^{31}P-NMR of phosphoramidite **19** in $CDCl_3$



Representative RP-HPLC chromatogram of PAGE-purified RNA transcript (**9**) and napthalimide modified RNA ON **20** synthesized by solid-phase synthesis method at 260 nm. Mobile phase A = 20 mM triethylammonium acetate buffer (pH 7.2), mobile phase B = acetonitrile. Flow rate = 1 mL/min. Gradient = 0–10% B in 10 min and 10–100% B in 20 min.



Chapter 5

Azide-modified nucleotides a versatile toolbox for posttranscriptional functionalization of RNA by bioorthogonal chemical reactions

5.1 Introduction

Chemical modification of RNA has become indispensible in the study of its structure and function and in the development of nucleic acid-based diagnostic and therapeutic tools. [1-2] Typically, RNA labeling strategies based on solid-phase chemical synthesis and enzymatic methods are sufficient for most in vitro applications. However, analogous labeling strategies for cellular RNA are much less developed. In particular, paucity of efficient RNA imaging tools has been a major impediment in the study of cellular RNA biogenesis, localization and degradation, a combination of processes that tightly regulates gene expression. [3] Methods to visualize RNA commonly rely on metabolic labeling of RNA with ribonucleoside or ribonucleotide analogs such as BrU or BrUTP followed by immunostaining with fluorescent antibody for BrU.[4-5] However, these methods involve laborious assay setups and are not applicable to all cell types and tissue samples due to limited permeability of the antibodies. Endogenous RNA has also been visualized by using fluorescently-modified antisense oligonucleotide (ON) probes,[6-7] molecular beacons,[8] nucleic acid-templated reactions [9-10] and more recently, aptamer-binding fluorophores.[11] Apart from synthetic challenges in preparing the ON probes, these methods also suffer from poor membrane permeability and short half-life of the probes and background fluorescence due to non-specific interactions.[12]

Alternatively, postsynthetic functionalization by using chemoselective reactions (e.g., azide-alkyne cycloaddition, Staudinger ligation) has recently emerged as a valuable method to label glycans, proteins, lipids, and nucleic acids for a variety of applications.[13-18] In this methodology, a nucleoside containing an unnatural reactive group is incorporated into an ON sequence by chemical or enzymatic method. Further functionalization is achieved postsynthetically by performing a chemoselective reaction between the labeled ON and a probe containing the cognate reactive group. While DNA labeling and imaging techniques based on this strategy are well documented,[19-27] postsynthetic RNA manipulations are less prevalent[28-30] as methods developed for DNA often do not work for RNA due to its inherent instability. Moreover, the azide group, which participates in a wide range of bioorthogonal reactions in comparison to alkyne functionality, cannot be easily incorporated into nucleic acids by solid-phase ON synthesis protocols because most azide substrates are unstable in solution and undergo Staudinger-type reaction with phosphoramidite substrates.[31-33] Hence, except for a very few examples wherein the azide group has been incorporated into DNA, [34-35] these procedures mostly use easily accessible alkyne-modified nucleic acids, thereby making this postsynthetic modification method one-dimensional.[36-41]

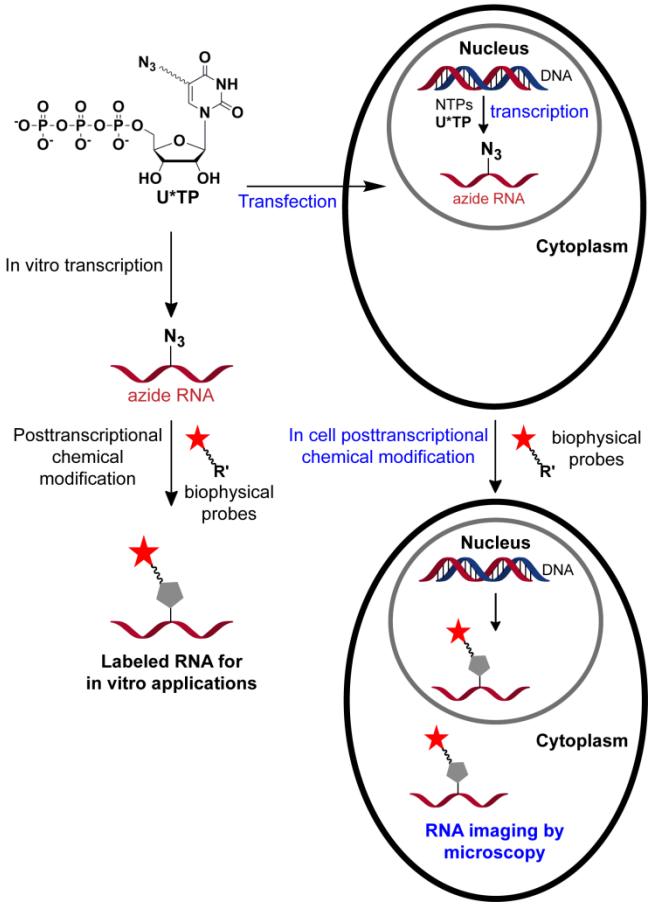


Figure 1. Schematic diagram illustrating the posttranscriptional chemical labeling of RNA transcripts in vitro and in cells by using azide-modified UTP analogs.

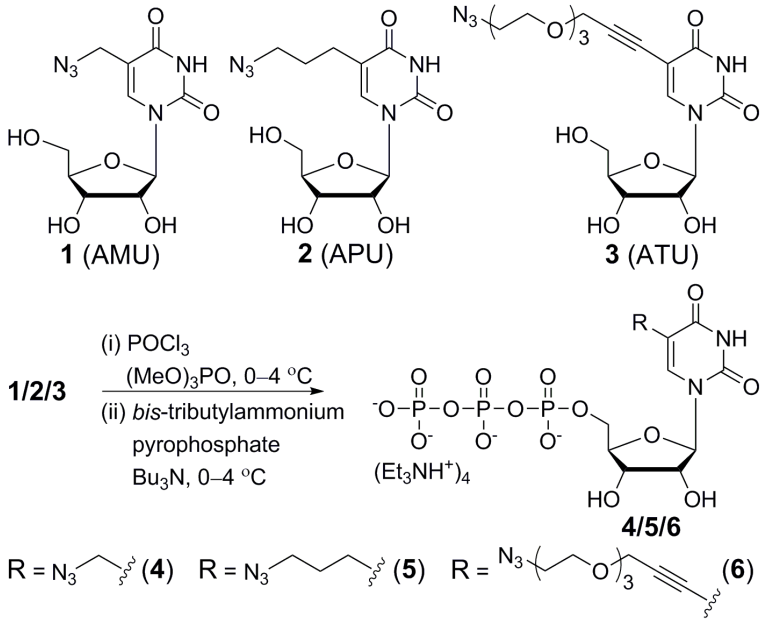
Owing to these practical problems in current labeling procedures and paucity of efficient RNA imaging tools, we sought to develop a robust and modular labeling tool that would enable the study of RNA *in vitro* as well as in cells. Towards this endeavour, we have recently reported the effective incorporation of an azide group into short RNA ONs by *in vitro* transcription reactions using 5-azidopropyl-modified UTP analogue.[42-43] The azide-modified RNA ONs were suitable for posttranscriptional chemical modification by copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) and Staudinger reduction reactions. Encouraged by these results we wanted to develop a small series of azide-modified nucleotide analogs, which would enable the detailed investigation of the utility of our azide labeling technique to functionalize RNA with biophysical probes by CuAAC, copper-free strain-promoted azide-alkyne cycloaddition (SPAAC) and azide-phosphine Staudinger ligation reactions both *in vitro* and in cells. Here, we describe the development of a versatile toolbox composed of azide-modified uridine triphosphates, which facilitates the direct incorporation of azide functionality into RNA transcripts by transcription reaction (Figure 1). The azide-

modified RNA is readily functionalized with biophysical probes in a modular fashion by CuAAC, copper-free strain-promoted azide-alkyne cycloaddition (SPAAC) and azide-phosphine Staudinger ligation reactions.

5.2 Result and Discussion

5.2.1 Synthesis of azide-modified UTP analogs

5-Modified pyrimidine ribonucleotides are known to serve as good substrates in in vitro transcription reaction catalyzed by T7 RNA polymerase.[44,46] Therefore, we assembled a toolbox containing 5-azido-modified UTP analogs (**4**–**6**) to perform posttranscriptional chemical labeling and imaging of RNA (Scheme 1). The ribonucleosides (**1**–**3**) were synthesized according to analogous procedures reported in the literature.[47] The corresponding modified triphosphate substrates (**4**–**6**) necessary for transcription reactions were obtained by first phosphorylating the nucleoside with $POCl_3$ followed by a reaction with *bis*-tributylammonium pyrophosphate (Scheme 1).[42]



Scheme 1. General scheme for the synthesis of azide-modified UTP analogs **4–6** from ribonucleosides **1–3**, respectively.[47]

5.2.2 Enzymatic incorporation of azide-modified UTP analogs in to RNA

The suitability of transcription reaction in producing azide-modified RNA was first evaluated in vitro by using a series of short promoter-template DNA duplexes and T7 RNA polymerase (Figure 2). Deoxyadenosine (dA) residue was placed in the coding region of the templates at

one or two sites to direct the incorporation of azide-modified UTP into RNA transcripts. The templates were also designed to contain a dT residue at the 5'-end to direct the incorporation of a unique A residue at the 3'-end of each transcript. Transcription reactions were performed with templates **T1**–**T5** in the presence of GTP, CTP, UTP/modified UTP and α-^{32}P ATP. The products were then resolved by polyacrylamide gel electrophoresis (PAGE) and phosphorimaged. In this experimental setup successful transcription reactions producing 3'-end radiolabeled full-length RNA transcripts will be detected and all failed reactions resulting in the formation of truncated transcripts will remain undetected.

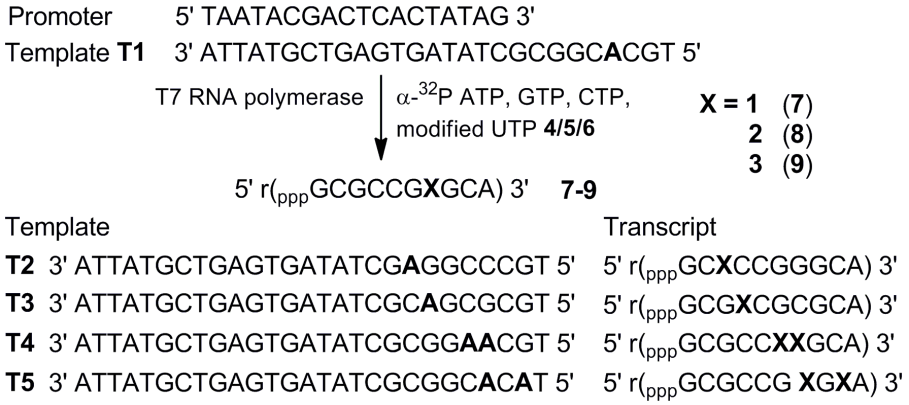


Figure 2. Incorporation of azide-modified UTP analogs (**4**–**6**) into RNA ONs by in vitro transcription reactions. Incorporation of **4**–**6** in the presence of DNA template **T1** will produce RNA transcripts **7**–**9**, respectively. Transcripts resulting from **T2**–**T5** are also shown.[47]

Transcription reactions performed in the presence of template **T1** and modified UTPs **4–6** resulted in very high yields of azide-modified full-length RNA transcripts **7–9**, respectively, in comparison to a control reaction in the presence of natural UTP (Figure 3, lane 2). The retarded mobility of modified transcripts as compared to control unmodified transcript indicated the incorporation of heavier nucleotides into transcripts (Figure 3, compare lanes 1 and 2). In addition to full-length product, non-templated random incorporation of nucleotides resulted in the formation of trace amounts of longer transcripts. This observation is not unusual as in vitro transcription reactions with short templates are known to produce N+1 and N+2 products.[48] Furthermore, absence of a band corresponding to full-length transcript in a control reaction performed with no UTP or modified UTP ruled out any misincorporation during the transcription reaction (Figure 3, lane 3). To test the preference of RNA polymerase, reactions were performed in the presence of 1:1 concentration of natural UTP and modified UTP. Rewardingly, the enzyme incorporated

both natural as well as modified UTPs reasonably well, indicating that the azide group could be potentially incorporated into cellular RNA transcripts by using cell's biosynthetic machinery (Figure 3, lane 4). Furthermore, reactions with templates **T2**–**T5** revealed that the modified uridine analogs could be incorporated into RNA transcripts near the transcription initiation site (TATA box) and at more than one site with moderate to high efficiency (Figure 3, lanes 6, 8, 10 and 12).

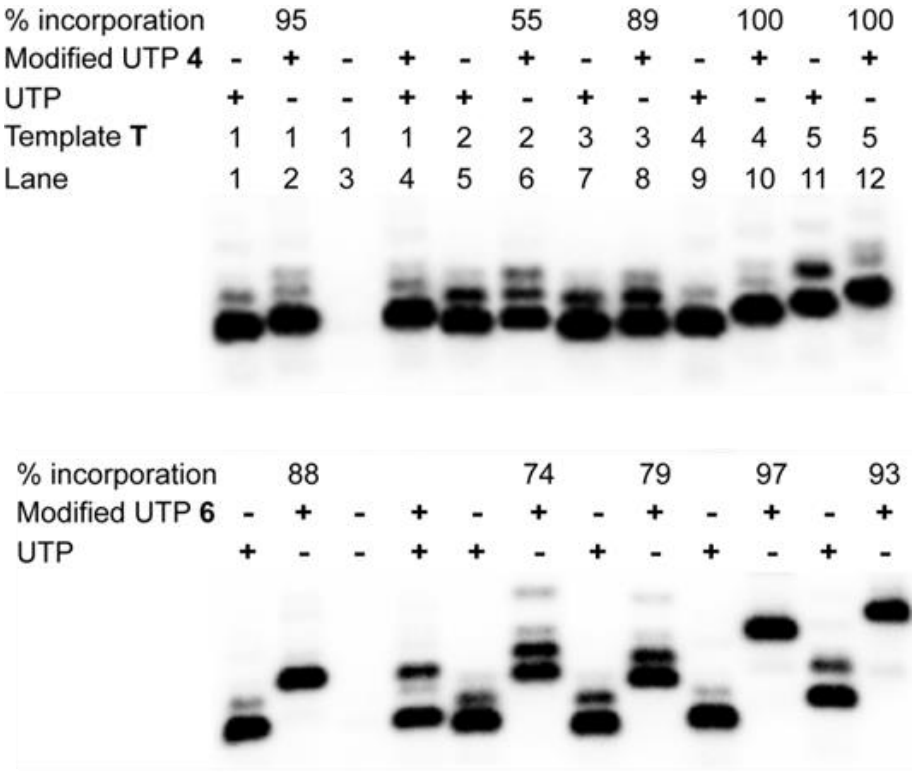


Figure 3. Incorporation of modified UTPs into RNA ONs by T7 RNA polymerase. Transcription reactions were performed with templates **T1**–**T5** in the presence of UTP and or modified UTPs **4** and **6**. Relative % incorporation of azide-modified UTPs into modified full-length transcript is given with respect to the amount of full-length transcript formed in the presence of natural NTPs.

To unambiguously confirm the incorporation of azide functionality into RNA ONs, large-scale transcription reactions were performed in the presence of template **T1** and modified UTPs **4**−**6**. Mass analysis of transcripts isolated after PAGE purification clearly confirmed the integrity of modified full-length RNA products (Table 1). In addition, HPLC analysis of ribonucleoside products obtained from digestion of modified transcripts clearly ascertained the incorporation of azide-modified uridine analogs into RNA (Figure 4, 5).

Table 1. ε_{260}, isolated yield and MALDI-TOF mass data of azide-modified RNA transcripts obtained from large-scale transcription reactions with template **T1**.

RNA Transcript	ε_{260} ($M^{-1}cm^{-1}$)	Isolated yield (nmol)[a]	Calcd. mass	Observed mass
7	90600	12	3469.9	3469.9
8	90133	14	3498.0	3497.2
9	84700	18	3670.2	3669.8

[a] Yield from large-scale transcription reactions (250 μL volume reaction).

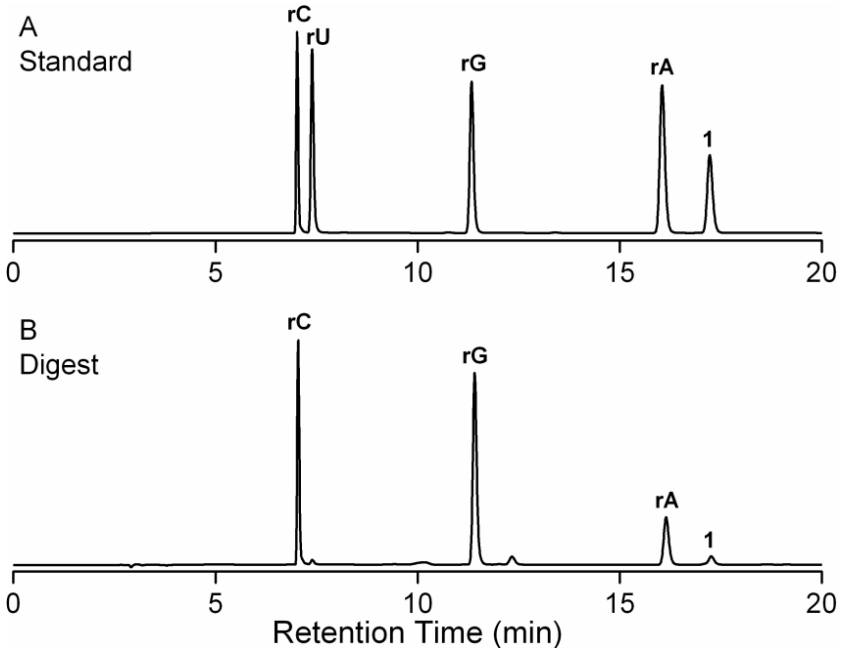


Figure 4. HPLC chromatogram of ribonucleoside products obtained from an enzymatic digestion of oligoribonucleotide **7** at 260 nm. (A) A mix of natural ribonucleosides and azide-modified ribonucleoside **1**. (B) ON 7 digest. Mobile phase A: 100 mM TEAA (pH 7.5); mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min. Mass analysis of HPLC fraction corresponding to ribonucleosides also confirmed the integrity of the ribonucleosides (data not shown). HPLC analysis was performed on Agilent Technologies 1260 Infinity.

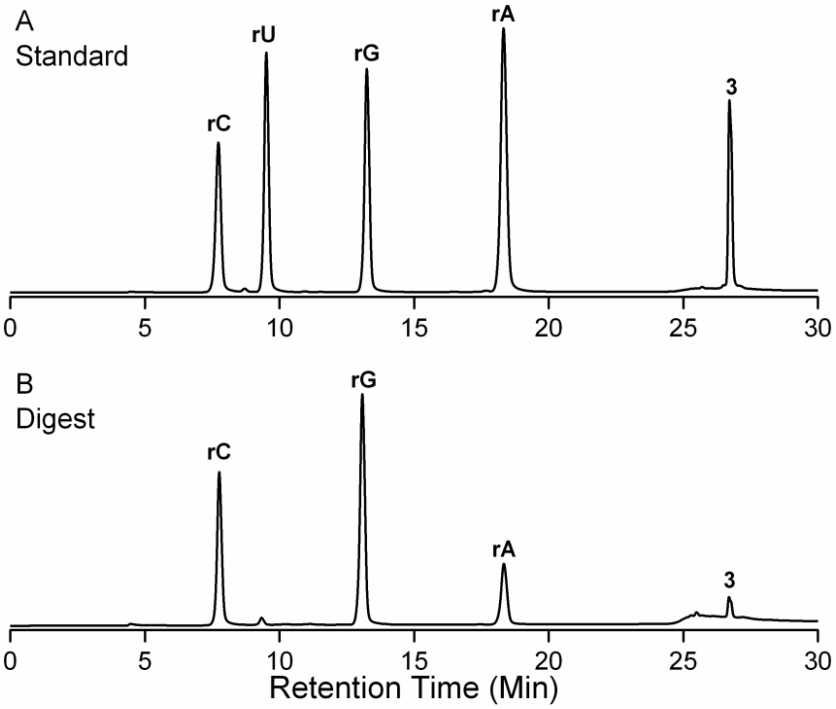


Figure 5. HPLC chromatogram of ribonucleoside products obtained from an enzymatic digestion of oligoribonucleotide **9** at 260 nm. (A) Natural ribonucleosides and azide-modified ribonucleoside **3** mix. (B) ON **9** digest. Mobile phase A: 100 mM TEAA (pH 7.5); mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min. Mass analysis of HPLC fraction corresponding to ribonucleosides also confirmed the integrity of the ribonucleosides (data not shown). HPLC analysis was performed on Dionex ICS 3000.

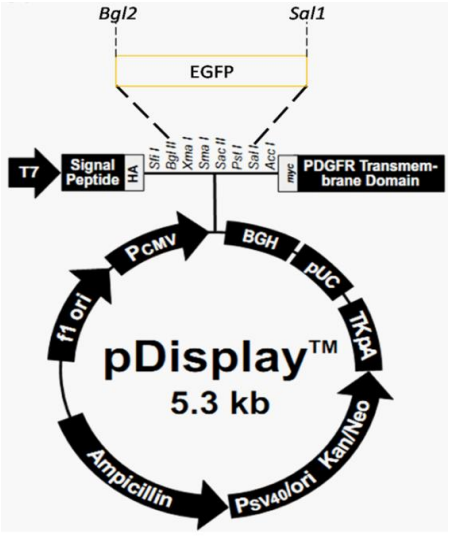


Figure 6. Schematic diagram of EGFPinpDisplay plasmid, which was linearized by using XhoI restriction enzyme. Transcription reactions with linearized plasmid in the presence of natural or modified UTPs will produce nearly 1.1 kb RNA transcript containing 258 uridine or modified uridine residues, respectively.[56]

The efficacy of RNA polymerase to incorporate azide groups at multiple sites in RNA sequences of biological relevance was evaluated by performing transcription reactions in the presence of linearized plasmid DNA that would generate nearly 1.1 kb RNA transcript (figure 6).[47] All modified UTPs (**4**−**6**) were found to be effectively incorporated by RNA polymerase to produce respective transcripts. As before, HPLC analysis of ribonucleoside products obtained from enzymatic digestion of transcripts confirmed the incorporation of modification into the transcripts (Figure 7). The fidelity of transcription reaction in the presence of modified UTPs (e.g., **4**) was further confirmed by reverse transcribing modified transcripts into cDNA, which was then PCR amplified and sequenced. The sequencing data showed 100% sequence identity with the template DNA used for the in vitro transcription reactions. These results clearly reveal that (i) the azide-modified UTPs are incorporated into RNA transcripts with high fidelity and (ii) the transcription products are recognized and copied by reverse transcriptase with high fidelity to generate cDNA, which can be PCR amplified and transcribed again in the presence of modified UTPs. This feature of nucleoside analogs could be highly beneficial in expanding the structural and functional diversity of RNA library used in the in vitro selection protocols.[49,50]

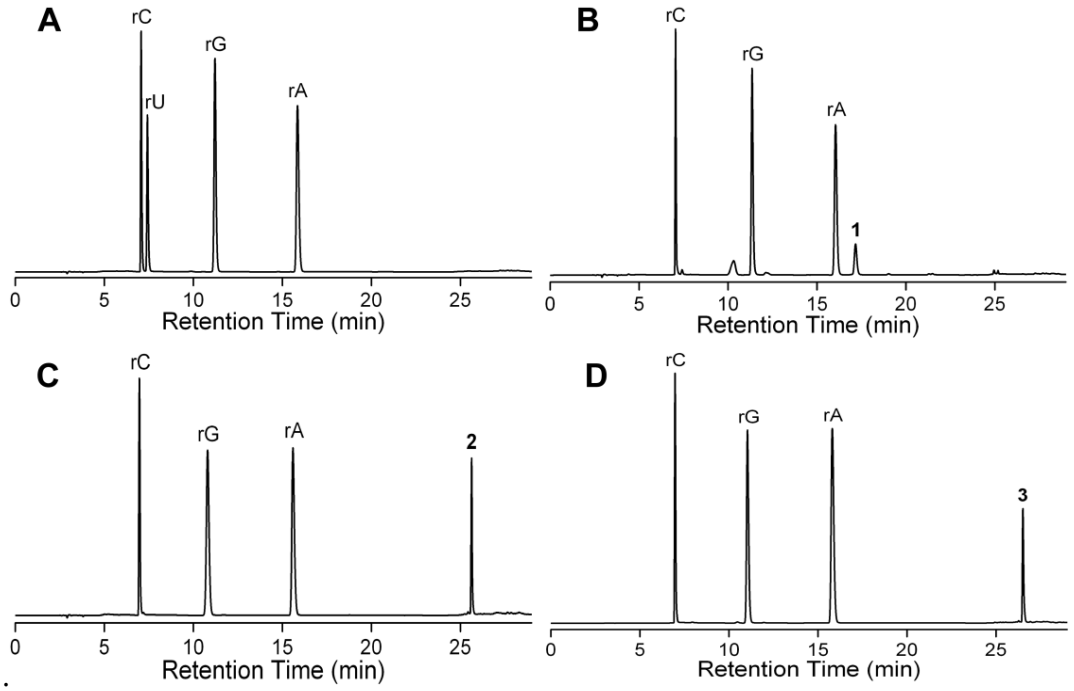


Figure 7. HPLC chromatogram of ribonucleoside products obtained from enzymatic digestion reactions of RNA transcripts, which were prepared by in vitro transcription reaction with linearized EGFPinpDisplay plasmid template and in the presence of **A**) natural UTP; **B**) AMUTP (**4**); **C**) APUTP (**5**); **D**) ATUTP (**6**). The presence of modified ribonucleosides (**1**–**3**) in RNA transcripts was confirmed by comparing the retention times with standard ribonucleoside mix.

5.2.3 Posttranscriptional chemical functionalization of azide-modified RNA ONs

The compatibility of azide-modified RNA ONs **7–9** to posttranscriptional chemical functionalization was evaluated by performing click and Staudinger ligation reactions in the presence of a variety of biophysical probes (Figure 8). First, the azide-modified RNA ONs **7–9** were subjected to CuAAC reaction with alkyne substrates in the presence of a water-soluble Cu(I) stabilizing ligand, tris-(3-hydroxypropyltriazolylmethyl)amine (THPTA), [36,51] $CuSO_4$ and sodium ascorbate at 37 °C. Typically, reaction mixture after 30 min was resolved by PAGE under denaturing conditions, and the band corresponding to the product was isolated. The cycloaddition reaction with naphthalimide-alkyne (**10**), Alexa594-alkyne (**11**) and biotin-alkyne (**12**) was completed in 30 min to afford the desired fluorescent- and biotin-labeled RNA ONs in moderate to good yields (Scheme 2, 10, Table 2). Reactions between ON **7–9** and biotin-alkyne substrate gave the expected clicked products and also an unknown slower migrating side–product in trace amounts (Figure 9A).

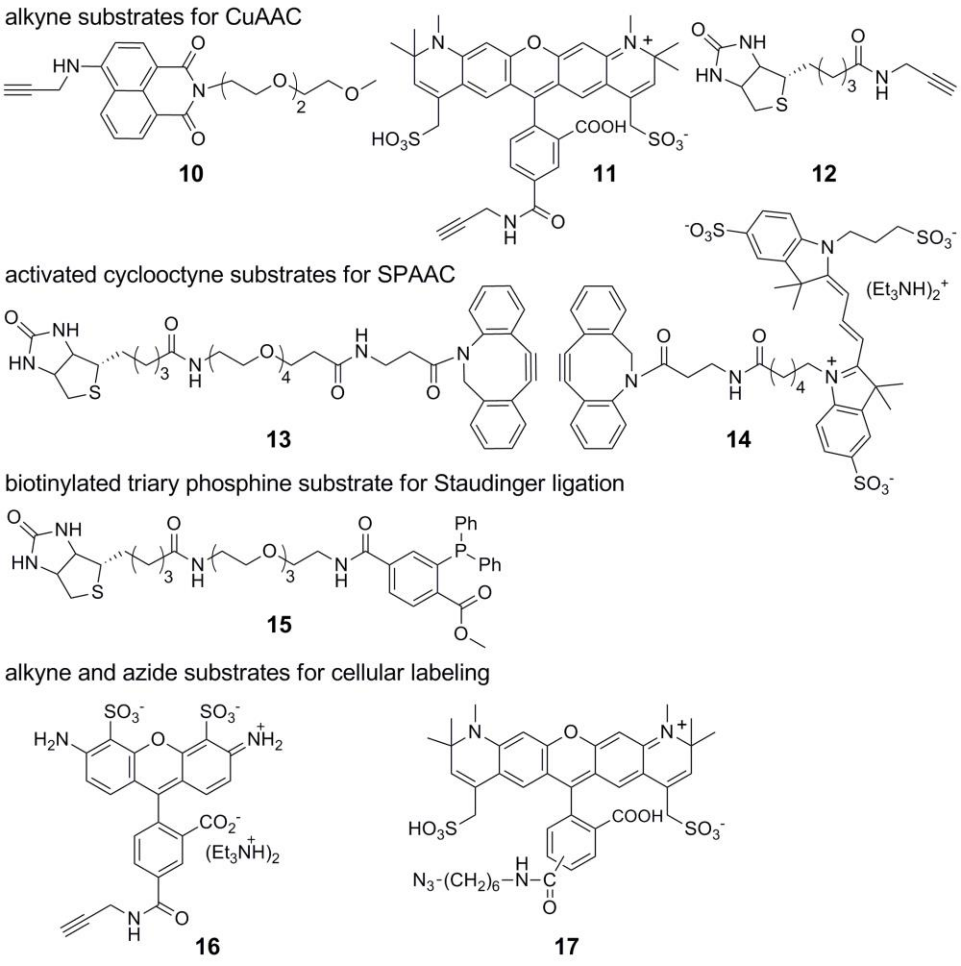
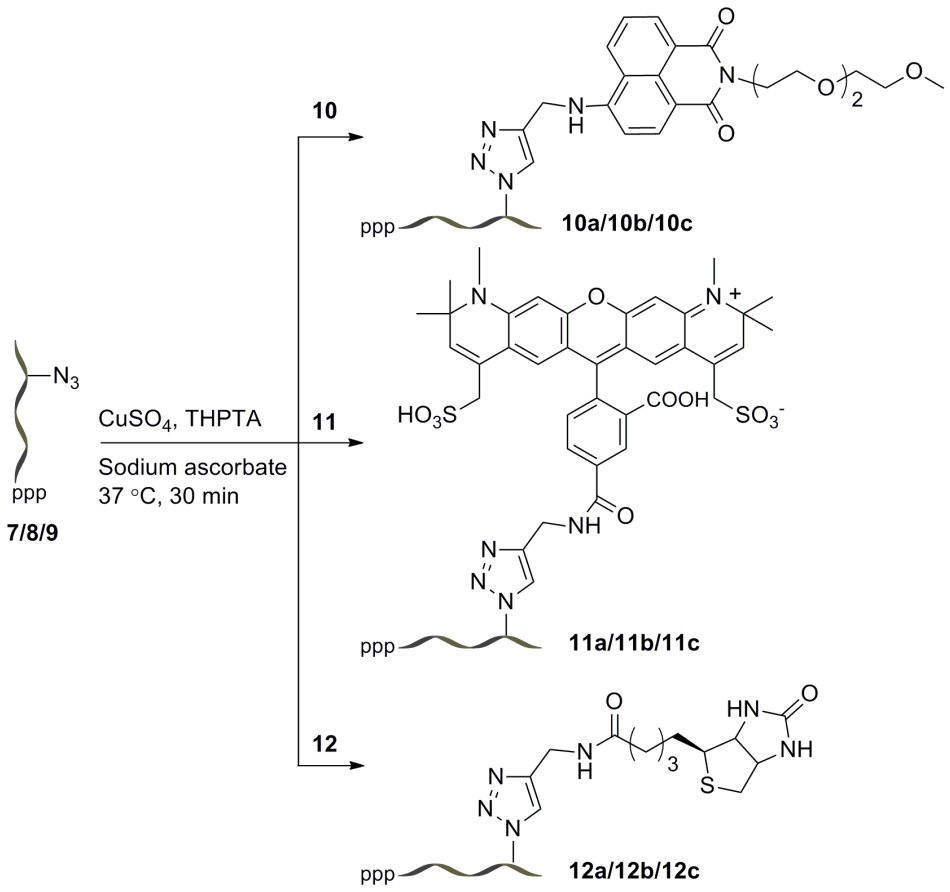
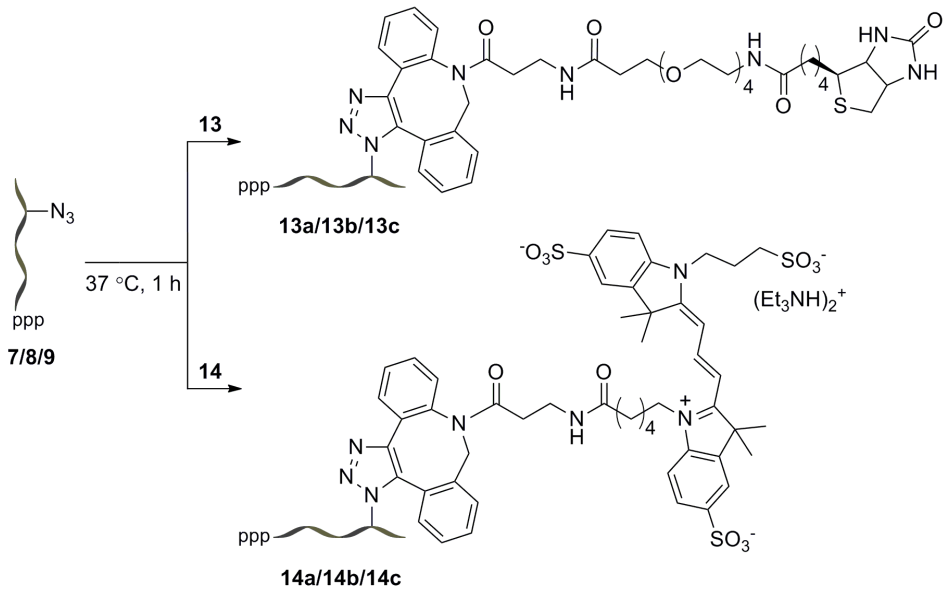


Figure 8. Building blocks for posttranscriptional chemical modification.[47]

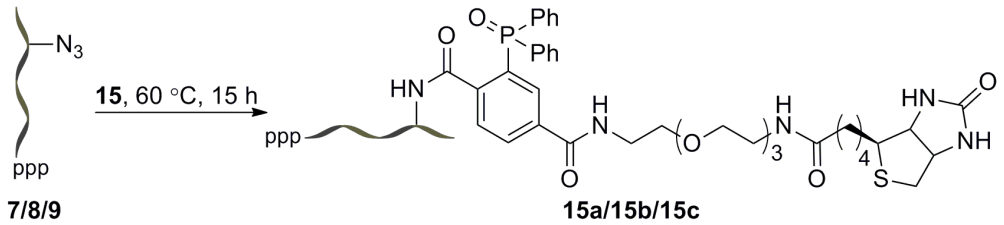


Scheme 2. Posttranscriptional chemical functionalization of azide-modified RNA transcripts **7**, **8** and **9** by CuAAC reaction using alkyne substrates **10–12**.



Scheme 3. Posttranscriptional chemical functionalization of azide-modified RNA transcripts **7**, **8** and **9** by copper–free SPAAC reaction with alkyne substrates **13** and **14**.

On the other hand, RNA ONs **7–9** underwent facile copper-free SPAAC reaction with commercially available cyclooctyne building blocks of biotin (**13**) and Cy3 (**14**) in 60 min at 37 °C (Scheme 3, 10, Table 2) to produce the clicked products in moderate to very high yields. In SPAAC reaction conditions, we did not detect any side–products. Next, the azide-modified ONs were subjected to Staudinger ligation reaction with biotinylated triaryl phosphine substrate **15**. Since the kinetics of the ligation reaction is slow,[13] the desired biotinylated RNA products (**15a-15c**) were formed in isolable amounts when the reaction was performed for 15 h at 60 °C (Scheme 4, Table 2). Some amount of unreacted azide-modified RNA transcript was noticed as the reaction did not proceed to completion (Figure 9C).



Scheme 4. Posttranscriptional chemical functionalization of azide-modified RNA transcripts **7**, **8** and **9** by Staudinger ligation reaction with triaryl phosphine substrate **15**

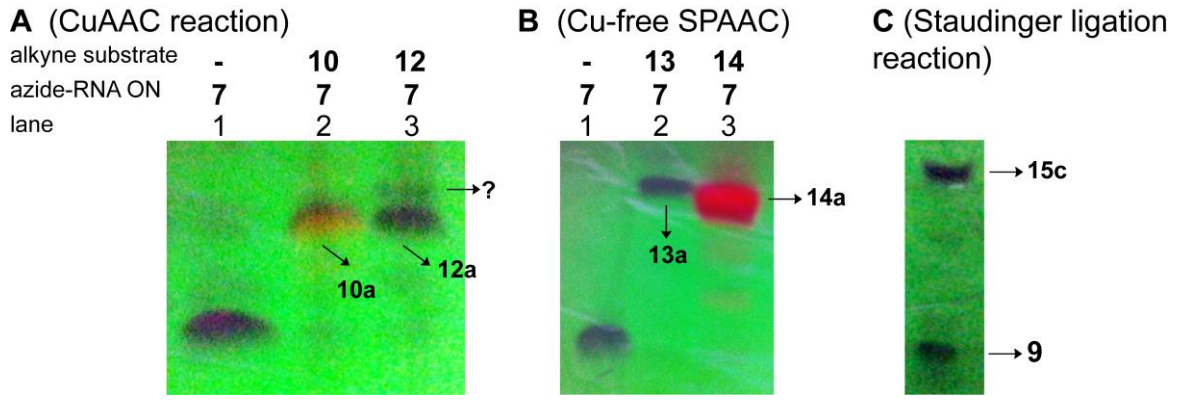


Figure 9. Representative UV shadowed (254 nm) images of polyacrylamide gels of RNA ON products obtained from posttranscriptional chemical modification by click reactions. **(A)** CuAAC reaction between azide-modified RNA ON **7** and alkyne substrates **10** and **12**. **(B)** Copper-free SPAAC reaction between azide-modified RNA ON **7** and alkyne substrates **13** and **14**. Bands corresponding to clicked products are indicated using arrows. The unknown slower migrating bands formed under CuAAC conditions are denoted using "?". Mass analysis indicates that the slower migrating bands could probably be the sulfoxide form of the biotinylated RNA products (data not shown). Conversion of biotin to biotin sulfoxide has been reported earlier.[7] (**C**) Staudinger ligation reaction between RNA ON **9** and biotinylated triaryl phosphine substrate **15**. Some amount of unreacted azide-modified RNA transcript was noticed as the reaction did not proceed to completion.

Notably, ATU (**3**) labeled RNA ON **9** containing a longer tetraethyleneglycol spacer gave the best yield with all substrates as compared to ONs **7** and **8** containing shorter linkers. Under these posttranscriptional modification conditions we observed minimum degradation of RNA (Figure 9). Typically, 15 nmol reaction scale yielded 4–12 nmol of the labeled product after gel electrophoretic purification. The integrity of all clicked and ligated RNA products was confirmed by mass analysis (Table 2). Taken together, these results clearly demonstrate that this posttranscriptional chemical labeling methodology is simple, modular, and generates labeled RNA ONs in sufficient amounts for biophysical analysis.

Table 2. Yield and mass data of posttranscriptionally functionalized RNA ONs **7**–**9**. See Schemes 5–7 for the chemical structure of the products.

Entry	RNA ON	Substrate	Product	Yield (nmol)	Isolated yield (%)[a]	MALDI-TOF mass analysis of products (*m/z*)	
						Calculated mass	Observed mass
1	**7**	**10**	**10a**	6	40	3867.4	3866.6
2	**8**	**10**	**10b**	8	53	3895.4	3895.7
3	**9**	**10**	**10c**	10	67	4066.6	4067.2
4	**7**	**11**	**11a**	9	60	4230.8	4230.6
5	**8**	**11**	**11b**	10	67	4258.8	4258.8
6	**9**	**11**	**11c**	11	73	4431.0	4431.3
7	**7**	**12**	**12a**	5	33	3751.3	3751.3
8	**8**	**12**	**12b**	5	33	3779.4	3778.8
9	**9**	**12**	**12c**	5	33	3951.5	3950.5
10	**7**	**13**	**13a**	8.4	56	4219.8	4220.0
11	**8**	**13**	**13b**	8.8	59	4247.9	4248.2
12	**9**	**13**	**13c**	10.2	68	4420.1	4419.7
13	**7**	**14**	**14a**	9.6	64	4454.1	4453.5
14	**8**	**14**	**14b**	9.7	65	4482.2	4482.4
15	**9**	**14**	**14c**	12.4	83	4654.4	4654.7
16	**7**	**15**	**15a**	4	27	4112.8	4112.9
17	**8**	**15**	**15b**	4	27	4140.8	4141.1
18	**9**	**15**	**15c**	7	47	4313.0	4312.5

[a] All reactions were performed on 15 nmol scale of modified RNA ONs. Yields reported are with respect to the RNA products isolated after PAGE purification.

5.3 Conclusions

We have developed a robust and modular posttranscriptional chemical functionalization protocol to label RNA *in vitro* with biophysical probes by using a novel toolbox made of azide-modified UTP analogs. These modified UTPs, synthesized in few steps, are effectively incorporated by RNA polymerases to endow RNA with azide functionality. Further, the azide-modified transcripts are conveniently labeled posttranscriptionally by CuAAC, SPAAC and Staudinger ligation reactions with a variety of substrates ranging from fluorescent probes to affinity tags.

One of my colleagues Anupam Sawant has taken forward this azide modified UTPs and evaluated their metabolic incorporation efficiency into cellular RNA. He has successfully demonstrated the incorporation of azide-modified UTP analog (AMUTP **4**) by endogenous RNA polymerases, which is the first example of selective labeling of cellular RNA transcripts with azide groups. The selective labeling of RNA with azide group enabled the imaging of newly transcribing RNA in fixed and in live cells by click reactions. It is expected that this modular and practical chemical labeling methodology will provide a new platform to study RNA *in vitro* and inside the cells.

5.4 Experimental Section

5.4.1 Materials

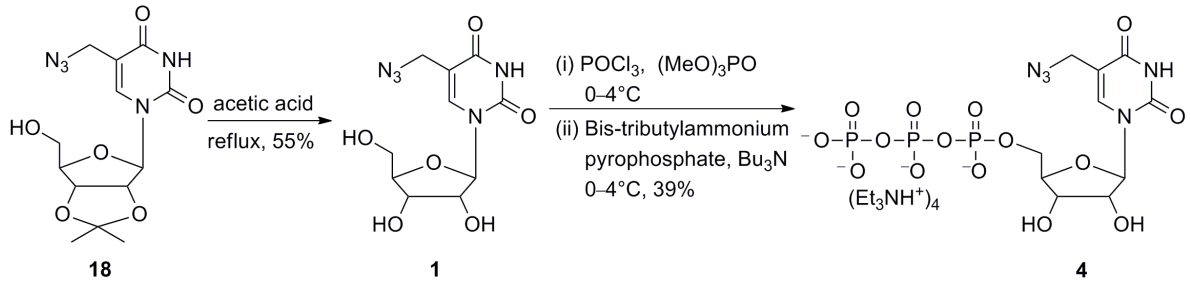
Uridine, 5-iodouridine, tetrakis(triphenylphosphine) palladium (0), copper(I) iodide, paraformaldehyde, 4-bromo-1,8-naphthalic anhydride, diphenylphosphine, 1-methyl 2-iodoterephthalate, triethylene glycol monomethyl ether, tetraethylene glycol, propargylamine, sodium azide, sodium ascorbate, $CuSO_4$, biotin, DEAE Sephadex A-25 resin and snake venom phosphodiesterase I were obtained from Sigma-Aldrich. $POCl_3$ was purchased from Across Organics and freshly distilled prior to use. Fluorescent dyes Alexa594-alkyne **11** and calf intestinal alkaline phosphatase (CIP) were acquired from Invitrogen (Life Technologies). Activated cyclooctyne derivatives **13** and **14** were obtained from Click Chemistry Tools. Radiolabeled α-^{32}P ATP (2000 Ci/mmol) was purchased from the Board of Radiation and Isotope Technology, Government of India. T7 RNA polymerase, ribonuclease inhibitor (RiboLock), NTPs, RNase A and RNase T1 were obtained from Fermentas Life Science. DNA ONs were purchased from Integrated DNA Technologies, Inc., purified by gel electrophoresis under denaturing condition and desalted using Sep-Pak Classic C18 cartridges

(Waters Corporation). Chemicals for preparing all buffer solutions were purchased from Sigma-Aldrich (BioUltra grade). Autoclaved water was used in all biochemical reactions and HPLC analysis.

5.4.2 Instrumentation

NMR spectra were recorded on a 400 MHz Jeol ECS-400. All MALDI-MS measurements were recorded on an Applied Biosystems 4800 Plus MALDI TOF/TOF analyzer, MicroMass ESI-TOF and Water Synapt G2 High Defination mass spectrometer. Absorption spectra were recorded on PerkinElmer, Lambda 45 UV-Vis and Shimadzu UV-2600 spectrophotometers. Reversed-phase flash chromatographic (C18 Redi*SepRf* column) purifications were carried out using Combi Flash *Rf* Teledyne ISCO, RP-HPLC analyses were performed using Agilent Technologies 1260 Infinity and Dionex ICS 3000. Phosphorimages were recorded on a Typhoon Trio+, GE-Healthcare phosphorimager.

5.4.3 Synthesis



Scheme 5. Synthesis of 5-azidomethyl UTP (AMUTP **4**) from corresponding ribonucleoside **1**.

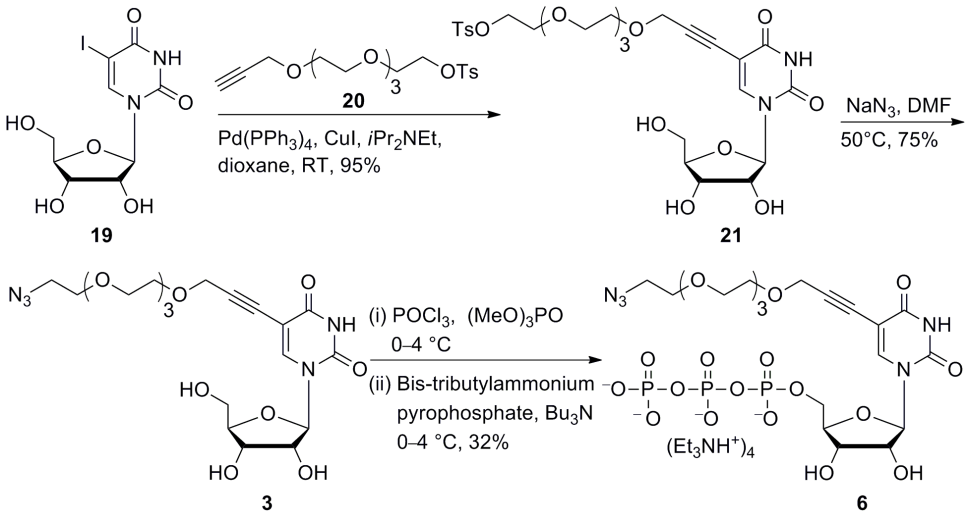
5-(azidomethyl) uridine (AMU 1): 5-(azidomethyl)-2′, 3′-O-isopropylideneuridine **18** [53] (0.80 g, 2.4 mmol) was dissolved in 50% aqueous acetic acid (16 ml) and refluxed for 4 h. The reaction mixture was then allowed to come to RT and the solvent was evaporated. The crude residue was purified by silica gel column chromatography (0–12% methanol in dichloromethane) to afford the product **1** as white solid (0.39 g, 55%). TLC (CH_2Cl_2:MeOH = 85:15) R_f = 0.30; ^{1}H NMR (400 MHz, d_6-DMSO): δ (ppm) 11.57 (s, 1H), 8.08 (s, 1H), 5.77 (d, J = 5.2 Hz, 1H), 5.41 (br, 1H), 5.13 (br, 2H), 4.05–4.02 (m, 3H), 3.97 (t, J = 4.4 Hz, 1H), 3.84 (dd, J = 7.4 Hz, 3.4 Hz, 1H), 3.65 (d, J = 12 Hz, 1H), 3.55 (d, J = 12 Hz, 1H); ^{13}C NMR (100 MHz, d_6-DMSO): δ (ppm) 162.9, 150.5, 140.1, 108.3, 87.9, 84.8, 73.6, 69.7, 60.8, 47.0;

HRMS: (*m/z*) Calculated for $C_{10}H_{12}N_5O_6$ $[M–H]^-$ = 299.0788, found = 299.0792; ε_{260} = 9300 $M^{-1}cm^{-1}$.

5-(azidomethyl) uridine-5′-triphosphate (AMUTP 4): To an ice cold solution of 5-(azidomethyl) uridine **1** (92 mg, 0.27 mmol, 1.0 equiv) in trimethyl phosphate (1.0 mL) was added freshly distilled $POCl_3$ (63 µL, 0.68 mmol, 2.5 equiv). The solution was stirred for 12 h at ~4 °C. A solution of *bis*-tributylammonium pyrophosphate (0.5 M in DMF, 2.7 mL, 5.0 equiv) and tributylamine (0.65 mL, 2.71 mmol, 10 equiv) was rapidly added under ice cold condition. The reaction was quenched after 30 min with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, 15 mL) and was washed with ethyl acetate (20 mL). The aqueous layer was evaporated and the residue was purified first on DEAE sephadex-A25 anion exchange column (10 mM–1 M TEAB buffer, pH 7.5) followed by reversed-phase flash column chromatography (C18 Redi*SepRf*, 0–40% acetonitrile in 50 mM triethylammonium actetate buffer, pH 7.5, 40 min). Appropriate fractions were lyophilized to afford the desired triphosphate product **4** as a tetratriethylammonium salt (113 mg, 39%); ^{1}H NMR (400 MHz, D_2O): δ (ppm) 7.96 (s, 1H), 5.82 (d, *J* = 3.6 Hz, 1H), 4.30 (br, 2H), 4.13–4.02 (m, 5H); ^{13}C NMR (100 MHz, D_2O): δ (ppm) 165.1, 151.7, 140.8, 109.8, 88.4, 83.4, 73.8, 69.5, 64.8, 47.1; ^{31}P NMR (162 MHz, D_2O): δ (ppm) –10.98 (br), –21.76 (br); HRMS: (*m/z*) Calculated for $C_{10}H_{15}N_5O_{15}P_3$ $[M–H]^-$ = 537.9777, found = 537.9796.

Tosyl-propargyloxy-tetra-ethylene glycol conjugated uridine (21): To a mixture of 5-iodouridine **19** (0.40 g, 1.08 mmol, 1.0 equiv), tosyl-propargyloxy-tetraethylene glycol **20** [54] (0.627 g, 1.62 mmol, 1.5 equiv), CuI (0.02 g, 0.11 mmol, 0.1 equiv) and tetrakis(triphenylphosphine) palladium(0) (0.075 g, 0.065 mmol, 0.06 equiv) in anhydrous dioxane (10 ml) was added *N*,*N*-diisopropylethylamine (0.75 ml, 4.32 mmol, 4.0 equiv) slowly. The reaction mixture was stirred at RT for 12 h. Solvent was evaporated and the resulting syrup was purified by silica gel column chromatography (0–10% methanol in dichloromethane) to give product **21** as faint orange semisolid (0.645 g, 95%). TLC (CH_2Cl_2:MeOH = 86:14) R_f = 0.73; ^{1}H NMR (400 MHz, d_6-DMSO): δ (ppm) 11.66 (s, 1H), 8.33 (s, 1H), 7.78 (d, *J* = 8.4 Hz, 2H), 7.48 (d, *J* = 8.4 Hz, 2H), 5.75 (d, *J* = 6.4 Hz, 1H), 5.42 (d, *J* = 5.2 Hz, 1H), 5.22 (t, *J* = 4.8 Hz, 1H), 5.08 (d, *J* = 5.2 Hz, 1H), 4.33 (d, *J* = 4.8 Hz, 2H), 4.10 (t, *J* = 4.2 Hz, 2H), 4.05 (dd, *J* = 10.0, *J* = 5.2 Hz, 1H), 3.98 (dd, *J* = 9.6, *J* = 4.8 Hz, 1H), 3.86 (dd, *J* = 7.0, *J* = 2.6 Hz, 1H), 3.58–3.52 (m, 10H), 3.49 (s, 2H), 3.44 (br, 4H),

2.42 (s, 3H); ^{13}C NMR (100 MHz, d_6-DMSO): δ (ppm) 161.5, 149.7, 144.9, 144.2, 132.4, 130.2, 128.1, 127.6, 125.5, 97.8, 88.8, 88.3, 84.8, 78.4, 73.9, 71.0, 70.0, 69.8, 69.7, 69.6, 69.5, 69.3, 68.6, 67.9, 60.3, 58.2, 21.1; HRMS: (*m/z*) Calculated for $C_{27}H_{36}N_2O_{13}SK$ $[M+K]^+$ = 667.1575, found = 667.1573.

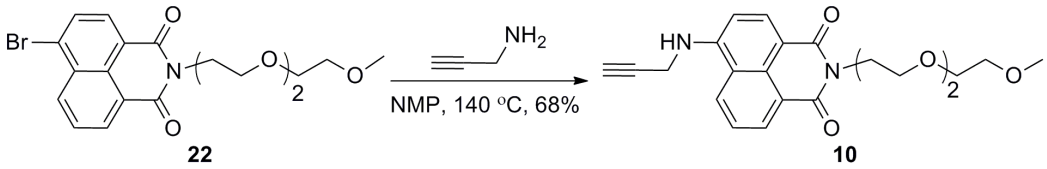


Scheme 6. Synthesis of azide-modified ribonucleoside ATU **3** and corresponding triphosphate ATUTP **6**.

Azido-propargyloxy-tetraethylene glycol conjugated uridine (ATU 3): To a solution of compound **21** (0.60 g, 0.96 mmol, 1.0 equiv) in DMF (6 ml) was added sodium azide (0.124 g, 1.92 mmol, 2.0 equiv). The reaction mixture was stirred at 50 °C for 4 h, then it was cooled to RT and solvent was evaporated. The resulting crude product was purified by silica gel column chromatography (0–12% methanol in dichloromethane) to afford the ribonucleoside **3** as a pale brown semi-solid (0.360 g, 76 %). TLC (CH_2Cl_2:MeOH = 86:14) R_f = 0.65; ^{1}H NMR (400 MHz, d_6-DMSO): δ (ppm) 11.64 (s, 1H), 8.32 (s, 1H), 5.75 (d, *J* = 4.4 Hz, 1H), 5.40 (br, 1H), 5.20 (t, *J* = 4.2 Hz, 1H), 5.06 (br, 1H), 4.33 (s, 2H), 4.05 (t, *J* = 4.6 Hz, 1H), 3.98 (t, *J* = 4.8 Hz, 1H), 3.83 (dd, *J* = 7.2, *J* = 2.8 Hz, 1H), 3.61–3.52 (m, 16H), 3.39 (t, *J* = 4.8 Hz, 2H); ^{13}C NMR (100 MHz, d_6-DMSO): δ (ppm) 161.6, 149.7, 144.2, 97.8, 88.8, 88.3, 84.8, 73.9, 69.8, 69.7, 69.6, 69.4, 69.2, 68.6, 60.3, 58.2, 50.0; HRMS: (*m/z*) Calculated for $C_{20}H_{29}N_5O_{10}K$ $[M+K]^+$ = 538.1551, found = 538.1558; ε_{260} = 3400 $M^{-1}cm^{-1}$.

Azido-propargyloxy-tetraethylene glycol conjugated uridine-5′-triphosphate (ATUTP 6): To an ice cold solution of azide-modified ribonucleoside **3** (90 mg, 0.18 mmol, 1.0 equiv)

in trimethyl phosphate (1 mL) was slowly added freshly distilled $POCl_3$ (42 μL, 0.45 mmol, 2.5 equiv). The reaction mixture was stirred for 5 h at ~4 °C. A solution of *bis*-tributylammonium pyrophosphate (0.5 M in DMF, 1.9 mL, 5.2 equiv) and tributylamine (0.47 mL, 1.98 mmol, 11 equiv) was rapidly added under ice cold condition. The reaction was quenched after 30 min with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.6, 15 mL) and was extracted with ethyl acetate (20 mL). The aqueous layer was evaporated and the residue was purified first on DEAE sephadex-A25 anion exchange column (10 mM–1 M TEAB buffer, pH 7.6) followed by reversed-phase flash column chromatography (C18 Redi*SepRf*, 0–40% acetonitrile in 100 mM triethylammonium actetate buffer, pH 7.5, 50 min). Appropriate fractions were lyophilized to afford the desired triphosphate product **6** as a tetratriethylammonium salt (66 mg, 32%). ^{1}H NMR (400 MHz, D_2O): δ (ppm) 7.98 (s, 1H), 5.75 (d, *J* = 4.4 Hz, 1H) 4.26 (br, 2H), 4.20 (br, 2H), 4.08 (br, 3H), 3.62 (t, *J* = 3.8, 2H), 3.55–3.51 (m, 12H), 3.30 (t, *J* = 4.6 Hz, 2H); ^{13}C NMR (100 MHz, D_2O): δ (ppm) 164.3, 150.7, 144.8, 99.2, 89.9, 88.8, 83.4, 77.2, 73.6, 69.5, 69.4, 69.2, 68.7, 65.0, 58.5, 50.1; ^{31}P NMR (162 MHz, D_2O): δ (ppm) −10.71 (br), –22.06 (br); HRMS: (*m/z*) Calculated for $C_{20}H_{32}N_5O_{19}P_3K$ $[M+K]^+$ = 778.0541, found = 778.0549.



Scheme 6. Synthesis of alkyne substrate **10** for CuAAC reaction, NMP = *N*-methyl-2-pyrrolidone.

4-(Prop-2-yn-1-ylamino-*N*-(triethyleneglycol monomethyl ether)-1,8-naphthalimide (10): A solution of 4-bromo-napthalimide derivative **22** [55] (0.200 g, 0.47 mmol, 1.0 equiv), propargylamine (60 μL, 0.94 mmol, 2.0 equiv) in *N*-methyl-2-pyrrolidone (2 mL) was heated at 140 °C for 12 h. Reaction mixture was diluted with water (10 mL) and the product was extracted with ethyl acetate (3×15 mL). The organic layer was dried over sodium sulphate and evaporated to dryness. The crude product was purified by silica gel chromatography (50–80% ethyl acetate in petroleum ether) to give product **10** as a lemon yellow solid (0.127 g, 68%). TLC (ethyl acetate) R_f = 0.50; ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) 8.44 (d, *J* = 7.2 Hz, 1H), 8.39 (d, *J* = 8.4 Hz, 1H), 8.07 (d, *J* = 8.8 Hz, 1H), 7.52 (t, *J* = 8 Hz, 1H), 6.72 (d, *J* =

8.4 Hz, 1H), 5.78 (s, 1H), 4.40 (t, *J* = 6.0 Hz, 2H), 4.20–4.18 (m, 2H), 3.84 (t, *J* = 6.0 Hz, 2H), 3.74 (t, *J* = 4.6 Hz, 2H), 3.65 (t, *J* = 4.8 Hz, 2H), 3.59 (t, *J* = 4.8 Hz, 2H), 3.43 (t, *J* = 4.8 Hz, 2H), 3.27 (s, 3H), 2.33 (t, *J* = 2.4 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) 164.7, 164.2, 148.3, 134.1, 131.2, 129.5, 126.4, 125.1, 122.9, 120.6, 111.5, 105.2, 79.1, 72.7, 72.0, 70.7, 70.6, 70.3, 68.3, 59.1, 39.0, 33.4; HRMS: *m/z* Calculated for $C_{22}H_{24}N_2O_5K$ $[M+K]^+$ = 435.1322, found = 435.1329.

5.4.4 Enzymatic incorporation of azide-modified UTP 4 and 6

Transcription reactions with α-^{32}P ATP: The promoter-template duplexes (5 μM) were assembled by heating T7 RNA polymerase consensus promoter DNA sequence and DNA template (**T1**–**T5)** in TE buffer (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.8) at 90 °C for 3 min. The solution was allowed to come to room temperature slowly and then placed on an ice bath for 20 min, and stored at –40 °C. The transcription reactions were carried out at 37 °C in 40 mM Tris-HCl buffer (pH 7.9) containing 250 nM annealed promoter-template duplexes, 10 mM $MgCl_2$, 10 mM NaCl, 10 mM of dithiothreitol (DTT), 2 mM spermidine, 1 U/μL RNase inhibitor (Riboblock), 1 mM GTP, CTP, UTP and or modified UTP **4**/**6**, 20 μM ATP, 5 μCi α-^{32}P ATP and 3 U/μL T7 RNA in a total volume of 20 μL. The reaction was quenched after 3.5 h by adding 20 μL of loading buffer (7 M urea in 10 mM Tris-HCl, 100 mM EDTA, 0.05% bromophenol blue, pH 8), heated for 3 min at 75 °C followed by cooling the samples on an ice bath. The samples (4 μL) were loaded on a sequencing 18% denaturing polyacrylamide gel and electrophoresed. The radioactive bands were phosphorimaged and then quantified using the GeneTools software from Syngene to determine the relative transcription efficiencies. Relative percentage incorporation of azide-modified ribonucleoside triphosphates **4**–**6** into full-length transcripts has been reported with respect to the amount of full-length transcript formed in the presence of natural NTPs. All reactions were performed in duplicate and the errors in yields were ≤ 4%.

Large-scale transcription reactions: Large-scale transcription reactions were performed using DNA template **T1**. Each reaction (250 μL) was performed in the presence of 2 mM GTP, CTP, ATP, and modified UTP **4**/**5**/**6**, 20 mM $MgCl_2$, 0.4 U/μL RNase inhibitor (Riboblock), 300 nM annealed template and 800 units T7 RNA polymerase. After incubating for 12 h at 37 °C, the reaction volume was reduced to 1/3 by speed vac. 50 μL of the loading buffer was added and the sample was loaded on a preparative 20% denaturing polyacrylamide

gel. The gel was UV shadowed, appropriate band was cut out, extracted with 0.3 M sodium acetate and desalted using Sep-Pak classic C18 cartridge.

5.4.5 Enzymatic digestion of RNA ONs 7 and 9

~4 nmol of the modified oligoribonucleotide transcripts **7** and **9** were treated with snake venom phosphodiesterase I (0.01 U), calf intestinal alkaline phosphatase (10 μL, 1 U/μL), and RNase A (0.25 μg) in a total volume of 100 μL in 50 mM Tris-HCl buffer (pH 8.5, 40 mM $MgCl_2$, 0.1 mM EDTA) at 37 °C. ON **7** was incubated for 5 h and ONs **9** was incubated for 12 h. After this period, RNase T1 (0.2 U/μL) was added, and the samples were incubated for another 1 h (**7**) and 4 h (**9**) at 37 °C. The ribonucleoside mixture obtained from the digest was analyzed by reversed-phase HPLC using Phenomenex-Luna C18 column (250 x 4.6 mm, 5 micron) at 260 nm.

5.4.6 In vitro transcription with linearized plasmid DNA template in the presence of modified UTPs 4−6

The EGFP gene was PCR amplified from the pEGFPC1 vector (Clontech) (forward primer 5' gcgcgcagatctATGGTGAGCAAGGGCGAGGAGCTGTTC 3', reverse primer 5' gcgcgcgtcgacatccatttcaacCTTGTACAGCTCGTCCATGCC GAGAGT 3') under standard PCR conditions with Platinum Pfx Polymerase (Invitrogen). The PCR product was cloned in the pDisplay vector (Invitrogen) in the Bgl2 and Sal1 restriction enzyme sites using standard restriction enzyme-ligation mediated cloning protocols to give GFPinpDisplay (Figure 6).[56] The construct was linearized with XhoI and purified by phenol-chloroform extraction.

The transcription reactions were carried out in 40 mM Tris-HCl buffer (pH 7.9) containing 500 ng of linearized EGFPinpDisplay plasmid template, 1 mM GTP, CTP, ATP, UTP and or modified UTPs **4/5/6**, 10 mM $MgCl_2$, 10 mM NaCl, 10 mM of dithiothreitol (DTT), 2 mM spermidine, 1 U/μL RNase inhibitor (Riboblock) and 40 units T7 RNA polymerase in a total volume of 25 μL. Samples were incubated at 37 °C for 3h. LiCl (7.5 M, 12.5 μL) was added and the reaction solutions were kept at –20 °C for 1 h followed by centrifugation (12000 rpm) at 4 °C for 15 min. The RNA pellets formed from each reaction were washed with 70% aqueous ethanol, dried and dissolved in 20 μL of nuclease free water. The presence of modified ribonucleosides (**1**–**3**) in each transcript was confirmed by enzymatic digestion followed by HPLC analysis of ribonucleoside products obtained from the digest (See Figure 7)

5.4.7 Posttranscriptional modification of azide-modified RNA ONs 7–9 by CuAAC reaction

A solution of THPTA (3.3 μL, 90 mM), $CuSO_4$ (3.3 μL, 45 mM) and sodium ascorbate (3.3 μL, 90 mM) was added to an aqueous solution of azide-modified RNA ONs **7/8/9** in water (27 μL, 0.55 mM). Stock solutions (7.5 mM) of alkyne substrates **10**, **11** and **12** were prepared in DMSO. Alkyne substrates (10 μL, 7.5 mM) were added to individual reaction mix and the reaction volume was adjusted to 50 μL by adding water. The final concentration of reaction components was the following: THPTA (6.0 mM), $CuSO_4$ (3.0 mM), sodium ascorbate (6.0 mM), RNA ON (0.30 mM, 15 nmol), alkyne substrate (1.5 mM) and DMSO (20%). The reaction mixtures were incubated at 37 °C for 30 min and purified by PAGE under denaturing conditions. The bands corresponding to clicked products were cut and extracted in 0.5 M ammonium acetate and desalted using Sep-Pak classic C18 cartridge. For structure of clicked products see Scheme 2 and for yields and mass data see Table 2.

5.4.8 Posttranscriptional modification of azide-modified RNA ONs 7–9 by copper-free SPAAC reaction with cyclooctyne building blocks of biotin (13) and Cy3 (14)

A solution of azide-modified RNA ONs **7/8/9** in water (23 μL, 0.65 mM) was mixed with alkyne substrate **13** or **14** (4.5 μL, 10 mM) dissolved in DMSO. Total volume of the reaction was adjusted to 37.5 μL by adding 10 μL of water such that the final concentration of ONs was 0.40 mM (15 nmol) and alkyne substrates was 1.2 mM and percentage DMSO was 12. The individual reaction mixture was incubated at 37 °C for 1 h and the clicked product was purified by PAGE under denaturing conditions. The bands corresponding to clicked product was cut and extracted in 0.5 M ammonium acetate and desalted using Sep-Pak classic C18 cartridge. For structure of clicked products see Scheme 3 and for yields and mass data see Table 2.

5.4.9 Posttranscriptional modification of azide-modified RNA ONs 7–9 by Staudinger ligation reaction with biotinylated triaryl phosphine substrate 15

A solution of azide-modified RNA ONs **7/8/9** in water (27 μL, 0.55 mM) was mixed with PBS buffer (10 μL, 100 mM). The biotinylated phosphine substrate **15** (10 μL, 15 mM) in DMSO was then added to the above solution and mixed well.[45] The reaction volume was adjusted to 50 μL by adding water. The final concentration of reaction components was the following: RNA ONs (0.30 mM, 15 nmol), **15** (3.0 mM) and DMSO (20%). The reaction

mixture was incubated at 60 °C for 15 h and was purified by PAGE under denaturing conditions. Bands corresponding to the ligated products were cut and extracted in 0.5 M ammonium acetate and desalted using Sep-Pak classic C18 cartridge. For structure of products see Scheme 4 and for yields and mass data see Table 2.

5.5 Refrences

[1] D. R. Corey, *J. Clin. Invest.* **2007**, *117*, 3615−3622.

[2] F. Wachowius, C. Höbartner, *ChemBioChem* **2010**, *11*, 469−480.

[3] M. Baker, *Nature Methods* **2012**, *9*, 787−790.

[4] G. Haukenes, A.-M. Szilvay, K. A. Brokstad, K.-H. Kalland, *BioTechniques* **1997**, *22*, 308−312..

[5] D. Cmarko, P. J. Verschure, T. E. Martin, M. E. Dahmus, S. Krause, R. X.-D. Fu, van Driel, S. Fakan, *Mol. Biol. Cell* **1999**, *20*, 211−223.

[6] A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, *Science* **1998**, *280*, 585−590.

[7] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, *Nature Methods* **2008**, *5*, 877−879.

[8] S. Tyagi, *Nature Methods* **2009**, *6*, 331−338.

[9] Z. Pianowski, K. Gorska, L. Oswald, C. A. Merten, N. Winssinger, *J. Am. Chem. Soc.* **2009**, *131*, 6492−6497.

[10] R. M. Franzini, E. T. Kool, *J. Am. Chem. Soc.* **2009,** *131*, 16021−16023.

[11] J. S. Paige, K. Y. Wu, S. R. Jaffrey, *Science* **2011**, *333*, 642−646.

[12] G. Bao, W. J. Rhee, A. Tsourkas, *Annu. Rev. Biomed. Eng.* **2009**, *11*, 25−47.

[13] J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.* **2005**, *1*, 13–21.

[14] A. G. Neef, C. Schultz, *Angew. Chem. Int. Ed*. **2009**, *48*, 1498−1500.

[15] S. S. van Berkel, M. B. van Eldijk, J. C. M. van Hest, *Angew. Chem. Int. Ed.* **2011**, *50*, 8806−8827; *Angew. Chem*. **2011**, *123*, 8968−8989.

[16] C. D. Spicer, B. G. Davis, *Nat. Commun*. 2014, 5:4740 doi: 10.1038/ncomms5740.

[17] P. M. E. Gramlich, C. T. Wirges, A. Manetto, T. Carell, *Angew. Chem. Int. Ed*. **2008**, *47*, 8350−8358; *Angew. Chem*. **2008**, *120*, 8478−8487.

[18] A. H. El-Sagheer, T. Brown, *Chem. Soc. Rev.*, **2010**, *39*, 1388−1405.

[19] A. Salic, T. J. Mitchison, *Proc. Natl. Acad. Sci. U.S.A.*, **2008**, *105*, 2415−2420.

[20] P. M. E. Gramlich, S. Warncke, J. Gierlich, T. Carell, *Angew. Chem. Int. Ed*. **2008**, *47*, 3442−3444; *Angew. Chem*. **2008**, *120*, 3491−3493.

[21] Y. Xu, Y. Suzuki, M. Komiyama, *Angew. Chem. Int. Ed*. **2009**, *48*, 3281−3284; *Angew. Chem*. **2009**, *121*, 3331−3334.

[22] C. Beyer, H.-A. Wagenknecht, *Chem. Commun.* **2010**, *46*, 2230−2231.

[23] G. Qing, H. Xiong, F. Seela, T. Sun, *J. Am. Chem. Soc*. **2010**, *132*, 15228−15232.

[24] A. B. Neef, N. W. Luedtke, *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108*, 20404−20409.

[25] J. Šečkutė, J. Yang, N. K. Devaraj, *Nucleic Acids Res.* **2013**, *41*, e148.

[26] L. Lercher, J. F. McGouran, B. M. Kessler, C. J. Schofield, B. G. Davis, *Angew. Chem. Int. Ed.* **2013,** *52*, 10553−10558.

[27] U. Rieder, N. W. Luedtke, *Angew. Chem. Int. Ed.* **2014**, *53*, 9168−9172; *Angew. Chem.* **2014**, *126*, 9322−9326.

[28] C. Y. Jao, A. Salic, *Proc. Natl. Acad. Sci. U.S.A.*, **2008**, *105*, 15779−15784.

[29] M. Grammel, H. Hang, N. K. Conrad, *ChemBioChem* **2012**, *13*, 1112−1115.

[30] D. Qu, L. Zhou, W. Wang, Z. Wang, G. Wang, W. Chi, B. Zhang, *Anal. Biochem.* **2013**, *434*, 128−135.

[31] T. Wada, A. Mochizuki, S. Higashiya, H. Tsuruoka, S. Kawahara, M. Ishikawa, M. Sekine, *Tet. Lett*. **2001**, *42*, 9215−9219.

[32] G. Pourceau, A. Meyer, J.-J. Vasseur, F. Morvan, *J. Org. Chem*. **2009**, *74*, 6837−6842.

[33] T. Santner, M. Hartl, K. Bister, R. Micura, *Bioconjugate Chem*. **2014**, *25*, 188−195.

[34] S. H. Weisbrod, A. Marx, *Chem. Commun*. **2007**, 1828−1830

[35] A. B. Neef, N. W. Luedtke, *ChemBioChem* **2014**, *15*, 789−793.

[36] A. H. El-Sagheer, T. Brown, *Proc. Natl. Acad. Sci. U.S.A*, **2010**, *107*, 15329−15334.

[37] E. Paredes, S. R. Das, *ChemBioChem* **2011**, *12*, 125−131.

[38] T. Ishizuka, M. Kimoto, A. Sato, I. Hirao, *Chem. Commun.* **2012**, *48*, 10835−10837.

[39] K. J. Phelps, J. M. Ibarra-Soza, K. Tran, A. J. Fisher, P. A. Beal, *ACS Chem. Biol*. **2014**, *9*, 1780−1787.

[40] A. Samanta, A. Krause, A. Jäschke, *Chem. Commun.* **2014**, *50*, 1313−1316.

[41] J. M. Holstein, D. Schulz, A. Rentmeister, *Chem. Commun.* **2014**, *50*, 4478−4481.

[42] H. Rao, A. A. Tanpure, A. A., Sawant, S. G. Srivatsan, *Nat. Protocols* **2012**, *7*, 1097–1112.

[43] H. Rao, A. A. Sawant, A. A. Tanpure, S. G. Srivatsan, *Chem. Commun.* **2012**, *48*, 498–500.

[44] S. G. Srivatsan, Y. Tor, *J. Am. Chem. Soc.* **2007**, *129*, 2044–2053

[45] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.

[46] A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J.* **2011**, *17*, 12820–12827.

[47] See experimental section for details.

[48] J. F. Milligan, O. C. Uhlenbeck, *Methods Enzymol*. **1989**, *180*, 51–62.

[49] N. K. Vaish, A. W. Fraley, J. W. Szostak, L. W. McLaughlin, *Nucleic Acids Res.* **2000**, *28*, 3316−3322.

[50] M. Kimito, R. Yamashige, K.-I. Matsunaga, S. Yokoyama, I. Hirao, *Nat. Biotech*. **2013**, *31*, 453−458.

[51] A. A. Kislukhin, V. P. Hong, K. E. Breitenkamp, M. G. Finn, *Bioconjugate Chem*. **2013**, *24*, 684−689.

[52] K. Upadhya, I. K. Khattak, B. Mullah, *Nucleoside, Nucleotides Nucleic Acids,***2005**, *24,* 919–922.

[53] K. Seio, T. Wada, K. Sakamoto, S. Yokoyama, M. Sekine, *J. Org. Chem.* **1998**, *63,* 1429−1443.

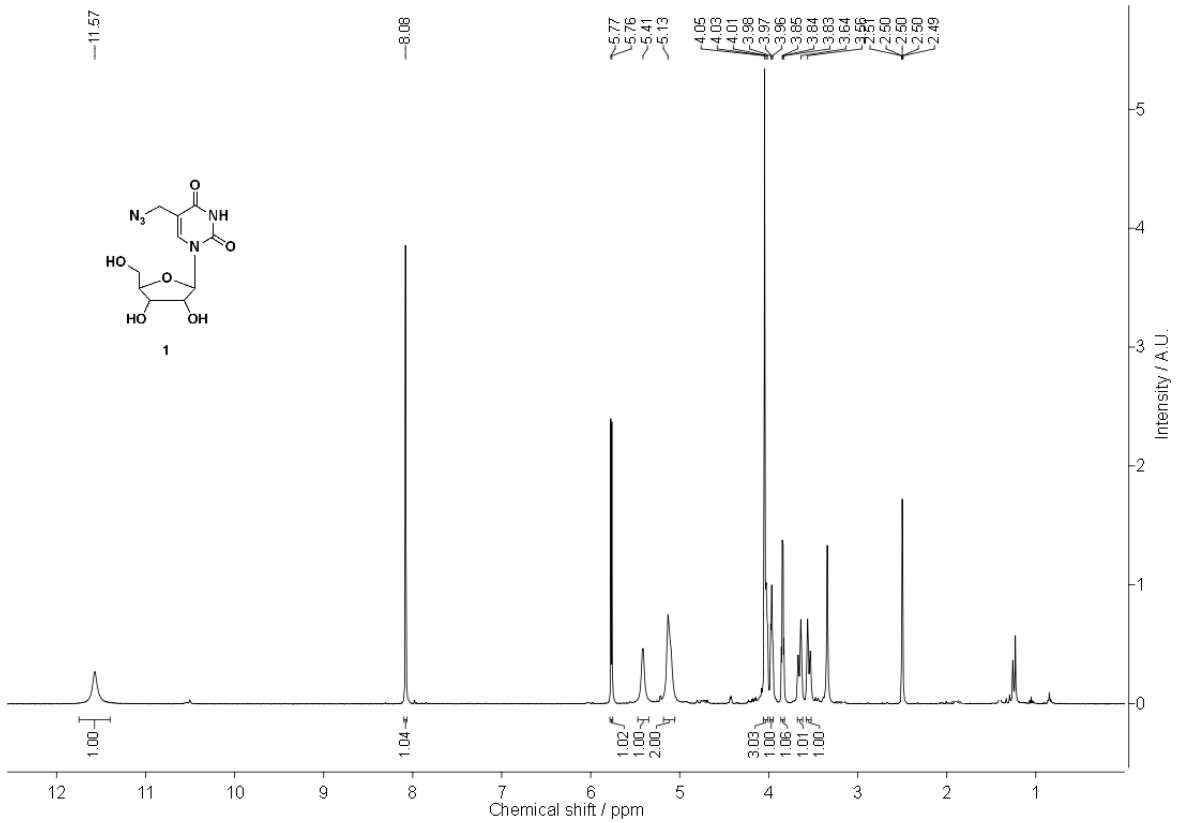
[54] S. Zhu, J. Zhang, G. Vegesna, F-T. Luo, S. Green, H. Liu, *Org. Lett.* **2011**, *13*, 438–41.

[55] A. A. Tanpure, S. G. Srivatsan, *ChemBioChem* **2014**, 15, 1309–1316.

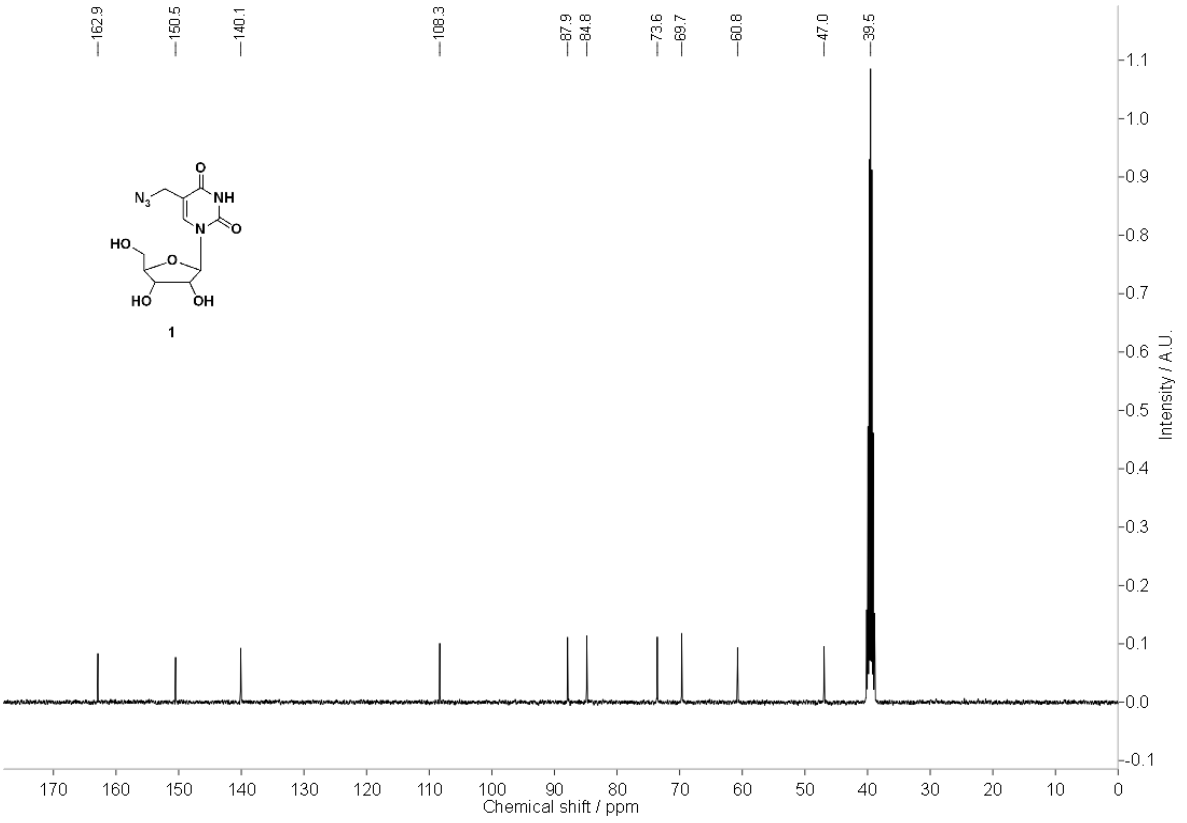
[56] J. Sambrook, D.W. Russell, *Molecular Cloning - A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2001**, pp. 4.1−4.86

5.6 Appendix-V: Characterization data of synthesized compounds

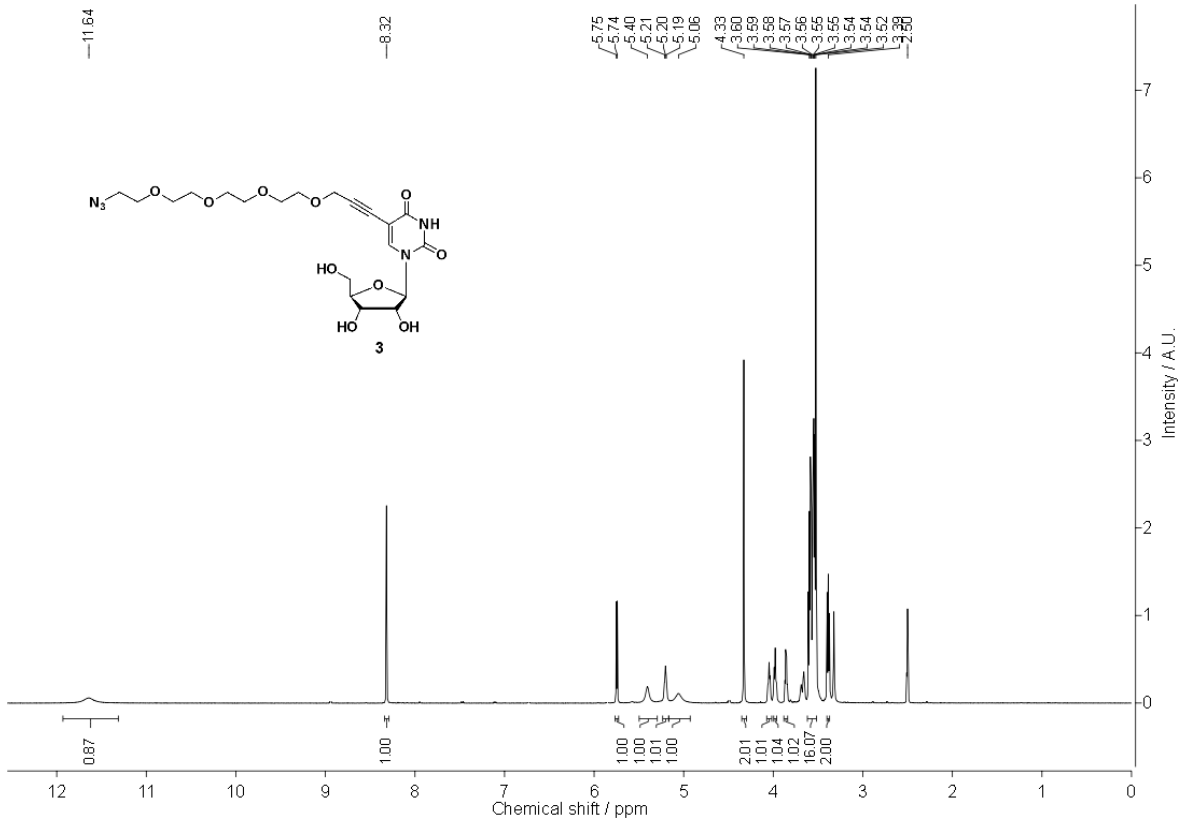
^{1}H-NMR of azide-modified ribonucleoside **1** in d_6- DMSO



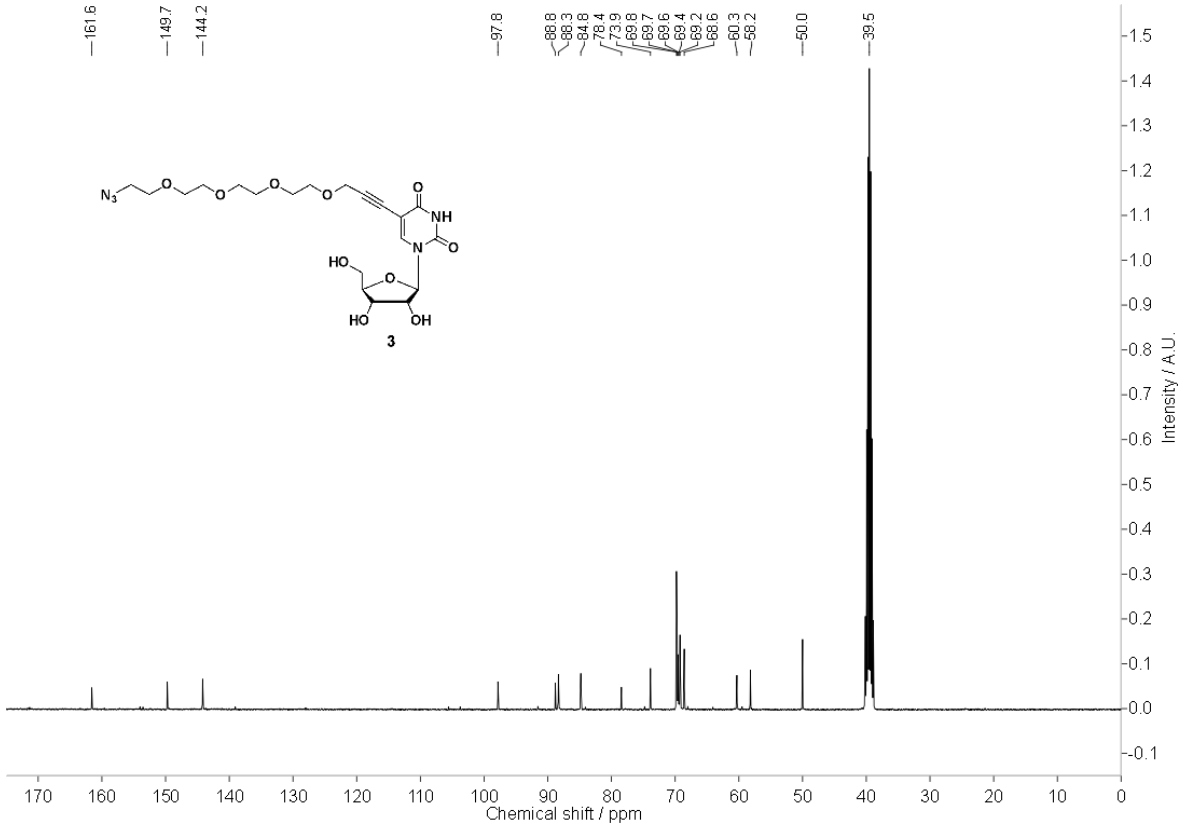
^{13}C-NMR of azide-modified ribonucleoside **1** in d_6- DMSO



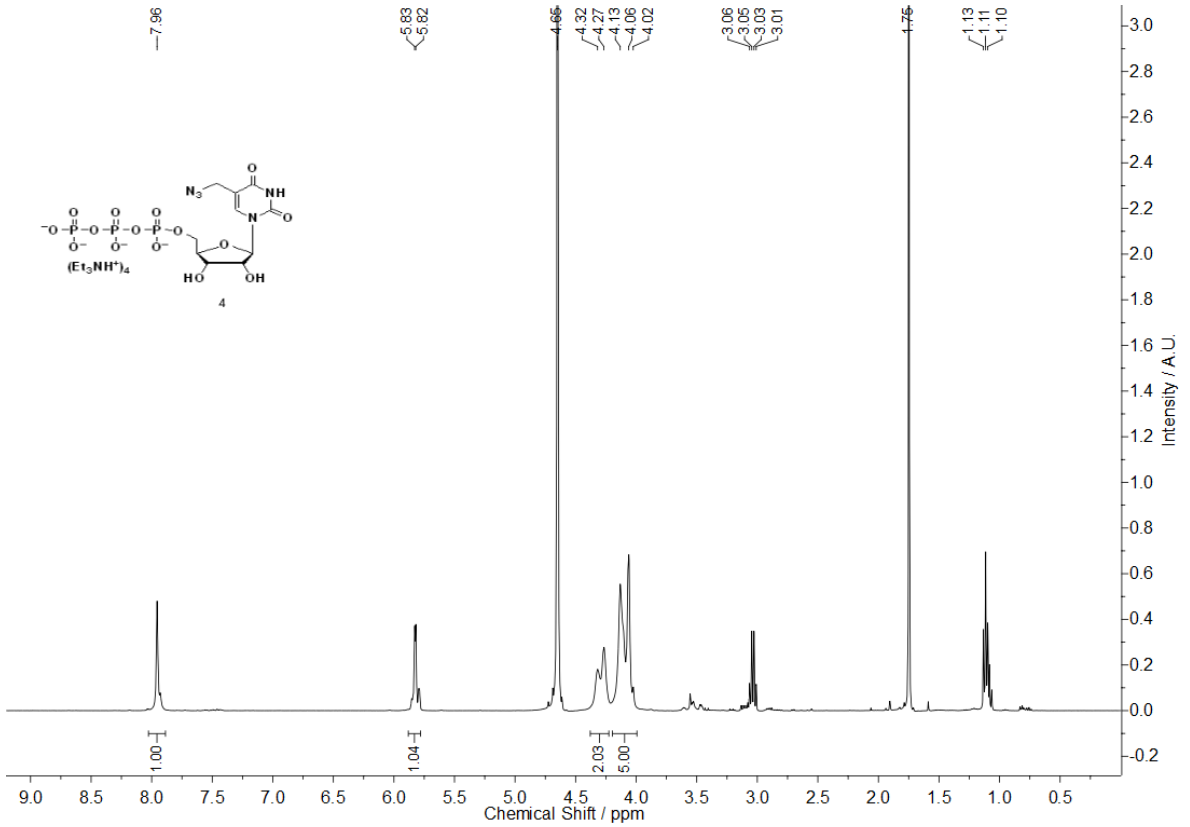
^{1}H-NMR of azide-modified ribonucleoside **3** in d_6-DMSO



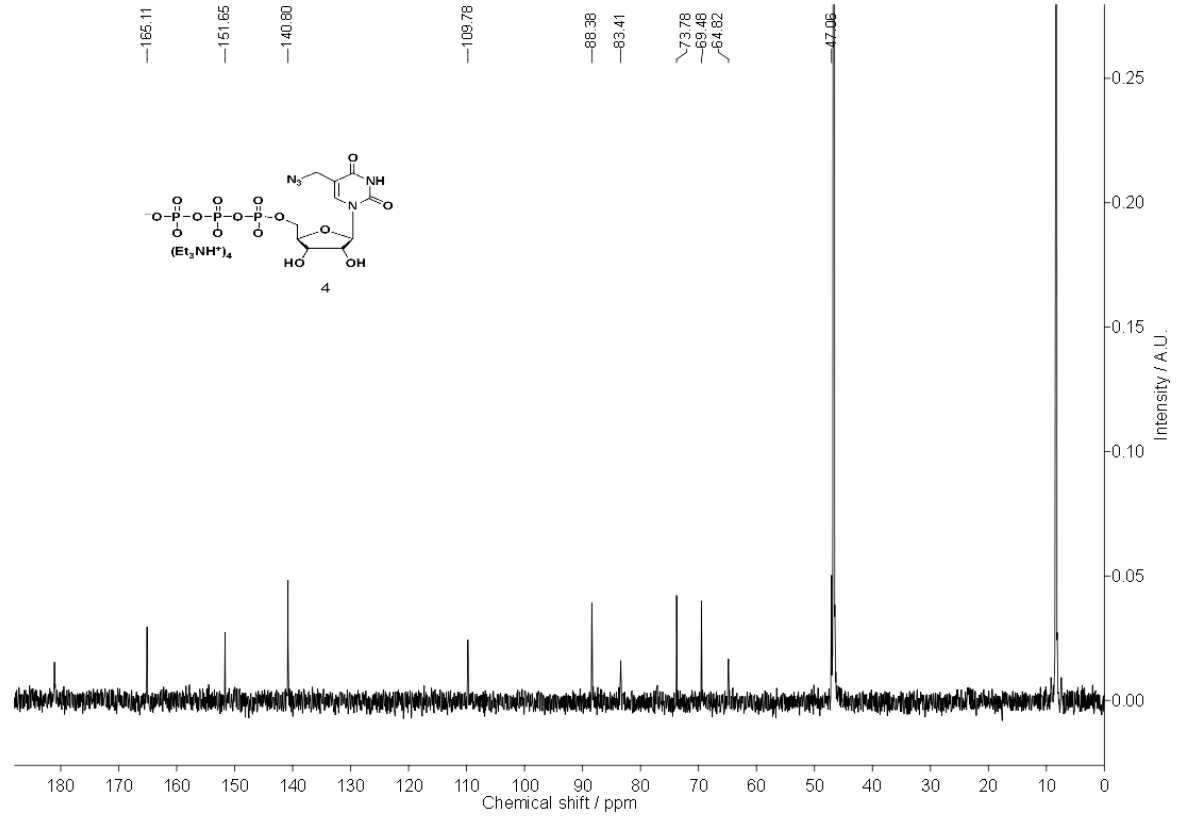
^{13}C-NMR of azide-modified ribonucleoside **3** in d_6- DMSO



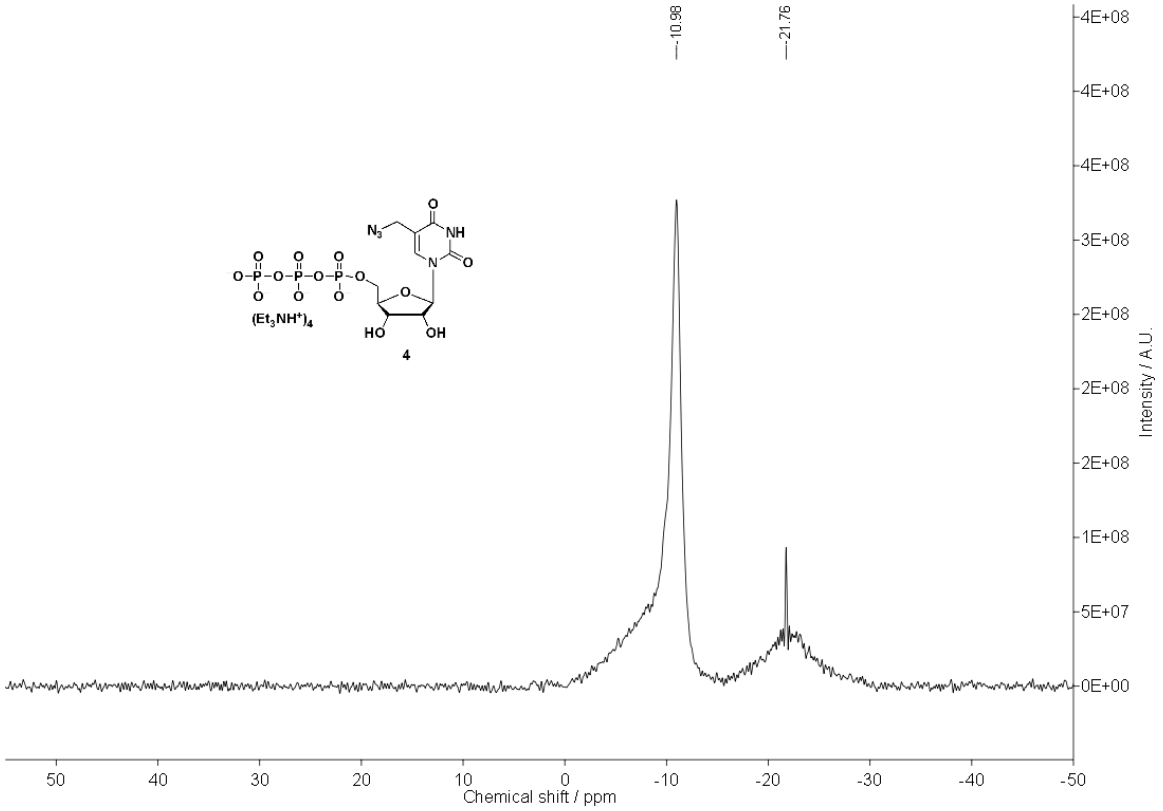
^{1}H-NMR of azide-modified ribonucleoside triphosphate **4** in D_2O. Trace amounts of triethylammonium acetate buffer is also present.



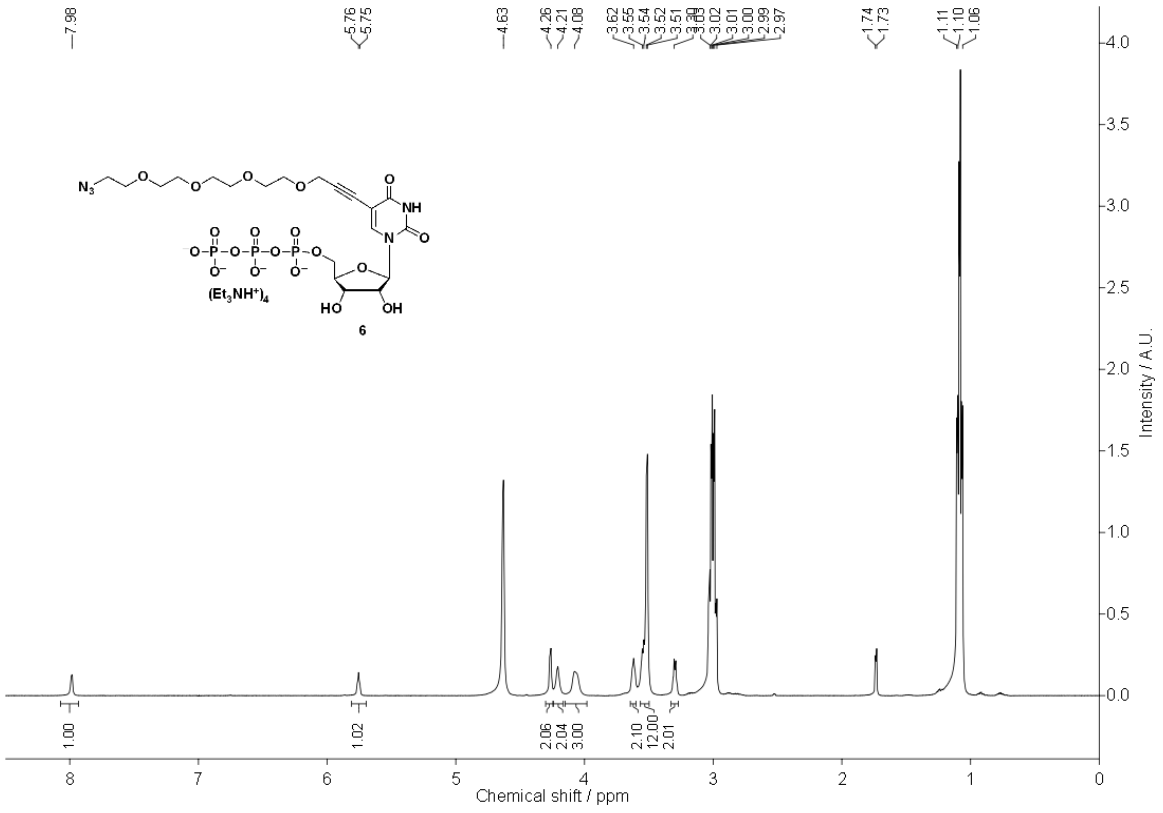
^{13}C-NMR of azide-modified ribonucleoside triphosphate **4** in D_2O. Trace amounts of triethylammonium acetate buffer is also present.



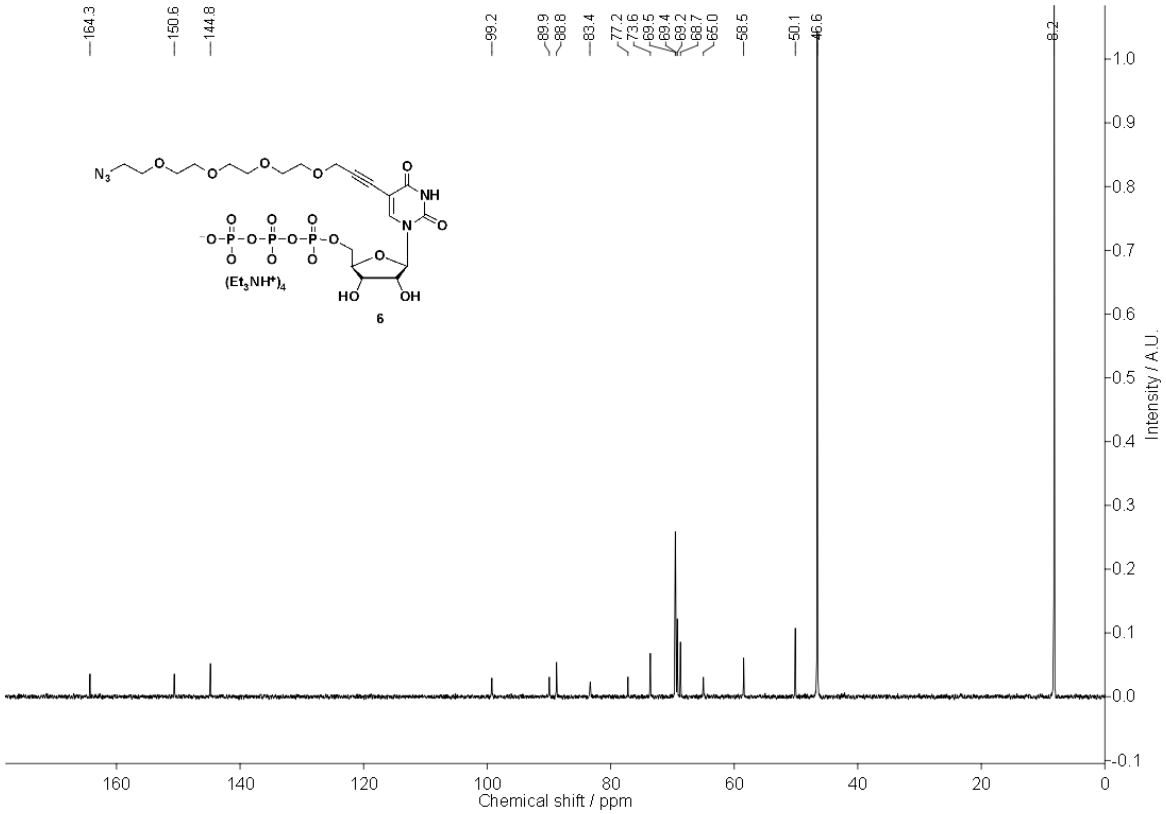
^{31}P-NMR of azide-modified ribonucleoside triphosphate **4** in D_2O.



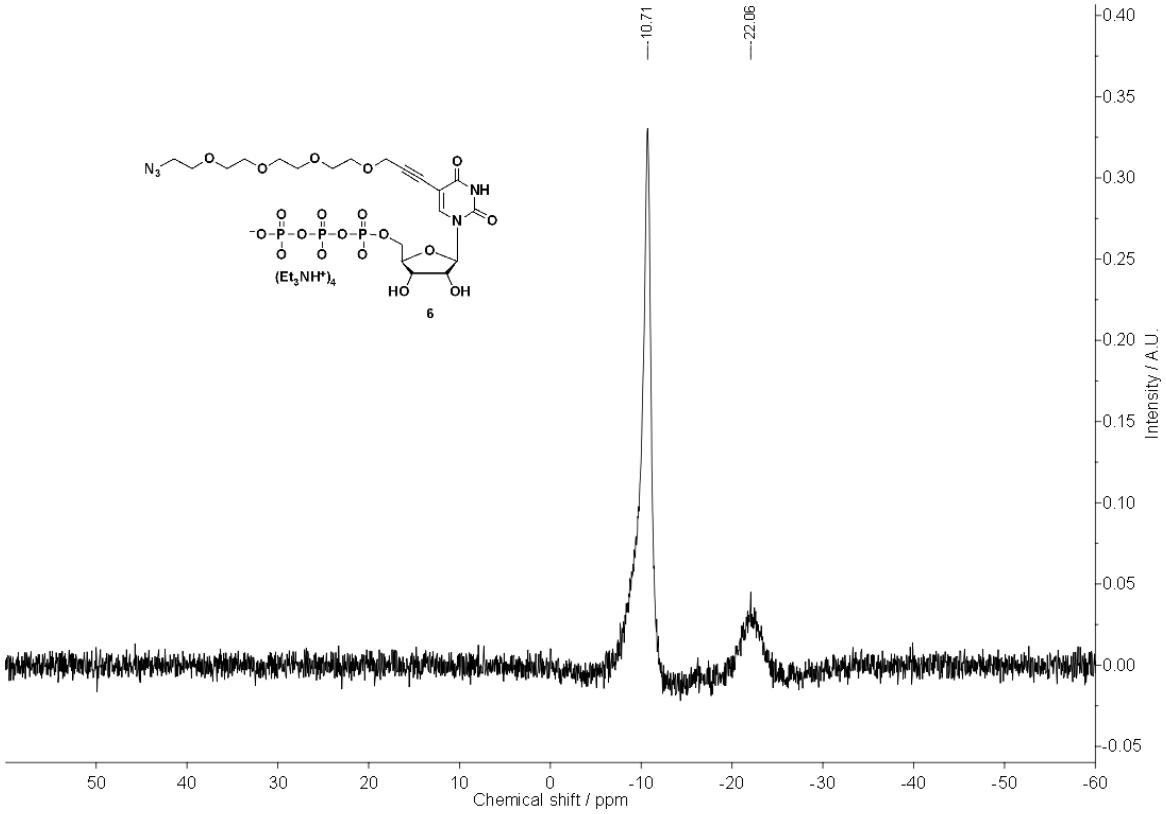
^{1}H-NMR of azide-modified ribonucleoside triphosphate **6** in D_2O. Trace amounts of triethylammonium acetate buffer is also present.



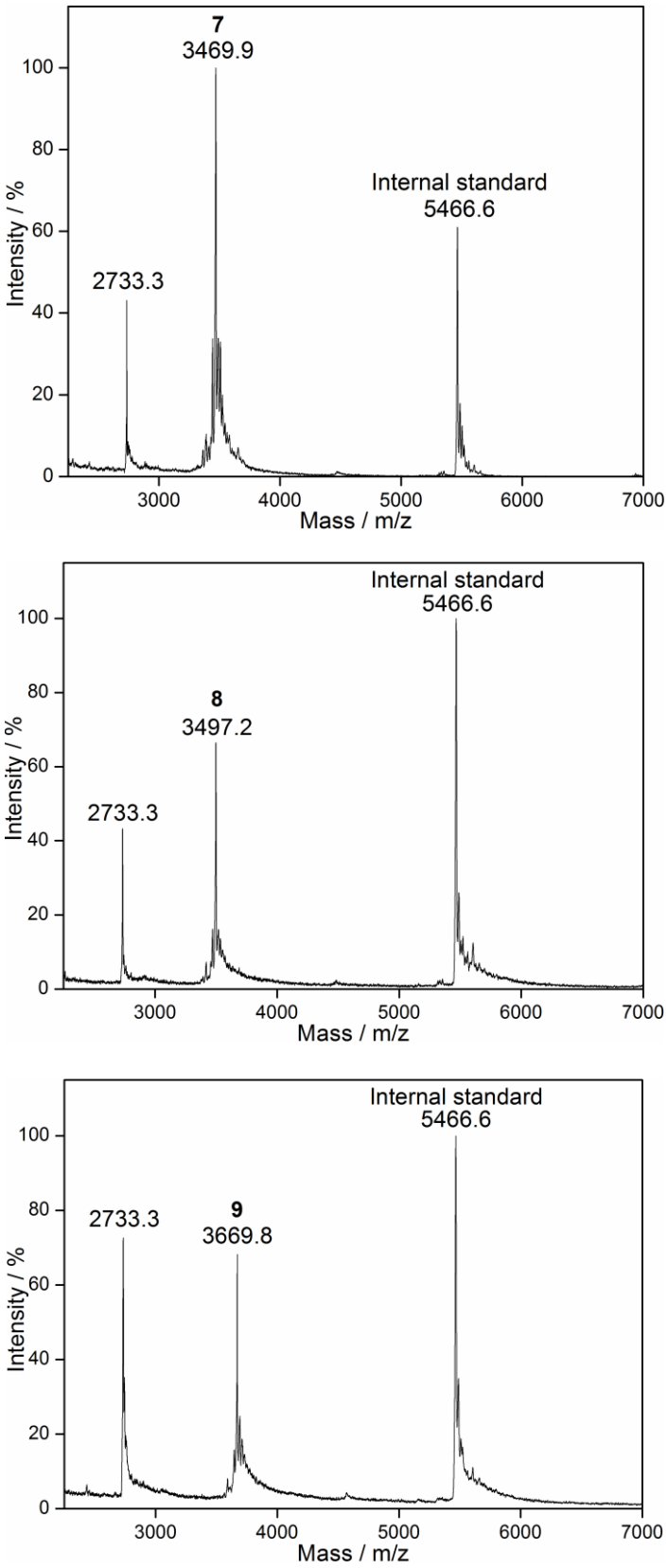
^{13}C-NMR of azide-modified ribonucleoside triphosphate **6** in D_2O. Trace amounts of triethylammonium acetate buffer is also present.



^{31}P-NMR of azide-modified ribonucleoside triphosphate **6** in D_2O



MALDI-TOF mass spectra of azide-modified transcripts **7**–**9** calibrated relative to the +1 and +2 ions of an internal 18-mer DNA ON standard (m/z for +1 and +2 ions are 5466.6 and 2733.3, respectively).



Final Conclusions

The main objective of this thesis was to develop biophysical tools to investigate the structure, function and recognition properties of nucleic acids using tailor-made nucleoside analogue probes. Towards this endeavour, we were successful in developing microenvironment-sensitive fluorescent pyrimidine ribonucleoside analogues by conjugating heterobicycles at the 5-position of uracil. In particular, 5-benzofuran-modified uridine was found to be highly emissive and sensitive to it surrounding environment. Importantly, when incorporated into DNA and RNA oligonucleotides (ONs), it retained reasonable fluorescence and was highly sensitive to its neighbouring base environment. Its probe-like properties within DNA and RNA ONs enabled the development of fluorescence assays to (i) detect DNA and RNA lesions (e.g., abasic sites), (ii) monitor nucleic acids- and topology-specific binding of ligands to G-quadruplexes, and (iii) detect the formation of non-canonical nucleic acid structures such as i-motif. We were also successful in developing another novel nucleoside analogue based on the Lucifer chromophore, with excitation and emission in the visible region and a high quantum yield. The triphosphate and phosphoramidite derivative of naphthalimide-modified nucleoside acts as a good substrate for the synthesis of fluorescent RNA ONs by *in vitro* transcription reaction as well as by solid-phase ON synthesis protocol, respectively. Importantly, the emissive nucleoside incorporated into ON duplexes exhibits appreciable fluorescence efficiency, and is also responsive to changes in its flanking bases and base pair substitutions.

In parallel, we synthesized a small series of azide-modified UTP analogues to establish a methodology to label RNA *in vitro* and in cells by using bioorthogonal chemical reactions. We were able to incorporate azide functionality into RNA by *in vitro* transcription reactions and further developed a modular posttranscriptional chemical functionalization protocol to label RNA *in vitro* with variety of biophysical probes. One of these analogues, 5-azidomethyl uridine AMUTP served as an efficient substrate for endogenous RNA polymerases. This is the first example of selective labeling of cellular RNA transcripts with azide groups. The selective labeling of RNA with azide group enabled the imaging newly transcribing RNA in fixed and in live cells by click reactions.

Taken together, our results highlight the potential of these base-modified nucleoside analog as useful probes in the study of nucleic acids *in vitro* as well as inside the cells.